Automatic Pollen Grains Counter

Aysha Kadaikar, Benjamin Guinot
Laboratoire d'Aérologie
Toulouse, France
{aysha.kadaikar, benjamin.guinot}@gmail.com

Maria Trocan, Frédéric Amiel, Patricia Conde-Cespedes
Institut Supérieur d'Electronique de Paris
Paris, France
{maria.trocan, frederic.amiel, patricia.conde-cespedes}@isep.fr

Gilles Oliver, Michel Thibaudon
Réseau National de Surveillance Aérobiologique
Brussieu, France
gilles.oliver@rnsa.fr, michel.thibaudon@wanadoo.fr

Roland Sarda-Estève, Dominique Baisnée
Laboratoire des Sciences du Climat et de l'Environnement
Saint-Aubin, France
{roland.sarda-esteve, dominique.baisnee}@lsce.ipsl.fr

***Abstract*—This paper deals with the problem of counting the amount of pollen grains (density of pollen) in a view acquired on a daily basis by a dedicated device. The grains are stuck on a ribbon which is analyzed by a microscope. This task is currently performed by a human operator who has to analyze quite 300 microscopic slides representing 10% of the ribbon. This task requires a particularly high concentration. Moreover, the viewer has to identify the pollen grains among water drops and soot spots. We propose an algorithm composed of four main steps to perform this task automatically. The basic detection of pollen grains, relying on their shape and color, is completed by pre-processing and post-processing operations to handle specific cases like broken grains or grains at the border of images. Finally the efficiency is improved by using a neural network to refine the results. Our automatic counter has been compared with the pollen density obtained by the manual counting and our program has proved its high accuracy.**

***Index Terms*—pollen counting, pollen features, neural network, transfer learning**

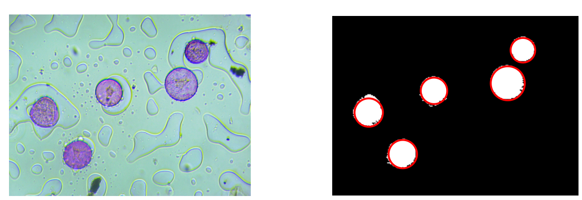
I. Introduction

Interest in pollen counting and classification has been growing during the last decades, as it is used mainly for allergy prevention purpose and also for other applications, such as paleoclimatic reconstruction and palynology [1], [2]. In fact, according to different studies, between 8 and 25 % of the world population suffers from allergic rhinitis highly caused in particular by pollen [3], [4]. However pollen analysis is a difficult task requiring in most cases a highly qualified human operator capable of counting and identifying the different pollen grains. This latter uses a microscope to analyze the collected pollen samples that have been previously deposited on an observation slide. One can imagine the heaviness of the task because of the huge number of pollen grains that can be found in a single slide depending on the seasons. Another difficulty of pollen analysis is that the collected samples not only contain pollens grains but also many others particles which are present in the ambient air like soot, pollution particles, plant residues, mildew or water droplets. These particles add a difficulty in the analyzing process as some of them can have an appearance similar to pollen grains by either their shape, their color or both of them. Sometimes pollution particles or water droplets layering with a pollen grain can mistake the pollen taxon identification. Finally, the pollens grains can be found folded or broken, which makes difficult their identification and also their counting.

An example of a microscope slide view containing pollen grains is given in Fig. 1(a). The pollen grains have been colored in purple using fuchsine, substance putting them and their features in evidence. Note that some plant residues can also be colored during the coloration process.

In this paper, a new method is proposed to automatize the pollen grains counting. The algorithm not only uses the color and shape of the pollen grains to detect them, but it also takes into account some pollen features in order to make it more robust to errors induced by broken pollen grains. The proposed method finally uses a neural network to retain and count only pollen grains.

The rest of the paper is organized as follows: Section II presents the developed method counting the number of pollen grains on a given microscope slide view. Section III discusses the experimental results. Finally, Section IV draws the conclusion and give insight of further future works.



(a) (b)

Fig. 1: Example of microscope pollen slide view I (a) and associated detection (b).

II. Pollen Grains Counter

This section describes the different steps of the proposed Hybrid Counting (HC) algorithm. It is mainly composed of four steps: (i) a pre-processing to detect pollen grains at the image border; (ii) a detection of the pollen grains based on their shape and their color; (iii) a post-processing to correct some specific counting errors and (iv) the usage of a neural network to improve the detection and counting accuracy. The algorithm composed of the succession of only the first three steps is denoted HC-algorithm. Moreover, we refer to the overall algorithm using the four steps including the usage of the neural network as the HCA-algorithm. The developed algorithm works on a given sharp microscope slide view obtained using the process described in [5].

The first subsection focuses on the basic detection of the pollen grains based on their color and shape. Then, we present the pre and post-processing steps increasing the counting robustness of the proposed algorithm. Finally, we describe the usage of a neural network which improves the algorithm accuracy.

A. Detection of pollen grains based on color and shape.

Considering a sharp microscope slide view (denoted I) as represented in Fig. 1(a), the HC-algorithm firstly aims at detecting all the pollen grains present in the view. As they have been previously colored in purple using fuchsine during the slide making process, the HC-algorithm focus on detecting the presence of purple pixels. A transformed image is firstly created by converting the original image into the HSV domain. Then the transformed image is filtered to keep only a range of purple pixels. Finally the filtered image is binarized where white pixels correspond in fact to the purple pixels in the original image.

Next, a succession of morphological closing and opening operations are applied on the binary image. The resulting image is then filled to get ready for the shape detection.

At this step, the HC-algorithm keeps only the regions larger than a threshold as potential pollen grains, very small regions are assumed to be noise. Then the HC-algorithm starts the shape detection using the Hough transform to detect shapes that are more or less circular (the most common shape of pollen grains). This step allows to eliminate air particles other than pollen grains that have not been eliminated yet and which have a shape different from pollen grains. An example of the pollen detection is given in Fig. 1(b).

Finally the HC-algorithm sums the number of circles detected as the number of pollen grains in the view.

B. Additional robustness tools

The HC-algorithm includes a pre-processing step to consider the case of detecting and counting non-entire pollen grains at the border of the images. This process is described in the first subsection. Moreover, after the color and shape based detection, the HC-algorithm performs a post-processing step to correct errors induced by specific cases, such as broken pollen grains for example. The post-processing is described in the second subsection.

1) Pre-processing: counting the pollen grains at the image border: One difficulty in pollen grains counting concerns the non-entire grains at the border of the image. If the visible part of the pollen is too small, there is a risk that the Hough transform does not detect them. To overcome this problem, we have expanded the images by adding the symmetric of the four borders to the original images. We have considered just a few numbers of pixels to build these symmetric parts, in order to avoid duplicating already fully visible pollen grains. An example of the image enlargement is given in Fig. 2. To avoid duplicate pollen counting, we consider a decision rule to know whether the pollen grains located inside the added border should be counted or not, as represented in Fig. 3. If the center of the detected pollen is located beyond the borders of the original image, the algorithm considers that this pollen has already been counted and that the considered pollen grain is a duplicate one. In Fig. 3 example, the pollens located in the extended border in the case (a) and (b) will not be counted again. Only the case (c) will be counted as a pollen grain.

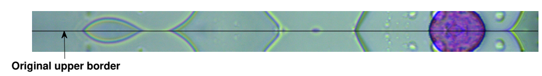


Fig. 2: Image extension at the top of the image.

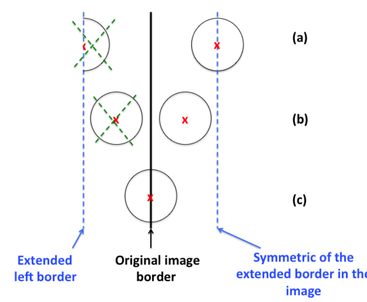


Fig. 3: Counting decision rule at the image border.

2) Post-processing: counting error correction: This subsection concerns some specific cases inducing error counting: the case of cypress pollen grains that are often found broken and the case of pine pollen grain that has an original shape different from the common circle shape of other pollen grains.

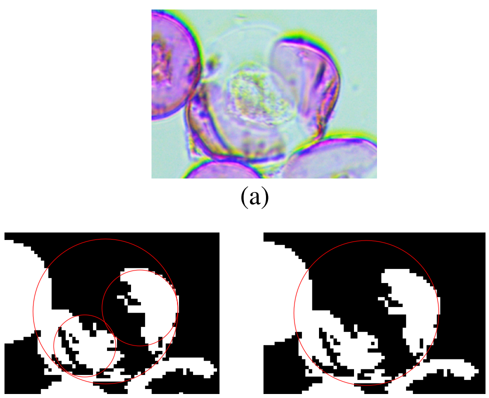
Indeed, concerning the broken pollen grains, each broken part should not be counted like different pollens. An example of a broken cypress pollen grain is given in Fig. 4 (a). This case should be distinguished from the case where two pollen grains appear side by side (cf Fig. 5 (a)). In both cases, applying the Hough transform results in detecting three circles (see Fig. 4 (b) and Fig. 5 (b) respectively) but in the first case it should be counted as a single pollen grain, while in the second case, the algorithm should count two pollen grains.

In a similar way, the pine pollen can also induce a counting error: depending on the angle it lies on the microscope slide, it appears to be composed of three main circles. An example of pine pollen is given in Fig. 6 (a). In this case also the three detected circles (see Fig. 6 (b)) should be counted as a single pollen grain.

Thus we propose to follow some basic decisions rules to detect a broken cypress pollen, a pine pollen or two side-by-side pollen grains. These rules depends on the position of the circles and their consistency, and are defined as follows in the case the algorithm finds two side-by-side circles included in one bigger circle:

- if the number of white pixels in each of the two small circles is below a pre-defined threshold and if the rest of the big circle is almost empty, weare considering a broken cypress pollen grain,
- if the number of white pixels in each of the two small circles is above a pre-defined threshold and if the rest of the big circle is almost full, we are considering a pine pollen grain,
- if the number of white pixels in each of the two small circles is above a pre-defined threshold and if the rest of the big circle is almost empty, we are considering two side-by-side pollen grains.

Following these basic rules leads to a better detection and counting of the number of pollen grains as shown in Fig. 5 (c), Fig. 4 (c) and Fig. 6 (c) where the wrong detections have been corrected. Indeed making these assumptions could results in some errors in counting, but in most cases the experimental results have shown the benefit of these decision rules.

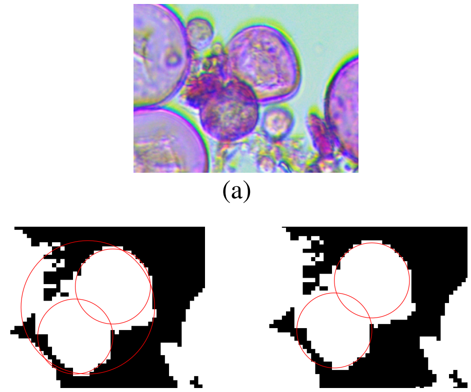


(a) (b) (c)

Fig. 4: Original view of a broken cypress (a), associated detection (b) and associated corrected detection (c).

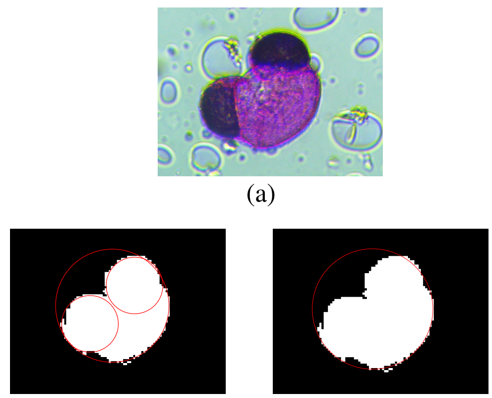
C. *Usage of neural network*

Finally the HC-algorithm detects all the pollen grains present in a view. But in some cases, it also counts some elements which are not pollen. These errors occur with some particles that have similar color and shape to that of the pollen grains. To handle this problem we performed transfer learning using the Alexnet [6] neural network by modifying only the last three layers of the network to adapt it with our binary classification. We had previously trained the neural network: the pollen grains images detected by the HC-algorithm have been used to establish a set of segmented images labeled as ”pollen” or ”not pollen”. The neural network was trained with 700 images for each set. Then, we used this trained neural



(a) (b) (c)

Fig. 5: Original view of side-by-side pollen grains (a), associated detection (b) and associated corrected detection (c).



(a) (b) (c)

Fig. 6: Original view of a pine pollen grain (a), associated detection (b) and associated corrected detection (c).

network to classify all the new images resulting from the HC-algorithm pollen detection. The application of the HC-algorithm followed by the usage of the neural network to classify the detected images is called the HCA-algorithm and it improves the counting accuracy.

A summary of the algorithms steps is given in Fig. 7.

III. Experimental Results

This section first discusses the performance of the HC-algorithm before and after using Alexnet [6] and and also by comparing the results to the manual counting of the pollen grains in the slides. Then we discuss the computation complexity and processing times of the algorithms.

1) Performance of the HC-algorithm and the HCA-algorithm: To build our image dataset, we have considered two microscope slides containing particles present in the ambient air, each one acquired in a different French city (Narbonne and Vichy). For each slide, we have acquired a great number of images covering around 10% of the slide

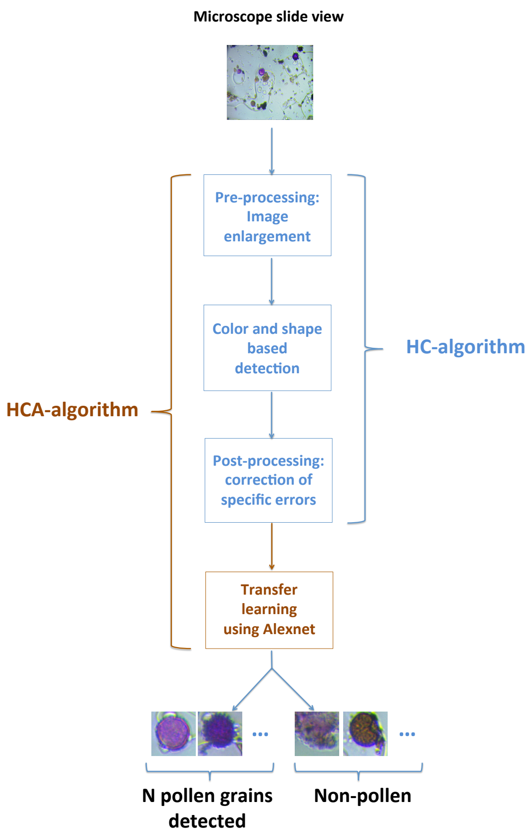


Fig. 7: Main steps of the HC-algorithm and the HCA-algorithm.

(corresponding to 3 full horizontal lines of the slide). Table I gives an overview of the HC-algorithm performance before using Alexnet and after using Alexnet (under the column "HCA-algorithm"), in comparison to the manual pollen grains counting for the two cities.

TABLE I: Number of counted pollen grains

City	HC-algorithm	HCA-algorithm	Manual counting
Narbonne	426	205	200
Vichy	107	73	79

One can see that the HC-algorithm before using Alexnet has found more pollen grains than the manual counting for all cities. This is due to the noise present in the slides looking similar to pollen grains. However, the problem of over-counting is solved by adding the binary classification step using Alexnet. The performance of the HC-algorithm using Alexnet is very close to the manual counting.

2) Computational complexity and processing time: Considering microscope slide images of size $n \times n$ pixels, the HC-algorithm alone has a complexity of $O(n^4)$ mostly induced by the search of circles performed using the Hough transform. The HCA-algorithm includes an additional step of transfer learning in comparison to the previous algorithm. To estimate the complexity of this additional step, we consider the time required for just one pass of forward propagation in the Alexnet network. One pass consists of mainly two operations: matrix multiplication and activation function calculation. The former is the most time consuming. The number of layers of Alexnet being fixed, the computational complexity of the this last step can be estimated as $O(n^3)$ [7]. Thus, the complexity of the HCA-algorithm remains $O(n^4)$ because of the previous steps of the algorithm.

For the experimentations, we have used a computer with an Intel Core i5 processor of 2.4Ghz and a RAM of 4GB. The simulations have been performed under Matlab (version 2018b). It takes an average of 4200 seconds to detect the pollen grains present in a portion covering 10% of a slide. The HCA-algorithm just takes in average 100 more seconds to classify the detected images while the training alone of the neural network in the conditions described in section III-2 lasted about 3400 seconds. The processing time of the HCA-algorithm remains reasonable considering this application as it takes one week to collect the pollen samples on a slide and just a bit more than 1 hour to count the pollen grains using the HCA-algorithm.

IV. Conclusion

In this paper we presented a new method to count the number of pollen grains on a microscope slide image. The presented scheme relies on the shape, the color and other features of the pollen grains to count. The proposed algorithm has shown globally good performance in comparison to the manual counting. In future works, we will consider the identification of the detected pollen grains.

V. Acknowledgment

The authors would like to gratefully acknowledge the support of the National Research Agency and the STAE foundation under the auspices of the Saint-Exupery Technological Research Institute without which the present study could not have been completed.

References

[1] Li, Ping, and J. Flenley, "Pollen texture identification using neural networks", Grana, vol. 38.1, pp. 59-64, 1999.
[2] K. A. Holt and K. D. Bennett. "Principles and methods for automated palynology". New Phytologist, 203(3), 735-742, 2014.
[3] R. Pawankar, C. E. Baena-Cagnani, J. Bousquet, G. W. Canonica, A. A. Cruz., M. A. Kaliner, ... and K. Henley, "State of world allergy report 2008: allergy and chronic respiratory diseases", World Allergy Organization Journal, vol. 1(1), S4, 2008.
[4] M. S. Dykewicz and D. L. Hamilos. "Rhinitis and sinusitis", Journal of Allergy and Clinical Immunology, 125(2), S103-S115, 2010.
[5] A. Kadaikar, M. Trocan, F. Amiel, P. Conde-Cespedes, B. Guinot, R. Sarda Estève, D. Baisnée and G. Oliver, "Sharp image detection for microscope pollen slides observation", in the proceedings of ACIIDS, 2019.
[6] A. Krizhevsky, I. Sutskever and , G. E. Hinton. "Imagenet classification with deep convolutional neural networks", In Advances in neural information processing systems, pp. 1097-1105, 2012.
[7] F. Kasper, "Computational complexity of neural networks." https://kasperfred.com/posts/computational-complexity-of-neural-networks. Date last accessed: 02.26.2019.