Journal of Engineering Sciences

Vol 13, Issue 06, JUNE/2022

ISSN NO:0377-9254

VEHICLE PATTERN RECOGNITION USING MACHINE & DEEP LEARNING TO PREDICT CAR MODEL

1Mr.G.D.K.Kishore (Head of the Department) (Department of IT), Kishore.galla1@gmail.com

2Ch.Kavya Sri, B.Tech,Dept of IT, kavyach7788@gmail.com

3B.Amrutha, B.Tech,Dept of IT, amruthabonu0812@gmail.com

4M.Jai Indra Teja, B.Tech,Dept of IT, mi4233715@gmail.com

5K. Hemanth Kumar, BTech,Dept of IT, hkorikani@gmail.com

ABSTRACT: There has been an increase in the application of computer vision techniques. Police departments around the world rely on various types of artificial intelligence to help them catch criminals and avoid big traffic accidents. Although advances in vehicle recognition based on deep learning are being made, there are concerns about the current performance of the state-of-the-art systems. Detection systems, for example, are bound by a wide variety of scales. Vehicle detecting experts are trying to come up with a solution. This thesis proposes a more accurate model for detecting moving vehicles. SVM, KNN, CNN, and linear regression are some of the machine learning algorithms used in this study to predict automobile models based on known inference and patterns. Among other models, CNN consistently outperforms the competition in terms of accuracy.

We'll cover SVMs, KNNs, CNNs, and linear regression in this part.

1. INTRODUCTION

In the last several years, the field of intelligent transportation systems (ITS) has produced a wide range of innovative concepts and solutions. In an ITS, cars, transportation systems, and drivers all work together to achieve these goals. In order to improve public health and safety while traveling, governments, corporations, and academic institutions work together to develop and implement cutting-edge ITS. To a great extent it is anticipated that developments in image processing and computer vision would have a significant impact on ITS technologies including vehicle surveillance and driver assistance systems as well as traffic control and monitoring. Identifying the manufacturer and model of a vehicle can only be done in real time using human interpretation and license plate recognition in traditional car identification systems. The primary difference between the two approaches is that they are both vulnerable to risks and confined by a wide range of variables. With so many different automotive models and make variations, it might be tough to tell them apart. Keeping track of several displays, recording vehicle arrivals and departures, and only searching for a specific make and model is time-consuming and challenging for people. New considerations for automatic identification of such characteristics are opened when it comes to preventing excessive warnings to drivers owing to LPR failure and identifying fake number plates and unlawful vehicles, having the proper tolls in the toll collection systems, and so on. Recent interest in vehicle model and make recognition systems has resulted in an increase in efforts to improve the current system of vehicle detection and classification. Classification of vehicle makes and models, such as Vehicle Makes and Model Recognition, is essential to determining the particular type of vehicle (VMMR). However, despite these approaches' promised effectiveness, only a restricted number of car models and manufacturers are tested.

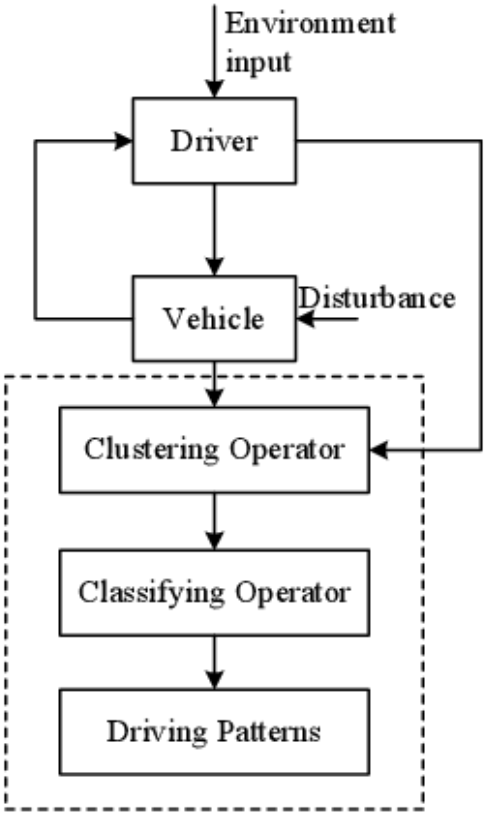

Fig.1: Example figure

Vehicle detection and current traffic conditions are essential for safe, accident-free driving, as well as for autonomous driving and tracking (Jazayeri et al. 2011). Real-time perception is critical to the development of autonomous vehicles. For these reasons, image processing can be used. Image processing has become a popular topic in the photogrammetric discipline, and a wide range of applications benefit from this. This method is used to extract useful information from the raw data of photos. Images can be processed in a variety of ways, including detecting or classifying items in the image and extracting useful information. Finding and categorizing objects in photographs has been the subject of much debate in computer vision and photogrammetry literature (Yang et al., 2019). Aerial and ground-based video can be used to compare new object detection technologies and identify which one is best suited for a specific data set in the future to be launched. And new models will be developed based on a thorough examination of existing approaches, which reveal their strengths and weaknesses. This study aims to investigate object detection methods for aerial and ground-based video. Unmanned aerial vehicle (UAV) footage was also incorporated into the research. Because of the photos' incredibly high quality, it's impossible to tell which vehicles are in them while looking at UAV shots. These topics were also examined in the research. The usefulness of several methods for finding a vehicle in a video frame has been evaluated. Precision and average Intersection over Union were measured through research (IoU). Real-time apps were used to test the procedures on actual data. The findings are shown in tables. Using deep learning to detect vehicles in real time has been found to be possible.

2.LITERATURE REVIEW

2.1 Faster R-CNN: Towards realtime object detection with region proposal networks:

Prediction algorithms are employed by the most advanced object detection networks. SPPnet and Fast R-CNN have demonstrated that the detection networks' region proposal computation is the bottleneck for reducing the time it takes to operate these detection networks. This work introduces RPNs, which share full-image convolutional characteristics with the detection network and hence enable practically cost-free region proposals. RPNs. A RPN simultaneously predicts object borders and objectness ratings at each location. Fast R-CNN can employ the high-quality region proposals produced by end-to-end RPN training. The RPN component tells the unified network where to look when RPN and Fast R-CNN are combined into a single network using the increasingly popular terminology of neural networks with "attention" processes. Only 300 suggestions per image were needed to achieve state-of-the-art object detection accuracy on PASCAL VOC 2007, 2012, and MS COCO datasets utilizing our 5fps (containing all phases) GPU detection approach for the extremely deep VGG-16 model. This year's ILSVRC and COCO competition winners relied heavily on Faster R-CNN and RPN. Everyone can now access and use free open-source software.

2.2 Robotics and vision: a collision Introducing the KITTI data collection.

In the field of mobile robotics and self-driving car research, a VW station wagon has provided a unique source of data. With sensors ranging from high-resolution color and grayscale stereo cameras, to a Velodyne 3D laser scanner, to very precise GPS/IMU inertial navigation systems, we captured six hours of traffic scenarios at 10-100 Hz sampling rates. Real-world traffic circumstances, ranging from highways through rural areas to dense urban areas filled with both static and dynamic things, are portrayed in the

scenarios. Besides supplying rectified and raw image sequences, our data is calibrated, synchronized and time stamped. Our dataset's object labels are used as 3D tracklets in online benchmarks for tasks like as object identification, stereo, optical flow, and more. In this article, we go over our recording platform, data format, and utility options.

2.3 Real-Time, Unified Object Detection: All It Takes Is One Look.

You can use YOLO to detect objects in an innovative method. Object detection classifiers have been repurposed for use in detection. Object detection is done by regressing to geographically distinct bounding boxes and their corresponding class probabilities. It's the first time that a single neural network has been able to accurately forecast class probabilities and bounding boxes from images. It is possible to fine-tune the entire detection pipeline as a single network. Our unified architecture is awe-inspiring in terms of speed. This model processes real-time images at 45 frames per second. Even though the network is scaled down, it still manages to get a mAP that is twice as high as other real-time detection systems. YoLO makes more localization mistakes, but it is less prone than cutting-edge systems to foresee false detections when nothing exists. Finally, YOLO is able to represent a large variety of objects. The Picasso Dataset and the People-Art Dataset both show that it outperforms all other detection techniques, including DPM and R-CNN.

A single-shot detector with many boxes.

Photos are recognized by a single deep neural network. Bounded box output space is divided into a sequence of boxes with varying aspect ratios and scaling depending on the feature position. This strategy is referred to as SSD. Each object category in each default box is given a score by the network, which generates a form adjustment for the box in order to better match the object. There are various feature maps, each with a distinct level of detail, to accommodate a wide range of items. Our SSD solution is easier to implement since it does not involve the production of object proposals or subsequent pixel or feature resampling. SSD may be trained rapidly and readily into detection systems thanks to this method. SSD is as accurate as techniques that involve an additional object proposal phase on the PASCAL VOC, MS COCO, and ILSVRC datasets while also being substantially faster and providing a unified framework for both training and inference. Even when the input image size is less, SSD's accuracy is greater to that of other single-stage techniques. When compared to other storage technologies, SSDs exceed them all by a wide margin. 58 frames per second for 300300 input and 75.1 percent for 500500 input for the R-CNN model in VOC2007 tests, respectively.

For ImageNet categorization, we use deep convolutional neural networks.

In the ImageNet LSVRC-2010 competition, we trained our convolutional neural network to classify the 1.2 million high-quality photographs into 1000 distinct categories. We achieved error rates of 37.5 percent and 17.0 percent, significantly lower than the previous state of the art, as evaluated by the test data. The neural network's 60 million parameters and 650,000 neurons are made up of five convolutional layers, some followed by max-pooling layers, and three fully connected layers with a final 1000-way softmax. To reduce training time, we used non-saturating neurons and a powerful GPU convolution function. It was shown that reducing overfitting in the fully-connected layers might be accomplished by employing a regularization algorithm called "dropout." In the ILSVRC-2012 competition, a variation of this model came in first place with a top-5 test error rate of 15.3 percent, beating out the second-best entry's 26.2 percent.

Deep Convolutional Networks for Large-Scale Image Recognition, Section 2.6

In the context of large-scale image recognition, convolutional network depth is investigated as a consideration. A study of rising-depth networks using an architecture with extremely small (3x3) convolution filters shows that raising the depth to 16-19 weight layers can be achieved a significant improvement over the prior-art configurations. To achieve first and second place in the 2014 ImageNet Challenge, our team used these insights to their work. Our representations' cutting-edge results can be applied to a wide variety of datasets. Research into deep visual representations in computer vision may continue with the release of two of our best-performing ConvNet models.

3.IMPLEMENTATION

Since the beginning of vehicle detection, researchers have proposed a variety of traditional methods. Approaches are judged on the basis of their handcrafted qualities. The HOG and Haar-

like properties of the histogram are the most commonly used. One of the first real-time detectors to achieve a competitive degree of accuracy is a cascaded detector. The part-based model technique has two well-known models: support vector machines (SVMs) and deformable part-based models (DPMs) [15].

In the current system, there are a number of issues:

a drop in production

First and foremost, there are concerns about the length of time it takes.

What I'm proposing is the mechanism outlined below:

Although advances in vehicle recognition based on deep learning are being made, there are concerns about the current performance of the state-of-the-art systems. Detection systems, for example, are bound by a wide variety of scales. Vehicle detecting experts are trying to come up with a solution. This thesis proposes a more accurate model for detecting moving vehicles. SVM, KNN, CNN, and linear regression are some of the machine learning algorithms used in this study to predict automobile models based on known inference and patterns. Among other models, CNN consistently outperforms the competition in terms of accuracy.

The system's advantages are as follows:

Foresight that's better

productivity at a high level

Third, it saves you time.

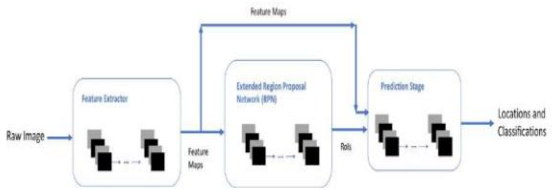

Fig.2: System architecture

A variety of machine learning approaches are being used in this effort to forecast automobile models based on known inference factors and pattern features. The photos below show some of the images from the Stanford automobile dataset that we used in this research.

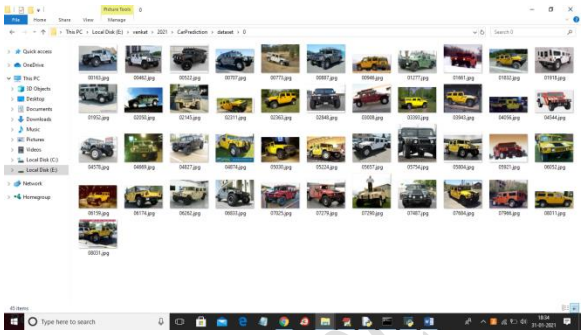

Fig.3: Dataset

4. ALGORITHM

SVM (KNN), CNN (CNN), and Linear Regression are examples of regression methods.

The supervised machine learning approach of k-nearest neighbors (KNN) can be used to address classification and regression problems. Using a collection of data, a machine learning model may predict the outcome of the experiment. KNN is one of the most basic forms of machine learning algorithms for categorization. The classification of data points is based on the classification of their neighbors.

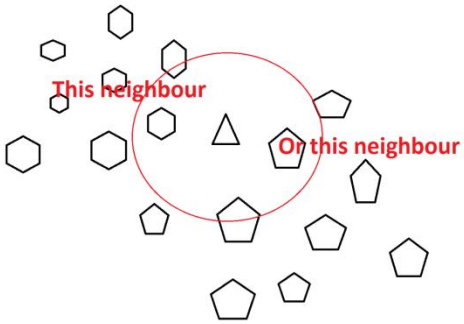

Fig.4: KNN model

In order to classify new data points, KNN uses the similarity measure of previously stored data points. A tomato and banana dataset might be given as an example of how this might be done. It will be possible to save similar measurements in KNN, such as the shape and color. New objects are compared to existing ones to see if their color (red or yellow) and shape are the same or different.

Algorithms for machine learning It is possible to forecast the likelihood of an event using LINEAR REGRESSION and supervised learning. It is used to perform a regression task. Regression uses a set of independent variables to calculate a target variable's anticipated value. Correlation analysis and forecasting are two of this tool's most prevalent applications.

Fig.5: Linear regression model

SVM:

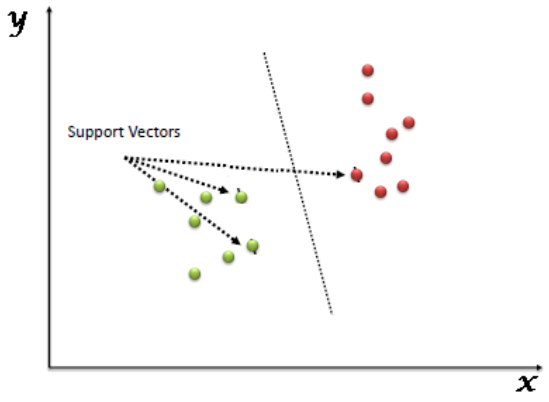

Fig.6: SVM model

FThe kernel trick, for instance, tampers with the data you supply to it. Classifiers can no longer recognize any of the input data when a large number of features are introduced into them. The process is quite similar to unraveling a DNA strand. A simple vector may appear at first glance, but the kernel technique transforms it into a far more complex set of data that cannot be understood by looking at it in a spreadsheet.. The SVM method is better equipped to build an ideal hyperplane when it has access to a larger dataset.

We explain how to detect picture tampering using deep learning. The suggested algorithm is divided into two parts: learning and testing. Sequentially illustrated learning and testing are shown. The original image is altered in order to create a new image. There is a training and a testing set built from the image. The proposed CNN model is then used to analyze the training data. Weights are then recalculated using back propagation. Testing the accuracy of a model is done by applying the trained model to a test set. When it comes to image processing, CNN is a deep learning model inspired by animal visual cortex organization that uses a grid pattern to automatically and adaptively learn spatial hierarchies of attributes. Convolutionary, pooling, and fully-connected layers make up the majority of the building pieces that go into creating a CNN. Classification is based on the features derived from this completely linked layer. Convolution and pooling are used to extract features in the first two layers. A layer of convolution is used in CNN, which is a type of linear operation composed of a succession of mathematical operations, such as convolution.

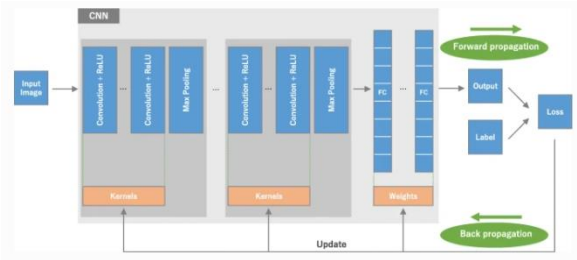

Fig.7: CNN model

4. EXPERIMENTAL RESULTS

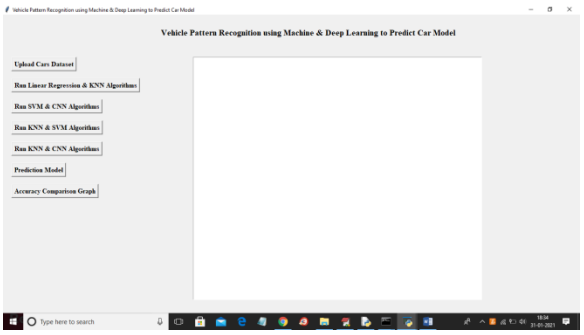

Fig.8: Upload cars dataset

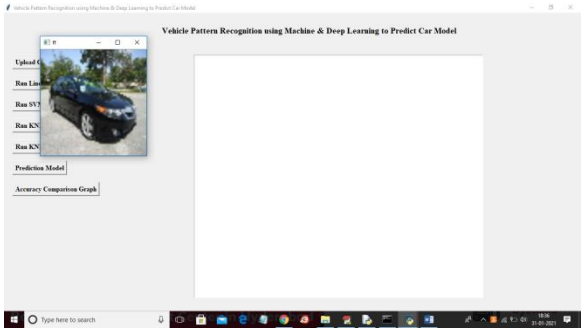

Fig.9: Dataset folder

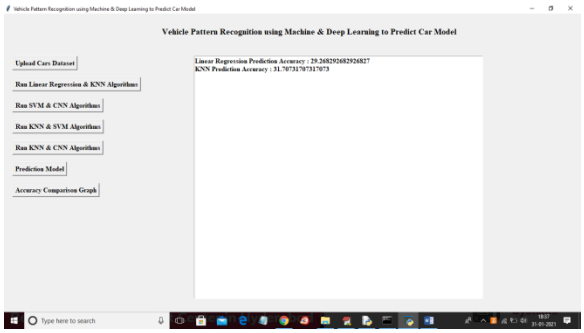

Fig.10: Run linear regression &knn algorithm

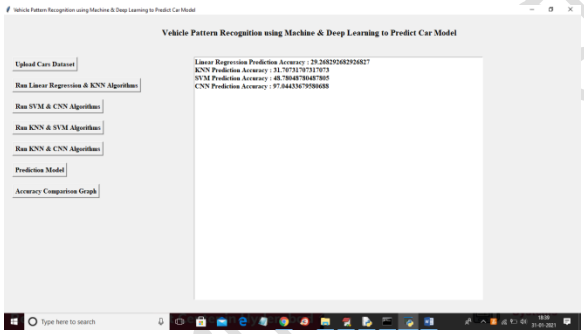

Fig.11: Run svm&cnn algorithm

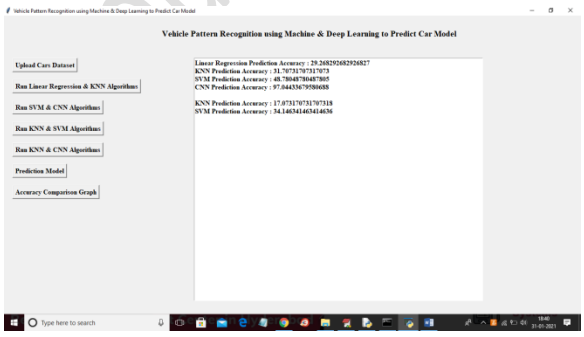

Fig.12: Run knn&svm algorithm

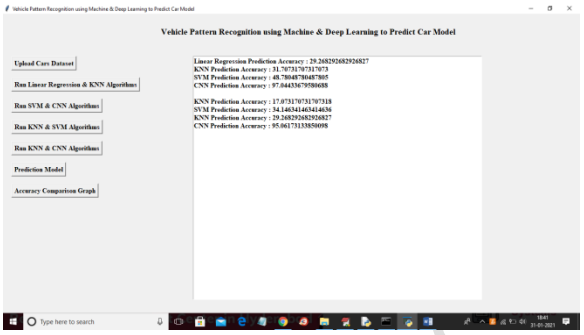

Fig.13: Run knn&cnn algorithm

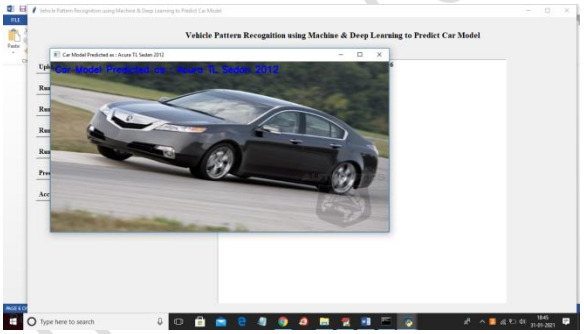

Fig.14: Prediction model

Fig.15: Accuracy comparison graph

6. CONCLUSION

This thesis introduces and improves the essential components of deep learning for vehicle detection. SVM, KNN, CNN, and linear regression are just a few of the machine learning techniques we're employing in this research to make predictions about future car models. Deep learning algorithms excel in detecting automobiles, as evidenced by test results and real-world applications. The new model has resulted in an increase in performance and a reduction in testing time.

Future-Oriented Insight

For example, in the field of image processing, convolutional neural networks perform well and are new. When working with huge datasets, CNN outperforms other machine learning approaches. We've come up with an outstanding model for two-dimensional vehicle detection, which means that we compute the two-dimensional bounding box of a specific object. The use of three-dimensional vehicle detection, which is an excellent technology for autonomous driving, is on the rise. Methods like YOLO or Faster R-CNN are frequently used to compute the 2D bounding box and then to identify the dimensions and orientation of the items that are employed in locating 3D objects while detecting vehicles in three dimensions (3D). A lot of people adopt this approach.

REFERENCES

[1] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. "Faster R-CNN: Towards realtime object detection with region proposal networks." In Advances in Neural Information Processing Systems, pp. 91-99, 2015.

[2] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. "Vision meets robotics: The KITTI dataset." The International Journal of Robotics Research, 32, no. 11, 2013: 1231-1237.

[3] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick and Ali Farhadi. "You Only Look Once: Unified, Real-Time Object Detection." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.

[4] Wei Liu, DragomirAnguelov, DumitruErhan, Christian Szegedy, Scott E. Reed, ChengYang Fu and Alexander C. Berg. "SSD: Single Shot MultiBox Detector." European Conference on Computer Vision (ECCV), 2016.

[5] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. "ImageNet classification with deep convolutional neural networks." In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12), F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Eds.), Vol. 1. Curran Associates Inc., USA, 1097-1105.

[6] Karen Simonyan and Andrew Zisserman. "Very Deep Convolutional Networks for LargeScale Image Recognition." Computing Research Repository (CoRR), abs/1409.1556, 2014.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun. "Deep Residual Learning for Image Recognition." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.

[8] Christian Szegedy, Wei Liu, YangqingJia, Pierre Sermanet, Scott E. Reed, DragomirAnguelov, DumitruErhan, Vincent Vanhoucke and Andrew Rabinovich. "Going deeper with convolutions." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 1-9.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun. "Identity Mappings in Deep Residual Networks." European Conference on Computer Vision (ECCV), 2016.

[10] Andrej. 2015. "Neural Networks Part 1: Setting Up the Architecture." Notes for CS231n Convolutional Neural Networks for Visual Recognition, Stanford University.39

[11] Brody Huval, Tao Wang, SameepTandon, Jeff Kiske, Will Song, Joel Pazhayampallil, MykhayloAndriluka, Pranav Rajpurkar, Toki Migimatsu, Royce Cheng-Yue, Fernando Mujica, Adam Coates, and Andrew Y. Ng. An empirical evaluation of deep learning on highway driving. Computing Research Repository (CoRR), abs/1504.01716, 2015.

[12] Pierre Sermanet, David Eigen, Xiang Zhang, Michaël Mathieu, Rob Fergus, and Yann LeCun. "Overfeat: Integrated recognition, localization and detection using convolutional networks." arXiv preprint arXiv:1312.6229, 2013.

[13] Girshick Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587, 2014.

[14] NavneetDalal and Bill Triggs. "Histograms of oriented gradients for human detection." In International Conference on Computer Vision & Pattern Recognition (CVPR'05), vol. 1, pp. 886-893. IEEE Computer Society, 2005.

[15] Johan AK. Suykens, and JoosVandewalle. "Least squares support vector machine classifiers." Neural Processing Letters, 9, no. 3, 1999: 293-300.