

# Deep CNNBased Detection for Tea Clone Identification

Ade Ramdan<sup>a,\*</sup>, Endang Suryawati<sup>a</sup>, R. Budiarianto Suryo Kusumo<sup>a</sup>,  
Hilman F. Pardede<sup>a</sup>, Oka Mahendra<sup>b</sup>, Rico Dahlan<sup>c</sup>, Fani Fauziah<sup>d</sup>, Heri Syahrian<sup>d</sup>

<sup>a</sup>Research Center for Informatics

<sup>b</sup>Technical Implementation Unit for Instrumentation Development

<sup>c</sup>Research Center for Electronics and Telecommunication

<sup>a,b,c</sup>Indonesian Institute of Sciences

Jl. Sangkuriang Kampus LIPI

Bandung, Indonesia

<sup>d</sup>Research Institute for Tea and Cinchona

Kampung Gambung - Pasirjambu

Bandung, Indonesia

---

## Abstract

One factor affecting the quality of tea is the selection of plant material that would be planted on the field. Clonal selection is a common way to produce tea with better quality. However, as a natural cross pollination species, tea often consists of various clones or progenies of cross-pollinated process. This commonly occurs on plantations owned by smallholder farmers. To produce a consistent quality tea, the clones or progenies need to be identified. Usually, human experts distinguish the plants from leaves by visual inspection on the physical attributes of the leaves, such as the textures, the bone structures, and the colors. It is very difficult for non-experts or common farmers to do such identifications. In this, we propose a deep learning-based identification of tea clones. We apply deep convolutional neural network (CNN) to identify 3 types of tea clones of Gambung series, a series of tea clones developed at Research Institute of Tea and Cinchona. Our study indicates that the performance of the CNN systems are affected by the depth of the convolutional layers. VGGNet, a popular CNN architectures with 16 layers, achieves better accuracy compared to AlexNet, a CNN with 6 layers.

**Keywords:** Convolutional neural network, deep learning, gambung clone series, tea clones identification.

---

## I. INTRODUCTION

Tea (*Camellia sinensis* (L.) O. Kuntze) is one important agricultural commodity in Indonesia. It is a popular beverage commodity of the world trade. The world demand for tea is increasing year by year, making tea becomes one of the export commodities which affects the country foreign exchange of the country. The importance of this commodity cannot be overlooked and it is important to produce high quality tea products.

Good farming is one aspect that able to increase the production and the quality of the tea products. One of the most important parts of the cultivation of tea is the preparation of the planting materials. Scientists and researchers at Research Institute for Tea and Cinchona have developed various tea clones for the tea yielding [1]. It is called the GMB (Gambung) Clone series. Currently, there are eleven types of GMB series.

However, it is common to find various types of tea clones on the plantations that are owned by smallholder farmers. As tea is a natural cross pollination species, the progenies may also have different characters resulted by the cross-pollinated process. Each section of the plants with different tea clones needs to be identified to maintain and produce consistent quality.

Usually, identification is conducted by experts with visual inspection by examining physical sizes, textures, bone structures, and the color of the leaves [2]. It is difficult for non-expert or even farmers to do the identification. But, it is impossible to provide enough experts to do the identifications manually as the plantations are spread in a very large area. Therefore, a better way to identify the tea clones is needed.

Machine learning is a solution to provide automatic identification for tea clones. Given adequate number of data, we may train machine learning methods to produce models for identifications for future data predictions. Several studies of using tea leaves for identification of their categories, diseases, and pests with conventional machine learning methods such as the use of K-Nearest Neighbor (K-NN) and Support Vector Machine (SVM) have been carried out, and they have been shown to make a good performance for simple cases [3]-[7]. However, conventional machine learning methods usually work by designing suitable features that are good to discriminate the class labels [8]-[12]. However, these processes often require complicated operations. Other approaches use data extracted from expensive equipment to achieve acceptable performances [12]-[13].

Deep learning is currently a prominent technology for object identification in computer visions [14]-[15]. Studies show it is now possible to train good identification systems using simple features or images from usual and inexpensive camera devices. The new findings on Convolutional Neural Network (CNN) have

---

\* Corresponding Author.

Email: ader001@lipi.go.id

Received: July 30, 2019

; Revised: October 14, 2019

Accepted: October 28, 2019

; Published: December 31, 2019

© 2019 PPET - LIPI

opened up new opportunities in computer vision research and the possibility of using it for tea clones identifications. CNN has been shown to achieve good results in classifying and image detection [16]-[17]. Other than that, CNN success is seen in ImageNet Challenge ILSVRC [18]. It can be regarded as state-of-the-art for many object recognition tasks [19].

Some usages of deep learning technology, particularly CNN, have been reported in past studies. In [20], the use of deep convnet was able to recognize 13 different types of plants. A study reported in [21] trained and compared AlexNet and GoogleNet to identify 26 plant diseases of fourteen species of plants. A study in [22], CNN with 12 layers was applied for the identification and classification of three categories of tea, where the results of the experiment showed that CNN could achieve high accuracy. A study in [23] used plantVillage dataset to train and compare the performance of several popular CNN architectures, namely VGGNet 16-19, Inception-V3, and ResNet. A study in [24] tested the depth factor AlexNet and VGGNet to the performance and compared them with GoogleNet in detecting diseases of tomato plants.

In 2014, at ImageNet Challenge, Karen Simonyan, *et al.* [25] investigated the depth effect of a convolutional network (convnet) for its accuracy. The results showed that a significant increase in classification accuracy could be achieved by applying a depth of up to 16-19 weight layers [25]. Using conventional ConvNet architecture with a much-improved depth could achieve good performance in the ImageNet Challenge [19], [26].

In this study, we propose Deep Convolutional Neural Networks (CNN) to identify tea clones of GMB series. We use two popular CNN architectures: AlexNet and VGGNet16. We evaluate the effect of their depths on their performance to identify the clones of the tea from the leaves images. Our contributions in this paper are as follows: First, we collect a new dataset for tea clones identification. The dataset contains images of several clones of the Gambung series. Second, we apply CNN for identifications of tea clones and we investigate the effect of the depth of the convolutional layers to the performance of the identification systems, which to the best of our knowledge, has never been done.

The remainder of this paper is organized as follows. Section II describes the architectures of CNN used in the study. Section III describes the experimental setup. In Section IV, we present the results and discussion. In the final section, our research was summarized.

## II. PROPOSED METHOD

The basic architectures of CNN are illustrated in Figure 1. Meanwhile, the details of the architectures of AlexNet [19] and VGGNet [25] used for the identification of the clones of tea leaves are shown in Figure 2 and Figure 3.

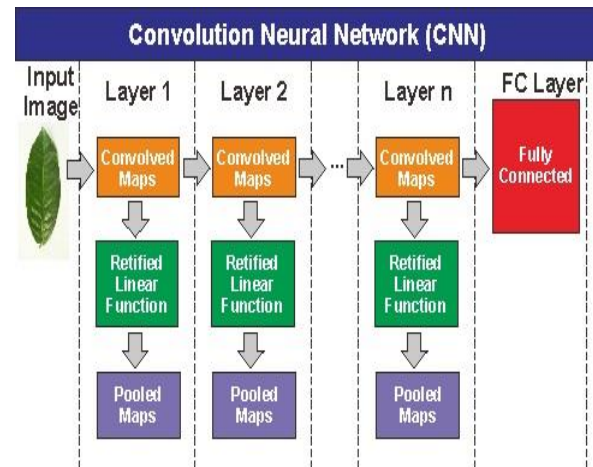


Figure 1. Architecture of CNN

### A. AlexNet

AlexNet, proposed by Krizhevsky, *et al.* [19] in 2012, is one of popular and pioneer deep convolutional neural networks (CNN). AlexNet has arguably a simpler architecture compared to many other recent architectures. The AlexNet architecture that is used in this study consists of five convolution layers, followed by two fully connected layers, and then the final output layer is softmax layer as shown in Figure 2. The number and size of the filter on each layer are as follows: 64 of  $11 \times 11$  filters on the first layer, 256 of  $5 \times 5$  filters in the second layer, 384 of  $3 \times 3$  filters on the third layer, 384 of  $3 \times 3$  filters on the fourth layer, 256 of  $3 \times 3$  filters on the fifth layer, two fully connected layers, and softmax classifier. Each layer or multiple convolutional layers are followed by a pooling layer that serves to reduce the spatial input by a down-sampling operation.

### B. VGGNet

VGGNet is also a popular deep convolutional neural networks (CNN). VGGNet is developed by K. Simonyan, *et al.* [25]. It is similar to AlexNet, but VGGNet has more convolutional layers. Additionally, VGGNet and AlexNet have other main difference. VGGNet uses small filter ( $3 \times 3$ ) throughout the layers while AlexNet uses various sizes of filters:  $11 \times 11$ ,  $5 \times 5$ , and  $3 \times 3$  filters. VGGNet architecture used in this study consists of 13 convolution layers, 2 fully connected layers, and ends with softmax classifier as shown in Figure 3. Convolutional layers on the first and the second layers use 64 of  $3 \times 3$  filters. The third and fourth layers use 128 of  $3 \times 3$  filters. The fifth, sixth, and seventh layers use 256 of  $3 \times 3$  filters. The eighth layer until the thirteenth layers use 512 of  $3 \times 3$  filters. The next two layers are two fully connected layers and the final layer is softmax classifier. Similar to AlexNet architecture, VGGNet architecture are also followed by pooling layer after convolutional layers that take the max value of the inputs (max pooling).

The architectures in this study are developed using TensorFlow machine learning frameworks with Python programming language.

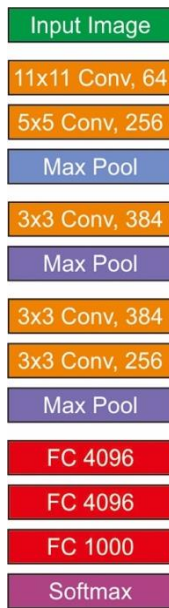


Figure 2. Architecture of AlexNet

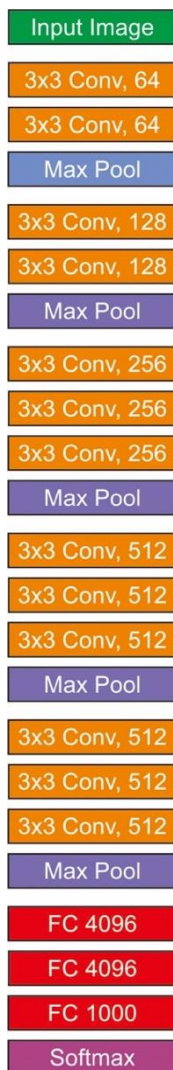





Figure 3. Architecture of VGGNet16

TABLE 1  
DATA SUMMARY CLONE BY CATEGORY

Clone Category	Sample of Image Data	Numbers of Image
3		598
9		699
11		669

### C. Dataset

In this study, we collected 1966 images of Gambung clones tea leaves dataset from three categories. The images are captured using various DSLR and smartphone cameras at Research Institute for Tea and Chinchona. We conducted the experiments in two scenarios: using two types of clones and three types of clones. For the first scenario, we use GMB 3 and GMB 9 from GMB clones series. There are a total of 1297 images for this scenario, comprise of 598 images of Gambung 3 clone and 699 images of Gambung 9 clone. For the second scenario, we use three clones of Gambung 9, 3, and 11. There are 669 images of Gambung 11 as shown in Table 1.

For the experiments, the dataset is divided into three subsets: the training set, validation set, and test set. We use 80% of the data as the training data, 10% of the data as the validation data, and the remaining is for testing. The images from the dataset are resized to 64×64 pixels in these experiments. For the setting parameters in training, we use Adam [27]-[29] for an adaptive learning rate with initial learning rates are varied with two values:  $10^{-4}$  and  $10^{-5}$ . We also varies the number of training epochs: 50 and 100 number of epochs for the experiments.

### III. RESULT AND DISCUSSION

Many studies view deep learning methods as feature learning methods [30]-[32]. Given simpler features, such as RGB as in this study, the complex architectures allow them to find complex non-linear relations between the data and the target labels. Each layer learns different abstraction of data and their local correlations and while the data are passed through the next layers, the networks may learn their global correlations. This is one advantage of using CNN for images.

As shown in Figure 4, we see that CNN produces completely different abstraction of the data given their RGB values. The input images are read by convolution layers by using a filter called the convolved maps. Then, features maps are passed through a non-linearity unit which is Rectified Linear Units (RELU) and processed by a pooling layer. Furthermore, on pooling layer, each

feature map is downsampled with pooling max method to get the greatest value of the features maps called pooled maps. In a sense, the max pooling would extract the edges of the data. After performing convolution and unification on the last layer, the results are incorporated into a fully connected layer for transforming the data dimension so that the data can be classified linearly. It is clear that CNN can capture the information about the edges and structures of the data, a common trait that is usually used by experts to perform identification using visual inspections.

The experimental results for the GMB clone series are shown in Table 2 and 3. AlexNet and VGGNet architectures achieve good accuracy (above 80%) when we use  $10^{-4}$  as learning rate, confirming that CNN is promising for clone identification task.

We notice that the learning rate setting is influential for deep neural network training to achieve good performance. The optimal learning rate is highly dependent on the architecture models and datasets. AlexNet and VGGNet models with two and three data classes produce high accuracy in 100 epochs with optimal learning rate setting  $10^{-4}$ , indicating that the learning rate needs to be selected carefully to find optimal performance. Table 3 shows that the highest accuracy for AlexNet (96.15%) is achieved when the learning rate is set at  $10^{-4}$ , while for VGGNet, the highest accuracy is achieved at 97.31%.

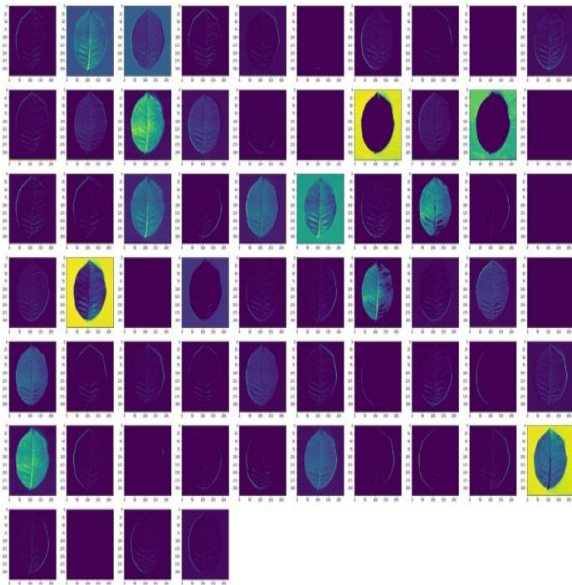


Figure 4. Visualization of convolution layer output

TABLE 2  
PERFORMANCE OF DIFFERENT CNN ARCHITECTURE MODELS FOR IDENTIFICATION OF TEA CLONES GMB SERIES ON TRAINING AND TESTING DATASET WITH A VALUE OF 50 EPOCHS

LR	Scenario	AlexNet		VGGNet16	
		Accuracy (%)	Loss	Accuracy (%)	Loss
$10^{-4}$	1	95.38	0.16	96.15	0.18
	2	81.98	1.29	82.49	1.05
$10^{-5}$	1	93.46	0.34	96.92	0.17
	2	72.59	1.44	73.95	0.81

TABLE 3  
PERFORMANCE OF DIFFERENT CNN ARCHITECTURE MODELS FOR IDENTIFICATION OF TEA CLONES GMB SERIES ON TRAINING AND TESTING DATASET WITH A VALUE OF 100 EPOCHS.

LR	Scenario	AlexNet		VGGNet16	
		Accuracy (%)	Loss	Accuracy (%)	Loss
$10^{-4}$	1	96.15	0.29	97.31	0.29
	2	81.47	1.15	85.79	1.69
$10^{-5}$	1	96.15	0.19	96.92	0.25
	2	78.43	1.50	79.70	1.59

We notice that the learning rate setting is influential for deep neural network training to achieve good performance. The optimal learning rate is highly dependent on the architecture models and datasets. AlexNet and VGGNet models with two and three data classes produce high accuracy in 100 epochs with optimal learning rate setting  $10^{-4}$ , indicating that the learning rate needs to be selected carefully to find optimal performance. Table 3 shows that the highest accuracy for AlexNet (96.15%) is achieved when the learning rate is set at  $10^{-4}$ , while for VGGNet, the highest accuracy is achieved at 97.31%.

However, we observe that while VGGNet achieves better performance, it may be overfitted as shown in Figure 5, 6, 7 and 8. The results of the AlexNet progressions of the loss and accuracy for each epoch are shown in Figure 5 and 6. The model is quite quickly converging around 40 epochs and remains stable in both scenarios. The loss of AlexNet suggests that the model is in good fit condition with its training loss and validation loss values decreased to a stable point and fluctuate with only a very small values.

The progression of the loss and the accuracy of the VGGNet model are shown in Figure 7 and 8. There is a big difference between accuracy and loss. It shows that the model has good accuracy but its loss curve suggest the model is in an overfit condition. The training loss value continues to decline to a stable point. The validation value also declines to a stable point, but tend to increase again after epoch 40, so that it has a big gap with the training loss value. This may affect the robustness of the methods that we would like to investigate in future studies.

From these observations, we may say that while the deeper convolutional network might increase loss because the gradient may increase dramatically in back propagation process when many layers of convolutional networks are used. The number of epochs should be evaluated carefully to avoid overfitting or some regularizations may be applied to overcome the problems. A certain automatic stopping conditions such as early stopping could also be implemented.

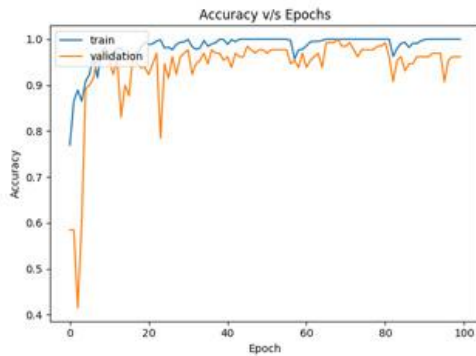


Figure 5. AlexNet Accuracy

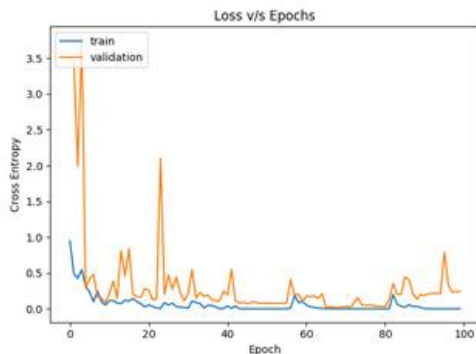


Figure 6. AlexNet Loss

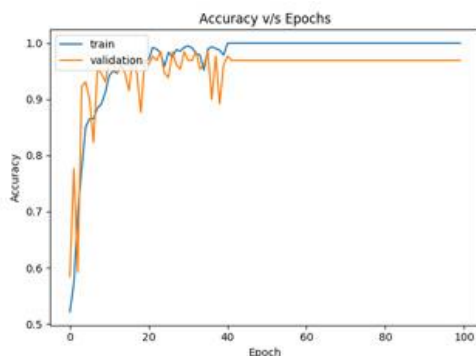


Figure 7. VGGNet Accuracy

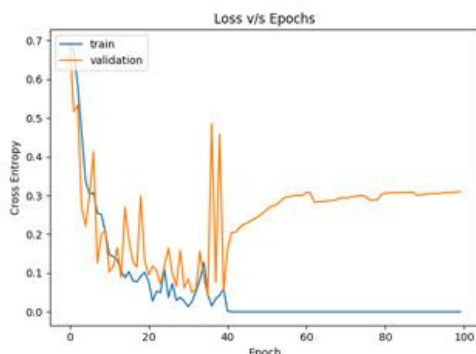


Figure 8. VGGNet Loss

The study suggests that while the identification may benefit from having more convolutional layers, several precautions may be needed to avoid overfitting. As shown in our experiments of using 5 layers in AlexNet and 13 layers in VGGNet indicated that the deeper a network, its optimization capabilities can be improved, but they tend to overfit more easily. It shows that the

optimal depth of a convolutional network helps to increase classification accuracy.

## CONCLUSION

In this study, we develop a clone detection systems for tea using Deep Convolutional Neural Networks (CNN). Two architectures of CNN, VGGNet and AlexNet are applied. The results proved that the depth of the AlexNet and VGGNet model influences their ability to produce a good model performance with VGGNet shows to be slightly better. With deeper network, we can improve and produce a high accuracy but it is prone to overfitting. While AlexNet has networks that are not so deep compared to VGGNet, it can produce high accuracy with more consistent loss, resulting in a fitting or good fit model.

Our study also points out that the optimization algorithm such as Adam may depend on the initial learning rate setting to get good performance. Setting the right parameters for CNN architecture, especially for VGGNet and AlexNet models to get an optimal training time and accurate identification, is still a challenging problem. Therefore, future research directions can be focused on researching a method to automate the process of setting CNN parameters. Developing identifications for more types of clones is also in our future plan.

## ACKNOWLEDGEMENT

The study is partially funded by INSINAS grant from Indonesian Ministry of Research, Technology, and Higher Education. The experiments in this paper are conducted at High Performance Computing Facilities - Research Center for Informatics, Indonesian Institute of Sciences (LIPI). We thank our fellow researchers in the Research Center for Informatics LIPI and Research Institute for Tea and Cinchona (PPTK) Gambung which has provided assistance in this study.

## REFERENCES

- [1] Dedi Soleh Effendi, M. Syakir, M. Yusron, and Wiratno, "Budidaya dan Pasca Panen The," Pusat Penelitian dan Pengembangan Perkebunan, 2010.
- [2] (2019) Gambung website [online]. Available: <https://www.gamboeng.com/pages/detail/2015/59/146>
- [3] Xueyan Wu, Jiquan Yang, and Shuihua Wang, "Tea category Identification Based on Optimal Wavelet Entropy and Weighted K-Nearest Neighbors Algorithm," *Journal Multimedia Tools and Applications*, vol. 77, no. 3, pp. 3745-3759, 2018.
- [4] Shuihua Wang, Preetha Phillips, Aijun Liu, and Sidan Du, "Tea Category Identification using Computer Vision and Generalized Eigenvalue Proximal SVM," *Journal Fundamenta Informaticae*, vol. 151, no. 1-4, pp. 325-339, 2017.
- [5] Shuihua Wang, Xiaojun Yang, Yudong Zhang, Preetha Phillips, Jianfei Yang, and Ti-Fei Yuan, "Identification of Green, Oolong and Black Teas in China via Wavelet Packet Entropy and Fuzzy Support Vector Machine," *Journal Entropy*, vol. 17, no. 10, pp. 6663-6682, 2015.
- [6] B. Karunamoorthy, and D. Somasundereswari, "A Defect Tea Leaf Identification Using Image Processing," *Journal Przegląd Elektrotechniczny*, vol. 89, no. 9, pp. 318-320, 2013.
- [7] Christophorus Candra Kusumadewa and Supatman, "Identifikasi Citra Daun Teh Menggunakan Metode Histogram untuk Deteksi Dini," *Journal Multimedia & Artificial Intelligence*, vol. 2, no. 1, pp. 27-36, 2018.
- [8] Qing Xia, Hao-Dong Zhu, & Yong Gan, and Li Shang, "Plant

- Leaf Recognition Using Histograms of Oriented Gradients,” in *International Conference on Intelligent Computing*, vol. 8589, 2014, pp. 369-374.
- [9] D. G. Tsolakidis, D. I. Kosmopoulos, and G. Papadourakis, “Plant leaf recognition using zernike moments and histogram of oriented gradients,” in *Hellenic Conference on Artificial Intelligence*, 2014, pp. 406-417.
- [10] Voncarlos A., Alceu S. B., Andre L. B., Alessandro L. K., and Rosane F., “Multiple Classifier System for Plant Leaf Recognition,” in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp 1880-1885.
- [11] Yogesh Dandawate and Radha Kokare, “An automated approach for classification of plant diseases towards development of futuristic decision supportsystem in indian perspective,” in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2015, pp. 794-799.
- [12] Anne-Katrin Mahlein, “Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping,” *Plant Disease*, vol. 100, no. 2, pp. 241-251, 2016.
- [13] Godliver Owomugisha, Friedrich Melchert, Ernest Mwebaze, John A. Quinn, and Michael Biehl, “Machine Learning for Diagnosis of Disease in Plants using Spectral Data,” in *International Conference on Artificial Intelligence (ICAI)*, 2018, pp. 9-15.
- [14] Yu Sun, Yuan Liu, Guan Wang, and Haiyan Zhang, “Deep Learning for Plant Identification in Natural Environment,” *Computational Intelligence and Neuroscience*, vol. 2017, 2017.
- [15] Ajeet Ram Pathak, Manjusha Pandey, and Siddharth Rautaray, “Application of Deep Learning for Object Detection,” in *Procedia Computer Science*, vol. 132, pp. 1706-1717, 2018.
- [16] T. L. I. Sugata, and C. K. Yang, “Leaf App: Leaf Recognition with Deep Convolutional Neural Networks,” in *IOP Conference Series: Materials Science and Engineering*, vol. 273, 2017.
- [17] Agustinus Kristiadi and Pranowo, “Deep Convolutional Level Set Method for Image Segmentation,” *Journal of ICT Research and Applications*, vol. 11, no. 3, pp. 284-298, 2017.
- [18] Russakovsky, O., Deng, J., Su, H. et al., “ImageNet Large Scale Visual Recognition Challenge,” *International Journal of Computer Vision*, vol. 115, no. 3, pp 211-252, 2015.
- [19] Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Neural Information Processing Systems*, pp. 1097-1105, 2012.
- [20] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic, “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,” *Journal Computational Intelligence and Neuroscience*, vol. 2016, 2016.
- [21] Sharada P. Mohanty, David P. Hughes and Marcel Salathé, “Using Deep Learning for Image-Based Plant Disease Detection,” *Frontiers in Plant Science*, 2016.
- [22] Yu-Dong Zhang, Khan Muhammad, and Chaosheng Tang, “Twelve-layer Deep Convolutional Neural Network with Stochastic Pooling for Tea Category Classification on GPU Platform,” *Journal Multimedia Tools and Applications*, vol. 77, no. 17, pp. 22821-22839, 2018.
- [23] Guan Wang, Yu Sun, and Jianxin Wang, “Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning,” *Journal Computational Intelligence and Neuroscience*, vol. 2017, 2017.
- [24] Endang Suryawati, Rika Sustika, R. Sandra Yuwana, Agus Subekti, and Hilman Pardede, “Deep Structured Convolutional Neural Network for Tomato Diseases Detection,” in *International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2018, pp. 385-390.
- [25] Karen Simonyan and Andrew Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” arXiv preprint arXiv:1409.1556, 2014.
- [26] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, “Backpropagation Applied to Handwritten Zip Code Recognition,” *Neural Computation*, vol. 1, no. 4, pp. 541-551, 1989.
- [27] Diederik P. Kingma and Jimmy Lei Ba, “Adam: A Method for Stochastic Optimization,” arXiv preprint arXiv:1412.6980, 2014.
- [28] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks,” arXiv preprint, 2017.
- [29] Kelvin Xu, Jimmy Lei Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Richard S. Zemel, and Yoshua Bengio, “Show, Attend and Tell: Neural Image Caption Generation with Visual Attention,” in *International Conference on Machine Learning*, 2015, pp. 2048-2057.
- [30] Yoshua Bengio, Aaron Courville, and Pascal Vincent, “Representation Learning: A Review and New Perspectives,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, 2013.
- [31] Yan Xu, Tao Mo, Qiwei Feng, Peilin Zhong, Maode Lai, and Eric I-Chao Chang, “Deep Learning of Feature Representation with Multiple Instance Learning for Medical Image Analysis,” in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 1626-1630.
- [32] Frédéric Li, Kimiaki Shirahama, Muhammad Adeel Nisar, Lukas Köping, and Marcin Grzegorzec, “Comparison of Feature Learning Methods for Human Activity Recognition Using Wearable Sensors,” *Sensors*, vol. 18, no. 2, 679, 2018