International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I7P235

Volume 70 Issue 7, 339-347, July 2022

Original Article

Evolutionary Optimization with Deep Learning-Driven Visual Place Recognition for Seasonal Variant Environment

P. Sasikumar[1], S. Sathiamoorthy[2]

[1,2]*Department of Computer and Information Science, Annamalai University, Annamalai Nagar, India*

[1]mailtoausasikumar@gmail.com

Received: 06 May 2022 Revised: 08 July 2022 Accepted: 15 July 2022 Published: 26 July 2022

Abstract - *Visual Place Recognition (VPR) defines the process of identifying the same place despite considerable variations in appearances and viewpoints. VPR is a major element of Spatial Artificial Intelligence, allowing robotic and intelligent augmentation platforms to perceive and understand the real world. Long-term navigation in varying environments is a challenging problem in VPR owing to the distinct appearances of places with significant variations at different times of day, months, and seasons. Recently, authors had to work on advanced deep learning techniques to address this issue. This paper presents a novel Remora Optimization with Deep Learning-Driven Visual Place Recognition for Seasonal variant Environment, named ROADL-VPRSI model. The proposed ROADL-VPRSI model employs a pretrained capsule network (CapsNet) model to learn the image descriptors. Besides, ROA is applied to adjust the hyperparameters involved in the CapsNet model, such as learning rate, batch size, and the number of hidden layers. Next, the feature vectors are transformed into binary codes to minimize the computational complexity for image matching. Finally, the Minkowski distance-based similarity measurement process is carried out to recognize the places effectively. The experimental validation of the ROADL-VPRSI model is performed using a benchmark dataset, and the results are inspected under several measures. The comparative study highlighted the betterment of the ROADL-VPRSI model over recent methods.*

Keywords - *Visual places recognition, Computer vision, Similarity measurement, Remora optimization algorithm.*

1. Introduction

Visual place recognition (VPR) mainly focuses on helping a vision or a robot-related navigation system decide if it is located in a previously visited location. It is also considered the most difficult and significant issue in computer vision and robotics. Such domains have noticed an increase in the utilization of VPR for several application areas over the past decades. For instance, in a visual Simultaneous Localization and Mapping (SLAM) technique [1], place identification, also termed loop closure detection (LCD), is the main element. The localization mistakes caused by visual odometry (VO) could be minimized, eliminating an unclear map of the strange ambiance [2]. Whereas VPR has gained interest and was widely learned in robotics societies and computer vision, there exist several open problems to be handled. VPR is difficult, particularly in uncontrolled outdoor atmospheres over longer periods. Images captured by a robot at a particular place may vary from those captured on a first pass via the same location. It is because of environmental elements, namely variations in the changing weather conditions, the season of the year, day and night cycles, light variations during the day, and completely geometric factors, namely variations in viewpoint amongst 2 traverses. But, large-scale navigation in varying environments applies important challenges to VPR because a robot unavoidably experiences serious environmental changes [3]. These variations split into viewpoint changes and conditions.

Deep learning (DL) is considered a subdomain of machine learning (ML) that endeavors to study high-level abstractions in data using hierarchical structures. It was broadly implied in conventional artificial intelligence fields, like, transfer learning, computer vision, natural language processing [4], semantic parsing, and so on. There are primarily 3 significant causes for the flourishment of DL today: the drastically raised chip processing capabilities (e.g., GPU unit), the importantly lowered cost of computing hardware, and the developments in the ML methods [5]. The current trend in VPR study is admired by the greatest achievement of Convolution Neural Networks (CNN) in various computer vision works [6]. General methods reject its completely linked layers and utilize the output of their middle and recent convolutional layers for encrypting rich semantic information, which could be a strength to numerous image variations. CNNs mirror, in many ways, biological vision [7]. Moreover, executing at the same levels as human beings in various identification works, they act in a

hierarchical style, resembling few extend how the visual cortex processes information [8, 9]. The latest works towards CNNs and a few earlier methods of the visual system recommend that DNNs are superior prevailing methods for describing representations of spatial layout.

This paper presents a novel Remora Optimization with Deep Learning-Driven Visual Place Recognition for Seasonal variant Environment, named ROADL-VPRSI model. The proposed ROADL-VPRSI model employs a pretrained capsule network (CapsNet) model to learn the image descriptors. ROA is being exploited as a hyperparameter optimizer for the CapsNet model. Then, the feature vectors are transformed into binary codes. Finally, the Minkowski distance-based similarity measurement process is carried out to recognize the places effectively. The experimental validation of the ROADL-VPRSI model is performed using a benchmark dataset, and the outcomes are inspected under several measures.

2. Literature Review

The authors in [10] developed a VPR model for the varying environment using the voting model. Garg et al. [11] presented a triplet loss formulation at which point the distance metric depends on sequence matching measured by a single image. Then, similar metrics are utilized to mine data during the training process that assists the optimization algorithm by choosing harder negatives and proper positives. Khaliq et al. [12] introduced a light-weighted visual place recognition technique, able to achieve higher performance with lower computation cost and viable for mobile robotics in point appearance changes and significant view. Chen et al. [13] presented a multiscale context flexible system to evaluate the significance of spatial regions in the feature map. The algorithm is trained dedicatedly for recognizing places and identifying regions of interest. A widespread experiment was conducted to verify the efficiency, demonstrating that the presented method could perform effectively compared to the benchmark dataset.

Zhu et al. [14] projected an Attention-related Pyramid Aggregation Network (APANet) trained for recognizing places. A major element of APANet, the spatial pyramid pooling, could efficiently encode multiple-size buildings encompassing geo-information. Zaffar et al. [15] developed a cognition-inspired agnostic architecture to build a map for recognizing Visual Place. Peng et al. [16] designed a Semantic Reinforced Attention Learning Network (SRALNet), where the inferred interest benefitted from data-driven finetuning and semantic prior. The contribution relies two-fold upon. (1) With the interpretability of the local weighted system, a semantic constraint initialization is presented; thus, the local attention is strengthened using semantic priors. (2) To conquer misleading local features, an interpretable local weighted system is presented based on hierarchical feature distribution. Zhu et al. [17] presented a methodology based on CNN by placing images into a pretrained network structure for automatic learning of image descriptors.

3. The Proposed Model

In this study, a new ROADL-VPRSI model was enhanced to recognize the places invariant to seasonal variations. The ROADL-VPRSI model derived a useful set of features using the CapsNet model. Moreover, the ROA is applied to adjust the hyperparameters involved in the CapsNet model, such as learning rate, batch size, and the number of hidden layers. Afterward, the feature vectors are transformed into binary codes. Finally, the Minkowski distance-based similarity measurement process is carried out to recognize the places effectively. Fig. 1 illustrates the overall process of the ROADL-VPRSI technique.

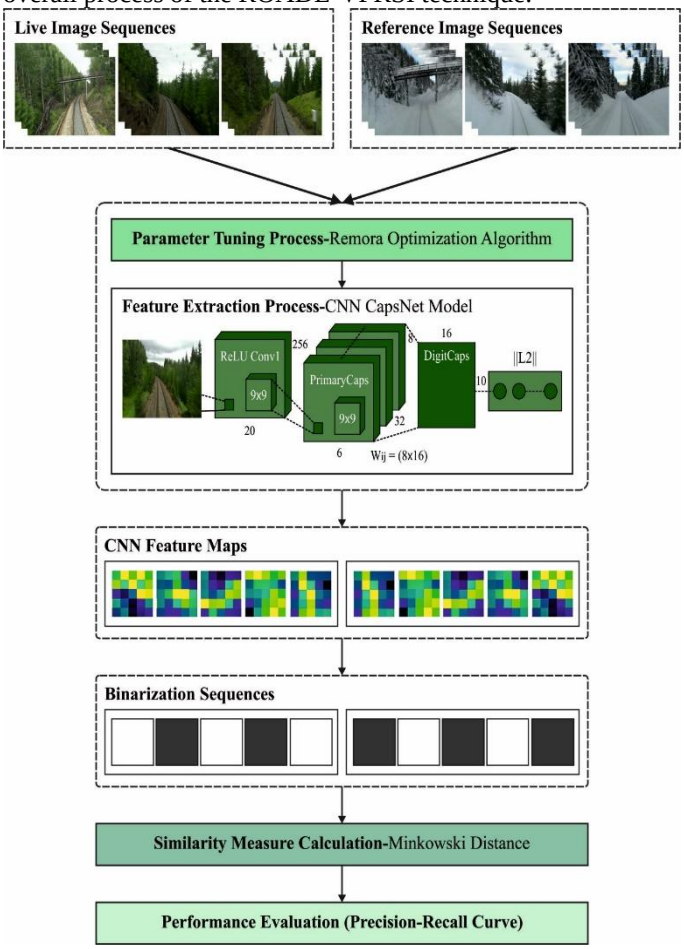

Fig. 1. Overall process of ROADL-VPRSI technique

3.1. CapsNet-based Feature Extraction

A capsule comprises a set of neurons where the output is inferred as distinct features of a similar entity and forms the activation vector. Every capsule comprises a pose matrix that characterizes the existence of a certain object situated at a pixel, and activation probability signifies the vector length. On rotating the image, e.g., the activation vector changes, but

its length remains unchanged [18]. In the presented framework, the authors applied a CancerCaps layer (that is, a capsule signifies two classes of histological images: cancerous/non-cancerous) and a primary capsule (squashed and reshaped output of the final convolution layer). Where u_i Indicates the i-th output capsule; the prediction for j-th parent capsule can be defined in the following equation.

$$\hat{u}_{\frac{j}{i}} = W_{ij} u_j \tag{1}$$

Now, $\hat{u}_{\frac{j}{i}}$ indicates the output prediction vector of the j-th capsules, W_{ij} and represents the weight matrixes that need to be learned during the backward pass. The Softmax operation is utilized for calculating the coupling coefficient c_{ij} Based on the amount of conformation amongst the capsules and also the parent capsule is called "iterative dynamic routing" as follows

$$c_{ij} = \frac{\exp\,(b_{ij})}{\sum_k \exp\,(b_{ik})} \tag{2}$$

In the equation, b_{ij} signifies the log-likelihood and fixed to zero initially, capsule i needs to be combined with j-th capsules using the agreement method. It is shown in the following equation

$$s_j = \sum_i c_{ij}\, \hat{u}_{\frac{j}{i}} \tag{3}$$

At last, the nonlinear squashing function is utilized to normalize the capsule's output vector by preventing them from exceeding 1. Thus, the length is characterized as the possibility that the capsule detects a provided feature. All the capsule's last output can be defined by its primary vector value as follows.

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|} \tag{4}$$

Now s_j indicates the overall input to j-th capsules and v_j represents the output. Based on the agreement between v_j and $\hat{u}$. Log probability needs to be upgraded in the routing process. Thus, the upgraded log-likelihood can be evaluated by using the following equation.

$$b_{ij} = b_{ij} + \hat{u}_{\frac{j}{i}} . \mathrm{v}_j \tag{5}$$

The routing coefficient is improved by using a dynamic routing process to j-th parent Capsule with factor $\hat{u}_{\frac{j}{i}} . v_j$. Subsequently, additional data is transmitted through child to parent capsules where the outcome v_j is similar to the prediction $\hat{u}_{\frac{j}{i}}$

An effective method to minimize the computation cost for image matching is to change the feature vector into binary codes that are compared using the Minkowski distance. The elements are normalized into an 8-bit integer (0-255) and acquire the integer feature F_{cn}^{inT}, as shown in (6). At that time, F_{cn}^{inT} can be transformed into a binary feature F_{cnn}^{bin}.

$$F_{cnn}^{inT} = \frac{F_{cn} - \min\,(F_{cn})}{\max\,(F_{cn}) - \min\,(F_{cn})} \times 255. \tag{6}$$

3.2. Hyperparameter Optimization

The ROA is applied for optimally tuning the CapsNet method's hyperparameters, such as learning rate, batch size, and the number of hidden layers. The ROA uses the biological features of remora for completing the optimization approach, that is, the parasitic behavior. Remora is capable of attaching themselves to whales, swordfish, or other animals. By using the host, remora could easily attain food. Accordingly, ROA adopted part of the location update mode of WOA and SFO for local and global searching. Notably, ROA employs an integer argument H (0 or 1) to define selecting the strategy of WOA or SFO. For a certain range, the ROA has the advantage of an optimization algorithm while resolving the optimization problem. The ROA employs the SFO approach for conducting the global searching that depends on the elite technique applied in the swordfish algorithm [19]. In the following, the location update equation is given:

$$V_i(T+1) = X_{besC}(T) - \left(rand \times \left(\frac{X_{best(T)} + X_{rand(T)}}{2} \right) - X_{rand}(T) \right) \tag{7}$$

Whereas $V_i(t+1)$ indicates the candidate location of ith remora. $X_{best}(t)$ represent the present optimal location. $X_{rand}(t)$ denotes an arbitrary location of remora. t implies the existing amount of iteration. Also, rand denotes an arbitrary value between 0 and 1. Furthermore, remora changes the host based on their experience. Here, a new candidate location is produced by using the following equation:

$$V_i'(T+1) = V_i(T+1) + randn \times \left(V_i(T+1) - X_i(T) \right) \tag{8}$$

Now $V_i'(t+1)$ indicates the candidate location of the i-th remora. $X_i(t)$ signifies the preceding location of i-th remora. Besides, $randn$ is utilized for producing arbitrary values. Also, remora attaches itself to the humpback whales for food. Therefore, remora have the motion features of humpback whales. The WOA approach is applied in ROA to implement the local searching. In particular, the bubble-net

attacking technique applied in WOA is used. The location update formula is given in the following equation:

$$V_i(T+1) = D \times e^a \times cos(2\pi a) + X_{best}(t) \quad (9)$$

$$D = |X_{besT}(t) - X_i(t)| \quad (10)$$

$$a = rand \times (b-1) + 1 \quad (11)$$

$$b = -\left(1 + \frac{t}{T}\right) \quad (12)$$

Here D signifies the distance between remora and food. As per the Eqs. (11) and (12), it is noted that a indicates an arbitrary value ranging from -2 and 1. Also, b decrease linearly from -l to -2. Furthermore, to enhance the solution quality, the remora produces a small step with the help of the encircling prey model in WOA as:

$$X_i(t+1) = V_i(t+1) + A \times D' \quad (13)$$

$$A = 2 \times B \times rand - B \quad (14)$$

$$B = 2 \times \left(1 - \frac{t}{T}\right) \quad (15)$$

$$D' = V_i(T+1) - C \times X_{besT}(T) \quad (16)$$

The equation $X_i(t+1)$ denotes the recently produced location of the i-th remora. C represents the remora factor, which is fixed to 0.1 in ROA. Fig. 2 depicts the flowchart of ROA.

Algorithm 1: The pseudocode of ROA

Begin
Initialize the remora population size (N) and maximal amount of iterations (T)
Initialize the position of each search agent $X_i(i = 1,2,3,\dots,N)$
Set the remora factor C
Main loop {
While $(t \le T)$
Compute the fitness of remora
find the optimal location and bestFitness, X_b
Compute the a, b, A, B
For the ith remora
If H(i) $= 0$
Create location V_i using Eq (3)
Else if H(i) $= 1$
Create location V_i using Eq (7)
End if
Create candidate location V_i' using Eq (8)
$iff(V_i') < f(V_i)$
$X_i = V_i^{\mathrm{l}}$
$H(i) = round(rand)$
else
Upgrade location X_i by Eq (13)
End if
End for
$t = t + 1$
End While}
Return the best fitness, X_b

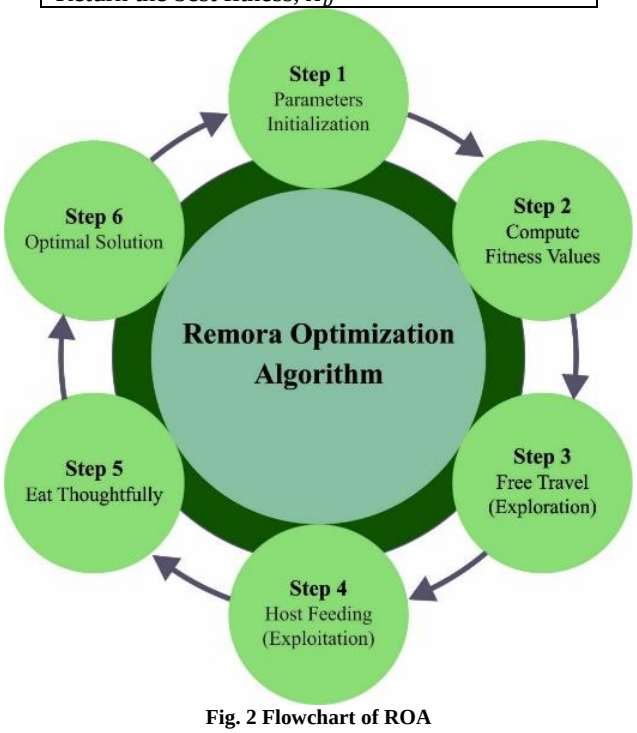

Fig. 2 Flowchart of ROA

3.3. Matching process

For effectively matching the visual places, the Minkowski distance is utilized to measure the similarity between two input images. The Minkowski metric or distance is a metric in a normed vector space that is treated as the generalization of a Manhattan and Euclidean distance. The Minkowski distance of order p (whereas p refers to the integer) amongst 2 points

$$X = (x_1, x_2, \dots, x_n) and\ Y = (y_1, y_2, \dots, y_n) \in R^n \quad (17)$$

$$D(X,Y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1\backslash p} \quad (18)$$

4. Results and Discussion

The experimental validation of the ROADL-VPRSI model is tested using the City Centre dataset [20] and Nordland dataset [21]. The former data is commonly employed for loop closure and place recognition. The second dataset includes samples under long-term conditions using a monocular camera. A few sample images are illustrated in Fig. 3.

A widely employed measure is the Precision-Recall (PR) curve which offers more details about the efficiency of the ROADL-VPRSI model.

Fig. 3 a) Reference Images b) Live Images

Table 1. Precision-recall analysis of ROADL-VPRSI technique on City Centre Dataset

	Precision (%)				
Recall Values	**20**	**40**	**60**	**80**	**100**
ROADL-VPRSI	99.56	98.05	91.11	79.13	36.98
Pool5 Fuse(ave7)	95.18	87.99	57.58	30.76	13.51
Pool5 Fuse(ave4)	92.54	85.84	68.6	51.35	17.10
Pool5 Fuse(ave2)	89.43	78.89	66.2	51.11	19.98
Pool5 Fuse(Max7)	83.92	77.22	70.51	55.90	15.67
Pool5 Fuse(Max4)	90.15	73.86	64.52	54.23	20.22
Pool5 Fuse(Max2)	85.60	87.28	73.62	52.31	21.18
Conv 5_3	94.70	88.23	75.3	58.54	19.02

Table 1 and Fig. 4 investigate precision-recall curves obtained by the ROADL-VPRSI model on the test city centre dataset. The experimental values indicated that the ROADL-VPRSI model had effectual outcomes over others. For instance, with a recall value of 20, the ROADL-VPRSI model has offered increased precision of 99.56%, whereas the Pool5 Fuse(Max7) has shown reduced precision of 83.92%. With a recall value of 60, the ROADL-VPRSI technique has an obtainable higher precision of 91.11%, whereas the Pool5 Fuse(Max7) has outperformed lower precision of 70.51%. Along with that, with a recall value of 80, the ROADL-VPRSI approach has accessible enhanced precision of 79.13%, whereas the Pool5 Fuse(Max7) has exhibited reduced precision of 55.90%. Moreover, with a recall value of 100, the ROADL-VPRSI algorithm has offered increased precision of 36.98%, whereas the Pool5 Fuse(Max7) has portrayed minimal precision of 15.67%.

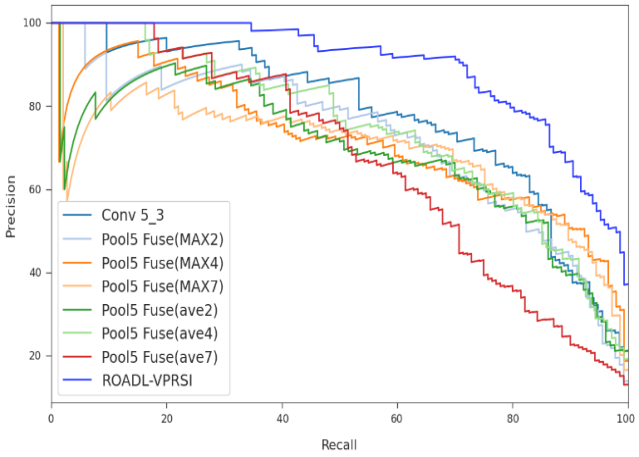

Fig. 4 Precision-recall analysis of ROADL-VPRSI technique on City Centre Dataset

Tables 2-3 report a detailed comparative outcome of the ROADL-VPRSI model with existing models under different season variations.

Fig. 5a illustrates the precision-recall outcomes of the ROADL-VPRSI model and existing approaches under the spring vs. summer season. The figure indicated that the ROADL-VPRSI model has increased precision values over other approaches. For instance, with a recall value of 20, the ROADL-VPRSI model has provided higher precision of 93.43%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP models have reduced precision of 80.31%, 82.06%, 88.83%, and 79.43% correspondingly. Also, with a recall value of 100, the ROADL-VPRSI approach has provided higher precision of 28.93%, whereas the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP techniques have reached lower precision of 24.34%, 28.27%, 24.12%, and 25.43% correspondingly.

Fig. 5b showcases the precision-recall outcome of the ROADL-VPRSI technique and existing algorithms under the spring vs. fall season. The figure revealed that the ROADL-VPRSI model had enhanced precision values over other approaches. For instance, with a recall value of 20, the ROADL-VPRSI model has provided superior precision of 97.11%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP methods have reduced precision of 95.08%, 93.49%, 100%, and 91.69%, respectively. Likewise, with a recall value of 100, the ROADL-VPRSI approach has provided higher precision of 35.85%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP models have lower precision of 23.42%, 22.06%, 22.51%, and 22.29% correspondingly.

Table 2. Precision-recall analysis of ROADL-VPRSI technique on Nordland Dataset

Precision (%) (Spring vs. Summer)					
Recall	**ROADL-VPRSI**	**VPR-LTSE**	**Seq-SLAM Model**	**ABLE Model**	**FAB-MAP Model**
0	100.00	100.00	100.00	100.00	100.00
20	93.43	80.31	82.06	88.83	79.43
40	92.33	83.15	82.93	76.59	61.07
60	86.21	76.15	64.57	56.91	54.51
80	70.25	63.04	52.76	42.92	45.11
100	28.93	24.34	28.27	24.12	25.43
Precision (%) (Spring vs. Fall)					
Recall	**ROADL-VPRSI**	**VPR-LTSE**	**Seq-SLAM Model**	**ABLE Model**	**FAB-MAP Model**
0	100.00	100.00	100.00	100.00	100.00
20	97.11	95.08	93.49	100.00	91.69
40	90.56	86.71	91.01	88.30	82.87
60	87.39	84.68	73.15	81.29	68.63
80	79.25	73.15	61.62	59.36	43.54
100	35.85	23.42	22.06	22.51	22.29
Precision (%) (Summer vs. Fall)					
Recall	**ROADL-VPRSI**	**VPR-LTSE**	**Seq-SLAM Model**	**ABLE Model**	**FAB-MAP Model**
0	100.00	100.00	100.00	100.00	100.00
20	91.16	81.50	90.06	99.06	85.24
40	85.24	87.43	92.04	84.80	88.97
60	88.53	89.85	80.41	39.14	84.14
80	78.87	67.46	51.43	30.14	55.16
100	49.89	42.65	33.65	23.99	24.21

Table 3. Comparative analysis of ROADL-VPRSI technique on Nordland Dataset

Precision (%) (Spring vs. Winter)					
Recall	**ROADL-VPRSI**	**VPR-LTSE**	**Seq-SLAM Model**	**ABLE Model**	**FAB-MAP Model**
0	100.00	100.00	100.00	100.00	100.00
20	85.60	80.58	73.27	91.09	90.63
40	88.34	81.03	68.93	78.06	70.75
60	82.18	76.92	63.67	61.84	53.39
80	73.72	59.56	49.05	44.25	33.29
100	32.37	30.77	25.29	24.38	24.38
Precision (%) (Summer vs. Winter)					
Recall	**ROADL-VPRSI**	**VPR-LTSE**	**Seq-SLAM Model**	**ABLE Model**	**FAB-MAP Model**
0	100.00	100.00	100.00	100.00	100.00
20	85.60	83.24	99.52	89.85	82.77
40	85.60	91.50	96.22	71.44	73.33
60	85.84	85.84	87.49	54.69	65.78
80	76.16	69.55	66.01	34.39	43.83
100	45.48	43.12	44.30	17.87	23.77
Precision (%) (Fall vs. Winter)					
Recall	**ROADL-VPRSI**	**VPR-LTSE**	**Seq-SLAM Model**	**ABLE Model**	**FAB-MAP Model**
0	100.00	100.00	100.00	100.00	100.00
20	87.74	92.98	69.26	79.75	58.54
40	89.10	86.60	62.42	70.63	49.65
60	81.58	75.88	52.84	62.42	41.67
80	65.84	53.30	38.02	46.91	31.18
100	25.02	23.42	24.11	22.74	22.51

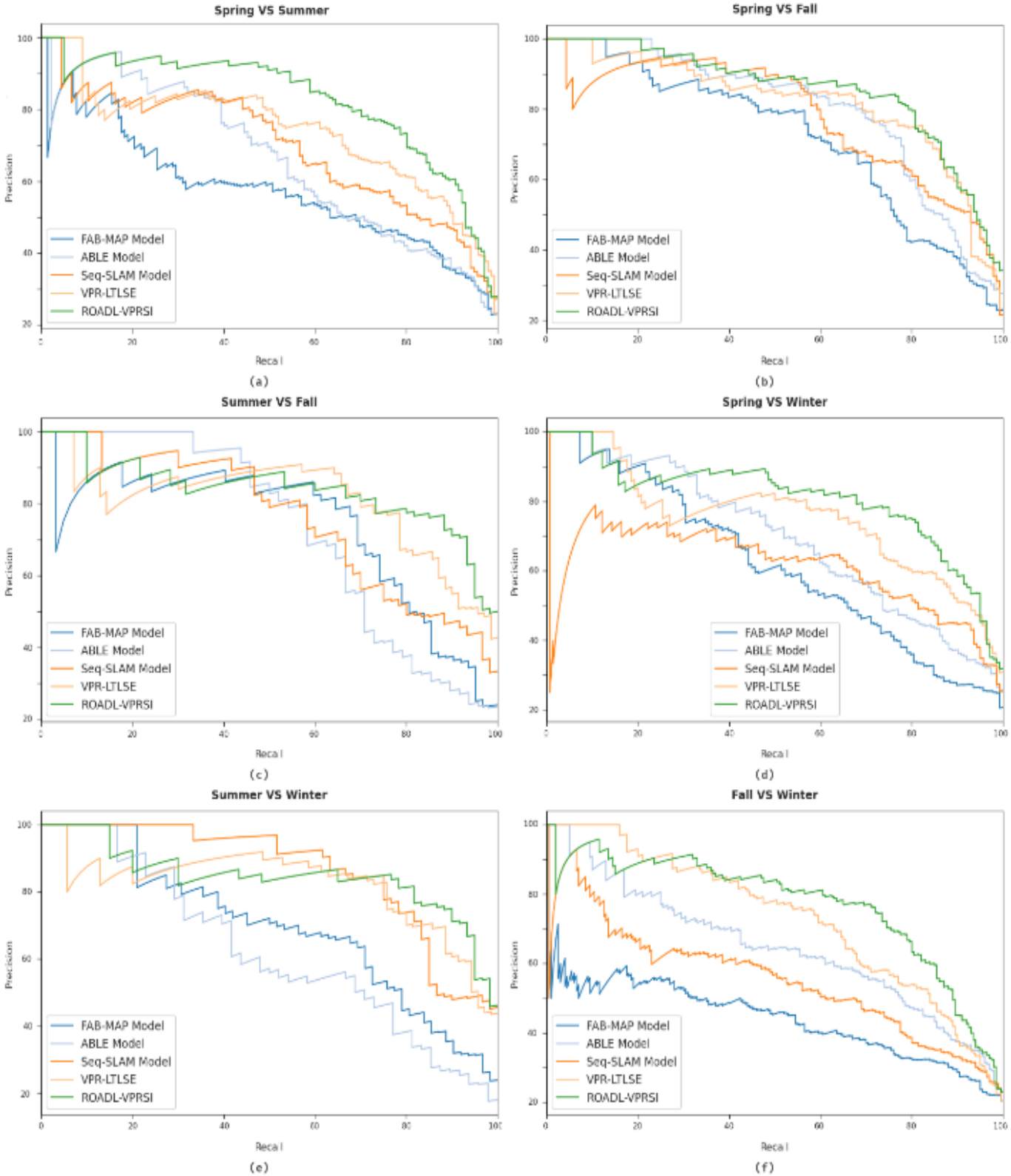

Fig. 5 Comparative analysis of ROADL-VPRSI technique on Nordland Dataset (a) spring vs. summer, (b) spring vs. fall, (c) summer vs. fall, (d) spring vs. winter, (e) summer vs. winter, and (f) fall vs. winter

Fig. 5c demonstrates the precision-recall outcomes of the ROADL-VPRSI approach and existing methods under the summer vs. fall season. The figure revealed that the ROADL-VPRSI model has increased precision values over other approaches. For instance, with a recall value of 20, the ROADL-VPRSI model has provided higher precision of 91.16%, whereas the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP models have obtained reduced precision of 81.50%, 90.06%, 99.06%, and 85.24% respectively. With a recall value of 100, the ROADL-VPRSI technique has accessible maximal precision of 49.89%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP models have achieved lower precision of 42.65%, 33.65%, 23.99%, and 24.21% correspondingly.

Fig. 5d depicts the precision-recall outcomes of the ROADL-VPRSI model and existing approaches under the spring vs. winter season. The figure indicated that the ROADL-VPRSI algorithm had enhanced precision values over other methods. For instance, with a recall value of 20, the ROADL-VPRSI approach has provided higher precision of 85.60%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP models have reduced precision of 80.58%, 73.27%, 91.09%, and 90.63% correspondingly. In addition, with a recall value of 100, the ROADL-VPRSI methodology has provided higher precision of 32.37%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP models have reduced precision by 30.77%, 25.29%, 24.38%, and 24.38%, respectively.

Fig. 5e defines the precision-recall outcome of the ROADL-VPRSI approach and existing methods under the summer vs. winter season. The figure indicated that the ROADL-VPRSI model has increased precision values over other approaches. For instance, with a recall value of 20, the ROADL-VPRSI system has increased precision by 85.60%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP techniques have reduced precision of 83.24%, 99.52%, 89.85%, and 82.77% correspondingly. In addition, with a recall value of 100, the ROADL-VPRSI model has provided higher precision of 45.48%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP models have reduced the precision by 43.12%, 44.30 %, 17.87%, and 23.77%, respectively correspondingly.

Fig. 5f illustrates the precision-recall outcomes of the ROADL-VPRSI model and existing approaches under the fall vs. winter season. The figure revealed that the ROADL-VPRSI algorithm has superior precision values over other approaches. For the sample, with a recall value of 20, the ROADL-VPRSI technique has provided higher precision of 87.74%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP approaches have obtained minimal precision of 92.98%, 69.26%, 79.75%, and 58.54% correspondingly. Eventually, with a recall value of 100, the ROADL-VPRSI system provided higher precision of 25.02%. In contrast, the VPR-LTSE, Seq-SLAM, ABLE, and FAB-MAP techniques reached lower precision of 23.42%, 24.11%, 22.74%, and 22.51% correspondingly.

Then observing the tables mentioned earlier and figures, it can be apparent that the ROADL-VPRSI technique is gained higher performance over other approaches.

5. Conclusion

In this study, a new ROADL-VPRSI model was enhanced to recognize the places invariant to seasonal variations. The ROADL-VPRSI model derived a useful set of features using the CapsNet model. Moreover, the ROA is applied to adjust the hyperparameters involved in the CapsNet model appropriately, namely learning rate, batch size, and the number of hidden layers. Afterward, the feature vectors are transformed into binary codes. Finally, the Minkowski distance-based similarity measurement process is carried out to recognize the places effectively. The experimental validation of the ROADL-VPRSI model is performed using a benchmark dataset, and the outcomes are inspected over several measures. The comparative study highlighted the betterment of the ROADL-VPRSI model over recent methods.

References

[1] Garg, S., Fischer, T. and Milford, M, " Where is Your Place, Visual Place Recognition?," 2021. *Arxiv Preprint Arxiv:2103.06443*.

[2] Garg, S., Suenderhauf, N. and Milford, M., 2019. Semantic–Geometric Visual Place Recognition: A New Perspective for Reconciling Opposing Views. *the International Journal of Robotics Research*, P.0278364919839761, 2019.

[3] Zaffar, M., Garg, S., Milford, M., Kooij, J., Flynn, D., Mcdonald-Maier, K. and Ehsan, S, "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.

[4] Sumitra, N., Lap, P., and Kunalai, P, "Improving the Prediction of Rotten Fruit Using Convolutional Neural Network," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 69, 2021.

[5] Amit, V., Upadhyay, V.G., Tripathi, M.M, "Development of Artificial Intelligent Techniques for Shortterm Wind Speed Forecasting," *International Journal of Engineering Trends and Technology (IJETT)*, vol.69, 2021.

[6] Rajendra.P, Pusuluri V.N.H, Gunavardhana Naidu.T, "the Performance of Various Optimizers In Machine Learning," *International Journal of Engineering Trends and Technology (IJETT)*, vol.69, no.7,2021.

[7] Garg, S. and Milford, M, "Fast, Compact and Highly Scalable Visual Place Recognition Through Sequence-Based Matching of Overloaded Representations," In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3341-3348, 2020.

[8] Schlegel, D. and Grisetti, G, "HBST: A Hamming Distance Embedding Binary Search Tree for Feature-Based Visual Place Recognition," *IEEE Robotics and Automation Letters*, vol.*3*, no.4, pp.3741-3748, 2018.

[9] Zhang, X., Wang, L. and Su, Y, "Visual Place Recognition: A Survey From Deep Learning Perspective," *Pattern Recognition*, vol.*113*, pp.107760, 2021.

[10] Masone, C. and Caputo, B, " A Survey on Deep Visual Place Recognition," *IEEE Access*, vol.*9*, pp.19516-19547, 2021.

[11] Garg, S., Vankadari, M. and Milford, M, "Seqmatchnet: Contrastive Learning with Sequence Matching for Place Recognition & Relocalization," *In Conference on Robot Learning , PMLR*, pp. 429-443, 2022.

[12] Khaliq, A., Ehsan, S., Chen, Z., Milford, M. and Mcdonald-Maier, K, "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, 2019.

[13] Chen, Z., Liu, L., Sa, I., Ge, Z. and Chli, M, "Learning Context Flexible Attention Model for Long-Term Visual Place Recognition," *IEEE Robotics and Automation Letters*, vol.3, no.4, pp.4015-4022, 2018.

[14] Zhu, Y., Wang, J., Xie, L. and Zheng, L, "October. Attention-Based Pyramid Aggregation Network for Visual Place Recognition," *In Proceedings of the 26th ACM International Conference on Multimedia*, pp. 99-107, 2018.

[15] Zaffar, M., Ehsan, S., Milford, M. and Mcdonald-Maier, K.D, "Memorable Maps: A Framework for Re-Defining Places In Visual Place Recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol.22, no.12, pp.7355-7369,2020.

[16] Peng, G., Yue, Y., Zhang, J., Wu, Z., Tang, X. and Wang, D, "May. Semantic Reinforced Attention Learning for Visual Place Recognition," *In 2021 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, pp.13415-13422, 2021.

[17] Zhu, J., Ai, Y., Tian, B., Cao, D. and Scherer, S, "Visual Place Recognition In Long-Term and Large-Scale Environment Based on CNN Feature," In *2018 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, pp.1679-1685, 2018.

[18] Panigrahi, S., Das, J. and Swarnkar, T, "Capsule Network-Based Analysis of Histopathological Images of Oral Squamous Cell Carcinoma," *Journal of King Saud University-Computer and Information Sciences*, 2020.

[19] Zheng, R., Jia, H., Abualigah, L., Wang, S. and Wu, D, "An Improved Remora Optimization Algorithm with An Autonomous Foraging Mechanism for Global Optimization Problems," *Mathematical Biosciences and Engineering*, vol.*19*, no.4, pp.3994-4037, 2022.

[20] M. Cummins and P. Newman, "FAB-MAP: Probabilistic Localization and Mapping In the Space of Appearance," *International Journal of Robotics Research (IJRR)*, vol. 27, no.6,pp. 647-665, 2008

[21] N Sünderhauf, P. Neubert, and P. Protzel, "Are We There Yet? Challenging Seqslam on A 3000 Km Journey Across All Four Seasons," *In Workshop on Long-Term Autonomy At the IEEE International Conference on Robotics and Automation (W-ICRA)*, Karlsruhe, 2013.