Content-based Medical Image Retrieval based on Deep Features Expansion

Metwally Rashad
dept. Computer Science
Computers & Artificial Intelligence
Benha, Egypt
metwally.rashad@fci.bu.edu.eg

Ibrahem Afifi
dept. Information Systems
Computers & Artificial Intelligence
Benha, Egypt
ibrahim.afify@fci.bu.edu.eg

Mohamed Abdelfatah
dept. Information Systems
Computers & Artificial Intelligence
Benha, Egypt
mohamed.abdo@fci.bu.edu.eg

Abstract—The collections of various digital image databases have significantly grown and many users have recognized that finding and recovering important images from large collections is a difficult task. Where the success of any image retrieval system is heavily dependent on the feature extraction capacity of the feature descriptor, therefore successful and effective retrieval method has been developed to provide an effective and rapid search and retrieval process. We present a unique deep learning-based approach for extracting high-level and compact features from medical images in this paper. To capture the discriminative features of medical images, we use Residual Networks (ResNets), a popular multi-layered deep neural network. The query is then broadened by reformulating the query image using the mean values for deep features from each database class's top-ranking images. Two publicly available databases in various forms were used to evaluate the performance of our technique. These studies demonstrated the benefits of our proposed strategy, with retrieval accuracy greatly improved.

Index Terms—Medical image retrieval, Deep learning technology, Residual Networks

I. Introduction

In recent years, large images and multimedia content repositories have been created by the rapid growth of digital, multimedia, and storage computers. The advances in digital storage and information delivery benefit clinical and diagnostic studies. Diagnostic and investigative imaging systems in hospitals produce an immense amount of images and thus increase the production of medical image collections. The implementation of an efficient framework for medical imaging is therefore important to help clinicians access such large databases. Many of the algorithms for the automated analysis of medical images were proposed in the literature to promote the creation and management of these vast databases of medical images [1]. Also, an effective means of supplementing diagnostics and treatment of different diseases and an efficient management tool for managing vast volumes of data may be provided by using the CBMIR. The process of access, management, and extraction from such large collections is highly complicated without these systems. Also, the retrieval of medical images based on textual details like tags and manual annotations are poor in productivity since they require manpower, clinical skills, and time. Medical image retrieval systems are therefore essential, and can automatically retrieve and identify images based on representations of the features resulting from the images themselves. It complements the clinical decision support systems, research, and clinical trials for the purpose of searching in large repositories for relevant information.

According to recent research, one of the most active medical image processing research domains is Content Based Medical Image Retrieval (CBMIR). This is due to the fact that numerous medical methods, such as ultrasound (US), magnetic resonance imaging (MRI), X-ray, and CT, are rapidly expanding to make a precise diagnosis and histopathological images, which show a clear image of diseases and the affected tissue while maintaining the underlying tissue architecture [2]. Any CBMIR approach is designed with two basic phases in mind: feature extraction (offline phase) and similarity measurement calculations (online phase) [3]. The main structure of a CBMIR system is depicted in Fig. 1. Many enhancements have been built for the CBMIR system to improve its efficacy and retrieval performance.

Matching pairs to the image are the core aspect of any CBMIR system, in which an image is compared to a database image to determine similarity. Traditional approaches analyze the textures, shapes, colors, and spatial organization of medical images to extract low-level information. These are all low-level qualities that frequently do not effectively convey semantic conceptions in images. Using these attributes for retrieval usually yields unsatisfactory results. To achieve the best retrieval accuracy, deep learning can be used as a feature extraction method in addition to a new retrieval method, as in our method, where deep learning reflects the deep architectural human brain [4], and using many hidden layers to represent and collecting the crucial high-resolution image features.

II. Related Works

CBIR is an image search method that seeks to identify images most comparable to a specific query, where the search for images is based on their features, whether low

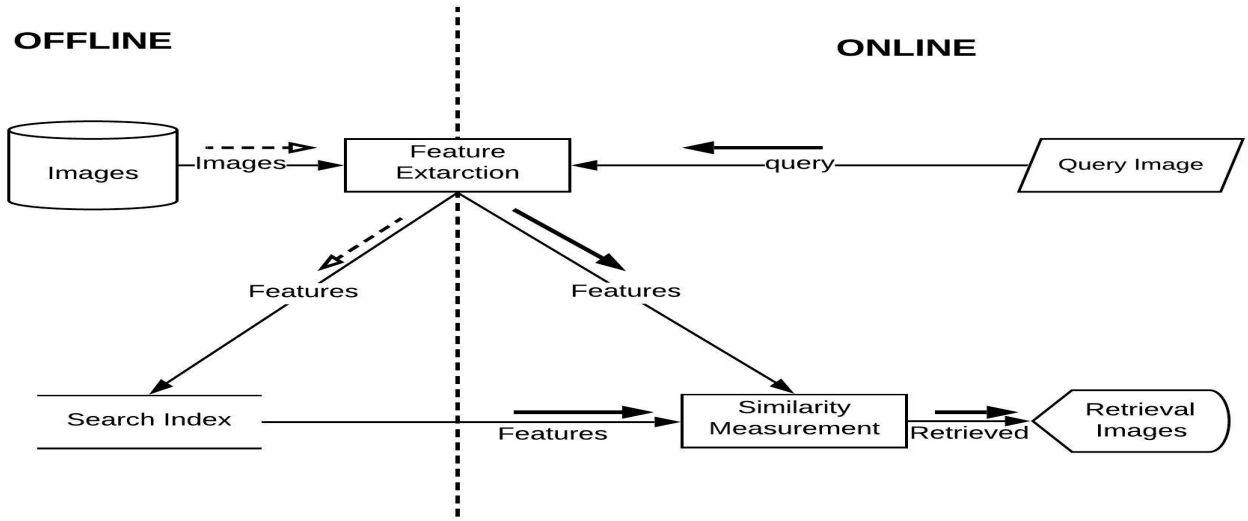

Fig. 1: The Framework of the CBMIR.

levels, such as colors, texture, or high level. CBIR's success depends mainly on the selected features [5] where the images are represented as features with a high dimension. The metrics of distance like Euclidean Distance (ED) are used to measure the similarity between the query and images stored in the databases. The encoding of image data in terms of features and the selection of a similarity measurement technique are thus the most significant components of CBIR systems. Although several researchers have researched these subjects widely. The most challenging difficulty in CBIR systems is still closing the semantic gap. It is information that is lost in the representation of an image from high-level semantic to low-level characteristics novel where this gap between the visual information of the human visual system (HVS) and the imaging system happens. Either integrating knowledge unique to the field or using some techniques of machine learning to create intelligent systems that can be equipped to behave like HVS, will reduce this gap.

The most contemporary and popular ways for retrieving medical images in recent years have centered on local features. One of the local features that have been widely used for medical picture indexing and retrieval is the Local Binary Pattern (LBP) descriptor, which was proposed by [6].

Also, the medical images are available in a variety of formats. Our study was limited to only two types of medical images: MRI and histopathological images. The authors in [7] suggested a new method works on the pixel in the center and the neighbors around it by encoding the relationship between them into a 3D plane created from a 2D image in five specified directions using a multiresolution Gaussian filter bank. Furthermore, they also suggested the color SS-3D-LTP, which views the RGB spaces as three planes of 3D volume. In [8], the authors presented an image reconstruction network (IR-Net), in which the input image is encoded into a collection of features before being reconstructed from these features too. The feature learning in (IR-Net) is split into two phases: the first phase consists of an IR-Net (image reconstruction network) designed to reproduce the provided input image using a new encoder-decoder deep network. The provided medical scan is encoded by IR-Net into a series of features, which are then utilized to reproduce the input image using a unique decoder network. When they successfully recreate the input image from the encoded features, they have proved that the image is considered strongly represented by it.

For the histopathological images, metric learning from radiology records [9] functions well, but it requires a lot of computing power. Only for smaller datasets does the unsupervised region suggestion technique [10] for histopathology image retrieval give excellent retrieval accuracy. Although encoding images as binary codes with deep learning-based hashing [11] improves the performance of retrieval for histopathology image collections, it limits total performance and loss of the incurs quantization. The authors in [12] presented a scheme that down-samples the image to different scales, in which for each down-sampled the LTP is used for split into concentrated patches, lower, and upper. The scheme defines the images on a different scale. In order to achieve the high accuracy of the retrieval of histopathological imagery, they also have applied power-law normalization and a vector of locally aggregated descriptors (VLAD) to get the histogram images.

III. Deep Features Expansion Method

Deep learning has grown in popularity in recent years, with several promising implementations in a wide range of fields [13]. Despite the fact that the structure of several deep learning has been suggested and implemented, the fundamental concept remains the same: deep learning is the representation of features and learning technique that

focuses on vast volumes of raw image data and can apply various levels of representation.

Residual Networks (ResNets) [14] are utilized in this paper for medical image retrieval. Identity mapping is used in ResNets, a form of neural network. This means that the input to one layer is transmitted directly or indirectly to another. Can be easily trained without increasing the percentage of training errors. Using identity mapping, ResNets aid in addressing the vanishing gradient problem. We used the ResNet-50 as a features extractor composed of 50 layers in deep, which is considered a pre-trained DCNN (Deep Convolutional Neural Network) on the ImageNet database [15] that contains millions of labeled images, and also the ResNet-101, which also a pre-trained DCNN on the ImageNet that is composed of 101 layers in deep.

In the field of information retrieval, the deep features expansion method is considered the most efficient and also simplest in used expansion process based on top-rank images, in which the expanded is automatically occurs to the query for obtaining its top-ranked neighbors, the chances of obtaining the highest semantically related output are improved. The basic idea behind this method is that the images that are most related to the query are retrieved and then used to extend and reformulate new query images with information from these most relevant images. The expanding process is based on the top images provided by a single query. Based on what we've said, the values of the mean derived for each feature of images that have a top-ranked are thought to be quite effective. We used this method very effectively to improve the results at the end, where instead of taking the images with top-ranking from the database in total, we took from each class the images with top-ranking and then applied the scenario of calculating the value for the mean of the features of these images with top-ranked from each class. This mean-value expansion method is also employed for the first time in deep features (ResNets).

This type of expansion is used for ResNet features. As demonstrated in Fig. 1, our method uses the mean values of the ResNet feature for the top-ranked images following a rapid search utilizing one image Query of ResNet Feature (QRF) on all ResNet features in the database. Each image has a ResNet Feature Vector (RFV) dimension of 1000. In each class, the value of the mean for each feature of ResNet is derived from the top 10 related images. For the number of classes in the database, this technique will generate the same number of DFENQ (Deep Features Expansion for New Query), for example, if we have five classes then we will have five DFENQ. The most similar DFENQ to the query is then chosen as the final DFENQ (FDFENQ), and the FDFENQ is used in the final search operation. Table I shows a simple numerical example of how to construct DFENQ.

TABLE I: DFENQ based on mean values.

	F1	F2	F3	F4	F5	.....	F1000
Img1	0.32	0.67	0.39	3.24	2.75	.....	3.22
Img2	0.49	0.26	0.18	3.02	3.47	.....	2.83
Img3	-0.54	-0.66	0.42	2.74	3.08	.....	3.66
Img4	0.21	0.27	-0.04	2.92	3.54	.....	2.94
Img5	-0.23	-0.43	0.18	2.71	2.45	.....	3.28
DFENQ	0.05	0.022	0.226	2.926	3.058	.....	3.186

IV. Experimental Results

In our tests, we exploit every image in the database as an image query, and the returned image is only relevant if it is in the same category as the image query. There are three performance measurements are used to assess each retrieval method: the Average Precision Retrieval (ARP), the Average Retrieval Rate (ARR), and the F_{score}.

As previously stated in the methodology's fundamental structure, there are two searching procedures: a rapid search for each class in the database classes using a single query image taken from the images. Then comes the last search, which makes use of the final DFENQ (FDFENQ). Similarity measures are calculated with the ED (Euclidean Distance)for both rapid and final search.

$$ED\left(X,Y\right)=\sqrt{\sum_{i=1}^{n}\left(x_i-y_i\right)^2} \tag{1}$$

A. Image Model Databases

We used two publicly available image databases with different formats, for testing purposes.

The first database is the Open Access Series (OASIS) with MRI [16], where the OASIS-MRI is considered a type of medical database created by the Image Studies Open Access Series (OASIS) with MRI, where a sectional range of 421 topics between 18 and 96 years can be studied and analyzed within the original databases. The authors of [17] divided these photos into four classes, having 124, 102, 89, and 106 images in each class, as shown in Fig. 2.

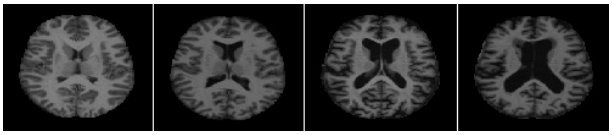

Fig. 2: Image from each class of the OASIS-MRI database.

The second is the KIMIA Path960 database [18] comprises 20 histopathological image classes created by the epithelial, muscle, and connective tissue set. In addition, the images of each class have certain characteristics in which there are high differences in intra classes and certain similarities between different classes. The KIMIA Path960 database contains 960 Images, where there are 48 images in each class as shown in Fig. 3.

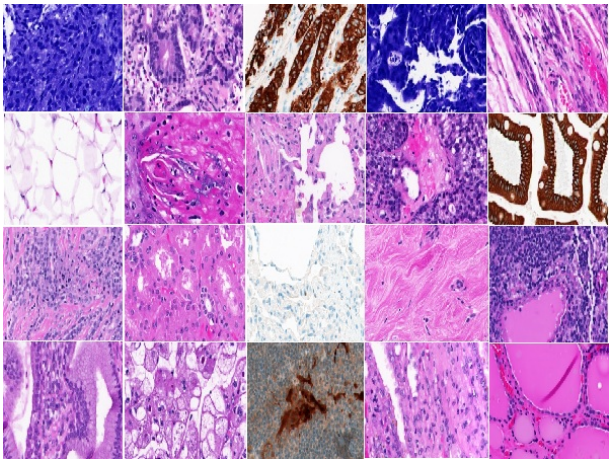

Fig. 3: Image from each class of the KIMIA Path960 database.

B. Retrieval Performance on OASIS-MRI Database

We have compared our method with ResNet-50 as a features extractor and our method with ResNet-101 as another features extractor. The efficiency of our method with the two types of features extractor has achieved high performance compared to IR-Net [8] in addition to some other modern methods. The result of the top 10 images as presented in Table II in terms of ARP and Table III present the result of the group-wise in terms of ARP. The proposed method with ResNet-50 achieved high performance on the top 10 images compared to all other methods and in terms of average group-wise, the proposed method with ResNet-50 exceed the IR-Net method with 6.04% while the proposed method with ResNet-101 exceeds the IR-Net method with 4.58%. Fig. 4 presents the query results of our method with ResNet-50.

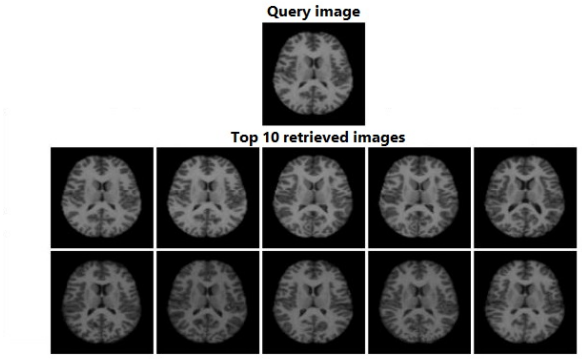

Fig. 4: Retrieved images for an OASIS-MRI database query using our method with ResNet-50.

C. Retrieval Performance on KIMIA Path960 Database

The Path960 KIMIA database is used in this experiment. We compared our method with two types of features extractor against a modern medical image retrieval method on a recent histopathological, LTP+VLAD [12] in addition to some other methods as explained in [12]. The retrieval performance is presented in Tables IV, V, and VI in terms of ARP, ARR, and F_{score} respectively. It is clear from the tables that our method with ResNet-50 is superior to other methods on top 5, 10, 15, and 20 images and also exceed the proposed method with ResNet-101 on top 5, 15, and 20 images exception on top 10 images. The result of the retrieval for our method with ResNet-50 are improved by 2.14%, 1.94%, 6.74%, 9.42% respectively in terms of ARP, and improved by 0.23%, 0.42%, 2.15%, 4.01% respectively in terms of ARR, also improved by 0.41%, 0.69%, 3.26% ,5.62% respectively in terms of F_{score} as compared with the LTP+VLAD method. Fig. 5, 6, 7, and 8 present the query results of the top 5, 10, 15,and 20 respectively for the proposed method with ResNet-50 over the KIMIA Path960 database which appears that the images that retrieved for query are from the same category of the query.

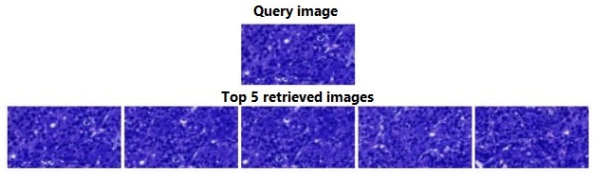

Fig. 5: Retrieving the top 5 images for a KIMIA Path960 database query using our method with ResNet-50.

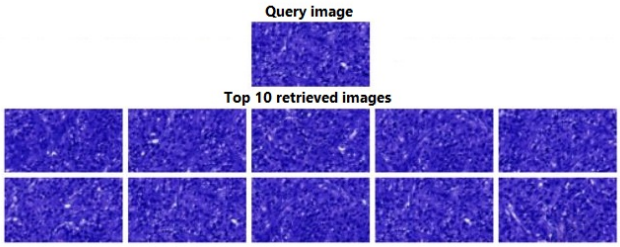

Fig. 6: Retrieving the top 10 images for a KIMIA Path960 database query using our method with ResNet-50.

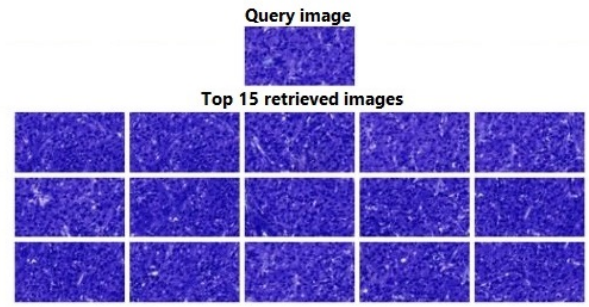

Fig. 7: Retrieving the top 15 images for a KIMIA Path960 database query using our method with ResNet-50.

TABLE II: Comparison of retrieval accuracy for the top 10 matches of the method suggested and other modern methods in an OASIS database.

Method	Top1	Top2	Top3	Top4	Top5	Top6	Top7	Top8	Top9	Top10
LBP [6]	100	72.92	61.20	56.89	53.25	50.04	48.63	47.21	45.87	44.66
SS3D [7]	100	68.53	56.77	51.54	47.89	45.68	43.77	42.37	41.44	40.45
AlexNet [19]	100	80.88	73.56	68.41	65.46	62.98	61.83	60.54	59.46	58.36
ResNet [14]	100	78.50	71.18	67.99	64.94	62.87	60.54	59.26	58.38	57.48
VGG-16 [20]	100	76.60	68.88	63.84	61.24	59.62	58.23	57.60	56.03	55.06
IR-Net [8]	100	83.37	78.62	76.37	74.68	73.44	72.21	71.62	70.89	70.45
Proposed method with ResNet – 101	**100**	**100**	**99.13**	**94.00**	**88.84**	**84.56**	**80.59**	**78.27**	**76.43**	**74.73**
Proposed method with ResNet – 50	**100**	**100**	**99.45**	**95.78**	**92.02**	**87.81**	**83.71**	**81.09**	**78.57**	**76.32**

TABLE III: Comparison of retrieval accuracy for group-wise on OASIS database in terms of ARP.

Method	Group1	Group2	Group3	Group4	Average
LBP [6]	55.08	35.20	32.70	51.60	43.64
SS3D [7]	44.19	39.02	35.73	41.42	40.08
AlexNet [19]	68.87	43.73	41.01	74.72	57.08
ResNet [14]	69.11	45.59	44.27	66.42	56.35
VGG-16 [20]	70.73	44.61	38.09	61.04	53.62
IR-Net [8]	77.10	55.29	57.19	88.40	69.49
Proposed method with ResNet – 101	**83.87**	**57.35**	**72.02**	**83.02**	**74.07**
Proposed method with ResNet – 50	**84.76**	**65.69**	**68.09**	**83.58**	**75.53**

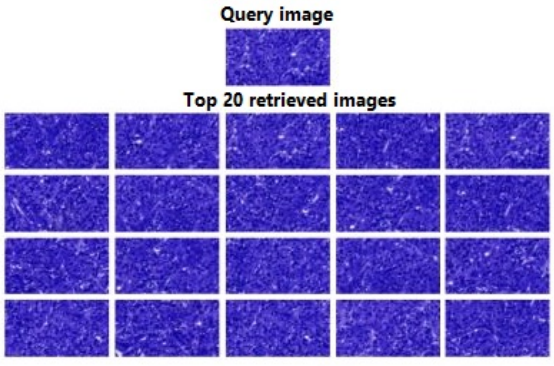

Fig. 8: Retrieving the top 20 images for a KIMIA Path960 database query using our method with ResNet-50.

TABLE IV: ARP for various methods on the KIMIA Path960 database for the top 5, 10, 15, and 20 retrieved images.

Method	Top5	Top10	Top15	Top20
LBP [6]	88.74	84.05	80.50	75.89
AlexNet [19]	93.10	91.04	85.01	81.98
VGG-16 [20]	94.12	92.40	85.41	82.24
DL [21]	94.51	92.56	85.57	81.89
LTP+VLAD [12]	96.92	95.00	88.57	84.65
Proposed method with ResNet – 101	**99.04**	**96.96**	**94.69**	**92.89**
Proposed method with ResNet – 50	**99.06**	**96.94**	**95.31**	**94.07**

V. Conclusion

We proposed a new medical image retrieval approach in this research. The proposed method is based on top-ranking images extracted using pre-trained models of deep learning techniques such as ResNet-50 and ResNet-101 for extracting high-level features from medical images,

followed by reformulating the query image using mean values of ResNet features extracted from each database class. The user is just required to enter the query image; otherwise, this approach is deemed fully automated. This approach was evaluated on two different types of standard and publically available databases: OASIS-MRI and KIMIA Path960. The results demonstrate great advancement and high efficiency when compared to other modern approaches.

TABLE V: ARR for various methods on the KIMIA Path960 database for the top 5, 10, 15, and 20 retrieved images.

Method	Top5	Top10	Top15	Top20
LBP [6]	9.44	17.88	25.69	32.29
AlexNet [19]	9.90	19.37	27.13	34.88
VGG-16 [20]	10.01	19.66	27.26	34.99
DL [21]	10.05	19.69	27.31	34.85
LTP+VLAD [12]	10.31	20.21	28.27	36.02
Proposed method with ResNet − 101	**10.54**	**20.63**	**30.22**	**39.53**
Proposed method with ResNet − 50	**10.54**	**20.63**	**30.42**	**40.03**

TABLE VI: F_{score} for various methods on the KIMIA Path960 database for the top 5, 10, 15, and 20 retrieved images.

Method	Top5	Top10	Top15	Top20
LBP [6]	17.06	29.49	38.95	45.30
AlexNet [19]	17.89	31.94	41.13	48.94
VGG-16 [20]	18.24	19.66	32.42	49.09
DL [21]	18.17	32.47	41.41	48.89
LTP+VLAD [12]	18.64	33.32	42.86	50.54
Proposed method with ResNet − 101	**19.05**	**34.02**	**45.82**	**55.46**
Proposed method with ResNet − 50	**19.05**	**34.01**	**46.12**	**56.16**

References

[1] M. Mizotin, J. Benois-Pineau, M. Allard, and G. Catheline, "Feature-based brain mri retrieval for alzheimer disease diagnosis," pp. 1241–1244, 2012.

[2] M. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, and B. Yener, "Histopathological image analysis: A review," IEEE Reviews in Biomedical Engineering, vol. 2, pp. 147–171, 2009.

[3] M. Owais, M. Arsalan, J. Choi, and K. R. Park, "Effective diagnosis and treatment through content-based medical image retrieval (CBMIR) by using artificial intelligence," Journal of Clinical Medicine, vol. 8, no. 4, p. 462, apr 2019.

[4] J. Wan, D. Wang, S. C. H. Hoi, P. Wu, J. Zhu, Y. Zhang, and J. Li, "Deep learning for content-based image retrieval," in Proceedings of the 22nd ACM international conference on Multimedia. ACM, nov 2014.

[5] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," Pattern recognition, vol. 40, no. 1, pp. 262–282, 2007.

[6] 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, jan 1996.

[7] 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, feb 2015.

[8] R. Pinapatruni and 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, pp. 1319–1326, mar 2020.

[9] J. Ramos, T. T. Kockelkorn, I. Ramos, R. Ramos, J. Grutters, M. A. Viergever, B. van Ginneken, and A. Campilho, "Content-based image retrieval by metric learning from radiology reports: application to interstitial lung diseases," IEEE journal of biomedical and health informatics, vol. 20, no. 1, pp. 281–292, 2014.

[10] Y. Ma, Z. Jiang, H. Zhang, F. Xie, Y. Zheng, H. Shi, Y. Zhao, and J. Shi, "Generating region proposals for histopathological whole slide image retrieval," Computer methods and programs in biomedicine, vol. 159, pp. 1–10, 2018.

[11] M. Sapkota, X. Shi, F. Xing, and L. Yang, "Deep convolutional hashing for low-dimensional binary embedding of histopathological images," IEEE journal of biomedical and health informatics, vol. 23, no. 2, pp. 805–816, 2018.

[12] K. Sukhia, M. Riaz, A. Ghafoor, S. Ali, and N. Iltaf, "Content-based histopathological image retrieval using multi-scale and multichannel decoder based LTP," Biomedical Signal Processing and Control, vol. 54, p. 101582, sep 2019.

[13] M. Wainberg, D. Merico, A. Delong, and B. J. Frey, "Deep learning in biomedicine," Nature Biotechnology, vol. 36, no. 9, pp. 829–838, oct 2018.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, jun 2016.

[15] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," jun 2009.

[16] D. S. Marcus, A. F. Fotenos, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open access series of imaging studies: Longitudinal MRI data in nondemented and demented older adults," Journal of Cognitive Neuroscience, vol. 22, no. 12, pp. 2677–2684, dec 2010.

[17] S. Murala and Q. M. J. Wu, "Local mesh patterns versus local binary patterns: Biomedical image indexing and retrieval," IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 3, pp. 929–938, may 2014.

[18] M. D. Kumar, M. Babaie, S. Zhu, S. Kalra, and H. R. Tizhoosh, "A comparative study of CNN, BoVW and LBP for classification of histopathological images," nov 2017.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems, vol. 25, pp. 1097–1105, 2012.

[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[21] J. Wang, P. Liu, M. F. She, S. Nahavandi, and A. Kouzani, "Bag-of-words representation for biomedical time series classification," Biomedical Signal Processing and Control, vol. 8, no. 6, pp. 634–644, nov 2013.