SEAGULL OPTIMIZATION WITH DEEP LEARNING DRIVEN CONDITION INVARIANT VISUAL PLACE RECOGNITION MODEL

P. Sasikumar

Department of Computer and Information Science,
Annamalai University, Annamalai Nagar,
India
mailtoausasikumar@gmail.com

S. Sathiamoorthy

Department of Computer and Information Science,
Annamalai University, Annamalai Nagar,
India
ks_sathia@yahoo.com

Abstract

Visual place recognition (VPR) is most important topic in the computer vision (CV) and robotics community, whereas the task is for efficiently and accurately recognize the place of a provided query image. VPR still remains an open problem because of the several manners whereas the presence of real-world locations can change. As condition-invariant and viewpoint-invariant features were important features to long-term robust visual place-detection, the accomplishment of deep learning (DL) approaches from the CV domain triggered a kind of primary studies as its efficacy for VPR utilizing generic features in networks which are trained for another kind of detection tasks. In this aspect, this article develops a Seagull Optimization with Deep Learning Driven Condition Invariant Visual Place Recognition (SGODL-CIVPR) model. The presented SGODL-CIVPR model majorly involves the recognition of live places with respect to reference images. To accomplish this, the SGODL-CIVPR technique utilizes deep convolutional neural network based EfficientNet model to produce features, and the hyperparameter tuning process takes place using the SGO algorithm. Besides, convolutional wavelet neural network (CWNN) with deep convolutional generative adversarial network (DCGAN) is used to map features to low dimensional space for improving invariant properties and decrease dimensions concurrently. The performance validation of the SGODL-CIVPR model is tested using different datasets and the results demonstrate better performance over other methods.

***Keywords*: Visual place recognition; Deep learning; Computer vision; Seagull optimization; Condition invariant**

1. Introduction

Visual place recognition (VPR) represents the capability of a computer system to define whether it has formerly stayed in a location with the help of visual data. Carrying out reliable and very resilient VPR is a fundamental characteristic of autonomous robotic navigation as the Simultaneous Localization and Mapping (SLAM) system is depending on loop-closure mechanism for map corrections [1]. Even though the VPR problems are well established, still it remains challenging to reliably perform since there are a number of difficulties that should be addressed [2]. Initially, a revisited place might look completely distinct from once it was first seen and recorded because of many changing conditions such as distinct viewpoints, seasonal changes, dynamic elements, illumination levels, or any grouping of these factors [3]. Also, it is feasible for various places to look similar, particularly within similar environments, an error called perceptual aliasing [4]. Primarily applied for computer vision (CV) tasks, Convolution Neural Network (CNN) based model has been pioneered in the VPR fields during the past few years, obtaining remarkable achievements on different data sets [5].

Typically, VPR is developed as an image matching process that is classified into two steps. The initial step, which is called loop closure recognition, chooses candidate where map image is characterized by global descriptor and a matching procedure among the current robot view and the map images is implemented interms of image similarity [6, 7]. In case of substantial appearance variation caused by, for example, the day-night, dynamic objects, and season change, hand-engineered descriptor frequently fails to identify places as locally keypoint descriptor might considerably change with the condition-dependent appearance [8]. In recent times, researcher workers have developed to use CNN for extracting features for detecting loop closures in largescale environments. In the beginning, pre-trained classification CNN is used straightforwardly for extracting dense local feature map that serves as the visual feature for visual place representation [9]. But as a result of higher dimension and incapability to adapt to crowded environments, an end-to-end training method with a pooling layer and a feature extractor has been suggested, for example, NetVLAD, average pooling, max pooling, and generalized-mean pooling [10].

Keetha et al. [11] examined a new technique for deducing 2 key kinds of efficacy for VPR: the service of visual cues `specific' to environments, and to specific locations. The author utilizes contrastive learning principles for estimating either the environment- or place-specific efficacy of Vector of Locally Aggregated Descriptors (VLAD) cluster from unsupervised approach that is then utilized for guiding local feature corresponding with keypoint selective. Hong et al. [12] established an easy but overlooked baseline approach which equivalents the target label distribution by post-processed the model forecast trained by cross entropy loss and softmax function. While this technique exceeds recent approaches to benchmark datasets, it is enhanced by directly separating the source label distribution in the model forecast from the trained phase.

Gupta et al. [13] presented a model to create pseudo-concepts from the lack of true concept labels. The authors employ the pre-trained DCNN based structures whereas activation maps (filter responses) in convolution layers were assumed as primary cues to pseudo-concepts. The non-significant activation map was eliminated by employing the presented filter-specific threshold-based technique which lead to elimination of non-prominent models in data. The authors in [14] projected a new hybrid method which generates a higher efficiency primary equal hypothesis generator utilizing short learned sequential descriptor that allows selecting control sequential score aggregation utilizing single image learned descriptor. In [15], a new deep distance learning infrastructure for visual place detection was presented. But in-depth study of several constraints of distance connection from the visual place detection problems, the multi-constraint loss function was presented for optimizing the distance constraint connections from the Euclidean space.

This article develops a Seagull Optimization with Deep Learning Driven Condition Invariant Visual Place Recognition (SGODL-CIVPR) model. The presented SGODL-CIVPR model majorly involves the recognition of live places with respect to reference images. To accomplish this, the SGODL-CIVPR technique utilizes deep convolutional neural network based EfficientNet model to produce features, and the hyperparameter tuning process takes place using the SGO algorithm. Besides, convolutional wavelet neural network (CWNN) with deep convolutional generative adversarial network (DCGAN) is used to map features to a low dimensional space for improving invariant properties and decrease dimensions concurrently. The performance validation of the SGODL-CIVPR model is tested using different datasets.

2. The Proposed Model

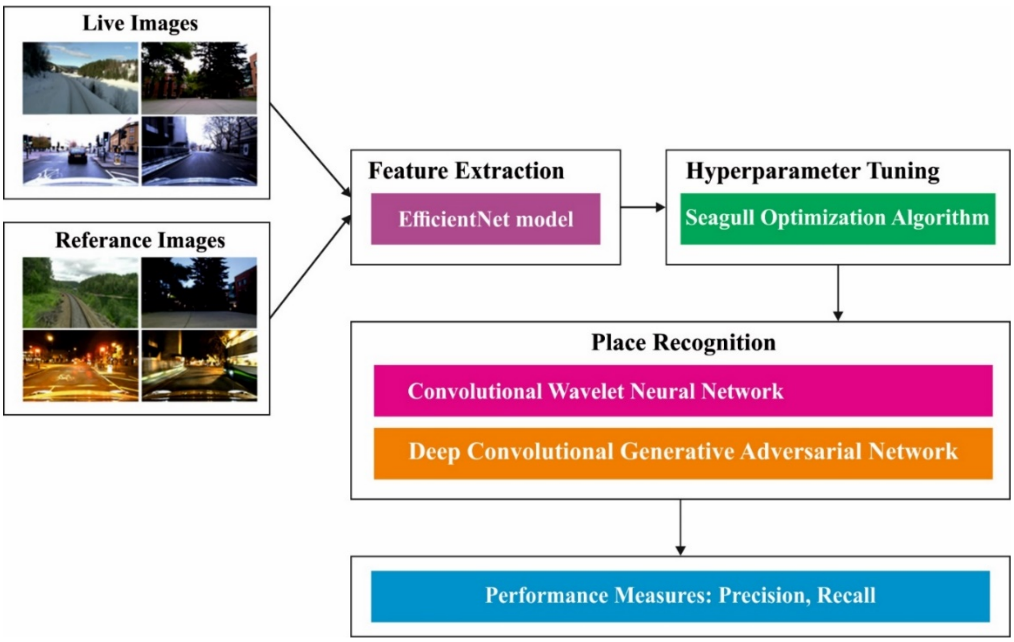

Fig. 1. Block diagram of SGODL-CIVPR approach

In this article, a novel SGODL-CIVPR system was introduced for the recognition of live places with respect to reference images. At the initial stage, the SGODL-CIVPR technique utilizes deep convolutional neural network based EfficientNet model to produce features, and the hyperparameter tuning process takes place using the SGO algorithm. Moreover, the CWNN with DCGAN is used to map features to low dimensional space for improving invariant properties and decreasing dimensions concurrently. Fig. 1 illustrates the block diagram of SGODL-CIVPR approach.

2.1. *Optimal EfficientNet based Feature Extraction*

In this study, the EfficientNet model is initially applied to produce feature maps [16]. CNN performance needs scaling from dimensional—depth, width, or resolution. Usually, to enhance model accuracy, the dimensional were investigated arbitrarily. Afterward, analysing the influence of scaling from distinct dimensional, their research suggests implementing a further principled method nearby model scaling—by scaling all the dimensions uniformly with set scaling coefficients. Afterward defining the set scaling coefficients, the baseline network was scaled for achieving superior accuracy and enhanced efficiency. While the performance of network dependent on not only scaling but also on the fundamental baseline network, it is also established a novel baseline network optimized for both accuracy and efficiency (FLOPS). The resultant structure utilizes a somewhat superior version of mobile inverted bottleneck convolutional (MBConv) that is then scaled up for obtaining the family of methods recognized as EfficientNets. But the ICML has illustrated the efficacy of utilizing EfficientNet-B0 and their enhanced accuracy on ImageNet dataset, it can aim for establishing that presented structure transmissions to skin lesion classifier dataset.

For improving the performance of the EfficientNet model, the SGO based hyperparameter tuning process is used in this study [17]. Seagulls are sea birds that could be projected complete world are termed as Laridae. The prey place can be computed by seagulls on the basis of knowledge. The migration and attacking strategies are major processes in the seagulls. Relocation can be described as a cyclical change of seagull around to calculate the richest and most plentiful energy sources and the behavior of seagulls are given below,

- Seagulls travel in a group all over relocation and may be varied in the earlier position to prevent accidents.
- They fly in a group t the way of optimum existence seagulls.
- The rest of the seagulls are upgraded with the position associated with the relevant seagulls.
- The mathematical modelling to attack the prey and migration is shown below,

Migration

In relocation method, the process starts with the seagull changing from one location to another location. During migration, the seagull needs to compensate for three conditions.

Avoid the collision: In SGO, the collision must be avoided amongst the neighbors, and the parameter is applied for computing optimum searching agent location.

$$Cs = A \times Ps(X) \tag{1}$$

In Eq. (20), A is represented as movement features of searching agent, X is represented as existing iteration, Ps is denoted as the existing location of searching agent, and Cs is denoted as searching agent location that didn't collision with residual searching agent and it is given in the following,

$$A = Fc - (X \times \left(\frac{FC}{Maxiteration}\right) \tag{2}$$

Where

$$X = 0_{,}1_{,}2_{,} \dots, t\ max\ iteration \tag{3}$$

Now, Fc is determined as value 2, A is decreased linearly from Fc to 0, and Fc represents frequency control of parameter.

The movement toward optimal neighboring direction:

While avoiding the Collison amongst neighbors, the searching agent is located near the movement of the optimal neighbors.

$$Ms = B \times \left(Pb(X) - PS(X)\right) \tag{4}$$

In Eq. (4), B is denoted as randomizes that is accountable for effectual balancing amongst exploration and exploitation, $Pb(X)$ is defined as optimal fit searching agent, and PS (X) is denoted as searching agent and Ms is represented as searching agent location.

$$B = 2 \times S = A^2 \times RD \tag{5}$$

In Eq. (5), RD denoted as an arbitrary value lies within [0, 1].

Remain nearer to the optimal searching agent:

Eventually, the searching agent could update the position based on the optimal searching agent that is shown below,

$$DS = |Cs + Ms| \tag{6}$$

In Eq. (6), DS represented the distance amongst the searching agents and better fit searching agents.

Attacking of prey

The exploitation aim is to exploit the history and experience of search method. During migration, the seagull changes the attack angle condition. They are maintaining their altitude based on wings and weight. While attacking the prey, the spiral drive features might occur in the air and it is formulated as,

$$X' = R * \cos(K) \tag{7}$$

$$Y' = R * \sin(K) \tag{8}$$

$$Z' = R * K \tag{9}$$

$$R = U * e^{KV} \tag{10}$$

Now, e refers to the base of natural logarithm, u and v indicated as spiral shape constant, k denote an arbitrary value within $[0 \leq k \leq 2\pi]$, and radius of each turn of spiral is represented as R. The updating procedure is calculated in the following,

$$Ps(X) = (DS * X' * Y' * Z') + Pbs(X) \tag{11}$$

In Eq. (11), $Pbs(X)$ is defined as an optimum solution and describes the position of other search agents. Fig. 2 displays the flowchart of SGO technique.

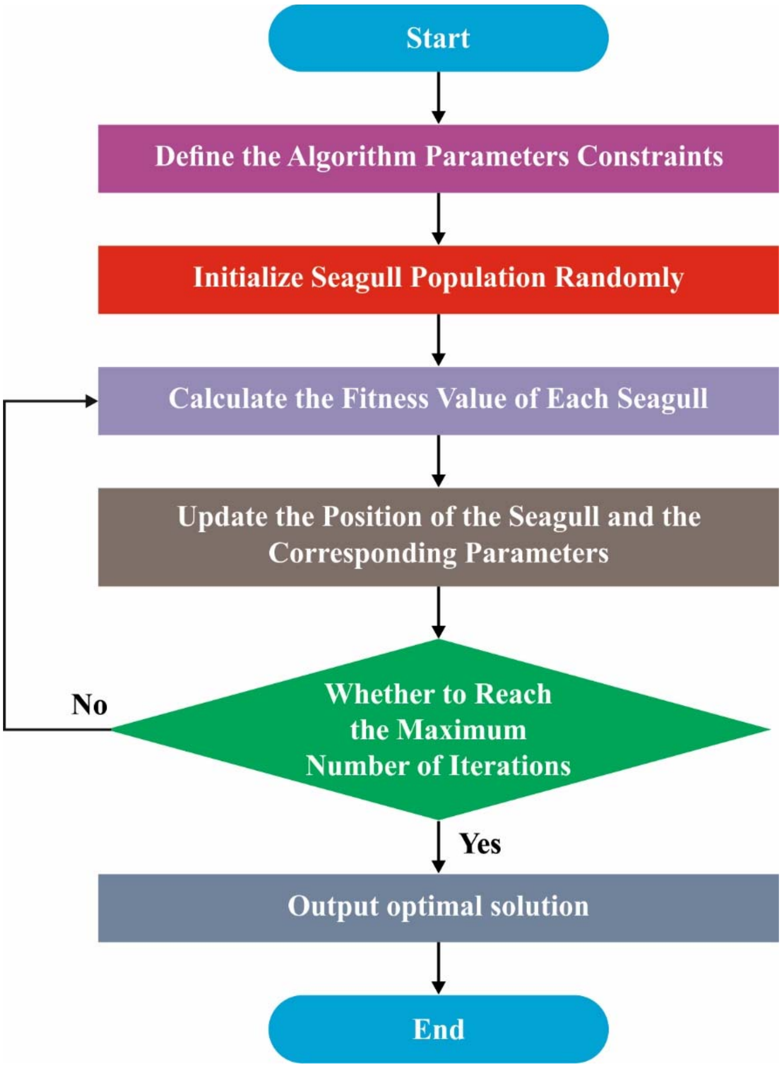

Fig. 2. Flowchart of SGO technique

2.2. *Design of CWNN-DCGAN Model*

Next, into feature extraction process, the CWNN with DCGAN is used to map features to a low dimensional space for improving invariant property and decreasing dimension concurrently [18]. A deep CWNN is introduced for VPR and it is increased by a DCGAN to improve the sample size for the class. The patch using $\lambda \times 2\lambda$ size is carefully chosen, which centered on the pixel of different class pseudo labels in an Image I_1. In the same way, Additional patch is attained in a similar position from the Image I_2. These two patches are concatenated to a novel patch using $2\lambda \times 2\lambda$ size as a pseudo label training instance of the changed class. According to the above method, each pseudo-label training sample is generated, signified as $\{P_{\tau_1}^{c}\}_{\tau_1=1}^{\tau_1=N_1}$ and $\{P_{\tau_2}^{uc}\}_{\tau_2=1}^{\tau_2=N_2}$, in which, $P_{\tau_1}^{c}, P_{\tau_2}^{uc} \in \mathbb{R}^{2\lambda\times 2\lambda}$, respectively represent the changed and unchanged patches. In the same way, hard patches $\{P_q^h\}_{q=1}^{q=N_h}$ with respect to hard pixels are obtained. Subsequently, a DCGAN takes place for enriching the pseudo label training sample of the changed class. The $D(.)$ discriminator is constructed for distinguishing fake images from real and false images and random noise is given to the generator $G(.)$ for generating false images. With an objective $max\text{-}min$ function, a convolution generative method $G(.)$ is constructed as follows:

$$V(G,D) = L^f + L^r \tag{12}$$

$$= E[\log(1 - D(G(\varphi)))] + E[\log(D(P_{\tau_1}^{c}))]$$

$$G^* = \arg\min_G \max_D V(G,D) \tag{13}$$

The trained generator G^* is takes place for expanding the dataset to $\{P_\tau^c\}_{\tau=1}^{\tau=N_T}$, in which $N_T - N_1$ image patch is produced as image patches. $2N_T$ instance is chosen for forming a training dataset $\{P_\tau^c, P_\tau^{uc}\}_{\tau=1}^{\tau=N_T}$,. Each pseudo-label training label and instance are given to the CWNN for place recognition.

3. Results and Discussion

This section inspects the performance validation of the SGODL-CIVPR model using Nordland, UACampus, and RobotCar datasets [19-21]. For verifying the entire efficiency of image descriptor, it can follow the general estimation metric that is dependent upon the top KNN amongst every database descriptor to query one. The equivalent is assumed effective if the correct equal occurs in top K nearest pairs. K will be fixed to 1, 5, and 10 from the experience. Precision-Recall is another key estimation metric from VPR. To provide corresponding pairs and threshold with respect to cosine similarity amongst image descriptors, it takes the amount of false positives, false negatives, and true positives.

Table 1 and Fig. 3 offer the overall VPR outcomes of the SGODL-CIVPR model on distinct results. The obtained values depicted that the SGODL-CIVPR model has demonstrated enhanced results under all datasets. For instance, with Nordland dataset, the SGODL-CIVPR model has attained average AP of 98.97, recall@1 of 97.60%, recall@5 of 99.65%, and recall@10 of 99.93%. At the same time, with UA Campus dataset, the SGODL-CIVPR approach has reached average AP of 99.93, recall@1 of 98.49%, recall@5 of 99.58%, and recall@10 of 99.73%. Moreover, with RobotCar (dbNight vs. qAutumn) dataset, the SGODL-CIVPR technique has accomplished average AP of 99.03, recall@1 of 94.06%, recall@5 of 96.80%, and recall@10 of 95.33%. Finally, with RobotCar (dbNight vs. qSnow) dataset, the SGODL-CIVPR methodology has achieved average AP of 98.60, recall@1 of 92.59%, recall@5 of 94.86%, and recall@10 of 95.24%.

No. of Runs	AP	Recall@1	Recall@5	Recall@10
Nordland				
Run-1	99.18	98.00	99.55	99.90
Run-2	98.93	97.07	99.53	99.92
Run-3	98.98	97.44	99.71	99.94
Run-4	99.17	97.67	99.70	99.96
Run-5	98.59	97.82	99.78	99.95
Average	98.97	97.60	99.65	99.93
UACampus				
Run-1	99.95	98.91	99.42	99.78
Run-2	99.94	98.02	99.63	99.65
Run-3	99.91	98.72	99.70	99.73
Run-4	99.96	98.40	99.55	99.72
Run-5	99.90	98.40	99.62	99.77
Average	99.93	98.49	99.58	99.73
RobotCar (dbNight vs. qAutumn)				
Run-1	99.15	94.95	96.14	94.98
Run-2	98.78	94.48	95.94	95.31
Run-3	99.07	94.36	97.26	95.12
Run-4	99.11	93.20	97.78	95.65
Run-5	99.06	93.31	96.89	95.58
Average	99.03	94.06	96.80	95.33
RobotCar (dbNight vs. qSnow)				
Run-1	98.27	92.10	95.75	95.32

Run-2	98.75	93.43	94.85	94.90
Run-3	98.77	91.81	93.71	95.62
Run-4	98.34	92.09	95.42	95.10
Run-5	98.86	93.50	94.56	95.28
Average	98.60	92.59	94.86	95.24

Table 1. Result analysis of SGODL-CIVPR approach with distinct runs and measures

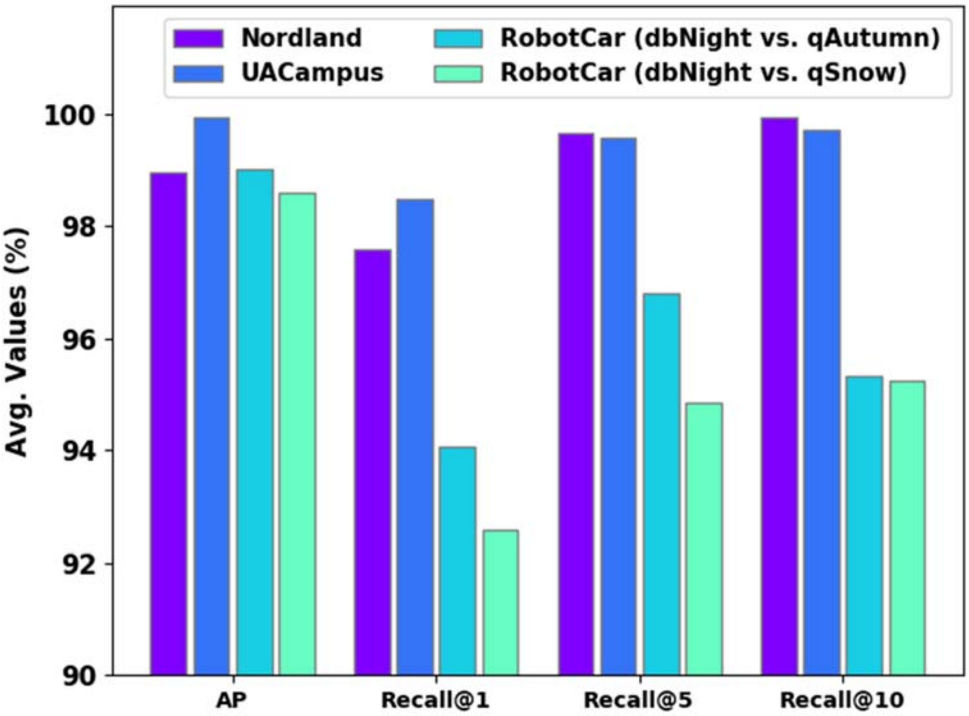

Fig. 3. Average analysis of SGODL-CIVPR approach with distinct measures

Table 2 and Fig. 4 demonstrate the comparative outcomes of the SGODL-CIVPR model on Nordland dataset [22]. The results indicated that the NetVLAD model has offered reduced performance over other models whereas the VGG16 model has accomplished certainly increased outcomes. Followed by, the AlexNet model has demonstrated moderately improved recognition efficiency whereas the CAE-AlexNet and CAE-VGG16 models have shown reasonable performance. However, the SGODL-CIVPR model has depicted effectual outcomes over other models with maximum AP, recall@1, recall@2, and recall@10 of 98.97%, 97.60%, 99.65%, and 99.93% respectively.

Nordland Dataset				
Methods	**AP**	**Recall@1**	**Recall@5**	**Recall@10**
SGODL-CIVPR	98.97	97.60	99.65	99.93
NetVLAD	40.25	27.25	47.26	57.64
AlexNet	95.57	90.94	97.50	99.17
VGG16	81.49	63.34	82.00	86.49
CAE-AlexNet	98.31	96.11	99.47	99.87
CAE-VGG16	97.88	94.59	99.01	99.74

Table 2. Comparative analysis of SGODL-CIVPR approach with existing algorithms under Nordland dataset

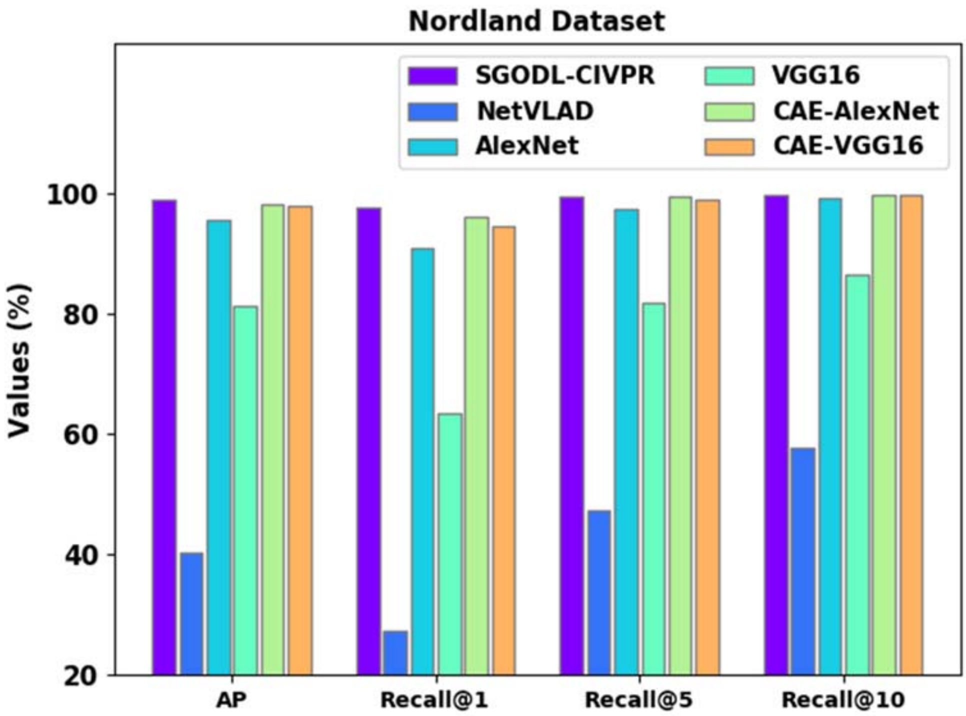

Fig. 4. Comparative analysis of SGODL-CIVPR approach under Nordland dataset

An obvious precision-recall investigation of the SGODL-CIVPR system on Nordland dataset is depicted in Fig. 5. The figure exposed that the SGODL-CIVPR method has resulted in higher values of precision-recall values under distinct classes.

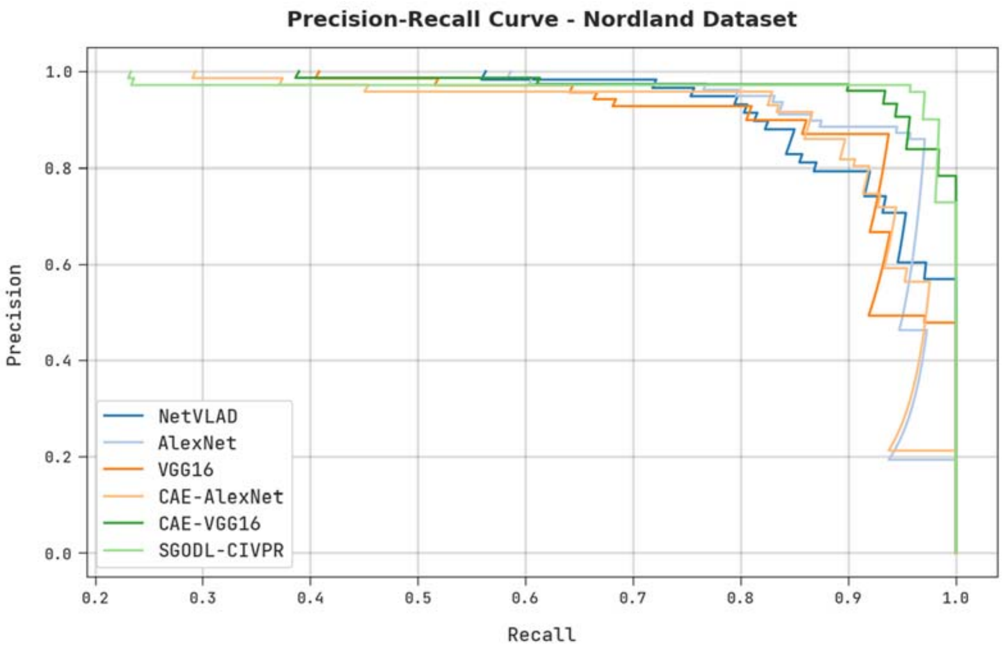

Fig. 5. Precision-recall analysis of SGODL-CIVPR approach under Nordland dataset

Table 3 and Fig. 6 demonstrate the comparative outcomes of the SGODL-CIVPR system on NetVLAD dataset. The results indicated that the NetVLAD approach has offered lower performance over other models whereas the VGG16 technique has attained certainly increased outcomes. Afterward, the AlexNet approach has established moderately improved recognition efficiency whereas the CAE-AlexNet and CAE-VGG16 algorithms have exhibited reasonable performance. But, the SGODL-CIVPR algorithm has showcased effectual outcome over other models with maximum AP, recall@1, recall@2, and recall@10 of 99.93%, 98.49%, 99.58%, and 99.73% respectively.

NetVLAD Dataset				
Methods	**AP**	**Recall@1**	**Recall@5**	**Recall@10**
SGODL-CIVPR	99.93	98.49	99.58	99.73
NetVLAD	74.45	67.47	78.75	83.75
AlexNet	99.33	93.30	97.48	98.51
VGG16	99.67	93.44	97.02	98.57
CAE-AlexNet	99.85	97.76	99.50	99.59
CAE-VGG16	99.92	96.80	99.26	99.53

Table 3. Comparative analysis of SGODL-CIVPR approach with existing algorithms under NetVLAD dataset

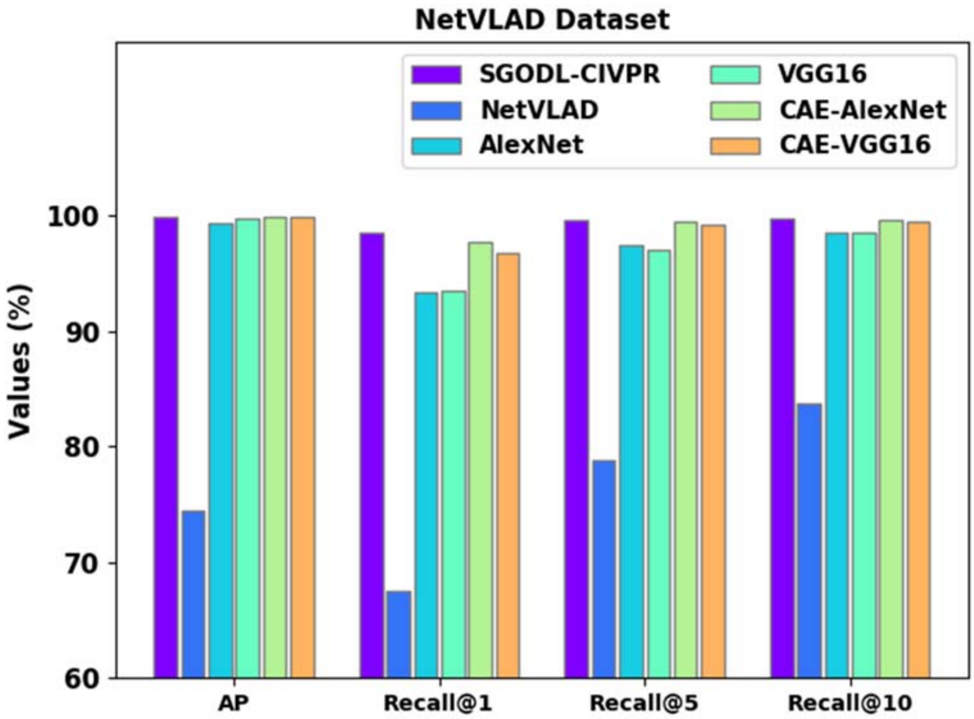

Fig. 6. Comparative analysis of SGODL-CIVPR approach under NetVLAD dataset

A clear precision-recall investigation of the SGODL-CIVPR system on NetVLAD dataset is depicted in Fig. 7. The figure stated that the SGODL-CIVPR approach has resulted in improved values of precision-recall values under various classes.

Table 4 and Fig. 8 demonstrate the comparative outcomes of the SGODL-CIVPR system on RobotCar (dbNight vs qAutumn) dataset. The results indicated that the NetVLAD technique has offered reduced performance over other models whereas the VGG16 approach has accomplished certainly increased outcomes. Followed by, the AlexNet algorithm has depicted moderately higher recognition efficiency whereas the CAE-AlexNet and CAE-VGG16 models have outperformed reasonable performance. Eventually, the SGODL-CIVPR model has depicted effectual outcomes over other approaches with higher AP, recall@1, recall@2, and recall@10 of 99.03%, 94.06%, 96.80%, and 95.33% correspondingly.

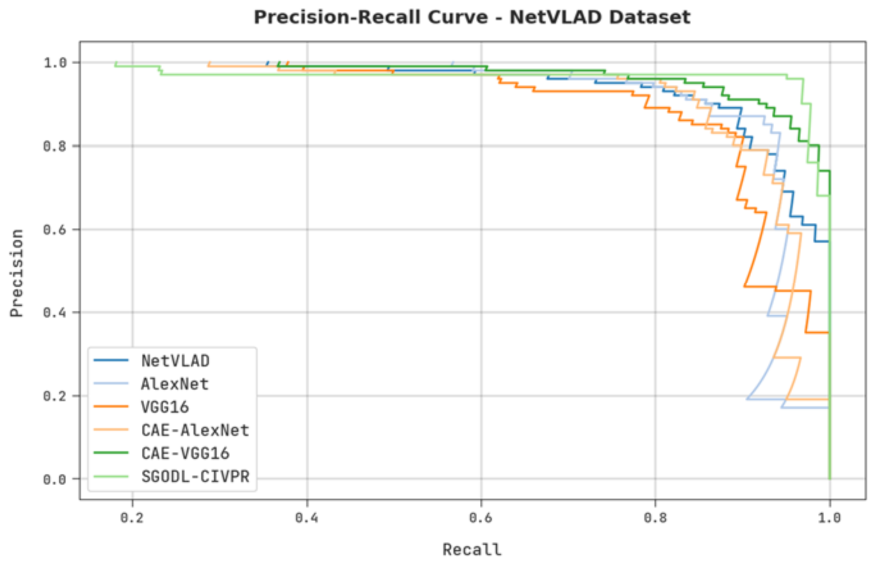

Fig. 7. Precision-recall analysis of SGODL-CIVPR approach under NetVLAD dataset

RobotCar (dbNight vs. qAutumn) Dataset				
Methods	**AP**	**Recall@1**	**Recall@5**	**Recall@10**
SGODL-CIVPR	99.03	94.06	96.80	95.33
NetVLAD	93.29	75.89	87.32	91.40
AlexNet	94.93	82.71	87.88	90.69
VGG16	86.53	64.33	71.95	75.32
CAE-AlexNet	95.10	83.25	88.48	90.21
CAE-VGG16	98.77	88.19	91.36	92.74

Table 4. Comparative analysis of SGODL-CIVPR approach with existing algorithms under RobotCar (dbNight vs. qAutumn) dataset

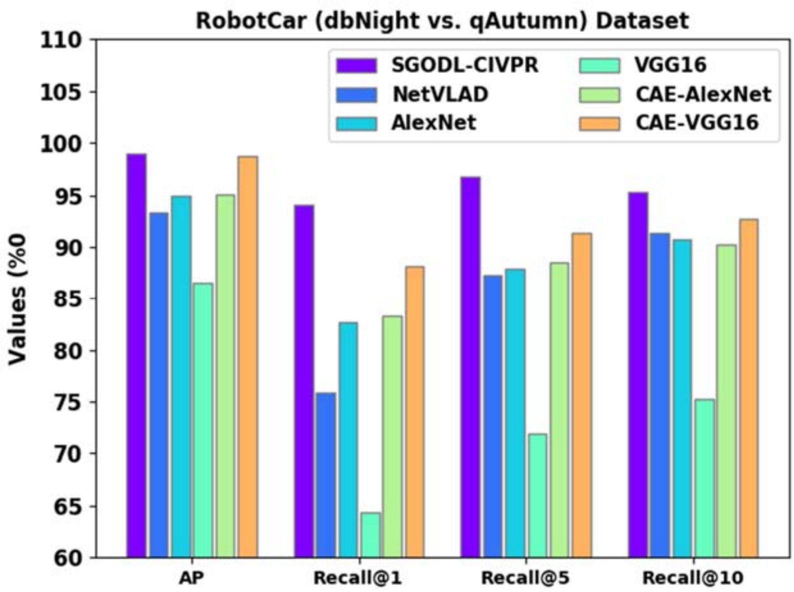

Fig. 8. Comparative analysis of SGODL-CIVPR approach under RobotCar (dbNight vs. qAutumn) dataset

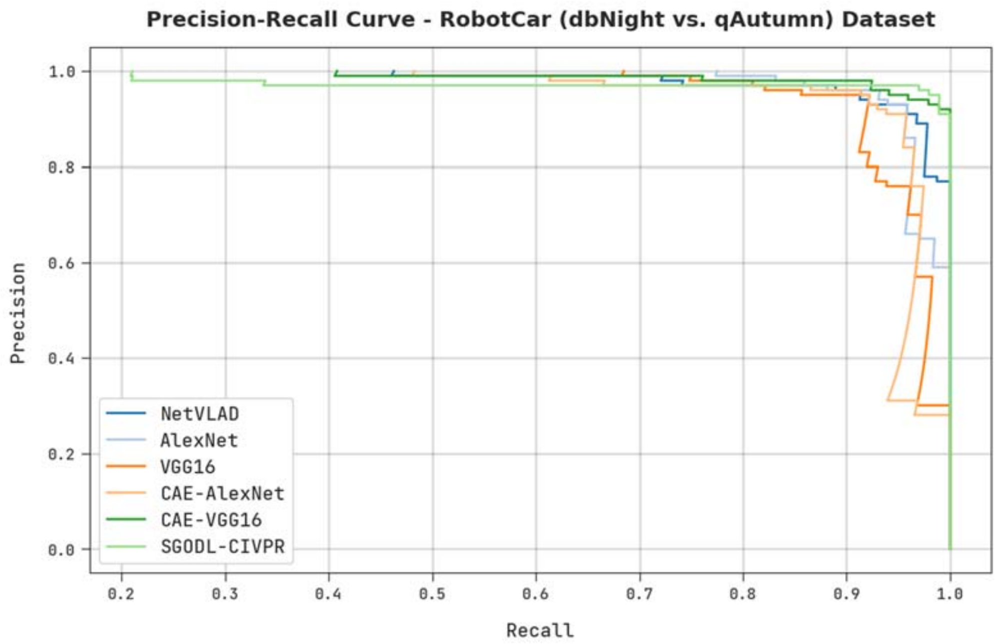

Fig. 9. Precision-recall analysis of SGODL-CIVPR approach under RobotCar (dbNight vs. qAutumn) dataset

An obvious precision-recall analysis of the SGODL-CIVPR algorithm on RobotCar (dbNight vs. qAutumn) dataset is demonstrated in Fig. 9. The figure stated that the SGODL-CIVPR system has resulted in higher values of precision-recall values under various classes.

Table 5 and Fig. 10 showcase the comparative outcomes of the SGODL-CIVPR methodology on RobotCar (dbNight vs qSnow) dataset. The outcomes exposed that the NetVLAD approach has obtainable lesser performance over other approaches whereas the VGG16 system has accomplished certainly improved outcomes.

RobotCar (dbNight vs. qSnow) Dataset				
Methods	**AP**	**Recall@1**	**Recall@5**	**Recall@10**
SGODL-CIVPR	98.60	92.59	94.86	95.24
NetVLAD	89.27	69.15	81.63	84.94
AlexNet	65.80	45.34	54.11	59.40
VGG16	80.94	52.38	58.30	63.44
CAE-AlexNet	94.99	72.97	79.70	82.46
CAE-VGG16	97.48	86.12	89.01	90.68

Table 5. Comparative analysis of SGODL-CIVPR approach with existing algorithms under RobotCar (dbNight vs. qSnow) dataset

Next, the AlexNet model has demonstrated moderately improved recognition efficiency whereas the CAE-AlexNet and CAE-VGG16 techniques have outperformed reasonable performance. At last, the SGODL-CIVPR model has depicted effectual outcome over other techniques with maximal AP, recall@1, recall@2, and recall@10 of 98.60%, 92.59%, 94.86%, and 95.24% respectively.

A clear precision-recall analysis of the SGODL-CIVPR technique on RobotCar (dbNight vs. qSnow) dataset is demonstrated in Fig. 11. The figure revealed that the SGODL-CIVPR methodology has resulted in higher values of precision-recall values under all classes. From the experimental outcomes, it is assured that the SGODL-CIVPR model has demonstrated maximum recognition performance over other models.

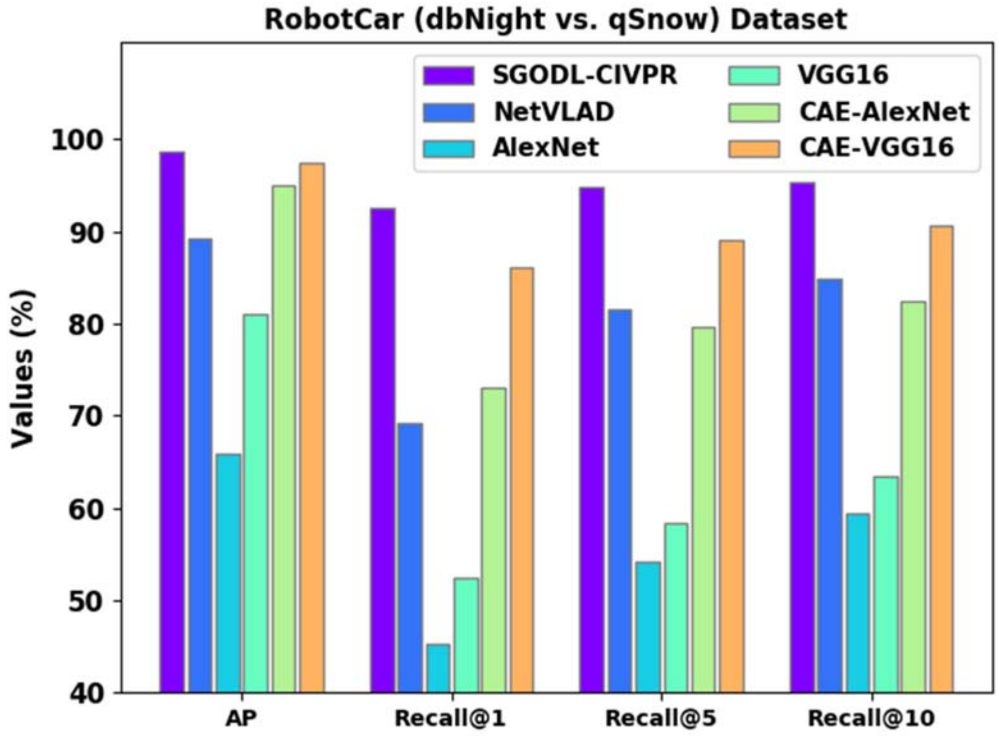

Fig. 10. Comparative analysis of SGODL-CIVPR approach under RobotCar (dbNight vs. qSnow) dataset

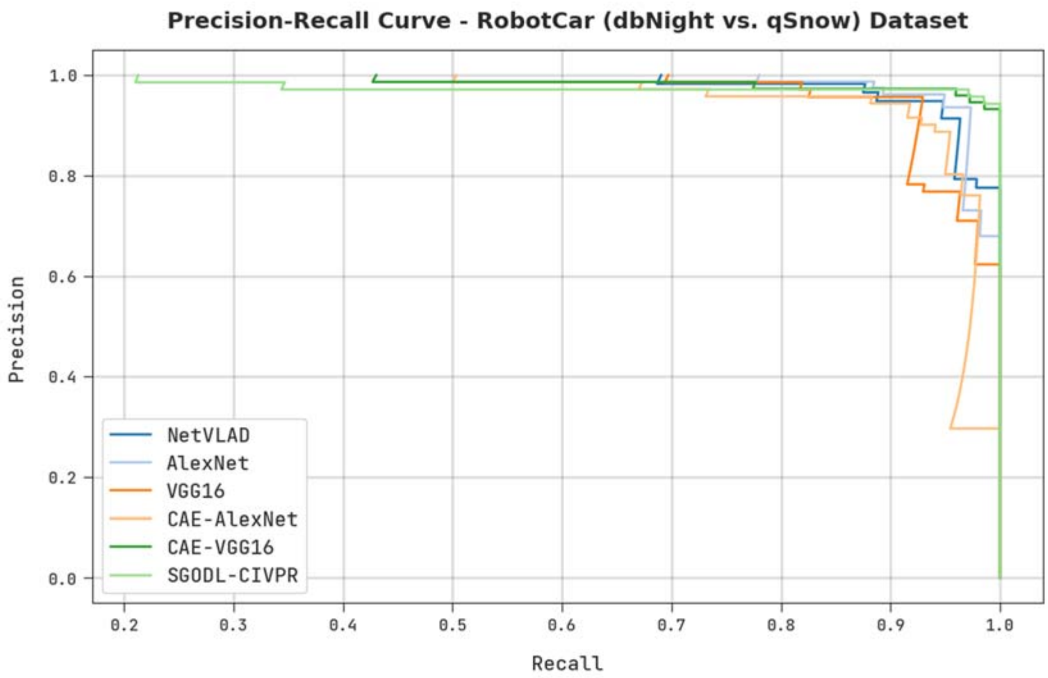

Fig. 11. Precision-recall analysis of SGODL-CIVPR approach under RobotCar (dbNight vs. qSnow) dataset

4. Conclusion

In this article, a novel SGODL-CIVPR algorithm was introduced for the recognition of live places with respect to reference images. At the initial stage, the SGODL-CIVPR technique utilizes deep convolutional neural network based EfficientNet model to produce features, and the hyperparameter tuning process takes place using the SGO algorithm. Moreover, the CWNN with DCGAN is used to map features to a low dimensional space for improving invariant properties and decreasing dimensions concurrently. The performance validation of the SGODL-CIVPR model is tested using different datasets. A wide ranging comparison study ensured the enhanced performance of the SGODL-CIVPR technique over other methods. Thus, the SGODL-CIVPR technique can be considered for effectual VPR in real time changing environments.

Conflicts of interest

The authors have no conflicts of interest to declare

References

[1] 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.
[2] Jamal, M.A., Brown, M., Yang, M.H., Wang, L. and Gong, B., 2020. Rethinking class-balanced methods for long-tailed visual recognition from a domain adaptation perspective. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7610-7619).
[3] Zhang, X., Wang, L. and Su, Y., 2021. Visual place recognition: A survey from deep learning perspective. *Pattern Recognition*, *113*, p.107760.
[4] Masone, C. and Caputo, B., 2021. A survey on deep visual place recognition. *IEEE Access*, *9*, pp.19516-19547.
[5] Zhao, Y., Chen, W., Tan, X., Huang, K. and Zhu, J., 2022, June. Adaptive logit adjustment loss for long-tailed visual recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 36, No. 3, pp. 3472-3480).
[6] Schubert, S., Neubert, P. and Protzel, P., 2021. Fast and Memory Efficient Graph Optimization via ICM for Visual Place Recognition. In *Robotics: Science and Systems*.
[7] Schubert, S., Neubert, P. and Protzel, P., 2021, May. Beyond ann: Exploiting structural knowledge for efficient place recognition. In *2021 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 5861-5867). IEEE.
[8] Ye, C., Zhang, F., Mu, L., Gao, Y. and Liu, Y., 2021. Urban function recognition by integrating social media and street-level imagery. *Environment and Planning B: Urban Analytics and City Science*, *48*(6), pp.1430-1444.
[9] Zhang, D., Wei, S., Li, S., Wu, H., Zhu, Q. and Zhou, G., 2021, May. Multi-modal graph fusion for named entity recognition with targeted visual guidance. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 35, No. 16, pp. 14347-14355).
[10] 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.
[11] Keetha, N.V., Milford, M. and Garg, S., 2021. A hierarchical dual model of environment-and place-specific utility for visual place recognition. IEEE Robotics and Automation Letters, 6(4), pp.6969-6976.
[12] Hong, Y., Han, S., Choi, K., Seo, S., Kim, B. and Chang, B., 2021. Disentangling label distribution for long-tailed visual recognition. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 6626-6636).
[13] Gupta, S., Sharma, K., Dinesh, D.A. and Thenkanidiyoor, V., 2021. Visual semantic-based representation learning using deep CNNs for scene recognition. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 17(2), pp.1-24.
[14] Garg, S. and Milford, M., 2021. SeqNet: Learning descriptors for sequence-based hierarchical place recognition. IEEE Robotics and Automation Letters, 6(3), pp.4305-4312.
[15] 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.
[16] Koonce, B., 2021. EfficientNet. In *Convolutional neural networks with swift for tensorflow* (pp. 109-123). Apress, Berkeley, CA.
[17] Karunkuzhali, D., Meenakshi, B. and Lingam, K., 2022. An Adaptive Fuzzy C Means with Seagull Optimization Algorithm for Analysis of WSNs in Agricultural Field with IoT. *Wireless Personal Communications*, pp.1-22.
[18] Zhang, X., Su, H., Zhang, C., Gu, X., Tan, X. and Atkinson, P.M., 2021. Robust unsupervised small area change detection from SAR imagery using deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *173*, pp.79-94.
[19] Olid, D., F´acil, J.M., Civera, J.: Single-view place recognition under seasonal changes. arXiv preprint arXiv:1808.06516 (2018)
[20] Liu, Y., Feng, R., Zhang, H.: Keypoint matching by outlier pruning with consensus constraint. In: 2015 IEEE International Conference on Robotics and Automation (ICRA), pp. 5481–5486 (2015). IEEE
[21] Maddern, W., Pascoe, G., Linegar, C., Newman, P.: 1 year, 1000 km: The oxford robotcar dataset. The International Journal of Robotics Research 36(1), 3–15 (2017).
[22] Ye, H., Chen, W., Yu, J., He, L., Guan, Y. and Zhang, H., 2022. Condition-Invariant and Compact Visual Place Description by Convolutional Autoencoder. *arXiv preprint arXiv:2204.07350*.

Authors Profile

P.Sasikumar received B.Sc degree in Computer Science as well as M.C.A degree from Annamalai University and M.Phil degree from Annamalai University, India. He joined the Department of Computer Science and Engineering, Annamalai University in the year 2004. At present he is with the Department of Computer and Information Science, Annamalai University. He has almost 18 years of teaching experience. He is currently pursuing Ph.D degree in the Department of Computer and Information Science, Annamalai University. His research interests include Image processing and Visual places image retrieval.

He working as a Professor in the Department of Computer Science & Engineering, Annamalai University, Chidambaram, Tamil Nadu, India. She has 28 years of Teaching Experience. She has published papers in 58 International Journals,19 National and 57 International Conferences and 1 Book has been published. Area of Interest: Image Processing and Computer Graphics.