Lex Baker

6/13/22

The Impact of Deep Learning on Autonomous Vehicles

Self-driving cars, once considered an impossible futuristic idea, now exist a step away from reality. Rapid technological developments have made them possible, and vehicles that provide driving assistance are becoming more commonplace. The final step, however, requires technology still in the making, with new breakthroughs happening daily. That technology is deep learning, a subset of machine learning and artificial intelligence. It’s capable of picking apart inputs across hundreds of layers to correctly identify an object out of thousands of possibilities. The mathematical reasoning behind these algorithms is incredibly fascinating, and they are constantly being improved upon, with new models and methods of learning boosting accuracy and shortening response times.

What exactly deep learning constitutes depends on several factors, such as the model complexity and number of hidden layers. Machine learning, a level simpler in terms of complexity from deep learning, involves computers learning and adjusting parameters to calculate something they weren’t explicitly programmed to know or do. Deep learning, in comparison, involves artificial neural networks, which are modeled after the human brain, that have anywhere from dozens to thousands of hidden layers, each of which have their own set of parameters, known as weights. These weights are adjusted based on how accurate the last iteration was and eventually can give impressively accurate results. The reason why deep learning models rose in popularity is that they can learn what to look for in raw data rather than being told, giving them a much wider range of application. For autonomous vehicles, the ability to process images and videos means computers can finally recognize road signs, traffic lights, and obstacles like a human would, making them adaptable outside of carefully controlled environments. This adaptability will lead to widespread adoption of self-driving vehicles as people can use them wherever they would normally need to travel.

One of the main issues with self-driving cars is obstacle detection. A vehicle moving a 20 mph has much more time to react to a deer running across the road than a vehicle travelling at 65 mph. A current popular method for cars to “see” the world around them is through LIDAR, which stands for **LI**ght **D**etection **A**nd **R**ange. These systems are well-tested and provide lots of data, but are expensive and struggle in certain conditions, such as heavy rain. Instead, the use of regular cameras might be a more cost effective and useful option. Regular cameras, however, would require deep learning to pick up on features around them and gauge their distance and size. Luckily, many recent developments in artificial intelligence have been on models that work with images and videos. Deep residual networks in particular mean that the thousands of hours of data that companies like Telsa accumulate can be used without accuracy plateauing. After all, a system for a task with such dangerous implications needs to have a detection model with an accuracy as close to 100% as possible.

Another problem with autonomous vehicles is traffic flow prediction, which requires both spatial and temporal analysis and predictions. In a closed environment with only cars that can communicate is manageable but adding in the randomness of human drivers means that calculations will need to be quick and precise. Once again, deep learning can help, as when provided with enough data, a model could accurately predict the best course of action, even with hundreds of rapidly changing variables to account for.

In all, the field of autonomous vehicles is closely intertwined with deep learning, and as one advances forward so will the other. With billions of dollars in funding across the world, new developments are sure to rapidly change the field, bringing about that last step needed to make self-driving vehicles a reality for daily use.

Citations and References:

* S. Kuutti, R. Bowden, Y. Jin, P. Barber and S. Fallah, "A Survey of Deep Learning Applications to Autonomous Vehicle Control," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 2, pp. 712-733, Feb. 2021
* Neha Sharma, Vibhor Jain, Anju Mishra, An Analysis Of Convolutional Neural Networks For Image Classification, Procedia Computer Science, Volume 132, 2018, https://doi.org/10.1016/j.procs.2018.05.198.
* S. Yang, W. Wang, C. Liu and W. Deng, "Scene Understanding in Deep Learning-Based End-to-End Controllers for Autonomous Vehicles," in IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 49, no. 1, pp. 53-63, Jan. 2019, doi: 10.1109/TSMC.2018.2868372.
* S. Mandal, S. Biswas, V. E. Balas, R. N. Shaw and A. Ghosh, "Motion Prediction for Autonomous Vehicles from Lyft Dataset using Deep Learning," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), 2020, pp. 768-773, doi: 10.1109/ICCCA49541.2020.9250790.
* H. -H. Jebamikyous and R. Kashef, "Autonomous Vehicles Perception (AVP) Using Deep Learning: Modeling, Assessment, and Challenges," in IEEE Access, vol. 10, pp. 10523-10535, 2022, doi: 10.1109/ACCESS.2022.3144407.
* C. -s. Oh and J. -m. Yoon, "Hardware Acceleration Technology for Deep-Learning in Autonomous Vehicles," 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), 2019, pp. 1-3, doi: 10.1109/BIGCOMP.2019.8679433.
* LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* 521**,**436–444 (2015). https://doi.org/10.1038/nature14539
* Krizhevsky, A., Sutskever, I. & Hinton, G. ImageNet classification with deep convolutional neural networks. In Proc. Advances in Neural Information Processing Systems 25 1090–1098 (2012).
* Farabet, C., Couprie, C., Najman, L. & LeCun, Y. Learning hierarchical features for scene labeling. IEEE Trans. Pattern Anal. Mach. Intell. 35, 1915–1929 (2013).
* Mikolov, T., Deoras, A., Povey, D., Burget, L. & Cernocky, J. Strategies for training large scale neural network language models. In Proc. Automatic Speech Recognition and Understanding 196–201 (2011).
* Kaggle. Higgs boson machine learning challenge. Kaggle https://www.kaggle.com/c/higgs-boson (2014).
* Sutskever, I. Vinyals, O. & Le. Q. V. Sequence to sequence learning with neural networks. In Proc. Advances in Neural Information Processing Systems 27 3104–3112 (2014).
* Bottou, L. & Bousquet, O. The tradeoffs of large scale learning. In Proc. Advances in Neural Information Processing Systems 20 161–168 (2007).
* Duda, R. O. & Hart, P. E. Pattern Classification and Scene Analysis (Wiley, 1973).
* Schölkopf, B. & Smola, A. Learning with Kernels (MIT Press, 2002).
* Bengio, Y., Delalleau, O. & Le Roux, N. The curse of highly variable functions for local kernel machines. In Proc. Advances in Neural Information Processing Systems 18 107–114 (2005).
* Werbos, P. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. PhD thesis, Harvard Univ. (1974).
* Parker, D. B. Learning Logic Report TR–47 (MIT Press, 1985).
* Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. Nature 323, 533–536 (1986).
* Glorot, X., Bordes, A. & Bengio. Y. Deep sparse rectifier neural networks. In Proc. 14th International Conference on Artificial Intelligence and Statistics 315–323 (2011).
* Dauphin, Y. et al. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. In Proc. Advances in Neural Information Processing Systems 27 2933–2941 (2014).
* Choromanska, A., Henaff, M., Mathieu, M., Arous, G. B. & LeCun, Y. The loss surface of multilayer networks. In Proc. Conference on AI and Statistics http://arxiv.org/abs/1412.0233 (2014).
* Hinton, G. E. What kind of graphical model is the brain? In Proc. 19th International Joint Conference on Artificial intelligence 1765–1775 (2005).
* Hinton, G. E., Osindero, S. & Teh, Y.-W. A fast learning algorithm for deep belief nets. Neural Comp. 18, 1527–1554 (2006).
* Bengio, Y., Lamblin, P., Popovici, D. & Larochelle, H. Greedy layer-wise training of deep networks. In Proc. Advances in Neural Information Processing Systems 19 153–160 (2006).
* Ranzato, M., Poultney, C., Chopra, S. & LeCun, Y. Efficient learning of sparse representations with an energy-based model. In Proc. Advances in Neural Information Processing Systems 19 1137–1144 (2006).
* Hinton, G. E. & Salakhutdinov, R. Reducing the dimensionality of data with neural networks. Science 313, 504–507 (2006).
* Sermanet, P., Kavukcuoglu, K., Chintala, S. & LeCun, Y. Pedestrian detection with unsupervised multi-stage feature learning. In Proc. International Conference on Computer Vision and Pattern Recognition http://arxiv.org/abs/1212.0142 (2013).
* Raina, R., Madhavan, A. & Ng, A. Y. Large-scale deep unsupervised learning using graphics processors. In Proc. 26th Annual International Conference on Machine Learning 873–880 (2009).
* LeCun, Y. et al. Handwritten digit recognition with a back-propagation network. In Proc. Advances in Neural Information Processing Systems 396–404 (1990).
* Felleman, D. J. & Essen, D. C. V. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47 (1991).
* Fukushima, K. & Miyake, S. Neocognitron: a new algorithm for pattern recognition tolerant of deformations and shifts in position. Pattern Recognition 15, 455–469 (1982).
* Bottou, L., Fogelman-Soulié, F., Blanchet, P. & Lienard, J. Experiments with time delay networks and dynamic time warping for speaker independent isolated digit recognition. In Proc. EuroSpeech 89 537–540 (1989).
* Simard, D., Steinkraus, P. Y. & Platt, J. C. Best practices for convolutional neural networks. In Proc. Document Analysis and Recognition 958–963 (2003).
* Vaillant, R., Monrocq, C. & LeCun, Y. Original approach for the localisation of objects in images. In Proc. Vision, Image, and Signal Processing 141, 245–250 (1994).
* Nowlan, S. & Platt, J. in Neural Information Processing Systems 901–908 (1995).
* Lawrence, S., Giles, C. L., Tsoi, A. C. & Back, A. D. Face recognition: a convolutional neural-network approach. IEEE Trans. Neural Networks 8, 98–113 (1997).
* Ciresan, D., Meier, U. Masci, J. & Schmidhuber, J. Multi-column deep neural network for traffic sign classification. Neural Networks 32, 333–338 (2012).
* Tompson, J., Goroshin, R. R., Jain, A., LeCun, Y. Y. & Bregler, C. C. Efficient object localization using convolutional networks. In Proc. Conference on Computer Vision and Pattern Recognition http://arxiv.org/abs/1411.4280 (2014).
* Hadsell, R. et al. Learning long-range vision for autonomous off-road driving. J. Field Robot. 26, 120–144 (2009).
* Farabet, C., Couprie, C., Najman, L. & LeCun, Y. Scene parsing with multiscale feature learning, purity trees, and optimal covers. In Proc. International Conference on Machine Learning http://arxiv.org/abs/1202.2160 (2012).
* Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. J. Machine Learning Res. 15, 1929–1958 (2014).
* Sermanet, P. et al. Overfeat: integrated recognition, localization and detection using convolutional networks. In Proc. International Conference on Learning Representations http://arxiv.org/abs/1312.6229 (2014).
* Girshick, R., Donahue, J., Darrell, T. & Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proc. Conference on Computer Vision and Pattern Recognition 580–587 (2014).
* Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proc. International Conference on Learning Representations http://arxiv.org/abs/1409.1556 (2014).
* Farabet, C. et al. Large-scale FPGA-based convolutional networks. In Scaling up Machine Learning: Parallel and Distributed Approaches (eds Bekkerman, R., Bilenko, M. & Langford, J.) 399–419 (Cambridge Univ. Press, 2011).
* Bengio, Y. Learning Deep Architectures for AI (Now, 2009).
* Montufar, G. F., Pascanu, R., Cho, K. & Bengio, Y. On the number of linear regions of deep neural networks. In Proc. Advances in Neural Information Processing Systems 27 2924–2932 (2014).
* Bengio, Y., Simard, P. & Frasconi, P. Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Networks 5, 157–166 (1994).
* Sutskever, I. Training Recurrent Neural Networks. PhD thesis, Univ. Toronto (2012).
* Pascanu, R., Mikolov, T. & Bengio, Y. On the difficulty of training recurrent neural networks. In Proc. 30th International Conference on Machine Learning 1310–1318 (2013).
* Weston, J., Bordes, A., Chopra, S. & Mikolov, T. Towards AI-complete question answering: a set of prerequisite toy tasks. http://arxiv.org/abs/1502.05698 (2015).
* Kavukcuoglu, K. et al. Learning convolutional feature hierarchies for visual recognition. In Proc. Advances in Neural Information Processing Systems 23 1090–1098 (2010).
* Ranzato, M., Mnih, V., Susskind, J. M. & Hinton, G. E. Modeling natural images using gated MRFs. IEEE Trans. Pattern Anal. Machine Intell. 35, 2206–2222 (2013).
* Bengio, Y., Thibodeau-Laufer, E., Alain, G. & Yosinski, J. Deep generative stochastic networks trainable by backprop. In Proc. 31st International Conference on Machine Learning 226–234 (2014).
* Kingma, D., Rezende, D., Mohamed, S. & Welling, M. Semi-supervised learning with deep generative models. In Proc. Advances in Neural Information Processing Systems 27 3581–3589 (2014).
* Ba, J., Mnih, V. & Kavukcuoglu, K. Multiple object recognition with visual attention. In Proc. International Conference on Learning Representations http://arxiv.org/abs/1412.7755 (2014).
* Mnih, V. et al. Human-level control through deep reinforcement learning. Nature 518, 529–533 (2015).
* Bottou, L. From machine learning to machine reasoning. Mach. Learn. 94, 133–149 (2014).
* Vinyals, O., Toshev, A., Bengio, S. & Erhan, D. Show and tell: a neural image caption generator. In Proc. International Conference on Machine Learning http://arxiv.org/abs/1502.03044 (2014).

This essay needs 50 references, but those can be obtained by reading at least 4 to 5 review papers on deep learning in autonomous vehicles. Basically, sum up the concepts explained in the review papers, essentially creating an ultimate review paper. 2 to 3 pages recommended, but the references are the important part.