Research article

A comparative study of Machine Learning-based classification of Tomato fungal diseases: Application of GLCM texture features

Chimango Nyasulu [a,d,*], Awa Diattara [a], Assitan Traore [b], Cheikh Ba [a],
Papa Madiallacké Diedhiou [c], Yakhya Sy [c], Hind Raki [d,e],
Diego Hernán Peluffo-Ordóñez [d,e,f]

[a] *LANI (Laboratoire d'Analyse Numérique et Informatique), University of Gaston Berger, BP:234, Saint-Louis, 32000, Saint-Louis, Senegal*
[b] *Business & Decision, Grenoble, 38000, Grenoble, France*
[c] *UFR des Sciences Agronomiques d'Aquaculture et des Technologies Alimentaires, University of Gaston Berger, BP:234, Saint-Louis, 32000, Saint-Louis, Senegal*
[d] *College of Computing, Mohammed VI Polytechnic University, Lot 660, Hay Moulay Rachid, Ben Guerir, 43150, Ben Guerir, Morocco*
[e] *Faculty of Engineering, Corporación Universitaria Autónoma de Nariño, Pasto, 520001, Nariño, Colombia*
[f] *SDAS Research Group, Lot 660, Hay Moulay Rachid, Ben Guerir, 43150, Ben Guerir, Morocco*

ARTICLE INFO

ABSTRACT

Globally, agriculture remains an important source of food and economic development. Due to various plant diseases, farmers continue to suffer huge yield losses in both quality and quantity. In this study, we explored the potential of using Artificial Neural Networks, K-Nearest Neighbors, Random Forest, and Support Vector Machine to classify tomato fungal leaf diseases: Alternaria, Curvularia, Helminthosporium, and Lasiodiplodi based on Gray Level Co-occurrence Matrix texture features. Small differences between symptoms of these diseases make it difficult to use the naked eye to obtain better results in detecting and distinguishing these diseases. The Artificial Neural Network outperformed other classifiers with an overall accuracy of 94% and average scores of 93.6% for Precision, 93.8% for Recall, and 93.8% for F1-score. Generally, the models confused samples originally belonging to Helminthosporium with Curvularia. The extracted texture features show great potential to classify the different tomato leaf fungal diseases. The results of this study show that texture characteristics of the Gray Level Co-occurrence Matrix play a critical role in the establishment of tomato leaf disease classification systems and can facilitate the implementation of preventive measures by farmers, resulting in enhanced yield quality and quantity.

1. Introduction

Agriculture remains a primary source of food and economic development in Africa [1]. It offers employment to over 70% of the populace and contributes over 12% to the gross domestic product [1]. Mostly, agricultural production is rain-fed and relies on seasonal rainfall in the continent. However, due to many factors, like climate change, rising food prices have intensified the already

---

existing problem of food insecurity in Africa [2]. According to a recent report [3], in sub-Saharan Africa (SSA), about 12% of the population cannot meet the basic required food consumption.

Tomato is one of the major crops and important vegetables, has many beneficial traits and increasing demand on the market [4]. Tomato fruit can be consumed in different forms: fried, processed, fresh salad vegetable, stewed, sauce, baked, and as an ingredient in different dishes [4]. Nevertheless, the production of tomatoes is affected by many diseases among others. To produce healthy tomatoes and realize maximum yield proper disease management is required [5]. Early and accurate crop disease detection ensures the quality and quantity of crop yield [6]. Identifying various crop diseases and knowing how to manage them is a continuous challenge for farmers in Africa. Commonly African farmers depend on the naked eye to detect and interpret visual symptoms of crop diseases. However, this approach is inefficient in several aspects including its lack of accuracy, labor-intensive nature, dependence on expertise, and time-consuming process, among other limitations [7]. Different technologies such as Machine Learning (ML) have played a key role in addressing real-world agricultural challenges in recent years, leading to enhanced production [7]. For example, ML algorithms have been employed to enhance accuracy and efficiency in crop disease detection.

Several researchers have used different ML models for the detection and classification of different crop diseases. A study by [8] proposed a method to identify leaf disease in tomato plants using features: Hu Moments, Haralick, Local Binary Pattern, and Color Histograms. In their study, Random Forest (RF) and Decision Tree (DT) classifiers were used to classify tomato leaf diseases. Classification accuracy results for Random Forest and Decision Tree were 94% and 90% respectively. In a study by [9], an Artificial Neural Network (ANN) classifier was used to classify cotton and tomato leaf diseases based on Contrast, Standard Deviation, Energy, Homogeneity, Correlation, Mean, and Variance features. A higher overall accuracy of 92.5% was reported. In another study by [10], using Gray Level Co-occurrence Matrix (GLCM) texture features, a Support Vector Machine (SVM) classifier was used to classify infected tomato leaves. A higher accuracy of 99.83% was obtained using the linear kernel function.

[11] investigated the performance of classifiers: SVM, Discriminant Classifier (DC), K-Nearest Neighbors (KNN), Logistic Regression (LR), Naive Bayes (NB), DT, and RF for image-based classification of rice plant diseases using color features from rice blast, sheath blight, leaf blight, and healthy leaves classes. The SVM classifier achieved the highest classification accuracy of 94.65%. A study by [12] explored the use of Multinomial Logistic Regression (MLR), KNN, NB, and Multi-Class SVM to classify sunflower crop diseases leaf spot, rust, powdering mildew, and healthy leaves using color and texture features: Standard Deviation, Energy, Mean, Coarseness, Homogeneity, and Contrast. Multinomial Logistic Regression achieved the highest average accuracy of 92.57%. All the classifiers achieved an accuracy of 100% for healthy leaves. [13] proposed a multi-layer SVM method with a linear kernel function to classify potato leaf diseases, late blight, early blight, and healthy leaves based on GLCM texture features. The SVM classifier achieved an overall accuracy, precision, recall, and F1-score of 95.99%, 96.12%, 96.25%, and 96.16% respectively.

Despite the use and effectiveness of many ML methods for detecting and classifying different crop diseases, no study has been conducted to demonstrate that they allow for the classifying of tomato leaf fungal diseases Alternaria, Curvularia, Helminthosporium, and Lasiodiplodi. Small differences between symptoms of these diseases make it difficult to use the naked eye to obtain better results in detecting and distinguishing these diseases. Therefore, the main goal of this study was to investigate Machine Learning classifiers: ANN, KNN, RF, and SVM for the classification of tomato leaf fungal diseases: Alternaria, Curvularia, Helminthosporium, and Lasiodiplodi using GLCM texture features: Angular Second Moment (ASM), Correlation, Contrast, Dissimilarity, Energy, and Homogeneity.

2. Dataset

Under the supervision of agricultural experts with over five years of experience, an in-house experiment was conducted in Saint-Louis, Senegal to construct a tomato leaves dataset. The experiment comprised the following major steps: obtaining pathogenic strains, isolation and identification of microorganisms, verification of pathogenicity, and data acquisition and pre-processing. The data was divided into 75% for training the model and 25% for testing.

I **Pathogenic strains:**
   Two fungi Lasiodiplodia and Alternaria were extracted from leaf lesions of eggplant while Curvularia and Helminthosporium were extracted from leaf lesions of tomato. Alternaria disease leaves appear yellowish and coffee-colored from the leaf tips and margins of leaf petiole [14]. Lasiodiplodi disease shows small yellow blotches and brownish lesions on bottom leaves in its early stages, leaves turn brown and shrivel under severe conditions [15]. Curvularia disease shows light to dark brown spots on leaves in the early stages and brown spots with dark edges and necrotic lesions with leaves looking dry and wilted under severe infection [16]. Helminthosporium disease shows small lesions with yellow leaf tissues around leaf blades and straw-colored leaves characterizing severe cases [17].

II **Isolation and identification of microorganisms:**
   Small explants were cut from parts of leaves that showed characteristic symptoms of fungal diseases [18], then were dipped in 10% bleach for 30 seconds to clear superficial germs and rinsed with sterile water. Explants were inoculated in Petri dishes that contained the Potato Dextrose Agar (PDA) culture medium to determine the explant's susceptibility to fungal symptoms. To obtain uncontaminated samples for identifying the pathogenic agent, the dishes with mycelium were transferred to fresh dishes containing the same culture media after a 72-hour incubation period. This process ensured the isolation of pure cultures for further analysis [18].

III **Verification of pathogenicity:**
   To check the hypersensitivity of the isolated fungi, 100 tomato seedlings were planted, and a six-day-old suspension of spores was prepared from the margin of cultures of the fungi Lasiodiplodi, Alternaria, Curvularia, and Helminthosporium. With three
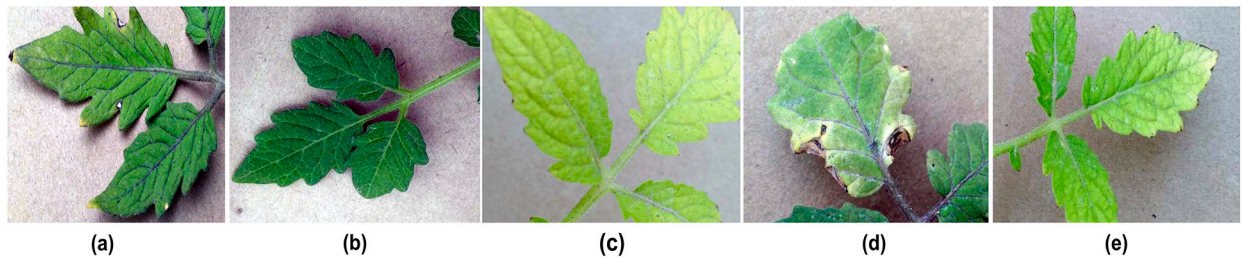
Fig. 1. Sample images for each class ((a) Alternaria, (b) Healthy, (c) Curvularia, (d) Helminthosporium, (e) Lasiodiplodi).

repetitions, 10 ml of the suspension spray was inoculated on the leaves of healthy tomatoes. For experimental control, three plants were inoculated with sterile water [19]. Inoculated plants were placed in polystyrene boxes to maintain room temperature and humidity. Agricultural experts assessed the symptoms every three days.

IV **Data acquisition and pre-processing:**

A smartphone was used to collect color images of healthy leaves and symptomatic leaves belonging to 5 classes containing 3,000 images in each class. All images were resized to 224 x 224.

Fig. 1(a) presents a sample of Alternaria class images, Fig. 1(b) presents a sample of Healthy class images, Fig. 1(c) presents a sample of Curvularia class images, Fig. 1(d) presents a sample of Helminthosporium class images, and Fig. 1(e) presents a sample of Lasiodiplodi class images.

2.1. GLCM texture feature extraction

Images can be expressed using different features such as texture. The texture of an image is determined by the spatial variation of the pixel intensity [20]. The GLCM method was used to extract 144 texture features in this study. The GLCM is a second-order statistical technique for texture analysis [20]. The GLCM method describes the texture of an image by calculating the frequency of occurrence of pairs of pixels with a specific value and in a defined spatial relationship in the image. Matrix element P (i, j) is the ratio of the number of times two pixels separated by a distance of Δx, Δy, and Δz, respectively, occur in a particular region of interest. The intensity of one region is 'i' and the intensity of the other region is 'j'. Matrix element P(i, j) contains second-order statistical values for the transition between gray levels 'i' and 'j' at a distance of d and an angle of θ) [20]. Each gray-scale image generated each textural feature, with distances at 1, 3, and 5 and angles at 0, 45°, 90°, 135°, 180°, 225°, 270° and 315°. The following features were extracted: Angular Second Moment, Contrast, Correlation, Dissimilarity, Energy, and Homogeneity. The ASM measures disorders in the textures of the image and is given by equation (1). Contrast measures local variations in an image by determining the difference between the maximum and minimum values of a group of adjacent pixels and is given by equation (2). Correlation measures correlation from -1 to 1 of neighboring pixels over the entire image and is given by equation (3). Dissimilarity is the average difference between the gray levels of an image and is given by equation (4). Energy measures the disorder of texture in an image by considering the degree of pixel pair repetitions and is given by equation (5). The energy value is large with highly correlated image pixels. Homogeneity measures how much of a difference between two pairs of objects in the same image. If the two objects have the same gray tone, then the homogeneity value will be higher. Mathematically, homogeneity is given by equation (6).

Given the following notation, where N represents the number of gray levels and $P(i, j)$ represents the normalized gray scale value of the kernel at positions i and j, with a sum of 1, the following features are considered:

$$Angular\,Second\,Moment = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j)^2, \tag{1}$$

$$Contrast = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i-j)^2, \tag{2}$$

$$Correlation = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i)(\sigma_j)}}, \tag{3}$$

$$Dissimilarity = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j) x \mid i - j \mid, \tag{4}$$

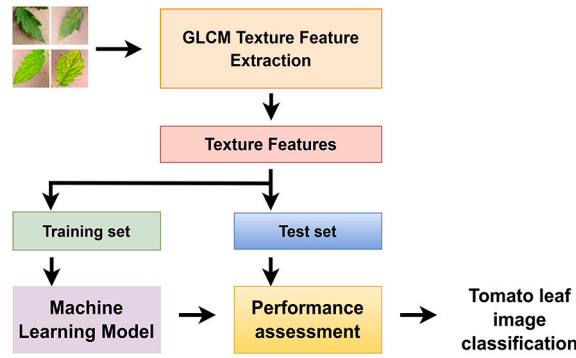$$Energy = \sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j)^2}, \tag{5}$$

and

Fig. 2. Summary of the steps followed in the study implementation.

$$Homogeneity = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{P(i,j)}{1+(i-j)^2}. \tag{6}$$

3. Methods

The implementation of this study was based on the principles of the Knowledge Discovery in Databases (KDD) process. The KDD process is the organized and iterative process of finding useful features and understandable patterns from datasets [21]. The main stages of KDD are data selection, pre-processing, transformation, mining, evaluation, and interpretation [21]. The workflow of the entire study is presented in Fig. 2.

3.1. Machine learning models

In this study, four ML classifiers were explored: ANN, KNN, RF, and SVM. The classifiers were chosen because in the reviewed studies these classifiers have performed well in classifying crop leaf diseases based on different features. The grid search method was used for hyperparameter tuning. Grid search is used to find the optimal hyperparameters of a model to achieve the best results. Table 1 presents a summary of the parameters we used. The SVM is an ML algorithm designed for classification and regression. The SVM aims at finding the superlative decision boundary referred to as a hyperplane in an N-dimensional feature space that correctly categorizes/classifies the data points. [22] presents a detailed explanation of the SVM. Random Forest classifier is among the most widely used classifiers. Random Forest classifier is based on random decision forest. It is an ensemble-based supervised learning algorithm for classification and regression. The algorithm works by creating several trees known as decision trees during the training of the model. The classifier outputs the class that has been selected by a majority of the trees. [23] presents a detailed explanation of the Random Forest classifier.

The ANN takes inspiration from the functioning of the human brain and utilizes interconnected nodes, known as neurons, which are arranged in multiple hidden layers to process data. The basic structure of an ANN is composed of an input layer, a hidden layer, and an output layer. The ANN architecture we used is presented in Fig. 3. [24] presents a detailed explanation of the ANN classifier. The K-Nearest Neighbors classifier is based on the KNN-supervised algorithm which works for both classification and regression. The underlying assumption in the KNN algorithm is that similar things exist near each other. The algorithm runs several times to select the optimal number of nearest neighbors denoted as K which reduces the number of errors when making predictions. For classification, the output is the label that is highly represented around a given data point. [25] presents a detailed explanation of the KNN classifier. This study was done with Python 3.80.5. [26].

3.2. Evaluation of the models

We considered metrics like Accuracy, Precision, Recall, F1-score, and the Area Under the ROC Curve (AUC) [27] to evaluate the models. By considering the following notation: TP is True Positives, TN is True Negatives, FP is False Positives, and FN is False Negatives, mathematically, Accuracy is given by equation (7), Precision is given by equation (8), Recall is given by equation (9) and F1-score is given by equation (10).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \tag{7}$$

$$Precision = \frac{TP}{TP+FP}, \tag{8}$$

$$Recall = \frac{TP}{TP+FN}, \tag{9}$$

and

Table 1
Model parameter setting.

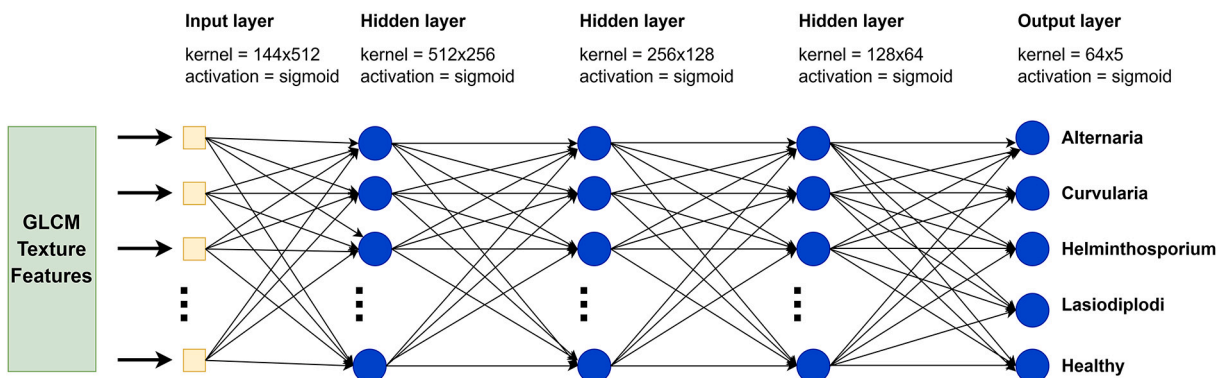| Model	Parameters	Values
ANN	activation	sigmoid
	optimizer	adam
	loss	categorical_crossentropy
	metrics	accuracy
	batch_size	128
	epochs	300
KNN	n_neighbors	1
	metric	minkowski
	p	2
	weights	distance
RF	n_estimators	100
	criterion	entropy
	max_depth	20
	min_samples_leaf	1
SVM	C	50
	class_weight	1: 3
	gamma	0.005
	probability	True

Fig. 3. ANN architecture for tomato fungal disease classification.

$$F1 - score = 2 \left(\frac{Precision \times Recall}{Precision + Recall} \right). \tag{10}$$

The AUC is a graph that shows how well the model can tell the difference between different classes by measuring the degree of separation of the model using the False Positive Rate and the True Positive Rate [27].

4. Results and discussion

Table 2 presents the Accuracy, Precision, Recall, and F1-score performance of ANN, KNN, SVM, and RF classifiers on classification of tomato leaf fungal diseases: Alternaria, Curvularia, Helminthosporium, and Lasiodiplodi using GLCM texture features: Energy, Correlation, Dissimilarity, Homogeneity, Contrast, and Angular Second Moment. The best result for each evaluation metric is highlighted.

The results show that ANN got an overall higher accuracy of 94% and average scores of 93.6% for Precision, 93.8% for Recall, and 93.8% for F1-score. Generally, among the disease classes, Lasiodiplodi got the highest average scores for Precision 89.75%, Recall 89.25%, and F1-score 89.25%. This could be attributed to the fact that there are more pronounced symptoms appearing at the edge of the leaves, with dark and damaged plant tissue. Hence it could be easier for the classifiers to make more correct positive predictions relative to the total positive predictions. Overall, the Healthy class got the highest average scores in Precision 91.75%, Recall 95.25%, and F1-score 93.75%. This could be attributed to the fact that the Healthy class is easily distinguishable from other classes hence a classifier can find a higher number of positive instances. Our results compare to other results reported by different researchers. In [12], Multinomial Logistic Regression achieved the highest overall accuracy of 92.57% for the classification of sunflower crop diseases using texture features. All the classifiers performed well in classifying healthy leaves. In another study by [9], ANN achieved the highest overall accuracy of 92.5% for the classification of cotton and tomato leaf diseases using texture features. In a study by [8], RF and DT classifiers were used to classify tomato leaf diseases using Color Histograms, Hu Moments, Haralick, and Local Binary Pattern features. The accuracy results were 94% for Random Forest and 90% for Decision Tree. Table 3 presents a summary comparison of the results obtained in this study and other studies on the parameter of accuracy.

Table 2

Summary of Accuracy, Precision, Recall, and F1-score performance of ANN, KNN, RF, and SVM classifiers.

Model	Accuracy	Class	Precision	Recall	F1-score
ANN	**94**	**Alternaria**	0.92	0.94	0.93
		Curvularia	0.93	0.91	0.92
		Helminthosporium	0.92	0.93	0.93
		Lasiodiplodia	0.95	0.93	0.94
		Healthy	**0.96**	**0.98**	**0.97**
KNN	89	**Alternaria**	0.86	0.85	0.85
		Curvularia	0.87	0.84	0.86
		Helminthosporium	0.88	0.86	0.87
		Lasiodiplodia	0.89	0.91	0.90
		Healthy	0.92	0.97	0.95
RF	86	**Alternaria**	0.84	0.86	0.85
		Curvularia	0.83	0.79	0.81
		Helminthosporium	0.86	0.82	0.84
		Lasiodiplodia	0.87	0.88	0.88
		Healthy	0.89	0.94	0.92
SVM	86	**Alternaria**	0.85	0.90	0.87
		Curvularia	0.85	0.80	0.82
		Helminthosporium	0.85	0.86	0.85
		Lasiodiplodia	0.88	0.85	0.87
		Healthy	0.90	0.92	0.91

Table 3

Comparison of the results obtained in this study and other studies on the parameter of accuracy.

Authors	Objective	Models	Accuracy (%)
Our study	Classify tomato leaf diseases	ANN	94
		KNN	88
		RF	86
		SVM	86
[8]	Classify tomato leaf diseases	RF	94
		DT	90
[9]	Classify tomato and cotton leaf diseases	ANN	92.5
[11]	Classify rice diseases	SVM	94.65
		RF	92.52
		DC	92.34
		KNN	91.39
		DT	83.18
		LR	75.85
		NB	75.72
[12]	Classify sunflower leaf diseases	MLR	92.57
		Multi-Class SVM	92.15
		KNN	89.32
		NB	89.07

The performance of the models was further analyzed by examining the confusion matrix to understand both successes and failures. This analysis helped identify the specific cases where the models excelled or became confused. Fig. 4, Fig. 5, Fig. 6, and Fig. 7 illustrate the correct and incorrect classification results of the models through confusion matrices.

The confusion matrices indicate that the models achieved a higher number of correct classifications overall, particularly for the Healthy class. Noteworthy misclassifications were observed in Fig. 4 where the ANN model correctly classified 684 samples out of the total 735 samples for Helminthosporium. Thus, an accuracy of 93%. In Fig. 5, the KNN model correctly classified 635 samples out of the total 735 samples for Helminthosporium. Thus, an accuracy of 86.3%. In Fig. 6, the RF model correctly classified 604 samples out of the total 735 samples for Helminthosporium. Thus, an accuracy of 82.1%. In Fig. 7, the SVM model correctly classified 610 samples out of the total 735 samples for Curvularia. Thus, an accuracy of 82.9%. Generally, the confusion matrices imply that the models confused samples originally belonging to Helminthosporium with Curvularia, the classification boundary between these classes was not well learned by the classifiers. For example, in Fig. 6 62 images of Helminthosporium were classified as Curvularia, and in Fig. 7 59 images of Curvularia were classified as Helminthosporium.

Additionally, we used the AUC as seen in Fig. 8, Fig. 9, Fig. 10 and Fig. 11 to visualize how well the models performed.
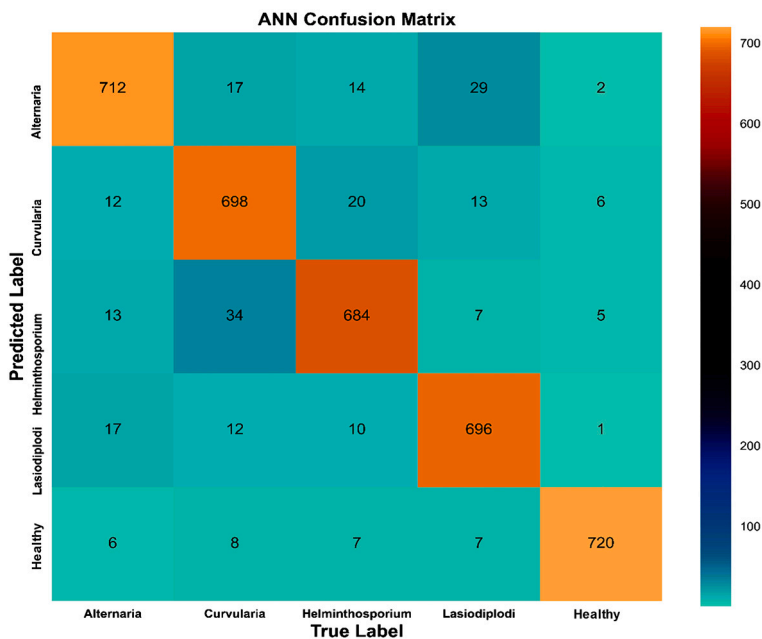
Fig. 4. ANN confusion matrix summarizing correct and incorrect classification.
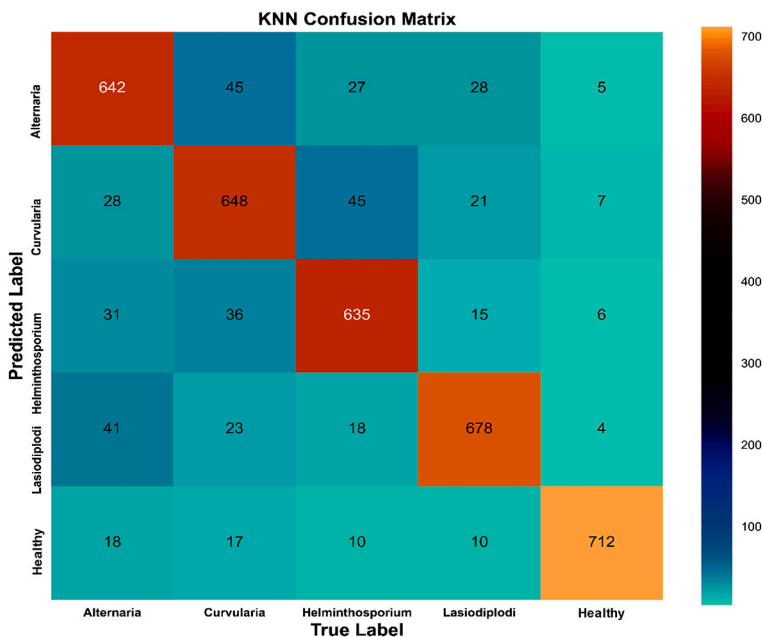

Fig. 5. KNN confusion matrix summarizing correct and incorrect classification.

Comparatively, the AUC for the Healthy class in all the models is higher (0.98 and above). However, the AUC for Curvularia is lower in all the models, this signifies that the models struggled to distinguish Curvularia from other classes. Despite the models achieving higher accuracies, there are several drawbacks of GLCM texture features.

GLCM texture features are limited to textures that are visually separable, images without large amounts of noise, and a high correlation between Haralick features like contrast, entropy, correlation, and homogeneity [28]. The GLCM algorithm, on the other hand, is simple to implement and has produced very impressive results in many applications like crop disease classification and medical image analysis and classification [28].

In several practical scenarios, the amount of image data available to train a model is limited. In most cases, a small set of samples can be used to model the relationship between training data size and model performance. This is particularly true in cases where
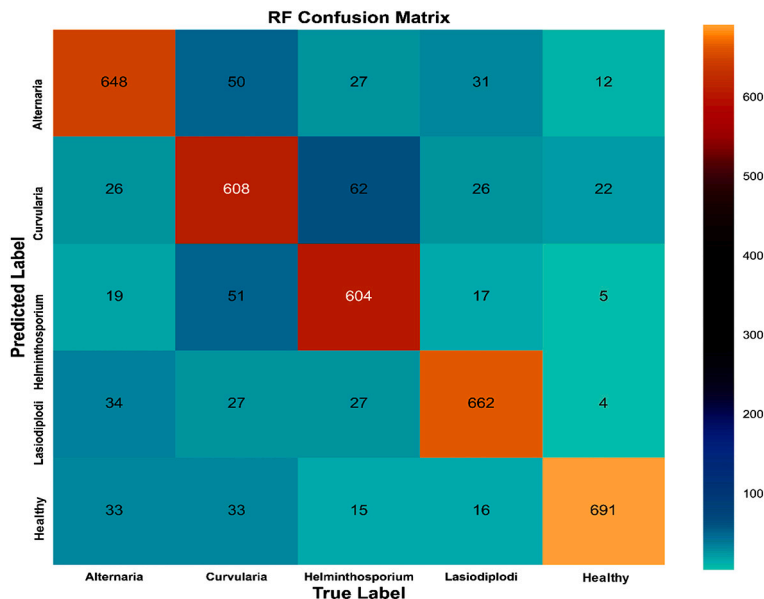
Fig. 6. RF confusion matrix summarizing correct and incorrect classification.
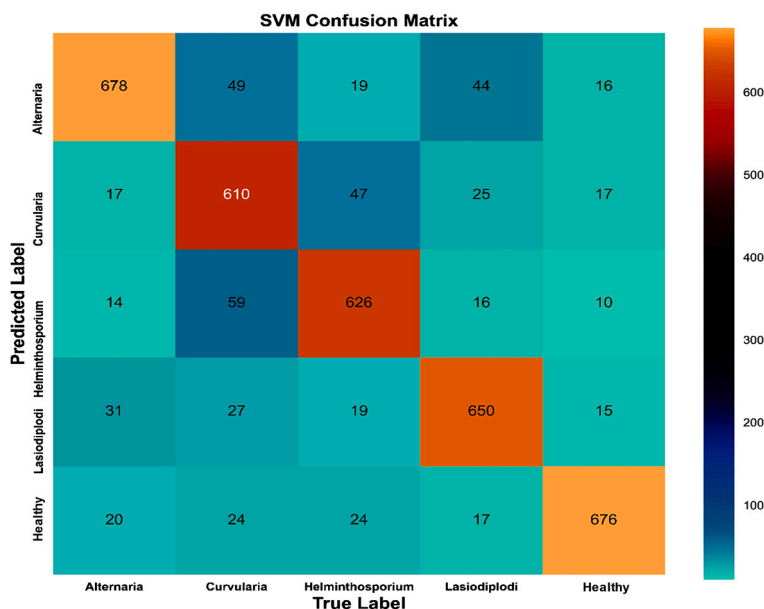

Fig. 7. SVM confusion matrix summarizing correct and incorrect classification.

dataset creation is expensive. This study used a balanced subsampling scheme to determine the minimum training sample required by the model as presented in Fig. 12.

The ANN model was trained on four random training set subsample splits of 25% (2812 images), 50% (5625 images), 75% (8437 images), and the entire training set (11250 images). The model was evaluated on an independent test set and recorded the accuracy values. In Fig. 12, we can observe that despite using only 25% of the training set, the model achieved an overall accuracy of 82%. However, model accuracy increased as the training set also increased. This implies that more images are needed to efficiently train the models.

5. Conclusion and future works

In this study, we explored the potential of using GLCM-based texture features: ASM, Correlation, Contrast, Dissimilarity, Energy, and Homogeneity for the classification of tomato fungal diseases: Alternaria, Curvularia, Helminthosporium, and Lasiodiplodi using the classifiers: ANN, KNN, RF, and SVM. Each textural feature was generated from each gray-scale image with distance at 1, 3, and
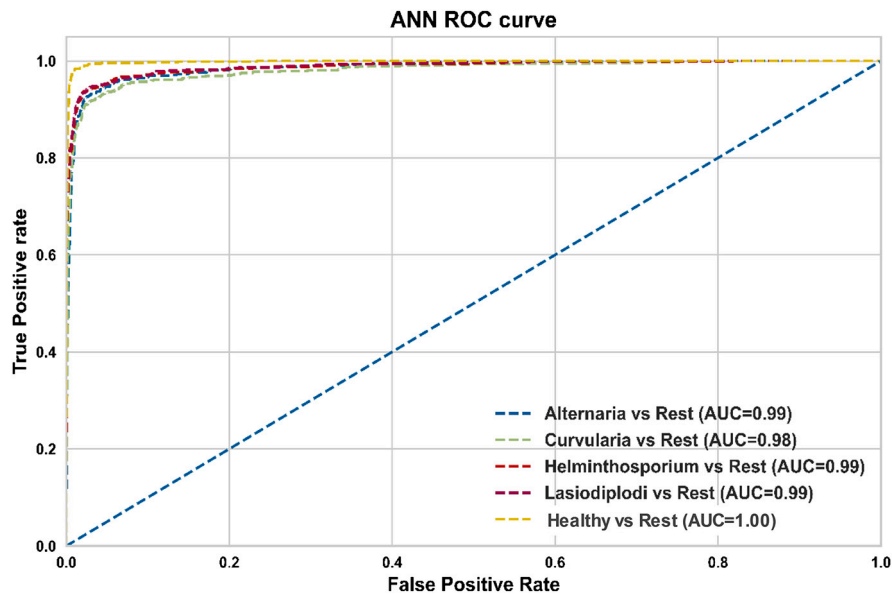
ANN ROC curve

Fig. 8. ANN Receiver Characteristic Operator curve.
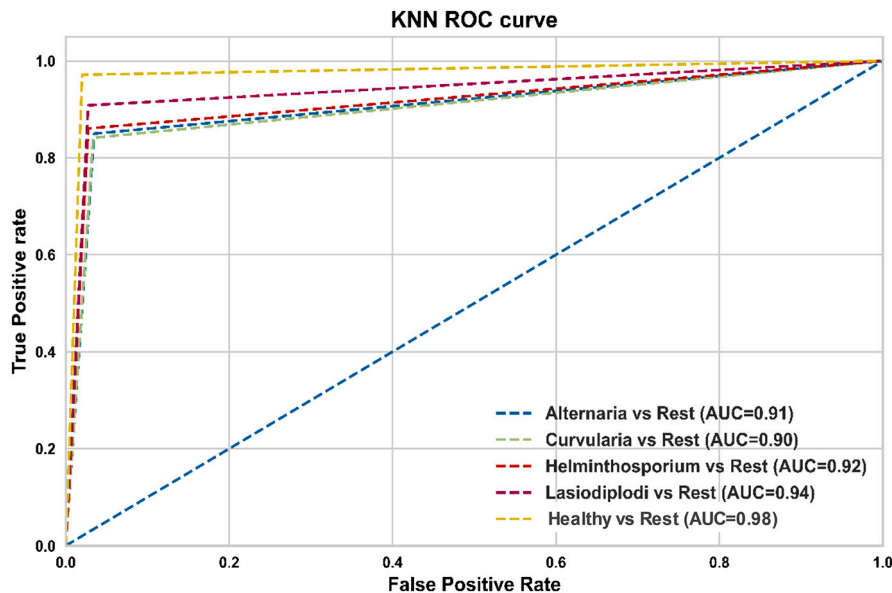
KNN ROC curve

Fig. 9. KNN Receiver Characteristic Operator curve.

5 and angle at 0, 45°, 90°, 135°, 180°, 225°, 270° and 315°. The ANN outperformed other classifiers with an overall accuracy of 94% and average scores of 93.6% for Precision, 93.8% for Recall, and 93.8% for F1-score. Generally, the models confused samples originally belonging to Helminthosporium with Curvularia. The extracted texture features show great potential to classify the different fungal tomato diseases.

To improve on the current study, in our future work we intend to explore classification using Convolutional Neural Networks. Currently, several studies related to image processing are using Convolutional Neural Networks methodologies to obtain better and more reliable results.

CRediT authorship contribution statement

Chimango Nyasulu: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Validation, Visualization. **Awa Diattara:** Conceptualization, Formal analysis, Methodology, Writing – review & editing, Supervision. **Assitan Traore:** Conceptualization, Formal analysis, Methodology, Writing – review & editing, Supervision. **Cheikh Ba:** Funding acquisition, Project administration, Supervision, Writing – review & editing. **Papa Madiallacké Diedhiou:** Conceptualization, Data curation, Supervi-
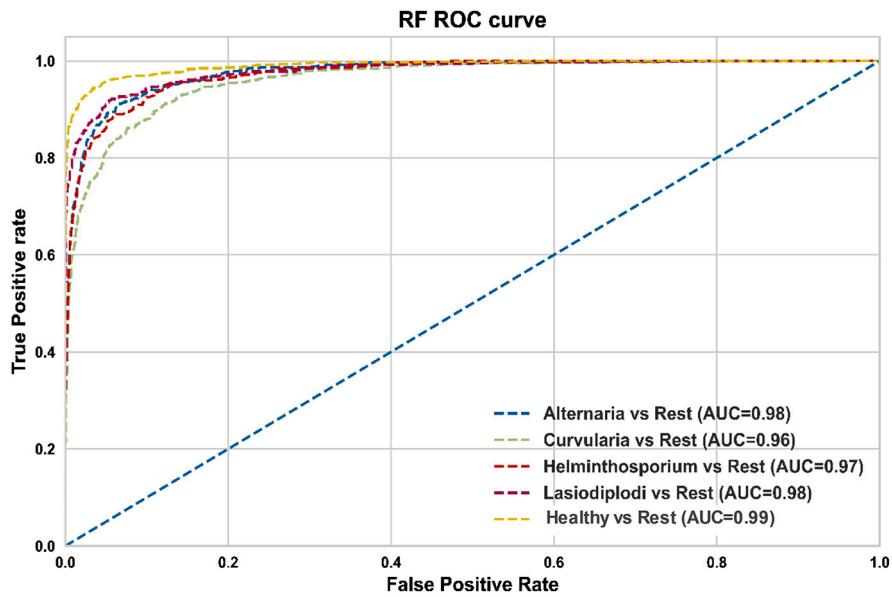
Fig. 10. RF Receiver Characteristic Operator curve.
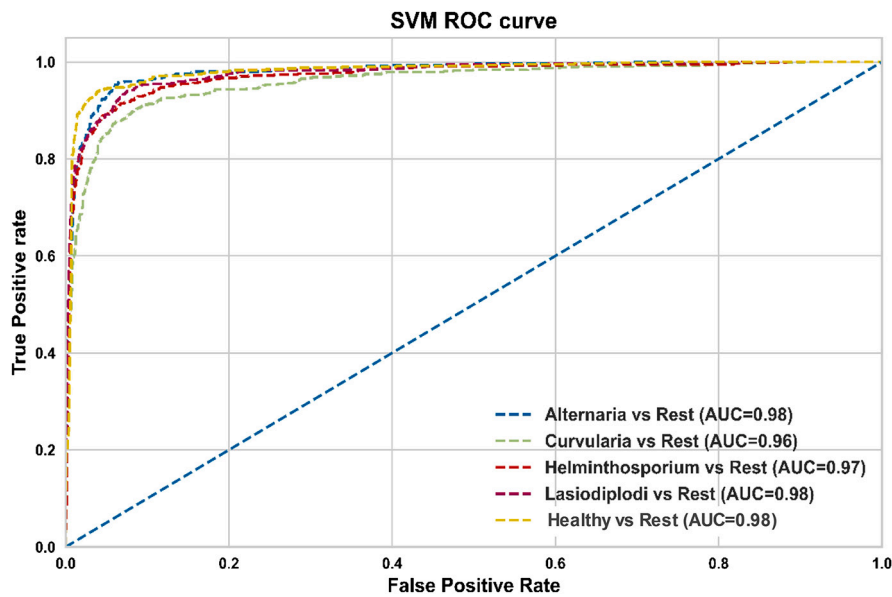

Fig. 11. SVM Receiver Characteristic Operator curve.

sion, Writing – review & editing. **Yakhya Sy:** Conceptualization, Data curation, Methodology, Validation, Writing – review & editing. **Hind Raki:** Methodology, Validation, Visualization, Writing – review & editing. **Diego Hernán Peluffo-Ordóñez:** Methodology, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

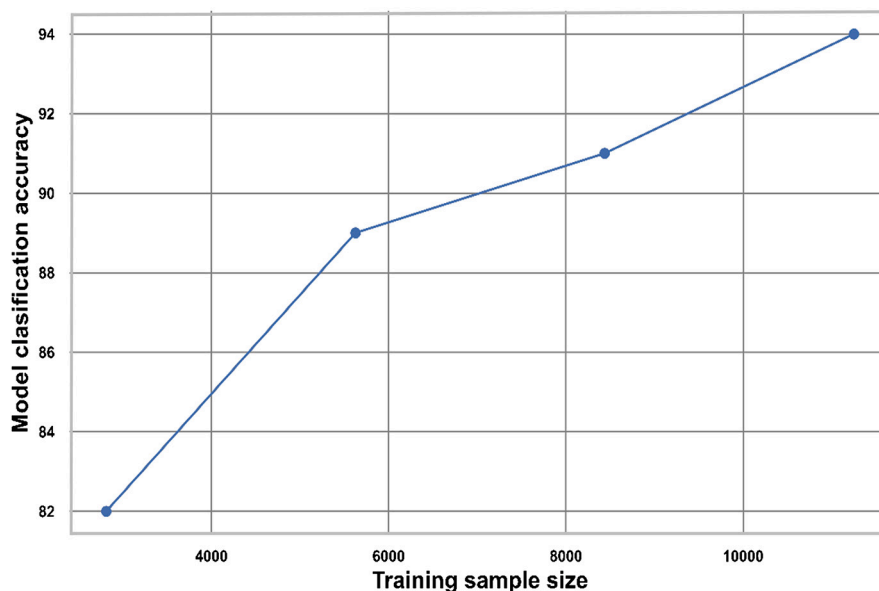## Data availability

Data will be made available on request.

Fig. 12. ANN model accuracy vs sample size.

Acknowledgement

References

[1] S.F. Ashebir, Review of the role of orphan crops in food security, Res. World Agric. Econ. 3 (2) (2022) 1–14, https://doi.org/10.22004/ag.econ.321873.

[2] Crop Diseases: the Nightmare of Every Farmer [Online]. Available: https://www.agrivi.com/blog/crop-diseases-the-nightmare-of-every-farmer/. (Accessed 14 November 2022).

[3] D.M.S. Baptista, M. Farid, D. Fayad, L. Kemoe, L.S. Lanci, P. Mitra, T.S. Muehlschlegel, C. Okou, J.A. Spray, K. Tuitoek, F.D. Unsal, Climate Change and Chronic Food Insecurity in Sub-Saharan Africa, Departmental Papers, 2022.

[4] M.S. Rahman, D.C. Acharjee, Impact of off-season summer tomato cultivation on income and food security of the growers, in: Agricultural Economics, 2020, pp. 1–11.

[5] M. Tall, M.B. Sylla, I. Diallo, et al., Projected impact of climate change in the hydroclimatology of Senegal with a focus over the Lake of Guiers for the twenty-first century, Theor. Appl. Climatol. 129 (2017) 655–665, https://doi.org/10.1007/s00704-016-1805-y.

[6] H. Orchi, M. Sadik, M. Khaldoun, On using artificial intelligence and the Internet of things for crop disease detection: a contemporary survey, Agriculture 12 (1) (2021), https://doi.org/10.3390/agriculture12010009.

[7] M.E. Chowdhury, T. Rahman, A. Khandakar, et al., Automatic and reliable leaf disease detection using deep learning techniques, AgriEngineering 3 (2) (2021) 294–312, https://doi.org/10.3390/agriengineering3020020.

[8] J. Basavaiah, A. Arlene Anthony, Tomato leaf disease classification using multiple feature extraction techniques, Wirel. Pers. Commun. 115 (1) (2020) 633–651, https://doi.org/10.1007/s11277-020-07590-x.

[9] C.U. Kumari, S.J. Prasad, G. Mounika, Leaf disease detection: feature extraction with K-means clustering and classification with ANN, in: 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), IEEE, 2019, pp. 1095–1098.

[10] U. Mokhtar, N.E. Bendary, A.E. Hassenian, et al., SVM-based detection of tomato leaves diseases, in: Intelligent Systems, Springer, Cham, 2015, pp. 641–652.

[11] V.K. Shrivastava, M.K. Pradhan, Rice plant disease classification using color features: a machine learning paradigm, J. Plant Pathol. 103 (2021) 17–26, https://doi.org/10.1007/s42161-020-00683-3.

[12] L.S. Pinto, A. Ray, M.U. Reddy, P. Perumal, P. Aishwarya, Crop disease classification using texture analysis, in: 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE, 2016, pp. 825–828.

[13] A. Singh, H. Kaur, Potato plant leaves disease detection and classification using machine learning methodologies, IOP Conf. Ser., Mater. Sci. Eng. 1022 (1) (2021) 012121, https://doi.org/10.1088/1757-899X/1022/1/012121, IOP Publishing.

[14] L.Y. Kokaeva, A.F. Belosokhov, L.Y. Doeva, E.S. Skolotneva, S.N. Elansky, Distribution of Alternaria species on blighted potato and tomato leaves in Russia, J. Plant Dis. Prot. 125 (2018) 205–212, https://doi.org/10.1007/s41348-017-0135-3.

[15] D. Zhang, K. Zhu, X. Shen, et al., Two interacting ethylene response factors negatively regulate peach resistance to Lasiodiplodia theobromae, Plant Physiol. (2023) kiad279, https://doi.org/10.1093/plphys/kiad279.

[16] H.A.S. AbdElfatah, N.M. Sallam, M.S. Mohamed, H.M.K. Bagy, Curvularia lunata as new causal pathogen of tomato early blight disease in Egypt, Mol. Biol. Rep. 48 (2021) 3001–3006, https://doi.org/10.1007/s11033-021-06254-8.

[17] M.M. El-Fawy, M.A.A. El-Said, Effect of foliar application of some zinc and phosphorus sources on controlling Helminthosporium leaf spot disease and production of sesame, J. Plant Prot. Pathol. 9 (3) (2018) 201–207, https://doi.org/10.21608/jppp.2018.41386.

[18] S. Barberini, C. Forti, M. Laura, et al., An optimized protocol for in vitro regeneration of Ocimum basilicum cv, FT Italiko, Horticulturae 9 (3) (2023) 407, https://doi.org/10.3390/horticulturae9030407.

[19] Z. Li, R. Paul, T. Ba Tis, et al., Non-invasive plant disease diagnostics enabled by smartphone-based fingerprinting of leaf volatiles, Nat. Plants 5 (8) (2019) 856–866, https://doi.org/10.1038/s41477-019-0476-y.

[20] P.K. Mall, P.K. Singh, D. Yadav, Glcm based feature extraction and medical x-ray image classification using machine learning techniques, in: 2019 IEEE Conference on Information and Communication Technology, IEEE, 2019, pp. 1–6.

[21] C.K. Nwagu, O.C. Omankwu, H. Inyiama, Knowledge discovery in databases (KDD): an overview, Int. J. Comput. Sci. Inf. Secur. 15 (12) (2017) 13–16.

[22] M.A. Chandra, S.S. Bedi, Survey on SVM and their application in image classification, Int. J. Inf. Technol. 13 (2021) 1–11, https://doi.org/10.1007/s41870-017-0080-1.

[23] N. Iqbal, R. Mumtaz, U. Shafi, S.M.H. Zaidi, Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms, PeerJ Comput. Sci. 7 (2021), https://doi.org/10.7717/peerj-cs.536.

[24] M. Pawelczyk, S. Bielawski, J.V.D. Heuvel, T. Richter, G. Kasneci, Carla: a python library to benchmark algorithmic recourse and counterfactual explanation algorithms, arXiv preprint, arXiv:2108.00783, 2021, https://doi.org/10.48550/arXiv.2108.00783.

[25] K. Taunk, S. De, S. Verma, A. Swetapadma, A brief review of nearest neighbor algorithm for learning and classification, in: 2019 International Conference on Intelligent Computing and Control Systems (ICCS), IEEE, 2019, pp. 1255–1260.

[26] Python Software Foundation, Python programming language, https://www.python.org/. (Accessed 15 May 2023).

[27] D.M. Powers, Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation, arXiv preprint, arXiv:2010.16061, 2020, https://doi.org/10.48550/arXiv.2010.16061.

[28] A. Humeau-Heurtier, Texture feature extraction methods: a survey, IEEE Access 7 (2019) 8975–9000, https://doi.org/10.1109/ACCESS.2018.2890743.