AN ENSEMBLE-BASED DEEP LEARNING APPROACH FOR EARLY AND ACCURATE WHEAT DISEASE DETECTION

#11635

[1]Tagel Aboneh, [1]Prof. Abebe Rorissa, [2]Prof. Ramasamy Srinivasagan
[1]Addis Ababa Science and Technology University
[2]Professor & Director, School of Information Sciences Associate Dean for Faculty Affairs
College of Communication and Information the University of Tennessee, Knoxville

ABSTRACT

Crop diseases are the primarily cause for yield loss and a factor for food security issue around the globe. Crop diseases caused by pathogens pose a significant threat to global food security, the challenge become worst particularly in developing countries like Ethiopia. Rapid population growth and accurate disease identification is crucial for timely intervention and minimizing crop losses. However, traditional methods often rely on expert analysis, which can be time-consuming and resource intensive. The state of the art in agriculture employed AI enable crop diseases early detection technologies to support the agriculture domain area. Currently machine learning based solutions plays significant role to detect and classify crop diseases as early as possible. In this study, we proposed ensemble-based deep learning approaches for crop diseases classification purpose. Ensemble Deep Learning is a cutting-edge technique in machine learning that combines the strengths of multiple deep learning models to achieve superior performance compared to any individual model. The proposed model leverages the strengths of pre-trained models such as ResNet50, EfficientNetB4, DenseNet, ViT-Base and VGG19. To train the base models, more 23,000 crop images are acquired from various sources. The trained models are combined using ensemble learning method adjustable weighted average techniques to create a robust and generalizable model.We done model performance assessment based on optimization, scalability, and mitigation of model drifting issues to enhance the overall generalizability. From experimental results, the proposed ensemble model demonstrated a promising performance with 99.48% for training data and 99.23% for validation accuracy respectively. This research signifies a crucial step towards developing a practical and reliable tool for early crop disease detection in resource-constrained environments.

INTRODUCTION

The rapid population growth, a constant decline in arable land size per capital, and dynamic environmental change are the main constraints in the agriculture process. On the other hand, fighting against plant diseases is crucial activity in the agriculture sector to maintain crop productivity. In this regards, researchers are attempting new methods and technology to support the identification of crop diseases(Jasim & Al-Tuwaijari, 2020). The research findings in the domain area reveal that the application of technology would enhance agriculture production quality. In this regard, classical farming approaches, resources optimization, dynamic weather condition, severity of different pathogens are the main cause for y(Habiba & Islam, 2021)ield variation in the case of Ethiopia. In this study, we proposed ensemble based deep learning approach to classify crop diseases into the respective categories. Based on different literature review reports, about 20 to 40\% yield loss because of crop diseases. This is a significant effect which demand

appropriate action by stakeholder. On the other hand, climate change also a critical factor for crop production and a cause for food insecurity issue. Thus, the sector demands an AI enabled system at least to minimize its consequence on food security.

CONTRIBUTION

The proposed ensemble-based deep learning model holds significant potential to empower farmers and agricultural stakeholders with the ability to rapidly identify and address crop diseases, ultimately contributing to improved food security and agricultural sustainability. The proposed model are converted smart phone application to support smallholder farmers. An ensemble-based deep learning approaches are designed to develop generalizable models to improve the limitations of crop diseases early detection(Fuentes et al., 2017). The proposed model can be used for others crop disease detection and reference for wheat disease management work. Similarly, the proposed model can be used as decision support tool for different stakeholder in the domain area.

MATERIALS AND METHODS

In this research, an empirical research approach has been used to implement an ensemble deep learning framework to efficiently detect plant diseases(Reddy et al., 2021). In this work our main focus is to build generalizable(Ferreira et al., 2020) machine learning model for the purpose of crop diseases detection as early as possible. Figure 4.4 below illustrate the proposed system general system architecture.

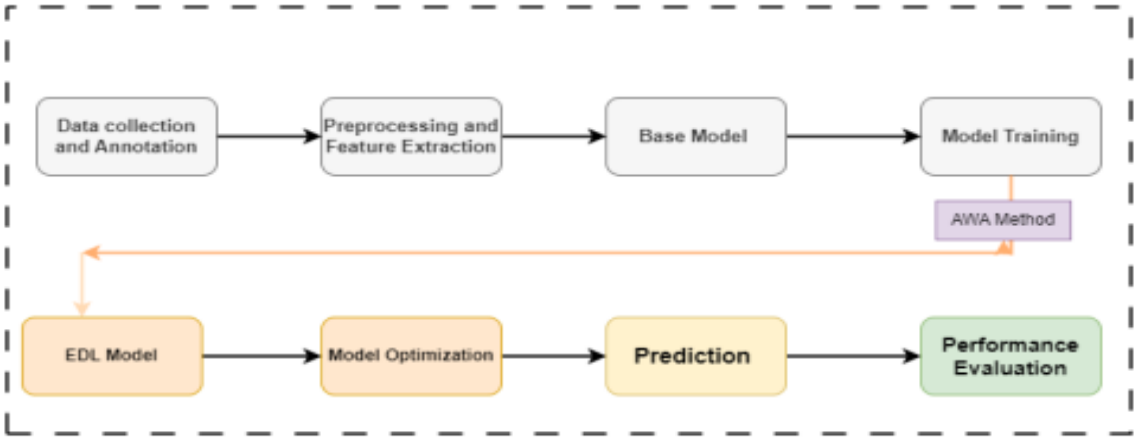

Figure 1. Ensemble based deep learning model data flow

DATASETS

To implement the experiment work, we have collected more than 25 thousand image datasets from kaggle repository and 1500 wheat image (RGB) dataset from our previous research study entitle with ’computer vision approach for wheat diseases classification using GPU infrastructure’. The datasets are structure into 4 major categories namely corn, wheat, potatoes and tomato. These categories are further classified into 20 classes (Figure 2).

SN	D. Main categories	Diseases Subcategories	Classes
1	Tomato Diseases	Yellow-leaf-curv-virus	1
		Mosaic-virus	2
		Target-spot	3
		Spider-mites	4
		Septoria-leaf-spot	5
		Leaf-mold	6
		Late-blight	7
		Healthy-tomato	8
2	Corn Diseases	North-leaf-blight	9
		Health-corn	10
		Common-rust	11
		Cercospora-leaf-spot	12
3	Potato Diseases	late-blight	13
		healthy	14
		early-blight	15
		Bacterial-spot	16
4	Wheat Diseases	Healthy weath	17
		leaf rust	18
		yellow rust	19
		stem rust	20

Figure 2.

Data processing for feature extraction (Jasim & Al-Tuwaijari, 2020) and selection are the important task before build the proposed model. In this regard, we have covered dimensional, removal of the least relevant features, image normalization, formatting, removal of poor-quality images, re-scaling or image resizing, and cropping of irrelevant parts of the image. Similarly, re-scaling pixel intensities values ranging from 0 to 255. Furthermore, we transformed the data by re scaling and setting the dimensions of the images at $224 \times 224 and channel = 3$ to standardize the data set. We have used well annotated crop image data to train our model, and the data sources are organized into training, testing and validation dataset. Then, we have selected five different deep learning model as a base learner namely ResNet50, EfficientNetB4, DenseNet, ViT-Base and VGG19 models (Rasti & Bleakley, 2020) (Figure 3). We have considered the following criteria, such as the size and quality of the data set, computational resources, disease types and crops, and accuracy requirements to select the base models. In line with ensemble learning method, the issue of computation infrastructure is very critical. To solve this challenge, we have used NVIDIA GeforceRTX3036 GPU facility to handle computational cost. Adjustable weighted average method has been used to readjust the weight of individual based models based on the validation loss accuracy.

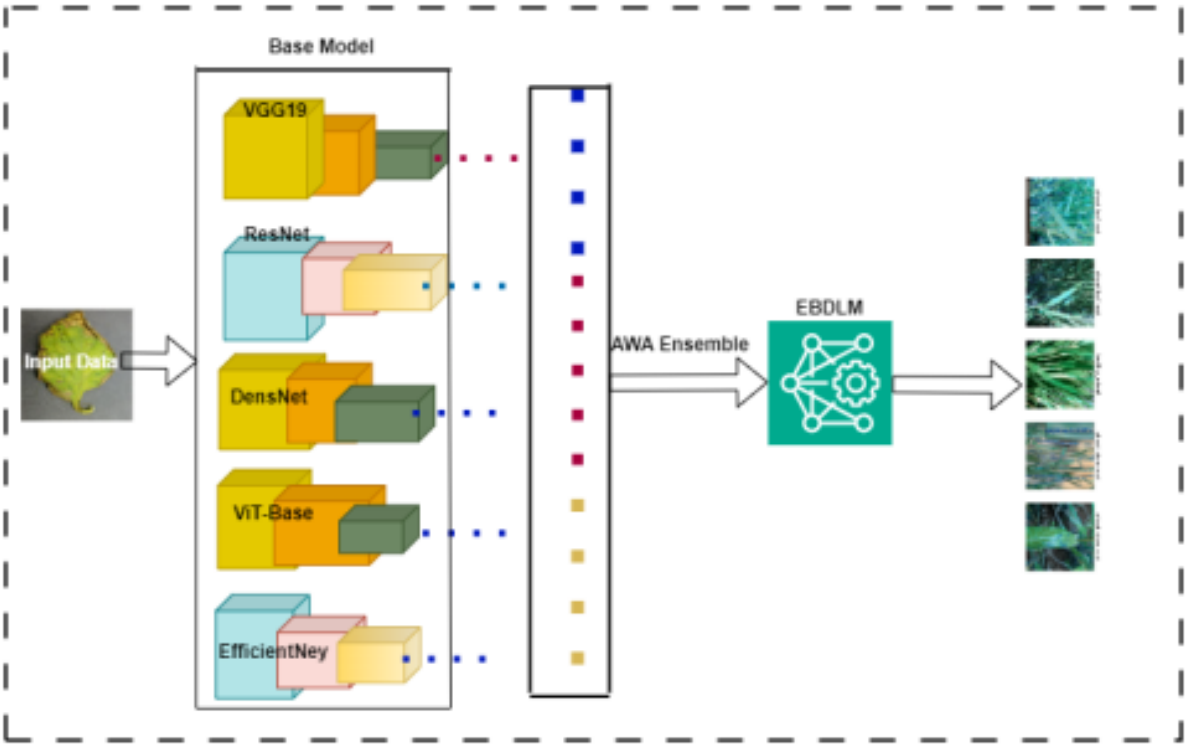

Figure 3.

In case of ensemble learning method, evaluating pairwise correlation between base models is very important before proceeding to build an ensemble model. In this regard, we computed the pairwise correlation between the output probabilities or predicted labels of each model. On Figure 4, we illustrated the model crop diseases prediction accuracy of 99.48% training and 99.23% validation accuracy respectively. From the experiment results, we can conclude that the proposed model has learned the underlying patterns in the on the unseen dataset perfectly.

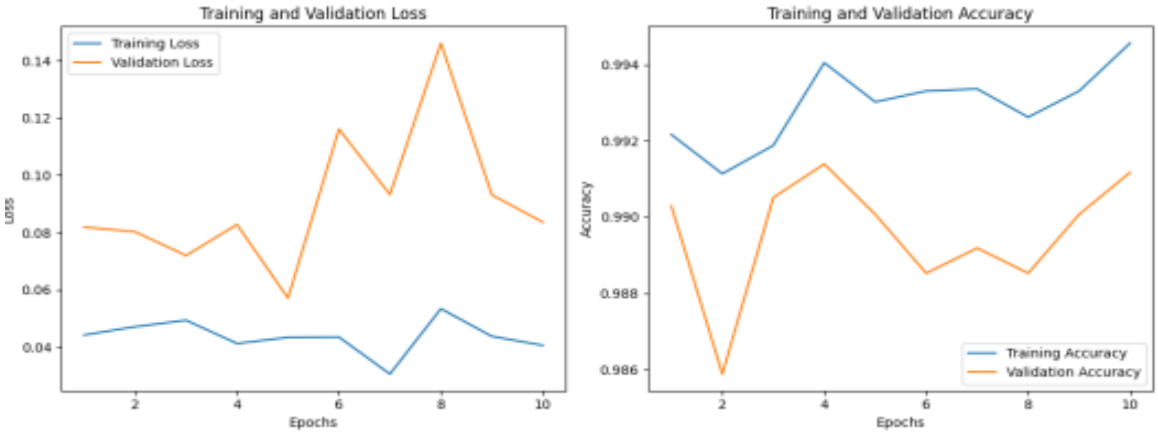

Figure 4. Ensemble model training and validation accuracy

We have used confusion matrix to understand the distribution of predictions across different classes which helps us to make informed decisions, such as adjusting the threshold for specific classes or focusing efforts on improving the model's performance for critical classes. Similarly,

we have computed model drifting based the validation dataset to assess how the model's performance has changed over time. Finally, we have made comparison analysis between the proposed model and similar studies in the domain area.

Ensemble learning is a powerful technique in machine learning where multiple models are combined to produce a single superior model.

CONCLUSION

Crop diseases also remain as a major threat to food security. However, rapid disease detection remains a home job for many developing countries such as Ethiopia. Different evidence reveal that proportion of yield loss due crop diseases is significant. Currently, machine learning based solutions plays significant role to detect and classify crop diseases as early as possible. In this study, we proposed ensemble-based deep learning approaches for crop diseases classification purpose. The underlying principle is that the collective decision of the committee tends to exhibit superior overall accuracy compared to any individual member. From the experiment results, the proposed model classifies the type of crop diseases with optimal accuracy. To further justify model's performance, we have used different scalability assessment on the proposed model.

REFERENCES

Ferreira, J. R., Armando Cardona Cardenas, D., Moreno, R. A., De Fatima De Sa Rebelo, M., Krieger, J. E., & Antonio Gutierrez, M. (2020). Multi-View Ensemble Convolutional Neural Network to Improve Classification of Pneumonia in Low Contrast Chest X-Ray Images. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2020-July, 1238–1241. https://doi.org/10.1109/EMBC44109.2020.9176517

Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors (Switzerland), 17(9). https://doi.org/10.3390/s17092022

Habiba, S. U., & Islam, M. K. (2021). Tomato Plant Diseases Classification Using Deep Learning Based Classifier from Leaves Images. 2021 International Conference on Information and Communication Technology for Sustainable Development, ICICT4SD 2021 - Proceedings, 82–86. https://doi.org/10.1109/ICICT4SD50815.2021.9396883

Jasim, M. A., & Al-Tuwaijari, J. M. (2020). Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques. Proceedings of the 2020 International Conference on Computer Science and Software Engineering, CSASE 2020, 259–265. https://doi.org/10.1109/CSASE48920.2020.9142097

Rasti, S., & Bleakley, C. J. (2020). Crop growth stage estimation prior to canopy closure using deep learning algorithms. 6. https://doi.org/10.1007/s00521-020-05064-6

Reddy, D. B. R. S., Madhavi, D. G. B., Lakshmi, C. H. S., Nagendra, D. K. V., & Sridevi, D. R. (2021). Detection of Disease in Maize Plant Using Deep Learning. Alinteri Journal of Agriculture Sciences, 36(2), 82–88. https://doi.org/10.47059/alinteri/v36i2/ajas21118

More references are available in the main content ….