2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC 2020), 8 August 2020, Shah Alam, Malaysia

Real-Time Facemask Recognition with Alarm System using Deep Learning

Sammy V. Militante
College of Engineering & Architecture
University of Antique
Sibalom, Antique, Philippines
https://orcid.org/0000-0003-1962-9450

Nanette V. Dionisio
College of Arts and Sciences
University of Antique
Sibalom, Antique, Philippines
nanette_dionisio@yahoo.com

***Abstract*— In the background of the COVID-19 pandemic, institutions such as the academy suffer a great deal from practically closed globally if the current situation is not going to change. COVID-19 also known as Serious Acute Respiratory Syndrome Corona virus-2 is an infectious disease that is released from an infected sick person who speaks, sneezes, or coughs by respiratory droplets. This spreads quickly through close contact with anyone infected, or by touching objects or surfaces affected with a virus. There's still currently no vaccine available to protect against COVID-19 and preventing exposure to the virus seems to be the only method to safeguard ourselves. Wearing a facemask that covers the nose and mouth in a public setting and often washing hands or the use of at least 70% alcohol-based sanitizers is one way to avoid being exposed to the virus. Amid the advancement of technology, Deep Learning has proven its effectiveness in recognition and classification through image processing. The research study uses deep learning techniques in distinguishing facial recognition and recognize if the person is wearing a facemask or not. The dataset collected contains 25,000 images using 224x224 pixel resolution and achieved an accuracy rate of 96% as to the performance of the trained model. The system develops a Raspberry Pi-based real-time facemask recognition that alarms and captures the facial image if the person detected is not wearing a facemask. This study is beneficial in combating the spread of the virus and avoiding contact with the virus.**

Keywords—Real-Time Facemask Recognition, Alarm System, Raspberry pi, Deep Learning

I. INTRODUCTION

Everyone has been affected by the COVID-19 coronavirus pandemic on a global scale. It crippled the economic growth of the entire nation around the world [1]. Coronavirus disease 2019 (COVID-19) is an emerging respiratory disease caused by severe acute respiratory syndrome coronavirus 2 or SARS-CoV2 [2]. As of June 10, 2020, the virus reached nearly eight million infected patients and half million died from the virus [3]. To combat the transmission of the virus [4], there are enforced protocols set by the World Health Organization (WHO) like compulsory wearing of face mask [5], observing strict social distancing in public places [6] and washing of hands or sanitizing hands with disinfectants frequently [5]. There are studies conducted that wearing facemask [7] is important to prevent the spread of the virus [8],[9],[10]. Research studies show the effectiveness of N95 and surgical masks in preventing virus transmission are 91% and 68% respectively [11]. Wearing these masks will effectively disrupt airborne viruses so that such infections can not reach a human being's respiratory system and it is an inexpensive way to mitigate fatalities and respiratory infection disorders. Nevertheless, the efficacy of facemasks in preventing disease transmission in the public has generally been lessened due to inadequate facemask use. It is essential to develop an automatic detection for wearing facemask which will provide individual protection and prevent the local epidemic.

Deep learning advancement with the integration of computer vision offers the breakthrough in development in countless fields of technology [12]. Deep neural networks (DNNs) as the main component of deep learning methods is everything it offers including object detection, image classification, and image segmentation [13],[14]. Convolutional neural networks (CNNs) is one of the principal models of DNN is generally used in computer vision tasks [15]. After training the model, CNNs can identify and classify facial images even with minor differences using their overwhelming feature extraction ability and store image pattern details.

In this research study, deep learning techniques are applied to construct a classifier that will collect images of a person wearing a face mask and not from the database and differentiate between these classes of facemask-wearing and not facemask-wearing [16]. The artificial neural network has been demonstrated to be a vigorous procedure for feature extraction from unprocessed data [17],[18]. This study proposes the use of a convolutional neural network to design the facemask classifier and to include the effect of the number of the convolutional neural layer on the prediction accuracy. This project is implemented in a Raspberry Pi using OpenCV TensorFlow and Python programming language.

Raspberry Pi (RPi) is designed as a Chip System (SoC) where the critical circuits such as the Central Processing Unit (CPU), the Graphics Processing Unit (GPU), input, and output are carried by a single circuit board. The GPIO pins provide an essential element to help enable the RPi to be accessible to hardware programming for controlling electronic circuits and data processing on input/output devices. Add a power adapter, keyboard, mouse, and monitor that works on the Raspberry Pi in compliance with the HDMI connector. New models are available to interact via WiFi to the internet. The RPi can be run using the Raspbian operating system. It has a pre-installed Python programming language.

We can summarize our main contributions as follows.

1) Create a system for classifying facial images using CNN.
2) Employ deep learning methods in the identification of wearing facemask conditions.

II. RELATED WORKS

Convolutional networks are also undoubtedly used for undertakings of the classification task. Standard architectures like AlexNet[19] and VGGNet [20] are packed or contain a load of the convolutional layer. AlexNet the winner of the ImageNet LSVRC-2012 competition comprises 5 convolution layers plus 3 fully connected layers while VGGNet is an improvement of AlexNet, progressively shares large kernels with several 3x3 kernels. New architectures like ResNet [21] implement an accelerated link on training accuracy which allows for much deeper networks to avoid overload processing. These architectures are often applied in image recognition frameworks for original feature-extraction. Our proposed study uses the architectural features of VGG-16 as the foundation network for face recognition and the Fully-Convolutional segmentation network. The VGG-16 network is fairly robust in extracting features and less costly in computational terms. Nevertheless, most segmentation architectures consecutively rely on input image downsampling and upsampling, Fully Convolutional Networks [22], [23], [24] are indeed the discreet and expressively comprehensive segmentation method.

Deep Learning is comprised of a very enormous number of neural networks that use the multiple cores of a processor of a computer and video processing cards to manage the neural network's neuron which is categorized as a single node [25], [26], [27]. Deep learning is used in numerous applications because of its popularity especially in the field of medicine and agriculture. Its application includes identification, detection, and recognition of diseases of both human-related, animals, and plants [17],[18], detection and grading of fruit images [25],[26], image capturing robots like face recognition through attendance system [27].

As for the convolution process portrayed in fig. 1, it begins through the extracted input image along with its features using a filter of 3x3 across with a stride of 1 as mentioned to be convolution (Conv). The resulting output from the Conv process produces a featured map through the dot product of the preceding Conv layer. Each featured map keeps precise details of the original image to establish a specific input and it will down-sampled through the ReLU method to go on other values intact and downgrade negative values to zero values. Additional down-sampling procedure following each Conv named as max-pooling decreases the values into half of its original value by simply selecting the max values only from the matrix of kernel [18]. Providing the primary clues in identifying a precise image for flexible handling of resources is the work of the pooling layer. The pooled-features is distributed and flattened in the fully connected layers (FC layers) that translates the activation from one or zero values. Then, the Softmax-activation function produces probabilities through its neural networks in classifying input data[17].

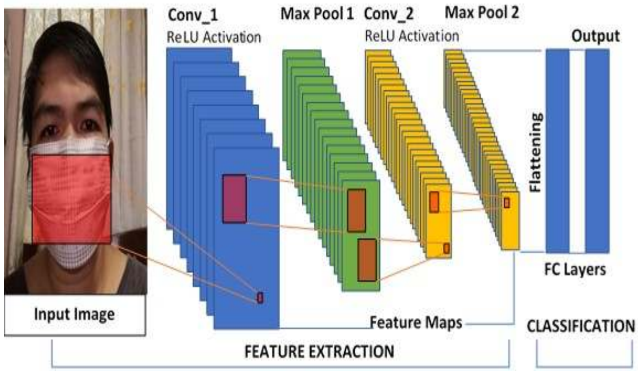

Fig. 1 Convolution Process

III. METHODOLOGY

The process of CNN is to identify and categorize images from learned features. It is very effective in a multi-layered structure when obtaining and assessing the necessary features of graphical images. The outline of the proposed method for the identification of facemasks is shown in fig. 2 to fig. 4. It describes the system proposed composed entirely of acquisitions of images as shown in fig. 2. Data collection consists of a person wearing a facemask and not wearing a facemask in fig. 3 and a CNN architecture classification in fig. 4.

Fig. 2. Image Acquisition for the proposed computer vision to detect facemask with a person wearing and not wearing a facemask

Fig. 3. Annotated Image data collection of the proposed computer vision to detect facemask with a person wearing and not wearing a facemask

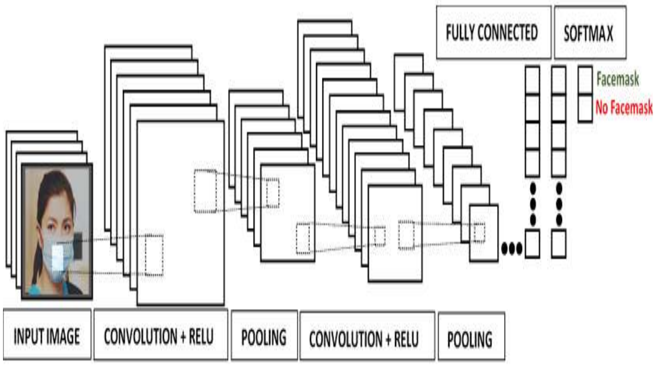

Fig. 4. Classification process using convolutional neural networks.

Artificial neural networks replicate the functioning of the human brain incorporating mathematical equations via neurons and linked networks. The ANN 's value is to be learned in the supervised learning process. CNN [28] is the development of mainstream artificial neural networks, typically based on applications with recurring trends in a variety of fields such as image recognition or imaging. ANNs key characteristic is that it reduces the number of neurons needed as opposed to conventional feedforward neural networks with the technique used in the layering.

Convolutional Neural Network (CNN) is a structured deep learning process that plays a groundbreaking push for a variety of applications focusing on computer vision and image-based applications [29]. The fields in which CNN is prevalently used are facial recognition, object recognition, classification of images [30],[31], etc. The components of the CNN model are shown in Fig. 5. Such component usages differ according to the design of the network.

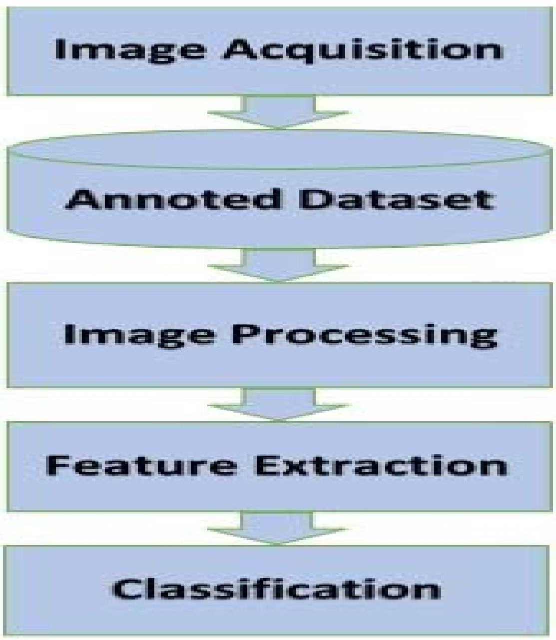

Fig. 5. Real-time Facemask Recognition Framework

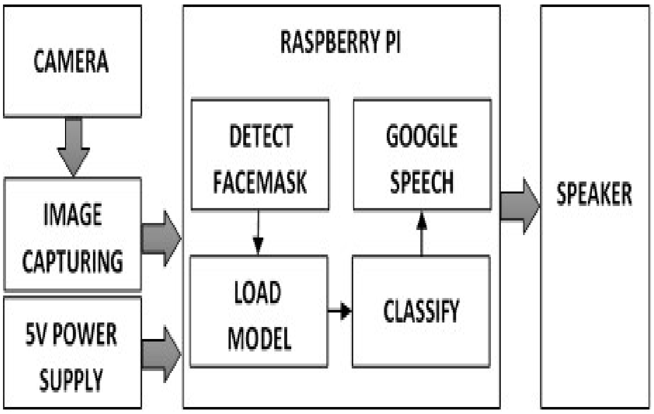

Fig. 6. Raspberry Pi Block Diagram

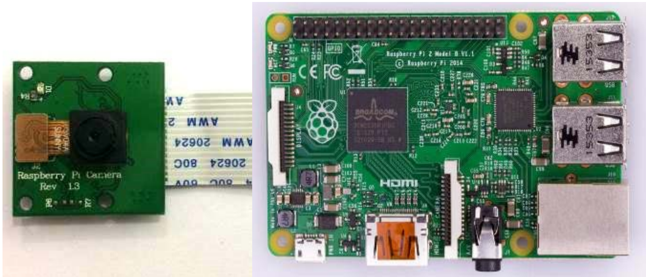

Fig. 7. The Raspberry Pi Camera Rev 1.3 board (left) and Raspberry Pi 2 Model B V1.1 (right).

A. Image Acquisition

The first step of the real-time facemask recognition system is image acquisition. High-quality images of the person posing with facemask wearing and not wearing facemask are obtained through digital cameras, cellphone cameras, or scanners.

B. Annotated Dataset Collection

A Knowledge-based dataset is created by proper labeling of the captured images with unique classes.

C. Image Processing

The obtained images that will be engaged in a pre-processing step is further enhanced specifically for image features during processing. The segmentation process divides the images into several segments and utilized in the extraction of facemask covered areas in the person's face from the background.

D. Feature-Extraction

This section involves the convolutionary layers that obtain image features from the resize images and is also joined after each convolution with the ReLU. Max and average pooling of the feature extraction decreases the size. Ultimately, both the convolutional and the pooling layers act as purifiers to generate those image characteristics.

E. Classification

The final step is to classify images, to train deep learning models along with the labeled images to be trained on how to recognize and classify images according to learned visual patterns. The authors used an open-source implementation via the TensorFlow module, using Python and OpenCV including the VGG-16 CNN model. The authors applied a supervised model of learning, with training and test sets divided to 80% for instruction and 20% for research. Three metrics were used to measure the model's performance: accuracy, training time, and learning error. In the conduct of experiments, the input parameters were set equally to 224 according to its input image width and height, batch size during training is set to 64 images and 100 iterations is set to the number of epochs. ADAM optimizer with the learning rate of 0.0001 is set for optimization. The study applies 12,500 images per class and this data is enough to train a deep learning model. Consequently, the data augmentation technique is also applied to intensify the quantity of image-data by simply employing rotation, rescaling, scrolling, and zooming procedures. With the rescale of 1./255, the dropout rate is set at 50 percent and used the same parameters for data augmentation. It leads to a multiplier for each pixel of the image, with the horizontal flip, nearest fill-mode, 0.3 factor for zoom range, 0.2 width shift range, horizontal and vertical shift factors, and 20 rotation range.

In the experiment the researchers conducted, they utilized a gaming laptop computer with an Intel Processor Core i7-8750 2.20 GHz base speed, Graphics Card NVIDIA Geforce GTX 1050 4 GB, the memory size of RAM is 6GB Kingston DDR4 2667 MHz, and the primary storage is SSD 256 GB. The trained CNN model was transferred to the Raspberry Pi to test its performance in detecting people wearing a facemask or not wearing. After the experiment was accomplished, the trained model along with other application program was transferred in the storage device of Raspberry-pi. The system runs in a Raspberry-pi as represented in the diagram shown in fig. 6 and fig. 7 respectively.

IV. Results and Discussion

The 96% validation accuracy was achieved during the training of the CNN model. This is the highest recorded rate after several experiments conducted with batch size set to 64 and 100 iterations for epochs as illustrated in fig. 8 and fig. 9 for performance tests results from visualization through accuracy and loss. Figures 10 and 11 display different test results on the performance of the model in detecting persons wearing a facemask or not wearing. The left image in fig. 10 shows the result of the detection of a person in the image not wearing facemask with a 53.69% rate and on the right image shows a person not wearing a facemask with a 91.22% accuracy rate. This will prompt the system to set an alarm. The left image in fig. 11 shows a person in the image wearing a facemask with an accuracy rate of 98.94% and the image in the right of fig. 11 shows a person wearing a facemask with a 71.84% accuracy rate.

	precision	recall	f1-score	support
NOfacemask	0.95	1.00	0.97	69
facemask	1.00	0.89	0.94	36
accuracy			0.96	105
macro avg	0.97	0.94	0.96	105
weighted avg	0.96	0.96	0.96	105

Fig. 8. Accuracy result of the trained model.

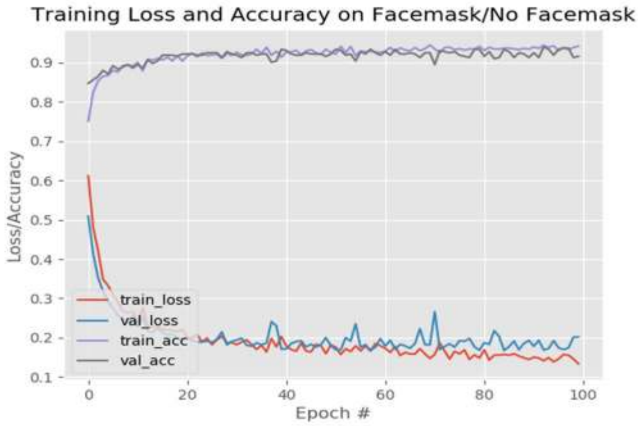

Fig. 9. The result of the Accuracy/Loss performance test during model training.

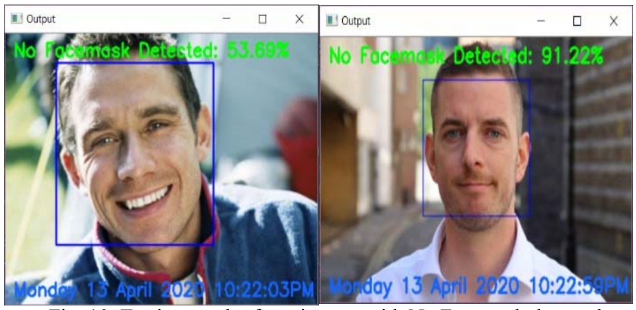

Fig. 10. Testing result of test images with No Facemask detected.

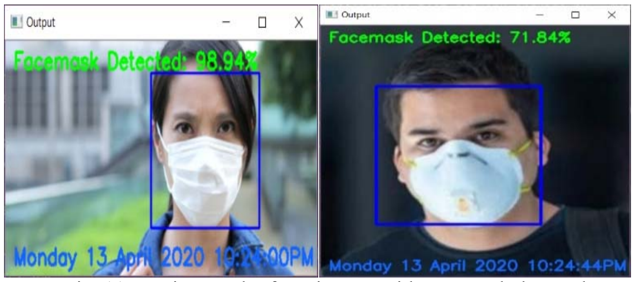

Fig. 11. Testing result of test images with Facemask detected.

V. Conclusion and Recommendations

This paper manuscript presented a study on real-time facemask recognition with an alarm system through deep learning techniques by way of Convolutional Neural Networks. This process gives a precise and speedily results for facemask detection. The test results show a distinguished accuracy rate in detecting persons wearing a facemask and not wearing a facemask. The trained model was able to perform its undertaking using the VGG-16 CNN model achieving a 96% result for performance accuracy. Moreover, the study presents a useful tool in fighting the spread of the COVID-19 virus by detecting a person who wears a facemask or not and sets an alarm if the person is not wearing a facemask.

Future works include the integration of physical distancing, wherein the camera detects the person wearing a facemask or not and at the same time measures the distance between each person and creates an alarm if the physical distancing does not observe properly. The integration of several models of CNNs and compare each model with the highest performance accuracy during training to increase the performance in detecting and recognizing people wearing facemasks is suggested. Also, the researchers recommend a different optimizer, enhanced parameter settings, fine-tuning, and using adaptive transfer learning models.

ACKNOWLEDGMENT

The authors express their gratitude to the University of Antique with its ever-supportive university president, Dr. Pablo S. Crespo, Jr., and the Vice President for RECETS, Dr. Nanette V. Dionisio in the conduct of this project study.

REFERENCES

[1] Yu, P., Zhu, J., Zhang, Z., & Han, Y. (2020). A Familial Cluster of Infection Associated With the 2019 Novel Coronavirus Indicating Possible Person-to-Person Transmission During the Incubation Period. The Journal of infectious diseases, 221(11), 1757–1761. https://doi.org/10.1093/infdis/jiaa077

[2] Chavez, S., Long, B., Koyfman, A. & Liang, S. Y. Coronavirus Disease (COVID-19): A primer for emergency physicians. Am J Emerg Med, https://doi:10.1016/j.ajem.2020.03.036 (2020).

[3] World Health Organization. Coronavirus disease 2019 (COVID-19) Situation Report– 142, 2020, [cited 10 June 2020], https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200610-covid-19-sitrep-142.pdf?sfvrsn=180898cd_6

[4] Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D. Y., Chen, L., & Wang, M. (2020). Presumed Asymptomatic Carrier Transmission of COVID-19. JAMA, 323(14), 1406–1407. Advance online publication. https://doi.org/10.1001/jama.2020.2565

[5] Centers for Disease Control and Prevention. Interim Infection Prevention and Control Recommendations for Patients with Suspected or Confirmed Coronavirus Disease 2019 (COVID-19) in Healthcare Settings. 2020 [cited 5 June 2020]. https://www.cdc.gov/coronavirus/2019-ncov/hcp/infection-control-recommendations.html

[6] Korea Centers for Disease Control and Prevention. Infection Prevention and Control Recommendations for Patients with Suspected or Confirmed Coronavirus Disease 2019 (COVID-19) in Healthcare Settings [in Korean]. 2020 [cited 5 June 2020]. http://ncov.mohw.go.kr/duBoardList.do?brdId=2&brdGubun=28

[7] Sim, S. W., Moey, K. S. & Tan, N. C. The use of facemasks to prevent respiratory infection: a literature review in the context of the Health Belief Model. Singapore Med J 55, 160-167, https://doi:10.11622/smedj.2014037 (2014).

[8] Cowling, B. J. et al. Facemasks and hand hygiene to prevent influenza transmission in households: a cluster randomized trial. Ann Intern Med 151, 437-446, https://doi:10.7326/0003-4819-151-7-200910060-00142 (2009).

[9] Tracht, S. M., Del Valle, S. Y. & Hyman, J. M. Mathematical modeling of the effectiveness of facemasks in reducing the spread of novel influenza A (H1N1). PLoS One 5, e9018, https://doi:10.1371/journal.pone.0009018 (2010).

[10] Jefferson, T. et al. Physical interventions to interrupt or reduce the spread of respiratory viruses. Cochrane Database Syst Rev, CD006207, https://doi:10.1002/14651858.CD006207.pub4 (2011).

[11] Feng, S., Shen, C., Xia, N., Song, W., Fan, M., & Cowling, B. J. (2020). Rational use of face masks in the COVID-19 pandemic. The Lancet. Respiratory medicine, 8(5), 434–436. https://doi.org/10.1016/S2213-2600(20)30134-X

[12] LeCun, Y., Kavukcuoglu, K., Farabet, C. & Ieee. in 2010 Ieee International Symposium on Circuits and Systems IEEE International Symposium on Circuits and Systems 253-256 (2010).

[13] Zhang, K. P., Zhang, Z. P., Li, Z. F. & Qiao, Y. Joint Face Detection and Alignment Using Multi-task Cascaded Convolutional Networks. Ieee Signal Proc Let 23, 1499-1503, https://doi:10.1109/Lsp.2016.2603342 (2016).

[14] Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. Communications of the Acm 60, 84-90, https://doi:10.1145/3065386 (2017).

[15] Cifuentes-Alcobendas, G. & Dominguez-Rodrigo, M. Deep learning, and taphonomy: high accuracy in the classification of cut marks made on fleshed and defleshed bones using convolutional neural networks. Sci Rep 9, 18933, https://doi:10.1038/s41598-019-55439-6 (2019).

[16] Qin, B. & Li, D. (2020), Identifying Facemask-wearing Condition using Image Super-resolution with ClassificationNetwork to Prevent COVID-19. https://doi.org/10.21203/rs.3.rs-28668/v1

[17] S. V. Militante, B. D. Gerardo, and N. V. Dionisio, "Plant Leaf Detection and Disease Recognition using Deep Learning," 2019 IEEE Eurasia Conference on IoT, Communication and Engineering (ECICE), Yunlin, Taiwan, 2019, pp. 579-582, https://doi: 10.1109/ECICE47484.2019.8942686.Sammy Sugarcane

[18] S. V. Militante, B. D. Gerardo, and R. P. Medina, "Sugarcane Disease Recognition using Deep Learning," 2019 IEEE Eurasia Conference on IoT, Communication, and Engineering (ECICE), Yunlin, Taiwan, 2019, pp. 575-578, https://doi: 10.1109/ECICE47484.2019.8942690.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[20] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” CoRR, vol. abs/1409.1556, 2014.

[21] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.

[22] K. Li, G. Ding, and H. Wang, “L-fcn: A lightweight fully convolutional network for biomedical semantic segmentation,” in 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Dec 2018, pp. 2363–2367.

[23] X. Fu and H. Qu, “Research on semantic segmentation of high-resolution remote sensing image based on full convolutional neural network,” in 2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE), Dec 2018, pp. 1–4.

[24] S. Kumar, A. Negi, J. N. Singh, and H. Verma, “A deep learning for brain tumor MRI images semantic segmentation using fcn,” in 2018 4th International Conference on Computing Communication and Automation (ICCCA), Dec 2018, pp. 1–4.

[25] Militante, S. Fruit Grading of Garcinia Binucao (Batuan) using Image Processing. International Journal of Recent Technology and Engineering (IJRTE), vol. 8 issue 2, pp. 1829-1832, 2019

[26] M. G. Lim and J. H. Chuah, "Durian Types Recognition Using Deep Learning Techniques," 2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia, 2018, pp. 183-187, https://doi: 10.1109/ICSGRC.2018.8657535.

[27] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” 2017.

[28] K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” Comput. Electron. Agric., vol. 145, no. September 2017, pp. 311–318, 2018.

[29] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” Comput. Electron. Agric., vol. 147, no. July 2017, pp. 70–90, 2018.

[30] S. V. Militante and B. D. Gerardo, "Detecting Sugarcane Diseases through Adaptive Deep Learning Models of Convolutional Neural Network," 2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Kuala Lumpur, Malaysia, 2019, pp. 1-5,

https://doi: 10.1109/ICETAS48360.2019.9117332.

[31] S. V. Militante, "Malaria Disease Recognition through Adaptive Deep Learning Models of Convolutional Neural Network," 2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Kuala Lumpur, Malaysia, 2019, pp. 1-6, https://doi: 10.1109/ICETAS48360.2019.9117446.