Indonesian Journal of Electrical Engineering and Computer Science
Vol. 37, No. 1, January 2025, pp. 551~559
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v37.i1.pp551-559

Convolutional neural network-based strategies for efficient content-based image retrieval

Chinnathambi Kamatchi[1], Rathiya Rajendran[2], Kopperundevi Nagarajan[3], Brinda Palanisamy[4], Deepika Jeyabalan[5], Rama Subramanian Paperananthamurugesan[6]
[1]Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India
[2]Information Technology, Dr. N.G.P. Institute of Technology, Coimbatore, India
[3]Department of School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India
[4]Department of Computer Science and Engineering, Veltech High Tech Dr Rangarajan Dr Sakunthala Engineering College, Avadi, India
[5]Department of Electronics and Communication Engineering, RajaRajeswari College of Engineering, Ramohalli Cross, Bangalore, India
[6]Department of Computer Science and Engineering, P.S.R Engineering College, Sivakasi, India

Article Info

Article history:

Received Jul 11, 2024
Revised Sep 11, 2024
Accepted Sep 29, 2024

Keywords:

CBIR
CNNs
Computer vision
Deep learning
Feature extraction
Image retrieval
Transfer learning

ABSTRACT

Recent years have seen a meteoric rise in the usage of enormous image databases due to advancements in multimedia technologies. One of the most critical technologies for image processing nowadays is image retrieval. This study uses convolutional neural networks (CNNs) for content-based image retrieval (CBIR). With the ever-growing number of digital photos, practical methods for retrieving these images are crucial. CNNs are incredibly efficient in many computer vision applications. Improving the efficacy and precision of image retrieval systems is the primary goal of our research into using deep learning. The paper starts with a thorough analysis of the current state of CBIR methods and the difficulties they face. Afterwards, it explores CNN's design and operation, focusing on CNN's capacity to learn hierarchical features from images autonomously. This paper also looks at how the model performs when it alters its hyperparameters, transfer learning techniques, and CNN topologies. The insights obtained from these experiments enhance the comprehension of the elements impacting CNN effectiveness in CBIR. Finally, our study shows that CNNs can change the game for image search by transforming CBIR systems. This research adds to the expanding body of information about using cutting-edge deep learning algorithms to make image retrieval more efficient and accurate.

Corresponding Author:

Brinda Palanisamy
Department of Computer science and Engineering
Veltech High Tech Dr Rangarajan Dr Sakunthala Engineering College
Avadi, India
Email: brindhaassistantprofessor@gmail.com

1. INTRODUCTION

In the current age of rapid digital information growth, efficiently accessing and retrieving pertinent visual content has become a crucial priority. The role of content-based image retrieval (CBIR) systems is significant in meeting the increasing need for efficient image search methods. The demand for sophisticated ways to categorize and retrieve images based on their content has grown significantly due to the rise of image-centric applications and the rapid expansion of digital image collections. Historically, CBIR techniques have heavily relied on manually designed features and similarity measures. However, these methods have faced challenges in effectively capturing images' intricate and diverse visual information. In recent years, computer vision has significantly transformed with the emergence of deep learning techniques [1]-[6], specifically convolutional neural networks (CNNs). This breakthrough has had a profound impact on the way computer vision tasks are approached and executed. CNNs have demonstrated remarkable proficiency in automatically acquiring hierarchical features from images. This ability allows them to effectively identify and comprehend intricate patterns and representations previously complex to capture using traditional methods.

Xia *et al.* [7] proposed a way to enable CBIR on encrypted photos while protecting sensitive data from being exposed to cloud servers in [7]. To begin, each image is represented by its feature vector. The next step in improving search speed is to develop the pre-filter database using location-sensitive hashing. After that, the protected k-nearest neighbor (KNN) model ensures the feature vector's security. To outsource the retrieval and storage of privacy-preserving images in a massive archive, Ferreira *et al.* [8] suggested a protected architecture. The IES-CBIR model is the foundation of the offered approach; the new image encryption system exhibits CBIR characteristics. This allows for private, encrypted CBIR search and storage with no compromise to security. A method that enables CBIR on encrypted photos without exposing sensitive data to cloud servers was introduced by Xia *et al.* [9]. To begin, each image is represented by an extracted feature vector. The next step in improving search efficiency is to develop the pre-filter table using location-sensitive hashing.

In addition, the protected KNN model encrypts the feature vector, and standard stream cyphers encrypt the image pixels. In addition, it is essential to consider the possibility that the user conducting the legal query may share the obtained image with someone they shouldn't. To prevent this, a watermark-based protocol could be suggested. Data owners can now safely outsource their image datasets and CBIR services to the cloud thanks to a privacy-preserving technology created by Xia *et al.* [10]. This method prevents the cloud server from gaining access to the actual database content. We use the local feature for image representation, and for similarity evaluation, we turn to earth mover's distance (EMD). An linear programming (LP) problem is the EMD computation. A new CBIR approach for integrating texture and colour features was developed by Nazir *et al.* [11]. Colour data extraction makes use of the color histogram (CH).

Using the CNN as a basis, Saritha *et al.* [12] suggested a method for image retrieval that integrates texture characteristics, edge histograms, colour histograms, and edge directions. Using the proposed method, we were able to get from the image database images that included certain content. An approach to image retrieval based on YCbCr colour using a clever edge histogram and discrete wavelet transform was presented in [13]. The framework for CBIR is made more efficient using this technique. A system that integrates the feature descriptors scale-invariant feature transform (SIFT) and binary robust invariant scalable key points (BRISK) was suggested by Sharif *et al.* [14]. Using image properties like texture, shape, and colour, Sampathila *et al.* [15] presented a method for CBIR using magnetic resonance imaging (MRI) images. Brain MRI images comparable to the query image are located and represented by the suggested method's extensive database. An approach encompassing all of an image's structural information, as retrieved from the image using descriptors like gray-level co-occurrence matrix (GLCM) and local binary patterns, was proposed in the paper [16].

The focus of this research paper is to explore the field of CBIR and investigate the potential of CNNs in improving the precision and speed of image retrieval systems. The primary objective of this study is to explore the possibility of deep learning techniques in extracting significant features from images [17]. Doing so aims to address the limitations commonly associated with conventional CBIR methods. The proposed model leverages the recent advancements in neural network architectures and transfer learning strategies to enable the development of a highly reliable CBIR system. The following sections thoroughly examine current CBIR methodologies, focusing on highlighting their respective strengths and limitations. The subsequent conversation shifts towards a comprehensive exam of CNNs, providing a detailed analysis of their structure and the underlying mechanisms that render them highly suitable for image tasks. The current research intends to add to what is already known about CBIR by presenting a novel approach that makes use of deep learning methods. This innovative approach holds significant potential in addressing the various obstacles encountered in retrieving images from heterogeneous and ever-changing datasets.

Through the implementation of rigorous experimentation and thorough comparative analyses, our objective is to present empirical evidence showcasing the efficacy of the CNN-based CBIR model that has been proposed. The experiments conducted in this study have yielded valuable insights that have the potential to advance the field of image retrieval. Additionally, these findings have broader implications for applying deep learning in various computer vision tasks. As the examination of our methodology progresses and the experimental results are presented, it becomes apparent that the incorporation of CNNs exhibits significant potential in transforming the domain of CBIR systems.

2. MATERIALS AND METHOD

2.1. Problem formulation

Let D represent a database of images, where each image is denoted as I_i and associated with a unique index i. The goal of CBIR is to develop a retrieval function R that, given a query image Q, identifies and ranks the images in D based on their content similarity to Q. Mathematically, the retrieval function R can be defined as (1):

$$R(Q, I_i) \rightarrow [0,1] \quad (1)$$

where $R(Q, I_i)$ denotes the similarity score between the query image Q and the image I_i in the database. The objective is to design a function R that maximizes the similarity score for relevant images while minimizing it for irrelevant ones. The traditional CBIR methods [18] often rely on handcrafted features and similarity metrics, which can be represented as (2).

$$R_{trad}(Q, I_i) = Sim_{trad}(F(Q), F(I_i)) \quad (2)$$

Here, $F(Q)$ and $F(I_i)$ are feature representations of the query image and the database image, respectively, and Sim_{trad} is a traditional similarity metric. In contrast, our proposed approach utilizes CNNs for feature extraction. The CNN-based retrieval function [19] is given by (3):

$$R_{CNN}(Q, I_i) = Sim_{CNN}\big(H(Q), H(I_i)\big) \quad (3)$$

where $H(Q)$ and $H(I_i)$ represent the high-level feature representations extracted by the pre-trained CNN for the query image and the database image, respectively. Sim_{CNN} is a similarity metric tailored for CNN-based features. The primary problem is optimizing the parameters of the CNN and the retrieval function R_{CNN} to achieve superior retrieval accuracy and robustness [20] compared to traditional CBIR methods. This involves addressing challenges such as adapting the CNN to diverse and complex image datasets, optimizing hyper-parameters, exploring transfer learning techniques, and investigating the impact of varying CNN topologies. The formulation of this problem sets the stage for exploring the transformative potential of CNNs in revolutionizing content-based image retrieval systems.

2.2. Dataset

The Caltech256 dataset [21] is a collection of 30,607 images of various objects. These images were sourced from Google image search and PicSearch.com. To make sure the images are relevant and of high quality, a classification system consisting of 257 categories was employed. These images were then evaluated by human experts. The Figure 1 shows the sample images from the Caltech256 dataset.

Figure 1. Sample images from the Caltech256 dataset

3. PROPOSED METHOD

The proposed methodology seeks to improve CBIR systems by utilizing CNNs to extract features. The methodology employed in this research involves a series of essential steps. These steps include the utilization of pre-trained CNNs, the extraction of hierarchical features, and the development of a robust retrieval function. The ResNet-50 architecture [22] was selected as the pre-trained CNN for feature extraction in this study. ResNet-50 is a well-known CNN architecture that is widely recognized for its deep structure and utilization of residual connections. These residual connections enable the network to effectively learn hierarchical features, which is crucial for tasks such as image classification and object detection. The incorporation of residual connections in ResNet-50 enhances the network's ability to propagate gradients through the layers, mitigating the vanishing gradient problem commonly encountered in deep networks. This characteristic of ResNet-50 contributes to its success in various computer vision applications, where the extraction of hierarchical features is essential for accurate and robust predictions. Figure 2 depicts the suggested methodology's overall working flow.

The pre-trained ResNet-50 model is employed to extract features from every image present in the database. In order to conduct our research, authors will obtain feature representations, denoted as $H(I_i)$, for each image. In this research, the approach involves removing the fully connected layers and utilizing the output obtained from the last convolutional layer. In order to create a fused representation, the low-level and high-level features extracted from the intermediate layers of ResNet-50 are concatenated. Table 1 shows the parameter details of the proposed model.

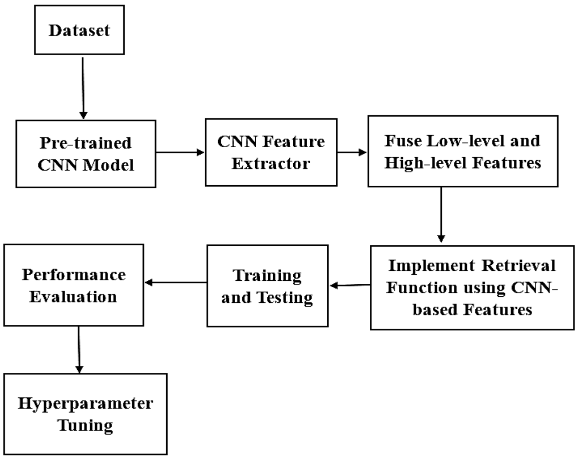

Figure 2. overall processing flow of the proposed methodology

Table 1. Parameter details of the proposed model

Model: sequential		
Layer (type)	Output shape	Param#
ResNet50 (Functional)	(None,7,7,2048)	23587712
Flatten (Flatten)	(None,100352)	0
Dense (Dense)	(None,1024)	102761472
Dropout (Dropout)	(None,1024)	0
Dense_1 (Dense)	(None,10)	10250
Total params:	12,63,49,434	
Trainable params:	10,27,71,722	
Non-trainable params:	2,35,77,712	

This process results in a fused representation denoted as $H'(I_i)$, for each individual image. In this research, author aims to develop a similarity metric specifically designed for ResNet-50-based features. The ResNet-50 architecture has gained significant popularity in computer vision tasks due to its deep convolutional layers and skip connections, enabling it to capture intricate visual patterns effectively. However, existing similarity metrics may not be optimized for ResNet-50-based features, potentially leading. To investigate the effectiveness of different similarity measures in capturing content similarity, various metrics such as cosine similarity and Euclidean distance were employed in this experiment. The objective was to determine the most optimal metric for accurately measuring the similarity between different content items. By comparing the results of these measures, insights were gained into their respective strengths and weaknesses in capturing content similarity.

$$R_{ResNet-50}(Q, I_i) = SimR_{esNet-50}\big(H'(Q), H'(I_i)\big) \tag{4}$$

Integrating the ResNet-50-based similarity metric into the retrieval function is crucial to this research. By incorporating this metric, they strive to improve efficiency and accuracy of the retrieval system. The ResNet-50 model, a widely used deep learning architecture, is the foundation for calculating the similarity between images. This metric enables us to find out how visually comparable the query images are. The retrieval function $R_{ResNet-50}$ is designed to rank images by measuring their content resemblance to the query image. This is achieved by utilizing ResNet-50 characteristics, known for their ability to capture and represent image content effectively.

The hyperparameter tuning that was carried out on the proposed ResNet-50-based CBIR model is detailed in Table 2. Several hyperparameters are available, such as the learning rate, batch size, feature dimension, epochs, dropout rate, and weight decay. The tuned values reflect the optimized configurations attained through tuning, whereas the original values served as the beginning points for investigation. In particular, the learning rate was adjusted from 0.001 to 0.0001 to encourage more consistent convergence throughout training. The generalizability of the results was improved by increasing the batch size from 32 to 64. The feature dimension, an important efficiency-influencing parameter, was lowered from 2048 to 1024. For better model performance, the number of epochs for the training iterations was increased from 50 to 100. The dropout rate was reduced from 0.5 to 0.3 as a countermeasure to overfitting. An additional regularization term, weight decay [23], was changed from 0.0001 to 0.00001. By adjusting various parameters [24], the model is fine-tuned to achieve its best performance in CBIR, considering the relative importance of convergence, generalization, and efficiency.

Table 2. Hyper parameter tuning of the proposed model

Hyperparameter	Initial value	Tuned value
Learning rate	0.001	0.0001
Batch size	32	64
Feature dimension	2048	1024
Epochs	50	100
Dropout rate	0.5	0.3
Weight decay	0.0001	0.00001

4. PERFORMANCE EVALUATION

4.1. Performance evaluation

The suggested model’s performance indicators are thoroughly examined in Table 3, which gives a snapshot of its effectiveness in a classification task. With an impressive value of 0.98, the model demonstrates remarkable accuracy in classifying cases across all categories. The model’s reliability in properly recognizing positive events is highlighted by its extraordinarily high precision of 0.99, which measures its positive predictive value. Another metric that highlights the model’s effectiveness in catching most real positive instances is recall, which measures the sensitivity or true positive rate. It is reported at 0.99. A robust F1-score of 0.98 indicates a balanced performance between recall and precision, the harmonic means of the two metrics. In addition, the receiver operating characteristic area under the curve (ROC-AUC) strongly indicates the model’s discriminatory capacity, which reaches a significant value of 0.99. According to these high-performance measures, the suggested model shows exceptional accuracy in instance classification, with sensitivity and precision above 98%. It is critical to consider the use case and dataset in context when interpreting these conclusions and to be aware of the evaluation process’s inherent limits and biases.

To see how well the model did during training, Figure 3 shows the accuracy graph for image-based content retrieval. Impressively, the training accuracy reaches a remarkable 100%, showing that the model has successfully internalized the training dataset and achieved flawless classification on the training samples. However, the testing accuracy of 0.99 demonstrates that the model can successfully apply its findings to new data while retaining a high degree of accuracy on the testing set. The fact that the model’s training and testing accuracy converged at such high numbers indicates that it learnt the data’s fundamental patterns and can discriminate between classes effectively.

Figure 4 shows the results of an image-based content retrieval loss graph that tracks how the training and testing losses change over time. The training loss of 0.00000002 shows that the model has achieved a very low level of error on the training set. The fact that the model keeps its loss low on the testing set

0.000005 in this case is evidence of its good generalizability to fresh data. It appears the model has not overfitted the training data and can successfully generalize its learnt representations to unseen samples, since the training and testing losses are closely aligned.

Table 3. Result analysis of the proposed model

Metric	Value
Accuracy	0.98
Precision	0.99
Recall	0.99
F1-score	0.98
ROC-AUC	0.99

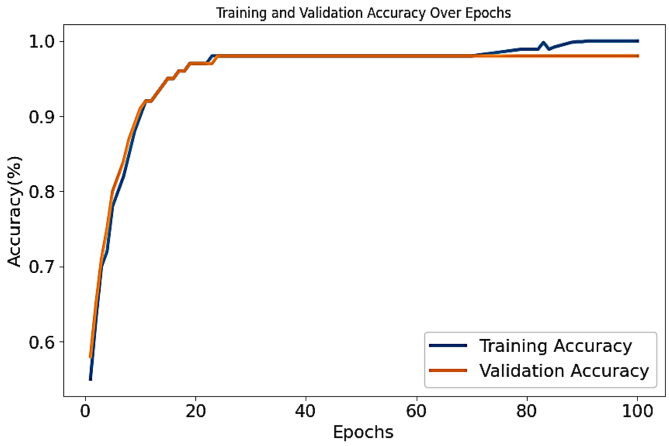

Figure 3. Accuracy graph for the image-based content retrieval

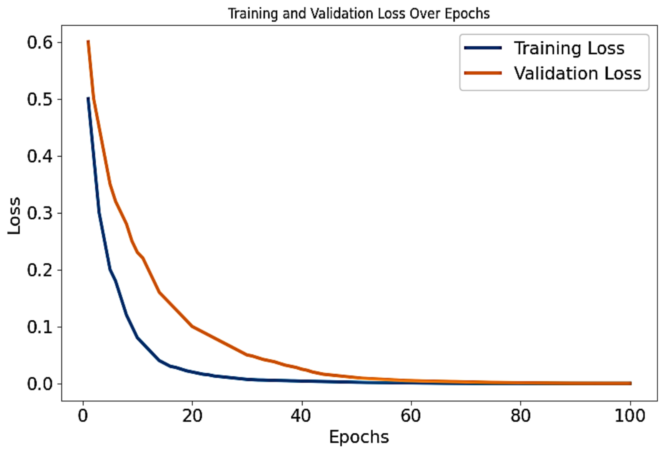

Figure 4. Loss graph for the image-based content retrieval

With DenseNet, a simple CNN, and visual geometry group (VGG)-16 as baseline models, Table 4 presents a thorough evaluation of the proposed ResNet50 model. You can see how well the models did on the task at hand by looking at the table, which displays important performance indicators including recall, accuracy, precision, and F1-score.

With an impressive accuracy of 0.98, the suggested ResNet50 model clearly excels at accurately classifying instances across all classes. Furthermore, the model can achieve high precision in positive predictions, effectively capture most positive cases, and maintain a balanced trade-off between recall and precision, as seen by the precision, recall, and F1-score values of 0.99, 0.99, and 0.98, respectively. When compared, DenseNet, CNN, and VGG16 all show impressive results. DenseNet's accuracy is 0.97, with precision at 0.98, recall at 0.97, and F1-score at 0.97. Accuracy, recall, and F1-score values for the basic CNN and VGG16 are in the 0.94 to 0.96 range, whereas they are marginally lower at 0.95 and 0.96, respectively. Figure 5, which is a vital part of our research findings, shows how the suggested ResNet50 model compares to baseline models in terms of performance. The image summarizes all of the evaluation measures in one place, making it easy to compare and contrast the models and get a feel for how well they work.

Table 4. Result comparison of the proposed model with baseline model

Model	Accuracy	Precision	Recall	F1-score
ResNet50 [25]	0.98	0.99	0.99	0.98
DenseNet [26]	0.97	0.98	0.97	0.97
CNN [27]	0.95	0.96	0.95	0.94
VGG16 [28]	0.96	0.97	0.96	0.95

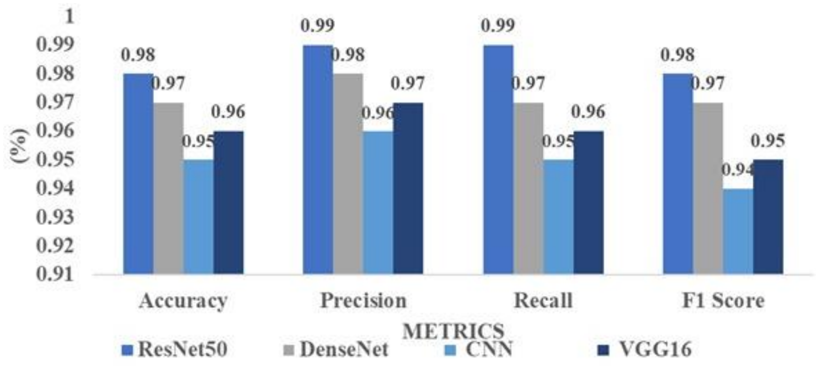

Figure 5. Result comparison of the proposed model with baseline model

5. CONCLUSION

This study extensively examines CBIR systems, with a specific focus on those that utilize deep learning architectures. The primary emphasis of our investigation revolves around the proposed ResNet50 model. Our findings indicate that this model exhibits superior performance compared to baseline models such as DenseNet, which is a conventional CNN, and VGG16.The ResNet50 model has demonstrated a high level of accuracy, as evidenced by its impressive value of 0.98. The statistical data, in conjunction with the ResNet50 model's impressive recall, precision, and F1-score, provides evidence supporting its efficacy and efficiency as a solution for CBIR tasks. The model is considered to be at the forefront of the field due to its ability to effectively handle diverse and complex datasets, as well as its autonomous capability to learn hierarchical features from images. There is an increasing amount of work on how to improve image retrieval systems using deep learning techniques, and this study adds to that body of knowledge. The ResNet50 model is widely recognized for its remarkable accuracy, positioning it as a highly desirable choice for applications that necessitate sophisticated CBIR capabilities. Although the author have made great strides in this subject, there are still many interesting directions that could be explored in the future. ResNet50's performance could be improved across domains with the use of fine-tuning methodologies and additional investigation into transfer learning approaches. Improving the suggested CBIR system's scalability, generalizability, and user-adaptability can be achieved by studying ensemble approaches, large-scale deployments, and user-centric innovations.

REFERENCES

[1] X. Zhang, C. Bai, and K. Kpalma, "OMCBIR: offline mobile content-based image retrieval with lightweight CNN optimization," *Displays*, vol. 76, p. 102355, 2023, doi: 10.1016/j.displa.2022.102355.
[2] M. Tzelepi and A. Tefas, "Deep convolutional learning for content based image retrieval," *Neurocomputing*, vol. 275, pp. 2467–2478, 2018, doi: 10.1016/j.neucom.2017.11.022.
[3] A. Latif *et al.*, "Content-based image retrieval and feature extraction: a comprehensive review," *Mathematical Problems in Engineering*, vol. 2019, no. 1, p. 9658350, 2019, doi: 10.1155/2019/9658350.
[4] M. K. Kelishadrokhi, M. Ghattaei, and S. Fekri-Ershad, "Innovative local texture descriptor in joint of human-based color features for content-based image retrieval," *Signal, Image and Video Processing*, vol. 17, no. 8, pp. 4009–4017, 2023, doi: 10.1007/s11760-023-02631-x.
[5] S. F. Salih and A. A. Abdulla, "An effective bi-layer content-based image retrieval technique," *Journal of Supercomputing*, vol. 79, no. 2, pp. 2308–2331, 2023, doi: 10.1007/s11227-022-04748-1.

[6] R. Gomathi, S. Logeswari, S. Jothimani, S. N. Sangeethaa, S. A. Sangeetha, and V. LathaJothi, "MEFNet-Micro expression fusion network based on micro-attention mechanism and 3D-CNN fusion algorithms," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 6, pp. 113–122, 2023, doi: 10.22266/ijies2023.1231.10.

[7] Z. Xia, N. N. Xiong, A. V. Vasilakos, and X. Sun, "EPCBIR: an efficient and privacy-preserving content-based image retrieval scheme in cloud computing," *Information Sciences*, vol. 387, pp. 195–204, 2017, doi: 10.1016/j.ins.2016.12.030.

[8] B. Ferreira, J. Rodrigues, J. Leitao, and H. Domingos, "Privacy-preserving content-based image retrieval in the cloud," in *Proceedings of the IEEE Symposium on Reliable Distributed Systems*, 2016, vol. 2016-Janua, pp. 11–20, doi: 10.1109/SRDS.2015.27.

[9] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, and K. Ren, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2594–2608, 2016, doi: 10.1109/TIFS.2016.2590944.

[10] Z. Xia, Y. Zhu, X. Sun, Z. Qin, and K. Ren, "Towards privacy-preserving content-based image retrieval in cloud computing," *IEEE Transactions on Cloud Computing*, vol. 6, no. 1, pp. 276–286, 2018, doi: 10.1109/TCC.2015.2491933.

[11] A. Nazir, R. Ashraf, T. Hamdani, and N. Ali, "Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor," in *2018 International Conference on Computing, Mathematics and Engineering Technologies: Invent, Innovate and Integrate for Socioeconomic Development, iCoMET 2018 - Proceedings*, 2018, vol. 2018-Janua, pp. 1–6, doi: 10.1109/ICOMET.2018.8346343.

[12] R. R. Saritha, V. Paul, and P. G. Kumar, "Content based image retrieval using deep learning process," *Cluster Computing*, vol. 22, pp. 4187–4200, 2019, doi: 10.1007/s10586-018-1731-0.

[13] R. Ashraf *et al.*, "Content based image retrieval by using color descriptor and discrete wavelet transform," *Journal of Medical Systems*, vol. 42, no. 3, pp. 1–12, 2018, doi: 10.1007/s10916-017-0880-7.

[14] U. Sharif, Z. Mehmood, T. Mahmood, M. A. Javid, A. Rehman, and T. Saba, "Scene analysis and search using local features and support vector machine for effective content-based image retrieval," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 901–925, 2019, doi: 10.1007/s10462-018-9636-0.

[15] N. Sampathila, Pavithra, and R. J. Martis, "Computational approach for content-based image retrieval of K-similar images from brain MR image database," *Expert Systems*, vol. 39, no. 7, p. e12652, 2022, doi: 10.1111/exsy.12652.

[16] M. Garg and G. Dhiman, "A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants," *Neural Computing and Applications*, vol. 33, no. 4, pp. 1311–1328, 2021, doi: 10.1007/s00521-020-05017-z.

[17] Y. Xing, B. J. Meyer, M. Harandi, T. Drummond, and Z. Ge, "Multimorbidity content-based medical image retrieval and disease recognition using multi-label proxy metric learning," *IEEE Access*, vol. 11, pp. 50165–50179, 2023, doi: 10.1109/ACCESS.2023.3278376.

[18] N. M. Hai, T. Van Lang, and T. The Van, "Improving the efficiency of semantic image retrieval using a combined graph and SOM model," *IEEE Access*, vol. 11, pp. 140646–140659, 2023, doi: 10.1109/ACCESS.2023.3333678.

[19] X. Li, J. Yu, S. Jiang, H. Lu, and Z. Li, "MSViT: training multiscale vision transformers for image retrieval," *IEEE Transactions on Multimedia*, vol. 26, pp. 2809–2823, 2024, doi: 10.1109/TMM.2023.3304021.

[20] G. Song, K. Huang, H. Su, F. Song, and M. Yang, "Deep ranking distribution preserving hashing for robust multi-label cross-modal retrieval," *IEEE Transactions on Multimedia*, vol. 26, pp. 7027–7042, 2024, doi: 10.1109/TMM.2024.3358995.

[21] "Caltech vision lab," *www.vision.caltech.edu*. https://www.vision.caltech.edu/Image_Datasets/Caltech256/.

[22] J. H. Kim, A. Poulose, and D. S. Han, "CVGG-19: customized visual geometry group deep learning architecture for facial emotion recognition," *IEEE Access*, vol. 12, pp. 41557–41578, 2024, doi: 10.1109/ACCESS.2024.3377235.

[23] X. Jia, X. Feng, H. Yong, and D. Meng, "Weight decay with tailored adam on scale-invariant weights for better generalization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 5, pp. 6936–6947, 2024, doi: 10.1109/TNNLS.2022.3213536.

[24] S. Kannan, D. Prabakaran, S. D. Kumar, and S. Sivaram, "A deep learning-based convolution neural networks to forecast wind energy," in *ICRTEC 2023 - Proceedings: IEEE International Conference on Recent Trends in Electronics and Communication: Upcoming Technologies for Smart Systems*, 2023, pp. 1–6, doi: 10.1109/ICRTEC56977.2023.10111917.

[25] I. Issaoui, M. A. Alohali, W. Mtouaa, F. A. Alotaibi, A. Mahmud, and M. Assiri, "Archimedes optimization algorithm with deep learning assisted content-based image retrieval in healthcare sector," *IEEE Access*, vol. 12, pp. 29768–29777, 2024, doi: 10.1109/ACCESS.2024.3367430.

[26] C. Liu, W. Ding, C. Cheng, C. Tang, J. Huang, and H. Wang, "DenseHashNet: a novel deep hashing for medical image retrieval," *IEEE Journal of Radio Frequency Identification*, vol. 6, pp. 697–702, 2022, doi: 10.1109/JRFID.2022.3209986.

[27] L. Wang, X. Qian, Y. Zhang, J. Shen, and X. Cao, "Enhancing sketch-based image retrieval by CNN semantic re-ranking," *IEEE Transactions on Cybernetics*, vol. 50, no. 7, pp. 3330–3342, 2020, doi: 10.1109/TCYB.2019.2894498.

[28] Q. Qin, L. Huang, Z. Wei, K. Xie, and W. Zhang, "Unsupervised deep multi-similarity hashing with semantic structure for image retrieval," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 7, pp. 2852–2865, 2021, doi: 10.1109/TCSVT.2020.3032402.

BIOGRAPHIES OF AUTHORS

Dr. Chinnathambi Kamatchi is working as assistant professor in the Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu. He is Doctorate in Computer Science having 7 years of research experience. He has published more research articles in international journals. He can be contacted at email: drchinnathambik@veltech.edu.in.

Rathiya Rajendran received her B.E degree in Computer Science and Engineering from Ponjesly College of Engineering, Kanyakumari, India in 2011 and M.E degree in Computer Science and Engineering from Noorul Islam University, Kanyakumari, India in 2013. Currently pursuing Ph.D. under Anna University, Chennai, India. Her research interest includes machine learning and deep learning. Published 6 papers in international conference and three patent publications. And also, a life time member of ISTE. She can be contacted at email: vr.rathiya@gmail.com.

Dr. Kopperundevi Nagarajan is currently working as assistant professor at Vellore Institute of Technology Vellore. She completed her research the year 2019 under Anna University Chennai, in the specialization of Information Communication Engineering. She has authored one book, published more than 15 papers and three patents. Acting as a reviewer in Textile Research Journal and Soft Computing has reviewed nearly 10 papers. Her area of interest includes image processing, neural networks, fuzzy systems machine learning and deep learning and have guided more than 15 UG and PG projects. She can be contacted at email: kopperundevi.n@vit.ac.in.

Brinda Palanisamy is working as an assistant professor VelTech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College from Mar 2022 to till date. She can be contacted at email: brindhaassistantprofessor@gmail.com.

Dr. Deepika Jeyabalan working as an associate professor in the department of ECE of RajaRajeswari College of Engineering, Bangalore. She has completed her B.E. and M.Tech in ACS College of Engineering. She pursued her Doctorate in Philosophy (Engineering) at Visvesvaraya Technological University, Belgavi, India. She is a recognized Research supervisor under VTU in RRCE Research center. Her area of research is on mobile ad-hoc networks. She can be contacted at email: deepika7193@gmail.com.

Mr. Rama Subramanian Paperananthamurugesan is currently working as an assistant professor, in the Department of Computer Science and Engineering, P.S.R. Engineering College. He is having 7 years of teaching experience He has handled interesting roles and responsibilities including Teaching in various engineering Institutions. Areas of interest and specialization includes networks, data structure, database management systems, machine learning, deep learning, and data analytics. He has published 3 papers, 2 conference articles so on. He can be contacted at email: ramboga29@gmail.com.