Astrobiology

Astrobiology Manuscript Central: http://mc.manuscriptcentral.com/astrobiology

Signatures of life detected in images of rocks using Neural Network analysis demonstrate new potential for searching for biosignatures on the surface of Mars

Journal:	*Astrobiology*
Manuscript ID	AST-2022-0034.R1
Manuscript Type:	Research Articles (Papers)
Date Submitted by the Author:	n/a
Complete List of Authors:	Corenblit, Dov; Universite Clermont Auvergne, Decaux, Olivier; MAD-Environnement Delmotte, Sébastien; MAD-Environnement Toumazet, Jean-Pierre; Université Clermont Auvergne Arrignon, Florent; MAD-Environnement André, Marie-Françoise; Université Clermont Auvergne Darrozes, José; Université Paul Sabatier, Earth Sciences Davies, Neil ; Cambridge University Julien, Frédéric; Université Paul Sabatier Otto, Thierry; CNRS, ECOLAB Ramillien, Guillaume; Géosciences Environnement Toulouse Roussel, Erwan; Universite Clermont Auvergne Steiger, Johannes; Universite Clermont Auvergne Viles, Heather ; University of Oxford
Keyword:	Biofilms, Biosignatures, Earth Mars, Life Detection
Manuscript Keywords (Search Terms):	Astrobiology, Biogeomorphology, Microbially induced sediment structures, Biosignatures, Neural Network, Mars

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Signatures of life detected in images of rocks using Neural Network analysis demonstrate new potential for searching for biosignatures on the surface of Mars

Dov Corenblit[1,2*], Olivier Decaux[3], Sébastien Delmotte[3], Jean-Pierre Toumazet[1], Florent Arrignon[3], Marie-Françoise André[1], José Darrozes[4], Neil S. Davies[5], Frédéric Julien[2], Thierry Otto[2], Guillaume Ramillien[4], Erwan Roussel[1], Johannes Steiger[1], Heather Viles[6]

[1]Université Clermont Auvergne, CNRS, GEOLAB – F-63000 Clermont-Ferrand, France.

[2]CNRS, Laboratoire écologie fonctionnelle et environnement, Université Paul Sabatier, CNRS, INPT, UPS, F-31062 Toulouse, France.

[3]MAD-Environnement, 23 rue de la République, 31560 Nailloux, France.

[4]Université Paul Sabatier, CNRS/IRD, GET – F-31062 Toulouse, France.

[5]Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, United Kingdom.

[6]School of Geography and the Environment, University of Oxford, South Parks Road, Oxford, OX1 3QY, United Kingdom.

*Corresponding author: dov.corenblit@uca.fr

Abstract

Microorganisms play a role in the construction or modulation of various types of landforms. They are especially notable for forming microbially induced sedimentary structures (MISS). Such microbial structures have been considered to be amongst the most likely biosignatures that might be encountered on the Martian surface. Twenty-nine algorithms have been tested with images taken during a laboratory experiment for testing their performance in discriminating mat cracks (MISS) from abiotic mud cracks. Among the algorithms, Neural Network types produced excellent predictions with similar precision of 0.99%. Following that step, a Convolutional Neural Network (CNN) approach has been tested to see if it can conclusively detect MISS in images of rocks and sediment surfaces taken at different natural sites where present and ancient (fossil) microbial mat cracks and abiotic desiccation cracks were observed. The CNN approach showed excellent prediction of biotic and abiotic structures from the images (global precision, sensitivity and specificity: respectively 0.99, 0.99 and 0.97). The key areas of interest of the machine matched well with human expertise for distinguishing biotic and abiotic forms (in their geomorphological meaning). The images indicated clear differences between the abiotic and biotic situations expressed at three embedded scales: texture (size, shape and arrangement of the grains constituting the surface of one form), form (outer shape of one form) and pattern of forms arrangement (arrangement of the forms over few square meters). The most discriminative components for biogenicity were the border of the mat cracks with their tortuous enlarged and blistered morphology more or less curved upwards, sometimes with thin laminations. In order to apply this innovative biogeomorphological approach to the images obtained by rovers on Mars, the main physical and biological sources of variation in abiotic and biotic outcomes must now be further considered.

Key words: Astrobiology; Biogeomorphology; Microbially induced sediment structures; Biosignatures; Neural Network; Mars.

1. Introduction

Advances in the understanding of the environmental context of the earliest life on Earth, ca. 3.8 Ga ago, and of the ability of modern extremophile microorganisms to cope with extreme conditions caused by salt, acidity, temperature, pressure or radiation, hint that the search for life beyond Earth could ultimately be successful (Landis, 2001; Cady and Noffke, 2009; Grotzinger *et al.*, 2014; Westall *et al.*, 2015; Ibarra and Corsetti, 2016; Cabrol, 2018; Damer and Deemer, 2019; Longo and Damer, 2020; Chacon-Baca *et al.*, 2021). The search for past or present signatures of life on telluric planets, their moons and on asteroids is primarily focussed on turning up direct evidence for fossilized microorganisms, biologically influenced minerals, or organic chemical or stable isotopic biomarkers at the surface or in the atmosphere (Ehlmann *et al.*, 2008; van Zuilen, 2008; Marshall *et al.*, 2017; Huang *et al.*, 2018; Webster *et al.*, 2013; Limaye *et al.*, 2018; McMahon *et al.*, 2018; Schwieterman *et al.*, 2018; Cockell and McMahon, 2019).

Several types of microorganisms (archaea, bacteria, and eukaryota domains, *e.g.*, protozoa, unicellular algae, and unicellular fungi) grow intimately with rock and sediment, and derive shelter, nutrients and water from them (Huang *et al.*, 2020). They are also known to leave many kinds of detectable traces on rocks or soft sediment during mineral precipitation and transformation, weathering, erosion and deposition processes, from micro to regional scales (Naylor *et al.*, 2002; Carter and Viles, 2005; Naylor, 2005; Viles, 2008, 2012; Noffke, 2010; Hays *et al.*, 2017). It has been suggested that the environments of Mars during the Noachian period (4.1-3.6 Ga), including primitive ocean, craters and playa lakes, volcanic aquifers, hot springs, and hydrothermal seafloors, could have been suited to the evolution of microorganisms. If this were the case then they must have affected surface and subsurface geomorphological characteristics of the planet and it is thus appropriate to search the Noachian sedimentary record of Mars for biosignatures (Naylor, 2005; Cady and Noffke, 2009; Schon *et al.*, 2012; Noffke *et al.*, 2013; Westall *et al.*, 2015; Corenblit *et al.*, 2019; Joseph *et al.*, 2020; Rizzo, 2020; Bosak *et al.*, 2021; Noffke, 2021).

The recognition of signatures of extraterrestrial life is only possible if analogous signatures can first be identified in Earth's sedimentary record (Cady, 2001; Cady *et al.*, 2004; McLoughlin *et al.*, 2007; Cady and Noffke, 2009; Noffke, 2000, 2009, 2010; Baucon *et al.*, 2017; Corenblit *et al.*, 2019). This approach is based on reasoning via abductive inference, with the assumption that the same or equivalent biogeomorphological processes result in categories of landforms displaying similar characteristics (Gilbert, 1886; Baker, 2008; Corenblit *et al.*, 2019).

Among the potential microbial biosignatures on Mars, microbially induced sedimentary structures (MISS) are promising candidates (Cady *et al.*, 2004; Schieber *et al.*, 2007; Noffke, 2010, 2021; Hickman-Lewis *et al.*, 2022). They are sedimentary substrates that are textured and patterned by the adhesive and cohesive properties of microbial mats and biofilms (Noffke *et al.*, 2001). They are presently widespread on Earth in several environments at the interface between water and land, including fluvial, marine, lacustrine and hypersaline settings (Stal, 2003; Thomas *et al.*, 2013; Cuadrado and Pan, 2018; Maisano *et al.*, 2019) and have a terrestrial fossil record that extends back for billions of years (Noffke, 2010; Carmona *et al.*, 2012; Davies *et al.*, 2017; Lepot, 2020; Davies and Shillito, 2021). The microbial mat communities responsible for generating MISS variably include bacteria, archaea, protozoans, algae, and fungi, and the structures formed are similarly diverse, ranging from millimetre to decimetre scales (Chacon-Baca *et al.*, 2021) and including an array of laminar structures, microbially induced wrinkles, mat chips, palimpsest ripples, roll-up forms, gas domes and mat cracks (Eriksson *et al.*, 2007; Porada and Bouougri, 2007; Noffke, 2010; Noffke *et al.*, 2013; Davies *et al.*, 2016).

Rovers and orbiters are presently acquiring large sets of images of the surface of Mars (Kwan, 2021), including imagery of bedding planes that can be considered true substrates (Davies and Shillito, 2021; Mangold *et al.*, 2021): fossilized remnants of the Martian lithic surface from the Noachian, which may once have been colonized by microbial life. The analysis of each image requires individual expertise from geobiologists and biogeomorphologists, and thus a new key challenge is accurate detection of those returned images that present the best potential of revealing biotic morphological signatures. The expertise applied to large sets of images is inevitably finite and can additionally be subjected to interpretation biases. To circumvent these human-derived pitfalls, an objective automatic high-throughput recognition procedure is desirable in order to establish first-pass classifications of images and to isolate those with the greatest potential to show morphological signatures of life.

When image data have adjacency structures that can be recognised by the human brain, an Artificial Neural Network can be trained to use these structures to emphasize local relationships between areas of interest. Neural Networks are now recognized as extremely efficient in object recognition and classification on images, even of poor definition (LeCun *et al.*, 1998; 2015, Rawat and Wang, 2017; Geirhos *et* *al.*, 2018; Shrestha and Mahmood, 2019; Zhao *et al.*, 2019). Such techniques have been successfully used in geology and geomorphology for automatic mineral and landform recognition (Du *et al.*, 2019; Liu *et al.*, 2019; Li *et al.*, 2020) and in biology for biofilm characterisation (Buetti-Dinh *et al.*, 2019; Dimauro *et al.*, 2020). These studies have demonstrated the potential for Neural Networks to exhibit superior performance in discrimination compared to the visual interpretation of human experts.

The aim of this study is to develop and test a Neural Network for the detection of MISS-like structures on images from Earth's sedimentary record. The method shows promise to be used to identify potential signatures of life in rock records related to ancient microbial activity on the surface of Mars. In this paper, we provide a detailed analysis based on empirical field observations of present and ancient mat cracks, complemented by a laboratory experiment, to test all the advantages and limits of Neural Network in identifying potential MISS.

2. Material and Methods

2.1. MISS-type: mat cracks

Among the variety of MISS, mat cracks have been selected as a pilot in this study because they represent a type of modern and ancient MISS that are both widespread and most easily distinguished from their abiotic counterparts. These structures are related to microbial mats colonizing sediment surfaces in damp muddy siliciclastic depositional systems, creating an elastic surficial membrane that subsequently fails and tears in a brittle fashion as a result of shrinking during intervals of drying (Tanner, 1978). Their visual aspects reflect sequences of mat growth during submersion and mat destruction under desiccation leading to more or less tick-curled and blistered crack margins (Eriksson *et al.*, 2007; Noffke, 2010), and prolonged drying may lead to the successive development of several variable morphological characteristics (Davies *et al.*, 2017). In contrast, desiccation of cohesive sediment that lacks an elastic surface membrane tends to form more regular polygonal cracks that are geometrically straight and do not typically exhibit blistered crack margins (Goehring *et al.*, 2015; Li and Zhang, 2010; Noffke, 2010).

Furthermore, small to large-scale polygonal cracking patterns with potential desiccation cracks have been observed by both rovers and orbiters on Mars, conclusively indicating past conditions of lake/ocean drying, but with present uncertainty about the role of microbial involvement (Hiesinger and Head, 2000; Seibert and Kargel, 2001; Mangold, 2005; Grotzinger *et al.*, 2005, 2014; El-Maarry *et al.*, 2013; 2014; Edgar *et al.*, 2018; Stein *et al.*, 2018).

Visual detection has inherent limitations as a tool to unequivocally identify life signatures in rocks. As stressed by Noffke (2021) and others (*e.g.*, Cockell and McMahon, 2019; Westall *et al.*, 2021), the identification of MISS-biogenicity without any ambiguity requires a multiproxy approach with information about (i) the environmental situation of a candidate structure and its association with other abiotic structures; (ii) the external macroscopic morphology (visual detection); (ii) the internal micro-forms and textures (identification using light and/or electron microscope); (iii) chemical or isotopic signals (confirmation of biotic activity). The method proposed here is devoted to automatically detect potential MISS structures that should be targeted for further analyses.

2.2. Laboratory experiment

In order to directly compare, under controlled conditions, the formation of desiccation cracks and their morphological particularities with and without induced biofilm, a laboratory experiment was undertaken in a ventilated greenhouse in Toulouse, France. The experiment was designed to test intra- and intergroup variability in the geomorphological response of different types of substrate texture without (control) and with a cyanobacteria biofilm (treatment). Machine learning was used on the images to establish a classification of biotic and abiotic classes with perfect knowledge about the biotic origin and the initial abiotic conditions. The objectives were (i) to test if the strictly physical and biological configurations would result in distinct and well-marked visual signatures; (ii) to confirm that visual differences between the treatments observed by the expert can be well captured by machine learning; and (iii) to compare and test the performance of a Neural Network for discriminating biotic and abiotic images and to validate its more in-depth use for the *in situ* analyses of present and ancient MISS from Earth.

Biofilm was collected from the river bed of the Garonne River, in Toulouse, France (43°34’25’’N; 1°26’04’’E; 135 m a.s.l.), during summer low flows, and directly implanted in experimental plastic trays. Twenty-four rectangular open plastic trays of 29 cm * 19 cm * 7 cm were used. For abiotic control, four trays were filled with a

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

30 mm sand layer (size < 4 mm) and on top a clay layer of 3 mm (smectite clay: green montmorillonite); and four trays with a layer of 40 mm of clay only. The same design was used for biotic treatment but with the implantation of a 3 mm biofilm on top (a total of eight trays). A further four trays were filled with a 30 mm sand layer only and a 2 mm biofilm at the top; another four trays were filled solely with a 2 mm biofilm. All the trays were fully submerged and then subjected to progressive evaporation and desiccation for three months (July-September) under normal conditions of temperature and sun exposure (assumed to be roughly equal to external conditions because of ventilation and the absence of a filter on the greenhouse glass); July: mean T°C = 21.7, max = 35.6, min = 13.1, sun exposure = 261 h; August: mean T°C = 21.8, max = 35.5, min = 12.0, sun exposure = 219 h; September: mean T°C = 20.7, max = 33.4, min = 12.2, sun exposure = 191 h (data collected at Toulouse-Blagnac climatic station, 43°38’19’’N; 1°21’36’’E; 152 a.s.l.). The substrates remained immersed for about 2-3 weeks during July. After full desiccation in September, images were taken of each tray (one vertical image at the nadir and four oblique images at the cardinal points; a total of 120 images). A digital Canon Ixus 180 camera was used for image acquisition (colour: sRGB; resolution:180 dpi).

The performance of a Neural Network in discriminating images from biotic and abiotic samples in the experiment was compared to other categories of machine learning algorithms using the Matlab (MathWorks^{TM}) simulation environment (R2021b). All the tests of the classifiers were conducted using the machine learning toolbox (classification learner) which contains a collection of twenty-nine machine learning algorithms belonging to eight main families of classifier: Support Vector Machine (SVM); Decision Trees (DT); K-Nearest Neighbours (K-NN); Discriminant Analysis (DA); Ensemble methods; Naive Bayes; Kernel Approximation and Neural Network (Mahesh, 2020; Matlab, 2022).

SVM methods use different mathematical functions to determine the boundaries between classes. SVMs are known to perform well with unstructured and semi-structured data such as images, but the results are highly dependent on the kernel function used and the learning time increases greatly with the size of the dataset. The types of SVMs tested here were: Linear, Quadratic, Cubic, Fine Gaussian, Medium Gaussian and Coarse Gaussian. In Decision Trees, classes are predicted by choosing branches of a tree from the root to the leaf nodes. DT are easy to implement and their results are easy to interpret. They are very efficient and fast for small data sets, but this time increases rapidly with the size of the data sets, making them less relevant for dealing with large volumes of data. The types of DT tested were: Fine, Medium and Coarse. K-Nearest Neighbours is based on classifying the data according to the class of its K-nearest neighbours. K-NN is easy to implement and works much faster than the other algorithms because it does not use the training dataset to learn and improve; it only stores the training dataset and learns from it when making predictions. However, it is known to be inefficient with large datasets because computing the distance between each new point and the old ones takes a lot of time. Types of K-NN that were used are: Fine, Medium, Coarse, Cosine, Cubic, and Weighted. Discriminant analysis is based on finding combinations of features that characterize or separate classes. DA is generally effective for classification into two or more classes, but this method may fail if the classes involved are highly intermingled in terms of the distribution of their descriptive parameters. Both Linear and Quadratic analyses were used. Ensemble classification uses the combination of two or more classification methods to improve their individual performance. Different types were used: Bagged Trees, Boosted Trees, Subspace Discriminant, Subspace KNN and RUSBoosted Trees. Naive Bayes classifiers are based on probabilistic classification applying Bayes' theorem with strong (naive) assumptions of independence between features. Naive Bayes analysis is simple and easy to implement. It does not require large training datasets and works quickly, but it assumes that all features are independent, which is rarely the case in real-world settings; thus, it can lead to misclassification. Two types were available for our tests: Gaussian and Kernel. Finally, five types of Neural Network-based classifiers were tested: Narrow medium, Wide, Bilayered and Trilayered. Neural Networks have many advantages, notably that they are very efficient even with very large data sets, but they have the disadvantage of functioning for the user as a black box, giving no information on the process that led to the classification. In our case, the fact of mastering the implementation of the Neural Network and of being able to access the internal layers made it possible to partially overcome this drawback. The tested configurations correspond to Neural Networks with a reduced number of neurons in parallel for the first intermediate layers for the 'narrow' classifier, a higher number for the 'medium' one and a much higher number for the 'wide' network classifier. The two-layer network offers greater flexibility by allowing the width of the first two inner layers of the network to be defined for the Bilayered classifier, and the first three layers for the Trilayered one. In our case, the parameters chosen were those proposed by default by the model, *i.e.*, a value of 10 for the size of the first layer of the Narrow classifier, 25 for the Medium classifier and 100 for the Wide classifier. For the Bilayered classifier, the size of the first two layers was set to 10 and that of the first three layers to 10 for the Trilayered classifier.

In order to increase the number of images for the classification, a data augmentation procedure was developed with the Matlab simulation environment and performed for each of the 120 initial images with a subdivision into four sub-images. Successive rotations of the original images were applied by steps of 10°. For one initial image, we obtained 1 + 4 + 35 = 40 images (*i.e.*, 120 + 40 * 120 = 4,920 images). A correct balance between the number of abiotic images (1,640) and biotic images (3,280) was obtained for training and testing.

The procedure began with the labelling of images into their category (based on expert analysis). The BagOfFeatures function of Matlab was used to provide an encoding scheme representing the large collection of images using a sparse set of 'visual word' histograms (O'Hara and Draper, 2011; Nanni and Lumini, 2013). Five hundred features were thus automatically extracted from each image to allow their classification.

A default method for the training and testing was used with the following steps: the whole dataset was cut into 5 equal subgroups; the model was trained with the data of 4 of the 5 subgroups, and the 5th subgroup was used for testing. The procedure was repeated five times, changing the images belonging the test group and to the 4 training group each time. The purpose of the procedure was an increased training consistency.

2.3. *In situ* images acquisition

The images of present-day features were collected in back-barrier tidal flats of the coastal zone of the Mediterranean Sea, near the village of Peyriac-de-Mer, France (Fig.1; 43°05'13''N; 2°57'33''E; 0 m a.s.l.). The prospected back-barrier tidal flats were dominated by mud-sized (clay) siliciclastic sediments in most of the studied sites. Surfaces were colonized by biofilms, as expected in brackish-water peritidal settings (Gerdes *et al.*, 2000). Microbial mats of different types were observed, depending on local habitat conditions and location on the supratidal gradient of salinity. The dominant type of MISS features that were observed corresponded to large epibenthic mats that had cracked more or less under desiccation following subaerial emergence. Images were collected at different dates, hours in the day, angles and in various locations to capture a wide range of light conditions, habitat types (type of substrate, water saltiness and submersion frequency and duration) and biological (type of microbial consortium) conditions. 2,000 images of mat cracks and 2,000 abiotic desiccation cracks for control were taken in different locations at various sites (Fig. 1). Each location in a study site corresponded to an area of ca. 4 m² showing a homogenous distribution and density of mat cracks or desiccation cracks . Five images were taken for each location with one vertical image at the nadir and four oblique images at the cardinal points. At each location, five images were taken for several small groups of a few forms following the same procedure with the same form at the nadir. A set of 2,000 images of ancient mat cracks (*i.e.*, fossilised mat cracks of ca. 250 Ma) and 1,000 images of ancient desiccation cracks were collected in the Permian Salagou formation in the Lodève Basin (Michel *et al.*, 2015), France, near the village of Octon (Fig. 1; 43°39’16’’N: 3°18’10’’E; 125 m a.s.l.). In order to confirm the biotic origin of the mat cracks three samples of fossilised cracks were collected in the field in different sites and analysed in the lab. MISS generally have specific micro-textures in thin sections and this has been recognized as a formal criterion for biogenicity (Noffke, 2009; Davies *et al.*, 2016). The three samples were sawn vertically in order to produce views in thin cross-sections for detailed visual analyses of the rock/fossilised mat structure under a light microscope (Olympus CX40 coupled with a high-definition colour camera head DS-Fi2). The analyses of the thin cross-sections showed the occurrence of fossilised tortuous biotic filaments (Fig. 2a) and dark mat layers with a fine textural fabric associated with a cyanobacteria activity (Fig. 2b).

All the images were collected with a digital Canon Ixus 180 camera (colour: sRGB; resolution:180 dpi).

2.4. Convolutional Neural Network procedure

To explore more deeply the performance of Neural Networks for detecting mat cracks in the field, a Convolutional Neural Network (CNN) procedure was developed and tested with the images of present and ancient mat-cracks (biotic) and desiccation cracks (abiotic) forming in back-barrier tidal flats (see Table 1 for definition of the technical words related to Artificial Intelligence). Beside the overall goal of classification accuracy, three objectives were defined to orientate the technical choices of model development: (i) quality of calibration; (ii) robustness to perturbations; (iii) ability of generalization. On the one hand, these three points depend on the technical choices of implementation among a wide variety of modern techniques. On the other hand, points (ii) and (iii) depend both on the initial construction (images in learning phase and test phase coming from different photoshoot sequences and sites, presence of artefacts such as debris or lichens on the test set) and the pretreatment of images (augmentation procedures).

Both present and ancient biotic and abiotic images were used in the CNN procedure. For ancient MISS, the images most likely associated with mat cracks were used for the learning procedure. Based on this assumption, consolidated with the observations that ancient and present MISS show common features, two classes were implemented as outputs of the model: Biotic or Abiotic. First, as the images originate from different locations and were taken in different conditions with the same camera, a standardization of the radiance was carried out following Jonnalagedda *et al.* (2021). Second, the images were divided into tiles of 1,024 * 1,024 pixels with no particular focus on the represented forms. The procedure resulted in 54,995 images of 1,024 * 104 pixels, with 34,858 biotic images (*i.e.*, mat cracks) and 20,137 abiotic images (desiccation cracks). Some images contained extraneous elements such as lichens, and more or less agglomerated small twigs, dust and gravels. The images with such artefacts were kept to ensure that the model is able to focus on mat cracks despite the occurrence of the other elements mentioned. This set of images was divided into three subsets: a learning subset of 38,166 images (70%), a validation subset of 11,333 images (20%), and a testing subset of 5,496 images (10%). To avoid correlations due to the fact that several images of a given mat crack were taken with different angles, no images coming from the same shooting sequence were distributed in the different subsets. By doing so, we ensured that the testing set contained images de-correlated from the images that were used in the learning/validating phases. Classical oversampling by duplication on the minority class (Abiotic) was performed to obtain a 50/50 distribution of biotic and abiotic samples in the learning set. Subsampling was stratified to ensure that there were ancient and present biotic and abiotic images in the learning, validation and testing sets.

As the number of available images was not sufficient to train a CNN from scratch, a transfer learning approach was adopted using a ResNet50 (*i.e.*, CNN that is 50 layers deep pre-trained on ImageNet; He *et al.*, 2016). Resnet CNNs have proven their high capacities to deal with a wide variety of cases (classification, detection and localisation). ResNet50 is a good compromise between performance and computational needs (He *et al.*, 2016). The training was done on a bottleneck block 4 on the classification layer and in freezing the deepest layers. This is classical in transfer learning, considering that the deepest layers of the CNN learn to identify very primary and simple features that are common to a very wide variety of objects and high-level layers more complex forms specific to some classes of objects.

The Jensen-Shannon Divergence Consistency Loss was retained as a loss function because of its demonstrated ability to improve the robustness of the training and the stability of the predictions on new inputs (Zheng *et al.*, 2016; Hendrycks *et al.*, 2020). Here, it was based on the binary cross-entropy with label smoothing of 0.1 (Müller *et al.*, 2019). SGD (stochastic gradient descent) with momentum algorithm was used to train the model (Qiang, 1999) with a weight decay of 0.03. For setting the learning rate, a scheduler based on the cyclical learning rates method was used. It improves the classification accuracy in fewer iterations (Smith, 2017).

The training was done on 17 epochs with mini-batches of 32 images. Beyond 17 epochs, the model is over-trained. Following Howard (2018), progressive learning by resizing was performed to accelerate and improve the learning phase: the network was initially trained with smaller resized images (128 pixels) and then on larger images (128, 192 and 224 pixels, respectively on 2, 5 and 10 epochs). An algorithm of random resize crop coupled to the Augmix augmentation process (Hendrycks *et al.*, 2020) was used. As an example, a first minibatch of 32 images arrives and the images are resized to 128 + 4 pixels (or 192 + 16 or 224 + 32). Then, a ROI (region of interest) is randomly chosen (size of 0.1 to 1) on each image and interpolated to generate a 128-pixel image. Augmix (severity = 2; chain.width = 2; chain.depth = 4), consisting of a sequence of image transformations (Hendrycks *et al.*, 2020) was then applied to the resulting image to obtain a triplet of images: the original one and two augmented ones. The model provides as an output the most probable class for a given image associated to a probability that represents the level of confidence of correctness of the classification. In a very simple fashion, we consider that the confidence threshold for belonging to a given class is equal to 0.5.

CNNs and Neural Networks, in general, are black boxes and are excellent to predict but not to explain. During the last decade, a high number of methods were developed to understand why CNNs perform so well and to visualize the features of the images that drive the model decision (Palafox *et al.*, 2017; Noh, 2021; Mani *et al.,* 2022). In the present study, as the searched landforms are not easily distinguishable objects that develop in environments with a lot of visual artefacts, it was fundamentally important to be able to visualize what parts of the images allowed the model to classify biotic and abiotic images. Two of the mainly used visualization methods but based on very different algorithms were chosen to ensure that the visualisation results are consistent. The first approach consisted in representing on the predicted image a heat map localizing the areas with the highest weight for the model decision for one given predicted class. To do that, Grad-CAM++ was implemented, based on a weighted combination of the positive partial derivatives of the last convolutional layer feature maps with respect to a specific class score (Chattopadhyay *et al.*, 2018). Grad-CAM++ was preferred to Grad-CAM because of its better precision and better ability to separate the different areas of importance. The other method was Score-CAM and was based on the linear combination of the feature maps weighted with the corresponding prediction (Softmax) score obtained by a forward pass of these feature maps in the CNN (Wang *et al.*, 2020). To visualize in high definition the pixels that were considered for the decision of the model, Guided Grad-CAM++ and Guided-Score-CAM were implemented, consisting of a combination of ReLU backpropagation with respect to Grad-CAM++ or Score-CAM activation maps (Zeiler and Fergus, 2014). These guided methods allow visualization of the detected features (gradients, forms and patterns) that are important in the prediction of a given class.

Model performances were evaluated in terms of: (i) calibration quality using ECE (expected calibration error) and reliability diagrams. This is based on the idea that, for

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

example, given 100 predictions with a probability of 0.8 for each to belong to a class, 80 will be correctly predicted (Guo *et al.*, 2017). The diagram compares the average probability of predictions to the expected accuracy, calculated on several intervals' bins of prediction confidence (it should be equal in the case of a perfect calibration). ECE is equal to the weighted average gap between expected accuracy and average confidence. A perfect calibration should lead to an ECE of 0; (ii) precision (ratio of good predictions); (iii) AUC (area under curve), sensitivity and specificity as it is a two-class classification. To do that, Biotic is considered here as a positive and Abiotic as a negative class. Using visualization tools, the predictions on the testing set were visually analysed to identify the features that drove the decision-making and to assess whether they corresponded to the features used by humans to do the same classification. Particular attention was given to the images with twigs, lichens, dust and gravels to ensure that these objects didn't influence the classification, and on the misclassified images to understand the sources of errors. Finally, one important question is whether the calculated probabilities associated with the classifications can be used to assess the level of uncertainty of the presence of biotic form. To tackle this issue, we calculated the probability distributions associated with the different prediction types: true biotic, true abiotic, false biotic and false abiotic.

The model and all associated procedures were developed using the language Python with the library Pytorch (v. 1.10.0).

3. Results

3.1. Laboratory experiment and classification of images

After full desiccation, the laboratory experiment indicated well-marked differences between the sets of controls and treatments (Fig. 3). The variation of sediment texture (grain size) and organisation (arrangement and thickness of the layers) significantly affected the geometry of desiccation cracks. Controls (A, B) and treatments (C, D, E, F) showed homogenous responses among the four replicates. The final morphology of the treatments with biofilm (C, D, E, F) varied with the initial sediment structure and texture with more or less fragmented and curved mat cracks.

The biotic signature was expressed at three embedded scales (Fig. 4): (i) the micro-textural aspects of the forms; (ii) the geometrical characteristics of the forms; and (iii) the arrangement of the forms. Differences between abiotic and biotic signatures did not systematically occur at the three spatial scales and instead varied dependent on the controls and the treatments. In the treatments with sand and clay (A, C), the biotic signature was expressed congruently at the three scales; whereas in the treatments with clay (B, D), the biotic signature only occurred as a texture (Fig. 5).

Visual discrimination and homogeneity in the replicates were confirmed by the learning machine classifications that were found to be very coherent with human expertise (Fig. 6). Among the twenty-nine algorithms that were tested, the Neural Network types performed best and showed excellent predictions with similar precision of 0.99% (Table 2). Most of the classification errors concerned the biotic categories (Fig. 6). The rare confusions between biotic and abiotic classes were mostly related to Sand_Clay and Sand_Clay_Bio; and Clay and Clay_Bio. It is most likely similarity in form that explains the confusions.

3.2. Convolutional Neural Network with present and ancient MISS

The CNN classification of *in situ* abiotic and biotic (present and ancient) images showed excellent results (Table 3). The overall precision in the test was 99% for an AUC of 0.99. Considering the Biotic class as the positive class and Abiotic as the negative one, the model sensitivity reached 99.3% and the specificity 97.4%. In detail, the sensitivity was equal to 98.4% for the ancient mat cracks for a specificity of 99.8%. This means that the model missed a few biotic images but produced only one false biotic prediction. Most of these misclassified biotic images were due to subsampled tiles coming from biotic images incorporating abiotic sub-areas. For the present-day mat cracks, the sensitivity was 99.5% with a specificity of 96.1% with few false biotic predictions (30 images representing 3.9% of abiotic images): among these 30 false biotic, 24 images belonged to the same original photograph and were characterized by the presence of mineral flocs without any cracks. That mean that such floc texture cannot by distinguished by itself from the learnt biotic structures without the occurrence of cracks on the image.

Quality of calibration was checked using ECE and reliability diagram. As shown in figure 7, the gap between expected accuracy and level of confidence was very small with an ECE of 0.06. This denotes a good quality of calibration of the model.

Moreover, the majority of correct predictions showed a level of confidence above 0.8, which can be used to assess the level of uncertainty of the presence of biotic structure in the images. More precisely, figure 8 shows the distribution of probabilities associated with each type of prediction. It appears that good predictions were associated with a peak of probability greater than 0.9, and a majority of values above 0.8, with a good separation with false prediction curves. Indeed, false predictions were associated with probabilities mainly lower than 0.7 except the false biotic predictions curve that showed a bimodal distribution with the highest peak above 0.9. But this peak was exclusively associated with the 24 images coming from the same original image we evoked before. Removing these images make the distribution unimodal with values lower than 0.65. Figure 8 shows that a biotic or an abiotic prediction with a probability lower than 0.8 can be suspected to be erroneous but the level of uncertainty of correctness is very low when the probability is greater than 0.9. In a more general way, the value of probability associated to the biotic class predictions (equal to 1-probability of being abiotic) showed a good potential to be used to assess the probability for a given image to contain MISS signatures.

To ensure that good performances were related to the mapping of relevant features, a visual examination of the activation mappings (GradCAM++ and ScoreCAM) was processed to identify the features underlying the model decisions. First, the two visualization methods consistently led to highlighting the same activation areas and providing the same high-definition representation of the discriminative features. Second, the detailed visual examination of the classification outputs image by image for the four modalities (*i.e.*, present and ancient Biotic and Abiotic; examples provided in Fig. 9) showed that the most discriminative components for biogenicity were the border of the mat cracks with their tortuous enlarged and blistered morphology more or less curved upwards, sometimes with thin laminations (Fig. 9a, c). Centimetre circular or oval rips and mat chips (for a detailed description see Noffke, 2010) with rounded edges were also discriminative forms targeted by the model. The model did not focus on the more homogenous inner surface of the mat cracks, nor on the complicated artefacts caused by biotic features related to small twigs, lichens and small shells (tested on one image). The most discriminative components for the abiotic images were the straight narrow desiccation cracks with smooth shapes and the one with a regular slight curvature.

The most angular edges with T, Y or a cross pattern were very discriminant (Fig. 9b, d). Both the largest cracks and the inner fold of the desiccation cracks were not discriminative.

In addition, as images coming from different photoshoot sequences were carefully separated between the learning and test sets and as the images were homogenised and augmented, the presence of hidden features in the images other than their visible content (hidden correlations between images coming from the same shoot sequence) were excluded.

4. Discussion

4.1. Convolutional Neural Network: a tool for the automatic detection of MISS-like structures

Overall, the results showed that Neural Networks, and in particular CNN, are robust and very promising for establishing high-throughput automatic classification of images which are most likely to show structures with a biotic signature, such as MISS-like structures. The detailed visual examination of the classification outputs image by image showed that the key areas of interest of the model corresponded to those used by human expertise to discriminate biotic and abiotic forms, *i.e.*, the borders of the cracks.

The expert analysis image by image of the few false positive and negative outputs revealed three error sources. (i) Certain types of MISS structures that were observed in the field differed from the mat crack features that were dominantly used in this study. These structures were marginally used in the learning procedure. In such a case, when thin biofilms occurred in low densities on an image, they were not detected and the image was classified as abiotic. This result points to the importance of using the largest possible set of images of the wider possible range of MISS-types with various density in the learning procedure. (ii) The images were subsampled for the testing with a focus on one form or a detail of a form. In certain cases, subsamples originating from one image that was classified as biotic by the expert were (correctly) identified as an image showing an abiotic sample. This points to potential issues when subsampling images that were previously identified as biotic. Labelling for the learning procedure should be done after subsampling. (iii) In certain cases, fortuitous convergent patterns between abiotic and biotic features occurred

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

leading to false-positive or negative outputs. False positives were related mostly to a combination of factors including local agglomeration of sands, small gravel, lichens and shade effects mimicking the biotic morphology of a mat crack border or a mat chip. False negatives occurred in cases where mat cracks presented linear and narrow cracks very similar to those produced under abiotic conditions. Overall, these observations suggest that the learning and testing procedures should also be performed with a large set of complicated images. Tricky images with abiotic components mimicking biotic signatures should also be used for testing the performance of the machine. The combination with the learning procedure should increase the robustness and fineness of the recognition procedure. However, it should be noted that the presence of false negatives or false positives was extremely rare. That shows the efficiency and the robustness of the CNN classification.

The field and laboratory observations of present and ancient mat cracks showed that signatures of microorganisms in sediment can be well marked and congruently occur at three embedded scales: texture (*i.e.*, the size, shape and arrangement of the grains constituting the surface), form (*i.e.*, the outer shape of an individual form) and pattern or arrangement of forms (*i.e.*, arrangement of the forms over a few square meters). In the CNN procedure, the borders of the forms were the most discriminative. However, the discriminative objects are dependent on the resolution scale that is used. In this study, forms at the centimetre scale were focused upon without using macro-resolution providing information about the millimetre micro-textural aspects of the form nor to the spatial pattern of forms arrangement at meter scale. We stress that the CNN procedure is applicable at the three different spatial scales that were identified here as relevant interdependent indicators. Consequently, each of the three scales should be used in the CNN procedure for the algorithm to detect on an image a characteristic signature of microorganism activity potentially expressed in the micro-texture, the form and/or the pattern of forms arrangement. All the necessary equipment for producing large sets of images from micro- to landscape scales from the surface of Mars is available on the rovers and orbiters (Edgett *et al.*, 2003; Bibring *et al.*, 2005; Josset *et al.*, 2017; Vago *et al.*, 2017; Farley *et al.*, 2020; Bell *et al.*, 2021; Bhardwaj *et al.*, 2021; Wiens *et al.*, 2021).

4.2. Inherent limits in dissociating biotic and abiotic signatures

This study showed that biotic signatures of microorganisms can be distinguished from abiotic signatures, potentially from micro (µm to few mm) to mesoscales with rovers (few m to tens of m), and potentially at macro-scales with orbiters (tens to hundreds of m). However, the laboratory greenhouse experiment showed that each of the treatments with the same biofilm behaved differently, highlighting the dependency of the biogeomorphological response on initial geomorphological conditions related here to the different types of sediment layers. This result, combined with the observations in the field, suggests that slight variations in physicochemical conditions (*e.g.*, sediment texture, layer thickness and disposition, the saltiness of the water, frequency and duration of the cycles of immersions/emersions) may result in singular modulations of the texture, form and pattern of mat cracks and desiccation cracks. The modern desiccation cracks and mat cracks also exhibited several diverse features that were related to biological attributes of biofilms (*e.g.*, strong variation in microorganism consortiums, thickness, texture, and planar extension of the biofilms) as well as attributes representative of physical-biological interaction such as the combination of mat type and sediment textures and the historic frequency and duration of aqueous submersion (Fig. 10).

A further inherent limitation is that our study has concentrated on only one subset of MISS, amongst a huge diversity of ancient types that are known from Earth's historical rock record. Each individual subset, such as the ancient desiccated MISS used in this article, encompasses a huge diversity of patterns (Fig. 11; *e.g.*, Gerdes *et al*., 2000; Schieber *et al*., 2007; Noffke, 2010; Davies *et al*., 2016). CNN shows great potential for searching for the signature of life in rock, but the procedure must consider a multitude of abiotic and biotic outcomes and be aware that many of these may be unknown or not thought about, and with or without diagnostic criteria available (Davies *et al*., 2018). Only with such a holistic approach, and no preconceived fixation on a possible biotic origin, can the risk of false-negative and positive interpretations be mitigated.

Furthermore, we stress that even with a very extensive database of abiotic and biotic images, making a formal distinction between purely abiotic structures and MISS based solely on expert visual expertise may remain challenging in certain cases. It is now recognized that biogenic forms originating from microbial activity generally are not unequivocal (McLoughlin *et al*., 2008; Noffke, 2009; Davies *et al*., 2016).

Microbes can modulate landform characteristics without creating a biogenic signature that is unequivocally distinguishable from abiotic signatures (Dietrich and Perron, 2006; Corenblit *et al.*, 2011; Davies *et al.*, 2016), resulting in equifinality of morphology that may always be underdetermined in some instances (Davies *et al.*, 2020). Abiotic erosional and tectonic processes can also overprint primary microbial sedimentary structures (Davies and Shillito, 2021), resulting in composite fossilised patterns that are not simple to interpret and have an innately greater diversity of form.

Based on the recognition of equifinality of biosignatures, Davies *et al.* (2016) stressed the importance of increasing our ability to identify abiotic processes that can create morphologically similar features than true MISS. The quality of expert determination of the true biotic structures on the images that are used for the learning procedure is essential. We stress the fact that the use of MISS as an analogue for the search for signatures of life in rocks on Mars thus requires, in the first instance, an increased appreciation of abiotic processes that create similar morphologies on Earth.

Davies *et al.* (2016) pointed out that the use of morphological criteria may be done with pragmatic consideration of the set of localised circumstantial evidence in support of the expert interpretation on a case-by-case basis. The classification scheme for images should include biotic, abiotic and problematic forms. Davies *et al.* (2016) proposed such a classification of landform biogenicity that could be adapted for expert category assignment in the deep-learning procedures (Fig. 12): (A) known to be abiotic in origin; (B) known to be microbial in origin, (ab) where there is uncertainty, and (Ab) or (Ba) where there is uncertainty but one interpretation is favoured. Such a classification could be used in CNN procedures with large training sets of images and accurate labelling for supervised learning. The model output probabilities show good potential to assess the uncertainty of the presence of biotic signatures and this result is a way to be developed in the final objective of associating for a given image a probability to correspond to one of the five Davies' classes.

In the laboratory experiment, curved desiccation cracks occurred in the Sand_Clay abiotic control but the curved structures that developed in the biotic treatment Sand_Clay_Bio were larger (Fig. 5). Furthermore, the texture of the biotic curved cracks was granular and one of the abiotic curved cracks remained smooth. The

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

general pattern of arrangement of forms also varied between the abiotic control and biotic treatment with the size of the forms. These results indicate that improving our capacity to identify biotic origin in ab, Ba and Ab (Davies' problematic categories) may depend principally on our capacity to identify and understand the key textural, geometric and paternal geomorphological parameters and the spatial scales that are affected by life from micro- to mesoscales. Formally establishing abiotic and biotic signatures on images will require the more extensive use of laboratory experiments combined with empirical observation along spatial gradients of exposure to submersion and saltiness and at various time periods during desiccation. .

5. Conclusion

The use of CNN to automatically detect a signature of life in rock and sediment offers great potential and needs to be fully explored and developed for targeting biogenic landforms. It should help make the first classifications of images with the best potential from large sets of Martian images.

However, our experimental results and our observations in the field also highlighted all the complexity and variability of the signatures of life in rocks and sediment. Much caution needs to be taken during the image labelling procedure in order to decrease as much as possible false positive or negative outputs. The use of Earth as a biogeomorphological analogue provides good reasons to hypothesize that ancient microbial life may have given rise to a variety of biogenic landforms on Mars. The technique developed and tested with MISS in this article is aimed to be used for other kinds of biogenic landforms also. In that perspective, we propose the need for a web-based platform for a standardized database of images of the three main categories of biogenic landforms found on Earth with possible analogues on Mars in rocks, sediments and ice, *i.e.*, biodeposition, bioweathering and bioturbation . Such a database of images would improve our capacity to detect potential biogeomorphological signatures of life in rocks on Mars and other telluric planets over the wider range of possibilities.

The question if CNN procedures can reach the ability to detect extremely fine details (textures, forms or patterns) for discriminating biotic from abiotic confounding signatures remains open. The possibility that human and artificial expertise may reach an inherent limit is probable. As pointed out by Noffke (2009, 2010, 2021), the confirmation of biotic origin in rock records requires integrative multiproxy approaches combining geomorphological, chemical and isotopic analyses . As shown by Hickman-Lewis *et al*. (2022), the combination of textural (visual) and chemical analyses using WATSON (imaging), SHERLOC (deep-UV Raman and fluorescence spectroscopy), and PIXL (X-ray lithochemistry) equipment's of the Perseverance Rover is particularly promising for discriminating organic and mineral signatures within rock laminated organic-bearing samples.

Acknowledgments

We thank the *Observatoire Midi-Pyrénées* (OMP), Toulouse, for their financial support provided to the ExoBioGeo project. The first author also thanks the CNRS for his half-time delegation to *Laboratoire Écologie Fonctionnelle et Environnement* in 2020/21 and 2021/22. CNN developments on *ex situ* images were financed by MAD-Environnement. We are grateful to Daniel Vizcaïno for is help in the palaeontological prospections in the southern *Montagne Noire*, France; Fabienne de Parseval, François Maubé and Julien Berger from GET, Toulouse, for rock sampling preparation and analysis; and Marius Alleb for his help in the *in situ* experiment. We are also grateful to the two anonymous reviewers for their constructive comments.

Authorship confirmation/contribution statement

D.C., O.C., S.D., J.-P.T., F.A., M.-F.A., J.D., N.S.D., F.J., T.O., G.R., E.R., J.S. and H.V. wrote the paper.

Authors' Disclosure Statement

The authors declare no conflict of interest.

Funding Statement

This paper was partially supported by the A01 Project of the *Observatoire Midi-Pyrénées* (OMP), Toulouse.

References

Baker VR (2008) Planetary landscape systems: a limitless frontier. *Earth Surface Processes and Landforms* 33:1341-1353.

Baucon A, de Carvalho CN, Barbieri, *et al.* (2017) Organism-substrate interactions and astrobiology: Potential, models and methods. *Earth-Science Reviews* 171:141-180.

Bell JF, Maki JN, Mehall GL, *et al.* (2021). The Mars 2020 perseverance rover mast camera zoom (Mastcam-Z) multispectral, stereoscopic imaging investigation. *Space science reviews* 217:1-40.

Bhardwaj A, Sam L and Gharehchahi S (2021) Four decades of understanding Martian geomorphology: Revisiting Baker's 'The geomorphology of Mars'. *Progress in Physical Geography: Earth and Environment* 45:979-989.

Bibring JP, Langevin Y, Gendrin A, *et al.* (2005) Mars surface diversity as revealed by the OMEGA/Mars Express observations. *Science* 307:1576-1581.

Bosak T, Moore KR, Gong J, *et al.* (2021) Searching for biosignatures in sedimentary rocks from early Earth and Mars. *Nature Reviews Earth & Environment*:1-17.

Buetti-Dinh A, Galli V, Bellenberg S, *et al.* (2019) Deep neural networks outperform human expert's capacity in characterizing bioleaching bacterial biofilm composition. *Biotechnology Reports* 22:e00321.

Cabrol NA (2018) The coevolution of life and environment on Mars: An ecosystem perspective on the robotic exploration of biosignatures. *Astrobiology* 18:1-27.

Cady SL (2001) Paleobiology of the ancient microbes, extreme environments, and the origin of life. *Advances in Applied Microbiology* 50:1-35.

Cady SL and Noffke N (2009) Geobiology: Evidence for early life on Earth and the search for life on other planets. *GSA Today* 19:4-10.

Cady SL, Farmer JD, Grotzinger JP, *et al.* (2004) Morphological biosignatures and the search for life on Mars. *Astrobiology* 3:351-368.

Carmona NB, Ponce JJ, Wetzel A, *et al.* (2012) Microbially induced sedimentary structures in Neogene tidal flats from Argentina: paleoenvironmental, stratigraphic and taphonomic implications. *Palaeogeography Palaeoclimatology Palaeoecology* 9:353-355.

Carter NEA and Viles HA (2005) Bioprotection explored: the story of a little known Earth surface process. *Geomorphology* 67:273-281.

Chacon-Baca E, Santos A, Sarmiento AM, *et al.* (2021). Acid Mine Drainage as Energizing Microbial Niches for the Formation of Iron Stromatolites: The Tintillo River in Southwest Spain. *Astrobiology* 21:443-463.

Chattopadhyay A, Sarkar A, Howlader P, *et al.* (2018) Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks. In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp 839-847.

Coates AJ, Jaumann R, Griffiths AD, *et al.* (2017) The PanCam Instrument for the ExoMars Rover. *Astrobiology* 17:511-541.

Cockell CS and McMahon S (2019) Lifeless Martian samples and their significance. *Nature Astronomy* 3:468-470.

Corenblit D, Baas ACW, Bornette G, *et al.* (2011) Feedbacks between geomorphology and biota controlling Earth surface processes and landforms: a review of foundation concepts and present understandings. *Earth-Science Reviews* 106:307-331.

Corenblit D, Darrozes J, Julien F, *et al.* (2019) The search for a signature of life on Mars: a biogeomorphological approach. *Astrobiology* 19:1279-1291.

Cuadrado DG and Pan J (2018) Field observations on the evolution of reticulate patterns in microbial mats in a modern siliciclastic coastal environment. *Journal of Sedimentary Research* 88:24-37.

Damer B and Deamer DW (2019) The hot spring hypothesis for an origin of life. *Astrobiology* 20:429-452.

Davies NS and Shillito AP (2021) True substrates: The exceptional resolution and unexceptional preservation of deep time snapshots on bedding surfaces. *Sedimentology* 68:3307-3356.

Davies NS, Liu AG, Gibling MR, *et al.* (2016) Resolving MISS conceptions and misconceptions: a geological approach to sedimentary surface textures generated by microbial and abiotic processes. *Earth-Science Reviews* 154:210-246.

Davies NS, Liu AG, Gibling MR, *et al.* (2018) Reply to comment on the paper by Davies *et al.* "Resolving MISS conceptions and misconceptions: A geological approach to sedimentary surface textures generated by microbial and abiotic processes". *Earth-Science Reviews* 176:384-386.

Davies NS, Shillito AP and McMahon WJ (2017) Short-term evolution of primary sedimentary surface textures (microbial, abiotic, ichnological) on a dry stream bed: modern observations and ancient implications. *Palaios* 32:125-134.

Davies NS, Shillito AP, Slater BJ, *et al.* (2020) Evolutionary synchrony of Earth's biosphere and sedimentary-stratigraphic record. *Earth-Science Reviews* 201:102979.

Dietrich WE and Perron JT (2006) The search for a topographic signature of life. *Nature* 439:411-418.

Dimauro G, Deperte F, Maglietta R, *et al.* (2020) A Novel Approach for Biofilm Detection Based on a Convolutional Neural Network. *Electronics* 9:881.

Du L, You X, Li K, *et al.* (2019) Multi-modal deep learning for landform recognition. *ISPRS Journal of Photogrammetry and Remote Sensing* 158:63-75.

Edgar LA, Gupta S, Rubin DM, *et al.* (2018) Shaler: in situ analysis of a fluvial sedimentary deposit on Mars. *Sedimentology* 65:96-122.

Edgett KS, Williams RM, Malin MC, *et al.* (2003) Mars landscape evolution: Influence of stratigraphy on geomorphology in the north polar region. *Geomorphology* 52:289-297.

Ehlmann BL, Mustard JF, Fassett CI, *et al.* (2008) Clay minerals in delta deposits and organic preservation potential on Mars. *Nature Geoscience* 1:355-358.

El-Maarry MR, Pommerol A and Thomas N (2013) Analysis of polygonal cracking patterns in chloride-bearing terrains on Mars: Indicators of ancient playa settings. *Journal of Geophysical Research: Planets* 118:2263-2278.

El-Maarry MR, Watters W, McKeown NK, *et al.* (2014) Potential desiccation cracks on Mars: A synthesis from modeling, analogue-field studies, and global observations. *Icarus* 241:248-268.

Eriksson PG, Porada H, Banerjee S, *et al.* (2007) Mat-destruction features. Atlas of microbial mats features preserved within the siliciclastic rock record. *Atlases in Geosciences* 2:76-105.

Farley KA, Williford KH, Stack KM, *et al.* (2020) Mars 2020 mission overview. *Space Science Reviews* 216:1-41.

Geirhos R, Temme CRM, Rauber J, *et al.* (2018) Generalisation in humans and deep neural networks. In 2018 Advances in Neural Information Processing Systems, edited by S Bengio, H Wallach, H Larochelle, K Grauman, N Cesa-Bianchi, and R Garnett, Volume 31. Curran Associates, Red Hook, NY, pp 7549-7561.

Gerdes G, Klenke T and Noffke N (2000) Microbial signatures in peritidal siliciclastic sediments: a catalogue. *Sedimentology* 47:279-308.

Gilbert GK (1886) The inculcation of scientific method by example. *American Journal of Science* 31:284-299.

Goehring L, Nakahara A, Dutta T, *et al.* (2015) Desiccation cracks and their patterns: Formation and Modelling in Science and Nature. John Wiley & Sons.

Grotzinger JP, Arvidson RE, Bell Iii JF, *et al.* (2005) Stratigraphy and sedimentology of a dry to wet eolian depositional system, Burns formation, Meridiani Planum, Mars. *Earth and Planetary Science Letters* 240:11-72.

Grotzinger JP, Sumner DY, Kah LC, *et al.* (2014) A habitable fluvio-lacustrine environment at Yellowknife Bay, Gale Crater, Mars. *Science* 343:6169.

Guo C, Pleiss G, Sun Y, *et al.* (2017). On calibration of modern neural networks. In International Conference on Machine Learning, PMLR, pp. 1321-1330.

Hays LE, Graham HV, Des Marais DJ, *et al.* (2017) Biosignature preservation and detection in Mars analog environments. *Astrobiology* 17:363-400.

He K, Zhang X, Ren S, *et al.* (2016) Deep Residual Learning for Image Recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778.

Hendrycks D, Mu N, Cubuk ED, *et al.* (2020) AugMix: A Simple Data Processing Method to Improve Robustness and Uncertainty. *ArXiv*, abs/1912.02781.

Hickman-Lewis K, Moore, KR, Hollis, J JR, *et al.* (2022). *In situ* Identification of Paleoarchean Biosignatures Using Colocated Perseverance Rover Analyses: Perspectives for *in situ* Mars Science and Sample Return. *Astrobiology*.

Hiesinger H and Head III JW (2000) Characteristics and origin of polygonal terrain in southern Utopia Planitia, Mars: results from Mars Orbiter Laser Altimeter and Mars Orbiter Camera data. *Journal of Geophysical Research: Planets* 105:11999-12022.

Howard J (2018) Training imagenet in 3 hours for 25 minutes. Available online at https://www.fast.ai/2018/04/30/dawnbench-fastai/

Huang J, Salvatore MR, Edwards CS, *et al.* (2018) A Complex fluviolacustrine environment on Early Mars and its astrobiological potentials. *Astrobiology* 18:1081-1091.

Huang W, Ertekin E, Wang T, *et al.* (2020) Mechanism of water extraction from gypsum rock by desert colonizing microorganisms. *Proceedings of the National Academy of Sciences* 117:10681-10687.

Ibarra Y and Corsetti FA (2016) Lateral Comparative Investigation 617 of Stromatolites: Astrobiological Implications and Assessment of Scales of Control. *Astrobiology* 16:271-81.

Jonnalagedda P, Surprenant R, Droser M, *et al.* (2021) SPACESeg: Automated Detection of Bed Junction Morphologies Indicating Signs of Life in Ediacaran Period. In 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 2095-2104.

Joseph RJ, Graham L, Büdel B, *et al.* (2020) Mars: Algae, Lichens, Fossils, Minerals, Microbial Mats, and Stromatolites in Gale Crater. *Journal of Astrobiology and Space Science Reviews* 3:40-111.

Josset JL, Westall F, Hofmann BA, *et al.* (2017) The Close-Up Imager onboard the ESA ExoMars Rover: objectives, description, operations, and science validation activities. *Astrobiology* 17:595-611.

Kwan C (2021) A Brief Review of Some Interesting Mars Rover Image Enhancement Projects. *Computers* 10:111.

Landis GA (2001) Martian Water: Are There Extant Halobacteria on Mars? *Astrobiology* 1:161-164

LeCun Y, Bengio Y and Hinton G (2015) Deep learning. *Nature* 521:436-444.

LeCun Y, Bottou L, Bengio Y, *et al.* (1998) Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86:2278-2324.

Lepot K (2020) Signatures of early microbial life from the Archean (4 to 2.5 Ga) eon. *Earth-Science Reviews* 209: 103296.

Li JH and Zhang LM (2010) Geometric parameters and REV of a crack network in soil. *Computers and Geotechnics* 37:466-475.

Li S, Xiong L, Tang G, *et al.* (2020) Deep learning-based approach for landform classification from integrated data sources of digital elevation model and imagery. *Geomorphology* 354:107045.

Limaye SS, Mogul R, Smith DJ, *et al.* (2018) Venus' spectral signatures and the potential for life in the clouds. *Astrobiology* 18:1181-1198.

Liu C, Li M, Zhang Y, *et al.* (2019) An enhanced rock mineral recognition method integrating a deep learning model and clustering algorithm. *Minerals* 9:516.

Longo A and Damer B (2020) Factoring origin of life hypotheses into the search for life in the solar system and beyond. *Life* 10:52.

Mahesh B (2020) Machine Learning Algorithms-A Review. *International Journal of Science and Research* 9:381-386.

Maisano L, Cuadrado DG and Gomez EA (2019) processes of MISS-formation in a modern siliciclastic tidal flat, Patagonia (Argentinia). *Sedimentary Geology* 381:1-12.

Mangold N (2005) High latitude patterned grounds on Mars: Classification, distribution and climatic control. *Icarus* 174:336-359.

Mangold N, Gupta S, Gasnault O, *et al.* (2021) Perseverance rover reveals an ancient delta-lake system and flood deposits at Jezero crater, Mars. *Science* 374:711–717.

Mani VRS, Saravanaselvan A and Arumugam N (2022). Performance comparison of CNN, QNN and BNN deep neural networks for real-time object detection using ZYNQ FPGA node. *Microelectronics Journal* 119:105319.

Marshall CP, Marshall AO, Aitken JB, *et al.* (2017) Imaging of Vanadium in Microfossils: A New Potential Biosignature. *Astrobiology* 17:1069-1076.

Matlab (2020) Statistics and machine learning toolbox. Available online at: https://fr.mathworks.com/help/pdf_doc/stats/stats.pdf

McLoughlin N, Brasier MD, Wacey D, *et al.* (2007) On biogenicity criteria for endolithic microborings on early Earth and beyond. *Astrobiology* 7:10-26.

McMahon S, Bosak T, Grotzinger JP, *et al.* (2018) A field guide to finding fossils on Mars. *Journal of Geophysical Research Planets* 123:1012-1040.

McMahon WJ, Davies NS, Liu AG and Went DJ (2022) Enigma variations: characteristics and likely origin of the problematic surface texture Arumberia, as recognized from an exceptional bedding plane exposure and the global record. *Geological Magazine* 159:1-20.

Michel LA, Tabor NJ, Montañez IP, *et al.* (2015) Chronostratigraphy and paleoclimatology of the Lodève Basin, France: evidence for a pan-tropical aridification event across the Carboniferous–Permian boundary. *Palaeogeography, Palaeoclimatology, Palaeoecology* 430:118-131.

Müller R, Kornblith S and Hinton GE (2019) When Does Label Smoothing Help? *Advances in neural information processing systems* 32.

Naylor LA (2005) The contributions of biogeomorphology to the emerging field of geobiology. *Palaeogeogr Palaeoecol* 219:35-51.

Naylor LA, Viles HA and Carter NEA (2002) Biogeomorphology revisited: looking towards the future. *Geomorphology* 47:3-14.

Noffke N (2000) Extensive microbial mats and their influences on the erosional and depositional dynamics of a siliciclastic cold water environment (Lower Arenigian, Montagne Noire, France). *Sedimentary Geology* 136:207-215.

Noffke N (2009) The criteria for the biogeneicity of microbially induced sedimentary structures (MISS) in Archean and younger, sandy deposits. *Earth-Science Reviews* 96:173-180.

Noffke N (2010) Microbial Mats in Sandy Deposits from the Archean Era to Today. Heidelberg, Germany, Springer Verlag.

Noffke N (2021) Microbially induced sedimentary structures in clastic deposits: implication for the prospection for fossil life on Mars. *Astrobiology* 21:866-892.

Noffke N, Christian D, Wacey D, *et al.* (2013) Microbially induced sedimentary structures recording an ancient ecosystem in the ca. 3.48 billion-year-old Dresser Formation, Pilbara, Western Australia. *Astrobiology* 13: 1103-1124.

Noh SH (2021) Performance comparison of CNN models using gradient flow analysis. *Informatics* 8:53.

Palafox LF, Hamilton CW, Scheidt SP, *et al*. (2017) Automated detection of geological landforms on Mars using Convolutional Neural Networks. *Computers & geosciences* 101:48-56.

Porada H and Bouougri EH (2007) Wrinkle structures—a critical review. *Earth-Science Reviews* 81:199-215.

Qian N (1999) On the momentum term in gradient descent learning algorithms. *Neural networks: the official journal of the International Neural Network Society* 12:145-151.

Rawat W and Wang Z (2017) Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation* 29:2352-2449.

Rizzo V (2020) Why should geological criteria used on Earth not be valid also for Mars? Evidence of possible microbialites and algae in extinct Martian lakes. *International Journal of Astrobiology* 19:283-294.

Schieber J, Bose PK, Eriksson PG, *et al.* (Eds.). (2007). Atlas of microbial mat features preserved within the siliciclastic rock record. Elsevier.

Schon S, Head J and Fassett C (2012) An overfilled lacustrine system and progradational delta in Jezero crater, Mars: Implications for Noachian climate. *Planetary and Space Science* 67:28-45.

Schwieterman EW, Kiang NY, Parenteau MN, *et al.* (2018) Exoplanet Biosignatures: A review of remotely detectable signs of life. *Astrobiology* 18:663-708.

Seibert NM and Kargel JS (2001) Small-scale Martian polygonal terrain: Implications for liquid surface water. *Geophysical Research Letters* 28:899-902.

Shrestha A and Mahmood A (2019) Review of deep learning algorithms and architectures. *IEEE Access* 7:53040-53065.

Smith LN (2017) Cyclical Learning Rates for Training Neural Networks. In 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 464-472.

Stal LJ (2003) Microphytobenthos, their extracellular polymeric substances, and the morphogenesis of intertidal sediments. *Geomicrobiology Journal* 20:463-478.

Stein N, Grotzinger J, Schieber J, *et al.* (2018) Desiccation cracks provide evidence of lake drying on Mars, Sutton Island member, Murray Formation, Gale Crater. *Geology* 46:515-518.

Tanner PWG (1978) Desiccation structures (mud cracks, etc.). In Encyclopedia of Sediments and Sedimentary Rocks, edited by GV Middleton, MJ Church, M Coniglio, LA Hardie and FJ Longstaffe, Encyclopedia of Earth Sciences Series, Springer, Dordrecht.

Thomas K, Herminghaus S, Porada H, *et al.* (2013) Formation of Kinneyia via shear-induced instabilities in microbial mats. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371:20120362.

Vago JL, Westall F, Coates AJ, *et al.* (2017) Habitability on early Mars and the search for biosignatures with the ExoMars Rover. *Astrobiology* 17:471-510.

van Zuilen M (2008) Stable isotope ratios as a biomarker on Mars. *Space Science Reviews* 135:221-232.

Viles HA (2008) Understanding dryland landscape dynamics: do biological crusts hold the key? *Geography Compass* 2: 899-919.

Viles HA (2012) Microbial geomorphology: A neglected link between life and landscape. *Geomorphology* 157-158:6-16.

Wang H, Wang Z, Du M, *et al.* (2020) Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 111-119.

Webster CR, Mahaffy PR, Atreya SK, *et al.* (2013) Low upper limit to methane abundance on Mars. *Science* 342:355-357.

Westall F, Foucher F, Bost N, *et al.* (2015) Biosignatures on Mars: what, where, and how? Implications for the search for martian life. *Astrobiology* 15:998-1029.

Westall F, Hickman-Lewis K, Cavalazzi B, *et al.* (2021) On biosignatures for Mars. *International Journal of Astrobiology* 20:377-393.

Wiens RC, Maurice S, Robinson SH, *et al.* (2021) The SuperCam instrument suite on the NASA Mars 2020 rover: body unit and combined system tests. *Space Science Reviews* 217:1-87.

Zeiler MD and Fergus R (2014, September). Visualizing and understanding convolutional networks. In European conference on computer vision, Springer, Cham, pp. 818-833.

Zhao ZQ, Zheng P, Xu ST, *et al.* (2019) Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems* 30:3212-3232.

Zheng S, Song Y, Leung T, *et al.* (2016) Improving the Robustness of Deep Neural Networks via Stability Training. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4480-4488.

Tables

Table 1. Definition of technical words commonly used in Artificial Intelligence.

Word	Definition
ANN or NN	Artificial Neural Network: network of calculation nodes (neurons) fully or partially connected to each other to process information. It is inspired from biological neural system. A large number of architectures exists and is dedicated to address prediction/classification/decision problems. Learning process can be supervised (learning inputs are pre-classified) or unsupervised (learning inputs are not classified).
AUC	Area Under Curve: classic estimate of the predictive power of a model compared to a random model. It is based on the representation of the sensitivity curve of the model (capacity to predict true positives) according to its specificity (capacity to predict true negatives).
Augmentation	Technique consisting of generating new images from an initial dataset by introducing minor changes (*e.g.*, rotation, flipping, shifting, distortion, noise) that aims to consolidate the learning process by introducing variations and increase the size of the dataset.
CNN	Convolutional Neural Network: particular architecture of supervised neural network dedicated to image processing. It is based on a succession of convolutional layers and normalization layers, connected to detect key patterns and features in an image and to classify it. Basically, the deepest layers detect simple and basic shapes (*e.g.*, lines, curves, circles, dots) and the successive upper successive layers detect more and more complex patterns. It is widely used in image recognition and classification.
Cross-entropy	Cross-entropy measures the difference between the predicted probability of belonging to a given class and the true probability. It is used in classification modelling problems, especially in deep learning.
ECE	Expected Calibration Error: estimation of the accuracy of the calibration of a model based on the difference between the expected accuracy and the average probability of predictions. A perfect calibration leads to a null ECE.
Epoch	An epoch is a cycle of training during which all the data from the learning dataset are presented to the model.
Grad-CAM++	Gradient Weighted Class Activation Map (++ is a variation of): mapping of activated neurons given an image and a prediction class. Visually represented by a heatmap, it is used to find in an image which areas contribute the most to the classification output. It is calculated as the importance of the feature maps of the output

	layer weighted by the average gradient for the considered class and rescaled to the original image.
Guided Grad-Cam / Score-Cam	Combination of Grad-Cam or Score-Cam activation map with Guided Backpropagation activation map to visualize the image features (fine-grained important pixels) that participates to the final prediction of the considered class. Guided Backpropagation is a gradient-based technique which computes the gradient of the target output with respect to the input, but gradients of ReLU functions are overridden so that only non-negative gradients are backpropagated.
Loss	Objective function of NN models that quantifies the difference between the predicted outcome of a model and the expected outcome. Here, cross-entropy is used as the loss function.
ReLU	Rectified Linear Unit: function of activation of neurons used in neural network modelling.
ResNetxx (here ResNet50)	Class of CNN with a number of layers equal to x (ResNet50=50 layers) that was pre-trained on ImageNets database. ResNets are used for transfer learning.
Score-CAM	Gradient Weighted Class Activation Map: Alternative to Grad-CAM to generate mapping of activated neurons given an image and a prediction class. Visually represented by a heatmap, it is used to find in an image which areas contribute the most to the classification output. It is based on the linear combination of feature maps of the output layer scaled to the image size and weighted by their Softmax score obtained by passing forward into the model.
SGD	Stochastic Gradient Descent: method for optimizing an objective function of a model based on the local approximation of the actual gradient in iteratively using random subsets of data. The use is generalized in deep learning modelling as it is proved to be more efficient to address large datasets.
Transfer Learning	Based on the idea that the first layers of CNN detect basic forms and the deeper layers more complex patterns that are more specific to the addressed problem (*e.g.*, to classify animal species), transfer learning uses CNN that were previously trained on a large number of images (several millions) to adapt it to a new classification problem: the pre-trained deepest layers (basic forms detection) are frozen and only the upper layers (complex patterns of basic forms) are trained on the new dataset. It solves the issues of dataset size needed to learn a CNN from scratch and of computational resources (energy, calculation power).

974 **Table 2.** Performances of the five Neural Network classification models in test.

Parameter	Narrow Neuronal Network	Medium Neuronal Network	Wide Neuronal Network	Bilayered Neuronal Network	Trilayered Neuronal Network
AUC	0.99	0.99	0.99	0.99	0.99
Precision	0.99	0.99	0.99	0.99	0.99
True biotic	3,280	3,279	3,280	3,279	3,274
False biotic	4	1	1	1	2
True abiotic	1,636	1,639	1,639	1,639	1,638
False abiotic	0	1	0	1	6
Sensitivity	1	0.99	1	0.99	0.99
Specificity	0.99	0.99	0.99	0.99	0.99

975

976 **Table 3.** Performances of the CNN classification model in test.

Parameter	Global	Ancient	Present
AUC		0.99	
Precision	0.99	0.99	0.99
True biotic	4,262	749	3,513
False biotic	31	1	30
True abiotic	1,172	440	732
False abiotic	31	12	19
Sensitivity	0.99	0.98	0.99
Specificity	0.97	0.99	0.96

977

Figure legends

Fig. 1. Location map of the two study areas for image collection. Present mat cracks (MISS) and desiccation cracks (abiotic) were collected near the village of Peyriac-de-Mer and ancient mat cracks and desiccation cracks near the village of Octon in the Permian (ca. 250 Ma) Salagou basin, France. The 10 cm scale applies to the four images. Photographs: D. Corenblit.

Fig. 2. Analyses performed on a vertical thin cross-section of fossilised microbial mat fabrics (Salagou Permian basin, France; ca. 250 Ma) under light microscope. a) elongated segmented and tortuous biotic filaments; b) Multi-layered carbon-rich mat interlaced with coarser mineral layers and overlying a quartz-dominated layer. The organic carbon-rich biotic laminae are in dark brown and have a finer texture than the mineral layers.

Fig. 3. Geomorphological patterns after full desiccation related to different control and treatment configurations. Two abiotic controls (A, B) and for biotic treatments (C, D, E, F). A: Sand_Clay; B: Clay; C: Sand_Clay_Bio; D: Clay_Bio; E: Sand_Bio; F: Bio (Bio: biofilm).

Fig. 4. Three key embedded spatial scale of geomorphological expression of MISS, from left to right: micro-texture (the size, shape and arrangement of the grains constituting the surface), form (outer shape of one form) and pattern of forms arrangement (arrangement of the forms over few square meters). The laboratory experiment suggests that signatures of microorganisms in rocks may potentially occur at three embedded spatial scales and that the detection algorithm should be able to detect biotic signatures at the three scales. The 5 cm scale only applies to the central image.

Fig. 5. Variations in desiccation cracks and mat crack features according to sediment texture and to the presence (biotic treatment) and absence (abiotic control) of a biofilm. The differences between abiotic and biotic signatures did not systematically occur at the three spatial scales identified in figure 3. It varied depending on the controls and the treatments, indicating that the detection algorithm may be capable of identifying a biotic signature at a given spatial scale independently from another. The beginning of bacterial colonization seen on the clay surface of B category at the end

of the experiment illustrates the difficulty of keeping strictly abiotic conditions in long standing experiments.

Fig. 6. Accuracy of the predictions and sources of error for the Wide Neural Network; The algorithm showed here exceptionally good classifications. The successful classifications are indicated in blue and the mis-classifications in tan; confusion between abiotic and biotic classes are indicated in red.

Fig. 7. Assessment of the quality of calibration using reliability diagram and ECE. A perfect calibration should lead to the alignment of the expected accuracy for each confidence interval on the diagonal. ECE represents the average gap and should be equal to 0 in the case of a perfect calibration.

Fig. 8. Distributions of probabilities associated with the correct and incorrect predictions. For false biotic predictions, the second peak of the bimodal distribution above 0.9 is exclusively related to the false predictions of images coming from the same original image characterized by mineral flocs without cracks. Removing this image leads to a single peak around 0.6. For a question of readability, the density was scaled for each type (with a maximal value by class rescaled to 1).

Fig. 9. Visualisation of the interest areas of the Neural Network (GRAD-CAM++ and Score Cam; for all cases the real sRGB image is shown at the right. **(a)** In the ancient mat cracks (MISS). **(b)** In the ancient desiccation cracks (abiotic); the parallel lines are fossilised ripple marks. **(c)** In the present mat cracks (MISS). **(d)** In the present desiccation cracks (abiotic). The heatmaps show the importance of the different areas in the final prediction of a given class (red is the most important areas) and the grey images show fine-grained pixels patterns that participate to the prediction (important image features).

Fig. 10. Examples (not exhaustive) of present MISS features observed in the study sites of Peyriac-de-Mer in relation to variation in biofilm type (x axis) and desiccation intensity (y axis). From left to right, the first column shows to the formation of tick polygonal cracks in a succession of microbial mats developing on a sandy flat; the second column shows the formation of thin elongated shrinkage cracks with margins rolling up and evolving into mat curls. The last column shows a tick and resistant microbial mat resembling a blistered ‘elephant skin’ that can be turned over locally by the current. Photographs: D. Corenblit.

Fig. 11. Examples (not exhaustive) of ancient desiccated MISS features preserved in the rock record. a) Square cracks with reticulate MISS on individual plates. Mesoproterozoic Copper Harbor Formation, Michigan, USA. b) Desiccation cracks within sediment bearing reticulate MISS. Neoproterozoic Diabaig Formation, Scotland. c) Detail of Arumberia fabric on desiccated plate (see McMahon *et al.*, 2022). Neoproterozoic Synalds Formation, Shropshire, England. d) Desiccation cracks with microbial bubbles on plates. Cambrian Port Lazo Formation, Brittany, France. e) Sand filled cracks within sand host sediment, implying cohesion of granular media. Ordovician Graafwater Formation, Western Cape, South Africa. f) Desiccation cracks associated with MISS wrinkle marks and stromatolites. Mississippian Hastings Formation, Nova Scotia, Canada. Photographs: N.S. Davies.

Fig. 12. Davies *et al.*'s classification of sedimentary surface on a gradient of uncertainty (from left to right) about the biotic origin. Adapted from Davies *et al.* (2016).

1056 **Abbreviations Used**

1057 MISS = microbially induced sedimentary structures

1058 sRGB = standard red green blue

1059

Figures

Figure 1

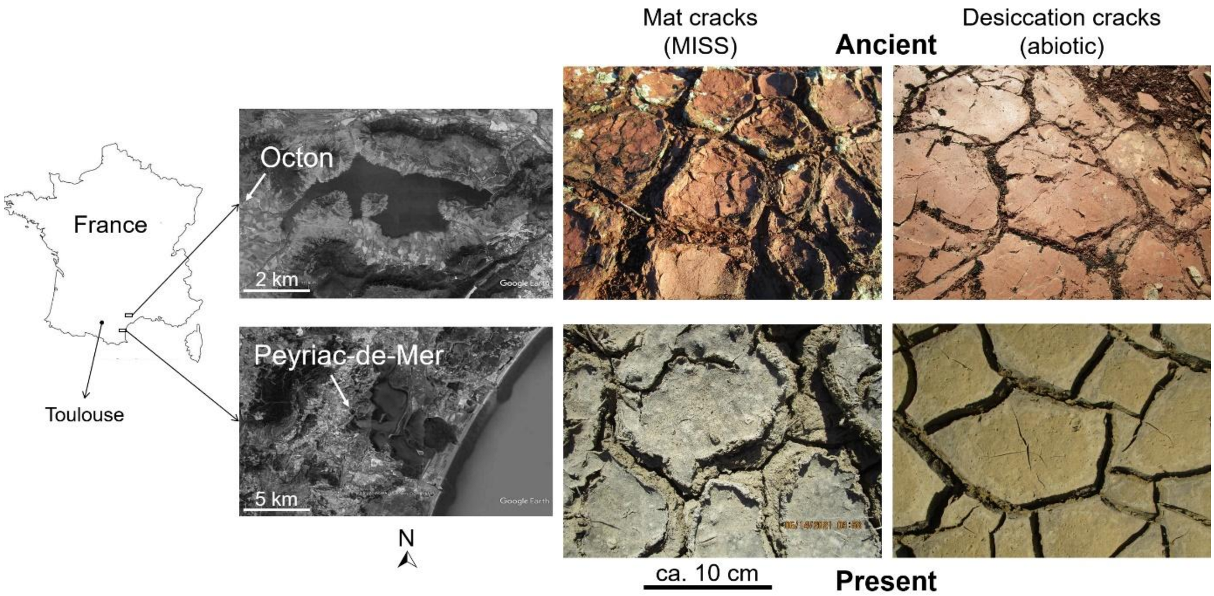
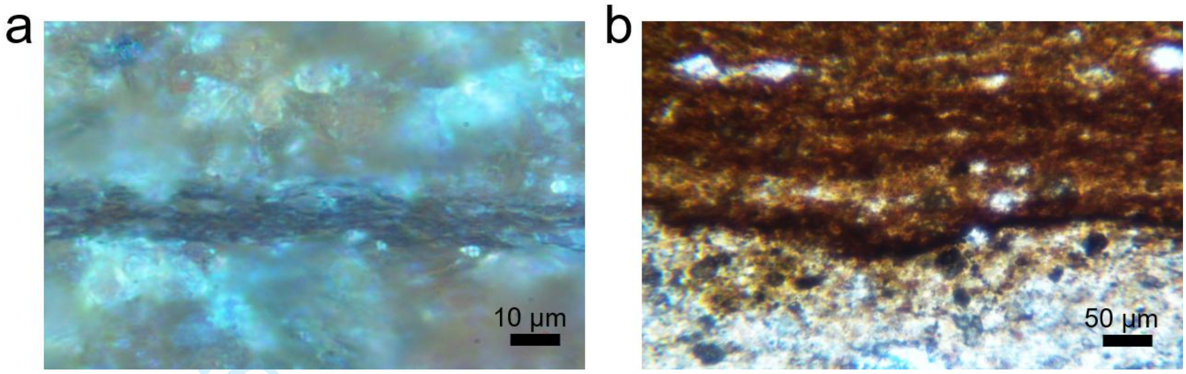
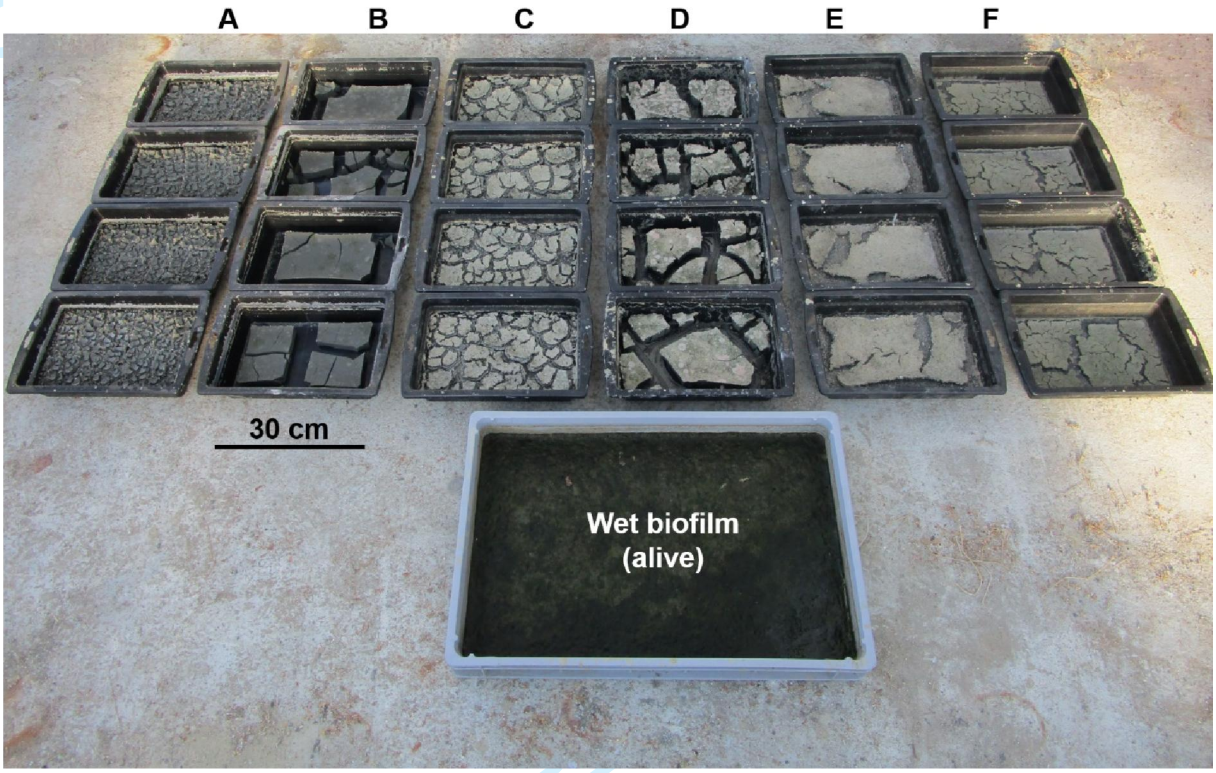


Figure 2



1067 Figure 3



1068

1069

Figure 4

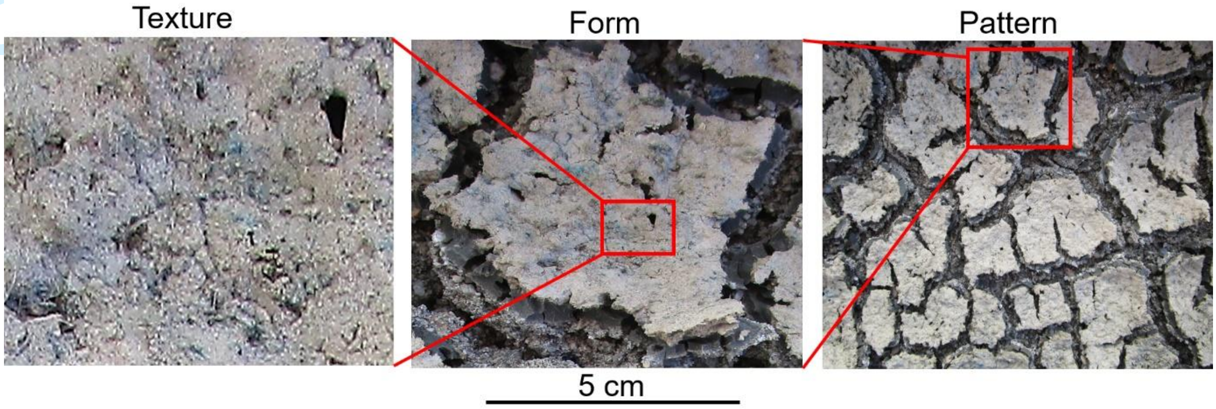


Figure 5

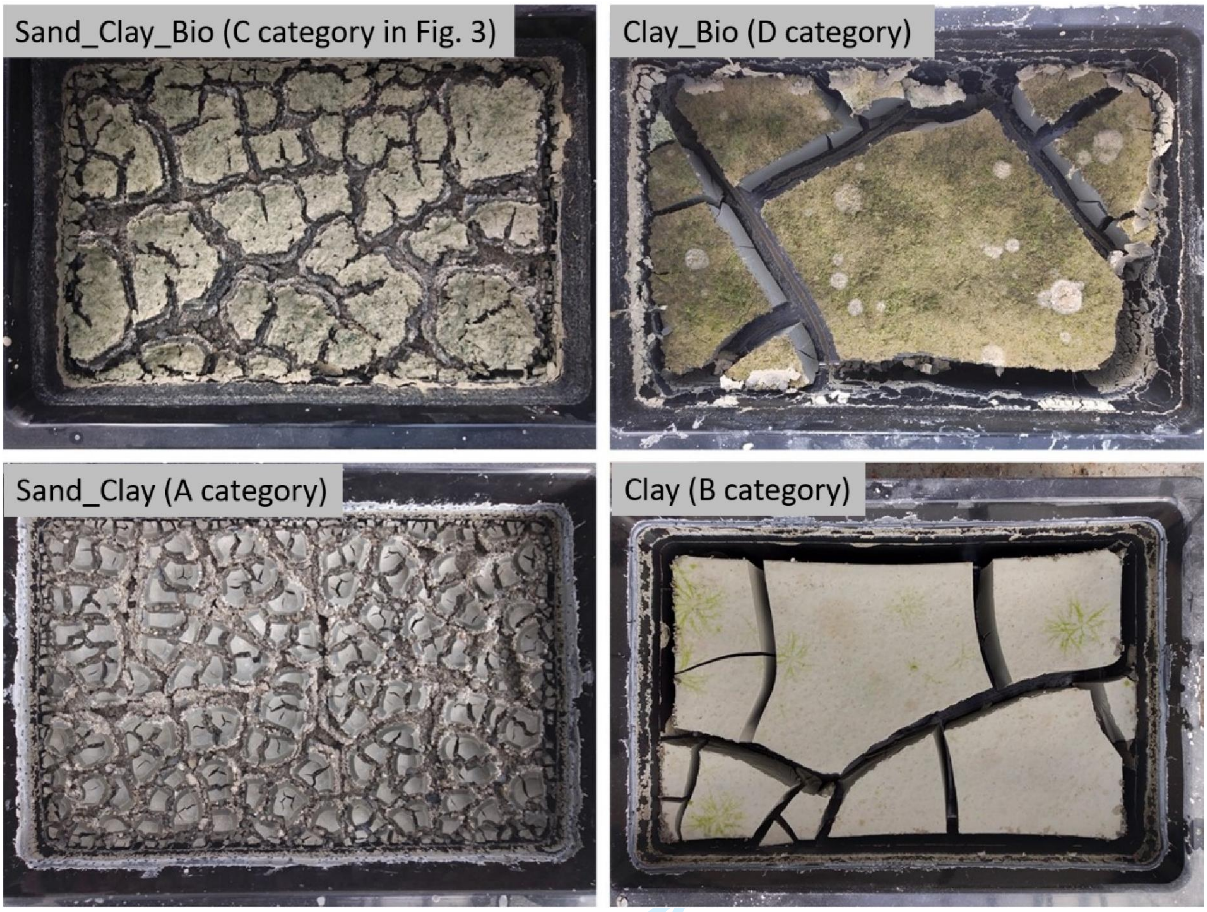
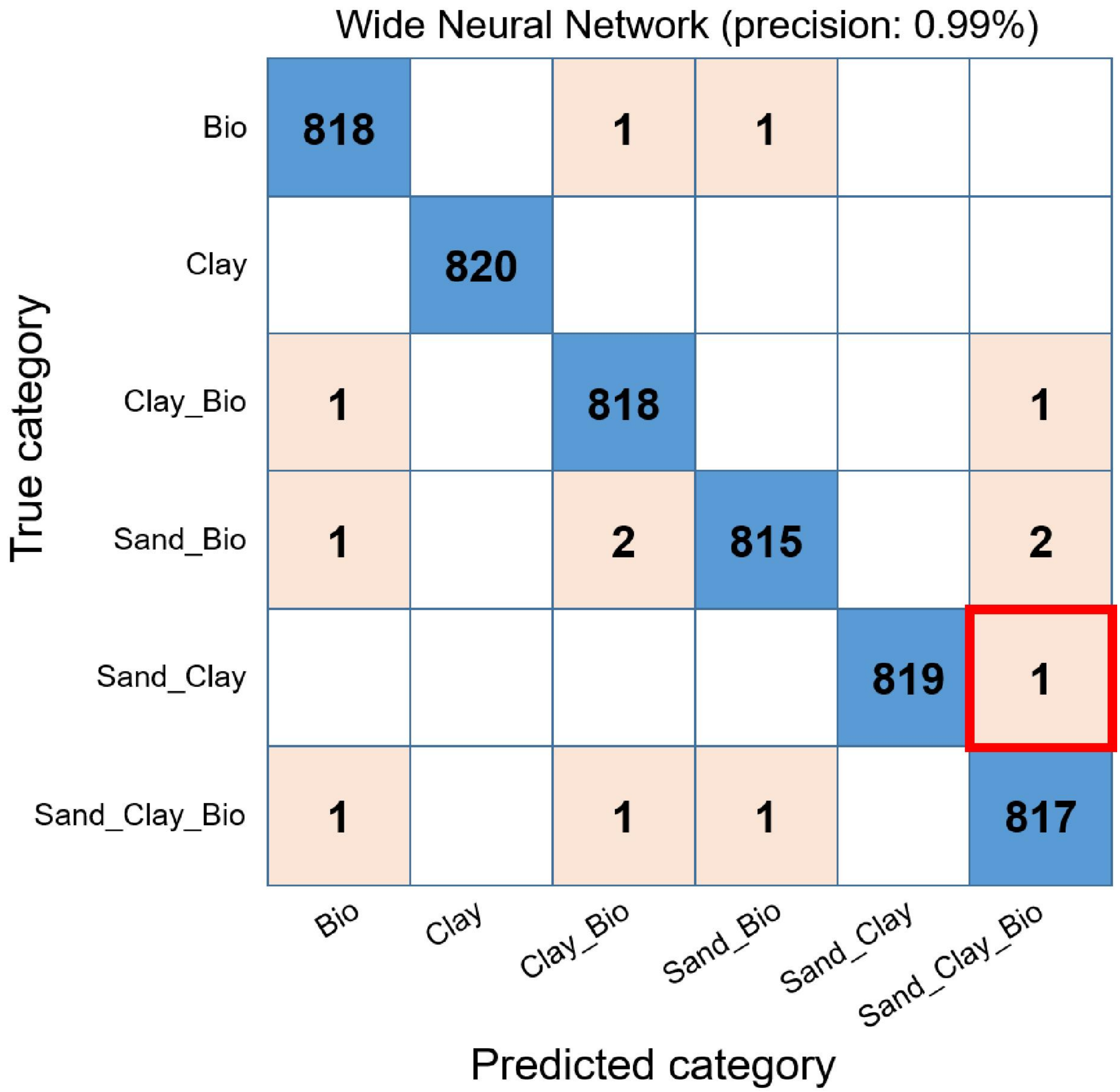
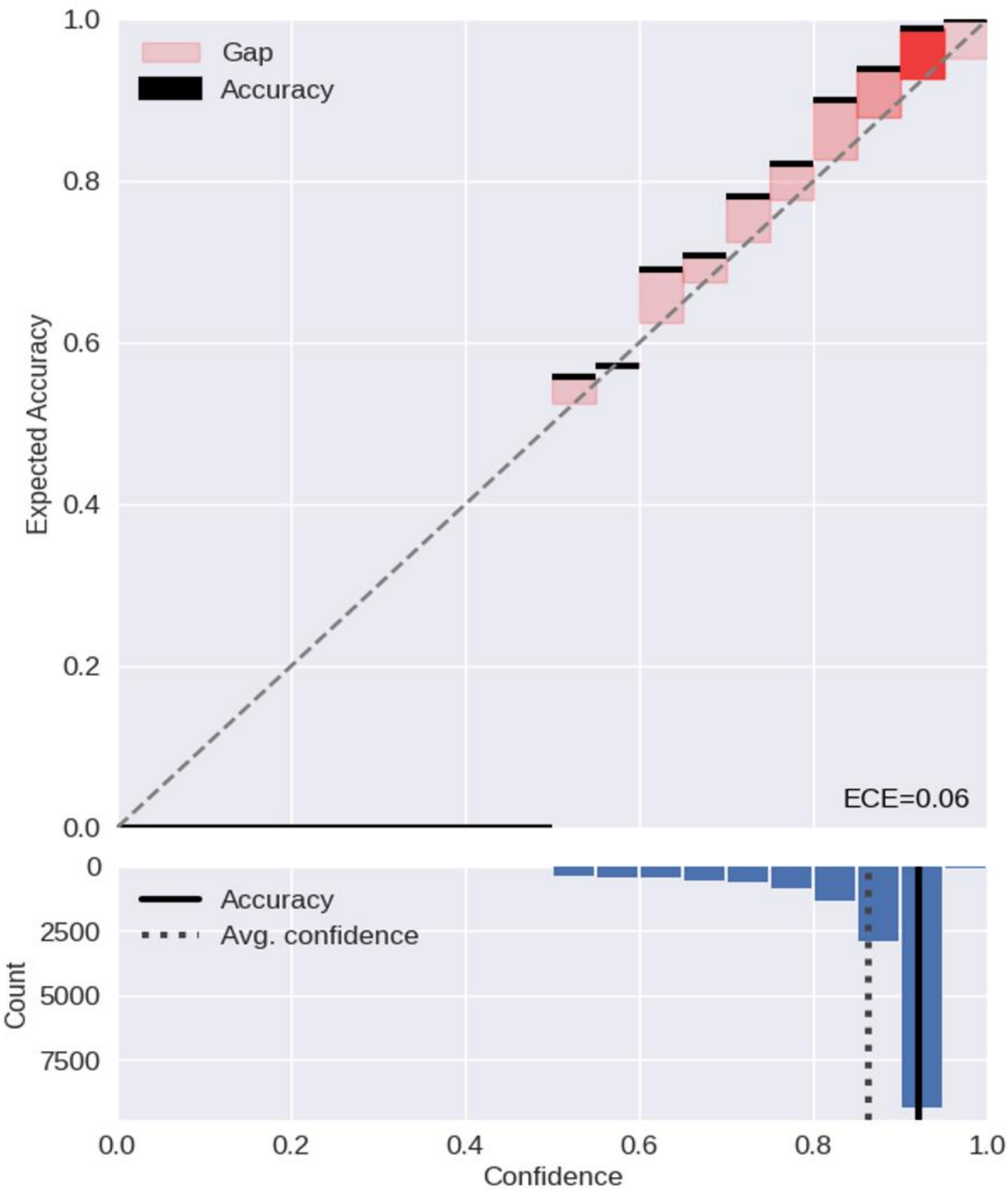

Sand_Clay_Bio (C category in Fig. 3)
Clay_Bio (D category)
Sand_Clay (A category)
Clay (B category)

Figure 6

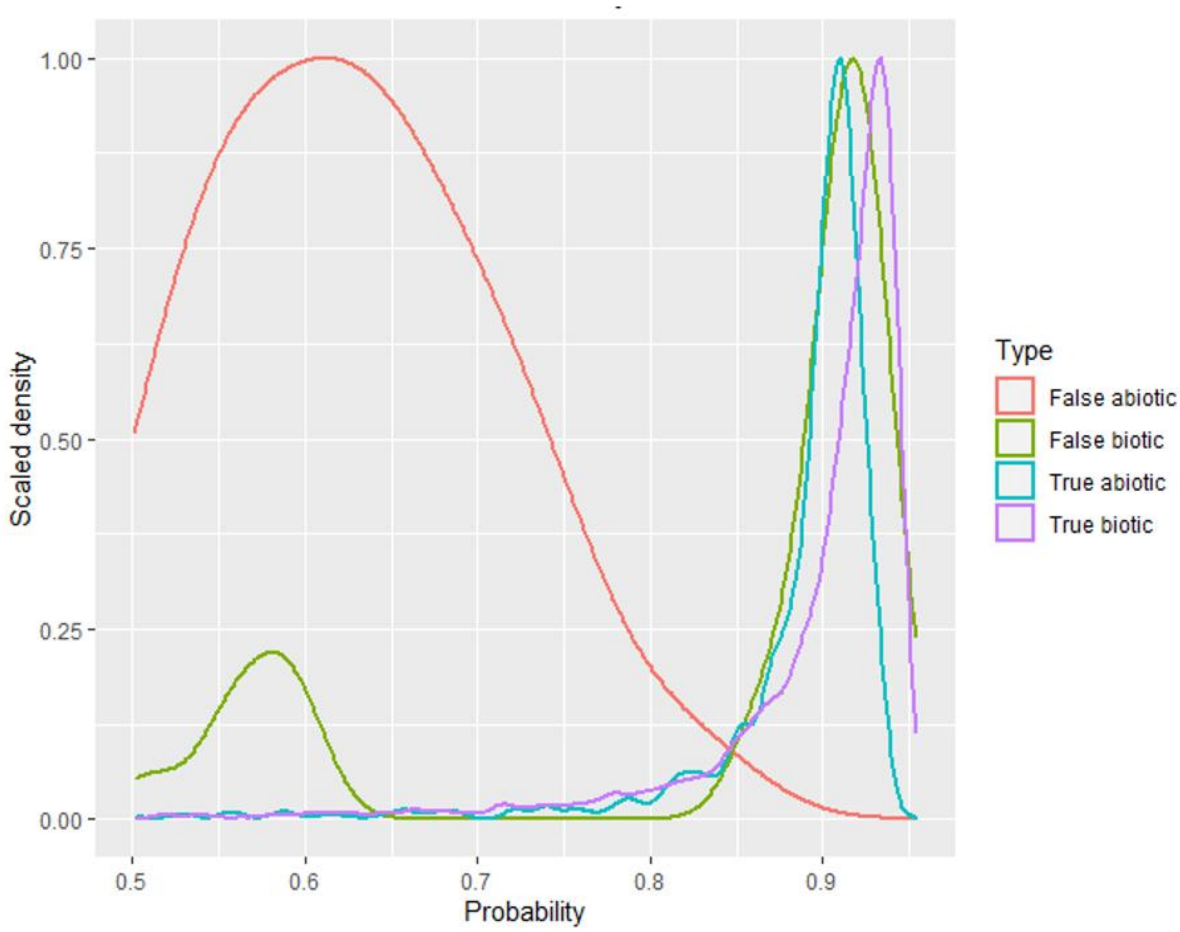


Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

Figure 7


Gap
Accuracy
ECE=0.06
Expected Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
Avg. confidence
Count
0
2500
5000
7500
0.0
0.2
0.4
0.6
0.8
1.0
Confidence

1082 Figure 8



1083

1084

Figure 9

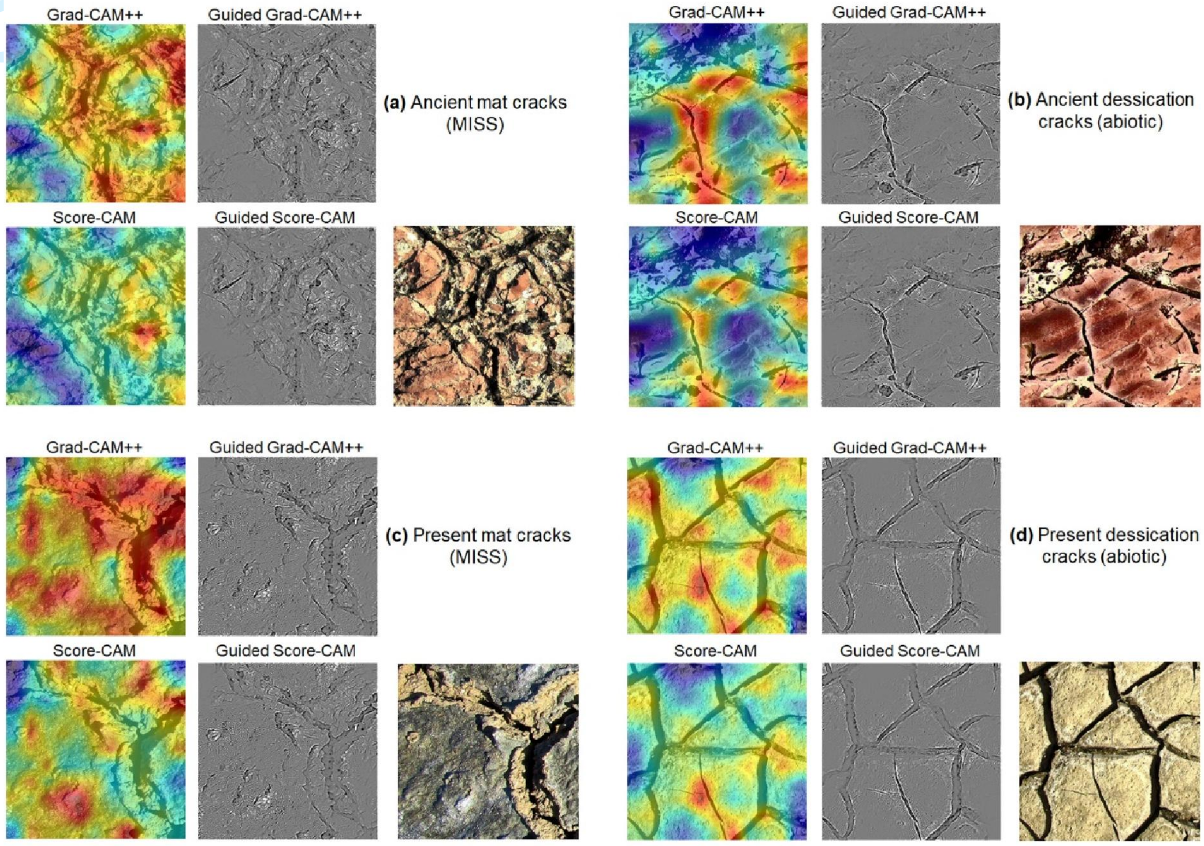
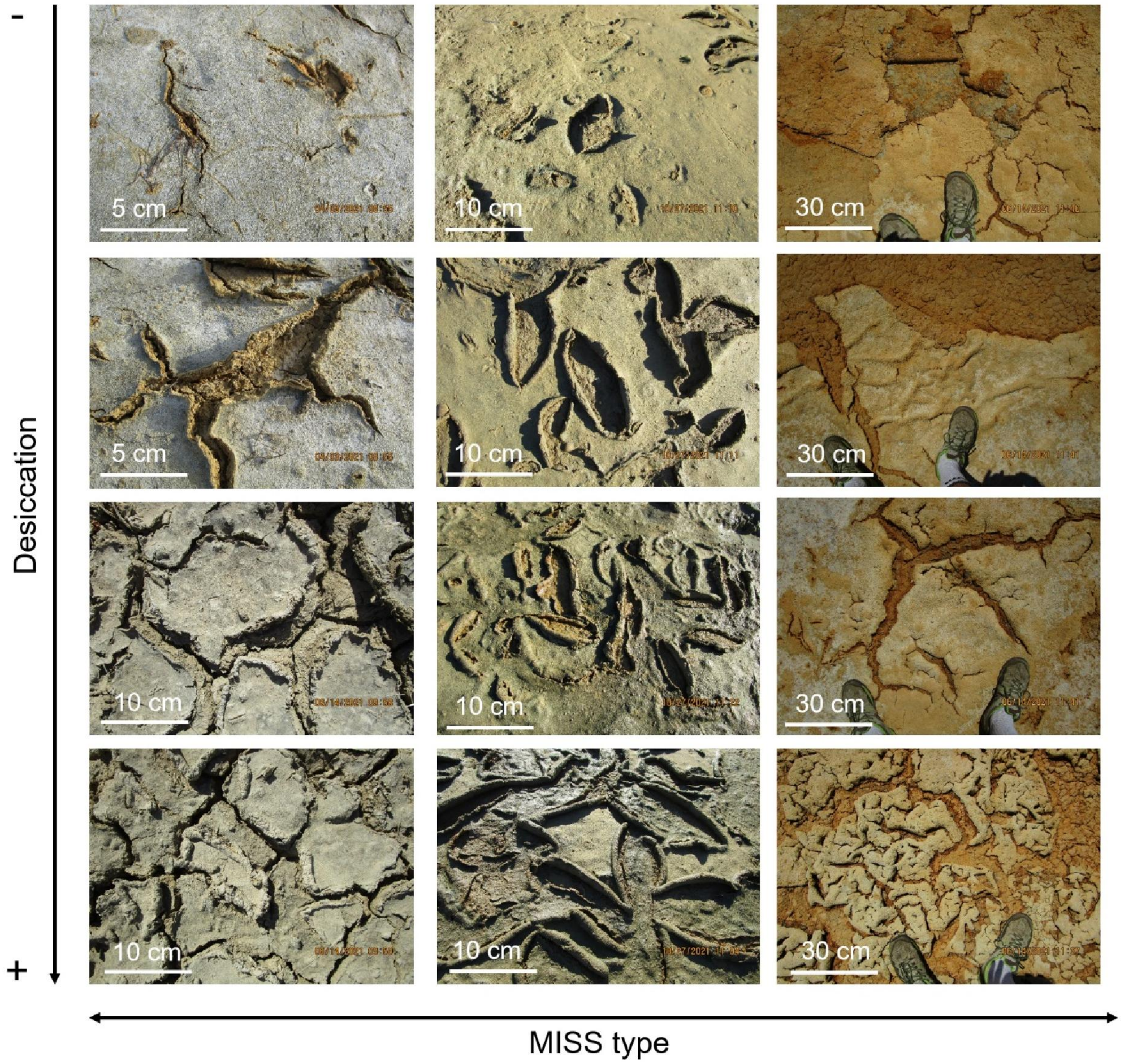
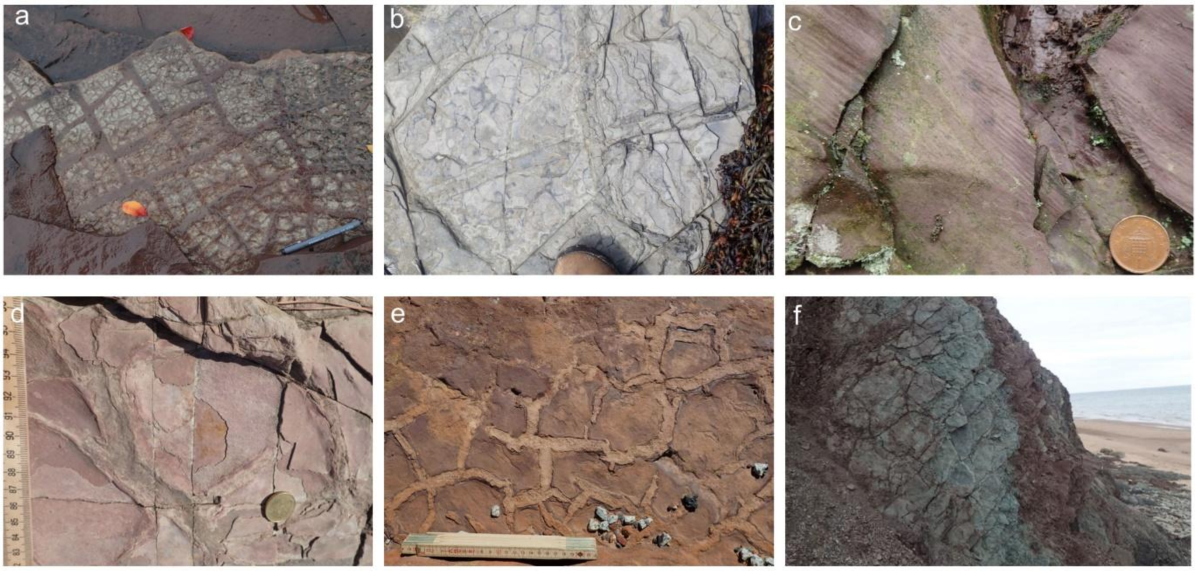

Grad-CAM++
Guided Grad-CAM++
(a) Ancient mat cracks (MISS)
Score-CAM
Guided Score-CAM
Grad-CAM++
Guided Grad-CAM++
(b) Ancient dessication cracks (abiotic)
Score-CAM
Guided Score-CAM
Grad-CAM++
Guided Grad-CAM++
(c) Present mat cracks (MISS)
Score-CAM
Guided Score-CAM
Grad-CAM++
Guided Grad-CAM++
(d) Present dessication cracks (abiotic)
Score-CAM
Guided Score-CAM

Figure 10


-
+
Desiccation
5 cm
5 cm
10 cm
10 cm
10 cm
10 cm
10 cm
10 cm
30 cm
30 cm
30 cm
30 cm
MISS type

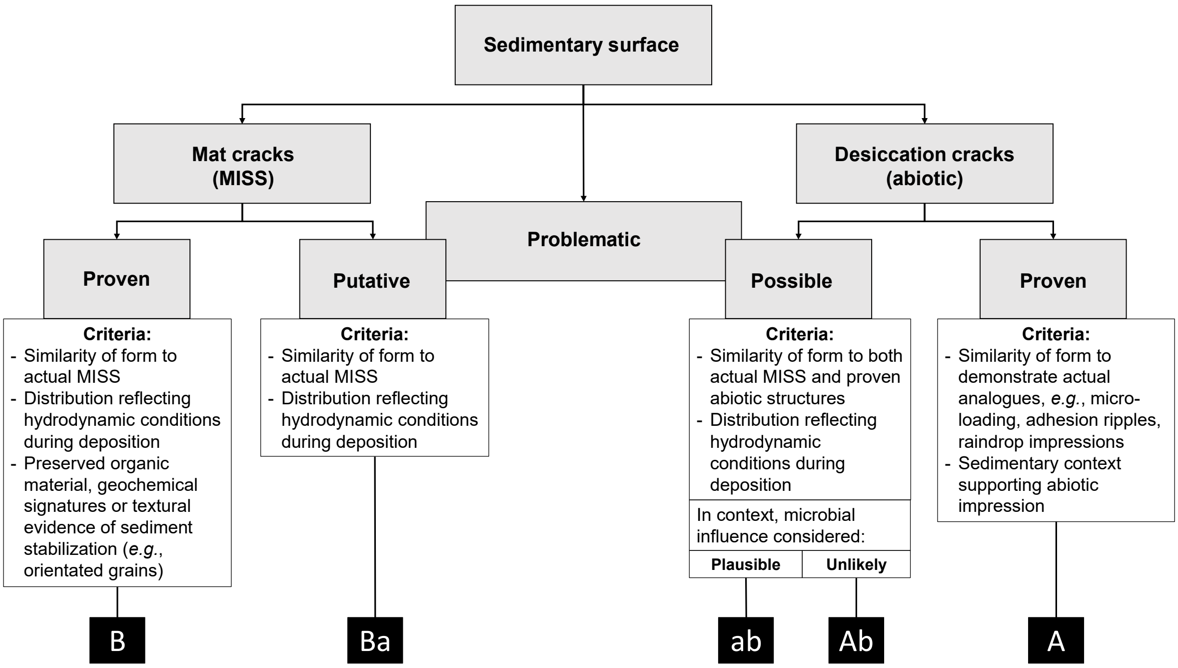
1091 Figure 11



1092

1093

1094 Figure 12



1095