# VEHICLE DETECTION & TRACKING

**Aditya Aggarwal , Shreya Pandey**

Any machine that moves people or goods qualifies as a vehicle. Vehicles include automobiles, bicycles, buses, aircraft, space shuttles, and a wide variety of others. A practical application of machine learning principles is vehicle detection and vehicle type recognition, which is directly applicable for a variety of operations in a traffic surveillance system and contributes to an intelligent traffic surveillance system. The processing of automatic vehicle identification and recognition using datasets of static images will be covered in this study. The surveillance system features moving vehicle detection and recognition, vehicle count, and permission verification with the organisation. Considering that algorithms are crucial to any machine learning programme, it is crucial that we select the optimum model for our project.

## INTRODUCTION:-

A CCTV camera is a crucial component of a system for intelligent traffic surveillance. Simply said, it is the automated process of keeping track of traffic in a specific area and identifying cars for future action. The acquired photographs can give law enforcement and other public tracking services useful information, such as the licence plate number of the car, its time and movements, the driver's information, etc., all of which may lead to evidence of a crime or other unanticipated or tragic incidents.

In the past, humans manually processed photos. In reality, this technology is still in use in India, whereas KSA has also adopted automated machines—CCTVs—that run continuously and take quick action by signalling. Human error and other factors that impact living things have long made manual labour slower and less effective. Using a variety of techniques, many creative minds have created specific intelligent traffic control systems.

The image shown above is the result of our study model's simple vehicle detection and recognition in part 1. All things are simply observed, but only those that are vehicles are recognised and categorised. When cameras are deployed at busy intersections, they scan the area and identify the vehicles as indicated.

In our research, live data is captured by smart sensors like cameras and sent to a cloud server, where it is first processed, cleaned, and then evaluated.

We would need a model that would produce the greatest outcomes for ITSS1 in order to move on. The effective model is chosen, then it is tested and trained using the examined data. The vehicles are then identified and detected after this is completed.

Because no person can imagine living without a daily commute by car, traffic accidents are regarded as a common occurrence.

It is crucial for everybody engaged in a traffic collision to portray potential defect scenarios in an objective, technical, legal, and scientific manner. Thus, researchers pay close attention to and place a lot of emphasis on vehicle identification and tracking systems.

Deep learning, which is at the heart of artificial intelligence technology, has made significant strides in recent years. Deep learning is used as the primary technology in the field of computer vision. The foundational technical support for several features, including target tracking, semantic segmentation, and unmanned driving, is provided by target detection technology, one of the key technologies in the field of computer vision.

Three components make up an unmanned driving system: perception, planning, and control. Perception is the process of taking in one's surroundings. It can be divided into perception of things and perception of people, specifically. Perception of things also includes perception of moving objects and impediments. The computer vision technologies are mostly employed in the unmanned driving system to detect the surrounding environment using the target detection technology. Of course, additional sensors are used to detect the surroundings. The development of autonomous parking systems represents how a vehicle perceives its surroundings. The autonomous parking of automobiles without human involvement is made possible by a variety of scientific technologies, including machine vision .

Vehicle target detection technology based on deep learning is one particular application area for target detection technology that is employed in unmanned driving systems as well as intelligent transportation systems. The numerous preserved vehicles have placed a significant burden on the traffic roads.

By applying vehicle detection technology to identify the appropriate information about road vehicles, an intelligent transportation system can be developed. Then, communication, computer, and artificial intelligence technologies may be used to control the entire transportation system. Extracting road information and intelligently dispatching vehicles can be used to increase traffic efficiency and relieve traffic pressure.

Although having more vehicles makes it easier for people to move, it also causes some issues. Traveling is no longer convenient due to traffic, and the need for intelligent transportation has also garnered considerable attention. The development of computer vision has advanced significantly in recent years thanks to the growth of deep learning and its extensive use in the area. Computer vision technology is also extensively used in everyday life, such as in automatic driving, facial recognition, and image segmentation. Vehicle detection is necessary for both unmanned driving and intelligent transportation.

Vehicle detection is a type of target detection technology that shares the same primary task objectives as target detection, which are divided into target location and target categorization. Vehicle identification algorithms can be divided into three primary categories: those based on prior knowledge, those based on shallow machine learning, and those based on deep learning.

## Traditional Vehicle Detection Algorithm

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 01 | January - 2023 Impact Factor: 7.185 ISSN: 2582-3930

The target features for a traditional vehicle detection algorithm must be manually designed. The detection algorithm among them recognises the vehicle through the lines, shadows, or other edge features of the vehicle body, and then extracts the position information of the vehicle in the image to accomplish the purpose of vehicle detection through the difference between the grey levels of the shadow at the bottom of the vehicle and the surrounding pixels of the vehicle body.

The algorithm will be misjudged, making it unsuitable for high-precision detection. The effect of detection by this method will, however, be greatly influenced by different lights because the grey value of its image will be greatly influenced by light, and the shadow of non-vehicle objects similar to the shape and size of the vehicle will also be identical to that of the vehicle. Artificially crafted car features can also be used to detect vehicles by detecting their brake lights.

Vehicle detection can be accomplished by detecting vehicle brake lights, but this method has significant drawbacks due to the influence of light, so the detection performance is not ideal.

Vehicle detection technology based on shallow machine learning, which realises vehicle recognition by combining machine learning algorithm on the basis of vehicle characteristics, is another component of traditional vehicle detection methods.

### Vehicle Detection Algorithm Based on Deep Learning

The algorithm for detecting vehicles using deep learning is the same as the algorithm for detecting targets using deep learning. The methods used are mainly divided into two categories, namely one-stage and single-stage target detection, and two-stage target detection can be used to determine whether a region is a candidate region or a recommended region. One network may completely finish one-stage target detection by feeding it photos and then producing bounding boxes and categorization labels. Two separate networks work together to complete a two-stage target identification process by creating suggested regions for input images and then passing those regions to a classifier for categorization.

R-CNN, Fast R-CNN, and Faster R-CNN are three examples of two-stage target identification algorithms. The main advantage of this approach is that it generates a set number of candidate targets based on local recommendations before processing them with a convolution neural network. The R-CNN first uses candidate regions to perform sparse sampling on the input original picture, then uses CNN to extract features from the candidate regions, and then uses SVM to classify the extracted information (support vector machine). R-CNN has significantly improved target detection accuracy compared to the conventional target detection algorithm, and the algorithm's restrictions have shrunk significantly.

**convolutional neural network (CNN) is a type of neural network that has at least one convolution layer**. We use them for obtaining local information, for instance, from neighbor pixels in an image, and to reduce the overall complexity of the model in terms of the number of parameters.
Within Deep Learning, a Convolutional Neural Network or CNN is **a type of artificial neural network, which is widely used for image/object recognition and classification**. Deep Learning thus recognizes objects in an image by using a CNN.
a convolutional neural network (CNN/ConvNet) is **a class of deep neural networks**, most commonly applied to analyze visual imagery.

**Layers in CNN**

There are five different layers in CNN

- Input layer
- Convo layer (Convo + ReLU)
- Pooling layer
- Fully connected(FC) layer
- Softmax/logistic layer
- Output layer

**Input Layer**

Input layer in CNN should contain image data. Image data is represented by three dimensional matrix as we saw earlier. You need to reshape it into a single column.

**Convo Layer**

Convo layer is sometimes called feature extractor layer because features of the image are get extracted within this layer. First of all, a part of image is connected to Convo layer to perform convolution operation as we saw earlier and calculating the dot product between receptive field(it is a local region of the input image that has the same size as that of filter) and the filter. Result of the operation is single integer of the output volume.

Convo layer also contains ReLU activation to make all negative value to zero.

**Pooling Layer**

Pooling layer is used to reduce the spatial volume of input image after convolution. It is used between two convolution layer.

There is no parameter in pooling layer but it has two hyperparameters — Filter(F) and Stride(S).

In general, if we have input dimension W1 x H1 x D1, then

$$W2 = (W1-F)/S+1$$

$$H2 = (H1-F)/S+1$$

D2 = D1

Where W2, H2 and D2 are the width, height and depth of output.

**Fully Connected Layer(FC)**

Fully connected layer involves weights, biases, and neurons. It connects neurons in one layer to neurons in another layer. It is used to classify images between different category by training.

**Softmax / Logistic Layer**

Softmax or Logistic layer is the last layer of CNN. It resides at the end of FC layer.

**Output Layer**

Output layer contains the label which is in the form of one-hot encoded .

DEEP LEARNING:-

**Skikit-learn** was built on top of two Python libraries – NumPy and SciPy and has become the most popular Python machine learning library for developing machine learning algorithms.

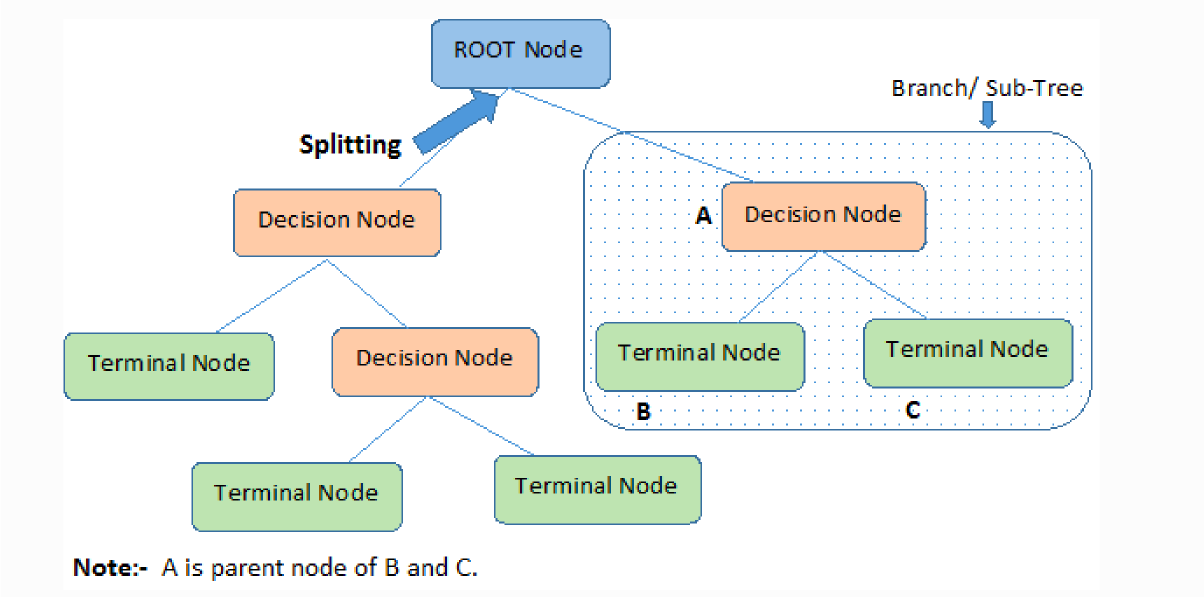
DECISION TREE ALGO:-

A decision tree is **a non-parametric supervised learning algorithm, which is utilized for both classification and regression tasks**. It has a hierarchical, tree structure, which consists of a root node, branches, internal nodes and leaf nodes.

Decision trees are extremely useful for data analytics and machine learning because **they break down complex data into more manageable parts**. They're often used in these fields for prediction analysis, data classification, and regression.

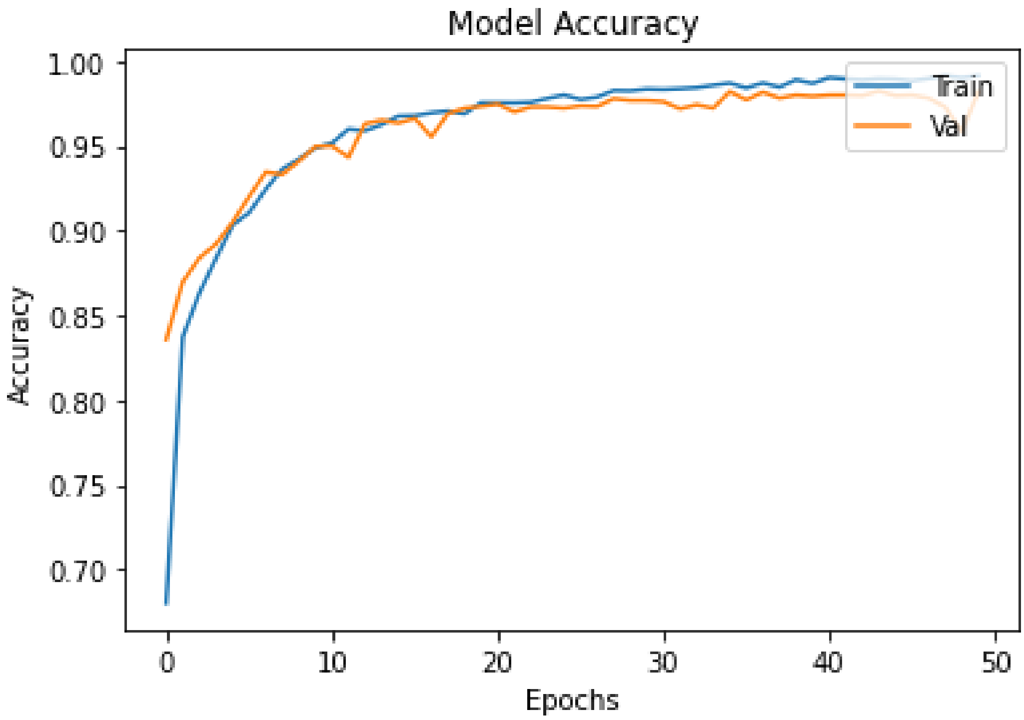
A decision tree is **one of the supervised machine learning algorithms**. This algorithm can be used for regression and classification problems — yet, is mostly used for classification problems. A decision tree follows a set of if-else conditions to visualize the data and classify it according to the conditions.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 01 | January - 2023 Impact Factor: 7.185 ISSN: 2582-3930

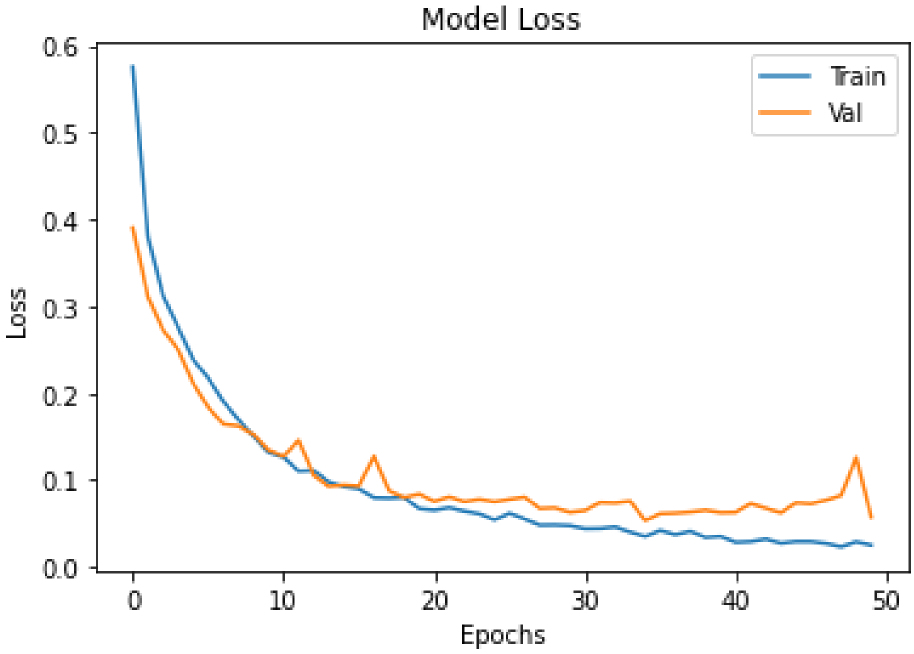



**RESULTS :-**

The working project has an accuracy of over 90% .



**Visualizing Power loss :-**




**REFERANCES :-**

[1] S. Selva Nidhyananthan, ... K. Gopalakrishnan, in Cognitive Systems and Signal Processing in Image Processing, 2022

[2] Al-Smadi, M., Abdulrahim, K., Salam, R.A. (2016). Traffic surveillance: A review of vision based vehicle detection, recognition and tracking. *International Journal of Applied Engineering Research*, *11*(1), 713–726.

[3] Radhakrishnan, M. (2013). Video object extraction by using background subtraction techniques for sports applications. *Digital Image Processing*, *5*(9), 91–97.

[4] Qiu-Lin, L.I., & Jia-Feng, H.E. (2011). Vehicles detection based on three-frame-difference method and cross-entropy threshold method. *Computer Engineering*, *37*(4), 172–174.

[5] Liu, Y., Yao, L., Shi, Q., Ding, J. (2014). Optical flow based urban road vehicle tracking. In *2013 Ninth International Conference on Computational Intelligence and Security*.

[6] Park, K., Lee, D., Park, Y. (2007). Video-based detection of street-parking violation. In International Conference on Image Processing. https://www.tib.eu/en/search/id/BLCP\%3ACN066390870/Video-based-detection-ofstreet-parking-violation, (Vol. vol. 1. IEEE, Las Vegas, pp. 152–156).

[7] Han, D., Leotta, M.J., Cooper, D.B., Mundy, J.L. (2006). Vehicle class recognition from video-based on 3d curve probes. In 2005 IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and *Surveillance*.

https://doi.org/10.1109/vspets.2005.1570927. IEEE.