International Journal of Data Science and Artificial Intelligence (IJDSAI)
Volume 02, Issue 05, September – October (2024)

RESEARCH ARTICLE

DEEP-FIR: DEEP LEARNING BASED BUTTERFLY OPTIMIZED REGRESSION NETWORK FOR FAST IMAGE RETRIEVAL

S. Caroline [1, *], Vemulapalli Sri Lakshmi Harshitha [2]

[1] Assistant professor, Department of Electronics and Communication Engineering, St. Xavier's Catholic College of Engineering, Nagercoil, Tamil Nadu, India
[2] Under Graduate Student, BE Computer Science and Design, RMK Engineering College (an autonomous college), Kavaraipettai, Tamil Nadu, India
[*]*Corresponding e-mail: caroline@sxcce.edu.in*

Abstract – Image retrieval is a fundamental task in computer vision, aims to retrieve relevant images from large-scale databases based on user queries. However, the existing image retrieval systems are their susceptibility to inaccuracies when dealing with semantic gaps among low-level and high-level features leading to mismatches in retrieved results. This work proposes a novel Deep learning (DL) based DEEP-FIR approach to enhance image retrieval efficiency by integrating bio-inspired optimization and deep learning network. Initially, the input query images are pre-processed with weighted median filter to remove the noisy distortions. The proposed method employs an advanced DL-based Regression network that extracts the low-level features like colour and texture, with high-level semantic features. By fusing extracted features, the retrieval system is the leading to more accurate and discriminative representations of images. Additionally, butterfly mating optimization (BMO) algorithm is utilized for boosting performance by calculating comparison between query and database images to specific retrieval tasks. Experimental results on benchmark datasets establish the efficiency of the proposed DEEP-FIR approach with an overall accuracy of 97.8%. The proposed DEEP-FIR approach increases the overall accuracy of 3.16%, 2.35%, and 4.70% for DL-CNN, Multi-view and CBIR-CNN respectively.

Keywords *– Image retrieval, Deep learning, Weighted median filter, Butterfly mating optimization, RegNet.*

1. INTRODUCTION

Image retrieval is a field within computer vision and information retrieval focused on the task of searching and regaining images from large databases based on their content, rather than just their metadata [1]. It involves developing algorithms and techniques to effectively find images that are visually similar or semantically related to a given query image [2]. This is particularly valuable in various applications like image organization, multimedia databases, and visual search engines [3]. The goal is to enable users to efficiently locate relevant images by automatically analyzing their visual content, including colors, textures, shapes, and sometimes even the context in which they appear [4]. Techniques employed in image retrieval often leverage features extracted from images using deep learning, traditional computer vision methods, or a combination of both to compare and rank images based on their comparison to the query [5].

In order to capture the complex and high-dimensional representations that are inherent in images, traditional approaches frequently relied significantly on handmade features and shallow learning algorithms [6]. Deep learning (DL) models, on the other hand, have completely changed this environment by automatically deriving hierarchical features from raw pixel data, allowing for a more complex comprehension and portrayal of images [7]. The feature extraction network, often implemented to transform input images into high-dimensional feature vectors that encode rich semantic information [8]. These feature vectors encapsulate various visual attributes such as shapes, textures, and spatial relationships, enabling the network to capture both low-level and high-level image characteristics [9]. Therefore, the main aim to provide insight into the progression of image retrieval by examining deep learning-based methods catalysing innovation within this dynamic field [10]. The key contributions of this work are pointed as:

- This study employs novel deep learning based NLD-net for retrieving top relevant breast cancer histopathology images.
- DL-based RegNet is used to extracts the low-level features and high-level features to more accurate and discriminative representations of images.
- BMO algorithm is utilized for boosting performance by calculating similarity between query and database images to specific retrieval tasks.
- The efficacy of the suggested model was determined using the network parameters viz.,

F1 score, precision, recall, accuracy, and specificity.

The rest of this work was pre-arranged into following sections. Section-2 summarizes the recent works of image retrieval sector, Section-3 introduces our suggested image retrieval system in technical aspect, Section-4 holds the experimental setup and the findings of the proposed model and Section-5 enfolds with conclusion and future work.

2. LITERATURE REVIEW

Many researchers have published studies for retrieval methods in the recent days using the advances in ML and DL approaches that can be found in a variety of literature works.

In 2023, Naeem et al. [11] introduced DL methodology for image retrieval. This technique combines cutting-edge methods with a CNN that can do gradient computation, auto-correlation, filtering, scaling and localization. In the first step, photos from adjacent rectangular key spots are automatically correlated. Subsequently, intensity determination based on local gradients facilitates image smoothing. Furthermore, feature vectors are dimensionally reduced using PCA to represent color channels, which are then combined with VGG features. In order to improve retrieval performance, these attributes are finally combined with spatial color coordinates.

In 2022 Keisham and Neelima [12] introduced CBIR and deep search and rescue (SAR), algorithm for efficient image retrieval. This algorithm has DNN-based SAR framework, starts with noise removal using the FAPG filter, followed by extracting color, shape, and texture features to compute feature vectors. These were grouped using the sunflower optimization technique after being integrated into a single feature using average and weighted average approaches. The DNN-SAR optimization ensures accurate retrieval of relevant photos based on comprehensive feature representations.

In 2021, Singh., et al. [13] developed a deep DL model tailored for CBIR using facial image data. Their approach incorporated max-pooling for feature reduction and convolution layer activation for feature representation. The significance of employing this method for CBIR was examined, focusing on two pivotal components: categorization and image retrieval via similarity-based approaches. For classification, a quadruple convolution layer model was proposed, while the Euclidean distance measure was utilized to compute similarity between images.

In 2021, Karthi and Kamath [14] introduced a DNN-based approach for classifying perspectives and retrieving images based on their content, particularly effective for medical image retrieval. Additionally, a method was devised to mitigate variance across different scan types by classifying body part orientation views. Initially, acquired features were utilized for class label prediction and subsequently for modelling the feature space to compute retrieval similarity. This approach was evaluated in terms of error score, demonstrating superior practical utility compared to the prior works.

In 2020 Yan., et al., [15] introduced the multi-view deep neural network based on hash learning, devising an efficient retrieval model that notably enhances retrieval efficacy. Our research presents a supervised multi-view hash model employing neural networks to enhance multi-view data representation. This proposed method actively explores inter-view relationships, thereby influencing the optimization trajectory of the entire network through the utilization of an efficient view stability assessment technique. Moreover, to harness the advantages of both convolution and multi-view approaches, several multi-data fusion techniques are devised within the Hamming space.

From this survey, the existing works were struggle to scale efficiently when applied to large-scale image databases, potentially leading to performance degradation or increased computational complexity. Despite integrating DL techniques, some methods still heavily rely on handcrafted features, which may not capture the full complexity and variability of real-world image data. Since, a novel DEEP-FIR approach is proposed to enhance image retrieval efficiency by integrating bio-inspired optimization and deep learning network.

3. PROPOSED METHODOLOGY

In this section, a novel DEEP-FIR approach is presented to retrieve the top most images from the available dataset based on the user query. The proposed DEEP-FIR system is displayed in Figure.1.

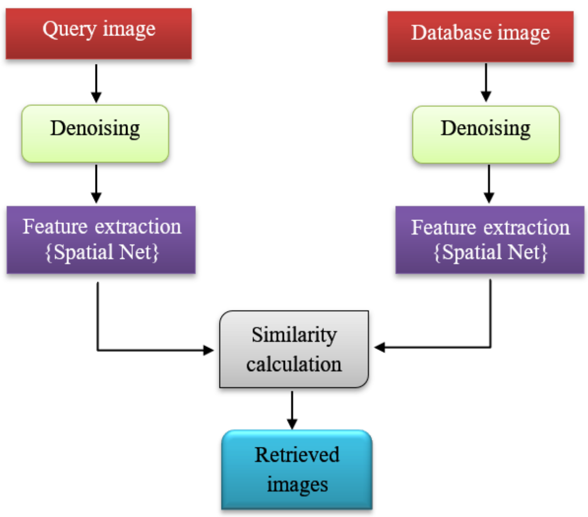

Figure 1. The proposed DEEP-FIR methodology

3.1. Dataset description

This research utilizes two primary datasets: corel1k and corel5k, which are extensively employed in the realm of image retrieval. The corel1k dataset comprises 1000 photos categorized into 10 distinct classes, including dinosaurs and buses, with each category containing 100 images. On the other hand, corel5k boasts fifty diverse image categories such as race cars, beaches, pets, and airplanes, each comprising 100 images, amounting to a total of 5000 images akin to the corel1k dataset. For the experiment, we will designate 10 randomly chosen images from each category in both corel1k and corel5k as query images, while the remaining images from corel5k will serve as the training set.

3.2. Data pre-processing

An adaptive hybrid genetic algorithm-based pre-processing method for query images is the weighted median filter. With fewer iterations needed to produce a more accurate denoised query image, this approach makes it easier to modify denoising parameters to accommodate changing lighting conditions and directions. Genetic cross and mutation probabilities are dynamically adjusted with respect to the stable-state regional population density in order to achieve higher precision. The method used to determine the weighted median filter was as follows:

$$f(i,j)_{wm} = median\{ w\ (i,j\ ) \times f\ (\ i,j\ )\ ,f\ (\ r,s\ )|f\ (\ r,s\ )\ \in N^{o}_{p(i,j)}\} \quad (1)$$

where $w(i,j) \times f(i,j)$ represents gray value when the weight of $f(i,j)$ is $w(i,j)$, $f(i,j)_{wm}$ is the center weighted median gray value, $f(i,j)$ represents the gray value of pixel $p(i,j)$, $N^{o}_{p(i,j)}$ represents the hollow neighborhood of pixel $p(i,j)$, and $f(r,s)$ represents the gray value corresponding to all pixels in $N^{o}_{p(i,j)}$.

3.3. Feature extraction network

The proposed method employs DL-based RegNet that extracts the low-level features like colour and texture, with high-level semantic features. By fusing extracted features, the retrieval system is the leading to more accurate and discriminative representations of images. RegNet is data-driven model with the unknown input-to-output mappings adaptively without supposing bias or past knowledge. Instead of using fixed parameters like depth and width, a quantized linear function controlled by the chosen parameters was used to establish the architecture of the RegNet model. In this experiment, stride (s=1) maximum pooling was used. As a result, there were 128, 64, and 32 units. The maximum input size allowed by the framework is 224*224, which is the typical RegNet system data size. The suggested structure makes use of flatten level, dense layers, and nine conv2d layers with filter widths of 128*64*32. For 224*224 drawings, there are a total of 194*903*073 training variables. Following extraction, the block widths were calculated using the following formula.

$$V_i = \omega_0 + \omega_x . i\ \ for\ 0 \leq i \leq l \quad (2)$$

The breadth of each block increased by a factor of x for each succeeding block. The authors then calculated *ri* after including a new parameter, 0, which the user had set:

$$\omega_i = \omega_0 . \omega_c{}^{ri} \quad (3)$$

The authors finally rounded ri and computed the quantized per-block widths in order to quantize V_i. Each phase's breadth was efficiently computed by summing all of the identically sized units that go into creating it. In this case, l, ω_0, ω_x, and ω_c is stand for depth, initial width, initial slope, and width parameter, respectively, which define a RegNet by defining the RegNet design area.

3.4. Similarity calculation

The Butterfly Mating Optimization (BMO) algorithm is employed for calculating similarity in image retrieval tasks. BMO utilizes an architecture that balances exploration and exploitation to efficiently explore the solution space, drawing inspiration from the mating behavior of butterflies. In the context of image retrieval, BMO operates by iteratively adjusting the parameters of the similarity measure based on the mating process of butterflies, allowing for the identification of relevant images that closely match the query.

The BOA algorithm mirrors the foraging and mating behaviours of butterflies. Butterflies possess chemical receptors across their bodies, acting as sensory nodes. These receptors enable butterflies to perceive the aroma of food and flowers, as well as to select optimal mates. As butterflies migrate, they emit varying scent intensities. Leveraging this scent, the BOA algorithm orchestrates the movement of its agents, akin to butterflies, through the search space.

When a butterfly fails to detect other butterflies' scents within its vicinity, it resorts to exploitation, locally searching by relocating to a randomly chosen spot, influenced by scent strength. Conversely, when a butterfly detects the scent of a promising mate, it engages in exploration, globally seeking the best match. Equation (4) encapsulates this scent dynamics, depicting it as a function of incentive intensity within the BOA algorithm.

$$S = cI^{z} \quad (4)$$

where c is the sense mode with values across the range [0, 1], z is the power exponent that relies on the sensory modality with values over the range [0, 1], and I is the stimulus intensity of the perfume that the butterfly emits. Controlling the amount of aroma absorption by stimulus intensity is another duty. Two primary equations are used to update the positions of butterflies in BOA based on *S* value. Equation (5) represents the global search in the first equation, and equation (6) represents the local search in the following way:

$$B_i^{t+1} = B_i^t + (r^2 \times G' - B_i^t) \times S_i \quad (5)$$

$$B_i^{t+1} = B_i^t + (r^2 \times B_j^t - B_k^t) \times S_i \quad (6)$$

where *G'* value shows current iteration, S_i is the smell of the i th butterfly, *r* represents a random value, B_j^t represents the jth butterfly, and B_k^t represents the kth butterfly of the available solutions space. By leveraging the BMO algorithm, image retrieval systems were effectively computing similarity scores between query images and database images, facilitating accurate and efficient retrieval of visually similar images.

4. RESULTS AND DISCUSSION

This section uses Matlab-2020b to implemented the experimental fallouts and evaluate the efficacy of the suggested DEEP-FIR approach. Moreover, the comparison offers a detailed description and analysis of the total accuracy rate besides the effectiveness of the suggested DEEP-FIR approach. Additionally, the proposed DEEP-FIR method is contrasted with traditional deep learning models.

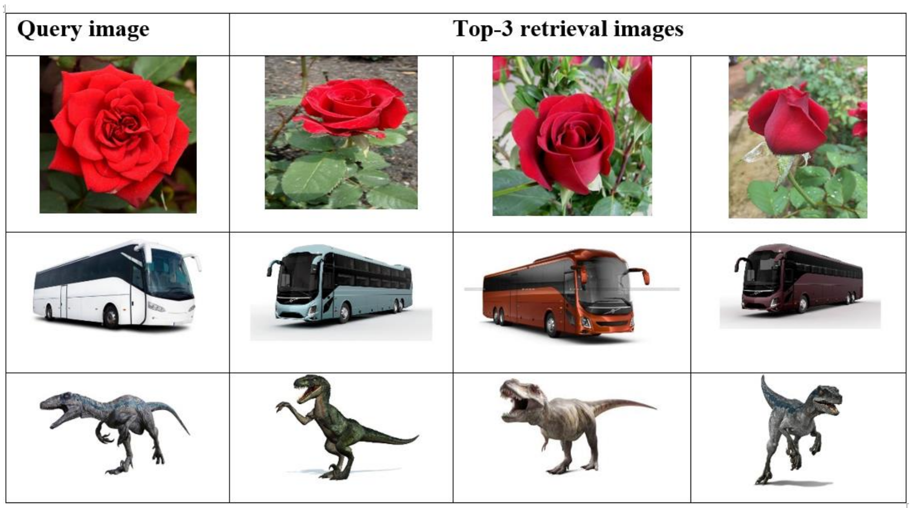

Figure 2. The proposed DEEP-FIR approach with the sample of images

The consequences of the suggested DEEP-FIR technique are exposed in Figure 2, which uses a sample of images from the public dataset to identify the top three pertinent images. Using the network characteristics of and specificity (S), precision (P), precision (P), accuracy (A), F1 score (F1) and recall (R)the efficacy of the proposed DA-CNN was computed.

$$A = \frac{(TP+FP)}{(TP+TN+FN+FP)} \tag{7}$$

$$P = \frac{TP}{TP+FP} \tag{8}$$

$$S = \frac{TN}{TN+FP} \tag{9}$$

$$R = \frac{TP}{TP+FN} \tag{10}$$

$$F1 = 2\left(\frac{P*R}{P+R}\right) \tag{11}$$

where TP and TN represents true positives and negatives of the images, FP and FN specifies false positives and negatives of the images. In this analysis, the competence of proposed and existing models was estimated using different metrics. The comparison assessment was competed among the proposed RegNet approach with different classification techniques. The comparison of traditional classification networks is illustrated in table.1.

Table 1. Comparison of traditional DL models

Models	Precision	Recall	F1 score	Specificity	Accuracy
AlexNet	87.3	84.5	87.6	85.4	84.5
LeNet	88.5	87.2	86.4	90.2	92.5
RegNet	97.2	97.5	97.7	96.8	97.8

Table 1 presents a comparison of various conventional DL networks, identifying the best classification accuracy achieved. Though, despite their utilization, classic DL networks didn't yield superior outcomes in comparison to the suggested RegNet. The suggested DCNN increases the overall accuracy by 11.0%, and 9.87% improved than AlexNet and LeNet respectively.

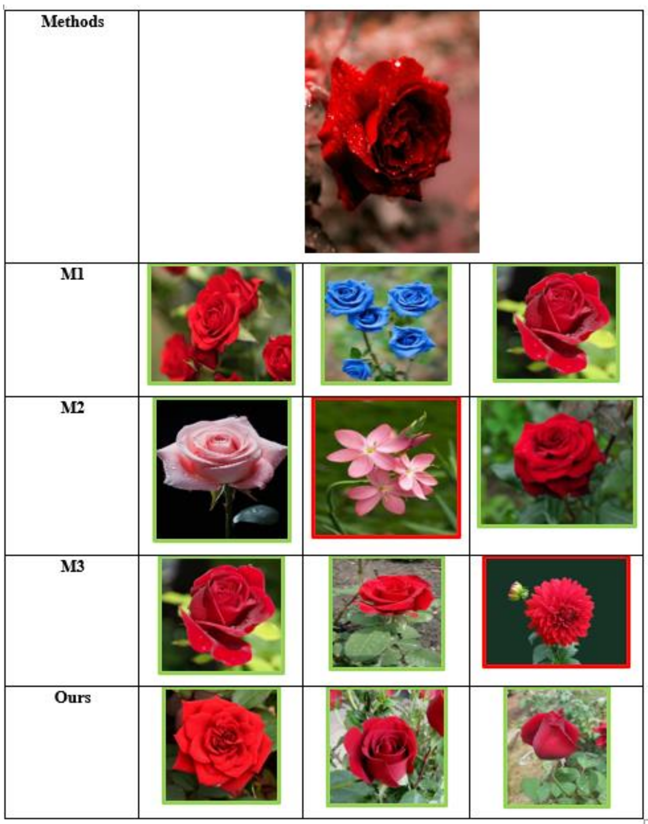

Figure 3. The proposed DEEP-FIR result

Moreover, the retrieval outcomes are examined for the query image to further evaluate the robustness of the suggested DEEP-FIR approach. The three examples were searched, one correctly classified and the other misclassified. Figure. 3 illustrate the retrieval results generated by the proposed DEEP-FIR and comparative methods like DL-CNN [11], Multi-view DNN [15] and CBIR-CNN [13] respectively. The top 3 images returned by M1, M2, M3, and our approach for the query are depicted with the remaining images being relevant, while the green boxes highlight irrelevant images. This highlights the resilience of our approach to misclassification, contrasting with the susceptibility of other methods to errors in categorization.

Table 2. Comparison of accuracy with the existing works

Authors	Methodologies	Accuracy (%)
Naeem et al., (2023) [11]	DL-CNN	94.7
Yan et al., (2020) [15]	Multi-view DNN	95.5
Singh et al., (2021) [13]	CBIR-CNN	93.2
Proposed	DEEP-FIR	97.8

Table.2 demonstrates that for better outcomes, the TP values of the image must constantly be larger and the FN values should remain as minimal as practicable. The proposed DEEP-FIR approach increases the overall accuracy of 3.16%, 2.35%, and 4.70% for DL-CNN [11], multi-view [15] and CBIR-CNN [13]. Comparing the proposed framework to the current networks, the latter have not yet achieved a satisfactory degree of accuracy.

5. CONCLUSION

This work presents DEEP-FIR approach to enhance image retrieval efficiency by integrating bio-inspired optimization and DL network. The network parameters F1 score, recall, accuracy, precision, specificity were used to assess the efficacy of the proposed DEEP-FIR technique. The proposed DEEP-FIR technique is effective, as shown by experimental findings on benchmark datasets, which show an overall accuracy of 97.8%. which was superior than the state-of-the-art approaches. The suggested DEEP-FIR method rises the overall accuracy of 3.16%, 2.35%, and 4.70% for DL-CNN, Multi-view and CBIR-CNN respectively. This research contributes to advancing image retrieval systems, offering practical solutions for real-world applications in fields like CBIR systems. Despite notable progress, challenges such as scalability, generalizability, and interpretability persist. In order to further improve the efficacy and efficiency of image retrieval systems, future research may essence on tackling these issues while investigating cutting-edge concepts like multimodal fusion, attention processes, and self-supervised learning.

CONFLICTS OF INTEREST

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

FUNDING STATEMENT

Not applicable.

ACKNOWLEDGEMENTS

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

REFERENCES

[1] W. Zhou, H. Guan, Z. Li, Z. Shao and M.R. Delavar, "Remote sensing image retrieval in the past decade: Achievements, challenges, and future directions", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 1447-1473, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] W.M. Shaban, "Early diagnosis of liver disease using improved binary butterfly optimization and machine learning algorithms", *Multimedia Tools and Applications*, vol. 83, no. 10, pp. 30867-30895, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Y. Su, L.M. Ang, K.P. Seng and J. Smith, "Deep Learning and Neural Architecture Search for Optimizing Binary Neural Network Image Super Resolution", *Biomimetics*, vol. 9, no. 6, pp. 369, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] K. Sundaramoorthi and M. Kamarasan, "A Chaotic Butterfly Optimization Driven Deep Learning Approach for Tomato Leaf Disease Detection and Classification", *In 2024 International Conference on Computing and Data Science (ICCDS)*, pp. 1-6, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Y. Wang, K. Zhao, Y. Hao and Y. Yao, Short-term wind power prediction using a novel model based on butterfly optimization algorithm-variational mode decomposition-long short-term memory. *Applied Energy*, vol. 366, pp. 123313, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] V. Suganya, S. Selvi, N. Ashokkumar and S. Prema, "An Automated Lion-Butterfly Optimization (LBO) based Stacking Ensemble Learning Classification (SELC) Model for Lung Cancer Detection", *Iraqi Journal For Computer Science and Mathematics*, vol. 4, no. 3, pp.87-100, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] R. Yasmin, A. Das, L.J. Rozario and M.E. Islam, "Butterfly detection and classification techniques: A review", *Intelligent Systems with Applications*, vol. 18, pp. 200214, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] S. Kumar, J. Pradhan and A.K. Pal, "Adaptive tetrolet based color, texture and shape feature extraction for content-based image retrieval application", *Multimedia Tools and Applications*, vol. 80, no. 19, pp. 29017-29049, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] S. Dhingra and P. Bansal, "Relative examination of texture feature extraction techniques in image retrieval systems by employing neural network: an experimental review", *In Proceedings of International Conference on Artificial Intelligence and Applications: ICAIA 2020*, pp. 337-349, Springer Singapore, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] H. Chugh, S. Gupta and M. Garg, "Image retrieval system-an integrated approach", *In IOP Conference Series: Materials Science and Engineering*, vol. 1022, no. 1, pp. 012040, 2021. IOP Publishing. [CrossRef] [Google Scholar] [Publisher Link]

[11] A. Naeem, T. Anees, K.T. Ahmed, R.A. Naqvi, S. Ahmad and T. Whangbo, "Deep learned vectors' formation using auto-correlation, scaling, and derivations with CNN for complex and huge image retrieval", *Complex & Intelligent Systems*, vol. 9. no. 2, pp. 1729-1751, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] N. Keisham and A. Neelima, "Efficient content-based image retrieval using deep search and rescue algorithm", *Soft Computing*, vol. 26, no. 4, pp. 1597-1616, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] P. Singh, P.N. Hrisheekesha and V.K. Singh, "CBIR-CNN: content-based image retrieval on celebrity data using deep convolution neural network". *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on*

Computer Science), vol. 14, no. 1, pp. 257-272, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] K. Karthik and S.S. Kamath, "A deep neural network model for content-based medical image retrieval with multi-view classification", *The Visual Computer*, vol. 37, no. 7, pp. 1837-1850, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] C. Yan, B. Gong, Y. Wei and Y. Gao, "Deep multi-view enhancement hashing for image retrieval", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 4, pp. 1445-1451, 2020. [CrossRef] [Google Scholar] [Publisher Link]

AUTHORS

S. Caroline received her B.E. degree in Electronics and Communication Engineering from Jayamatha Engineering College, Manonmaniam Sundaranar University, Tirunelveli in 2001. She obtained her M.E degree from College of Engineering, Guindy in 2004, specializing in Power Electronics and Drives. She is in teaching profession for the past eighteen years. She has published more than 10 Journal publications. She has presented more than 20 papers in National and International Conferences. She has authored a book titled Electric Circuits and Electron Devices. She has been a reviewer for Journal of Scientific Computing. Currently, she is working as an Assistant Professor in St Xavier's Catholic College of Engineering, Nagercoil. Her area of interest is Control Systems, Power Electronics and Digital Electronics. She is a Life member of ISTE and a Professional Member of IET.

Vemulapalli Sri Lakshmi Harshitha is an undergraduate student specializing in Computer Science and Design at RMK Engineering College, an esteemed autonomous institution known for its rigorous academic environment. With a strong passion for technology and innovation, Harshitha is dedicated to exploring the intersections of computer science and design, aiming to create impactful solutions through her work. Her academic journey is marked by a commitment to excellence and a keen interest in developing skills that bridge technical knowledge with creative problem-solving. Outside of her studies, she actively engages in collaborative projects and extracurricular activities, fostering a well-rounded approach to her education.

Arrived: 18.09.2024
Accepted: 20.10.2024