Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/38/20230537

Review of machine learning in traffic accident detection

Tengxiang Li

Maynooth International Engineering College, Fuzhou University, Fujian, China

tengxiang.li.2021@mumail.ie

Abstract. With the development and wide application of New Energy Vehicles (NEVs) in the past decade, the chip and power supply method for Machine Learning (ML) intervention provides a good hardware premise, while the 5G technology under the auspices of large-scale rapid data processing is also a prerequisite for the development of Auto Driving, moreover the current traffic accident prediction algorithms have a more critical role in the development of the field of Auto Driving. In this paper, I will consider these aspects as shown below. On the one hand, Anticipation of Traffic Accidents (ATAs) can be well used in today's Autonomous Driving has not yet been popularized, the existing warning information into the corresponding automated driving intervention signals to achieve the possibility of improving the safety of automated driving to increase the possibility of future use of automated driving. On the other hand, the use of ML in the ATAs can be better in the complex traffic environment in a timely manner for the accident warning to achieve a certain degree of reduction in the rate of traffic accidents, in order to optimize the traffic and reduce the economic loss of people. At the same time, in this review, we will propose the combination of data from in-vehicle cameras and road surveillance cameras to analyze the current development of autonomous driving.

Keywords: Machine Learning, Traffic Accident Prediction, Auto Driving, Combination Data.

1. Introduction

With the rapid advancements in communication and related technology, particularly in the fields of ML and sensor, the transportation landscape is undergoing a transformative shift. Among the groundbreaking innovations, the concept of autonomous driving has emerged as a paradigm shift that holds the potential to revolutionize urban mobility.In the relevant papers published in recent years, most of the algorithms related to traffic accident prediction are mainly based on the third-person surveillance video as a dataset for the design of algorithms, which are mainly used in the field of traffic management, while algorithms based on the in-vehicle camera traffic accident prediction algorithm design is relatively rare, and the combination of the two for the coordinated decision-making to ensure the accuracy of the accident prediction and thus to improve the safety of the automated driving coefficient.

1.1. Auto Driving

Autonomous driving, also referred to as self-driving or driverless technology, refers to vehicles' ability to navigate and operate without human intervention, relying on a combination of sensors, algorithms, and real-time data analysis. Besides, it is susceptible to a variety of practical factors such as weather conditions, vehicle speed, traffic flow, roadway ups and downs, etc. In addition, Driving situations

DOI: 10.54254/2755-2721/38/20230537

obey a long-tailed distribution [1], the vast majority of driving situations occur within a limited set of common circumstances, while the causes of most accidents are an endless array of rare and unpredictable events.

2. Method

2.1. Machine Learning

Machine Learning, a subset of artificial intelligence, empowers systems to learn from data and adapt their performance over time.[2] Its capabilities span diverse applications, including complex data analysis, pattern recognition, and predictive modeling. One of the significant applications of ML lies in accident prediction, a critical aspect of modern transportation management. By analyzing large datasets, ML algorithms can uncover hidden relationships and trends, allowing us to make informed decisions and predictions.

2.2. Unsupervised learning

Unsupervised Learning in Traffic Accident Detection can increase the speed of detection, which can be done simply using three modules including preprocessing, candidate selection, and backtracking anomaly detection.[3] In the field of machine learning, unsupervised learning is an important branch. Unlike supervised learning, unsupervised learning does not require labeled training data, but rather discovers patterns and structures from the data.[4] It can be better used in traffic accident detection for anomaly detection i.e. by identifying anomalous traffic conditions and thus predicting them, cluster analysis i.e. by identifying patterns in different environments and thus analyzing the probabilistic scenarios of accidents, downscaling and visualization are good for data analysis.

2.3. Preprocessing

Preprocessing in machine learning refers to the set of operations and transformations applied to raw or unprocessed data before it is used as input to a machine learning algorithm. The main purpose of preprocessing is to clean, transform, and organize the data in a way that makes it more suitable and effective for training and evaluating machine learning models. Preprocessing steps are typically performed to address issues like noise, missing values, outliers, and feature scaling, among others.

Especially in the unsupervised learning the result would be largely affected by the noise of input data, which cause there are worse indicators in the resulting loss function. Most of the preprocessing in Accident anticipation would contain the Background Modeling to distinguishing between moving objects and the stationary objects(Useless information such as the building). Road segmentation to distinguish the road instead of recognized it as the cause of accident. Objection detection which would use some vision algorithms like YOLO or other deep learning-base methods.[3]

2.4. YOLOv3

The YOLO algorithm, based on deep learning, which utilizes deep learning models to process images and achieve precise, efficient, and real-time object detection. Besides, One of the pioneering advancements in computer vision is the YOLOv3 (You Only Look Once version 3) algorithm. YOLOv3 revolutionizes object detection by enabling real-time, accurate identification of objects in both images and video streams.[5] Unlike traditional methods that involve multiple stages of processing, YOLOv3 adopts a unified approach that directly predicts object bounding boxes and class probabilities in a single pass.

YOLOv3 has high accuracy for target detection and can handle large amount of data at the same time, despite the fact that a higher version is now available, it still has a high level of hot use we can notice that in Fig 1. On the other hand, the algorithm has the ability to recognize different targets for detection, and can perform well for classification of pedestrians, vehicles, bicycles, motorcycles,it would able to reach multi-objective detection, and it’s Usage Scenarios has good generalization ability to migrate from road surveillance to in-vehicle cameras.

DOI: 10.54254/2755-2721/38/20230537

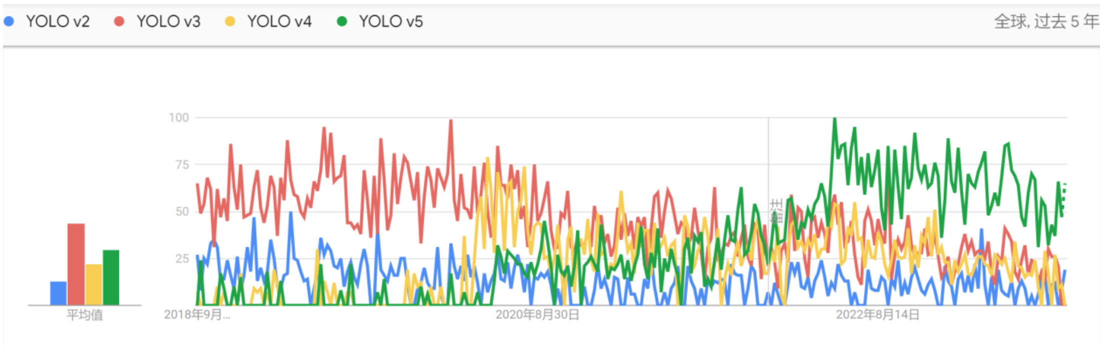

Figure 1. Trend for YOLO version in the past five years in Google browser.

2.5. *Deep leaning*

The fundamental principle of deep learning is artificial neural networks, particularly deep neural networks (DNNs), which comprise an input layer, multiple hidden layers, and an output layer. Each layer consists of numerous neurons that are interconnected with all the neurons in the preceding layer. During the training process, the connection strength (weight) between neurons is automatically adjusted to enable pattern recognition and correlation identification within data. The training of neural networks in deep learning typically involves updating the network weights using the backpropagation algorithm. The word ‘deep’ refers to the large number of hidden layers that compose the neural network.[6] Deep learning can be employed for the acquisition of models pertaining to both normal and abnormal events, subsequently facilitating accident determination.

2.6. *Convolutional Neural Networks*

Convolutional Neural Networks (CNNs) emulate the hierarchical structure of the human visual system, mirroring how our brain processes visual information in a layered manner. By utilizing artificial neurons arranged in layers, CNNs progressively extract features from raw data, capturing various levels of abstraction ranging from basic edges and textures to intricate shapes and patterns. The architecture comprises three key components: Convolutional Layer, which applies filters to input images for detecting specific features such as edges, corners, and textures; Pooling Layer, responsible for downsampling feature maps to reduce computational complexity while preserving crucial information;[7] and Fully Connected Layer that establishes connections between every neuron in the preceding layer with those in the subsequent layer to facilitate complex decision-making.

2.7. *3D Convolutional Neural Networks*

Similar to 2D Convolutional Neural Networks (CNNs), 3D DNNs employ convolution, pooling, and fully connected layers; however, they exhibit distinct operations and hierarchical structures in processing three-dimensional data. This enables 3D DNNs to effectively capture features within the spatial domain, facilitating more precise analysis of video, motion, medical imaging, and other multidimensional datasets.[8] Primarily utilized for analyzing intricate temporal-spatial data, they offer enhanced capabilities for comprehension and interpretation.

2.8. *Recurrent Neural Networks*

The salient characteristic of recurrent neural networks (RNNs) is their ability to pass information between different time steps through their recurrent structure. Each input at a given time step not only includes the current data but also the hidden state from the previous time step, enabling RNNs to capture temporal relationships in sequences and model patterns and dependencies within them. [9] However, deep RNNs are prone to vanishing gradients during backpropagation, which can result in minimal weight updates in shallow layers and limit the network’s capacity for learning and capturing certain types of data. Additionally, exploding gradients can occur when gradients become too large during backpropagation, leading to unstable training that hinders proper model convergence. These issues are often related to the length of analyzed time series data; however, they are less likely to arise in applications with reasonable analysis of temporal dimensions. Therefore, this algorithm may be effectively applied for accident prediction purposes or LSTM may be utilized as a means of mitigating these challenges.

2.8.1. LSTM

The Long Short-Term Memory (LSTM) is a specialized form of recurrent neural network (RNN) that addresses the challenges associated with vanishing gradients and long-term dependencies. LSTM introduces gate mechanisms to selectively retain and forget information, comprising three distinct units: the input gate regulates the inclusion of new cell information, determining its significance; the forget gate manages the extent to which old information in the cell state should be disregarded, influencing its retention level; and finally, the output gate adjusts dynamic adjustments based on inputs. By effectively capturing long-term dependencies within sequences, LSTM mitigates issues such as vanishing or exploding gradients.

2.9. Bayesian deep neural networks

In Bayesian deep neural networks, the weights and parameters of the model are treated as random variables rather than fixed values. This allows us to estimate the distribution of weights and obtain uncertainty information about them through Bayesian inference methods. The role of Bayesian Neural Networks (BNNs) in this paper is to enhance the quality of learned relationship features and introduce predictive uncertainty. By employing Bayesian methods, the model can effectively capture and estimate uncertainties in latent relationship representations, thereby improving the accuracy and robustness of traffic accident prediction models while reducing false alarm rates.[9]

2.10. Tracking Aided anomaly detection

Tracking-assisted anomaly detection is a methodology that leverages target tracking techniques to enhance the detection of anomalies. The fundamental concept is that normal scenes typically exhibit well-defined motion patterns and behavioral rules, while abnormal situations often result in atypical trajectories or behaviors of the targets.[7] By tracking the targets and analyzing their motion patterns, we can establish a baseline for determining whether the targets are in an anomalous state. Subsequently, by integrating information such as the target’s trajectory and appearance features, we can effectively identify anomalies within the scene. Rapid changes observed in targets indicate a heightened likelihood of accidents occurring, which distinguishes it from other existing anomaly detection algorithms that necessitate extensive training with abnormal data samples. Besides, these abnormal data would be hard to receive. However, this algorithm circumvents excessive reliance on training data volume, thereby contributing to improved accuracy in detecting anomalies. This technique can serve as a supplementary tool for most algorithms employed in anomaly detection.

3. Discussion and analysis

Currently, the focus of research lies in utilizing convolutional neural networks and deep learning algorithms for predicting traffic accidents. There have also been attempts to combine the YOLOv3 algorithm with unsupervised learning, which requires more extensive noise reduction processing of videos. However, the preprocessing of these videos and their application scenarios are still quite limited. Most training and application datasets currently used are from road surveillance cameras. To meet China’s development needs for autonomous driving technology, it is necessary to shift towards a first-person perspective using onboard cameras instead of third-person perspective road monitoring. This will enable personalized scenario analysis and provide algorithmic support for autonomous driving while reducing economic pressure by minimizing the number of sensors required to better analyze complex road conditions.

Regarding specific implementation, the initial step involves gathering a substantial amount of data sets; however, presently there is a limited availability of data sets from in-car cameras. To supplement

DOI: 10.54254/2755-2721/38/20230537

the dataset, YouTube's API can be utilized. Furthermore, when constructing the model to enhance accuracy and effectiveness in traffic accident prediction, other influential factors such as traffic flow, vehicle speed, weather conditions, and road conditions should be taken into account. Methodologically speaking, auxiliary accident prediction through tracking-assisted anomaly detection can be achieved by performing object recognition using YOLOv3. This primarily employs recurrent neural networks and convolutional neural networks for traffic accident prediction functionality. Lastly, Bayesian deep neural networks can be employed to estimate predictive uncertainty and enhance model accuracy.

4. Future work

Firstly, the models will undergo further optimization and enhancement, specifically focusing on time series models, convolutional neural networks, and ensemble learning models. Additionally, integration of multiple data sources will be conducted for comprehensive analysis, incorporating weather data to predict various traffic accidents including natural disasters. This may involve accessing authorized vehicle surveillance or satellite imagery for collaborative analysis of vehicle accident warnings, aiming to proactively anticipate and prevent chain-reaction collisions. The accident warning systems of different vehicles will be systematically combined. Through accurate accident prediction, this data can facilitate vehicle attitude correction to avoid accidents and optimize autonomous driving. Lastly, selected personal vehicle data will be shared with traffic management departments through video authorization in order to provide decision support for traffic management and guide the optimization of transportation resources.

4.1. Fusion analysis of weather data

The algorithms presented in this review are based on the analysis of in-vehicle cameras, as these camera systems demonstrate reliable performance under typical weather conditions, particularly on sunny days [10]. However, their pixel quality, brightness levels, visibility, and other analyzed parameters exhibit greater variations during adverse weather conditions such as rain, snow, and fog.Therefore, the purpose of considering these complex parameters is to reduce their interference with the algorithm

4.1.1. Rain

The influence of rain could be solved by a algorithm of base on a model to remove rain from videos. The approach involves segmenting the video into rain and non-rain regions utilizing photometric and dynamics models. Subsequently, the pixels corresponding to rain in each frame are substituted with an estimation of the background intensity, effectively eliminating a significant portion of the rainfall from the video.[11]

4.1.2. Fog

The presence of foggy weather directly impacts video visibility, necessitating adjustments to the algorithmic intervention index. However, addressing this issue solely through video modifications proves challenging. Therefore, in such scenarios, we can initially access the OpenWeatherMap or Tomorrow.io APIs to analyze visibility conditions and subsequently adjust the vehicle's corresponding warning state parameters. Additionally, we can examine road information from widely used navigation software in China or utilize the HERE Technologies API to evaluate traffic flow conditions and provide appropriate feedback support for vehicles.

5. Conclusion

An overview of the application of machine learning in traffic accident prediction, which focuses on introducing the mainstream algorithms that are currently used to perform traffic accident prediction in most of the existing papers. On the one hand, the feasibility, advantages and disadvantages of these algorithms are analyzed and explained in the paper so people who are new to the field can get a good understanding of the strengths and weaknesses of each algorithm through this paper. On the other hand,

DOI: 10.54254/2755-2721/38/20230537

the paper analyzes the application scenarios as well as the training sets of the existing algorithms, firstly, it proposes to transfer the road monitoring application scenarios of the mainstream algorithms to in-vehicle monitoring cameras in order to scater to automated driving requirements. secondly, corresponding solutions are proposed to address issues related to training set quantity. Finally, the paper puts forward an outlook on the future development of this application, which mainly includes the analysis of the model, data fusion, application, and social effect.

References

[1] Yao, Y., Xu, M., Wang, Y., Crandall, D. J., & Atkins, E. M. 2019. Unsupervised traffic accident detection in first-person videos. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) pp. 273-280.
[2] Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet], 9(1), 381-386.
[3] Doshi, K., & Yilmaz, Y., 2020. Fast Unsupervised Anomaly Detection in Traffic Videos. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 624-625.
[4] Pang, G., Yan, C., Shen, C., Hengel, A. V. D., & Bai, X. 2020. Self-trained deep ordinal regression for end-to-end video anomaly detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition pp. 12173-12182.
[5] Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A Review of Yolo algorithm developments. Procedia Computer Science, 199, 1066-1073.
[6] Zantalis F. Koulouras G. Karabetsos S. Kandris D.A Review of Machine Learning and IoT in Smart Transportation. Future Internet 2019, 11(4): 94.
[7] Wang, G., Yuan, X., Zheng, A., Hsu, H. M., & Hwang, J. N. 2019. Anomaly Candidate Identification and Starting Time Estimation of Vehicles from Traffic Videos. In: CVPR workshops pp. 382-390.
[8] Bao, W., Yu, Q., & Kong, Y. (2021). Drive: Deep reinforced accident anticipation with visual explanation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision pp. 7619-7628.
[9] Bao, Wentao. Yu, Qi. Kong, Yu., 2020. Uncertainty-based traffic accident anticipation with spatio-temporal relational learning. In: MM ‘20: The 28th ACM International Conference on Multimedia. Seattle WA USA. 2682-2690.
[10] Zang S, Ding M, Smith D, et al.(2019). The impact of adverse weather conditions on autonomous vehicles: How rain, snow, fog, and hail affect the performance of a self-driving car. In IEEE vehicular technology magazine, 14(2): 103-111.
[11] Garg, K., & Nayar, S. K. (2004, June). Detection and removal of rain from videos. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. (Vol. 1, pp. I-I). IEEE.