ML for Autonomous Driving Car

Dr. Rajiv Misra, Professor
Dept. of Computer Science & Engg.
Indian Institute of Technology Patna
rajivm@iitp.ac.in

Preface

After completion of this lecture you will know the following:

- Understanding of Autonomous Vehicles
- Role of Edge computing in Automotive Industry
- How ML is trained in Self-driving cars?
- Use Case of LSTM model for self-driving cars

Autonomous Vehicles: Introduction

Autonomous vehicles (AVs) have attracted a significant amount of interest in recent years. According to a report released by the US state Department of Transportation, "Self-Driving-Cars can reduce 90% of Traffic Deaths".

A big chunk of major Automobile companies is trying to develop Self-Driving-Cars. Some big players are Tesla, Waymo, even Google is developing Self Driving Cars which has no presence in the automobile sector, have invested a huge amount of money, manpower and engineering capabilities in developing such systems.

Designing policies for an autonomous driving system is particularly challenging due to demanding performance requirements in terms of both making safe operational decisions and fast processing in real-time.

Edge Computing in Automotive

(AVs)

Historically, the adoption of computing (be it cloud or edge) and software in automotive has trailed the in-general adoption in other industries.

Cloud computing has been around for a while in many industries and many forms. But, vehicle telematics became one of the top use cases adopted in automotive somewhere in 2008.

Autonomous vehicles AV Connected Cars → Communication { Vehicle to vehicle comm, Vehicle to any

Connected vehicles will continue to evolve at an exponential rate with V2V and V2X communication. This generates a large volume of data (every connected vehicle will generate data up to 4TB/day). How to handle, process, analyse the large amounts of data and make critical decisions quickly and efficiently?

Automobile makers are focused on leveraging edge computing to address these ever-evolving challenges. A group of cross-industry global players has formed the Automotive Edge Computing Consortium (AECC) to drive best practices for the convergence between the vehicle and computing ecosystem.

Edge Commu – Latency – milliseconds

When driving a vehicle, milliseconds matter. Autonomous vehicles are no different, even though it may be your AI that drives them. AI = data + compute, and you want your compute to be as close to your data as possible. Enter edge computing.

Edge Computing: Self-Driving Car Sensors

Given its real-time data processing capabilities, edge computing has naturally established itself as a pillar in autonomous vehicle technology. However, this data isn't generated by the computer but rather by the multitude of sensors that comprise an autonomous vehicle's peripheral "eyes" and "ears."

tesla

Sensor topology can vary widely amongst autonomous vehicles, even within the same sector.

Most self-driving sensors are fundamentally similar - they collect data about the world around them to help pilot the vehicle. For example, the Nuro vehicle contains cameras, radar, Lidar, and thermal cameras to provide a complete, multi-layered view of the vehicle's surroundings.

① Stop
② acceleration Pedal

Currently, a Tesla utilize eight cameras,12 , and a forward radar system, but rely much more heavily on camera visuals than Nuro vehicles. Google's Waymo Driver primarily relies on Lidar and uses cameras and radar sensors to help map the world around it.

Self-Driving Car: Requirements

Autonomous driving vehicles require two in-vehicle computing systems. One computer processes a large amount of sensed data and images collected by cameras and sensors. And a second computer to analyze processed image data and make intelligent and quick decisions for the vehicle.

- **Pre-processing collected data.** Autonomous vehicles have video cameras and a variety of sensors like ultrasonic, LiDAR, and radar to become aware of their surroundings and the internals of the vehicle. This data coming from different vehicle sources must be quickly processed through data aggregation and compression processes. An in-vehicle computer needs to have multiple I/O ports for receiving and sending data.
- **Secure network connectivity.** The in-vehicle computing solution must remain securely connected to the Internet to upload the pre-processed data to the cloud. In this case, having multiple wireless connections for redundancy and speed is crucial. High-speed connectivity is also vital for continuous deployments of vehicle updates or "push" updates like location, on-road conditions, and vehicle telematics.
- **High-performance computing.** Autonomous vehicles may generate approximately 1 GB of data every second. Gathering and sending a fraction of that data (for instance, 5 minutes of data) to a cloud-based server for analysis is impractical and quite challenging due to limited bandwidth and latency. Autonomous driving systems shouldn't always rely on network connectivity and cloud services for their data processing. Self-driving vehicles need real-time data processing to make crucial quick decisions according to their surroundings. In-vehicle edge computing is essential for reducing the need for network connectivity (offline decision-making) and for increasing decision-making accuracy.a

How Machine Learning Trains AI in Self-Driving Cars

The value of the sensor data collected in all self-driving cars and vehicles depends on the compute methodologies downstream of the sensors themselves. In many ways, the most valuable intellectual property of companies like Tesla, Waymo, Aurora Innovations, and Nuro is the software and data infrastructure built to process and action the sensor data.

Today, all autonomous vehicles on the road utilize edge computing AI programs, which are often trained using data center machine learning models. Autonomous car machine learning models are only made possible by the incredible computing power of modern data centers capable of hundreds of petaflops.

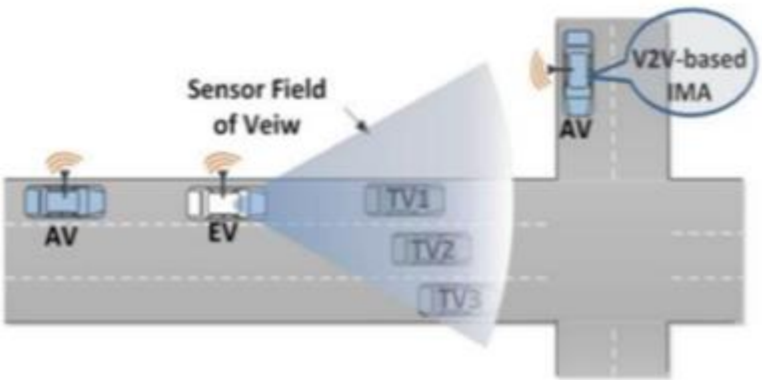
The computing requirements of these vast machine learning models well exceed the computing power of edge computers. Given this information, data centers are often used to form algorithms deployed for edge.

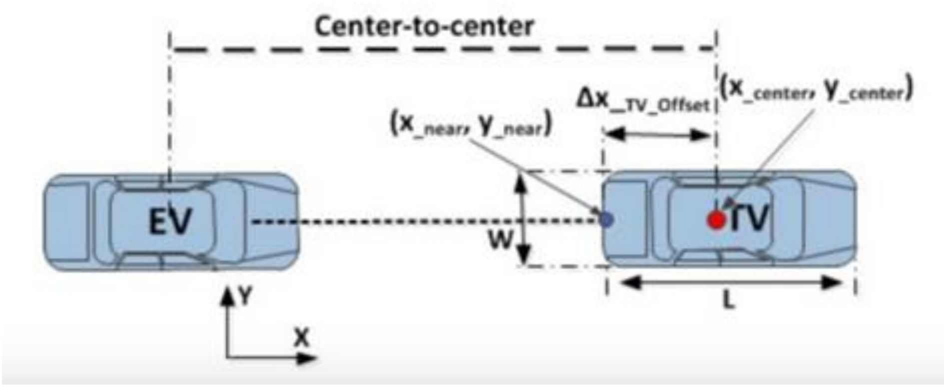
The problem of self-driving-car can be seen as a Regression Problem.

Training an AI algorithm is similar; it takes hundreds of compute hours on a high-power data center. Yet once that algorithm is learned, it can quickly and accurately utilize that algorithm using much less computing power.

Machine learning in autonomous driving

Kalman Filter, In real-life autonomous driving, the machine will deal with the same information from different sensors, such as Lidar, Radar, MEC signals and V2V Communications. This information will always have discrepancies with each other, and Kalman filter can help us to get a relatively reliable answer according to these two sets of information.





Machine learning in autonomous driving

Lidar, Radar, and Cameras ML is an important part of autonomous driving. A self-driving vehicle usually has multiple sensors, including cameras, lidar, and radar sensors. The machine learning module will tell the vehicle what to do with different information. For example, the car needs to stop when there were pedestrians, and the machine must be able to tell the difference between actual pedestrians and pictures of human. Additionally, distances camera sets cannot precisely measure distance or work at night. Lidar sensors usually emit high-frequency signals, and those high-frequency signals could be used for positioning and 3D modelling, being able to tell the difference between actual human and pictures of human. Radar is a low energy cost solution for positioning because the radio wave it emits is usually with low frequency. Low-frequency wave cannot depict the detailed 3D shape, but it is enough for positioning. However, cameras are still needed because neither lidar nor radar can identify colors. signal

Vehicle-to-Vehicle Communication. Communication (V2V) technology can increase the accuracy of autonomous driving prominently. When multiple cars are sharing their information, they can calibrate according to their relative positions.

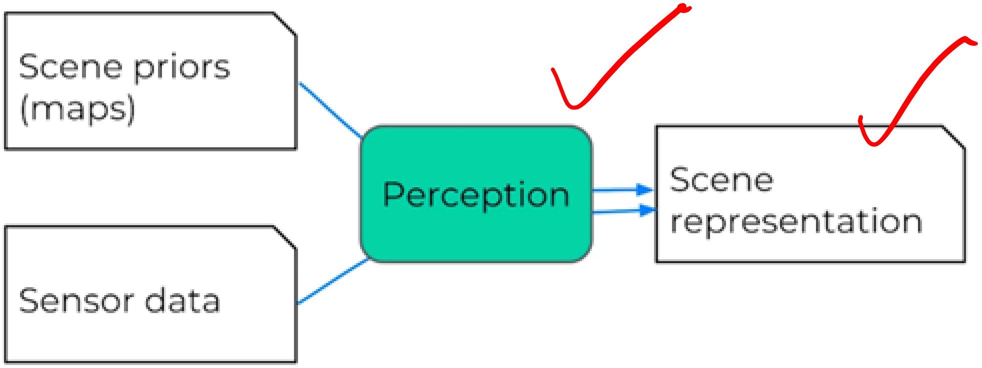
Key component of ML for self driving cars

Perception: a core element of what the self-driving car needs to build an understanding of the world around it using two major inputs:

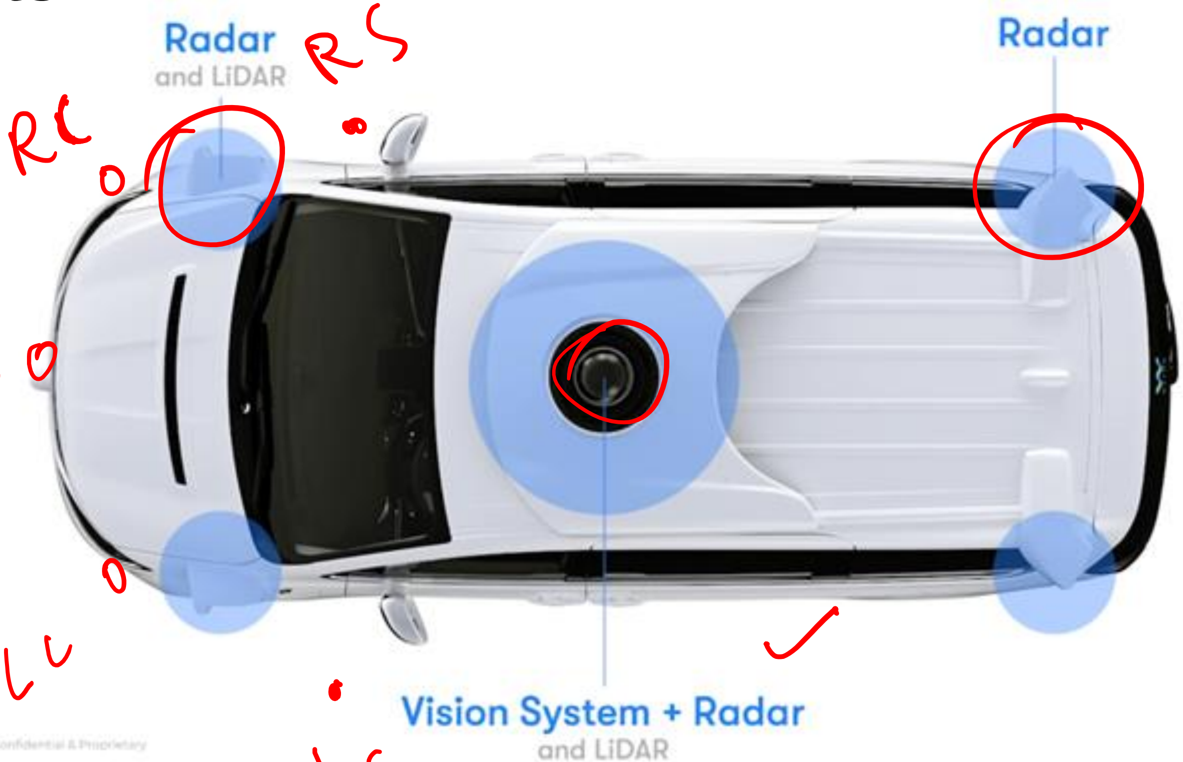
- Scene Prior, and
- Sensor Data

Scene Prior, is prior on the scene. For example it would be a little silly to recompute the actual location of the road, interconnectivity of the intersections of every intersection. Things you can pre-compute in advance and save your onboard computing for all the tasks that are more critical which is often referred to as the mapping exercise.

Sensor, the signal that's going to tell you what is not like, what you mapped and the things like traffic light right or green, where are the pedestrians and the cars what are you doing.

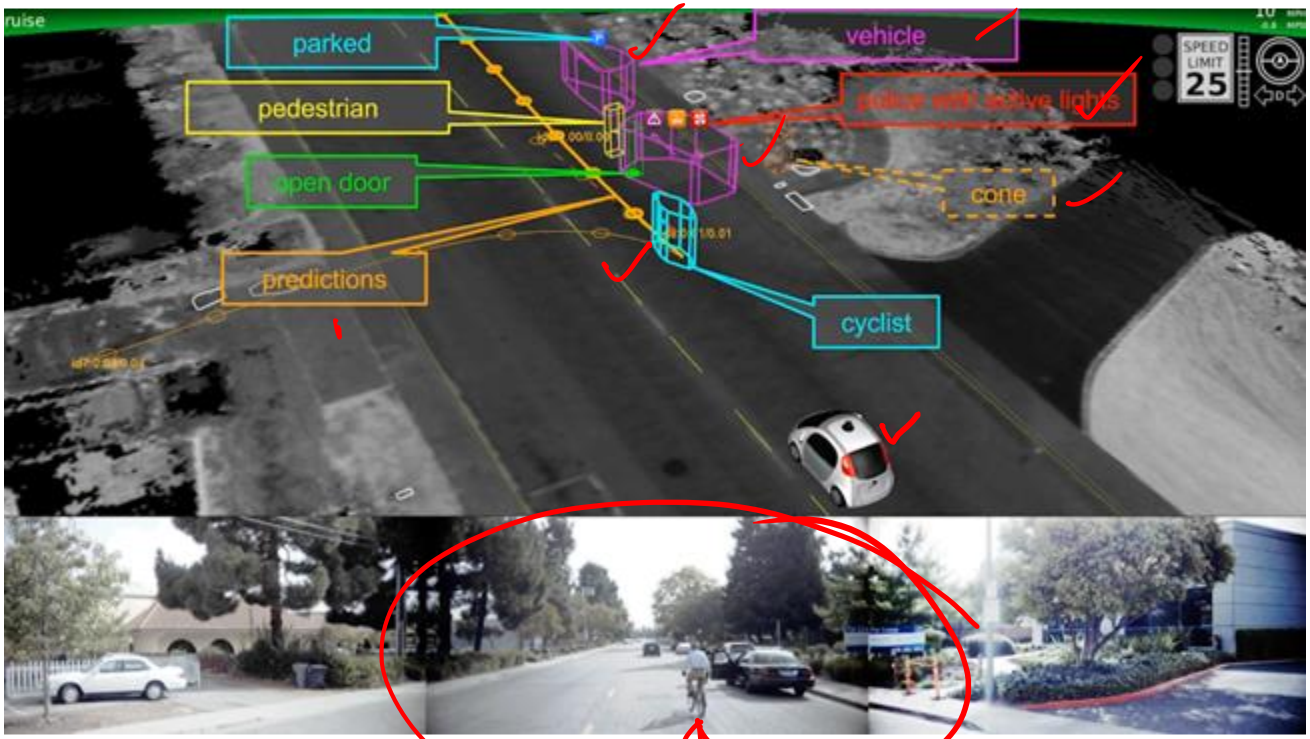


Components


Radar
and LiDAR
Radar
Vision System + Radar
and LiDAR
RS
RC
FC
LC
LS

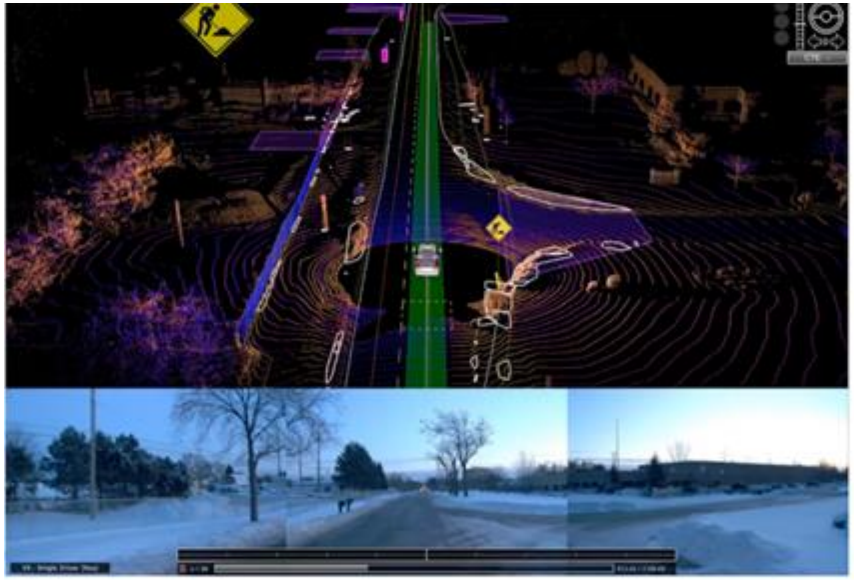
Key component of ML for self driving cars

Scene Representation



Key component of ML for self driving cars

Perform semantic object segmentation



Key component of ML for self driving cars

Perform finer classification of objects

School Bus

Police Car

Fire Truck

Ambulance

Key component of ML for self driving cars

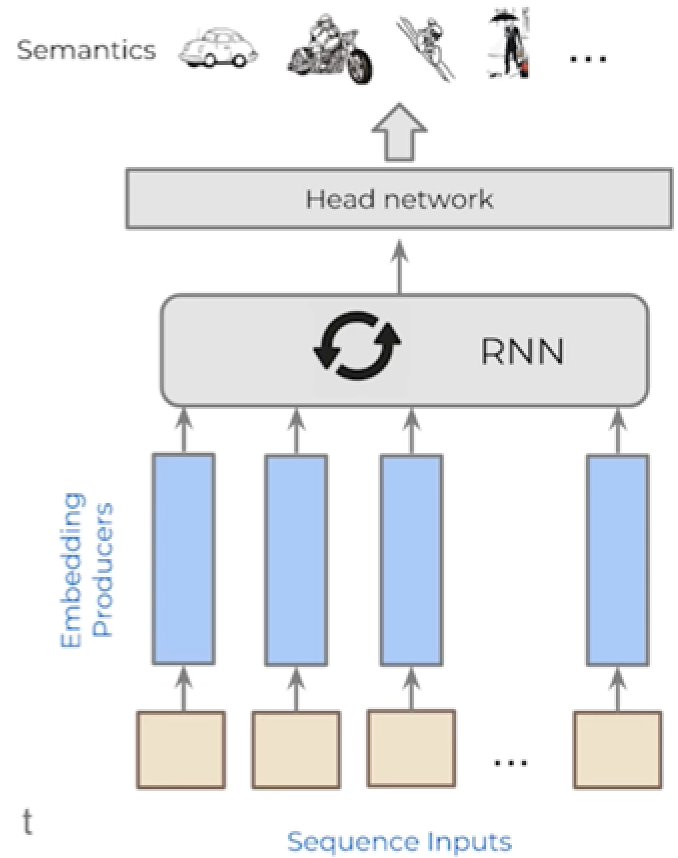
Time tracking using embeddings/RNN

Now the vector representations of different objects will be tracked over time.

A common technique that you can use is a recurrent neural networks that essentially are networks that will build a state that gets better and better as it gets more observation sequential observations of for the pattern.

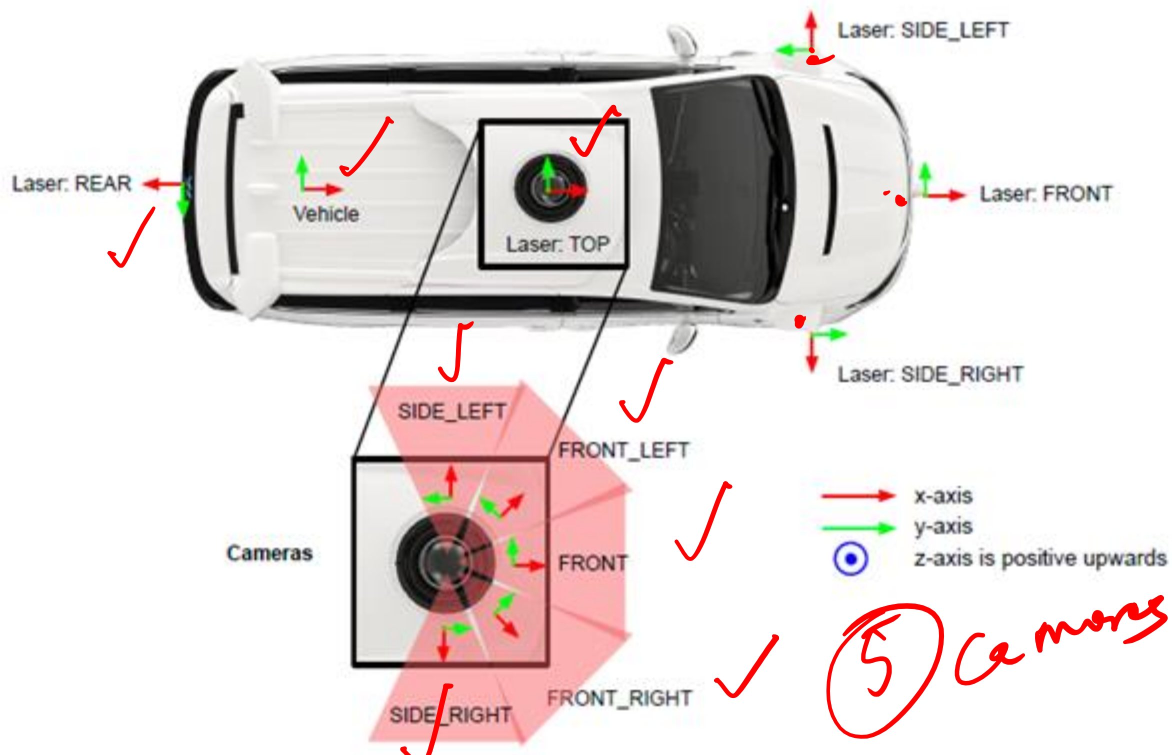
Once semantic representation and coding in an embedding for the pedestrian, the car under it and the model will track that over time and build a state of a good understanding of what's going on in the scene.

The vector representation combined with recurrent neural networks is a common technique to achieve this.



Data for training ML models in Self-Driving Cars

Waymo Open Dataset is the largest, richest and most diverse AV datasets ever published for academic research Sun et al. (2019). This dataset, collected from Waymo level-5 autnomous vehicles in various traffic conditions, comprise radar, lidar and camera data from 1000 20-second segments with labels. We will introduce details about the Waymo dataset, as well as how the data is preprocessed before being fed into several machine learning models.



5 cameras

Data for training ML models in Self-Driving Cars

Labels refer to kinematics and spatial parameters of objects, which are represented as bounding boxes. Specifically, one kind of labels, type, is classified into pedestrian, vehicle, unknown, sign and cyclist categories. Detailed information is provided for each label, among which we especially pay attention to the coordinates of the bounding boxes, velocities v, and accelerations a in the subsequent feature extraction step.

Coordinate Systems three coordinate systems are provided in this dataset: global frame, vehicle frame, and sensor frame. Some raw features are represented in unintended coordinate systems. In order to maintain consistency, it is crucial to transform data into the correct coordinate system. The dataset also provides vehicle pose VP, a 4 × 4 row matrix, to transform variables from one coordinate system to another.

Acceleration Computation Because one's instant acceleration of is not directly available in the dataset, the "ground truth" for training and evaluation needs to be computed by velocity differences.

Data for training ML models in Self-Driving Cars

Data Size: According to the data format, 1000 segments are packed into multiple compressed files (tars) with a size of 25 GB each. In our experiments, 32 training tars are used as the training set and 8 validation tars are used as the testing set. The total number of videos extracted from the segments is 45000.

Image embedding there are five cameras installed on the AV, facing towards front, front-left, frontright, side-left, and side-left respectively. These images reflect the time-series information of the moving vehicle with relatively smoother variation than numerical data, which helps to prevent spiky prediction between consecutive frames.

Use Case: LSTM model for self driving cars

Acceleration Prediction model

Basic Model with 12 Features, One of the straightforward ways to build the acceleration prediction model is to treat 12 basic features as the input of the model. The ”encoder-decoder” architecture proposed for trajectory prediction in SS-LSTM is a suitable architecture for the acceleration prediction problem as the acceleration curve is a trajectory based on past experiences.

encoder decoder → accelerations for next 5 frames

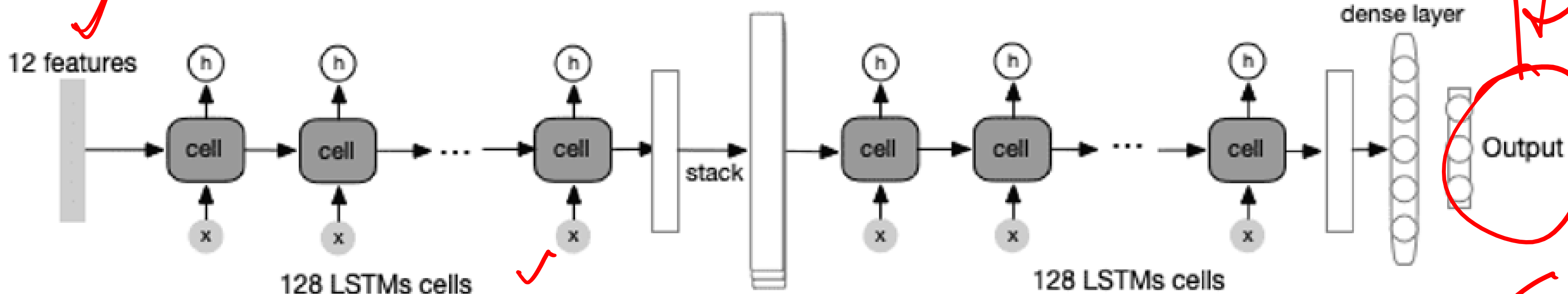
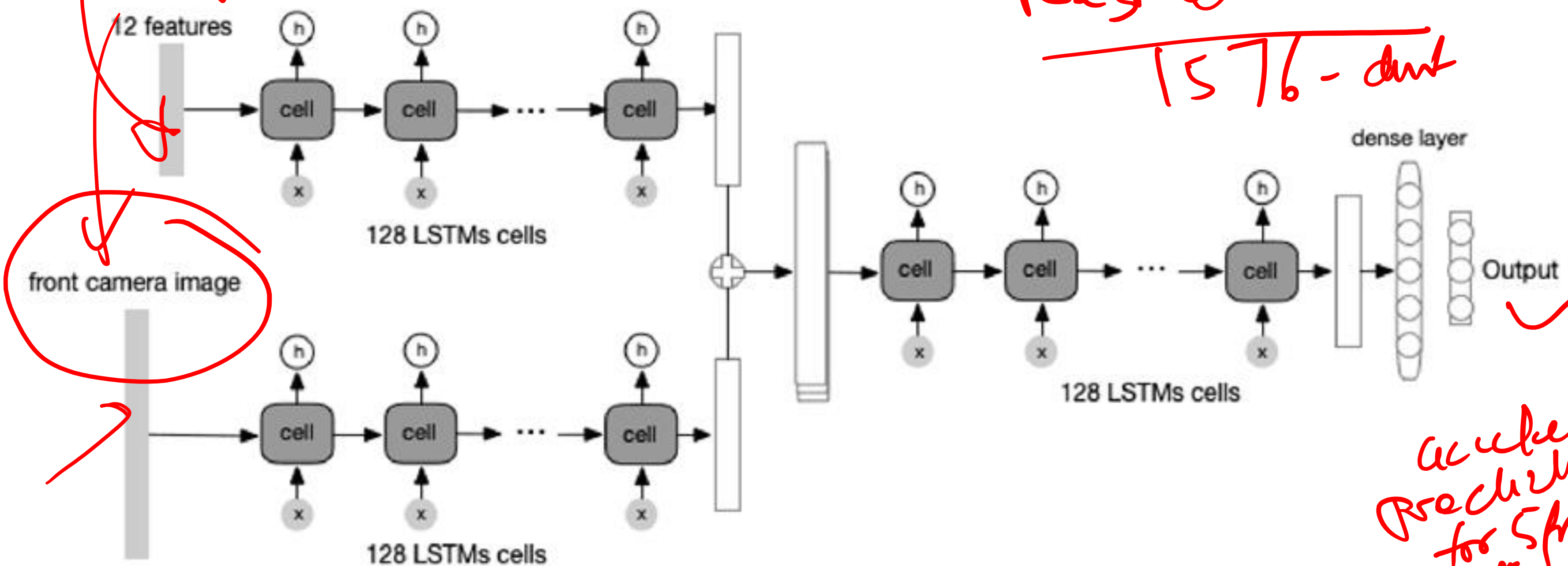


Figure 4: Given one video clip with a frame-length of 10, the input is the vector consists of 12 features from these 10 frames. The output is the acceleration for the next 5 frames starting from the end of the video clip. The ”encoder” module contains 128 LSTM-cells and the ”decoder” module contains 128 LSTM-cells

Use case: LSTM model for self driving cars

Advanced Model with Image Inputs, The architecture of such an advanced model is similar to the previous basic model. An "encoder-decoder" structure is maintained to learn the information hidden in the input features. The difference is that the front camera images are treated as additional inputs.



ResNet152V2

1576-dim

acceleration prediction for 5 frames

Figure 5: Noticed that the "image" input is actually a vector containing key image content. The first channel input is the 12 features for the observed 10 frames. The second channel input is the 1576-dimensional representation of front camera images from these 10 frames. Such representation is extracted from the second last output of a pre-trained Resnet152v2. The output is the acceleration for the future 5 frames. The "encoder" module contains 128 LSTM-cells and the "decoder" module contains 128 LSTM-cells

Use Case: LSTM model for self driving cars

Advanced Model with more Image Inputs

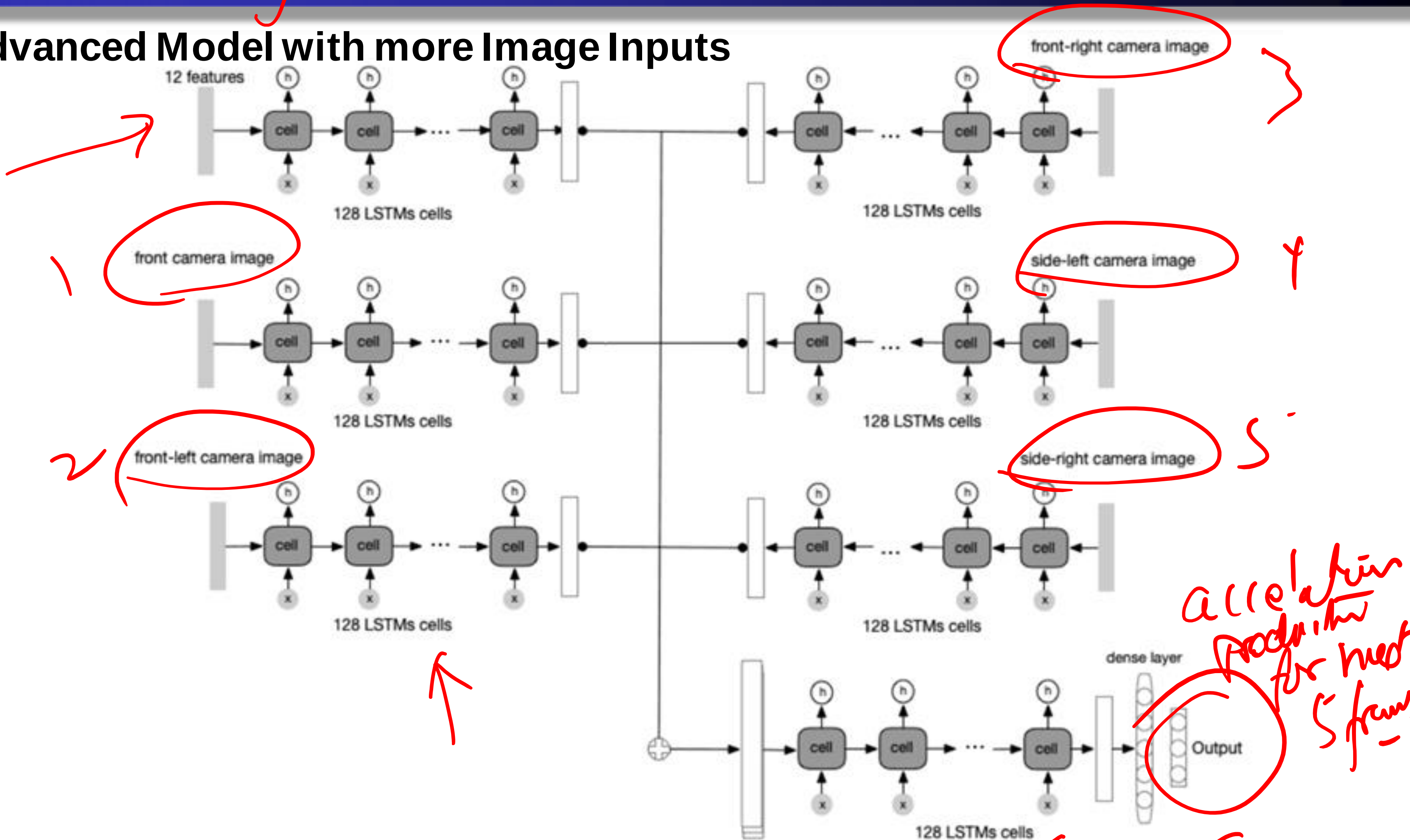


Figure 6: The first channel input is the 12 features for the observed 10 frames. The rest of the input is the 1576-dimensional representation of camera image from different views in the observed 10 frames. The output is the acceleration for the future 5 frames. All "encoder" modules contain 128 LSTM-cells and all "decoder" modules contain 128 LSTM-cells

Use Case: LSTM model for self driving cars

Comparison of results with other state-of-the-art methods

Models	MAE X	MAE Y
NN	0.4014	0.4312
CNN	0.3272	0.3123
NN+CNN	0.2985	0.2802
XGBoost	0.3502	0.3537
Light Gradient Boosting	0.3459	0.3444
Stacked Linear Regressor	0.3623	0.3222
LSTM with 12 features	0.3179	0.3088
LSTM with front camera	0.1379	**0.1278**
LSTM with all cameras	**0.1327**	0.1363

Future trend of autonomous driving

Like other intelligent industries of IIoT, autonomous driving is also reducing the total energy consumption. Gasoline has been the primary fuel for all kinds of vehicles, and natural gas storage only has about 52 years left, with current consumption levels. If the natural gas demands increased, natural gas could run out faster. So, the energy crisis is existing all the time.

First of all, the rise of autonomous driving cars can improve the energy efficiency of private-owned cars. Usually, an average family car can reach its maximum speed at about 200 to 250 km/h, but the city's usual speed limit is usually about 60km/h. That means the engine displacement of nowadays cars are mostly excessive. However, high engine displacement is necessary because faster cars are always safer because driver can overtake or change lane faster. If autonomous vehicles took the places of private-owned vehicles. In that case, it is pointless to use bigger and faster cars because autonomous driving cars are much more reliable than human drivers.

Future trend of autonomous driving

Secondly, auto-driving vehicles could reduce the natural gas dependency. As this paper mentioned before, smaller cars do not need potent energy resource, and electricity will be enough for most auto-driving vehicles. The popularization of auto-Driving cars is also an excellent opportunity to accept renewable energy over traditional energy sources, which will do good to the global climate as well.

Last but not least, when autonomous driving vehicles replaced private cars, parking issues will be solved, people will have bigger house and living areas because no garage is needed. There will be no traffic congestion as routes will be pre-scheduled to ensure efficiency. Long-distance deliverance will be more reliable because the auto-driving vehicle will never be tried.

Lecture Summary

- Different concepts of Autonomous Vehicles
- How Edge computing is important in Automotive Industry?
- How ML is trained in Self-driving cars?
- Use Case of LSTM model for self-driving cars


THANK YOU!