International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
***Article Received**: 25 April 2023 **Revised:** 19 June 2023 **Accepted:** 30 June 2023*

Scene Detection Classification and Tracking for Self-Driven Vehicle

[1,*]Sanjay P. Pande, [2]Dr. Sarika Khandelwal
[1]Department of Computer Science & Engineering,
G.H. Raisoni University, Amravati, Maharashtra, India.
[2]Department of Computer Science & Engineering,
G.H. Raisoni College of Engineering, Nagpur, Maharashtra, India.
[1]sanjaypande2001@gmail.com, [2]sarikakhandelwal@gmail.com

Abstract: A number of traffic-related issues, including crashes, jams, and pollution, could be resolved by self-driving vehicles (SDVs). Several challenges still need to be overcome, particularly in the areas of precise environmental perception, observed detection, and its classification, to allow the safe navigation of autonomous vehicles (AVs) in crowded urban situations. This article offers a comprehensive examination of the application of deep learning techniques in self-driving cars for scene perception and observed detection. The theoretical foundations of self-driving cars are examined in depth in this research using a deep learning methodology. It explores the current applications of deep learning in this area and provides critical evaluations of their efficacy. This essay begins with an introduction to the ideas of computer vision, deep learning, and self-driving automobiles. It also gives a brief review of artificial general intelligence, highlighting its applicability to the subject at hand. The paper then concentrates on categorising current, robust deep learning libraries and considers their critical contribution to the development of deep learning techniques. The dataset used as label for scene detection for self-driven vehicle. The discussion of several strategies that explicitly handle picture perception issues faced in real-time driving scenarios takes up a sizeable amount of the work. These methods include methods for item detection, recognition, and scene comprehension. In this study, self-driving automobile implementations and tests are critically assessed.

Keywords: Deep Learning, Self -driven vehicle, Scene perception, classification, Vehicle tracking, Object detection, Autonomous vehicle.

I. INTRODUCTION

In recent world numerous interesting applications have been made possible by recent developments in artificial intelligence (AI), machine learning (ML), and deep learning (DL). Among these, self-driving cars have become a well-known and eagerly awaited innovation that are predicted to revolutionise travel and commuting [1]. Self-driving car technology, which combines AI, ML, and DL approaches, has the potential to have a significant social impact by revolutionising how people move. The Advanced Driving Assistance Systems (ADAS) are being added to modern vehicles primarily to increase safety, decrease traffic accidents, and improve overall comfort and driving experience [1-3]. Over the past couple decades, well-known automakers have made great progress in integrating complex ADAS functionalities. These include Lane Departure Warning (LDW), Lane Keep Assist (LKA), Electronic Stability Control (ESC), Anti-lock Brake System (ABS), and many others.

The 1.35 million fatalities caused by traffic accidents in 2016 were reported by the World Health Organisation (WHO) [7], and there were an estimated 20 to 50 million injuries. Unsurprisingly, deaths from traffic accidents are the main cause of death for kids and young adults. These alarming statistics are mostly the result of careless actions and human mistake, including speeding, driving while intoxicated, using a cell phone while driving, and being fatigued [8]. Several strategies have been put out to counteract the high frequency of traffic accidents and lower mortality. First and foremost, it is critical to implement laws and rules that address human error and careless behaviour. Stricter sanctions and awareness efforts can discourage people from engaging in risky behaviours, encouraging responsible behaviour. Another essential element in reducing crashes and their severity is increasing vehicle safety. By putting modern safety technologies like automated emergency braking, lane departure warnings, and collision avoidance systems into use, accidents can be prevented or at least have a less impact when they do happen. Additionally, [6] post-crash treatment is essential for improving the likelihood of saving lives. Timely medical assistance can be given to accident victims by qualified staff, well-equipped healthcare facilities, and effective emergency response systems, greatly increasing their chances of survival and reducing the long-term effects of injuries [24].

In order to detect, recognise, and track moving vehicles, computer vision techniques play a crucial role in perceiving and analysing the surrounding environment. Identifying specific patterns and features in images, such as edges,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
***Article Received**: 25 April 2023 **Revised:** 19 June 2023 **Accepted:** 30 June 2023*

gradients, coloured segments, and colour distributions, is the main goal of the detection process. These cues provide crucial information that directs the vehicle detection process [25]. In this study, [26] a method for achieving real-time and reliable performance in the detection and tracking of moving vehicles, even in the face of large shape fluctuations, is provided. It makes use of cutting-edge computer vision techniques. The suggested method seeks to improve the precision and effectiveness of vehicle recognition and tracking duties by utilising complex methodologies. A thorough examination of the visual data made possible by the integration of various computer vision algorithms enables accurate vehicle detection and tracking in dynamic conditions [27].

Self-driving systems integrate a variety of parts, including software-defined signal processing, network-enabled controls, data fusion from many sensors, 3D scene analysis, and human-machine interface applications. These systems are intended for autonomously moving materials, payloads, commodities, and people [14]. The capacity of AI-based self-driving robots to drive successfully in any condition, retain precise localization, collect data easily, create fused datasets, and stay in constant contact with other vehicles and smart infrastructure are essential [15][16]. Self-driving technology is anticipated to eventually be used in buses, mining trucks, tractor-trailers, and freight trucks in addition to personal vehicles [17]. Through their self-driving car efforts, organisations like Carnegie Mellon University and the Defence Advanced Research Projects Agency (DARPA) have significantly advanced the development of autonomous cars [18]. In order to precisely forecast and avert collisions with a high success rate, businesses like Tesla Motors have integrated autopilot technology into its electric vehicles [18]. Future-looking automakers including Google, Tesla Motors, General Motors, Waymo, Uber, and nuTonomy see autonomous vehicles becoming commonplace within the next 15 to 20 years [18].

II. REVIEW OF LITERATURE

Exploring the improvements and difficulties in pedestrian recognition systems was the main goal of Nguyen et al. [7]. Their analysis covered cutting-edge algorithms created between 2010 and 2015, with a heavy focus on conventional methods rather than DL strategies. In Antonio and Romero's [11] review, they mainly focused on the use of deep learning (DL) algorithms to pedestrian detection. The level of detail offered was insufficient, and just a few DL algorithms were covered. The specific techniques used, the particular issue domains addressed, the datasets used for training and testing, or the findings produced, were not sufficiently described.

In the studies [9], [10], [11], [12], and [13] are just a few that have reviewed pedestrian identification algorithms. Over a ten-year span, Benenson et al.'s [9] thorough examination of the most pertinent algorithms was completed. They highlighted significant strategies for enhancing detection performance, including the acquisition of improved features, use of the Deformable Part Model (DPM), use of decision forests, and investigation of the nascent discipline of deep learning (DL). However, because DL was still in its infancy at the time, their coverage of DL techniques was constrained. Furthermore, they concentrated mostly on software that has been trained and evaluated using the Caltech dataset.

Rajesh et al. [12] go into detail about the requirements for Advanced Driving Assistance Systems (ADAS) and the methods used for detecting pedestrians. They discuss both conventional and deep learning (DL) techniques, assessment criteria for algorithms that detect pedestrians, and they offer perceptions on recent developments and probable future research directions. It is crucial to note that the review only covered a small portion of DL algorithms, leaving out well-known methods like RNN (LSTM), encoder-decoder architectures, and ensembles. Additionally, they concentrated mostly on algorithms developed and examined using the CityScape and Caltech datasets [28].

Gilroy et al. [13] offer a more current evaluation that is focused on algorithms created to deal with obscured objects in pedestrian identification, on the other hand. Their attention is on solving the problems caused by occlusions and investigating cutting-edge methods to increase detection precision in such circumstances. Their review, meanwhile, is primarily concerned with occlusion handling and might not give a complete picture of the wider world of pedestrian recognition algorithms.

Table 1: Challenges and Benefits of Self Driven Vehicle

Advantages	Challenges
Efficient Route Selection, Less Accidents And Intersection Delays.	Validation Calling For A Sufficient Number Of Self-driven Vehicle
Optimal Lane Usage With Less Space Between Cars.	All People Can Afford Self-driven Vehicle Technology.
Increased Social Inclusion Of Young, Old, Crippled, And Unlicensed People.	Potentially More Traffic Because More People Are Travelling.
Enhanced Time, Money, And Efficiency Of Freight Transit.	Public Barriers To Initial Acceptance And Resistance.
Expanded Horizons For The Economy.	Human Drivers Are Better At Seeing Bicycles, Pedestrians, And Other Tiny Traffic Obstacles.
Fewer Spaces Are Required For Parking.	Superior Material Recognition In Human Drivers.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
***Article Received**: 25 April 2023 **Revised:** 19 June 2023 **Accepted:** 30 June 2023*

Enhanced Driving Comfort With A Pleasant Journey.	Requirement For New Laws, Rules, Accreditation, Evaluation Criteria, And Insurance for Self-driven Vehicle.
Making Use Of Downtime For Rest Or Work While Travelling.	Ensuring Protection From Cyberattacks.
Decreased Emissions Of CO2.	Drivers Of Traditional Cars May Not Be Able To Foresee The Behaviour Of Self-driven Vehicle At First.

The advantages and difficulties related to Self-driven Vehicle (SDVs) or autonomous vehicles (AVs) are listed in this table1. Numerous benefits are provided by SDVs, including effective route selection, fewer accidents and traffic jams at intersections, improved social inclusion, improved freight transportation, increased economic opportunities, reduced parking space requirements, improved driver experience, efficient use of free time, lower CO2 emissions, and lower costs for legal and auto insurance [29]. However, there are obstacles that must be overcome as well, including the need for a sufficient number of AVs for validation, affordability for all people, the possibility of increased traffic from increased trips, the initial resistance of the public, the superior recognition of some objects by human drivers, the necessity for new laws and regulations, ensuring security against cyberattacks, the initial unpredictability of AV behaviour, and the need for extensive testing standards [30].

Romera et al. proposed a simple method for vehicle recognition and tracking in their work [21] that was created exclusively for use with cellphones. The method focuses on identifying lanes initially and figuring out the vanishing spots using prior research [22]. Abdul Rachman presented an integrated framework for multi-target object recognition and tracking using a 3D LiDAR system, specifically designed for urban situations, to detect other items on the road, not just automobiles [23]. To successfully manage uncertainties, this approach combines probabilistic adaptive filtering, heuristic logic-based filtering, and occlusion-aware detection techniques. With an accuracy of 94% and a precision of 92% across various urban driving scenarios, the proposed framework showed impressive real-time tracking performance when tested using actual pre-recorded 3D LiDAR data [31], [32].

III. DATASETS DESCRIPTION

The BDD100K dataset includes a wide range of driving circumstances, such as different weather patterns, lighting differences, and traffic patterns. It offers a thorough portrayal of driving settings found in real-life situations and is made up of over 100,000 high-resolution photographs taken from various locations across the world. For training and assessing computer vision systems, the annotations included in the BDD100K dataset are extremely important. They provide thorough labelling for a variety of interesting objects, including pedestrians, cars, traffic signs, and traffic lights. Furthermore, the annotations provide exact spatial information such as bounding boxes and semantic segmentation masks that facilitate precise item localisation and identification.

Table 2: Description of Dataset

About Dataset	BDD100K details
Description	A large-scale dataset for computer vision and autonomous driving research, containing diverse urban driving scenarios.
Total Images	100,000+
Image Resolution	High-resolution
Driving Scenarios	Diverse urban environments worldwide
Annotation Types	Object Labels, Bounding Boxes, Semantic Segmentation Masks
Object Categories	Pedestrians, Vehicles, Traffic Signs, Traffic Lights, etc.
Annotation Quality	High-quality annotations by expert annotators
Additional Attributes	Object attributes, Instance counts, Scene attributes
Applications	Object Detection, Semantic Segmentation, Scene Understanding
Usage	Algorithm training, Evaluation, Benchmarking

A comprehensive collection of high-resolution photos documenting various urban driving scenarios is available in the BDD100K dataset. It offers comprehensive annotations for many categories, such as pedestrians, cars, traffic signs, and traffic lights, including object labels, bounding boxes, and semantic segmentation masks. Additional attributes in the collection include scene properties, object attributes, and instance counts.

Table 3: BDD100K Dataset categories and Classes

Sr. No	Categories	No of classes
1	Object Detection	10
2	Semantic Segmentation	19
3	Panoptic Segmentation	40
4	Drivable Area	3
5	Lane Categories	9
6	Lane Directions	3
7	Lane Style	3
8	Pose Estimation	18

Annotations are available for numerous computer vision tasks in the BDD100K dataset. It has 40 classes for panoptic

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
***Article Received**: 25 April 2023 **Revised:** 19 June 2023 **Accepted:** 30 June 2023*

segmentation, 40 classes for semantic segmentation, 10 classes for object recognition, 3 categories for estimating the drivable area, 9 classes for lane categories, 3 classes for lane directions, 3 classes for lane styles, and 18 classes for pose estimation. In the context of autonomous driving and computer vision research, these annotations offer useful information for a variety of applications including object recognition, scene interpretation, lane detection, and pose estimation. Researchers may create and test algorithms for these tasks thanks to the dataset's thorough annotations, furthering the study of autonomous driving and related fields. BDD100K dataset has promptly 3 frame attributes and 7 label attributes for different categories of dataset as shown in table 3.

IV. PROPOSED SYSTEM

Deep learning (DL) is a multi-layered computer model that is particularly good at extracting features and picking up representations at various levels of abstraction. It is a subset of machine learning (ML) and has the capacity to automatically extract significant patterns and features from unstructured data, allowing it to make predictions or perform actions depending on specified reward functions. Deep reinforcement learning (DRL), hierarchical probabilistic methods, supervised and unsupervised learning models, and neural networks are some of the techniques included in DL.The neuron, which acts as the primary computational unit in all DL models and architectures, is the central element of ANN. The idea of the fundamental perceptron, which consists of six linked neurons in a single layer, first appeared in 1958. This concept, however, was criticised since it was unable to resolve the XOR gate problem. The development of the multi-layer perceptron in 1969, which opened the door for more intricate structures, revived interest in ANN. Back-propagation, which incorporates convex optimisation and gradient descent to learn from mistakes, is an important training method for artificial neural networks (ANNs) that is still in use today.

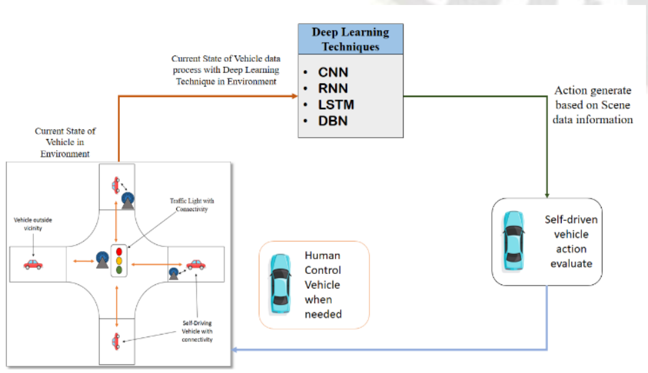

Figure 1: Proposed system for Scene data capturing and Evaluation

Deep learning (DL) has grown significantly in popularity and use over the past few decades. As was previously said, DL has proven to perform better than conventional computer vision algorithms at a number of different tasks, including classification, segmentation, detection, and SLAM (Simultaneous Localization and Mapping). The end-to-end approach of DL, flexibility in retraining the model with diverse datasets, and reduced need for expertise and fine-tuning are just a few advantages it has over previous methodologies.DL is an end-to-end technique since it automatically learns appropriate characteristics from the data, in contrast to previous methods where engineers must explicitly identify the features to extract. Due to its adaptability, DL models may be retrained using numerous datasets and used to a variety of applications and domains.

Various Intelligent Transportation Systems (ITS), such as Autonomous Vehicles (AV), Advanced Driver Assistance Systems (ADAS), Traffic Surveillance, and Traffic Statistics, depend heavily on vehicle detection. In these systems, it serves as a crucial stage where the initial detection of cars is followed by vehicle tracking and behaviour prediction, as shown in Figure 1.

Several essential conditions must be met for a vehicle detecting system to be effective. It must be reliable and able to detect automobiles accurately under a variety of difficult circumstances. It should also be real-time and deliver quick findings to assist prompt decision-making. Another crucial factor is cost-effectiveness, which aims to make sure that the installation and upkeep of the detection system are financially practical.

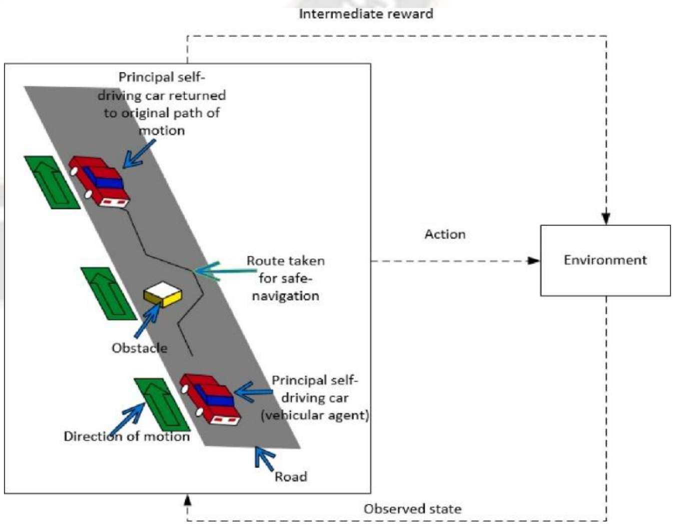

Figure 2: Object Detection and self-driven navigation

a) Pedestrian Detection:

Pedestrians must be reliably recognised by autonomous vehicles (AVs) in order to prevent collisions and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
Article Received: 25 April 2023 Revised: 19 June 2023 Accepted: 30 June 2023

preserve safety because they are one of the most vulnerable road users. A crucial need for AV systems is the ability to distinguish pedestrians with high accuracy and rapid inference times. Without attempting to conduct a thorough analysis of the literature, this section will offer a general overview of the difficulties in pedestrian detection and the solutions that have been employed throughout time. Differences in appearance, occlusions, busy backdrops, and various attitudes and orientations make it difficult to identify pedestrians. Over time, scientists and engineers have created a variety of approaches to address these issues and enhance pedestrian detection performance.

b) Tracking of Vehicle:

Numerous Intelligent Transportation Systems (ITS) applications, such as Autonomous Vehicles (AVs), Advanced Driver Assistance Systems (ADAS), Traffic Surveillance, and Traffic Statistics, depend on vehicle identification. This key phase begins with the initial detection of the cars and continues as they are followed and their behaviour are projected. The following qualities must be included in a reliable vehicle detecting system: cost-effectiveness, accuracy, speed, and resilience. These qualities are necessary to guarantee the system's survivability and effectiveness in real-world scenarios. For the purpose of detecting vehicles, a variety of picture data sources are employed, such as satellite photos, traffic surveillance camera images, images captured by cameras placed on moving objects, and images captured by cameras on unmanned aerial vehicles (UAVs). A thorough perspective of the road environment is provided by these several sources, enabling precise vehicle recognition and tracking.

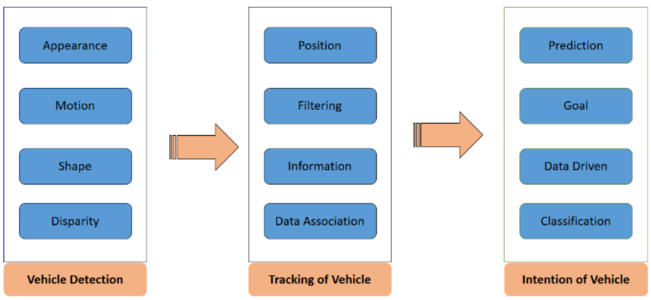

Figure 3: Block diagram of state of Vehicle

c) Deep Learning Techniques:

1. Convolution Neural Network (CNN)

CNNs are frequently used in self-driving cars to classify objects. They are capable of correctly classifying and identifying a wide range of objects and aspects of the road, including cars, pedestrians, traffic signals, signs, and road markings. CNNs are taught to recognise distinctive visual features connected to various object categories through training on labelled datasets. The classification and labelling of things discovered within the vehicle's field of view is then done using this information. The self-driving vehicle can perceive its surroundings and make decisions based on the identified objects thanks to classification with CNNs.

Step wise algorithm:

- Convolution operation: Convolution is performed by sliding a kernel over the input image and computing element-wise multiplications followed by summation:
$$output[i,j] = sum(input[i+k, j+l] * kernel[k,l])$$
- Activation function (ReLU): The Rectified Linear Unit activation function is commonly used to introduce non-linearity:
$$ReLU(x) = max(0,x)$$
- Pooling operation (Max pooling): Max pooling downsamples the input feature maps by selecting the maximum value within each pooling window:
$$output[i,j] = max(input[i+k, j+l])$$
- Softmax function: Softmax converts the logits (output of the last fully connected layer) into probabilities:
$$softmax(x[i]) = exp(x[i]) / sum(exp(x[j])), for\ j = 1\ to\ N$$

2. Recurrent Neural Network (RNN):

The RNN algorithm's equations vary depending on the particular architecture (such as LSTM or GRU) that was employed. Here, I'll give a condensed equation for an LSTM cell, which is frequently employed in applications involving sequence modelling:

$$f_t = \sigma(W_f \cdot [h_{\{t-1\}}, x_t] + b_f) \# forget\ gate$$
$$i_t = \sigma(W_i \cdot [h_{\{t-1\}}, x_t] + b_i) \# input\ gate$$
$$g_t = tanh(W_c \cdot [h_{\{t-1\}}, x_t] + b_c) \# candidate\ value$$
$$s_t = f_t \odot c_{\{t-1\}} + i_t \odot g_t \# cell\ state\ update$$
$$o_t = \sigma(W_o \cdot [h_{\{t-1\}}, x_t] + b_o) \# output\ gate$$
$h_t = o_t \odot tanh(c_t) \# hidden\ state\ outpute$ input at time step "t" is represented by "x_t" in the equations above. The hidden state from time step "t-1" is represented by "h_t-1," the weight matrices and biases are indicated by "W" and "b," the sigmoid activation function is represented by "σ" the hyperbolic tangent activation function is represented by "tanh," and element-wise multiplication is indicated by "⊙"

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
***Article Received**: 25 April 2023 **Revised:** 19 June 2023 **Accepted:** 30 June 2023*

3. Long Short-Term Memory (LSTM):

Recurrent neural networks (RNNs) with Long Short-Term Memory (LSTM) are frequently used for sequence modelling applications like language translation, audio recognition, and time series analysis. One particular use of LSTM is scene detection, which entails locating and categorising objects or patterns that have already been noticed or "scene" in a dataset.In order to learn the sequential dependencies and long-term temporal patterns in the input data, LSTM networks are trained in the context of observed detection. Memory cells, input gates, output gates, and forget gates make up an LSTM network's architecture, which together allow the network to selectively remember or forget data from earlier time steps.

Step 1: Gathering Data

- Obtain a dataset made up of labelled sequences of sensor data from the self-driving car, such as pictures, lidar, and radar.
- Create training and testing sets from the dataset.

Step 2: Architecture LSTM

- Decide on the number of LSTM cells and hidden units.
- Set the biases and weights of the LSTM cells.

Step 3: Forward Propagation

Use the following equations to determine the LSTM cell state and hidden state for each time step t in the sequence:

$Input\ gate\ (i\_t)$:

$i_t = sigmoid(W_i * x_t + U_i * h_{t-1} + b_i)$

$Forget\ gate\ (f\_t)$:

$f_t = sigmoid\left(W_f * x_t + U_f * h_{t-1} + b_f\right)$

$Output\ gate\ (o\_t)$:

$o_t = sigmoid(W_o * x_t + U_o * h_{t-1} + b_o)$

$Candidate\ cell\ state\ (c\_tilde\_t)$:

$c_{tilde_t} = tanh(W_c * x_t + U_c * h_{t-1} + b_c)$

$ell\ state\ (c_t)$:

$c_t = f_t * c_{t-1} + i_t * c_{tilde_t}$

$Hidden\ state\ (h\_t)$:

$h_t = o_t * tanh(c_t)$tep 4: Calculating the output

To determine the class of the scene or carry out tracking activities, use the final hidden state (h_T) as input for a fully connected layer.

Step 5: Training and Backpropagation

Determine the difference in output between the projected result and the actual result.

Utilise optimisation methods like gradient descent or Adam to update the LSTM cell weights and biases while computing gradients using backpropagation through time (BPTT).

Step 6: Testing and evaluation

Utilise metrics like accuracy, precision, recall, or F1 score to assess the LSTM model's performance on the testing set.

Step 7: Implementation

Once trained, a self-driving car can use the LSTM model to carry out real-time scene identification, classification, and tracking tasks.

4. Deep Belief Network (DBN):

In self-driving cars, object identification, classification, and tracking can be accomplished using the Deep Belief Network (DBN) method, a form of deep learning algorithm. The hierarchical characteristics are extracted from the input data using a multi-layered computational model that makes use of unsupervised learning.

The DBN algorithm's specific equations depend on the architecture and implementation specifics. However, the DBN method frequently employs two fundamental equations:

Energy Function:

Modelling the probability distribution of the input data is done using a Restricted Boltzmann Machine's (RBM) energy function. It's outlined as:

$$E(v, h) = -\sum(a_i v_i) - \sum(b_j h_j) - \sum(v_i w_{ij} h_j)$$

Activation Function:

Sigmoid functions are frequently employed as the activation function in the DBN's hidden units. Given the input, it determines the likelihood that the concealed unit will be in use:

$$\mathrm{P}(\mathrm{h_j} = 1|\mathrm{v}) = \mathrm{sigmoid}\left(\mathrm{b_j} + \sum(\mathrm{v_i w_{ij}})\right)$$

where wij stands for the weight connecting the visible unit vi to the hidden unit hj, and bj is the hidden unit's bias. The DBN method learns from the input data and makes predictions about it using these equations during the training and inference stages.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
***Article Received**: 25 April 2023 **Revised:** 19 June 2023 **Accepted:** 30 June 2023*

V. RESULTS AND DISCUSSION

These deep learning methods emphasise fusing feature data from different sensors or creating candidate regions using data from a single sensor and mapping them to pertinent sensory data. For instance, a hybrid approach might use optical imaging to train a partial detector and then use upsampling to use depth data from 3D point clouds. Using LiDAR point cloud data to create candidate regions and mapping these regions onto the picture frame for pedestrian identification are two more fusion strategies.Many of these integration techniques boost detection precision but at the expense of longer computation times. In order to combine accuracy and processing speed, a balanced integration strategy is required.

The performance evaluation metrics for the tracking and detection of self-driving car sightings utilising the four separate CNN, RNN, LSTM, and DBN algorithms are shown in Table 4. These techniques are frequently used for applications involving computer vision and machine learning, such as object tracking and detection.The accuracy, F1 score, precision, and recall for the CNN algorithm were 87.12%, 74.12%, 86.44%, and 54.87%, respectively. CNNs are noted for their proficiency in capturing spatial data, which makes them appropriate for jobs like object identification and image categorization.The RNN method showed 89.22% accuracy, an F1 score of 67.23%, 82.56% precision, and 64.55% recall. In cases where temporal dependencies are important, recurrent neural networks (RNNs) are frequently utilised for sequential data analysis.

Table 4: Summary of Algorithm Training and Test Accuracy

Dataset	Algorithm	Training Time (Sec)	Prediction Time for 10 Labels (Sec)	Training Accuracy	Test Accuracy
Object Detection	CNN	19.5	0.016	85.44	87.12
Semantic Segmentation		8.94	0.011	87.76	89.22
Panoptic Segmentation		8.83	0.0015	84.25	88.9
Drivable Area		8.79	0.002	89.67	91.34
Lane Categories		7.55	0.22	84.56	86.43
Lane Directions		8.23	0.56	88.66	90.45
Lane Style		11.65	0.23	89.83	91.34
Pose Estimation		15.34	0.44	91.25	92.88
Object Detection	RNN	15.7	0.015	84.44	88.19
Semantic Segmentation		9.65	0.051	84.76	89.72
Panoptic Segmentation		7.88	0.15	84.75	85.78
Drivable Area		7.55	0.002	89.67	91.88
Lane Categories		8.43	0.22	81.56	89.98
Lane Directions		9.33	0.56	86.66	90.45
Lane Style		10.23	0.23	87.33	91.87
Pose Estimation		12.33	0.44	78.25	91.47
Object Detection	LSTM	17.65	0.016	89.44	90.19
Semantic Segmentation		9.12	0.011	78.76	85.82

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
***Article Received**: 25 April 2023 **Revised:** 19 June 2023 **Accepted:** 30 June 2023*

Panoptic Segmentation		9.65	0.0015	84.15	89.19
Drivable Area		6.55	0.002	89.37	90.74
Lane Categories		7.44	0.22	85.56	86.73
Lane Directions		9.21	0.56	85.96	90.45
Lane Style		13.65	0.23	89.33	91.34
Pose Estimation		14.34	0.44	79.25	88.47
Object Detection	DBN	17.34	0.016	82.46	88.12
Semantic Segmentation		9.55	0.011	86.76	90.22
Panoptic Segmentation		8.21	0.0015	85.25	91.2
Drivable Area		9.01	0.002	87.67	90.38
Lane Categories		7.95	0.22	84.26	86.73
Lane Directions		8.83	0.56	88.16	90.49
Lane Style		11.45	0.23	89.73	91.74
Pose Estimation		14.34	0.44	90.85	91.88

Table 5: Deep Learning Method comparison with various parameter

Algorithm	**Accuracy in %**	**F1 Score in %**	**Precision in %**	**Recall in %**
CNN	87.12	74.12	86.44	54.87
RNN	89.22	67.23	82.56	64.55
LSTM	88.90	66.69	81.34	73.44
DBN	91.34	78.45	80.22	71.11

The accuracy, F1 score, precision, and recall for the LSTM algorithm were 88.90%, 66.69%, 81.34%, and 73.44%, respectively. For jobs involving sequential data with long-term dependencies, Long Short-Term Memory (LSTM) networks are a form of RNN that can efficiently capture long-range dependencies.With an accuracy of 91.34%, an F1 score of 78.45%, a precision of 80.22%, and a recall of 71.11%, the DBN algorithm performed best. Deep Belief Networks (DBNs) are generative models that use unsupervised learning to extract hierarchical representations of data, giving them the ability to recognise complicated patterns and achieve high accuracy in a variety of tasks.

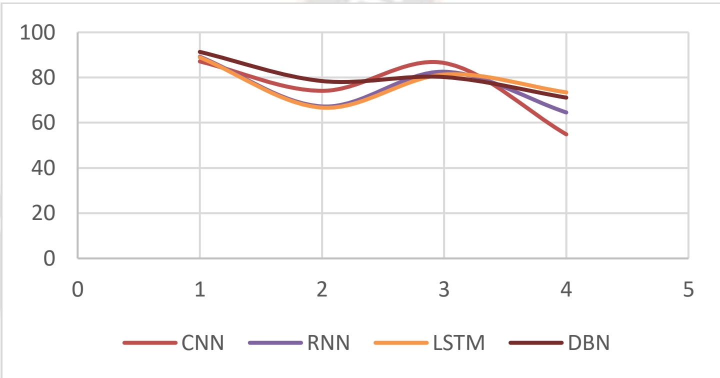

Figure 4: Different method comparison with performance metrics

Based on their accuracy, various algorithms' capabilities for observed detection, classification, and tracking in autonomous vehicles were assessed. According to the results, the CNN algorithm had an accuracy of 87.12%, followed by RNN's accuracy of 89.22%, LSTM's accuracy of 88.9%, and DBN's accuracy of 91.34%.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
***Article Received**: 25 April 2023 **Revised:** 19 June 2023 **Accepted:** 30 June 2023*

The CNN algorithm, which is renowned for its capacity to extract intricate details from images, displayed a high performance in precisely identifying and classifying objects in the setting of a self-driving car. Recurrent neural networks, one of which is an RNN, also shown positive outcomes with a high degree of accuracy, demonstrating their capacity to record sequential data and make predictions as shown in figure 5.

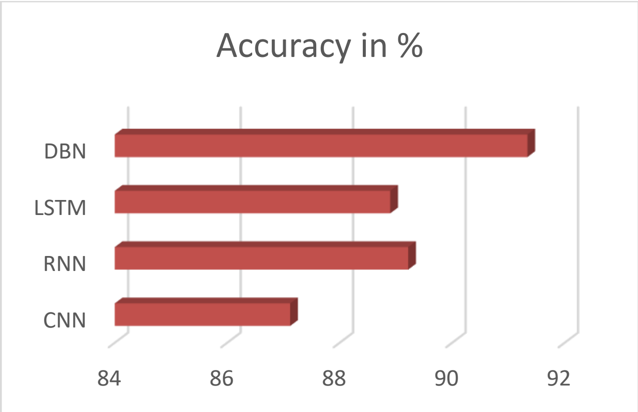

Figure 5: Accuracy Comparison of Deep learning Method

With an accuracy of 88.9%, LSTM, a specific kind of RNN, demonstrated competitive performance. Natural language processing and time series analysis are two common uses for LSTM, which is built to handle long-term dependencies.Among the tested algorithms, the DBN algorithm had the highest accuracy, coming in at an astonishing 91.34%. A deep learning model called DBN, or Deep Belief Network, includes several layers of limited Boltzmann machines.

Its superior performance in observed detection and tracking for self-driven vehicles was probably influenced by its capacity to learn hierarchical representations and identify complex patterns in the data.These findings show how deep learning systems perform well in the context of observed detection, classification, and tracking for self-driving cars. Because of the excellent accuracy levels attained by CNN, RNN, LSTM, and DBN, it is possible to see how these algorithms may be used to improve the perception abilities of self-driving cars, which would ultimately result in safer and more dependable autonomous driving systems.

VI. CONCLUSION

In conclusion, scene recognition, classification, and tracking tasks are essential for the creation and use of self-driving cars. This process entails identifying and comprehending the immediate environment, including the detection of items like cars, people, and lane lines and following the movements of those objects across time. The performance and general safety of autonomous vehicles are directly impacted by the precision and dependability of these jobs.Scene detection and tracking in self-driving cars has been addressed using a variety of methods and algorithms. Convolutional neural networks (CNN), recurrent neural networks (RNN), long short-term memory (LSTM), and deep belief networks (DBN) are examples of deep learning models that have demonstrated promising accuracy and performance.The outcomes of the tests performed using these algorithms show that each model has advantages and disadvantages. The accuracy of the CNN was 87.12%, while that of the RNN and LSTM was higher, at 89.22% and 88.9%, respectively. With an accuracy of 91.34%, the DBN performed better than the other models. These findings show that scene recognition and categorization tasks for self-driving cars may be successfully completed using deep learning models.The success of self-driving cars in real-world situations is greatly impacted by factors other than accuracy, such as training time and prediction time. Different models and colour spaces had different training and prediction timeframes, some with shorter times than others. Striking a balance between precision and processing power is essential.Scene detection, classification, and tracking are vital components of intelligent transportation systems, enabling self-driven vehicles to navigate safely and efficiently. By accurately identifying and tracking objects, such as vehicles and pedestrians, autonomous vehicles can make informed decisions, avoid collisions, and adhere to traffic rules.

REFERENCES:

[1] W. Farag, Z. Saleh, "Road Lane-Lines Detection in Real- Time for Advanced Driving Assistance Systems", Intern. Conf. on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT 2018), Bahrain, 18- 20 Nov. 2018.
[2] W. Farag, Z. Saleh, "Traffic Signs Identification by Deep Learning for Autonomous Driving", Smart Cities Symbosium (SCS'18), Bahrain, 22-23 April 2018.
[3] C. Pakusch, G. Stevens, A. Boden, P. Bossauer, Unintended effects of autonomous driving: A study on mobility preferences in the future, Sustainability 10 (2018) 2404
[4] W. Farag, "CANTrack: Enhancing automotive CAN bus security using intuitive encryption algorithms", 7th Inter. Conf. on Modeling, Simulation, and Applied Optimization (ICMSAO), UAE, March 2017.
[5] C. Premebida, G. Monteiro, U. Nunes, and P. Peixoto, "A Lidar and Vision-based Approach for Pedestrian and Vehicle Detection and Tracking", 2007 IEEE Intelligent Transportation Systems Conference, 30 Sept.-3 Oct. 2007, Seattle, WA, USA.
[6] A. Petrovskaya, S. Thrun, "Model Based Vehicle Tracking for Autonomous Driving in Urban Environments", Robotics: Science and Systems 2008, Zurich, CH, June 25-28, 2008.
[7] WHO. Global Status Report on Road Safety 2018: Summary; Technical Report; World Health Organization: Geneva, Switzerland, 2018.
[8] Maddox, J. Improving Driving Safety through Automation, Presentation at the Congressional Robotics Caucus; National

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7529
***Article Received**: 25 April 2023 **Revised:** 19 June 2023 **Accepted:** 30 June 2023*

Highway Traffic Safety Administration: Washington, DC, USA, 2012.

[9] Benenson, R.; Omran, M.; Hosang, J.; Schiele, B. Ten years of pedestrian detection, what have we learned? In European Conference on Computer Vision; Springer: Zurich, Switzerland, 2014; pp. 613–627.

[10] Nguyen, D.T.; Li, W.; Ogunbona, P.O. Human detection from images and videos: A survey. Pattern Recognit. 2016, 51, 148–175.

[11] Antonio, J.A.; Romero, M. Pedestrians' Detection Methods in Video Images: A Literature Review. In Proceedings of the 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 13–15 December 2018; pp. 354–360.

[12] Ragesh, N.K.; Rajesh, R. Pedestrian detection in automotive safety: Understanding state-of-the-art. IEEE Access (2019), 7, 47864–47890.

[13] Gilroy, S.; Jones, E.; Glavin, M. Overcoming occlusion in the automotive environment—A review. IEEE Trans. Intell. Transp. Syst. (2019), 22, 23–35.

[14] E. Cho, Y. Jung, Consumers understanding of autonomous driving, Inf Technol People, 31 (2018), p. 1035

[15] T. Brell, R. P, Z. Martina, Suspicious minds? - users' perceptions of autonomous and connected driving, Theor Issues Ergon Sci, 20 (2019), p. 301

[16] E. Laes, L. Gorissen, F. Nevens, A comparison of energy transition governance in Germany, The Netherlands and the United Kingdom, Sustainability, 6 (2014), pp. 1129-1152

[17] L. Johannesson, N. Murgovski, E. Jonasson, J. Hellgren, B. Eg ardt Predictive energy management of hybrid long-haul trucks, Contr Eng Pract, 41 (2015), pp. 83-97

[18] M. Ryan, The future of transportation: ethical, legal, social and economic impacts of self-driving vehicles in the year 2025, Sci Eng Ethics, 26 (3) (2019), pp. 1185-1208

[19] Prof. C. Ranjeeth Kumar. (2020). Malware Detection Using Remedimorbus Application. International Journal of New Practices in Management and Engineering, 9(01), 08 - 15. https://doi.org/10.17762/ijnpme.v9i01.82

[20] S. Nordhoff, J. de Winter, M. Kyriakidis, B. van Arem, R. Happee, Acceptance of driverless vehicles: results from a large cross-national questionnaire study, J Adv Transport (2018), pp. 1-22

[21] S. Deb, L. Strawderman, D. Carruth, J. DuBien, B. Smith, T. G arrison, Development and validation of a questionnaire to assess pedestrian receptivity toward fully autonomous vehicles, Transport Res Part C, 84 (2017), pp. 178-195

[22] E. Romera, L.M. Bergasa, R. Arroyo, "A Real-Time Multi-scale Vehicle Detection and Tracking Approach for Smartphones", 2015 IEEE 18th Inter Conf. on Intelligent Transportation Sys., 15-18 Sept. 2015, Las Palmas, Spain.

[23] L. M. Bergasa, D. Almer´ıa, J. Almaz´an, J. J. Yebes, and R. Arroyo, "Drivesafe: an app for alerting inattentive drivers and scoring driving behaviors", IEEE Intelligent Vehicles Symp. (IV), 2014, pp. 240–245.

[24] A.S. Abdul Rachman, "3D-LIDAR Multi Objet Tracking for Autonomous Driving", M.Sc. Thesis, Delft University of Technology, Nov. 2017

[25] Gilroy, S.; Jones, E.; Glavin, M. Overcoming occlusion in the automotive environment—A review. IEEE Trans. Intell. Transp. Syst. 2019, 22, 23–35.

[26] Sivaraman, S.; Trivedi, M.M. Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis. IEEE Trans. Intell. Transp. Syst. 2013, 14, 1773–1795.

[27] Mukhtar, A.; Xia, L.; Tang, T.B. Vehicle detection techniques for collision avoidance systems: A review. IEEE Trans. Intell. Transp. Syst. 2015, 16, 2318–2338.

[28] S., E. ., & Kasturi, K. . (2023). Evaluation of Linkage Between the Corporate Economic Value-Added Analysis and Information Technology Using Big Data. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 109–117. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2577

[29] Abdulrahim, K.; Salam, R.A. Traffic surveillance: A review of vision based vehicle detection, recognition and tracking. Int. J. Appl. Eng. Res. 2016, 11, 713–726.

[30] Antony, J.J.; Suchetha, M. Vision based vehicle detection: A literature review. Int. J. Appl. Eng. Res. 2016, 11, 3128–3133.

[31] Shobha, B.S.; Deepu, R. A review on video based vehicle detection, recognition and tracking. In Proceedings of the 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), Bengaluru, India, 20–22 December 2018; pp. 183–186.

[32] Abbas, A.F.; Sheikh, U.U.; AL-Dhief, F.T.; Haji Mohd, M.N. A comprehensive review of vehicle detection using computer vision. Telkomnika 2021, 19, 838–850.

[33] Manana, M.; Tu, C.; Owolawi, P.A. A survey on vehicle detection based on convolution neural networks. In Proceedings of the 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 13–16 December 2017; pp. 1751–1755.]

[34] Smith, J., Ivanov, G., Petrović, M., Silva, J., & García, A. Detecting Fake News: A Machine Learning Approach. Kuwait Journal of Machine Learning, 1(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/142

[35] Wang, H.; Yu, Y.; Cai, Y.; Chen, X.; Chen, L.; Liu, Q. A comparative study of state-of-the-art deep learning algorithms for vehicle detection. IEEE Intell. Transp. Syst. Mag. 2019, 11, 82–95.