

---

## Research and Applications

# Transfer learning for mobile real-time face mask detection and localization

Francesco Mercaldo<sup>1,2</sup> and Antonella Santone<sup>1</sup>

<sup>1</sup>Department of Medicine and Health Sciences “Vincenzo Tiberio,” University of Molise, Campobasso, Italy, and <sup>2</sup>Institute for Informatics and Telematics, National Research Council of Italy, Pisa, Italy

Corresponding Author: Francesco Mercaldo, Institute for Informatics and Telematics, National Research Council of Italy, Pisa, Italy; francesco.mercaldo@unimol.it, francesco.mercaldo@iit.cnr.it

Received 15 January 2021; Revised 25 February 2021; Editorial Decision 4 March 2021; Accepted 5 March 2021

### ABSTRACT

**Objective:** Due to the COVID-19 pandemic, our daily habits have suddenly changed. Gatherings are forbidden and, even when it is possible to leave the home for health or work reasons, it is necessary to wear a face mask to reduce the possibility of contagion. In this context, it is crucial to detect violations by people who do not wear a face mask.

**Materials and Methods:** For these reasons, in this article, we introduce a method aimed to automatically detect whether people are wearing a face mask. We design a transfer learning approach by exploiting the Mobile-NetV2 model to identify face mask violations in images/video streams. Moreover, the proposed approach is able to localize the area related to the face mask detection with relative probability.

**Results:** To assess the effectiveness of the proposed approach, we evaluate a dataset composed of 4095 images related to people wearing and not wearing face masks, obtaining an accuracy of 0.98 in face mask detection.

**Discussion and Conclusion:** The experimental analysis shows that the proposed method can be successfully exploited for face mask violation detection. Moreover, we highlight that it is working also on device with limited computational capability and it is able to process in real time images and video streams, making our proposal applicable in the real world.

**Key words:** face mask, deep learning, artificial intelligence

---

## INTRODUCTION

The severe acute respiratory syndrome Coronavirus-2 (SARS-CoV-2) is the name given to the new coronavirus discovered in 2019.<sup>1</sup> COVID-19 is the name given to the disease associated with this new kind of virus. SARS-CoV-2 is a new coronavirus strain that has not previously been identified in humans.

Some coronaviruses can be transmitted from person to person, usually after close contact with an infected patient, such as between family members or in a healthcare setting.<sup>2</sup>

The new coronavirus, responsible for the COVID-19 respiratory disease, can also be transmitted from person to person through close contact with a probable or confirmed case.

Current evidence suggests that SARS-CoV-2 spreads from person to person:

- directly;
- indirectly (through contaminated objects or surfaces);
- by close contact with infected persons through secretions from the mouth and nose (saliva, respiratory secretions, or droplets).

When a sick person coughs, sneezes, talks or sings, these secretions are released from the mouth or nose. People who are in close contact (less than 1 meter) with an infected person can become infected if the droplets enter the mouth, nose, or eyes.<sup>3</sup>

Preventive measures are: 1) maintain a physical distance of at least 1 meter, 2) wash your hands frequently, and 3) wear a mask.

Sick people can release infected droplets on objects and surfaces (called fomites) when they sneeze, cough, or touch surfaces (tables, handles, handrails). By touching these objects or surfaces, other people can become infected by touching their eyes, nose or mouth with contaminated (not yet washed) hands.<sup>4</sup>

This is why it is essential to wash hands<sup>5</sup> properly and regularly with soap and water or alcohol-based product and to clean surfaces frequently.<sup>6</sup>

Moreover, to avoid the spread of the pandemic, it is mandatory for people to always wear face masks.<sup>7</sup> These must be worn in indoor places other than private houses and also in all outdoor places, except in cases where, due to the characteristics of the place or the factual circumstances, the condition of isolation is continuously guaranteed.<sup>8</sup>

With the aim of ensuring the safety of the people and places we daily frequent in the COVID-19 pandemic, in this article we present a method to automatically detect whether people are or are not wearing face masks.

The main aim of the proposed approach is real-time detection (from both video and/or image streams) of face mask use or nonuse. To this purpose, we exploit deep learning techniques; in particular, transfer learning is considered to build an accurate model to detect people who are not wearing face masks (even if there are multiple people in the image). It will also localize people within the image and/or video stream, associating a detection accuracy percentage for each individual person detected.

Below are the distinctive points of the proposed approach:

- This approach will automatically and silently detect whether people are wearing face masks;
- We propose a method aimed to understand the reason behind the classifier prediction, making the proposed method explainable, that is, by automatically detecting and drawing a bounding box around the area of interest (thus providing the analyst the area of the image that brought the model to output a certain prediction);
- We used transfer learning, more specifically, we based the proposed architecture on top of the MobileNetV2 network to efficiently work on a wide variety of devices with limited resources (eg, smartphones, tablets, Google Coral, and Raspberry Pi devices);
- The proposed method runs completely in real time, working on both images and live video streams, making it easily implemented in devices such as surveillance cameras;
- We evaluated an extended dataset composed of 4095 images (2165 of people wearing face masks and 1930 of people not wearing face masks);
- The proposed approach achieves an accuracy of 0.98.

The article continues in the following way: in the Face Mask Detection Method section, we present the proposed approach for mobile real-time face mask detection; in the Experimental Analysis section, we present the study we conducted to assess the effectiveness of the proposed method; in the Discussion section, we present the state-of-the-art literature in the face mask detection context; and, finally, in the last sections, conclusions and future work are discussed.

## FACE MASK DETECTION METHOD

In this section we present the proposed method for real-time face mask detection. Our approach is based on transfer learning, an

artificial intelligence technique that serves to adapt an artificial intelligence to a task other than the one for which it was initially trained. The fundamental knowledge learned by an artificial intelligence in a given domain can be directly reapplied to another domain by simply “retuning,” thus avoiding retraining it from scratch.

In this article we experiment with the MobileNetV2<sup>9</sup> model (ie, a deep convolutional neural network composed of 53 layers). To train this network, researchers employed the ImageNet<sup>10</sup> database composed of more than 1 million images. It is able to classify images into 1000 object categories (eg, keyboard, mouse, pencil, animals, etc). As a result, the network has learned feature-rich representations for a wide range of images. The network has an image input size of 224x224.

In [Figure 1](#) the workflow of the proposed approach is depicted.

As shown from [Figure 1](#) the method we propose relies on 2 main steps: 1) training, to generate a model, and 2) validation, to evaluate the model obtained in the previous step.

For model generation we firstly consider a Face Mask dataset, composed of several images representing some people wearing face masks and other not wearing face masks. We manually inspected the images belonging to the dataset in order to accurately annotate each image with a “mask” or “no mask” label. This is an important task, since labeling errors lead to noise in the dataset, which is then reflected in a decay of performance in the proposed approach.

Once a dataset composed by an adequate amount of data is obtained, in the Model Building, we consider a deep learning network designed by the authors, exploiting architecture proposed by the MobileNetV2 network of Google. MobileNetV2 is a model designed to be performed primarily on mobile and low capability devices (eg, Raspberry Pi) to ensure portability and speed of execution—at the expense, however, of the general accuracy in phase of detection.<sup>11</sup> Basically, this is a neural network aimed at image classification, but, with the application of the Single Shot Multibox Detector (SSD) detector, it was converted to the object Detection task. The architecture of this network is based on the same one as the VGG-16 network,<sup>12</sup> but removing the fully connected layers. The reasons this network was used as a basis are its excellent performance in image classification and its previous success in problems in where the transfer learning technique helped improve results.<sup>13</sup> Instead of fully connected layers, a set of auxiliary convolutional layers has been implemented in order to extract the feature for multiple scales and progressively decrease the size of the input for each following layer.

The snippet in [Figure 2](#) shows a Python pseudocode showing the proposed network.

In row 2 of the snippet in [Figure 2](#) we load the base model (ie, the MobileNetV2 network with pretrained ImageNet weights). ImageNet is a large database of images created for use in the field of computer vision, specifically, object recognition. The dataset consists of more than 14 million images that have been manually annotated with the indication of the objects they represent and the bounding box that delimits them. This is 1 of the advantages of using transfer learning: that is, to “inherit” a trained network with a very large dataset of generic images to create a specialized model on a more specific task. In our case, the generic task (performed by the MobileNetV2 network) represents the generic object detection from image, while the specific one is the detection of people wearing (or not wearing) face masks.

Rows from the 6 to 11 are related to the layers we added for the specialized task. In detail we consider following layers:

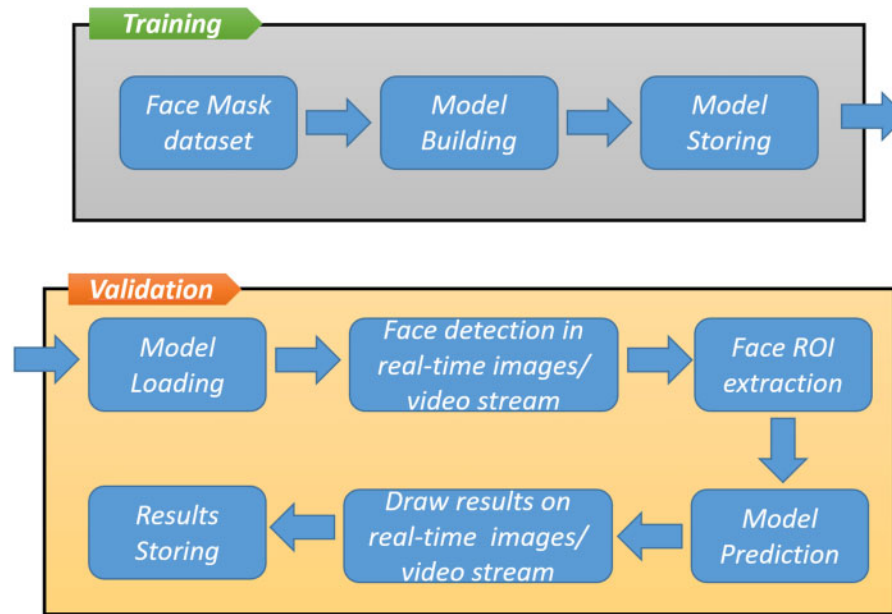


Figure 1. The workflow of proposed approach.

```

1 #loading the MobileNetV2 model
2 baseModel = MobileNetV2(weights="imagenet", include_top=False,
3   input_tensor=Input(shape=(224, 224, 3)))
4 #building the head of the model and place it on top on the
5 #MobileNetV2 model
6 headModel = baseModel.output
7 headModel = AveragePooling2D(pool_size=(7, 7))(headModel)
8 headModel = Flatten(name="flatten")(headModel)
9 headModel = Dense(128, activation="relu")(headModel)
10 headModel = Dropout(0.5)(headModel)
11 headModel = Dense(2, activation="softmax")(headModel)
12 # place the model on top of the MobileNetV2 model
13 model = Model(inputs=baseModel.input, outputs=headModel)
  
```

Figure 2. Python pseudocode of the proposed network.

- AveragePooling2D: pooling is basically the task related to “downscaling” the image obtained from the previous layers. It can be compared to shrinking an image to reduce its pixel density;
- Flatten: reshape the tensor to have a shape that is equal to the number of elements contained in the tensor;
- Dense: represents a normal layer consisting of  $n$  neurons (in this case from 128), in practice the classic scheme of the artificial neural network in which the inputs are weighed and, together with the bias, are transferred through the activation function to the output;
- Dropout: as a regularizer which randomly sets half of the activations to the fully connected layers to zero during training. It has improved the generalization ability and largely prevents overfitting;
- Dense: in this case, we consider as a final layer a Dense layer with 2 neurons (1 for the “mask” prediction and the second 1 for the “no mask” prediction).

With the last row in the snippet shown in Figure 2, we place the layers we added on the top of the MobileNetV2 model. In this way

we consider the MobileNetV2 network (with its training on the ImageNet dataset) with the additional 5 layers we described above, aimed to make a binary prediction (as shown from the last Dense layer with 2 neurons).

Once generated, we store the model we obtained using architecture we designed. This model represents the knowledge of the proposed approach for face mask detection and localization. The storing of the generated model terminates the Training step.

Table 1 shows the details for the layers we added in terms of output shape and parameters.

Once the model is stored, we test the model effectiveness in the Validation step. The model is loaded into the memory and, subsequently, we detect whether in the images/video streams there is/are faces (ie, face detection in real-time images/video streams). If the model found a face, it draws the region of interest (ROI) (face ROI extraction) around the face. In this way the model can focus only on the part of the image/video under analysis, thereby avoiding the rest of the image/video. We highlight that the proposed model is able to

**Table 1.** Model architecture in numbers

Type	Output Shape	Parameters
Average Pooling2D	(1, 1, 1280)	0
Flatten	(None, 1280)	0
Dense	(None, 128)	163 968
Dropout	(None, 128)	0
Dense	(None, 2)	258

detect even more faces contextually present in the same image/video stream. Once we marked only the parts of the image related to faces, these parts are then input to the model that will output a prediction for each face image (Model Prediction in Figure 1). The model outputs a certain prediction (ie, “mask” or “no mask”) with a certain probability (from 0 to 100%). Thus, we draw on the input images/video stream the ROI, the label, and the prediction probability (to understand the degree of confidence with which the model has predicted a certain label) and the results (ie, the images/video with ROI, label and probability prediction) are stored.

We consider the proposed method successful for its ability to automatically draw the ROI in the image (ie, by using the bounding box). This approach makes it possible to visualize the area of the image under analysis (and responsible for a certain prediction) in order to both evaluate the effectiveness of the proposed model (to understand whether it is correctly considering a person in the image under analysis) and automatically identify a subject who may not be wearing a mask.

## EXPERIMENTAL ANALYSIS

In this section we describe the experiment we conducted in order to demonstrate the effectiveness of the proposed method.

We obtained a dataset composed of 2165 images representing people with face masks and 1930 images of people without face masks for a total of 4095 images. The images were obtained from the RMFD dataset (<https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset>) and from a Kaggle repository (<https://www.kaggle.com/prithwirajmitra/covid-face-mask-detection-dataset>).

From the implementation point of view, we used: Python programming language, TensorFlow (<https://www.tensorflow.org/>), the Google library for artificial intelligence experiments providing a plethora of supervised and unsupervised algorithms; and Keras, a library for neural network management. Keras works as an interface at a higher level of abstraction than other similar lower-level libraries

and supports TensorFlow as a back end. The machine used to run the experiments and to take measurements was an Intel Core i7 8th gen, equipped with 2GPU and 16Gb of RAM; and we used the Microsoft Windows 10 operating system. For research purposes, the source code developed by the authors is freely available (<https://mega.nz/file/AM93lKhA#nOryc32RZV1oYjAj9hnTPp0Lv1vME-frigbl3K-NDulw>), providing also the model generated by the proposed deep learning network.

Figures 3 and 4 show 2 examples of the detection different from that provided by the proposed approach.

As shown from the example in Figure 3, the proposed approach is able to correctly detect different types of face masks; in fact, in Figure 3, both the masks (ie, the white and the black ones) are detected with a probability equal to 100%. The aim of this example is to demonstrate that the proposed method is resilient to the type and color of the face mask (in fact also the face mask is different with respect to the face masks in Figure 3).

The second example in Figure 4 demonstrates that the proposed approach is resilient to facial expressions. We consider an image related to a woman with 6 different facial expressions and the related facial expressions with the face mask, where the proposed approach successfully detects all the sub images where the face mask is worn.

In Figure 5 another example of detection is shown.

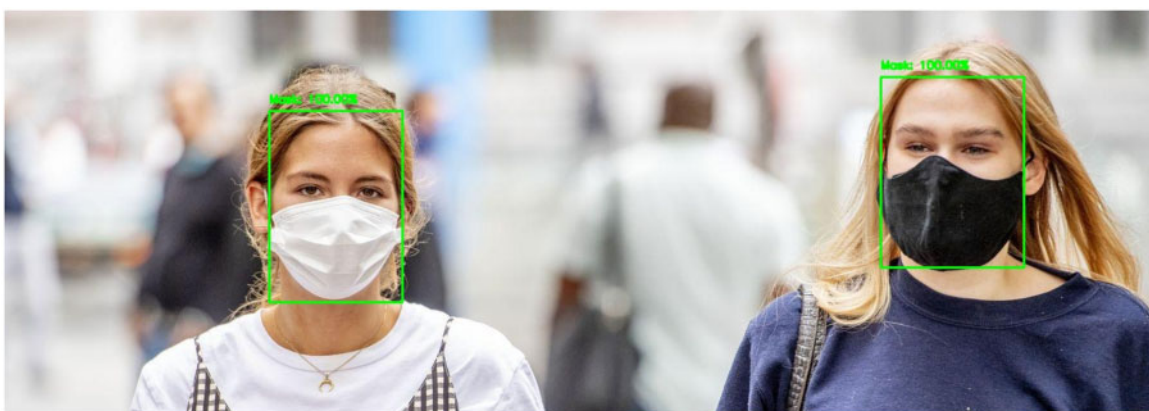
Differently from the detection example shown in Figure 4, in Figure 5 we show the detection obtained from images of different people wearing different face masks. It is interesting to highlight that even when people are using face masks with different colors, the proposed method is able to detect the face mask correctly.

In Figure 6 we show an example of a frame obtained from a video stream, demonstrating that the proposed method can effectively be embedded into video surveillance cams.

As seen in Figure 6, the proposed method is able to detect all three people wearing face masks. In particular, the proposed method is even able to detect the girl on the left with her face slightly lowered. In the frame there are also several people not wearing face masks: As evidenced from the red bounding box in Figure 6, the proposed method is able to rightly identify infractions.

To assess the effectiveness of the proposed approach, we take into account the accuracy and the loss metrics.

The accuracy is defined as closeness degree related to the measurements of a quantity to that quantity’s true value: It is basically the fraction of the predictions that are correct, and it is computed as the sum of true positives and negatives divided by all the evaluated images:

**Figure 3.** An example of detection.

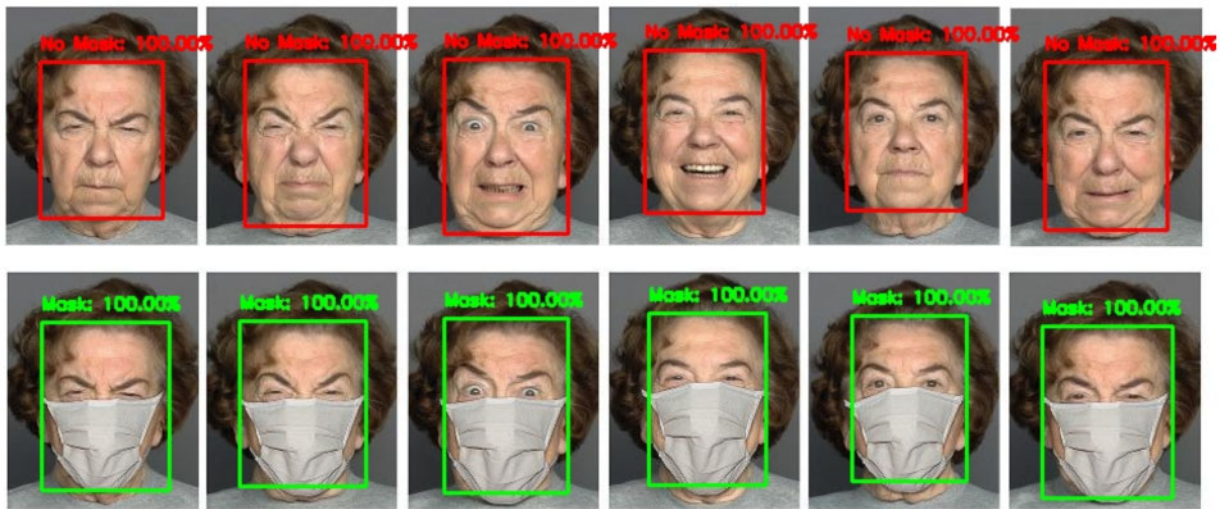


Figure 4. A second example of detection.

$$\text{Accuracy} = \frac{\text{tp} + \text{tn}}{\text{tp} + \text{fn} + \text{fp} + \text{tn}}$$

where tp indicates the number of true positives, tn indicates the number of true negatives, fn indicates the number of false negatives, fp indicates the number of false positives.

The loss metric represents a quantitative measure of how much the predictions differ from the assigned label. For definition, it is inversely proportional to the model correctness.

The loss interpretation is how well the model is doing for the training and validation sets: It is basically a summation of the errors made for each image in training or validation sets.

In a nutshell, the supervised learning of a neural network is done like any other machine learning: A training dataset is presented to the network, the network output is compared with the desired output, an error vector is generated, and corrections are applied to the network based on it, usually using a back propagation algorithm. The groups of training data that are treated together before applying

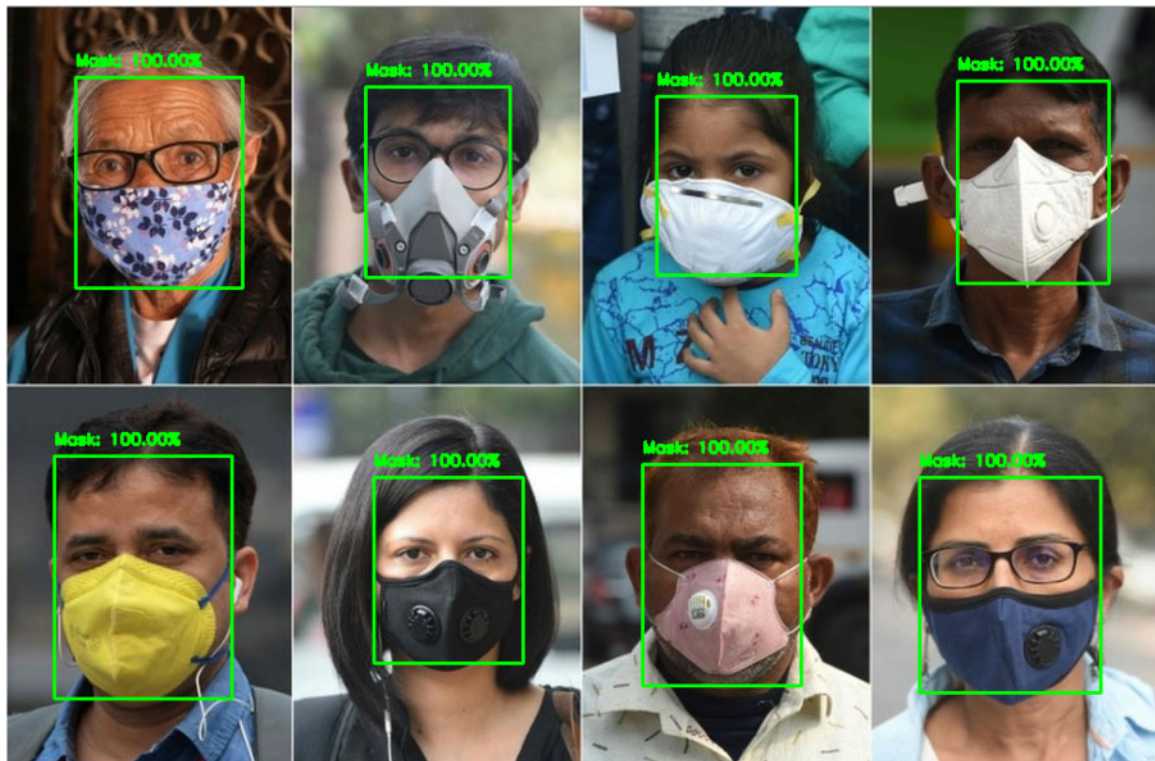
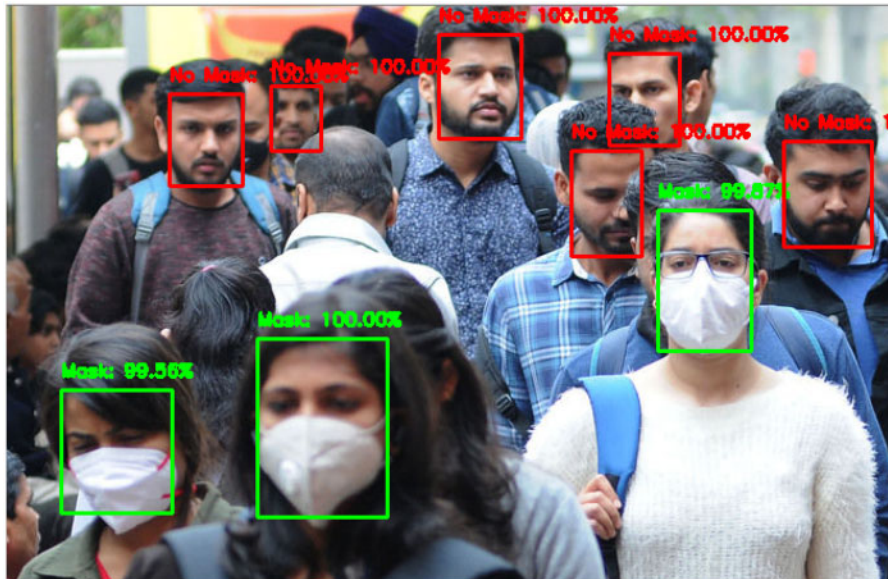


Figure 5. A third example of detection.



**Figure 6.** A fourth example of detection.

the corrections are called epoch (epochs). We set the epoch number equal to 20. In [Figure 7](#), there are the accuracy and the loss trends for the training and for the evaluation step.

[Table 2](#) shows the results we obtained for each epoch: with the label train loss, we indicate the loss values for the training; with the val loss, the loss values for the evaluation; with train acc, the accuracy values for the training step; and with val acc, the values of the accuracy for the evaluation step.

In the last epoch, that is, the twentieth one, with regard to the evaluation task, the accuracy we obtain is equal to 0.98 and the loss equal to 0.03.

In [Figure 7](#), we show the trends for accuracy and loss and for the training and the evaluation steps. On the x axis, there is the number of considered epochs (from 0 to 20), while on the y axis we represent the accuracy and loss values for a certain loss (ranging from 0 to 1). From an ideal point of view, we expected an accuracy equal to 1 and a loss equal to 0.

As shown in [Figure 7](#), the accuracy trends are increasing while the loss trend is decreasing. This is symptomatic that, along the 20 epochs, the network is learning the distinctive features of mask and no mask images.

We also provide details about the time performance analysis (ie, the time required by the proposed method to generate the mask/no mask detection with the related green/red bounding box). On the machine considered for the experimental analysis, on average, it took the proposed method 4.7 seconds to process a never-seen image.

## DISCUSSION

The use of a face mask has become mandatory, given the COVID-19 pandemic. Consequently, researchers have begun to study ways to automatically detect mask-wearing violations; but, at least at the time of this writing, there are still few contributions in this area. In this section, we report the efforts produced by the research community so far and highlight the novelty of our proposed contribution.

For instance, Loey et al exploit artificial intelligence, in particular the ResNet50 deep model in combination with the support vector machine learning classifier, to predict if people in an image under analysis

are wearing face masks or not.<sup>14</sup> The difference between this work and our proposal is represented by the model adopted: as a matter of fact, we consider a transfer learning approach based on the MobileNetV2 model for fake mask detection from images and video streams, making our method able to run on devices with limited resources, for example, mobile devices and webcams. Thus, the proposed method is applicable to both types of devices.

Researchers Chowdary et al<sup>15</sup> experimented with the InceptionV3 deep learning model for face mask detection. The effectiveness of this method is evaluated using a dataset composed of 1570 images (785 of people wearing face masks and 785 images of people without face masks), while we evaluate the proposed approach with 4095 different images. Moreover, the proposed approach is able to generate the prediction on mobile and embedded platforms.

Chen et al<sup>16</sup> propose the adoption of machine learning techniques, in particular the K-nearest neighbors supervised classification algorithm for the discrimination between people wearing and not wearing face masks. They reach an accuracy of 0.87, while the proposed method obtains an accuracy of 0.98.



**Figure 7.** Experimental analysis results.

**Table 2.** Experimental analysis evaluation

Epoch	train_loss	train_acc	eval_loss	eval_acc
1	0.5491	0.7650	0.1467	0.9756
2	0.1887	0.9498	0.0839	0.9841
3	0.1072	0.9702	0.0636	0.9829
4	0.0783	0.9797	0.0606	0.9853
5	0.0620	0.9851	0.0527	0.9853
6	0.0611	0.9833	0.0513	0.9866
7	0.0670	0.9828	0.0475	0.9866
8	0.0513	0.9879	0.0435	0.9853
9	0.0529	0.9841	0.0432	0.9878
10	0.0555	0.9835	0.0427	0.9866
11	0.0537	0.9852	0.0393	0.9878
12	0.0338	0.9899	0.0409	0.9878
13	0.0467	0.9851	0.0388	0.9890
14	0.0509	0.9825	0.0356	0.9866
15	0.0315	0.9930	0.0374	0.9878
16	0.0374	0.9882	0.0377	0.9902
17	0.0312	0.9914	0.0354	0.9878
18	0.0355	0.9877	0.0382	0.9902
19	0.0314	0.9908	0.0379	0.9902
20	0.0299	0.9911	0.0351	0.9890

## CONCLUSION AND FUTURE WORK

The COVID-19 pandemic has radically changed our habits. While waiting for the spread of the virus to decrease, it is necessary to observe a series of rules, such as the use of face masks in public places. The aim of the following work is to provide a method for the automatic detection of violations of persons not wearing face masks. In detail, our approach relies on the adoption of transfer learning to detect whether there are people wearing/not wearing face masks in images and video streams. The proposed method is able to work also on devices with limited computation capabilities as, such as smartphones and webcams, making the proposed approach actually implementable in a real-world context.

With regard to the future research lines, we plan to reinforce the proposed approach by exploiting a series of transfer learners with the aim of increasing performance. Moreover, we plan to adopt activation maps to highlight areas symptomatic of the detection, making the proposed approach more interpretable.<sup>17,18</sup> That is, while the proposed method is able to draw the bounding box around the people's faces, activation maps can be helpful to detect which part of the face contributes to the detection. In this way it will be possible to achieve finer-grain detection.

## FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## AUTHOR CONTRIBUTIONS

FM and AS contributed to the design of the proposed method, to the experimental analysis and to all the aspects of the article, from draft writing to final approval.

## DATA AVAILABILITY STATEMENT

The Python source code underlying this article is available at: <https://mega.nz/file/AM93IKhA#nOryc32RZV1oYjAj9hnTPp0Lv1vMEfrigl3K-NDulw>

The datasets were derived from sources in the public domain: RMTD, <https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset>, and Kag-

gle, <https://www.kaggle.com/prithwirajmitra/covid-face-mask-detection-dataset>

## CONFLICT OF INTEREST STATEMENT

None declared.

## REFERENCES

- Hu B, Huang S, Yin L. The cytokine storm and covid-19. *J Med Virol* 2021; 93 (1): 250–6.
- Ahmed F, Bukhari SAC, Keshtkar F. A deep learning approach for covid-19 viral pneumonia screening with x-ray im-ages. *Digit Gov: Res Pract* 2021; 2 (2): 1–12.
- Brunese L, Martinelli F, Mercaldo F, San-Tone A. Machine learning for coronavirus covid-19 detection from chest x-rays. *Procedia Comput Sci* 2020; 176: 2212–21.
- Pham TD. Classification of covid-19 chest x-rays with deep learning: new models or fine tuning? *Health Inf Sci Syst* 2021; 9 (1): 2–11.
- Mieth L, Mayer MM, Ho Mann A, Buchner A, Bell R. Do they really wash their hands? Prevalence estimates for personal hygiene behaviour during the covid-19 pandemic based on indirect questions. *BMC Public Health* 2021; 21 (1): 1–8.
- Kavitha M, Jayasankar T, Venkatesh PM, et al. Covid-19 disease diagnosis using smart deep learning techniques. *J Appl Sci Eng* 2021; 24 (3): 271–7.
- Islam SMD, Mondal PK, Ojong N, et al. Water, sanitation, hygiene and waste disposal practices as covid-19 response strategy: insights from Bangladesh. *Environ Dev Sustain* 2021: 1–22.
- Rekha HS, Behera HS, Nayak J, Naik B-N. Deep learning for covid-19 prognosis: a systematic review. In: Sekhar GC, Behera HS, Nayak J, Naik B, Pelusi D. (eds). *Intelligent Computing in Control and Communication*. Singapore: Springer; 2021: 667–87.
- Sandler M, Howard A, Zhu M, Zhmoginov A, and Chen LC. MobileNetV2: inverted residuals and linear bottlenecks. In proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; June 18–23, 2018; Salt Lake City, UT.
- Deng J, Dong W, Socher R, Li-Jia L, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. In proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition; June 20–25, 2009; Miami, FL.
- Nagrath P, Jain R, Madan A, Arora R, Kataria P, Hemanth J. Ssdmv2: a real time DNN-based face mask detection system using single shot multi-box detector and MobileNetV2. *Sustain Cities Soc* 2021; 66: 102692.
- Ulloa C, Dora MB, Renza D. Video forensics: identifying colorized images using deep learning. *Appl Sci* 2021; 11 (2): 476.
- Ferdous RH, Arifeen MM, Tipu Sultan E, Mamun SA. Performance analysis of different loss function in face detection architectures. In Proceedings of International Conference on Trends in Computational and Cognitive Engineering; June 23–25, 2021; Texas.
- Loey M, Manogaran G, Taha MHN, Khalifa NEM. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement (Lond)* 2021; 167: 108288.
- Chowdary GJ, Punn NS, Sonbhadra SK, Agarwal S. Face mask detection using transfer learning of InceptionV3. In proceedings of the International Conference on Big Data Analytics; September 18–20, 2020, Honolulu, HI.
- Yuzhen C, Menghan H, Chunjun H, et al. Face mask assistant: Detection of face mask service stage based on mobile phone. *arXiv Preprint arXiv:2010.06421* 2020.
- Luca B, Mercaldo F, Reginelli A, San-Tone A. Explainable deep learning for pulmonary disease and coronavirus covid-19 detection from x-rays. *Comput Methods Prog Biomed* 2020; 196: 105608.
- Iadarola G, Martinelli F, Mercaldo F, Santone A. Towards an interpretable deep learning model for mobile malware detection and family identification. *Comput Security* 2021; 105: 102198. page