



Development of Pest Control Model (Using Tomato as Case Study)

¹Adeola, O. O., ²Morakinyo, T. O., ³Alowosile, O. Y., ¹Kilani, S. O., ¹Ajayi, O. D. and ¹Kehinde, F. A.

¹Department of Computer Science, Oyo State College of Agriculture and Technology, Igboora.

²Department of Computer Engineering, Oyo State College of Agriculture and Technology, Igboora.

³Department of Computer Science, Abraham Adesanya Polytechnic, Ijebu Igbo, Ogun State.

Corresponding Email: ooadeola@gmail.com; Phone Number: +234 703 326 1184

Abstract

Farmers who practice automation usually use a variety of technologies to boost agricultural productivity, but these have not proven successful. In this work, a Convolutional Neural Network (CNN) model was utilized to forecast tomato health and identify the pest management strategies. The dataset for tomato leaves was obtained from KAGGLE and used for the research. The dataset includes ten (10) different categories for both good and ill tomato leaf diseases. The steps taken to complete the task using the CNN model were: Dataset collection and preparation, image pre-processing, shape feature extraction, building a model to control tomato pests, training, evaluating, and testing the model. On the established dataset, the model was tested. CNN's average recognition accuracy was 97.7%, whereas KNN's accuracy was 76.45%, ANN's accuracy was 70.8%, SVM's accuracy was 79.95%, the naive bayes model was 54.54%, and the accuracy of the logistic regression model was 60.92%. This initiative has produced successful experimental outcomes. It is very important for tomato pest prediction and control to use image recognition studies of tomato pests using CNN model. It is necessary to investigate fresh algorithms that can spur creativity. The model still needs to be effectively optimized in a variety of additional ways.

Keyword: Tomato; CNN; Model; Agriculture; Pest.

Introduction:

Agriculture and food manufacturing are two important industries for a nation survival. Growth in expansion is a crucial factor in the export of high-quality agricultural and food products. (Bhadane, 2019). The expansion of economic earnings in processing nations is mostly influenced by export supply chain profits and market demand. During the production of most foods, feeds, and natural fibers essential to man, pests are continuously destroyed the process. All types of crops, and plants such as flowers, ornamentals, and lawn grasses, are damaged by pests by attacking the roots, stems, leaves, and fruiting parts. Pests are unwanted organisms that are nuisance to man or domestic animals, and can cause injury to humans, animals, plants, and property (Boissard, 2020). Pests can reduce yield and/or quality of harvested products ranging from field crops, fruits and vegetables, to lawns, trees and golf courses. Pests have an impact on agricultural produce growth and quality. The rise in hunger and poverty demonstrated how negatively agriculture had been impacted (Gautam, 2018). Farmers who practice automation usually used a variety of technologies to boost agricultural productivity, but these have not proven successful (Kaya,

2018). In this study, machine learning tare used to forecast tomato health and determine the pest management strategies that threatens the crops so has to reduce the damages.

Material and Methods:

Fact Finding Method: Many techniques were used in the gathering and collection. The project's data and information come from an interview and the Tomato Leaf dataset from KAGGLE.

Interview Method: This were derived by consulting the market authorities of Towobowo Market, Igboora, Reliable facts were gotten based on the questions posed to them by the researcher.

Dataset: The dataset was taken from a popular and reliable data center known as KAGGLE. The repository of tomato leaf images contains 10,000 records total across 10 classes, although all tests only employ a subset of 2 or 3 features. For insect detection, the computational time is less than (5) five minutes.

Using CNN Model for Tomato Leaf Classification

The repository of tomato leaf images in the KAGGLE dataset comprises 10,000 records total across 10 classes; however, all experiments will use a subset of 2 or 3 attributes. The computational time is less than (5) five minutes for insect detection.

The approach used to achieve this task using CNN model were;

1. **Collect and prepare a dataset:** The dataset for tomato leaves was taken from KAGGLE, and each image was classified as either healthy or diseased. The dataset was then divided into training and testing sets.
2. **Image Pre-Processing:** Following the collection of the data, picture pre-processing was used to carry out important activities such as resizing the image, introducing noise, filtering, HSV, Grey image, Binary image, complement image, edge detection segmentation, and dilatation of the image.
3. **Shape Feature Extraction:** In image processing, the feature extraction begins with a preliminary position of calculated values and constructed into another values (features) considered to be communicative and non-significant, facilitating the consequent studying and normalization steps and later dominating to enhanced human interpretations. Feature extraction is connected to proportional decrease.
4. **Building a model for Tomato Leaf classification:** the Convolutional Neural Network (CNN) Architecture was takes place in preprocessed image as input and output prediction for each image. The CNN consist of several convolutional layers followed by pulling and fully connected layers.
5. **Train the CNN model:** The training set of tomato leaf images was used to train the CNN model. The model then develops its pattern recognition skills while being trained on images of healthy or diseased tomato leaves.
6. **Evaluate the CNN Model:** Using the test set of photographs of tomato leaves and assessed the effective performance of the CNN to further evaluate the model performance, computation measures like accuracy, precision, recall, and f1 score are also carried out
7. **Fine-tune the CNN model:** By adjusting the CNN model's hyperparameters, including the number of layers, the number of filters in each layer, and the learning rate, the model was made more precise. Retrain the model, then assess its performance once.
8. **Test the CNN model:** After CNN model is trained, the model is tested and it gives higher accuracy and performance. The accuracy of CNN model gives 97.7% accuracy.

Model Design:

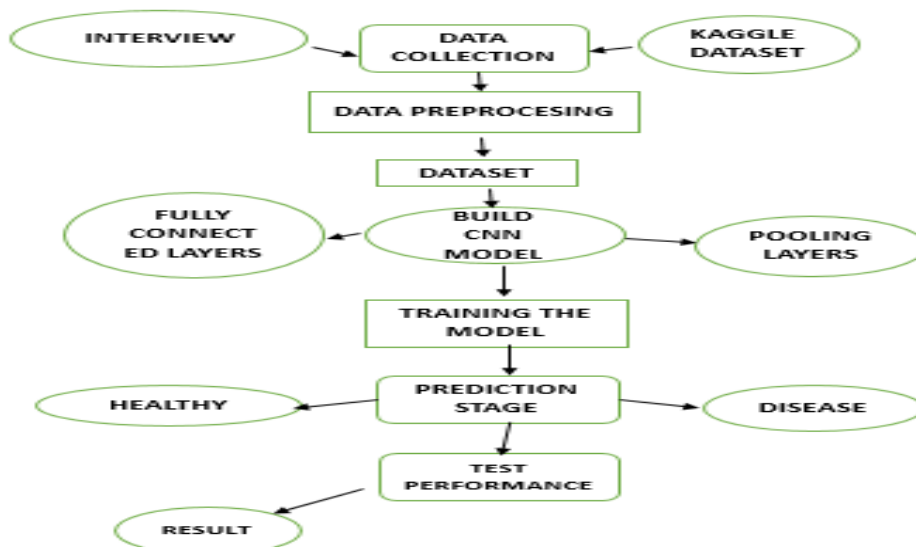


Figure 1: Model Design

Results and Discussion:

A class of neural networks known as convolutional neural networks, or CNNs or ConvNets, are particularly adept at processing input with a grid-like layout, like images. A

binary representation of visual data is a digital image. It is made up of a grid-like arrangement of pixels, each of which has a pixel value to indicate how bright and what color it should be.

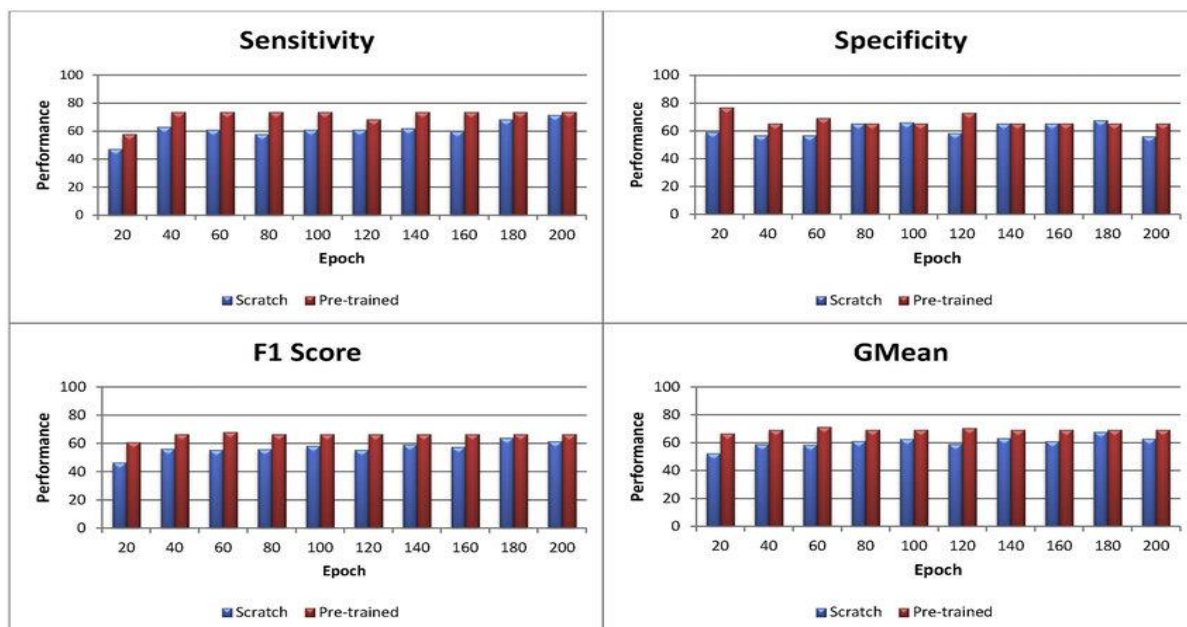


Fig 4.2: CNN graph

Validation of the Proposed Model

The k-fold cross-validation approach was used to examine the model's validation and estimate how well our algorithm performed on the tomato picture dataset. To divide the dataset,

we specify a 10 k-value. The k-fold approach was implemented using the Scikit-Learn machine learning toolkit, which returned a list of scores calculated for each of our ten folds.

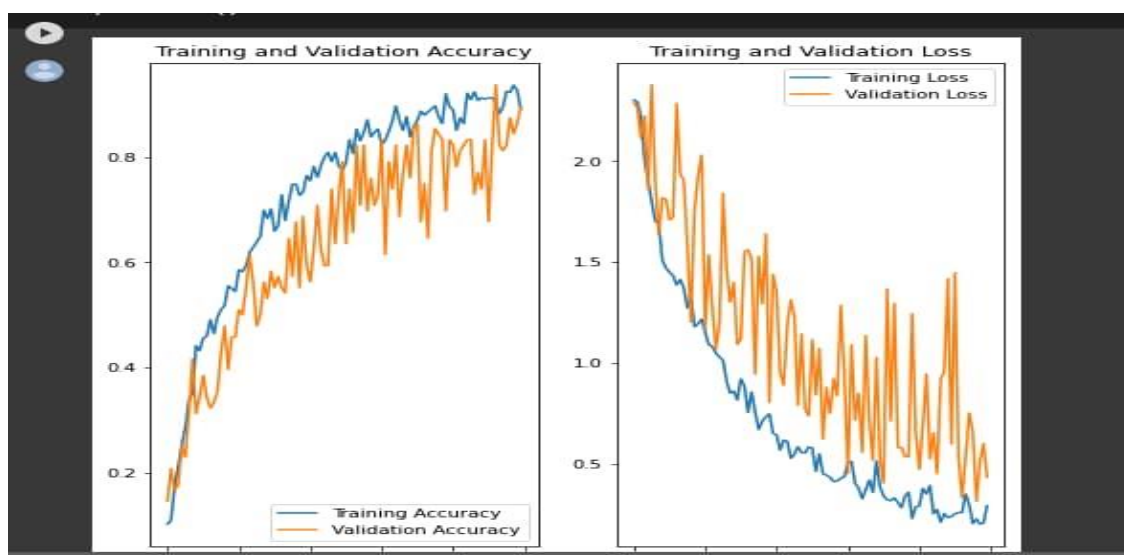


Fig 4.1. Result Graph of the Model

The performance of our model during the training and validation phases for the identification and classification of tomato leaf diseases is shown in Fig. 4.1. A training accuracy of 99.0% was reached by the results.

The training process took place in the JUPYTER environment for 423 (sec). The suggested model classified leaf diseases with a validation accuracy of 85.7%.



Fig 4.2. Confusion matrix of the proposed

The confusion matrix discovered during the examination of the suggested model was displayed in Figure 4.2. The true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values attained for each class assessed are displayed in the confusion matrix. The confusion matrix displayed the results, which show that the suggested model had an accuracy of 97.7% in predicting half of the classes that were assessed using the KAGGLE dataset. The model obtained better results than most of the studies suggested in the literature for the other classes. For instance, Kang and Kaya (2015) with the outcome of their model performances of 84%, Kayci (2018) with the model prediction of 76.3%, Li and Yin (2021) with the accuracy of 78% and Zhao and Wang (2019) with the outcome of 95%

Comparison of the Performance of Various Models

Datasets are extracted from the KAGGLE dataset with 10 Classes, each class containing

1000 photos, and several models are compared with our model. For the 10 classes of tomato leaf pest identified using nine size features and predicted by Random Forest and ANN, respectively, the classification accuracy is 86.9% and 70.8%. Due to feature dependence and equal weighting, the results of the Gaussian naive Bayes classifier and logistic regression had lower accuracy of 54.54% and 60.92%, respectively. As the author used SVM classifiers and moment-invariant characteristics to reach an accuracy of 79.95% for 10 classes. The Euclidean distance metric and the number of neighbors is employed in KNN to achieve an accuracy of 76.43%. CNN's model then achieved 97.7%. Time required for processing ten classes of image datasets using different classifier methods. As opposed to the proposed CNN, other algorithms including ANN, SVM, KNN, Naive Bayes, Random Forest, and logistic regression required less computation time.

Table 4.4: Results of Comparison of the Performance of Various Models

Model	Input	Precision	Recall	F1	speed (ms)	Accuracy (%)
CNN	10000	0.9999	0.9588	0.9950		97.7
KNN	10000	0.9069	0.9130	0.9133	10s	76.45
ANN	10000	0.8905	0.8891	0.8898	25s	70.8
Random Forest	10000	0.9332	0.9222	0.9325	23s	86.9
Model	Input	Precision	Recall	F1	speed (ms)	Accuracy (%)



SVM	10000	0.9331	0.9324	0.9327	15s	79.95
Naïve bayes	10000	0.9537	0.9548	0.9542	46s	54.54
Logistic regression	10000	0.9798	0.9816	0.9807	67s	60.92

Conclusion and Recommendation:

From these findings, CNNs are used to swiftly and effectively diagnose tomato infections, which are challenging for human experts due to the intricacy and irrationality of symptoms. The model's ability to accurately classify different disease types can help farmers control and stop the spread of the illness. In terms of agricultural technology, the creation of a pest control model employing CNNs for tomato disease detection is a promising advancement.

It is very important for tomato pest prediction and control that this project's image recognition research on tomato pests has produced good testing findings. It is necessary to investigate fresh algorithms that can spur creativity. The model still needs to be optimized in a variety of additional ways that need to be tried.

References

- Bhadane, G., Sharma, S. and Nerkar, V.B. 2019. Early pest identification in agricultural crops using image processing techniques. *Int. J. Electr. Electron. Comput. Eng.* Vol. 2: pp 77–82. Retrieved Jan., 17, 2012, from <https://www.researchgate.com>.
- Boissard, P., Martin, V. and Moisan, S. 2020. A cognitive vision approach to early pest detection in greenhouse crops. *Comput. Electron. Agric.* Vol. 62: pp 81–93. Retrieved march, 20, 2016, from <https://www.researchgate.com>.
- Bradshaw, C.J., Leroy, B., Bellard, C., Roiz, D., Albert, C., Fournier, A., Barbet-Massin, M., Salles, J.M., Simard, F. and Courchamp, F. 2019. Massive yet grossly underestimated global costs of invasive insects. *Nat. Commun.* Vol. 7: pp 29–86. Retrieved march, 17, 2015, from <https://www.researchgate.com>.
- Fraser, A. 2019. Land grab/data grab: Precision agriculture and its new horizons. *J. Peasant Stud.* Article 46. Retrieved may, 8, 2021, from <https://www.researchgate.com>.
- Gautam, M.P., Singh, H., Kumar, S., Kumar, V., Singh, G. and Singh, S.N. 2018. Diamondback moth, *Plutella xylostella* (Linnaeus) (Insecta: Lepidoptera: Plutellidae) a major insect of cabbage in India: A review. *J. Entomol. Zool. Stud.* Vol. 6: pp 1394–1399. Retrieved jan., 20, 2018, from <https://www.researchgate.com>.
- Kang, S.H., Cho, J.H. and Lee, S.H. 2014. Identification of butterfly based on their shapes when viewed from different angles using an artificial neural network. *J. Asia-Pac. Entomol.* Vol. 17: pp 143–149. Retrieved April, 20, 2011, from <https://www.researchgate.com>.
- Li, F. and Yin, X. 2021. Automatic identification of butterfly species based on HoMSC and GLCMoIB. *Vis. Comput.* Vol. 34: pp 1525–1533. Retrieved Nov., 18, 2018, from <https://www.researchgate.com>.
- Kaya, Y., Kayci, L. and Tekin, R. 2018. A computer vision system for the automatic identification of butterfly species via gabor-filter-based texture features and extreme learning machine: GF+ELM. *TEM J.* Vol. 2: pp 13–20. Retrieved Sep., 21, 2012, from <https://www.researchgate.com>.
- Zhao, Y.; Wang, Y., Wang, J., Hu, Z., Lin, F. and Xu, M. 2019. GMM and DRLSE Based Detection and Segmentation of Pests: A Case Study. In *Proceedings of the 2019 4th International Conference on Multimedia Systems and Signal Processing*, Guangzhou China, 10–12 May 2019; pp. 62–66.