

Original Article

Visual Place Recognition Model using Deep Learning with Arithmetic Optimization Algorithm

S. Senthamizhselvi¹, A. Saravanan²

^{1,2}Department of Computer & Information Science, Annamalai University, Annamalai Nagar

¹Corresponding Author : selvikarthi2002@gmail.com

Received: 30 April 2023

Revised: 27 June 2023

Accepted: 17 July 2023

Published: 31 July 2023

Abstract - Visual Place Recognition (VPR) has engrossed the interest of many researchers in several domains like robotics and Computer Vision (CV). It is the process of identifying a location visited previously, depending on visual input like videos or images. DL techniques have witnessed a potential to solve this task. One method for VPR utilizing deep learning is Convolutional Neural Networks (CNNs) to extract features in visual input. The CNN is trained on a large dataset of videos or images, with each video or image corresponding to a diverse location. The CNN extract discriminative features for each location, allowing it to detect previously visited locations depending on their visual appearance. Therefore, this study presents an Arithmetic Optimization Algorithm with Deep Learning-Driven Robust Visual Place Recognition (AOADL-VPR) technique. The AOADL-VPR technique's purpose is to recognise the visual places using the DL model properly. In the AOADL-VPR technique, Gaussian Filtering (GF) based pre-processing is performed to remove the noise. Meanwhile, the MobileNet-v2 model is utilized for generating a feature vector set. Furthermore, the AOA is exploited to adjust the hyperparameter values of the MobileNet-v2 model. At last, Minkowski Distance is exploited for effectual similarity measurement between two images, thereby recognising the places. A series of experimental analyses can be performed to ensure the improved performance of the AOADL-VPR approach. The simulation outcomes portrayed the enhancements of the AOADL-VPR system on the place recognition process.

Keywords - Visual Place Recognition, Deep learning, Arithmetic Optimization Algorithm, MobileNetv2, Similarity measurement.

1. Introduction

The purpose of VPR is to assist a vision-based navigating system or robot in deciding if it finds early places visited [1]. It remains to be a challenging issue in the CV and robotics domain. Recently, these domains have surged VPR usage for many applications [2]. While VPR has gained significant interest and has been studied broadly in robotics and CV communities, many open problems remain. The problem with VPR is two-fold [3, 4]. Primarily, many prevailing VPR techniques are based on appearance [5]. However, the appearance of similar locations can drastically vary with various illumination situations, distance, standpoints, background clutter-first and occlusion, seasons [6], and false place recognition grants interference for localizing methods that reduce the precision and result in catastrophic localizing failure for navigating systems. Therefore, it could be challenging to recognize a similar location correctly once it goes through appearance changes [7, 8].

Initial visual descriptors are divided into global and local descriptors [9]. Local descriptors, by mining the feature near

all interest points, define the image that reveals robustness against viewpoint variations but undergoes appearance changes, whereas global descriptors generate one compact feature vector and define the image as a whole [10]; thus, it has benefits in illumination and appearance invariance but is ineffective in handling viewpoint changes [11]. A CoHOG impressive handcrafted feature-based work uses entropy mapping for extracting Regions of Interest (ROI). It utilizes the Histogram of Oriented Gradients (HOG) descriptor to produce cooperative regional depictions [12, 13]. Currently, the concentration on region-based VPR has changed to methods based on learning [14, 15], particularly CNN, because of its most significant achievements in classification and image retrieval. The great potential of joining CNN methods and local region descriptors is confirmed by the preliminary outcomes in the VPR [16]. However, CNN-based descriptors need more computing resources, like hardware-based acceleration utilizing a GPU unsuitable for resource-limited devices [17, 18].

This study presents an Arithmetic Optimization Algorithm with Deep Learning-Driven Robust Visual Place



Recognition (AOADL-VPR) technique. The AOADL-VPR technique's purpose is to recognise the visual places using the DL model properly. In the AOADL-VPR technique, Gaussian Filtering (GF) based pre-processing is performed to remove the noise. Meanwhile, the MobileNet-v2 model is utilized for generating a feature vector set. Furthermore, the AOA is exploited to adjust the hyperparameter values of the MobileNet-v2 model. At last, Minkowski Distance is exploited for effectual similarity measurement between two images, thereby recognising the places. A series of experimental analyses can be performed to ensure the improved performance of the AOADL-VPR approach.

2. Related Works

Zaffar et al. [19] presented a training-independent computation effective model relevant to the HOG descriptor to attain current performance in VPR. The idea behind the CoHOG model is dependent upon complex scanning and area-related feature extraction employed by CNN. The technique achieved successful VPR in varying environments.

The authors [20] introduced a learning-based solution, a CNN, to make image-level representations invariant to weather and lighting. The author introduced a system that executes visual localization through only image-level representation calculated from a series of images. The author constructed the Gaussian Process Particle Filter structure, presenting two enhancements that enable localization while utilizing databases that cover large areas and strengthening the behaviour when dealing with incorrect initialization or particle deprivation of the filter. Eventually, the author developed two new general-purpose modules for convolution neural structures. Firstly, the author presented the CNN-COSFIRE module for the image recognition task. CNN-COSFIRE extends and adapts the COSFIRE structure for its inclusion in CNN architectures.

Schubert et al. [21] build a design on graph optimizing technique for recognizing where the graph has been utilized for modelling additional structural knowledge. A succeeding Non-Linear Least Squares Optimization (NLSQ) is used to enhance place recognition performance. This study addressed the high memory and extended run-time used to gain better place recognition performance faster on more significant issues. The author devised a new graph optimization process based on Iterated Conditional Modes (ICM).

In addition, the author inspected the improved cost function for a graph edge. Yang et al. [22] presented a simple yet potential technique named Multi-Scale Sliding Window (MSW) to generate landmarks to enhance the place-recognizing outcome. Unlike traditional techniques of generation of the landmark that is commonly based on identifying objects whose size dispersions were not equal and so may not be productive in realizing viewpoint

invariance and shift-invariance. Xie et al. [23] modelled hybrid networking to capture the point cloud descriptors and images to sort out the place detection issue. This contribution has been shortened as (1) framing a compact fusion structure that captured both the intense depiction of the imagery and 3D point cloud, adopting the cut-short approach in point cloud global factor combination to enrich the detection outcome, (2) a studying appropriate metric to explain the resemblance of this fused global featuring. Islam et al. [24] aim to use Independent Component Analysis (ICA) and AE Convolutional to comprehend the route over the robot.

3. The Proposed Model

The present article introduces a novel AOADL-VPR approach to accurately and automatically recognize visual places. The purpose of the AOADL-VPR approach is to recognise the visual places using the DL model appropriately. In the AOADL-VPR technique, four sub-processes are involved: GF-based noise elimination, MobileNet-v2 feature extractor, AOA-based hyperparameter tuning, and Minkowski Distance-based visual recognition. Fig. 1 demonstrates the comprehensive workflow of the AOADL-VPR technique.

3.1. GF Based Pre-processing

In this work, the GF removes the noise in the input images. It is a technique used for smoothing or blurring images. It is named after the Gaussian function, which is used as a kernel or filter for the operation. GF is a kind of linear filter; this implies that it executes a linear transformation to every pixel from the image [25]. The fundamental stages contained in GF are as follows:

- Select the size of the filtering kernel: A primary stage is to select the size of the kernel, which is classically a square/rectangular area centred nearby the pixel being filtered.
- Compute the weighted for filtering kernel: The next step is to estimate the weighted to the filter kernel dependent upon the Gaussian function. The weight can be estimated depending on the distance of every pixel in the centre of the kernels.
- Convolution of the filtering kernel with the image: The filtering kernel is then convoluted using a sliding window system. For every pixel from the image, the equivalent value in the filtered image can be computed as a weighted average of adjacent pixels.

It effectively eliminates higher-frequency noise in an image but maintains the lower-frequency modules.

3.2. Feature Extraction using MobileNetv2 Model

Here, the MobileNetv2 model is applied for the extraction of feature vectors. MobileNet is a CNN-based model broadly exploited for categorizing images [26, 27].

The primary benefit of utilizing the MobileNet structure is that method requires somewhat lesser computational work than the typical CNN method, which generates the appropriate for functioning on a mobile device and computer with lesser computational abilities. The MobileNet technique is a basic design that takes part of the convolution layer to distinguish the feature that depends on two managing features that switch between the parameter accuracy and latency effectually. The MobileNet technique seems to help diminish the network dimension. MobileNet structure was equally effective with minimal feature counts like Palmprint Recognition. The simple design depends on various abstraction layers, the element of various convolutions that implemented that quantized configuration that calculates a common convolution issue in-depth. The convolution of 1x1 was termed pointwise convolution. To create in-depth, platforms can take extraction layer with infrastructures in-depth and point with typical ReLU.

The feature vector mapping of dimensional $F_m \times F_m$ and the filtering is dimensional $F_s \times F_s$ input variable was represented by p , and the resultant variable can be

represented as q . For the essential extracting layers of design, a total computation work can be defined by the variable c_e and is evaluated by the subsequent in Eq. (1):

$$c_e = F_s \cdot F_s \cdot \omega \cdot \alpha F_m \cdot \alpha F_m + \omega \cdot \rho \cdot \alpha F_m \cdot \alpha F_m \quad (1)$$

The multiplier value was context-precise, and for a research study, the value of multipliers ω is assumed between the limit $[1 - n]$. The significance of the variable resolution's multiplier distinguished by α is that 1. The calculational works can be accepted by the variable $cost$ evaluated by Eq. (2) as:

$$cost_e = F_s \cdot F_s \cdot \omega \cdot \rho \cdot F_m \cdot F_m \quad (2)$$

The projected system integrates the pointwise and depth-wise convolutional and can be limited by the variable of depletion represented by the variable d is evaluated by Eq. (3) as:

$$d = \frac{F_s \cdot F_s \cdot \omega \cdot \alpha F_m \cdot \alpha F_m + \omega \cdot \rho \cdot \alpha F_m \cdot \alpha F_m}{F_s \cdot F_s \cdot \omega \cdot \rho \cdot F_m \cdot F_m} \quad (3)$$

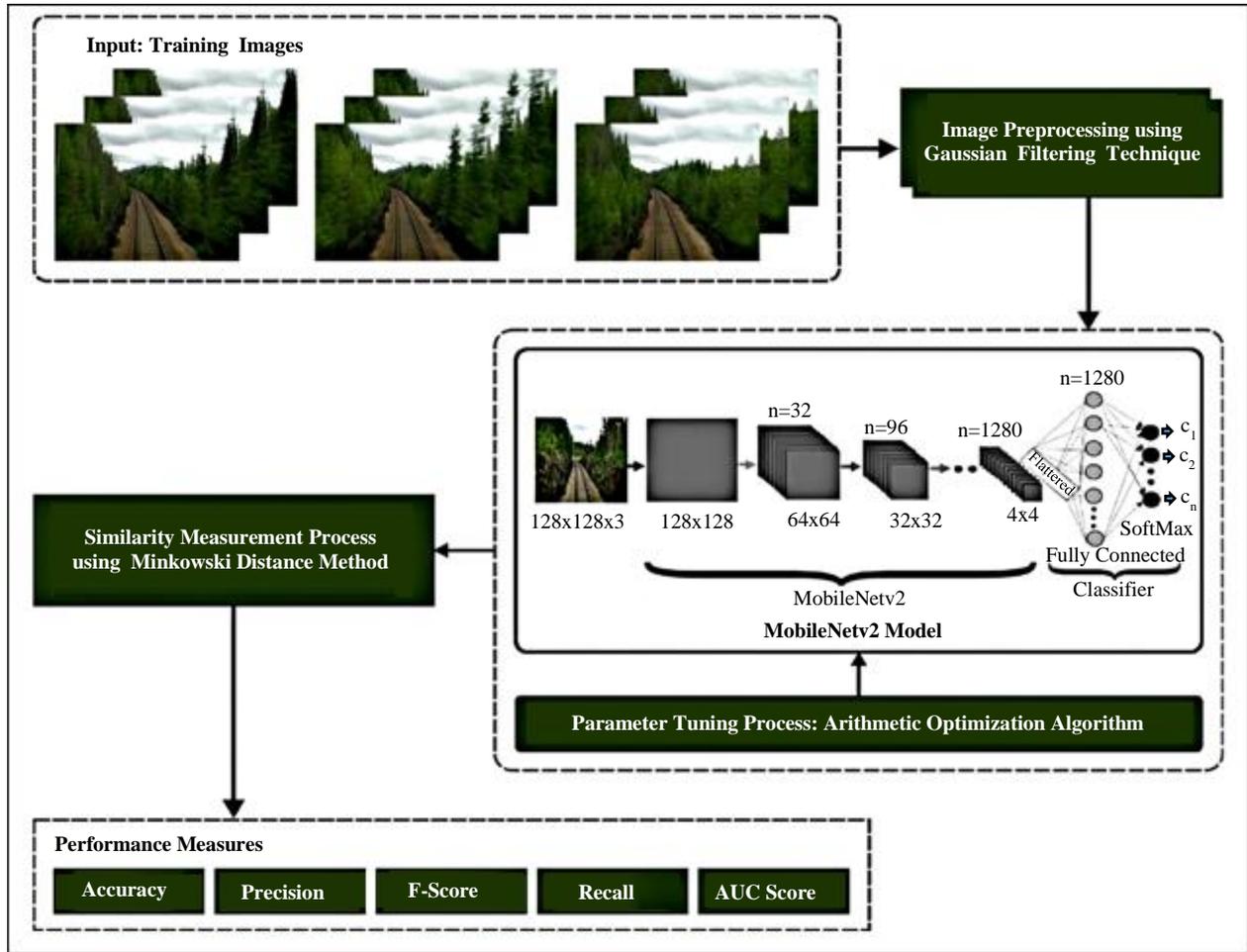


Fig. 1 Comprehensive procedure of the AOADL-VPR approach

During the presented method, the input dimensional of images is 224x224x3. The first two values (224x224) represent the height and width of images. Such values are continuously superior to 32. The third value proposes that it take three input channels. The presented structure consists of 32 filters; the filter size was 3x3x3x32.

The rule under the MobileNet structures is to additional complex convolution layers, in which every layer contains a convolution layer of dimensional 3x3 that buffers the input dataset, follow by a convolution layer of dimensional 1x1 pointwise which integrates these filter parameters to create a novel element. This model simplifies and makes it more significant than the standard convolution method.

3.3. Hyperparameter Tuning

The AOA is used to modify the hyperparameter values of the MobileNetv2 model optimally. AOA is an MH method that relies on a primary function to search for the optimum solution [28]. Like other MH approaches, it initiates by the nearly optimal or best-obtained solution and the random number of candidate alternatives (X). For the AOA to start working, the search phase must be chosen first (exploitation or exploration). In the succeeding search stage, the Math Optimizer Accelerated (MOA) was applied and determined by expression as follows. Fig. 2 portrays the steps comprised in AOA.

$$MOA(t) = \text{Min} + t \times \left(\frac{\text{Max} - \text{Min}}{T} \right) \quad (4)$$

The parameter t indicates the existing recurrence and ranges from 1 to the maximal admissible amount of epochs (T). The term indicates the lowest and most significant values of accelerating function, Max and Min. To determine an ideal choice, AOA's exploration agent examines the study scope at arbitrary positions across many regions, utilizing two primary search approaches (multiply and divide techniques defined in Eq. (5)).

$$x_{i,j}(t+1) = \begin{cases} Xb_j \div (M_{OP}) \times (UL_j \times \mu + LB_j), r2 > 0.5 \\ Xb_j \times M_{OP} \times (UL_j \times \mu + LB_j), \text{Otherwise} \end{cases} \quad (5)$$

In Eq. (5), $UL_j = UB_j - LB_j$. In these scenarios, $x_i(t+1)$ characterizes the i -th solution in the ensuing recurrence, $x_{i,j}(t)$ signifies the j -th position of the i th the solution from the newest iteration, and Xb_j signifies the j -th location in the optimum technique. e denotes a small integer number. The j -th place is minimal and maximal bounds are represented as UB_j and LB_j , correspondingly. The $\mu = 0.5$ procedure parameter regulates the behaviour of the search.

$$M_{OP}(t) = 1 - \frac{t^{\frac{1}{\alpha}}}{T^{\frac{1}{\alpha}}} \quad (6)$$

In Eq. (6), $M_{OP}(t)$ in Eq. (6) characterizes the probability of Math Optimizer (M_{OP}). The present iteration is characterized as t , whereas the overall amount of iterating is (T). The exploitation precision through iteration is determined by the sensitivity variable $\alpha = 5$.

It is essential to perform the exploitation phase by only studying if $r1$ is lesser than the present $MOA(t)$ quantity (Eq. (4)). In AOA, the exploitation operator (addition and subtraction) discovers the investigation scope intensely through some populated regions and method to generate a solution depends on two primary search methods (addition and subtraction) that are modelled in Eq. (7).

$$x_{i,j}(t+1) = \begin{cases} Xb_j - M_{OP} \times (UL_j \times \mu + LB_j), r3 > 0.5 \\ Xb_j + M_{OP} \times (UL_j \times \mu + LB_j), \text{Otherwise} \end{cases} \quad (7)$$

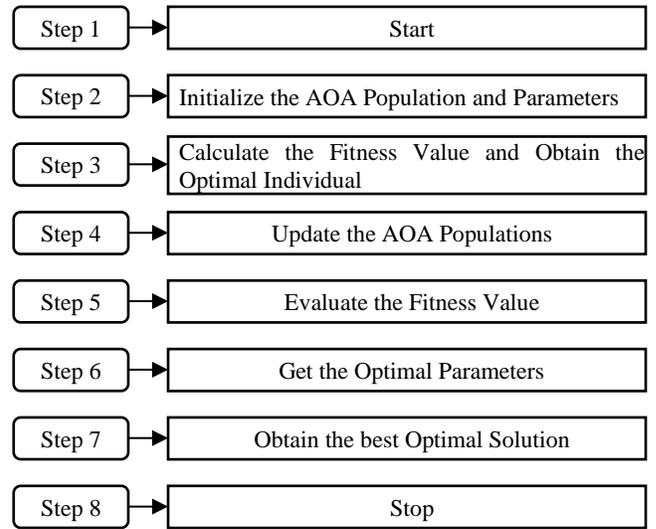


Fig. 2 Steps involved in AOA

3.4. Place Recognition Process

Finally, the Minkowski distance is applied to recognize places via a similarity measurement process. Minkowski Distance is the generalization method of Euclidean and Manhattan Distance, and the equation is given as:

$$D = [\sum_{i=1}^n |p_i - q_i|^p]^{1/p} \quad (8)$$

At this point, p signifies the sequence of norms, n implies the count of dimensional, and p_i and q_i Stand for the data points.

4. Results and Discussion

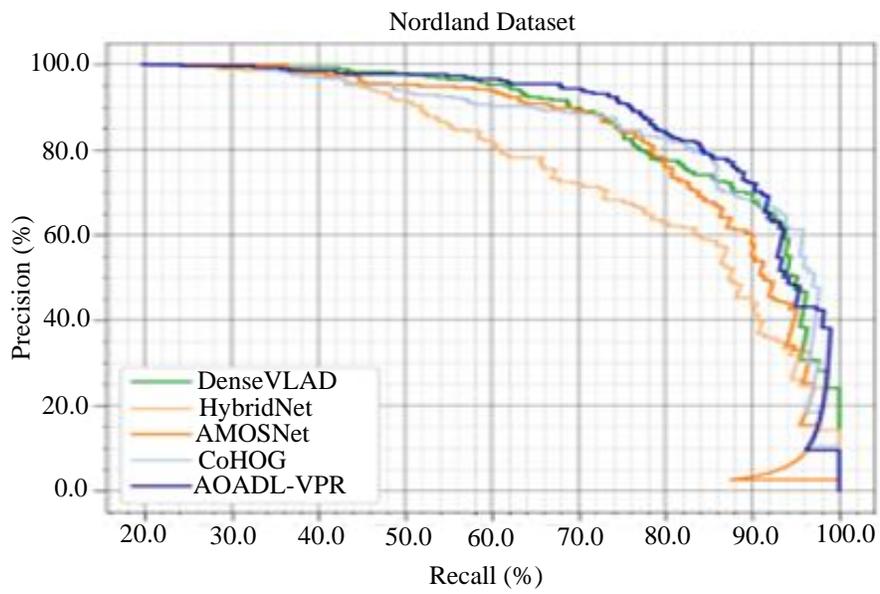
The investigational validation of the AOADL-VPR model is examined under discrete accomplishment measures. Figure 3 demonstrates the sample imaging. Table 1 and Figure 4 exhibit a complete precision-recall investigation of the AOADL-VPR method with the related techniques [29].

Table 1. Precision-recall examination of the AOADL-VPR model with present techniques under four datasets

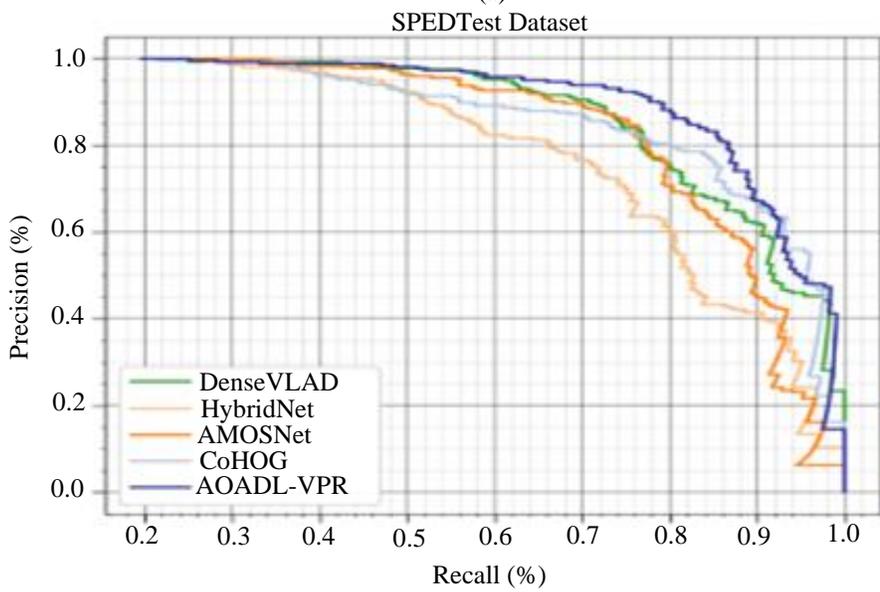
Recall (%)	Precision (%)				
	AOADL-VPR	CoHOG	AMOSNet	HybridNet	DenseVLAD
Nordland Dataset					
0	100.00	100.00	100.00	100.00	100.00
10	100.00	100.00	100.00	100.00	100.00
20	100.00	100.00	100.00	100.00	99.88
30	98.63	100.00	100.00	98.81	98.88
40	97.26	97.13	97.05	94.05	96.63
50	97.26	94.67	97.50	84.52	94.38
60	97.26	92.00	96.25	77.38	89.89
70	93.15	88.00	90.00	75.00	78.65
80	84.93	82.67	77.50	70.24	73.03
90	63.01	62.67	57.50	54.76	53.93
100	0.00	0.00	0.00	0.00	0.00
SPEDTest Dataset					
0	100.00	100.00	100.00	100.00	100.00
10	100.00	100.00	100.00	100.00	100.00
20	100.00	100.00	100.00	100.00	98.88
30	98.63	100.00	100.00	97.62	98.58
40	98.63	97.33	98.75	94.05	96.63
50	97.26	93.33	97.50	88.10	91.01
60	94.52	89.33	96.25	82.14	85.39
70	94.52	88.00	90.00	78.57	79.78
80	83.56	77.33	73.75	66.67	67.42
90	64.38	64.00	60.00	54.76	55.06
100	0.00	0.00	0.00	0.00	0.00
Synthia Night to Fall Dataset					
0	100	100.00	100.00	100.00	100.00
10	100	100.00	100.00	100.00	100.00
20	100	100.00	100.00	100.00	98.86
30	100	100.00	98.75	96.43	97.75
40	98.53	98.67	97.50	92.86	96.63
50	98.39	92.00	96.25	83.33	92.13
60	98.40	84.00	96.25	73.81	79.78
70	96.70	80.00	83.75	71.43	71.91
80	96.70	69.33	71.25	63.10	62.92
90	65.28	52.00	48.75	46.43	43.82
100	0.00	0.00	0.00	0.00	0.00
Living Room Dataset					
0	100.00	100.00	100.00	100.00	100.00
10	100.00	100.00	100.00	100.00	100.00
20	100.00	100.00	100.00	100.00	100.00
30	100.00	100.00	100.00	98.81	98.88
40	98.63	98.67	97.50	95.24	97.75
50	97.26	94.67	97.50	84.52	94.38
60	97.26	88.00	96.25	77.38	83.15
70	91.78	84.00	86.25	73.81	75.28
80	87.67	80.00	77.50	66.67	68.54
90	73.97	61.33	57.50	50.00	52.81
100	0.00	0.00	0.00	0.00	0.00



Fig. 3 Sample images



(a)



(b)

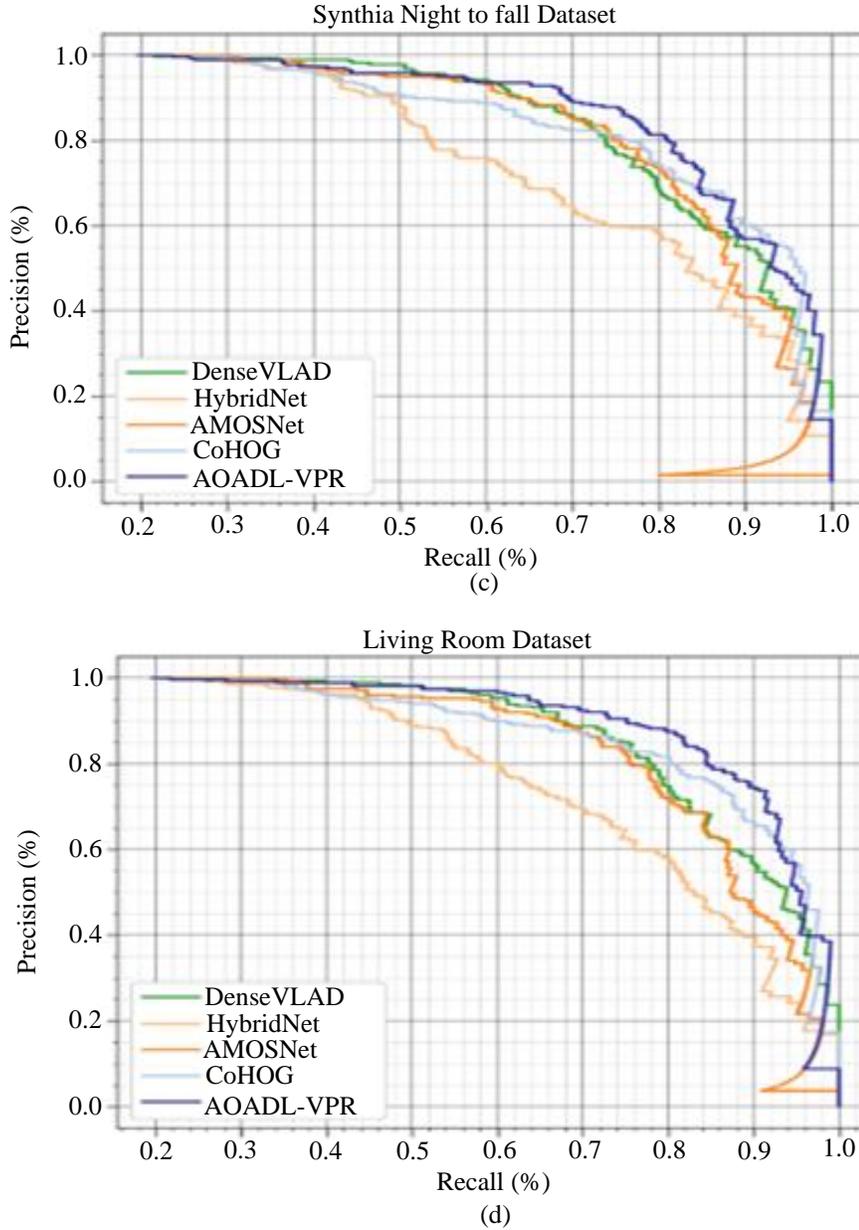


Fig. 4 Precision-recall examination of AOADL-VPR model with four datasets (a) Nordland, (b) SPED test, (c) Synthia night to fall, and (d) Living room

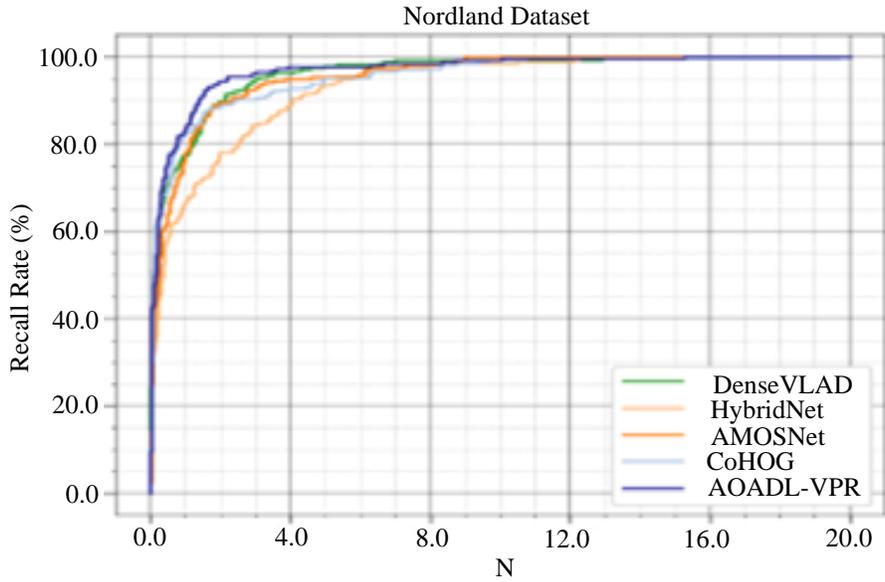
The investigational outputs stated that the AOADL-VPR method had depicted more significant outcomes. For instance, with $reca_l$ of 40%, the AOADL-VPR approach has given a superior $prec_n$ of 97.26%, whereas the CoHOG, AMOSNet, HybridNet, and DenseVLAD approaches have reached lesser $prec_n$ of 97.13%, 97.05%, 94.05%, and 96.63% subsequently. Followed by $reca_l$ of 70%, the AOADL-VPR approach has offered a maximal $prec_n$ of 93.15%, whereas the CoHOG, AMOSNet, HybridNet, and DenseVLAD approaches have reached lesser $prec_n$ of 88%, 90%, 75%, and 78.65% subsequently. In the meantime, with $reca_l$ of 90%, the AOADL-VPR approach has presented a superior $prec_n$ of 63.01%, whereas the CoHOG, AMOSNet,

HybridNet, and DenseVLAD approaches have accomplished lessened $prec_n$ of 62.67%, 57.50%, 54.76%, and 53.93% correspondingly.

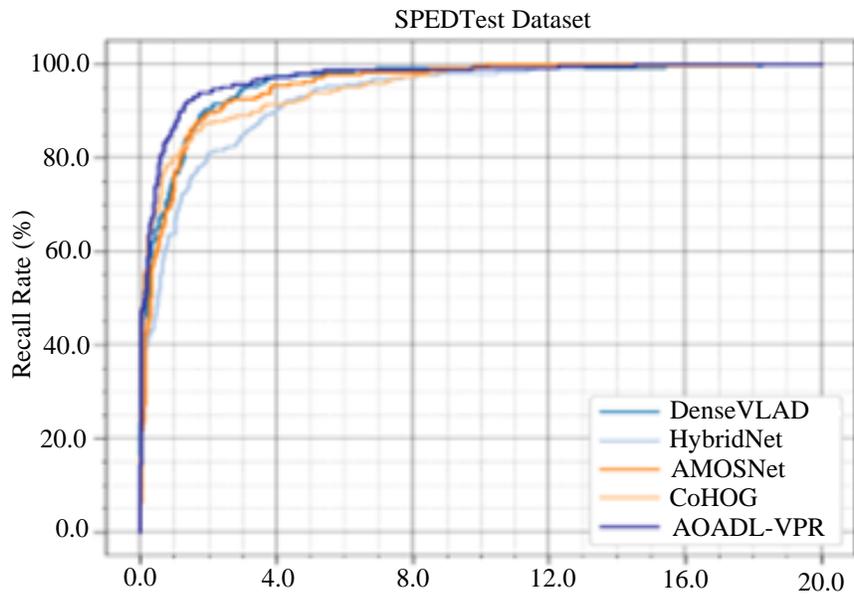
Table 2 and Fig. 5 exhibits a complete RR value of the AOADL-VPR algorithm with current techniques. The outcomes denoted that the AOADL-VPR algorithm has given an outcome in enhanced RR values. With the four steps and the Norland dataset, the AOADL-VPR approach has attained an enhanced value of RR of 93.15%. In contrast, the CoHOG, AMOSNet, HybridNet, and DenseVLAD approaches have reached reduced RR values of 88%, 91.25%, 73.81%, and 91.01%. Afterwards, with the 12 steps

and the Norland dataset, the AOADL-VPR method attained an enhanced RR value of 98.63%. In contrast, the CoHOG, AMOSNet, HybridNet, and DenseVLAD methods achieved reduced values of RR of 93.33%, 96.25%, 79.76%, and 96.63% subsequently. At the same time, with the four steps and SPEDTest dataset, the AOADL-VPR approaches have obtained an enhanced value of RR of 93.15%, whereas the CoHOG, AMOSNet, HybridNet, and DenseVLAD approaches reached lesser values of RR of 80%, 81.25%, 67.86%, and 76.40% respectively. Along with that, with the ten steps and SPEDTest dataset, the AOADL-VPR methodology has reached an improved value of RR of

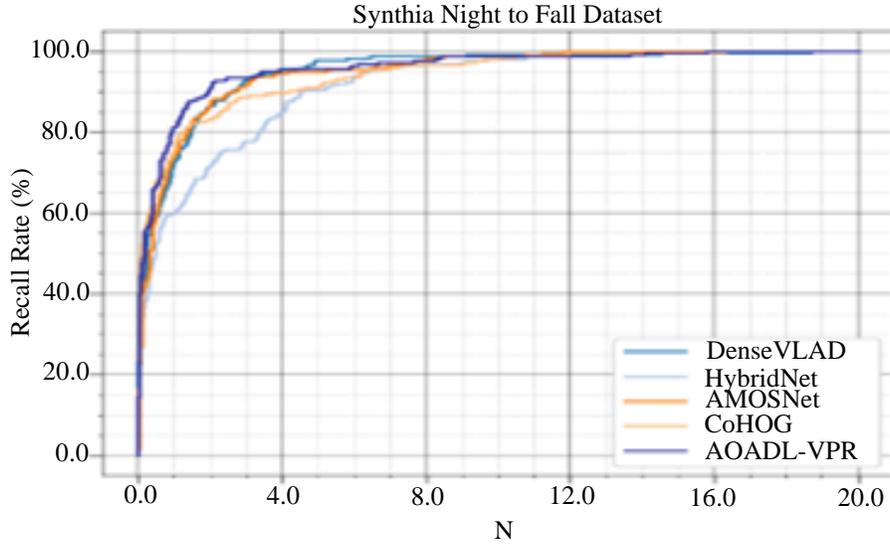
95.89%. In contrast, the CoHOG, AMOSNet, HybridNet, and DenseVLAD methodologies have reached reduced RR values of 85.33%, 85%, 73.81%, and 83.15% subsequently. Finally, with the 16 steps and SPEDTest dataset, the AOADL-VPR model has attained an improved value of RR of 100%. In contrast, the CoHOG, AMOSNet, HybridNet, and DenseVLAD models have reached reduced RR values of 90.67%, 93.75%, 77.38%, and 87.64% subsequently. Inclusive relative research of the AOADL-VPR method with recent approaches is made in Table 3. Figure 6 represents the AUC_{score} results of the AOADL-VPR method with current methods on the SPEDTest and Nordland datasets.



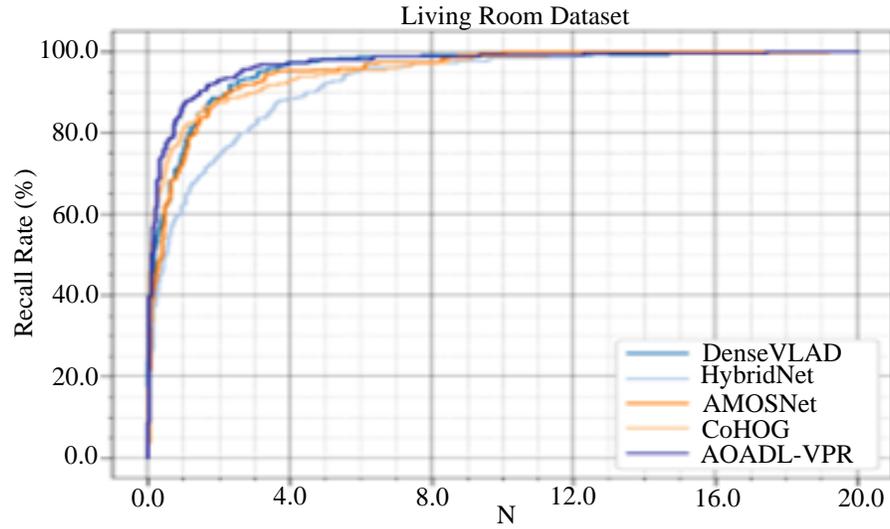
(a)



(b)



(c)



(d)

Fig. 5 Recall rate investigation of the AOADL-VPR model with four datasets (a) Nordland, (b) SPEDTest, (c) Synthia night to fall, and (d) Living room

The results indicate that the AOADL-VPR technique reaches increasing AUC_{score} Values. For instance, on the SPEDTest dataset, the AOADL-VPR technique obtains a higher AUC_{score} of 95.17%, whereas the CoHOG, AMOSNet, HybridNet, and DenseVLAD methods accomplish lesser AUC_{score} of 47.90%, 91.40%, 90.30%, and 84.80% correspondingly. Meanwhile, on the Nordland dataset, the AOADL-VPR technique attains a superior AUC_{score} of 79.62%, whereas the CoHOG, AMOSNet, HybridNet, and DenseVLAD techniques accomplish lesser AUC_{score} of 10.46%, 30%, 17.50%, and 12.95% correspondingly. Figure 7 signifies the AUC_{score} outcomes of the AOADL-VPR model with current systems on the

Living room and Synthia datasets. The results indicate that the AOADL-VPR model reaches increasing AUC_{score} Values. For instance, on the Living room dataset, the AOADL-VPR method obtains a superior AUC_{score} of 99.56%, whereas the CoHOG, AMOSNet, HybridNet, and DenseVLAD methods reached a reduced AUC_{score} of 85.50%, 98%, 97%, and 99.25% respectively. Meanwhile, on the Nordland dataset, the AOADL-VPR approach attains greater AUC_{score} of 99.35%, whereas the CoHOG, AMOSNet, HybridNet, and DenseVLAD approaches realize lesser AUC_{score} of 79.50%, 89%, 91.30%, and 98.80% respectively. These results show the enhanced recognition results of the AOADL-VPR technique.

Table 2. Recall rate investigation of the AOADL-VPR model with present techniques under four datasets

N (steps)	Recall Rate (%)				
	AOADL-VPR	CoHOG	AMOSNet	HybridNet	DenseVLAD
Nordland Dataset					
0	0.00	0.00	0.00	61.54	0.00
2	42.47	13.33	2.50	10.71	32.58
4	93.15	88.00	91.25	73.81	91.01
6	94.52	89.33	92.50	75.00	92.13
8	95.89	90.67	93.75	76.19	94.38
10	97.26	92.00	95.00	77.38	95.51
12	98.63	93.33	96.25	79.76	96.63
14	100.00	94.67	97.50	80.95	97.75
16	100.00	97.33	100.00	84.52	100.00
18	100.00	98.67	100.00	86.90	0.00
20	100.00	100.00	100.00	100.00	95.00
SPEDTest Dataset					
0	0.00	0.00	0.00	0.00	0.00
2	47.95	17.33	5.00	8.33	29.21
4	93.15	80.00	81.25	67.86	76.40
6	94.52	81.33	82.50	70.24	80.90
8	95.89	85.33	85.00	73.81	83.15
10	95.89	85.33	85.00	73.81	83.15
12	97.26	88.00	87.50	75.00	85.39
14	98.63	89.33	91.25	76.19	86.52
16	100.00	90.67	93.75	77.38	87.64
18	100.00	96.00	98.75	83.33	93.26
20	100.00	100.00	100.00	100.00	100.00
Synthia Night to Fall Dataset					
0	0.00	0.00	0.00	0.00	0.00
2	43.84	13.33	30.00	5.95	26.97
4	87.67	81.33	87.50	69.05	71.91
6	89.04	82.67	88.75	70.24	73.03
8	91.78	84.00	90.00	71.43	76.40
10	95.89	90.67	96.25	75.00	80.90
12	97.26	92.00	97.50	76.19	82.02
14	98.63	93.33	98.75	77.38	83.15
16	100.00	94.67	100.00	80.95	84.27
18	100.00	94.67	100.00	80.95	84.27
20	100.00	100.00	100.00	100.00	100.00
Living Room Dataset					
0	0.00	0.00	0.00	0.00	0.00
2	42.47	14.67	3.75	14.29	29.21
4	89.04	86.67	86.25	63.10	73.03
6	95.89	93.33	97.50	72.62	83.15
8	97.26	94.67	98.75	73.81	84.27
10	98.63	94.67	98.75	73.81	84.27
12	100.00	97.33	98.75	76.19	88.76
14	100.00	98.67	100.00	77.38	91.01
16	100.00	100.00	100.00	100.00	100.00
18	100.00	100.00	100.00	100.00	100.00
20	100.00	100.00	100.00	100.00	100.00

Table 3. AUC_{Score} analysis of the AOADL-VPR approach with other techniques

AUC Score (%)					
Datasets	AOADL-VPR	CoHOG	AMOSNet	HybridNet	DenseVLAD
SPED Test	95.17	47.90	91.40	90.30	84.80
Nordland	79.62	10.46	30.00	17.50	12.95
Living Room	99.56	85.50	98.00	97.00	99.25
Synthia	99.35	79.50	89.00	91.30	98.80

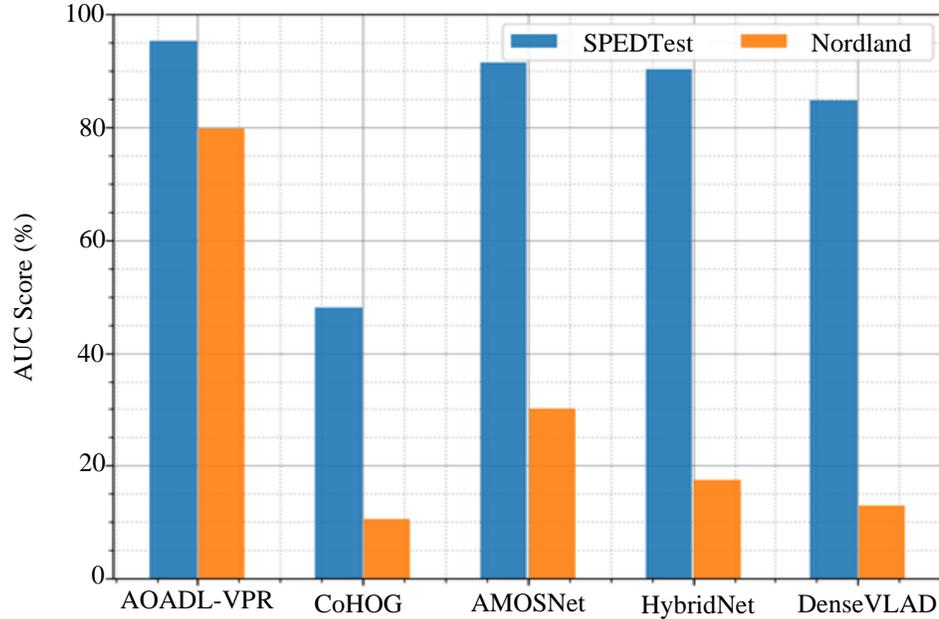


Fig. 6 AUC_{Score} investigation of the AOADL-VPR approach under SPEDTest and Nordland datasets

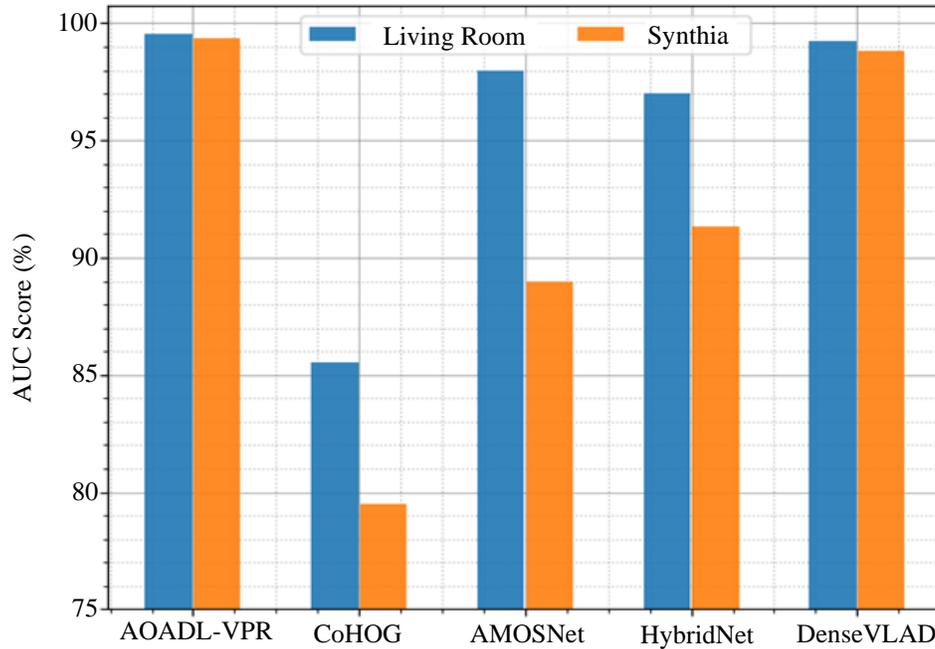


Fig. 7 AUC_{Score} investigation of the AOADL-VPR method under the datasets (a) Living room and (b) Synthia

5. Conclusion

The present article introduces a new AOADL-VPR technique to accurately and automatically recognize visual places. The AOADL-VPR technique's purpose is to recognise the visual places using the DL model properly. In the AOADL-VPR technique, four sub-processes are involved: GF-based noise elimination, MobileNet-v2 feature extractor, AOA-based hyperparameter tuning, and

Minkowski Distance-based visual recognition. Applying the AOA helps in the optimal hyperparameter selection of the MobileNet-v2 technique. A series of experimental analyses can be performed to ensure the enhanced accomplishment of the AOADL-VPR approach. The simulation outcomes portrayed the enhancements of the AOADL-VPR algorithm on the place-recognizing procedure. In the future, the AOADL-VPR system's recognition performance will be improved by using hybrid meta-heuristic optimizers.

References

- [1] Bruno Arcanjo et al., "An Efficient and Scalable Collection of Fly-Inspired Voting Units for Visual Place Recognition in Changing Environments," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2527-2534, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] Ahmad Khaliq et al., "A Holistic Visual Place Recognition Approach using Lightweight CNNs for Significant Viewpoint and Appearance Changes," *IEEE Transactions on Robotics*, vol. 36, no. 2, pp. 561-569, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Jurica Maltar, Ivan Marković, and Ivan Petrović, "Visual Place Recognition using Directed Acyclic Graph Association Measures and Mutual Information-Based Feature Selection," *Robotics and Autonomous Systems*, vol. 132, p. 103598, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Baifan Chen et al., "Hierarchical Visual Place Recognition Based on Semantic-Aggregation," *Applied Sciences*, vol. 11, no. 20, pp. 1-14, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] Dong Xu et al., "A Heterogeneous 3D Map-Based Place Recognition Solution using Virtual LiDAR and a Polar Grid Height Coding Image Descriptor," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 183, pp. 1-18, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Weiqi Zhang et al., "Learning Second-Order Statistics for Place Recognition Based on Robust Covariance Estimation of CNN Features," *Neurocomputing*, vol. 398, pp.197-208, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Bruno Ferrarini et al., "Binary Neural Networks for Memory-Efficient and Effective Visual Place Recognition in Changing Environments," *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2617-2631, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] Luis G Camara, and Libor Přeučil, "Visual Place Recognition by Spatial Matching of High-Level CNN Features," *Robotics and Autonomous Systems*, vol. 133, p. 103625, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Mihnea-Alexandru Tomiță et al., "Convsequential-Slam: A Sequence-Based, Training-Less Visual Place Recognition Technique for Changing Environments," *IEEE Access*, vol. 9, pp. 118673-118683, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Tariqul Islam, Sheikh Rabiul Islam, and Mahbubur Rahman, "Learning Condition-Invariant Scene Representations for Place Recognition across the Seasons using Auto-Encoder and ICA," *Journal of Electrical and Computer Engineering*, vol. 2022, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] T. Rajendran et al., "A Deep Learning Based Methodological Analysis for Breast Cancer Classification," *SSRG International Journal of Electronics and Communication Engineering*, vol. 10, no. 6, pp. 52-68, 2023. [[CrossRef](#)] [[Publisher Link](#)]
- [12] Kuniaki Noda et al., "Audio-Visual Speech Recognition using Deep Learning," *Applied Intelligence*, vol. 42, pp. 722-737, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] G. N. Srikanth, and M. K. Venkatesha, "SEOA DRN: Social Exponential Optimization Algorithm Based Deep Residual Network for Visual Speech Recognition," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 1, pp. 90-105, 2023. [[CrossRef](#)] [[Publisher Link](#)]
- [14] Anandh Nagarajan, and Gopinath M P, "Hybrid Optimization-Enabled Deep Learning for Indoor Object Detection and Distance Estimation to Assist Visually Impaired Persons," *Advances in Engineering Software*, vol. 176, p. 103362, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] M. Muruga Lakshmi, and S. Thayammal, "Ship Detection in Medium-Resolution SAR Images using Deep learning," *SSRG International Journal of Electronics and Communication Engineering*, vol. 8, no. 5, pp. 1-5, 2021. [[CrossRef](#)] [[Publisher Link](#)]
- [16] Sara Khosravi, and Abdolrah Chalechale, "Chimp Optimization Algorithm to Optimize a Convolutional Neural Network for Recognizing Persian/Arabic Handwritten Words," *Mathematical Problems in Engineering*, vol. 2022, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] S. Suma Christal Mary et al., "Selfish Herd Optimization with Improved Deep Learning Based Intrusion Detection for Secure Wireless Sensor Network," *SSRG International Journal of Electronics and Communication Engineering*, vol. 10, no. 4, pp. 1-8, 2023. [[CrossRef](#)] [[Publisher Link](#)]
- [18] M. Usharani et al., "An Optimized Deep Learning Model Based PV Fault Classification for Reliable Power Generation," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 9, no. 9, pp. 23-31, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [19] Mubariz Zaffar et al., “CoHog: A Light-Weight, Compute-Efficient, and Training-Free Visual Place Recognition Technique for Changing Environments,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1835-1842, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Manuel Lopez Antequera, “*Computer Vision Techniques for Calibration, Localization and Recognition*,” University of Groningen, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Stefan Schubert, Peer Neubert, and Peter Protzel, “Fast and Memory Efficient Graph Optimization via ICM for Visual Place Recognition,” *Robotics: Science and Systems*, 2021. [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Bo Yang et al., “Landmark Generation in Visual Place Recognition using Multi-Scale Sliding Window for Robotics,” *Applied Sciences*, vol. 9, no. 15, pp. 1-17, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Shaorong Xie et al., “Large-Scale Place Recognition Based on Camera-LiDAR Fused Descriptor,” *Sensors*, vol. 20, no. 10, pp. 1-21, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [24] Md. Tariqul Islam et al., “Convolutional Auto-Encoder and Independent Component Analysis Based Automatic Place Recognition for Moving Robot in Invariant Season Condition,” *Human-Centric Intelligent Systems*, vol. 3, pp. 13-24, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [25] V. Nyemeesha, and B. Mohammed Ismail, “Implementation of Noise and Hair Removals from Dermoscopy Images using Hybrid Gaussian Filter,” *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 10, pp. 1-10, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [26] Parvathaneni Naga Srinivasu et al., “Classification of Skin Disease using Deep Learning Neural Networks with MobileNet V2 and LSTM,” *Sensors*, vol. 21, no. 8, pp. 1-27, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Sagnik M, and Ramaprasad P, “Comparative Study of Convolutional Neural Networks,” *SSRG International Journal of Electronics and Communication Engineering*, vol. 6, no. 8, pp. 18-21, 2019. [[CrossRef](#)] [[Publisher Link](#)]
- [28] Mohamed Abd Elaziz et al., “Medical Image Classifications for 6G IoT-Enabled Smart Health Systems,” *Diagnostics*, vol. 13, no. 5, pp. 1-26, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [29] Mubariz Zaffar et al., “VPR-Bench: An Open-Source Visual Place Recognition Evaluation Framework with Quantifiable Viewpoint and Appearance Change,” *International Journal of Computer Vision*, vol. 129, no. 7, pp. 2136-2174, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]