

Review

Driver Distraction Using Visual-Based Sensors and Algorithms

Alberto Fernández ^{1,*}, Rubén Usamentiaga ², Juan Luis Carús ¹ and Rubén Casado ²

¹ Grupo TSK, Technological Scientific Park of Gijón, 33203 Gijón, Asturias, Spain; juanluis.carus@grupotsk.com

² Department of Computer Science and Engineering, University of Oviedo, Campus de Viesques, 33204 Gijón, Asturias, Spain; rusamentiaga@uniovi.es (R.U.); rcasado@lsi.uniovi.es (R.C.)

* Correspondence: alberto.fernandez@grupotsk.com; Tel.: +34-984-29-12-12; Fax: +34-984-39-06-12

Academic Editor: Gonzalo Pajares Martinsanz

Received: 14 July 2016; Accepted: 24 October 2016; Published: 28 October 2016

Abstract: Driver distraction, defined as the diversion of attention away from activities critical for safe driving toward a competing activity, is increasingly recognized as a significant source of injuries and fatalities on the roadway. Additionally, the trend towards increasing the use of in-vehicle information systems is critical because they induce visual, biomechanical and cognitive distraction and may affect driving performance in qualitatively different ways. Non-intrusive methods are strongly preferred for monitoring distraction, and vision-based systems have appeared to be attractive for both drivers and researchers. Biomechanical, visual and cognitive distractions are the most commonly detected types in video-based algorithms. Many distraction detection systems only use a single visual cue and therefore, they may be easily disturbed when occlusion or illumination changes appear. Moreover, the combination of these visual cues is a key and challenging aspect in the development of robust distraction detection systems. These visual cues can be extracted mainly by using face monitoring systems but they should be completed with more visual cues (e.g., hands or body information) or even, distraction detection from specific actions (e.g., phone usage). Additionally, these algorithms should be included in an embedded device or system inside a car. This is not a trivial task and several requirements must be taken into account: reliability, real-time performance, low cost, small size, low power consumption, flexibility and short time-to-market. The key points for the development and implementation of sensors to carry out the detection of distraction will also be reviewed. This paper shows a review of the role of computer vision technology applied to the development of monitoring systems to detect distraction. Some key points considered as both future work and challenges ahead yet to be solved will also be addressed.

Keywords: driver distraction detection; visual-based sensors; image processing

1. Introduction

According to the most recent published World Health Organization (WHO) report, it was estimated that, in 2013, 1.25 million people were killed on the roads worldwide, making road traffic injuries a leading cause of death globally [1]. Most of these deaths happened in low- and middle-income countries, where rapid economic growth has been accompanied by an increased motorization and therefore, road traffic injuries. In addition to deaths on the roads, up to 50 million people incur non-fatal injuries each year as a result of road traffic crashes, while there are additional indirect health consequences associated with this growing epidemic. Road traffic injuries are currently estimated to be the ninth leading cause of death across all age groups globally, and are predicted to become the seventh leading cause of death by 2030 [1].

Distracted driving is a serious and growing threat to road safety [1]. Collisions caused by distracted driving have captured the attention of the US Government and professional medical organizations during the last years [2]. The prevalence and identification as a contributing factor in crashes is seen as an *epidemic* of American roadways, in words of Ray LaHood, when he was US Secretary of Transportation in 2012 [3]. There is not an exact figure regarding statistics about accidents caused by inattention (and its subtypes) since studies are made in different places, different time frames and therefore, different conditions. The studies referenced below show both the different statistics about inattention in general and those recorded when produced by distraction and fatigue in particular. These authors have estimated that distraction and inattention account for somewhere between 25% and 75% of all crashes and near crashes [4–8].

The trend towards increasing the use of in-vehicle information systems (IVISs) is critical [9] because they induce visual, manual and cognitive distraction [10] and may affect driving performance in qualitatively different ways [11]. Additionally, the advancement and prevalence of personal communication devices has exacerbated the problem during these last years [12]. All these factors can lead to the increment of the number of tasks subordinate to driving activity. These tasks, namely secondary tasks, which may lead to distraction [13], include eating, drinking, the act of taking something or tuning the radio or the use of cell phones and other technologies. The secondary tasks that take drivers' eyes off the forward roadway [14,15] reduce visual scan [16] and increase cognitive load may be particularly dangerous [13]. For example, the use of cell phones while driving, according to naturalistic studies [17], causes thousands of fatalities in the United States every year [18,19].

The purpose of this paper is the analysis of the state-of-the-art regarding the detection of drivers' distraction. The scope of the paper can be seen in Figure 1 and is commented as follows. The main methods for face detection, face tracking and detection of facial landmarks are summarized in Section 2 because they are a key component in many of the video-based inattention monitoring systems. In Sections 3–5, the main algorithms for biomechanical, visual and cognitive distraction detection are reviewed, respectively. Additionally, in Section 6, there are some algorithms detecting mixed types of distraction and, hence, are also reviewed. The relationship between facial expressions and distraction is also explored in Section 7. The key points for the development and implementation of sensors to carry out the detection of distraction will be considered in Section 8. In Section 9, the key ones to test and train driving monitoring systems are summarized. Privacy issues related to camera sensors are commented in Section 10. Lastly, conclusions, future aspects and challenges ahead will be considered in Section 11.

With the objective of introducing the scope and limitations of this review, some key aspects have been briefly introduced as follows. Driver distraction is just one form of inattention, which occurs when drivers divert their attention away from the driving task to focus on another activity. Therefore, a “complete” solution should consider all aspects of inattention. At least, the system should detect both distraction and drowsiness as the main contributing factors in crashes and near-crashes. As stated before, in this work, only distraction algorithms are summarized but one must not forget that other forms of inattention should be taken into account. Moreover, the use of on-board sensors already available in the vehicle to analyze driver behaviour is a low-cost and powerful alternative to the vision-based monitoring systems [20,21]. However, these systems should not be treated like different alternatives, because they can be used together (fused) in order to obtain indicators for monitoring [22]. Hence, for the sake of completeness, in this paper review only “purely” vision-based monitoring systems have been reviewed.

One of the challenges in decreasing the prevalence of distracted drivers is that many of them report that they believe they can drive safely while distracted [23]. However, for example, in connection with the use of mobile phones while driving, there is a great deal of evidence interacting with mobile devices, such as sending messages or engaging in conversations, which can impair driving performance because this interaction can create distraction. Moreover, a recent research showed that phone notifications alone significantly disrupted performance, even when drivers did not directly

interact with a mobile device during the task [24]. Another study suggests that people in general can reduce both inattention and hyperactivity symptoms simply by silencing the smartphones and avoiding notifications [25]. Therefore, it is clear that drivers should not use and notice the presence of the smartphones inside the car while driving. It should be pointed out that distraction generation is a very complex process and is scarcely addressed here. We recommend some research papers that focused on driver distraction generation: Angell et al. [26] focused on the process of cognitive load in naturalistic driving; Liang et al. [27] addressed the adaptive behaviour of the driver under task engagement and their results on visual, cognitive and combined distraction; Caird analyzed the effects of texting on driving [28]. In the context of intelligent vehicles, Ohn et al. [29] highlights the role of humans by means of computer vision techniques.

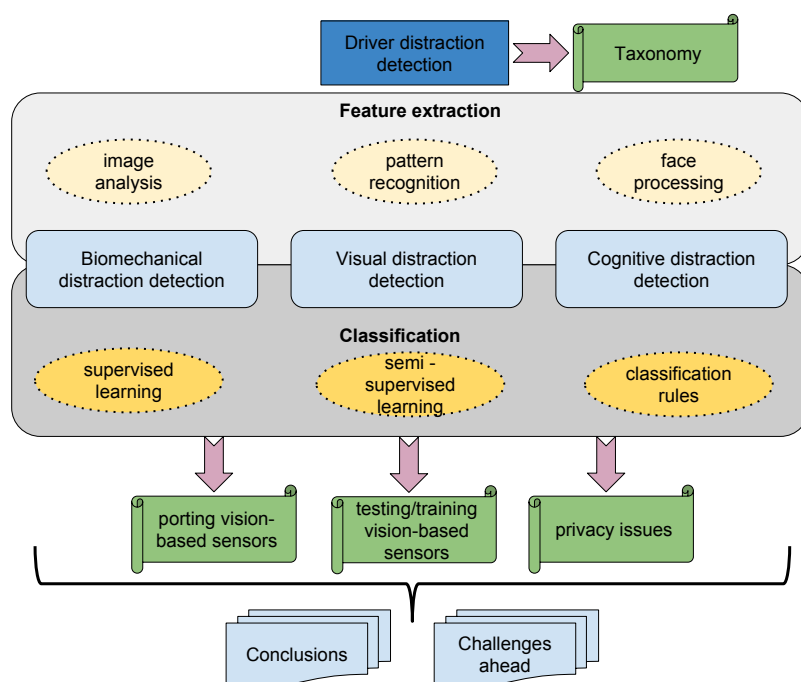


Figure 1. Scope of the present work.

1.1. Taxonomy

Both distraction and inattention have been inconsistently defined and the relationship between them remains unclear [30]. The use of different, and sometimes inconsistent, definitions of driver distraction can create a number of problems for researchers and road safety professionals [31]. Inconsistent definitions across studies can make the comparison of research findings difficult or impossible, can also lead to different interpretations of crash data and, therefore, to conclude different estimates of the role of distraction in crashes. This problem can be further seen in these recent works [32–35]. Many definitions have been proposed in order to define distraction [5,7,8,31]. Regan et al. [35] proposed a taxonomy of both driver distraction and inattention in which distraction is conceptualized as just one of several factors that may give rise to inattention. They concluded that driver inattention means “insufficient or no attention to activities critical for safe driving”. They defined driver distraction as “the diversion of attention away from activities critical for safe driving toward a competing activity, which may result in insufficient or no attention to activities critical for safe driving”. The definition proposed here is almost identical to that coined for driver distraction by Lee et al. [31].

It is acknowledged that the taxonomy proposed by Reagan et al. [35] suffers from “hindsight bias”, that is, the forms of driver inattention proposed are derived from studies of crashes and critical incidents in which judgements have been made after the fact about whether or not a driver was

attentive to an activity critical for safe driving [35]. Driving consists of a variety of sub-tasks and it may not be possible to attend to all at the same time. Determining which sub-task is more important (and the driver, thus, should attend to) can often only be determined after the fact (i.e., after a crash or incident occurred) and, hence, this attribution of inattention is somewhat arbitrary [36]. Additionally, the dynamics of distraction [37], which identifies breakdowns on interruption as an important contributor to distraction should also be considered as part of this taxonomy, and hence, timing and context have implications on the algorithm design that should be taken into account.

1.2. Methodology

Papers addressed in this review are within the topic of distraction detection using vision-based systems. The search and review strategy is described below. A comprehensive review of the English language scientific literature was performed. It encompassed the period from 1 January 1980 to 31 August 2016. The following databases were used: EBSCO, ResearchGate, ScienceDirect, Scopus, Pubmed, Google Scholar and Web of Knowledge. Search terms related to driver distraction were employed combining all of them: driver, visual, cognitive, manual, biomechanical, vision, vision-based, impairment, distraction, distractions, review, task, tasks, inattention, performance, phone, sms, vehicle, problem, looking, face, head, pose, glasses, illumination, self-driving, tracking, sensors, image, traffic, safety, facts, privacy, issues, porting, taxonomy. Many items were returned from the search criteria shown before. These were, then, reviewed using the following criteria. Exclusion criteria were obviously non-relevant papers or from medical, electronic, networking, marketing and patent topics. Only publications from peer-reviewed English language journals were considered for inclusion. Additionally, reviewed papers were ordered by the number of references in order to include all relevant papers. Finally, in order to get the latest published papers, search filters were applied for this purpose. Search filters were applied to get publications only from years 2015 and 2016. References and bibliographies from the selected papers identified were examined to determine potentially additional papers. A total of approximately 1500 publications were revised in the review process.

2. Face and Facial Landmarks Detection

A common face processing scheme in many inattention monitoring systems, which can be seen in Figure 2, includes the following steps:

- Face detection and head tracking. In many cases a face detection algorithm is used as a face tracking one. In other cases, a face detection algorithm is used as an input for a more robust face tracking algorithm. When the tracking is lost, a face detection call is usually involved (that is why in Figure 2 these steps are placed inside the same external orange box).
- Localization of facial features (e.g., eyes). Facial landmarks localization is usually performed, but it should be noted that, in some cases, no specific landmarks are localized. So, in such cases, estimation of specific cues are extracted based on anthropometric measures from both face and head.

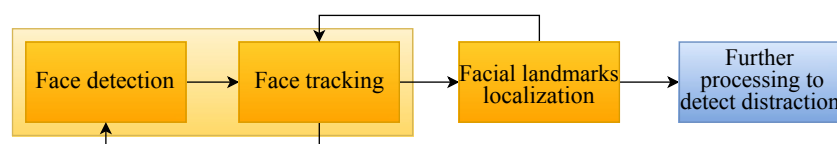


Figure 2. Common steps in most distraction monitoring systems.

2.1. Face Detection

Viola-Jones [38] have made object detection practically feasible in real world applications, which contains three main ideas that make possible to build and run in real time: the integral image,

classifier learning with AdaBoost, and the attentional cascade structure [39]. This framework is used to create state-of-the-art detectors (e.g., face detector [40]), available, for example, in OpenCV library. However, this framework turned out to be really time-consuming [41]. Moreover, cascade detectors work well on frontal faces but sometimes, they fail to detect profile or partially occluded faces.

One possible solution is to use the standard approach for human detection [42], which can also be used for face detection [39]. This approach is based on the Histogram of Oriented Gradients (HOG), which is a feature descriptor used in computer vision and image processing for the purpose of object detection. This approach can be trained with less images and faster [43]. Deep Learning approaches can also be used for face detection. For example, in [44], a deep learning approach, called DP2MFD, is used. DP2MFD detects faces at multiples scales, poses and occlusion by integrating deep pyramid features with Deformable Parts Models (DPMs). Experiments were carried out on four publicly available unconstrained face detection datasets, which demonstrated the effectiveness of the approach. However, this face detector was tested on a machine with 4 cores, 12 GB RAM, 1.6 GHz processing speed and it took about 26 s. Consequently, complex features may provide better discrimination power than Haar-like features for the face detection task. However, they generally increase the computational cost [44].

Some modifications to the Viola-Jones algorithm have been proposed [45,46] to speed up the algorithm. For example, in [45], different optimization techniques to speed up the Viola-Jones detector for embedded smart camera applications have been discussed. In their paper, skin colour information is integrated with the Viola-Jones detector in order to reduce the computation time. PICO (Pixel Intensity Comparison-based Object detection) is another modification of the standard Viola-Jones object detection framework, which scans the image with a cascade of binary classifiers at all reasonable positions and scales [46]. This algorithm can achieve competitive results at high processing speed. This is especially evident on devices with limited hardware support for floating point operations. PICO outperforms the other two OpenCV detectors in terms of accuracy and processing speed.

Since driver face monitoring system should work in all light conditions, lighting and camera selection is one of the most important stage in the design of the system. Lighting devices not only should provide enough light in environment, but they also should not hurt his/her eyes. For example, learning-based methods (e.g., Viola-Jones algorithm or PICO) can also be used for face detection in Infrared (IR) images [47].

Position of the camera inside the car is another key factor in the detection rate. For example, in [48], if the camera is installed under the front mirror of the car, face detection has 85% accuracy. But if it is installed on the dashboard, face detection reaches up to 93%. This is because they used the Viola-Jones face detector, which is trained to distinguish faces that are tilted up to about 45° out of plane (towards a profile view) and up to about 15° in plane. Therefore, if the camera is installed on the dashboard, the captured images will contain frontal or near-frontal faces. In [49], the camera was placed over the steering wheel column for two reasons: a) it facilitates the estimation of gaze angles, such as pitch, which is relevant for detecting distraction, and b) from a production point of view, it is convenient to integrate a camera into the dashboard. On the downside, when the wheel is turning, there will be some frames in which the drivers face may be occluded by the steering wheel. However, the driver is seldom very sleepy or inattentive to traffic while turning the steering wheel.

2.2. Face Tracking

Head pose estimation can be defined as the ability to infer the orientation of a person's head relative to the view of a camera and different studies have reported statistics showing consistent range of head motion [50], which (see Figure 3) can be decomposed in:

- Sagittal flexion/extension, i.e., forward to backward movement of the neck usually from -60° to 70° , which can be characterized by pitch angle.
- Axial rotation, i.e., right to left rotation of the head usually from -80° to 75° , which can be characterized by yaw angle.

- Lateral bending, i.e., right to left bending of the neck usually from -41° to 36° , which can be characterized by roll angle.

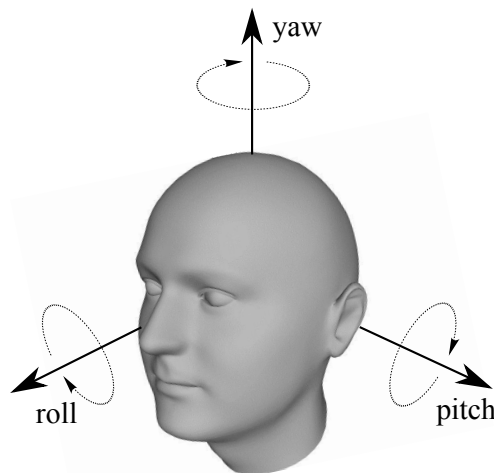


Figure 3. Head pose can be decomposed in pitch, yaw and roll angles.

Many vision-based algorithms for pose estimation have shown good performance when the head is near frontal, which is 95% of the time. But it is during those 5% of the time when interesting events, which are critical for safety, will occur [51]. Furthermore, as face orientation while driving is normally frontal, if the driver faces in other directions for a long period of time, this is probably due to fatigue or inattention [52]. Hence, a key component for a monitoring system based on face tracking is the ability to robustly and continuously operate even during large head movements. However, face tracking remains a challenging vision problem and, hence, a system for a continuous estimation of head movement is needed. On the other hand, as many head tracking algorithms have shown good performance when the head is near frontal, it can be concluded that the driver is looking away when tracking is unavailable. This information could be an alternative approach instead of adding more cameras to increase the range of the tracker.

Accordingly, numerous research works and publications have been trying to perform face tracking using a single camera and they are discussed as follows. Head pose estimation methods based on geometric approach using facial landmark and its 3D correspondences [49,53–56] can provide a good estimation and operate in real-time. For example, in [53], 3D pose estimation is achieved based on the position of the eyes and the mouth. A similar approach is proposed in [54], where only three points (eye centers and the middle point between the nostrils) are used to estimate continuous head orientation and gaze direction. Very closed to this approach, in [55], at least four prominent facial features are extracted from the face. After that, their correspondence on a 3D generic-face model is used to estimate head orientation. Oyini et al. [56] proposed the visual analysis of head position using a single camera aligning and scaling the 3D head model of the face according to the position and distance between the two eyes of the face in the 2D image. Another interesting approach recently published is [49], where a 3D head pose estimation system is proposed. This system is based on the 49 tracked 2D facial landmarks from Supervised Descent Method (SDM) tracker [57].

Other options include the combination of information [58–60], using for example, several classifiers [58,59] or combining 2D and 3D algorithms [60]. Asthana et al. [58] developed a system able to handle 3D pose variations up to $\pm 45^\circ$ in yaw and $\pm 30^\circ$ in pitch angles combining four different face detectors based on Viola-Jones framework. The drawback of this approach is that it requires four classifiers in order to track the face so it increases the execution time and memory requirements. In [59], the system consists of three interconnected modules, which detects drivers'

head, provides initial estimates of head pose, and continuously tracks its position and orientation in six degrees of freedom. Pelaez et al. [60], combined 2D and 3D algorithms to provide head pose estimation and regions of interest identification based on 3D information from a range imaging camera.

Alternatively, more than a camera can be used to implement the tracking [51,61–63], that is, a distributed camera system is commonly used, where two or more cameras can be located inside the car cockpit. Following this line of research, in [61], they proposed a distributed camera framework for gaze estimation using head pose dynamics based on the algorithm proposed in [51]. They predict three gaze zones: right, front and left. In [51], a continuous head movement estimator (CoHMEt) is proposed, which independently tracks the head in each camera, and their outputs are further analyzed to choose the best perspective and corresponding head pose. When tracking is lost, due to either the loss of facial point detection or the rejection of the estimated points, reinitialization is performed using a scoring criterion. In [62], they also used a two-camera system to overcome challenges in head pose estimation, which allows for continuous tracking even under large head movements, as proposed in [51]. Therefore, following the setup of [51], a two-camera system can provide a simple solution in order to improve tracking during large head movements. Two cameras are also used in [63] for head pose estimation. Head pose is tracked over a wide operational range in the yaw rotation angle using both camera perspectives.

For a quantitative evaluation over the algorithms the Mean Absolute Error (MAE) is commonly used. Best results for the different algorithms can be seen in Table 1, where different databases are used. For example, in [49,56] the public database of Boston University (BU) is used to evaluate the performance of the proposed head pose estimation scheme. Some other algorithms used naturalistic on-road data set [59]. Moreover, some algorithms achieving good performance did not indicate any database [60]. LISA-P Head Pose database [55] introduces head pose data from on-road daytime and nighttime drivers of different age, race and gender, with continuous ground truth measurements and manual annotation of facial features. Therefore, this database can be used to compare head pose algorithms and head behaviour studies. The LISA-P Head Pose Database consists of 14 video sequences of drivers in on-road driving environment in natural and spontaneous conditions. The video sequences were collected at a frame rate of 30 frames per second, with a 640×480 pixel resolution.

Table 1. Mean Absolute Error (MAE) (in degrees) of face tracking algorithms comparison working in an automobile environment.

Algorithm	Roll(°)	Yaw(°)	Pitch(°)
La Cascia et al. [64]	9.8	4.7	2.4
Oyini et al. [56] average results 1 camera	5.3	3.9	5.2
Oyini et al. [56] uniform illumination 1 camera	4.8	3.8	3.9
Oyini et al. [56] varying illumination 1 camera	5.3	5.1	6.3
Vicente et al. [49] 1 camera	3.2	4.3	6.2
Pelaez et al. [60] 1 Kinect device	2.7	3.8	2.5
Murphy et al. [59] 1 camera	2.4	4.7	3.4
Tawari et al. [51] (MPS + POS) 1 camera	3.0	8.2	7.6
Tawari et al. [51] (MPS + POS) 2 cameras	3.8	7.0	8.6
Tawari et al. [51] (MPS + POS) 3 cameras	3.5	5.9	9.0
Tawari et al. [51] (CLM + POS) 1 camera	3.4	6.9	9.3
Tawari et al. [51] (CLM + POS) 2 cameras	3.6	5.7	8.8
Tawari et al. [51] (CLM + POS) 3 cameras	2.7	5.5	8.5

Based on the results from Table 1, in [56], the MAE decreased by an average of 1.3° due to illumination variations. In [51], the best performance of 3.9% failure rate, which is the percentage of the time that the system output is unreliable, is achieved with the three-camera view compared with that of over 15% for the single view, which is a significant improvement.

2.3. Location of Facial Features

The detection of facial features (also called landmarks) is an essential part of many face monitoring systems. The problem of the precise and robust detection of facial landmarks has drawn a lot of attention during this decade. State-of-the-art methods include tree models [65,66], DPM [67], SDM [57], explicit shape regression [68] or learning local binary features [69]. A comprehensive survey of facial feature point detection can be seen here [70]. All the above listed research suffers more or less from a lack of verification and performance analysis with a realistic variation in lighting conditions. Therefore, further research should be performed in order to adapt these algorithms to the traffic research in general and to the drivers' monitoring systems in particular. Difficulties for proper detection of drivers' facial features are mainly due to the non-uniformity of light sources, asymmetric shades on their face and eye regions, or rapid changes in light intensity during real-world driving due to shadows caused by buildings, bridges, trees, or, for example, when entering or leaving a tunnel [71].

Eyes, as one of the most salient facial features reflecting individuals' affective states and focus of attention [72], have become one of the most remarkable information sources in face analysis. Eye tracking serves as the first step in order to get glance behaviour, which is of most interest because it is a good indicator of the direction of the driver's attention [73]. Glance behaviour can be used to detect both visual and cognitive distraction [74]. It has also been used by many studies as an indicator of distraction while driving [75] and has been evaluated in numerous ways [73]. Therefore, both eye detection and tracking form the basis for further analysis to get glance behaviour, which can be used for both cognitive and visual distraction.

Eye tracking data is typically captured through the use of a vehicle instrumented with an in-vehicle eye tracker system. On one hand, complex systems consist of single or multiple cameras directed at the driver's face. As the number of face cameras increases, so does the ability of the system to capture larger and more dramatic head movements of the driver. On the other hand, simpler systems consisting of one or two cameras are usually less expensive and easier to install than more complex systems. For example, in [76], a comparison of eye tracking systems with one and three cameras using Smart Eye technology [77] is performed. The system uses a single standard camera of VGA resolution together with IR flash illuminators. The three-camera system used is the Smart Eye Pro [77], which has similar properties as the one-camera system, but it also facilitates gaze direction in full 3D.

Eye detection is required before eye region processing. Eye detection methods can be divided into two general categories: (1) methods based on imaging in IR spectrum; and (2) feature-based methods. A literature survey on robust and efficient eye localization in real-life scenarios can be seen in [72], and a review on eye localization in car environment can be seen in [78].

Methods based on imaging in IR spectrum, which are commonly called "hardware-based" approaches, rely on IR illuminators to generate the bright pupil effect to driver head pose and gaze estimation. These methods use two ring-type IR light-emitting diodes: one located near the camera optical axis and the other located far from it. This approach is often used to detect visual distraction. In contrast to these methods, in [79], the authors use a progressive camera and only one on-axis lighting source [80]. In this situation, the camera always produces images with bright pupils and image processing techniques are applied to detect pupils. Based on thresholding techniques, the possible pupils can be selected. An appearance model, trained using Principal Component Analysis (PCA) and Support Vector Machine (SVM), is exploited to verify the final pupils. To increase the robustness against eyeglasses, the Generalized Symmetry Transform (GST) is incorporated achieving a recognition rate of 99.4% and 88.3% for users not wearing and wearing eyeglasses, respectively.

Regarding feature-based methods, different techniques are commonly applied. Image binarization [81], projection [82,83], face anthropometric properties of the face [84], individual classifiers [85] or particle filtering [86] can be used to detect driver's eyes. For example, in [86], an algorithm for eyes tracking based on particle filtering is proposed. Their method works with a low-cost IR camera device at a low frame rate. They used a single particle filter to track both eyes at the same time. Evaluation was carried out in a driving simulator with five users achieving an average accuracy of 93.25%. In [85],

two individual classifiers based on Haar-like features, one for the head and another for both eyes, were used. They tested face and eye detection in their research vehicle in daylight conditions achieving a hit rate of 97.2% for eye detection and a false alarm of 4.6%.

All in all, the task of accurate eye localization is challenging due to the high degree of eyes appearance variability: facial expression variations, occlusion, pose, lighting and other imaging conditions and quality [72], are frequently encountered in car environments. Another problem that is scarcely addressed in the literature is that, in strong sunlight, the driver tends to squint, which makes, even more difficult to track the eyes. To mitigate these deficiencies, different approaches can be adopted. Sigari et al. [82] proposed to extract symptoms of hypo-vigilance based on eye-region processing but without explicit eye detection stage. Flores et al. [84] proposed a combination of algorithms in order to deal with illumination conditions for both day and night. Rezaei et al. [71] used a methodology to enhance the accuracy, performance and effectiveness of Haar-like classifiers, especially for complicated lighting conditions. These authors also proposed ASSAM [87], which is based on the asymmetric properties of the driver's face due to illumination variations. A good solution is also to use a "divide and conquer" strategy to handle different variations at different stages [72].

3. Biomechanical Distraction

In connection with biomechanical detection and recognition using computer vision techniques, we can find two approaches. The first one involves hands secondary tasks recognition involving hands action, while the second one is based on hands tracking and information.

3.1. Secondary Tasks Involving Biomechanical Distraction

Zhao et al. [88–91] proposed different machine learning approaches to detect predefined driving postures, where four predefined postures were considered: (1) grasping the steering wheel; (2) operating the shift lever; (3) eating; and (4) talking on a cellular phone, which are recorded from the passenger seat, that is, from the right profile view of the driver. Yan et al. [92] proposed a combination of the Motion History Image (MHI) and POHG, and the application of Random Forest (RF) classifier for driving actions recognition. Trying to improve the accuracy of the aforementioned approach, the same authors included a Convolutional Neural Network (CNN) [93], which was tested over three datasets covering four driving postures: (1) normal driving; (2) responding to a cell phone call; (3) eating; and (4) smoking. For fair comparison, Yan et al. [93] re-implemented aforementioned state-of-the-art approaches [88–91] and carried out experiments on other two popular vision descriptor approaches (PHOG [94] and SIFT [95]). Classification accuracy of all of these methods can be seen in Table 2 evaluated on the Southeast University (SEU) driving posture dataset [88].

In connection with secondary tasks recognition, different computer vision algorithms have been proposed in order to detect cell phone usage of the driver while driving [96–100]. High recognition rates are usually obtained (from 86.19% to 95%) using very different approaches. Computer vision techniques seem to be the best approach for this task, whose results can be seen in Table 3, compared to other non-computer vision algorithms relying on inertial sensors of the mobile phone [101]. Best results are obtained by the algorithm proposed by Xu et al. [99], which consists of two stages: first, the frontal windshield region localization using DPM; next, they utilized Fisher vectors (FV) representation to classify the driver's side of the windshield into cell phone usage violation and non-violation classes. The proposed method achieved about 95% accuracy with a dataset of more than 100 images with drivers in a variety of challenging poses with or without cell phones.

It can be concluded that many different computer vision and machine learning techniques can be used to recognize predefined postures involving hand gestures. The CNN model offered a better performance than other approaches but with some limitations. The algorithm needs high computational resources making difficult to be applied in some conditions with common hardware architecture (e.g., embedded systems). Moreover, training a CNN needs a large amount of data, which is also difficult to obtain in some scenarios.

Table 2. Classification accuracy evaluated on the Southeast University (SEU) driving posture dataset [88].

Algorithm	Features	Classifier	Average Accuracy (%)
Zhao et al. [88]	Homomorphic filtering, skin-like regions segmentation and Contourlet Transform (CT)	RF	90.63
Zhao et al. [89]	Geronimo-Hardin-Massopust (GHM) multiwavelet transform	Multiwavelet Transform	89.23
Zhao et al. [90]	Histogram-based feature description by Pyramid Histogram of Oriented Gradients (PHOG) and spatial scale-based feature description	Perceptron classifiers	94.20
Zhao et al. [91]	Homomorphic filter, skin-like regions segmentation, canny edge detection, connected regions detection, small connected regions deletion and spatial scale ratio calculation	Bayes classifier	95.11
Bosch et al. approach [94]	PHOG	SVM	91.56
Lowe et al. approach [95]	SIFT	SVM	96.12
Yan et al. [93]		CNN	99.78

Table 3. Computer vision algorithms to detect cell phone usage. High recognition rates are usually obtained using very different approaches.

Algorithm	Features	Classifier	Recognition Rate (%)
Zhang et al. [96]	Features from the driver's face, mouth and hand	Hidden Conditional Random Fields (HCRF)	91.20
Artan et al. [97]	Image descriptors extracted from a region of interest around the face	SVM	86.19
Berri et al. [98]	Percentage of the Hand and Moment of Inertia	FV	91.57
Xu et al. [99]	DPM	FV	95
Seshadri et al. [100]	Raw pixels and HOG features	Real AdaBoost, SVM, RF	93.86

3.2. Hands Information

Hand detection is a challenging problem as human hands are highly deformable and are also exposed to different illumination conditions [102]. One approach for object detection relies on a sliding-window, where a model is learned based on positive samples (i.e., hands in different poses) of fixed size and negative samples with no hands. A classifier is then used to learn a classification rule. In order to detect hands at different scales, this scheme can be applied on hand images at different sizes. But a sliding window-based approach trained on hand instances was shown to be prone to false positive detection rates [103]. A recent common approach to improve the results is the assumption that hands can only be found in a small and predefined set of regions [103,104].

As opposed to training a model for hand shape or appearance and running a sliding window detector, two different approaches are analyzed in [103] taking into account three activity classes:

(1) two hands on the wheel; (2) hands on the instrument panel and (3) hand on the gear shift. The motion-cue-based hand approach uses temporal accumulated edges in order to maintain the most reliable and relevant information motion and then, it is fitted with ellipses in order to produce the location of the hands. The static-cue-based approach uses features in each frame in order to learn a hand presence model for each of the three regions and a second-stage classifier (SVM) produces the final activity classification. Martin et al. [104] also constraint the problem of hands detection to a number of regions of interest. They used HOG at different scales. Afterwards, a SVM is used to learn a hand presence in each of the three regions and ‘two hands on the wheel’ model for the wheel region. A similar approach is proposed in [102], training a linear SVM model for each region using a different set of descriptors. Ohn et al. [62] incorporated hand gestures in order to study preparatory motions before a maneuver had been performed, training a hand detector using fast feature pyramids. Gradient and colour channels are extracted for each patch image. They used CIE-LUV colour channels because they worked better compared to RGB and HSV. Afterwards, an AdaBoost classifier was applied in order to learn the features from the hands and finally, they trained a SVM-based detector using HOG features to differentiate the left hand from the right one. Later on, Ohn et al. [105] also explored the use of a pyramidal representation for each region of interest using HOG finding that edge features are particularly successful in the task of hands detection.

In order to compare these algorithms, a dataset of synchronized RGB and depth videos collected in an operating vehicle was proposed [106]. The CVRR-HANDS 3D dataset was designed in order to study natural human activity under difficult settings (background, illumination, occlusion) containing three subsets: (1) hand localization; (2) hand and objects localization; and (3) 19 hand gestures for occupant-vehicle interaction. Five regions of interest were considered: (1) wheel; (2) lap; (3) hand rest; (4) gear; and (5) instrument panel. Recognition rates from some of these previous algorithms using this database can be seen in Table 4.

Table 4. Hands recognition in different regions inside the car using CVRR-HANDS 3D dataset [106].

Algorithm	Features	Classifier	Regions	Recognition Rate (%)
Ohn et al. [106]	RGB data	SVM	5	52.1
Ohn et al. [106]	RGB combined with depth data	SVM	5	69.4
Martin et al. [104]	Hands cues	SVM	3	83
Martin et al. [104]	Hands and head cues	SVM	3	91
Ohn et al. [105]	Hands cues	SVM	3	90
Ohn et al. [105]	Hands and head cues	SVM	3	94

Summarizing, a common approach is to recognize if the hands are positioned in one of the established areas (wheel, gearbox and so on) and to track them over time. It could be considered that the steering wheel is the critical area because it is where hands should remain most of the time while driving. If hands remained in a non-critical zone for a certain period of time, which could be different for each of the non-critical areas, an alarm would be created to warn drivers to lay their hands in the correct position.

Hand Disambiguation

There is another interesting problem to solve related to hands detection that needs further research: hand disambiguation [107]. Once hands are detected, it is crucial to ensure that the hands belong to the driver. Both hand disambiguation and hand activity detection should be studied and considered together in order to infer final, clear and unambiguous results.

4. Visual Distraction

Visual distraction is often related to the on-board presence of electronic devices such as mobile phones, navigation or multimedia systems, requiring active control from the driver. It can also be

related to the presence of salient visual information away from the road causing spontaneous off-road eye glances and momentary rotation of the head. A 2006 report on the results of a 100-car field experiment [4] showed that almost 80% of all crashes and 65% of all near-crashes involved drivers looking away from the forward roadway just prior to the incident.

Engagements in visually distracting activities divert drivers' attention from the road and cause occasional lapses, such as imprecise control of the vehicle [108], missed events [28], and increasing reaction times [108]. Visual time sharing between the driving task and a secondary task reveals that the glance frequency to in-car devices is correlated to the task duration, but the average glance duration does not change with task time or glance frequency [109]. Drivers do not usually increase the glance duration for more difficult or longer tasks but rather increase the accumulated visual time sharing duration by increasing the number of glances away from the road [110]. As both single long glances and accumulated glance duration have been found to be detrimental for safety [110–112], a driver distraction detection algorithm based on visual behaviour should take both glance duration and repeated glances into account [113].

On one hand, high-resolution cameras placed throughout the cabin are needed to view the driver's eyes from all head positions and at all times. Several economic and technical challenges of integrating and calibrating multiple cameras should be tackled to achieve this. Technically, eye orientation cannot always be measured in vehicular environments because eye region can be occluded by (1) sunlight reflections on eyeglasses; (2) the eye blink of the driver; (3) a large head rotation; (4) sunglasses; (5) wearing some kind of mascaras; (6) direct sunlight; (7) hats, caps, scarves; or (8) varying real world illumination conditions.

On the other hand, many security systems do not require such detailed gaze direction but they need coarse gaze direction to reduce false warnings [114,115]. For example, forward collision warning (FCW) systems need not only exterior observations but interior observations of the driver's attention as well to reduce false warnings (distracting and bothering the driver), that is, coarse gaze direction can be used in order to control the timing of warning emission when the system detects that the driver is not facing forwards.

Taking into account that errors in facial feature detection greatly affect gaze estimation [116], many researchers have measured coarse gaze direction by using only head orientation with the assumption that coarse gaze direction can be approximated by head orientation [117]. Head pose is a strong indicator of a driver's field of view and his/her focus of attention [59]. It is intrinsically linked to visual gaze estimation, which is the ability to characterize the direction in which a person is looking [118]. However, it also should be noted that drivers use a time-sharing strategy when engaged in a visual-manual task where the gaze is constantly shifted between the secondary task and the driving scene for short intervals of time [119] and often position the head in between the two involved gaze targets and only uses the eyes to quickly move between the two targets. In this situation, a face tracking algorithm would recognize this as a distracted situation based on head position, but the driver is constantly looking the road ahead. Therefore, in an ideal situation, both driver gaze tracking and eyes-off-road should be detected together [49].

In short, visual distraction can be categorized into two main approaches as it can be seen in Figure 4. In the first approach, which can be called "coarse", researchers measured the coarse gaze direction and the focus of attention by using only head orientation with the assumption that the coarse gaze direction can be approximated by the head orientation. In the second approach, which can be called "fine", researchers considered both head and eye orientation in order to estimate detailed and local gaze direction.

Moreover, considering its operating principles, visual distraction systems can be grouped in two main categories: hardware- and software-based methods. Additionally, some systems can combine these two approaches and therefore, a third category can also be considered, as seen in Figure 4.

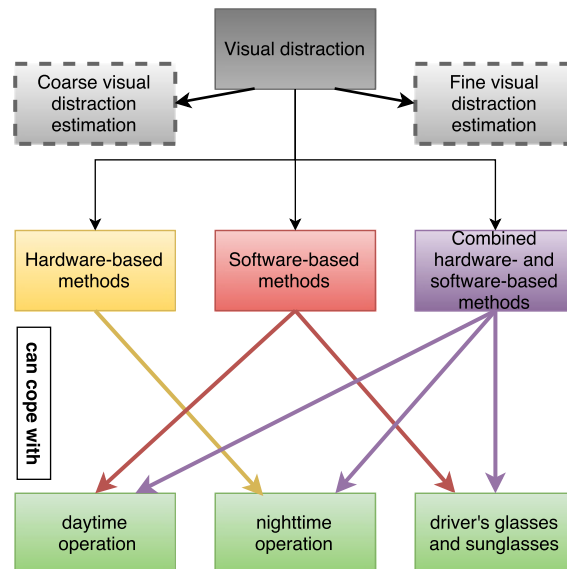


Figure 4. Visual distraction algorithms categorization.

4.1. Hardware-Based Methods to Extract Gaze Direction

Hardware-based approaches to head pose and gaze estimation rely on Near Infrared (NIR) illuminators to generate the bright pupil effect. These methods use two ring-type IR light-emitting diodes: one located near the camera's optical axis and the other located far from it [120–126]. The light source near the camera optical axis makes a bright pupil image caused by the red-eye effect, and the other light source makes a normal dark pupil image. The pupil was, then, easily localized by using the difference between bright and dark pupil images. Ji et al. used the size, shape, and intensity of pupils, as well as the distance between the left and right pupil, to estimate the head orientation. Specifically, the authors used the pupil-glint displacement to estimate nine discrete gaze zones [121,122], a geometric disposition of the IR LEDs similar to that of Morimoto et al. [120] and two Charge Coupled Device (CCD) cameras embedded on the dashboard of the vehicle. In connection with the CCD cameras, the first one is a narrow angle camera, focusing on the driver's eyes to monitor eyelid movement while the second one is a wide angle camera focusing on his/her head to track and monitor head movement. Based on this work, Gu et al. [124] proposed a combination of the Kalman filtering with the head motion to predict the features localization and used Gabor wavelet in order to detect the eyes constrained to the vicinity of predicted location. Another existent approach proposed by Batista et al. used dual Purkinje images to estimate a driver's discrete gaze direction [125]. A rough estimation of the head-eye gaze was described based on the position of the pupils. The shape of the face is modeled with an ellipse and the 3D face pose is recovered from a single image assuming a ratio of the major and minor axes obtained through anthropometric face statistics. In this method, further research is necessary in order to improve the accuracy of the face orientation estimation, which is highly dependent on the image face ellipse detection.

The aforementioned NIR illumination systems work particularly well at night. The major advantage of these methods is the exact and rapid localization of the pupil. However, performance can drop dramatically due to the contamination introduced by external light sources [126,127]. In addition, during daytime, sunlight is usually far stronger than NIR light sources and hence, the red-eye effect may not occur. Moreover, these methods could not work with drivers wearing glasses because the lenses create large specular reflections and scatter NIR illumination [127–129]. While the contamination due to artificial lights can easily be filtered with a narrow band pass filter, sunlight contamination will still exist [126]. Furthermore, such systems are vulnerable to eye occlusion caused by head rotation and blinking [114].

4.2. Software-Based Methods to Extract Gaze Direction

Combining facial feature locations with statistical elliptical face modelling, Batista et al. [83] presented a framework to determine the gaze of a driver. To determine the gaze of the face, an elliptical face modelling was used taking the eye's pupil locations to constraint the shape, size and location of the ellipse. The proposed solution can measure yaw head rotation over $[-30^\circ, +30^\circ]$ interval and pitch head rotation over $[-20^\circ, +20^\circ]$ interval.

Furthermore, despite the technical challenges of integrating multiple cameras, Bergasa et al. [130] proposed a subspace-based tracker based on head pose estimation using two cameras. More specifically, the initialization phase was performed using the Viola and Jones algorithm [40] and a 3D model of the face was constructed and tracked. In this work, head pose algorithm, which was the base for visual distraction estimation, could track the face correctly up to $[-40^\circ, +40^\circ]$.

A limitation of the software-based methods is the fact that they cannot often be applied at night [126,131]. This has motivated some researchers to use active illumination based on IR LEDs, exploiting the bright pupil effect, which constitutes the basis of these systems [126,131] (explained in previous section), or combine both methods, which can be seen in the next section.

4.3. Hardware- and Software-Based Methods to Extract Gaze Direction

Lee et al. [114] proposed a system for both day and night conditions. A vision-based real-time gaze zone estimator based on a driver's head orientation composed of yaw and pitch is proposed. The authors focused on estimating a driver's gaze zone on the basis of his/her head orientation, which is essential in determining a driver's inattention level. For night conditions, additional illumination to capture the driver's facial image was provided. The face detection rate was higher than 99% for both daytime and nighttime.

The use of face salient points to track the head was introduced by Jimenez et al. [132], instead of attempting to directly find the eyes using object recognition methods or the analysis of image intensities around the eyes. The camera was modified to include an 850 nm band-pass filter lens covering both the image sensor and the IR LEDs in order: (a) to improve the rejection of external sources of IR radiation and reduce changes in illumination and (b) to facilitate the detection of the pupils, because the retina is highly reflective of the NIR illumination of the LEDs. An advantage of salient points tracking is that the approach is more robust to the eyes occlusion whenever they occur, due to the driver's head or body motion.

Later on, the same authors extended their prior work in order to improve non-invasive systems for sensing a driver's state of alert [133]. They used a kinematic model of the driver's motion and a grid of salient points tracked using the Lukas-Kanade optical flow method [132]. The advantage of this approach is that it does not require one to directly detect the eyes, and therefore, if the eyes are occluded or not visible from the camera when the head turns, the system does not lose the tracking of the eyes or the face, because it relies on the grid of salient points and the knowledge of the driver's motion model. Experiments involving fifteen people showed the effectiveness of the approach with a correct eyes detection rate of 99.41% on average. It should be noted that this work is focused on sensing the drivers' state of alert, which is calculated measuring the percentage of eyelid closure over time (PERCLOS), and it is not focused on distraction detection.

Eyes Off the Road (EOR) detection system is proposed in [49]. The system collects videos from a CCD camera installed on the steering wheel column and tracks facial features. Using a 3D head model, the system estimates the head pose and gaze direction. For night time operation, the system requires an IR illumination. The proposed system does not suffer from the common drawbacks of NIR based systems [121,122,125], because it does not rely on the bright pupil effect. The system works reliably with drivers of different ethnicities wearing different types of glasses. However, if the driver is wearing sunglasses, it is not possible to robustly detect the pupil. Thus, to produce a reliable EOR estimation in this situation, only head pose angles are taken into account.

Cyganek et al. [134] proposed a setup of two cameras operating in the visible and near infra-red spectra for monitoring inattention. In each case (visible and IR) two cascade of classifiers are used. The first one is used for the detection of the eye regions and the other for the verification stage.

Murphy-Chutorian et al. used Local Gradient Orientation (LGO) and Support Vector Regression (SVR) to estimate the driver's continuous yaw and pitch [135]. They used head pose information extracted from a LGO and SVR to recognize drivers' awareness. The algorithm was further developed in [59] by introducing a head tracking module built upon 3D motion estimation and a mesh model of the driver's head. There is a general weakness here as the tracking module may easily diverge from face shapes that are highly different to the given mesh model.

4.4. Driver Distraction Algorithms Based on Gaze Direction

In these previous Sections 4.1–4.3, gaze direction is extracted using different methods. The next step is to detect distraction using gaze direction regardless of the type of method used to extract this information, and hence, is commented as follows.

Many software-based methods have been proposed in order to detect visual distraction, many of which, rely on "course" information extracted from visual cues [114,136–139]. Hattori et al. [136] introduced a FCW system using drivers' behavioural information. Their system determines distraction when it detects that the driver is not looking straight ahead. Following this approximation, an Android app [137] has been developed to detect and alert drivers of dangerous driving conditions and behaviour. Images from the front camera of the mobile phone are scanned to find the relative position of the driver's face. By means of a trained model [38] four face related categories were detected: (1) no face is present; (2) facing forwards, towards the road; (3) facing to the left and (4) facing to the right. Another related system is proposed by Flores et al. [138] where, in order to detect distraction, if the system detects that the face position is not frontal, an alarm cue is issued to alert the driver of a danger situation. Lee et al. [114] proposed a vision-based real-time gaze zone estimator based on a driver's head orientation composed of yaw and pitch. This algorithm is based on normalized histograms of horizontal and vertical edge projections combined with an ellipsoidal face model and a SVM classifier for gaze estimation. In the same research line but in a more elaborated fashion, Yuding et al. [139] used machine vision techniques to monitor the driver's state. The face detection algorithm is based on detection of facial parts. Afterwards, the facial rotation angle is calculated based on the analysis of the driver's head rotation angles. When the angle of facial orientation is not in a reasonable range and lasts for a relatively long time, it can be thought that the driver is distracted and warning information will be provided.

Additionally, other software-based approaches rely on "fine" information considering both head and eye orientation in order to estimate distraction [83,130,140,141]. Pohl et al. [140] focused on estimating the driver's visual distraction level using head pose and eye gaze information with the assumption that the visual distraction level is non-linear: visual distraction increased with time (the driver looked away from the road scene) but nearly instantaneously decreased (the driver re-focused on the road scene). Based on the pose and eye signals, they established their algorithm for visual distraction detection. Firstly, they used a Distraction Calculation (DC) to compute the instantaneous distraction level. Secondly, a Distraction Decision-Maker (DDM) determined whether the current distraction level represented a potentially distracted driver. However, to increase the robustness of the method, also the robustness of the eye and head tracking device to adverse lighting conditions has to be improved.

Bergasa et al. [126] presented a hardware- and software-based approach for monitoring driver vigilance. It is based on a hardware system, for real time acquisition of driver's images using an active IR illuminator and a software implementation for real time pupil tracking, ocular measures and face pose estimation is proposed. Finally, driver's vigilance level is determined from the fusion of the measured parameters into a fuzzy system. The authors yielded an accuracy percentage close to 100% both at night and for users not wearing glasses. However, the performance of the system decreases

during daytime, especially in bright days, and at the moment, the system does not work with drivers wearing glasses [126].

Recently, Lee et al. [141] evaluated four different vision-based algorithms for distraction under different driving conditions. These algorithms were chosen for their ability to distinguish between distracted and non-distracted states using eye-tracking data [141]. The resulting four algorithms, summarized in Table 5, are commented next:

1. Eyes off forward roadway (EOFR) estimates distraction based on the cumulative glances away from the road within a 6-s window [7].
2. Risky Visual Scanning Pattern (RVSP) estimates distraction by combining the current glance and the cumulative glance durations [142].
3. “AttenD” estimates distraction associated with three categories of glances (glances to the forward roadway, glances necessary for safe driving (i.e., at the speedometer or mirrors), and glances not related to driving), and it uses a buffer to represent the amount of road information the driver possesses [143–145].
4. Multi distraction detection (MDD) estimates both visual distraction using the percent of glances to the middle of the road and long glances away from the road, and cognitive distraction by means of the concentration of the gaze on the middle of the road. The implemented algorithm was modified from Victor et al. [146] to include additional sensor inputs (head and seat sensors) and adjust the thresholds for the algorithm variables to improve robustness with potential loss of tracking.

Table 5. AUC comparisons by algorithm across tasks.

Task	Algorithm			
	RVSP	EOFR	AttenD	MDD
Arrows	0.67	0.75	0.71	0.87
Bug	0.78	0.87	0.80	0.86

Considering the results of the ROC curves, AUC values, accuracy and precision, it is apparent that a trade-off exists between ensuring distraction detection and avoiding false alarms, which complicates determining the most promising algorithm. More specifically, the MDD algorithm showed the best performance across all evaluation metrics (accuracy, precision, AUC). Although the EOFR algorithm had promising AUC values, the AttenD algorithm often yielded better accuracy and precision. Additionally, the RVSP algorithm consistently yielded the lowest values for both accuracy and precision, but yielded a slightly higher AUC value than AttenD. All of the algorithms succeeded in detecting distraction well above chance detection ($AUC = 0.5$). The performance of the algorithms varied by task, with little difference in performance for the looking and reaching task (bug) but more stark differences for the looking and touching (arrows). The AUC for each task for each algorithm is provided in Table 5.

5. Cognitive Distraction

Cognitive distraction is a critical area of concern with regard to driver distraction, particularly as related to tasks of listening and conversing, but also, as related to spontaneously occurring processes like daydreaming or becoming lost in thought, which may occur really often on long drives. The term “cognitive load” can be defined as any workload imposed on a driver’s cognitive processes [26]. There are several types (and subtypes) of scenarios where cognitive load may occur during (see Figure 5), and therefore, affect driving. For further information, the reader may refer to [26]. These include:

1. Cognitive load imposed by secondary tasks undertaken while driving.
2. Cognitive load associated with the driver's internal activity.
3. Cognitive load arising from the driving task itself.

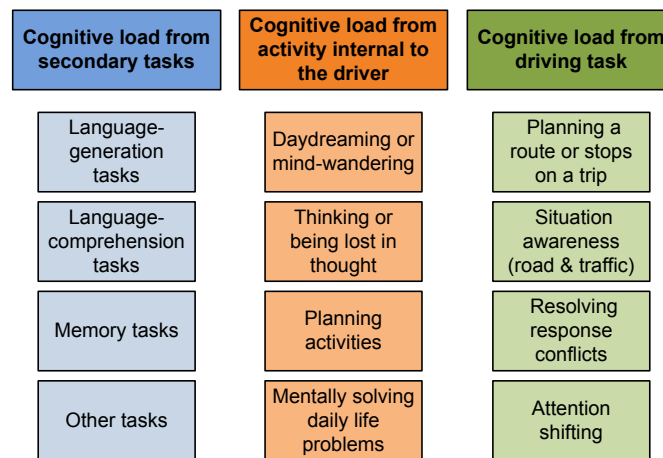


Figure 5. Classification of main types and subtypes of cognitive load while driving.

5.1. Behavioral and Physiological Indicators of Cognitive Load

The research literature documents several types of measures associated with periods of cognitive load. Secondary tasks imposing cognitive load lead to: (1) a high percentage of glances on the forward road and; (2) unusually long glances on the forward road. These two metrics together have been found to be uniquely indicative of cognitive loads [110,147]. Moreover, a narrowing of the spatial extent of scanning [147] is also produced, which is reflected in slightly fewer glances to locations where the mirrors, the speedometer and the areas peripheral to the road centre are located [26,148]. As a result, cognitive load may cause an increasing gaze concentration towards the middle of the road [11].

An eye-gaze pattern could be used to differentiate the action of only driving from driving under the influence of any cognitive task [147]. Drivers under cognitive distraction had fewer saccades per unit time, which was consistent with less exploration of the driving environment [149]. Saccades may be a valuable index of mental workload [150]. In fact, the standard deviations of both eye and head movement could be suitable for detecting cognitive distraction causing gaze concentration and slow saccades when drivers are looking forward [151]. A higher blink rate and a shrink in visual searching range were observed when the driver was cognitively distracted [152]. Kircher et al. [144] indicated the percentage of time the driver spent observing the road ahead, which is called the percentage road center (PRC) of gaze direction, was more than 92% under cognitive distraction in a field study.

Therefore, both glance and blink measures can be used to detect cognitive distraction. For example, He et al. [153] have observed that mind-wandering has effects on glance patterns and blink rates similar to those observed in periods of cognitive secondary task load. During mind-wandering, there is also an increasing concentration of gaze on the forward road with concomitant narrowing of scanning, longer glances on the forward road, and changes in blink rate [26]. Results from [154] suggested that performance data and oculomotor scanning behavior may allow the detection of drivers' mind wandering episodes before they are recognized by the driver himself/herself, potentially providing interventions to detect inattentiveness and alert drivers. Blink rate seems to be a promising indicator of cognitive processing [27]. However, there are measurement issues that may affect how successfully it can be applied in discriminating different types of task loading during driving. For example, there are some questions left about whether it offers sufficient sensitivity when extracted from real world data acquired from a complex task like driving wherein there are inherent temporal variations in driving task load [26,27].

Physiological measures can also be used to detect cognitive load. The average value of pupil diameter is suitably used as a physiological feature for detection of cognitive distraction [155]. When cognitive loads such as arithmetic or conversation were imposed to the subjects, dilation of pupils occurred by acceleration of the sympathetic nerve, resulting in an increase of diameter of pupils [156]. The average value of pupil diameter caused by cognitive loads, such as arithmetic, increased by 13.1% compared with ordinary driving [156]. The tests were performed in a driver simulator in controlled settings. Further experiments are required in a naturalistic setting. Moreover, additional works highlighted the difficulty in estimating cognitive load using pupil diameter during a dialogue task [157] or in different lighting conditions [158].

The same limitation applies to other physiological measures, as Heart Rate (HR), which tends to increase as cognitive task load raises [159]. The traditional method to quantify these physiological measures is by wearing physiological sensors. However, HR measurements can be acquired using computer vision techniques, and consequently, special care has been taken reviewing HR information. Additionally, it is considered a good indicator of fatigue, stress and cognitive load.

By means of the use of HR information the cognitive state of a driver can be monitored [160] in controlled settings. Changes in HR have been noted during certain driving tasks [161]. Similarly, Apparies et al. [162] showed that HR and Heart Rate Variability (HRV) may serve as early indicators of fatigue. In general, HRV specifically measures mental workload, while HR measures physical one [163]. HRV analysis is a strong indicator of mental stress or workload caused by driving tasks [162,164,165]. Experiments carried out in a driving simulator by Zhao et al. [166] found that human heart rates violently fluctuate during a mental stress situation. Ostlund et al. [167] and Miller et al. [165] identified both HR and HRV as promising indicators of the driver's stress level, by increasing HR and decreasing HRV [165,167]. Physiological measures, such as HR and skin conductance level, tend to increase as cognitive task load increases [159].

There are some research works able to extract HR and HRV from face video images in real time from human faces [168–170]. Eulerian Video Magnification framework [171] can be also used to obtain human pulse from a video sequence [172]. In [173], the authors described an approach offering a non-invasive, non-contact means of cardiac monitoring. Once the HRV time series are extracted, feature generation, feature selection and classification should be performed. The conventional method that uses Fast Fourier Transform (FFT) analysis on HRV is 2-min long. In [174], a new method developed by using wavelet-based feature and SVM for classification uses only 1-min HRV signals. Moreover, this method increases accuracy, sensitivity and specificity compared to FFT-based results.

Therefore, concerning cardiovascular measures, they have been reported to be sensitive to mental workload changes and both HR and HRV are widely adopted mental workload measures because they are easy to use and provide fundamental information about the autonomic nervous system [175]. Most methods [168–170] enable low-cost, non-contact cardiovascular activity monitoring using regular RGB cameras by analyzing minute skin color changes caused by periodic blood flow. Nevertheless, for automotive applications, these methods can encounter difficulties under different illumination conditions [176]. In [176], the authors proposed an artifact reduction method, which is caused by lighting variation. Another option is to use an IR-based camera system suitable for automotive applications [177].

To conclude this section, the use of physiological parameters can be used to monitor the cognitive state of the driver. Many of these parameters have been described in controlled settings, but further experiments are required to validate their capability in naturalistic conditions. The main algorithms in this matter are included in the next section.

5.2. Algorithms

Zhang et al. [178] used a decision tree approach to estimate drivers' cognitive workload from eye gaze-related features and driving performance. Liang, Reyes, et al. [179] showed that the SVM models can also detect cognitive distraction. The model's accuracy and sensitivity increased with

window size, suggesting that using longer periods to summarize the data made the distraction signal easier for the models to detect. The conclusion was that the best models were obtained using 40-s window size. Additionally, Liang, Lee, et al. [180] also used Bayesian Network (BN) models and found that they could identify cognitive load reliably for simulator data, and also found that Dynamic Bayesian Networks (DBNs), which considered time dependencies of driver's behaviour, gave a better performance than static BN models. This fact suggests that time-dependent relationship is critical in estimating the cognitive state of the driver. However, to train DBN models, longer training sequences are necessary to obtain more accurate and sensitive models. The results obtained in [180] using BNs, which stated that window size did not affect model performance, clearly conflict with those of Liang et al. [179], which found that larger window sizes improved the detection of cognitive distraction, although another data mining method, SVM, was applied in that study. An additional work from Liang et al. [181] compared SVMs, SBNs, and DBNs in detecting cognitive distraction using the best parameter settings from the same dataset used in the previous two studies [179,180]. DBNs produced the most accurate and sensitive models compared to SBN and SVM. Based on the comparisons of SVMs and BNs, Liang et al. [27,182,183] used a hierarchical layered algorithm, which incorporated both a DBN and a supervised clustering algorithm, to identify feature behaviors when drivers were in different cognitive states. This layered algorithm includes a DBN algorithm at the higher level to model the time-dependent relationship of driver behavior and a supervised clustering algorithm at the lower level to identify feature behaviors. The layered algorithm overcomes the disadvantages of DBNs and significantly improves computational efficiency in training and prediction. Miyaji et al. [184] proposed an approach to detect eye and head movement tracked via standard deviation and categorized features for pattern recognition by using AdaBoost method to detect distraction. The authors compared the performance achieved by both SVM and AdaBoost in estimating cognitive workload, finding that AdaBoost could achieve higher accuracy. Additionally, Miyaji et al. [156] introduced a mixed method by applying a SVM and an AdaBoost classifier for three parameters: (1) heart rate; (2) visual information (standard deviation of both gaze and head rotation angles) and (3) pupil diameter to assess the level of the driver's vigilance. Recently, a new machine learning tool, Extreme Learning Machine (ELM) [185,186]), has gained much attention due to its simple structure, high generalization capability, and fast computational speed. For example, in [187], ELM and SVM were applied to detect drivers' workload using eye movement, as well as eye movement combined with driving performance data. The results suggested that both methods can detect drivers' workload at high accuracy, but ELM outperformed SVM in most cases.

The results of all the works mentioned so far can be summarized in Table 6. Common features include the use of eye gaze-related features, driving performance, pupil diameter features and HR. It should also be noted that very good results can be obtained using only eye gaze-related features. Additionally, many supervised machine learning techniques have been proposed so far: decision trees, SVM, BN, DBN, AdaBoost or ELM.

All these distraction detection systems are based on supervised learning, meaning that the training of such systems need to be "supervised" by human experts by providing a target set for training data containing distraction status. The supervised learning paradigm is only suitable for early stage research and may not be suitable for implementation in real driving cases because of the huge cost and difficulty of creating a target distraction status set, which would require additional subjective ratings by the driver [115], post-processing by the experimentalists [56], or additional computation based on data from other sources [179]. For example, in a recent study [188], labeling drivers' distraction state involves the development of Graphical User Interface (GUI), the training of external evaluators, and the actual labeling time, which is approximately 21.5 h of manpower (43 min per one of the 30 evaluator) to label the entire video segments. For naturalistic driving, where the driver voluntarily decides which tasks to perform at any time, the labeling process can become infeasible. On the other hand, data without known distraction states (unlabeled data) can be collected readily, e.g., from drivers' naturalistic driving records.

Table 6. Supervised algorithms for cognitive distraction detection.

Algorithm	Features	Classifier	Accuracy (%)
Zhang et al. [178]	Eye gaze-related features and driving performance	Decision Tree	81
Zhang et al. [178]	Eye gaze-related features	Decision Tree	80
Zhang et al. [178]	Pupil-diameter features	Decision Tree	61
Zhang et al. [178]	Driving performance	Decision Tree	60
Liang, Reyes, et al. [179]	Eye gaze-related features and driving performance	SVM	83.15
Liang, Reyes, et al. [179]	Eye gaze-related features	SVM	81.38
Liang, Reyes, et al. [179]	driving performance	SVM	54.37
Liang, Lee, et al. [180]	Eye gaze-related features and driving performance data	DBNs	80.1
Miyaji et al. [156]	Heart rate, Eye gaze-related features and pupil diameter	AdaBoost	91.5
Miyaji et al. [156]	Eye gaze-related features	SVM	77.1 (arithmetic task)
Miyaji et al. [156]	Eye gaze-related features	SVM	84.2 (conversation task)
Miyaji et al. [156]	Eye gaze-related features	AdaBoost	81.6 (arithmetic task)
Miyaji et al. [156]	Eye gaze-related features	AdaBoost	86.1 (conversation task)
Yang et al. [187]	Eye gaze-related features and driving performance data	ELM	87.0
Yang et al. [187]	Eye gaze-related features and driving performance data	SVM	82.9

With the purpose of tackle these deficiencies, Unsupervised and Semi-Supervised algorithms can be used. For example, in [12], Semi-Supervised Extreme Learning Machine (SS-ELM) is proposed for drivers' distraction detection. SS-ELM outperformed supervised ELM in both accuracy and model sensitivity, suggesting that the proposed semi-supervised detection system can extract information from unlabeled data effectively to improve the performance. SS-ELM based detection system has the potential of improving accuracy and alleviating the cost of adapting distraction detection systems to new drivers, and thus, more promising for real world applications. However, several points are unclear from these preliminary results [12] further explored in [189], where the Semi-Supervised Learning (SSL) paradigm is introduced to real time detection of distraction based on eye and head movements.

In [189], two graph-based SSL methods were compared with supervised learning methods. These algorithms are detailed as follows. Laplacian Support Vector Machine (LapSVM), which is an extension of SVMs to SSL under manifold regularization framework [190], and SS-ELM were compared with three supervised learning methods (static BN with Supervised Clustering (SBN-SC) [180,183], ELM and SVM) and one low-density-separation-based method (Transductive SVM (TSVM) [191]). To capture realistic eye and head movements patterns, data from an on-road experiment were used. By utilizing unlabeled data, the graph-based semi-supervised methods reduced the labeling cost and improved the detection accuracy. The highest accuracy of 97.2% and G-mean of 0.959 were achieved by SS-ELM. The benefits of using SSL methods increased with the size of unlabeled data set showing that by exploring the data structure without actually labeling them, extra information to improve models performance can be obtained.

It is worth noting that cognitive distraction detection is only performed in "laboratory settings" and not in real conditions. In real life situations, when the driver is under cognitive load (e.g., mind wandering): (1) he is alone and does not interact with anybody; (2) he is also the only one who can decide whether or not to activate the attentional processing of distractive thoughts [192];

and (3) drivers are likely to be performing multiple tasks at the same time (e.g., talking on the mobile phone and listening to music). Moreover, there are two main limitations intrinsic to laboratory-based studies. First of all, most of these studies require that the execution of predefined tasks last for no more than some minutes. In our opinion, such experiments make it very difficult, if not impossible, to infer, for instance, the long-term effectiveness of for example, warning signals, monotonous driving (in general, real driving), on the basis of the results of experiments that are typically so short; And secondly, the drivers are abnormally vigilant to the driving task because they are being observed [193]. In connection with this point, the use of physiological parameters, which form the basis for cognitive distraction detection, have also been extracted in controlled settings and not in real conditions.

6. Mixing Types of Distraction

There are some algorithms trying to detect mixing types of distraction, whose results can be seen in Table 7. In [194], facial features are extracted to detect both visual and cognitive distractions. Binary classifiers (normal vs distracted) are built for visual and cognitive distraction detection. Gaze and Action Units (AU) features are useful in order to detect visual distractions, while AU features are particularly important for cognitive distractions. It should be pointed out that the cognitive tasks considered in this study are closely related to talking activities.

Table 7. Mixing types of distraction detection algorithms.

Algorithm	Features	Classifier	Average Accuracy (%)
Li et al. [194]	AU and head pose	LDC (visual distraction) and SVM (cognitive distraction)	80.8 (LDC), 73.8 (SVM)
Craye et al. [195]	eye behaviour, arm position, head orientation and facial expressions using both color and depth images	Adaboost and HMM	89.84 (Adaboost), 89.64 (HMM)
Liu et al. [196]	Head and eye movements	SVM, ELM and CR-ELM	85.65 (SVM), 85.98 (ELM), 86.95 (CR-ELM)
Ragab et al. [197]	arm position, eye closure, eye gaze, facial expressions and head orientation using depth images	Adaboost, HMM, RF, SVM, CRF, NN	82.9 (RF—type of distraction detection), 90 (RF—distraction detection)

Liu et al. [196] applied Cluster Regularized Extreme Learning Machine (CR-ELM) for detecting mixing types of distraction. Compared with the traditional ELM, CR-ELM introduces an additional regularization term penalizing large covariance of training data within the same clusters in the output space. CR-ELM, ELM and SVM were compared to predict mixing types of distraction. They simulated the mixing types of distraction by combining two types of distracting driving activities (a visual task and a cognitive one). CR-ELM showed lower error rate on most of the 11 subjects (see Table 7).

There are other approaches trying to merge both RGB and depth images to get the features to be used by the algorithms [195,197]. Craye et al. [195] extracted features from face and body using both color and depth images in order to build a distraction system, which is composed of four sub-modules: eye behaviour (gaze and blinking), arm position, head orientation and facial expressions. The information from these modules are merged together using two different classification techniques: Adaboost classifier and Hidden Markov Model (HMM). A set of video sequences was collected to test the system. Five distractive tasks were recorded and manually labelled for training and evaluation. HMM outperforms Adaboost for most drivers. Finally, a vision-based driver distraction is investigated using several machine learning techniques in [197]. IR and Kinect cameras were used in this system, where five visual cues were calculated: arm position, eye closure, eye gaze, facial expressions and

head orientation. These cues were fed into a classifier, such as Adaboost, HMM, RF, SVM, Conditional Random Field (CRF) or NN, in order to detect and recognize the type of distraction.

7. The Relationship between Facial Expressions and Distraction

Facial expressions can be described at different levels [198]. A widely used description is Facial Action Coding System (FACS) [199], which is a human-observer-based system developed to capture subtle changes in facial expressions. With FACS, these expressions are decomposed into one or more AUs [200]. AU recognition and detection have attracted much attention recently [201]. Meanwhile, psychophysical studies indicate that basic emotions have corresponding universal facial expressions across all cultures [202]. This is reflected by most current facial expression recognition systems attempting to recognize a set of prototypic emotional expressions including disgust, fear, joy, surprise, sadness and anger [201], which can be helpful in predicting driving behaviour [203].

Therefore, in this work, main facial expression works in the driving environment are described in accordance with the two aforementioned levels (FACS and prototypic emotional expressions) and how they are related with distraction.

On one hand, in connection with FACS and distraction while driving, the reference work is the one proposed by Li et al. [194]. The authors performed the analysis of driver's facial features under cognitive and visual distractions. In addition to the obvious facial movement associated with secondary tasks such as talking, they hypothesized that facial expression can play an important role in cognitive distraction detection. They studied the top five features (from a total of 186 features) to predict both cognitive and visual distraction. For cognitive distraction, the most important features to consider are: (1) head yaw; (2) Lip Corner Depressor (AU15); (3) Lip Puckerer (AU18); (4) Lip Tightener (AU23) and (5) head roll. For visual distraction, the most important features to consider are: (1) Lip Tightener (AU23); (2) jaw drop (AU26); (3) head yaw; (4) Lip Suck (AU28) and (5) Blink (AU45). The results indicated that gaze and AU features are useful for detecting visual distractions, while AU features are particularly important for cognitive distractions. It should be pointed out that since the cognitive tasks considered in this study are closely related to talking activities, their future work will include the analysis of other cognitive tasks (e.g., thinking or solving math problems).

On the other hand, in connection with prototypic emotional expressions, there are some works trying to study how these emotions affect behaviour.

The relationship between emotion and cognition is complex, but it is widely accepted that human performance is altered when a person is in any emotional state. It is really important to fully understand the impact of emotion on driving performance because, for example, roadways are lined with billboard advertisements and messages containing a lot of different emotional information. Moreover, the distracting effects of emotions may come in other forms, such as cell phone, passenger conversations, radio information or texting information [204]. For example, Chan et al. [204] conducted a study to examine the potential for distraction from emotional information presented on roadside billboards. The findings in this study showed that emotional distraction: (a) can seriously modulate attention and decision-making abilities and have adverse impacts on driving behavior for several reasons and (b) can impact driving performance by reorienting attention away from the primary driving task towards the emotional content and negatively influence the decision-making process. In another study with a similar line of work, Chan et al. [205] showed that emotion-related auditory distraction can modulate attention to differentially influence driving performance. Specifically, negative distractions reduced lateral control and slowed driving speeds compared to positive and neutral distractions.

Some studies have shown that drivers who are more likely to become angry (e.g., those with high trait anger rates) tend to engage in more aggressive behavior on the road, which can result in negative outcomes such as crashes [206]. Moreover, anger negatively influences several driving performance and risky behaviors such as infractions, lane deviations, speed, and collisions [207].

In conclusion, aggressiveness and anger are emotional states that extremely influence driving behaviour and increase the risk of accident. However, a too low level of activation (e.g., resulting from emotional states like sadness) also leads to reduced attention and distraction as well as prolonged reaction time and, therefore, lowers driving performance [208]. On this basis, research and experience have demonstrated that being in a good mood is the best precondition for safe driving and that happy drivers produce fewer accidents [209]. In other words, *happy drivers are better drivers* [208,210]. Facial expression and emotion recognition can be used in advanced car safety systems, which, on one hand, can identify hazardous emotional drivers' states that can lead to distraction and, on the other, can provide tailored (according to each state and associated hazards) suggestions and warnings to the driver [211].

8. Sensors

Once the algorithms for distraction detection have been designed and implemented, the next step is to port them to an embedded device or system to be executed inside the car. However, porting a vision-based algorithm is not a straightforward step and some key factors should be taken into account. Furthermore, there is not a standard implementation platform, so different alternatives have been proposed by both the scientific community and the industry.

8.1. Porting a Vision Algorithm to an Embedded Automotive System

The implementation of computer vision applications in automotive environments is not straightforward because several requirements must be taken into account: reliability [212,213], real-time performance [213–215], low-cost [216–219], spatial constraints [217,219], low power consumption [220], flexibility [219], rapid prototyping [215,221], design requirements [217] and short time to market [217]. Therefore, there must be a trade-off among these design requisites [217]. Moreover, there is not a commonly accepted hardware and software platform, so different solutions have been proposed by the industry and the scientific community. Last but not least, some driver distraction guidelines and test procedures for all applications to be used while driving should be considered [222], and so should ADAs.

One approach can rely on the use of microprocessors, which incorporates the functions of a computer's central processing unit (CPU) on a single integrated circuit (IC). For example, in [223], a vision-based system for monitoring the loss of attention, tested under day and night driving conditions, is proposed. The algorithm was cross-validated using brain signals and finally, implemented on a Single Board Computer (SBC). Another example is presented in [224], where a vehicle was equipped with a USB camera connected to the system in order to track the driver's eyes for fatigue detection.

A similar approach is the use of digital signal processors (DSPs) [225], which can perform multiplications and additions in a single cycle and have parallel processing capabilities. DSPs have been used in image and audio signal processing when the use of microcontrollers was not enough. These processors were used in [215], where an optimized vision library approach for embedded systems was presented. VLIB is a software library that accelerates computer vision applications for high-performance embedded systems. By significantly speeding up pixel-intensive operations, the library provides more headroom for innovative algorithms, and enables processing of more channels at higher resolutions. Authors optimized the library for the Texas Instruments C64x/C64x+ DSP cores. Karuppusamy et al. [226] proposed an embedded implementation of facial landmarks detection based on both Viola-Jones face detector and facial landmarks detection using extended Active Shape Model (ASM) [227]. However, DSPs imply a much higher cost compared with other options such as field-programmable gate arrays (FPGAs) [228].

Another option is to use hardware implementation, since it can achieve a much better computational performance, where two types are commonly used namely, FPGA and ASIC. A FPGA is an integrated circuit designed to be configured by a customer or a designer after manufacture.

FPGAs take advantage of high speed operations, especially for parallelizable operations achieving good performance in face monitoring applications [229–232]. For example, several well-known algorithms have been used and optimized for this field of application, such as: (a) spatial and temporal filtering, motion detection and optical flow analysis [229] or (b) gray scale projection, edge detection with Prewitt operator and complexity functions [230]. Additionally, the use of Application-Specific Integrated Circuits (ASIC), which is an IC customized for a particular use rather than intended for general-purpose use, has also been considered [233]. FPGAs have an important advantage over ASICs: they are reconfigurable, which gives them some of the flexibility of software. ASICs are only used for high volume manufacturing and long series due to higher initial engineering cost.

Developing the whole application in hardware is a cumbersome task, so hybrid solutions have appeared combining both software and hardware implementations. The work in [234] describes a System on a Chip (SOC) designed to support a family of vision algorithms. However, this system uses an ASIC, so it cannot be completely reconfigured. This important drawback makes impossible to update the device. A generic embedded hardware and software architecture was proposed to design and evaluate ADAS vision applications [221]. Although the system is useful to test some vision applications, the performance obtained in the case study showed that the system is not powerful enough to run more complex applications including I/Os management, vehicle communications or more demanding vision applications. In [219], a reconfigurable embedded vision system reaching the requirements of ADAS applications is presented. A SOC, which is formed by an FPGA with a dual core ARM, is prepared to be easily reconfigured. A lane departure warning system was implemented in the case study obtaining a good computational performance. The obtained computational time allows the system to include another more complex vision algorithm running in parallel. In [235], they proposed an approach to predict performances of image processing algorithms on different computing units of a given heterogeneous SOC.

Despite the fact that in recent years some authors have been trying to propose some architectures in order to achieve some key factors in embedded ADAS systems inside a vehicle [219,221,235], these efforts do not seem to be sufficient to reach the requirements stated before. The use of microprocessors in embedded computer vision-based systems has experienced a significant growth in recent years. Moreover, current premium cars implement more than 90 Electronic Control Units (ECU) with close to 1 Gigabyte embedded software code [236]. In 2018, 30% of the overall vehicle cost is predicted to stem from vehicle electronics [237]. The independence of different applications (with different criticality levels) running on the same platform must be made evident. Therefore, the development of embedded automotive systems has become quite complex. To that end, the use of standards and frameworks is indispensable.

8.2. Commercial Sensors

8.2.1. Smart Eye

The Smart Eye [77] system is a well-suited head and gaze tracking method for the demanding environment of a vehicle cabin and flexible to cope with most research projects. It consists of a multi-camera system running on a single PC and on a single algorithm. The system is scalable from 2 up to 8 cameras allowing 360° head and eye tracking. A typical configuration inside a vehicle cabin is composed of four cameras with two IR lightings, located on the dashboard on either side of the steering wheel. Smart Eye offers a sampling rate of 60 Hz (up to 8 cameras) or 120 Hz (up to 4 cameras). The field of view, depending on the number of cameras, is in the range of 90°–360°. The data output includes over 145 values covering, among others, gaze, eyelid, pupilometry and head tracking, raw and filtered gaze, blinks, fixations and saccades. Smart Eye has been used in several driver assistance and inattention systems, such as [76,143–145,238].

8.2.2. EyeAlert

EyeAlert [239], cited in several publications [128,240], has been conceived to detect driver inattention using computer vision and to generate a warning signal in case of dangerous situation. The EyeAlert system focuses entirely on the driver's alertness levels or inattention to the road ahead, regardless of the time of the day or the weather conditions. Three models are available:

- EyeAlert EA410 detects both distracted and fatigue driving. The EA410 has a highly integrated IR camera, a computer, an image processing unit and an alarm. The EA410 technology is protected by over ten patents. The system will also respond in case the driver does not focus on driving.
- EyeAlert EA430 with GPS detects both distracted and fatigue driving. Moreover, a minimum speed threshold is programmed into the internal GPS to prevent false alarms in urban environments.
- EyeAlert EA450 with Data detects both distracted and fatigue driving. Additionally, minimum speed threshold, sensitivity, volume and data can be remotely programmed. The minimum speed and sensitivity controls allow the reduction of false alarms in highway and urban environments.

8.2.3. Seeing Machines

Seeing Machines [241] builds image-processing technology that tracks the movement of a person's eyes, face, head, and facial expressions. It monitors fatigue and distraction events in real-time and uses IR technology to provide fatigue and distraction monitoring at any time of the day. The system can also combine multiple camera sensors to detect a wider range of movements. The Seeing Machines' system continuously measures operator eye and eyelid behaviour to determine the onset of fatigue and micro sleeps and delivers real-time detection and alerts. The system has been used in many different driver assistance and inattention systems [11,142,151,156,242–244].

8.2.4. Visage Technologies AB

Visage Technologies AB [245] provides a state-of-the-art commercial head tracker based on feature-point detection and tracking of the nose boundary and eye regions. Visage SDK finds and tracks the face and facial features, including gaze direction, in video sequences in real time. It provides pupil coordinates, 3D gaze direction as well as (with a calibration step) screen-space gaze point. Visage Technologies also features support for embedded systems like FPGA and IR light tracking for poor lighting conditions.

8.2.5. Delphi Electronics Driver Status Monitor

Delphi Electronics, a major automotive electronics manufacturer, developed a single camera Driver Status Monitor (DSM) [246]. By detecting and tracking the driver's facial features, the system analyzes eye-closures and head pose to infer his/her fatigue or distraction. This information is used to warn the driver and to modulate the actions of other safety systems. The system includes the use of NIR illumination, an embedded processing unit, as well as the camera (resolution of 640×480 pixels).

8.2.6. Tobii Technologies

Tobii Technologies develops Tobii's eye-tracking technology for integration into volume products such as computers, computer games, virtual reality and cars. The Tobii platform consists of two-camera sensors, placed at different angles, and operating at IR frequencies to eliminate interference from external light. The system can distinguish whether the driver's eyes are open or closed or if the driver has turned his/her head. The sensors work even when the driver is wearing glasses or sunglasses. By observing the specifics of eyelid closure, in combination with eye gaze patterns, an active safety system powered by Tobii's eye tracking sensor can reliably detect if a driver is falling asleep and warns him/her properly. Moreover, Tobii Technologies provides the Tobii EyeChip, which is a dedicated eye tracking SOC ASIC.

8.2.7. SensoMotoric Instruments

SensoMotoric Instruments GmbH (SMI) [247] is a German company, whose eye tracking solutions can measure head position and orientation, gaze direction, eyelid opening, and pupil position and diameter. Eye trackers use a sampling rate of 120 Hz for head pose and gaze measurement, 120 Hz for eyelid closure and blink measurement, and 60 Hz for combined gaze, head pose, and eyelid measurement. It also provides PERCLOS information for drowsiness detection. It is a computer-based system and needs user calibration. In [248], SensoMotoric was used to recognize the pupil in each image in order to measure horizontal and vertical eye movements.

8.2.8. Automobile Manufacturers

Nissan introduces its new Driver Attention Alert system with the 2016 Nissan Maxima, which was unveiled at the New York International Auto Show [249]. The automaker has highlighted that the new system will be able to detect drowsy and inattentive driving and it will alert the driver about the situation by giving visual and audible warning. Ford's Driver Alert [250] seems only to detect drowsiness but not distraction. The Driver Alert system comprises a small forward-facing camera connected to an on-board computer. The camera is mounted on the back of the rear view mirror and is trained to identify lane markings on both sides of the vehicle. When the vehicle is on the move, the computer looks at the road ahead and predicts where the car should be positioned relative to the lane markings. Then, it measures where the vehicle actually is and, if the difference is significant, the system issues a warning. The Saab Driver Attention Warning System [251] detects visual inattention and drowsy driving. The system uses two miniature IR cameras integrated with Smart Eye technology [77] to accurately estimate head pose, gaze, and eyelid status. When a driver's gaze is not located inside the primary attention zone (which covers the central part of the frontal windshield) for a predefined period, an alarm is triggered. Toyota has equipped their luxury Lexus models with their Driver Monitoring System [252]. The system permanently monitors the movement of the driver's head when looking from side to side using a NIR camera installed at the top of the steering wheel column. The system is integrated into Toyota's pre-crash system, which warns the driver when a collision is likely to happen. In 2007, Volvo Cars introduced Driver Alert Control to alert tired and non-concentrating drivers [253,254]. Based on the idea that the technology for monitoring eyes is not yet sufficiently mature and human behavior varies from one person to another, Volvo Cars developed the system based on car progress on the road. It is reported that Driver Alert Control monitors the car movements and assesses whether the vehicle is driven in a controlled or uncontrolled way. More recently, a Hyundai concept car (the Hyundai HCD-14) incorporates Tobbi Technologies to track the eyes [255].

9. Simulated vs. Real Environment to Test and Train Driving Monitoring Systems

The development of the computer vision algorithm only represents one part of all the cycle of the product design. One of the hardest tasks is to validate the whole system with the wide variety of driving scenarios [256]. In order to complete the whole "process development" of the vision-based ADAS, some key points are presented.

In order to monitor both the driver and his/her driving behaviour, several hardware and software algorithms are being developed, but they are tested mostly in simulated environments instead of in real driving ones. This is due to the danger of testing inattention in real driving environments [21]. Experimental control, efficiency, safety, and ease of data collection are the main advantages of using simulators [257,258]. Some researches have validated that driving simulators can create driving environment relatively similar to road experiments [259–261]. However, some considerations should be taken into account since simulators can produce inconsistent, contradictory and conflicting results. For example, low-fidelity simulators may evoke unrealistic driving behavior and, therefore, produce invalid research outcomes. One common issue is that real danger and the real consequences

of actions do not occur in a driving simulator, giving rise to a false sense of safety, responsibility, or competence [262]. Moreover, simulator sickness symptoms may undermine training effectiveness and negatively affect the usability of simulators [262].

A study on distraction in both simulated and real environment was conducted in [11] and it was found out that the driver's physiological activity showed significant difference. Engstorm et al. [11] stated that physiological workload and steering activity were both higher under real driving conditions compared to simulated environments. In [257], the authors compared the impact of a narrower lane using both a simulator and real data, showing that the speed was higher in the simulated roads, consistent with other studies. In [263], controlled driving yielded more frequent and longer eye glances than the simulated driving setting, while driving errors were more common in simulated driving. In [167], the driver's heart rate changed significantly while performing the visual task in real-world driving relative to a baseline condition, suggesting that visual task performance in real driving was more stressful.

After the system is properly validated in a driver simulator, it should be validated in real conditions as well, because various factors including light variations and noise can also affect the driver's attention. The application on a real moving vehicle presents new challenges like changing backgrounds and sudden variations of lighting [264]. Moreover, a useful system should guarantee real time performance and quick adaptability to a variable set of users and to natural movements performed during driving [264]. Thus, it is necessary to make simulated environments appear more realistic [203].

To conclude, in most previous studies, independent evaluations using different equipment and conditions (mainly simulated environments) resulted in time-consuming and redundant efforts. Moreover, inconsistency in the algorithm performance metrics makes it difficult to compare algorithms. Hence, the only way to compare most algorithms and systems is the metrics provided by each author when comparing their values, but with scarce information about the used images and conditions. Public data sets covering simulated and real driving environments should be released in the near future, as stated by some authors previously [203].

10. Privacy Issues Related to Camera Sensors

Although there is a widespread agreement for intelligent vehicles to improve safety, the study of driver behaviour to design and evaluate intelligent vehicles requires large amounts of naturalistic driving data [265]. However, in current literature, there is a lack of publicly available naturalistic driving data largely due to concerns over individual privacy. It also should be noted that a real-time visual-based distraction detection system does not have to save the video stream. Therefore, privacy issues are mostly relevant in research works where video feed is collected and stored to be studied at a later stage, for example in the large naturalistic studies conducted in the US.

Typical protection of the individuals' privacy in a video sequence is commonly referred as "de-identification" [266]. Although this fact will help protect the identities of individual drivers, it impedes the purpose of sensorizing vehicles to control both drivers and their behaviour. In an ideal situation, a de-identification algorithm would protect the identity of drivers while preserving sufficient details to infer their behaviour (e.g., eye gaze, head pose or hand activity) [265].

Martin et al. [265,267] proposed the use of de-identification filters to protect the privacy of drivers while preserving sufficient details to infer their behaviour. Following this idea, a de-identification filter preserving only the mouth region can be used for monitoring yawning or talking and a de-identification filter preserving eye regions can be used for detecting fatigue or gaze direction, which is precisely proposed by Martin et al. [265,267]. More specifically, the authors implemented and compared de-identification filters made up of a combination of preserving eye regions for fine gaze estimation, superimposing head pose encoded face masks for providing spatial context and replacing background with black pixels for ensuring privacy protection. A two-part study revealed that human facial

recognition experiment had a success rate well below the chance while gaze zone estimation accuracy disclosed 65%, 71% and 85% for One-Eye, Two-Eyes and Mask with Two-Eyes, respectively.

Fernando et al. [268] proposed to use video de-identification in the automobile environment using personalized Facial Action Transfer (FAT), which has recently attracted a lot of attention in computer vision due to its diverse applications in the movie industry, computer games, and privacy protection. The goal of FAT is to “clone” the facial actions from the videos of a person (source) to another one (target) following a two-step approach. In the first step, their method transfers the shape of the source person to the target subject using the triangle-based deformation transfer method. In the second step, it generates the appearance of the target person using a personalized mapping from shape changes to appearance changes. In this approach video de-identification is used to pursue two objectives: (1) to remove person-specific facial features and (2) to preserve head pose, gaze and facial expression.

11. General Discussion and Challenges Ahead

The main visual-based approaches reviewed in this paper are summarized in Table 8 according to some key factors.

A major finding emerging from two recent research works reveals that just-driving baselines may, in fact, not be “just driving” [26,269], containing a considerable amount of cognitive activity in the form of daydreaming and lost-in-thought activity. Moreover, eye-gaze patterns are somewhat idiosyncratic when visual scanning is disrupted by cognitive workload [27]. Additionally, “look-but-failed-to-see” impairment under cognitive workload is an obvious detriment to traffic safety. For example, Strayer et al. [270] found that recognition memory for objects in the driving environment was reduced by 50% when the driver was talking on a handsfree cell phone, inducing failures of visual attention during driving. Indeed, visual, manual and cognitive distraction often occur simultaneously while driving (e.g., texting while driving and other cell-phone reading and writing activities). Therefore, the estimates of crash risk based on comparisons of activities to just-driving baselines may need to be reconsidered in light of the possible finding that just-driving baselines may contain the aforementioned frequent cognitive activity. As a result, for example, secondary tasks effects while driving should be revised [269]. Accordingly, as detecting driver distraction depends on how distraction changes his/her behavior compared to normal driving without distraction, periods with minimal or no cognitive activity should be identified in order to train the distraction detection algorithms.

Additionally, computer vision techniques can be used, not only for extracting information inside the car, but also for extracting information outside the car, such as traffic, road hazards, external conditions of the road ahead, intersections, or even position regarding other cars. The final step should be the correlation between the driver’s understanding and the traffic context. One of the first works trying to fuse “out” information (visual lane analysis) and “in” information (driver monitoring) is the one proposed by Apostoloff et al. [271], pointing out the benefits of this approach. Indeed, visual lane analysis can be used for “higher-order tasks”, which are defined by interacting with other modules in a complex driver assistance system (e.g., understanding the driver’s attentiveness—distraction—to the lane-keeping task [272]). Hirayama et al [273] focused on temporal relationships between the driver’s eye gaze and the peripheral vehicles behaviour. In particular, they concluded that the timing when a driver gazes towards the overtaking event under cognitive distraction is later than that under the neutral state. Therefore, they showed that the temporal factor, that is, timing, of a reaction is important for understanding the state by focusing on cognitive distraction in a car-driving situation. Additionally, Rezaei et al. [87] proposed a system correlating the driver’s head pose to road hazards (vehicle detection and distance estimation) by analyzing both simultaneously. Ohn et al. [274] proposed a framework for early detection of driving maneuvers using cues from the driver, the environment and the vehicle. Tawari et al. [275] provided early detection of driver distraction by continuously monitoring driver and surround traffic situation. Martin et al. [276] focused on intersections and studied the interaction of head, eyes and hands as the driver approaches

a stop-controlled intersection. In this line work of research, Jain et al. [277] deal with the problem of anticipating driving maneuvers a few seconds before the driver performs them.

Table 8. Summary of visual-based approaches to detect different types of driver distraction.

Approach	Distraction Detection Approaches			Real Conditions	Operation	
	Manual	Visual	Cognitive		Daytime	Nighttime
Zhao et al. [88]	✓	✗	✗	✗	✓	✗
Zhao et al. [89]	✓	✗	✗	✗	✓	✗
Zhao et al. [90]	✓	✗	✗	✗	✓	✗
Zhao et al. [91]	✓	✗	✗	✗	✓	✗
Bosch et al. [94]	✓	✗	✗	✗	✓	✗
Lowe et al. [95]	✓	✗	✗	✗	✓	✗
Yan et al. [92]	✓	✗	✗	✗	✓	✗
Yan et al. [93]	✓	✗	✗	✗	✓	✗
Zhang et al. [96]	✓	✗	✗	✗	✓	✗
Artan et al. [97]	✓	✗	✗	✓	✓	✓
Berri et al. [98]	✓	✗	✗	✓	✓	✗
Xu et al. [99]	✓	✗	✗	✓	✓	✓
Seshadri et al. [100]	✓	✗	✗	✓	✓	✗
Ohn et al. [106]	✓	✗	✗	✓	✓	✗
Martin et al. [104]	✓	✗	✗	✓	✓	✗
Ohn et al. [105]	✓	✗	✗	✓	✓	✗
Morimoto et al. [120]	✗	✓	✗	✓	✗	✓
Ji et al. [121]	✗	✓	✗	✓	✗	✓
Ji et al. [122]	✗	✓	✗	✓	✗	✓
Ji et al. [123]	✗	✓	✗	✓	✗	✓
Gu et al. [124]	✗	✓	✗	✓	✗	✓
Batista et al. [125]	✗	✓	✗	✓	✗	✓
Bergasa et al. [126]	✗	✓	✗	✓	✓	✓
Lee et al. [114]	✗	✓	✗	✓	✓	✓
Vicente et al. [49]	✗	✓	✗	✓	✓	✓
Cyganek et al. [134]	✗	✓	✗	✓	✓	✓
Donmez et al. [142]	✗	✓	✗	✗	✓	✗
Klauer et al. [7]	✗	✓	✗	✓	✓	✓
Kircher et al. [143]	✗	✓	✗	✓	✓	✓
Kircher et al. [144]	✗	✓	✗	✓	✓	✓
Kircher et al. [145]	✗	✓	✗	✓	✓	✓
Victor et al. [146]	✗	✓	✗	✓	✓	✗
Zhang et al. [178]	✗	✗	✓	✗	✓	✗
Liang et al. [179]	✗	✗	✓	✗	✓	✗
Liang et al. [180]	✗	✗	✓	✗	✓	✗
Liang et al. [181]	✗	✗	✓	✗	✓	✗
Liang et al. [27]	✗	✗	✓	✗	✓	✗
Liang et al. [182]	✗	✗	✓	✗	✓	✗
Liang et al. [183]	✗	✗	✓	✗	✓	✗
Miyaji et al. [184]	✗	✗	✓	✗	✓	✗
Miyaji et al. [156]	✗	✗	✓	✗	✓	✗
Yang et al. [187]	✗	✗	✓	✗	✓	✗
Li et al. [194]	✗	✓	✓	✗	✓	✗
Craye et al. [195]	✓	✓	✗	✗	✓	✗
Liu et al. [196]	✗	✓	✓	✗	✓	✗
Ragab et al. [197]	✓	✓	✗	✗	✓	✗

There are many factors that can modulate distraction. For example, as discussed in Section 7, emotional information can modulate attention and decision-making abilities. Additionally, numerous studies link highly aroused stress states with impaired decision-making capabilities [278], decreased situational awareness [279], and degraded performance, which could impair driving ability [280]. Another driver state, often responsible for traffic violations and even road accidents that can lead to distraction,

is confusion or irritation, as it is related to loss of self-control and, therefore, loss of vehicle control, which can be provoked by non-intuitive user interfaces or defective navigation systems as well as by complex traffic conditions, mistakable signs and complicated routing. Moreover, the amount of information that needs to be processed simultaneously during driving is a source of confusion especially for older people [281], who have slower perception and reaction times. Just like stress, confusion or irritation leads to impairment of driving capabilities including driver's perception, attention, decision making, and strategic planning. Nervousness corresponds to a level of arousal above the "normal" one, which best suits to the driving task [211]. It is an affective state with negative impact both on decision-making process and strategic planning. Nervousness can be induced by a variety of reasons either directly related to the driving task like novice drivers or by other factors like personal/physical conditions [211].

The system should be validated, firstly, in a driver simulator and afterwards, in real conditions, where various factors including variations in lighting and noise can also affect both the driver's attention and the performance of the developed algorithms. Therefore, public data sets covering simulated and real driving environments should be released. The driver's physiological responses could be different in a driver simulator from those in real conditions [11,167,257,263]. Hence, while developing an inattention detection system, the simulated environment must be a perfect replica of the real environment. However, they are normally used in research and simulated scenarios, but not in real ones, due to the problems of vision systems working in outdoor environments (lighting changes, sudden movements, etc.). Moreover, they do not work properly with users wearing glasses and may need high computational requirements.

Data-driven applications will require large amount of labeled images for both training and testing the system. Both manual data reduction and labeling of data are time-consuming and they are also subject to interpretation of the reductionist. Therefore, to deal with this problem, two approaches are emerging from the literature: (1) unsupervised or semi-supervised learning and (2) automatic data reduction. For example, in connection with the first approach, Liu et al. [189] commented the benefits of SSL methods. Specifically, they explained the benefits of using SSL increased with the size of unlabeled data set showing that by exploring the data structure without actually labeling them, extra information to improve models performance can be obtained. On the other hand, there has been a hype in data reduction using vehicle dynamics and looking outside on large scale naturalistic driving data [282–284], and looking in at the driver [285].

In many distraction detection systems, the use of commercial sensors is usually performed [77,239,241,245–247]. We understand that the reason from this is twofold: these systems are well-established solutions offering both head and gaze tracking in the car environment and the efforts of the investigation can be focused to detect and predict distraction from the outputs from these commercial sensors instead of developing a new sensor from the very beginning. These commercial sensors can operate using one camera [239,245–247], two cameras [241] or even up to 8 cameras [77] placed all over the vehicle cabin. What we find missing is some research works trying to compare these commercial sensors in order to highlight the pros and cons of each one. Also, missing from the literature is the comparison between a new sensor and a commercial one trying to offer a competitive solution from the sake of the research community.

Author Contributions: Alberto Fernández, Rubén Usamentiaga, Juan Luis Carús have made the review of the already existing methods and have written the paper as well. Alberto Fernández, Rubén Usamentiaga and Rubén Casado have designed paper concept and structure and have revised its content.

Conflicts of Interest: The authors declare no conflict of interest.

References

- World Health Organization. Global Status Report on Road Safety 2015. Available online: http://apps.who.int/iris/bitstream/10665/189242/1/9789241565066_eng.pdf?ua=1 (accessed on 2 July 2016).
- Llerena, L.E.; Aronow, K.V.; Macleod, J.; Bard, M.; Salzman, S.; Greene, W.; Haider, A.; Schupper, A. An evidence-based review: Distracted driver. *J. Trauma Acute Care Surg.* **2015**, *78*, 147–152.
- Carsten, O.; Merat, N. Protective or not? (visual distraction). In Proceedings of the 2015 4th International Conference on Driver Distraction and Inattention, Sydney, Australia, 9–11 November 2015.
- Dingus, T.A.; Klauer, S.; Neale, V.; Petersen, A.; Lee, S.; Sudweeks, J.; Perez, M.; Hankey, J.; Ramsey, D.; Gupta, S.; et al. *The 100-Car Naturalistic Driving Study, Phase II—Results of the 100-Car Field Experiment*; National Highway Traffic Safety Administration: Washington, DC, USA, 2006.
- Ranney, T.A.; Mazzae, E.; Garrott, R.; Goodman, M.J. NHTSA driver distraction research: Past, present, and future. In *Driver Distraction Internet Forum*; National Highway Traffic Safety Administration: Washington, DC, USA, 2000; Volume 2000.
- Klauer, S.G.; Neale, V.L.; Dingus, T.A.; Ramsey, D.; Sudweeks, J. Driver inattention: A contributing factor to crashes and near-crashes. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Orlando, FL, USA, 26–30 September 2005; SAGE Publications: Thousand Oaks, CA, USA; Volume 49, pp. 1922–1926.
- Klauer, S.G.; Dingus, T.A.; Neale, V.L.; Sudweeks, J.D.; Ramsey, D.J. *The Impact of Driver Inattention on Near-Crash/crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data*; National Highway Traffic Safety Administration: Washington, DC, USA, 2006.
- Talbot, R.; Fagerlind, H. Exploring inattention and distraction in the SafetyNet accident causation database. *Accid. Anal. Prev.* **2009**, *60*, 445–455.
- Bennakhi, A.; Safar, M. Ambient Technology in Vehicles: The Benefits and Risks. *Procedia Comput. Sci.* **2016**, *83*, 1056–1063.
- Lee, J.D. Driving safety. In *Review of Human Factors*; Nickerson, R.S., Ed.; Human Factors and Ergonomics Society: San Francisco, CA, USA, 2006; pp. 172–218.
- Engström, J.; Johansson, E.; Östlund, J. Effects of visual and cognitive load in real and simulated motorway driving. *Transp. Res. Part F Traffic Psychol. Behav.* **2005**, *8*, 97–120.
- Liu, T.; Yang, Y.; Huang, G.B.; Lin, Z. Detection of Drivers' Distraction Using Semi-Supervised Extreme Learning Machine. In *Proceedings of ELM-2014*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 2, pp. 379–387.
- Simons-Morton, B.G.; Guo, F.; Klauer, S.G.; Ehsani, J.P.; Pradhan, A.K. Keep your eyes on the road: Young driver crash risk increases according to duration of distraction. *J. Adolesc. Health* **2014**, *54*, S61–S67.
- Dingus, T.A.; Neale, V.L.; Klauer, S.G.; Petersen, A.D.; Carroll, R.J. The development of a naturalistic data collection system to perform critical incident analysis: An investigation of safety and fatigue issues in long-haul trucking. *Accid. Anal. Prev.* **2006**, *38*, 1127–1136.
- Harbluk, J.L.; Noy, Y.I.; Trbovich, P.L.; Eizenman, M. An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. *Accid. Anal. Prev.* **2007**, *39*, 372–379.
- Recarte, M.A.; Nunes, L.M. Mental workload while driving: Effects on visual search, discrimination, and decision making. *J. Exp. Psychol. Appl.* **2003**, *9*, 119–137.
- Klauer, S.G.; Guo, F.; Simons-Morton, B.G.; Ouimet, M.C.; Lee, S.E.; Dingus, T.A. Distracted driving and risk of road crashes among novice and experienced drivers. *N. Engl. J. Med.* **2014**, *370*, 54–59.
- Bergmark, R.W.; Gliklich, E.; Guo, R.; Gliklich, R.E. Texting while driving: The development and validation of the distracted driving survey and risk score among young adults. *Inj. Epidemiol.* **2016**, *3*, doi:10.1186/s40621-016-0073-8.
- Administration, N.H.T.S. Traffic Safety Facts: Distracted Driving 2014. Available online: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812260> (accessed on 2 July 2016).
- Carmona, J.; García, F.; Martín, D.; Escalera, A.D.L.; Armingol, J.M. Data fusion for driver behaviour analysis. *Sensors* **2015**, *15*, 25968–25991.
- Sahayadhas, A.; Sundaraj, K.; Murugappan, M. Detecting driver drowsiness based on sensors: A review. *Sensors* **2012**, *12*, 16937–16953.

22. Daza, I.G.; Bergasa, L.M.; Bronte, S.; Yebes, J.J.; Almazán, J.; Arroyo, R. Fusion of optimized indicators from Advanced Driver Assistance Systems (ADAS) for driver drowsiness detection. *Sensors* **2014**, *14*, 1106–1131.
23. Hoff, J.; Grell, J.; Lohrman, N.; Stehly, C.; Stoltzfus, J.; Wainwright, G.; Hoff, W.S. Distracted driving and implications for injury prevention in adults. *J. Trauma Nurs.* **2013**, *20*, 31–34.
24. Stothart, C.; Mitchum, A.; Yehnert, C. The attentional cost of receiving a cell phone notification. *J. Exp. Psychol. Hum. Percept. Perform.* **2015**, *41*, 893–897.
25. Kushlev, K.; Proulx, J.; Dunn, E.W. Silence Your Phones: Smartphone Notifications Increase Inattention and Hyperactivity Symptoms. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 1011–1020.
26. Angell, L.; Perez, M.; Soccolich, S. *Identification of Cognitive Load in Naturalistic Driving*; Virginia Tech Transportation Institute: Blacksburg, VA, USA, 2015.
27. Liang, Y. Detecting Driver Distraction. Ph.D. Thesis, University of Iowa, Iowa City, IA, USA, May 2009.
28. Caird, J.K.; Johnston, K.A.; Willness, C.R.; Asbridge, M.; Steel, P. A meta-analysis of the effects of texting on driving. *Accid. Anal. Prev.* **2014**, *71*, 311–318.
29. Ohn-Bar, E.; Trivedi, M.M. Looking at Humans in the Age of Self-Driving and Highly Automated Vehicles. *IEEE Trans. Intell. Veh.* **2016**, *1*, 90–104.
30. Toole, L.M. Crash Risk and Mobile Device Use Based on Fatigue and Drowsiness Factors in Truck Drivers. Ph.D. Thesis, Virginia Tech, Blacksburg, VA, USA, October 2013.
31. Lee, J.D.; Young, K.L.; Regan, M.A. *Defining Driver Distraction*; CRC Press: Boca Raton, FL, USA, 2008.
32. Foley, J.P.; Young, R.; Angell, L.; Domeyer, J.E. Towards operationalizing driver distraction. In Proceedings of the 7th International Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Bolton Landing, NY, USA, 17–20 June 2013; pp. 57–63.
33. Young, R.A. A Tabulation of Driver Distraction Definitions. Available online: <https://www.yumpu.com/en/document/view/22871401/a-tabulation-of-driver-distraction-definitions-toyota/5> (accessed on 27 October 2016).
34. Young, R. Cognitive distraction while driving: A critical review of definitions and prevalence in crashes. *SAE Int. J. Passeng. Cars Electron. Electr. Syst.* **2012**, *5*, 326–342.
35. Regan, M.A.; Hallett, C.; Gordon, C.P. Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accid. Anal. Prev.* **2011**, *43*, 1771–1781.
36. Engström, J.; Monk, C.; Hanowski, R.; Horrey, W.; Lee, J.; McGehee, D.; Regan, M.; Stevens, A.; Traube, E.; Tuukkanen, M.; et al. *A Conceptual Framework and Taxonomy for Understanding and Categorizing Driver Inattention*; European Commission: Brussels, Belgium, 2013.
37. Lee, J.D. Dynamics of driver distraction: The process of engaging and disengaging. *Ann. Adv. Autom. Med.* **2014**, *58*, 24–32.
38. Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, USA, 8–14 December 2001; Volume 1.
39. Zafeiriou, S.; Zhang, C.; Zhang, Z. A Survey on Face Detection in the wild: Past, present and future. *Comput. Vis. Image Underst.* **2015**, *138*, 1–24.
40. Viola, P.; Jones, M.J. Robust real-time face detection. *Int. J. Comput. Vis.* **2004**, *57*, 137–154.
41. Jensen, O.H. Implementing the Viola-Jones Face Detection Algorithm. Ph.D. Thesis, Technical University of Denmark, Lyngby, Denmark, September 2008.
42. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), Miami, FL, USA, 20–25 June 2005; Volume 1, pp. 886–893.
43. Dlib. Make Your Own Object Detector. Available online: <http://blog.dlib.net/2014/02/dlib-186-released-make-your-own-object.html> (accessed on 25 November 2015).
44. Ranjan, R.; Patel, V.M.; Chellappa, R. A Deep Pyramid Deformable Part Model for Face Detection. 2015, arXiv preprint arXiv:1508.04389.
45. Wang, Q.; Wu, J.; Long, C.; Li, B. P-FAD: Real-time face detection scheme on embedded smart cameras. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2013**, *3*, 210–222.

46. Markuš, N.; Frljak, M.; Pandžić, I.S.; Ahlberg, J.; Forchheimer, R. A Method for Object Detection Based on Pixel Intensity Comparisons Organized in Decision Trees. 2013, arXiv preprint arXiv:1305.4537
47. Reese, K.; Zheng, Y.; Elmaghraby, A. A comparison of face detection algorithms in visible and thermal spectrums. In Proceedings of the International Conference on Advances in Computer Science and Application, New Delhi, India, 25–27 May 2012.
48. Abtahi, S.; Omidyeganeh, M.; Shirmohammadi, S.; Hariri, B. YawDD: A yawning detection dataset. In Proceedings of the 5th ACM Multimedia Systems Conference, Singapore, 19–27 March 2014; pp. 24–28.
49. Vicente, F.; Huang, Z.; Xiong, X.; de la Torre, F.; Zhang, W.; Levi, D. Driver Gaze Tracking and Eyes Off the Road Detection System. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 2014–2027.
50. Murphy-Chutorian, E.; Trivedi, M.M. Head pose estimation in computer vision: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 607–626.
51. Tawari, A.; Martin, S.; Trivedi, M.M. Continuous Head Movement Estimator for Driver Assistance: Issues, Algorithms, and On-Road Evaluations. *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 818–830.
52. Zhu, Z.; Ji, Q. Real time and non-intrusive driver fatigue monitoring. In Proceedings of the 2004 7th International IEEE Conference on Intelligent Transportation Systems, Washington, DC, USA, 3–6 October 2004; pp. 657–662.
53. Garcia-Mateos, G.; Ruiz, A.; Lopez-de Teruel, P.E.; Rodriguez, A.L.; Fernandez, L. Estimating 3D facial pose in video with just three points. In Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'08), Anchorage, AK, USA, 23–28 June 2008; pp. 1–8.
54. Kaminski, J.Y.; Knaan, D.; Shavit, A. Single image face orientation and gaze detection. *Mach. Vis. Appl.* **2009**, *21*, 85–98.
55. Martin, S.; Tawari, A.; Murphy-Chutorian, E.; Cheng, S.Y.; Trivedi, M. On the design and evaluation of robust head pose for visual user interfaces: Algorithms, databases, and comparisons. In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Portsmouth, NH, USA, 17–19 October 2012; pp. 149–154.
56. Oyini Mbouna, R.; Kong, S.G.; Chun, M.G. Visual analysis of eye state and head pose for driver alertness monitoring. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 1462–1469.
57. Xiong, X.; de la Torre, F. Supervised descent method and its applications to face alignment. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Portland, OR, USA, 25–27 June 2013; pp. 532–539.
58. Asthana, A.; Marks, T.K.; Jones, M.J.; Tieu, K.H.; Rohith, M. Fully automatic pose-invariant face recognition via 3D pose normalization. In Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain, 6–13 November 2011; pp. 937–944.
59. Murphy-Chutorian, E.; Trivedi, M.M. Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness. *IEEE Trans. Intell. Transp. Syst.* **2010**, *11*, 300–311.
60. Peláez, G.; de la Escalera, A.; Armingol, J. Head Pose Estimation Based on 2D and 3D Information. *Phys. Procedia* **2014**, *22*, 420–427.
61. Tawari, A.; Trivedi, M.M. Robust and continuous estimation of driver gaze zone by dynamic analysis of multiple face videos. In Proceedings of the 2014 IEEE Intelligent Vehicles Symposium Proceedings, Ypsilanti, MI, USA, 8–11 June 2014; pp. 344–349.
62. Ohn-Bar, E.; Tawari, A.; Martin, S.; Trivedi, M.M. Predicting driver maneuvers by learning holistic features. In Proceedings of the 2014 IEEE Intelligent Vehicles Symposium Proceedings, Ypsilanti, MI, USA, 8–11 June 2014; pp. 719–724.
63. Ohn-Bar, E.; Tawari, A.; Martin, S.; Trivedi, M.M. On surveillance for safety critical events: In-vehicle video networks for predictive driver assistance systems. *Comput. Vis. Image Underst.* **2015**, *134*, 130–140.
64. La Cascia, M.; Sclaroff, S.; Athitsos, V. Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3D models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 322–336.
65. Zhu, X.; Ramanan, D. Face detection, pose estimation, and landmark localization in the wild. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, USA, 16–21 June 2012; pp. 2879–2886.

66. Uricár, M.; Franc, V.; Hlavác, V. Facial Landmark Tracking by Tree-based Deformable Part Model Based Detector. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Santiago, Chile, 11–18 December 2015; pp. 10–17.
67. Uricár, M.; Franc, V.; Thomas, D.; Sugimoto, A.; Hlavác, V. Real-time multi-view facial landmark detector learned by the structured output SVM. In Proceedings of the 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Ljubljana, Slovenia, 4–8 May 2015; Volume 2, pp. 1–8.
68. Cao, X.; Wei, Y.; Wen, F.; Sun, J. Face alignment by explicit shape regression. *Int. J. Comput. Vis.* **2014**, *107*, 177–190.
69. Ren, S.; Cao, X.; Wei, Y.; Sun, J. Face alignment at 3000 fps via regressing local binary features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1685–1692.
70. Wang, N.; Gao, X.; Tao, D.; Li, X. Facial feature point detection: A comprehensive survey. 2014, arXiv preprint arXiv:1410.1037
71. Rezaei, M.; Klette, R. Adaptive Haar-like classifier for eye status detection under non-ideal lighting conditions. In Proceedings of the 27th Conference on Image and Vision Computing New Zealand, Dunedin, New Zealand, 26–28 November 2012; pp. 521–526.
72. Song, F.; Tan, X.; Chen, S.; Zhou, Z.H. A literature survey on robust and efficient eye localization in real-life scenarios. *Pattern Recognit.* **2013**, *46*, 3157–3173.
73. Kircher, K.; Ahlström, C. *Evaluation of Methods for the Assessment of Minimum Required Attention*; Swedish National Road and Transport Research Institute (VTI): Linköping, Sweden, 2015.
74. Victor, T. Keeping Eye and Mind on the Road. Ph.D. Thesis, Uppsala University, Uppsala, Sweden, 2005.
75. Flannagan, C.A.; Bao, S.; Klinich, K.D. *Driver Distraction From Cell Phone Use and Potential for Self-Limiting Behavior*; University of Michigan: Ann Arbor, MI, USA, 2012.
76. Ahlstrom, C.; Dukic, T. Comparison of eye tracking systems with one and three cameras. In Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research, Eindhoven, The Netherlands, 24–27 August 2010.
77. SmartEye. Smart Eye Pro. Available online: <http://smarteye.se/products/smart-eye-pro/> (accessed on 3 January 2016).
78. Sigari, M.H.; Pourshahabi, M.R.; Soryani, M.; Fathy, M. A Review on Driver Face Monitoring Systems for Fatigue and Distraction Detection. *Int. J. Adv. Sci. Technol.* **2014**, *64*, 73–100.
79. Zhao, S.; Grigat, R.R. Robust eye detection under active infrared illumination. In Proceedings of the 2006 IEEE 18th International Conference on Pattern Recognition (ICPR 2006), Washington, DC, USA, 20–24 August 2006; Volume 4, pp. 481–484.
80. Zhao, S.; Grigat, R.R. An automatic face recognition system in the near infrared spectrum. In *Machine Learning and Data Mining in Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 437–444.
81. Smith, P.; Shah, M.; Lobo, N.D.V. Determining driver visual attention with one camera. *IEEE Trans. Intell. Transp. Syst.* **2003**, *4*, 205–218.
82. Sigari, M.H. Driver hypo-vigilance detection based on eyelid behavior. In Proceedings of the IEEE 7th International Conference on Advances in Pattern Recognition (ICAPR'09), Kolkata, India, 4–6 February 2009; pp. 426–429.
83. Batista, J. A drowsiness and point of attention monitoring system for driver vigilance. In Proceedings of the 2007 IEEE Intelligent Transportation Systems Conference (ITSC 2007), 30 September–3 October 2007; pp. 702–708.
84. Flores, M.J.; Armingol, J.M.; de la Escalera, A. Driver drowsiness warning system using visual information for both diurnal and nocturnal illumination conditions. *EURASIP J. Adv. Signal Process.* **2010**, *2010*, 438205.
85. Rezaei, M.; Klette, R. Simultaneous analysis of driver behaviour and road condition for driver distraction detection. *Int. J. Image Data Fusion* **2011**, *2*, 217–236.
86. Craye, C.; Karray, F. Multi-distributions particle filter for eye tracking inside a vehicle. In *Image Analysis and Recognition*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 407–416.
87. Rezaei, M.; Klette, R. Look at the driver, look at the road: No distraction! No accident! In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 129–136.

88. Zhao, C.; Zhang, B.; He, J.; Lian, J. Recognition of driving postures by contourlet transform and random forests. *Intell. Transp. Syst. IET* **2012**, *6*, 161–168.
89. Zhao, C.; Gao, Y.; He, J.; Lian, J. Recognition of driving postures by multiwavelet transform and multilayer perceptron classifier. *Eng. Appl. Artif. Intell.* **2012**, *25*, 1677–1686.
90. Zhao, C.H.; Zhang, B.L.; Zhang, X.Z.; Zhao, S.Q.; Li, H.X. Recognition of driving postures by combined features and random subspace ensemble of multilayer perceptron classifiers. *Neural Comput. Appl.* **2013**, *22*, 175–184.
91. Zhao, C.; Zhang, B.; He, J. Vision-based classification of driving postures by efficient feature extraction and bayesian approach. *J. Intell. Robot. Syst.* **2013**, *72*, 483–495.
92. Yan, C.; Coenen, F.; Zhang, B. Driving posture recognition by joint application of motion history image and pyramid histogram of oriented gradients. *Int. J. Veh. Technol.* **2014**, *2014*, 719413.
93. Yan, C.; Coenen, F.; Zhang, B. Driving posture recognition by convolutional neural networks. *IET Comput. Vis.* **2016**, *10*, 103–114.
94. Bosch, A.; Zisserman, A.; Munoz, X. Representing shape with a spatial pyramid kernel. In Proceedings of the 6th ACM International Conference on Image and Video Retrieval, Hyderabad, India, 6–12 January 2007; pp. 401–408.
95. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.
96. Zhang, X.; Zheng, N.; Wang, F.; He, Y. Visual recognition of driver hand-held cell phone use based on hidden CRF. In Proceedings of the 2011 IEEE International Conference on Vehicular Electronics and Safety (ICVES), Beijing, China, 10–12 July 2011; pp. 248–251.
97. Artan, Y.; Bulan, O.; Loce, R.P.; Paul, P. Driver cell phone usage detection from HOV/HOT NIR images. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Columbus, OH, USA, 23–28 June 2014; pp. 225–230.
98. Berri, R.A.; Silva, A.G.; Parpinelli, R.S.; Girardi, E.; Arthur, R. A Pattern Recognition System for Detecting Use of Mobile Phones While Driving. 2014, arXiv preprint arXiv:1408.0680.
99. Xu, B.; Loce, R.P. A machine learning approach for detecting cell phone usage. In Proceedings of the IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics, San Francisco, CA, USA, 4 March 2015.
100. Seshadri, K.; Juefei-Xu, F.; Pal, D.K.; Savvides, M.; Thor, C.P. Driver Cell Phone Usage Detection on Strategic Highway Research Program (SHRP2) Face View Videos. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Boston, MA, USA, 7–12 June 2015.
101. Li, Y.; Zhou, G.; Li, Y.; Shen, D. Determining driver phone use leveraging smartphone sensors. *Multimed. Tools Appl.* **2015**, *1*, doi:10.1007/s11042-015-2969-7.
102. Ohn-Bar, E.; Trivedi, M. In-vehicle hand activity recognition using integration of regions. In Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IV), Gold Coast City, Australia, 23–26 June 2013; pp. 1034–1039.
103. Ohn-Bar, E.; Martin, S.; Trivedi, M.M. Driver hand activity analysis in naturalistic driving studies: Challenges, algorithms, and experimental studies. *J. Electron. Imaging* **2013**, *22*, 041119.
104. Martin, S.; Ohn-Bar, E.; Tawari, A.; Trivedi, M.M. Understanding head and hand activities and coordination in naturalistic driving videos. In Proceedings of the 2014 IEEE Intelligent Vehicles Symposium Proceedings, Ypsilanti, MI, USA, 8–11 June 2014; pp. 884–889.
105. Ohn-Bar, E.; Martin, S.; Tawari, A.; Trivedi, M. Head, eye, and hand patterns for driver activity recognition. In Proceedings of the IEEE 2014 22nd International Conference on Pattern Recognition (ICPR), Stockholm, Sweden, 24–28 August 2014; pp. 660–665.
106. Ohn-Bar, E.; Trivedi, M.M. The power is in your hands: 3D analysis of hand gestures in naturalistic video. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Portland, OR, USA, 23–28 June 2013; pp. 912–917.
107. Lee, S.R.; Bambach, S.; Crandall, D.J.; Franchak, J.M.; Yu, C. This hand is my hand: A probabilistic approach to hand disambiguation in egocentric video. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Columbus, OH, USA, 23–28 June 2014; pp. 557–564.
108. Drews, F.A.; Yazdani, H.; Godfrey, C.N.; Cooper, J.M.; Strayer, D.L. Text messaging during simulated driving. *Hum. Factors J. Hum. Factors Ergon. Soc.* **2009**, *51*, 762–770.

109. Tsimhoni, O.; Arbor, A. *Time-Sharing of a Visual in-Vehicle Task While Driving: Findings from the Task Occlusion Method*; University of Michigan, Transportation Research Institute: Ann Arbor, MI, USA, 2003.
110. Victor, T.W.; Harbluk, J.L.; Engström, J.A. Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transp. Res. Part F Traffic Psychol. Behav.* **2005**, *8*, 167–190.
111. Donmez, B.; Boyle, L.N.; Lee, J.D. Safety implications of providing real-time feedback to distracted drivers. *Accid. Anal. Prev.* **2007**, *39*, 581–590.
112. Klauer, S.G.; Guo, F.; Sudweeks, J.; Dingus, T.A. *An Analysis of Driver Inattention Using a Case-Crossover Approach on 100-Car Data: Final Report*; National Highway Traffic Safety Administration: Washington, DC, USA, 2010.
113. Ahlstrom, C.; Kircher, K.; Kircher, A. A gaze-based driver distraction warning system and its effect on visual behavior. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 965–973.
114. Lee, S.J.; Jo, J.; Jung, H.G.; Park, K.R.; Kim, J. Real-time gaze estimator based on driver's head orientation for forward collision warning system. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 254–267.
115. Wollmer, M.; Blaschke, C.; Schindl, T.; Schuller, B.; Farber, B.; Mayer, S.; Trefflich, B. Online driver distraction detection using long short-term memory. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 574–582.
116. Neale, V.L.; Dingus, T.A.; Klauer, S.G.; Sudweeks, J.; Goodman, M. *An Overview of the 100-Car Naturalistic Study and Findings*; National Highway Traffic Safety Administration: Washington, DC, USA, 2005.
117. Boyraz, P.; Yang, X.; Hansen, J.H. Computer vision systems for context-aware active vehicle safety and driver assistance. In *Digital Signal Processing for in-Vehicle Systems and Safety*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 217–227.
118. Hammoud, R.I.; Wilhelm, A.; Malawey, P.; Witt, G.J. Efficient real-time algorithms for eye state and head pose tracking in advanced driver support systems. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), San Diego, CA, USA, 20–25 June 2005; Volume 2.
119. Tivesten, E.; Dozza, M. Driving context and visual-manual phone tasks influence glance behavior in naturalistic driving. *Transp. Res. Part F Traffic Psychol. Behav.* **2014**, *26*, 258–272.
120. Morimoto, C.H.; Koons, D.; Amir, A.; Flickner, M. Pupil detection and tracking using multiple light sources. *Image Vis. Comput.* **2000**, *18*, 331–335.
121. Ji, Q.; Yang, X. Real time visual cues extraction for monitoring driver vigilance. In *Computer Vision Systems*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 107–124.
122. Ji, Q.; Yang, X. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real Time Imaging* **2002**, *8*, 357–377.
123. Ji, Q. 3D Face pose estimation and tracking from a monocular camera. *Image Vis. Comput.* **2002**, *20*, 499–511.
124. Gu, H.; Ji, Q.; Zhu, Z. Active facial tracking for fatigue detection. In Proceedings of the 6th IEEE Workshop on Applications of Computer Vision (WACV 2002), Orlando, FL, USA, 3–4 December 2002; pp. 137–142.
125. Batista, J.P. A real-time driver visual attention monitoring system. In *Pattern Recognition and Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 200–208.
126. Bergasa, L.M.; Nuevo, J.; Sotelo, M.; Barea, R.; Lopez, M.E. Real-time system for monitoring driver vigilance. *IEEE Trans. Intell. Transp. Syst.* **2006**, *7*, 63–77.
127. Hansen, D.W.; Ji, Q. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 478–500.
128. D'Orazio, T.; Leo, M.; Guaragnella, C.; Distanto, A. A visual approach for driver inattention detection. *Pattern Recognit.* **2007**, *40*, 2341–2355.
129. Nuevo, J.; Bergasa, L.M.; Sotelo, M.; Ocaña, M. Real-time robust face tracking for driver monitoring. In Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC'06), Toronto, ON, Canada, 17–20 September 2006; pp. 1346–1351.
130. Bergasa, L.M.; Buenaposada, J.M.; Nuevo, J.; Jimenez, P.; Baumela, L. Analysing driver's attention level using computer vision. In Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems (ITSC 2008), Beijing, China, 12–15 October 2008; pp. 1149–1154.
131. Flores, M.J.; Armingol, J.M.; De la Escalera, A. Driver drowsiness detection system under infrared illumination for an intelligent vehicle. *IET Intell. Transp. Syst.* **2011**, *5*, 241–251.
132. Jimenez-Pinto, J.; Torres-Torriti, M. Face salient points and eyes tracking for robust drowsiness detection. *Robotica* **2012**, *30*, 731–741.

133. Jiménez-Pinto, J.; Torres-Torriti, M. Optical flow and driver's kinematics analysis for state of alert sensing. *Sensors* **2013**, *13*, 4225–4257.
134. Cyganek, B.; Gruszczyński, S. Hybrid computer vision system for drivers' eye recognition and fatigue monitoring. *Neurocomputing* **2014**, *126*, 78–94.
135. Murphy-Chutorian, E.; Doshi, A.; Trivedi, M.M. Head pose estimation for driver assistance systems: A robust algorithm and experimental evaluation. In Proceedings of the 2007 IEEE Intelligent Transportation Systems Conference (ITSC 2007), Bellevue, WA, USA, 30 September–3 October 2007; pp. 709–714.
136. Hattori, A.; Tokoro, S.; Miyashita, M.; Tanaka, I.; Ohue, K.; Uozumi, S. Development of Forward Collision Warning System Using the Driver Behavioral Information. Available online: <http://papers.sae.org/2006-01-1462/> (accessed on 26 October 2016).
137. You, C.W.; Lane, N.D.; Chen, F.; Wang, R.; Chen, Z.; Bao, T.J.; Montes-de Oca, M.; Cheng, Y.; Lin, M.; Torresani, L.; et al. Carsafe app: Alerting drowsy and distracted drivers using dual cameras on smartphones. In Proceedings of the 11th Annual International Conference on Mobile Systems, Applications, and Services, Taipei, Taiwan, 25–28 June 2013; pp. 13–26.
138. Flores, M.J.; Armingol, J.M.; de la Escalera, A. Real-time warning system for driver drowsiness detection using visual information. *J. Intell. Robot. Syst.* **2010**, *59*, 103–125.
139. Yuying, J.; Yazhen, W.; Haitao, X. A surveillance method for driver's fatigue and distraction based on machine vision. In Proceedings of the 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE), Changchun, China, 16–18 December 2011; pp. 727–730.
140. Pohl, J.; Birk, W.; Westervall, L. A driver-distraction-based lane-keeping assistance system. *Proc. Inst. Mech. Eng. Part I J. Syst. Control Eng.* **2007**, *221*, 541–552.
141. Lee, J.; Moeckli, J.; Brown, T.; Roberts, S.; Victor, T.; Marshall, D.; Schwarz, C.; Nadler, E. Detection of driver distraction using vision-based algorithms. In Proceedings of the 23rd International Conference on Enhanced Safety of Vehicles, Seoul, Korea, 27–30 May 2013.
142. Donmez, B.; Boyle, L.N.; Lee, J.D. Mitigating driver distraction with retrospective and concurrent feedback. *Accid. Anal. Prev.* **2008**, *40*, 776–786.
143. Kircher, K.; Kircher, A.; Ahlström, C. *Results of a Field Study on a Driver Distraction Warning System*; Swedish National Road and Transport Research Institute (VTI): Linköping, Sweden, 2009.
144. Kircher, K.; Ahlstrom, C.; Kircher, A. Comparison of two eye-gaze based real-time driver distraction detection algorithms in a small-scale field operational test. In Proceedings of the 5th International Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Big Sky, MT, USA, 22–25 June 2009; pp. 16–23.
145. Kircher, K.; Kircher, A.; Claezon, F. *Distraction and Drowsiness—A Field Study*; Swedish National Road and Transport Research Institute (VTI): Linköping, Sweden, 2009.
146. Victor, T. The Victor and Larsson (2010) distraction detection algorithm and warning strategy. *Volvo Technol.* **2010**, *1*, 0–6.
147. Angell, L.S.; Auflick, J.; Austria, P.; Kochhar, D.S.; Tijerina, L.; Biever, W.; Diptiman, T.; Hogsett, J.; Kiger, S. *Driver Workload Metrics Task 2 Final Report*; National Highway Traffic Safety Administration: Washington, DC, USA, 2006.
148. Recarte, M.A.; Nunes, L.M. Effects of verbal and spatial-imagery tasks on eye fixations while driving. *J. Exp. Psychol. Appl.* **2000**, *6*, 31.
149. Harbluk, J.L.; Noy, Y.I.; Eizenman, M. *The Impact of Cognitive Distraction on Driver Visual Behaviour and Vehicle Control*; Transport Canada: Ottawa, ON, Canada, 2002.
150. May, J.G.; Kennedy, R.S.; Williams, M.C.; Dunlap, W.P.; Brannan, J.R. Eye movement indices of mental workload. *Acta Psychol.* **1990**, *75*, 75–89.
151. Miyaji, M.; Kawanaka, H.; Oguri, K. Driver's cognitive distraction detection using physiological features by the adaboost. In Proceedings of the IEEE 12th International IEEE Conference on Intelligent Transportation Systems (ITSC'09), St. Louis, MO, USA, 4–7 October 2009; pp. 1–6.
152. Yang, Y. The Effects of Increased Workload on Driving Performance and Visual Behaviour. Ph.D. Thesis, University of Southampton, Southampton, UK, December 2011.
153. He, J.; Becic, E.; Lee, Y.C.; McCarley, J.S. Mind wandering behind the wheel performance and oculomotor correlates. *Hum. Factors J. Hum. Factors Ergon. Soc.* **2011**, *53*, 13–21.

154. He, J. Identify Mind-Wandering Behind the Wheel. Ph.D. Thesis, University of Illinois at Urbana-Champaign, Champaign, IL, USA, December 2010.
155. Kahneman, D.; Tursky, B.; Shapiro, D.; Crider, A. Pupillary, heart rate, and skin resistance changes during a mental task. *J. Exp. Psychol.* **1969**, *79*, 164–167.
156. Miyaji, M.; Kawanaka, H.; Oguri, K. Effect of pattern recognition features on detection for driver's cognitive distraction. In Proceedings of the 2010 13th International IEEE Conference on Intelligent Transportation Systems (ITSC), Funchal, Portugal, 19–22 September 2010; pp. 605–610.
157. Heeman, P.A.; Meshorer, T.; Kun, A.L.; Palinko, O.; Medenica, Z. Estimating cognitive load using pupil diameter during a spoken dialogue task. In Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Eindhoven, The Netherlands, 27–30 October 2013; pp. 242–245.
158. Pfleging, B.; Fekety, D.K.; Schmidt, A.; Kun, A.L. A Model Relating Pupil Diameter to Mental Workload and Lighting Conditions. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 5776–5788.
159. Mehler, B.; Reimer, B.; Coughlin, J.F. Physiological reactivity to graded levels of cognitive workload across three age groups: An on-road evaluation. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, San Francisco, CA, USA, 27 September–1 October 2010; SAGE Publications: Thousand Oaks, CA, USA; Volume 54, pp. 2062–2066.
160. Partin, D.L.; Sultan, M.F.; Thrush, C.M.; Prieto, R.; Wagner, S.J. Monitoring Driver Physiological Parameters for Improved Safety. Available online: <http://papers.sae.org/2006-01-1322/> (accessed on 26 October 2016).
161. Liu, S.H.; Lin, C.T.; Chao, W.H. The short-time fractal scaling of heart rate variability to estimate the mental stress of driver. In Proceedings of the 2004 IEEE International Conference on Networking, Sensing and Control, Taipei, Taiwan, 21–23 March 2004; Volume 2, pp. 829–833.
162. Apparies, R.J.; Riniolo, T.C.; Porges, S.W. A psychophysiological investigation of the effects of driving longer-combination vehicles. *Ergonomics* **1998**, *41*, 581–592.
163. Wickens, C.D.; Lee, J.D.; Liu, Y.; Gordon-Becker, S. *Introduction to Human Factors Engineering*; Pearson Education: Upper Saddle River, NJ, USA, 1998.
164. Mulder, L. Measurement and analysis methods of heart rate and respiration for use in applied environments. *Biol. Psychol.* **1992**, *34*, 205–236.
165. Miller, E.E. Effects of Roadway on Driver Stress: An On-Road Study using Physiological Measures. Ph.D. Thesis, University of Washington, Seattle, WA, USA, July 2013.
166. Zhao, C.; Zhao, M.; Liu, J.; Zheng, C. Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. *Accid. Anal. Prev.* **2012**, *45*, 83–90.
167. Östlund, J.; Nilsson, L.; Törnros, J.; Forsman, Å. Effects of cognitive and visual load in real and simulated driving. *Transp. Res. Part F Traffic Psychol. Behav.* **2006**, *8*, 97–120.
168. Poh, M.Z.; McDuff, D.J.; Picard, R.W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Opt. Express* **2010**, *18*, 10762–10774.
169. Poh, M.Z.; McDuff, D.J.; Picard, R.W. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Trans. Biomed. Eng.* **2011**, *58*, 7–11.
170. Fernandez, A.; Carus, J.L.; Usamentiaga, R.; Alvarez, E.; Casado, R. Unobtrusive health monitoring system using video-based physiological information and activity measurements. In Proceedings of the IEEE 2015 International Conference on Computer, Information and Telecommunication Systems (CITS), Paris, France, 11–13 June 2015; pp. 1–5.
171. Wu, H.Y.; Rubinstein, M.; Shih, E.; Guttag, J.V.; Durand, F.; Freeman, W.T. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.* **2012**, *31*, 65.
172. Chambino, P. Android-Based Implementation of Eulerian Video Magnification for Vital Signs Monitoring. Ph.D. Thesis, Faculdade de Engenharia da Universidade do Porto, Porto, Portugal, July 2013.
173. Balakrishnan, G.; Durand, F.; Guttag, J. Detecting pulse from head motions in video. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Washington, DC, USA, 23–28 June 2013; pp. 3430–3437.
174. Li, G.; Chung, W.Y. Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier. *Sensors* **2013**, *13*, 16494–16511.

175. Solovey, E.T.; Zec, M.; Garcia Perez, E.A.; Reimer, B.; Mehler, B. Classifying driver workload using physiological and driving performance data: Two field studies. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, ON, Canada, 26 April–1 May 2014; pp. 4057–4066.
176. Lee, D.; Kim, J.; Kwon, S.; Park, K. Heart rate estimation from facial photoplethysmography during dynamic illuminance changes. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milano, Italy, 25–29 August 2015; pp. 2758–2761.
177. Jeanne, V.; Asselman, M.; den Brinker, B.; Bulut, M. Camera-based heart rate monitoring in highly dynamic light conditions. In Proceedings of the 2013 IEEE International Conference on Connected Vehicles and Expo (ICCVE), Las Vegas, NV, USA, 2–6 December 2013; pp. 798–799.
178. Zhang, Y.; Owechko, Y.; Zhang, J. Driver cognitive workload estimation: A data-driven perspective. In Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems, Washington, DC, USA, 3–6 October 2004; pp. 642–647.
179. Liang, Y.; Reyes, M.L.; Lee, J.D. Real-time detection of driver cognitive distraction using support vector machines. *IEEE Trans. Intell. Transp. Syst.* **2007**, *8*, 340–350.
180. Liang, Y.; Lee, J.D.; Reyes, M.L. Nonintrusive detection of driver cognitive distraction in real time using Bayesian networks. *Transp. Res. Rec. J. Transp. Res. Board* **2007**, *2018*, 1–8.
181. Liang, Y.; Lee, J.D. Comparing Support Vector Machines (SVMs) and Bayesian Networks (BNs) in detecting driver cognitive distraction using eye movements. In *Passive Eye Monitoring: Algorithms, Applications and Experiments*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 285–300.
182. Liang, Y.; Lee, J.D. Using a Layered Algorithm to Detect Driver Cognitive Distraction. In Proceedings of the seventh International Driving Symposium on Human Factors in Driver assessment, Training, and Vehicle Design, New York, NY, USA, 17–20 June 2013; pp. 327–333.
183. Liang, Y.; Lee, J.D. A hybrid Bayesian Network approach to detect driver cognitive distraction. *Transp. Res. Part C Emerg. Technol.* **2014**, *38*, 146–155.
184. Miyaji, M.; Danno, M.; Kawanaka, H.; Oguri, K. Driver's cognitive distraction detection using AdaBoost on pattern recognition basis. In Proceedings of the IEEE International Conference on Vehicular Electronics and Safety (ICVES 2008), Columbus Ohio, OH, USA, 22–24 September 2008; pp. 51–56.
185. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501.
186. Huang, G.B.; Zhou, H.; Ding, X.; Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2012**, *42*, 513–529.
187. Yang, Y.; Sun, H.; Liu, T.; Huang, G.B.; Sourina, O. Driver Workload Detection in On-Road Driving Environment Using Machine Learning. In *Proceedings of ELM-2014*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 2, pp. 389–398.
188. Li, N.; Busso, C. Predicting Perceived Visual and Cognitive Distractions of Drivers With Multimodal Features. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 51–65.
189. Liu, T.; Yang, Y.; Huang, G.B.; Yeo, Y.; Lin, Z. Driver Distraction Detection Using Semi-Supervised Machine Learning. *IEEE Trans. Intell. Transp. Syst.* **2015**, *17*, 1–13.
190. Belkin, M.; Niyogi, P.; Sindhvani, V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *J. Mach. Learn. Res.* **2006**, *7*, 2399–2434.
191. Joachims, T. Transductive inference for text classification using support vector machines. In Proceedings of the Sixteenth International Conference on Machine Learning (ICML), Bled, Slovenia, 27–30 June 1999; Volume 99, pp. 200–209.
192. Lemercier, C.; Pêcher, C.; Berthié, G.; Valéry, B.; Vidal, V.; Paubel, P.V.; Cour, M.; Fort, A.; Galéra, C.; Gabaude, C.; et al. Inattention behind the wheel: How factual internal thoughts impact attentional control while driving. *Saf. Sci.* **2014**, *62*, 279–285.
193. Ho, C.; Gray, R.; Spence, C. To what extent do the findings of laboratory-based spatial attention research apply to the real-world setting of driving? *IEEE Trans. Hum. Mach. Syst.* **2014**, *44*, 524–530.
194. Li, N.; Busso, C. Analysis of facial features of drivers under cognitive and visual distractions. In Proceedings of the 2013 IEEE International Conference on Multimedia and Expo (ICME), San Jose, CA, USA; 15–19 July 2013; pp. 1–6.
195. Craye, C.; Karray, F. Driver distraction detection and recognition using RGB-D sensor. 2015, arXiv preprint arXiv:1502.00250.

196. Liu, T.; Yang, Y.; Huang, G.B.; Lin, Z.; Klanner, F.; Denk, C.; Rasshofer, R.H. Cluster Regularized Extreme Learning Machine for Detecting Mixed-Type Distraction in Driving. In Proceedings of the 2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC), Las Palmas de Gran Canaria, Spain, 15–18 September 2015; pp. 1323–1326.
197. Ragab, A.; Craye, C.; Kamel, M.S.; Karray, F. A Visual-Based Driver Distraction Recognition and Detection Using Random Forest. In *Image Analysis and Recognition*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 256–265.
198. Tian, Y.L.; Kanade, T.; Cohn, J.F. Handbook of face recognition. In *Ch Facial Expression Analysis*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 487–519.
199. Ekman, P.; Friesen, W. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*; Consulting Psychologists: San Francisco, CA, USA, 1978.
200. Tian, Y.; Kanade, T.; Cohn, J.F. Facial expression recognition. In *Handbook of Face Recognition*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 487–519.
201. Shan, C.; Gong, S.; McOwan, P.W. Facial expression recognition based on local binary patterns: A comprehensive study. *Image Vis. Comput.* **2009**, *27*, 803–816.
202. Ekman, P.; Friesen, W.V.; Press, C.P. *Pictures of Facial Affect*; Consulting Psychologists Press: Palo Alto, CA, USA, 1975.
203. Kang, H.B. Various approaches for driver and driving behavior monitoring: A review. In Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops (ICCVW), Sydney, Australia, 1–8 December 2013; pp. 616–623.
204. Chan, M.; Singhal, A. The emotional side of cognitive distraction: Implications for road safety. *Accid. Anal. Prev.* **2013**, *50*, 147–154.
205. Chan, M.; Singhal, A. Emotion matters: Implications for distracted driving. *Saf. Sci.* **2015**, *72*, 302–309.
206. Deffenbacher, J.L.; Lynch, R.S.; Filetti, L.B.; Dahlen, E.R.; Oetting, E.R. Anger, aggression, risky behavior, and crash-related outcomes in three groups of drivers. *Behav. Res. Ther.* **2003**, *41*, 333–349.
207. Jeon, M.; Walker, B.N.; Gable, T.M. Anger effects on driver situation awareness and driving performance. *Presence Teleoper. Virtual Environ.* **2014**, *23*, 71–89.
208. Eyben, F.; Wöllmer, M.; Poitschke, T.; Schuller, B.; Blaschke, C.; Färber, B.; Nguyen-Thien, N. Emotion on the road—Necessity, acceptance, and feasibility of affective computing in the car. *Adv. Hum. Comput. Interact.* **2010**, *2010*, 263593.
209. James, L. *Road Rage and Aggressive Driving: Steering Clear of Highway Warfare*; Prometheus Books: Amherst, NY, USA, 2000.
210. Grimm, M.; Kroschel, K.; Harris, H.; Nass, C.; Schuller, B.; Rigoll, G.; Moosmayr, T. On the necessity and feasibility of detecting a driver's emotional state while driving. In *Affective Computing and Intelligent Interaction*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 126–138.
211. Katsis, C.D.; Rigas, G.; Goletsis, Y.; Fotiadis, D.I. Emotion Recognition in Car Industry. In *Emotion Recognition: A Pattern Analysis Approach*; Wiley: New York, NY, USA, 2015; pp. 515–544.
212. Stein, F. The challenge of putting vision algorithms into a car. In Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Providence, RI, USA, 16–21 June 2012; pp. 89–94.
213. Nieto Doncel, M.; Arróspide Laborda, J.; Salgado Álvarez de Sotomayor, L.; García Santos, N. Video-Based Driver Assistance Systems. Ph.D. Thesis, Robert Bosch SRL, Cluj Napoca, Romania, October 2008.
214. Kim, K.; Choi, K. SoC Architecture for Automobile Vision System. In *Algorithm & SoC Design for Automotive Vision Systems*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 163–195.
215. Dedeoğlu, G.; Kisačanin, B.; Moore, D.; Sharma, V.; Miller, A. An optimized vision library approach for embedded systems. In Proceedings of the 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Colorado Springs, CO, USA, 20–25 June 2011; pp. 8–13.
216. Satzoda, R.K.; Lee, S.; Lu, F.; Trivedi, M.M. Snap-DAS: A Vision-based Driver Assistance System on a Snapdragon TM Embedded Platform. In Proceedings of the 2015 IEEE Intelligent Vehicles Symposium (IV), Seoul, Korea, 28 June–1 July 2015.
217. Nieto, M.; Otaegui, O.; Vélez, G.; Ortega, J.D.; Cortés, A. On creating vision-based advanced driver assistance systems. *IET Intell. Transp. Syst.* **2014**, *9*, 59–66.

218. Pelaez, C.; Garcia, F.; de la Escalera, A.; Armingol, J. Driver Monitoring Based on Low-Cost 3-D Sensors. *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 1855–1860.
219. Velez, G.; Cortés, A.; Nieto, M.; Vélez, I.; Otaegui, O. A reconfigurable embedded vision system for advanced driver assistance. *J. Real Time Image Process.* **2014**, *10*, 1–15.
220. Forster, F. Heterogeneous Processors for Advanced Driver Assistance Systems. *ATZelektronik Worldw.* **2014**, *9*, 14–18.
221. Anders, J.; Mefenza, M.; Bobda, C.; Yonga, F.; Aklah, Z.; Gunn, K. A hardware/software prototyping system for driving assistance investigations. *J. Real Time Image Process.* **2013**, *11*, 1–11.
222. Young, R.; Zhang, J. Safe Interaction for Drivers: A Review of Driver Distraction Guidelines and Design Implications. Available online: <http://papers.sae.org/2015-01-1384/> (accessed on 26 October 2016).
223. Dasgupta, A.; George, A.; Happy, S.; Routray, A. A vision-based system for monitoring the loss of attention in automotive drivers. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 1825–1838.
224. Krishnasree, V.; Balaji, N.; Rao, P.S. A Real Time Improved Driver Fatigue Monitoring System. *WSEAS Trans. Signal Process.* **2014**, *10*, 146.
225. Veeraraghavan, H.; Papanikolopoulos, N.P. *Detecting Driver Fatigue Through the Use of Advanced Face Monitoring Techniques*; University of Minnesota: Minneapolis, MN, USA, 2001.
226. Karuppusamy, S.; Jerome, J.; Shankar, N. Embedded implementation of facial landmarks detection using extended active shape model approach. In Proceedings of the IEEE 2014 International Conference on Embedded Systems (ICES), Coimbatore, India, 3–5 July 2014; pp. 265–270.
227. De Marsico, M.; Nappi, M.; Riccio, D.; Wechsler, H. Robust face recognition for uncontrolled pose and illumination changes. *IEEE Trans. Syst. Man Cybern. Syst.* **2013**, *43*, 149–163.
228. Malinowski, A.; Yu, H. Comparison of embedded system design for industrial applications. *IEEE Trans. Ind. Inform.* **2011**, *7*, 244–254.
229. Moreno, F.; Aparicio, F.; Hernández, W.; Paez, J. A low-cost real-time FPGA solution for driver drowsiness detection. In Proceedings of the 29th Annual Conference of the IEEE Industrial Electronics Society, Roanoke, VA, USA, 2–6 June 2003; Volume 2, pp. 1396–1401.
230. Wang, F.; Qin, H. A FPGA based driver drowsiness detecting system. In Proceedings of the IEEE International Conference on Vehicular Electronics and Safety, Xian, China, 14–16 October 2005; pp. 358–363.
231. Sanz, R.; Salvador, R.; Alarcon, J.; Moreno, F.; López, I. Embedded Intelligence on Chip: Some FPGA-Based Design Experiences. Available online: <http://www.intechopen.com/books/pattern-recognition-recent-advances/embedded-intelligence-on-chip-some-fpgabased-design-experiences> (accessed on 26 October 2016).
232. Samarawickrama, M.; Pasqual, A.; Rodrigo, R. FPGA-based compact and flexible architecture for real-time embedded vision systems. In Proceedings of the 2009 International Conference on Industrial and Information Systems (ICIIS), Peradeniya, Sri Lanka, 28–31 December 2009; pp. 337–342.
233. Mielke, M.; Schafer, A.; Bruck, R. Asic implementation of a gaussian pyramid for use in autonomous mobile robotics. In Proceedings of the 2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS), Seoul, Korea, 7–10 August 2011; pp. 1–4.
234. Stein, G.P.; Rushinek, E.; Hayun, G.; Shashua, A. A computer vision system on a chip: A case study from the automotive domain. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, Miami, FL, USA, 20–25 June 2005; p. 130.
235. Saussard, R.; Bouzid, B.; Vasiliu, M.; Reynaud, R. Towards an Automatic Prediction of Image Processing Algorithms Performances on Embedded Heterogeneous Architectures. In Proceedings of the 2015 44th International Conference on Parallel Processing Workshops (ICPPW), Beijing, China, 1–4 September 2015; pp. 27–36.
236. Ebert, C.; Jones, C. Embedded software: Facts, figures, and future. *Computer* **2009**, *42*, 42–52.
237. Macher, G.; Stolz, M.; Armengaud, E.; Kreiner, C. Filling the gap between automotive systems, safety, and software engineering. *E I Elektrotech. Informationstech.* **2015**, *132*, 1–7.
238. Jin, L.; Niu, Q.; Jiang, Y.; Xian, H.; Qin, Y.; Xu, M. Driver sleepiness detection system based on eye movements variables. *Adv. Mech. Eng.* **2013**, *5*, 648431.
239. Lumeway. EyeAlert Distracted Driving and Fatigue Warning Systems. Available online: <http://www.lumeway.com/EA.htm> (accessed on 3 January 2016).

240. Craye, C.; Rashwan, A.; Kamel, M.S.; Karray, F. A Multi-Modal Driver Fatigue and Distraction Assessment System. *Int. J. Intell. Transp. Syst. Res.* **2015**, *14*, 1–22.
241. SeeingMachines. Advanced Driver Fatigue and Distraction Detection. Available online: <http://www.seeingmachines.com/> (accessed on 6 January 2016).
242. Fletcher, L.; Apostoloff, N.; Petersson, L.; Zelinsky, A. Vision in and out of vehicles. *IEEE Intell. Syst.* **2003**, *18*, 12–17.
243. Fletcher, L.; Loy, G.; Barnes, N.; Zelinsky, A. Correlating driver gaze with the road scene for driver assistance systems. *Robot. Auton. Syst.* **2005**, *52*, 71–84.
244. Friedrichs, F.; Yang, B. Camera-based drowsiness reference for driver state classification under real driving conditions. In Proceedings of the 2010 IEEE Intelligent Vehicles Symposium (IV), San Diego, CA, USA, 21–24 June 2010; pp. 101–106.
245. VisageTechnologies. Face Tracking and Analysis. Available online: <http://www.visagetechologies.com/> (accessed on 6 January 2016).
246. Edenborough, N.; Hammoud, R.; Harbach, A.; Ingold, A.; Kisačanin, B.; Malawey, P.; Newman, T.; Scharenbroch, G.; Skiver, S.; Smith, M.; et al. Driver state monitor from delphi. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2005; Volume 2, pp. 1206–1207.
247. SMI. Eye and gaze tracking systems. Available online: www.smivision.com (accessed on 6 January 2016).
248. Juhola, M.; Aalto, H.; Joutsijoki, H.; Hirvonen, T.P. The classification of valid and invalid beats of three-dimensional nystagmus eye movement signals using machine learning methods. *Adv. Artif. Neural Syst.* **2013**, *2013*, doi:10.1155/2013/972412.
249. Nissan. Nissan Introduces Driver Attention Alert for Passenger Safety. Available online: <http://automobiletechnology.automotive-business-review.com/news/nissan-introduces-driver-attention-alert-for-passenger-safety-020415-4546074> (accessed on 6 March 2016).
250. Ford. Ford Technology: Driver Alert, Lane Departure Warning. Available online: <http://technology.fordmedia.eu/documents/newsletter/FordTechnologyNewsletter082010.pdf> (accessed on 6 March 2016).
251. Nabo, A. Driver attention-dealing with drowsiness and distraction. Available online: <http://smarteye.se/wp-content/uploads/2015/01/Nabo-Arne-IVSS-Report.pdf> (accessed on 26 October 2016).
252. Ishiguro, H.; Hayashi, T.; Naito, T.; Kasugai, J.; Ogawa, K.; Ohue, K.; Uozumi, S. Development of facial-direction detection sensor. In Proceedings of the 13th Its World Congress, London, UK, 8–12 October 2006.
253. Volvo. Volvo Cars Introduces New Systems for Alerting Tired And Unconcentrated Drivers. Available online: <http://www.mobileye.com/wp-content/uploads/2011/09/Volvo.Driver.Alert.pdf> (accessed on 6 January 2016).
254. MobileEye. MobileEye Advanced Vehicle Technologies Power Volvo Car’s Driver Alert Control (DAC) System. Available online: <http://www.mobileye.com/wp-content/uploads/2011/09/MobileEyeAdvanceVehicleTechnologiesPowerVolvo.pdf> (accessed on 6 January 2016).
255. SAE. Hyundai HCD-14 Genesis Concept Previews Eye-Tracking, Gesture-Recognition Technologies. Available online: <http://articles.sae.org/11727/> (accessed on 17 January 2016).
256. Velez, G.; Otaegui, O. Embedded Platforms for Computer Vision-based Advanced Driver Assistance Systems: A Survey. 2015, arXiv preprint arXiv:1504.07442.
257. Rosey, F.; Auberlet, J.M.; Moisan, O.; Dupré, G. Impact of Narrower Lane Width. *Transp. Res. Rec. J. Transp. Res. Board* **2009**, *2138*, 112–119.
258. Konstantopoulos, P.; Chapman, P.; Crundall, D. Driver’s visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers’ eye movements in day, night and rain driving. *Accid. Anal. Prev.* **2010**, *42*, 827–834.
259. Auberlet, J.M.; Rosey, F.; Anceaux, F.; Aubin, S.; Briand, P.; Pacaux, M.P.; Plainchault, P. The impact of perceptual treatments on driver’s behavior: From driving simulator studies to field tests-First results. *Accid. Anal. Prev.* **2012**, *45*, 91–98.
260. Johnson, M.J.; Chahal, T.; Stinchcombe, A.; Mullen, N.; Weaver, B.; Bedard, M. Physiological responses to simulated and on-road driving. *Int. J. Psychophysiol.* **2011**, *81*, 203–208.

261. Mayhew, D.R.; Simpson, H.M.; Wood, K.M.; Lonero, L.; Clinton, K.M.; Johnson, A.G. On-road and simulated driving: Concurrent and discriminant validation. *J. Saf. Res.* **2011**, *42*, 267–275.
262. De Winter, J.; Happee, P. Advantages and Disadvantages of Driving Simulators: A Discussion. In Proceedings of the Measuring Behavior, Utrecht, The Netherlands, 28–31 August 2012; pp. 47–50.
263. Bach, K.M.; Jæger, M.G.; Skov, M.B.; Thomassen, N.G. Evaluating driver attention and driving behaviour: Comparing controlled driving and simulated driving. In Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction, Swindon, UK, 1–5 September 2008; pp. 193–201.
264. Masala, G.; Grosso, E. Real time detection of driver attention: Emerging solutions based on robust iconic classifiers and dictionary of poses. *Transp. Res. Part C Emerg. Technol.* **2014**, *49*, 32–42.
265. Martin, S.; Tawari, A.; Trivedi, M.M. Toward privacy-protecting safety systems for naturalistic driving videos. *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 1811–1822.
266. Newton, E.M.; Sweeney, L.; Malin, B. Preserving privacy by de-identifying face images. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 232–243.
267. Martin, S.; Tawari, A.; Trivedi, M.M. Balancing Privacy and Safety: Protecting Driver Identity in Naturalistic Driving Video Data. In Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, New York, NY, USA, 17–19 September 2014; pp. 1–7.
268. la Torre, F.D. Vision-Based Systems for Driver Monitoring and De-Identification. Available online: <http://fot-net.eu/wp-content/uploads/sites/7/2015/09/Fernando-de-la-Torre.pdf> (accessed on 30 April 2016).
269. Young, R. Revised Odds Ratio Estimates of Secondary Tasks: A Re-Analysis of the 100-Car Naturalistic Driving Study Data. Available online: <http://papers.sae.org/2015-01-1387/> (accessed on 26 October 2016).
270. Strayer, D.L.; Drews, F.A.; Johnston, W.A. Cell phone-induced failures of visual attention during simulated driving. *J. Exp. Psychol. Appl.* **2003**, *9*, 23–32.
271. Apostoloff, N.; Zelinsky, A. Vision in and out of vehicles: Integrated driver and road scene monitoring. *Int. J. Robot. Res.* **2004**, *23*, 513–538.
272. Shin, B.S.; Xu, Z.; Klette, R. Visual lane analysis and higher-order tasks: A concise review. *Mach. Vis. Appl.* **2014**, *25*, 1519–1547.
273. Hirayama, T.; Mase, K.; Takeda, K. Analysis of temporal relationships between eye gaze and peripheral vehicle behavior for detecting driver distraction. *Int. J. Veh. Technol.* **2013**, *2013*, doi:10.1155/2013/285927.
274. Ohn-Bar, E.; Tawari, A.; Martin, S.; Trivedi, M.M. Vision on wheels: Looking at driver, vehicle, and surround for on-road maneuver analysis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, USA, 24–27 June 2014; pp. 185–190.
275. Tawari, A.; Sivaraman, S.; Trivedi, M.M.; Shannon, T.; Toppelhofer, M. Looking-in and looking-out vision for urban intelligent assistance: Estimation of driver attentive state and dynamic surround for safe merging and braking. In Proceedings of the 2014 IEEE Intelligent Vehicles Symposium Proceedings, Dearborn, MI, USA, 8–11 June 2014; pp. 115–120.
276. Martin, S.; Rangesh, A.; Ohn-Bar, E.; Trivedi, M.M. The rhythms of head, eyes and hands at intersections. In Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV), Gothenburg, Sweden, 19–22 June 2016; pp. 1410–1415.
277. Jain, A.; Koppula, H.S.; Raghavan, B.; Soh, S.; Saxena, A. Car that knows before you do: Anticipating maneuvers via learning temporal driving models. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 13–16 December 2015; pp. 3182–3190.
278. Baddeley, A.D. Selective attention and performance in dangerous environments. *Br. J. Psychol.* **1972**, *63*, 537–546.
279. Vidulich, M.A.; Stratton, M.; Crabtree, M.; Wilson, G. Performance-based and physiological measures of situational awareness. *Aviat. Space Environ. Med.* **1994**, *65*, 7–12.
280. Helmreich, R.L.; Chidester, T.R.; Foushee, H.C.; Gregorich, S.; Wilhelm, J.A. How effective is cockpit resource management training. *Flight Saf. Dig.* **1990**, *9*, 1–17.
281. Ball, K.; Rebok, G. Evaluating the driving ability of older adults. *J. Appl. Gerontol.* **1994**, *13*, 20–38.
282. Satzoda, R.K.; Trivedi, M.M. Drive analysis using vehicle dynamics and vision-based lane semantics. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 9–18.

283. Satzoda, R.K.; Gunaratne, P.; Trivedi, M.M. Drive quality analysis of lane change maneuvers for naturalistic driving studies. In Proceedings of the 2015 IEEE Intelligent Vehicles Symposium (IV), Seoul, Korea, 18 June–1 July 2015; pp. 654–659.
284. Kusano, K.D.; Montgomery, J.; Gabler, H.C. Methodology for identifying car following events from naturalistic data. In Proceedings of the 2014 IEEE Intelligent Vehicles Symposium Proceedings, Dearborn, MI, USA, 8–11 June 2014; pp. 281–285.
285. Martin, S.; Ohn-Bar, E.; Trivedi, M.M. Automatic Critical Event Extraction and Semantic Interpretation by Looking-Inside. In Proceedings of the 2015 IEEE 18th International Conference on Intelligent Transportation Systems, Las Palmas, Spain, 15–18 September 2015; pp. 2274–2279.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).