Algerian Journal of Chemical Engineering

ISSN:2773-3068

Journal homepage: http://www.journal.acse.science/index.php/ajce/index

Original Article

Clustering and discernment of Algerian bee pollen using an image analysis system

A. Rebiai[a], B. Ben Seghir[b,c], H. Hemmami[c]*, S. Zeghoud[d] , T. Siham[d], I. Kouadri[e], H. Terea[f], F. Brahmia[d]

[a]*Department of Chemistry, Faculty of Exact Sciences, University of Echahid Hamma Lakhdar, El Oued 39000, Algeria.*

[b,c]*Laboratory of Industrial Analysis and Materials Engineering (LAGIM). University 8 May 1945 Guelma, BP 401, Guelma 24000, Algeria.*

[c]*Department of Process Engineering, Faculty of Technology, University of Echahid Hamma Lakhdar, El Oued 39000, Algeria.*

[d]*Laboratory Valorisation and Technology of Saharan Resources (VTRS). University of El-Oued, BP 789, El-Oued, 39000, Algeria.*

[e]*Laboratory of Applied Chemistry (LCA), Department of Material Sciences, University May 8, 1945 Guelma, BP 401, Guelma 24000, Algeria.*

[f]*Laboratory valorization and promotion of saharan resources university of kasdi merbah ouargla 30000 Algeria.*

ARTICLE INFO

Article history:
Received 02 May 2021
Revised 03 june 2021
Accepted 15 July 2021

Keywords:
Algerian bee pollen;
Microscope images;
Classification; HCA;
Properties modeling.
.

ABSTRACT

In this paper, we suggest a framework for multi-focal image classification and identification, the methodology being demonstrated on microscope pollen images (image processing and classification techniques). The framework is intended to be generic and based on a brute force-like approach aimed to be efficient not only on any kind, and any number, of pollen images (regardless of the pollen type), but also on any kind of multi-focal images. Microscope images information obtained from bee pollen samples (72 samples) of different floral origin from various Algerian counties were used to formulate a method for rapid classification using Hierarchical Cluster Analysis (HCA). Both stages of the framework's pipeline are planned to be used in an automated fashion. First, the optimum focus is chosen using the absolute gradient method. Then, pollen grains are collected using a coarse-to-fine method involving both clustering and morphological techniques. Finally, features are extracted and selected using a generalized method, and their classification is checked with using HCA. Our findings indicate that HCA meets the demands for automatic pollen detection making it an alternative method for research concerning pollen.

1. Introduction

Bee and honey hive related products, along with other natural resources of essential amino acids, flavones, vitamins, polyphenols, and enzymes, enjoy considerable traction on the social marketplace. Pollen, bee-bread, royal jelly and bee-venom are all indicated by an outstanding anti-allergenic sequence[1].

Classification of pollen grains has become a costly analytical process involving the detection and classification of features by a professionally qualified palynologist. Still the most detailed and reliable method. But it does obstruct scientific progress, taking significant time and resources [2].Recent improvements in the instruments used to collect, process and analyze fluorescence signals have now allowed the classification and counting of pollen grains[3].

Those problems can be solved by automated identification of pollen grains, generating strictly objective results faster[4]. Such an instrument will prove invaluable in flora studies. For Flenley these advantages were obvious[5].At that moment, however, the idea was intractable. Mainly because of limitations on the

* *Corresponding author.*
E-mail address: Email: hemmami.h@gmail.com
Peer review under responsibility of University of Echahid Hamma Lakhdar.
http://dx.doi.org/ 10.5281/zenodo.5105756

technology. Nowadays, technology is no longer an obstacle, and thanks to computer vision the device being addressed is a reality[6].

It is generally easy and unambiguous to classify the grains as fluorescent; however, often pollen grains show intermediate fluorescence, so their classification is difficult. Another type of classification error occurs when the grain fluorescence in a microscope field is uniform because the ratio of fluorescent to non-fluorescent grains is either very large or very small. In such situations, a slight variation in the intensity of the fluorescence will allow the operator to take some grains into a classification[7].

However, this method is comparatively less reliable as a subjective classification tool. Various spectral techniques, such as Fourier transform infrared (FTIR) spectroscopy, Raman spectroscopy (RS), and fluorescence spectroscopy (FS), have so far put a lot of effort into improving the performance of pollen classification [8-10].

According to Pinnick et al., Fluorescence is a useful index for differentiating among biological and non-biological airborne particles; hence, fluorescence microscopy is a practical method for pollen grain investigation [7, 11, 12].

A pollen analysis system requires not only a pollen grain location in the picture but also a botanical type classification of the pollen grain. Fluorescence images of bee pollen grains can be used to identify pollens, so the fluorescence technique is an important method for classifying pollens [3, 7, 13].

The application of computational image processing and machine learning algorithms to recognise and characterise disease patterns on digitised tissue slides [14]. In the field of breast cancer pathology, a variety of computational imaging methods have been recently applied for problems such as I detection of mitoses [15-17] , tubules[18], nuclei[19], (ii) association of quantitative histologic image features and molecular features of breast cancer aggressiveness [20], and (iii) recognition of histologic image features that are predictive of breast cancer outcome and survival [21]. Deep learning has also been used to identify image-based plant diseases [22].

We depend on this study to classify the variables (pictures, Algerian bee pollen) according to specific properties in different groups with chemometrics analysis, and they are arranged inside hierarchical clusters, where the variables with similar characteristics are positioned within one cluster distinguished by features that are different from the rest of the other clusters. We may also find the degree of similarity in the form between the samples by their components using a hierarchical.

2. Materials and methods

2.1. Materials

Ethanol (C_2H_6O, M = 46.07 g /mol, 99.8%) manufactured bya company Honeywell. Distilled water (H_2O) prepared in the laboratory.

1.2. Apparatus

Sensitive scale (EXPLORER) (0.1 mg) made (OHRUS), Optical Microscope equipped with a digital camera (OPTIKA B-350), Scanner.

1.3. Bee pollen sampling

Seventy two samples of bee Pollen were collected from various locations and states of Algeria (Figure 1), where they were collected by specialists in beekeeping, in a time spanning between 2016 - 2018. Table 1 shows a summary about the various collected propolis samples.

1.4. Sample Preparation

2 mg of each sample of different pollen was weighed and placed in test tubes, we added 1 ml of ethanol (C_2H_6O) "Ethanol (alcohol) was chosen to avoid the spread of bacteria in the medium". After an hour, we take the remaining residue from the decomposition process and place it on a glass slide, add drops of distilled water to dilute the samples, then cover the slide with a coverslip, and add drops of oil to "clarify the images better", and place them in an optical microscope equipped with a computer and a camera to capture different Pictures of pollen zoom 1000 times (Figure 2).

1.5. Unsupervised analysis of bee pollen images

Deep learning [23] has revolutionised the field of biomedical image processing. Conventional methods have used problem-specific algorithms to represent images with manually designed features, such as cell morphology, count, strength, and texture [24].

Feature learning with deep convolutionary neural networks is implicit, and network training usually focuses on specific tasks, such as mammography detection of breast cancer[14], subcellular protein localization [25], or plant disease detection [22]. Training a deep network normally involves a large number of images, which limits its usefulness.

Here, we use the Orange Data Mining (Orange3-3.13.0-Python36 Pro 2018. University of Ljubljana, Slovenia) visual programming toolbox to simplify the study of bee pollen images by incorporating deep-learning embedding, machine learning processes and data visualisation.

1.6. Hierarchical cluster analysis (HCA)

Cluster analysis is really the process of grouping objects into clusters which have meaning in the context of a specific problem. Clustering methods are unsupervised types of analysis, as there are no a priori concepts of cluster membership [26]. HCA helped classify the samples analysed into groups of similar characteristics [27].

In this work, we checked that the best results were obtained using a metric based on Cosine's correlation coefficient and the average linkage process.

Fig. 1. Geographical locations from which bee pollen samples were obtained.

Table 1. A summary of Propolis samples.

Code		Region	Forest Cover	Date of harvest	Source
P1	A1	Bouira	Intensive	2017	*Acer negundo* L
	J1				*Acer opalus* subsp.
	O1				*Anemonastrum narcissiflorum* L
	N1				*Ajuga reptans* L
P2	J2	Mtija	Intensive	2017	*Ajuga reptans* L
	O2				*Soybean*
	Js2				*Spotted yellow loosestrife*
	B2				*Red sand-spurrey*
P3	J3	Skikda	Intensive	2017	*Pink corydalis*
	O3				*Pearly everlasting*
	B3				*Caliculé Leatherleaf*
	V3				*Canada fly honeysuckle*
P4	J4	Constantine	Intensive	2017	*Trembeling aspen*
	O4				*Common storksbill*
	A4				*Leatherleaf*
	Js4				*Crocus sativus* L
P5	J5	Tipaza	Intensive	2017	*Bitter Wintercress*
	Jo5				*Birch*
	Js5				*Common ragweed*
	O5				*Buckwheat*
	V5				*European columbine*
	N5				*Brunet's milk-vetch*
	R5				*Holly*
P6	J6	El-Bayadh	Intensive	2017	*Mexican dock*
	O6				*Plantain lily*
	Jo6				*Meadow geranium*
	N6				*Tatarian honeysuckle*
P7	J7	Tipaza	Intensive	2017	*Common wormwood*
	O7				*Everlasting pea*
	Js7				*Garlic mustard*
	N7				*European columbine*
	R7				*Bitter wintercress*
	Rs7				*Round-leaved dogwood*

P8	J8	Bouira-Boumerdès	Intensive	2017	*European bistort*
	O8				*Basswood*
	Js8				*Wild sarsaparilla*
	R8				*Brunet's milk-vetch*
	B8				*Wild sarsaparilla*
	N8				*Prostrate knotweed*
P9	J9	Laghouat, Blida and Médéa	Average density	2017	*Creeping buttercup*
	O9				*Broad fruited burred*
	B9				*Northern marsh yellowcress*
	Bn9				*American beech*
	R9				*Staghorn sumac*
	N9				*Tall meadow-rue*
P10	J10	Tizi-Ouzou	Intensive	2017	*Bird's-eye speedwell*
	O10				*Agropyron caninum* L
	B10				*Large flowered barrenwort*
	R10				*Benoîte du Canada White avens*
	V10				*Amélanchier Serviceberry*
P11	J11	Boumerdès	Intensive	2017	*Siberian pea shrub*
	O11				*Pearly everlasting*
	Js11				*Dill*
	R11				*Spotted jewelweed*
	V11				*Purslane speedwell*
P12	J12	Tizi-Ouzou	Intensive	2017	*Bitter wintercress*
	O12				*Birch*
	Js12				*Alder*
	B12				*Black knapweed*
	R12				*Creeping bugleweed*
	N12				*Garlic mustard*
P13	A13	EL-Oued	Not dense	2017	*Zygophyllum album* L
					Genista saharae Coss & Dur
	W13				*Eucalyptus*
	J13				*Mathiolalivida* DC
					Phoenix dactylifera L
	O13				*Anacyclus valentinus* L
	Js13				*Launaeare sedifolia* O.K
	B13				*Anacyclus valentinus* L
					Launaeare sedifolia O.K
	V13				*Brassica oleracea var.viridis* L
	Vs13				*Brassica oleracea var.viridis* L
					Mathiola livida DC
	VIO13				*Malcomia aegyptiaca* spr
	R13				*Retama raetamEucalyptus*
					Genista saharae Coss & Dur
	N13				*Retama raetam*

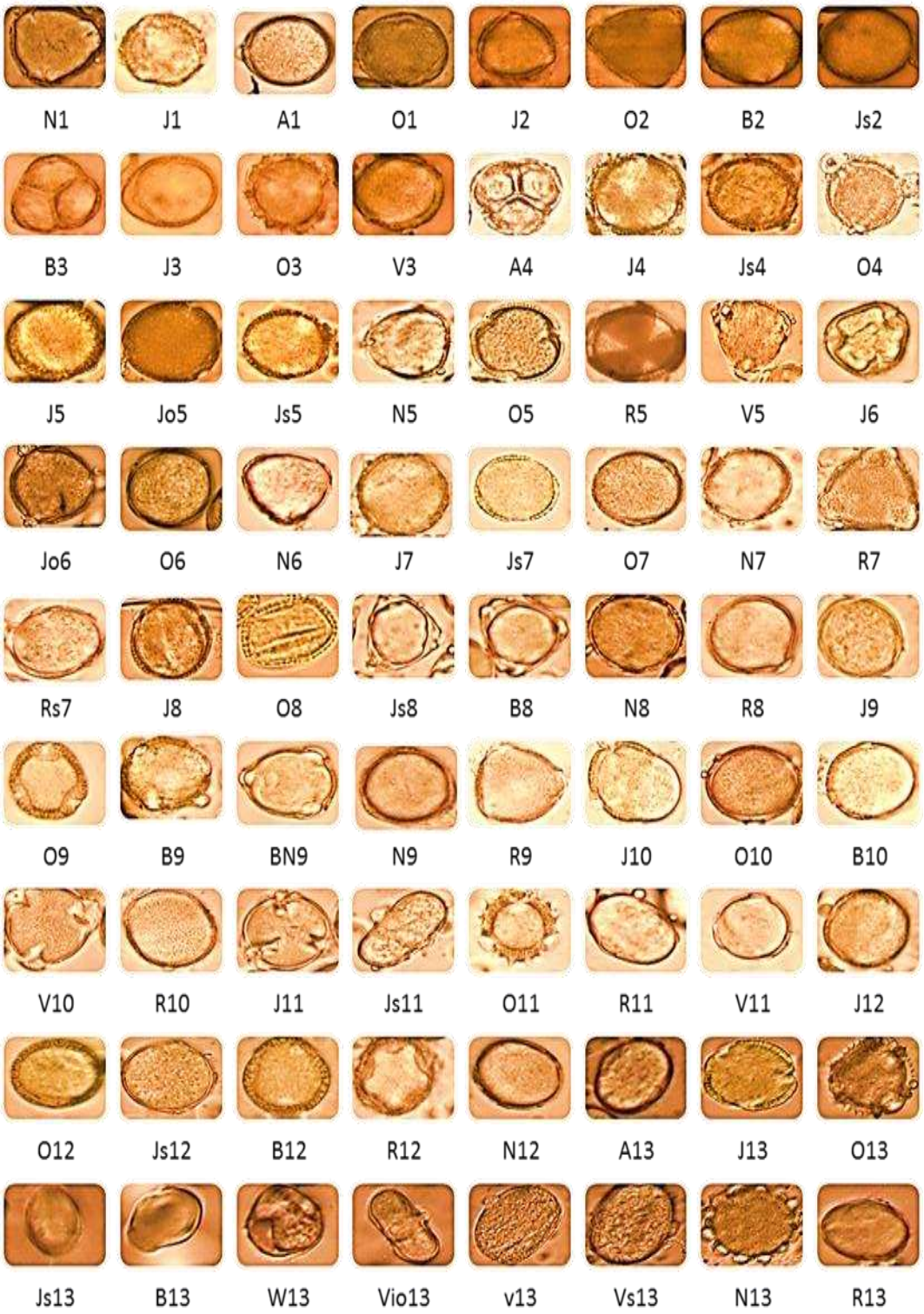

Fig. 2. Captured images of the optical microscope samples of pollen.

3. Results and discussion

Pollen authentication is a more specific problem in literature where there is limited data to model pollen types. One-class classification is an appropriate machine learning paradigm to deal with this problem.

While there are restricted discernment methods for recognising pollen types in macroscopic images the majority of the current methods for analysing bee pollen and its origin are applied to microscopic pollen grains images, The first works on recognising pollen grains by optical microscopes were provided by *France et al. (2000)* and *Boucher et al. (2002)* where some discriminative features of various pollen taxa were detected and classified.

Pollen authentication is a more complex issue in literature where there is insufficient data to model pollen forms. One-class classification is an effective machine learning model to deal with this problem.

Cluster analysis easily classifies data into groups which helps to show similarities and is commonly used for rapid differentiation and classification of data.

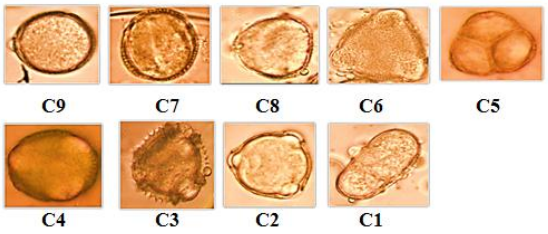

Fig. 3. Different images of pollen grains whose non-homogeneous background features

The model is taught using an open database of samples images containing grand a lot species. Experiments with a plant dataset show that the proposed model is significantly better than other classification methods. High classification accuracy makes the model very useful for supporting the plant recognition system for working in real conditions.

The particular attention has attributed to the understanding of the mechanisms underlying Microscope images classification of bee pollen. The influence of bee species, color of bee pollen, plant origin, geographic location, and season of collection it is directly related to the quality of the samples.

4. Conclusions

This paper has introduced the problem of automatic classification for bee pollen samples of different floral origin from various Algerian counties, where we adopted a standard methodology multi-focal image classification and authentication.

We showed the results of applying the image processing algorithms to obtain the on any kind of multi-focal images properties of the pollen. Then, we tested the different one-class classification models based on HCA. The use of the presented standard methodology drastically reduce the time and effort spent by experts to several seconds and can be used as an standard method for macroscopically rejecting unknown pollen loads. Future work can be devoted to apply a more interpretable multi-classification system.

In order to determined similarities and differences between bee pollen samples based on their Microscope images profile were established by the application of multivariate discriminate analysis. This method was proven to be a useful tool to study the relationships between bee pollen according to the Geographical area and to determine the importance of the Geographical area and plant origin on the bee pollen classification.

Acknowledgments

The authors wish to thank Mr. Ali Tliba from Laboratory Valorization and Technology of Saharan Resources (VTRS), University of El Oued, staff.

Conflict of Interest

The authors declare that they have no conflict of interest

References

[1] Pasupuleti, V. R., Sammugam, L., Ramesh, N., and Gan, S. H. Honey, propolis, and royal jelly: a comprehensive review of their biological actions and health benefits. Oxidative Medicine and Cellular Longevity (2017). https://doi.org/10.1155/2017/1259510

[2] Stillman, E., and Flenley, J. R. The needs and prospects for automation in palynology. J Quaternary Science Reviews 15 (1996) 1-5.

[3] Mitsumoto, K., Yabusaki, K., and Aoyagi, H. Classification of pollen species using autofluorescence image analysis. J Journal of bioscience bioengineering 107 (2009) 90-94.

[4] Flenley, J. The problem of pollen recognition. J Problems in Picture Interpretation (1968) 141-145.

[5] Flenley, J. Some prospects for palynology in the South-West Pacific Region, J Massey University Faculty of Social Sciences Occasional Papers (1990).

[6] Del Pozo-Baños, M., Ticay-Rivas, J. R., Cabrera-Falcón, J., Arroyo, J., Travieso-González, C. M., Sánchez-Chavez, L., Pérez, S. T., Alonso, J. B., and Ramírez-Bogantes, M. Image processing for pollen classification. In Biodiversity Enrichment in a Diverse World (2012). http://dx.doi.org/10.5772/47527

[7] Aronne, G., Cavuoto, D., and Eduardo, P. Classification and counting of fluorescent pollen using an image analysis system. J Biotechnic Histochemistry 76 (2001) 35-40.

[8] Bağcıoğlu, M., Zimmermann, B., and Kohler, A. A multiscale vibrational spectroscopic approach for identification and biochemical characterization of pollen. J PLoS One (2015) 10.

[9] Dell'Anna, R., Lazzeri, P., Frisanco, M., Monti, F., Campeggi, F. M., Gottardini, E., and Bersani, M. Pollen discrimination and classification by Fourier transform infrared (FT-IR) microspectroscopy and machine learning. J Analytical bioanalytical chemistry 394 (2009) 1443-1452.

[10] Ivleva, N., Niessner, R., and Panne, U. Characterization and discrimination of pollen by Raman microscopy. J Analytical bioanalytical chemistry 381 (2005) 261-267.

[11] Pinnick, R. G., Hill, S. C., Nachman, P., Pendleton, J. D., Fernandez, G. L., Mayo, M. W., and Bruno, J. G. Fluorescence particle counter for detecting airborne bacteria and other biological particles. J Aerosol Science Technology 23 (1995) 653-664.

[12] Fonseca, A. E., Westgate, M. E., and Doyle, R. T. Application of fluorescence microscopy and image analysis for quantifying dynamics of maize pollen shed. J Crop Science 42 (2002) 2201-2206.

[13] France, I., Duller, A., Duller, G., and Lamb, H. A new approach to automated pollen analysis. J Quaternary Science Reviews 19 (2000) 537-546.

[14] Cruz-Roa, A., Gilmore, H., Basavanhally, A., Feldman, M., Ganesan, S., and Shih, N. Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. Sci Rep. 2017; 7: 46450, Epub 2017/04/19. https://doi.org/10.1038/srep46450 PMID: 28418027.

[15] Mori, K., Sakuma, I., Sato, Y., Barillot, C., and Navab, N. Medical Image Computing and Computer-Assisted Intervention--MICCAI 2013: 16th International Conference, Nagoya, Japan, September 22-26, 2013, Proceedings, Part III, (2000) 8151, Springer.

[16] Janowczyk, A., Chandran, S., and Madabhushi, A. Quantifying local heterogeneity via morphologic scale: Distinguishing tumoral from stromal regions. Journal of pathology informatics (2013) 4.

[17] Irshad, H., Jalali, S., Roux, L., Racoceanu, D., Hwee, L. J., Le Naour, G., and Capron, F. Automated mitosis detection using texture. SIFT features and HMAX biologically inspired approach, Journal of pathology informatics (2013) 4.

[18] Basavanhally, A., Yu, E., Xu, J., Ganesan, S., Feldman, M., Tomaszewski, J., and Madabhushi, A. Incorporating domain knowledge for tubule detection in breast histopathology using OAcA (tm) Callaghan neighborhoods. In Proceedings of SPIE- The International Society for Optical Engineering, SPIE, P. O. BOX 10 Bellingham WA 98227-0010 USA (2011).

[19] Monaco, J., Hipp, J., Lucas, D., Smith, S., Balis, U., and Madabhushi, A. Image segmentation with implicit color standardization using spatially constrained expectation maximization: Detection of nuclei.

In International Conference on Medical Image Computing and Computer-Assisted Intervention, (2012) 365-372, Springer.

[20] Basavanhally, A., Feldman, M., Shih, N., Mies, C., Tomaszewski, J., Ganesan, S., and Madabhushi, A. Multi-field-of-view strategy for image-based outcome prediction of multi-parametric estrogen receptor-positive breast cancer histopathology: Comparison to Oncotype DX.Journal of pathology informatics (2011) 2.

[21] Beck, A., Sangoi, A., Leung, S., Marinelli, R., Nielsen, T., van de Vijver, M., West, R., van de Rijn, M., and Koller, D. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. Sci Transl Med 3 (2011) 108-113.

[22] Mohanty, S., and Hughes, D. Salathé Marcel. Using deep learning for image-based plant disease detection. Frontiers in Plant Science 7 (2016) 1419.

[23] Bengio, Y., and LeCun, Y. Scaling learning algorithms towards AI, Large-scale kernel machines (2007) 34, 1-41.

[24] Godec, P., Pančur, M., Ilenič, N., Čopar, A., Stražar, M., Erjavec, A., Pretnar, A., Demšar, J., Starič, A., and Toplak, M. Democratized image analytics by visual programming through integration of deep models and small-scale machine learning, Nature communications 10 (2019) 1-7.

[25] Kraus, O. Z., Grys, B. T., Ba, J., Chong, Y., Frey, B. J., Boone, C., and Andrews, B. J. Automated analysis of high content microscopy data with deep learning. Molecular systems biology 13 (2017) 924.

[26] Dell'Anna, R., Lazzeri, P., Frisanco, M., Monti, F., Campeggi, F. M., Gottardini, E., and Bersani, M. Pollen discrimination and classification by Fourier transform infrared (FT-IR) microspectroscopy and machine learning. Analytical and bioanalytical chemistry 394 (2009) 1443-1452.

[27] Kaškonienė, V., Kaškonas, P., and Maruška, A. Volatile compounds composition and antioxidant activity of bee pollen collected in Lithuania. Chemical Papers 69 (2015) 291-299.

Recommended Citation

A. Rebiai, B. Ben Seghir, H. Hemmami, S. Zeghoud , T. Siham, I. Kouadri, H. Terea, F. Brahmia. (2021) Clustering and discernment of Algerian bee pollen using an image analysis system. *Algerian Journal of Chemical Engineering*, *02 (2021) 41–48. http://dx.doi.org/ 10.5281/zenodo.5105756*