Driver drowsiness detection using Behavioral measures and machine learning techniques: A review of state-of-art techniques

Mkhuseli Ngxande  
CSIR Defence, Peace Safety and Security  
Optronic Sensor Systems,  
Email: mngxande@csir.co.za

Jules-Raymond Tapamo  
School of Electrical, Electronic and Computer Engineering,  
University of Kwa-Zulu Natal,  
Durban, South Africa  
Email: tapamoj@ukzn.ac.za

Michael Burke  
Mobile Intelligent Autonomous Systems  
Modelling and Digital Sciences  
Council for Scientific and Industrial Research  
South Africa  
Email: mburke@csir.co.za

Abstract—This paper presents a literature review of driver drowsiness detection based on behavioral measures using machine learning techniques. Faces contain information that can be used to interpret levels of drowsiness. There are many facial features that can be extracted from the face to infer the level of drowsiness. These include eye blinks, head movements and yawning. However, the development of a drowsiness detection system that yields reliable and accurate results is a challenging task as it requires accurate and robust algorithms. A wide range of techniques has been examined to detect driver drowsiness in the past. The recent rise of deep learning requires that these algorithms be revisited to evaluate their accuracy in detection of drowsiness. As a result, this paper reviews machine learning techniques which include support vector machines, convolutional neural networks and hidden Markov models in the context of drowsiness detection. Furthermore, a meta-analysis is conducted on 25 papers that use machine learning techniques for drowsiness detection. The analysis reveals that support vector machine technique is the most commonly used technique to detect drowsiness, but convolutional neural networks performed better than the other two techniques. Finally, this paper lists publicly available datasets that can be used as benchmarks for drowsiness detection.

Keywords—Drowsiness Detection; facial expression; Machine learning; behavioral measures.

I. INTRODUCTION

There is substantial statistical evidence that points to driver drowsiness as a primary cause of road accidents all over the world. Driving for lengthy periods of time can lead to accidents if rest is not taken. The World Health Organization (WHO) have shown that South Africa among African regions has the highest road traffic accident fatalities of about 26.6 % per 100 000 population [1]. Moreover, 1,700 people died on South African roads in the festive season of 2016 alone, a 5% increase on 2015 [2] season. The transport minister of South Africa released a report on the statistics of 2014-2015 annual year, which reveals that 80% of road accidents involve adult males between the ages of 19 and 34 [3]. Furthermore, the minister added that women are most likely to die in road accidents as passengers, especially on public transport. In addition, statistics showed that the top three causes of road accidents in South African roads include distracted drivers (for example, a driver on a phone call), speeding, and driving under the influence of alcohol [4]. Statistics of road accidents as per category of road crashes and casualties are shown in Fig.1.

These incidents have led researchers around the world to investigate methods for early warning drowsiness detection and warning. In addition, many countries and government officials are paying attention to the implementation of solutions to improve driving safety.

Fig. 1. Accidents as per category of road crashes and casualties.

Drowsiness or sleepiness can be described as a biological state where the body is in-transition from an awake state to a sleeping state. At this stage, a driver can lose concentration and be unable to take actions such as avoiding head-on collisions or braking timeously. There are obvious signs that suggest a driver is drowsy, such as:

- Frequently yawning.
- Inability to keep eyes open.
- Swaying the head forward.
- Face complexion changes due to blood flow.
A number of studies recommend countering drowsiness by taking naps between trips, consuming caffeine (coffee, energy drinks etc.), or driving with company [5] [6].

There are various measures to determine the level of driver drowsiness. These measures can be grouped into three categories:

I. Physiological Measures,
II. Vehicle-based Measures, and
III. Behavioral Measures.

In the first category, measurements are obtained by accessing driver’s conditions through the addition of electronic devices onto the skin. This includes Electroencephalography (EEG), Electrocardiography (ECG) and Electrooculogram (EOG) [7] [8] [9]. Although these devices yield highly accurate results, they are not widely accepted because of their practical limitations. For the second category, a driver’s drowsiness is analyzed based on vehicle control systems, which could include steering wheel movements, braking patterns, and lane departure measurements [10]. Steering wheel measurements tend to yield better results than other vehicle-based methods [11]. Vehicle-based methods are non-invasive, but may not be as reliable in detecting drowsiness accurately because they are dependent on the nature of the road and the driver’s driving skills. The last category consists of behavioral or computer vision measures that tend to be reliable than vehicle-based because they focus on the person rather than the vehicle. Furthermore, behavioral measures are non-invasive and more practical than physiological measures. Here, information is obtained by using cameras to detect slight changes in driver’s facial expressions. As behavioral measures are non-invasive in nature, they are becoming a popular way of detecting drowsiness [12].

Existing reviews have been conducted in order to understand advancements in driver drowsiness detection systems. The authors of [13] reviewed the use of head movement-based detection for driver drowsiness. They covered general measures that can be used to detect drowsiness in a driver and provided a comparative analysis of various drowsiness detection systems. More recently, [14] conducted a survey of car safety systems. This included an analysis of signs of drowsiness and various techniques used to measure these signs, with a range of driver drowsiness detection systems reviewed. In addition, [15] presented a survey of driver fatigue-drowsiness detection systems. This work focused on methods that can be used to prevent road accidents and designs for drowsiness detection. Jill and Chisty [16] presented a review of driver drowsiness detection systems. They focused on reviewing existing (2015) drowsiness detection techniques, with an emphasis on pre-processing techniques that can be used on different systems, for example, the circular Hough transform and the Lab color space.

While a large number of reviews have been conducted around driver drowsiness detection, the field has advanced and there is a need for a review of machine learning approaches applied to drowsiness detection. This is particularly relevant given recent advances in deep learning. This paper attempts to address this need by assessing behavioural methods that are based on machine learning techniques for the classification stage of drowsiness detection, as shown in fig.2. This paper provides information on a set of machine learning techniques that one can use to make reliable and precise decisions for driver drowsiness detection systems. The remainder of this paper is organized as follows: A general framework for behavioral driver drowsiness detection using machine learning techniques is described in Section II. Section III. provides a review of metrics that are used in driver drowsiness detection and decision-making techniques. Section IV. gives meta-analysis results and lists publicly available datasets which can be used as benchmarks for drowsiness detection task. Finally, conclusions are provided in Section V.

II. DRIVER DROWSINESS DETECTION PROCESS

Behavioral methods measure levels of drowsiness through the use of mounted cameras in the car to observe facial features such as eye state, head movement, blinking rate and yawning. Most researchers follow a general process to extract facial features from the camera feed. After obtaining these features, further processing is applied to determine the level of drowsiness, typically by applying machine learning techniques such as Support Vector Machines (SVM), Convolutional Neural Networks (CNN) or Hidden Markov Models (HMM). These techniques are trained using features and labelled outputs to build models that can be used for drowsiness prediction. The most challenging part of this process is finding a large dataset that covers the expected variability across races and different skin pigments. This is a particular challenge due to security and confidentiality issues that arise when publishing datasets for academic and commercial use. Figure 2 shows a common framework used for most driver drowsiness detection approaches.

Facial features that are typically extracted from a driver’s face include the following:

1. **Eye closure analysis**: The eye state is an important feature that is widely used to determine drowsiness in the driver. Methods that are used to measure the level of drowsiness include the Percentage of the eye closure (PERCLOS) [17] and eye aspect ratio (EAR). EAR is the ratio between the height and width of the eye and was introduced by Soukupova and Cech in 2016 [18]. In contrast, PERCLOS is the percentage of eye closure over a period of time. The primary difference between the two is that EAR classifies the ratio of the eye as it decreases whereas PERCLOS classifies whether the eye is open or closed.

2. **Eye blink rate**: Methods that measure the blinking rate use the frequency of eye-blinks to measure drowsiness. The normal blinking rate per minute is roughly 10. When the driver is drowsy, the blinking rate decreases.

3. **Yawning analyses**: Yawning can be caused by fatigue or boredom, in drivers it indicate that they might fall asleep while driving. Methods can measure the widening of the mouth to detect yawning traits in the
driver by tracking mouth shape and position of lip corners [19].

4. Facial expression analysis: This approach makes use of a combination of more than one facial feature to detect drowsiness in a driver. This includes features such as wrinkles in the forehead and extreme head poses [20] [21].

 Various measures are used in different studies for detecting a face and extracting features from the video feed. Unfortunately, most of these studies use differing datasets that may favor their own algorithms. This is due to the lack of standardised datasets that can be used as a benchmark. As a result, it is hard to compare approaches by simply evaluating reported accuracies. Machine learning techniques to classify different levels of drowsiness are now discussed, along with a review of measures that form a driver drowsiness detection system.

A. Support Vector Machines (SVM)

SVMs are supervised learning methods for classification and regression [23] [24]. SVMs were firstly introduced by Boser, Guyo, and Vapnik in 1992 [25]. SVMs attempt to find a hyperplane that separates training data into pre-defined classes. In the driver drowsiness field, SVMs are primarily used to learn to classify different states of the driver from labelled data. A great deal of work has attempted to utilize the capabilities of SVMs in the detection of drowsiness. A number of measures have been used as features to determine a driver’s level of drowsiness using SVMs. A comparison of these measures is presented in Table I and the approaches included briefly described below.

The authors of [26] proposed a fully automatic system that is capable of detecting driver drowsiness. For face detection and eye extraction, the well-known Haar feature algorithm was used [26]. An SVM was then trained to classify when eyes are open or closed and to trigger an alarm. Similarly, [27] proposed a system that can also detect driver drowsiness and distraction. Here, the Viola and Jones algorithm was used for face detection and color histograms with Local Binary Patterns (LBP) applied to track the face over frames. The system achieved an accuracy of 100% in face detection, but a potential downfall of this approach is the low frame rate achievable, which could result in missed facial expressions.

B. Hidden Markov Model (HMM)

HMMs are a statistical model used to make predictions about hidden states based on observed states defined by probabilities. HMMs were developed in the late 1960’s and early 1970’s by Leonard Baum and colleagues [28]. Today, HMMs have a widespread use in applications such face expression recognition, gene annotation, modeling DNA sequence errors, and computer virus classification [29] [30].

Table II shows the range of features and approaches used by HMM-based drowsiness detectors, but Zhang et al [31] and...
Choi et al. [21] omitted some of the required information for making comparison and are not used for meta-analysis step.

The authors of [32] proposed a new facial feature by using changes in wrinkles detected by calculating the local edge intensity on the face. They used an Infra-red (IR) camera to eliminate illumination changes and allow for operation in both day and night conditions. Unfortunately, this system can yield false results when is used on older people because they have deeper wrinkles. In contrast, [33] implemented HMM techniques for eye tracking based on color and geometrical features. For illumination elimination, authors used a two-level Lloyd-max quantization intended to be robust to illumination changes [33]. Unfortunately, this system is designed for indoor conditions and it fails to detect the face if the driver is not facing forward.

### TABLE I. DRIVER DROWSINESS DETECTION THROUGH SVMs

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Measure</th>
<th>Classifiers</th>
<th>Frame per second(fps)</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. J.AL-Anzy et al.</td>
<td>2015</td>
<td>Eye closure</td>
<td>Haar features with SVM</td>
<td>60</td>
<td>99.74</td>
</tr>
<tr>
<td>M. Sabet et al. [27]</td>
<td>2012</td>
<td>Eye state</td>
<td>SVM</td>
<td>25</td>
<td>98.4</td>
</tr>
<tr>
<td>L. Pauly and D. Sankar</td>
<td>2015</td>
<td>Eye state</td>
<td>HOG and SVM</td>
<td>5</td>
<td>91.6</td>
</tr>
<tr>
<td>A. Punitha et al. [35]</td>
<td>2014</td>
<td>Eye state</td>
<td>SVM</td>
<td>15</td>
<td>93.5</td>
</tr>
<tr>
<td>B. N. Manu [36]</td>
<td>2016</td>
<td>Eye closure and Yawning</td>
<td>SVM with Linear kernel</td>
<td>15</td>
<td>94.58</td>
</tr>
</tbody>
</table>

### TABLE II. HMM TECHNIQUE ON DRIVER DROWSINESS DETECTION

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Metric</th>
<th>Classifiers</th>
<th>Frame per second(fps)</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang et al. [31]</td>
<td>2015</td>
<td>Eye state</td>
<td>HMM</td>
<td>N/A*</td>
<td>95.9</td>
</tr>
<tr>
<td>A. Bagci and R. Ansari [33]</td>
<td>2004</td>
<td>Eye state</td>
<td>HMM</td>
<td>3</td>
<td>99.7</td>
</tr>
<tr>
<td>Patt et al. [37]</td>
<td>2007</td>
<td>Eye Blink</td>
<td>HMM</td>
<td>25</td>
<td>95.7</td>
</tr>
<tr>
<td>I. H. Choi et al. [21]</td>
<td>2016</td>
<td>Eye state and Head position</td>
<td>HMM</td>
<td>16–20</td>
<td>N/A*</td>
</tr>
<tr>
<td>E. Tadesse et al. [38]</td>
<td>2014</td>
<td>Eye closure and other features</td>
<td>HMM and SVM</td>
<td>20</td>
<td>97</td>
</tr>
<tr>
<td>Y. Sun et al. [39]</td>
<td>2013</td>
<td>Eye blinks</td>
<td>SVM and HMM</td>
<td>61</td>
<td>90.99</td>
</tr>
</tbody>
</table>

### C. Convolutional Neural Network (CNN)

CNN’s are similar to an ordinary neural network which is also made up of neurons that consist of learnable weights [40]. CNN’s make use of layers of spatial convolutions that are well suited for images, which exhibit strong spatial correlations. CNN’s have proven successful in areas such as image recognition, video analysis, and classification [41]. Yann Le Cun and Yoshua Bengio were the first to apply CNN’s in computer vision [42], but the excellent performance of CNNs in computer vision was only made apparent in 2012, when deep convolutional neural networks showed excellent results in object recognition [43]. Table III gives a brief review of CNN based methods in the detection of drowsiness.

[44] proposed an algorithm for driver drowsiness detection using representation learning. Here, the popular Viola and Jones algorithm was used to detect the faces. Images were cropped to 48*48 square images and fed into the first layer of the network which consisted of 20 filters. The whole network contains two layers. The output of the CNN was passed to a softmax layer for classification. This system did not consider head pose changes and as a result can fail. However, the authors of [45] used a 3D deep Neural Network to obtain more accurate results. Here, the face is tracked by a combination of a Kernelized Correlation filter with a Kalman filter [45] for robust face tracking. The extracted face regions are then passed to a 3D-CNN which is followed by a gradient boosting machine for classification. This system works well even if the driver is changing head position [44].

### TABLE III. CNN TECHNIQUE ON DRIVER DROWSINESS DETECTION

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Metric</th>
<th>Methods</th>
<th>Classifiers</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>F. Zhang et al.</td>
<td>2017</td>
<td>Eye state</td>
<td>AdaBoost, LBF and PERCLOS</td>
<td>CNN</td>
<td>95.18</td>
</tr>
<tr>
<td>K. Dwivedi et al.</td>
<td>2014</td>
<td>Visual features</td>
<td>Viola and Jones algorithm</td>
<td>CNN with softmax Layer</td>
<td>78</td>
</tr>
<tr>
<td>A. George and A. Routray [47]</td>
<td>2016</td>
<td>Eye gaze</td>
<td>Viola and Jones algorithm</td>
<td>CNN</td>
<td>98.32</td>
</tr>
<tr>
<td>B. Reddy et al.</td>
<td>2017</td>
<td>Eye state</td>
<td>Eye state and mouth</td>
<td>MTCNN and DDDN</td>
<td>91.6</td>
</tr>
</tbody>
</table>

### IV. META-ANALYSIS

Although a great deal of work has been conducted to date there is significant room to improve driver drowsiness detection systems. The primary challenge identified through this review is that each of the reviewed systems used different datasets to achieve their goals and as a result are not easily compared. In addition, the datasets used to test these systems tend to be limited and are typically captured in controlled environments, which can lead to failure in real-world situations.

In an attempt to provide a fair comparison, a Meta-analysis was performed using 25 papers collected for this literature study. The collected papers primarily use classification accuracy to compare the performance of systems. Performance estimation revealed that CNNs yielded more accurate results when compared to SVMs and HMMs. A non-parametric Skillings-Mack test was conducted and rendered a Chi-square value of 6.66 which was significant $p = 0.035709$, indicating...
that there is a difference in performance between the methods compared. Tests were performed on reported accuracies obtained on the ULg Multimodality Drowsiness Database(DROZY) [49], ZJU Eye blink Database [50], Yawn Detection Dataset(YawnDD) [51], Eye-Chimera [52] and the NTHU-drowsy driver detection video dataset [53]. Figure 3 shows boxplots of the accuracies obtained for each method, along with the raw data [27], [37], [46]-[48], [54]-[59] and associated databases. It appears as if CNN’s outperform other approaches, but it is difficult to compare HMM’s and SVMs due to insufficient data. It should be noted that these tests are based on reported accuracies, which may be biased or favour methods proposed in the reporting papers.

Among all literature, SVM’s were the most used classifiers, followed by CNNs and HMMs. It is also clear that there has been a rise in CNN use in the field of driver drowsiness since 2012, accompanying the rise in deep learning elsewhere in computer vision.

Machine learning techniques such as SVM, CNN, and HMM are reviewed in this paper. Unfortunately, it is extremely difficult to compare these approaches, as there is a limited number of standardized datasets that currently exist to do so. In an attempt to remedy this, a meta-analysis was performed. This analysis highlighted the performance of CNNs, which outperformed other approaches, but also showed that there is a need for larger datasets and standard benchmarking measures for drowsiness detection. Future work will focus on the creation of a suitable dataset that covers a wide range of different races for making more reliable drowsiness comparisons.

**REFERENCES**


