New frontiers of mineralogical and structural data analysis in the era of machine learning: tools for modern petrography

Alberto D'Agostino

Dipartimento di Scienze Biologiche, Geologiche ed Ambientali, Università degli Studi di Catania, Corso Italia 57, Catania, 95129, Italia

DOI:10.19276/plinius.2023.01.005

INTRODUCTION

In modern petrography an increasingly significant role is played by automated numerical computations, which determines a higher demand of quantitative analysis. While petrographers' experience still plays a huge role in the recognition of specific features in rocks, it may also lead to misinterpretations driven by subjectiveness. Hence, numerous statistical analyses can be applied on rocks to objectively extract quantitative information through advanced software solutions. However, while machines outperform humans in long and recursive tasks, computer-driven analyses can be misleading and may introduce other kinds of biases if not properly overseen by an expert operator. Therefore, petrologists' experience is still fundamental to validate machine-generated data and identify petrogenetic processes.

Machine learning (ML) algorithms have been widely experimented to support petrography, by processing several types of petrographic data, such as bulk-rock chemistry (e.g., Ren et al. 2019; Itano et al., 2020) and optical thin section images (e.g., Su et al., 2020; Visalli et al. 2021). Optical scans of rocks thin sections, however, can sometimes be a complex type of input to process, and although efforts have been made towards the realization of databases of optical thin section images (e.g., Quinn et al., 2011), an open-access archive of standardized and labeled data is missing. X-ray elemental maps are yet another type of input data that significantly benefits from the application of ML algorithms. Unlike punctual chemical analyses, the information is not scattered and prevents possible biases introduced by the choice of points locations. Therefore, powerful tools have been developed to quantify X-ray maps (e.g., Lanari et al., 2014; Arganda-Carreras et al., 2017; Ortolano et al., 2018). Generally, the acquisition of X-ray elemental maps and BSE maps is an efficient and relatively cheap process; however, an online structured database of labelled X-ray maps or BSE maps is again missing. In this scenario, X-Min Learn is here presented. This stand-alone software allows an interactive and friendly application of ML techniques for automatically identifying minerals from EDS and WDS X-ray maps and any other type of multi-spectral image data, including, for example, BSE maps. It also allows the development of custom ML models for specific research needs in a friendly way, meeting the needs of users not experienced in computer coding.

Nevertheless, usually fabric plays a central role in the final properties of the lithotypes as well. In this view, ArcStereoNet, which has already been published in Ortolano et al. (2021), is also presented. This ArcGIS® toolbox, useful for the statistical analysis and projection of structural data, allows the comparison of oriented data from the outcrop scale to the thin section scale within the same ArcGIS® project. ArcStereoNet allows the application of statistical analysis on stereographic projections and rose diagrams, while also taking full advantage of default GIS mapping processes, expanding the potential of other pioneering tools (e.g., Kociánová & Melichar, 2016).

X-MIN LEARN

The current free software for the classification of X-ray maps (e.g., Lanari et al., 2014; Arganda-Carreras et al., 2017; Ortolano et al., 2018) implement unsupervised and supervised classifiers. The latter are usually trained on specific user-drawn regions of interest (ROIs) and then employed to classify the full sample. While this approach (referred to as "lazy learning") often yields very accurate results, it may also introduce dangerous user-driven biases (Nickerson, 1998; Mehrabi et al., 2021). Lazy learners just memorize the information and then use it to classify new data by comparison (Hendrickx & Van Den Bosch, 2005). On the other hand, "eager learners" process the training data to extract a function that describes the relationships occurring between the features of the training data, hence producing a generalized enough model that can foresee pixel variance within the same mineral classes.

X-Min Learn (XML) implements both unsupervised and lazy supervised classifiers as well. In addition to that, however, it also includes a collection of interactive

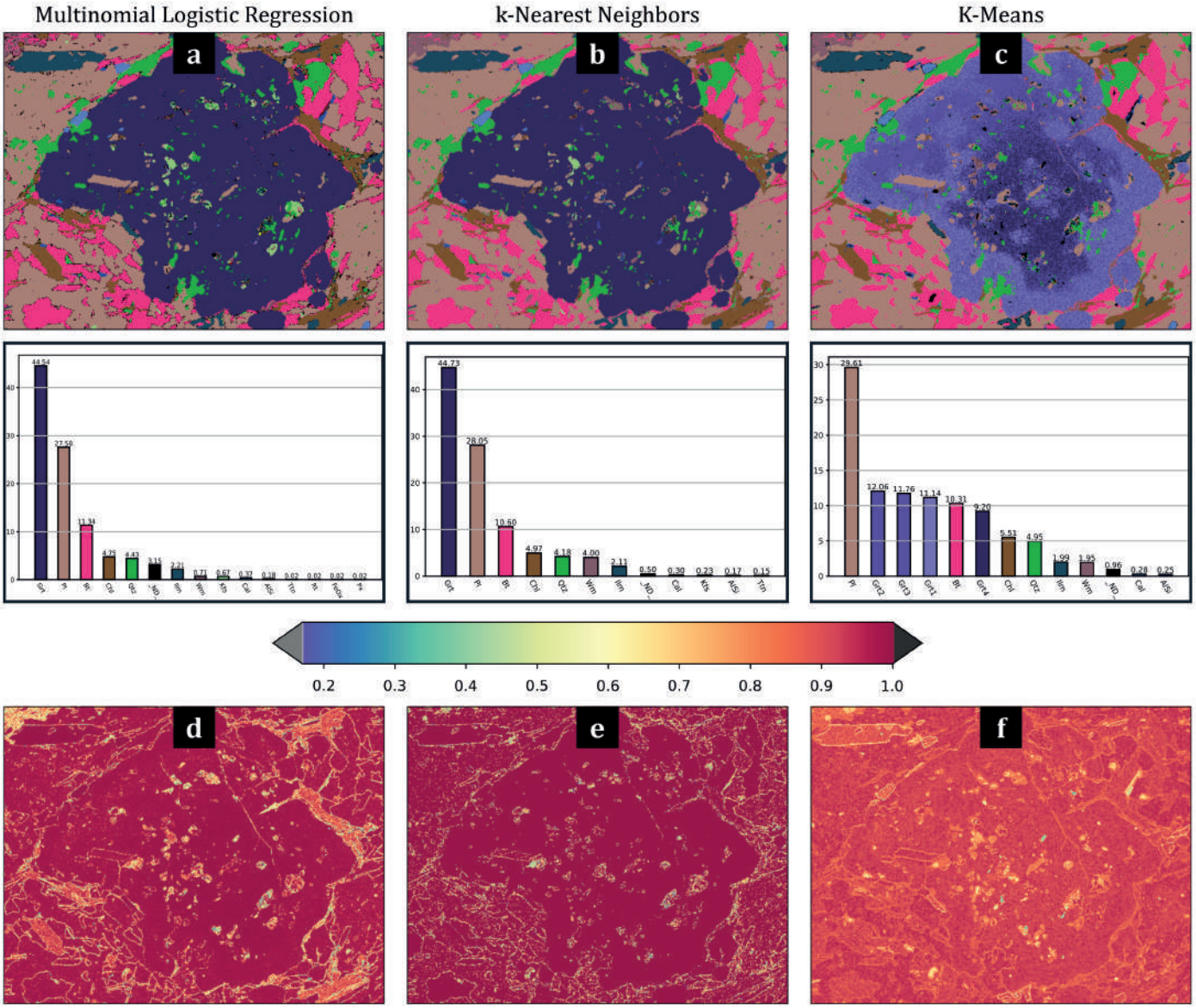

Figure 1 Sample classification using the three classifiers of X-Min Learn: **(a)** MLR, **(b)** k-NN, **(c)** K-Means with **(d, e, f)** their linked probability maps.

tools for a step-by-step development of custom eager ML models within its "developer's toolkit", simplifying the compilation of ground truth datasets and including statistics and graphics useful for the inspection of learning processes. After a custom model is stored, it can be employed to classify new samples automatically, not requiring users to trace additional training ROIs.

Classifiers

X-Min Learn provides three ML algorithms to identify minerals: K-Means (MacQueen, 1967), an unsupervised classifier, K-Nearest Neighbours (k-NN – Cover & Hart, 1967) a lazy supervised classifier and Multinomial Logistic Regression (MLR – Bridle, 1990) a user-trainable eager supervised classifier. They all generate a mineral map and an associated probability map, this latter displaying the classification confidence score of each pixel (Fig. 1d-f). Probability maps are useful for highlighting pixels located at the boundary of two different minerals or near fractures, displaying a mixed chemical composition (i.e., mixed pixels). This score can be used as a rejection factor to exclude low confidence pixels, providing a stronger user control.

In order to compare the classifiers' performance, they are here tested on SEM-EDS X-ray maps collected from a thin section of a metamorphic rock. The classification results are displayed in Fig. 1. The time required for the computation was comparable for each classifier (i.e., under 10 seconds). Both k-NN and K-Means, however, required additional time to either draw training ROIs or set the appropriate number of clusters and label them. The *MLR* model was previously trained using other metamorphic rocks samples collected from different outcrops and lithotypes, and therefore did not require additional training time, achieving a fully automatic mineral classification (Fig. 1a). It assigned some pixels to mineral classes that are not truly occurring in the analyzed sample (i.e., FeOx, Px, and Rt, namely iron oxide, pyroxene, and rutile). These are noisy pixels, that the model assumes as true mineral classes. Although some of them may just be excluded with a confidence threshold, this issue can be fixed with post-processing operations. The overall re-

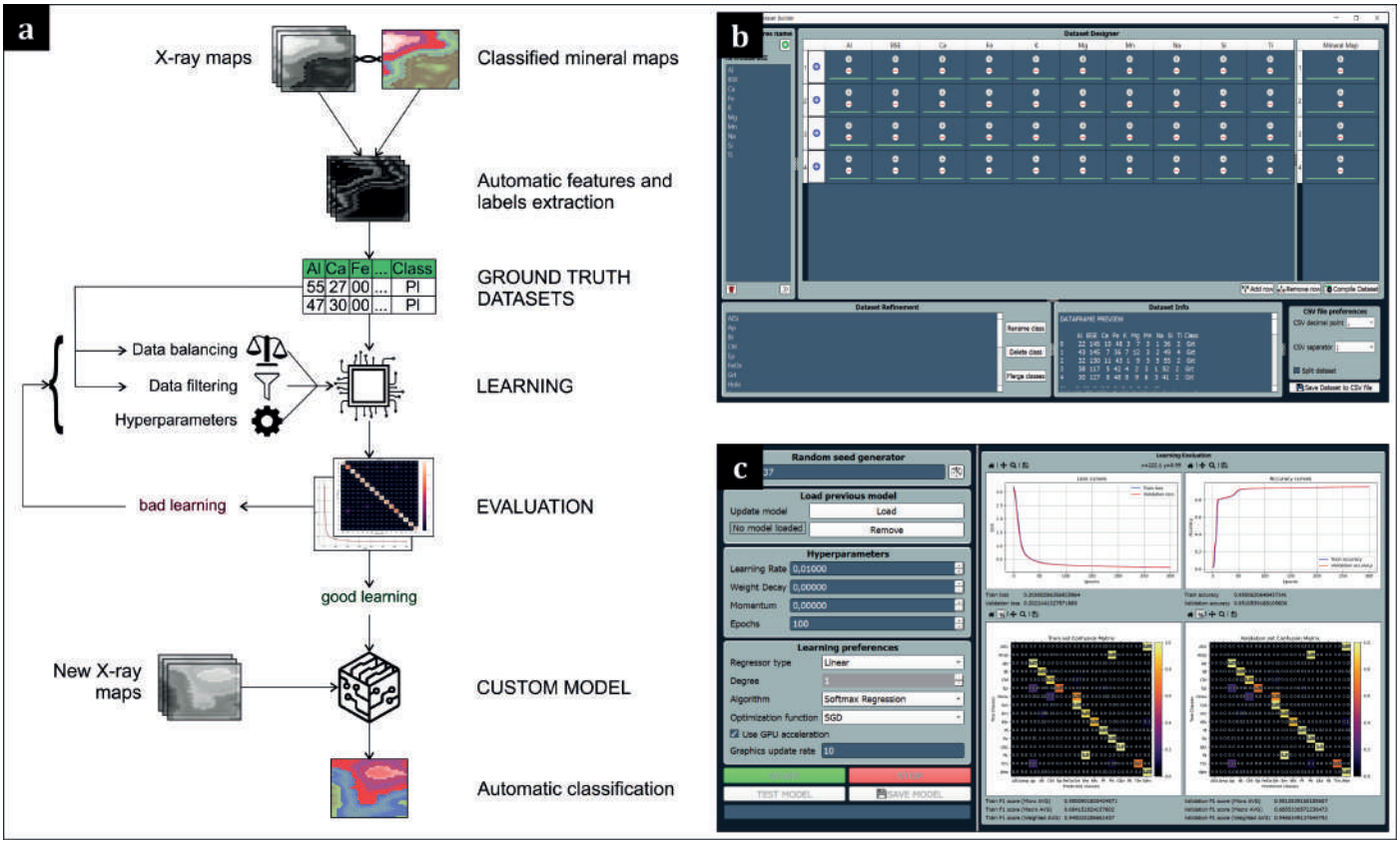

Figure 2 ***a)*** *Workflow of the developer's toolkit, that allows users to build custom ML models with interactive tools, such as the Dataset Builder* ***(b)*** *and the Model Learner* ***(c)***

sult is in accordance with that of the other classifiers. The *k-NN* classifier was steered by the operator, who traced several training ROIs on the occurring mineral phases, and, therefore, the mineral map displays the expected classes (Fig. 1b). While this behaviour is appreciated because provides a stronger user control over the classification, it can also be a dangerous source of sampling and confirmation biases (Nickerson, 1998; Mehrabi et al., 2021). The main drawback is that tracing the training ROIs is time consuming and is required for each analysis. The *K-Means* differs from the previous classifiers as it employs an unsupervised learning strategy to cluster the data into a K number of classes defined by the user. Although being the most unbiased of the available classifiers, the main drawback of K-Means is that its output classes are not labelled with mineral names, thus the result must be interpreted. Furthermore, it can sometimes struggle with uneven sized clusters, making it not the best choice when analyzing rocks with very imbalanced mineral amounts. Indeed, since garnet displays a high number of pixels (majority class), K-Means separates its zonation patterns in different clusters (i.e., Grt1, Grt2, Grt3, Grt4) at the expenses of minority classes such as titanite (Ttn) and K-feldspar (Kfs) that are not detected (Fig. 1c).

A sub-class analysis can also be achieved with XML (e.g., identification of mineral zonation patterns), by recursively applying the algorithms to already classified mineral maps. The data is automatically masked in order to focus the sub-phase classification only on the selected class. This process can be repeated an unlimited number of times, also using a sequence of different classifiers.

Phase Refiner

The Phase Refiner allows the removal of noisy pixels (i.e., small classification errors) produced by a classifier, which can negatively affect further analysis on mineral maps (e.g., sub-phase identifications). It provides image processing algorithms in two modalities: basic mode and advanced mode. The basic mode allows the application of a maximum frequency filter to smoothen the entire mineral map, while the advanced mode provides binary morphological operations to refine each mineral class individually. In both modes, users can set the sliding window size and shape to obtain different refinement results. In advanced mode ROIs can be traced to restrict the refinement only to specific areas of the sample and several algorithms can be consecutively applied on different phases, guaranteeing a complete user control over the final result. Six different algorithms are available: Erosion, Dilation, Opening, Closing, Erosion + Reconstruction and Fill Holes. Refined and validated mineral maps could, in turn, be employed as ground truth data to train new ML models.

Developer's toolkit

X-Min Learn is the first mineral-oriented software that allows users to train mineral classifiers by providing a collection of interactive tools for a step-by-step development of custom eager ML models, using fully classi-

fied and validated samples as training data (Fig. 2). A lazy classifier trained with user-drawn ROIs is prone to be affected by confirmation bias (Nickerson, 1998), as operators are led to modify the ROIs multiple times until the model generates a result that aligns with their initial hypothesis. Fully classified samples, on the other hand, provide training data with a larger intraclass variance, since all their pixels are processed rather than just those contained within arbitrary ROIs. The evaluation of XML's custom classifiers, moreover, is not based on the result of a specific classification (i.e., another possible source of confirmation bias), but rather on graphics and statistics during the learning process. Finally, another advantage of providing the instruments to build custom ML classifiers is that users can freely choose how to train their models, potentially developing highly specialized classifiers. The developer's toolkit gathers tools for ground truth datasets management (Dataset Builder – Fig. 2b) and for the actual training of custom machine learning models (Model Learner – Fig. 2c).

Dataset Builder

The Dataset Builder automatizes the compilation of a human readable, machine friendly, standardized dataset from validated examples of already classified data (i.e., ground truth data). Each instance (i.e., row) of such dataset is populated with the features (the numerical values stored in the input maps' pixels) and the corresponding labels (mineral class stored in the classified mineral maps). Users can also operate several dataset refinement operations (e.g., renaming, deleting and merging mineral classes). The dataset is saved as a CSV file; this file format was chosen for both its wide compatibility and its popularity among most users.

Model Learner

The Model Learner allows users to stepwise build new custom ML models after having compiled a ground truth dataset with the Dataset Builder. Custom models can be tailored to solve various tasks, from recognizing the most common mineral classes in different rock types, to identifying intra-class variabilities of individual mineral species. Users may also develop models for classifying specific lithotypes or even samples. Moreover, the Model Learner allows users to update their models with new training data, refining them over time.

The adopted ML algorithm is the Multinomial Logistic Regression (Eq. 1), which is paired with the Cross-Entropy loss function for the quantitative evaluation of the model's errors (Eq. 2) and the gradient descent algorithm for the model's optimization (Eq. 3). Here ϑ are the model's internal parameters, x the input features, K the number of mineral classes, κ the class id (with $\kappa \in [0, K)$), y the true mineral class, N the train set size, ε the learning iteration, η the learning rate, λ the weight decay and μ the momentum.

$$h_{\vartheta}^{(\kappa)}(x) = \frac{e^{\vartheta^{(\kappa)T}x}}{\sum_{i=0}^{K-1} e^{\vartheta^{(i)T}x}} \quad (1)$$

$$\mathcal{L}(\vartheta) = -\frac{1}{N}\sum_{i}^{N}\sum_{\kappa=0}^{K-1} 1\{y_i = \kappa\}\log(h_{\vartheta}^{(\kappa)}(x_i)) \quad (2)$$

$$\vartheta_j^{\varepsilon} = \vartheta_j^{\varepsilon-1} - \eta \cdot \left(\frac{\partial \mathcal{L}_{\vartheta_j}}{\partial \vartheta_j} + \mu\vartheta_j^{\varepsilon-1}\right) - \lambda\eta\vartheta_j^{\varepsilon-1} \quad (3)$$

The learning operations consist of an empirical fine tuning of the model's hyperparameters (i.e., ε, η, λ, μ) performed by users with the aim of optimizing the model's performance. Such operations are reiterated several times with different settings, until a satisfactory performance is achieved. The tool provides interactive instruments for the evaluation of the performance, such as learning scores, loss curves and confusion matrices (Fig. 2c). The automatic pseudo-randomizations that are operated during the learning session are controlled by a random seed, that can be set for reproducibility purposes. Hence, two learning sessions with same input data and hyperparameters settings will always produce the identical result if the same seed is adopted.

Various pre-processing operations can optionally be performed on the train set, such as polynomial features mapping for the identification of non-linear patterns in the data (Theodoridis & Koutroumbas, 2006) and balancing operations. These last consist of over- and under-sampling algorithms as implemented in the open-source Python library imbalanced-learn (Lemaître et al., 2017 and references therein). They are particularly useful to address the problem of imbalanced datasets (Kaur et al., 2019), which occurs when training data contains different amounts of examples per class, that negatively impact the accuracy of supervised ML classifiers. This issue frequently affects mineralogical datasets, as the amount of minerals extremely differs in rocks depending on the mineral species.

Case study

An application of XML for the extraction of quantitative petrographic data from a natural rock is here described. The collected sample (GC29) consists of a late-Variscan amphibolite from the Aspromonte Unit, NE Sicily (Cirrincione et al., 2015 and references therein). From a thin section of GC29, three micro-domains were selected, depicting three relict eo-Variscan garnets surrounded by symplectitic micro-structures. WDS X-ray elemental maps and BSE maps were collected and then analyzed with XML.

The first micro-domain was classified with the k-NN classifier (Fig. 3a). The training ROIs were validated with WDS punctual chemical data and optical microscope analysis. A custom ML model was trained with the de-

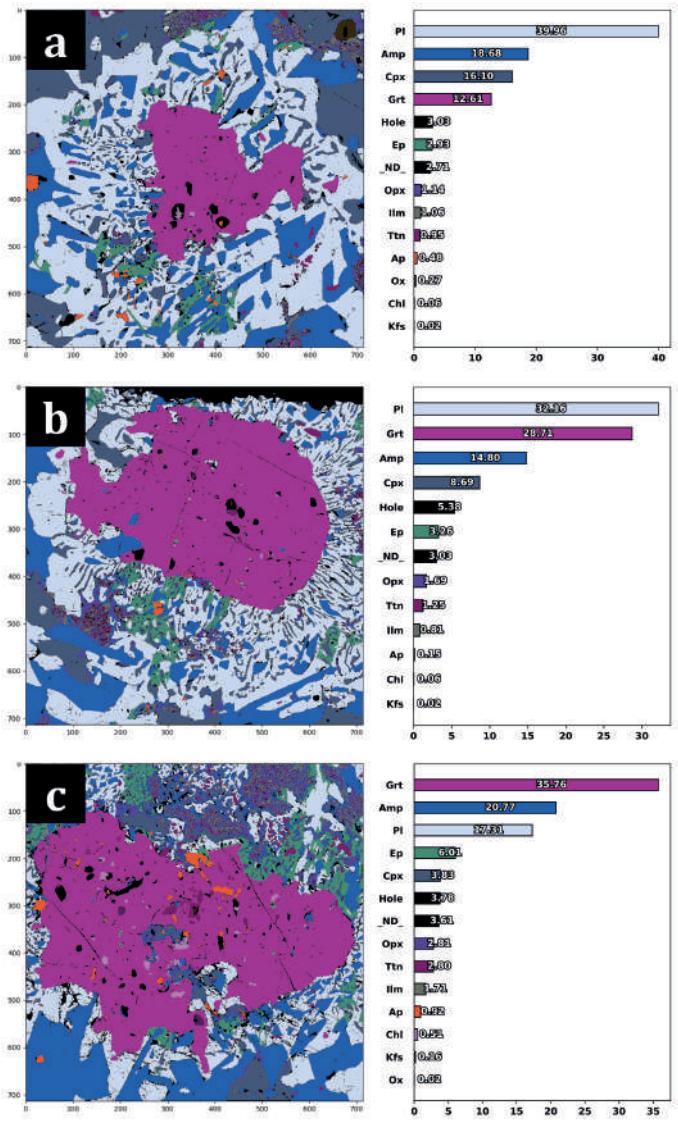

Figure 3 *Classification of three micro-domains from GC29 sample. The first micro-domain* ***(a)*** *was classified with k-NN and then processed with the developer's toolkit to generate a custom ML model, which allowed a fully automatic classification of the other two* ***(b, c).***

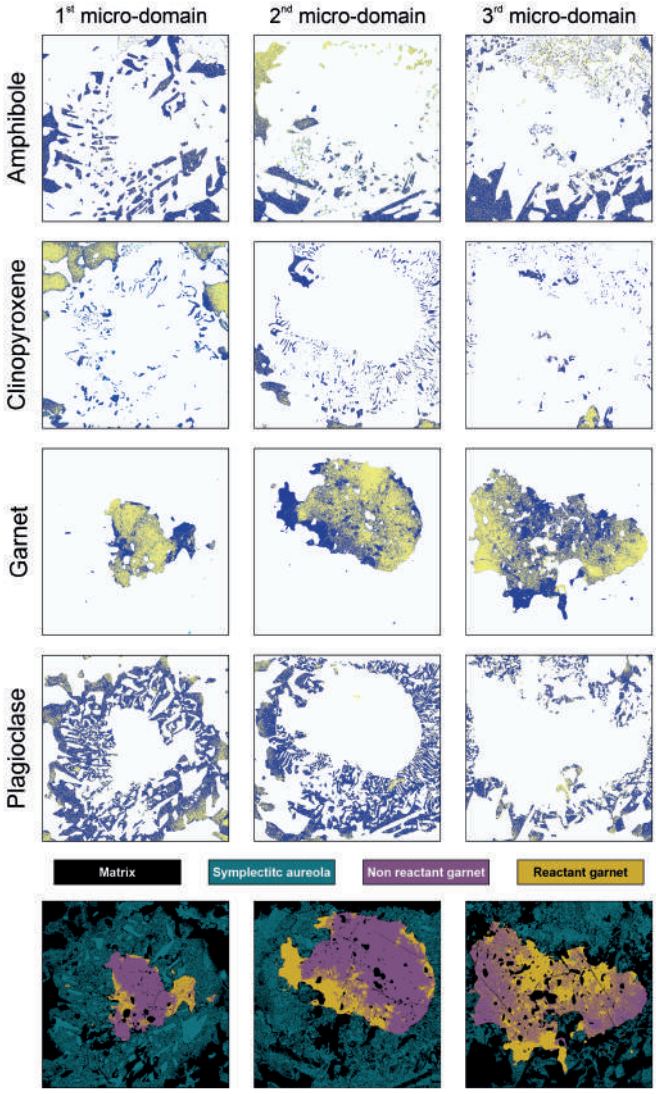

Figure 4 *Sub-phase identification achieved with K-Means for amphibole, clinopyroxene, garnet and plagioclase on each micro-domain. The estimated reaction rate of each garnet porphyroblast is, respectively, 0.78, 0.55 and 0.38*

veloper's toolkit using the first micro-domain as ground truth data. Thus, the model was tailored for the classification of the sample GC29, achieving an accuracy of 98.8%. Consequently, the other two micro-domains were automatically classified with such model (Fig. 3b,c). Finally, K-Means classifier was applied on amphibole, clinopyroxene, garnet and plagioclase classes to detect the occurrence of sub-phases in each micro-domain. This allowed the identification of mineral zonation patterns, that were interpreted as the effect of the symplectitic reaction, and the estimation of the reaction rate of garnet porphyroblasts (Fig. 4). This preliminary procedure can be helpful to infer the effective reactant volumes and, in turn, to identify the effective bulk chemistry (Zuluaga et al., 2005) and obtain more reliable pseudo-sections and phase diagrams.

ARCSTEREONET

ArcStereoNet (ASN) is a Python-toolbox that adds geological-oriented tools to ArcGIS®, allowing the projection and the statistical analysis of georeferenced oriented data, preserving its coordinates in real space. If paired with ad hoc tools like Micro-Fabric Analyzer (MFA – Visalli et al., 2021), ASN can also process micro-structural data collected from thin sections of rocks. Available statistical functions include contouring, cluster and girdle analysis and mean vectors extraction. A new clustering algorithm (i.e., MEAD – Mean Extractor from Azimuthal Data) is also included in the toolbox. The algorithms can be compared simultaneously, allowing a more reliable interpretation of the data distribution. ASN requires the following information: azimuth angle (dip-direction, strike, trend), dip angle, data format type and data label, which can automatically retrieve by accessing the attribute tables of ESRI® shapefiles. The dataset management operations provided by ArcGIS® can be leveraged to manage and standardize the data.

Tools and algorithms

Three tools are included in ASN: Stereoplots, Rose Diagrams and Graph To Hyperlink, respectively useful to generate stereographic projections and rose diagrams

and to connect them with the geographic position of the data. *Stereoplots* tool generates lower hemisphere equal area or equal angle azimuthal projections, displaying cyclographic traces, and/or poles for planar data, and/or points for linear elements. It allows the overlaying of contours and the comparison of multiple clustering algorithms, as well as the extraction of mean planes or vectors for each cluster. *Rose Diagrams* tool allows the identification of clusters and mean vectors as well. Rose diagrams can also be weighted based on a user-selected parameter included in the input shapefile. *Graph To Hyperlink* tool is useful to link the plots generated by the previous tools to the spatial position of the data in the map. Positions corresponds to the mean latitude and longitude coordinates of plotted data. Users can click on each position to show a popup window displaying the hyperlinked plot.

ASN provides unsupervised algorithms for the recognition of recurring patterns in the data, allowing users to group the data based on orientation similarities (i.e., clustering) and then to extract the average value of each identified group (i.e., mean vector extraction). There are four available algorithms in the Stereoplots tool (i.e., MEAD, MEAD + Fisher, K-Means, Bingham) and one in the Rose Diagrams tool (MEAD). *MEAD* was created to provide a more user-controlled clustering algorithm. Azimuth and dip tolerance parameters can be fine-tuned to steer the algorithm towards the preferred clustering behaviour, also allowing MEAD to discard spurious (i.e., noisy) data. A slightly modified version of MEAD is implemented in the Rose Diagrams tool to address some specific requirements of such types of projections. *MEAD + Fisher* is an alternative version of MEAD, where the mean vector extraction process is carried out by the Fisher function (Fisher et al., 1993). It yields three statistics: R value (mean vector magnitude), confidence radius (mean vector confidence) and K value (dispersion factor). *K-Means* (MacQueen, 1967) strategy differs from MEAD mainly for the selection of the starting cluster centroid and for the absence of user-controlled tolerance parameters. *Bingham* (Bingham, 1974) does not perform a clustering process, but rather finds the best fit plane of a girdle-like distribution pattern. A full description and comparison of the algorithms is provided in Ortolano et al. (2021).

Case study

ArcStereoNet was tested by extracting, analyzing and comparing meso-structural (outcrop scale) and micro-structural (thin section scale) oriented data collected at Palmi (SW Calabria) within the Palmi Shear Zone (Fazio et al., 2017; Ortolano et al., 2020; Ortolano et al., 2021), integrated with meso-structural data virtually collected after an aerial photogrammetry survey. The Stereoplots tool was employed to analyze both the manually and the virtually collected meso-structural data (Fig. 5),

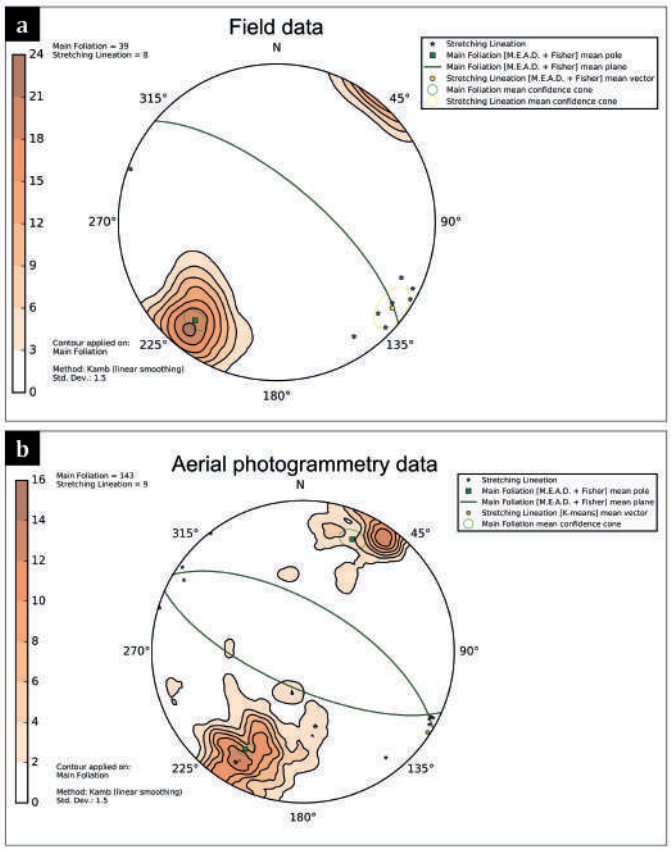

Figure 5 *Modified after Ortolano* et al. *(2021). Meso-structural data (main foliations and stretching lineations) collected* ***(a)*** *manually and* ***(b)*** *after aerial photogrammetry survey*

while the Rose Diagrams tool allowed the analysis of micro-structural data (Fig. 6), which was extracted with the Micro Fabric Analyzer tool (Visalli et al., 2021) from thin sections of samples collected at the same outcrops. The analysis was performed on the minerals belonging to porphyroclastic domains, highlighting their preferred orientations. Here, pre-kinematic clasts behave as rigid phases during sub-simple shearing plastic deformation events (see Ortolano et al., 2020 for further details).

A correlation between the orientation of micro-structural and meso-structural data can be observed; the mean main foliation planes (outcrop scale) and the porphyroclasts (thin section scale) are oriented WNW – ESE. The porphyroclasts, which are surrounded by quartz-rich weak layers, determining a relatively small rheology contrast, facilitate wing formation, producing greater resistance to the mylonitic flow and, in turn, clearer evidence of formation of sub-simple shear kinematic indicators. Aerial photogrammetry data is in accordance with field collected data, enlarging the amount of available meso-structural data, previously scarce because of the impervious terrain. This further confirms the versatility of ASN in processing data collected from different sources and at different scales.

CONCLUSIONS

In the era of digitalization and big data collection, petrography can benefit from the application of data science techniques such as machine learning algorithms. In

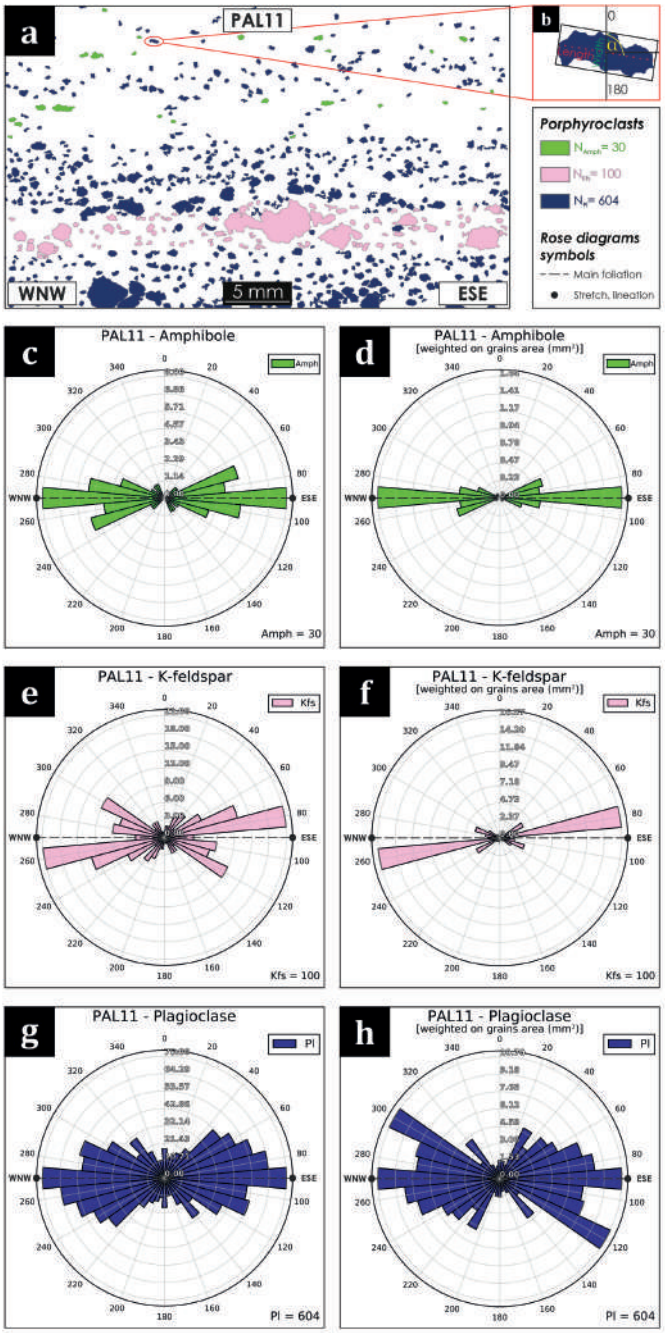

Figure 6 *From Ortolano* et al. *(2021). Micro-structural data extracted with MFA (Visalli* et al.*, 2021) after* ***(a)*** *thin sections digitalization (sample "PAL11"), using the* ***(b)*** *minimum bounding geometry of each single grain. The displayed rose diagrams of amphibole, K-feldspar and plagioclase porphyroclasts are both unweighted* ***(c, e, g)*** *and weighted* ***(d, f, h)*** *on grains' cumulative area (in mm²).*

this view, two new computer software, useful for the quantitative investigation of the mineralogy and the fabric of rock samples, were introduced: a) X-Min Learn, that provides customizable machine learning algorithms to identify rocks minerals from thin section multi-spectral data and b) ArcStereoNet, that allows the statistical analysis of structural oriented data within the ArcGIS® environment. By encouraging users towards an aware application of the provided algorithms, both software allow the derivation of reliable interpretations and constraints, representing an important contribution towards the increasingly pressing demand of achieving quantitative results in petrography. This is at the service of the most diverse facets of geosciences, from the resolution of petrological problems to micro-structural ones, passing through the field of geomaterial analysis.

REFERENCES

Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K.W., Schindelin, J., Cardona, A. & Sebastian Seung, H. (2017): Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinform.*, **33**, 2424-2426.

Bingham, C. (1974): An antipodally symmetric distribution on the sphere. *Ann. Stat.*, **2**, 1201-1225.

Bridle, J.S. (1990): Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. *Neurocomputing*, **68**, 227-236.

Cirrincione, R., Fazio, E., Fiannacca, P., Ortolano, G., Pezzino, A. & Punturo, R. (2015): The Calabria-Peloritani Orogen, a composite terrane in Central Mediterranean; its overall architecture and geodynamic significance for a pre-Alpine scenario around the Tethyan basin. *Period. Mineral.*, **84**, 701-749.

Cover, T.M. & Hart, P.E. (1967): Nearest neighbor pattern classification. *IEEE T. Inform. Theory*, **13**, 21–27.

Fazio, E., Ortolano, G. & Cirrincione, R. (2017): Eye-type folds at the Palmi shear zone (Calabria, Italy). *Int. J. Earth Sci.*, **106**, 2039-2040.

Fisher, N.I., Lewis, T. & Embleton, B.J. (1993): Statistical analysis of spherical data. *Cambridge university press*, 344.

Hendrickx, I. & Van Den Bosch, A. (2005): Hybrid algorithms with instance-based classification. *Lect. Notes Comput. Sc.*, **16**, 158-169.

Itano, K., Ueki, K., Iizuka, T. & Kuwatani, T. (2020): Geochemical discrimination of monazite source rock based on machine learning techniques and multinomial logistic regression analysis. *Geosciences*, **10**, 63.

Kaur, H., Pannu, H.S. & Malhi, A.K. (2019): A systematic review on imbalanced data challenges in machine learning: Applications and solutions. *ACM Comput. Surv.*, **52**, 1-36.

Kociánová, L. & Melichar, R. (2016): OATools: An ArcMap add-in for the orientation analysis of geological structures. *Comput. Geosci.*, **87**, 67-75.

Lanari, P., Vidal, O., De Andrade, V., Dubacq, B., Lewin, E., Grosch, E.G. & Schwartz, S. (2014): XMapTools: A MATLAB©-based program for electron microprobe X-ray image processing and geothermobarometry. *Comput. Geosci.*, **62**, 227-240.

Lemaître, G., Nogueira, F. & Aridas, C.K. (2017): Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *J. Mach. Learn. Res.*, **18**, 559-563.

MacQueen, J. (1967): Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical*

Statistics and Probability, **1**, 281-297.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K. & Galstyan, A. (2021): A survey on bias and fairness in machine learning. *ACM Comput. Surv.*, **54**, 1-35.

Nickerson, R.S. (1998): Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.*, **2**, 175-220.

Ortolano, G., Visalli, R., Godard, G. & Cirrincione, R. (2018): Quantitative X-ray Map Analyser (Q-XRMA): A new GIS-based statistical approach to Mineral Image Analysis. *Comput. Geosci.*, **115**, 56-65.

Ortolano, G., Fazio, E., Visalli, R., Alsop, I., Pagano, M. & Cirrincione, R. (2020): Quantitative microstructural analysis of mylonites formed during Alpine tectonics in the western Mediterranean realm. *J Struct. Geol.*, **131**, 103956.

Ortolano, G., D'Agostino, A., Pagano, M., Visalli, R., Zucali, M., Fazio, E., Alsop, I. & Cirrincione, R. (2021): ArcStereoNet: a new ArcGIS® toolbox for projection and analysis of meso-and micro-structural data. *ISPRS Int. J. Geo-Inf.*, **10**, 50.

Quinn, P., Rout, D., Stringer, L., Alexander, T., Armstrong, A. & Olmstead, S. (2011): Petrodatabase: an on-line database for thin section ceramic petrography. *J. Archaeol. Sci.*, **38**, 2491-2496.

Ren, Q., Li, M., Han, S., Zhang, Y., Zhang, Q. & Shi, J. (2019): Basalt tectonic discrimination using combined machine learning approach. *Minerals*, **9**, 376.

Su, C., Xu, S.J., Zhu, K.Y. & Zhang, X.C. (2020): Rock classification in petrographic thin section images based on concatenated convolutional neural networks. *Earth Sci. Inform.*, **13**, 1477-1484.

Theodoridis, S. & Koutroumbas, K. (2006): Pattern recognition. *Elsevier*, 856.

Visalli, R., Ortolano, G., Godard, G. & Cirrincione, R. (2021): Micro-Fabric Analyzer (MFA): A new semiautomated ArcGIS-based edge detector for quantitative microstructural analysis of rock thin-sections. ISPRS *Int. J. Geo-Inf.*, **10**, 51.

Zuluaga, C.A., Stowell, H.H. & Tinkham, D.K. (2005): The effect of zoned garnet on metapelite pseudosection topology and calculated metamorphic PT paths. *Am. Mineral.*, **90**, 1619-1628.