

A terrain data collection sensor box towards a better analysis of terrains conditions

Kouame Yann Olivier Akansie¹, Rajashekhar C. Biradar¹, Karthik Rajendra¹, Geetha D. Devanagavi²

¹School of Electronics and Communication Engineering, REVA University, Bangalore, India

²School of Computing and Information Technology, REVA University, Bangalore, India

Article Info

Article history:

Received Jan 25, 2024

Revised Mar 26, 2024

Accepted Apr 17, 2024

Keywords:

Data collection method

Dataset creation

Image classification

Robotic terrain perception

Environment perception

ABSTRACT

Autonomous mobile robots are increasingly used across various applications, relying on multiple sensors for environmental awareness and efficient task execution. Given the unpredictability of human environments, versatility is crucial for these robots. Their performance is largely determined by how they perceive their surroundings. This paper introduces a machine learning (ML) approach focusing on land conditions to enhance a robot's locomotion. The authors propose a method to classify terrains for data collection, involving the design of an apparatus to gather field data. This design is validated by correlating collected data with the output of a standard ML model for terrain classification. Experiments show that the data from this apparatus improves the accuracy of the ML classifier, highlighting the importance of including such data in the dataset.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Kouame Yann Olivier Akansie

School of Electronics and Communication Engineering, REVA University

Rukmini knowledge park, Kattigenahalli, Yelahanka, Bangalore, 560064, India

Email: akansiekouame@gmail.com

1. INTRODUCTION

Robots that can move around autonomously in their surroundings without human assistance are known as mobile robots. These machines may be built for various jobs, from straightforward surveillance and inspection to intricate manipulation and assembly. They generally sense and move through their surroundings using various sensors, including cameras, lidar, sonar, and infrared sensors. Their performances highly depend on how they perceive their surroundings, as poor perception can lead to poor decision-making. The next factor that may affect their performance is the result of decisions made from environmental perception, such as motion planning. In this regard, researchers proposed techniques based on semantic perception, deep learning, visual perception, and multi-sensor perception to ensure that autonomous mobile robots understand their environment accurately.

Semantic segmentation-based perception is a technique that involves segmenting images into semantically meaningful regions to comprehend the environment. Wu *et al.* [1] proposed an object simultaneous localization and mapping (SLAM) framework that integrates visual sensors, such as cameras, for association, mapping, and high-level tasks in robotics. Similarly, Nan *et al.* [2] developed a joint object detection and semantic segmentation model using visual sensors to enhance robot perception capabilities. Betoño *et al.* [3] applied semantic segmentation for developing an indoor navigation system, relying on visual sensors like cameras. These approaches primarily utilize cameras for capturing images and performing semantic segmentation, enabling robots to perceive and understand their surroundings in real-time. Additionally, Several researchers demonstrated the use of semantic environment modeling for vision-based global localization and autonomous navigation, respectively [4], [5], further highlighting the significance of

semantic segmentation in robotic perception tasks. Furthermore, Zhang *et al.* [6] explored intelligent collaborative localization among air-ground robots, leveraging semantic segmentation with visual sensors to enhance environment perception for industrial applications.

Transitioning to machine learning (ML)-based environment perception techniques, Ginerica *et al.* [7] proposed a vision dynamics learning approach to robotic navigation in unstructured environments, incorporating various sensors for perception tasks. Singh *et al.* [8] presented an efficient deep learning-based semantic mapping approach utilizing monocular vision, while Bena *et al.* [9] developed a safety-aware perception system for autonomous collision avoidance in dynamic environments, leveraging sensor fusion techniques. Sultana *et al.* [10] developed a vision-based robust lane detection and tracking system, suggesting the use of camera sensors for lane detection tasks. Teixeira *et al.* [11] explored deep learning for underwater visual odometry estimation, potentially employing underwater imaging sensors for navigation. Bekiarski [12] discussed visual mobile robot perception for motion control, which may involve camera sensors. Kowalewski *et al.* [13] focused on semantic mapping and object detection, indicating the utilization of various sensors for mapping and localization tasks. Ran *et al.* [14] addressed scene perception-based visual navigation in indoor environments, likely involving camera sensors for scene understanding and navigation.

Moving on to multisensor or sensor fusion-based environment perception techniques, Ge *et al.* [15] introduced an object localization system using monocular cameras and laser ranging sensors, highlighting fusion of visual and range data. Xia *et al.* [16] presented a visual-inertial SLAM method, leveraging visual and inertial sensors for robust navigation and mapping. Xie *et al.* [17] proposed a method for moving object segmentation and detection in dynamic environments, likely incorporating red green blue-depth (RGB-D) sensors alongside visual cameras. Surmann *et al.* [18] explored deep reinforcement learning for autonomous navigation, indicating potential sensor fusion techniques. Guo *et al.* [19] addressed autonomous navigation in dynamic environments with multi-modal perception uncertainties, suggesting fusion of data from multiple sensors. Luo [20] presented a multi-sensor-based strategy learning approach with deep reinforcement learning, integrating data from various sensors. Huang *et al.* [21] proposed a multi-modal perception-based navigation method using deep reinforcement learning, indicating fusion of data from multiple sensors. Nguyen *et al.* [22] discussed autonomous navigation in complex environments with a deep multimodal fusion network, highlighting sensor fusion. Feng *et al.* [23] addressed deep multi-modal object detection and semantic segmentation, likely incorporating data from multiple sensors. Braud *et al.* [24] focused on robot multimodal object perception and recognition, suggesting integration of information from multiple sensors. Lastly, Yue *et al.* [25] explored day and night collaborative dynamic mapping based on multimodal sensors, indicating fusion of data for mapping in various lighting conditions. From this review, we identify a diverse range of techniques and sensors used in environment perception, providing insights for designing hardware and methodologies for robotic applications.

2. METHOD

This study endeavors to propose a comprehensive system and methodologies for terrain data collection, aimed at augmenting the performance of conventional ML image classifiers. The devised data gathering procedure stands as a crucial mechanism for acquiring indispensable data, establishing a direct nexus between the collected data and the efficacy of conventional ML image classifiers. However, it is noteworthy that the analysis in this research is confined to a limited subset of parameters. Future investigations could expand upon this by incorporating additional sensors into the design to enhance the sensing system's capabilities, thus addressing this constraint. Furthermore, this research refrains from introducing mathematical models to elucidate the relationship between the performance of traditional ML image classifiers and the parameters scrutinized. The subsequent sections of this paper delineate the exhaustive design and development of the hardware from multifarious perspectives, outlining the proposed methodology for terrain data collection. Additionally, pertinent experiments are conducted to validate the effectiveness of the developed hardware.

2.1. Dataset creation system design

The data collection system consists of the sensor box and the control circuit for the whole mechanism. The sensor box is a combination of sensors used to collect crucial data like terrain images, surrounding temperature, humidity, pressure, light intensity, and speed. The sensor box should be mounted on a robot, with all the sensors facing downward, toward the terrain to enable data collection and further terrain analysis. The system should be trained first before any terrain analysis can be performed. The training session requires data to be collected, which requires the use of a control structure for dataset collection. Obtaining a quality dataset implies some parameters to be considered based on their effect on the data gathered. In our case, such parameters include the height from which the data is collected, the angle the sensors make with the terrain, the light intensity, and many more. The sensor box itself should be designed to get many such parameters.

2.1.1. Sensor box design

a) System design

Sensors are essential for robots to understand their environment. Data collected by the sensors is used in many algorithms for different tasks. In our proposal, we use two cameras to achieve a stereovision, some laser distance sensors, an accelerometer and a gyroscope, a light sensor, a humidity, temperature, and pressure sensors. Each of these sensors is integrated to achieve different tasks. For instance, the robot will get a clear state of the terrain on which it is standing from the analysis of data collected by the sensor box. Figure 1 shows the block diagram of the sensor box. In the diagram, all the sensors are interfaced with a microcontroller for internal control. Since the cameras require post-processing, a powerful microcomputer with a small form factor should be used for processing the images, rather than using a microcomputer on board, the camera's data will be accessed through USB by the main computer controlling the robot. The microcontroller will run an ML algorithm that uses data captured for further analysis. In addition to the sensors enumerated previously, the sensor box features a laser diode, a white LED, and a servo motor. Each integrated sensor plays a specific role in understanding the navigated terrain. The cameras work as a stereo camera and take a live video of the terrain on which the robot stands and feeds it to the main computer, which analyses it, and detects the type of terrain using a pre-trained ML model.

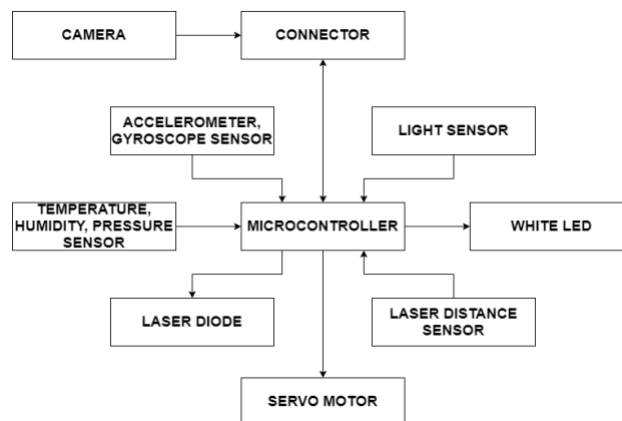


Figure 1. Sensor box block diagram

Upon detecting the type of terrain, the robot would be able to adjust its locomotion in real-time. Using cameras only would not be enough to get accurate data and provide an accurate analysis, as the images captured could be affected by the distance between the camera and the terrain, the light intensity while capturing the images, and the angle at which the images are captured. Hence, such parameters should be considered during the dataset collection itself to increase the terrain analysis accuracy that would take place in further processes. In order to detect whether the sensors are parallel to the terrain, an accelerometer and gyroscope sensor, which serve as an inertial measurement unit, are used. If the whole sensor box is not leveled due to the irregularity of the terrain, the microcontroller adjusts it through an in-built servo motor, while the inertial measurement unit records the unbalanced angles for adjustments. However, the adjustment angle will be limited to some point since there is a critical point over which the whole structure will lose stability and fall. A temperature, humidity, and pressure sensor are used to get the same data mentioned in the sensor name. Such parameters would affect the terrain and cause the terrain analysis algorithm to misbehave and output inaccurate results. Gathering these parameters during the dataset creation matters as they are used for tuning the terrain analysis algorithm to get more accurate predictions. A laser diode projects a red dot on the terrain as a reference point during image capture. A combination of inertial measurement unit and camera images with the red dot reference point determines the speed and direction of motion. The light intensity while capturing images plays a crucial role during the image-capturing phase. A low-light environment can alter the quality of the image captured and mislead the terrain analysis algorithm. The light intensity can be determined using a light sensor, which can be used to switch on an LED strip in case of low light detection. Another parameter influencing the terrain analysis output is the height at which the images are captured. This distance, known as the clearance from the ground, can be collected with the help of a laser distance sensor. The sensor box returns JSON data containing each sensor's value through the I2C bus. Finally, the collected data during the training session can be used to train an ML model for terrain data analysis. Figure 2 shows the image of the sensor box developed for that purpose.

Since the type of sensors used affects the data quality of data collected, it's important to mention the sensors used in the design of the sensor box.

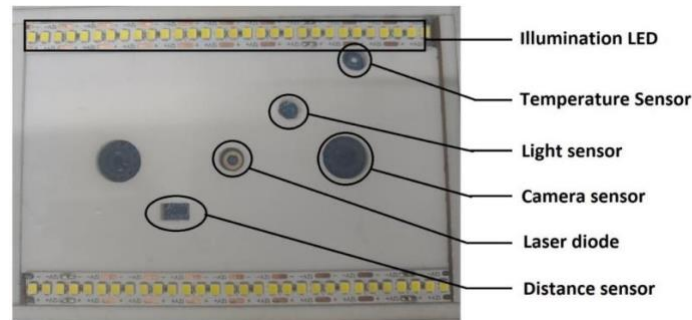


Figure 2. Photos of the actual sensor box

The sensor box is a compact device which has 2 RGB Logitech C270 cameras of 720 pixels, 30 frame per second spaced by 7 centimetres and inclined at 45 degrees inside the box. The cameras provide a stereo vision to the module for further depth perception of the navigable surface. The camera sensors are aided with a temperature and humidity sensor DHT22, a light intensity sensor BH1750 and a laser distance sensor VL53L0X. For accurate and quality data gathering, high end sensors can be used to replace each sensors, based upon the scenario. The proposed sensor box's specifications are consigned in Table 1.

Table 1. Sensor box's specifications

Number of cameras	2
Maximum resolution	720P/30FPS
Mega pixel	0.9
Field of view	55° (each)
Distance sensor	VL53L0X, 1mm resolution, 200cm maximum range
Light sensor	BH1750, 65535 lux maximum range
Temperature and humidity sensor	Temperature: 0.1°C./ $\pm 0.5^\circ\text{C}$ / $-40^\circ\text{C} \sim 80^\circ\text{C}$. Humidity: 0.1% RH./ $\pm 2\%$ RH (25°C) / 0% RH $\sim 99.9\%$ RH.
Lighting	5050 LEDs, 70000 lux maximum
Overall size	130x100x40 mm

b) System calibration

The device is used to gather data that will be used to train an ML model for classifying different terrains and conduct some more experiments. Therefore, ensuring that the sensor box gets accurate data is essential. The same procedure is used to calibrate the sensors integrated into the sensor box. Figure 3 shows some arrangements made for the calibration process of each integrated sensor. Most of the parameters linked with the sensors can be calibrated with a mobile phone since mobile phones use similar sensors. Using a mobile application called physics tools with the mobile phone configured to monitor the desired parameters and placed on top of the sensor box mounted on the carrying apparatus, the values gathered by the sensor box are collected and cross-checked with the values obtained from the mobile app. The following parameters: accelerometer, gyroscope, inclination, light intensity, and GPS coordinates are checked individually, and the values obtained for each parameter from both the sensor box and mobile phone are collected and plotted to find out the behavior of the data. The next step is the actual calibration, which uses a polynomial regression to get a polynomial function that satisfies the values from both devices. Figure 3 shows different views of the data collection apparatus with tools meant for calibration. Figure 3(a) shows the arrangement used for calibrating the ground clearance, using a measuring tape. Figure 3(b) shows the arrangement for the sensor box calibration and Figure 3(c) shows an inner view of the apparatus from which we can see the sensing part of the sensor box. Data gathered by the sensor for different ground clearances are compared with the actual ground clearance from the measuring tape. Similarly, a thermometer calibrates the temperature sensor following the same procedure. Finally, a polynomial function is fitted using data gathered to calibrate each sensor integrated into the sensor box. The fitted polynomials are used in the firmware of the sensor box to correct the raw data and

output calibrated values. Once each sensor is calibrated, a test is run to ensure that the values supplied by the sensor box are accurate enough to be trusted for subsequent operations.

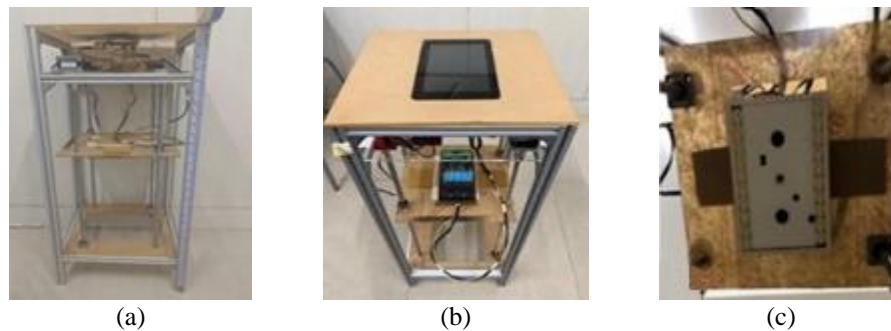


Figure 3. Images of the arrangements for the calibration of the sensor box: (a) arrangement for ground clearance calibration, (b) arrangement for temperature, humidity, accelerometer and gyroscope calibration, and (c) arrangement for depth image and light sensing calibration

2.1.2. Control structure design

The sensor box is a collection of sensors brought together to collect essential data that help robots understand their environments, especially the terrains on which they move. A sensor box should be mounted on a robot to enable such capabilities. Since the training phase does not require the sensor box to be installed on an actual robot, a structure should be built to help us collect the training dataset. The structure should be easily movable, provide the ability to set different heights for the sensor box to capture images, and facilitate data collection through an automated control system. The proposed apparatus has two parts: a mechanical frame that provides the necessary movements and a control system that automates the data-collecting process.

a) Mechanical frame design

The mechanical setup frame in Figure 4 is made from aluminium extrusion profiles, acrylic sheets, sliding and threaded rods and nuts, and motors. The apparatus height is 75 cm, with a surface of 25 cm by 25 cm. The top of the structure is mounted with a 5mm thick acrylic sheet to bear the control system, stepper motors, and sliding nuts. The moving part (Z-axis moving frame) is made of the same material of less surface, with sliding and threaded nuts to allow the frame to move up and down, considering the rotation of the stepper motors. The sensor box is mounted on the Z-axis moving frame for motion control. This movement is due to the accuracy of collecting data at different ground clearances. Indeed, at different heights, the results might get affected and lead to some errors in the interpretation.

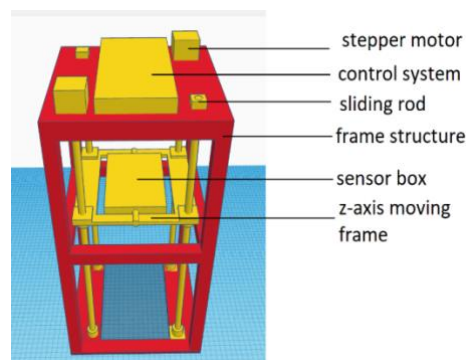


Figure 4. 3D rendering of the mechanical frame of the dataset collection apparatus

Allowing the same image to be taken at different heights helps to understand the effects of height on the collected data and therefore provides a solution. The direct application of such practice can be seen in robots, which do not keep the exact clearance from the ground while accomplishing specific tasks. This may be due to the irregularity of terrains, any obstacle to be avoided, or a pre-set user-defined ground clearance. It

is, therefore, essential to consider such parameters while collecting data for better accuracy. The whole structure can be carried easily and moved toward different locations for data collection.

b) The control system

The block diagram involving the components used in the control system is depicted in Figure 5. The control system facilitates automated data collection through physical switches or a remote application, employing an electronic circuit with an embedded computer to execute a data collection algorithm in Figure 6. Utilizing a mini-computer like the Raspberry Pi 4 allows for image capture and processing, enabling testing and refinement of machine-learning models based on collected data. The system features a touchscreen interface for user setup, with critical parameters including terrain location and image capture angles, essential for accurate model generation. The addition of a GPS module facilitates location coordinate recording. Actuation mechanisms, such as stepper motors, manage setup adjustments for data collection conditions, ensuring robust power supply for portable operation. The algorithm for data collection, depicted in a flowchart, iteratively captures images and environmental data, allowing for variations in conditions like ground clearance height and illumination status to enhance data sensitivity and accuracy. This iterative process enables comprehensive data capture for interpretation and analysis across different terrains and conditions.

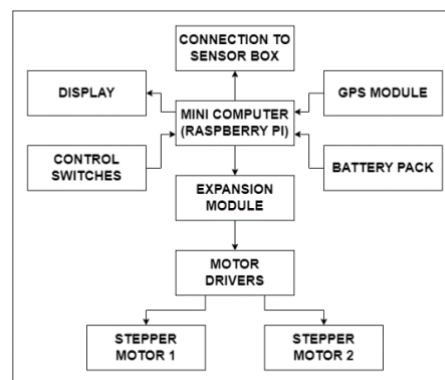


Figure 5. Control system of the dataset collection setup

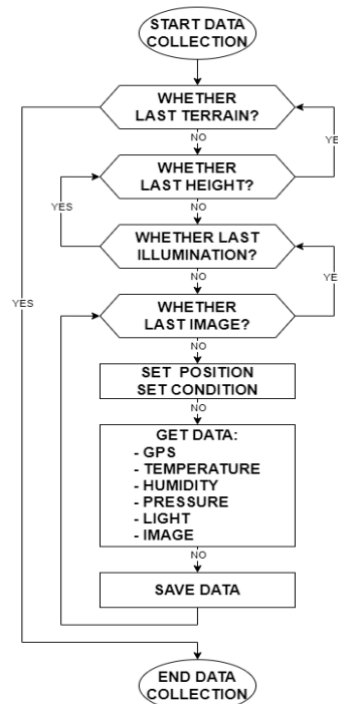


Figure 6. Flowchart for data collection algorithm on the control system

2.2. Data collection method

2.2.1. Identification of terrains

The proposed terrain data collection process aims to standardize and streamline the identification and classification of various terrains based on their distinct features and properties, with the ultimate goal of facilitating effective navigation for robots. Beginning with the meticulous selection of the working environment, thorough inspections are conducted to identify and catalog distinguishing features such as color, pattern, material composition, and overall appearance of the terrain surfaces. These meticulously noted features are then systematically utilized to classify terrains into classes and subclasses, taking into account their potential impact on robot locomotion and navigation. For instance, terrains may be grouped based on characteristics such as color consistency, texture variations, and surface irregularities. This comprehensive classification process not only aids in organizing the terrain data but also serves to highlight the nuances and complexities inherent in different terrain types. Subsequently, meticulous data collection planning is undertaken to estimate the requisite number of subclasses and images needed for each terrain category. This strategic planning phase not only ensures the systematic initialization of the terrain dataset but also lays the groundwork for effective and efficient data collection methodologies. Finally, the culmination of the process involves the automated execution of the meticulously devised data collection plan, with the primary objective of establishing robust correlations between terrain properties and their suitability for robot navigation. By systematically linking the identified terrain features with their corresponding navigation challenges and opportunities, this process, summarized in Figure 7, endeavors to enhance the overall navigation capabilities of robots operating in diverse environmental settings, by providing an efficient data collection method for dataset establishment and ML model training.

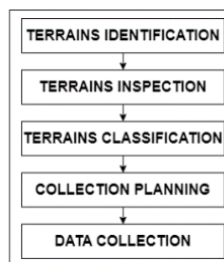


Figure 7. Terrain data collection process

2.2.2. Data collection phase

Once there is a ground for starting the actual terrain's data collection, the data collection apparatus can be configured to automate the data collection process, as depicted in Figure 8. Before starting any data collection, the terrain data collection process should be followed. Once the terrains are identified, inspected, classified, and the data collection execution is planned, the actual data collection phase can start. More information needs to be provided to the experimental setup before collecting data. The setup provides the ability to start a new collection or resume a collection. In case of a new collection, the name and description of the collection should be provided. This creates a dedicated folder in the system's memory for storing the data that will be collected. The folder's name is the same as the collection's name provided by the user. The folder contains an Excel file to store sensor data, a text file that holds the description of the collection and information related to it, and finally, a folder that stores all the stereo images captured by the cameras of the sensor box. When the user has to resume a previous collection, the existing folder gets opened, and data get stored in that folder. During a fresh new data collection process, information like the collection's name, a description, and the number of terrains that will be investigated should be provided to initialize the system. Upon providing such information, the labels for each terrain should be specified along with the collection parameters. Collection parameters specify the parameters to be considered while collecting field data. Such parameters are the date, time, location, clearance heights, illumination, temperature, pressure, and humidity. The system can be set to take the same image under different conditions involving height clearance and illumination. For more specific research, any of these parameters can be varied while others remain constant to study the effect of that particular parameter on the data collected under certain conditions. Once the initial setup is done, the actual data collection can start with respect to the algorithm depicted in Figure 6. Once data for all terrains is collected considering the initial setup provided, the output can be accessed in an Excel file. The file can be used further for subsequent analysis and ML algorithm training based on the goal to reach.

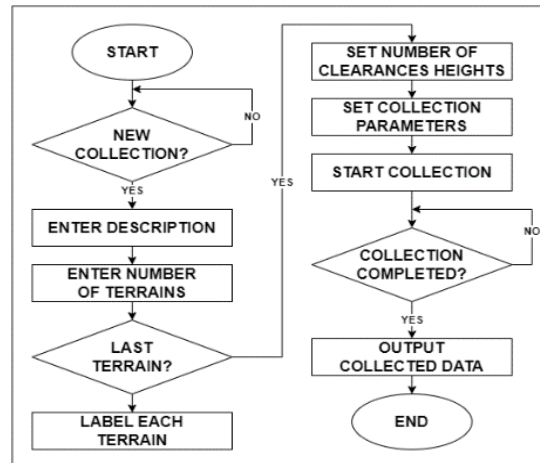


Figure 8. Data collection flowchart

2.3. Terrain data collection and hardware design validation experiments

2.3.1. Terrain data collection

The experiments are designed to validate the proposed hardware design by demonstrating the efficacy of the sensors in capturing relevant data that influences the classification outcomes of a conventional ML classifier. To achieve this objective, we will adhere to the prescribed terrain data collection process, beginning with the acquisition of data followed by training a ML classifier. Subsequently, we will evaluate any fluctuations in the similarity percentage as the data captured by the hardware sensors undergo variations. Any observed alterations in the similarity percentage would substantiate our hypothesis, thereby affirming the validity of the proposed design and methodology.

a) Terrain identification step

Considering the Terrain data collection process, the initial action to perform is to identify a place for conducting the experiments. For this study, the test environment selected is REVA University and its surroundings, located in Bangalore, India. Applying the procedure dictated by this stage helped in identifying the following terrains in the test environment:

- Solid, shiny flat tiles of different colors
- Solid, not shiny flat rock tiles with patterns
- Solid, rugged rock tiles
- Less practicable small-height flowers
- Different types of grass
- Different sizes of concrete tiles, flat concrete surfaces
- Soft irregular sandy ground, irregular soft sand, and stone ground
- Asphalt

b) Terrain inspection and classification steps

The next step in the process is easier from the observations and details mentioned during the terrain's identification step. It could be observed that the environment has different terrains with different properties. For example, some tiles are shiny, while others are not. All the tiles have different colors and patterns, which can be used for further classification. While certain tiles are flat and shiny, some others are rugged and not shiny. Some surfaces are just made of concrete with a pretty flat aspect. With such kinds of terrains come asphalts, which share the same properties with concrete surfaces from a certain point of view but can be easily distinguished by their colors and textures. Another group concerns sandy grounds, which can have different colors and a sufficient number of stones. Last but not least, we have grasses, which can be of different types but share the same green color most of the time. From these details, classifying the terrains becomes easy. However, the classification should be done based on the potential performances of robots on such terrains. For example, Flat, regular, hard, and not shiny terrains would be preferred over soft and irregular grounds for better locomotion performances. The proposed classification for this experiment is shown in Figure 9. As the classification is done, the collection planning is the next phase to work on.

c) Data collection step

Data collection can be done following the same order as in the classification. The test environment features many buildings with shiny and not shiny flat, and hard tiles. Generally, the same type of tiles is used in all the buildings. Hence, a single building is sufficient for collecting data of such types. Consequently, shiny

and not shiny flat and hard tiles data can be collected within a single building. On the other hand, not shiny, rugged, and hard tiles, as well as concrete surfaces and asphalts, are available outdoors at the entrances of different buildings. Such data can be collected once the first land data collection is completed. Finally, sandy and grassy lands can be investigated because it is available further away than previous lands. The main experiment is creating a dataset of terrains within REVA University and its surroundings with the help of the sensor box mounted on the apparatus. The dataset collection apparatus is initialized with seven terrains classes that can be further subdivided into subclasses based on specific features. All the parameters for data collection (date, time, illumination, GPS location, ground clearance, accelerometer, gyroscope, temperature, humidity, and light intensity) are enabled, and 300 images of each terrain are collected at a constant ground clearance, i.e., 40 cm, with illumination enabled all the time, providing an average of 50 lux in an indoor environment with artificial lights on. All the images are taken with no inclination in both x-axis and y-axis, with the sensor box being parallel to the target surface. The data collection experiment is conducted during the daytime to maximize the quality of images taken and ease the overall process. Figure 10 shows the data collection apparatus on different at 40 cm clearance during the data collection process once the setup has been initialized.

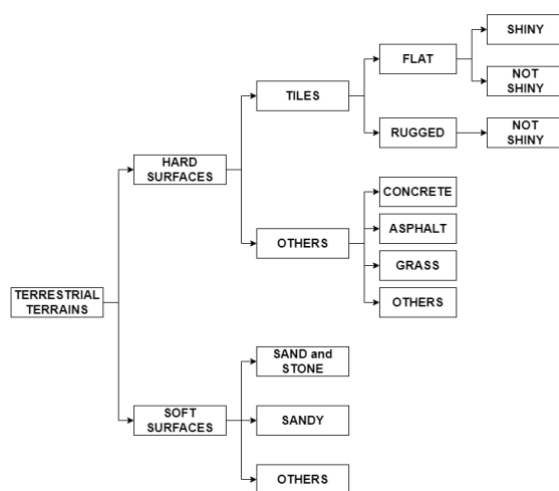


Figure 9. Classification of available terrains in the test environment



Figure 10. Apparatus at different locations on different terrains

The collected images are used to train a basic ML classifier that can classify terrain from images captured by the cameras. Data augmentation is enabled to increase the overall number of images and thereby the prediction accuracy. Few network models have been tested against a single test image to determine the best model to be use for training. By comparing the similarity percentage of different models after being tested with the same test image, the efficientNet-V2-bo-21k provided the highest accuracy percentage. The training is done with a batch size of 16, keeping 50 epochs with 6 steps per epoch.

2.3.2. Investigation of the correlation between sensor data and similarity percentage

To validate the data collection process and method, a single terrain—hard, flat, shiny tile—was selected for further testing. A few images of this terrain were collected, keeping some parameters constant while varying a single parameter to investigate its impact on the ML classifier's accuracy. These experiments involved recording how accurately the ML classifier performed under different conditions, establishing a relationship between the varying parameter and prediction accuracy through the similarity percentages obtained from testing different images. In the next sections, we will study the effects of shooting angle, ground clearance, and ambient light on similarity percentage.

a) Effects of shooting angle on similarity percentage

The first parameter to investigate is the shooting angle or angle at which the images are taken with respect to the surface. The apparatus is inclined first towards one axis with the ground clearance kept constant (40 cm) while recording the similarity percentage of the ML classifier at different angles for the same terrain. The same experiment is repeated for the second axis to get the similarity percentage or detection accuracy percentage variation along that axis. Nine images are collected for each axis, keeping a quasi-constant angle variation that varies between 0 to 25 degrees. An illustration of the experiment is given in Figure 11.

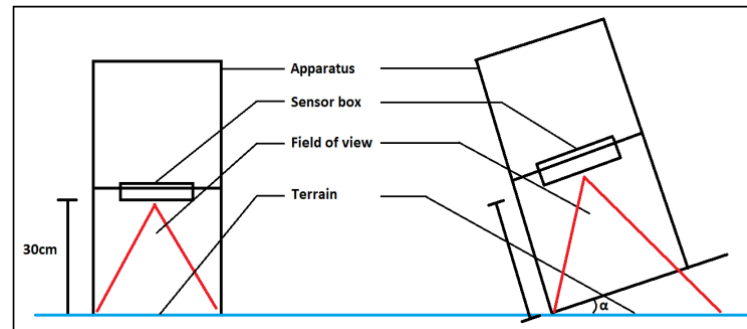


Figure 11. Detection accuracy percentage variation against shooting angle experiment illustration

b) Effects of ground clearance variations on similarity percentage

The second parameter to study is ground clearance and its implications on ML classifier accuracy. The experiment establishes a link between ground clearance level and terrain identification accuracy by taking images of the same surface with different ground clearance levels while maintaining the equipment in the same place. The ground clearance is adjusted in 2 cm increments from 16 cm to 54 cm, and photos of the selected terrain are collected each time. A total of 21 images are taken by the sensor box parallel to the terrain with no inclination at constant lighting of 50 lux. Figure 12 depicts an example of the experiment.

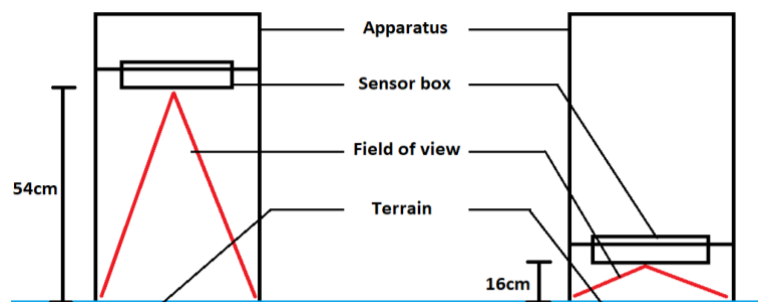


Figure 12. Detection accuracy percentage variation against various ground clearance levels

c) Effects of ambient light on similarity percentage

The last experiment concerns the amount of light available near the sensors while taking a capture of the same terrain. The apparatus is covered with an opaque matter up to some level to allow a certain amount of ambient light near the sensor while taking a photo of the terrain at a constant ground clearance level (40cm). The amount of light allowed around the sensors is kept constant while the illumination LEDs are controlled to provide more or less light in the area of interest. A total of 7 images of the same terrain are captured varying the light intensity from 15 lux to 50 lux. The captured images are later tested with the pre-trained model to get the detection accuracy percentage of each image subjected to different lighting conditions. It is important to note that the experiments related to the parameter's investigation rely on the terrain classification model obtained from the main experiment. Also, the number of data gathered to establish the relationships depends on each experiment. In each case, the maximum amount of data is taken to maximize the experiments outcomes.

3. RESULTS AND DISCUSSIONS

3.1. Results

The objective of the conducted experiments is to demonstrate the effectiveness of the proposed dataset collection apparatus in acquiring meaningful data pertaining to diverse terrain types within a specified environment. While existing environment perception techniques often employ sensor fusion to enhance accuracy, they typically focus on overall surroundings rather than specific terrain properties and their impact on robotic performance. The proposed approach seeks to investigate the influence of distinct parameters on the detection accuracy of an image classifier trained for terrain detection and identification. This process involves initial data collection, followed by classifier training, and subsequent analysis of parameter effects. Data collection entails capturing numerous images of similar terrains alongside sensor data such as GPS coordinates,

accelerometer and gyroscope readings, temperature, humidity, pressure, and light intensity, all stored in an Excel database for further analysis and classifier training. Each terrain-specific dataset is segregated within the Excel file, with separate sheets created for different ground clearances if required. The prediction accuracy achieved using the efficientnet-V2-bo-21k network after training is 98%. Experimental results detailing the relationship between model prediction accuracy and various parameters are depicted graphically in Figures 13 to 16, with outcomes contingent on experimental conditions. The reported parameter values reflect accuracy percentages derived from testing a single terrain image under diverse conditions.

Figures 13 and 14 illustrate the correlation between the shooting angle, represented by two axes, and the terrain detection percentage. Another critical factor influencing detection accuracy is the distance between the sensor and the terrain. Illustrated in Figure 15, the detection accuracy percentage of terrain is depicted against the ground clearance height, which refers to the distance between the robot's chassis and the surface it traverses. Increasing the distance between the camera sensors and the surface enhances the field of view but diminishes the capability to capture intricate details. These parameters play a pivotal role in terrain detection based on the clearance set during the collection of the training dataset. Therefore, accounting for ground clearance during dataset collection facilitates establishing a relationship with test data gathered at various ground clearance levels, thereby mitigating potential losses resulting from clearance variation and augmenting detection accuracy.

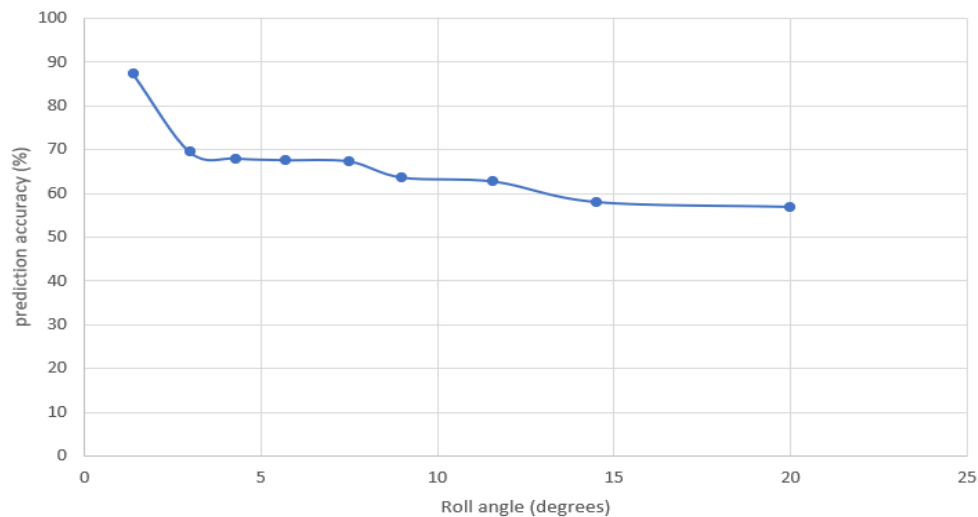


Figure 13. Plot of percentage accuracy of the ML classifier against roll inclination angle

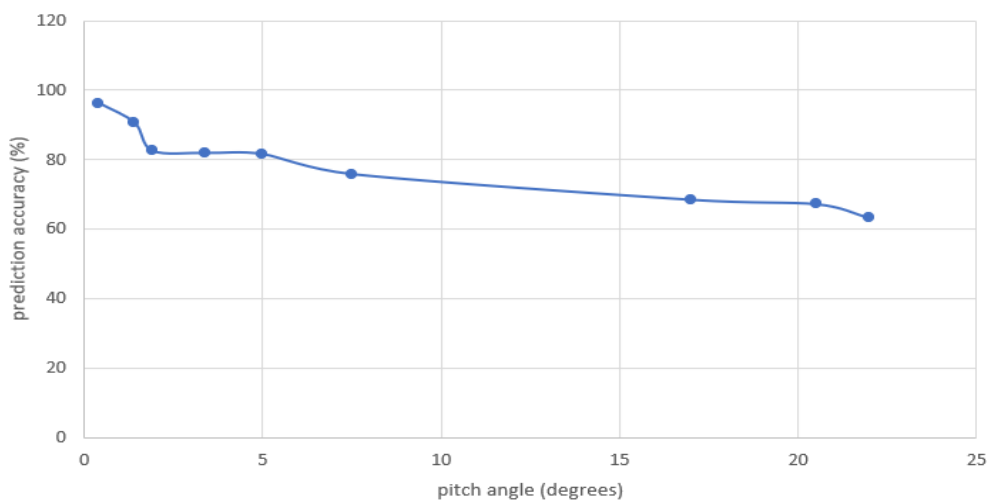


Figure 14. Plot of percentage accuracy of the ML classifier against pitch inclination angle

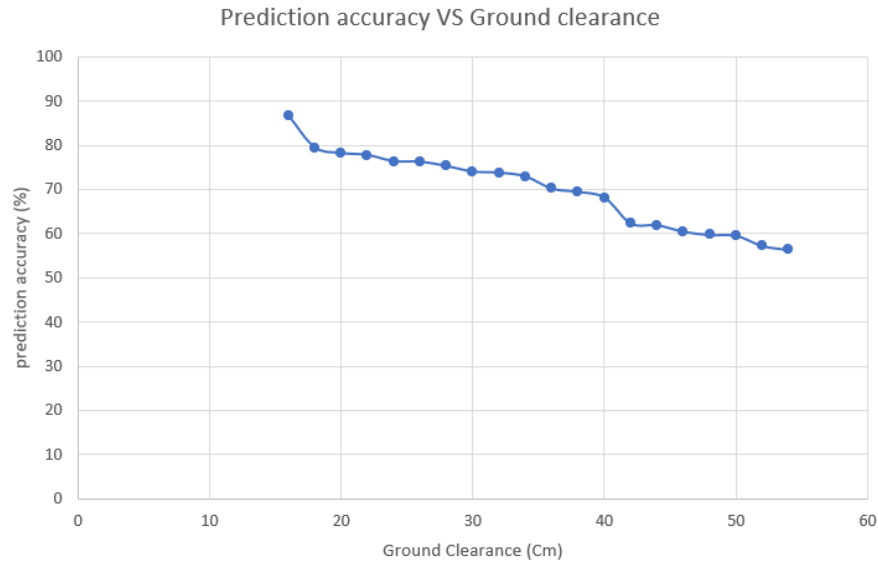


Figure 15. Plot of percentage accuracy of the ML classifier against multiple ground clearances

The experiments also examined the effect of light intensity surrounding the sensor box. Since the sensor box is positioned at the bottom of the robot, with its sensors directed downward parallel to the surface being navigated, the amount of available light during terrain image capture is likely to be limited. It is crucial to assess how varying light conditions affect terrain detection using the available sensors. Figure 16 illustrates the ML classifier's accuracy percentage relative to the surrounding light intensity. The graph demonstrates a notable rise in detection accuracy as the light intensity increases. This observation highlights the significant impact of lighting conditions on terrain detection via ML models.

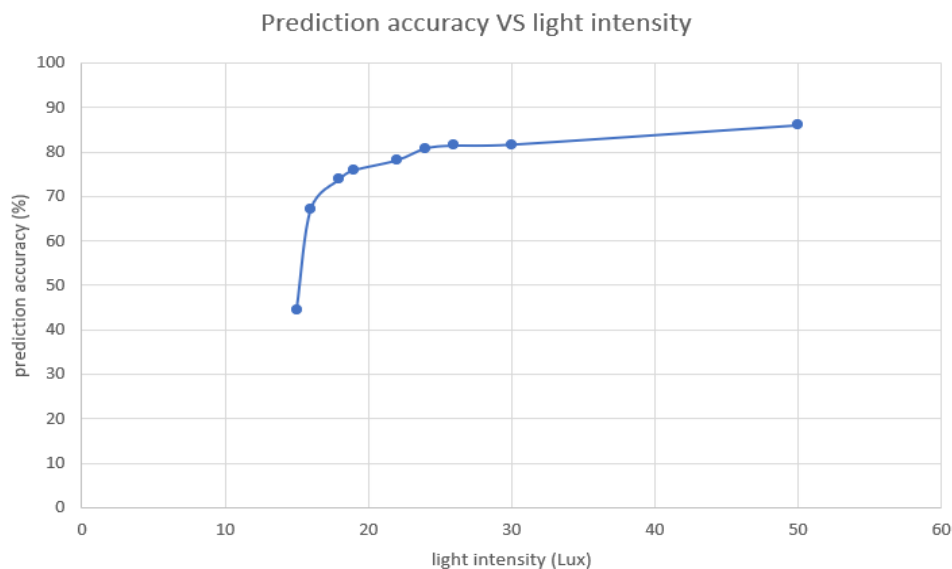


Figure 16. Plot of percentage accuracy of the ML classifier against the surrounding light intensity

3.2. Discussions

Based on the findings outlined in the preceding section, we can delve into the effects of the analyzed parameters on the performance of a conventional ML image classifier. Given the mobile nature of the robot, it navigates diverse and irregular terrains, resulting in variations in shooting angles dictated by the terrain's irregularities. Consequently, the sensor box may not consistently align parallel to the terrain, potentially influencing detection accuracy. This influence is evident in Figures 13 to 14, illustrating a decrease in detection

accuracy percentage with changes in shooting angles. Initially, the similarity percentage is at its peak at a ground clearance of 40 cm and no inclination. However, as inclination angles in both roll and pitch axes are altered, there is a noticeable decline in the similarity percentage for the chosen terrain in the experiment. This decline is attributed to conditions during dataset collections, including fixed ground clearance, inclination, light intensity, humidity, and temperature. The subsequent parameter examined was ground clearance, aimed at discerning its impact on the pre-trained ML image classifier's performance. Proximity to the terrain alters the field of view, potentially affecting detection accuracy. While reduced depth focuses on surface details, increasing depth widens the field of view, reducing accuracy. Notably, performance degradation is observed from 30 cm to 54 cm, indicating the critical role of ground clearance in detection accuracy. However, while the similarity percentage diminishes with rising ground clearance, the decline gradually underscores the importance of ground clearance in training and beyond. Lastly, the investigation into light intensity or ambient light revealed its significant influence on camera detection ability and subsequent results, particularly concerning images. Figure 16 graphically depicts the experiment's outcome, indicating an expected rise in detection accuracy with increased light intensity. Excessive light may negatively impact classifier performance, depending on training data conditions.

The outcomes of implementing the proposed terrain dataset collection process demonstrate its effectiveness in identifying, classifying, and compiling terrain data within a designated environment. The resultant Excel file, generated after the actual data collection experiment, serves as concrete evidence of the reliability and efficiency of the terrain dataset collection apparatus employed in the trials. These findings underscore the indispensable role of integrated sensors in crafting the data collection apparatus, with each parameter linked to the sensors utilized significantly influencing the detection accuracy of the ML terrain classification model. The significance of these experiments is delineated by two primary components: the sensor box and the terrain dataset collection setup. The sensor box, housing an array of sensors for capturing essential data such as terrain photos, ambient temperature, humidity, light intensity, and ground clearance, forms the cornerstone for acquiring terrain data for subsequent analysis. Each sensor within the box fulfills a pivotal function in terrain perception and analysis, with the accelerometer and gyroscope aiding in maintaining the sensor box's alignment parallel to the ground and monitoring image capture angles. Furthermore, incorporated LEDs facilitate clear imaging in low-light conditions. At the same time, the humidity sensor offers insights into terrain moisture levels, thereby enabling a deeper analysis of a robot's locomotion performance under varied conditions. On the other hand, the dataset collection setup simulates ground clearance from the robot's base while simultaneously recording GPS location. Image clarity and detail size are contingent upon the distance between the terrain and sensors and ambient light intensity, necessitating data collection at diverse clearances and illumination levels to establish correlations that enhance terrain identification accuracy. Additionally, the setup incorporates independent circuitry for data capture and storage, integrating a GPS module to provide location data for each captured image, thereby enabling effective terrain identification based on GPS coordinates. However, it is essential to acknowledge that the outcomes are subject to the type of hardware and sensors employed. The quality of sensors utilized may either enhance or diminish the results. As specific parameters have been identified as critical in terrain detection and classification, future work would involve proposing a model that leverages the effects of these parameters to bolster the detection accuracy of the ML classifier.

4. CONCLUSION

The current trend in computer science technology is artificial intelligence and ML, which gives rise to an infinite possibility of developing specific applications. ML relies on a dataset used to train an ML algorithm, which later gives a model that can be used to get insights about data similar to the training dataset. This paper uses the same approach for a dataset collection process for a suitable environment perception through terrain perception. The authors proposed a methodology and method for terrain dataset collection that can be used further to train a ML algorithm and get a working terrain analysis model. The terrain analysis model's accuracy depends on the data quality used for training the ML algorithm. The dataset collection apparatus proposed by the authors gathers data concerning different parameters that impact a standard ML model. Such parameters include ground clearance, humidity, light intensity, and shooting angle. The investigation of the impact of such parameters on the detection accuracy of terrains through the conducted experiments revealed that each sensor integrated into the design is worth it. Hence, the type of data generated by the data collection apparatus helps improve the accuracy percentage of ML classifier model. The end goal of the proposed methodology is to focus on the land on which a robot stands in real-time and get insights about the land's conditions to adjust the locomotion the robot uses to move rather than focusing on the surroundings. It's important to note that some more sensors can be added to sensor box to get more data thereby more insight into the terrain on which a robot stands and move. In the future, the collected data will be used to create first

mathematical model that correlate the investigated parameters with a classical ML image classifier for improving its performance, and further a terrain perception model, which could help a robot understand the terrain on which the robot stands in real-time and later give inputs to a motion planner that modifies the dynamic and locomotion to be adopted in real time.

ACKNOWLEDGEMENTS




This work was supported and funded by REVA University, Bangalore, INDIA under the University seed funding with the reference RU:EST:EC:2022/41 granted on 04-03-2022.

REFERENCES




- [1] Y. Wu *et al.*, "An object SLAM framework for association, mapping, and high-level tasks," *IEEE Transactions on Robotics*, vol. 39, no. 4, pp. 2912–2932, 2023, doi: 10.1109/TRO.2023.3273180.
- [2] Z. Nan *et al.*, "A joint object detection and semantic segmentation model with cross-attention and inner-attention mechanisms," *Neurocomputing*, vol. 463, pp. 212–225, 2021, doi: 10.1016/j.neucom.2021.08.031.
- [3] D. T. -F. -Betoño, E. Zulueta, A. S. -Chica, U. F. -Gamiz, and A. S. -Aguirre, "Semantic segmentation to develop an indoor navigation system for an autonomous mobile robot," *Mathematics*, vol. 8, no. 5, 2020, doi: 10.3390/MATH8050855.
- [4] S. Se, D. G. Lowe, and J. J. Little, "Vision-based global localization and mapping for mobile robots," *IEEE Transactions on Robotics*, vol. 21, no. 3, pp. 364–375, 2005, doi: 10.1109/TRO.2004.839228.
- [5] S. H. Joo *et al.*, "Autonomous navigation framework for intelligent robots based on a semantic environment modeling," *Applied Sciences*, vol. 10, no. 9, 2020, doi: 10.3390/app10093219.
- [6] J. Zhang, R. Liu, K. Yin, Z. Wang, M. Gui, and S. Chen, "Intelligent collaborative localization among air-ground robots for industrial environment perception," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9673–9681, 2019, doi: 10.1109/TIE.2018.2880727.
- [7] C. Ginerica, M. Zaha, L. Floroian, D. Cojocaru, and S. Grigorescu, "A vision dynamics learning approach to robotic navigation in unstructured environments," *Robotics*, vol. 13, no. 1, 2024, doi: 10.3390/robotics13010015.
- [8] A. Singh, R. Narula, H. A. Rashwan, M. A. -Nasser, D. Puig, and G. C. Nandi, "Efficient deep learning-based semantic mapping approach using monocular vision for resource-limited mobile robots," *Neural Computing and Applications*, vol. 34, no. 18, pp. 15617–15631, 2022, doi: 10.1007/s00521-022-07273-7.
- [9] R. M. Bena, C. Zhao, and Q. Nguyen, "Safety-aware perception for autonomous collision avoidance in dynamic environments," *IEEE Robotics and Automation Letters*, vol. 8, no. 12, pp. 7962–7969, 2023, doi: 10.1109/LRA.2023.3322345.
- [10] S. Sultana, B. Ahmed, M. Paul, M. R. Islam, and S. Ahmad, "Vision-based robust lane detection and tracking in challenging conditions," *IEEE Access*, vol. 11, pp. 67938–67955, 2023, doi: 10.1109/ACCESS.2023.3292128.
- [11] B. Teixeira, H. Silva, A. Matos, and E. Silva, "Deep learning for underwater visual odometry estimation," *IEEE Access*, vol. 8, pp. 44687–44701, 2020, doi: 10.1109/ACCESS.2020.2978406.
- [12] A. Bekiariski, "Visual mobile robots perception for motion control," *Intelligent Systems Reference Library*, vol. 29, pp. 173–209, 2012, doi: 10.1007/978-3-642-24693-7_7.
- [13] S. Kowalewski, A. L. Maurin, and J. C. Andersen, "Semantic mapping and object detection for indoor mobile robots," *IOP Conference Series: Materials Science and Engineering*, vol. 517, no. 1, 2019, doi: 10.1088/1757-899X/517/1/012012.
- [14] T. Ran, L. Yuan, and J. B. Zhang, "Scene perception based visual navigation of mobile robot in indoor environment," *ISA Transactions*, vol. 109, pp. 389–400, 2021, doi: 10.1016/j.isatra.2020.10.023.
- [15] H. Ge, T. Wang, Y. Zhang, and S. Zhu, "An object localization system using monocular camera and two-axis-controlled laser ranging sensor for mobile robot," *IEEE Access*, vol. 9, pp. 79214–79224, 2021, doi: 10.1109/ACCESS.2021.3084153.
- [16] L. Xia, D. Meng, J. Zhang, D. Zhang, and Z. Hu, "Visual-inertial simultaneous localization and mapping: Dynamically fused point-line feature extraction and engineered robotic applications," *IEEE Transactions on Instrumentation and Measurement*, 2022, doi: 10.1109/TIM.2022.3198724.
- [17] W. Xie, P. X. Liu, and M. Zheng, "Moving object segmentation and detection for robust RGBD-SLAM in dynamic environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 2021, doi: 10.1109/TIM.2020.3026803.
- [18] H. Surmann, C. Jestel, R. Marchel, F. Musberg, H. Elhadj, and M. Ardani, "Deep reinforcement learning for real autonomous mobile robot navigation in indoor environments," *arXiv-Computer Science*, pp. 1–7, 2020, doi: 10.48550/arXiv.2005.13857.
- [19] H. Guo, Z. Huang, Q. Ho, M. Ang, and D. Rus, "Autonomous navigation in dynamic environments with multi-modal perception uncertainties," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2021-May, pp. 9255–9261, 2021, doi: 10.1109/ICRA48506.2021.9561965.
- [20] M. Luo, "Multi-sensor based strategy learning with deep reinforcement learning for unmanned ground vehicle," *International Journal of Intelligent Networks*, vol. 4, pp. 325–336, 2023, doi: 10.1016/j.ijin.2023.11.003.
- [21] X. Huang, H. Deng, W. Zhang, R. Song, and Y. Li, "Towards multi-modal perception-based navigation: a deep reinforcement learning method," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4986–4993, 2021, doi: 10.1109/LRA.2021.3064461.
- [22] A. Nguyen, N. Nguyen, K. Tran, E. Tjiputra, and Q. D. Tran, "Autonomous navigation in complex environments with deep multimodal fusion network," *IEEE International Conference on Intelligent Robots and Systems*, pp. 5824–5830, 2020, doi: 10.1109/IROS45743.2020.9341494.
- [23] D. Feng *et al.*, "Deep multi-modal object detection and semantic segmentation for autonomous driving: datasets, methods, and challenges," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1341–1360, 2021, doi: 10.1109/TITS.2020.2972974.
- [24] R. Braud, A. Giagkos, P. Shaw, M. Lee, and Q. Shen, "Robot multimodal object perception and recognition: synthetic maturation of sensorimotor learning in embodied systems," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 2, pp. 416–428, 2021, doi: 10.1109/TCDS.2020.2965985.
- [25] Y. Yue *et al.*, "Day and night collaborative dynamic mapping in unstructured environment based on multimodal sensors," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 2981–2987, 2020, doi: 10.1109/ICRA40945.2020.9197072.

BIOGRAPHIES OF AUTHOR






Kouame Yann Olivier Akansie    obtained the B.E. (2018) degree in Electrical and Electronics Engineering from East Point College of Engineering and Technology, India. He further obtained the M.Tech. (2020) degree in Digital Communications and Networking from REVA University, India. He is currently a Ph.D. scholar at REVA University, India, within the school of Electronics and Communication Engineering. His thesis is about the design and development of autonomous hybrid wheel-legged robot for terrestrial navigation. He can be contacted at email: akansiekouame@gmail.com.






Dr. Rajashekhar C. Biradar    completed his B.E. (Electronics and Communication Engineering) in 1990, M.E. (Digital Electronics) in 1997 from Karnataka University Dharwad, India, and Ph.D. in 2011 under Visvesvaraya Technological University (VTU), Belgaum, India. He is currently working as professor and director of school of ECE, REVA university, Bangalore, India. To his credit, he has many national/international journal and conference publications. His research areas include multicast routing in mobile ad hoc networks, wireless Internet, group communication in MANETs, agent technology. He is a member of IEEE (USA), member of IETE (MIETE, India), member of ISTE (MISTE, India), member of IE (MIE, India) and member of ACM. He can be contacted at email: rcbiradar@reva.edu.in.



Dr. Karthik Rajendra    received M.Tech. degree from the Visvesvaraya Technological University, India and Ph.D. degree from VIT University, India. His Ph.D. thesis research work was carried out at one of the labs of Center for Nano electronics, Indian Institute of Technology – Bombay, India. He was the Professor at Department of Electronics and Communication and Dean R&D at MLR Institute of Technology, Hyderabad. Earlier, he was working as a faculty member at VIT University, Vellore. He received best researcher award from VIT University for his contribution to Nano dielectrics in 2013 and 2014. Also, he received best researcher award in 2017 to 2019 at MLR Institute of Technology. His current area of research includes fabrication and modeling of nano electronic or optoelectronic material-based devices, microwave antennas, medical image processing, and transformation in engineering education. At present, he is guiding 1 PhD research scholar. He has completed 5 Sponsored research projects worth Rs. 44 Lakhs. At present he has 2 ongoing research projects funded by DST, Govt. of India worth Rs. 3 Crores. He has published more than 110 research papers in reputed journals and conferences, 4 book chapters and filed 3 patents. He is one of the co-designers for developing a nano-size high performance capacitor in 2013 and 2020. He can be contacted at email: karthik.rajendra@reva.edu.in.



Dr. Geetha D. Devanagavi    received her Ph.D., M.Tech., and B.E. degrees in 2014, 2005, and 1993, respectively. She is currently working as associate Professor in the school of Computing and Information Technology at Reva University, Bangalore, India. She has 24 years of teaching experience. Her research interests include wireless sensor networks, network security, and computer networks. She can be contacted at email: dgeetha@reva.edu.in.