Deep Learning in Biometrical Studies

Elżbieta Kubera
Department of Applied Mathematics and Computer Science
University of Life Sciences in Lublin
elzbieta.kubera@up.lublin.pl

Objectives

Image recognition aims at automatization of tasks performed by human visual system. Every year, the ImageNet challenge shows the current state of knowledge in the field of image recognition, and the winners of the competition have been deep learning models (DL) over the past few years.

The main sources of allergens include pollen grains (Holgate et al. 2001). Information on the amount of these allergens in the air comes from pollen monitoring, where Hirst-type samplers actively suck in pollen grains from the air (Hirst, 1952), and traps them on a sticky tape. Each tape must be analyzed by appropriately trained staff — pollen grains are identified and counted under microscope. It is difficult and time consuming task, because the differences in the morphological structure of pollen grains of some taxa are very small and finding them largely depends on the position of the grain on the tape. Therefore, there is a need to automate or facilitate this process.

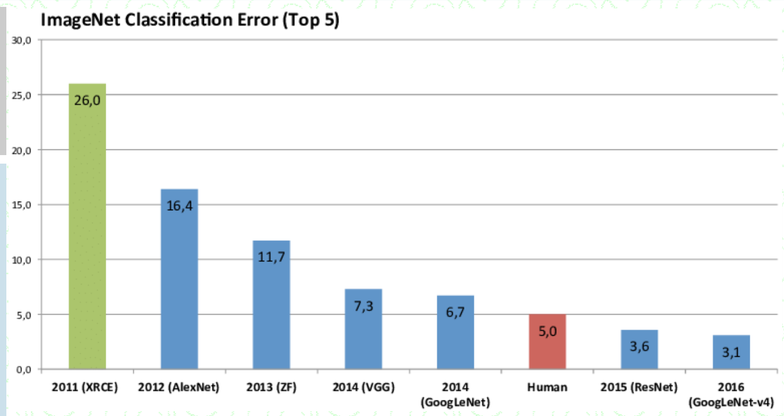


Figure 1 ImageNet Results Timeline (von Zitzewitz, 2017). After 2012 every single winner have used a Convolutional Neural Network as a model.

In this work, deep learning (which is also called deep neural networks) is used with biometrical data to recognize pollen taxa based on microscopic images. The deep learning methods are able to classify objects directly from the images, so no feature extraction or additional annotation is needed before model building. In this case, the pollen expert's work during data preparation can be minimized.

In the case of Poland, in spring it is birch pollen (*Betula*) that shows the strongest allergenic properties (Rapiejko, 2008; Piotrowska and Kubik-Komar, 2012). Hazel (*Corylus*) and alder (*Alnus*) belong to the same family (Betulaceae) and cause allergic cross reactions (Valenta et al. 1991, Rapiejko, 2008). Pollen seasons of these plants partially overlap and thus their pollen grains may be recorded during the same time. The aim of this work was to discriminate between the three investigated types of pollen (birch, alder and hazel) based on their microscope images.

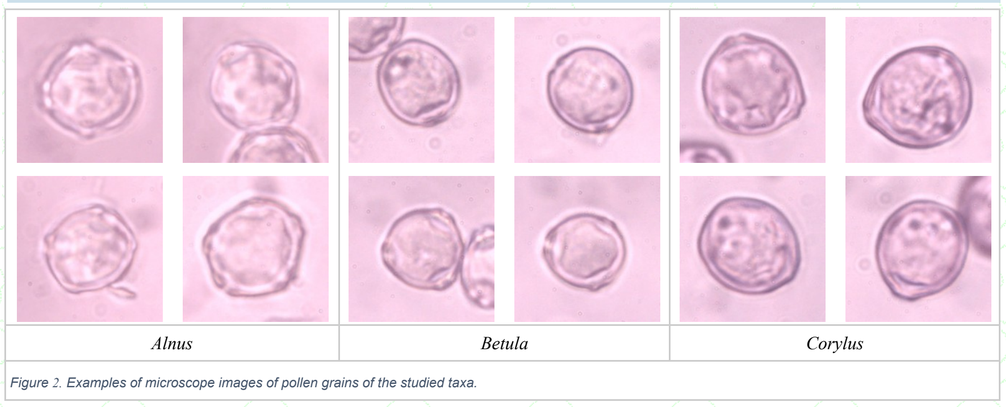


Figure 2. Examples of microscope images of pollen grains of the studied taxa.

Methods

The deep learning model is a kind of artificial neural network — it consists of layers of neurons. In a traditional – shallow neural network there are usually only three cascaded fully connected layers – one input, one output and one hidden layer of neurons, while deep networks are built of many layers of different types. The set and order of layers is called the network model, and is closely related to the specific purpose of the network. Convolutional Neural Network (CNN) (LeCun et al, 2015) is a widely used architecture of a deep neural network in multi-object recognition from images. This network is capable to perform classification directly from the image, so no feature extraction is needed before building the model. The weights of CNN are treated as features of the objects, and they are extracted across the whole image. This is why the objects in images can be shifted in the scene and still detectable by the network, making this type of network useful for object recognition in photographs.

CNN has many types of layers. The first layer is the input layer, which is initialized with the image pixel data. The first hidden layer in the CNN models is the convolutional layer that creates set of activation maps for the filters applied on the image. Each filter is a set of neurons that search for patterns in the image through the convolution. A single neuron of the convolutional layer performs the operation of filtering a small square of the image underlying it with a filter with the mask of the same size as the square of the image (eg 3x3 pixels). Filter masks, or convolutional layer weights, are initiated with random values, and during network training, they are modified so that the network recognizes objects with the best possible quality. The entire set of neurons gives location data of the given pattern in the image. CNNs typically uses multiple filters in parallel, each scanning for a different pattern. Neurons from different filters have different weights and search for different patterns. The neurons in the given filter have the same weights and bias to allow the filter to look for the same pattern in different sections of the image. The convoluted input is then send to the next layers.

Network training involves calculating the cost value from the training data, which is usually the difference between the predicted and the actual output. The rate at which cost changes with respect to weight or bias is called gradient. The layer gradient is a product of gradients from later layers. When the gradient is small, the network will train slowly. Gradients are much smaller in earlier layers, so deep learning requires very long training, and accuracy is often very low. This training problem is called vanishing gradient. To prevent this, the next layer after the convolutional layer is a common fully connected layer with Rectified Linear Unit (ReLU) function used as activation. After ReLU the activations are typically pooled in an adjacent layer to reduce the dimensionality of weights and biases, and get only promising patterns.

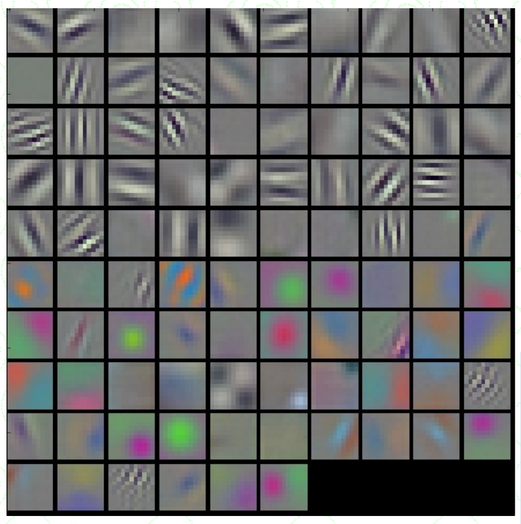
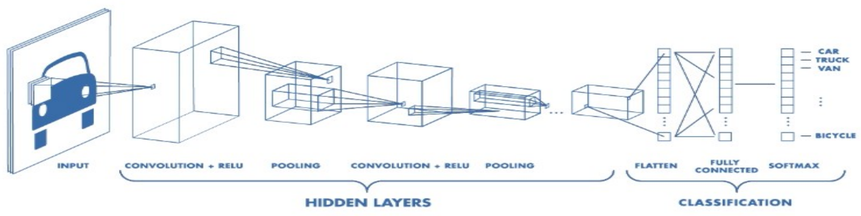


Figure 3 Typical-looking filters for color images on the first convolutional layer . Source: http://cs231n.github.io/understanding-cnn/

Figure 4 Architecture of a CNN. — Source: https://www.mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--148951276-5771.html



CNNs that are in use usually have an architecture with repeated sets of layers. Set1 is a convolutional layer followed by ReLU. This set can be repeated several times, and the repeated structure is followed by a pooling layer. This resulting combination forms Set2, which is repeated a few more times. Then there is a fully connected (FC) layer that acts as classifier. The last layer is output layer with the number of neurons equal to the number of classes, and the output values in these neurons show the probabilities that the object belongs to the class. Such architecture allows network to continuously build a complex pattern from simple ones, while reducing computational costs with dimensionality reduction. CNNs use the backpropagation method for the network training.

CNNs are a powerful tool, but they require millions of labeled data points for training. In convolutional networks early layers detect simple patterns and later layers recombine them, just as in the facial recognition example: early layers detect the edges in the image and later layers use these results to form facial features.

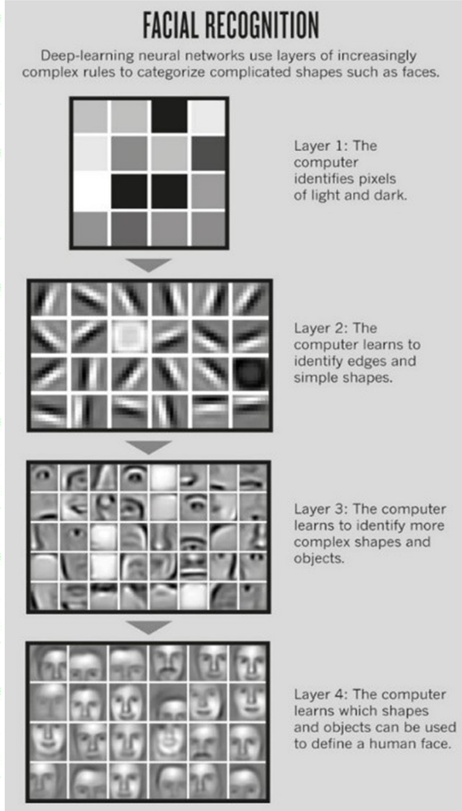


Figure 5 Abstraction levels in facial recognition. (Andrew Ng : http://www.nature.com/news/computer-science-the-learning-machines-1.14481)

Results

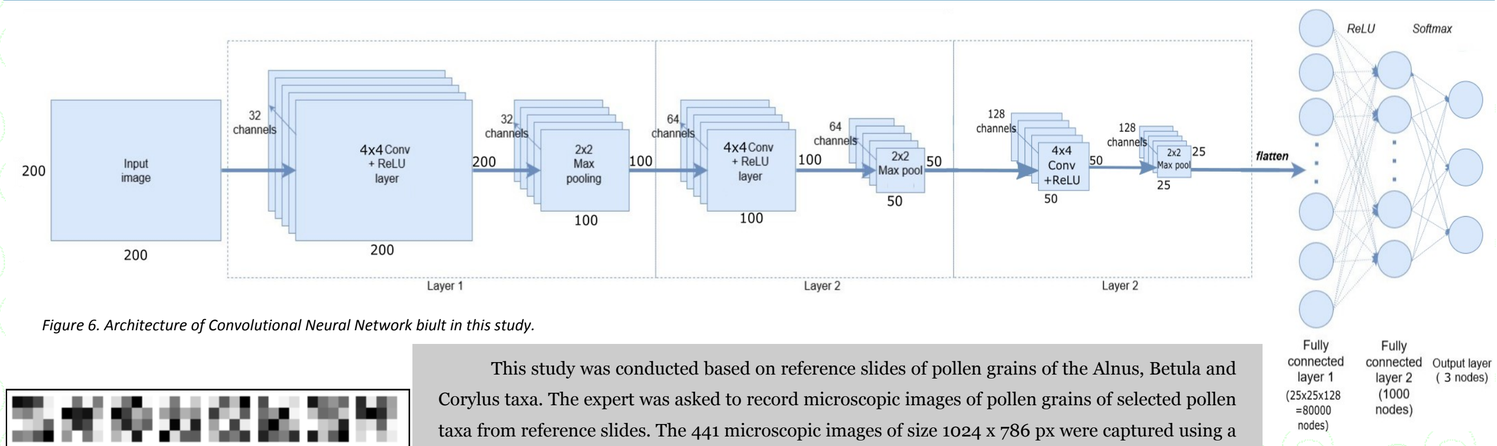


Figure 6. Architecture of Convolutional Neural Network biult in this study.

This study was conducted based on reference slides of pollen grains of the Alnus, Betula and Corylus taxa. The expert was asked to record microscopic images of pollen grains of selected pollen taxa from reference slides. The 441 microscopic images of size 1024 x 786 px were captured using a Nikon Eclipse E400 biological microscope at a magnification of 600x. The collection of 1274 images cropped manually to the size of 200 x 200 px was obtained from this data, each image representing one pollen grain in the center of the image, which may partially overlap with other grains of the same taxon. There were 406 images of Alnus grains, 435 images of Betula grains, and 433 images of Corylus grains. Some of the pictures were out of focus and some ot them contained overlapping grains. Each picture was labeled by the taxon name, and there was no multi-label objects in collection. The set of pictures was fed into the convolutional neural network (CNN) consisting of three convolutional layers, each with 4x4 filters. Figure 6 shows the structure of the network. The final masks of convolutional filters in each layer are shown in Figure 7.

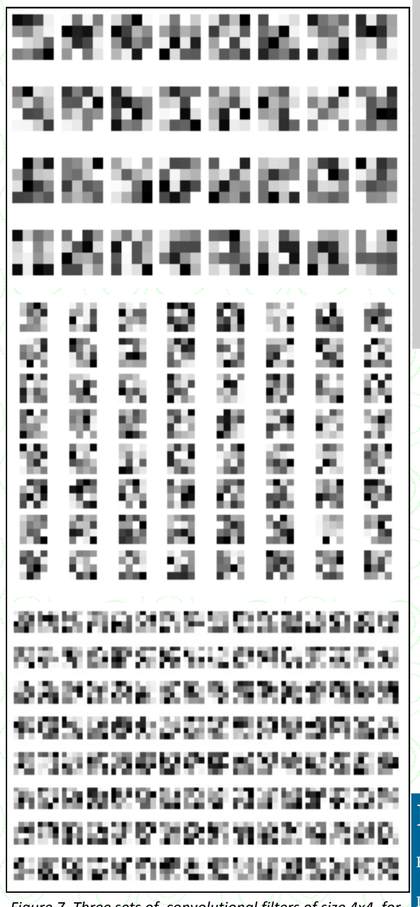


Figure 7. Three sets of convolutional filters of size 4x4, for first, and third layer, after 100 epochs of training the CNN.

The training process was performed using 1224 randomly chosen images and test set contained the remaining 50 samples. Adam optimizer was used to train network. The 100 epochs of training took 3 hours and 18 minutes on the mobile workstation equipped with 32 GB of memory and Intel® Core™ i7 processor with 4 cores, 8 threads and a 2.4GHz clock. The training history is presented in Figure 8. Additionally two analogical models were built, with filter sizes 3x3 and 5x5, and the results are compared in the table below.

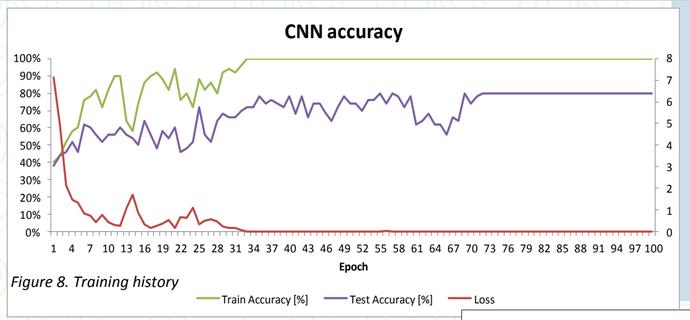


Figure 8. Training history

Figure 9. The comparison of CNN accuracies with respect to the filter size

	Filter size 3x3	Filter size 4x4	Filter size 5x5
Final accuracy on test set (after 100 epochs)	68%	**80%**	72%
The best accuracy on test set	76%	**80%**	78%

Conclusions

The results show that the deep learning method has achieved 80% of accuracy in the classification of taxa, which is a fairly good result in the difficult task of classifying objects directly from the images. The pollen grains analyzed in this study differ from each other primarily in their shape and number of pores; the alder pollen grain has 5 pores, sometimes 4 or 6, whereas the birch and hazel pollen grains – 3 (only 2 are sometimes visible, depending on the position). Unlike the birch pollen grain, the hazel pollen grain is more triangular in polar view and has less protruding pores. The studied pollen grains also differ in oncus size. Microscopic visibility of pores and onci greatly depends on the position of the grain on the tape. While manual observation the researcher can control the focus setting of the microscope in order to observe the grain at different depths and thus better perceive the characteristic structural elements. When analyzing a single microscope image, it is much more difficult to correctly describe these characteristics. Future works on the procedure of the automated identification of pollen taxa should also include automated microscope control and focus adjustment to particular grains.

Discussion

A pilot study on pollen recognition of the same three taxa was conducted on the basis of morphological features measured manually (Kubik-Komar et al., 2018). The values of these features were read under the microscope for 225 pollen grains (75 grains for each of the analyzed taxa). The decision tree was used as a classifier, and 94% accuracy was achieved using the best subset of features. The actually presented research consist in recognition pollen taxa by means of deep neural networks, directly from microscopic images of pollen grains, without the necessity of earlier feature extraction, and an accuracy of 80% on the test set has been achieved.

Many of previous works on taxa discrimination were carried out as the image-based classification using neural networks. Li et al. (1999) took into consideration pollen grains of 4 plant types typical of New Zealand. The authors themselves admit that these grains can be easily discriminated due to their shape, but in the article they differentiated them only by the texture features, obtaining 100% of correctly classified instances. France et al. (2000) were the first to create an automated system for detection and classification of pollen grains of Polemonium caeruleum, Nymphaea alba, and Crataegus monogyna, obtaining the average classification success rate at a level of 83%. The three above-mentioned species produce pollen grains that clearly differ from one another, in particular in terms of their exine and aperture structure. Tello-Mijares et al. (2016) addressed the problem of automated analysis of pollen material collected in Mexico (12 plant taxa typical for this area), obtaining a high pollen recognition rate, as much as 96%. Most of the investigated pollen grains strongly differ both in shape and in size, but among them there are also grains that are similar to one another in terms of their structure. A direct comparison of the percentage values is unreliable, due to the different levels of similarity between the taxa investigated in these studies as well as different classifiers or feature vectors.

Future work will include the further development of the network architecture, and image database of pollen grains , which can probably increase accuracy of the classification.

References

France, I, Duller, A, Duller, G, Lamb, H. (2000) A new approach to automated pollen analysis. Quaternary Science Reviews, 18, 537–536.
LeCun Y., Bengio Y., Hinton G. 2015. Deep learning. Nature. 521 (7553): 436–444
Li, P. & Flenley, J. R. 1999. Pollen texture identification using neural networks. – Grana 38: 59–64. ISSN 0017–3134
Holgate, S.T, Church, M.K. and L. M. Lichtenstein (Eds.) (2001). Allergy, (2nd Ed.), 3-16, London, Mosby International Ltd.
Hirst J.M.: 1952, An automatic volumetric spore trap. Annals of Applied Biology 39, 257–265.
Kubik-Komar A, Kubera E, Piotrowska-Weryszko K. (2018) Selection of morphological features of pollen grains for chosen tree taxa. Biology Open (2018):bio.031237.Web.1 Aug.2018
Kubik-Komar A, Kubera E, Piotrowska-Weryszko K, Weryszko-Chmielewska E. (2015.) Application of decision tree algorithms for discriminating among woody plant taxa based on the pollen season characteristics. Archives of Biological Sciences, 67 (4), 1127-1135.
Kawashima S, Thibaudon M, Matsuda S, Fujita T, Lemonis N, Clot B, Oliver G (2017) Automated pollen monitoring system using laser optics for observing seasonal changes in the concentration of total airborne pollen. Aerobiologia 33 (3), pp 351–362.
Piotrowska, K. , Kubik-Komar, A. (2012). The effect of meteorological factors on airborne Betula pollen concentrations in Lublin (Poland). Aerobiologia, Vol. 28 (4), 467-479.
Rapiejko, P. (2008). Alergeny pyłku roślin / Allergens of plants pollen. Medical Education, 62-69.
Santiago Tello-Mijares and Francisco Flores (2016). A Novel Method for the Separation of Overlapping Pollen Species for Automated Detection and Classification. Computational and Mathematical Methods in Medicine, vol. 2016, Article ID 5689346.
Valenta R, Duchene M, Pettenburger K, Sillaber C, Valent P, Bettelheim P, Breitenbach M, Rumpo DH, Kraft D, Scheiner O (1991) Identification of profilin as a novel pollen allergen: IgE autoreactivity in sensitized individuals. Science, 253, 557-560.
von Zitzewitz, Gustav. (2017). Survey of neural networks in autonomous driving. ADVANCED SEMINAR SS 2017 . Avaialable from ResearchGate 06/09/2018.