

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2019.Doi Number

Robust detection of fatigue parameters based on infrared information

Carlos M. Travieso-González^{1,2*}, Jesús B. Alonso-Hernández^{1,2}, José M. Canino-Rodríguez², Santiago T. Pérez-Suárez^{1,2}, David Sánchez-Rodríguez^{1,3}, Antonio G. Ravelo-García^{1,2,4*}

¹Institute for Technological Development and Innovation in Communications, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, 35017, Spain.

²Department of Signals and Communications, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, 35001, Spain.

³Department of Telematics Engineering, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, 35001, Spain.

⁴Madeira Interactive Technologies Institute (ITI/Larsys/M-ITI), 9020-105 Funchal, Portugal

Corresponding author: (e-mail: antonio.ravelo@ulpgc.es; carlos.travieso@ulpgc.es).

ABSTRACT Driver fatigue is one of the major causes of traffic accidents, and this need has increased the amount of driver fatigue detection systems in vehicles in order to reduce human and material losses. This work puts forward an approach based on capturing near-infrared videos from a camera mounted inside the vehicle. Then, from the captured images and using image-processing techniques the eyes are detected. Next, features are extracted from eye images using several transforms and finally, the system detects if there is fatigue or not using a SVM as classifier. Throughout the recording, eye position is tracked with a low computational time and fatigue is analysed based on the percentage of eyelid closure. This approach has been developed on two public datasets. Our experiments were able to reach an eye recognition rate of up to 96.87% and our results showed that SVM with RBF kernel were 99.66% accurate on one of the databases used for the system training. This approach shows promising results in comparison with the state of the art and deep learning approaches in order to be implemented in real conditions.

INDEX TERMS eye detection and identification, infrared imaging, fatigue parameters, pattern recognition

I. INTRODUCTION

Nowadays, society is developing and progressing rapidly and demand for new technologies, systems and devices focus heavily on safety. Safety has been defined as a condition reached through the systematic process of identifying risks and developing methods to minimize these risks. These preventive measures have been refined and safety systems can be divided into active and passive systems. Active safety systems try to avoid accidents, while passive safety systems look for ways to minimize the effects on people involved in accidents.

The tendency to incorporate electronic systems for vehicle safety is an irrefutable fact [1]. Despite the significant decrease in mortality rates in traffic accidents, drowsiness and sleep disorders are an important factor and can cause up to 10% of all accidents [2]. Symptoms of fatigue while driving are the following [3]: decreased attention and concentration, decreased reflexes, changes in perception, changes in body movements, feeling of physical discomfort, headaches, dizziness, and pain in the neck, back and arms, etc. The National Highway Traffic Safety Administration of USA

reported that about 825 people die a year in over 730 accidents caused by drowsy driving [4].

Currently, there is a large body of research studying the use of different methods to detect fatigue in drivers. Some of them are based on control systems located on the steering wheel and pedal when the driver falls asleep. There are also systems that track the lane markings on the road to prevent the vehicle from running off it.

The interest in developing methods capable of measuring the degree of fatigue or alertness in drivers it is not new, as it dates back to the 70s. But it was not until mid-90s when it began to develop. In [5], authors state the three main reasons for the development of these technologies as the following: fatigue is a persistent risk for professional drivers, these drivers are under pressure to reach their destinations regardless tiredness, and finally the driver cannot correctly assess their condition due to fatigue's adverse effects.

Most of the information that is needed while driving is visual. Hence, techniques based on monitoring ocular activity can better assess the driver's state. The basic premise is that eye behavior can provide the necessary information on driver

alertness. Increased blinking, the speed of pupil movement or the percentage of eyelid closure known as PERCLOS [6] are variables to analyze. In recent years, numerous studies have focused on the use of cameras to detect facial features, especially eye measurements. These techniques are more practical as they are non-invasive and computer vision is growing rapidly. In addition, studies have shown that the PERCLOS parameter has consistent, valid and high rates of reliability compared to PVT (Psychomotor Vigilance Test) [7] when detecting and predicting fatigue in drivers.

Researchers have explored different methods to develop drowsiness detection systems. Image processing and computer vision were among the most popular methods, where drowsiness is detected via eyelid movement tracking, face and eye detection, facial expressions, head movements, and eye behavior characteristics [8-10].

The main proposal in this article is a system capable of monitoring the degree of fatigue in drivers, as in [11], which calculates parameters that have a direct relation to the level of driver fatigue. This system is based on driver imaging by using a near-infrared camera located inside the vehicle, using public datasets. First, the images are processed so that the eye position can be reliably located. Subsequently, methods are used which predict eye movement and update their position, so as to reduce computation time. Finally, the level of driver fatigue is assessed by measuring eye activity. In order to achieve this, the following steps are carried out: study the methodology used, design a recording system, implement a database, develop software to detect and track eyes in a video sequence, and calculate parameters to determine the driver alertness. Lastly, the results are compared with other research showing that the proposed fatigue monitoring system is effective and that the computational load required by our system is appropriate.

The rest of the paper is organized as follows: Section 2 describes the proposed fatigue detection system. Section 3 shows the results from the verification and eye tracking stages. In section 4 a discussion is held on how to best apply the system found in the results. Finally, the conclusions derived from this work are presented in Section 5.

II. PROPOSED SYSTEM AND METHODOLOGY USED

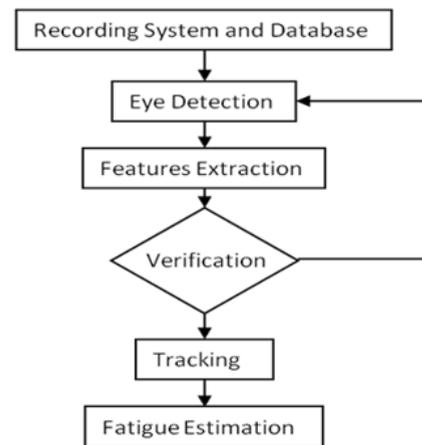
The proposed fatigue monitoring approach is shown in figure 1, with a robust prototype. The system is made up of different modules where each module has a specific task such as automatic eye location, feature extraction and selection to estimate the degree of fatigue. It is important to mention that, although the image sequences processed may be recorded by the camera, it can also use video sequences from databases. After obtaining the images, the next task is to automatically detect the position of the driver's eyes.

A. RECORDING SYSTEM AND DATABASES

A recording system is used to obtain sufficiently clear and bright images of drivers. It consists of a video camera equipped to record in the near-infrared (NIR) range and as well

as a subsystem equipped with Infrared Emitting Diodes (IRED) to ensure proper lighting under different ambient conditions, without affecting the driver's vision.

FIGURE 1. Proposed fatigue Monitoring System.



The characteristics of this system and the obtained images are the following:

- System image capture: common camera with a 352×288 pixel CMOS sensor. The camera filter is replaced by an optical filter to block out visible light (400-700 nm). Thus, the spectral response of this filter strongly attenuates wavelengths below 780 nm and allows longer wavelengths for transmission components. The camera is located 40-75 cm from the person. Furthermore, the lighting system is made up of a ring of IREDs with the camera in the centre.
- Images: video clips with 640×480 pixels per frame, 8-bit colour and a recording speed of 30 fps.

In this paper, the authors have also used two public databases from other research groups for two purposes, under their conditions:

- To create a specific database with images of people's eyes for training and testing, which is extracted from two public databases.
- To study the quality of algorithms used to automatically locate the position of eyes from a facial image.

The databases used for this purpose are TUNIR [12] and CBR NIR [13]. Both databases have different images, hence the efficiency of the eye detection algorithm from images with different resolutions can be checked.

The most significant characteristics of the TUNIR database (Figure 2) are the following:

- Composed of 64 video sequences of different people recorded in NIR wavelengths. People belong to different nationalities.
- Recording System: a common camera with a silicon sensor and a filter that removes daylight; a lighting system made up of 12 Aluminium Gallium Arsenide (AlGaAs) diodes emitting at 875nm. Furthermore, the camera position is 0.39 - 0.78m from the people with an

inclination angle $\alpha=30^\circ$

- Images features: 320×240 pixels of resolution and recorded at a rate of 30 fps in 'AVI' format.



FIGURE 2. Example of capture video sequences of TUNIR database.

The CBSR NIR database was developed in 2007 by Li, et al. in order to study facial recognition under varying lighting conditions [13]. Therefore, it can also be used for other purposes such as eye detection.

The main features of this database are:

- 3940 images of 197 different people.
- Recording system: a sensitive NIR camera and a normal camera placed on the same vertical and spaced about 12 cm apart between the two axes. The lighting system consists of 18 NIR LEDs ring. Vertical axis connecting the two chambers is located 70 cm from the person and is perpendicular to the line of sight (see figure 3).
- Images features: 640×480 pixel resolution, encoded with 8 bits grayscale, compressed and in bmp format.

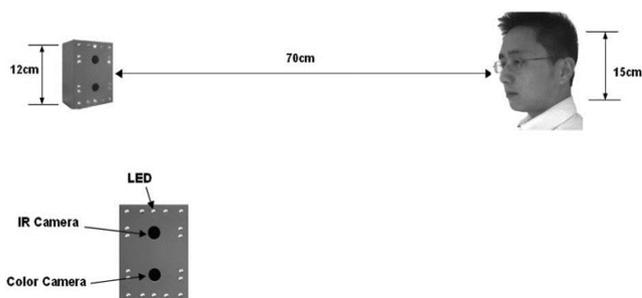


FIGURE 3. Schematic recording system database NIR CBSR.

Figure 4 is a sample of the images that make up the CBSR NIR database.



FIGURE 4. Example of images in CBSR database.

Moreover, in the verification stage, images are required to train and test the classifiers, in order to discern whether it has properly located the position of the eyes. Therefore, from previous databases a set of samples for these training and

testing phases are obtained. Then the two sample types have been obtained:

- Eye images are taken from pupil images of people who appear in the database. Passing these images through the classifier should identify them positively.
- Non-eye images with random contents are used to train the classifier.

As a result of using the images available in the TUNIR database [12], a new dataset has been created with the following characteristics (Figure 5):

- 300 images of eyes
- 64 non-eyes images of different people
- 31×31 pixels
- Image format: bitmap



FIGURE 5. Images of eyes and non-eyes obtained from the TUNIR database.

Similarly, the following dataset has been collected from the CBSR NIR database (Figure 6):

- 3152 images of eyes
- 107 non-eyes images of different people, 16 images per person
- 61x61 pixels

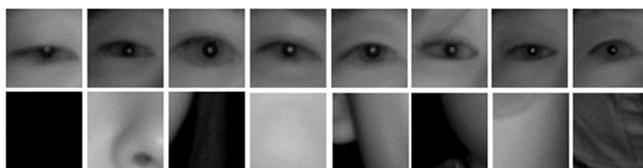


FIGURE 6. Images of eyes and non-eyes obtained from the CBSR NIR database.

B. EYE DETECTION STAGE

The first stage of the system is to implement image detection techniques to get the eye position. The proposed algorithms to perform this task are: valleys detection, the Circle Hough Transform and brightness detection [14].

1) VALLEYS DETECTION

The method of finding valleys is to calculate the $v_1(x,y)$ and $v_2(x,y)$ values. These values are calculated with the equations below.

$$v_1(x,y) = C_{11}\phi_{11} + C_{12}\left(\frac{I(x-2,y)+I(x+2,y)}{2}\right) - S_{11}(x,y) \quad (1)$$

$$v_2(x,y) = C_{21}\phi_{21} + C_{22}\left(\frac{I(x-3,y)+I(x+3,y)}{2}\right) - S_{22}(x,y) \quad (2)$$

First, valley intensity is calculated. For this purpose, the morphological close operation is applied to the original

image. The result is subtracted from the original image to obtain the intensity matrix valley. The next step is to compute the mean value for each pixel in the image. For this, the values $\phi_{11}(x,y)$, $\phi_{21}(x,y)$, $S_{11}(x,y)$ and $S_{22}(x,y)$ are calculated from the grayscale image and the valley intensity previously obtained.

Mean values of pixels around the point (x,y) of the image are $S_{11}(x,y)$ and $S_{22}(x,y)$. The only difference between them is the size of the matrix taken for the average, 3×3 and 5×5 pixels respectively. The values $\phi_{11}(x,y)$ and $\phi_{21}(x,y)$ are calculated likewise as $S_{11}(x,y)$ and $S_{22}(x,y)$, with the exception that the mean values are obtained from the intensity valley. Subsequently, C_{11} , C_{12} , C_{21} and C_{22} coefficients are adjusted to improve the algorithm.

Once all the values are found, the $v_1(x,y)$ and $v_2(x,y)$ vectors are computed. The points where the values are simultaneously higher correspond to regions of interest where eyes can be found.

2) HOUGH TRANSFORM

Circular Hough Transform (CHT) based algorithm is used to find circles in images. This approach is useful because of its robustness in the presence of noise, occlusion and irregular lighting. This algorithm involves three steps:

- 1) Accumulator Array Computation: The first step of the CHT is to compute the whitest points, closest to each other, to draw the surrounding circle.
- 2) Center Estimation: The second step of this algorithm is to detect and calculate the center within the white points calculated in the first step.
- 3) Radius Estimation: The last step of CHT is to calculate the radius of the white points. This step can be carried out by computing the farthest white point from the center, and then, draw the circle around the calculated points.

3) BRIGHTNESS DETECTION

To locate the position of the eyes, they are illuminated with light emitting diodes. The image of the driver is eroded using a structural element. In this way, a very similar image to the original is got, but where the brightest spots are dimmed. Finally, the difference between the result and the original image is obtained.

These three methods were applied on two datasets in order to check the best proposal for the segmentation process, before the feature extraction and classification steps. The best method was brightness detection and only these accuracies will be shown, which using brightness detection.

C. FEATURE EXTRACTION AND SELECTION

As a result of applying the above eye detection methods a matrix of the same size as the original image is obtained. The highest values are likely to coincide with the position of an eye. From these results, images should then be taken around the highest values and passed to the next stage. The regions of interest should be analyzed to validate the result. This phase consists in extracting their characteristics and verifying if those regions correspond to an eye. Three

transformations have been made: The Discrete Cosine Transform (DCT), the Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) [14].

1) DISCRETE COSINE TRANSFORM

The DCT bidimensional equation, for an image of size $M \times N$, is exposed below.

$$D(i,j) = \frac{1}{\sqrt{MN}} C(i)C(j) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} p(x,y) \cos \left[\frac{(2x+1)\pi i}{2M} \right] \cos \left[\frac{(2y+1)\pi j}{2N} \right] \quad (3)$$

For the previous expression $i = 0, 1, 2, \dots, M-1$ and $j = 0, 1, 2, \dots, N-1$; where $p(x,y)$ is the intensity value for the pixel with the (x,y) position, $C(u) = 1/\sqrt{2}$ for $u = 0$, and $C(u) = 1$ for $u > 0$.

The energy of the DCT is located near the low frequency components. One of the main advantages of this transformation is that it eliminates redundancy between neighbouring pixels. Most of the DCT coefficients information is included in the low-frequency coefficients, providing information about the pupil, iris, sclera and the eyelids.

The DCT is applied to images that have already been processed in the detection and segmentation phases, as the CBSR NIR database and TUNIR database. Moreover, the size of the original image is not reduced with this transformation. Thus, in the case of the CBSR NIR database, the DCT output will remain at 61×61 pixels; in the case of the TUNIR database, the output will be 31×31 pixels. After choosing the size of the square matrix, as the number of coefficients that will be used, the matrix must be converted into a feature vector. There are several alternatives to carry out this step: zigzag algorithm, columns concatenation and rows concatenation. According to previous studies, the algorithm with best results is the rows concatenation.

2) THE WAVELET TRANSFORM

The definition of the wavelet transforms for a continuous function $f(t)$ is given by:

$$F(a,b) = \int f(t) \Psi_{a,b}(t) dt \quad (4)$$

In the above equation, a and b are scale parameter and translation parameter respectively. These parameters vary continuously on the real axis. The functions $\Psi_{a,b}(t)$ are the wavelets functions. Wavelet functions are defined as:

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \quad (5)$$

DWT is interesting for image analysis because of the decomposition of image frequencies between high and low frequencies, where the noisy and particularized can be smoothed out, facilitating the detection processes. For this purpose, by using a DWT whose relations in time ($m = 0, 1, 2, \dots$) and scale ($n = 0, 1, 2, \dots$) are dyadic according to the below expression:

$$\Psi_{m,n}(k) = 2^{\frac{-n}{2}} \Psi \left(2^{\frac{-n}{2}} k - mb_0 \right) \quad (6)$$

In the continuous time equation, the scaling factor a , has been replaced by a power of two. Thus, the frequency axis corresponds to octaves decomposition. For the selected DWT a multiresolution analysis approach has been chosen. This analysis is useful for studying information contained in an image.

A pyramidal algorithm has been used which applies a one dimensional DWT in each image dimension, first in one dimension and then in the other. If an image of size $M \times N$ is assumed, the DWT is applied first to dimension m , and then to n . The transformation is $DWT_m(DWT_n(x[m,n]))$.

The DWT decomposes the image into subbands that are located in frequency and space. A wavelet transform is performed after passing the image through a series of filters. The high-pass filter (wavelet function) and the low-pass filter (scalar function) are finite impulse filters. In other words, each output only depends on a finite portion of the input. The filtered outputs are then subsampled by a factor of 2 in the horizontal direction. These signals are then filtered by another identical pair of filters in the vertical direction. The signal is decomposed into four subbands called HH, HL, LH and LL.

Each subband is interpreted as a smaller version of the original image and represents different properties. Firstly, it can assume that the effect of the different facial expressions may be mitigated by eliminating high frequency components. Secondly, only low frequency components are enough for eye recognition.

After applying the DWT, the image size is reduced. The type of wavelet family must be selected for appropriate performances. Experiments have been carried out with four function families. Firstly, the Haar function wavelet was used because of its simplicity. Secondly, the Daubechies function was used, in particular the db1 function, which is similar to Haar function. Furthermore, two other wavelets belonging to the Biorthogonal function were also used. These families were Bior 3.3 and Bior 4.4. It was tested from one to four iterations. After obtaining the reduced image, it is transformed to a vector, in order to get the input feature for the classification step. The feature vector was created by rows concatenation algorithm.

3) PRINCIPAL COMPONENT ANALYSIS

The main objective of the Principal Component Analysis is to transform a number of correlated variables into a smaller number of uncorrelated ones. These new variables are called principal components or factors, and they will be a linear and independent combination of the original variables. It is important to note that if the input values do not have zero mean, then the mean value must be calculated and subtracted.

Therefore, the number of principal components used for CBSR NIR and TUNIR databases are limited to 3152 and 293, just half the number of samples of the respective datasets. The experiments took into account the main

components limit allowed in each case. After applying the algorithm to obtain the output matrix, it is observed that the dimensions are the number of samples multiplied by the number of principal components. For CBSR NIR database, it is 3152 multiplied by the number of principal components. For TUNIR database, it is 293 multiplied by the number of principal components. Once the corresponding output matrix is obtained, the feature vector is created for the classifier. In this case, the algorithm used is columns concatenation.

These methods will be applied on two public datasets and after experiments will be chosen the best option.

D. VERIFICATION

A supervised classifier was applied. The choice is based on the great number of data, and therefore, a support vector machine (SVM) classifier was implemented for the discrimination of the outputs obtained from the different feature extractor modules or discriminative common vectors for both public datasets.

The SVM is a multiclass type with Radial Basis Function (RBF) as kernel.

$$K(x, y) = e^{-\gamma(x-y)^2} \quad (7)$$

The gamma (γ) value is estimated as the mean value of the samples and the average size of the sample vectors. The optimal value of gamma (γ) is estimated by searching in the range from 0 to 1 for each of the experiments.

In the experiment, the classifiers were trained using half of the samples chosen randomly from each class and the remaining samples were used for the test phase.

Another parameter that has been set is the parameter c : the cost of the training. Various values of c are taken (5, 10, 50, 100), obtaining various results for each experiment and taking the best in each case.

E. TRACKING

For each frame the eye position is tracked from their initial position obtained during the verification stage. This strategy reduces processing time. Different tracking strategies were evaluated: block matching, optical flow and extended Kalman filter [15].

1) BLOCK MATCHING

The Tracking by Block Matching uses the current frame and a reference image where the position of the eyes is known. Each object is tracked individually, and each eye is treated independently.

First, a macroblock is defined in the reference image focused on the eye. Then, a search area is defined in the current frame using the position occupied by the eye in the reference image. The technique is based on finding the macroblock in the current frame's search region that is the most similar to the reference image's macroblock. For this, Summation of Absolute Difference (SAD) is calculated in the searching area between macroblocks. For each comparison, a measure of image distortion is obtained. The macroblock, which has

greater similarity to the reference, will lower distortion further and find the motion vector.

2) OPTICAL FLOW

The principal objective of optical flow estimation is to separate the moving eyes from the rest of the face and generate optical flow field vector for the eyes. Optical flow calculates the motion between two frames, which are taken at different time intervals for every pixel in the frame. It is a discriminative method to track each eye, and it is a binary classification problem in a local image region.

3) EXTENDED KALMAN FILTER

Extended Kalman Filter (EKF) is one of the most common and popular filtering approaches in nonlinear target tracking and state estimation. It includes state estimation of a nonlinear dynamic system, parameters estimation for nonlinear system identification and dual estimation where both states and parameters are estimated simultaneously. However, EKF simply linearizes all nonlinear functions to the first order by using the Taylor series expansion. At the same time, EKF may cause more errors for the nonlinear system while estimating system state and its variance. Moreover, the linearization may lead to a divergence of the filtering process. In a nonlinear mismatched model and limited applications scope, EKF filter will lead the divergence problem of state estimation.

The three methods have been applied on two public datasets but it only will be the best results with independence of the method.

F. FATIGUE ESTIMATION

Once the above steps are performed, a system has been implemented to detect and know the position of the driver's eyes, who is being recorded with a near-infrared camera (see figure 7).

Next, when the eye position is known, it is possible to calculate parameters in direct relation to the level of driver fatigue [11] such as the eye movement, the percentage of eye closure (PERCLOS) or the speed opening and closing the eyes.

It should be also mentioned that once the eyes are located, it can indicate other facial features, such as the nose or mouth [11]. Figure 8 shows an image of the CBSR NIR database, in which the eyes, nostrils and four points around the mouth are located.

In the state-of-the-art, there are a number of parameters that have been studied to determine a person's fatigue [11]. For example, from the points associated with the contour of the mouth, a parameter to calculate the frequency of the driver's yawns is defined. The position of the nose, a feature that remains static in the face, can be used to estimate the movement of the face and detect a fatigued driver nod.

While there are many parameters that can be used to estimate the condition of the driver, the parameters related to the movement of the eyes and eyelids are the most interesting to study [11]. This approach has been addressed and implemented here.

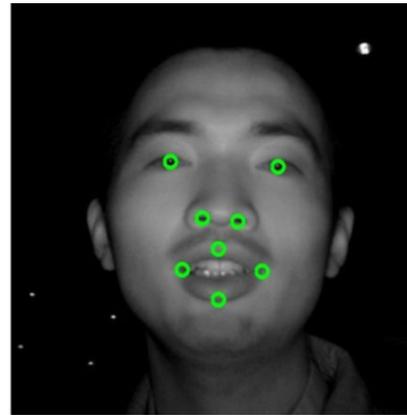


FIGURE 7. Points of interest in the face to determine the degree of fatigue.

III. EXPERIMENTS AND RESULTS

To compare the performance of the methods employed, those are characterized by means of the most relevant parameters for the proposed system. These parameters are the recognition rate, (defined as the percentage of times that the eye position is detected correctly) and the runtime (or the average time spent on performing detection).

During the experiments all methods shown in this work were used, but in the following tables and figures will be observed the best results. These results were reached with a segmentation based on brightness detection. For the feature extraction, all methods have been included.

The computational times of the detection method used are specified in the results. This runtime also includes the subsequent segmentation process.

During this experimentation phase, simulations were performed using both the images of the TUNIR and CBSR NIR databases [12-13]. Many simulations were implemented and results with low accuracies and high computational time have been omitted, to highlight the most interesting results. The results for the two proposed methods are shown in the next section.

In order to validate the robustness of the proposal a cross validation method based on a 50% hold-out strategy was implemented for all experiments. The quality measure is based on identification accuracy. It will show the best proposal in order to be implemented in a future real application.

Tables 1, 2 and 3 show the maximum recognition rate for each experiment. A training percentage of 50% from TUNIR and SBSR NIR has been selected as training and test datasets respectively. The samples chosen for training belong to people different to those used in the training phase in an attempt to select different samples.

Test times, present in those tables, show test time per sample. The training times are related to training models. All times were displayed in milliseconds.

A. VERIFICATION STAGE - TUNIR DATABASE

Table I shows the results obtained with SVM classifier and figure 8 shows the evolution of the accuracy according to the gamma value for the TUNIR database.

For DCT features, Table I shows the accuracy, training time and testing time for the different number of DCT parameters (49-900). Regardless of the number of parameters, results obtained are quite stable. From 64 parameters, that is, a matrix of 8x8 DCT coefficients, accuracy reached the maximum value of 99.66%. Figure 8.a shows a graph with recognition rates obtained with the DCT where similar responses for different experiments can be seen.

When looking at the results obtained with the wavelet module, four cases can be differentiated: Haar, db1, Bior 3.3 and Bior 4.4 presented in section 2.3.2. Figure 8.b presents the accuracy of Haar wavelet for different gamma values and a similar accuracy is obtained for the first three parameters number (97.26%) but four parameters numbers affects the performance (96.23%). For DWT Daubechies (DWT db1) a maximum accuracy (97.26%) for 256 parameters is also reached.

DWT Bior 3.3 for 1156, 400, 169 and 100 parameters are presented in Figure 8.d. Slightly worse results are obtained compared to previous features getting the best result for 361 parameters, which was 96.58% accurate. For DWT Bior 4.4, the accuracy reached 96.92% for 400 parameters.

Regarding PCA analysis, the best result is found with 200 coefficients showing an accuracy rate of 97.95%.

TABLE I
RESULTS FOR SVM CLASSIFIER AND TUNIR DATABASE FOR TRAINING AND TESTING PURPOSE WITH DIFFERENT FEATURE SELECTION TECHNIQUES

Features	Number of features	Accuracy (%)	Training time (ms)	Testing time (ms)
DCT	49	98.97	217.9	0.7
	64	99.66	242.8	0.8
	81	99.66	275.9	0.9
	100	99.66	317.2	1.1
	144	99.66	427.3	1.5
	900	99.66	2461.4	1.9
DWT Haar	256	97.26	636.4	2.1
	64	97.26	252.2	0.8
	16	97.26	177.8	0.6
DWT db1	4	96.23	144.3	0.5
	256	97.26	669.1	2.2
	64	97.26	311.2	1.0
DWT Bior 3.3	16	97.26	191.6	0.6
	4	96.23	178.3	0.6
	361	96.58	1331.4	4.4
	169	95.89	716.9	2.4
DWT Bior 4.4	100	95.21	504.7	1.7
	64	90.41	480.7	1.6
	400	96.92	1184.3	3.9
	196	96.58	641	2.1
PCA	121	95.55	442.9	1.5
	100	94.86	371.4	1.2
	10	94.89	146.3	0.5
	20	96.23	172	0.6
	50	96.92	235.7	0.8
	100	97.6	331.1	1.1
	150	97.95	447.6	1.5
	200	97.95	592.9	2

B. VERIFICATION STAGE - CBSR NIR DATABASE

The main difference between the CBSR NIR and TUNIR databases is that the images in the first are twice the size of the second, and that CBSR NIR also presents lower illumination.

Table II shows the results obtained with the SVM classifier and figure 9 the evolution of the accuracy according to the gamma value for the CBSR NIR database exploring the same previously studied features.

For DCT, for example, there is an accuracy rate of 99.05% (see Table II) for 64 and 81 DCT parameters when choosing an input array of 8x8 or 9x9. Despite this variation in the number of parameters, very similar results due to the features robustness are obtained.

According to DWT Haar, the results obtained with the first three parameters number (956, 256 and 64) are similar reaching a 97.26% recognition rate. For DWT db1, the best result obtained is 99.11% with 64 parameters.

For DWT Bior 3.3 a recognition rate of 99.18% was achieved. The number of input parameters is 400 and the average test time was 4.3 ms for each sample. For DWT Bior 4.4 the first three parameters numbers again provided very similar results and dropped to 98.54% when the number of parameters is reduced. The highest recognition rate obtained was 99.08%. This result was reached with 20 coefficients with a test time of 0.5 ms.

TABLE II
RESULTS FOR SVM CLASSIFIER AND CBSR NIR DATABASE FOR TRAINING AND TESTING PURPOSE WITH DIFFERENT FEATURE SELECTION TECHNIQUES

Features	Number of features	Accuracy (%)	Training time (ms)	Testing time (ms)
DCT	49	98.98	4382	0.7
	64	99.05	4958	0.8
	81	99.05	6028	0.9
	100	98.98	7140	1.1
	144	98.89	7914	1.2
	900	98.95	7805	1.2
DWT Haar	961	97.26	636.4	2.1
	256	97.26	252.2	0.8
	64	97.26	177.8	0.6
DWT db1	16	96.23	144.3	0.5
	961	97.26	31.508	5.2
	256	97.26	11.321	1.8
DWT Bior 3.3	64	97.26	5.070	0.8
	16	96.23	3.072	0.5
	1156	99.11	59.710	9.6
	400	99.18	25.413	4.3
DWT Bior 4.4	169	99.11	14.6625	2.3
	100	98.16	11.215	1.8
	1225	99.05	56.303	9.1
PCA	484	99.08	21.444	3.4
	225	99.08	12.363	1.9
	144	98.54	10.4579	1.6
	10	98.7	2817	0.4
PCA	20	99.05	3398	0.5
	50	98.83	5808	0.9
	100	98.83	7355	1.2
	150	98.79	9330	1.5
	200	98.89	11313	1.8

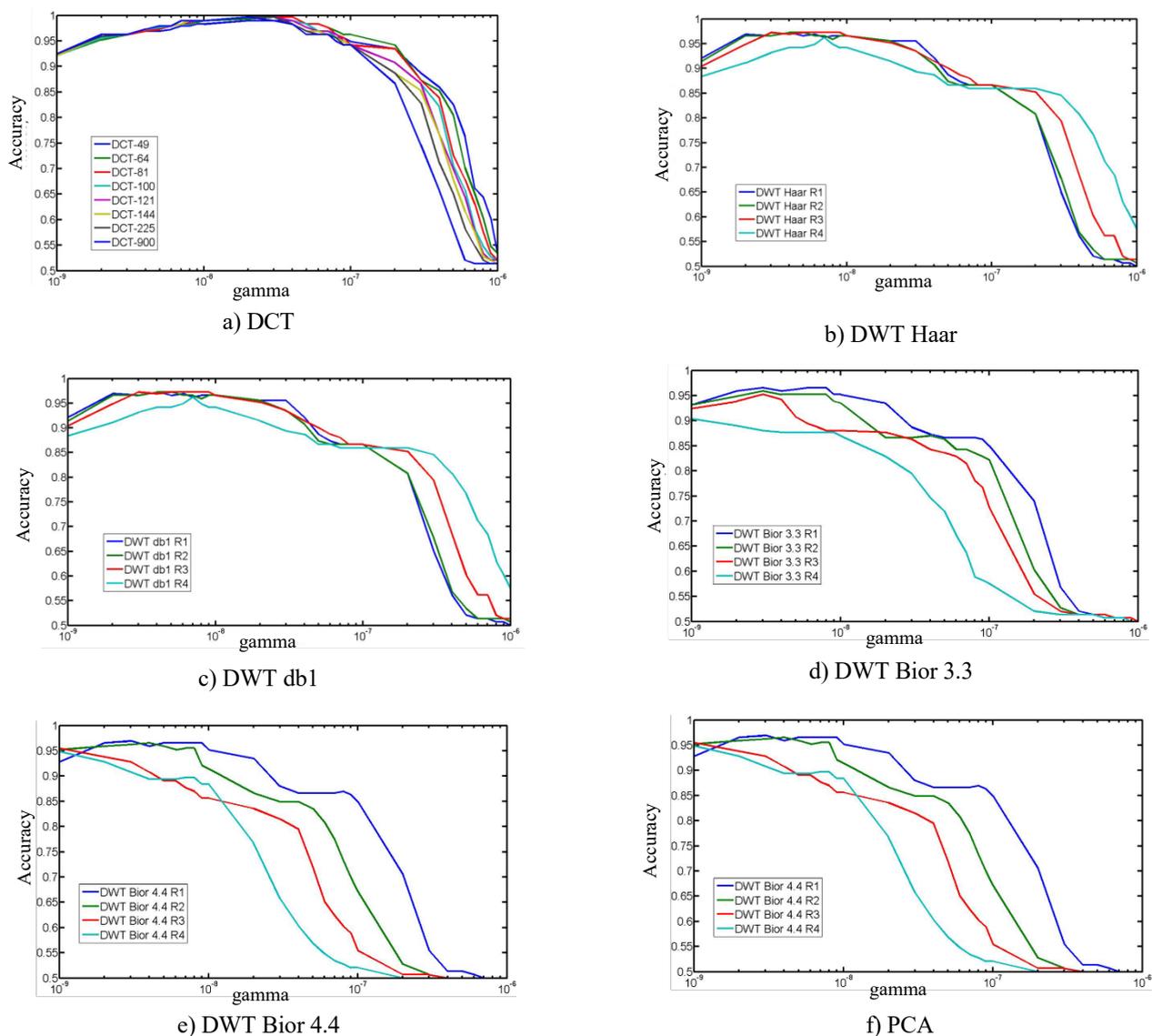


FIGURE 8. Accuracy with SVM for different feature extraction techniques and gamma values with TUNIR dataset.

C. VERIFICATION STAGE - COMBINING TUNIR AND CBSR NIR DATABASES

The main reason to combine classifiers was to check the robustness of the classifiers as the dataset used for training is different to the one used to validate the solution.

It should be noted that the eye image dataset generated from TUNIR and CBSR NIR present different sizes, 31x31 and 61x61 pixels respectively. Furthermore, the lighting in TUNIR data base is more powerful.

Testing and training images are adjusted in size and the histograms of the test images are equalized compared to an average histogram used in training. The results classify the image database CBSR NIR models using SVM generated by the TUNIR training database.

Recognition rates found with DCT are quite good. Despite using different images to train and test, the recognition rate is quite high, with a maximum of 98.73%. On the other hand, using PCA a comparable result with a maximum accuracy of

97.92% is obtained. These results can be observed in Table III and Figure 10.

D. TRACKING STAGE

In the tracking stage, the working of the proposed methods was tested. The results with block matching methods and optical flow were not particularly remarkable. The estimation error of the eye movement increased frame to frame, due to accumulate error of using those estimations as samples. Instead, employing the extended Kalman filter (EKF), the estimated position for each frame is corrected, and the best results with this method are obtained.

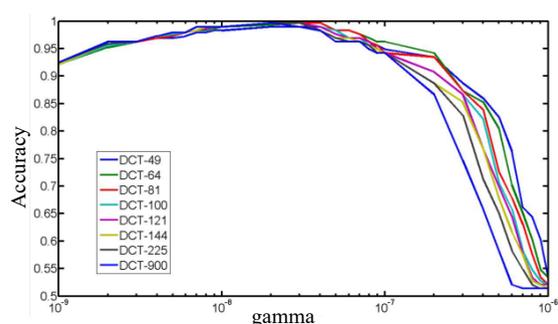
Figure 11 shows an example of using the EKF method to keep the eyes in a sequence of the TUNIR dataset, for 45 frames, and taking approximately four frames per second. As can be observed, the position estimated by the Kalman filter and calculated by the methods of automatic eye detection is similar to the actual position of the eyes in the estimated samples.

The averaged time to track between two frames using the extended *Kalman* filter and using the brightness detection method to calculate the position of the eyes is 16.8 ms. Using the Circle Hough Transform method, the average execution time is 38.9 ms.

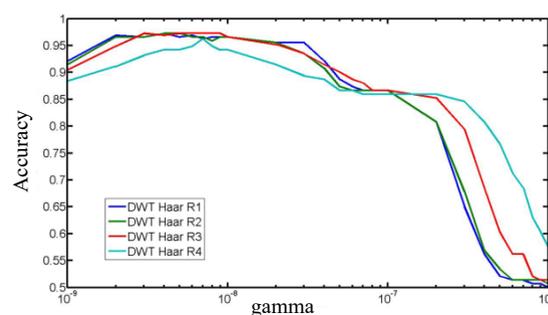
Execution times using the Kalman filter have been significantly reduced vs. using eye location methods for each frame and in all areas of the image. In Table IV, the averaged times are shown; in particular, the Hough transform is 128.1 ms and the brightness detection is 254.8 ms. The improvements of the computational time are 69.6% and 93.4%, respectively.

TABLE III
RESULTS FOR SVM CLASSIFIER WITH TUNIR DATABASE AS TRAINING AND CBSNR NIT AS TESTING DATABASE WITH DCT AND PCA FEATURES.

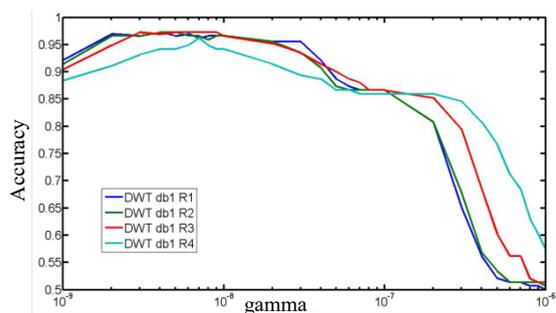
Features	Number of features	Accuracy (%)	Training time (ms)	Testing time (ms)
DCT	49	98.79	217.9	0.7
	64	98.78	242.8	0.8
	81	98.83	275.9	0.9
	100	98.79	317.2	1.1
	121	97.54	383.4	1.3
	144	97.46	427.3	1.5
	225	98.86	583.2	2.6
PCA	900	97.40	3461.4	11.9
	10	92.80	2438	0.3
	20	96.21	3173	0.3
	50	97.45	5461	0.5
	100	97.8	6929	0.9
	150	97.87	8775	1.1
	200	97.92	10157	1.8



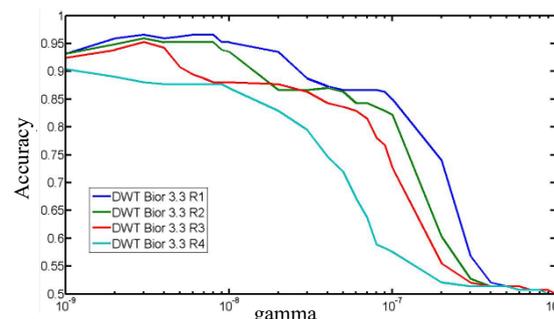
a) DCT



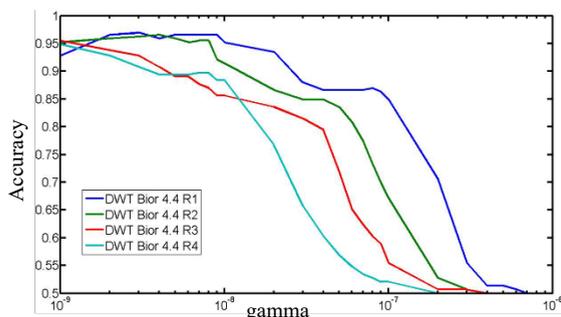
b) DWT Haar



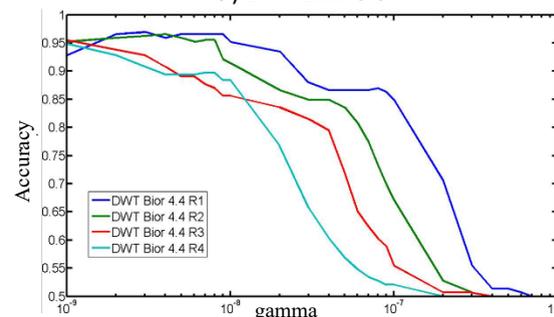
c) DWT db1



d) DWT Bior 3.3



e) DWTBior 4.4



f) PCA

FIGURE 9. Accuracies for different feature extraction techniques and CBSR NIR dataset.

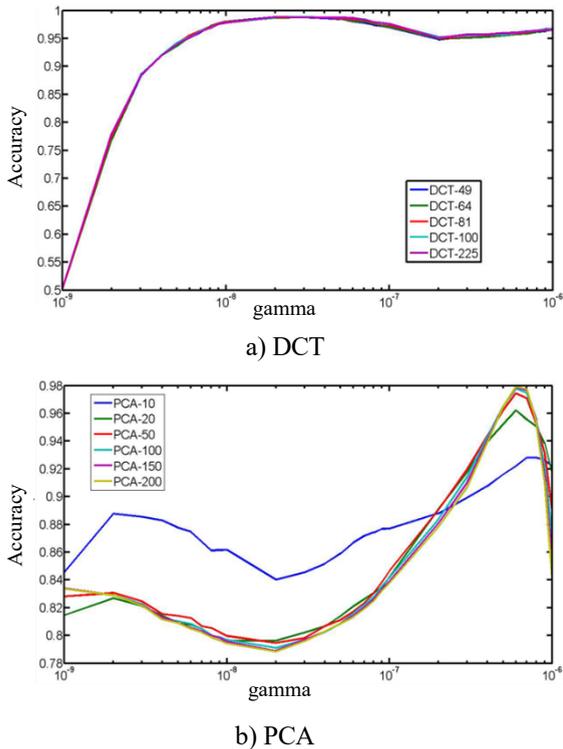


FIGURE 10. Accuracies for PCA and DCT combining TUNIR and CBSR NIR datasets.

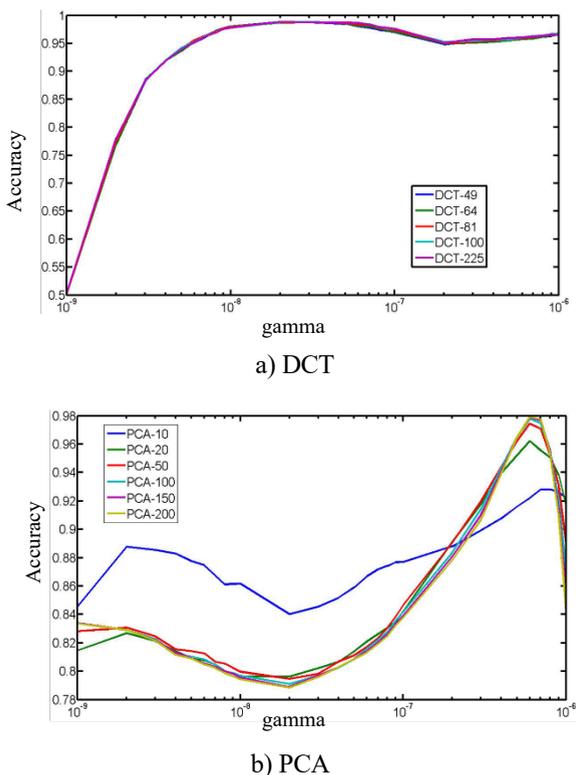


FIGURE 11. Estimated position by EKF vs. calculated position of the eyes.

IV. DISCUSSION

This section presents the system with the best results for each database. In figure 12 the proposal for the TUNIR dataset is showed.



FIGURE 12. Summary of the benefits of the proposed system with the TUNIR database.

As can be seen, the best performing system for TUNIR database consists of the following modules:

- Localization stage: Brightness detection method. The eye detection rate with this method is 94%.
- Classification stage: DCT-64 + SVM. The highest recognition rate was obtained by combining feature extraction using the DCT with 64 parameters and SVM. Recognition rate achieved was 99.66% and test time per sample was 0.8 ms.
- Tracking stage: Extended Kalman Filter.

In Figure 13, the shown system works with the CBSR NIR database;

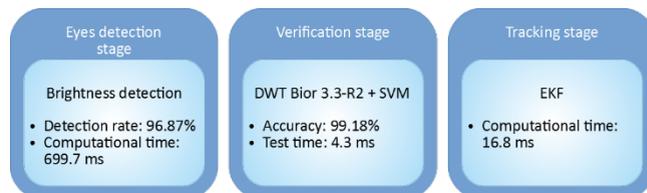


FIGURE 13. Summary of benefits of the proposed database CBSR NIR system.

The best performance is achieved using:

- Location stage: Brightness detection method. The detection rate when applying the brightness detection method in the CBSR NIR database images reached 96.87% and a running time of 699.7 ms.
- Classification stage: DWT Bior 3.3 R2 + SVM. Best performance is obtained when using biorthogonal wavelet transform, namely 3.3 Bior wavelet and SVM-based classifier. Recognition rate of 99.18% was achieved with a time of test per sample of 4.3 ms.
- Tracking stage: Extended Kalman Filter with a computational time of 16.8 ms.

From the results, it is concluded that there is a significant difference between the eye detection rates of the two databases. The best results when applying the detection methods proposed in the database are found for CBSR NIR. This may be due to two factors determined by the recording system: lighting and distance of individuals with the camera.

The images of the TUNIR database are brighter than those in the CBSR NIR database. Furthermore, the position of the people recorded varies and, at times, are very close to the camera. The resulting images have excess light which makes eye detection more difficult. Figure 14 shows the difference between these images.

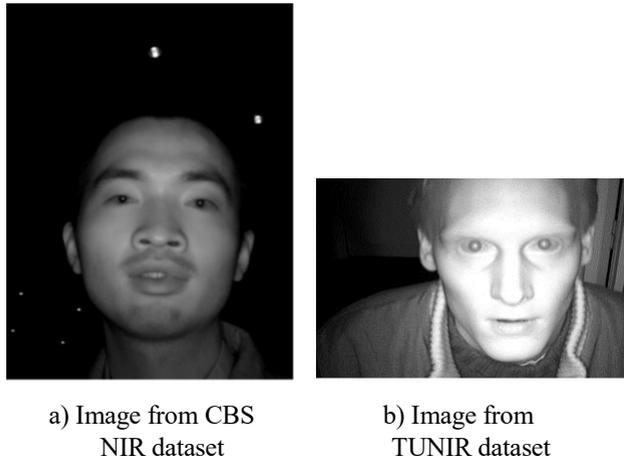


FIGURE 14. Difference of illumination and resolution between the two used datasets.

Another difference between the two databases is resolution. The resolution of the CBSR NIR database is greater than the TUNIR database. This size difference allowed us to test the algorithms with different resolution images and is key to understanding the results of the classifiers when the databases are crossed.

When SVM models were crossed, better results were found when samples from CBSR NIR database were tested with a model SVM, and trained with TUNIR database. Sample eye images of the CBSR NIR database have twice the resolution of the TUNIR dataset. Adapting samples to the SVM model works better, when the sample image has a higher resolution and its size is reduced. But the result is not the same when the size of the sample image is increased from a lower resolution.

The state-of-the-art on eye detection and location is large and diverse, and many references employ different methods, but they are all different to this work, although some works have similar blocks. From all these references, the following have been highlighted as they have been consulted in this work. The results are shown in Table V, indicating the percentages obtained in the detection and verification phases, although in some of them it has not been possible to obtain information of the two stages.

Observing the previous tables, it is deduced that these proposals are very different to each other. It should be noted that the Yingyu Ji *et al.* [16] eye detection system has a 98.42% recognition rate. The obtained results confirm this proposal has a good behavior using brightness detection combined with a classifier based on SVM, which is superior to the remaining articles analyzed.

TABLE V
RESULTS FROM OTHER REFERENCES FOR EYE DETECTION STAGE

Eye detection	Recognition Rate
[16]: Eye recognition and yawning analysis (CNN)	98.42%
[17]: Pupil center detection (Adaboost & Adaptive Thresholding)	97.2%
This proposal - CBSR NIR: Eye detection (DWT)	96.87%
[18]: Face detection and eyes detection (CNN)	95.00%
This proposal – TUNIR: Eye detection (DCT)	94.00%

TABLE VI
RESULTS FROM OTHER REFERENCES FOR VERIFICATION STAGE

Classification system	Recognition Rate
This proposal – TUNIR: Support Vector Machine with RBF kernel	99.66%
This proposal – CBSR NIR: Support Vector Machine RBF kernel	99.18%
PERCLOS + Convolutional Neural Networks [18]	98,81%
Convolutional Neural Networks (CNN) [19]	98.06%
SVM and Adaboost [20]	96.5%-95.4%
Support Vector Machine with RBF kernel [21]	93.50%
Convolutional Neural Networks (RNN) [22]	92.00%
Support Vector Machine with RBF kernel [17]	86.58%
Adaboost [23]	80.19%
Haar [24]	80.00%

Table VI shows results of different research works for the fatigue detection, and where the current proposal reaches a very good accuracy. This table shows a comparison vs. different approaches, including methods based on deep learning. They are applied in different datasets for visible and infrared data, and the proposal described in this work yields the best accuracy for two different public datasets. For both datasets, the final proposal is similar which authors propose to use a brightness detection for the eye detection stage, and a Extended Kalman Filter for the tracking stage. Moreover, for the verification stage, a different approach, DCT or Wavelet, for each dataset were applied to obtain the input features for a SVM classifier. DCT have a bit better results, the difference of accuracy is 0,48%, and better than other recent approaches. Therefore, this combination gives the best accuracy for a possible prototype.

The key of this proposal is its robustness, which reaches the best accuracy during the verification process of the eye activity.

V. CONCLUSIONS

Various methods have been studied for automatic eye detection by testing the performance of different features and SVM classifier. Furthermore, the eye tracking optimization have been studied.

Brightness detection proved to be the best method to locate eye position in the images, with high detection rates and with an acceptable computational cost. For simulations with the TUNIR database, an accuracy of 94% and total running time of 254.8 ms was reached. With the CBSR NIR database results were better, with an accuracy of 96.87% and 699.7 ms runtime.

In the verification phase, high recognition rates were also yielded. For both databases, the use of DCT and SVM showed a good and robust result with accuracy over 99%. Specifically, a 99.66% accuracy rate is obtained by the DCT and TUNIR dataset and 99.05% for CBSR NIR. With this last dataset, the best result is found with the 3.3 Bior DWT with a 99.18%. This approach has been compared with different research works, some of them applying deep learning methods, yielding the current proposal better accuracies than the state-of-the-art, as it is shown in Table VI.

In the monitoring stage, best results were obtained using the Extended Kalman filter updating the position of the eyes in each frame for long sequences of frames. It was possible to reduce the execution time 93.4% compared to times obtained in the detection phase. The average time to calculate the current position of the eyes was 16.8 ms, which would update the eye position over 50 times per second.

The innovation of this proposal is a robust combination of methods for the detection of fatigue parameters with infrared information, as it were shown after experiments, and with very high accuracies for those fatigue parameters. Another important aspect of this work is the big variability of samples, users with glasses, different head positions and different position of the camera, as can be observed in Tables 2 and 4.

In future works, the approach will be implemented on real conditions in order to show the robustness of this work.

ACKNOWLEDGMENT

This work has been supported by the Agreement between Universidad Antonio Nebrija and University of Las Palmas de Gran Canaria.

REFERENCES

- [1] Spanish Traffic Agency, "Elaboración de un estudio sobre los dispositivos que existan en la actualidad y se estén utilizando para la monitorización del comportamiento de los conductores durante el ejercicio de la conducción, las variables registradas, exactitud y fiabilidad de las mismas," Technical Report., General Directorate of Traffic, Ministry of Interior, Government of Spain. 2014
- [2] J. M. Owens, T. A. Dingus, F. Guo, Y. Fang, M. Perez, J. McClafferty, B. Tefft, "Prevalence of Drowsy Driving Crashes: Estimates from a

- Large-Scale Naturalistic Driving Study", AAA Foundation for Traffic Safety, 2018
- [3] J. F. May and C. L. Baldwin, "Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies" *Transp. Res. F, Psychol. Behav.*, vol. 12, no. 3, pp. 218–224, 2009. DOI: 10.1016/j.trf.2008.11.005
- [4] National Center for Statistics and Analysis, "Drowsy Driving 2015 (Crash Stats Brief Statistical Summary. Report No. DOT HS 812 446)", National Highway Traffic Safety Administration, 2017.
- [5] I. D. Brown, "Prospects for technological countermeasures against driver fatigue", *Accident Analysis and Prevention*, vol. 29, no. 4, pp. 525 – 531, 1997. DOI: 10.1016/s0001-4575(97)00032-8
- [6] R. Carroll, "Ocular measures of driver alertness: Technical conference proceedings", Federal Highway Administration, Office of Motor Carrier and Highway Safety, Washington, DC, FHWA Tech. Rep, 1999.
- [7] L. Staplin, K. Gish, L. Decina, and R. Brewster, *Ocular measures of driver alertness: technical conference proceedings: April 26-27, 1999, Hyatt-Dulles Hotel, Herndon, Virginia.* U.S. Dept. of transportation, Federal Highway Administration, 1999.
- [8] A. Koesdwiady, R. Soua, F. Karray and M. S. Kamel, "Recent Trends in Driver Safety Monitoring Systems: State of the Art and Challenges," in *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 4550-4563, 2017. DOI: 10.1109/TVT.2016.2631604
- [9] B. Mandal, L. Li, G. S. Wang, ET AL.: "Towards detection of bus driver fatigue based on robust visual analysis of eye state", *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no 3, pp. 545–557, 2017. DOI: 10.1109/TITS.2016.2582900
- [10] J. Xu, J. Min and J. Hu, "Real-time eye tracking for the assessment of driver fatigue", *Healthcare technology letters*, vol. 5, no 2, pp. 54–58, 2018. DOI: 10.1049/htl.2017.0020
- [11] G. Sikander and S. Anwar, "Driver Fatigue Detection Systems: A Review," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2339-2352, June 2019. DOI: 10.1109/TITS.2018.2868499
- [12] S. Zhao, R. Kricke, and R.-R. Grigat, "TUNIR: A multi-modal database for person authentication under near infrared illumination," in *6th WSEAS International Conference on Signal Processing, Robotics and Automation (ISPRRA 2007)*, vol. 3, (Corfu, Greece), pp. 173–178, 2007.
- [13] S. Z. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images," *IEEE Transactions on Pattern Analysis and Machine Intelligence (Special issue on Biometrics: Progress and Directions)*, vol. 29, no. 4, pp. 627–639, 2007. DOI: 10.1109/TPAMI.2007.1014
- [14] R. C. Gonzalez and R. E. Woods, "Digital Image Processing", Upper Saddle River, N.J.: Prentice Hall, 2008.
- [15] S. Challa, M. R. Morelande, D. Mušicki, and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge: Cambridge University Press, 2011.
- [16] Y. Ji, S. Wang, Y. Zhao, J. Wei and Y. Lu, "Fatigue State Detection Based on Multi-Index Fusion and State Recognition Network," in *IEEE Access*, vol. 7, pp. 64136-64147, 2019. DOI: 10.1109/ACCESS.2019.2917382
- [17] R. O. Mbouna, S. G. Kong, M. G. Chun: "Visual analysis of eye state and head pose for driver alertness monitoring", in *IEEE Trans. Intell. Transp. Syst.*, 14, (3), pp. 1462–1469, 2013. DOI: 10.1109/TITS.2013.2262098
- [18] B. K. Savaş and Y. Becerikli, "Real Time Driver Fatigue Detection System Based on Multi-Task ConNN," in *IEEE Access*, vol. 8, pp. 12491-12498, 2020.
- [19] W. H. Gu, Y. Zhu, X. D. Chen, L. F. He and B. B. Zheng, "Hierarchical CNN-based real-time fatigue detection system by visual-based technologies using MSP model," in *IET Image Processing*, vol. 12, no. 12, pp. 2319-2329, 12 2018. DOI: 10.1049/iet-ipr.2018.5245
- [20] B. Fatima, A.R. Shahid, S. Ziauddin, A.A. Safi and H. Ramzan. "Driver Fatigue Detection Using Viola Jones and Principal Component Analysis," in *Applied Artificial Intelligence*, Taylor & Francis, vol. 332, no 6, pp. 456-483, 2020.
- [21] A. Punitha, M. K. Geetha, A. Sivaprakash: Driver fatigue monitoring system based on eye state analysis, *Int. Conf. Circuit, Nagercoil, India*, pp. 1405–1408, March 2014.

- [22] Y. Ed-Doughmi, N. Idrissi and Y. Hbali. "Real-Time System for Driver Fatigue Detection Based on a Recurrent Neuronal Network", in *Journal Imaging* vol. 6, no 3, 2020.
- [23] Fan, Y. Sun, B. Yin: "Driver fatigue detection based on AdaBoost LBP", *Journal of Information and Computational Science.*, vol. 5, no 1, pp. 61–68, 2009.
- [24] V. B. Hemadri, U. P. Kulkarni: "Detection of drowsiness using fusion of yawning and eyelid movements", In: Unnikrishnan S., Surve S., Bhoir D. (eds) *Advances in Computing, Communication, and Control*. ICAC3 2013. Communications in Computer and Information Science, vol 361. pp. 583–594, Springer, Berlin, Heidelberg, 2013



CARLOS M. TRAVIESO-GONZALEZ (S'94–M'97–F'02) received the M.Sc. degree in 1997 in Telecommunication Engineering at Polytechnic University of Catalonia (UPC), Spain; and Ph.D. degree in 2002 at University of Las Palmas de Gran Canaria (ULPGC-Spain). He is Full Professor on Signal Processing and Pattern Recognition and Head of Signals and Communications Department at ULPGC, teaching from 2001 on subjects on signal processing and learning theory. He is co-author of 4 books, co-editor of 24 Proceedings

Book, Guest Editor for 8 JCR-ISI international journals and up to 24 book chapters. He has over 450 papers published in international journals and conferences (80 of them indexed on JCR – ISI - Web of Science). He has published 7 patents and has been supervisor on 9 PhD Theses and 130 Master Thesis. His research lines are biometrics, biomedical signals and images, data mining, classification system, signal and image processing, machine learning, and environmental intelligence. He has researched in more than 51 International and Spanish Research Projects, some of them as head researcher. He is founder of The IEEE IWOB conference series and President of its Steering Committee, of The InnoEducaTIC conference series; and of The APPIS conference series. He is evaluator of project proposals for European Union (H2020), for Medical Research Council (MRC – UK), Spanish Government (ANECA), for ANR (France), for DAAD (Germany), for Argentinian Government and for Colombian Institutions. He has been reviewer in different indexed international journals (<75) and conferences (<240) since 2001. He is Associate Editor for Computational Intelligence and Neuroscience journal (Hindawi – Q2 JCR-ISI), and for Entropy journal (MDPI – Q2 JCR-ISI). He has been General Chair in 15 Internacional Conferences. He got The awards in Catedra Telefonica 2017, Catedra Telefonica 2018 and Catedra Telefonica 2019 on Knowledge Transfer Modality, and Catedra Telefonica 2020 on Research on COVID-19.



JESÚS B. ALONSO-HERNÁNDEZ received the Telecommunication Engineer degree in 2001 and the Ph.D. degree in 2006 from University of Las Palmas de Gran Canaria (ULPGC-Spain) where he is an Associate Professor in the Department of Signal and Communications from 2002. His research interests include signal processing in biocomputing, biometrics, nonlinear signal processing, recognition systems, audio characterization and data mining. He has been guest editor of special issues for Journal as

Cognitive Computation, Neural Computing and Applications, or Neurocomputing.



JOSÉ MIGUEL CANINO-RODRÍGUEZ. He received the Telecommunication Engineering Diploma and the Science Teacher Diploma at the University of Las Palmas de Gran Canaria, Spain (ULPGC,1984), the Physical Science degree (National University of Distance Education – UNED-, Spain, 1988) and the PhD degree (ULPGC, He is currently a Professor at ULPGC since 1997, where he has taught several subjects related to electromagnetic propagation, navigation systems and intelligent systems. Currently, His research work has been focused on Multi-Agent Systems and Matching Learning Applications for intelligent navigation environments.



SANTIAGO T. PÉREZ was born in Las Palmas de Gran Canaria, Spain in 1965. He received the B.S. and M.S. degrees in telecommunication engineering from the University of Las Palmas de Gran Canaria, in 1991 and 2002; and the Ph.D. degree in cybernetics and telecommunication from the same university in 2015.

From 1989 to 1991, he worked as technical engineer for Telettra España S. A. From 1992 to 1999, he has been an assistant professor; from 1999 he has been an associate professor for Signals and Communications Department in the mentioned university. He is the author of two chapters of books, more than 30 conferences papers and 8 international journals. His research interests include Spread Spectrum Communications, Radio Frequency Identification and neural network implementations on Field Programmable Gate Arrays.



DAVID DE LA CRUZ SÁNCHEZ-RODRÍGUEZ received the M.Sc. degree in 1999 in Telecommunication Engineering at University of Las Palmas de Gran Canaria (ULPGC), Spain; and Ph.D. degree in 2006 at University of Las Palmas de Gran Canaria. He is an Associate Professor from 2000 in ULPGC, teaching subjects on computer networking, wireless communications and computer programming. His research lines are wireless network protocols, wireless devices indoor localization and tracking,

data mining and machine learning, wireless sensor networks and Internet of Things technologies.

ANTONIO G. RAVELO-GARCÍA (S'96–M'06–F'12) was born in Las Palmas de Gran Canaria, Spain, in 1973. He received the M.S. and Ph.D. degrees in telecommunication engineering from the Universidad de Las Palmas de Gran Canaria, Spain. He is currently a Docent and a Researcher with the Department of Signal and Communications and the Institute for Technological Development and Innovation in Communications, Universidad de Las Palmas de Gran Canaria. He has participated in multiple research projects and publications. His current research interests include biomedical signal processing, nonlinear signal analysis, data mining, and intelligent systems in multimedia, communications, environment, health, and wellbeing.