

Facemask Detection Algorithm on COVID Community Spread Control using EfficientNet Algorithm

Vivekanadam Balasubramaniam,

Faculty of Computer Science and Multimedia,
Lincoln University College,
Kota Bharu, Malaysia.
Email: vivekanandam@lincoln.edu.my

Abstract: Facemask has become mandatory in all COVID-infected communities present across the world. However, in real-life situations, checking the facemask code on each individual has become a difficult task. On the other hand, Automation systems are playing a widespread role in human community to automate different applications. As a result, it necessitates the need to develop a dependable automated method to monitor the facemask code to benefit humans. Recently, deep learning algorithms are emerging as a fast growing application, which has been developed for performing huge number of analysis and detection process. Henceforth, this paper proposes a deep learning based facemask detection process for automating the human effort involved in monitoring process. This work utilizes an openly available facemask detection dataset with 7553 images for the training and verification process, which is based on CNN driven EfficientNet architecture with an accuracy of about 97.12%.

Keywords: Face detection, facemask detection, COVID 19, CNN, EfficientNet.

1. Introduction

The community spread of COVID can be controlled in two ways, the first and simple way without any cost investment is social distancing. The second and efficient way is coverage of nose and mouth with a facemask [1]. Though, due to discomfort of wearing masks, the people are not using it properly. So, manual governance has been applied on some places for ensuring the appropriate utilization of facemasks. It is always difficult to monitor them manually and most

of the time the efficiency of monitoring process may get reduced due to general human errors [2].

The automation systems can deal this problem by reducing the manual human errors. Moreover, the automation system requires certain sensors, controllers and output devices. The recent automation systems are also equipped with deep learning based algorithms on their microcontroller architecture to deliver a better performance improvement [3]. The only problem in equipping a microcontroller with deep learning model is training process. The training process of deep learning algorithm requires more number of training data. The efficiency of a deep learning algorithm may vary with respect to the number of images included in the learning phase [4].

The deep learning algorithms are widely used for performing prediction, estimation and classification tasks. Here, the prediction and estimation requires huge number of attributes to satisfy the application requirement [5, 6]. The number of attributes can't be same for all the applications. Selection of attributes for some particular application plays a major role in the similar estimation and prediction processes. Several soft computing techniques are mostly employed in such cases for estimating the optimum attributes to train the deep learning model.

In the case of classification process, the basic requirement is the number of classes, which have to be identified from the classification process. Therefore, the training process must include the data samples on each class with almost equivalent numbers among them. The imbalance samples of training data are manageable by implementing the augmentation kind of process. Collection of enough dataset for several studies is not a huge burden nowadays. Lot of dataset websites is widely open for public to download the datasets without any costs for the research purpose [7, 8].

Another challenge in employing deep learning model is the identification of right neural network architecture for a specific application. The performance between the deep learning models differs with respect to the nature of a deep learning model. Therefore, several trials may initiate the development of a particular deep learning filter for the specific application. Figure 1 explores some of the most widely used deep learning models for big data analytics. The

performance of deep learning models can also be improved or reduced with the adoption of a preprocessing step [9]. The preprocessing steps are widely used for implementing the noise removal process in medical images and satellite images. However, the preprocessing step may results in missing some useful information from the training images. Therefore, the efficiency may get reduced on some applications [10]. Apart from noise removal, the preprocessing steps are widely used for marking out the region of interest in the sample image.

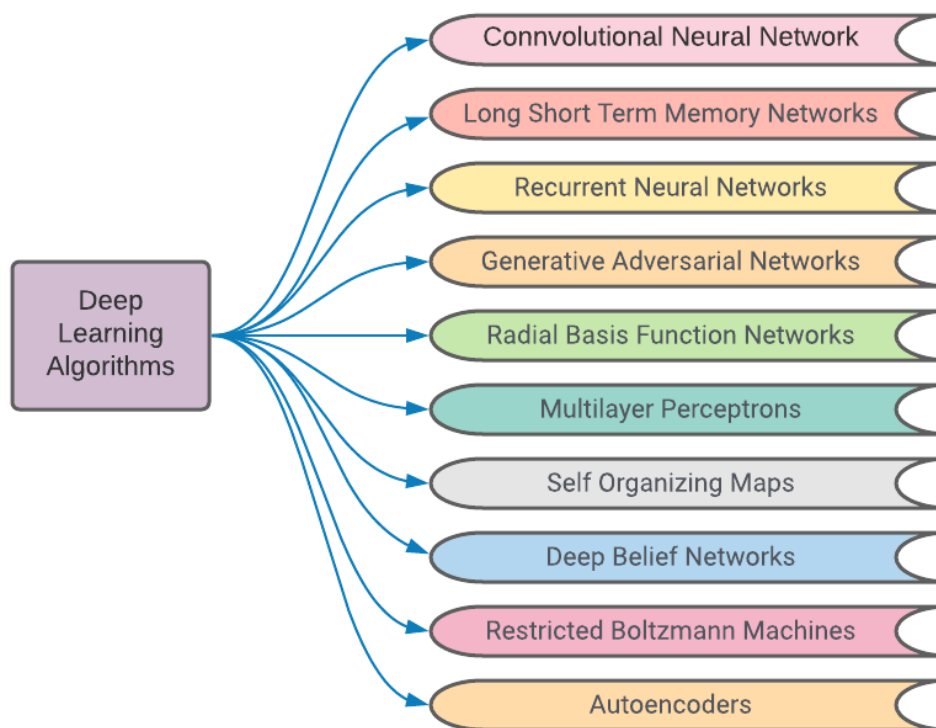


Figure 1. Types of Deep Learning Models

2. Literature Work

A hybrid deep learning algorithm was proposed based on support vector machine architecture to perform facemask classification process and further the feature extraction process was employed with ResNet50 architecture [11]. The performance of the such hybrid model is verified with three different datasets called real world masked face dataset, simulated masked face dataset and labeled faces in the wild. The accuracy attainment on the three datasets is

99.64%, 99.49% and 100% respectively. Similarly, a hybrid learning technique based on hidden markov model and support vector machines was developed to perform speech recognition process [12]. The performance of the developed hybrid model is verified with a regular hidden markov model with the same dataset for proving its efficiency. The test method includes word and phonetic units, where in both models the performances of hybrid model gives better accuracy over the regular model. The performance attainments on word unit with hybrid markov model are 76% and 82% for phonetic unit. The performance of hybrid model on both units is 88% and 94% for the hybrid model.

Again a facemask detection algorithm was employed with ResNet50 architecture for feature extraction process [13] and the model is equipped with YOLO V2 model for the classification process. The experimental analysis of the proposed model was verified with two dataset namely medical mask datasets and face mask datasets. The images from the datasets are labelled and undergone with a data augmentation process for increasing the available sample counts from the datasets. The average precision achievement of the proposed model is 81%. The ResNet50 algorithm models are widely used for many applications. A deep net model based technique was presented to identify COVID infections from radiograph images using ResNet50 architecture [14]. The performances of ResNet50 was verified here with InceptionV3 and InceptionResNetV2, where the accuracy of ResNet50 reaches up to 98% on a Github dataset, which consists of various lung disorder X-ray images.

A deep neural network based facemask detection system was designed to perform the classification process on light weighted embedded devices like Raspberry pi and NVIDIA nano [15]. The system utilizes single shot multibox detector for face recognition and it equips MobileNetV2 model for the face mask classification process. The dataset utilized in the work has images collected from various open source datasets and the final dataset is uploaded to Github website for future verification purpose. The total number of images available in the final dataset contains 5521 samples of faces with and without facemasks. The accuracy achievement of this model is 92.64% with f1 score of 0.93. The mobileNet classifier model was also used for binary classification of detecting vacant parking slots as smart application [16]. The system was developed to detect the vacant place from the recorded video on the parking slots. The

MobileNet model is employed in this work because of its light weighted nature on operation. The video informations are collected from FOSCAM FI9800P protocol of about 4 weeks data recorded with time frame of 24/7. The video covers almost 50 parking slots in that particular location. The experimental results explores that the process speed has been satisfied with the MobileNet algorithm.

A detection algorithm was made using R-CNN for identifying the effective usage or wearing model of facemasks by the users [17]. An experiment was conducted in the work with a self-made dataset of about 8635 images on various style of proper and improper usage of facemasks. The workflow is presented with multiple context feature extraction process with different spatial levels. The mean average precision attained by the proposed model on this work is 84.1%. A comparative analysis was made between the deep learning models like CNN, RNN and FNN on tumor investigation process [18]. The performance was analyzed using Wisconsin original dataset and taken 7564 images for the training process and 3456 images for the testing process. The evaluation indicates that, the CNN was performing better than the other two neural network architectures on the same dataset. The outcome of the RNN gets satisfied over the FNN and it gives 25% of reduced prediction over the CNN.

An experimental work based on R-CNN and YOLO V3 was made with a facemask detection dataset namely widerface of about 7500 images approximately [19]. The system marks up the human faces with a green and red color bounding boxes for specifying the people wear masks and avoiding masks. However, the average precision of R-CNN is 62%, where for YOLO V3 the precision value is 55%. Among the two neural network architecture, YOLO V3 leverages better inference time over the R-CNN. A smart security system was developed to identify the suspicious activity on secured areas [20]. The system is enabled with a PIR sensor for detecting the human activity in a particular area. If a movement is identified on a particular area, the camera connected over the place will capture the human faces for recognition process. The algorithm based on KNN is employed in this work for implementing face recognition process on authorized and unauthorized people. The system was designed to give alertness through an alarm system if it identifies an unauthorized movement. The accuracy on face recognition achieves 98.45% with the KNN classifier.

A deep learning based detection technique was developed to identify facemask presence on a human face using deep residual learning model equipped with feature pyramid network for human body identification. The system is extended with a multitask CNN for identifying the human faces. Finally, a regular CNN classifier is employed for categorizing the human faces with mask and without mask [21]. The accuracy attainment of the proposed model is 81.3% with a recall rate of 99.2%. An image classification algorithm based on CNN and SVM was designed to present the abnormal incidents from a video surveillance [22]. The experimental work of the proposed hybrid model reaches an accuracy of 90% over 69% and 72% than the usual CNN and SVM. A self-made dataset was developed to verify the accuracy attainment of YOLO model on facemask identification process. The proposed work is verified with four numbers of tiny YOLO architecture and four numbers of regular YOLO architecture [23]. The experimental analysis projects the tiny YOLO V3 model, which performs better in the verified self-made dataset.

3. Proposed Methodology

3.1 Dataset description

The dataset utilized in the work is taken from [24] kaggle website and it contains 7553 numbers of RGB images. The images are originally collected from github website of about 1776 count and the balance images are collected through web search engine. Therefore the total images available in the utilized dataset are 3725 faces covered with masks and 3828 faces covered without masks. From both the samples, the images are taken out randomly as 70% for the training process and 30% for the verification process. So that the images count for training is 2793 from the mask faces and 2679 from fresh faces. Similarly, 932 images, where taken to verify faces with masks and 1149 samples without masks.

3.2 EfficientNet

EfficientNet is a type of convolution neural network, which has been widely used for performing image classification process. The architecture of EfficientNet varies from B0 to B7. However, the EfficientNet with B0 architecture leverages better accuracy when compared to the other CNN based pre-trained models [25] in terms of accuracy and required feature attributes.

The model is actually suitable for images with lesser resolution because the depth wise scaling operation of the EfficientNet is improved when compared to others. The floating point operations are also improved in terms of architecture for obtaining better accuracy through a multi-objective architecture network.

An uneven group of residual blocks with count of seven is employed in the EfficientNet architecture based on squeeze and excitation principle. A swish activation process is also merged in the squeeze principle for addressing the operation of zeroing the negative values presented in the image during the classification training process. Swish is a combination of linear activation with sigmoid operations. The squeeze and excitation process is taken into account for avoiding equal weightages among the image pixel information. Therefore, the thinking and analyzing capacity of the neural network architecture is improved due to the reading variable information.

3.3 VGG 19

VGG 19 is also a type of convolutional neural network architecture, which consists of 19 deep layers. The preprocessing flow of the VGG 19 architecture equalizes the pixel values between 0 and 255 to develop a new value of image by analyzing the obtained adjusted value with the mean image values of the given input image for training. The VGG 19 has been made up of 16 layers of convolution process and 3 layers of fully connected operations. The concept of pooling layer is also utilized in the architecture same as that of CNN.

3.4 Inceptionv3

The inception networks are developed to improvise the computational efficiency of the convolutional neural network. The issues of over fitting are also rectified in the inception model by implementing the convolutional layer of 7x7. To avoid the loss of information, further layers in the Inceptionv3 are going deeper into the architecture. A batch normalization process is employed as an auxiliary network of fully connected layers. Therefore, the Inceptionv3 reduces the error rate to extend over its previous models.

3.5 ResNet50

ResNet50 is also called as residual network, which is developed from convolutional neural network with deep network connections of 50 layers. The information from the training samples are forwarded through residual blocks with 5 stages. The residual blocks are employed with 3 layers for convolution module. An identity connection is made in the ResNet50 model by connecting the upcoming layers through serial and hop connection.

4. Experimental Setup

The workflow followed in the experimental analysis among the classification algorithms discussed in the previous section is projected in the figure 2. The images from the dataset are preprocessed at the beginning before splitting the samples for the training and testing process. Because the images available in the dataset are not in a uniform size, therefore the image samples are resized to 224x224 sizes which are suitable for most of the image classifiers. The images are further taken into consideration for training process by randomly considering the samples from each class. First, the training images are forwarded to the classifier network through supervised learning approach. Therefore, the training accuracy of the analysis model can be improved when compared to the unsupervised learning process. After the learning process, the testing images are forwarded to the trained model in a queue for determining its classes. From the predicted output, the confusion matrix chart is obtained by the algorithm for evaluating the accuracy, precision, recall and f1 score of the verified models. All the classifiers utilized in the work are pre-trained models, which are utilized from Google Colab platform.

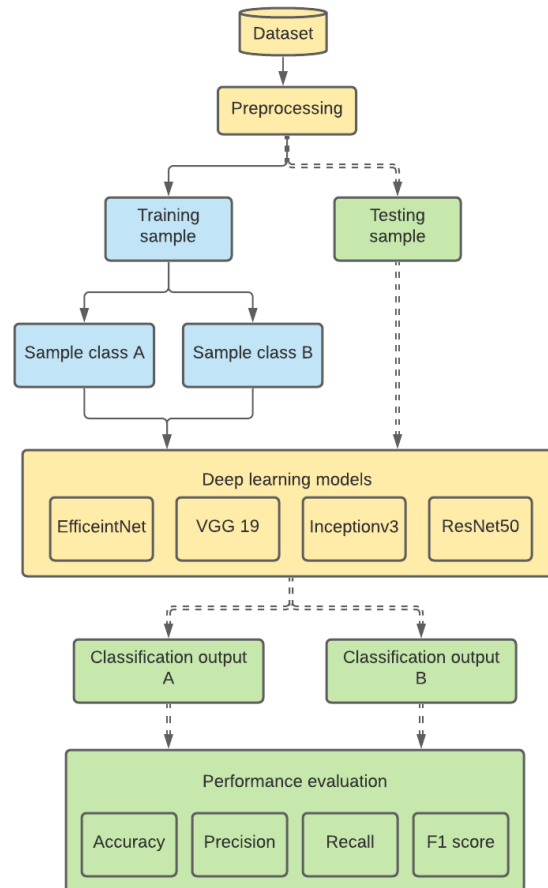


Figure 2. Workflow of the Proposed Model

The performance analysis is presented in table 1 and figure 3. The experimental analysis explores that, the EfficientNet algorithm performs better in terms of precision. It means that, the negative predictions are completely lesser in the EfficientNet algorithm over the other models. The accuracy of the EfficientNet is also getting improved due to such better precision attainments. The performance of ResNet50 comes up next to the EfficientNet because of its deep network connections. Similarly, the performance of VGG19 is comparatively reduced due to its reduced network layers over the ResNet50 algorithm. The Inceptionv3 model is employed with a novel information gathering network algorithm comes almost equal to the ResNet50 in terms of accuracy and recall. However, the performance of Inceptionv3 may go beyond ResNet50 when there is a change in testing and training samples. The major thing which has to be specified in the

work is about the training time. The training time of the EfficientNet is comparatively better because of its efficient training accuracy achievement with lesser epochs.

Table 1. Performance Analysis among the Classifiers

Algorithms	EfficientNet	VGG 19	Inceptionv3	ResNet50
Accuracy	97.12	92.74	95.19	95.87
Precision	99.10	95.40	96.74	97.85
Recall	94.42	89.06	92.38	92.81
F1 Score	96.70	92.12	94.51	95.26

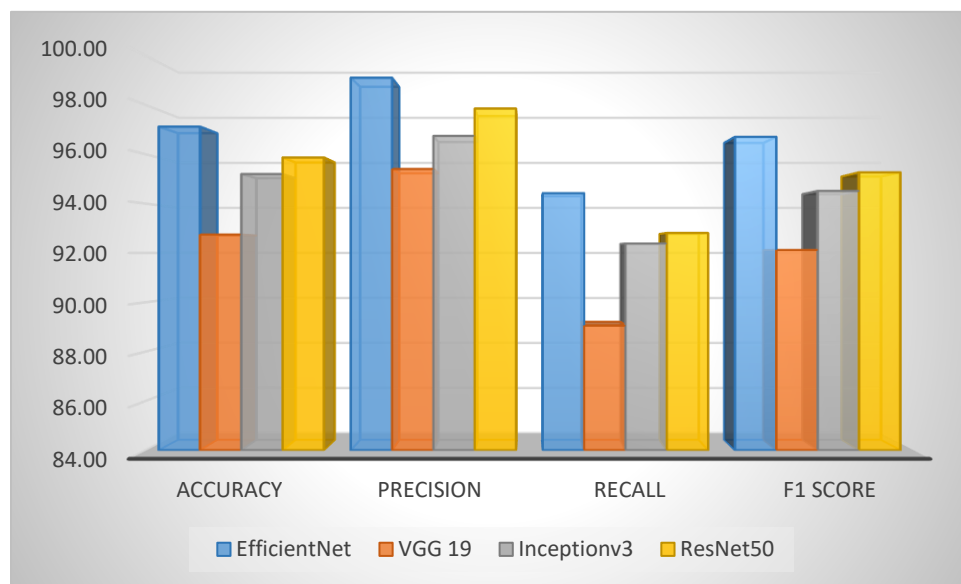


Figure 3. Comparison chart on the performances of classifier models

Conclusion

This paper suggests a deep learning algorithm called EfficientNet for a face mask detection process due to its better precision and accuracy rate. The proposed algorithm can be applied for real-time operations with Raspberry pi and camera peripherals. The efficiency of the system can be improved when it is trained with huge number of data samples. For obtaining

better observation on real-time applications, the camera has to be kept at a good lighting place for efficiently observing the information available in the image.

References

- [1] Cheng, Vincent Chi-Chung, Shuk-Ching Wong, Vivien Wai-Man Chuang, Simon Yung-Chun So, Jonathan Hon-Kwan Chen, Siddharth Sridhar, Kelvin Kai-Wang To et al. "The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2." *Journal of Infection* 81, no. 1 (2020): 107-114.
- [2] Adam, Edriss Eisa Babikir. "Evaluation of Fingerprint Liveness Detection by Machine Learning Approach-A Systematic View." *Journal of ISMAC* 3, no. 01 (2021): 16-30.
- [3] Senthilkumar, D., C. Akshayaa, and D. George Washington. "Efficient Deep Learning Approach for Multi-label Semantic Scene Classification." In *International Conference on Image Processing and Capsule Networks*, pp. 397-410. Springer, Cham, 2020.
- [4] Hariharakrishnan, Jayaram, and N. Bhalaji. "Adaptability Analysis of 6LoWPAN and RPL for Healthcare applications of Internet-of-Things." *Journal of ISMAC* 3, no. 02 (2021): 69-81.
- [5] Niranjana, D. K., and N. Rakesh. "Smart Surveillance System by Face Recognition and Tracking Using Machine Learning Techniques." In *Computational Vision and Bio-Inspired Computing*, pp. 1-14. Springer, Singapore, 2021.
- [6] Adam, Edriss Eisa Babikir. "Survey on Medical Imaging of Electrical Impedance Tomography (EIT) by Variable Current Pattern Methods." *Journal of ISMAC* 3, no. 02 (2021): 82-95.
- [7] Kakati, Munindra, and Parismita Sarma. "Human Pose Detection: A Machine Learning Approach." In *International Conference On Computational Vision and Bio Inspired Computing*, pp. 8-18. Springer, Cham, 2019..
- [8] Vijayakumar, T., Mr R. Vinothkanna, and M. Duraipandian. "Fusion based Feature Extraction Analysis of ECG Signal Interpretation—A Systematic Approach." *Journal of Artificial Intelligence* 3, no. 01 (2021): 1-16.

- [9] Huang, Qiubo, and Chun Ji. "Face Detection Based on Image Stitching for Class Attendance Checking." In International Conference on Image Processing and Capsule Networks, pp. 31-43. Springer, Cham, 2020.
- [10] Ancy, C. A., and Maya L. Pai. "Brain Tumour Three-Class Classification on MRI Scans Using Transfer Learning and Data Augmentation." In Computational Vision and Bio-Inspired Computing, pp. 41-56. Springer, Singapore, 2021.
- [11] Loey, Mohamed, Gunasekaran Manogaran, Mohamed Hamed N. Taha, and Nour Eldeen M. Khalifa. "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic." *Measurement* 167 (2021): 108288.
- [12] Manoharan, Samuel, and Narain Ponraj. "Analysis of Complex Non-Linear Environment Exploration in Speech Recognition by Hybrid Learning Technique." *Journal of Innovative Image Processing (JIIP)* 2, no. 04 (2020): 202-209.
- [13] Loey, Mohamed, Gunasekaran Manogaran, Mohamed Hamed N. Taha, and Nour Eldeen M. Khalifa. "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection." *Sustainable cities and society* 65 (2021): 102600.
- [14] Dhaya, R. "Deep net model for detection of covid-19 using radiographs based on roc analysis." *Journal of Innovative Image Processing (JIIP)* 2, no. 03 (2020): 135-140.
- [15] Nagrath, Preeti, Rachna Jain, Agam Madan, Rohan Arora, Piyush Kataria, and Jude Hemanth. "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2." *Sustainable cities and society* 66 (2021): 102692.
- [16] Kamel, Khaled, S. Smys, and Abul Bashar. "Tenancy Status Identification of Parking Slots Using Mobile Net Binary Classifier." *Journal of Artificial Intelligence* 2, no. 03 (2020): 146-154.
- [17] Zhang, Jun, Feiteng Han, Yutong Chun, and Wang Chen. "A Novel Detection Framework About Conditions of Wearing Face Mask for Helping Control the Spread of COVID-19." *IEEE Access* 9 (2021): 42975-42984.

- [18] Vijayakumar, T. "Neural network analysis for tumor investigation and cancer prediction." *Journal of Electronics* 1, no. 02 (2019): 89-98.
- [19] Singh, Sunil, Umang Ahuja, Munish Kumar, Krishan Kumar, and Monika Sachdeva. "Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment." *Multimedia Tools and Applications* (2021): 1-16.
- [20] Chen, Joy Iong Zong. "Smart Security System for Suspicious Activity Detection in Volatile Areas." *Journal of Information Technology* 2, no. 01 (2020): 64-72.
- [21] Snyder, Shay E., and Ghaith Husari. "Thor: A Deep Learning Approach for Face Mask Detection to Prevent the COVID-19 Pandemic." In *SoutheastCon 2021*, pp. 1-8. IEEE, 2021.
- [22] Ranganathan, G. "A Study to Find Facts Behind Preprocessing on Deep Learning Algorithms." *Journal of Innovative Image Processing (JIIP)* 3, no. 01 (2021): 66-74.
- [23] Kumar, Akhil, Arvind Kalia, Kinshuk Verma, Akashdeep Sharma, and Manisha Kaushal. "Scaling up face masks detection with YOLO on a novel dataset." *Optik* 239 (2021): 166744.
- [24] www.kaggle.com/omkargurav/face-mask-dataset
- [25] Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." In *International Conference on Machine Learning*, pp. 6105-6114. PMLR, 2019.

Authors Biography

Vivekanadam Balasubramaniam is a professor in the Department of Computer Science and Multimedia at Lincoln University College, in Malaysia. His major area of research are machine learning, neural network algorithms, image processing, video and signal processing, cloud computing, deep learning, artificial intelligence, object recognition, complex feature extraction and vision graphics