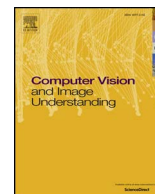




Contents lists available at ScienceDirect

# Computer Vision and Image Understanding

journal homepage: [www.elsevier.com/locate/cviu](http://www.elsevier.com/locate/cviu)

## A multi-agent system for the classification of gender and age from images<sup>☆</sup>

Alfonso González-Briones\*, Gabriel Villarrubia, Juan F. De Paz, Juan M. Corchado

University of Salamanca, BISITE Research Group, Edificio I+D+i, Salamanca 37007, Spain



### ARTICLE INFO

MSC:  
41A05  
41A10  
65D05  
65D17

#### Keywords:

Facial recognition  
Automatic age estimation  
Automatic gender estimation  
Preprocessing of images  
Multi-agent system

### ABSTRACT

The automatic classification of human images on the basis of age range and gender can be used in audiovisual content adaptation for Smart TVs or marquee advertising. Knowledge about users is used by publishing agencies and departments regulating TV content; on the basis of this information (age, gender) they are able to provide content that suits the interests of users. To this end, the creation of a highly precise image pattern recognition system is necessary, this may be one of the greatest challenges faced by computer technology in the last decades. These recognition systems must apply different pattern recognition techniques, in order to distinct gender and age in the images. In this work, we propose a multi-agent system that integrates different techniques for the acquisition, preprocessing and processing of images for the classification of age and gender. The system has been tested in an office building. Thanks to the use of a multi-agent system which allows to apply different workflows simultaneously, the performance of different methods could be compared (each flow with a different configuration). Experimental results have confirmed that a good preprocessing stage is necessary if we want the classification methods to perform well (Fisherfaces, Eigenfaces, Local Binary Patterns, Multilayer perceptron). The Fisherfaces method has proved to be more effective than MLP and the training time was shorter. In terms of the classification of age, Fisherfaces offers the best results in comparison to the rest of the system's classifiers. The use of filters has allowed to reduce dimensionality, as a result the workload was reduced, a great advantage in a system that performs classification in real time.

### 1. Introduction

In recent years, facial recognition has become an active field of research that covers various disciplines, such as Biometry (Alim et al., 2017; Mehra and Charaya, 2016; Surekha et al., 2017), information security (Robertson and Burton, 2016), law enforcement and surveillance (Robertson et al., 2016), smartcards or controlled access (Frikha et al., 2016; Lin et al., 2016). Some facial recognition systems have focused on achieving very specific objectives, such as classifying the gender and the age of a person or recognizing facial expression from a human image. This topic has inspired many researchers to look for diverse solutions, however, the feasibility of these proposals has been tested mainly in controlled environments (The number of faces and the position of the face in front of the camera is fixed, lighting conditions are controlled). For this reason, it is necessary to demonstrate that classifying age and obtaining gender is worthwhile in general and realistic situations, that is to say, in environments whose conditions can vary. The process of determining age and gender would grant machines the ability to make decisions on the basis of the values of these factors, such as decision making in the recommendation of advertisements on

marquees or the authorization to see certain contents on Smart TVs. These decisions can be applied to a variety of cases, from access control, human-machine interaction and person identification to data mining and organization (Lanitis, 2010). This problem can be divided into two, the classification of images by gender and the classification of images according to the age group. There are several common factors used in the age and gender classification methods the classification results of such complex problems can be improved by using techniques that have already been tested. Existing classification systems are successful if the image obtained meets a given standard (high quality, occlusion-free, no background images with a neutral facial expression). The majority of these methods experience uncontrolled adjustments in their configuration as described in Shan (2010).

From the range of facial recognition techniques, we can identify Eigenfaces and Fisherfaces (Lu et al., 2013) and Dandpat and Meher (2013) and Artificial Neural Networks (Parkhi et al., 2015; Sun et al., 2015; Yang et al., 2016) and Levi and Hassner (2015) to be highly successful and commonly used techniques. What both of these techniques have in common is that they are based on homogenized images in which the faces are aligned in order to abstract the features and make a

<sup>☆</sup> The name of the Editor in chief Dr. I Bloch.

\* Corresponding author.

E-mail address: [alfonsogb@usal.es](mailto:alfonsogb@usal.es) (A. González-Briones).

correct classification. Many facial classification approaches have been taken when trying to solve this problem: from using neural networks, to the identification of features and even several mathematical techniques that allow to reduce the dimensions of the photographs facilitating the calculations. Although advances have been numerous in all of these technologies, it is still possible to improve these techniques which, in general, require a number of controlled conditions to achieve high effectiveness. As to the classification of images by age group, many image representation methods have been studied, such as anthropometric models, active appearance models (AMM) or aging pattern subspace. Extensive reviews of these age representation methods can be found in Han et al. (2013), Geng et al. (2013) and Panis and Lanitis (2014). Perhaps, among the pioneering studies in age classification are those proposed by Kwon and da Vitoria Lobo (1999), Lanitis et al. (2002) and Guo et al. (2009).

In this work we aim to accomplish the following goals: to study and develop a system that is capable of classifying gender and age from images; to estimate the extent to which the preprocessing stage influences (different face crops) on the performance of classification techniques; to apply diverse filters and observe whether they increase the accuracy of image classification. On the contrary to the previously described works, the proposed method performs the classification of gender and age automatically, without placing any restrictions on the condition of the image. A multi-agent system (MAS) is a system composed of multiple intelligent entities called agents that coordinate, communicate, interact and cooperate with each other. Multi-agent systems can be used to solve problems that are difficult or impossible to solve for an individual agent or a monolithic system. Moreover, they can propose a variety of solutions to the same problem. By making use of the multiple features of a MAS, an architecture has been designed. Agents are a part of this architecture and they carry out a variety of tasks: they acquire images and recognize patterns in them, they perform the training using the database, moreover they incorporate newly obtained images to the training set in order to relate the features of the individuals on the images with a particular age group and gender. The developed system has been tested in an office building and positive results have been obtained. In this work we present the preliminary results of the proposed method and the conclusions drawn from conducting this research work.

This paper is organized as follows: Section 2 reviews related studies, Section 3 describes the proposed architecture, Section 4 presents the case study in which the platform is applied and the results of this implementation, and lastly, conclusions are discussed in Section 5.

## 2. Related work

In recent years, considerable advancements have been made in the area of facial recognition. With the aim of achieving robust results in the classification of images, diverse techniques have been proposed for their handling. The first study focusing on the estimation of age was proposed by Kwon and da Vitoria Lobo (1999), since then, many different studies have emerged, each using a different technique and different ways of treating input data. Over the last few years, there has also been an increase in studies that leverage neural networks for face detection, classification of gender and estimation of age from images. From the simplest examples, which use a single multilayer network to which the images are passed, to other more complex ones, which have also been tested successfully by different researchers. We have, for example, the solution devised by Agarwal et al. (2010) consisting of 3-layer neural networks. Sharma et al. (2013) proposes an efficient method which has been shown to provide good results, even in the presence of slight appearance variations due to lighting and expression. Although Dehshibi and Bastanfard (2010) have conducted studies to deal with this issue, the focus of the research is on the application of mathematical and / or statistical algorithms to the data set.

Despite the great variety of methods and advances in this field, the

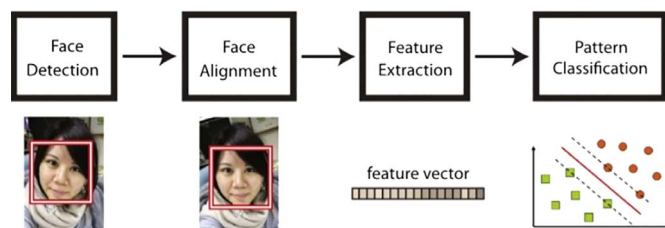


Fig. 1. Pipeline for the classification of gender and age range in images.

vast majority of them have one thing in common: they depend on a process of learning or training, a crucial stage that will define the subsequent performance of our classifier. In Fig. 1 we can see the definition of a basic system used to classify gender and age in images (Chu et al., 2013). Correlation techniques were the first to be used in face detection, also called the template matching approach, its basis is the comparison of an image using a training database to classify it with its nearest neighbor in the image space (Pentland et al., 1994). This strategy has several variations, in some cases different templates of one individual's image are used, with their position changed or close-ups on certain parts of their face. Pentland et al. (1994), the authors of this technique, admit that some problems may emerge when the images are not taken with the same main light source, something that does not happen with other methods such as Fisherfaces (Özgül and Özbilen, 2014). Other methods use Fourier transform to make comparisons in the frequency domain, thus making the recognition independent of light (Banerjee and Datta, 2014). While these authors used it for facial recognition and not to distinguish gender, it could possibly work if the method was adapted.

Further studies on facial recognition have in some way continued those on age estimation and soon, classification algorithms such as Eigenfaces or Fisherfaces were applied in this area. These algorithms have to learn about the features of different age groups in order to distinguish between them. This task is carried out by the Eigenfaces method which is based on the analysis of principal components, however, it is not fully compatible with unsupervised statistical model. Nevertheless, Fisherfaces is much more convenient for this task since it uses class specific linear projection, as demonstrated by Dadi and Mohan (2015) and Wang and Jia (2016). However, previous works were not only concerned with applying, improving and combining algorithms or developing new ones, they also focused on the extraction of information from faces, as proposed by Choi et al. (2011), one of the first more complete systems, which includes a combination of global and local features. There are other less complete proposals, such as the research enunciated by Jana et al. (2012) which calculates biometric proportions and angles formed by the different features: eyes, nose and mouth. However, this type of research is usually supplemented with the extraction of features which only perform well in headshots with uniform light, where the individual's hair does not cover their forehead and they do not wear glasses for proper eye and eyeball detection (Jana et al., 2015).

A typical, holistic method called linear discriminant analysis (LDA), is very successful in the classification of high-dimensional patterns and in facial recognition since the appearance of the Fisherfaces method. All current facial recognition systems are based on LDA. However, when LDA is used in facial recognition, the problem of small sample size (SSS) always emerges, where the number of training samples  $N$ , is much smaller than the size of the facial image. To date, many researchers have proposed different ways of solving this problem, Fisherfaces presented by Swets and Weng (1996) and Belhumeur et al. (1997) has been one of the most famous of these proposals, it uses Principal Component Analysis (PCA) for the mapping of original data on a low-dimensional subspace in order to eliminate the uniqueness of  $S_w$ . It has been shown to be better than Eigenfaces, particularly in images that vary in lighting and facial expressions. Unlike Eigenfaces which calculates a special

eigenvector for each image from the training set, Fisherfaces only calculates a special eigenvector and eigenvalue for each person. Hence, it is important to carry out a study on the two most successful recognition techniques for determining age group and gender in human images. These techniques are Fisherfaces (Belhumeur et al., 1997) and Neural Networks (Levi and Hassner, 2015). They are a derivation of one of the most known principal components analysis (PCA) methods, the Eigenvectors. However, this technique works in such way that it enables to classify the different age groups more effectively; its uses the facial features that differentiate the groups from each other effectively. At the same time this technique minimizes the influence of features that differentiate individuals who are in the same age group. Since correlation or template matching classification methods have a high computational complexity, this classification problem is tackled using dimensionality reduction techniques, Eigenfaces extraction being one of the most known; Reducing dimensionality and expressing data according to a new, adequate base enables us to express our data without losing valuable information for the classification. PCA (Bro and Smilde, 2014); this technique is used in the calculation of Eigenfaces, it consists in calculating the main elements of the image, called Eigenvectors. The Eigenvectors are orthogonal to each other permitting us to express the content of an image as a matrix, using its Eigenvectors as a base, instead of using the X and Y axis. These Eigenvectors are n-dimensional, which is why they tend to be referred to as Eigenfaces or Eigenpictures. The problem with this method is that it not only considers the elements that differentiate different age groups, instead, differentiating elements within each age group, such as lighting are also identified (Imran et al., 2015). The effect produced by light when using Eigenfaces can be eliminated with the use of the Fisherfaces method. The main difference between both methods is that Fisherfaces make use of the database using not only the information contained in the photos themselves but also metadata. The use of tags helps create a more reliable method since it informs us on the age group that an image belongs to, reducing the dimensionality of the feature space. The Eigenfaces method is a combination of the Fisherfaces and Fisher's Linear Discriminant (Fisher, 1936). The FLD uses a linear combination of features measured in classes to classify them. The use of this method permits us to eliminate the influence of light in the images, particularly the problem of the direction of the main focus.

For classification to be carried out correctly, it is very important to compile a good training set, once the face in an image is detected the use of an eye detection classifier is indispensable. It is crucial for the position of the face to be the same in every image during the process of training and classification. These eye detection classifiers also allow for the detection of facial features such as the eyes, nose or ears and are later used for determining and extracting the individual's facial features (Lai and Ko, 2014; Majumder et al., 2014). The preprocessing stage plays a fundamental role in further classification since the extracted face features are used to determine the gender or the age group that a person belongs to.

Since a number of algorithms and techniques has to be applied in the various stages of the image classification process; Image processing, pre-processing, dimensionality reduction, training, filters, classifier, etc. It is necessary for the system which brings together this whole functionality to be capable of communicating and organizing without the intervention of a user. Due to the numerous advantages of multi-agent systems they are a typically chosen approach for the creation of systems which require autonomy, decentralization and collaborative communication for the accomplishment of more complex tasks. They allow for an integrated adaptation to new tasks through the incorporation of agents whose role is to perform these tasks. Multi-agent systems are also notable for automating these processes and it is therefore common to find that they are employed in the literature for a wide variety of purposes, such as the management of heterogeneous sensor networks (Bajo et al., 2015), data analysis in bioinformatics (González-Briones et al., 2015; Ramos et al., 2017) or efficient street

lighting control (De Paz et al., 2016). In the area of facial recognition, we should mention works that have made use of a multi-agent system for facial recognition in images, such as the one conducted by Wang and Jia (2016) or Shaikh et al. (2016) who used a multi-agent system in his proposal of algorithms for the recognition of facial expressions. In the next section, we describe our proposal, a multi-agent architecture, specifically developed for the acquisition and classification of images in real time.

### 3. Proposed system

The proposed system consists of a multi-agent architecture, specifically developed for the acquisition and classification of images in real time. The agents from the proposed architecture carry out all the required tasks, from the initial headshot to the classification of gender and age range. The three facial recognition algorithms that have the best results in the reviewed literature (Fisherfaces, Eigenfaces, LBP, MLP) have been modeled as independent agents in the architecture, as well as other agents for the extraction of facial features, reduction of dimensionality and those employed for the classification of gender and age in the different ranges (Shyam and Singh, 2015; Taigman et al., 2014). In order to apply the algorithms correctly and obtain a high success rate in the classification results, it is essential to pre-process the images. This is why this architecture must contain agents responsible for performing these pre-processing tasks and for applying different filters to obtain input data, for eliminating unnecessary data and for the classification process. The operational structure encompassing this process can be seen in Fig. 2.

#### 3.1. Preprocessing

Pre-processing is one of the most important stages when it comes to the correct functioning of the classifiers that will be used. In general, the image pre-processing stage comprises of the phases shown in Fig. 3. The images obtained by the cameras are scanned to detect the presence of a human face, once this process is done, the eyes are detected using the Haar feature-based cascade classifier (Viola and Jones, 2001). According to Sharifara et al. (2014) there are a lot of reasons for which the use of features is preferable to the use of pixels, one of them is that features encode the knowledge of the training set. Speed is another reason; working with features is much faster than working with systems based on pixels. Working with features is based on reducing the computational complexity by representing the rectangular features through its integral. These classifiers have been trained with a series of close ups of particular features (eyes in this case) scaled to the same size (20x20) along with a series of arbitrary images that will allow the system to adapt and not give false positives. Once we have the coordinates of the eyes in the image, we can identify the facial oval, crop it and rotate it (if necessary) so that it is perfectly aligned with the eyes. The area to be cropped will be square, centered on point  $(X_c, Y_c)$  of the line that is perpendicular to the segment joining the two eyes and which originates at the midpoint  $(P_{mx}, P_{my})$  of that segment. In addition, this area will be rotated according to the angle of inclination of the head ( $\alpha$ ). Considering that the coordinate system is the Cartesian with the origin point located in the upper left corner, we can determine how the center of the area in Fig. 4, that is to be cropped, is calculated.

In Fig. 5 we can identify how the images that have different angles and sizes are converted into completely homogenous and normalized facial ovals. In addition, we can see how some images will have to be rotated while others will not. The range of positions is wide, however whenever the image is a headshot, it is possible to align it. This is because a perfect alignment with the eyes is crucial in our training methods. As explained at the beginning of this work, the goal of this study is to determine how different ways of cropping an image influence the classification process. In Figs. 6 and 7 we can look at the entire process that an image goes through to finally obtain the different crop

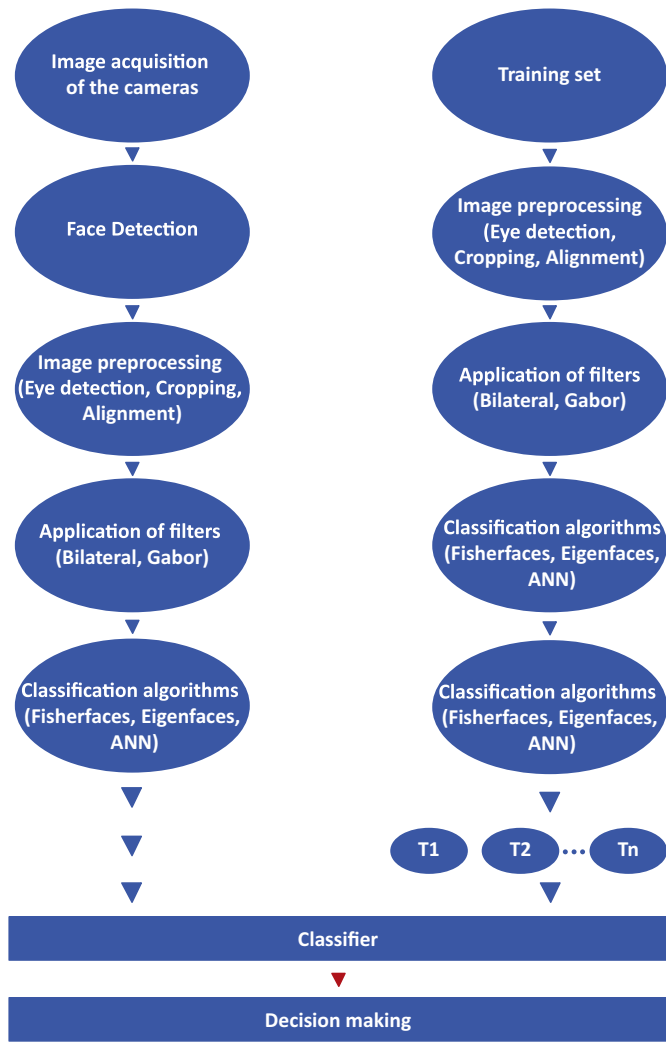


Fig. 2. Operational structure of the proposed system.



Fig. 3. Steps in pre-processing stage.

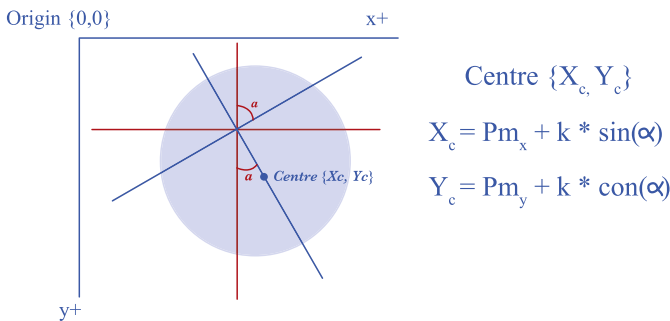


Fig. 4. Calculation of the center of the picture to cut in the image.

types that vary in their centering and in the length of the side of the square. The first step, as mentioned above, is the calculation of the midpoint of the face. This step is carried out by using the coordinates obtained from the photograph during the eye detection process. In Fig. 6 we can observe the process of rotating the image according to its angles and the coordinates of the eyes. Finally, we can notice in Fig. 7



Fig. 5. Images before and after pre-processing.

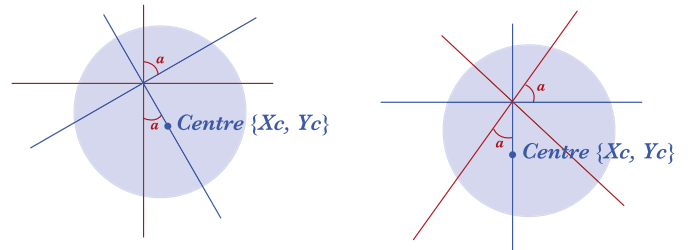


Fig. 6. Facial oval rotation process.

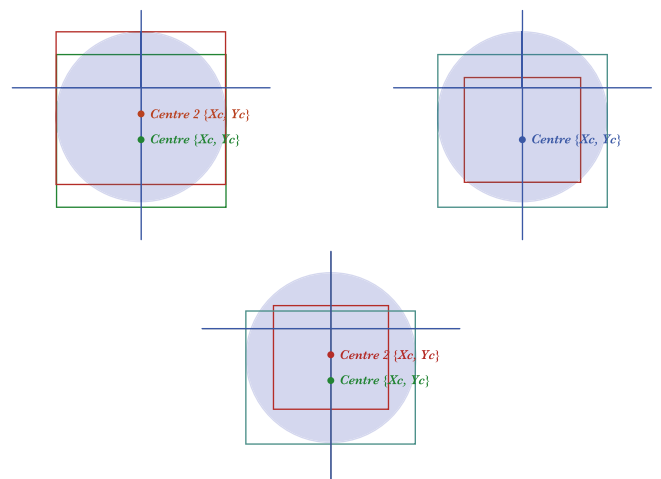


Fig. 7. Face cropping process.

how to crop the different areas, depending on the center and the side of the square that is to be cropped. In Fig. 8 we can detect the four vertical crops to prepare the four different sets of preprocessing images. It can be seen that the difference in facial features, in particular the ratio between the width and height of the face, can make one of the image sets more suitable than another. Additionally, we can assess that the larger the area of the image, the more features will fit into it, however, if the area is small we will be missing bottom pieces which can distort our results. It is therefore of vital importance to achieve a balance between these two factors. This balance will be difficult to achieve since there are considerable differences between individuals in terms of face size, even when the cut is proportional.

### 3.2. Filtering the image

One of the most common problems that we encounter when analyzing images obtained by capture devices is the compression, quantization and sensor noise of the capture sensor. One of the methods used for reducing this noise is to spatially filter the image (smoothing), although this may contribute to the loss of some details in the image.



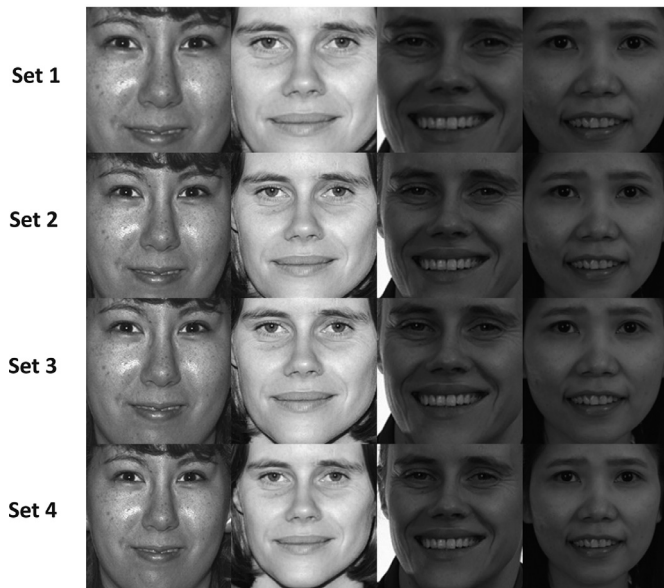


Fig. 8. Different cuts of the same facial oval.

**Bilateral filter:** Bilateral filtering smooths the images without altering the edges, by means of a non-linear combination of the close values of the image. The method is non-iterative, local and simple. It combines gray levels or colors, depending on their geometric proximity and their photometric similarity. It prefers close values to distance values in both domain and range. In contrast to filters operating in three separate color bands, a two-sided filter can enforce the underlying perception metrics in the CIE-Lab color space, smooth the colors and preserve the edges to suit human perception. Also, in contrast to standard filtering, bilateral filtering does not produce ghost colors along edges in color images, and reduces the ghost colors that appear in the original image (Cho et al., 2014).

**Gabor filter:** The Gabor filter is a linear filter whose impulse response is a sinusoidal function multiplied by a Gaussian function (Rai and Khanna, 2014). They are almost pass band functions. The main advantage of introducing the Gaussian envelope is that the Gabor functions are located in both the spatial and the frequency domain, unlike the sinusoidal functions, which are perfectly localized in the frequency domain and completely delocalized in the spatial (sinusoidal functions cover the entire space) (Ahonen et al., 2006; Liu and Wechsler, 2002). Therefore, these functions are more suitable for representing a signal together in both domains. Gabor is a bandpass filter in 2D, if we assign a certain frequency and direction we obtain a reduction in the noise at the same time the address of the original image is preserved (Zhang et al., 2007). The main disadvantage of applying this filter is that the images in order to be classified must have a defined orientation and frequency.

The Bilateral and Gabor filters, focus on the need of obtaining a higher percentage of success in the classification. The system includes these two filters because they give good results for different brightness, contrast, saturation or sharpness settings.

### 3.3. Classification of gender

One of the main goals of this work is to study different techniques that allow for a robust recognition of gender in images. For this, the architectures reconnaissance layer is made up by agents that model these two approaches. The first approach is the use of the previously mentioned Fisherfaces, whose implementation proposal is based on the OpenCV framework. The process of creating a classifier using the Fisherfaces implementation can be seen in Fig. 9. The second approach will be modeled by agents in charge of an Artificial Neuronal Network,



Fig. 9. Process of creating a classifier using Fisherfaces.

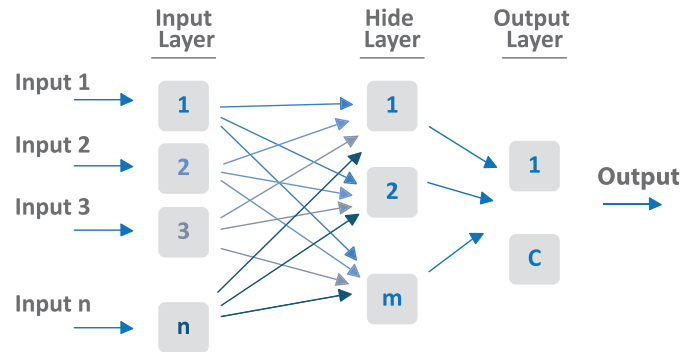


Fig. 10. Multilayer perceptron structure.

specifically by a multilayer perceptron. In the preprocessing, the image will be handled as a  $64 \times 64$  array with values in  $[0, 255]$  corresponding to the grayscale values of each of the pixels in the image. The pixels are prepared for entry into the network, having to be normalized in the interval  $[-1, 1]$ . The structure designed is represented in Fig. 10, where  $n$  is the number of inputs,  $m$  is the number of hidden layers of the multilayer perceptron, and  $c$  is the number of possible outputs. In the classification of gender images,  $n = m$ , number of pixels in the image (4096) and  $c = 2$ , corresponding to the groups to be classified; male and female (Fig. 11).

### 3.4. Classification in age ranges

Unlike the works that have been described previously, in this work we propose a new method for the automatic classification of age without any restriction on the images. Our research is focused on the development of a pattern recognition scheme that does not rely heavily on geometry and calculus as deformable templates. The use of Fisherfaces seems to be a suitable method for face recognition, due to its capacity, simplicity, speed and learning (Delgado-Gomez et al., 2009). The agents in charge of classifying into age groups, interact with agents in charge of the pre-processing of images and the database for training. Workflow management agent coordinates the pre-processing, training and classification stages by communicating with the agents that belong to these layers. In Fig. 12, we can observe how this agent communicates with the Classifier Agent in order to train the different classifiers with the current image database (in each classification process the correctly labeled image is added to the database). When the Image Capturing Agent detects a face, Workflow management agent sends the image to the agents that make up the preprocessing layer, afterwards, the same agent sends the preprocessed image to a classifier, in this case the Fisherfaces classifier.

A Haar feature-based cascade classifier trained to recognize human eyes in images will be used. Then, the agents from the pre-processing layer will decide whether it is necessary to align the image and compare it with the Fisherfaces training mode, with which each of the obtained images will be compared through the camera system, so that the image is classified according to the way it adjusts to the model of the different training sections. This process of classification in age ranges is identical to the process of classifying images by gender as shown in Fig. 9, independent of training for this purpose.

### 3.5. Multi-agent system

A multi-agent system was used because it was necessary for the

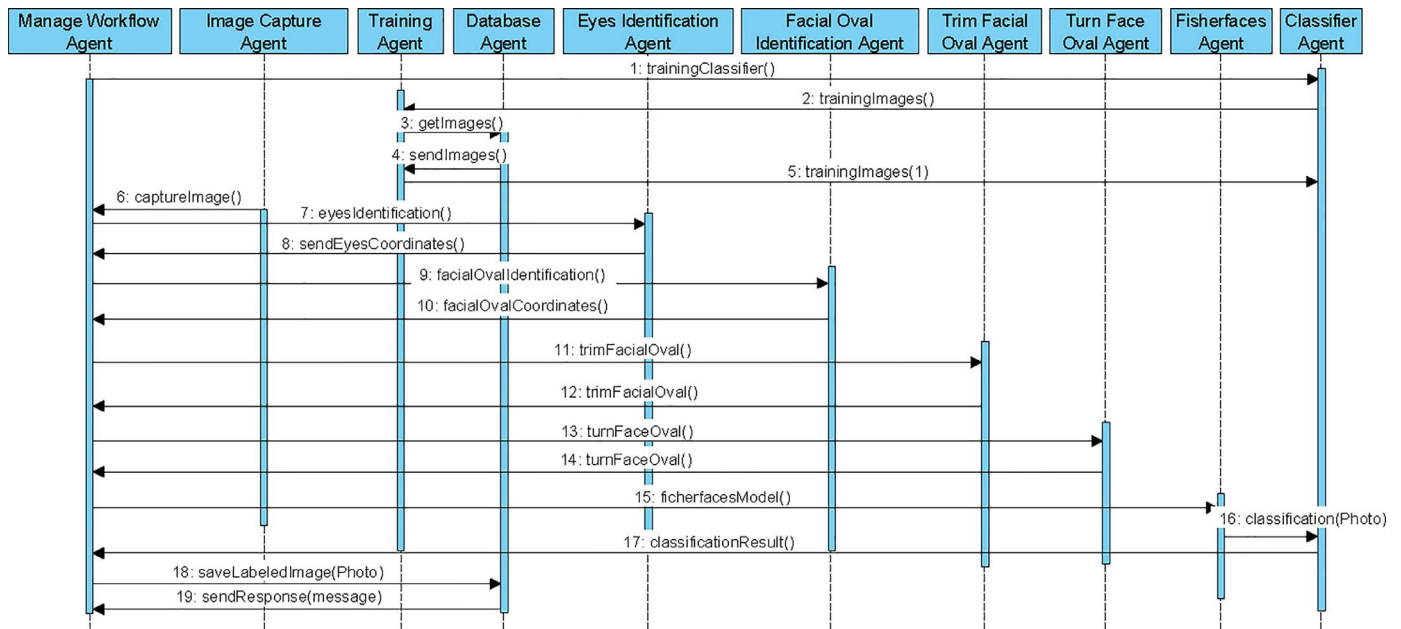


Fig. 11. Sequence diagram of a possible communication between agents.

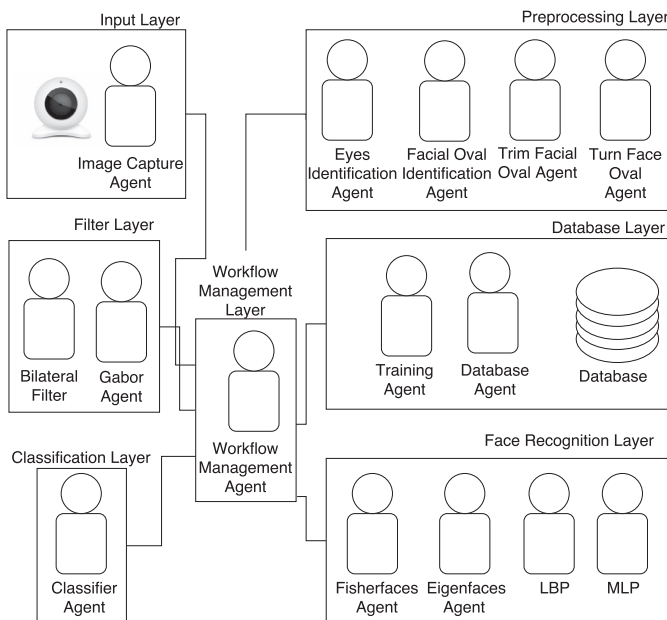


Fig. 12. Multi-agent system architecture.

architecture to communicate autonomously, necessary for the carrying out of tasks and accomplishment of the established objectives (Villarrubia et al., 2014). The multi-agent system organizes the operation of the agents in the system; it distributes roles among the agents which are part of this architecture. They are grouped in seven layers according to the affinity of activities performed by each one as shown in Fig. 12. These activities include acquiring images and recognizing patterns in them, performing the training using the database and incorporating new, tagged images obtained in the training set, once the system has a correct confirmation on their classification. This confirmation is carried out by the system itself when images of the same person are continually obtained in real time by the web cameras. Once the system obtains fifteen images of the same person and all the classifications indicate the same result in gender or age range, the image is used in subsequent classifications.

The Input Layer consists of an agent that controls the webcam for

continuous image collection in real time. If a human face is found in these images, they are sent to the pre-processing layer where it is detected if the image has two eyes, if not it is discarded. It is important for the image to contain both eyes to be able to align it, using trigonometry if necessary (in most cases, since otherwise it would not be possible to compare it). Once aligned, the image is cropped to the desired size, the background of the image is discarded, eliminating extra, unnecessary information. Agents reduce the amount of information in the images by applying filters like the bilateral filter or Gabor filter forming the Filter layer. The Face recognition layer contains agents that develop the Fisherfaces, Eigenfaces, Local Binary Pattern (LBP) and Multilayer Perceptron (MLP) approaches with which training and classification processes are conducted. The Classification Layer contains the last mean faces which are compared with the captured image. In the Database layer, there are agents responsible for the database and agents related to the training performed with the faces stored in the database. In the database, images of correctly labeled faces by age and gender, with which the system will train. Each new image is incorporated into the system once its correct classification is confirmed. This confirmation consists of ten images being taken from the same person and the same classification result being obtained with an interval of 0.5s between each image. The Workflow Management Layer is responsible for communication between layers, it is also the layer that decides whether the filters in the filter layer should be used, this decision is made on the basis of whether the result has improved or not, when conditions previous conditions did not did good results (lighting, contrast, saturation, brightness or sharpness).

#### 4. Results

In order to work with the system developed with the Jade framework, it is necessary to load the image database. The FERET database was used in order to avoid the cold start problem. The FERET database is the standard database in the evaluation of facial recognition algorithms. The FERET database contains images that were collected/ taken by the cameras under semi-controlled conditions in order to evaluate the algorithms. This base of images has 14051 images with different angles. To conduct this study, we trained the Haar feature-based cascade classifier using a subset consisting of frontal images and only of individuals without glasses. This classifier allows us to detect the eyes of an individual in the pre-processing stage. The classifier was trained

with images of faces that did not wear any glasses because this allowed us to obtain better eye detection results in the preprocessing stage, for those who wore glasses and those who did not. When the classifier was trained with images of people who wore glasses, the results were not satisfactory.

Therefore, the initial image database used by the system has had a total of 930 images (465 images per gender). Preparing the database properly is essential, each image must be correctly labeled by its gender and age group and previously pre-processed, in order to train in the best possible way, in exactly the same conditions as in the classification. The groups in each of the iterations for the different methods and sets of images do not vary. For the classification of both gender and age in this case study has focused on the oval of the faces. While selecting the oval of a face image may seem simple, great differences can be found depending on the crop that is made. In the trimming stage, four different workbench have been created from the original database, specifically, four different versions will be created for analyzing if there are differences between the variations of the trimmed area with respect to the success rate.

Although the training of the system was carried out with images taken under controlled conditions (FERET database), the system has used this training in a real environment (without modifying the lighting conditions, without fixing the position of the face in front of the system's camera, without fixing the distance between the face and the system's camera, etc.). There was fluorescent lighting, OSRAM L 36 W/20 Cool White with MAZDA S1 Single 220-240V 4-65W starter, where the camera used by the system was located. However, the number of lumens varied during the case study due to the entry of natural light from outside, through the windows and doors.

#### 4.1. Classification of gender

The developed system has enabled us to carry out a comparative study of both methods (Fisherfaces and MLP), in different situations, in order to analyze their efficiency. The comparison between the two methods and the four separate sets of differently pre-processed images, specifically changing the way the oval of the face was cropped. The four sets correspond to the four vertical crops (pre-processed images) shown in Fig. 8. The agent that develops the neural network has been modeled using the Java Encog framework. This agent uses a multilayer perceptron as shown in Fig. 12 with a Sigmoid activation function and the Resilient Propagation function as a training method. As explained above, the use of the neural network involves an additional stage when dealing with images. Since the images are a triplicate bitmap, due to the PPM format in which the FERET database comes, to use them as the input of the neural network we will have to transform them into a 64x64 one-dimensional array with values in [0, 255] corresponding to the grayscale values of each of the pixels in the image. In addition, we do not pass the grayscale values directly, but instead convert them into the interval [-1, 1]. Each neuron from the output layer corresponds to one of the two genders, the output of the neuron that is the closest to 1 will be the one that marks the gender chosen by the network. The testing of both models was performed using the 5x2 cross-validation method, which allowed to evaluate the significance of the presented techniques (Pinzon et al., 2013). The four sets correspond to the four vertical crops (pre-processed images) shown in Fig. 8. With the use of this method, we will be certain that the data obtained from model testing are not dependent on the chosen training and test subsets. In addition, it allows us to make a comparison in the third phase between the different cropping methods and versions of cropped face images. In this paper we have successfully implemented and tested two main methods that are currently used in the task of gender classification, Fisherfaces and ANN, the success rate of which can be seen in Tables 1 and 2.

Each of these tables contain four columns, one for each data set corresponding to the different face trimming methods. For each dataset

**Table 1**  
Fisherfaces success rate.

	Set 1		Set 2		Set 3		Set 4	
	a	b	a	b	a	b	a	b
Iteration 1	0.903	0.881	0.896	0.894	0.883	0.888	0.892	0.896
Iteration 2	0.901	0.886	0.898	0.879	0.901	0.875	0.883	0.890
Iteration 3	0.892	0.875	0.896	0.883	0.886	0.879	0.883	0.886
Iteration 4	0.883	0.883	0.847	0.879	0.858	0.858	0.873	0.898
Iteration 5	0.907	0.881	0.892	0.883	0.886	0.858	0.898	0.892
Avg.	0.889		0.885		0.879		0.889	

**Table 2**  
ANN success rate.

	Set 1		Set 2		Set 3		Set 4	
	a	b	a	b	a	b	a	b
Iteration 1	0.866	0.894	0.855	0.888	0.877	0.894	0.864	0.877
Iteration 2	0.836	0.870	0.868	0.881	0.862	0.858	0.883	0.832
Iteration 3	0.864	0.860	0.894	0.862	0.883	0.851	0.883	0.836
Iteration 4	0.853	0.862	0.845	0.886	0.855	0.860	0.840	0.851
Iteration 5	0.896	0.817	0.847	0.855	0.843	0.832	0.840	0.888
Avg.	0.862		0.868		0.861		0.858	

a 5x2 cross validation has been performed, for this reason each dataset has five rows and two columns. The first column corresponds to the percentage of success as a result of training with 50% of the data belonging to subset B, the second column is the percentage of success when training with subset A and the prediction with subset B. The first column corresponds to the percentage of success as a result of training with 50% of the data belonging to subset B, the second column is the percentage of success when training with subset A and the prediction with subset B.

To test the two methods, we used the Mann–Withney test by setting an  $\alpha$  level equal to 0.05 with a sample size of  $n_1 = 10$ ,  $n_2 = 10$ . We establish a hypothesis  $H_0$  - the median of the set of fisherfaces and the median of the ANN set are the same for the same preprocessed image. A p-value of less than 0.05 (significance value) was obtained for each for the four comparisons, so  $H_0$  was rejected. Then we proceeded to establishing  $H_1$  - The median of the set fisherfaces is greater than the median of the ANN set. For this second comparison, we obtained in the four comparisons a p-value of less than 0.05 for the four comparisons, so  $H_1$  was accepted. Therefore, the best classification for the created datasets, can be obtained using the Fisherfaces method.

The sets of images for which the same method was used were compared. The Kruskal–Wallis test was applied to the results of the image sets for the two methods. To begin, we established an  $\alpha$  level equal to 0.05 and we proposed a null hypothesis.  $H_0$  - The four samples do not have significant differences. We calculated the p-value for Fisherfaces to be equal to 0.2825 and the p-value for ANN to be equal to 0.7267. Both p-values are greater than the  $\alpha$  level of significance equal to 0.05. We can conclude therefore, that there are no significant differences between the different sets created for the test, Fisherfaces and the neural network provide the same results.

#### 4.2. Classification of age range

In the problem of classification by age ranges, the following age ranges have been established: children (< 18 years), youth (19–30 years), adults (31–70 years) and the elderly (> 70 years). The faces that make up these training and classification groups are all headshots. There are 115 images in the childrens group (not belonging to FERET), 310 in the youth group, 310 in the adult group and 310 in the elderly peoples group. As the FERET database does not contain images of children these have been manually collected, so the number of

**Table 3**  
Success rate.

	Without filter	Bilateral filter	Gabor filter
Fisherfaces	83,62%	91,93%	93,60%
Eigenfaces	77, 81%	75,16%	76,86%
LBP	74,62%	77,34%	76,10%

childrens faces is not very high. The system estimates the age in ten times in a interval for each individual that appears on the input image, in this way, the most obtained age range in the ten estimates will be the chosen one, causing the system to have a greater degree of success in real time. To verify the rate of the systems accuracy, a cross validation was performed with the data that was used in the training.

This test was performed with Fisherfaces, Eigenfaces and Local Binary Pattern (LBP), and Gabor and bilateral filtering were applied to both training and testing data. It was possible to verify how the Eigenfaces method is not acceptable in classification tasks related to age, since it considers light adjustment factors. Table 3 shows Fisherfaces as the most suitable method, and how after applying a filter the success rate increases slightly. Next, the confusion matrices of the different methods are shown. It can be seen that in the different matrices of confusion, the success rate is lower in the estimation of youth and adults. In Tables 4–6 show the confusion matrix of the different methods.

## 5. Conclusions

Different classification methods could be compared in this work thanks to the use of a multi-agent system designed for their execution. This technology based on agents allows to greatly improve the results of the global system, as different agents are in charge of carrying out specific roles in the processes that occur in the system. Each agent applies a different methodology and the best solution to the common goal pursued by the agents in the system is obtained. For example, when choosing the most suitable recognition algorithm (Fisherface, Eigenface, LBP, ANN) or the implemented filters (Gabor, Sobel). The system allows to easily configure different workflows and analyze the

**Table 4**  
Fisherfaces confusion matrix.

	Children	Youth	Adults	Elderly
Children	69	0	46	0
Youth	0	310	0	0
Adults	0	62	248	0
Elderly	0	62	124	124

**Table 5**  
Eigenfaces confusion matrix.

	Children	Youth	Adults	Elderly
Children	92	0	0	23
Youth	0	248	62	0
Adults	0	93	186	31
Elderly	0	93	62	155

**Table 6**  
LBP confusion matrix.

	Children	Youth	Adults	Elderly
Children	92	0	0	23
Youth	0	155	155	0
Adults	0	0	217	93
Elderly	0	31	186	93

efficiency of different configurations. For this reason, it is possible to apply different preprocessing methods and different types of classifiers. Moreover, new configurations can be added to the design and evaluated for inclusion in future versions of the system. On the basis of the results shown in Tables 1 and 2, we reach to the conclusion that the preprocessing stage is very important; when the preprocessing stage was conducted correctly, the performance of the two proposed classification methods also improved. A well-trimmed face shape had the greatest influence on the quality of the preprocessed image. Therefore, the quality of the preprocessed images is key for the good performance of classification methods. From comparing the four sets of images created in this work (two per each classification method), we can verify that each of them reaches a sufficiently good degree of abstraction. We can also conclude that the features required for the training of the classifiers were contained in each of these sets. For this reason, superfluous information that could harm the preprocessing quality, were disregarded by creating different cropping areas that contained only one of all the features in all picture sets, this was more than enough for training a new classifier.

With regards to the problem of gender classification, the classifiers used (Fisherfaces and MLP) are fully dependent on the preprocessing stage, unlike other methods in the literature. The Fisherfaces classifier has proved to be significantly more effective in the classification of gender than neural networks and it required less training time. As to age group classification, these classifiers were chosen for their wide and well organized image bases. Although the scientific community has made efforts to develop new algorithms that use new feature extraction methods, biometric calculations and complementation between different system types and algorithms, Fisherfaces and MLP continue offering acceptable results which improve with the use of filters. The images of individuals who were on the verge of being in one group or another, for example a teenager that could also be classified as an adult, could be included in both groups. However, individuals who could be classified into the two groups posed a problem to the system. This was solved by deleting the headshots that were classified into both groups, although this was not a simple task as the images from the database were classified on the basis of personal perception. It is quite striking that the estimation of childrens age range has a great success rate in the two methods although the database is much lower than with any other age ranges. This is because the eigenfaces generated by the eigenvectors from the childrens group, have a greater difference in the projection that characterizes the facial image of the childrens group, in comparison to the other groups. More specifically, the sum all the facial features of different weights from the childrens group makes up an eigenface whose distance is greater because of the changes that occur in the structure and characteristics of the face. In the teenagers, adults and elderly peoples groups the facial changes are minor (such as, appearance of wrinkles and spots or the darkening of the skin).

The images were obtained from the FERET database. Racial variability and anatomical differences can make the task of recognizing gender difficult even for humans. We were aware that these differences could also influence the training of our classifier, therefore, we presented different solutions that helped solve this problem. For example, cropping that reduces the area of an image, leaving us only with the most differentiating facial features or the use of different classifiers whose choice depended on the origin of the individual. The performance of the system also improved when different classification methods were combined.

This work also analyzed how the use of different filters influenced the results, this analysis allowed us to understand which type of filtering aided the performance of the classification methods. When applying the dimensionality reduction technique, the Gabor filter and the Bilateral filter allow fewer dimensions to be obtained, thus reducing the work load; something that is very valuable in a system that classifies in real-time and has a constant inflow of images to compute and which should not be accumulated.



## Acknowledgments

This work has been supported by the Spanish Government through the project SURF (grant TIN2015-65515-C4-3-R) and FEDER funds. The research of Alfonso González-Briones has been co-financed by the European Social Fund (Operational Programme 2014–2020 for Castilla y León, EDU/128/2015 BOCYL).

## References

- Agarwal, M., Agrawal, H., Jain, N., Kumar, M., 2010. Face recognition using principle component analysis, eigenface and neural network. *Signal Acquisition and Processing*, 2010. ICSAP'10. International Conference on. IEEE, pp. 310–314.
- Ahonen, T., Hadid, A., Pietikainen, M., 2006. Face description with local binary patterns: application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (12), 2037–2041.
- Alim, M.A., Baig, M.M., Mehboob, S., Naseem, I., 2017. Method for secure electronic voting system: face recognition based approach. *Second International Workshop on Pattern Recognition*. 10443. International Society for Optics and Photonics, pp. 104430H.
- Bajo, J., De Paz, J.F., Villarrubia, G., Corchado, J.M., 2015. Self-organizing architecture for information fusion in distributed sensor networks. *Int. J. Distrib. Sensor Netw.* 2015.
- Banerjee, P.K., Datta, A.K., 2014. Class specific subspace dependent nonlinear correlation filtering for illumination tolerant face recognition. *Pattern Recognit. Lett.* 36, 177–185.
- Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J., 1997. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (7), 711–720.
- Bro, R., Smilde, A.K., 2014. Principal component analysis. *Anal. Methods* 6 (9), 2812–2831.
- Cho, H., Lee, H., Kang, H., Lee, S., 2014. Bilateral texture filtering. *ACM Trans. Graphics (TOG)* 33 (4), 128.
- Choi, S.E., Lee, Y.J., Lee, S.J., Park, K.R., Kim, J., 2011. Age estimation using a hierarchical classifier based on global and local facial features. *Pattern Recognit.* 44 (6), 1262–1281.
- Chu, W.-S., Huang, C.-R., Chen, C.-S., 2013. Gender classification from unaligned facial images using support subspaces. *Inf. Sci.* 221, 98–109.
- Dadi, H.S., Mohan, P.K., 2015. Performance evaluation of eigen faces and fisher faces with different pre-processed data sets. *Int. J. Adv. Res. Comput. Eng. Technol. (IJARCE)* 4 (5), 2110–2116.
- Dandpat, S.K., Meher, S., 2013. Performance improvement for face recognition using pca and two-dimensional pca. *Computer Communication and Informatics (ICCCI)*, 2013 International Conference on. IEEE, pp. 1–5.
- De Paz, J.F., Bajo, J., Rodríguez, S., Villarrubia, G., Corchado, J.M., 2016. Intelligent system for lighting control in smart cities. *Inf. Sci.* 372, 241–255.
- Dehshibi, M.M., Bastanfard, A., 2010. A new algorithm for age recognition from facial images. *Signal Process.* 90 (8), 2431–2444.
- Delgado-Gomez, D., Fagertun, J., Ersbøll, B., Sukno, F.M., Frangi, A.F., 2009. Similarity-based fisherfaces. *Pattern Recognit. Lett.* 30 (12), 1110–1116.
- Fisher, R.A., 1936. The use of multiple measurements in taxonomic problems. *Ann. Hum. Genet.* 7 (2), 179–188.
- Frikha, T., Siala, Y., Louati, M., Abid, M., 2016. Use of ridgelets, curvelets application for face recognition: Case study: smart identity card. *Advanced Technologies for Signal and Image Processing (ATSIP)*, 2016 2nd International Conference on. IEEE, pp. 393–397.
- Geng, X., Yin, C., Zhou, Z.-H., 2013. Facial age estimation by learning from label distributions. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (10), 2401–2412.
- González-Briones, A., Ramos, J., De Paz, J.F., Corchado, J.M., 2015. Multi-agent system for obtaining relevant genes in expression analysis between young and older women with triple negative breast cancer. *J. Integr. Bioinf. (JIB)* 12 (4), 1–14.
- Guo, G., Mu, G., Fu, Y., Huang, T.S., 2009. Human age estimation using bio-inspired features. *Computer Vision and Pattern Recognition*, 2009. CVPR 2009. IEEE Conference on. IEEE, pp. 112–119.
- Han, H., Otto, C., Jain, A.K., 2013. Age estimation from face images: human vs. machine performance. *Biometrics (ICB)*, 2013 International Conference on. IEEE, pp. 1–8.
- Imran, M., Miah, M., Rahman, H., Bhowmik, A., Karmaker, D., 2015. Face recognition using eigenfaces. *Int. J. Comput. Appl.* 118 (5).
- Jana, R., Datta, D., Saha, R., 2015. Age estimation from face image using wrinkle features. *Procedia Comput. Sci.* 46, 1754–1761.
- Jana, R., Pal, H., Chowdhury, A.R., 2012. Age group estimation using face angle. *IOSR J. Comput. Eng. (IOSRJCE)* 7 (5), 35–39.
- Kwon, Y.H., da Vitoria Lobo, N., 1999. Age classification from facial images. *Comput. Vision Image Understanding* 1 (74), 1–21.
- Lai, C.-C., Ko, C.-H., 2014. Facial expression recognition based on two-stage features extraction. *Optik-Int. J. Light Electron Opt.* 125 (22), 6678–6680.
- Lanitis, A., 2010. Facial age estimation. *Scholarpedia* 5 (1), 9701.
- Lanitis, A., Taylor, C.J., Cootes, T.F., 2002. Toward automatic simulation of aging effects on face images. *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (4), 442–455.
- Levi, G., Hassner, T., 2015. Age and gender classification using convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. pp. 34–42.
- Lin, W.-H., Wang, P., Tsai, C.-F., 2016. Face recognition using support vector model classifier for user authentication. *Electron. Commerce Res. Appl.* 18, 71–82.
- Liu, C., Wechsler, H., 2002. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Trans. Image Process.* 11 (4), 467–476.
- Lu, C.-Y., Min, H., Gui, J., Zhu, L., Lei, Y.-K., 2013. Face recognition via weighted sparse representation. *J. Visual Commun. Image Represent.* 24 (2), 111–116.
- Majumder, A., Behera, L., Subramanian, V.K., 2014. Emotion recognition from geometric facial features using self-organizing map. *Pattern Recognition* 47 (3), 1282–1293.
- Mehra, S., Charaya, S., 2016. Enhancement of face recognition technology in biometrics. *Int. J. Sci. Res. Educ.* 4 (08).
- Özdil, A., Özbilen, M.M., 2014. A survey on comparison of face recognition algorithms. *Application of Information and Communication Technologies (AICT)*, 2014 IEEE 8th International Conference on. IEEE, pp. 1–3.
- Panis, G., Lanitis, A., 2014. An overview of research activities in facial age estimation using the fg-net aging database. *European Conference on Computer Vision*. Springer, pp. 737–750.
- Pinzon, C.I., De Paz, J.F., Herrero, A., Corchado, E., Bajo, J., Corchado, J.M., 2013. idMAS-SQL: intrusion detection based on MAS to detect and block SQL injection through data mining. *Inf. Sci.* 231, 15–31.
- Parkhi, O.M., Vedaldi, A., Zisserman, A., et al., 2015. Deep face recognition. *BMVC.* 1. pp. 6.
- Pentland, A., Moghaddam, B., Starner, T., et al., 1994. View-based and modular eigen-spaces for face recognition. *CVPR.* 94. pp. 84–91.
- Rai, P., Khanna, P., 2014. A gender classification system robust to occlusion using gabor features based (2d) 2 pca. *J. Visual Commun. Image Represent.* 25 (5), 1118–1129.
- Ramos, J., Castellanos-Garzón, J.A., González-Briones, A., de Paz, J.F., Corchado, J.M., 2017. An agent-based clustering approach for gene selection in gene expression microarray. *Interdiscip. Sci. Comput. Life Sci.* 9 (1), 1–13.
- Robertson, D.J., Burton, A.M., 2016. Unfamiliar face recognition: Security, surveillance and smartphones. *J. Homeland Defense Secur. Inf. Anal. Center* 14–21.
- Robertson, D.J., Noyes, E., Dowsett, A.J., Jenkins, R., Burton, A.M., 2016. Face recognition by metropolitan police super-recognisers. *PLoS One* 11 (2), e0150036.
- Shaikh, W., Shinde, H., Sharma, G., 2016. Face recognition using multi-agent system. *System* 1, 3.
- Shan, C., 2010. Learning local features for age estimation on real-life faces. *Proceedings of the 1st ACM International Workshop on Multimodal Pervasive Video Analysis*. ACM, pp. 23–28.
- Sharifara, A., Rahim, M.S.M., Anisi, Y., 2014. A general review of human face detection including a study of neural networks and haar feature-based cascade classifier in face detection. *Biometrics and Security Technologies (ISBAST)*, 2014 International Symposium on. IEEE, pp. 73–78.
- Sharma, P., Arya, K., Yadav, R.N., 2013. Efficient face recognition using wavelet-based generalized neural network. *Signal Process.* 93 (6), 1557–1565.
- Shyam, R., Singh, Y.N., 2015. Identifying individuals using multimodal face recognition techniques. *Procedia Comput. Sci.* 48, 666–672.
- Sun, Y., Liang, D., Wang, X., Tang, X., 2015. Deepid3: face recognition with very deep neural networks. [arXiv:1502.00873](https://arxiv.org/abs/1502.00873).
- Surekha, B., Nazare, K.J., Raju, S.V., Dey, N., 2017. Attendance recording system using partial face recognition algorithm. *Intelligent Techniques in Signal Processing for Multimedia Security*. Springer, pp. 293–319.
- Swets, D.L., Weng, J.J., 1996. Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Pattern Anal. Mach. Intelligence* 18 (8), 831–836.
- Taigman, Y., Yang, M., Ranzato, M., Wolf, L., 2014. Deepface: closing the gap to human-level performance in face verification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1701–1708.
- Villarrubia, G., De Paz, J.F., De La Prieta, F., Bajo, J., 2014. Hybrid indoor location system for museum tourist routes in augmented reality. *Information Fusion (FUSION)*, 2014 17th International Conference on. IEEE, pp. 1–8.
- Viola, P., Jones, M., 2001. Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. 1 IEEE. I–I
- Wang, J., Jia, L., 2016. Eigenfaces vs. fisherfaces: recognition using class specific linear projection.
- Yang, J., Ren, P., Chen, D., Wen, F., Li, H., Hua, G., 2016. Neural aggregation network for video face recognition. [arXiv:1603.05474](https://arxiv.org/abs/1603.05474).
- Zhang, B., Shan, S., Chen, X., Gao, W., 2007. Histogram of gabor phase patterns (hgpp): a novel object representation approach for face recognition. *IEEE Trans. Image Process.* 16 (1), 57–68.