Autoencoder and CNN for Content-based Retrieval of Multimodal Medical Images

Suresh Kumar J S[1], Maria Celestin Vigila S[2]
Research Scholar, Department of Computer Science and Engineering,
Noorul Islam Centre for Higher Education, Kumaracoil, Kanyakumari, Tamil Nadu, India[1]
Associate Professor, Department of Information Technology,
Noorul Islam Centre for Higher Education, Kumaracoil, Kanyakumari, Tamil Nadu, India[2]

***Abstract*—Content-Based Medical Image Retrieval (CBMIR) is a widely adopted approach for retrieving related images by the comparison inherent features present in the input image to those stored in the database. However, the domain of CBMIR specific to multiclass medical images faces formidable challenges, primarily stemming from a lack of comprehensive research in this area. Existing methodologies in this field have demonstrated suboptimal performance and propagated misinformation, particularly during the crucial feature extraction process. In response, this investigation seeks to leverage deep learning, a subset of artificial intelligence for the extraction of features and elevate overall performance outcomes. The research focuses on multiclass medical images employing the ImageNet dataset, aiming to rectify the deficiencies observed in previous studies. The utilization of the CNN-based Autoencoder method manifests as a strategic choice to enhance the accuracy of feature extraction, thereby fostering improved retrieval results. In the ImageNet dataset, the results obtained from the proposed CBMIR model demonstrate notable average values for accuracy (95.87%), precision (96.03%) and recall (95.54%). This underscores the efficacy of the CNN-based autoencoder model in achieving good accuracy and underscores its potential as a transformative tool in advancing medical image retrieval.**

Keywords*—Medical image retrieval; multiclass medical images; artificial intelligence; deep learning; convolutional neural network; autoencoder

I. Introduction

Content Based Medical Image Retrieval (CBMIR) plays a pivotal part in modern healthcare, leveraging the developments in Deep Learning (DL) to enhance efficiency and accuracy of diagnosing and treating various medical conditions. As medical imaging technologies continue to evolve, the vast amount of digital medical images generated necessitates robust and intelligent retrieval systems. DL, a subset of Artificial Intelligence (AI), has emerged as a transformative force in the field, offering unprecedented capabilities in feature extraction and pattern recognition. This research delves into the area of CBMIR, exploring the application of DL techniques to navigate and retrieve relevant information from extensive medical image databases [1]. The integration of DL models not only streamlines the retrieval process but also contributes to the overall improvement of diagnostic accuracy and clinical decision-making. In this comprehensive examination, we delve into the key methodologies, challenges, and breakthroughs associated with CBMIR using DL, shedding light on the promising future it holds for the medical community [2].

The fusion of cutting-edge technology and healthcare exemplifies a synergy that has the potential to revolutionize patient care and medical research. Through an exploration of various DL architectures and their adaptations to the intricacies of medical images, this work provides implications for the future of CBMIR. In navigating the intricate landscape of medical data, DL proves to be an invaluable tool, offering a paradigm shift in how medical professionals access, analyze, and leverage the wealth of information embedded in medical images [3]. The integration of DL into CBMIR not only addresses the challenges posed by the sheer volume of data but also opens avenues for novel insights, early disease detection, and personalized treatment strategies. As we embark on this exploration, it is evident that the combination of medical imaging and DL is poised to redefine the landscape of healthcare, paving the way for more precise diagnoses, timely interventions, and improved patient outcomes [4].

DL can automatically obtain hierarchical representations from images and provide a compelling solution to the complex task of CBMIR. By leveraging Convolutional Neural Network (CNN) and other sophisticated architectures, DL models can discern subtle patterns and relationships within medical images that may elude traditional retrieval methods [5]. The synergy between the depth of neural networks and the intricacies of medical image content enables the extraction of high-level features crucial for accurate retrieval and analysis. Heterogeneity of imaging modalities, ranging from X-rays and MRIs to CT scans and beyond [6] is a serious issue. DL algorithms, through transfer learning and domain adaptation, demonstrate their adaptability to diverse imaging sources, promising a unified framework for efficient retrieval across modalities [7].

Autoencoders, with their capacity to learn compact and informative representations of input data, are examined for their utility in extracting latent features from medical images [8]. Whether applied to X-rays, MRIs, or CT scans, autoencoders demonstrate their versatility in capturing intrinsic features, thereby enhancing the robustness of CBMIR systems across a spectrum of medical image types. An in-depth discussion on the potential synergy between autoencoders and other DL architectures, such as CNN, is presented [9]. The combination of these models provides a comprehensive framework for medical image retrieval, where autoencoders contribute to feature extraction and CNNs leverage these extracted features for accurate CBMIR. Furthermore, the paper addresses the interpretability of autoencoder-based models, highlighting how the encoded representations can be harnessed for visualizing and understanding the salient features within medical images [10]. This interpretative aspect not only fosters trust in the model's decision-making process but also facilitates the identification of clinically relevant patterns that may elude traditional CBMIR techniques. The major contribution of the proposed work includes:

- Integration of two powerful techniques, autoencoder and CNN, to tackle content-based retrieval of multimodal medical images.
- Provides a solution to efficiently search and retrieve images based on their content, enabling better diagnosis, treatment planning, and research in healthcare.
- Enhances retrieval accuracy and efficiency compared to traditional methods by leveraging deep learning techniques.

II. Literature Review

CBMIR witnessed transformative advancements with the integration of DL techniques. In the area of multimodal image representation, the effective retrieval of relevant medical images plays a pivotal role in ensuring accurate and timely diagnoses. This survey examines the recent developments and contributions of DL methods, specifically focusing on CBMIR for multimodal diagnosis images.

Ozturk [11] introduced an approach for radiological image retrieval by employing deep features extracted through CNN. The study demonstrates an enhancement in retrieval performance, showcasing the potential of DL in streamlining radiological diagnosis. The automated feature extraction process proves crucial in improving the efficiency of the diagnostic workflow, providing valuable insights for the integration of DL in medical imaging applications. Liu et al. [12] propose an innovative technique utilizing autoencoder architectures for feature extraction in cross-modality image retrieval. The research highlights the versatility of autoencoders in handling various medical imaging modalities, showcasing improved performance and robustness. By minimizing misinformation, this work contributes significantly to the reliability and accuracy of cross-modality image retrieval, offering potential advancements in diagnostic capabilities across diverse imaging technologies. Cai et al. [13] conducted a comprehensive comparative analysis of multiple CNN architectures for medical image retrieval. The findings reveal substantial variations in retrieval accuracy based on the selected network architecture, providing critical insights for practitioners in choosing optimal models for specific medical imaging applications. This study underscores the influence of network architecture on the performance of medical image retrieval systems, aiding informed decision-making for the development and implementation of DL technologies in clinical settings.

Li et al. [14] analyzed the robustness of CNN-based autoencoders in the realm of CBMIR. The study evaluates the performance of these models across various medical imaging modalities and assesses their ability to handle noisy or low-quality images. By investigating the robustness of CNN-based autoencoders, the research contributes valuable insights into the reliability of these models in real-world clinical scenarios. The findings provide guidance on the potential challenges and opportunities in deploying such models for CBMIR applications. Guan et al. [15] concentrate on improving the interpretability of features extracted by CNN in CBMIR. The study introduces methodologies for visualizing and understanding critical features, enhancing transparency in the decision-making process. This research marks a crucial step toward building trust in DL models and refining their interpretative capabilities for real-world medical applications. Shen et al. [16] explored the application of federated learning for privacy-preserving CBMIR. This research underscores the potential of federated learning in maintaining data security while advancing CBMIR capabilities.

Liu et al. [17] focused on investigating semi-supervised DL approach in CBMIR. This work demonstrates the potential of leveraging unlabeled data to enhance model performance, addressing challenges associated with limited labeled datasets. By incorporating semi-supervised learning, the study contributes to the adaptability of CBMIR systems across diverse clinical scenarios. Bouchareb et al. [18] delve into ethical considerations associated with AI-driven diagnostic imaging. The study emphasizes transparency, accountability, and the mitigation of biases in the deployment of DL models for diagnostic purposes. By addressing ethical concerns, this research contributes to responsible AI practices in the evolving landscape of CBMIR. Swati et al. [19] provide a broader perspective by exploring applications of DL in precision medicine for medical imaging. The study highlights the potential for personalized treatment strategies based on CBMIR results, showcasing the transformative impact of DL in tailoring medical interventions to individual patient needs. Jaiswal et al. [20] focused on exploring the applications of transfer learning in the context of CBMIR. By adapting knowledge learned from one domain to another, transfer learning proves to be a valuable strategy for addressing challenges associated with limited labeled medical image datasets. The study provides insights into the potential of transfer learning to improve the generalization capabilities of DL models in the medical imaging domain. This research contributes to the ongoing efforts in making medical image retrieval systems more adaptable and effective in diverse clinical settings.

This review showcased the evolution of CBMIR in multimodal diagnosis, emphasizing the transition from traditional CBMIR to sophisticated DL models. The integration of CNN-based Autoencoders presents a promising avenue for addressing challenges in feature extraction and enhancing overall performance. This research continues to explore innovative methodologies and datasets to advance the capabilities of DL in CBMIR for multimodal diagnosis.

III. Materials and Methods

In this research, we present a novel CBMIR system designed specifically for multimodal diagnosis images associated with various diseases. Our primary focus is to overcome limitations identified in previous studies, and to achieve this, we employ the CNN-based autoencoder methodology. The rationale behind adopting the CBMIR approach lies in its demonstrated efficacy in optimizing both the feature extraction process and the learning phase. This strategic utilization aims to rectify and minimize inaccuracies that may have been prevalent in earlier research endeavors. The CNN-based Autoencoder method plays a crucial role in amending misinformation issues that might have arisen in the feature extraction process during previous studies. By leveraging the power of DL, this approach not only refines the accuracy of feature extraction but also contributes to an overall improvement in the system's performance in essence; this research positions the CBMIR system, enhanced by the CNN-based autoencoder method, as a robust solution for retrieving multimodal diagnosis images. By addressing and mitigating misinformation concerns, we aim to contribute to the advancement of CBMIR methodologies, fostering precision and reliability in the context of subclass dataset categorization. The process flow of proposed methodology is depicted in Fig. 1.

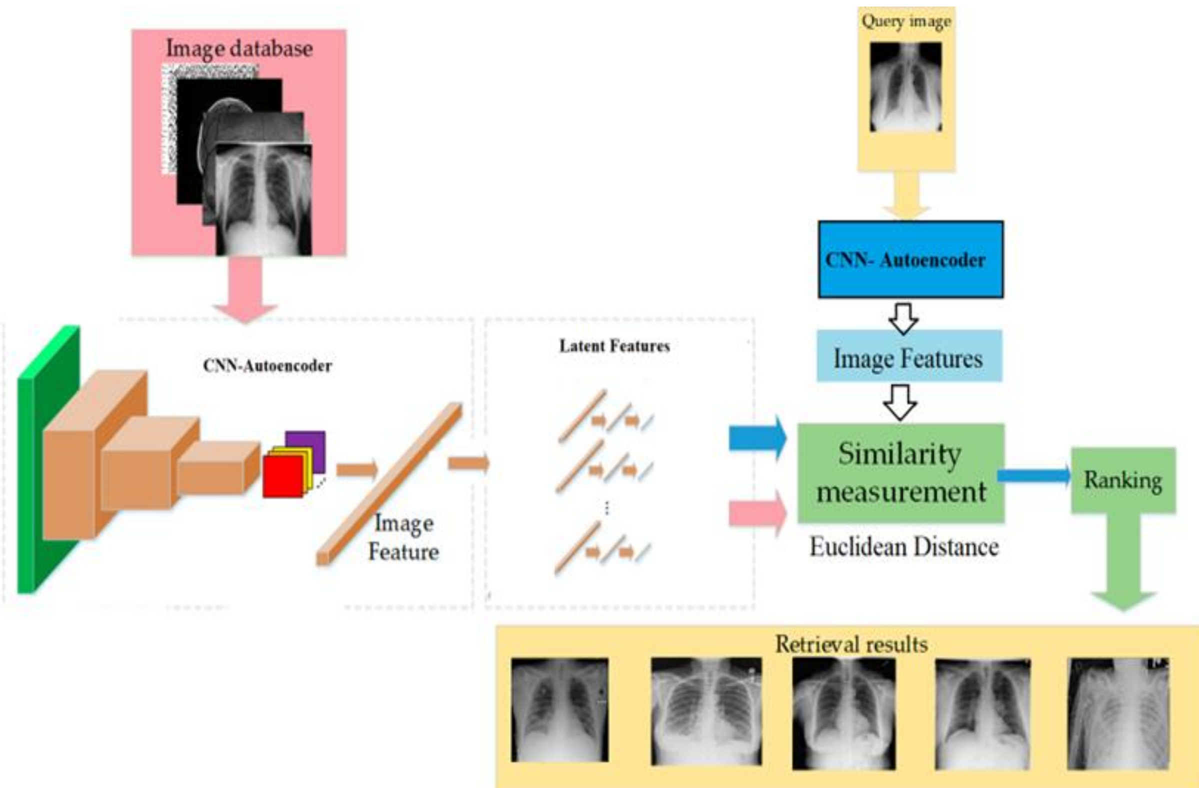

Fig. 1. Basic block diagram of proposed method.

A. Dataset Description

The dataset comprising around 50,000 2D images, sourced from diverse and easily accessible open-access medical databases like ImageNet. The primary objective of this compilation is to enable the differentiation among Magnetic Resonance Imaging (MRI), X-ray Electroencephalograph (EEG), and OCT. ImageNet stands as a vast image repository that has played a pivotal role in propelling forward the realms of computer vision and DL research. Notably, the dataset encompasses images of varying data sizes, providing a comprehensive and diverse set of examples for training and evaluation purposes in the context of multimodal medical image analysis. The dataset has been categorized into four primary classes, encompassing X-ray, MRI, OCT, and EEG modalities. A visual representation of a sample image from the dataset, utilized in the present study, is depicted in Fig. 2. This division into distinct classes serves as a foundational structure for the dataset, facilitating a nuanced exploration and analysis of varied medical imaging modalities.

B. Proposed Model Architecture

The Autoencoder process involves three key stages: Encoder, Decoder, and the computation of the Calculating Function and Optimization Errors. During the encoding phase, input data undergoes a transformation into smaller dimensions, often referred to as compression. Employing Conv2D (2D Convolution Layer), the input image is transformed to 48 nodes (latent dimension). This latent representation must then be converted back to initial state. Decoding process utilizes transpose operation to generate the reconstructed image. Loss calculation for each function is iteratively performed to determine the function with the lowest loss value. In this research, the selected loss function is Mean Square Error (MSE). Fig. 3 illustrates the structure of designed autoencoder.

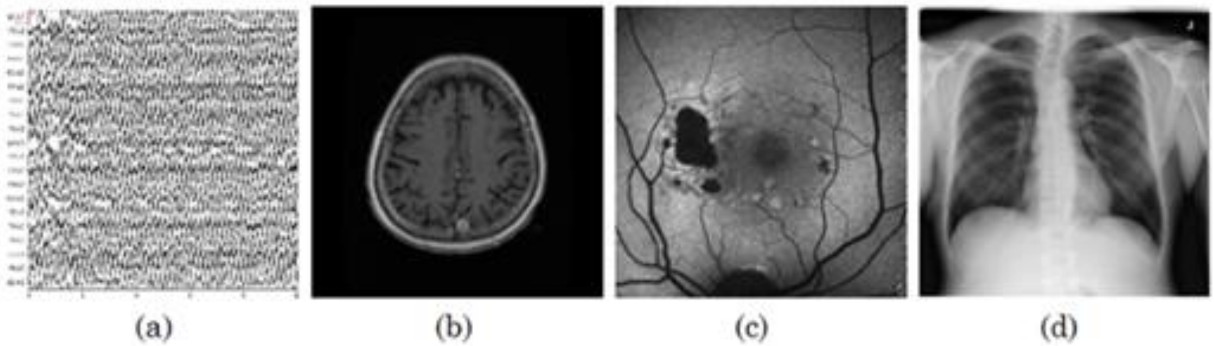

Fig. 2. Sample images from dataset (a) EEG, (b) MRI, (c) OCT, (d) X-Ray.

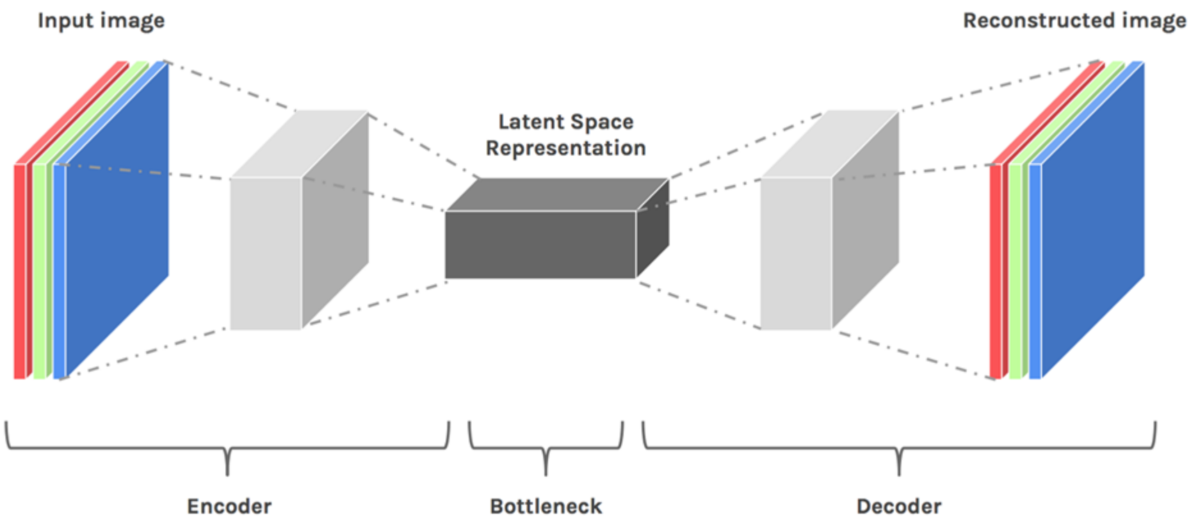

Fig. 3. Structure of the proposed autoencoder.

The commencement of the training process in this investigation involves segregating images into three subsets: training (80%), validation (10%), and testing (10%). Employing the CNN- Autoencoder model, the original image is transformed into a reconstructed image. The Autoencoder involve in the extraction of output image to reconstruct the input image, facilitating a comparison with the original input image [21]. Following this, the learning process initiates, and through numerous iterations, the optimal model is obtained and saved, featuring the lowest loss value in relation to the extraction and retrieval stages. The training data includes the segmented data for training, while the validation data evaluates the developed model. Additional explanation of phases in the training process is provided in Fig. 4.

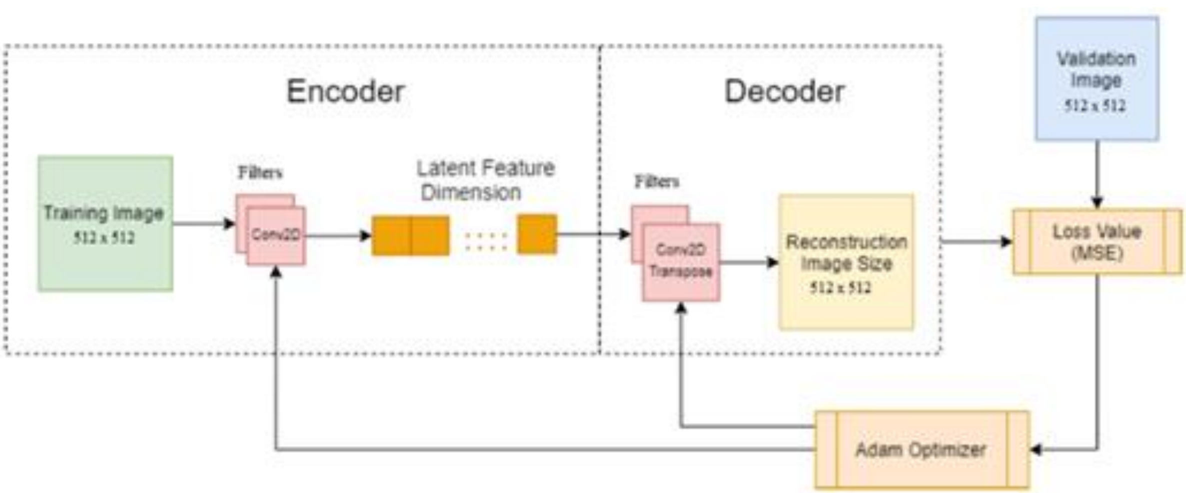

Fig. 4. Process flow of training.

During the encoder stage, the input data undergoes compression. Conv2D, with modified default parameters, employs a stride of 2 for convolution, along with "same" padding value. This ensures even distribution of zero-value padding convolution. With 128 filters having 3x3 kernels, Leaky Relu is applied, which is recognized for its effectiveness in overcoming gradient loss issues in DL. Leaky Relu is preferred over its predecessors, sigmoid and Tanh, for its simplicity and improved performance [22]. The Leaky Relu function is defined by the following equation.

$$f(x) = \begin{cases} x, x \geq 0 \\ ax, x < 0 \end{cases} \quad (1)$$

The LeakyRelu activation function incorporates a parameter known as alpha, representing the negative slope coefficient, and in this study, it is set to 0.2.

In a parallel arrangement to the encoder, the decoder employs a convolution method with 128 filters and a filter having 3x3 kernels, maintaining the same activation function but with a distinct approach to processing. The decoder employs transposed data from the latent dimension to produce the reconstructed image. Upon the conclusion of both the encoder and decoder stages, the procedure progresses to the error calculation and optimization phase. In this phase, iterative loss calculations are performed for each function, aiming to pinpoint the function with the minimum loss value [23]. The chosen loss function for this study is the Mean Square Error (MSE), calculated through Eq. (2). Here, pi is the predicted image and yi is the actual image.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(p_i - y_i)^2 \quad (2)$$

In the process of optimizing CNNs, the weighted values and biases of the convolutional layers undergo updates through the utilization of the Adam's algorithm. This algorithm, known for its efficiency in optimization tasks, plays a crucial role in adjusting these parameters. To initiate the learning process, a standard learning rate of 1e-3 is employed, serving as a foundational value for the optimization algorithm. Moreover, a learning rate scheduler is integrated into the training procedure, incorporating a decay parameter set at a rate of 2e-5. This scheduler dynamically adjusts the learning rate during the training epochs, enabling a gradual decrease in the learning rate over time. This adaptive learning rate scheduling contributes to the model's stability and convergence during the optimization process. By leveraging these techniques, the convolutional layers continually adapt their weighted values and biases to improve its ability to obtain meaningful features. The thoughtful integration of optimization strategies, such as the Adam's optimization algorithm and learning rate scheduling, is essential for achieving robust and effective performance in CNN training.

Feature extraction involves the retrieval of distinctive attributes from an image, which are subsequently scrutinized for subsequent processes. Following this, the acquired features are recognized to establish distinctions between images [24]. As depicted in Fig. 5, extraction of features is executed by utilizing the stored optimal model.

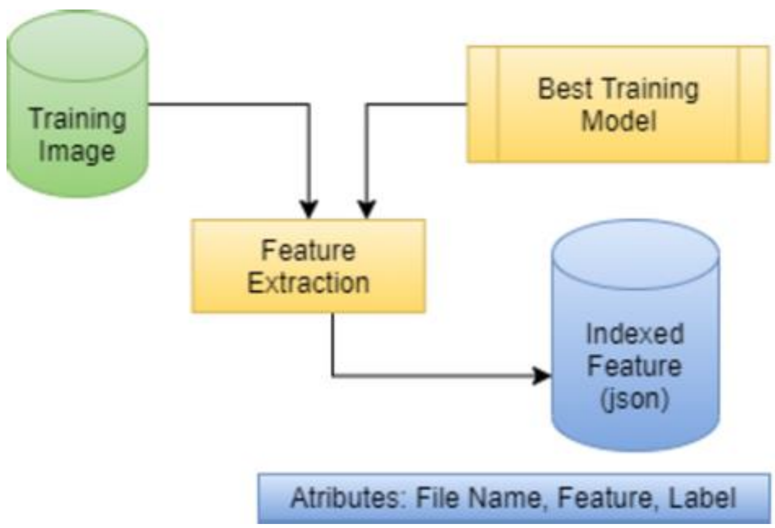

Fig. 5. Feature extraction process.

The outcomes of this feature extraction process are systematically cataloged to create a “.json” (JavaScript Object Notation) file. This file not only includes the extracted features but also incorporates the image file names and labels associated with each image. The compilation of annotated images serves as a comprehensive database, pivotal for subsequent comparisons during the retrieval stage.

C. Image Retrieval

The retrieval process involves seeking identical images with reference to query image. The data, previously segregated into test data, undergoes feature extraction, mirroring the procedure in the preceding stage. Following the successful extraction of images, calculations ensue to determine the resemblances of features in test images and those indexed in the training set, saved in the “.json” file. In this work, the Euclidean Distance is specifically applied to quantify the distance similarity between two image vectors. A detailed depiction of the retrieval process is provided in Fig. 6.

Euclidean Distance is utilized for efficiently calculating the similarity distance between two vectors, irrespective of their dimensionality—be it two-dimensional, three-dimensional, or beyond [25]. The comparison of the two feature vectors extracted from the features involves calculating the distance between them. This meticulous process ensures a robust evaluation of the similarity between images, employing a widely recognized distance calculation method within the domain of image retrieval. The Euclidean Distance is expressed through the following equation, frequently employed for measuring the distance between vectors [26]. Here de is the Euclidean distance, qi is the vector query image and bi is the vector present in train data.

$$d_e = \sqrt{\sum_{i=1}^{n}(q_i - b_i)^2} \quad (3)$$

Following the computation of distances between each image in the test data and all images within the training data, the resulting distance values are subjected to comparison. A smaller Euclidean value indicates a higher degree of similarity between images. As a result, the image possessing the smallest distance value is included in the retrieval process. In essence, this retrieval methodology prioritizes images that exhibit the closest proximity in terms of features, as quantified by the Euclidean distance. The selection of the top five most similar images enhances the comprehensiveness and accuracy of the retrieval process, providing a nuanced understanding of image similarity.

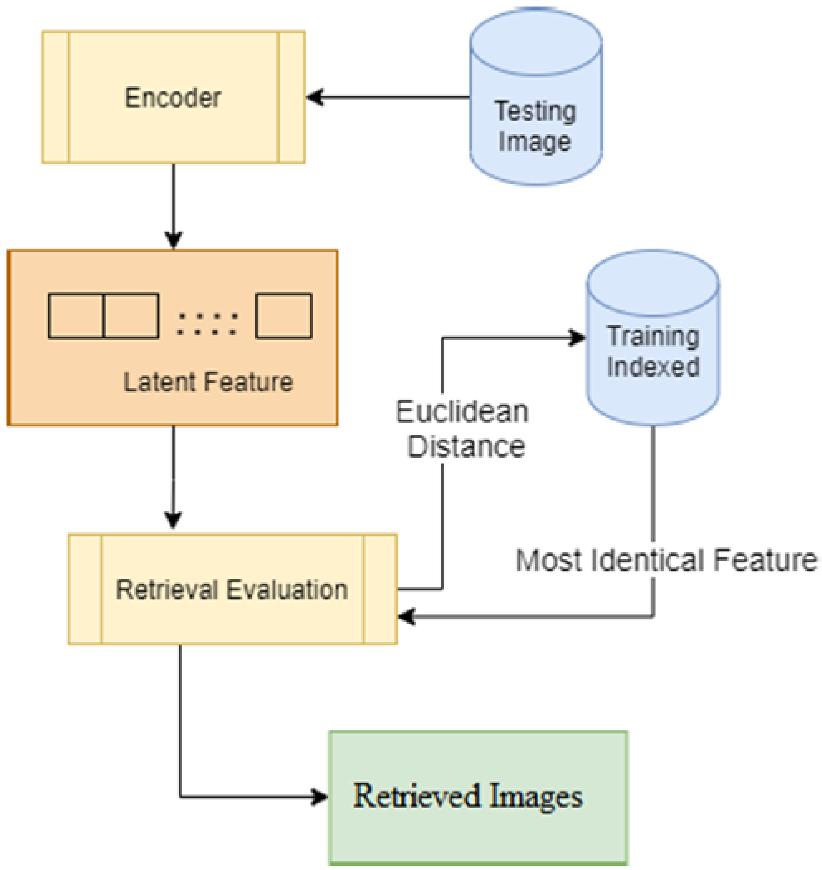

Fig. 6. Image retrieval process.

IV. Results and Discussion

The evaluation stage is crucial in assessing the effectiveness of the system developed within this research [27]. Precision and recall metrics are actively employed to gauge the success rate of the system, providing insightful measures of its performance. Initially, the retrieval process flow begins by taking an image from the test data as input. Utilizing the saved encoder model from the training phase, the system generates latent features and proceeds to the retrieval evaluation stage. In this stage, Euclidean distances are calculated based on the indexed .json file, employing information obtained from the previous feature extraction process flow. The ultimate output of this process includes precision and recall values, which serve as key indicators of the system's retrieval performance. Eq. (4) and (5) outline the formulas employed for calculating precision and recall in this context. Here Ir represents the count of retrieved images, n represents the count of images captured and M denotes relevant images present in the database.

$$Precision = \frac{I_r}{n} \quad (4)$$

$$Recall = \frac{I_r}{M} \quad (5)$$

As depicted in Fig. 7, the training phase during the testing of the dataset in all categories exhibits a gradual reduction in loss values up to epoch 35. The cumulative results across all categories yield loss values below 0.2, indicating a successful retrieval process. This signifies the efficacy of the training process in minimizing errors, enabling a robust and accurate retrieval of images across various imaging modalities.

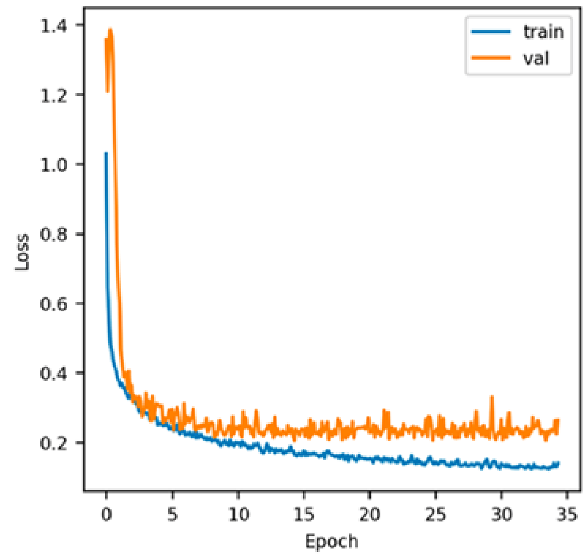

Fig. 7. Loss plot of proposed model.

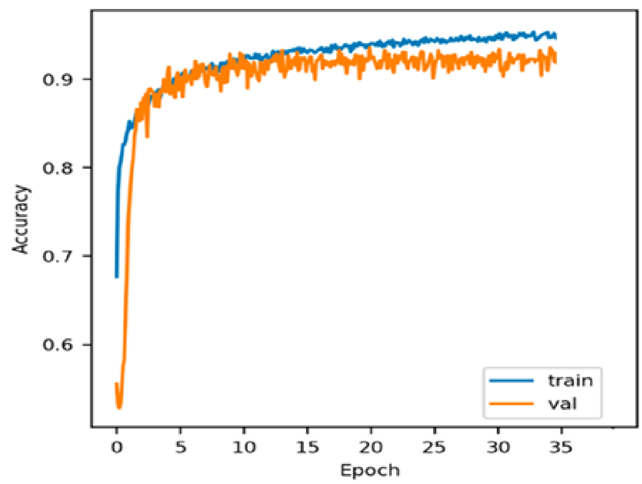

Fig. 8. Accuracy plot of proposed model.

As illustrated in Fig. 8, the training phase showcases a steady rise in accuracy values for all categories up to the 5th epoch, maintaining a consistent level thereafter. The collective outcomes across all categories yield an accuracy value surpassing 94%, underscoring the success of the retrieval process. This highlights the effectiveness of the training process in minimizing errors, facilitating a resilient and precise retrieval of images across diverse imaging modalities.

Upon conducting image retrieval using the implemented CBMIR model, the evaluation process provides crucial insights into the system's effectiveness, quantified through accuracy, precision and recall values. These metrics serve as key indicators of the system's ability to accurately retrieve relevant medical images based on content. The comparison of system performance involves assessing the proposed CBMIR system against varying number of retrieved images. This comparative analysis is conducted separately for each dataset category, with the results meticulously tabulated in Table I.

The evaluation results presented in Table I affirm that CNN-autoencoder based CBMIR model has achieved notable success in delivering suitable outcomes. The effectiveness of this model is demonstrated through its capability to enhance precision and recall values, signifying improvements in the accuracy and completeness of the image retrieval process. Overall, this evaluation adds empirical evidence to the merit of the proposed system in the context of CBMIR. Fig. 9 visually represents the performance of the CBMIR system proposed in this study.

TABLE I. PERFORMANCE EVALUATION

Retrieved Images (IR)	Accuracy (%)	Precision (%)	Recall (%)
20	98.34	97.84	98.42
40	97.56	97.15	97.15
60	96.11	95.48	94.87
80	94.35	95.66	94.24
100	93.00	94.00	93.00

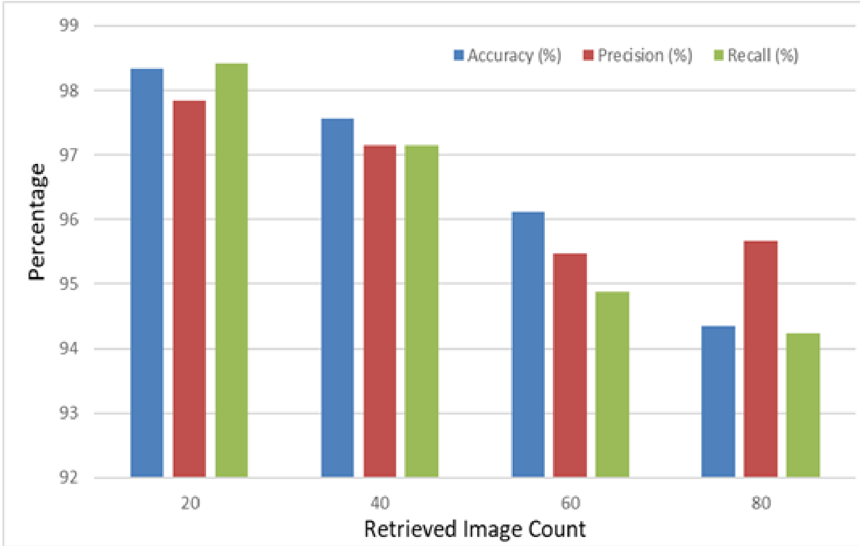

Fig. 9. Performance of proposed CBIMR model.

The evaluation outcomes reveal the effectiveness of the proposed approach, showcasing satisfactory performance with an average accuracy of 95.87%, precision of 96.03% and recall of 95.54%. Conversely, the results are comparatively optimal for all modalities of images in the dataset. Nevertheless, upon comparison with various parameters the performance gradually decreases while increasing the number of images to be retrieved.

To analyze the effectiveness of the constructed model, a thorough examination of their image retrieval performance is essential. The assessment of proposed CBMIR model's retrieval performance is conducted across multiple methodologies. The comparison of efficiency among existing models is presented in Table II, employing carefully chosen performance metrics,

TABLE II. PERFORMANCE COMPARISON WITH EXISTING METHODS

Methodology	Accuracy (%)	Precision (%)	Recall (%)
DCNN	88.42	87.48	88.62
AlexNet	87.65	87.51	87.75
VGG16	92.34	91.84	93.14
ResNet50	93.74	92.27	92.57
Proposed Method	95.87	96.03	95.54

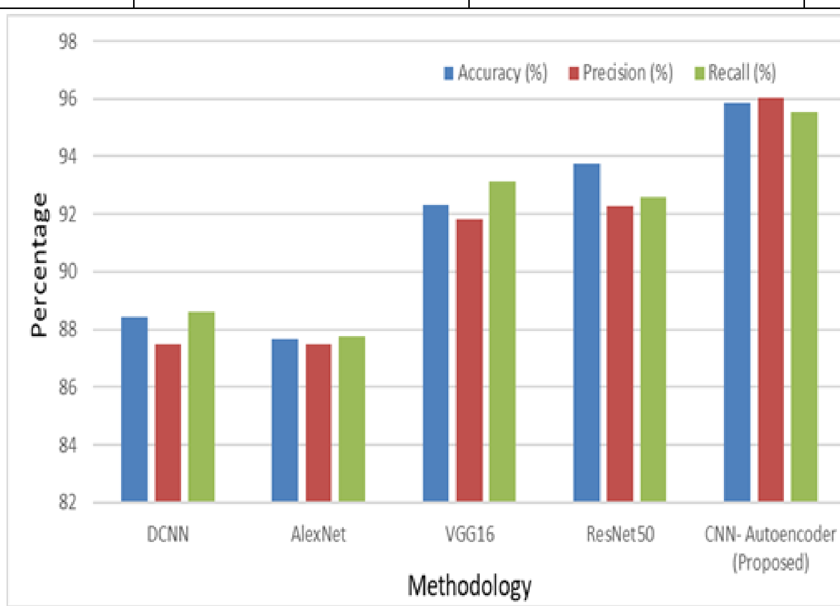

Fig. 10. Performance comparison.

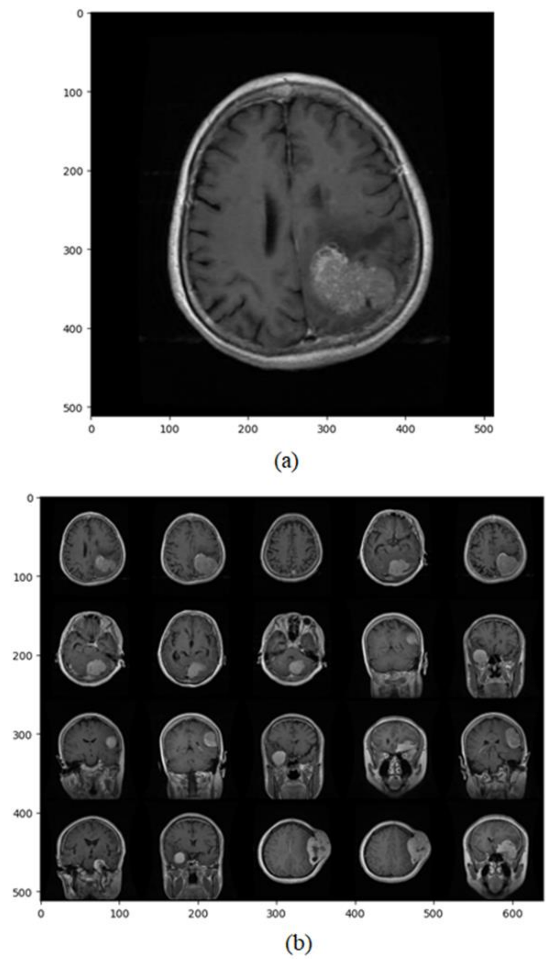

Fig. 11. Experimental results (a) Query image, (b) Retrieved images.

In the assessment of retrieval accuracy, the proposed CNN-autoencoder model distinguishes itself with an impressive score of 95.87%. Among the pre-trained models, VGG16 attains an accuracy rate of 92.34%, ResNet50 achieves 93.74%, and AlexNet displays an accuracy of 87.65%. Notably, the proposed model outperforms its nearest competitor, the ResNet50 model, by a notable margin, achieving a superior accuracy that is 2.13% higher. Moving to precision, the proposed CNN-autoencoder model excels with an impressive precision rate of 96.03%. In contrast, ResNet50 achieves 92.27% precision, VGG16 records 91.84%, and AlexNet obtains 87.51%. Fig. 10 provides a visual assessment of proposed model with state-of-the-art CBMIR models.

The precision of the proposed model surpasses that of the ResNet50 model by 3.76%, reinforcing its superiority. Proposed CNN-autoencoder achieves an exceptional recall value of 95.54%, outperforming all other models in this metric. In comparison, VGG16 achieves a recall rate of 93.14%, ResNet50 reached 92.57%, and AlexNet also records 87.75% in recall. Notably, the proposed model's recall surpasses the ResNet50 model by a significant margin of 2.97%, highlighting its superiority in retrieving relevant images. The proposed model demonstrates highest accuracy in medical image retrieval. For the visual evaluation of the proposed model the retrieval result obtained from the proposed model is illustrated in Fig. 11.

V. CONCLUSION

This research work introduced an innovative approach CBMIR specifically tailored for multimodal diagnosis images. Leveraging the CNN-based autoencoder method, the proposed system incorporates a learning process. This learning process is strategically designed to mitigate misinformation during the feature extraction phase, aiming to refine and improve upon the performance observed in previous works. This method is intended to overcome challenges associated with feature extraction and subsequently enhance the overall efficiency of CBMIR. By applying a learning mechanism within the autoencoder framework, the system adapts and refines its ability to accurately represent and extract meaningful features from multimodal diagnosis images. The results obtained from the evaluation of this method demonstrate notable average accuracy of 95.87%, precision of 96.03% and recall of 95.54%. This work contributes to the advancing field of CBMIR. The proposed system stands out for its ability to harness DL methodologies to address challenges in feature extraction, thereby achieving superior performance in comparison to existing methods.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to all those who contributed to the completion of this research paper. I extend my heartfelt thanks to my supervisor, my family, my colleagues and fellow researchers for their encouragement and understanding during the demanding phases of this work.

REFERENCES

[1] Qayyum, A., Anwar, S. M., Awais, M., & Majid, M. (2017). Medical image retrieval using deep convolutional neural network. Neurocomputing, 266, 8-20.

[2] Li, Z., Zhang, X., Müller, H., & Zhang, S. (2018). Large-scale retrieval for medical image analytics: A comprehensive review. Medical image analysis, 43, 66-84.

[3] Dubey, S. R., Singh, S. K., & Singh, R. K. (2016). Multichannel decoded local binary patterns for content-based image retrieval. IEEE transactions on image processing, 25(9), 4018-4032.

[4] Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical image analysis using convolutional neural networks: a review. Journal of medical systems, 42, 1-13.

[5] Ker, J., Wang, L., Rao, J., & Lim, T. (2017). Deep learning applications in medical image analysis. Ieee Access, 6, 9375-9389.

[6] Wang, G., Li, W., Zuluaga, M. A., Pratt, R., Patel, P. A., Aertsen, M., ... & Vercauteren, T. (2018). Interactive medical image segmentation using deep learning with image-specific fine tuning. IEEE transactions on medical imaging, 37(7), 1562-1573.

[7] Suzuki, K. (2017). Overview of deep learning in medical imaging. Radiological physics and technology, 10(3), 257-273.

[8] Quellec, G., Cazuguel, G., Cochener, B., & Lamard, M. (2017). Multiple-instance learning for medical image and video analysis. IEEE reviews in biomedical engineering, 10, 213-234.

[9] Xia, Z., Wang, X., Zhang, L., Qin, Z., Sun, X., & Ren, K. (2016). A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing. IEEE transactions on information forensics and security, 11(11), 2594-2608.

[10] Wei, X. S., Luo, J. H., Wu, J., & Zhou, Z. H. (2017). Selective convolutional descriptor aggregation for fine-grained image retrieval. IEEE transactions on image processing, 26(6), 2868-2881.

[11] Ozturk, S. (2020). Stacked auto-encoder based tagging with deep features for content-based medical image retrieval. Expert Systems with Applications, 161, 113693.

[12] Liu, X., Wang, M., Zha, Z. J., & Hong, R. (2019). Cross-modality feature learning via convolutional autoencoder. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 15(1s), 1-20.

[13] Cai, Yiheng, et al. "Medical image retrieval based on convolutional neural network and supervised hashing." IEEE access 7 (2019): 51877-51885.

[14] Li, S., Dai, W., Zheng, Z., Li, C., Zou, J., & Xiong, H. (2021). Reversible autoencoder: A CNN-based nonlinear lifting scheme for image reconstruction. IEEE Transactions on Signal Processing, 69, 3117-3131.

[15] Guan, A., Liu, L., Fu, X., & Liu, L. (2022). Precision medical image hash retrieval by interpretability and feature fusion. Computer Methods and Programs in Biomedicine, 222, 106945.

[16] Shen, Z., Ding, F., Yao, Y., Bhardwaj, A., Guo, Z., & Yu, K. (2022). A privacy-preserving social computing framework for health management using federated learning. IEEE Transactions on Computational Social Systems.

[17] Liu, Quande, et al. "Semi-supervised medical image classification with relation-driven self-ensembling model." IEEE transactions on medical imaging 39.11 (2020): 3429-3440.

[18] Bouchareb, Y., Khaniabadi, P. M., Al Kindi, F., Al Dhuhli, H., Shiri, I., Zaidi, H., & Rahmim, A. (2021). Artificial intelligence-driven assessment of radiological images for COVID-19. Computers in biology and medicine, 136, 104665.

[19] Swati, Z. N. K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., & Lu, J. (2019). Content-based brain tumor retrieval for MR images using transfer learning. IEEE Access, 7, 17809-17822.

[20] Jaiswal, A., Gianchandani, N., Singh, D., Kumar, V., & Kaur, M. (2021). Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning. Journal of Biomolecular Structure and Dynamics, 39(15), 5682-5689.

[21] Chen, M., Shi, X., Zhang, Y., Wu, D., & Guizani, M. (2017). Deep feature learning for medical image analysis with convolutional autoencoder neural network. IEEE Transactions on Big Data, 7(4), 750-758.

[22] Liu, Y., Wang, X., Wang, L., & Liu, D. (2019). A modified leaky ReLU scheme (MLRS) for topology optimization with multiple materials. Applied Mathematics and Computation, 352, 188-204.

[23] Li, D., Deng, L., Gupta, B. B., Wang, H., & Choi, C. (2019). A novel CNN based security guaranteed image watermarking generation scenario for smart city applications. Information Sciences, 479, 432-447.

[24] Maksoud, E. A. A., Barakat, S., & Elmogy, M. (2019). Medical images analysis based on multilabel classification. In Machine Learning in Bio-Signal Analysis and Diagnostic Imaging (pp. 209-245). Academic Press.

[25] Patel, S. P., & Upadhyay, S. H. (2020). Euclidean distance based feature ranking and subset selection for bearing fault diagnosis. Expert Systems with Applications, 154, 113400.

[26] Prasetyo, H., & Akardihas, B. A. P. (2019). Batik image retrieval using convolutional neural network. Telkomnika (Telecommunication Computing Electronics and Control), 17(6), 3010-3018.

[27] Alsmadi, M. K. (2020). Content-based image retrieval using color, shape and texture descriptors and features. Arabian Journal for Science and Engineering, 45(4), 3317-3330