Machine Vision in Data Science

Driver Drowsiness Detection System

Dimitra Kouvara | Nikitas Mahmudis | Stelios Mathioulakis | Panagiotis Pastroudis

Driver Drowsiness Detection System

Abstract

In this report, the unique role and significant importance of machine vision is realized through the motive of describing the implementation of a machine vision safety assurance system in a car. Approaching the problem of car accidents due to factors, such as the driver, in a systemic way, an archetype of a cause-and-effect loop was evolved, which indicates that a different aspect of the problem must be addressed. Regardless of the advanced technologies that have been enhanced on a car, since the driver's supervision is critical, even on cars with auto-pilot systems, the burden of accidents should not fall explicitly on the automated systems, but also to the supervisor. So, in other words, in this report the idea of *supervising the supervisor* is presented. Nevertheless, such an implementation is applicable not only in cars, but also on every occasion where the human supervision is critical, mandatory or of high importance. The system's objective is to prevent a car driver to lose focus from the driving process due to drowsiness. The key elements, that this system utilizes to detect drowsiness, are the driver's eyes closing or showing signs of fatigue, the head tilting, and the mouth yawning. Two separate system design methodologies are discussed and a final machine vision implementation is tested, evaluated and reported. Several obstacles and efficiency limitations are analyzed as far as the tools, components, logical steps and calculation functions of the final overall approach. In the end, final conclusions are reported and further ideas for future work are presented.

Keywords

Sleepiness detection, Drowsiness detection, Facial landmark detection

Table of Contents

Introduction - Current research

During the last few decades, machine vision has evolved into a mature field embracing a wide range of real-life applications including vehicle guidance, traffic monitoring and control, surveillance, automated inspection, robot assembly, signature verification, biometric measurement, and analysis of remotely sensed images, offering tremendous advancements (Davies, 2004). Meanwhile, the mobile technology with built-in cameras, the affordable and easily accessible computing power, the widely available hardware designed for computer vision and analysis, as well as the new algorithms (like convolutional neural networks) have converged to bring about a renaissance in machine vision.

The emerging role of machine vision in the mapping and localization process is illustrated at the example of autonomous driving. In the study of Goyal et al (2020), autonomous driving car systems having the ability to identify optical information in order to use in continuous choice are presented. Moreover, recognition and classification schemes allow the identification of specific objects in a traffic scene. The work of detection and classification of cars at city crossroads presented by Messelodi et al (2005) is a characteristic example. Machine vision applications are not exclusively used in car-related technology. Machine vision has become sophisticated in medical field as well. Through deep learning and pattern recognition processes, the early diagnosis of brain tumor (Lang et al, 2016), breast, colon, and lung cancer, as well as the detection of genetic mutations are possible. Many applications using machine vision technology have also been developed in agricultural sectors, such as land-based and aerial-based remote sensing for natural resources assessments. Che et al (2002) state that machine vision techniques can aid agricultural production by detecting deceases and defects. Furthermore, according to Aydin and Othman (2017), machine vision techniques are widely established in face detection implementations to address security management.

In the automotive industry, there have been several papers released, pertaining to systems that leverage artificial intelligence applications for this domain. For example, one such case is the continuous authentication of a vehicle driver (Derman & Salah, 2018) or the driver face recognition from images that come from intelligence traffic monitoring cameras (Hu et al., 2019). The former paper tries to address the face detection problem by tackling severe illumination variations, coupled with a single sample (a single image is captured per traffic violation case) problem.

In the area of vehicle safety, specifically, the machine vision research community has been prolific. This study specifically addresses the automatic detection of sleepiness/fatigue of a driver, through machine vision systems and applications. Before the advent and widespread of neural network technologies, attempts at solving this problem were made by leveraging more *traditional* image processing techniques. For instance, Eriksson and Papanikotopoulos (1997) proposed a simple and non-intrusive system that would alert the driver if fatigue signed were detected. They used simple approaches, like edge detection (to locate, first the driver's head and then the eye areas), pixel values' histogram analysis, etc.

Barr et al. (2005) argue that the Federal Motor Carrier Safety Administration (FMCSA), the trucking industry, highway safety advocates, and transportation researchers have identified driver drowsiness as a high priority commercial vehicle safety issue, hence it is a key factor that contributes to the occurrence of a traffic accident. Barr et al. (2005) also suggest that significant advances in video camera and computer processing technologies coupled with robust, non-invasive eye detection and tracking systems have made it possible to characterize and monitor a driver's state of alertness in real time under all types of driving conditions and then reported a comprehensive survey of startups and research groups that were active on the area of driver fatigue detection, at that time.

Conducted studies have shown that around 20% of road accidents are caused due to driver fatigue. Driver drowsiness detection is a car safety technology induced by several car manufactures. More explicitly, detection systems that monitor the steering pattern and the position of the vehicle by lane monitoring camera, as well as driver facial monitoring and physiological measurement by body sensors (brain activity, heart rate, breathing, skin conductance, muscle activity, etc.) methods have been developed.

This computer vision project focuses on the facial monitoring of the driver by camera and on the detection of typically visual signs of drowsiness (difficulty focusing, frequent blinking, or heavy eyelids, yawning repeatedly or rubbing of the eyes, trouble keeping the head up).

Materials and Methods

Sigari et al. (2013) propose a system that can detect, not only fatigue but also distraction signals. In their solution, face template matching and horizontal projection of top-half segment of face image are used to extract hypovigilance symptoms from face and eye, respectively. Head rotation is a symptom to detect distraction that is extracted from face region. The extracted symptoms from eye region are percentage of eye closure, eyelid distance changes with respect to the normal eyelid distance, and eye closure rate.

The recognition of drowsiness's signs in a real-time video stream require the use of a camera. In this case, a built-in webcam acquires frames that are used as input to perform visual searches. The main objective of the implemented algorithm is the detection of the facial features, and more specifically, of the ones that indicate drowsiness. These can either be facial expressions and movements, such as head tilting, yawning, eye blink duration, or even skin blemishes as in bags under the eyes.

Haar Cascade Technique

The first attempt of face recognition was through the classic machine learning based approach of Haar cascades. This technique is considered to be effective in face recognition, as its features can sufficiently detect edges and lines. Pre-trained classifiers provided by OpenCV are used for the front face, profile, eyes, and smile. This implementation approach consists of two parts. The first objective is to achieve an optimized or improved detection of each facial feature, by altering the grayscale spread. The alterations can be blurring, smoothing, or increasing the contrast through histogram equalization. Applying those, did not improve the Haar cascade template. Additional rules were developed to boost the Haar cascade algorithm, such as logical assumptions. More specifically, the eyes and the mouth are expected to be within the face region, and the mouth below the eyes' region. Furthermore, assuming that the facial features were successfully detected, the second objective is to acquire the specific coordinates of the sub-features, such as the eyelid and lip edges, utilizing methods, such as erosion, dilation of several iterations, closing, opening, thresholding techniques, and combining them with Sobel and Canny algorithms.

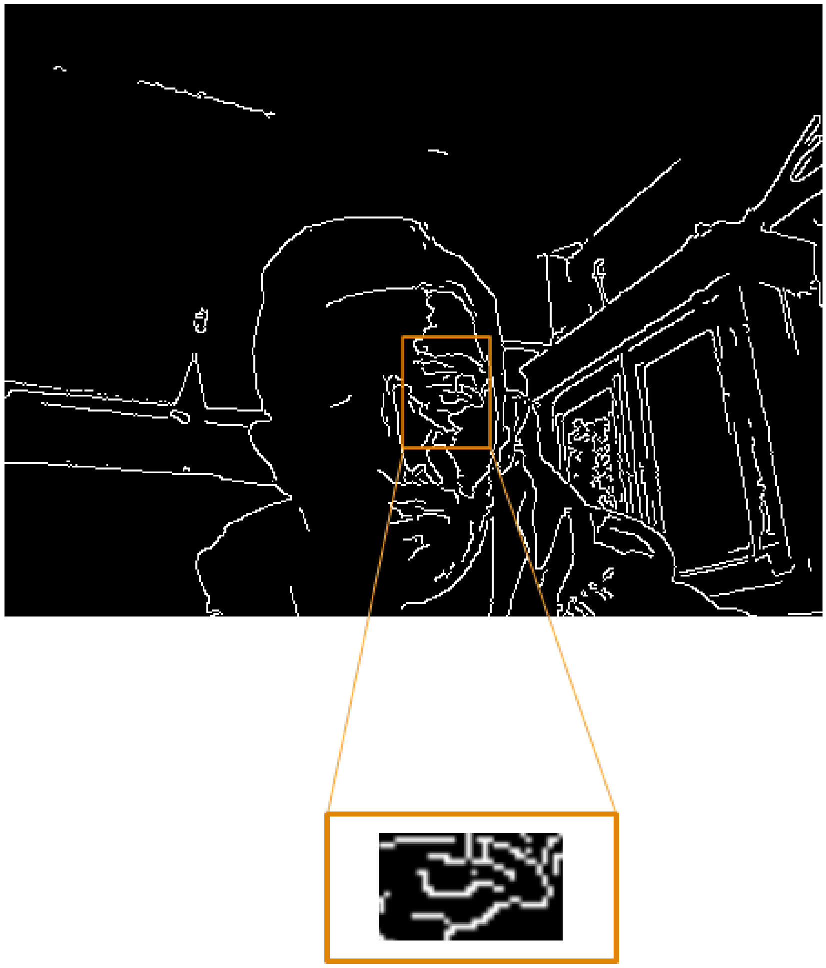


Figure 1: *A frame with Canny edge detection, no smoothing, threshold 120-160 and aperture size 3*

To evaluate the results of the above trials, sampling methods of acquiring images are developed for each pre-processing step.

A major drawback of the previous approach is, that in order to achieve high accuracy on the detection of the facial features, much more image pre-processing is required. This is due to the fact that the Haar features have to be determined manually, resulting to a limitation in object detection, such as partially covered faces by wearing sunglasses. In addition, Haar cascade is able to detect only the facial features on the image inside the ROI (region of interest) box in order to control the environment. In the case where the face is partially out of the ROI, and that is actually a desired position of the head, as tilting to one side is a sign of drowsiness, it cannot be detected. Overall, it seems that the Haar cascade technique is rather inefficient in terms of sensitivity. Since the goal of the project is the drowsiness detection of drivers in real-time, the selection of the facial detector demands a more delicate handling.

Facial Landmark Detectors

The nature of the problem reveals several issues that need to be addressed and solved: the changes in daylight's illumination, a relative face-camera pose (head orientation), image resolution, motion dynamics, the possibility that the object wears either glasses or sunglasses, etc. Based on all the above-mentioned issues, the usage of a proper facial detector is quite vital to the effectiveness of a drowsiness detection system. The solution is offered from the robust real-time facial landmark detectors. As they are publicly available pre-trained on ‘in-the-wild datasets’, they are robust to varying illumination, various facial expressions, and moderate non-frontal head rotations. These state-of-the-art detectors, that are able to capture most of the characteristic points on a human face image, including eye corners and eyelids, inner lips, formulate a regression problem, where a mapping from an image into landmark positions is learned.

Proposed method

One of the most accurate methods on detecting facial parts is by model training. The dlib library offers a facial landmark detector, which produces 68 coordinates each of which translate into specific facial structures. It should be noted that there are plenty of facial databases able to cover large variations, including different subjects, poses, illumination, occlusions etc. However, according to Sagonas et al. (2016), so far, the existed annotations face several restrictions:

- The facial mapping points correspond to a relatively small subset of the overall images.
- In some cases, due to human fatigue, the accuracy of the annotations is a bit low.
- The annotation models differ in terms of the number of facial landmarks.

These restrictions make comparisons between different methods almost infeasible. In an attempt to overcome this problem, the shape detector was trained on a semi-automatic annotation technique that was employed to re-annotate most existing facial databases under a unified protocol.

To retrieve alternative groupings of the regarding landmarks, a custom dictionary can be defined. Using this dictionary, the pixels’ indexes of interest can easily be extracted.

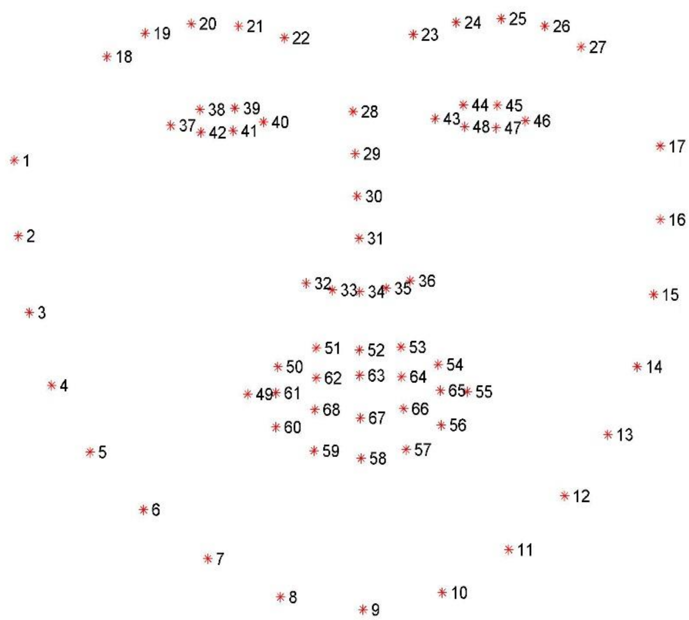


Figure 2: *The 68 mapping points used for the facial point annotations. (retrieved December 20, 2020, from* *<u>https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/</u>)*

The sketch is used for the visualization of the mapping between the facial features and the 68 (x, y)-coordinates:

- The jaw can be accessed through points [0, 17],
- The right eyebrow through points [17, 22],
- The left eyebrow through points [22, 27],
- The nose using [27, 36],
- The right eye using [36, 42],
- The left eye via [42, 48],
- The mouth via [48, 68], and more specifically, the upper lip via [48, 55], the lower lip via [54, 61], and the inner lips via [60, 68].

The main objective of this methodology is the movement detection of the facial features. Such indications have been implemented utilizing these landmarks. For each facial feature, the Euclidean distance of the moving over the stable ones is being calculated, extracting scalar quantities. For example, a closed eye is detected by computing the ratio of distances, between the vertical and the horizontal landmarks. The aspect ratio quantifies the facial feature's change (eye closing, mouth opening, head tilting etc.).

Description of features

In our methodology, we have identified key areas of interest in the face that are likely to reveal symptoms of fatigue and drowsiness. These are the eyelids, the mouth, the tilting motion of the head, and the so-called bag region under the eyes. The methodology behind of each of these key featured is further described below.

Eye closure

Any changes in the frequency and duration of the eye blinks, as well as episodes of slow eye closure are considered to be ocular indications that the subject is suffering from sleep deprivation and circadian rhythm effects (Åkerstedt et al, 2005; Tucker et al, 2005). One of the most characteristic signs of real-time drowsiness is the eye closure.

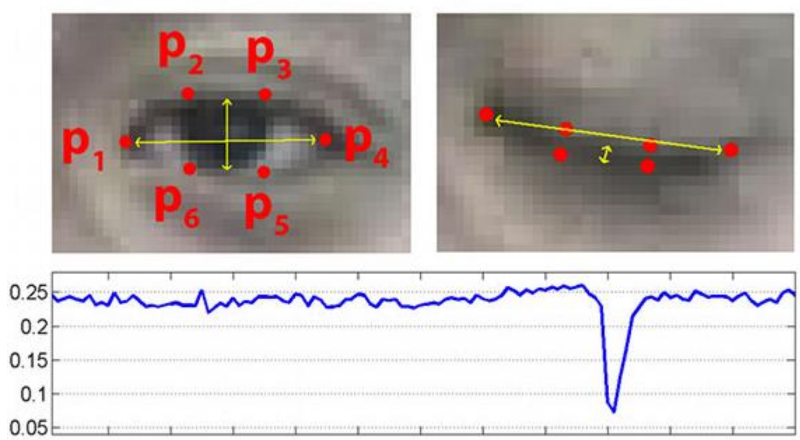


Figure 3: *The line-chart shows the progression of the eye aspect ratio value as the eyelid closes/blinks. Notice the downward spike in eye aspect ratio when the eye is fully closed (Cech & Soukupova, 2016)*

For every frame of the video sequence, the eye aspect ratio (EyeAR) between the vertical and horizontal landmarks can be computed using the following formula:

$$EyeAR = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||}$$

where $p_1, \ldots, p_6$ are the 2D landmark locations (Cech & Soukupova, 2016).

Figure 4: *Eye Aspect Ratio (EyeAR) is shown by facial landmarks for (a) the right and (b) the left eye (Savaş & Becerikli, 2018)*

The $p_1, \ldots, p_6$ are equivalent to the points [36, 42] for the right, and [42, 48] for the left eye of the 68 facial annotations. The EyeAR function is calculated separately for each eye (left aspect ratio and right aspect ratio).

The eye aspect ratio maintains an almost constant value in the case where the eyes remain open, while whenever blinks occur the value decreases. Since eye blinking is performed by both eyes synchronously, the EyeAR of both eyes is averaged:

$$avg\ aspect\ eye\ =\ (left\ aspect\ ratio\ +\ right\ aspect\ ratio)\ /\ 2$$

Note that this heuristic can vary from person to person, since not all individuals share the same facial features in terms of geometry, and so, a small variance is expected. However, it is fully invariant to a uniform scaling of the image and in-plane rotation of the face.

Once the aspect ratio has been computed, a threshold, indicating whether the eyes are closing or not, should be set. The selection of the threshold is based on the study conducted by Zhang et al. (2017). According to their study, the number of blinks is between 15-30 per minute, 0.25 ~ 0.3 seconds every time when driver is awake:

If eye blink detection < 0.25:

    eye closed

Else: eye opened

This frequency gives the optimum result for the system. For every frame in the video sequence that has a value of eye aspect ratio lower than the above-mentioned threshold, eye closure is occurring. A number of 35 consecutive frames is considered sufficient to safely conclude that the subject has his/her eyes closed, and therefore he/she is feeling drowsiness. It should be noted, however, that ideally the number of consecutive frames should be replaced by an actual time frame. This is because cameras can video stream in various frames per second, hence the number of consecutive seconds does not always translate to a specific time frame. However, for this study the simpler approach of defining only the number of consecutive frames was followed.

Yawning

Yawning is one of the natural responses to being drowsy. By definition, it is a mostly involuntary process of opening the mouth and breathing in deeply, filling the lungs with air. There is a significant challenge in its detection, since the mouth may also be open as a result of speaking, or even singing. Many research teams deal with yawning detection focusing their methods on geometric features of the mouth (Fan et al 2007). In the current project, yawning detection is being approached in the same logic as the eye closure computing the aspect ratio between specific facial landmarks of the mouth for the inner and outer lips.

The "outer" lips are meant as the external zone of the mouth (the structure surrounding the oral aperture), where vermillion is included. In Figure 5, the landmarks of the "outer" lips can be recognized, providing the reader with a clearer notion as to which elements of the oral area are considered in the followed analysis.

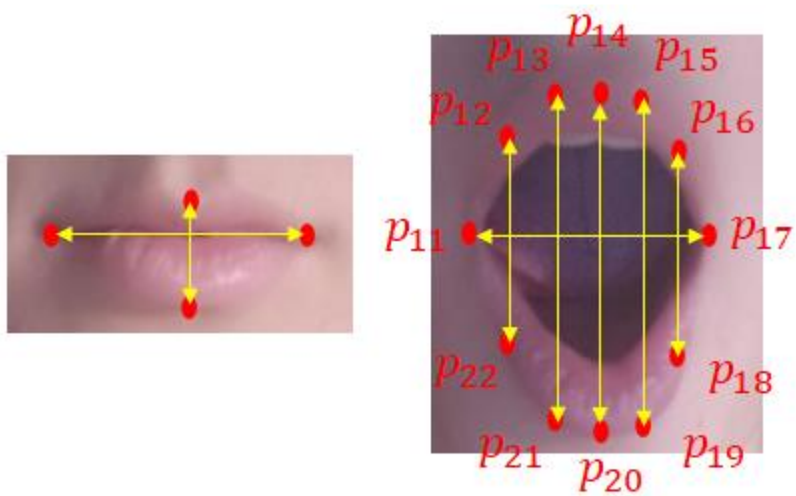


Figure 5: *Detection of (a) close and (b) open mouth with landmarks pi for the "outer" lips*

For every frame of the video sequence, the mouth aspect ratio (MAR) between the vertical and horizontal landmarks can be computed using the following formula:

$$MAR = \frac{\|p_{12} - p_{22}\| + \|p_{13} - p_{21}\| + \|p_{14} - p_{20}\| + \|p_{15} - p_{19}\| + \|p_{16} - p_{18}\|}{2\|p_{11} - p_{17}\|}$$

where $p_{11}, \ldots, p_{22}$are the 2D landmark locations

The "inner" lips are meant as the internal zone of the mouth. When the lips are closed, the oral commissure, which, by definition, is the position where the lateral aspects of the vermillion of the upper and lower lips join, is basically the horizontal distance of the landmarks depicted with the red points, in Figure 6. The distance between the vertical landmarks seems to be almost zero, as the two landmarks coincide.

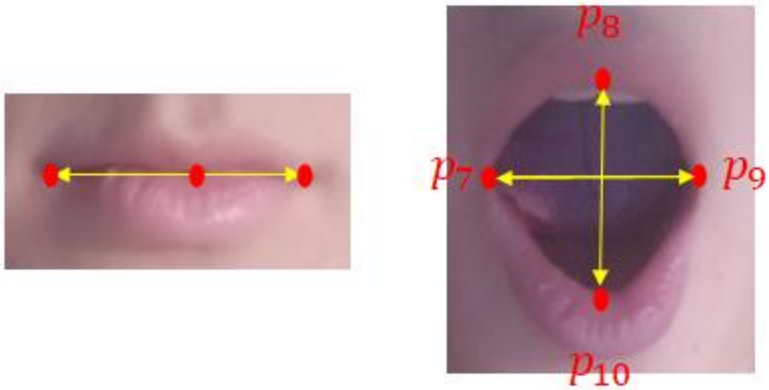


Figure 6: *Detection of (a) close and (b) open mouth with landmarks p_i for the "inner" lips*

For every frame of the video sequence, the inner mouth aspect ratio (InMAR) between the vertical and horizontal landmarks can be computed using the following formula:

$$InMAR = \frac{\|p_8 - p_{10}\|}{\|p_7 - p_9\|}$$

where $p_7, \ldots, p_{10}$ are the 2D landmark locations

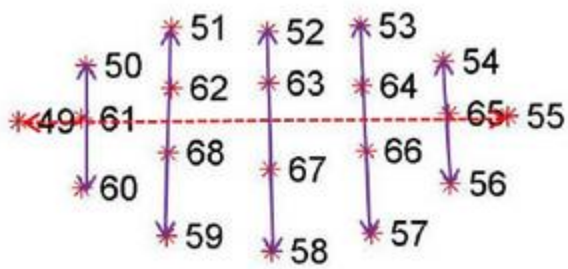


Figure 7: *Mouth Aspect Ratio (MAR) is shown by facial landmarks for the "outer" lips (Savaş & Becerikli, 2018)*

The points $p_7, \ldots, p_{22}$ are equivalent to the points [48, 68] for the mouth of the 68 facial annotations. Due to the differentiation of the physical characteristics between individuals, the oral apertures, and as a result the mouth aspect ratio do not have a single common value. For the case of the "inner" lips, it can be assumed that the aspect ratio's value is close to zero, when the mouth is closed. This value is not absolute and any rare cases i.e., deformations, and extreme cases of buck teeth causing a natural opening of the mouth are excluded.

For the reasons mentioned above, the selection of the thresholds for both "outer" and "inner" lips is confirmed through observation. The thresholds are the indicators that yawning occurs, meaning that any value less than these predetermined values indicate that the mouth is opened.

If “outer” lips detection < 1.6 or “inner” lips detection < 0.3:

mouth closed

Else: mouth opened

Head tilting

The methodology behind detecting head titling as a sign of fatigue is straightforward and simple. By using the nose landmarks, we overlay a line on top of the nose using points 28-34. At the same time, we can construct a reference line starting from point 34 and rising vertically. This line acts as a reference to the Y axis (Figure 8).

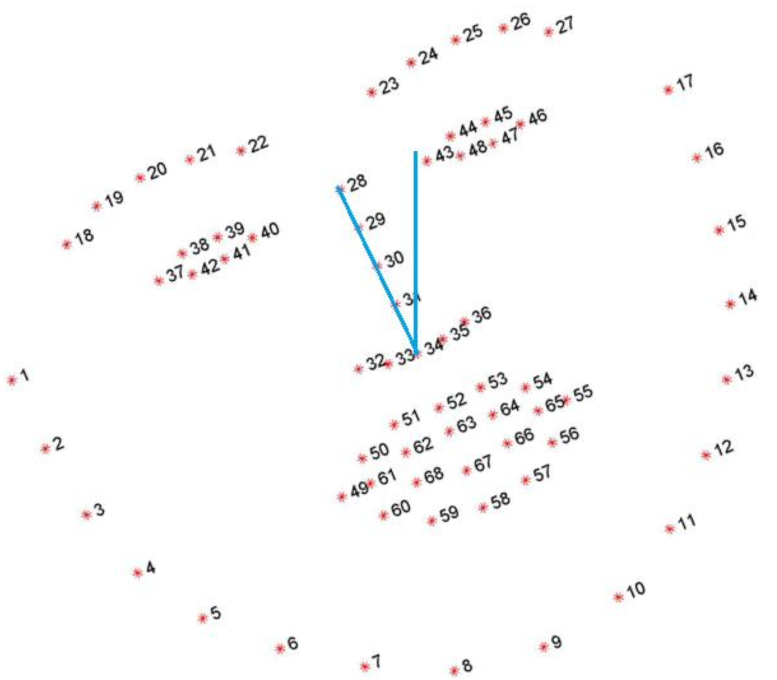
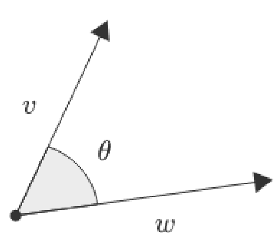


Figure 8*: Illustration of the logic behind the head tilting detector. The blue lines represent the vectors used in the cosine similarity formula (original image retrieved December 20, 2020, from https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/)*

Finally, we can treat those two lines as vectors in a 2-dimensional space and, by using the cosine similarity formula, compute the angle between the lines (Figure 9).



$$\theta = \arccos\left(\frac{v \cdot w}{|v||w|}\right)$$

Figure 9: *Formula used for the computation of the angle between the lines of Figure 8 (retrieved December 21, 2020, from https://wumbo.net/formula/angle-between-two-vectors-arc-cosine/)*

Bags detection

A facial phenomenon that can be used as signal of drowsiness are the so called bags. The bags refer to puffiness caused by an accumulation of fluid under the skin beneath the eyes. The lack of sleep causes blood vessels to leak and mix with those fluids, leading to the formation of dark circles. In order to achieve the identification of eye bags the following steps were followed:

Once the coordinates of the eyes have been located, we can utilize these landmarks to mark a specific area below the eyes, where the indicative skin color occurs. In addition, the fact that each individual has each own style/habits when driving should be considered. For example, some drive by being closer to the steering wheel, while others in a more relaxed way by having a wider distance. Moreover, the relative distance between each individual and the steering wheel can vary during driving. We are able to locate the eye bags distance from the eyes, considering the height of the nose. We assume that the height of each eye bag area is equal to the height of the nose multiplied by 0.2. The next step of the procedure is to capture the average value of the pixels in the region under the eyes and compare it with another part of the face. For that reason, a specific region in the forehead is selected and the average value of pixels is captured. By comparing these two average grayscale values, we are able to conclude that if the difference between the specific area of interest and the forehead is more than 60, then this is an indication of tired eyes. The selection of the threshold is based on the difference from the picture below (Fig. 10), where it is equal to 58.54.

Figure 10: *The image that was used as an extreme example of a person having bags under their eyes*

Results

By applying the methodology described in the previous section, a drowsiness detection system was developed using Python (reference file app.py inside the root folder of the project). The system assumes that a camera is set in front of the driver of the car, on top of the dashboard. Below are screenshots of the system that depict the various drowsiness detection triggers in action (Figures 11 through 13).

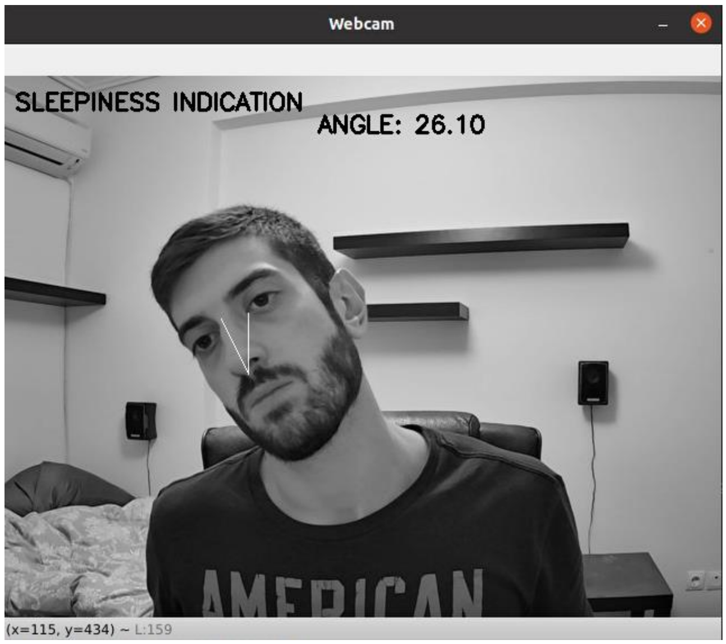


Figure 11*: Head tilting detection gets triggered*

Figure 12: *Detection of eyes starting to close as a sign of drowsiness*

Figure 13: *Detection of yawning*

For each of these triggers, a respective sound alert is being played on a separate Python process, so that the main process of the system does not freeze. Once the driver stops showing signs of fatigue then the alerts stop. The app also allows switching between debug mode and normal mode. In detail, you can set the parameter debug_mode to True (in app.py) to trigger helpful information of what computations take place behind the scenes. The screenshots of Figures 11 through 13 have this setting activated.

Regarding the initial experimentation with Haar cascades, the screenshot below shows some feedback of the bounding boxes of facial structures, as predicted by the Haar cascades method.

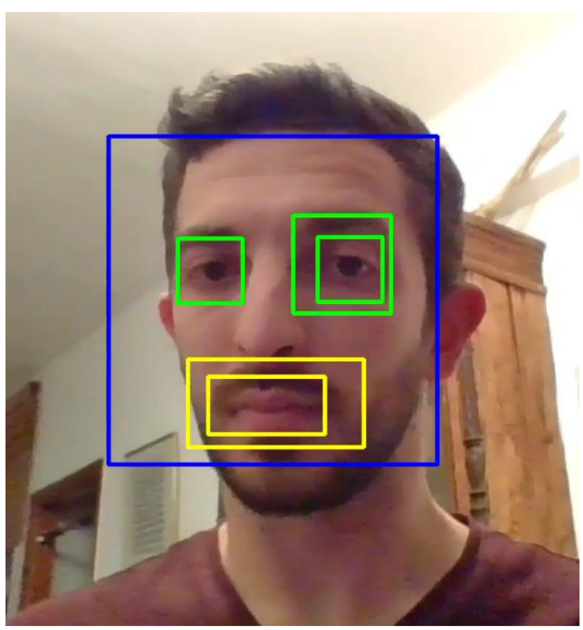


Figure 14: *Initial experimentation with Haar cascades*

Conclusions

This project reflects an effort to address a possible solution to a problem of a complicate system. Such monitoring tools and support systems can be part of an environment for several cases, processes or systems like “black boxes” that the human engagement is critical, such as plane piloting, heavy machinery driver, surgical operations, vessel shipping or in general, situations that the human supervision factor is mandatory. Furthermore, even in a high automated robotic system environment, the human factor is not only a matter of high importance, but also, according with Choon Yue Wong and Gerald Seet (2017), the level and quality of human engagement can induce high supervisor workload and diminish situation awareness. So, we can associate and approach the problem of “Who supervise the supervisor?” with the systemic archetype, “Shifting the Burden” of Daniel H. Kim (1992). The structure of this systemic template indicates that a problem symptom is “solved”, by applying a symptomatic solution, which diverts attention away from a more fundamental solution. In analogy with self-driving vehicles, the automation systems are very well trained, solving symptomatic problems like following the white lines of the road or break if an obstacle occurs, but these systems can fail and that is why manufactures declare that the human supervision is mandatory, which is the fundamental problem that needs to be approached.

The initial attempt of using Haar cascades was not effective mainly since it was not consistent enough to detect the human face. The fact that every few frames some non-relevant squares of templates over the frame were detected as a face, or in every few frames no face was detect drove us to follow another method. In addition, the templates for the facial parts were not so stable and consistent. So, even we had proceeded with edge detection and connected component analysis to locate the coordinates of the landmark points that we are interested on, this implementation would not be able to provide stable and efficient solution. We found out that Haar Cascades are too sensitive on the environment alterations. Preprocessing the frames with techniques such as blurring or histogram equalization, did not improve its performance, not even by enhancing additional logic restrictions regarding the topology of the facial parts. The facial landmark detector was overall successful to detect one or more faces in each frame and furthermore the facial parts such as the eyes, the nose and the mouth. Model training by images, was proved to be significantly more accurate.

Future work

On the subject of future work, there are some alternative methods that could have been tested and some aspects of this project that can be improved. An alternative method or approach on detecting a face could have been to initialize the process of identifying the face through edge detection, bypassing the Haar Cascades. We noticed that some geometrical shapes that have been marked as edges after applying the Canny algorithm could have been utilized to be matched as facial parts. Using a template to identify these shapes, combing them as a system of shapes and calculate the distance among them, could have provide some useful coordinates of facial features. Considering some points of improvement, although the facial landmark detector was overall pretty accurate, there are better ways of dealing with the “bags” under the eye. A better solution would have been to train a binary classifier that would specifically detect “bags” by using a labeled dataset that would contain images labeled as having “bags” and images labeled as not having “bags”. One other feature that could complement the system could have been a detector of distractions during driving. Also, due to lack of time and equipment, the tests were not made in an actual car. If an actual car was used, a camera would have to be installed on the dashboard of the car in order to record the driver. In that case we would have to also deal with nighttime condition using a special camera for that purpose. Cameras enhanced with infrared are able to provide a fare collection of frames even by night, despite of the fact that additional preprocessing on each frame would be required.

References

Eriksson, M., & Papanikotopoulos, N. P. (1997, November). Eye-tracking for detection of driver fatigue. In *Proceedings of Conference on Intelligent Transportation Systems* (pp. 314-319). IEEE.

Sagonas, C., Antonakos, E., Tzimiropoulos, G., Zafeiriou, S., & Pantic, M. (2016). 300 faces in-the-wild challenge: Database and results. *Image and vision computing*, *47*, 3-18.

Sigari, M. H., Fathy, M., & Soryani, M. (2013). A driver face monitoring system for fatigue and distraction detection. International journal of vehicular technology, 2013.

Hu, C. H., Zhang, Y., Wu, F., Lu, X. B., Liu, P., & Jing, X. Y. (2019). Toward driver face recognition in the intelligent traffic monitoring systems. IEEE Transactions on Intelligent Transportation Systems.

Derman, E., & Salah, A. A. (2018, May). Continuous real-time vehicle driver authentication using convolutional neural network-based face recognition. In 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018) (pp. 577-584). IEEE.

Barr, L., Howarth, H., Popkin, S., & Carroll, R. J. (2005). A review and evaluation of emerging driver fatigue detection measures and technologies. National Transportation Systems Center, Cambridge. US Department of Transportation, Washington.

Choon Yue Wong and Gerald Seet (2017, June), Workload, awareness and automation in multiple-robot supervision, International Journal of Advanced Robotic Systems

Daniel H. Kim (1992), System Archetypes 1, Diagnosis Systemic Issues and Designing High Leverage Interventions, Pegasus Communications Inc.

Åkerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2005). Impaired alertness and performance driving home from the night shift: a driving simulator study. Journal of sleep research, 14(1), 17-20.

Tucker, A. J., & Johns, M. W. (2005). The duration of eyelid movements during blinks: changes with drowsiness. Sleep, 28(Suppl), A122.

Cech, J., & Soukupova, T. (2016). Real-time eye blink detection using facial landmarks. Cent. Mach. Perception, Dep. Cybern. Fac. Electr. Eng. Czech Tech. Univ. Prague, 1-8.

Zhang, F., Su, J., Geng, L., & Xiao, Z. (2017, February). Driver fatigue detection based on eye state recognition. In 2017 International Conference on Machine Vision and Information Technology (CMVIT) (pp. 105-110). IEEE.

Savaş, B. K., & Becerikli, Y. (2018, October). Real time driver fatigue detection based on SVM algorithm. In 2018 6th International Conference on Control Engineering & Information Technology (CEIT) (pp. 1-4). IEEE.

Fan, X., Yin, B. C., & Sun, Y. F. (2007, August). Yawning detection for monitoring driver fatigue. In 2007 International Conference on Machine Learning and Cybernetics (Vol. 2, pp. 664-668). IEEE.

Ahmad, I., & Pothuganti, K. (2020, August). Design & implementation of real time autonomous car by using image processing & IoT. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 107-113). IEEE.

Aydin, I., & Othman, N. A. (2017, September). A new IoT combined face detection of people by using computer vision for security application. In 2017 International Artificial Intelligence and Data Processing Symposium (IDAP) (pp. 1-6). IEEE.

Messelodi, S., Modena, C. M., & Zanin, M. (2005). A computer vision system for the detection and classification of vehicles at urban road intersections. Pattern analysis and applications, 8(1-2), 17-31.

Lang, R., Zhao, L., & Jia, K. (2016, October). Brain tumor image segmentation based on convolution neural network. In 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI) (pp. 1402-1406). IEEE.

Chen, Y. R., Chao, K., & Kim, M. S. (2002). Machine vision technology for agricultural applications. Computers and electronics in Agriculture, 36(2-3), 173-191.