ARTIFICIAL INTELLIGENCE USING IN VISUAL SLAM METHODS – A REVIEW

Rajay Vedaraj I.S[1], Jino.S.Ganesh[2], Niranjen swarup[3], Madhan Kumar[4], Raviteja Tirumalapudi[5]
[1]Associate Professor, School of Mechanical Engineering, Vellore Institute of Technology,
[1]rajay@vit.ac.in.
[2]PG Scholar, School of Mechanical Engineering, Vellore Institute of Technology,
[2]jinoganesh00@gmail.com.
[3]PG Scholar, School of Mechanical Engineering, Vellore Institute of Technology,
[3]swarupnranjen97@gmail.com.
[4]PG Scholar, School of Mechanical Engineering, Vellore Institute of Technology
[4]madhansmith13@gmail.com.
[5]Research Scholar, School of Mechanical Engineering, Vellore Institute of Technology,
[5]tirumala.raviteja@vit.ac.in

Abstract

SLAM (Simultaneously Localization and Mapping) is widely used in many ways; they are used in robots, autonomous cars, mobile, underwater robots, and unmanned aerial vehicles, etc. a lot of applications use slam technology. Visual SLAM is a solution for the ground robot, and another type of robot for the localization and mapping was in this field for the past decade lot research. In the paper, we go to see the visual SLAM method direct and indirect way with progress how deep learning used to solve the dynamic problem in the environment mapping techniques which are applied to VSLAM are reviewed. Then finally, AI will be the developed which can affect similar human.

Keywords: *Deep learning, Mobile robots, Visual SLAM*

1 Introduction:

Before SLAM, we need navigation for a robot, the autonomous car when they need to reach a destination, so they need to map the path and plan the direction way to reach their destination; apart from mapping the robot first need to know their location and pose of the robot. [5] Now need to solve three problem localization, mapping, and path planning. Localization is to estimate where our robot is located in which direction and pose in the environment, which is determined by the sensors. Mapping is used for generating a replicated the surrounding environment in 3d model or 2d model for robot and where the surrounding is mapped. Still, the robot needs to plan for the path and choose the optimized path for the robot. At the same time, these problems are studied separately until the IEEE robotics and automation conference in 1986 were proposed a concept of SLAM (Simultaneously Localization and Mapping). SLAM can solve the problem that a robot can recognize its location in the unknown environment and form that point itself gradually to construct a continuous map of the environment. These done by various sensor devices here. We consider an external sensor camera, then the concept is known as visual SLAM.

Visual SLAM framework consists of the four models:

- Frontend visual odometry
- Back-end optimization

- Loop closure detection and

- Mapping.

Visual odometry is responsible for elementary estimating the robot location of the robot and their pose where the robot is the position of the map point from the frame to frame for every time position change. Then the place keeps update in the map point. The back-end optimization is responsible for the robot receiving the pose and direction these are updated by checking the position with the landmark measure the distance when the robot moves with distance in the path then using visual the robot understands the distance traveled this is done by the back-end optimization operation which measures by visual odometry and maximum a posteriori estimation. The robot should know the place mapping with a loop close mean when the same frame is updated; it needs to see this place was pre visited so for the purpose loop closure detection at recognizing which help the mapping and registered algorithm to obtain a more accurate and consistent result. Finally, the mapping is responsible for the regenerating of the map according to the camera pose and reference frame.

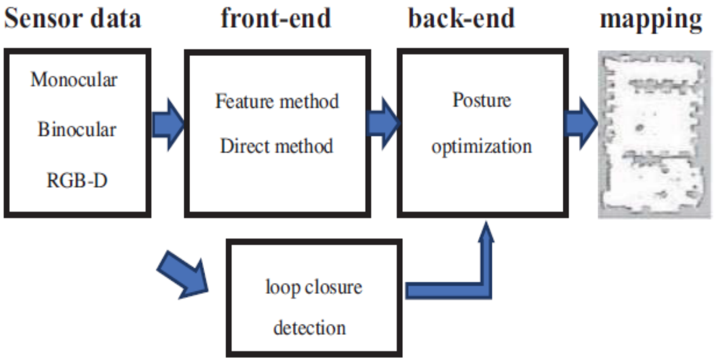

Fig-1 Framework of visual SLAM

The application of visual SLAM is when the system is not enough with the other sensor unit, and then the visual system is used. What special about the visual view is it consists of more information about the environment; for example, if a proximity sensor is used, it will measure the distance, acceleration, position. Still, these are not sustainable in real-time this sensor information can be assumed as the person is blindfold to walk; then, the condition makes it very hard to decide active status. The same state takes place in the robot when the robot is fixed static robot in industry or pick and place where it can use another sensor to know the system's feedback, but when the robot is dynamic, and environment can't be mapped and find its location. Where visual SLAM is will be used to know the environment. When the changes are occurring in the background. So then after the robot need to update so using the camera capture video the image frame per frame, the is used to calculate the distance the robot has moved [7]

Many visual SLAM systems have disadvantaged the system has failed to the external environment, in dynamic environment consist of too many or few salient features when image have extensive scale environment data takes long time process and during unpredictable movements of the camera where partial or if total obstruction of the sensor occurs.

As animal and human cases navigate in the environment is done by many factors where we can understand every object, and we remember path by the features, landmark, many number factors where we replicated in robot [22]. AI is the most research development field technology in many fields where the camera using image processing. The deep learning is used for the path planning, they took the concept from human to how humans avoid obstacles and path planning when they find the obstacle, and with optimal path distance, they reach the destination [31].

2 Geometry Based Vısual Slam

a) Feature-based Method

Theory the visual SLAM based on the geometric approach. Where the extract the features of the object from the image capture from the camera the image processing where optimizes the result using the filters where different ways of managing the information that solves the VSLAM problems most common techniques are EKF (Extended Kalman Filter) this method is the improved version of the classic KF (Kalman filter) this model for the nonlinear systems by mean of linearization process. The unknown the environment which is 3D structure is where the data to obtain by the camera simultaneously detections of the 3D structure and motion of the robot where the location is placed estimation of the pose with the landmark and estimate the distance the pixel measure concerning the time frame of the video of the image where the input video image frame by frame the travel distance by the robot the estimated by the geometry of 6-degree -of -freedom (DoF) motion of the camera and 3D position of the feature motion is same and based method distance formula is used by the initial to final frame time. Where to map in a closed environment is easy compared to the dynamic environment that is a massive problem with the SLAM and requires significant computation with an increase of the broad environment due to the increment of feature points. It is difficult to achieve real-time calculation. [17-20] for solving the MonoSLAM, PTAM heavy computation is divided into two categories: one is taking and other mappings two are executed in parallel processing, so this method saves the computational cost.

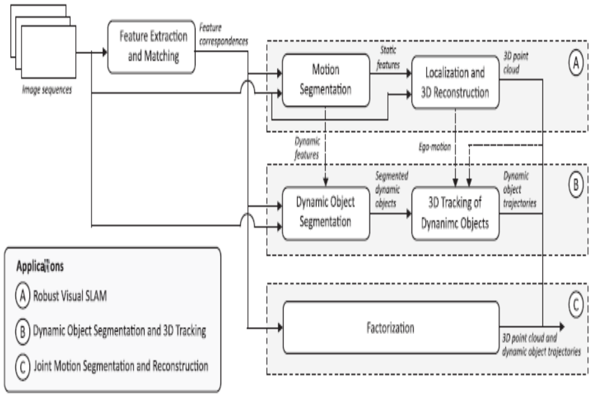
Image sequences
Feature Extraction and Matching
Feature correspondences
Motion Segmentation
Static features
Localization and 3D Reconstruction
3D point cloud
A
Dynamic features
Ego-motion
Dynamic Object Segmentation
Segmented dynamic objects
3D Tracking of Dynanimc Objects
Dynamic object trajectories
B
Factorization
3D point cloud and dynamic object trajectories
C
Applications
A Robust Visual SLAM
B Dynamic Object Segmentation and 3D Tracking
C Joint Motion Segmentation and Reconstruction

Fig 2 Feature extraction from the image sequences.

3 different type ai slam using deep learning:

DeepLearning is a sturdy branch of machine learning. This deep learning is a technique which teaches the computer what to do like the nature of human, i.e., learning by its mistakes. Deep learning is one of the primary technologies used in the driverless cars for detecting the road signs, for automatic parking, or for distinguishing pedestrians from the lamp post. In the recent century, consumer devices like phones, tablets, televisions, hand-free speakers like Alexa, Google smart speakers also use the techniques of Deep Learning.

A computer system performs classification tasks directly from the images like Google lens, texts, or sounds using Deep Learning. Deep Learning techniques using technology can achieve state of the art accuracy, which even can exceed human can perform in his level of performance. These technologies are trained by using large sets of labeled data and neural network architecture, which contains multiple layers. There are various neural networks and deep learning models used for a variety of complicated tasks:

a) Ann (artificial neural networks) for regression andclassification:

Artificial neural Network are the computing systems which was inspired by the biological neural networks that learn to perform tasks by example and mistakes, generally without being programmed with specific task rules. It is based on a collection of connected units or nodes, which are known as artificial neurons. An artificial neuron receives the signal and processes it and can send signal neurons connected to it [31] the paper represents SLAM online navigation based on Artificial Neural Networks. This ANN can solve the optimization problem of the super nonlinear equation. Thus ANN can be used for solving the nonlinear problems of magnetic field positioning. It stimulates the human brain and nervous system from the perspective of information processing, establishes a mathematical model, and consists of many complex interconnected neurons. In this study [31], three hidden layers of MLP in combination with the supervised learning method called the error backpropagation algorithm. The architecture of MLP ANN is shown in fig.

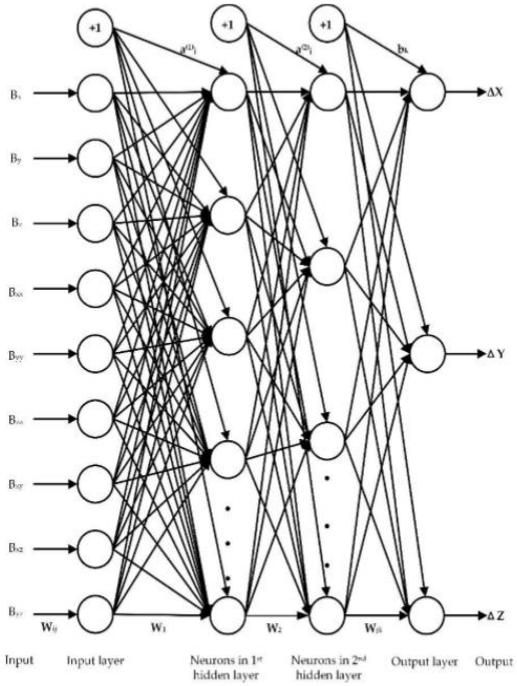

Fig-3 Schematic diagram of the proposed Multilayer Perception Artificial Neural Network (MLP ANN) model.

i) Optimization OfMlp-Ann:

MLP-ANN are optimized by the unsupervised optimization method (or learning method) which is the backward error propagation (BP) algorithm. This back propagation of MLP-ANN requires three processes which are

1. Data pre-processing

2. Training data

3. Validating data

In this paper, they used the hyperbolic tangent activation function. The sample data for the neural networks are taken from the magnetic field data and AUV of the magnetic beacon were connected in the field and the positions of the beacon. Thus, the sample data is obtained by the data processing process.After the collection of the sample data set, the inputs are taken from the magnitude of the magnetic field and the magnetic beacons relative positions and Autonomous underwater vehicle, which is used for the training process involves the optimization of the neural weights by adjusting the internal parameters of MLP-ANN which are used for the optimal global solution for the highly nonlinear objective equations and for minimizing the value of the error function. The no. of connection weights and the no. of hidden layers and the neurons in each hidden layer are established in the architecture of MLP-ANN. These hyper-parameters are typically determined through a trial and error method by comparison with the performance of different types of MLP ANN architecture, as shown in the Figure below.

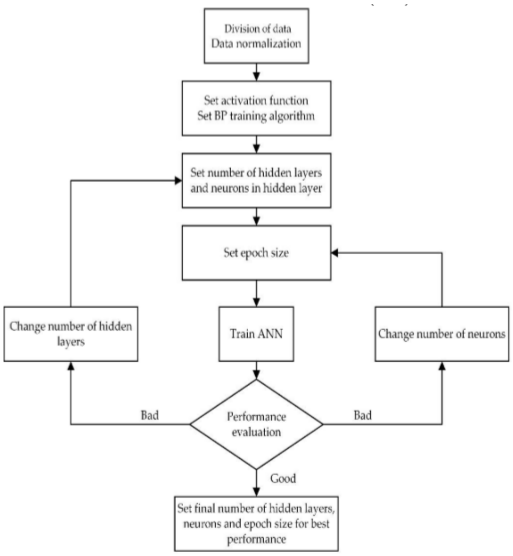
Division of data
Data normalization
Set activation function
Set BP training algorithm
Set number of hidden layers
and neurons in hidden layer
Set epoch size
Change number of hidden layers
Train ANN
Change number of neurons
Bad
Performance evaluation
Bad
Good
Set final number of hidden layers, neurons and epoch size for best performance

Fig-4.MLP-ANN training process for the best performance.

b) Rnn (recurrent neural networks) for time series analysis:

The feature-based method is used the various feature detection to detect the original features. Some of the features are, for example, corner detectors proposed by the author C. A. V. Hernández and F. A. P. Ortiz, "A corner detector algorithm for feature extraction in simultaneous localization and mapping" where they used shi – tomansi detector techniques which they compare with Harris corner, trajkovic, SIFT keypoint extractor where all the algorithm is compared with a standard image and also with OGM image maps with 40 sets of maps with the same algorithm mentioned above all the result tabled data was each technique detection of corner and time taken for the detection is calculated where the data is useful for the choose the optimal corner detection, and he finalized the result where Shi-Tomasi detector is the best techniques in terms of the quality and number of features extraction.[14] FAST, SURF(Speeded up Robust Features), BRIEF, BRISK, KAZE, A-KAZE, feature matching algorithm ANN(approximate nearest neighbor), some of the binary feature matching such as ORB-SLAM and FEAK. BBF, Spill-tree, hierarchical K-means tree, Randomized KD-trees [6].

The direct method of VO (visual odometry) is directly dependent on the pixel intensity value of a picture, so which can reduce the minimizes errors in the sensor space, where there is no need for the tracking and feature matching. The previous paper used the many SLAM algorithm, such as a direct monocular camera, stereo camera [17]. For the dynamic situation, they use of SfM (Structure from Motion) and use nonlinear least-squares estimation other methods, such methods like DTAM, PTAM were a lot of GPU parallelism required for the processing so the massive computation demand can be reduced. Many researchers are going on to attempt to apply deep learning methods on the visual odometry problem, which is categorized into two types supervised and unsupervised learning methods.

4 CNN-Slam

The CNN (Convolutional Neural Network) which is used for predicting the depths in the maps with the help of a deep neural network which can be deployed for accurate and dense monocular reconstruction. The maps which has dense depths that are naturally fused with depth measures obtained from direct monocular SLAM are predicted by CNN. This allows for all sorts of smart bots, such as those constrained to perform a given task [3, 10].

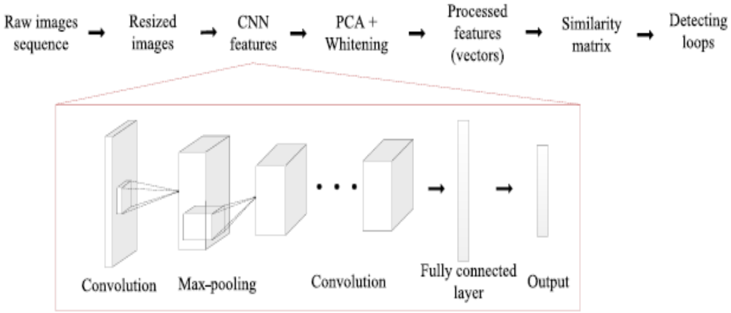

Fig.6 The architecture of CNN.

a) Depth prediction using CNN-SLAM

Structure from motion and simultaneous localization mapping which are coined as the umbrella names for active research area in the field of computer visions and robotics for the goal of 3D scene reconstruction and camera pose estimation from 3D and imaging sensors. Recently, real time SLAM methods aimed at fusing maps together which are obtained by moving depth sensors that have witnessed increased popularity which are being implemented for navigation and mapping of several types of robots, autonomous agents, drones, AI based vacuum cleaner, etc.

CNN explains and proves the potential of high resolution with good accuracy of the maps depth even under the absence of monocular cues to drive the depth estimation task. The other aspects of CNN explain that the same network architecture can be applied for different high dimensional regression tasks. One typical example is semantic segmentation that explains the feature of their framework where it uses pixel wise labels which effectively and coherently fuse sematic label with dense SLAM.The combination of SLAM and CNN to perform a high-level performance which are used for Augmented reality.

The Augmented Reality uses the SLAM algorithm in which they provide the geometric position because with the SLAM algorithms we can build a 3D maps of an environment which we want by allowing them to tracking the location and position of the camera in that environment. The algorithms estimate the sensor which can be built into the camera or cell phone or goggles which helps in Modeling the environment to create a map. By knowing the sensor's position in the device or robot or any device in which the sensor is planted and the direction/ position of the sensor combies with the generated 3D map of the environment lets the device (and the user looking through the device) move through the environment in reality. By adding deep learning/CNNs for perception the SLAM provides the location of where the camera/sensor position in the environment and a 3D model of the environment, to be aware and recognizing obstacles and the walls or any kinds of objects present in the environment can be identified by deep learning algorithms like CNNs. CNN's, the current method for implementing deep neural networks for vision, complement SLAM algorithms in AR systems by enhancing the user's AR experience or adding new capabilities to the AR system.

CNN will provide an accurate object recognition tasks which include location of the object and classification like identifying the image class i.e., distinguishing the type of animals like dogs or cats or birds and also if dogs means what type of breeds like Labrador or German Shepherd based on pre-training of the neural network's coefficients. The SLAM helps the camera to move through an environment without running into objects or obstacles, CNNs can identify that the object is a wall, refrigerator, or desk, and will highlight where it is in the field. Popular CNN graphs will be used for real time object detection that includes the classification and localization with the help of YOLO v2, Faster R-CNN, and Single-shot multi-box detector. Even theCNN’s object detection graphs are specialized to detect faces or hands. With the help of CNN based facial detection and recognition, AR systems can be programmed to be made to add a name and social media information above a person's face in the AR environment. By the help of CNN which detects the user's hands that helps the game developers to place a device or instrument needed in the game player's virtual hand which helps in recognizing the hand's existence which are more accessible than determining hand positioning. Some CNN based solutions will require a depth camera output as well as RGB sensor output for training and also to give a CNN graph.

CNN can also be applied for semantic segmentation which only cares about the pixels in an image. The semantic segmentation is concerned about every pixel in an image. In an automotive scene, a semantic segmentation CNN which will label all the pixels of the sky, road, buildings, private cars as a group, which is critical for self-driving car navigation. Applying this semantic segmentation in AR, it can detect and identifies the ceilings, walls, the floor, furniture, or other objects in the space. Thus, the Semantic knowledge of a scene by using the CNN that enables the realistic interactions between the real objects and virtual objects.

b) Hardware implementations for high-performance systems

In this hardware implementation both SLAM and CNN algorithms requires a significant amount of computations per camera captured image per frame. Making a seamless environment for the AR user’s which merges the real world with the virtual without significant latency requires a video frame rate of 20-30 frames per second (fps) which means AR system has about 33 to 40ms to capture, process, render, and display results to the user. The frame rate is faster when it completes the task faster thus AR feels it to be more natural.Considering single camera SLAM system for an SoC, computational efficiency and memory optimization are both critical design concerns. If the camera captures a 4k image at 30 fps, that means 8,294,400 pixels a frame or 248,832,000 pixels a second need to be stored and processed.

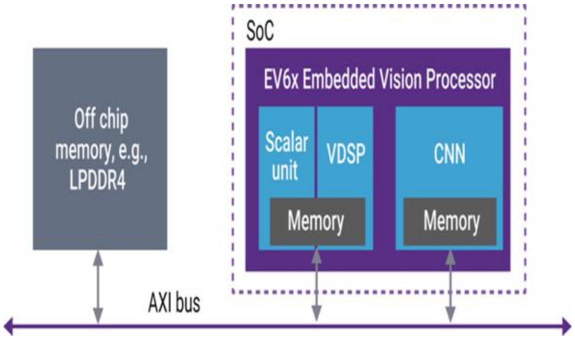

Fig.7 CNN proposed system architecture

5 loop closure using deep learning

The primary purpose of the loop closure is to give the robot the ability to know the place is the same scene. In other words, the robot can recognize the area that if it's been there before or not. These kinds of issues have been one of the most significant obstacles of large scale SLAM and recover from creating errors again and again. Then the perception aliasing is the two different places that may be considered the same, which represents the problem even when the cameras used as sensors due to the repetitive nature of the environment, e.g., corridor, similar architecture elements are areas with lots of bushes. Proper closed-loop detection cannot return any false positives and obtain less false negatives. P Newman and K Ho [33] proposed this loop closure with the help of an experiment using a small ATRV Jnr mobile robot that was driven around the building, which contains a large loop or extremely challenging environment for contemporary SLAM algorithm. The vehicle camera kept a constant orientation in vehicle coordinates, looking forward and slightly to the right. The camera captures images every two seconds and written to disk. That vehicle is equipped with a standard SICK laser, in which the output was also logged along with the odometry from the wheel encoders.

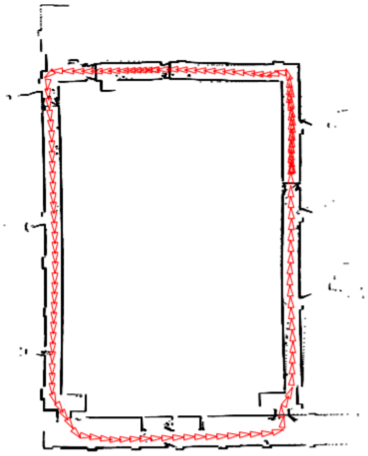

Fig.8 The map of the test area after closed looping

The above Figure shows the final map of the study [33] after applying the loop closing constraint. The marginal covariance on each of the vehicle poses decreases and a crisp result, which would be the case of any SLAM algorithm.

6 Conclusion

Thus, the SLAM, which is used in the AI robots becomes much easier for them as they use the process of deep learning and help the robots to move safe and short easy distance to the destination without being hit any obstacle between them. The robots, by applying the deep learning techniques they travel safer and quicker. According to the studies we took, it has been clearly explained that the SLAM with the help of Deep Learning the robot will drive faster and safer as the AI features have been improving faster in the current generation. In the future, the SLAM will be much quicker and more reliable for the robot to predict the map of the new area without making many mistakes and predicting the direction much faster due to their brain, i.e., the Artificial Intelligence, which is considered the mind of the robots. The recently improved technology of mapping is used in the autonomous car for predicting their direction to travel and the safer and faster distance for them to travel to the destination. In the future, the process of unsupervised learning will be a better process for the robot further to consolidate the deep learning contribution of the SLAM.

References

1. A. M. Azri, S. Abdul-Rahman, R. Hamzah, Z. A. Aziz, and N. A. Bakar, “Visual analytics of 3D LiDAR point clouds in robotics operating systems,” Bull. Electr. Eng. Informatics, vol. 9, no. 2, pp. 492–499, 2020.
2. Z. Zhao, Y. Mao, Y. Ding, P. Ren, and N. Zheng, "Visual-Based Semantic SLAM with Landmarks for Large-Scale Outdoor Environment," Proc. - 2nd China Symp. Cogn. Comput. Hybrid Intell. CCHI, 2019, pp. 149–154, 2019.
3. X. Wang, “Autonomous Mobile Robot Visual SLAM Based on Improved CNN Method,” IOP Conf. Ser. Mater. Sci. Eng., vol. 466, no. 1, 2018.
4. A. Singandhupe and H. La, “A Review of SLAM Techniques and Security in Autonomous Driving,” Proc. - 3rd IEEE Int. Conf. Robot. Comput. IRC 2019, no. 19, pp. 602–607, 2019.
5. C. Duan, S. Junginger, J. Huang, K. Jin, and K. Thurow, “Deep Learning for Visual SLAM in Transportation Robotics: A review,” Transp. Saf. Environ., vol. 1, no. 3, pp. 177–184, 2019.
6. A. Li, X. Ruan, J. Huang, X. Zhu, and F. Wang, “Review of vision-based simultaneous localization and mapping,” Proc. 2019 IEEE 3rd Inf. Technol. Networking, Electron. Autom. Control Conf. ITNEC 2019, no. Itnec, pp. 117–123, 2019.
7. M. Zeilinger, R. Hauk, M. Bader, and A. Hofmann, “Design of an Autonomous Race Car for the Formula Student Driverless (FSD),” Proc. OAGM ARW Jt. Work., pp. 3–8, 2017.
8. G. Bresson, Z. Alsayed, L. Yu, and S. Glaser, “Simultaneous Localization and Mapping: A Survey of Current Trends in Autonomous Driving,” IEEE Trans. Intell. Veh., vol. 2, no. 3, pp. 194–220, 2017.
9. B. M. Potts and J. B. Reid, “Variation in the Eucalyptus gunnii-archeri complex. II.* The origin of variation,” Aust. J. Bot., vol. 33, no. 5, pp. 519–541, 1985.
10. M. Etxeberria-Garcia, M. Labayen, M. Zamalloa, and N. Arana-Arexolaleiba, “Application of Computer Vision and Deep Learning in the railway domain for autonomous train stop operation,” Proc. 2020 IEEE/SICE Int. Symp. Syst. Integr. SII 2020, pp. 943–948, 2020.
11. M. Colosi et al., “Plug-and-Play SLAM: A Unified SLAM Architecture for Modularity and Ease of Use,” 2020.
12. R. Giubilato, S. Chiodini, M. Pertile, and S. Debei, "An evaluation of ROS-compatible stereo visual SLAM methods on a Nvidia Jetson TX2," Meas. J. Int. Meas. Confed., vol. 140, no. April, pp. 161–170, 2019.
13. P. Muehlfellner, P. Furgale, W. Derendarz, and R. Philippsen, “Evaluation of fisheye-camera based visual multi-session localization in a real-world scenario,” IEEE Intell. Veh. Symp. Proc., pp. 57–62, 2013.
14. C. A. V. Hernández and F. A. P. Ortiz, “A corner detector algorithm for feature extraction in

simultaneous localization and mapping," J. Eng. Sci. Technol. Rev., vol. 12, no. 3, pp. 104–113, 2019.

15. A. Karimian, Z. Yang, and R. Tron, "Statistical Outlier Identification in Multi-robot Visual SLAM using Expectation Maximization," 2020.
16. H. Liu, G. Liu, G. Tian, S. Xin, and Z. Ji, "Visual SLAM based on dynamic object removal," IEEE Int. Conf. Robot. Biomimetics, ROBIO 2019, pp. 596–601, 2019.
17. S. Saeedi, M. Trentini, H. Li, and M. Seto, "Multiple-robot Simultaneous Localization and Mapping - A Review 1 Introduction 2 Simultaneous Localization and Mapping : problem statement," J. F. Robot., vol. 33, no. 1, pp. 3–46, 2016.
18. V. Ilci and C. Toth, "High definition 3D map creation using GNSS/IMU/LiDAR sensor integration to support autonomous vehicle navigation," Sensors (Switzerland), vol. 20, no. 3, 2020.
19. S. A. Scherer and A. Zell, "Efficient onbard RGBD-SLAM for autonomous MAVs," IEEE Int. Conf. Intell. Robot. Syst., pp. 1062–1068, 2013.
20. Z. Rozsa, M. Golarits, and T. Sziranyi, "Localization of Map Changes by Exploiting SLAM Residuals," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 12002 LNCS, pp. 312–324, 2020.
21. Z. Javed and G.-W. Kim, "A Comparative Study of Recent Real-Time Semantic Segmentation Algorithms for Visual Semantic SLAM," pp. 474–476, 2020.
22. G. Ros, A. Sappa, D. Ponsa, and A. Lopez, "Visual SLAM for Driverless Cars: A Brief Survey," IEEE Work. Navig. Perception, Accurate Position. Mapp. Intell. Veh., pp. 1–6, 2012.
23. J. Fuentes-Pacheco, J. Ruiz-Ascencio, and J. M. Rendón-Mancha, "Visual simultaneous localization and mapping: a survey," Artif. Intell. Rev., vol. 43, no. 1, pp. 55–81, 2012.
24. M. R. U. Saputra, A. Markham, and N. Trigoni, "Visual SLAM and structure from motion in dynamic environments: A survey," ACM Comput. Surv., vol. 51, no. 2, 2018.
25. T. Taketomi, H. Uchiyama, and S. Ikeda, "Visual SLAM algorithms: A survey from 2010 to 2016," IPSJ Trans. Comput. Vis. Appl., vol. 9, 2017.
26. J. Ni, T. Gong, Y. Gu, J. Zhu, and X. Fan, "An Improved Deep Residual Network-Based Semantic Simultaneous Localization and Mapping Method for Monocular Vision Robot," Comput. Intell. Neurosci., vol. 2020, 2020.
27. D. S. Chaplot, D. Gandhi, S. Gupta, A. Gupta, and R. Salakhutdinov, "Learning to Explore using Active Neural SLAM," 2020.
28. J. von Falkenhausen and Q. Liu, Fusion of raw sensor data for testing applications in autonomous driving. 2020.
29. S. Wen, Y. Zhao, X. Yuan, Z. Wang, D. Zhang, and L. Manfredi, "Path planning for active SLAM based on deep reinforcement learning under unknown environments," Intell. Serv. Robot., vol. 13, no. 2, pp. 263–272, 2020.
30. N. Botteghi, B. Sirmacek, K. A. A. Mustafa, M. Poel, and S. Stramigioli, "On Reward Shaping for Mobile Robot Navigation: A Reinforcement Learning and SLAM Based Approach," 2020.
31. W. Gao, D. Hsu, W. S. Lee, S. Shen, and K. Subramanian, "Intention-Net: Integrating Planning and Deep Learning for Goal-Directed Autonomous Navigation," no. Figure 1, pp. 1–10, 2017.
32. G. Hou et al., "A novel underwater simultaneous localization and online mapping algorithm based on neural network," ISPRS Int. J. Geo-Information, vol. 9, no. 1, 2019.
33. P. Newman and K. Ho, "SLAM- Loop closing with visually salient features," Proc. - IEEE Int. Conf. Robot. Autom., vol. 2005, pp. 635–642, 2005.