Face Recognition System on Mobile Device based on Web Service approach

Angelo Galiano[#1], Alessandro Massaro[#2], Donato Barbuzzi[#3], Mattia Legrottaglie[#4], Valeria Vitti[#5], Leonardo Pellicani[#6], Vitangelo Birardi[#7]

[#]Dyrecta Lab, Research Institute
Via V. Simplicio 45, 70014 Conversano (BA) - Italy

Abstract— **This paper proposes a new application for face recognition task on mobile device. The originality of the work is based on the web service, able to store the normalized distances in the database, to verify the correct building of the classifier's model and to match the test normalized distances with the training normalized distances already stored in the database. This innovative technique is named Bridge Approach (BA). The Bridge Approach allows both to speed up the processing and to use the face recognition system in all locations where an internet network is available. The application uses the OpenCV library for dynamic targeting of face, eyes, nose and mouth and WEKA library for the classification purpose. Experimental results are carried out on 500 face images belonging to 10 individuals (50 images each). Finally, the main goal of the work is to show the performances in terms of computational time and classifier's accuracy.**

Keywords— **Face Recognition System, Web Service Architecture, OpenCV, WEKA, Mobile device.**

I. INTRODUCTION

Face Recognition on mobile platforms represents an open challenge in the scientific community and many commercial applications are developed [1]-[3]. More specifically, these systems are widely used in the security [4]-[6] and marketing [7] fields, despite the main problems related to the: *image quality*, *illumination* and *pose variation*. In this paper, we discuss the face recognition based on 25 different distances, considering 14 landmarks between: eyes, nose and mouth. This system is simple and it can be run on a normal Android device.

Another important aspect about the implementation of this mobile face recognition system is the execution of some functions outside the mobile processor (Bridge Approach). Some authors have already considered the use of web service architecture as support to the calculation [8]. So, the operations are faster respect to the processing on the mobile device and the system is able to increase the biometric landmarks in order to improve the recognition accuracy at a lower computational cost.

As already discussed in the literature, best performance depends on features, classifiers and data distribution between samples in the training and test set [9]-[13].

Different algorithms can be used for the face detection and for face recognition. More specifically, the Haar-like features are extracted for face detection, using OPENCV tool [14], and normalized distances between landmarks in the: eyes, nose and mouth are used for face recognition.

All experiments have been carried out on real face images acquired by the built-in camera of the mobile device. A 7-NN classifier is used [15].

The paper is organized as follows: Section 2 presents an overview on the state of the art in face recognition system. Section 3 describes the method. The operating conditions and experimental results are presented in Section 4. Finally, conclusion is reported in Section 5.

II. FACE RECOGNITION: OVERVIEW

In literature [16], a face recognition system includes two main tasks: *face detector* and *face recognizer*. The difference between the two is that the face detector categorizes the incoming feature vector to one of the two image classes: "face" images and "non-face" images, while the face recognizer classifies the feature vector (assuming it is from a "face" image) as "Name1's face", "Name2's face", or some other person's face that is already registered in the database. Both belong to the same framework for face recognition purpose. More specifically, a face recognition system includes also a feature extractor that transforms the pixels of the facial image into a useful vector representation, in order to use a classifier that searches the database to find the best match to the incoming face image. Fig. 1 shows the main phases in Face Recognition process, including the face detector and face recognizer. In particular:

1. Face Detection;
2. Feature Extraction;
    a. Normalization;
3. Face recognition.

Initially a Face Recognition system consists in the image acquisition and pre-processing phases. More specifically, the pre-processing process consists in the image standardization in terms of: size, pose, illumination and so on. The pre-processed image is necessary for the reorientation and slight variations purposes. Indeed, the recognition requires that reference image and all other trained images are the same in terms of pose, orientation, rotation, scale, size and so on. Then, the first step of the system (see Fig. 1) consists in Face Detection. This process decides which pixels in the image is part of the face and which are not. Successively, in the detected face, some methods are used in order to detect eyes, nose and mouth. More specifically, standard feature templates are used. After face detection, the step of feature extraction can take place. In feature extraction, the biometric distances are calculated and normalized respect to the face size. The third step consists in the Face Recognition step (see Fig. 1). In the last phase, the biometric template of the suspect face is compared with the biometric template of each face present in the database. We recognize the face when we get a match between these two biometric templates.

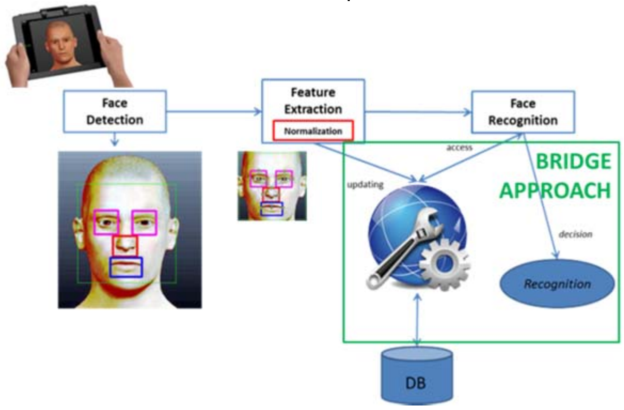


Fig. 1 Face Recognition system phases

In this work, the Bridge Approach (BA) is used in order to speed up the image processing and reduce the computational costs for the face recognition purpose. In particular, the web service:

1. Stores the normalized distances obtained from the Euclidean distances between two specific landmarks;
2. Verifies the correct building of the classifier's model and saves it in a specific file;
3. Matches the test sample with each other samples stored in the database, respect to the constructed model.

An example of pieces of the code is illustrated in appendix.

Finally, as already discussed in the state of the art [8], new challenges of face recognition systems are investigated. More specifically, the following three problems are analyzed:

- Low quality of the resolution;
- Illumination;
- Angulation, pose variation and perspective (or distance from the built-in camera of the mobile device).

The primary requirement in a face recognition system is the high quality of the acquired image. More specifically, the high quality it is important for the feature extraction phase that represents the crucial role of the system. The second challenge is related to the illumination problem. Some faces appear differently due to the change in lighting and the Fig. 2 shows the same face image in different illuminations.

Fig. 2 Face images in different illuminations

The 3D faces modeling, in this paper, have been implemented by "MakeHuman" tool. Face, eyes, nose and mouth detections on these face modeling was performed by means of the implemented code.

The illumination problem can be solved using different approaches. The most commonly used techniques are: Heuristic approach, Statistical approach, Light-modeling approach, Model-based approach, Multi-Spectral Imaging (MSI) known in literature.

The third challenge is related to the pose variation (see Fig. 3). Usually, the training data used by face recognition systems is composed by frontal face images of individuals because the frontal view contains more specific information respect to the profile or other pose angle images. The problem appears when the system has to recognize a rotated face using this frontal view training data. User needs together multiple views of an individual in a face database.

Fig. 3 Database of a person with different pose

The pose problem has been divided in to three categories:

1. Simple case. The dataset contains training images with small rotation angle ;
2. Normal case. The dataset contains training image pairs, frontal and rotated images;
3. Difficult case. The dataset contains only individual's frontal images.

Following are the most relevant approaches to pose problem: Multi-image approach, Single-model based approach and Geometric approach.

III. METHOD

This face recognition system used a real dataset (or numerical feature vector set) in order to test the system robustness. Here, we discuss the overall procedure.

Let:

- $P = \{x_k | k = 1,2, \dots, |P|\}$ be the training image set. In particular, the real images are acquired through the built-in camera of the mobile device and each face is labeled.
- $P' = \{y_z | z = 1,2, \dots, |P'|\}$ be the test image set. In particular, P' is used to test the system.
- A be the classifier (or recognizer);
- $F_i(x_k) = (F_{i,1}(x_k), F_{i,2}(x_k), \dots, F_{i,r}(x_k), \dots, F_{i,R}(x_k))$ be the numeral feature vector used by A to represent the specific pattern $x_k \in P$ (with R numeral features);
- KB be the knowledge base of A.

Initially, in the first stage (s=1), the classifier's knowledge base A is empty. Therefore, it is initially defined as:

$$KB = \{\emptyset\} \quad (1)$$

Successively, the set P of labeled samples is provided to system for learning.
So,

$$KB = \{F(x_k) | k = 1,2, \dots, |KB|\} \quad (2)$$

Therefore, KB is the numeral feature vector set of the classifier for the K patterns that belongs to the knowledge base (KB).
P' is considered as set of real cases for testing in order to avoid biased or too optimistic results.

The face recognition system in the Algorithm 1 is detailed.

Algorithm 1: Face Recognition

```
  Captured images, using the built-in camera of
the mobile device:
     •   Dynamic targeting of face, eyes, nose and
         mouth;
     •   Acquisition of 50 correct detections;

  TRAINING PHASE

   1.  FOR each captured image: P={x1,…,xi,...,
       xK}, the system:
            a.  Normalizes the image size;
            b.  Applies the Histogram Equalization;
            c.  Detects the face;
            d.  Detects the main landmarks face;
            e.  Extracts the features;
                   i.  Normalizes the distances
                       between two particular
                       points;
  END FOR

  Adds all the detected distances in the database,
using the BA based on web service architecture;

  Captured images, using the built-in camera of
the mobile device:
     •   Dynamic targeting of face, eyes, nose and
         mouth.
     •   Acquisition of 60 images;

  TEST PHASE

   2.  FOR each captured image: P'={y1,…,yj,...,
       yz}, the system:
            a.  Normalizes the image size;
            b.  Applies the Histogram Equalization;
            c.  Detects the face;
            d.  Detects the main landmarks face;
            e.  Extracts the features;
                   i.  Normalizes the distances
                       between two particular
                       points;
  END FOR

  Compares this feature vector with each other
stored in the database, using the classifier model
built through the BA;

  IF (declared label != classification label)
  Sends an alert (log file), associating the
classification probability;
```

The final goal of this work is both to detect constantly the individual's face in front of mobile device and to send an alert to the responsible if another individual's face is recognized. In order to constantly detect the individual's face, in test phase, 60 facial images are acquired.

Furthermore, in order to solve the illumination problem, the Histogram Equalization method was used.

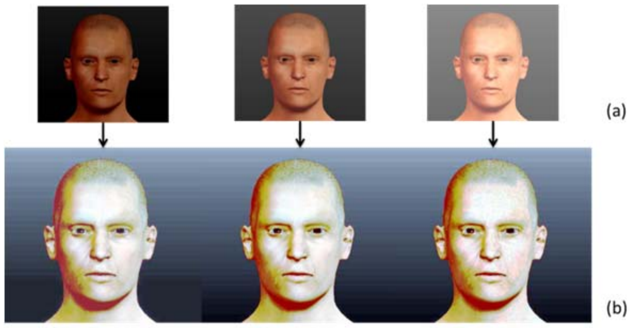


Fig. 4 Histogram Equalization

Fig. 4 shows an example of Histogram Equalization method. More specifically, the Fig. 4 (a) shows the images of individual's face under different lighting conditions and Fig. 4 (b) shows the histogram equalization images.

Let consider the discrete gray scale input image $X = x(i,j)$, with L discrete levels, where $x(i,j)$ represents the intensity levels of the image at the spatial domain (i,j). Let histogram of Image X is $H(X)$ that represents the transfer function of the filtering process.

Now the probability density function $pdf(X)$ can be defined as:

$$pdf(X_k) \text{ or } p(X_k) = {}^{n_k}/_{N} \quad (3)$$

Where, $0 \leq k \leq (L-1)$
L is the total number of gray levels in the image;
N is the total number of pixels in the image;
n_k is the total number of pixels with the same intensity level k.

From the $pdf(X)$ (3) the cumulative distribution function $cdf(X_i)$ is defined as:

$$cdf(X_i) \text{ or } p(X_i) = \sum_{i=0}^{k} p(X_i) \quad (4)$$

Note that $cdf(X_{L-1}) = 1$ from eq. (3) and (4).

Histogram Equalization increases the global contrast of many images, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to gain a higher contrast. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values. Moreover it's a scheme that maps the input image into the entire dynamic range $[X_0, X_{L-1}]$ by using the cumulative distribution function as a transform function. Let's define the transform function $f(X)$ using cumulative distribution function $cdf(X_i)$ as:

$$f(X) = X_0 + (X_{L-1} - X_0) * cdf(X_i) \quad (5)$$

Then the output image of histogram equalization $Y = y(i, j)$ can be expressed as:

$$Y = f(X) \qquad (6)$$

or equivalently by

$$Y = \{f(x(i,j)) | \forall x(i,j) \in X\} \qquad (7)$$

The above describe the histogram equalization on gray scale image. However it can be also adopted on color image by applying the same method separately to the Red, Green and Blue component of the RGB color image [17].

It's possible to observe (see Fig. 4 (b)) that histogram equalization images are similar respect to the original image captured under different lighting conditions.

Finally, in order to overcome the perspective problem, the distances were normalized. Fig. 5 shows the processing of three different images in perspective. More specifically, each distances in the face image has been normalized respect to the height (or width) of the face.

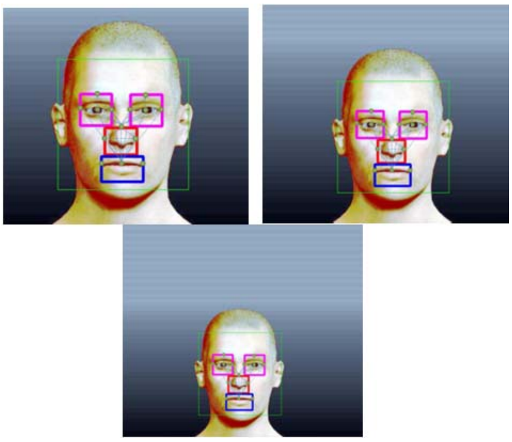


Fig. 5 Images in perspective

In Cartesian coordinates, if $P = (p_1, p_2, \ldots, p_{14})$ and $Q = (q_1, q_2, \ldots, q_{14})$ are two points in Euclidean n-space, the distance (d) from P to Q (or from Q to P) is given by the Pythagorean formula:

$$d(P,Q) = d(Q,P) = \sqrt{(q_1 - p_1)^2 + \cdots + (q_{14} - p_{14})^2} = \sqrt{\sum_{i=1}^{14} (q_1 - p_1)^2} \qquad (8)$$

This distance has normalized respect to the height (or width) of the face. Height and width are equal because the detection face is a square:

$$nd(P,Q) = d(P,Q)/height(face) \qquad (9)$$

So, the normalized distances $nd(X, Y)$ have been stored in the database as features for the recognition purpose. An example of normalized distances is illustrated in appendix.

IV. RESULTS

Experimental results are reported in terms both of correlation score (CS) and of recognition total cost time. More specifically, 25 distances based on 14 face landmarks (see Fig. 6 where face, eyes, nose and mouth was detected by the implemented code) and 7-NN classifier are used for face recognition purpose.

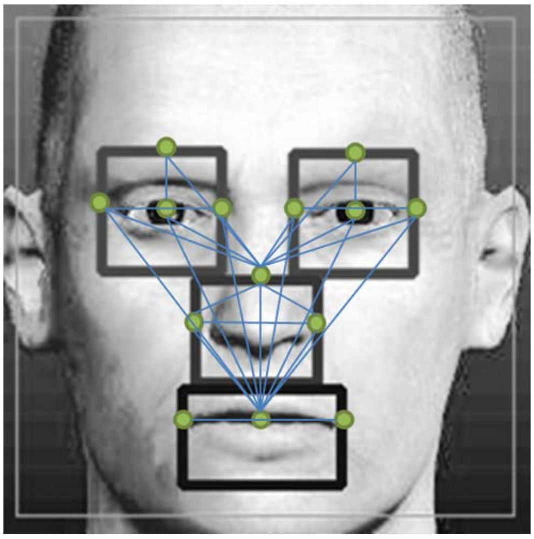


Fig. 6 Face landmarks and distances used for face recognition

Moreover, a smartphone Blackview BV2000 based on Android 5.1 is adopted. It is powered by 64-bit MediaTek MT6735 quad-core processor based on Cortex A53 as well as Mali T720 GPU and 1GB of RAM memory. So, given the mobile device, the angle variation (see Fig. 7) and perspective must be considered in order to optimize the recognition rate.

A TOMCAT platform has been used for the implementation of the web service control.

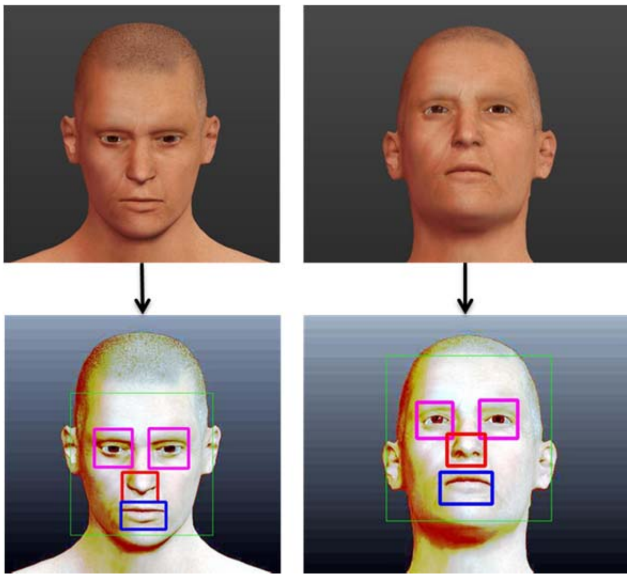


Fig. 7 : Face detection by different angles

Tables I and II show the results obtained by the face recognition discussed above. In particular, these are averaged respect to the three tests measured for each sample (Alessandro, Vito, Valeria, Donato, Lorenza, Davide, Letizia, Cosimo, Annalisa and Mattia) and the mean values of Classification Accuracy (CA) and Classification Occurrence (CO) are considered.

Table I: Classification Accuracy

		Predicted class									
		Alessandro	Vito	Valeria	Donato	Lorenza	Davide	Letizia	Cosimo	Annalisa	Mattia
Actual Class	Alessandro	0.72	0.29	0.00	0.00	0.19	0.00	0.47	0.00	0.00	0.0
	Vito	0.43	0.57	0.00	0.48	0.00	0.00	0.00	0.00	0.68	0.0
	Valeria	0.00	0.00	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.0
	Donato	0.00	0.43	0.00	0.50	0.00	0.00	0.89	0.43	0.00	0.7
	Lorenza	0.49	0.00	0.00	0.00	0.29	0.00	0.50	0.00	0.00	0.0
	Davide	0.57	0.00	0.78	0.00	0.00	0.72	0.00	0.00	0.00	0.0
	Letizia	0.43	0.71	0.00	0.00	0.00	0.00	0.64	0.00	0.00	0.0
	Cosimo	0.38	0.19	0.00	0.23	0.00	0.00	0.00	0.41	0.49	0.0
	Annalisa	0.19	0.40	0.00	0.10	0.00	0.00	0.00	0.33	0.73	0.0
	Mattia	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.9

Table II: Classification Occurrence

		Predicted class									
		Alessandro	Vito	Valeria	Donato	Lorenza	Davide	Letizia	Cosimo	Annalisa	Mattia
Actual Class	Alessandro	50	2	0	0	0	0	1	0	0	0
	Vito	1	11	0	6	0	0	0	0	4	0
	Valeria	0	0	26	0	0	0	0	0	0	0
	Donato	0	3	0	13	0	0	22	1	0	3
	Lorenza	3	0	0	0	1	0	1	0	0	0
	Davide	7	0	2	0	0	14	0	0	0	0
	Letizia	1	1	0	0	0	0	2	0	0	0
	Cosimo	1	1	0	3	0	0	0	13	4	0
	Annalisa	1	1	0	0	0	0	0	1	9	0
	Mattia	0	0	0	0	0	0	0	0	0	49

To combine both measures into a single score, we propose the use of a new correlation score (CS) that combines the both classification accuracy and classification occurrence. The idea behind this score is the observation that classification occurrence is not dependent of the recognition rate. To counteract this effect, we multiply the classification occurrence with the classification accuracy, thus CS=CO*CA/100. A higher correlation score indicates a higher probability of correct recognition (see Table III).

Table III: Correlation Score

	Predicted class									
	Alessandro	**Vito**	**Valeria**	**Donato**	**Lorenza**	**Davide**	**Letizia**	**Cosimo**	**Annalisa**	**Mattia**
Alessandro	0.36	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0
Vito	0.00	0.06	0.00	0.03	0.00	0.00	0.00	0.00	0.03	0.0
Valeria	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.0
Donato	0.00	0.01	0.00	0.07	0.00	0.00	0.20	0.00	0.00	0.0
Lorenza	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.0
Davide	0.04	0.00	0.02	0.00	0.00	0.10	0.00	0.00	0.00	0.0
Letizia	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.0
Cosimo	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.05	0.02	0.0
Annalisa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.0
Mattia	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.4

In general, this face recognition system based on web service architecture had a classification accuracy of 80%. In particular, for the actual class "Donato", the system produced a prediction error with "Letizia", whose correlation difference is 0.13 respects to the actual class. Instead, the system confused the Lorenza's face with "Letizia" and "Alessandro". These results have been influenced by real operating conditions, such as:

- Camera distortion and noise
- Complex background
- Illumination
- Translation, rotation, scaling, and occlusion
- Facial expression
- Makeup and hair style

Finally, Fig. 8 shows a comparative analysis of the cost time of Face Verification employed by our Bridge Approach (BA) and another technique known in literature [18].

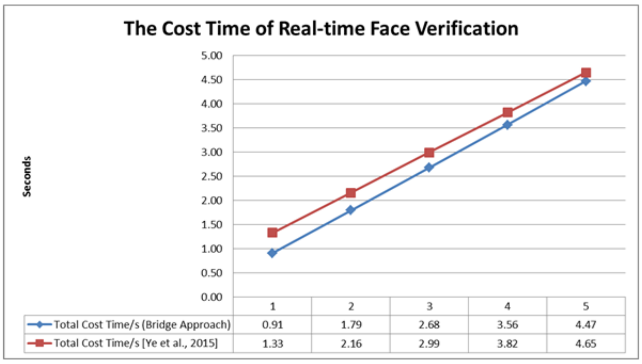


Fig. 8 : Face detection by different angles

Our Bridge Approach (BA) is more efficient respect to the state-of-the-art of 0.31s.

V. Conclusions

Face Recognition System on Mobile Device based on Web Service approach has been introduced, according to the monitoring of an individual in front of a mobile device on Android platform. More specifically, some methods have been introduced in order to overcome the problems related both to the different illumination conditions and to the perspective/pose variation respect to the built-in camera on mobile device.

The experimental results show the system efficiency, using real face images in terms both of correct rate and total cost time of verification. Finally the use of a correlation score helps to better understand the person in front of the device, in order to send an alert to the responsible if another individual's face is recognized respect to the declared person. Clearly, despite the 25 features (normalized distances), further acquisitions will allow to obtain a more classification accuracy.

In the future, this system will be easily implemented in IoS systems for the mobile face recognition.

Acknowledgment

This work has been developed within the framework of Research Project titled: "Applicativo software mobile per il riconoscimento di persona mediante algoritmi identificativi di punti corrispondenti a tratti somatici rilevanti" (or Mobile software tool for face recognition by means of identification algorithms of points corresponding to the somatic relevant features). Also the authors gratefully acknowledge Vito Antonacci, Bachir Boussahel, Veronica Calati, Giacomo Fanelli, Davide Donato Romagno and Antonio Semeraro for their support.

References

[1] Ng, C. C., Yap, M. H., Costen, N., & Li, B. (2015, October). Will Wrinkle Estimate the Face Age?. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on* (pp. 2418-2423). IEEE.

[2] Perez, Sarah. "Lambda labs is launching a facial recognition api for google glass." Techcrunch, May 23 (2013).

[3] Gorodnichy, D., Granger, E., Radtke, P., Authority, C. S., & Meunier, P. (2014). Survey of commercial technologies for face recognition in video. *CBSA, Border Technology Division, Tech. Rep, 22*.

[4] Dandashi, Amal, and Walid Karam. "Biometrics security and experiments on face recognition algorithms." Computational Intelligence for Security and Defence Applications (CISDA), 2012 IEEE Symposium on. IEEE, 2012.

[5] Peter, K. J., Nagarajan, G., Glory, G., Devi, V. V. S., Arguman, S., & Kannan, K. S. (2011, April). Improving ATM security via face recognition. In Electronics Computer Technology (ICECT), 2011 3rd International Conference on (Vol. 6, pp. 373-376). IEEE.

[6] Holat, Recep, and Selman Kulac. "ID identification by using face detection and recognition systems." Signal Processing and Communications Applications Conference (SIU), 2014 22nd. IEEE, 2014.

[7] Nakano, Miyoko, F. Yasukata, and Minoru Fukumi. "Marketing data collection from face images using neural networks." Nonlinear Signal and Image Processing, 2005. NSIP 2005. Abstracts. IEEE-Eurasip. IEEE, 2005.

[8] Ramchandra, Ambika, and Ravindra Kumar. "Overview of Face Recognition System Challenges." International Journal Of Scientific & Technology Research 2.8 (2013).

[9] Tolba, A. S., A. H. El-Baz, and A. A. El-Harby. "Face recognition: A literature review." *International Journal of Signal Processing* 2.2 (2006): 88-103.

[10] Barbuzzi, D., Impedovo, D., & Pirlo, G. (2012, September). Benchmarking of update learning strategies on digit classifier systems. In *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on* (pp. 35-40). IEEE.

[11] Barbuzzi, D., Pirlo, G., & Impedovo, D. (2014). About retraining rule in multi–expert intelligent system for semi–supervised learning using SVM classifiers. *International Journal of Signal and Imaging Systems Engineering*, *7*(4), 245-251.

[12] Barbuzzi, D., Pirlo, G., Uchida, S., Frinken, V., & Impedovo, D. (2015, August). Similarity-based regularization for semi-supervised learning for handwritten digit recognition. In *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on* (pp. 101-105). IEEE.

[13] Impedovo, S., & Barbuzzi, D. (2014). Instance Selection for Semi-Supervised Learning in Multi-Expert Systems: A Comparative Analysis. *Journal of Next Generation Information Technology*, *5*(4), 61.

[14] Laganière, R. (2011). *OpenCV 2 Computer Vision Application Programming Cookbook: Over 50 recipes to master this library of programming functions for real-time computer vision*. Packt Publishing Ltd.

[15] Starovoitov, V. V., Samal, D. I., & Briliuk, D. V. (2002, October). Three approaches for face recognition. In *The 6-th International Conference on Pattern Recognition and Image Analysis, Velikiy Novgorod, Russia* (pp. 707-711).

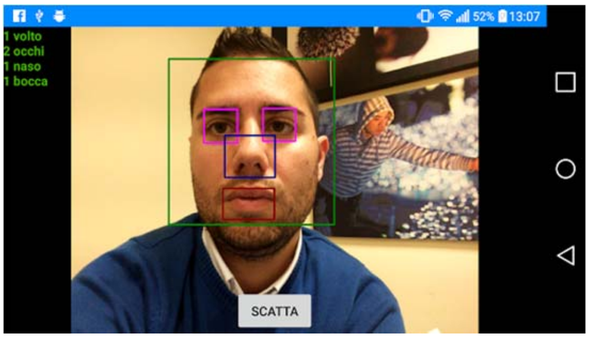
[16] Lin, Shang-Hung. "An introduction to face recognition technology." Informing Science 3.1 (2000): 1-8.

[17] Patel, Omprakash, Yogendra PS Maravi, and Sanjeev Sharma. "A comparative study of histogram equalization based image enhancement techniques for brightness preservation and contrast enhancement." arXiv preprint arXiv:1311.4033 (2013).

[18] Ye, Pinjie, Mengmeng Yu, and Minghui Wu. "Implementation: Mobile Face Identity Authentication System on Android Platforms." International Journal of Security and Its Applications 9.1 (2015): 51-60.

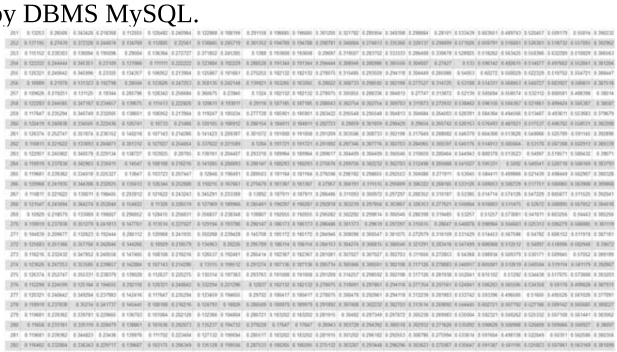
Appendix

Below is illustrated the real face recognition application by means the use of a mobile device.



Example of Face Recognition using Mobile Device.

Below are illustrated the extracted features of the above face. Each raw represent the extracted features from a single detection. The features are stored into a DB created by DBMS MySQL.



Example of Extracted Features of the Mattia's face.

Below we report the code related to the calls to the web service.

```
wsm=new
WsManager(MainActivity.this,"response","http://
192.168.0.39:8080/FaceRecognition/ws_test",WsMa
nager.METHOD_POST);
wsm.addParam("action",2);
wsm.run();
```

Call to the web service to verify the correct building of the classifier's model

```
wsm=new
WsManager(MainActivity.this,"response","http://
192.168.0.39:8080/FaceRecognition/ws_test",WsMa
nager.METHOD_POST);
wsm.addParam("action",WS_ACTION_TRAINING);
wsm.addParam("nome",nome);
wsm.addParam("r1", json.toString())
wsm.run();
```

Call to the web service to store the normalized distances in the database

```
wsm=new
WsManager(MainActivity.this,"response","http://
192.168.0.39:8080/FaceRecognition/ws_test",WsMa
nager.METHOD_POST);
wsm.addParam("action",WS_ACTION_TEST);
wsm.addParam("nome",nome);
wsm.addParam("r1", json.toString())
wsm.run();
```

Call to the web service to match the test normalized distances with all normalized distances in the database