Comparison Study of VGG16 And Inceptionv3 Transfer Learning Models For Classification of Plant Leaf Disease

Arya Rajagopal[1], Lekshmi R Nair[2], Dr.Priya S[3]
[1, 2, 3] Dept of Computer Engineering
[1, 2, 3] College of Engineering Cherthala

Abstract- *Early diagnosis of plant diseases carried out by professionals in laboratory trials is usually not suitable for quick and affordable performance. Using deep learning, leaf images are used as information input. Training deep learning models need large, hard-to-come datasets to perform the task to reach optimal results. In this study, the Plant Village dataset was used totaling 2700 training data and 300 validation data. Data were trained using 100 epoch iterations operating the transfer learning process with the VGG16 and InceptionV3 models. Founded on testing using 150 IVEGRI data, the VGG16 model can generalize data reasonably than InceptionV3. VGG16 by tuning block-3 using parameters 4096x2 and Dropout 0.4 shows the best performance with an average score of 1 precision, an average recall of 1, an average f1- score of 1, and 100 % accuracy. Then, with the same parameters, the Inception-v3 model with tuning in the mixed6 inception module shows the best performance with an average score of 0.93 precision, an average recall of 0.92, an average f1-score of 0.92, and an average accuracy of 92%.*

Keywords- Deep Learning, Transfer Learning, VGG16, InceptionV3, Potato Leaf Diseases Classification.

I. INTRODUCTION

Agriculture is just not helpful for human feeding or earning it is much more like energy and global warming. Leaf disease has been affecting many aspects in the field of agriculture mainly they are production, quality and quantity. India is a country which is dependent on agriculture. Leaf disease detection can be helpful for the farmers. Research works in smart computing surrounding to identify the disease using the pictures of leaves. Several problems are to be identified which are given as follows. Detecting the diseased leaf, to measure area affected by the disease, identifying the boundary of affected area by disease, finding out the colour of the affected area and what exactly causes the disease i.e., by insects, rust, nematodes etc., Diseases on the leaves are mainly viral, bacterial, fungal. Plant disease is one of the important factors which causes significant reduction in the quality and quantity of plant production. Detection and classification of plant diseases are important task to increase plant productivity and economic growth. Detection and classification are one of the interesting topics and much more discussed in engineering and IT fields. Identification of the plant diseases is the key to preventing the losses in the yield and quantity of the agricultural product. The studies of the plant diseases mean the studies of visually observable patterns seen on the plant.

Health monitoring and disease detection on plant is very critical for sustainable agriculture.

The classification of plant diseases is a very important job for producers or farmers. By using deep learning for image- based leaf disease classification, plant diseases can be detected accurately and at an affordable cost, because to classify using experts may be very expensive. Classification using vision technology has been widely used to increase accuracy and reduce work costs. The Convolutional Neural Network method is proposed by many researchers to classify data in the form of images with large data input. Examples of CNN models that are widely used by researchers today are AlexNet [4], VGGNet [5], GoogleNet [6], Inceptionv3 [7], ResNet [8], and DenseNet[9] followed. All of these CNN algorithms are the result of the ImageNet Large Scale Visual Recognition Challange (ILSVRC) competition which annually creates and searches for new algorithms for object detection and classification in recognizing ImageNet objects containing 1000 classes [10].

II. RELATED WORK

Research related to the classification of plant diseases had been carried out by several previous studies. Rakhmawati et al [1] used 300 training images and 90 test images consisting of 3 classes of potato leaves using GLCM and Color Moment for feature and color extraction and then SVM classifier. The results of this analysis get an accuracy of up to 80%. Also, in the study, it was concluded that the pattern of diseased leaves greatly influenced identification. Thus, the choice of non-disease or healthy leaves should be more selective with no spots at all. Then other study related to plant diseases was discussed by Arya [14].

This study uses 2 architectures, namely CNN with 3 layers of feature learning and 2 layer classification and the AlexNet model with transfer learning. The dataset used in this study is from the PlantVillage website and is taken in real-time from the GBPUAT plantation. The total data amounted to 4004 consisting of 4 classes, 2 classes of mango plants, and 2 classes of potato plants. 80:20, 3523 random splits were performed for training data and validation, while 481 was for testing. The results of the study obtained an accuracy of 90.85% for CNN and 98.33% for AlexNet. Research on grape leaves using trans- fer learning was conducted by Gangwar et al [15] using the

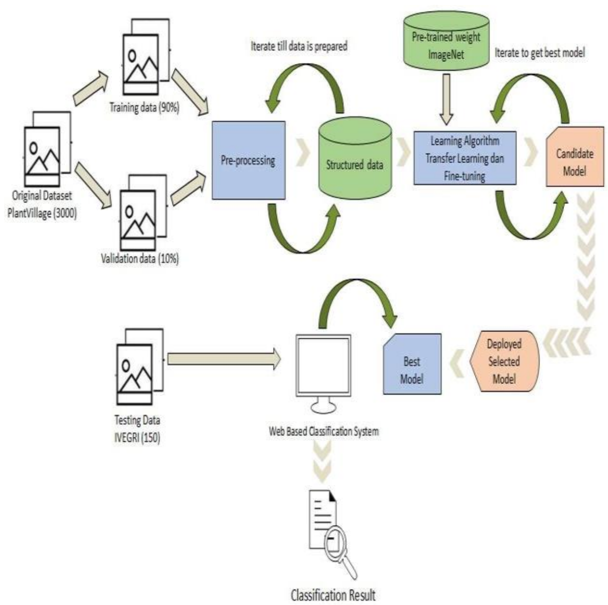


Fig. 1.Block diagram of proposed potato leaf disease classification system research.

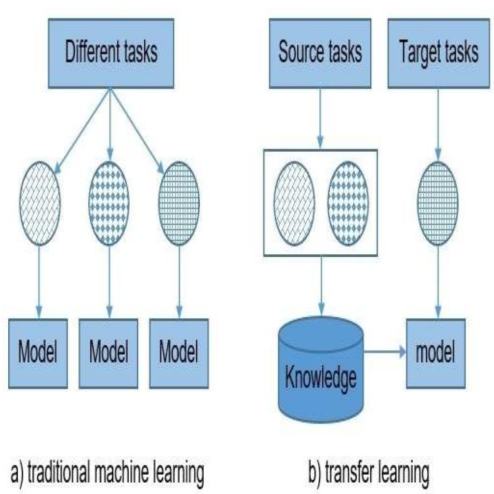


Fig. 2. Differences in traditional learning and transfer

InceptionV3 model with SVM classifier, logistic regression, and neural network. This study uses a PlantVillage dataset totaling 4062 consisting of 4 classes where the training data is 3209 and the test data is 853. The best classification results for InceptionV3 with logistic regression classifier achieve an accuracy of 99.4%. Agarawal, M. et al [16] conducted a study on potato leaves with a PlantVillage dataset of 3000 datasets using the caviar approach [17] to train the model. that. The training accuracy was 99.47% and testing accuracy was 98% using the CNN architecture using four layers with 32, 16, and 8 filters in each layer against 150 test data

III. THE PROPOSED WORK

In this study, there are several stages to identify potato leaf disease. Figure 1 shows the general stages in this study to classify potato leaf disease. At the training stage, the transfer learning method uses the pre-trained VGG16 and Inceptionv3 models found in the hard repository. Both models have previously been trained using the ImageNet large dataset, where previously trained models on the ImageNet large dataset can help solve computational problems and relatively few datasets at the training stage [18].

A. Data Preprocessing

In the training and testing stages, the image input size must be the same as the input size applied to the model. In this study, the image is resized to 224x224. Then scale the pixel value (between 0 and 255) to the interval [0, 1], this operation is because neural networks prefer to handle small input values [18].

B. Data Augmentation

To generalize the data, data augmentation is applied during the training process. Data augmentation is an operation toprevent overfitting by generating more training data. The goal is that the model will not see the same image twice and be able to adapt to real- world problems. This operation is performed using Keras with the ImageDataGenerator function by apply- ing a transformation operation consisting of rotation, width and height shifting, zoom, and horizontal flip. Augmentation was performed on training data and not performed on data validation [18].

IV. TRANSFER LEARNING

Transfer learning is a learning technique method using pre- trained neural networks by taking part in a model that has been trained to be reused in recognizing new models [12]. As shown in Figure 2, there is a source task and a target task, the knowledge gained from the source task model training process is transferred to handle the target task, in this case, the trained weight model from a large dataset of ImageNet images is transferred to recognize potato leaf disease.

A. Transfer learning vgg16

VGG16 was developed by the Visual Geometry Group at the University of Oxford [5]. This model won the ILSVRC contest as the second winner in image classification and the winner of image localization in 2014 in recognizing 1000 ImageNet object classes. VGG16 is designed to reduce the large kernel size on AlexNet 1x1 and 5x5 which is replaced by several 3x3 kernels with 1 stride which is useful for extracting complex features with low computation. VGG16 has 5 convolutional blocks consisting of 2 to 3 convolutional layers with ReLu activation. At the end of each block, MaxPooling 2x2 with 2 strides is used. Further descriptions of the VGG16 architecture have been described in the literature [5]. In this study, the pre- trained VGG16 convolutional base was maintained, then after the 5th block convolution layer, Global Average Pooling was

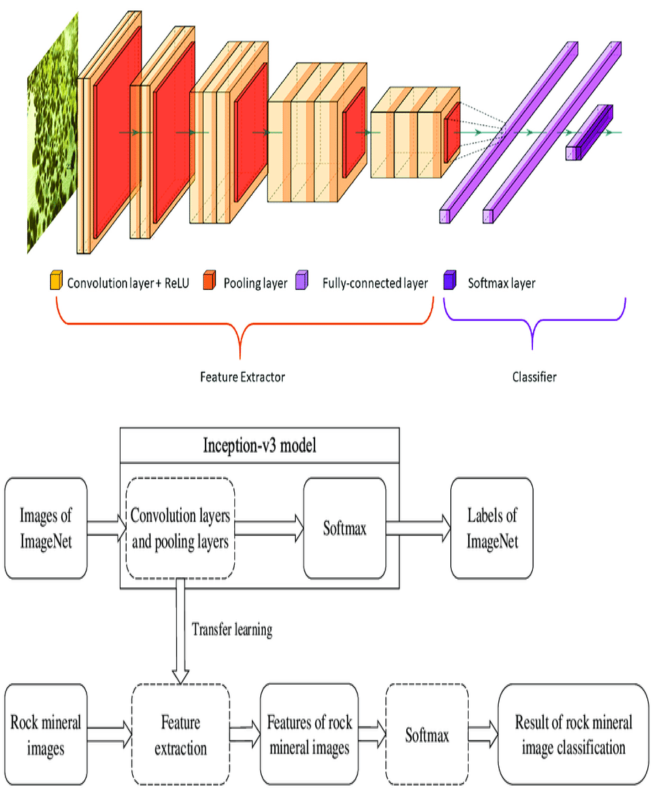


Fig. 3. Transfer learning architecture of VGG16.

used. The top layer in this study was modified using three ful connected layers with dropouts. where the last fully connect layer uses the softmax classifier. Figure 4 shows the VGG architectural model proposed in this study.

B. Transfer Learning InceptionV3

The Inception-v3 model is the development of t GoogleNet or Inception-v1 model developed in research [which won the ImageNet Large Scale Visual Recogniti Challenge (ILSVRC) contest in the image classification cat gory in 2014. Then it was refined by adding batch normaliza-tion (BN) called Inception-v2 model. With the development of additional factorization ideas for convolutional operations, this architecture is named after the Inception-v3 model which won the first runner-up in the ISLVRC contest in 2015 in rec-ognizing 1000 ImageNet objects classes. Inception-v3 consists of 5 basic convolutional layers (stem) consisting of conv2d0 to conv2d4 where each convolutional operation is followed by ReLu activation and Batch Normalization. Then followed by 11 inception modules consisting of mixed0 modules to mixed10 modules. The 11 Inceptionv3 module blocks are designed with 1x1, 3x3, 1x3, 3x1, 5x5, 1x7, and 7x7 convolutional kernels. In this study, the Inceptionv3 convolutional base was maintained, while the upper layer in this study was modified consisting of GAP and three dense layers where the last dense layer used the softmax classifier. A complete description of Inceptionv3 is described in the literature [7]. In Figure 5 is the InceptionV3 architectural model proposed in this study.

D. Drop Out

Dropout is a regularization technique that can help CNN withstand overfitting and also speed up the training process [20]. The way it works is to temporarily remove hidden layers as well as visible.layers that are randomly located in the network and redirect to more trained neurons to reduce interdependence learning in each neuron [19]. Figure 6 below is an example of an Artificial Neural Network before and after the dropout process.

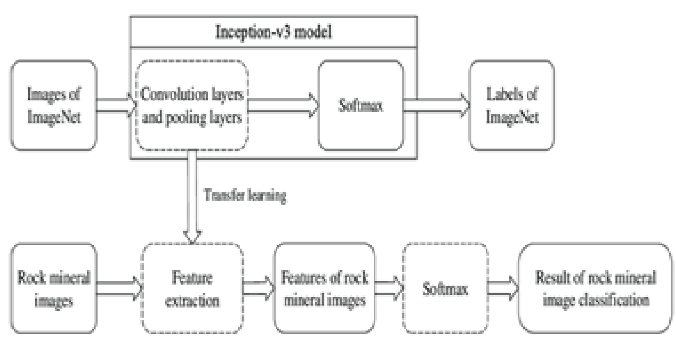


Fig. 4. Transfer Learning Architecture of Inception

V3.

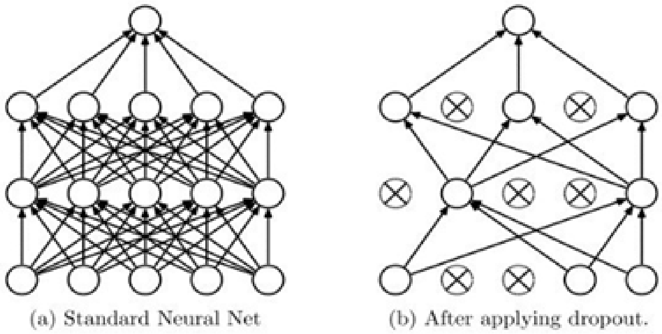


Fig.5.Drop Out.

D. Performance Measures.

To evaluate the performance of the classification model and object recognition in machine learning, deep learning, and information retrieval, evaluation based on precision, recall, f1- score, and accuracy is used to test the model performance of the test dataset test process [21].

Accuracy is the ratio of correct predictions to the overall data. To calculate accuracy, we use Equation 1. precision is the part of the object that is predicted to be correct, calculated using Equation 2. Recall / Sensitivity is used to find out how accurate is the model's performance to classify correctly, or in other words, how many times the model misclassifies false negative using Equation 3. The F1-score is the comparison of the mean precision and recall, in other words summarizing the classification performance with a single metric representing precision and recall. The F1-score is calculated using Equation

.Where TP (True Positive) is positive data that is proven true, FP (False Positive) is negative data that is proven true, TN (True Negative) is positive data that is not proven true, and FN (False Negative) is negative data that is not proven right.

V. EXPERIMENTS AND RESULT

In this section, the results obtained are displayed based on the method used. The first machine used for training has specifications of a Xeon processor, 25 GB RAM,

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (1)$$

$$precision = \frac{TP}{TP + FP} \quad (2)$$

$$recall = \frac{TP}{TP + FN} \quad (3)$$

Fig. 6. Equation of f1 score,Precision,Recall

Nama Class	Sample of image	Num of training	Num. of validation
Early Blight		900	100
Late Blight		900	100
Healthy		900	100
Total		2700	300

Fig. 7. Traing and Validation of Data

Nvidia Tesla P400, and storage of 147 GB from the co-lab.research.google.com server [22]. The second machine used to test the model has the 7th generation Intel i5 specification, 12GB RAM, Nvidia GeForce 940MX, and 1 TB HDD memory specification. Experiment using Tensorflow and Keras to implement the VGG16 and Inceptionv3 pre-trained models.

A. Dataset

This study uses a PlantVillage training dataset obtained from research [10]. The image used in this research is a potato leaf that has an image resolution of 256x256. The training data amounted to 3000 with 3 classes, each class consisting of 1000 pictures. Training data is divided into 90:10 for training data and validation data. While the test data is data obtained from the Research Group on Germplasm Breeding and Germination, the Indonesian Vegetable Research Institute, totaling 150 data taken using a Nikon D3200 DSLR camera and smartphone from various angles and using white cardboard as a background. The amount of data in this study is shown in Table 1 and Table 2.

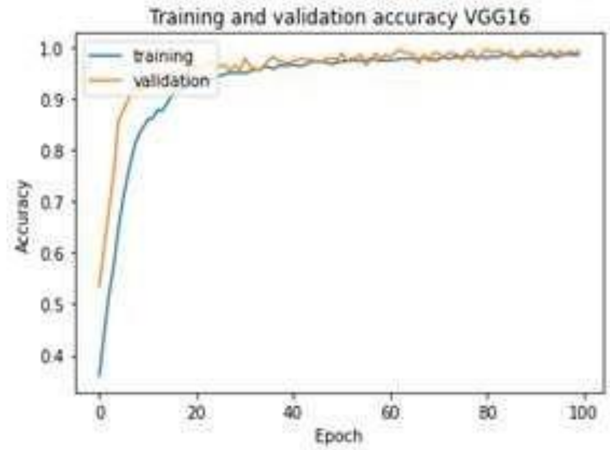
Class	Sample	Num. of image
Early Blight		50
Late Blight		50
Healthy		50
Total		150

Fig.8.Testing Data

B. Experimental parameters

The top layer of the VGG16 and InceptionV3 models in this study was carried out by tuning trial and error on the parameters to obtain the best performance for training the data. The tuning performed on the VGG16 model is replacing Flatten with Global Average Pooling to produce one feature map for each output. Whereas in the InceptionV3 model, which originally had one FC layer, in this study tuning was carried out by adding one FC layer to the top layer so that the performance of the two models based on MLP on the same top layer could be compared.

After using GAP, two FC layers followed with dropouts. The use of GAP aims to reduce prone overfitting when using a fully connected layer and summarize spatial information to speed up the training process [23]. The fully connected layer is useful for getting a feature map then using it to classify images into the appropriate labels. The FC layer has a function to represent image features into appropriate labels, therefore this study also looks for the performance of the number of nodes/neurons in the upper layer. On the other hand, the FC layer tends to experience overfitting especially when large neural networks are trained on relatively few datasets [24]. Because of that, after the dense layer, a dropout parameter is used. This research applies dropout parameter tuning in both models starting from not applying it at all then applying dropout 0.2, 0.3, 0.4, and 0.5.Fine-tuning of the model is done by experimenting with different numbers of neurons and dropouts. Each network is tuned backward. In VGG16, the first tuned block is the 5th block, then the backward tuning until the entire network is the 1st block. Whereas the InceptionV3 model has 11 inception model blocks, backward tuning is performed starting from the mixed10 inception module then backward tuning to the entire basic convolutional network. This adjustment is done to find at what layer the model gets the best performance or converges faster (training accuracy And validation are not corrugated).



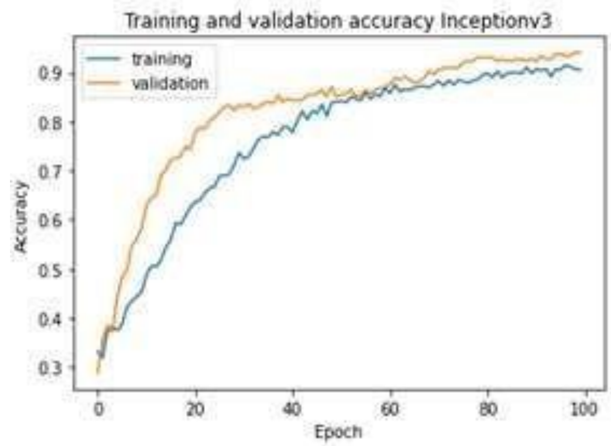


Fig. 9. Training and validation Accuracy of VGG16.

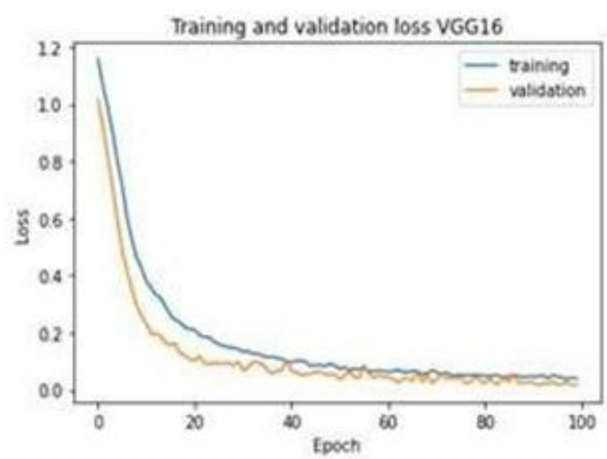


Fig. 10. Training and validation loss of VGG16.

VI. CONCLUSION

In this study, it was successful in implementing Deep Learning using the Transfer Learning method with the pre-trained VGG16 and InceptionV3 models for the classification of plant leaf disease. The training process is carried out using the same top layer MLP. In the VGG16 model, fine- tuning the last two blocks can provide good performance because of the association that sublayer features have common features such as edge features that can be used for all types of datasets. Whereas the top layer feature has more specific features, it means the bottom layer can be frozen or not retrained. In deep models such as InceptionV3, retraining the entire network results in reduced performance, because the deeper the architecture gets, the more parameters to be trained increase and the potential for overfitting increases. It is concluded in this study that the VGG16 model can generalize data better than InceptionV3. Obtained the highest accuracy in both models using the parameters Size of FC 4096x2 and Dropout 0.4 on the top layer where VGG16 got 100% accuracy and InceptionV3 92%. The use of Size of FC 4096x2 makes the representation of image features more and helps the model to recognize more features but requires a lot of memory. As the dropout rate increases, the accuracy will increase because it reduces the learning interdependence in each neuron. The use of a dropout rate of 0.5 gives a decrease in performance because too much weight is removed, in this study a trial and error experiment was carried out.

In the future, research can be developed by adding data on types of vegetable plant diseases. And in its implementation, this research is expected to help the agricultural industry in maintaining vegetable crops.

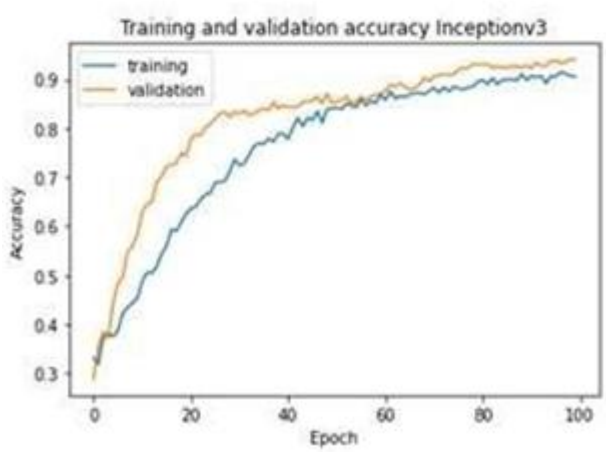


Fig.11. Training and validation Accuracy of InceptionV3.

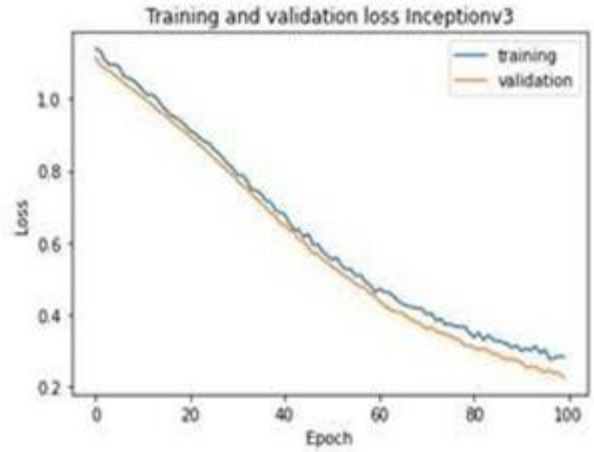


Fig.12. Training and validation loss of InceptionV3.

and InceptionV3 92%. The use of Size of FC 4096x2 makes the representation of image features more and helps the model to recognize more features but requires a lot of memory. As the dropout rate increases, the accuracy will increase because it reduces the learning interdependence in each neuron. The use of a dropout rate of 0.5 gives a decrease in performance because too much weight is removed, in this study a trial and error experiment was carried out.

In the future, research can be developed by adding data on types of vegetable plant diseases. And in its implementation, this research is expected to help the agricultural industry in maintaining vegetable crops.

REFERENCES

[1] P. U. Rakhmawati, Y. M. Pranoto, and E. Setyati, “Klasifikasi Penyakit Daun Kentang Berdasarkan Fitur Tekstur Dan Fitur Warna Menggu- nakan Support Vector Machine,” Semin. Nas. Teknol. dan Rekayasa 2018, pp. 1–8, 2018..

[2] Gunadi, A. K. Karjadi, and S. Sirajuddin, “Pertumbuhan dan Hasil Beberapa Klon Kentang Unggul Asal International Potato Center di Dataran Tinggi Malino, Sulawesi Selatan,” J. Hortik., vol. 24, no. 2, p. 102, 2014..

[3] A. Dimyati, “Research priorities for potato in Indonesia,” Proc. CIP- Indonesia Res. Rev. Work. held Bogor, Indones., pp. 15–19, 2003.

[4] A. Krizhevsky and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” pp. 1–9, 2012.

[5] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–14, 2015.

[6] C. Szegedy et al., “Going deeper with convolutions,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp. 1–9, 2015.

[7] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2015- Decem, pp. 2818–2826, 2015.

[8] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 770–778, 2015.

[9] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-Janua, pp. 2261–2269, 2017.

[10] E. C. Too, L. Yujian, S. Njuki, and LYingchun, “A comparative study of fine- tuning deep learning models for plant disease identification,” Comput. Electron. Agric., vol. 161, no. October 2017, pp. 272– 279, 2019

[11] M. Shu, “Deep learning for image classification on very small datasets using transfer learning,” 2019

[12] S. J. Pan and Q. Y. Fellow, “A Survey on Transfer Learning,” pp. 1–15, 2010

[13] G. Geetharamani and A. P. J., “Identification of plant leaf diseases using a nine-layer deep convolutional neural network,” Comput. Electr. Eng., vol. 76, pp. 323–338, 2019

[14] S. Arya and R. Singh, “A Comparative Study of CNN and AlexNet for Detection of Disease in Potato and Mango leaf,” IEEE Int. Conf. Issues Challenges Intell. Comput. Tech. ICICT 2019, no. Dl, 2019.

[15] N. Gangwar, D. Tiwari, A. Sharma, M. Ashish, A. Mittal, and G. K. Vishwavidyalaya, “Grape Leaf Diseases Classification using Transfer Learning,” pp. 3171–3177, 2020

[16] Agarwal M., Sinha A., Gupta S.K., Mishra D., Mishra R. (2020) Potato Crop Disease Classification Using Convolutional Neural Network. In: Somani A., Shekhawat R., Mundra A., Srivastava S., Verma V. (eds) Smart Systems and IoT: Innovations in Computing. Smart Innovation, Systems and Technologies, vol 141. Springer.

[17] R. Serrano-Gotarredona et al., “CAVIAR: A 45k neuron, 5M synapse, 12G connects/s AER hardware sensory-processing- learning-actuating system for high-speed visual object recognition and tracking,” IEEE Trans. Neural Networks, vol. 20, no. 9, pp. 1417–1438, 2009.

[18] F. Chollet, Deep Learning with Python. 2018.

[19] M. Elgendy, “Deep Learning for Vision Systems,” p. 475, 2019.

[20] N. Srivastava, G. Hinton, A. Krizhevsky, stskever, I., and R. Salakhut- dinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” Journal Mach. Learn. Res., vol. 299, no. 3–4, pp. 345–350, 2014

[21] A. H. Abas, N. Ismail, A. I. M. Yassin, and M. N. Taib, “VGG16 for plant imageol. 7, no. 4, pp. 90–94, 2018.

[22] V. M. Da Nobrega, T. Nepomuceno, G. Bin Bian, V. H. C. De Albuquerque, and P. P. R. Filho, “Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications,” IEEE Access, vol. 6, pp. 61677–61685, 2018.

[23] en, and S. Yan, “Network in network,” 2nd Int. Conf. Learn. Represent. ICLR 2014 - Conf. Track Proc., pp. 1–10, 2014.

[24] W. Gunaya, and I. K. G. D. Putra, “Handwriting identification using deep convolutional neural network method,” Telkomnika (Telecommunication Comput. Electron. Control., vol. 18, no. 4, pp. 1934– 1941, 2020.

[25] Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?,” Adv. Neural Inf. Process. Syst., vol. 4, no. January, pp. 3320– 3328, 2014. Suja Radha, “Leaf Disease Detection using Image Processing,” Article in Journal of Chemical and Pharmaceutical Sciences, March 2017.

[26] U.K Jaliya, Pranay Patel, “A Survey on Plant Leaf Disease Detection,” International Journal for Modern Trends in Science and Technology, April 2020.

[27] Rekha Chahar, “A Segmentation Improved Robust PNN Model for Disease Identification in Different Leaf Images,” 1st IEEE International Conference on Power Electronics, Intelligent Control and Energy Sys- tems (ICPEICES-2016).

[28] R. Newlin Shebiah S. Ananthi, S. Vishnu Varthini, “Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features,” CIGR Journal, March 2013.

[29] oja Pius Adewale Owolawi “Deep Learning Based on NAS Net for Plant Disease Recognition Using Leave Images,” 2018.

[30] la Pranathi, Kandiraju Sai Ashritha, Nagaratna B. Chittaragi, Shashidhar G. Koolagudi, “Tomato Leaf Disease Detection using Convolutional Neural Networks,” Proceedings of 2018 Eleventh International Confer- ence on Contemporary Computing (IC3), 2018.