Medical Image Retrieval using End-to-End Convolution Neural Network

Pradnya Maske, J. A. Kendule
SVERI's College of Engineering, Pandharpur, India
pmaske21@gmail.com

Abstract- Content-based medical image retrieval (CBMIR) is a challenging research area since long time. It aims at retrieving set of medical images having similar content to that of the input query medical image. In this paper, we propose, an effective feature learning approach to extract content-relevant features for medical image retrieval. The proposed approach is divided into two stages among which first stage encodes the input medical image into set of features followed by the decoder which reconstructs the input medical image from the encoded features. While, the second stage make use of the extracted features by encoder for index matching and retrieval task. This, two stage approach effectively improved the accuracy of the CBMIR system. We have utilized the existing benchmark datasets to validate the proposed approach for CBMIR. Extensive evaluation shows that the proposed approach outperforms the other existing methods for CBMIR.

***Keywords*—CBIR, Feature extraction, encoder-decoder**

1 Introduction

Image processing improves representational details in the image and makes it much more suitable towards autonomous interpretation by the computer. The human eye is confined to an electromagnetic spectrum visual band. In comparison, algorithms for image processing can process pictures taken from the entire electromagnetic spectrum. Because of advanced technologies, which are used for the processing of images, broad database repositories are developed and used in different tasks. At the same time, biomedical data are increasing increasingly as medical scanners such as MRI, CTs etc. can be easily used for patient diagnosis. Owing to the increase of medical data, smart health services should be provided both in urban and rural areas. It is a difficult job to efficiently search, access, retrieve required medical image and/or related medical history from a large-scale dataset. An automatic algorithm must therefore be established to address the limitations described above. Content-based medical image retrieval is an effective solution to search, access, and retrieve a necessary medical image from a large-scale medical dataset. In this article, we introduce a deep end-to-end network for retrieval of medical images with relevant content to the input query medical image. We addressed the current research in the field of image retrieval in the following section.

2 Literature Survey

Content-based image retrieval (CBIR) has large scientific, biomedical and industry applications. The latest methods are [1–4] for image retrieval and classification. Extraction of features is a very necessary and vital step in any natural and medical image retrieval system. The overlapping and non-overlapping area or pixel-based techniques extract local and global features. The efficient and effective characteristics are taken from spatial or transform domain data. The global feature extraction method, on the other hand, uses the entire image for feature extraction. The most likely methods for feature extraction in CBIR are local patterns performing the mathematical operation on the center pixel and its surrounding pixel of a particular image block.

Color is a prominent characteristic of any image. The explanation is that the human visual system is more responsive to the colors of the image. In the area of CBIR, Swain *et al.* [5] have proposed approach of color histogram to estimate the distance metric of the histogram intersection between images. The color histogram method has been implemented in [6]. In [7], the color correlogram based feature extraction method is suggested with the aid of color distribution and color interaction for the image and video retrieval. It is a difficult job to achieve high retrieval accuracy through the color characteristics in the presence of huge data. Person *et al.* [8] proposed a shape feature-based retrieval system using a Fourier descriptor. A shape based improved method based on a Fourier descriptor was proposed in [8]. However, the technology focused on geometric transformation is highly susceptible to noise and size. The improved Fourier descriptor proposed by [9] to surmount the limits of the existing Fourier descriptor.

Color and shape feature-based methods had their own limitations. The textural features are another essential component of an image and play a key role in extracting relevant and effective features from images. The key methods used to characterize an image textures in the image-processing field are in general, statistical, spectral and structural. For textural

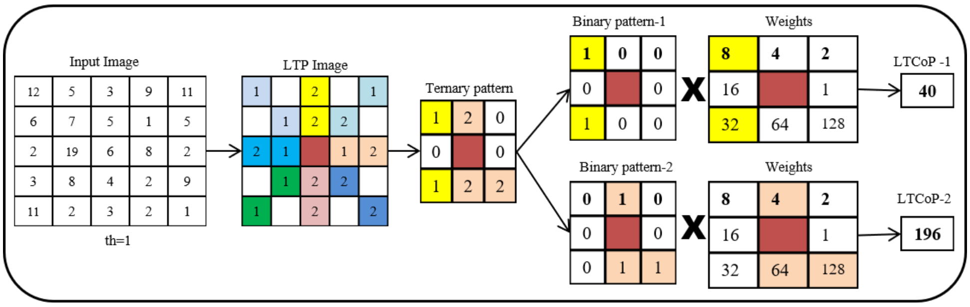

Figure 1: Procedure for LTCoP values calculation

feature extraction [10, 11] defines transform domain techniques. The grayscale ralationship within a pixel group is used to extract textual features known as Local Binary Pattern for image re-trieval [12].

Zhenhua *et al*. [13] have proposed rotational invariant LBP for feature extraction with application to texture classification. In [14], author proposed completed LBP (CLBP) for texture classification. In medical image analysis, [15, 16] explored the concept of LBP for feature extraction with application to medical image classification. LBP defines a first-order derivative only. In [17], the higher-order local derivative based feature extraction approach is proposed for image retrieval and face recognition. Further, Murala *et al*. [18]have proposed the pixel-level edge correlation based feature extraction for image indexing and retrieval task.

Local ternary pattern (LTP) [19] is the extension of the LBP as it considers three quantization level to encode the pixel intensity. Murala *et al.* [20–24] have proposed several local feature discriptor for effective feature extraction with application to CBIR. In [20], sub-block vocabulary based color histogram and spatial orientation trees are proposed for feature extraction, in [21] binary wavelet transform based feature extraction is proposed for medical as well as natural image retrieval, in [22] directional horizontal and vertical derivative based features are extracted for natural image retrieval. Further, the concept of 2×2 non-overlapping motif grids is proposed by Jhanwar *et al* [25] for CBIR. They cosidered motif occurance probability as image features for CBIR. Lin *et al.* [26] combined the color, texture and motif-based features for feature extraction with application to CBIR. In [27], authors encoded pixel intensity and its directional neighbourhood information to represent image into its abstract version for CBIR. Similarly, the directional co-occurrence based approach is proposed in [23] to extract effective directional information with the help of local ternary pattern for image retrieval. Further, Vipparthi *et al*. [28] used logical XoR operation with different motif representation used in [25] for CBIR named as directional local motif XoR pattern. The different researchers make use of different features like spherical local directional features [29], directional magnitude local triplet features [30], local energy oriented pattern [31], multi-dimensional multi-directional edge based feature extraction [32], etc for CBIR.

Most of the current methods used a hand-crafted feature-descriptor approach that has its own limitations. Since hand-crafted descriptors are based on certain hand-written rules and can not represent complex and reliable characteristics as the human brain does. To the other hand, convolution neural networks are highly inspired by the work of the human brain. CNN-based solution extracts more robust attributes and is more effective for nearly all computer vision applications like foreground-background segmentation [33–38], image enhancement [39–42] etc. Krizhevsky *et al* [43] designed a deep network for object recognition. Their deep network known as AlexNet is able to extract robust features and thus many researchers utilized it for feature extraction. In [44], authors proposed VGGNet for object recognition. Unlike AlexNet, VGGNet is deeper architecture having spatially smaller sized convolution filters. Various researchers followed these basic CNN architectures and utilized them for various computer vision applications.

In this paper, we propose a convolution neural network based approach for content-based medical image retrieval (CBMIR). Highlights of the proposed approach for CBMIR are listed as follows:

1. An end-to-end content-based medical image retrieval system is proposed.

2. An approach of feature extraction through encoder-decoder deep CNN is proposed.

3. Two benchmark datasets are used to validate the proposed approach for content-based medical image retrieval.

Rest of the manuscript is organised as, Section 1 introduces the general CBMIR system. Section 2 discusses the existing CBMIR approaches. The proposed approach for CBMIR is discussed in the Section 3. Section 4 and 5 discusses the training details and experimental analysis respectively. Finally, Section 6 concludes the proposed approach for CBMIR.

3 Existing Local Feature Descriptor for CBIR

Here, we have discussed about the existing baseline local feature descriptor approach used widely for feature extraction in CBIR.

3.1 Local Binary Patterns (*LBP*)

For texture classification, Ojala *et al*. [12] introduced the new local feature extraction technique known as Local Binary Pattern (LBP). LBP feature extraction is carried out with the help of gray scale relationship among group of pixels. The LBP values are derived by comparing center pixel value and its neighbours. Mathematically calculated using Eq. (1)-(2).

$$LBP_{p,q} = \sum_{n=1}^{p} 2^{(n-1)} \times f_1(K(I_n) - K(I_c)) \tag{1}$$

$$f_1(x) = \begin{cases} 1 & x \geq 0 \\ 0 & else \end{cases} \tag{2}$$

where, K is input image, I_c intensity at reference pixel, I_n is surrounding pixel values in the radius q, p is pixels to be considered having radius q for gray scale relationship $(q > 0,\ q \in N)$.

3.2 Local Ternary Co-occurrence Pattern (*LTCoP*)

Murala *et al*. [45] proposed Local Ternary Co-occurrence Pattern (LTCoP) for biomedical image retrieval. By using gray value of reference pixel and its surrounding neighbourhood pixels, the ternary edges present in image are calculated as shown in Figure 1. In LTCoP, the first order derivatives with references to center pixel is calculated using Eq. (3)-(4) as given below.

$$R_{p,q}(g_i) = R_{p,q}(g_i) - R_{p,q}(g_c) \tag{3}$$

$$R_{p,q+1}(g_i) = R_{p,q+1}(g_i) - R_{p,q}(g_c) \tag{4}$$

where, R is input image, g_c is reference pixel gray value, g_i represent the intensity value of all neighbourhood pixels and p is pixels to be considered having radius q for gray scale relationship $(q > 0,\ q \in N)$ After first order derivative calculation, the sign of derivatives are encoded as follows:

$$R^1_{p,q}(g_i) = f_1(R_{p,q}(g_i)) \tag{5}$$

$$R^1_{p,q+1}(g_i) = f_1(R_{p,q+1}(g_i)) \tag{6}$$

Eq. (5)-(6) gives the three value quantized image and these three valued images are used to to find the co-occurrences value between the elements as shown in Figure 1. Detail information of LTCoP is available in [45].

4 Proposed Approach

Here, we have discussed about the proposed approach for content-based medical image retrieval. In any computer vision application, feature extraction plays a vital role. To make an end-to-end CNN based approach for CBMIR, we make use of encoder-decoder type network architecture. The proposed approach for CBMIR is divided into two stages. In first stage, encoder encodes the input image into set of features followed by the decoder which reconstruct back the input image from extracted features. While, in second stage, we make use of the encoder features for index matching and retrieval task.

As discussed in Section 2, VGGNet [44] is a popular choice for the feature extraction. Here, we have utilized the trained weight parameters of the VGG16 to design the encoder of the proposed network. *i.e.* we initialized the weight parameters of the proposed encoder by trained weight parameters of the VGG16. The encoder network used in the proposed approach is shown in the Figure 2 (a). The basic building block of the encoder network is shown in Figure 2 (c) which consists of combination of convolution and rectified linear unit (ReLU layer). Followed by this basic architecture, obtained feature maps processed through the max-pooling layer. Similar to the traditional VGG16 network architecture, we have used smaller *i.e.* 3×3 spatial sized convolution filters in entire network. Obtained encoder feature maps are processed through the fully connected layers.

In decoder network, at each level, we concatenated the preceding layer feature maps with the respective encoder level feature map and up-sampled them by a factor of 2. Further, these up-sampled feature maps are then processed through the convolution layer. Similar to the encoder network, decoder network has basic building block as shown in Figure 2 (d) which consists of combination of convolution and ReLU layer. Combination of the encoder and decoder completes the proposed network for image reconstruction. The proposed encoder-decoder network for image reconstruction is shown in the Figure 2 (a). As shown in the Figure 2, input medical image is fed to the encoder to extracts set of features followed by the decoder network to reconstruct the medical image from the encoded features. It is evident from the Figure 2 that the proposed network produces near perfect medical image to that of input medical image from set of encoded features. It is therefore validate that the set of encoded features have prominent information about the input medical image. Thus,

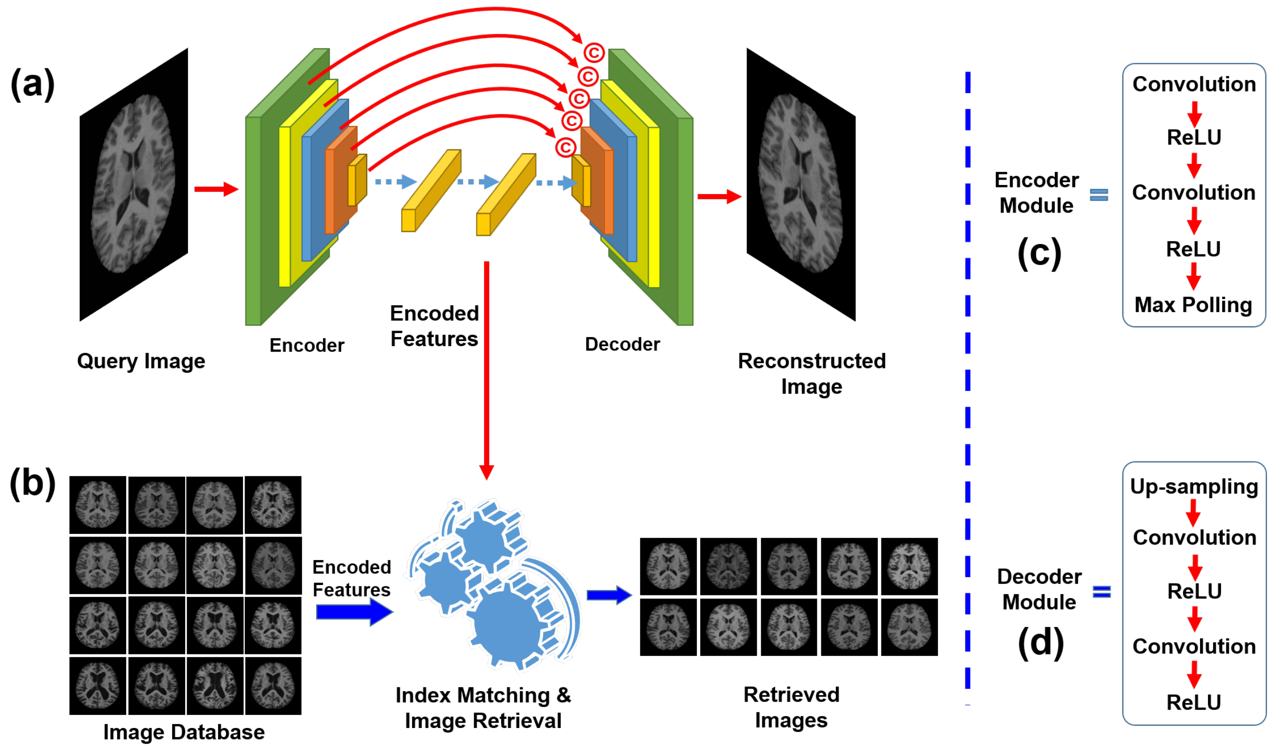

Figure 2: Proposed approach for Content-based Medical Image Retrieval. Note: *To understand the minor details, please maximize the figure.*

we make use of the encoded features to represent the input medical image and make use of them in index matching and retrieval task.

As shown in the Figure 2 (b) encoded features are taken out and utilized in the index matching and retrieval task. We follow the traditional way for index matching and retrieval task as discussed below.

4.1 Index Matching and Retrieval Task

Let us consider a database (DB) having N number of images. Each image from the database is fed to the trained encoder-decoder network for feature extraction. f_{DB_p} represents the set of features extracted from the p^{th} image from the database (DB) by the encoder network. Now, let us consider a query medical image Q. As per the procedure, query image is fed to the encoder-decoder network and set of features (f_Q) extracted by the encoder network are considered as query image features.

To select the n top ranked images similar to the given query image, similarity between f_{DB_p} and f_Q is calculated. We make use of $d1$ distance measure to compute the similarity between query image and database images. Mathematical formulation of the $d1$ distance measure used in this work is given by Eq. (7)

$$D(Q, DB_p) = \left| \frac{f_{DB_p} - f_Q}{1 + f_{DB_p} + f_Q} \right| \tag{7}$$

where, Q is the query image, DB_p is database image, f_{DB_p} is feature vector of p^{th} image from the database, f_Q is feature vector of query image.

Figure 2 (b) shows the process of index matching and retrieval. Sample retrieved medical images for a given query medical image using the proposed approach are shown in Figure 2 (b).

5 Training Details

The proposed encoder-decoder network is trained for image reconstruction. As discussed earlier in Section 4, we have considered the trained weight parameters of the VGG16 as an initial weight parameters for the encoder network. While, weight parameters of the decoder network are initialized randomly. With this, we have trained the the encoder-decoder network on BraTS-2015 database for the image reconstruction. In total, 14014 brain MRI slices from the BraTS-2015 database are considered in training set. As, the proposed encoder-decoder network is trained for reconstruction of the input medical image from its encoded features, input and ground truth for the encoder-decoder network is same. Weight

Table 1: Retrieval accuracy comparison of proposed method and other existing state-of-the-art methods on OASIS-MRI database in terms of ARP with top 10 matches considered. Note: *PM - Proposed Method*

Method	1	2	3	4	5	6	7	8	9	10
LBP [12]	100	72.92	61.20	56.89	53.25	50.04	48.63	47.21	45.87	44.66
LTP [46]	100	73.63	63.42	57.84	53.87	51.46	49.37	47.77	46.45	45.04
SS3D [47]	100	68.53	56.77	51.54	47.89	45.68	43.77	42.37	41.44	40.45
LTCoP [45]	100	73.87	62.79	57.66	54.49	51.58	49.37	47.57	46.93	45.82
LTrP [22]	100	70.67	59.62	52.55	48.69	47.43	45.81	44.30	43.34	42.52
MDMEP [32]	100	81.47	73.87	70.19	67.84	66.71	65.52	63.9	63.37	62.49
AlexNet [43]	100	80.88	73.56	68.41	65.46	62.98	61.83	60.54	59.46	58.36
ResNet [48]	100	78.50	71.18	67.99	64.94	62.87	60.54	59.26	58.38	57.48
PM	**100**	**92.38**	**90.08**	**89.13**	**88.50**	**87.75**	**87.11**	**86.81**	**86.64**	**86.18**

Table 2: Group-wise retrieval accuracy comparison in terms of ARP on OASIS-MRI database.

Method	Group 1	Group 2	Group 3	Group 4	Average
LBP [12]	55.08	35.20	32.70	51.60	43.64
LTP [46]	52.90	37.06	35.73	51.32	44.25
SS3D [47]	44.19	39.02	35.73	41.42	40.08
LTCoP [45]	50.08	41.08	34.16	55.19	45.12
LTrP [22]	52.26	37.35	34.04	43.21	41.75
MDMEP [32]	69.52	50.59	48.31	77.64	62.49
AlexNet [43]	68.87	43.73	41.01	74.72	57.08
ResNet [48]	69.11	45.59	44.27	66.42	56.35
Proposed Method	**88.30**	**80.50**	**82.30**	**93.80**	**86.22**

parameters of the proposed encoder-decoder network for image reconstruction are updated on Matlab-18 platform with learning rate=0.00001.

6 Experimental Analysis

In this Section, we have discussed about the experimental analysis *i.e.* validation of the proposed approach for CBMIR. We have considered two benchmark databases namely CT ROI and OASIS for the analysis. The performance of the proposed and existing approaches on these databases for CBMIR is compared. The parameters Precision (P), Recall (R), Average Retrieval Precision (ARP) and Average Retrieval Rate (ARR) are used to measure the retrieval accuracy obtained using proposed and existing approaches. Mathematical formulation of evaluation parameters is given by Eq. (8-11).

Table 3: Retrieval accuracy comparison of proposed method and other existing state-of-the-art feature descriptor in terms of ARP with top 10 matches considered on VIA/I-ELCAP database. Note: *PM - Proposed Method*

Method	1	2	3	4	5	6	7	8	9	10
LBP [12]	100	88.95	84.27	81.65	78.98	77.00	75.70	74.36	73.03	72.01
LTP [46]	100	91.90	88.93	85.98	84.08	82.28	81.04	79.81	78.57	77.46
SS3D [47]	100	74.25	64.27	58.70	54.78	52.55	50.47	48.68	47.44	46.23
LTCoP [45]	100	93.80	91.10	88.75	86.80	85.35	84.13	82.60	81.31	80.28
LTrP [22]	100	81.60	74.73	70.55	67.70	65.32	63.49	61.96	60.88	59.62
MDMEP [32]	100	81.60	74.73	70.55	67.70	65.32	63.49	61.96	60.88	59.62
AlexNet [43]	100	99.60	99.13	98.63	98.06	97.43	96.56	95.85	94.89	93.88
ResNet [48]	100	98.35	96.60	94.88	93.38	91.98	90.27	88.68	87.22	85.78
PM	**100**	**100**	**99.97**	**99.95**	**99.94**	**99.91**	**99.89**	**99.83**	**99.78**	**99.74**

$$Precision : P = \frac{N_R \cap N_{RT}}{n_{RT}} \tag{8}$$

Table 4: Retrieval accuracy comparison of proposed method and other existing state-of-the-art feature descriptor in terms of ARR with top 10 matches considered on VIA/I-ELCAP database. Note: *PM - Proposed Method*

Method	1	2	3	4	5	6	7	8	9	10
LBP [12]	10	17.79	25.28	32.66	39.49	46.20	52.99	59.49	65.73	72.01
LTP [46]	10	18.38	26.68	34.39	42.04	49.37	56.73	63.85	70.71	77.46
SS3D [47]	10	14.85	19.28	23.48	27.39	31.53	35.33	38.94	42.70	46.24
LTCoP [45]	10	18.76	27.33	35.50	43.40	51.21	58.89	66.08	73.18	80.28
LTrP [22]	10	16.32	22.42	28.22	33.85	39.19	44.44	49.57	54.79	59.62
MDMEP [32]	10	16.32	22.42	28.22	33.85	39.19	44.44	49.57	54.79	59.62
AlexNet [43]	10	19.92	29.74	39.45	49.03	58.46	67.59	76.68	85.40	93.88
ResNet [48]	10	19.67	28.98	37.95	46.69	55.19	63.19	70.94	78.50	85.78
PM	**10**	**20**	**29.99**	**39.98**	**49.97**	**59.95**	**69.87**	**79.83**	**89.81**	**99.74**

$$ARP = \frac{1}{DB}\sum_{n=1}^{DB} P(I_n)\bigg|_{n\leq 10} \tag{9}$$

$$Recall: R = \frac{N_R \cap N_{RT}}{n_R} \tag{10}$$

$$ARR = \frac{1}{DB}\sum_{n=1}^{DB} R(I_n)\bigg|_{n\geq 10} \tag{11}$$

where, N_R is total number of relevant images present in the database, N_{RT} is total number of retrieved image similar to query image from database, $N_R \cap N_{RT}$ gives retrieved images which are from query image category, n_R is total relevant images present in the database for query image, n_{RT} is total images retrieved using similarity measure, I_n is n^{th} query image and total number of images in database is denoted by DB.

6.1 Retrieval Accuracy on OASIS-MRI Database

This experiment is evaluated on OASIS-MRI database, which is publicly available for medical image retrieval. This database consists of 421 MRI scans captured between 18 and 96 years aged patients. Using different medical terms and easy experimental purpose, these images are divided into four groups based on shape of ventricular in the images. Group 1 to 4 has 124, 102, 89, 106 images respectively. We have used entire database for testing and there is no overlap with training set MRI slices. Table 1 summarizes the top 10 images retrieval accuracy of the proposed method and existing approaches on OASIS-MRI database in terms of ARP. Also, the group-wise ARP comparison of the proposed approach with other existing state-of-the-art methods on OASIS-MRI dataset is given Table 2. From Table 1 and Table 2, it is clear that the proposed method outperforms the other existing deep learning and non-deep learning approaches by a large margin.

6.2 Retrieval Accuracy on VIA/I-ELCAP CT Database

In this experiment, VIA/I-ELCAP dataset is utilized to evaluate the performance of proposed approach for CBMIR. VIA/I-ELCAP is one of the most preferred database for CBMIR technique. It consists of 1000 images divided into 10 categories with 100 images per category. The top 10 image retrieval accuracy of the proposed and existing methods is given in Table 3 and Table 4 in terms of ARP and ARR respectively. Existing deep learning based approach is able to achieve 93% accuracy. Proposed approach achieved 99% accuracy which shows the robustness of the proposed approach for CBMIR. We give this credit to the proposed feature learning approach.

7 CONCLUSION

In this paper, we have proposed a new approach for content-based medical image retrieval. The proposed approach comprise two stages. Among which stage 1 processes input medical image through the trained encoder-decoder network which encodes the input medical image into set of features followed by the its reconstruction using the decoder network. The near perfect reconstruction of the input medical image from the set of encoded features witnessed a robustness of the proposed encoder network to extract robust features. In second stage, we have considered set of encoded features are compared with the database features vectors for index matching and retrieval task. the performance of the proposed approach is tested on existing two benchmark databases namely VIA/I-Elcap CT database, OASIS-MRI and compared against the existing methods for CBMIR. The performance comparison show that the proposed approach outperforms the other existing methods for content-based medical image retrieval.

References

[1] L. Zheng, A. W. Wetzel, J. Gilbertson, and M. J. Becich, "Design and analysis of a content-based pathology image retrieval system," *IEEE transactions on information technology in biomedicine*, vol. 7, no. 4, pp. 249–255, 2003.

[2] G. Scott and C.-R. Shyu, "Knowledge-driven multidimensional indexing structure for biomedical media database retrieval," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 3, pp. 320–331, 2007.

[3] X. Xu, D.-J. Lee, S. Antani, and L. R. Long, "A spine x-ray image retrieval system using partial shape matching," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 100–108, 2008.

[4] B. André, T. Vercauteren, A. M. Buchner, M. B. Wallace, and N. Ayache, "Learning semantic and visual similarity for endomicroscopy video retrieval," *IEEE Transactions on Medical Imaging*, vol. 31, no. 6, pp. 1276–1288, 2012.

[5] M. J. Swain and D. H. Ballard, "Color indexing," *International journal of computer vision*, vol. 7, no. 1, pp. 11–32, 1991.

[6] ——, "Indexing via color histograms," in *Active Perception and Robot Vision*. Springer, 1992, pp. 261–273.

[7] M. A. Stricker and M. Orengo, "Similarity of color images," in *Storage and Retrieval for Image and Video Databases III*, vol. 2420. International Society for Optics and Photonics, 1995, pp. 381–393.

[8] E. Persoon and K.-S. Fu, "Shape discrimination using fourier descriptors," *IEEE Transactions on systems, man, and cybernetics*, vol. 7, no. 3, pp. 170–179, 1977.

[9] Y. Rui, A. C. She, and T. S. Huang, "Modified fourier descriptors for shape representation-a practical approach," in *Proc of First International Workshop on Image Databases and Multi Media Search*. Citeseer, 1996, pp. 22–23.

[10] M. Kokare, P. K. Biswas, and B. N. Chatterji, "Texture image retrieval using new rotated complex wavelet filters," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 35, no. 6, pp. 1168–1178, 2005.

[11] M. Kokare, P. Biswas, and B. Chatterji, "Rotation-invariant texture image retrieval using rotated complex wavelet filters," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 36, no. 6, pp. 1273–1282, 2006.

[12] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[13] Z. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using lbp variance (lbpv) with global matching," *Pattern recognition*, vol. 43, no. 3, pp. 706–719, 2010.

[14] ——, "A completed modeling of local binary pattern operator for texture classification," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657–1663, 2010.

[15] M. M. Pawar, S. N. Talbar, and A. Dudhane, "Local binary patterns descriptor based on sparse curvelet coefficients for false-positive reduction in mammograms," *Journal of healthcare engineering*, vol. 2018, 2018.

[16] A. A. Dudhane and S. N. Talbar, "Multi-scale directional mask pattern for medical image classification and retrieval," in *Proceedings of 2nd International Conference on Computer Vision & Image Processing*. Springer, 2018, pp. 345–357.

[17] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor," *IEEE transactions on image processing*, vol. 19, no. 2, pp. 533–544, 2010.

[18] S. Murala, R. Maheshwari, and R. Balasubramanian, "Local maximum edge binary patterns: a new descriptor for image retrieval and object tracking," *Signal Processing*, vol. 92, no. 6, pp. 1467–1479, 2012.

[19] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE transactions on image processing*, vol. 19, no. 6, pp. 1635–1650, 2010.

[20] S. Murala, R. Maheshwari, and R. Balasubramanian, "Expert system design using wavelet and color vocabulary trees for image retrieval," *Expert Systems with Applications*, vol. 39, no. 5, pp. 5104–5114, 2012.

[21] ——, "Directional binary wavelet patterns for biomedical image indexing and retrieval," *Journal of Medical Systems*, vol. 36, no. 5, pp. 2865–2879, 2012.

[22] ——, "Local tetra patterns: a new feature descriptor for content-based image retrieval," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2874–2886, 2012.

[23] S. Murala and Q. J. Wu, "Local ternary co-occurrence patterns: a new feature descriptor for mri and ct image retrieval," *Neurocomputing*, vol. 119, pp. 399–412, 2013.

[24] S. Murala, Q. J. Wu, R. Maheshwari, and R. Balasubramanian, "Modified color motif co-occurrence matrix for image indexing and retrieval," *Computers & Electrical Engineering*, vol. 39, no. 3, pp. 762–774, 2013.

[25] N. Jhanwar, S. Chaudhuri, G. Seetharaman, and B. Zavidovique, "Content based image retrieval using motif cooccurrence matrix," *Image and Vision Computing*, vol. 22, no. 14, pp. 1211–1220, 2004.

[26] C.-H. Lin, R.-T. Chen, and Y.-K. Chan, "A smart content-based image retrieval system based on color and texture feature," *Image and Vision Computing*, vol. 27, no. 6, pp. 658–665, 2009.

[27] S. K. Vipparthi and S. Nagar, "Directional local ternary patterns for multimedia image indexing and retrieval," *International Journal of Signal and Imaging Systems Engineering*, vol. 8, no. 3, pp. 137–145, 2015.

[28] ——, "Expert image retrieval system using directional local motif xor patterns," *Expert Systems with Applications*, vol. 41, no. 17, pp. 8016–8026, 2014.

[29] A. B. Gonde, P. W. Patil, G. M. Galshetwar, and L. M. Waghmare, "Volumetric local directional triplet patterns for biomedical image retrieval," in *2017 Fourth International Conference on Image Information Processing (ICIIP)*. IEEE, 2017, pp. 1–6.

[30] G. M. Galshetwar, P. W. Patil, A. B. Gonde, L. M. Waghmare, and R. Maheshwari, "Local directional gradient based feature learning for image retrieval," in *2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS)*. IEEE, 2018, pp. 113–118.

[31] G. Galshetwar, L. Waghmare, A. Gonde, and S. Murala, "Local energy oriented pattern for image indexing and retrieval," *Journal of Visual Communication and Image Representation*, vol. 64, p. 102615, 2019.

[32] G. Galshetwar, L. M. Waghmare, A. B. Gonde, and S. Murala, "Multi-dimensional multi-directional mask maximum edge pattern for bio-medical image retrieval," *International Journal of Multimedia Information Retrieval*, vol. 7, no. 4, pp. 231–239, 2018.

[33] P. Patil and S. Murala, "Fggan: A cascaded unpaired learning for background estimation and foreground segmentation," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2019, pp. 1770–1778.

[34] P. W. Patil, O. Thawakar, A. Dudhane, and S. Murala, "Motion saliency based generative adversarial network for underwater moving object segmentation," in *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019, pp. 1565–1569.

[35] P. W. Patil, A. Dudhane, S. Murala, and A. B. Gonde, "A novel saliency-based cascaded approach for moving object segmentation," in *International Conference on Computer Vision and Image Processing*. Springer, 2019, pp. 311–322.

[36] P. W. Patil, K. M. Biradar, A. Dudhane, and S. Murala, "An end-to-end edge aggregation network for moving object segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 8149–8158.

[37] P. W. Patil and S. Murala, "Msfgnet: A novel compact end-to-end deep network for moving object detection," *IEEE Transactions on Intelligent Transportation Systems*, 2018.

[38] P. Patil, S. Murala, A. Dhall, and S. Chaudhary, "Msednet: multi-scale deep saliency learning for moving object detection," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2018, pp. 1670–1675.

[39] A. Dudhane and S. Murala, "C2msnet: A novel approach for single image haze removal," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018.

[40] A. Dudhane, H. Singh Aulakh, and S. Murala, "Ri-gan: An end-to-end network for single image haze removal," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0–0.

[41] A. Dudhane and S. Murala, "Ryf-net: Deep fusion network for single image haze removal," *IEEE Transactions on Image Processing*, vol. 29, pp. 628–640, 2019.

[42] A. Dudhane, P. Hambarde, P. Patil, and S. Murala, "Deep underwater image restoration and beyond," *IEEE Signal Processing Letters*, vol. 27, pp. 675–679, 2020.

[43] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[44] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[45] S. Murala and Q. J. Wu, "Local ternary co-occurrence patterns: a new feature descriptor for mri and ct image retrieval," *Neurocomputing*, vol. 119, pp. 399–412, 2013.

[46] J. Ren, X. Jiang, and J. Yuan, "Relaxed local ternary pattern for face recognition," in *2013 IEEE international conference on image processing*. IEEE, 2013, pp. 3680–3684.

[47] S. Murala and Q. J. Wu, "Spherical symmetric 3d local ternary patterns for natural, texture and biomedical image indexing and retrieval," *Neurocomputing*, vol. 149, pp. 1502–1514, 2015.

[48] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.