

University of the Philippines - Cebu

Deep Learning Model on Real-time
Face Mask Detection

Fayne Roxenne A. Bagaipo

Prof. Dhong Fhel K. Gom-os

Word count: 1219

December 10, 2021

1 Introduction

Deep learning is composed of three or more neuron layers and is also a subset of machine learning. It is an attempt to copy the workaround of a human brain, able to cluster data and predict outcomes with incredible accuracy (Education 2020). To be more accustomed to their differences, let's look at the definition of traditional programming, machine learning, and deep learning. Regular/Traditional Programming is where users give computers explicit instructions using several programming languages to describe exactly how they want things to work. Machine Learning is learning from statistics. We program the machine to look at large amounts of data and then make guesses from the data. The results in this are proportional to how much data is you have and what features you have implemented. Deep Learning is where you make the computer learn from the data without knowing what features are going to be important.

As of early 2019, a deadly pandemic struck several countries deeming people's normal ways of living dangerous in fear of contracting COVID-19 the virus. Everyone around the world is instructed to wear face masks as a way of preventing the contraction of the mentioned virus. Almost two years have gone by and yet the situation seems to not have dwindled as much in terms of new cases of infections and deaths pertaining to the virus. In line with this, our only best option is to opt for vaccination, social distance, and wear our face masks when out in public. There are proven results as to how wearing a mask has a very low risk of spread compared to those who do not.

In public crowded places such as malls, groceries, airports, and other establishments, certain security protocols are observed. Security and sanitation patrols may need to constantly remind people to wear their masks which can prove tedious at times. Having said this, instead of increasing the burden of frontliners and further risk of exposure, adapting to this new normal can be made easier using technological advancements. Methods of deep learning is extremely useful in this regard as we can instruct it to learn from a given dataset of people who put on face masks as a means of protection and people who don't. From the dataset, the computer will be trained to differentiate people who are wearing their masks from the people who are not wearing them. This method will also be implemented in real time using OpenCV to perform these computer vision tasks. The use of MobileNetV2 also allows a cheaper and lightweight

alternative to the other more expensive deep learning methods available.

In a world where living in the pandemic is the new norm, observations made in this study include: (a) Importance of face masks as means of prevention during the virus pandemic; (b) Manual inspection of the public on whether they are abiding to the usage of face masks is too laborious; (c) Implementing a face mask detection in real-time with the use of deep learning methods whilst also choosing cost-efficient tools.

2 Analysis and Discussion

There have been several deep learning adaptations on face detection and recognition. It has evolved to a certain point where it is now used for security and privacy measures. A variety of these DL methods have reduced the detection accuracy and computational time, therefore allowing even more potential breakthroughs for automated image processing applications. As of the COVID-19 outbreak, these deep learning adaptations on face detection have evolved to cater to the detection of people wearing masks. Therefore, needing further discussion on the literature regarding machine learning based methods for detecting faces and detection of face masks.

A paper on this by Loey et al. (2021b), developed an identification method able to categorize three conditions for wearing a mask: masked, improperly masked, and unmasked. Their proposed model had two components using ResNet50, a deep transfer learning model and a Support Vector Machine (SVM), a classical supervised algorithm. In the paper, the performance used three traditional classifiers for model improvement (SVM, Decision Tree, and Ensemble classifiers). The results showed that the SVM classifier outperformed with the least training time consumed for all three datasets, achieving a 99.64% (RMFD), 99.49% (SMFD), and a 100% (LFW) accuracy.

For Nagrath et al. (2021), an SSDMNV2 model was designed using Tensorflow and Keras libraries, MobileNetV2 architecture as a classifier framework, Single Shot Multibox Detector as their detector and OpenCV DNN. Their use of multibox allowed for multiple objects to be captured in one shot making it comparable to the technique of another deep learning method, YOLO giving it a faster speed and higher accuracy. The technique deployed gave a score of 92.64%.

Meanwhile, Qin & Li (2020) developed a new face mask wearing condition identification method called SRCNet, that combines classification networks and image super-resolution. The method consists of an algorithm with four main steps for pre-processing of the image, cropping and detection of the face, and the condition identification. The SRCNet study outperformed other conventional deep learning image classification methods, achieving a 98.70% detection accuracy.

A paper from Hussain & Al Balushi (2020), proposed a facial detection and recognition model to identify facial features and authenticate them. Three different phases were implemented in a sequential process, a human face is detected in the first phase, the input captured is then brought to the second phase to be analyzed on the existing features with the support of a Keras CNN model. In the last phase, the face undergoes authentication and classified on whether it is projecting a certain emotion. These three objective models were validated and designed with an accuracy of 88%.

Another paper from Loey et al. (2021a)) conducted a deep learning algorithm for detecting masks using YOLO-v2 with the use of a ResNet50 model to achieve high results on discernment. It follows the transfer learning as the first paper for employing a feature extraction process based on ResNet50 but detection on face mask was employed using YOLO-v2. An average precision of 81% was shown using this model.

Similarly, Li et al. (2020) proposed the use of YOLO-v3 for better speed and accuracy on object detection. However, when used on face detection, several problems came to light such as the detection of smaller faces. To adapt, a loss function, Softmax was chosen instead in maximizing the segregation of inter-class features, and which also helped in decreasing the features in dimension to improve speed. YOLO-v3 proved to perform better than other existing techniques while maintaining speed and accuracy.

Furthermore, Wang et al. (2020) proposed datasets with masked faces for three types: MFDD, RMRFD, and SMFD. These datasets were done for the improvement of detecting face masks using deep learning methods. From the three datasets, the RFMRD achieved a 95% accuracy.

3 Conclusion

Among the literature presented regarding detection of faces and detection of face masks, several key concerns such as optimization issues, data heavy applications, and energy inefficiency are involved but the usage of varying deep learning methods allowed for comparability in speed, precision, and accuracy. Apart from this, the design and deployment of sturdy models on deep learning are factors in the enhancement of mask detection accuracy on publicly available datasets and in real-time. The proposal of a lightweight and cost-effective model for automatic or real-time detection in times of COVID-19 safety purposes will be employed.

References

Education, I. C. (2020), ‘What is deep learning?’.

URL: <https://www.ibm.com/cloud/learn/deep-learning>

Hussain, S. A. & Al Balushi, A. S. A. (2020), A real time face emotion classification and recognition using deep learning model, in ‘Journal of Physics: Conference Series’, Vol. 1432, IOP Publishing, p. 012087.

Li, C., Wang, R., Li, J. & Fei, L. (2020), Face detection based on yolov3, in ‘Recent Trends in Intelligent Computing, Communication and Devices’, Springer, pp. 277–284.

Loey, M., Manogaran, G., Taha, M. H. N. & Khalifa, N. E. M. (2021a), ‘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**, 102600.

URL: <https://www.sciencedirect.com/science/article/pii/S2210670720308179>

Loey, M., Manogaran, G., Taha, M. H. N. & Khalifa, N. E. M. (2021b), ‘A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the covid-19 pandemic’, *Measurement* **167**, 108288.

URL: <https://www.sciencedirect.com/science/article/pii/S0263224120308289>

Nagrath, P., Jain, R., Madan, A., Arora, R., Kataria, P. & Hemanth, J. (2021), ‘Ssdmnv2: A real time dnn-based face mask detection system using single shot multibox detector and mobilenetv2’, *Sustainable cities and society* **66**, 102692.

Qin, B. & Li, D. (2020), ‘Identifying facemask-wearing condition using image super-resolution with classification network to prevent covid-19’, *Sensors* **20**(18), 5236.

Wang, Z., Wang, G., Huang, B., Xiong, Z., Hong, Q., Wu, H., Yi, P., Jiang, K., Wang, N., Pei, Y. et al. (2020), ‘Masked face recognition dataset and application’, *arXiv preprint arXiv:2003.09093* .