THE ANNALS OF "DUNĂREA DE JOS" UNIVERSITY OF GALATI
FASCICLE III, 2017, VOL. 40, NO. 1, ISSN 2344-4738, ISSN-L 1221-454X
ELECTROTECHNICS, ELECTRONICS, AUTOMATIC CONTROL, INFORMATICS

MACHINE LEARNING TECHNIQUES FOR IMAGE RECOGNITION APPLICATIONS

George Ciubucciu*, Razvan Solea*, Adrian Filipescu*

**Department of Automation and Electrical Engineering,*
"Dunarea de Jos" University Galati, Romania
(email:ciubucciu_george@yahoo.com, Razvan.Solea@ugal.ro, Adrian.Filipecu@ugal.ro)

Abstract: In the present paper is presented an algorithm to detect sleepiness of the driver. The purpose of this system is to locate, track and analyze face and eyes to prevent falling asleep, or the inattention of the driver working under different lighting conditions and, more importantly, in real time. Real-time processing is an absolute condition for the proper functioning of this system. The image acquisition is achieved with the help of a web camera, which is positioned on the dashboard of the car. For the face recognition task was used the support vector machine models. This is an artificial intelligence technique applied in the field of artificial vision, which is part of the supervised learning techniques.

Keywords: Image processing, Artificial Intelligence, Support vector machine, Real-time processing, Algorithm.

1. INTRODUCTION

The face plays a key role in social relationships, communication and emotional identity. The human ability to recognize faces is remarkable: we can recognize thousands of faces learned throughout life and identify familiar faces from one look even after years. This capacity is quite robust, despite great changes in visual stimulus due to visualization, expression, aging and other disturbing factors such as the presence of glasses or hair styling changes.

Face recognition is a difficult task because of the numerous sources of variability in real conditions. These include among others: orientation of the face relative to the camera, the lighting, facial expressions, the time between sampling times of distinct images and demographics (race, age, and gender). There is a need to identify the human face images in many applications such as recognizing people, counting people, identifying the head position for driving assistance systems, etc.

For the face recognition task was used the SVM models. This is an artificial intelligence technique applied in the field of artificial vision, which is part of the supervised learning techniques.

Road accidents caused by human errors cause many deaths and injuries worldwide. To help reduce this phenomenon, a new Advanced System Driver Assistance (ADAS) has been developed to automatically detect driver drowsiness based on visual information and other sensors. Data taken from either the camera and / or other sensors is processed by a computer specialist on board the car using artificial sight algorithms or even artificial intelligence.

The purpose of this system is to locate, track and analyze face and eyes to prevent falling asleep, or the inattention of the driver working under different

lighting conditions and, more importantly, in real time. Real-time processing is an absolute condition for the proper functioning of this system.
ADAS is part of active safety systems that interact with drivers to help them avoid traffic accidents, and its purpose is to help reduce road accidents by using new technologies.
In this case, vehicle security research is focused on the driver's analysis, in particular case - sleepiness and distraction are studied extensively (Hong, Chung, K, 2005).
It has been estimated that sleepiness causes between 10% and 20% of the traffic accidents resulting in deaths and wounded [11], while in the case of trucks statistics show that 57% of fatal accidents is due to fatigue [2,22]. Author Pai-Yuan Tsai, *et al.* 2009 say that 30% of all road accidents have been caused by drowsiness or inattention. The Royal Society for the Prevention of Accidents (see https://www.rospa.com/road-safety/advice/drivers/fatigue/road-accidents/) presents statistics in which 20% of all accidents are caused by fatigue and lack of concentration. In the United States sleepiness is responsible for 100,000 road accidents and the related costs are about 12,000 million dollars. In Germany, one in four road accidents has its origin in drowsiness, in England 20% of all road accidents are caused by drowsiness. Every year in the European Union more than 40,000 deaths and 1.4 million injured because of car accidents are affirmed (International Road Traffic Database).
It is important to use new technologies and build systems to monitor drivers, and prevent them falling asleep or inattention. Drivers in a state of drowsiness have several visual features that can be detected on the human face. These are:
• blink frequency, moving and closing eyes, head inclination, facial expressions.
Considering these, artificial vision is the right technology to handle this problem. This paper presents a warning system for driver's sleepiness or inattention. The purpose of this system is to automatically intervene with a beep when the driver falls asleep or when he is not paying attention to traffic.
It has to be underlined that in the case of studies (Bekiaris, *et al* 1997) it was stated that those who experienced an accident or those who were very close to causing an accident were interested in such a system.
In this regard, the use of artificial vision technologies to identify the state of somnolence or inattention of the driver and understanding his / her behavior is an active field of research.

Fig.1. The concept of Toyota's driver supervision system(see http://cherup.yonsei.ac.kr/research/research_driver_monitoring(090602).html)

To analyze driver drowsiness in recent years have been built several systems. D'Orazio *et al.* 2004 proposed an eye detection algorithm that searches the eyes across the image, assuming that the iris is always darker than the sclera, and based on a circle identification algorithm such as Hought transformation and other scale constraints. The driver's eyes are located. Next, the result is processed by a neural network that is capable of classifying eye-images and images that are not human eyes. This system is able to classify your eyes as open or closed.
Smith *et al.* 2003 presents a driver's visual analysis system based on the global motion estimation. The algorithm considers motion and color statistics to track the person's head and classifies rotation in all directions of the head.
Er. Manoram Vats1 and Er. Anil Garg2, 2012 built a system using fatigue detection neural networks that optimize detection using the ANN technique.
Horng *et al.* 2004 presented a system using a HIS skin color model for face detection, eye-tagging information, and dynamic eye-matching template. Using eye color information, the eye state is identified and the condition of the driver is determined, for example, if he or she is asleep or alert.
Numerous authors have built artificial vision algorithms. Cristiani A *et al.* 2010; Garg R. *et al.* 2009; Pai-Yuan Tsai *et al.* 2010 proposes a portable device equipped with an electrode.
Dong, W. and Wu, X. 2005 presented a driver fatigue detection system that uses a skin color model based on the bivariate normal distribution and the Cb and Cr components of the YCbCr color space. After locating the eyes, calculate the fatigue index using the eyelid distance to classify the eyes open and the eyes closed, if the eyes are locked over five consecutive frames, the driver is considered to be moaning. This paper does not aim to identify behavior prior to sleepiness.
Ji *et al.* (2004) presented a system of sleepiness detection based on the infrared illumination to stereo (two cameras). This system locates the position of the eye using image differences between the two cameras. The system calculates the frequency eyelid

and look to build two indices of sleepiness: PERCLOS (percent while closing eyes) and AECS (average velocity of eye closure).

The models SVM (support vector machine) were invented by the russian scientist Vladimir Vapnik. SVM was initially used for classification issues, such as OCR (optical character recognition) and object recognition applications. Also, remarkable results have been obtained for data mining problems.
Learning algorithms based on support vector have the main aim of determining a function $\hat{f}$, which approximates a continuous function f by means of vectors of the form <xi, yi> with i = 1..n where the vectors $\mathbf{x}_i \in \Re^m$ represent the input data of the model and the values $y_i \in \Re$ represent the values of the function f for these input data.
In the case of object recognition, the SVM technique is limited to finding the objects for which it was trained and eliminating those that are not searched. The use of SVM is done in two stages (Fig.1): training and querying the algorithm. In the training phase, using databases representing both so-called positive images - that is, those objects that are being searched, as well as negative images - objects that are not searched, the search algorithm is trained. In the second stage, the algorithm is interrogated and the algorithm will find in the current image if the object is searched. Note that both the positive and negative image images extract features to be used later. I point out that the current search image should not be previously in the database.
Figure 2 shows the basic principle of SVM techniques for recognizing images. We identify faces in videos. This SVM technique gives us the ability to separate the set of objects that may exist in a frame in the video stream. The SVM technique involves learning both positive and negative examples. For example, for Face Detection, we need to train hundreds of pictures representing faces and hundreds of images that are different than faces, such as furniture, cars, diverse landscapes, homes, and so on.
The learning process of SVM is limited to finding the parameters of a hyperplan that:
- Separate perfectly the positive examples and the negative examples
- maximize the edge and maximize the distance of the closest
- positive example and the distance of the closest negative example to the hyperplane found.

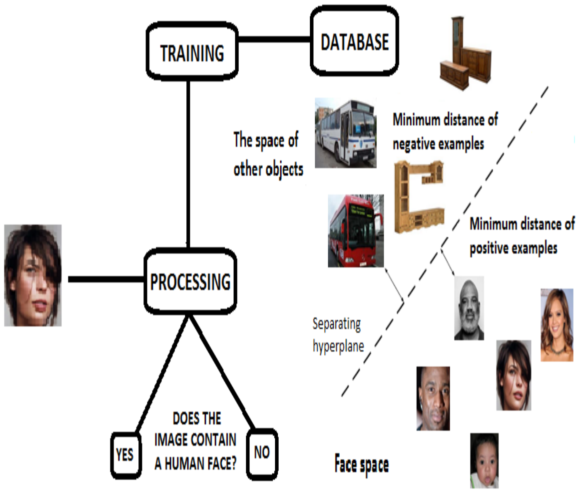

Fig.2. The principle of training for SVM models and the use of SVM models for face recognition

2. DEVELOPED ALGORIHM FOR DROWSINESS AND INATTENTION

2.1. Scenarios to be identified

Next, there is presented the developed driver warning system for drowsiness and inattention. The developed algorithm can be ported to any PC with the image processing function.

We considered several possible scenarios in which the driver may end up as a result of drowsiness or lack of attention from driving as follows (Fig. 3):

- The driver has eyes on the left or right of the car, a situation frequently encountered, in which case the driver's eyes are directed at objects on the street.
- The driver is headed for conversation with passengers on the back seat, a situation that is extremely dangerous, with a higher degree than other situations.
- The driver has a direct look at the car audio system controls, a common occurrence.
- The driver has turned his head towards the passenger on the right, a situation very common.
- The driver falls asleep at the wheel, a particularly dangerous situation that usually leads to fatal accidents.
- The driver is looking over the rearview mirror for a long time.

Fig.3. Possible scenarios

2.2. Hardware architecture used on the system

The experimental hardware system (Fig.4.) consists of: Genius Hd Camera, Atom 1.6 Ghz Laptop Processor, 2 Gb RAM Memory.
If it detects that the driver falls asleep or is distracted in another direction than the leading triggers a visual and audible alarm that warns the driver.

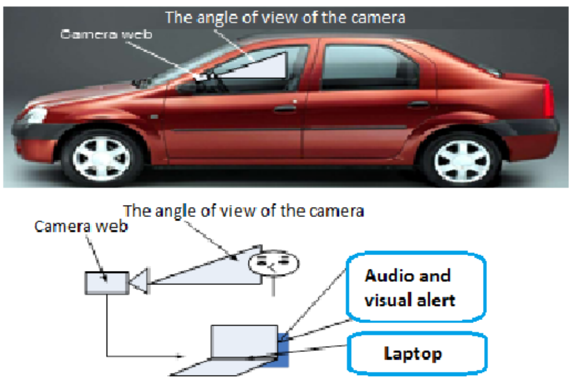

Fig.4. The proposed hardware system. (see https://commons.wikimedia.org/wiki/File:Dacia_Logan_Facelift_front_-_PSM_2009.jpg/)

2.3. Stage of training using SVM

Below we present steps for using SVM techniques using a picture database:

- Preprocessing images in the database - the collected images must be the same size. In our case, I chose 20X20 pixels for eye-catching and 40X40 pixels with face-ups. The training with the SVM algorithm was carried out in the training phase for the creation of the database.
- Separation database training and testing sets - we proceeded to this strategy in that SVM algorithm was not trained with images of drivers that we tested algorithm.
- The choice of input representation - we chose to represent the color histogram information drive image.
- Choosing negative images (images that are not faces or eyes) to drive the image non-sides and non-eye (negative examples). In this sense, we have chosen images that can appear in traffic and can be seen through the window: trees, cars, etc.
- Testing and evaluation of performance - at this stage were used three test drivers. Although this number is not considerably enough, it is sufficient to prove the reliability of the algorithm.

As shown in the figure below (Fig.5.), in the database use phase resulting from the SVM algorithm training, the camera provides the flow of frames to be processed.

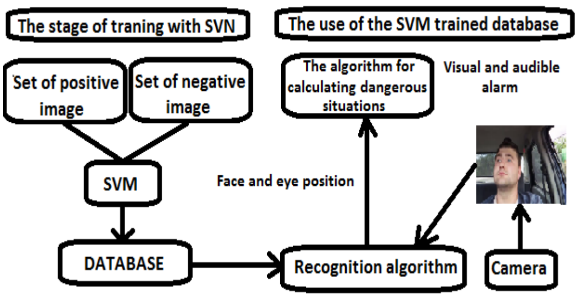

Fig.5. Stages of training and use of the database

2.4. The sleep detection software algorithm

This paper presents a system to detect sleepiness of the driver that works on gray scale images. The scheme of the system is shown in Fig. 5 where the following issues are considered: face detection, eye detection, face tracking, eye tracking, detecting sleepiness, distraction detection. Each of these components will be explained in the following subsections.
For face detection, this system uses the SVM object detector, which is a learning machine for detecting visual objects. It uses three important aspects to make an object-based, effective, full-frame object detector

and cascading classifier. Each of these elements is important for image processing efficiently and in real-time with 90% accurate detection. An important aspect of this method is its robustness in changing light conditions. Continuing with the description of the algorithm, when the driver's face is detected, it is framed in a rectangle. Rectangle size comes from identifying the face database that was created for this task, and is not exact, but sufficient given task.

Locating the eye position is a difficult task because different features define the same eye depending, for example, on the image area where the color of the iris appears, but the main issue during driving is changing ambient light conditions.

Once the face was localized by the RI rectangle in the previous section, using the anthropometric properties of the face [13], which are derived from the face database analysis, two rectangles containing the eyes are obtained. First, this system uses the RI_L system for the left eye rectangle and RI_R for the right eye rectangle.

After the previous step, the exact position of each eye is searched by incorporating pixel gray level information. The basic idea is to get a random sample of pixels that are part of the eye area, and then adjust the parametric model. Figure 6 shows this procedure, in which a random sample is drawn into (a) and an elliptical pattern is adjusted in (b). In this case, the state of the eye is independent, that is, it can be opened or closed. To extract the random sample the following algorithm is proposed.

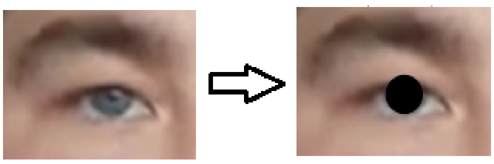

Fig.6. The results of the maximization algorithm on the spatial distribution of eye pixels using iris identification.

Generating the image J by the following equation:

$$(1)\quad J(x,y)=\frac{I(x,y)-m}{\sigma}$$

Where: m and σ are the mean and respective standard deviation. These parameters are calculated over the rectangles of the localized eyes.

$$(2)\quad K(x,y)=\begin{cases} J(x,y)-256*\delta_1 \; if \; J(x,y)\geq 0 \\ 256*\delta_2+J(x,y) \; if \; J(x,y)<0 \end{cases}$$

Starting from the pixels extracted from the B, G and L images, it is possible to obtain the random sample mentioned above. This sample has an elliptical shape and an elliptical pattern has been adjusted over it using the maximization algorithm (EM) [26]. The center of the ellipse has been a great deal of attention because it allows the exact position of the center of the eye. The ellipse axis determines the width and height of the eyes.

The eye detection procedure includes (Fig.7.): after inserting a facial image, pre-processing is the first step performed by image binarization. The top and sides of the face are detected to find the eye area. Using the sides of the face, the center of the face is found, which will be used as a reference when comparing the left eye and the right eye. The bottom-up movement of the face, the horizontal averages (average intensity for each coordinate y) in the face area is calculated. Large changes in environments are used to define the area of the eyes.

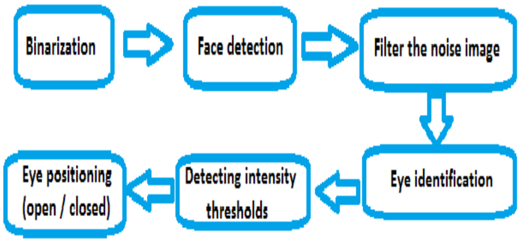

Fig.7. Stages of face and eye identification

Next, face detection is performed using the SVM-driven database. To identify the eyes in the image representing the face, the next step is to filter the noise (Fig.8.). This facilitates the identification of the eyes in the image that represents the face. Eye identification is the stage where the position in the image is found.

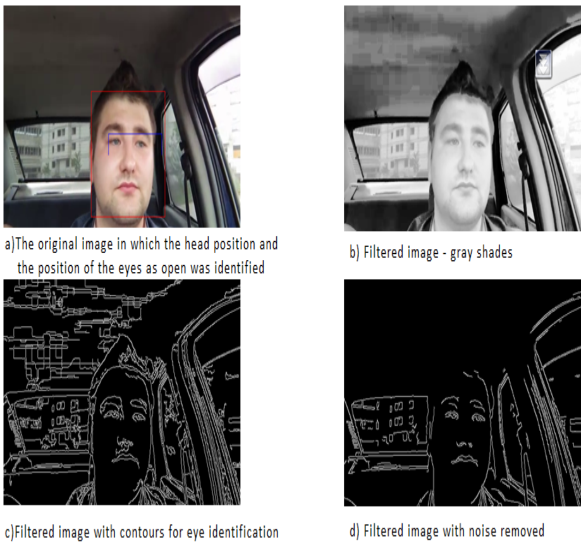

Fig.8. Examples of face identification

The figure 9 shows the algorithm for detecting the dangerous situation. If more than 5 seconds are detected the head is not pointed forward, then the alarm is activated. The algorithm that decides if the head is not straight forward is simplistic. If the symmetry is not detected between the eye position or if the face is not detected then the algorithm decides.

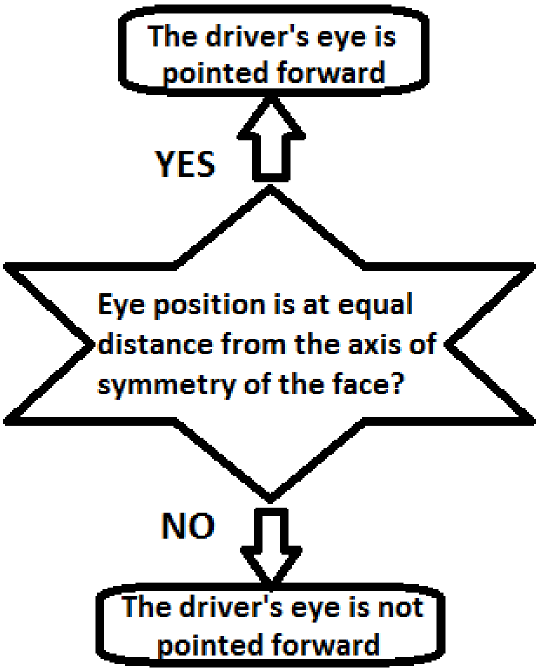

Fig.9. The algorithm that decides whether the driver is looking forward

The next step in the eye detection function is to determine the top and side faces of the driver's face. This is important because finding the contour of the face restricts the area where the eyes are, making it easier to locate the position of the eyes (Fig.10.). The first step is to find the top of the face. The following algorithm describes how to find the real starting point on the face that will be used to find the top of the face.

1. Starting from the top increases the coordinate until a white pixel is found. Consider the left side of the face.
2. If the original white pixel is followed by 30 pixels more white, the x increment is retained until a black pixel is found.
3. Count the black pixels, after which the pixels found in step 2, if a series of 30 black pixels are found on the right side of the face
4. The new coordinate value x (x1) is the middle point of the left and the right.

   To identify sleepiness by eye analysis, it is necessary to know their condition: open or closed, over time, and develop a time analysis, that is, measure the time for each condition.

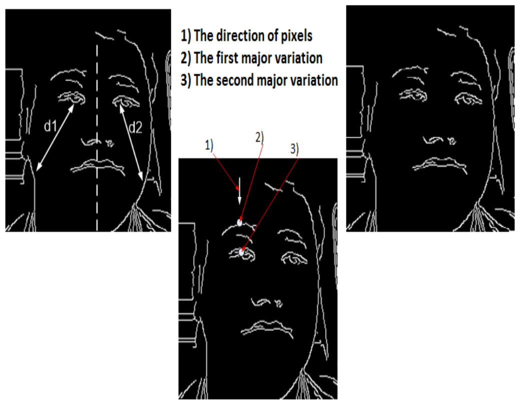

Fig.10. The steps for finding the variations corresponding to the eyebrow and the eye (symmetrical eye positioning)

Classification of the open and closed state is complex due to the changing shape of the eye, among other factors, position change and face rotation, and blink variations and illumination. All these aspects make it difficult to analyze the eyes in a reliable way. For the problems that have been exposed, it was used a supervised classification method for this task, in this case, a support vector machine (SVM). Figure 11 shows the proposed scheme for checking the state of the eyes.

Table 1. Achieved values of the correct detection rate according to the number of frames

Driver	Number of frames analyzed	Correct detection rate
S1	950	92%
S2	100	93%
S3	400	95%

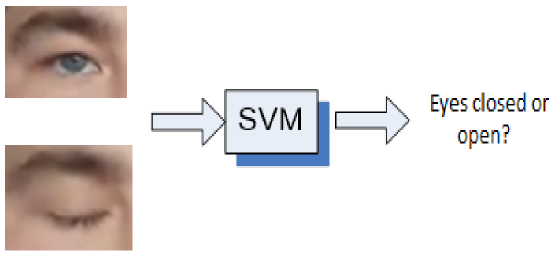

Fig.11. SVM scheme for eye check

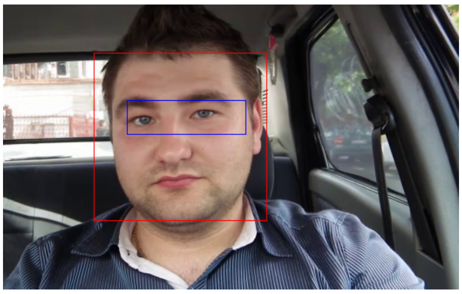

Fig.12. Instant results of the algorithm

We also considered the possibility of training through SVM algorithm if eyes are open or closed, but we have not achieved satisfactory results. The inconvenience we hit was that we did not have a closed eye image database. Several training data is required when training with the SVM algorithm (Fig.12).

Eye status (if opened or closed) is determined by the distance between the first two intensity changes found in the above step. When the eyes are closed, the distance between the y coordinates of intensity changes if it is larger compared to the stage when the eyes are open.
When there are 5 consecutive eyes closed, the alarm is activated, and the driver is warned to wake up. Involves many successive frames closed to avoid including cases of eye closure due blinking.

2.5. Eye tracking

There are a number of reasons for eye tracking. One aspect is the need to follow the face and eyes continuously from one frame to another. A reason is to meet real-time requirements that reduce the search space. The tracking process was developed using the condensation algorithm (CA), in conjunction with SVM for tracking the face with the eye tracking matching template.
The condensation algorithm, which was proposed by Blake, A. and Isard, M. to track active contours, was implemented using a stochastic approach. CA combines samples with a dynamic pattern that is regulated by the state equation.

$Xt = f(Xt\text{-}1, \xi t)$ (10),

Where:

Xt is the state of time t, f (•) is a linear equation and depends on the previous state, plus a white noise. The purpose is to estimate the Xt state vector by means of the observation system which are the achievements of the Zt stochastic process governed by the measurement equation:

$ZT = h(Xt, \eta t)$ (11)

Zt is the system measurement at time t, h (•) is a linear equation which links the other stage, plus a white noise. The ξt and ηt processes are each white noise and are independent of each other. Also, these processes are generally non-Gaussian and multi-modal.

2.6. Face tracking

Previously, it was mentioned that the VJ method has problems in detecting the faces when deviating from the nominal position and orientation, thus, to correct this disadvantage, the face tracking system has been developed. To prove this shortcoming, Fig. 9 shows several times when the VJ method does not find the driver's face and Fig.10 shows an extended example if the actual position and VJ position are represented in a frame sequence. The actual position was obtained by manual recoveries
The main problem of the VJ method is that it is able to locate the human face when it is in the front position of the camera. This shortcoming leads to a distrust of the driver's analysis system throughout the driving process, which is extremely dynamic, for example, when looking at the mirror. Thus, an effective tracker was implemented to correct this problem using the Lukas Kanade algorithm, in conjunction with the SVM technique.

The frequency of incidents for the tests performed over an hour showed some situations - an incident where the driver did not look in the direction of driving, as shown in the chart below (Fig.13.). We find that over the limit of 4 s there were 5 potentially dangerous situations.

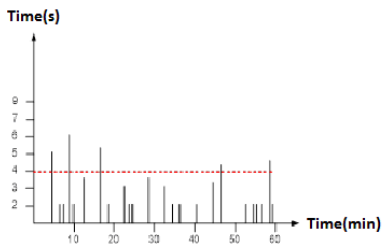

Fig. 13. The duration and number of incidents over 60 minutes of driving

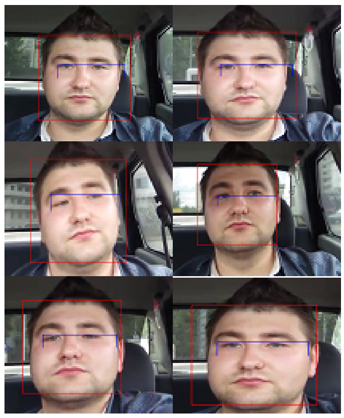

Fig.14. The detection algorithm of the face and the eyes of the driver

3. CONCLUSIONS

SVM is a technique that allows learning to recognize objects. Recognition of the human face is a difficult task in different environmental conditions. SVM technique has been used for image recognition application of other authors and promises by the combined use with other algorithms to provide accurate and fast results.

Approximately 60 ms is needed for a frame on an Atom processor at 1.6 Ghz, which allows the implementation of the algorithm developed for the desired application, that allows for real-time processing.

The algorithm could not be ported in this form on an embedded processor with a 300-800 Mhz frequency, as it requires processing resources superior to the capabilities of these types of devices.

The developed algorithm allows running on an intel atom with 1.8 Ghz frequency, the right processor for a computer installed on the machine.

The developed system works very well in variable light conditions.

This algorithm will be improved and used to acquire images from a webcam to supervise a flexible mechatronics line.

Machine learning algorithms are powerful when resolving some particular problems such as artificial vision.

4. REFERENCES

Bekiaris, E., Petica, S. and Brookhuis, K.A. (1997). Driver needs and public acceptance regarding telematic in-vehicle emergency control aids. Paper presented at the 4th World Congress on ITS, Berlin.

Blake A, Isard M, The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion, issue April (1998) pp. 25-37 Published by Springer London

Cristiani A., Porta M., Gandini D., Bertolotti G.M., Serbedzija N., "Driver Drowsiness Identification by Means of Passive Techniques for Eye Detection and Tracking", Self-Adaptive and Self-Organizing Systems Workshop (SASOW), 2010 Fourth IEEE International Conference on pp. 142-146, 27-28 Sept. 2010.

D'Orazio T., Leo M. and Distante A.: Eye Detection in Faces Images Vigilante System. IEEE Intelligent Vehicles Symposium University of Parma, Italy June, 14-17 (2004)

Dong, W., Wu, X.: Driver fatigue detection based on the distance of eyelid. In: IEEE Int.Workshop VLSI Design & Video Tech., Suzhou, China (2005)

Er. Manoram Vats1 and Er. Anil Garg2, Detection and security system for drowsy driver by using artificial neural network technique. International Journal of Applied Science and Advance Technology June 2012, Vol. 1, No. 1, pp. 39-43

Garg R., Gupta V., Agrawal V., "A Drowsy Driver Detection and Security System", Ultra Modern Telecommunications and Workshops, 2009. ICUMT '09 on, pp. 1-8, 12-14 Oct. 2009.

Hong, Chung,K.: Electroencephalographic study of drowsiness in simulated driving with sleep deprivation. International Journal of Industrial Ergonomics. (Volume 35, Issue 4, April 2005, Pages 307-320.)

Horng W., Chen C. and Chang Y.: Driver Fatigue Detection Based on Eye Tracking and Dynamic Template Matching. Proceedings of the IEEE International Conference on Networking, Sensing & Control (2004)

Ji, Q., Zhu, Z., Lan, P.: Real time nonintrusive monitoring and prediction of driver fatigue. IEEE Trans. Veh. Technol. 53(4), 1052–1068 (2004). doi:10.1109/TVT.2004.830974

Pai-Yuan Tsai; Weichih Hu; Kuo, T.B.J.; Liang-Yu Shyu, "A Portable Device for Real Time Drowsiness Detection Using Novel Active Dry Electrode System", Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE, pp. 3775-3778, 3-6 Sept. 2009.

Smith, P.,& Shah, M.,& da Vitoria Lobo, N. , "Determining driver visual attention with one camera" Intelligent Transportation Systems, IEEE Transactions, Dec, 2003