INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RSEARCH ARTICLE

An Investigation of Distance Measures for Development of Effective Content Based Tumor Image Retrieval System

T Anitha[1], **T Kalaiselvi**[1]*

1 Department of Computer Science and Applications, The Gandhigram Rural Institute (Deemed to be University), Gandhigram, 624 302, Tamil Nadu, India

OPEN ACCESS

Received: 29-08-2023

Accepted: 05-03-2024

Published: 25-04-2024

Editor: ICDMMMDE-2023 Special Issue Editors: Dr. G. Mahadevan & Prof. Dr. P. Balasubramaniam

Citation: Anitha T, Kalaiselvi T (2024) An Investigation of Distance Measures for Development of Effective Content Based Tumor Image Retrieval System. Indian Journal of Science and Technology 17(SP1): 40-44. https://doi.org/10.17485/IJST/v17sp1.283

* **Corresponding author**.

kalaiselvi.gri@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Background/Objectives: The MRI has proven to be extremely effective in detecting tumors, with millions of images created each day throughout the world. To find similar images from a vast collection, Content-Based Tumor Image Retrieval (CBTIR) technology has been used to analysis the medical image. In the traditional retrieval methods, retrieving a similar image from the large database is crucial task. To overcome this issue we developed deep learning based retrieval method. **Methods:** This research offers a retrieval approach based on predefined ResNet models for quick and accurate image retrieval. We tested various prominent ResNet models with different distance similarity metrics, and the best option was determined by this work. **Findings:** After the various evaluation of ResNet models with varied distance measures on the CE-MRI data set, ResNet50 model applied with Hamming distance yields 99.33% of retrieval precision. **Novelty:** This work used predefined ResNet models with the combination of Distance similarity metrics to achieve more accurate results on medical image retrieval compared to the other conventional methods.

Keywords: Content Based Image Retrieval; Tumor Retrieval; Hamming Distance; Euclidean Distance; Minkowski

1 Introduction

Medical imaging technology generates a large volume of data in every single day, causing an interruption in the clinical progression. For this reason, scientists have demonstrated concentration in content-based image retrieval systems (CBIR), such as MRI, X-ray, and CT[(1)]. Radiologists may struggle to retrieve MRI images manually from a vast collection of images with similar structures or looks. This will be governed by the radiotherapist's availability and ability to review and retrieve tumor portion images. For large quantities of data manual retrieval procedure is inefficient, tough to replicate, and it take more time. To rectify this issue, automated CBIR could index archive photos with minimal involvement from radiologists. The motivation of this research is retrieving tumor portion images. The CBIR technology retrieves the specific tumor images based on the query image. For the new instance, the radiotherapist retrieved similar or the same images and previous diagnosis details from the repository.

The CBIR mechanism has a pair of additional crucial procedures: feature extraction and distance measurement. For the feature extraction process, several kinds of approaches were used, one is conventional machine-learning methods. It has two variants of features. The local feature [2] depends on image texture and intensity levels such as First and second-order statistics, gray-level co-occurrence matrix (GLMC), etc. Since a multitude of tumor images have identical characteristics but vary in certain characteristics like boundaries, uniformity, and size. In the other categories, the global features are extracted. They also extracted statistical features from images [3] like scale-invariant feature transformation [4].

Several strategies for retrieving images from a huge dataset or repository have been proposed by many researchers. Rao et al. [5] used a neural network (NN) to construct an image retrieval algorithm and the Canny steerable texture filter(CSTF) as a classifier. Image noise is reduced mostly through the Modified Kuan Filter and increasing the image contrast. Then, the features of the images were extracted using CSTF. They attained 99.81 % of an average precision value for each image evaluated. Sampathila et al. [6] stated an approach using the K-Nearest Neighbour (KNN)algorithm to retrieve the tumor images. They achieved 95.5 % of average accuracy for images of the OASIS publicly available dataset.

Madhu et al. [7] utilized a GLCM, DWT, and principal component analysis to develop a CBIR system for retrieving medical images. The test outcomes were derived from several sorts of medical imaging and were accurate to 99%. Deepak et al. [8] devised a method for retrieving tumor MR images by GoogleNet. The assessment results were achieved using Figshare datasets with an accuracy rate of 97.3 %. Swati et al. [9] developed a system for retrieving brain MR images utilizing a deep CNN, VGG19, to compute the resemblance between the query and source images. Using the CE-MRI image dataset, the test results reached 96 % as a precision value compared to previous approaches.

Sikandar et al. [10] devised a novel retrieval system that employs a transfer learning technique and incorporates ResNet50 and VGG16 pre-trained deep learning models, as well as one machine learning model, KNN. The test results were obtained with 100% precision. Rashad et al. [11] utilized query expansion to establish a retrieval approach for medical images. They employed a pre-trained AlexNet and VGG-19 model to fetch medical image features. They evaluated OASIS-MRI freely obtainable dataset, yielding an 84% precision level. Garg and Dhiman [12] devised a retrieval approach based on the GLCM and a local binary pattern to retrieve image information.. A particle swarm optimization-based feature selector is used specifically for decreasing the number of features employed during the classification process. The Coral dataset was used in the assessment process and their precision levels are 82% and 92%. We presented a CBTR methodology for retrieving tumor images from a massive repository utilizing ResNet models with Euclidean and Hamming distance measurement for similarity assessment.

The remaining portions of the paper's details are listed below. Section 2 dealt with the proposed work. Section 3 depicts the CBTIR system's performance and outcomes. Finally, Section 4 conveys the CBTIR system's conclusion.

2 Methodology

We introduced a CBTIR method for retrieving exact tumor images from the database. In this paper, we estimate the resemblance between the stored image and the query image using a transfer learning and Euclidean and hamming distance. ResNet is a pre-trained CNN that was used to fetch tumor portion characteristics from MRI tumor images. We utilize different varient of ResNet models for the retrieval process with the similarity measures. From that models ResNet50 predefines model produced better retrieval results than other ResNet models. ResNet50 is a ResNet model version contain a max and average pool layer and 48 convolution layers. ResNets were first used for image recognition before being expanded to image categorization, object recognition, and location. ResNet50's structure is depicted in Figure 1. We extracted brain tumor images from the datasets using ResNet50.

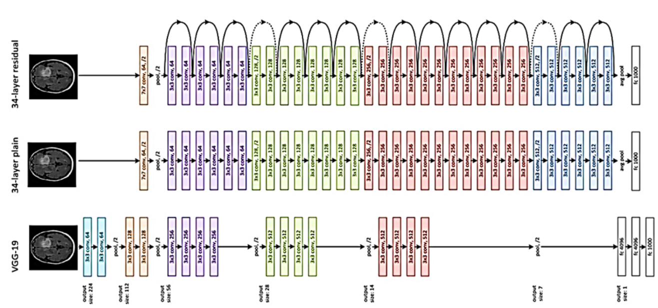

Fig 1. The architecture of the ResNet50

The proposed work is depicted in Figure 2. The purpose of the CBTIR system is to fetch images from the database that is comparable to the query image. At first, features of tumor images in the training database are extracted and stored. The retrieved features are then saved in a repository. Following that, fetch the query image's features from the test dataset. These two extracted features are compared to calculate the similarity measure using different techniques like Euclidean distance (ED), Manhattan distance, Minkowski Distance, and Hamming distance (HD). HD is the best distance-measure technique.

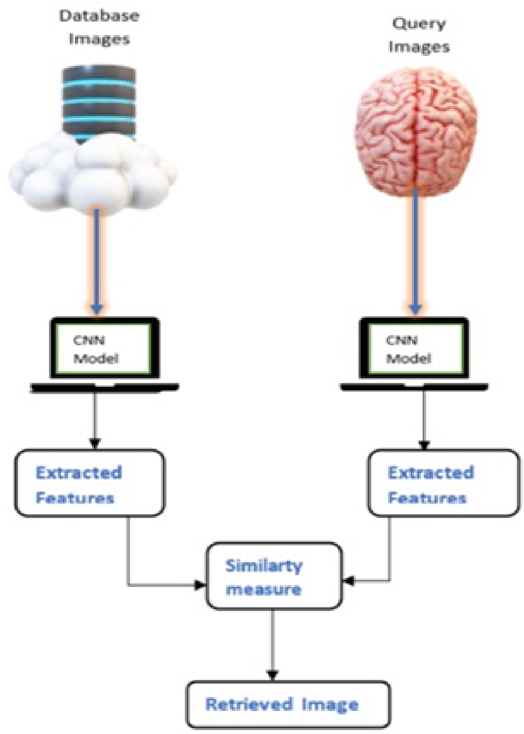

Fig 2. The flowchart of the proposed work

In the proposed investigation, we employ a transfer learning concept with all ResNet model versions and evaluate the results on the CE-MRI dataset [1]. We examine all ResNet models, including ResNet50, ResNet50v2, ResNet101, ResNet101v2, ResNet152, and ResNet152v2, using Euclidean (ED) and hamming distance (HD) metrics to determine which model produces the best results for the CBTIR system. This model extracts image features automatically for testing as well as training images. The similarity measure is particularly significant in this procedure since it will be used to get similar images from the repository [13].

3 Results and Discussion

To test the models, initially, we set a number of images to retrieve as 30. Row 1 of Figure 3 depicts the ResNet50 model's output. In this illustration, Figure 3 (a) stands for the query image, and Figure 3 (b) stands for the retrieved images comparable to the query image. The ResNet50 model retrieved 20 MRI brain tumor images with similar shapes from the 30 images.

The visual representation of the ResNet101 model's output is shown in Figure 3 row 2. In this diagram, Figure 3 (c) denotes the user image, and b indicates the tumor portion images that were retrieved. From the obtained 30 images that are comparable to the query image selected by the user the ResNet101 model retrieved 3 accurate tumor images and 27 MRI brain images with identical shapes.

The visual representation of the ResNet152 model's output is shown in Figure 3 row 4 column 1. The query image is denoted by Figure 3 (g) whereas the retrieved tumor images are represented by Figure 3 (h). The ResNet152 model obtained a large number of precisely matching images that have a comparable shape to the request image. Figure 3 row 4 column 2 and column 3 depicts a graphical representation of the ResNet101V2 and ResNet152V2 respectively. These models retrieved only a few similar images. As a result, we figured out that these models are unsuitable for the tumor retrieval processes.

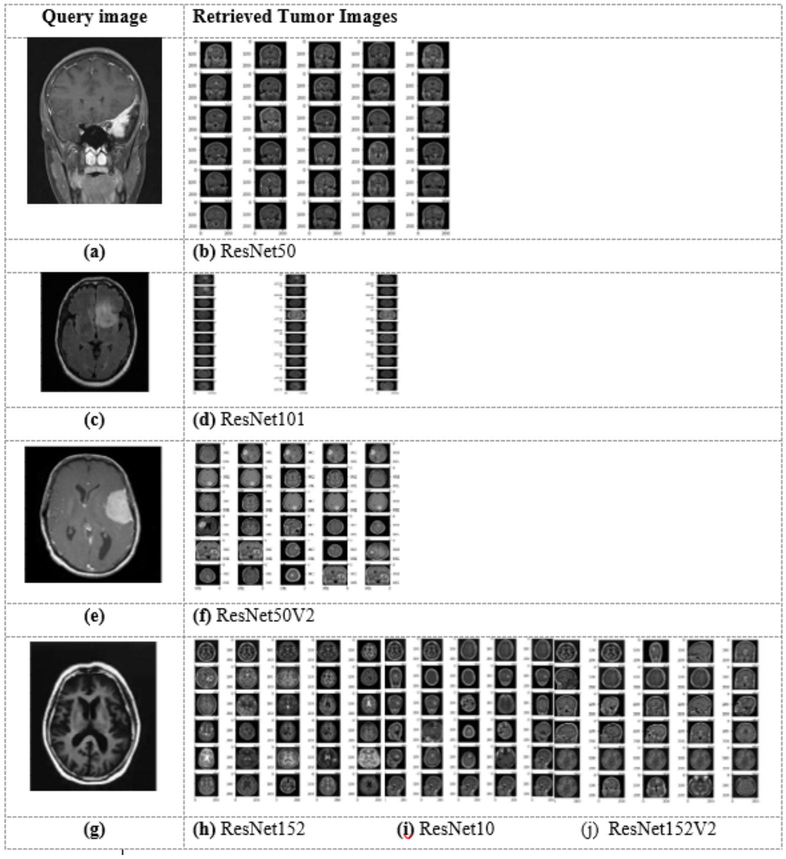

Fig 3. Visual representation of the proposedResNet models

Figure 3, row 3 gives the retrieval results of the ResNet50v2. This model finds 40 equivalent images. When compared to the remaining ResNet models, the ResNet50 and ResNet50V2 will produced better outcomes. The percentage of retrieval precision is more important than a quality validation. We calculate the retrieval precision value using the following equation [14]. The precision value for each ResNet model listed in Table 1 are calculated for the distance measure ED and HD respectively.

$$precision = \frac{No.of.releveant\ \ retrieved\ images}{No.of\ retrieved\ images}$$

Table 1. Retrieval Precision value

Models	Precision For HD (%)	Precision For ED (%)
ResNet50	99	90
ResNet101	853	72
ResNet152	62	70
ResNet50V2	89	85
ResNet101V2	54	49
ResNet152 V2	45	40

Table 1 proved that the ResNet50 generates higher precision of above 90% for ED and 99% for HD. Finally, a few existing methods result were compared with our proposed method and given in Table 2.

From Tables 1 and 2, we figured out that ResNet50 performs better than the other ResNet models for the CBTIR system with HD as an index for similarity measure calculation.

Table 2. Comparison of the precision value with existing methods.

Methods	Similarity metrics	% Precision
Garg and Dhiman [12]	ED	85
Monowar et al. [15]	ED	83
Abdullah [16]	HD	98
Proposed ResNet50	ED	**90**
Proposed ResNet50	HD	**99**

4 Conclusion

In this study, we implemented a CBTIR method to retrieve tumor images from a large database. Accessing similar or the same images from the wide database is a crucial process. To overcome this issues researchers produced a variety of CBIR methods using variety of measures. Deep learning methods have an excellent scope in all methodologies due to their rapid retrieval process. Based on the results of this investigation, we developed an automatic CBTIR system utilizing CNN and ResNet50 models with HD metrics. In the future, we provide a rapid method to retrieve exact tumor images with similar grades and diagnostic information.

Declaration

Presented in 9th INTERNATIONAL CONFERENCE ON DISCRETE MATHEMATICS AND MATHEMATICAL MODELLING IN DIGITAL ERA (ICDMMMDE-2023) during March 23-25, 2023, Organized by the Department of Mathematics, The Gandhigram Rural Institute (Deemed to be University), Gandhigram - 624302, Dindigul, Tamil Nadu, India. ICDMMMDE-23 was supported by GRI-DTBU, CSIR.

References

1) Swati ZNK, Zhao Q, Kabir M, Ali F, Ali Z, Ahmed S, et al. Content-Based Brain Tumor Retrieval for MR Images Using Transfer Learning. *IEEE Access*. 2019;7:17809–17822. Available from: https://dx.doi.org/10.1109/access.2019.2892455.
2) Hameed IM, Abdulhussain SH, Mahmmod BM. Content-based image retrieval: A review of recent trends. *Cogent Engineering*. 2021;8(1):1927469–1927469. Available from: https://dx.doi.org/10.1080/23311916.2021.1927469.
3) Latif A, Rasheed A, Sajid U, Ahmed J, Ali N, Ratyal NI, et al. Content-Based Image Retrieval and Feature Extraction: A Comprehensive Review. *Mathematical Problems in Engineering*. 2019;2019:1–21. Available from: https://dx.doi.org/10.1155/2019/9658350.
4) Wang Y, Liu F, Pang Z, Hassan A, Lu W. Privacy-preserving content-based image retrieval for mobile computing. *Journal of Information Security and Applications*. 2019;49:102399–102399. Available from: https://dx.doi.org/10.1016/j.jisa.2019.102399.
5) Rao RV, Prasad TJC. An efficient content-based medical image retrieval based on a new Canny steerable texture filter and Brownian motion weighted deep learning neural network. *The Visual Computer*. 2023;39(5):1797–1813. Available from: https://dx.doi.org/10.1007/s00371-022-02446-w.
6) Sampathila N, Pavithra, Martis RJ. Computational approach for content-based image retrieval of K-similar images from brain MR image database. 2022. Available from: https://doi.org/10.1111/exsy.12652.
7) Madhu, Kumar R. A hybrid feature extraction technique for content based medical image retrieval using segmentation and clustering techniques. *Multimedia Tools and Applications*. 2022;81(6):8871–8904. Available from: https://dx.doi.org/10.1007/s11042-022-11901-8.
8) Deepak S, Ameer PM. Retrieval of brain MRI with tumor using contrastive loss based similarity on GoogLeNet encodings. *Computers in Biology and Medicine*. 2020;125:103993–103993. Available from: https://dx.doi.org/10.1016/j.compbiomed.2020.103993.
9) Swati ZNK, Zhao Q, Kabir M, Ali F, Ali Z, Ahmed S, et al. Content-Based Brain Tumor Retrieval for MR Images Using Transfer Learning. *IEEE Access*. 2019;7:17809–17822. Available from: https://dx.doi.org/10.1109/access.2019.2892455.
10) Sikandar S, Mahum R, Alsalman A. A Novel Hybrid Approach for a Content-Based Image Retrieval Using Feature Fusion. *Applied Sciences*. 2023;13(7):4581–4581. Available from: https://dx.doi.org/10.3390/app13074581.
11) Rashad M, Afifi I, Abdelfatah M. RbQE: An Efficient Method for Content-Based Medical Image Retrieval Based on Query Expansion. *Journal of Digital Imaging*. 2023;36(3):1248–1261. Available from: https://dx.doi.org/10.1007/s10278-022-00769-7.
12) Garg M, Dhiman G. A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants. *Neural Computing and Applications*. 2021;33(4):1311–1328. Available from: https://dx.doi.org/10.1007/s00521-020-05017-z.
13) Patel B, Yadav K, Ghosh D. State-of-Art: Similarity Assessment for Content Based Image Retrieval System. *2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC)*. 2020;p. 1–6. Available from: https://doi.org/10.1109/iSSSC50941.2020.9358899.
14) Ayyachamy S, Khened AV, Krishnamurthi M. Medical image retrieval using ResNet-18. *Medical imaging 2019: imaging informatics for healthcare, research, and applications*. 2019;10954:233–241. Available from: https://doi.org/10.1117/12.2515588.
15) Monowar MM, Hamid MA, Ohi AQ, Alassafi MO, Mridha MF. AutoRet: A Self-Supervised Spatial Recurrent Network for Content-Based Image Retrieval. *Sensors*. 2022;22(6):2188–2188. Available from: https://dx.doi.org/10.3390/s22062188.
16) Abdullah SM, Jaber MM. Deep learning for content-based image retrieval in FHE algorithms. *Journal of Intelligent Systems*. 2023;32(1). Available from: https://dx.doi.org/10.1515/jisys-2022-0222.