Morphology exploration of pollen using deep learning latent space

James A. Grant-Jacob*, Michalis N. Zervas and Ben Mills

Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, UK

*J.A.Grant-Jacob@soton.ac.uk

Abstract

The structure of pollen has evolved depending on its local environment, competition, and ecology. As pollen grains are generally of size 10-100 microns with nanometre-scale substructure, scanning electron microscopy is an important microscopy technique for imaging and analysis. Here, we use style transfer deep learning to allow exploration of latent w-space of scanning electron microscope images of pollen grains and show the potential for using this technique to understand evolutionary pathways and characteristic structural traits of pollen grains.

1. Introduction

Land plants have their origin in the mid-Palaeozoic era[1]. It is estimated that around 350,000 species of flowering plants have evolved[2], producing pollen grains with sizes ranging from approximately 10-100 microns depending on the species, and with substructure on the scale of nanometres[3]. Pollen takes a variety of shapes, such as trilobal, spherical or hexagonal, and the surface of pollen (the exine) can have features such as apertures[4] that play a role in mechanisms such as germination and harmomegathy[5]. Imaging of pollen grains is therefore an important technique, as it offers a vital indicator in the health of agriculture crop[6] and the local environmental[7].

Current understanding of pollen traits and evolution generally involves manual observation, with various methods of imaging pollen grains used to identify their external and internal structure. These include fluorescence microscopy[8] and electron microscopy[9], while modelling using analytical methods has been used for exploring the formation of pollen grain apertures[10]. However, these approaches can be time consuming and modelling biological systems can be challenging. In recent years, developments in graphics processing units (GPUs) and deep learning algorithms have unlocked a new paradigm of large-scale data driven research[11]. Relevant palynology, deep learning has been used for pollen identification via visible light microscopy[12] and fibre optic-based sensing[13], for imaging of pollen grains from scattering patterns[14], and for transforming images of dehydrated pollen grains into images of hydrated pollen grains[15].

This manuscript describes the application of deep learning to artificially generate scanning electron microscope (SEM) images of pollen grains and then interpolate between the generated images in latent w-space to potentially allow the exploration of transformations from one pollen species to another, as conceptualised in figure 1. The latent w-space is a multi-dimensional space that enables close positional representation of data that are similar in the space externally to the neural network (i.e. training data). Since evolution generally results in incremental changes to the visual appearance of pollen grains, and that the neural network clusters pollen grain types based on their visual appearance in latent w-space, interpolation and extrapolation across latent w-space could be comparable to the evolutionary path between pollen species in a phylogenetic tree. We also show that w-space latent vectors can be determined that can allow characteristics, such as the pollen grain size, to be increased or decreased in the generated images, potentially allowing the understanding of species morphology evolution.
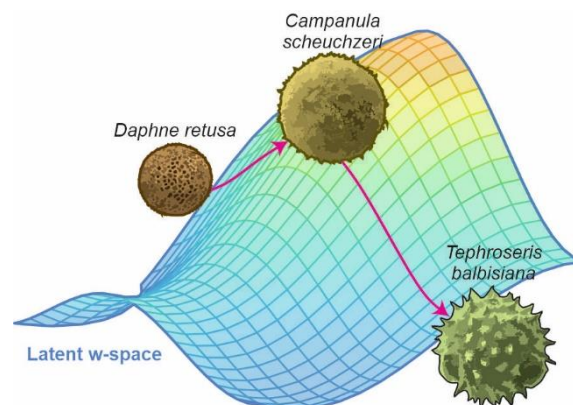
Figure 1. Exploring pollen characteristics using image transformation in latent w-space. A neural network transforms a generated image of a pollen grain into another generated image of a pollen grain, via interpolation in latent w-space.

2. Experimental methods

2.1 Data acquisition

All SEM images were downloaded from PalDat[16], with each image file associated with a number and the species, taxa, genus and imaging method stored in an array. Only rehydrated single pollen grain images were used for training the neural network. A total of 2775 images were used in training the neural network, which consisted of 778 different genus of pollen. Each image was cropped and resized to 512×512 pixels, with the scale text removed but the scale bar preserved to allow for the neural network to generate images with scale bars and thus provide information of the pollen grain size. The scale bar in each image represents 10 µm.

2.2 Neural network

The style transfer deep learning neural network known as StyleGAN[17] was used to generate synthetic SEM images of pollen grains, where the appearance of the generated pollen grains was based on the training data. Specifically, we used the StyleGAN 2 neural network that is freely available on GitHub at https://github.com/NVlabs/stylegan2-ada-pytorch.git. The StyleGAN neural network architecture is a type of generative adversarial network architecture that uses a non-linear mapping network to transform a random vector into an intermediate latent w-space vector, where vector elements are correlated with different visual features in the generated image (see figure 2). This enables latent w-space to permit smooth transitions between different generated images. The mapping network transforms a 1×512 random vector into a 1×512 latent w-space vector, and the generation network transforms this vector into a 512×512 resolution generated image.
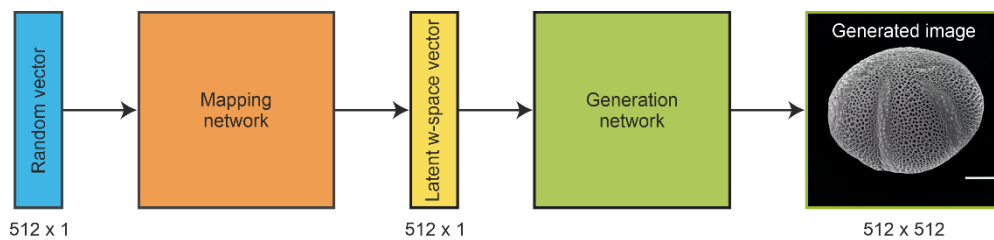

Figure 2. Schematic of the StyleGAN neural network.

StyleGAN has been used for generating and interpolating between different photographs. As an example, figure 3 shows the interpolation between a seed image showing a field and a seed image showing a forest. The sequence of images demonstrates the transition from yellow grass into green trees and a red sky into blue sky. The network used to generate these data was a pretrained StyleGAN 2 network trained on landscape images (https://github.com/ThisJustin-code/pretrained-gan-landscapes-256.git).

Figure 3. Sequence of images from seed 1 to seed 2 showing the interpolation from a field to a forest.

The neural network was trained on a Windows 10 workstation, with a 64 AMD core CPU, 512 GB RAM, 4 × NVIDIA A5000 GPUs (24 GB memory each), using CUDA-enabled Python. The neural network was trained for 7000 kimg (thousands of real images shown to the discriminator) at 39 sec/kimg, taking approximately 3.5 days. Training parameters included a learning-rate of 0.0025, a non-saturating logistic loss function, batch-size of 32 and the ADAM optimiser[18].

2.3 Image generation

W-space seeds were used as input to the generative neural network to produce 512×512-pixel images, which contain generated 10 µm scale bars. Interpolation between different w-space seeds was achieved by calculating the w-space vector from w-space seed 1 to w-space seed 2 and adding fractions of this vector to the coordinate of w-space seed 1. As the w-space seeds are random vector coordinates in 512-dimension w-space, the interpolation from one w-space seed to another w-space seed is unique and can be mathematically described as the position of w-space seed 1 added to a fraction of a vector from w-space seed 1 to w-space seed 2 (calculated as w-space seed 2 minus w-space seed 1). Therefore, by taking the average of many w-space vectors from generated images of small pollen grains to generated images of large pollen grains, a w-space vector for 'increasing size' (size vector) can be determined. This vector can then be added (or subtracted) to any w-space seed to increase or decrease the size of the pollen in the generated image. The vector for pollen with no spikes to spikes (spike vector) was also determined.
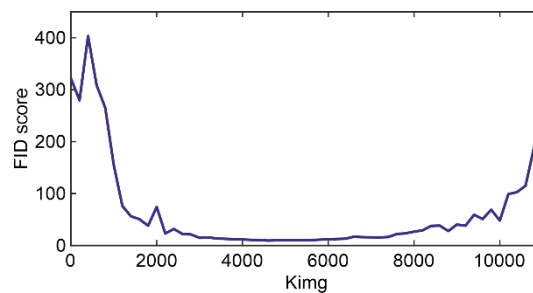

Figure 4. FID score vs kimg.

The Fréchet Inception Distance (FID) score[19], which quantifies the distance between feature vectors for generated images and training images, was used to compare the distributions of the generated images and the training images, with a score of zero indicating that the two sets of data have identical distributions. The FID score was calculated every 200 kimg and monitored during training, reaching 250 for 11000 kimg (see figure 4), and there was visible overfitting since there was no variation between the generated images, unlike for 7000 kimg.

3. Results and discussion

Figure 5 show five pairs of generated pollen images, along with the associated generated images for ten equal interpolation steps, and labels that correspond to species that the pollen grains in the images resemble. The pollen species are identified via manual comparison with the actual pollen grain images used in training and present in PalDat database. In (a), no images resembling additional species appear to be generated between *Ipomoea* and *Tephroseris balbisiana*, which is likely due to their similar structure. In (b), two visually different pollen grains are used as w-space seed images (pollen resembling *Thymus praecox* for w-space seed 1 and pollen resembling *Pinus strobus* for w-space seed 2), and notably, an image of pollen with characteristics similar to *Tillandsia hildae* is generated during interpolation. Of note, the pollen grains increase in size from seed 1 to seed 2. Similar observations occur in (c) where an image with characteristics similar to an additional species (*Campanula scheuchzeri*) appears to be produced between w-space seed 1 (*Daphna retusa*) and w-space seed 2 (*Tephroseris*), where the interpolated pollen grain has a size mid-way between the start and end pollen grains. For figure 5(d), images of pollen with characteristics distinct from the w-space seeds are generated and resemble the species *Liriope graminifolia*. In (e), images of pollen grains with similar characteristics to the species namely *Pachysandra terminalis* and *Crepis aurea* are generated during interpolation.

Current methodology of classifying pollen is generally via structural examination. As such, if the generated pollen grains match an evolutionary path in a phylogenetic tree, then the neural network could assist in classifying the taxa or family of pollen. For example, in figure 5(a), no additional species appear to be generated when interpolating between *Ipomoea* and *Tephroseris balbisiana*, which is likely due to their similar structure. Based on comparison with a phylogenetic tree, these two species are relatively close in evolutionary terms, being part of the *Asterids* clade.

Since interpolation takes place in latent w-space, the seed images must also be generated in latent w-space, meaning the initial seeds images may not always resemble any actual species due to accuracy of generation. Since evolutionary changes to morphology can be incremental[20], it is worth considering that the images between two

species could infer ancestral characteristics, and transitions between the species could aid in understanding characteristic development. However, owing to convergent evolution, pollen images may be visually similar to other species but aligned on a different evolutionary pathway.
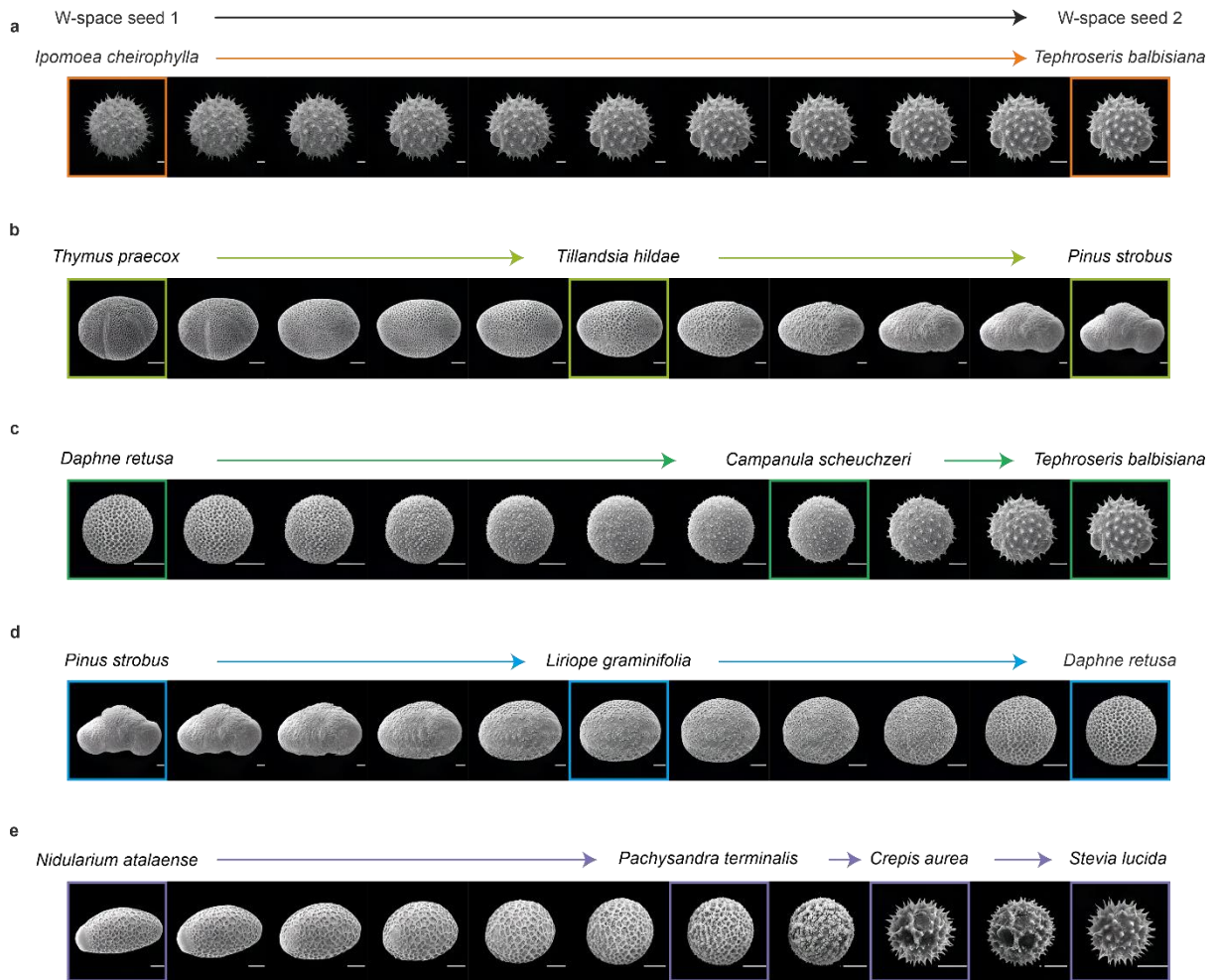

Figure 5. Generated SEM images corresponding to interpolation in latent w-space between five pairs of w-space seed 1 to w-space seed 2. The coloured boxes are the identified resembled species of pollen grains. Scale bars are 10 μm.

We carried out vector arithmetic on different w-space seeds to see how the generated images change. This could be used to understand how certain characteristics could evolve in certain environments, such as increasing humidity, which could change the pollen grain size. As shown in figure 6(a), we add $-100\%, -50\%, 0\%, +50\%$ and $+100\%$ of the spike vector onto the vector w-space seed images of (i) *Stevia lucida* and (ii) *Lithops gracilidelineata*, where the w-space seed is labelled as 0%. When the spike vector is subtracted from the w-space seeds, the spikes on the generated pollen grains get smaller and eventually disappear, and conversely as the spike vector is added to the w-space seeds, the spikes become more prominent. The resultant images appear to resemble species of pollen different from the input seeds. In figure 6(b), percentages of the size vector were added onto the w-space seeds of (i) *Stevia lucida* and (ii) *Fritillaria raddeana*. As the *Stevia lucida* gets larger, the apertures become less visible, and the spikes change in appearance. As the *Fritillaria raddeana* becomes smaller, the grain becomes more circular and resembles a different species, potentially indicating a relationship in latent w-space between size and shape. This could perhaps be used to infer properties of the internal structure or morphology adaptations associated with increasing size. It has been shown that the water availability of pollen during germination can affect the size of pollen[21], and so by determining the relationship between size and hydration, one could predict how the morphology of a pollen grain may change under different conditions (which could aid in classification and predicting viability). In practice, vectors for any type of visual transformation could be explored.
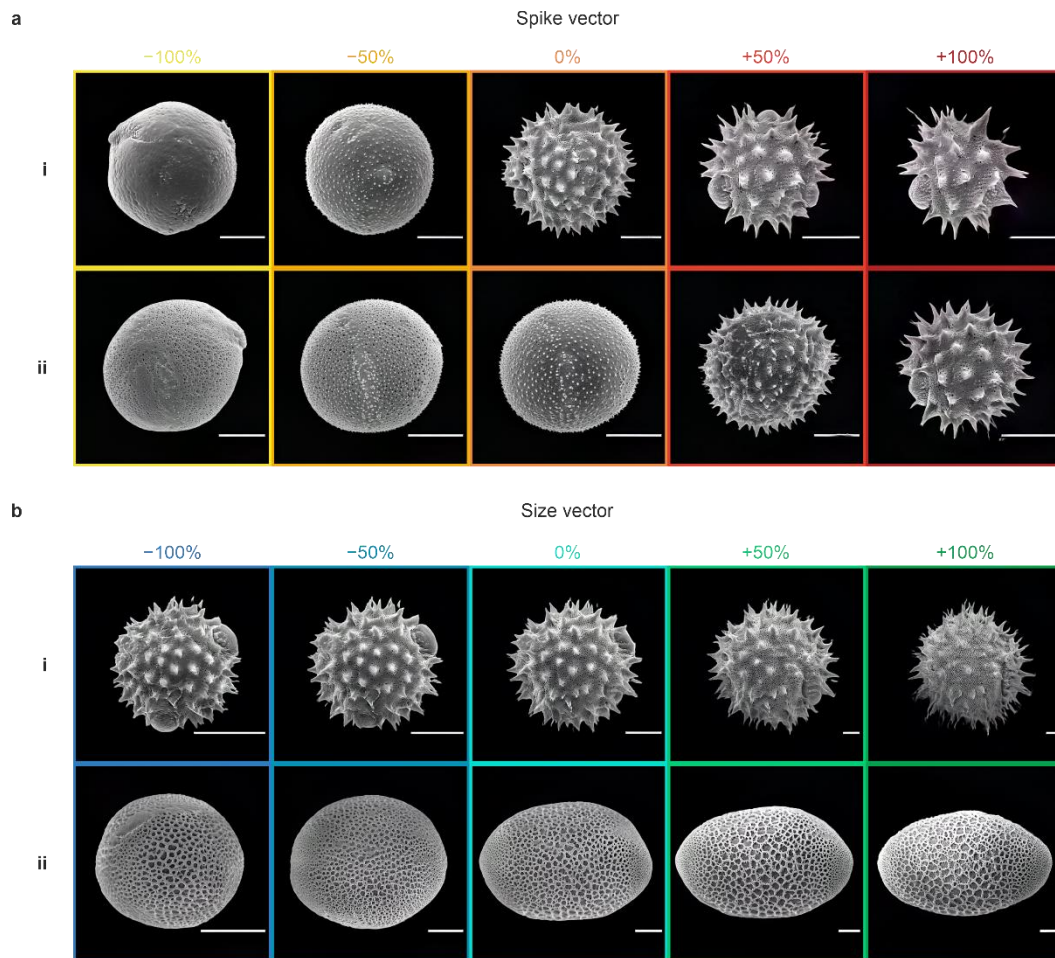
Figure 6. Vector manipulation in latent w-space. Showing (a) no spike-to-spike vector added to w-space seeds corresponding to (i) *Stevia lucida* and (ii) *Lithops gracilidelineata*, and (b) small-to-large vector added to w-space seeds corresponding to (i) *Stevia lucida* and (ii) *Fritillaria raddeana*. The scale bars are 10 µm.

4. Conclusion

In summary, interpolation and extrapolation in latent w-space of a style transfer generative neural network enabled generation of images of pollen grains of species different to the w-space seed images. The incremental interpolation pathways hint that the clustering of pollen grain characteristics in multidimensional latent w-space could be correlated with evolutionary pathways and hence this technique could be applied to evolutionary research. Since the internal structure of a pollen grain can impact on the external morphology, imaging modalities such as transmission electron microscopy could also be useful to explore. This technique could also be used to simulate the visual appearance of pollen grains based on potential climate changes (i.e., increased humidity leading to changes in pollen grain size). This methodology could be applied to photographs of animals or X-ray scattering patterns of viruses.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5258/SOTON/D2356.

Conflicts of interest

The authors declare no conflict of interest.

ORCiDs

James A. Grant-Jacob https://orcid.org/0000-0002-4270-4247

Michalis N. Zervas https://orcid.org/0000-0002-0651-4059

Ben Mills https://orcid.org/0000-0002-1784-1012

References

[1]   P. Kenrick, P. R. Crane, *Nature 1997 389:6646* **1997**, *389*, 33.
[2]   L. N. Joppa, D. L. Roberts, S. L. Pimm, *Proceedings of the Royal Society B: Biological Sciences* **2011**, *278*, 554.
[3]   H. Halbritter, S. Ulrich, F. Gr\'\imsson, M. Weber, R. Zetter, M. Hesse, R. Buchner, M. Svojtka, A. Frosch-Radivo, in *Illustrated Pollen Terminology*, Springer, **2018**, pp. 97–127.
[4]   C. Mert, *HortScience* **2009**, *44*, 519.
[5]   A. Matamoro-Vidal, C. Prieu, C. A. Furness, B. Albert, P. H. Gouyon, *New Phytologist* **2016**, *209*, 376.
[6]   P. C. Fernandez-Mensaque, F. J. G. Minero, J. Morales, C. Tomas, *Aerobiologia (Bologna)* **1998**, *14*, 185.
[7]   T.-C. Lau, X. Lu, R. T. Koide, A. G. Stephenson, *Plant Cell Environ* **1995**, *18*, 169.
[8]   J. Atlagić, S. Terzić, A. Marjanović-Jeromela, *Ind Crops Prod* **2012**, *35*, 88.
[9]   A. P. Coutinho, A. M. Dinis, *Plant Systematics and Evolution* **2009**, *283*, 79.
[10]   Y. Zhou, A. A. Dobritsa, *Plant Science* **2019**, *288*, 110222.
[11]   Y. LeCun, Y. Bengio, G. Hinton, *Nature* **2015**, *521*, 436.
[12]   V. Sevillano, J. L. Aznarte, *PLoS One* **2018**, *13*.
[13]   J. A. Grant-Jacob, S. Jain, Y. Xie, B. S. Mackay, M. D. T. McDonnell, M. Praeger, M. Loxham, D. J. Richardson, R. W. Eason, B. Mills, *Journal of Physics: Photonics* **2019**, *1*, 44004.
[14]   J. A. Grant-Jacob, M. Praeger, M. Loxham, R. W. Eason, B. Mills, *Environ Res Commun* **2020**, *2*, 75005.
[15]   J. A. Grant-Jacob, M. Praeger, R. W. Eason, B. Mills, *IOP SciNotes* **2022**, *3*, 024001.
[16]   M. Weber, S. Ulrich, *Grana* **2017**, *56*, 257.
[17]   T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, T. Aila, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* **2020**, 8107.
[18]   D. P. Kingma, J. Ba, *arXiv preprint arXiv:1412.6980* **2014**.
[19]   M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, S. Hochreiter, *Adv Neural Inf Process Syst* **2017**, *30*.
[20]   S. A. Newman, *Philosophy of Complex Systems* **2011**, 335.
[21]   H. Fatmi, S. Mâalem, B. Harsa, A. Dekak, H. Chenchouni, *Web Ecol* **2020**, *20*, 19.