Multimedia Tools and Applications (2023) 82:25107–25130
https://doi.org/10.1007/s11042-023-14417-x

# Fully convolutional neural networks for LIDAR–camera fusion for pedestrian detection in autonomous vehicle

J Alfred Daniel[1] · C Chandru Vignesh[2] · Bala Anand Muthu[3] · R Senthil Kumar[4] · CB Sivaparthipan[3] · Carlos Enrique Montenegro Marin[5]

Received: 8 August 2022 / Revised: 24 September 2022 / Accepted: 21 January 2023 /
Published online: 1 February 2023


**Abstract**
Pedestrian detection appears to be an integral part of a vast array of vision-based technologies, ranging from item recognition and monitoring via surveillance cameras to, more recently, driverless cars or autonomous vehicles. Moreover, due to the rapid development of Convolutional Neural Networks (CNN) for object identification, pedestrian detection has reached a very high level of performance in dataset training and evaluation environment in autonomous vehicles. In order to attain object identification and pedestrian detection, a sensor fusion mechanism named Fully Convolutional Neural networks for LIDAR–camera fusion is proposed, which combines Lidar data with multiple camera images to provide an optimal solution for pedestrian detection. The system model proposes a separate algorithm for image fusion in pedestrian detection. Further, architecture and framework are designed for Fully Convolutional Neural networks for LIDAR–camera fusion for Pedestrian detection. In addition, a fully functional algorithm for Pedestrians detection and identification is proposed to precisely locate the pedestrian in the range of 10 to 30 m. Finally, the proposed model's performance is evaluated based on multiple parameters such as Precision, Sensitivity, Accuracy, etc.; hence the proposed system model has proven to be effective comparatively.

**Keywords** Convolutional neural networks · LIDAR–camera · Fusion for pedestrian · Detection in autonomous · Vehicle

## 1 Introduction

Pedestrian detection is the foundation for behavior assessment and target acquisition in various industries, from self-driving automobiles to surveillance cameras. Pedestrian detection seems to be a critical activity that must be completed precisely and consistently in terms of enhancing levels of automation. A pedestrian detection system employs cutting-edge computer sensors to

✉ C Chandru Vignesh
profchandruvignesh@gmail.com

Extended author information available on the last page of the article

find pedestrians near the vehicle [1]. With this knowledge, the car's onboard computer may warn the driver if any people or other moving objects like cyclists are present close to the path of the car supposed to follow. However certain autonomous vehicles were on the industry, subsequent crashes involving automated vehicles operated by their autonomous car emphasize the critical need for additional research and evaluation [15]. Pedestrian detection is crucial in various real-world scenarios, including self-driving cars, robotics routing, and surveillance footage. Self-driving vehicles are equipped with sensors, cameras, and radars to assist them navigate to their destination and understand their surroundings [16]. The three main sensors that self-driving cars rely on perform together like the eyes and brain of a human. They enable the vehicle to see its surroundings clearly as a whole. They assist the automobile in determining the position, motion, and three-dimensional (3D) forms of nearby objects. Self-driving cars should be capable of detecting pedestrians in urban environments. If the vehicle is heading to crash with a pedestrian, the appropriate mechanisms must be engaged to safeguard the pedestrian; similarly, if the car seems due to collision with some other car or perhaps a barrier, the appropriate mechanisms should indeed be engaged to safeguard the commuters. Different sensor technologies, including Light Detection and Ranging (LiDAR), Millimeter-Wave Radar (MWR), cameras, and sensor fusion, can be used to develop pedestrian detection systems. This system's primary parts are acceleration sensors, a control unit, and an actuator mechanism. The acceleration sensors are located in the vehicle's front bumper and it notices movement when a pedestrian goes up against the front bumper. Furthermore, current sensors such as radar used in adaptive cruise control cannot discriminate between identified obstructions. Detecting pedestrians using computer vision is a critical but complex concept for several automation systems [5]. Moreover, this is one of the most frequently explored issues in computer vision that have seen significant development over the past. As with many other computer vision challenges over the past few years, deep learning-based approaches have permitted tremendous improvement in identifying pedestrians.

Pedestrian detection seems to be a critical component of a wide variety of vision-based technologies, spanning object detection and monitoring through surveillance footage and, most recently, automated vehicles. Autonomous cars rely on perception of their environment to enable safe and efficient driving performance. In order to properly identify nearby items including pedestrians, cars, traffic signs, and obstacles, this perception system employs object detection algorithms [4]. Also, with the fast advancement of Convolutional Neural Networks (CNN) in object recognition, pedestrian detection has attained extremely high efficiency inside the standard single-dataset training and assessment environment. There is a scarcity of heterogeneity within surroundings and situations and a comparatively limited number of people, limiting the efficacy of existing pedestrian trackers. Nearly all other current pedestrian detection approaches were indeed based on machine learning, and their efficiency seems to be extensively dependent on the amount of substance and reliability of data currently offered. Furthermore, there is an indication that perhaps the effectiveness of many other computer vision tasks such as image categorization, object recognition, and image fragmentation continue to improve even after billions of samples are accumulated. The input image is fragmented into segments, which are arbitrarily formed homogenous sections. The size of the fragments varies with the level of information in each part of the image, and the borders of the image serve as the boundaries of the fragments.

Most available datasets enabling pedestrian identification in autonomous vehicles would have at least three significant drawbacks. To begin, organizations have a finite number of discrete pedestrians. Furthermore, regions get a modest pedestrian density, which means that only a tiny fraction of pedestrians face geometric distortion. Numerous investigations have acquired increased early focus in multi-sensor information merging, and cohesion techniques consist of sensors including such camera, LiDAR, and RADAR in recent times while detection systems enabling environment perception accomplished. The complementing multimodal sensors, for instance, are critical for robotics and autonomous technology. Pedestrians' recognition, specifically in autonomous vehicles, should be rapid and precise enough to allow the management capacity to react and make necessary modifications. Nevertheless, contemporary related works outline an exact pedestrian recognition framework based on range and vision sensing equipment, albeit significant challenges remain unresolved from the perspective of an autonomous vehicle.

Based on the sensor modalities employed, numerous methodologies have been suggested for object proximity estimate. Every sensor conceptual framework is competent for experiencing the world from a unique viewpoint. It is restricted by the ability to detect particular item attribute information; furthermore, the primary restriction in interpreting numerous images within real-time and measurement and estimate inaccuracies. For instance, a significant shortcoming with existing object identification systems is the inability to discern the item's range. Most precisely, vision-based techniques seem to be more robust and reliable in detecting pedestrians however fall short of adequately measuring the pedestrians' distance.

In comparison, although LiDAR-based approaches are extraordinarily robust and precise in calculating the distance to an entity, they remain restricted in their capacity to classify the entity. LiDAR is a surveying technique that determines the distance to a target by projecting a pulsed laser light at the target and detecting the reflected pulses using a sensor. It is a well-known sensor that may offer 3D information about an item in the form of a point cloud, allowing for the identification of forms and the localization of objects.In contrast, just several ways are built on a vision-based camera's capacity to detect the surroundings in three dimensions. Thus, considering both the camera's and LiDAR sensor's capabilities and limitations, the sensor-fusion-based technique seems optimal for detecting pedestrians at such a range from the vehicle's present location. Sensor fusion has the capacity to combine data from many radars, lidars, and cameras to generate a single model or image of the world surrounding a vehicle. Because it balances the strengths of the many sensors, the final model is more accurate.

Pedestrian distance prediction employs an objective information fusion approach between a LiDAR sensor and a 3D camera to determine the object's distance, which would be required to allow an autonomous vehicle to go along a regular road by detecting its surroundings exceptionally precisely. Additionally, sensor fusion techniques were widely classified into several groups depending upon the information layers employed in fusion, such as low-level fusion, high-level fusion, and feature-level fusion [12]. Low-level data fusion integrates data from several sources to generate new data. The fusion of data is expected to be more insightful and synthetic than the original inputs. It combines raw sensor measurements depending on the sensor's physical state, while feature-level synthesis pulls any exceptional functionality from raw data across a sequence of data preprocessing methodologies. Feature-level fusion involves the integration of feature sets generated from several biometric sources into a single feature set using suitable feature normalization, transformation, and reduction strategies. During high-level fusion, every sensor conducts pedestrians' recognition or tracking technique autonomously while doing fusion. When compared to low-level fusion, high-level fusion is the different. Each sensor individually runs a tracking algorithm and generates an item list. The fusion model then links these things together and combines the sensor-independent objects at a high level, track-to-track. At the same time, each fusion strategy has well-established benefits and drawbacks compared to the others, and low-level fusion looks like the most appropriate option for autonomous cars due to its actual performance and much more accurate data fusion. Additionally, by fusing different types of sensors, it's indeed able to offer a wealth of data about an entity's kind, elevation, breadth, and proximity from its present location in the external environment.

Though sensor fusion technologies are being used across various industries, the problem of sensor harmonization has become more critical. Expressly, increasing advancement of fusion technologies incorporating 3Dcameras and LiDAR for autonomous cars necessitates an exact location coordinate of 3Dcamera and LiDAR sensor, incorporating attitude and orientation information. This may have been performed by solving extrinsic configuration constraints to determine the appropriate conversion matrix among sensor devices. Thus, a correct matching connection between specific sensors is essential to distinguish distinctive features inside the LiDAR point cloud and the spatial camera information. One of the most critical features that autonomous vehicle should possess is perception, enabling the autonomous car to perceive its surroundings and distinguish and categorize the pedestrians it observes. The automobile's ability to make appropriate judgments is contingent upon its ability to identify barriers immediately.

Consequently, the autonomous car must view and categorize traffic signals, people, street signs, sidewalks, parking areas, and highways, among other things [8]. Maybe not all, but it should also determine the correct distances between itself and the entities inside its direct proximity. Perception encompasses as much as detecting and categorizing; this helps the process to analyze the proximity and determine whether to slow down. The camera gives the vehicle perception, allowing it to perform many activities such as categorization, fragmentation, and localization. The cameras would have to feature high-resolution and adequately portray the landscape. The 3D cameras were combined to provide a 360-degree image of a surrounding area to ensure that the automobile gets video data across all directions. Those cameras have a wide field of view of up to 200 m and a close-range vision for further precise perception. Autonomous vehicles are equipped with sensors to do picture categorization, pedestrians' recognition, fragmentation, and positioning Fig. 1.

Moreover, Simultaneous localization and mapping (SLAM) techniques calculate the location and alignment of every item adjacent concerning the initial picture, assisting in the classification of roadways, people, and other entities. SLAM, or simultaneous localization and mapping, is a technique for mapping an environment while simultaneously locating a moving mobile robot from inside. SLAM is a technique and a component system that requires an odometry source and a global perception system. SLAM allows AR apps to detect 3D objects and sceneries, monitor the world in real time, and overlay digital interactive augmentations. Localization algorithms within autonomous vehicles determine the vehicle's location and inclination during navigation. The automobile may build predictive models about the pedestrians in its vicinity using various information representation techniques. A deep learning system may interpret various data throughout development, including cloud data points and images using RADARs and Lidars. A very similar concept, however, during interpretation, may assist the automobile in preparing for all conceivable maneuvers, including braking, stopping, decelerating, and lane changes, amongst many others. Moreover, deep learning's purpose for self-driving vehicles would be to effectively comprehend complicated tasks, locate themselves in their surroundings, increase perception, and trigger kinematics motions. This also assures traffic safety and uncomplicated commuting of the autonomous vehicle.

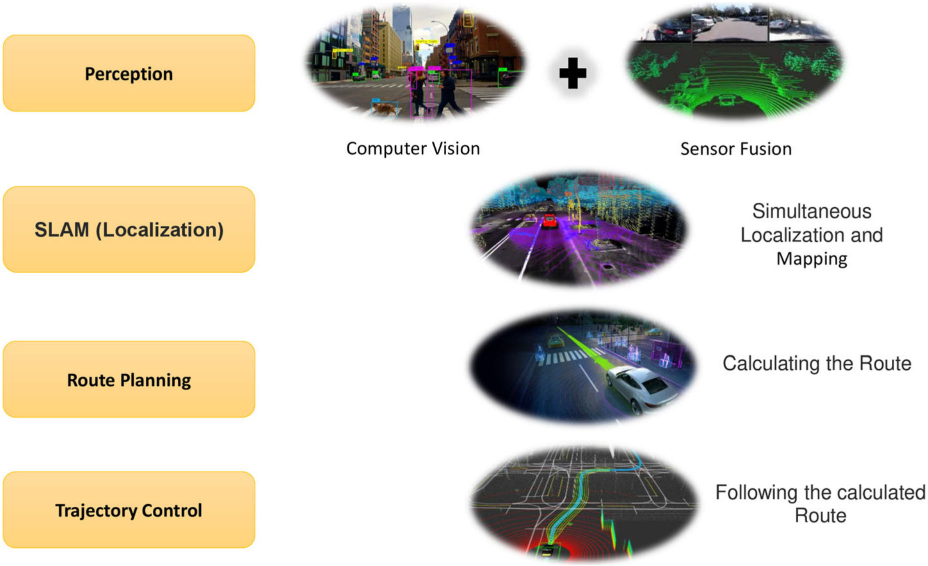


**Fig. 1** Process in Sensor fusion mechanism in Autonomous Vehicle

Therefore, to address the difficulties mentioned above, we propose an empirical model for Fully Convolutional Neural networks for LIDAR–camera fusion for Pedestrian detection in Autonomous vehicles.

Our contributions in this manuscript are threefold:

(i) first, we propose a competent architecture with a system model for fully convolutional neural networks for lidar–camera fusion for pedestrian detection in autonomous vehicles
(ii) in addition to that, functional algorithms have been designed to demonstrate the functionality of the sensor fusion mechanism for Pedestrians' detection and identification using Lidar sensor and 3D camera images.
(iii) furthermore, the performance analysis of the proposed model is evaluated and estimated based on multiple parameters.

## 2 Related works

Generalizable detectors demonstrate how poorly existing pedestrian detectors handle even minor domain alterations when evaluating datasets [10]. The manuscript attributes the article's investigation of restricted generality to two main characteristics: the technique used and the existing data sources. Concerning the approach, the research demonstrates which bias inherent in existing pedestrian detector concept decisions is the primary contributor to the restricted generality. Most contemporary pedestrian detectors were tuned to a specific dataset, achieving excellent efficiency in a specific training and validation workflow while degrading effectiveness if examined across datasets. As a result of its modular architecture, object classification detection outperforms state-of-the-art pedestrian detectors during cross-dataset evaluations. Concerning the information, researchers demonstrate that perhaps the autonomous vehicle metrics are generally repetitive, lacking in various situations, and congested with pedestrians.

As a result, they are benchmarking collected via web crawling that comprises a variety of different and complex situations and function as an essential resource of pre-training in creating a much more comprehensive representation. As a result, the article offers an approach for gradual fine-tuning that enhances generality. Furthermore, the manuscript examines current Transformer Systems using a backbone network to evaluate generalization.

By merging 3D camera images and LIDAR point clouds, the research team has established a deep learning technique for road recognition [2]. A spatial and temporal fragmented point cloud is envisioned over a 3D camera image and afterward adequately and accurately to provide a series of dense two-dimensional images storing spatial features. Numerous fully convolutional neural networks are then utilized to recognize roads, potentially employing input through a single sensor device via several fusing techniques: earlier, delayed, and the recently suggested crossover fusion. Meanwhile, the previous two fusion methodologies incorporate multidimensional data at a predetermined description level. The cross-fusion FCN was introduced to understand where and how to incorporate information directly from data; it is achieved through generative model cross-connections among both the LIDAR and camera computation sections [22].

Autonomous cars must obtain precise and real-time information from numerous sources in proximity, ensuring the commuter and vehicle's security in various circumstances [23]. 3D Lidar sensors would retrieve the location and feature extraction techniques about an item immediately inside their sensing range, while image cameras are best suited for object identification. Researchers thus provide a unique object recognition and identification research strategy combining the complementary information acquired through two different sensing devices. First, researchers use 3D LIDAR knowledge to generate object recognition recommendations. A convolutional neural network (CNN) is used to progressively recognize objects when these possibilities have been projected into the visual space and ROIs have been identified. Multi-scale extracted characteristics from ROIs are achieved by combining characteristics from the final three layers of both CNN.

Automated vehicles use their senses and perceptions of their surroundings [6]. Automated vehicles use a range of sensors, including LiDAR, radar, ultrasonic sensors, and 3d cameras, to monitor their surroundings. The heterogeneous data sources concurrently capture various physical properties of the surroundings. Such sensing heterogeneity and complexity must be favorably used for the trustworthy and reliable perception of the ecosystem using sensor data fusion. Furthermore, such multidimensional sensor data sources exhibit significant differences in temporal and geographical precision, data type, and geometric orientation. Future perception algorithms can take advantage of the variety provided through multisensory perception. The data streams should be geographically, spatially, and chronologically synchronized. The research investigation would analyze the difficulty of fusing both outcomes of such a LiDAR scanner with a wide-angle monochrome imaging system. Both output signals of the LiDAR scanner and, indeed, the photodetector have varying spatial resolutions and must be harmonized. After spatially aligning the two sensor outputs using a mathematical approach, using Gaussian Process regression-based resolution comparison technique is utilized to extrapolate the incomplete information with measurable ambiguity.

Vehicle detection is critical for driverless cars' surroundings awareness [18]. Although conventional vision-based vehicle identification systems were insufficiently precise, particularly with tiny and obstructed objects, lidar-based approaches were effective at identifying barriers however are time-consuming and provide a poor classification performance for different targeted classifications. To address inherent inadequacies and maximize the benefits of lidar's depth details and vision's capacity to classify obstacles, this study presents a real-time vehicle recognition system that combines 3D camera and lidar point cloud metadata. Initially, obstacles were spotted utilizing the grid projection approach and lidar point cloud data. Therefore, the barriers are mapped to the picture, yielding multiple distinct regions of interest (ROIs). Afterward, the ROIs were extended and combined depending on the decision threshold. Furthermore, a strategy known as the YOLO framework is used for the ROI to identify autonomous vehicles. Experiments upon that KITTI dataset reveal that perhaps the suggested technique does indeed have a high detection rate and a fast response period.

Researchers present a roadway identification approach in this task predicated upon merging lidar and picture information within the context of maximum entropy fields [9]. Investigators projected lidar point clouds onto monocular imagery using cross calibrating to obtain sparse altitude pictures, then used a combined bilateral filtration to obtain high-resolution altitude pictures. Moreover, for every pixel inside the learning imagery corresponding to a Lidar point, researchers extracted the characteristics from Lidar point clouds and 3d camera images and combined them with the position information to develop the Adaboost classification. Moreover, using a Boolean random matrix architecture, all experimental pixels were categorized as either road or non-road. Researchers employ the Adaboost classifier ratings as substring perspective and the elevation number and color features of every pixel as the bilateral perspective throughout this conditional random matrix architecture.

Avoiding collisions is a crucial problem in various application scenarios, including Advanced Driver Assistance Systems (ADAS), factory automation, and industrial automation. ADAS makes use of internal sensors, including radar and cameras, to sense the environment around it. Depending on what it discovers, it either informs the driver or reacts automatically. The term "warning" is widely used in the name of ADAS functions that offer information [19]. Specific regions must indeed remain off-limits to autonomous vehicles inside an automated production context to safeguard humans and high-valued commodities. Enclosed zones may be isolated through mapping, such as GPS, or primarily through signals indicating another no boundary. Researchers present a unique demarcation approach wherein a manufacturing vehicle detects active signals that use a LiDAR and a separate 3d camera and uses model-predictive automation to avoid triggering a limited territory. The beacons are regular orange traffic cones affixed to a luminous vertical pole. Although the LiDAR can detect these transmitters, it is susceptible to false positives caused by other reflectors, including worker protection outfits. This article proposes a technique for decreasing false positive tracking using LiDAR by portraying their signals within a camera picture that used a deep learning approach and confirming the identification that used a neural network-learned estimation mostly from a 3d camera to the LiDAR region.

Currently, computer vision enabling object recognition is becoming more popular and widespread [21] in a wide variety of sectors, including surveillance, transportation, and commuter movement analyses [3]. The study examines the investigation of combining lidar and image sensors for pedestrian detection to achieve extraordinarily high detection rates. To mitigate both the false-positive rate and the interference issue associated with camera-based pedestrian recognition, researchers examine the overall geometry of the item using a 3D point cloud returned by a Lidar depth sensor. Lidar is a method of calculating ranges that involves using a laser to target an objects or a surface and measuring the time it takes for the reflected light to return to the receiver. The hypothesized Lidar/camera sensor fusion architecture balances both advantages and disadvantages of the two detectors, resulting in a more reliable detection technique than many others. All sensors provide region proposals, and candidates from sensors are indeed sent to the secondary categorization for confirmation.

Throughout this investigation, researchers present a unique 3D object detection method to achieve exact localization using the other LIDAR and cameras. The lidar-based 3d object recognition problem is generally defined as follows: given a point cloud of a scene developed by returning lidar points represented in the lidar coordinate frame, predict oriented 3d bounding boxes represented in the lidar coordinate frame corresponding to target characters in the scene [13]. To accomplish this, researchers develop a robust end-to-end trainable framework that uses continual convolution layers to merge 3d imaging and LIDAR image features of varying resolutions. The suggested continuous fusion layer seems capable of encoding simultaneously discrete-state and continuous geometric information [14]. Doing so enables users to create an end-to-end capable of learning 3d image detection methods based on numerous unique, trustworthy, and affordable sensors. The continuous fusion layer encapsulates all discrete-state picture characteristics and continuous geometrical metadata. It further enables users to create a new, dependable, and economical end-to-end trainable three-dimensional object detector depending on multiple detectors.

Several LiDAR 3d camera-enabled object detection has already been constructed using two dense neural networks to retrieve view-specific characteristics, but a single dense neural network-based 3D detector has not even been deployed [20]. To address this problem, this work initially provides an early-fusion technique for rapidly detecting 3D objects using both LiDAR and image information using a single backhaul, establishing an appropriate mix of effectiveness and precision. Researchers offer a single pixel fusion component that extracts point-wise characteristics features from the raw RGB pictures and merges those with their associated point cloud without a framework. The foundation that extracts RGB picture characteristics is eliminated to alleviate the high computational cost associated with this approach. Our technique begins by voxelizing a point cloud into a three-dimensional voxel grid and then employs two ways to minimize data redundancy throughout the process. Correspondingly, the first technique involves using a tiny voxel size, whereas the second approach involves projecting point cloud features like intensity or height information onto RGB pictures.

Despite the excellent depiction capability of CNN characteristics, significant advancements in pedestrian detection have already been made. Furthermore, reducing false-positive results for hard negative instances like foliage, traffic signals, and poles is problematic. Applying high-level semantics visual signals may eliminate several of these inescapable negatives [17]. Researchers present a geographical area CNN method for the research that uses contextual information to improve pedestrian recognition. Our technique expands the architecture towards Faster R-CNN recognition by integrating another network branching for semantic picture categorization [7]. The semantic framework is built to calculate additional higher-level conceptual characteristics combined with convolutional characteristics. Researchers utilize multi-resolution image features generated from several network layers to identify pedestrians of various sizes accurately [11]. The enhanced forestry was utilized for training and incorporated characteristics in a cascaded fashion when extracting hard negatives. The Caltech pedestrian dataset's investigation demonstrates that the conceptual network improves detection capability. The rest of the article is organized as follows. We briefly review some related work in Section 2. In Section 3, we discuss the overview of the proposed model Proposed Fully Convolutional Neural networks for LIDAR–camera fusion for Pedestrian detection Framework. The experimental results and discussion are elaborated in Section 4. Finally, in Section 7, we have the conclusion.

## 3 Proposed fully convolutional neural networks for LIDAR–camera fusion for pedestrian detection framework

As the input sensors, the LiDAR and the 3D camera have been incorporated into the design that has been proposed. Utilizing two sensors to determine the pedestrian region recommendations for the first classification is supposed to be the result of the design that has been proposed. Furthermore, a proposal that includes the segmentation and features of the prospective contenders in the LiDAR-based method is presented. The prospects can also be determined using the camera-based method provided by the suggested method simultaneously. Additionally, we build the area of interest (ROI) for pedestrian recognition by projecting LiDAR's region proposals into the image plane. Then, users implement the recommended object-matching technique to establish the recommended candidates list, which signifies those particular prominent single objects that do not coincide with others will be assigned as both-detected and will not be sent to the following classification. The potential occluded object candidates will be put through a second classification process.

The proposed pedestrian system model generated by object matching will be put through the contrary LiDAR and camera sensors to complete the classification. Moreover, if an individual is lying just on the wall, segmentation would consider them as a non-pedestrian entity; however, it will be recognized by the cameras, and individuals who fall into this category will also be put through the subsequent classification Combining these two different types of sensors can boost the pace at which the target is detected. The second step in the classification process involves using a cropped image of a pedestrian to control whether a contender fits a pedestrian's profile. This primarily concentrates on the skipped target for each sensor and lowers the false-positive rate of a camera by analyzing with its 3D point cloud data inside a LiDAR-based technique.

Furthermore, the final result of pedestrian detection is given with high accuracy through LIDAR and Camera imagery fusion. The camera has been an excellent sensor that can produce region proposals, lane line positions, the state of traffic lights and signs, and a considerably better opportunity. One of the most important things to remember regarding a particular camera would be that it contains a 2D Sensor. LiDAR. A three-dimensional sensor gives off point clouds, each of which has an X, Y, and Z coordinate. Additionally, it is feasible to carry out various applications on 3d information, including executing machine learning techniques and neural networks.

Moreover, we have two possible scenarios in sensor fusion: early fusion with raw data and late fusion for the fusion of results. Fusing raw data is an integral part of early fusion. Consumers could, for instance, transfer the three-dimensional LiDAR point clouds onto the two-dimensional image. The LiDAR sensor data is shown as a dense three-dimensional point cloud. Pre-processing is required to cluster the point cloud, minimize the number of clusters, and refine the data before it can be utilized for classification. Besides that, consumers examine the 2D bounding boxes to see if the point clouds relate within them or otherwise. Furthermore, late fusion refers to the fusing of the data after independent detections have been made. Consumers could, for instance, transform the two-dimensional bounding boxes received from the camera into three-dimensional bounding boxes and then combine these with the two-dimensional bounding boxes produced from the LiDAR detection method. One of the most fundamental strategies for making attribute calculations easier for computer vision-based models is bounding box annotation. Autonomous cars are trained to recognize various items on the roadways, such as traffic, lights, lanes, potholes, and so on, using bounding box annotated images. In image processing projects, a 3D bounding box is a simulated rectangle that acts as a point of reference for object recognition and generates a collision box for that entity.

### 3.1 Image fusion

In the proposed model concerning route recognition, a critical activity must be completed precisely in terms of enhancing levels of autonomy in autonomous vehicles. For decision-making and secure trajectory planning, the crucial requirement is that the system knows about the road surface sections available for driving. Further, certain autonomous vehicles are currently on the market, and the subsequent accident on autonomous vehicles by its autopilot system articulated the critical need for additional study and validation. Autopilot navigates the area by utilizing a variety of sensors positioned throughout the vehicle. These systems use cameras, digital displays, radar/sonar, and other sensors to detect the surrounding area and maintain the car in its proper lane. When autopilot is enabled, the system may steer the vehicle, change lanes, modify speed, and regulate the brake function. The existing method uses either LIDAR sensors or a camera for route recognition. In the autonomous car, the camera offers dense information with a high frame rate under good lighting and reasonable climatic condition. On the other hand, the most common methods now utilized for spotting pedestrians include 2D cameras with LIDAR sensors. When the lighting is adequate, and the conditions are favorable, cameras can operate at a high frame rate and transmit a large amount of information across a considerable distance.

Calculations are made of the distances from the vehicle to every LIDAR spot. And checks every point in each linear sequence. As shown in Fig. 2, the points are considered part of the obstacle if their distance from one another is significantly greater than the mean distance throughout that sequence for pedestrian detection. The distance-focused approach determines the average intensity of the fused image inaccuracy by employing a quadratic equation in conjunction with RMSE, which stands for Root Mean Squared Error.

***Step 1:*** *The square root of the average of squared discrepancies between predicted and observed values.*

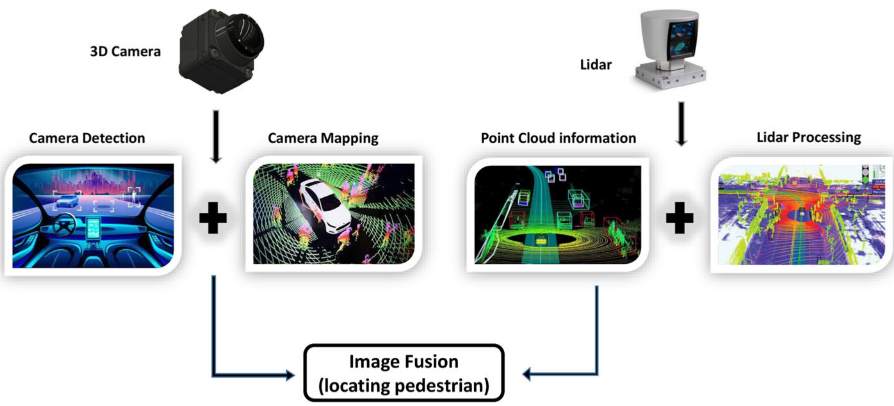


**Fig. 2** Proposed Fully Convolutional Neural networks for LIDAR–camera fusion for Pedestrian detection Framework

***Step 2:*** *The fused error is computed by squaring the actual prediction and observation.*

$$Fus_{Accurancy} = \sqrt{pre^2 + obs^2}$$

*Where obs is actual observation and pre is for prediction and*
***Step 3:*** *actual observation obs is obtained by detracting Mean Target by mean fused*

$$O = \phi_{tr} - \phi_{fi}$$

*Where $_{tr}$ means target region and $_{fi}$ means the fused image*
***Step 4:*** *The accuracy of predictions is determined by the standard deviation of the specific target image and the fused image.*

$$P = SD_{tr} - SD_{fi}$$

*Where $SD_{tr}$ is the standard deviation of the target image and $SD_{fi}$ is the standard deviation of the fused image.*
***Step 5:*** *Hence using correlation coefficients, the statistically significant association between the targeted and fused images is calculated as follows.*

$$Pyz = C(y, z) / \left(\sigma_y \sigma_z\right)$$

*Where Pyz denotes correlation coefficient product*
*C(y,z) covariance of y and z*
*σy standard deviation of y*
*σz standard deviation of z*
***Step 6:*** *The distances between the pedestrian and each LIDAR point are computed to determine whether or not the points form a linear sequence.*
***Step 7:*** *The obstacle is recognized if the distance between points is significantly greater than the sequence's mean distance.*

### 3.2 Fully convolutional neural networks for LIDAR–camera fusion for pedestrian detection

The proposed fully convolutional neural networks for lidar–camera fusion for pedestrian detection are illustrated in Fig. 2. To minimize overall processing time, the fusion approach builds a correlation between the 3D points from LiDAR and the object identified by a camera. Because sensor fusion increases robustness and detection accuracy while making up for the limitations of the individual sensors is attainable. Real-time information from 3D camera detection and camera mapping is predominant for camera detection. Using the camera mapping technique, users may project images onto 3D objects to give the impression that they were actually captured with a moderate camera movement. Further, for lidar data, the real-time lidar processing information and point cloud data are fused for better efficiencies. In the proposed model, the role of Point Clouds significantly impacts the image fusion mechanism in autonomous vehicles. Point clouds are collections of points that are used to characterize an item or surface. A point cloud may be produced using laser scanning technologies like LiDAR. There is enough information at each site to generate a 3D model or to combine it with information from other sources. A form or feature in three dimensions can be represented by a point cloud, a collection of such points. Every point has its own unique set of X, Y, and Z coordinates, in addition to other characteristics in some instances. In comparison, it is possible to conceptualize a point cloud as a collection of many points.

The proposed model discusses two fundamental challenges that are associated with the integration of sensor data. The geographical inaccuracy and quality variation in heterogeneous sensors are the two problems. Heterogeneous wireless sensor networks (heterogeneous WSNs) are made up of sensor nodes that have varying capabilities, such as computational power and sensing range. When compared to homogeneous WSN, heterogeneous WSN deployment and topology control are more difficult. In addition to being distinct from other contributions in the kinds of sensors that were utilized for data fusion, the dual goals that prompted us to produce this investigation are as follows: To begin, one of our primary goals is to devise a method for data fusion that is more resilient and which takes into consideration the uncertainty inherent in the fusion process. Because of this, the following perceptual tasks in an autonomous vehicle will be able to function with a higher degree of reliability. Second, we foresee a day in the not-too-distant future when self-driving cars can communicate with one another and share information gathered from their sensors. The outputs from the camera and the LiDAR can be combined to help overcome the constraints that each of them has on its own. The fusion yields reliable results for various applications, including determining the depth of an image's scene and locating objects in images. The fusion of data from a camera and a LiDAR can be accomplished in one of two ways: by combining the data findings. The combination of different types of data is referred to as data fusion, whereas the combination of different types of outcomes is referred to as output fusion.

The camera picture and the LiDAR point cloud are superimposed on one another during the fusion process to provide depth information for the pixels that make up the camera image. For this fusion, you will first need to locate the place where the camera and the LiDAR meet, and then you will need to assign the remaining points in the point cloud to the appropriate pixels in the picture. Upsampling, this output would then allow one to extract depth values for each pixel in the image. Whereas in output fusion, the fusion of results happens when, for example, we conduct object recognition in the camera picture and the LiDAR point cloud independently, and then we fuse the findings to boost our confidence. Separately from the LiDAR data processing, we process the camera picture to obtain an output, which we then test against the processed LiDAR output or vice versa. Because it has the potential to contribute to an increase in a system's dependability, this fusion approach is one of the most beneficial ones available.

### 3.3 Algorithm for pedestrians' detection and identification

We propose an algorithm for Pedestrians' detection and identification in autonomous vehicles. Furthermore, the LiDAR output's data points and the 360-Degree camera is geometrically aligned.

Additionally, to find the corresponding pixel in the camera output for each data point output by the LiDAR sensor. Moreover, forward displacement of the centers of LiDAR and Camera sensor and the Horizontal displacement of the centers of LiDAR and Camera sensor is estimated. Furthermore, this alignment aims to find the corresponding pixel in the camera output for each data point output by the LiDAR sensor Fig. 3.

> ***Step 1****: Geometrically align the data points of the LiDAR output and the 360-Degree camera.*

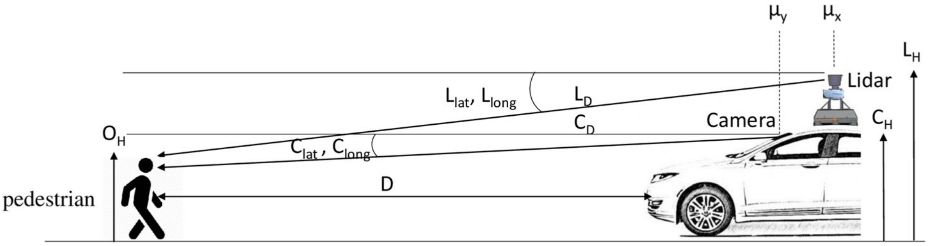


**Fig. 3** Proposed framework for Pedestrian detection and identification

***Step 2***: *To find the pixel in the camera output for each data point output by the LiDAR sensor.*
***Step 3***: *object of height at distance from the Autonomous Vehicle.*
***Step 4***: $\mu x$ = *Forward displacement of the centers of LiDAR and Camera sensor.*
***Step 5***: $\mu y$ = *Horizontal displacement of the centers of LiDAR and Camera sensor.*
***Step 6***: *OH* = *distance to the object O sensed by Autonomous Car.*
***Step 7***: *LD* = *distance to the object O sensed by LiDAR.*
***Step 8:*** *CD* = *distance to the object O sensed by Camera Sensor.*
***Step 9:*** *Llat, Llong* = *Latitude and longitude of object O as measured by the LiDAR.*
***Step 10:*** *Clat, Clong* = *Latitude and longitude of object O as measured by the Camera Sensor.*
***Step 11:*** *CH* = *Height of the camera from the ground.*
***Step 12:*** *LH* = *Height of the Lidar from the ground.*
***Step 13:*** *LD, Llat, Llong are the outputs of the LiDAR sensor.*
***Step 14:*** *This alignment aims to find the corresponding pixel in the camera output for each data point output by the LiDAR sensor.*

The proposed system model estimates the distance between the object (Pedestrian) and the camera using eq. 1. Moreover, to compute the pedestrian's vertical height, we use eq. 2. We calculate the camera's corresponding latitude from Eqs. 1 & 2 using eq. 3. Furthermore, through eq. 4, we calculate the horizontal displacement of the object, which is a pedestrian. Additionally, through eq. 1 and eq. 4, we calculate the longitude of the camera. Finally, using eqs. (3) and (5), we align the data points of the LiDAR and the camera.

To calculate the distance of object O

$$D=: \ LD \cos Llat.\cos Llong = r.COS\ Clat, \cos Clong + \mu x \tag{1}$$

To calculate the vertical height of the object O,

$$OH = LH - LD.\sin Llat = CH - r.\sin Clat \tag{2}$$

From 1 and 2 we can calculate the corresponding latitude of the camera Clat

$$\tan Clat = ((CH - LH) + LD.\sin Llong).\cos Clong / LD.\cos Llat.\cos Llong - \mu x \tag{3}$$

The horizontal displacement is calculated as

$$LD.\cos Llat.\sin Llong = r.\cos Clat.\sin Clong = \mu y \tag{4}$$

From 1 and 4 longitude of the camera Clong

$$tan\ Clong = LD.cos\ Llat.sin\ Llong + \mu y / LD.cos\ Llat.cos\ Llong - \mu x \tag{5}$$

The Eqs. (3) and (5) pave way to align the data points of the LiDAR and the camera.

## 4 Result and discussion

In the proposed model, the camera picture and the LiDAR point cloud are superimposed on top of one another during the fusion process in order to provide depth information for the pixels that make up the camera image. For this fusion, we will first need to locate the place where the camera and the LiDAR meet, and then we will need to assign the remaining points in the point cloud to the appropriate pixels in the picture.

Upsampling, this output would then allow one to extract depth values for each pixel in the image. Whereas in output fusion, the fusion of results happens when, for example, we conduct object recognition in the camera picture and the LiDAR point cloud independently, and then we fuse the findings to boost our confidence. Separately from the LiDAR data processing, we process the camera picture to obtain an output, which we then test against the processed LiDAR output or vice versa. Because it has the potential to contribute to an increase in a system's dependability, this fusion approach is one of the most beneficial ones available. Figure 4 illustrates the Fully Convolutional Neural network's real-time prediction and identification. Furthermore, the performance analysis of the proposed algorithm is estimated based on Precision, Sensitivity, Accuracy, and Matthew Correlation Coefficient (MCC) Table 1.

Precision refers to the proportion of relevant outcomes. In contrast, recall refers to the proportion of accurate cumulative evidence categorized adequately by the proposed algorithm. Consequently, the error is discovered, which must be reduced in future results. Figure 5 explains the performance evaluation of precision and recall for the retrieved dataset. It illustrates how the proposed approach fluctuates with distinct inputs. Furthermore, it evaluates the various attributes through obtaining outcomes, including random, poor er, good er, excel, and perf as mentioned in Table 1. Ten positive samples and ten negative samples are collected.

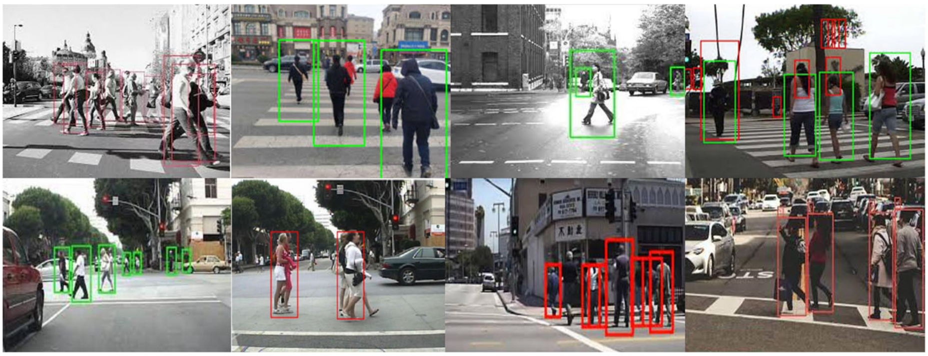

**Fig. 4** Proposed Fully CNN Pedestrians' detection and identification

**Table 1** Performance Metrics

| Level Name | Description |
|---|---|
| random | Random |
| poor_er | Poor early retrieval |
| good_er | Good Early retrieval |
| excel | Excellent |
| perf | Perfect |

Precision-Recall curves with 95% confidence bounds are shown in the results obtained by an error detection method. Furthermore, with 95% confidence intervals, there's fluctuation in performance metrics. The rate of precision and recall following error discovery is depicted in Fig. 6.

Figure 7 depicts the results of a performance evaluation of the suggested algorithm based on sensitivity to interference. Due to the perceived suggested model's high sensitivity maintains an overall sensitivity of more than 86%. Similarly, to precision evaluation, the sensitivity is inferred by splitting the correct positive (+VE) estimations by the overall number of correct positive (+VE). In addition, the accuracy of the proposed algorithm is analyzed quantitatively. All of the criteria, such as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), are maintained well above 82% of accuracy.

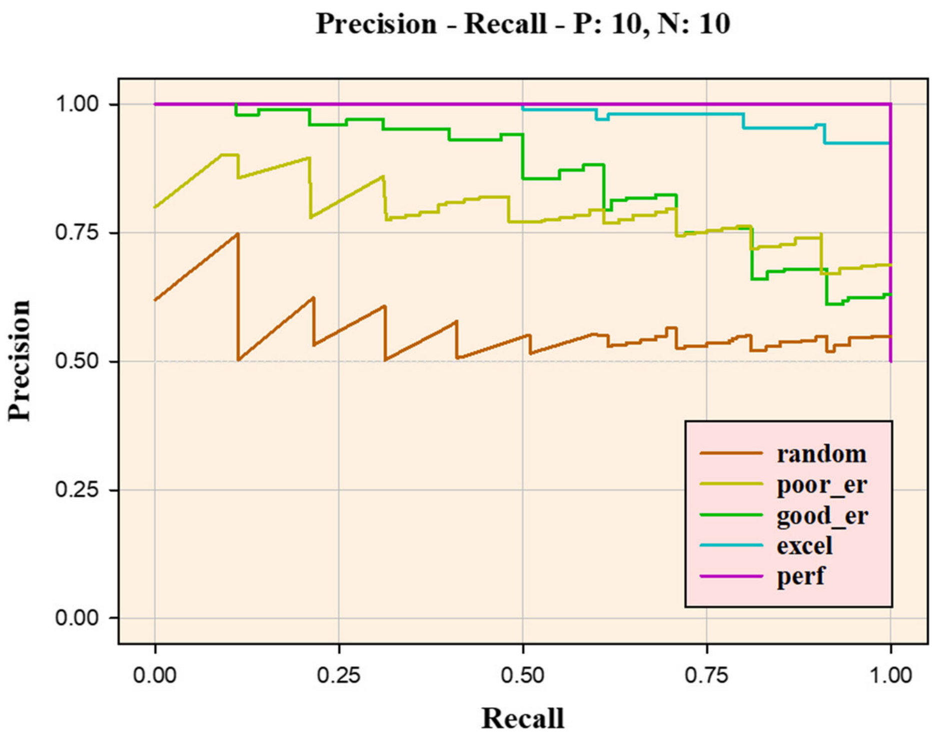


**Fig. 5** Proposed model evaluation based on Precision Vs Recall

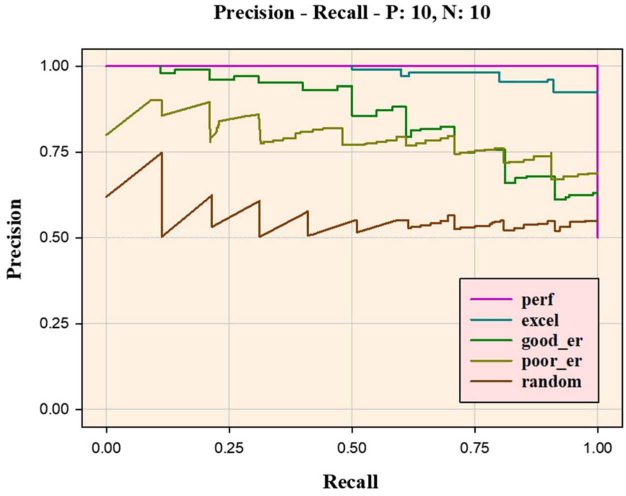


**Fig. 6** Proposed model evaluation based on Precision Vs. Recall after Error Deduction

Figure 8 shows the difference between the Error and Accuracy for different values by focusing on the normalized rank as a fixed variable. The normalized rank is used to determine the rate of accuracy. Scale values are derived using a normalized rank approach. The findings indicate that there are influences on the obtained scale values from both groups and situation. Each variable receives equal weights/importance through normalization, ensuring that no one variable may influence model performance just because they have larger numbers. Besides, the proposed algorithm uses normalized rank to minimize the error. As a result, the accuracy of the proposed algorithm is better comparatively.

The sensitivity of a prediction model is defined as the ratio between the amount of information that was accurately labeled as positive and the amount of genuinely positive data. Furthermore, the specificity of a classification model is defined as the ratio between the amount of data that was accurately classified as negative and the amount of data that was substantially negative. Moreover, accuracy is concerned with how many positives were

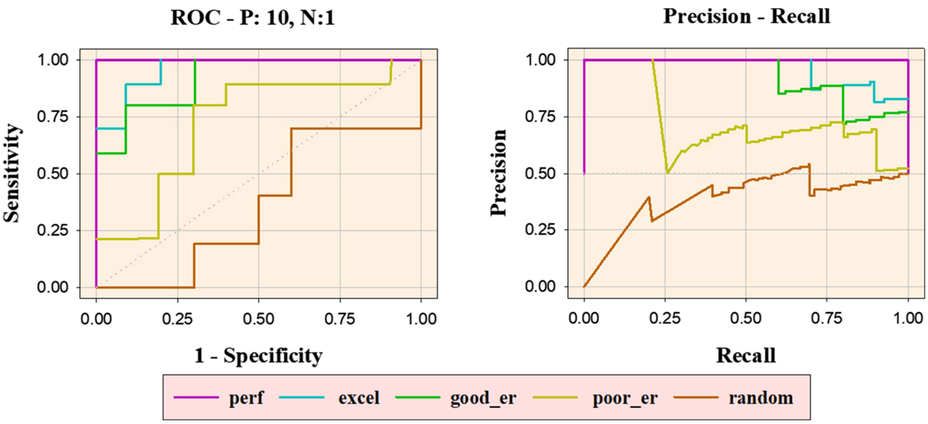


**Fig. 7** Proposed model evaluation based on Sensitivity vs Specificity and Precision Vs recall

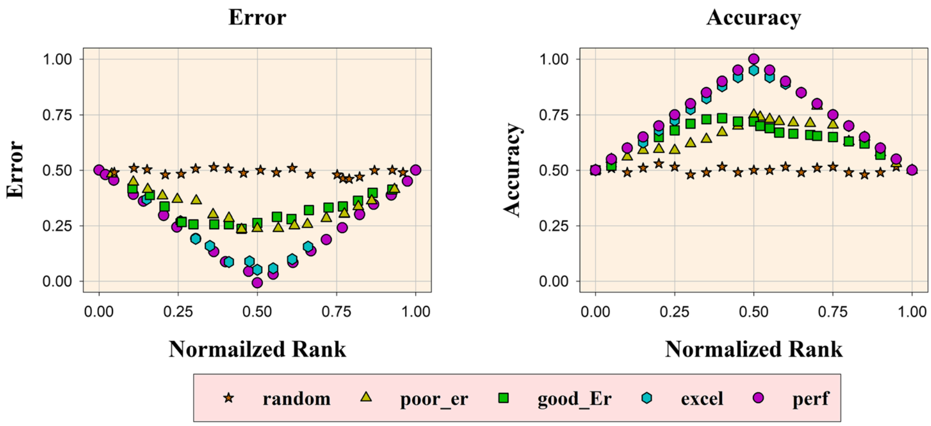


Fig. 8 Proposed model evaluation based on Error vs. Accuracy

accurately identified as positive among all the positives. This study evaluates all three parameters with a normalized rank, demonstrating that the suggested approach produces an effective result, as mentioned in Fig. 9.

A confusion matrix, commonly referred to as an error matrix, is a matrix that contains ambiguous information. The MCC method is used to determine how powerfully the error rate is dropping, and F-score is determined by the ratio of precision to recall. MCC, like the F-score, is a single-value statistic that summarizes the confusion matrix and is similar to it. The F1 score does not take into consideration the number of True Negatives. On the other hand, MCC is good enough to assist all four indicators in the confusion matrix for the proposed model, as mentioned in Fig. 10.

Furthermore, in dealing with normalized ranks against proposed predicted values, we use the autoplot function to plot normalized ranks against scores and labels. Autoplot is a general function for visualizing diverse data objects; it attempts to provide better default visuals and customizable options for each data type; it is faster and more comfortable to study genomic data than the low-level ggplot approach and it is much simpler and easier to generate relatively complicated images. Figure 11, depicts that the score fluctuates based on the performance metrics used, and the label values are random.

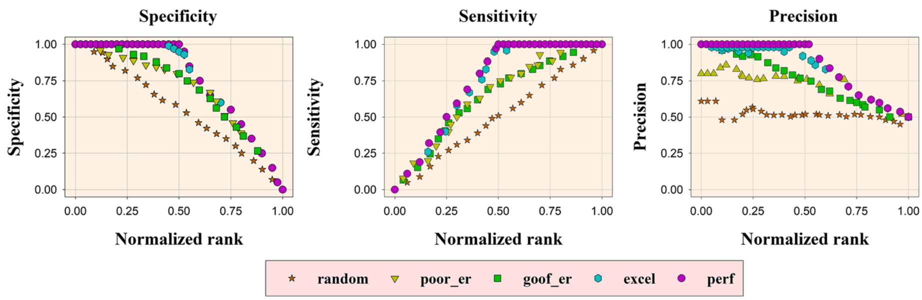


Fig. 9 Proposed model evaluation based on Specificity, Sensitivity, and Precision

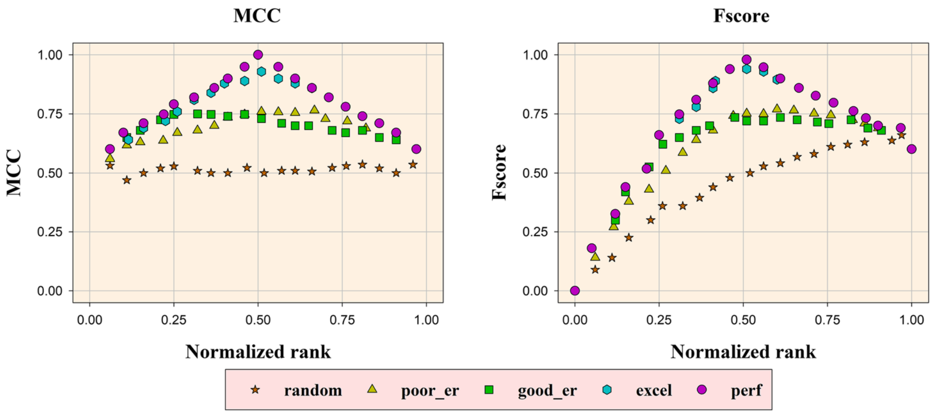


Fig. 10 Proposed model evaluation based on MCC & F-score

In the proposed model, we compute the F1 curve to determine the weighted average of recall and precision. Further, in the F1 curve, the values must be between 0 and 1, where 1 denotes the highest accuracy. As depicted in Fig. 12, the F1 curve value range is maintained, and the detection accuracy is attained in the proposed work.

P value is used to determine the point of rejection to offer the minimal significance level when null hypotheses are lowest or rejected. It is defined as the level confidence interval that ranges between 0 and 1, and if the p value is less, and there would be substantial probability of rejecting the null hypothesis as mentioned in Fig. 13. This would be done in order to find the threshold at which the null hypothesis gets rejected.

Precision-recall curves provide a concise summary of the trade-off between a predictive model's actual positive rate and positive predictive value. This trade-off occurs when different probability thresholds are used. As mentioned in Fig. 14, the proposed model evaluation based on PR_Curve is well maintained based on the actual positive rate and positive predicted value. In Figs. 15 and 16, the proposed model evaluation is based on comparative analysis based on metrics such as training dataset versus box_loss, obj_loss, and cls_loss, etc.,

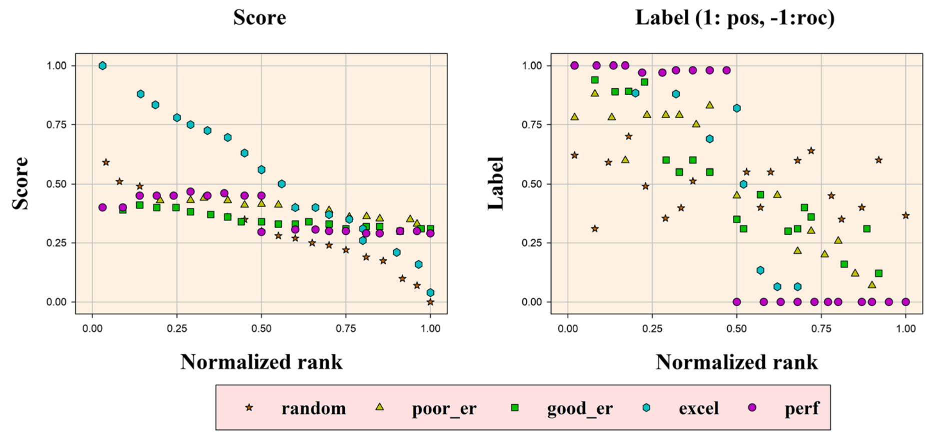


Fig. 11 Proposed model evaluation based on Normalized & Score & Label

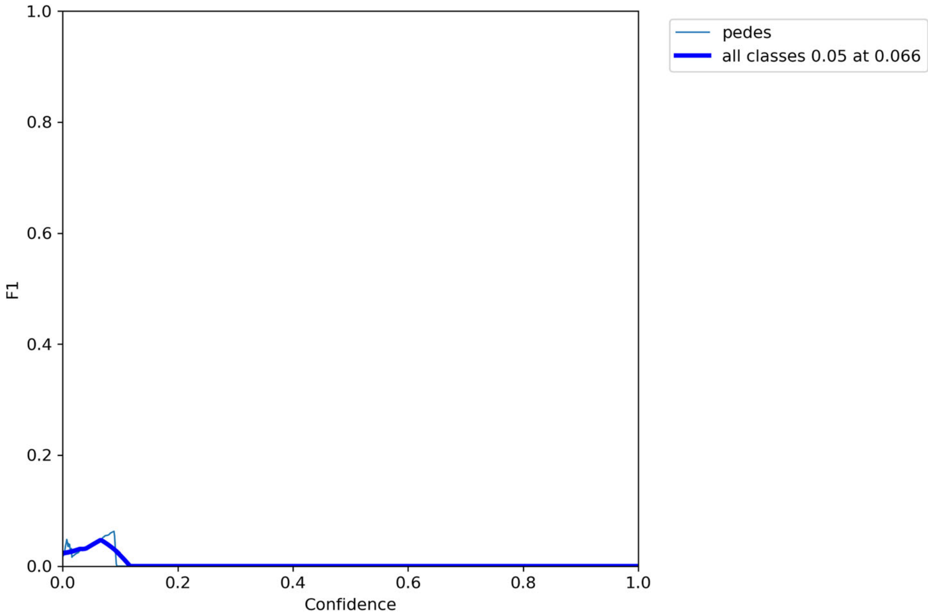


**Fig. 12** Proposed model evaluation based on F1_curve

Furthermore, the performance analysis of the proposed model is evaluated and estimated based on Precision, Sensitivity, Accuracy, and Matthew Correlation Coefficient (MCC). Matthew's correlation coefficient (MCC) is a statistical technique used for model assessment. Its purpose is to assess or quantify the difference between the expected and actual values, and it

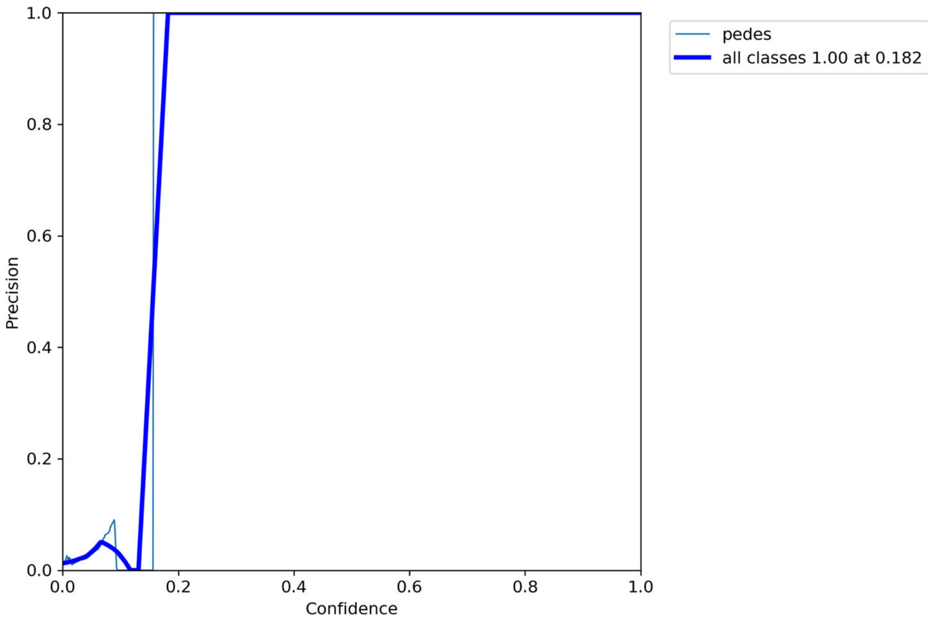


**Fig. 13** Proposed model evaluation based on P_Curve

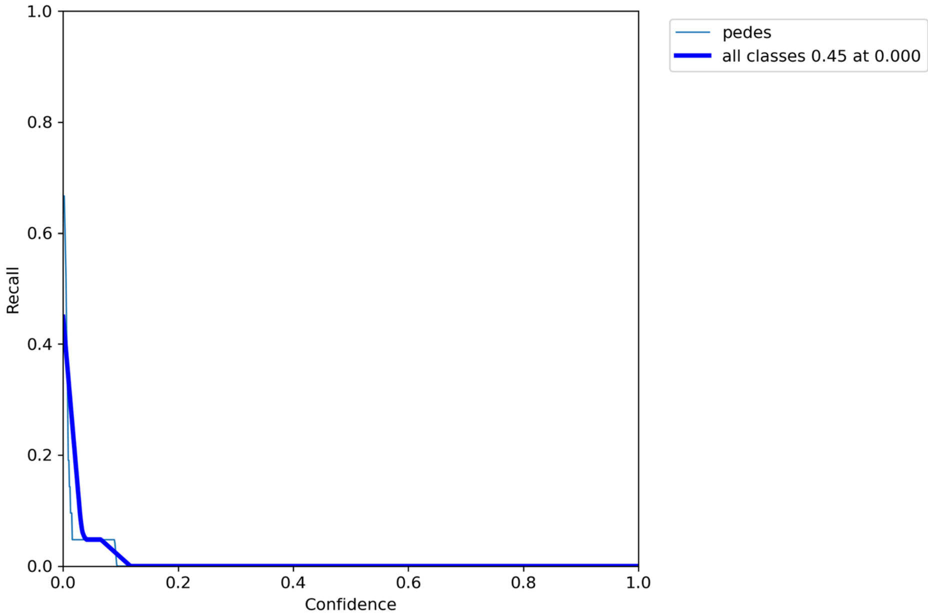


**Fig. 14** Proposed model evaluation based on PR_Curve

is comparable to chi-square statistics for a 2 × 2 contingency table. Additionally, we determine the F_Curve, P_Curve, R_Curve, and PR_Curve. Finally, we evaluate the comparative analysis based on metrics such as training dataset versus box_loss, obj_loss, cls_loss etc. Henceforth, the precision of the proposed method is calculated by no of correct positive (+VE)

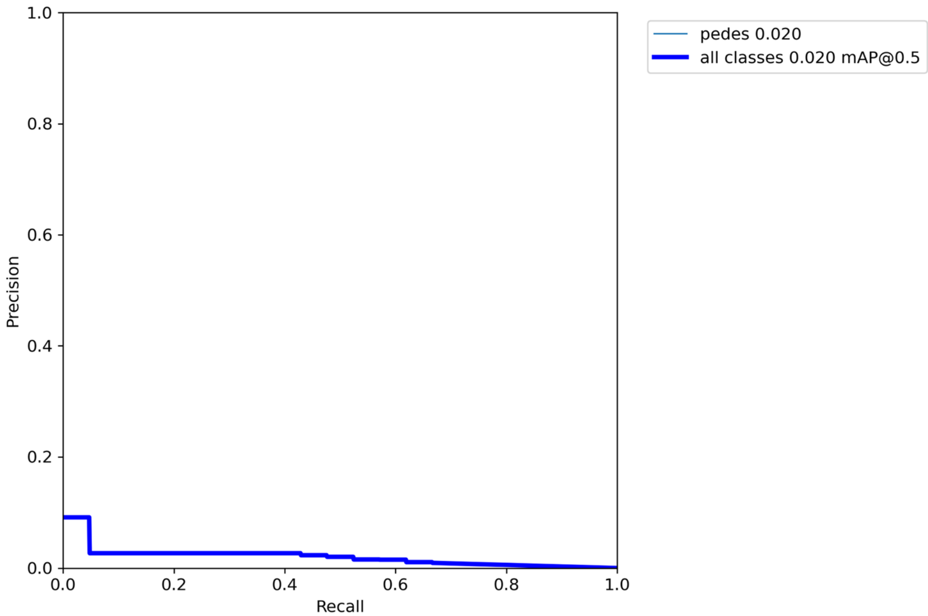


**Fig. 15** Proposed model evaluation based on R_Curve

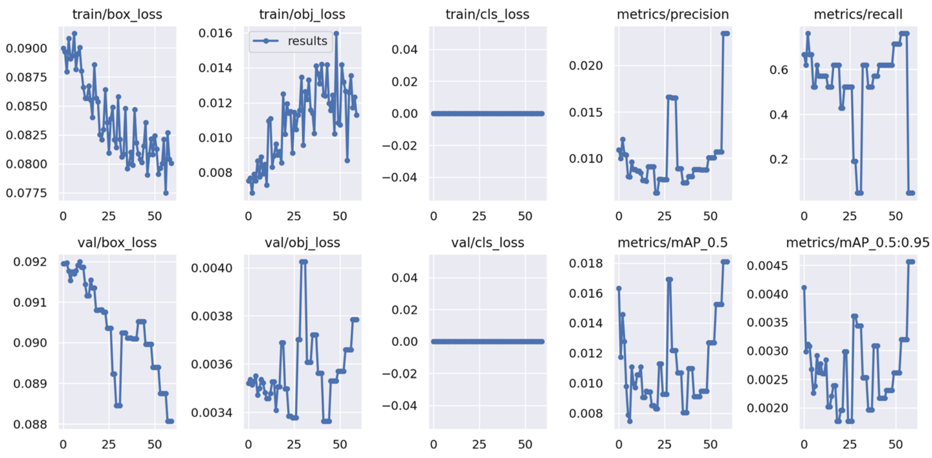


**Fig. 16** Proposed model evaluation based on Comparative analysis based on metrics

**Table 2** Fusion performance 10 to 30 m range

| ATTRIBUTES | LIDAR | CAMERA | FUSION |
|---|---|---|---|
| True Positive Rate (TPR) | 94.13% | 98.03% | 98.06% |
| False Positive Rate (FPR) | 26.35% | 0.00% | 6.54% |
| False Negative Rate (FNR) | 6.05% | 2.25% | 2.25% |

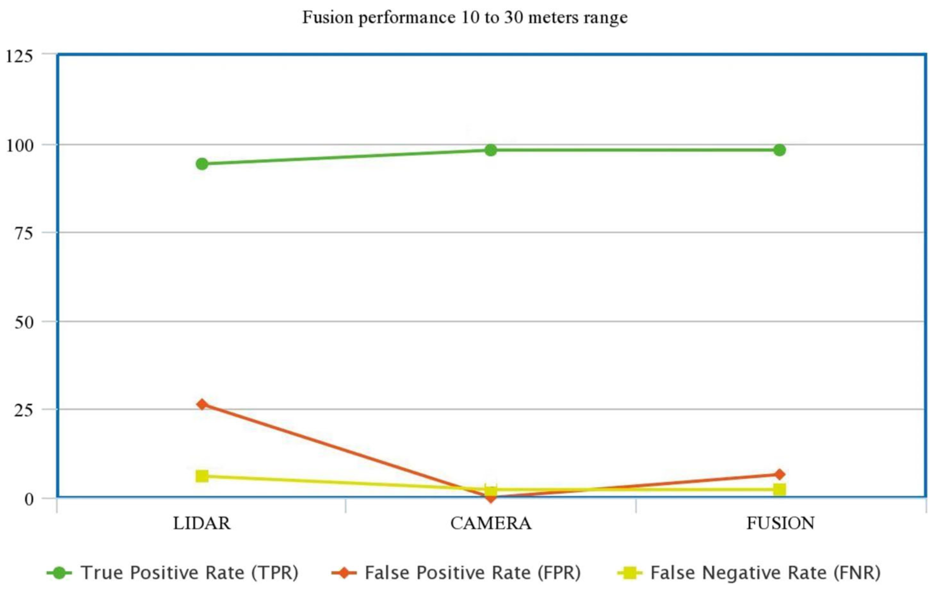


**Fig. 17** Fusion performance 10 to 30 m range

predictions by the total no of positive (+VE) predictions which is illustrated in Table 2: Fusion performance 10 to 30 m range (Fig. 17).

## 5 Conclusion

In the commencement of this paper, we highlighted the potential and constraints associated with successfully employing deep learning in autonomous vehicles (AV) for pedestrian detecting objection detection. Pedestrian detection appears to be an integral part of a vast array of vision-based technologies, ranging from object recognition and detection in an autonomous vehicle. Moreover, the fast adoption of Convolutional Neural Networks (CNN) for item recognition and pedestrian detection in autonomous cars has attained a very high level of performance in both training and assessment environments. A data fusion method called Fully Convolutional Neural networks for LIDAR–camera fusion, which integrates Lidar data with multiple camera pictures to produce an ideal pedestrian detection solution, is suggested to achieve the object identification that is pedestrian detection. A different technique is given in the system model for picture fusion in pedestrian detection. In addition, architecture and framework for Fully Convolutional Neural networks for LIDAR–camera fusion for Pedestrian identification are proposed. In addition, a fully functioning algorithm for detecting and identifying pedestrians is suggested to precisely identify pedestrians between 10 and 30 m away. In realization, the performance of the suggested model is assessed using a variety of characteristics, including precision, sensitivity, accuracy, and the Matthew correlation coefficient (MCC). Furthermore, we calculate the F-Curve, the P-Curve, the R-Curve, and the PR-Curve. In the end, we evaluate the based on a comparison analysis using metrics such as training dataset against box loss, obj loss, cls loss, etc. Henceforth, the proposed system model was shown to be effective in comparison.

**Author contribution** Alfred Daniel J, Chandru Vignesh C, Bala Anand Muthu, is responsible for designing the framework, analyzing the performance, validating the results, and writing the article. Sivaparthipan CB, Carlos Enrique Montenegro Marin, is responsible for collecting the information required for the framework, provision of software, critical review, and administering the process.

**Funding** Authors did not receive any funding.

**Data availability** Available on request.

**Code availability** Not applicable.

## Declarations

**Ethics approval** This article does not contain any studies with human participants performed by any of the authors.

**Informed consent** Not Applicable.

**Conflict of interest** Authors do not have any conflicts.

## References


1. Alfred Daniel J, Karthik S, Newlin Rajkumar M (2021, ISSN 0141-9331) Optimizing spectral efficiency based on poisson queuing model for procuring cognitive intelligence in vehicular communication. Elsevier Microprocess Microsyst 82:103819. https://doi.org/10.1016/j.micpro.2020.103819
2. Caltagirone L, Bellone M, Svensson L, Wahde M (2019) LIDAR–camera fusion for road detection using fully convolutional neural networks. Robot Auton Syst 111:125–131
3. Daniel A, Paul A, Ahmad A, Rho S (2016) Cooperative intelligence of vehicles for intelligent transportation systems (ITS). Wirel Pers Commun 87(2):461–484
4. Daniel A, Subburathinam K, Paul A, Rajkumar N, Rho S (2017) Big autonomous vehicular data classifications: towards procuring intelligence in ITS. Veh Commun 9:306–312
5. Daniel A, Subburathinam K, Muthu BA, Rajkumar N, Kadry S, Mahendran RK, Pandian S (2020) Procuring cooperative intelligence in autonomous vehicles for object detection through data fusion approach. IET Intell Transp Syst 14(11):1410–1417
6. De Silva V, Roche J, Kondoz A (2018) Robust fusion of LiDAR and wide-angle camera data for autonomous mobile robots. Sensors 18(8):2730
7. Girshick R, Donahue J, Darrell T, Malik J (2016) Region-based convolutional networks for accurate object detection and segmentation. IEEE Trans Pattern Anal Mach Intell 38(1):142–158. https://doi.org/10.1109/TPAMI.2015.2437384
8. Han T-T, Ge SS (2015) Styled-velocity flocking of autonomous vehicles: a systematic design. IEEE Trans Autom Control 60:2015–2030. https://doi.org/10.1109/TAC.2015.2400664
9. Han X, Wang H, Lu J, Zhao C (2017) Road detection based on the fusion of lidar and image data. Int J Adv Robot Syst 14(6):1729881417738102
10. Hasan I, Liao S, Li J, Akram SU, Shao L (2022) Pedestrian detection: domain generalization, CNNs, transformers and beyond. *ArXiv, abs/2201.03176*
11. Hou YL, Song Y, Hao X, Shen Y, Qian M, Chen H (2018) Multispectral pedestrian detection based on deep convolutional neural networks. Infrared Phys Technol 94:69–77. https://doi.org/10.1016/j.infrared.2018.08.029
12. Kim H, Liu B, Goh CY, Lee S, Myung H (2017) Robust vehicle localization using entropy-weighted particle filter-based data fusion of vertical and road intensity information for a large scale urban area. IEEE Robotics Autom Lett 2(3):1518–1524. https://doi.org/10.1109/LRA.2017.2673868
13. Liang M, Yang B, Wang S, Urtasun R (2018) Deep continuous fusion for multi-sensor 3d object detection. In: Proceedings of the European conference on computer vision (ECCV), pp. 641-656
14. Liu Z, Zhang W, Lin S, Quek TQS (2017) Heterogeneous sensor data fusion By deep multimodal encoding. IEEE J Sel Top Signal Process 11(3):479–491. https://doi.org/10.1109/JSTSP.2017.2679538
15. Paul A, Chilamkurti N, Daniel A, Rho S (2016) Intelligent vehicular networks and communications: fundamentals, architectures, and solutions. Morgan Freeman Publisher, Elsevier SN - 0128095466
16. Paul A, Daniel A, Ahmad A, Rho S (2017) Cooperative cognitive intelligence for internet of vehicles. *IEEE Syst J* 11(3):1249–1258. https://doi.org/10.1109/JSYST.2015.2411856
17. Tianrui L, Tania S (2018) Faster R-CNN for Robust Pedestrian Detection Using Semantic Segmentation Network. Front Neurorobot 12, ISSN=1662–5218. https://doi.org/10.3389/fnbot.2018.00064
18. Wang H, Lou X, Cai Y, Li Y, Chen L (2019) Real-time vehicle detection algorithm based on vision and LiDAR point cloud fusion. Journal of Sensors 2019:1–9
19. Wei P, Cagle L, Reza T, Ball J, Gafford J (2018) LiDAR and camera detection fusion in a real-time industrial multi-sensor collision avoidance system. Electronics. 7(6):84. https://doi.org/10.3390/electronics7060084
20. Wen L-H, Jo K-H (2021) Fast and accurate 3D object detection for Lidar-camera-based autonomous vehicles using one shared voxel-based backbone. IEEE Access 9:22080–22089. https://doi.org/10.1109/ACCESS.2021.3055491
21. Wu T-E, Tsai C-C, Guo J-I (2017) LiDAR/camera sensor fusion technology for pedestrian detection. 2017 Asia-Pacific signal and information processing association annual summit and conference (APSIPA ASC), pp. 1675-1678. https://doi.org/10.1109/APSIPA.2017.8282301
22. Zhao G, Xiao X, Yuan J, Wahng G (2014) Fusion of 3D-LIDAR and camera data for scene parsing. J Vis Commun Image Represent 25(1):165–183, ISSN 1047-3203. https://doi.org/10.1016/j.jvcir.2013.06.008
23. Zhao X, Sun P, Xu Z, Min H, Yu H (2020) Fusion of 3D LIDAR and Camera Data for Object Detection in Autonomous Vehicle Applications. *IEEE Sensors J* 20(9):4901–4913. https://doi.org/10.1109/JSEN.2020.2966034

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## Affiliations

**J Alfred Daniel[1] · C Chandru Vignesh[2] · Bala Anand Muthu[3] · R Senthil Kumar[4] · CB Sivaparthipan[3] · Carlos Enrique Montenegro Marin[5]**

J Alfred Daniel
85.alfred@gmail.com

Bala Anand Muthu
balavdy@gmail.com

R Senthil Kumar
senthil2481@gmail.com

CB Sivaparthipan
sivaparthipanece@gmail.com

Carlos Enrique Montenegro Marin
cemontenegrom@udistrital.edu.co

1 Department of Computer Science and Engineering, Faculty of Engineering, Karpagam Academy of Higher Education, Eachanari, India

2 School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

3 Department of Computer Science and Engineering, Tagore Institute of Engineering and Technology, Attur, India

4 Department of Computer Science and Engineering, Hindusthan Institute of Technology, Malumichampatti, India

5 District University Francisco José de Caldas, Bogotá, Colombia