

## Classification of tomato leaf diseases using MobileNet V2

Siti Zulaikha Muhammad Zaki, Mohd Asyraf Zulkifley, Marzuraikah Mohd Stofa, Nor Azwan  
Mohammed Kamari, Nur Ayuni Mohamed

Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti  
Kebangsaan Malaysia, 43600, UKM Bangi, Selangor, Malaysia

---

### Article Info

#### Article history:

Received Jan 20, 2020

Revised Mar 18, 2020

Accepted Apr 1, 2020

#### Keywords:

Deep learning classification

MobileNet V2

Tomato disease

---

### ABSTRACT

Tomato is a red-colored edible fruit originated from the American continent. There are a lot of plant diseases associated with tomatoes such as leaf mold, late blight, and mosaic virus. Tomato is an important vegetable crop that contributes to the world economically. Despite tremendous efforts in plant management, viral diseases are notoriously difficult to control and eradicate completely. Thus, accurate and faster detection of plant diseases is needed to mitigate the problem at the early stage. A computer vision approach is proposed to identify the disease by capturing the leaf images and detect the possibility of the diseases. A deep learning classifier is utilized to make a robust decision that covers a wide variety of leaf appearances. Compact deep learning architecture, which is MobileNet V2 has been fine-tuned to detect three types of tomato diseases. The algorithm is tested on 4,671 images from PlantVillage dataset. The results show that MobileNet V2 is able to detect the disease up to more than 90% accuracy.

*This is an open access article under the [CC BY-SA](#) license.*



---

### Corresponding Author:

Siti Zulaikha Muhammad Zaki,  
Department of Electrical, Electronic and Systems Engineering,  
Universiti Kebangsaan Malaysia,  
43600 UKM, Bangi Selangor, Malaysia.  
Email: A159764@siswa.ukm.edu.my

---

## 1. INTRODUCTION

Tomato (*Solanum Lycopersicum*), is one of the most widely consumed vegetable crops in Malaysia [1]. Tomato is regarded as a good nutritional source for vitamin C. However, tomato plants are prone to be infected by various diseases. Some of the disease pathogens are fungal organisms, while others are bacterial or viral [2]. To curb these diseases from spreading, various types of pesticides are used to kill the pathogens. The widespread use of these chemicals will pose harm towards human health as well as the environment [3]. Several popular tomato plant diseases are late blight, leaf mold, and mosaic virus. Late blight [4] is a potentially serious tomato disease caused by infestation of the phytophthora fungus. It causes lesions, which are small, dark and looks water-soaked spot. These leaf spots will quickly enlarge and a white mold will appear at the margins of the affected areas. Another fungus, the passalora fulva is the caused for leaf mold, which usually occurs on the older leaves that are closed to the soil where air circulation is poor and humidity is high. The initial symptom of this disease is pale green or yellowish spots on the upper leaf surface, which will enlarge and turn to be more distinctive yellow spots. The tomato mosaic virus [5] belongs to the tobamoviridae family, which is a pathogenic virus among the plants. The symptom can be observed at any stage of plant growth and it affects all parts of the plant.

Recognition and classification of plant diseases play a vital role in the field of agriculture. The quality, quantity, and productivity of the plants depend on the timely detection of the diseases. Therefore, an automatic system needs to be developed so that the detection process can be made autonomous

with the minimal human intervention [6]. This system will help the farmers in diagnosing the disease at the early stage and allow them to perform mitigation actions before it is too late. One of the ways to detect the diseases is by using a deep learning approach using a supervised learning technique. Shadow occlusion problem is not considered, where it can easily be removed using shadow removal [7]. In this paper, the parameters of MobileNet V2 network is trained to classify tomato leaf diseases into their respective classes, before validation test is done to measure the effectiveness of the proposed system. In this paper, a subset of PlantVillage dataset of three types of tomato diseases and healthy leaves are downloaded from the Kaggle platform, which includes 373 tomato mosaic virus images, 1756 late blight images, 952 leaf mold images, and 1590 healthy tomato images [8]. The dataset is split into two groups of training and testing subsets. To meet the input requirements of the MobileNet V2 model, the input image size is rescaled to 224×224 pixels.

## 2. LITERATURE REVIEW

### 2.1. Deep learning

Deep learning is the state-of-the-art machine learning method, which utilizes a complex network of artificial nodes with large amounts of hidden layers. Many of the techniques before the introduction of deep learning classify the task through semantic features information. Some examples of the semantic features are corners, edges, shapes, etc. A deep learning approach does not require the design of features ahead of time. These features are the results of optimum automatic learning. Therefore, this method is robust to various modes in the data as these features are not handcrafted [9]. Some examples of the applications of deep learning approach are in object tracking [10-11], disease screening [12-13], physiotherapy [14], face retrieval [15], and remote sensing [16].

### 2.2. Development convolutional neural network (CNN)

#### 2.3.1. GoogleNet

GoogLeNet is the winner of ILSVRC 2014, which has been proposed by [17]. The network allows parallel branches of convolutional neural networks with various kernel sizes. It contains 22 layers of network with more than 7,000,000 parameters. As for reference, AlexNet has only around 60,000,000 parameters, which is more than 10 times less number of trainable parameters. As a consequence, the complexity of GoogLeNet processing is also much lower than AlexNet's [18]. In general, GoogLeNet has also been proved to be consistently more accurate than AlexNet.

#### 2.3.2. Residual network (ResNet)

ResNet [19] was first introduced in 2015, where it has also won ILSVRC competition with an error rate of 3.57%. ResNet's high accuracy rate can be mainly attributed to the introduction of residual layers that allow the network to be designed deeper compared to the previous popular network architectures [20]. The residual layer or also known as identity mapping mitigates the problem of diminishing gradient in training a deep network, where the previous layer is fed to the later layers. The idea was to overcome the reduction of input features from the original learning feature that produces zero features [21].

#### 2.3.4. MobileNet V2

MobileNet is a deep learning architecture that focuses on the mobile platform where the computational resource is limited. An improved version, which is called MobileNet V2 [22] is then introduced by Google with slight modifications to the original version. The basis of the network still remains the same, which is separable convolution. MobileNet version 2 previously trained on ImageNet datasets has been used to extract fruit image features in [23]. The paper claimed that the parameters used have reduced from 4.24 millions to just 3.47 millions, but with better accuracy.

## 3. RESEARCH METHOD

### 3.1. Disease detection using MobileNet

MobileNet V2 is an improvement over MobileNet V1 [24]. Both of them still retain separable convolution as the core layer, where the number of parameters trained is much reduced compared to the full convolutional. The small requirement of the parameter number allows MobileNet V2 suitable for mobile phone applications, where the number of registers is much less compared to desktop. Separable convolution is divided into two distinct steps, which are depthwise convolution and pointwise convolution [25].

**3.1.1. Depthwise convolution**

Depthwise convolution is a reduced version of convolution, where each channel will undergo the process separately. An original convolution with height × width × channel input feature map is divided into several groups depending on the number of channels, which signifies the depth. The reduction in kernel size will also follow the same grouping. The depthwise convolution collects spatial features separately and thus the number of parameters needed also reduced significantly.

**3.3.2. Pointwise convolution**

Pointwise convolution is the opposite of the depthwise convolution. The width and height of the kernel are set to 1, but the depth will depend on the number of input channels. It will be cascaded right after depthwise convolution to represent the full convolution but with a much lesser number of parameters. It can be used to set the dimensions of the output channel features.

**3.4. The method based on MobileNet V2**

Figure 1 shows all components of the system where the PlantVillage dataset is used to train and test the proposed MobileNet V2-based tomato disease screening algorithm. 4671 images are extracted from the full dataset that includes 1590 images of healthy tomato leaves, 952 leaf mold images, 1756 late blight images, and 373 mosaic virus images. All leaves are captured individually without interference from other leaves as shown in Table 1. In these experiments, the images are randomly split into two subsets: 3471 training images and 1200 testing images. To meet the input requirement of the MobileNet V2 model, all images are rescaled to 224 × 224 pixels.

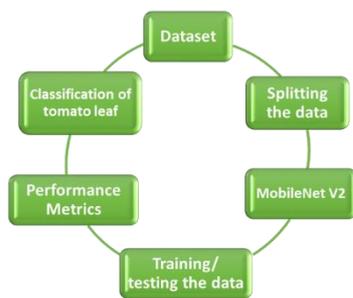


Figure 1. Main components of the tomato diseases detection using MobileNet V2

Table 1. Some samples of tomato leaf diseases

Class	Sample Image of tomato leaves		
Healthy Leaf			
Leaf Mold			
Late Blight			
Mosaic Virus			

After that, MobileNet V2 is trained from scratch through random initialization or using transfer learning techniques. Subsequently, a training progress chart is plotted to demonstrate the performance MobileNet V2 in classifying the tomato diseases. The performance metric of accuracy is then used to measure classification performance, which is divided into four classes healthy tomato leaves, late blight, leaf mold, and mosaic virus.

**4. RESULTS AND DISCUSSION**

In this paper, experimental results are tested using HP Intel core i7-3770 @ 3.9GHz CPU with 8 GB memory. No graphic processing unit is utilized where normal CPU-based TensorFlow is implemented using the Python platform. Several hyper-parameter configurations of the MobileNet V2 are tested that include

batch size, optimizer selection and learning rate. The testing process is done sequentially where the best setup of each hyper-parameter is tested separately to find the optimal setting.

**4.1. Optimization method**

The following Table 2 and Figure 2 shows the classification results of five different optimizers, namely Adagrad, Adam, SGD, RMSprop and Nadam. Among these five optimizers, Adagrad optimizer gives the best accuracy of 0.9434, followed by Adam and SGD optimizers with accuracy of 0.8996 and 0.8558, respectively. RMSProp and Nadam return a low accuracy in classifying the tomato plant diseases.

Table 2. Training and testing subset final accuracy and final loss of various optimizers

Optimizers	Final Accuracy	Final Loss
Adagrad	0.9434	0.1554
Adam	0.8996	0.2587
SGD	0.8558	0.4306
RMSprop	0.6282	4.3899
Nadam	0.0908	20.5444

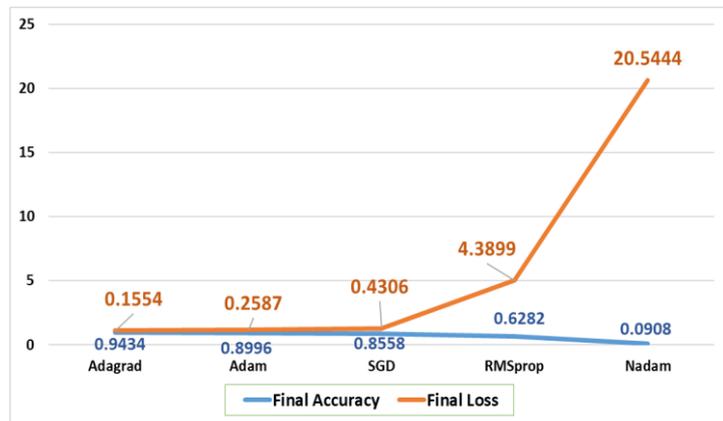


Figure 2. Testing accuracy and final loss of various optimizer methods

**4.2. Learning rate**

Learning rate is a hyper-parameter that controls how much the gradient error will be used to update the current weights. In this test, Adam optimizer is selected as the core for the testing because of the noticeable difference once the rate is varied. Three rates are tested, which are 0.01, 0.001 and 0.0001 as shown in Table 3. The best-performed learning rate given the same number of training epoch is recorded when a rate of 0.001 is used to train the MobileNet V2. The results in Figure 3 shows that an accuracy of 0.8996 is achieved when a learning rate of 0.001 is implemented.

Table 3. Final accuracy and final loss of various learning rates

Learning Rate	Final Accuracy	Final Loss
0.00001	0.7585	0.5957
0.0001	0.8996	0.2587
0.001	0.7606	3.3920

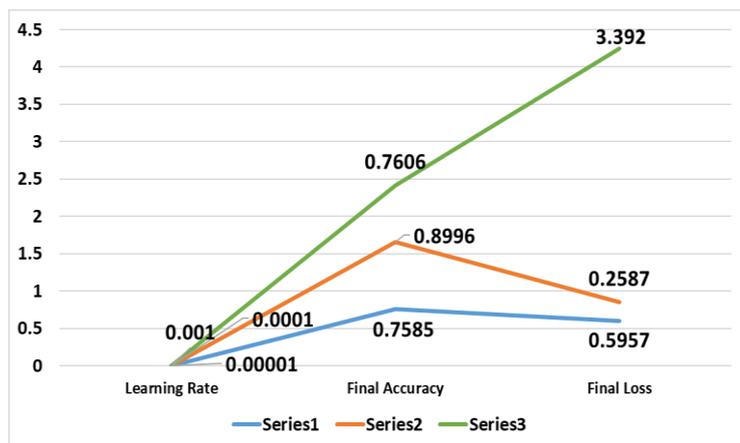


Figure 3. Final accuracy and final loss for various learning rates

**4.3. Training and Testing Subset**

Training data is the dataset used to train the MobileNet V2 (weights and biases in the case of standard CNN), while testing data is the sample that is used to evaluate the performance of the trained network. Inspired by [26], four ways of data division as shown in Table 4 are explored between training and

testing data, which are 9:1, 4:1, 7:3 and 3:2 ratios. The best classification result of the tomato disease classification is obtained when a split ratio of 4:1 is used between training and testing data with an accuracy value of 0.9562 as shown in Figure 4.

Table 4. Final accuracy and final loss between training and testing subset

Training	Testing	Final Accuracy	Final Loss
4203	468	0.9083	0.3380
3736	935	0.9562	0.1402
3269	1402	0.7612	0.8120
2802	1869	0.3866	0.8583

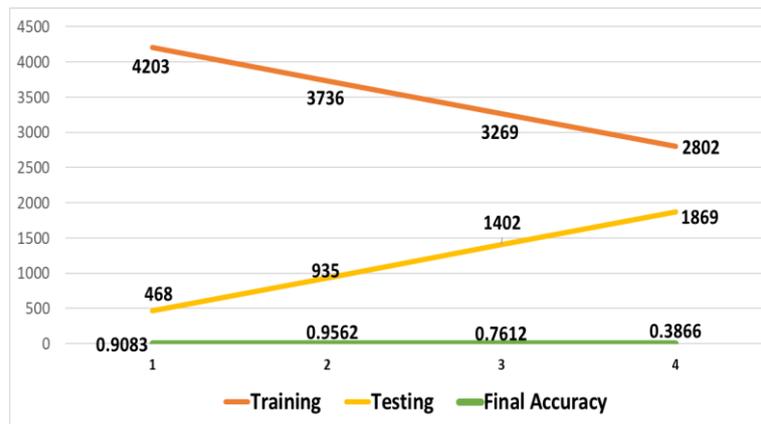


Figure 4. Final accuracy and final loss between training and testing subset

**4.4. Batch Size**

Batch size is a hyperparameter that controls the number of images is fed to the network for one training iteration. It allows local analysis of several images instead of an individual image. Less fluctuation in training error is observed when a batch size method is used, but a too-large batch size will result in over-generalization. Three batch sizes are explored that includes 16, 32 and 48. Table 5 shows the classification accuracy when the batch size is increased from 16 to 48. The best performance is obtained when the batch size of 16 is used with an accuracy value of 0.9594. Figure 5 also reveals that accuracy is decreasing once the size is increased from 16 to 48.

Table 5. Final accuracy and final loss of various batch sizes

Batch Size	Final Accuracy	Final Loss
48	0.7970	0.6940
32	0.8632	0.3912
16	0.9594	0.1265

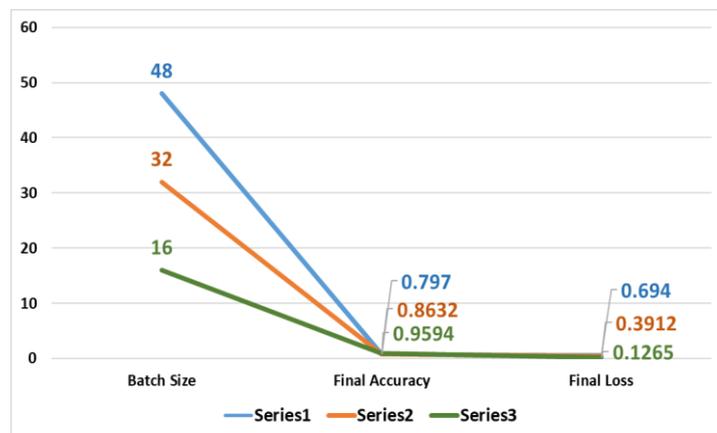


Figure 5. Final accuracy and final loss between batch size

## 5. CONCLUSION

In conclusion, MobileNet V2 has successfully implemented to classify various tomato plant diseases based on captured leaf images. The best classification performance is obtained when MobileNet V2 is trained using Adagrad with a batch size of 16. The experimental results also prove that a learning rate of 0.001 and data division of 4:1 ratio between training and testing deliver the most accurate classification performance. For future work, all classes in the PlantVillage will be explored instead of just three diseases.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge fundings from Universiti Kebangsaan Malaysia (GUP-2019-008) and Ministry of Education Malaysia (FRGS/1/2019/ICT02/UKM/02/1). The graphic processing unit used in this research is donated by NVIDIA Corporation (KK-2019-005).

## REFERENCES

- [1] S. Sato *et al.*, "The tomato genome sequence provides insights into fleshy fruit evolution," *Nature*, vol. 485, no. 7400, pp. 635–641, 2012.
- [2] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors (Switzerland)*, vol. 17, no. 9, 2017.
- [3] H. Durmus, E. O. Gunes, and M. Kirci, "Disease detection on the leaves of the tomato plants by using deep learning," *6th Int. Conf. Agro-Geoinformatics, Agro-Geoinformatics 2017*, 2017.
- [4] M. Nowicki, M. R. Foolad, M. Nowakowska, and E. U. Kozik, "Potato and Tomato Late Blight Caused by *Phytophthora infestans*: An Overview of Pathology and Resistance Breeding," *Plant Dis.*, vol. 96, no. 1, pp. 4–17, 2012.
- [5] I. M. Hanssen *et al.*, "Seed transmission of Pepino mosaic virus in tomato," *Eur. J. Plant Pathol.*, vol. 126, no. 2, pp. 145–152, 2010.
- [6] W. Zaka-ud-din, Muhammad Aziz, Sumair Rashid, Junaid Ismail, "Classification of Disease in Tomato Plants ' Leaf Using Image Segmentation and SVM," *Tech. Journal, Univ. Eng. Technol.*, vol. 23, no. 2, pp. 81–88, 2018.
- [7] M. A. Zulkifley, B. moran, "Enhancement of robust foreground detection through masked GreyWorld and color co-occurrence approach," *3rd IEEE International Conference on Computer Science and Information Technology*, vol. 4, pp. 131–136, 2010.
- [8] A. K. Rangarajan, R. Purushothaman, and A. Ramesh, "Tomato crop disease classification using pre-trained deep learning algorithm," *Procedia Comput. Sci.*, vol. 133, pp. 1040–1047, 2018.
- [9] Q. Xiang, G. Zhang, X. Wang, J. Lai, R. Li, and Q. Hu, "Fruit image classification based on Mobilenetv2 with transfer learning technique," *3rd International Conference on Computer Science and Application Engineering*, pp. 1-7, 2019.
- [10] Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, S. Nasrin, and V. K. Asari, "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches," *arXiv*, vol. abs/1803.01164, pp. 1-39, 2018.
- [11] M. A. Zulkifley and N. Trigoni, "Multiple-Model Fully Convolutional Neural Networks for Single Object Tracking on Thermal Infrared Video," *IEEE Access*, vol. 6, pp. 42790-42799, 2018.
- [12] M. A. Zulkifley, "Two streams multiple-model object tracker for thermal infrared video," *IEEE Access*, vol. 7, pp. 32383-32392, 2019.
- [13] S. B. Jadhav, "Convolutional Neural Networks for Leaf Image-Based Plant Disease Classification," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 8, no. 4, pp. 328–341, 2019.
- [14] M. A. Zulkifley, S. R. Abdani, and N. H. Zulkifley, "Pterygium-Net: a deep learning approach to pterygium detection and localization," *Multimedia Tools and Applications*, vol. 78, no. 24, pp. 34563-34584, 2019.
- [15] M. A. Zulkifley, N. A. Mohamed, and N. H. Zulkifley, "Squat Angle Assessment Through Tracking Body Movements," *IEEE Access*, vol. 7, pp. 48635-48644, 2019.
- [16] I. Hachchane, A. Badri, A. Sahel, and Y. Ruichek, "Large-Scale Image-To-Video Face Retrieval with convolutional neural network Features," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 1, pp. 40-45, 2020.
- [17] S. R. Abdani, and M. A. Zulkifley, "DenseNet with Spatial Pyramid Pooling for Industrial Oil Palm Plantation Detection," *2019 International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE)*, pp. 134–138, 2019.
- [18] 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, 2016.
- [19] B. Ashqar and S. Abu-Naser, "Image-Based Tomato Leaves Diseases Detection Using Deep Learning," *Int. J. Acad. Eng. Res.*, vol. 2, no. 12, pp. 10–16, 2018.
- [20] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front. Plant Sci.*, vol. 7, pp. 1–10, 2016.
- [21] A. Krizhevsky, "ImageNet Classification with Deep Convolutional Neural Networks," *ImageNet Classif. with Deep Convolutional Neural Networks Alex*, pp. 1–9, 2012.
- [22] W. Liu and K. Zeng, "SparseNet: A Sparse DenseNet for Image Classification," *ArXiv*, vol. abs/1804.05340, pp. 1-17, 2018.

- [23] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4510–4520, 2018.
- [24] G. Suresh, V. Gnanaprakash, and R. Santhiya, "Performance Analysis of Different CNN Architecture with Different Optimisers for Plant Disease Classification," *2019 5th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2019*, pp. 916–921, 2019.
- [25] L. Kaiser, A. N. Gomez, and F. Chollet, "Depthwise Separable Convolutions for Neural Machine Translation," *ArXiv*, vol. abs/1706.03059, pp. 1-10, 2017.
- [26] Z. Kadim, M. A. Zulkifley, N. Hamzah, "Deep-learning based single object tracker for night surveillance," in *IAES International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 4, 2020.

## BIOGRAPHIES OF AUTHORS



Siti Zulaikha Muhammad Zaki is a final year student at the Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia. Her research interest is deep learning application in computer vision.



Mohd Asyraf Zulkifley received his B.Eng. degree in mechatronics from International Islamic University Malaysia in 2008 and the Ph.D. degree in electrical and electronic engineering from The University of Melbourne in 2012. He was a sponsored Researcher at the Department of Computer Science, University of Oxford from 2016 to 2018. He is currently an Associate Professor at the Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia. His current research interests are visual object tracking and medical image analysis.



Marzuraikah Mohd Stofa is a final year student at the Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia. Her research interest is deep learning application in computer vision.



Nor Azwan Mohammed Kamari received B.Eng. from Meiji University, Japan, M.Sc. degrees from Ehime University, Japan and PhD degree from Universiti Teknologi Mara, Malaysia 2000, 2004 and 2016, respectively. He has authored and co-authored more than 30 technical papers published in the international journals, national and international conferences. His research interests include power system stability and artificial intelligence.



Nur Ayuni Mohamed received her Bachelor Degree of Engineering and Master of Science from University Kebangsaan Malaysia. Currently, she is a final year Ph.D. student at the Universiti Kebangsaan Malaysia, researching on deep learning-based fall detection system.