ARTIFICIAL INTELLIGENCE THEORY and APPLICATIONS

ISSN: 2757-9778 || ISBN : 978-605-69730-2-4

More information available at aita.bakircay.edu.tr

Face Mask Detection on LabVIEW

Gülyeter ÖZTÜRK*, Osman ELDOĞAN, Durmuş KARAYEL, Gökhan ATALI

Sakarya University of Applied Sciences, Faculty of Technology, Dept. of Mechatronic Engineering, Turkey

* Corresponding Author: Sakarya University of Applied Sciences, Faculty of Technology, Turkey
Tel.: +90 264 6160248. E-Mail: gulyeterozturk@subu.edu.tr

Publication Information

Keywords :

- COVID-19;
- Face mask;
- Deep learning;
- Faster R-CNN Inception V2;
- Transfer learning.

Category : Research Article

Received : 07.07.2021
Accepted : 25.09.2021

A B S T R A C T

The world is facing a huge health crisis due to the coronavirus pandemic (COVID-19). The World Health Organization (WHO) has issued that the most effective preventive measure against the rapid spread of coronavirus is wearing a mask and keeping social distance in public places and crowded areas. Various studies have proven that wearing a face mask significantly reduces the risk of viral transmission, and also provides a sense of protection for people. But it is difficult to monitor and control people manually, especially in crowded areas.

In this study, a deep learning model is proposed to automatically detect people wearing face masks or not. The pre-trained Faster R-CNN Inception V2 deep learning model is fine-tuned with the transfer learning method and trained and tested on the Simulated Masked Face Dataset (SMFD). The model trained in the TensorFlow environment is accurate enough to detect the face mask. Thus, face mask detection is performed with the interface created on LabVIEW and a safe working environment can be maintained by controlling security violations in public living areas under control.

1. Introduction

WHO declared COVID-19 as a global epidemic due to the high transmission rates and death of many infected people caused by this disease. In line with the recommendations of WHO, many countries have imposed rules on citizens to wear masks in public places, pay attention to hygiene, especially hand cleaning, and keep a distance between people.

One of the measures taken to slow the spread of COVID-19 is to wear masks in crowded environments and public places. Various studies have been conducted to identify people who do not wear masks. One of them is a French government initiative to detect passengers who are not wearing masks in subway stations. For this, artificial intelligence software has been created and applied to security cameras in Paris metro stations [1]. Applications such as predicting the distribution of epidemics, preventing its spread, and facilitating early diagnosis using data of COVID-19 have been implemented with the help of AI. Qin and Li [2] has been successful in identifying face mask through their SRCNet classification network. The study divides people into three categories: wearing masks correctly, wearing them incorrectly and not wearing them at all. Loey, Manogaran, Taha and Khalifa create a hybrid model that uses deep and classic machine learning for face mask detection [3]. As the first stage of the model, features are extracted using Resnet50. In the second stage, the acquired features classify the faces with masks by using decision trees, support vector machines (SVM) and ensemble algorithms. In this study, RMFD (Real-World Masked Face Dataset), SMFD, and LFW (Labeled Faces in the Wild) datasets are used and the conclusion is that SVM provides the best performance among the classifiers. In another study by Loey, Manogaran, Taha and Khalifa [4], training, validation, and testing processes are conducted by merging two data sets. The researchers who use Resnet50 and transfer learning methods for feature extraction do with YOLO V2 for feature extraction and mask recognition process. Li, Wang, Li and Fei [5] train the YOLO V3 deep learning model for facial recognition by using Winder Face and Celeba data sets. The model is evaluated in the FDDB database with an accuracy of 93.9%. Din, Javed, Bae and Yi [6] recommend a GAN based network architecture for removing a mask and reconstructing the area covered by the mask. Rodriguez, Mucientes and Brea [7] identify whether medical personnel is wearing the required surgical masks in the operating room and set off an alarm when non-masked personnel is detected. Ejaz, Islam, Sifatullah and Sarker [8] apply Principal Component Analysis (PCA) to face recognition on the faces of people with or without masks. It is observed that the accuracy of PCA in the recognition of faces without masks reach 96.25%, while the accuracy in the recognition of faces with masks decrease to 68.75%.

As mentioned above, at this stage of our lives in which COVID-19 has emerged, various mask testing studies have been conducted and are underway for further refinement. In this study, the Faster R-CNN Inception V2 deep learning model is used to detect face mask through transfer learning.

2. Methodology

Deep learning models extract features from data through a lot of layers contained in their structures, and through the different functions they perform in those layers. In Figure 1, CNN is shown to extract advanced features from the input image by performing convolution processing in the hidden layers for the face recognition process. In order for the network to extract good features from the data, the data set must have a large amount of data representing the problem. Since the number of faces with masks in this study is limited, a transfer learning method is used to enable the network to give more accurate results. The Faster R-CNN Inception V2 model, previously trained on the COCO dataset, has been retrained with the new dataset based on transfer learning.

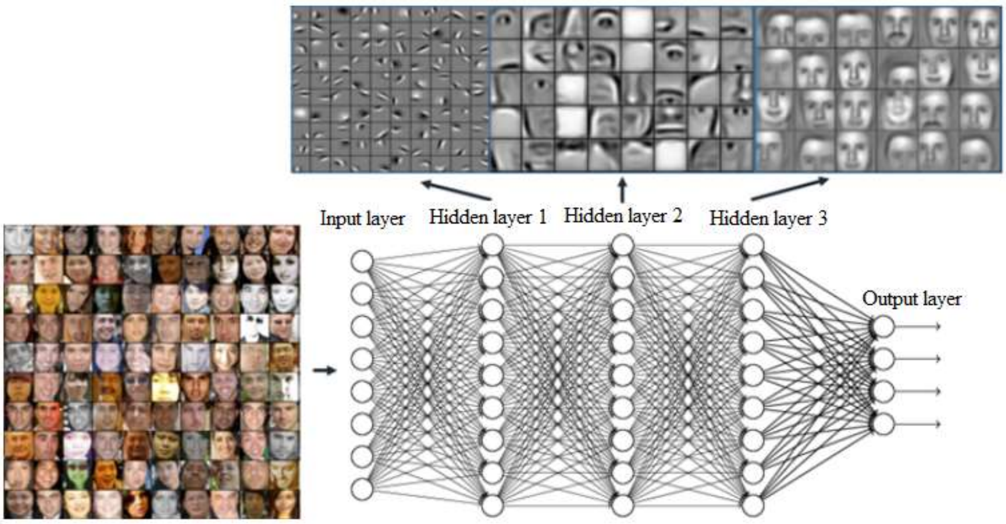


Figure 1. Features obtained for face recognition application in hidden layers of CNN [9].

The COCO dataset contains 328.000 images belonging to 91 object categories.

2.1 Dataset

In this study, a Simulated Masked Face Dataset (SMFD) [10] is used, which is consisting of 785 face images without masks and 785 simulated masked face images with masks that are added to faces in a simulated environment. Figure 2 shows images without masks and Figure 3 shows images of simulated masked faces.

Figure 2. Images without masks.

Figure 3. Simulated masked images.

Of the data set consisting of 1570 images, 80% is used for training and 20% for testing.

2.2 Transfer Learning

Deep neural networks extract features from data by using large amounts of data, and perform self-learning with these extracted features. Compared with other machine learning models, deep learning models with better performance in processes such as image classification and object detection require more computing power and large data sets, which make the research cost very high. In order for the network to be trained faster and with fewer datasets, a transfer learning approach is used. Through transfer learning, the weights of deep learning models previously trained on large data sets are transferred to the new research. Good performance can be obtained from models trained on small data sets through transfer learning.

People in the study are divided into those who are masked and those who are without masks (In Turkish masked: maskeli, without mask: maskesiz). Transfer learning is applied to the fine-tuned Faster R-CNN Inception V2 model, which has previously been trained on the COCO dataset. Trained and tested in the TensorFlow deep learning library, the model is available in the interface generated in LabVIEW.

2.3 Faster R-CNN Inception V2

Faster R-CNN [11] is composed of pre-trained Convolutional Neural Networks (CNN), Region Proposal Network (RPN), RoI pooling layer and a fully connected layer (FC), and its structure is shown in Figure 4. It uses the Inception V2 model as a CNN structure and extracts feature maps from the input image. The feature maps are transferred to RoI pooling layer, and to the RPN for regional proposal. In RPN, regions that may contain objects are provided to the RoI pooling layer by using three two-dimensional anchors of different shapes and sizes. In the ROI pooling layer, areas adjusted to the same size are allocated to the fully connected layer, and object classification and location detection are performed.

The Google team developer the Inception V2 model to reduce computational complexity of CNN and improve its performance. On an input image, the Inception V2 model performs multiple convolution operations in parallel. Features obtained by using different filters such as 1x1 3x3 5x5 are combined into one output and transferred to the next layer.

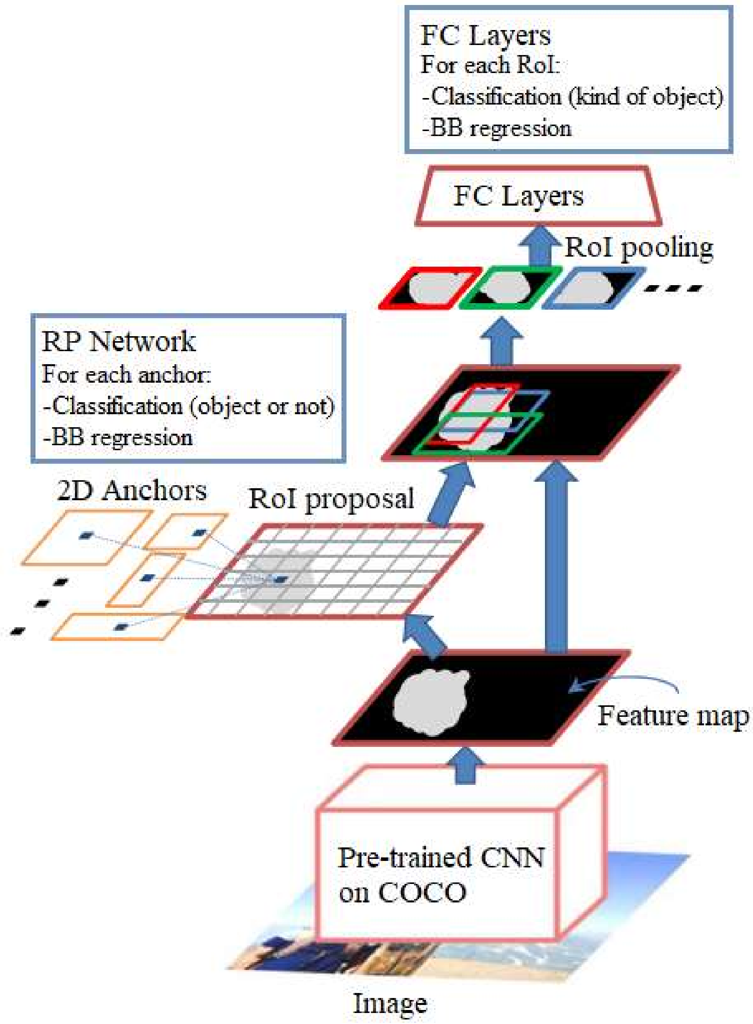
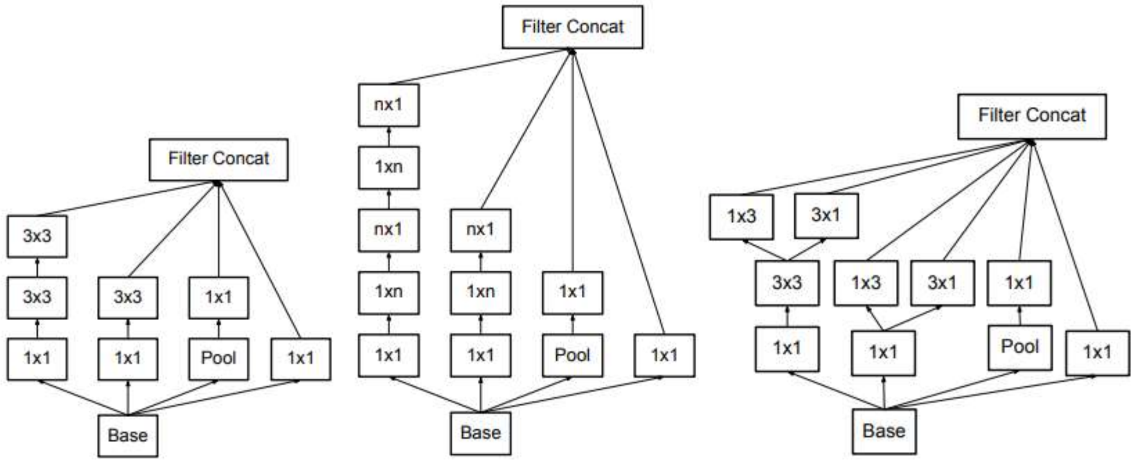


Figure 4. Faster R-CNN [12].

Figure 5 shows three modules of the Inception V2 model that are developed to reduce computational burden and improve performance.



a) A module b) B module c) C module

Figure 5. The modules used for Inception V2 network [13].

The architecture of the Inception V2 network is given in Table 1.

Table 1. The architecture of Inception V2 [13].

Type	**Patch size/Stride**	**İnput size**
Conv	3x3/2	299x299x3
Conv	3x3/1	149x149x32
Conv	3x3/1	147x147x32
Pool	3x3/2	147x147x64
Conv	3x3/1	73x73x64
Conv	3x3/2	71x71x80
Conv	3x3/1	35x35x192
3xInception	As in module A	35x35x288
5xInception	As in module B	17x17x768
2xInception	As in module C	8x8x1280
Pool	8x8	8x8x2048
Linear	Logits	1x1x2048
Softmax	Classifier	1x1x1000

3. Experiment

In this study, network training and testing are conducted with TensorFlow and OpenCV libraries using Python programming language in the Google CoLab environment. The pre-trained Faster-CNN Inception V2 model is built on the TensorFlow Object Detection API [14] and retrained with 1256 images. Figure 6 shows the basic flowchart used in training the model.

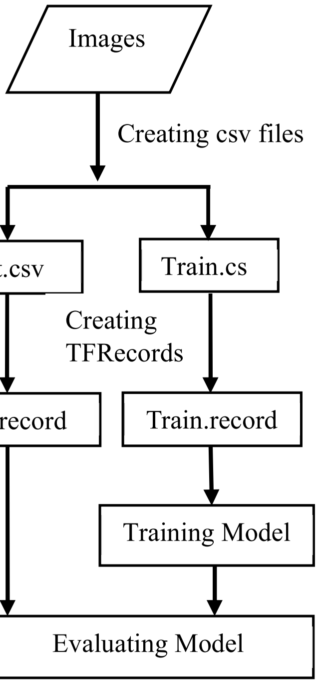


Figure 6. The flowchart of training of Faster R-CNN Inception V2 model.

The LabelImg program is used for labelling masked and unmasked images in the dataset. The XML extension file obtained as program output is later transformed into a CVS file containing information that belongs to each object. The CSV files are converted into a data format designed for TensorFlow, and the network training and testing operations are carried out.

Faster R-CNN Inception V2 learning value, batch size value, and optimization method momentum are selected as 0.002,1, and 0.9 respectively, and the classifier is selected as Softmax, and the training was performed. Figure 7 shows the error variation graph of the model trained with 12000 epoch.

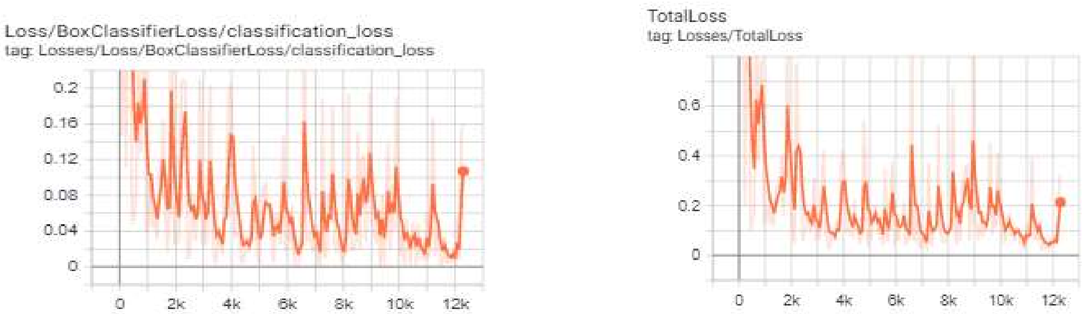


Figure 7. The error variation graph of the model trained with 12000 epoch.

The TensorFlow models make use of confusion matrix in the evaluation. In the confusion matrix shown in Figure 8, TP, FP, TN and FN respectively obtain the numbers of true positive objects, false positive objects, true negative objects and false negative objects.

		Predicted Class	
		Positive (P)	Negative (N)
Actual Class	Positive (P)	True Positive (TP)	False Negative (FN)
	Negative (N)	False Positive (FP)	True Negative (TN)

Figure 8. The confusion matrix used in the evaluation.

The calculation of the metric values resulting from the confusion matrix is given in Equation 1,2,3,4.

Accuracy:

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

Specificity:

$$\frac{TN}{TN+FP} \tag{2}$$

Recall:

$$\frac{TP}{TP+FN} \tag{3}$$

Precision:

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

The weight gained during the training of this model, which provides 85% accuracy, is transferred to the LabVIEW environment. An interface is created in LabVIEW, using Machine Vision and its submodule Deep Learning module. With this interface, the mask detection study could be used more easily by users. Figure 9 shows the block panel generated for face mask detection research through deep learning.

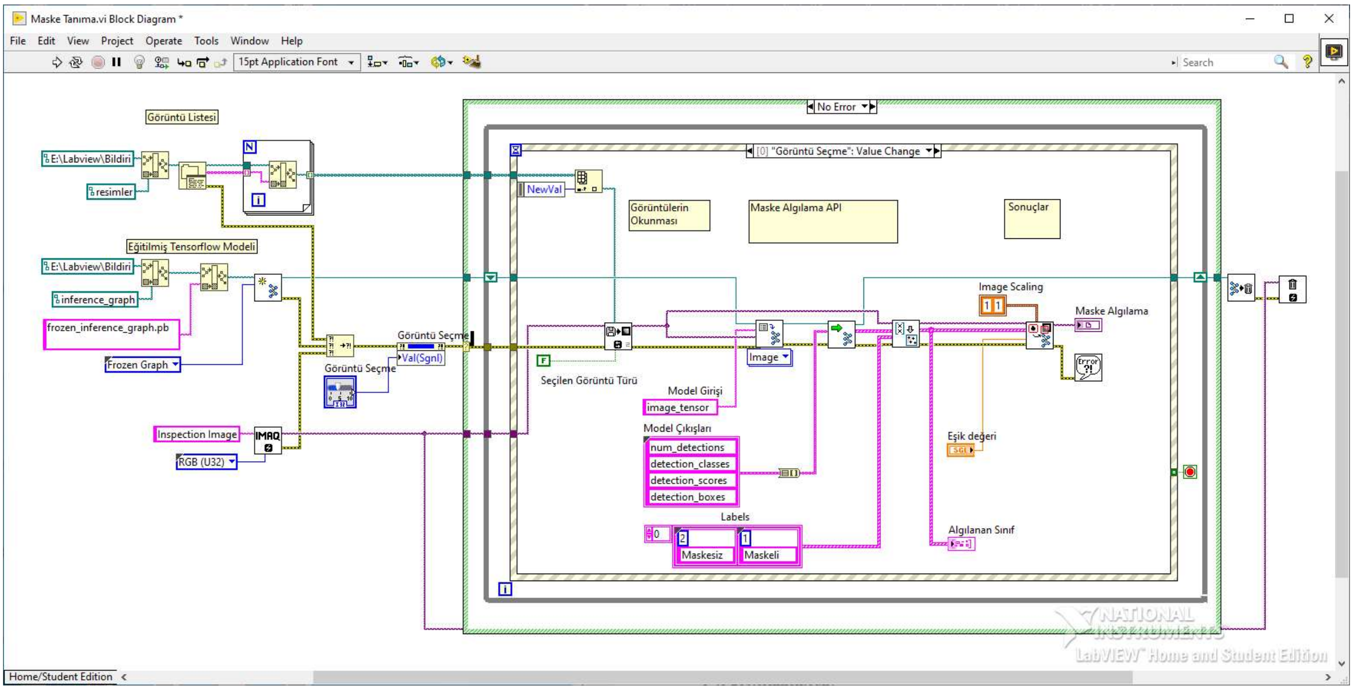


Figure 9. Block panel of mask detection.

In Figure 10 and Figure 11, the detection of whether a person in the image is with a mask and whether multiple people in the image are wearing masks is performed respectively.

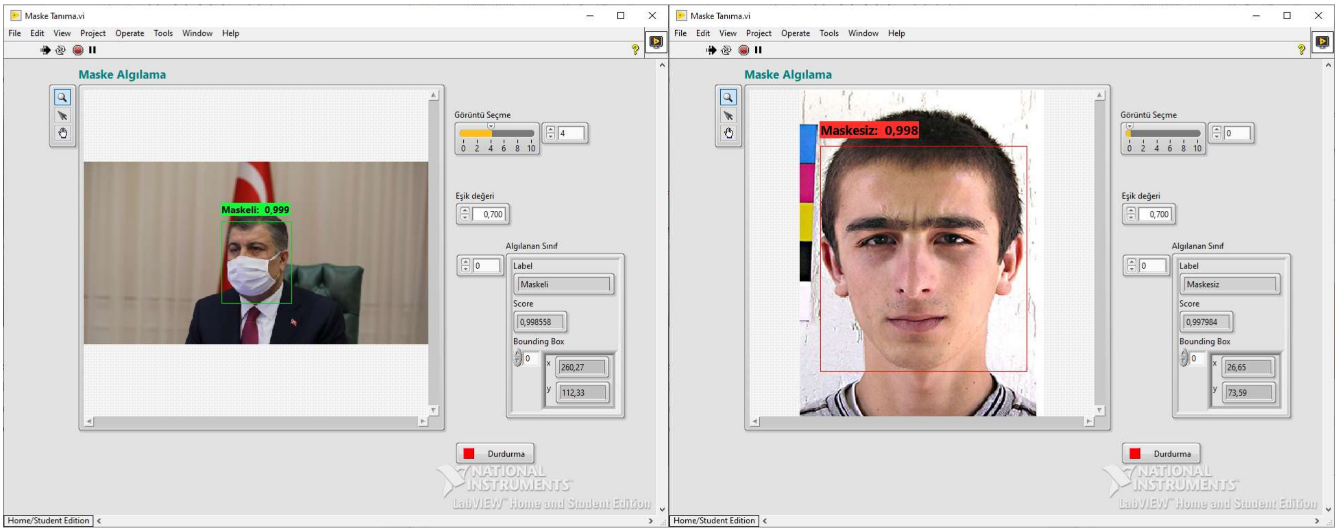


Figure 10. Mask detection for one person.

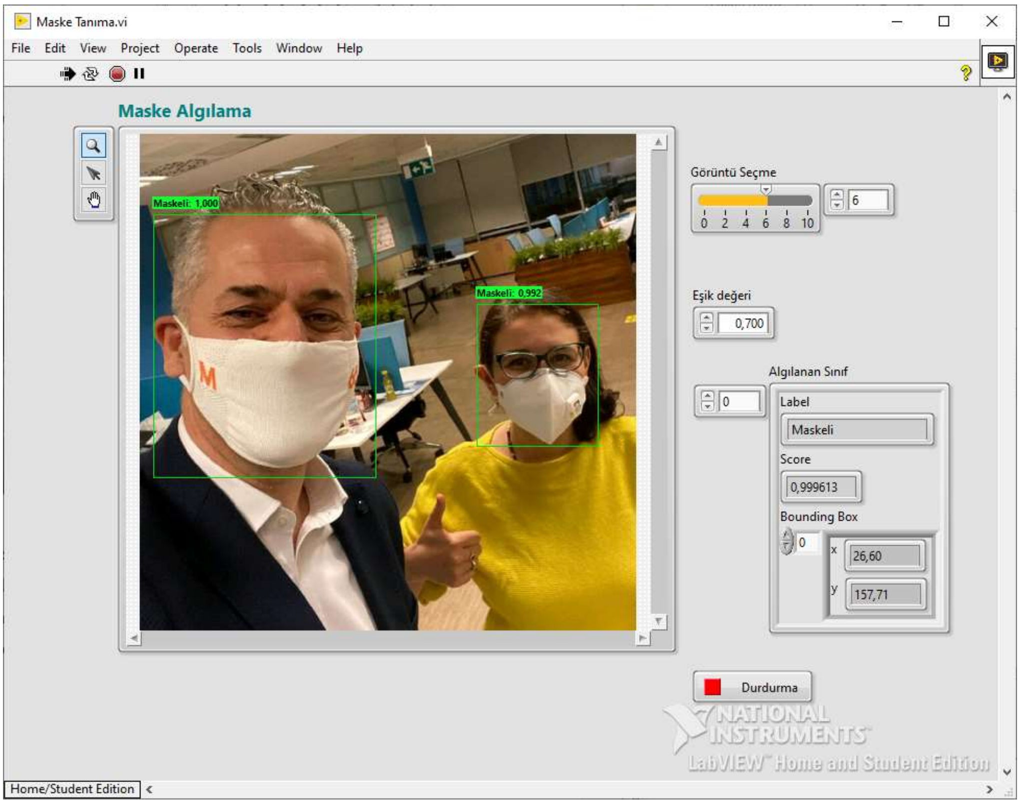


Figure 11. Mask detection on images of more than one person.

The dataset images we used in our study show people of different peoples and genders. Since the desired accuracy value cannot be obtained in the training of a network with 1570 images, the transfer learning method was used. Thus, better results have been obtained in classification by training the network with less data. In addition, masks in images have been created in the virtual environment, so there are not actually people wearing masks. After the training and testing processes of the network, an accuracy value close to the studies given in Table 2 has been obtained in classifying whether people wear masks in daily life.

Table 2. Performance comparison of different methods.

Reference	Methodology	Classification	Result Accuracy (%)
[2]	SRCNet	No facemask-wearing Incorrect facemask-wearing Correct facemask-wearing	98.00
[3]	Decision Trees with ResNet50 Support Vector Machine with RseNet50 Ensemble with ResNet50	With mask Without mask	**<u>100.00</u>**
[4]	YOLO V2 with ResNet50	-	81.00
[8]	Principal Component Analysis	Masked face Non-masked face	96.25

4. Conclusion

Facing a huge health crisis because of COVID-19, many countries in the world make it mandatory to wear masks in crowded and public places in order to slow the spread of Coronavirus. In this study, Faster R-CNN Inception V2 deep learning model is used to detect people who are masked and those who are without masks. The deep learning model previously trained in the COCO dataset is retrained on the SMFD dataset using the transfer learning method. As the outcomes of the face mask detection operated on the interface of LabVIEW, green bounding boxes and red bounding boxes are separately created around the faces of the people who wear masks and those who do not, and accuracy values and ‘masked’ and ‘without mask’ labels are displayed. Using this developed system, it is possible to detect whether one or more people in an image are wearing a mask and provide a safe environment by preventing violations of the mask- wearing rules.

mAP (Mean Average Precision) value is a metric to measure the accuracy of single-stage and two-stage object detection models like SSD and Faster R-CNN.

In this study, 85% accuracy rate (mAP) is achieved. For further research, a larger dataset or data sets can be used. Different studies could be conducted on classifying the masks that people wear or identifying people who wear masks.

References

[1] Paris tests face-mask recognition software on metro riders [online], Available: http://bloomberg.com/

[2] Qin, B., & Li, D. (2020). Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19. Sensors, 20(18), 5236.

[3] Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2020). 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.

[4] Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2020). 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.

[5] Li, C., Wang, R., Li, J., & Fei, L. (2020). Face detection based on YOLOv3. In Recent Trends in Intelligent Computing, Communication and Devices (pp. 277-284). Springer, Singapore.

[6] Din, N. U., Javed, K., Bae, S., & Yi, J. (2020). A novel GAN-based network for unmasking of masked face. IEEE Access, 8, 44276-44287.

[7] Nieto-Rodríguez, A., Mucientes, M., & Brea, V. M. (2015, June). System for medical mask detection in the operating room through facial attributes. In Iberian Conference on Pattern Recognition and Image Analysis (pp. 138-145). Springer, Cham.

[8] Ejaz, M. S., Islam, M. R., Sifatullah, M., & Sarker, A. (2019, May). Implementation of principal component analysis on masked and non-masked face recognition. In 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT) (pp. 1-5). IEEE.

[9] URL,"https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a", Date of access: 12/31/2020.

[10] Dataset, https://github.com/prajnasb/observations.

[11] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28, 91-99.

[12] Hachiya, H., Saito, Y., Iteya, K., Nomura, M., & Nakamura, T. (2018, October). 2.5 d faster r-cnn for distance estimation. In 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 3999-4004). IEEE.

[13] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

[14] URL,"https://www.tensorflow.org/overview/,"TensorFlow Overview, Date of access: 12/27/2020.

[15] URL, https://forums.ni.com/t5/Example-Code/Deep-Learning-Object-Detection-using-Vision-Development-Module/ta-p/3842985?profile.language=en, Date of access: 11/20/2020.