Performance Comparison of Convolutional Neural Network for COVID-19 Diagnosis

Suhaim Parvez Wadekar[1], Koon Meng Ang[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Tiong Hoo Lim[5], Chun Kit Ang[1], Sew Sun Tiang[1,*], Wei Hong Lim[1,*]

[1]*Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia*
[2]*Maynooth International Engineering College, Maynooth University, Maynooth, Co Kildare, Ireland*
[3]*Maynooth International Engineering College, Fuzhou University, Fujian, 350116, China*
[4]*Department of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun 248002, India*
[5]*Faculty of Engineering, Universiti Teknologi Brunei, Bandar Seri Begawan 1410, Brunei Darussalam*
E-mail: 1001748138@ucsiuniversity.edu.my, 1001436889@ucsiuniversity.edu.my, chinhong.wong@mu.ie, abhishek15491@gmail.com, lim.tiong.hoo@utb.edu.bn, angck@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my limwh@ucsiuniversity.edu.my

Abstract

COVID-19 has devastated the global healthcare system as well as the world economy with more than 600 million confirmed cases and 6 million deaths globally. A timely and accurate diagnosis of the disease plays a vital role in the treatment and preventative spread of disease. Recently, deep learning such as Convolutional Neural Networks (CNNs) have achieved extraordinary results in many applications such as medical classifications. This work focuses on investigating the performance of nine state-of-the-art architectures: Alexnet, Googlenet, Inception-v3, Mobilenet-v2, Resnet-18, Resnet-50, Shufflenet, Squeezenet and Resnet-50 RCNN for COVID-19 classification by comparing with performance metrics such as accuracy, precision, sensitivity, specificity and F-score. The datasets considered in current study are divided into three different classes namely Normal Chest X-Rays (CXRs), Pneumonia patient CXR and COVID-19 patient CXR. The results achieved shows that Resnet-50 RCNN achieved an accuracy, precision, sensitivity, specificity and F-score of 95.67%, 95.71%, 95.67%, 97.84% and 95.67% respectively.

Keywords: Chest X-ray, convolutional neural network, COVID-19 diagnosis, deep learning

1. Introduction

COVID-19 is a highly infectious and contagious viral disease caused by the SARS-COV-2 virus. Over the past three years, this disease has caused ravage throughout the world by creating catastrophic damage on global economic as well as irreversible losses of human life. This disease first emerged in Wuhan, China in December 2019. As of January 2023, a total of 667,510.073 infection cases and 6,707,374 death cases are reported by World Health Organization (WHO)[1].

Despite being considered as the golden test used for COVID-19 diagnosis, reverse transcription polymerase chain reaction (RT-PCR) is a time-consuming test and it could take up to 48 hours to confirm infected cases. Thus, it is necessary to explore for other efficient alternatives that can address the drawbacks of RT-PCR. Existing studies show that the feasibility of using clinical imaging techniques such as chest X-ray (CXR) images for early diagnosis of COVID-19 with promising accuracy level[2]. From Fig 1, notable opacities can be observed from the CXR images of COVID-19 patients due to the SARS-COV-2 virus induced pneumonia on their lungs. These opacities are considered as the important features used by deep learning or machine learning based computer-aided diagnosis (CAD) systems to perform rapid diagnosis of COVID-19 based on the CXR images of patients.

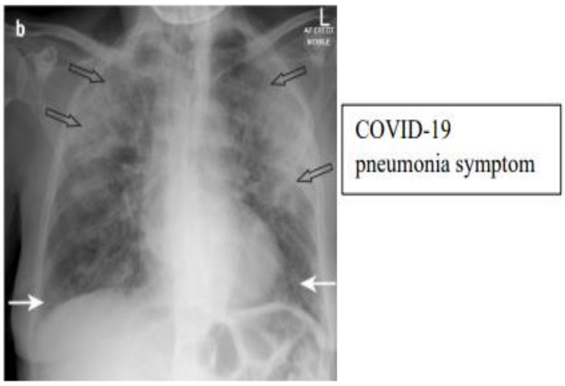

Figure 1: Opacities observed from the CXR images of COVID-19 patients

Convolutional neural network (CNN) is widely used to solve various real-world applications [2], [3], [4], [5], [6],[7] given its excellent capability to extract meaningful information from input sources and learn the nonlinear relationships between the input and expected outputs. The fundamental architecture of CNN consists of feature extractor (i.e., convolutional layers and pooling layers) and classifier (fully-connected layers). The breakthrough of deep learning researches motivated the development of various CNN architectures such as AlexNet[8], GoogleNet[9], Inception-v3[10], MobileNet-v2[11], ResNet-18[12], ResNet-50[12], ShuffleNet[13], SqueezeNet[14] and ResNet-50 RCNN[15].

While numerous works of COVID-19 diagnosis with deep learning technology were reported[16], the network performances of many existing architectures remain unexplored. This scenario opens an opportunity to further enhance the accuracy of COVID-19 diagnosis through the optimal selection of CNN architectures. Therefore, a comprehensive study is conducted in this paper to investigate the performance of previously mentioned network architectures in COVID-19 classification. The best CNN architecture is determined objectively via a set of statistical metrices and then implemented into an automated CAD system designed to identify COVID-19 diseases, which could be of great help to radiologist and other medical officers. Similar approach can also be used to design the automated CAD systems for other medical and non-medical applications.

2. Materials and Methods

2.1. *Datasets*

The CXR images used to train the abovementioned CNN architectures are obtained from a public database[17] contains 2700 images that can be equally divided into 3 classes, namely Normal, Pneumonia and COVID-19 cases. These images datasets are randomly divided into a ratio of 70:30 for training and testing sets, respectively. Before training the selected architecture with transfer learning for COVID-19 diagnosis, these input images are resized based on the requirements of respective network architecture as mentioned in Table 1. These input datasets are also pre-processed using data augmentation and converted into color images before the model training.

2.2. *Transfer Learning of Pretrained Networks*

It is nontrivial to train the CNN networks from scratch for specific tasks because it involves tremendous amount of resources (e.g., training time, input datasets and etc.). Transfer learning is a promising solution used to tackle these drawbacks due to its ability to transfer the knowledge from one or more domains and apply the knowledge to another domain with a different target task. During transfer learning process, learnable parameters of these pretrained network architectures (i.e., AlexNet, GoogleNet, Inception-v3, MobileNet-v2, ResNet-18, ResNet-50, ShuffleNet, SqueezeNet and ResNet-50 RCNN) are extracted and applied on the same type of network for different purposes. The original output layers of these pretrained networks are also replaced with the new output layers containing three classes (i.e., COVID-19, pneumonia and normal cases) and trained with new datasets in Section 2.1 for COVID-19 diagnosis.

Table 1. Image resolution used by each architecture

CNN	Resolution
AlexNet	227×227×3
GoogleNet	224×224×3
Inception-v3	299×299×3
MobileNet-v2	224×224×3
ResNet-18	224×224×3
ResNet-50	224×224×3
ShuffleNet	224×224×3
SqueezeNet	227×227×3
ResNet-50 RCNN	224×224×3

2.3. *Hyperparameters Settings*

Classification performances of pretrained networks used for COVID-19 can be governed by the hyperparameter settings of transfer learning process. Stochastic gradient descent (SGD) is used to train the selected architectures by minimizing the cross-entropy loss function. Five hyperparameters known as Momentum (MOM), Initial Learning Rate (ILR), Max Epoch (ME), Mini Batch Size (MBS), L2Regularization (L2) and Validation Frequency (VF) are considered. The best hyperparameter settings of each CNN architecture are determined based on trial and error and their values are presented in Table 2.

2.4. *Evaluation*

The performances of CNN architectures in COVID-19 diagnosis are evaluated using five metrics known as Accuracy, Sensitivity or Recall, Specificity, Precision and F1 score. These five metrics can be computed based on true positive (TP), true negative (TN), false positive (FP) and false negative (FN) results obtained from the confusion matrices of respective CNN architectures during the testing stage. Mathematical formulations of all performance metric are defined as follows.

Table 2: Hyperparameters used for all CNN architectures

CNN	MOM	ILR	ME	L2	MBS	VF
AlexNet	0.2	0.002	5	5e-4	32	35
GoogleNet	0.9	0.04	5	5e-4	32	35
Inception-v3	0.9	0.04	5	5e-4	32	35
MobileNet-v2	0.9	0.04	5	25e-4	32	35
ResNet-18	0.9	0.04	5	5e-4	32	35
ResNet-50	0.9	0.04	5	5e-4	32	35
ShuffleNet	0.9	0.04	5	25e-4	32	35
SqueezeNet	0.2	0.003	5	5e-4	32	35
ResNet-50 RCNN	0.9	0.04	5	5e-4	32	35

Accuracy represents the total numbers of correct predictions made by the CNN network, i.e.,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

The probability of a CNN network to obtain true positive results is quantified as sensitivity or recall, where

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

In contrary to sensitivity or recall, specificity refers to the probability of a CNN network in obtaining true negative results, i.e.,

$$Specificity = \frac{TN}{TP + FN} \tag{3}$$

Precision metric is used to evaluate the accuracy of positive results obtained by a CNN network out of all positive predictions made as defined below:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Finally, the F1 score represents the harmonic score evaluation of precision and sensitivity obtained by a CNN network, where

$$F1\ Score = 2 \times \left(\frac{Sensitivity \times Precision}{Sensitivity + Precision}\right) \tag{5}$$

Apart from the quantitative performance metrics, the performance of CNN networks in COVID-19 diagnosis is also evaluated qualitatively using the confusion matrix and receiver operating characteristic (ROC) curve.

3. Performance Evaluations of CNN Networks

The performances of all CNN networks (i.e., AlexNet, GoogleNet, Inception-v3, MobileNet-v2, ResNet-18, ResNet-50, ShuffleNet, SqueezeNet and ResNet-50 RCNN) to diagnose COVID-19, pneumonia and normal cases based on CXR images of patients are evaluated. All pretrained networks are trained with 70% of training datasets and tested with 30% of testing datasets. During the transfer learning process, SDG is used to adjust the learnable parameters of CNN models based on the hyperparameter settings reported in Table 2.

Fig 2 illustrates the confusion matrices used to qualitatively compare the classification performances all CNN architectures after completing their training and testing processes. The effectiveness of all selected CNN network architectures in COVID-19 diagnosis are also qualitatively analyzed based on ROC curves presented in Fig 3. Note that the ROC curve of an CNN network is constructed based on the corresponding TP and FP values obtained when solving the COVID-19 image datasets. Accordingly, ResNet-50 RCNN is observed to dominate most pretrained networks (i.e., GoogleNet, MobileNet-v2, ResNet-50 and ShuffleNet) by producing more correct classifications on the COVID-19, pneumonia and normal cases. Although AlexNet, ResNet-18 and SqueezeNet can perform slightly better than ResNet-50 RCNN by producing relatively more correct classification on pneumonia case, the latter network architecture can produce significantly better results when classifying the remaining two cases (i.e., COVID-19 and normal). Similar observations are found when comparing Inception-v3 and ResNet-50 RCNN, where the former network has slightly better result on COVID-19 case but the latter architecture has more competitive classification performance on the pneumonia and normal cases. Based on the qualitative analysis, ResNet-50 RCNN is found to be most reliably used for COVID-19 diagnosis.

Five performance metrics explained in Eqs. (1) to (5) are also used to quantitatively compare the performances of all pretrained networks in COVID-19 diagnosis as shown in Table 3, where the best results are highlighted with bold font. Accordingly, ResNet-50 RCNN can produce the best results in terms of accuracy, sensitivity, specificity, precision and F1 score when classifying the COVID-19, pneumonia and normal cases from CXR images. The quantitative results reported in Table 3 are consistent with the qualitative ones as shown in Fig 2 and Fig 3. ResNet-50 is another competitive classifier used to solve COVID-19 datasets by producing second best results for all performance metrics. Meanwhile, both of AlexNet and ShuffleNet are the worst-performing CNN models in COVID-19 diagnosis.

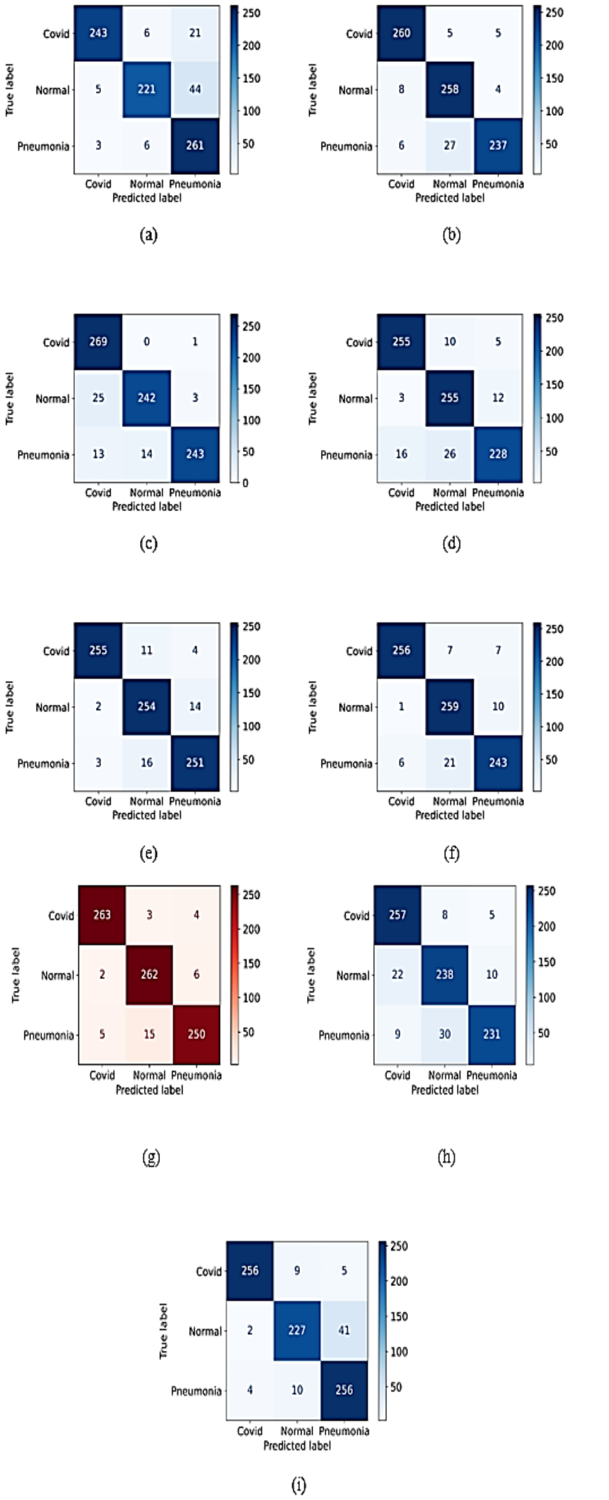

Figure 2: Confusion matrices of: (a) Alexnet, (b) Googlenet, (c) Inception-v3, (d) Mobilenet-v2, (e) Resnet-18, (f) Resnet-50, (g) Restnet-50 R-CNN, (h) Shufflenet and (i) Squeezenet in COVID-19 diagnosis using CXR images.

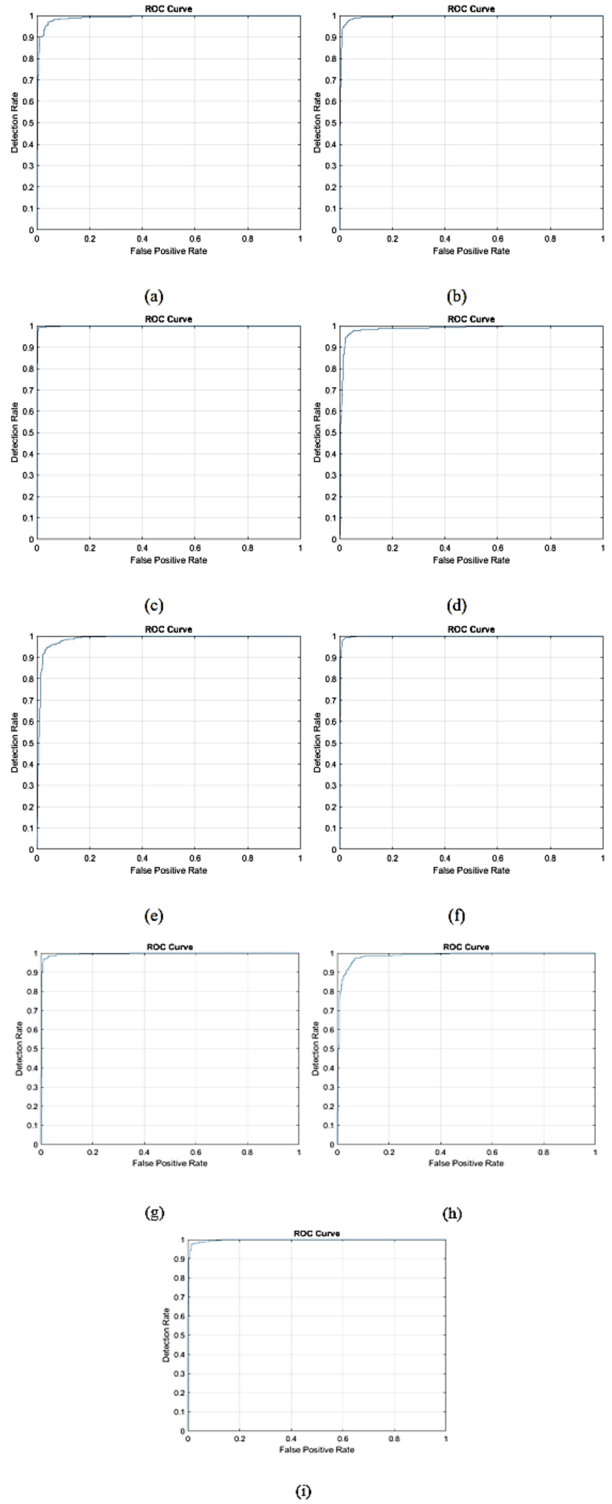

Figure 3: ROC curves of: (a) Alexnet, (b) Googlenet, (c) Inception-v3, (d) Mobilenet-v2, (e) Resnet-18, (f) Resnet-50, (g) Restnet-50 R-CNN, (h) Shufflenet and (i) Squeezenet in COVID-19 diagnosis using CXR images.

Table 3: Quantitative Performance Comparison Results

CNN	Acc.	Prec.	Sens.	Spec.	F1
Alexnet	89.50	90.50	89.50	94.75	89.58
Googlenet	93.21	93.39	93.21	96.60	93.19
Inception-v3	93.09	93.51	93.08	96.54	93.08
Mobilenet-v2	91.11	91.25	91.11	95.56	91.00
Resnet-18	91.73	92.10	91.73	95.86	91.75
Resnet-50	93.58	93.68	93.58	96.70	93.59
Shufflenet	89.63	89.79	89.63	94.81	89.61
Squeezenet	91.24	91.58	91.23	95.61	91.25
Resnet-50 RCNN	**95.67**	**95.71**	**95.67**	**97.84**	**95.67**

4. Conclusion

This paper proposes a deep learning-based CAD system by combining pretrained CNN models and transfer learning for the classifications of COVID-19, pneumonia and normal cases based on CXR images of patients. This current study aims to perform comprehensive analyses on the performances of AlexNet, GoogleNet, Inception-v3, MobileNet-v2, ResNet-18, ResNet-50, ShuffleNet, SqueezeNet and ResNet-50 RCNN to tackle COVID-19 diagnosis task based on different performance metrices. Transfer learning employed during the training process enables fast convergence of network without required long training time and large numbers of datasets. Simulation studies show that ResNet-50 RCNN achieves the best results with in terms of accuracy of 95.67%, precision of 95.71%, sensitivity of 95.67%, specificity of 97.84% and F1-score of 95.67%.

Acknowledgment

This research is supported by Ministry of Higher Education Malaysia (MOHE) under the Fundamental Research Grant Scheme with the project codes of FRGS/1/2019/TK04/UCSI/02/1 and FRGS/1/2020/TK0/UCSI/02/4. This research is also supported by UCSI Research Excellence & Innovation Grant (REIG) with project code of REIG-FETBE-2022/038.

References

1. COVID Live – Coronavirus Statistics Worldometer. https://www.worldometers.info/coronavirus/?utm_campaignomeAdvegas1?. Accessed 6 Jan 2023.
2. Berghout, T., Benbouzid, M., Muyeen, S. M., Bentrcia, T., Mouss L. Auto-NAHL: A Neural Network Approach for Condition-Based Maintenance of Complex Industrial Systems. *IEEE Access* **9**, 152829 – 152840 (2021).
3. Berghout, T., Benbouzid, M., Muyeen, S. M. Machine learning for cybersecurity in smart grids: A comprehensive review-based study on methods, solutions, and prospects. International Journal of Critical Infrastructure Protection, 38, 100547 (2022).
4. Alrifaey, M., Lim, W. H., Ang, C. K. A Novel Deep Learning Framework Based RNN-SAE for Fault Detection of Electrical Gas Generator. 21433-21442 (2021).
5. Alrifaey, M. *et al.* Hybrid Deep Learning Model for Fault Detection and Classification of Grid-Connected Photovoltaic System. *IEEE Access* **10**, 13852-13869 (2022).
6. Jdid, B., Hassan, K., Dayoub, I., Lim, W. H. & Mokayef, M. Machine Learning Based Automatic Modulation Recognition for Wireless Communications: A Comprehensive Survey. *IEEE Access* **9**, 57851-57873 (2021).
7. Jdid, B., Lim, W. H., Dayoub, I., Hassan, K. & Juhari, M. R. B. M. Robust Automatic Modulation Recognition Through Joint Contribution of Hand-Crafted and Contextual Features. *IEEE Access* **9**, 104530-104546 (2021).
8. Krizhevsky, A., Sutskever, I., Hinton., G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in neural information processing systems* (2012).
9. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1-9 (2015).
10. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2818-2826 (2016).
11. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L.C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510-4520, (2018).
12. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778 (2016).
13. Zhang, X., Zhou, X., Lin, M., Sun, J., ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In *2018 IEEE/CVF*

Conference on Computer Vision and Pattern Recognition, 6848-6856 (2018).
14. Iandola, F. N., Song, H., Matthew W. M., Khalid, A., William, J. D., Kurt, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. Preprint, submitted November 4 (2016).
15. Renjun, X., Junliang. Y., Yi, W., MengCheng, S. Fault Detection Method Based on Improved Faster R-CNN: Take ResNet-50 as an Example. *Geofluids*. (2022).
16. Siddiqui, S. et. al. Deep Learning Models for the Diagnosis and Screening of COVID-19: A Systematic Review. *SN Computer Science 3*, 1303 (2022).
17. Goel, T., Murugan, R., Mirjalili, S., Chakrabartty, D. K., OptCoNet: an optimized convolutional neural network for an automatic diagnosis of COVID-19. *Applied Intelligence* 51, 1351-1366 (2021).

Authors Introduction

Mr. Suhaim Parvez Wadekar

He received the B.Eng. degree in Mechatronic Engineering with Honours from UCSI University, Malaysia, in 2022. His research interests are swarm intelligence, machine learning and deep learning, and medical diagnosis system.

Mr. Koon Meng Ang

He received the B.Eng. degree in Mechatronic Engineering with Honours from UCSI University, Malaysia, in 2019. He is currently pursuing Ph.D. degree in UCSI University, Malaysia. His research interests are swarm intelligence, machine learning and deep learning.

Dr. Chin Hong Wong

He is a Lecturer in Maynooth International Engineering College at Fuzhou University in China. He received his PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2017. His research interests are Energy harvesting and control system.

Dr. Abhishek Sharma

He is a Research Assistant Professor at Graphic Era Deemed to be University in India. He received his PhD from University of Petroluem & Energy Studies in 2022. His research interests are artificial intellgence and power electronics.

Dr. Tiong Hoo Lim

He is a Senior Assistant Professor and Director of Planning Development Office at Universiti Teknologi Brunei in Brunie Darussalam. He received his PhD in Computer Science from University of York, United Kingdom. His research interests are artificial intelligence, advanced technology, competitive and forecasting analysis.

Dr. Chun Kit Ang

He is the Dean and Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Mechanical and Manufacturing Engineering from Universiti Putra Malaysia in 2014. His research interests are artificial intelligence, soft computing, robotics and mechatronics.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Wei Hong Lim

He is an Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.