JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

PLANT LEAF DISEASE CLASSIFICATION USING A LIGHTWEIGHT MODEL

[1]RAHAJANIAINA Andriamasinoro, [2]RATIARISON Adolphe Andriamanga
[1]Associate Professor, [2]Professor Emeritus
[1]Department of Mathematics Computer Science and Applications
[1]University of Toamasina, Toamasina, Madagascar

Abstract: Plant health take a place important in the life of humanity. Developing country like Madagascar has an economy based on the agriculture. Or, the majority of developing country face of lack agricultural infrastructures and specialists. The farmer uses archaic technical and this has an impact to the agricultural products and the environment. In fact, several plants are at the risk of extermination because of the diseases or human activity. So, it is crucial to setup an application for classifying plant leaf disease. The aim of this study is to perform a model for image classification that can help people indeed the farmer determine the kind of plant leaf disease based on lightweight device. Thus, we propose a lightweight deep learning-based model inspired from the architecture of SSDLiteX. The Balancing technical was used. After experiments, our model achieves 100% of accuracy and validation accuracy. The F1-score is 99.44%. The results demonstrate that the proposed model is reliable and answer the need of the agriculture and environment domains.

IndexTerms **- leaf disease, lightweight deep learning model, leaf classification.**

I. INTRODUCTION

Agriculture is a sector which has an impact on life and economy of the majority of the developing country. Like other African countries, 80% of Malagasy population are farmers. Although the efforts deployed by more organisations, food and nutritional insecurity affect a large part of the Malagasy population. Furthermore, the agricultural organization does not allow agricultural households – in particular small family farms, which are largely in the majority – to escape poverty. Indeed, agricultural activity generates low revenue and does not permit households to make the productive investments necessary for the development of agricultural operations.

There are numerous problems that the agriculture domain faces such as the drought, different kind of plant leaves diseases, lack of agricultural infrastructures and specialist. Thus, manual recognition disease was used more often. This process could be resulted in misdiagnosis and could lead to incorrect treatment. Furthermore, the attacks of plant disease cause a huge reduction in term of the quality and quantity of agricultural production. Therefore, it is essential to establish an efficient and durable solution.

Several deep learning-based models have been proposed in the literature for classifying leaves disease. These methods were divided into two categories: the first category was focused to classify one kind of plant. The second one was taking care of various species. The use of deep learning methods to assist specialists to detect plant disease could provide suitable treatment to the plant. Unfortunately, the most of them have been more complex and not suitable for any low-resource devices. Furthermore, several Malagasy farmers dispose a smartphone and capable to use it. Face to its situations, in this work we propose a lightweight deep learning-based model to classify plant disease that could be run on the lightweight devices.

The rest of this paper is organized as follows. Section 2 describes a brief literature review on the different plant diseases classification methods. Section 3 presents the material and methods. Results and discussion about the proposed method are discussed in section 4 and section 5 conclude this paper.

II. LITERATURE REVIEW

Paul S. et al. (2021) performed plant leaves classification using ResNet model as classifier. Seven datasets (Flavia, MalayaKew, Swedish, Middle European Woody Plants, Folio, Amazon Forestry, and Leaf Snap) were used for testing your model. conditional Generative Adversarial Networks (cGANs) was performed to achieve data augmentation. This technique was used to avoid overfitting problem. They used training/test set split of 75/25%. The results of the model show that the accuracy of all the datasets combined produced an average of 96.1% and best on the individual datasets.

Rinu and Manjula (2021) developed a model based on VGG16 for detection and classification of plant diseases. They used plant Village dataset downloaded from Kaggle website. The dataset contained 14 plant varieties formed by 38 plant classes. 54305 images were used for training. The data was preprocessed. After experiments, the model provided an accuracy around 94.8%.

Geetabai S. et al. (2024) in this paper, the authors used various pre-trained models such as AlexNet, VGG16, MobilNet V2, Inception V3, ResNet50 and EfficientNet. Two different optimization approaches (SGD and Adam) were applied to perform plant leaf disease classification. They used balanced dataset contained grayscale images of varying size. The model EfficientNet has achieved 97,5% accuracy, the best result compared to the other models.

Catur P. et al. (2020) applied a CNN architecture for classifying medicinal plant leaves. They used face94 dataset sourced from Dr. Libor Spacek. Numerous preprocess techniques were performed before the classification task. The experiment resulted an accuracy 86%, precision 22%, F1 score 23% and recall 23%.

In their work, Andrew N. et al. (2024) presented an ensemble learning based on two transfer learning architectures such as ResNet50 and MobileNetV3 small for classifying diseases in plants through their images. They used training/validation/test set split of 70/20/10%. Images dataset are resized to 224x224 pixels in order to be compatible with the models. Data augmentation was used to avoid overfitting and improving performance. The model achieved a maximum accuracy of 99,35%.

In paper Priyanka V. and Bhavani R. (2022), convolutional neural organization was utilized to classify and detect the type of leaf disease. The plant Village dataset were used. This one contains 5148 images grouped into 38 classes. Authors used Data processing and augmentation to improve the result. They used 80/20% train/test split distribution. The model gave an accuracy of 90,62% for 50 epochs.

Kiran S. et al. (2020) worked with a modified CNN architecture in order to detect and classify images. They released experiment on 17 plant species. Various image processing was used before classification task. This one used a combination of texture and color features as input. Their system obtained an accuracy around the 94,26%.

In Hersh M. et al. (2024), comparative studies based on pre-trained models such as ResNet-50 and MobileNetV2 was proposed. They used a local dataset formed by 2500 images of indoor plant to classify species of houseplant leaves. Images data included ten classes. Data augmentation was used to increase the size of dataset and reduce overfitting. They achieved an accuracy of 99% with data augmentation and 98,60% without the augmentation.

The authors in Konduru et al. (2023) produced also a comparative study based on Convolutional Neural Network (CNN) architectures, Alexnet and MobileNet to find which one provides the best results. The dataset utilized for this study consists of over 50,000 images. The MobileNet model has attained 99.69% training accuracy, 94.37% validation accuracy, 96% average precision, 96% recall, and an F1-score.

Mital V. et al. (2023) introduced a lightweight CNN architecture with fine-tuning to identify the best fit arguments for the model. They used images captured through the Samsung A52S mobile from the agricultural field of Jalgaon city combined with public plants PlantVillage dataset. they performed your experimentation by using 5 different distributions of train - test - validation 70-10-20%; 60-25-15%; 50-40-10%; 40-45-15% and 20-70-10%. Data augmentation was utilized to expand artificially and improve the visibility and diversity of dataset. Then, SoftMax activation function was used as output layer. The result of the model was dependent on the choice of train-test-validation set distribution. The experience shows that the 70-10-20% one yields maximum accuracy of 99,81%.

III. MATERIAL AND METHODS

III.1 Dataset description

Dataset contains nine species of plant leaves downloaded from Kaggle web site Sharma T. (2022). Each specie was already split in train, test and validation set. For this study, we were used five kinds formed by 23128 images and each variety has two classes. The Fig.1 shows samples leaves healthy images and the Fig.2 shows samples leaves diseases images.

Figure 1 Sample leaves healthy

Figure 2 Sample leaves diseases

III.2 Data preprocessing

Table 1 Data repartition before preprocessing

Plant	Disease name	Train	Test
Apple	Healthy	2259	251
	Scab Bell	2268	252
Pepper	Healthy	2236	249

	Bell_Bacterial	2151	240
Grape	Healthy	1903	212
	Leaf Blight	1936	216
Peach	Healthy	2067	216
	Bacterial	1944	230
Strawberry	Healthy	2052	228
	Leaf Scorch	1996	222

For achieve our goal, the dataset was rearranged. Then, among 23128 images, 10% were used for testing the model and 90% for training. For this later, 10% were utilized for validation. The table 1 gives the repartition of the dataset into training and testing set.

As we show in the table 1, the data in the training set is imbalanced. To overcome this problem, a balancing technique was performed rather than data augmentation for better results. During the balancing action, five transformations were achieved like a clockwise rotation with an angle choice randomly in range of 2 to 350 degree, a flipping right-left, a flipping top-bottom, an adjustment of contrast and adjustment of brightness. The choice of this one was done randomly. The Fig.3 and Fig.4 show some samples leave images and their copies in dataset after preprocessing.

Figure 3 Sample leaves

Figure 4 Sample copies of the leaves after preprocessing

III.3 Proposed Model

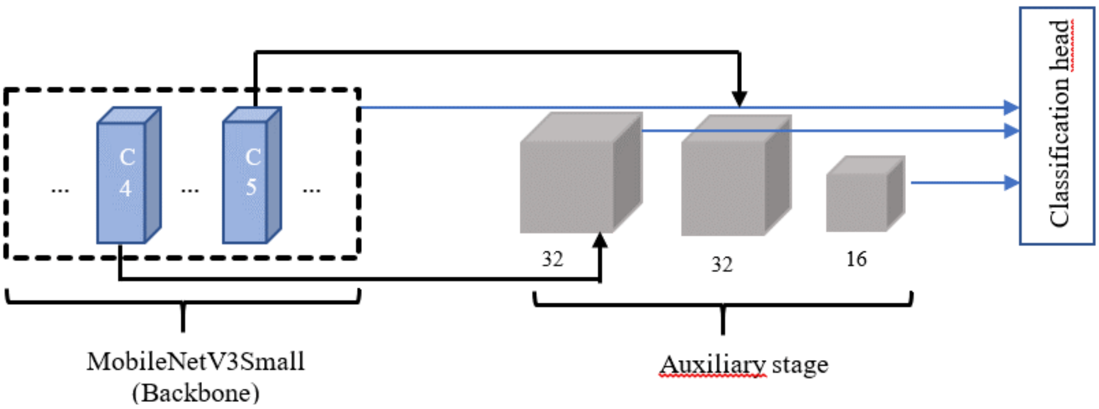

Figure 3 Model architecture

The goal of this work is to perform a lightweight model for classifying plant leaves diseases. In fact, transfer learning technique was used. Thus, the lighter version of MobileNetV3 called MobileNetV3Small was utilized as backbone. The SSDLiteX auxiliary stage composed of a 3 × 3 depthwise convolutional layer and a 1 × 1 convolutional layer was used with a reduction number of filters. Rather than used 256 filters, the proposed system was used 32, 32 and 16 filters at the first, the third and the last stage respectively. The last feature extractor layer that has an output stride of 16 called C4 is attached to the first layer of the auxiliary stage and the last one that has an output stride of 32 called C5 is attached to the second layer of auxiliary stage. These tricks contribute to the quality of the result. Our model has only 0.9 million (3.44 MB) of parameters and has a size of 6.17 MB. It outperforms categorically several lightweight models such as MobileNetV3Small that has 2.4 million of parameters and has a size of 9.83MB. Fig. 3 illustrates the architecture of our model.

IV. RESULTS AND DISCUSSION

In this part, we present the proposed model's experiment result. A batch size of 1 and an epoch of 50 gave great result. The experiments were carried out in Intel(R) Core (TM)i7-1255U CPU, 10 cores, 12 threads @2.30GHz, and 24 Gb RAM. The implementation was performed using TensorFlow and Keras. A shuffling technique with a seed of 1000 were performed. To evaluate the model, we refer to two metrics: the accuracy and F1-score. The accuracy gives the proportion of the correct prediction made by the model. Thus, it gives the ration for both true positives and true negatives to the true total number cases. F1-score on the over hand can be interpreted as a harmonic mean of the precision and recall.

The model achieves an accuracy and validation accuracy of 100%. The global F1-score is 99.44%. Its shows that the proposed model is more effective for the plant Village dataset. Fig.4 below illustrates the accuracy and loss of the model.

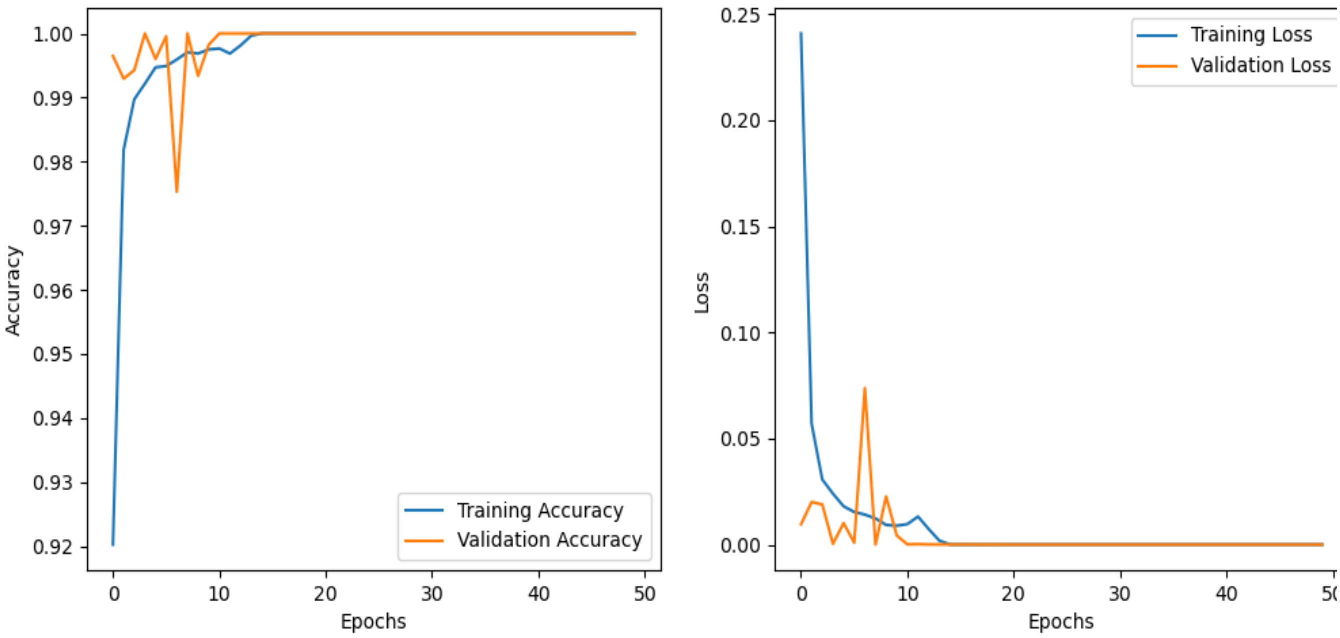

Figure 4: Accuracy and loss of our model

The confusion matrix below illustrates the global result of the classification task. It shows that only a slight misclassification on the ten classes.

$$\begin{bmatrix} 249 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 248 & 0 & 0 & 0 & 1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 244 & 2 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 1 & 1 & 238 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 212 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 216 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 230 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 216 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 228 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 222 \end{bmatrix}$$

V. CONCLUSION

In this paper we introduced a deep-learning based model demonstrating new state of the art in plant leaves diseases classification. A very lightweight model is presented. We have described our efforts to implement various techniques in order to have the optimal results. These ones demonstrate that the presented model get the highest accuracy and outperforms the state-of-the art recorded in this domain. It shows that the proposed model is more effective for this dataset. The size and other characteristics show that it could be deployed on a lightweight device such as a Smartphone, Thus, it could be also aided the peasants in your domain to know the kind of plant diseases

REFERENCES

[1] Andrew N. et al. (2024). Classification of plant leaf disease using ensemble deep transfer learning. International Journal of Intelligent Computing and Information Sciences, 24 (1): 69-88. http://dx.doi.org/10.21608/ijicis.2024.278335.1331

[2] Catur P. et al. (2020, December). Implementation of CNN for plant leaf classification. International Journal of Informatics and Computation, 2 (2), 1-9. http://dx.doi.org/10.35842/ijicom

[3] Geetabai S. et al. (2024). Classification of Various Plant Leaf Disease Using Pretrained Convolutional Neural Network On Imagenet. The Open Agriculture Journal, 18, 1-15. http://dx.doi.org/10.2174/0118743315305194240408034912

[4] Hersh M. et al. (2024). Houseplant leaf classification system based on deep learning algorithms. Journal of Electrical Systems and Information Technology, (1): 1-15. https://doi.org/10.1186/s43067-024-00141-5

[5] https://data.mendeley.com/datasets/tyw btsjrjv/1

[6] Kiran S. et al. (2020, November). Deep Learning for Plant Species Classification. Journal of Emerging Technologies and Innovative Research, 7 (11): 99-105. https://www.jetir.org/papers/JETIR2011016.pdf

[7] Konduru et al. (2023, october). Plant Disease Detection and Classification Using Deep Learning Models. Journal of Computing Sciences in Colleges, 39 (3): 66-75. https://dl.acm.org/doi/abs/10.5555/3636988.3637006

[8] Mital V. et al. (2023). High performance deep learning architecture for early detection and classification of plant leaf disease. Journal of Agriculture and Food Research, 14, 1-9. https://doi.org/10.1016/j.jafr.2023.100675

[9] Paul S. et al. (2021, December). A Deep Learning-Based Recognition Technique for Plant Leaf Classification. IEEE Access, 9, 162590 – 162613. http://dx.doi.org/10.1109/ACCESS.2021.3131726

[10] Priyanka V. and Bhavani R. (2022, July). Plant leaf disease classification using CNN. International Research Journal of Engineering and Technology, 9 (7): 772-777. https://www.irjet.net/archives/V9/i7/IRJET-V9I7148.pdf

[11] Rinu and Manjula (2021, September). Plant disease detection and classification using CNN. International Journal of Recent Technology and Engineering, 10 (03), 152-156. https://doi.org/10.35940/ijrte.C6458.0910321

[12] Sharma T. (2022). A broad dataset of diseased and healthy plant leaves for 9 different species. Kaggle. Accessed on 14/11/2024 at 14 :18 p.m.