International Journal of Formal Sciences: Current and Future Research Trends (IJFSCFRT)

ISSN (Print) , ISSN (Online)

https://ijfscfrtjournal.isrra.org/index.php/Formal_Sciences_Journal/index

Comparative Analysis of Machine Learning Algorithms Accuracy for Maize Leaf Disease Identification

Vincent Mbandu Ochango[a]*, Geoffrey Mariga Wambugu[b], John Gichuki Ndia[c]

[a,b,c]*Murang'a University of Technology, 75, Murang'a and 10200, Kenya*

[a]*Email: ochangovincent@gmail.com* , [b]*Email: gmariga@mut.ac.ke* , [c]*Email: jndia@mut.ac.ke*

Abstract

The number of data points predicted correctly out of the total data points is known as accuracy in image classification models. Assessment of the accuracy is very important since it compares the correct images to the ones that have been classified by the image classification models. Image classification accuracy is a challenge since image classification models classify images to the class they don't belong to hence there is an inaccurate relationship between the predicted class and the actual class which results in a low model accuracy score. Therefore, there is a need for a model that can classify the images with the highest accuracy. The paper presents image classification models together with the feature extraction methods used to classify maize disease images. The researcher used an augmented maize leaf disease dataset obtained from the Kaggle website. Features are extracted from maize disease images and passed to the machine learning classification algorithm to identify the possible disease based on the features detected using the feature extraction method. The maize disease images used include images of common rust, leaf spot, and northern leaf blight and healthy images. An evaluation was done for the feature extraction methods and the outcomes revealed Histogram of Oriented Gradients performed best with classifiers compared to KAZE and Oriented FAST and rotated BRIEF. The experimental outcome also indicated that the Artificial Neural Network model had the highest accuracy of 0.82 compared to Logistic Regression, K-Nearest Neighbors, Random Forest, Linear Support Vector Classifier, Decision Tree, and Support Vector Machine.

Keywords: Feature Descriptor; Gradient Direction; Gradient Magnitude; Machine Learning; Cross-Validation; Overfitting; Artificial Neural Network; Support Vector Machine.

1. Introduction

The purpose of this paper is to identify maize disease through feature extraction and classify the maize disease images from the features extracted using machine learning algorithms. The farmers are usually unable to detect diseases on their crops by just looking at them. This leads to damages that cost farmers a lot of money.

----

* Corresponding author.

Using captured images of crops to tell whether or not they are disease through image classification using a machine learning algorithm and if they are disease, the machine learning algorithm to tell the particular disease affecting the plant is the solution to this problem [1]. The farmer can then purchase the right medicine for their plants. From this research paper, the features from the images are extracted using ORB, HOG, and KAZE method, and once the features are extracted, they are passed to the machine learning image classification algorithm which can tell the particular maize disease affecting the crops [2]. A comparison of three methods was done and the HOG feature extraction method performed better with image classification algorithms hence the researcher decided to work out with HOG as a feature extraction method. HOG feature descriptor extracts key points from images and throws away information that is not useful and this is what is considered dimensionality reduction [3]. These key points are the ones that differentiate an image from the other images since they are unique for every image and clearly distinguish an image from the other images. The feature descriptor converts an image to a vector which is an array and this feature vector is an input value to the classification algorithms. Before the feature extraction method calculated the descriptor, the image window was resized to an aspect ratio of 1:2 and most probably 64 × 128 and this process was known as image preprocessing [3]. The main reason for resizing the image to 64 × 128 sizes was that when extracting features, the image needs to be divided into a patch of 8 × 8 and 16 ×16. The histogram of Gradient was calculated by first calculating the vertical and horizontal gradient which was achieved by applying filters to an image. A lot of unnecessary information such as colored background was removed by gradient image and only the shape and the edges of the image remained. Other feature descriptors usually recognize if an element in an image is an edge or not in the case of edge features but the proposed feature extraction method went further and extracted the magnitude and direction of the edges thus being able to provide the edge direction. Calculating gradient meant calculating the direction of x and y pixel values for the image. A patch was taken from an image and a gradient was calculated for the patch taken. The pixel matrix was generated for each small patch taken from an image [4]. For every pixel value in the matrix, the researcher calculated the change in x and y direction which was denoted by Gx and Gy respectively. And the process gave us the new matrices, one storing Gx and the other one storing Gy. The step that followed was to find the direction and magnitude of all elements in an image. And the process was done by calculating the Total Gradient Magnitude (T.G.M). And the following equation helped in calculating the total gradient magnitude;

T.G.M= √ [((Gy)2 + Gx)2 +]

The following mathematical equation shows how the direction of the pixel was calculated;

Ө=arctan (Gx / Gy)

Finally, the histogram was calculated for each pixel using the magnitude and the direction of each pixel [5]. The features extracted are the ones that acted as the input value for the proposed image classification model.

The image classification algorithms used include Logistic Regression, K-Nearest Neighbors, Artificial Neural Network, Linear Support Vector Classifier, Decision Tree, and Support Vector Machine. The comparative analysis of the image classification models was done to investigate which model classifies maize leaf diseases accurately. The comparative analysis is also done to find out which best performing models will be combined to come up with a stronger learner that can classify maize leaf disease with the highest accuracy [4]. The rest of this paper is structured as follows; Section II provides related work on feature extraction methods and image classification algorithms. Section III explains how the feature extraction was done, the hyperparameters used with classification algorithms, and how the cross-validation was done to reduce the overfitting of the classification algorithms. Section IV provides an explanation of the experimental results obtained and the best classifier. Section V the conclusion and future work.

2. Related Work

The use of feature extraction methods in this paper as the input to the machine learning algorithms to identify maize disease images is widely used. The metrics used to measure the classification of disease images differ for each machine learning algorithm. By using computer vision different methods are utilized in the identification of crop infections. Extracting features from images is one of the techniques that is used to detect diseases from the plant.

The maize disease images were classified by carrying out an experiment that uses ANN and Support Vector Machine. The two classifiers were trained based on image features extracted using a feature extraction method. The results demonstrated that SVM performed best with the image features extracted compared to ANN. The SVM classifier had an accuracy of 0.9217 and 0.874 for the ANN classifier [5].

According to [6], the researcher identified three kinds of cereal plants and classified them using machine learning algorithms. Specifically, they used two classifiers to classify jowar, maize, and wheat leaf diseases by using the fungal symptoms associated with each leaf disease. Normal, smut, powdery mildew, leaf spot, and leaf blight maize leaf disease were collected by the authors and were used in the experiment. The authors followed certain steps to identify and categorize fungal disease symptoms; acquired normal and fungal affected 750 JPG formatted images. After that, the images are preprocessed and then the image segmentation is done. The leaf disease images are used to extract features using the Color Co-occurrence matrix algorithm and the features extracted act as the input value to the machine learning algorithms and for program interface MATLAB tool was used. The classification accuracy for SVM and ANN machine learning algorithms used was 83.83% and 77.75% respectively. To identify and classify cereals' fungal disease the authors found out that the SVM algorithm is the best to use since it is more accurate than the Artificial Neural network. Other feature extraction methods and machine learning algorithms for classifying leaf disease images were recommended by the researcher as future work that needs to be done [7].

The machine learning algorithms experiment was conducted to categorize 5 types of maize crop diseases. The researchers collected 20 images for each category of maize leaf disease which aided in the experiment and all the images collected were used both for training and testing purposes [8]. However, the authors did not mention the five-leaf diseases used in the experiment. The images collected were scaled and normalized in terms of orientation and histogram equilibrium was used to convert them to 32 by 32 pixels with each image having a white background due to each pixel having a 255 gray level. The experiment was conducted by first collecting the images of five different types of leaf diseases by using digital cameras and then the images were segmented. Features were then extracted from images and passed onto the KNN algorithm which classified the features according to the respective leaf diseases hence producing class labels for each image feature. The experiment for image classification was done 50 times and the results showed the classification accuracy was above 80%. The researcher finally proposed that future work should be done by increasing the training data set and extracting key points from an image since the key points clearly distinguish an image from one another [8].

According to [9] did research in china farm area that classifies four types of maize leaf diseases and the researchers followed these steps; Under sunlight conditions, the researchers used digital cameras to collect JPG types of maize disease images and to obtain the information from the images they are converted to BMP format and later used a thresholding value to segment the images. The standard deviation and mean are calculated after the images are converted from RGB to HIS and the researchers finally classified the images using the GA-SVM algorithm. The maize leaf diseases were classified also using support vector machine and RBF kernel function. The machine learning algorithm which classified the maize leaf disease was measured in terms of precision and the GA-SVM algorithm had a precision of between 88.72% and 92.59% for each image for maize leaf disease and the support vector machine had a precision of between 69.63% and 90.09% [10].

Mohammad, Sayeed, and Billah developed a model in 2019 for plant disease detection using the HOG and LBP feature extraction method together with the Support Vector Machine image classification algorithm. A public dataset known as Flavia leaf dataset was used in their experiment both for training and testing. The support vector machine is used to classify the images and the image features are extracted using the Local Binary Pattern and Histogram of Oriented Gradient. The accuracy of 0.9125 was achieved when a hybrid of Local Binary Pattern and Histogram of Oriented Gradient was used to extract image features and the Support Vector Machine was used to classify the images. When the Local Binary Pattern was used to extract features and when the features extracted were used with the Support Vector Machine an accuracy of 0.406 was obtained and this indicated that the Local Binary Pattern feature extraction method performed badly with the Support Vector Machine classifier. Finally, the accuracy of 0.8531,0.8125,0.775 was obtained using a cell size of 8 by 8, 4 by 4, and 2 by 2 respectively when Support Vector Machine classifier was used together with a Histogram of Oriented Gradient as a feature extraction method. It was realized that when the Histogram of oriented gradient feature extraction method is used on an image cell size of 8 by 8 together with the Support Vector Machine classifier, the best accuracy is obtained. Hence the experimental results indicated that a hybrid of Local Binary Pattern and Histogram of Oriented Gradient for image feature extraction is more effective when used with Support Vector Machine Classifier [13]. Many experiments have been done on image classification using support vector machine algorithm and even accuracy and precision have been measured, further research needs to be done with other machine learning algorithms to be able to verify if really support vector machine is best to be used in image classification. Also, a further experiment needs to be done to assess the machine learning algorithms in terms of recall and f1-score. This has propelled the current study to increase the training dataset and use more algorithms to come to a conclusion on which algorithm is best when it comes to maize leaf disease classification. And the current study also wants to explore if there are other better algorithms than support vector machines when it comes to maize leaf disease classification.

3. Methodology

This chapter explains the data set used during the experiment together with the number of images used during training and testing. It also explains how the features were extracted and passed to the machine learning algorithms for classification.

3.1 Dataset Description

The research used an augmented maize disease dataset which is public and was obtained from the Kaggle website and the data set contained training and testing images. The dataset consisted of common rust, leaf spot, northern leaf blight disease images, and healthy images. The whole training dataset consisted of 7308 images,1634 images for leaf spot, 1907 images for common rust,1908 images for northern leaf blight, and 1859 for healthy leaf images. The whole testing dataset consisted of 1826 images, 407 images for leaf spot, 477 images for common rust, 477 images for northern leaf blight, and 465 for healthy leaf images. And because of time and limited resources, the researcher decided first to work with 200 images from the training data set which was used for training for each category of disease resulting in 800 images in total. This is because the process of generating features from each image takes a lot of time and consumes a lot of computer resources hence the researcher decided first to work with a total of 800 images. The testing data set had a total of 1826 images and 50 images were used for each category of maize leaf disease.

3.2 Numerical Feature Extraction

Feature extraction from the image was done using ORB, KAZE, and HOG feature extraction methods. The features extracted were in terms of integers and these integers were passed to the machine learning classifier algorithm [11]. The 4096 key points were extracted from each maize leaf disease category and acted as input values to the machine learning classifiers. These attributes extracted were key points to each image since the key point is a feature that is unique to an image and can be detected despite the change in the image. The features extracted by each method were passed to the classification algorithms to find out which feature extraction method performs better with machine learning algorithms. The machine learning algorithms were measured in terms of accuracy to determine which feature extraction method works best with them. Based on the results and analysis of the classification algorithms it was found out that HOG performs better with the classification algorithms. This made the researcher propose that indeed HOG methods are the best method to use while extracting maize leaf disease features compared to KAZE and ORB. HOG divides the image into smaller parts or an image patch of 8 × 8 cells and calculates the features for this every 8 × 8 cells which represent the histogram for the whole image. A matrix of 9 × 1 is usually obtained for each cell after generating the histogram from an image divided into 8 × 8 cell [12]. The histogram is finally normalized after extracting HOG features from 8 × 8 cells. Assuming a feature vector V, the noramalized features are calculated as shown below.

$V = [a1, a2, a3 \ldots a36]$

We determine the root of the sum of squares:

K= √ (a1)2 + (a2)2 + (a3)2 + … (a36)2

All the values in the vector V are divided with this value k

The normalized vector = $\left(\frac{a1}{k}, \frac{a2}{k}, \frac{a3}{k}, \ldots \frac{a36}{k} \right)$

And the normalized HOG features acted as the input value to the image classification models. The below algorithm illustrates how the HOG feature extraction method works.

Algorithm 1: Histogram of Oriented Gradient

Input: Image

Output: HOG features

Begin Algorithm

1. Read Image in the form of pixels
2. Extract the HOG features:
   - 2.1 Divide the image pixel window by 8x8 pixel cells and calculate the gradient components with respect to each pixel (x, y) in vertical and horizontal directions
   - 2.2 Calculate number of blocks in vertical and horizontal directions taking block step size of 8 pixels.
   - 2.3 Histogram of 9 gradient directions is then calculated for each cell and extract feature vectors i.e., HOG features
   - 2.4 Save all the features in a matrix

End Algorithm

Figure 3.1: HOG Feature Extraction Algorithm.

3.3 Hyperparameter Tuning

This was done to be able to work with a set of optimal hyperparameters for each classification algorithm. To work with classification algorithms, you need to set parameters for each algorithm before the process of learning begins. The penalty in logistic regression and loss in stochastic gradient descent is some of the examples of hyperparameters for the classification algorithm. The tuning strategy used for our case for optimizing hyperparameters for each classification algorithm was grid search [13]. Hyperparameter optimization according to research usually improves the performance of the machine learning algorithm. The grid search method used for hyperparameter tuning works by exhaustively searching through a specified set of hyperparameters. The optimal combination of parameters supplied is guaranteed by using a grid search and one of the major disadvantages of grid search is that it is computationally expensive and time-consuming. During the implementation of hyperparameter tuning is that before we ran the grid search method in the Jupiter notebook, we first defined our grid of parameters to search over [14].

3.4 Cross-Validation

The validation was done to test the classifiers if it performs well on the data that it has never seen before introducing the classifier to the test data set. Cross-validation is done to get an assurance that your classification algorithm works better and predicts correctly in case it is given data that it has never seen before. The method also helps you to know if the classification algorithm is either underfitting or overfitting the data.

Using part of the training data for validating your model usually results in an underfitting problem since there is never enough data for training your model. This, in turn, increases error induced by bias and we risk losing important trends in training data set and patterns which results after reducing the training data. Hence, we require K-Fold cross-validation which leaves the part of the data for validation and the other for training. The method puts together a k-1 subset to be used for training and a k subset to be used for testing or validation purposes. The total effectiveness of the model is obtained by averaging error estimation for all k trials. With this method, every data gets to be in training set k-1 times and gets to be in the testing set once. This reduces variance significantly since most of the data is used in the validation set and most of the data is used in the training set which reduces bias significantly. The effectiveness of this method is seen since it interchanges the test set or validation set with the training set [15].

Steps followed;

i. Unsystematically interchange the set of data.
ii. Break the set of the data into groups (K groups)
iii. For every distinctive group;
   a. Let k subset be used as a testing or validating set
   b. Let the other k-1 subset be put together to act as the training set
   c. Discard the model and withhold the evaluation score
iv. Obtain the total effectiveness of the model by averaging the error estimation of all the k trials.

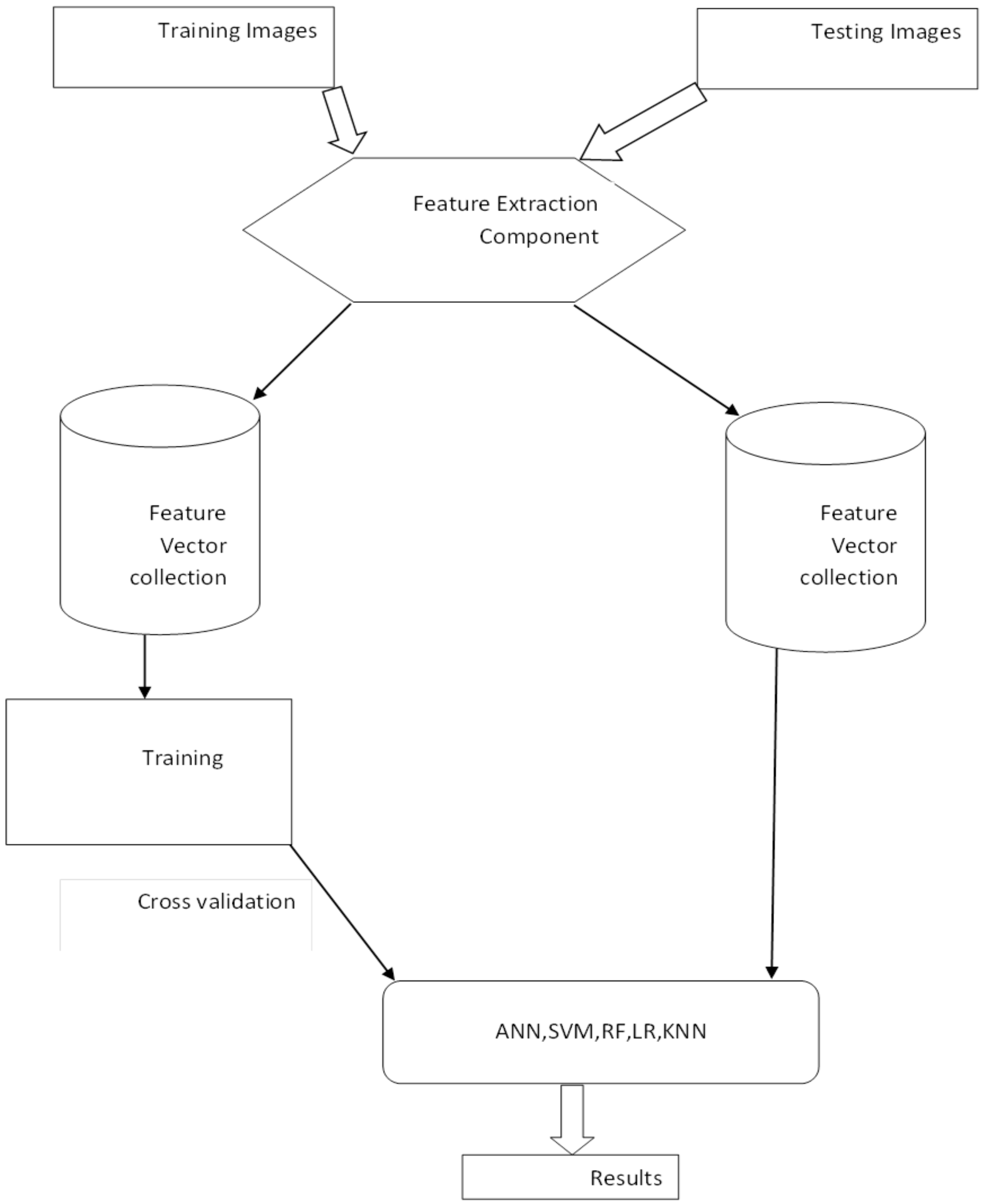

Figure 3. 2: Conceptual Model for Classifying Maize Leaf Diseases.

4. Experimental Results and Discussion

4.1 Feature Extraction

Fig 4.1 illustrates how features are generated using the ORB method for common rust disease image

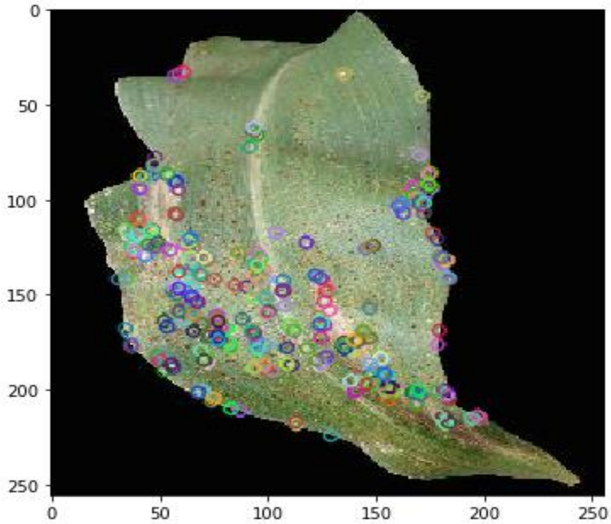

Figure 4. 1: Detecting key points from common rust disease image using ORB method.

Feature generation using the KAZE method for common rust disease image is shown in Fig 4.2.

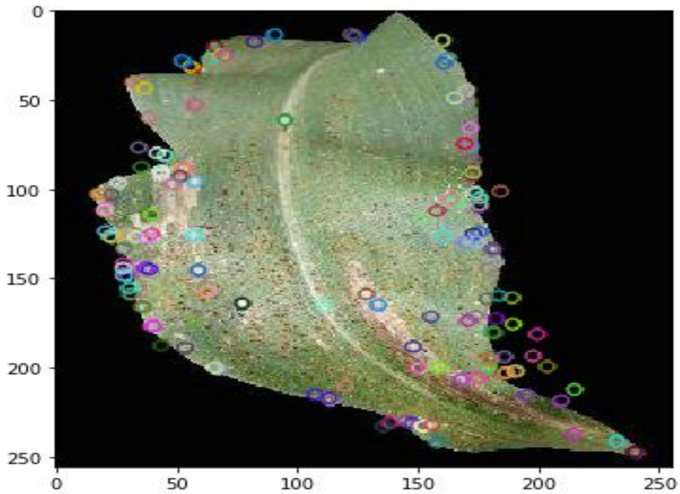

Figure 4.2: Detecting key points from common rust disease image using KAZE method.

The dataset had images for common rust, leaf spot, northern leaf blight disease, and images for healthy leaves. The features extracted from each category of maize leaf disease images were 4096 features and the research narrowed down to 200 images of each disease category since extracting features for the whole images in the dataset takes a lot of time. After the features were extracted using the three feature extraction methods, the machine learning algorithms were trained based on the features extracted. The training was done basically to associate each feature with their respective disease type and hence making the algorithms learn from the training data.

4.2 Imaage Classification Models

The image classification models used for the comparison of their performance were Random Forest, Decision Tree, Logistic Regression, K-Nearest Neighbors, Artificial Neural Network, Linear SVC, and Support Vector Machine. The researcher looked at the accuracy level of each classifier based on each feature extraction method and this was to be able to identify which feature extraction method performs better with the classification models. Table 1 shows the image classification accuracy values for various models based on features extracted using KAZE, ORB, and HOG, respectively.

Table 1: Classifier Accuracy using features generated by KAZE, ORB, and HOG methods.

Model	KAZE	ORB	HOG
Artificial Neural Network	0.716	0.443	0.82
Support Vector Machine	0.706	0.423	0.8
Logistic Regression	0.695	0.361	0.78
Linear SVC	0.690	0.361	0.715
Random Forest	0.675	0.376	0.78
K-Nearest Neighbors	0.609	0.376	0.685
Decision Tree	0.579	0.289	0.405
Average Accuracy:	0.667	0.376	0.712

As seen in Table 1, the base models perform so badly with the ORB feature extraction method. The HOG produces the best performance as shown by the average performance of 0.712, followed by KAZE at 0.667, and ORB at 0.376. Based on the HOG performance and Artificial Neural Network model, it was decided to work with the HOG method for feature extraction and Artificial Neural Network for image classification in future experiments.

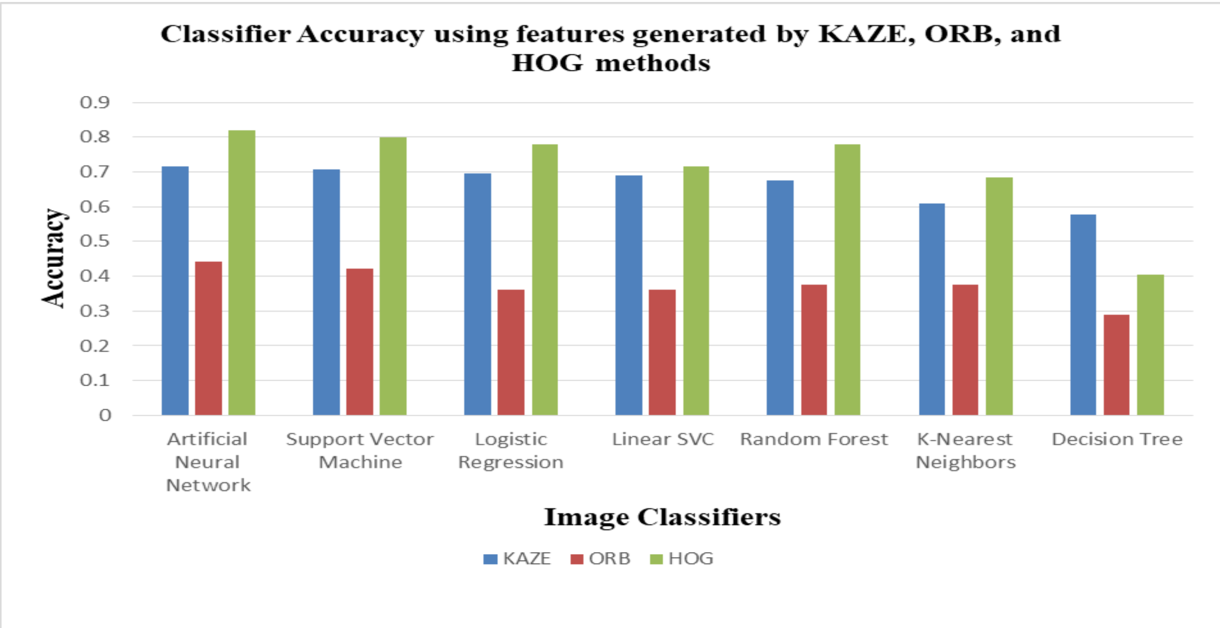

Figure 4.3 : Classifier Accuracy using features generated by KAZE, ORB, and HOG methods.

The results in Figure 4.3 above show Artificial Neural Network as the best-performing image classification model when subjected to the test data set and this is because the model generalizes the training data well since the data is distributed across the entire network.

4.3 Discussion

The researcher used the maize leaf disease dataset obtained from the Kaggle website and the training data set consisted of 7308 images. The whole testing dataset consisted of 1826 images. Therefore, using all the training data set images was a challenge since training the image classification models on all the images required a computer with very high specification and it would take a lot of time to train the models on all the images in the training dataset. To overcome this challenge the researcher used only the images that the computer used during the experiment supported. The experimental results shown above were mainly used to investigate and compare different image classification models and provide a way forward on the classification model to be used in identifying maize leaf diseases. Figure 4.1 and Figure 4.2 are used to demonstrate how feature extraction methods are used to extract key points from an image. These key points are the ones that act as input values to the machine learning algorithms. Since after the features are extracted, they are passed to the classification models. Table 1 shows how three feature extraction methods performed with different classification algorithms. As it can be seen from table 1 Artificial Neural Network performs best with HOG, KAZE, and ORB methods. The average accuracy for each feature extraction method on how it performed with different classification models was calculated and the results showed that HOG performed well. It performs well than the other feature extraction methods because;

- It mainly concentrates on the shape of the image and be able to offer the direction of the edges.
- The direction and magnitude of the edges are calculated after dividing the image into smaller parts called cells.
- The Histogram is developed from the smaller parts of the images by using the magnitude and direction of the pixel values.
- It is more accurate because the gradients are usually normalized since in the 8×8 cells some portion of the image usually appears bright than the other portion hence the gradients are normalized by taking 16×16 blocks which help in reducing the variation in light.

Due to HOG performance, it was concluded to work with it in future experimental work that involves feature extraction from maize leaf disease images.

Figure 4.3 shows that the Artificial Neural Network classifier emerged the best in terms of classifying the maize leaf disease images. The main reason why the Artificial Neural Network produced good results is that Artificial Neural Networks mostly uses unweighted averaging as a hybrid approach and this is achieved by using predicted unweighted average scores for all single learners and recognizing it as the predicted outcome. ANN has low variance, so taking the average predicted outcomes of multiple Artificial Neural Networks reduces the variance thus producing better results. The alpha parameter in Artificial Neural Network is used to constrain the weights to reduce the overfitting problem hence it is a regularization term. Increasing the value of the alpha parameter reduces the overfitting problem which makes the ANN perform well on unseen data. The forward propagation and backward propagation are done iteratively by ANN to adjust the weights on the network until the right weights are assigned on the network which in turn enables the network to make an accurate prediction. The tolerance parameter for ANN is used to indicate the point where the network has learned and hence training stops at this point indicating the network has reached convergence point hence the network cannot learn beyond this point. And the convergence point is when the Artificial Neural Network can accurately classify maize leaf disease images to the class they belong to. The major disadvantage of the algorithms is that it does not explain clearly through its functions how it arrives at the final output.

5. Conclusion

This paper provides a solution to farmers for them to be able to identify maize leaf diseases through image classification models. The research has found out which feature extraction method can perform better with image classification models. And based on the results the research found out that the HOG feature extraction method performs best with the classification algorithms compared to KAZE and ORB hence informing the researcher to work with the HOG method for future experiments. There was also a comparison of image classification models and the Artificial Neural Network emerged the best. The results indicated that the Artificial Neural Network had an accuracy score of 0.82. Based on the results the researcher proposed working with the HOG feature extraction method and the Artificial Neural Network when it comes to maize leaf disease identification since the classifier produces better results with HOG features. Further investigation needs to be done by increasing the training dataset and the testing data set and using a combination of two classifiers (Artificial Neural Network and Support Vector Machine) together with the HOG method and compare the results and make a new conclusion based on the increased dataset and the combination of the two classifiers.

References

[1] Amato, G., & Falchi, F. (2018, September). kNN based image classification relying on local feature similarity. In *Proceedings of the Third International Conference on SImilarity Search and APplications* (pp. 101-108).

[2] Bosch, A., Zisserman, A., & Munoz, X. (2019, October). Image classification using random forests and ferns. In *2017 IEEE 11th international conference on computer vision* (pp. 1-8). Ieee.

[3] Chen, P. Y., Huang, C. C., Lien, C. Y., & Tsai, Y. H. (2018). An efficient hardware implementation of HOG feature extraction for human detection. *IEEE Transactions on Intelligent Transportation Systems*, *15*(2), 656-662.

[4] Chen, Y. S., Chien, J. C., & Lee, J. D. (2016, September). KAZE-BOF-based large vehicles detection at night. In *2017 International Conference On Communication Problem-Solving (ICCP)* (pp. 1-2). IEEE.

[5] Foody, G. M., & Mathur, A. (2019). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote*

Sensing of Environment, *103*(2), 179-189.

[6] Gan, G., & Cheng, J. (2016, December). Pedestrian detection based on HOG-LBP feature. In *2014 Seventh International Conference on Computational Intelligence and Security* (pp. 1184-1187). IEEE.

[7] Geismann, P., & Schneider, G. (2021, June). A two-staged approach to vision-based pedestrian recognition using Haar and HOG features. In *2011 IEEE Intelligent Vehicles Symposium* (pp. 554-559). IEEE.

[8] Goh, K. S., Chang, E., & Cheng, K. T. (2017, October). SVM binary classifier ensembles for image classification. In *Proceedings of the tenth international conference on Information and knowledge management* (pp. 395-402).

[9] KIM¹, J. I. N. H. O., Kim, B. S., & Savarese, S. (2018). Comparing image classification methods: K-nearest-neighbor and support-vector-machines. In *Proceedings of the 6th WSEAS international conference on Computer Engineering and Applications, and Proceedings of the 2012 American conference on Applied Mathematics* (Vol. 1001, pp. 48109-2122).

[10] Kobayashi, T. (2020). BFO meets HOG: feature extraction based on histograms of oriented pdf gradients for image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 747-754).

[11]K Song, Z liu, et al. A Research of maize disease image recognition of Corn Based on BP Networks. Measuring Technology and Mechatronics Automation China 2018: 246-249.

[12] Li, W., Qian, Y., Loomes, M., & Gao, X. (2020, March). The application of KAZE features to the classification echocardiogram videos. In *International Workshop on Multimodal Retrieval in the Medical Domain* (pp. 61-72). Springer, Cham.

[13] Li, Y., & Cheng, B. (2019, August). An improved k-nearest neighbor algorithm and its application to high resolution remote sensing image classification. In *2009 17th International Conference on Geoinformatics* (pp. 1-4). Ieee.

[14] Millard, K., & Richardson, M. (2018). On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote sensing*, *7*(7), 8489-8515.

[15] PD Jagadeesh, Y Rajesh, et al. Classification of Fungal Disease Symptoms affected on Cereals using Color Texture Features. International Journal of Signal Processing, Image Processing and Pattern Recognition 2017; 6: 321-330.

[16] P Jagadeesh, R Yakkundimath, et al. SVM and ANN Based Classification of Plant Diseases Using

Feature Reduction Technique. International Journal of Interactive Multimedia and Artificial Intelligence 2018; 3: 6-14.

[17] Ramteke, R. J., & Monali, Y. K. (2018). Automatic medical image classification and abnormality detection using k-nearest neighbour. *International Journal of Advanced Computer Research*, *2*(4), 190-196.

[18] Sanchez-Morillo, D., González, J., García-Rojo, M., & Ortega, J. (2021, April). Classification of breast cancer histopathological images using KAZE features. In *International Conference on Bioinformatics and Biomedical Engineering* (pp. 276-286). Springer, Cham.

[19] Spyrou, E., Le Borgne, H., Mailis, T., Cooke, E., Avrithis, Y., & O'Connor, N. (2018, September). Fusing MPEG-7 visual descriptors for image classification. In *International Conference on Artificial Neural Networks* (pp. 847-852). Springer, Berlin, Heidelberg.

[20] Umbaugh, S. E., Wei, Y. S., & Zuke, M. (2019). Feature extraction in image analysis. A program for facilitating data reduction in medical image classification. *IEEE engineering in medicine and biology magazine*, *16*(4), 62-73.

[21] Xia, J., Ghamisi, P., Yokoya, N., & Iwasaki, A. (2021). Random forest ensembles and extended multiextinction profiles for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(1), 202-216.

[22] Xu, B., Ye, Y., & Nie, L. (2018, June). An improved random forest classifier for image classification. In *2012 IEEE International Conference on Information and Automation* (pp. 795-800). IEEE.

[23] Z Zhiyong, H Xiaoyang, et al. Image recognition of maize leaf disease based on GA-SVM. Chemical Engineering Transactions 2019; 46:199-204.