TECHNICAL NOTE

Deep Neural Networks in Perception Systems for Advanced Driver Assistance Systems

An Overview

Neural networks are critical building blocks to realize advanced driver assistance systems (ADAS) and are employed in the various tasks required to accomplish it, such as localization, path planning and perception. This Technical Note explains what neural networks are, how they function and the various techniques used for object detection and classification for perception systems. Further on, this document explores a commercially available solution and the value it brings to automotive perception engineering today.

What Are Neural Networks and How Do They Work?

Neural networks are inspired in their function and design by the human brain and are a subset of machine learning and artificial intelligence. A neural network is a computing system that operates using a series of algorithms and produces an output based on input data. The algorithms are expressed as mathematical functions, and artificial neural networks can learn from events and make decisions on similar events. These algorithms are especially useful in advanced driver assistance systems due to the wide range of situations presented to vehicles equipped with ADAS. Over the long term, the goal of using machine learning and neural networks is that if vehicles are trained and exposed to enough varied data, they will become robust and reliable in assisting humans with driving functions even when faced with previously unseen driving circumstances.

A feed-forward neural network is shown in Figure 1 for reference. Neural networks can be thought of as a system that, based on the input data, delivers a best-fit output (minimum error) by optimizing variables. A simplified example is provided below, wherein the system is provided raw data (number of gears vs. top-speed achieved) and delivers a best-fit line (in dotted blue). The variables that could be adjusted by the system, in this case, are the y-intercept and slope of the line. Neural networks are complex systems that output the best fit by optimizing the variables.

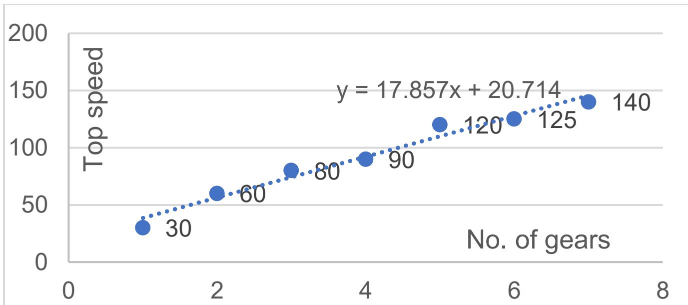


Figure 1 – Best-fit line given an input dataset

A neural network is shown in Figure 2 below. The raw data is fed into the input layer and the output is delivered through the output layer. The hidden layers represent the series of algorithms and activation functions that optimize the variables to provide the output. Each circle in each layer is called a "neuron" and every neuron has a value assigned to it. The line connecting one layer's neuron to another layer's neuron is called "weightage." Weightage is the relationship between two neurons with the magnitude of the weight representing the strength of the relationship.

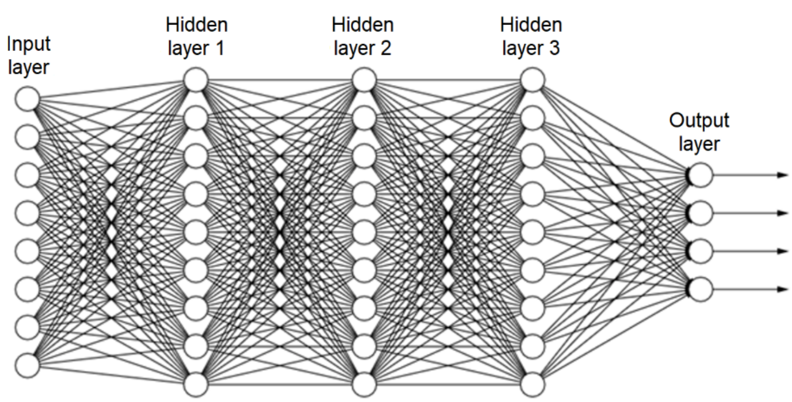


Figure 2 – Neural network architecture example

The neuron value in the previous layer and its associated weightage are processed under a mathematical function called the "activation function" to determine the value of a neuron in the next layer. As an example, the B1 neuron value is shown below.

$$B1 = f\ (A1.W1 + A2.W2 + A3.W3)$$

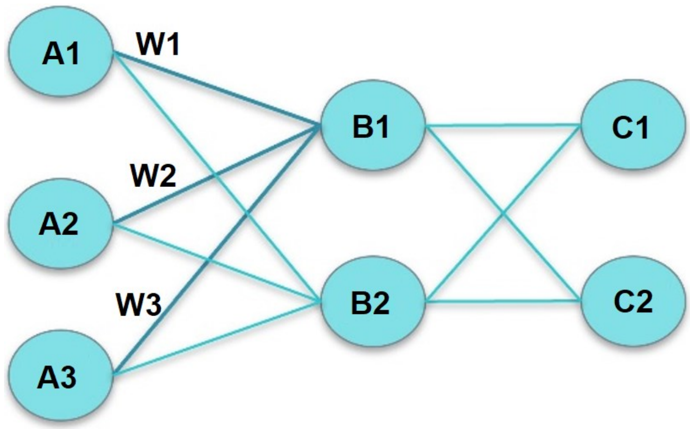


Figure 3 – Sample neural network with weights

The multiple hidden layers and activation function add non-linearity to the neural network model, enabling complex mathematical operations that better represent the complex behavior of the application and further lower the output error. The output error is calculated using the cost function, which is explained in further detail later. Some examples of activation functions are sigmoid, hyperbolic tangent (tanh) and rectified linear unit (ReLU) functions. The choice of which activation function to deploy for optimization depends on factors such as the network's learning efficiencies, computational performance and vanishing gradient problem.

In some instances, it is desirable to offset the activation function by adding a constant. This constant is known as "bias." Weights and biases are the parameters that are changed to fit the model and produce the best-fit line.

How Do Neural Networks Learn?

There are two techniques to train neural networks: supervised learning and unsupervised learning. Supervised learning refers to providing a training dataset in which an input is provided to the network and the correct answer is labeled. During training, the neural network adjusts the weights and biases to improve its performance on the training data and reduce the error. The networks are trained on large datasets, with the goal that what the network learns during this training exercise by changing weights and biases can apply to previously unseen data to produce the correct output. An example of supervised neural network training is by teaching the network that '1+1=2, 1+2=3, and 1+3=4'. Then the goal is for the network to correctly output the answer to 'What is 2+3?'.

Another method of training neural networks is unsupervised learning, wherein the correct output is not manually provided by a human but rather by an automated system that can accurately and reliably provide the true output.

A key element of improving a neural network's performance is understanding the difference between the produced value and the true value. The cost function quantifies the error between the produced and true output value. There are many techniques to construct a cost function, one of which evaluates the root mean square of the difference between the true value and the produced value. Next, the optimizer analyzes the error and changes the weights and biases to minimize this error. Once the cost function produces an output of minimum error (best-fit line), the network can be thought to have been trained and can be tested for performance against another dataset and finally deployed in real-world scenarios.

Neural Networks in Advanced Driver Assistance Systems

Machine learning and neural networks are used for various driving tasks such as localization, object behavior prediction, decision making, path planning and environmental perception. This piece focuses on environmental perception challenges and how deep neural network is used to recognize, classify and detect objects in the surrounding environment.

Sensor architecture for entry-level ADAS consists of radars and cameras for front- and surround-view systems. Higher autonomy levels will likely add LiDARs to the sensor mix. Radars and LiDARs utilize machine learning and neural networks for object detection and classification by training the network on point clouds.

Figure 4 – Perception system in action

The first stage of perception is object detection, and two techniques used to achieve it are anchor-based object detection and anchor-free object detection. Predefined bounding boxes as proposals of ground truth are called "anchors." In anchor-based object detection, anchors are assigned class labels with a label assignment strategy. For example, in a naive label assignment strategy, if the maximum IoU (intersection over union) of an anchor is greater than 0.5, the anchor is assigned the ground truth label.

Anchor-free object detection provides a 3D bounding box with respect to a fixed reference point in the image, the center of the ground truth, which must fall within a grid cell. The grid cell is, therefore, responsible for predicting the object’s width, height and deviation of the ground truth center from the grid cell center. Anchor-free object detection is more generalizable and can be easily extended to key-points detection. For an anchor-based object detection model to work well, suitable shape and size anchor boxes must be provided in training to achieve strong performance.

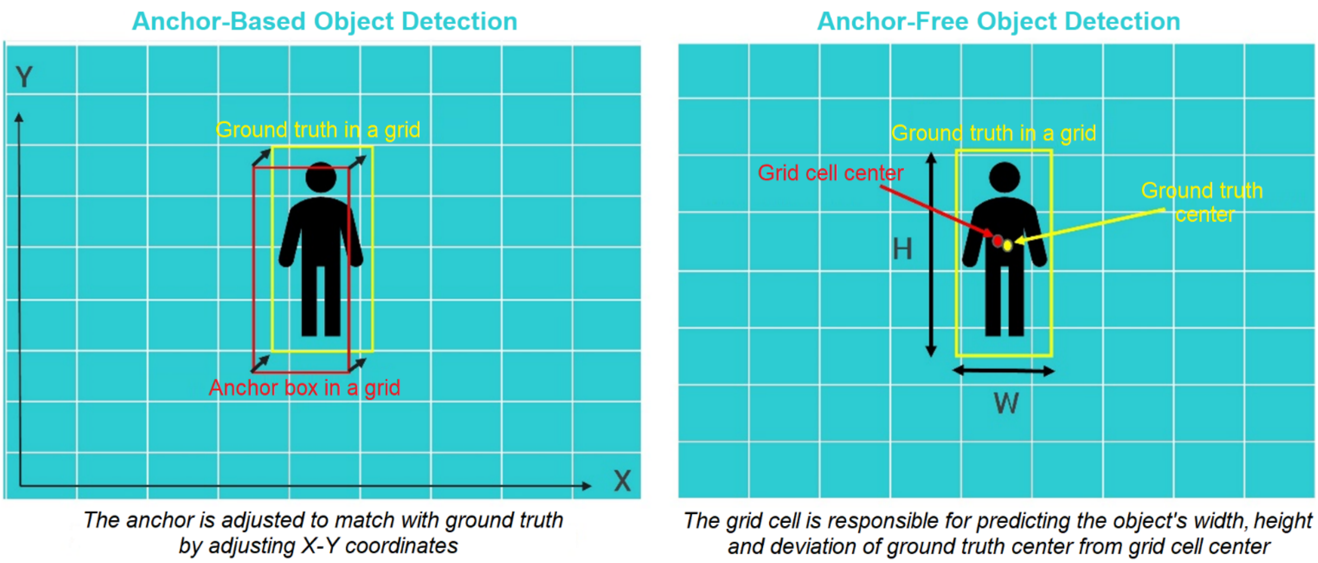


Figure 5 – Difference between anchor-based and anchor-free object detection

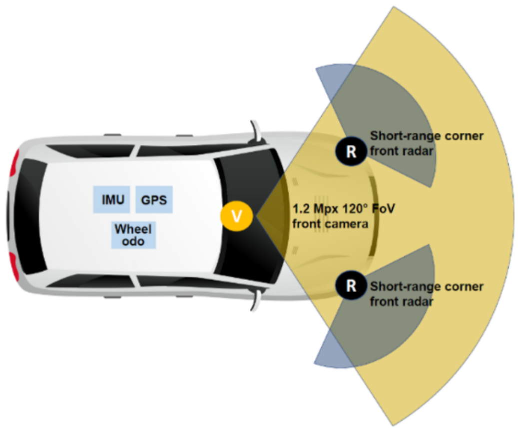
Once an object is detected, it must be classified to understand what that object is. Object detection and classification can be performed using a convolutional neural network (CNN). Convolutional neural network is a type of neural network with applications in computer vision, image recognition and classification systems. CNNs pass the input data through the neural network and then understand correlations and patterns that enable image recognition and classification.

Commercial Fusion and Perception Solution for ADAS and AD

LeddarVision, LeddarTech’s low-level sensor fusion and perception software, combines AI and computer vision technologies and deep neural networks with computational efficiency to scale up the performance of ADAS/AD sensors and hardware, which are essential to enable safe and reliable ADAS and AD.

LeddarTech enables customers to reap even greater performance and cost rewards as the platform is scalable and sensor-agnostic, due to its raw data fusion technology, unlike other solutions on the market. As a result, the customer controls the design by determining any camera, radar or LiDAR suite that best meets their application and performance requirements.

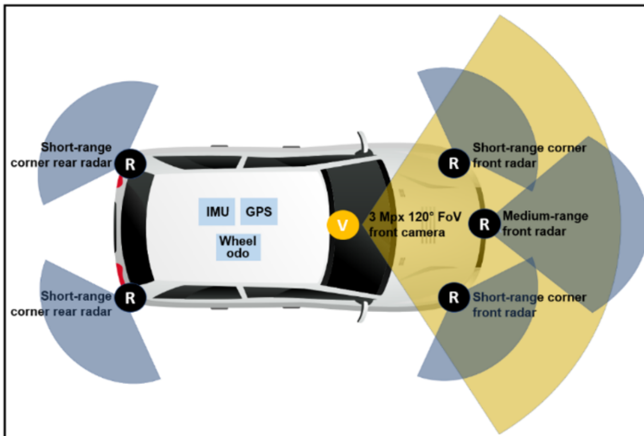
LVF-E



Hardware Platform

- Sensor configuration: 1V2R
- Front camera: 1.2 Mpx 120° FoV
- Front radars: 2 corner SRR
- SoC: TI TDA4L, 8 GB
- IMU, GPS, wheel speed: CAN

LVF-H



Hardware Platform

- Sensor configuration: 1V5R
- Front camera: 3 Mpx 120° FoV
- Front radar: 1 MRR
- Front/back radars: 4 corner SRR
- IMU, GPS, wheel speed: CAN

In December 2022, LeddarTech released two LeddarVision front-view products, namely **LeddarVision Front-View – Entry (LVF-E)** and **LeddarVision Front-View – High (LVF-H)**, enabling entry to premium L2/L2+ ADAS, highway ADAS and 5-star safety rating for NCAP 2025/GSR 2022. The products implement premium stacks that handle sensors' interface, offline and online calibration & diagnostics, sensors synchronization, sensors fusion, object detection and classification, extended to include unclassified objects and events (e.g., cut-in), continuous tracking and stabilization, free space detection, road model, comprehensive traffic signs detection, highway traffic light detection, vehicle odometry interface, ego-motion localization and global localization with HD Map input. **Critically, LVF-E and LVF-H accelerate the commercialization of ADAS and AD and democratize safety for all road users by delivering high performance and reduced costs.**

LeddarTech®

CANADA – USA – AUSTRIA – FRANCE – GERMANY – ITALY – ISRAEL – HONG KONG – CHINA

Head Office
4535, boulevard Wilfrid-Hamel, Suite 240
Québec (Québec) G1P 2J7, Canada
leddartech.com

Phone: + 1-418-653-9000
Toll-free: 1-855-865-9900

20230324 / TF ID 049488