IJCTA, 10(8), 2017, pp. 573-579 ISSN: 0974-5572

Developing a novel Algorithm for identifying Driver's behavior in ADAS using Deep Learning

Sakshi Virmani*, Shilpa Gite**

ABSTRACT

With the rapid advancement in today's world a novel deep learning solution, which could be an attempt of studying deep learning for driving behavior analysis. Maneuver anticipation accompaniments current Advance Driver Assistance Systems (ADAS) by preventing accidents and giving drivers more duration to acknowledge to road situations. We introduce a novel approach of analyzing driver's behavior by using Convolutional Neural Network (CNN) that would provide exceptional performance in less response time. We are also proposing to excerpt remarkable and accountable features illustrating complicated driving patterns through CNN.

Index Terms: ADAS, CNN, Deep learning, ITS, RNN

1. INTRODUCTION

Over the coming decagon requirement for advanced driver-assistance systems (ADAS)—has been increasing tremendously. Terminologies such as alley keeping, unseen area check, pre-crash systems etc. that assist drivers to take action petty moment before they carry through a critical maneuver. An extreme focal point of experimentation is to construct fully self-determining vehicles. If autonomous cars influence the market, ADAS technology could gain an advantage. To endorse relaxation and economy, McKinsey recent survey suggests that for the safety car purchasers are becoming even more fascinated in ADAS appliance that protect drivers and reduce accidents.

According to most sources, ADAS technology compared with other automotive systems has transformed the automotive sector from $5 billion to $8 billion. Many automobile purchasers are unfamiliar with the ADAS applications and hence buy other technologies. Through 2020 ADAS is proposed to show forceful momentum.

ADAS have gained progressively more attention with the rapid development of intelligent transportation systems (ITSs).For improved safety, productivity, environmental performance and to carry out data between systems, a current technology Intelligent Transport Systems (ITSs) that is applied to infrastructure and transport. ITSs is rapidly developing throughout the universe in the most powerful way.

Intelligent transport systems are decisive to execute instruction, data processing, exchanging information, and sensor technologies to automobiles (including trucks, cars, aircraft, trains and ships), carting infrastructure and transport users by applying various communication technologies and information to all passenger modes and freight transport to raise the effectiveness, sustainability, safety, efficiency, environmental performance, and resilience of the transportation. New services can be created by the combination of current technologies. ITS can tackle many growing exhalation and clogging problems and make transport safer.

* Symbiosis Institute of Technology, *Email: sakshi.virmani@sitpune.edu.in*

** Symbiosis Institute of Technology, *Email: shilpa.gite@sitpune.edu.in*

Recently deep learning that makes the process much more powerful, an increasingly popular section of machine learning, which takes us closer to one of its original goals: Artificial Intelligence. Research in this field tryout to accomplish finer representations, construct models to attain information from large-scale undefined data and to distinguish patterns in the data, then arrange and allocate them.

Advanced Driver Assistance Systems (ADAS) and driverless vehicles spot pedestrians, identify obstacles and other vehicles and analyze traffic patterns using computer vision focused on deep learning. Deep learning also integrates an ample volume of information aggregated by cameras, sensors, and toll gates etc.

2. RELATED WORK

Our work is relevant to numerous task done previously in understanding driver behavior, anticipating individual driver's action, Recurrent Neural Networks (RNNs) and Long-Short Term Memory for prediction. To improve human-robot association and anticipating individual driver's action several studies has been done. Anticipating individual driver's action for feature matching techniques and extracting data from videos in identifying driver's behavior have been proposed. Approaches used in these works handles data from different sensory fusion and model mortal appearance of individual action. Anticipation requires deep learning approach that handles temporal data and acquires knowledge for different sensory streams fusion.

Ashesh Jain et al. and Hema S Koppula et al. [1]have anticipated maneuvers and proposed a sensory fusion deep learning architecture that fuses data from multiple sensory streams. Their architecture abide of Recurrent Neural Network that needs Long Short-Term Memory (LSTM) units to grab long and short temporal dependencies. A novel loss layer for apprehension has also been used which prevents over-fitting.

They have evaluated their approach on a driving data set with 1180 miles of natural highway and anticipated maneuvers 3.5 seconds earlier with 90.5% precision and 87.45% recall. Eye gaze feature has not been used. For lane change, the algorithm only predicts the left and right lane changes i.e. this setting is only relevant for freeway driving where the preceding possibilities of lane changes are less. Wrong anticipations interpreted by an algorithm can also occur for different reasons such as overtakes, driver's interactions with fellow passengers or they're looking at the surroundings etc.

Ashesh Jain et al. and Avi Singh et al. [2] have introduced a sensory fusion architecture that anticipates and binds information from different sensory streams. They have used RNNs along with LSTM units to capture long temporal dependencies. An innovative loss layer has been introduced for the anticipation that obstructs over-fitting and animate early participation. Anticipated maneuvers certain seconds ahead they happen by increasing the precision to 84.5% and recall to 77.1% on a natural driving data set of 1180 miles. They have also improved head tracking and included driver's 3D head pose as a feature and increased the precision to 90.5% and recall to 87.4%. For lane change, this framework is applicable for freeway driving. They have not used Concurrent Neural Network (CNN) [3] that would give improved performance and accuracy.

Weishan Dong et al., Jian Li, Renjie Yao et al., Changsheng Li et al., Ting Yuan et al., Lanjun Wang et al.[3]. Characterizing Driving Styles with Deep Learning have analyzed Driving Styles Using Deep Learning Approach that transforms raw data of GPS into feature matrices and employed a deep neural network such as CNN and RNN to learn driving styles features. Only low level driving behaviors have been calculated and no explicit temporal context is modeled yet. It is expected to learn and extract higher levels of driving styles features such as input format through deep learning. Driving context such as road level, road shape and weather has not been taken into accounts that also influence driving behavior. Additional information such as trip length, trip shape etc. can also influence in identifying driver's behavior that may give better performance.

S.Gite et al. [21] proposed intelligent multi-sensor data fusion framework which generates context automatically as per current situation which is very useful in ITS. Identifying driver behavior has turned out to be a very critical thing in ADAS. Disha Bhatt [20] worked on driver behavior identification for lane changing scenario for ADAS, using Hidden Markov Model and found improved results in smart cities. They further used Dempster-Shafer Theory to reduce uncertainty in their behavior.

Yuichi Saito et al., Makoto Itoh et al., Toshiyuki Inagaki et al. Driver Assistance System with a Dual Control Scheme: Effectiveness of Identifying Driver Drowsiness and Preventing Lane Departure Accidents have developed a system that pursuits to carry through the safety control of the vehicle and description of driver's state. They have used a driving simulator fitted with the assistance system to explore the efficiency of identifying driver laziness and avoiding lane departure accidents.

Tobias Gindele et al., Sebastian Brechtel et al., and Rüdiger Dillmann et al. Learning Driver Behavior Models from Traffic Observations for Decision Making and Planning [6] have estimated and predicted traffic situations, Building persistent probabilistic models of drivers communication with the environment, hierarchical Dynamic Bayesian Model, various approaches, continuous Partially Observable Markov Decision Process.The system uses many concrete solutions for Probability and decision making and they all are tightly coupled, so this outputs a bit more time complexity.

3. RECURRENT NEURAL NETWORK AND CONVOLUTIONAL NEURAL NETWORK

Recurrent Neural Networks (RNNs) are popular in many Natural Language Processing (NLP) tasks. The idea behind RNNs is to make use of sequential information i.e. they execute the similar task for every element in a sequence, with the output being dependent on the previous computations. RNNs have a memory which acquires information about what has been computed yet. To process arbitrary sequences of inputs RNNs uses their internal memory. At every layer, new data is combined and that information is passed on for an indefinite number of networks. RNNs receives input and yields output at each step. RNNs makes use of a Long Short-term memory(LSTM) that learns which information they need to pass and what they need to reject.

Convolutional Neural Networks (CNNs) are effective in the area such as image recognition and classification. Convolutional Neural Network is successful in distinguishing objects, self-driving cars, images and traffic signs. CNN distributes arrays n arbitrary dimensions and performs mappings. CNNs are applied in many applications such as time series, picture or videos. CNNs are represented by translation invariance and local connectivity. In translation invariance, a neural weight is locked with reference to spatial translation and in local connectivity, neural connections only occur within spatially local regions. CNNs significantly achieve improved outcome than basic feed-forward networks because they take appropriate constraints in mapping function and are applicable when the input data is distributed temporally or spatially.

One of the most popular deep learning models is Convolutional Neural Network which learns visual patterns directly from image pixels. CNNs are efficient in image recognition and classification. They are advantageous in identifying objects, self-driving cars, images and traffic signs. ConvNets is derived from convolution operator. Its primary purpose is to excerpt features from the input image. Convolution learns image features using small squares of input data and extract the spatial relationship between pixels.

The first layer of the hierarchical structure is fed with the small chunk of the image and the information is passed onto other layers of the network. To generate salient features of the data discovered each layer is exploited with trainable filters and local neighborhood pooling operations. From training data set CNN learns features automatically and provides rotation, scaling and shifting as the hidden data or the local receptive field to access the elementary features such as corners and oriented edges.

Our work focuses on recognizing and analyzing driving behavior from the raw input image by extracting high-level features using convolutional layers and pooling.

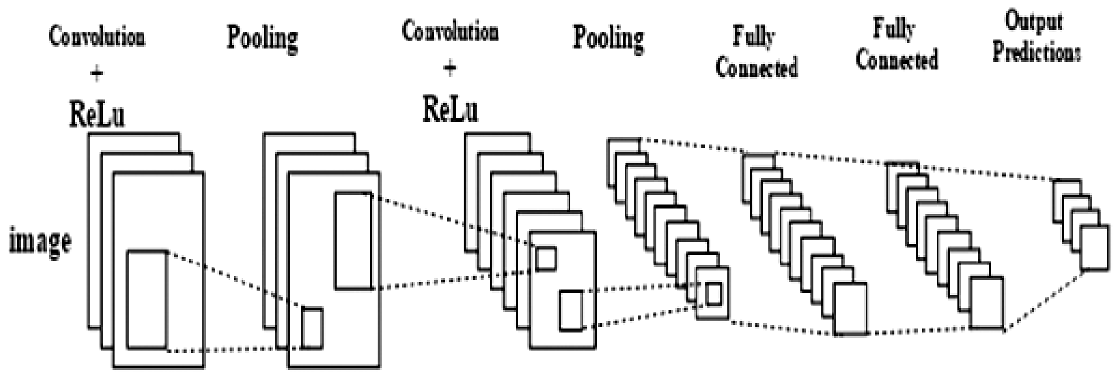

Figure 1: Convolutional Neural Network architecture

For analyzing driver's behavior, a novel approach of the CNN is proposed for exploring salient features that have trainable filters and pooling operations that are alternatively applied. A non-linear operation called ReLU (Rectified Linear Unit) has been used after every Convolution operation. All adverse pixel values in the feature map by zero are replaced by ReLU. The basic purpose of ReLU is to introduce non-linearity of real-world data.

Pooling reserve the very crucial information and scale down the dimensionality of every feature map. Pooling makes feature dimension more manageable and denser and reduces the number of parameters and controls overfitting.

The Fully connected layer is Multi-layer Perceptron which implies that every neuron in the preceding layer is linked with every neuron in the next layer. The output from the pooling layers and convolutional serves high-level features of the input image. The fully connected layer uses high-level features of convolution and pooling and its important feature is classifying the input image into different classes based on the training data set.

3.1. Cnn Application

(1) Image recognition: Convolutional Neural Networks are very much used in image recognition systems and have bring about the lowest error rate of 0.23 percent on MNIST database.[9] Another paper for image recognition and classification on using CNN observed the fastest learning process and have published the best results.[10] CNN when applied to facial recognition contributes a large deduction in error rate.[11]Another paper on 5600 stationary images authors achieved 97.6 percent recognition rate.[12]

(2) Video Analysis: As compare to image classification relatively less effort and time on applying CNNs for video analysis. Videos are more complex than images and hence some extensions have been explored. One approach is to perform convolutions equivalent in both time and space. [13]Another way is to combine different CNNs for the spatial and temporal stream. [14],[15]

(3) Natural Language Processing (NLP): CNNs have shown subsequently productive in NLP and have retain better outputs in semantic parsing [16], sentence modeling [17], prediction [18], and other traditional NLP tasks [19].

(4) Drug Discovery: CNNs have also been applied in discovery of drugs in identifying potential treatments that are effective and secure and can be used in predicting the collaboration between biological proteins and molecules. For structure-based rational drug design, Atoms discovered AtomNet, which is the first deep neural networks. [20]

(5) Fine Tuning: In order to avoid overfitting, CNNs require a huge amount of data set. A familiar approach is to train the network on a huge amount of data set from a relevant domain. After converging the network parameters, the next training step of using in-domain data to fine-tune the network weights is performed. The problem with less training sets is successfully removed when applying CNNs. [21]

3.2. Difference

Recurrent Neural Network (RNN)	*Convolutional Neural Network (CNN)*
RNN handles Arbitrary input and output lengths.	CNN take fixed size input and accomplish fixed size outputs.
It requires more data for processing.	It requires less data for processing.
RNN is a Complex model.	CNN is a Simple model.
RNN is ideal for text and speech.	CNN is ideal for Images and videos.
It is for NLP.	It is for Computer Vision and Image Recognition.

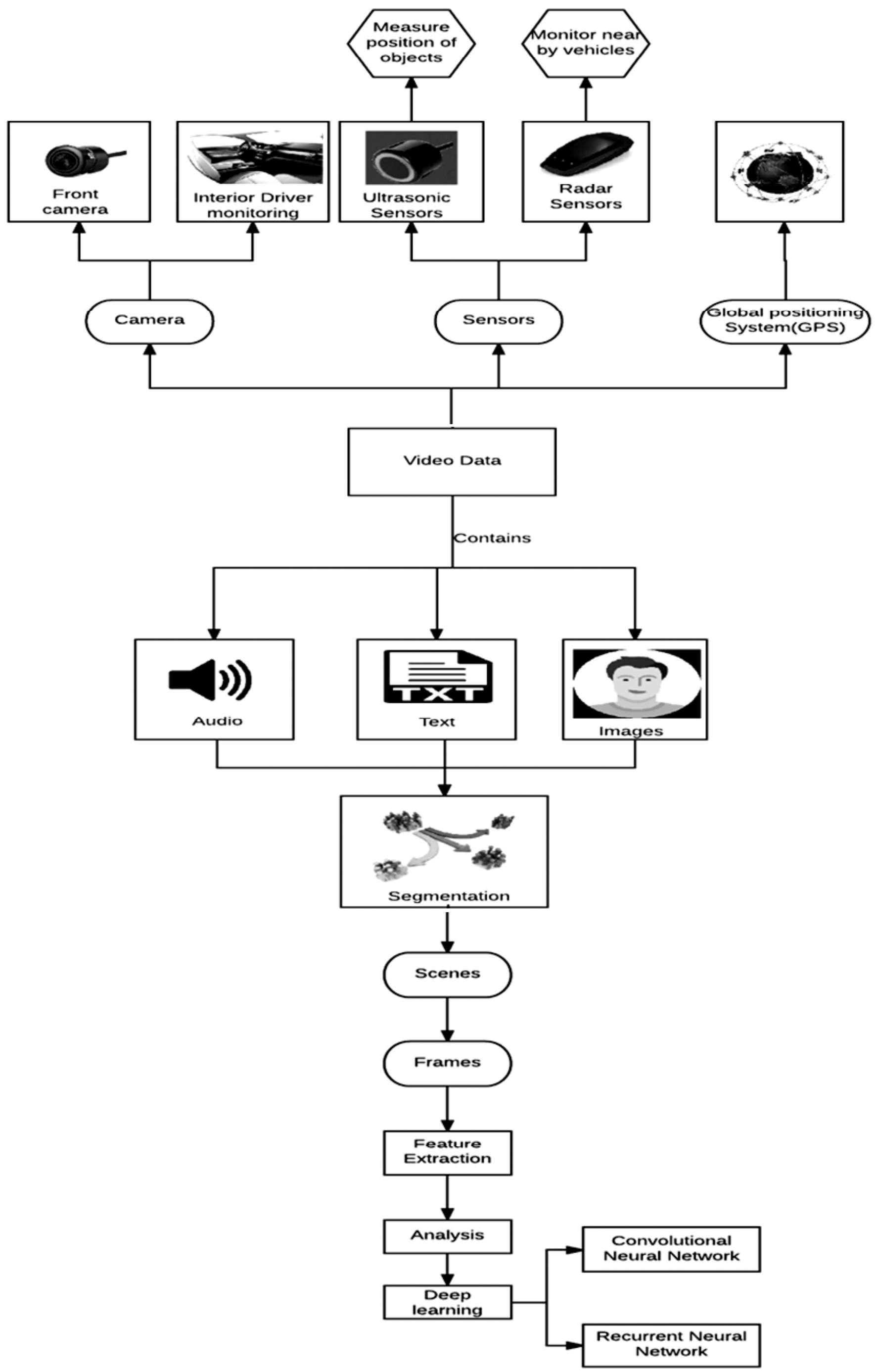

Figure 2: General Architecture of the proposed system

4. DIAGRAM EXPLANATION

Our technique is to model driver's intention and its driving context before the maneuvers. The information is collected from multiple sources like camera, sensors and GPS etc. to constitute driving context. We use the front camera that records videos outside the vehicle for excerpting lane change information and one camera is placed inside the vehicle that monitors the driver's head movements. This information allows us to track driver's mood and expressions that give additional reasoning on actions that the driver is likely to perform. For example, when the vehicle is in the left-most lane it is advised to keep straight or right lane change if the vehicle is coming closer. Additionally, the GPS provides current location and time that empower us to identify the forthcoming road artifacts such as highways, bridges, intersections etc. Videos that are recorded contains data, text, and images that are segmented into scenes and that are splitted into frames and feature extraction is done that give us consistent information from the driving situation and we introduce a deep learning model to knob the temporal aspects of the issue. Furthermore, analysis is done that requires Deep Learning.

5. WHY CNN?

By learning the communication between features of your data and few recognized patterns neural networks make improved predictions. For the presence of distinct features present in the authentic data, CNN acts as a detection filter. Large features can be observed relatively easy and effectively with the CNN earlier layers. For extracting smaller and deep features CNN later layers works more effectively. An ultra-specific classification is made by CNN last layer by linking all above layers specific data as input and gives better results.

In recent years, image classification, face recognition and identifying behavior deep neural network like CNN have achieved great performance. For learning units and to construct increasingly abstract and detailed information of image CNN has elongated layers of neurons.

For identifying specific feature or patterns present in the authentic image every layer acts as a filter. It is relevant to detect precise features located in the original image by such filter. Filters identify whether or not the image contains any such characteristics. Until the whole image is covered in detail the filter is shifted and applied at different positions in the image.

CNNs are successful in the image classification field. To solve ultra-complex problems CNN is applied where the input image is being converted to image data. CNN works better than existing methods and gives more accuracy. Learning features can be identified from time series data such as audio and speech from its two main characteristics of weight sharing and locality. Thus, for image recognition CNN has become very popular. For extracting driving style features from featured matrix requires CNN.

6. CONCLUSION

In this paper, we have proposed advanced driver assistance system for analyzing and predicting driver's behavior in real time traffic. Using CNN and RNN we could recognize the driver's behavior on the basis of certain parameters. Using deep learning we would be able to recognize uncertainty in driver's behavior. In order to avoid accidents and reduce risk while driving we will present a system that performs various maneuvers for avoiding accidents. We will try to implement our algorithm on traffic data sets that significantly gives better results in less execution time and gives better performance.

REFERENCES

[1] Ashesh Jain and Hema S Koppula " Brain4cars: Car that knows before You Do via Sensory Fusion Deep Learning Architecture".

[2] Ashesh Jain and Avi Singh " Recurrent Neural Networks for Driver Activity Anticipation via Sensory-Fusion Architecture".

[3] Weishan Dong , Jian Li, Renjie Yao , Changsheng Li , Ting Yuan , Lanjun Wang " Characterizing Driving Styles with Deep Learning".

[4] Chao Yan , Bailing Zhang " Driving Posture Recognition by Convolutional Neural Networks" .

[5] Yuichi Saito , Makoto Itoh , Toshiyuki Inagaki " Driver Assistance System With a Dual Control Scheme: Effectiveness of Identifying Driver Drowsiness and Preventing Lane Departure Accidents".

[6] Tobias Gindele , Sebastian Brechtel , and Rüdiger Dillmann "Learning Driver Behaviour Models from Traffic Observations for Decision Making and Planning".

[7] Yuichi Saito , Makoto Itoh , Toshiyuki Inagaki " Driver Assistance System With a Dual Control Scheme: Effectiveness of Identifying Driver Drowsiness and Preventing Lane Departure Accidents".

[8] Matthew Browne , Saeed Shiry Ghidaryn " Convolutional neural networks for image processing: an application in robot vision".

[9] Ciresan , Dan , Meier, Ueli; Schmidhuber, Jürgen (June 2012). "Multi-column deep neural networks for image classification". 2012 IEEE Conference on Computer Vision and Pattern Recognition. New York, NY: Institute of Electrical and Electronics Engineers (IEEE): 3642–3649.

[10] Ciresan, Dan , Ueli Meier; Jonathan Masci; Luca M. Gambardella; Jurgen Schmidhuber (2011). "Flexible, High Performance Convolutional Neural Networks for Image Classification" *(PDF). Proceedings of the Twenty-Second international joint conference on Artificial Intelligence-Volume Volume Two.* ***2****: 1237–1242.*

[11] Lawrence, Steve , C. Lee Giles; Ah Chung Tsoi; Andrew D. Back (1997). "Face Recognition: A Convolutional Neural Network Approach". Neural Networks, IEEE Transactions on. **8** (1): 98–113.

[12] Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda (2003). "Subject independent facial expression recognition with robust face detection using a convolutional neural network" (PDF). Neural Networks. **16** (5): 555–559.

[13] Baccouche, Moez; Mamalet, Franck; Wolf, Christian; Garcia, Christophe; Baskurt, Atilla (2011-11-16). "Sequential Deep Learning for Human Action Recognition". In Salah, Albert Ali; Lepri, Bruno. Human Behavior Understanding. Lecture Notes in Computer Science. **7065**. Springer Berlin Heidelberg. pp. 29–39.

[14] Karpathy , Andrej "Large-scale video classification with convolutional neural networks." IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014.

[15] Simonyan, Karen, Zisserman, Andrew (2014). "Two-Stream Convolutional Networks for Action Recognition in Videos".

[16] Grefenstette, Edward; Blunsom, Phil; de Freitas, Nando; Hermann, Karl Moritz (2014-04-29). "A Deep Architecture for Semantic Parsing".

[17] Kalchbrenner, Nal; Grefenstette, Edward; Blunsom, Phil (2014-04-08). "A Convolutional Neural Network for Modelling Sentences".

[18] Collobert, Ronan, and Jason Weston. "A unified architecture for natural language processing: Deep neural networks with multitask learning." Proceedings of the 25th international conference on Machine learning. ACM, 2008.

[19] Collobert, Ronan; Weston, Jason; Bottou, Leon; Karlen, Michael; Kavukcuoglu, Koray; Kuksa, Pavel (2011-03-02). "Natural Language Processing (almost) from Scratch".

[20] Disha Bhatt, Shilpa Gite (2016)."Novel Driver Behavior Model Analysis using Hidden Markov Model to increase Road Safety in Smart Cities". ACM Conference ICTCS '16.

[21] Wallach, Izhar; Dzamba, Michael; Heifets,Abraham (2015-10-09). "AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery".

[22] Durjoy Sen Maitra; Ujjwal Bhattacharya; S.K. Parui, "CNN based common approach to handwritten character recognition of multiple scripts," in Document Analysis and Recognition (ICDAR), 2015 13th International Conference on, vol., no., pp.1021-1025, 23-26 Aug. 2015.

[23] S. Gite and H. Agrawal. "On context awareness for multisensor data fusion in IOT". In Proceedings of the Second International Conference on Computer and Communication Technologies, pages 85-93.

[24] S. Sukode and S. Gite." Vehicle traffic congestion control & monitoring system in IoT. International Journal of Applied Engineering Research, 10(8):19513-19523.

[25] S. Sukode , S. Gite , H. Agrawal, Context Aware Framework in IoT: A survey. IJATCSE-Warse, vol 4, no. 1. (2015)