Journal of Human Centered Technology Vol. 1: No. 2 (2022) 7 – 16

Journal of Human Centered Technology

E-ISSN: 2821-3467
https://humentech.utm.my

Drone Image Based Illegal Crowd Detection for Covid-19 Disease Prevention Via Convolutional Neural Networks (CNNs) Transfer Learning

Ong Jia Eek[1], Muhammad Amir As'ari[1,2,3,*]

[1]BioMedical Instrumentation and Electronics Research Group (bMiE), Universiti Teknologi Malaysia, Johor Bahru, 81310, Malaysia

[2]Sport Innovation and Technology Center (SITC), Institute of Human Centered Engineering (iHumEn), Universiti Teknologi Malaysia, Johor Bahru, 81310, Malaysia

[3]School of Biomedical Engineering and Health Sciences, Universiti Teknologi Malaysia, Johor Bahru, 81310, Malaysia

*Corresponding Author amir-asari@utm.my

Received 06 April 2022; Accepted 14 May 2022; Available online 06 August 2022
https://doi.org/10.11113/humentech.v1n2.20

Abstract:
COVID-19 originated in Wuhan, China, in December 2019 and quickly became a global outbreak in January 2020. COVID-19 is a disease caused by SARS-CoV-2, which is a human transmission disease. Since it is a human transmission disease, thus mass gathering in public is not allowed to prevent the possible spread of COVID-19. However, the current monitoring technology, such as closed-circuit television (CCTV), only cover a limited area of the public and lack of mobility. Image classification is one of the approaches that can detect crowds in an image and can be done through either machine learning or deep learning approach. Recently, deep learning, especially convolutional neural networks (CNNs) outperform classical machine learning in image classification and the common approach for modelling CNN is through transfer learning. Thus, this study aims to develop a convolutional neural network that can detect illegal crowd gathering from offline drone view images through image classification using the transfer learning technique. Several models are used to train on the same dataset obtained, and the all-model performance is evaluated through a confusion matrix. Based on performance analysis, it shows that the ResNet50 model outperforms the VGG16 model and InceptionV3 model by achieving 95% test accuracy, 95% precision, 95% recall and 95% F1-score. In conclusion, it can be concluded that the deep learning approach uses a pre-trained convolutional neural network that can be used to classify object images in this study.

Keywords: Covid-19; Drone; Convolutional Neural Networks; Deep Learning; Crowd

1. Introduction

The first case of COVID-19 originated in Wuhan, China, in December 2019. COVID-19 resulted in a global outbreak in January 2020 [1]. Since then, cases of COVID-19 are detected in almost all countries around the world. Elderly people with severe illness are at higher risk for mortality while children rarely have symptoms, and they are less subjected to COVID-19. COVID-19 is a human transmission disease where this deadly virus can be transmitted from person to person directly or indirectly. The common symptoms of COVID-19 are cough, sore throat, fever, fatigue and headache [2]. The COVID-19 infected person may transmit the virus through coughing or sneezing. This pandemic has resulted in a massive global public health campaign done by each country to slow the spread of the virus [3]. Different organizations worldwide, such as CDC, WHO and Healthline, came out with guidelines to control the rapid spread of this deadly virus. The guidelines are: 1) Avoid social contact such as hugging, handshaking or touching; 2) Maintain social distancing by at least 1 meter in public places; 3) Wear a mask in public places at all times; 4) Stay at home if they are having symptoms such as fever, flu or coughing; and 5) Avoid coughing, sneezing in the open air.

Government and law enforcement authorities are active in preventing individuals' unnecessary movement in various cities, monitoring the spread and tracking individuals in compliance with the CDC guidelines. However, the government cannot control all locations, such as shopping malls, hospitals, government offices, and banks, and direct individuals to comply with the safety guidelines [4]. Several technologies are used in this pandemic to monitor and prevent illegal gathering. Closed-circuit television (CCTV) is an easily adaptable technology that allows detecting people wearing masks from an image or live feed fed to it [5]. However, CCTV has a limitation where it cannot cover all angles in a public place as it is fixed to monitor only one location. Hence, drone, one of technology which can be used in helping the authority to monitor the public movement in preventing the pandemic COVID-19.

A drone or Unmanned Aerial Vehicle (UAV) is an aircraft without a human pilot aboard. Drones have drastically raised attention in recent years due to their ability to perform a variety of applications beyond military applications to public and civil sectors [6]. Drones are increasingly popular for monitoring the environment with the equipment of additional hardware such as sensors and cameras. Recently, the drone has been used in many applications such as remote sensing [7], security surveillance [8-11], agriculture [12], inspection [13] and rescuing [14]. This study has been focused on establishing the illegal gathering prevention through crowd detection from the drone-based image.

In analyzing drone data, there are several techniques available for image classification. The state-of-art techniques in drone image analysis consist of both machine learning (ML) and deep learning (DL) [15]. Drone view image classification has been used primely in remote sensing such as building area estimation, crop monitoring, site inspection and protection of marine life. In general, the approach of existing work can be categorized into two, which are conventional machine learning and deep learning-based image classification.

Machine learning uses algorithms to make decisions based on data on what is learned while. Work by Mountrakis et al. [16] is one of the examples of implementation of machine learning called Support Vector Machine (SVM) [17] in drone applications. SVM can classify both linear and non-linear data, and it is an optimal hyper-plane based mathematical learning scheme. The ultimate idea behind SVMs is to use the optimal hyperplane to be used to classify of linearly separable patterns [18]. SVM performs classification by transforming the original training data into multidimensional space and constructing hyper-plane in higher dimensional. SVMs algorithm can simplify the problem by classifying the patterns into linear and non-linear [19]. As dimensionality of data increased, the classification problems have drastically increased. This results in higher computational cost and memories usage are required to solve the high dimensional data [20]. Feature extraction and feature selection are the common approaches in machine learning. However, during feature extraction, human intervention is still involved in deciding whether to follow or avoid some processing techniques. A drawback of feature extraction is that the information on the original feature contributes often lost along the process, and the combination of original features is usually not interpretable [20]. Recently, deep learning has been introduced to solve not only the human intervention in feature extraction but also improve the previous machine learning especially in dealing with big data.

Deep learning algorithm is a subset of machine learning algorithms. Deep learning relies on layers of artificial neural networks. Deep learning has rapid advancement due to the availability of data and drastically improvement in hardware technologies [21]. Different kinds of deep learning networks are introduced for various domains, such as convolutional neural networks for image processing [21]. In this particular study, a deep learning image classification model is built to detect a crowd in an image using convolutional neural networks. CNNs are the most common networks used with image classification in deep learning. CNNs were inspired by the human visual system proposed by LeCun and Bengio [22]. Convolutional neural networks (CNN) have become the dominant approach for visual object recognition due to the improvement of hardware and network structure.

One of the dominant approaches in modelling the CNN for image classification is transfer learning. Transfer Learning is a technique whereby a model is trained and developed for one task and is then re-used on a second related task. It refers to the situation in which what has been learned in one setting is exploited to boost optimization in another setting [23]. Transfer learning is commonly applied to a new dataset that is smaller than the original dataset used to train the pre-trained model [24]. In fact, transfer learning can leverage knowledge from pre-trained models and use it to solve similar problems. Through transfer learning, pre-trained models can be trained faster, more accurately and require less training data. There are several well-established CNN models also known as pre-trained model that can be used in transfer learning techniques such as Inception V3, VGG16 and ResNet50. Inception V3 is the winner of ILSVRC 2014 which achieved a top-5 error rate of 6.67%. The network consists of 22 layers in-depth and goes deeper in parallel paths with different receptive field sizes. The Inception V3 transfer learning model able to be utilized for improvement performance for most of the computer vision task on high quality, learned visual features [25].

VGG16 is a convolution neural network (CNN) architecture with convolution layers of 3x3 filter with stride one and always used same padding and max pool layer of 2x2 filter of stride 2 [26]. VGG16 does not have a large number of hyperparameter compared with other traditional convolution neural networks. VGG16 represent 16 layers of convolution neural network have weights and VGG16 consist of approximately 138 million parameters as it has a large network. The reduction in the number of trainable variables improves the learning rate and becomes more robust to overfitting [27]. However, the VGG16 model is large in file size due to large number of weights parameter and high inference time. Nevertheless, since VGG16 consists of a large number of parameters, it results in a longer training time and consumes a lot of time and computation power [28]. ResNet50 is a convolution neural network with 48 layers along with 1 Max Pool and 1 Average Pool layer. ResNet50 architecture solves vanishing gradient problems through having shortcut connection, and ResNet50 is a network build repeated with Residual block.

Recently, several works have focused on formulating the CNNs from drone view based images for developing several applications. For example, the work by Bidare et al. [29] proposed sidewalk detection toward autonomous drone navigation development, while work by Agrawal and Meleet [30] introduced natural disaster classification through the drone view image. Other than that, the study by Hafeez et al. [31] implemented CNNs model for monitoring the crop from the drone view image. However, little study has been done for modelling the CNNs for detecting detect illegal gathering crowds which is beneficial in preventing the spreading of Covid-19 virus infection. Thus, this project has been focused on classifying and detecting crowds from the drone view image by formulating the CNNs through transfer learning approach. Furthermore, the study has also evaluated several well-established CNNs such as Inception V3, VGG16 and ResNet50 as pre-trained models to detect illegal gathering crowds via the transfer learning approach.

2. Materials and Methods

2.1 Main framework

The aim is to develop automated image classification to detect crowds using deep pre-trained CNN models like ResNet50, InceptionV3 and VGG16. The performance of each model was evaluated using the confusion matrix. The overview of the project is shown in the block diagram in Figure 1.

2.2 Data preparation

Collecting raw drone images on static crowd drone view images and static non-crowd drone view images was being done. The drone dataset images are 400 in total and were divided into two classes: crowd and non-crowd. Each class have 200 images. The guideline for both crowd and no crowd dataset acquisition was as static images, top view drone images, and the images must be outdoor view.

The dataset was split according to the ratio of 0.7:0.2:0.1 for training, validation and test dataset. Both crowd and non-crowd datasets were resized to 244 x 244 pixels. The image augmentation was done to increase the data variation by adding slightly modified copies from existing data sets to generate the generalized model. Image augmentation only applies to training datasets and the image augmentation techniques used are shuffle and image rotation. Figure 2 shows the overall process executed in this data preparation while Figure 3 and 4 show the sample crowd and non-crowd images in the developed dataset.

2.3 Training phase

In this phase, the pre-trained convolutional neural networks called VGG16, InceptionV3 and ResNet50 were used to execute the transfer learning technique. The final layer of the transfer learning, which consisted of the classification layer, was removed and replaced with a new classifier layer. Training of the model on the new classification layer was done with the prepared crowd and non-crowd datasets. The flow chart for the training of the pre-trained transfer learning model is shown in Figure 5. Table 1 illustrates the training setup whereby several hyperparameters were initially set during the training phase for all mentioned models.

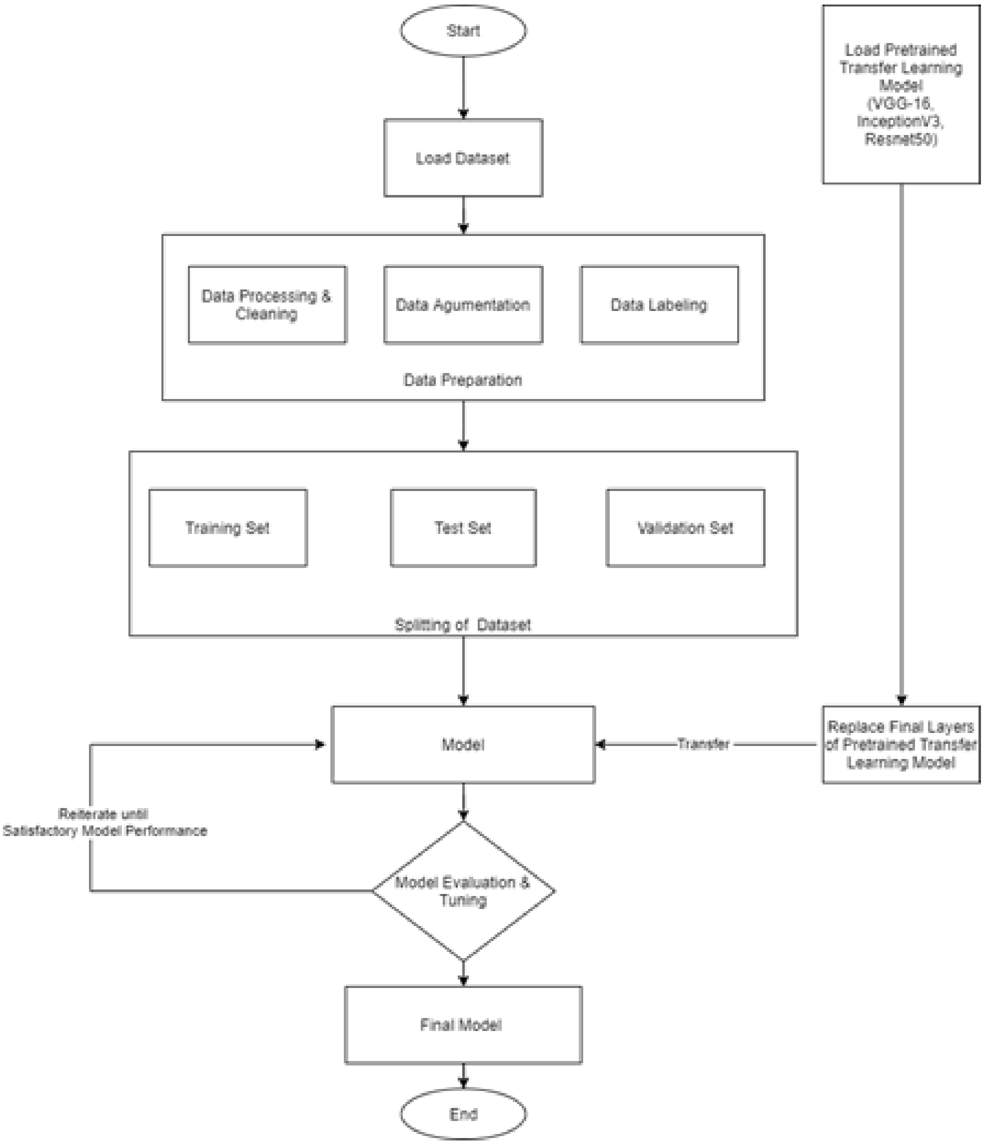

Figure 1. Main framework

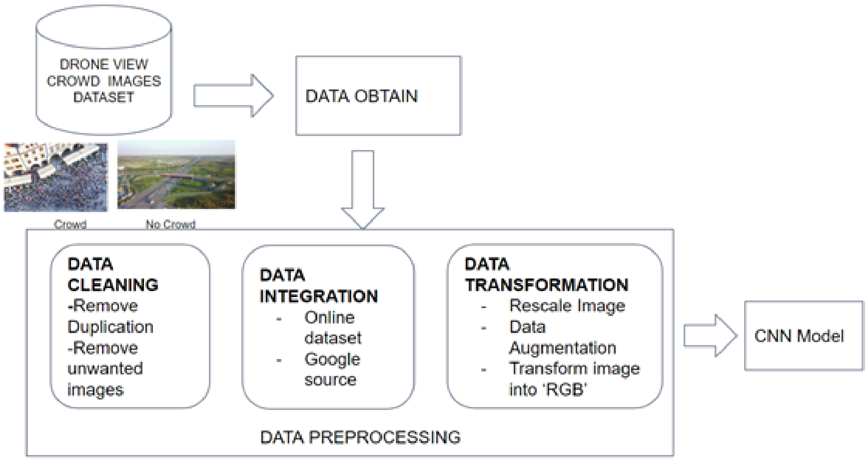

Figure 2. Flow chart of data preparation

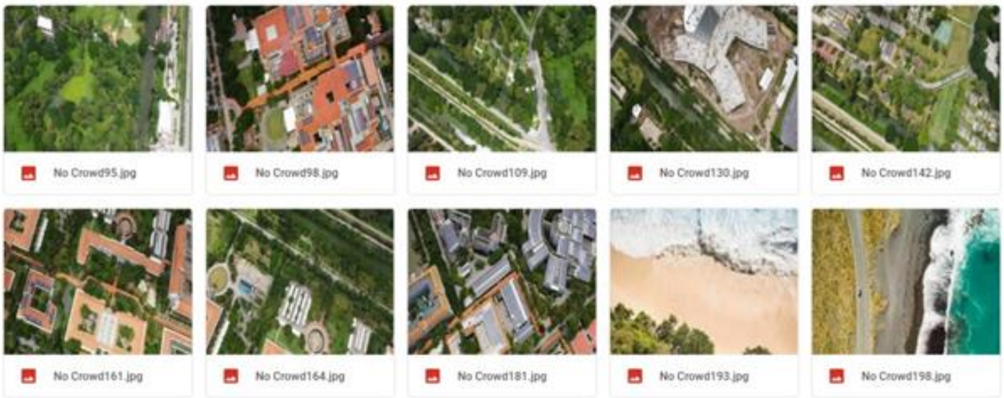

Figure 3. Sample images of non-crowd images

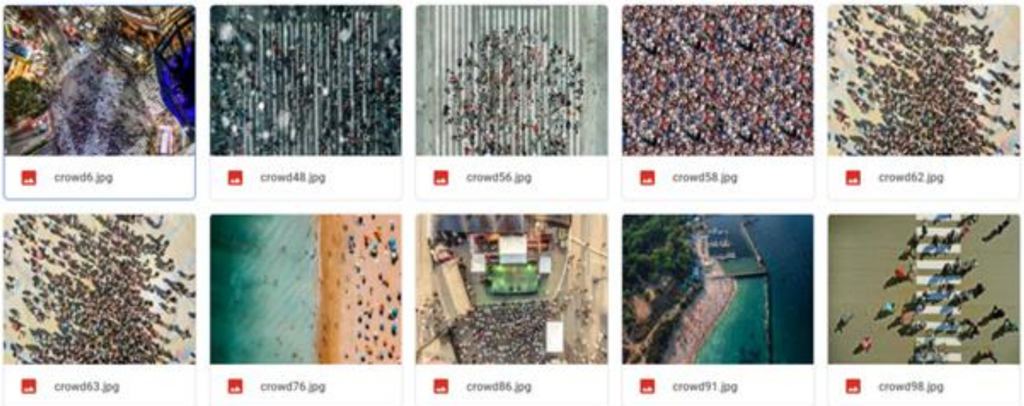

Figure 4. Sample image of crowd images

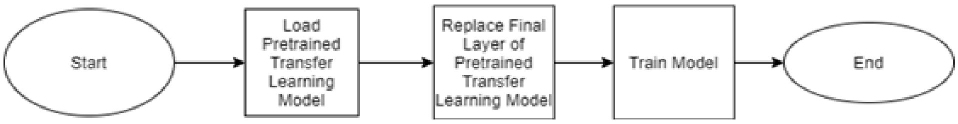

Figure 5. Flow Chart for training of pretrained transfer learning model

Table 1. Initial training setup for all models

Hyperparameter	**Value**
Training, validation and testing ratio	0.7:0.2:0.1
Number of epochs	30
Dropout rate	0.5
Batch size	5

2.4 Testing phase and performance measurement

There were four main parameters used to evaluate a classification model: accuracy, precision, recall, and F1-score. The confusion matrix has commonly been used to describe the performance of classification models on the test data needed to provide the four parameters mentioned earlier. The confusion matrix consists of 2 classes which are positive and negative. True positive (TP) refers to the number of predictions that correctly predict the positive class by the classifier as positive. In contrast, True negative (TN) is defined as the number of predictions that correctly predict the negative class by the classifier as negative. Other than that, False Positive (FP) refers to the number of predictions that incorrectly predict the negative class by the classifier as positive. False Negative (FN) refers to the number of predictions that incorrectly predict the positive class by the classifier as negative. The confusion matrix provides a simple yet efficient performance evaluation of the model. The value of each model for the confusion matrix had to be recorded down for further analysis. In addition, several parameters can be computed from the confusion matrix for binary classification.

$$accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (1)$$

$$precision = \frac{TP}{TP + FP} \quad (2)$$

$$recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 \quad score = \frac{2 \times recall \times precision}{recall + precision} \quad (4)$$

3. Results and Discussion

All three models, VGG16, ResNet50 and InceptionV3 were compared according to testing accuracy, precision, recall and F1-score. ResNet50 ranked 1st among all three models with a test accuracy of 95%, followed by VGG16 and InceptionV3 model ranked the last with only 65% in test accuracy (Table 2). It can be concluded that ResNet50 is the model with the best performance, while the InceptionV3 Model performs the worst among all three models in detecting crowd and non-crowd images. Figure 6, Figure 7 and Figure 8 illustrate the confusion matrix of all mentioned CNN models. As can be seen in Figure 6, only 2 out of 40 images are misclassified by the ResNet50 model. While InceptionV3 (see Figure 8) mostly fails on the non-crowd side, only eight out of 20 images are correctly classified, contributing to the poor accuracy result. However, VGG16 (Figure 7) manage to detect all the crowd images correctly while missing 5 out of 20 images as non-crowd images. As summarized in Table 2, ResNet50 outperforms other CNNs models in representing crowd detection with precision, recall, and F1 score of 95%.

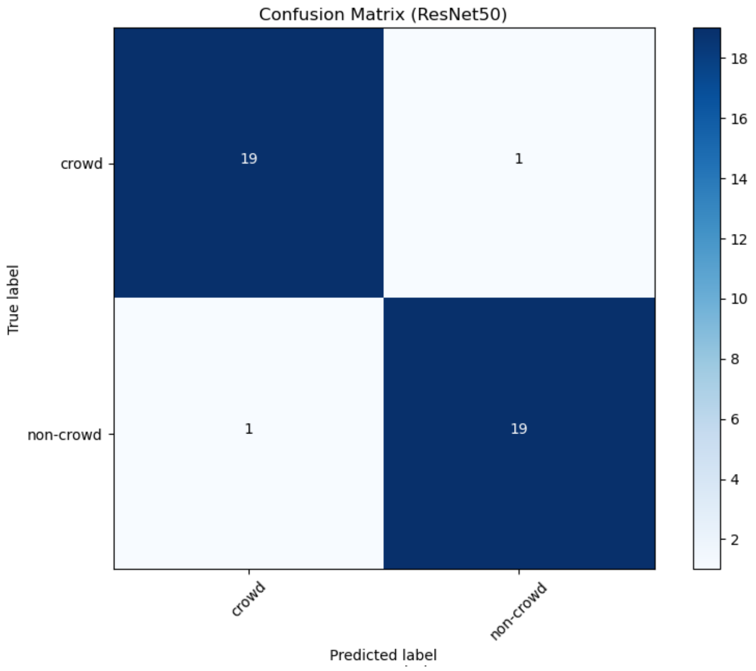

Figure 6. Confusion matrix of ResNet50 on testing dataset

Table 2. Summary of trained network performance

Model	**ResNet50**	**VGG16**	**InceptionV3**
Precision	95%	90%	70%
Sensitivity	95%	88%	65%
F1 score	95%	87%	63%
Accuracy	95%	87.5%	65%

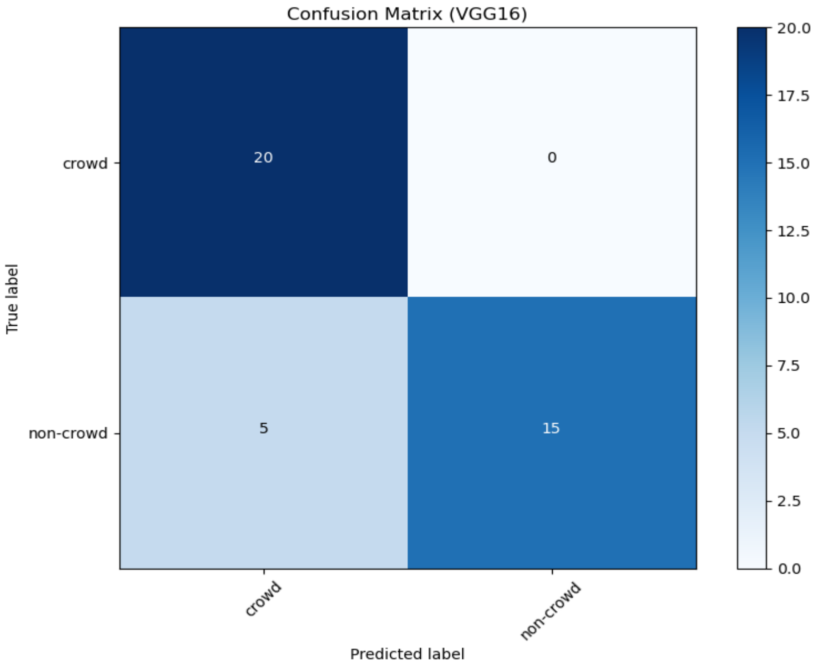

Figure 7. Confusion matrix of VGG16 on testing dataset

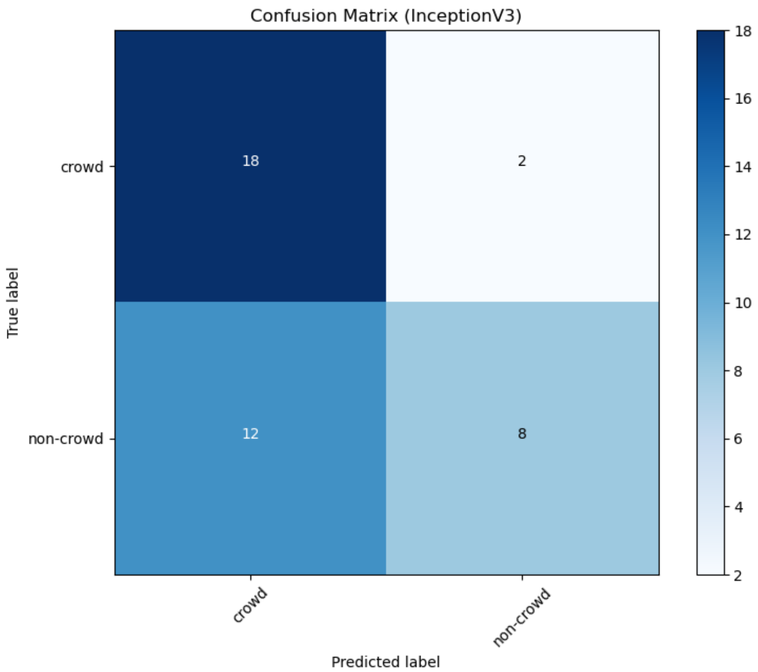

Figure 8. Confusion matrix of InceptionV3 on testing dataset

4. Conclusion

In conclusion, this study is to develop a transfer learning model that detects crowds through offline drone view images that can help local authorities in the COVID-19 pandemic. With the deep learning approach, there are three models, including VGG16, InceptionV3 and ResNet50, in developing a transfer learning model that can perform image classification and detect crowds through offline drone view images. Among all three models, ResNet50 outperform the other two transfer learning model with a test accuracy of 95%. Upon successful development of the convolutional neural network model, ResNet50 can perform image classification and detect crowds through offline drone view images, which can help local authorities in the COVID-19 pandemic. Therefore, it can be concluded that a deep learning approach by using the pre-trained convolutional neural network can be used to classify object images in this study.

There is a limitation in the number of datasets for the crowd and non-crowd images in this study. The total images for both classes are only 400 as only limited images can be found in the online sources. Therefore, the size of the dataset must be improved in the future to prevent overfitting, underfitting and improve the performance of the model. Nevertheless, through the improvement of the dataset, the model can classify a wide variety of different scenarios of the crowd and no crowd images. Besides, there is a limitation in model development; the model was trained in an ideal environment where the images are static, and the image's resolution is high. Trained models might perform worse than expected when adopted in drones and dealing with real-time images. Thus, the recommendation for future work is to collect real-time images from the drone and train the pre-trained model using the actual images from the drone. Finally, the developed models are established from the basic model configuration. However, for future work recommendations, exploration on fine-tune in different hyperparameters such as activation function and learning rate can be used to fine-tune the model and improve the model accuracy.

Acknowledgment

The authors would like to express their appreciation to Universiti Teknologi Malaysia (UTM) for endowing this research and the Ministry of Higher Education (MOHE), Malaysia for supporting this research work under Research Grant No. Q.J130000.2651.16J20.

References

[1] M. Ciotti, M. Ciccozzi, A. Terrinoni, W.-C. Jiang, C.-B. Wang and S. Bernardini, The COVID-19 pandemic, Critical Reviews in Clinical Laboratory Sciences, 2020, 57(6):365–388. https://doi.org/10.1080/10408363.2020.1783198.

[2] T. Struyf, J. J. Deeks, J. Dinnes, Y. Takwoingi, C. Davenport, M. M. Leeflang, R. Spijker, L. Hooft, D. Emperador, S. Dittrich, J. Domen, S. R. A. Horn, A. Van den Bruel and C. COVID-19 Diagnostic Test Accuracy Group, Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID - 19 disease, Cochrane Database of Systematic Reviews, 2020, 7(7):CD013665. https://doi.org/10.1002/14651858.CD013665.

[3] J. J. Van Bavel, K. Baicker, P. S. Boggio, V. Capraro, A. Cichocka, M. Cikara, M. J. Crockett, A. J. Crum, K. M. Douglas, J. N. Druckman, J. Drury, O. Dube, N. Ellemers, E. J. Finkel, J. H. Fowler, M. Gelfand, S. Han, S. A. Haslam, J. Jetten, S. Kitayama, D. Mobbs, L. E. Napper, D. J. Packer, G. Pennycook, E. Peters, R. E. Petty, D. G. Rand, S. D. Reicher, S. Schnall, A. Shariff, L. J. Skitka, S. S. Smith, C. R. Sunstein, N. Tabri, J. A. Tucker, S. van der Linden, P. van Lange, K. A. Weeden, M. J. A. Wohl, J. Zaki, S. R. Zion and R. Willeret, Using social and behavioural science to support COVID-19 pandemic response, Nature Human Behaviour, 2020, 4:460–471. https://doi.org/10.1038/s41562-020-0884-z.

[4] M. I. Uddin, S. A. A. Shah and M. A. Al-Khasawneh, A novel deep convolutional neural network model to monitor people following guidelines to avoid COVID-19, Journal of Sensors, 2020, 2020. https://doi.org/10.1155/2020/8856801.

[5] B. Roy, S. Nandy, D. Ghosh, D. Dutta, P. Biswas and T. Das, MOXA: A deep learning based unmanned approach for real-time monitoring of people wearing medical masks, Transactions of the Indian National Academy of Engineering, 2020, 5(3):509–518. https://doi.org/10.1007/s41403-020-00157-z.

[6] A. Bahabry, X. Wan, H. Ghazzai, H. Menouar, G. Vesonder and Y. Massoud, Low-altitude navigation for multi-rotor drones in urban areas, IEEE Access, 2019, 7:87716–87731. https://doi.org/10.1109/ACCESS.2019.2925531.

[7] J. Chen, G. Wang, L. Luo, W. Gong and Z. Cheng, Building area estimation in drone aerial images based on mask R-CNN, IEEE Geoscience and Remote Sensing Letters, 2020, 18(5):891–894. https://doi.org/10.1109/LGRS.2020.2988326.

[8] V. Puri, A. Nayyar and L. Raja, Agriculture drones: A modern breakthrough in precision agriculture, Journal of Statistics and Management Systems, 2017, 20(4):507–518. https://doi.org/10.1080/09720510.2017.1395171.

[9] D. Bloom, P. A. Butcher, A. P. Colefax, E. J. Provost, B. R. Cullis and B. P. Kelaher, Drones detect illegal and derelict crab traps in a shallow water estuary, Fisheries Management and Ecology, 2019, 26(4):311–318. https://doi.org/10.1111/fme.12350.

[10] M. Bonetto, P. Korshunov, G. Ramponi and T. Ebrahimi, Privacy in mini-drone based video surveillance, 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2015, 4:1–6. https://doi.org/10.1109/FG.2015.7285023.

[11] A. Al-Sheary and A. Almagbile, Crowd monitoring system using unmanned aerial vehicle (UAV), Journal of Civil Engineering and Architecture, 2017, 11(11):1014–1024. https://doi.org/10.17265/1934-7359/2017.11.004.

[12] P. Radoglou-Grammatikis, P. Sarigiannidis, T. Lagkas and I. Moscholios, A compilation of UAV applications for precision agriculture, Computer Networks, 2020, 172:107148. https://doi.org/10.1016/j.comnet.2020.107148.

[13] R. Ashour, T. Taha, F. Mohamed, E. Hableel, Y. Abu Kheil, M. Elsalamouny, M. Kadadha, K. Rangan, J. Dias, L. Seneviratne and G. Cai, Site inspection drone: A solution for inspecting and regulating construction sites, 2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS), 2016, 1–4. https://doi.org/10.1109/MWSCAS.2016.7870116.

[14] R. Tariq, M. Rahim, N. Aslam, N. Bawany and U. Faseeha, Dronaid: A smart human detection drone for rescue, 15th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT), 2018, 33–37. https://doi.org/10.1109/HONET.2018.8551326.

[15] S. Bhatnagar, L. Gill and B. Ghosh, Drone image segmentation using machine and deep learning for mapping raised bog vegetation communities, Remote Sensing, 2020, 12(16):2602. https://doi.org/10.3390/rs12162602.

[16] G. Mountrakis, J. Im and C. Ogole, Support vector machines in remote sensing: A review, ISPRS Journal of Photogrammetry and Remote Sensing, 2011, 66(3):247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001.

[17] C. Cortes and V. Vapnik, Support-vector networks, Machine Learning, 1995, 20(3):273–297. https://doi.org/10.1007/BF00994018.

[18] C. Zhang and Z. Zhang, A survey of recent advances in face detection, Microsoft Corporation, MSR-TR-2010-66, 2010.

[19] S. Karamizadeh, S. M. Abdullah, M. Halimi, J. Shayan and M. javad Rajabi, Advantage and drawback of support vector machine functionality, 2014 International Conference on Computer, Communications, and Control Technology (I4CT), 2014, 63–65. https://doi.org/10.1109/I4CT.2014.6914146.

[20] A. Janecek, W. Gansterer, M. Demel and G. Ecker, On the relationship between feature selection and classification accuracy, New Challenges for Feature Selection in Data Mining and Knowledge Discovery, 2008, 90–105.

[21] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M. L. Shyu, S. C. Chen and S. S. Iyengar, A survey on deep learning: Algorithms, techniques, and applications, ACM Computing Surveys (CSUR), 2018, 51(5):1–36. https://doi.org/10.1145/3234150.

[22] Y. LeCun and Y. Bengio, Convolutional networks for images, speech, and time series, in: The Handbook of Brain Theory and Neural Networks, ACM Digital Library: USA, 1998, 255–258.

[23] Y. Gao and K. M. Mosalam, Deep transfer learning for image-based structural damage recognition, Computer-Aided Civil and Infrastructure Engineering, 2018, 33(9):748–768. https://doi.org/10.1111/mice.12363.

[24] M. Hussain, J. J. Bird and D. R. Faria, A study on cnn transfer learning for image classification, in UK Workshop on Computational Intelligence, Springer: USA, 2018, 191–202. https://doi.org/10.1007/978-3-319-97982-3_16.

[25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, Rethinking the inception architecture for computer vision, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, 2818–2826. https://doi.org/10.48550/arXiv.1512.00567.

[26] R. Thakur, Step by step VGG16 implementation in Keras for beginners, Medium-Towards Data Science, 2019. https://towardsdatascience.com/step-by-step-vgg16-implementation-in-keras-for-beginners-a833c686ae6c (accessed Nov. 8, 2021).

[27] A. Anwar, Difference between AlexNet, VGGNet, ResNet and Inception, Medium-Towards Data Science, 2019. https://towardsdatascience.com/the-w3h-of-alexnet-vggnet-resnet-and-inception-7baaaecccc96 (accessed Nov. 8, 2021).

[28] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv: Computer Vision and Pattern Recognition, 2014. https://doi.org/10.48550/arXiv.1409.1556.

[29] M. Bidare, A. Srivastav and T. Khuu, CNN-based robust sidewalk identification for autonomous drone applications, Computer Science, 2020.

[30] T. Agrawal and M. Meleet, Classification of natural disaster using satellite & drone images with CNN using transfer learning, International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2021, 1–5. https://doi.org/10.1109/ICSES52305.2021.9633803.

[31] A. Hafeez, V. Tiwari, V. K. Verma, A. S. Ansari, M. A. Husain, S. P. Singh and A. N. Khan, Crop monitoring and automatic Wweed detection using drone, International Conference on Control, Automation, Power and Signal Processing (CAPS), 2021, 1–4. https://doi.org/10.1109/CAPS52117.2021.9730682.