IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⋇ Impact Factor 8.102 ⋇ Peer-reviewed / Refereed journal ⋇ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.12508

Driver Profile and Driving Pattern Recognition for Road Safety Assessment: Main Challenges and Future Directions

Chandana R A[1], Ravikiran R[2]

Dept. of ECE, SJCIT, Chickballapur, India[1]

Dept. of ECE, SJCIT, Chickballapur, India[2]

Abstract: This study reviews the Artificial Intelligence and Machine Learning approaches developed thus far for driver profile and driving pattern recognition, representing a set of macroscopic and microscopic behaviors respectively, to enhance the understanding of human factors in road safety, and therefore reduce the number of crashes. It provides a definition of the two scientific fields in terms of safety, and identifies the most efficient approaches used regarding methodology, data collection and driving metrics. Results show that K-means and Neural Networks are the most commonly used methodologies for driver profile identification, and Dynamic Time Warping for driving pattern detection. Most studies discovered driver profiles related to aggressiveness, considering mainly speed and acceleration as driving metrics. Based on the gaps and challenges identified, this paper provides a new framework for combining microscopic and macroscopic driving behavior analysis, instead of examining them separately as is the state-of-theart. Such combined results can potentially improve the development of traffic risk models, which could be exploited in applications that monitor drivers in real-time and provide feedback. These models will represent human behavior more accurately, which can eventually lead to the recognition of "optimal" human driving patterns that Automated Vehicles (AV) could 'mimic' to become safer.

I. INTRODUCTION

According to WHO, road traffic injuries cause the premature death of over 80,000 people every year and therefore constitute a major public health problem in the WHO European Region. Approximately 2.7 million people are seriously injured each year in road crashes. These cause a substantial economic loss to society: up to 3% of the gross domestic product of any given country. The main cause of road crashes is persistently attributed to human factors, with a percentage of 65%–95% . It is therefore crucial to deeply understand those factors in order to suggest new effective approaches to shape safe driving behaviors The review of this article was arranged by Associate Editor Chongfeng Weirecovery after surgery, or track the movements of athletes for performance analysis. Driver behavior analytics contribute to this direction through the monitoring of driver behavior in real time and fine resolution.

They have important applications in several business fields including insurance, autonomous vehicles and road network management. In the current era of naturalistic driving, Big Data availability and advances in modelling techniques, there are considerable opportunities for statistical, econometric, Machine Learning (ML) and Artificial Intelligence (AI) applications as a basis for driving behavior analysis.Considerable opportunities are also present in terms of the usage of new data such as driver physiological indicators, variables of driving time and conditions, congestion, road surface and environment conditions, detailed weather and spatial information . The recognition of existing driver profiles and driving patterns could be an approach that takes into account all those factors and contributes to the understanding of driving behavior for the improvement of road safety.

Driver profiles and driving patterns are obviously related to the way the driver interacts with the environment. However, it is not always possible or available to collect related data through experiments. A driver profile is defined as a group of drivers having similar driving behaviors and characteristics, whereas driving pattern is a specific driving behavior that is repetitively occurring by one or more drivers. incorporated in driving behavior analysis studies is the driving "pulse", which is defined as the time period that a vehicle is in motion, bounded by two adjacent stops profile and driving pattern recognition will be reviewed separately in terms of definitions, methodologies and data used. The primary focus is to reveal the best practices, identify future directions for driver profile and driving pattern identification for safety assessment, and determine what each field could potentially "learn" from the other and how both fields can be optimally integrated.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁂ Impact Factor 8.102 ⁂ Peer-reviewed / Refereed journal ⁂ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.12508

It is clarified that this research focuses only on studies related to driver profiles and driving patterns and not generally on the broader concept of individual driving behavior. For more details on studies related to personalized driving behavior analysis, e.g., using reinforcement learning, readers could refer to Some of these studies have developed more advanced simulation and co-simulation platforms to analyze personalized behavior and support personalized driving research. Moreover, it focuses on safety-related studies, and excludes studies on the relationship between driving patterns and, e.g., efficiency, sustainability. Although these are important aspects of driver profiles and driving patterns detection they will not be examined herein. The rest of the paper is structured as follows: Section II describes the methodological approach followed for this literature review; Section III presents the results of the review; the synthesis of the results is performed in Section IV; the conclusions drawn from this analysis and the new framework proposed are revealed in Section V. This research is based on the Rhapsody H2020 research project.

II. METHOD

A RESEARCH The selection steps are illustrated in Fig. 1. Initially, the combined keyword search lead to the identification of a total number of 7,662 studies. Regarding the date range, the key relevant studies published within the past two decades were only taken into consideration – since, one of the pioneer studies that followed the "driving pulse" approach for the identification of driving patterns was [11], which was published in 2002. This review ended in late June 2022 and included research published before June 2022. The date criterion filtered out approximately 6,259 papers. A number of approximately 291 duplicate records were found and removed reaching thus the number of 1,112 studies. The research question addressed in this research are: RQ1) What are the existing methodologies for the identification of driver profiles in terms of safety? RQ2) What are the existing methodologies for driving pattern recognition in terms of safety and what is their focus? RQ3) What are the Artificial Intelligence (AI) techniques applied in these field? RQ4) Which are the data sources and driving metrics used in the analyses? RQ5) What are the main issues and challenges research has encountered, which are the gaps? RQ6) What are the future opportunities and synergies arising in these two fields?

B. KEYWORDS AND SOURCE SELECTION

The search terms used were "driver profile" and "driving pattern". The keyword "driving style" was also tested, and it was observed that this term is usually mentioned in studies focusing on driver recognition and not to the identification of repetitive driving patterns, which is the focus of this research. The keyword "driving behavior" returned an excessively high number of studies on all human factors related to driving. A combination of the words "safe" AND ("driver" OR "driving" OR "profile" OR "pattern") was finally selected. The online databases from which scientific literature was selected included the Scopus, Science Direct, Google scholar and IEEE Xplore search engines.

C.SEARCH STRATEGY AND STUDY SELECTION CRITERIA

The key papers presented in this systematic literature review (SLR) are selected based on some defined search criteria and filters applied in sequence. These criteria included the paper language, which should be English, the relevance of the keywords to the research topic of this review, the publication date range considered for the reviewed articles and access to full papers drawn from this analysis and the new framework proposed are revealed in Section.

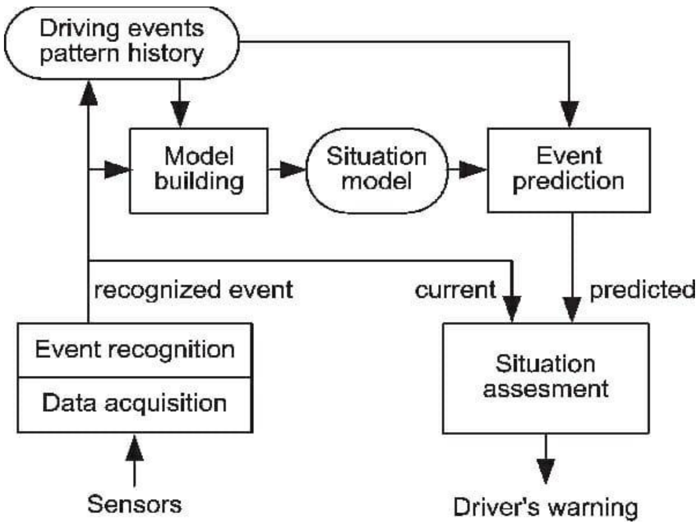

Fig 1: Driving pattern recognition

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.12508

The selection steps are illustrated in Fig. 1. Initially, the combined keyword search lead to the identification of a total number of 7,662 studies. Regarding the date range, the key relevant studies published within the past two decades were only taken into consideration – since, one of the pioneer studies that followed the "driving pulse" approach for the identification of driving patterns was which was published in 2002. This review ended in late June 2022 and included research published before June 2022. The date criterion filtered out approximately 6,259 papers. A number of approximately 291 duplicate records were found and removed reaching thus the number of 1,112 studies. At this point, it was noticed that there is ambiguity in the way the terms 'driver profiles" and "driving patterns" are used in literature. For instance, in some studies driver profiling is defined as the driving behaviour observed in each recorded time step whereas in others, driver profile is a group of drivers that have common behavioral characteristics. '‘Driving style" is used for driver recognition and classification purposes rather than for driving pattern recognition.Other studies refer to "personalized driving behavior", but mostly focus on the identification of driverspecific patterns for the purpose of providing personalized feedback and/or adjustment of ADAS functionalities. It is therefore highlighted that in this study: • Driver profile is defined as a group of drivers having similar driving behaviour and characteristics. • A driving pattern can be defined as a driving behaviour characteristic, such as a driving man over like a harsh braking event, that is occurring repetitively either by the same driver or by different drivers in a population. Hence, in our study driving pattern is a more microscopic aspect than personalized driving behaviour. More specifically, driver profile detection investigates the existence of i) groups of people that behave similarly while driving, ii) common macroscopic characteristics among drivers, iii) methodologies to identify the "strongest" behavioural characteristics that separate these drivers into different groups. Driving pattern recognition investigates the existence of i) repetitive patterns in the data, ii) anomalous patterns in the data, iii) patterns that are uniquely representative of the data, iv) methodologies to separate the data naturally into different regimes. All the above were taken into consideration as criteria to filter out those studies that were not relevant to the primary focus of this research. This filtering was applied to the remaining 1,112 articles based on their title and/or abstract and determined whether an article was relevant to the research questions of this review. This narrowed down the pool of studies to 204. Subsequently, the exact filtering procedure was followed based on the full paper, which leads to the final number of 20 papers for driver profile identification and 26 papers for driving pattern recognition for which, the data collection and analysis methodologies will be reviewed.

III. RESULTS

This section presents the main features of the selected studies on driver profile and driving pattern recognition, in terms of the methodological approaches, the most preferred data sources and driving metrics used in the analyses. The ML and AI techniques applied so far in these fields and their main findings are presented in Table 1 and Table 4 and will also be discussed below. The data collection methodologies and driving metrics used for each driver profiling study can be found in Table 2 and Table 3 respectively, whereas Table 5 and Table 6 illustrate the data collection methodologies and driving metrics used in driving pattern recognition studies.a driver profile identification studies Reference proposed a methodological framework for the evaluation of driving safety efficiency based on Data Envelopment Analysis (DEA). This ML approach was tested on a sample of 56 drivers and resulted in the identification of three groups of drivers namely the non-efficient, weakly efficient and most efficient drivers. Results indicated that inefficient drivers present considerable differences in driving characteristics compared to the groups of weakly efficient and most efficient drivers with the difference of the two latter being less significant. It was found that the number of harsh braking events is an attribute is considered more significant for the characterization of a driver as aggressive or not.

The percentage of speeding and the mobile phone usage were also identified as key factors for the estimation of the safety efficiency index of a driver. A methodology to classify driving behaviour as normal or aggressive on a route level was developed.To this end, authors considered a hybrid AI classification methodological framework and employed both Recurrent Neural Networks (RNN)-guided time-series encoding and rule-guided event detection. It was shown that despite the fact that both long-short-term memory (LSTMs) and Gated recurrent units (GRUs) achieved a similar accuracy in the driving behaviour classification task, GRUs were more efficient at the training stage. Their conclusions regarding driver profiling were that although all drivers drive 'normally' at the time-slice level, the time-slices characterized as 'aggressive' appear more frequently when aggressive drivers are driving. Data used for the analyses were derived from the UAH (University of Alcala) naturalistic driving dataset, which is a public in collection of data captured by Drive Safe, which is a driving monitoring smartphone application recording driving behaviour of various testers in different environments [24]. The data sample included 6 drivers and used recordings of 500 minutes total. Aiming to characterize acceleration, braking and turning as medium, high, low datasets used in the two previous studies, [40] attempted also to classify vehicle type among different driving types. The authors followed three popular classification approaches, i.e., k-nearest neighbours (k-NN), Support Vector Machines (SVM) and decision trees to train their models and test the results using the OBD experiment data.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁘ Impact Factor 8.102 ⁘ Peer-reviewed / Refereed journal ⁘ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.12508

They concluded that based on the proposed features, the decision tree approach achieved the highest classification accuracy and it outperformed RNN-based approaches. With the goal to measure the similarities among individual driving patterns, proposed a classification model based on a combination of ML and AI techniques to recognize the individuals' stable driving patterns. In order to confirm that drivers have their own driving pattern, the authors applied a hierarchical clustering analysis that was performed using two approaches.

The first one used the Dynamic time warping (DTW) method to measure the similarity between two original time series data, whereas the second one used the measure of the Euclidean distance based on a bag-of-pattern (BOP) representation. Consequently, a bi-directional LSTM layer was used and six driver classes were specified. In this research, the time-series data was expressed symbolically using a method called Symbolic Aggregate approximation (SAX), which is used for mapping a symbol or an alphabet to equal-sized segments of time-series data. It was shown that stable driving patterns vary with drivers and driving events, meaning that, e.g., driver A's driving pattern may be similar to that of driver B in a sudden-stop event but similar to driver C in the curve section. This methodology was applied on data collected from a driving simulator experiment with 6 participants who drove a test road of 6.7 km that can be divided into five sections according to the events. Driving metrics were collected with a frequency of 20 Hz and included speed, acceleration, the depth of brake and acceleration pedal, the angle of the steering wheel and the distance from the road lane and the lead vehicle.

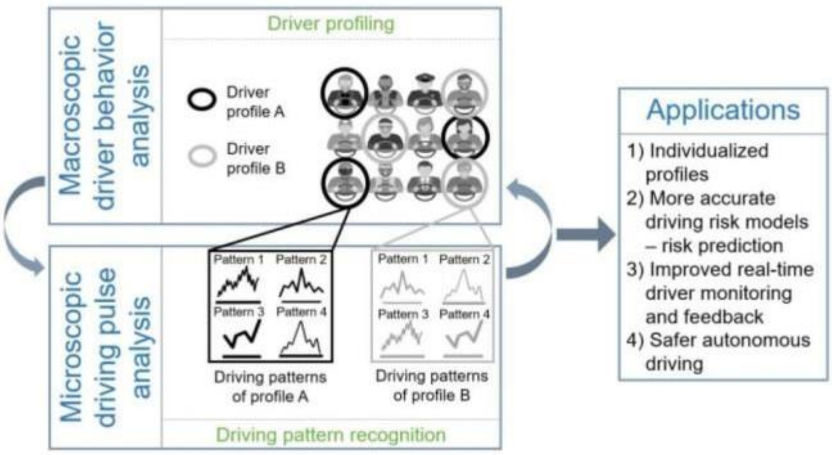

Fig 2: Driving pattern Recognition

IV. DISCUSSION

A. MAIN FINDINGS

Among the Artificial Intelligence methodologies that are used for driver profile identification, it seems that K-means is most commonly used, followed by NN-based models. The extended use of NN approaches also appears in recent studies [63]. Statistical and optimization methodologies are also utilized, whereas PCA is used in several studies to reduce dimensionality of the datasets used. On the other hand, since driving pattern recognition involves mainly analyses of time-series data, the majority of the studies reviewed herein make use of NN-based models. Several different NN approaches are used, mainly from the family of RNN, such as LSTM, GRU, standard RNN, CNN and fuzzy NN. Those are employed both for supervised (classification) and unsupervised tasks (clustering) depending on the type of the dataset. Moreover, DTW is a methodology that is very much preferred for pattern detection that usually complements other clustering methodologies such as Hierarchical and K-Medoid clustering.

It was found that classification methodologies such as SVM, k-NN, and decision trees are exploited as a standalone approach for pattern identification more frequently than clustering methodologies, e.g., support vector clustering. Recent studies also confirm the efficiency of the combination of clustering and classification approaches [43]. Semi-supervised learning methodologies are also adopted by two studies. Other methodologies that were found useful for pattern detection were the HMM, DBN and the SAX method, which performs time-series segmentation by creating symbolic time- series. The common data analytics methodologies that was utilized across studies of both fields studied were the NN-based approaches. The main common characteristic of these studies is that they worked with time-series data, where the performance of RNN is proved to be significantly high.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁒ Impact Factor 8.102 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.12508

B. MAIN CHALLENGES

The main research challenges identified during the conduction of this review are outlined below and discussed afterwards: • The terms "driver profile" and "driving pattern" are used in an ambiguous way
• Absence of a robust methodology for the identification of driver profiles and the recognition of driving patterns • Absence of a methodological approach combining both driver profiles and driving patterns • The quality of the data collected through the naturalistic experiments It was noticed in this review that the terms driver profile and especially driving pattern are used in an ambiguous way in the literature (see Section II-C), for different scopes and analysis purposes. Another important research gap discovered is that there is no robust methodology to identify macroscopic driver safety profiles and microscopic driving safety patterns and understand their relationship with road risk. The methodologies developed so far have focused on grouping behaviors without identifying clear connections with safety. This would be extremely important as it would enable the provision of feedback to decrease driving risk even in real-time. It would also assist in predicting the future states of driving behavior that is entering into a new driver profile, having acquired knowledge on how the behavior of other drivers of the same profile was evolved in the past.

C.SUGGESTED FUTURE DIRECTIONS

The future directions suggested to tackle the challenges discussed earlier are outlined below and discussed afterwards: • Standardization of the terminology for driver profiles and driving patterns • Examination of the concept of "driving pulse" • Application of a combined macroscopic and microscopic approach for driver profiling and pattern detection • Focus on the collection of best-quality data that represent the efficient metrics identified by this review The target of this review was not to analyze individualized driving behavior but rather to focus on micro and macroscopic methods and characteristics that could be used to identify common behaviors among drivers. It is therefore very important to explain the terminology and make clear what is the research objective in this case. From a road safety perspective, this is a research gap that was answered in this study and is summarized in the following paragraph. Based on our research, driver profile can be defined as a group of drivers having similar driving behaviors and characteristics (e.g., aggressive/ non-aggressive, cautious/ distracted/ normal). A driving pattern on the other hand is a specific driving behavior that is repetitively occurring by one or different drivers, and this should be identified at a more disaggregate level, i.e., over very short time frames (within seconds) of driving. On a separate note, the term "personalized driving behavior" refers to the investigation and analysis of the behavior of an individual driver, e.g., for driver detection or for providing personalized feedback [22]. A significant contribution towards this direction is the concept of the driving pulse that was relatively recently introduced [11]. This is based on the concept that each trip should be segmented into shorter time-series in order to investigate the relationships among these segments and how they evolve over time either during the same trip or among trips of the same driver.

V. CONCLUSION

This research thoroughly reviewed the AI and ML approaches used thus far in driver profile and driving pattern recognition studies for traffic safety analysis purposes. The objective was to identify the best approaches in terms of methodology and data collection, and propose future directions to enhance the understanding of the macroscopic and microscopic aspects of driving behavior and therefore, road safety. One of the main findings of this study, was the ambiguity in the definition of the two scientific fields. It also discovered the most efficient driving metrics that should be used in similar research and that data collection frequency should depend on the level of analysis. Moreover, it indicated that the levels of analysis that is used to identify groups of common behaviors, could be categorized as macroscopic, mesoscopic (e.g., [7], [9]) and microscopic depending on the level of information they use. The absence of a clear methodological framework for the identification of macroscopic driver profiles and microscopic driving patterns is suggested to be tackled through a methodology that combines macroscopic and microscopic driving metrics in order volume 4, 2023 driver profile and driving pattern recognition to capture how these two aspects interact with each other as well as using the different approach of the "driving pulse" for microscopic research.

REFERENCES

[1] European Regional Status Report on Road Safety, WHO Europe, Geneva, Switzerland, 2019.
[2] E. Papadimitriou, C. Schneider, J. A. Tello, W. Damen, M. L. Vrouenraets, and A. ten Broeke, "Transport safety and human factors in the era of automation: What can transport modes learn from each other?" Accid. Anal. Prevent., vol. 114, Sep. 2020, Art. no. 105656.
[3] D. Bhatt et al., "CNN variants for computer vision: History, architecture, application, challenges and future scope," Electronics, vol. 10, no. 20, p. 2470, 2021.
[4] Patel et al., "DBGC: Dimension-based generic convolution block for object recognition," Sensors, vol. 22, no. 5, p. 1780, 2022. [5] W. Weidner, F. W. G. Transchel, and R. Weidner, "Telematic driving profile classification in car insurance pricing," Ann. Actuarial Sci., vol. 11, no. 2, pp. 213–236, 2017.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⋇ Impact Factor 8.102 ⋇ Peer-reviewed / Refereed journal ⋇ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.12508

[5] Ellison, S. P. Greaves, and M. C. Bliemer, "Driver behaviour profiles for road safety analysis," Accid. Anal. Prevent., vol. 76, pp. 118–132, Mar. 2015

[6] I. Tselentis, E. I. Vlahogianni, and G. Yannis, "Driving safety efficiency benchmarking using smartphone data," Transp. Res. C, Emerg. Technol., vol. 109, pp. 343–357, Dec. 2019.

[7] Papadimitriou, A. Argyropoulou, D. I. Tselentis, and G. Yannis, "Analysis of driver behaviour through smartphone data: The case of mobile phone use while driving," Safety Sci., vol. 119, pp. 91–97, Nov. 2019.

[8] I. Tselentis, E. I. Vlahogianni, and G. Yannis, "Temporal analysis of driving efficiency using smartphone data," Accid. Anal. Prevent., vol. 154, May 2021, Art. no. 106081.

[9] G. Mantouka, E. N. Barmpounakis, and E. I. Vlahogianni, "Identifying driving safety profiles from smartphone data using unsupervised learning," Safety Sci., vol. 119, pp. 84– 90, Nov. 2019.

[10] Y. Liaw, K. P. Bethune, and C. S. Kim, "Time-series field trip data analysis using adaptive recognition approach. Analysis on driving patterns and vehicle usage for electric vehicles," in Proc. 19th Electr. Veh. Symp. (EVS-19), 2002, pp. 19–23.

[11] J. I. Obuhuma, H. O. Okoyo, and S. O. McOyowo, "Driver behaviour profiling using dynamic Bayesian network," Int. J. Modern Educ. Comput. Sci., vol. 10, no. 7, pp. 50–59, 2018.