IJITCE

ISSN 2347- 3657

International Journal of
Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

SNNO: 2347–3657
Volume 9,Issue2 April ,2021

Fusions of Deep Learning Techniques for Monitoring Vehicle-Based Traffic Abnormalities

MUNMUN SWAIN
.

ABSTRACT

In order to ensure the safety of drivers and passengers in vehicles, as well as a necessary first step toward fully automated driving, video-based anomalous driving behaviour identification is becoming more popular. This tough identification job may be greatly helped by the notable generalisation potential of advanced deep learning models as well as the vast quantities of video clips that are essential for completely training these data-driven deep learning models New deep learning-based fusion models, inspired by the widely discussed densely connected convolutional network (DenseNet), are presented in this research to tackle the difficult issue of detecting anomalous driving behaviour from video. WGD, WGRD, and AWRDN are the names given to these three novel deep learning-based fusion models, which are all based on the concept of wide group densely (WGD) networks. When developing its DenseNet-based model structure, WGD takes significant deep learning model challenges, such as depth, breadth, and cardinality, into account. Remaining networks with superpositions of preceding levels is an essential concept in the WGRD and AWGRD since they are more complex. Three new models' efficacy is being tested with a slew of trials. In this video-based aberrant driving behaviour detection investigation, a comprehensive comparison of various prominent deep learning models revealed their superiority.

INDEX TERMS

Images, driving, aberrant driving, and tightly linked convolutional networks are all examples of artificial intelligence.

INTRODUCTION

At this point, it is generally accepted that high-resolution movies are becoming more and more popular in visual applications. Multiple high-definition cameras are required for video surveillance, for example, to cover a wide area. Working together, they may more easily identify the moving target (e.g.behaviour or even hypothetical purpose) by identifying it [1], [2], re-identifying it [3], [4], and tracking it [5], [6]. When it comes to today's security concerns, the use of high-resolution cameras to catch both noticeable and subtle changes in a target person's emotions in real time [7], [8] has a considerable influence. It is clear from the preceding descriptions that, for the time being, it is possible to acquire and store a considerable volume of high-resolution movies. It's difficult to make high-level choices based on vast quantities of low-level video clips, but that's not the most difficult problem to solve. Videos of drivers in their automobiles are given special attention in this research. Detecting anomalous driving behaviour (i.e. patterns) in drivers is an important high-level decision. The first step to achieving completely autonomous driving is the identification of aberrant driving behaviours. Safety is unquestionably the most important consideration when it comes to autonomous driving. Drivers' actions must be strictly regulated to prevent any possible mishap, as is well-known. As a result, many high-resolution cameras installed in the driver's car may be used to monitor the driver's condition in real time. Additionally, films taken with high-resolution cameras often need fast analysis to identify whether the driver's state remains abnormal or not. In light of the above, it's clear that accurate detection of anomalous driving behaviour, as well as the ability to recognise it quickly, are in great demand. An additional need for the automated identification of improper driving behaviour is a high-speed wireless transmission of high-quality videos, which is essential for this purpose [9]–[23].

ASSISTANT PROFESSOR, Mtech,
Department of CSE
Gandhi Institute for Technology,Bhubaneswar

Detecting inappropriate driving conduct frequently needs a clear and authoritative definition. Normal driving is defined by the International Organization for Standardization (ISO) as a state in which the capacity of a motorist to drive is compromised because the driver is preoccupied with anything other than the task of driving. The three primary types of anomalous driving behaviour may be summarised as follows: In the first case, the driver engages in potentially distracting behaviours like as smoking, drinking, eating, adjusting the temperature of the vehicle's air conditioning system, etc., to ensure their own personal well-being while driving. There are a number of things that may be done while driving, such as applying make-up, shaving, talking on the phone or using other needless equipment, to keep the driver from being distracted. The third category includes driving behaviour that is influenced by the surrounding environment, such as long-term attention to unexpected occurrences outside the vehicle, or caring for children. The usage of cell phones while driving has already become a key role in modern aberrant driving, as seen by the aforementioned abnormal driving behaviours. When making a phone call while driving, researchers discovered that the driver's attention was diverted by 20 percent. Even more severely, if the call is crucial, the driver will be distracted by up to 37 percent, which would increase the risk of a collision by 23 times [24]. Mobile phone usage while driving is thus included in the list of dangerous driving behaviours that may be detected by a computer algorithm in this research. Three innovative deep learning-based fusion models are presented to complete the video-based anomalous driving behaviour detection objective in this work, which uses a single visible-light camera to capture high-resolution footage of the driver. DenseNet (densely linked convolutional networks) was the inspiration for this study's fundamental architecture for deep learning models, which was first suggested in 2017 and later awarded the best paper award at CVPR [25]. It is generally accepted that DenseNet is a relatively new convolutional neural network (CNN)-based deep learning architecture, and it has significant merits for achieving state-of-the art performance in several well-known classification challenges (e.g., CIFAR and SVHN databases) with fewer parameters.. Although it is difficult to train even inside a model's structure that is very complex, the residual network makes it possible. Deep learning fusion methods are also used in this work to create three innovative fusion models for the first time in order to achieve the video-based anomalous driving challenge. Wide group dense (WGD) network, wide group residual dense (WGRD) network, and alternative wide group residual dense (AWGRD) network are the three novel fusion models described in this research. When developing its DenseNet-based model structure, WGD takes significant deep learning model challenges, such as depth, breadth, and cardinality, into account. Since residual networks containing superpositions of previously applied layers are integrated into both WGRD and AWGRD, these methods are more complex than their predecessors. This paper is laid out as follows. Section II provides a short overview of current popular deep learning experiments and related work in aberrant driving detection. There are three new fusion models that use deep learning in Section III. Section IV includes numerous experiments and in-depth studies based on a huge aberrant driving database. Deep learning models are compared against one other in a statistically rigorous manner to demonstrate their superiority against newer, less well-known models. Section V concludes this investigation. The following are some of the study's most important takeaways. It is the first time that the newly announced DenseNet algorithm has been used to a tough video-based anomalous driving behaviour identification problem.. First and foremost, the new models provided in this work, including the major increase of WGD's breadth and cardi nality as well as the sophisticated integration of ResNet and DenseNet in WGRD and AWGRD, are crucial. Third, this study's anomalous driving behaviour identification issue was solved with the use of numerous experiments and detailed analysis.

RELATED WORKS

Following, aberrant driving identification and deep learning approaches, which are closely connected to this work, are highlighted.. Two recent advancements have been briefly explored, with their advantages and disadvantages highlighted.

ABNORMAL DRIVING DETECTION

The literature on automated abnormal driving behaviour detection may be described as follows: there are often three detection systems. An electro-oculogram, electro-encephalogram, respiratory alterations, and variations in blood flow are only a few examples of sensors used in the first method [27], [28]. Face-to-face comparisons are used in the second (i.e., changes in eye movement, mouth movement, head movement, hand features, etc.). It is possible to identify the driver's hand pressure [30], steering time, and braking behaviour [31] using the steering wheel's motion characteristics. Furthermore, sensing human physiological signals offers excellent real-time performance and great accuracy, but its primary benefit of impacting drivers' regular driving behaviour cannot be overlooked. Because of this, the physiological signals of human beings range widely from person to person and from one setting to another. To develop objective and quantifiable criteria for identifying human physiological signals, it is difficult. The gaze direction of the eyes is directly linked to normal/abnormal driving habits, hence eye areas are commonly highlighted in face detections. The proportion of eyelid closure over pupil over time (PERCLOS) [32] is prominent among detecting techniques based on the eyes. It is technically possible to express the proportion of time spent with closed eyes as a percentage of total time in Equation

$$PERCLOS = \frac{Eye\ closing\ time}{Detection\ period} \times 100\%$$

(1).

A motorist is generally deemed to be in an abnormal driving condition when the proportion of time that their eyes are closed exceeds 70% [32]. This detection approach, although accurate and efficient, has a number of major flaws that must be addressed. To begin with, the eyes of drivers with varying physical characteristics and driving styles might seem quite different. One extreme example is the practise of not closing one's eyes when sleeping, resulting in a high rate of PERCLOS false positives. A second problem is that the eyes would be unable to identify difficult difficulties, such as sudden head motions. It is clear that the preceding settings are not ideal for detecting inappropriate driving behaviour just on eyes. Detection methods based on the steering wheel and human physiological signals have many similarities. Using a single visible-based camera, high-resolution recordings of the driver are taken, and the automated identification of anomalous driving behaviour based on these movies is achieved using sophisticated deep learning algorithms. A sensor that relies on physiological signals or information from the steering wheel to identify potential problems, such as undesirable high-variance, will not have to be used at all.

RECENT DEVELOPMENTS IN DEEP LEARNING AND ITS POPULAR UTILIZATIONS

It's worth noting that as computing power and massive amounts of data become more readily available, deep learning approaches are gaining in favour. One may broadly classify most modern deep learning models as either generative or discriminant, depending on how they learn. For example: VAE (varia tional auto-encoder), GAN (generative adversarial network), GLOW (generative flow), etc., are all examples of deep generative learning models that try to duplicate "fake-but-realistic" data from actual data. For classification and discrimination applications, deep discriminative learning models are most often used. Many well-known deep discriminative learning models have won notable international vision contests (e.g., ILSVRC, COCO, etc.). AlexNet, VGG, GoogleNet, ResNet, etc. are all examples of deep discriminative learning models. Recent deep learning models show the following patterns. Deep learning models are becoming more complex, ensuring that they can perform very well in generalisation. Their model structures are also becoming more and more complex. As an example, the breadth of many modern deep learning models is steadily increasing. According to [40], a standard "narrow" single-channel 1001-layer ResNet model can generalise just as well as a large 40-layer ResNet model, but the narrow model requires 1/8 the time to train. Since deep learning models are becoming more complex, their cardinality has also risen sharply. As the number of isolated pathways in a deep learning model is commonly referred to as 'cardinality,' these paths tend to have the same topological structure, resulting in many modern deep learning models being multi channel-based. Therefore, the depth, the wideness, and the cardinality of modern deep learning models are frequently carefully considered while constructing their structures. Also worth noting is the growing popularity of various contemporary deep learning model designs, such as CapsNet [42], the capsule architecture in CapsNet, and so on. In order to carry out the video-based aberrant driving behaviour detection job, this work incorporates the three concerns described above into the design structures of the three newly developed deep learning fusion models. Current vision-based applications have made extensive use of a number of the most recent deep learning methods. New and advanced deep learning architectures have been suggested and used in the popular vision-based detection and tracking sector. When it comes to visual attention prediction, researchers have developed a new visual attention network that incorporates both global and local saliency. Using an action-prediction network powered by continuous deep Q-learning, a novel object tracking hyper-parameter optimization approach was developed in [44]. In order to obtain expressive deep latent features from Siamese networks in order to complete the same object tracking challenge, [45] presented a unique triplet loss. In [46], a new multi-channel ResNet model based on deep learning was used for tracking purposes. According to contemporary video and image-based detection applications [47–48], notions such as saliency and attention have become more popular [47–48]. Video-based anomalous driving behaviour identification is the focus of this investigation.

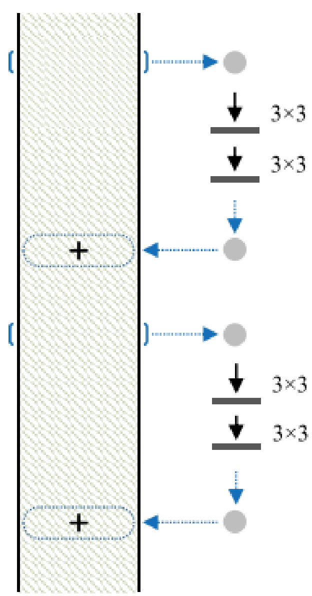

FIGURE 3. An illustration of the core residual architecture in ResNet

5) DENSELY CONNECTED CONVOLUTIONAL NETWORK

(DENSEnET) There are three novel deep learning-based fusion models in this work that were inspired by DenseNet. DenseNet is a large family of models that contains various well-known models as High way Network [53], GoogleNet, and so on. DenseNet is more comprehensive than ResNet. That's because ResNet simply combines the outputs of two adjacent layers, as shown in Figure 3, but DenseNet must add the current layer to all of its preceding layers, as shown in Figure 4. (i.e., as illustrated in Figure 4). The ResNet algorithm is prone to (L 1) direct connections when L layers are present (i.e., one direct connection appears between two adjacent layers). DenseNet, on the other hand, will succeed in

FIGURE 13. Example cases in which WGRD correctly classifies driving patterns but ResNet cannot do.

FIGURE 14. Example cases in which WGD correctly classifies driving patterns but DenseNet cannot do.

International Journal of
Information Technology & Computer Engineering

SNNO: 2347–3657
Volume 9,Issue2 April ,2021

FIGURE 15. Example cases in which AWGD correctly classifies driving patterns but ResNet cannot do

FIGURE 16. Example cases in which AWGD correctly classifies driving patterns but DenseNet cannot do.

V. CONCLUSION

In today's world, video-based improper driving behaviour identification is critical since it is both reliable and automated. For one thing, it's an important step toward completely autonomous driving (as defined by the US Department of Transportation's National Highway Traffic Safety Administration), and therefore it has widespread acceptance. This work is the first to offer three new deep learning-based fusion models for detecting anomalous driving behaviour using video footage. These new models are based on the popular DenseNet, which has been around for a while now. Among WGD's key concerns are the depth, the breadth, and the cardinality of current deep learning models. There, WGD's breadth and cardinality both grow noticeably. Since residual networks containing superpositions of previously applied layers are integrated into both WGRD and AWGRD, these methods are more complex than their predecessors. This concept is very useful in the video-based identification of anomalous driving behaviour, since it provides a thorough description of the latent information in the form of superpositions of preceding layers. Using the standard Kaggle state farm distracted driver detection dataset and thorough comparisons with many other prominent deep learning models, extensive studies have shown that new deep learning-based fusion models are better in both efficacy and efficiency. Studies to detect inappropriate driving behaviour based on "global + local" latent variables will be done in the future. Deep learning methods based on customised mobile chips will also be researched for detecting aberrant driving behaviour.

REFERENCES

[1] W. Cao, J. Yuan, Z. He, Z. Zhang, and Z. He, ‘‘Fast deep neural networks with knowledge guided training and predicted regions of interests for realtime video object detection,’’ IEEE Access, vol. 6, pp. 8990–8999, 2018.

[2] H. Shuai, Q. Liu, K. Zhang, J. Yang, and J. Deng, ‘‘Cascaded regional spatio-temporal feature-routing networks for video object detection,’’ IEEE Access, vol. 6, pp. 3096–3106, 2018.

[3] A. Nanda, P. K. Sa, S. K. Choudhury, S. Bakshi, and B. Majhi, ‘‘A neuromorphic person re-identification framework for video surveilance,’’ IEEE Access, vol. 5, pp. 6471–6482, 2017.

[4] L. Sun, Z. Jiang, H. Song, Q. Lu, and A. Men, ‘‘Semi-coupled dictionary learning with relaxation label space transformation for video-based person re-identification,’’ IEEE Access, vol. 6, pp. 12587–12597, 2018.

[5] Y. Wu, Y. Sui, and G. Wang, ‘‘Vision-based real-time aerial object localization and tracking for UAV sensing system,’’ IEEE Access, vol. 5, pp. 23969–23978, 2017.

[6] S.-H. Lee, M.-Y. Kim, and S.-H. Bae, ‘‘Learning discriminative appearance models for online multi-object tracking with appearance discriminability measures,’’ IEEE Access, vol. 6, pp. 67316–67328, 2018.

[7] M. S. Hossain and G. Muhammad, ‘‘An emotion recognition system for mobile applications,’’ IEEE Access, vol. 5, pp. 2281–2287, 2017.

[8] Z. Pan, X. Yi, and L. Chen, ‘‘Motion and disparity vectors early determination for texture video in 3D-HEVC,’’ Multimedia Tools Appl., to be published. doi: 10.1007/s11042-018-6830-7.

[9] J. Wang, Z. Zhang, B. Li, S. Lee, and R. S. Sherratt, ‘‘An enhanced fall detection system for elderly person monitoring using consumer home networks,’’ IEEE Trans. Consum. Electron., vol. 60, no. 1, pp. 23–29, Feb. 2014.

[10] Z. Zhang, X. Guo, and Y. Lin, ‘‘Trust management method of D2D communication based on RF fingerprint identification,’’ IEEE Access, vol. 6, pp. 66082–66087, 2018.

[11] J. Wang, Y. Cao, B. Li, H.-J. Kim, and S. Lee, ‘‘Particle swarm optimization based clustering algorithm with mobile sink for WSNs,’’ Future Gener. Comput. Syst., vol. 76, pp. 452–457, Nov. 2017.

[12] Z. Xue, J. Wang, G. Ding, Q. Wu, Y. Lin, and T. A. Tsiftsis, ‘‘Deviceto-device communications underlying UAV-supported social networking,’’ IEEE Access, vol. 6, pp. 34488–34502, 2018.

[13] J. Wang, J. Cao, R. S. Sherratt, and J. H. Park, ‘‘An improved ant colony optimization-based approach with mobile sink for wireless sensor networks,’’ J. Supercomput., vol. 74, no. 12, pp. 6633–6645, Dec. 2018.

[14] Y. Tu, Y. Lin, J. Wang, and J.-U. Kim, ‘‘Semi-supervised learning with generative adversarial networks on digital signal modulation classification,’’ Comput. Mater. Continua, vol. 55, no. 2, pp. 243–254, 2018.

[15] J. Wang, Y. Gao, X. Yin, F. Li, and H.-J. Kim, ‘‘An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks,’’ Wireless Commun. Mobile Comput., vol. 2018, Dec. 2018, Art. no. 9472075. [16] J. Sun et al., ‘‘A multi-focus image

fusion algorithm in 5G communications,'' Multimedia Tools Appl., vol. 3, pp. 1–20, Feb. 2018.

[17] J. Wang, C. Ju, Y. Gao, A. K. Sangaiah, and G.-J. Kim, ''A PSO based energy efficient coverage control algorithm for wireless sensor networks,'' Comput. Mater. Continua, vol. 56, no. 3, pp. 433–446, Sep. 2018.

[18] Y. Lin, X. Zhu, Z. Zheng, Z. Dou, and R. Zhou, ''The individual identification method of wireless device based on dimensionality reduction and machine learning,'' J. Supercomput., vol. 5, pp. 1–18, Dec. 2017.

[19] E. B. Tirkolaee, A. A. R. Hosseinabadi, M. Soltani, A. K. Sangaiah, and J. Wang, ''A hybrid genetic algorithm for multi-trip green capacitated arc routing problem in the scope of urban services,'' Sustainability, vol. 10, no. 5, p. 1366, 2018.

[20] W. Qidi, L. Yibing, L. Yun, and Y. Xiaodong, ''The nonlocal sparse reconstruction algorithm by similarity measurement with shearlet feature vector,'' Math. Problems Eng., vol. 2014, Mar. 2014, Art. no. 586014.