

# Article

# Facial Recognition System for People with and without Face Mask in Times of the COVID-19 Pandemic

Jonathan S. Talahua <sup>1,2</sup>, Jorge Buele <sup>1</sup> , P. Calvopiña <sup>3</sup> and José Varela-Aldás <sup>1,4,\*</sup> 

<sup>1</sup> SISAu Research Group, Facultad de Ingeniería y Tecnologías de la Información y la Comunicación, Universidad Tecnológica Indoamérica, Ambato 180103, Ecuador; jonathan.talahua029@comunidadunir.net (J.S.T.); jorgebuele@indoamerica.edu.ec (J.B.)

<sup>2</sup> Escuela Superior de Ingeniería y Tecnología, Universidad Internacional de La Rioja, 26006 Logroño, Spain

<sup>3</sup> Departamento de Eléctrica y Electrónica, Universidad de las Fuerzas Armadas ESPE, Sangolquí 171103, Ecuador; jpcalvopina1@espe.edu.ec

<sup>4</sup> Department of Electronic Engineering and Communications, University of Zaragoza, 44003 Teruel, Spain

\* Correspondence: josevarela@uti.edu.ec

**Abstract:** In the face of the COVID-19 pandemic, the World Health Organization (WHO) declared the use of a face mask as a mandatory biosafety measure. This has caused problems in current facial recognition systems, motivating the development of this research. This manuscript describes the development of a system for recognizing people, even when they are using a face mask, from photographs. A classification model based on the MobileNetV2 architecture and the OpenCv's face detector is used. Thus, using these stages, it can be identified where the face is and it can be determined whether or not it is wearing a face mask. The FaceNet model is used as a feature extractor and a feedforward multilayer perceptron to perform facial recognition. For training the facial recognition models, a set of observations made up of 13,359 images is generated; 52.9% images with a face mask and 47.1% images without a face mask. The experimental results show that there is an accuracy of 99.65% in determining whether a person is wearing a mask or not. An accuracy of 99.52% is achieved in the facial recognition of 10 people with masks, while for facial recognition without masks, an accuracy of 99.96% is obtained.

**Keywords:** convolutional neural networks; face mask; facial recognition; COVID-19



**Citation:** Talahua, J.S.; Buele, J.; Calvopiña, P.; Varela-Aldás, J. Facial Recognition System for People with and without Face Mask in Times of the COVID-19 Pandemic. *Sustainability* **2021**, *13*, 6900. <https://doi.org/10.3390/su13126900>

Academic Editors:  
Carlo Maria Medaglia,  
Marco Barbanera and  
Guido Marseglia

Received: 30 March 2021  
Accepted: 5 June 2021  
Published: 18 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In recent decades, facial recognition has become the object of research worldwide [1–5]. In addition, with the advancement of technology and the rapid development of artificial intelligence, very significant advances have been made [6,7]. For this reason, public and private companies use facial recognition systems to identify and control the access of people in airports, schools, offices, and other places [8–12]. On the other hand, with the spread of the COVID-19 pandemic, government entities have established several biosafety regulations to limit infections [13–15]. Among them is the mandatory use of face masks in public places, as they have been shown to be effective in protecting users and those around them [16–19].

As the spread of the virus occurs through physical contact, conventional recognition systems (such as fingerprints) or typing a password on a keyboard become insecure. Thus, facial recognition systems are the best option, as they do not require physical interaction as in other cases. However, the use of the face mask within these systems has represented a great challenge for artificial vision [20], because at the time of facial recognition, half of the face is covered and several essential data are lost. This clearly denotes the need to create algorithms that recognize a person when they are wearing a face mask [21]. This has made it necessary to implement new strategies to achieve robustness in the current systems [22].

In this sense, convolutional neural networks (CNN) belong to a set of techniques grouped under the so-called deep learning [23–25]. Thus, over the years, this technology

has been adapted to the needs of the human being, as established in [26,27], developing applications in various fields of knowledge, such as agriculture [28], military area [29], and medicine [30], among others. The contribution of this type of neural network has also been applied to analyze dental images, and this is technically described in the review of [31]. In [32], a system that analyzes medical images is proposed, through selective data sampling, that detects hemorrhages in color images. On the other hand, in [33], a technical review of the contributions of the CNN in the mammographic breast cancer diagnosis (MBCD) is shown. Although there are several related investigations, they are still in the initial stages, with the clear objective of providing robust tools in the future. In [34], a review is described that seeks to identify the chronological advancement of CNN in brain magnetic resonance imaging (MRI) analysis.

Although its use in medicine is not recent, it has now been directed particularly to applications related to COVID-19 [35–37]. Various methods and procedures are used for the diagnosis of this disease, and one of them is the review of computed tomography scans. For this reason, [38] suggests a rapid and valid method based on AI (10 CNN) for the detection of pulmonary affections. The results show a sensitivity, specificity, and precision of over 99%. Similarly, in [39], a system for automatic disease diagnosis using the EfficientNet architecture is described. The results denote an average accuracy of over 96%, validating the contribution in situations of health crisis. Taking X-rays is another way to identify the virus affectations in the patient's chest [40,41]. Given this, in [42] a deep learning method based on nCOVnet networks is presented for detection, the results of which show an accuracy of between 98% and 99%. Something similar is done in [43], where chest X-ray images are analyzed and the various training techniques of the networks are compared, obtaining an accuracy of 98.33% when using ResNet-34.

Although in most cases CNNs are used in the diagnosis of COVID-19, they can also be used in other applications, as part of contagion prevention measures [44]. In [45], a system is presented that allows people to be monitored when entering and being inside a certain place, and to evaluate if they are complying with the established biosecurity measures. In the event that this is not complied with, other people can be informed to exercise caution and health personnel to apply the respective measures. They have also been used to develop detection systems for the proper use of face masks. For this reason, in [46], a system is proposed that differentiates the people who use a mask or not with the algorithms RCNN, Fast RCNN, and Faster RCNN with an accuracy of 93.4%. In [47], the VGG-16 CNN model is used to implement a detection system with an accuracy rate of 96%. Similarly, in [48] they propose the SSDMNv2 model based on the MobileNetV2 architecture, which has an accuracy of 92.64% when performing the experimental tests.

On the other hand, [49] describes a system for the detection of face masks using a support vector machine (SVM) algorithm. The datasets are the Real-World Masked Face Dataset (RMFD), the Simulated Masked Face Dataset (SMFD), and the Labeled Faces in the Wild (LFW). The results show an accuracy of 99.64% with SVM in RMFD, 99.49% in SMFD, and 100% in LFW. In [50], InceptionV3 transfer learning is used, obtaining an accuracy of 99.92% during training and 100% during tests with SMFD data. In [51], a method to identify the correct use of masks is defined by combining classification networks and super-resolution of images (SRCNet). An accuracy of 98.70% is achieved, surpassing conventional image classification methods of this type.

The problem of facial recognition due to the use of face masks during the COVID-19 pandemic has caused new horizons to be explored in artificial intelligence, representing a challenge for researchers, which has motivated the development of ocular recognition systems, as a parallel response. In [52], a facial recognition system using eye information and CNN trained by ImageNet is presented. The results present an accuracy of between 90–95%. Similarly, [53] provides a facial recognition system using SVM with three databases (UBIPr, Color FERET, and Ethnic Ocular). Performance tests show a yield of approximately 92%.



In continuity with the works described in the bibliography, this document presents a facial recognition system for people regardless of whether they use a face mask or not. For this purpose, the work has been organized into four sections, Section 2 contains the materials and methods, Section 3 shows the results, and Section 4 presents the discussion.

## 2. Materials and Methods

### 2.1. Description of the Problem

The effects of COVID-19 on the global economy can be seen with the naked eye, as the confinement of people in the homes brings with it less production and slows down the commercial dynamism. However, it should be noted that in situations of health crisis such as the one that continues to be experienced, it is relevant to put people's health before any productive activity. That is why biosecurity measures and social distancing protocols have been implemented to limit the spread of this dangerous virus. As well as the capacity in public institutions, industries and other establishments has been limited, highlighting the so-called telework (in certain cases). Thus, companies have implemented various methodologies, strategies, and techniques to protect the integrity and health, both when entering and staying in face-to-face work sessions. As previously mentioned, CNN have been an important technological tool during this pandemic. Although most approaches have been taken towards the diagnosis of the disease, monitoring and prevention has also been covered.

Today, the use of a personal face mask is a mandatory preventive measure. Keeping the mouth, nose, and cheeks covered has now made people only recognizable by their eyes, eyebrows, and hair, which is a problem for the human eye, which tends to find similarities in several faces that have similar features. This problem also affects computer systems, as facial recognition systems are now very common. They are used to unlock the smartphone, access sensitive applications, and to enter certain places. Current systems usually process information from the entire face of the person, which is why technology must adapt to these new conditions. All this is done with the purpose of maintaining the biosecurity of the user, but giving them the opportunity to continue with the activities as naturally as possible.

The literature has shown that there are systems that seek to identify whether people use it properly. These works have had very good results. However, facial recognition using biosecurity material has not yet been explored. All of this motivated the present investigation, in which a detection system with two approaches is presented. The first is to develop a face classifier, starting from a database of people with and without a mask. The second describes a facial recognition algorithm in controlled environments, which allows for personnel to be identified automatically, without removing the face mask. This can be implemented as an access system to an institution or a home, but at a low cost. This is ensured by using open-source programming software and simple features that reduce computational expense. For this reason, the possibility of improving the adaptability of current facial recognition systems, in the face of new circumstances, has been established as a starting hypothesis.

### 2.2. Requirements

The programming language used here is Python. For optimal operation, a high-processing equipment (GPU) is needed. However, we received no external financing, so we chose to work with free Google servers, which are available in Google Colab. Another of the essential requirements is to have the necessary databases in order to carry out the training and obtain the classification and recognition models. Taking into account that building the database of these databases requires a high investment of time when working with artificial intelligence and especially with convolutional neural networks.

Additionally, it is necessary to develop a consent form for the people who will allow for taking photographs for the facial recognition algorithm database. This is necessary because there are currently no databases for the recognition of people with face masks. For

this, it is necessary to rely on Art 6.1d of the European General Data Protection Regulation (RGPD), in connection with article 46.3 of the LOU. Here, it is mentioned that the data of a person will be treated in accordance with the exercise of public powers conferred on the universities as responsible for the treatment of the data of the students. As well as biometric data ((article 9.2.a) of the RGPD), in which consent will be needed so that it can be part of the exams where facial recognition techniques are applied. The collection, filing, processing, distribution, and dissemination of these information data will require the authorization of the owner and the mandate of the law.

### 2.3. System Development

It is proposed to design a system that is capable of identifying a person's face, even if it is with or without a mask. For the system to work properly, it is necessary to use two databases: the first is for classifier training and consists of a large number of images of people who wear a face mask and others who do not. The second is used for training the facial recognition system, and here there are people with and without the biosafety material (face mask). The input data are obtained either from an image, or a video and the architecture used is MobileNet, with the aim of having a better precision and robustness. This project is divided into three stages, which are described below.

#### 2.3.1. First Stage

This stage focuses on finding the location and dimension of one or more faces, regardless of whether or not they wear a mask, within an image. For this, the OpenCV Deep Learning-based face detection model is used and, as a result, the region of interest (ROI) is obtained, which contains data such as the location, width, and height of the face.

#### 2.3.2. Second Stage

A diagram of the operation of the second stage is shown in Figure 1. This is where the classifier training is performed to detect faces with a mask and without a mask. For this, the "Real-World-Masked-Face-Dataset" database available on Git-Hub is used. Unzipping the files makes available a large number of images of people of Asian origin wearing a mask. From this database, the training of the classifier of the first stage is carried out.

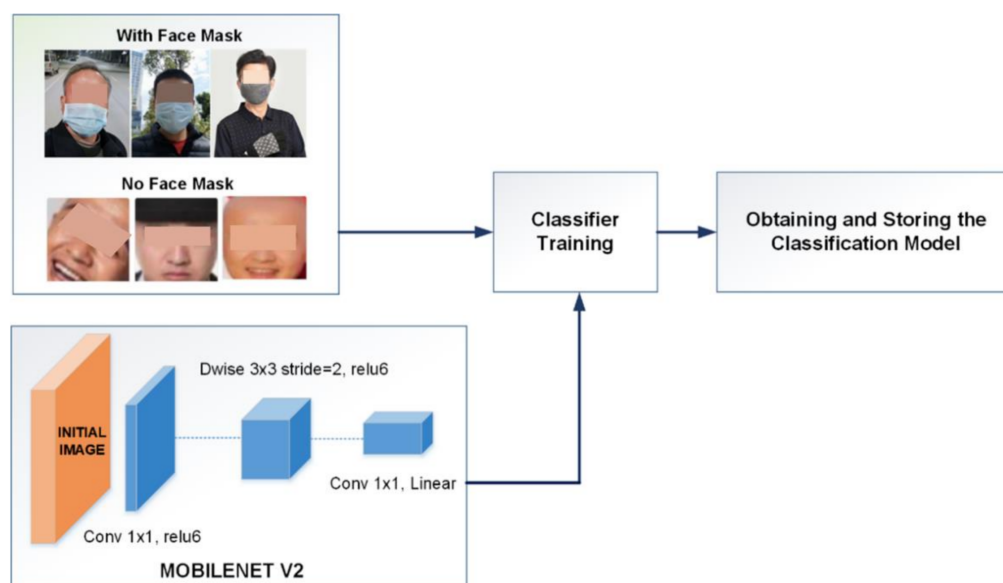
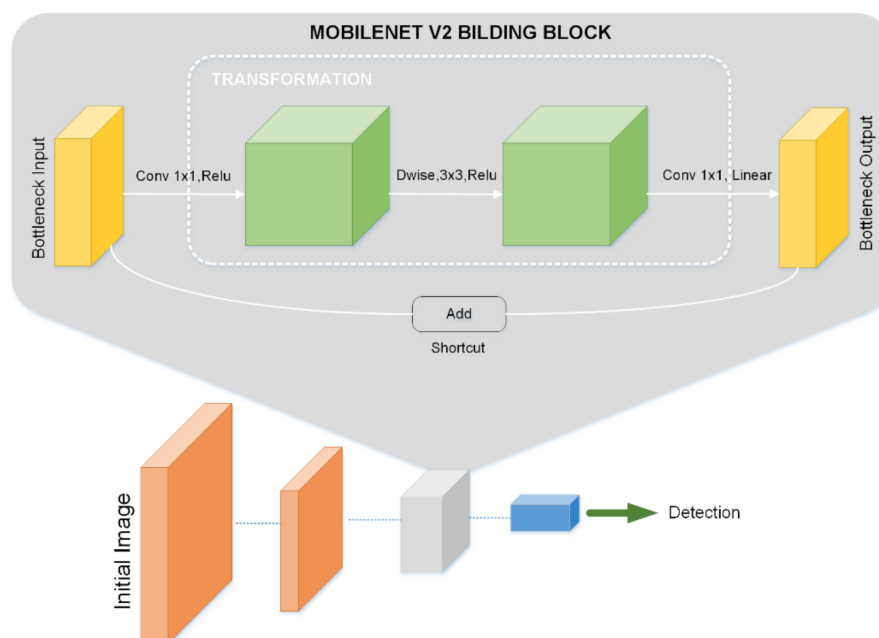


Figure 1. Classifier training block diagram.

For the classifier, MobileNet V2 architecture is used, as it uses smaller models with a low latency and low parameterization power. This improves the performance of mobile models in multiple tasks and benchmarks, resulting in a better accuracy. It also retains

the simplicity and does not require any special operator to classify multiple images and various detection tasks for mobile applications. MobileNetV2 is 35% faster at object detection compared with the first version, when combined with Single Shot Detector Lite. It is also more accurate and requires 30% fewer parameters, as the Bottleneck encodes the intermediate inputs and outputs, as shown in Figure 2.



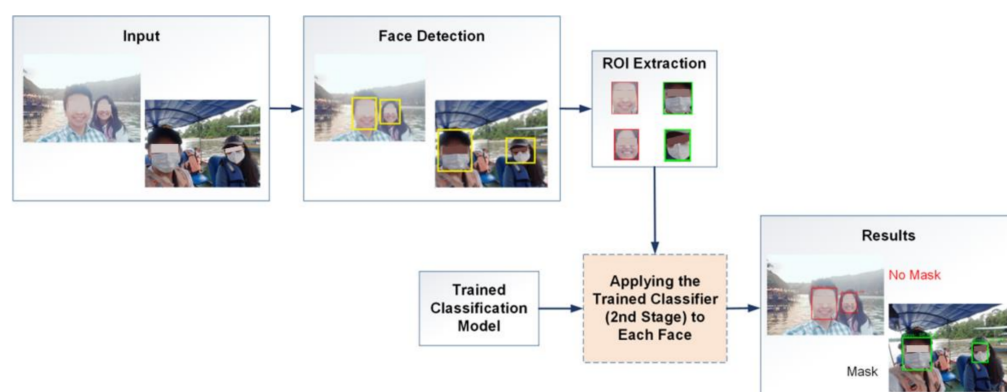
**Figure 2.** Architecture of MobileNetV2.

At the same time, the inner layer encapsulates the model's ability to transform of lower-level concepts (pixels) to higher-level descriptors (image categories). This is available in the "Tensorflow" library in Python, and through "Transfer Learning", a change is made in the last layer of the convolutional neural network. It should be noted that this architecture is selected because it is a model that does not require excessive computational expense and is therefore efficient in terms of processing speed. For the training of the first classifier presented in Figure 1, the input data of the neural network come from a scaling of the color images to a size of  $224 \times 224$  pixels. The architecture used comprises a max-pooling layer ( $7 \times 7$ ), a flatten layer; a hidden layer of 128 neurons with a "ReLU" activation function, a Dropout of 0.5, an output layer with two neurons, and a Softmax activation function. Regarding the training of the classifier, the configurations from Table 1 are used.

**Table 1.** Settings used in the classifier training.

Parameter	Detail
Learning rate	1e-4
Epochs	20
Batch size	32
Optimizador:	Adam
Loss function	Binary Cross Entropy

At this point (first and second stage together), we have information about where the face is and whether or not it has a face mask. Therefore, it is possible to pass this information to the third stage, which is where the recognition task is carried out. Figure 3 shows a diagram of the operation of the second stages.



**Figure 3.** Block diagram for face detection with and without a face mask.

### 2.3.3. Third Stage

Once the face of the person has been identified, in the third stage, facial recognition is carried out, for which a set of own observations is used that is built based on the faces of various people. For the construction of the set of observations, a balance is sought in terms of gender, namely, five women and five men from whom the images are obtained. Figure 4 shows the set of faces using a mask and Figure 5 shows without a face mask.



**Figure 4.** Set of observations using a face mask.



**Figure 5.** Set of observations without using a face mask.



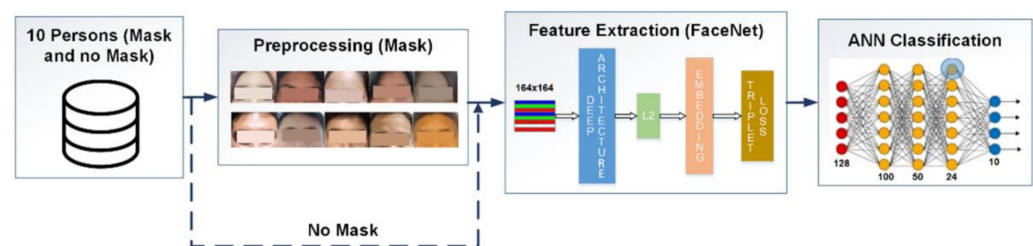
## 2.4. Training of Facial Recognition Models

The procedure to obtain the images is as follows: (i) during a week, daily videos of the face of each individual are obtained, seeking to capture different angles and different environmental conditions (lighting changes). (ii) From the videos obtained, images are captured at different moments in order to build a set of observations with images. (iii) The images where the face is not found in the entirety are eliminated. At the end of the procedure, a total of 13,359 images are obtained; 7067 (52.9%) with a face mask and 6292 (47.1%) without a face mask. In practice, acquiring so many images of a face requires a short video recording of a few seconds showing different views of the face. Regarding the identification of the images used for the recognition of people, the images have been labeled from left to right, as follows:

- Women: W1, W2, W3, W4, and W5 (Vicky, Mela, Damaris, Ale, and Yaritza, respectively).
- Men: M1, M2, M3, M4, and M5 (Oscar, Jorge, Pablo, John, and Jonathan respectively).

In this way, when a person is recognized in an image, the name information can be accessed and placed in the image. Therefore, once the necessary data have been obtained, the recognition model is trained. The two facial recognition models follow the architecture of Figure 6, which is briefly explained below. To do this, from this set of observations, the facial recognition model has two approaches:

1. The first model aimed at recognizing people using a face mask.
2. The second model aimed at recognizing people not using a face mask.



**Figure 6.** Block diagram of the training of the two face recognition models (mask and no mask).

**Database:** Set of observations of the faces of people using a mask and without using a mask for both approaches of the third stage.

**Preprocessing:** For the facial recognition model of people using a face mask, only 3/5 of the upper part of the face is extracted. This in order to discriminate the mask that the person is using (as in the experimental tests, this information is not useful for the recognition model). Whereas, for the facial recognition model without a face mask, the image of the full face is used. In both cases, the resulting image is scaled to a size of  $164 \times 164$  pixels.

**Characteristics extraction:** The FaceNet model is used to extract the most essential characteristics of the face. This model extracts the most essential features from the input image, in this case a face, and returns a vector of 128 features. The input of the network is an image with a human face and, using a deep metric learning technique, it calculates the metric to generate vectors of real characteristics [17]. This network is a model belonging to PyTorch, which can generate neural networks similar to Caffe [18]. The image is inserted into the network, then passes through the neural network and obtains the embedding of each face represented by  $f(x) \in R^d$ . This method attempts to ensure that the image of a specific person ( $x_i^a$ ) is closer to all of the images of the same person ( $x_i^p$ ) and away from images of other people ( $x_i^n$ ). Equation (1) shows the calculation of the loss ( $L$ ), where  $\alpha$  is the margin applied between positive and negative pairs [54]. It receives an image of

$164 \times 164$  pixels as the input data, and a vector of characteristics of 128 elements called “face embedding” is obtained at the output.

$$L = \sum_i^N \left[ \|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right] \quad (1)$$

ANN classification: Once the vector of characteristics has been obtained, any machine learning classification model can be applied; in this case, a feedforward multilayer perceptron is chosen.

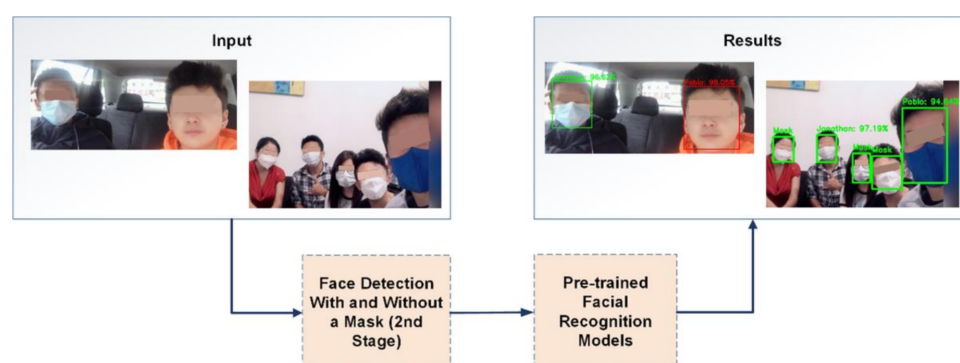
A feedforward multilayer perceptron (ANN) is chosen, because it does not need a large amount of data, it has been demonstrated that an ANN is a universal approximator with excellent results, and it does not require as much computational cost. After investigating the literature, it is best to use convolutional neural networks, but at this stage it is not necessary, as the best characteristics have been extracted by FaceNet and it only uses a vector of characteristics, instead of a raw image.

The architecture of our simple neural network applied to the two approaches is as follows:

- 128 neurons in the initial layer (size of the feature vector—face embedding)
- 100 neurons in a hidden layer with built-in ReLu activation
- 50 neurons in a hidden layer with built-in ReLu activation
- 24 neurons in a hidden layer with built-in ReLu activation
- 10 neurons in the output layer with a Softmax activation function
- The following configurations are used for classifier training:
- Epochs: 15
- Batch Size: 32
- Optimizer: Adam
- Loss Function: MSE

## 2.5. Application of the Facial Recognition System

Once the facial recognition models have been trained, they are applied following the results obtained in the second stage. In this way, in addition to defining whether or not the person uses the biosafety material, it is possible to know their identity. This as long as the face is within the selected database. These obtained models are applied to the previously identified faces, and a label is returned with the name and the probability of a match in the face. The procedure is described in Figure 7.



**Figure 7.** Block diagram of the implemented face recognition system.

## 2.6. Implementation Costs

In the market, there are a number of high-processing equipment oriented to the development of data science and deep learning, and because of a lack of resources for acquiring equipment like this, the free servers that Google provides through the GoogleColab platform are used.

The characteristics of a computer that Google offers for free are as follows:

- 12 GB RAM
- 50 GB disk
- Duration of 12 h, if it is used for more than 90 min, it disconnects.
- Using a random GPU.

In these conditions, the implementation costs do not exceed 400 euros, as it uses a portable computer with basic characteristics to send the Python programming of the cloud computer to execute.

The cost of a computer with similar characteristics in the international market, including a 4.1 Ghz processor, has an average price of 1500 euros, and if a monitor and camera with an approximate value of 220 euros is added to this, the complete equipment reaches a cost of 1720 euros.

### 3. Results

The experiments carried out have the purpose of demonstrating the potential use of the system, so tests are carried out using the recall metrics, Precision, F1-score, and the corresponding macro avg and weighted avg. The objective of using these metrics is to evaluate the system from different perspectives. Recall and precision indicate the ability of the model to correctly detect true positives. Recall also considers the false negatives detected, and the precision of the false positives detected by the model. False positives, in this case of face detection with masks, occur when an object is labeled as a face. For example, the system frames a plant as a face with a face mask, as it is false that a plant is a positive face. The reasons this can occur in our system are numerous, which is why hard work is needed in order to collect a large database so that the model being trained can better distinguish the desired classes (faces). False negatives occur when a face is not detected in the first stage, because the face has covered areas that make classification difficult; in this proposal, this initial classifier is a pre-developed tool. On the other hand, the F1-score provides a global measure of the system's performance, it is a combination of precision and recall (in a single value), with 0 being low performance and 1 being the best possible performance (all cases detected correctly).

By considering the macro avg metric, sd can get a general idea of the average of all of the experiments, while the weighted avg establishes an average measure of all of the experiments, but considering the number of observations of each class. Thus, in the event that a class has a higher score, the final weighted avg score will not be affected by it, but will give a value of importance to each score depending on the amount of observation. When considering these metrics, what is sought is to verify the robustness of the method by classifying both classes. In the second stage, the face classifier training with a mask and faces without a mask takes a period of approximately 10 h, and in the third stage the extraction of face embeddings takes approximately 5 h and the ANN-based classifier training takes 10 min.

#### 3.1. Face Classification—First and Second Stages

For the training of face classification, 20% of the data is used for testing and the rest are used to train the model; Figure 8 shows the graph corresponding to the training. The convergence obtained from the model occurs approximately at 10 epochs and an accuracy of 99.6% is reached for training data. Finally, analyzing the trained model with the test data, the confusion matrix is obtained (Figure 9), where the following values are observed: true positives (VP) of 440, false negatives (FN) of 1, false positives (FP) of 4, and true negatives (VN) of 996. When comparing both graphs, it can be seen that there is a correspondence in terms of the high accuracy achieved and the low detection of FN and FP during training. Equations (2)–(5) describe how the accuracy, precision, sensitivity, and specificity have

been calculated, respectively. In addition, it has an accuracy of 99.65%, precision of 99.09%, sensitivity of 99.77%, and specificity of 99.6%.

$$\text{Accuracy} = (\text{VP} + \text{VN}) / \text{Total} \quad (2)$$

$$\text{Precision} = \text{VP} / (\text{VP} + \text{FP}) \quad (3)$$

$$\text{Sensitivity} = \text{VP} / (\text{VP} + \text{FN}) \quad (4)$$

$$\text{Specificity} = \text{VN} / (\text{VN} + \text{FP}) \quad (5)$$

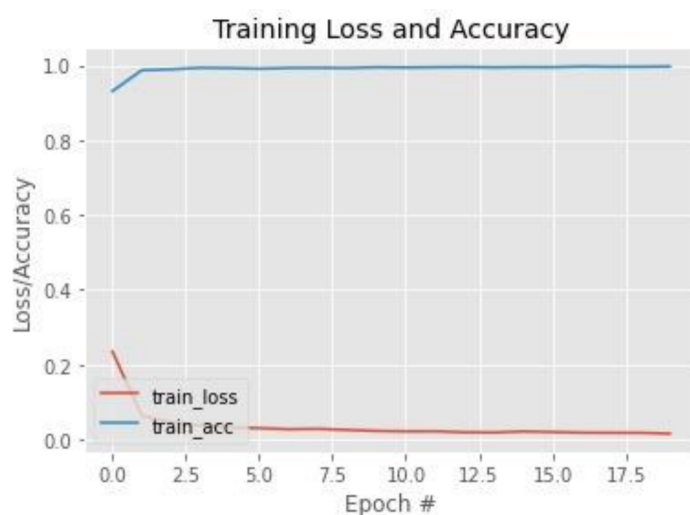


Figure 8. Training loss vs. accuracy classifier.

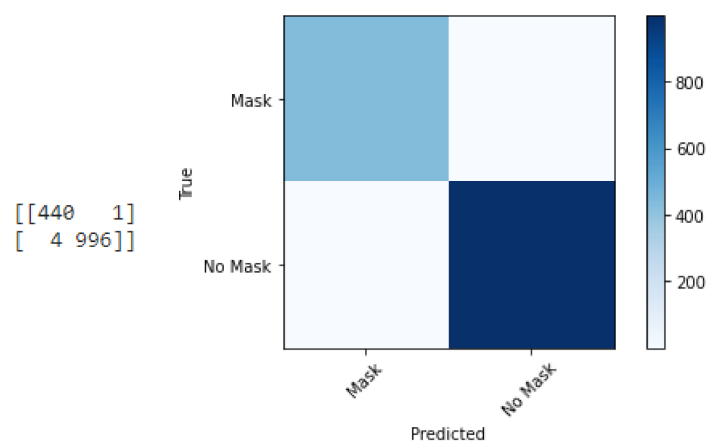


Figure 9. Confusion matrix classifier.

With the values obtained during the training, the accuracy, precision, sensitivity, and specificity are calculated—the results of which are seen in Table 2. The values obtained confirm the good performance of the classifier to detect when there is or is not a mask on the face. It can be noted that the F1-score for both classes (mask and no mask) has a value close to one that indicates the good performance of the model. In addition, when observing the weighted avg, it can be verified that for both classes, the model works correctly, regardless of the amount of data per class.

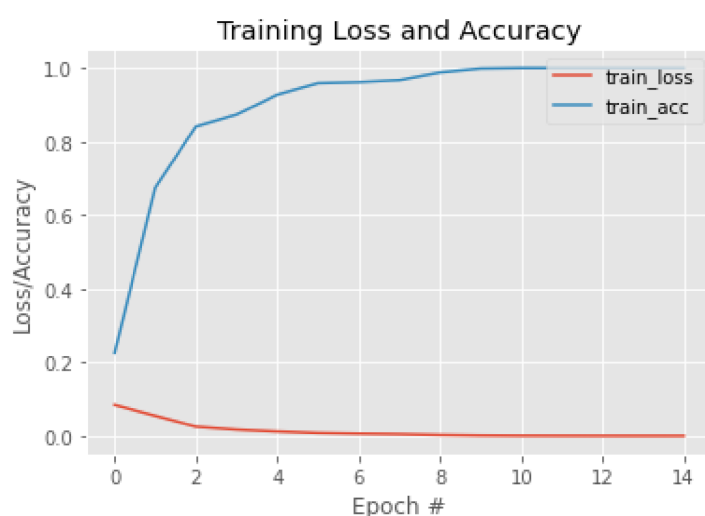


**Table 2.** Metrics for each class (with face mask).

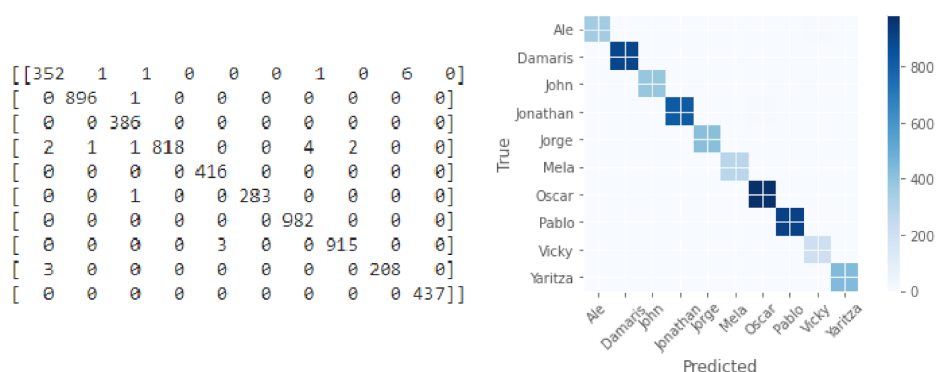
	Precision	Recall	F1-Score
Face mask	0.99	1.00	0.99
No face mask	1.00	1.00	1.00
Accuracy			1.00
Macro avg	0.99	1.00	1.00
Weighted avg	1.00	1.00	1.00

### 3.2. Facial Recognition—Third Stage

For the facial recognition model, 20% of the data is used in training and 80% in testing in order to avoid overtraining in the model. Figure 10 shows the graph corresponding to facial recognition training with a face mask.

**Figure 10.** Training loss vs. accuracy in a facial recognition (FR) model with face mask.

In this model, convergence is reached at approximately 10 epochs, with an accuracy of 98%. When evaluating the model with the test data, the accuracy is 99.52% and the confusion matrix that is observed in Figure 11 is obtained. Table 3 shows the metrics obtained from the facial recognition model with a face mask.

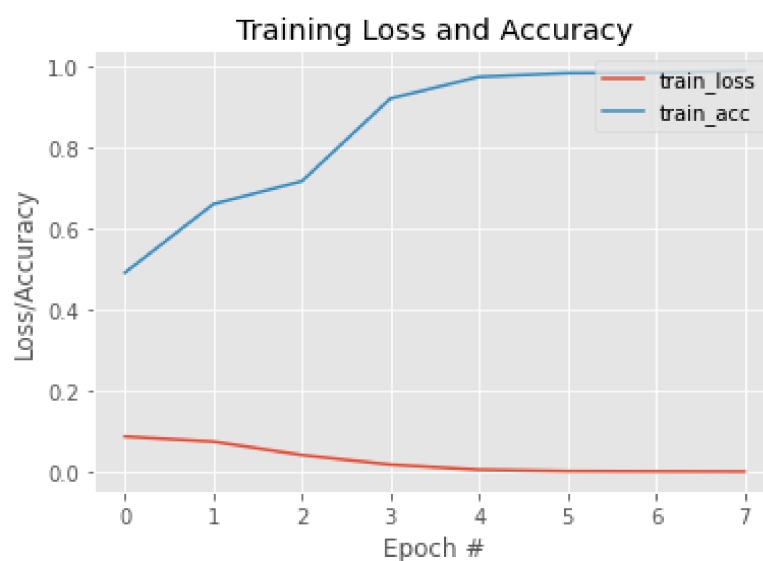
**Figure 11.** Confusion matrix in FR model with a face mask.

**Table 3.** Metrics for each class (with face mask).

	Precision	Recall	F1-Score
W4	0.99	0.98	0.98
W3	1.00	1.00	1.00
M4	0.99	1.00	0.99
M5	1.00	0.99	0.99
M2	0.99	1.00	1.00
W2	1.00	1.00	1.00
M1	0.99	1.00	1.00
M3	1.00	1.00	1.00
W1	0.97	0.99	0.98
W5	1.00	1.00	1.00
Accuracy			1.00
Macro avg	0.99	0.99	0.99
Weighted avg	1.00	1.00	1.00

When observing the F1-score, it can be seen that, in almost all cases, the score is closest to 1. When comparing this result with respect to the weighted avg, it can be seen that a good performance occurs in all classes, regardless of the amount of data that exists in each class.

On the other hand, in Figure 12, the training graph of the facial recognition model of people without a mask is shown.

**Figure 12.** Training loss vs. accuracy in FR model without a face mask.

In training, an accuracy of 99% is obtained. When evaluating the model with the test data without a face mask, an accuracy of 99.96% is obtained and the confusion matrix is observed in Figure 13.

As can be seen in Table 4, the model successfully classifies the face of a certain person. When comparing the results obtained in the macro avg and the weighted avg of the F1-score, an adequate functioning is corroborated, which is not influenced by the number of observations in a class. It should be noted that, if the person in the image is in the database, the name and the probability of success are placed on the label. On the other hand, if a face is not in the database, it is still detected, but a label of “mask” or “no mask” is added (because the name is not known), referring to whether said person is using or not using mask. In addition, it should be considered that the confidence percentage used in the system to accept if a face belongs to a person is 60%. The final results obtained from the

complete system in operation are shown below in Figure 14. The average time to run the entire system is 0.8399 s, which is acceptable for use in a practical application.

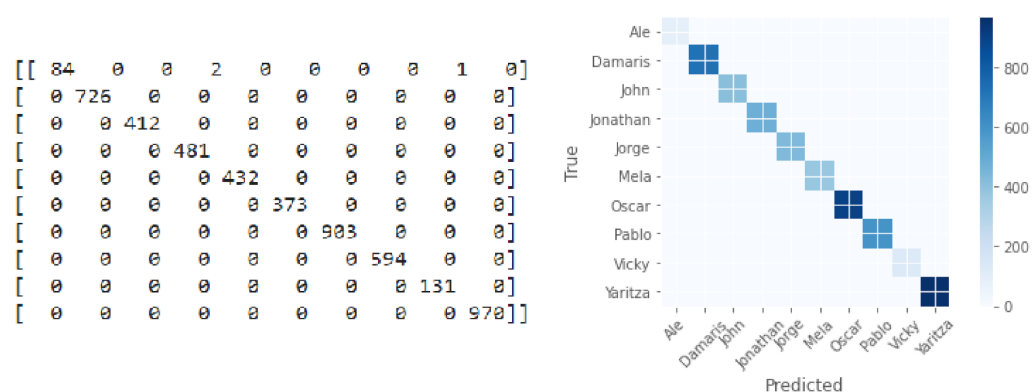


Figure 13. Confusion matrix for the FR model without mask.

Table 4. Metrics for each class (without face mask).

	Precision	Recall	F1-Score
W4	1.00	0.97	0.98
W3	1.00	1.00	1.00
M4	1.00	1.00	1.00
M5	1.00	1.00	1.00
M2	1.00	1.00	1.00
W2	1.00	1.00	1.00
M1	1.00	1.00	1.00
M3	1.00	1.00	1.00
W1	0.99	1.00	1.00
W5	1.00	1.00	1.00
Accuracy			1.00
Macro avg	1.00	1.00	1.00
Weighted avg	1.00	1.00	1.00

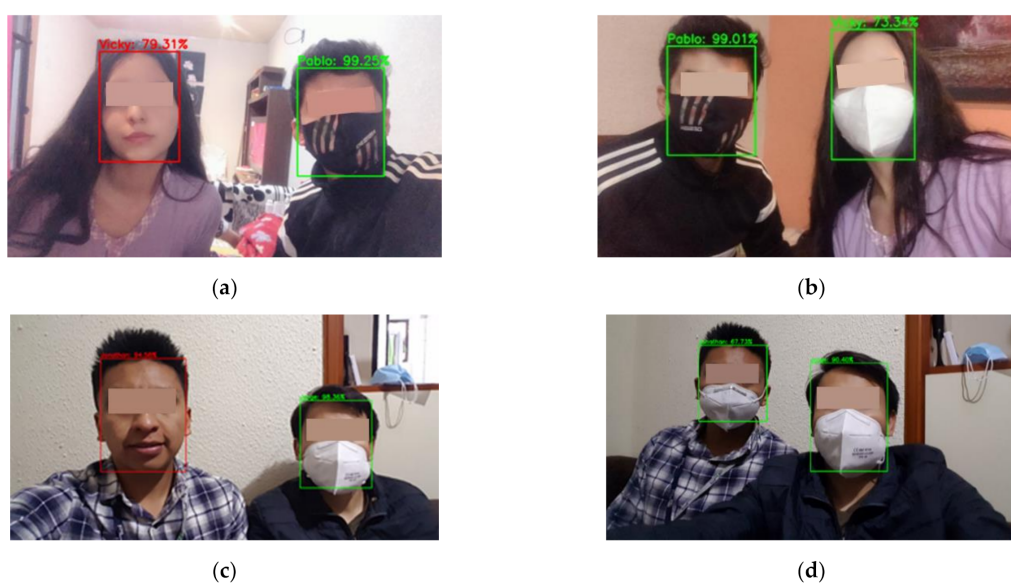
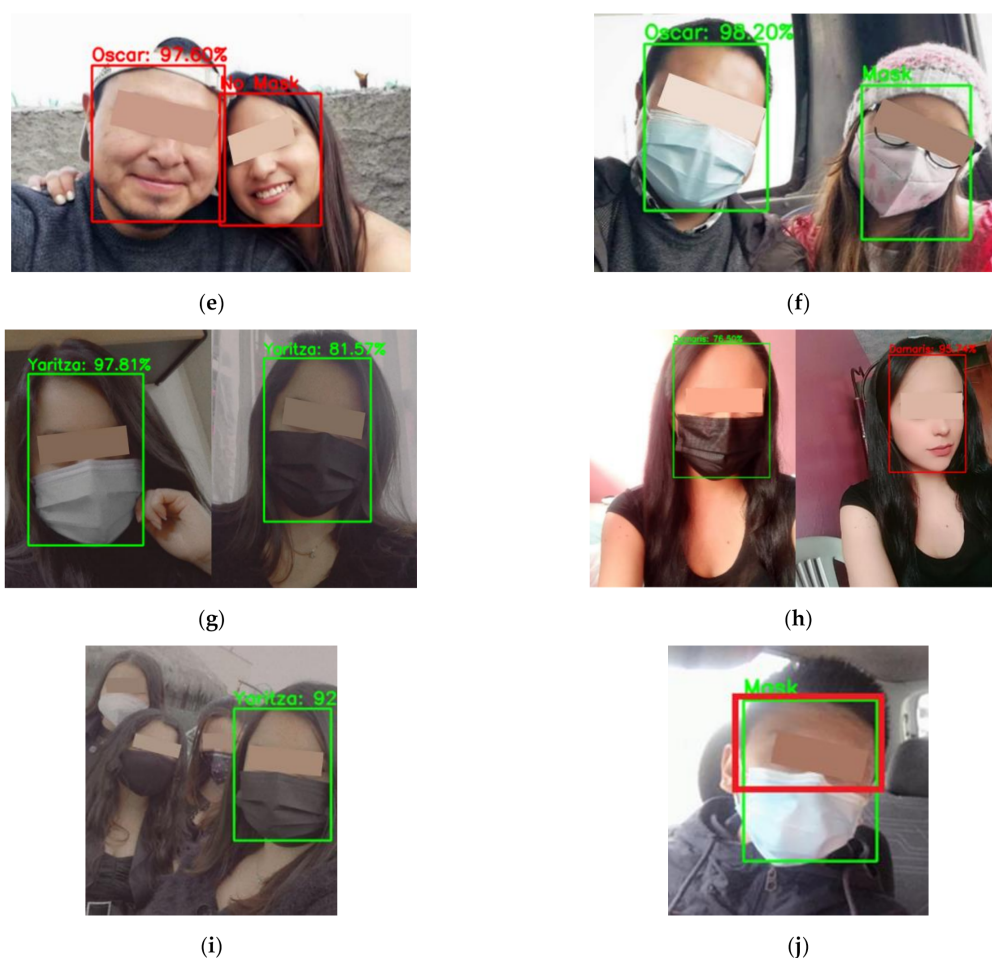


Figure 14. Cont.



**Figure 14.** Functionality test for (a) test 1, (b) test 2, (c) test 3, (d) test 4, (e) test 5, (f) test 6, (g) test 7, (h) test 8, (i) fail 1, and (j) fail 2.

#### 4. Discussion

As time passes worldwide, the COVID-19 pandemic poses greater challenges for humans. Over the past two years, leaders in biometric systems, such as FLIR Systems, Zkteco, among others, have launched proposals for updated access control and temperature detection. They also highlight the products previously launched by the companies SenseTime, Telpo, and Herta. All of this is focused mainly on temperature control in airports and high-traffic places, with which the development of easy recognition systems has been relegated.

As previously mentioned, the regular use of face masks has directly affected facial recognition systems. This has altered the usual development of activities and delayed innovation processes. For example, Apple's Face ID is designed to identify users based on the mouth, nose, and eyes, which must be fully visible. In the same way, other important companies worldwide, such as Go from Amazon and Walmart's "Store of the Future", are affected, as they use this technology to interact with their customers. There are independent investigations as in [55], where Google resources have been used to develop systems for detecting the use of masks, but with still low yields.

As can be seen in the bibliography presented, during the advancement of this pandemic, CNNs have been used as a priority for the diagnosis of the disease. Various standard training methods have been used, such as EfficientNet [39] and y ResNet-34 [43], as well as other proprietary ones such as nCOVnet [42]. Similarly, for face mask use detection systems, several methods have been tested, such as RCNN [46], VGG-16 [47], MobileNetV2 [48], SVM [49], InceptionV3 [50], and y SRCNet [51]. All of them have been analyzed and it is



determined that using Mobile-NetV2 would have efficient results, with a low consumption of computational resources. In addition, when comparing the efficiency of the various systems, it is found that this proposal is competitive, as it has 99.65%, only surpassed by [49,50] (100% in both cases). From the point of view of facial recognition, a similar proposal has not been found, which clearly shows the added value offered. By having a precision greater than 99%, the starting hypothesis is corroborated, establishing variability in this type of system and also adding robustness. Table 5 presents a comparison of this proposal with the other works.

**Table 5.** Comparison of this proposal with other works.

	Current Approach	[46]	[47]	[48]	[49]	[50]	[51]
Classification mask and no mask	X	X	X	X	X	X	X
Facial recognition	X						
Social distance measurement		X					
Network type	MobileNetV2	R-CNN	VGG-16	MobileNetV2	SVM	InceptionV3	SRCNet
software	Python	MATLAB	Python	Python	MATLAB	Not specified	MATLAB
Prosecution	Google Colab	Not specified	Computer				
Entry	images	Real time video	images	images	images	images	
Image or video capture distance	0.5 metros	3 metros	Not specified	Not specified	Not specified	Not specified	Not specified
Database size	13,359 images	Not specified	12,500 images	5521 images	10,000 images	1570 images	3835 images
Calculation time	0.84 s	Not specified	Not specified	Not specified	0.03 s (only classifier)	Not specified	0.1 s
Precision	99.96% and 99.52%	93.4%	96%	92.64%	100%	100%	98.7%
Limitations/Observations	Problems with lighting and programming code that could be optimized.	Problems with white skin tones that are mistaken for the face mask	Adjusting the settings requires a review.	Low performance.	high computational consumption.	Reduced database. It does not have facial recognition.	Reduced database. The images used do not have lighting changes.

Other proposals in face detection have also reviewed, but with the ocular recognition variant, as seen in [52,53]. However, it should be clarified that focusing on the human eye represents a different approach than the one analyzed in this document. In addition, this would require the acquisition of other hardware and software components that would increase the requirements presented. This document seeks to capture the upper part of the face, i.e., eyes, eyebrows, forehead, and hair. Thus, it has been shown that this approach can be successful.

On the subject of convolutional neural networks, the most important thing is to have a good database (thousands of observations). This is not a rule, but the larger the database, the better the resulting model, as, in a certain way, more characteristics are extracted and a model that represents them can be generalized. It does not matter if there are many or few classes (people); for each class (person), a considerable amount of observations is needed (images, database, etc.). This type of architecture is ideal when it comes to artificial vision, as it is demonstrated in the literature that convolutional neural networks are leading the advances in artificial vision, which is why this architecture is chosen, in order to have the

best possible results, as the topic of facial recognition is an area where the system should fail as little as possible.

The small number of people is not a problem, because the proposal is aimed at access control, which is a structured environment, in order for a restricted area to be limited to 10 or fewer people with entry authorization. In addition, the literature reviewed presents cases with the same performance for differentiating people with and without a face mask. The architecture has a heavy component in Transfer Learning (MobileNet) and is used to train the first classifier (second stage); this model is the one that classifies whether a face is using a mask or not. This has nothing to do with the classes for facial recognition, as facial recognition uses a fairly simple and lightweight architecture such as a feedforward multilayer perceptron.

Although the tools used in this work use existing models, the authors did not find a proposal that uses them in combination with a similar application. Another contribution is the use of cloud computing for this architecture, offering a low-cost system.

## 5. Conclusions and Future Works

This prototype system allows for the facial recognition of people with and without a mask, and could be used as a low computational consumption proposal for personnel access control. The two models of this system are tested with images, thus achieving better precision and optimization for each model. The face of someone found in the database is successfully classified to provide the name tag and probability of success. The three stages of the system allowed for the relevant characteristics of a person's face to be extracted, and thus use a simple neural network for the classification task. In this sense, the use of a "Face Embeddings" as input to the neural network obtained satisfactory results in the experiments carried out.

During the training of the third stage, it is possible to notice that there is an over-adjustment, this fact is due to the fact that the database is built for this stage with a few participants, although it is composed of several images, and does not exist much variability. However, the system shows potential to be used in differentiated facial recognition applications. It should be noted that if a face is not found in the database, it will be detected, but the tag "mask" or "no mask" will be added, which refers to whether the person is wearing a mask. It should be considered that the level of confidence used in the system to accept if a face belongs to someone is 0.6. When defining whether people are wearing a mask or not, the accuracy is 99.65%. When evaluating the facial recognition model with the test data of people who do not use a mask, an accuracy of 99.96% is obtained, and for those who use a mask, an accuracy of 99.52% is obtained. In this way, the basis for future research that can expand the study in this field is established.

In the bibliographic review, the use of MATLAB has been evidenced as an alternative proposal that could provide a lower number of false positives that should be evaluated. It is also proposed to investigate new extraction architectures that can be compared with FaceNet, and to thus choose the best one. One important thing to note is that the system has difficulty detecting certain faces when wearing a mask. This problem is due to the fact that initially, the Open CV Deep Learning based face detector was being used and it is not designed to detect faces with masks. It could also be observed that the face recognition stage is not robust when the detected face presents a certain angle of inclination. However, this is not a problem of great impact, as this application is oriented to access control and at this point, the person must maintain a firm and straight posture in front of the device that acquires the image.

Therefore, as a future work, merging the first and second stages in the same model and creating an own algorithm that directly searches for a face with and without a mask should be considered. This avoids first using the Open CV face detector and then classifying faces with and without mask. This will further reduce the processing time and make the model more robust. In addition, it is proposed that in the future, a comparative study of the models used for the transfer of learning can be carried out in order to determine the best

model and network trained in unfavorable evaluation conditions. Once the final models have been trained, they can be compressed and deployed on low-cost embedded devices such as Raspberry Pi or mobile devices.

**Author Contributions:** Conceptualization, J.S.T., J.B. and J.V.-A.; methodology, J.S.T. and P.C.; software, J.S.T. and P.C.; validation, J.S.T., J.B. and J.V.-A.; formal analysis, J.S.T., J.B. and J.V.-A.; investigation, J.S.T. and J.B.; resources, J.S.T., J.B. and J.V.-A.; data curation, J.S.T., J.B. and J.V.-A.; writing—original draft preparation, J.S.T. and J.B.; writing—review and editing, J.B. and J.V.-A.; visualization, J.S.T., J.B. and J.V.-A.; supervision, J.B. and J.V.-A.; project administration, J.S.T., P.C. and J.V.-A.; funding acquisition, J.V.-A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research and the APC was funded by Universidad Tecnológica Indoamérica (Ecuador).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Adjabi, I.; Ouahabi, A.; Benzaoui, A.; Taleb-Ahmed, A. Past, present, and future of face recognition: A review. *Electronics* **2020**, *9*, 1188. [CrossRef]
2. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep Learning for Computer Vision: A Brief Review. *Comput. Intell. Neurosci.* **2018**, *2018*. [CrossRef]
3. Chakraborty, B.K.; Sarma, D.; Bhuyan, M.K.; MacDorman, K.F. Review of constraints on vision-based gesture recognition for human-computer interaction. *IET Comput. Vis.* **2018**, *12*, 3–15. [CrossRef]
4. Ko, B.C. A brief review of facial emotion recognition based on visual information. *Sensors* **2018**, *18*, 401. [CrossRef] [PubMed]
5. Egger, M.; Ley, M.; Hanke, S. Emotion Recognition from Physiological Signal Analysis: A Review. In Proceedings of the Electronic Notes in Theoretical Computer Science, Yangzhou, China, 14–17 June 2019; Elsevier B.V.: Larnaca, Cyprus, 2019; Volume 343, pp. 35–55. [CrossRef]
6. Qiu, S.; Liu, Q.; Zhou, S.; Wu, C. Review of artificial intelligence adversarial attack and defense technologies. *Appl. Sci.* **2019**, *9*, 909. [CrossRef]
7. Dang, K.; Sharma, S. Review and comparison of face detection algorithms. In Proceedings of the 7th International Conference Confluence 2017 on Cloud Computing, Data Science and Engineering, Noida, India, 12–13 January 2017; pp. 629–633.
8. Cook, C.M.; Howard, J.J.; Sirotnin, Y.B.; Tipton, J.L.; Vemury, A.R. Demographic Effects in Facial Recognition and Their Dependence on Image Acquisition: An Evaluation of Eleven Commercial Systems. *IEEE Trans. Biom. Behav. Identity Sci.* **2019**, *1*, 32–41. [CrossRef]
9. Dargan, S.; Kumar, M. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Syst. Appl.* **2020**, *143*, 113114. [CrossRef]
10. Jeon, B.; Jeong, B.; Jee, S.; Huang, Y.; Kim, Y.; Park, G.H.; Kim, J.; Wufuer, M.; Jin, X.; Kim, S.W.; et al. A facial recognition mobile app for patient safety and biometric identification: Design, development, and validation. *JMIR mHealth uHealth* **2019**, *7*, e11472. [CrossRef]
11. Gonzalez-Sosa, E.; Fierrez, J.; Vera-Rodriguez, R.; Alonso-Fernandez, F. Facial soft biometrics for recognition in the wild: Recent works, annotation, and COTS evaluation. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 2001–2014. [CrossRef]
12. Galterio, M.G.; Shavit, S.A.; Hayajneh, T. A review of facial biometrics security for smart devices. *Computers* **2018**, *7*, 37. [CrossRef]
13. Karthik, K.; Aravindh Babu, R.P.; Dhama, K.; Chitra, M.A.; Kalaiselvi, G.; Alagesan Senthilkumar, T.M.; Raj, G.D. Biosafety Concerns During the Collection, Transportation, and Processing of COVID-19 Samples for Diagnosis. *Arch. Med. Res.* **2020**, *51*, 623–630. [CrossRef]
14. Souza, T.M.L.; Morel, C.M. The COVID-19 pandemics and the relevance of biosafety facilities for metagenomics surveillance, structured disease prevention and control. *Biosaf. Heal.* **2020**, *3*, 1–3. [CrossRef]
15. Ortiz, M.R.; Grijalva, M.J.; Turell, M.J.; Waters, W.F.; Montalvo, A.C.; Mathias, D.; Sharma, V.; Renoy, C.F.; Suits, P.; Thomas, S.J.; et al. Biosafety at home: How to translate biomedical laboratory safety precautions for everyday use in the context of COVID-19. *Am. J. Trop. Med. Hyg.* **2020**, *103*, 838–840. [CrossRef]
16. Mills, M.; Rahal, C.; Akimova, E. Face masks and coverings for the general public: Behavioural knowledge, effectiveness of cloth coverings and public messaging. *R. Soc.* **2020**, 1–37. Available online: <https://royalsocietypublishing.org/doi/full/10.1098/rsos.200114> (accessed on 5 February 2021).
17. Ngan, M.; Grother, P.; Hanaoka, K. Ongoing Face Recognition Vendor Test (FRVT)—Part 6A: Face recognition accuracy with masks using pre-COVID-19 algorithms. *Natl. Inst. Stand. Technol.* **2020**. [CrossRef]

18. Cheng, V.C.C.; Wong, S.C.; Chuang, V.W.M.; So, S.Y.C.; Chen, J.H.K.; Sridhar, S.; To, K.K.W.; Chan, J.F.W.; Hung, I.F.N.; Ho, P.L.; et al. The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2. *J. Infect.* **2020**, *81*, 107–114. [[CrossRef](#)] [[PubMed](#)]
19. Liu, X.; Zhang, S. COVID-19: Face masks and human-to-human transmission. *Influenza Other Respi. Viruses* **2020**, *14*, 472–473. [[CrossRef](#)] [[PubMed](#)]
20. Rahman, A.; Hossain, M.S.; Alrajeh, N.A.; Alsolami, F. Adversarial Examples—Security Threats to COVID-19 Deep Learning Systems in Medical IoT Devices. *IEEE Internet Things J.* **2020**, *8*, 9603–9610. [[CrossRef](#)]
21. Mundial, I.Q.; Ul Hassan, M.S.; Tiwana, M.I.; Qureshi, W.S.; Alanazi, E. Towards facial recognition problem in COVID-19 pandemic. In Proceedings of the 2020 4th International Conference on Electrical, Telecommunication and Computer Engineering, ELTICOM 2020—Proceedings, Medan, Indonesia, 3–4 September 2020; pp. 210–214.
22. Ting, D.S.W.; Carin, L.; Dzau, V.; Wong, T.Y. Digital technology and COVID-19. *Nat. Med.* **2020**, *26*, 459–461. [[CrossRef](#)]
23. Yao, G.; Lei, T.; Zhong, J. A review of Convolutional-Neural-Network-based action recognition. *Pattern Recognit. Lett.* **2019**, *118*, 14–22. [[CrossRef](#)]
24. McCann, M.T.; Jin, K.H.; Unser, M. Convolutional neural networks for inverse problems in imaging: A review. *IEEE Signal Process. Mag.* **2017**, *34*, 85–95. [[CrossRef](#)]
25. Rawat, W.; Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* **2017**, *29*, 2352–2449. [[CrossRef](#)]
26. Aloysius, N.; Geetha, M. A review on deep convolutional neural networks. In Proceedings of the 2017 IEEE International Conference on Communication and Signal Processing, ICCSP 2017, Chennai, India, 6–8 April 2017; Volume 2018, pp. 588–592.
27. Al-Saffar, A.A.M.; Tao, H.; Talab, M.A. Review of deep convolution neural network in image classification. In Proceedings of the 2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications, ICRAMET 2017, Jakarta, Indonesia, 23–24 October 2017; Volume 2018, pp. 26–31.
28. Kamilaris, A.; Prenafeta-Boldú, F.X. A review of the use of convolutional neural networks in agriculture. *J. Agric. Sci.* **2018**, *156*, 312–322. [[CrossRef](#)]
29. Yang, Z.; Yu, W.; Liang, P.; Guo, H.; Xia, L.; Zhang, F.; Ma, Y.; Ma, J. Deep transfer learning for military object recognition under small training set condition. *Neural Comput. Appl.* **2019**, *31*, 6469–6478. [[CrossRef](#)]
30. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629. [[CrossRef](#)]
31. Schwendicke, F.; Golla, T.; Dreher, M.; Krois, J. Convolutional neural networks for dental image diagnostics: A scoping review. *J. Dent.* **2019**, *91*, 103226. [[CrossRef](#)] [[PubMed](#)]
32. Van Grinsven, M.J.J.P.; Van Ginneken, B.; Hoyng, C.B.; Theelen, T.; Sánchez, C.I. Fast Convolutional Neural Network Training Using Selective Data Sampling: Application to Hemorrhage Detection in Color Fundus Images. *IEEE Trans. Med. Imaging* **2016**, *35*, 1273–1284. [[CrossRef](#)]
33. Zou, L.; Yu, S.; Meng, T.; Zhang, Z.; Liang, X.; Xie, Y. A Technical Review of Convolutional Neural Network-Based Mammographic Breast Cancer Diagnosis. *Comput. Math. Methods Med.* **2019**, *2019*. [[CrossRef](#)]
34. Bernal, J.; Kushibar, K.; Asfaw, D.S.; Valverde, S.; Oliver, A.; Martí, R.; Lladó, X. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: A review. *Artif. Intell. Med.* **2019**, *95*, 64–81. [[CrossRef](#)]
35. Hassantabar, S.; Ahmadi, M.; Sharifi, A. Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches. *Chaos Solitons Fractals* **2020**, *140*, 110170. [[CrossRef](#)]
36. Lalmuanawma, S.; Hussain, J.; Chhakhuak, L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. *Chaos Solitons Fractals* **2020**, *139*, 110059. [[CrossRef](#)]
37. Mukherjee, H.; Ghosh, S.; Dhar, A.; Obaidullah, S.M.; Santosh, K.C.; Roy, K. Deep neural network to detect COVID-19: One architecture for both CT Scans and Chest X-rays. *Appl. Intell.* **2020**, 1–13. [[CrossRef](#)]
38. Ardakani, A.A.; Kanafi, A.R.; Acharya, U.R.; Khadem, N.; Mohammadi, A. Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. *Comput. Biol. Med.* **2020**, *121*, 103795. [[CrossRef](#)] [[PubMed](#)]
39. Marques, G.; Agarwal, D.; de la Torre Díez, I. Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network. *Appl. Soft Comput. J.* **2020**, *96*, 106691. [[CrossRef](#)] [[PubMed](#)]
40. Chowdhury, N.K.; Rahman, M.M.; Kabir, M.A. PDCOVNet: A parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images. *Heal. Inf. Sci. Syst.* **2020**, *8*, 27. [[CrossRef](#)] [[PubMed](#)]
41. Toraman, S.; Alakus, T.B.; Turkoglu, I. Convolutional capsnet: A novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. *Chaos Solitons Fractals* **2020**, *140*, 110122. [[CrossRef](#)]
42. Panwar, H.; Gupta, P.K.; Siddiqui, M.K.; Morales-Menendez, R.; Singh, V. Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. *Chaos Solitons Fractals* **2020**, *138*, 109944. [[CrossRef](#)]
43. Nayak, S.R.; Nayak, D.R.; Sinha, U.; Arora, V.; Pachori, R.B. Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study. *Biomed. Signal Process. Control* **2021**, *64*, 102365. [[CrossRef](#)]
44. Yu, X.; Lu, S.; Guo, L.; Wang, S.H.; Zhang, Y.D. ResGNet-C: A graph convolutional neural network for detection of COVID-19. *Neurocomputing* **2020**. [[CrossRef](#)]



45. Uddin, M.I.; Shah, S.A.A.; Al-Khasawneh, M.A. A Novel Deep Convolutional Neural Network Model to Monitor People following Guidelines to Avoid COVID-19. *J. Sens.* **2020**, *2020*. [[CrossRef](#)]
46. Meivel, S.; Indira Devi, K.; Uma Maheswari, S.; Vijaya Menaka, J. Real time data analysis of face mask detection and social distance measurement using Matlab. *Mater. Today Proc.* **2021**. [[CrossRef](#)]
47. Militante, S.V.; Dionisio, N.V. Real-Time Facemask Recognition with Alarm System using Deep Learning. In Proceedings of the 2020 11th IEEE Control and System Graduate Research Colloquium, ICSGRC 2020—Proceedings, Shah Alam, Malaysia, 8 August 2020; pp. 106–110.
48. Nagrath, P.; Jain, R.; Madan, A.; Arora, R.; Kataria, P.; Hemanth, J. SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustain. Cities Soc.* **2021**, *66*, 102692. [[CrossRef](#)] [[PubMed](#)]
49. Loey, M.; Manogaran, G.; Taha, M.H.N.; Khalifa, N.E.M. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Meas. J. Int. Meas. Confed.* **2021**, *167*, 108288. [[CrossRef](#)] [[PubMed](#)]
50. Chowdary, G.J.; Punna, N.S.; Sonbhadra, S.K.; Agarwal, S. Face mask detection using transfer learning of inceptionV3. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; In Proceedings of the International Conference on Big Data Analytics, Sonepat, India, 15–18 December 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 81–90.
51. Qin, B.; Li, D. Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network to Prevent COVID-19. *Sensors* **2020**, *20*, 5236. [[CrossRef](#)] [[PubMed](#)]
52. Alonso-Fernandez, F.; Diaz, K.H.; Ramis, S.; Perales, F.J.; Bigun, J. Facial Masks and Soft-Biometrics: Leveraging Face Recognition CNNs for Age and Gender Prediction on Mobile Ocular Images. *IET Biom.* **2021**. [[CrossRef](#)]
53. Kumari, P.; Seeja, K.R. A novel periocular biometrics solution for authentication during Covid-19 pandemic situation. *J. Ambient Intell. Humaniz. Comput.* **2021**, *1*, 3. [[CrossRef](#)]
54. Schroff, F.; Kalenichenko, D.; Philbin, J. FaceNet: A unified embedding for face recognition and clustering. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 815–823.
55. Wang, X.; Chen, Z.; Wei, B.; Ling, M. Application of Pruning Yolo-V4 with Center Loss in Mask Wearing Recognition for Gymnasiums and Sports Grounds of Colleges and Universities. In Proceedings of the 2020 IEEE 6th International Conference on Computer and Communications, ICC 2020, Chengdu, China, 11–14 December 2020; pp. 1373–1377.