IMPLEMENTATION OF A REAL TIME FACE RECOGNITION SYSTEM BASED ON NEURAL NETWORKS

Jose Luis Medrán del Río

Tutor: Antonio Albiol Colomer

Trabajo Fin de Máster presentado en la Escuela Técnica Superior de Ingenieros de Telecomunicación de la Universitat Politècnica de València, para la obtención del Título de Máster en Ingeniería Telecomunicación

Curso 2018-19

Valencia, 10 de Septiembre de 2019

Técnica Superior de Ingeniería de Telecomunicación
at Politècnica de València
D. Camino de Vera, s/n, 46022 Valencia
96 387 71 90, ext. 77190
etsit.upv.es


UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA


Nagoya
Institute of
Technology


TELECOM ESCUELA
TÉCNICA VLC SUPERIOR
DE INGENIERÍA DE
TELECOMUNICACIÓN

Abstract

In recent years, the use of neural networks for all kind of applications has experienced a sharp rise due to the improvement in computers and the growth of computing power that let us implement new techniques that were discarded before because they were too resource consuming.

Facial recognition systems have been largely researched in the last decade in two main sectors, in public security departments, they are employed thanks to the ability to detect and recognize “wanted” people, and in SNS, Social Networking Services, being Google and Facebook the two main leads of this sector with DeepFace and FaceNet respectively.

In this master thesis we use a neural network approach in order to implement a real time face recognition system with a user-oriented interface using Python. Versatility has been the main focus for this project letting the user create their own target groups for identification.

Performance tests have been performed on the model employed in this project and on the real time implementation. The results obtained show a state-of-the-art performance.

Resumen

En los últimos años, el empleo de redes neuronales en todo tipo de aplicaciones ha experimentado un fuerte aumento debido a la mejora en los equipos informáticos y al crecimiento de la potencia de cómputo, lo que ha permitido implementar nuevas técnicas que antes se descartaban porque consumían demasiados recursos.

Los sistemas de reconocimiento facial se han investigado en gran medida en la última década en dos sectores principales, en los departamentos de seguridad pública, donde se emplean gracias a la capacidad de detectar y reconocer a personas buscadas, y en SNS, servicios de redes sociales, siendo Google y Facebook los dos líderes principales de este sector con sus implementaciones, DeepFace y FaceNet respectivamente.

En esta tesis de máster, utilizamos un enfoque basado en redes neuronales para implementar un sistema de reconocimiento facial en tiempo real con una interfaz orientada al usuario basada en Python. La versatilidad ha sido el foco principal de este proyecto, permitiendo al usuario crear diferentes datasets en los que realizar la identificación.

Se han realizado pruebas de rendimiento sobre el modelo empleado en este proyecto y sobre la implementación en tiempo real presentando unos resultados que demuestran un alto rendimiento.

Resum

En els últims anys, l'ús de xarxes neuronals en tot tipus d'aplicacions ha experimentat un fort augment a causa de la millora en els equips informàtics i al creixement de la potència de còmput, el que ha permès implementar noves tècniques que abans es descartaven perquè consumien massa recursos.

Els sistemes de reconeixement facial s'han investigat en gran mesura en l'última dècada en dos sectors principals, en els departaments de seguretat pública, on es fan servir gràcies a la capacitat de detectar i reconèixer a persones buscades, i en SNS, serveis de xarxes socials, sent Google i Facebook els dos líders principals d'aquest sector amb les seves implementacions, DeepFace i FaceNet respectivament.

En aquesta tesi de màster, utilitzem un enfocament basat en xarxes neuronals per implementar un sistema de reconeixement facial en temps real amb una interfície orientada a l'usuari basada en Python. La versatilitat ha estat el focus principal d'aquest projecte, permetent a l'usuari crear diferents datasets en els quals fer la identificació.

S'han realitzat proves de rendiment sobre el model emprat en aquest projecte i sobre la implementació en temps real presentant uns resultats que demostren un alt rendiment.


UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA


Nagoya
Institute of
Technology


TELECOM ESCUELA
TÉCNICA VLC SUPERIOR
DE INGENIERÍA DE
TELECOMUNICACIÓN

Acknowledgments

I would like first to thank Antonio Albiol Colomer of the Communications Department on the Telecommunications School at Universitat Politècnica de València. Even if this project was carried out outside the UPV, your expert advice and encouragement throughout this project was just what I needed to successfully complete this master thesis.

I would like to thank my supervisor at Nagoya Institute of Technology, Sako Shinji sensei. The door to Sako-sensei's office was always open whenever I ran into a trouble spot or had a question about my research or my stance at Japan. I never ran into a problem when staying at Nagoya because of his invaluable support.

I would also like to thank the rest of the members of Sako-sensei's group of investigation. They made me feel like at home and supported me whenever I needed. Specifically, the help provided by Taniguchi during the first days was really appreciated because there were a lot of things I didn't know, and he took the time to explain them for me.

To my colleagues from Cosmo Village, I never thought before going to Japan that I would find such good people and good friends. I hope we can meet again in the future, because I will always remember the friendship and time we spent together.

Y por supuesto, quisiera dar las gracias a mis padres, quienes me han dado todo en esta vida y gracias a ellos estoy hoy aquí, después de tantos años de trabajo, terminando esta tesis. Su fortaleza, sacrificio y ayuda han sido indispensables para que yo sea la persona que soy hoy y a ellos se lo debo todo

Gracias ありがとうございます Thank you


UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA


Nagoya
Institute of
Technology


TELECOM ESCUELA
TÉCNICA VLC SUPERIOR
DE INGENIERÍA DE
TELECOMUNICACIÓN

Contents

Chapter 1: Introduction

The main objective of this master thesis is the implementation of a user-oriented real time face recognition system using Python including a GUI interface made with Kivy and Tkinter, two of the most commons and developed Python GUI packages. In order to accomplish this goal, we are going to follow the next scheme:

In the first place, we will introduce many recognition systems based on different biological traits and provide a historical context about the task of identifying someone automatically, focusing on the chosen system for this project, the facial recognition system.

Afterwards, we will explain how facial recognition systems can be used in many different environments and the different problems that can be resolved by its use.

Then, we will study the different concerns or technical problems that can be observed at present. We present two approaches, the first from an ethical point of view, how facial recognition systems affect people, and the second one from a technical angle, presenting the challenges that this technology needs to overcome.

The next step will be to study the different existing techniques to implement facial recognition systems before start talking about the technique used in this project and related work that has been used for the creation of this master thesis.

Later, different kind of tests have been performed to evaluate first, the quality of the model used as the core for this project, and then, the real time facial recognition system that has been implemented in this master thesis.

Finally, we will provide some conclusions about this master thesis and explain some technical problems that were faced during its development, and a future line of research very closely related to this project that we think it is worthwhile exploring.

Chapter 2: Identifications systems and historic context

2.1 Biometrics

A facial recognition system is an advanced technology that uses biometrics to identify or verify a person by mapping facial features from a photo or a video and comparing and analyzing them with the biometrical patterns already stored in a database to find a match.

The term “biometrics” is derived from the Greek words “bio” (life) and “metrics” (to measure). It refers to metrics related to human characteristics, and the biometrics authentication systems, such as the facial recognition systems, use them to automatically identify/verify people by the means of computer science.

Biometric identifiers [1] are different biological traits unique to a sole individual that let us perform an identification. These identifiers can be differentiated between physiological features, such as fingerprint, palm veins, face, DNA, palm print, hand geometry, iris or retina, and behavioral features like signature, voice or gait.

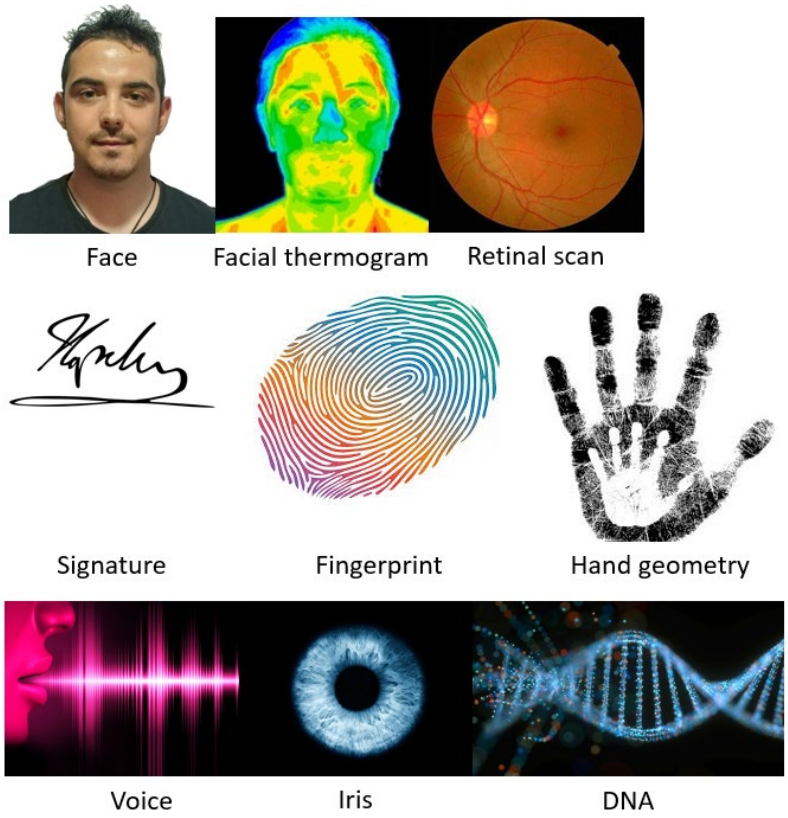


Figure 2.1.1. Examples of different biometric characteristics

2.2 Historical context

Historically [2], people have used this sort of biometric identifiers in all kind of applications. There is evidence that in a cave estimated to be at least 31.000 years old, the prehistoric paintings that cover the walls of the cave are surrounded by numerous handprints by way of showing individuality. There are also clay tablets that include fingerprints that record Babylonian business. Joao de Barros, a Spanish explorer and writer, exposed that in the ancient China, the merchants used fingerprints to close business deals.

There are examples like these throughout all history. But it is not until mid-1800s, with the industrial revolution and the rapid growth of cities that it caused, that the need to identify people surged.

This way, two kinds approaches arose. The first was the Bertillon system which originated in France in 1870 by Alphonse Bertillon, a measuring system that consisted in the recording of several body dimensions, physical descriptions and photographs that let them sort them out by height, arm length or any other need. Nevertheless, this system was not robust enough and collapsed in 1903 because the impossibility of distinguish between identical twins that were condemned to the US Penitentiary at Leavenworth.

The other approach, employed in Europe, Asia and America, consisted in the use of fingerprints by police departments. In this case, the first robust system originated in India by Azizul Haque for Edward Henry, Inspector General of Police of Bengal, in 1896. This system, called the Henry System, was the basis of modern-day AFIS (Automated Fingerprint Identification System) classification methods that were used until the 1990s. Recently, the Henry Classification System has been replaced by ridge flow classification approaches.

In 1936, Ophthalmologist Frank Burch studied the concept of using iris patterns as a method to recognize an individual. Following this approach, in 1985 Drs. Leonard Flom and Aran Safir, ophthalmologists, conclude that there are not two irises alike and in 1987 developed and patented an iris identification concept. Nevertheless, they did not have the means to perform it, neither the algorithm nor the method.

And in 1995 the US Defense Nuclear Agency with Irscan prepared a prototype unit based on the patented technology.

In the case of facial recognition [2,3,4], the one we will focus on this work, this is the easiest and oldest way to recognize each other. Nevertheless, this approach is not so straightforward if we are talking about machines.

In 1960s, Woodrow W. Bledsoe under contract to the US Government developed the first semi-automated face recognition system. This system required the user to locate the key common features on the photographs, such as eyes, ears nose, hairline and mouth. The process in which the recognition took place was based on calculated distances and ratios to a common reference point that were recorded in a database. This could be used later to compare the new people measurements to the database data to find a match.

In 1970s, Goldstein, Harmon and Lesk achieved greater precision in the semi-automated face recognition system using 21 specific subjective markers such as lip thickness and hair color to automate face recognition. Nevertheless, the problem of manually inputting the locations of the features remained.

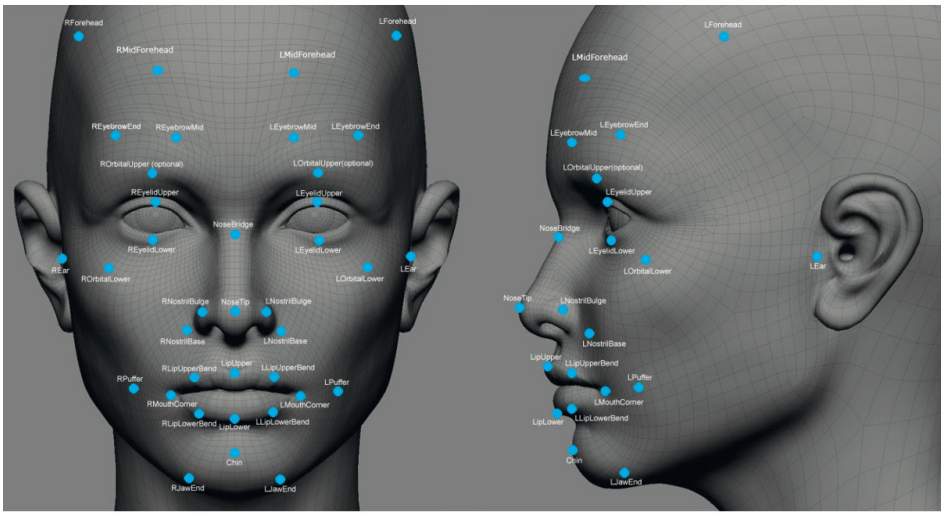


Figure 2.2.1. The 21 facial markers used by Goldstein, Harmon and Lesk.

In 1988, Kirby and Sirovich developed the eigenface approach, which consisted in the use of principle component analysis (PCA) on a collection of face images to form a set of basis features. This approach surged in order to obtain a low-dimensional representation of face images.

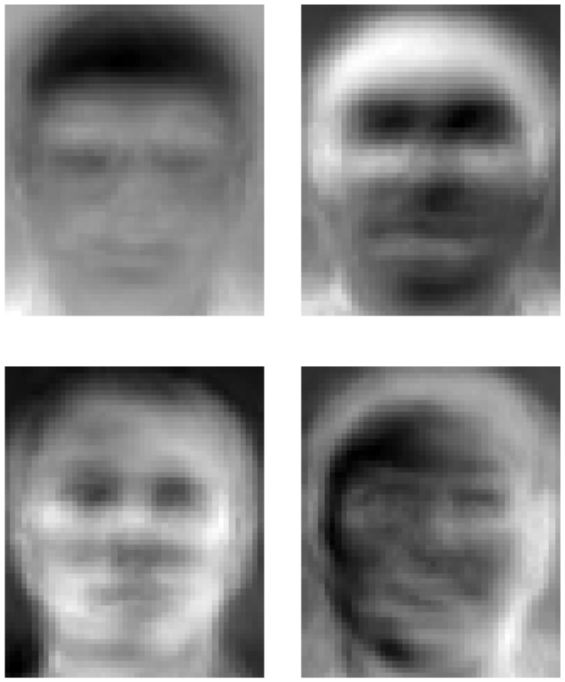


Figure 2.2.2. Some eigenfaces from AT&T Laboratories Cambridge. [4]

In 1991, Turk and Pentland kept working on the eigenface approach, and showed a way of calculating the eigenvectors of a covariance matrix in a way that the computers of the time could do it in near real time. This was a huge step in face recognition systems.

Between 1993-2000s the Defense Advanced Research Products Agency (DARPA) and the DoD Counterdrug Technology Development Program Office sponsored the FacE REcognition Technology (FERET) program with the objective of encouraging the development of face recognition algorithms and technology. This program consisted in the development of a facial image's database. This database was updated on 2003 including 24 bits high resolution images. There are 2416 representing 856 people.

In January 2001, a face recognition system was used at the Super Bowl with the objective of identifying "wanted" individuals at the stadium. The results were not good enough since no "wanted" individuals were recognized, but several innocent fans were falsely identified. The main problem at this moment was that facial recognition system were not prepared to operate in multitudes since the computing requirements were excessive.

In 2004, in order to encourage the development of new algorithms and techniques to improve the face recognition, the US Government arranged a sponsored challenger, the Face Recognition Grand Challenge (FRGC), were researchers analyzed the provided data to solve the problems and share ideas for the future.

In 2009, the Pinellas County Sheriff's Office established a forensic database that allowed the police officers to access the photographic archives from Department of Highway Safety and Motor Vehicles (DHSMV).

As of 2010, the Social networking services, leaded by Facebook started to implement face recognition techniques to help the users identify the people at their uploaded images. Starting from this moment people started worrying about privacy issues that could come from the implementation of this technique.

In 2011, the Panamá Government, associated with the US Government, implemented a pilot program of face recognition at the Tocumen airport in order to reduce the illegal drug smuggling. This led them to arrest several suspects from Interpol.

From that moment, facial recognition systems have been implemented by police departments to facilitate the identification of suspects.

In daily life, this kind of systems have also been constantly developed. Since 2017, when iPhone X was launched by Apple, this technology has reached society and is being used by more and more people every day, not only on their smartphones, but also in all kind of IoT implementations.

Chapter 3: General applications

In the last decade, facial recognition systems have been implemented at all stages of life, from police departments to homes for very different purposes. This is due to the new technologies and the greater computing power that allows us to implement this system in real life.

The general applications [22] that come from the employment of this technology are as follows:

- **Payments.** Companies want payments to be easy in order to speed up the process and to make the clients make customers not think much about spending money. After the rise of online shopping and contactless cards the next step is to use selfie payment applications that use the facial verification system to confirm a payment. It only needs the customer to face the smartphone to confirm a payment, that way the customers would not even need to use or carry their cards.
- **Access.** Facial verifying systems can be used to replace passcodes, key locks and old ways of providing verification to your home, car, smartphone or other consumer electronic devices.
- **Banking.** Facial recognition systems can be used at ATM to provide reliable multi-factor authentication that does not affect the customer experience.
- **Public security.** The US Customs and Border Protection have been using face recognition technology for many years to ensure that unwanted visitors do not enter the country and to recognize criminals trying to leave the country. This technology is not only passively used at the borders, but also the US Federal Bureau of Investigation, FBI, is attempting to employ the faces from the driver's licenses to build a database that could include half of the faces of the US population. This could be useful as a way to track criminals across the country.
- **Marketing.** Organizations have taken advantage of this technology as a useful and cost-effective way to segment their campaigns using facial recognition systems in order to determine the gender, age and other demographic data of a client to be able to offer more concrete and "refined" advertising.

- **Healthcare.** Thanks to facial recognition, medical professionals can analyze certain characteristics of the patient and detect numerous diseases, including some rare ones. Obviously, this fact cannot neglect the intervention of actual doctors, so it should be taken as an aid.
- **Entertainment.** The film industry has also used facial recognition systems to study the emotional reactions of the film audience in order to know if their emotions match what the content creators want. Therefore, after receiving these comments, they can satisfy the pleasure of the spectators.

These are only some the multiple ways to implement facial recognition systems to our daily life. This technology is still in development so it has an incredulously high growth rate, so it is expected than in the next decade there are even more ways, and more direct at that, in which this technology will impact the society.

Chapter 4: Concerns and challenges

Nevertheless, not everything is going smooth about the implementation of facial recognition systems. Even if this technology surged with the objective of making or lives more comfortable and safer, there are serious privacy concerns that come from the use of this technology.

It is true that this technology can help security forces to identify criminals, but in order to do that the system has to identify all the people in its field of action.

One of the main causes of concern is that there is no law that regulate only this kind of technology so there really isn't a consensus about what the security departments or other companies can and can't do. This way, at present, the law that regulates this data is the general protection regulations established in Organic Law 15/1999, of December 2013, on the Protection of Personal Data and in Regulation (EU) 2016/679 of the European Parliament and of the Council, of April 27, 2016 (General data protection regulation), whose article 4-14 defines as biometric data those personal data obtained from a specific technical treatment, related to the physical, physiological or behavioral characteristics of a natural person that allow or confirm the Unique identification of said person, such as facial images or fingerprint data.

Therefore, Article 9-1 of the General Data Protection Regulation states that the processing of biometric data intended to uniquely identify a natural person is prohibited, since the treatment of photographs is considered the processing of personal data when the fact of being treated with specific technical means allows the univocal identification or authentication of a natural person.

At this point, there are two major users of this kind of data: the security forces and the common companies. In the first group, as we have talked about before in the applications of this technology, facial recognition has become increasingly important for security, as it has been implemented in some airports, stations and places with crowds of people. But here comes the second major cause of concern, the reliability of the data. For example, the FBI's facial recognition program is questioned as considered potentially dangerous due to its high margin of error, while it has been labeled as racist by society sectors because it increases its failures with black people [7], this concept is, of course, misused, because technology is not

racist, the problem is that the databases are not complete enough for all races. And it is in this plane where facial recognition generates more doubts from the legal point of view, since the privacy of people can be seriously threatened for the sake of national security, since it is a subject excluded from the scope of the General Regulation of Data Protection by the statement (16).

Notwithstanding the foregoing, the personal data files of the Ministry of the Interior must be governed by Order INT / 2287/2014, dated November 25; and, in any case, article 23-3 of Law 36/2015, of September 28, on National Security establishes that a situation of interest for national security cannot in any case imply the suspension of fundamental rights and public liberties of citizens, so that the rights to privacy and the image protected by article 18 of the Spanish Constitution and Organic Law 1/1982, of May 5, on Civil Protection of the Right to Honor, must be respected , to personal and family privacy and the image itself.

In the case of commercial uses, such as in Facebook or other Social Networking Services, the employment of this facial data must be done under explicit consent of the user. In cases like this, the fact that the user has given his express consent for the treatment of the image would exclude the violation of his rights, since it would be within the exceptions provided, respectively, in article 2-2 of the Organic Law mentioned above 1/1982, and in Article 9-2-a) of the General Data Protection Regulation, although the user must be correctly informed about the processing of his personal data, including those acquired from his facial features.

The other big reason of concern is, as has been briefly mentioned before, the reliability of the results. An independent study [6] from the University of Essex (United Kingdom), has stated that the London police facial recognition system is wrong in 81% of cases when suspects are reported. London police have tested the system in up to 10 different locations, but with 42 alerts, the system has only been successful with 4 suspects. Also, as we have commented before, this kind of systems are used in USA too, in the case of the algorithm used by FBI, a report presented in 2017 at the United States House of Representatives [7] denoted that the facial recognition software used by the FBI could be a potential danger for US citizens, since its margin of failure is too high and fails even more with black people.

The issue with the reliability may be a surprise for some people, since the test results that are shown in most papers are incredible. For example, FaceNet [8], the solution presented by Google, presents a classification accuracy of 98.87% on the Labeled Faces in the Wild (LFW) dataset. DeepFace [9], another well-known method, implemented by Facebook engineers, achieve a classification accuracy of 97.35% on the same dataset. The reason for this difference in behavior is that the test conditions and the real life are vastly different.

And this bring us to the next point of discussion, the challenges this technology will have to overcome in the future to become a real option in high specs applications, and they are closely related with what the concerns we have explained before.

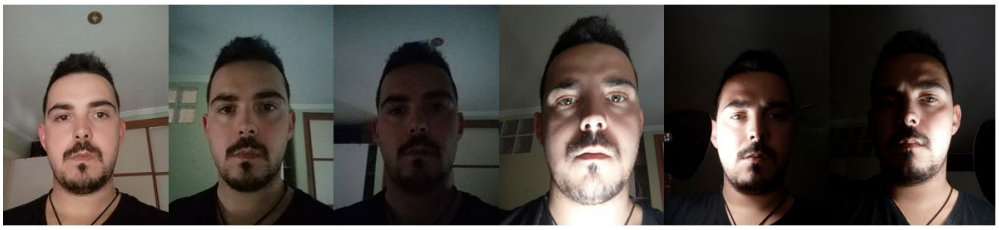
The point concerning the privacy issues is without a doubt something that must be faced by the political organizations in a global scale, there must be laws that find a middle point between security and privacy, but this is not something that should slow the developments of new techniques or approaches. We, as researches have the mission to develop new and better options that improve the ultimate results, and that is what we should concern about at this level.

And for that matter, we have to face the technical challenges that make this technology have so different results between in real life and test performances. These challenges, as commented before, are related to the difference between the test conditions and the real-life environment. This include two factors [10], the image acquisition and the imaging conditions and can be further developed in several individual challenges:

- **Illumination variation.** Lightning and camera characteristics affect to some degree the digital print of the human face. This changes in lightning can come from artificial means like room lightning or because of natural causes such as the time of the day. This can be controlled to some point employing image preprocessing techniques.
- **Pose/viewpoint.** The images of a face vary because of the relative camera face pose and some facial features such as the eyes or nose may become partially or wholly occluded affecting the performance. Usually, in tests, the faces are recorded in nice and controlled ways, but in real life, the

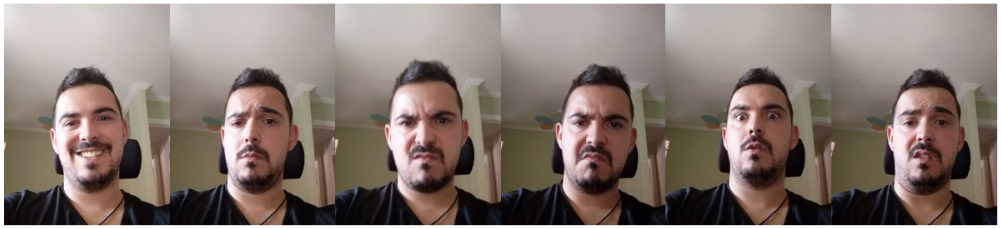
distance and angle to the camera are not controlled at all, and this is one of the main causes for the difference between the performance in real life and in tests.

- **Ageing and wrinkles.** Ageing can be natural (because of age progression) and artificial (using makeup tools). In both cases, ageing and wrinkles can severely affect the performance of face recognition methods. Nevertheless, this is not a point that impact so much the performance because it is easily faced if the database is updated regularly.
- **Facial expression/facial style.** Facial expressions, especially, the overreacted ones, affect directly to the image that is received of the face, and they can be very different between each other. A rich database can prevent this problem.
- **Occlusion.** This point affect at the quality of the image perceived in a way that this can be partially occluded by other objects, so we do not have a clear image to process. There is not much we can do about it except think carefully about where to locate the cameras so there are not many problems in the future because of this matter.



a

b



c

d

Figure 4.1. Examples of challenges: a) Illumination variation; b) Pose/viewpoint; c) Facial expression/facial style; d) Occlusion

Chapter 5: Face Recognition

5.1 Techniques

There are many techniques for recognizing faces from images/video, next we will explain some of these techniques [11,12] and then focus on the main theme of this report, the system we have developed.

- **Template-based Approaches.** Consist on comparing an input image with difference set of templates. For this method, statistical tools such as Principal Component Analysis (PCA), Support Vector Machines, Linear Discriminant Analysis, Independent Component Analysis can be used for the construction of the templates. The recognition function uses correlation or distance measures algorithms to find a match between the features extracted and the stored templates.
- **Holistic Approaches.** This method considers the entire face as a unique feature for detection and recognition, so individual characteristics such as eyes, mouth and nose are omitted when comparing similarities.
- **Model-based Approaches.** The image is treated as a high-dimensional vector. In this method statistical techniques are very useful to derive a space of features through the use of image distribution. The human face is modelled and is expected to be compared with the training set. It is necessary to adjust the sample given in the model, and different model parameters are very useful for recognizing different face images.
- **Statistical approach.** Each image is represented in terms of n characteristics that are displayed as a point in n dimensional spaces and are compared to the values of the database images. This is very expensive computationally, so it is very important to choose the correct statistical tool.
- **Neural Network Approach.** This is the latest approach and the one used in this work. There are many ways to apply this concept to face recognition and they get very good results because Neural Networks simulate the way neurons work in the human brain, achieving high performance when trained correctly.

Before beginning to explain in detail the method used in this work let's review the typical flowchart in a face detection and recognition system.

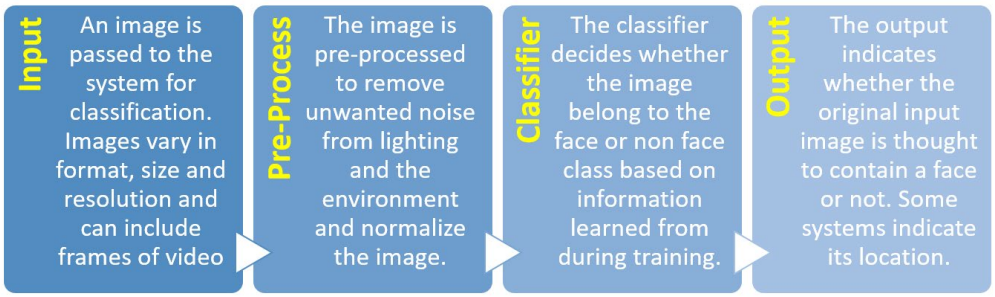


Figure 5.1.1. Face Detection System flowchart.

That is the general scheme for this kind of systems, nevertheless, changes may be applied depending on the specific method used.

5.2 Our approach

5.2.1 Related work

In our case, the Facial Recognition System follows the Neural Network approach and is inspired by the work done by David Sandberg and presented to the public on GitHub [13]. The work presented by David Sandberg is inspired from the GitHub source from OpenFace [15] and is a TensorFlow implementation of the face recognizer FaceNet [8] from Google with some ideas from Deep Face [14] from the Visual Geometry Group at Oxford. The project has been developed under Windows 10 with Python 3.6 with Tensorflow r1.7.

The method presented in [8], FaceNet, done by Google engineers, is based on learning a Euclidean embedding per image using a deep convolutional network. The network is trained in a such a way that the squared L2 distances in the embedding space directly correspond to face similarity: the faces of the same person have small distances and the faces of different people have large distances.

Once this embedding has been produced, the recognition task becomes easier as it becomes a k-NN classification problem. This method was originally developed with the intention of developing a unified system for verification (is this the same person), recognition (who is this person) and clustering (find common people between these faces). Nevertheless, the objective of this project focusses only on the recognition

in real time. Nonetheless, the addition of these two other tasks would be easy to implement since once the embeddings have been produced, the face verification simply involves thresholding the distance between two embeddings and clustering can be achieved using off-the-shelf techniques such as k-means or agglomerative clustering.

FaceNet directly trains its output to be a compact 128-D embedding using a triplet-based loss function based on large margin nearest neighbor (LMNN). The triplets consist of two matching face thumbnails and a non-matching face thumbnail and the loss aims to separate the positive pair from the negative by a distance margin. The thumbnails are tight crops of the face area, no 2D or 3D alignment, other than scale and translation is performed.

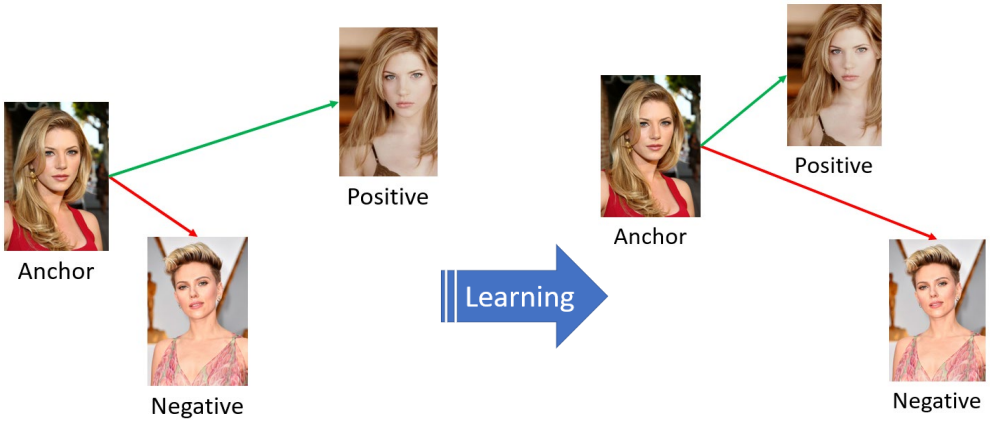


Figure 5.2.1.1. Triplet-based loss function.

5.2.2 Our model

The model used for this project has not been trained by us, we have used the one presented in [8]. There are two reasons for this:

- It is a state-of-the-art model in terms of performance. The accuracy presented on the LFW dataset is 99.65%.
- This model has been trained using the VGGFace2. This data set contains +9000 identities with +3.3 million images. To train the model with this data set, a Nvidia Pascal Titan X GPU has been used together with a high-end CPU. In our case we did not have that kind of means and for matters related to deep learning, a good GPU is essential. According to [8] the time needed to train the model was 10 hours, we also tried to train our own model, but we

had to reject this idea since for our means, this time would have been 31 days even using a smaller dataset.

This model has not been trained using the triplet-based loss function as in FaceNet but a softmax loss function. Since the use of the triplet-based loss function can be problematic and quite complicated, they decided to do so. And the performance, as we can see in Figure 5.2.2.1, is not worse than that of FaceNet.

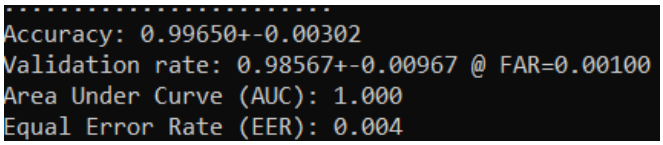


Figure 5.2.2.1. Validation on LFW.

Another difference is the architecture of the convolutional neural network used, in this case, the Inception-ResNet-v1 model was used instead of the non-ResNet version used on FaceNet. As noted in [13], the use of the ResNet architecture solves several convergence problems that are shown when training in the CASIA / Facescrub data sets, also the results obtained in LFW show improvements.

In facial recognition applications, there are two kind of target populations. The first is a large-scale group of the order of the population of a city, in these cases there are no problems when training the model and the same population can be used for that task. However, the second target group is smaller, usually on the order of 50 to 100 people, belonging to a company or similar. For these cases the training cannot be done on the same population because millions of data are required to obtain a well-trained model.

The best way of solving this problem is employing the one-shot learning technique. One-shot learning aims to learn information about object categories from one, or only a few, training images. The model still needs to be trained on millions of data, but the dataset can be any, but of the same domain. In this way, we can use the frozen model presented in [13] and be sure that the results obtained will be of excellent quality.

Chapter 6: FaceReco – Real time implementation

We have developed a real time face recognition application able to switch easily between different datasets thanks to the one-shot learning technique, this way it can be employed for different purposes. As noted before the development has been done with Python under Windows 10 and for the GUI, kivy has been used.

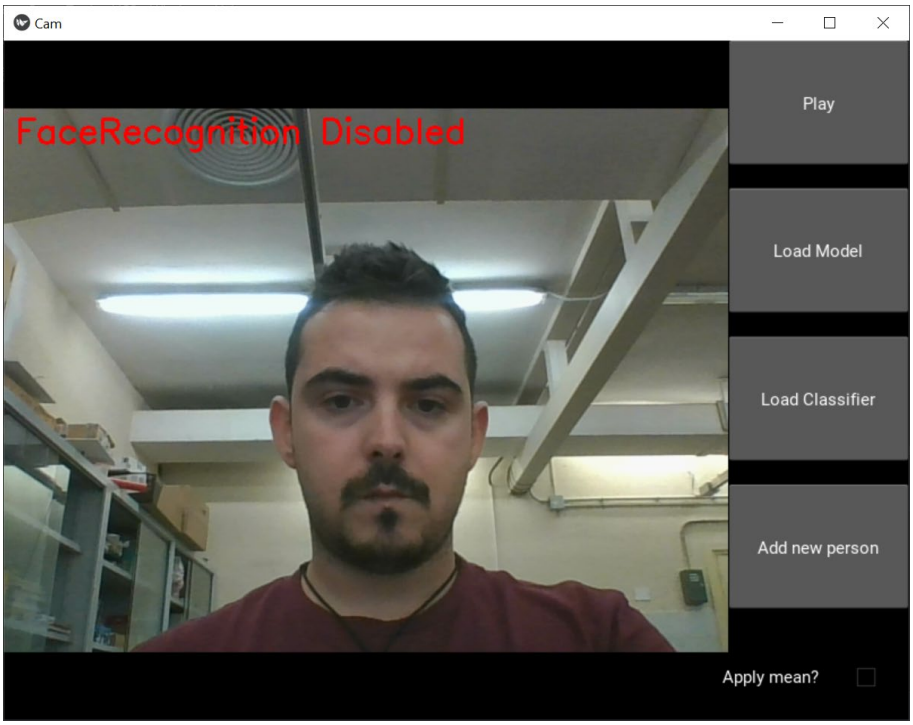


Figure 6.1. FaceReco GUI.

The operation process followed by our application is as follows:

Figure 6.2. Working flow.

Those are the three stages in which can be divided the recognition process.

INPUT

We use a camera to capture in real time the images that will be processed before the recognition stage. The resolution of the camera is not important, the program

will capture the image at the natural resolution of the camera and later will be pre-processed to follow the requirements needed by the recognizer.

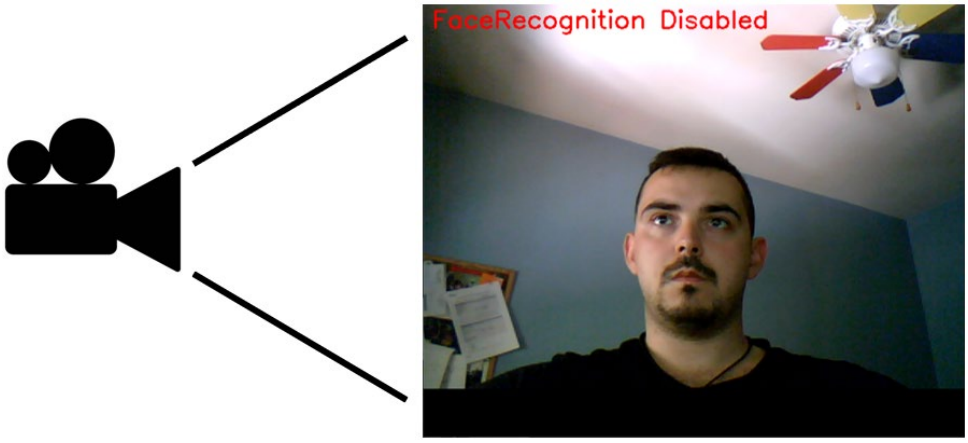


Figure 6.3. Input images.

MTCNN

The second stage represents the preprocessing stage. Nevertheless, this typically important process for projects related to images has been kept to the minimum in this project. The only work done in this stage is the detection of the faces that will be tightly cropped to be adjusted to the same size as the faces used during training, 160x160 pixels. This work is done by the Multi-task Cascaded Convolutional Networks (MTCNN), which is essentially several convolutional networks strung together in order to detect landmarks that are used to detect individual faces. This give us a bounding box around the face that will be cropped and resized to the requirements before entering the recognition stage. No alignment will be done in contrast with other face recognition systems in order to reduce the complexity and increase the versatility of the application.

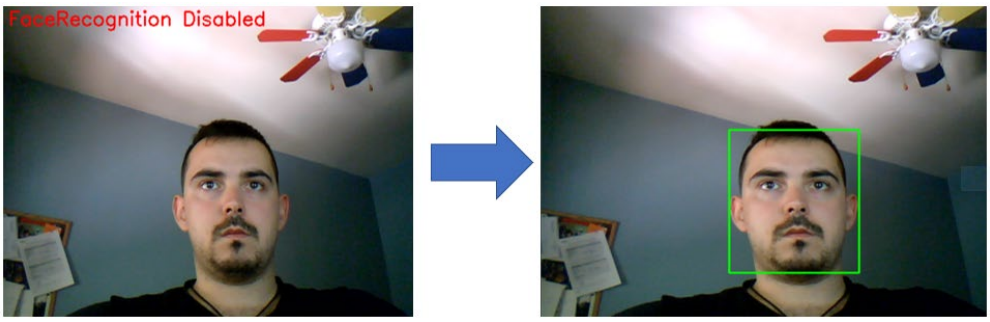


Figure 6.4. Bounding box introduced by MTCNN.

This same work is done before training the model on the dataset selected for training.

Figure 6.5. Examples of MTCNN on LFW images.

RECOGNITION

The last stage of the process is the recognition. As explained before, for this task we will build a classifier from the original frozen model employing the one-shot learning technique. This way, we will be able to difference between target populations in a fast and reliable way. This is completely necessary for this kind of applications because the full training time of a model from scratch requires a lot of time, for a higher end computer with the latest GPU, this time is up to 10 hours, so this is not a reliable way to achieve a real time face recognition system, since for every new class/person that we want to include in the system, the model would need to be trained again. In contrast, the use of a classifier greatly speeds up the process, so the training of the classifier is only a matter of seconds or minutes. So, we can have one different classifier for each target population we want to recognize.

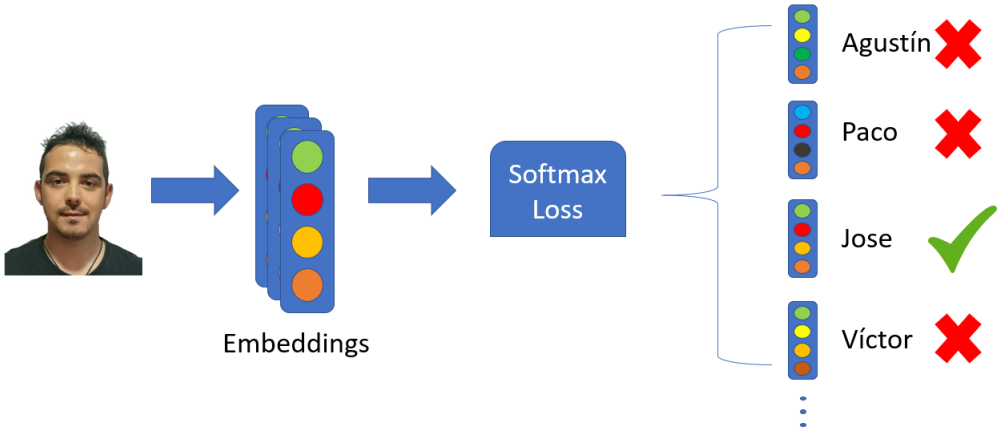


Figure 6.6. Recognition stage.

For this classifier to be built, the user will have to select the folder which contains the classes/people to be selected as part of the system. Each class should have at least 10 images so the classifier can be correctly trained. It is very important that all the classes have a similar number of samples and that they are taken in a similar way, since if it is very unbalanced this could affect the performance.

These images can be obtained directly from the application by taking photos or the user can select a folder that include those photos previously taken, the images used to train the classifier must contain only the class/person that want to be included in the system so there are no problems when locating the face. Another good practice would be to use different expressions in each photo so the system can be more robust to changes on the face.

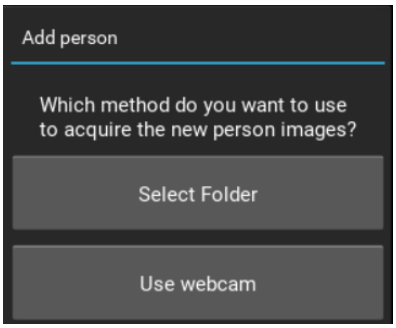


Figure 6.7. Class input methods.

When one person is introduced into the dataset the system will preprocess the images with MTCNN so the faces are cropped and ready to train the classifier.

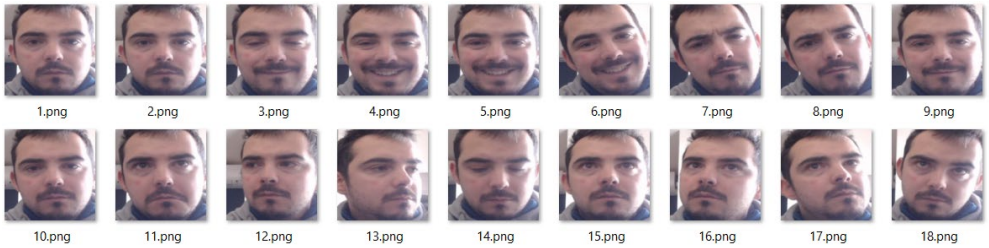


Figure 6.8. Folder containing the class images ready to train the classifier after being processed.

When the class is correctly introduced to the database the system will ask the user to check that the images have been properly preprocessed, this step is due to that the MTCNN is an automatic face detector and there are times when the face can be incorrectly detected, this fact could affect the performance so we think that a fast manual check can be useful.

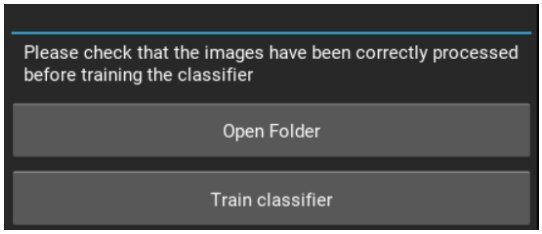


Figure 6.9. System inquiry after adding a class to the database.

If everything is correct the user can choose to train the classifier with the new class or abandon the process to add a new person or whatever the user requires.

The system will calculate embeddings for the inputted images in real time and compare them with the ones in the dataset obtaining a confidence probability factor that will be used to choose the class/person recognized. This confidence factor is a very important value in the system and has to be regulated with a threshold in order to avoid false positives, since the system will output the class with the highest confidence probability factor if a threshold is not set, so unknown people will not be ruled out but the system would use the class with the highest probability.

Chapter 7: Performance

When talking about performance we have to differentiate between the performance of the frozen model used on the project which we briefly mentioned before, and the performance of the application.

The first one has to be considered employing datasets and we can obtain a quantitative value of this performance, nevertheless, the performance in a real time environment is much more difficult to test but we will try to offer an adequate view of the results.

7.1 Validation on LFW dataset

Labeled Faces in the Wild is one of the most common datasets designed for studying the problem of unconstrained face recognition. It has become in one of the most used when talking about performance of a model. This database was created and maintained by researchers at the University of Massachusetts. 13.233 images of 5.749 people were detected and centered by the Viola Jones face detector and collected from the web. 1.680 of the people pictured have two or more distinct photos in the dataset.

In order to test the model with the dataset we have to pre-process it using MTCNN to obtain the 160x160 pixels images with the face cropped at the center.

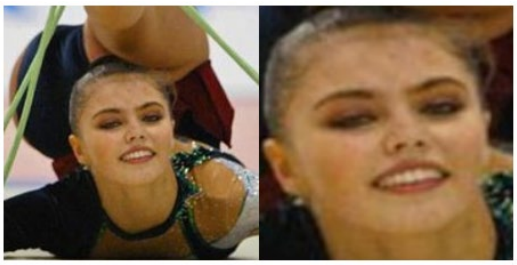


Figure 7.1.1. LFW image pre and post processed.

Then, we employ a script provided by [13] designed for testing datasets to obtain the performance results of the model. This script let you select the model to test and the dataset to be employed. It will divide the dataset in batches that will be used to train a temporal classifier that will be used to validate image pairs while recording the statistics.

The results obtained for the LFW dataset are shown in the Table 7.1.1:

	Model from [13]	FaceNet [8]	DeepFace [9]
Accuracy	0.99650+-0.00252	0.9963	0.9887
True Positive Rate	0.98367+-0.00948 @ FPR=0.00100		
Area Under Curve	1.000		
Equal Error Rate	0.004		

Table 7.1.1. Performance results in LFW dataset.

As we have commented before, FaceNet and DeepFace are two of the best facial recognition systems at the moment, employed by Google and Facebook respectively. So, looking at the results of the model we have used to build our real time face recognition system we see that the performance obtained is the same of a state-of-the-art model.

Before we comment about the different parameters used to evaluate our model let's define first what do they mean. In the first place we have to know that in machine learning and specifically the problem of statistical classification, a confusion matrix [16], also known as error matrix, is used to evaluate the performance of an algorithm, it is a table that distribute the different possible outcomes when classifying an event.

		ACTUAL CLASS	
		POSITIVE	**NEGATIVE**
PREDICTED CLASS	**POSITIVE**	TRUE POSITIVE A positive was correctly predicted	FALSE POSITIVE A negative was taken as positive
	NEGATIVE	FALSE NEGATIVE A positive was taken as negative	TRUE NEGATIVE A negative was correctly predicted

Table 7.1.2. Confusion matrix.

Considering the parameters presented at Table 2 we can define many performance measures but in facial recognition systems the most important values are:

- **Accuracy.** A statistical measure that defines how well a binary classification test correctly identifies or excludes a class. That is, the accuracy [17] is the proportion of true results (both true positives and true negatives) among the total number of cases examined.

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{True\ Positives} + \boldsymbol{True\ Negatives}}{\boldsymbol{True\ Positives} + \boldsymbol{True\ Negatives} + \boldsymbol{False\ Positives} + \boldsymbol{False\ Negatives}}$$

- **True Positive Rate (Sensitivity).** Measures the proportion of actual positives that are correctly identified as such.

$$\textbf{Sensitivity} = \frac{\boldsymbol{True\ Positives}}{\boldsymbol{True\ Positives} + \boldsymbol{False\ Negatives}}$$

- **False Positive Rate (Specificity).** Specificity measures the proportion of actual negatives that are correctly identified as such.

$$\textbf{Specificity} = \frac{\boldsymbol{False\ Positives}}{\boldsymbol{False\ Positives} + \boldsymbol{True\ Negatives}}$$

The measured value presented in Table 1 (TPR@FPR=0.001) means the rate that faces are successfully accepted when the rate that faces are incorrectly accepted is 0.001.

- **Area Under Curve (AUC).** A ROC curve [18] (Receiver Operating Characteristic) is a graph that shows the performance of a classification model at all classification thresholds. A ROC curve represents TPR against FPR at different classification thresholds. Reducing the classification threshold classifies more elements as positive, so that both false positives and true positives will increase.

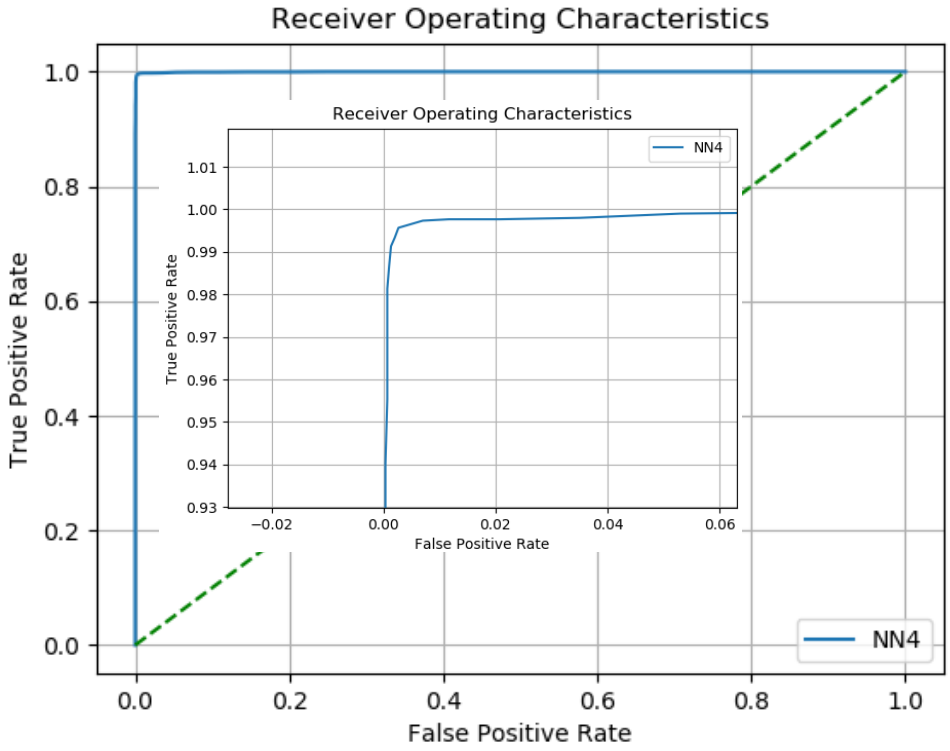


Figure 7.1.2. ROC Curve. It is observed how the TPR is almost 1 for all the range.

AUC measures the entire two-dimensional area below the full ROC curve. The AUC provides an aggregate measure of performance at all possible classification thresholds. AUC of a classifier is equal to the probability that the classifier will rank a randomly chosen positive example higher than a randomly chosen negative example.

The AUC ranges from 0 to 1. A model whose predictions are 100% incorrect has an AUC of 0.0; another whose predictions are 100% correct has an AUC of 1.0.

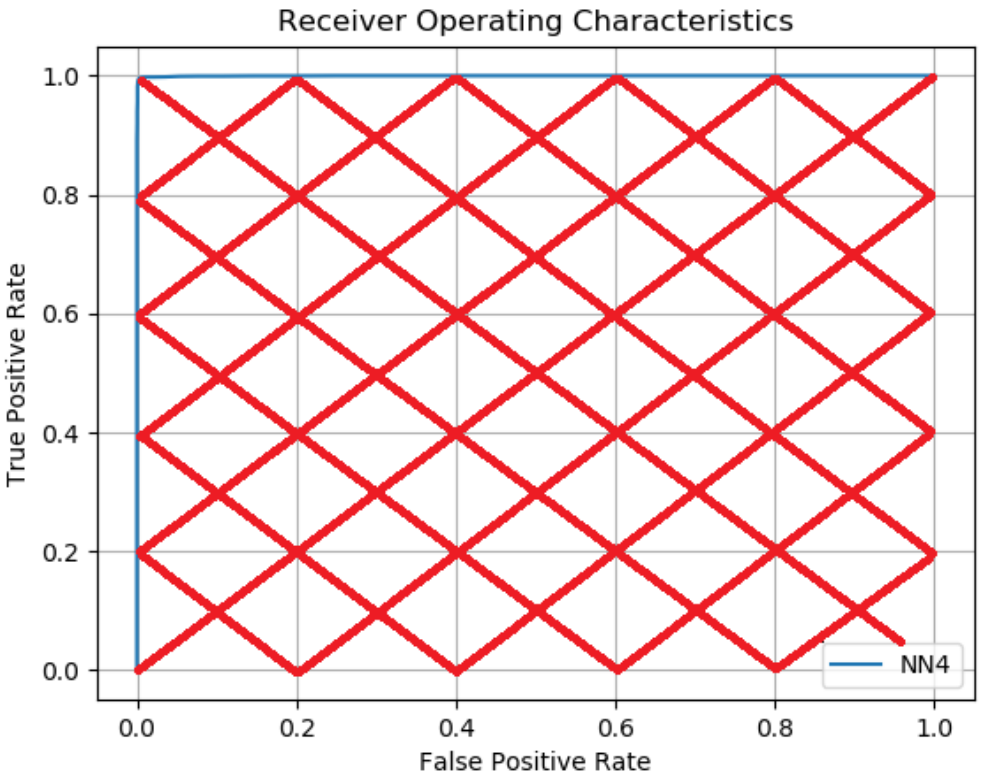


Figure 7.1.3. Area Under Curve for LFW dataset test. It is observed how everything is under the curve giving a AUC = 1.

The AUC is convenient for the following two reasons:

- The AUC is invariable with respect to the scale. It measures how well the predictions are classified, rather than their absolute values.
- The AUC is invariable with respect to the classification threshold. It measures the quality of the model predictions, regardless of which classification threshold is chosen.

Due to these reasons the AUC is maybe the best way to measure the quality of a facial recognition model.

- **Equal Error Rate**. The EER [19] is related to the sensitivity and the specificity and is an algorithm that helps us to determine the threshold value for the TPR and the FPR. When both rates are equal, the common value is referred to as the equal error rate. The value indicates that the proportion of false acceptances is equal to the proportion of false rejections. The lower the equal error rate value, the higher the accuracy of the system.

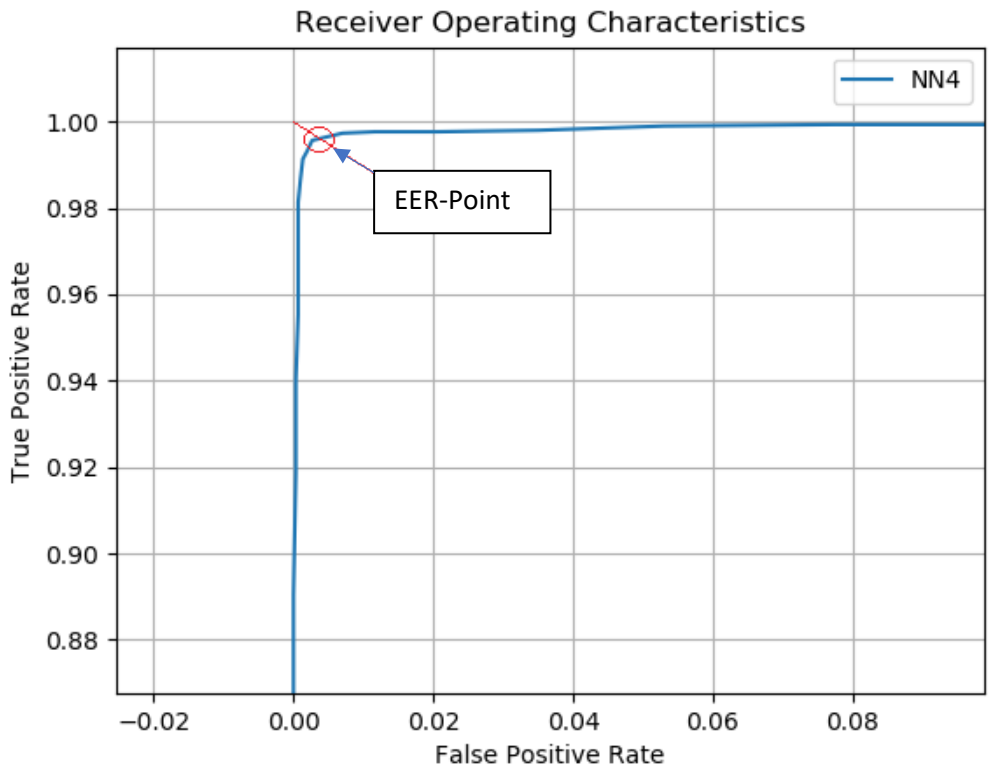


Figure 7.1.4. Equal error rate for LFW dataset test. It is observed how close to zero is the EER value, meaning the recognition is almost perfect.

Once we understand these parameters, we can comment on the results obtained by validating the model in the LFW dataset.

These results are outstanding, the accuracy is almost 100% and more important, when the False Positive Rate is kept at 0.1% the True Positive Rate is 98.37%, this means that the model almost never makes mistakes on this dataset. This can also be seen with the Area Under Curve that shows a value of 1.000.

The results obtained make us think that the system works very nicely and reaffirm our thoughts of using this model for our real time face recognition system.

7.2 Validation on CASIA-WebFace subset.

In order to achieve a better understanding and confirm the quality of the model, we ran another validation test, this time with a random subset of the CASIA-WebFace dataset.

This dataset is composed of 10575 unique people with 494.414 images in total. It is the second largest public dataset available for face verification and recognition

problems. Nevertheless, it is too big to run in our computer, so we created a subset of this dataset composed by 720 classes and 64.184 images. The validation test using this dataset took 12 hours and the results obtained can be seen in Table 3

	Model from [13]
Accuracy	0.91215 +- 0.00230
True Positive Rate	0.65549 +- 0.00474 @ FPR = 0.00099
Area Under Curve	0.951
Equal Error Rate	0.099

Table 7.2.1. Performance results in CASIA-WebFace custom subset.

Looking at the results we observe that differ a little with the obtained before in the LFW dataset.

The accuracy is still over 90% which is a very good value but the True Positive Rate when the False Positive Rate is 0.0099% has been reduced to 65%. This is worse than before, but it is not a bad value because the FPR target is very low. This TPR value indicates how many image pairs the detector can correctly identify as the same identity while keeping the FPR (FPR, i.e. the probability that two images of different identities are the same) to 0.099%. Since the detector needs to be more certain, the TPR will always be lower than the accuracy.

Nevertheless, although these are good results, we wanted to find the reasons to explain the difference in performance between this custom dataset and the the LFW dataset.

To reach an explanation we proceeded to review both datasets to see possible reasons and found that the CASIA-WebFace dataset, or at least, the subset we randomly chose presented several possible reasons to explain the results:

- **Noisy dataset.** There are many low-resolution images, so this will difficult the detection and verification of the classes.

- **Mismatched images.** We found many mismatched images in many of classes in the subset used. There were people wrongly placed on the classes folders and even men were found in female classes and vice versa.
- **Images from very different ages.** The images found for many famous people were very apart in the time in the same class, we found differences of 30 or more years between images of the same people. This causes problems at the verification stage because there are cases in which the difference between images is too big, and the scope of the class is too large.
- **Side pictures.** Another possible reason is the presence of many side pictures in the chosen subset. This kind of images will not be correctly classified by the model because it has been trained mainly in front or slightly rotated faces. The side pictures are too different and do not contain the information needed to perform an accurate identification.
- **Presence of more than one person in an image.** There are cases that the image presents two or more people. This is not acceptable to train a classifier because we are trying to recognize only one person and this fact affects the performance.

Figure 7.2.1. Examples of the reasons found for the performance difference.

As seen in Figure 7.2.1, the problems mentioned before can happen all of them in the same class, this will affect the good behavior of the system and could explain the results obtained in the tests. This is not a rare case or an exception to the dataset. We do not have to try hard to find more of these issues with the dataset.

Figure 7.2.2 More mistakes in the dataset.

In almost all the classes there are some kind of problems of the previously mentioned. The performance does not decrease further because it is a dataset that has an average of 50 images for class, so the impact of individual errors does not alter the performance too much.

After exploring and studying the dataset, which is widely used in applications related to face recognition, needs to be cleaned. The presence of mistaken pictures will affect the performance of whatever application the user is developing.

Taking into account the possible reasons presented before, we can explain the performance difference between both datasets, so the results obtained, that were pretty accurate to begin with, should be taken as even better because these reasons affect very much to the True Positive Rate as this is a parameter very sensible to single recognition mistakes since the False Positive Rate target is kept very low in order to have a very safe recognition system.

7.3 Real-Time recognition on closed set (few people).

To evaluate the performance of the real time face recognition system we have performed a series of test in different scenarios.

In the first place, we have supposed a group target with few people and have tested the ability to recognize objective people and discard unwanted people.

7.3.1 Recognizing

The group target this time is composed of 4 people, 2 men and 2 women. The objective in this test is to correctly identify one person of this group. Different situations will be tested in order to evaluate the robustness of the system.

The members of this group can be seen at Figure 7.3.1.1.

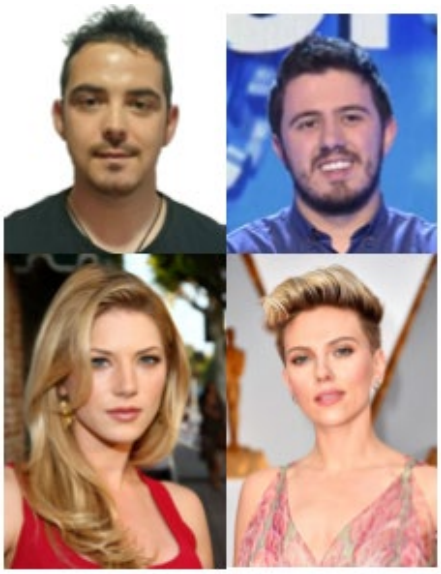


Figure 7.3.1.1. Target group for closed set test.

For this kind of tests, it is difficult to obtain a quantitative value for the performance. So, for this case, we will show different images that will serve as samples of different scenarios.

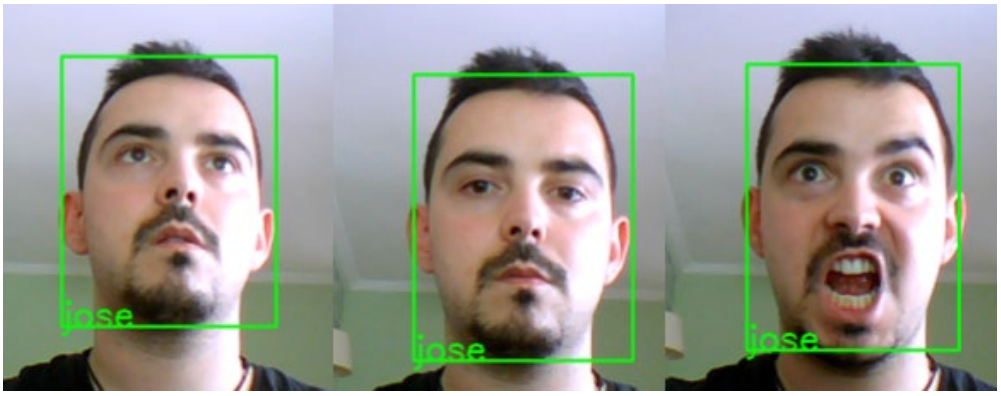


Figure 7.3.1.2. Test results – Clean line of sight, different expressions.

We observe from this captures that the system is robust when facing different kind of expressions or when looking to another places. Nevertheless, not always the line of sight to the face is clean and can be partially blocked by objects such as glasses or our own hands.

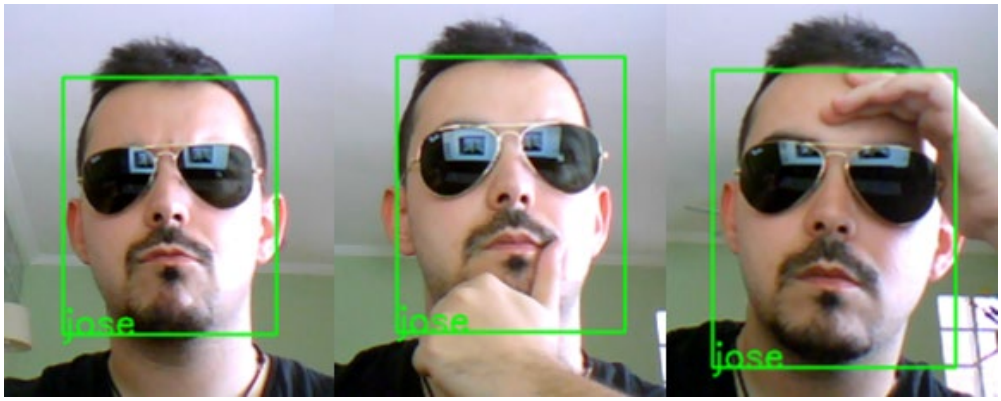


Figure 7.3.1.3. Test results – Partially blocked line of sight.

The previous results have not been affected by the inclusion of new elements like the glasses or the hands covering part of the face. The system keeps offering accurate predictions not mistaking at any moment the person to recognize.

Next, we present two extreme cases when little of the face is shown to the camera.

Figure 7.3.1.4. Test results – Probing the limits.

Nevertheless, this can be considered as the system's limit when facing few subjects, because the recognition is not constant and there are frame where no face is detected.

7.3.2 Not recognizing

For the next test we will replace the person called “jose” for another one to test the capacity of the system to not recognize unwanted people.

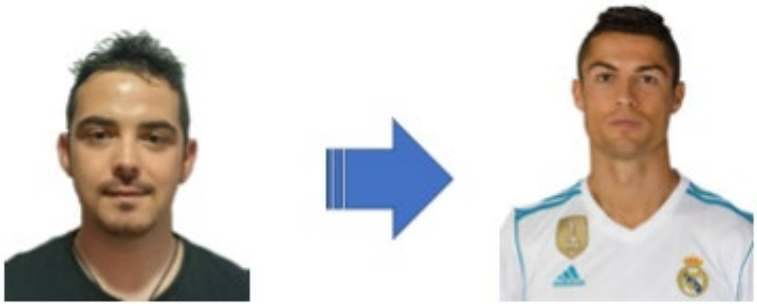


Figure 7.3.2.1. Changing the subject.

Figure 7.3.2.2. Test results – Testing the capability to not recognize people.

In this test we have probed the viability to use this system as auxiliary in scenarios where few people are needed to be recognized. The system does not present any problem in recognizing the people it has to recognize or in not recognizing people who should not be recognized.

We have to note that when using few people, the confidence probability factor is very high because the different between the classes is greater. This factor stays at 0.80 when faced directly and at 0.65 when trying to force the recognition. For this test the lower threshold limit to stop trying to recognize someone has been set to 0.6.

7.4 Real-Time recognition on open set (many people).

This test will be performed using 30 classes, simulating a medium sized company. The objectives and the planned test are the same as in the previous case.

In figure 7.4.1, we observe the people that compose the group target.

Figure 7.4.1. Target group for open set test.

7.4.1 Recognizing

The first part of test will consist on recognizing the target.

Figure 7.4.1.1. Test results – Clean line of sight, different expressions.

The system keeps recognizing the target, nevertheless, this time the margins are narrower so there are frames in which the system makes mistakes confusing the objective. This happens because when there are many people, the confidence probability factor decreases a lot, from 0.6 that we saw in the previous test to 0.1, so the difference between people is narrower which incites eventual mistakes. In

order to improve this aspect, we have implemented a mean-based strategy between 5 frames, so the system gives the prediction based on the last 5 frames, selecting the class that has more appearances. This makes the system more robust because it faces directly the problem we encountered. Nevertheless, there is a drawback to this method. The maximum number of recognized people in every frame is reduced to 1, because it is not possible to keep an accurate tracking if people keeps entering and leaving the camera area. Even so, we think one person is enough if we want a more robust system, so two versions have been developed, one with the mean-based strategy for one target, and one without it for multiple targets, in which the recognition is done.

Nevertheless, the use of the mean-based strategy does not resolve the inherent problem, that the confidence probability factor is very similar between subjects. The target person to recognize is almost always in first place but the difference with the rest of classes is very small. This fact affects overall to the selection of the confidence probability factor, this makes more difficult to correctly select a good value to this parameter that let us differentiate between “wanted” and “unwanted” people. We have chosen 0.09 as limit value after probing with different values. After applying linear regression to these two cases, we have defined a range of values to automatically change the confidence factor limit based on the number of classes that contains the target group.

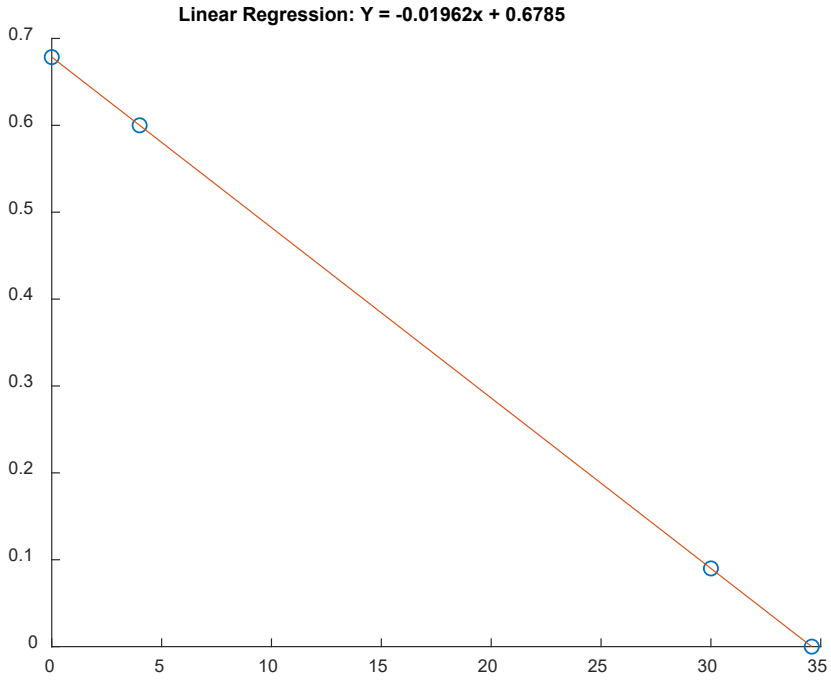


Figure 7.4.1.2. Approximated curve.

This cannot be taken as definitive and we consider that further studies should be done about the confidence probability factor.

Continuing with tests, the next one will probe the limit when aiming at a subject within a crowded group target. This limit is much lower than in the previous case, when the system was able to correctly recognize the target even in very difficult situations.

Figure 7.4.1.3. Test results – Probing the limits.

We note that for two similar frames, the system is not able to carry a constant recognition when wearing glasses, frequently switching between the target and the "NotRecognized" classes.

7.4.2 Not recognizing

To test the system when it should not recognize a person as target, we have removed the target class we have used until now.

Figure 7.4.2.1. Class removed.

So, after removing the class we test again the system over the new group target without the class removed

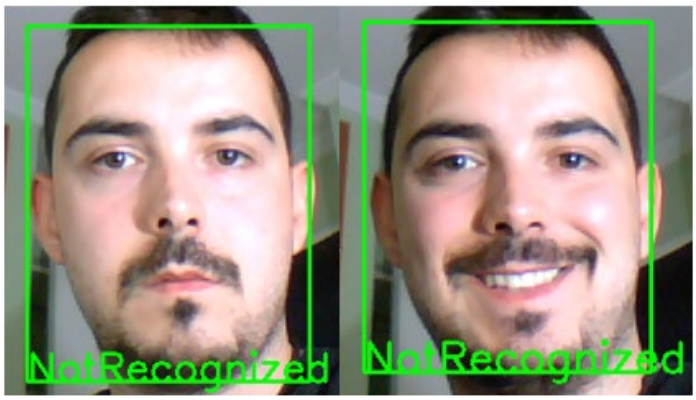


Figure 7.4.2.2. Test results – Testing the capability to not recognize people.

There are not any problems when testing the capability. The mean-based strategy helps the system maintain constant recognition by avoiding the introduction of incorrect recognitions.

Chapter 8: Conclusions and Future Lines of Research

Once this thesis is concluded, we can affirm that the results obtained for our real time face recognition system are very successful. We have developed a very versatile system that allows the user to switch between classifiers (group of people to recognize) and models with ease. We provide the model presented in [13] as the default to use since this is a state-of-the-art model at the height of the best models used by Google [8] or Facebook [9]. Nevertheless, we do not exclude the possibility that the user wants to use their own model to compare results or the possibility that another model will be presented in the future even better than the current one. We always support versatility and do not want to inhibit the imagination of the user, so we did everything possible to not reduce any possibility to the user whenever possible.

This thesis has allowed us to improve our knowledge in a field that we have not studied before, becoming better people and researchers while we acquire new techniques and possibilities to face future problems. We believe that this has been an extremely good source experience for our student career and will be of use for our next steps, whatever they may be.

In this thesis we have concluded that real time face recognition systems are a reality not only for government agencies or major multinationals companies, but for the average user or company this approach can be a new turning point to, for example, control attendance of little to medium sized companies now that new Spanish law [21] for attendance control forces companies to keep track of employees' working hours. This can be a very good time to invest in new technologies in such a way that the worker is not affected since the system would be automatic and could also ensure the rights of the worker since this method cannot be avoided as it happens with RFID tags or similar whose security is at least questionable. This would be a good method to face overworking issues that Spain suffers right now. In 2018, according to the Economically Active Population Survey [19], 330 million of extra hours were made, which suppose the creation of 180.000 new jobs.

Our approach is not a final product focused in one sole purpose such as the commented before, but this project is intended to serve as a middle point or general

implementation to provide future users of a starting point to implement their own ideas.

In the future line we propose, the extra work needed to adjust is not much since it would mainly consist on including a log system and link it to the database that would be the same as the one for the classifier.

Back to the main topic of this project, we have to discuss some issues or points to improve in future works regardless the future line proposed. We have faced several issues with the graphic interface used for our Python implementation. In our case, after exploring different options, we found the best solution to be Kivy since is more modern with some minor Tkinter implementations for specific tasks. Nevertheless, after finishing this project we can assert that this kind of GUI interfaces are not fully developed and present some issues that can affect the user interactions. For this project, we do not face any issues that hinder too much the user, but there can be issues such as losing focus on the application interface after a task is finished or after the user is asked to input text requiring the user to click on the application screen to regain the focus again.

Also, the confidence probability factor needs further studies to better understand its value. We believe that, in order to improve the performance, the system administrator has to ensure that all the classes have near the same number of images and are taken in a similar manner. Otherwise, this could affect the performance of the system.

Apart from this, it is needed to remember that there are a couple of concerns with this technology that need to be faced in the next years, so these techniques keep growing and entering more in our daily lives.

- **Privacy.** As commented before, this is a topic that needs to be confronted in a global scale and governors and law enforcers have the obligation to find a middle point between data preservation and technological advancements because this affects the researchers and the evolution speed.
- **Technical.** It has been demonstrated that at this moment, this technology works really well when used in the right environment, when the camera faces directly to the objective, but it is severely affected when the angle

changes, and this is a technical issue that needs to be corrected in the future in hope of the advancement of this technology. In addition, the fact that proven systems such as the FBI and the London Police make mistakes can be explained since they are systems whose target groups are in the millions, making the recognition very difficult since the difference between the classes will be very small.

Excluding these two issues, we consider that this technology has come to stay, and we think it is a technology full of potential that will make our life easier and above all, safe.

References

[1] "Biometric Identification", Jain, A.; Hong, L. and Pankanti, S. Communications of the ACM, Volume 43 Issue 2, p. 91–98. February 2000.

[2] "History of Biometrics", Stephen Mayhew, July 2018 https://www.biometricupdate.com/201802/history-of-biometrics-2

[3] "Una breve historia del reconocimiento facial", Retrieved from https://medium.com/@spot_blog/una-breve-historia-del-reconocimiento-facial-vision-blog-5a76fdfe4865, March 2018,

[4] Image retrieved from: https://en.wikipedia.org/wiki/Eigenface

[5] "5 Applications of Facial Recognition Technology", Laura Cox, July 2017, https://disruptionhub.com/5-applications-facial-recognition-technology/

[6] "Facial recognition technology used by the Met Police is wrong in four out of five cases and could breach human rights laws, a new report warns", Chris Dyer, July 2019, Retrieved from: https://www.dailymail.co.uk/news/article-7211141/Facial-recognition-technology-used-Met-Police-wrong-four-five-cases.html

[7] Hearings from Committee on Oversight and Reform of the USA: https://oversight.house.gov/legislation/hearings/facial-recognition-technology-part-1-its-impact-on-our-civil-rights-and and https://oversight.house.gov/legislation/hearings/facial-recognition-technology-part-ii-ensuring-transparency-in-government-use

[8] "FaceNet: A Unified Embedding for Face Recognition and Clustering", Florian Schroff, Dmitry Kalenichenko, James Philbin, June 2015.

[9] "DeepFace: Closing the Gap to Human-Level Performance in Face Verification", Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, Lior Wolf, September 2014

[10] "Face recognition: challenges, achievements and future directions" M. Hassaballah, Saleh Aly, ET Comput. Vis., Vol. 9, Iss. 4, pp. 614–626, 2015.

[11] "Challenges and Advances in Human Face Recognition from Real Time Video", Ranjana Dahake, M.U. Kharat, Priti Lahane. International Journal of Advanced Trends in Computer Science and Engineering, Volume 5, No.6, November – December 2016.

[12] "Face Recognition techniques and approaches: a survey", Muhammad Naeem, Imran Qureshi, Faisal Azam. Sci.Int.(Lahore),27(1),301-305,2015.

[13] FaceNet repo from David Sandberg: https://github.com/davidsandberg/facenet

[14] "Deep Face Recognition", Omkar M. Parkhi, Andrea Vedaldi, Andrew Zisserman, 2015.

[15] OpenFace repo from cmusatyalab: https://github.com/cmusatyalab/openface

[16] Retrieved from: https://en.wikipedia.org/wiki/Confusion_matrix

[17] "Basic principles of ROC analysis" Metz, CE. Semin. Nucl. Med. 8 (4): 283–98. PMID 112681, October 1978.

[18] Machine Learning Crash Course with TensorFlow APIs from Google: https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=es-419

[19] Retrieved from: https://www.webopedia.com/TERM/E/equal_error_rate.html

[20] Instituto Nacional de Estadística – Encuesta de Población Activa 2018: https://www.ine.es/dyngs/INEbase/es/categoria.htm?c=Estadistica_P&cid=1254735976595

[21] BOE: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2019-2861

[22] "5 Applications of Facial Recognition Technology", Laura Cox, 2017, July 13th. https://disruptionhub.com/5-applications-facial-recognition-technology/