International Journal of Research in Engineering, Science and Management
Volume-2, Issue-11, November-2019
www.ijresm.com | ISSN (Online): 2581-5792

An Overview of Deep Learning in Autonomous Driving

Rohit Pandey[1], Nishant Jethwa[2], Parth Joshi[3]
[1,2,3]*B.Tech. Student, Department of Computer Engineering, Mukesh Patel School of Technology Management and Engineering, Mumbai, India*

***Abstract*: Today, and potentially for a long time to come, the full driving role is too complicated to be entirely formalized as a sensing robotic system that can be directly addressed by model-based or learning-based approaches to achieve maximum unregulated vehicle autonomy. Localization, navigation, scene awareness, vehicle control, route optimization and higher-level design decisions related to the development of autonomous vehicles remain full of open challenges. This is particularly true for unconstrained, real-world operations where the allowable error margin is extremely small and the number of edge cases is extremely large. Until these problems are solved, people will remain an important part of the driving mission, tracking the AI process from just over 0 % to just fewer than 100% of the driving.**

***Keywords*: Autonomous Driving, Deep Learning, Automation**

1. Introduction

The fact that human beings are bad drivers, is well documented in popular culture all over the globe. While this idea is often over-dramatized, there is some truth to it in that we're at times distracted, drowsy, drunk, drugged, and irrational decision makers. This paper focuses on the analysis of driving behavior using deep learning and how it is done to collect data, which will be used to improve the autonomous driving technology. This paper acts as a review paper for MIT-AVT study research paper. This study is still going on and on a much larger scale than when it was started. Volunteers drive their cars in day-to-day environment, and MIT researchers collect the data from that drive with the help of various devices that are discussed below. To date, there are more than 160 participants, >15500 days of participation, 51000 miles and 6.5 billion frames of data collected. This does not mean, however, that a perception-control system that drives better than the average human driver is easy to design and develop. The 2007 DARPA Urban Challenge was a landmark achievement in robotics and autonomous vehicle technologies, when 6 of the 11 autonomous vehicles in the finals successfully navigated an urban environment to reach the finish line, with the first place finisher traveling at an average speed of 15 mph.

2. Sensors and datasets

Autonomous cars are becoming the new face of automobile industry. To track their surroundings, these cars use different sensors. Radar, Lidar, SONAR, and GPS among others are some of the sensors and tech used. A host of systems and technologies are used in autonomous driving that help control the car such as Bayesian simultaneous localization and mapping (SLAM), Real-time locating system (RTLS) and Deep neural network or simply Deep Learning.

Deep learning is a branch of machine learning which uses neural networks that have many layers or that seeks to shape hierarchies of data representation on the actual structure of the hierarchy with minimal input from a human being. Deep Learning or simply deep learning has many properties that can be used in automation of driving, such as being able to automatically learn complex mapping functions, image classification and recognition, speech recognition, etc.

In cars deep learning is used to process raw data provided by Lidar, SONAR and other sensors. Using this information, required inputs are given to the car which is on the on the road. Now to leverage the power of deep learning to extract human behavior from a raw video or to extract any sorts of data, large-scale annotated datasets are required. Late on, these datasets are used to train deep neural network for object detection, image detection and other uses. It is then perfected to be used for its application in autonomous driving. Some of the data sets used to create driving behavior algorithms are:

COCO: Developed by Microsoft, this dataset is specifically sued for object detection and accurate object location. There are two steps involved, in the first step, object localization is done marking an object by a box and in second step, instance segmentation is done, for which precise image masks are needed. This whole dataset contains over 200,000 images.

KITTI: This dataset sets challenging benchmarks for SLAM, stereo vision and 3d object detection, captured by driving around in rural areas and highway of Karlsruhe. This features driving scenarios over 6 hours, tracked at 10-100 Hz using various types of cameras and sensors. It also proposes good truth for 3d scene flow by collecting 400 dynamic scenes from raw datasets and augmenting them with semi-dense scene flow ground truth.

Cityscapes: Dataset focused on understanding the urban street scenes semantically. It features large sets of stereo video recorded in pixel level from 50 different cities and textual labeling at example level. It also contains 20,000 partially segmented images with coarse annotations.

CamVid: This dataset features videos with frame-wise semantic labels, captured from perspective of driving automobile. The ground truth labels in this dataset associate each pixel with one of the 32 different semantic classes. It also enables research on topics such as pedestrian detection and label propagation.

Design of control systems in the driving domain have benefited greatly from learning-based approaches that leverage large-scale data collection in order to construct models that generalize over the edge cases of real-world operation. The software visualization is also the foundation of a good study of naturalistic driving. Here, the hardware such as camera, sensors, etc. and software that performs the data collection. All these components work continuously to record all sensors and data streams by time-stamping them, capture and store HD video from a set of cameras, collect vehicle telemetry from The vehicle's CAN (Control Area Network) buses have remote cellular connectivity to identify when a device failure occurs and to be discreet and elegant so that it does not impact the overall driving experience.

So the application of software engineering, data processing, distributed computing, deep learning techniques and other techniques and technologies are being utilized to create autonomous vehicles in rapiDeep Learning changing transporting systems. Autonomous vehicles will revolutionize the way humans travel and transport goods. Though there is still couple of years of research before the first fully autonomous vehicle arrives.

3. Structure and goals

The governing principle underlying the design of all hardware, low-level software, and higher-level data processing performed in this AVT study is: continual, relentless innovation, while maintaining backward compatibility. From the beginning, we chose to operate at the cutting-edge of data collection, processing, and analysis approaches. This meant trying a lot of different approaches and developing completely new ones: from sensor selection and hardware design to the robust time-critical recording system and the highly sophisticated data pipeline described in. It's a philosophy that allowed us to scale quickly and find new solutions at every level of the system stack.

As previously noted, the medium duration (one month long) NDS is conducted using MIT-owned vehicles, while the long duration (over 1 year) NDS is conducted in a vehicle owned by subject. Participants are divided into primary and secondary drivers. Warning labels on windows to advise non-consented passengers and drivers of the ongoing data collection, and coordinate with project staff for system maintenance and data retrieval. Recruitment is conducted through flyers, social networks, forums, online referrals, and word of mouth. Primary drivers are paid for their time involved in vehicle instrumentation, appointments for system maintenance, data recovery, and questionnaires completion.

Participants in the medium duration (one month long) NDS7are provided with introductions to the fleet vehicles in the form of an approximately 1.5-hour long training session. This session is intended to introduce drivers to the physical characteristics of the vehicle, and provide a sufficient understanding of vehicle features in order to support safe use of advanced technologies. Participants are provided with a study overview by a researcher and presented with manufacturer produced videos or information packets on one or more of the basic and advanced features available in the vehicle. Their devices are matched with the car and are given the opportunity to perform some voice commands. (e.g. making a phone call, entering a destination). Next, the role, activation and use of the following features are given more detailed overviews:

- Adaptive Cruise Control (ACC)
- Pilot Assist (from Volvo)
- Super Cruise (from Cadillac)
- Forward Collision Alert Warning / City Safety (from Volvo)
- Automatic Emergency Braking
- Lane Departure Warning (LDW)
- Lane Keep Assist (LKA)
- Blind Spot Monitor

Following this stationary in-vehicle training, participants are provided with an on-road training drive on a multi-lane highway. A driving session on the highway takes at least 30 minutes to provide realistic exposure to real-world setting systems. Participants were encouraged to use the researcher and ask questions while checking the systems during the learning ride. Three questionnaire batteries and a semi-structured interview are used to capture medium-length (one month long) NDS self-report information. A semi-structured interview is conducted in person between a research associate and the study participant at the end of the one-month naturalistic driving period, and lasts approximately 30-60 minutes. It consists of predefined questions focusing on initial reactions to the vehicle, experience during the training drive, how training affected their understanding of the technologies, and driver perceptions of the technologies. Naturalistic driving data and automated in-depth training analysis of these data provide observations, recommendations and well-founded scenarios as to the road to safe and effective integration of artificial intelligence into modern and future vehicle systems. In such autonomous vehicle engineering, raw data and a high level understanding of human actions and system performance are of importance. They are:

- Car manufacturers
- Suppliers for automotive parts
- Insurance companies
- Technology companies
- Government agencies
- Education and academic organizations

When the path forward is full of uncertainty, risks, potentially costly misaligned investments, and paradigm shifts, open innovation provides more value than closed competition. At this moment in time, autonomous vehicle technology is a space where competitors win by collaborating, sharing high-level insights and large-scale, real-world data. High-level measures such as system use and system performance can be used to inform the design, development and validation of future vehicle systems. In and out of the car, video recording can be used to build systems for vision, monitoring, scheduling, driver sensing, and driver aid.

The backbone of a successful naturalistic driving study is the hardware and low-level software that performs the data collection. In the study conducted by MIT, that role is served by a system named RIDER. It was designed and continuously developed to satisfy the following goals and requirements:
Time-stamped Asynchronous Sensor Recording: Record all sensors and data streams in a way that each sample of data (no matter its frequency or data source) is time-stamped using a centralized, reliable time-keeper. In other words, data has to be time-stamped in a way that allows perfect synchronization of multiple data streams in post-processing.
High-Definition Video: Capture and record 3 to 6 cameras at 720p (2.1 megapixels) resolution. One of the most important design decisions of the entire study was the choice of camera positions, resolution, and compression.

- *CAN Bus:* Collect vehicle telemetry from the Controller Area Network (CAN) bus(es) of the vehicle. Each vehicle has different ports and bus use rules, with little publicly available information on message ID mapping or message quality. Raw CAN messages must be registered in such a manner that the essential information is stored within those messages even if they cannot be decoded at the time of compilation.
- *Remote Cellular Connectivity:* Low-bandwidth, infrequent communication of system status via a cellular connection in order to detect when RIDER system malfunction occurs.
- *Discrete and Elegant Appearance:* Parts of the system that are noticeable from inside or outside the car should have a minimal shape factor and visual design features that do not distract from the vehicle's overall appearance or affect the overall driving experience.
- *Camera Mounting is Robust but Removable:* Mounting must be consistent, reliable, and removable designed specifically for each vehicle’s interior physical characteristics.

4. Conclusion

The application of state-of-the-art embedded system programming, software engineering, data processing, distributed computing, computer vision and deep learning techniques to the collection and analysis of large-scale naturalistic driving data in the MIT-AVT study seeks to break new ground in offering insights into how human and autonomous vehicles interact in the rapid Deep Learning changing transportation system. This research introduces the methodology behind the study of MIT-AVT aimed at identifying and encouraging the next wave of naturalistic driving studies.

Acknowledgement

The authors would like to thank all the people associated with MIT-AVT study and the broader driving and artificial intelligence research community for their valuable content and information, which helped in making this review paper.

References

[1] L. Fridman, MIT Autonomous Vehicle Technology Study: arXiv Apr 2019.
[2] W. H. Organization, Global status report on road safety 2015. World Health Organization, 2015.
[3] V. V. Dixit, S. Chand, and D. J. Nair, “Autonomous vehicles: disengagements, accidents and reaction times,” PLoS one, vol. 11, no. 12, p. e0168054, 2016.
[4] M. Buehler, K. Iagnemma, and S. Singh, The DARPA urban challenge: autonomous vehicles in city traffic. springer, 2009, vol. 56.
[5] V. V. Dixit, S. Chand, and D. J. Nair, “Autonomous vehicles: disengagements, accidents and reaction times,” PLoS one, vol. 11, no. 12, p.e0168054, 2016.
[6] F. M. Favar`o, N. Nader, S. O. Eurich, M. Tripp, and N. Varadaraju, “Examining accident reports involving autonomous vehicles in california,” PLoS one, vol. 12, no. 9, p. e0184952, 2017.
[7] R. Tedrake, “Underactuated robotics: Algorithms for walking, running, swimming, flying, and manipulation (course notes for mit 6.832),” 2016.
[8] M. R. Endsley and E. O. Kiris, “The out-of-the-loop performance problem and level of control in automation,” Human factors, vol. 37, no. 2, pp. 381–394, 1995.
[9] B. Reimer, “Driver assistance systems and the transition to automated vehicles: A path to increase older adult safety and mobility?” Public Policy & Aging Report, vol. 24, no. 1, pp. 27–31, 2014.
[10] K. Barry, “Too much safety could make drivers less safe,” July2011. https://www.wired.com/2011/07/active-safety-systems-could-create-passive-drivers/
[11] V. L. Neale, T. A. Dingus, S. G. Klauer, J. Sudweeks, and M. Goodman, “An overview of the 100-car naturalistic study and findings,” National Highway Traffic Safety Administration, Paper, no. 05-0400, 2005.
[12] T. A. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. Perez, J. Hankey, D. Ramsey, S. Gupta et al., “The100-car naturalistic driving study, phase ii-results of the 100-car field experiment,” Tech. Rep., 2006.
[13] S. G. Klauer, T. A. Dingus, V. L. Neale, J. D. Sudweeks, D. J. Ramseyet al., “The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data,” 2006.
[14] K. L. Campbell, “The shrp 2 naturalistic driving study: Addressing driver performance and behavior in traffic safety,” TR News, no. 282, 2012.
[15] T. Victor, M. Dozza, J. B¨argman, C.-N. Boda, J. Engstr¨om, C. Flannagan, J. D. Lee, and G. Markkula, “Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk,” Tech. Rep., 2015.
[16] M. Benmimoun, F. Fahrenkrog, A. Zlocki, and L. Eckstein, “Incident detection based on vehicle can-data within the large scale field operational test (eurofot),” in 22nd Enhanced Safety of Vehicles Conference (ESV 2011), Washington, DC/USA, 2011.
[17] L. Fridman, P. Langhans, J. Lee, and B. Reimer, “Driver gaze region estimation without use of eye movement,” IEEE Intelligent Systems, vol. 31, no. 3, pp. 49–56, 2016.
[18] L. Fridman, J. Lee, B. Reimer, and T. Victor, “Owl and lizard: patterns of head pose and eye pose in driver gaze classification,” IET Computer Vision, vol. 10, no. 4, pp. 308–313, 2016.

[19] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, "Labelme: a database and web-based tool for image annotation," International Journal of computer vision, vol. 77, no. 1, pp. 157–173, 2008.

[20] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes data set for semantic urban scene understanding," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 3213–3223.

[21] R. R. Knipling, "Naturalistic driving events: No harm, no foul, novalidity," in Driving Assessment 2015: International Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design. Public Policy Center, University of Iowa Iowa City, 2015, pp. 196–202.

[22] R. R. Knipling, "Crash heterogeneity: implications for naturalistic driving studies and for understanding crash risks," Transportation Research Record: Journal of the Transportation Research Board, no. 2663, pp.117–125, 2017.

[23] L. Fridman, B. Jenik, and B. Reimer, "Arguing machines: Perception control system redundancy and edge case discovery in real-world autonomous driving," arXiv preprint arXiv:1710.04459, 2017.

[24] V. Shankar, P. Jovanis, J. Aguero-Valverde, and F. Gross, "Analysisof naturalistic driving data: prospective view on methodological paradigms," Transportation Research Record: Journal of the Transportation Research Board, no. 2061, pp. 1–8, 2008.

[25] N. Kalra and S. M. Paddock, "Driving to safety: How many miles odriving would it take to demonstrate autonomous vehicle reliability?" Transportation Research Part A: Policy and Practice, vol. 94, pp. 182–193, 2016.

[26] I. Good fellow, Y. Bengio, and A. Courville, Deep learning. MIT press, 2016.

[27] R. Hartley and A. Zisserman, Multiple view geometry in computer vision. Cambridge university press, 2003.

[28] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE,2009, pp. 248–255.

[29] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, "Introduction to word net: An on-line lexical database," International Journal of lexicography, vol. 3, no. 4, pp. 235–244, 1990.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.