

# Image reconstruction using a refined Res-UNet model for medical image retrieval

Pinapatruni Rohini<sup>1</sup>, C. Shoba Bindu<sup>1</sup>

<sup>1</sup> Department of Computer Science and Engineering, JNTUA College of Engineering, Anantapuramu, Andhra Pradesh, 515002, India

## ABSTRACT

Measurement of medical data and extraction of medical images with various sensory systems have become a crucial part of diagnosing and treating various diseases. In this contest measurement technology plays a key role in diagnosis. Medical experts usually use previous case studies to identify and deal with the current medical condition. In this context, an expert required to explore the large medical database to search relevant images for the required analysis. Searching such large database and retrieving an image efficiently becomes very tedious task. Therefore, this paper proposed a measurement-based image reconstruction using the refined Res-UNet Framework for Medical Image Retrieval (MIR) for managing such crucial tasks and reduce complexities. The proposed two stage framework consists of an image reconstruction using Res-UNet and index similarity matching of query image with history images. Res-UNet is a vanilla combination of ResNet50 as an encoder that gives latent information of input image, and the decoder from UNet reconstructs the image using latent information. Further, those latent features match similar image features and retrieve indexed images from the medical image database. The efficacy of proposed method was confirmed on benchmark medical image databases such as ILD and VIA/ELCAP-CT for MIR. The proposed framework outperforms the existing methods in the task of MIR.

**Section:** RESEARCH PAPER

**Keywords:** Medical image retrieval; Res-UNet Framework; image reconstruction; index matching

**Citation:** Pinapatruni Rohini, C. Shoba Bindu, Image reconstruction using a refined Res-UNet model for medical image retrieval, Acta IMEKO, vol. 11, no. 1, article 32, March 2022, identifier: IMEKO-ACTA-11 (2022)-01-32

**Section Editor:** Md Zia Ur Rahman, Koneru Lakshmaiah Education Foundation, Guntur, India

**Received** December 25, 2021; **In final form** February 17, 2022; **Published** March 2022

**Copyright:** This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Corresponding author:** Pinapatruni Rohini, e-mail: [rohini11@gmail.com](mailto:rohini11@gmail.com)

## 1. INTRODUCTION

The proliferation of digital image measurement devices in medical field induces image-based diagnosis as a profound approach to acquire relevant treatment information for healthcare applications. The measured images are very susceptible and diverge demanding safe and efficient database storage. However, due to rapidly growing medical database, efficiently accessing the information from the database is utmost important for necessary analysis and processing of patient's records.

In this context, a structured model is required to handle this huge database with effective image information retrieval. In this context, Content-based medical image retrieval (CBMIR) measurement method proved to be an effectual method to access the required medical image from the database efficiently. CBMIR follows two steps: i) It performs feature extraction utilizing hand crafted techniques or learning based models, ii) Performs the index matching task for image retrieval. Existing works utilized

hand-crafted techniques for the retrieval of the images from the database resembling the query medical image. But the method is not effective when processing with a huge database.

In past years, the image information like shape, colour, texture, are used as local feature descriptor to retrieve the image is proposed in [1]-[5]. In [6], [7] author extracted the features called Local Binary Pattern (LBP) using the Gray-scale information between the group of pixels for texture classification. In [8], the local ternary patterns descriptor is proposed which is an added version of LBP descriptor, used for face recognition. The descriptor exploits three quantization levels improving the pixels encoding mechanism. Murala et al. [9] proposed spherical and local direction-based feature descriptors for medical and natural image retrieval. A directional edge local ternary pattern based on reference pixel is proposed by Vipparthi et al. [10]. In [11], a three-stage local feature descriptor is proposed with multi- dimension and multi-direction for image retrieval. Moreover, directional mask is utilized to extract the directional edges. Lastly, the image is retrieved using the maximum directional edge-based features. In further work, a

reference block gradient directions-based feature descriptor is proposed for image retrieval [12]. A wavelet-based 3D circular difference pattern is proposed for medical image retrieval. Threshold-based LBP and local adjacent neighbourhood average difference information is combined for medical image retrieval in [13].

Previous works utilized hand crafted feature descriptors for feature extraction in CBMIR modelling. However, the hand-crafted feature descriptors own to the limitation of relying on hand-written assumptions [14], which limits the efficiency of the overall model. These features become less effective in encoding the robust and complex features. Currently, learning based feature descriptors are resolving the above limitation by utilizing the power of learning complex features based on their different model architectures. Convolution neural networks (CNN) shows remarkable results in the defined area because of the high learning ability of these network. Many works are proposed utilizing the CNN network architecture for diverse applications like objection detection, image classification, anomaly detection etc. [15]-[18]. The feature learning capability is also utilized by researchers in medical domain for medical image analysis, medical image segmentation, image de-hazing etc., [15], [19]-[22]. However, even after the remarkable performance on small network, the CNN based architectures perform badly when the depth of the network increases giving rise to the vanishing gradient issues. To resolve the above limitation, a residual learning (ResNet) based architecture is proposed in [23] for image classification. Motivated from this, an image retrieval approach is proposed based on image-to-image reconstruction for image retrieval.

GANs are utilized with Image-to-image translation approach provides outperforming results. This inspired the researchers to utilize GAN network for various diverse computer vision applications like image enhancement [15], moving object segmentation [16], depth estimation [24], etc. GAN is composed of two component generator and discriminator. Adversarial

training includes the training of both generator and discriminator. In [25], an inception architecture is performed which significantly improved the performance with relatively low computational cost.

Motivated from the significant performance of adversarial learning and inception architecture in diverse fields, this paper proposed refined network architecture utilizing residual network for image reconstruction based on measurement technology. Our network removes the skip connection which makes the reconstructed image more robust in addition to the high-quality reconstruction. The paper defines the following as major contribution:

1. An end-to-end refined content-based image reconstruction learning framework is proposed for medical image retrieval.
2. Proposed a novel Res-UNet framework for reconstruction of input medical image exploits novel pretrained ResNet50 model with CNN up-sampler.
3. Robust features generated using ResNet encoder for index matching and retrieval of image from the database.
4. This paper utilized two benchmark datasets to quantify proposed Res-UNet in CBMIR.

## 2. PROPOSED WORK

A novel content-based medical image retrieval framework named refined Res-UNet Model is proposed for effective and efficient image retrieval as shown in Figure 1. The framework incorporates two functioning blocks: Reconstruction Block and Indexing Block. The reconstruction block is inspired by the standard Pix2Pix [26] architecture where the UNet architecture comprises an encoder and decoder. Proposed framework includes ResNet50 [23] network as an encoder with five blocks. Firstly, the input image is processed to the ResNet50 block of the novel proposed model, which performs the encoding to

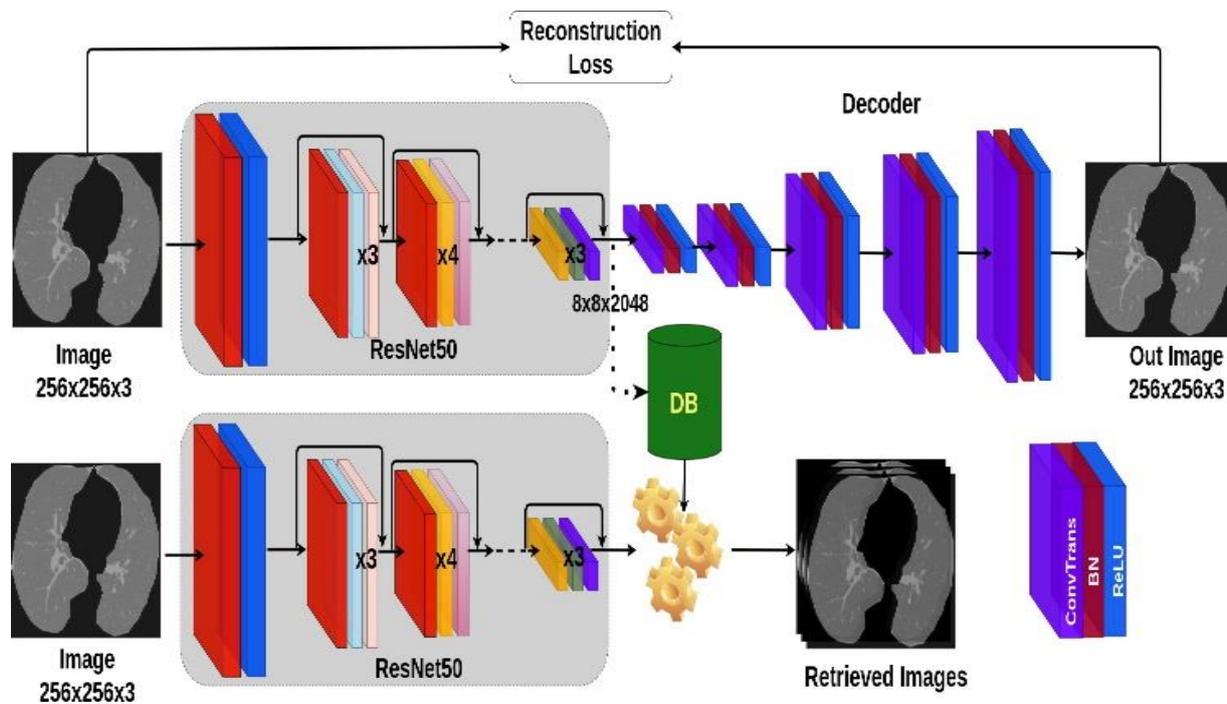


Figure 1. Refined Res-UNet Framework for CBMIR.

Stage 1: Reconstruct image using Res-UNet.

Stage 2: Encoder ResNet50 feature extraction for index matching and image retrieval.

generate the set of features taking advantage of self-attention process. Further, the features are decoded to reconstruct the original input medical image. The accurate and robust reconstruction of the original image from the encoded features justifies the conceptual representation of input image as encoded features [27], [28]. Hence, the encoded features are further processed for index matching and retrieval task. For the retrieval of best matches from the available database corresponding to any input query image, conventional index matching and retrieval module is used. The proposed model functioning is processed in three steps as follows:

- i) Refined Image Reconstruction Network
- ii) Robust Feature Generation
- iii) Medical Image Retrieval Module

### 2.1. Refined Image Reconstruction Network

The deployment of generative adversarial networks (GANs) for various problems like object segmentation, image enhancement, image super-resolution, image reconstruction, depth estimation, etc. GANs are utilized to produce the synthetic data resembling with the original data. The algorithmic architecture of GAN consists of two neural networks that are called generator and discriminator. The generator module generates unique plausible examples from the sample domain and the discriminator classifies it as real or fake content. Motivated by the successful results from GANs in literature, we proposed Res-UNet model for image restoration with self-attention for robust feature learning. The proposed network initially encodes the input into encoded features exploiting self-attention process. Consequently, the features are decoded back to restore the original image.

As shown in Figure 1, the proposed model designed with an encoder and decoder part. For the encoder part ResNet is utilized with its 50-layer architecture named ResNet50. The ResNet50 [23] network is considered till layer  $conv5\_x$ . The encoder part takes input ( $I_{256 \times 256 \times 3}$ ) as image and output a latent feature vector ( $Feat_{8 \times 8 \times 2048}$ ). The input/output relation of the proposed encoder is defined in (1) as follows:

$$Feat_{8 \times 8 \times 2048} = ResNet_{50}(I_{256 \times 256 \times 3}) . \quad (1)$$

Encoder is followed by a decoder while training the model. Decoder includes five layers each of which is comprises of convoTranspose layer, batch normalization layer and ReLU layer. Decoder up-samples the feature vector of the input image. The output of encoder ( $Feat_{8 \times 8 \times 2048}$ ) is fed to the decoder to which finally outputs the reconstructed image. The output from the decoder is given in (2) as follows:

$$Out = CT_{3 \times 3}(Feat)_{(2,2,2,2,2)}^{(512,256,128,64,3)} , \quad (2)$$

where  $CT_{3 \times 3}(\cdot)_S^F$  represents ConvTranspose and  $F$ : 512, 256, 128, 64, 3 is the number of filters used at each layer with corresponding stride  $S$ : 2, 2, 2, 2, 2.

### Reconstruction Loss

Proposed Res-UNet is an unsupervised learning technique, which generates output image as an input replica, i.e., reconstruction of input image. Hence to learn proposed architecture for reconstruction of image. We utilize the L1 loss as shown in (3) as follows

$$loss = I_{256 \times 256 \times 3} - Out , \quad (3)$$

where  $I_{256 \times 256 \times 3}$  is an input image and out is reconstructed image of  $I$ .

### 2.2. Robust Feature Generation

The accuracy of the image retrieval model highly depends on different representation of features through essential feature extraction method. Two of the major methods for feature extraction are hand-crafted feature descriptors and learning based feature descriptor. Hand-crafted feature extractors are rule based descriptors like SS3D [9], LTCoP [29], LTrP [15]. The learning based feature descriptors make use of the characteristic of input image like shape, size, texture, to extract the features (AlexNet [30], ResNet [23], VGG-16 [31]).

The proposed Res-UNet model uses the learning-based feature descriptor for image reconstruction process. Figure 2 represents the performance of the model for reconstruction of the image from the abstract features of an encoder. The results justify the ability of the encoder in the model for generating effective features from the input medical image. Hence, the

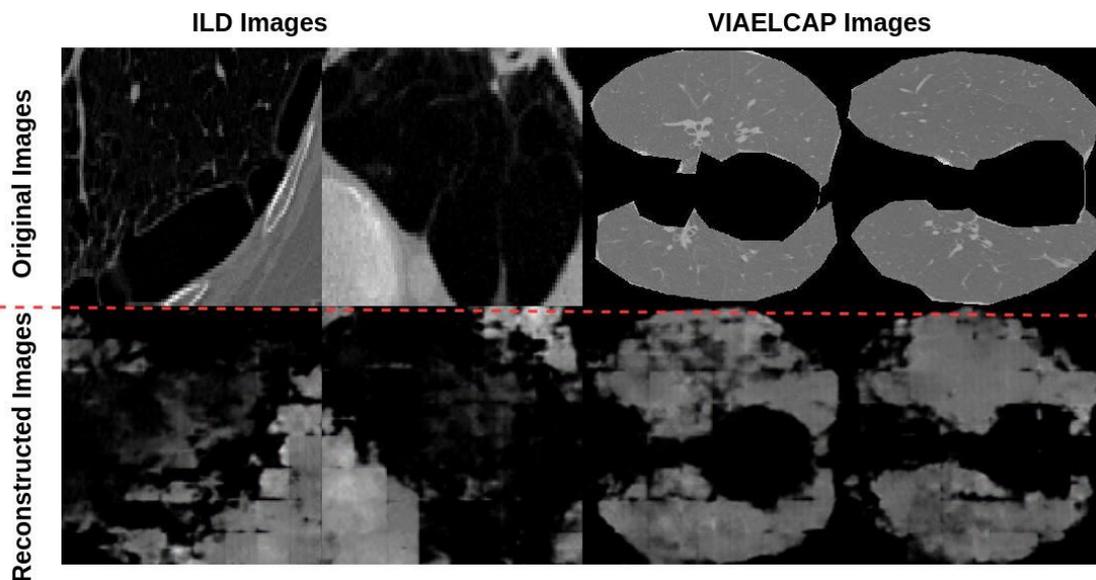


Figure 2. Upper row depicts the original database image. Lower row depicts the reconstructed image using proposed Res-UNet model. ILD and VIA/ELCAP-CT databases are used in column 1-2 and column 3-4 respectively.

output from the encoder, ( $Feat_{2 \times 2 \times 2048}$ ) is utilized as robust features of the corresponding input image. Further, these features are used for index matching and retrieval tasks from the database. In Figure 1, the lower side depicts the indexing and image retrieval from the extracted features of the query image.

### 2.3. Medical Image Retrieval Module

Let's consider,  $vf_Q = [vf_{Q1}, vf_{Q2}, \dots, vf_{QN}]$  defines the feature vector of input query image. Further, the feature vector of  $i^{th}$  image in datasets  $DB$  is,  $vf_{DBi} = [vf_{DBi1}, vf_{DBi2}, \dots, vf_{DBiN}]$ ,  $i = (1, 2, \dots, N)$  represents dataset images. For retrieving the similar images from database as the query image, a similarity index  $D$  is used between the input query image  $vf_Q$  and each image of the dataset  $vf_{DBi}$ . For our model similarity evaluation,  $D1$  distance is used for query image as most of the existing state-of-the-art methods have used the same. The  $D$  distance is given as in (4)

$$D(Q, DB) = \sum_{n=1}^N \left| \frac{vf_{DBm,n} - vf_{Qn}}{1 + vf_{DBm,n} + vf_{Qn}} \right|, \quad (4)$$

where,  $Q$  is the query image,  $N$  is the length of feature vector,  $DB$  is database image,  $vf_{DBm,n}$  is  $n^{th}$  feature of  $n^{th}$  image in the database,  $vf_{Qn}$  is  $n^{th}$  feature of query image.

### 3. TRAINING DETAILS

BraTS-2015 [32] database is considered for training the proposed model. The database consists of 220 high-grade gliomas and 54 low-grade gliomas. Database has scans of  $T1$ -weighted ( $T1$ ),  $T1$ -weighted imaging with gadolinium enhancing contrast ( $T1c$ ),  $T2$ -weighted ( $T2$ ), and fluid attenuated inversion recovery scans. 14,415 brain Magnetic Resonance Imaging (MRI) slices are taken from the entire BraTS-2015 database and 4155 slices are selected to train the model over 20 epochs. The time taken for training is 553 s per epoch and the batch size is 1.

## 4. RESULTS AND DISCUSSION

This section provides the elaborated performance analysis of the proposed model Res-UNet and gives a comparison with the conventional hand-crafted and learning based algorithm in the literature. For comparison, the work considered three benchmark medical image database with different modalities like MRI and Computer Tomography (CT).

### 4.1. Retrieval Accuracy on ILD Database

Two image sets interstitial images and clinical data for lung disease are examined for performing the experimental study from the integrated Interstitial Lung Disease (ILD) dataset [33]. ILD is known to be the finest database which is used for research and studying of lung disease with radiological and clinical data. The dataset is utilized as a referral resource to train new radiologists. The efficiency of ILD diagnosis can be improved by using a reduced invasive X-ray methodology, as X-ray and CT of a particular patient is kept in the same resource. The standard ILD dataset includes CT scans of lungs for 130 patients which has underlined disease regions by the experienced Radiologists for each ILD [33]. A smallpart of the dataset is available with 658 regions with marking of patches/regions for different classes such as fibrosis (187 regions), micro-nodules (173 regions), and healthy (139 regions), ground glass (106 regions), emphysema (53 regions). Table 1 provides the comparison evaluation of the proposed method and existing methods on the metric average retrieval precision applied on the topmost 10 images from the ILD dataset. Table 2 provides the comparison on the metric group-wise average retrieval precision on the ILD dataset. Results from Table 1 and Table 2 proves the out-performance of the proposed model from all the conventional methods.

Table 1. Comparison evaluation of the proposed method and existing methods on the metric average retrieval precision applied on the topmost 10 images from the ILD dataset.

| Method       | 1          | 2         | 3            | 4            | 5            | 6            | 7            | 8            | 9            | 10           |
|--------------|------------|-----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| SS3D [9]     | 100        | 83.21     | 76.85        | 72.34        | 69.27        | 66.44        | 64.48        | 62.71        | 61.43        | 60.40        |
| LTCoP [29]   | 100        | 84.95     | 76.29        | 71.77        | 67.44        | 64.79        | 62.57        | 60.83        | 59.14        | 58.13        |
| LTrP [34]    | 100        | 85.26     | 77.56        | 71.96        | 68.54        | 65.02        | 62.48        | 60.75        | 59.08        | 57.61        |
| MDMEP [11]   | 100        | 86.17     | 79.43        | 75.87        | 72.58        | 70.09        | 68.32        | 67.21        | 66.18        | 65.4         |
| AlexNet [30] | 100        | 88.91     | 83.99        | 79.6         | 76.66        | 74.54        | 72.38        | 71.12        | 69.79        | 68.45        |
| ResNet [23]  | 100        | 89.67     | 84.75        | 81.31        | 78.84        | 76.55        | 75.51        | 74.34        | 73.44        | 72.67        |
| VGG-16 [31]  | 100        | 89.21     | 83.94        | 80.24        | 77.96        | 75.94        | 74.4         | 73.21        | 72.26        | 71.34        |
| Res-UNet     | <b>100</b> | <b>98</b> | <b>97.33</b> | <b>97.25</b> | <b>97.20</b> | <b>97.00</b> | <b>96.85</b> | <b>96.75</b> | <b>96.66</b> | <b>96.60</b> |

Table 2. Comparison on the metric group-wise average retrieval precision on the ILD dataset.

| Method       | Group 1     | Group 2    | Group 3     | Group 4      | Group 5     | Average      |
|--------------|-------------|------------|-------------|--------------|-------------|--------------|
| SS3D [9]     | 47.17       | 72.62      | 53.21       | 52.82        | 75.68       | 60.30        |
| LTCoP [29]   | 62.64       | 69.95      | 53.96       | 48.72        | 78.97       | 62.85        |
| LTrP [34]    | 54.72       | 72.41      | 50.94       | 53.85        | 77.51       | 61.88        |
| MDMEP [11]   | 75.09       | 70.37      | 68.11       | 43.08        | 79.56       | 67.24        |
| AlexNet [30] | 65.28       | 83.10      | 65.85       | 48.72        | 82.64       | 69.12        |
| ResNet [23]  | 71.70       | 88.24      | 61.51       | 46.67        | 85.13       | 70.65        |
| VGG-16 [31]  | 73.96       | 87.49      | 57.17       | 57.44        | 83.22       | 71.86        |
| Res-UNet     | <b>93.3</b> | <b>100</b> | <b>97.3</b> | <b>96.09</b> | <b>97.3</b> | <b>96.79</b> |

Table 3. Comparison evaluation of the proposed method and existing methods on the metric average retrieval precision applied on the topmost 10 images from the VIA/I- ELCAP dataset.

| Method       | 1          | 2         | 3            | 4            | 5           | 6            | 7            | 8           | 9            | 10          |
|--------------|------------|-----------|--------------|--------------|-------------|--------------|--------------|-------------|--------------|-------------|
| SS3D [9]     | 100        | 74.25     | 64.27        | 58.70        | 54.78       | 52.55        | 50.47        | 48.68       | 47.44        | 46.23       |
| LTCoP [29]   | 100        | 93.80     | 91.10        | 88.75        | 86.80       | 85.35        | 84.13        | 82.60       | 81.31        | 80.28       |
| LTrP [34]    | 100        | 81.60     | 74.73        | 70.55        | 67.70       | 65.32        | 63.49        | 61.96       | 60.88        | 59.62       |
| MDMEP [11]   | 100        | 81.60     | 74.73        | 70.55        | 67.70       | 65.32        | 63.49        | 61.96       | 60.88        | 59.62       |
| AlexNet [30] | 100        | 99.60     | 99.13        | 98.63        | 98.06       | 97.43        | 96.56        | 95.85       | 94.89        | 93.88       |
| ResNet [23]  | 100        | 98.35     | 96.60        | 94.88        | 93.38       | 91.98        | 90.27        | 88.68       | 87.22        | 85.78       |
| VGG-16 [31]  | 100        | 94.55     | 90.33        | 86.88        | 84.16       | 81.85        | 79.76        | 77.78       | 75.96        | 74.23       |
| Res-UNet     | <b>100</b> | <b>99</b> | <b>98.66</b> | <b>98.75</b> | <b>98.6</b> | <b>98.33</b> | <b>98.42</b> | <b>98.5</b> | <b>98.55</b> | <b>98.6</b> |

Table 4. Comparison evaluation of the proposed method and existing methods on the metric average retrieval rate applied on the topmost 10 images from the VIA/I-ELCAP database.

| Method       | 1         | 2           | 3           | 4           | 5           | 6         | 7           | 8           | 9           | 10          |
|--------------|-----------|-------------|-------------|-------------|-------------|-----------|-------------|-------------|-------------|-------------|
| SS3D [9]     | 10        | 14.85       | 19.28       | 23.48       | 27.39       | 31.53     | 35.33       | 38.94       | 42.70       | 46.24       |
| LTCoP [29]   | 10        | 18.76       | 27.33       | 35.50       | 43.40       | 51.21     | 58.89       | 66.08       | 73.18       | 80.28       |
| LTrP [34]    | 10        | 16.32       | 22.42       | 28.22       | 33.85       | 39.19     | 44.44       | 49.57       | 54.79       | 59.62       |
| MDMEP [11]   | 10        | 16.32       | 22.42       | 28.22       | 33.85       | 39.19     | 44.44       | 49.57       | 54.79       | 59.62       |
| AlexNet [30] | 10        | 19.92       | 29.74       | 39.45       | 49.03       | 58.46     | 67.59       | 76.68       | 85.40       | 93.88       |
| ResNet [23]  | 10        | 19.67       | 28.98       | 37.95       | 46.69       | 55.19     | 63.19       | 70.94       | 78.50       | 85.78       |
| VGG-16 [31]  | 10        | 18.91       | 27.10       | 34.75       | 42.08       | 49.11     | 55.83       | 62.22       | 68.36       | 74.23       |
| Res-UNet     | <b>10</b> | <b>19.8</b> | <b>29.6</b> | <b>39.5</b> | <b>49.3</b> | <b>59</b> | <b>68.9</b> | <b>78.8</b> | <b>88.7</b> | <b>98.6</b> |

#### 4.2. Retrieval Accuracy on VIA/I-ELCAP-CP Database

The proposed model is evaluated for CBMIR on the VIA/I-ELCAP-CT [35] dataset. This dataset is widely used for the CBMIR model evaluation. The dataset comprises 1000 images splitted into 10 categories having 100 images per category. Table 3 depicts the comparison result in terms of average retrieval precision on the uppermost 10 images of database. Likewise, Table 4 provides the average retrieval recall comparison between the proposed and existing methods. Table 3 and Table 4 shows that the proposed method exceeds in the retrieval performance. The accuracy of conventional approaches reached till 93 %. However, the proposed method shows 99 % accuracy for CBMIR. The produced results can be justified as the power of network in adversarial feature learning.

#### 5. CONCLUSION

In this paper, a novel content-based Res-UNet framework is proposed for reconstruction of the input medical image and performs an efficient image retrieval task. The framework operates in two phases, training phase and inference phase. In the training phase, the model utilizes an encoder and a decoder to reconstruct the input medical image. In our proposed work, ResNet50 is utilized as an encoder to perform the encoding of feature vectors. In phase 2, we only utilize the encoder part to generate the feature vector of the input image. Further, utilizing those features, similar images are retrieved from the database using a similarity metric. The performance evaluation of the proposed model is done on the two benchmark datasets that are ILD and VIA/ELCAP-CT. Comparison results shows the outperformance of the proposed model as compared to the conventional methods.

#### REFERENCES

- [1] Mandal Murari, Mallika Chaudhary, Santosh Kumar Vipparthi, Subrahmanyam Murala, Anil Balaji Gonde, Shyam Krishna Nagar, ANTIC: ANTIthetic isomeric cluster patterns for medical image retrieval and change detection, IET Computer Vision, vol. 13, no.1, 2019, pp. 31-43.  
DOI: [10.1049/iet-cvi.2018.5206](https://doi.org/10.1049/iet-cvi.2018.5206)
- [2] Vipparthi Santosh Kumar, Subrahmanyam Murala, Anil Balaji Gonde, Q. M. Jonathan Wu, Local directional mask maximum edge patterns for image retrieval and face recognition, IET Computer Vision, vol.10, no.3, 2016, pp. 182-192.  
DOI: [10.1049/iet-cvi.2015.0035](https://doi.org/10.1049/iet-cvi.2015.0035)
- [3] Vipparthi Santosh Kumar, Subrahmanyam Murala, S. K. Nagar, Anil Balaji Gonde, Local Gabor maximum edge position octal patterns for image retrieval, Neurocomputing, vol.167, 2015, pp. 336-345.  
DOI: [10.1016/j.neucom.2015.04.062](https://doi.org/10.1016/j.neucom.2015.04.062)
- [4] Vipparthi Santosh Kumar, S. K. Nagar, Expert image retrieval system using directional local motif XoR patterns, Expert Systems with Applications, vol. 41, no. 17, 2014, pp. 8016-8026.  
DOI: [10.1016/j.eswa.2014.07.001](https://doi.org/10.1016/j.eswa.2014.07.001)
- [5] Vipparthi Santosh Kumar, Shyam Krishna Nagar, Multi-joint histogram based modelling for image indexing and retrieval, Computers & Electrical Engineering, vol. 40, no. 8, 2014, pp. 163-173.  
DOI: [10.1016/j.compeleceng.2014.04.018](https://doi.org/10.1016/j.compeleceng.2014.04.018)
- [6] D. G. Lowe, Object recognition from local scale-invariant features, Proceedings of the seventh IEEE international conference on computer vision, IEEE, vol. 2, 1999, pp. 1150-1157.  
DOI: [10.1109/ICCV.1999.790410](https://doi.org/10.1109/ICCV.1999.790410)
- [7] Ojala Timo, Matti Pietikainen, Topi Maenpaa, Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on pattern analysis and machine intelligence, vol. 24, no.7, July 2002, pp. 971-987.  
DOI: [10.1109/TPAMI.2002.1017623](https://doi.org/10.1109/TPAMI.2002.1017623)
- [8] Tan Xiaoyang, Bill Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE

- transactions on image processing, vol. 19, no. 6, 2010, pp. 1635-1650.  
DOI: [10.1007/978-3-540-75690-3\\_13](https://doi.org/10.1007/978-3-540-75690-3_13)
- [9] Murala Subrahmanyam, Q. M. Jonathan Wu, Spherical symmetric 3D local ternary patterns for natural, texture and biomedical image indexing and retrieval, *Neurocomputing*, vol. 149, 2015, pp. 1502-1514.  
DOI: [10.1016/j.neucom.2014.08.042](https://doi.org/10.1016/j.neucom.2014.08.042)
- [10] Vipparthi Santosh Kumar, S. K. Nagar, Directional local ternary patterns for multimedia image indexing and retrieval, *International Journal of Signal and Imaging Systems Engineering*, vol. 8, no. 3, 2015, pp. 137-145.  
DOI: [10.1504/IJSISE.2015.070485](https://doi.org/10.1504/IJSISE.2015.070485)
- [11] G. M. Galshetwar, L. M. Waghmare, A. B. Gonde, 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, 2018, pp. 231-239.  
DOI: [10.1016/j.bbe.2020.07.011](https://doi.org/10.1016/j.bbe.2020.07.011)
- [12] G. M. Galshetwar, L. M. Waghmare, A. B. Gonde, S. Murala, Local energy oriented pattern for image indexing and retrieval, *Journal of Visual Communication and Image Representation*, vol. 64, 2019, pp. 102615.  
DOI: [10.1016/j.jvcir.2019.102615](https://doi.org/10.1016/j.jvcir.2019.102615)
- [13] Biswas Ranjit, Sudipta Roy, Debraj Purkayastha, An efficient content-based medical image indexing and retrieval using local texture feature descriptors, *International Journal of Multimedia Information Retrieval*, vol. 8, no. 4, 2019, pp. 217-231.  
DOI: [10.1007/s13735-019-00176-9](https://doi.org/10.1007/s13735-019-00176-9)
- [14] Vipparthi Santosh Kumar, Subrahmanyam Murala, Shyam Krishna Nagar, Dual directional multi-motif XOR patterns: A new feature descriptor for image indexing and retrieval, *Optik*, vol. 126, 2015, pp. 1467-1473.  
DOI: [10.1016/j.jilco.2015.04.018](https://doi.org/10.1016/j.jilco.2015.04.018)
- [15] Akshay Dudhane, Kuldeep M. Biradar, Prashant W. Patil, Praful Hambarde, Subrahmanyam Murala, Varicolored image de-hazing, *Proc. of the IEEE/CVF conference on computer vision and pattern recognition*, Seattle, WA, USA, 13-19 June 2020, pp. 4563-4572.  
DOI: [10.1109/CVPR42600.2020.00462](https://doi.org/10.1109/CVPR42600.2020.00462)
- [16] Prashant W. Patil, Kuldeep M. Biradar, Akshay Dudhane, Subrahmanyam Murala, An end-to-end edge aggregation network for moving object segmentation, *Proc. of the IEEE/CVF conference on computer vision and pattern recognition*, Seattle, WA, USA, 13-19 June 2020, pp. 8146-8155.  
DOI: [10.1109/CVPR42600.2020.00817](https://doi.org/10.1109/CVPR42600.2020.00817)
- [17] Kuldeep Marotirao Biradar, Ayushi Gupta, Murari Mandal, Santosh Kumar Vipparthi, Challenges in time-stamp aware anomaly detection in traffic videos, 2019, 8 pp.  
DOI: [10.48550/arXiv.1906.04574](https://doi.org/10.48550/arXiv.1906.04574)
- [18] Kuldeep Biradar, Sachin Dube, Santosh Kumar Vipparthi, DEAREST: Deep Convolutional aberrant behavior detection in real-world scenarios, *IEEE 13th International Conference on Industrial and Information Systems (ICIIS)*, IEEE Rupnagar, India, 1-2 December 2018, pp. 163-167.  
DOI: [10.1109/ICIINF.2018.8721378](https://doi.org/10.1109/ICIINF.2018.8721378)
- [19] E. S. Kumar, C. Shoba Bindu, Medical Image Analysis using Deep Learning: A Systematic Literature Review, *International Conference on Emerging Technologies in Computer Engineering*, Springer, vol 985, 2019, pp. 81-97.  
DOI: [10.1007/978-981-13-8300-7\\_8](https://doi.org/10.1007/978-981-13-8300-7_8)
- [20] R. Pinapatruni, C. S. Bindu, Learning image representation from image reconstruction for a content-based medical image retrieval, *Signal, Image and Video Processing*, vol. 14, no.7, 2020, pp. 1319-1326.  
DOI: [10.1007/s11760-020-01670-y](https://doi.org/10.1007/s11760-020-01670-y)
- [21] Praful Hambarde, Sanjay Talbar, Abhishek Mahajan, Satishkumar Chavan, Meenakshi Thakur, Nilesh Sable, Prostate lesion segmentation in MR images using radiomics based deeply supervised U-Net, *Biocybernetics and Biomedical Engineering*, vol. 40, no. 4, 2020, pp. 1421-1435.  
DOI: [10.1016/j.bbe.2020.07.011](https://doi.org/10.1016/j.bbe.2020.07.011)
- [22] Olaf Ronneberger, Philipp Fischer, Thomas Brox, U-net: Convolutional networks for biomedical image segmentation, *International Conference on Medical image computing and computer-assisted intervention*, Springer, Cham, 2015.  
DOI: [10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [23] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.  
DOI: [10.1109/cvpr.2016.90](https://doi.org/10.1109/cvpr.2016.90)
- [24] Praful Hambarde, Akshay Dudhane, Prashant W. Patil Subrahmanyam Murala, Abhinav Dhall, Depth estimation from single image and semantic prior, *IEEE International Conference on Image Processing (ICIP)*, Abu Dhabi, United Arab Emirates, 25-28 October 2020, IEEE, pp. 1441-1445.  
DOI: [10.1109/ICIP40778.2020.9190985](https://doi.org/10.1109/ICIP40778.2020.9190985)
- [25] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alexander A. Alemi, Inception-v4, Inception-ResNet and the impact of residual connections on learning, *Thirty-first AAAI conference on artificial intelligence*, vol. 31, no. 1, 2017. Online [Accessed 24 March 2022] <https://ojs.aaai.org/index.php/AAAI/article/view/11231>
- [26] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros, Image-to-image translation with conditional adversarial networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 5967-5976.  
DOI: [10.1109/CVPR.2017.632](https://doi.org/10.1109/CVPR.2017.632)
- [27] Murala Subrahmanyam, Q. M. Jonathan Wu, 3D local circular difference patterns for biomedical image retrieval, *International Journal of Multimedia Information Retrieval*, vol. 8, no. 2, pp. 115-125.  
DOI: [10.1016/j.neucom.2014.08.042](https://doi.org/10.1016/j.neucom.2014.08.042)
- [28] Vipparthi Santosh Kumar, Shyam Krishna Nagar, Integration of color and local derivative pattern features for content-based image indexing and retrieval, *Journal of the Institution of Engineers (India): Series B*, vol. 96, no. 3, 2015, pp. 251-263.  
DOI: [10.1007/s40031-014-0153-5](https://doi.org/10.1007/s40031-014-0153-5)
- [29] Murala, Subrahmanyam Q. M. Jonathan Wu, Local ternary co-occurrence patterns: a new feature descriptor for MRI and CT image retrieval, *Neurocomputing*, vol. 119, 2013, pp. 399-412.  
DOI: [10.1016/j.neucom.2013.03.018](https://doi.org/10.1016/j.neucom.2013.03.018)
- [30] Krizhevsky Alex, Ilya Sutskever, Geoffrey E. Hinton, Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, vol. 25, 2012, pp. 1097-1105.  
DOI: [10.1145/3065386](https://doi.org/10.1145/3065386)
- [31] Simonyan Karen, Andrew Zisserman, Very deep convolutional networks for large-scale image recognition, *ICLR-2015*. Online [Accessed 24 March 2022] <https://arxiv.org/abs/1409.1556>
- [32] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, K. H. Maier-Hein, Brain tumor segmentation and radiomics survival prediction, Contribution to the BRATS 2017 challenge, *International MICCAI Brainlesion Workshop*, Springer, pp. 287-297.  
DOI: [10.1007/978-3-319-75238-9\\_25](https://doi.org/10.1007/978-3-319-75238-9_25)
- [33] Adrien Depeursinge, Alejandro Vargas, Alexandra Platon, Antoine Geissbuhler, Pierre-Alexandre Poletti, Henning Müller, Building a reference multimedia database for interstitial lung diseases, *Computerized medical imaging and graphics*, vol. 36, no. 3, 2012, pp. 227- 238.  
DOI: [10.1016/j.compmedimag.2011.07.003](https://doi.org/10.1016/j.compmedimag.2011.07.003)
- [34] Murala Subrahmanyam, R. P. Maheshwari, R. Balasubramanian, Local tetra patterns: a new feature descriptor for content-based image retrieval, *IEEE transactions on image processing*, vol. 21, no. 5, 2012, pp. 2874-2886.  
DOI: [10.1109/TIP.2012.2188809](https://doi.org/10.1109/TIP.2012.2188809)
- [35] Via/i-elcap database. Online [Accessed 10 March 2019] <http://www.via.cornell.edu/lungdb.html>