International Journal of Advanced Engineering and Nano Technology (IJAENT)
ISSN: 2347-6389 (Online), Volume-11 Issue-12, December 2024

Meta-Analysis and Review of Artificial Intelligence (AI) and Deep Learning Algorithms on Autonomous Vehicles (Avs) Via Vision-Based System: Current Trends, Issues, and Future Direction

Fidelis Nfwan Gonten, AY Gital, Datti Useni Emmanuel, Larson Kefas Suwa, Mudimka Jahota Yerima

Abstract: ***The invention of autonomous vehicles (AVs) and their use in transportation have been substantially accelerated by technological developments in artificial intelligence (AI) and deep learning Algorithms. Vision-based systems are a crucial part of AVs for detecting their surroundings and making the right decisions. At the same time, they are in motion, thanks to massive data from numerous sensor devices and sophisticated computing power. They understand how AI and deep learning functions in AV systems are crucial in achieving the objective of full automation, or self-driving, systems. Previous studies have done a fantastic job of looking into various facets of using AI and deep learning in AV production. Nevertheless, few studies have provided a comprehensive analysis of existing methods for integrating AI in AVs to the research community. This paper offers a systematic review of the most important papers in this field of research. It seeks to close the knowledge gap by providing state-of-the-art practices, challenges, and future direction. Its specific goal is to examine the various algorithms, models, and techniques applied to AVs by enhancing AI and deep learning for effective vision, navigation, and location in making decisions. It looks into the methods now in use to determine the potential applications of AI and the difficulties and problems that come with putting them into practice. This study offers more insights into possible opportunities for utilizing AI and deep learning in conjunction with other developing technologies, based on an examination of current practices and technological advancements. Big data, high computing power, and high-resolution navigation; expanded simulation platforms through a vision-based system.***

Keywords: ***Autonomous Vehicle, Segmentation, Sensor, Artificial Intelligence, Deep Learning, Vision-Based System***

Manuscript received on 25 August 2024 | Revised Manuscript received on 03 December 2024 | Manuscript Accepted on 15 December 2024 | Manuscript published on 30 December 2024.
***Correspondence Author(s)**
Fidelis Nfwan Gonten*, Department of Computer Science, Abubakar Tafawa Balewa University Bauchi, Bauchi, Nigeria. Email ID: gontenlur@gmail.com, ORCID ID: 0009-0001-1037-6031
Professor Abdulsalam Ya'u Gital, Department of Computer Science, Abubakar Tafawa Balewa University Bauchi, Bauchi, Nigeria. Email ID: aygital@gmail.com
Mr. Datti Useni, Department of Computer Science, Abubakar Tafawa Balewa University Bauchi, Bauchi, Nigeria. Email ID: dattiemmanueluseni@gmail.com
Mr. Larson Suwa, Department of Computer Science, Abubakar Tafawa Balewa University Bauchi, Bauchi, Nigeria. Email ID: larsonsuwa@gmail.com
Mr. Mudimka Jahota Yerima, Department of Computer Science/Statistics, Plateau State Polytechnic, Plateau, Nigeria. Email ID: gontenlur@gmail.com

I. INTRODUCTION

The ability of autonomous vehicle technological facilities to enhance the vision-based system through the application of AI and deep learning in detecting traffic signs, pathways, framework delivery, pedestrians and lane detection, etc., has tremendously evolved and transformed computer vision and vision-based architectures or models. The major achievement recorded was the use of neural networks and the number of layers built with the capability to mimic the human brain as against the traditional use of Histogram of oriented gradient (HOG), Viola-Jones algorithm, etc (Goyal, et al., 2020).

Recently, AI-based picture recognition has progressed to the point where it is similar to humans for basic tasks like identifying photos.(He et al., 2016) [19] The study of how computers can comprehend digital photos or movies and automate operations that the human visual system can accomplish is known as computer vision (CV). The word "CV" refers to AI-based image recognition and image sensing technology because this field works with all aspects of real-world computer information gathering. (Hashimoto et al., 2018).With the aid of a vision-based system, several forms of information can be extracted from a single image depending on factors like morphology, contrast, colour, size shape, etc. The information can be perceived and enhanced through the use of mathematical models. The majority of these mathematical techniques can be used to streamline the process of data perception and processing by utilizing computer vision libraries. Similar to how the human brain functions, a vision-based system may process, alter, and display images to comprehend visual data (Ribeiro et al., 2023) [33].

In various driving environments, autonomous cars need an accurate localization and mapping solution. SLAM, or simultaneous localization and mapping, technology is a well-researched solution in this regard. For perception and localization, cameras and light detection and ranging (LIDAR) sensors are frequently employed. But after ten or twenty years of development, the LIDAR-SLAM technique doesn't appear to have evolved all that much. When compared to LIDAR-based systems, visual SLAM offers the benefits of low cost and simple installation together with powerful scene recognition capabilities (Cheng et al., 2022).

Our study's goal was to perform a thorough analysis of nearly the most significant papers on the use of AI and deep learning (DL) in building a vision-based system for autonomous vehicles that are related to object detection, vehicle classification, image segmentation, object-based image analysis, scene classification, image fusion, and image registration. We compiled the key scientific findings mentioned in the literature and used a meta-analysis to identify and classify the papers about DL and vision-based systems. Lastly, a critical assessment, techniques, issues, forecast for further research, and conclusion are provided.

II. BACKGROUND OF AUTONOMOUS VEHICLE (AV)

The ability of autonomous vehicle technological facilities to enhance the vision-based system through the application of AI and deep learning in detecting traffic signs, pathways, framework delivery, pedestrians and lane detection, etc., has tremendously evolved and transformed computer vision and vision-based architectures or models. The major achievement recorded was the use of neural networks and the number of layers built with the capability to mimic the human brain as against the traditional use of Histogram of oriented gradient (HOG), Viola-Jones algorithm, etc. (Goyal, et al., 2020).

Recently, AI-based picture recognition has progressed to the point where it is similar to humans for basic tasks like identifying photos.(He et al., 2016) The study of how computers can comprehend digital photos or movies and automate operations that the human visual system can accomplish is known as computer vision (CV). The word "CV" refers to AI-based image recognition and image sensing technology because this field works with all aspects of real-world computer information gathering. (Hashimoto et al., 2018) [18]. With the aid of a vision-based system, several forms of information can be extracted from a single image depending on factors like morphology, contrast, colour, size shape, etc. The information can be perceived and enhanced through the use of mathematical models. The majority of these mathematical techniques can be used to streamline the process of data perception and processing by utilizing computer vision libraries. Similar to how the human brain functions, a vision-based system may process, alter, and display images to comprehend visual data (Ribeiro et al., 2023).

In various driving environments, autonomous cars need an accurate localization and mapping solution. SLAM, or simultaneous localization and mapping, technology is a well-researched solution in this regard. For perception and localization, cameras and light detection and ranging (LIDAR) sensors are frequently employed. But after ten or twenty years of development, the LIDAR-SLAM technique doesn't appear to have evolved all that much. When compared to LIDAR-based systems, visual SLAM offers the benefits of low cost and simple installation together with powerful scene recognition capabilities (Cheng et al., 2022).

III. MONITORING AND SENSOR DEVICES

Geographic Positioning Systems (GPS) have grown into an integral component in transportation and other businesses as well as daily life. The GPS units triangulate their position using signals from satellites in orbit to determine their global coordinates. (Bagloee et al., 2016) [4]. To pinpoint a car's location on the road, these coordinates are compared with the topography of the network of roads. There are still significant GPS inaccuracies in many meters. (Miura & Kamijo, 2015) [29]. Because automotive radar scenarios are distinct from military radar scenarios, new signal-processing techniques must be developed and derived. New performance constraints and the reformulation of vehicular radar positions offer a chance to create novel signal-processing techniques. By updating standard methods for signal processing in automotive radar (Bilik et al., 2019) [6].

Innovative technologies, such as navigational visual algorithms, obstacle detection and avoidance, and aerial decision-making, can be developed to address challenges with aerial perception through the integration of computer vision technology into unmanned aerial vehicles (UAVs). A wide range of uses for UAVs have been developed using this specialized technology, going beyond traditional military and defence applications. (Al-Kaff et al., 2018) [2]. In various driving environments, autonomous cars need an accurate localization and mapping solution. (Cheng et al., 2022). SLAM, or simultaneous localization and mapping, technology is a well-researched solution in this regard. For perception and localization, cameras and light detection and ranging (LIDAR) sensors are frequently employed. But after ten or twenty years of development, the LIDAR-SLAM technique doesn't appear to have evolved all that much. When compared to LIDAR-based systems, visual SLAM offers the benefits of low cost and simple installation together with powerful scene recognition capabilities(Cheng et al., 2022) [11]. (Nguyen et al., 2021) Suggests a productive method, based on artificial intelligence (AI) platforms for data processing, to maximize data transmission and reception in UAV-CSS systems. An initial backdrop frame and an updated background are generated by the algorithm and supplied to the server. After dividing the scene's region of interest, it sends only the modifications.

The autonomous car is capable of self-control, environment recognition, and driving at a level comparable to that of a human driver. To develop this type of system, the following tasks were completed: To detect curve lanes, deep learning is applied for lane detection, based on fully Convolutional Neural Networks (CNNs); (c) The cubic spline interpolation method is used for path generation, based on Global Positioning System (GPS) data, where the distance between two adjacent path generation points is the same. This real-time lane detection system is proposed, based on vision system functions, using webcam cameras. A sliding mode control system is suggested for the steering control of the vehicle. (Dai & Lee, 2020) [13].

IV. VISION-BASED ALGORITHM FOR IMAGE PROCESSING

The use of machine learning and Neural network (AI) enhanced the driver assistance model and self-driven vehicles in which, self-driven functions are categories, conducted in simulation and HIL system. This process

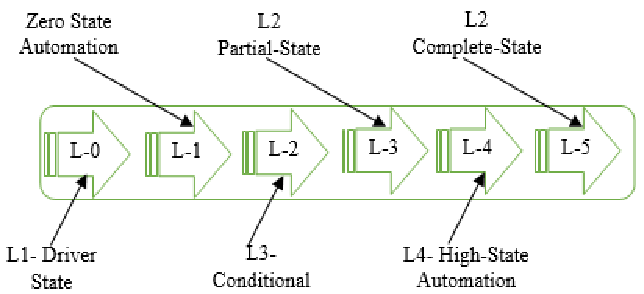
Retrieval Number: 100.1/ijaent.I048411090924
DOI: 10.35940/ijaent.I0484.11121224
Journal Website: www.ijaent.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

also initiated the AI core fed with test instances, conditions for boundary scenarios, and test examples to conduct simulation and virtual assessment of the environment (Harsha et al., 2017). (Goyal et al., 2020) analyzed a self-driven vehicle algorithm used to identify various images, through the application of the CNN algorithm to handle complicated calculators. (Ramakrishnan 2021) applied CNN with the traffic-based model to accelerate driving in dense traffic environments and to check the algorithm performance. In the simulation on the traffic-based model with highway and traffic hyperparameter tunning, the proposed model achieved 7.18% free of error for prediction.

A vision-based technique was employed in image processing to extract specific details and an optimal path for the vehicle to accelerate on the pathway through the environment. To generate the dataset, they simulate the environment, which consists of a car and track. With the generated data set, training was conducted to attain the optimal neural network architecture; through the simulator, tests were conducted and depicted via test results in an autonomous car on the road and track. The model performed a 23-lap round before moving in a different direction, but the neural network model helped maintain its track (Ribeiro et al., 2023). develops a 1/10 scale remote-controlled car (RC car) to represent an automated car. The model is made up of the following hardware and software components: CNN (convolutional neural network), monocular vision algorithm, Haar cascade classifier, Raspberry Pi Board model B+, Pi camera, Arduino, and an Ultrasonic sensor (Thadeshwar et al., 2020) [38]. (Naz et al., 2022) [30] introduced onboard sensors like laser, radar, lidar, Global Positioning System (GPS), and vehicular communication networks are used by autonomous cars to gather scene information. The acquired data is then applied to several path design and management strategies to enable the vehicles to operate independently in intricate surroundings. Modern AI algorithms are used by autonomous cars to locate themselves in both familiar and unfamiliar areas. Motion control, path planning, and perception are more applications of AI systems.

(Naz et al., 2022) Provided six levels of automation, ranging from zero-state automation to complete-state automation. This can be illustrated in Figure 1.



[Fig.1: Six Levels of Automation] [30]

In (Stanchev & Geske, 2016), A standardized classification system has been suggested by the National Highway Traffic Safety Administration (NHTSA):

Table 1: States of Autonomous Vehicles

Level	State
0	At all times, the driver has total control over the vehicle.
1	Certain car systems, such as automated braking or electronic stability control, are automated.
2	Automating at least two controls simultaneously, such as adaptive cruise control and lane maintenance is possible.
3	Under specific circumstances, the driver can completely relinquish control of all safety-critical operations. When the situation calls for the driver to regain control, the vehicle recognizes it and gives them a "sufficiently comfortable transition time" to do so.
4	The driver is not expected to operate the car at any point during the trip; it handles all safety-critical tasks. This car might incorporate empty cars since it would be in complete charge of all operations, including parking.

V. ARTIFICIAL INTELLIGENCE

The ultimate idea that will enable machines to carry out any kind of work independently and in any circumstance is artificial intelligence. Applications like machine learning (ML), which fall under the general category of artificial intelligence, are being utilized more and more to create solutions in several domains, including self-driving cars. Recently, methods like deep learning have gained popularity due to their exceptional performance on a range of tasks, including decision-making and image identification. (Mallozzi et al., 2019) [27].

Different machine learning algorithms must be applied to multiple car components for fully autonomous vehicles to function. Deep neural networks, or DNNs, have gained popularity recently as a primary method for handling several significant tasks about the perception component of autonomous driving. Examples of this include semantic segmentation of images, road detection, and item recognition and categorization. (Mallozzi et al., 2019). Because DNNs are strong techniques and can learn intricate representations from unprocessed data, they can do away with the need for hand-crafted features, which is challenging for other machine-learning techniques. Perception, prediction, and planning are the three primary responsibilities to be considered while creating an autonomous agent, or an entity that can make decisions based on its observations. You can use machine learning in each of them as shown in Table 1. (Sallab et al., 2017) [34].

Table 2: Three Main Actions for Creating AVs

Perception	Right now, deep learning is mostly used for perception [9]. It entails taking raw data from the vehicle's sensors and extracting semantic knowledge from it. It is present in autonomous vehicles in various contexts, including object tracking and detection, pedestrian identification, road sign recognition, position and mapping, and intelligent vehicle assistance functions like friction, destination, and traffic flow prediction (Chavez-Garcia & Aycard, 2015; Dollar et al., 2011) [14].

	Recurrent neural networks can be used to create models that forecast the future states of the environment (Pinheiro & Collobert, 2014) [32]. Additionally, using the current state as a basis, supervised learning can be utilized to forecast the near future (Shalev-Shwartz et al., 2016) [35]. From raw input signals, such as a pixel from the front-facing camera, convolutional neural networks can be utilized to directly forecast the steering angle of the car (Bojarski et al., 2016) [7].
Planning	A series of actions is planned by combining perception and prediction. For this reason, frameworks like reinforcement learning are useful. It works with recurrent neural networks as well as deep neural networks. An agent that uses reinforcement learning makes decisions about its surroundings to maximize a monetary reward (Sutton & Barto, 1999) [37]. These incentives are given to the agent by the environment in exchange for each action it gets. The agent monitors the incentives and applies them to enhance its policy for making decisions. Learning by demonstration is another model for teaching a system how to do tasks. In this paradigm, policies are taught by an instructor (Argall et al., 2009) [3].

VI. DEEP LEARNING ALGORITHM

Deep learning, also referred to as deep machine learning, is a subfield of machine learning that is partially based on structured algorithms. Its goal is to use multiple layers of linear and non-linear transformation functions, graphs, multiples of deep graphs, and multiples of deep processing layers to create a high-level abstraction model from the given data. (Vishnukumar et al., 2017) [40].

Table 3: Some Deep Learning Algorithm

References	Algorithm	Approached
(Ahmed et al., 2020) [1].	K-Means	Is an unsupervised technique that divides datasets without labels or classifications into pre-established clusters. It assigns a centroid to each cluster and iteratively seeks to minimize the total sum of the distances between the centroids of the clusters and their respective data points.
(Miglani & Kumar, 2019) [28]	RCNN	Object detection is the application of this kind of neural network. It is favoured over traditional CNN since it uses less space than CNN when detecting an object in an image because the object does not appear the same number of times. To locate the borders and labels of each object in an image and build a boundary box around it, RCNN employs a selective searching technique. To determine the bounding box's precise coordinates, a linear regression model is ultimately applied. RCNN is utilized in AVs to recognize objects, pedestrians, and traffic signs
(Li et al., 2020) [23]	YOLO	Joseph Redmon created this CNN-based algorithm within the Darknet framework, and it is an effective real-time object detection tool. After that, YOLO was the basis for a series of object detectors in computer vision, including YOLOv2, YOLOv3, and YOLOv4. Identifiers and classes of objects are handled by this single-stage detector in a single network pass. Models based on YOLO are effective and simple to implement.
(Dai & Lee, 2020)	SLAM	Based on data from LiDAR or RADAR sensors, SLAM is primarily used to estimate the relative position of static objects in an environment. By serving as an odometer, RADAR-SLAM can offer velocity information and even carry out localization using information from maps. Using a monocular or stereo camera, LiDAR-SLAM, also known as Visual SLAM, tracks the features of successive images while calculating their relative direction and translation. AVs employ SLAM for Path Planning, Motion Control, and occasionally even Pedestrian Detection [24].
(Miglani & Kumar, 2019)	CNN	Because this kind of neural network is so good at using the convolution function to extract distinguishing information from images, it is utilized for image processing. It uses several hidden layers to extract high-level features from a 2D input. Following input, it uses the spatial arrangement of the input pixels to find patterns in the images that are valuable. CNN can be easily implemented since pre-processing is not necessary. They are employed in AVs for pedestrian recognition and path planning.

Table 4: Performance of Some Deep Learning Algorithm

References	Proposed Algorithm	Objects Detected	Approach	Performance/Results
Goyal et al 2020	YOLO	Vehicle, traffic light, lane, people, traffic signs	The approach used a Convolutional neural model to distinguish highlights though the related neural layers provide navigations and likelihood. The proposed system used 24 layers in progress with an additional 2 related layers.	The Yolo algorithm performed better and faster in real-time instances for detecting objects.
(Ramakrishnan 2021)	CNN		The traffic-based model was developed through the Tensor Flow framework with Python programming language for simulation. Various parameters were installed for highway traffic.	
(Nguyen et al., 2021) [31]	Video Surveillance Processing Algorithm	Moving Vehicles	The model developed a background frame from initial, enhanced to finish, and updated the server. The area of interest from the object in motion is split in the environment and transmits the changes only.	The model was designed to improve the reception and passing of data in UAV CSS based on AI with 80% memory space and data transmission.

VII. APPROACH FOR DETECTING LANES

The main development here is to review various algorithms applied for lane detection through lane images. (Goyal et al., 2020) developed the model to predict the gradient change through colour intensity amid adjacent images (pixels) and further change images' colour to grayscale. Thereafter, images were smoothed

Retrieval Number: 100.1/ijaent.I048411090924
DOI: 10.35940/ijaent.I0484.11121224
Journal Website: www.ijaent.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

through the Gaussian model, noise was reduced to enhanced detection for different edges with edge detection technique. Lastly, the images are masked in triangular form which forms the region of interest.

VIII. DATASET

At the moment, there are a lot of datasets accessible for autonomous car segmentation tasks; some of them are introduced for instance segmentation, as discussed in Table 3, while others are connected to semantic segmentation. (Stanchev & Geske, 2016) [36].

Table 5: Description of Some Datasets used in AVs

Reference	Dataset	Description
(Krajewski et al., 2018) [22]	High-D dataset	Approximately 110,000 corrected trajectories of various vehicles, including cars and trucks, are included in the collection. These paths are taken from drone footage that was shot above German motorways at 25 frames per second and with a resolution of 4096 x 2160 pixels. The dataset contains the trajectory ID, speed, acceleration, longitudinal position, distance to the leader, and ID of the current leader for each specific vehicle trajectory.
(Cordts et al., 2016) [12]	The city-Space	the premium pixel-level semantic segmentation dataset, gathered in about 50 German cities and the surrounding nations, to comprehend urban street scenes. The collection comprises five thousand 1024 × 2048 pixel-level annotated photos that show intricate metropolitan settings under various weather situations, with variable backdrops and scene arrangements.
(Caesar et al., 2020) [8]	NuScenes	is a massive 3D object detection dataset that was just released to aid in AD scene comprehension. Moving cars fitted with a variety of specialized sensors are used to gather the dataset in Singapore's Holland Village, Queenstown, and One North as well as Boston's South Boston and Seaport. Thirteen sensors make up the car-mounted suite: one IMU sensor, five long-range radar sensors running at 77 GHz with a 13 Hz capture frequency, six RGB cameras with 1600 ×900 resolution and a 20 Hz capture frequency, and one LiDAR sensor.
(Geiger et al., 2012) [16]	Semantic-KITTI	is a large outdoor scene dataset designed for panoptic and point cloud semantic segmentation of road scenery, encompassing freeways, residential areas, and city traffic? It consists of 43552 point-by-point re-annotated 3D scans for the KITTI Vision Odometry Benchmark dataset that were produced using an automotive LiDAR sensor. There are 22 unique sequences in this dataset, which are divided between test and training-validation subsets. Twenty-three,511,3D scans from sequences 0 to 10 make up the training-validation set, while twenty,351,3D scans from sequences 11 to 21 make up the test set.
(Geiger et al., 2012)	KITTI	With a size of 1240x376 pixels, the dataset consists of 14999 RGB stereo image pairs (which include the picture and its accompanying ground truth). A training set (7841 samples) and a test set (7518 samples) comprise the full dataset. The training set is then divided into two subsets: the test set (3769 samples), which is mostly utilized for validation, and the training subset (3712 samples).
(Lin et al., 2014) [25]	COCO	One of the most popular databases made available by Microsoft is the Common Objects in Context (COCO) dataset, which is extensively utilized for object captioning, semantic and object instance segmentation, and object recognition. The dataset includes 330,000 photos with over 200,000 labelled instances, human pose estimates, 250,000 individuals with key points, and 1,500,000 object instances divided into 80 different types. The image data is gathered from several sources, such as Flickr users who have contributed amateur photos with searchable tags and pertinent item images from the PASCAL VOC collection [17].
(Everingham et al., 2010) [15]	PASCAL VOC	One of the hardest datasets available to the public is PASCAL VOC (Visual Object Classes), which is used for object recognition, segmentation, and image classification. The VOC dataset is divided into two portions, VOC 2007 and VOC 2012, much like the COCO dataset. 9,962 images and their accompanying annotations altogether, divided into three subsets: 2501 training, 2510 validation, and 4951 testing images, are included in the VOC 2007 release. 22,531 photos total from the VOC 2012 release are split up into three subsets for training, validation, and testing, namely 5,717, 5,823, and 10,991 images.

IX. TECHNIQUES OF DATA EXTRACTION FROM ACADEMIC SOURCES

Several scholarly datasets on autonomous vehicles through AI and deep learning approaches were obtained based on relevant and sufficient papers. The review papers underwent sufficient scrutiny using basic investigations and other diverse methodologies. As seen in Tbale3, the research methodology used in the review paper searches through a variety of academic databases for pertinent works, including Springer, IEEE Explore, Scopus, Web of Science, Science Direct, Google Scholar, and the ACM Digital Library.

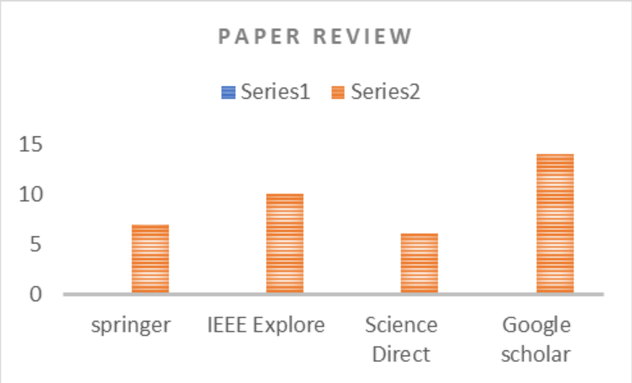
X. TERMINOLOGY FOR SEARCHES

One of the most efficient search keywords was established by carefully and methodically choosing the core search terms. A significant research paper was located using the following keywords in a reliable academic database: Autonomous Vehicle, AI and Deep Learning, vision-based, and Autonomous Vehicle using AI and Deep Learning.

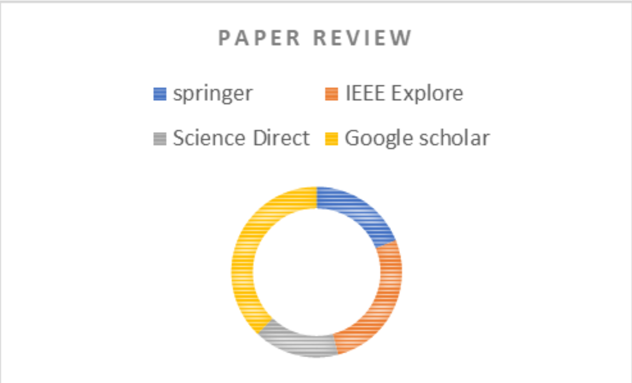
Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Table 6: Summary of Search Terms

S/N	Sources of Academic Papers	Domain Site	No. of Review Papers
1.	Springer	http://www.springer.com/	7
2.	IEEE Explore	http://www.ieeexplore.ieee.org/	14
3.	Science Direct	http://www.sciencedirect.com/	6
4.	Google Scholar	http://www.scholars.google.com/	14
		Total of Review Papers	**41**



[Fig.2: Histogram of Papers Review]



[Fig.3: Doughnut of Papers Review]

XI. FUTURE DIRECTION

There has been a noticeable increase in research interest recently on the prediction of pedestrians and other moving objects—like cars—future whereabouts (Hu et al., 2020) [20]. Considerable effort has been paid to semantic segmentation utilizing frame-based visual data, and significant advancements have been made in the previous two years [42]. While there are a lot of robust techniques available for frame-level segmentation, their computational efficiency is compromised because they are primarily focused on reaching higher accuracy levels [43].

Thus far, segmentation techniques have been the main method used to approach scene understanding. However, attention can be shifted to more advanced forms of vehicular cognition, like events-based scene interpretation. (Zhang et al., 2021) [41]. One intriguing approach, for example, is to analyze the events for scene parsing [44]. This way, the surrounding occurrences—like a bike on the vehicle lane or a person crossing the street—can better assist and encourage autonomous vehicles to make more educated decisions. (Chen et al., 2020) [10]. The key takeaway from this is that to comprehend a scene fully, one needs to investigate other metrics and ascertain the links among recognised items in both space and time, (Ullah et al., 2020) [39].

XII. CHALLENGES FACED BY AUTONOMOUS VEHICLES

Since the real world is filled with a plethora of unexpected events and situations due to our departure from the conventional notion of an automobile, automated driving represents a significant paradigm shift from conventional vehicles. As a result, it is no longer a request but rather a fundamental necessity in the field of testing and validation (Vishnukumar et al., 2017). It is already widely acknowledged that manufacturing autonomous cars in large quantities is not the only important step; testing and validation of the vehicles are also necessary to ensure that they perform as intended and meet safety regulations (Vishnukumar et al., 2017).

Because there are more lines of code in software every day, updates are taking too long. To address this issue, an over-the-air approach with High processing capacity and speed is necessary for processing sensor devices. GPUs, CPUs, and FPGA are necessary for high-end computing devices. (Ma et al., 2020) [26]. AI requires computation that is beyond the capabilities of traditional CPUs. As a result, many researchers create AVs using GPUs. The drawback of GPUs is that they require ten times as much electricity as FPGA. Google created TPU and performs 15–20 times better than GPUs (Jouppi et al., 2017) [21].

(Bathla et al., 2022) [5] Provided some challenges:

- The accuracy of the sensor data used as input signals determines how well AI systems operate. Sensor problems influence AI approaches' input.
- Cars can communicate with other vehicles, roadside equipment, and other vehicles thanks to vehicle-to-everything (V2X) technology. For those in academics and business, privacy protection and safe communication between parties remain top priorities in AV.
- Because there are more lines of code in software every day, updates are taking too long. To address this issue, an over-the-air approach was implemented; however, several attacks are noted when software is being updated.
- Cars can communicate with other vehicles, roadside equipment, and other vehicles thanks to vehicle-to-everything (V2X) technology. For those in academics and business, privacy protection and safe communication between parties remain top priorities in AV.
- Because there are more lines of code in software every day, updates are taking too long. To address this issue, an over-the-air approach was implemented; however, several attacks are noted when software is being updated.
- To reduce perception errors, it will be challenging to precisely locate, classify, and detect things in the outside environment. Error perception is one of the

Retrieval Number: 100.1/ijaent.I048411090924
DOI: 10.35940/ijaent.I0484.11121224
Journal Website: www.ijaent.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

difficult problems in AV safety.

- An accurate, reliable, and efficient decision-making system ought to be built with the ability to react appropriately and quickly to the external environment. Thorough and rigorous software and hardware system testing should be carried out to lower the decision error.
- Examine the system's behaviour in a variety of scenarios and settings to prevent failure.
- Researchers' top issue is cybersecurity for AV. How safe it is to use wireless communication. The public's perception of the developing AV technology might be greatly influenced by serious worries about security and safety.

XIII. CONCLUSION

An extensive examination of artificial intelligence and deep learning in autonomous vehicles via vision-based systems, and driving experiences has been carried out by this work. According to observations, safety regulations for autonomous systems are lacking, and artificial intelligence (AI) is a key consideration when creating vision-based guidelines for futuristic autonomous systems. Additionally, a comparative review of numerous research on autonomous systems reveals that the development of autonomous systems requires the integration of two or more cutting-edge technologies, including artificial intelligence, blockchain, Internet of Things, cloud computing, fog computing, and edge computing. Here, the attention is focused on how artificial intelligence and deep learning track the actions and travels of the car with computer vision.

XIV. ACKNOWLEDGMENT

I would also like to express my gratitude to the members of the research committee members for their constructive criticism and valuable feedback, which contributed to enhancing the quality of this research.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been sponsored or funded by any organization or agency. The independence of this research is a crucial factor in affirming its impartiality, as it has been conducted without any external sway.
- **Ethical Approval and Consent to Participate:** The data provided in this article is exempt from the requirement for ethical approval or participant consent.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Authors Contributions:** The authorship of this article is contributed equally to all participating individuals.

REFERENCES

1. Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. Electronics, 9(8), 1295. Doi: https://doi.org/10.3390/electronics9081295
2. Al-Kaff, A., Martin, D., Garcia, F., de la Escalera, A., & Armingol, J. M. (2018). Survey of computer vision algorithms and applications for unmanned aerial vehicles. Expert Systems with Applications, 92, 447-463. DOI: https://doi.org/10.1016/j.eswa.2017.09.033
3. Argall, B. D., Chernova, S., Veloso, M., & Browning, B. (2009). A survey of robot learning from demonstration. Robotics and autonomous systems, 57(5), 469-483. DOI: https://doi.org/10.1016/j.robot.2008.10.024
4. Bagloee, S. A., Tavana, M., Asadi, M., & Oliver, T. (2016). Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. Journal of modern transportation, 24, 284-303. DOI: https://doi.org/10.1007/s40534-016-0117-3
5. Bathla, G., Bhadane, K., Singh, R. K., Kumar, R., Aluvalu, R., Krishnamurthi, R., Kumar, A., Thakur, R., & Basheer, S. (2022). Autonomous vehicles and intelligent automation: Applications, challenges, and opportunities. Mobile Information Systems, 2022(1), 7632892. DOI: https://doi.org/10.1155/2022/7632892
6. Bilik, I., Longman, O., Villeval, S., & Tabrikian, J. (2019). The rise of radar for autonomous vehicles: Signal processing solutions and future research directions. IEEE Signal Processing Magazine, 36(5), 20-31. DOI: https://doi.org/10.1109/MSP.2019.2926573
7. Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L. D., Monfort, M., Muller, U., & Zhang, J. (2016). End-to-end learning for self-driving cars. arXiv preprint arXiv:1604.07316. DOI: https://doi.org/10.48550/arXiv.1604.07316
8. Caesar, H., Bankiti, V., Lang, A. H., Vora, S., Liong, V. E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., & Beijbom, O. (2020). nuscenes: A multimodal dataset for autonomous driving. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, Doi: https://doi.org/10.1109/CVPR42600.2020.01164
9. Chavez-Garcia, R. O., & Aycard, O. (2015). Multiple sensor fusion and classification for moving object detection and tracking. IEEE Transactions on Intelligent Transportation Systems, 17(2), 525-534. DOI: https://doi.org/10.1109/TITS.2015.2479925
10. Chen, G., Cao, H., Conradt, J., Tang, H., Rohrbein, F., & Knoll, A. (2020). Event-based neuromorphic vision for autonomous driving: A paradigm shift for bio-inspired visual sensing and perception. IEEE Signal Processing Magazine, 37(4), 34-49. DOI: https://doi.org/10.1109/MSP.2020.2985815
11. Cheng, J., Zhang, L., Chen, Q., Hu, X., & Cai, J. (2022). A review of visual SLAM methods for autonomous driving vehicles. Engineering Applications of Artificial Intelligence, 114, 104992. DOI: https://doi.org/10.1016/j.engappai.2022.104992
12. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., & Schiele, B. (2016). The cityscapes dataset for semantic urban scene understanding. Proceedings of the IEEE conference on computer vision and pattern recognition, DOI: https://doi.org/10.1109/CVPR.2016.350
13. Dai, Y., & Lee, S.-G. (2020). Perception, planning, and control for self-driving systems based on onboard sensors. Advances in Mechanical Engineering, 12(9), 1687814020956494. DOI: https://doi.org/10.1177/1687814020956494
14. Dollar, P., Wojek, C., Schiele, B., & Perona, P. (2011). Pedestrian detection: An evaluation of the state of the art. IEEE transactions on pattern analysis and machine intelligence, 34(4), 743-761. DOI: https://doi.org/10.1109/TPAMI.2011.155
15. Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2010). The Pascal visual object classes (VOC) challenge. International journal of computer vision, 88, 303-338. DOI: https://doi.org/10.1007/s11263-009-0275-4
16. Geiger, A., Lenz, P., & Urtasun, R. (2012). Are we ready for autonomous driving? The Kitti vision benchmark suite. 2012 IEEE Conference on Computer Vision and Pattern Recognition, DOI: https://doi.org/10.1109/CVPR.2012.6248074
17. Harding, J., Powell, G., Yoon, R., Fikentscher, J., Doyle, C., Sade, D., Lukuc, M., Simons, J., & Wang, J. (2014). Vehicle-to-vehicle communications: the readiness of V2V technology for application. https://www.automotivesafetycouncil.org/wp-content/uploads/2017/01/Readiness-of-V2V-Technology.pdf
18. Hashimoto, D. A., Rosman, G., Rus, D., & Meireles, O. R. (2018). Artificial intelligence in surgery: promises and perils. Annals of Surgery, 268(1), 70-76. DOI:

https://doi.org/10.1097/SLA.0000000000002693
19. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, DOI: https://doi.org/10.1109/CVPR.2016.90
20. Hu, A., Cotter, F., Mohan, N., Gurau, C., & Kendall, A. (2020). Probabilistic future prediction for video scene understanding. Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVI 16, DOI: https://doi.org/10.48550/arXiv.2003.06409
21. Jouppi, N. P., Young, C., Patil, N., Patterson, D., Agrawal, G., Bajwa, R., Bates, S., Bhatia, S., Boden, N., & Borchers, A. (2017). In-datacenter performance analysis of a tensor processing unit. Proceedings of the 44th Annual International Symposium on Computer Architecture, DOI: https://doi.org/10.1145/3079856.3080246
22. Krajewski, R., Bock, J., Kloeker, L., & Eckstein, L. (2018). The high dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems. 2018 21st International Conference on Intelligent Transportation Systems (ITSC), DOI: https://doi.org/10.1109/ITSC.2018.8569552
23. Li, C., Wang, R., Li, J., & Fei, L. (2020). Face detection based on YOLOv3. Recent Trends in Intelligent Computing, Communication and Devices: Proceedings of ICCD 2018, DOI: https://doi.org/10.1007/978-981-13-9406-5_34
24. Liden, D. (2013). What Is a Driverless Car? WiseGeek. Retrieved, 11. https://www.scribd.com/document/324783105/Driverless-Car-Ieee-Format
25. Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, Doi: https://doi.org/10.48550/arXiv.1405.0312
26. Ma, Y., Wang, Z., Yang, H., & Yang, L. (2020). Artificial intelligence applications in the development of autonomous vehicles: A survey. IEEE/CAA Journal of Automatica Sinica, 7(2), 315-329. DOI: https://doi.org/10.1109/JAS.2020.1003021
27. Mallozzi, P., Pelliccione, P., Knauss, A., Berger, C., & Mohammadiha, N. (2019). Autonomous vehicles: state of the art, future trends, and challenges. Automotive systems and software engineering: State of the art and future trends, 347-367. DOI: https://doi.org/10.1007/978-3-030-12157-0_16
28. Miglani, A., & Kumar, N. (2019). Deep learning models for traffic flow prediction in autonomous vehicles: A review, solutions, and challenges. Vehicular Communications, 20, 100184. DOI: https://doi.org/10.1016/j.vehcom.2019.100184
29. Miura, S., & Kamijo, S. (2015). GPS error correction by multipath adaptation. International Journal of Intelligent Transportation Systems Research, 13, 1-8. DOI: https://doi.org/10.1007/s13177-013-0073-9
30. Naz, N., Ehsan, M. K., Amirzada, M. R., Ali, M. Y., & Qureshi, M. A. (2022). Intelligence of autonomous vehicles: A concise revisit. Journal of Sensors, 2022(1), 2690164. https://doi.org/10.1155/2022/2690164
31. Nguyen, M. T., Truong, L. H., & Le, T. T. (2021). Video surveillance processing algorithms utilizing artificial intelligence (AI) for unmanned autonomous vehicles (UAVs). MethodsX, 8, 101472. DOI: https://doi.org/10.1016/j.mex.2021.101472
32. Pinheiro, P., & Collobert, R. (2014). Recurrent convolutional neural networks for scene labelling. International conference on machine learning, Doi: https://doi.org/10.48550/arXiv.1306.2795
33. Ribeiro, I. A., Ribeiro, T., Lopes, G., & Ribeiro, A. F. (2023). End-to-end approach for autonomous driving: a supervised learning method using computer vision algorithms for dataset creation. Algorithms, 16(9), 411. DOI: https://doi.org/10.5220/0007575908330839
34. Sallab, A. E., Abdou, M., Perot, E., & Yogamani, S. (2017). Deep reinforcement learning framework for autonomous driving. arXiv preprint arXiv:1704.02532. DOI: https://doi.org/10.2352/ISSN.2470-1173.2017.19.AVM-023
35. Shalev-Shwartz, S., Ben-Zrihem, N., Cohen, A., & Shashua, A. (2016). Long-term planning by short-term prediction. arXiv preprint arXiv:1602.01580. Doi: https://doi.org/10.48550/arXiv.1602.01580.
36. Stanchev, P., & Geske, J. (2016). Autonomous cars. History. State of art. Research problems. Distributed Computer and Communication Networks: 18th International Conference, DCCN 2015, Moscow, Russia, October 19-22, 2015, Revised Selected Papers 18, DOI: https://doi.org/10.1007/978-3-319-30843-2_1
37. Sutton, R. S., & Barto, A. G. (1999). Reinforcement learning: An introduction. Robotica, 17(2), 229-235. DOI: https://doi.org/10.1016/S0893-6080(99)00098-2
38. Thadeshwar, H., Shah, V., Jain, M., Chaudhari, R., & Badgujar, V. (2020). Artificial intelligence-based self-driving car. 2020 4th International Conference on Computer, Communication, and Signal Processing (ISP), DOI: https://doi.org/10.1109/ICCCSP49186.2020.9315223
39. Ullah, A., Muhammad, K., Hussain, T., Lee, M., & Baik, S. W. (2020). Deep LSTM-based sequence learning approaches for action and activity recognition. In Deep Learning in Computer Vision (pp. 127-150). CRC Press. DOI: https://doi.org/10.1201/9781351003827-5
40. Vishnukumar, H. J., Butting, B., Müller, C., & Sax, E. (2017). Machine learning and deep neural network—Artificial intelligence core for lab and real-world test and validation for ADAS and autonomous vehicles: AI for efficient and quality test and validation. 2017 Intelligent Systems Conference (IntelliSys), DOI: https://doi.org/10.1109/IntelliSys.2017.8324372
41. Zhang, J., Yang, K., & Stiefelhagen, R. (2021). ISSAFE: Improving semantic segmentation in accidents by fusing event-based data. 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), DOI: https://doi.org/10.1109/IROS51168.2021.9636109
42. Shaik, I., Nittela, S. S., Hiwarkar, Dr. T., & Nalla, Dr. S. (2019). K-means Clustering Algorithm Based on E-Commerce Big Data. In International Journal of Innovative Technology and Exploring Engineering (Vol. 8, Issue 11, pp. 1910–1914). Doi: https://doi.org/10.35940/ijitee.K2121.0981119
43. Verma, Dr. P. K., & Dr. Preety. (2020). Application of K-Means Algorithm to Mapping Poverty Outline by Province in India. In International Journal of Recent Technology and Engineering (IJRTE) (Vol. 8, Issue 6, pp. 1045–1049). Doi: https://doi.org/10.35940/ijrte.F7357.038620
44. Maheswari, Dr. K. (2019). Finding Best Possible Number of Clusters using K-Means Algorithm. In International Journal of Engineering and Advanced Technology (Vol. 9, Issue 1s4, pp. 533–538). Doi: https://doi.org/10.35940/ijeat.A1119.1291S419

AUTHORS PROFILE

Fidelis Nfwan Gonten, Msc., PhD in View (Computer Science, With a strong focus on prolific research (in AI) and global learner development, I am deeply committed to contributing to the research community and advancing the overall body of knowledge through a systematic approach. My expertise lies in building vision-based systems using advanced deep-learning algorithms. Over the past 13 years, I have been actively engaged in the teaching and learning process. Through extensive and rigorous studies, I am well-prepared and equipped to establish and uphold computing standards at a high level.

Professor Abdulsalam Ya'u Gital, Department of Computer Science, Abubakar Tafawa Balewa University, Bauchi Nigeria Verified email at atbu.edu.ng an expert on Modelling and Simulation, Cloud Computing, Collaborative Virtual Environment, Computer Communication and with_1219 citations. Currently a dean faculty of computing

Mr. Datti Useni, Msc and PhD. In view. I held from Gung village in Pankshin local government, Plateau State of Nigeria. I was born forty-seven (47) years old, married to Mrs. Datti Nanpan Plang, and blessed with three (3) kids. I reflect a varied personality including ambition and qualities of generosity and thoughtfulness. I am also a well-determined and vigorous individual, yet pleasantly calm. I encourage fighting for what you desire and believe in and doing it through God because nothing great comes easy and with God everything is possible.

Mr. Larson Suwa is a dedicated and passionate individual. Bagged BTech and Msc. in Computer Science from Abubakar Tafawa Balewa University Bauchi, Bauchi Nigeria. A seasoned researcher and dedicated and experienced lecturer with Global College of Education Jos, Plateau state Nigeria. Strong foundation in technology and a drive for continuous learning, committed to expanding my knowledge through intensive education pursuit.

Mr. Mudimka Jahota Yerima is A dedicated and passionate individual. Bagged National Diploma in Computer Science from Plateau State Polytechnic Barkin Ladi, Plateau State, Nigeria, Bauchi Nigeria currently a student with the same institution studying Statistical science.