

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/375671963>

# Automated identification of soil Fungi and Chromista through Convolutional Neural Networks ☆

Article in *Engineering Applications of Artificial Intelligence* · January 2024

DOI: 10.1016/j.engappai.2023.107333

CITATIONS

0

READS

27

5 authors, including:



**Karol Struniawski**

Warsaw University of Life Sciences - SGGW

9 PUBLICATIONS 7 CITATIONS

SEE PROFILE



**Ryszard Kozera**

Warsaw University of Life Sciences - SGGW

124 PUBLICATIONS 925 CITATIONS

SEE PROFILE



**Pawel Trzcinski**

Research Institute of Horticulture in Skierniewice

47 PUBLICATIONS 362 CITATIONS

SEE PROFILE



**Anna Lisek**

The National Institute of Horticulture Research

48 PUBLICATIONS 255 CITATIONS

SEE PROFILE



Contents lists available at ScienceDirect

# Engineering Applications of Artificial Intelligence

journal homepage: [www.elsevier.com/locate/engappai](http://www.elsevier.com/locate/engappai)

## Automated identification of soil Fungi and *Chromista* through Convolutional Neural Networks<sup>☆</sup>

Karol Struniawski<sup>a,\*</sup>, Ryszard Kozera<sup>a,b</sup>, Pawel Trzcinski<sup>c</sup>, Anna Lisek<sup>c</sup>, Lidia Sas Paszt<sup>c</sup>

<sup>a</sup> Institute of Information Technology, Warsaw University of Life Sciences - SGGW, ul. Nowoursynowska 159, Warsaw, 02-787, Mazowieckie, Poland

<sup>b</sup> School of Physics, Mathematics and Computing, The University of Western, Australia, 35 Stirling Highway, Perth, WA 6009, Western Australia, Australia

<sup>c</sup> Department of Microbiology and Rhizosphere, The National Institute of Horticultural Research, Pomologiczna 18, Skierniewice, 96-100, Lodzkie, Poland

### ARTICLE INFO

#### Keywords:

Soil microorganisms  
Agriculture  
Horticulture  
Machine vision  
Machine learning  
Convolutional neural networks

### ABSTRACT

The identification of soil microorganisms plays a crucial role in agriculture and horticulture, as it enables the monitoring of beneficial species and early detection of pathogens. In this study, we propose a system that utilizes machine vision and machine learning techniques, specifically Convolutional Neural Networks, to automate the identification of different fungi and *Chromista* based on microscopic images and morphological traits. Our system aims to provide a cost-effective and efficient method for pathogen detection, improving the overall health and productivity of agricultural systems. We conducted experiments using a dataset of soil microorganisms and the performance of the classifier was evaluated using precision, recall and F1-score measures. Despite challenges such as class imbalance and imperfect subimage retrieval, the classifier achieved promising results, with an overall precision of 82% indicating the high accuracy of correctly predicted positive instances across all classes. Furthermore, the incorporation of a majority voting scheme significantly improved the classifier's performance, addressing the issue of underrepresented classes. The enhanced results demonstrated an average precision and F1-score of 97%. Our work highlights the potential of CNNs in soil microorganism identification and paves the way for future research to expand the dataset and to incorporate a wider range of microorganism genera.

### 1. Introduction

The use of soil fungi as a natural fertilizer is a growing area of interest due to their potential in increasing crop yield and in replacing chemical fertilizers, which are still widely used in agriculture and horticulture (Han et al., 2021; Fan et al., 2023). Certain soil fungi and *Chromista* have a mutualistic relationship with plants, where they exchange nutrients with their host plant. These beneficial fungi, known as mycorrhizal fungi, play a crucial role in promoting the health and growth of many plant species (Smith and Read, 2010). Mycorrhizal fungi can enhance nutrient uptake, improve plant growth and increase resistance to environmental stressors e.g. drought and extreme temperatures. However, some soil fungi can be pathogenic and cause significant crop losses (Ma et al., 2013). Therefore, it is crucial to accurately identify soil microorganisms to monitor beneficial species and to detect pathogens early.

Microbial identification typically involves a combination of morphological, phenotypic and molecular methods. Morphological identification is the most cost-effective approach and involves microscopic observation of the microorganism, where a microbiology specialist identifies the organism based on characteristic fragments (Watanabe, 2010). However, identification based on morphological traits alone is limited to the genus level due to the subtle differences between species. Therefore, microbiologists typically use a combination of different methods for accurate identification.

In this study, we propose the use of machine vision and machine learning techniques, specifically Convolutional Neural Networks (CNN), to identify different fungi and *Chromista* based on microscopic images and their morphological traits. This system aims to automate the identification process, making it faster and more efficient. While the current system is limited to genus-level identification, it can be enhanced in future by incorporating phenotypic and molecular methods.

<sup>☆</sup> This research was supported by The National Centre for Research and Development, Poland within the framework of the project BIOSTRATEG, grant number BIOSTRATEG3/344433/16/NCBR/2018.

\* Corresponding author.

E-mail addresses: [karol.struniawski@sggw.edu.pl](mailto:karol.struniawski@sggw.edu.pl) (K. Struniawski), [ryszard\\_kozera@sggw.edu.pl](mailto:ryszard_kozera@sggw.edu.pl) (R. Kozera), [pawel.trzcinski@inhort.pl](mailto:pawel.trzcinski@inhort.pl) (P. Trzcinski), [anna.lisek@inhort.pl](mailto:anna.lisek@inhort.pl) (A. Lisek), [lidia.sas@inhort.pl](mailto:lidia.sas@inhort.pl) (L. Sas Paszt).

<https://doi.org/10.1016/j.engappai.2023.107333>

Received 16 June 2023; Received in revised form 22 August 2023; Accepted 17 October 2023

Available online 22 October 2023

0952-1976/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

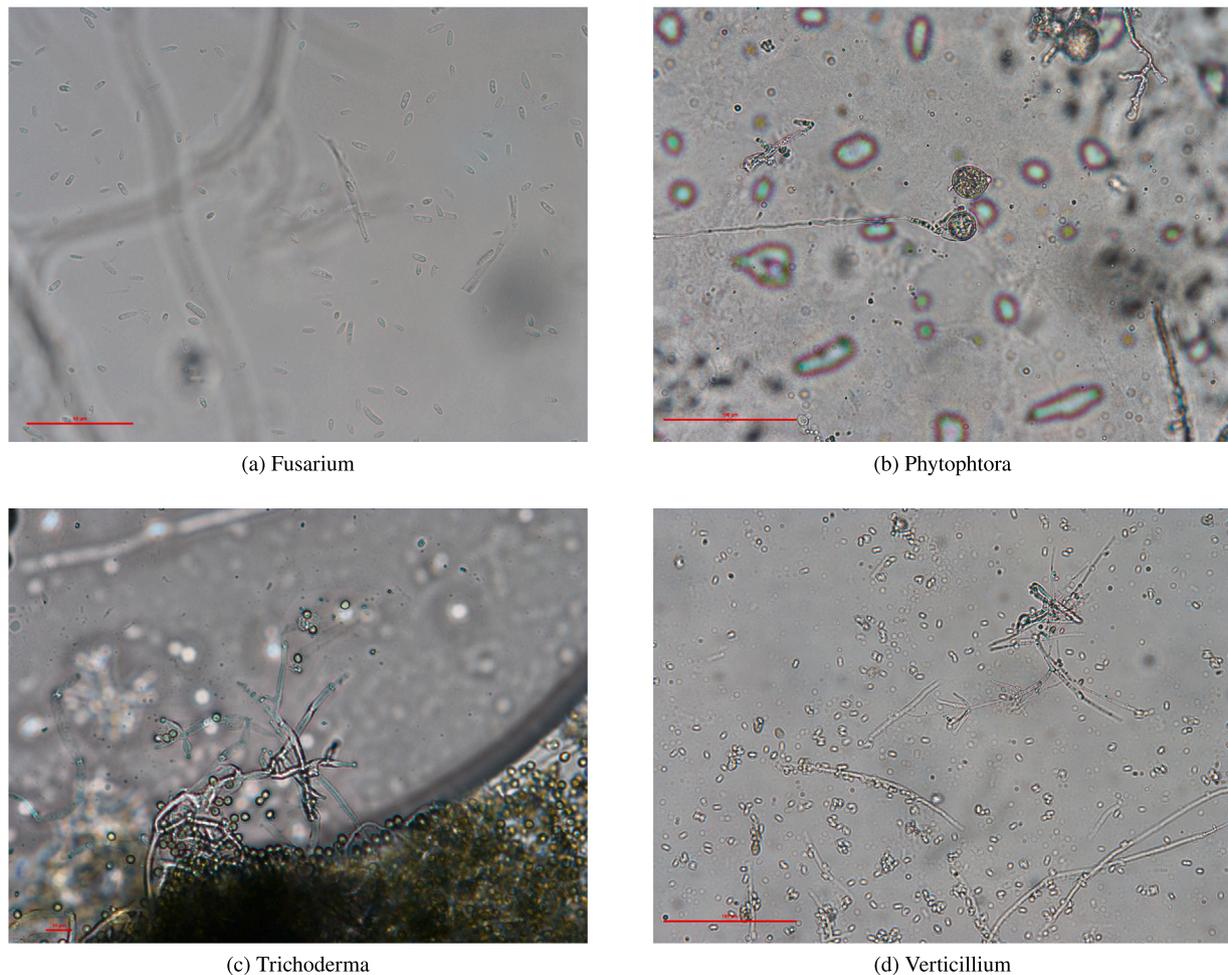
(a) *Fusarium*(b) *Phytophthora*(c) *Trichoderma*(d) *Verticillium*

Fig. 1. Microscopic images of cultivated samples obtained from the dataset sourced from *Symbio-Bank* - the microorganism collection at The National Institute of Horticultural Research in Skierniewice, Poland.

Such integration would enable identification at the species level, albeit at a higher cost.

The main objective of our system is to provide in future highly accurate results that can be utilized by farmers to quickly detect potential pathogens and to supply solutions that can be implemented immediately upon the soil's sample delivery to local laboratories. With its low-cost laboratory equipment and full automation process such system becomes highly desirable for the agricultural industry worldwide. System's ability to quickly detect potential threats can minimize the risk of crop losses and can improve the overall productivity of agricultural setups. This aspect is particularly significant taking into consideration recommendations of the United States Environmental Protection Agency that emphasize the importance of implementing methods to enhance crop resistance and to employ intelligent systems addressing the impacts of climate change on agriculture and food supply (U.S. Environmental Protection Agency, 2023). The detection of Mycorrhizal microorganisms enables precise tracking of natural fertilizer application, monitoring their quantity and effects on the soil. This paper describes our initial work in constructing such a system using CNNs and we believe that the system described above represents the ultimate goal for the future.

Identification of soil fungi and *Chromista* using conventional image processing methods (Gonzalez and Woods, 2018) poses significant challenges due to their intricate and diverse structures, which exhibit non-homogeneity unlike soil bacteria (Watanabe, 2010). These structures encompass various components such as hyphae, phialides, micro and macroconidia, conidia cells, zoospores and more. To address the

latter our research explores a dataset sourced from *Symbio-Bank*, which houses a collection of microorganisms at The National Institute of Horticultural Research in Skierniewice, Poland. Within this collection, we specifically select, cultivate and capture microscopic images of microorganisms belonging to the genera *Fusarium*, *Trichoderma*, *Verticillium* and *Chromista* of the *Phytophthora* genus. It is important to note that each image in the dataset was assumed to contain only one type of microorganism, obtained through a rectification process that isolates the monoculture from the sample. The dataset comprises 134 images of microorganisms cultivated in Potato Dextrose agar (Merck 1.10130.0500) at a temperature of 26 degrees Celsius. Incubation periods varied, with *Fusarium*, *Verticillium* and *Phytophthora* microorganisms requiring 7–10 days, while the *Trichoderma* strain took approximately 4–5 days to produce conidial spores.

The biomass was collected from the colonies using sterile swabs and spread over the microscope slide. A small amount of distilled water was added and the material was covered with a coverslip. The images were taken using a Nikon 80i microscope and no manual post-processing was performed. Therefore, the images may contain artifacts like light reflections or air bubbles under the coverslip, which could resemble the microorganism's structures. The dataset comprised 57 *Fusarium*, 26 *Trichoderma*, 31 *Verticillium* and 20 *Phytophthora* images, which were obtained with different magnification levels for each of the microorganism genera. Fig. 1 presents sample images from the dataset. A portion of the whole dataset is provided as supplementary material (Struniewicz, 2023).

The work starts with the conducting a literature review to identify the current state-of-the-art solutions in the area in question and to determine the direction addressed in our experiments.

Wahid et al. (2018) presented a transfer learning approach where the Inception CNN was retrained using a dataset consisting of five different bacteria species microscopic images known to be harmful to human health. Increasing the total number of samples, multiple bacteria samples were extracted from each image through a manual cropping. However, a major limitation of the system was its reliance on human involvement in the manual input image processing, which rendered the system as a hybrid and not fully automated, thus making the task more laborious.

Zhang et al. (2017) utilized CNNs to cervical cell classification. The proposed approach involves segmenting single cell instances from the original image, which can be challenging due to the location of the nucleus within the cell not always being in the center position. In the case of images containing only cells and background, the task involves pre-segmentation by retrieving subimages based on the observable nucleus within the cell boundaries. These subimage patches are then exploited to train the CNN for cervical cell classification without the need for accurate cell segmentation, as the CNN with a large dataset is recognized as a better generalizer even with inaccurate segmentation. This approach can potentially yield better robustness of the classifier. The limitation of the system is its strict dependency on a coarse nucleus center being provided. To overcome this limitation, the researchers employed transfer learning by the pre-trained CNN on the ImageNet.

Upon examining the research conducted by Ong et al. (2020), it was observed that they employed transfer learning with the ResNet50 CNN model to detect *Fusarium* fungi in laboratory-cultivated samples. The images were captured using a microscope, utilizing three distinct microscopy configurations: Brightfield, Darkfield and Fluorescent, where each configuration provided unique image characteristics. The authors trained three separate CNN models, where each model is specialized in classifying one of the three image modes. To classify the different modes, the authors employed the respective ResNet50 model and a logistic regression classifier was utilized to obtain the final prediction. This study represents a rare example of using CNNs for soil fungi identification, focusing on binary classification for detecting the presence of *Fusarium* in the images. In contrast, our research aims to differentiate various fungi and *Chromista* genera from one another. Noteworthy, the study by Ong et al. employed a manual segmentation method, while our objective is to develop a fully automated system.

## 2. Methods

The paper provides a comprehensive description of the subimage retrieval process and the application of CNN for the classification task. The subimage retrieval process is elaborated upon, highlighting the methodology employed to extract relevant fragments from the original images. Then, the paper delves into the CNN approach used for classifying the retrieved subimages, discussing the architecture and techniques employed to achieve accurate classification results.

### 2.1. Subimages retrieval

The microscopic images used in our research contain scale labels in the left-down corner, which are added automatically by the microscope's software. These labels come in three different colors – black, red, or white – selected based on image metrics. However, the presence of these scale labels may introduce distortions during the subimage retrieval process, ultimately impacting performance of the predictor. To mitigate this issue, we masked the label's area using manual thresholding techniques (Gonzalez and Woods, 2018). The threshold values were determined empirically and the mask's area was defined as the set of pixels meeting the following conditions:  $r < 30$  and  $g < 30$  and  $b < 30$  (black label), or  $r > 150$  and  $g < 125$  and  $b <$

125 (red label), or  $r > 240$  and  $g > 240$  and  $b > 240$  (white label), where  $r$ ,  $g$  and  $b$  denote the red, green and blue components of the RGB pixel  $I = [r, g, b]$ , respectively. The resulting binary mask identifies the label area with a value of 1, while the rest of the image is assigned a value of 0. Finally, the label's area is filled with the mean color of the image and merged with the input image to create a new, label-free image for analysis.

In our study, we present an image preprocessing pipeline for enhancing the quality of microscopic images to facilitate subsequent analysis. First, we converted the RGB image to grayscale using the COLOR\_BGR2GRAY mode provided by the popular cv2 package in Python. To improve the quality of the image, we applied the widely-used Gaussian blur technique (Gonzalez and Woods, 2018), which effectively smooths out small-scale variations in the image while preserving larger-scale features. A Gaussian kernel with a window size of  $5 \times 5$  was used in our approach, which has been shown to be an effective choice for reducing image noise while preserving important features (Párraga et al., 1998).

Microscopic images often suffer from low contrast, which can make it difficult to discern the relevant features and structures in the image. To address this, a variety of contrast enhancement techniques have been proposed in the literature, which modify the spatial characteristics of the image to improve its visibility and facilitate image analysis (Cakir et al., 2018). In our work, we employ the Contrast-limited Adaptive Histogram Equalization (CLAHE) method, which has been shown to be effective in enhancing the visibility of structures in low-contrast images. This method has been successfully applied in a variety of applications, such as improving the accuracy of mammographic image classification (Alshamrani et al., 2023) and detecting pneumonia using CNNs (Tjoa et al., 2022). By applying CLAHE to our microscopic images, we aim to improve the performance of our machine learning-based subimage retrieval method by increasing the distinguishability between different fungi structures. On the grayscale image obtained in the previous step the CLAHE method is applied using a clip limit of 0.1 and a tile grid size of  $8 \times 8$ . The clip limit is a threshold value that limits the maximum pixel intensity value in the histogram, while the tile grid size refers to the size of the tile used for local histogram equalization that affects the level of local contrast enhancement.

Segmentation of microscope images is a challenging task due to uneven illumination, lightning variations and noise (Pham et al., 2000). Global thresholding techniques like the Otsu method (Otsu, 1979) are commonly used for segmentation, but they may produce unreliable results when the image is complex and its histogram cannot be segmented by a single threshold value. Our literature review reveals that local adaptive thresholding techniques, such as the one proposed by Wen-Nung Lie and other researchers (Lie, 1995), have been found to be superior to global thresholding techniques for segmenting microscopic images (Khan et al., 2015; Dave and Upla, 2017). Therefore, we applied the Adaptive Image Thresholding algorithm, implemented in Python's cv2 adaptiveThreshold function, with a 9-pixel block size that determines the size of the neighborhood area and a constant of  $C = 213$ . The threshold value is a Gaussian-weighted sum of the neighborhood values minus the constant C. The adaptive threshold method creates a matrix of thresholds equal to the size of the input image that are based on the local neighborhood of the image pixels. The details of implementation can be found in the original paper (Lie, 1995).

It is important to highlight the significant challenges associated with soil microorganism images, including low contrast, out-of-focus objects, sample contamination and image impurities such as light reflection from the coverslip. It should be noted once the sample is observed under a microscope, the microorganisms exist in a three-dimensional space and the focus of the camera is limited to a specific level. As a result, objects located at different depths may appear out of focus. To tackle these challenges and further enhance the quality of the images, we incorporate morphological operations and filling algorithms into our pipeline. This approach aligns with other studies in the field, which

involve converting microscopic images to grayscale, preprocessing, segmenting and employing morphological operations to eliminate small objects or fill gaps within objects (Mohamad et al., 2014). By applying these techniques, we aim to improve image quality and ensure accurate analysis of the microorganisms in our research.

In our system, we begin by dilating the images using an  $11 \times 11$  kernel, followed by applying a flood fill algorithm and subsequently performing open and close operations with a  $5 \times 5$  kernel. We then perform a second flood fill and obtain the final image, from which sub-images can be retrieved in subsequent steps. The algorithm used in our study is based on the implementation described in Soille (1999).

In the subsequent stage, the image is labeled by identifying all objects, also referred to as connected components, using the label function from the `skimage` module. Two pixels are considered connected if they are neighbors and have the same value. It is important to note that we are still operating on the binary image mask, where the pixels can be neighbors in a 1- or 2-connected sense and the value denotes the maximum number of orthogonal hops to consider a pixel as a neighbor (Fiorio and Gustedt, 1996). Consequently, the labeled image is obtained with  $k$  objects identified, where the image comprises of pixels with values of either 0, representing the background, or 1 up to  $k$ , identified as the objects. Next, utilizing the `regionprops` function from Python's `scikit-image` module, we are able to compute the parameters of each object. Here, we used the solidity, bounding box and area measures. Solidity is the ratio of the region's area to the area of its convex hull, which is the smallest convex polygon that can contain the region (Ilonen et al., 2018). The bounding box is the smallest box that contains the entire object, defined by its pixel coordinates on the image and the area is defined as the number of pixels in the region. It is noteworthy that solidity, along with other region properties, is recognized as an accurate descriptor of objects in an image (Vinnett et al., 2022; Acuña et al., 2016; Bailey et al., 2005). The output of the labeling and region properties calculation is a list of all objects in the image with assigned measures.

In the next stage, our objective was to accurately select objects that are fragments of microorganisms, which posed a difficult challenge due to the heterogeneity of soil fungi and *Chromista* organisms, which cannot be identified based on a single cell but instead rely on specific morphological features such as conidiophores, sporodochia, pycnidia, sclerotia, mycelia, etc. Wijayawardene et al. (2020). Retrieving single bacterial cells in microscopic images is straightforward since they are represented by uniformly sized circular cells. Consequently, recent works focus on filtering out objects with solidity lower than a predetermined threshold and retrieving  $n$  objects closest to the median area of the objects (Struniawski et al., 2022). However, this approach did not yield satisfactory results in our dataset. After conducting several experiments, we propose a pipeline that specifies a padding parameter  $p$  (in our case,  $p = 3$ ) to address the issue of losing crucial information on object edges during processing images phase. To accomplish this goal, we remove all objects that would exceed the image's size after padding and those with solidity lower than 0.8. We also leave out images with standard deviation of pixel luminosity less than 5 to eliminate low contrast objects, which could be mistaken for background or out-of-focus fragments.

The next challenge is encountered when the number of retrieved objects is significantly less than the maximum target number of instances (200 in our case), usually due to uneven lighting conditions or the entire image being out of focus. In such instances, we avoid filtration based on solidity and instead selecting 100 objects with the highest solidity value. Furthermore, we sort objects based on the covered area and select a maximum of 200 objects with the largest area. Once the filtering process is complete, we retrieve single images by cutting the area around the object's bounding box with additional padding. Next we resize the sub-image's mask to match its subimage size by eroding it with a kernel of size  $(2p + 1) \times (2p + 1)$  to mitigate the effect of the enlarged boundaries that may contain background information.

This erosion operation helps to remove the background information consistently throughout the boundary of the object. Finally, we filter the object's final image by taking the Hadamard product (Szeliski, 2011) of the sub-image with its corresponding mask, followed by placing such image in the center of  $224 \times 224$  image with all pixel values set to 0 (black image). Employing the Hadamard product of the object with its corresponding mask alleviates the impact of overlapping objects that may arise following the retrieval process from the original image. This strategic approach guarantees that each subimage exclusively encompasses a single object, a direct outcome of the labeling procedure applied to the input image. The result is a black background image with the color object in the center. It is worth noting that in literature, better classification performance is achieved by using color object images with a black background, which eliminates irrelevant background information and reduces unnecessary distortions (Konopka et al., 2022; Struniawski et al., 2022; Gao et al., 2017). The data flow of the algorithm is presented in Fig. 2.

The generalization abilities of the system forms a significant challenge. Although the aforementioned algorithm demonstrates accurate subimage retrieval for certain genera, its performance may be suboptimal for others. Our studies have revealed that setting the threshold value of solidity filtration to 0.9, instead of 0.8, led to unsatisfactory outcomes of sub-images generation for most images of *Phytophthora* genera. Consequently, the filtration conditions must be as weak as possible to increase the robustness of sub-image computations. However, this approach may result in sub-images that are not actually fragments of microorganisms, but represent in fact noise or background. To address this issue, we have incorporated a majority voting rule after classification, which is further discussed in details (see Section 2.4.2).

After analyzing the generated dataset, we can observe that the system achieves accurate retrieval of subimages for all the examined genera. The filtration process parameters, including image processing kernel sizes and shapes, threshold values, padding and the maximum number of retrieved objects, can be adjusted using a new dataset of soil microorganisms. The stepwise retrieval process is illustrated in Fig. 3, while Fig. 4 showcases exemplary retrieved sub-images for each genus. Additional subimages can be found in the supplementary materials (Struniawski, 2023). The presented system signifies a significant progress towards fully automated analysis of soil microorganisms. It is important to note that the sizes of objects in the images vary due to differences in magnification levels employed during image acquisition. The dataset utilized for training Machine Learning methods for single instance retrieval comprises a total of 10,175 images, including 4278 images of *Fusarium*, 1285 images of *Phytophthora*, 2352 images of *Trichoderma* and 2260 images of *Verticillium*.

## 2.2. ResNet50 neural network

In our research, we have utilized Residual Network 50 (ResNet50) as a Machine Learning classification model. ResNet50 is a type of CNN that has demonstrated outstanding performance across various domains. Unlike traditional machine learning methods, which rely on manually crafting features in advance (called also as handcrafted features) and feeding them into a chosen classifier such as Support Vector Machines, Random Forests, or Multilayer Perceptron (Hanbal et al., 2020).

In classical ML approaches, the manual determination of features requires domain experts to carefully consider the task's characteristics and evaluate the selected features. This process often leads to suboptimal accuracy, if chosen features are believed to adequately represent the observed phenomenon. It typically involves multiple iterations of feature determination, followed by the application of feature selection methods to form a final set of features. These selection methods aim to choose or filter features that exhibit strong correlation with the target variable and weak correlation with each other. In contrast, features can be automatically calculated from the data using Convolutional

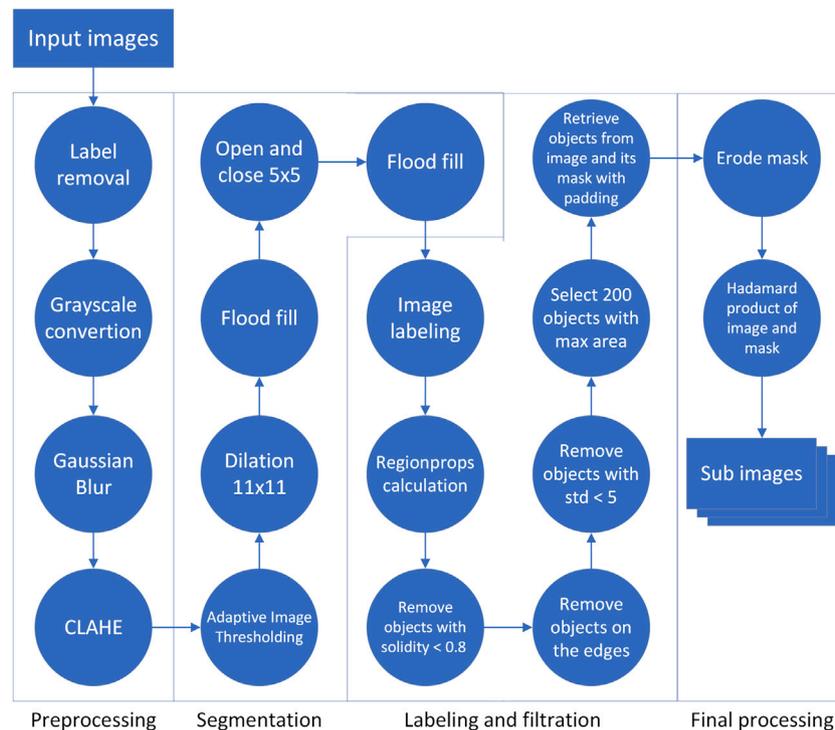


Fig. 2. The data flow diagram illustrating the proposed algorithm for sub-images retrieval.

Neural Networks. CNNs employ convolutional layers with learnable filters to convolve input data, capturing relevant patterns and spatial relationships automatically (Yamashita et al., 2018). For image data, 2D convolutional layers are commonly used, resulting in 2D feature maps. CNNs also incorporate pooling layers, which downsample feature maps and promote translation invariance. This characteristic reduces the network's sensitivity to the precise spatial location of features.

CNNs consist of multiple building blocks, including convolution layers, pooling layers and, in the final block, feature maps are flattened into vectors. This process converts each observation in the dataset into a vector of values that can be recognized as features. Typically, a fully connected (dense) feedforward network with a backpropagation algorithm is employed to train the weights within the dense network, as well as the filter coefficients in the convolutional layers. The training process minimizes a given loss function, enabling the network to learn discriminative representations from the data. The depth of a CNN plays a crucial role in its performance on various tasks and significant advancements have been achieved by employing very deep models (Simonyan and Zisserman, 2015; He et al., 2015). However, increasing the depth of the network by stacking more blocks and expanding the number of layers introduces challenges such as dealing with vanishing or exploding gradient and degradation problems. The vanishing and exploding gradient issues can be addressed through normalized training initialization techniques (LeCun et al., 1998). On the other hand, the degradation problem arises when adding extra layers to a deep architecture leads to higher training error.

To overcome these challenges ResNets have been introduced (He et al., 2016). They utilize residual connections, also referred to as skip or shortcut connections that allow the network to bypass one or more convolutional layers and directly add the original input to the output of these layers. By incorporating residual connections, ResNets mitigate the degradation problem and enable more effective training.

In this work, we utilize ResNet50, a variant of the ResNet architecture, which has shown remarkable performance in similar tasks (Ong et al., 2020). The ResNet50 implementation employed in this study is provided by the TensorFlow module. ResNet50 is composed of a total of 50 layers (this is the root of 50 in its name), with 48 convolutional

layers, one max pool layer and one average pool layer. The architecture consists of five residual blocks, each containing nested convolutional layers, which can be viewed as a combination of convolution, batch normalization and activation operations. While it is not feasible to depict the entire ResNet architecture within the constraints of this paper, we have included it as Supplementary Material (Struniawski, 2023). Fig. 5 showcases a sample of the second residual block with the third convolutional block, where all the aforementioned operations with residual shortcut connection, are visible.

### 2.3. Transfer learning

Transfer learning is a popular approach in Machine Learning where a model developed for one task is utilized for another, transferring the knowledge gained from the first task to the second one. The idea of transfer learning dates back to 1976 when it was first applied to the perceptron algorithm (Bozinovski and Fulgosi, 1976). CNNs have become increasingly complex, leading to resource-intensive training and transfer learning has become a widely used technique to address this issue. Huge CNNs are typically trained on high-performance computing systems such as clusters, supercomputers, or cloud-based platforms that are optimized for deep learning workloads, with multiple GPUs that parallelize computations and speed up the training process (Huerta et al., 2020). For instance, some of the largest CNN models, including VGG-16, ResNet and Inception, were trained on the ImageNet Large Scale Visual Recognition Challenge dataset, which comprises 1.2 million high-resolution images and took several days to weeks to train. Training a large CNN model from scratch is usually infeasible due to the long training time and high cost. Therefore, we utilized the ResNet50, which had been previously trained on the ImageNet dataset (Deng et al., 2009).

In spite of the fundamental disparity between the ImageNet dataset and microscopic or X-ray images, transfer learning has demonstrated remarkable efficacy in diverse applications, including the classification of COVID-19 from chest X-ray images (Hossain et al., 2022) and the classification of Malaria Cell-Images (Arrabally and Juliet, 2019). The utilization of the extensive knowledge amassed from the ImageNet

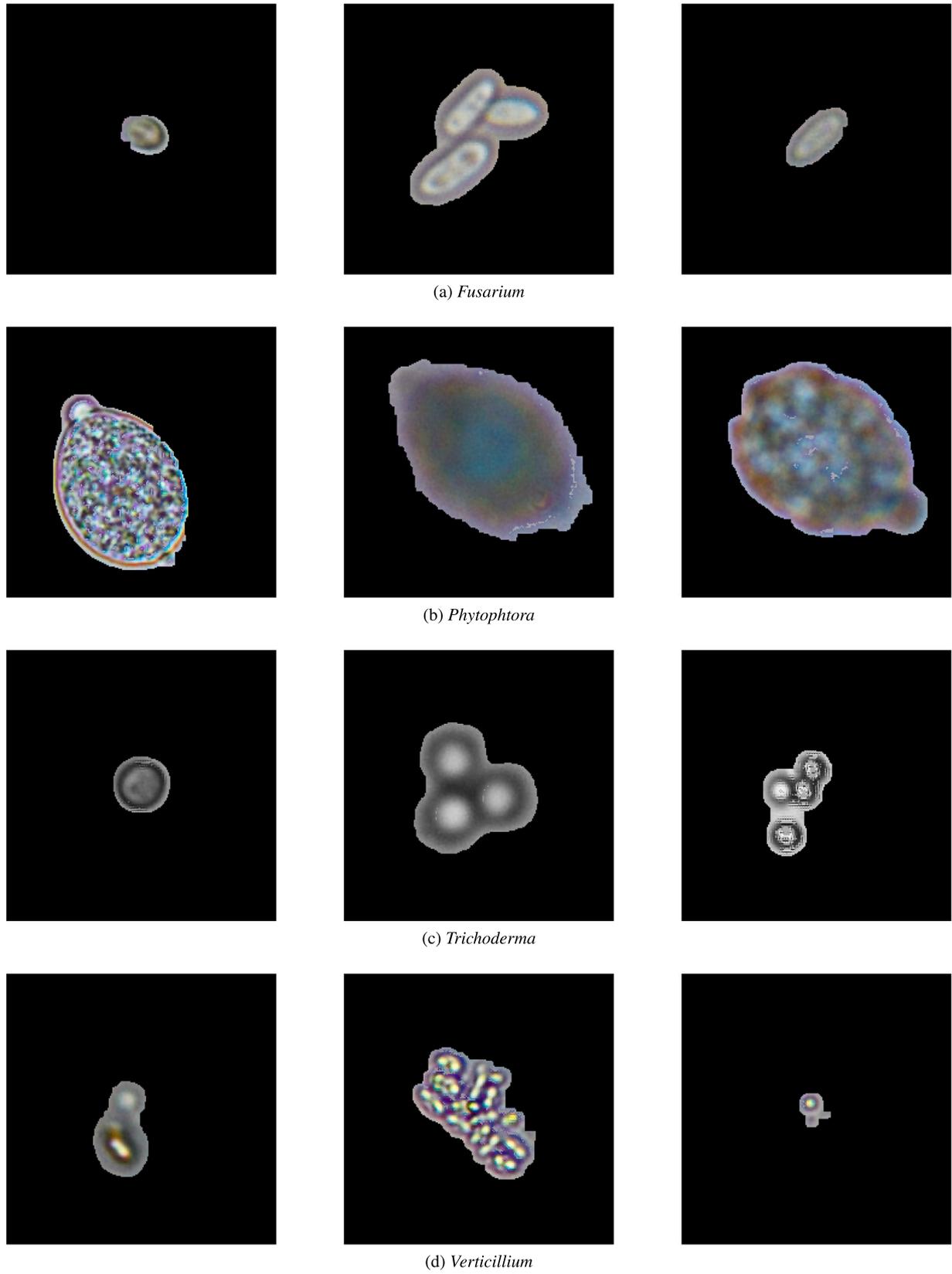


Fig. 3. The sample sub images retrieved from the original dataset from a given genera of soil microorganism.

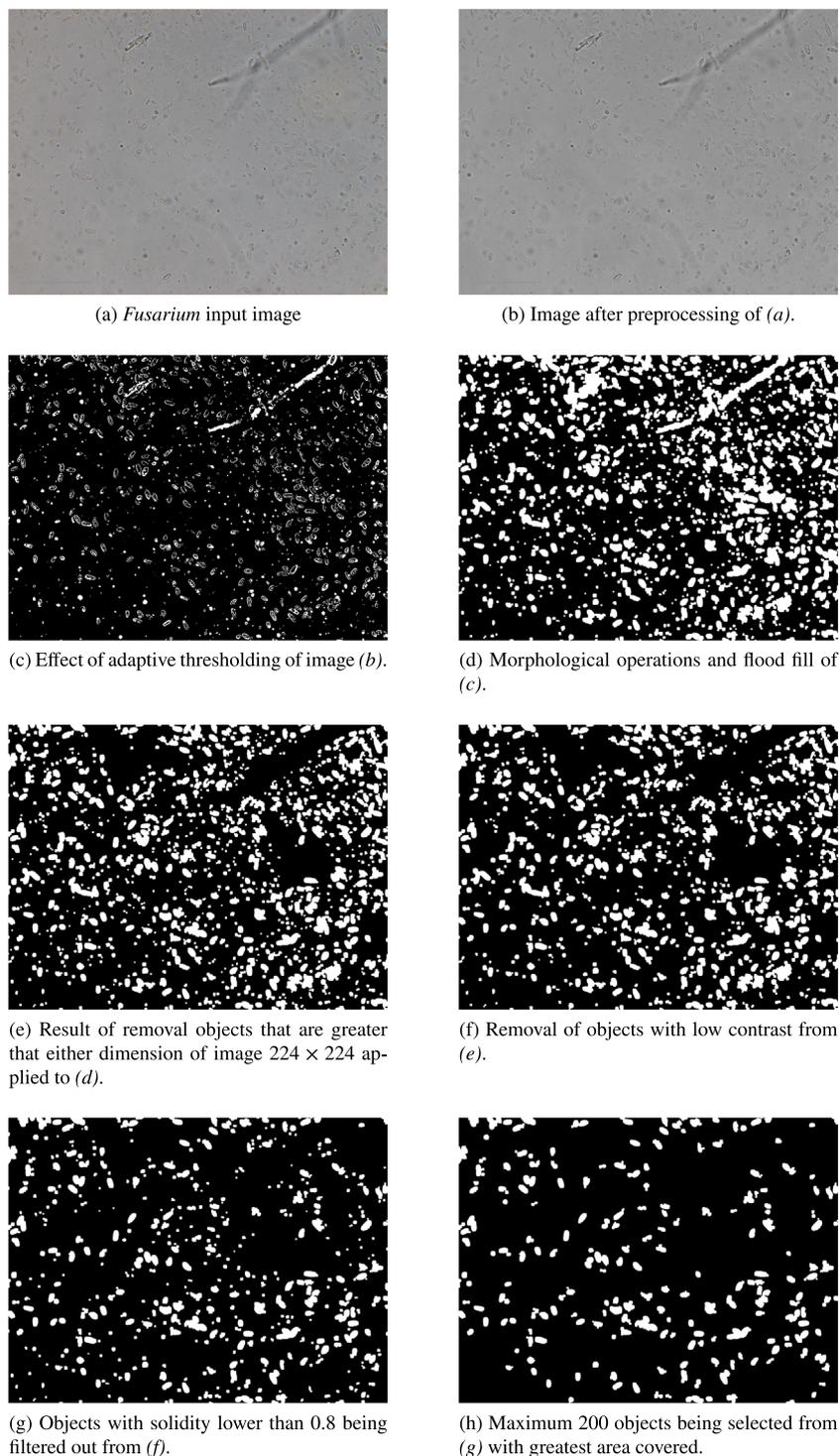


Fig. 4. Retrieval process of subimages from *Fusarium* microscopic image.

dataset, spanning a wide range of object categories, has proven the remarkable breakthroughs in these respective domains. By harnessing the acquired knowledge from ImageNet, researchers are able to achieve significant advancements in ML area.

In the context of transfer learning, several techniques can be employed. One approach is to initialize the network with pre-trained weights, such as ResNet50 weights established during training on the ImageNet dataset. Network is then exposed to new images and training is performed in batches. The objective is to fine-tune the network's weights slightly, allowing it to adapt to a new task, such as soil microorganism classification. Another strategy involves freezing specific

blocks within the network during training, preventing their weights from being updated through backpropagation.

Determining which blocks to freeze is a crucial decision. As information propagates through the network, the convolutional layers learn to detect higher-level features that become increasingly specific to the given task (Yamashita et al., 2018). Therefore, it is advisable to freeze the initial layers responsible for low-level features that are not task-specific. In our experiments, we chose to freeze all blocks except for the fifth residual block, which includes all convolutional layers named in the ResNet50 topology starting from *conv5\_block1\_1\_conv*. This allows us to fine-tune the high-level feature maps, tailoring them to our specific

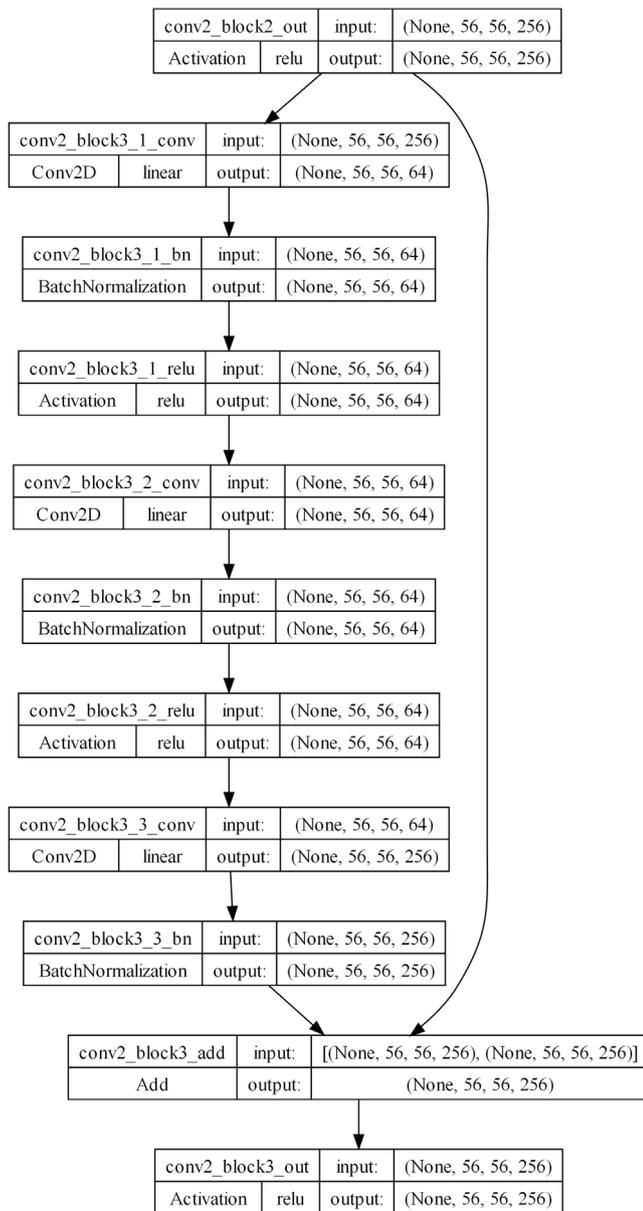


Fig. 5. Fragment of the ResNet50 network topology.

task. By selectively freezing the appropriate blocks and enabling fine-tuning of high-level features, our system can effectively leverage the knowledge gained from pre-training on ImageNet while adapting to the soil microorganism classification task.

## 2.4. Model training and evaluation process

To validate the statistical significance of our application, we provide a concise overview of the training and evaluation process using transfer learning with ResNet50.

### 2.4.1. Training

The dataset consisting of 134 microscopic images is partitioned into distinct training, validation and testing sets, with a ratio of 0.6:0.1:0.3 respectively. The training set is employed to feed the network with data during the training process, while the current training loss is evaluated on the validation dataset. The testing set is reserved for assessing the performance of the trained model. Upon partitioning the

dataset, subimages are generated, resulting 7121 images in the training set, 1020 in the testing set and 2034 in the validation set. In total, the dataset comprises 10,175 subimages.

For the training process, a configuration is established utilizing the ResNet50 network. The trainable fifth residual block is employed, followed by average pooling and flattening operations. On top of this, a classifier is constructed, consisting of a dense layer with 512 neurons and a dropout rate of 0.3. During training, the dropout mechanism randomly drops 30% of the connections between neurons, thereby preventing overfitting of the classifier. Furthermore, the same elements are added once more, leading to an output dense layer with 4 neurons, corresponding to the number of target classes associated with the soil microorganism genera. In other words topology of the classifier can be described as 512-DropOut-512-DropOut-4. A *softmax* activation function is applied to the output layer to obtain probability distributions over the classes to select affiliation the given class by the highest probability value for a given output neuron. Please refer to Fig. 6 for a visual representation of the classification architecture.

During the training of the network, the categorical cross-entropy loss function is employed, which is a common choice for multi-class classification tasks (LeCun et al., 2015). The utilized optimizer is *Adam*, with a learning rate set to 0.001 (Kingma and Ba, 2014). The training process is conducted with a batch size of 64 images over 1000 epochs. It is worth mentioning that the model was also evaluated for 5000 epochs, but it yielded inferior results. It should be noted that the input dataset was too large to be loaded directly into both the RAM and GPU memory. In this context, the Dataset functionality within the TensorFlow module played a pivotal role, closely followed by the preprocessing step of normalizing the loaded sub-images (ranging from [0; 255] to [0; 1]) for each color channel. The Dataset functionality facilitates the random retrieval of data from memory, allowing for the feeding of batches during the fitting process of the CNN. The whole process is performed on PC computer with Ryzen 9 3900X CPU, 64 GB DDR4 3600MHz RAM and Nvidia RTX 3090 24 GB VRAM GPU using CUDA and cuDNN.

### 2.4.2. Evaluation

In this study, the trained model from the previous step was subjected to evaluation on the testing set. The classification of individual subimages reflects the effectiveness of the subimage retrieval and CNN training processes. The practical objective is to classify the entire image into the appropriate class, providing an answer to the question of which microorganism is present in the sample. To achieve this, the results obtained from assigning subimages to classes are combined using a majority voting scheme. To illustrate the latter, let us consider the case of the first test image  $I_t^1$  with 150 subimages retrieved. Next, images are fed as input to the pre-trained network in batches, resulting in predictions for each subimage. For instance, the results indicate that 100 subimages are assigned to class  $C_0$ , 20 subimages to  $C_1$ , 10 subimages to  $C_2$  and 20 subimages to  $C_3$  respectively. By employing the majority voting strategy, the final classification of image  $I_t^1$  is determined based on the class with the highest number of “votes” from the subimages. In this example,  $I_t^1$  is classified as  $C_0$  since it received 100 “votes” for class  $C_0$ .

### 2.5. Preliminary experiments

In the results section, we showcase the outcomes of the most successful setup noted during our experimentation process. Prior to this, we conducted a series of experiments aimed at refining the optimal model configuration. Our initial objective was to determine the most effective method of sub-image retrieval. The comprehensive process, elucidated in Section 2.1 and illustrated in the data flow diagram in Fig. 2, is an outcome of preliminary trials where we assessed various dataset retrieval approaches. In the course of this experimentation, we generated 18 distinct datasets, focusing on factors such as solidity and

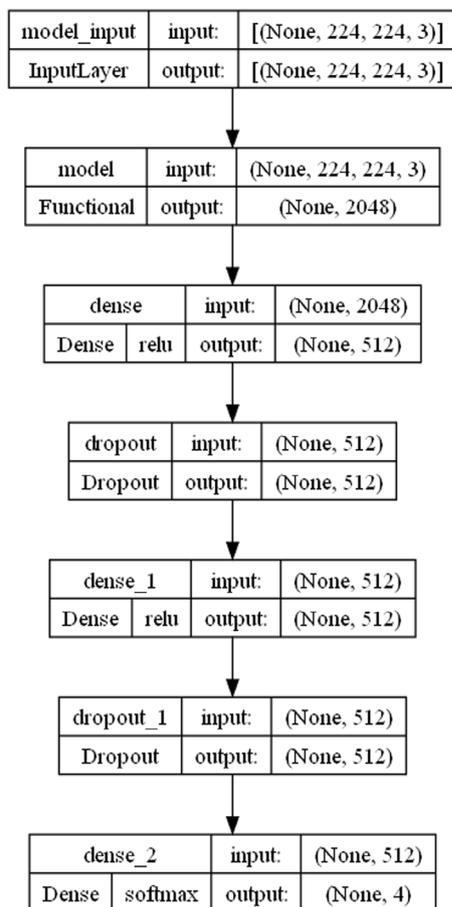


Fig. 6. The ResNet50 classifier, where model\_input states as input layer, model as ResNet50 feature extractor and then two classification layers with 512 neurons each and 0.3 dropout rate.

standard deviation filtration thresholds, the number of retrieved sub-images, selection criteria for the designated number of sub-images (including methods like choosing images closest to the median object size or selecting top objects with the highest covered area), and the evaluation of whether a second flood-fill process yielded improved results. Given the impracticality of presenting the detailed results for various sets, including thorough explanations of their nuanced distinctions, we chose to focus on the best sub-image retrieval method.

Subsequently, we proceeded to construct the model and were confronted with the task of fine-tuning hyperparameters for CNN compilation. This encompassed considerations such as the choice of loss function, optimizer along with learning rate, topology of the top layer (classification layers atop convolutional layers), dropout layer rates, batch size, and the selection of the block from which weights would be made trainable for Transfer Learning. Our experimentation encompassed the following configurations:

- v1: trainable from *conv5\_block1\_1\_conv*, top topology: Dense(512, *relu*) x Dropout(0.3) x Dense(512, *relu*) x Dropout(0.3), Categorical Crossentropy loss function, Adam(learning rate 0.001) optimizer, batch size of 32.
- v2: v1 trainable from *conv4\_block1\_1\_conv*.
- v3: v1 trainable from *conv3\_block1\_1\_conv*.
- v4: v3 with batch size of 64.
- v5: v3 with batch size of 256.
- v6: v3 with batch size of 128 and topology of the top layer: 256 x 0.3 x 256 x 0.3.

Table 1

The classifier performance results on different variants of models.

Model version	Precision	Recall	AUC	TOP 1 ACC
v1	0.710	0.698	0.875	0.706
v2	0.706	0.705	0.834	0.705
v3	0.684	0.684	0.801	0.684
v4	0.672	0.658	0.845	0.665
v5	0.674	0.659	0.841	0.668
v6	0.688	0.672	0.855	0.682
v7	0.694	0.652	0.858	0.672
v8	0.667	0.658	0.844	0.662
v9	0.458	0.458	0.645	0.458
v10	0.692	0.672	0.857	0.680
v11	0.699	0.696	0.853	0.698
v12	0.686	0.680	0.844	0.683
v13	0.693	0.685	0.834	0.691
v14	0.459	0.314	0.673	0.365
v15	0.694	0.688	0.846	0.692
v16	0.000	0.000	0.649	0.421
v17	0.704	0.700	0.861	0.702
v18	0.692	0.685	0.849	0.690

- v7: v3 with batch size of 128 and topology of the top layer: 256 x 0.3 x 512 x 0.3.
- v8: v3 with batch size of 128 and topology of the top layer: 1024 x 0.3 x 1024 x 0.3.
- v9: v1 trainable from *conv1\_block1\_1\_conv*.
- v10: v3 with topology of the top layer: 512 x 0.2 x 512 x 0.2.
- v11: v3 with topology of the top layer: 512 x 0.4 x 512 x 0.4.
- v12: v1 with topology of the top layer: 512 x 0.4 x 512 x 0.4.
- v13: v1 with topology of the top layer: 512 x 0.2 x 512 x 0.2.
- v14: v1 with batch size of 256.
- v15: v1 with batch size of 128.
- v16: v1 with *Adam*(learning rate 0.1).
- v17: v1 with *Adam*(learning rate  $1e-5$ ).
- v18: v1 with *elu* activation function.

The outcomes are provided in Table 1. Among the variants, the most favorable results in terms of Area Under the Curve (AUC) were obtained for v1 and it is this variant that was chosen for the model detailed in Section 2.4.1.

### 3. Results

Upon conducting experiments, the performance of the classifier was evaluated using various measures on the testing set, including precision, recall and F1-score for different classes, as well as overall performance. The results are summarized in Table 2. The testing set consisted of a total of 5448 subimages; however, it is important to note that the classes were unbalanced. Specifically, class 1 was represented by 644 images, while class 0 contained 2360 images. As a consequence, the classifier exhibited lower performance in distinguishing class 1, with a precision of 62% compared to 86% for class 0, which is overrepresented in the dataset.

To mitigate the negative impact of class imbalance, we initially attempted to address the issue by undersampling the images. This involved randomly selecting 644 subimages from each class, resulting in a balanced dataset of exactly 644 images per class that yielded poor performance by the classifier, prompting us to retain the original dataset.

The overall precision (that indicates the accuracy of correctly predicted positive instances across all classes) achieved by the classifier was 82%, which is considered promising given the challenges inherent in the experiment. It includes imperfect subimage retrieval, which may include objects not representing actual microorganism fragments. Additionally, the input microscopic images were captured at different magnification levels and contained impurities such as reflections

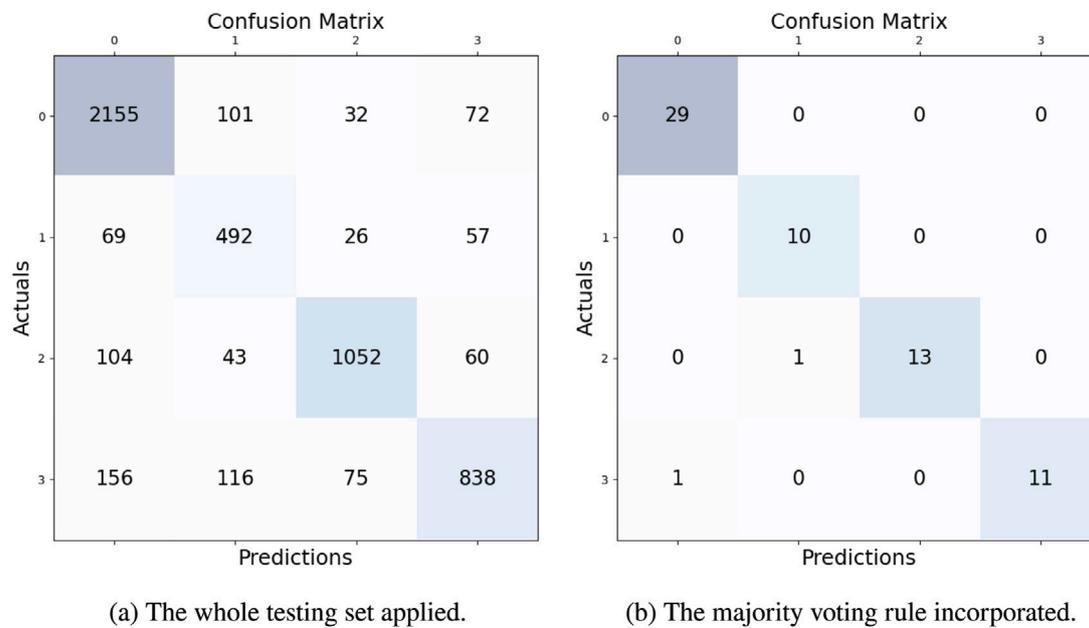


Fig. 7. Confusion matrices of the classifier performance results on the whole testing set and upon incorporating majority voting rule on the subimages to merge results obtaining the final classification to the given class.

Table 2  
The classifier performance results on applied directly the testing test.

	Precision	Recall	F1-score	Support
0	0.86	0.91	0.88	2360
1	0.62	0.78	0.69	644
2	0.89	0.83	0.86	1259
3	0.81	0.68	0.74	1185
Micro avg	0.82	0.82	0.82	5448
Macro avg	0.80	0.80	0.79	5448
Weighted avg	0.83	0.82	0.82	5448
Samples avg	0.82	0.82	0.82	5448

Table 3  
The classifier performance results upon incorporating majority voting rule on the subimages to merge results obtaining the final classification to the given class.

	Precision	Recall	F1-score	Support
0	0.97	1.00	0.98	29
1	0.91	1.00	0.95	10
2	1.00	0.93	0.96	14
3	1.00	0.92	0.96	12
Micro avg	0.97	0.97	0.97	65
Macro avg	0.97	0.96	0.96	65
Weighted avg	0.97	0.97	0.97	65
Samples avg	0.97	0.97	0.97	65

from the coverslip. Despite these impediments, the obtained result is promising.

Further insight into the classifier’s performance can be gained from the confusion matrix, as shown in Fig. 7. The matrix reveals the classification outcomes for each class, reaffirming the subpar performance observed for class 1.

The incorporation of the majority voting scheme, as detailed in the previous section, has yielded significant improvements in the performance of the classifier. The results on the testing set are depicted in Table 3. Notably, the primary issue observed in the previous model, namely the poor performance of class 1 due to its underrepresentation, has been effectively addressed. The precision for class 1 has now reached 91%, while the average precision and F1-score have significantly improved to 97%, indicating superior performance. These advancements are also evident in the confusion matrix shown in Fig. 7.

The utilization of the majority voting scheme on the retrieved subimages has greatly enhanced the classifier’s ability to accurately differentiate input images. This improvement stems from the enhanced generalizability of the classifier, where minor misclassifications in individual subimages do not significantly impact the overall image classification. The classifier exhibits robustness in handling faulty classifications, which aligns with the true objective of our task: accurately identifying the genera of microorganisms present in the image. Given that monocultures are observable in the images, the classification of subimages becomes less relevant, with the focus shifting to assigning the overall image to a specific genera. It should be noted that applying the same scheme directly to polycultures of microorganisms in the sample may require additional methods, such as utilizing graphs with probabilities assigned to the edges.

One aspect of concern in the results is the limited size of the testing set. The preparation of samples for our experiment is a non-trivial and laborious task, involving a purification process, cultivation of samples under specific conditions for several days and image preparation. Consequently, we acknowledge the need to expand the dataset in future work and to incorporate a wider range of microorganism genera. The evaluation results involves assessing the accuracy and loss function values throughout the training process, ensuring that overfitting does not occur. To alleviate the risk of overfitting, we have incorporated dropout in the dense layers of our CNN. As illustrated in Fig. 8, there is no evidence of overfitting. Both the accuracy on the training set and the validation set steadily increase over the course of the epochs, particularly beyond 500 epochs. Additionally, no significant decrease in training accuracy is observed as the training progresses, indicating that our model successfully avoids overfitting.

A significant fact is that computations were also conducted without employing the Transfer Learning. However, the results obtained without Transfer Learning exhibited similarities to those achieved with the latter, even though there was a notably extended computation time. This can be attributed to the requirement of training all blocks of ResNet50 for each epoch, as well as starting with random weights rather than utilizing weights pre-trained on ImageNet. The generated outcomes underscore the superiority of the Transfer Learning.

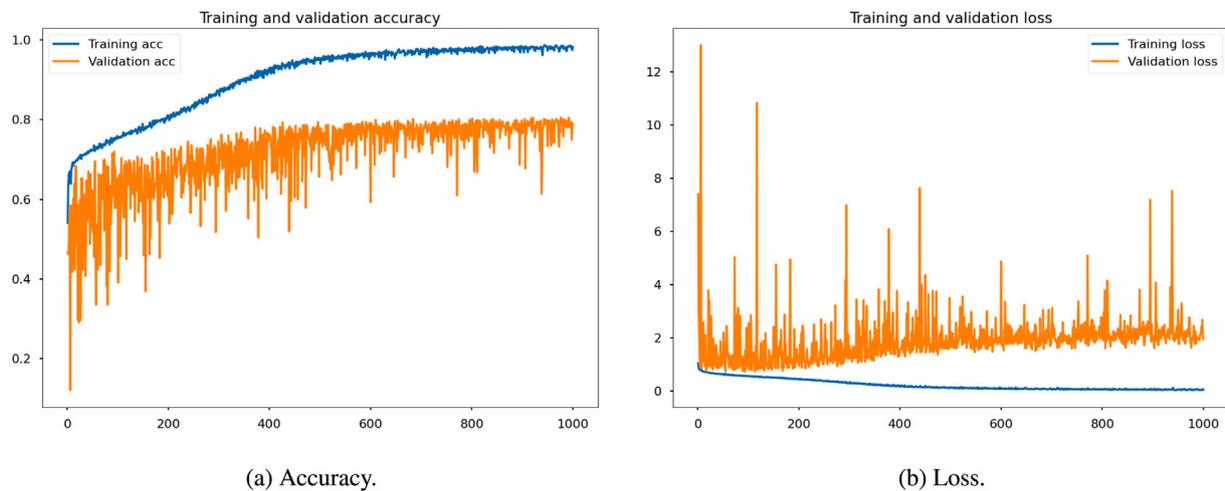


Fig. 8. Accuracy and loss function during training of the examined model on training and validation sets.

#### 4. Conclusions

In conclusion, this study proposes the use of machine vision and machine learning techniques, specifically CNN for the automated identification of different fungi and *Chromista* based on their microscopic images and morphological traits. The system aims to address the limitations of traditional identification methods, such as morphological analysis, which are often restricted to genus-level identification, by automation the process, making it faster and more efficient.

The developed system was evaluated using a dataset of soil microorganisms obtained from the *Symbio-Bank* at The National Institute of Horticultural Research.

The dataset (Struniawski, 2023) comprised images of four different genera: *Fusarium*, *Trichoderma*, *Verticillium* and *Phytophthora*. Despite challenges related to imperfect subimage retrieval, variations in magnification levels and impurities in the images, the classifier achieved a promising overall precision of 82%.

Furthermore, majority voting scheme was incorporated, leading to significant improvements in classifier performance. The precision for the underrepresented class 1, increased to 91% from 62%, while the overall average precision and F1-score improved to 97%. These enhancements demonstrate the system's ability to accurately differentiate between microorganism genera, aligning with the primary objective of identifying the overall image's genera.

The evaluation of the model's accuracy and loss function values throughout the training process revealed no evidence of overfitting. The accuracy on the training and validation sets steadily increased over the epochs, with no significant decrease in training accuracy observed. This indicates that the dropout mechanism implemented in the dense layers of the CNN effectively prevented overfitting.

For future extension of this work, it is important to expand the dataset, to incorporate additional microorganism genera and to explore the integration of phenotypic and molecular methods to achieve species-level identification. The system has the potential to revolutionize this area in the agricultural industry by providing cost-effective and efficient pathogen detection, thereby minimizing the risk of crop losses and improving overall agricultural productivity.

In summary, the combination of machine vision, machine learning techniques and CNNs presented in this study represents a significant step towards developing an automated system for accurate identification of soil microorganisms. With further advancements and integration of complementary methods, the system can serve as a valuable and feasible tool for farmers, laboratories and researchers in the agricultural domain.

#### CRedit authorship contribution statement

**Karol Struniawski:** Prepared and programmed the classification tool, Wrote the main manuscript text, Work in the algorithm conceptual phase. **Ryszard Kozera:** Wrote the main manuscript text, Work in the algorithm conceptual phase. **Pawel Trzcinski:** Microscopic images. **Anna Lisek:** Microscopic images. **Lidia Sas Paszt:** Experimental concept.

#### Declaration of competing interest

The author(s) declare no competing interests.

#### Data availability

Exemplary data is shared on the Zenodo repository cited in the proposed paper, the full dataset is available to be shared on reasonable request.

#### Acknowledgments

This research was supported by The National Centre for Research and Development, Poland within the framework of the project BIOS-TRATEG, grant number BIOSTRATEG3/344433/16/NCBR/2018.

#### References

- Acuña, C., Vinnett, L., Kuan, S., 2016. Improving image analysis of online bubble size measurements with enhanced algorithms. In: *IMPC Proc.* pp. 26–28.
- Alshamrani, K., Alshamrani, H.A., Alqahtani, F.F., Almutairi, B.S., 2023. Enhancement of mammographic images using histogram-based techniques for their classification using CNN. *Sensors* 23 (1), <http://dx.doi.org/10.3390/s23010235>.
- Arrabally, S.B.R., Juliet, S., 2019. Transfer Learning with ResNet-50 for Malaria cell-image classification. In: *ICCS Proc.* pp. 0945–0949. <http://dx.doi.org/10.1109/ICCSProc.2019.8697909>.
- Bailey, M., Gomez, C., Finch, J., 2005. Development and application of an image analysis method for wide bubble size distributions. *Miner. Eng.* 18 (12), 1214–1221. <http://dx.doi.org/10.1016/j.mineng.2005.07.019>.
- Bozinovski, S., Fulgosi, A., 1976. The influence of pattern similarity and transfer learning upon the training of a base perceptron B2. In: *Proceedings of Symposium Informatica*. pp. 3–121–5.
- Cakir, S., Kahraman, D., Cetin-Atalay, R., Cetin, A., 2018. Contrast enhancement of microscopy images using image phase information. *IEEE Access PP*, <http://dx.doi.org/10.1109/ACCESS.2018.2796646>.
- Dave, I.R., Upla, K.P., 2017. Computer aided diagnosis of Malaria disease for thin and thick blood smear microscopic images. In: *SPIN*. pp. 561–565. <http://dx.doi.org/10.1109/SPIN.2017.8050013>.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. In: *CVPR Proc.* pp. 248–255.

- Fan, H., Zhang, Y., Li, J., Jiang, J., Waheed, A., Wang, S., Rasheed, S.M., Zhang, L., Zhang, R., 2023. Effects of organic fertilizer supply on soil properties, tomato yield, and fruit quality: A global meta-analysis. *Sustainability* 15 (3), <http://dx.doi.org/10.3390/su15032556>.
- Fiorio, C., Gustedt, J., 1996. Two linear time Union-Find strategies for image processing. *TCS* 154 (2), 165–181. [http://dx.doi.org/10.1016/0304-3975\(94\)00262-2](http://dx.doi.org/10.1016/0304-3975(94)00262-2).
- Gao, Z., Wang, L., Zhou, L., Zhang, J., 2017. HEp-2 cell image classification with Deep Convolutional Neural Networks. *IEEE J. Biomed. Health Inform.* 21 (2), 416–428. <http://dx.doi.org/10.1109/JBHI.2016.2526603>.
- Gonzalez, R.C., Woods, R.E., 2018. *Digital Image Processing*, fourth ed. Pearson.
- Han, J., Dong, Y., Zhang, M., 2021. Chemical fertilizer reduction with organic fertilizer effectively improve soil fertility and microbial community from newly cultivated land in the Loess Plateau of China. *Appl. Soil Ecol.* 165, 103966. <http://dx.doi.org/10.1016/j.apsoil.2021.103966>.
- Hanbal, I.F., Ingosan, J.S., Oyam, N.A.A., Hu, Y., 2020. Classifying wastes using Random Forests, Gaussian Naïve Bayes, support vector machine and multilayer perceptron. *IOP Conf. Ser.: Mater. Sci. Eng.* 803 (1), 012017. <http://dx.doi.org/10.1088/1757-899X/803/1/012017>.
- He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In: *ICCV Proc.* pp. 1026–1034. <http://dx.doi.org/10.1109/ICCV.2015.123>.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: *CVPR Proc.* <http://dx.doi.org/10.48550/arXiv.1512.03385>.
- Hossain, M.B., Iqbal, S.H.S., Islam, M.M., Akhtar, M.N., Sarker, I.H., 2022. Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images. *IMU* 30, 100916. <http://dx.doi.org/10.1016/j.imu.2022.100916>.
- Huerta, E.A., Khan, A., Davis, E., Bushell, C., Gropp, W.D., Katz, D.S., Kindratenko, V., Koric, S., Kramer, W.T.C., McGinty, B., McHenry, K., Saxton, A., 2020. Convergence of artificial intelligence and high performance computing on NSF-supported cyberinfrastructure. *J. Big Data* 7 (1), <http://dx.doi.org/10.1186/s40537-020-00361-2>.
- Ilonen, J., Juránek, R., Eerola, T., Lensu, L., Dubská, M., Zemčík, P., Kälviäinen, H., 2018. Comparison of bubble detectors and size distribution estimators. *Pattern Recognit. Lett.* 101, 60–66. <http://dx.doi.org/10.1016/j.patrec.2017.11.014>.
- Khan, M.B., Nisar, H., Ng, C.A., Lo, P.K., Yap, V.V., 2015. Local adaptive approach toward segmentation of microscopic images of activated sludge flocs. *J. Electron. Imaging* 24 (6), 061102. <http://dx.doi.org/10.1117/1.JEI.24.6.061102>.
- Kingma, D., Ba, J., 2014. Adam: A method for stochastic optimization. In: *ICLR Proc.* <http://dx.doi.org/10.48550/arXiv.1412.6980>.
- Konopka, A., Struniawski, K., Kozera, R., Trzciński, P., Sas-Paszt, L., Lisek, A., Górnik, K., Derkowska, E., Gluszek, S., Sumorok, B., Frać, M., 2022. Classification of soil bacteria based on machine learning and image processing. In: *ICCS Proc.* pp. 263–277. [http://dx.doi.org/10.1007/978-3-031-08757-8\\_23](http://dx.doi.org/10.1007/978-3-031-08757-8_23).
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521 (7553), 436–444. <http://dx.doi.org/10.1038/nature14539>.
- LeCun, Y., Bottou, L., Orr, G.B., Müller, K.-R., 1998. Efficient BackProp. In: *Neural Networks: Tricks of the Trade*. pp. 9–50. [http://dx.doi.org/10.1007/3-540-49430-8\\_2](http://dx.doi.org/10.1007/3-540-49430-8_2).
- Lie, W.-N., 1995. Automatic target segmentation by locally adaptive image thresholding. *IEEE Trans. Image Process.* 4 (7), 1036–1041. <http://dx.doi.org/10.1109/83.392347>.
- Ma, L.-J., Geiser, D.M., Proctor, R.H., Rooney, A.P., O'Donnell, K., Trail, F., Gardiner, D.M., Manners, J.M., Kazan, K., 2013. *Fusarium pathogenomics*. *Annu. Rev. Microbiol.* 67 (1), 399–416. <http://dx.doi.org/10.1146/annurev-micro-092412-155650>.
- Mohamad, N., Jusoh, N., Zaw, Z., Win, S., 2014. Bacteria identification from microscopic morphology: A survey. *IJSCAI* 3, 1–12. <http://dx.doi.org/10.5121/ijscai.2014.3201>.
- Ong, J.D.L., Abigan, E.G.T., Cajucom, L.G., Abu, P.A.R., Estuar, M.R.J.E., 2020. Ensemble convolutional neural networks for the detection of microscopic fusarium oxysporum. In: *ISVC Proc. Part I*. pp. 321–332. [http://dx.doi.org/10.1007/978-3-030-64556-4\\_25](http://dx.doi.org/10.1007/978-3-030-64556-4_25).
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* 9 (1), 62–66.
- Párraga, C.A., Brelstaff, G., Troscianko, T., Moorehead, I.R., 1998. Color and luminance information in natural scenes. *J. Opt. Soc. Amer. A* 15 (3), 563–569. <http://dx.doi.org/10.1364/JOSAA.15.000563>.
- Pham, D., Xu, C., Prince, J., 2000. A survey of current methods in medical image segmentation. *Annu. Rev. Biomed. Eng.* 2, 315–337. <http://dx.doi.org/10.1146/annurev.bioeng.2.1.315>.
- Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. In: *ICLR Proc.* pp. 1–14. <http://dx.doi.org/10.48550/arXiv.1409.1556>.
- Smith, S.E., Read, D.J., 2010. *Mycorrhizal Symbiosis*. Academic Press.
- Soille, P., 1999. *Morphological Image Analysis: Principles and Applications*. Springer-Verlag, pp. 173–174. <http://dx.doi.org/10.1007/978-3-662-03939-7>.
- Struniawski, K., 2023. *Sample Microscopic Images of Soil Microorganism*. Zenodo, <http://dx.doi.org/10.5281/zenodo.7965200>.
- Struniawski, K., Konopka, A., Kozera, R., 2022. Identification of soil bacteria with machine learning and image processing techniques applying single cells' region isolation. In: *ESM 2022*. pp. 76–81.
- Szeliski, R., 2011. *Computer Vision: Algorithms and Applications*. Springer Science & Business Media.
- Tjoa, E.A., Yowan Nugraha Suparta, I.P., Magdalena, R., Kumalasari CP, N., 2022. The use of CLAHE for improving an accuracy of CNN architecture for detecting pneumonia. *SHS Web Conf.* 139, 03026. <http://dx.doi.org/10.1051/shsconf/202213903026>.
- U.S. Environmental Protection Agency, 2023. *Climate change impacts on agriculture and food supply*. <https://www.epa.gov/climateimpacts/climate-change-impacts-agriculture-and-food-supply>.
- Vinnett, L., Cornejo, I., Yianatos, J., Acuña, C., Urriola, B., Guajardo, C., Esteban, A., 2022. The correlation between macroscopic image and object properties with bubble size in flotation. *Minerals* 12 (12), <http://dx.doi.org/10.3390/min12121528>.
- Wahid, M.F., Ahmed, T., Habib, M.A., 2018. Classification of microscopic images of bacteria using deep convolutional neural network. In: *ICECE*. pp. 217–220. <http://dx.doi.org/10.1109/ICECE.2018.8636750>.
- Watanabe, T., 2010. *Pictorial Atlas of Soil and Seed Fungi*, third ed. CRC Press.
- Wijayawardene, N.N., et al., 2020. Outline of Fungi and fungus-like taxa. *Mycosphere* 11 (1), 1060–1456. <http://dx.doi.org/10.5943/mycosphere/11/1/8>.
- Yamashita, R., Nishio, M., Do, R.K.G., Togashi, K., 2018. Convolutional neural networks: an overview and application in radiology. *Insights into Imag.* 9 (4), 611–629. <http://dx.doi.org/10.1007/s13244-018-0639-9>.
- Zhang, L., Lu, L., Nogues, I., Summers, R.M., Liu, S., Yao, J., 2017. DeepPap: Deep convolutional networks for cervical cell classification. *IEEE J. Biomed. Health Inform.* 21 (6), 1633–1643. <http://dx.doi.org/10.1109/JBHI.2017.2705583>.