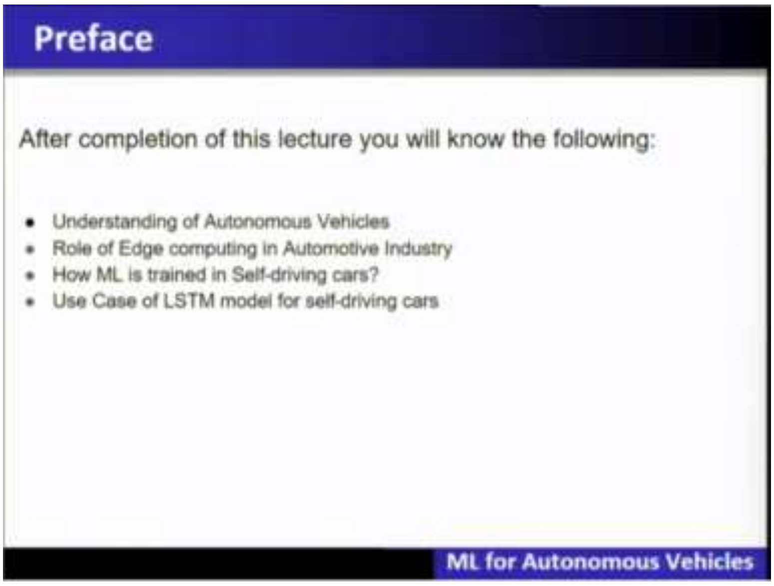
Foundation of Cloud IoT Edge ML
Professor Rajiv Misra
Department of Computer Science and Engineering
Indian Institute of Technology, Patna
Lecture No. 22
ML for Autonomous Driving Car

I am Doctor Rajiv Misra from IIT, Partner. The topic of this lecture is Machine Learning for Autonomous Driving Cars.

(Refer Slide Time: 0:23)



In this particular lecture, we will cover the following topics: understanding of autonomous vehicle technology, then role of IoT and edge computing in automotive industry such as autonomous cars, how machine learning model is used to train the self-driving cars intelligence and use case of LSTM based model training for self-driving cars, functionalities.

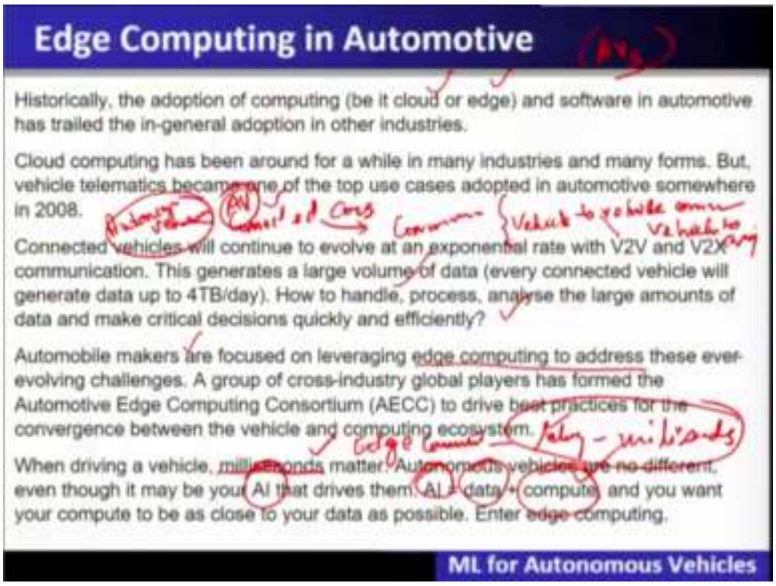
So, this particular lecture we consider that self-driving car is a example of an IoT that is the car is an internet of thing device which is also connected. And as you know that with the advent of edge computing, it is now possible that the driving decisions can be made with the help of AI and therefore, it is about self-driving cars. How LSTM model can be used for giving the intelligence for self-driving cars.

(Refer Slide Time: 1:34)

So, let us introduce this particular aspect of IoT device that is called autonomous vehicles. So, here in the picture you can see these are the different pictures of Tesla autonomous vehicle and Waymo autonomous vehicles. So, autonomous vehicles nowadays is under the development and have attracted the significant amount of interest in the recent years. So, according to the report released by US State Department of Transportation, self-driving cars can reduce 90 percent of traffic deaths. So, therefore, it becomes a boom on the road traffic scenarios.

So, the big chunk of major automobile companies are trying to develop self-driving car and some of the E players in building the self-driving car are Tesla, Waymo, even though Google is also developing the self-driving cars and has invested lot of money, manpower and developing the capabilities towards this self-driving car or autonomous vehicle technologies. Now designing policies for autonomous driving system is a challenging, but due to the demand, performance requirements in terms of both the safety and the operational decisions and fast processing in real time, lot of development is happening around the interesting development that is in the form of autonomous vehicles.

(Refer Slide Time: 3:09)



Now, edge computing will facilitate this autonomous vehicles AVs. So, trace down and see the topic of our lecture, how it is connected in realizing this autonomous vehicle design. So, historically adoption of computing that is whether it is a cloud or the edge and the software in automotive technologies has in general adopted in other industries also. So, therefore, cloud computing, as you know, that it has been around for a while in many industries, many firm. But, vehicle telemetrics also has become one of the top use case nowadays to be adopted.

So, connected vehicles will be continued to evolve at an exponential rate with vehicle to vehicle communication and vehicle to infrastructure, everything communication. So, this V2V means vehicle to vehicle communication and V2X means vehicle to everything. So, these are two forms of communication giving rise to the connected cars or connected vehicles. But now we are going beyond this and now we are interested in looking about autonomous vehicles.

So, autonomous vehicles or AVs are the newer development where the driver or the driving is also made to be autonomous. That is AI will drive the car. So, therefore, this particular device, which is called self-driving cars or autonomous vehicle, is going to generate lot of data and often require the use of these technologies that is called IoT technologies. That is the use of edge sensors, then internet bandwidth and the computing at the edge and at the cloud, everything is required for the development of autonomous vehicle on the roads.

So, why it is required is to handle process and analyse the information which autonomous vehicle used to get through lot of sensors. So, we will introduce about them. So, these kind of the developments of what we have kept here the part of the course that is IoT, cloud edge and machine learning. All this is are the requirements to build the autonomous vehicles. So, these automobile makers are now focusing on leveraging the edge computing to address these challenges, which was now giving away for the development in this particular manner.

So, when driving a vehicle, the milliseconds matters and there comes the edge computing in the way to help this, that is, to reduce the latency in the range of milliseconds. So, this is most important factor, not only the computation, but the computation has to be completed within the millisecond time for the development of self-driving car. And we have seen in the previous lectures that how this edge computing will bring into this kind of fact where latency can be reduced to milliseconds in the range.

So, autonomous vehicles are no different and even though it may be your AI; artificial intelligence that drives the car. So, artificial intelligence as you know, that requires a lot of data and machine learning to do the model training for that lot of compute power is needed. So, to perform the machine learning model training, often the cloud comes as defacto standard. Why? Because cloud will provide enormous of compute for doing the AI such as CPU, GPU and (())(7:25) and many other such processing units, which will facilitate the model training of the enormous or a huge amount of information.

When it comes to autonomous vehicle, this information is even more to process and thereby entering this AI on edge devices or edge computing. So, autonomous vehicle becomes an edge device and how the artificial intelligence training can perform so that the car can be driven autonomously.

(Refer Slide Time: 7:57)

So, when it comes to the edge computing, let us see what are the data which is required to be computed very close to the device that is inside the car or very close to the car. So, edge computing here will facilitate in this. Now given this requirement of a real-time data processing capabilities, edge computing has a edge and therefore, it is one of the basic pillars for developing this autonomous vehicle technology. However, this particular data is not generated by the computer, but it is generated with the help of multitude of sensors which are equipped with the autonomous vehicles, which acts as the eyes and ears of the drivers, and that is now embedded with the car.

So, sensor technology can vary widely amongst the different models of autonomous car and even within the same sector. So, most of these self-driving car sensors are fundamentally similar. Often, they collect the data around of around the world to help the pilot or an autopilot of the vehicle to drive. So, these sensors are the camera, radar, lidar, thermal sensors and so on. There are many other type of sensors which requires to capture the outside environment situation, understand, interpret it, and then take the decision in two terms.

One is, one is the steering like you have seen in the car, and the other is the acceleration which the pedal in the car gives. So, two entities to have this car autonomous without the assistance of driver that is with the help of artificial intelligence. So, this particular decision has to be converged into these two aspects that how the steering has to be operated autonomously and what will be the acceleration that is the pedal. So, these two things are the decision when manual driver used to take.

So, therefore, how these particular models that is IoT, edge cloud and AI together can solve these two problems, which we will see. These are the different technologies are used in different kind of models. So, therefore, Tesla company uses for its car eight different cameras and there are 12 features and it also uses the input from the radar sensors and much more heavily on a camera visuals is used in this kind of Tesla's car.

Google's Waymo driver primarily relies on lidar sensor and often uses camera and radar sensors. These three different type of devices that is the sensors will help to create a map around that particular car and thereby, it will be fed into artificial intelligence model to take the decisions or to suggest the decision in the two forms, how the car is to be operated on the steering and what will be that, how the acceleration is to be taken.

(Refer Slide Time: 11:44)

So, see the self-driving car, different requirements. So, autonomous driving vehicle requires two kinds of computing systems. One computer processes large amount of sensor data and images which are collected by the camera and sensor. This is one of the most important use case of an IoT and edge computing has to be enabled to do this.

The second kind of thing is that once these images are ingested, that is the data is ingested into the edge or a cloud system, then intelligence computation has to be made to take the quick decision for the vehicle to drive autonomously. Therefore, these are the different steps. First is that the data which is collected has to be pre-processed on the fly. So, autonomous vehicles, how the cameras, variety of sensors like ultrasonic, lidar, radar, all these sensors will sense the surrounding environment and then it will ingest the data into the system. So, all these things we have already covered in the lecture, so we will only give the references.

So, this data coming from different vehicle source may be quickly processed through the data aggregation and compression techniques through the edge and through the cloud. So, the computing needs requires receiving and sending these kind of data, whether it is IO ports or data ingestion, maybe MQTT or Kafka connectors will ingest the data into the cloud computing system and cloud, you know that it is a highly centralized set of resources and there these kind of data will be now fed into an artificial intelligence model for influencing.

So, after the influencing that particular data or the results will be fed to the logic or the actuators and these actuators will drive the car. Second important thing is about the secure network connectivity. Now in this particular process, collecting the data, sending to the edge or the cloud, doing the computation, that all requires good network connectivity. So, internet is very much needed nowadays. You may be seeing that 5G 6G satellite and many other type of, or a wifi, there are many other type of internet connectivity is possible.

So, thereby, this wireless connectivity and a speed is very much required for secure network connectivity needed for this vehicle. And high speed connectivity is vital for continuous deployment of the vehicle updates or push the updates like location on road condition, vehicle telemetry and so on.

Third important thing is high performance computing is needed to be enabling which will be used to process this particular data which is ingested into the system from the car sensors like camera sensors, radar sensors, lidar sensors, ultrasonic, and many other sensors will bring the data into the system, requires high performance computing. So, autonomous vehicle may generate approximately 1 gigabyte of data every second gathering and sending a fraction of data to the cloud-based server for the analysis.

And therefore, all these different resources are needed such as a bandwidth, computing power to compute this. And then only this autonomous driving systems can do the influencing part and it has to be done in a real time processing that is these data has to be considered as a hot data path analytics. So, in vehicle edge computing is very much essential to reduce the dependency on the network outages or network failures. So, this is going to increase the decision making and the accuracy of this particular driving self-driving car.

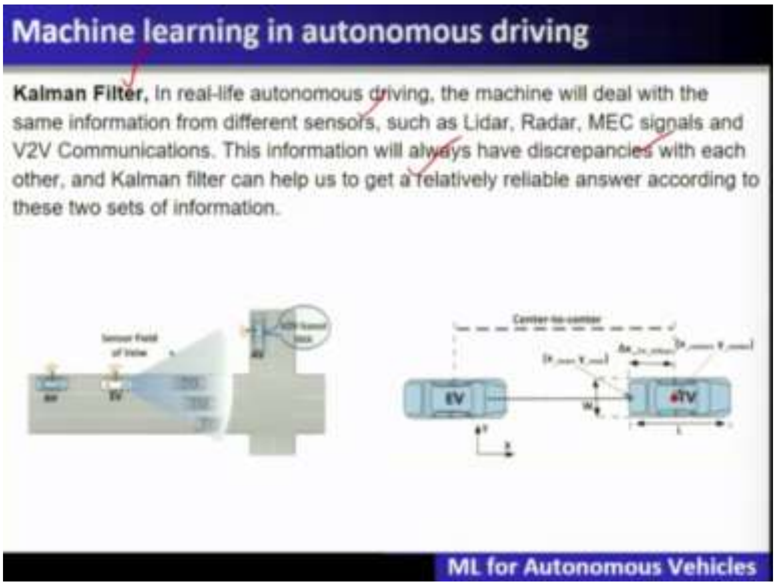
(Refer Slide Time: 16:02)

Now let us see how the machine learning is used to train the artificial intelligence for driving this self-driving car. So, the sensor data which is collected in the self-driving cars and the vehicle depends upon the methods computing technologies and the sensors itself for understanding and making the decisions. So, in many ways, most of the valuable intellectual properties of the companies like Tesla, Waymo and they are developing as the software and the data infrastructure to process and perform the actions based on the sensor data.

So, all the autonomous vehicle on the road now are you have, they have to utilize edge computing AI on this particular sensor data and they use to train the model in the data center using various machine learning models using machine learning algorithms. So, these models will be used in the autonomous car learning models and are the only way that is possible with the help of computing power of modern data centers, that is the cloud and the edge and it will be capable of doing this computation of a better flops.

So, the computing requirement of these vast machine learning models will exceed the computing power of edge computing. All things are required. So, the problem of self-driving car, we can see that it is a regression problem and regression computation has to be done through the training on an ai more algorithms. And it is similar that it takes hundreds of hours on a high performance computing. Yet once the algorithm is learned, it can quickly and accurately utilize the algorithm with a much less power. That is called influencing part of this algorithm.

(Refer Slide Time: 17:59)



Let us see some of the electronics part which is needed here in autonomous driving Kalman filter in real life autonomous driving. The machines will deal with some information coming from different sensors such as lidar, radar signals vehicle to vehicle communication. And this information will always have dependency with each other. Therefore, Kalman filter will help us to get related answer according to these different set of information.

(Refer Slide Time: 18:32)

Now let us see more detail about different kind of devices which are required in the autonomous vehicle for autonomous driving decisions required to be made with the help of machine learning. So, lidar, radar, cameras are the sensor data, which is, and machine learning algorithm is an important part of self-driving car or autonomous driving. Self-driving vehicles usually has multiple sensors, so camera, lidar and radar are sensors are a few to name with. And also, the machine learning module will tell the vehicle what to do with different kind of information which is being gathered.

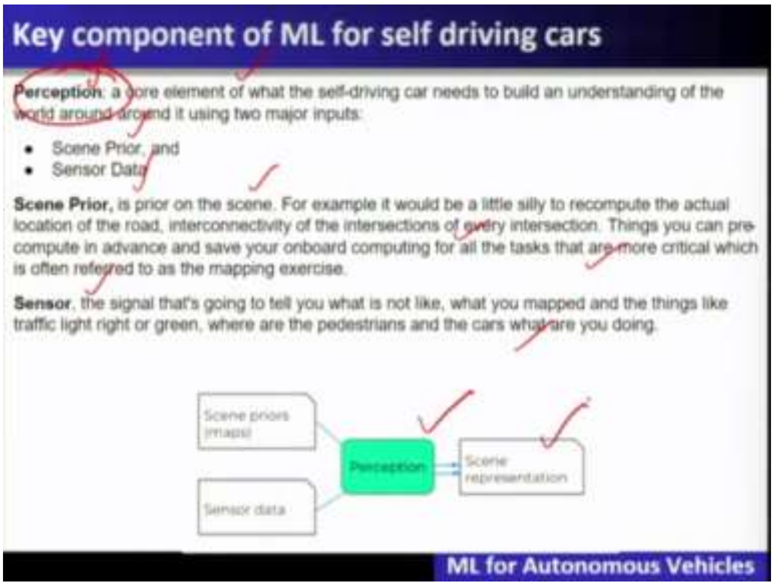
So, for example, the car needs to stop when there is a pedestrian comes in the front of the trajectory or the path of a car and the machine must be able to tell the difference between the actual pedestrians and the picture of a human. So, these are some of the things which artificial intelligence has to be trained and learned to work with the true positives, not with the false positives.

So, additionally, the camera sets cannot precisely measure the distance or work at the night, therefore lidar sensors are there to measure the distance or the depth usually emit the high frequency signals. And this high frequency signal could be used for positioning and 3D modelling, being able to tell the difference between the actual human and the picture of a human in this way.

So, radar is a low energy, a low energy cost solution for positioning because the radio wave trans is emits is usually is with the low frequency, low frequency wave cannot depict the detailed 3D shape, but it is enough for positioning. So, you can see here lidar is for the distance measurement whereas the radar is for positioning information and the cameras are still needed because in addition to that lidar and radar, they cannot identify the colours.

Now you can tell in the roadside condition colours means that signals the signal operates with a red means stop, green and orange, stop ready and go. So, to find, follow these signals and follow the road indication or the road signs, all these requires a separate technologies like, like camera, which has the capabilities to work with the colours also. So, now coming to the communication earlier, you know that vehicle to vehicle communication was quite established technology and now with the increase of accuracy which is needed for autonomous driving and the vehicles is, is very much required to be improved upon and that is being used over here.

(Refer Slide Time: 21:29)



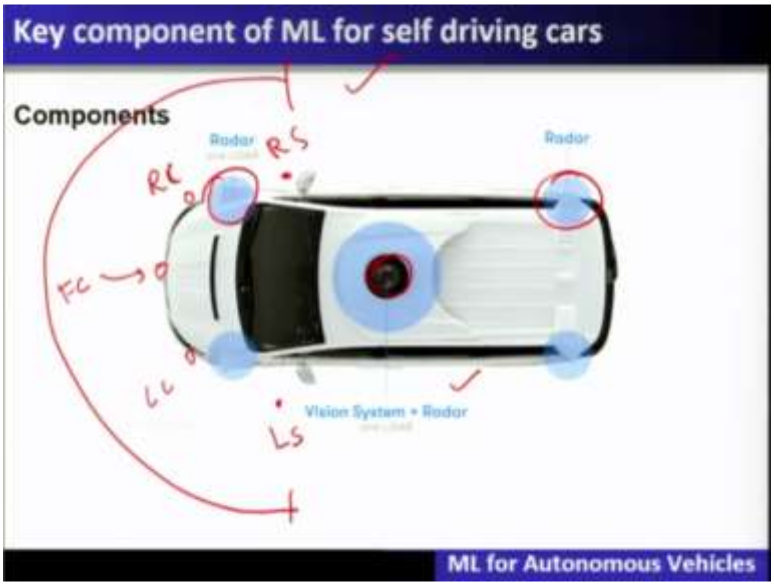
Now let us see what are the key components in the self-driving car and how machine learning is used to solve this particular problem and we are going in more detail about the use of machine learning in self-driving cars. So, perception is a key element or a factor which is required in self-driving car implementation. So, perception means a core element of what the self-driving car needs to build to understand of the world around it using two major inputs.

So, perception is done with the help of scene, prior scene, prior means all the points, all the signs, road signs which are there in that city that is called scene prior will become a point of perception. And also, the sensor data that is current position, current situation of the environment with the help of sensor data. So, perception is captured using the scene prar and the sensor data together.

Now scene prior is a prior on the scene that is, it would be like signs on the road or the location of the road interconnectivity, intersection or every other (sec) intersection or a road dividers. These are all scene priors which are not going to change, but the car is moving on it. So, it has to see that is why it is called scene prior. So, things you can pre-compute scene prior can be pre-computed in advance and save this onboard computing for all the tasks which are more critical.

Whereas, sensors will give the real time or the actual information at that instant of time. So, the signals out from the sensors is going to tell you what is not like or what you mapped or the things like traffic light, green and pedestrians and other things on the road which are also moving in, in the same trajectory as the car is moving. So, all this is enabled with the help of sensors. So, this forms very important concept that is called the perception has to be captured for taking the self-driving decisions. So, scene prior and sensor data together will form the perception and will construct about the scene representation and that is going to be useful.

(Refer Slide Time: 24:00)

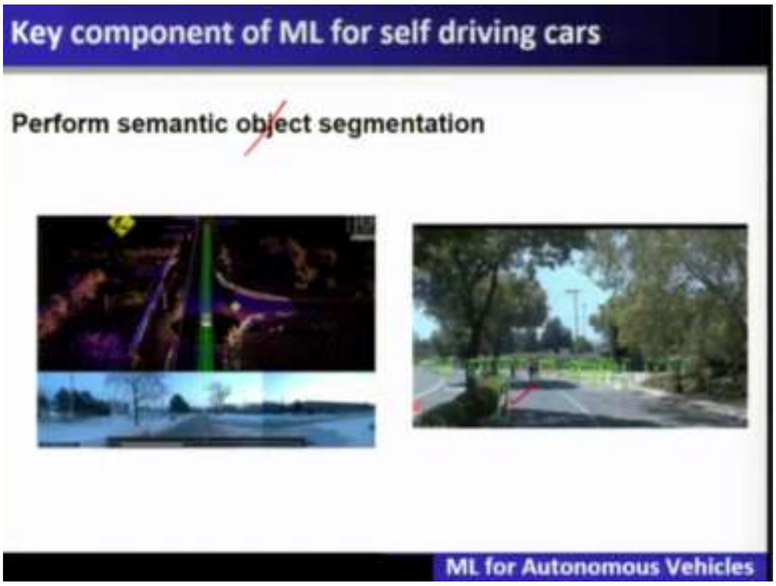


So, let us see this particular diagram. In this diagram. This is the top view of Waymo self-driving car, which requires the sensor on the backs radar on the backside. In the front side, there is a radar R and lidar and there is lidar and a vision system and a radar r that is lidar and lot of cameras are embedded in the car. That is the front camera. And this is the left camera and this is the right camera and there are two side cameras are there. This is right side camera and this is left side camera; 5 camera will be capturing the, the view which is required to understand the current situation and that is being taken out by the camera lidar and raar sensors.

(Refer Slide Time: 24:58)

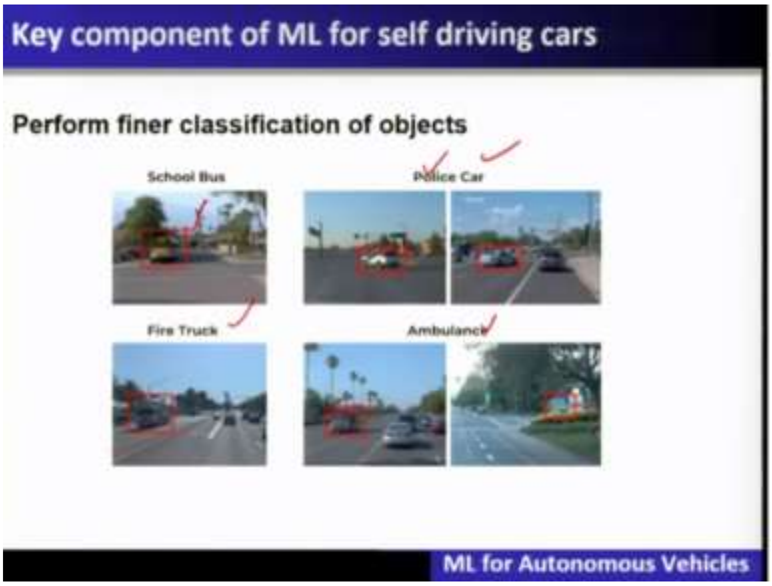
So, let us see how the scene representation is being finally built out of these particular sensors. So, for example, in the car you see that the, this is the position of a car which radar will give the localization. In this particular position of a car, you can see various tagged objects are there, this is the cyclist coming on and these are all different pedestrians. This is the parking, this is the vehicle, this is the police traffic car and so on. So, you can see that this entire scenario where cyclist is also there, all these things has to be captured and this is called the scene representation.

(Refer Slide Time: 25:36)



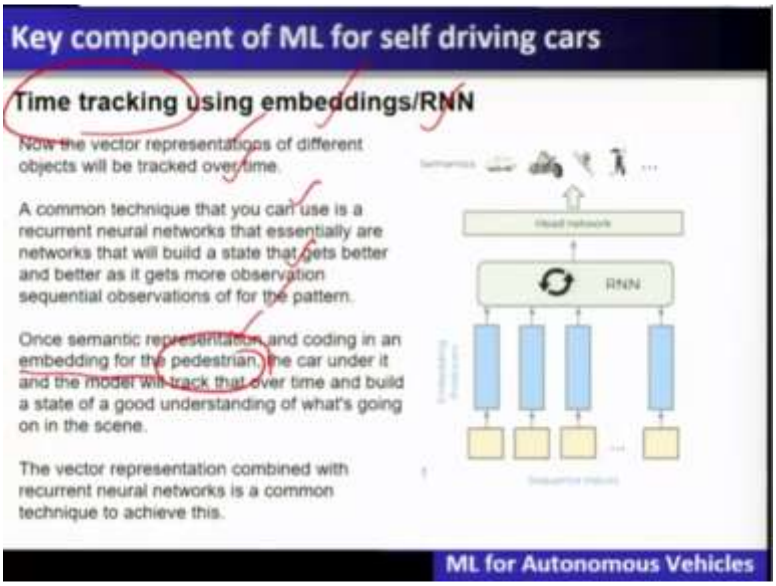
Then comes to perform the semantic object segmentation. You know that in the image processing and object detection. So, object segmentation and object classification, all these are part of that deep learning. So, deep learning technologies that is the machine learning can be utilized for doing this object segmentation. So, this is the example of an object segmentation which has to be done here in this car.

(Refer Slide Time: 26:09)



Now another thing is that once the object segmentation is complete, then you have to do a finer filter classification of this object. So, what this object which are being identified, whether it is a school bus, so the self-driving car has to take a different decision if the vehicle coming in front or going ahead is a school bus or it's a police car or it is an ambulance or it is a fire truck. So, it depends upon what kind of objects they are being classified and how what decisions has to be taken as per the conditions.

(Refer Slide Time: 26:47)

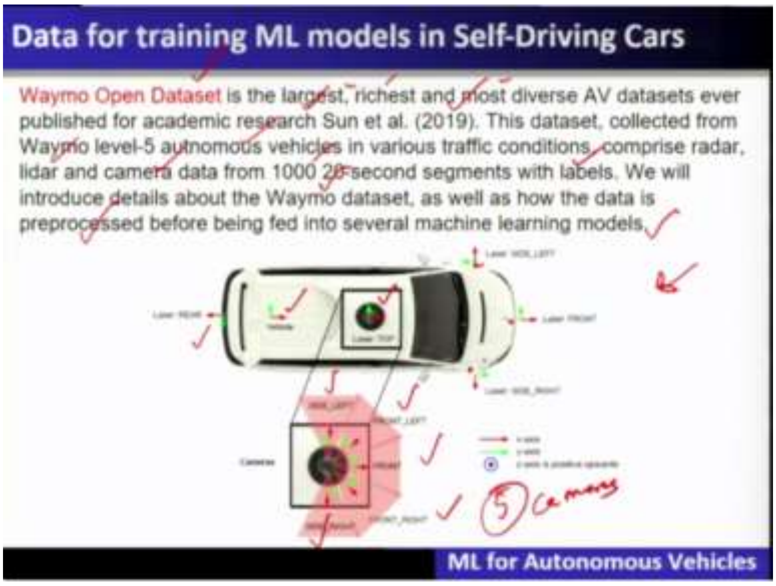


Now another important thing is that once we have shown that how you can use these kind of sensors for scene representation, then the next thing is about the time tracking and how this time tracking is to be done with the deep learning or machine learning and using embedding. And that is to be done with the technology which is called recurrent neural network.

So, in this recurrent neural network for time tracking, when you say embedding, embedding means representing the entire rep, entire picture or entire scene in the form of a vector of different objects which can be tracked over the time. So, a common technique that you can use is a recurrent neural network that essentially are the networks that will be built with a state and gets better and better as it gets more observation. And that is sequential observation of the pattern.

Once the semantic representation and the coding in an embedding of for a pedestrian is done, the car under it and the model will track over the time and build the state of a good understanding what's there in the scene, what's happening in the scene. So, this vector representation often calls EM embedding is combines the recurrent neural network is a common technique. To achieve this, we will have this estate information embedded within it in the vector.

(Refer Slide Time: 28:18)



So, let us understand about one use case, which is called a Waymo open dataset, which is will label, which is released for CAD machines and researchers. They are for to use this dataset for doing the experiment in IoT cloud, IoT edge and machine learning, we require this kind of dataset and we will also show you the model which and the, the issues with the self-driving car request to be solved for better accuracy and the greater development with the speedy development of this entire technology.

Waymo more open dataset is the largest, richest and the most diverse autonomous vehicle dataset ever published for academic research. It was just released in 2019 and this dataset is collected from Waymo level 5 autonomous vehicle in various traffic conditions comprised of radar, lidar and camera sensor data from thousands of twenty second segments with the labelling that is how this particular dataset can be used for model training. It is a supervised model training can be done.

So, we will introduce these details of Waymo dataset as well as how this dataset is pre-processed before being fed to several machine learning model. So, again, to explain this Waymo dataset, we have to see the, the placement of different sensors in the car. So, this is a Waymo car. You can see here that 1, 2, 3 different cameras are there and these particular devices are also, it could be the laser and then it will be on top will be the lidar laser. And this vehicle also will have this particular position.

So, the camera, if you see, there are 5 cameras you can find out which is placed here on the top of this car side left, front left front, front, right side right. So, 5 different cameras in the front side view it will be able to take.

(Refer Slide Time: 30:35)

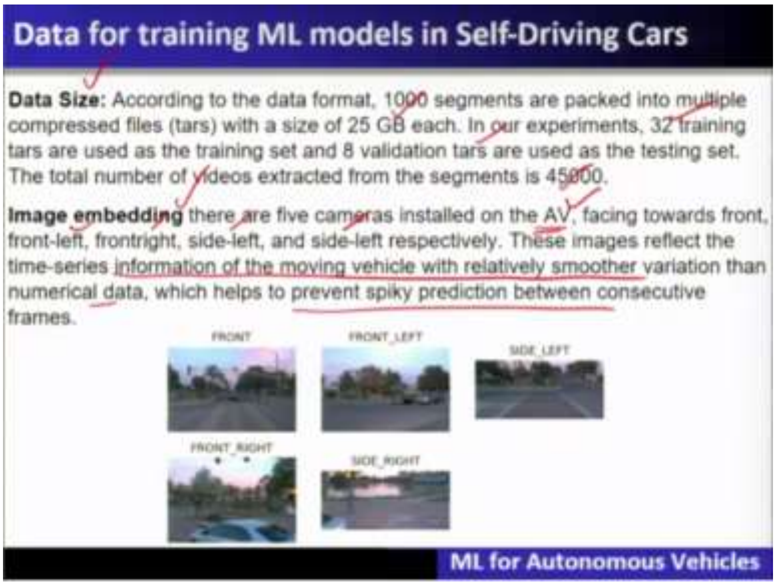
Now let us see, when you say supervised or learning or a dataset with a label. So, what do you mean by the label? So, what is there in into the label which is being prepared by Waymo open dataset? Labels refer to kinematics and special parameters of the objects which are represented as the bounding box. Specifically, one kind of label type is classified into the pedestrian, whether it is a pedestrian, whether it is a vehicle or it is an unknown sign, cyclist, all these categories in one such level, the or type of object is, is being classified.

The detailed information is provided for each level among which we especially pay attention to the coordinates of the bounding box that is the velocity and accelerations a in the subsequent feature extraction steps. So, these are all presented in the database. With the help of labels, you will be able to train the models, how, what action has to be done and how the vehicle, how, how the self-driving car has to act based on these things.

Then coordinate system. There are three coordinate system provided in this dataset. Global frame, then vertical frame and the sensor frame. So, some raw features are also represented in the unintended coordinate system and in order to maintain the consistency, it is crucial to transform the data into the correct coordinate system. So, data dataset is also provided the vehicle posed to transform the vehicle from one coordinate system to another coordinate system.

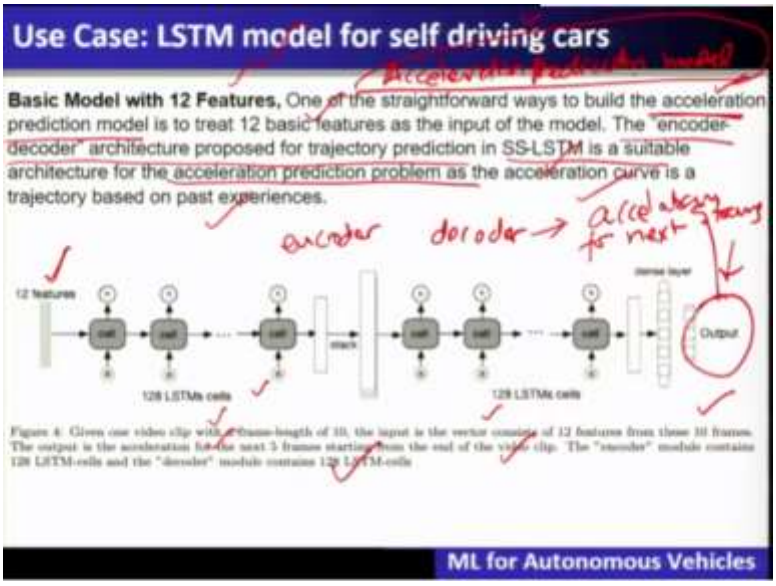
Now how to do the acceleration computation with all this information with the Waymo dataset, because once instant acceleration is not directly available in the dataset, but the groundwork for the training and evaluation needs to be computed by velocity differences. So, velocity information is there by taking the velocity differences you can find out this particular factor or a feature.

(Refer Slide Time: 32:42)



Now dataset size, according to the data format, 1000 segments are packed into the multiple compressed files with the size of 25 GB each. And in this experiment 32 training tars are used for training set and 8 validation tars are used for testing sets and total number of videos extracted is 45,000. So, image embedding, if you talk about that there are 5 cameras installed in an autonomous vehicle, which we have seen facing towards front, front left, front right, side left, side right, respecting these images reflect the time series information of a moving vehicle with relatively smoother variation than the numerical data which helps to prevent spiky predictions between the consecutive frame.

(Refer Slide Time: 33:44)



So, you can see that this is the picture of a front camera. This is the picture of a front left, the side, left front, right side, right? So, these will give the complete information. Now let us see how this particular sensor data from autonomous vehicle will be used to train the model that is LSTM model example we are taking, there are many other functionality which require to be created out of this particular dataset. So, basic model with 12 features. One is state forward way to build the acceleration prediction model. So, we are now focusing on acceleration prediction model for self-driving car to treat these 12 basic features as the input to the model.

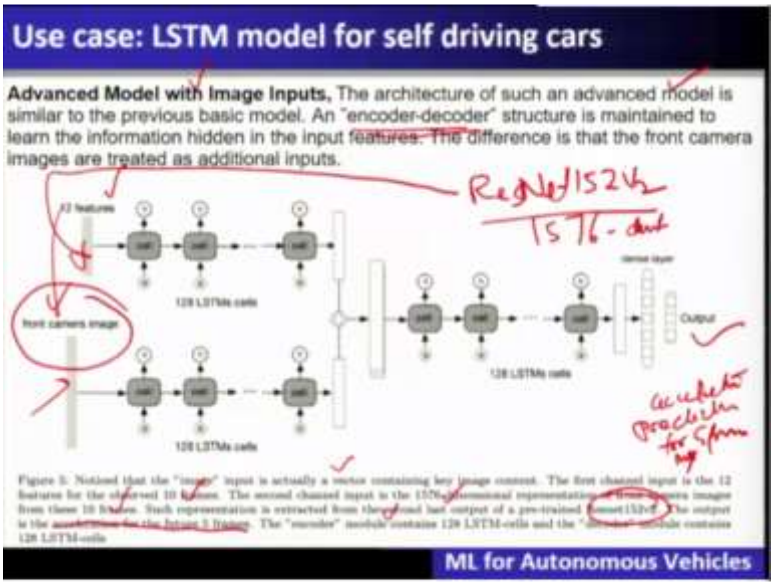
So, for this we need an encode and decoder architecture which proposed for the trajectory prediction in SS-LSTM is a suitable architecture for acceleration protection problem. We are not going in detail of this architecture. We have already covered this LSTM in in prior, in previous lectures as an acceleration curve and trajectory based post experience.

So, here in this particular figure you can see there are 12 features which are fed into the 128 LSTM cells. And this is good enough given one video clip of frame 10, frame length 10. Each input is the vector consist of 12 feature from 10 different frames. So, these 12 features are obtained. Now to obtain these 12 features out of video clip of frame land 10 can be done using CNN technologies and flatten that CNN model into extracting these 12 feature and fed to the LSTM. That is 1 28 LSTM cells.

So, the output is the acceleration for the next 5 frames starting from the end of the video clip. So, this will be output will be that acceleration. So, this acceleration will be computed using this particular model that we have seen that the driver has to take two decision. One is the acceleration using the paddle and second is the steering wheel. So, so this LSTM model for self-driving car, we have shown that how the acceleration prediction model can do the computation for acceleration for the next 5 frames.

So, encoder and quarter model contains 128 LSTM cells and decoder model also contains 128 LSTM cells. And finally, there will be a dense layer and the output will be the regression that is in terms of the acceleration.

(Refer Slide Time: 36:37)

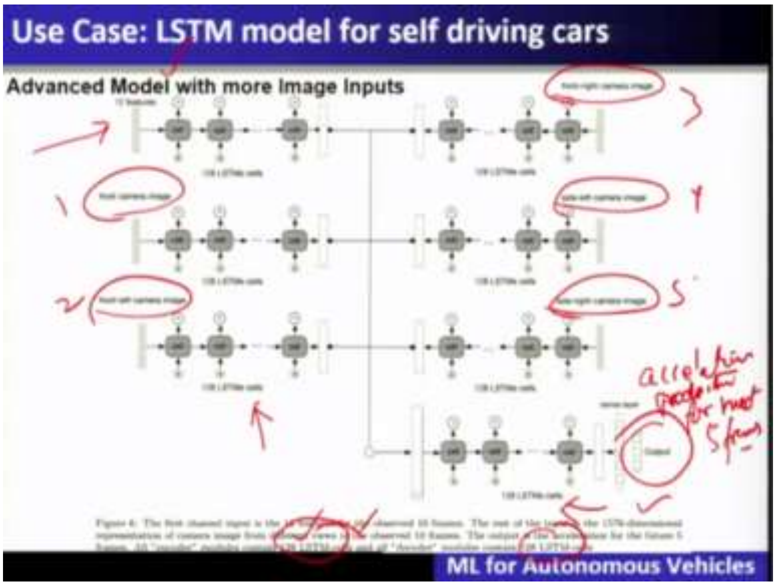


Now let us see the advanced model with the image inputs. So, this particular example shows that there is a input camera feeds. So, notice that the camera input is actually a vector containing the, the key image contents. So, first channel input is 12 features here for observe 10 frames. The second channel input is 1576 dimensional representation of the front camera image from these 10 frames. So, such representation is extracted from the second last output of a pre-train resonate 152 version 2. 5

So, here to come to this particular features and give the input, it requires the use of ResNet pre-trained model 152 V2 and this will extract this 1576 dimensional representation of the front camera and the output will be the acceleration prediction for the next 5 frames. So, here the architecture of such an advanced model is very much required, is similar to the previous model.

So, here the encoder decoder is structure is maintained to learn the information hidden in the input features. And the difference is the front camera images are treated as additional points, additional inputs.

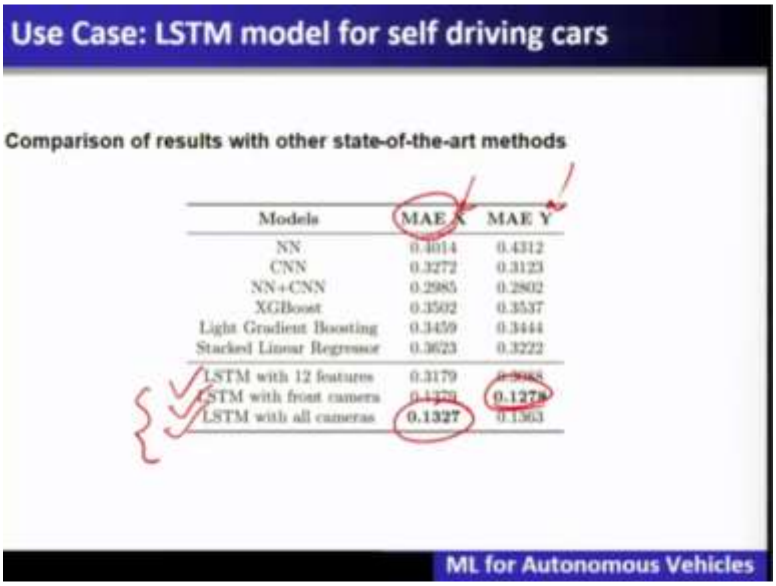
(Refer Slide Time: 38:14)



So, let us go ahead with more advanced model with more input image inputs. So, you can see that all the 6 cameras, 5 cameras, front camera images will be fed and front left camera images will be fed front right camera images will be fed side left, side right, so 1, 2, 3, 4, 5. So, five camera input is now integrated together in this particular system and it is going to give you the input in 1576 dimensional representation of the camera image from different views. And you can see that here this will be observed and the output is the acceleration for the next 5 frames.

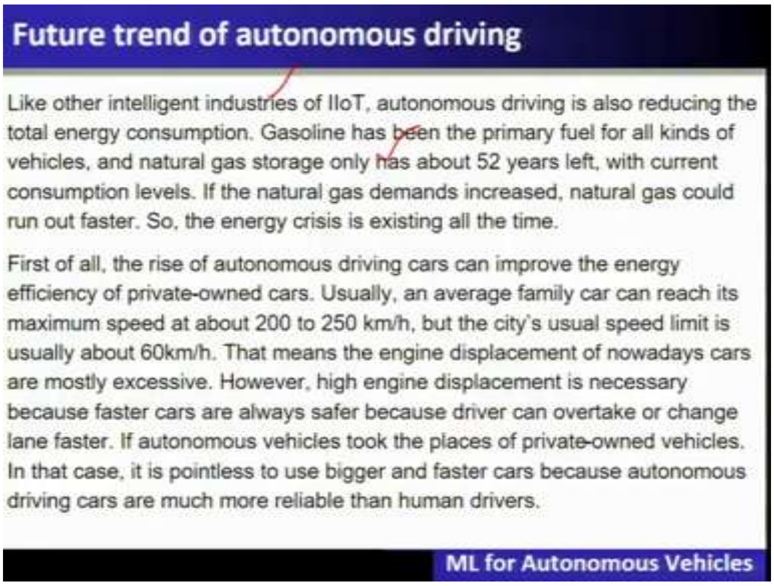
So, all the encoded module must contain 128 LSTM cells and decoder module also contains 128 LSTM cells iron. This will be able to predict the acceleration prediction for the next 5 frames with the help of 5 different cameras placed over here.

(Refer Slide Time: 39:29)



Now if you compare these results with the other state-of-the-art methods, what we have shown you, this is the LSTM with 12 features, LSTM with front cameras and LSTM with all the cameras. So, here you can see this particular metric called mean absolute error in terms of x in terms of y is quite reasonably obtained with a and this particular error is quite less so in terms of very less error. So, this LSTM with the front camera and LSTM with all cameras are the technologies or the models which can be used here for making the acceleration prediction for self-driving cars.

(Refer Slide Time: 40:12)

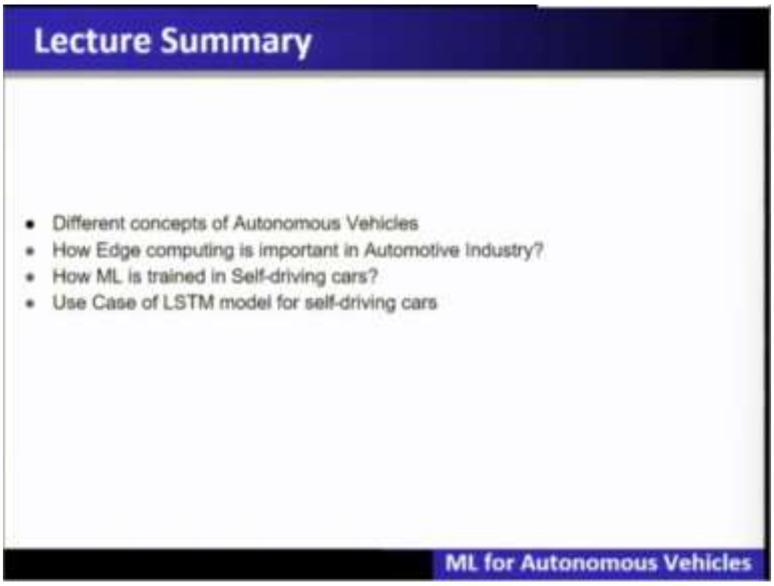


So, let us see the future trends of autonomous driving. So, (indus) so intelligent industries or industrial IoTs and autonomous driving is reducing the energy conjunction and you see that this is going to happen, why, because this autonomous driving car will improve the energy efficiency and safety of the car.

(Refer Slide Time: 40:41)

So, this autonomous drive vehicle will replace the private car and also will solve the parking issues. People will have the bigger house and bigger living areas and no separate houses that is called garages are needed and no traffic congestion. The routes will be, and the road traffic will be without congestion and will have this efficiency in the road transport.

(Refer Slide Time: 41:06)



So, let us summarize this entire lecture. We have seen various concepts of autonomous vertical iron, how edge computing IoT, cloud and machine learning together is applied. We have also seen how the edge computing is very, very important in the self-driving car industry and how the machine learning is trained for the self-driving, car driving and its functionality such as steering operations and the pedal operation that is called acceleration prediction.

For that, we have shown the use of LSTM model for these kind of functionalities using various sensor data. Let us say that we have shown the use of cameras, radar, lidar, all that sensors we have applied and used in a LSTM based model for taking the self-driving car (())(42:03). Thank you. Thank you very much.