

HOSTED BY



Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: [www.sciencedirect.com](http://www.sciencedirect.com)

Original article

# GAN based simultaneous localization and mapping framework in dynamic environment

Bo Sun<sup>a,b,\*</sup>, Song Gao<sup>c</sup>, Hongmei Zi<sup>d</sup>, Qiong Wu<sup>e</sup><sup>a</sup>Academic Affairs Office, Xi'an Technological University, Xi'an 710021, China<sup>b</sup>School Of Mechatronic Engineering, Xi'an Technological University, Xi'an 710021, China<sup>c</sup>School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, China<sup>d</sup>Library, Xi'an Technological University, Xi'an 710021, China<sup>e</sup>School of Computer Science and Engineering, Xi'an Technological University, Xi'an 710021, China

## ARTICLE INFO

## Article history:

Received 10 January 2022

Revised 21 July 2022

Accepted 25 August 2022

Available online 5 September 2022

## Keywords:

Localization

Mapping

Cloud

Generative Adversarial Network and Aquila

Optimizer

## ABSTRACT

**Objective:** Because of interference from dynamic objects, the traditional simultaneous localization and mapping (SLAM) framework execute poorly while operational in a dynamic environment. The Dynamic-SLAM model is introduced by considering the merits of deep learning in object discovery. At the semantic level, the dynamic objects in the new detection are detected to construct the prior knowledge with an SSD object detector. Among the radar and the LiDAR, the synchronization and conversion calibration makes the multi-sensor fusion. Localizing using a camera in a dynamic environment is more challenging because the localization process takes place in moving segments. This further results in unstable pose estimation. A large amount of information of the visual point cloud and high precision of the laser radar information enhances the accuracy of real-time positioning thereby attaining grid map and 3-D point cloud map.

**Methods:** Here, we have proposed Generative Adversarial Network (GAN) with Aquila Optimizer for moving object detection. The LiDAR measurements check the Radar outcomes. The targeted moving objects are determined via Doppler velocity from the radar and their exact location and mass can be estimated with LiDAR and the proposed GAN-AO approach. Hence the GAN-based AO approach is used to segment the objects inside the point clouds. The arrangements of point clouds are made in a particular range to multiply the vertical points with the laser channel number. If the same objects are identified then the angles are analyzed among the image vectors and then labeled identically for the same points of the object. In addition to this, Velocity compensation is made to estimate the actual moving target from the world frame. This is due to the fact that the estimated velocity of the mmW-radar is in the radial direction with respect to the touching sensors.

**Results:** The investigation is conducted between different state-of-art methods like Deep Learning (DL), Generative Adversarial Network (GAN), artificial neural network (ANN), Deep neural network (DNN) and propose methods. From this analysis, the proposed method provided 93.8% detection accuracy than other existing methods like DL, ANN, GAN and DNN respectively.

**Conclusions:** While comparing to the state-of-art techniques, the proposed method demonstrated superior performance results in terms of tracking, detection, root mean square error (RMSE) and accuracy.

© 2022 The Authors. Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

\* Corresponding author at: Academic Affairs Office, School Of Mechatronic, Xi'an Technological University, Xi'an 710021, China.

E-mail address: [bosun12021@163.com](mailto:bosun12021@163.com) (B. Sun).

Peer review under responsibility of King Saud University.



## 1. Introduction

Visual Synchronization Localization is a process in wireless sensor networks to determine the location of sensor nodes. Under the field of the sensor in which they deployed, the Localization algorithm was designed (Ding et al., 2018; Zhang et al., 2019; Labbé and Michaud, 2019; Fuentes-Pacheco et al., 2015; Tsintotas et al., 2022). Some sensor nodes are static and some are dynamic, based

<https://doi.org/10.1016/j.jksus.2022.102298>

1018-3647/© 2022 The Authors. Published by Elsevier B.V. on behalf of King Saud University.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

on their nature localization algorithm was built. Localizing using a camera (Zheng et al., 2022; Ruan et al., 2022; Chen et al., 2021; Wang et al., 2021a; Wang et al., 2021b; Wang et al., 2021c; Feng, 2021; Lee et al., 2021; Zhao et al., 2022; Sodhi et al., 2022) in a dynamic environment is more challenging because the localization takes place in moving segments and leads to unstable pose estimation. In this proposed work, we use GAN with an Aquila optimizer for Visual Synchronization Localization and Mapping depending upon the visual point cloud information and laser in a dynamic environment (Shastri et al., 2022; Lindqvist et al., 2021; Venator et al., 2021; Rajendran et al., 2021; Wang et al., 2021a; Wang et al., 2021b; Wang et al., 2021c; Shin et al., 2019). The proposed scheme uses laser and visual point cloud information (Zhao et al., 2022; Ghaffari et al., 2019) to support Visual Synchronization Localization to optimize pose. A large amount of information on the visual point cloud and high precision of the laser radar information enhances the accuracy of real-time positioning thereby attaining grid map and 3-D point cloud map (Morales and Kassas, 2017).

To improve the performance of the Visual Synchronization Localization, we have to combine Radar and Lidar (Schultz and Zarzycki, 2021) to effectively localize the dynamic objects. The static objects are canceled via the measurements of Lidar and a visual point clouds from Lidar are segmented. The input data for the Lidar based Visual Synchronization Localization (Shao et al., 2015) was the filtered visual point clouds. Our proposed method works efficiently in real-time. The accuracy and robustness of Visual Synchronization Localization were improved by Laser and visual point cloud information on the other hand the dynamic objects can be efficiently identified (Yu et al., 2020; Xiao et al., 2019). The major contributions of this study are summarized as follows:

- The dynamic objects in the new detection are detected to construct the prior knowledge with an SSD object detector at the semantic level.
- For moving object detection, we used Generative Adversarial Network (GAN) with Aquila Optimizer.
- Doppler velocity from the radar detects the targeted moving objects and their exact location and mass can be estimated with LiDAR and the proposed GAN-AO approach.

The remaining sections of this article are summarized as: Section 2 illustrated the literature survey. Section 3 explains the system model and the GAN with AO is formulated in section 4. Section 5 delineates the proposed methodology and the experimental results are discussed in section 6. At last, section 7 terminates the article.

## 2. Related works

By fusing mmW-radar and LiDAR, Dang et al. (2020) suggested an effective method for the reduction of the power of dynamic environment on SLAM. The mapping and localization accuracy was enhanced. Due to doppler effect, the moving objects are eliminated via localization and segmentation with efficient moving object detection. Different real world scenarios were used to validate the performance of moving object elimination. The quick dynamic object detections were provided via calibration and synchronization. Fed the resulting point clouds to SLAM and thereafter removed moving objects.

The mapping system and robust visual-Lidar simultaneous localization were suggested by Qian et al. (2021) for unmanned aerial vehicles. From point clouds via clustering, extract the plane features and more stable line. The least-squares iterative

closest point algorithm calculates the relation pose among both consecutive frames. At a lower frequency, construct the texture information with 3-D map thereby referring to the elimination of pose. For unmanned aerial vehicles (UAV), the higher precise and robust mapping and localization with higher cost were achieved.

A robust framework was suggested by Wang et al. (2021a), Wang et al. (2021b), Wang et al. (2021c) for simultaneous mapping and localization by multiple non-repetitive scanning Lidars. The map alignment configures the transformation between two lidars depending upon the rigidity assumption of geometric structure. Time synchronizes the original information from various lidars. While estimating lidar odometry, send all the feature candidates. For enhanced loop detection, integrate the novel place descriptor and remove the dynamic objects. An experimental result verifies large motion and feature-less scenario performances.

Xiao et al. (2019) suggested deep learning (DL) based semantic monocular visual mapping and localization in a dynamic environment. In the following thread, the selective tracking algorithm via feature points of dynamic objects was processed thereby constructing a feature-oriented visual SLAM organization. Compared to the original SSD network, the recall rate of the system is raised from 82.3 % to 99.8 % respectively. In a real-world dynamic environment, an accurate environmental map was localized. But, the environment was more complex and more time-consuming.

## 3. System model

The system model of the suggested method is depicted in Fig. 1. The multi-sensor fusion can be made with the time synchronization and conversion calibration among the radar and the LiDAR. From the LiDAR dense and precise clouds are produced (Nguyen and Le, 2013). Then the segmentation using the GAN-AO approach is conducted to detect the objects from the surroundings. However, the targeted objects are analyzed and detected by using this approach and then the outcomes are collected on Radar. Meanwhile, the outcomes of Radar are checked with the help of LiDAR measurements and thus filtered the ghost targets effectively (Dubayah and Drake, 2000). The targeted moving objects can be detected by Doppler velocity from the radar and their exact location and mass can be estimated with LiDAR and the proposed GAN-AO segmentation approach.

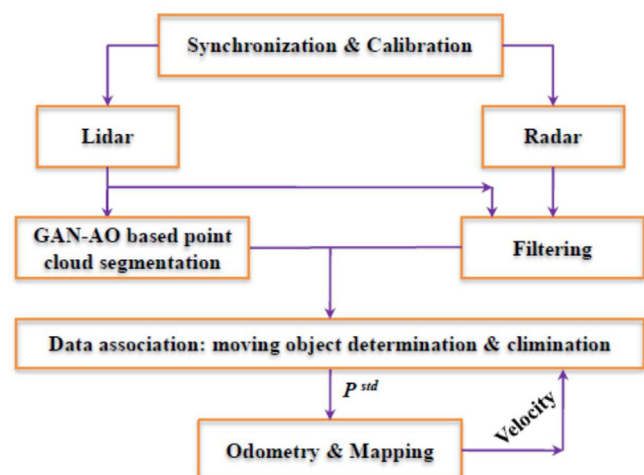


Fig. 1. System model.

#### 4. Formulation of Generative Adversarial Network and Aquila Optimizer

This section describes the background of the adopted GAN and AO approaches for segmentation.

##### 4.1. Generative Adversarial network (GAN)

The discriminator  $D$  and generator  $G$  are the two sub-networks of the GAN network. For training, provides the ground truth data representation is generated via generator attempts. The generator data produces the true ground truth information from data, which are differentiated from the discriminator attempts. According to the distribution  $P$ , the binary segmentation map  $Y$  is produced via the map  $F : X \rightarrow Y$ . From the generator sub-network, the ground truth mask  $s$  represented as  $Y$ . The discriminators are indicated via segmenting map into 0 to 1 values and input data comprises the pair of discriminator maps  $\{X, Y\}$ .

Equation (1) expresses the GAN for segmentation and the following equation expresses the objective function.

$$GAN_L(F, D) = F_{X,Y \sim P_{data}(X,Y)}[\log D(X, Y)] + E_{X \sim P_{data}(Y)}[\log(1 - D(X, F(Y)))] \quad (1)$$

To make the right decision, the discriminator ( $D$ ) subnetwork is trained via minimization of  $D(X, D(X))$  and the maximization of  $D(X, Y)$ . While making the right decision, the outputs are generated via the generator sub-network ( $F$ ). The secondary loss functions of binary cross-entropy ( $A_L(F)$ ) are delineated in addition to the GAN objective function (Abdollahi et al., 2021).

$$A_L(F) = F_{X,Y \sim P_{data}(X, Y)}[-Y \cdot \log F(X) + (1 - Y) \cdot \log(1 - F(Y))] \quad (2)$$

The GAN objective functions with the optimal results are delineated as.

$$F^* = \arg \min_D \left[ \max_D GAN_L(F, D) \right] + \gamma A_L(F) \quad (3)$$

The weighting parameters are  $\gamma$ . The map generates the spatial resolution are increases via a generator. At last, the practical probability outputs are achieved based on both backward and forward passes. However to enhance the segmentation capabilities we have adopted Aquila optimization. This will help to tune up the parameters.

##### 4.2. Aquila optimization

The parameters of GAN are tuned using the Aquila Optimization (AO) (AlRassas et al., 2021) which is a swarm-based approach. The estimation of the best solution can be acquired by,

$$H = \begin{bmatrix} h_{1,1} & \dots & h_{1,j} & h_{1,Dim-1} & h_{1,Dim} \\ h_{2,1} & \dots & h_{2,j} & \dots & h_{2,Dim} \\ \dots & \dots & h_{i,j} & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{N-1,1} & \dots & h_{N-1,j} & \dots & h_{N-1,Dim} \\ h_{N,1} & \dots & h_{N,j} & h_{N,Dim-1} & h_{N,Dim} \end{bmatrix} \quad (4)$$

The dimensionality of the segmentation problem in the dynamic object detection is given as  $Dim$ .  $N$  is the entire amount of solutions to the problem. The best solution can be determined as  $H$ . Here, the attained value while performing the  $i^{th}$  solution is represented as  $H_i$ .

$$H_{ij} = rand \times (U_j - L_j) + L_j, i = 1, 2, \dots, N; j = 1, 2, \dots, Dim \quad (5)$$

The randomly generated number can be given as  $rand$  and falls under the range of 0 to 1,  $L_j$  is the  $j^{th}$  lower boundary and  $U_j^{th}$  upper boundary is indicated as  $UC_j$ .

##### (i) Numerical Expression of AO

The hunting behavior of the AO can be classified as (i) Expanded exploration (ii) Encircling (iii) Expanded exploitation and (iv) Narrowed exploitation.

##### (ii) Expanded Exploration (H1)

The targeted ghosts are chosen with the characteristics of Aquila selecting the hunting area for the hunting.

$$H_1(t+1) = H_{best}(t) \times \left(1 - \frac{t}{T}\right) + (H_M(t) - H_{best}(t) * rand) \quad (6)$$

The solution obtained after the first search application is denoted as  $H_1(t+1)$ . However, the best previous solution can be given as  $H_{best}(t)$ . That can be used to analyze the exact object. The exploration can be managed by observing the term  $(1 - \frac{t}{T})$  for each iteration and  $H_M(t)$  is used to find the position average value of the ongoing solutions and can be evaluated as,

$$H_M(t) = \frac{1}{N} \sum_{i=1}^N H_i(t), \forall j = 1, 2, \dots, Dim \quad (7)$$

The maximum number of iterations is represented as  $T$ .

##### 4.3. Encircling (H2)

According to the behavior of Aquila encircling and attacking the prey, the targeted ghosts are encircled and analyzed for detection. This is termed as Aquila contour flight which involves a short glide attack and can be given as,

$$H_2(t+1) = H_{best}(t) \times Levy(D) + H_R(t) + (y - x) * rand \quad (8)$$

The solution produced by the encircling process can be given as  $H_2(t+1)$ .  $Levy(D)$  has been indicated as levy flight distribution function and the space dimensionality is given as  $D$ . however, the randomly acquired solution ranges between 1 and  $N$  is given as  $H_R(t)$ . Then the levy flight distribution function can be estimated as,

$$Levy(D) = \sigma \times \frac{u \times \beta}{|v|^{\frac{1}{\beta}}} \quad (9)$$

here,  $\sigma$  is a constant value. The randomly selected numbers that lie between 0 and 1 is given as,  $u$ , and  $v$ . Then the  $\beta$  value is calculated as,

$$\beta = \left( \frac{\Gamma(1 + \delta) \times \sin(\frac{\pi\delta}{2})}{\Gamma(\frac{1+\delta}{2}) \times \delta \times 2^{\frac{\delta-1}{2}}} \right) \quad (10)$$

here,  $\delta = 1.6$  and the constant values  $x$  and  $y$  are evaluated as,

$$y = a \times \cos(\theta) \quad x = a \times \sin(\theta) \quad (11)$$

where  $a = a_1 + U \times D_1$  and  $\theta = -\omega \times D_1 + \theta_1$ ,  $\theta_1 = \frac{3 \times \pi}{2}$ ; and  $r_1$  falls between the values 1 and 20, and  $U$  is a constant value is equal to 0.00564.  $D_1$  is the integer value of  $Dim$ . The constant  $\omega$  is given as 0.006.

#### 4.4. Expanded exploitation (H3)

This follows the segmentation of objects based on the identified exact location and is performed slowly and carefully. It can be determined as,

$$H_3(t+1) = (H_{best}(t) - H_M(t)) \times \alpha - rand + ((U - L) \times rand + L) \times v \quad (12)$$

The solutions obtained using the third stage are given as  $H_3(t+1)$ . The parameters used to change the exploitation stage can be denoted as  $\alpha$  and  $v$  and determined as 0.1. The lower and upper boundaries of the problem can be denoted as L and U.

#### 4.5. Narrowed exploitation (H4)

In this stage, object segmentation can be made with the help of a walk and grab prey approach. It can be evaluated as,

$$H_4(t+1) = F \times H_{best}(t) - (G_1 \times H(t) \times rand) - G_2 \times Levy(D) + rand \times G_1 \quad (13)$$

The quality function is indicated as F and the obtained solution in this stage is given as  $H_4(t+1)$ . G1 is the motion tracker and G2 is the flight slope between the first and last location.

To enhance the segmentation accuracy of GAN we have merged it with the AO approach and Fig. 2 demonstrates the proposed segmentation methods with their working mechanism.

### 5. Proposed GAN-AO for the elimination of moving objects

This section expresses the process of Generative Adversarial Network and Aquila Optimizer for the elimination of moving objects.

#### 5.1. Point cloud segmentation

During the segmentation process, the earth plane from the point clouds is removed. While performing this estimation is inaccurate and does not match with the SLAM (Bailey and Durrant-Whyte, 2006). Hence the GAN-based AO approach is used to segment the objects inside the point clouds. The arrangements of point clouds are made in a particular range to multiply the vertical points with the laser channel number. If the same objects are identified then the angles are analyzed among the image vectors and then labeled identically for the same points of the object.

#### 5.2. Filtering outcomes from the radar

The object detected using the radar includes noises because of propagation features and multi-path effect of EM waves. Hence to analyze a negative impacts of the inclusion of false alarms a verification strategy is stated along with LiDAR measurements to detect the targeted ghost. The targeted points are converted into a coordinate frame of LiDAR-based on the calibration (Dawkins et al., 2001) outcomes.

#### 5.3. Association of data

Velocity compensation is made to estimate the actual moving target from the world frame. This is due to the fact that the estimated velocity of the mmW-radar is in the radial direction with respect to the touching sensors. The Doppler velocity (Berger, 1957)  $V_t^{sensor}$  is acquired from the radar for each target. Then the absolute velocity  $V_t^A$  of the target as of the current static frame is determined by the integration of sensor velocity  $V_{sensor}^{world}$  along with the outcomes of SLAM and can be given as,

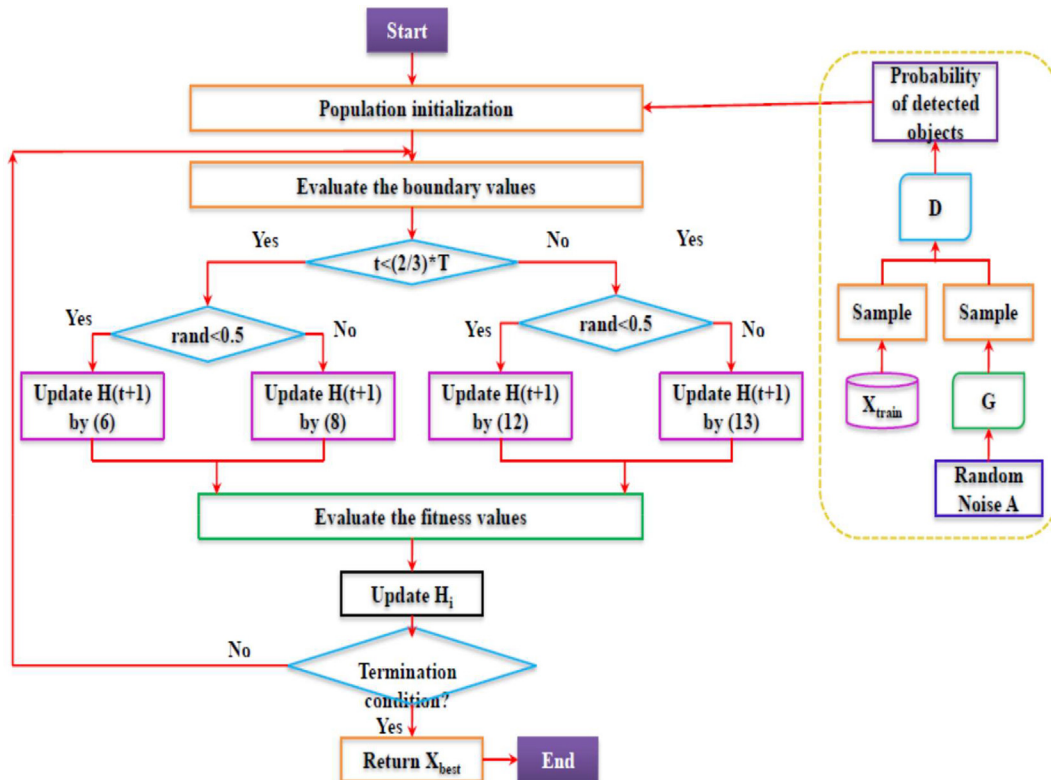


Fig. 2. Proposed GAN-AO segmentation approach.

$$V_t^A = P_{world}^{Sensor} V_{sensor}^{world} + V_t^{sensor} \tag{14}$$

here, the inverse of current arrangements of the sensor in world fame can be indicated as  $P_{world}^{Sensor}$ . Then the status of the moving object can be evaluated.

### 6. Experimental results

This section validates the performance of GAN-based Aquila Optimizer for dynamic object detection. The mmWradar (Delphi ESR) with the LiDAR (VLP-16) equips the wheeled robot based on the TUM RGB-D benchmark dataset (Sturm et al., 2012). There is 30°×360° of FOV (field of view) with 100 m for LiDAR detection range. At mid-range, a wide of 90° FOV is combined by radar and the scenarios are recorded through a camera. The baseline as open-source code LOAM conducts a series of experiments and real-time implementations are performed using Robot Operating System (ROS). Fig. 3 explains the sample dataset images.

The performance analysis based on tracking and detection results with respect to various experimentations is plotted in Fig. 4. There are eight experimentation conditions with the tracking and detection performance is evaluated in this graph. Next to the enhancement measures, the transplant results are evaluated based on the RGB-D benchmark dataset. From this investigation, the experimentation conditions are increased with decreasing detection d tracking results are obtained.

The root means square error (RMSE) performance analysis is delineated in Fig. 5. The experimentation conditions are varied from 1 to 8 with respect to varying error values. The RMSE values are varied in each experiment condition. From this, we have obtained 0.23, 0.018, 0.20, 0.22, 0.21, 0.21, 0.25 and 0.15 RMSE output values are obtained based on the experimentation conditions from 1 to 8.

Table 1 delineates the operation performance evaluation based on three examination time results of ORB-SLAM2, Dynamic SLAM and Improvements. When compared to ORB-SLAM2, the dynamic SLAM accuracy is 7.84 % by taking the RMSE as the standard.

Fig. 6 expresses the performance analysis based on the number of frames with varying numbers. In this experiment, we have taken three conditions such as static, and dynamic with both static and dynamic conditions. In this experiment, the average number of feature points in every keyframe becomes 3090 in order to highlight the dynamic feature point's proportions. There are 47.96 % of static feature points with 74.11 % dynamic feature points proportions obtained.

The comparative analysis of accuracy results are plotted in Table 2. This investigation is conducted between different state-of-art methods like Deep Learning (DL) (Xiao et al., 2019), Generative Adversarial Network (GAN), Artificial neural network (ANN), Deep neural network (DNN) and proposed methods (Bojaj et al., 2021; anghera et al., 2021). From this analysis, the proposed

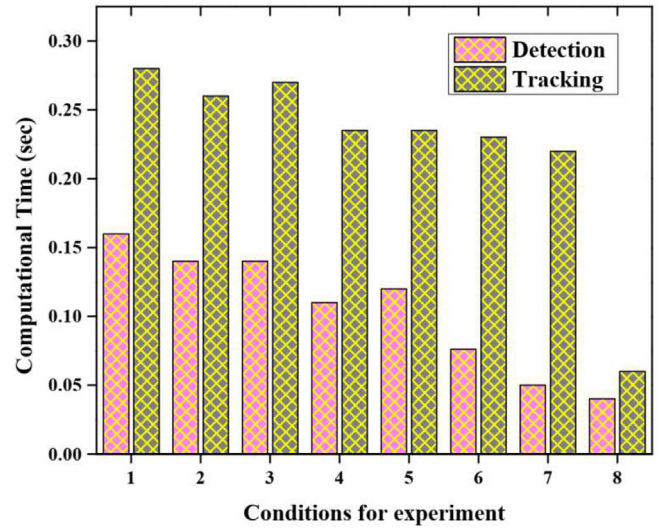


Fig. 4. Performance analysis based on tracking and detection results.

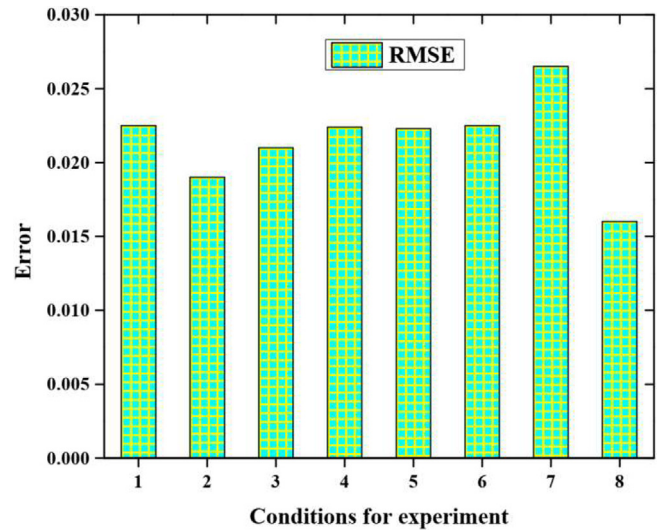


Fig. 5. Performance analysis in terms of RMSE.

Table 1  
Three examination times for operation performance evaluation.

Time (s)	ORB-SLAM2	Dynamic SLAM	Improvements
<b>Median</b>	0.50	0.044	12 %
<b>Mean</b>	0.50	0.045	10 %
<b>Total</b>	42.90	38.40	10.49 %

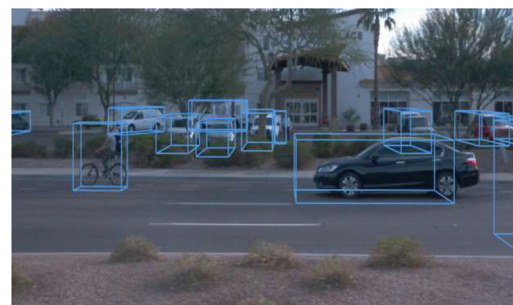
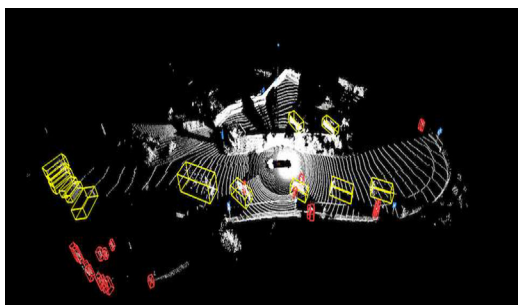


Fig. 3. Sample dataset images based on dynamic moving objection detection.

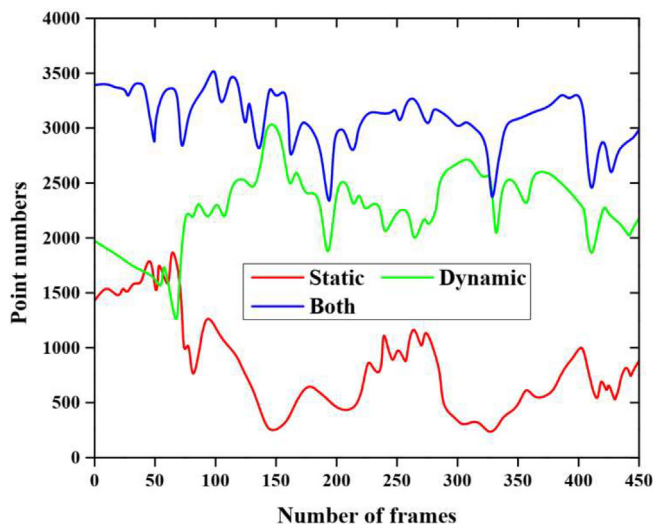


Fig. 6. Performance analysis based on number of frames with varying numbers.

Table 2  
Comparative analysis of accuracy.

Methods	Accuracy
DL	86 %
GAN	90 %
DNN	90.4 %
ANN	89.45 %
Proposed	93.8 %

method provided 93.8 % detection accuracy than other existing methods like DL, ANN, GAN and DNN respectively.

## 7. Conclusion

In this paper, the Visual Synchronization Localization and mapping of the dynamic object are detected using laser and visual cloud point information. Localizing using a camera in a dynamic environment is more challenging because the localization takes place in moving segments and leads to unstable pose estimation. To overcome this problem, the proposed work uses GAN with an Aquila optimizer in a dynamic environment. In this experiment, we have taken three conditions such as static, dynamic with both static and dynamic conditions. The average number of feature points in every keyframe becomes 3090 to highlight the dynamic feature point's proportions. There are 47.96 % of static feature points with 74.11 % dynamic feature points proportions are obtained. The proposed technique offers 93.8 % detection accuracy than other existing methods like DL, ANN, GAN and DNN. In the future research, various other parameters namely the execution time as well as computational cost will be discussed and taken into consideration.

## Funding

Key R & D plan project of Shaanxi Province in 2021 "digital evaluation system of whole dentition based on coded structured light measurement" (Item No: 2021GY-005).

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

Abdollahi, A., Pradhan, B., Sharma, G., Maulud, K.N.A., Alamri, A., 2021. Improving road semantic segmentation using generative adversarial network. *IEEE Access* 9, 64381–64392.

AlRassas, A.M., Al-qaness, M.A., Ewees, A.A., Ren, S., Abd Elaziz, M., Damaševićius, R., Krilavičius, T., 2021. Optimized ANFIS model using Aquila Optimizer for oil production forecasting. *Processes* 9 (7), 1194.

Anghera, S., Neb, T., Kaur, S.P., Madan, A., Singh, B., 2021. Assessment and Role of Social Media in Dental Education. *SPR* 1 (2), 52–57.

Bailey, T., Durrant-Whyte, H., 2006. Simultaneous localization and mapping (SLAM): Part II. *IEEE Rob. Autom. Mag.* 13 (3), 108–117.

Berger, F.B., 1957. The nature of Doppler velocity measurement. *IRE Trans. Aeronaut. Navig. Electron.* 3, 103–112.

Bojaj, G., Agahi, R., Hoxha, I., 2021. Treatment of the first COVID-19 case in kosovo and management of the pandemic. *SPR* 1 (3), 58–62.

Chen, M., Tang, Y., Zou, X., Huang, Z., Zhou, H., Chen, S., 2021. 3D global mapping of large-scale unstructured orchard integrating eye-in-hand stereo vision and SLAM. *Comput. Electron. Agric.* 187, 106237.

Dang, X., Liang, X., Li, Y., Rong, Z., 2020. Moving objects elimination towards enhanced dynamic SLAM fusing LIDAR and MmW-radar. In: 2020 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM). IEEE, pp. 1–4.

Dawkins, C., Srinivasan, T.N., Whalley, J., 2001. Calibration. *Handbook of econometrics* 5, 3653–3703.

Ding, X., Wang, Y., Li, D., Tang, L., Yin, H., Xiong, R., 2018. Laser Map Aided Visual inertial Localization in Changing Environment. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 4794–4801.

Dubayah, R.O., Drake, J.B., 2000. Lidar remote sensing for forestry. *J. Forest.* 98 (6), 44–46.

Feng, T., 2021. Deep Learning for Depth, Ego-Motion, Optical Flow Estimation, and Semantic Segmentation. University of Essex. Doctoral dissertation.

Fuentes-Pacheco, J., Ruiz-Ascencio, J., Rendón-Mancha, J.M., 2015. Visual simultaneous localization and mapping: a survey. *Artif. Intell. Rev.* 43 (1), 55–81.

Ghaffari, M., Clark, W., Bloch, A., Eustice, R.M., Grizzle, J.W., 2019. Continuous direct sparse visual odometry from RGB-D images. *arXiv preprint arXiv:1904.02266*.

Labbé, M., Michaud, F., 2019. RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *J. Field Rob.* 36 (2), 416–446.

Lee, K.M.B., Kong, F.H., Cannizzaro, R., Palmer, J.L., Johnson, D., Yoo, C. and Fitch, R., 2021. Decentralised Intelligence, Surveillance, and Reconnaissance in Unknown Environments with Heterogeneous Multi-Robot Systems. *arXiv preprint arXiv:2106.09219*.

Lindqvist, B., Agha-Mohammadi, A.A., Nikolakopoulos, G., 2021. ExplorAtion-RRT: A Multi-objective PATH Planning and ExplorAtion Framework for Unknown and Unstructured Environments. In: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 3429–3435.

Morales, J.J., Kassas, Z.M., 2017, September. Distributed signals of opportunity aided inertial navigation with intermittent communication. In Proceedings of the 30th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2017) (pp. 2519–2530).

Nguyen, A., Le, B., 2013, November. 3D point cloud segmentation: A survey. In 2013 6th IEEE conference on robotics, automation and mechatronics (RAM) (pp. 225–230). IEEE.

Qian, J., Chen, K., Chen, Q., Yang, Y., Zhang, J., Chen, S., 2021. Robust visual-lidar simultaneous localization and mapping system for UAV. *IEEE Geosci. Remote Sens. Lett.*

Rajendran, G., Uma, V. and O'Brien, B., 2021. Unified robot task and motion planning with extended planner using ROS simulator. *J. King Saud Univ.-Comp. Inf. Sci.*

Ruan, X., Guo, P., Huang, J., 2022. A semantic OctoMap mapping method based on CBAM-PSPNet. *J. Web Eng.*, 879–910

Schultz, A.C., Zarzycki, A., 2021. The anthropomorphism of intelligence. *Technol. Architecture+ Design.* 5 (2).

Shao, Y., Gan, Z., Epifanovsky, E., Gilbert, A.T., Wormit, M., Kussmann, J., Lange, A.W., Behn, A., Deng, J., Feng, X., Ghosh, D., 2015. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* 113 (2), 184–215.

Shastri, A., Valecha, N., Bashirov, E., Tataria, H., Lentmaier, M., Tufvesson, F., Rossi, M., Casari, P., 2022. A review of millimeter wave device-based localization and device-free sensing technologies and applications. *IEEE Commun. Surv. Tutorials.*

Shin, D.W., Ho, Y.S., Kim, E.S., 2019. Loop closure detection in simultaneous localization and mapping using descriptor from generative adversarial network. *J. Electron. Imaging* 28, (1) 013014.

Sodhi, P., Dexheimer, E., Mukadam, M., Anderson, S., Kaess, M., 2022, January. Leo: Learning energy-based models in factor graph optimization. In Conference on Robot Learning (pp. 234–244). PMLR.

Sturm, J., Engelhard, N., Endres, F., Burgard, W., Cremers, D., 2012. A Benchmark for the Evaluation of RGB-D SLAM Systems. In: 2012 IEEE/RSJ international conference on intelligent robots and systems. IEEE, pp. 573–580.

Tsintotas, K.A., Bampis, L., Gasteratos, A., 2022. The revisiting problem in simultaneous localization and mapping: a survey on visual loop closure detection. *IEEE Trans. Intell. Transp. Syst.*

Venator, M., Aklanoğlu, S., Bruns, E., Maier, A., 2021. Enhancing collaborative road scene reconstruction with unsupervised domain alignment. *Mach. Vis. Appl.* 32 (1), 1–16.

- Wang, J., Jia, X., Zhang, T., Ma, N., Meng, M.Q.H., 2021a. Deep neural network enhanced sampling-based path planning in 3D space. *IEEE Trans. Autom. Sci. Eng.*
- Wang, Y., Lou, Y., Zhang, Y., Song, W., Huang, F., Tu, Z., 2021c. A robust framework for simultaneous localization and mapping with multiple non-repetitive scanning lidars. *Remote Sensing*. 13 (10), 2015.
- Wang, K., Ma, S., Ren, F., Lu, J., 2021b. SBAS: Salient bundle adjustment for visual SLAM. *IEEE Trans. Instrum. Meas.* 70, 1–9.
- Xiao, L., Wang, J., Qiu, X., Rong, Z., Zou, X., 2019. Dynamic-SLAM: Semantic monocular visual localization and mapping based on deep learning in dynamic environment. *Rob. Auton. Syst.* 117, 1–16.
- Yu, J., Xue, H., Liu, B., Wang, Y., Zhu, S., Ding, M., 2020. Gan-based differential private image privacy protection framework for the internet of multimedia things. *Sensors* 21 (1), 58.
- Zhang, Y., Zhang, H., Xiong, Z., Sheng, X., 2019. A VisuAI SLAM System With LAser Assisted OptimizAtion. In: 2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM). IEEE, pp. 187–192.
- Zhao, C., Tang, Y., Sun, Q., 2022. Unsupervised monocular depth estimation in highly complex environments. *IEEE Trans. Emerg. Topics Comput. Intelligence.*
- Zheng, X., Gan, H., Liu, X., Lin, W., Tang, P., 2022. 3D Point Cloud Mapping Based on Intensity Feature. In: *Artificial intelligence in China*. Springer, Singapore, pp. 514–521.