JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Intelligent Visual Place Recognition using Seagull Optimization Algorithm with Deep Transfer Learning Model

[1]S. Senthamizhselvi and [2]A. Saravanan
[1]Research Scholar, Department of Computer and Information Sciences, Annamalai University.
[2]Assistant Professor / Programmer, Department of Computer and Information Sciences, Annamalai University.

Abstract

Visual place recognition (VPR) employing deep learning (DL) is a cutting-edge method that can significantly advance the domain of computer vision (CV). Leveraging deep neural networks (DNN), this technology allows machines to autonomously recognize and find particular landmarks or places within visual scenes with excellent speed and accuracy. VPR allows machines to recognize their environments, producing better decisions about their interactions and activities with the surroundings. This study designs an Intelligent Visual Place Recognition using Seagull Optimization Algorithm with Deep Transfer Learning (IVPR-SOADTL) method. The IVPR-SOADTL suggest an incorporated approach to improve performance of visual place recognition. We leverage a MixNet system for feature extraction, optimizing its hyperparameters employing the Seagull Optimization Algorithm (SOA), and use Manhattan Distance as the similarity measurement metric. For further improving the model's performance, we implements the SOA, a nature-inspired optimization method, to fine-tune the MixNet hyperparameters. Manhattan Distance is chosen as the similarity measurement metric for its capability to take both horizontal and vertical spatial relationships among feature vectors. We calculate our IVPR-SOADTL system on benchmark datasets and compare it against existing methods. The results exhibit that our combined architecture substantially increases accuracy of visual place recognition, outperforming existing approaches. An extensive comparison study stated the improved performance of the IVPR-SOADTL model over other methodologies.

Keywords: Visual places recognition; Transfer learning; Deep learning; MixNet model; Feature extraction; Seagull Optimization Algorithm

1. Introduction

In unstructured, dynamic and challenging atmospheres, the autonomous system functioning requires more localization ability robustness to grow up odometry faults [1]. The common method could be used to improve robustness utilizing a recognition engine that is a tool that utilizes revisited scene recognition to recover the robot's position in localization failure conditions or to resolve the expected odometry [2]. The classification engine is based upon a visual sensing technique which is classified as Visual Place Recognition (VPR). VPR method must have to show differences in environment, viewpoint and invariance over lighting to get improved robustness [3]. The trajectory condition with appearance variation due to different seasons of the year i.e. winter & summer or day & night periods has been raised VPR into challenging tasks in robotic vision. VPR-related visual clues are not uniformly distributed across an image, so focusing on essential areas was an ideal way to improve VPR performance that was opposed to irrelevant regions [4]. For example, detecting a street act using features derived from static structures such as road symbols could exist fraudulent data into place detection [5].

By inspiring such achievements, many researchers investigated the Convolutional Neural Network (CNN) of feature efficiency executed to the VPR problem [6]. From several layers of the CNN technique, the authors have enlarged features and then related them in comparison to numerous present structure-related VPR processes [7]. This shows that the CNN central layer overtakes any other method in the result of single image matching. The extended technology VPR executes crucial tasks in robot navigation owing to an object in which a single place endures significant presence dissimilarities for seasonal, weather and then illumination [8]. Recently, in Computer Vision (CV) the Deep Learning (DL) method has initiated a huge range of assessments on exactly how to make a feature representation from the CNNs which is vigorous to such kind of differences [9]. To improve the CNN performance, the CV community put efforts into constructing deeper & complex network structures as well as towards an improved insight into how the Daedalian structure executes in dissimilar occasions and stimuli [10].

In [11], a deep Distance Learning structure for VPR was designed. By studying numerous controls of distance association from VPR problems, the multiconstraint loss method was developed to minimize the distance constraint influences from the Euclidean space. The different kinds of CNN are likely AlexNet, VGGNet as well as other user-defined networks for deriving more individual features. In [12], the deep metric learning method is incorporated with enhancing removing feature as well as Similarity Metric could be used to train end-wise networks similarly for place recognition performance and handling & altering the event over time. A dynamic enhanced SM was developed to support the discrimination ability as well as to figure out the similarity betwixt descriptors of appearance pairs which may be removed in a CNN.

Park et al. [13] projected a new and trivial CNN method for the VPR approach. The designed method exactly focuses on the embedding technique. To decrease the computational effort of the network, the authors have proposed an FCN system with some layers and filters. Instantly, the developed network absorbs a vector space while their distance is comparable to place similarity with DML. Mao et al. [14] proposed a novel method to build up a multi-scale feature pyramid & project 2 approaches for enhancing place recognition capability. The initial process is used for gaining a unique mapping feature which is both local and semi global existence. In the second method, learn an attention process in the feature pyramid to weight up the three-dimensional grid on new mapping features. Both methods have integrated the multiscale features from the pyramid to destroy the local feature however, notify the harms from 2 various methods.

Zhu et al. [15] proposed an original method which depends on CNN, the place image has to pre-train network structure for gaining learned image descriptor, and with numerous roles of binarization, fusion and pooling to optimize them. After that, place recognition similarity is predictable with the Hamming distance of place sequences. Pei et al. [16] determined a fast as well as dependable method for employing BoW of structural lines established by Extended Line Band Descriptors that includes Pose evaluation capability, which is termed ELBDP. In addition to that, the robust structure only employs one fixed of threatened facts and a single structural line to estimate the comparative pose betwixt image pairs that were designed.

This study designed an Intelligent Visual Place Recognition using Seagull Optimization Algorithm with Deep Transfer Learning (IVPR-SOADTL) approach. The IVPR-SOADTL propose an integrated approach to enhance visual place recognition performance. We leverage a MixNet model for feature extraction, optimizing its hyperparameters using the Seagull Optimization Algorithm (SOA), and employ Manhattan Distance as the similarity measurement metric. Manhattan Distance is chosen as the similarity measurement metric for its ability to capture both horizontal and vertical spatial relationships between feature vectors. A wide-ranging comparative analysis indicated the enhanced performance of the IVPR-SOADTL system over other techniques.

2. The Proposed Model

In this study, we have developed an automated IVPR-SOADTL technique for the VPR. The introduced IVPR-SOADTL method automatically identifies the visual places accurately and efficiently. It involves a 3-stage method namely place recognition, parameter tuning, and feature extraction.

2.1. Feature Extraction: MixNet Model

MixNet is a deep neural network (DNN) framework developed to equalize computational performance and model accuracy in CV tasks, especially image classification [17]. It attains this with the help of a new technique named "mixed depthwise separable convolutions." In traditional convolutional neural networks (CNNs), convolutional layers utilize filters (kernels) for scanning input data and extracting features. Depthwise separable convolutions can be a highly effectual alternative, dividing the convolution into 2 phases: First one is depthwise convolution, which processes all input channels individually, and another one is pointwise convolution, which integrates the processed channels. MixNet moves to the next stage by combining various kernel depths and sizes within the depthwise separable convolutions.

Now, simplified mathematical formula to represent a depthwise separable convolution process:

$$\begin{aligned} &Depthwise\ Separable\ Convolution(x) \\ &= Pointwise\ Convolution\big(Depthwise\ Convolution(x)\big) \end{aligned} \qquad (1)$$

In MixNet, the "MixConv" method integrates numerous convolutional filters with differing kernel sizes within a single layer. This enables the model to take features at various measures effectively, improving its capability for identifying patterns or objects in images. MixNet techniques were accessible in different sizes, permitting users to select the proper trade-off between computational efficiency and model complexity dependent upon their particular desires and hardware restraints. Generally, MixNet is confirmed efficient in accomplishing superior performance in image classification process when improving computational resources.

2.2. Hyperparameter Tuning: SOA Technique

The SOA has employed as a hyperparameter optimizer. SOA is a swarm intelligence approach stimulated by the biological behavior of seagull population namely attack and migration [18]. The improved and standard SOA is effectively resolved complicated power grid optimizations namely power quality improvement and economic planning, which shows the tremendous potential to resolve the problems of OAO algorithm.

The BSOA is effectively used for exploring the possible DG allocation scheme of OAO problems targeted at decreasing the DAEL and APL aims. BSOA complete the migration of seagull individual based on Eqs. (2)−(4), later performs the attack and foraging based on Eqs. (5)−(9), to attain the efficient location of seagull population.

$$Ds = \left|A^{*}P_{s}(k) + B^{*}\left(P_{best}(k) - P_{s}(k)\right)\right| \tag{2}$$

$$A = f_{c} - \left(\frac{k^{*}f_{c}}{k_{\max}}\right) \tag{3}$$

$$B = 2^{*}A^{2*}f_{rnd} \tag{4}$$

Now A shows the movement behavior of searching agent and B is to balance among exploration and exploitation. $P_s(k)$ represents the existing location of seagull individual at k^{th}iterations and P_{best} indicates the optimum location of seagull population. $k_{\max}$ denotes the number of maximal iteration and f_c is accountable for the dynamic change of A. f_{rand} refers to random integer in [0,1].

$$x = R_{sp} * \cos\theta \tag{5}$$

$$y = R_{sp} * \sin\theta \tag{6}$$

$$z = R_{sp} * \theta \tag{7}$$

$$R_{sp} = u * e^{\theta v} \tag{8}$$

$$P_{s}(k) = Ds * x * y * z + P_{best}(k) \tag{9}$$

Individual seagull perform spiral attack behaviors by continuously varying the angle θ and radius R_{sp} after completing migration behavior. The location of individual seagull in spiral motion is represented by x, y and z. θ is a random value in $[0,2\pi]$ and the two coefficients viz., u and v, are used for controlling R_{sp}.

The use of BSOA to resolve OAO problem is demonstrated by simulation experiment on typical DNs. But in comparison to the present approaches, the performance of BSOA in resolving allocation problem of DN has room for additional improvements. Based on the features of OAO, ISOAE which incorporates the re-update process of weak seagull and the reserve population of elite seagull is introduced. Fig. 1 illustrates the flowchart of SOA.

While resolving the OAO problem, the optimization objectives like APL have been changed as the fitness of ISOAE to define the weak individual seagulls. Every individual seagulls viz., allocation system of OAO, such as access node, capacity, and power factor of N_{dg} DGs.

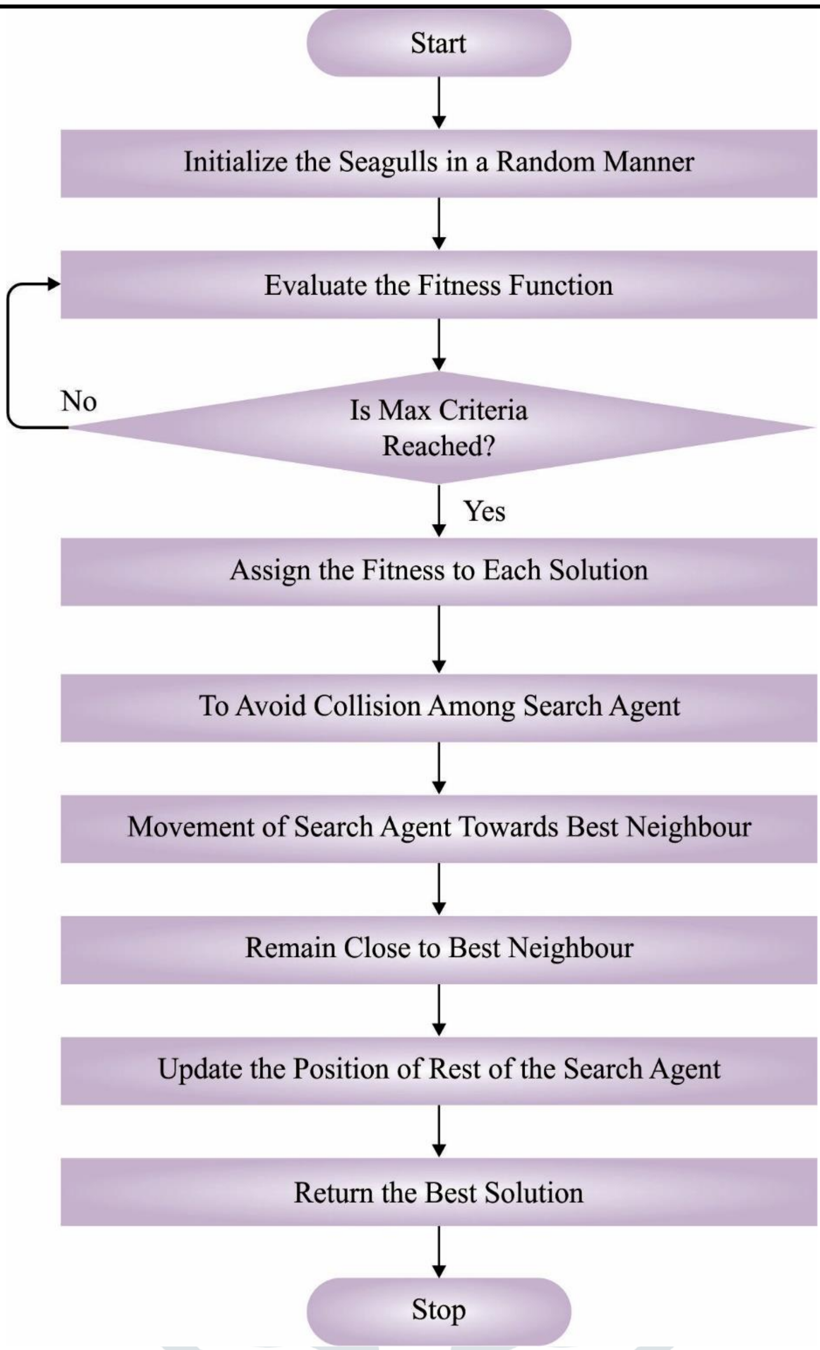

Fig. 1. Flowchart of SOA

Mainly, ISOAE update the inferior allocation scheme depends on the published elite scheme to attain a better DG allocation scheme than BSOA. Here, N_p is the size of seagull population and $10\%N_p$ individual seagulls with a great APL are described as inferior allocation schemes. The elite reserve population comprises potential DG access nodes. The inferior scheme selects the DG position from the elite reserve population and arbitrarily generates the DG capacity and power factor, thereby developing the re-exploration population of ISOAE. The updated population at the existing iteration can be defined by incorporating the re-exploration population and the residual seagull population, which removes lesser individuals. The method with minimum APL attained after the maximum iteration has been the final-adopted DG allocation system defined by the proposed ISOAE. The objective of ISOAE is to resolve the problem of OAO with APL reduction.

2.3. Similarity Measurement: Manhattan Distance

Manhattan Distance also called taxicab distance or L1 distance is an easy but, efficient metric to determine the dissimilarity or similarity between a 2 data points, frequently employed in different domains comprising ML, image analysis, and computer science [19]. It computes the distance among two points in a grid-like path, where the path contains right-angle movements (same as a taxi navigating city blocks). For two points (x1, y1) and (x2, y2), the Manhattan Distance is calculated by applying the given mathematical form:

$$Manhattan\ Distance = |x1 - x2| + |y1 - y2| \quad (10)$$

This equation measures the sum of the absolute variances among the related coordinates of a 2 points with the horizontal (x) and vertical (y) axes. Therefore, Manhattan Distance determines the "travel distance" required for moving from one point to another point by only moving at grid lines. This can be specifically beneficial while handling both features and data, which have been grid-like or when it is significant to consider vertical and horizontal movements independently. In applications such as recommendation or image processing systems, Manhattan Distance works as a robust computation for measuring similarity and dissimilarity among data points.

3. Results and Discussion

The simulated validation of the IVPR-SOADTL model is studied on different performance measurements. Fig. 2 demonstrates the sample images.

Fig. 2. Sample Images

Table 1 depicts the comparative AUC_{score} performance of IVPR-SOADTL algorithm with other recent methodologies on a 4 datasets [20].

Table 1 AUC_{score} analysis of IVPR-SOADTL approach with other methods under four datasets

AUC Score (%)					
Datasets	**IVPR-SOADTL**	**CoHOG**	**AMOSNet**	**HybridNet**	**DenseVLAD**
SPEDTest	96.61	47.90	91.40	90.30	84.80
Nordland	81.12	10.46	30.00	17.50	12.95
Living Room	99.62	85.50	98.00	97.00	99.25
Synthia	99.41	79.50	89.00	91.30	98.80

Fig. 3 illustrates the AUC_{score} analysis of IVPR-SOADTL algorithm with existing methods on SPEDTest database. The simulation values imply that IVPR-SOADTL approach has resulted in increased values of AUC_{score}. Based on AUC_{score}, the IVPR-SOADTL algorithm has gained increased value of AUC_{score} of 96.61% while the CoHOG, AMOSNet, HybridNet, and DenseVLAD approaches have attained less values of AUC_{score} of 47.90%, 91.40%, 90.30% and 84.80% respectively.

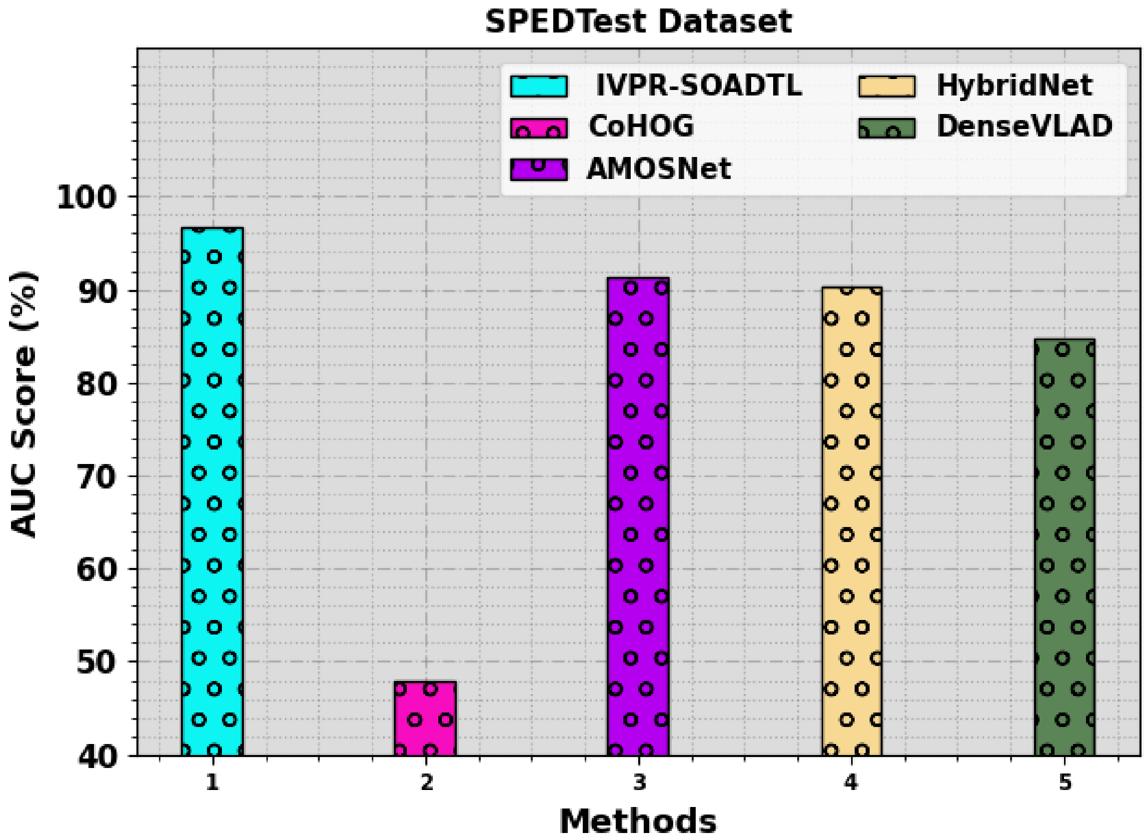

Fig. 3. AUC_{score} analysis of IVPR-SOADTL approach on SPEDTest dataset

Fig. 4 represents the AUC_{score} analysis of IVPR-SOADTL technique with existing models on Nordland dataset. The experimental values indicate that IVPR-SOADTL system has resulted in raised values of AUC_{score}. According to AUC_{score}, the IVPR-SOADTL algorithm has achieved an improved value of AUC_{score} of 81.12% whereas the CoHOG, AMOSNet, HybridNet, and DenseVLAD methodologies obtained lower values of AUC_{score} of 10.46%, 30.00%, 17.50% and 12.95% correspondingly.

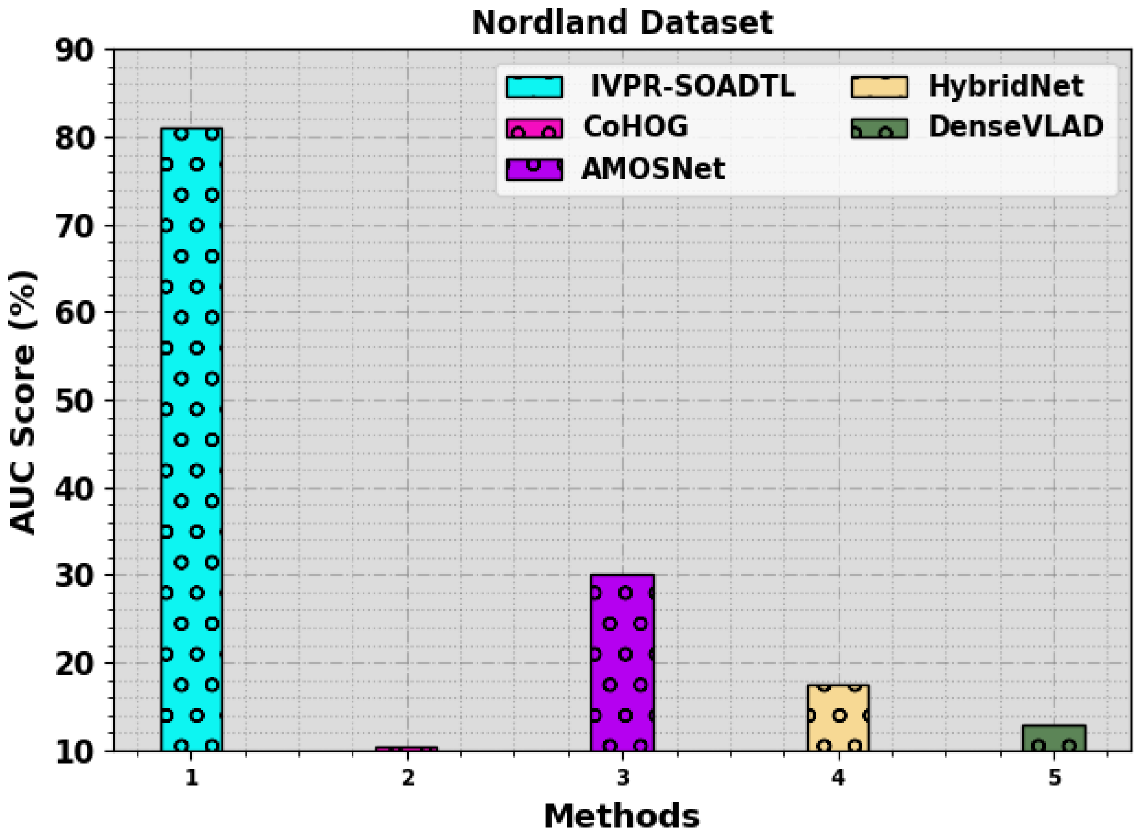

Fig. 4. AUC_{score} analysis of IVPR-SOADTL approach on Nordland dataset

Fig. 5 shows the AUC_{score} analysis of IVPR-SOADTL method with existing techniques on Living Room dataset. The outcome values denote that IVPR-SOADTL models have resulted in raised values of AUC_{score}. Additionally, Based on AUC_{score}, the IVPR-SOADTL approach has reached an improved value of AUC_{score} of 99.62% but the CoHOG, AMOSNet, HybridNet, and DenseVLAD systems acquired minimum values of AUC_{score} of 85.50%, 98.00%, 97.00% and 99.25% individually.

Fig. 6 exhibiting the AUC_{score} analysis of IVPR-SOADTL system with existing techniques on Synthia dataset. The simulation values represent that IVPR-SOADTL approach leads to improved values of AUC_{score}. Moreover, Based on AUC_{score}, the IVPR-SOADTL models have attained an increased value of AUC_{score} of 99.41% but the CoHOG, AMOSNet, HybridNet, and DenseVLAD techniques acquired minimum values of AUC_{score} of 79.50%, 89.00%, 91.30% and 98.80% respectively.

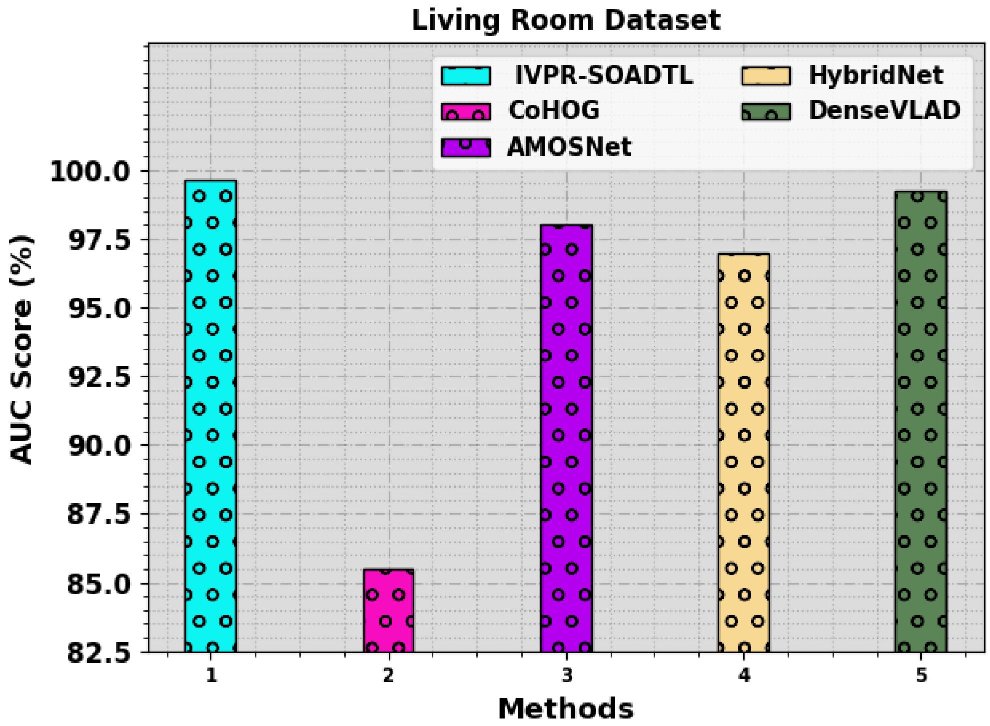

Fig. 5. AUC_{score} analysis of IVPR-SOADTL approach on Living room dataset

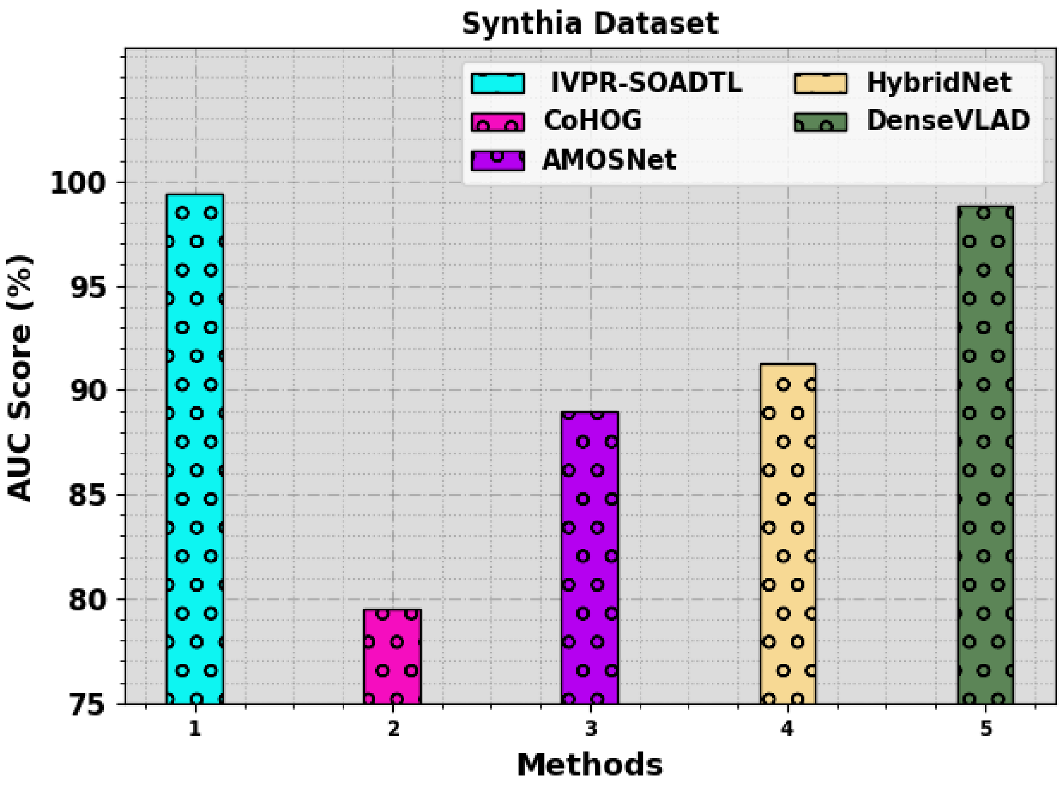

Fig. 6. AUC_{score} analysis of IVPR-SOADTL approach on Synthia dataset

4. Conclusion

In this study, a novel design and development of an Intelligent IVPR-SOADTL model. The IVPR-SOADTL propose an integrated approach to enhance visual place recognition performance. We leverage a MixNet model for feature extraction, optimizing its hyperparameters using the SOA, and employ Manhattan Distance as the similarity measurement metric. Manhattan Distance is chosen as the similarity measurement metric for its ability to capture both horizontal and vertical spatial relationships between feature vectors. A wide-ranging comparative analysis indicated the enhanced performance of the IVPR-SOADTL model over other methods.

References

[1] Zhang, X., Wang, L. and Su, Y., 2021. Visual place recognition: A survey from deep learning perspective. *Pattern Recognition*, *113*, p.107760.

[2] Garg, S., Suenderhauf, N. and Milford, M., 2022. Semantic–geometric visual place recognition: a new perspective for reconciling opposing views. *The International Journal of Robotics Research*, *41*(6), pp.573-598.

[3] Lee, A.J. and Kim, A., 2021, September. EventVLAD: Visual place recognition with reconstructed edges from event cameras. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 2247-2252). IEEE.

[4] Zaffar, M., Ehsan, S., Milford, M. and McDonald-Maier, K., 2020. Cohog: A light-weight, compute-efficient, and training-free visual place recognition technique for changing environments. *IEEE Robotics and Automation Letters*, *5*(2), pp.1835-1842.

[5] Hausler, S., Garg, S., Xu, M., Milford, M. and Fischer, T., 2021. Patch-netvlad: Multi-scale fusion of locally-global descriptors for place recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 14141-14152).

[6] Camara, L.G. and Přeučil, L., 2020. Visual place recognition by spatial matching of high-level CNN features. *Robotics and Autonomous Systems*, *133*, p.103625.

[7] Huang, G., Chen, A., Gao, H. and Yang, P., 2021, September. SMCN: Simplified mini-column network for visual place recognition. In *Journal of Physics: Conference Series* (Vol. 2024, No. 1, p. 012032). IOP Publishing.

[8] Masone, C. and Caputo, B., 2021. A survey on deep visual place recognition. *IEEE Access*, *9*, pp.19516-19547.

[9] Cait, K., Wang, B. and Lu, C.X., 2022, May. AutoPlace: Robust Place Recognition with Single-chip Automotive Radar. In *2022 International Conference on Robotics and Automation (ICRA)* (pp. 2222-2228). IEEE.

[10] Neubert, P., Schubert, S., Schlegel, K. and Protzel, P., 2021, July. Vector Semantic Representations as Descriptors for Visual Place Recognition. In *Robotics: Science and Systems*.

[11] Chen, L., Jin, S. and Xia, Z., 2021. Towards a Robust Visual Place Recognition in Large-Scale vSLAM Scenarios Based on a Deep Distance Learning. *Sensors*, *21*(1), p.310.

[12] Zhao, C., Ding, R. and Key, H.L., 2019, September. End-to-end visual place recognition based on deep metric learning and self-adaptively enhanced similarity metric. In *2019 IEEE International Conference on Image Processing (ICIP)* (pp. 275-279). IEEE.

[13] Park, C., Jang, J., Zhang, L. and Jung, J.I., 2018, January. Light-weight visual place recognition using convolutional neural network for mobile robots. In *2018 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-4). IEEE.

[14] Mao, J., Hu, X., He, X., Zhang, L., Wu, L. and Milford, M.J., 2018. Learning to fuse multiscale features for visual place recognition. *IEEE Access*, *7*, pp.5723-5735.

[15] Zhu, J., Ai, Y., Tian, B., Cao, D. and Scherer, S., 2018, June. Visual place recognition in long-term and large-scale environment based on CNN feature. In *2018 IEEE Intelligent Vehicles Symposium (IV)* (pp. 1679-1685). IEEE.

[16] Pei, L., Liu, K., Zou, D., Li, T., Wu, Q., Zhu, Y., Li, Y., He, Z., Chen, Y. and Sartori, D., 2019. IVPR: An instant visual place recognition approach based on structural lines in Manhattan world. *IEEE Transactions on Instrumentation and Measurement*, *69*(7), pp.4173-4187.

[17] Wang, Y., Yan, J., Yang, Z., Zhao, Y. and Liu, T., 2021. Optimizing GIS partial discharge pattern recognition in the ubiquitous power internet of things context: A MixNet deep learning model. *International Journal of Electrical Power & Energy Systems*, *125*, p.106484.

[18] Qian, J., Peng, Y., Zheng, H. and Wang, X., Application of Improved Seagull Optimization Algorithm on Optimal Allocation Optimizations of Distributed Generation.

[19] Sasikumar, P. and Sathiamoorthy, S., 2023. Optimal Densely Connected Networks with Pyramid Spatial Matching Scheme for Visual Place Recognition. In *Pervasive Computing and Social Networking* (pp. 123-137). Springer, Singapore.

[20] Zaffar, M., Garg, S., Milford, M., Kooij, J., Flynn, D., McDonald-Maier, K. and Ehsan, S., 2021. Vpr-bench: An open-source visual place recognition evaluation framework with quantifiable viewpoint and appearance change. *International Journal of Computer Vision*, *129*(7), pp.2136-2174.