Received: 18 August 2023 | Accepted: 8 January 2024

DOI: 10.1111/2041-210X.14293

Methods in Ecology and Evolution

RESEARCH ARTICLE

Deep learning- and image processing-based methods for automatic estimation of leaf herbivore damage

Zihui Wang[1] | Yuan Jiang[2] | Abdoulaye Baniré Diallo[3] | Steven W. Kembel[1]

[1]Département des sciences biologiques, Université du Québec à Montréal, Montréal, Québec, Canada

[2]School of Ecology, Sun Yat-sen University, Guangzhou, China

[3]Département d'informatique, Université du Québec à Montréal, Montréal, Québec, Canada

Correspondence
Zihui Wang
Email: wang.zihui@courrier.uqam.ca

Steven W. Kembel
Email: kembel.steven_w@uqam.ca

Funding information
Natural Sciences and Engineering Research Council of Canada; Canada Research Chairs; Chinese Scholarship Council

Handling Editor: Sarah Goslee

Abstract

1. Quantifying the intensity of leaf herbivory pressure is crucial for understanding the interaction between plants and herbivores in both applied and basic science. Visual estimates and digital analysis have been commonly used to estimate leaf herbivore damage but are time-consuming which limits the amount of data that can be collected and prevent answering big picture questions that require large-scale sampling of herbivory pressure. Recent developments in deep learning have provided a potential tool for automatic collection of ecological data from various sources. However, most applications have focused on identification and counting, and there is a lack of deep learning tools for quantitative estimation of leaf herbivore damage.
2. Here, we trained generative adversarial networks (GANs) to predict the intact status of damaged leaves and applied image processing technique to estimate the area and percentage of leaf damage. We first described procedures for collecting leaf images, training GAN models, predicting intact leaves and calculating leaf area, with a Python package provided to enable hands-on application of these procedures. Then, we collected a large leaf data set to train a universal deep learning model and developed an online app *HerbiEstim* to allow direct use of pretrained models to estimate herbivory damage of leaves. We tested these methods using both simulated and real leaf damage data.
3. The procedures provided in our study greatly improved the efficiency of leaf herbivore damage estimation. Our test demonstrated that the reconstruction of damaged leaf image resembled the ground-truth image with a similarity of 98.8%. The estimation of leaf herbivore damage exhibited a high accuracy with an averaged root mean square error of 1.6% and had a general applicability to different plant taxa and leaf shapes.
4. Overall, our work demonstrated the feasibility of applying deep learning techniques to quantify leaf herbivory intensity. The use of GANs allows automatic estimation of leaf damage, representing a major advantage of the method. The Python package and the online app with pre-trained models will facilitate the use of our method for the analysis of large data sets of plant–herbivore interactions.

KEYWORDS
data collection, deep learning, generative adversarial network, image processing, leaf herbivore damage

1 | INTRODUCTION

Herbivory of plants by insects is a key biotic interaction that has been intensively investigated in both basic and applied science. Numerous theories and applications have been established on the importance of insect herbivores in plant ecology, evolution, crop production and ecosystem functioning, such as the green world hypothesis in explaining plant biomass (Hairston et al., 1960), the Janzen-Connell hypothesis in maintaining plant diversity (Janzen, 1970), the enemy release hypothesis in species invasion (Keane & Crawley, 2002), evolutionary arms races in plant trait evolution (Brodie, 1999) and agricultural practices of pest monitoring and control (Myers & Sarfraz, 2015). Testing hypotheses related to herbivory and applying knowledge in agricultural practices relies on efficient and accurate estimation of herbivory, defined here as the degree to which herbivores consume plant leaves. However, current methods for quantifying leaf herbivore damage are largely manual, which is time-consuming and potentially biased (Getman-Pickering et al., 2020; Machado et al., 2016; Xirocostas et al., 2022). Automatic estimation of leaf herbivore damage would be extremely useful in this regard but remains challenging.

Early studies often estimated leaf herbivore damage by eye or using a transparent grid with leaf damage classified into predefined categories, for example, slightly (0%–25%), moderately (25%–50%) and heavily damaged (>50%) (Coley, 1983; Kogan et al., 1977). However, the data sets generated from visual estimates of herbivory are low resolution and potentially biased with low reproducibility, despite the development of tools that train researchers to reduce the bias of visual estimation (Xirocostas et al., 2022). Recent studies have taken advantage of digital image analysis to quantify the percentage of leaf damage, in which the leaf boundary was retraced (in the case that damages occur on leaf edges) and holes were refilled to estimate the area of damaged leaf tissue (Neves et al., 2014; Sam et al., 2020). This method is expected to provide an accurate estimate of leaf damage, but is limited by the operating time associating with image scanning and analysing, for example, 10–20 s for scanning and 40–70 s for image processing per leaf using ImageJ software (O'Neal et al., 2002). Although mobile software such as LeafByte and Bioleaf have been developed to improve the efficiency of image collection and processing, the method still relies on manual operation which has limited the ability to collect large data sets of herbivore intensity for tens of thousands of leaves or more (Getman-Pickering et al., 2020; Machado et al., 2016).

Recent advances in deep learning techniques have facilitated the collection of ecological data by automatically extracting information from various sources such as images and audio recordings (Christin et al., 2019). For example, deep learning using convolutional neural networks (CNNs) has been applied to identify and count plant and animal species from photos, which generated large data sets in a rapid and automatic way (Ferreira et al., 2020; Norouzzadeh et al., 2018; Tabak et al., 2019). Although there have been attempts to apply CNNs in the estimation of leaf damage (da Silva et al., 2019), the precision of these approaches was low because many features of CNNs were designed for classification rather than quantitative estimation. An alternative technique is generative adversarial networks (GANs), which has recently been applied to reconstruct damaged plant leaves (Hussein et al., 2021; Silva et al., 2022; Villacis-Llobet et al., 2020). Different from CNNs whose application mainly focuses on image classification, objective detection and segmentation tasks, GANs handle various computer vision tasks such as image to image translation, image synthesis and semantic image editing (Creswell et al., 2018). Specifically, image to image translation allows training models based on paired images (as input and output, respectively), which can be used to predict images that resemble the source image (Isola et al., 2017). GANs provided a potential solution for quantifying leaf herbivore damage by comparing leaf area of damaged versus reconstructed intact leaves.

Training GAN models requires a huge amount of training data of paired images that include damaged and intact leaves (Ferreira et al., 2020). Collecting damaged leaves and reconstructing the intact status to get training data can be time-consuming as one has to manually trace the damaged parts of the leaf and refill the holes to make it intact (O'Neal et al., 2002). An alternative way is to collect healthy leaves and simulate artificial leaf holes that resemble the leaf damage observed in reality (da Silva et al., 2019). This can be achieved by cutting intact healthy leaves with randomly shaped polygons and circles. Although such random cutting simplifies leaf damage patterns, it has been shown to be effective to train GAN models with good performance in reconstructing damaged leaves (Hussein et al., 2021). However, the collection of training data requires considering the variety of leaf shapes since the reconstruction of leaf boundaries is based on learning the leaf shape of plants (Creswell et al., 2018).

In this study, we used deep learning to reconstruct damaged leaves to their intact status and applied image processing techniques to estimate the area and percentage of leaf damage by herbivores. We first described a complete procedure including collection and preprocessing of leaf images, generating artificial leaf images, training GAN models, reconstructing damaged leaves and measuring leaf area. A Python package was introduced to enable hands-on application of this procedure to build researchers' own deep learning models. Then, we collected healthy leaves from 229 plant species that commonly occur in North America, east Asia and Europe and trained a universal GAN model based on the leaves of these species. We also developed an online Shiny app 'HerbiEstim' to enable the direct use of pretrained models in the estimation of leaf herbivore damage. We tested our methods using both artificially damaged leaves and real leaves collected in the field. Overall, we showed that image processing- and deep learning-based methods are a state-of-the-art solution for automatic quantification of leaf herbivore damage, featured with high efficiency, accuracy and reproducibility. The Python package and online Shiny app provided easy to use tools that will facilitate the collection of large data sets in studying leaf–herbivore interactions.

2 | MATERIALS AND METHODS

2.1 | Image processing- and deep learning-based procedures for estimating leaf herbivore damage

2.1.1 | Image processing improves the efficiency of image collection

Scanning or photographing multiple leaves in a single image (with no overlap among leaves) is much faster than taking single leaf images. We applied image processing techniques to automatically extract leaves from multiple leaf images into single leaf images in order to improve the efficiency of image collection. To do that, we first applied a blur process to dilute the salt-and-pepper style noise in the image and remove the impure background using Otsu's automatic thresholding (Bangare et al., 2015). Then, we performed canny edge detection based on automatically determined thresholds to find the external contours of individual leaves (Xu et al., 2017). The external contours depicted individual leaf segments which were extracted and assigned to separate images. We placed individual leaf objectives in the middle of a blank squared image with a constant margin to standardize the display of leaf images (Figure 1). The leaf processing was conducted with OpenCV (version 4.2.0) and Python (version 3.8.5).

2.1.2 | Generate artificial leaf damage as training data set

We used image processing to generate artificial leaf damage from healthy (intact) leaves by randomly removing leaf pixels under different scenarios (da Silva et al., 2019). The healthy leaf images were collected and standardized to a size of 256×256 pixels. We considered two simple cutting scenarios, polygons and circles, and assumed that the two kinds of damage can occur simultaneously on a healthy leaf (Hussein et al., 2021). For the polygonal damage, we first selected a random pixel from within the leaf boundary as the central point and a random radius ranging from 5 to 20 pixels. We defined a start angle as 0 to draw the first pixel based on the radius and trigonometry. We then changed the radius by adding a random value ranging from −4 to 5 pixels and also added the angle by a random value ranging from 10° to 40°. According to the new radius and angle, we drew another pixel. As such, we continued to modify the radius, add the angle and draw pixels until the angle exceeded 360°. We jointed all the pixels in lines sequentially to generate a polygon. This process was repeated a random number of times ranging from 1 to 6 to simulate leaf damage of different degrees. For the circular damage scenario, we selected a random number (ranging from 5 to 20) of leaf pixels and draw circles with random radius ranging from 1 to 5 pixels to simulate tiny but dense holes on leaves (Figure 2). The parameters used in this method can generate different levels of leaf damage that range from 0% to 80% with a frequency distribution resembling that of leaf damage in reality (Avila-Sakar et al., 2003).

2.1.3 | GAN models

GAN models are generative models characterized by training a pair of networks, namely a generator G and a discriminator D (Creswell et al., 2018). The G network generates 'forgery' images with the aim to make them realistic while the D network is responsible for discriminating if the 'forgery' image from G is different from the real one. The two networks are trained simultaneously and interactively, that is, the generator G keeps updating the ability to produce forgeries of better quality with the 'guide' from discriminator D until the forgery and real images are indistinguishable. Because the generator G has no direct access to real images, it generates images by mapping data from a latent space to the space of the image, expressed as $G{:}\, G(z) \rightarrow R^{|y|}$, where $z \in R^{|z|}$ represents a sample from the latent space (random noise vector) and $y \in R^{|y|}$ is a sample in the image space. However, in the case of image to image translation, the 'forgery' image is based on not only latent variable z but also the input image. Thus, we applied conditional GANs called *pix2pix*, an extension of GANs that include a label x as a parameter in the input of G, expressed as $G(z, x) \rightarrow R^{|y|}$, where x is the information content of input image (Isola et al., 2017).

To train the *pix2pix* model, the damaged and intact leaves are paired and combined as a single image with the damaged leaf on the right (denoted as x) and the intact leaf on the left (denoted as y). During the training process, the information on damaged leaf x is passed to generator G, which maps the noise data z conditional on the input image x and generates a fake image y'. The generated y' conditional on x (i.e. the $y' + x$ pair) is passed to discriminator D, which performs a x-conditional discrimination on y' and the real image y (i.e. compare the $y + x$ and $y' + x$ pairs). The interaction between G and D networks generates loss functions including the adversarial loss and the L1 loss that help improve the prediction ability of G until the generated fake image is indistinguishable from the real image (Figure 3).

2.1.4 | Estimation of leaf area and percentage of leaf damage

The trained *pix2pix* models can be used to reconstruct the intact status of plant leaves. To improve the accuracy of the reconstruction,

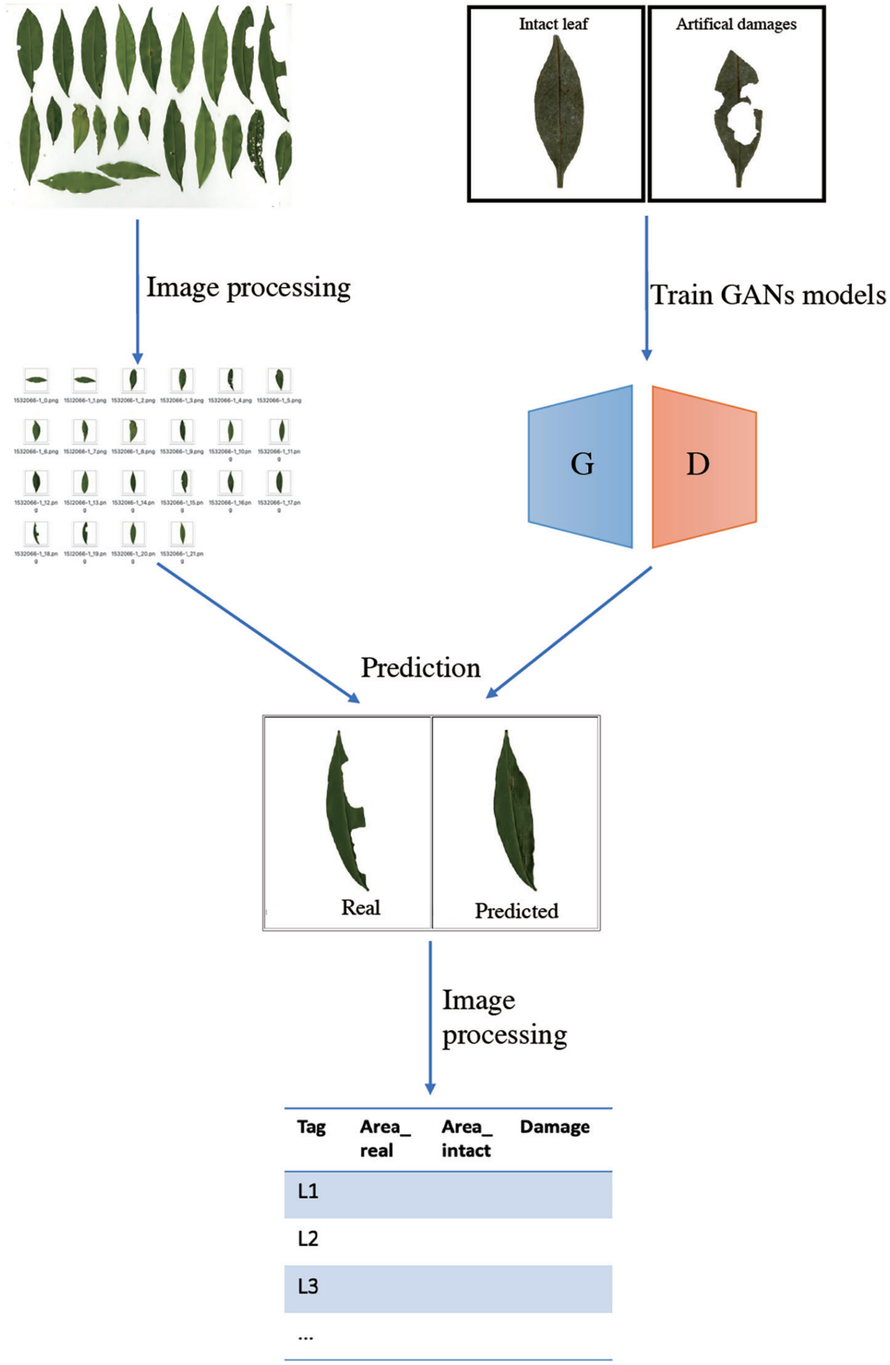


FIGURE 1 Overview of the complete procedure for leaf damage estimation using deep learning and image processing techniques.

we resized single leaf images to the same size as the training data (i.e. 256 × 256 pixels) and recorded the resized ratio for each leaf image. After reconstruction, we used image processing to calculate the number of leaf-belonging pixels for both damaged and reconstructed leaves, and multiplied it by the square of the resized ratio to obtain the size of pixels in the raw image. Based on the resolution of the scanned raw image, that is, dots per linear inch (dpi), we estimated the leaf area of damaged and reconstructed leaves and calculated the percentage of leaf area loss. The method can also be applied to photographed images without scales, in which case only the percentage of leaf damage is estimated.

2.2 | Application development

2.2.1 | A Python package

The Python package includes five main functions 'split', 'synthetic', 'train', 'predict' and 'calculation' (see Appendix S1; Wang, 2024). The 'split' function takes multiple leaf images as input, extracts leaves to separate images and returns standardized single leaf images. The 'synthetic' function can automatically generate training data which take intact leaves as input and return images with a pair of leaves including an intact leaf on the left and an artificially damaged leaf

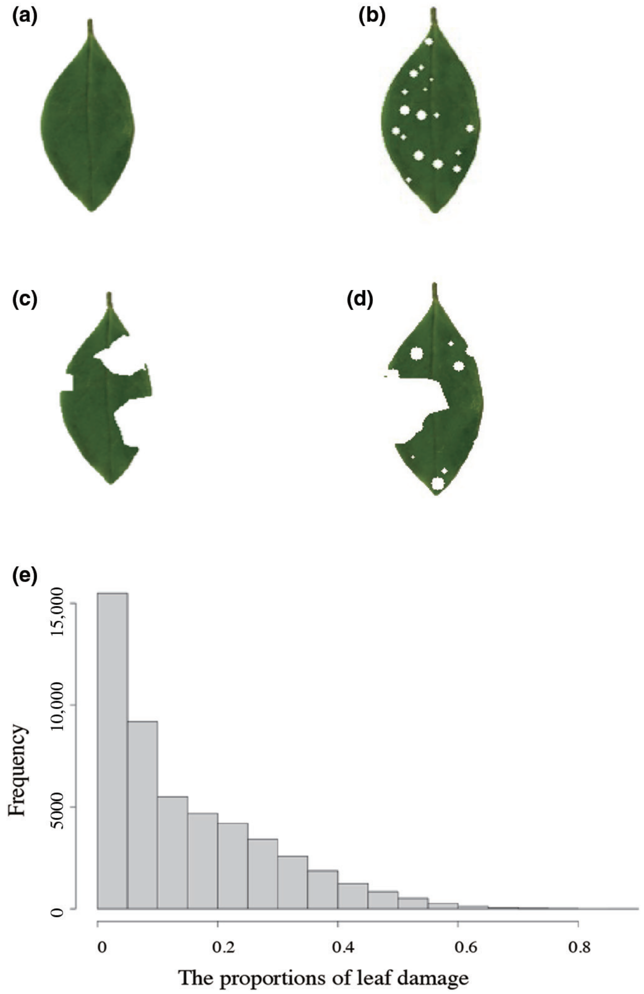


FIGURE 2 The simulation of artificially damaged leaves (a–d) and the frequency distribution of proportions of leaf damage (e). The damaged leaves were generated by cutting intact leaves (a) with random circles (b), polygons (c) and both (d).

on the right of the image. The 'calculation' function takes single leaf images as input, calculates the number of leaf-belonging pixels and estimates leaf area based on image resolution. This function returns a csv document that records the leaf area of both damaged and intact leaves, as well as the percentage of leaf damage. Lastly, the 'train' and 'predict' functions were used to launch pix2pix code developed by Zhu et al. (2017) to train the GAN model and predict the intact status of damaged leaves, respectively (Isola et al., 2017; Zhu et al., 2017).

2.2.2 | Pretrained deep learning models

The procedures and Python package allow researchers to train deep learning models based on training data from their study systems. To facilitate the general use of our methods, we collected intact plant leaves from the field and also from published plant leaf data sets including Swedish leaf data set (Söderkvist, 2001) and LeafSnap data set (Kumar et al., 2012), in order to train a universal model that can be generally applied to plant taxa and leaf shapes. A total of 6667 intact leaves from 229 plant species that commonly occur in North America, east Asia and Europe were collected to include a variety of leaf shapes and colours (i.e. Swedish data set: 300 images from 15 species; LeafSnap data set: 2701 images from 140 plant species; and our collected data: 3666 images from 74 plant species, also see Table S1). These leaves were visually checked to ensure clearness and intactness, and the leaves of each species were collected randomly to represent the diverse leaf shapes within plant species. The collected images from various databases differed in background, quality and size, and thus, we standardized the leaf images to the same size with blank backgrounds using image processing.

We took 80% of these leaf images (5337) to generate 50,000 randomly damaged leaves as training data to train a universal pix2pix model. The network weights were randomly initialized from a Gaussian distribution with a mean of 0 and a standard deviation of 0.02 with a batch size of 1. The generators and discriminators were trained using the Adam optimizer with a learning rate of 0.0002 and momentum of 0.5. The model was trained for 50 epochs at constant learning rate and another 50 epochs at linearly decaying learning rate until zero. Other network parameters were left at default settings (Zhu et al., 2017). The model training was implemented with PyTorch (version 2.0) in Python (version 3.8.4) on a high-performance computer equipped with an NVIDIA GeForce RTX 3060 GPU.

We also built plant species-specific models for five plant species: *Acer mono*, *Ormosia glaberrima*, *Quercus mongolica*, *Randia canthioides* and *Ulmus japonica*. For each species, we used 200 healthy intact leaves to simulate 5000 artificially damaged leaves as training data. The setting of training species-specific models was the same as the universal model.

2.2.3 | Online Shiny app: *HerbiEstim*

To make the methods accessible to users without powerful computing platform, we developed an online Shiny app *HerbiEstim* to estimate the leaf area and leaf damage based on pretrained deep learning models. In the app, users upload scanned or photographed leaf images and provide the dpi of images (if available) and the app will reconstruct the intact status for each leaf and calculate the area and percentage of leaf damage. After running, the app shows the actual and reconstructed leaves in pairs so that researchers can visually inspect the reconstruction of leaves to control data quality. The estimated leaf area and percentage of leaf damage are shown in the app and available for download in csv file format. The app estimates the percentage of leaf damage at a speed of 0.25 s per leaf. Apart from the universal model trained in our study, we also included case-specific pretrained models in the Shiny app, which are specific for leaf shapes or plant taxa. The Shiny app encourages researchers to share their pretrained models so that others that work with the same plant taxa or leaf shapes can potentially use these models.

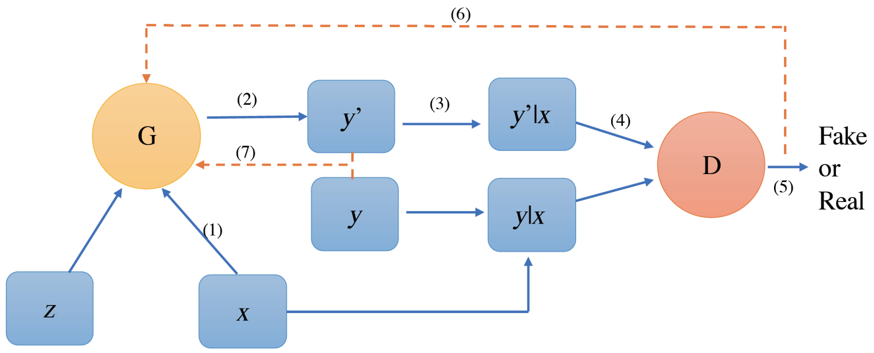


FIGURE 3 The overall structure of pix2pix model. The model is composed of two major elements, the generator G and the discriminator D. G is used to generate a fake image y' (process 2) from a latent space z (or noise source) based on the information of input images x (process 1, here x image represents the damaged leaf). Then, the real image y (y is the intact status of x leaf) and the fake image y' (generated by G to resemble an intact leaf) will be paired with x, respectively, leading to x-conditional y and y' (process 3). After that, the D will perform x-conditional discrimination on real image y and fake image y' (process 4) and have an evaluation on whether y and y' are distinguishable or not (process 5). Meantime, the D will compute adversarial loss, which will be used to tune the G (process 6). To note, the model calculates L1 loss between y and y', which also serve as a tuning for G (process 7). The aim of these processes is to make the G learn to generate a fake y' that is so identical to the real image y that even the D can't tell. The D serves as a judge and also a guide that help improve the performance of G.

2.3 | Test and case study

2.3.1 | Test using artificial leaf damage

We tested our methods using the remaining set of 20% of collected leaf images (1332), which are mostly from the same plant species as the training data set. Based on these leaves, 5000 damaged leaves were simulated as the testing data set (and thus the ground-truth images are available for the damaged leaves). We also collected 500 leaves for a variety of plant species that are not included in the training data set, by randomly selecting leaves from a published leaf data set (imageCLEF2012, https://www.imageclef.org/2012/plant). Based on these leaves, 1000 leaf damage images were artificially generated as the second testing data set to evaluate the performance of our model when applied to a plant species not used for the training. We considered the performance of our model from two perspectives: How similar is it between the reconstructed intact leaf and the ground-truth image (i.e. image similarity); and how close is the predicted leaf damaged compared to ground-truth value (i.e. the accuracy of leaf damage estimates). To quantify the similarity between two images, we calculated the structural similarity index measure (SSIM). This metric compares the luminance, contrast and structure between two images and gives an overall score that range from 0 to 1, with 1 indicting a perfect match between two images and 0 indicating the worst match (Wang et al., 2004). To evaluate the accuracy of leaf damage estimates, we calculated the root mean square error of estimated leaf damage (RMSE) by the following equations

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum(\text{Predicted} - \text{GroundTruth})^2},$$

where the predicted and ground-truth percentage of leaf damage are included, and n represents the number of tested images.

2.3.2 | Test using real leaf damage

We randomly collected 311 leaves from five plant species: *A. mono*, *O. glaberrima*, *Q. mongolica*, *R. canthioides* and *U. japonica*, that is, around 60 leaves per plant species in the field and scanned the leaves. We used LeafByte software to manually estimate the percentage of leaf damage for each leaf image (Getman-Pickering et al., 2020). Specifically, the software allowed us to trace leaf boundaries for the leaves that were damaged and refill holes to calculate the 'restored' leaf area. The percentage of leaf damage was calculated based on the actual and restored leaf area. We calculated the RMSE to compare the leaf damage estimated by our method (both universal and species-specific models) with that estimated by LeafByte.

3 | RESULTS

3.1 | Efficiency of proposed procedures

Scanning multiple leaves in one image and splitting it to single-leaf images saved more than half of the time in leaf image collection compared to single-leaf image scanning. Generating artificially damaged leaves as training data is also fast which allows simulation of tens of thousands of images in a few minutes. The time required for training GAN models varied depending on the size of training data and the number of epochs but generally took a few hours, for example, 4h for 5000 images of training data and 100 epochs with a single NVIDIA GeForce RTX 3060 GPU. The leaf reconstruction and calculation of leaf area took around 0.25s per leaf on an Intel Core i5 CPU computer. Overall, these procedures can achieve a very high efficiency allowing estimates for millions of leaves in less than a day, which reduced labour and saved time compared with the weeks or months that would be required for manual digital analysis of this number of leaves.

3.2 | Accuracy of pretrained GAN models

The universal model accurately reconstructed artificially damaged leaves with an averaged similarity of 98.8% ± 1.2% (SSIM) between reconstructed leaves and ground truth (Figure 4). With real leaf damage, most of the leaf reconstruction were satisfying (Figure 4), and only a few heavily and irregularly damaged leaves were not reasonably reconstructed (Figure 5). Overall, the universal model achieved a high accuracy in estimating the percentage of artificial leaf damage, and such high accuracy was consistent whether the targeted plant species was included or not included in the training data set (Figure 6a,e). The model also showed a high accuracy in the test with real leaf damages, which revealed a close estimation with the manually digital analysis (RMSE = 1.6%, Figure 6c).

The reconstruction of artificially damaged leaves with plant species-specific models demonstrated a similarity of 98.4% ± 1.1% (SSIM) to the ground-truth image. The prediction of artificially damaged leaves based on species-specific model was close to the ground-truth value (RMSE = 2.4%), and the prediction for real leaf damage was also comparable to the manually digital analysis with RMSE of 1.6% (Figure 6b,d).

4 | DISCUSSION

Deep learning techniques have the potential to revolutionize data collection in ecology (Christin et al., 2019). Biotic interactions between plants and insect herbivores are one of the research areas

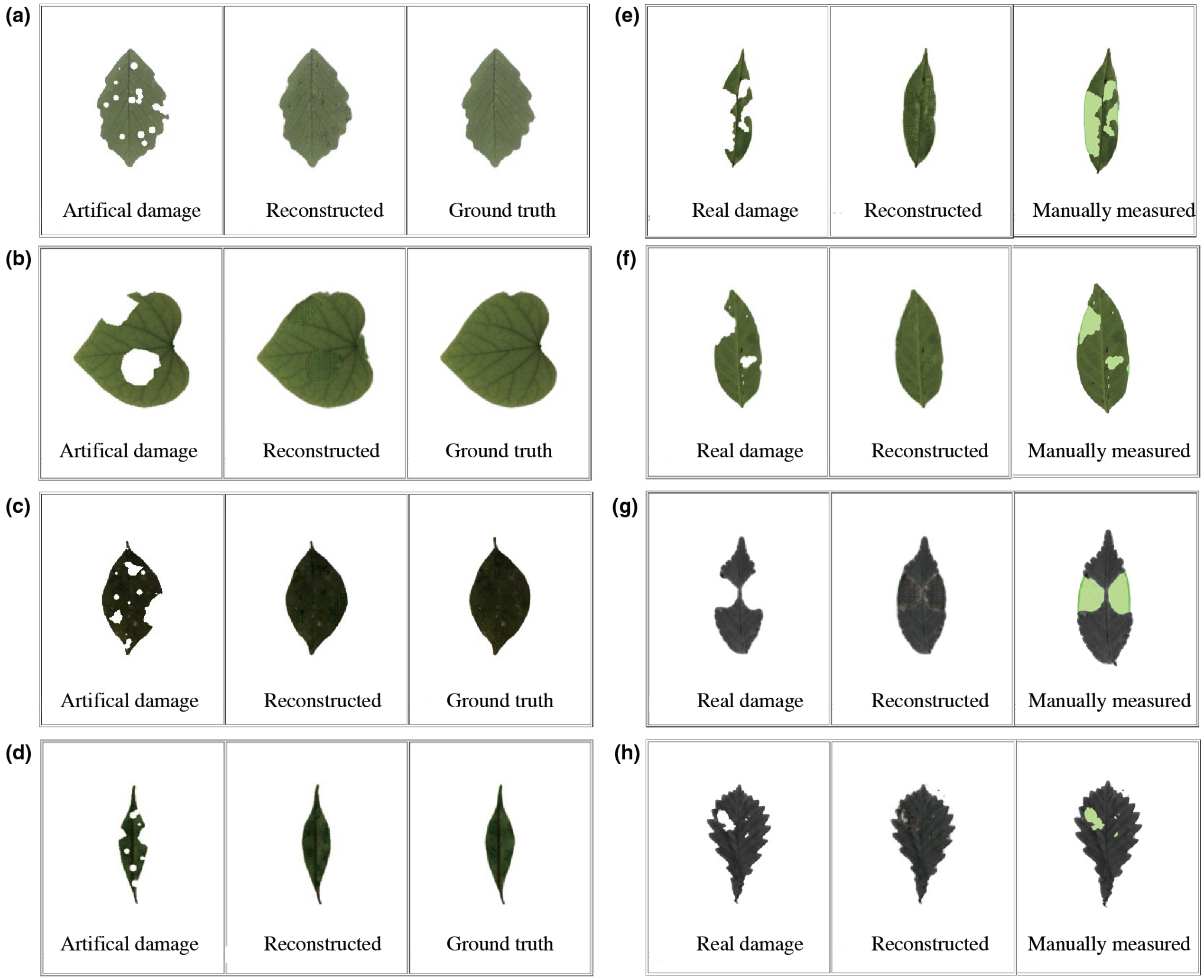


FIGURE 4 Examples of leaf reconstruction using the pretrained universal model for artificially damaged leaves (a–d) and real leaf damage data (e, f). The artificial leaf damage was generated from intact leaves which were taken as the ground truth, and real leaf damage data were manually estimated using LeafByte software to measure leaf damage. Both green and dark leaf colours were included. Plant species in (a–h) are *Quercus robur*, *Tilia americana*, *Neolitsea phanerophlebia*, *Magnolia championii*, *Ormosia glaberrima*, *Randia canthioides*, *Ulmus japonica*, *Quercus mongolica*, respectively.

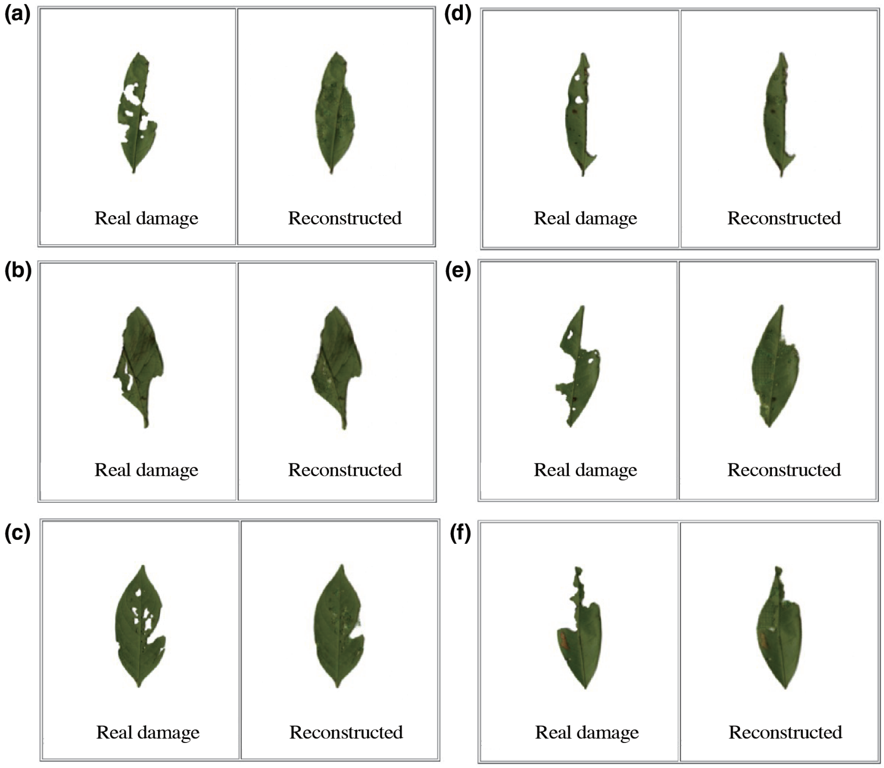


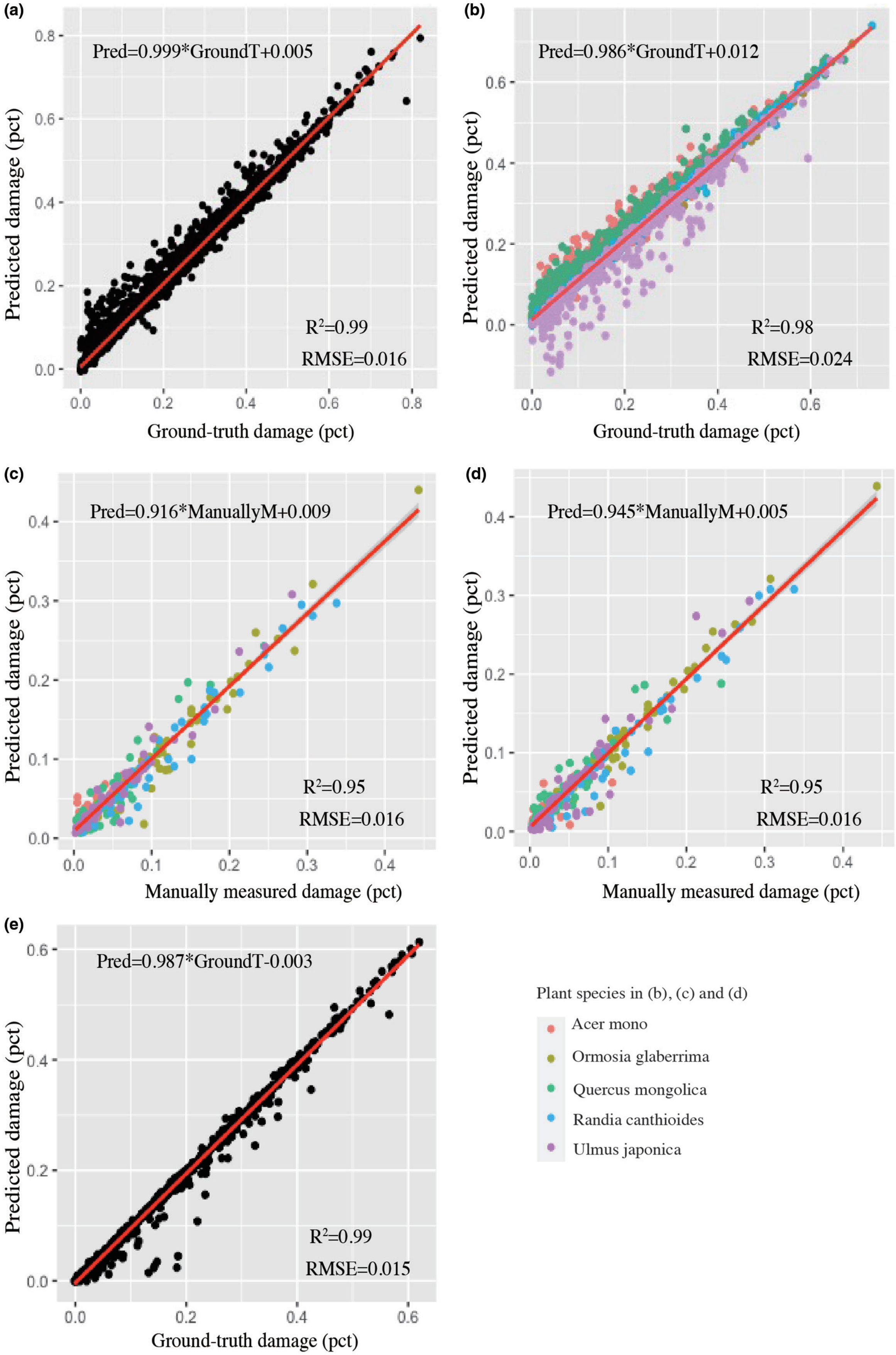
FIGURE 5 Examples of failures in leaf reconstruction using the pretrained universal model. (a–f) are the real leaf damages of *Ormosia glaberrima* and their unsatisfying reconstructions. The inappropriate leaf reconstructions primarily occur on irregularly or heavily damaged leaves.

that have been suffering from a lack of fast and automatic methods for data collection (Williams & Abbott, 1991). In our study, we combined image processing and deep learning techniques to build a pipeline for automatic estimation of leaf herbivore damage from images. We also developed a Python package and an online Shiny app with pretrained deep learning models to enable practical use of our methods. We demonstrated a high accuracy of the method in quantifying leaf damage which is comparable to manual estimation using software such as LeafByte (Getman-Pickering et al., 2020). Our study provided a practical solution for automated assessment of leaf herbivory damage, which will facilitate collecting large data sets of herbivory intensity to study plant–herbivore interactions in both applied and basic science.

Both CNN- and GAN-based deep learning models can be used to quantify the degree of leaf damage but with different computational approaches (da Silva et al., 2019; Hussein et al., 2021). CNNs can directly predict leaf damage from leaf images (i.e. return a label indicating the degree of leaf damage), while GANs can predict the intact status of damaged leaves (i.e. return a reconstructed leaf image that can be used to calculate leaf area and the degree of leaf damage) (da Silva et al., 2019; Hussein et al., 2021). We suggest that GANs are superior to CNNs for leaf damage estimation for several reasons. First, GANs generate intact leaf images that allow visual checks of the reconstruction of damaged leaves. Researchers can pick out the unsuccessful reconstructions by a quick visual skim through images and manually determine their leaf damage to improve the overall accuracy of the method (Isola et al., 2017; Zhu et al., 2017). In contrast, the CNNs directly return values of estimated leaf damage without information on the confidence of the prediction. Second, GANs can achieve higher accuracy than CNNs because CNNs are designed to classify objectives into categories instead of for quantitative estimation. For example, previous attempts to use CNNs in quantifying leaf damage has revealed an averaged bias (RMSE) of 4.6% while in our study we showed an averaged bias of 1.6% (da Silva et al., 2019). Lastly, we showed that a relatively small training data set (3000–5000 images) is enough to train GANs with good performance in reconstructing damaged leaves, comparing to training data sets of tens of thousands of images required for training CNNs, which consequentially reduces the time and computing resources needed for model training (da Silva et al., 2019).

We showed that GAN models can achieve a high accuracy in reconstructing the intact status of leaves and estimating leaf damage.

FIGURE 6 Tests of the universal model (a, c, e) and plant species-specific models (b, d) using artificial testing data (a, b, e) and real testing data (c, d). In (a, b, e) the predicted proportion of leaf damage was correlated with the ground truth (a, 5000 leaf damages generated from the same plant species as the training data set; b, 5000 leaf damages generated for five plant species; and e, 1000 leaf damages generated from different plant species than the training data set). In (c, d) the predicted proportion of leaf damage was correlated with manually measured leaf damage (c, 311 real leaf images from 5 plant species predicted by the universal model; and d, the same 311 images predicted by species-specific models). The equation indicating the slope and intercept of these relationships (Pred: predicted; GroundT: ground-truth; ManuallyM: manually measured), the explained variance (R^2) and the root mean square error (RMSE) between estimates were shown.


(a)
Pred=0.999*GroundT+0.005
R²=0.99
RMSE=0.016
Predicted damage (pct)
Ground-truth damage (pct)
(b)
Pred=0.986*GroundT+0.012
R²=0.98
RMSE=0.024
Predicted damage (pct)
Ground-truth damage (pct)
(c)
Pred=0.916*ManuallyM+0.009
R²=0.95
RMSE=0.016
Predicted damage (pct)
Manually measured damage (pct)
(d)
Pred=0.945*ManuallyM+0.005
R²=0.95
RMSE=0.016
Predicted damage (pct)
Manually measured damage (pct)
(e)
Pred=0.987*GroundT-0.003
R²=0.99
RMSE=0.015
Predicted damage (pct)
Ground-truth damage (pct)
Plant species in (b), (c) and (d)
Acer mono
Ormosia glaberrima
Quercus mongolica
Randia canthioides
Ulmus japonica

We considered both universal models and case-specific models (i.e. plant species-specific models). The universal model was trained based on artificially damaged leaves of multiple woody plant species and thus can be applied to estimate leaf damage for different plant taxa and leaf shapes. The universal model achieved good accuracy in leaf damage estimation with an averaged RMSE of 1.6%, which is comparable to the error of manual damage estimation with LeafByte and ImageJ (Getman-Pickering et al., 2020), and is more accurate than visual estimates of leaf damage (Johnson et al., 2016). Although the universal model showed a general applicability to different plant taxa and leaf shapes, it is not guaranteed that the model can predict intact leaves for all plant taxa. For example, herbaceous or fern species could have different leaf shape and/or leaf damage patterns compared to the woody plants that we used for model training, and therefore, the performance of the universal model may be reduced when applied to non-woody plant life forms. In the case where universal model is not applicable, researchers can train their own case-specific models using the procedures and packages we provided (see Supporting Information). Our tests have shown that the effort required for training case-specific model is reasonable, and the case-specific model can achieve an accuracy comparable with the universal model.

While our methods offer a general solution for estimating leaf herbivore damage, they do come with several limitations that will necessitate future development and optimization. First, the reconstruction of heavily damaged leaves can sometimes yield unsatisfactory results (see Figure 5), which is understandable as tracing the boundaries of heavily damaged leaves is difficulty even for manual estimation. As a result, the method may underestimate leaf damage at higher damage levels, as indicated by the slope of the relationship between predicted and ground-truth herbivory (the coefficient of the slope is slightly less than 1, Figure 6). To address this limitation, we strongly recommend conducting a visual inspection of leaf images to identify and rectify unsatisfactory reconstructions. This may require a modest additional effort but can significantly enhance overall accuracy and data quality. Fortunately, heavily damaged leaves represent only a small fraction of total leaf damage in natural ecosystems. For instance, in a study conducted in natural forests (unpublished data, Zihui Wang), only less than 5% of leaf images failed in damage reconstruction with our model. Second, our method primarily focused on leaf chewing damage (i.e. complete removal of leaf materials) and regular damage shapes (such as circles and polygons); therefore, it may struggle with reconstructing irregular leaf damage and damage associated with colour changes. The challenges associated with reconstructing irregular leaf damage can potentially be mitigated by optimizing the 'synthetic' function to generate training data featuring irregular damage shapes such as the linear damage caused by leaf miners (da Silva et al., 2019). However, it is worth noting that our method is not applicable to leaf damage that only changes leaf colours (i.e. yellowish spots caused by sap feeders) or damage that only shows on the upper or lower side of leaves (i.e. slight damage caused by some leaf miners). Researchers should evaluate the dominant leaf damage type in their study system before applying this method.

Another limitation is that our method requires clear leaf images with a blank background, containing only leaves without overlaps among them. This may limit the collection of in situ leaf images (i.e. without detaching leaves from the plant). To address this, researchers can adopt practices such as using blank paper as a background when scanning or photographing leaves and preprocessing images containing non-leaf features or non-blank backgrounds before applying the method. Lastly, the universal model was trained based on simple leaves and separated leaflets of compound leaves, and therefore, it cannot directly predict whole compound leaves. It is possible to train a model based on whole compound leaves, but such a model may be highly specific since the shape of compound leaves can vary greatly among plant taxa.

By combining deep learning and image processing, we provided a state-of-the-art solution for automatic estimation of leaf herbivore damage. The method will allow researchers to quantify leaf herbivory intensity in complex systems with much less effort compared to traditional methods. Moreover, the method and tools can be potentially modified to solve other problems related to damage estimation such as leaf disease damage or butterfly wing damage. We hope that our method will motivate the collection of large data sets of plant–leaf herbivore interaction to address big picture questions in plant ecology, entomology, agronomy and forestry.

AUTHOR CONTRIBUTIONS

Zihui Wang and Steven W. Kembel conceived the ideas; Zihui Wang designed methodology and wrote the codes with supports from Steven W. Kembel and Abdoulaye Baniré Diallo; Yuan Jiang collected leaf images and manually measured leaf damage data; Zihui Wang led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We thank Dr. Burhan Rashid Hussein and Dr. Owais Ahmed Malik from Universiti Brunei Darussalam for their help in the coding of 'synthetic' function. We thank Amanda Boatswain Jacques from Université du Québec à Montréal for the suggestions in developing deep learning models. We thank the Natural Sciences and Engineering Research Council of Canada CREATE Computational Biodiversity Science and Services (BIOS^2) training programme for the training on image processing tools.

CONFLICT OF INTEREST STATEMENT

The authors state no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14293.

DATA AVAILABILITY STATEMENT

The program and it source code used in the manuscript were released at https://github.com/ZihuiWang1/HerbiEstim and archived in Zenodo at https://zenodo.org/records/10514389; the online website application of the method is at http://herbiestim.online; the image data supporting the training and testing of deep learning models are available at Figshare https://doi.org/10.6084/m9.figshare.23982735.v1.

ORCID

Zihui Wang https://orcid.org/0000-0002-2512-7361
Yuan Jiang https://orcid.org/0000-0001-8704-3676

REFERENCES

Avila-Sakar, G., Leist, L. L., & Stephenson, A. G. (2003). Effects of the spatial pattern of leaf damage on growth and reproduction: Nodes and branches. *Journal of Ecology*, *91*(5), 867–879. https://doi.org/10.1046/j.1365-2745.2003.00819.x

Bangare, S. L., Dubal, A., Bangare, P. S., & Patil, S. T. (2015). Reviewing Otsu's method for image thresholding. *International Journal of Applied Engineering Research*, *10*(9), 21777–21783. https://doi.org/10.37622/ijaer/10.9.2015.21777-21783

Brodie, E. D. (1999). Predator-prey arms races: Asymmetrical selection on predators and prey may be reduced when prey are dangerous. *Bioscience*, *49*(7), 557–568. https://doi.org/10.2307/1313476

Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, *10*(10), 1632–1644. https://doi.org/10.1111/2041-210x.13256

Coley, P. D. (1983). Herbivory and defensive characteristics of tree species in a lowland tropical forest. *Ecological Monographs*, *53*(2), 209–234. https://doi.org/10.2307/1942495

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, *35*(1), 53–65. https://doi.org/10.1109/msp.2017.2765202

da Silva, L. A., Bressan, P. O., Gonçalves, D. N., Freitas, D. M., Machado, B. B., & Gonçalves, W. N. (2019). Estimating soybean leaf defoliation using convolutional neural networks and synthetic images. *Computers and Electronics in Agriculture*, *156*, 360–368. https://doi.org/10.1016/j.compag.2018.11.040

Ferreira, A. C., Silva, L. R., Renna, F., Brandl, H. B., Renoult, J. P., Farine, D. R., Covas, R., & Doutrelant, C. (2020). Deep learning-based methods for individual recognition in small birds. *Methods in Ecology and Evolution*, *11*(9), 1072–1085. https://doi.org/10.1111/2041-210x.13436

Getman-Pickering, Z. L., Campbell, A., Aflitto, N., Grele, A., Davis, J. K., & Ugine, T. A. (2020). LeafByte: A mobile application that measures leaf area and herbivory quickly and accurately. *Methods in Ecology and Evolution*, *11*, 215–221. https://doi.org/10.1111/2041-210x.13340

Hairston, N. G., Smith, F. E., & Slobodkin, L. B. (1960). Community structure, population control, and competition. *The American Naturalist*, *94*(879), 421–425. https://doi.org/10.1086/282146

Hussein, B. R., Malik, O. A., Ong, W.-H., & Slik, J. W. F. (2021). Reconstruction of damaged herbarium leaves using deep learning techniques for improving classification accuracy. *Ecological Informatics*, *61*, 101243. https://doi.org/10.1016/j.ecoinf.2021.101243

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). *Image-to-image translation with conditional adversarial networks*. 2017 IEEE conference on Computer Vision and Pattern Recognition (CVPR), 5967–5976. https://doi.org/10.1109/cvpr.2017.632

Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, *104*(940), 501–528. https://doi.org/10.1086/282687

Johnson, M. T. J., Bertrand, J. A., & Turcotte, M. M. (2016). Precision and accuracy in quantifying herbivory. *Ecological Entomology*, *41*(1), 112–121. https://doi.org/10.1111/een.12280

Keane, R. M., & Crawley, M. J. (2002). Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution*, *17*(4), 164–170. https://doi.org/10.1016/s0169-5347(02)02499-0

Kogan, M., Turnipseed, S. G., Shepard, M., Oliveira, E. B. D., & Borgo, A. (1977). Pilot insect pest management program for soybean in Southern Brazil1. *Journal of Economic Entomology*, *70*(5), 659–663. https://doi.org/10.1093/jee/70.5.659

Kumar, N., Belhumeur, P. N., Biswas, A., Jacobs, D. W., Kress, W. J., Lopez, I. C., & Soares, J. V. B. (2012). Leafsnap: A computer vision system for automatic plant species identification. *Lecture Notes in Computer Science*, *7573*, 502–516. https://doi.org/10.1007/978-3-642-33709-3_36

Machado, B. B., Orue, J. P. M., Arruda, M. S., Santos, C. V., Sarath, D. S., Goncalves, W. N., Silva, G. G., Pistori, H., Roel, A. R., & Rodrigues, J. F. (2016). BioLeaf: A professional mobile application to measure foliar damage caused by insect herbivory. *Computers and Electronics in Agriculture*, *129*, 44–55. https://doi.org/10.1016/j.compag.2016.09.007

Myers, J. H., & Sarfraz, R. M. (2015). Impacts of insect herbivores on plant populations. *Annual Review of Entomology*, *62*(1), 207–230. https://doi.org/10.1146/annurev-ento-010715-023826

Neves, F. S., Silva, J. O., Espírito-Santo, M. M., & Fernandes, G. W. (2014). Insect herbivores and leaf damage along successional and vertical gradients in a tropical dry Forest. *Biotropica*, *46*(1), 14–24. https://doi.org/10.1111/btp.12068

Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(25), E5716–E5725. https://doi.org/10.1073/pnas.1719367115

O'Neal, M. E., Landis, D. A., & Isaacs, R. (2002). An inexpensive, accurate method for measuring leaf area and defoliation through digital image analysis. *Journal of Economic Entomology*, *95*(6), 1190–1194. https://doi.org/10.1603/0022-0493-95.6.1190

Sam, K., Koane, B., Sam, L., Mrazova, A., Segar, S., Volf, M., Moos, M., Simek, P., Sisol, M., & Novotny, V. (2020). Insect herbivory and herbivores of *Ficus* species along a rain forest elevational gradient in Papua New Guinea. *Biotropica*, *52*(2), 263–276. https://doi.org/10.1111/btp.12741

Silva, M. C., Bianchi, A. G. C., Ribeiro, S. P., & Oliveira, R. A. R. (2022). Bringing deep learning to the fields and forests: Leaf reconstruction and shape estimation. *SN Computer Science*, *3*(3), 195. https://doi.org/10.1007/s42979-022-01082-4

Söderkvist, O. (2001). *Computer vision classifcation of leaves from swedish trees (Master's Thesis)*. Linköping, Sweden: Linkoping University.

Tabak, M. A., Norouzzadeh, M. S., Wolfson, D. W., Sweeney, S. J., Vercauteren, K. C., Snow, N. P., Halseth, J. M., Salvo, P. A. D., Lewis, J. S., White, M. D., Teton, B., Beasley, J. C., Schlichting, P. E., Boughton, R. K., Wight, B., Newkirk, E. S., Ivan, J. S., Odell, E. A., Brook, R. K., … Miller, R. S. (2019). Machine learning to classify animal species in camera trap images: Applications in ecology. *Methods in Ecology and Evolution*, *10*(4), 585–590. https://doi.org/10.1111/2041-210x.13120

Villacis-Llobet, J., Lucio-Troya, M., Calvo-Navarro, M., Calderon-Ramirez, S., & Mata-Montero, E. (2020). A first glance into reversing senescence on herbarium sample images through conditional generative adversarial. *Communications in Computer and Information Science*, *1087*, 438–447. https://doi.org/10.1007/978-3-030-41005-6_30

Wang, Z. (2024). HerbiEstim: A deep learning- and image processing-based software for automatic estimation of leaf herbivore damage (v1.0.0). *Zenodo*. https://doi.org/10.5281/zenodo.10514389

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600–612. https://doi.org/10.1109/tip.2003.819861

Williams, M. R., & Abbott, I. (1991). Quantifying average defoliation using leaf-level measurements. *Ecology*, *72*(4), 1510–1511. https://doi.org/10.2307/1941126

Xirocostas, Z. A., Debono, S. A., Slavich, E., & Moles, A. T. (2022). The ZAX herbivory trainer—Free software for training researchers to visually estimate leaf damage. *Methods in Ecology and Evolution*, *13*(3), 596–602. https://doi.org/10.1111/2041-210x.13785

Xu, Z., Baojie, X., & Guoxin, W. (2017). *Canny edge detection based on open CV*. 2017 13th IEEE international conference on Electronic Measurement & Instruments (ICEMI), 53–56. https://doi.org/10.1109/icemi.2017.8265710

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). *Unpaired image-to-image translation using cycle-consistent adversarial networks*. 2017 IEEE International Conference on Computer Vision (ICCV), 2242–2251. https://doi.org/10.1109/iccv.2017.244

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Table S1. The detailed information about plant leaf images used for training the universal model.

Appendix S1. A user guide for the Python package 'HerbiEstim'.

How to cite this article: Wang, Z., Jiang, Y., Diallo, A. B., & Kembel, S. W. (2024). Deep learning- and image processing-based methods for automatic estimation of leaf herbivore damage. *Methods in Ecology and Evolution*, *00*, 1–12. https://doi.org/10.1111/2041-210X.14293