Journal of Healthcare Engineering

Vision Based Computing Systems for Healthcare Applications

Lead Guest Editor: Subrahmanyam Murala
Guest Editors: Santosh K. Vipparthi and Zahid Akhtar

Vision Based Computing Systems for Healthcare Applications

Journal of Healthcare Engineering

Vision Based Computing Systems for Healthcare Applications

Lead Guest Editor: Subrahmanyam Murala
Guest Editors: Santosh K. Vipparthi and Zahid Akhtar

Editorial Board

Contents

Hindawi
Journal of Healthcare Engineering
Volume 2019, Article ID 9581275, 2 pages
https://doi.org/10.1155/2019/9581275

Editorial

Vision Based Computing Systems for Healthcare Applications

Subrahmanyam Murala,[1] **Santosh Kumar Vipparthi**,[2] **and Zahid Akhtar**[3]

[1]*Department of Electrical Engineering, Indian Institute of Technology Ropar, Roopnagar, India*
[2]*Department of Computer Science and Engineering, Malaviya National Institute of Technology, Jaipur, India*
[3]*Department of Computer Science, University of Memphis, Memphis, USA*

Correspondence should be addressed to Subrahmanyam Murala; subbumurala@iitrpr.ac.in

Received 3 December 2018; Accepted 4 December 2018; Published 14 February 2019

Vision-based systems for healthcare applications utilize computer vision techniques to provide intelligent support to patients and medical practitioners for common needs to sophisticated technical assistance. With the ease of availability and advancements in sensing, mobile devices, and computing power, development of advanced computer vision techniques for assistive healthcare is actively being pursued both by researchers and industries. These systems are built with a combination of pattern recognition, computer vision, machine learning, sensing devices, medical records, expert suggestions, and feedbacks from the medical practitioners. Various vision-based systems are being used in biomedical image analysis, patient monitoring, gesture recognition, gait analysis, advanced rehabilitation, and human psychology, i.e., mental behaviour analysis and lie detection/microexpression analysis.

Pattern recognition is one of the essential parts of the vision-based analytics system. The efficacy of any pattern recognition technique primarily depends on two aspects: feature extraction and classification. As a result, extensive efforts have been made in designing robust, accurate feature descriptors and developing relevant classification techniques. Srensen et al. [1] quantitatively analyzed pulmonary emphysema in computed tomography (CT) images of the lungs using local binary patterns (LBPs). Similarly, Mandal et al. [2] designed ANTithetic Isomeric Cluster Patterns (ANTIC) for robust query-based retrieval of CT images. For analysis of MRI images, numerous variants of LBP have also been proposed in the literature. Vipparthi et al. [3] proposed local directional mask maximum edge patterns for MRI image retrieval. Furthermore, Murala and Wu [4] used local mesh patterns for MRI image indexing and retrieval. The researchers have also used feature descriptors to grade the retinopathy images for diabetic retinopathy detection. The retinopathy grading is specified based on the number of haemorrhages, microaneurysms, and the sign of neovascularization. The absence of these abnormalities constitutes a normal image. Mandal et al. [2] distinguished these features using ANTIC feature descriptors.

Computer vision approaches are also being used for patient behaviour analysis through emotion recognition, gait analysis, and gesture recognition. Various human-computer interaction systems have been developed for improved patient diagnosis and communication between the patient and doctors.

Recent advances in deep learning techniques [5] have led to development of a more intelligent computer-aided detection and analysis system. Ronneberger et al. [6] designed a U-Net convolutional neural network (CNN) for biomedical image segmentation. In addition, Falah et al. [7] employed virtual reality- (VR-) based visualization and training environments in the delivery of anatomy teaching transferring the learning experience.

The scope of visual analytics in healthcare applications is becoming more relevant due to the rise of deep learning techniques, leading to very high accuracy in various computer vision tasks such as image classification, object detection, and semantic segmentation. The potential of these deep networks in medical image analysis has already been tested by the recent work. However, there are many important and interesting challenges such as contextual understanding of visual information and interpretability of deep learning models. We hope that the readers will find interesting research for vision-based healthcare applications in this special issue.

Conflicts of Interest

We declare that there are no conflicts of interest regarding the publication of this special issue.

Subrahmanyam Murala
Santosh Kumar Vipparthi
Zahid Akhtar

References

[1] L. Srensen, S. B. Shaker, and M. De Bruijne, "Quantitative analysis of pulmonary emphysema using local binary patterns," *IEEE Transactions on Medical Imaging*, vol. 29, no. 2, pp. 559–569, 2010.

[2] M. Mandal, M. Chaudhary, S. K. Vipparthi, S. Murala, A. B. Gonde, and S. K. Nagar, "ANTIC: ANTithetic isomeric cluster patterns for medical image retrieval and change detection," *IET Computer Vision*, 2018, In press.

[3] S. K. Vipparthi, A. B. Gonde, Q. M. Jonathan Wu, and S. Murala, "Local directional mask maximum edge patterns for image retrieval and face recognition," *IET Computer Vision*, vol. 10, no. 3, pp. 182–192, 2016.

[4] S. Murala and Q. M. J. Wu, "Local mesh patterns versus local binary patterns: biomedical image indexing and retrieval," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 3, pp. 929–938, 2014.

[5] S. Hoo-Chang, H. R. Roth, M. Gao et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of International Conference Medical Image Computing Computer-Assisted Intervention*, pp. 234–241, Munich, Germany, October 2015.

[7] J. Falah, S. Khan, T. Alfalah et al., "Virtual Reality medical training system for anatomy education," in *Proceedings of Science and Information Conference*, pp. 752–758, London, UK, August 2014.

Hindawi
Journal of Healthcare Engineering
Volume 2019, Article ID 9610212, 21 pages
https://doi.org/10.1155/2019/9610212

Research Article

Development of a Stand-Alone Independent Graphical User Interface for Neurological Disease Prediction with Automated Extraction and Segmentation of Gray and White Matter in Brain MRI Images

Ayush Goyal,[1] **Sunayana Tirumalasetty**,[1] **Gahangir Hossain**,[1] **Rajab Challoo**,[1] **Manish Arya**,[2] **Rajeev Agrawal**,[2] **and Deepak Agrawal**[3]

[1]*Texas A&M University-Kingsville, Kingsville, Texas, USA*
[2]*G. L. Bajaj Institute of Technology and Management, Greater Noida, UP, India*
[3]*All India Institute of Medical Sciences, New Delhi, India*

Correspondence should be addressed to Ayush Goyal; ayush.goyal@tamuk.edu

Received 22 March 2018; Accepted 16 September 2018; Published 14 February 2019

Guest Editor: Subrahmanyam Murala

This research presents an independent stand-alone graphical computational tool which functions as a neurological disease prediction framework for diagnosis of neurological disorders to assist neurologists or researchers in the field to perform automatic segmentation of gray and white matter regions in brain MRI images. The tool was built in collaboration with neurologists and neurosurgeons and many of the features are based on their feedback. This tool provides the user automatized functionality to perform automatic segmentation and extract the gray and white matter regions of patient brain image data using an algorithm called adapted fuzzy *c*-means (FCM) membership-based clustering with preprocessing using the elliptical Hough transform and postprocessing using connected region analysis. Dice coefficients for several patient brain MRI images were calculated to measure the similarity between the manual tracings by experts and automatic segmentations obtained in this research. The average Dice coefficients are 0.86 for gray matter, 0.88 for white matter, and 0.87 for total cortical matter. Dice coefficients of the proposed algorithm were also the highest when compared with previously published standard state-of-the-art brain MRI segmentation algorithms in terms of accuracy in segmenting the gray matter, white matter, and total cortical matter.

1. Introduction

Recent advances in neuropathology have significantly facilitated research into the underlying physiology in the advancement of cognitive impairment. This disorder is related to irregular protein buildup in the cerebrum, which prompts neuronal impairment in the synapses, nerve cells, and axons. Research has shown that anatomical changes start much before any symptomatic indications. The impairment starts in the medial temporal lobe, which contains the entorhinal cortex and hippocampus regions responsible for memory and motion and progresses to the neocortex region responsible for sensory perception, reasoning, and motor commands [1, 2]. The delayed symptoms of dementia are due to dissipation of cognitive reserve in terms of numbers of undamaged neurons, which result in loss of memory function only when decreasing below a certain limit.

1.1. Screening of Dementia. Memory loss related cognitive impairment precedes extensive damage in the temporal lobe, which over a limit is classified as Alzheimer's [3]. Early detection of predementia cognitive impairment through brain scans can facilitate therapy for slowing the progression of Alzheimer's or dementia starting early on. This is the motivation for early screening of cognitive impairment from brain images.

1.2. Brain MRI Cortical Measurements. Past research has demonstrated that assessment of cognitive impairment is feasible from measurement of the change or decrease in size of the cortical and hippocampal regions in anatomical MR images. Additionally, anatomical extraction, segmentation, and measurement of these regions in the medial temporal lobe from brain MR images can differentiate dementia from vascular or neurological degeneration. Furthermore, evaluation of the advancement of dementia via estimation of the rate of atrophy from the measurement of the above-mentioned regions in the medial temporal lobe may be utilized to examine the efficacy of drugs administered to patients for the treatment of Alzheimer's from their MRI scans across several dates [4–6]. Structural anatomical changes in regions of the medial temporal lobe measured from brain MRI scans of patients taken over several dates can be used to estimate the rate of atrophy.

1.3. Measurement of Rate of Atrophy. Brain MR image-based measurement of the rate of atrophy in gray cortical matter is considered a legitimate marker of dementia or cognitive impairment. Atrophy of gray matter is an inescapable result of the degeneration of neuronal cells. The size or volume of cortical tissue is correlated to cognitive function, and the decrease in size or volume is proportional to the degree of cognitive deficiencies. Change in size of regions in the cortex maps to structural and anatomical change such as deposition of neurofibrillary tangle [7, 8] and neuropsychological deficiencies [9, 10]. The first noticeable changes in structural MRI occur along the polysynaptic connections between the hippocampus, entorhinal cortex, and posterior cingulate cortex, due to atrophy caused by excessive protein buildup. This atrophy of the synapses is expected and predictable, given the memory loss in patients in early stage cognitive impairment [11, 12]. At a more advanced stage, neuronal degeneration in temporal, frontal, and parietal lobes results in impairments in speech, motion, and behavior [13, 14]. Complete cerebral [15–19], entorhinal cortical [20], hippocampal [9,21–23], and temporal [24, 25] lobe volume atrophy rates and swelling of ventricular regions [9, 21, 23, 26] both correspond to cognitive impairment, validating their legitimacy as measures of dementia. The utilization of an image-based measure of cognitive degeneration requires that its progression is known at the diverse phases of the disorder and that its association with other imaging-based markers is accounted for. The rate of atrophy varies with the severity of the dementia, from mild cognitive impairment all the way to Alzheimer's disease. In the advanced stages of dementia, anatomical measures have higher sensitivity to atrophy than biomarkers of the proteins imaged or analyzed in cerebrospinal fluid samples [18, 27]. In the presymptom to mild impairment phases, amyloid protein biomarkers show more anomalies than anatomical or structural measures [16,28–33]. Anatomical degeneration occurs at both the macro (tissue level) and micro (axon, dendrites, myelin, and neuron level) scales. These atrophic changes are quantifiable through MR spectroscopy, magnetization transfer, fiber tracking, and diffusion weighted imaging [34–38]. Functional and tissue perfusion MR imaging also can be employed as screening measures [39–42] but require more extensive clinical validation as of present.

1.4. Brain Segmentation. The neocortex of the cerebrum in the MRI scan can be divided into regions after image registration. To perform this, a labeling procedure can be utilized to label each pixel in the cortex to be in one of the regions [43]. Also, clustering based pixel classification algorithms can be used to segment the amygdala and hippocampus regions, which are no neocortical. Algorithms used for the segmentation can either be region growing or pixel clustering using the similarities and disparities in pixel intensities and neighborhood connectivity [44].

1.5. Cortex Thickness Measurements. In patient brain MRI images, the thickness of the cerebral cortex is one of the important parameters used for assessing dementia. The cerebral cortex thickness is measured as the orthogonal distance across the edges between the gray matter and white matter and cerebrospinal fluid. The thickness is measured at each point axially across the full cortical mantle [45–47]. Validation of cortical thickness estimation method has been accepted by means of histological [48] and manual estimations [49]. Cortex thickness measurement is across gray and white matter edges and hence requires gray and white matter segmentation. This paper presents development of an algorithm for automatically calculating the gray and white matter region boundaries postsegmentation, after which computation of the thickness and volume of the cortex for assessing dementia or cognitive impairment in patients can be done. Future work would entail validating the accuracy of the cortex thickness measurements using distance across the boundaries and testing for robustness of the algorithm over varying image acquisition systems with changing scanner type, signal to noise ratio, and number of MRI slices captured [50, 51].

2. Background and Motivation

Nowadays, with the increase in patients with brain abnormalities, analyzing the patient's brain MRI images by extracting diagnostic features and other clinical information is the most challenging task for doctors or neurologists in the field of biomedical image processing. This work presents automatic segmentation of gray and white matter regions as anatomical features in brain MRI images. Changes in the size or volume of these regions can be correlated to changes in cerebral structure in patients with Alzheimer's, dementia, cognitive impairment, or other neurological disorders.

Segmentation of MRI images is used in many biomedical applications to effectively measure and visualize the patient's brain anatomical structures [51]. An important aspect in analyzing the brain MRI image is extracting gray and white matter regions, tumors, or lesions, which is possible through the segmentation process. Figure 1 below shows the segmented white matter (green boundary) and gray matter (blue boundary). After segmentation of a diseased patient's

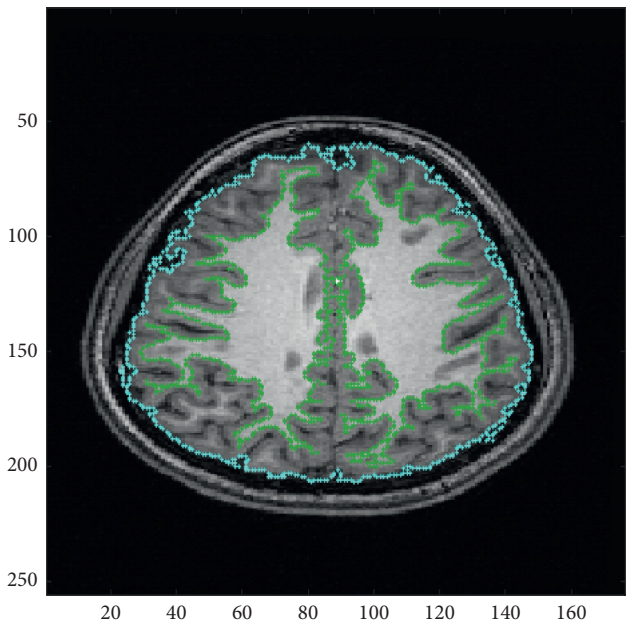


Figure 1: Segmented white matter (green boundary) and gray matter (blue boundary). Gray matter consists of the cortex, and its size can be measured after segmentation of the gray matter.

MRI image, the data extracted from the multidimensional image give the information about the tumor size, type (benign or malignant), and position. This can facilitate and assist neurologists in treatment planning [52].

Initially, manual segmentation techniques were used by neurologists which are time consuming and vulnerable to human errors. Therefore, several techniques were introduced for segmentation of MRI images into regions of interest. They are classified as threshold-based, region-based, pixel classification-based, and model-based techniques. These fully automatic segmentation methods are determined by the computer without any human intervention.

In this research work, automatic segmentation of regions in cerebrum is performed using adapted fuzzy c-means algorithm (FCM), which is one among various pixel classification techniques, combined with connected component analysis. FCM is one amongst many predominantly used techniques for tumor segmentation and other regions in especially brain MRI scans as it gives efficient results while analyzing nonhomogeneous tumored brain MRI images [53]. This is a unique method that can also be used for noisy image segmentation to produce efficient results. Fuzzy c-means clustering is grouping similar data objects or components within the same cluster and dissimilar data objects within other clusters. In biomedical image processing, the term data object is nothing but pixels of an image. The same concept is implemented in this research to build a structured framework to automatically segment these cerebral regions in multidimensional brain MRI images.

Segmentation of various brain tissues is an important aspect to analyze brain image data, study patient's anatomical structure, and assist neurologists in treatment planning. Segmentation has various real-time applications such as data compression and visualization that helps neurologists to provide patient's information for surgical planning. This process of brain segmentation identifies regions of interest such as tumors, lesions, and other abnormalities. It can also be used to measure the increase or decrease in volume of tissue to measure growth of a tumor [6]. Magnetic resonance imaging (MRI) and computed tomography (CT) technologies to generate scans of internal brain structures have been increasingly used nowadays to detect tumor or any other abnormalities in human brain. These technologies make it almost compulsory for any neurologist or radiological experts to use computers in the field of medical sciences. The major goal of brain MRI segmentation is to separate the brain image into a set of important, meaningful, similar, and nonoverlapping regions having identical properties such as texture, color, intensity, or depth. The result obtained is the segmentation of each homogenous regions which are identified by labels also describing the region boundaries [7]. A typical MRI image study of one patient may require 100 or more images to be analyzed. This would be a tedious task for neurologists who have knowledge in the field to perform manual segmentation for each of the 100 images.

Nowadays, MRI imaging is used in many medical applications, especially for brain imaging to obtain clinical information and analyze patient's data. It is because, MR imaging is efficient and produces accurate results while detecting brain abnormalities of patient's brain during initial stages of any disease when compared to a CT scan. This increase in the use of MR imaging led to introducing many unsupervised automated segmentation techniques that enable managing and analyzing huge data of a patient which are in the form of an image.

Based on the repetition time (TR) and time to echo (TE), MRI scans are classified into two different sequences for scans. These scans are named as T2-weighted and T1-wieghted scans. These scans are generated depending on the time of echo (TE) and repetition time (TR) values. T2-weighted images are obtained by longer TE and TR times whereas T1-weighted images are obtained by shorter TE and TR times. The brightness and contrast of these scans are determined by T1 and T2 properties of brain tissue accordingly. The human brain contains tissues with large amounts of fat content that appear bright in MRI images. The parts of the brain which are filled with fluid appear dark in the MRI image. In our research, T1-weighted images are used because of high resolution and clarity [54].

Since the last decade, many researchers have developed advanced technologies in the field of brain MRI segmentation to detect tumors or segment brain MRI images. Even though many algorithms exist, they are not available as software packages or downloadable software and thus inaccessible to medical researchers, neurologists, surgeons, or doctors in the hospital. Even those implemented in software packages are expensive and only affordable to high-end hospitals, or do not offer the feature of automatic segmentation [55–71], or are not easy to use. However, in this, we present and publish a free-to-use graphical computational software tool that automatically performs the brain

MRI image segmentation as a stand-alone application with a user-friendly easy-to-use graphical user interface and functions as a neurological disease prediction framework and disease detection tool. It is freely available to any medical student, academician, researcher, technician, nurse, doctor, neurologist, or surgeon in any country in any part of the world who accesses this paper. It is packaged in a stand-alone independent GUI, which can load medical images in any format (NIfTI, DICOM, PNG, TIF, JPG, etc.) and help neurologists to perform various automatic segmentations to analyze the patient's data. Specifically, the thickness of the cortex plays an important role in determining the severity level of dementia or cognitive impairment.

The work herein presents a method using the gray-to-white matter thickness ratio computed from the brain MRI slices of the patient as part of the development of a software platform-based computational tool for aiding neurologists in assessing anatomical and functional changes in cerebral structure from brain MRI scans of neurological patients. This GUI also enables user to perform various other actions like segmentation of brain MRI images as masks, segmented regions, or boundaries.

2.1. Aims and Objectives. The aims and objectives of this research paper are listed below:

(1) To develop an automatic brain segmentation tool that can be used by neurologists for analyzing patient's brain image data

(2) To predict neurological disease using automated segmentation to extract clinical information from the images

(3) To compare automatic segmentation and manual tracings performed by experts for validation purpose

2.2. Step-by-Step Procedure. The stepwise procedure of this research paper is defined as follows:

(1) Perform fully automatic segmentation of gray and white matter regions in brain images for disease prediction

(2) Build a graphical computational tool for assisting neurologists

(3) Validation of automatic segmentation with manual tracings by experts

3. Pixel Classification Techniques

3.1. Clustering Algorithms. Clustering is the grouping of objects into different clusters. In other words, the set of data is divided into subsets. Each subset should have some common property like distance, size, etc. According to the similarity measures of these data subsets, they are assigned to similar clusters. There are various clustering techniques such as fuzzy *c* clustering, each of which has their own benefits.

3.1.1. K-Means Algorithm. The *k*-means method is one of the most widely used clustering-based algorithm for image processing. In this algorithm, an image dataset is considered which is divided into subsets or group of data. Each group of data is called cluster which is partitioned accordingly. Each cluster will have data members and cluster centroid. A point in the cluster is defined as a centroid if it has minimized sum of distances from all the data members to that point. This *k*-means is a repetitive and iterative algorithm because of which can minimize the sum of distances from all the data members to centroid and over all other clusters of the dataset. Let us assume an image data that has $a * b$ resolution and k be the number of clusters of that image data. Also, the pixels of the image be $P(a, b)$ and c be the center point of the cluster [70, 71]. The *k*-means algorithm can be determined as follows:

After initializing the number of clusters and centroid of each cluster, compute the Euclidean distance with below formula:

$$\text{Euclidean distance} = |P(a,b) - C(k)|. \tag{1}$$

In equation (1), $P(a, b)$ is the input pixel at data member point (a, b) of the input image, and $C(k)$ as in equation (2) is center for kth cluster.

After the calculation of distance from each pixel, determine the nearest center to all the pixels and assign the pixels to the center based on the calculated distance. Next step after assigning the pixel is to calculate again the center position of the kth cluster using the following formula:

$$C(k) = \frac{1}{K} \sum P(a,b). \tag{2}$$

This process of computing position of centroid is repeated iteratively until error value or tolerance value is satisfied. *K*-means clustering is easy to implement and simple to understand but it also has some backlogs because of poor quality of final segmentation as the centroid value here depends on the initial value selected. This algorithm may sometimes fail as the initial value is based on the human assumptions. Therefore, many other algorithms are introduced to overcome these drawbacks.

3.1.2. Fuzzy c-Means Algorithm. Fuzzy *c*-means clustering algorithm is the one among the most widely used methods in which the dataset is classified into clusters having similar data objects. That is, each cluster will have similar type of pixels [72]. This classification into clusters is based on the intensity values of pixels. Therefore, similar pixels are grouped into similar clusters. In this algorithm, each pixel may belong to one or more clusters unlike in *k*-means algorithm. Each pixel in the image dataset will have membership value that determines the degree of share of that pixel or data point on every cluster of that image. From this, we can build a membership matrix that has all the membership values of all the pixels of all the clusters of that image. Also, we can define the fuzzy *c*-means algorithm in other words as it processes segmentation using unique pixel classification technique in assumption that each pixel may be

allowed to be present in one or more classes with value of membership that is between 0 and 1. Assume a dataset of s number where $X = \{x_1, x_2, \ldots, x_n\}$. This algorithm divides the dataset into group of fuzzy clusters according to some criteria or some condition. This grouping of data into clusters is an iterative and continuous process till all the pixels are given at least one membership of clusters based on some objective function. Given below is the objective function of fuzzy c-means clustering algorithm:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^m \left\| x_i - c_j \right\|^2. \quad (3)$$

In equation (3), m here is a fuzzy parameter which defines the fuzziness of the clusters and u_{ij} as in equation (5) is the membership degree of cluster C_j which is the center of the cluster as in equation (4). The first step of the algorithm for fuzzy c-means clustering is to specify the number of clusters of the dataset and the matrix for the membership function of all data members of the dataset [73]. The next step is to compute the center of each cluster using the formula below:

$$C_j = \frac{\sum_{j=1}^{n} u_{ij}^m x_i}{\sum_{j=1}^{n} u_{ij}^m}. \quad (4)$$

After the center calculation, one should determine the error or cost value and evaluate if it is less than the threshold value so that to improve the previous iteration of the function. If the error value is satisfactory then it is further processed to cluster the data. If the error value is not satisfactory, membership matrix is continuously updated till the results are satisfactory to obtain final segmentation with improved level of quality. Below is the condition to compute the relation with membership function:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left[d_{ij} / d_{kj} \right]^{(2/(m-1))}}. \quad (5)$$

There are many other segmentation algorithms among which this fuzzy c-means algorithm is more suitable to analyze patient's data through segmentation process. In this research work, we use an adaptive fuzzy c-means clustering algorithm for segmentation of gray and white matter regions in brain MRI images.

4. Brain MRI Segmentation

Past literature presents reduction (measured as atrophy rate) of cortex volume as a valid measure for dementia from patient MRI scans. The estimation of atrophy rate requires measurement of the gray and white matter regions in the brain MRI images of the patient. In the proposed method, the gray and white matter are automatically segmented using a form of adaptive modified pixel clustering methods such as k-means or fuzzy c-means clustering, which will cluster the pixels by labeling them (based on their intensities) to belong to the gray matter, white matter, cerebrospinal fluid, or background [74]. The adaptive clustering methods are modified by running them separately for the gray and white matter and postprocessing with connected region labeling to separately label the gray and white matter regions.

4.1. Image Acquisition. The patient's brain MRI image and neurological data used in this research work were obtained from the Image and Data Archive (IDA) powered by Laboratory of Neuro Imaging (LONI) provided by the University of Southern California (USC) and also from the Department of Neurosurgery at the All India Institute of Medical Sciences (AIIMS), New Delhi, India. The data were anonymized as well as followed all the ethical guidelines of the participating research institutions.

4.2. Segmentation Methodology. The methodology for segmenting the gray and white matter used in this research is illustrated in Figure 2. The first step is the removal of the skull outline from the brain MRI images with the Hough transform. Fuzzy c-means clustering is next applied on the skull outline removed brain MRI image slice to obtain separate clustered image slices for the gray and white matter regions. These clustered gray and white matter images are divided into connected regions using connected component labeling. The largest two connected regions are heuristically the gray and white matter regions. The binary extracted gray and white matter images can be used as masks, which when applied to the original brain MRI image produces the final segmented gray and white matter regions with the original pixel intensities [75]. The skull outline removal using the Hough transform is shown in Figure 3. The detected skull outline is removed to obtain only the cerebral cortex in the MRI image slice. This cerebral cortex image slice is used in the fuzzy c-means clustering step of the procedure.

In this paper, we present a framework for neurological disease prediction and decision making for patients of cognitive impairment, dementia, or Alzheimer's disease based on automatic segmentation of gray and white matter regions as anatomical features in brain MRI images. Changes in the size or volume of these regions can be correlated to changes in cerebral structure in patients with Alzheimer's, dementia, cognitive impairment, or other neurological disorders. Specifically, the thickness of the cortex plays an important role in determining the severity level of dementia or cognitive impairment [76]. The work herein presents a method using the segmentation of gray and white matter from the brain MRI slices of the patient as part of the development of a software platform-based computational tool for aiding neurologists in assessing anatomical and functional changes in cerebral structure from brain MRI scans of neurological patients. The aforementioned tool can be implemented as a software package that can be installed in the computational platforms in the neurology department or division of hospitals. In its final implementation and deployment, this tool would predict neurological disease type and severity after automatically processing the brain MRI or CT images with the abovementioned algorithms and displaying the highlighted gray and white matter regions in the brain CT or MRI images [77].

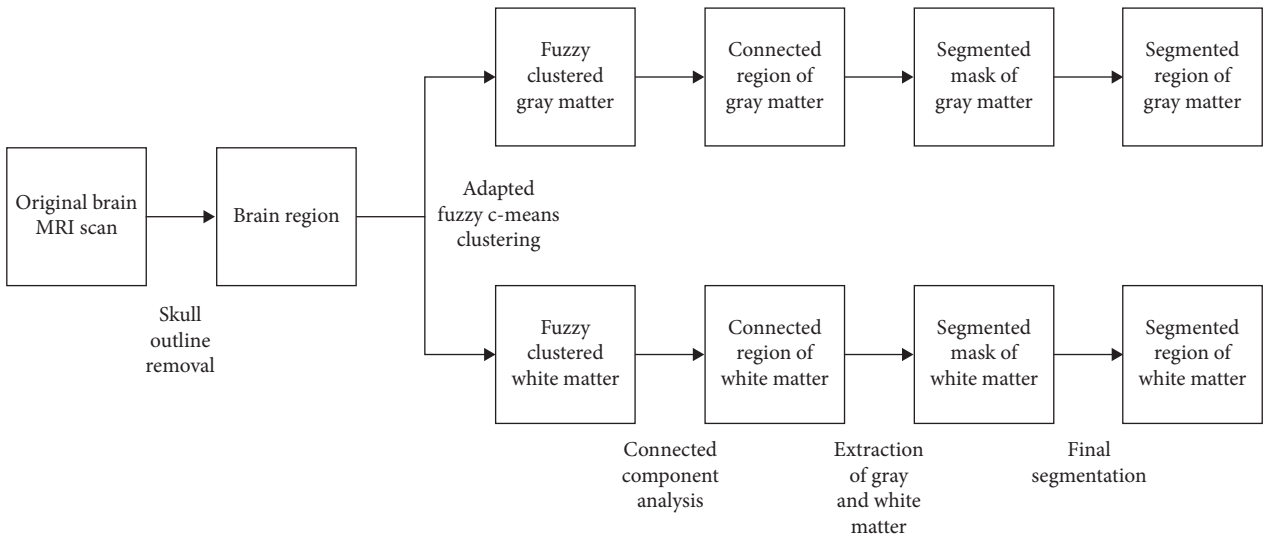


Figure 2: Block diagram of this paper's proposed fully automatic brain MRI gray and white matter segmentation procedure.

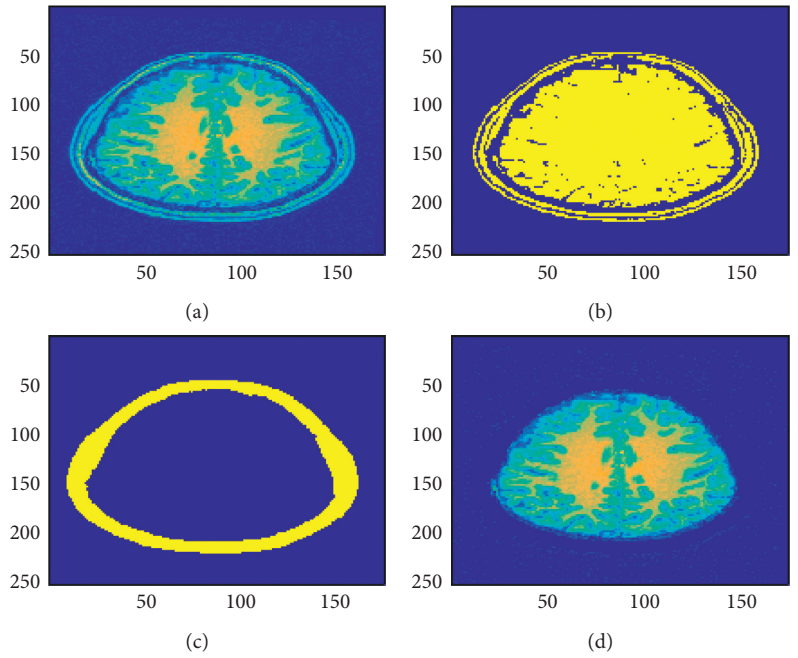


Figure 3: Skull outline detection in brain MRI images: (a) original MRI image slice; (b) thresholded MRI image slice; (c) detected skull outline; (d) skull outline removed.

In the field of medical image processing, the most challenging task to any neurologist or a doctor or a scientist is to detect the patient's disease by analyzing the patient's clinical information. Patient's data is extracted and analyzed to detect the abnormalities and to measure the illness of the disease which helps a medical practitioner to cure the disease at its early stages [78]. Extraction of brain abnormalities in brain MRI images is performed by segmentation of gray and white matter regions in patient's brain MRI images. After segmentation is performed, patient's clinical data such as the area of the cortex, size of tumor, type of tumor (malignant or benign) and position of tumor are determined which helps a doctor to take early decisions for surgery or treatment to cure any brain disease.

During initial days, these segmentation techniques were performed manually by subject matter experts or neurological experts which consumes time and effort of neurological specialists in the field. The segmentation results obtained from the manual segmentation techniques may not be accurate due to vulnerable and unsatisfactory human errors which may lead to inappropriate surgical planning. Therefore, it has become very much necessary for a neurologist or an academician or a researcher to introduce automatic segmentation [79, 80] techniques, which give

accurate segmentation results. These segmentation techniques that are performed automatically are of two types typically known as semiautomatic and fully automatic segmentation techniques. In a semiautomatic segmentation process, partial segmentation is performed automatically and then the results thus obtained are checked by neurological experts to modify for obtaining final segmentation results. In a fully automatic segmentation technique, there is no need for manual checking by neurological experts which minimizes his time and effort. These fully automatic segmentation techniques are classified as threshold-based, region-based, pixel classification-based, and model-based techniques which are determined by the computer without any human participation.

This research work presents the segmentation of various regions that are segmented automatically using a technique called fuzzy *c*-means algorithm (FCM), which is a pixel classification technique followed by component labeling technique which is used widely in biomedical image processing to perform fully automatic segmentation in brain MRI images [81].

Over the past few years, a set of techniques were introduced for automatic image segmentation among which fuzzy *c*-means (FCM) clustering method yields both gray matter and white matter regions more homogenously which can efficiently remove noisy spots when compared to other segmentation techniques. Figure 2 shows the detailed description of the segmentation process as a block diagram.

Therefore, this technique can be used to segment noisy brain MRI images obtaining accurate, reliable, and robust results. Also, unlike other techniques, this can be used for both single-featured and multifeatured information analysis with spatial data. This automated unsupervised technique can be used to perform segmentation to achieve feature analysis, clustering, and classifier designs in fields of astronomy, target recognition, geology, medical imaging, and image segmentation [9]. A set of data points constitutes to form an image that has similar or dissimilar regions. This algorithm helps to classify the similar data points into similar clusters by grouping them based on some similarity criteria. In medical image processing field, image pixels are highly correlated as they may have same characteristics or feature data to its next or immediate neighbor. In this method, spatial information of neighboring pixels is highly considered while performing clustering. This paper presents a technique for clustering of brain MRI image slices into different classes followed by component labeling using knowledge-based algorithm. The steps in the fully automatic segmentation algorithm are as follows.

4.3. Skull Outline Detection. The preliminary step in our research is to extract the skull outline from an MRI image slice as it is not our region of interest. Also, these quantitative studies especially in living organisms of brain MRI images usually will have a preparatory processing in which the part of the brain itself is isolated from the external brain regions and no-brain tissues which are not required for brain analysis. This process of skull outline detection and removal is called skull stripping. This helps us to focus more on the actual brain itself [10]. In this stage, many superfluous and nonbrain tissues such as fat, skin, and skull in brain images had been detected and removed using Hough Transform which is an image feature extraction tool in digital image processing. This Hough transform technique for skull outline detection helps to find unwanted points or data objects of an image with different shapes such as circular and elliptical using voting procedure in a parameter space. These generalized Hough transform techniques are used to detect an arbitrary shape at a given position and scale. In this technique, in a parametric space of an MRI image, parametric shapes are detected by tracing the acquisition of various points in the space. If in an image a shape like circle and elliptical exists, all its points are mapped in the parametric space grouping them together around the parametric values forming clusters which correspond to that shape [11]. The result obtained in this step is shown in Figure 3.

4.4. Adaptive Fuzzy c-Means Clustering. After the skull outline detection and removal, internal part of the brain is clustered into different regions. Clustering is a well-known and widely used technique for pattern classification and image segmentation purposes in the field of medical sciences. In this process, similar data objects or pixels are grouped into similar clusters. Usually, medical images tend to have more noise due to its internal and external factors. During the segmentation process, the medical images having noise generate inefficient results, and it is difficult to analyze anatomical structures of patient's brain [12]. This may lead to inappropriate diagnosis and treatment planning. Therefore, to avoid inaccurate results during segmentation process, several types of image segmentation techniques were introduced by the researchers and neurologists to achieve accurate results during segmentation of regions in an MRI image of a patient. These techniques can perform segmentations equally for noise MRI images [13–18]. Among them, fuzzy *c*-means clustering methods are widely used techniques in MRI segmentation as they have substantial advantages comparatively because of uncertainty present in brain MRI image data. To enhance features of fuzzy *c*-means algorithm, in our research, adaptive fuzzy *c*-means clustering algorithm is used as it minimizes computational errors [19].

4.5. Connected Component Labeling. In the next step, the clustered image is subjected to connected component labeling based on connectivity. Deriving and labeling positions of several disjoint and connected components in brain MRI image is a very essential step in segmentation process [20]. In any medical image, pixels which are positioned together as connected components will have similar values for their intensities. Connected component labeling method scans the image, pixel-by-pixel to first detect the connected component pixels and then it extracts connected pixel regions which are adjacent to one another. These pixels which positioned together will have same set of intensity values [21–25]. After all groups have been extracted, each pixel component is labeled according to component it was assigned to. In our research, we use 8-connectivity measures for connected component labeling.

4.6. Final Segmentation Mask after Removing Noise. The final step is to obtain actual segmented gray and white matter regions by overlaying gray matter and white matter masks on original MRI image to remove all pixels which background and only keep the pixels in the foreground or regions of interest in the original image [26]. This method enhances the distinction of gray and white matter regions and allows more accurate segmentation results. The algorithm presented herein works for gray and white matter segmentation as well as tumor segmentation in brain MRI images. Figure 4 below shows the results on a sample patient specimen brain MRI image obtained from the abovementioned fuzzy *c*-means clustering followed by the connected component labeling to extract the cerebral regions as masks [27, 28]. When these masks are applied to the original image, final gray and white matter regions segmentation or tumor segmentation results are obtained. The results thus obtained are shown in Figure 4 below for a normal patient brain MRI image. As this method is also applicable for tumor segmentation, Figure 5 shows the results of tumor segmentation applying this work's proposed algorithm on a tumor brain MRI image.

The segmentation results for a brain tumor patient's brain MRI images are shown below. The figures below show a sample brain MRI image of a patient brain with a tumor. These figures demonstrate that the algorithm developed herein for detection of gray and white matter regions works well for tumor detection and segmentation of the tumor section in a patient's brain as well. As mentioned earlier, in our segmentation methodology after skull outline detection, we perform adapted fuzzy *c*-means clustering followed by the connected component labeling to extract the gray and white matter regions as masks for gray and white matter segmentation or to extract the brain region and tumor regions as masks for tumor segmentation and identification.

The results of the automatic segmentation algorithm for tumor identification and segmentation on a sample patient's tumor brain MRI image are shown below in this section. The first step was skull outline removal (see Figure 6), and the final segmentation results of this brain tumor MRI image are shown in Figure 5.

Table 1 shows the comparison of different brain MRI segmentation methods [81, 82] based upon pixel classification and clustering classified by the region of interest being segmented.

5. Segmentation Tool

To process, extract and analyze the patient's image data, a neurologist or a researcher requires a computational tool that can perform all the required functions automatically minimizing the cost, effort, and time. These software tools are widely used nowadays in almost all the hospitals to detect patient's disease by analyzing patient-specific information and to provide patient-specific medical care at early stages of the disease [29]. These days, software engineers and programmers have been actively developing tools which are used in medical fields to assist neurologists, scientists, doctors, and academicians to analyze patient specific information. This research work herein presents an independent standalone graphical computational tool which is developed for assisting neurologists or researchers in the field to perform automatic segmentation of gray and white matter regions in brain MRI images [30, 31]. This software application is built using a neurological disease prediction framework for diagnosis of neurological disorders like dementia, impairment, brain injury, lesions, or tumors in patient's brain. This tool provides the user to perform automatic segmentation and extract the gray and white matter regions of patient's brain image data using an algorithm called adapted fuzzy *c*-means (FCM) [32]. In this research work, we also present the methodology used to obtain segmentation, in which patient's images are subjected to fuzzy *c*-means clustering followed by connected component labeling technique.

The entire process of feature extraction, classification, preprocessing, and segmentation [33] is developed as a graphical computational tool with a user interface (GUI). This application built is a stand-alone graphical user interface (GUI) that will load the brain MRI images from the local computers of neurologists on the click of a button and then segment out [34–37] the gray and white matter regions in the brain MRI images upon just the click of buttons and display the results as a mask, color images, or as the boundaries of those two cerebral regions. The developed GUI system assists neurologists or any user making it easy to upload patient's brain image from his local computer, viewing and obtaining the results in very less time, reducing efforts due to manual tracings by the experts [38–42]. The GUI has the following features:

(1) Automatized segmentation of brain MRI images is provided as a stand-alone independent software package.

(2) It is freely accessible to all researchers in the medical field and neurologists, radiologists, and doctors in any part of the world.

(3) It is user-friendly and easy to use.

(4) It automatically segments the brain images and so no manual tracing is required by the user. This tool allows timely efficient segmentation of the brain MRI images so that the neurologists' or neurosurgeons' precious time is used efficiently and not wasted on manual segmentation.

(5) It is developed to support several medical image datatypes (NIfTI, DICOM, PNG, etc.).

(6) Neurological disease prediction framework can be provided in this software tool.

(7) The tool was developed in collaboration with neurosurgeons and neurologists at the All India Institute of Medical Sciences (AIIMS) and hence it has the expert neurological feedback and opinion of doctors implemented in it.

Below are the three screenshots which show running the GUI for loading the brain MRI image (Figure 7), viewing the gray and white matter segmented regions (Figure 8), viewing the gray and white matter extracted masks (Figure 9), and viewing the gray and white matter region boundaries (Figure 10).

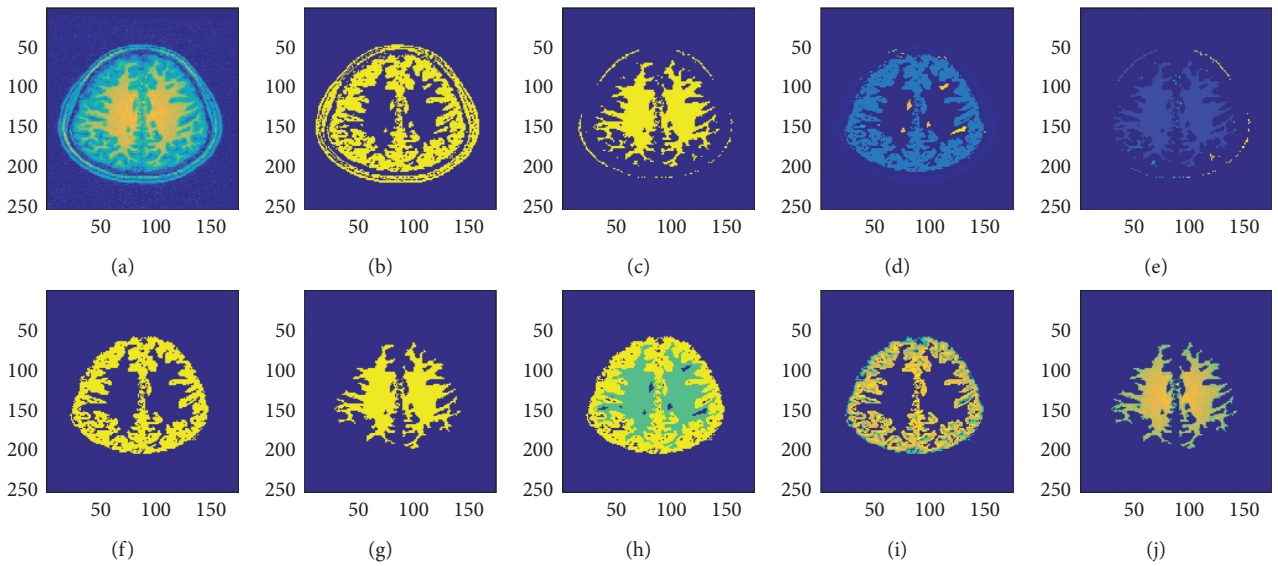


Figure 4: Fully automatic gray and white matter segmentation in brain MRI images (for a sample patient specimen image). (a) Original MRI frame. (b) Fuzzy gray matter. (c) Fuzzy white matter. (d) Connected gray matter. (e) Connected white matter. (f) Segmented gray matter. (g) Segmented white matter. (h) Gray and white matter. (i) Gray matter mask. (j) White matter mask.

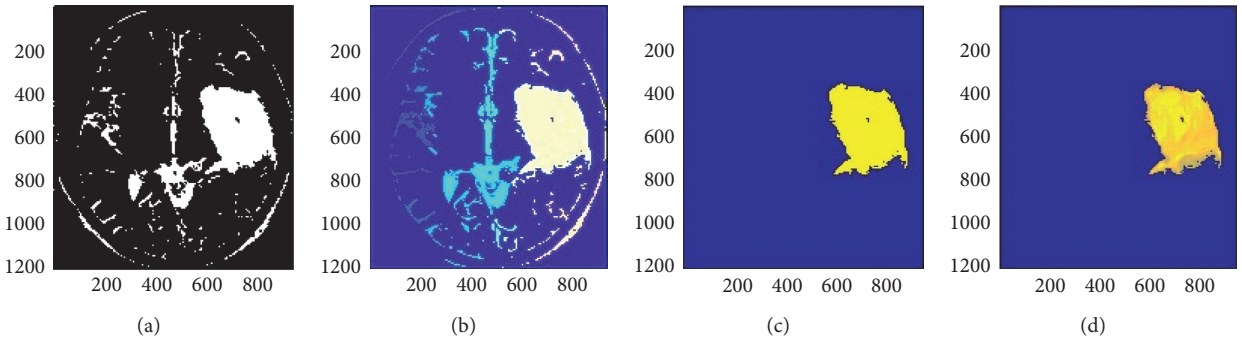


Figure 5: Tumor in brain region segmentation in a sample tumor brain MRI image. The brain MRI image after performing fuzzy c-means and connected regions operations is shown along with the final segmented tumor region and mask using the fully automatic procedure for tumor segmentation from the brain segmentation. This shows that the method proposed in this paper successfully works for tumor segmentation and identification along with gray and white matter segmentation. Thus, brain tumor segmentation is another application of this paper's proposed algorithm along with gray and white matter region segmentation. (a) Fuzzy tumor region. (b) Connected tumor region. (c) Segmented tumor region. (d) Tumor region mask.

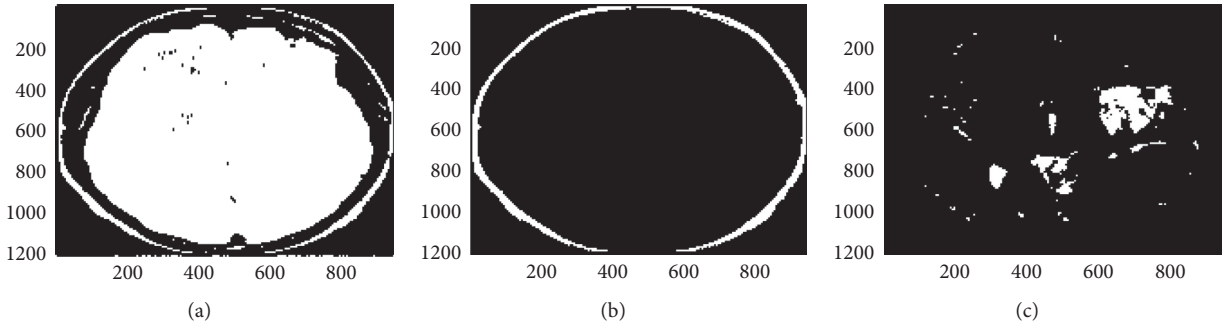


Figure 6: Skull outline detection in brain MRI image with tumor. (a) Threshold MRI image Slice. (b) Detected skull outline. (c) Skull outline removed.

Table 1: Comparison of different brain MRI segmentation methods [81, 82] along with method proposed by the authors [83] based upon pixel classification and clustering classified by the region of interest being segmented.

Region of interest	Method	Procedure
Brain tumors	k-means + fuzzy c-means	Pixel intensity k-means followed by pixel intensity and membership-based fuzzy c-means clustering with preprocessing using median filters and postprocessing using feature extraction and approximate reasoning
Brain lesions	Fuzzy c-means with edge filtering and watershed	Pixel intensity and membership-based fuzzy c-means with preprocessing using thresholding techniques and postprocessing using edge filtering and watershed techniques
Gray and white matter regions	Adaptive fuzzy c-means (proposed method in this work)	Pixel intensity and membership-based fuzzy c-means clustering with preprocessing using elliptical Hough transform and postprocessing using connected region analysis

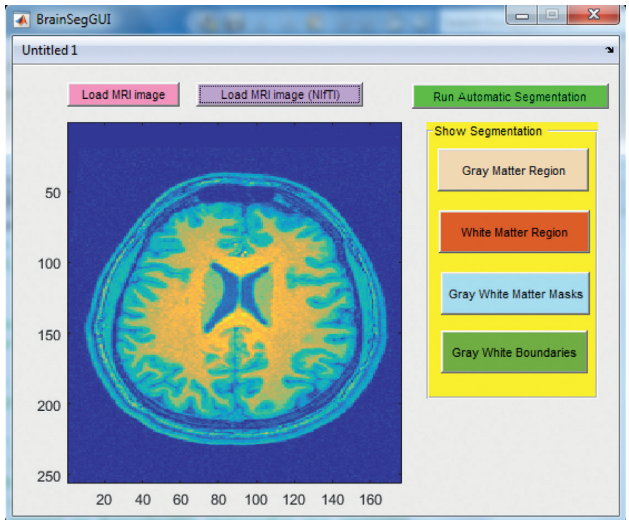
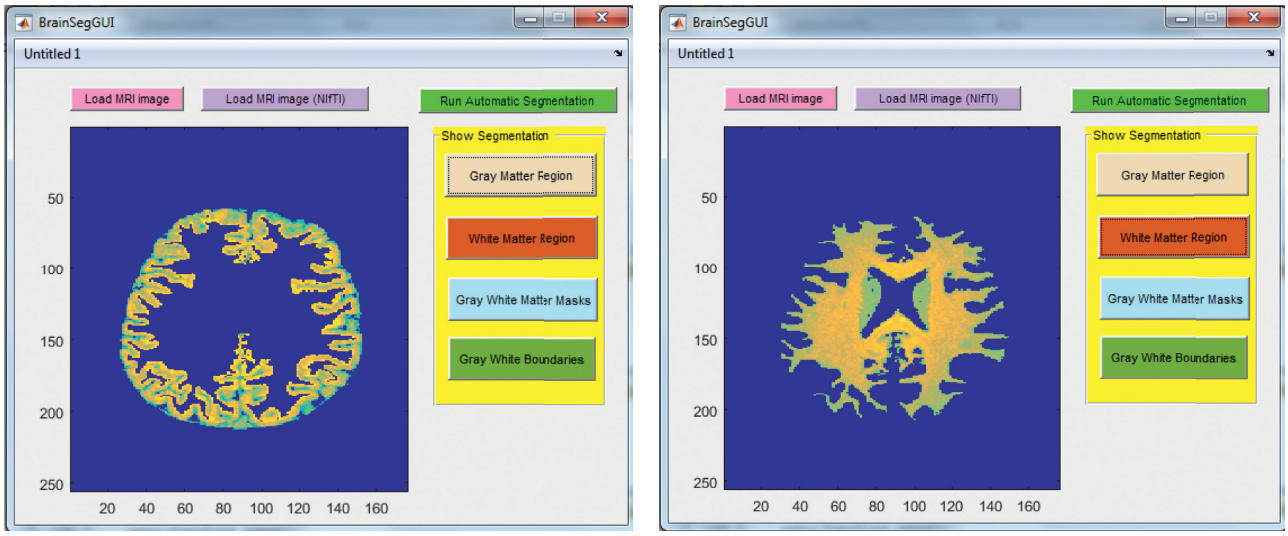


Figure 7: Screenshot of the graphical user interface (GUI) designed and developed in this work for automatic brain MRI image processing. Step shown here is to load the MRI image (NIfTI in this case) upon the click of the "Load MRI image" or "Load MRI image (NIfTI)" button depending upon the image type.



(a) (b)

Figure 8: Screenshots of the graphical user interface (GUI) designed and developed in this work for automatic brain MRI image processing. Steps shown here are to show extracted gray (a) and white (b) matter regions upon the click of the "Gray Matter Region" (a) and "White Matter Region" (b) buttons, respectively.

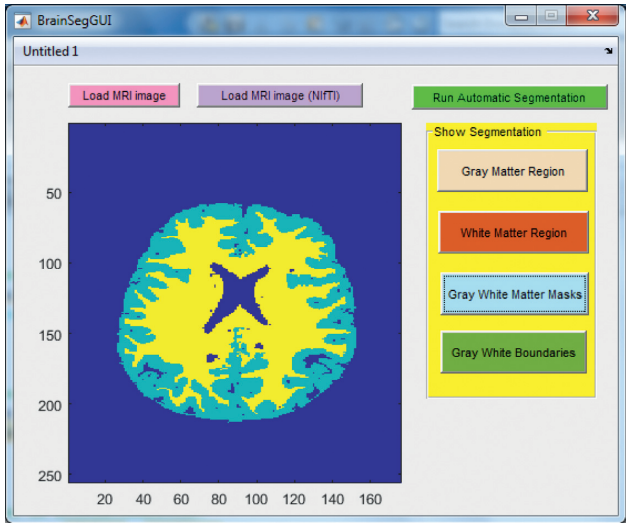


Figure 9: Screenshot of the graphical user interface (GUI) designed and developed in this work for automatic brain MRI image processing. Step shown here is to show the gray and white matter masks upon the click of the "Gray White Matter Masks" button.

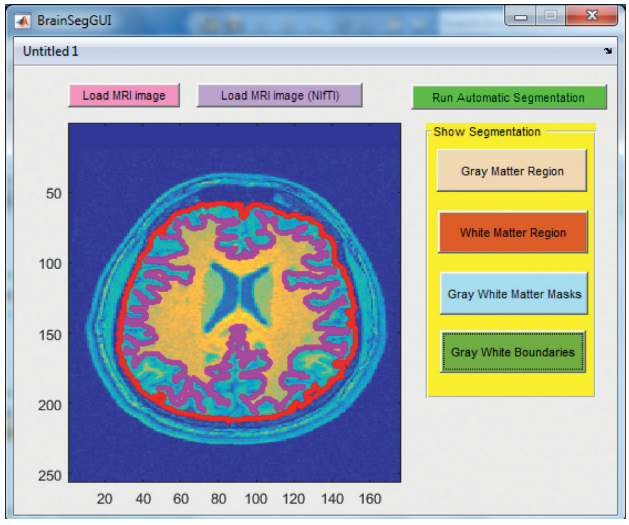


Figure 10: Screenshot of the graphical user interface (GUI) designed and developed in this work for automatic brain MRI image processing. Step shown here is to show the gray matter boundary (shown as a red colored contour) and white matter boundary (shown as a magenta colored contour) superimposed on the original brain MRI image upon the click of the "Gray White Boundaries" button.

6. Manual Segmentation

In this section, the accuracy of the proposed automatic segmentation methodology of the white and gray matter regions was validated against manual neurological tracing-based segmentation by experts. The validation of the automatic segmentation of gray and white matter regions in patient brain MRI images using adapted fuzzy c-means clustering followed by the connected labeling is done by verifying against the manual segmentation by neurologist experts shown in Figure 11.

We have also performed validation of the automatic segmentation of gray and white matter and tumors in tumor brain MRI images using adapted fuzzy c-means clustering combined with the connected component labeling and this is validated by the manual segmentation by experts, an example of which is shown in Figure 12.

7. Validation

This validation compares the manual and automatic segmentation of five patient brain MRI images statistically using the Dice coefficient as a similarity measure [79, 80, 84–87]. Figures 13, 14, and 15 show the sample manual and automatic segmentation of three of the patients. For this purpose, a total of five MRI scans of different patients were used to validate the automatic segmentation proposed in this paper by comparison against manual segmentation by neurological experts for each patient's MRI image by calculating the [89–95] Dice coefficient between the automatic and manual

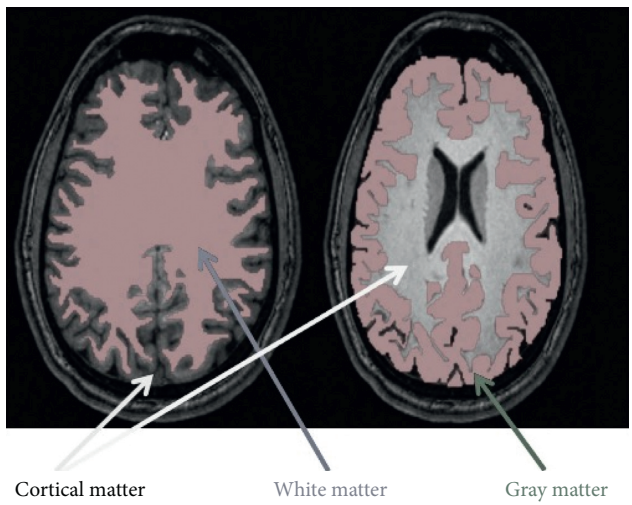


Figure 11: Sample manual segmentation (labeling) by neurologist expert of the gray and white matter regions in brain MRI images: white matter region (left) and gray matter region (right).

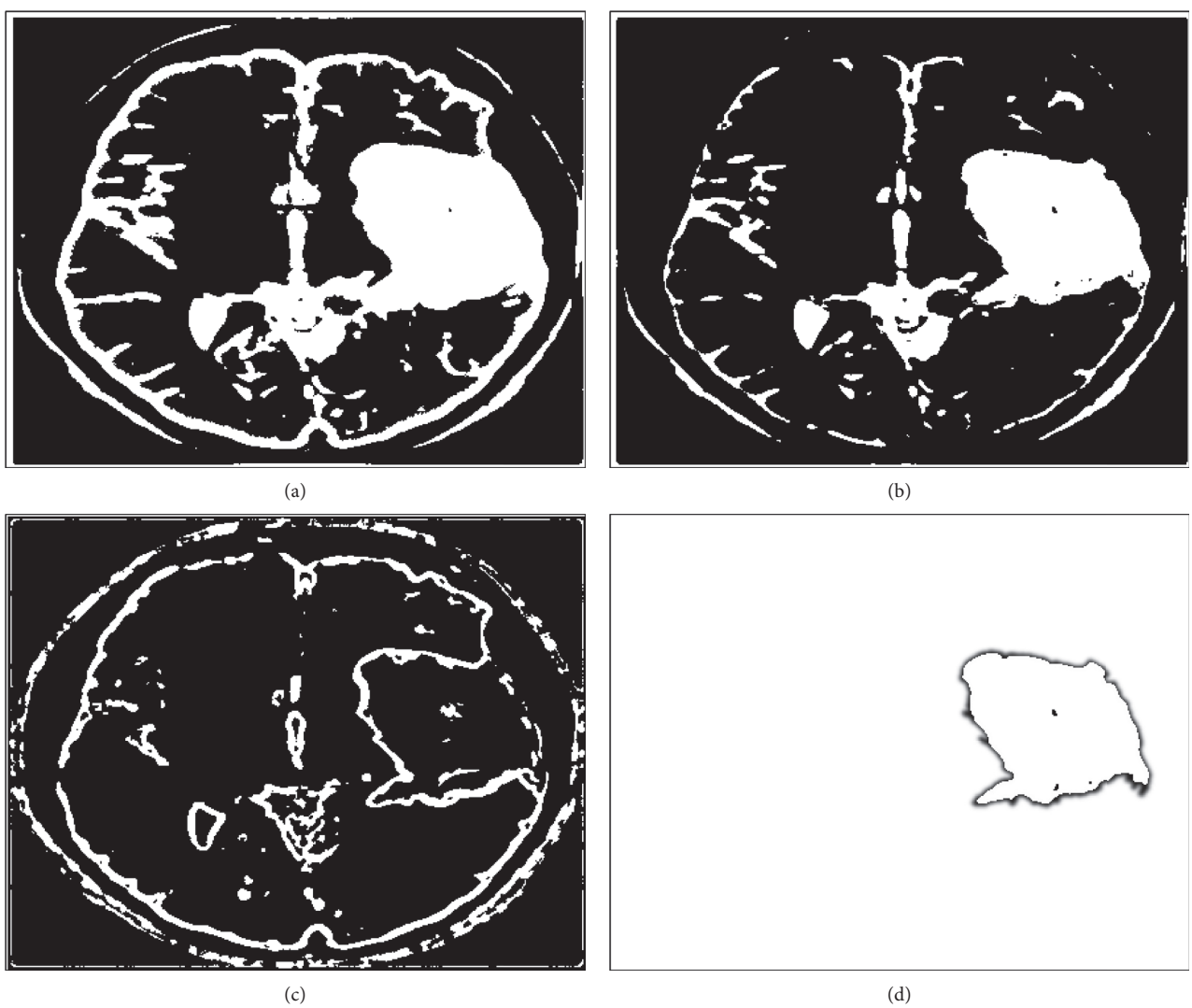


Figure 12: Example of steps in segmentation (tracing) by expert of the gray and white matter regions in brain tumor MRI images in a sample patient brain MRI image.

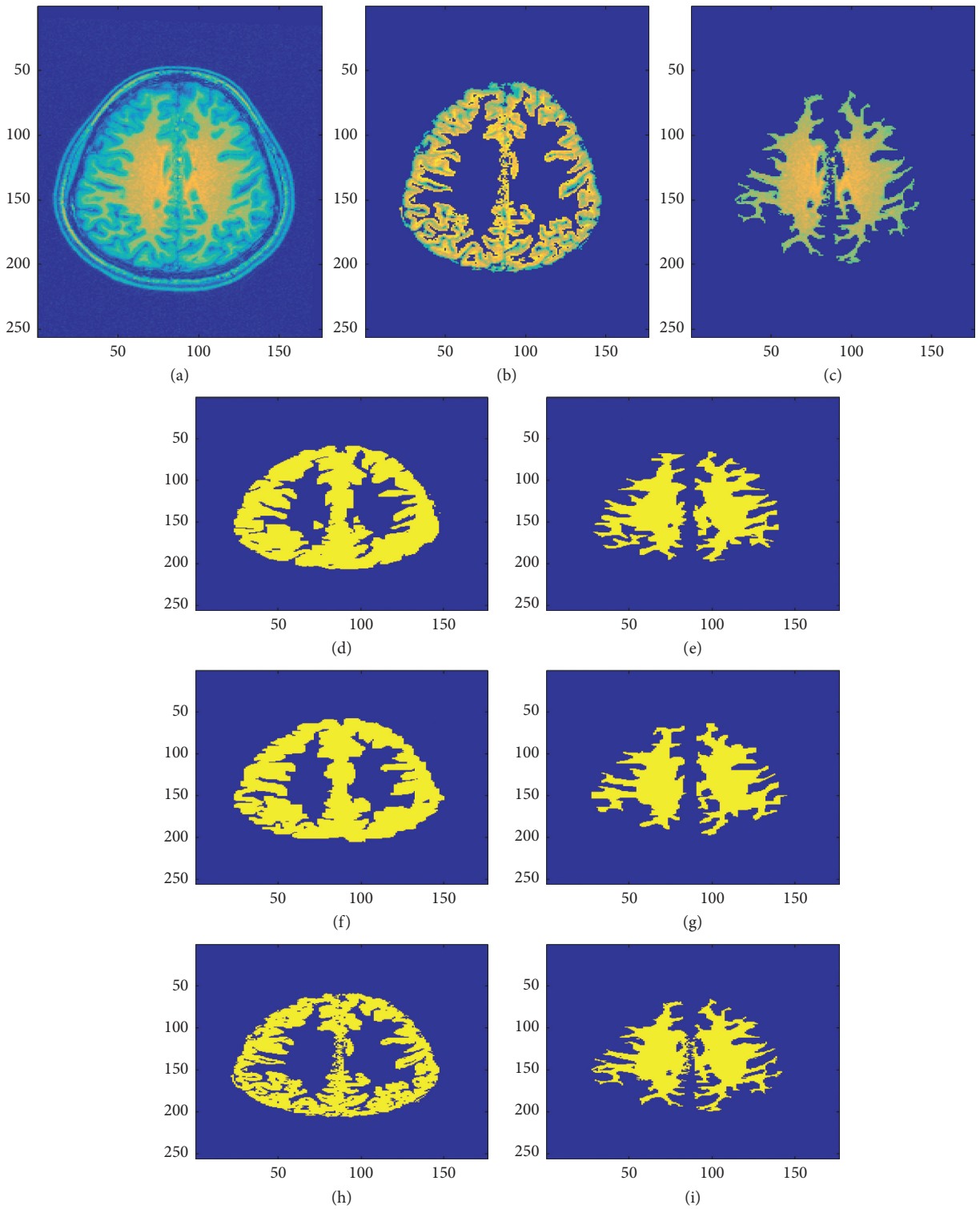


Figure 13: Visual comparison of two manual expert tracing-based and automatic segmentation (using the fully automatic segmentation method presented in this paper) results of sample patient 1 brain MRI image (see last row of Table 2 and Figure 16 for validation results that show the high accuracy and low error of the automatic segmentation method proposed in this research as compared to the two manual expert tracing-based segmentation results). (a) Original brain MRI image. (b) Gray matter region in original image. (c) White matter region in original image. (d) Gray matter: manual segmentation 1. (e) White matter: manual segmentation 1. (f) Gray matter: manual segmentation 2. (g) White matter: manual segmentation 2. (h) Gray matter region: automatic segmentation. (i) White matter region: automatic segmentation.

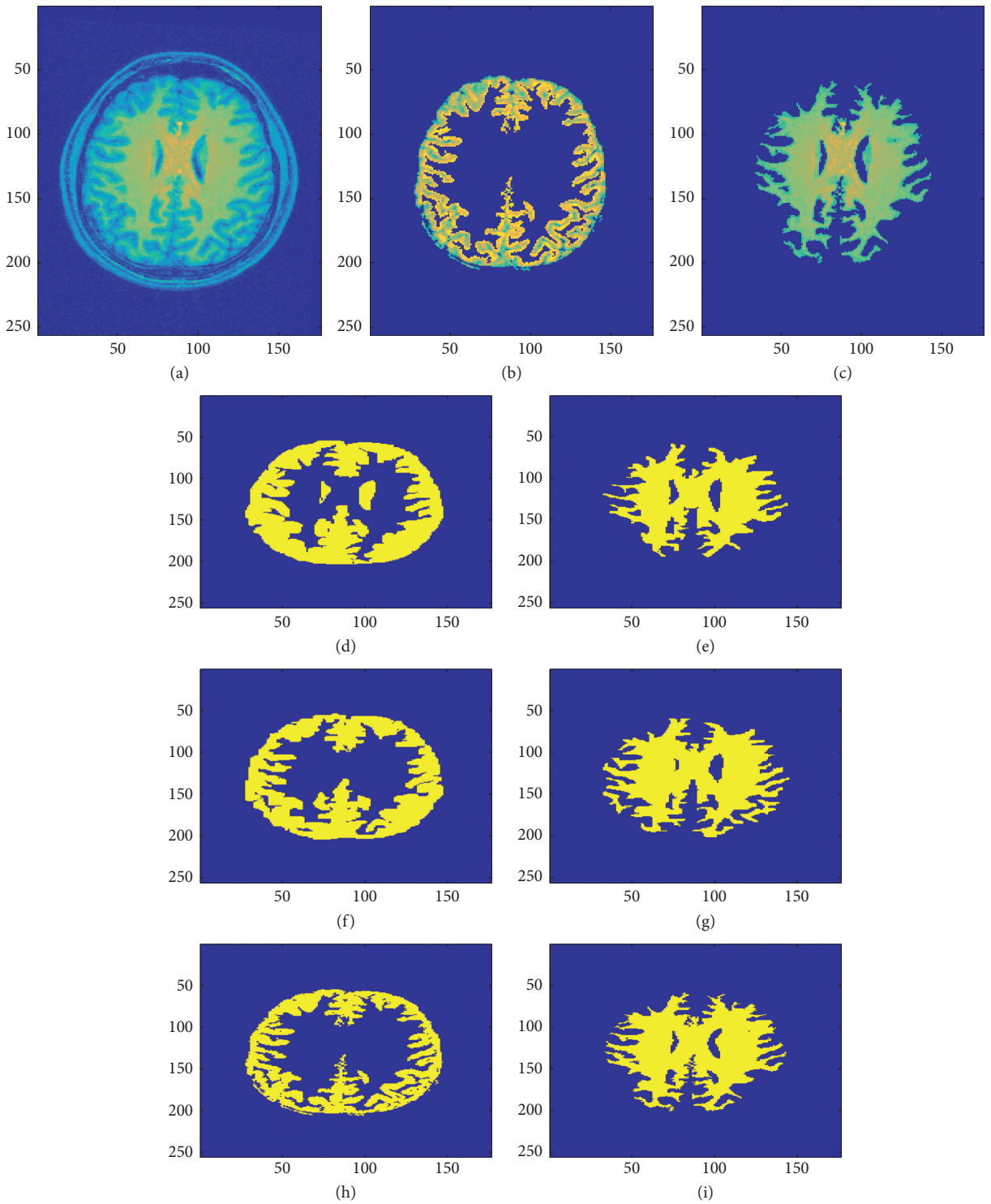


Figure 14: Visual comparison of two manual expert tracing-based and automatic segmentation (using the fully automatic segmentation method presented in this paper) results of sample patient 2 brain MRI image (note the difference between the two manual segmentations of the gray matter: one including and the other excluding portion(s) of the cerebrospinal fluid region; this shows the robustness of the proposed automatic segmentation algorithm to still have high validity even when considering error, taking human manual error into account; see last row of Table 2 and Figure 16 for validation results that show the high accuracy and low error of the automatic segmentation method proposed in this research as compared to the two manual expert tracing-based segmentation results). (a) Original brain MRI image. (b) Gray matter region in original image. (c) White matter region in original image. (d) Gray matter: manual segmentation 1. (e) White matter: manual segmentation 1. (f) Gray matter: manual segmentation 2. (g) White matter: manual segmentation 2. (h) Gray matter region: automatic segmentation. (i) White matter region: automatic segmentation.

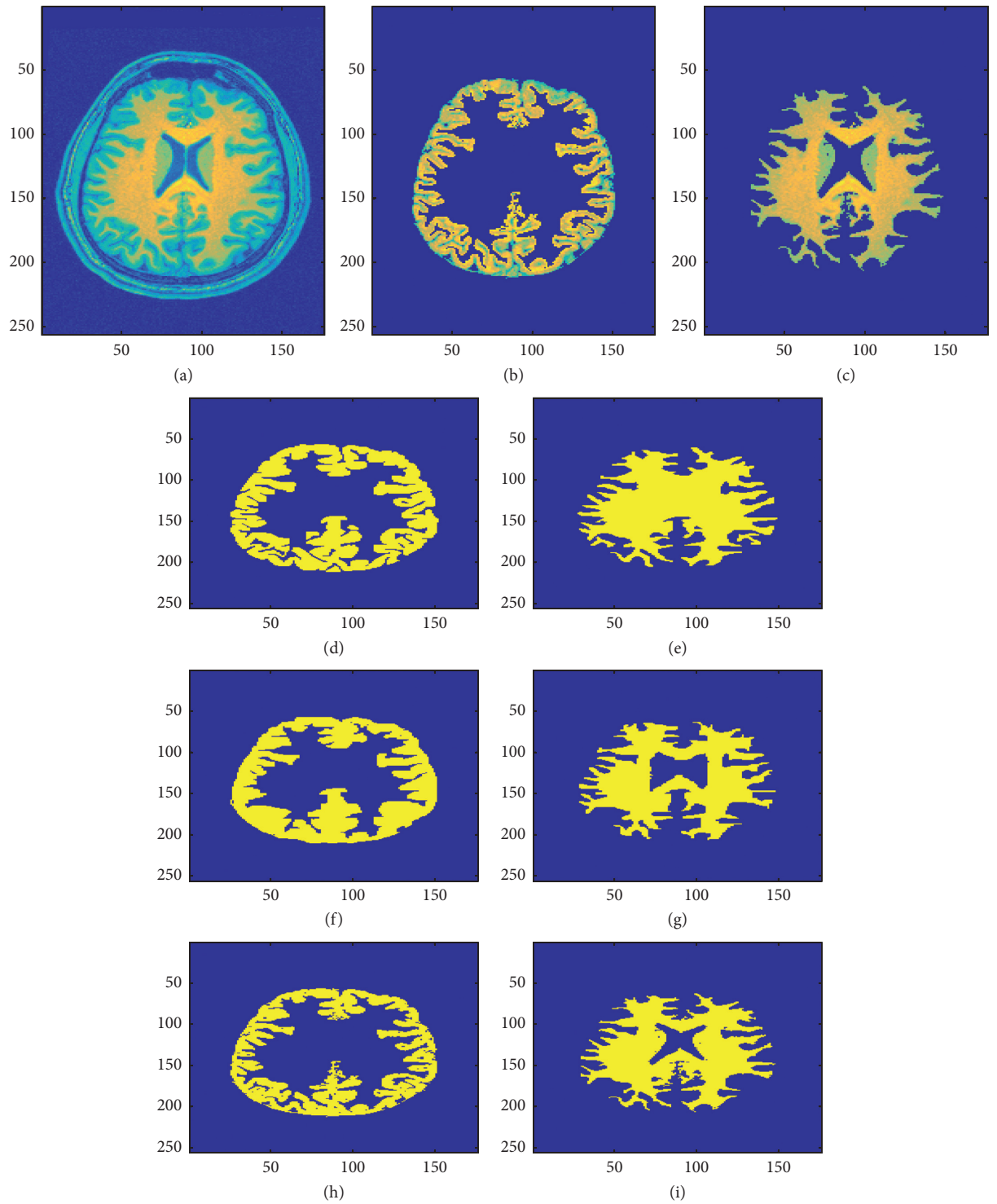


Figure 15: Visual comparison of two manual expert tracing-based and automatic segmentation (using the fully automatic segmentation method presented in this paper) results of sample patient 3 brain MRI image (see last row of Table 2 and Figure 16 for validation results that show the high accuracy and low error of the automatic segmentation method proposed in this research as compared to the two manual expert tracing-based segmentation results). (a) Original brain MRI image. (b) Gray matter region in original image. (c) White matter region in original image. (d) Gray matter: manual segmentation 1. (e) White matter: manual segmentation 1. (f) Gray matter: manual segmentation 2. (g) White matter: manual segmentation 2. (h) Gray matter region: automatic segmentation. (i) White matter region: automatic segmentation.

segmentation for each of the patient brain MRI images. For each patient brain MRI image, manual segmentation was performed three times by experts. The Dice coefficients are calculated between all the manual and automatic segmentation for each patient brain MRI image. Figure 16 shows the box plots of the Dice coefficients calculated as the similarity

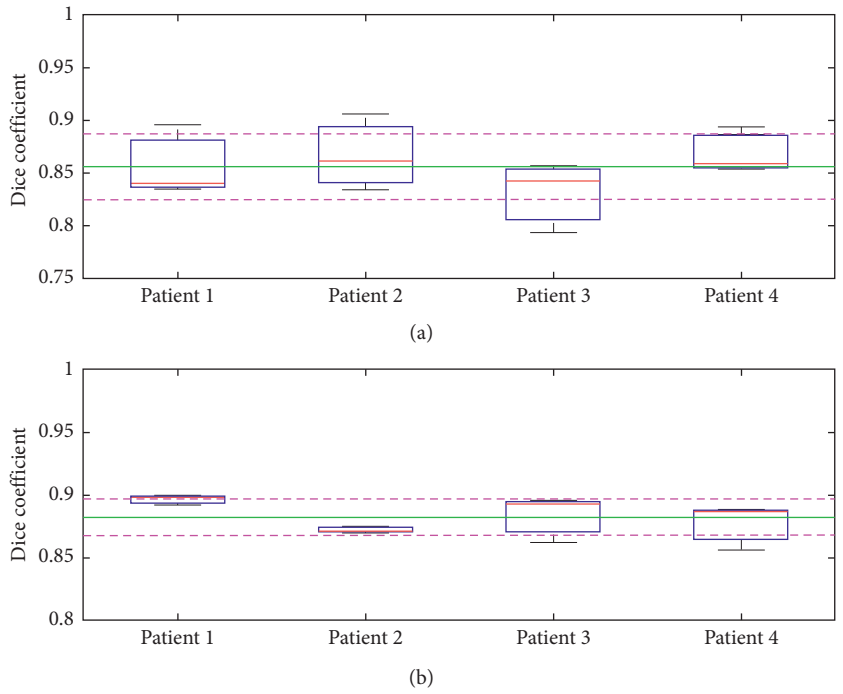


Figure 16: Box plots for Dice coefficients to compare manual and automatic segmentation of brain MRI images of 5 patients. Overall mean of the Dice coefficient is represented as a green line and standard deviation is represented as the dashed purple lines. (a) Comparison between automatic and manual segmentations of gray matter. (b) Comparison between automatic and manual segmentations of white matter.

measure to compare manual and automatic segmentation of the brain MRI images for the five sample patients.

The box plots in Figure 16 show the minimum, first quartile, median, third quartile, and maximum values of the distribution of Dice coefficients computed between each pair of manual and automatic segmentation for each patient. Each patient's brain MRI image was automatically segmented by the algorithm proposed in this research work and was manually traced three separate times by experts (three manual segmentations) [96–102]. So, several Dice coefficients were calculated between each of the manual segmentations by expert tracing and the automatic segmentation for each patient.

One of the challenging tasks in medical imaging sciences is to extract the gray and white matter from MRI brain images. In our research, we have used adaptive fuzzy c-means algorithm in which pixels are classified based on intensity and membership-based fuzzy c-means clustering with preprocessing using elliptical Hough transform and postprocessing using connected region analysis. Table 2 shows the average Dice coefficient values for the similarity measures between the manual expert tracings and the automatic segmentations of gray matter, white matter, and total cortical matter results of the proposed algorithm presented in this paper, compared with previously used standard state-of-the-art methods for brain MRI segmentation. The proposed algorithm presented in this work has the highest Dice coefficient similarity measures for gray, white, and total cortical matter segmentation when compared with other previously published standard state-of-the-art brain MRI segmentation methods.

8. Future Work

Future research in this work will further investigate gray white matter ratio as a marker of cognitive impairment or dementia. The advantage of this proposed future idea is that it will not require a sequence of MRI scans over several dates but will rather be able to predict severity of cognitive impairment or dementia from a single MRI scan.

The motivation of this work is that this idea is implemented in this proposed user-friendly software platform with an easy-to-use graphical user interface for neurologists to automatically quantify severity of dementia or cognitive impairment from a single structural MRI scan of a patient brain. In future, the proposed algorithm will be applied on larger datasets of brain MR images for gray and white matter extraction, which can be validated by experts. Further, neurological disease classification can be done based on volume ratio of gray and white matter for different MRI images.

The idea proposed herein is that the machine learning or model-based prediction algorithm that is developed can calculate the cognitive impairment level as the distance from the regression line, which here is the curve fitted to the scatter data points in the gray white matter ratio to age plot from previously published research.

Figure 17 shows a depiction of the neurological disease prediction and decision-making framework developed in this work for prediction of cognitive impairment level. The patient image data and metadata containing the age and medical history are also employed. A model-based prediction or machine learning algorithm can be used to output

Table 2: Performance and accuracy comparison of the authors' proposed automatic brain MRI segmentation algorithm [83] with previous algorithms [88] using Dice coefficients as similarity measure estimated between manual expert tracings and automatic algorithm-based segmentation.

Methods	Procedure	Average of Dice coefficients (gray matter)	Average of Dice coefficients (white matter)	Average of Dice coefficients (total cortical matter)
K-means	Statistical distance-based *k*-means clustering with preprocessing using median filters	0.70	0.71	0.71
Intensity-based fuzzy *c*-means	Pixel intensity and membership-based fuzzy *c*-means clustering with preprocessing using median filters	0.71	0.79	0.75
Adaptive fuzzy *c*-means with preprocessing and postprocessing (proposed method in this work)	Pixel intensity and membership-based fuzzy *c*-means clustering with preprocessing using elliptical Hough transform and postprocessing using connected region analysis	0.86	0.88	0.87

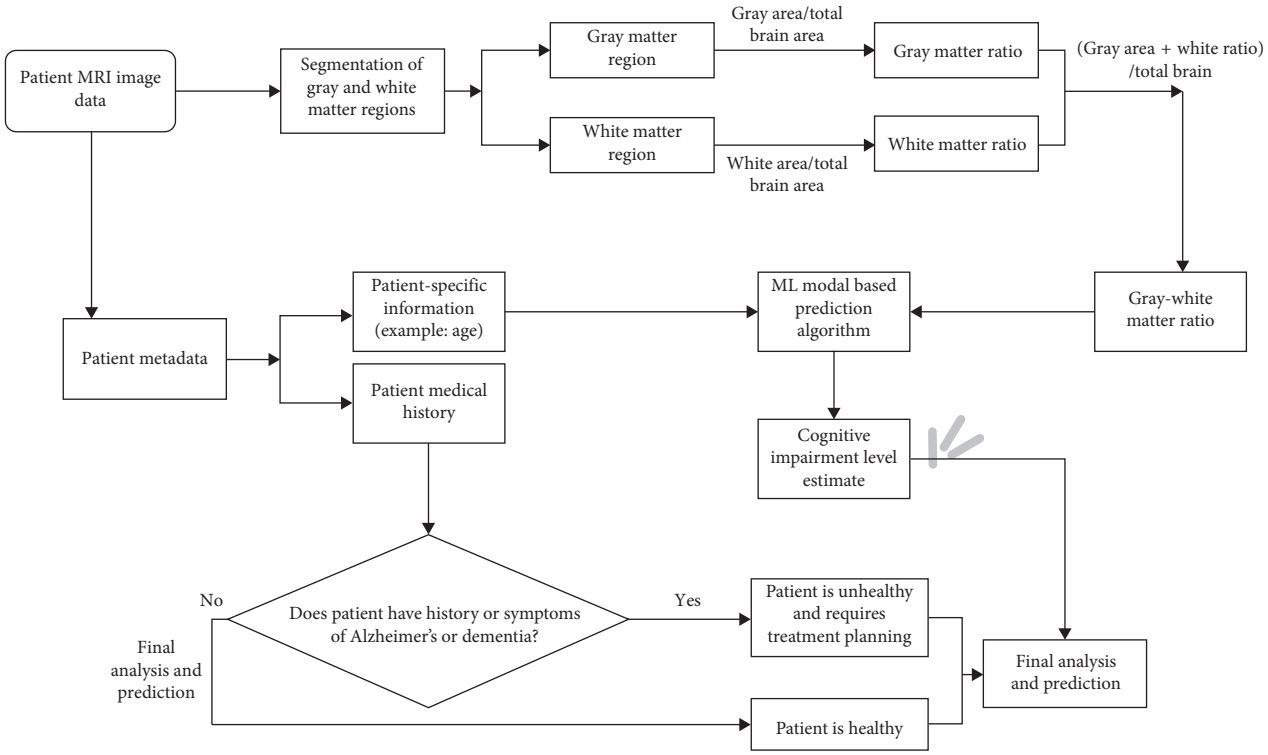


Figure 17: Neurological disease prediction and decision-making framework for determining cognitive impairment level based on gray and white matter ratio and patient data.

the prediction based on the input parameters, namely, age and gray-white matter ratio. This algorithm can be based on previous research published on the correlation between age and gray and white matter ratios.

As proposed in this work, the average thickness and volume measurements of the neocortical and nonneocortical regions between the boundaries of the white and gray matter regions, the aggregate of the parts of the regions in both the left and right hemispheres, can be used as the measures with which the cognitive impairment or dementia is quantitatively assessed for a patient, based on their brain MRI scan.

As shown in Figure 17, based on the work proposed in this research paper, a neurological disease detection and decision-making framework can be developed with segmentations of the gray and white matter regions to determine the level of atrophy or degeneration in the cortical matter and assess the severity of dementia or cognitive impairment in a neurologically diseased patient.

9. Conclusion

The research presented in this work facilitates efficient and effective automatic segmentation of gray and white matter regions from brain MRI images, which has several clinical neurological applications. A fully automatic segmentation methodology using elliptical Hough transform along with pixel intensity and membership-based adapted fuzzy *c*-means clustering followed by connected component labeling

and region analysis has been implemented in this research to perform segmentation of gray and white matter regions in brain MRI images. The algorithm was tested and verified for several sample brain MRI images including patient brain MRI images having tumor sections. The algorithm implemented in this research acquired higher accuracy in the results when compared to other previous state-of-the-art algorithms that have been published so far. Manual segmentations were performed by neurological experts for several patient brain MRI images. These manual segmentations were used to compare and validate with the results obtained from the automatic segmentations in this research work. Validations were performed by calculating several Dice coefficient values between the automatic segmentation results and the manual segmentation results. The Dice coefficient values are similarity measures that are represented statistically using box plots in this research. The average of the Dice coefficient values obtained was higher for the algorithm proposed and implemented in this work when compared to other methodologies that have been published so far in the medical field to automatically segment gray and white matter regions in brain MRI images. The automatized computational segmentation tool developed in this research can be employed in hospitals and neurology divisions as a computational software platform for assisting neurologist in detection of disease from brain MRI images after MRI segmentation. This tool obviates manual tracing and saves the precious time of neurologists or radiologists. This research presented herein is foundational to a neurological disease prediction and disease detection framework, which in the future, with further research work, can be developed and implemented with a machine learning model-based prediction algorithm to detect and calculate the severity level of the disease, based on the gray and white matter region segmentations and estimated gray and white matter ratios to the total cortical matter, as outlined in this research.

Data Availability

The data can be provided to the readers from the corresponding author upon request and can also be sent to them along with the code and software to test out and see the results for themselves.

Ethical Approval

The patient's brain MRI image and neurological data used in this research work were obtained from the Image and Data Archive (IDA) powered by Laboratory of Neuro Imaging (LONI) provided by the University of Southern California (USC) and also from the Department of Neurosurgery at the All India Institute of Medical Sciences (AIIMS), New Delhi, India. The data were anonymized as well as followed all the ethical guidelines of the ethical and institutional review boards of all the participating research institutions. The images, image acquisition, and image processing followed all the ethical guidelines of the institutional review boards of the University of Southern California (USC), National Institutes of Health (NIH), National Institute of Biomedical Imaging and Bioengineering (NIBIB), and All India Institute of Medical Sciences (AIIMS).

Disclosure

An earlier initial version of this research work was presented as a poster at the Texas A&M University System 14th Annual Pathways Student Research Symposium on November 2-3, 2017, at Tarleton State University, Stephenville, Texas, USA.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank and acknowledge the neurologists at the All India Institute of Medical Sciences (AIIMS) and the Image and Data Archive (IDA) powered by Laboratory of Neuro Imaging (LONI) provided by the University of Southern California (USC) for providing brain MRI patient data and for sharing the neurological data in this project.

References

[1] B. C. Dickerson, D. H. Salat, J. F. Bates et al., "Medial temporal lobe function and structure in mild cognitive impairment," *Annals of Neurology*, vol. 56, no. 1, pp. 27–35, 2004.

[2] P. J. Visser, P. Scheltens, F. R. J. Verhey et al., "Medial temporal lobe atrophy and memory dysfunction as predictors for dementia in subjects with mild cognitive impairment," *Journal of Neurology*, vol. 246, no. 6, pp. 477–485, 1999.

[3] G. W. Small, A. La Rue, S. Komo, A. Kaplan, and M. A. Mandelkern, "Predictors of cognitive change in middle-aged and older adults with memory loss," *American Journal of Psychiatry*, vol. 152, no. 12, pp. 1757–64, 1995.

[4] M. E. Shenton, C. C. Dickey, M. Frumin, and R. W. McCarley, "A review of MRI findings in schizophrenia," *Schizophrenia Research*, vol. 49, no. 1, pp. 1–52, 2001.

[5] B. Fischl, D. H. Salat, E. Busa et al., "Whole brain segmentation," *Neuron*, vol. 33, no. 3, pp. 341–355, 2002.

[6] I. Despotović, B. Goossens, and W. Philips, "MRI segmentation of the human brain: challenges, methods, and applications," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 450341, 23 pages, 2015.

[7] M. W. Weiner, D. P. Veitch, P. S. Aisen et al., "The Alzheimer's disease neuroimaging initiative: a review of papers published since its inception," *Alzheimer's & Dementia*, vol. 9, no. 5, pp. e111–e194, 2013.

[8] J. C. Tamraz, C. Outin, M. F. Secca, and B. Soussi, *MRI Principles of the Head, Skull Base and Spine: A Clinical Approach*, Springer Science & Business Media, Berlin, Germany, 2013.

[9] B. P. Rourke, "Arithmetic disabilities, specific and otherwise," *Journal of Learning Disabilities*, vol. 26, no. 4, pp. 214–226, 2016.

[10] A. Sehgal and R. Agrawal, "Entropy based integrated diagnosis for enhanced accuracy and removal of variability in clinical inferences," in *Proceedings of 2014 International*

Conference on Signal Processing and Integrated Networks (SPIN), pp. 571–575, IEEE, Noida, Uttar Pradesh, India, February 2014.

[11] A. L. Guillozet, S. Weintraub, D. C. Mash, and M. M. Mesulam, "Neurofibrillary tangles, amyloid, and memory in aging and mild cognitive impairment," *Archives of Neurology*, vol. 60, no. 5, pp. 729–736, 2003.

[12] S. Sneha and R. Agrawal, "Towards enhanced accuracy in medical diagnostics—a technique utilizing statistical and clinical data analysis in the context of ultrasound images," in *Proceedings of 2013 46th Hawaii International Conference on System Sciences (HICSS)*, pp. 2408–2415, January 2013.

[13] S. B. Chapman, R. N. Rosenberg, M. F. Weiner, and A. Shobe, "Autosomal dominant progressive syndrome of motor-speech loss without dementia," *Neurology*, vol. 49, no. 5, pp. 1298–1306, 1997.

[14] J. R. Petrella, R. E. Coleman, and P. M. Doraiswamy, "Neuroimaging and early diagnosis of Alzheimer disease: a look to the future," *Radiology*, vol. 226, no. 2, pp. 315–336, 2003.

[15] P. A. Steen, L. A. Newberg, J. H. Milde, and J. D. Michenfelder, "Nimodipine improves cerebral blood flow and neurologic recovery after complete cerebral ischemia in the dog," *Journal of Cerebral Blood Flow & Metabolism*, vol. 3, no. 1, pp. 38–43, 2016.

[16] P. A. Steen, S. E. Gisvold, J. H. Milde et al., "Nimodipine improves outcome when given after complete cerebral ischemia in primates," *Anesthesiology*, vol. 62, no. 4, pp. 406–414, 1985.

[17] W. L. Lanier, K. J. Stangland, B. W. Scheithauer, J. H. Milde, and J. D. Michenfelder, "The effects of dextrose infusion and head position on neurologic outcome after complete cerebral ischemia in primates," *Anesthesiology*, vol. 66, no. 1, pp. 39–48, 1987.

[18] T. Persson, B. O. Popescu, and A. Cedazo-Minguez, "Oxidative stress in Alzheimer's disease: why did antioxidant therapy fail?," *Oxidative Medicine and Cellular Longevity*, vol. 2014, Article ID 427318, 11 pages, 2014.

[19] C. Pantofaru and M. Hebert, *A Comparison of Image Segmentation Algorithms*, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA, 2005.

[20] Y. H. Wang, *Tutorial: Image Segmentation*, National Taiwan University, Taipei, Taiwan, 2010.

[21] J. A. F. Costa and J. G. de Souza, "Image segmentation through clustering based on natural computing techniques," in *Image Segmentation*, IntechOpen, London, UK, 2011.

[22] S. Arumugadevi and V. Seenivasagam, "Comparison of clustering methods for segmenting color images," *Indian Journal of Science and Technology*, vol. 8, no. 7, pp. 670–677, 2015.

[23] M. H. Zafar and M. Ilyas, "A clustering based study of classification algorithms," *International Journal of Database Theory and Application*, vol. 8, no. 1, pp. 11–22, 2015.

[24] M. K. Siddiqui and S. Naahid, "Analysis of KDD CUP 99 dataset using clustering based data mining," *International Journal of Database Theory and Application*, vol. 6, no. 5, pp. 23–34, 2013.

[25] M. E. Celebi, H. A. Kingravi, and P. A. Vela, "A comparative study of efficient initialization methods for the k-means clustering algorithm," *Expert Systems with Applications*, vol. 40, no. 1, pp. 200–210, 2013.

[26] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using *K*-means clustering algorithm and subtractive clustering algorithm," *Procedia Computer Science*, vol. 54, pp. 764–771, 2015.

[27] H. Li, H. He, and Y. Wen, "Dynamic particle swarm optimization and *K*-means clustering algorithm for image segmentation," *Optik*, vol. 126, no. 24, pp. 4817–4822, 2015.

[28] R. Jensi and G. W. Jiji, "Hybrid data clustering approach using k-means and flower pollination algorithm," 2015, http://arxiv.org/abs/1505.03236.

[29] S. B. Belhaouari, S. Ahmed, and S. Mansour, "Optimized *K*-means algorithm," *Mathematical Problems in Engineering*, vol. 2014, Article ID 506480, 14 pages, 2014.

[30] S. Khanmohammadi, N. Adibeig, and S. Shanehbandy, "An improved overlapping *k*-means clustering method for medical applications," *Expert Systems with Applications*, vol. 67, pp. 12–18, 2017.

[31] A. Halder, S. Pramanik, and A. Kar, "Dynamic image segmentation using fuzzy *C*-means based genetic algorithm," *International Journal of Computer Applications*, vol. 28, no. 6, pp. 15–20, 2011.

[32] A. M. Ali, G. C. Karmakar, and L. S. Dooley, "Review on fuzzy clustering algorithms," *Journal of Advanced Computations*, vol. 2, no. 3, pp. 169–181, 2008.

[33] N. Dhanachandra and Y. J. Chanu, "A survey on image segmentation methods using clustering techniques," *European Journal of Engineering Research and Science*, vol. 2, no. 1, pp. 15–20, 2017.

[34] J. M. Mendel, R. I. John, and F. Liu, "Interval type-2 fuzzy logic systems made simple," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 6, pp. 808–821, 2006.

[35] L. Ma, Y. Li, S. Fan, and R. Fan, "A hybrid method for image segmentation based on artificial fish swarm algorithm and fuzzy *c*-means clustering," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 120495, 10 pages, 2015.

[36] O. M. Rotman, B. Kovarovic, C. Sadasivan, L. Gruberg, B. B. Lieber, and D. Bluestein, "Realistic vascular replicator for TAVR procedures," *Cardiovascular Engineering and Technology*, vol. 9, no. 3, pp. 339–350, 2018.

[37] P. Datta, A. Gupta, and R. Agrawal, "Statistical modeling of B-mode clinical kidney images," in *Proceedings of 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*, pp. 222–229, IEEE, Greater Noida, Uttar Pradesh, India, November 2014.

[38] P. A. Steen, L. A. Newberg, J. H. Milde, and J. D. Michenfelder, "Cerebral blood flow and neurologic outcome when nimodipine is given after complete cerebral ischemia in the dog," *Journal of Cerebral Blood Flow & Metabolism*, vol. 4, no. 1, pp. 82–87, 2016.

[39] O. Steward and S. A. Scoville, "Cells of origin of entorhinal cortical afferents to the hippocampus and fascia dentata of the rat," *Journal of Comparative Neurology*, vol. 169, no. 3, pp. 347–370, 1976.

[40] S. J. Lupien, M. de Leon, S. de Santi et al., "Cortisol levels during human aging predict hippocampal atrophy and memory deficits," *Nature Neuroscience*, vol. 1, no. 1, pp. 69–73, 1998.

[41] F. Nicoletti, M. J. Iadarola, J. T. Wroblewski, and E. Costa, "Excitatory amino acid recognition sites coupled with inositol phospholipid metabolism: developmental changes and interaction with alpha 1-adrenoceptors," in *Proceedings of the National Academy of Sciences*, vol. 83, no. 6, pp. 1931–1935, 1986.

[42] W. F. Styler, S. Bethard, S. Finan et al., "Temporal annotation in the clinical domain," *Transactions of the Association for Computational Linguistics*, vol. 2, pp. 143–154, 2014.

[43] N. Geschwind and W. Levitsky, "Human brain: left-right asymmetries in temporal speech region," *Science*, vol. 161, no. 3837, pp. 186-187, 1968.

[44] M. A. Warner, T. S. Youn, T. Davis et al., "Regionally selective atrophy after traumatic axonal injury," *Archives of Neurology*, vol. 67, no. 11, pp. 1336–1344, 2010.

[45] C. R. Jack Jr., D. S. Knopman, W. J. Jagust et al., "Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers," *Lancet Neurology*, vol. 12, no. 2, pp. 207–216, 2013.

[46] G. B. Frisoni, N. C. Fox, C. R. Jack Jr., P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nature Reviews Neurology*, vol. 6, no. 2, pp. 67–77, 2010.

[47] N. K. Roberts, "The journal: the next 5 years," *Journal of Insurance Medicine*, vol. 32, pp. 1–4, 2000.

[48] M.-H. Choi, H.-S. Kim, S.-Y. Gim et al., "Differences in cognitive ability and hippocampal volume between Alzheimer's disease, amnestic mild cognitive impairment, and healthy control groups, and their correlation," *Neuroscience Letters*, vol. 620, pp. 115–120, 2016.

[49] L. C. Silbert, H. H. Dodge, L. G. Perkins et al., "Trajectory of white matter hyperintensity burden preceding mild cognitive impairment," *Neurology*, vol. 79, no. 8, pp. 741–747, 2012.

[50] H. Shinotoh, H. Shimada, S. Hirano et al., "Longitudinal [11C]PIB PET study in healthy elderly persons, patients with mild cognitive impairment, and Alzheimer's disease," *Alzheimer's & Dementia*, vol. 7, no. 4, p. S224, 2011.

[51] M. Dumont and M. F. Beal, "Neuroprotective strategies involving ROS in Alzheimer disease," *Free radical Biology and Medicine*, vol. 51, no. 5, pp. 1014–1026, 2011.

[52] F. J. Rugg-Gunn and M. R. Symms, "Novel MR contrasts to reveal more about the brain," *Neuroimaging Clinics of North America*, vol. 14, no. 3, pp. 449–470, 2004.

[53] M. A. Greenough, J. Camakaris, and A. I. Bush, "Metal dyshomeostasis and oxidative stress in Alzheimer's disease," *Neurochemistry international*, vol. 62, no. 5, pp. 540–555, 2013.

[54] D. N. Loy, J. H. Kim, M. Xie, R. E. Schmidt, K. Trinkaus, and S.-K. Song, "Diffusion tensor imaging predicts hyperacute spinal cord injury severity," *Journal of Neurotrauma*, vol. 24, no. 6, pp. 979–990, 2007.

[55] E. M. Haacke and Z. Kou, *Development of Magnetic Resonance Imaging Biomarkers for Traumatic Brain Injury*, Wayne State University, Detroit, MI, USA, 2014.

[56] P.-H. Yeh, T. R. Oakes, and G. Riedy, "Diffusion tensor imaging and its application to traumatic brain injury: basic principles and recent advances," *Open Journal of Medical Imaging*, vol. 2, no. 4, pp. 137–161, 2012.

[57] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet, "MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders," *Radiology*, vol. 161, no. 2, pp. 401–407, 1986.

[58] P. T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*, Oxford University Press, Oxford, UK, 1993.

[59] B. R. Rosen, J. W. Belliveau, J. M. Vevea, and T. J. Brady, "Perfusion imaging with NMR contrast agents," *Magnetic Resonance in Medicine*, vol. 14, no. 2, pp. 249–265, 1990.

[60] R. R. Edelman, B. Siewert, D. G. Darby et al., "Qualitative mapping of cerebral blood flow and functional localization with echo-planar MR imaging and signal targeting with alternating radio frequency," *Radiology*, vol. 192, no. 2, pp. 513–520, 1994.

[61] N. Gordillo, E. Montseny, and P. Sobrevilla, "State of the art survey on MRI brain tumor segmentation," *Magnetic Resonance Imaging*, vol. 31, no. 8, pp. 1426–1438, 2013.

[62] S. Suhag and L. M. Saini, "Automatic detection of brain tumor by image processing in matlab," in *Proceedings of 10th SARC-IRF International Conference*, pp. 45–48, New Delhi, India, May 2015.

[63] A. Naveen and T. Velmurugan, "Identification of calcification in MRI brain images by k-means algorithm," *Indian Journal of Science and Technology*, vol. 8, no. 29, 2015.

[64] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, and Y. Pan, "A survey of MRI-based brain tumor segmentation methods," *Tsinghua Science and Technology*, vol. 19, no. 6, pp. 578–595, 2014.

[65] C. Tsai, B. S. Manjunath, and R. Jagadeesan, "Automated segmentation of brain MR images," *Pattern Recognition*, vol. 28, no. 12, pp. 1825–1837, 1995.

[66] K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, and T.-J. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," *Computerized Medical Imaging and Graphics*, vol. 30, no. 1, pp. 9–15, 2006.

[67] M. Padurariu, A. Ciobica, R. Lefter, I. Lacramioara Serban, C. Stefanescu, and R. Chirita, "The oxidative stress hypothesis in Alzheimer's disease," *Psychiatria Danubina*, vol. 25, no. 4, p. 409, 2013.

[68] D. Antolovic, *Review of the Hough transform method, with an implementation of the fast Hough variant for line detection*, Department of Computer Science, Indiana University, 2008.

[69] N. Kumar and M. Nachamai, "Noise removal and filtering techniques used in medical images," *Indian Journal of Computer Science and Engineering*, vol. 3, no. 1, pp. 146–153, 2012.

[70] P. Melin, C. I. Gonzalez, J. R. Castro, O. Mendoza, and O. Castillo, "Edge-detection method for image processing based on generalized type-2 fuzzy logic," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 6, pp. 1515–1525, 2014.

[71] C. Jayalakshmi and K. Sathiyasekar, "Analysis of brain tumor using intelligent techniques," in *Proceedings of 2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)*, pp. 48–52, May 2016.

[72] K. K. L. Wong, J. Tu, R. M. Kelso et al., "Cardiac flow component analysis," *Medical Engineering & Physics*, vol. 32, no. 2, pp. 174–188, 2010.

[73] E. A. Zanaty, "An approach based on fusion concepts for improving brain Magnetic Resonance Images (MRIs) segmentation," *Journal of Medical Imaging and Health Informatics*, vol. 3, no. 1, pp. 30–37, 2013.

[74] E. A. Zanaty and S. Ghoniemy, "Medical image segmentation techniques: an overview," *International Journal of Informatics and Medical Data Processing*, vol. 1, no. 1, pp. 16–37, 2016.

[75] E. A. Zanaty and A. Afifi, "A watershed approach for improving medical image segmentation," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 16, no. 12, pp. 1262–1272, 2013.

[76] E. A. Zanaty, "An adaptive fuzzy C-means algorithm for improving MRI segmentation," *Open Journal of Medical Imaging*, vol. 3, no. 4, p. 125, 2013.

[77] M. B. Dillencourt, H. Samet, and M. Tamminen, "A general approach to connected-component labeling for arbitrary

image representations," *Journal of the ACM*, vol. 39, no. 2, pp. 253–280, 1992.

[78] K. Wu, E. Otoo, and A. Shoshani, "Optimizing connected component labeling algorithms," in *Proceedings of Medical Imaging 2005: Image Processing*, vol. 5747, pp. 1965–1977, International Society for Optics and Photonics, San Diego, CA, USA, February 2005.

[79] K. Suzuki, I. Horiba, and N. Sugie, "Linear-time connected-component labeling based on sequential local operations," *Computer Vision and Image Understanding*, vol. 89, no. 1, pp. 1–23, 2003.

[80] M. D. Sinclair, J. Lee, A. N. Cookson, S. Rivolo, E. R. Hyde, and N. P. Smith, "Measurement and modeling of coronary blood flow," *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, vol. 7, no. 6, pp. 335–356, 2015.

[81] A. Muda, N. Saad, S. Bakar, S. Muda, and A. Abdullah, "Brain lesion segmentation using fuzzy C-means on diffusion-weighted imaging," *ARPN Journal of Engineering and Applied Sciences*, vol. 10, no. 3, pp. 1138–1144, 2015.

[82] J. Selvakumar, A. Lakshmi, and T. Arivoli, "Brain tumor segmentation and its area calculation in brain MR images using *K*-mean clustering and fuzzy *C*-mean algorithm," in *Proceedings of 2012 International Conference on Advances in Engineering, Science and Management (ICAESM)*, pp. 186–190, Nagapattinam, Tamil Nadu, India, March 2012.

[83] A. Goyal, M. K. Arya, R. Agrawal, D. Agrawal, G. Hossain, and R. Challoo, "Automated segmentation of gray and white matter regions in brain MRI images for computer aided diagnosis of neurodegenerative diseases," in *Proceedings of 2017 International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT)*, pp. 204–208, Aligarh,India, November 2017.

[84] B. S. Sikarwar, M. Roy, P. Ranjan, and A. Goyal, "Automatic disease screening method using image processing for dried blood microfluidic drop stain pattern recognition," *Journal of Medical Engineering & Technology*, vol. 40, no. 5, pp. 245–254, 2016.

[85] B. S. Sikarwar, M. K. Roy, P. Priya Ranjan, and A. Ayush Goyal, "Imaging-based method for precursors of impending disease from blood traces," in *Advances in Intelligent Systems and Computing*, pp. 411–424, Springer, Singapore, 2016.

[86] B. S. Sikarwar, M. K. Roy, P. Ranjan, and A. Goyal, "Automatic pattern recognition for detection of disease from blood drop stain obtained with microfluidic device," in *Advances in Intelligent Systems and Computing*, vol. 425, pp. 655–667, Springer, Berlin, Germany, 2015.

[87] A. Bhan, D. Bathla, and A. Goyal, "Patient-specific cardiac computational modeling based on left ventricle segmentation from magnetic resonance images," in *International Conference on Data Engineering and Communication Technology*, pp. 179–187, Springer, Singapore, 2017.

[88] V. Deepa, C. C. Benson, and V. L. Lajish, "Gray matter and white matter segmentation from MRI brain images using clustering methods," *International Research Journal of Engineering and Technology (IRJET)*, vol. 2, no. 8, pp. 913–921, 2015.

[89] V. Ray and A. Goyal, "Automatic left ventricle segmentation in cardiac MRI images using a membership clustering and heuristic region-based pixel classification approach," in *Advances in Intelligent Systems and Computing*, pp. 615–623, Springer, Cham, Switzerland, 2015.

[90] M. Chhabra and A. Goyal, "Accurate and robust Iris recognition using modified classical Hough transform," in *Information and Communication Technology for Sustainable Development*, pp. 493–507, Springer, Singapore, 2017.

[91] A. Goyal and V. Ray, "Belongingness clustering and region labeling based pixel classification for automatic left ventricle segmentation in cardiac MRI images," *Translational Biomedicine*, vol. 6, no. 3, 2015.

[92] M. Roy, B. Singh Sikarwar, M. Bhandwal, and P. Ranjan, "Modelling of blood flow in stenosed arteries," *Procedia Computer Science*, vol. 115, pp. 821–830, 2017.

[93] A. Bhan, A. Goyal, N. Chauhan, and C. W. Wang, "Feature line profile based automatic detection of dental caries in bitewing radiography," in *Proceedings of 2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)*, pp. 635–640, Delhi, India, September 2016.

[94] A. Bhan, A. Goyal, M. K. Dutta, K. Riha, and Y. Omran, "Image-based pixel clustering and connected component labeling in left ventricle segmentation of cardiac MR images," in *Proceedings of 2015 7th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 339–342, Brno, Czech Republic, October 2015.

[95] V. Ray and A. Goyal, "Image-based fuzzy c-means clustering and connected component labeling subsecond fast fully automatic complete cardiac cycle left ventricle segmentation in multi frame cardiac MRI images," in *Proceedings of 2016 International Conference on Systems in Medicine and Biology (ICSMB)*, pp. 36–40, Kharagpur, India, January 2016.

[96] A. Goyal, J. van den Wijngaard, P. van Horssen, V. Grau, J. Spaan, and N. Smith, "Intramural spatial variation of optical tissue properties measured with fluorescence microsphere images of porcine cardiac tissue," in *Proceedings of Annual International Conference of the IEEE Proceedings of Engineering in Medicine and Biology Society EMBC 2009*, pp. 1408–1411, Minneapolis, MN, USA, September 2009.

[97] P. Sharma, S. Sharma, and A. Goyal, "An MSE (mean square error) based analysis of deconvolution techniques used for deblurring/restoration of MRI and CT Images," in *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, p. 51, Udaipur, India, March 2016.

[98] A. Goyal, D. Bathla, P. Sharma, M. Sahay, and S. Sood, "MRI image based patient specific computational model reconstruction of the left ventricle cavity and myocardium," in *Proceedings of 2016 International Conference on Computing, Communication and Automation (ICCCA)*, pp. 1065–1068, Greater Noida, India, April 2016.

[99] S. J. Verzi, C. M. Vineyard, E. D. Vugrin, M. Galiardi, C. D. James, and J. B. Aimone, "Optimization-based computation with spiking neurons," in *Proceedings of 2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 2015–2022, Anchorage, AK, USA, May 2017.

[100] M. S. Atkins and B. T. Mackiewich, "Fully automatic segmentation of the brain in MRI," *IEEE Transactions on Medical Imaging*, vol. 17, no. 1, pp. 98–107, 1998.

[101] M. G. Wagner, C. M. Strother, and C. A. Mistretta, "Guidewire path tracking and segmentation in 2D fluoroscopic time series using device paths from previous frames," in *Proceedings of Medical Imaging 2016: Image Processing*, vol. 9784, p. 97842B, International Society for Optics and Photonics, San Diego, CA, USA, February 2016.

[102] C. Amiot, C. Girard, J. Chanussot, J. Pescatore, and M. Desvignes, "Spatio-temporal multiscale Denoising_newline of fluoroscopic sequence," *IEEE Transactions on Medical Imaging*, vol. 35, no. 6, pp. 1565–1574, 2016.

Hindawi
Journal of Healthcare Engineering
Volume 2019, Article ID 2374645, 12 pages
https://doi.org/10.1155/2019/2374645

Research Article

An Iterative Method for Estimating Nonlinear Elastic Constants of Tumor in Soft Tissue from Approximate Displacement Measurements

Maryam Mehdizadeh Dastjerdi,[1] **Ali Fallah**,[1] **and Saeid Rashidi**[2]

[1]*Department of Biomedical Engineering, Amirkabir University of Technology (AUT), Tehran 15875-4413, Iran*
[2]*Faculty of Medical Sciences and Technologies, Science and Research Branch, Islamic Azad University, Tehran 1477893855, Iran*

Correspondence should be addressed to Ali Fallah; afallah@aut.ac.ir

Received 24 March 2018; Revised 24 June 2018; Accepted 12 July 2018; Published 6 January 2019

Academic Editor: Zahid Akhtar

Objectives. Various elastography techniques have been proffered based on linear or nonlinear constitutive models with the aim of detecting and classifying pathologies in soft tissues accurately and noninvasively. Biological soft tissues demonstrate behaviors which conform to nonlinear constitutive models, in particular the hyperelastic ones. In this paper, we represent the results of our steps towards implementing ultrasound elastography to extract hyperelastic constants of a tumor inside soft tissue. *Methods*. Hyperelastic parameters of the unknown tissue have been estimated by applying the iterative method founded on the relation between stress, strain, and the parameters of a hyperelastic model after (a) simulating the medium's response to a sinusoidal load and extracting the tissue displacement fields in some instants and (b) estimating the tissue displacement fields from the recorded/simulated ultrasound radio frequency signals and images using the cross correlation-based technique. *Results*. Our results indicate that hyperelastic parameters of an unidentified tissue could be precisely estimated even in the conditions where there is no prior knowledge of the tissue, or the displacement fields have been approximately calculated using the data recorded by a clinical ultrasound system. *Conclusions*. The accurate estimation of nonlinear elastic constants yields to the correct cognizance of pathologies in soft tissues.

1. Introduction

According to the World Health Organization report, cancer is one of the principal morbidity and mortality agents throughout the world, with approximately 8.8 million deaths (nearly 1 in 6 deaths) in 2015 and 70% increase in the number of new cases over the next two decades. The cancer statistics imply the requisite to extend medical scrutiny to improve cancer prevention, early correct diagnosis, meticulous screening, and effective treatment and reduce the invasiveness and costs of applied techniques.

Since the first introduction of ultrasound (US) imaging in clinical practice in the 1970s, ultrasonography and other US modalities, for example, Doppler imaging and state-of-the-art elastography imaging methods, which provide the information related to the tissue acoustic impedance, vascular flow, and tissue mechanical characteristics or variables such as its stiffness or strain, respectively, have been extensively utilized for medical diagnoses [1]. The US imaging is recognized a noninvasive, safe, easy-to-use, low-cost, and widely accessible imaging modality for visualizing in vivo tissues. Elastography approach has currently been regarded a promising alternative to invasive medical procedures, for example, the biopsy, to characterize tissue abnormalities.

The wide variety of strategies that are being employed to quantify and image mechanical properties of biological tissues are recognized as elastography or elasticity imaging techniques with reference to their similar premise [2]:

(1) The in vivo tissue is being deformed by a specified external or internal load or motion.

(2) The response of tissue is being recorded by the use of a standard clinical imaging system, such as the US or magnetic resonance imaging (MRI) system.

(3) The mechanical characteristics of tissue are estimated through the assessment of tissue displacement fields.

The alterations in the microstructure of tissue as a consequence of pathophysiological phenomena would change the mechanical properties of tissue; for instance, the increase in the stromal density of cancerous tissue would cause the increment in its Young's modulus [3, 4]. The outcomes of numerous experimental research studies carried out by Krouskop et al. [5], Samani et al. [6, 7], Lyshchik et al. [8], Soza et al. [9], Hoyt et al. [10], Schiavone et al. [11], O'Hagan et al. [12], and Moran et al. [13], to mention but a few, have confirmed the relation between tissue structures and macroscopic mechanical features which are being evaluated, quantified, and/or imaged by employing the palpation, elastography, digital rectal examination, and such like methods.

The precise determination of mechanical characteristics of understudy tissue by the use of an elastography technique would undoubtedly necessitate realistically appraising or modeling the tissue manners, specifically its nonlinear response to the stimulation. Hyperelasticity theory is one of the constitutive theories that have been manipulated to model the nonlinear constitutive demeanor demonstrated by biological soft tissues. A variety of hyperelastic models, for instance the well-known Neo-Hookean, Mooney–Rivlin, Yeoh, and Polynomial models, have been recommended for this purpose [14–17]. In comparison with the studies involving the linear elasticity imaging, the number of research studies conducted to image the nonlinear features of tissues is limited [2].

With the aim of diagnosing a tumor inside the understudy tissue correctly, we have utilized an iterative method, as explicated in the next section, to accurately estimate the Mooney–Rivlin hyperelastic parameters of the tumor inside the tissue. The displacement fields inside the tumor have been analyzed to extract its hyperelastic parameters. An iterative technique has been employed since it has been assumed that no initial knowledge of the tumor was accessible except the displacement fields inside the tumor. The displacement fields inside the tumor have been extracted from the simulated/recorded radio frequency (RF) signals using the cross correlation-based method. The response of an abnormal tissue to a sinusoidal load (with low frequency to negate the inertia) has been simulated by applying the finite element (FEM) software package ABAQUS. The RF signals have been simulated by the use of the Field II US Simulation Program. In brief, in this paper, we scrutinize the diagnosis of tumor through its hyperelastic parameters in the conditions where there is no primary perception of the tumor and the displacement fields inside the tumor are estimated imprecisely.

2. Materials and Methods

2.1. Hyperelasticity Theory. Constitutive theories take the improvement of mathematical models, also known as constitutive equations, into consideration in order to provide the possibility to minutely describe the behavioral characteristics of materials. Constitutive theories in continuum mechanics deal with formulating material models that are [18, 19]

(a) on the basis of some mechanical universal principles

(b) in accordance with experimental observations

Thoughtful consideration of soft tissue's model would result in the realistic prediction of its behavior such that it could be verified by experimental observations. Nonlinear constitutive manners that have been observed from soft tissues in numerous in vivo and ex vivo experimental research studies could be modeled by the use of hyperelastic models [15, 16, 20].

The hyperelastic constitutive laws deal with modeling materials with nonlinear elastic behaviors in reaction to large strains. The nonlinearities that are the consequences of (a) the material behavior and (b) the significant change in the shape of material are both regarded in the constitutive theory of hyperelastic materials. Hyperelastic materials are generally described by specific forms of strain energy density (stored energy) functions. While characterizing the homogeneous material's absorbed energy due to its deformation, the strain energy function, W, is defined as a function of deformation gradient, $\mathbf{F}$ [21, 22]

$$W = W(\mathbf{F}). \tag{1}$$

It is considered that B_r and B represent, respectively, the reference or undeformed configuration, which refers to the situations where no load is exerted to the material and the deformed configuration, which is relevant to the situations where the material is under load and therefore it may alter with time, t. In addition, it is assumed that $\mathbf{X}$ and $\mathbf{x}$, respectively, correspond to the position vectors of a material point in the reference and deformed configurations, B_r and B. The time-dependent deformation of material, that is, the motion of material point, from B_r to B could be described by the function $\boldsymbol{\chi}$, which, for each t, (a) is an invertible function and (b) satisfies proper regularity conditions as follows [23]:

$$\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t). \tag{2}$$

The deformation gradient tensor, $\mathbf{F}$, is defined as

$$\mathbf{F} = \mathrm{Grad}\mathbf{x}, \tag{3}$$

with Cartesian components

$$F_{i\alpha} = \frac{\partial x_i}{\partial X_\alpha} \quad i, \alpha \in \{1, 2, 3\}, \tag{4}$$

where Grad, x_i, and X_α refer to the gradient operator in the configuration B_r and components of $\mathbf{x}$ and $\mathbf{X}$, respectively, while the general convention,

$$J \equiv \det \mathbf{F} > 0, \tag{5}$$

is satisfied. Due to the local invertibility of deformation, $\mathbf{F}$ should be nonsingular. The unique polar decomposition of $\mathbf{F}$ is defined as

$$\mathbf{F} = \mathbf{RU} = \mathbf{VR}, \tag{6}$$

where the tensor $\mathbf{R}$ is an appropriate orthogonal tensor and the tensors $\mathbf{U}$ and $\mathbf{V}$ are symmetric positive-definite tensors known as the right and left stretch tensors, respectively. Equation (7) represents the spectral decompositions of tensors $\mathbf{U}$ and $\mathbf{V}$:

$$\begin{aligned}\mathbf{U} &= \sum_{i=1}^{3} \lambda_i \mathbf{u}^{(i)} \otimes \mathbf{u}^{(i)}, \quad \lambda_i > 0, i \in \{1,2,3\},\\ \mathbf{V} &= \sum_{i=1}^{3} \lambda_i \mathbf{v}^{(i)} \otimes \mathbf{v}^{(i)}, \quad \lambda_i > 0, i \in \{1,2,3\},\end{aligned} \tag{7}$$

where each λ_i refers to one of the principal stretches, $\mathbf{u}^{(i)}$ and $\mathbf{v}^{(i)}$ are the unit eigenvectors of $\mathbf{U}$ and $\mathbf{V}$ known as the Lagrangian and Eulerian principal axes, and $\otimes$ is the sign of tensor product [23, 24].

The tensor function $\mathbf{H}$ is regarded as the function of material response in the configuration B_r with respect to the nominal stress, $\mathbf{S}$, that is, the transpose of the first Piola-Kirchhoff stress. The following equation for the nominal stress, $\mathbf{S}$,

$$\mathbf{S} = \mathbf{H}(\mathbf{F}) = \frac{\partial W}{\partial \mathbf{F}}, \tag{8}$$

is validated for an unconstrained homogeneous hyperelastic material. While the material is incompressible, the arbitrary hydrostatic pressure, p, which is the Lagrange multiplier associated with the material incompressibility, modifies the relation for the nominal stress, $\mathbf{S}$, as

$$\mathbf{S} = \frac{\partial W}{\partial \mathbf{F}} - p\mathbf{F}^{-1}, \quad \det \mathbf{F} = 1. \tag{9}$$

With regard to the relation between the nominal stress tensor, $\mathbf{S}$, and Cauchy stress tensor, $\boldsymbol{\sigma}$,

$$\mathbf{S} = J\mathbf{F}^{-1}\boldsymbol{\sigma}, \tag{10}$$

the Cauchy stress tensor, $\boldsymbol{\sigma}$, could be calculated through (11) in which the symmetric tensor function $\mathbf{G}$ denotes the function of material response in the configuration B_r associated with the Cauchy stress tensor, $\boldsymbol{\sigma}$,

$$\boldsymbol{\sigma} = \mathbf{G}(\mathbf{F}) = J^{-1}\mathbf{F}\frac{\partial W}{\partial \mathbf{F}}. \tag{11}$$

With respect to the arbitrary hydrostatic pressure, p, defined previously, for an incompressible material, the relation for the Cauchy stress tensor, $\boldsymbol{\sigma}$, modifies as [23–25]

$$\boldsymbol{\sigma} = \mathbf{F}\frac{\partial W}{\partial \mathbf{F}} - p\mathbf{I}, \quad \det \mathbf{F} = 1. \tag{12}$$

First, second, and third invariants of $\mathbf{F}$, known as the strain invariants of deformation, which make provision for mapping the area and volume between the deformed configuration, B, and reference configuration, B_r, are computed through

$$\begin{aligned}I_1 &= \operatorname{tr}(\mathbf{F}) = F_{11} + F_{22} + F_{33},\\ I_2 &= \frac{1}{2}\left(F_{ij}F_{ij} - F_{ii}F_{jj}\right),\\ I_3 &= \det(\mathbf{F}) = J,\end{aligned} \tag{13}$$

for an unconstrained isotropic elastic material. The left Cauchy-Green deformation tensor, $\mathbf{B}$, and its principal invariants are calculated as follows:

$$\begin{aligned}\mathbf{B} &= \mathbf{F}\mathbf{F}^{T},\\ I_1^{\mathbf{B}} &= \operatorname{tr}(\mathbf{B}),\\ I_2^{\mathbf{B}} &= \frac{1}{2}\left[\left(I_1^{\mathbf{B}}\right)^2 - \operatorname{tr}\left(\mathbf{B}^2\right)\right],\\ I_3^{\mathbf{B}} &= \det(\mathbf{B}) \equiv (\det \mathbf{F})^2.\end{aligned} \tag{14}$$

For incompressible materials, a slightly different set of principal invariants of $\mathbf{B}$, as represented, is generally employed:

$$\begin{aligned}\overline{I}_1^{\mathbf{B}} &= \frac{I_1^{\mathbf{B}}}{J^{2/3}},\\ \overline{I}_2^{\mathbf{B}} &= \frac{I_2^{\mathbf{B}}}{J^{4/3}},\\ J_{el} &= \sqrt{\det(\mathbf{B})}.\end{aligned} \tag{15}$$

The right Cauchy–Green deformation tensor, $\mathbf{C}$, and its principal invariants are computed similarly.

The Cauchy stress tensor for an unconstrained isotropic elastic material is computed in terms of strain invariants, I_1, I_2, and I_3, as follows:

$$\begin{aligned}\boldsymbol{\sigma} &= \alpha_0\mathbf{I} + \alpha_1\mathbf{B} + \alpha_2\mathbf{B}^2,\\ \alpha_0 &= 2I_3^{1/2}\frac{\partial W}{\partial I_3},\\ \alpha_1 &= 2I_3^{-1/2}\left(\frac{\partial W}{\partial I_1} + I_1\frac{\partial W}{\partial I_2}\right),\\ \alpha_2 &= -2I_3^{-1/2}\frac{\partial W}{\partial I_2},\end{aligned} \tag{16}$$

as the result of the absence of α_0 (because $I_3 = 1$) and the presence of p for incompressible materials, the Cauchy stress tensor changes to

$$\boldsymbol{\sigma} = -p\mathbf{I} + \alpha_1\mathbf{B} + \alpha_2\mathbf{B}^2, \tag{17}$$

for the forenamed materials [21–24], which simplifies to

$$\boldsymbol{\sigma} = -p\mathbf{I} + 2\frac{\partial W}{\partial I_1}\mathbf{B} + 2\frac{\partial W}{\partial I_2}\left(I_1\mathbf{B} - \mathbf{B}^2\right). \tag{18}$$

A variety of stored energy functions have been introduced in the literature that could be employed to model

the nonlinear elastic behavior of soft tissues precisely [26, 27], from the popular long-standing Neo-Hookean model (originated by Treloar in 1943 [28]) and Mooney–Rivlin model (proposed by Rivlin et al. in 1951 [29, 30]) to the state-of-the-art models, for instance, the ones introduced by Limbert in 2011 [31], Nolan et al. in 2014 [32], and Shearer in 2015 (for modeling ligaments and tendons) [33], although this is hitherto an active field of study in the material and biomedical sciences.

Being pertinent to model the behavior of a wide range of materials, for instance the soft tissues and polymers, the Mooney–Rivlin model is one of the conventional hyperelastic models in the literature [34–39]. Two historical attributes have made this model a distinguished one: (1) it is one of the primarily introduced hyperelastic models; (2) it could meticulously predict the nonlinear demeanor observed from some materials, specifically the isotropic rubber-like materials. The most general version of this model has been defined based on (the linear combination of) the first and second strain invariants of deformation, I_1 and I_2. The Mooney–Rivlin strain energy function is expressed as

$$W = C_{10}(I_1 - 3) + C_{01}(I_2 - 3) + \frac{1}{D}(J-1)^2, \tag{19}$$

where C_{10} and C_{01} are the material hyperelastic constants, D is the constant related to the material volumetric response (i.e., the material bulk modulus), and J is the determinant of deformation gradient tensor, $\mathbf{F}$. In addition, regarding the initial shear modulus of material, μ_0, the relation

$$C_{10} + C_{01} = \frac{1}{2}\mu_0, \tag{20}$$

links the two hyperelastic parameters [26, 36]. For incompressible materials, the Mooney–Rivlin strain energy function simplifies to (since $J = 1$)

$$W = C_{10}(I_1 - 3) + C_{01}(I_2 - 3). \tag{21}$$

2.2. Soft Tissue Simulation. To estimate nonlinear elastic parameters of an unidentified tumor using the proposed technique, we have simulated a simplified 3D breast tissue geometry utilizing the FEM software ABAQUS (Dassault Systèmes Simulia Corp., Johnston, RI, USA). The breast tissue is comprised of three partial tissues, fat, fibroglandular, and tumor, located consecutively from outside to inside. We have applied the Mooney–Rivlin hyperelastic model to evaluate the deformation of simulated breast tissue induced by the external excitation.

For estimating hyperelastic parameters of the tumor, we applied a sinusoidal load with frequency of 0.1 Hz (the very low frequency to ignore the inertia effects) to the simulated tissue and registered its response to extract the displacement fields inside the tumor. It is feasible to estimate displacement quantities inside the in vivo tissue by the use of some conventional medical imaging systems such as the US imaging or MRI system. We discuss the methods of calculating the displacements inside the tissue from RF signals or images recorded by the US imaging system in Section 2.4. In order to provide the accessibility to displacement values at some sequential moments, the US images or RF signals should be continuously saved for a period of time.

2.3. Simulation of US RF Signals and Images. When a tissue is inspected with the US imaging system, the tissue is scanned with respect to the probe; in other words, when the tissue is compressed with the probe, the presented features of the tissue in the image seemingly move upward, although in point of fact they would move downward. In furtherance of simulating the postcompression RF signals and B-mode images in the probe coordinate system, instead of the phantom's surface in contact with the probe, the surface in front of the probe was assigned to be moving [40–42].

The RF signals and B-mode images of the simulated phantom while responding to the sinusoidal load have been simulated using the Field II US Simulation Program (A MATLAB® toolbox for US field simulation) [43]. The nodal displacement measurements of the phantom in response to the sinusoidal load have been applied to simulate the postdeformation RF signals and B-mode images in varied deformation states. With the view to simulating the RF signals and B-mode image correlated with a particular deformation state, the correspondent nodal displacement values achieved by the finite deformation analysis have been linearly interpolated to compute the positions of scatterers corresponding to the specified deformation state.

2.4. Estimation of Displacement Field. In addition to the US elastography imaging techniques, motion tracking algorithms have been applied in various US-based methods, for example, blood flow imaging, thermal strain imaging, phase-aberration correction, strain compounding, and temperature imaging, to name a few. The prominence of clinical applications of motion tracking methods has contributed to the significant accrual in the number of relevant investigations and the proposal of a multitude of techniques including phase-domain tactics, time-domain (1D) or space-domain (2D) procedures, and spline-based methods [44].

Among the propounded techniques, the cross-correlation algorithm is known as the gold standard of motion estimation. In this technique, the displacement quantities are computed through searching the locations of the maximums of cross-correlation values between corresponding axial-lateral grids of small windows (with high overlap) in the pre- and postdeformation frames recorded from the medium. The shifts between the pre- and postdeformation windows quantify the displacement field in the medium [44, 45].

The cross-correlation algorithm with guided search has been applied to initially estimate the displacement quantities in the tumor at the selected step times from the simulated pre- and postdeformation RF signals and B-mode images, and afterwards evaluate the errors of displacement estimates. The calculated displacement values of proximate regions (i.e., previous samples and lines) have been exploited to reduce the search area and therefore significantly decrease the computational expense of the cross-correlation algorithm. By the use of simulated RF signals, although we

devoted more time to calculate the displacement values (because of their higher sampling rate) compared with the US images, we achieved more accurate displacement estimates with higher spatial resolution.

2.5. Estimation of Hyperelastic Parameters. Regarding the aforementioned explanation related to the finite strain theory (also known as large deformation theory), when a uniaxial stress, $\boldsymbol{\sigma}$, is applied to the material, the deformation gradient tensor, $\mathbf{F}$, could be computed as

$$\mathbf{F} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \tag{22}$$

where λ_1, λ_2, and λ_3 are the principal stretches with respect to the set of coordinate axes (corresponding with $x_i = \lambda_i X_i$, $i = 1,2,3$). The principal invariants, I_1, I_2, and I_3, could be stated further in terms of principal stretches as follows:

$$\begin{aligned} I_1 &= \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \\ I_2 &= \lambda_1^2 \cdot \lambda_2^2 + \lambda_2^2 \cdot \lambda_3^2 + \lambda_3^2 \cdot \lambda_1^2, \\ I_3 &= \lambda_1^2 \cdot \lambda_2^2 \cdot \lambda_3^2. \end{aligned} \tag{23}$$

If $\boldsymbol{\lambda}$ symbolizes the stretch parallel to the stress applied to the medium in line with the first coordinate axis (that is equal to the λ_1 set), two assumptions, (1) the equality of deformations in the two other coordinate axes and (2) the incompressibility of the medium ($I_3 = 1$), result in simplifying the relation of the Cauchy stress tensor (18), to [35, 46]

$$\boldsymbol{\sigma} = 2\left(\boldsymbol{\lambda}^2 - \boldsymbol{\lambda}^{-1}\right)\left(\frac{\partial W}{\partial I_1} + \frac{1}{\boldsymbol{\lambda}}\frac{\partial W}{\partial I_2}\right). \tag{24}$$

As explicated in the previous sections, the displacement field inside the in vivo tumor could be measured in a noninvasive way. With the purpose of estimating Mooney–Rivlin hyperelastic parameters of the tumor just by using the displacement quantities, we manipulate (25), that is, the stress-stretch relation of the hyperelastic model,

$$\boldsymbol{\sigma} = 2\left(\boldsymbol{\lambda}^2 - \boldsymbol{\lambda}^{-1}\right)\left(C_{10} + \boldsymbol{\lambda}^{-1} C_{01}\right). \tag{25}$$

The uniaxial load applied to the medium in alignment with the first coordinate axis practically generates the axial strain in the same direction; therefore, the strain and relation between the stress and strain could be expressed as [47, 48]

$$\begin{aligned} \boldsymbol{\varepsilon} &= \boldsymbol{\lambda} - 1, \\ \boldsymbol{\sigma} &= 2\left((\boldsymbol{\varepsilon} + 1)^2 - (\boldsymbol{\varepsilon} + 1)^{-1}\right)\left(C_{10} + (\boldsymbol{\varepsilon} + 1)^{-1} C_{01}\right). \end{aligned} \tag{26}$$

With respect to the stress-strain relation, (26), the iterative algorithm propounded for estimating Mooney–Rivlin hyperelastic parameters of a completely unknown interior tissue (i.e., tumor) could be described in the following steps:

(1) Image the tissue and its adjacent mediums by the use of clinical US imaging system before/while applying a sinusoidal load with low frequency (to annul the inertia) to the exterior medium. In other words, record the relevant US RF signals or images.

(2) Extract the displacement fields inside the tissue at some sequential step times (i.e., eight consecutive instants) from the recorded US RF signals or images.

(3) Simulate the tissue and its neighboring mediums by making use of one of the FEM softwares corresponding to

   (a) the recorded predeformation US images
   (b) the loading specifications
   (c) the boundary conditions

   It is assumed that just the tumor and its mechanical characteristics are unidentified. The mechanical parameters of almost all healthy soft tissues have been reported in the literature. The parameters have been predominantly estimated by performing meticulous in vivo or ex vivo experiments.

(4) Consider the elastic modulus, E, and Poisson's ratio, v, of the simulated tumor, named elastic tumor, equal 1 Pa and 0.5, respectively.

(5) Compute the displacement fields inside the particular elastic tumor at the same consecutive step times by dint of the selected FEM software.

(6) Calculate the real elastic modulus of the understudy tumor, E_{realt}, with the help of the MATLAB® software (The MathWorks, Inc., Natick, Massachusetts, USA) using

   (a) the axial displacement quantities of several points of the tumor at some consecutive instants (based on the achieved outcomes, the axial displacement values of twelve points of the tumor at eight step times), $\mathbf{Y}_{\text{realt}}$
   (b) the axial displacement values of the identical points of the elastic tumor at the same moments, $\mathbf{D}$
   (c) the relation [47–49]

$$E_{\text{realt}} = \frac{\mathbf{D}^T \mathbf{D}}{\mathbf{D}^T \mathbf{Y}_{\text{realt}}}. \tag{27}$$

(7) Specify a set of strains, $\boldsymbol{\varepsilon}$, and compute the correspondent stresses, $\boldsymbol{\sigma}$, using (28) in accordance with (in the first iteration):

   (a) The strain field in the tumor could be roughly approximated from the displacement estimates.
   (b) The achieved results imply the selection of a set of high strains in the first iteration since the strain values are being modified in some steps of the proposed iterative algorithm; therefore, the strain values could be reduced uniformly.
   (c) Slight changes in the stress values, for instance by the use of the normal distribution function, might cause the stress and strain values to conform more effectively to the Mooney–Rivlin hyperelastic model assigned to the tumor,

$$\boldsymbol{\sigma} = E_{\text{realt}} \boldsymbol{\varepsilon}. \tag{28}$$

(8) Compute hyperelastic parameters of the Mooney–Rivlin model, C_{10} and C_{01}, using the stress and

strain sets and the relation between stress and strain, as represented,

$$\boldsymbol{\sigma} = 2\Big((\boldsymbol{\varepsilon}+1)^2 - (\boldsymbol{\varepsilon}+1)^{-1}\Big)\Big(C_{10} + (\boldsymbol{\varepsilon}+1)^{-1}C_{01}\Big), \tag{29}$$

with the help of MATLAB algorithms, for instance, the regression algorithms.

(9) Assign the estimated hyperelastic parameters to the simulated tumor and compute the displacement fields inside the tumor at the determinate step times by means of the FEM software.

(10) Calculate the elastic constant of the simulated tumor, E_{estt}, as explained previously in step 6, using the axial displacement quantities (of the appointed points at the selected step times) of the simulated tumor, $\mathbf{Y}_{\text{estt}}$, and the elastic tumor, $\mathbf{D}$ (determined in step 6), by applying the relation

$$E_{\text{estt}} = \frac{\mathbf{D}^T\mathbf{D}}{\mathbf{D}^T\mathbf{Y}_{\text{estt}}}. \tag{30}$$

(11) Appraise the estimated hyperelastic parameters for the tumor through considering the following:

(a) The error of the computed axial displacement values of the selected points at the specified moments, $\mathbf{Y}_{\text{estt}}$, by comparing them with the correspondent displacement quantities estimated from the recorded US RF signals or images, $\mathbf{Y}_{\text{realt}}$

(b) The error of the elastic constant calculated for the tumor, E_{estt}, by comparing it with the real elastic modulus of the understudy tumor, E_{realt}, estimated in step 6

The errors of the hyperelastic parameters estimated for the tumor, by comparing them with the real hyperelastic parameters of the tumor, could not be considered because it has been assumed that the tumor is entirely obscure.

(12) Alter the set of strains specified in step 7 (based on the above explanation, decrease them regularly) and repeat steps 7 to 12. By reducing the strain values steadily, the error of the estimated elastic parameter for the tumor and the error of the calculated displacement quantities in the tumor are decreasing below the defined tolerance values, as illustrated,

$$\begin{aligned} \left\| \mathbf{Y}^k_{\text{estt}} - \mathbf{Y}_{\text{realt}} \right\| &\le e_{\text{displacement}}, \\ \left\| E^k_{\text{estt}} - E_{\text{realt}} \right\| &\le e_{\text{elastic}}, \end{aligned} \tag{31}$$

where e_{elastic} and $e_{\text{displacement}}$ are the tolerance values and k represents the number of iterations of the algorithm. By decreasing the strain values chosen with regard to the mentioned conditions, the strain and stress values successively adjust more to the Mooney–Rivlin stress-strain relationship of the tumor.

3. Results

3.1. Soft Tissue Simulation. The breast tissue (with the dimensions of $100 \times 60 \times 20\,\text{mm}^3$), simulated using the FEM software ABAQUS, has been depicted in Figure 1. The breast tissue consists of three partitions, namely, fat, fibroglandular, and tumor. The Mooney–Rivlin hyperelastic model has been applied to obtain the response of simulated breast tissue to the external sinusoidal load. The Mooney–Rivlin material constants of the named breast tissues have been presented in Table 1. Since we have utilized the elastic parameter of the tumor for estimating its hyperelastic parameters, we have additionally reported the elastic parameters of the named breast tissues in Table 1. The linear and nonlinear elastic parameters have been supposed to be constant throughout each tissue partition. This set of hyperelastic parameters has been broadly utilized in the literature to simulate the breast tissue [38, 50–54].

The mesh considered for the simulated phantom consists of 183783 second-order (quadratic) tetrahedral hybrid elements (C3D10H) with 259813 nodes. The convergence analyses have warranted the accuracy of the simulation results. With reference to the explanations in Section 2.3, the number of nodes in the simulated medium should significantly be increased to precisely calculate the positions of scatterers after applying the load to the medium. With regard to the boundary conditions and the load applied to the tissue (represented in Figure 1), the postcompression RF signals and B-mode images have been simulated in the probe coordinate system. Two snapshots of the response of the simulated breast tissue to the sinusoidal load have been demonstrated in Figure 2.

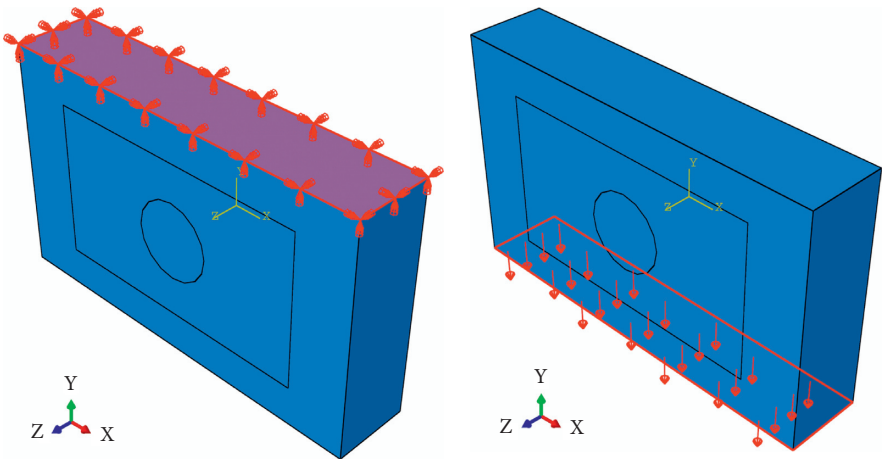
3.2. Simulation of US RF Signals and Images. The RF signals and B-mode images of part (with the dimensions of $50 \times 60 \times 10\,\text{mm}^3$) of the simulated breast tissue which encircles the tumor, as illustrated in Figure 3, have been simulated using the Field II US Simulation Program. In the Field II US Simulation Program,

(a) The properties considered to model the probe array and simulate the US RF signals and images are as follows:

(1) Linear array (with 64 active elements)
(2) Transducer center frequency of 3.5×10^6 Hz
(3) Sampling frequency of 100×10^6 Hz
(4) Transmit focus of 70 mm (in depth)
(5) Element's width (the distance between the elements or the pitch of the probe array) of 0.44 mm (equal to the wavelength)
(6) Element's height of 5 mm
(7) Element's kerf of 0.05 mm
(8) Lateral spatial spacing of 0.08 mm (512 scan lines in the image)

(b) With regard to the elastic and hyperelastic parameters of the tumor, for scatterers which have resided within the tumor, the amplitudes are set to zero.

The postdeformation RF signals and B-mode images in eight deformation states of the phantom (corresponding to



(a) (b)

Figure 1: The simulated breast tissue comprises three parts, namely, fat, fibroglandular, and tumor (from outside to inside).

Table 1: The elastic and Mooney–Rivlin hyperelastic constants of breast tissues [38].

Hyperelastic and elastic parameters	Fat	Fibroglandular	Tumor
C_{10} (Pa)	2000	3500	10000
C_{01} (Pa)	1333	2333.3	6667
E (kPa)	20	35	100

eight sequential step times: 7.75 s, 8.00 s, 8.25 s, 8.50 s, 8.75 s, 9.00 s, 9.25 s, and 9.50 s, while responding to the sinusoidal load) have been simulated using the displacements of the phantom's nodes computed by the finite deformation analysis. The nodal displacement values correlated with a particular deformation state have been linearly interpolated to compute the positions of scatterers and simulate the correspondent RF signals and B-mode images thereafter. Two simulated postdeformation B-mode images (pre- and postdeformation images) associated with the step times of 8.00 s and 9.00 s have been represented in Figure 3.

3.3. Estimation of Hyperelastic Parameters. After simulating the pre- and postdeformation RF signals and B-mode images correlated with the defined deformation states, the cross-correlation algorithm with guided search, as briefly described in Section 2.4, has been employed to initially estimate displacement quantities inside the tumor at the selected step times and afterwards evaluate the errors of displacement estimates.

The suggested iterative algorithm, comprehensively explicated in Section 2.5, has been applied to extract the Mooney–Rivlin hyperelastic parameters of the tumor from the axial displacement values of some points of tumor at the specified step times. As represented in Table 2, precise estimates of hyperelastic parameters of the tumor have been achieved. The automatic iteration of the algorithm would be feasible through bilaterally connecting the MATLAB and FEM softwares.

4. Discussion

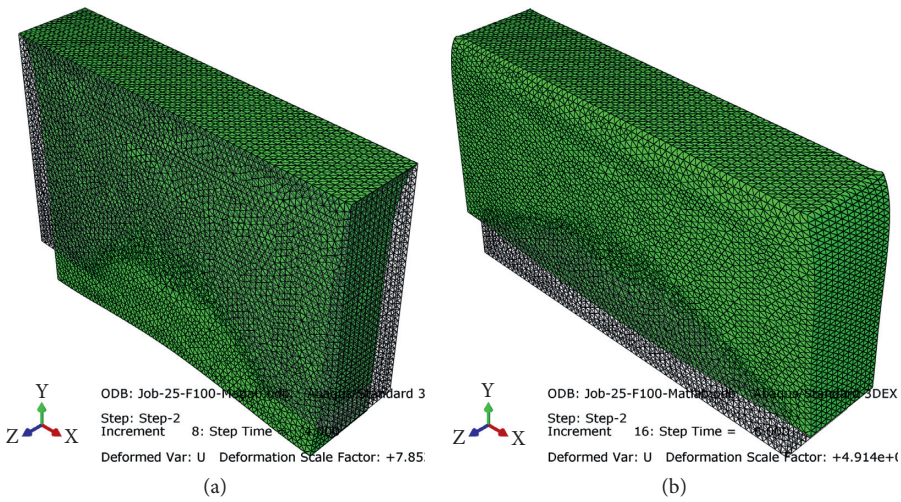
In the majority of diversified approaches proffered for estimating elastic parameters of soft tissues, particularly the nonlinear ones, for instance, the techniques proposed by MacManus et al. [55], Esmaeili et al. [56], Omidi et al. [57], Roy and Desai [58], Liu et al. [59], Boonvisut and Çavuşoğlu [60], and Wang et al. [61], to mention but a few, the alterations (of precise values) of at least two deformation variables, which are associated with the mechanical characteristics of soft tissues, have been exploited. The assessment of recommended techniques would reveal that the direct dependencies of methodologies to deformation variables except the displacement (and strain) have impelled the researchers to carry out experiments on ex vivo tissues, or perform invasive procedures to precisely measure the variables; consequently, the emphasis of recent investigations should be on advancing noninvasive methods with the capability to accurately estimate nonlinear elastic parameters of tissues.

The hyperelastic constitutive theory takes two types of nonlinearities perceived in responses of soft tissues, into consideration [62, 63]:

(a) The material nonlinearity of the stress-strain relation, known as the physical nonlinearity

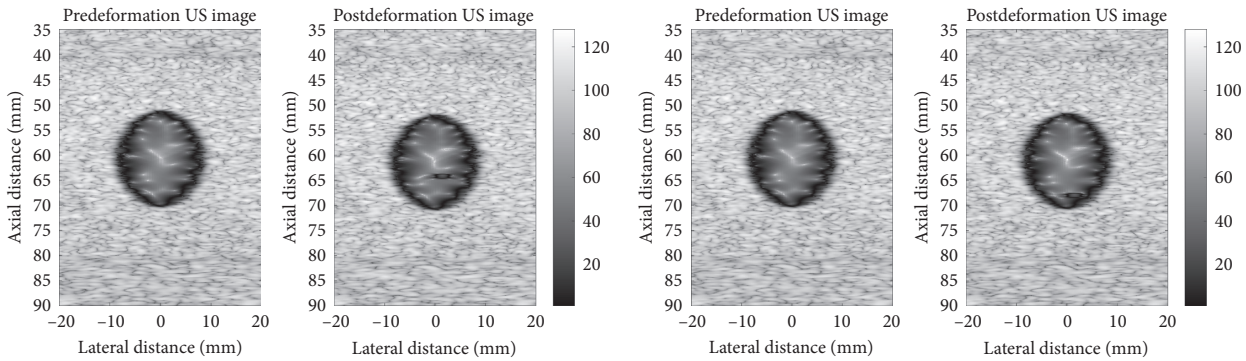
(b) The nonlinearity of the strain-displacement relation, called the geometrical nonlinearity

consequently, it has been regarded as one of the best practical theories for formulating mechanical behaviors of soft tissues. To the best of our knowledge, amongst the strategies proposed for quantifying hyperelastic parameters of materials, two methods founded on the displacement fields inside and on the boundary of the medium (i.e., phantoms) which have been introduced by Mehrabian and Samani [26, 27, 64] and Hajhashemkhani and Hematiyan [47, 48], respectively, could be utilized to noninvasively reconstruct hyperelastic parameters of in vivo tissues. The displacement field inside the understudy medium could be extracted from RF signals or images recorded by a clinical



(a) (b)

Figure 2: The responses of the simulated breast tissue at (a) $t = 4.00$ s and (b) $t = 8.00$ s after starting to apply the sinusoidal load.



(a) (b)

Figure 3: The simulated pre- and postdeformation B-mode images of breast phantom, based on its states at (a) $t = 8.00$ s and (b) $t = 9.00$ s after starting to apply the sinusoidal load.

Table 2: The elastic and hyperelastic parameters estimated for the tumor.

Estimated elastic and hyperelastic parameters		
2 estimates		
E_{realt} (kPa)	88908.41	
Error of E_{realt} (%)	11.09	
C_{10} (Pa)	9426.98	10005.05
C_{01} (Pa)	6702.30	5992.82
Error of C_{10} (%)	5.73	0.05
Error of C_{01} (%)	0.53	10.11
E_{estt} (kPa)	88711.93	88658.27
Error of E_{estt} (%)	0.22	0.28
Error of $\mathbf{Y}_{estt}$ (%)	0.28	0.35

US imaging system, for instance, by using the conventional cross-correlation method.

In the technique recommended by Mehrabian and Samani [26, 27, 64], the displacement quantities of a large number of contiguous points inside the medium should be utilized to calculate the defined coefficient matrix which correlates the stress distribution computed for the tissue (with the help of a finite element model of the tissue deformation) to its hyperelastic parameters. The displacement values at some boundary points of the understudy medium have been manipulated by Hajhashemkhani and Hematiyan [47, 48] for characterizing its nonlinear material constants.

On account of the explanations provided by Mehrabian and Samani [26, 27, 64] and Hajhashemkhani and Hematiyan [47, 48] and the results achieved through the implementation of their methods (part of them published in our paper [65]), it has been realized that precise estimates of hyperelastic parameters of the understudy medium could be attained through the following:

(a) Accurately calculating the displacement quantities, respectively, in a multitude of adjacent points of the medium and in several boundary points, which might not be possible using registered US images or RF signals

(b) Applying proper regularization techniques, for instance, the Tikhonov regularization, Truncated

Singular Value Decomposition (SVD), and Wiener Filtering methods

(c) Even considering appropriate initial guesses of the hyperelastic parameters, as indicated by Hajhashemkhani and Hematiyan [47, 48], Aghajani et al. [66], and Kim and Srinivasan [67]

The limitations of the techniques propounded with the aim of reconstructing nonlinear elastic parameters of in vivo soft tissues persuade us to concentrate on developing a more practical method with consistent results. Our primary upshot represented in Section 3 confirms that accurate estimates of hyperelastic parameters of the understudy tissue (i.e., tumor) could be obtained on the basis of the displacement values of some points inside the tissue, which has been excited by a low frequency sinusoidal load, even when there is no prior knowledge of the tissue.

The displacement quantities of the selected points might be computed approximately, for instance, by the use of the cross-correlation technique as a consequence of recording low-quality US images or RF signals or other attributes; therefore, we have evaluated the consistency of calculated values for the hyperelastic constants by applying errors with normal distribution to the measured displacement fields in the tissue stimulated by the sinusoidal load. At this point, the average errors of the displacement values estimated for the selected points from the simulated US RF signals and images using the cross-correlation algorithms (without/with guided search with respect to the calculated displacements of previous lines or samples) have been regarded. The achieved results have been demonstrated in Table 3.

It should be considered that the imprecise estimates of displacement fields inside the tumor affect all the computed parameters and errors, even the real elastic modulus of the understudy tumor; therefore, the results presented in Table 3 could not be compared. Similar to the case where the displacement values of the appointed points are exact,

(a) the error of calculated axial displacement values of the selected points at the specified moments, $\mathbf{Y}_{\text{estt}}$

(b) the error of elastic constant computed for the tumor, E_{estt}

(as explained in step 11 of the proposed algorithm in Section 2.5) have been considered except in the situations where the displacement errors are significant.

The convergence of the aforementioned errors to values, which might not be small errors, specifies the best estimates of hyperelastic parameters when the displacement values are highly inaccurate, as represented,

$$\begin{aligned} \left\| \mathbf{Y}_{\text{estt}}^{k} - \mathbf{Y}_{\text{estt}}^{k-1} \right\| &\leq e'_{\text{displacement}}, \\ \left\| E_{\text{estt}}^{k} - E_{\text{estt}}^{k-1} \right\| &\leq e'_{\text{elastic}}, \end{aligned} \tag{32}$$

where e'_{elastic}, $e'_{\text{displacement}}$, and k are, respectively, the specified tolerance values and the number of iterations of the algorithm. Provided that the set of strains (required in step 7 of the algorithm described in Section 2.5) is selected properly based on the displacement fields calculated for the tumor, precise estimates of tumor's hyperelastic parameters could be obtained. With regard to the outcomes summarized in Table 3, it is deduced that the suggested method is strongly resistant to the displacement errors.

Table 3: The hyperelastic parameters estimated for the tumor using inaccurate displacement measurements.

Estimated hyperelastic parameters	Inaccurate displacement measurements			
	Error 2%	Error 5%	Error 8%	Error 10%
C_{10} (Pa)	9386.78	9484.46	9815.78	9858.99
C_{01} (Pa)	6889.50	6817.80	6437.30	6465.64
Error of C_{10} (%)	6.13	5.16	1.84	1.41
Error of C_{01} (%)	3.34	2.26	3.45	3.02
Error of E_{estt} (%)	0.53	0.39	1.89	2.29
Error of $\mathbf{Y}_{\text{estt}}$ (%)	1.77	4.59	6.28	7.96

The US RF signals recorded by means of the Antares Siemens system (Issaquah, WA) at the center frequency of 6.67 MHz from an elastography phantom (CIRS elastography phantom, Norfolk, VA) have been utilized to evaluate the suggested method experimentally. The signals were registered via a VF10-5 linear array at a sampling rate of 40 MHz by Rivaz et al. to assess the performance of the proposed real-time static elastography techniques which were based on the analytic minimization of regularized cost functions. Young's moduli of the lesion and surrounding medium have been reported 56 kPa and 33 kPa, respectively, while the phantom is under compression [68, 69].

The enhanced cross-correlation algorithm, in that the search regions were minimized with respect to the estimated displacements of previous lines or samples, has been employed to compute the axial displacement field in the compressed phantom. The Kalman filtering, introduced by Rivaz et al. [68], has been applied to calculate the strain field in the compressed phantom from the displacement measurements. Minor differences between the displacement fields estimated by the use of the enhanced cross-correlation algorithm and analytic minimization method validate the results of the former technique.

The US images of the phantom constructed from the recorded RF signals and the estimated displacement and strain fields have been represented in Figure 4. Following the instructions in Section 2.5, the elastic and hyperelastic parameters of the lesion could be calculated from the estimated axial displacement and strain fields in the lesion. The percent error of the elastic parameter computed for the lesion, on the basis of the explanations in step 6 of the algorithm, is 15.06%; in other words, E_{realt} has been estimated 47565.72 Pa. The relation between stress, strain, and the parameters of the Mooney–Rivlin hyperelastic model, C_{10} and C_{01}, has been manipulated to compute the hyperelastic parameters of the lesion. The values of 6871.65 Pa and 1020.00 Pa have been obtained for the mentioned parameters.

On the basis of the achieved results summarized in Sections 3 and 4, it is concluded that the main objectives that have been accomplished in this paper are as follows:

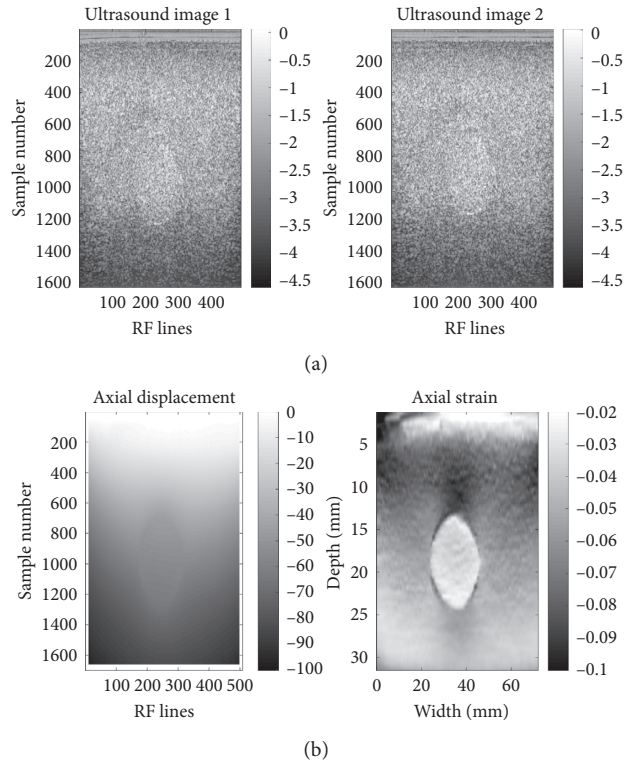


Figure 4: (a) Two US images of the phantom and (b) the estimated axial displacement and strain fields. The recorded RF signals of the elastography phantom (by Rivaz et al. [68, 69]), before/after it was compressed, have been used to create the US images and compute the axial displacement and strain fields.

(a) The identification of an entirely unknown tissue (i.e., tumor) in a soft tissue through its nonlinear elastic parameters

(b) The feasibility to use imprecise measurements of displacement in the tumor, which are extracted from the signals or images recorded from the soft tissue while responding to a sinusoidal stimulus (with low frequency)

5. Conclusion

In this paper, the noninvasive diagnosis of tumors in soft tissues, such as the breast, through their nonlinear elastic parameters has been evaluated by the use of a novel iterative algorithm founded on the principle of US elastography technique. The achieved results could be undoubtedly considered the validation of the precise estimation of hyperelastic constants of an undiagnosed pathology which has been accomplished:

(a) By manipulating the relation between stress, strain, and the parameters of a hyperelastic model as explicated in the paper

(b) Based on the response of tissue to the sinusoidal load with low frequency, indeed the displacement quantities of a few points of tissue at certain step times

The displacement fields inside the tissue could be noninvasively computed from the data recorded by the employment of conventional medical imaging modalities, for instance, the RF signals or images registered by the US imaging system. Even by processing approximate displacement measurements, accurate estimates of the material constants could be obtained. The competency of the proposed method to estimate nonlinear elastic constants of normal and abnormal in vivo tissues will be further appraised in the future research.

Data Availability

Since the research is still in progress, the authors have decided to make data available upon request.

Disclosure

The funding had no role in the study design, data collection, or analysis.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] J. L. Gennisson, T. Deffieux, M. Fink, and M. Tanter, "Ultrasound elastography: principles and techniques," *Diagnostic and Interventional Imaging*, vol. 94, no. 5, pp. 487–495, 2013.

[2] A. A. Oberai, N. H. Gokhale, S. Goenezen et al., "Linear and nonlinear elasticity imaging of soft tissue in vivo: demonstration of feasibility," *Physics in Medicine and Biology*, vol. 54, no. 5, pp. 1191–1207, 2009.

[3] J. Palacio-Torralba, D. W. Good, S. A. McNeill, R. L. Reuben, and Y. Chen, "Histology-based homogenization analysis of soft tissue: application to prostate cancer," *Journal of the Royal Society Interface*, vol. 14, no. 129, article 20170088, 2017.

[4] M. M. Dastjerdi and A. Mahloojifar, "Optical elastography: from theory to clinical applications," *Journal of Lasers in Medicine*, vol. 5, no. 3-4, pp. 6–20, 2008.

[5] T. A. Krouskop, T. M. Wheeler, F. Kallel, B. S. Garra, and T. Hall, "Elastic moduli of breast and prostate tissues under compression," *Ultrasonic Imaging*, vol. 20, no. 4, pp. 260–274, 1998.

[6] A. Samani, J. Zubovits, and D. Plewes, "Elastic moduli of normal and pathological human breast tissues: an inversion-technique-based investigation of 169 samples," *Physics in Medicine and Biology*, vol. 52, no. 6, pp. 1565–1576, 2007.

[7] A. Samani and D. Plewes, "An inverse problem solution for measuring the elastic modulus of intact ex vivo breast tissue tumours," *Physics in Medicine and Biology*, vol. 52, no. 5, pp. 1247–1260, 2007.

[8] A. Lyshchik, T. Higashi, R. Asato et al., "Elastic moduli of thyroid tissues under compression," *Ultrasonic Imaging*, vol. 27, no. 2, pp. 101–110, 2005.

[9] G. Soza, R. Grosso, C. Nimsky, P. Hastreiter, R. Fahlbusch, and G. Greiner, "Determination of the elasticity parameters of brain tissue with combined simulation and registration," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 1, no. 3, pp. 87–95, 2005.

[10] K. Hoyt, B. Castaneda, M. Zhang et al., "Tissue elasticity properties as biomarkers for prostate cancer," *Cancer Biomarkers*, vol. 4, no. 4-5, pp. 213–225, 2008.

[11] P. Schiavone, F. Chassat, T. Boudou, E. Promayon, F. Valdivia, and Y. Payan, "In vivo measurement of human brain elasticity using a light aspiration device," *Medical Image Analysis*, vol. 13, no. 4, pp. 673–678, 2009.

[12] J. J. O'Hagan and A. Samani, "Measurement of the hyperelastic properties of 44 pathological ex vivo breast tissue samples," *Physics in Medicine and Biology*, vol. 54, no. 8, pp. 2557–2569, 2009.

[13] R. Moran, J. H. Smith, and J. J. Garcia, "Fitted hyperelastic parameters for human brain tissue from reported tension, compression, and shear tests," *Journal of Biomechanics*, vol. 47, no. 15, pp. 3762–3766, 2014.

[14] W. Maurel, Y. Wu, D. Thalmann, and N. M. Thalmann, *Biomechanical Models for Soft Tissue Simulation*, Springer-Verlag, Berlin, Heidelberg, Germany, 1998.

[15] L. A. Mihai, L. Chin, P. A. Janmey, and A. Goriely, "A comparison of hyperelastic constitutive models applicable to brain and fat tissues," *Journal of the Royal Society Interface*, vol. 12, no. 110, article 20150486, 2015.

[16] A. Samani and D. Plewes, "A method to measure the hyperelastic parameters of ex vivo breast tissue samples," *Physics in Medicine and Biology*, vol. 49, no. 18, pp. 4395–4405, 2004.

[17] S. Avril and S. Evans, *Material Parameter Identification and Inverse Problems in Soft Tissue Biomechanics*, Springer International Publishing AG, Cham, Switzerland, 2017.

[18] I.-S. Liu, "Basic principles of constitutive theories," in *Continuum Mechanics*, pp. 63–95, Springer-Verlag, Berlin, Heidelberg, Germany, 1st edition, 2002.

[19] I.-S. Liu, "Constitutive theories: basic principles," in *Continuum Mechanics, Encyclopedia of Life Support Systems (EOLSS)*, vol. 1, p. 198, UNESCO Publications, Oxford, UK, 2011.

[20] Y. Payan and J. Ohayon, *Biomechanics of Living Organs: Hyperelastic Constitutive Laws for Finite Element Modeling*, Academic Press, Cambridge, MA, USA, 1st edition, 2017.

[21] A. F. Bower, *Applied Mechanics of Solids*, CRC Press Book, Taylor & Francis Group, London, UK, 2009.

[22] A. F. Bower, *Continuum Mechanics*, School of Engineering, Brown University, Providence, RI, USA, 2012, http://www.brown.edu/Departments/Engineering/Courses/En221/.

[23] G. A. Holzapfel and R. W. Ogden, *Biomechanics of Soft Tissue in Cardiovascular Systems*, Springer-Verlag, Vol. 441, Springer-Verlag, Wien, Austria, 2003.

[24] C. A. Felippa, *Nonlinear Finite Element Methods*, University of Colorado, Boulder, CO, USA, 2001, https://www.colorado.edu/engineering/cas/courses.d/NFEM.d/Home.html.

[25] R. M. Hackett, *Hyperelasticity Primer*, Springer International Publishing AG, Cham, Switzerland, 2016.

[26] H. Mehrabian, "Soft tissue hyperelastic parameter reconstruction for breast cancer assessment," *School of Graduate and Postdoctoral Studies*, M.Sc. Thesis, The University of Western Ontario, London, ON, Canada, 2008.

[27] H. Mehrabian, G. Campbell, and A. Samani, "A constrained reconstruction technique of hyperelasticity parameters for breast cancer assessment," *Physics in Medicine and Biology*, vol. 55, no. 24, pp. 7489–7508, 2010.

[28] L. R. G. Treloar, "The elasticity of a network of long-chain molecules. I," *Transactions of the Faraday Society*, vol. 39, pp. 36–41, 1943.

[29] M. Mooney, "A theory of large elastic deformation," *Journal of Applied Physics*, vol. 11, no. 9, pp. 582–592, 1940.

[30] R. S. Rivlin and D. Saunders, "Large elastic deformations of isotropic materials. VII. Experiments on the deformation of rubber," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 243, no. 865, pp. 251–288, 1951.

[31] G. Limbert, "A mesostructurally-based anisotropic continuum model for biological soft tissues-decoupled invariant formulation," *Journal of the Mechanical Behavior of Biomedical Materials*, vol. 4, no. 8, pp. 1637–1657, 2011.

[32] D. Nolan, A. Gower, M. Destrade, R. Ogden, and J. McGarry, "A robust anisotropic hyperelastic formulation for the modelling of soft tissue," *Journal of the Mechanical Behavior of Biomedical Materials*, vol. 39, pp. 48–60, 2014.

[33] T. Shearer, "A new strain energy function for the hyperelastic modelling of ligaments and tendons based on fascicle microstructure," *Journal of Biomechanics*, vol. 48, no. 2, pp. 290–297, 2015.

[34] P. Martins, R. Natal Jorge, and A. Ferreira, "A comparative study of several material models for prediction of hyperelastic properties: application to silicone-rubber and soft tissues," *Strain*, vol. 42, no. 3, pp. 135–147, 2006.

[35] S. Zaeimdar, "Mechanical characterization of breast tissue constituents for cancer assessment," *School of Mechatronic Systems Engineering*, M.Sc. Thesis, Faculty of Applied Sciences, Simon Fraser University, BC, Canada, 2014.

[36] R. Tobajas, E. Ibartz, and L. Gracia, "A comparative study of hyperelastic constitutive models to characterize the behavior of a polymer used in automotive engines," in *Proceedings of 2nd International Electronic Conference on Materials, Sciforum Electronic Conference Series*, Chaired by Prof. Maryam Tabrizian, Biomedical Engineering Department, McGill University, Montreal, Canada, 2016.

[37] F. Safshekan, M. Tafazzoli-Shadpour, M. Abdouss, and M. B. Shadmehr, "Mechanical characterization and constitutive modeling of human trachea: age and gender dependency," *Materials*, vol. 9, no. 6, p. 456, 2016.

[38] H. M. Yin, L. Z. Sun, G. Wang, T. Yamada, J. Wang, and M. W. Vannier, "ImageParser: a tool for finite element generation from three-dimensional medical images," *BioMedical Engineering OnLine*, vol. 3, no. 1, p. 31, 2004.

[39] L. Ruggiero, H. Sol, H. Sahli, S. Adriaenssens, and N. Adriaenssens, in *Conference Proceedings of the Society for Experimental Mechanics Series, Published in Mechanics of Biological Systems and Materials*, vol. 2, T. Proulx, Ed., pp. 19–32, Springer, New York, NY, USA, 2011.

[40] S.-Y. Sun, B. W. Anthony, and M. W. Gilbertson, "Trajectory-based deformation correction in ultrasound images," in *Proceedings of SPIE, vol. 7629, Medical Imaging 2010: Ultrasonic Imaging, Tomography, and Therapy*, SPIE, Bellingham, WA, USA, 2008.

[41] S.-Y. Sun, "Deformation correction in ultrasound imaging in an elastography framework," *Department of Electrical Engineering and Computer Science*, M.Sc. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2010.

[42] O. Goksel, R. Zahiri-Azar, and S. E. Salcudean, "Simulation of ultrasound radio-frequency signals in deformed tissue for validation of 2d motion estimation with sub-sample accuracy," in *Proceedings of 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, pp. 2159–2162, Lyon, France, August 2007.

[43] J. A. Jensen, "Field: a program for simulating ultrasound systems," in *Proceedings of 10th Nordic-Baltic Conference on Biomedical Imaging, Published in Medical & Biological Engineering & Computing*, vol. 34, no. 1, pp. 351–353, 1996.

[44] J. Luo and E. E. Konofagou, "A fast normalized cross-correlation calculation method for motion estimation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 57, no. 6, pp. 1347–1357, 2010.

[45] A. Fenster and J. C. Lacefield, *Ultrasound Imaging and Therapy*, CRC Press Book, Taylor & Francis Group, London, UK, 2015.

[46] M. Rackl, "Curve fitting for Ogden, Yeoh and polynomial models," in *Proceedings of the ScilabTEC, 7th International Scilab Users Conference, International Partnership Committee (IPC)*, Scilab Enterprises, Paris, France, 2015.

[47] M. Hajhashemkhani, "Identification of material parameters of isotropic and anisotropic hyperelastic materials," M.Sc. thesis, School of Mechanical Engineering, Shiraz University, Shiraz, Iran, 2013.

[48] M. Hajhashemkhani and M. R. Hematiyan, "Determination of material parameters of isotropic and anisotropic hyper-elastic materials using boundary measured data," *Journal of Theoretical and Applied Mechanics*, vol. 53, no. 4, pp. 895–910, 2015.

[49] M. R. Hematiyan, A. Khosravifard, Y. C. Shiah, and C. L. Tan, "Identification of material parameters of two-dimensional anisotropic bodies using an inverse multi-loading boundary element technique," *Computer Modeling in Engineering and Sciences*, vol. 87, no. 1, pp. 55–76, 2012.

[50] T.-C. Shih, D. Liu, J.-H. Chen et al., "Computer simulation of breast compression based on segmented breast and fibroglandular tissues on MRI," in *Proceedings of International Society for Magnetic Resonance in Medicine (ISMRM) 17*, HI, USA, 2009.

[51] T.-C. Shih, J.-H. Chen, D. Chang et al., "Quantitative analysis of projection breast density changes at different compression angles based on 3D MRI," in *Proceedings of International Society for Magnetic Resonance in Medicine (ISMRM) 18*, Stockholm, Sweden, 2010.

[52] T.-C. Shih, J.-H. Chen, D. Liu et al., "Computational simulation of breast compression based on segmented breast and fibroglandular tissues on magnetic resonance images," *Physics in Medicine and Biology*, vol. 55, no. 14, pp. 4153–4168, 2010.

[53] Y.-L. Liu, P.-Y. Liu, M.-L. Huang, J.-T. Hsu, R.-P. Han, and J. Wu, "Simulation of breast compression in mammography using finite element analysis: a preliminary study," *Radiation Physics and Chemistry*, vol. 140, pp. 295–299, 2017.

[54] E. M. Diniz, "Simulaçao de cirurgia mamária usando Elementos Finitos com modelos reconstruıdos a partir de mamografias," M.Sc. thesis, Centro de Ciências Exatas e Tecnologia, Universidade Federal do Maranhão, São Luis, MA, Brazil, 2011.

[55] D. B. MacManus, B. Pierrat, J. G. Murphy, and M. D. Gilchrist, "Mechanical characterization of the P56 mouse brain under large-deformation dynamic indentation," *Scientific Reports*, vol. 6, no. 1, p. 21569, 2016.

[56] A. R. Esmaeili, M. Keshavarz, and A. Mojra, "Optimization of hyperelastic model parameters of soft tissue based on genetic algorithm utilizing experimental mechanical dataset," *Journal of Modares Mechanical Engineering*, vol. 15, no. 9, pp. 134–140, 2015.

[57] E. Omidi, L. Fuetterer, S. R. Mousavi, R. C. Armstrong, L. E. Flynn, and A. Samani, "Characterization and assessment of hyperelastic and elastic properties of decellularized human adipose tissues," *Journal of Biomechanics*, vol. 47, no. 15, pp. 3657–3663, 2014.

[58] R. Roy and J. P. Desai, "Determination of mechanical properties of spatially heterogeneous breast tissue specimens using contact mode atomic force microscopy (AFM)," *Annals of Biomedical Engineering*, vol. 42, no. 9, pp. 1806–1822, 2014.

[59] H. Liu, K. Sangpradit, M. Li, P. Dasgupta, K. Althoefer, and L. D. Seneviratne, "Inverse finite-element modeling for tissue parameter identification using a rolling indentation probe," *Medical & Biological Engineering & Computing*, vol. 52, no. 1, pp. 17–28, 2014.

[60] P. Boonvisut and M. C. Çavuşoğlu, "Estimation of soft tissue mechanical parameters from robotic manipulation data," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 5, pp. 1602–1611, 2013.

[61] Z. Wang, Y. Liu, G. Wang, and L. Sun, "Elastography method for reconstruction of nonlinear breast tissue properties," *International Journal of Biomedical Imaging*, vol. 2009, Article ID 406854, 9 pages, 2009.

[62] I. D. Breslavsky, M. Amabili, and M. Legrand, "Nonlinear vibrations of thin hyperelastic plates," *Journal of Sound and Vibration*, vol. 333, no. 19, pp. 4668–4681, 2014.

[63] I. D. Breslavsky, M. Amabili, and M. Legrand, "Physically and geometrically non-linear vibrations of thin rectangular plates," *International Journal of Non-Linear Mechanics*, vol. 58, pp. 30–40, 2014.

[64] H. Mehrabian and A. Samani, "An iterative hyperelastic parameters reconstruction for breast cancer assessment," in *Proceedings of SPIE*, vol. 6916, *Medical Imaging 2008: Physiology, Function, and Structure from Medical Images*, SPIE, Bellingham, WA, USA, 2008.

[65] M. M. Dastjerdi, A. Fallah, and S. Rashidi, "Detection of lesions in soft tissues based on nonlinear constitutive model," in *Proceedings of 22nd IEEE Iranian Conference on Biomedical Engineering (ICBME)*, pp. 363–368, Tehran, Iran, 2015.

[66] A. Aghajani, M. Haghpanahi, and T. Nikazad, "The ultrasound elastography inverse problem and the effective criteria," *Proceedings of the Institution of Mechanical Engineers (IMechE), Part H: Journal of Engineering in Medicine*, vol. 227, no. 11, pp. 1203–1212, 2013.

[67] J. Kim and M. A. Srinivasan, "Characterization of viscoelastic soft tissue properties from in vivo animal experiments and inverse FE parameter estimation," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 599–606, Springer-Verlag, Berlin, Heidelberg, Germany, 2005.

[68] H. Rivaz, E. Boctor, M. Choti, and G. Hager, "Real-time regularized ultrasound elastography," *IEEE Transactions on Medical Imaging*, vol. 30, no. 4, pp. 928–945, 2011.

[69] H. Rivaz, E. Boctor, P. Foroughi, R. Zellars, G. Fichtinger, and G. Hager, "Ultrasound elastography: a dynamic programming approach," *IEEE Transactions on Medical Imaging*, vol. 27, no. 10, pp. 1373–1377, 2008.

Hindawi
Journal of Healthcare Engineering
Volume 2018, Article ID 5940436, 16 pages
https://doi.org/10.1155/2018/5940436

Research Article

Local Binary Patterns Descriptor Based on Sparse Curvelet Coefficients for False-Positive Reduction in Mammograms

Meenakshi M. Pawar,[1] **Sanjay N. Talbar**,[2] **and Akshay Dudhane**[2]

[1]*Department of Electronics and Telecommunication, SVERI's College of Engineering, Pandharpur, Solapur, Maharashtra, India*
[2]*Department of Electronics and Telecommunication Engg., S.G.G.S.I.E. & T, Nanded, Maharashtra, India*

Correspondence should be addressed to Meenakshi M. Pawar; mmpawar@coe.sveri.ac.in

Received 6 April 2018; Revised 18 June 2018; Accepted 8 August 2018; Published 25 September 2018

Academic Editor: Santosh K. Vipparthi

Breast Cancer is the most prevalent cancer among women across the globe. Automatic detection of breast cancer using Computer Aided Diagnosis (CAD) system suffers from false positives (FPs). Thus, reduction of FP is one of the challenging tasks to improve the performance of the diagnosis systems. In the present work, new FP reduction technique has been proposed for breast cancer diagnosis. It is based on appropriate integration of preprocessing, Self-organizing map (SOM) clustering, region of interest (ROI) extraction, and FP reduction. In preprocessing, contrast enhancement of mammograms has been achieved using Local Entropy Maximization algorithm. The unsupervised SOM clusters an image into number of segments to identify the cancerous region and extracts tumor regions (i.e., ROIs). However, it also detects some FPs which affects the efficiency of the algorithm. Therefore, to reduce the FPs, the output of the SOM is given to the FP reduction step which is aimed to classify the extracted ROIs into normal and abnormal class. FP reduction consists of feature mining from the ROIs using proposed local sparse curvelet coefficients followed by classification using artificial neural network (ANN). The performance of proposed algorithm has been validated using the local datasets as TMCH (Tata Memorial Cancer Hospital) and publicly available MIAS (Suckling et al., 1994) and DDSM (Heath et al., 2000) database. The proposed technique results in reduction of FPs from 0.85 to 0.02 FP/image for MIAS, 4.81 to 0.16 FP/image for DDSM, and 2.32 to 0.05 FP/image for TMCH reflecting huge improvement in classification of mammograms.

1. Introduction

Breast cancer is the most common cancer disease among women across worldwide. It is the leading cause of deaths for women suffering from cancer disease in India. It is estimated that breast cancer cases in India would reach to as high as 1,797,900 by 2020 [1]. Rising rate of incidences can cause high mortality. This is due to lack of awareness about breast screening, late reporting, and insufficient medical access [2]. This fact brings a concern and necessity that screening for breast cancer is prudent in its early stage to confirm longer survival. Among all techniques, namely, mammography, tomosynthesis, ultrasonography, computed tomography, and magnetic resonance, mammography is the most reliable and accepted modality by radiologist for preliminary examination of breast cancer due to cost benefits and accessibility [3–5]. The diagnosis of breast cancer using mammogram by radiologist varies from expert to expert as symptoms are misinterpreted or overlooked, due to the tedious task of screening mammograms. Study reveals that 10% to 30% of the visible cancers on mammograms are overlooked, and only 20% to 30% of biopsies are positive [6–8]. Biopsies are traumatic in nature and costly; therefore, computer aided detection and diagnosis (CAD) systems combined with expert radiologists' experience would provide more comprehensive diagnosis [9]. Detailed survey about the research in the design of CAD systems has been given in next section.

2. Literature Survey

The design and development of CAD system is an important progressive area of research for contrast enhancement for better visualization and clarification [10–12], pectoral

muscle removal, segmentation for better delineation of region of interest (ROI), extraction of features, and classification [13, 14]. The segmentation method is classified as region based, contour-based, and clustering method [15]. The region and contour-based methods are popularly used by many researchers. Görgel et al. [16] developed Local Seed Region Growing-Spherical Wavelet Transform (LSRG–SWT) algorithm using local dataset and MIAS [17] with classification accuracy of 94% and 91.67%, respectively. Pereira et al. [18] presented segmentation and detection of masses in mammogram using wavelet transform and genetic algorithm that provides FP rate of 1.35 FP/image and sensitivity of 95% using DDSM [19]. Rouhi et al. [20] studied segmentation using region growing, Cellular Neural Network (CNN), and ANN. The result of classification varied from 80 to 96%, which is the main weakness of their study. Berber et al. [21] proposed Breast Mass Contour Segmentation (BMCS) approach and showed 6 FPR for local dataset. Hybrid level set segmentation method [22] based on combination of region growing and level set was used to segment tumor. The results showed that the sensitivity varied from 78 to 100% due to the presence of artifact in the MIAS database. The difficulties in region and contour-based segmentation methods are the appropriate initialization of seed point and contour position.

Several researchers have implemented clustering method like K-means and Fuzzy C-means (FCM) for breast abnormality segmentation [3, 23]. However, they have limitations in terms of learning abilities. Learning-based techniques such as Self-organizing map (SOM) [24] have been successfully used in medical image segmentation [25]. The success of SOM in medical image segmentation has inspired the researcher to choose it for mammogram segmentation. Many of the times the tumor-segmented regions are not the abnormal tissues (cancerous region), and they are known as false positives (FPs). This FP consumes much time of radiologists and results into unnecessary biopsies. Thus, reducing the FPs is an open research problem and various researchers have proposed FP reduction algorithms to improve the specificity of the CAD systems [5, 9, 23, 26–31]. Usually, FP reduction algorithm is postprocessing step of a CAD system with two stages namely: Feature extraction and Classification. Various methods have been developed for feature extraction based on wavelets [8, 18, 32], curvelet [33, 34], Gabor [35, 36], morphological descriptors [20], textural analysis [26, 27, 30, 32], histogram [4, 5, 7, 29, 37–40], etc. The segmentation error can reduce the performances of morphological descriptor. When Gray Level Co-occurrence Matrix (GLCM) from normal and abnormal region in dense mammogram is same, texture descriptor overlaps that leads to more number of FPs [37]. Ojala et al. proposed local binary patterns (LBPs) [41] for textural feature extraction which works well in feature extraction as compared to morphological descriptor and GLCM-based textural descriptor. LBP descriptor can be considered as local microstructures, namely, edges, flat areas, spots, etc. Variants of LBP have been proposed by various researchers to achieve rotation and intensity invariant features. Also, LBP is computationally efficient and extracts robust features; therefore, LBP descriptors have been widely applied in FP reduction and classification methods for mammogram images [29, 37, 39, 40]. However, LBP descriptor does not provide the directional information of local micropattern. Therefore, transform technique such as curvelet combined with LBP was used to extract features. Various curvelet-based approaches have been proposed in the literature [8, 33, 34, 42] which conclude that curvelet outperforms as compared to wavelet transform.

In this work, novel method of extracting sparse curvelet subband coefficients by incorporating the knowledge of irregular shape of masses as they appear in sparse matrix and calculating LBP features has been presented. Therefore, this paper presents scheme as follows:

(1) Preprocessing of mammogram image for contrast enhancement using local entropy maximization-based image fusion algorithm and removal of background noise

(2) Cluster-based segmentation of mammograms using SOM and extract tumor regions, i.e., ROI)

(3) FP reduction: extraction of sparse curvelet subband coefficients and computation of LBP descriptor to classify true positives and false positives to improve performance of CAD system using MIAS [17], DDSM [19], and Tata Memorial Cancer Hospital (TMCH) datasets.

The organization of paper is as follows: Sections 1 and 2 illustrate the introduction and literature review on automatic segmentation and extraction of abnormal masses (i.e., tumor region) as well as FP reduction methods. Section 3 presents the proposed methodology for SOM based segmentation of mammograms followed by novel false positive reduction in detail. Section 4 depicts the experimental results and discussions on three benchmark datasets. Finally, Section 5 concludes the proposed approach for accurate extraction of abnormal masses (i.e., tumor region) by excluding the FPs.

3. Methodology

The block schematic of proposed integrated method for automatic detection of breast cancer using sparse curvelet coefficient-based LBP descriptor has been shown in Figure 1.

3.1. Preprocessing. The mammogram images are low-dose x-ray images so they have poor contrast and suffer from noises. The preprocessed mammogram image as shown in Figures 2(a)–2(d) represents preprocessing of mammogram, and Figures 2(e)–2(g) represents SOM clustering and ROI extraction.

3.1.1. Local Entropy Maximization-Based Image Fusion: Contrast Enhancement. The contrast enhancement of the mammogram is performed using local entropy maximization [12] for better segmentation. Here, original image is given to the contrast limited adaptive histogram equalization (CLAHE) algorithm to get the second input to our image

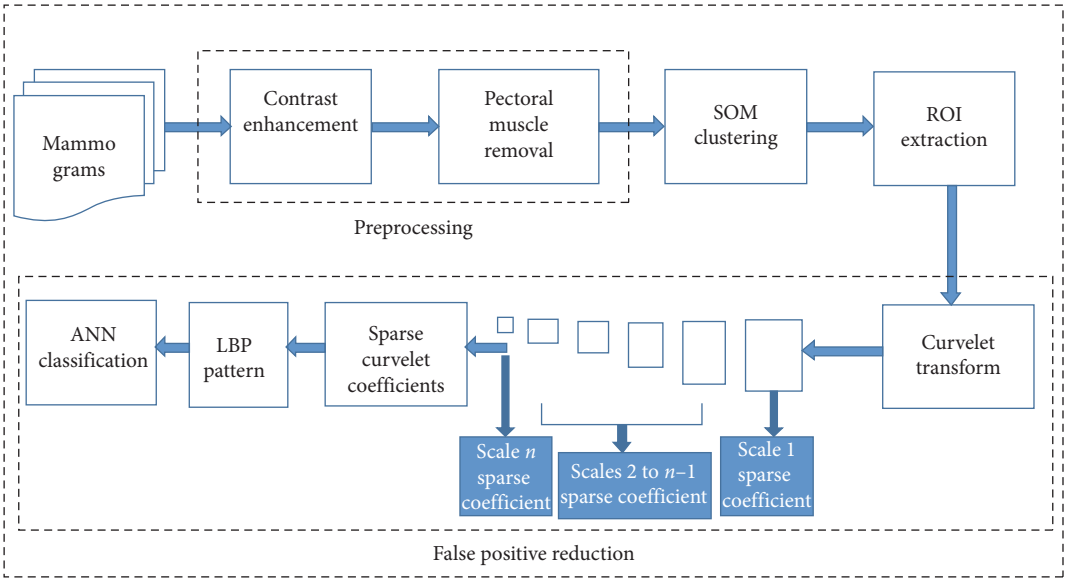


Figure 1: Schematic architecture for automatic breast cancer detection.

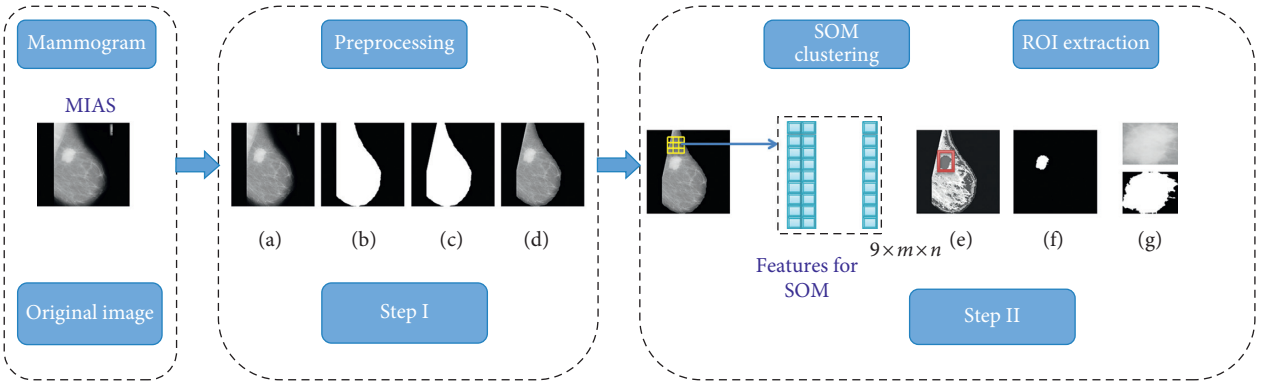


Figure 2: Steps for mammogram processing (a) enhanced mammogram, (b) binary mask, (c) pectoral removal, (d) pectoral removed mammogram, (e) clustered image, (f) cluster of interest, and (g) ROI extraction.

fusion algorithm. Further, original image along with the CLAHE has been given to the image fusion algorithm. Procedure of the image fusion has been given in Algorithm 1. We have used local entropy as a fusion rule given by the following equation:

$$\mathrm{ENT} = -\sum_{k=0}^{255} p(k)\log(p(k)), \tag{1}$$

where ENT is the local entropy and $p\_\mathrm{org}(k)$ and $p\_\mathrm{CLAHE}(k)$ are the probability of k^{th} pixel from 5×5 sliding window [12]. Here, both high frequency components from original mammogram and CLAHE mammogram have been fused using maximum entropy criteria. Figure 3(b) presents contrast-enhanced mammogram using local entropy maximization-based image fusion.

3.1.2. Pectoral Muscle Removal. Pectoral muscle suppression has been performed by defining rectangle as suggested in [14] (Figure 3(c)). It illustrates the rectangle (ABDC) and fixes the points G and has intensity variation and joins them for pectoral muscle suppression. Figure 3(d) illustrates pectoral muscle removed image to avoid discrepancies in the algorithm because of similar intensities present between pectoral muscle and masses.

3.2. SOM Clustering. SOM is a special type of neural network designed to map the input image of size $N_x \times N_y$ to M clusters based on their characteristic features [25]. For SOM, the image (I) is converted into a feature vector $f = \{f_1, f_2, \ldots, f_m\}$, where m is the number of features. In this experiment, we have trained SOM with $M = 4$ clusters using $p = 9$ neighbourhood features such as given a centre pixel (g_c) in the image, the neighbourhood features are computed as given in the following equation:

$$F(1, p) = g_p, \quad p \in [1, n], \tag{2}$$

where n is the number of neighbourhood (3×3 window), g_p is the neighbourhoods, and F is the feature vector

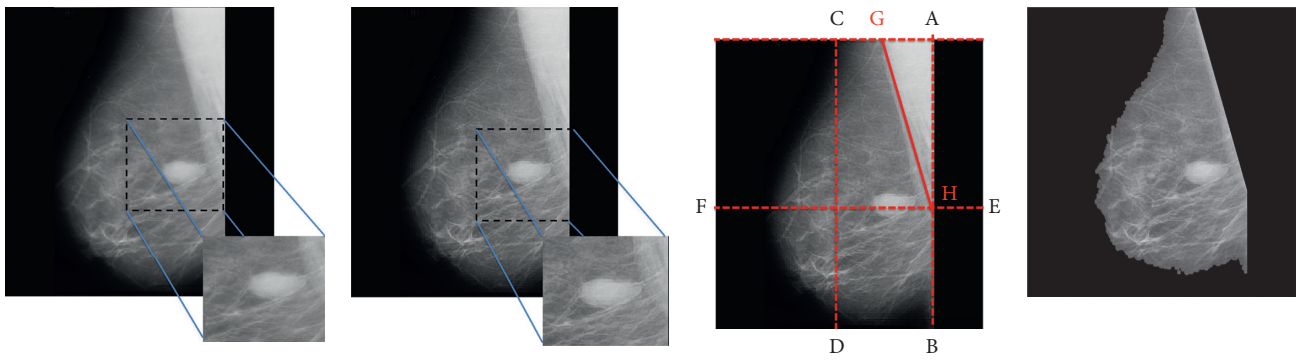


Figure 3: Preprocessing. (a) Original image from MIAS database. (b) Contrast-enhanced mammogram using local entropy maximization. (c) Process of pectoral muscle removal. (d) Pectoral muscle removed mammogram.

corresponds to centre pixel g_c. The selection of 3×3 window pixel is based on [43] to capture local details.

At the start, weight vector $W_i = \{w_{i1}, w_{i2}, \ldots, w_{im-1}\}$ is random and updated as the network learns. The minimum Euclidean distance $\|f - W_i\|$ is described as the best matching component or winner node $\|f - W_c\|$ and described as

$$\|f - W_c\| = \min_i\{\|f - W_i\|\}. \tag{3}$$

Weight vector for winning output neuron and its neighboring neurons are updated as

$$W_i(t+1) = W_i(t) + N_{ci}(t)\big(f(t) - W_i(t)\big), \tag{4}$$

where $t = 1, 2, \ldots$ is time coordinate. The function $N_{ci}(t)$ is the neighbourhood kernel function and expressed as

$$N_{ci}(t) = \eta(t)\exp\left(-\frac{\|m_c - m_i^2\|}{2\sigma^2(t)}\right), \tag{5}$$

where $\eta(t)$ is the learning rate, $\sigma(t)$ is a width of kernel that corresponds to neighbourhood neurons around node c and m_c and m_i corresponds to location vectors of nodes c and i.

Figures 4(a) and 4(b) represent cluster map and cluster boundaries marked on mammogram. After the several observations for known areas, it was empirically noticed that number of pixels of range or pixel level threshold (PLT based on pixel count in TP) as 450 to 31,500; 16,000 to 2,00,000; and 4,000 to 2,00,000 consist of abnormality for MIAS, DDSM, and TMCH database, respectively, which is verified from the expert. The size of the tumor is varying because of the mammogram size of 1024×1024 pixels for MIAS, 2728×3920 pixels to 4608×6048 pixels for DDSM, and 2294×1914 or 4096×3328 pixels for TMCH datasets. Therefore, cluster regions below or above the specified threshold are discarded and the remaining region is marked as true positive (TP) as shown in Figure 4. Figure 4(a) shows the clustered image using SOM; Figure 4(b) shows the cluster boundaries marked on original image.

We can see that there are many FPs along with TP (marked by pink color) which are reduced using pixel level threshold (PLT based on pixel count in TP) as explained above. Figure 4(c) shows the filtered result using PLT.

3.3. ROI Extraction. After SOM clustering (initial segmentation), the next step is to classify the detected regions into TP and FP by using proposed local sparse curvelet features (LSCF) followed by ANN classifier. To do so, initially, we have extracted ROIs from detected regions by SOM clustering and manually categorized into TP and FP. We collected these ROIs from three different datasets according to their maximum height and maximum width using connected components e.g., region marked in Figure 4(c). Therefore, their patch size is different as shown in Figure 5, ROIs for MIAS, DDSM, and TMCH dataset. Further, these extracted patches have been used to train the ANN for the task of FP reduction.

3.4. False-Positive (FP) Reduction. After ROI extraction, FP reduction algorithm performs computation of proposed local sparse curvelet features (LSCF) followed by ANN classifier.

3.4.1. Proposed Algorithm. LBP [43] was proposed as LBP descriptor computation at circular neighbourhood which is called as uniform LBP (ULBP) descriptor and expressed as

$$\mathrm{ULBP}_{(P,R)} = \begin{cases} \sum_{n=1}^{P-1} S & \text{if } U\big(LBP_{(P,R)}\big) \le 2, \\ P+1 & \text{otherwise,} \end{cases} \tag{6}$$

where

$$U\big(\mathrm{LBP}_{(P,R)}\big) = \big|S(g_{P-1} - g_C) - S(g_0 - g_C)\big| + \sum_{n=1}^{P-1}\big|S(g_P - g_C) - S(g_{P-1} - g_C)\big|. \tag{7}$$

Computation of LBP based on actual shape of mass according to sparse matrix has been shown in Figure 6, where

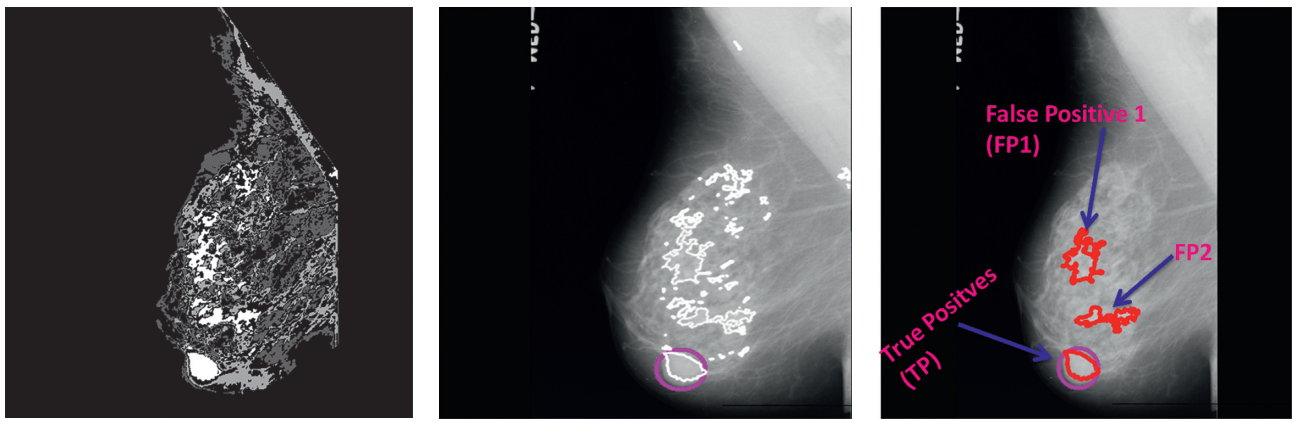


Figure 4: FP reduction by thresholding (a) clustered image, (b) clusters boundaries marked on original image, and (c) clusters after thresholding.

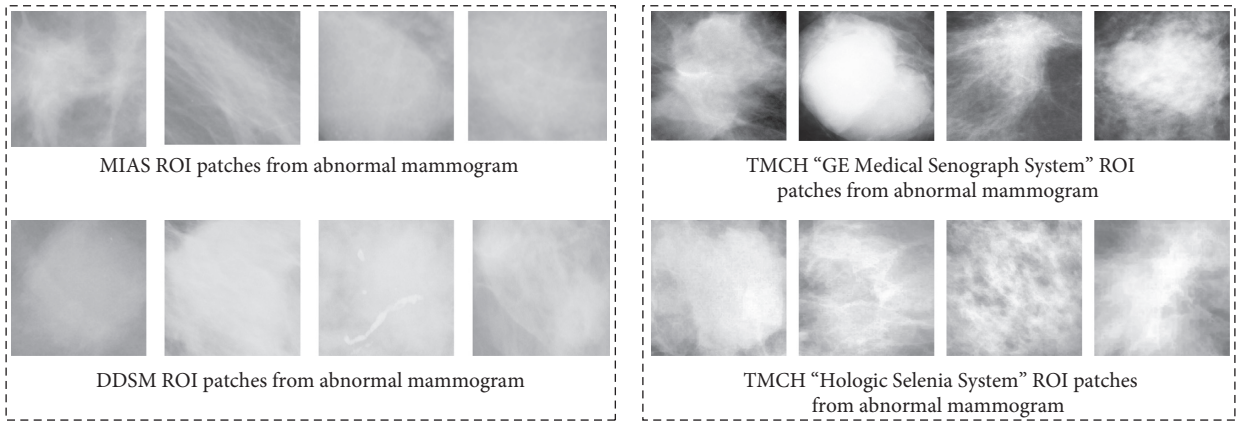


Figure 5: Variable sizes ROIs from MIAS, DDSM, and TMCH datasets.

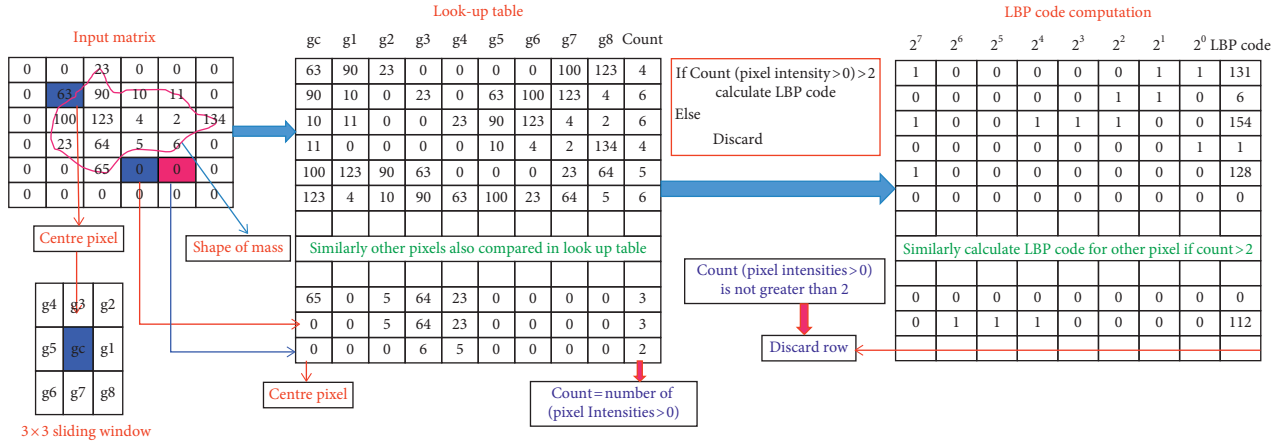


Figure 6: Lookup table approach for LBP computation from shape of mass in ROI.

it takes pixels related to shape of mass which are called as foreground pixels and rejects the other pixels called as background pixels. The proposed algorithm uses foreground pixels only for LBP computation, and this will tend to number of pixel reduction in LBP computations. Therefore, identification of foreground and background pixels is an important step which is performed using lookup table approach. The identification of foreground and background pixel is based on number of nonzero pixels in the lookup table, i.e., if count of sliding window nonzero pixels is greater than 2, count($p(i, j)$) > 2 is identified as foreground and LBP is estimated. On the other hand, if count of sliding window nonzero pixels is less than 2, count($p(i, j)$) < 2 is identified as background and LBP would not be estimated and rejected from lookup table.

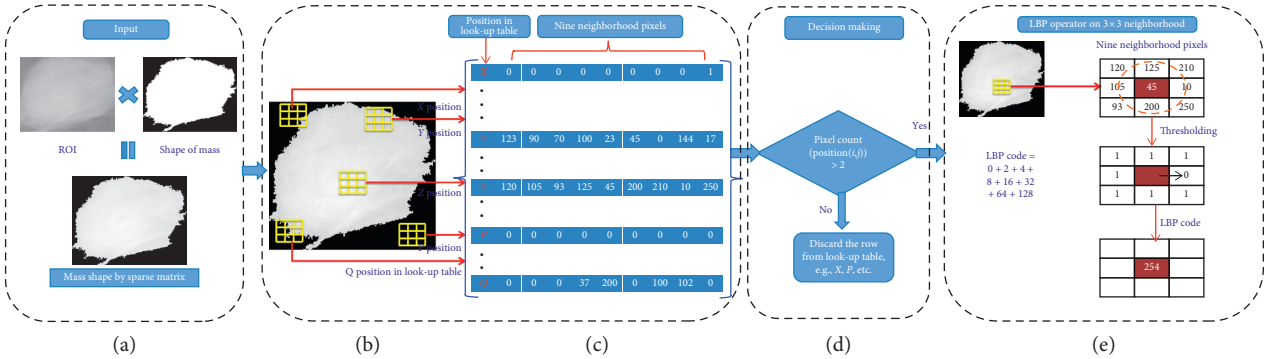


FIGURE 7: Process for computation of LBP descriptor from shape of mass in ROI. (a) Original image, (b) 3×3 window for selection of foreground pixels, (c) lookup table, (d) decision making process, (e) LBP computation from selected foreground pixels.

Nonzero pixels provide actual shape of mass and are taken for LBP computations. Graphical representation of proposed algorithm for LBP descriptor computation using foreground pixels has been given in Figure 7 and the algorithm has been described in Algorithm 2.

3.4.2. The Fast Discrete Curvelet Transform (FDCT). The authors [44] have introduced computationally simple and efficient Fast Discrete Curvelet Transform (FDCT). We have preferred wrapping-based FDCT approach in proposed work, as it is faster. The curvelet coefficients $C^{\mathrm{D}}(j,l,k)$ represented by scale j, angle l, and spatial location k can be written as

$$C^{\mathrm{D}}(j,l,k_1,k_2) = \sum_{n_1=1}^{n_1=N_1} \sum_{n_2=1}^{n_2=N_2} I[n_1,n_2]\varphi^{\mathrm{D}}_{j,l,k_1,k_2}[n_1,n_2]. \quad (8)$$

Figure 8 illustrates LBP code computation based on sparse curvelet coefficients; ROI decomposes using curvelet transform with scale orientations l of 16° and scale of 2 as the database consists of minimum ROI size of 25×22 pixels. Curvelet transform with scale orientations l of 16° and scale of 2 produces $1 + 16 = 17$ different subbands based on subband division. Further, each curvelet subband coefficients have been represented using lookup table using 3×3 sliding window, and if the row in the lookup table identifies foreground coefficient, then LBP is computed with radius $R = 1$ and $P = 8$ neighboring pixels as shown in Algorithm 2; total 58 LBP features have been obtained from foreground curvelet subband coefficients. Therefore, total 986 LBP features have been extracted from 17 curvelet subbands. It can be observed from Figure 8, curvelet subbands also provide shape of mass in 16 different directions so that the directional information can be associated with LBP features. Kanadam et al. [3] used concept of sparse ROI; similarly, we have extended it for sparse curvelet subband and LBP features computation.

3.5. Classification. In this work, we have analyzed extracted ROI from mammogram using normal-abnormal, benign-malignant, and normal-malignant classes with ANN, SVM, and KNN classifiers. The detailed description of ANN classifier has been given in [45, 46]. To evaluate performance of the proposed system, we have used 3-fold cross validation where database is randomly divided into three sets and accuracy is calculated for each set. The final accuracy of the system is average of accuracy of each of three sets. However, it will not be fair to compare 3-fold cross validation result of SVM and KNN classifier with ANN, because ANN classifier is tested on only one set of images (33% for training, 33% for testing, and 33% for validation). Thus, to do fair comparison, we have trained ANN using input layer (986 neuron) over three different sets (which are considered in SVM and KNN) and calculated its average accuracy. Our proposed false positive reduction algorithm illustrates in Figures 9(a)–9(c). Algorithm 3 summarizes flow of the proposed method for FP reduction in mammograms.

4. Experimental Results and Discussions

The proposed method has been tested and validated using three classifiers and three clinical mammographic image datasets.

4.1. Data Sets

4.1.1. Mammographic Image Analysis Society (MIAS) Database. The mini-MIAS [17] database consists of 322 mammograms, each having 1024×1024 pixels and annotated like background tissue character, class, severity, center of abnormality, and radius of circle for abnormality. This database includes 64 benign, 51 malignant, and 207 normal cases, which have been taken for experimentation.

4.1.2. Digital Database for Screening Mammography (DDSM). The DDSM [19] dataset consists of 2500 studies and is composed of cranial-caudal (CC) and mediolateral-oblique (MLO) views of mammographic image for left and right breast, annotated with ACR breast density, type of abnormality, and ground truth. Randomly selected 150 abnormal and 100 normal cases from both HOWTEK and LUMISYS scanner of 12 bits per pixel resolution have been subjected for experimentation.

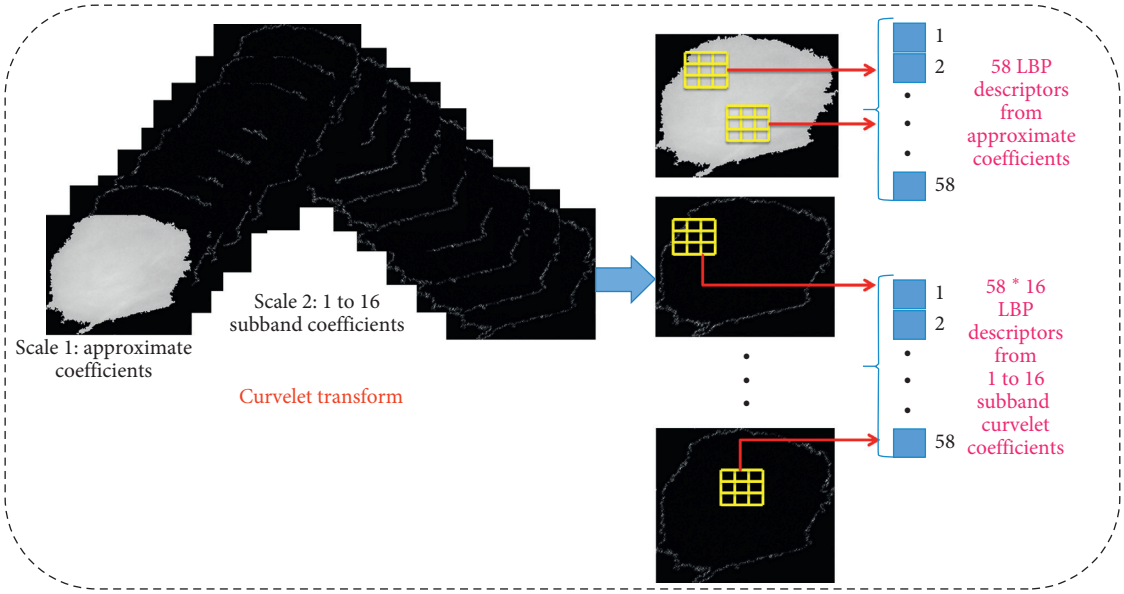
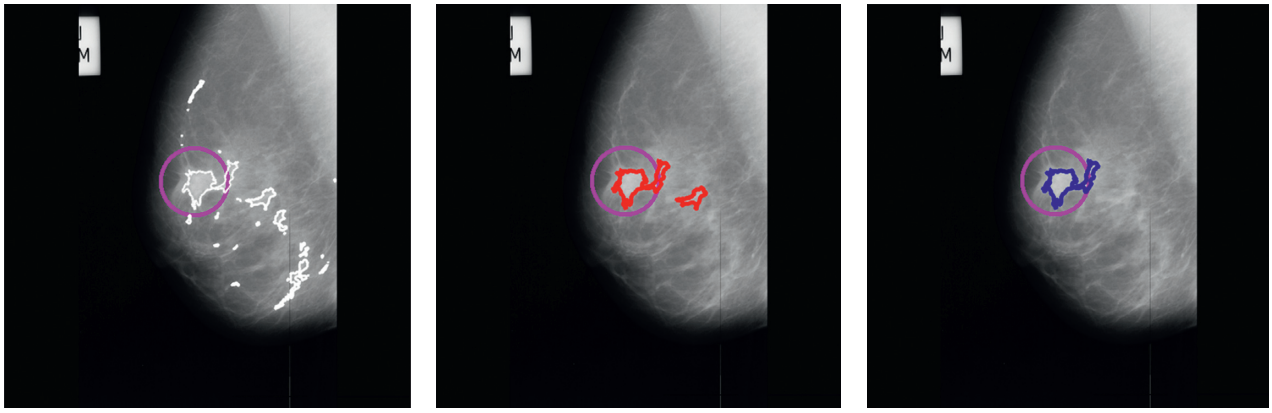


Figure 8: LBP code computation using sparse curvelet subband coefficients.



(a) (b) (c)

Figure 9: (a) FP reduction by clusters marked on original image, (b) FP reduction by thresholding, (c) FP reduction by sparse curvelet coefficient-based LBP, and ANN.

4.1.3. The Tata Memorial Cancer Hospital (TMCH). This dataset [47] contains 360 full-field digital mammograms (FFDMs) comprising 180 CC views and 180 MLO views from right and left breast acquired from 90 randomly selected patients. It is composed of 180 verified malignant and 180 normal breast images. It uses biopsy proven breast cancer patients' pathological data approved by the Institutional Research Ethics Committee of Tata Memorial Centre Hospital (TMCH), Mumbai, India. The ground truth marking on each abnormal mammogram is performed manually using the Histopathological Reports (HPR) of the respective patients and expert radiologist from TMCH, Mumbai. Approximately 35 patients are examined using "Hologic Selenia System" (Scanner1) gives 16-bit.

The remaining 55 patients were examined with "GE Medical Senograph System" (Scanner2) providing 8-bit true color mammogram image in DICOM format of 4096×3328 or 2294×1914 pixels each measuring size $50 \times 50\ \mu m^2$.

4.2. Segmentation Evaluation and ROI Extraction. The segmentation using SOM that detects suspicious mass regions is considered as TP whereas from nonmass is taken as FP. From Table 1, it is clear that total suspicious ROI (including TP & FP) of 381 for MIAS, 1343 for DDSM, and 1009 for TMCH have been taken for evaluation our proposed algorithm for FP reduction.

From extracted ROIs, the minimum patch size is 25×22 pixels whereas the maximum size is 1152×1356 pixels. Tables 2 and 3 represent curvelet subband coefficients from 17 subbands, and reduced coefficients based on lookup table approach are used to calculate LBP features. It has been observed during experimentation that the curvelet coefficients on an average are reduced for sparse LBP by 14%, 32%, 33%, and 34% for MIAS, DDSM, TMCH: Scanner1, and TMCH: Scanner2, respectively. It may be noticed that reduction in curvelet coefficients for every ROI is not fixed. It completely

(1) Load input image (img1)
(2) Apply CLAHE algorithm and obtain enhanced image (img2)
(3) Decompose img1 and img2 up to 3 level of decomposition using Discrete Wavelet transform (DWT)
(4) Use maximum local entropy rule for fusion of img1 and img2 for high frequency subbands
(5) Take inverse DWT to obtain the fused image

Algorithm 1: Image fusion for contrast enhancement.

```
Input: I(m, n); m = no. of rows and n = no. of column
Output: LBP features
Initialize: Radius R = 1 and neighborhood pixels P = 8
        Mask = [1 2 4 8 16 32 64 128 0]
        Sliding window coordinates: k = −1 : 1
        Count = 1 //number of pixels in I(m, n)
for i = 1 to m do
        for j = 1 to n do
                //prepare local circular window
                I_local = I(i + k, j + k)
                center_pixel = I(i, j)
                //Arrange local neighborhoods of I(i, j) pixels in a row col = 1 : 9
                Lookup_table (i, col) = reshape(I_local [7, 17])
                //count number of pixels greater than zero
                a = length(find(Lookup_table > 0))
                //select pixel position from lookup-table for computation of LBP
                if a > 2

                  LBP_code(count,:) = I_local > center_pixel
                  count = count + 1

                end
        end
end
//compute histogram of LBP codes
LBP_descriptor = LBP_descriptor/count
//scale invariant
```

Algorithm 2: Algorithm for LBP feature computation based on shape of mass in ROI as.

(1) Load input image (img1)
(2) Apply CLAHE algorithm and obtain enhanced image (img2)
(3) Process img1 and img2 and obtain enhanced image using procedure given in Algorithm 1
(4) Remove pectoral muscle using proposed approach (Section 3.1.2)
(5) Extract neighbourhood features for each pixel and apply SOM clustering
(6) Obtain clustered image and separate out the tumorous cluster
(7) Extract detected regions *i.e.,* ROI's from clustered result
(8) Extract Sparse Curvelet Coefficients (Subband) up to 2 level from each ROI
(9) Extract Sparse LBP code for each subband and obtain a combined feature vector for each ROI
(10) Classify each ROI into tumorous and nontumorous class *i.e.,* TP and FP respectively
(11) Map each TP region on original mammogram (img1)
(12) end

Algorithm 3: Summary of proposed method for FP reduction in mammograms.

depends upon the shape of the ROI as per the sparse matrix. Tables 2 and 3 do not represent exact reduction in pixels for complete database, but they exhibit pixel reduction for sample mammograms.

4.3. Classifier Evaluation and False-Positive Reduction. From Figures 10–13, the best classification accuracy of 98.57 % has been obtained for MIAS in benign versus malignant classification, whereas 98.70% for DDSM, 98.30% for

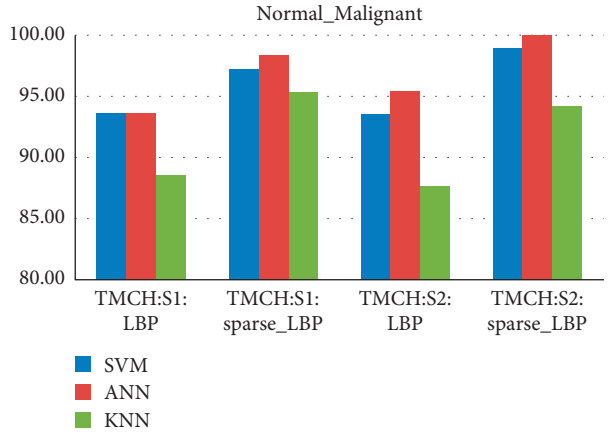


Figure 10: Average classification rate for TMCH dataset.

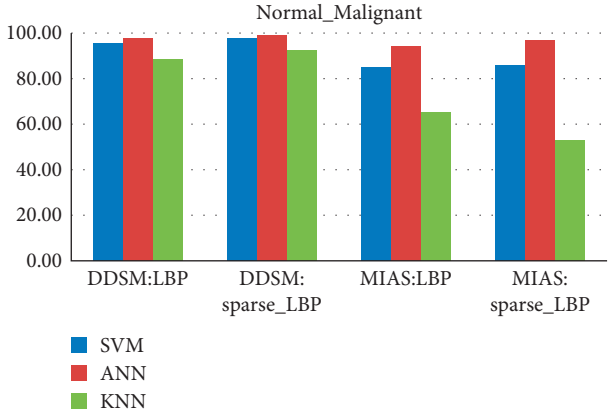


Figure 11: Average classification rate for MIAS and DDSM dataset.

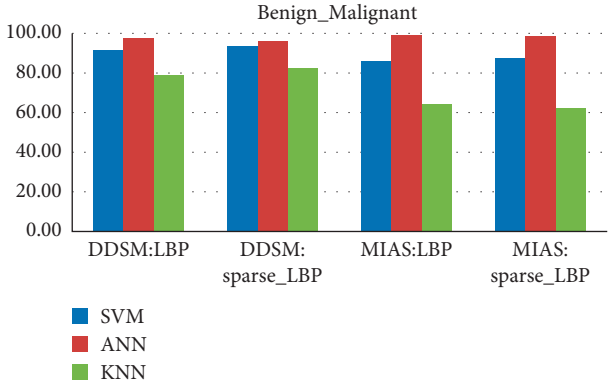


Figure 12: Average classification rate for MIAS and DDSM dataset.

TMCH: Scanner1, and 100% for TMCH: Scanner2 classification accuracies have been obtained in normal versus malignant classification. The classification performance of ANN has improved from 6% to 43% for different databases as compared to KNN classifier, whereas there is little improvement about 7% compared with SVM classifier. The performances of both proposed sparse LBP and LBP computation on curvelet subbands are nearly same; therefore, the proposed algorithm can be efficiently implemented in CAD system with lesser number of curvelet coefficients.

Data augmentation has been used for some classes to maintain balance between two classes, to improve performance, and to learn more powerful model. Table 4 explains the FP reduction with the use of curvelet-based LBP features and ANN. It has been observed that FP reduced from 0.85 to 0.02 FP/image in MIAS, 4.81 to 0.02 FP/image in DDSM and 2.32 to 0.13 FP/image in TMCH.

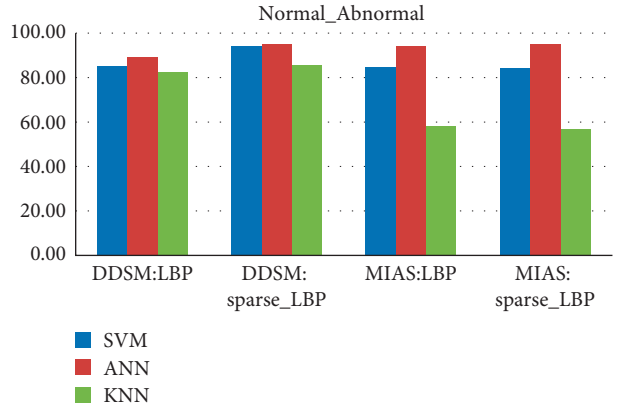


Figure 13: Average classification rate for MIAS and DDSM dataset.

Table 1: Result of SOM segmentation.

Dataset used	Result of SOM clustering and threshold			Total (#) images	TPR (true-positive rate) = TP/#lesions	FPPI (false-positive per image) = FP/#images
	Mass Segmented (TP)	Lost	Segmented nonmass (FP)			
MIAS	108	7	273	322	(108/115) = 0.94	(273/322) = 0.85
DDSM	140	10	1203	250	(140/150) = 0.93	(1203/250) = 4.81
TMCH	172	8	837	360	(172/180) = 0.95	(837/360) = 2.32

Table 2: Reduction in curvelet coefficients for sample mammograms from MIAS and DDSM dataset.

Sr. No.	MIAS				DDSM			
	ROI Size	Total number of curvelet coefficients from subbands	Total number of selected curvelet coefficients from subbands	% reduction in curvelet coefficients	ROI Size	Total number of curvelet coefficients from subbands	Total number of selected curvelet coefficients from subbands	% reduction in curvelet coefficients
1	124 × 138	1,03,911	77,133	25.77	192 × 187	2,16,729	1,68,333	22.33
2	179 × 138	1,50,123	1,26,142	15.97	294 × 291	5,18,267	3,66,680	29.25
3	51 × 116	36,815	33,421	9.22	145 × 207	1,81,663	1,48,765	18.11
4	83 × 83	42,449	36,653	13.65	169 × 168	1,71,873	1,54,752	9.96
5	84 × 76	39,115	35,815	8.44	182 × 248	2,72,517	2,19,822	19.34
6	74 × 83	37,767	34,448	8.79	213 × 349	4,49,783	3,01,359	33.00
7	53 × 64	20,969	18,621	11.20	578 × 412	14,33,195	6,62,072	53.80
8	70 × 44	18,899	16,610	12.11	215 × 219	2,86,429	2,42,935	15.18
9	80 × 66	32,409	29,454	9.12	420 × 428	10,82,461	5,01,209	53.70
10	69 × 86	36,783	33,552	8.78	203 × 307	3,75,871	2,66,829	29.01
11	59 × 116	42,019	38,442	8.51	226 × 262	3,57,209	2,76,763	22.52
12	81 × 101	50,637	46,122	8.92	159 × 194	1,87,563	1,48,741	20.70
13	41 × 84	21,427	18,907	11.76	718 × 686	29,61,127	7,62,561	74.25
14	69 × 141	60,641	52,427	13.54	409 × 550	13,52,439	9,04,829	33.10
15	60 × 62	22,925	20,475	10.69	524 × 375	11,84,671	7,66,522	35.30
16	96 × 101	59,647	54,235	9.07	311 × 275	5,17,433	3,97,737	23.13
17	136 × 139	1,13,727	66,352	41.56	319 × 320	6,14,129	4,12,606	32.81
18	55 × 94	31,359	28,305	9.74	313 × 447	8,42,855	5,10,482	39.43
19	157 × 140	1,32,373	1,07,000	19.17	291 × 517	9,07,903	6,37,412	29.79
20	156 × 130	1,23,007	90,834	26.15	370 × 837	18,64,889	8,98,236	51.83
Average		58,850	48,247	14	Average	7,88,950	4,37,432	32

Similarly, Table 5 shows the reduction in FPs as 0.85 to 0.01 FP/image for MIAS, 4.81 to 0.03 FP/image for DDSM, and 2.32 to 0.00 FP/image for TMCH using sparse curvelet coefficient-based LBP features. The results show the effectiveness of sparse curvelet coefficient-based LBP and ANN. From Table 6, the best value of AUC = 0.99 is obtained in benign versus malignant classification for MIAS, AUC = 0.98 in benign versus malignant in case of

Table 3: Reduction in curvelet coefficients for sample mammograms from TMCH Scanner1 and Scanner2 dataset.

	TMCH: Scanner 1: "GE Medical Senograph System"				TMCH: Scanner 2: "Hologic Selenia System"			
Sr. no.	ROI size	Total number of curvelet coefficients from subbands	Total number of selected curvelet coefficients from subbands	% reduction in curvelet coefficients	ROI size	Total number of curvelet coefficients from subbands	Total number of selected curvelet coefficients from subbands	% reduction in curvelet coefficients
1	459 × 412	11,40,617	7,94,885	30.31	291 × 278	4,89,243	3,17,014	35.20
2	548 × 513	16,93,403	11,17,873	33.99	545 × 246	8,09,627	5,31,604	34.34
3	415 × 303	7,58,323	4,45,585	41.24	560 × 483	16,30,583	10,68,439	34.47
4	645 × 495	19,51,443	11,45,580	41.29	782 × 510	24,00,137	12,75,073	46.87
5	437 × 691	18,16,651	9,85,120	45.77	87 × 141	75,773	67,871	10.43
6	812 × 500	24,39,937	12,87,065	47.25	311 × 185	3,48,565	2,40,546	30.99
7	468 × 379	10,66,333	7,10,242	33.40	262 × 348	5,50,303	2,82,821	48.61
8	673 × 582	23,55,589	17,30,915	26.52	610 × 440	16,14,515	8,42,876	47.79
9	250 × 201	3,04,513	2,35,670	22.61	949 × 391	22,27,209	13,59,338	38.97
10	525 × 488	15,44,691	11,61,942	24.78	365 × 385	8,46,523	6,04,473	28.59
11	488 × 779	22,87,547	16,11,385	29.56	393 × 247	5,85,063	4,90,474	16.17
12	1434 × 966	83,26,581	37,42,277	55.06	341 × 301	6,18,111	4,10,542	33.58
13	348 × 421	8,81,701	6,16,501	30.08	523 × 702	22,06,097	8,00,057	63.73
14	460 × 530	14,67,227	7,99,885	45.48	370 × 284	6,32,539	4,23,727	33.01
15	398 × 450	10,78,441	8,28,064	23.22	344 × 202	4,18,955	3,02,427	27.81
16	247 × 272	4,04,401	3,44,822	14.73	264 × 188	2,99,997	2,31,926	22.69
17	411 × 305	7,57,657	4,05,919	46.42	233 × 247	3,46,983	2,67,686	22.85
18	286 × 344	5,93,155	4,48,566	24.38	370 × 291	6,50,701	4,86,125	25.29
19	417 × 207	5,23,477	4,29,755	17.90	680 × 483	19,79,543	10,52,793	46.82
20	463 × 458	12,75,021	8,57,608	32.74	202 × 266	3,24,295	2,15,042	33.69
Average		16,33,335	9,84,983	33	Average	9,52,738	5,63,543	34

Table 4: Number of ROIs resulted in FP reduction using curvelet-based LBP (without sparse) & ANN classification at training and validation stage.

Class	Dataset used	Benign/malignant mass			Nonmass/benign mass			Total (#) images	TPR (true-positive rate) = TP/#lesions	FPPI (false-positive per image) = FP/# images
		Previous stage	Selected (TP)	Lost (FN)	Previous stage	Selected (TN)	Lost (FP)			
Normal vs abnormal	MIAS	108 ∗ 2 = 216	203	13	273	257	16	315	(203/216) = 0.94	(16/315) = 0.05
	DDSM	140 ∗ 4 = 560	465	95	1203	1095	108	240	(465/560) = 0.83	(108/240) = 0.45
Benign vs malignant	MIAS	49	49	0	59	57	2	108	(49/49) = 1.00	(2/108) = 0.02
	DDSM	46 ∗ 2 = 92	91	1	94	91	3	140	(91/92) = 0.99	(3/140) = 0.02
Normal vs malignant	MIAS	49 ∗ 4 = 196	184	12	273	254	19	256	(184/196) = 0.94	(19/256) = 0.07
	DDSM	46 ∗ 4 = 184	180	4	1203	1143	60	146	(180/184) = 0.98	(60/146) = 0.41
	TMCH: Scanner1	107 ∗ 4 = 428	416	12	605	551	54	217	(416/428) = 0.97	(54/217) = 0.25
	TMCH: Scanner2	65 ∗ 4 = 260	255	5	232	214	18	135	(255/260) = 0.98	(18/135) = 0.13

*Augmentation of image.

DDSM, AUC = 0.94 in normal versus malignant in case of TMCH: Scanner1, and AUC = 0.96 in normal versus malignant classification in TMCH: Scanner2 using ANN and curvelet subband-based LBP features. The worst performance of AUC = 0.53 for MIAS is obtained with the proposed algorithm using KNN classifier as shown in Table 7. Similarly, from Table 7, the best value of AUC = 0.98 is obtained in TMCH: Scanner1, AUC = 1 is obtained in TMCH: Scanner2 database for normal versus malignant classification, AUC = 0.98 in benign versus malignant classification is attained in MIAS database, and AUC = 0.98 is achieved for normal versus malignant classification in DDSM database using ANN classifier for sparse curvelet subband-based LBP features.

However, from Table 7, it should be noted that the performance of proposed algorithm is the best using ANN classifier. Figure 14 represents automated CAD system for breast cancer diagnosis with sample mammograms.

Table 8 provides comparative study of methods developed for breast tissue classification. The proposed method provides best results in terms of AUC and reduction of number of FPs as 0.85 to 0.01 FP/image for MIAS, 4.81 to 0.03 FP/image for DDSM, and 2.32 to 0.00 FP/image for TMCH. The earlier reported work uses the fixed patch size-based approach which

Table 5: Number of ROIs resulted in FP reduction using sparse curvelet coefficient-based LBP & ANN classification at training and validation stage.

Class	Dataset used	Benign/malignant mass			Nonmass/benign mass			Total (#) images	TPR (true-positive rate) = TP/#lesions	FPPI (false-positive per image) = FP/# images
		Previous stage	Selected (TP)	Lost (FN)	Previous stage	Selected (TN)	Lost (FP)			
Normal vs abnormal	MIAS	108 ∗ 2 = 216	201	15	273	265	8	315	(201/216) = 0.93	(8/315) = 0.02
	DDSM	140 ∗ 4 = 560	516	44	1203	1155	48	240	(516/560) = 0.92	(48/240) = 0.2
Benign vs malignant	MIAS	49	48	1	59	59	1	108	(48/49) = 0.98	(1/108) = 0.01
	DDSM	46 ∗ 2 = 92	89	3	94	89	5	140	(89/92) = 0.97	(5/140) = 0.03
	MIAS	49 ∗ 4 = 196	192	4	273	259	14	256	(192/196) = 0.98	(14/256) = 0.05
	DDSM	46 ∗ 4 = 184	182	2	1203	1167	36	146	(182/184) = 0.99	(36/146) = 0.25
Normal vs malignant	TMCH: Scanner1	107 ∗ 4 = 428	424	4	605	593	12	217	(424/428) = 0.99	(12/217) = 0.05
	TMCH: Scanner2	65 ∗ 4 = 260	260	0	232	232	0	135	(260/260) = 1.00	(0/135) = 0

*Augmentation of image.

Table 6: Performance evaluation of curvelet-based LBP descriptor algorithm.

Dataset	Classification Classifier	Normal-malignant			Normal-abnormal			Benign-malignant		
		Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
MIAS	ANN	0.94	0.93	0.94	0.94	0.94	0.94	1.00	0.97	0.99
	SVM	0.85	0.85	0.85	0.83	0.86	0.85	0.88	0.84	0.86
	KNN	0.67	0.63	0.65	0.58	0.57	0.58	0.62	0.68	0.63
DDSM	ANN	0.98	0.95	0.95	0.83	0.91	0.85	0.99	0.97	0.98
	SVM	0.97	0.88	0.92	0.71	0.91	0.83	0.94	0.89	0.92
	KNN	0.96	0.64	0.87	0.67	0.90	0.80	0.87	0.73	0.79
TMCH: Scanner1	ANN	0.97	0.91	0.94	—	—	—	—	—	—
	SVM	0.96	0.91	0.94	—	—	—	—	—	—
	KNN	0.98	0.82	0.89	—	—	—	—	—	—
TMCH: Scanner2	ANN	0.98	0.92	0.96	—	—	—	—	—	—
	SVM	0.97	0.90	0.94	—	—	—	—	—	—
	KNN	0.92	0.83	0.88	—	—	—	—	—	—

Table 7: Performance evaluation of proposed algorithm.

Dataset	Classification Classifier	Normal-malignant			Normal-abnormal			Benign-malignant		
		Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
MIAS	ANN	0.98	0.95	0.96	0.93	0.97	0.95	0.97	1.00	0.98
	SVM	0.88	0.83	0.85	0.85	0.82	0.84	0.84	0.92	0.87
	KNN	0.55	0.51	0.53	0.55	0.63	0.56	0.61	0.67	0.61
DDSM	ANN	0.99	0.97	0.98	0.92	0.96	0.93	0.97	0.95	0.96
	SVM	0.99	0.92	0.96	0.89	0.96	0.92	0.94	0.92	0.93
	KNN	0.98	0.73	0.92	0.74	0.90	0.82	0.89	0.77	0.83
TMCH: Scanner1	ANN	0.99	0.98	0.98	—	—	—	—	—	—
	SVM	0.98	0.96	0.97	—	—	—	—	—	—
	KNN	0.99	0.92	0.95	—	—	—	—	—	—
TMCH: Scanner2	ANN	1.00	1.00	1.00	—	—	—	—	—	—
	SVM	1.00	0.98	0.99	—	—	—	—	—	—
	KNN	0.96	0.92	0.94	—	—	—	—	—	—

limits the automatic CAD system scope whereas proposed system provides complete solution to CAD system right from automatic tumor patch segmentation to reduction in FPs and final representation of mammogram with TP marked on it. It will drastically reduce the radiologist work by location tumor directly on mammogram.

5. Conclusion

A fully automatic CAD system, which can accurately locate the tumor on a mammogram and reduces FPs, has been proposed. The developed CAD system consists of preprocessing, SOM clustering, ROI extraction,

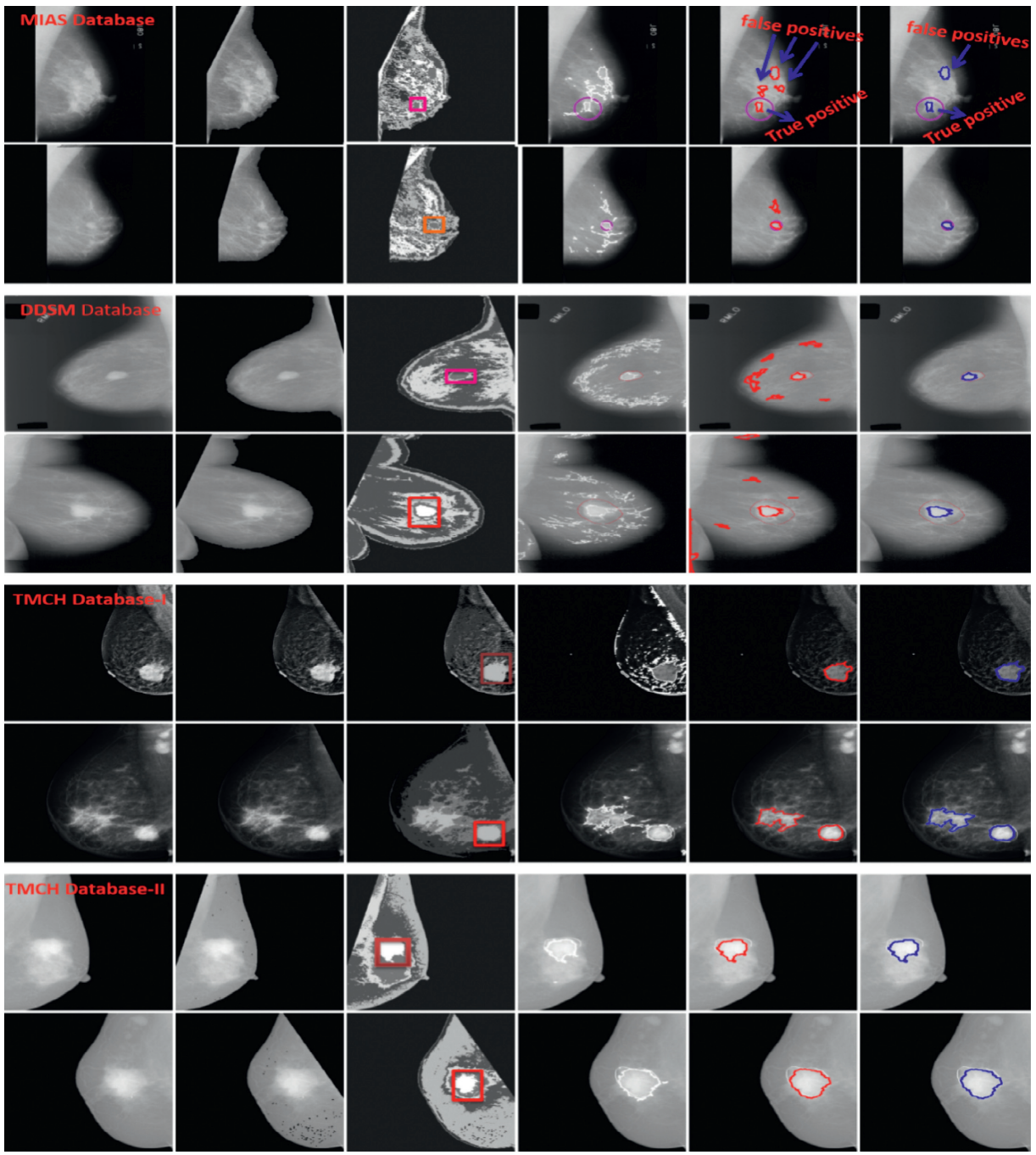


Figure 14: Representation of fully automatic CAD system for breast cancer using (a) sample mammograms from MIAS, DDSM, and TMCH datasets, (b) preprocessed mammograms, (c) clustered image, (d) TP and FP marked on mammogram, (e) TP marked by thresholding, (f) TP marked by using LBP descriptor based on sparse curvelet coefficients.

sparse LBP feature computation based on sparse Curvelet coefficients, and finally, FP reduction using ANN classifier.

The proposed algorithm presents a novel concept of extraction of curvelet coefficients according to irregular shape of mass is called as sparse curvelet coefficients and computation of LBP. The analysis proves that the FPs are reduced significantly from 0.85 to 0.01 FP/image for MIAS, 4.81 to 0.03 FP/image for DDSM and 2.32 to 0.00 FP/image for TMCH. The ANN classifier showed best results as AUC = 0.98 and accuracy = 98.57% for MIAS in benign-malignant classification, AUC = 0.98 and accuracy = 98.70% for DDSM in normal-malignant classification, AUC = 0.98 and accuracy = 98.30% for TMCH: Scanner1, and AUC = 1 and accuracy = 100% for TMCH: Scanner2 in normal-malignant classification as compared with SVM and KNN

Table 8: Comparison of classification accuracy, AUC, and FP/image values from different approaches in breast cancer diagnosis.

Author	Database	Method	Classifier	Result	AUC	FP/image
Eltoukhy et al. [33]	MIAS	Biggest curvelet coefficients as a feature vector	Euclidean classifier	94.07%	—	—
Eltoukhy et al. [42]				98.59	—	—
Eltoukhy et al. [8]			SVM	97.3	—	—
Dhahbi et al. [34]	Mini-MIAS	Curvelet moments	KNN	91.27	—	—
	DDSM			86.46	—	—
Bruno et al. [4]	DDSM	Curvelet + LBP	SVM	85	0.85	—
			PL	94	0.94	—
da Rocha et al. [40]	DDSM	LBP	SVM	88.31	0.88	—
Kanadam and Chereddy [3]	MIAS	Sparse ROI	SVM	97.42	—	—
Pereira et al. [18]	DDSM	Wavelet and Wiener filter	Multiple thresholding, wavelet, and GA	—	—	1.37
Liu and Zeng [29]	DDSM, FFDM	GLCM, CLBP, and geometric features	SVM	—	—	1.48
De Sampaio et al. [39]	DDSM	LBP	DBSCAN	98.26		0.19
Zyout et al. [30]	DDSM	Second order statistics of wavelet coefficients (SOSWC)	SVM	96.8	0.97	0.018
	MIAS			95.2	96.6	0.029
Casti et al. [31]	DDSM	Differential features	Fisher linear discriminant analysis (FLDA)	—	—	1.68
	MIAS					2.12
	FFDM					0.82
Proposed method	MIAS	LBP based on sparse curvelet subband coefficients	ANN	98.57	0.98	0.01
	DDSM			98.70	0.98	0.03
	TMCH: Scanner1			98.30	0.98	0.05
	TMCH: Scanner2			100	1	0

classifier. The performance of LBP features and LBP features based on sparse curvelet coefficients are nearly same which show that the proposed algorithm is suitable for cancer breast tissue diagnosis.

In future, the reduced curvelet coefficients can be used to extract local ternary patterns and other local descriptor and local directional patterns, etc. The present work deals with mammogram with single mass; this can be further extended for multiple mass models with multiple LBP features based on sparse curvelet coefficients.

Data Availability

In this research, we have used two publicly available datasets MIAS and DDSM. These datasets can be found here in [17] and [19]. The third database is collected from the local hospital Tata Memorial Cancer Hospital, Mumbai, which can be found at http://eureka.sveri.ac.in/ or available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The TMCH database for this work was given by Department of Radiodiagnosis, Tata Memorial Cancer Hospital, Mumbai.

References

[1] S. Malvia, S. A. Bagadi, U. S. Dubey, and S. Saxena, "Epidemiology of breast cancer in Indian women," *Asia-Pacific Journal of Clinical Oncology*, vol. 13, no. 4, pp. 289–295, 2017.

[2] A. Gupta, K. Shridhar, and P. Dhillon, "A review of breast cancer awareness among women in India: cancer literate or awareness deficit?," *European Journal of Cancer*, vol. 51, no. 14, pp. 2058–2066, 2015.

[3] K. P. Kanadam and S. R. Chereddy, "Mammogram classification using sparse-ROI: a novel representation to arbitrary shaped masses," *Expert Systems with Applications*, vol. 57, pp. 204–213, 2016.

[4] D. O. T. Bruno, M. Z. do Nascimento, R. P. Ramos et al., "LBP operators on curvelet coefficients as an algorithm to describe texture in breast cancer tissues," *Expert Systems with Applications*, vol. 55, pp. 329–340, 2016.

[5] M. Hussain, "False-positive reduction in mammography using multiscale spatial Weber law descriptor and support vector machines," *Neural Computing and Applications*, vol. 25, no. 1, pp. 83–93, 2014.

[6] M. M. Pawar and S. N. Talbar, "Genetic fuzzy system (GFS) based wavelet co-occurrence feature selection in mammogram classification for breast cancer diagnosis," *Perspectives in Science*, vol. 8, pp. 247–250, 2016.

[7] C. Muramatsu, T. Hara, T. Endo, and H. Fujita, "Breast mass classification on mammograms using radial local ternary patterns," *Computers in Biology and Medicine*, vol. 72, pp. 43–53, 2016.

[8] M. M. Eltoukhy, I. Faye, and B. B. Samir, "A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation,"

Computers in Biology and Medicine, vol. 42, no. 1, pp. 123–128, 2012.

[9] Y. Li, H. Chen, Y. Yang, L. Cheng, and L. Cao, "A bilateral analysis scheme for false positive reduction in mammogram mass detection," *Computers in Biology and Medicine*, vol. 57, pp. 84–95, 2015.

[10] A. Gandhamal, S. Talbar, S. Gajre, A. F. M. Hani, and D. Kumar, "Local gray level S-curve transformation—a generalized contrast enhancement technique for medical images," *Computers in Biology and Medicine*, vol. 83, pp. 120–133, 2017.

[11] S. Anand and S. Gayathri, "Mammogram image enhancement by two-stage adaptive histogram equalization," *Optik-International Journal for Light and Electron Optics*, vol. 126, no. 21, pp. 3150–3152, 2015.

[12] M. M. Pawar and S. N. Talbar, "Local entropy maximization based image fusion for contrast enhancement of mammogram," *Journal of King Saud University-Computer and Information Sciences*, 2018.

[13] K. Ganesan, U. R. Acharya, K. C. Chua, L. C. Min, and K. T. Abraham, "Pectoral muscle segmentation: a review," *Computer Methods and Programs in Biomedicine*, vol. 110, no. 1, pp. 48–57, 2013.

[14] I. K. Maitra, S. Nag, and S. K. Bandyopadhyay, "Technique for preprocessing of digital mammogram," *Computer Methods and Programs in Biomedicine*, vol. 107, no. 2, pp. 175–188, 2012.

[15] A. Oliver, J. Freixenet, J. Martí et al., "A review of automatic mass detection and segmentation in mammographic images," *Medical Image Analysis*, vol. 14, no. 2, pp. 87–110, 2010.

[16] P. Görgel, A. Sertbas, and O. N. Ucan, "Mammographical mass detection and classification using local seed region growing–spherical wavelet transform (lsrg–swt) hybrid scheme," *Computers in Biology and Medicine*, vol. 43, no. 6, pp. 765–774, 2013.

[17] J. Suckling, J. Parker, D. Dance et al., *The Mammographic Image Analysis Society Digital Mammogram Database, International Congress Series*, Exerpta Medica, England, UK, 1994.

[18] D. C. Pereira, R. P. Ramos, and M. Z. do Nascimento, "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm," *Computer Methods and Programs in Biomedicine*, vol. 114, no. 1, pp. 88–101, 2014.

[19] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. Kegelmeyer Jr., "The digital database for screening mammography," in *Proceedings of the 5th International Workshop on Digital Mammography*, M. J. Yaffe, Ed., Medical Physics Publishing, 2001, ISBN 1-930524-00-5.

[20] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation," *Expert Systems with Applications*, vol. 42, no. 3, pp. 990–1002, 2015.

[21] T. Berber, A. Alpkocak, P. Balci, and O. Dicle, "Breast mass contour segmentation algorithm in digital mammograms," *Computer Methods and Programs in Biomedicine*, vol. 110, no. 2, pp. 150–159, 2013.

[22] R. Rouhi and M. Jafari, "Classification of benign and malignant breast tumors based on hybrid level set segmentation," *Expert Systems with Applications*, vol. 46, pp. 45–59, 2016.

[23] L. A. Salazar-Licea, J. C. Pedraza-Ortega, A. Pastrana-Palma, and M. A. Aceves-Fernandez, "Location of mammograms ROI's and reduction of false-positive," *Computer Methods and Programs in Biomedicine*, vol. 143, pp. 97–111, 2017.

[24] T. Kohonen, "The self-organizing map," *Neurocomputing*, vol. 21, no. 1, pp. 1–6, 1998.

[25] A. Demirhan and İ. Güler, "Combining stationary wavelet transform and self-organizing maps for brain MR image segmentation," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 2, pp. 358–367, 2011.

[26] X. Lladó, A. Oliver, J. Freixenet, R. Martí, and J. Martí, "A textural approach for mass false positive reduction in mammography," *Computerized Medical Imaging and Graphics*, vol. 33, no. 6, pp. 415–422, 2009.

[27] G. B. Junior, S. V. da Rocha, M. Gattass, A. C. Silva, and A. C. de Paiva, "A mass classification using spatial diversity approaches in mammography images for false positive reduction," *Expert Systems with Applications*, vol. 40, no. 18, pp. 7534–7543, 2013.

[28] N. Vállez, G. Bueno, O. Déniz et al., "Breast density classification to reduce false positives in CADe systems," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 2, pp. 569–584, 2014.

[29] X. Liu and Z. Zeng, "A new automatic mass detection method for breast cancer with false positive reduction," *Neurocomputing*, vol. 152, pp. 388–402, 2015.

[30] I. Zyout, J. Czajkowska, and M. Grzegorzek, "Multi-scale textural feature extraction and particle swarm optimization based model selection for false positive reduction in mammography," *Computerized Medical Imaging and Graphics*, vol. 46, pp. 95–107, 2015.

[31] P. Casti, A. Mencattini, M. Salmeri et al., "Contour-independent detection and classification of mammographic lesions," *Biomedical Signal Processing and Control*, vol. 25, pp. 165–177, 2016.

[32] S. Beura, B. Majhi, and R. Dash, "Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer," *Neurocomputing*, vol. 154, pp. 1–14, 2015.

[33] M. M. Eltoukhy, I. Faye, and B. B. Samir, "A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram," *Computers in Biology and Medicine*, vol. 40, no. 4, pp. 384–391, 2010.

[34] S. Dhahbi, W. Barhoumi, and E. Zagrouba, "Breast cancer diagnosis in digitized mammograms using curvelet moments," *Computers in Biology and Medicine*, vol. 64, pp. 79–90, 2015.

[35] U. Raghavendra, U. R. Acharya, H. Fujita, A. Gudigar, J. H. Tan, and S. Chokkadi, "Application of Gabor wavelet and locality sensitive discriminant analysis for automated identification of breast cancer using digitized mammogram images," *Applied Soft Computing*, vol. 46, pp. 151–161, 2016.

[36] K. Ganesan, U. R. Acharya, C. K. Chua, L. C. Min, K. T. Abraham, and K.-H. Ng, "Computer-aided breast cancer detection using mammograms: a review," *IEEE Reviews in Biomedical Engineering*, vol. 6, pp. 77–98, 2013.

[37] M. Abdel-Nasser, H. A. Rashwan, D. Puig, and A. Moreno, "Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern," *Expert Systems with Applications*, vol. 42, no. 24, pp. 9499–9511, 2015.

[38] S. K. Wajid and A. Hussain, "Local energy-based shape histogram feature extraction technique for breast cancer diagnosis," *Expert Systems with Applications*, vol. 42, no. 20, pp. 6990–6999, 2015.

[39] W. B. de Sampaio, A. C. Silva, A. C. de Paiva, and M. Gattass, "Detection of masses in mammograms with adaption to breast density using genetic algorithm, phylogenetic trees, LBP and

SVM," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8911–8928, 2015.

[40] S. V. da Rocha, G. B. Junior, A. C. Silva, A. C. de Paiva, and M. Gattass, "Texture analysis of masses malignant in mammograms images using a combined approach of diversity index and local binary patterns distribution," *Expert Systems with Applications*, vol. 66, pp. 7–19, 2016.

[41] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[42] M. M. Eltoukhy, I. Faye, and B. B. Samir, "Breast cancer diagnosis in digital mammogram using multiscale curvelet transform," *Computerized Medical Imaging and Graphics*, vol. 34, no. 4, pp. 269–276, 2010.

[43] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[44] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Modeling & Simulation*, vol. 5, no. 3, pp. 861–899, 2006.

[45] A. Dudhane, G. Shingadkar, P. Sanghavi, B. Jankharia, and S. Talbar, "Interstitial lung disease classification using feed forward neural networks," in *Proceedings of Advances in Intelligent Systems Research*, vol. 137, pp. 515–521, 2017.

[46] A. A. Dudhane and S. N. Talbar, "Multi-scale directional mask pattern for medical image classification and retrieval," in *Proceedings of 2nd International Conference on Computer Vision & Image Processing,* Indian Institute of Technology Roorkee, Roorkee, India, 2018.

[47] http://eureka.sveri.ac.in/.

Hindawi
Journal of Healthcare Engineering
Volume 2018, Article ID 1930357, 13 pages
https://doi.org/10.1155/2018/1930357

Research Article

On the Development of Virtual Reality Scenarios for Computer-Assisted Biomedical Applications

Eder H. Govea-Valladares,[1] Hugo I. Medellin-Castillo,[1] Jorge Ballesteros,[2] and Miguel A. Rodriguez-Florido[2]

[1]*Facultad de Ingeniería, Universidad Autónoma de San Luis Potosí, 78290, San Luis Potosí, Mexico*
[2]*Instituto Tecnológico de Canarias, Playa Pozo Izquierdo s/n, 35119, Las Palmas, Islas Canarias, Spain*

Correspondence should be addressed to Hugo I. Medellin-Castillo; hugoivanmc@uaslp.mx

Received 24 January 2018; Accepted 17 July 2018; Published 30 August 2018

Academic Editor: Subrahmanyam Murala

The modelling of virtual environments and scenarios is an important area of research for the development of new computer-assisted systems in the areas of engineering and medicine, particularly in the area of biomechanics and biomedical engineering. One of the main issues while designing a virtual environment is the level of realism, which depends on the computing capacity and the level of accuracy and usefulness of the generated data. Thus, the dilemma is between the aesthetic realism and the information utility. This paper proposes a methodology to develop low-cost and high-quality virtual environments and scenarios for computer-aided biomedical applications. The proposed methodology is based on the open-source software Blender and the Visualization Toolkit libraries (VTK). In order to demonstrate the usability of the proposed methodology, the design and development of a computer-assisted biomedical application is presented and analysed.

1. Introduction

In the early days, scientific research was based on observations of natural and physical phenomena. However, in the last years, the research scope has included the modelling and simulation of physical phenomena by means of computer technologies. The sight has been the most developed sense in computer simulations and virtual environments, leading to the origin of the scientific visualization concept [1]. By means of scientific visualization, it is possible to transform mathematical data into 2D or 3D images or vice versa [2], allowing the communication and understanding of large amount of information efficiently. It also allows the visualization of physical phenomena that are not possible to visualize by common methods, such as experimental observations. Moreover, the representation of reality is also possible by means of animations, modelling, and graphic renderings.

Virtual reality (VR) can be described as a set of technologies that enable people to interact with a virtual environment beyond reality [3]. VR takes advantage of the computer technological development and scientific visualization to create a virtual world [4]. The use of VR has become very popular because it offers a high level of realism and immersion but requires advanced computing technologies capable of processing large amounts of scientific data and graphics [5]. VR has been used in different areas such as engineering, medicine, education, entertainment, astronomy, archaeology, and arts. In the area of medicine, virtual environments are created to enable the interaction with the human body anatomy [6]. The practice of medicine is a complex decision-making process that requires knowledge, experience, and manual abilities [7]; practitioner abilities are gained by training and experience, which is a slow process that may take several years. In order to get experience and abilities, a medical apprentice must be the protagonist of his/her training but considering as the main priority the avoidance of risks and unnecessary inconveniences for the patient [8]. Consequently, the use of VR and computer technologies in medicine has become an important tool for students and practitioners to understand and confirm

concepts and to improve surgical skills and for experienced surgeons to make more precise diagnosis and plan the surgery [9–11]. One of the main applications of VR and computer technologies in medicine has been the development of computer-assisted surgery and simulation systems [12, 13].

In the area of engineering, VR applications include the design and evaluation of components and prototypes before construction [14–16], the manufacturing planning of components [12, 17–19], and assembly training [20]. On the other hand, VR applications in the area of art include virtual sculpting [21], reconstruction and preservation of buildings [22, 23], and development of environments, structures, and scenarios for the film industry [24].

A main issue of VR and computer-assisted applications is the design and development of the virtual environment (VE), which comprises the modelling of virtual objects, including their geometry and surface characteristics such as colours and textures. However, the amount of data to be processed by the VR application increases as the level of realism of the virtual environment increases, affecting the performance of the application. Thus, the dilemma is between the largest amount of data to increase the quality of the results, and the minimum use of computer resources with an acceptable level of realism.

This work presents a methodology to design and develop high-quality virtual scenarios with a high degree of realism for medical applications. The methodology is based on the open-source software Blender and VTK, leading to a low-cost development. The proposed approach is intended to be used in the development of computer-assisted biomedical applications, such as surgery planning, simulation, and training.

2. Literature Review

The solid modelling developed in the mid-1970s [25] is a set of mathematical principles for modelling objects, such as solid or hollow shapes delimited by a mesh, using computational methods [26], that is, creating digital models of physical objects of the real world. The main feature of solid modelling is that it is focused on the surface characteristics of the object. The constructive solid geometry (CSG) considers the modelling of solid objects using Boolean operations, which is useful for tasks requiring mathematical precision [27, 28]. On the other hand, Boundary Representation (B-Rep) [29] connects vertices with lines to create faces, allowing the generation of complex geometries where the level of accuracy depends on the amount of elements in the mesh [27, 28].

Several applications of solid modelling in the area of medicine can be found in the literature, for instance, the 3D modelling of a femur by using the software called 3ds Max [30]. One example of designing virtual scenarios for surgical planning was presented by Domínguez-Quintana et al. [31]. Snyder et al. [32] presented an investigation to compare the impact of training with or without supervision using a virtual reality surgical simulator of laparoscopy and endoscopy. Debes et al. [33] compared the training effectiveness between a virtual simulator and a training video in laparoscopic surgery. A successful virtual environment requires that virtual parts behave as the parts in the real world [34].

Virtual reality (VR) technologies can be used to enhance the performance of surgical simulators by providing a virtual environment where users can get the feeling of immersion in a real environment, in addition to more intuitive cues such as collisions between virtual objects, collisions with obstacles, friction, inertia, restitution, 3D rendering, and sound [35]. Moreover, virtual environments can be improved by incorporating haptic technologies to provide the user with the sense of touch. Haptics allows natural manipulation of virtual objects by enabling the user with the feeling of collisions, forces, weight, and inertia of virtual objects. In this way, haptic-enabled computer-assisted medical applications are more intuitive, accurate, and efficient than conventional computer-aided systems in medicine [11].

In general, a VR application in medicine comprises five main modules: (1) model reconstruction module, to generate 3D models from medical data such as CT and MRI images, (2) visualization module, responsible for the graphics rendering of the virtual environment, (3) manipulation module, to provide the interaction between the user and the virtual environment, (4) simulation module, responsible for the physical based behaviour of the virtual environment and objects, and (5) data module, responsible for processing, analysing, and exporting the medical data.

There have been several research works reported in the literature focusing on the development and analysis of computer-aided VR applications in medicine. However, few works have addressed the development of virtual environments for such applications [10]. A virtual environment has a great impact on the performance of the application; therefore, its importance is high. Sometimes, it is necessary to sacrifice the level of realism to prioritize the data processing and time response of the system. Nowadays, there are several tools such as modelling software, computers with high computing power, and measurement devices, to develop virtual scenarios. However, these tools can be very expensive and limited to those users with financial capability to such tools. In addition, very few works in the literature have focused on reducing the costs of developing virtual environments.

3. Methodology

A new methodology to create virtual reality scenarios (environments) is proposed as shown in Figure 1. This methodology has been implemented using the open-source software Blender 2.49, the Visualization Toolkit VTK 5.6, and the Python 2.7 programming language in Windows operating system.

The main steps of the proposed methodology are as follows:

(1) *Create scene*. Modelling starts in Blender and the first step is to generate the virtual models to be placed on the stage. 3D models can be imported or created using the commands, primitive objects, and Boolean operations of Blender.

(2) *Materials and textures*. Textures are then assigned to each element of the models. The texture corresponds to the type of material and the visual aspect that they

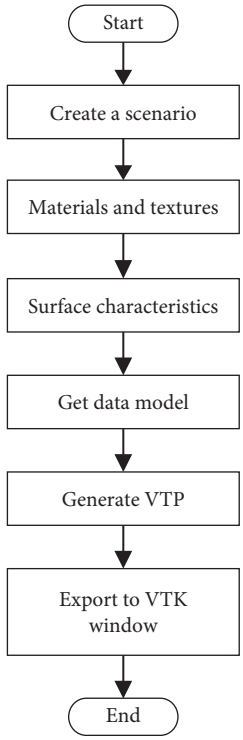


Figure 1: Methodology to create virtual reality scenarios.

have in the real world. This process of adding textures is done by using standard images or by assigning colour to each part in Blender.

(3) *Surface characteristics*. To boost realism, it is necessary to provide the VE with visual features like lighting, shadows, reflections, transparencies, and more. This process can be carried out in Blender.

(4) *Get data model*. Once the scene has been completed, the next step is to extract the information corresponding to orientation, rotation, and location of each object in the scene.

(5) *Generate VTP*. In this step, the virtual scenario is converted into a VTP file. The VTP (VTK Polygonal Data) is a VTK file that contains the polygonal data of the 3D model.

(6) *Export to VTK window*. Finally, the VTP file is used to import all the elements of the scene into a VTK window in the external application being developed. The algorithm to export the Blender scene to a VTK window of the external application is shown in Figure 2.

In order to show the level of realism that can be obtained using the proposed methodology, two virtual scenarios were developed. These scenarios are described in the next paragraphs.

3.1. Jaw Articulator. The first medical virtual scenario corresponds to a jaw articulator, which is a mechanical device that represents the human jaw joints, and that is used to simulate and adjust the motion of the jaw physical models. The jaw articulator is used for oral and maxillofacial surgery planning. Figure 3(a) shows the real-world articulator used as reference. Each part of the articulator was created using primitive objects and Boolean operations in Blender 2.49. Texture characteristics such as colour and visualization properties such as transparency level, reflection, shadows, and so on were also added to the different parts of the virtual model, Figure 3(b). A 3D model of a human jaw was obtained from medical image reconstruction and imported into the virtual scenario where the articulator was created. Also an image type texture was added to simulate the real bone texture. Finally, a smoothing filter was applied to obtain a smooth surface free of imperfections due to the mesh. Figure 3(c) shows the virtual jaw articulator and the final virtual scenario corresponding to a hospital.

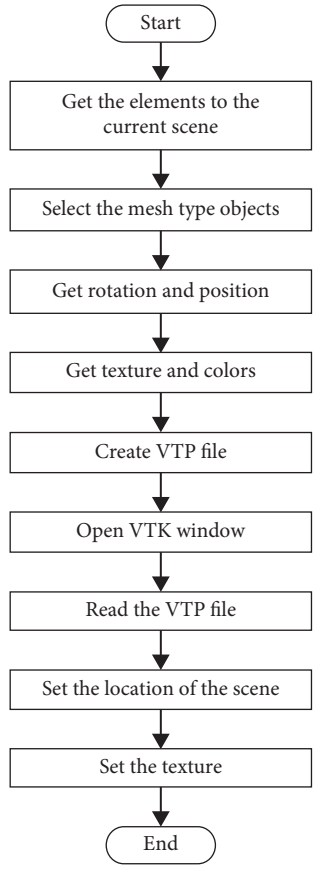
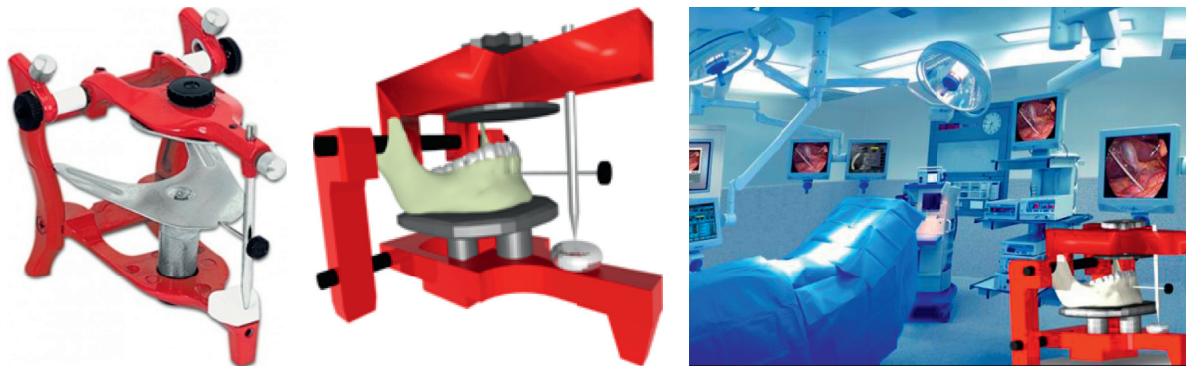


Figure 2: Algorithm to export the Blender scene to a VTK application.

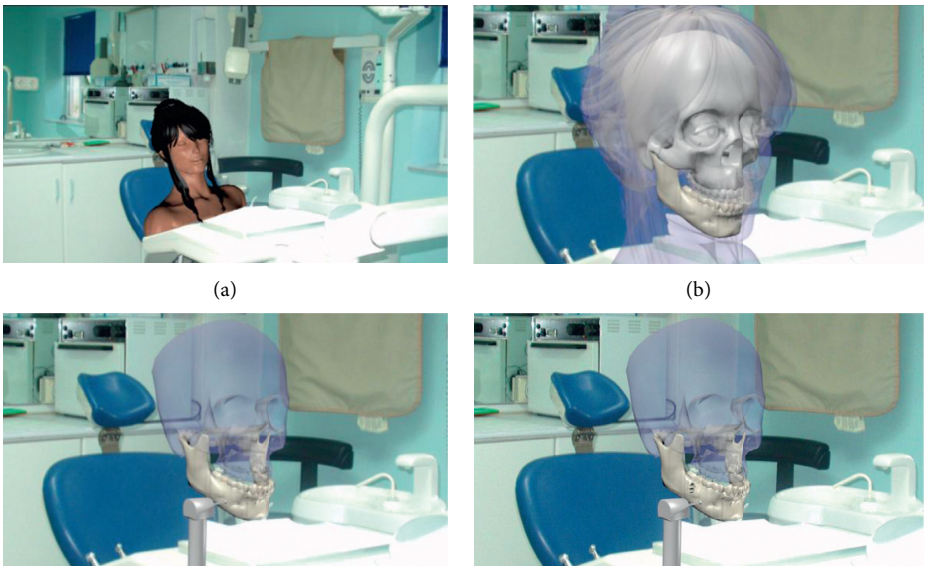
From Figure 3, it can be observed the lighting effects on the elements' surface, the texture applied to the mandible, and some surface properties to simulate the light reflection. It can also be observed that there are some elements that simulate different types of materials, such as plastics and metals, resulting in objects with transparency and very high gloss. Moreover, metallic elements give a more realistic visualization.

3.2. Surgical Simulator. Biomedical engineering makes use of virtual environments to develop surgical simulators for planning and training various clinical procedures. Thus, the second virtual scenario corresponds to an orthognathic surgery simulator to perform dental surgeries or procedures in a virtual dental office environment. Figure 4(a)



(a) (b) (c)

Figure 3: Jaw articulator. (a) Real articulator. (b) Virtual articulator. (c) Virtual scenario.



(a) (b)

(c) (d)

Figure 4: Medical surgical simulator. (a) Dental room. (b) Transparent layers of virtual patient. (c) Virtual tool for bone cutting. (d) Results of virtual bone cutting.

shows the virtual scene corresponding to a dental room with a patient sitting in a dental unit. Textures, lights, and images can be used to increase the level of realism. Transparency can be also used to observe internal details such as the bone structure, Figure 4(b). Dental surgical tools, such as mills or drills, can be modelled and added to the virtual scenario to perform virtual cuts on the patient jaw bone, Figure 4(c). The advantage of this type of virtual scenarios is that the practitioner can perform the surgical procedures as many times as necessary in order to practice or to plan the real surgery procedure. Figure 4(d) shows the footprint of the virtual cut performed on the virtual mandible.

4. Case Study

In order to show the details and usability of the proposed methodology for the development of VR scenarios for computer-aided biomedical applications, a Virtual Osteotomy Simulator System (VOSS) for 3D osteotomy simulation and training was developed and evaluated.

4.1. System Description. The VOSS general architecture is shown in Figure 5 and comprises four main modules:

(1) Visualization module, responsible for carrying out the graphic rendering of virtual objects, tools, and virtual environments

(2) Osteotomy module, responsible for computing bone cuts and enabling virtual osteotomies

(3) Manipulation module, to allow the 3D free-form movement and manipulation of virtual objects and surgical tools

(4) Data exportation module, to export any information regarding the osteotomy simulation (e.g., STL models)

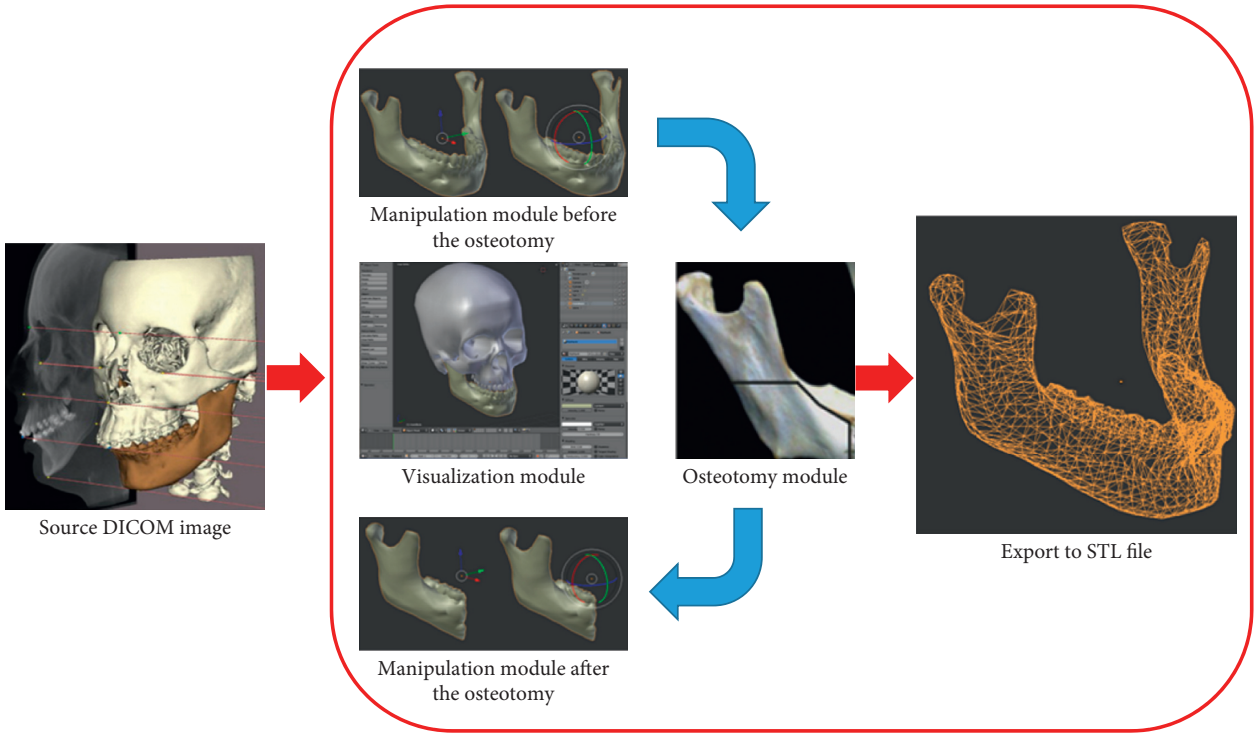


Figure 5: VOSS architecture.

The VOSS was implemented using Python 2.7 and Blender 2.59 in a Workstation with a dual-core AMD processor (3.0 GHz/dual core), 4 GB of RAM, and a NVIDIA Quadro FX3500 PCI Express graphics card. The main capabilities of the VOSS are as follows:

(i) Virtual reality environment and real-time response

(ii) 3D visualization of anatomical models and tools, including textures and transparency

(iii) 3D free manipulation and interaction of virtual cutting tools, bones, and bone fragments

(iv) Simulation of single and multiple osteotomies

(v) Free-form cutting path to perform osteotomies

(vi) Free camera manipulation

(vii) Automatic scaling of models.

The GUI of the VOSS is shown in Figure 6. The biomodels can be imported into VOSS as STL or 3DS file formats, which can be generated from medical images (e.g., DICOM images) or 3D scanning.

4.2. Virtual Scenario. Figure 7 shows the particular methodology used to define the virtual reality environment of the VOSS. This methodology comprises the following steps in Blender:

(1) Create a scenario. Generate the virtual environment (lights, cameras, and background image)

(2) Load skull and jaw biomodels. Import the skull and jaw models as STL or 3DS file formats

(3) Add texture to bone. Add texture to the skull and jaw bones by means of images. Since the skull is used only as reference and visual support, its texture can be set as transparent

(4) Model cutting tools. Create the surgical cutting tools (saw and drill) in Blender or import them as STL or 3DS file formats

(5) Add texture to tools. In order to reproduce the real appearance of surgical cutting tools, provide them with texture

(6) Create sensors. Generate sensors for the user to control and manipulate objects in the virtual environment by means of the keyboard and/or mouse buttons

(7) Create controllers. Generate controllers to specify the action to be executed after the activation of a sensor

(8) Create actuators. Generate actuators to perform the movement or manipulation of the virtual objects according to the sensors. Movements can be either linear or rotational.

The addition of textures to biomodels or tools can be made by means of images with standard file formats (e.g., bmp, jpg, jpeg, and png). The movement and manipulation of virtual models in VOSS is made by means of a numeric keypad, an alphanumeric keypad, a standard computer mouse, or a 3D computer mouse. These movements are defined by sensors,

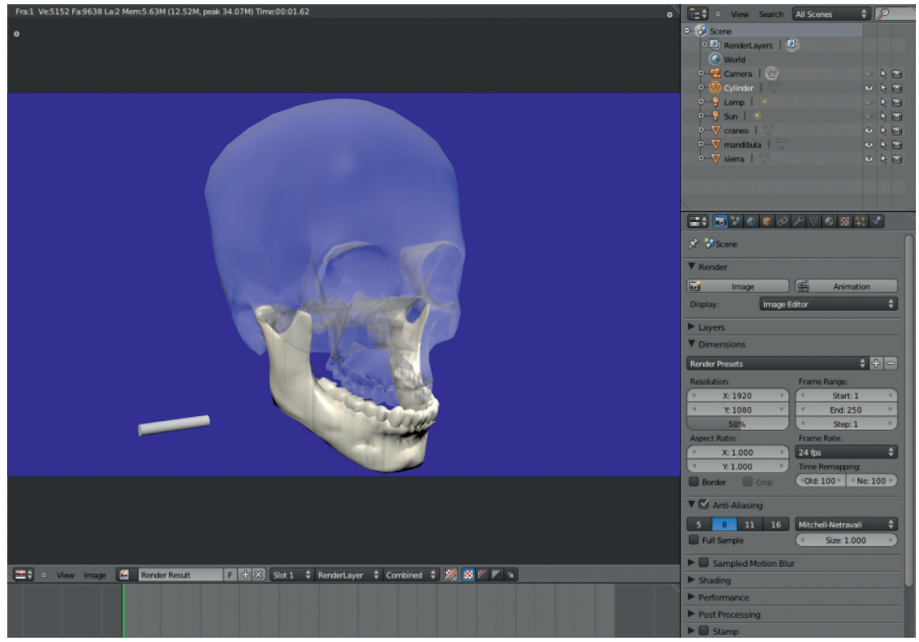


Figure 6: VOSS graphic user interface (GUI).

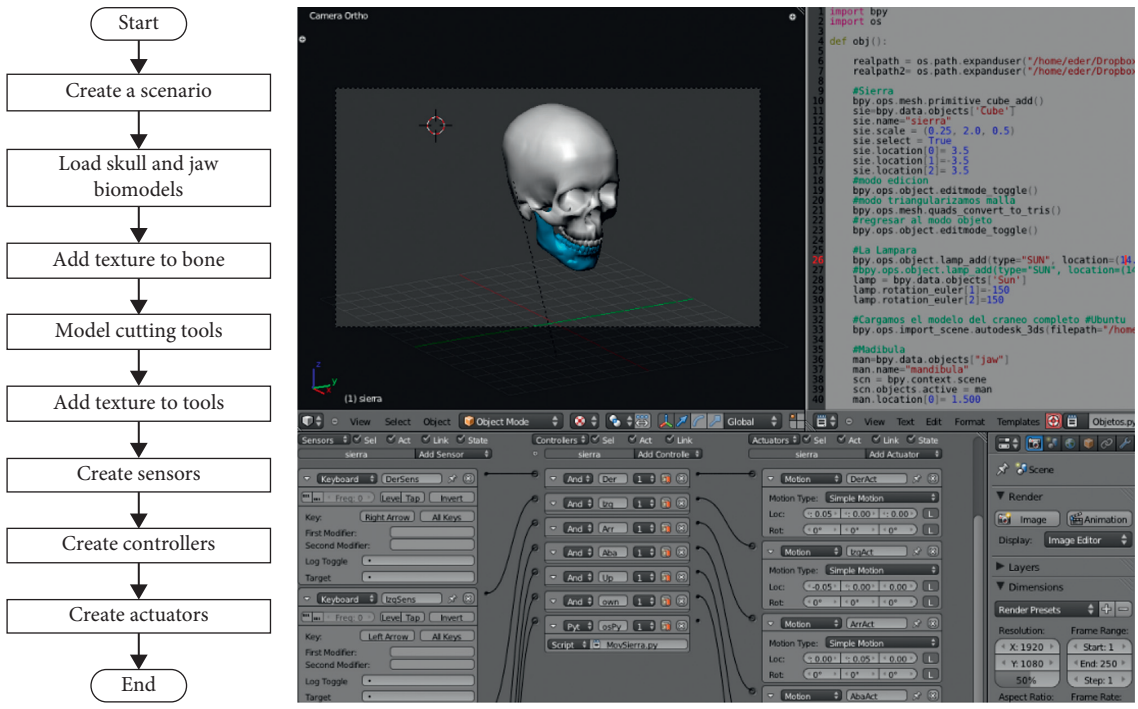


Figure 7: VOSS virtual reality environment. (a) Methodology. (b) Implementation.

controllers, and actuators, which are created in the Logic mode of the Game Engine in Blender. A sensor is a function for the user to control objects in the virtual environment by means of the keyboard and/or mouse buttons. On the other hand, a controller is a function used to define the action to be executed after the activation of a sensor. Finally, an actuator executes the movements of the virtual objects according to the sensors.

4.3. Virtual Osteotomy. The virtual osteotomy procedure implemented in the VOSS corresponds to a Bilateral Sagittal Split Osteotomy Ramus Mandibular (BSSROM) of a human mandible. The aim is to perform bone cutting operations on virtual models of human jaws, that is, to simulate the work of a maxillofacial surgeon when correcting bone malformations. Figure 8 shows the general virtual osteotomy procedure in VOSS.

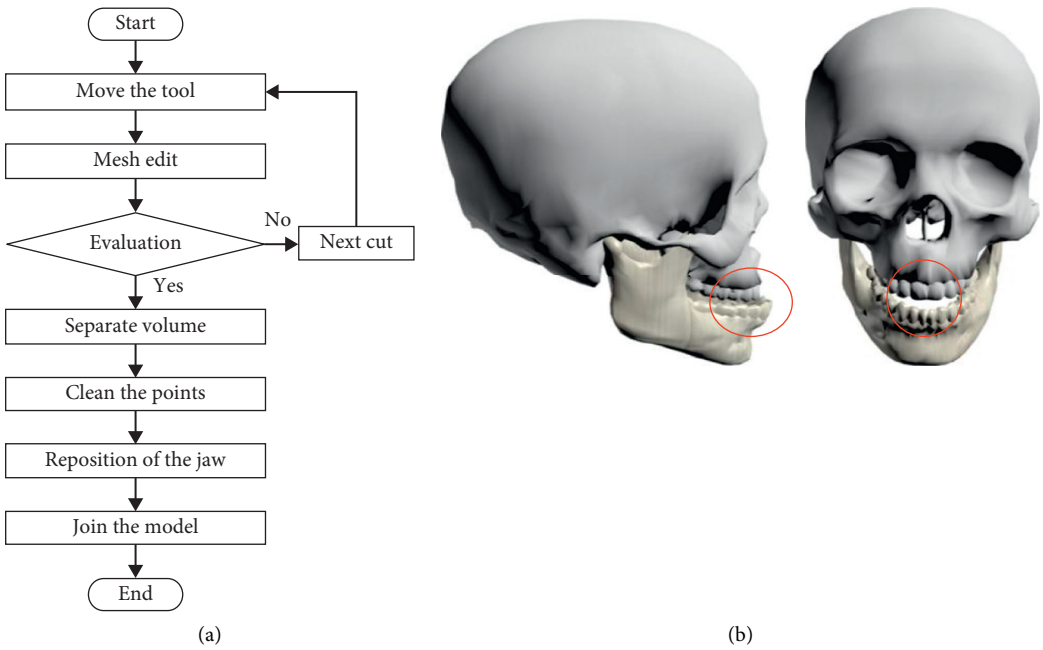


Figure 8: Virtual osteotomy. (a) Procedure. (b) Virtual model with malformation.

The osteotomy procedure begins by selecting a cutting tool and placing it at the position where the first cut is meant to be made, Figure 9. Once the tool is positioned, the cut is performed and it can be repeated as many times as necessary to make a longitudinal cut along the jaw. The user is able to freely move the tool in any 3D path while performing the cut. Figure 10(a) shows the simulation of a vertical cut using a drill, while Figure 10(b) shows the simulation of a cut operation using a sagittal saw. Similar to the real procedure, the virtual jaw can be separated into two fragments (jaw 1 and jaw 2), which can be moved or manipulated independently. Figure 11 shows the bone fragments after the bone separation. The virtual cutting and separation of biomodels are performed by means of the Boolean operations in Blender.

After the mandible splitting, the user is able to manipulate and relocate the jaw fragments in order to reduce or eliminate the bone defect or malformation. Once the mandible fragments are relocated at the correct position, a Boolean operation is carried out to join the jaw fragments. Figure 12 shows the last movement and final relocation of the mandible.

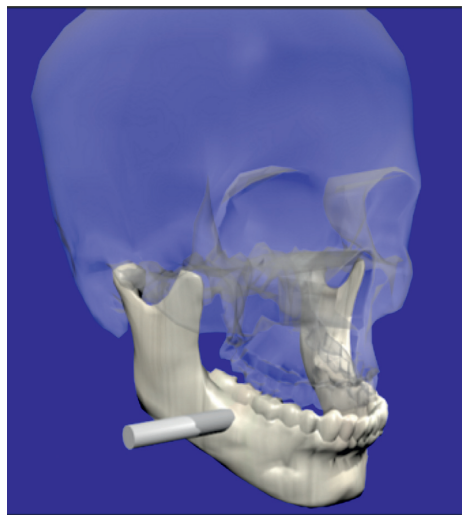
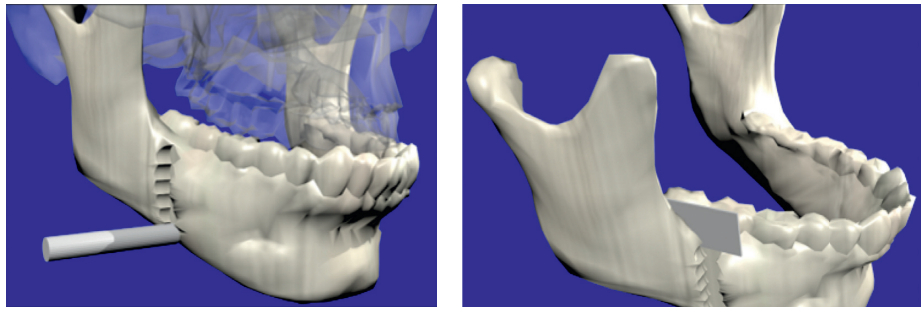


Figure 9: Initial positioning of the tool.

4.4. Virtual Osteotomy Training. A biomedical application to evaluate the proposed virtual osteotomy approach as a training tool was developed in the programming language C++ using the Microsoft Foundation Classes (MFC) of Visual Studio 2010, the Visualization Toolkit libraries (VTK) for graphics rendering, and the H3DAPI haptic rendering software development platform for the manipulation of virtual objects. The application provides the user with force feedback by means of a haptic device. Additionally, the application supports different types of haptic devices, including the Phantom Omni from Sensable and the Falcon from Novint.

The overall experimental methodology used to evaluate the virtual osteotomy training is shown in Figure 13. A total of nine students of the Oral and Maxillofacial Surgery Postgraduate Program of the "Hospital Central Dr. Ignacio Morones Prieto" in San Luis Potosi, Mexico, were selected. These participants were selected because previous knowledge of the maxillofacial surgical procedures was required. The nine participants were divided into three groups with three persons in each group:



(a) (b)

Figure 10: Cutting process. (a) Drilling. (b) Sawing.

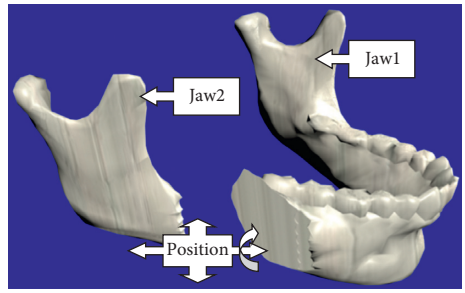
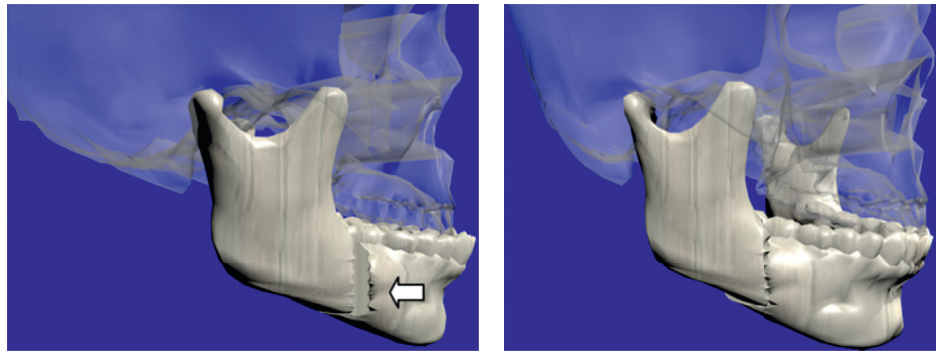


Figure 11: Repositioning of mandible fragments, jaw 1 and jaw 2.



(a) (b)

Figure 12: Final step of the virtual osteotomy. (a) Relocation of mandible fragments. (b) Joining of the mandible.

(i) *Group I*. No virtual training. This group of participants carried out the real osteotomy procedure without previous virtual training.

(ii) *Group II*. Virtual training without force feedback. Before carrying out the real osteotomy procedure, this group of participants undertook virtual training but with no force feedback.

(iii) *Group III*. Virtual training with force feedback. Before carrying out the real osteotomy procedure, this group of participants undertook virtual training with force feedback.

Two different osteotomy procedures were considered in the evaluation: mentoplasty (or chin cut) and sagittal osteotomy (or branch cut) of a human mandible. The mentoplasty procedure comprises one main cut on the chin, while the sagittal osteotomy procedure comprises three cuts: sagittal exterior, sagittal superior, and sagittal interior. Figure 14 shows these four cuts marked by an experienced maxillofacial surgeon on a real human mandible. These cutting trajectories were used as reference for evaluating the usability of virtual training.

At the beginning, all participants were informed about the general background related to the experiments, the conditions in which they would be working, and the experimental procedure. Then, all participants received a verbal explanation about the osteotomy procedures under consideration and the cuts required, allowing them to ask questions and receive further explanation. Participants of Groups II and III were instructed on the use of the virtual system and the virtual osteotomy procedure, given them the opportunity to familiarise themselves with the

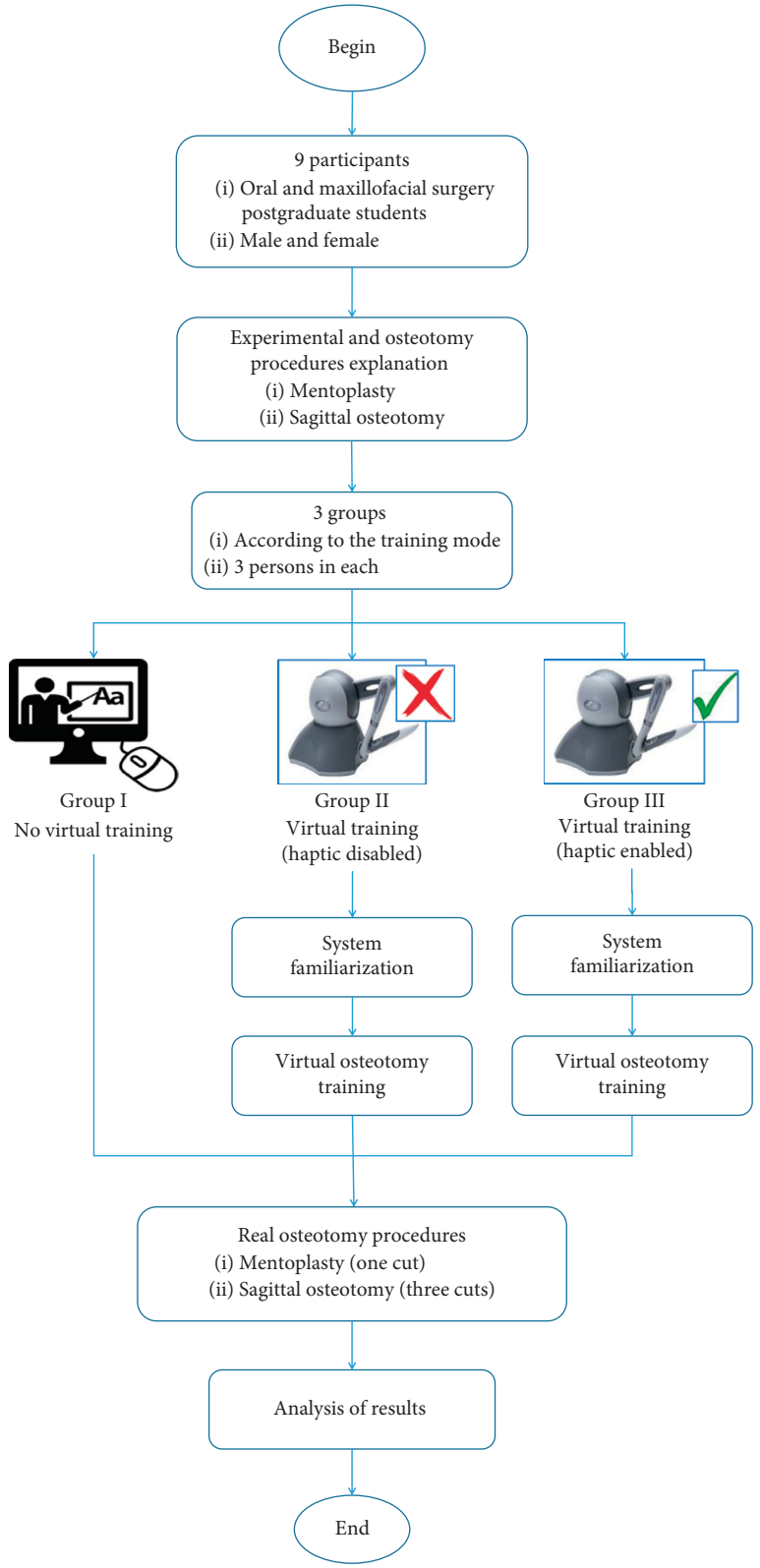
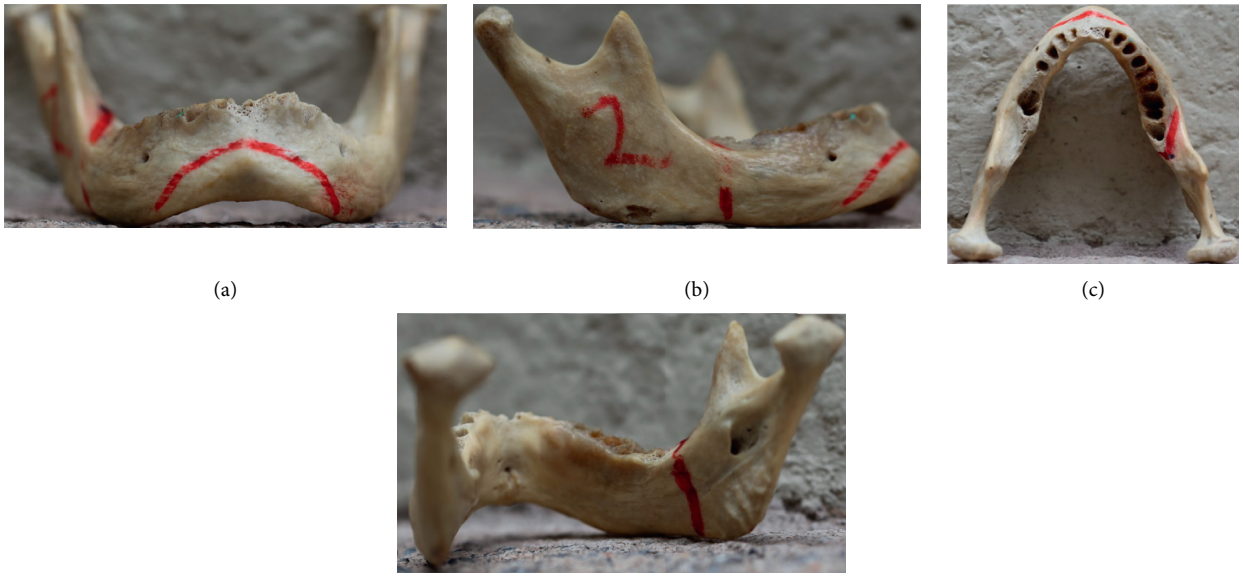


Figure 13: Experimental methodology to evaluate the virtual osteotomy training.

system for twenty minutes before undertaking virtual training.

Virtual osteotomy training for participants of Groups II and III consisted in the realization of the two osteotomy procedures in the virtual environment. Participants of Group II were able to manipulate the virtual surgical tools by means of the haptic device but without receiving force feedback, whereas participants of Group III did receive force



(a) (b) (c)

(d)

Figure 14: Osteotomy procedures marked by an expert on a real human mandible. (a) Mentoplasty. (b) Sagittal exterior. (c) Sagittal superior. (d) Sagittal interior.

feedback during the manipulation of the virtual surgical tools.

Virtual training was carried out on virtual models in the virtual environment by means of a Phantom Omni haptic device. On the other hand, the real osteotomy procedures were carried out on physical 3D prototypes, made of PLA in a 3DTouch printer from 3D systems®, and using a conventional high-speed milling tool, as shown in Figure 15.

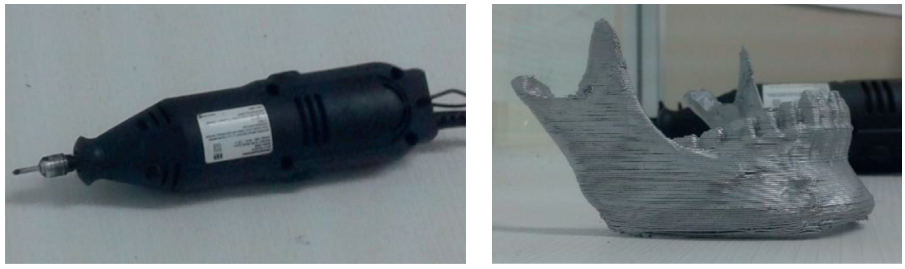
4.5. Results and Discussion. Figure 16 shows the results of the four virtual osteotomies performed by one participant of Group II. On the other hand, Figure 17 shows the real osteotomy procedure being performed by one participant of Group I.

To evaluate the effectiveness of virtual training, all participants of Groups I, II, and III were observed during the real osteotomy procedure executions, and the time to complete the task was measured for all participants. The cutting error was also evaluated by comparing the cutting trajectory performed by each user, with the "ideal" cutting trajectory performed by an experienced maxillofacial surgeon, as shown in Figure 18(a). The cutting error was defined as the difference between the two trajectories, red area of Figure 18(b). This error was quantified in a CAD software.

Table 1 summarizes the results obtained for Groups I, II, and III during the real osteotomy procedures. The values reported in this table correspond to the average values obtained by each group of participants. From these results, it is observed that participants of Group I showed average task completion times of 432 s and 389 s for the mentoplasty and sagittal real osteotomy procedures, respectively. On the other hand, participants of Group II performed these procedures in 308 s and 317 s, respectively, and participants of Group III carried out these real tasks in 120 s and 241 s, respectively. These results suggest that participants that undertook virtual training first (Groups II and III) had superior performance, in terms of time, than those who did not train (Group I). In other words, participants who trained virtually completed the real osteotomy procedures faster than those who did not train. Moreover, participants who virtually trained with haptic force feedback (Group III) completed the real osteotomy procedures faster than those who virtually trained but without force feedback.

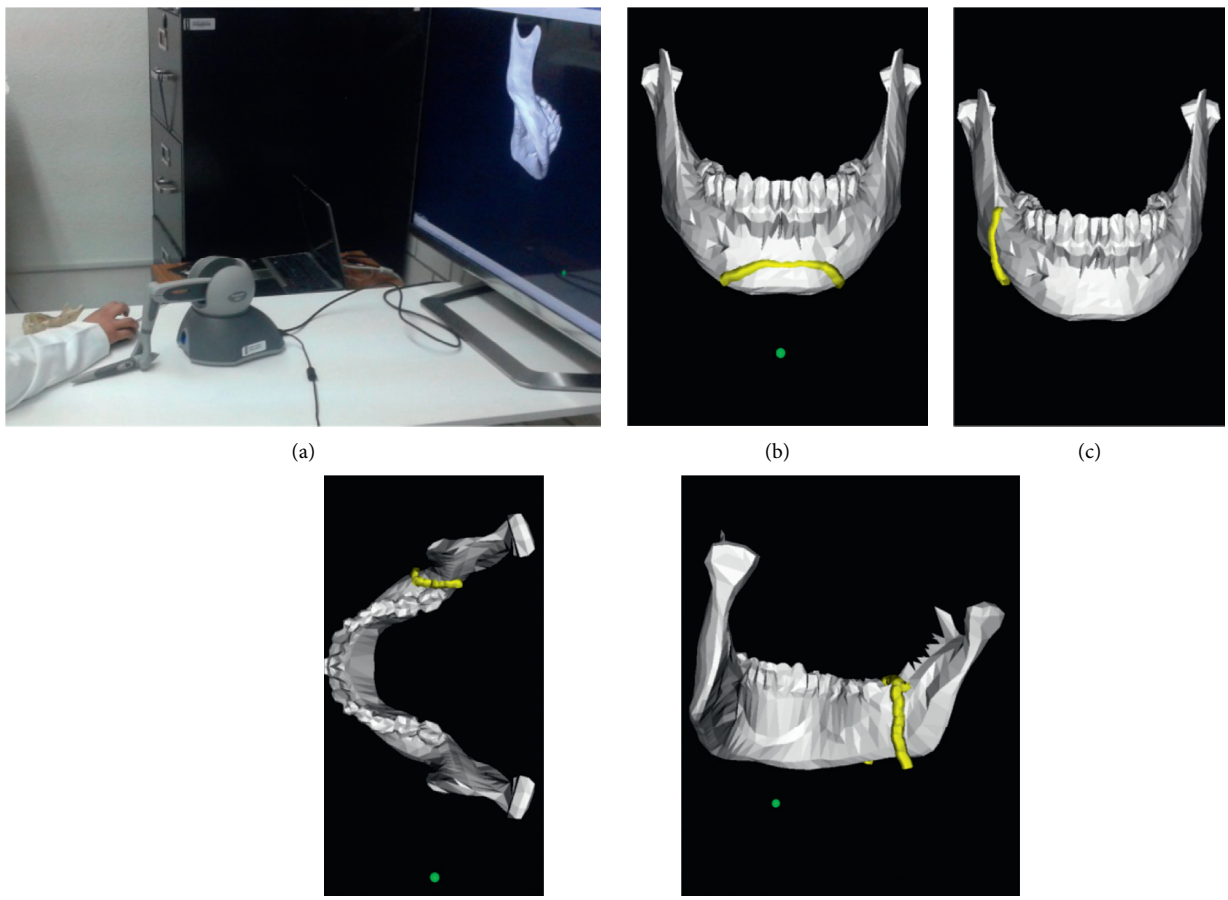
Regarding the cutting error, participants of Group I exhibited an average cutting error of 24.2% and 27.7% for the mentoplasty and sagittal real osteotomy procedures, respectively, whereas participants of Group II exhibited an average cutting error of 14.6% and 21.2%, respectively, and participants of Group III obtained an average cutting error of 6.4% and 4.9%, respectively. These results clearly evidence than participants of Group II and Group III, who undertook a virtual training period first, achieved a more accurate mandible cutting during the real osteotomy procedure than those who did not train first (Group I). Furthermore, participants of Group III who trained with force feedback exhibited smaller cutting errors than participant of Group II who trained but without force feedback.

Therefore, it can be said that the proposed virtual osteotomy training procedure is a feasible approach to improve the performance and skills of the participants. A faster and more accurate osteotomy procedure was achieved by subjects who undertook virtual training first than those who did not train virtually first. Additionally, the use of haptic force feedback during virtual training has showed to enhance the virtual training procedure; a better performance is achieved by subjects that train with haptic force feedback than subjects that train with no haptic force feedback. Thus, the usability of the



(a) (b)

Figure 15: Real osteotomy implements. (a) Manual high-speed milling tool. (b) Physical jaw.



(a) (b) (c)

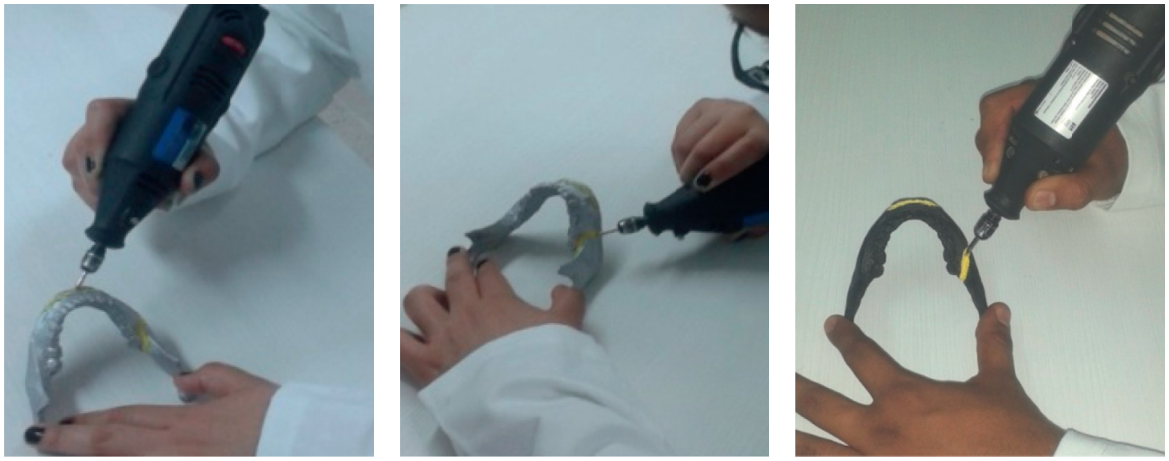
(d) (e)

Figure 16: Virtual osteotomy training. (a) Participant. (b) Mentoplasty. (c) Sagittal exterior. (d) Sagittal superior. (e) Sagittal interior.

proposed methodology to design and develop virtual scenarios for biomedical applications has been demonstrated.

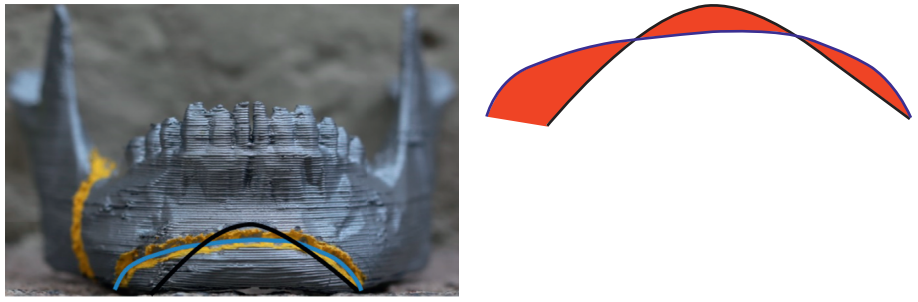
5. Conclusions

A new methodology to develop virtual reality environments for biomedical applications has been presented in this paper. This methodology represents a low-cost solution for the development of virtual environments with a high level of realism and with physical characteristics very close to the real devices. A case study corresponding to a virtual osteotomy simulator was developed using the proposed methodology. The results of this virtual osteotomy training biomedical application have demonstrated the effectiveness of the virtual system; users increased their abilities and skills to perform real osteotomy procedures. Therefore, the



(a) (b) (c)

Figure 17: Real osteotomy procedure. (a) Mentoplasty. (b) Sagittal exterior. (c) Sagittal superior.



(a) (b)

Figure 18: Error evaluation (black line: ideal trajectory, blue line: participant trajectory). (a) Front picture. (b) Error measurement (red area).

Table 1: Time and error results of the real osteotomy procedures.

Participants	Mentoplasty		Sagittal	
	RO time (seconds)	Cutting error (%)	RO time (seconds)	Cutting error (%)
Group I	432	24.2	389	27.7
Group II	308	14.6	317	21.2
Group III	120	6.4	241	4.9

RO: Real Osteotomy.

usability of the proposed methodology and approach has been validated. Future work considers a more comprehensive analysis of the effect of biomedical virtual training on the performance of subjects. This analysis will include a larger number of participants, modelling and rendering of real cutting forces, and comparison with traditional training.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the financial support from CONACYT (National Science and Technology Council of Mexico), research grant CB-2010-01-154430, and from Secretariat of Public Education (SEP) of Mexico. They also thank the Instituto Tecnológico de Canarias ITC for the support provided to this research.

References

[1] C. Wang and H. W. Shen, "Information theory in scientific visualization," *Entropy*, vol. 13, no. 1, pp. 254–273, 2001.

[2] P. Du, W. Li, S. Sang, L. Wang, and X. Zhou, "Application of 3D visualization concept layer model for coal–bed methane indez system," *Procedia Earth and Planetary Science*, vol. 1, no. 1, pp. 977–981, 2009.

[3] P. BergLeif and M. VanceJudy, "Industry use of virtual reality in product design and manufacturing: a survey," *Virtual Reality*, vol. 21, no. 1, pp. 1–17, 2016.

[4] I. Kartiko, M. Kavakli, and K. Cheng, "Learning science in a virtual reality application: the impacts of animated-virtual actors visual complexity," *Computers and Education*, vol. 55, no. 2, pp. 881–891, 2010.

[5] O. Rübel, S. Ahern, E. W. Bethel et al., "Coupling visualization and data analysis for knowledge discovery from multi-dimensional scientific data," *Procedia Computer Science*, vol. 1, no. 1, pp. 1757–1764, 2010.

[6] C. A. Ruiz, F. Montagut, J. Yeison, and E. Heidenreich, "Algorithm for virtual modeling of organs by medical images," *Vector*, vol. 4, pp. 14–26, 2009.

[7] G. Vázquez-Mata, "Realidad virtual y simulación en el entrenamiento de los estudiantes de medicina," *Educación Médica*, vol. 11, no. 1, pp. 29–31, 2008.

[8] G. Subhas, A. Gupta, and V. K. Mittal, "Necessity for improvement in endoscopy training during surgical residency," *The American Journal of Surgery*, vol. 199, no. 3, pp. 331–335, 2010.

[9] L. Hanna, "Simulated surgery: the virtual reality of surgical training," *Surgery*, vol. 28, no. 9, pp. 463–468, 2010.

[10] E. Govea and H. I. Medellín-Castillo, "Design and development of virtual reality environments for biomedical and engineering applications," in *Proceedings of ASME International Mechanical Engineering Congress and Exposition, Volume 14: Emerging Technologies; Safety Engineering and Risk Analysis; Materials: Genetics to Structures*, Houston, TX, USA, November, 2015.

[11] H. I. Medellín-Castillo, E. H. Govea-Valladares, C. N. Pérez-Guerrero, J. Gil-Valladares, T. Lim, and M. Ritchie James, "The evaluation of a novel haptic-enabled virtual reality approach for computer-aided cephalometry," *Computer Methods and Programs in Biomedicine*, vol. 130, pp. 46–53, 2016.

[12] S. Girod, E. Keeve, and B. Girod, "Advances in interactive craniofacial surgery planning by 3D simulation and visualization," *International Journal of Oral and Maxillofacial Surgery*, vol. 24, no. 1, pp. 120–125, 1995.

[13] V. N. Palter and T. P. Grantcharov, "Virtual reality in surgical skills training," *Surgical Clinics of North America*, vol. 90, no. 3, pp. 605–617, 2010.

[14] J. D. Cardona, M. A. Hidalgo, H. Castán, F. Rojas, D. Borro, and H. Jaramillo, *Realidad Virtual y Procesos de Manufactura*, Programa Editorial Universidad Autonoma de Occidente, Rosemead, CA 91770, USA, 1st edition, 2007.

[15] AA. Kadir, X. Xu, and E. Hammerle, "Virtual machine tools and virtual machining—a technological," *Robotics and Computer-Integrated Manufacturing*, vol. 27, no. 3, pp. 494–508, 2011.

[16] J. Nomura and K. Sawada, "Virtual reality technology and its industrial applications," *Annual Reviews in Control*, vol. 25, pp. 99–109, 2001.

[17] G. Gonzalez-Badillo, H. Medellin-Castillo, T. Lim, J. Ritchie, and S. Garbaya, "The development of a physics and constraint-based haptic virtual assembly system," *Assembly Automation*, vol. 34, no. 1, pp. 41–55, 2014.

[18] S. Jayaram, H. I. Connacher, and K. W. Lyons, "Virtual assembly using virtual reality techniques," *Computer Aided Design*, vol. 29, no. 8, pp. 575–584, 1997.

[19] A. Pérez and L. A. Sanz, "Virtual reality simulation applied to a numerical control milling machine," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 1, no. 3, pp. 143–154, 2007.

[20] E. Gallegos-Nieto, H. I. Medellín-Castillo, G. González-Badillo, T. Lim, and J. Ritchie, "The analysis and evaluation of the influence of haptic-enabled virtual assembly training on real assembly performance," *The International Journal of Advanced Manufacturing Technology*, vol. 89, no. 1, pp. 581–598, 2017.

[21] L. Stanculescu, R. Chaine, and M. P. Cani, "Freestyle: sculpting meses with self-adaptive topology," *Computers and Graphics*, vol. 35, pp. 614–622, 2011.

[22] F. Bellotti, R. Berta, R. Cardona, and A. De Gloria, "An architectural approach to efficient 3D urban modeling," *Computers and Graphics*, vol. 35, no. 5, pp. 1001–1012, 2011.

[23] D. A. Guttentag, "Virtual reality: applications and implications for tourism tourism management," *Tourism Management*, vol. 31, no. 5, pp. 637–651, 2010.

[24] D. Di Sorte, M. Femminella, and G. Reali, "QoS-enabled multicast for delivering live events in a digital cinema scenario," *Journal of Network and Computer Applications*, vol. 32, pp. 314–344, 2009.

[25] G. Farin, J. Hoschek, and M. S. Kim, "Solid modeling. Handbook of computer aided geometric design," Tech. Rep. 2, Elsevier, North Holland, Netherlands, 2001.

[26] G. Elber, N. Patrikalakis, and P. Burnet, "Solid modeling theory and applications," *Graphical Models*, vol. 67, no. 5, pp. 371-372, 2005.

[27] A. A. G. Requicha, "Representations for rigid solids: theory, methods, and systems," *ACM Computing Surveys*, vol. 12, p. 4, 1980.

[28] A. Thakur, A. G. Banerjee, and S. K. Gupta, "A survey of CAD model simplification techniques for physics-based simulation applications," *Computer-Aided Design*, vol. 41, no. 2, pp. 65–80, 2009.

[29] R. V. Fleisig and A. D. Spence, "Techniques for accelerating B-rep based parallel machining simulation," *Computer-Aided Design*, vol. 37, no. 12, pp. 1229–1240, 2005.

[30] S. Filippi, B. Motyl, and C. Bandera, "Analysis of existing methods for 3D modelling of femurs starting from two orthogonal images and development of a script for a commercial software package," *Computer Methods and Programs in Biomedicine*, vol. 89, no. 1, pp. 76–82, 2008.

[31] L. Domínguez-Quintana, M. A. Rodríguez-Florido, J. Ruiz-Alzola, and D. Sosa, "Modelado 3D de escenarios virtuales realistas para simuladores quirúrgicos aplicados a la funduplicatura de nissen," *Revista Científica de la Sociedad Española de Informatica y Salud*, vol. 48, pp. 14–20, 2004.

[32] C. W. Snyder, M. J. Vandromme, S. L. Tyra, and M. T. Hawn, "Proficiency–based laparoscopic and endoscopic training with virtual reality simulators: a comparison of protored and independent approaches," *Journal of Surgical Education*, vol. 66, pp. 201–206, 2009.

[33] A. J. Debes, R. Aggarwal, I. Balasundaram, and M. B. Jacobsen, "A tale of two trainers: virtual reality versus a video trainer for acquisition of basic laparoscopic skills," *The American Journal of Surgery*, vol. 199, pp. 840–845, 2010.

[34] B. M. Howard and J. M. Vance, "Desktop haptic virtual assembly using physically based modeling," *Virtual Reality*, vol. 11, no. 4, pp. 207–215, 2007.

[35] T. Gutierrez, J. I. Barbero, M. Aizpitarte, A. R. Carrillo, and A. Eguidazu, "Assembly simulation through haptic virtual prototypes," in *Proceedings of the Third PHANTOM Users Group Workshop*, Cambridge, MA, USA, December 1998.

Hindawi
Journal of Healthcare Engineering
Volume 2018, Article ID 4168538, 11 pages
https://doi.org/10.1155/2018/4168538

Research Article

Deep Convolutional Neural Networks for Chest Diseases Detection

Rahib H. Abiyev and Mohammad Khaleel Sallam Ma'aitah

Department, of Computer Engineering, Near East University, North Cyprus, Mersin-10, Turkey

Correspondence should be addressed to Mohammad Khaleel Sallam Ma'aitah; mohammad.maaitah@neu.edu.tr

Received 29 January 2018; Revised 9 May 2018; Accepted 12 June 2018; Published 1 August 2018

Academic Editor: Subrahmanyam Murala

Chest diseases are very serious health problems in the life of people. These diseases include chronic obstructive pulmonary disease, pneumonia, asthma, tuberculosis, and lung diseases. The timely diagnosis of chest diseases is very important. Many methods have been developed for this purpose. In this paper, we demonstrate the feasibility of classifying the chest pathologies in chest X-rays using conventional and deep learning approaches. In the paper, convolutional neural networks (CNNs) are presented for the diagnosis of chest diseases. The architecture of CNN and its design principle are presented. For comparative purpose, back-propagation neural networks (BPNNs) with supervised learning, competitive neural networks (CpNNs) with unsupervised learning are also constructed for diagnosis chest diseases. All the considered networks CNN, BPNN, and CpNN are trained and tested on the same chest X-ray database, and the performance of each network is discussed. Comparative results in terms of accuracy, error rate, and training time between the networks are presented.

1. Introduction

Medical X-rays are images which are generally used to diagnose some sensitive human body parts such as bones, chest, teeth, skull, and so on. Medical experts have used this technique for several decades to explore and visualize fractures or abnormalities in body organs [1]. This is due to the fact that X-rays are very effective diagnostic tools in revealing the pathological alterations, in addition to its noninvasive characteristics and economic considerations [2]. Chest diseases can be shown in CXR images in the form of cavitations, consolidations, infiltrates, blunted costophrenic angles, and small broadly distributed nodules [3]. By analyzing the chest X-ray image, the radiologists can diagnose many conditions and diseases such as pleurisy, effusion, pneumonia, bronchitis, infiltration, nodule, atelectasis, pericarditis, cardiomegaly, pneumothorax, fractures, and many others [4].

Classifying the chest X-ray abnormalities is considered as a tedious task for radiologists; hence, many algorithms were proposed by researchers to accurately perform this task [5–7]. Over the past decades, computer-aided diagnosis (CAD) systems have been developed to extract useful information from X-rays to help doctors in having a quantitative insight about an X-ray. However, these CAD systems could not have achieved a significance level to make decisions on the type of conditions of diseases in an X-ray [2–4]. Thus, the role of them was left as visualization functionality that helps doctors in making decisions.

A number of research works have been carried out on the diagnosis of chest diseases using artificial intelligence methodologies. In [1], multilayer, probabilistic, learning vector quantization, and generalized regression neural networks have been used for diagnosis chest diseases. The diagnosis of chronic obstructive pulmonary and pneumonia diseases was implemented using neural networks and artificial immune system [8]. In [9], the detection of lung diseases such as TB, pneumonia, and lung cancer using chest radiographs is considered. The histogram equalization in image segmentation was applied for image preprocessing, and feedforward neural network is used for classification purpose. The above research works have been efficiently used in classifying medical diseases; however, their performance was not as efficient as the deep networks in terms of accuracy, computation time, and minimum square error

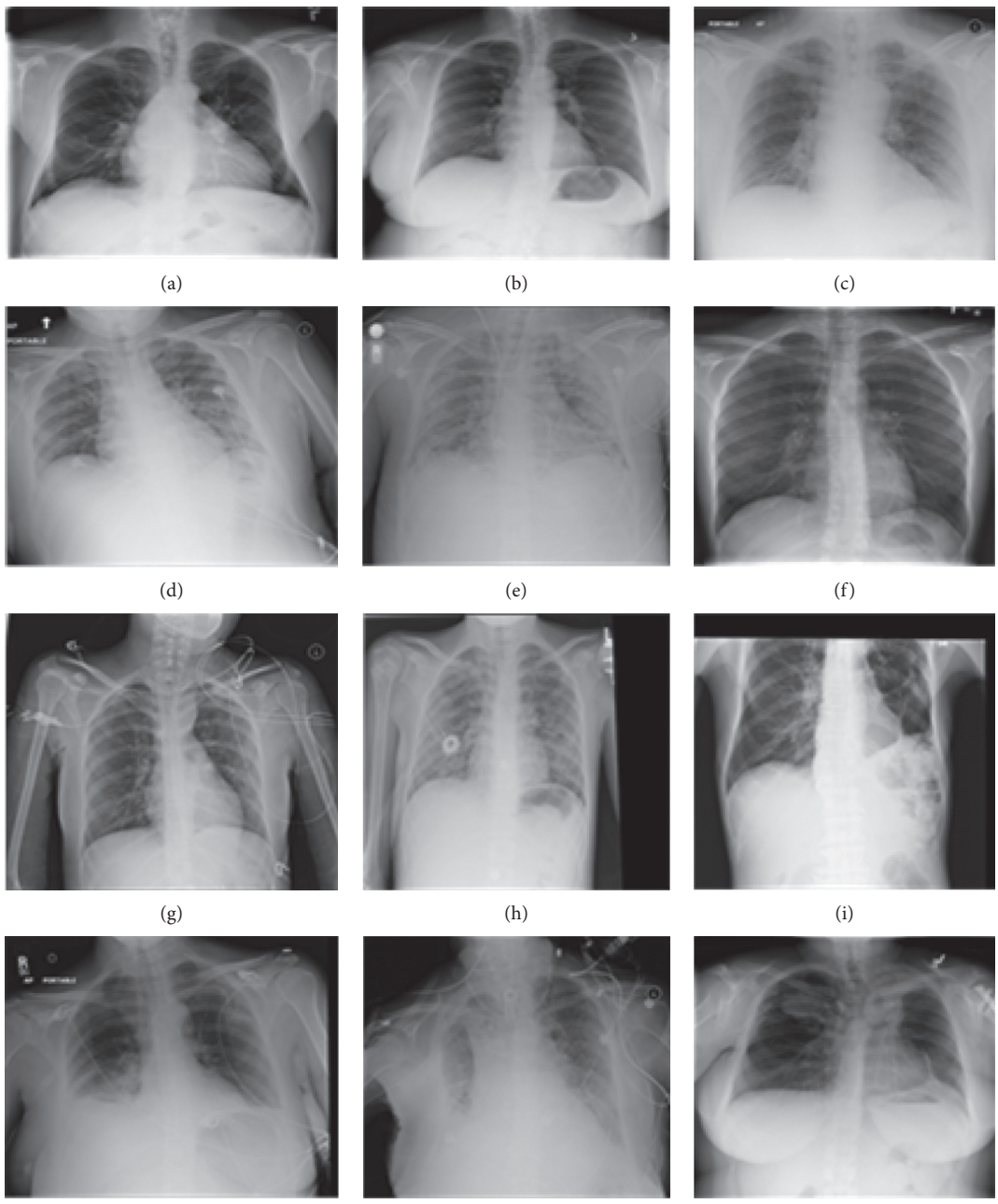


Figure 1: Chest X-ray diseases. (a) Atelectasis. (b) Cardiomegaly. (c) Consolidation. (d) Edema. (e) Effusion. (f) Emphysema. (g) Fibrosis. (h) Infiltration. (i) Mass. (j) Nodule. (k) Pneumonia. (l) Pneumothorax.

achieved. Deep learning-based systems have been applied to increase the accuracy of image classification [10, 11]. These deep networks showed superhuman accuracies in performing such tasks. This success motivated the researchers to apply these networks to medical images for diseases classification tasks, and the results showed that deep networks can efficiently extract useful features that distinguish different classes of images [12–15]. Most commonly used deep learning architecture is the convolutional neural network (CNN). CNN has been applied to various medical images classification due to its power of extracting different level features from images [11–15].

Having gone through the related research studies, in this paper, a deep convolutional neural network (CNN) is employed to improve the performance of the diagnosis of the chest diseases in terms of accuracy and minimum square error achieved. For this purpose, traditional and deep learning-based networks are employed to classify most common thoracic diseases and to present comparative results. Backpropagation neural network (BPNN), competitive neural network (CpNN), and convolutional neural network (CNN) are examined to classify 12 common diseases that may be found in the chest X-ray, that is, atelectasis, cardiomegaly, effusion, infiltration, mass, nodule, pneumonia, pneumothorax, consolidation, edema, emphysema,

and fibrosis (Figure 1). In this paper, we aim at training both traditional and deep network using the same chest X-ray dataset and evaluating their performances. The data used in the paper are obtained from the National Institutes of Health—Clinical Center [16]. The dataset contains 112,120 frontal-view X-ray images of 30,805 unique patients.

This paper is structured as follows: Section 2 presents the methodologies used for diagnosis chest diseases. A brief explanation of the BPNN, CpNN, and CNN is given. A description of the convolutional neural network used for diagnosis chest diseases and its operating principles are presented. Section 3 discusses the results of simulations of the networks used, in addition to the database description. A comparison of the performances of the networks used in simulations is given in Section 4, and Section 5 is the conclusion part of the paper.

2. Machine Learning for Diagnosis of Chest Diseases

2.1. Backpropagation Neural Network (BPNN). Backpropagation neural network (BPNN) is a multilayer feedforward neural network that uses a supervised learning algorithm known as error back-propagation algorithm. Errors accumulated at the output layer are propagated back into the network for the adjustment of weights [16–19]. Figure 2 depicts a conventional BPNN which consists of three layers: input, hidden, and output. As seen in Figure 2, there is no backward pass of computation except the operations used in training. All the functioning operations proceed in the forward direction during simulation.

The pseudocode algorithm for BPNN is given below [20].

(i) Network initialization: randomly choose the initial weights

(ii) Select first training pair

(iii) Forward computation that includes the following steps:

  (a) Apply the inputs to the network
  (b) Calculate the output for every neuron from the input layer, through the hidden layer(s), to the output layer
  (c) Calculate the error at the outputs

(iv) Backward computation

  (a) Use the output error to compute error signals for preoutput layers
  (b) Use the error signals to compute weight adjustments
  (c) Apply the weight adjustments

(v) Repeat Forward and Backward computations for other training pairs.

(vi) Periodically evaluate the network performance. Repeat Forward and Backward computations until the network converges on the target output.

To calculate outputs for each neuron based on the input pattern, the equations below can be used. The output of the j-th neuron for the pattern p is O_{pj}:

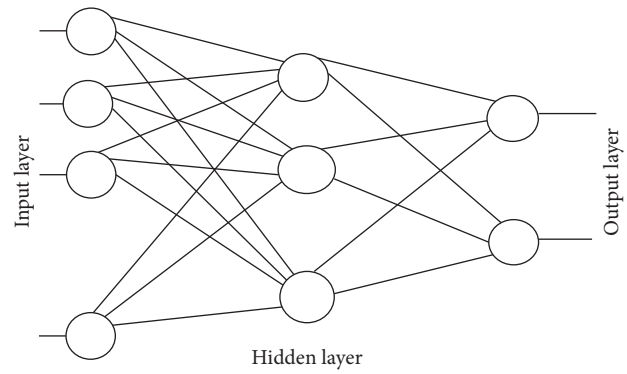


Figure 2: Backpropagation neural network.

$$O_{pj}\left(\text{net}_j\right) = \frac{1}{1 + e^{-\lambda \text{net}_j}},$$
$$\text{where} \quad \text{net}_j = b_j + \sum_k O_{pk} W_{kj}, \tag{1}$$

where k ranges over the input indices, W_{kj} is the weight on the connection from k-th input to j-th neuron, and b_j is the bias weight for the j-th output neuron.

To calculate the error signal at the output, the equations below can be used:

$$E = \frac{1}{2} \sum_{i=1}^{N} \left(T_{pj} - O_{pj}\right)^2, \tag{2}$$

where T_{pj} is the target value of the j-th output neuron for pattern p and O_{pj} is the actual output value of the j-th output neuron for pattern p.

The backpropagation algorithm is based on the gradient descent optimization method [20–22]. By determining the derivative of error, we can update the network parameters. The output neuron error signal d_{pj} is determined as follows:

$$d_{pj} = \left(T_{pj} - O_{pj}\right) O_{pj} \left(1 - O_{pj}\right). \tag{3}$$

To calculate the error signal for each hidden neuron, the equations below can be used.

The hidden neuron error signal δ_{pj} is given by

$$\delta_{pj} = O_{pj}\left(1 - O_{pj}\right) \sum_k \delta_{pk} W_{kj}, \tag{4}$$

where δ_{pk} is the error signal of a postsynaptic neuron k and W_{kj} is the weight of the connection from j-th hidden neuron to the k-th postsynaptic neuron [21].

To calculate and apply weight adjustments, the equations below can be used:

$$W_{ji}(t+1) = W_{ji}(t) - \gamma \Delta W_{ji}(t) + \beta\left(W_{ji}(t) - W_{ji}(t-1)\right), \tag{5}$$

where γ is the learning rate and β is the momentum. Here,

$$\Delta W_{ji}(t) = \delta_{pj} O_{pj}. \tag{6}$$

2.2. Competitive Neural Network. The competitive neural network is a simple neural network that consists of two

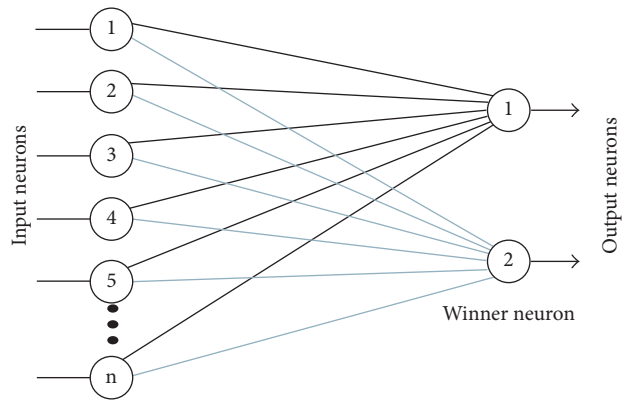


Figure 3: Competitive neural network.

layers and uses an unsupervised learning algorithm for training. The inputs of the network are features, and the outputs are the classes. The input layer is fully connected to the output layer. Each connection between input and output layers is characterized by weight coefficients. In every epoch, the neurons in the output layer compete among themselves when input features are applied to the network input [23–25]. The competitive neural network (Figure 3) relies fundamentally on the Hebbian learning rule. The distinction is the following: in competitive learning, output neurons have to compete among themselves to get activated, and only one neuron is activated at any time, as compared to Hebbian learning where more than one neuron can be activated or fired at any time.

These networks use a "winner-takes-all" strategy, where only the weights connected to the winner neuron are updated in a particular epoch, while other weights are not updated [24, 25]. This learning process has the resultant effect of increasingly strengthening the correlation between the inputs and the corresponding winner neurons during learning.

When the patterns are supplied to the input layer, the neurons in the output layer compete among themselves to be activated [23–25]. The rules used to update the weights of these networks are given below. For output winner neuron k, we have

$$\Delta w_{kj} = \eta\left(x_j - w_{kj}\right), \tag{7}$$

where η is the learning rate, x_j is the j-th input pattern, w_{kj} is the weight connection between j-th and k-th neurons, and Δw_{kj} is the computed weight change.

If k-th output neuron loses at epoch p, then

$$\Delta w_{kj} = 0. \tag{8}$$

Weight update for k-th neuron at epoch $(p+1)$ is achieved using the following equation:

$$w_{kj}(p+1) = w_{kj}(p) + \Delta w_{kj}. \tag{9}$$

2.3. Convolutional Neural Networks. Deep learning is a machine learning method inspired by the deep structure of a mammal brain [26]. The deep structures are characterized by multiple hidden layers allowing the abstraction of the different levels of the features. In 2006, Hinton et al. developed a new algorithm to train the neuron layers of deep architecture, which they called greedy layerwise training [12]. This learning algorithm is seen as an unsupervised single layer greedily training where a deep network is trained layer by layer. Because this method became more effective, it has been started to be used for training many deep networks. One of the most powerful deep networks is the convolutional neural network that can include multiple hidden layers performing convolution and subsampling in order to extract low to high levels of features of the input data [27–30]. This network has shown a great efficiency in different areas, particularly, in computer vision [28], biological computation [29], fingerprint enhancement [30], and so on. Basically, this type of networks consists of three layers: convolution layers, subsampling or pooling layers, and full connection layers. Figure 4 shows a typical architecture of a convolutional neural network (CNN). Each type of layer is explained briefly in the following sections.

2.3.1. Convolution Layer. In this layer, an input image of size $R*C$ is convolved with a kernel (filter) of size $a*a$ as shown in Figure 4. Each block of the input matrix is independently convolved with the kernel and generated a pixel in the output. The result of the convolution of the input image and kernel is used to generate n output image features. Generally, a kernel of the convolution matrix is referred to as a filter while the output image features obtained by convolving kernel and the input images are referred to as feature maps of size $i*i$.

CNN can include multiple convolutional layers, the inputs and outputs of next convolutional layers are the feature vector. There is a bunch of n filters in each convolution layer. These filters are convolved with the input, and the depth of the generated feature maps $(n*)$ is equivalent to the number of filters applied in the convolution operation. Note that each filter map is considered as a specific feature at a certain location of the input image [31–33].

The output of the l-th convolution layer, denoted as $C_j^{(l)}$, consists of feature maps. It is computed as

$$C_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{a_i^{(l-1)}} K_{i,j}^{(l-1)} * C_j^{(l)}, \tag{10}$$

where $B_i^{(l)}$ is the bias matrix and $K_{i,j}^{(l-1)}$ is convolution filter or kernel of size $a*a$ that connects the j-th feature map in layer $(l-1)$ with the i-th feature map in the same layer. The output $C_i^{(l)}$ layer consists of feature maps. In (10), the first convolutional layer $C_i^{(l-1)}$ is input space, that is, $C_i^{(0)} = X_i$.

The kernel generates feature map. After the convolution layer, the activation function can be applied for nonlinear transformation of the outputs of the convolutional layer:

$$Y_i^{(l)} = Y\left(C_i^{(l)}\right), \tag{11}$$

where $Y_i^{(l)}$ is the output of the activation function and $C_i^{(l)}$ is the input that it receives.

Typically used activation functions are sigmoid, tanh, and rectified linear units (ReLUs). In this paper, ReLUs

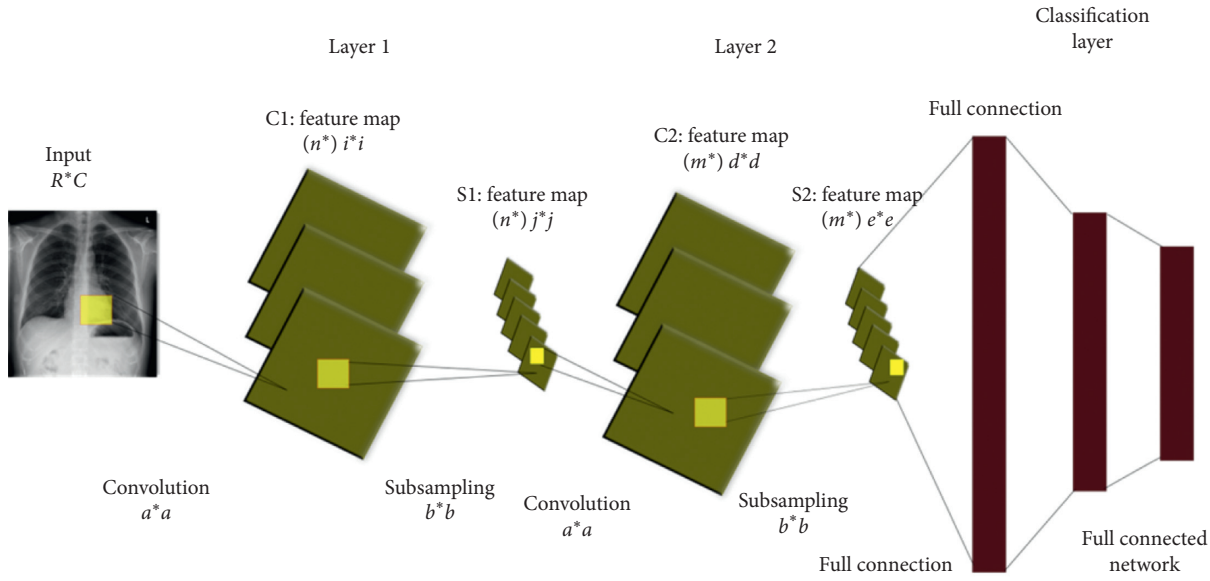


Figure 4: Convolutional neural network.

which is denoted as $Y_i^{(l)} = \max(0, Y_i^{(l)})$ are used. This function is popularly used in deep learning models due to its help in reducing the interaction and nonlinear effects. ReLU converts the output to 0 if it receives a negative input, while it returns the same input value if it is positive. The advantage of this activation function over other functions is the faster training because of the error derivative, which becomes very small in the saturating region; therefore, the updates of the weights almost vanish. This is called the vanishing gradient problem.

2.3.2. Subsampling Layer. The main aim of this layer is to spatially reduce the dimensionality of the features maps extracted from the previous convolution layer. To do so, a mask of size $b*b$ is selected as shown in Figure 4, and the subsampling operation between the mask and the feature maps is performed. Many subsampling methods were proposed such as averaging pooling, sum pooling, and maximum pooling. The most commonly used pooling is the max pooling, where the maximum value of each block is the corresponding pixel value of the output image. Note that a subsampling layer helps the convolution layer to tolerate rotation and translation among the input images.

2.3.3. Full Connection. The final layer of a CNN is a traditional feedforward network with one or more hidden layers. The output layer uses Softmax activation function:

$$y_i^{(l)} = f\left(z_i^{(l)}\right),$$
$$\text{where } z_i^{(l)} = \sum_{i=1}^{m_i^{(l-1)}} w_{i,j}^{(l)} y_i^{(l-1)}, \tag{12}$$

where $w_{i,j}^{(l)}$ are the weights that should be tuned by the complete fully connected layer in order to form the representation of each class and f is the transfer function which represents the nonlinearity. Note that the nonlinearity in the fully connected layer is built within its neurons, not in separate layers as in convolutions and pooling layers.

After finding output signals, the training of the CNN is started. Training is performed using the stochastic gradient descent algorithm [34]. The algorithm estimates the gradients using a single randomly picked example from the training set. As a result of training, the parameters of CNN are determined.

3. Simulations

In this section, the simulations of the above networks are described. Note that the BPNN and CpNN networks are trained using 620 out of 1000 images, and the rest is used for testing. The CNN is trained using 70% of 120,120 available data, and 30% are used for testing. The input images are of size 32×32 for the sake of reducing computation cost.

3.1. Simulation of Chest Diseases Using BPNN. Backpropagation neural network is based on a supervised learning algorithm, and they are very important and useful in pattern recognition problems [17, 19, 35]. The training of backpropagation networks includes the update of parameters in order to produce good classification results. Hence, in this paper, several experiments were conducted such that significantly accurate results can be obtained. For this aim, different number of hidden neurons, learning rate, and momentum are applied for obtaining better classification result.

The architecture of the designed backpropagation neural network for the image of size 32×32 is described in Figure 5.

Since the backpropagation network uses a supervised learning algorithm, it is, therefore, necessary that the training data could be labelled. The used training data have been labelled according to the 12 classes presented in the classification task. In training stage, different number of hidden neurons, learning rate, and momentum were experimented

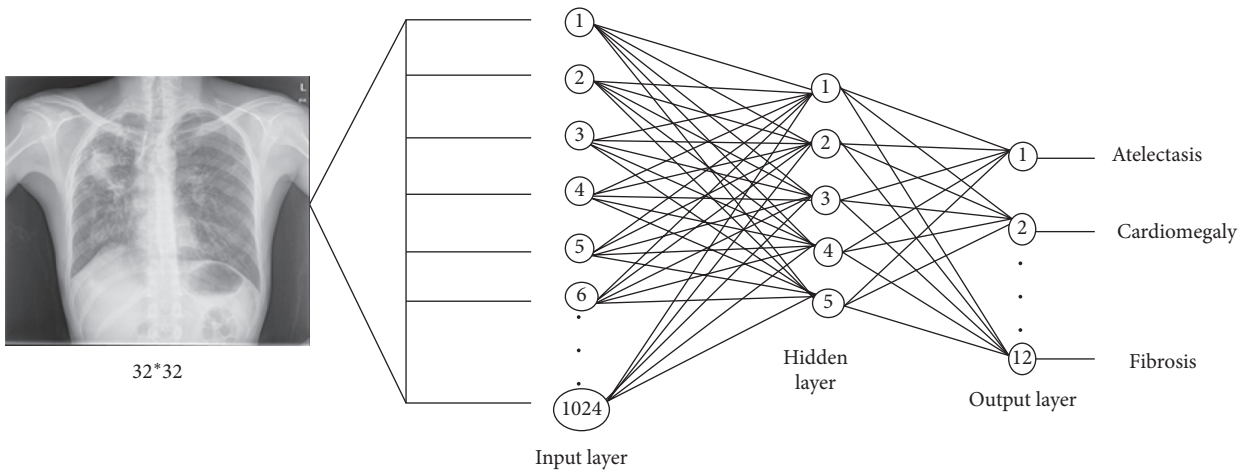


Figure 5: Backpropagation neural network.

Table 1: Training parameters for backpropagation networks (32×32 input pixels).

Networks	BPPN1	BPNN2	BPNN3	BPNN4
Training samples	620	**620**	620	620
Hidden neurons	20	**35**	45	60
Learning rate	0.010	**0.0045**	0.300	0.15
Momentum rate	0.040	**0.0072**	0.0504	0.0619
Activation function	Sigmoid	**Sigmoid**	Sigmoid	Sigmoid
Epochs	1000	**1000**	1256	1374
Training time (sec)	148	**156**	184	193
Mean squared error	0.0077	**0.0025**	0.0056	0.0096

for obtaining better classification result. Table 1 presents the used architectures of BPNN, denoted as BPNN1, BPNN2, BPPN3, and BPNN4.

Since there are 12 classes, 12 neurons have been used in the output layer of the network. The learning curve of BPNN2, which is the network with lowest achieved MSE (Table 1), is shown in Figure 6.

3.2. Simulation of Chest Diseases Using Competitive Neural Network (CpNN). In this section, a competitive neural network using an unsupervised learning algorithm is used for classification of chest diseases. Leveraging on the fact that such networks do not need manual labelling of training data, they save time for the labelling process. Figure 7 shows the architecture of the network used in this paper.

The competitive neural network has two layers designated for the input and output signals. The images are fed as input to the network, and the output neurons learn unique attributes or patterns of the images that differentiates one class from the others. The number of input neurons is 1024 (input image pixels), and the number of output neurons is 12 (number of output classes).

The training parameters of the networks used in this paper are given in Table 2. These competitive networks are trained using 32×32 pixels images. Since the network uses an unsupervised learning algorithm, there is no mean squared error goal to minimize.

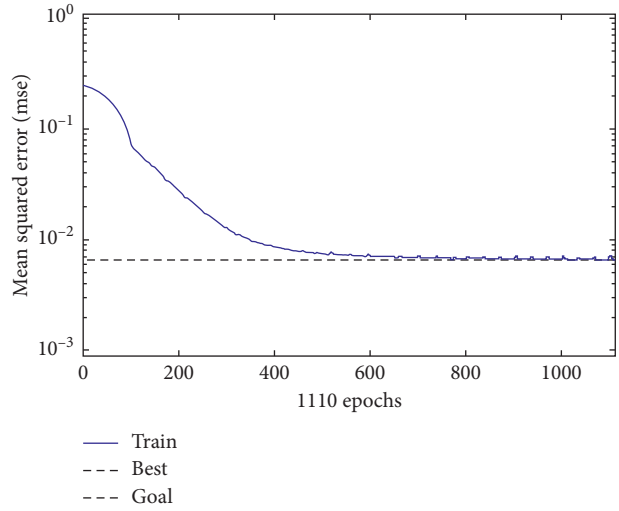


Figure 6: Learning curve for BPNN2.

3.3. Simulation of Chest Diseases Using Convolutional Neural Networks. In this section, the design of the convolutional neural network employed for the chest X-ray medical images are presented. The suitable values of learning parameters of the network are determined through experiments. Note that out of the obtained 120,120 images, 70% are used for training and 30% are used for validating the network.

The input images of the network are of size 32×32. The outputs are 12 classes. The proposed CNN includes 3 hidden layers. Table 3 shows the structure of the CNN and its learning parameters. Here, "Conv" represents a convolution layer, "BN" represents batch normalization, "FM" represents feature maps, and "FC" represents fully connected layer. Note that the filters of size 3×3 are used in all convolution operations with padding, while all pooling operations are performed using max pooling windows of size 2×2.

During simulation, the size of available training data and system specifications for constructing a model were taken into consideration. Thus, dropout training schemes and a batch normalization were employed, and the improvement in model generalization was achieved [24, 25]. Note that a minibatch

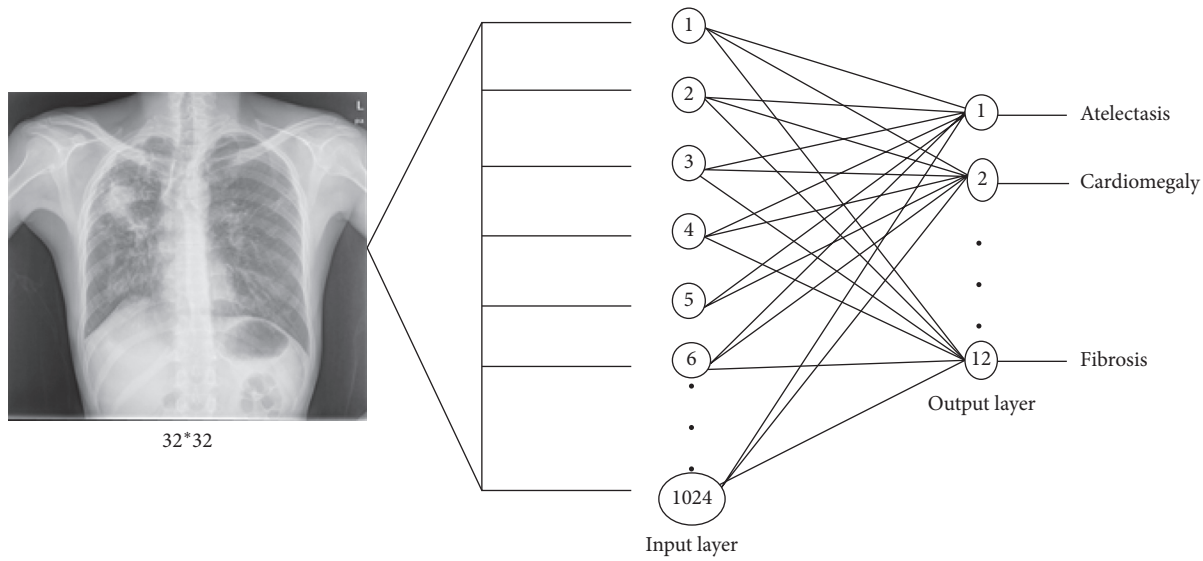


Figure 7: Competitive neural network.

Table 2: Training parameters for competitive neural network (32×32 input pixels).

Networks	CpNN1	CpNN2	CpNN3
Training samples	**620**	620	620
Learning rate	**0.0036**	0.05	0.1
Maximum epochs	**1000**	2000	4000
Training time (sec)	**300 secs**	434 secs	468

optimization of size 100 via stochastic gradient descent is employed [34] for training. In addition, a learning rate of 0.001 and 40,000 iterations are used for training of the CNN model.

The extraction of different levels of features of chest X-ray images in both convolution and pooling layer 1 is given in Figure 8. Figure 8(a) shows the learned filters (or kernels) at convolution layer 1 and Figure 8(b) at the pooling layer of the CNN.

Table 3: CNN training parameters.

Layers	Description	Values
Input layer	Input image	$32 \times 32 \times 1$ images with "zerocenter" normalization
Hidden layer 1	Conv1 + BN + ReLu	16 feature maps of size 10×10
	Pool1	2×2 kernel size with stride of 2
Hidden layer 2	Conv2 + BN + ReLu	32 feature maps of size 10×10
	Pool2	2×2 kernel size with stride of 2
Hidden layer 3	Conv3 + BN + ReLu	64 feature maps of size 10×10
Classification layer	FC	2 fully connected layers
	Softmax	2 units

4. Discussion of Results

The overall performances of the BPNN and CpNN are tested using 380 images. Table 4 shows the recognition rates obtained for the backpropagation networks using 32×32 pixels as the input image size.

It can be seen from the table that all the trained backpropagation neural networks (BPNNs) have different training and testing performances. BPNN2 achieved the highest recognition rate for both training and testing datasets compared to the other networks, that is, 99.19% and 89.57%, respectively.

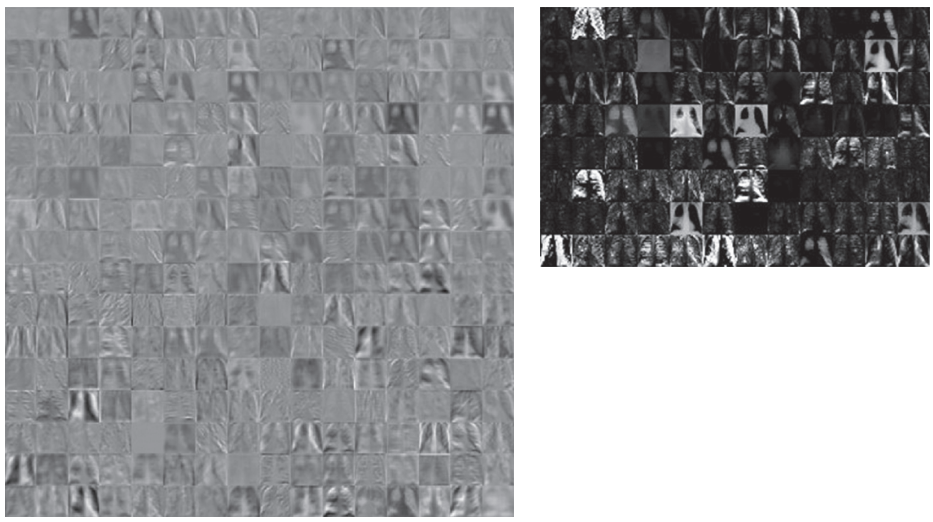
Competitive neural networks that use an unsupervised learning algorithm were also trained and tested using the same images. These networks are faster to train, considering that they have no desired outputs and therefore no error computations and back-pass of error gradients for weights update. The simulation results of the competitive networks using different learning rate and the number of maximum epochs are given in Table 5.

From the table, it can be seen that CpNN2 has the highest recognition rates for both training and test data. Furthermore, it can be seen that CpNN3 has a higher recognition rate than CpNN2 for the training data. Its performance on the test data is lower than CpNN2; that is, it can be stated that CpNN3 has lower generalization power as compared to CpNN2.

Furthermore, the convolutional neural network (CNN) designed for this classification task is also tested using 30% of the available chest X-ray images, and the results are shown in Table 6.

Overall, the performance of the three employed networks in terms of recognition rate, training time, and reached mean square error (MSE) is described in Table 7.

As shown in Tables 4 and 5, the networks behave differently during training and testing, and this is obviously due to the difference in the structures, working principles, and training algorithms of the three employed networks. Also in Table 7, the CNN has achieved the highest recognition rate for training and testing data, compared to other employed



(a) (b)

Figure 8: Learned filters: (a) convolution layer 1 and (b) pooling layer 1.

Table 4: Recognition rates for BPNNs on training and validation data (32×32 pixels).

Network models	Training data (70%)	Validation data (30%)
BPNN1	92.74%	87.42%
BPNN2	**99.19%**	**89.57%**
BPNN3	97.32%	84.36%
BPNN4	98.10%	85.24%

Table 5: Recognition rates for CpNNs using training and validation data (32×32 pixels).

Network models	Training data (70%)	Validation data (30%)
CpNN1	84.21%	81.40%
CpNN2	**85.23%**	**84.71%**
CpNN3	86.57%	76.25%

Table 6: Recognition rates for CNNs on training and validation data (32×32 pixels).

Network model	Training data (70%)	Validation data (30%)
CNN	100%	92.4%

networks. In contrast, this outperformance of CNN over other networks requires longer time and a larger number of learning iterations than that of BPNN2 and CpNN2. Moreover, it can be seen that the three networks have achieved a low MSE, whereas the CNN scored the lowest (0.0013). Furthermore, it is noted that the time needed for the CNN to converge is roughly higher than that of BPNN2 and CpNN2. Consequently, this is due to the depth of the structure of a convolutional neural network, which normally requires a long time, in particular, when the number of inputs is large. Nonetheless, this deep structure is the main factor in achieving a higher recognition rate compared to other networks such as BPNN and CpNN. Lastly, Figure 9 shows an example of the CNN testing paradigm.

Table 7: Performance of the BPNN, CpNN, and CNN.

Network models	Training time	Recognition rate	Reached MSE	Maximum number of iterations
BPNN2	630 secs	80.04%	0.0025	5000
CpNN2	300 secs	89.57%	0.0036	1000
CNN	2500 secs	92.4%	0.0013	40,000

The networks first take a chest X-ray as an input and output the probabilities of the classes.

A comparison of the developed networks with some earlier works is shown in Table 8. Firstly, it is seen that shallow (traditional) networks (BPNN and CpNN) could not achieve high recognition rates compared to other deep networks, which is obviously due to their deficiency in extracting the important features from input images. Moreover, it is noticed that the proposed deep convolutional neural network (CNN) achieved a higher recognition rate than other earlier research work such as CNN with GIST features [36]. The transfer learning-based networks are also used for chest X-rays classification such as VGG16 [37] and VGG19 [37]. They have gained lower generalization capabilities compared to the proposed network. These pretrained models [37] have very powerful features extraction capabilities since they were trained using a huge database, Image Net [38]. Note that, we compared the researches that provided explicitly achieved accuracies. The obtained results can show that applying deep CNNs to the problem of chest X-ray diseases is promising in a way that similar or confusing diseases could be correctly classified with good recognition rates.

5. Conclusion

In this paper, convolutional neural network (CNN) is designed for diagnosis of chest diseases. For comparative analysis, backpropagation neural network (BPNN) and

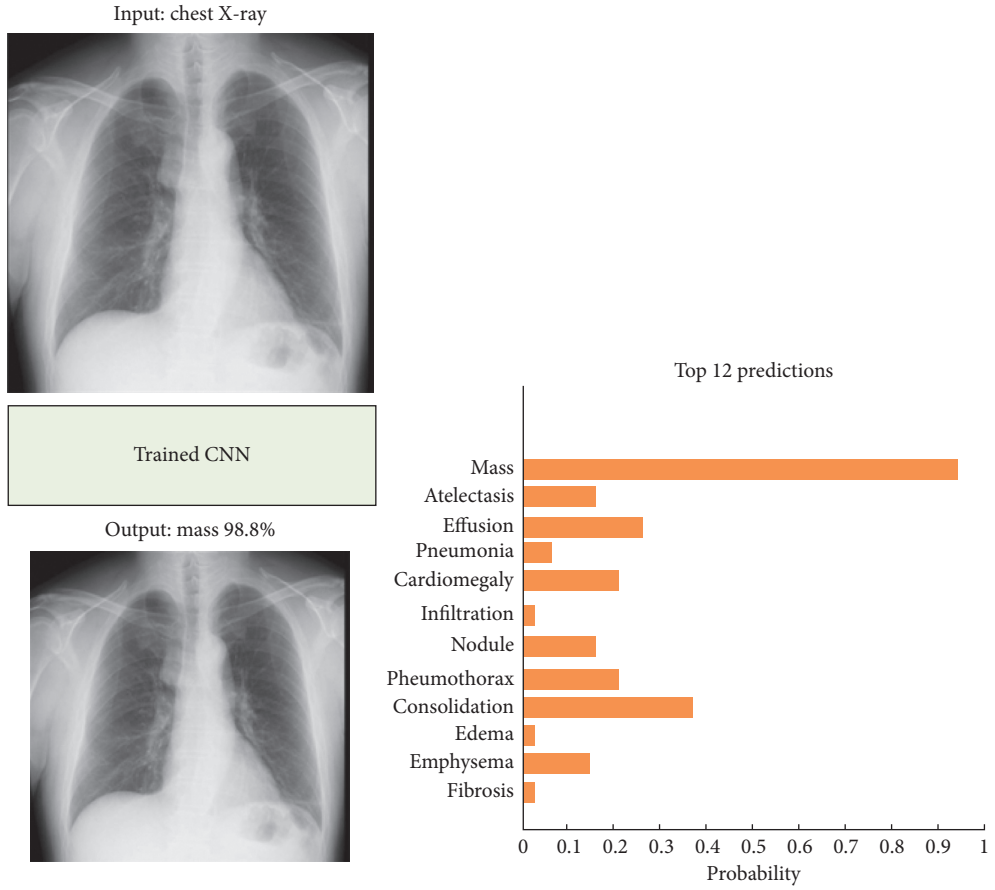


Figure 9: CNN final classification of chest X-rays with classes probabilities.

Table 8: Results comparison with earlier works.

Parameters	CNN	BPNN2	CpNN2	CNN with GIST [36]	VGG16 [37]	VGG19 [37]
Number of images	120,120	1000	1000	637	8100	8100
Accuracy	92.4%	80.04%	89.57%	92%	86%	92%

competitive neural network (CpNN) are carried out for the classification of the chest X-ray diseases. The designed CNN, BPNN, and CpNN were trained and tested using the chest X-ray images containing different diseases. Several experiments were carried out through training of these networks using different learning parameters and a number of iterations. In both backpropagation and competitive networks, it was observed that the input image of size 32×32 pixels showed good performance and achieved high recognition rates. Based on recognition rates, the backpropagation networks outperformed the competitive networks. Moreover, the competitive networks did not require manual labelling of training data as it was carried out for the backpropagation network. Furthermore, a CNN was also trained and tested using a larger dataset which was also used for training and testing of BPNN and CpNN. After convergence, it was noticed that the CNN was capable of gaining a better generalization power than that achieved by BPNN and CpNN, although required computation time and the number of iterations were roughly higher. This outperformance is mainly due to the deep structure of CNN that uses the power of extracting different level features, which resulted in a better generalization capability. The simulation result of proposed CNN is also compared with other deep CNN models such as GIST, VGG16, and VGG19. These networks have lower generalization capabilities and accuracies compared to the proposed network. The obtained results have demonstrated the high recognition rates of the proposed CNN.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] O. Er, N. Yumusak, and F. Temurtas, "Chest diseases diagnosis using artificial neural networks," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7648–7655, 2010.

[2] A. A. El-Solh, C.-B. Hsiao, S. Goodnough, J. Serghani, and B. J. B. Grant, "Predicting active pulmonary tuberculosis using an artificial neural network," *Chest*, vol. 116, no. 4, pp. 968–973, 1999.

[3] K. Ashizawa, T. Ishida, H. MacMahon, C. J. Vyborny, S. Katsuragawa, and K. Doi, "Artificial neural networks in chest radiography: application to the differential diagnosis of interstitial lung disease," *Academic Radiology*, vol. 11, no. 1, pp. 29–37, 2005.

[4] A. M. Dos Santos, B. B. Pereira, and J. M. de Seixas, "Neural networks: an application for predicting smear negative pulmonary tuberculosis," in *Proceedings of the Statistics in the Health Sciences*, Liège, Belgium, 2004.

[5] U. Avni, H. Greenspan, E. Konen, M. Sharon, and J. Goldberger, "X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words," in *Proceedings of IEEE Transactions on Medical Imaging*, Orlando, FL, USA, 2011.

[6] S. Jaeger, A. Karargyris, and S. Candemir, "Automatic tuberculosis screening using chest radiographs," in *Proceedings of IEEE Transactions on Medical Imaging*, London, UK, 2014.

[7] P. Pattrapisetwong and W. Chiracharit, "Automatic lung segmentation in chest radiographs using shadow filter and multilevel thresholding," in *Proceedings of 2016 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, Manchester, UK, 2016.

[8] O. Er, C. Sertkaya, F. Temurtas, and A. C. Tanrikulu, "A comparative study on chronic obstructive pulmonary and pneumonia diseases diagnosis using neural networks and artificial immune system," *Journal of Medical Systems*, vol. 33, no. 6, pp. 485–492, 2009.

[9] S. Khobragade, A. Tiwari, C. Y. Pati, and V. Narke, "Automatic detection of major lung diseases u sing chest radiographs and classification by feed-forward artificial neural network," in *Proceedings of 1st IEEE International Conference on Power Electronics. Intelligent Control and Energy Systems (ICPEICES-2016) 2016 IEEE*, pp. 1–5, Delhi, India, 2016.

[10] G. Litjens, T. Kooi, E. B. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[11] S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci, and N. Navab, "Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1313–1321, 2016.

[12] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[13] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in *Proceedings of Advances in Neural Information Processing Systems*, pp. 153–160, Vancouver, BC, Canada, 2007.

[14] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI," *Medical Image Analysis*, vol. 30, pp. 108–119, 2016.

[15] H.-C. Shin, K. Roberts, L. Lu, D. Demner-Fushman, J. Yao, and R. M. Summers, "Learning to read chest X-rays: recurrent neural cascade model for automated image annotation," Cornel University library, 2016, https://arxiv.org/abs/1603.08486.

[16] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chest X-ray 8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proceedings of IEEE CVPR 2017*, Honolulu, HI, USA, 2017.

[17] R. H. Abiyev and K. Altunkaya, "Neural network based biometric personal identification with fast iris segmentation," *International Journal of Control, Automation and Systems*, vol. 7, no. 1, pp. 17–23, 2009.

[18] A. Helwan and D. P. Tantua, "IKRAI: intelligent knee rheumatoid arthritis identification," *International Journal of Intelligent Systems and Applications*, vol. 8, no. 1, p. 18, 2016.

[19] A. Helwan and R. H. Abiyev, "Shape and texture features for identification of breast cancer," in *Proceedings of International Conference on Computational Biology 2016*, San Francisco, CA, USA, October 2016.

[20] M. Cilimkovic, *Neural Networks and Back Propagation Algorithm*, Institute of Technology Blanchardstown, Dublin, Ireland, 2015.

[21] A. Helwan, D. U. Ozsahin, R. Abiyev, and J. Bush, "One-year survival prediction of myocardial infarction," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 6, pp. 173–178, 2017.

[22] M. K. S. Ma'aitah, R. Abiyev, and I. J. Bus, "Intelligent classification of liver disorder using fuzzy neural system," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 12, pp. 25–31, 2017.

[23] M. A. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 13, no. 5, pp. 815–826, 1983.

[24] A. R. Abas, "Adaptive competitive learning neural networks," *Egyptian Informatics Journal*, vol. 14, no. 3, pp. 183–194, 2013.

[25] G. A. Barreto, J. C. M. Mota, L. G. M. Souza et al., "Competitive neural networks for fault detection and diagnosis in 3G cellular systems," in *Lecture Notes in Computer Science*, Springer, Berlin, Germany, 2004.

[26] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of Advances in Neural Information Processing Systems*, pp. 1097–1105, Lake Tahoe, Nevada, USA, 2012.

[28] F. Gao, Z. Yue, J. Wang, J. Sun, E. Yang, and H. Zhou, "A novel active semisupervised convolutional neural network algorithm for sar image recognition," *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 3105053, 8 pages, 2017.

[29] A. Rios and R. Kavuluru, "Convolutional neural networks for biomedical text classification: application in indexing biomedical articles," in *Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics*, pp. 258–267, ACM, Atlanta, GA, USA, 2015.

[30] J. Li, J. Feng, and C.-C. Jay Kuo, "Deep convolutional neural network for latent fingerprint enhancement," *Signal Processing: Image Communication*, vol. 60, pp. 52–63, 2018.

[31] J. Bouvrie, *Notes on Convolutional Neural Networks*, http://cogprints.org/5869/1/cnn_tutorial.pdf.

[32] S. Hussain, S. Muhammad Anwar, and M. Majid, "Segmentation of glioma tumors in brain using deep convolutional neural network," *Neurocomputing*, vol. 282, pp. 248–261, 2018.

[33] J. Gu, Z. Wang, J. Kuen, L. Ma et al., "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, 2018.

[34] R. G. J. Wijnhoven and P. H. N. de With, "Fast training of object detection using stochastic gradient descent," in *Proceedings of International Conference on Pattern Recognition (ICPR)*, pp. 424–427, Tsukuba, Japan, 2010.

[35] R. Abiyev and K. Altunkaya, "Iris recognition for biometric personal identification using neural networks," in *Lecture Notes in Computer Sciences*, Springer-Verlag, Berlin, Germany, 2007.

[36] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology detection using deep learning with non-medical training," in *Proceedings of Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium*, pp. 294–297, Brookly, NY, USA, 2015.

[37] M. T. Islam, M. A. Aowal, A. T. Minhaz, and K. Ashraf, "Abnormality detection and localization in chest x-rays using deep convolutional neural networks," 2017, https://arxiv.org/abs/1705.09850.

[38] O. Russakovsky, J. Deng, H. Su et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

Hindawi
Journal of Healthcare Engineering
Volume 2018, Article ID 1753480, 9 pages
https://doi.org/10.1155/2018/1753480

Research Article

Classification of Computed Tomography Images in Different Slice Positions Using Deep Learning

Hiroyuki Sugimori

Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan

Correspondence should be addressed to Hiroyuki Sugimori; sugimori@hs.hokudai.ac.jp

Received 21 March 2018; Accepted 22 May 2018; Published 16 July 2018

Academic Editor: Zahid Akhtar

This study aimed at elucidating the relationship between the number of computed tomography (CT) images, including data concerning the accuracy of models and contrast enhancement for classifying the images. We enrolled 1539 patients who underwent contrast or noncontrast CT imaging, followed by dividing the CT imaging dataset for creating classification models into 10 classes for brain, neck, chest, abdomen, and pelvis with contrast-enhanced and plain imaging. The number of images prepared in each class were 100, 500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, and 10,000. Accordingly, the names of datasets were defined as 0.1K, 0.5K, 1K, 2K, 3K, 4K, 5K, 6K, 7K, 8K, 9K, and 10K, respectively. We subsequently created and evaluated the models and compared the convolutional neural network (CNN) architecture between AlexNet and GoogLeNet. The time required for training models of AlexNet was lesser than that for GoogLeNet. The best overall accuracy for the classification of 10 classes was 0.721 with the 10K dataset of GoogLeNet. Furthermore, the best overall accuracy for the classification of the slice position without contrast media was 0.862 with the 2K dataset of AlexNet.

1. Introduction

In the field of computer vision, deep learning with a convolutional neural network (CNN) [1] can be used to attain precise general image classification. Recently, deep learning has been increasingly used in medical imaging [2–17]. Arguably, deep learning has several potential abilities, including object detection [9, 10] and image segmentation. Typically, medical images differ from general images in that medical images only depict human structures with no background structure other than that of a human body. Previously, some studies have reported the classification of the scan position with computed tomography (CT) imaging using deep learning [18, 19]; however, it evaluated the accuracy of classification only when detecting scan slice positions. Thus, research has recognized the necessity of the enhancement information using contrast media for tumor diagnosis [20, 21] because radiologists commonly refer to the slice position, organ structure status, and presence of an organ or tumor enhancement. Further, information concerning the position and presence of contrast media is one of the critical factors for the basic requirement for automatic diagnosis using deep learning. Fundamentally, deep learning requires several images [22] to create classification models, although human structures comprise variable organ structures with different sizes in each subject. Moreover, related works with deep learning in CT images have recently been reported dealing with detections of anatomies and tumors [23–26]. In these techniques, whole body CT images could not be used for deep learning because those images have to be classified as concerning region in advance. However, till date, no study has reported how many images are required for the classification of CT images, including contrast enhancement data. If the precise classification of CT images has done as a preprocessing, the automatic diagnosis using deep learning for whole body images will be more practical technique. Thus, this study aimed at elucidating the relationship between the number of CT images, including data

Table 1: Names of datasets and the number of images in each label.

	Names of datasets												
	0.1K	0.5K	1K	2K	3K	4K	5K	6K	7K	8K	9K	10K	Testing dataset
Brain (P)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Brain (CE)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Neck (P)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Neck (CE)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Chest (P)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Chest (CE)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Abdomen (P)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Abdomen (CE)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Pelvis (P)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Pelvis (CE)	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	100
Total number of images	1000	5000	10,000	20,000	30,000	40,000	50,000	60,000	70,000	80,000	90,000	100,000	1000
For training	900	4500	9000	18,000	27,000	36,000	45,000	54,000	63,000	72,000	81,000	90,000	-
For validation	100	500	1000	2000	3000	4000	5000	6000	7000	8000	9000	10,000	-

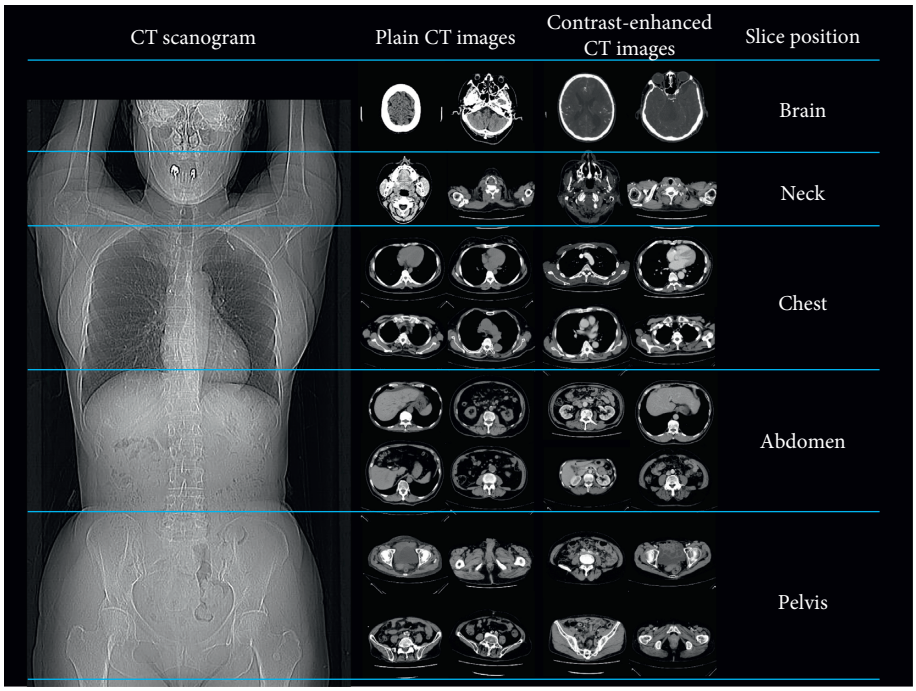


Figure 1: Range of the slice location and example image of 10 classes: the brain, neck, chest, abdomen, and pelvis.

concerning the accuracy of models and contrast enhancement for creating classification models.

2. Materials and Methods

2.1. Subjects. Totally, 1539 patients (males, 815; females, 724; mean age ± standard deviation (SD), 59.9 ± 18.5 years) who underwent contrast or noncontrast CT imaging of the brain, neck, chest, abdomen, or pelvis in January 2016 were included in the study. The study protocol was approved by the Ethics Committee of the Hokkaido University Hospital (Sapporo, Japan).

2.2. Datasets. The dataset of CT images for creating classification models was divided into 10 classes for the brain, neck, chest, abdomen, and pelvis with contrast-enhanced (CE) and non-contrast-enhanced CT, which was defined as plain (P). The number of images prepared for each class were 100, 500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, and 10,000; the datasets were named 0.1K, 0.5K, 1K, 2K, 3K, 4K, 5K, 6K, 7K, 8K, 9K, and 10K. We used these images of 90% of the data for training and 10% for validation for creating classification models. In addition, 1000 images from each class other than the datasets mentioned above were prepared for testing models. Table 1 presents the name of all such datasets with complete details.

The image ranges of each class were defined as follows: brain, slice from the anterior tip of the parietal bone to the foramen magnum; neck, slice from the foramen magnum to the pulmonary apex; chest, slice from the pulmonary apex to the diaphragm; abdomen, slice from the diaphragm to the

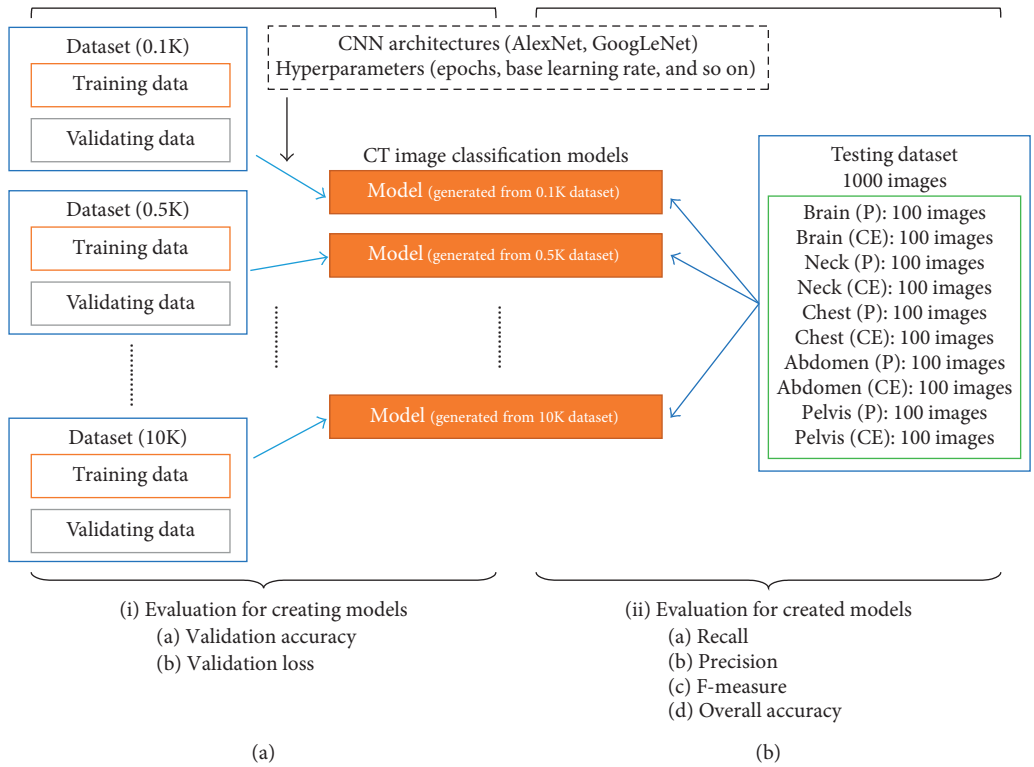


Figure 2: Schematic of the datasets, creating models, and validation of creating models. (a) Workflow of creating models and (b) workflow of the evaluation of created models. (i) Evaluation points for creating models and (ii) evaluation points for the created models.

top of an iliac crest; and pelvis, slice from the top of an iliac crest to the distal end of the ischium. Figure 1 indicates the image ranges for each class.

Furthermore, CE was defined as the state of intravascular injection of contrast media in the examination. We did not consider the timing of scans from the injection. Criteria for exclusion from the dataset were CT images with excessive magnification, the reconstruction kernel of the bone or lung, nothing (above the anterior tip of the parietal bone), and arms or legs only.

2.3. Preprocessing for Creating Models. We retrieved CT images from the Picture Archiving and Communication System. Next, to convert images for the training database, we converted these CT images from the digital imaging and communications in medicine (DICOM) to the joint photographic experts group (JPEG) file format using dedicated DICOM software (XTREK View; J-Mac System Inc., Sapporo, Japan). In addition, the window width and level of the DICOM image were used with preset values in the DICOM-tag. Next, we converted all JPEG images to grayscale 8-bit images sized 512×512 pixels, followed by sorting the converted JPEG files into particular folders according to image classes. Furthermore, we used the NVIDIA Deep Learning GPU Training System (NVIDIA DIGITS; NVIDIA Corporation, Santa Clara, CA), the conversion software of the training database, and an authoring software for deep learning. Finally, the database type was set to the lightning memory-mapped database (LMDB).

Table 2: Hyperparameters of training models.

	CNN architecture	
	AlexNet	GoogLeNet
Training epochs	30	
Snapshot interval	10	
Validation interval	1	
Random seed	None	
Batch size of training	128	32
Batch size of validation	32	16
Solver type	Stochastic gradient descent (SGD)	
Base learning rate	0.01	
Policy	Step down	
Step size (%)	33	
Gamma	0.1	

2.4. Training for Creating Models. Figure 2(a) outlines the training for creating models. We used the authoring software NVIDIA DIGITS for deep learning, a deep learning optimized machine with two GTX1080 Ti GPUs with 11.34 TFlops single precision, 484 GB/s memory bandwidth, and 11 GB memory per board. The convolutional architecture for fast feature embedding (Caffe) [27] constituted the deep learning framework, which was worked on the NVIDIA DIGITS. Here, we compared CNN architectures, which could be selected on the NVIDIA DIGITS, between the 16-layer AlexNet [28] and 22-layer GoogLeNet [29]. The training model hyperparameters were used as a default on the software (Table 2), and the maximum training epoch was set to 30. The initial learning rate was set at 0.01; it was later dropped by one-tenth following every 10 epochs of training.

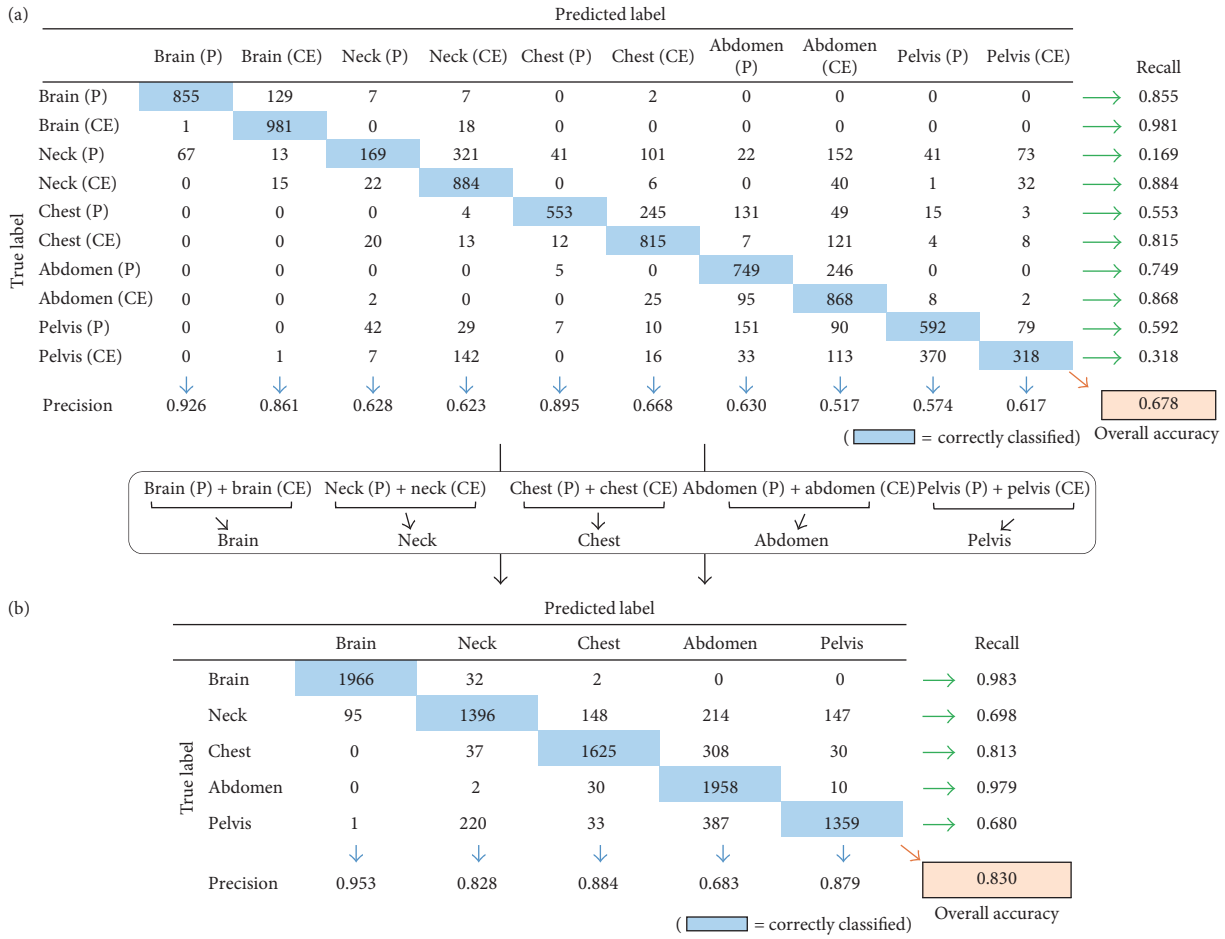


FIGURE 3: Computing different measures from the confusion matrix. The precision and recall of each class and the overall accuracy were calculated. (a) Confusion matrix with 10 classes and (b) confusion matrix of each slice location, calculated from (a).

In both CNN architectures, training was acquired three times for each dataset, followed by recording the best accuracy and loss of validation and calculating the mean value. We evaluated these results using datasets and CNN architectures, respectively. Moreover, the duration from the start of training to the complete creation of the model was assessed in each dataset.

2.5. Evaluation of Created Models. Figure 2(b) outlines the training for the evaluation of created models. The confusion matrix obtained by an independent dataset was intuitively a fair indicator of the performance of the created models because training accuracy was the only result that was repeatedly evaluated using the same dataset. Figure 3 shows examples of the confusion matrix. To evaluate the created models, we applied the dataset for the testing model (Section 2.2). For training with 10 classes, which presented as a 10×10 table, all performance measures were based on hundred numbers obtained by applying the classifier to the test dataset. Moreover, the confusion matrix comprised columns and rows corresponding to the predicted and the true image label, respectively. The same position was integrated for evaluating the positional detection ability, for example, brain (P) and brain (CE) as to the brain; the confusion matrix was created as a 5×5 table.

Here, the confusion matrix generated four parameters: precision, recall, F-measure, and overall accuracy. Precision was presented as a ratio of how many images were correctly predicted to produce the predicted labels; recall as a ratio presented how many of these were correctly classified. Furthermore, we defined F-measure as the harmonic mean of precision and recall:

$$\text{F-measure} = \frac{(2 \times \text{recall} \times \text{precision})}{(\text{recall} + \text{precision})}. \quad (1)$$

Overall accuracy was presented as the ratio of the number of correctly classified images in all test images. Furthermore, we evaluated the confusion matrix thrice based on the number of created models.

2.6. Statistical Analysis. Precision, recall, F-measure, and overall accuracy were presented as mean ± SD regardless of

Table 3: Accuracy and loss of validation for training models.

	CNN architecture	Dataset											
		0.1K	0.5K	1K	2K	3K	4K	5K	6K	7K	8K	9K	10K
Accuracy of validation	AlexNet	0.27	0.77	0.87	0.92	0.93	0.95	0.95	0.96	0.97	0.97	0.97	0.97
	GoogLeNet	0.40	0.78	0.93	0.96	0.96	0.97	0.98	0.97	0.97	0.98	0.98	0.98
Loss of validation	AlexNet	2.00	0.57	0.36	0.21	0.21	0.15	0.14	0.12	0.11	0.11	0.10	0.09
	GoogLeNet	1.59	0.55	0.24	0.12	0.12	0.10	0.09	0.10	0.09	0.08	0.08	0.07
Time for training (min)	AlexNet	0.8	2.1	3.8	7.3	10.7	14.2	17.4	21.3	24.7	27.9	31.8	34.8
	GoogLeNet	1.1	4.9	9.5	18.7	28.9	38.5	48.3	45.7	67.6	76.5	86.4	95.0

Table 4: Recall, precision, and F-measure for each class in AlexNet and GoogLeNet.

	Recall		Precision		F-measure	
	AlexNet	GoogLeNet	AlexNet	GoogLeNet	AlexNet	GoogLeNet
Brain (P)	0.71	0.77	0.81	0.80	0.74	0.78
Brain (CE)	0.77	0.79	0.72	0.70	0.73	0.73
Neck (P)	0.20	0.17	0.46	0.46	0.25	0.22
Neck (CE)	0.53	0.57	0.46	0.55	0.46	0.54
Chest (P)	0.44	0.49	0.68	0.67	0.52	0.55
Chest (CE)	0.71	0.67	0.63	0.62	0.65	0.64
Abdomen (P)	0.61	0.69	0.57	0.54	0.56	0.59
Abdomen (CE)	0.73	0.74	0.48	0.47	0.56	0.57
Pelvis (P)	0.52	0.52	0.48	0.53	0.48	0.51
Pelvis (CE)	0.42	0.40	0.52	0.60	0.44	0.47

Table 5: Recall, precision, F-measure, and overall accuracy for each dataset in AlexNet and GoogLeNet.

	CNN architecture	Dataset											
		0.1K	0.5K	1K	2K	3K	4K	5K	6K	7K	8K	9K	10K
Recall	AlexNet	0.20	0.39	0.52	0.53	0.61	0.67	0.61	0.67	0.67	0.65	0.64	0.62
	GoogLeNet	0.15	0.38	0.57	0.62	0.65	0.68	0.61	0.65	0.69	0.62	0.67	0.69
Precision	AlexNet	0.20	0.39	0.53	0.54	0.61	0.67	0.65	0.69	0.70	0.68	0.67	0.64
	GoogLeNet	0.15	0.41	0.58	0.62	0.65	0.68	0.63	0.66	0.71	0.64	0.70	0.70
F-measure	AlexNet	0.14	0.35	0.49	0.50	0.60	0.65	0.59	0.65	0.66	0.63	0.62	0.59
	GoogLeNet	0.11	0.35	0.54	0.61	0.64	0.67	0.59	0.63	0.67	0.60	0.65	0.68
Overall accuracy	AlexNet	0.20	0.39	0.52	0.53	0.61	0.66	0.61	0.67	0.67	0.65	0.63	0.62
	GoogLeNet	0.15	0.38	0.57	0.62	0.65	0.68	0.61	0.65	0.69	0.62	0.67	0.69

the dataset. In addition, the accuracy of validation for training, recall, precision, F-measure, and overall accuracy were evaluated at mean values higher than 0.80, 0.85, 0.90, or 0.95, and the loss of validation for training was evaluated at mean values less than 0.20, 0.15, 0.10, or 0.05. The time for training the model was evaluated as the mean time. The best overall accuracy was recorded from all datasets. Besides, the comparison of CNN architecture irrespective of the dataset was evaluated using the Mann–Whitney U test. Furthermore, we used the Steel–Dwass test for the multiple comparison of the slice position for the recall, precision, and F-measure and the comparison of the CNN architecture irrespective of the dataset. $P < 0.05$ was considered statistically significant.

3. Results and Discussion

3.1. Results

3.1.1. Evaluation of Training of Models. Table 3 presents the accuracy of validation, loss of validation, and time taken for training the model. Datasets from 4K to 10K on AlexNet and 2K to 10K on GoogLeNet had an accuracy of validation of >0.95. In addition, datasets of 10K on AlexNet and from 5K to 10K, except for 6K, on GoogLeNet had loss of validation of <0.05. The mean accuracy of validation for AlexNet and GoogLeNet was 0.87 ± 0.19 and 0.90 ± 0.16, respectively, and a significant difference was observed between CNN architectures ($P = 0.0027$). The mean loss of validation for AlexNet and GoogLeNet was 0.35 ± 0.52 and 0.27 ± 0.43, respectively, and a significant difference was observed between CNN architectures ($P = 0.0036$). Furthermore, the mean time of training model for AlexNet and GoogLeNet was 16.4 ± 11.45 minutes and 43.42 ± 31.78 minutes, respectively, and a significant difference was observed between CNN architectures ($P = 0.0003$).

3.1.2. Evaluation of Created Models

(1) Comparison of All 10 Classes. Table 4 presents the recall, precision, and F-measure for each class for both AlexNet and

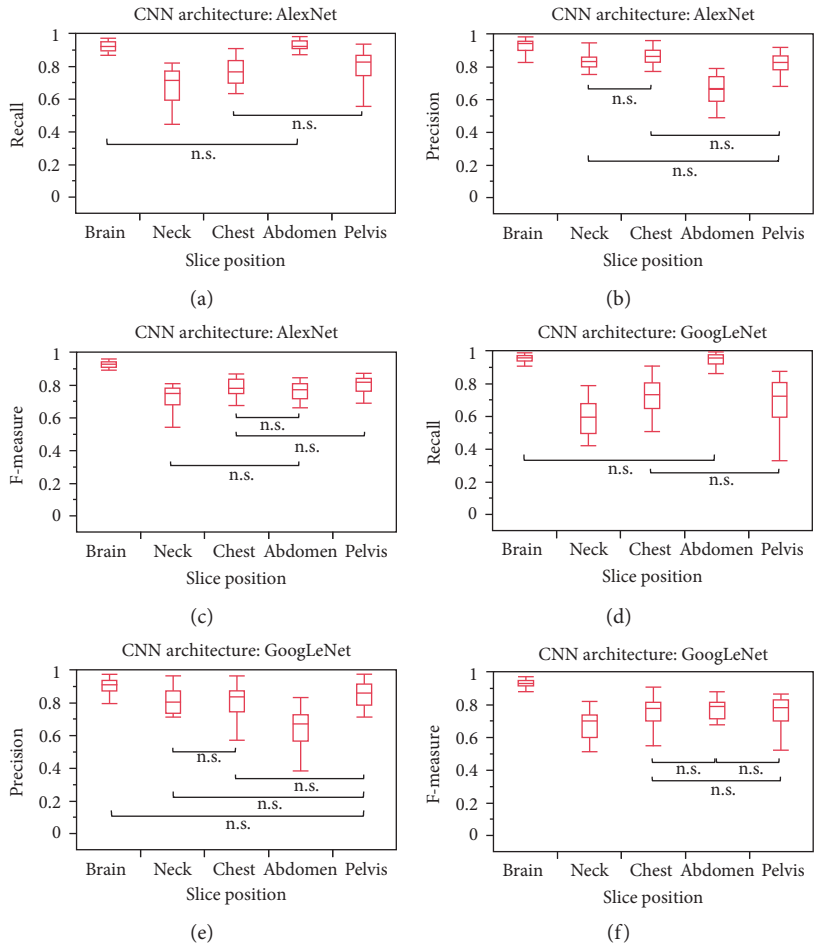


Figure 4: (a) Recall, (b) precision, and (c) F-measure for each slice position with AlexNet. (d) Recall, (e) precision, and (f) F-measure for each slice position with GoogLeNet (without specific marks indicates significant differences between each group; *n.s., no significant differences between the two groups, Steel–Dwass test).

GoogLeNet. The precision of the recall for brain (P) for AlexNet and GoogLeNet was >0.80; however, no parameters of the recall and F-measure were >0.80. Table 5 presents the recall, precision, F-measure, and overall accuracy for each dataset of AlexNet and GoogLeNet. No parameters in all datasets were >0.80. The mean recall for AlexNet and GoogLeNet was 0.56 ± 0.28 and 0.58 ± 0.28, respectively, and no significant difference was observed between CNN architectures ($P = 0.4052$). The mean precision for AlexNet and GoogLeNet was 0.58 ± 0.23 and 0.59 ± 0.22, respectively, and no significant difference was observed between CNN architectures ($P = 0.2496$). The mean F-measure for AlexNet and GoogLeNet was 0.54 ± 0.24 and 0.56 ± 0.24, respectively, and no significant difference was observed between CNN architectures ($P = 0.1918$). In addition, the mean overall accuracy for AlexNet and GoogLeNet was 0.56 ± 0.14 and 0.58 ± 0.16, respectively, and no significant difference was observed between CNN architectures ($P = 0.2601$). Furthermore, the best overall accuracy for the classification of 10 classes was 0.721, which was obtained with dataset #3 of 10K of GoogLeNet.

(2) Comparison of Each Slice Position. Figure 4 indicates the recall, precision, and F-measure for each slice position for AlexNet and GoogLeNet. The mean recall for the brain in GoogLeNet and the abdomen in AlexNet and GoogLeNet was >0.90 and that for the brain in AlexNet was >0.85. The mean precision for the brain in AlexNet and GoogLeNet was >0.85 and that for the neck and chest in AlexNet and the pelvis in AlexNet and GoogLeNet was >0.80. In addition, the mean F-measure for the brain in AlexNet and GoogLeNet was >0.85. Table 6 presents the recall, precision, F-measure, and overall accuracy for each dataset in AlexNet and GoogLeNet. Furthermore, the best overall accuracy for the classification of the slice position was 0.862, which was obtained with dataset #3 of 2K of AlexNet.

3.2. Discussion. In this study, a total of 12 datasets for the training models were compared for evaluating models. Regarding the training model, datasets with >4000 images/class on AlexNet and >2000 images/class on GoogLeNet could obtain an accuracy >0.95. Regarding validation loss, a higher number of images per slice were essential to reduce the loss. In addition, the training time in the CNN architecture of GoogLeNet was approximately double than that of AlexNet. The significant difference in training models

Table 6: Recall, precision, F-measure, and overall accuracy for each dataset in AlexNet and GoogLeNet.

	CNN architecture	Dataset											
		0.1K	0.5K	1K	2K	3K	4K	5K	6K	7K	8K	9K	10K
Recall	AlexNet	0.39	0.77	0.84	0.82	0.83	0.85	0.81	0.84	0.83	0.84	0.81	0.83
	GoogLeNet	0.34	0.72	0.78	0.79	0.82	0.82	0.75	0.79	0.81	0.75	0.81	0.83
Precision	AlexNet	0.43	0.79	0.85	0.84	0.84	0.86	0.82	0.85	0.86	0.85	0.84	0.83
	GoogLeNet	0.33	0.80	0.81	0.80	0.82	0.83	0.78	0.81	0.83	0.78	0.82	0.84
F-measure	AlexNet	0.32	0.77	0.83	0.81	0.83	0.85	0.80	0.84	0.83	0.84	0.82	0.83
	GoogLeNet	0.30	0.71	0.78	0.79	0.81	0.82	0.74	0.79	0.81	0.74	0.80	0.83
Overall accuracy	AlexNet	0.39	0.77	0.84	0.82	0.83	0.85	0.81	0.84	0.83	0.84	0.81	0.83
	GoogLeNet	0.34	0.72	0.78	0.79	0.82	0.82	0.75	0.79	0.81	0.75	0.81	0.83

between AlexNet and GoogLeNet occurred because the CNN structures of GoogLeNet had several layers in detail. Regarding the evaluation of created models with 10 classes, CT images of the brain (P) and brain (CE) exhibited higher recall and precision because the anatomy and size of the brain were different from that of the other parts of the body, indicating that the F-measure value was higher in the brain than in other organs. However, other classes of the recall and precision exhibited lower classes of the brain; in fact, the recall of the neck (P) was particularly lower than that of the others, perhaps, because the CT images of the neck were affected by the artifact from the metal of the artificial tooth. Reportedly, the recall of the classification could be lower because these artifacts from the tooth, which rendered differentiating the enhancement by contrast media difficult, affected images in the form of noise [30]. Typically, the difference between P and CE images could only be reflected in the vascular because of the few structures of the strong enhancement by contrast media in the neck. In addition, the lower recall of the neck and pelvis could be attributed to the similar symmetric structure of the iliac bone and scapula on the backside of the body. Precisely, the lower recall of the neck and pelvis was caused not only by the difference in the enhancement by contrast media but also the structure. The calcification of the iliac artery was one of the causes of confusion between P and CE images, because the CT value was higher at the calcification of the artery, which resembled the enhanced vascular. Reportedly, the calcification of the iliac artery, which is associated with some diseases [31, 32], was one of the causes of confusion between P and CE images, because the calcification of the artery was depicted like an enhanced vascular. Regarding the evaluation of the created models for each slice position, CT images of the brain exhibited higher recall and precision for the same reason as that for the 10 classes. As the recall of the abdomen was higher, it leaves no bone structure, except for the spine, in the range of the abdomen in this study. Of note, the lower recall of the neck and pelvis was formed by the similar symmetric structure because the contrast of the bone was primarily characterized as a feature of CT images. In addition, the features over the two regions were unclear because the borders between the neck and chest and between the chest and abdomen were overlapped in each position. In case of removing the overlapped range of images from the dataset, the recall and precision would probably be higher; this assumption was supported by the fact that the dataset of 1K had already exhibited an overall accuracy of approximately 0.8. Previously, some studies have reported the classification of CT images [18, 19]. However, the number of CT images in each class was one-tenth compared with the datasets in this study. Moreover, the test dataset for the validation of created models was 10,000 images of 1000 images per class; in this respect, this study differed from others. Regarding the CNN architecture, the time for the training model of AlexNet was faster than that for GoogLeNet, because the larger number of layers and complicated CNN structure of GoogLeNet affected the time for the training model. In addition, no significant differences existed between both of them regarding the evaluation of created models. Thus, AlexNet would be useful for being rapid and simple. This study has some limitations. First, we used the default CNN architecture and hyperparameters. Because the classification of general images with deep learning was sophisticated [28, 29], the major change in the CNN architecture for medical images might be unstable for the optimization of creating models. Second, the created models were evaluated using the hold-out method which was known as the simplest kind of cross validation. With regard to the cross validation, we focused on the simple tendency for the effect of the size of dataset though the k-fold cross validation has been reported [33]. The reason of choice for the hold-out method was that the small number of dataset had not enough the number of images because the k-fold method has to divide the data into k-equally sized folds. However, we could present the basic tendency for the size of the dataset in this study. Third, all images of the dataset in this study were original CT images. Thus, giving a variety to the feature of the dataset was necessary for this study, although the data augmentation [19] was critical for not only the increased number of images but also the repeated usage of similar images for the training. However, the number of patients in this study was approximately 1539 with >100,000 images; thus, we believe that the results of this study could serve as a reference for further investigation in the future.

4. Conclusions

This study elucidates the relationship between the number of CT images, including data about contrast enhancement for creating classification models, and accuracy of models. The time for training models of AlexNet was faster than that for

GoogLeNet. Furthermore, the best overall accuracy for the classification of 10 classes was 0.721 with the dataset of 10K of GoogLeNet, and the best overall accuracy for the classification of the slice position regardless of contrast media was 0.862 with the dataset of 2K of AlexNet.

Data Availability

The author has no permission for providing the created models and datasets currently. In order to provide data to a third-party organization in any way, after acquiring consent from the patients and preparing an experiment plan specifying the provision of providing data to the outside, it is necessary to obtain the approval of the ethics committee of the Hokkaido University Hospital. Unfortunately, the author cannot make the data publicly available in the prescribed form.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The author thanks the laboratory students Kazuya Sasaki, Hazuki Kobayashi, Kanako Kawahara, and Momoka Yamauchi for their help.

References

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998.

[2] M. U. Dalmiş, G. Litjens, K. Holland et al., "Using deep learning to segment breast and fibroglandular tissue in MRI volumes," *Medical Physics*, vol. 44, no. 2, pp. 533–546, 2017.

[3] Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, and B. J. Erickson, "Deep learning for brain MRI segmentation: state of the art and future directions," *Journal of Digital Imaging*, vol. 30, no. 4, pp. 449–459, 2017.

[4] P. Lakhani and B. Sundaram, "Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks," *Radiology*, vol. 284, no. 2, pp. 574–582, 2017.

[5] K. Yasaka, H. Akai, A. Kunimatsu, O. Abe, and S. Kiryu, "Liver fibrosis: deep convolutional neural network for staging by using gadoxetic acid–enhanced hepatobiliary phase MR images," *Radiology*, vol. 287, no. 1, pp. 146–155, 2018.

[6] A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 8–20, 2017.

[7] X. W. Gao, R. Hui, and Z. Tian, "Classification of CT brain images based on deep learning networks," *Computer Methods and Programs in Biomedicine*, vol. 138, pp. 49–56, 2017.

[8] M. Ali, D.-H. Son, S.-H. Kang, and S.-R. Nam, "An accurate ct saturation classification using a deep learning approach based on unsupervised feature extraction and supervised fine-tuning strategy," *Energies*, vol. 10, no. 12, p. 1830, 2017.

[9] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu et al., "Convolutional neural networks for medical image analysis: full training or fine tuning?," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.

[10] M. Havaei, A. Davy, D. Warde-Farley et al., "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017.

[11] A. Kalinovsky, V. Liauchuk, and A. Tarasau, "Lesion detection in ct images using deep learning semantic segmentation technique," *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-2/W4, pp. 13–17, 2017.

[12] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI," *Medical Image Analysis*, vol. 30, pp. 108–119, 2016.

[13] W. Zhang, R. Li, H. Deng et al., "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *NeuroImage*, vol. 108, pp. 214–224, 2015.

[14] C. Wachinger, M. Reuter, and T. Klein, "DeepNAT: deep convolutional neural network for segmenting neuroanatomy," *NeuroImage*, vol. 170, pp. 434–445, 2018.

[15] O. Emad, I. A. Yassine, and A. S. Fahmy, "Automatic localization of the left ventricle in cardiac MRI images using deep learning," in *Proceedings of 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 683–686, Milan, Italy, 2015, https://embc.embs.org/2015/.

[16] D. Forsberg, E. Sjöblom, and J. L. Sunshine, "Detection and labeling of vertebrae in MR images using deep learning with clinical annotations as training data," *Journal of Digital Imaging*, vol. 30, no. 4, pp. 406–412, 2017.

[17] H. K. van der Burgh, R. Schmidt, H. J. Westeneng, M. A. de Reus, L. H. van den Berg, and M. P. van den Heuvel, "Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis," *NeuroImage: Clinical*, vol. 13, pp. 361–369, 2017.

[18] J. Cho, K. Lee, E. Shin, G. Choy, and S. Do, "How much data is needed to train a medical image deep learning system to achieve necessary high accuracy?," in *International Conference on Learning Representations 2016*, San Juan, Puerto Rico, 2015.

[19] H. R. Roth, C. T. Lee, H.-C. Shin et al., "Anatomy-specific classification of medical images using deep convolutional nets," in *IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 101–104, Brooklyn Bridge, NY, USA, April 2015.

[20] K. T. Bae, "Intravenous contrast medium administration and scan timing at CT: considerations and approaches," *Radiology*, vol. 256, no. 1, pp. 32–61, 2010.

[21] D. J. Hamlin, F. A. Burgener, and J. B. Beecham, "CT of intramural endometrial carcinoma: contrast enhancement is essential," *American Journal of Roentgenology*, vol. 137, no. 3, pp. 551–554, 1981.

[22] J. Deng, A. C. Berg, K. Li, and L. Fei-Fei, "What does classifying more than 10,000 image categories tell us?," *Lecture Notes in Computer Science*, vol. 6315, no. 5, pp. 71–84, 2010.

[23] A. Masood, B. Sheng, P. Li et al., "Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images," *Journal of Biomedical Informatics*, vol. 79, pp. 117–128, 2018.

[24] X. Zhao, L. Liu, S. Qi, Y. Teng, J. Li, and W. Qian, "Agile convolutional neural network for pulmonary nodule classification using CT images," *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 4, pp. 585–595, 2018.

[25] X. Ren, L. Xiang, D. Nie et al., "Interleaved 3D-CNNs for joint segmentation of small-volume structures in head and neck CT images," *Medical Physics*, vol. 45, no. 5, pp. 2063–2075, 2018.

[26] M. Zreik, N. Lessmann, R. W. van Hamersvelt et al., "Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis," *Medical Image Analysis*, vol. 44, pp. 72–85, 2018.

[27] Y. Jia, E. Shelhamer, J. Donahue et al., "Caffe: convolutional architecture for fast feature embedding," in *Proceedings of the ACM International Conference on Multimedia-MM '14*, pp. 675–678, Orland, Florida, USA, 2014, http://acmmm.org/2014/.

[28] I. Dimitrovski, D. Kocev, I. Kitanovski, S. Loskovska, and S. Džeroski, "Improved medical image modality classification using a combination of visual and textual features," *Computerized Medical Imaging and Graphics*, vol. 39, pp. 14–26, 2015.

[29] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, Boston, MA, USA, June 2015.

[30] M. Kidoh, T. Nakaura, S. Nakamura et al., "Reduction of dental metallic artefacts in CT: Value of a newly developed algorithm for metal artefact reduction (O-MAR)," *Clinical Radiology*, vol. 69, no. 1, pp. e11–e16, 2014.

[31] J. Aalten, H. M. Dekker, J. A. D. Van Der Vliet, and A. J. Hoitsma, "Does a plain X-ray of the pelvis predict arterial complications in renal transplantation? A prospective study," *Nephrology Dialysis Transplantation*, vol. 26, no. 6, pp. 2007–2012, 2011.

[32] J. Blacher, A. P. Guerin, B. Pannier, S. J. Marchais, and G. M. London, "Arterial calcifications, arterial stiffness, and cardiovascular risk in end-stage renal disease," *Hypertension*, vol. 38, no. 4, pp. 938–942, 2001.

[33] A. Blum, A. Kalai, and J. Langford, "Beating the hold-out: bounds for K-fold and progressive cross-validation," in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory-COLT '99*, pp. 203–208, Santa Cruz, CA, USA, July 1999.

Hindawi
Journal of Healthcare Engineering
Volume 2018, Article ID 7574860, 8 pages
https://doi.org/10.1155/2018/7574860

Research Article

Feasibility of Kinect-Based Games for Balance Rehabilitation: A Case Study

Ines Ayed,[1,2] **Adel Ghazel**,[1] **Antoni Jaume-i-Capó**,[2] **Gabriel Moya-Alcover**,[2] **Javier Varona**,[2] **and Pau Martínez-Bueso**[3]

[1]*GresCom Lab, Ecole Supérieure des Communications de Tunis, Université de Carthage, Tunis, Tunisia*
[2]*Unitat de Gràfics, Visió per Computador i Intel·ligéncia Artificial, Department de Ciències Matemàtiques i Informàtica, Universitat de les Illes Balears, Palma, Spain*
[3]*Research Group on Evidence, Lifestyles and Health, Department of Nursing and Physiotherapy, University of the Balearic Islands, Palma, Spain*

Correspondence should be addressed to Ines Ayed; ines.ayed91@gmail.com

Received 23 March 2018; Revised 11 May 2018; Accepted 23 May 2018; Published 9 July 2018

Academic Editor: Zahid Akhtar

We aimed at determining the effects of prototype games on older adults attending a rehabilitation program in an elderly house in this work. We conducted an initial case study where two participants underwent a 5-week intervention. Feasibility was assessed by examining recruitment, adherence, and safety. The Tinetti balance test was used as pretest and posttest assessments. Results show that adherence was very high and no adverse effects were registered during the sessions. The included participants also reported enjoyment during the playtime and exhibited improvements in Tinetti scores. The findings suggest that game-based rehabilitation can be useful for improving balance in elderly people and can be incorporated in a fall prevention program.

1. Introduction

Falls are prevalent among elderly people. More than a third of people aged 65 years and over fall at least once per year [1, 2]. Risk of falls is even higher within institutional residents than among community dwelling older people, with an incidence of 1.5 falls per bed per year [3, 4]. In fact, housing type was determined as one of the extrinsic factors predisposing to risk of falls and fractures [5]. However, previous falls and deficits in gait and balance have been identified as significant risk factors as they were highly correlated with falling [2, 3, 6]. In fact, falls are the leading cause of injury, deaths, and hospital admissions for traumatic injuries among people aged 65 and older [1]. Adopting preventive measures and implementing effective fall prevention programs and interventions help us to reduce the risk of falls and minimize the substantial social and healthcare costs induced by falls [7, 8]. Moreover, decreasing fall-related injury incidences could significantly improve quality of life of older adults and their caregivers [9].

Most of the fall prevention programs include motor rehabilitation with exercise interventions. A fall prevention meta-analysis showed that exercise programs that challenged balance were associated with significant reductions in fall rates [8].

Recently, serious games have been proven to be effective in motor rehabilitation [10]. Serious games are games designed for a primary purpose different from the purpose of pure entertainment; the cognitive and motor activities required by these games attract the attention of users [11], and this helps to distract them from the task [12, 13]. Healthcare-related serious games can be deployed in treatment, recovery, and rehabilitation. It is demonstrated that these games are highly promising in maintaining motor skills since they help us to motivate the patients to adhere to repetitive and intensive therapy sessions [14]. In long-term rehabilitation programs, patient's demotivation is frequent because of the repetition and boredom of rehabilitative activities, which may cause the lack of focus or the abandonment of the program, and consequently the loss of the benefit and effectiveness of the therapy.

Elderly population was the target of many studies that examined the games' effectiveness in balance training [15–17]. Though much more research in this area is needed, published findings show that video games can be remedies for minimizing risk of falls [18, 19]. In addition, vision-based systems are proven to be suitable for elderly people as they are motivating and noninvasive [20–23]. While exist many vision-based interfaces, Microsoft Kinect gained much interest in rehabilitation in recent years by offering a natural human computer interaction. Kinect is a low-cost RGBD sensor that captures colour and depth data providing full-body tracking and gesture recognition. It enables the user to interact with the game without the need to use a controller device.

Several studies have investigated the feasibility of using Kinect with elderly participants employing commercial games to improve balance and gait [24, 25]. Nevertheless, usability tests conducted in [26] show that some commercially available games are not suitable for therapeutic purposes and provide negative auditory and visual feedback during game tasks. Thus many research studies proposed prototype games specified for rehabilitation and addressed to defined target groups. For instance, Hoda et al. [27] aimed at improving the movement and control of upper limbs of elderly stroke patients; Ofli et al. [28] suggested a set of exercises for the improvement of balance, flexibility, strength, and endurance of independently living older adults; and authors in [29] investigated the feasibility of a rehabilitation game for dynamic postural control of people with Parkinson's disease. Although there is evidence that serious games improve balance of elderly people in general, there is a lack of evidence about their effects, under the supervision of physiotherapists, on institutionalized older adults and their deployment in elderly houses and institutions. Accordingly, this paper is an exploratory and descriptive case study [30] about the impact that the games could have in the development of future rehabilitation games for the fall prevention programs for elderly people using vision-based technologies. Such works are preliminary studies that help in the design and preparation of randomised clinical trials [31]. They tend to describe the feasibility and potential effectiveness of the system proposed at a lower cost and shorter time.

The aim of this work was to investigate the feasibility and effectiveness of prototype Kinect-based serious games that focus on postural control and balance rehabilitation in elderly people. In the study, we used Microsoft Kinect as the vision-based interface, and we hypothesised that the games would have a positive impact on the participants attending the elderly house.

2. Materials and Methods

This section explains how we assessed the feasibility of Kinect-based games for balance rehabilitation in the elderly house. First, we present the serious game implementation. Second, we describe the case trial design.

The serious games developed for this study try to imitate exercises included in traditional physical therapy, such as reaching in different directions, small and large lateral steps, weight shifting to both sides, neck movements (flexion, extension, lateral flexion, and rotation), shoulder movements (flexion, extension, adduction and abduction), trunk movements (flexion, extension, lateral flexion, and rotation), knee movements (flexion and extension), and hip movements (flexion, extension, rotation, adduction and abduction) [32–34]. These, however, had the added value that the immersive virtual environment tried to hook the patient to the point of not focusing on the fact of being in a rehabilitation session. This rehabilitation method using serious games is proposed for patients with risk of falls in order to improve their balance.

2.1. Game Design and Implementation. A series of games focusing on the physical rehabilitation of the balance and postural control of elderly people were designed and developed according to the requirements and indications from physiotherapists [35]. In the design process, we adopted the guidelines for developing rehabilitative games in order to offer engaging exercises that predispose motivation, feedback, and game monitoring [36]. Safety of the games is also very important; thus, game parameters were configured to allow the physiotherapist to adapt the exercises for each user according to his needs and preferences. The developed games are the following:

(i) *Reach game*: Here, the users had to move their centre of mass (COM) in order to reach with their hands one of the five balls located on their user plan. Once the user touched a ball, it disappeared and reappeared after a determined time set by the physiotherapist according to the user's speed. The two symmetric items added previously on the level of hips were deleted because the participants found it very difficult to do weight shifting movements. To further adapt the games for the participants recruited, the balls were resized and their colours were set to black and red over a simple white background. Data like user id, session date, play duration, and distance from the sensor were automatically recorded and saved in an excel document for any later check or use.

(ii) *HitIt*: In this game, soccer balls fall randomly within the same plan of the user. To hit them, the user needs to make lateral steps, and touch them with his head when they are at his level. The use of any other part of the body, except the head, goes unperceived during the game play. The game can be also played in a seated position. The user has to make lateral movements of the trunk to be able to touch the elements with his head.

(iii) *WatchOut*: Here, subjects have to move laterally in order to escape falling eggs. Falling items fall randomly within the same plan of the user with adaptable falling rate and speed.

In all games, a background with a solid colour has been defined to keep the user focused on the game.

2.2. Trial Design. We present a case study to describe in detail a patient's episode of care and to assess a new initiative in a health service [30, 37]. We undertook an intrinsic case study to investigate the feasibility and effectiveness of prototype Kinect-based serious games on postural control and balance rehabilitation in a group of elderly people. And this was conducted into the specific context of an elderly house in Tunisia, where demographic changes have decreased fertility due to aging population. Persons aged more than 65 in the country constituted 8% of the total population in 2015; this percentage is considered the highest in Arab region. In 2050, it is predicted to reach 20 percent of the population [38].

2.2.1. Participants. Participants were chosen by a physiotherapist to participate in this study and were recruited in an elderly house. The inclusion criteria were as follows: (1) aged over 55 years; (2) ability to understand, learn, and follow simple instructions; and (3) voluntary agreement to participate in the clinical study. The exclusion criteria were as follows: (1) severe cognitive deterioration; (2) profound bilateral hearing loss with the use of hearing aids; (3) hemiplegia, dementia, or Parkinson; (4) serious or uncontrolled epilepsy; and (5) serious or recurring medical complications.

Total population was screened for eligibility in the elderly house of Manouba in Tunisia: 120 older adults. Among them, only 32 met the inclusion and exclusion criteria. The research team made a request to participate in the intervention to all adults who met the inclusion criteria. 15 subjects accepted to participate in the study but only 8 showed up for a trial session that lasted between 10 and 15 min. Those who joined the trial session liked the games, but for different reasons, some of them could not adhere to the trial sessions (one got a job and others had time incompatibility with doctor visit or their time of hanging out of the elderly house), so the final case study sample included only 3 subjects. One of the participants dropped in the middle of the intervention as she left the elderly house, and 2 completed the intervention and the pre- and postassessments. Subjects who accepted to participate were illiterate and were screened for cognitive impairment using Arabic-Mini Mental State Examination (A-MMSE) [39]. An A-MMSE score higher than 20 suggested that both participants had no severe cognitive impairment as defined in the exclusion criteria.

For accessibility reasons, the serious game program was provided to users with different conditions (education, history of falls, cognitive impairment level, and balance problems). The two subjects included in this case study [40] met also the following criteria:

(1) Subjects' capabilities were stable which means there would not be any evolvement in their capabilities during the study due to some pathology.

(2) Users were attending intervention sessions according to the study schedule, so they were not exercising out of this time. This criterion provided homogeneity to the context of study.

Table 1: Participants' characteristics.

	Participant 1	Participant 2
Age (years)	78	72
Sex	Female	Female
History of falls	2	0
Walking aids	Yes	Yes
A-MMSE	22/30	28/30

The characteristics of the participants are presented in Table 1. An informed consent was signed before the intervention.

2.2.2. Procedure. The intervention was conducted along a period of 5 weeks. Each participant underwent one 30 min session per day at a rate of 3 days a week. The rehabilitation program was divided as follows: 20 min was devoted for *Reach game*, 10 min for *HitIt*, and an extra 5 min was added for *WatchOut* game starting from session number 9. The duration of each game was set according to its understanding and its acceptability by the participants. A break time between 3 and 6 min was allowed each 5 min of play regarding the fatigue level.

Balance was assessed using the Tinetti balance test at the beginning (initial assessment) and at the end (final assessment) of the intervention. Adherence, game scores, and adverse events were noted along the intervention. The design of the clinical intervention is presented in Figure 1.

The system setting consisted of a personal computer and Microsoft Xbox Kinect sensor (Figure 2(a)). In the literature, a vision-based system that processes higher than 19 fps its response is considered real time [41]. All games were designed to process at least 25 fps, and this ensures a real-time interaction. The participants stood in front of the camera which detected their presence and allowed them interacting with virtual objects appearing on the screen. The Kinect camera was placed at a fixed height of 1 m and the average distance from the camera was 1.5 m. The games used were *HitIt* (Figure 2(b)), *Reach game* (Figure 2(c)), and *WatchOut* (Figure 2(d)).

2.2.3. Measurements. Gaming was assisted by an occupational therapist (OT) and monitored by the research team.

As stated before, all subjects were clinically evaluated prior to the intervention program and again at the conclusion to assess their balance with the Tinetti balance test.

The Tinetti test was performed by the OT who was not blinded to the intervention. Other measurements were registered in a case report form, by the research team or saved automatically by the system.

To assess the feasibility of using vision-based rehabilitation games for balance and postural control in the elderly house, the following outcome measurements were used:

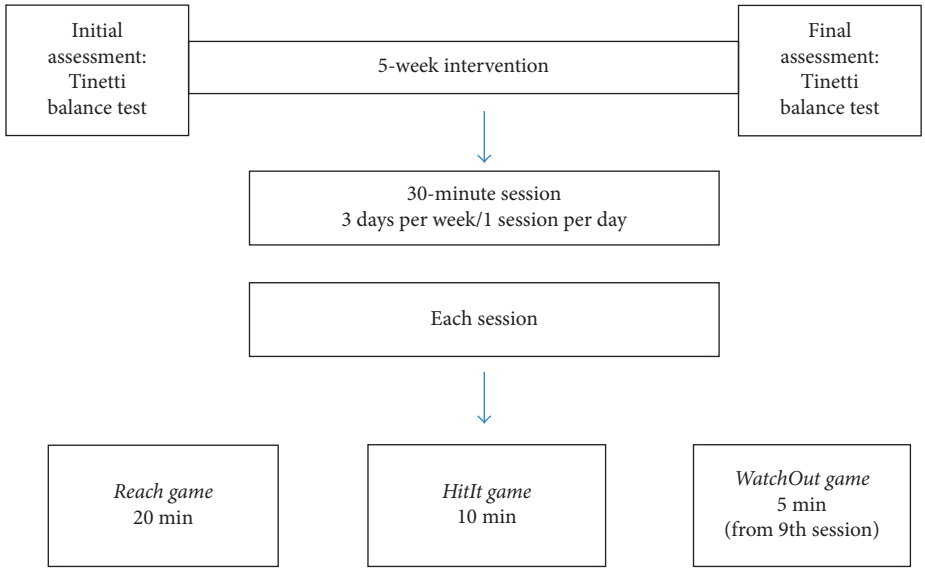


Figure 1: Intervention scheme.

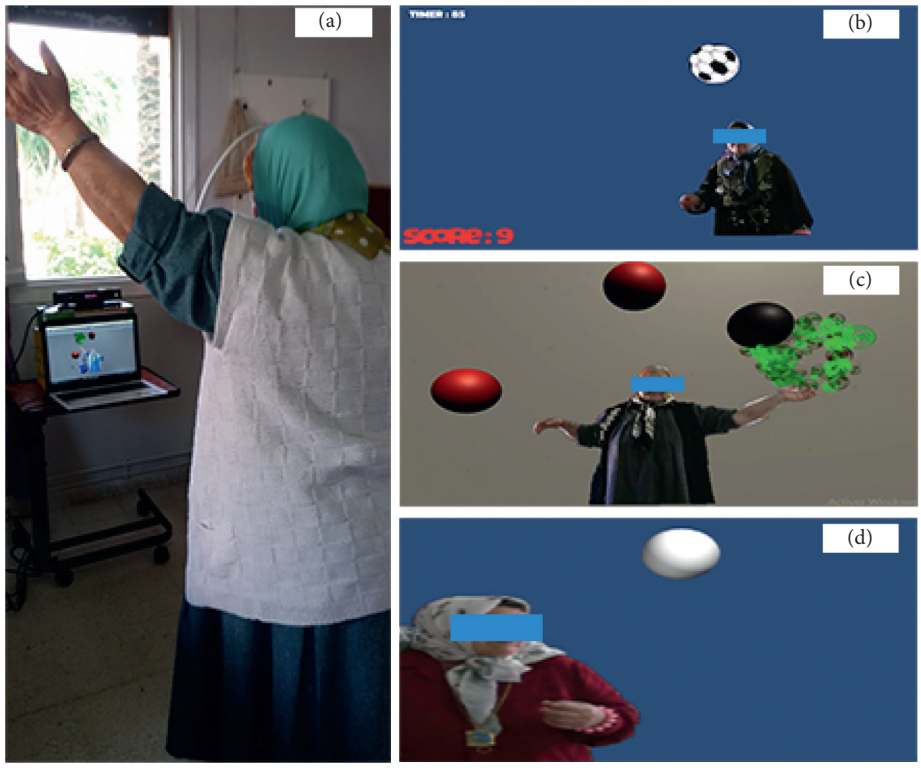


Figure 2: (a) System settings; (b) *HitIt* game; (c) *Reach game*; (d) *WatchOut* game.

(i) *Tinetti balance test:* The Tinetti assessment tool is a simple and easily administered test that measures a patient's gait and balance [42]. The test scoring is done on a three-point ordinal scale with a range of 0 to 2 where a score of 0 represents the most impairment, while 2 would represent independence of the patient. The individual scores are then combined to form three measures: an overall gait assessment score, an overall balance assessment score, and a gait and balance score. The maximum score for the gait component is 12 points. The maximum score for the balance component is 16 points. The maximum total score is 28 points. In general, patients who score below 19 are at a high risk for falls, while a score in the range of 19–24 indicates that the patient has a risk for falls [43].

Table 2: Tinetti balance test scores.

	Participant 1		Participant 2	
	Preassessment	Postassessment	Preassessment	Postassessment
Total score	15	19	14	17
Balance score	7	11	6	10
Gait score	8	8	8	7

(ii) *Adherence:* The research team registered the number and length of sessions. Planned playing time and actual playing time were recorded by the system.

(iii) *Game score:* The score of *HitIt* game was saved by the system at each session. Whenever the participant touched the ball with her head, one point was added to the score. The scores of the other games were not considered in this case since they did not truly reflect the progress of the participant. Besides, the *HitIt* game score was used to motivate the patients and create a kind of competitiveness between them.

(iv) *Position of the patients and healthcare assistant:* The position of the participants as seated or standing and the location of the therapist with respect to the participants were noted during intervention sessions.

(v) *Adverse events:* Events such as falls, fatigue, or any safety incidents requiring medical attention were also documented. Participants were orally asked for feedback during each session for any further improvements or suggestions.

3. Results

We are interested in reporting the results of the two participants who completed the intervention and performed the pre- and postassessments.

3.1. Balance. According to the Tinetti test scores, both participants showed a similar trend. As shown in Table 2, there was an improvement of 4 points in the balance section score, while almost no difference was noted between pre- and postmeasures in the gait section. In the preassessment, Participant 1 had a total score of 15/28 (balance: 7/16; gait: 8/12) and Participant 2 had a total score of 14/28 (balance: 6/16; gait: 8/12). In the postassessment, Participant 1 had a total score of 19/28 (balance: 11/16; gait: 8/12) and Participant 2 had a total score of 17/28 (balance: 10/16; gait: 7/12).

3.2. Adherence. Participants attended 86.6% of the sessions with an average 30 min length each. As depicted in Figure 3, Participant 1 had 363 min of a total playing time divided between *Reach game* (245 min), *HitIt* (123 min), and *WatchOut* (15 min). Similarly, Participant 2 spent around 415 min of a total playing time divided between *Reach game* (270 min), *HitIt* (125 min), and *WatchOut* (20 min).

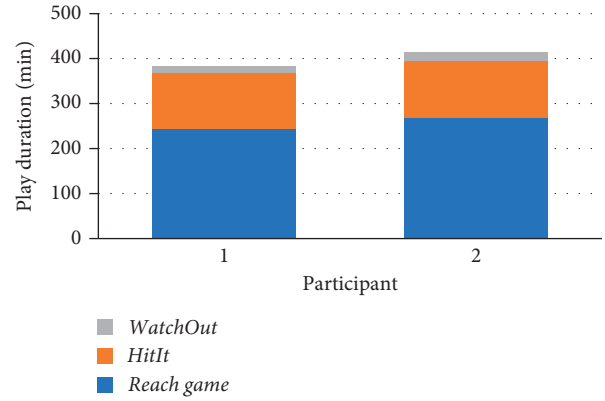


Figure 3: Time spent by each participant in each game.

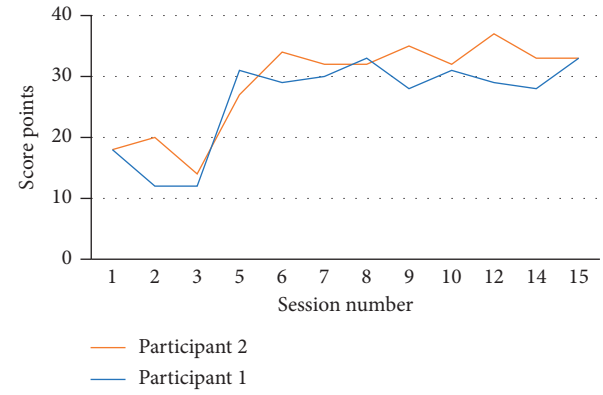


Figure 4: *HitIt* game score for the two participants along the experiment.

3.3. Game Score. The game score improved over time, although there were no significant differences over time between the two participants as shown in Figure 4.

3.4. Position of the Participant and Healthcare Assistant. In all sessions, the participants were standing during playtime, and they sat on a chair during the breaks. As for the therapist, she stood behind the participants the first sessions as they expressed fear of fall and needed more help in doing lateral steps in *HitIt* and *WatchOut* games or reaching objects in *Reach game*. By the end, the participants could stand without any support and the assistant stood on their left or right side, between one and three meters away, with the exception of Participant 1 who had a relapse after the fall accident and needed more support for restoring her confidence again.

3.5. Adverse Events. Participant 1 fell during the intervention period; precisely after the ninth session, it was on a weekend day, not in the training day because she did not use her ambulatory aids that day. She returned to assist the training sessions just two days after the fall accident. At the beginning, she felt unsecure and needed more assistance as someone had to stand behind her for reassurance, but sooner she started to gain confidence again. Besides, it is worth mentioning that she got tired easily in the postfall period, usually both participants reported fatigue after 5 or 6 min of training, but she asked for a break after 2 or 3 min. No serious adverse events were reported from Participant 2.

4. Discussion

As far as we reviewed, it is a first case study conducted in a Tunisian elderly house using Microsoft Kinect and serious games for balance rehabilitation and postural control.

The findings show that the recruitment was very low, and this is due to the following. First, the recruitment period was short which partially explains the low participation between the elderly residents of the institution.

Second, exergaming was a new concept for many older adults. In fact, some refused to participate thinking it would be difficult for them to understand the games looking to the technology used, and others considered themselves too old to play. Consequently, we had to change the terminology used to stimulate them to participate at least to a trial session. We believe that using terms as "exercises" and "training" rather than "games" and "playing" may increment the recruitment rate.

Third, disliking exergames was also associated with low recruitment and high dropout rate as shown in another study suggesting that elderly people go for paper-based instructions to practice self-regulated conventional exercise instead of using computer-based exergames [44].

Nonetheless, the results related to the two participants were quite promising. Regarding the Tinetti balance test, both participants have shown an improvement of 4 points in the total score over the study period; the increase in the score was noticed especially in the balance section. The Tinetti test measures balance and gait, but our system focused particularly on balance; this justifies the absence of significant differences in the gait section as well. However, with this improvement, Participant 1 moved from high risk for falls range to the risk for falls range, while participant 2 is still at high risk for falls [43, 45].

In addition, the adherence rate was very high, and participants reported enjoyment during play sessions. We assume that the prototype games provided engagement, fun, and motivation for the participants. Interest and enjoyment provided in a game-based environment was pointed out by many researchers [29, 46, 47]. Regarding to play duration, the time dedicated to play *WatchOut* game was much less than the playtime of *HitIt* game because they both implicated doing the same movements nevertheless it was sometimes confusing for participants whether to hit or to avoid falling objects. The game score in *HitIt* game was motivating and created competitiveness between the two participants as they were performing the exercises alternately at the same period of time. They were both eager to score more points by hitting more balls. Social interaction between the two participants had an impact on motivation as well. In this respect, the findings of Wu et al. highlighted the importance of social presence for motivating elderly people while exergaming [18, 48].

Though fatigue was daily reported every 5 or six minutes of play, the games were safe especially with the presence of physiotherapist around. The fall of one participant occurred out of the training sessions, and her return to the intervention program had implied two main points. The first point is that games were engaging and interesting. The second point is that games could be used to restore confidence after falls. Older adults minimize their activities and close up on themselves after fall, besides they express fear of falling again [49]. They usually develop "postfall syndrome" when they fall; in other terms, they become very anxious and cannot stand and walk without support [50]. For that reason, the participant needed more support and reassurance. Games may soften the impact of postfall syndrome and help in the reintegration of the faller into his community [51, 52].

The two subjects included in this case study met the criteria that would be accomplished by the future subjects. Conducting this study, we proved that for subjects that meet these criteria, the system and rehabilitative activities facilitated by the prototype games may be suitable and have allowed closely spotting and improving the issues that elderly subjects may face when interacting with the system providing more effective rehabilitation training for fall prevention.

The use of materials that were relatively small in size recommends a wider screen to facilitate the accessibility to elderly people. Besides, expanding the games repertoire by adding games that require movement and gait to work on static and dynamic balance was highly recommended by the physiotherapist for any future work.

5. Conclusion

We conducted a first case study in an elderly house where two participants underwent a 5-week intervention. During the training sessions, the included participants played three prototype games using Microsoft Kinect. The objective of the study was to investigate the feasibility of Kinect-based prototype games in a Tunisian elderly house and study its effectiveness on balance and postural control of elderly people. Consequently, aspects such as enjoyment, adherence, and adverse events were monitored. In addition, balance impact was assessed by the Tinetti balance test. In general, the participants reported enjoyment. They attended the majority of the training sessions and no adverse events were noted. Furthermore, there was an improvement in Tinetti balance test scores.

In conclusion, despite the study sample size was very small, the findings suggest that the use of Kinect-based prototype games for improving balance and postural control of elderly people living in a Tunisian elderly house is feasible and safe and can be integrated in fall prevention

programs. In future work, it is intended to conduct a randomized controlled study with a larger sample size to further study the effectiveness of these games.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors acknowledge the Ministerio de Economía, Industria y Competitividad (MINECO), the Agencia Estatal de Investigación (AEI), and the European Regional Development Funds (ERDF) for their support to the project TIN2016-81143-R (MINECO/AEI/ERDF, EU). Ines Ayed benefited from the fellowship FPI/2039/2017 from the Conselleria d'Educació, Cultura i Universitats, of the Govern de les Illes Balears. The authors thank the elderly house of Manouba in Tunisia for the facilities offered in order to conduct this case study.

References

[1] A. J. Campbell, J. Reinken, B. C. Allan, and G. S. Martinez, "Falls in old age: a study of frequency and related clinical factors," *Age and Ageing*, vol. 10, no. 4, pp. 264–270, 1981.

[2] J. M. Hausdorff, D. A. Rios, and H. K. Edelber, "Gait variability and fall risk in community-living older adults: a 1-year prospective study," *Archives of Physical Medicine and Rehabilitation*, vol. 82, no. 8, pp. 1050–1056, 2001.

[3] L. Z. Rubenstein and K. R. Josephson, "The epidemiology of falls and syncope," *Clinics in Geriatric Medicine*, vol. 18, no. 2, pp. 141–158, 2002.

[4] L. Z. Rubenstein, K. R. Josephson, and A. S. Robbins, "Falls in the nursing home," *Annals of Internal Medicine*, vol. 121, no. 6, pp. 442–451, 1994.

[5] T. D. L. Resende, "Analysis of extrinsic and intrinsic factors that predispose elderly individuals to fall," *CEP*, vol. 90, p. 900, 2000.

[6] M. E. Tinetti and C. Kumar, "The patient who falls: "it's always a trade-off"," *JAMA*, vol. 303, no. 3, pp. 258–266, 2010.

[7] J. A. Stevens, P. S. Corso, E. A. Finkelstein, and T. R. Miller, "The costs of fatal and non-fatal falls among older adults," *Injury Prevention*, vol. 12, no. 5, pp. 290–295, 2006.

[8] C. Sherrington, J. C. Whitney, S. R. Lord et al., "Effective exercise for the prevention of falls: a systematic review and meta-analysis," *Journal of the American Geriatrics Society*, vol. 56, no. 12, pp. 2234–2243, 2008.

[9] R. B. Silva, G. D. Eslick, and G. Duque, "Exercise for falls and fracture prevention in long term care facilities: a systematic review and meta-analysis," *Journal of the American Medical Directors Association*, vol. 14, no. 9, pp. 685–689, 2013.

[10] P. Rego, P. M. Moreira, and L. P. Reis, "Serious games for rehabilitation: a survey and a classification towards a taxonomy," in *Information Systems and Technologies (CISTI), 2010 5th Iberian Conference*, IEEE, pp. 1–6, Berlin, Germany, 2010.

[11] D. A. Norman, *The Invisible Computer*, MIT Press, Cambridge, MA, USA, 1998.

[12] E. Flores, G. Tobon, E. Cavallaro, F. Cavallaro, J. Perry, and T. Keller, "Improving patient motivation in game development for motor deficit rehabilitation," in *Proceedings of Int'l ACM Conference Advances in Computer Entertainment Technology (ACE '08)*, pp. 381–384, New York, NY, USA, 2008.

[13] A. Krichevets, E. Sirotkina, I. Yevsevicheva, and L. Zeldin, "Computer games as a means of movement rehabilitation," *Disability and Rehabilitation*, vol. 17, no. 2, pp. 100–105, 1995.

[14] A. Jaume-i-Capó, P. Martínez-Bueso, B. Moya-Alcover, and J. Varona, "Interactive rehabilitation system for improvement of balance therapies in people with cerebral palsy," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 2, pp. 419–427, 2014.

[15] A. Nawaz, N. Skjæret, J. L. Helbostad et al., "Usability and acceptability of balance exergames in older adults: a scoping review," *Health Informatics Journal*, vol. 22, no. 4, pp. 911–931, 2016.

[16] M. Van Diest, C. J. Lamoth, J. Stegenga, G. J. Verkerke, and K. Postema, "Exergaming for balance training of elderly: state of the art and future developments," *Journal of Neuroengineering and Rehabilitation*, vol. 10, no. 1, p. 101, 2013.

[17] H. Hawley-Hague, E. Boulton, A. Hall, K. Pfeiffer, and C. Todd, "Older adults' perceptions of technologies aimed at falls prevention, detection or monitoring: a systematic review," *International Journal of Medical Informatics*, vol. 83, no. 6, pp. 416–426, 2014.

[18] S. D. Choi, L. Guo, D. Kang, and S. Xiong, "Exergame technology and interactive interventions for elderly fall prevention: a systematic literature review," *Applied Ergonomics*, vol. 65, pp. 570–581, 2017.

[19] K. Balzer, M. Bremer, S. Schramm, D. Lühmann, and H. Raspe, "Falls prevention for the elderly," *GMS Health Technology Assessment*, vol. 8, 2012.

[20] H. R. Marston and S. T. Smith, "Interactive videogame technologies to support independence in the elderly: a narrative review," *Games for Health Journal*, vol. 1, no. 2, pp. 139–152, 2012.

[21] A. E. Staiano and R. Flynn, "Therapeutic uses of active videogames: a systematic review," *Games for Health Journal*, vol. 3, no. 6, pp. 351–365, 2014.

[22] L. McPhate, E. M. Simek, and T. P. Haines, "Program-related factors are associated with adherence to group exercise interventions for the prevention of falls: a systematic review," *Journal of Physiotherapy*, vol. 59, no. 2, pp. 81–92, 2013.

[23] L. Robinson, J. L. Newton, D. Jones, and P. Dawson, "Self-management and adherence with exercise-based falls prevention programmes: a qualitative study to explore the views and experiences of older people and physiotherapists," *Disability and Rehabilitation*, vol. 36, no. 5, pp. 379–386, 2014.

[24] J. Pompeu, L. Arduini, A. Botelho et al., "Feasibility, safety and outcomes of playing Kinect Adventures!™ for people with Parkinson's disease: a pilot study," *Physiotherapy*, vol. 100, no. 2, pp. 162–168, 2014.

[25] B. Seamon, M. DeFranco, and M. Thigpen, "Use of the Xbox Kinect virtual gaming system to improve gait, postural control and cognitive awareness in an individual with Progressive Supranuclear Palsy," *Disability and Rehabilitation*, vol. 39, no. 7, pp. 721–726, 2016.

[26] B. Lange, S. Flynn, and A. Rizzo, "Initial usability assessment of off-the- shelf video game consoles for clinical game-based motor rehabilitation," *Physical Therapy Reviews*, vol. 14, no. 5, pp. 355–363, 2009.

[27] M. Hoda, Y. Hoda, A. Hage, A. Alelaiwi, and A. El Saddik, "Cloud-based rehabilitation and recovery prediction system for stroke patients," *Cluster Computing*, vol. 18, no. 2, pp. 803–815, 2015.

[28] F. Ofli, G. Kurillo, S. Obdrzalek, R. Bajcsy, H. Jimison, and M. Pavel, "Design and evaluation of an interactive exercise

coaching system for older adults: lessons learned," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 1, pp. 201–212, 2016.

[29] B. Galna, D. Jackson, G. Schofield et al., "Retraining function in people with Parkinson's disease using the Microsoft Kinect: game design and pilot testing," *Journal of NeuroEngineering and Rehabilitation*, vol. 11, no. 1, p. 60, 2014.

[30] R. Yin, "Case study research, design and methods," in *Applied Social Research Methods Series*, Vol. 5, Sage, Thousand Oaks, CA, USA, 3rd edition, 2003.

[31] R. Rosenthal and R. L. Rosnow, *Essentials of Behavioral Research: Methods and Data Analysis*, McGraw-Hill, New York, NY, USA, 1991.

[32] M. Zhang and Y. Fan, *Computational Biomechanics of the Musculoskeletal System*, CRC Press, Boca Raton, FL, USA, 2014.

[33] T. T. Dao and M. C. Ho Ba Tho, "Biomechanics of the musculoskeletal system," *Biomechanics of the Musculoskeletal System*, pp. 1–35, 2014.

[34] D. A. Neumann, *Kinesiology of the Musculoskeletal System*, Foundations for Rehabilitation, St. Louis, MO, USA, 2013.

[35] I. Ayed, A. Ghazel, A. Jaume-i-Capo et al., "Fall prevention serious games for elderly people using RGBD devices," in *Proceedings of Games and Virtual Worlds for Serious Applications (VS-Games), 2016 8th International Conference*, pp. 1–3, IEEE, Barcelona, Spain, 2016.

[36] A. Jaume-I-Capó, B. Moyà-Alcover, and J. Varona, "Design issues for vision-based motor-rehabilitation serious games," in *Technologies of Inclusive Well-Being*, pp. 13–24, Springer, Berlin, Germany, 2014.

[37] R. E Stake, *The Art of Case Study Research*, Sage Publications Ltd., London, UK, 1995.

[38] A. Goujon, D. Weber, and E. Loichinger, *Demographic Profile of the Arab Region: Realizing the Demographic Dividend*, IIASA, New York, NY, USA, 2016.

[39] S. Al-Rajeh, A. Ogunniyi, A. Awada, A. Daif, and R. Zaidan, "Preliminary assessment of an Arabic version of the mini-mental state examination," *Annals of Saudi Medicine*, vol. 19, no. 2, pp. 150–152, 1999.

[40] J. Lazar, J. H. Feng, and H. Hochheiser, *Research Methods in Human-Computer Interaction*, Wiley, London, UK, 2010.

[41] J. Varona, A. Jaume-i-Capó, J. Gonzèlez, and F. J. Perales, "Toward natural interaction through visual recognition of body gestures in real-time," *Interacting with Computers*, vol. 21, no. 1-2, pp. 3–10, 2008.

[42] M. E. Tinetti, "Performance-oriented assessment of mobility problems in elderly patients," *Journal of the American Geriatrics Society*, vol. 34, no. 2, pp. 119–126, 1986.

[43] M. L. del Nogal, A. González-Ramírez, and A. Palomo-Iloro, "Evaluación del riesgo de caídas. protocolos de valoración clínica," *Revista Española de Geriatría y Gerontología*, vol. 40, no. 1, pp. 54–63, 2005.

[44] P. Oesch, J. Kool, L. Fernandez-Luque et al., "Exergames versus self-regulated exercises with instruction leaflets to improve adherence during geriatric rehabilitation: a randomized controlled trial," *BMC Geriatrics*, vol. 17, no. 1, p. 77, 2017.

[45] B. E. Maki, K. M. Sibley, S. B. Jaglal et al., "Reducing fall risk by improving balance control: development, evaluation and knowledge-translation of new approaches," *Journal of Safety Research*, vol. 42, no. 6, pp. 473–485, 2011.

[46] I. Pastor, H. A. Hayes, and S. J. Bamberg, "A feasibility study of an upper limb rehabilitation system using Kinect and computer games," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 1286–1289, IEEE, Piscataway, NJ, USA, 2012.

[47] T. A. Türkbey, Ş. Kutlay, and H. Gök, "Clinical feasibility of Xbox KinectTM training for stroke rehabilitation: a single-blind randomized controlled pilot study," *Journal of Rehabilitation Medicine*, vol. 49, no. 1, pp. 22–29, 2017.

[48] Z. Wu, J. Li, and Y. L. Theng, "Examining the influencing factors of exercise intention among older adults: a controlled study between exergame and traditional exercise," *Cyberpsychology, Behavior, and Social Networking*, vol. 18, no. 9, pp. 521–527, 2015.

[49] M. E. Tinetti, C. F. Mendes De Leon, J. T. Doucette, and D. I. Baker, "Fear of falling and fall-related efficacy in relationship to functioning among community-living elders," *Journal of Gerontology*, vol. 49, no. 3, pp. M140–M147, 1994.

[50] J. Murphy and B. Isaacs, "The post-fall syndrome: a study of 36 elderly patients," *Gerontology*, vol. 28, no. 4, pp. 265–270, 1982.

[51] K. Marivan, C. Boully, S. Benveniste et al., "Rehabilitation of the psychomotor consequences of falling in an elderly population: a pilot study to evaluate feasibility and tolerability of virtual reality training," *Technology and Health Care*, vol. 24, no. 2, pp. 169–175, 2016.

[52] P. Wargnier, E. Phuong, K. Marivan et al., "Virtual Promenade: a new serious game for the rehabilitation of older adults with post-fall syndrome," in *Serious Games and Applications for Health (SeGAH), 2016 IEEE International Conference*, pp. 1–8, IEEE, Orlando, FL, USA, 2016.

Hindawi
Journal of Healthcare Engineering
Volume 2018, Article ID 8961781, 13 pages
https://doi.org/10.1155/2018/8961781

Research Article

Gastric Pathology Image Classification Using Stepwise Fine-Tuning for Deep Neural Networks

Jia Qu,[1] **Nobuyuki Hiruta**,[2] **Kensuke Terai**,[2] **Hirokazu Nosato**,[3] **Masahiro Murakawa**,[1,3] **and Hidenori Sakanashi**[1,3]

[1]*Department of Intelligent Interaction Technologies, University of Tsukuba, Tsukuba 305-8573, Japan*
[2]*Department of Surgical Pathology, Toho University Sakura Medical Center, Sakura 285-8741, Japan*
[3]*Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8560, Japan*

Correspondence should be addressed to Jia Qu; s1330219@u.tsukuba.ac.jp

Received 23 March 2018; Revised 14 May 2018; Accepted 27 May 2018; Published 21 June 2018

Academic Editor: Santosh K. Vipparthi

Deep learning using convolutional neural networks (CNNs) is a distinguished tool for many image classification tasks. Due to its outstanding robustness and generalization, it is also expected to play a key role to facilitate advanced computer-aided diagnosis (CAD) for pathology images. However, the shortage of well-annotated pathology image data for training deep neural networks has become a major issue at present because of the high-cost annotation upon pathologist's professional observation. Faced with this problem, transfer learning techniques are generally used to reinforcing the capacity of deep neural networks. In order to further boost the performance of the state-of-the-art deep neural networks and alleviate insufficiency of well-annotated data, this paper presents a novel stepwise fine-tuning-based deep learning scheme for gastric pathology image classification and establishes a new type of target-correlative intermediate datasets. Our proposed scheme is deemed capable of making the deep neural network imitating the pathologist's perception manner and of acquiring pathology-related knowledge in advance, but with very limited extra cost in data annotation. The experiments are conducted with both well-annotated gastric pathology data and the proposed target-correlative intermediate data on several state-of-the-art deep neural networks. The results congruously demonstrate the feasibility and superiority of our proposed scheme for boosting the classification performance.

1. Introduction

Cancer is acknowledged as one of the top threats to human health. According to the International Agency for Research on Cancer (IARC), in 2012, there were approximately 14.1 million new cancer cases and 8.2 million deaths around the world [1]. This number is estimated to increase to 24 million by 2035, and the deaths will continually rise. Within the diagnostic methods, although advanced image diagnosis devices (computed tomography (CT), magnetic resonance imaging (MRI), ultrasound (US)) are evolving rapidly, for many kinds of cancers, pathological diagnosis is still realized as the gold standard to assess cancer's presence or absence, type, and malignance degree. However, the shortage of pathologists represents as great restriction to pathological diagnosis and causes social problems. In the United States, the lack of pathologist workforce is become more and more concerned [2]. In Japan, the number of pathologist normalized by the general population is even smaller than one-third of that in the United States (one pathologist per 19,000 people) [3]. The situation is even more severe in China. As reported, China has approximately one pathologist per 74,000 people [4]. Such severe shortage is now consequently leading to immense working burden on pathologists and possible errors and oversights in diagnosis. Compared with 2005, the number of all pathological diagnosis cases and intraoperative pathological diagnosis cases in Japan has, respectively, risen up to 1.7 times and 3 times by 2012. New technologies such as digital pathology has widely spread and facilitated faster and cheaper diagnosis since more than

a decade ago, due to its operational ease enabled by virtual microscopy [5, 6]. Nevertheless, since diagnosis correctness and pathologist workload alleviation remain challenges, further assistance based on advanced image classification technologies is expected to play a key role to facilitate more advanced computer-aided pathology diagnosis.

In earlier periods, conventional classification methods for pathology images including specified histologically concerned features or generalized texture image features are commonly adopted. The specific histologically concerned features, such as nuclei's area and nuclear-cytoplasmic ratio (N/C), are subtly calculated from unknown images [7]. These features are compared with predefined criteria to judge whether the target image is benign or malignant. Unfortunately, such process usually meets a big issue that it is a hard task to make adequate definition for the morphological characteristics, because cancerous cells usually lack control for regular division. Thus, shape extraction failures for cells could become a direct reason for classification failures. In contrast, the schemes using generalized texture image features appear to sustain more robustness to various cancerous appearances. As a customary way, many of the previous studies taking advantage of generalized texture feature have demonstrated their capability for respective tasks. One of the focused texture feature is grey-level co-occurrence matrix (GLCM). For example, Esgiar et al. [8] employed GLCM to obtain texture features corresponding to contrast, entropy, angular second moment, dissimilarity, and correlation from colon biopsy and employed linear discriminate analysis (LDA) and k-nearest neighbour algorithm (KNN) to realize the categorization of normal and cancerous colon mucosa. Likely, Diamond et al. [9] employed Haralick features (a kind of texture features developed from GLCM) for identifying tissue abnormalities in prostate pathology images. Another mighty rival is local binary patterns (LBPs). In the study of Masood and Rajpoot [10], a scheme consisting of LBP and support vector machines (SVMs) is proposed and demonstrated effective for colon pathology images. In another work, Sertel et al. [11] developed classification methods for neuroblastoma H&E-stained whole-slide images, using co-occurrence statistics and local binary patterns similar to the above study. A recent report by Kather et al. [12] gave a relatively comprehensive investigation of texture analysis for colorectal cancer histology image. Besides LBP and GLCM, lower-order and higher-order histogram features, Gabor filters, and perception-like features are involved as well. In our earlier studies [13], another texture features called higher-order local autocorrelation (HLAC) bonded with linear statistical models such as principal component analysis (PCA)-based subspace method were also demonstrated capable of indicating the anomaly degree of gastric pathology images. Apart from straightforward benign/malignant classification, some other methods in pathology image domain have been put forward with texture features as well, to settle similar classification-correlative tasks such as gland segmentation and grade estimation [14–19].

While all of these texture-feature-based approaches shown promising feasibility, intractable issues still existed between the research and practical application. One particular instance is that confirming how suitable the hand-crafted geometric features are for certain tasks is quite difficult [20]. Meanwhile, the uneven H&E staining among images brings adverse impact on classification performance so that it makes the tasks more challenging [21–23]. Since the dominative victory of the team using deep learning at ImageNet Large Scale Visual Recognition Competition (ILSVRC) 2012, many of the image recognition techniques have been replaced by deep learning using convolutional neural networks (CNNs) [24]. Due to more domain agnostic approach combining both feature discovery and implementation to maximally discriminate between the classes of interest [25], deep learning shows unprecedented adaptability for various kinds of images [26–28]. Accordingly, high hope is placed on deep learning to exert great power in pathology image and other medical image analysis fields [29–33]. Specifically within the pathology image domain, in addition to aspiration for more precise classification and segmentation [22, 34–38], deep learning has also been inspired for new patulous applications, such as stain normalization [39], assessment of tumor proliferation [40], and comprehensive multimodal mapping between medical images and diagnostic reports [30].

In comparison with computer vision for natural images, scarcity of training data along with accurate annotations in the medical image field has currently become a primary issue. Due to the necessity of medical doctor's knowledge and collaboration, acquisition procedures usually cost both expensive financial resources and workload. In consideration to this problem, many of the studies have evidenced that transfer learning using fine-tuning techniques for deep neural networks can boost the performance and alleviate the scarcity of training data in some degree by transferring a general neural network pretrained by large-scale image datasets (such as ImageNet), to a more specified one corresponding to more complicated target tasks [22, 41–43]. However, since pretraining image datasets possess fixed categories of contents and image size, the coefficient efficiency of the datasets and deep neural network is generally suboptimal for problems encountered in specific image classification domain [44]. In many cases, there are still gaps between the knowledge gained from pretrained tasks and different specified target domains.

Hence, in this paper, aiming to further alleviate the scarcity of well-annotated training data for gastric pathology image classification and enhance the performance in a rational way, we have proposed a novel scheme adopting two-stage fine-tuning approach for CNNs and introduced a new type of target-correlative intermediate datasets (called "medium-level" datasets, hereinafter). In addition to the "low-level" large-scale pretraining datasets owning enormous amount but little target specificity, and the "high-level" well-annotated pathology datasets directly related to the target task, the proposed "medium-level" datasets are produced based on tissue-wise and cell-wise information within pathology image domain. With the "medium-level" datasets, our scheme is supposed capable of making the deep neural networks imitating the perception manner of pathologists and acquiring pathology-related knowledge in advance, but with very limited extra cost in data annotation. In the

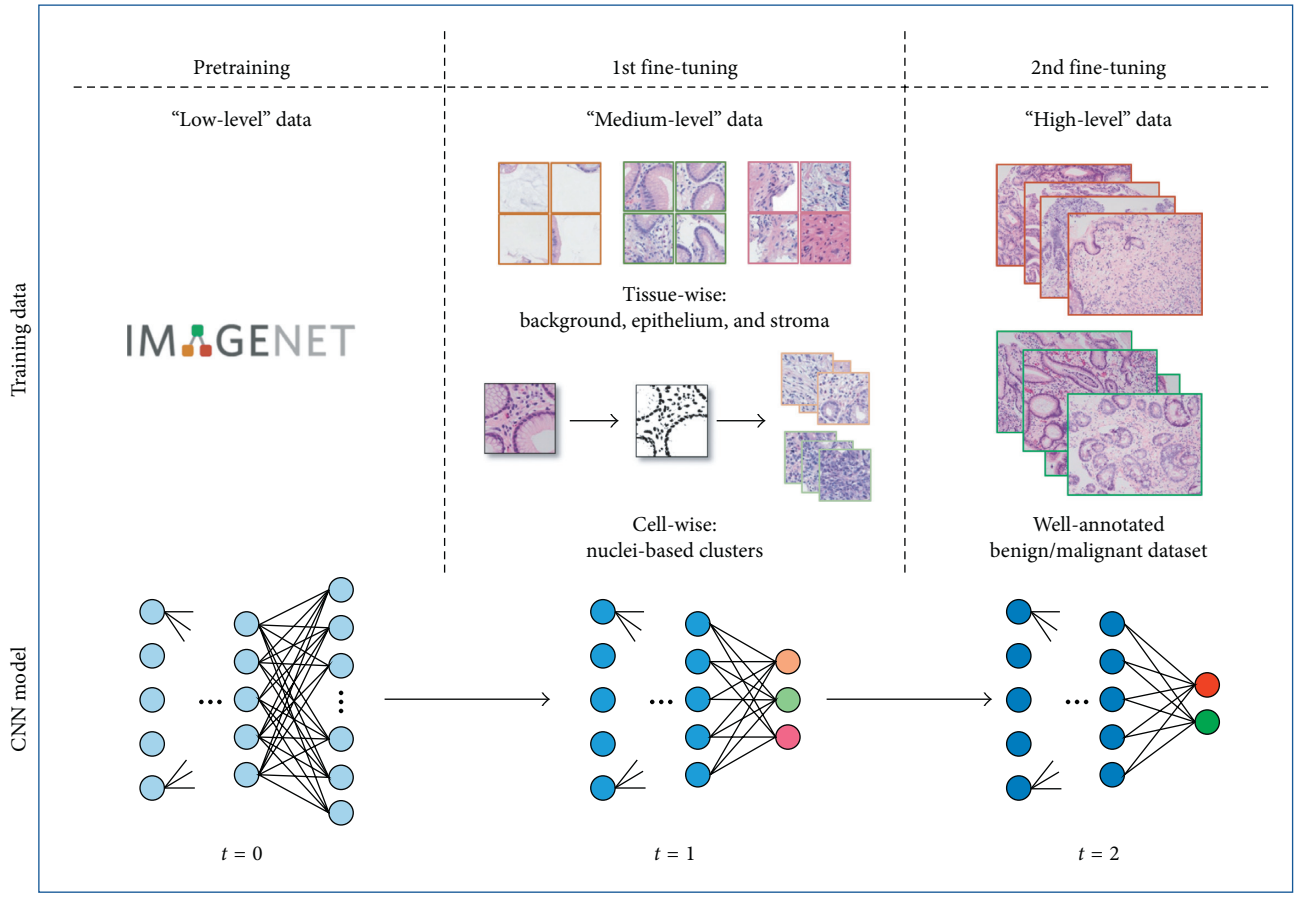


Figure 1: The proposed stepwise fine-tuning architecture. Apart from "low-level" pretraining datasets and well-annotated "high-level" datasets, "medium-level" data including tissue-wise data and cell-wise data are involved as well. In CNN models corresponding to training step $t = 0$, $t = 1$, and $t = 2$, blue nodes coloured darker denote more specified (deeper) representation for the pathology image classification task.

following parts, we will give the detailed materials of both the stepwise fine-tuning scheme and the datasets. After that, experiments will be organized to evidence our proposed scheme.

2. Materials and Methods

This section will be unfolded to several subtopics. Firstly, we will make a brief instruction of convolutional neural networks (CNNs) and the architectures adopted in this paper. Secondly, we will give some details about the proposed stepwise fine-tuning scheme (Figure 1) to clarify how it helps to improve CNN's performance and alleviate the insufficiency of well-annotated pathology training data. Thirdly, we will represent a feasible tissue-wise "medium-level" dataset including classes of background, epithelium, and stroma, which can be semiprofessionally annotated based on fundamental pathological knowledge with little pathologist's workload. Lastly, supposing to minimize the annotation cost and give an alternative solution to extremely insufficient annotating workforce, we are further inspired to put forward a solution by implementing full-automatic production of the cell-wise "medium-level" datasets which are yielded upon nuclei measurement and unsupervised clustering.

2.1. Convolutional Neural Network (CNN) Architecture. Among various approaches which have been studied in the field of image recognition and computer vision, convolutional neural network (CNN) is currently the most remarkable success. The prototype of CNN can be found in neocognitron [45] devised based on the neurophysiological findings on the visual cortex of living organism's brain. It is a neural network that alternately arranges a convolution layer corresponding to the cells for feature extraction and a pooling layer corresponding to the cells having a function to allow positional deviation hierarchically. Intuitively, it can be interpreted as a network that takes co-occurrence of adjacent features on different scales little by little and selectively gives information effective for identification to upper layers. Practically, refinement of such information is usually implemented by minimizing the cost function:

$$L(W) = \frac{1}{N}\sum_{i=1}^{N} l\left(y\left(x_i; W\right), y_i\right). \quad (1)$$

In (1), while $L(W)$ indicates the total cost (difference between prediction upon current configuration and the ground truth) over a dataset of N training samples, corresponding to weights W. y_i denotes the label of training data x_i. $y(x_i; W)$ is the predicted label of x_i, while l is the lost function. Practically, in order to evaluate the efficacy and generalization of the proposed stepwise fine-tuning scheme, in this paper, we employ several state-of-the-art CNN architectures including VGG-16 [28], AlexNet [24], and GoogLeNet (Inception V3, hereinafter) [46].

2.2. Making CNN Learn Pathology by Stepwise Fine-Tuning. The fine-tuning approach, which generally indicates retraining the pretrained CNN with the dataset corresponding to the target task, is widely adopted in image classification domain [47–52]. Since training a CNN strongly depends on initial parameters, it is significant to obtain appropriate parameter initialization as much as possible in order to prevent overfitted learning. Generally, the early layers within a CNN are in charge of acquiring relatively universal image features, which are considered analogous to the conventional texture features and applicable to many of related tasks, while the later layers are involving more specific information corresponding to the target task. Therefore, if the distributional difference between the datasets for pretraining and target is sufficiently correlated, one may fine-tune part of or all the layers to yield more desired results than those train from full scratch in many cases. However, in other situations if target tasks possess much different distribution compared with the pretraining datasets, effectiveness of initialization and fine-tuning may be largely restricted. Such issue is exactly posed in pathology image classification. In light of human's perception, pathology images usually have more complicated appearances than natural images (included in the pretraining dataset) since it is difficult to figure out the intuitionistic difference between benign and malignant images due to their color uniformity by H&E stain. From this perspective, we believe that it is necessary to make some substantial effort to fill up the "gap" between the tasks of pretraining dataset and well-annotated pathology dataset.

Therefore, in this paper, a novel conception is proposed if it is possible to make deep neural networks learn to understand pathology images in pathologist's way. With regard to the perception manner and learning progress of a pathologist, before drawing conclusions of benign or malignant for a pathology image, the pathologist should understand the difference of basic tissue-wise structures beforehand. Then, the pathologist may concentrate more on the detailed morphological characteristics, including but not limited to the spreading status and density of cells, degree of nucleus distortion, nucleus size, and nuclear-cytoplasmic ratio. On account of all of the above observations, the pathologist can finally make a benign or malignant judgement. To realize such conception, we build up a stepwise fine-tuning scheme to impart the intellectual pathology knowledge on deep neural networks, and introduce a new type of target-correlative "medium-level" dataset. Rather than directly driving the deep neural networks to learn how to identify benign or malignant, making it gain fundamental pathological knowledge from the "medium-level" datasets probably contributes to the task of higher difficulty. Therefore, based upon the in-being framework, training with the "medium-level" datasets is placed in the middle of the stepwise fine-tuning scheme [53]. As shown in Figure 1, at the beginning of the training progress $t = 0$, we have a pretrained network as initialization. The following step $t = 1$ is fine-tuning the network with our proposed "medium-level" dataset. After this step, the deep neural network is considered to possess more pathology-related weights. Finally, well-annotated benign/malignant images are used for the second time ($t = 2$) fine-tuning. In the lower part of the figure, corresponding to training step $t = 0$, $t = 1$, and $t = 2$, blue nodes coloured darker in CNN models denote more specified (deeper) representation for the pathology image classification task. When the number of output classes changes, definitely, the network architecture needs to be adjusted accordingly.

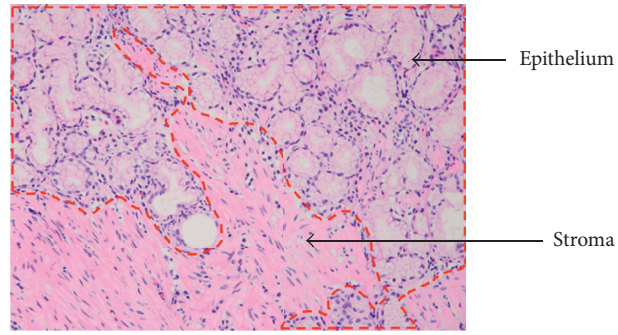


Figure 2: Stroma tissue and epithelium tissue appear in gastric pathology image.

2.3. Building Up a Low-Cost "Medium-Level" Dataset. As for the "medium-level" datasets, it is supposed necessary to meeting the prerequisite requirement that they could be acquired at much lower cost than the datasets annotated by pathologists and meanwhile charged with fundamental pathology related knowledge. Therefore, we firstly pay attention to seeking for a kind of pathology image datasets which can be made by "nonprofessional" workforce. In this paper, we present a feasible stroma-epithelium datasets which can be made by ourselves with only a little pathologist's direction. Epithelium and stroma are two tissue types that can be found in every organ [54]. In gastric pathology domain, epithelial tissues line the outer surfaces of gastric mucosa, while stroma tissue locates right under epithelium. Since cancer metastasis between epithelium and stroma is deemed as inextricable to cancer's progression [55], both of the two types of tissue are usually extracted during biopsy examinations and revealed in the pathology images. Although it is quite difficult to identify if the two types of tissue are in order for nonprofessional workers, actually, we find that stating the visual difference between them is quite an undemanding work. As shown in Figure 2, the upper area encircled by dashed contour indicates epithelium tissue, while the remaining lower area denotes

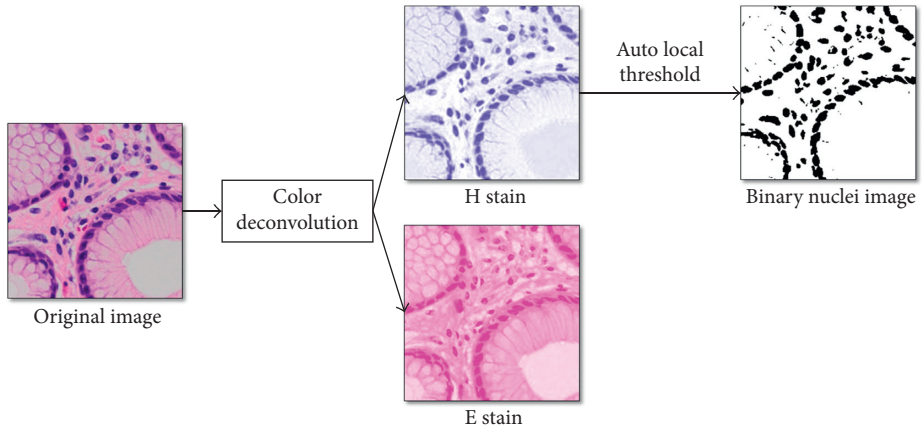


Figure 3: Image processing for making binary nuclei images.

stroma tissue. Because of their respective functions, obviously, epithelium tissue appears as organized arrangement, but stroma tissue seems scattered and disordered. Such distinct difference is considered quite beneficial to our "non-professional" work. Without pathologist's expensive annotation, it would be encouraging if these epithelium and stroma images can impart the pathology knowledge to the deep neural networks on the basis of our assumption. Practically, we have managed to collect a large amount of images including epithelium and stroma, segment them with stroma area and epithelium and manually, and separate into three classes (epithelium, stroma, and background).

2.4. Automatically Generating a "Medium-Level" Dataset. In this section, this paper moves one step further and presents another full-automatic "medium-level" dataset generating approach, corresponding to the situation when data annotation workforce is extremely insufficient. The proposed approach firstly applies a string of image processing techniques, before extracting low-dimension feature vectors indexing the measurement of morphological status of nuclei. Finally, we perform unsupervised clustering in order to form the "medium-level" dataset.

In order to realize the measurement of nuclei, it is necessary to generate a binary nuclei image instead of the original one. Thus, the first step is to implement color deconvolution to separate H&E- (hematoxylin and eosin-) stained pathology image into H and E component, since nuclei are usually richly visualized in H component. This paper adopts the algorithm proposed by Ruifrok and Johnston [56]. According to their work, although relative intensity in each of RGB channels depends on the concentration of stain in a nonlinear way, the optical density, $OD_c = -\log 10\,(I_c/I_0, c) = A * b_c$ (c is any of the RGB channels), is linear with the concentration of amount of stain A with absorption factor b_c and therefore can be used for separation of the contribution of multiple stains. Practically, I_c is the intensity of one of RGB channels and I_0, c is valued 1. Then, with the deconvolution matrix D denoted as (2), combination of hematoxylin, eosin, and DAB can be calculated as $C = D[y]$, where $[y]$ indicates vector of optical density in RGB space:

$$D = \begin{bmatrix} 1.88 & -0.07 & -0.60 \\ -1.02 & 1.13 & -0.48 \\ -0.55 & -0.13 & 1.57 \end{bmatrix}. \quad (2)$$

Next, we perform binarization on H (hematoxylin) channel, with the auto local threshold algorithm presented by Phansalkar et al. [57]. In this method, the threshold T is computed as

$$T = \mu^* \left(1 + p^* e^{-q^* \mu} + k^* \left(\left(\frac{\sigma}{r}\right) - 1\right)\right). \quad (3)$$

In (3), μ and σ are the local mean and standard deviation, respectively. $k = 0.25$, $r = 0.5$, $p = 2$, and $q = 10$ are constants with their recommended values. The image on the right in Figure 3 shows the result after banalization.

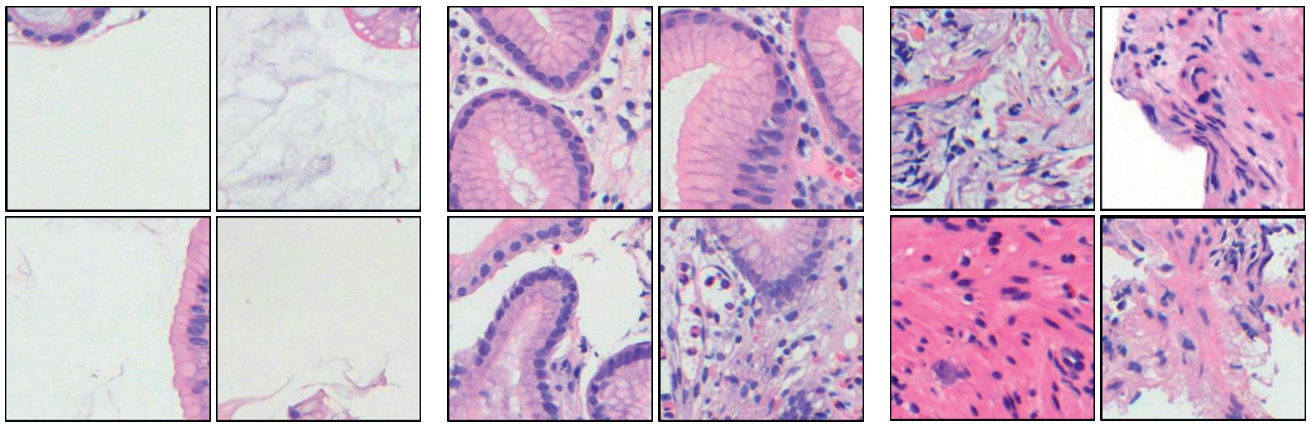
Afterwards, our proposed method adopted watershed algorithm for separating the conjoined or partly overlapped nuclei and then the contour detection method was employed, developed by Suzuki and Abe [58], in order to acquire a two-dimensional vector of total area and number of nuclei from each pathology image. After all, unsupervised K-means clustering is utilized to separate images into two classes. The clustered images are directly used as "medium-level" training data in the first time fine-tuning. Compared with tissue-wise dataset, the automatically generated dataset concentrates on cell-wise characteristics.

3. Results and Discussion

3.1. Experimental Procedures. In order to evaluate the effectiveness of our proposed stepwise fine-tuning scheme and the low-cost "medium-level" datasets, the experiments are progressively arranged in the following order. (1) We firstly give discussion on the performances of the two-stage fine-tuning scheme based on different CNN architectures when tissue-wise dataset is employed. (2) We then replace the tissue-wise dataset with automatically generated cell-wise dataset and validate the proposed scheme in the same way as (1). Separately in (1) and (2), we will further talk over how our proposed two-stage scheme performs upon

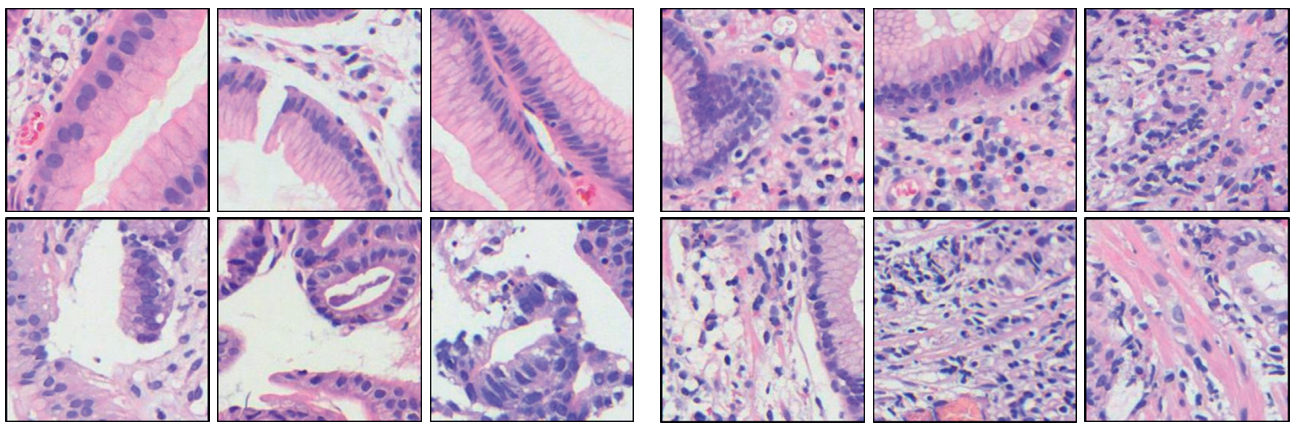
Table 1: Datasets used in experiments.

Data type	Category	Training	Validation	Test
Tissue-wise data	Background	15,000	1,000	—
	Epithelium	15,000	1,000	—
	Stroma	15,000	1,000	—
Cell-wise (nuclei) data	Cluster 1	6,905	767	—
	Cluster 2	5,811	646	—
Well-augmented data	Small	540 + 540	1620 + 1620	2,700 + 2,700
	Large	5400 + 5400		



(a) (b) (c)

Figure 4: "Nonprofessionally" annotated tissue-wise datasets. (a) Background. (b) Epithelium. (c) Stroma.



(a) (b)

Figure 5: Full-automatically generated cell-wise (nuclei) datasets. (a) Cluster 1. (b) Cluster 2.

well-annotated datasets in different sizes, in order to produce more evidence to support our conception.

3.2. Datasets. This paper employs three types of data including "low-level," "medium-level," and "high-level" data (Figure 1), respectively, used for the initialization, the first-stage fine-tuning, and the second-stage fine-tuning. In practice, ImageNet data [59] containing approximately 1.2 million images in 1,000 separate categories are customary utilized to initialize the CNN models. As to gastric pathology images, all the datasets utilized are illustrated in Table 1. By depicting maps of epithelium and stroma, we succeeded in collecting 48,000 tissue-wise patches (256×256) separated into "background," "epithelium," and "stroma" categories (Figure 4). In each category, 15,000 patches are used as training data, while the remained 1,000 patches are used for validation. From this dataset, we further obtained automatically generated cell-wise dataset in line with nuclei measurement. The cell-wise dataset consists of 7,672 and 6,457 patches in two clusters (Figure 5), after excluded those having too few nuclei or difficult to be extracted due to stain

Table 2: Performances of the proposed two-stage fine-tuning using tissue-wise data.

Data size	Scheme	CNN architecture											
		VGG-16				AlexNet				GoogLeNet (Inception V3)			
		AUC	ACC	Precision	Recall	AUC	ACC	Precision	Recall	AUC	ACC	Precision	Recall
Small	One stage	0.879	0.793	0.863	0.695	0.828	0.723	0.74	0.72	0.838	0.753	0.75	0.75
	Two stage (tissue)	0.914	0.829	0.865	0.781	0.844	0.761	0.76	0.76	0.877	0.772	0.79	0.77
Large	One stage	0.936	0.836	0.957	0.703	0.867	0.794	0.80	0.79	0.881	0.779	0.79	0.78
	Two stage (tissue)	0.963	0.881	0.869	0.898	0.920	0.837	0.84	0.84	0.934	0.862	0.86	0.86

inconformity. Experimentally, 90% of patches out of each cluster are used for training and the rest are used for validation. As shown in Figure 5, according to rough observation, the cluster on the left appears to have larger and fewer nuclei, while the cluster on the right side seems opposite with more but smaller nuclei. As to the well-annotated "high-level" datasets, in order to evaluate the efficacy and generalization of the proposed two-stage scheme, we have prepared well-annotated datasets in two different sizes. One is a small train dataset including 540 benign and 540 malignant patches. Another one is a nonaugmented large dataset of 5,400 benign and 5,400 malignant patches within which the small dataset is included. Except from the former datasets, we additionally use a validation dataset including 1,620 benign and 1,620 malignant patches to select the best model configuration, and a test dataset of 2,700 benign and 2,700 malignant patches to finally evaluate the performance in each optional case. It is noteworthy that there is no overlap between the "medium-level" datasets and the "high-level" datasets and no overlap among the training, validation, and test datasets.

4. Experimental Results

4.1. Stepwise Fine-Tuning Using Tissue-Wise Data. In this part, we will specifically discuss about our proposed stepwise fine-tuning when tissue-wise data are employed for the first-stage fine-tuning and well-annotated datasets are used for the second-stage fine-tuning. The performances are collected from the experiments performed with well-annotated pathology image datasets in different sizes and different deep neural network architectures. In this paper, we concurrently adopt AUC, ACC, precision, and recall as the evaluation criteria [60].

In Table 2, notably, in all of the couples of rival schemes, our proposed two-stage fine-tuning using tissue-wise data has yield promotion. Specifically, in the results of small data group, AUC value is raised by 0.035, 0.016, and 0.039, when we adopt CNN architectures VGG-16, AlexNet, and Inception V3, respectively. In the large data group, the corresponding AUC values are raised by 0.027, 0.053, and 0.053. Although the performance using smaller training data is expected to be more boosted, according to AUC values, we find that our proposed scheme has actually brought slightly more benefit to the large well-annotated dataset groups. Meanwhile, if we focus on ACC values, we are aware of the fact that the greatest improvement happens when our proposed scheme using Inception V3 is adopted upon the large dataset. The accuracy has remarkably increased from 0.779 to 0.862. Besides, precision and recall, which are commonly used for medical image classification, are appearing with the similar trend to AUC and ACC. As more intuitively illustrated in Figure 6, three CNN architectures combined with two datasets have produced six ROC Figures. The red curve denotes the two-stage scheme using tissue-wise "medium-level" dataset, while the green curve denotes the conventional one-stage scheme. It is clear at a glance, in each figure, that our proposed scheme possesses overwhelming area all along both the false-positive rate axis and true-positive rate axis. These results have proved that our proposed scheme is adaptable and rarely dependent on the amount of well-annotated data. In addition, a set of filtered response images exported by the stepwise trained network are displayed in Figure 7. Considering the practical and intuitive facility, we investigate on VGG-16 model and obtain some of the outcomes of the first convolutional layers from all of the convolutional blocks as shown.

In line with the common sense, larger training data yield larger absolute AUC value. Nevertheless, if we make comparison between the data size crosswise, we can observe that when we use the proposed two-stage scheme with the small dataset, our scheme has actually boosted the performance up to the level approaching to that when only one-stage fine-tuning is implemented with the large dataset which is 10 times that of the small one. By viewing the two rows "Small-Two stage (tissue)"and "Large-One stage" (right the two rows in the middle of Table 2), it is not hard to see that the differences of AUC values between the two rows are no larger than 0.023 (AlexNet), while the largest difference of ACC values is only 0.033 (AlexNet). These results can fully prove that the introduction of tissue-wise information has indeed contributed to making the deep neural networks "understand" pathology images better. Hence, it is reasoned to infer that our proposed two-stage fine-tuning scheme can be used as an alternative method to boost the performance when the number of well-annotated pathology image data is limited, but "nonprofessional" annotation is practicable in some degree.

4.2. Stepwise Fine-Tuning Using Cell-Wise Data. In the next step, we will follow the same procedures as above but substitute the cell-wise data which have been generated by our proposed automatic clustering approach for the tissue-wise data. To make a comprehensive comparison of the rival schemes, again in Table 3, we list up the results of the two schemes covering two sizes of well-annotated datasets and

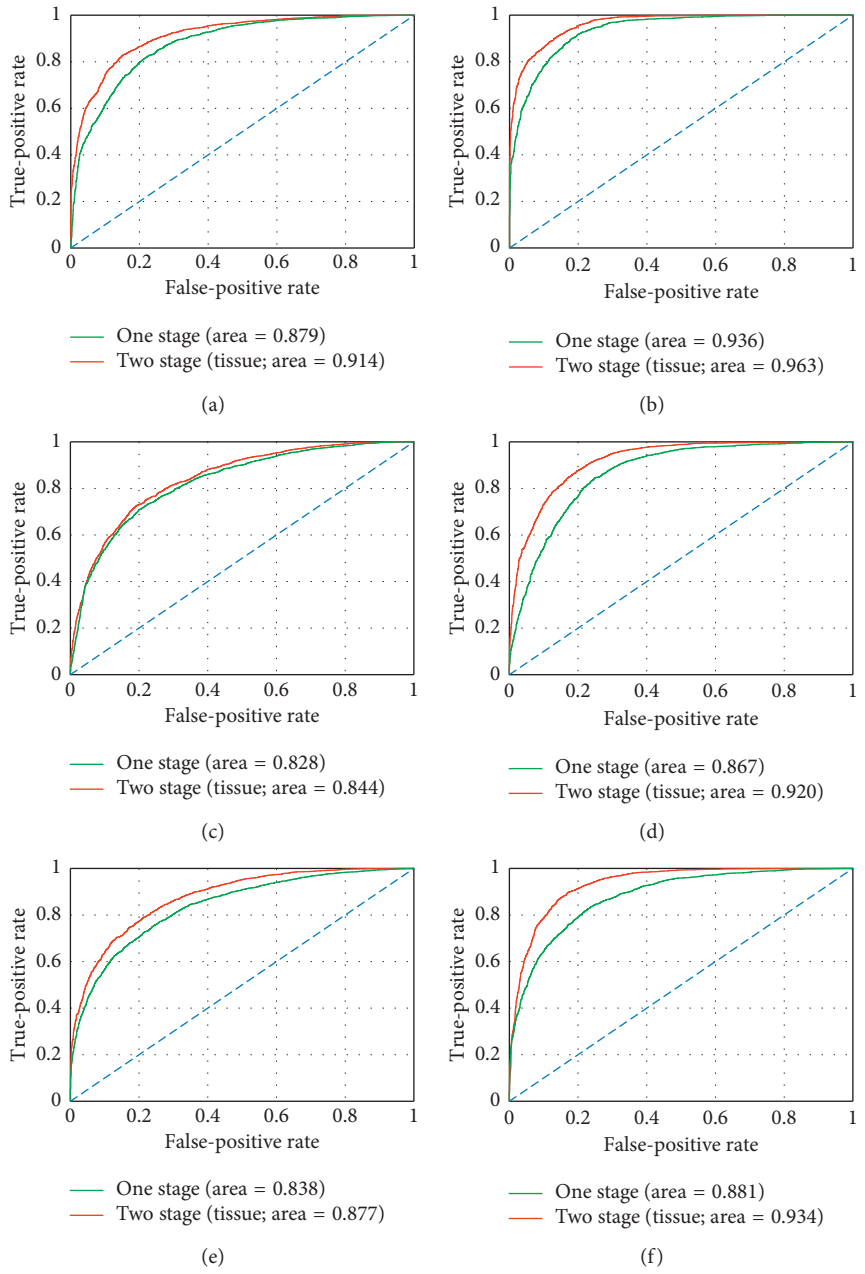


Figure 6: Performances of the proposed two-stage fine-tuning using tissue-wise data presented by ROC. (a) VGG-16 + small data. (b) VGG-16 + big data. (c) AlexNet + small data. (d) AlexNet + big data. (e) Inception V3 + small data. (f) Inception V3 + big data.

three types of CNN architectures. From Table 3, we can find that the proposed stepwise fine-tuning scheme using cell-wise (nuclei) datasets as well governs all couples of results. According to the AUC values, when small well-annotated dataset is employed, the fine-tuning by nuclei datasets has led to improvement by 0.023 (VGG-16), 0.024 (AlexNet), and 0.034 (Inception V3), respectively, along the horizontal direction. In the large data group, AUC values for the three networks are separately elevated by 0.029 (VGG-16), 0.048 (AlexNet), and 0.052 (Inception V3). In agreement with AUC values, ACC values have been largely raised after we introduce the nuclei dataset into the first fine-tuning stage, with whichever CNN architecture and well-annotated dataset. As to the indexes precision and recall, in each couple of the small data group, our proposed scheme yields enhancement for both precision and recall. However, the results depicted in the large dataset group show decline of precision, but on the contrary, substantial increase of recall as offset. In line with the definitions, a larger recall indicates less undetected anomaly (false negative), while a larger prevision indicates less overdetection (false positive). Practically, due to the high risk of undetected anomaly, a high recall is acceptable rather than a loss of it. In Figure 8, the ROC performances of the two-stage fine-tuning using nuclei "medium-level" dataset are presented. Likewise, we notice that our proposed two-stage stepwise fine-tuning

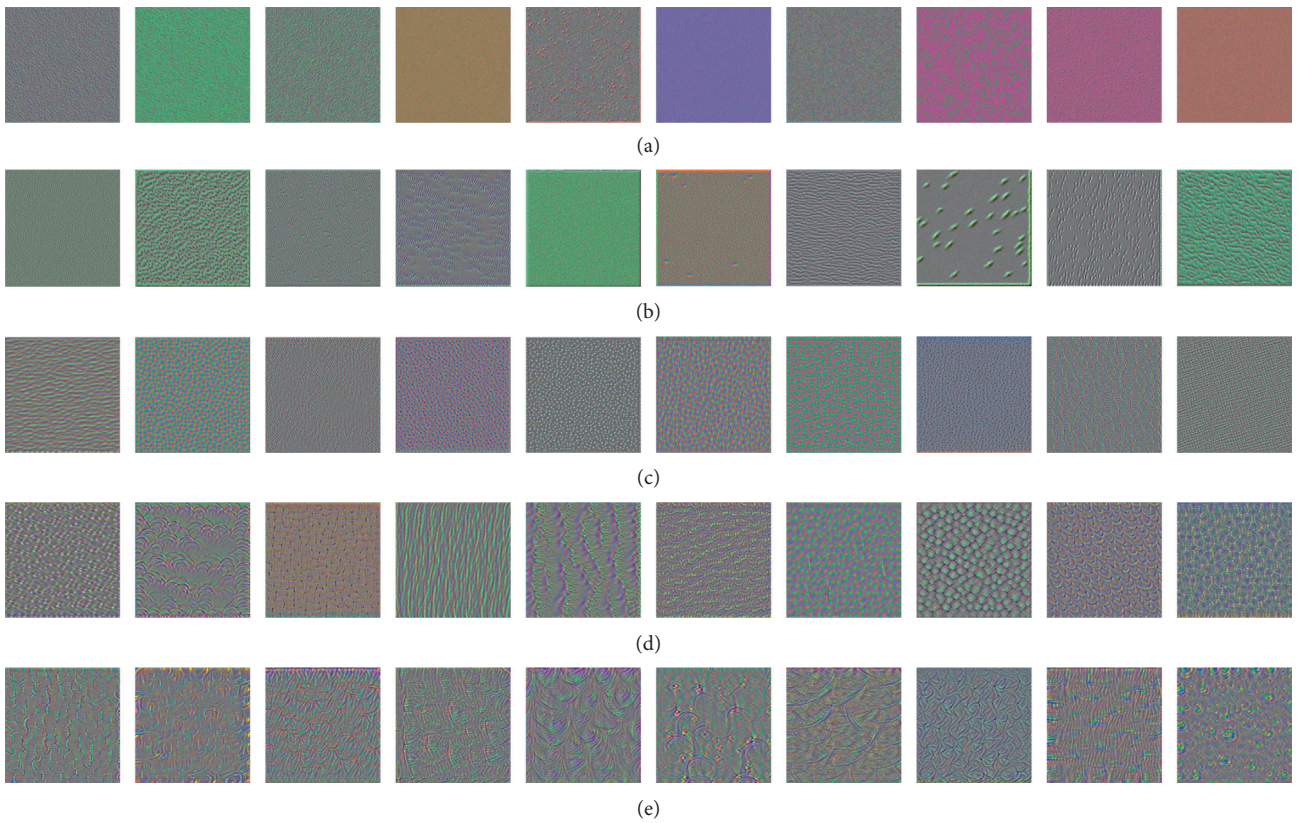


Figure 7: A set of filtered response images outputted by the stepwise trained VGG-16 model. (a) Block 1, Convl. (b) Block 2, Convl. (c) Block 3, Convl. (d) Block 4, Convl. (e) Block 5, Convl.

Table 3: Performances of the proposed two-stage fine-tuning using cell-wise (nuclei) data.

Data size	Scheme	CNN architecture											
		VGG-16				AlexNet				GoogLeNet (Inception V3)			
		AUC	ACC	Precision	Recall	AUC	ACC	Precision	Recall	AUC	ACC	Precision	Recall
Small	One stage	0.879	0.793	0.863	0.695	0.828	0.723	0.74	0.72	0.838	0.753	0.75	0.75
	Two stage (nuclei)	0.902	0.815	0.873	0.730	0.852	0.777	0.78	0.78	0.872	0.784	0.78	0.78
Large	One stage	0.936	0.836	0.957	0.703	0.867	0.794	0.80	0.79	0.881	0.779	0.79	0.78
	Two stage (nuclei)	0.965	0.89	0.881	0.901	0.915	0.834	0.84	0.83	0.933	0.862	0.86	0.86

scheme oversteps the state-of-the-art one-stage fine-tuning approach regardless of the scale of "high-level" dataset. By implementing the two-stage fine-tuning scheme, the performances of the state-of-the-art deep neural networks are all boosted up to a competitive level close to those achieved by the usual one-stage fine-tuning with a much larger well-annotated "high-level" dataset.

Moreover, if we compare with the two types of "medium-level" datasets utilized in the first-stage fine-tuning, there are actually hardly apparent differences observed between them. When we have a small well-annotated dataset, tissue-wise data seem to promote more for VGG-16 and Inception V3 models, while the AlexNet seems to profit more from the cell-wise data. Correspondingly, when we have a large well-annotated dataset, AlexNet and Inception V3 are both reinforced more effectively. Therefore, to sum up, all of these results have demonstrated that the proposed stepwise fine-tuning employing either of the tissue-wise dataset or the cell-wise dataset has successfully boosted the performance of the pretrained neural networks for gastric pathology image classification in various situations.

5. Conclusions

In this paper, aiming to alleviate the insufficient well-annotated training data for pathology image classification, we proposed a novel stepwise fine-tuning scheme and a kind of low-cost target-correlative intermediate data with which the deep neural networks are supposed able to acquire fundamental pathological knowledge in accordance with pathologist's perception manner. To match the demand of different data production capabilities, we proposed two practical approaches to build up the intermediate data. Both

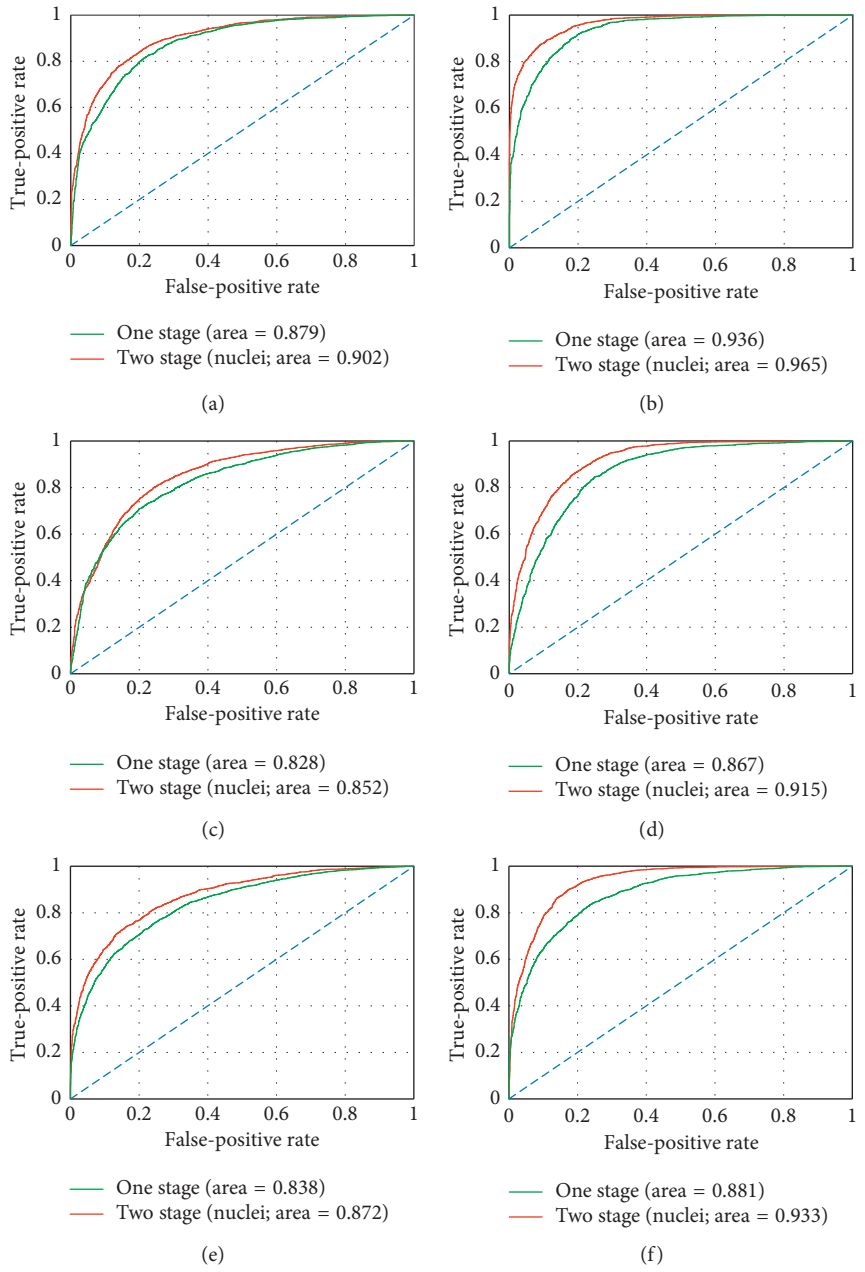


Figure 8: Performances of the proposed two-stage fine-tuning using cell-wise data presented by ROC. (a) VGG-16 + small data. (b) VGG-16 + big data. (c) AlexNet + small data. (d) AlexNet + big data. (e) Inception V3 + small data. (f) Inception V3 + big data.

of the two approaches incur no additional pathologist's workload. In the experiments, our proposed scheme exerted adequate efficacy for boosting the classification performance in respect of different possession situations of well-annotated pathology image data, and revealed high applicability for different state-of-the-art CNN architectures.

Last but not least, in this paper we regarded the final classification performance as the only assessment standard. To be objective, the effectiveness of the target-correlative intermediate data applied to the deep neural networks is expected to be concretely evaluated in the future work. Moreover, there is still necessity to discuss about the difference between the training framework using fine-tuning and other promising techniques such as multitask learning. Taking the proposed scheme as seed, we are looking forward to seeing that such kind of training scheme using low-cost target-correlative data can suggest solutions for more medical image classification tasks.

Data Availability

The data used in this paper were provided by Toho University Sakura Medical Center and are now currently managed by the National Institute of Advanced Industrial Science and Technology (AIST). However, as to the agreement on data usage, Toho University Sakura Medical Center only gained

consent from the patients "The anonymized image data are provided to AIST and can be used for image analysis experiments", which indicates only the right "to use." Meanwhile, at Toho University Sakura Medical Center, a research plan to provide data only to AIST was approved, so it is not permissible to provide data from AIST to the outside under the existing agreement framework. In order to provide data to a third-party organization in any way, after acquiring consent from the patients and preparing an experiment plan specifying the provision of providing data to the outside, it is necessary to obtain the approval of the Ethics Review Committee of both institutions. Since it is difficult to carry out all of these procedures this time, unfortunately, we cannot make the data publicly available in the prescribed form.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] J. Ferlay, I. Soerjomataram, R. Dikshit et al., "Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012," *International Journal of Cancer*, vol. 136, no. 5, pp. E359–E386, 2015.

[2] S. J. Robboy, S. Weintraub, A. E. Horvath et al., "Pathologist workforce in the United States," *Archives of Pathology and Laboratory Medicine*, vol. 137, no. 12, pp. 1723–1732, 2013.

[3] The Japanese Society of Pathology, *The Japanese Society of Pathology Guideline*, The Japanese Society of Pathology, Japan, 2015.

[4] *How Telemedicine Answers Global Pathology Demands*, 2018, https://proscia.com/blog/2015/07/14/global-crisis-digital-solution.

[5] A. O. Morrison and J. M. Gardner, "Microscopic image photography techniques of the past, present, and future," *Archives of Pathology and Laboratory Medicine*, vol. 139, no. 12, pp. 1558–1564, 2015.

[6] P. W. Hamilton, Y. Wang, and S. J. Mccullough, "Virtual microscopy and digital pathology in training and education," *Acta Pathologica, Microbiologica et Immunologica Scandinavica*, vol. 120, no. 4, pp. 305–315, 2012.

[7] M. Jondet, R. Agoli-Agbo, and L. Dehennin, "Automatic measurement of epithelium differentiation and classification of cervical intraneoplasia by computerized image analysis," *Diagnostic Pathology*, vol. 5, no. 1, p. 7, 2010.

[8] A. N. Esgiar, R. Naguib, B. S., Sharif et al., "Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa," *IEEE Transactions on Information Technology in Biomedicine*, vol. 2, no. 3, pp. 197–203, 1998.

[9] J. Diamond, N. H. Anderson, P. H. Bartels et al., "The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia," *Human Pathology*, vol. 35, no. 9, pp. 1121–1131, 2004.

[10] K. Masood and N. Rajpoot, "Texture based classification of hyperspectral colon biopsy samples using CLBP," in *Proceedings of International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1011–1014, Boston, MA, USA, July 2009.

[11] O. Sertel, J. Kong, H. Shimada et al., "Computer-aided prognosis of neuroblastoma on whole-slide images: classification of stromal development," *Pattern Recognition*, vol. 42, no. 6, pp. 1093–1103, 2009.

[12] J. N. Kather, C.-A. Weis, F. Bianconi et al., "Multi-class texture analysis in colorectal cancer histology," *Scientific Reports*, vol. 6, article 27988, 2016.

[13] J. Qu, H. Nosato, H. Sakanashi et al., "Cancer detection from pathological images using higher-order local autocorrelation," in *Proceedings of IEEE 2012 11th International Conference on Signal Processing*, pp. 1198–1201, Beijing, China, October 2012.

[14] M. Peikari, J. Zubovits, G. Clarke et al., "Clustering analysis for semi-supervised learning improves classification performance of digital pathology," in *Proceedings of International Conference On Medical Image Computing and Computer Assisted Intervention 2015: Machine Learning in Medical Imaging*, pp. 263–270, Munich, Germany, October 2015.

[15] A. Tabesh, M. Teverovskiy, H. Y. Pang et al., "Multifeature prostate cancer diagnosis and gleason grading of histological images," *IEEE Transactions on Medical Imaging*, vol. 26, no. 10, pp. 1366–1378, 2007.

[16] K. Nguyen, A. Sarkar, and A. K. Jain, "Structure and context in prostatic gland segmentation and classification," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 115–123, Nice, France, October 2012.

[17] P.-W. Huang and C.-H. Lee, "Automatic classification for pathological prostate images based on fractal analysis," *IEEE Transactions on Medical Imaging*, vol. 28, no. 7, pp. 1037–1050, 2009.

[18] J. Qu, H. Nosato, H. Sakanashi et al., "Computational cancer detection of pathological images based on an optimization method for color-index local auto-correlation feature extraction," in *Proceedings of IEEE 11th International Symposium on Biomedical Imaging*, pp. 822–825, Beijing, China, May 2014.

[19] K. Sirinukunwattana, D. Snead, and N. M. Rajpoot, "A novel texture descriptor for detection of glandular structures in colon histology images," in *Proceedings of SPIE Medical Imaging 2015: Digital Pathology*, Orlando, FL, USA, March 2015.

[20] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual Review of Biomedical Engineering*, vol. 19, no. 1, pp. 221–248, 2017.

[21] R. Marée, "The need for careful data collection for pattern recognition in digital pathology," *Journal of Pathology Informatics*, vol. 8, no. 19, 2017.

[22] H. Chen, X. Qi, and L. Yu, "DCAN: deep contour-aware networks for accurate gland segmentation," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2487–2496, Las Vegas, NV, USA, July 2016.

[23] B. Bejnordi, G. Litjens, N. Timofeeva et al., "Stain specific standardization of whole-slide histopathological images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 2, pp. 404–415, 2016.

[24] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of 25th International Conference on Neural Information Processing Systems*, vol. 1, pp. 1097–1105, Lake Tahoe, NV, USA, December 2012.

[25] A. Janowczyk and A. Madabhushi, "Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases," *Journal of Pathology Informatics*, vol. 7, no. 29, 2016.

[26] K. He, X. Zhang, S. Ren et al., "Deep residual learning for image recognition," in *Proceedings of IEEE Conference on*

Computer Vision and Pattern Recognition, pp. 770–778, Las Vegas, NV, USA, 2016.

[27] T. Durand, T. Mordan, N. Thome et al., "WILDCAT: weakly supervised learning of Deep ConvNets for image classification, pointwise localization and segmentation," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 642–651, Salt Lake City, UT, USA, 2017.

[28] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of International Conference on Learning Representations*, San Diego, CA, USA, May 2015.

[29] N. Dhungel, G. Carneiro, and A. P. Bradley, "Deep learning and structured prediction for the segmentation of mass in mammograms," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 605–612, Munich, Germany, October 2015.

[30] Z. Zhang, Y. Xie, F. Xing et al., "MDNet: a semantically and visually interpretable medical image diagnosis network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6428–6436, Beijing, China, China 2017.

[31] M. Havaei, A. Davy, D. W. Farley et al., "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017.

[32] G. Carneiro, J. C. Nascimento, and A. Freitas, "The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 968–982, 2012.

[33] Y. Yuan and M. Q. Meng, "Deep learning for polyp recognition in wireless capsule endoscopy images," *Medical Physics*, vol. 44, no. 4, pp. 1379–1389, 2017.

[34] Y. Xu, Z. Jia, L. Wang et al., "Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features," *BMC Bioinformatics*, vol. 18, no. 1, p. 281, 2017.

[35] L. Hou, D. Samaras, and T. M. Kurc, "Patch-based convolutional neural network for whole slide tissue image classification," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2424–2433, Las Vegas, NV, USA, 2016.

[36] J. Xu, X. Luo, G. Wang et al., "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," *Neurocomputing*, vol. 191, pp. 214–223, 2016.

[37] K. Sirinukunwattana, J. P. W. Pluim, H. Chen et al., "Gland segmentation in colon histology images: the GlaS Challenge Contest," *Medical Image Analysis*, vol. 35, pp. 489–502, 2017.

[38] Z. Han, B. Wei, Y. Zheng et al., "Breast cancer multi-classification from histopathological images with structured deep learning model," *Scientific Reports*, vol. 7, no. 1, 2017.

[39] F. Ciompi, O. Gessinnk, B. E. Bejnordi et al., "The importance of stain normalization in colorectal tissue classification with convolutional networks," in *Proceedings of the IEEE International Symposium in Biomedical Imaging*, Melbourne, VIC, Australia, 2017.

[40] M. Shah, C. Rubadue, D. Suster et al., "Deep learning assessment of tumor proliferation in breast cancer histological images," http://arxiv.org/abs/1610.03467.

[41] H. Chen, Q. Dou, D. Ni et al., "Automatic fetal ultrasound standard plane detectionusing knowledge transferred recurrent neural networks," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 507–514, Granada, Spain, September 2015.

[42] H.-C. Shin, H. R. Roth, M. Gao et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[43] J. Yosinski, J. Clune, Y. Bengio et al., "How transferable are features in deep neural networks?," in *Proceedings of the Annual Conference on Neural Information Processing Systems*, pp. 3320–3328, Barcelona, Spain, 2014.

[44] S. Christodoulidis, M. Anthimopoulos, and L. Ebner, "Multi-source transfer learning with convolutional neural networks for lung pattern analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 76–84, 2017.

[45] K. Fukushima and S. Miyake, "Neocognitron: a new algorithm for pattern recognition tolerant of deformations and shifts in position," *Pattern Recognition*, vol. 15, no. 6, pp. 455–469, 1982.

[46] C. Szegedy, V. Vanhoucke, S. Ioffe et al., "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, Las Vegas Valley, NV, USA, July 2016.

[47] P. Agrawal, R. Girshick, and J. Malik, "Analyzing the performance of multilayer neural networks or object recognition," in *Proceedings of European Conference on Computer Vision*, pp. 329–344, Zurich, Switzerland, September 2014.

[48] A. Babenko, A. Slesarev, A. Chigorin et al., "Neural codes for image retrieval," in *Proceedings of European Conference on Computer Vision*, pp. 584–599, Zurich, Switzerland, September 2014.

[49] B. Hariharan, P. Arbeláez, R. Girshick et al., "Simultaneous detection and segmentation," in *Proceedings of European Conference on Computer Vision*, pp. 297–312, Zurich, Switzerland, September 2014.

[50] S. Branson, G. Van Horn, S. Belongie et al., "Improved bird species categorization using pose normalized deep convolutional nets," in *Proceedings of British Machine Vision Conference*, Nottingham, UK, 2014.

[51] R. Girshick, J. Donahue, T. Darrell et al., "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, June 2014.

[52] Z. Zhou, J. Shin, L. Zhang et al., "Fine-tuning convolutional neural networks for biomedical image analysis: actively and incrementally," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7340–7351, Madison, WI, USA, June 2017.

[53] A. Suzuki, S. Suzuki, S. Kido et al., "A 2-staged transfer learning method with deep convolutional neural network for diffuse lung disease analysis," in *Proceedings of International Forum on Medical Imaging in Asia*, pp. 160–163, Quebec City, QC, Canada, September 2017.

[54] M. Guarino, P. Micheli, and F. Pallotti, "Pathological relevance of epithelial and mesenchymal phenotype plasticity," *Pathology–Research and Practice*, vol. 195, no. 6, pp. 379–389, 1999.

[55] B. S. Wiseman and Z. Werb, "Stromal effects on mammary gland development and breast cancer," *Science*, vol. 296, no. 5570, pp. 1046–1049, 2002.

[56] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution," *Analytical and*

Quantitative Cytology and Histology, vol. 23, no. 4, pp. 291–299, 2001.

[57] N. Phansalkar, S. More, A. Sabale et al., "Adaptive local thresholding for detection of nuclei in diversity stained cytology images," in *Proceedings of the International Conference on Communications and Signal Processing*, pp. 218–220, Calicut, India, February 2011.

[58] S. Suzuki and K. Abe, "Topological structural analysis of digitized binary images by border following," *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, pp. 32–46, 1985.

[59] O. Russakovsky, J. Deng, H. Su et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[60] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing and Management*, vol. 45, no. 4, pp. 427–437, 2009.