International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 06 | June - 2021 ISSN: 2582-3930

Distracted Driver Detection

Nikita Negi[1], Shivani Gupta [2], Nayanika Shukla[3], Tanvi Attri[4], Sharanya Chandran[5]

[1, 2,3,4,5] HMR Institute of Technology and Management, New Delhi, India.

---***---

Abstract - In the last several years, the number of traffic accidents has been steadily rising around the world. According to the National Highway Traffic Safety Administration's survey, distracted driving is responsible for approximately one out of every five car accidents. We are attempting to design a reliable and accurate method for recognizing and warning distracted drivers. We describe a CNN-based system that not only detects but also identifies the distracted driver, inspired by the efficiency of Convolutional Neural Networks in computer vision. The VGG-16 architecture has been updated for this task, and several regularization techniques have been implemented to improve accuracy. Our system exceeds earlier methods in the literature, obtaining an accuracy of 96.31 percent and processing 42 photos per second on GPU, according to experimental results. We also look at how dropout, L2 regularization, and batch normalization affect the system's performance. Following that, we show a modified version of our architecture that achieves 95.54 percent classification accuracy while reducing the amount of parameters from 140M in the original VGG-16 to only 15M.

***Key Words*:** Machine Learning (ML), Convolutional Neural Network (CNN), Distracted Driver, VGG-16.

1. INTRODUCTION

Driving is regarded as one of those tasks which requires deep understanding followed by thorough concentration and knowledge about a vehicle. Distracted driving refers to a kind of driving where a driver's concentration is diverted from the road due to some risky activities. The three major types of distractions found by various surveys and studies are: Visual distractions (when the driver's eyes are not on the road), manual distractions (the driver's hands off the wheel), and cognitive distractions (the driver's head off the driving task) [1].

We lost around 36,750 lives in car accidents in the year 2018 and 12% of those accidents were caused due to distracted driving, as per the data provided by National Highway Traffic Safety Administration (NHTSA) [2] respectively. Texting while driving is found to be the most common cause out of all the risky activities.

Many laws have been implemented in various regions that prohibits messaging, doing phone calls, and several other distractions while driving any vehicle. Car accidents caused by distracted driving could be prevented if the efforts of the governments are compiled together with technology such as computerised vision. The algorithm implemented by us automatically detects whether a driver is committing any risky activity and immediately alerts them to focus on the road. If a system like that is implanted in cars, many car accidents caused due to distracted driving could be prevented.

A deep learning neural network [3] which is designed to implement the data in the form of organised arrays such as photographs, is known as a Convolutional neural network (CNN) [4]. CNNs are widely used in computer vision and have become the modern way for implementing various visual applications like image classification, as well as natural language processing for text categorization.

CNN performs extremely well in fetching up patterns in the input images, such as lines, gradients, circles, or even eyes and faces [5]. CNN's are extremely strong for computer vision due to the above mentioned property only. CNN's can be implemented directly on a raw image without any preprocessing, unlike earlier computer vision algorithms [6].

1.1 Goals and Objective

There are three sorts of distractions:

- visual distractions (driving with one's eyes off the road)
- Cognitive distractions (taking the driver's focus off the job of driving).
- Manual distractions (hands off the steering wheel)

The goal of this study is to successfully anticipate what a motorist is doing in each of the images in the collection.

2. Literature Survey

The number of traffic accidents has been steadily growing over the last few years all across the world. According to a survey conducted by the National Highway

Traffic Safety Administration, distracted driving is responsible for approximately one in every five motor vehicle accidents. We strive to create an accurate and robust system for detecting distracted driving [7] and warning the driver. We describe a CNN-based system that not only detects the distracted driver but also determines the reason for distraction, inspired by the performance of Convolutional Neural Networks in computer vision. The VGG- Driving is regarded as one of those tasks which requires deep understanding followed by thorough concentration and knowledge about a vehicle [8]. Distracted driving refers to a kind of driving where a driver's concentration is diverted from the road due to some risky activities. The three major types of distractions found by various surveys and studies are: Visual distractions (when the driver's eyes are not on the road), manual distractions (the driver's hands off the wheel), and cognitive distractions (the driver's head off the driving task).

We lost around 36,750 lives in car accidents in the year 2018 and 12% of those accidents were caused due to distracted driving, as per the data provided by National Highway Traffic Safety Administration (NHTSA) respectively. Texting while driving is found to be the most common cause out of all the risky activities.

Many laws have been implemented in various regions that prohibits messaging, doing phone calls, and several other distractions while driving any vehicle. Car 16 architecture is adapted for this purpose, and several regularisation techniques are used to boost efficiency. Our system exceeds previous approaches in the literature, obtaining an accuracy of 96.31 percent and processing 42 photos per second on GPU. We also investigate the effects of dropout, L2 regularisation [9], and batch normalisation on system performance. Following that, we show a modified version of our architecture that achieves 95.54 percent classification accuracy while reducing the number of parameters from 140M in the original VGG-16 to only 15M.

3. Dataset Description

We use the dataset generated by Abouelnaga et al. in his article. The dataset is divided into 10 categories: safe driving, texting on mobile phones with either the right or left hand, chatting on mobile phones with either the right or left hand, changing radio, eating or drinking, hair and cosmetics, reaching behind and chatting to passengers. The data was gathered from 31 people from seven different nations, using four different automobiles and incorporating many permutations of the drivers and driving circumstances. Drivers, for example, are subjected to varying illumination conditions such as sunshine and shadows. The dataset consists of 17308 images divided into training sets and test sets [10]. We follow the same data distribution for true performance comparison.

4. Technical Approach

Deep Convolutional Neural Network is a form of Artificial Neural Network (ANN) that is inspired by the visual brain of animals [11]. CNNs have made significant progress in a variety of applications over the last few years, including image classification, object identification, action identification, natural language processing, and many more. Convolutional filters/layers, Activation functions, a Pooling layer, and a Fully Connected (FC) layer are the fundamental building components of a CNN-based system. A CNN is created by stacking these layers one on top of the other. Because of the availability of massive amounts of labelled data and computational power, CNNs have advanced at a breakneck pace since 2012. Various architectures like AlexNet, ZFNet, VGGNet, GoogLeNet, ResNet have established benchmarks in computer vision. In this paper, we explore the VGG-16.

5. Libraries Used

- PANDAS - *pandas* aims to be the fundamental high-level building block for doing practical, real world data analysis in Python [12]. Additionally, it has the broader goal of becoming the most powerful and flexible open source data analysis / manipulation tool available in any language.
- NUMPY - Numpy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays. It is the fundamental package for scientific computing with Python [13].
- PICKLE - The pickle module implements binary protocols for serializing and de-serializing a Python object structure. *"Pickling"* is the process whereby a Python object hierarchy is converted into a byte stream, and *"unpickling"* is the inverse operation, whereby a byte stream (from a binary file or bytes-like object) is converted back into an object hierarchy.
- SEABORN - Seaborn is one of the world's most regarded Python libraries that is purpose-built to create beautiful looking visualizations. Data visualization is easily performed in Seaborn.
- MATPLOTLIB - Matplotlib is one of the most popular Python packages used for data visualization. It is a cross-platform library for making 2D plots from data in arrays.
- SKLEARN - Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistent interface in Python.

- KERAS - Keras is a high-level API that is used to make deep learning networks easier with the help of a backend engine. Keras is easy to use and understand with python support so it feels more natural than ever [14].

6. Architecture

CNNs have shown to be incredibly useful in a variety of applications such as image classification, object identification, action identification, natural language processing, and many more. VGG Net is regarded as one of the most prominent CNN architectures in the literature. It reaffirmed the notion that networks should be both deep and simple. It performed admirably in both picture classification and localization tasks. In all thirteen convoluted layers, VGG employs 3 3 filters, a ReLU activation function [15], 2 2 max pooling with stride 2, and a categorical cross-entropy loss function [16]. We start with the pre-trained ImageNet model weights and then fine-tune all of the network's layers using our dataset. The main disadvantage of VGG-16 is the large number of parameters (almost 140M). Fully connected layers are computationally prohibitively costly and use the majority of these parameters. Furthermore, the network with completely linked layers can only be applied to fixed-size input. Replacing a fully linked layer with a convolution layer saves parameters and allows for variable input size. As a result, we construct a fully convolutional neural network by replacing dense layers with 1 1 convolutions. Figure 3 depicts the redesigned network architecture. The number of parameters has been decreased to 15M, which is just 11% of the original VGG-16 parameters. All the regularization parameters remain unchanged [17].

7. Implementation

We tested the realtimeness of the trained models on the Jetson TX1 board. Overall, all models may run in real-time, though with varying computation cycle periods. When the human driver becomes distracted, the TX1 board quickly sends a request to the M3 board to activate the voice alarm, which reminds the driver to focus on the driving job. In terms of performance, the VGG-16 has the highest frequency (14 Hz) but the poorest accuracy (86 percent). This is due to the VGG-16 having the most basic design, which is a sequential model. Other models with more complicated structures reach better levels of accuracy. The ResNet model [18] has the greatest accuracy of 92%. It is, however, the slowest of the four variants. This is due to the fact that the maximum complexity of the ResNet requires greater computing time. The GoogleNet approach achieves greater processing speed at the expense of accuracy (3 percent reduction). Even though the GoogleNet model is slightly slower than the AlexNet model [19] (11 vs 12 Hz), it is more accurate. As a result, in our driving simulation testbed, the GoogleNet model [20] is the best for distraction detection in terms of accuracy and speed.

8. CONCLUSION

Distracted driving is a big issue that contributes to an alarming number of accidents globally. Its detection is a critical system component in semi-autonomous vehicles. We showed a powerful vision-based system that detect distracted driving postures in this study. We will gather a new publicly available distracted driver dataset, which we will utilise to design and test our system. To reach a classification accuracy of 90%, our best model employs a genetically weighted ensemble of convolutional neural networks. We intend to offer a baseline performance against which future research may be benchmarked. We intend to offer a baseline performance against which future research may be benchmarked. We will also demonstrate that a simpler model may work in real-time while maintaining acceptable classification accuracy. The identification of faces, hands, and skin improved classification accuracy in our ensemble. In a real-time situation, however, their performance overhead far outweighs their value.

REFERENCES

1. N. Li and C. Busso, "Predicting Perceived Visual and Cognitive Distractions of Drivers With Multimodal Features," in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 1, pp. 51-65, Feb. 2015, doi: 10.1109/TITS.2014.2324414.

2. M. Pecht, A. Ramakrishnan, J. Fazio and C. E. Nash, "The role of the U.S National Highway Traffic Safety Administration in automotive electronics reliability and safety assessment," in IEEE Transactions on Components and Packaging Technologies, vol. 28, no. 3, pp. 571-580, Sept. 2005, doi: 10.1109/TCAPT.2005.851597.

3. E. Nishani and B. Çiço, "Computer vision approaches based on deep learning and neural networks: Deep neural networks for video analysis of human pose estimation," 2017 6th Mediterranean Conference on Embedded Computing (MECO), 2017, pp. 1-4, doi: 10.1109/MECO.2017.7977207.

4. D. Arora, M. Garg and M. Gupta, "Diving deep in Deep Convolutional Neural Network," 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 2020, pp. 749-751, doi: 10.1109/ICACCCN51052.2020.9362907.

5. S. Kido, Y. Hirano and N. Hashimoto, "Detection and classification of lung abnormalities by use of convolutional neural network (CNN) and regions with CNN features (R-CNN)," 2018 International Workshop on Advanced Image

Technology (IWAIT), 2018, pp. 1-4, doi: 10.1109/IWAIT.2018.8369798.

6. X. Zhang and S. Xu, "Research on Image Processing Technology of Computer Vision Algorithm," 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2020, pp. 122-124, doi: 10.1109/CVIDL51233.2020.00030.

7. M. Leekha, M. Goswami, R. R. Shah, Y. Yin and R. Zimmermann, "Are You Paying Attention? Detecting Distracted Driving in Real-Time," 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), 2019, pp. 171-180, doi: 10.1109/BigMM.2019.00-28.

8. B. Baheti, S. Talbar and S. Gajre, "Towards Computationally Efficient and Realtime Distracted Driver Detection With MobileVGG Network," in IEEE Transactions on Intelligent Vehicles, vol. 5, no. 4, pp. 565-574, Dec. 2020, doi: 10.1109/TIV.2020.2995555.

9. Q. Han, Y. Fan, J. Ge, X. Cui and D. Yu, "Training Ternary Neural Networks By Rectified L2 Regularization," 2019 IEEE 1st International Conference on Civil Aviation Safety and Information Technology (ICCASIT), 2019, pp. 231-234, doi: 10.1109/ICCASIT48058.2019.8973113.

10. B. Qin, J. Qian, Y. Xin, B. Liu and Y. Dong, "Distracted Driver Detection Based on a CNN With Decreasing Filter Size," in IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2021.3063521.

11. L. D. Zhang, L. Jia and W. X. Zhu, "Overview of traffic flow hybrid ANN forecasting algorithm study," 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), 2010, pp. V1-615-V1-619, doi: 10.1109/ICCASM.2010.5620414.

12. Y. Chen, M. Winslett, S. Kuo, Y. Cho, M. Subramaniam and K. E. Seamons, "Performance Modeling for the Panda Array I/O Library," Supercomputing '96:Proceedings of the 1996 ACM/IEEE Conference on Supercomputing, 1996, pp. 45-45, doi: 10.1109/SUPERC.1996.183547.

13. J. Ranjani, A. Sheela and K. P. Meena, "Combination of NumPy, SciPy and Matplotlib/Pylab -a good alternative methodology to MATLAB - A Comparative analysis," 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), 2019, pp. 1-5, doi: 10.1109/ICIICT1.2019.8741475.

14. P. Vidnerova and R. Neruda, "Evolving keras architectures for sensor data analysis," 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), 2017, pp. 109-112, doi: 10.15439/2017F241.

15. D. Stursa and P. Dolezel, "Comparison of ReLU and linear saturated activation functions in neural network for universal approximation," 2019 22nd International Conference on Process Control (PC19), 2019, pp. 146-151, doi: 10.1109/PC.2019.8815057.

16. Y. Ho and S. Wookey, "The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling," in IEEE Access, vol. 8, pp. 4806-4813, 2020, doi: 10.1109/ACCESS.2019.2962617.

17. Juanru Huang, Mei Dong and Shili Li, "A new method of regularization parameter estimation for source localization," Proceedings of 2011 IEEE CIE International Conference on Radar, 2011, pp. 1804-1808, doi: 10.1109/CIE-Radar.2011.6159922.

18. Z. Zhang, "ResNet-Based Model for Autonomous Vehicles Trajectory Prediction," 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), 2021, pp. 565-568, doi: 10.1109/ICCECE51280.2021.9342418.

19. T. Wang, Y. Zhao, L. Zhu, G. Liu, Z. Ma and J. Zheng, "Lung CT image aided detection COVID-19 based on Alexnet network," 2020 5th International Conference on Communication, Image and Signal Processing (CCISP), 2020, pp. 199-203, doi: 10.1109/CCISP51026.2020.9273512.

20. Y. Ibrahim, H. Wang and K. Adam, "Analyzing the Reliability of Convolutional Neural Networks on GPUs: GoogLeNet as a Case Study," 2020 International Conference on Computing and Information Technology (ICCIT-1441), 2020, pp. 1-6, doi: 10.1109/ICCIT-144147971.2020.9213804.