Pollen grain recognition through deep learning convolutional neural networks

Cite as: AIP Conference Proceedings **2425**, 140001 (2022); https://doi.org/10.1063/5.0081614
Published Online: 06 April 2022

Fernando C. Monteiro

Pollen Grain Recognition Through Deep Learning Convolutional Neural Networks

Fernando C. Monteiro

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal

Corresponding author:monteiro@ipb.pt

Abstract. Palynology is the study of pollen, in particular, the pollen's grain type, but the tasks of classification and counting of pollen grains are highly skilled and laborious. Despite the efforts made during the last decades, the manual classification process is still predominant. One of the reasons for that is the small number of taxa usually used in previous approaches. In this paper, we propose a new method to automatically classify pollen grains using a state-of-the-art deep learning technique applied to the recently published POLEN73S image dataset. Our proposal manages to classify up to 94% of the samples from the dataset with 73 different classes of pollen grains. This result, which surpasses all previous attempts in number and difficulty of taxa under consideration, gives good perspectives to achieve a perfect score in pollen recognition task even with a large number of pollen grain types.

INTRODUCTION

Pollen grain identification and certification classification have a remarkable role in a wide range of applications including crime scene investigation [1], aerobiological studies [2] as well as the botanical and geographical studies concerning origins of honey to prevent honey labelling fraud [3]. However, pollen analysis is time consuming, laborious and highly skilled work. Despite of the various efforts to develop approaches that allow the automatic identification and classification of pollen grains [4], the observation and discrimination of features from relevant entities performed by qualified experts is still predominant [5]. Many industries, including medical and pharmaceutical, rely on the accuracy of this manual classification process, which is reported to be around 67% [6].

In recent years, deep learning, in particular, convolutional neural networks (CNNs) have become the dominant machine learning approach in the field of computer vision and to be specific in image classification and recognition. While still a combination of classification based on key image features, this new approach involves a model determining and extracting the features itself, rather than them being predefined by human analysts. Several machine learning techniques have been developed for classifying pollen grain images [6, 7, 8]. In [7], Daood et al. present a model that learns not only the features but also the classifier itself from a deep learning neural network.This method achieved 94% classification rate on a dataset of 30 pollen types. Sevillano and Aznarte in [8] presented an example of pollen classification which applied transfer learning to the POLEN23E image dataset achieving an accuracy of over 95% on the 23 pollen classes. Recently, the authors used the same approach [6] to classify up to 98% of the samples from a dataset with 46 different classes of pollen grains.

The goal of this paper is to suggest an approach for automated pollen detection and recognition given digital images produced by microscopes in the recently published POLEN73S image dataset [9] which has 2555 images from 73 different pollen grain types. This dataset includes more than three times as many pollen types and images as the POLEN23E dataset. We propose to use a state-of-the-art deep learning convolutional neural network (DenseNet-201) to classify grain pollen images since this model has shown promising results in object recognition [10].

International Conference of Numerical Analysis and Applied Mathematics ICNAAM 2020
AIP Conf. Proc. 2425, 140001-1–140001-4; https://doi.org/10.1063/5.0081614
Published by AIP Publishing. 978-0-7354-4182-8/$30.00

APPROACH OVERVIEW

A CNN is especially suited for image processing as it makes use of 2D hidden ('convolutional') layers to convolve the features with the input data. The main strength of CNN is that it eliminates the need for feature extraction by automatically extracting the more discriminant features of a set of training images.

Constraints of practical problems such as the limited size of training data refrain the performance of deep CNNs trained from scratch to be satisfactory. Since there is so much work that has already been done on image recognition and classification, we can use transfer learning technique to solve the problem. With transfer learning, instead of starting the learning process from scratch, we can start from patterns that have been learned when solving a similar problem. This way we can take previous learnings and avoid starting from scratch.

We are interested in learning a model that recognizes a pollen grain in an image. If a large dataset of pollen grain images is available, it would be relatively straightforward to learn the discriminant features and distinguish the objects. This is useful for new applications that will have a very large number of output categories. However, due to the large amount of required annotated training data and the speed of learning, with present day hardware it usually takes a very long time to train these networks.

To overcome this issue, we take the approach of fine-tuning recognition, derived from transfer learning. Fine-tuning recognition works as follows: Starting with a pretrained model trained on a large number of images, the model is subsequently fine-tuned using a relatively small amount of labelled data available for the final goal. In our case, this refers to the training dataset referred earlier.

Transfer Learning

In deep learning, transfer learning is a technique whereby a CNN model is first trained on a large image dataset with a similar goal to the problem that is being solved. One or more layers from the trained model are then used in a new CNN trained with sampled images for the current task. This way, the learned features in re-used layers may be the starting point for the training process and adapted to predict new classes of objects. Transfer learning has the benefit of decreasing the training time for a CNN model and can result in lower generalization error due to the small number of images used in the training process.

The weights in re-used layers may be used as the starting point for the training process and adapted in response to the new problem. This usage treats transfer learning as a type of weight initialization scheme. This may be useful when the first related problem has a lot more labelled data than the problem of interest and the similarity in the structure of the problem may be useful in both contexts.

DenseNet Convolutional Neural Network

We performed our experiments using the DenseNet201 model [10] which had shown exceptional performance in terms of ImageNet database [11] classification. DenseNet is built from dense blocks and pooling operations, where each dense block is an iterative concatenation of previous feature maps as Fig. 1 illustrates. Within those blocks, the layers are densely connected together: each layer gets the input from all preceding layers and passes on its own feature maps to all subsequent layers. This extreme reuse of residuals creates deep supervision because every layer receives more supervision from the previous layer and thus loss function will react accordingly and due to this methodology which makes it a more powerful network.

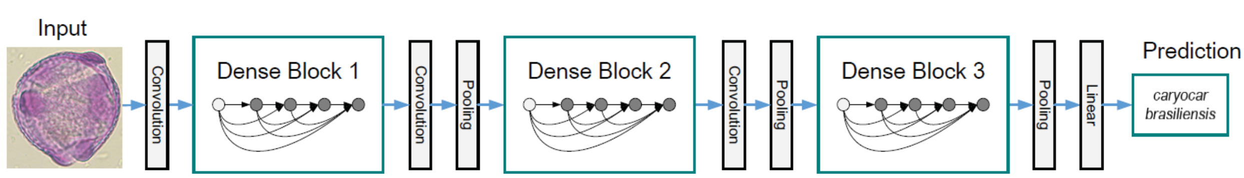

FIGURE 1. A deep DenseNet with three dense blocks. The layers between two adjacent blocks are referred to as transition layers and change feature map sizes via convolution and pooling. Adapted from [10]

Between each layer there is a concatenate layer where pre-activation batch normalization (BN), rectified linear units (ReLU), and 3×3 Convolution (Conv) are done with output feature maps. Since this concatenation operation used inside each dense block is not viable when the size of feature-maps changes, a transition layer which do convolution and max pooling is used between dense blocks.

After the last dense block, a global average pooling is performed by a fully connected layer and an activation function (ReLU) is applied. This activation function is a softmax function which normalizes output values from the last fully connected layer to target class probabilities, where each normalized value ranges between 0 and 1.

MATERIALS

The POLEN73S is an annotated image dataset, publicly available, for the Brazilian Savannah pollen types, according to its description in [9] contains pollen grain images taken with a digital microscope at different angles and distributed in 73 pollen types, containing 35 sample images for each pollen type except *gomphrena sp* that has 10, *trema micrantha* 34, and *zea mays* 29. So, in our experiment several synthetic images were generated through rotating and scaling the original images of these types in order to used the same number of samples for each pollen type which gives a total of 2555 pollen images.

The new POLEN73 dataset was split into a training dataset with 1788 images and a testing dataset with 767 images with equal number of samples for each type of pollen. As some of the original images have 4 dimensions we preprocessed those images in order to put all the images in the RGB format. No other changes have been made.

EXPERIMENTS AND DISCUSSION

The DenseNet201 has been tested with several learning rates, batch sizes and optimizers to find the best hyperparameter setup. All experiments described have been done using the same training and testing datasets.

We used a variety of data augmentation operations to increase the datasets diversity. This operation considerably increases the convolutional neural network training time but that also increases the accuracy accordingly. We decided to implement following augmentation methods: Horizontal and vertical flip, zooming, rotating and shifting. The range of every augmentation method was determined by empirical testing accordingly with the grain pollen dimension. Augmentation artifacts, for example as result of rotation, were filled in with a reflection of the original image content. The experiments performed better with the Adam than with the stochastic gradient descent with momentum (SGDM) optimizer, so, all the results were obtained with Adam optimizer.

The experiments were conducted on a computer with a NVIDIA RTX-2060 GPU with 6 GB of RAM with a total of 30 hours for each training process with 50 epochs. To properly evaluate the model performance, we use a testing set composed of images not seen by the model during the training process. Since they are completely unknown to the model, they allow us to anticipate its behaviour against new images. We obtained an accuracy of 98.4% during the training process and a value of 94.3% during the testing process. In the 73 pollen grain types of the dataset the model classified correctly all the images in 48 of them. In 16 of the types it failed only one image and in the other 9 it misclassified up to four images.

Figure 2 shows some examples of misclassified pollen grains. In the first row we have the input image and below the result of classification. As we can see, the model has misclassified the input images with very similar images of another type of pollen.

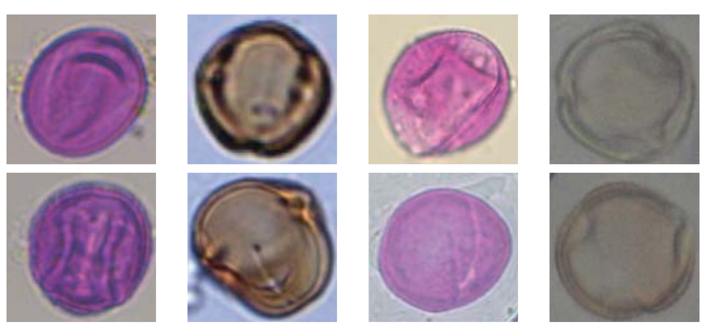

FIGURE 2. Example of misclassified pollen grains. First row: input image. Second row: misclassified identification of the pollen grain in first row.

Table 1 provides a summary table of previous studies, including class sizes and accuracy/success rates against our result. All the literature reviewed used a significantly smaller image dataset than the one used in this paper. As the number of pollen types is directly related to the classification performance, the results obtained in such simpler scenarios are not comparable to ours. The largest dataset, in terms of pollen types, was presented in [6], with 46 types, still smaller than the 73 pollen types dataset presented in this work.

TABLE 1. Previous attempts at automated classification of 10 or more pollen types using a CNN classifier, with number of taxa and highest reported correct classification rate (CCR). Adapted from [6]

Authors	Taxa	Reported CCR
Daood et al. (2016) [7]	30	94%
Sevillano and Aznarte (2018) [8]	23	97%
Khanzhina et al. (2018) [12]	11	96%
Menad et al. (2019) [13]	23	95%
Sevillano et al. (2020) [6]	46	98%
Proposed approach	73	94%

CONCLUSION

In this paper, we present an approach using deep learning convolutional neural networks to classify grain pollen images in the recently published POLEN73S image dataset which has 2555 images from 73 different pollen grain types. Using the transfer learning technique in the pretrained DenseNet network we achieved an accuracy up to 94% that represents one of the best successes rate, when weighted for the number of taxa, of any attempt at automated pollen analysis currently documented in the literature. These results confirmed the effectiveness of applying CNNs to pollen grain image recognition overcome human-like performance and inspired us to improve it staying under deep learning approach. The adjustment of the class weights used during training did improve the performance of the DenseNet gradually. Most likely, further improvements of deep learning techniques will achieve a perfect score in pollen recognition task even with large number of pollen types.

REFERENCES

[1] A. R. Laurence and V. M. Bryant, "Forensic palynology," in *Encyclopedia of Criminology and Criminal Justice*, edited by G. Bruinsma and D. Weisburd (Springer New York, New York, NY, 2014), pp. 1741–1754.

[2] G. D'Amato, L. Cecchi, S. Bonini, C. Nunes, I. Annesi-Maesano, H. Behrendt, G. Liccardi, T. Popov, and P. van Cauwenberge, Allergy **62**, 976–990 (2007).

[3] R. Ponnuchamy, V. Bonhomme, S. Prasad, L. Das, P. Patel, C. Gaucherel, A. Pragasam, and K. Anupama, PLoS ONE **9**, p. e101618 (2014).

[4] K. A. Holt and K. D. Bennet, New Phytololy **203**, 735–742 (2014).

[5] J. Buters, C. Antunes, A. Galveias, K. C. Bergmann, M. Thibaudon, C. Galán, C. Schmidt-Weber, and J. Oteros, Clinical and translational allergy **8** (2018).

[6] V. Sevillano, K. Holt, and J. L. Aznarte, PLoS ONE **15**, p. e0229751 (2020).

[7] A. Daood, E. Ribeiro, and M. Bush, "Pollen grain recognition using deep learning," in *Advances in Visual Computing* (Springer International Publishing, 2016), pp. 321–330.

[8] V. Sevillano and J. L. Aznarte, PLoS ONE **13**, p. e0201807 (2018).

[9] G. Astolfi and A. B. Gonçalves, POLLEN73S, https://doi.org/10.6084/m9.figshare.12536573.v1 (2020).

[10] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 1 (2017), pp. 2261–2269.

[11] ImageNet, http //www.image-net.org (2012).

[12] N. Khanzhina, E. Putin, A. Filchenkov, and E. Zamyatina, "Pollen grain recognition using convolutional neural network," in *Proceedings of European Symposyum of Artifitial Neural Networks* (2018), pp. 1–6.

[13] H. Menad, F. Ben-Naoum, and A. Amine, JERI (2019).