International Journal of Advanced and Applied Sciences, 9(12) 2022, Pages: 11-24

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Pre-trained convolution neural networks models for content-based medical image retrieval

Ali Ahmed [1,*], Alaa Omran Almagrabi [2], Ahmed Hamza Osman [2]

[1]Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, Saudi Arabia
[2]Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 20 January 2022
Received in revised form
15 May 2022
Accepted 16 August 2022

Keywords:
Pre-trained convolution neural networks
Features extraction
Medical image retrieval

ABSTRACT

Content-based image retrieval (CBIR) is a recent method used to retrieve different types of images from repositories. The traditional content-based medical image retrieval (CBMIR) methods commonly used low-level image representation features extracted from color, texture, and shape image descriptors. Since most of these CBMIR systems depend mainly on the extracted features, the methods used in the feature extraction phase are more important. Features extraction methods, which generate inaccurate features, lead to very poor performance retrieval because of semantic gap widening. Hence, there is high demand for independent domain knowledge features extraction methods, which have automatic learning capabilities from input images. Pre-trained deep convolution neural networks (CNNs), the recent generation of deep learning neural networks, could be used to extract expressive and accurate features. The main advantage of these pre-trained CNNs models is the pre-training process for huge image data of thousands of different classes, and their knowledge after the training process could easily be transferred. There are many successful models of pre-trained CNNs models used in the area of medical image retrieval, image classification, and object recognition. This study utilizes two of the most known pre-trained CNNs models; ResNet18 and SqueezeNet for the offline feature extraction stage. Additionally, the highly accurate features extracted from medical images are used for the CBMIR method of medical image retrieval. This study uses two popular medical image datasets; Kvasir and PH2 to show that the proposed methods have good retrieval results. The retrieval performance evaluation measures of our proposed method have average precision of 97.75% and 83.33% for Kvasir and PH2 medical images respectively, and outperform some of the state-of-the-art methods in this field of study because these pre-trained CNNs have well trained layers among a huge number of image types. Finally, intensive statistical analysis shows that the proposed ResNet18-based retrieval method has the best performance for enhancing both recall and precision measures for both medical images.

1. Introduction

Image retrieval is an important application of image processing; it has many usages such as architectural and engineering design, computer vision, face recognition, and military and medical diagnosis (Kokare et al., 2002). In the past, the first generation of image retrieval was based on the exact matching of text description with an image. This method was text-based image retrieval (TBIR), which had very poor performance. Two of the most famous earlier surveys in this field are found in Tamura and Yokoya (1984), and Chang and Hsu (1992). Most systems that work on this method need to annotate and index all the images and give them descriptive keywords. Therefore, such systems are considered very difficult to use and time-consuming, especially when the groups and classes of images are large. A summarization of the disadvantages of TBIR is found in Abioui et al. (2018). In the late 1990s, due to the shortcomings of TBIR, a new method of content-based image retrieval (CBIR) was developed to solve the previous limitations (Rui et al., 1999),

* Corresponding Author.
Email Address: aabdelrahim@kau.edu.sa (A. Ahmed)
https://doi.org/10.21833/ijaas.2022.12.002
Corresponding author's ORCID profile:
https://orcid.org/0000-0002-8944-8922

the latest method used in retrieving images from image repositories and websites reported in Smeulders et al. (2000), and Marques and Furht (2002). The main advantage of CBIR systems is their ability to represent images by their features or attributes extracted from different image descriptors such as color, texture, and shape. This step or process is known as an offline operation because it is performed first and independently of the exact retrieval process. Later, in the online retrieval stage, the similarity between the features or attributes of the query image and the image features stored in the database is measured through similarity coefficients or metrics. The basic idea of any CBIR system is to retrieve the most similar image according to the user query image in descending order, with the most likely images retrieved first. The primary key to the success of any content-based image retrieval system is the extraction of features from images in a particular domain. This stage is very important and results in numerical feature values stored separately in repositories or data storage to be used later for similarity computation (Liu et al., 2007). Due to their importance, researchers in the field of multimedia in general, and medical image retrieval in particular, have conducted extensive studies of methods of representation of features or characteristics and methods of measuring similarity, which mainly affects the accuracy and efficiency of content-based image retrieval systems (Ahmed, 2020; Ahmed and Malebary, 2020). Jiang and Kim (2021) enhanced the performance retrieval of CBIR by feature fusion and the discrete cosine transform method. Many studies found that the semantic gap between the features or characteristics of the images captured by imaging machines and the high-level features humans perceive is still large despite the broad research in this field. In the past two decades, a number of content-based image retrieval methods have been proposed to reduce the semantic gap. CBIR methods based on low-level features such as color, shape, and texture are explored in many studies such as Veltkamp and Tanase (2002), Müller et al. (2004), Oussalah (2008), Rajam and Valli (2013), and Bhagyalakshmi and Vijayachamundeeswan (2014). Colour descriptors are considered important sources for features used in CBIR systems. There are many color descriptors for content-based image retrieval that have been used including color histograms (Zhang et al., 2009), color moments, and colour coherent vector (CCV) (Kodituwakku and Selvarajah, 2004). Colour moments are widely used in content-based image retrieval and image classification (Ahmed and Sadig, 2019; Ahmed and Malebary, 2019; Ibrahim et al., 2019; Mohammed et al., 2020). A color histogram is defined as the distribution of the number of pixels for an image. It represents how image pixels are distributed by plotting the number of pixels at all levels of color intensity. The histogram shows details in shadows, mid-tones, and highlights, and can help determine whether the images are detailed enough to make a good correction. Colour moments are a group of functions used to extract a single numeric value from each image based on pixel density value; for example, mean, variance, and standard deviation. The main disadvantage of the color histogram method is that it does not consider the spatial information of pixels. Therefore, different images may have the same histogram distribution. To solve such limitation, a colour coherent vector (CCV) is proposed. In this method, each histogram bin is categorized into two main types: Coherent and incoherent. Coherent type grouped pixel values belong to the large informally colored region, while the incoherent category contains the other pixel values, comparison of different color features are found in Kodituwakku and Selvarajah (2004). The recent implementation of a local binary pattern (LBP) was successful for feature extraction, which was used for content-based image retrieval. Different versions of the local binary pattern (LBP) are found in Win et al. (2020), Anitha and Naresh (2021), and Ghahremani et al. (2021). Also indexing approach GPU-accelerated feature extraction-based systems were successfully implemented by Farruggia et al. (2014), Tsai et al. (2017), and Rundo et al. (2019). Some indirect methods could positively affect the retrieval process of medical images found in the literature. Unsupervised medical anomaly detection generative adversarial network (MADGAN) for MRI medical images was proposed by Han et al. (2021), and another similar approach SteGANomaly was proposed by Baur et al. (2020). Consideration of the hidden information detected by anomaly detection methods in the two previous studies could enhance and affect the retrieval process. All the traditional feature extraction methods mentioned above fail to reduce the semantic gaps for CBIR. This is due to inaccurate features resulting from the absence of domain or field-specific knowledge. These limitations lead to the exploration and use of pre-trained convolution neural network models, which are considered the latest methods for this purpose (Bhandi and Devi, 2019). There are many available models of pre-trained CNNs models such as ResNet18, SqueezeNet AlexNet, and GoogleNet. The main reason for the wide acceptance of the success of these models are least training time required in addition to their ability for learning transfer (Pham, 2021). The architectures and specifications of training parameters for transfer learning were found in Krizhevsky et al. (2012), Szegedy et al. (2015), and Iandola et al. (2016). Further discussion about the architectures and layers of these models is found in Gopalakrishnan et al. (2017). In the present study, we utilize the first two models as feature extraction tools for the first part of content-based medical image retrieval CBMIR. The rest of this paper is organized as follows. Section II provides more details about CBIR, semantic-based Image Retrieval (SBIR), and deep learning methods and their use in the medical and diagnosis-based domain. Section III explains the related works for the study. Section IV describes the proposed methodology. Section V presents the experimental results and summarised results and discussion. Finally, Section VI contains the conclusion and outlines future work.

2. CBIR, SBIR, and deep learning for medical image retrieval

In the last two decades, the development of several computing and advanced medical imaging devices as well as large-scale storage media has resulted in large repositories of medical image content and other media. Clinical and diagnostic studies and various medical research centers benefit from this medical content. Therefore, developing a high-performance medical image retrieval system is imperative, to support medical specialists in their studies, in addition to retrieving medical images for various purposes such as diagnostics. To facilitate the effective use of the many medical images in different and available databases, several algorithms for automatic analysis of medical images are proposed in the literature (Quellec et al., 2011; Rahman et al., 2011; Mizotin et al., 2012; Ponciano-Silva et al., 2013; Jiji and Raj, 2015). Content-based medical image retrieval (CBMIR) system, which is an extension of CBIR systems, can be considered an effective tool to assist in the diagnosis of various diseases (Zhang et al., 2016). Although many researchers have studied CBIR extensively, the semantic gap remains the biggest challenge and the most complex problem in CBIR systems (Wan et al., 2014). Recently, it was concluded that this problem can be solved and the gap minimized based on two solutions, which take into consideration part of the domain-specific knowledge and apply some machine learning methods to produce a more advanced and intelligent model. This could be trained and used as an alternative method to extract features for calculating similarity (Nair et al., 2021). The traditional CBIR system (also known as instance-based CBIR system) usually has better retrieval outcomes and good performance results when a database contains relatively fewer images. As the development of imaging techniques and the size of databases increase, these methods and systems have very poor retrieval performance. To address such a problem, researchers in this field introduced the semantic-based image retrieval (SBIR) system (Li et al., 2006; Pourghassem and Ghassemian, 2008; Guo et al., 2018), which has also the main disadvantage of miss-classification because it depends mainly on classifier use, and works in binary form. This is considered a rigid model with poor retrieval results. The second-generation development in CBIR-based systems to overcome the previous limitations of SBIR are machine learning-based systems, which have limited success and relatively good performance such as in Safavian and Landgrebe (1991), Suykens and Vandewalle (1999), Wang et al. (2008), and Huang et al. (2016). The major problem of machine learning ML approaches is that it has separate feature extraction stages using special feature extraction functions. The investigation and development of deep learning DL methods are one of the most important explorations in the area of AI. Deep learning methods and techniques provide more accurate ways to extract image features through the use of deep neural networks, and of their ability to learn better and obtain more accurate features than those extracted through traditional methods, which depend on the image descriptors of color, shape, and texture. Unlike the machine learning ML-based CBIR system and all its previous approaches to CBIR systems, deep learning DL-based CBIR systems do not require manual feature extraction. This benefit makes the DL approach the most powerful and effective method used for CBIR in many application areas in general and the medical field in particular. The most common deep learning DL-based CBIR systems used recently depended on VGG-16 (Simonyan and Zisserman, 2014), inceptions V3 (Szegedy et al., 2016), AlexNet (Krizhevsky et al., 2012), and ResNet V2 (Szegedy et al., 2017). The second benefit that a deep learning DL provides for successful use with medical images is that the visual features, which must be extracted from the medical images, should be unique and differ from those of the normal images in traditional CBIR systems. Thus, direct use of all CBIR, SBIR, and pre-deep learning DL technologies for medical image retrieval applications usually produces unsatisfactory and high false-rated retrieval results.

To address this problem, researchers have used different DL models (CNN models) to retrieve medical images (Sundararajan et al., 2019; Yadav and Jadhav, 2019; Kalaivani et al., 2020). A more recent survey on this topic is found by Ghahremani et al. (2021). The wide range of usage supports the hypothesis that DL techniques can succeed. Yadav and Jadhav (2019) successfully applied a deep convolutional neural network to classify pneumonia for chest X-ray images. In their study, they conclude that transfer learning approaches of VGG16 and InceptionV3 are more useful classification methods compared with SVM classifier with oriented fast and rotated binary robust independent elementary features (ORB) and capsule neural network. Deep learning DL methods are used also for extracting powerful features from DICOM images. In general, deep learning methods could be used for extracting powerful DICOM image features from both metadata and pictorial content; however, Kalaivani et al. (2020) focused on pictorial content only. The authors of this study first predict the retrieved images from a given query input image-based trained network, and then the related and similar images are retrieved from the image repository. This is a more useful idea leading to good retrieval performance compared with traditional neural networks, where a set of image features is used as input for the classification stage, and if some features are missed or inaccurate, then the retrieval system will be ineffective and produce poor results. In general, pre-trained CNNs models, the recent architecture of deep learning DL methods, could be used and applied to the classification or retrieval of medical images in two ways (Pham, 2020): Firstly, for feature extraction, as seen in previous studies (Yadav and Jadhav, 2019; Kalaivani et al., 2020); and, secondly, to use their architectures as in classification and retrieval to take advantage of their powerful training and transfer knowledge capabilities. This study utilizes and compares two of the most used pre-trained CNNs models; ResNet18 (He et al., 2016) and SqueezeNet (Gopalakrishnan et al., 2017) to be used as feature extraction tools for content-based medical images retrieval methods. The general framework of the proposed retrieval method in this study is shown in Fig. 1, and the main contribution of the paper is summarised as follows:

a) Enhance the retrieval performance of CBMIR methods by replacing the traditional mathematical methods with pre-trained feature extraction methods.
b) Develop an efficient feature extraction method based on medical images using recent and effective pre-trained CNN models: ResNet-18 and SqueezeNet.
c) Develop CBMIR methods, which have high retrieval performance with the assistance of accurate extracted features and effective Euclidean distance similarity measures.

The most important component or element in this retrieval method in Fig. 1 is the pre-trained CNNs module, which is used for extracting both database images and query images during two stages of offline and online processes. The rest of the proposed framework is quite similar to other general CBIR systems, which have some parts related to noise removal or any type of feedback process, in order to increase the retrieval performance. In this model, the main focus is the use of pre-trained CNNs, which are more efficient and accurate because this is based on the transfer learning method, and it was trained on a large scale and with a huge number of image classes.

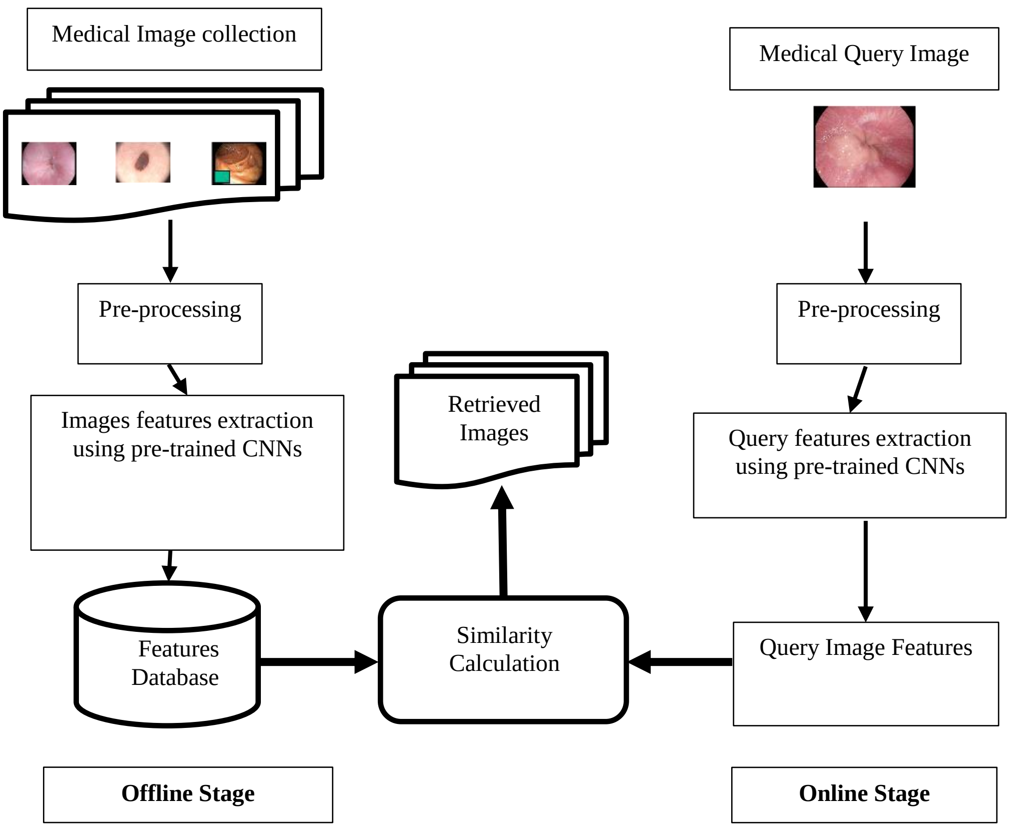


Fig. 1: Main framework of proposed CBMIR

3. Literature survey

Qayyum et al. (2017) used abstract deep convolution neural networks DCNN for large groups of images classified and labeled into various classes. Firstly, DCNN is trained for the classification of collected medical images, and then the learned feature representations in the form of numeric value vectors are used for CBMIR in the second phase. Sundararajan et al. (2019) performed a necessary pre-processing operation on Avascular Necrosis affected bone images, and then they use a deep belief convolutional neural network (DB-CNN) to represent the images' features and translated them into binary codes. Finally, a similarity measure is computed for the retrieval process. As in Ayyachamy et al. (2019), the ResNet18 deep CNN model is utilized for CBMIR for four different types of medical images: CT, MRI, MG, and PET. The authors of the study obtain the maximum power of the pre-trained ResNet18, the most well-known deep CNNs model. Recent review studies on descriptive frameworks and in-depth methods use indexing and retrieval of medical images for CBMIR (Das and Neelima, 2017; Tian and Fu, 2020). Some authors such as Khatami et al. (2018) argued that transformation processes for input images before feeding the CNNs are useful and produce significant retrieval results. In their study, they use the radon-based transform process for medical images that feed the CNNs. Their proposed method significantly increases the retrieval performance. Deepak and Ameer (2020) used GoogleNet with transfer learning for feature extraction, and then the Siamese Neural Network (SNN) is utilized with contrastive loss function for feature representation. At the similarity measure stage, Euclidean is applied in the lower dimensional feature space. Furthermore, SqueezeNet, a commonly used pre-trained deep CNN, is combined with Bayes optimization for the development of an AI model for COVID-19 image detection and retrieval (Ucar and Korkmaz, 2020).

Besides the range of successful usage of the deep CNN models, three main challenges remain for the implementation of these tools in the medical field as reported by Godasu et al. (2020). Godasu et al. (2020) reviewed and summarize the limitations into three main points: Over-parameterisation due to a large pool and parameters and different options of settings; high and expensive computation needs for experimentations; and insufficient availability of a well-classified and labeled medical dataset in this field of study.

4. Proposed methodology

The methodology of this study is based on the analysis and comparison of retrieval methods for three different approaches. The first retrieval method depends on the traditional approach that was used to retrieve medical images. The second and third methods depend on the use of previously pre-trained deep learning neural network tools to extract the features of medical images and then use similarity coefficients to calculate the similarity between the query images and those in the image database. In the first traditional approach, color and texture features were extracted from medical images, while the two modern pre-trained deep learning neural network tools (ReseNet18 and SqueezeNet) were used in the second and third methods, respectively.

The process of extracting image features is one of the most important stages for any content-based retrieval system. Therefore, the main focus of the proposed method is to use pre-trained CNNs tools because of their ability to extract accurate features, which leads to increasing the efficiency of retrieval and reducing the semantic gap. Pre-trained CNNs deep learning neural network (CNNs) tools are used to extract the features of the medical images in the offline phase, while the features of the query image are extracted in the online phase when the user submits the required query image as shown in Fig. 1.

Fig. 1 also shows the storing of the extracted features from all medical images in separate databases to be used in the second stage instead of the medical images themselves. This process is very important, and it helps to speed up the final search process. In the following sections, the three approaches are explained in detail.There are many successful models of pre-trained CNNs that are used for both classification and retrieval problems in the medical domain. Among these choices, SqueezeNet and ResNet18 are acceptable due to their simple architectures, the number of processing elements, and the feasible number of extracted features that do not need any type of dimension reduction. Using the exact architecture of pre-trained CNN models for transfer learning without major changes to their parameters is considered the simplest and most cost-effective method that results in acceptable retrieval performance due to the powerful and high capabilities of transfer learning. In this study, we do a more accurate and extensive way by choosing the percentage of 80% for training and 20% for testing. In each run non-repeated, 20% of testing images are chosen randomly and their features were saved into a separate database. After repeating this process five times we got the image features for all images in the dataset (each image represented by 512 features). This process was too similar to the cross-validation and resulted in accurate features that led to good retrieval performance.

4.1. Retrieval method based on traditional color and texture features

As mentioned earlier, feature extraction is one of the most important stages of any CBMIR system, and it plays a significant role in enhancing and increasing the retrieval performance, whether extracted through some mathematical functions as in traditional CBMIR systems or through deep neural network CNNs tools as in modern methods. Most of the old traditional CBMIR approaches depend on the properties of color and texture as the main descriptors for extracting the images' features. In this traditional model, which is used for comparison, we also extract color and texture features to represent all the images. All colored medical images are first converted from RGB representation to HSV representation because this color model is the best and most accurate representation of color images. From each channel of H, S, and V, six mathematical functions are used to extract six values from each channel; in total, 18 features are extracted from the color descriptor.

The six functions used here are provided by Maheshwary and Srivastava (2009) and are described in Eqs. 1-6. These color moments functions are the most well-known and are successfully used in many previous studies such as Ahmed (2020), and Ahmed and Malebary (2020). For any image in the database with the dimension of (M, N) let Vij be the density value of the pixel at the i^{th} and j^{th} column. Then, color moments functions used for feature extraction are defined as follows:

$$Mean\ m = \frac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N}V_{ij} \tag{1}$$

$$Variance\ v = \frac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N}(V_{ij}-m)^2 \tag{2}$$

$$Contrast = \frac{1}{M*N}(\sum_{i=1}^{M}\sum_{j=1}^{N}(V_{ij} - m)^2)^{1/2} \quad (3)$$
$$Skew = \frac{1}{M*N}(\sum_{i=1}^{M}\sum_{j=1}^{N}(V_{ij} - m)^3)^{1/3} \quad (4)$$
$$kurtosis = \frac{1}{M*N}(\sum_{i=1}^{M}\sum_{j=1}^{N}(V_{ij} - m)^4)^{1/4} \quad (5)$$
$$Smoothness = 1 - \frac{1}{(1+v)} \quad (6)$$

For texture features in this study, the well-known Grey-level co-occurrence matrix GLCM method proposed by Weszka et al. (1976) is used. The GLCM method analyses the grey levels for all pixel values of images after converting them from RGB color space to HSV model. Herefour mathematical functions are also used to extract four feature values from each channel of H, S , and V results, in total, 12 features. Finally, these features are combined and fused together with the 18 color features into a single vector and stored in the separated database repository. The following Eqs. 7-10 describe the functions used for texture properties. According to the GLCM method, for any image in the image database let P(i, j)=(P(i, j))/R is (i,j)th entry in the normalized matrix for unique grey levels of the quantized image (R is maximum gray value), Ng is the number of distinct grey levels of the quantized image and calculate μ and σ as the mean and standard deviation respectively. The functions used for texture feature extraction are shown in the following equations:

$$uniformity = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng}\{P(i,j)\}^2 \quad (7)$$
$$correlation = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng}\frac{(i-\mu)(j-\mu)p(i,j)}{\sigma^2} \quad (8)$$
$$contrast = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng}P(i,j)(i-j)^2 \quad (9)$$
$$entropy = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng}P(i,j)\,log\{P(i,j)\} \quad (10)$$

4.2. Retrieval method based on SqueezeNet model features

The SqueezeNet is one of the simplest and most efficient pre-trained CNNs tools, which can transfer learning capability. This model was developed by Gopalakrishnan et al. (2017) and it consists of 68 layers. The size of all input images must be 227×227×3. Before feeding the tool on this model with medical images for training and extracting the required features, it must resize according to the required size. Then, the input images are manipulated through 14 processing elements, whereby the model performs various convolution operations for all sides of the image to extract features with high accuracy. This model is shown in more detail in Fig. 2a. Finally, a total of 1000 features are extracted and saved, and stored to use later in the similarity measurement stage.

4.3. Retrieval method based on ResNet18 model features

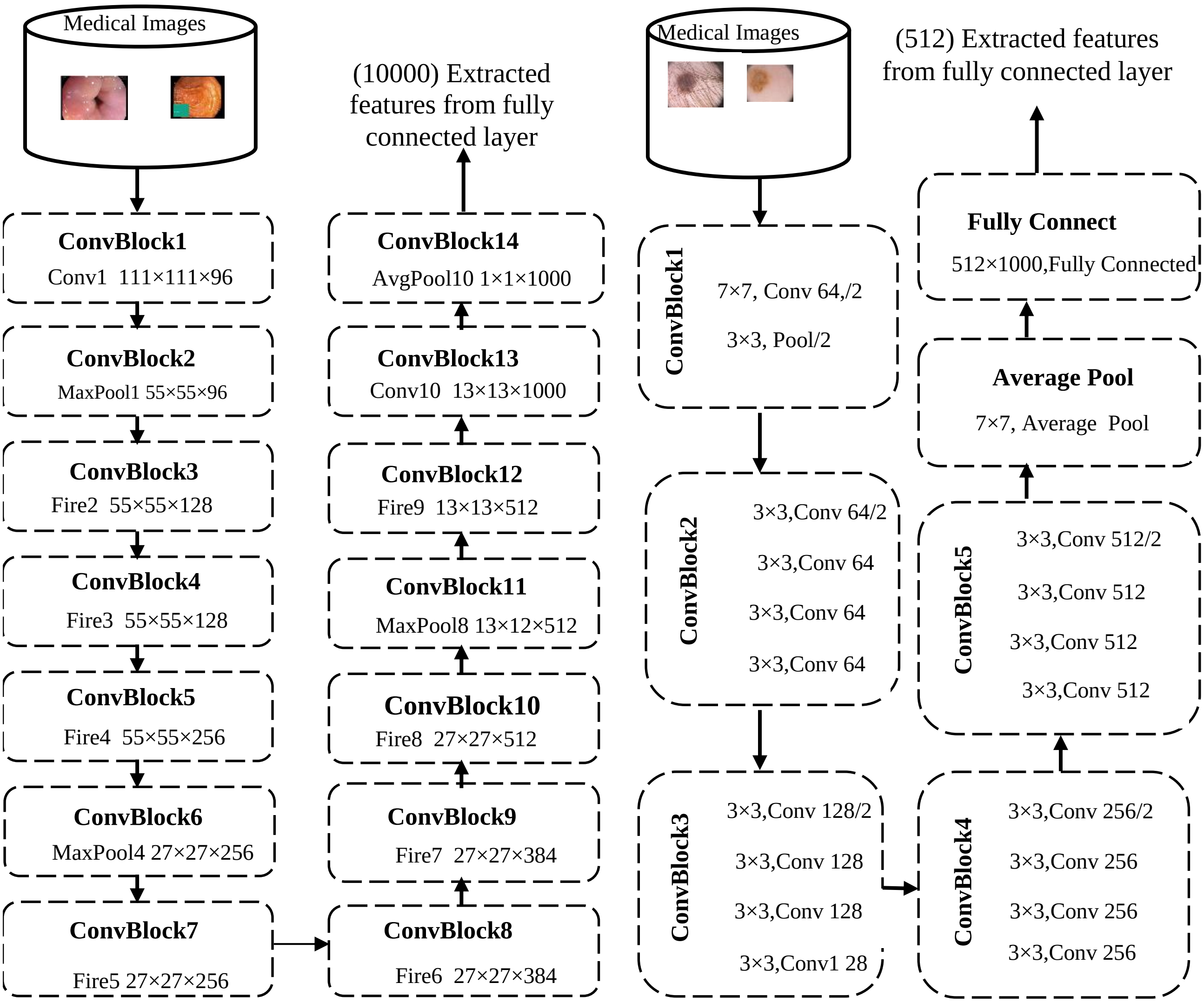
Our third retrieval model is based on the ResNet18 model, which is also a very efficient pre-trained CNNs model for feature extraction. The model was developed by Zhang et al. (2017), and the customized model for extracting medical image features for this study is shown in Fig. 2b. This model consists of 18 layers and five convolutional processing elements in addition to the last layer. After performing the required convolution operations, as shown in Fig. 2b, a set of 512 features was extracted and saved in the numerical database. This model of the pre-trained CNNs model and the previous model has the ability to transfer learning, as a result of the extensive training on the ImageNet database (Deng et al., 2009), which contains millions of images with hundreds of classes. The extensive training makes these models an optimal choice for feature extraction of medical images after performing the required tuning for the different parameters (Pham, 2020).

The successful transfer learning capability is not limited to the image retrieval area, but it has also been extended to support various areas of research in the medical image field such as image classification and clustering, object recognition, and image segmentation. In our previous study (Ahmed, 2021), we applied a data augmentation process which made the retrieval model very slow and required long computation times; however, in this study, only rescaling and resizing processes required by the input layer were applied which made our proposed model very simple with the minimum number of iterations and better performance with medical datasets.

4.4. Medical image datasets

The first dataset used in this study is Kvasir (Pogorelov et al., 2017), one of the most popular datasets in the area of medical imaging. It consists of 4000 images, annotated and verified by medical doctors (experienced endoscopists), and divided into eight classes (500 images for each class) showing anatomical landmarks, pathological findings, or endoscopic procedures in the GI tract. The number of images is suitable and sufficient to be used for different medical tasks in the medical field such as image retrieval, machine learning, and deep learning and transfer learning (Donahue et al., 2014; Quattoni et al., 2008; Abadi et al., 2016).

The anatomical landmarks are Z-line, pylorus, and cecum, while the pathological findings include esophagitis, polyps, and ulcerative colitis. All images are RGB images with different resolutions from 720×576 to 1920×1072 pixels, recent use of this dataset was found by Mahmud et al. (2021), and Yeung et al. (2021). The second dataset is PH2 images for melanoma detection and diagnosis. This image dataset, which was built through joint research between a group of universities and hospitals in Portugal (Mendonça et al., 2013), has a total of 200 RGB dermoscopic images divided into three classes: 80 common nevi, 80 atypical nevi, and 40 melanomas. Samples of images from the two datasets are shown in Fig. 3a and Fig. 3b for Kvasir and PH2 medical images respectively. For resizing process, first, all images are resized to match the input layer requirement of pre-trained model 224×224×3 for ResNet18 and 227×227×3 for SqueezeNet, and then all images are converted from RGB to gray-scale images with the same size as 224×224 and 224×224, for both models respectively. Rescaling is done implicitly by the convolution step and pixel values were changed according to the convolution process.



a b

Fig. 2: (a) SqueezeNet model for feature extraction; (b) ResNet18 model for feature extraction

4.5. Similarity measures

The second stage in any CBMIR system is similarity computation, which is considered an important online retrieval process because it results in a pool of similar images according to the user query image. In this study, Euclidean distance is used for the similarity measures. This simple similarity coefficient has good performance and is an efficient, standard, and well-known similarity measure used in many studies (Shirkhorshidi et al., 2015). For any vector Q of a user query image and any vector D from the stored database of numeric features values, the following equation calculates the Euclidean distance measure:

$$similarity\ score = \sqrt{\sum_{i=1}^{n}(Q_i - D_i)^2} \tag{11}$$

where, n is the number of dimensions of the Q and D vectors; the lowest value for this measure means the most similar images.

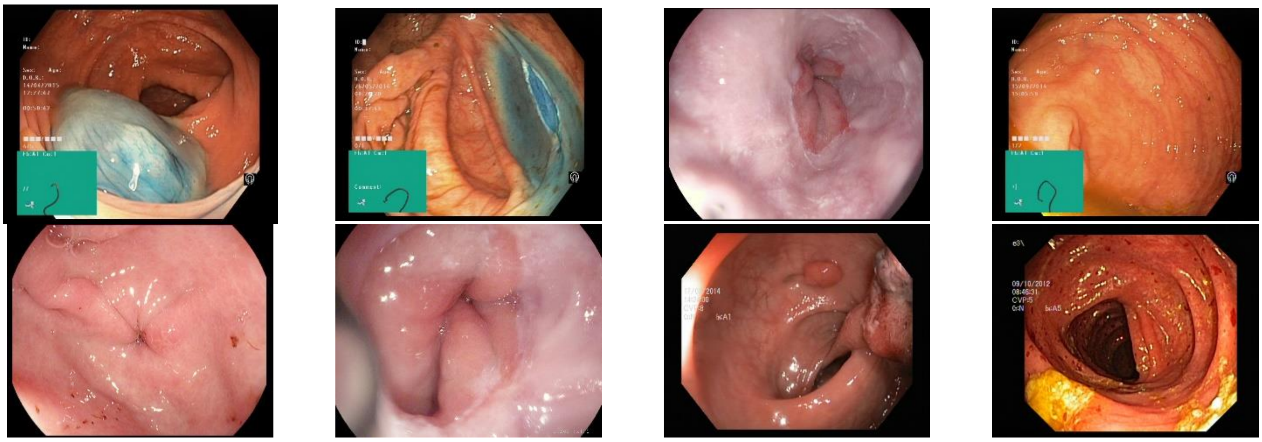
4.6. Performance evaluation

Recall and precision are two common evaluation performance metrics in the area of information retrieval in general and CBMIR in particular. In this study, these two metrics are used and their values calculated at different top-cut values from 5 to 100 retrieval images. The formula or equations of these metrics is shown below:

$$Precision = \frac{number\ of\ relevant\ images\ retrieved}{total\ number\ of\ retrieved\ images} \tag{12}$$

$$Recall = \frac{number\ of\ relevant\ images\ retrieved}{total\ number\ of\ relevant\ images} \tag{13}$$

All experiments are implemented using the MATLAB license version running on the Windows OS platform.



a

b

Fig. 3: (a) Image sample for Kvasir dataset; (b) Image sample for PH2 dataset

4.7. Based methods for comparison

For comparative purposes, this study uses two similar methods from the literature. The first method, as in Hu et al. (2021), modifies ResNet-18 to generate binary hash codes for Kvasir image retrieval, while the second method (Kasban and Salama, 2019) is based on wavelet optimization and adaptive block truncation coding. These two studies are recent and more related to our proposed method.

5. Results and discussion

In this study, the three retrieval methods of CBMIR systems are developed and evaluated. The first colour and texture method (CL and GLCM) is based on the use of traditional features extracted from colour and texture descriptors using well-known mathematical functions used previously in this field of study. As mentioned, this method is the basis upon which the results of our proposed two methods are compared. The second and third methods are SqueezeNet-based method and ResNet18-based, which depend on the use of pre-trained CNNs models to extract more accurate features to increase the retrieval efficiency. For each of the three experiments, 10 random images from each class re used as user queries, and then the average accuracy is calculated for each class. Finally, the overall average is calculated for all the classes for two medical image datasets.The overall retrieval results of the three retrieval methods for the two medical images from the two datasets are shown in Table 1. In Table 1, the average recall and precision for each class are shown, as the average precision for all classes of the three methods. It is clear that the retrieval precision of the ResNet18-based method is superior to the traditional retrieval method by 27.4% and 3.07% for the two datasets, respectively, and the SqueezeNet-based method is better than the traditional method. This supports the hypothesis of the more accurate features extracted based on these pre-trained models. The same observation could conclude from Table 2 which shows that both our proposed retrieval methods based on pre-trained CNNs models outperform the traditional method, and again the results of ResNet18-based method are better. All the results in Table 1 and Table 2 are calculated for the top 10 retrieval images as in many related studies. However, average precisions values for all classes of the two images datasets at different cut-off values (from 100 images to 5 images) are shown in Fig. 4a and Fig. 4b for Kvasir and PH2. These two figures show the relation of number of images retrieved and average precisions values. The upper red lines in both Figs. 4a and 4b prove that the ResNet18-based method has best retrieval performance. The effectiveness and efficiency of the proposed methods are also proven by using Kendall W concordance test (Sidney, 1988), one of the most important analytical tests. The test is applied in the present study; the medical images classes represent judges, while the average precision objects are considered judges and the average precisions of the three different methods are considered objects. The average precisions values are an input for this test, and the output are the Kendall coefficient and the associated level of significance. The result of this test is shown in Table 3 for the two medical images. The agreement or confidence is 75.5% for the Kvasir-1K dataset and 93.9% for the PH2 dataset. The ranking of the different methods in the corresponding rows in Table 3 shows that the ResNet18-based method and SqueezeNet-based method outperform the traditional CL and GLCM based method since their ranks are better. ResNet18-based method is also the best.

An important finding is shown in the samples of top retrieved images of some classes for two medical images in Fig. 5 and Fig. 6, in which our proposed two methods have success in retrieving all class images in the top 10 images in more than one class for Kvasir medical images and one of the three classes of PH2 medical images. Finally, Additional statistical certainty in terms of the mean, lower and upper bounds of the confidence intervals of the three retrieval methods is shown in Fig. 7a and Fig. 7b for the two medical images groups. Both figures indicate that the proposed methods are much better with obvious improvement.

Table 1: Recall and precision of Kvasir images

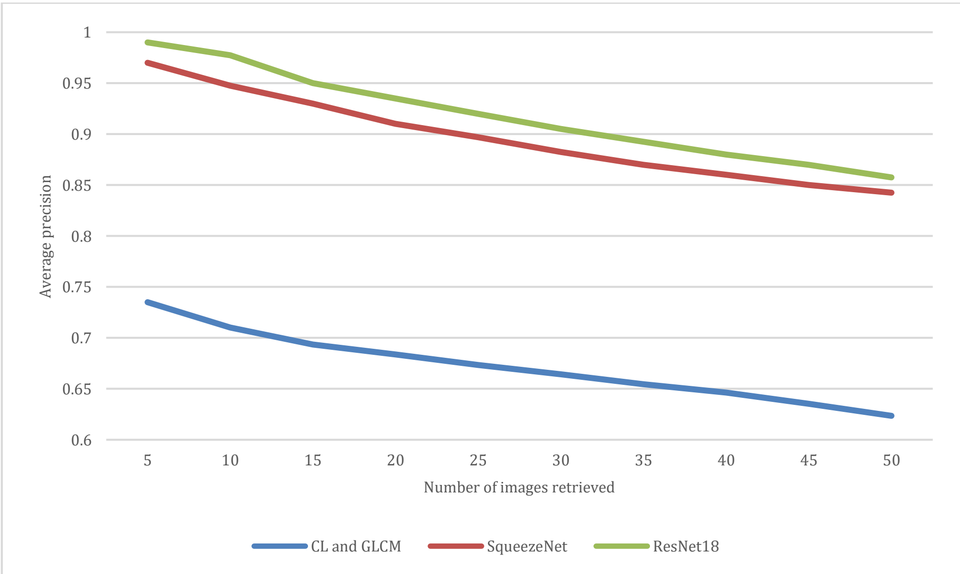
Classes	CL and GLCM		SqueezeNet		ResNet18	
	R	P	R	P	R	P
DLP	0.0156	0.7800	0.0196	0.9800	0.0192	0.9600
DRM	0.0128	0.6400	0.0188	0.9400	0.0200	1.0000
Esophagitis	0.0152	0.7600	0.0180	0.9000	0.0200	1.0000
Normal Caecum	0.0160	0.8000	0.0200	1.0000	0.0200	1.0000
Normal Pylorus	0.0092	0.4600	0.0192	0.9600	0.0196	0.9800
Normal Z Line	0.0200	1.0000	0.0200	1.0000	0.0200	1.0000
Polyps	0.0136	0.6800	0.0184	0.9200	0.0188	0.9400
Ulcerative Colitis	0.0112	0.5600	0.0176	0.8800	0.0188	0.9400
Average	0.0142	**0.7100**	0.01895	**0.9475**	0.01955	**0.9775**
Method1 (Hu et al., 2021)						0.9390
Method2 (Kasban and Salama, 2019)						0.6120

Table 2: Recall and precision of PH2 images

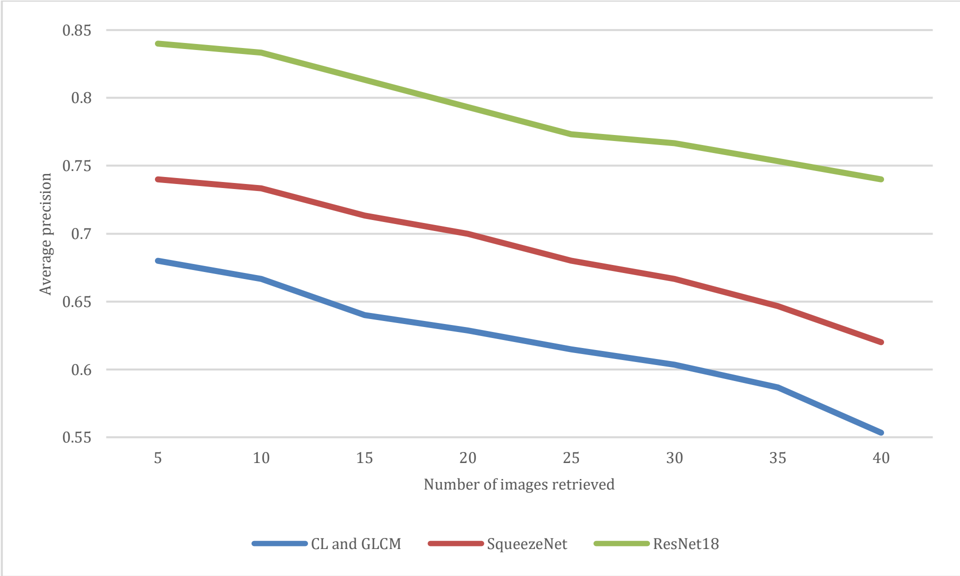
Classes	CL and GLCM		SqueezeNet		ResNet18	
	R	P	R	P	R	P
Normal	0.200	0.800	0.200	0.800	0.225	0.900
Atypical Nevus	0.150	0.600	0.175	0.700	0.200	0.800
Melanoma	0.150	0.600	0.175	0.700	0.200	0.800
Average	0.1667	**0.6667**	0.18333	**0.7333**	0.2083	**0.8333**

Table 3: Recall ranking of retrieval methods based on the Kendall W test for precision values

Data set	W	P	Ranking of Retrieval Methods
Kvasir	0.755	0.002	ResNet18 > SqueezeNet > CLR and GLCM
PH2	0.939	0.060	ResNet18 > SqueezeNet > CLR and GLCM



a



b

Fig. 4: Average precision at different top images for (a) Kvasir; (b) and PH2

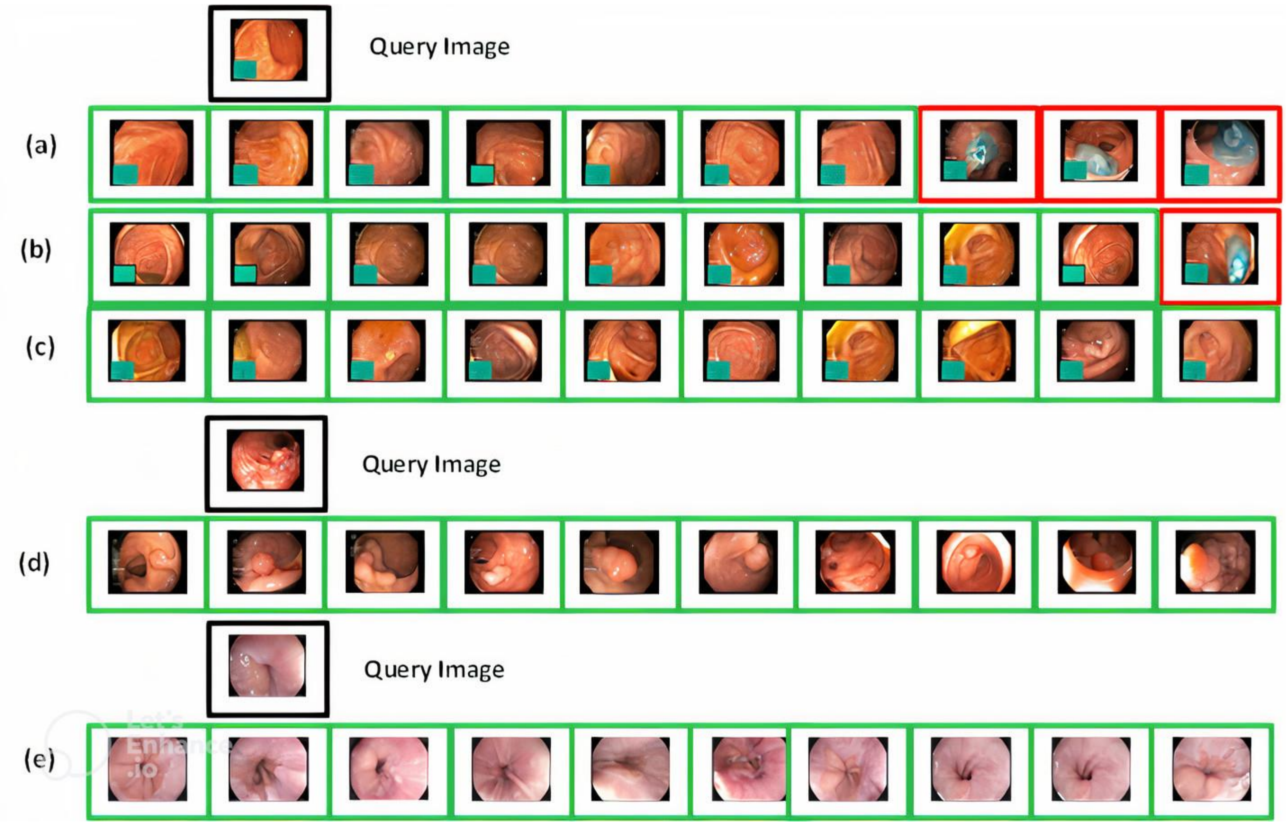


Fig. 5: Kvasir Image samples retrieved: (a) Normal Caecum class retrieved by CLR and GLCM; (b) Normal Caecum class retrieved by SqueezeNet; (c) Normal Caecum class retrieved by ResNet18; (d) Polyps class retrieved by ResNet18; (e) Normal Z Line class retrieved by three methods (red colors refer to false retrieved)

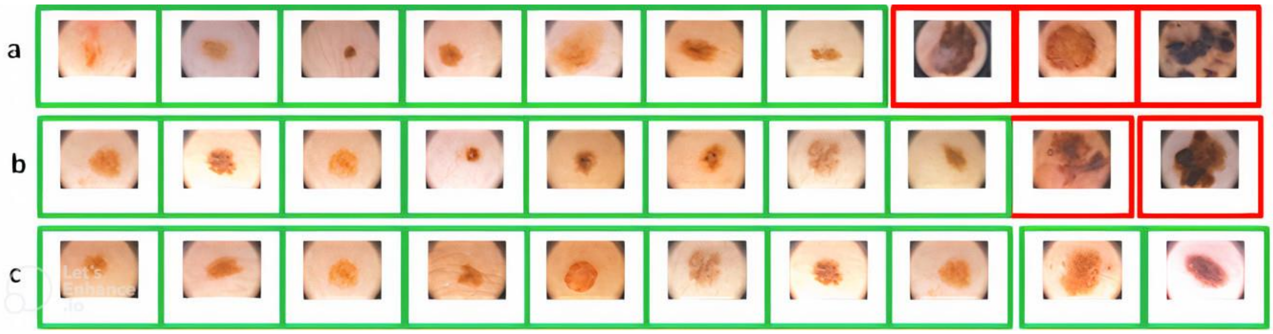
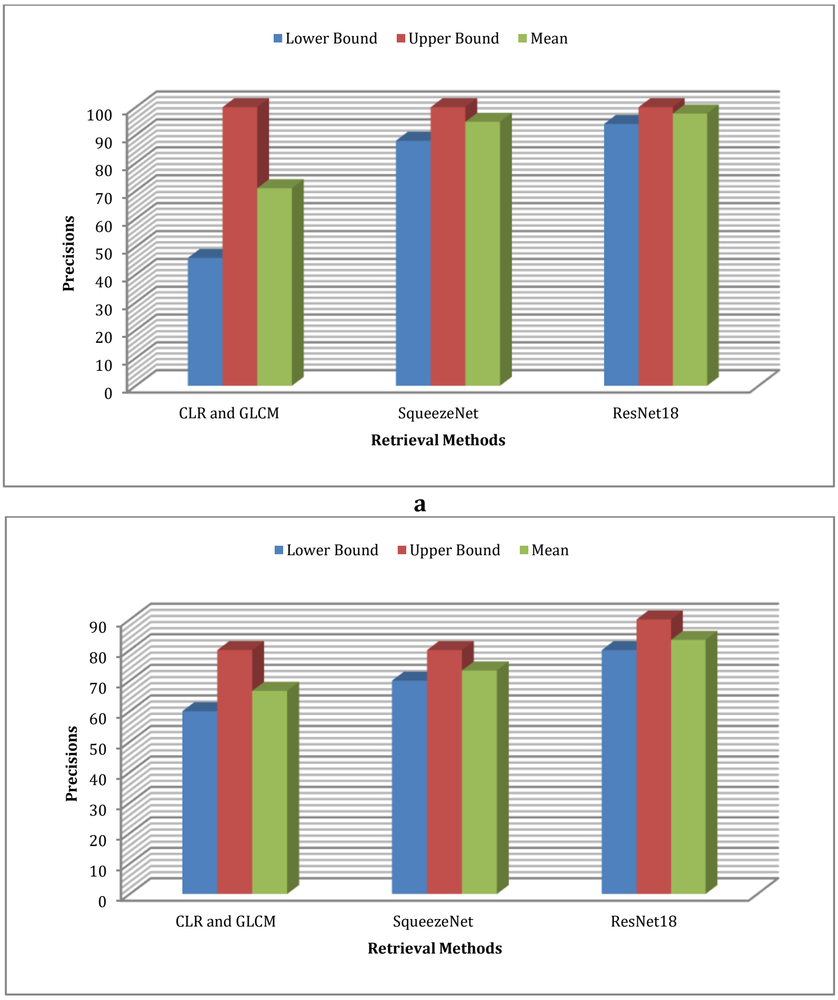


Fig. 6: PH2 Image samples retrieved: (a) Atypical Nevus class retrieved by CLR and GLCM ; (b) Atypical Nevus class retrieved by SqueezeNet; (c) Atypical Nevus class retrieved by ResNet18 (red colors refer to false retrieved)



b

Fig. 7: (a) Confidence bounds for Kvasir images ; (b) Confidence bounds for PH2 images

6. Conclusion

In this study, a pre-trained CNNs-based method was proposed to retrieve medical images using more accurate features than the traditional function-based features, which led to improving the precision of retrieval and reducing the error rate and the false positive percentage of retrieved images. The proposed method was based on the recent generation of pre-trained CNNs models in extracting image features. Two models of these networks were used: SqueezeNet-based retrieval and ResNet18-based retrieval. These two methods have proven high accuracy in retrieving medical images from two well-known and widely used medical image datasets. After implementation and comparing the results with traditional function-based retrieval methods, it is clear that the retrieval precision was improved from 71% to 94.75% using the SqueezeNet retrieval model and also it improved to 97.97% based on ResNet18 for Kvasir images. For PH2 images the achieved enhancement was from 66.67% to 73.33% and 83.33% using SqueezeNet and ResNet18 retrieval models respectively. Both of our proposed methods outperformed the traditional CL and GLCM traditional methods, and ResNet18 has better performance than SqueezeNet for both image datasets. These types of pre-trained CNNs, in addition, to extract highly accurate features, leads to high precision retrieval that can be used also for different types of medical images, because it is pre-trained and has the ability for learning transfer. The focus is currently on reducing computational time as well as dealing with different parameter options. Moreover, many of these network models are currently available, which increases the difficulty of choosing the most appropriate and effective for various types of images. Also, for future efforts for this method, more than one pre-trained CNNs model could be used for the features extraction stage. Then any type of combination approach for similarity measure could use such as the late fusion approach or relevant feedback information. Finally, to address the challenges of this successful and modern type of network, future research could focus on three important issues: Over-parameterization due to a large pool and parameters and the different options of settings, the high computational time required, and insufficient availability of well-labeled and balanced medical images.

Acknowledgment

The authors gratefully acknowledge the Deanship of Scientific Research (DSR) at King Abdulaziz University for their technical and financial support.

This research was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, under Grant No. (G: 146-830-1441093-830-1441).

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, and Ghemawat S (2016). TensorFlow: Large-scale machine learning on heterogeneous distributed systems. ArXiv Preprint ArXiv:1603.0446. https://doi.org/10.48550/arXiv.1603.04467

Abioui H, Idarrou A, Bouzit A, and Mammass D (2018). Automatic image annotation for semantic image retrieval. In the International Conference on Image and Signal Processing, Springer, Cherbourg, France: 129-137. https://doi.org/10.1007/978-3-319-94211-7_15

Ahmed A (2020). Implementing relevance feedback for content-based medical image retrieval. IEEE Access, 8: 79969-79976. https://doi.org/10.1109/ACCESS.2020.2990557

Ahmed A (2021). Pre-trained CNNs Models for content based image retrieval. International Journal of Advanced Computer Science and Applications, 12(7): 200-206. https://doi.org/10.14569/IJACSA.2021.0120723

Ahmed A and El Sadig OB (2019). Heterogeneous multi-classifier method based on weighted voting for breast cancer detection. International Journal on Advanced Science, Engineering and Information Technology, 7(4): 36-41.

Ahmed A and Malebary S (2019). Feature selection and the fusion-based method for enhancing the classification accuracy of SVM for breast cancer detection. International Journal of Computer Science and Network Security, 19(11): 55-60.

Ahmed A and Malebary SJ (2020). Query expansion based on top-ranked images for content-based medical image retrieval. IEEE Access, 8: 194541-194550. https://doi.org/10.1109/ACCESS.2020.3033504

Anitha K and Naresh K (2021). A content-based approach to medical image retrieval. In: Anbarasan K (Ed.), AI innovation in medical imaging diagnostics: 114-136. IGI Global, Hershey, USA. https://doi.org/10.4018/978-1-7998-3092-4.ch007 **PMid:35635523**

Ayyachamy S, Alex V, Khened M, and Krishnamurthi G (2019). Medical image retrieval using ResNet-18. In the Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications, International Society for Optics and Photonics, San Diego, USA, 10954: 1095410.

Baur C, Graf R, Wiestler B, Albarqouni S, and Navab N (2020). SteGANomaly: Inhibiting CycleGAN steganography for unsupervised anomaly detection in brain MRI. In the 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Lima, Peru: 718-727. https://doi.org/10.1007/978-3-030-59713-9_69

Bhagyalakshmi A and Vijayachamundeeswan V (2014). A survey on content based image retrieval using various operators. In the International Conference on Computer Communication and Systems, IEEE, Chennai, India: 18-23. https://doi.org/10.1109/ICCCS.2014.7068160

Bhandi V and Devi KS (2019). Image retrieval by fusion of features from pre-trained deep convolution neural networks. In the 1st International Conference on Advanced Technologies in Intelligent Control, Environment, Computing and Communication Engineering, IEEE, Bangalore, India: 35-40. https://doi.org/10.1109/ICATIECE45860.2019.9063814

Chang SK and Hsu A (1992). Image information systems: Where do we go from here? IEEE Transactions on Knowledge and Data Engineering, 4(5): 431-442. https://doi.org/10.1109/69.166986

Das P and Neelima A (2017). An overview of approaches for content-based medical image retrieval. International Journal of Multimedia Information Retrieval, 6(4): 271-280. https://doi.org/10.1007/s13735-017-0135-x

Deepak S and Ameer PM (2020). Retrieval of brain MRI with tumor using contrastive loss based similarity on GoogLeNet encodings. Computers in Biology and Medicine, 125: 103993. https://doi.org/10.1016/j.compbiomed.2020.103993 **PMid:32980778**

Deng J, Dong W, Socher R, Li LJ, Li K, and Fei-Fei L (2009). ImageNet: A large-scale hierarchical image database. In the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Miami, USA: 248-255. https://doi.org/10.1109/CVPR.2009.5206848

Donahue J, Jia Y, Vinyals O, Hoffman J, Zhang N, Tzeng E, and Darrell T (2014). DeCAF: A deep convolutional activation feature for generic visual recognition. ArXiv Preprint ArXiv:1310.1531. https://doi.org/10.48550/arXiv.1310.1531

Farruggia A, Magro R, and Vitabile S (2014). A text based indexing system for mammographic image retrieval and classification. Future Generation Computer Systems, 37: 243-251. https://doi.org/10.1016/j.future.2014.02.008

Ghahremani M, Ghadiri H, and Hamghalam M (2021). Local features integration for content-based image retrieval based on color, texture, and shape. Multimedia Tools and Applications, 80(18): 28245-28263. https://doi.org/10.1007/s11042-021-10895-z

Godasu R, Zeng D, and Sutrave K (2020). Transfer learning in medical image classification: Challenges and opportunities. In the 15th Annual Conference of the Midwest Association for Information Systems, Des Moines, USA: 1-6.

Gopalakrishnan K, Khaitan SK, Choudhary A, and Agrawal A (2017). Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. Construction and Building Materials, 157: 322-330. https://doi.org/10.1016/j.conbuildmat.2017.09.110

Guo Y, Jia X, and Paull D (2018). Effective sequential classifier training for SVM-based multitemporal remote sensing image classification. IEEE Transactions on Image Processing, 27(6): 3036-3048. https://doi.org/10.1109/TIP.2018.2808767 **PMid:29994442**

Han C, Rundo L, Murao K, Noguchi T, Shimahara Y, Milacski ZÁ, and Satoh SI (2021). MADGAN: Unsupervised medical anomaly detection GAN using multiple adjacent brain MRI slice reconstruction. BMC Bioinformatics, 22(Suppl 2): 31. https://doi.org/10.1186/s12859-020-03936-1 **PMid:33902457 PMCid:PMC8073969**

He K, Zhang X, Ren S, and Sun J (2016). Deep residual learning for image recognition. In the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Las Vegas, USA: 770-778. https://doi.org/10.1109/CVPR.2016.90 **PMid:26180094**

Hu H, Zheng W, Zhang X, Zhang X, Liu J, Hu W, and Si J (2021). Content-based gastric image retrieval using convolutional neural networks. International Journal of Imaging Systems and Technology, 31(1): 439-449. https://doi.org/10.1002/ima.22470

Huang K, Li S, Kang X, and Fang L (2016). Spectral–spatial hyperspectral image classification based on KNN. Sensing and Imaging, 17: 1. https://doi.org/10.1007/s11220-015-0126-z

Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, and Keutzer K (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. In the 5th International Conference on Learning Representations, Toulon, France: 1-13.

Ibrahim AO, Ahmed A, Abdu A, Abd-alaziz R, Alobeed MA, Saleh AY, and Elsafi A (2019). Classification of mammogram images using radial basis function neural network. In the International Conference of Reliable Information and Communication Technology, Springer, Johor, Malaysia: 311-320. https://doi.org/10.1007/978-3-030-33582-3_30

Jiang D and Kim J (2021). Image retrieval method based on image feature fusion and discrete cosine transform. Applied Sciences, 11(12): 5701. https://doi.org/10.3390/app11125701

Jiji GW and Raj PSJD (2015). Content-based image retrieval in dermatology using intelligent technique. IET Image Processing, 9(4): 306-317. https://doi.org/10.1049/iet-ipr.2013.0501

Kalaivani N, Devi D, Sophia S, Manimaran N, and Kowsalya P (2020). Medical image retrieval using convolution neural networks. IOP Conference Series: Materials Science and Engineering, 994(1): 012038. https://doi.org/10.1088/1757-899X/994/1/012038

Kasban H and Salama DH (2019). A robust medical image retrieval system based on wavelet optimization and adaptive block truncation coding. Multimedia Tools and Applications, 78(24): 35211-35236. https://doi.org/10.1007/s11042-019-08100-3

Khatami A, Babaie M, Tizhoosh HR, Nazari A, Khosravi A, and Nahavandi S (2018). A radon-based convolutional neural network for medical image retrieval. International Journal of Engineering, 31(6): 910-915.

Kodituwakku SR and Selvarajah S (2004). Comparison of color features for image retrieval. Indian Journal of Computer Science and Engineering, 1(3): 207-211.

Kokare M, Chatterji BN, and Biswas PK (2002). A survey on current content based image retrieval methods. IETE Journal of Research, 48(3-4): 261-271. https://doi.org/10.1080/03772063.2002.11416285

Krizhevsky A, Sutskever I, and Hinton GE (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25: 1097-1105.

Li J, Allinson N, Tao D, and Li X (2006). Multitraining support vector machine for image retrieval. IEEE Transactions on Image Processing, 15(11): 3597-3601. https://doi.org/10.1109/TIP.2006.881938 **PMid:17076417**

Liu Y, Zhang D, Lu G, and Ma WY (2007). A survey of content-based image retrieval with high-level semantics. Pattern Recognition, 40(1): 262-282. https://doi.org/10.1016/j.patcog.2006.04.045

Maheshwary P and Srivastava N (2009). Prototype system for retrieval of remote sensing images based on color moment and gray level co-occurrence matrix. International Journal of Computer Science Issues, 3: 20-23.

Mahmud T, Paul B, and Fattah SA (2021). PolypSegNet: A modified encoder-decoder architecture for automated polyp segmentation from colonoscopy images. Computers in Biology and Medicine, 128: 104119. https://doi.org/10.1016/j.compbiomed.2020.104119 **PMid:33254083**

Marques O and Furht B (2002). Content-based image and video retrieval. Volume 21, Springer Science and Business Media, Berlin, Germany. https://doi.org/10.1007/978-1-4615-0987-5

Mendonça T, Ferreira PM, Marques JS, Marcal AR, and Rozeira J (2013). PH 2-A dermoscopic image database for research and benchmarking. In the 35th annual international conference of the IEEE Engineering in Medicine and Biology Society, IEEE, Osaka, Japan: 5437-5440. https://doi.org/10.1109/EMBC.2013.6610779 **PMid:24110966**

Mizotin M, Benois-Pineau J, Allard M, and Catheline G (2012). Feature-based brain MRI retrieval for Alzheimer disease diagnosis. In the 19th IEEE International Conference on Image Processing, IEEE, Orlando, USA: 1241-1244. https://doi.org/10.1109/ICIP.2012.6467091

Mohammed A, Ahmed A, Mohammed W, Viju GK, and Taha M (2020). Mammogram images classification using linear discriminant analysis. International Research Journal of Engineering and Technology, 7(6): 6656-6662.

Müller H, Michoux N, Bandon D, and Geissbuhler A (2004). A review of content-based image retrieval systems in medical applications-Clinical benefits and future directions. International Journal of Medical Informatics, 73(1): 1-23. https://doi.org/10.1016/j.ijmedinf.2003.11.024 **PMid:15036075**

Nair LR, Subramaniam K, PrasannaVenkatesan GKD, Baskar PS, and Jayasankar T (2021). Essentiality for bridging the gap between low and semantic level features in image retrieval systems: An overview. Journal of Ambient Intelligence and Humanized Computing, 12(6): 5917-5929. https://doi.org/10.1007/s12652-020-02139-z

Oussalah M (2008). Content based image retrieval: Review of state of art and future directions. In the 1st Workshops on Image Processing Theory, Tools and Applications, IEEE, Sousse, Tunisia: 1-10. https://doi.org/10.1109/IPTA.2008.4743799

Pham TD (2020). A comprehensive study on classification of COVID-19 on computed tomography with pretrained convolutional neural networks. Scientific Reports, 10: 16942. https://doi.org/10.1038/s41598-020-74164-z **PMid:33037291 PMCid:PMC7547710**

Pham TD (2021). Classification of COVID-19 chest X-rays with deep learning: New models or fine tuning? Health Information Science and Systems, 9: 2. https://doi.org/10.1007/s13755-020-00135-3 **PMid:33235710 PMCid:PMC7680558**

Pogorelov K, Randel KR, Griwodz C, Eskeland SL, de Lange T, Johansen D, and Halvorsen P (2017). Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In the 8th ACM on Multimedia Systems Conference, Association for Computing Machinery, Taipei, Taiwan: 164-169. https://doi.org/10.1145/3083187.3083212

Ponciano-Silva M, Souza JP, Bugatti PH, Bedo MV, Kaster DS, Braga RT, and Traina AJ (2013). Does a CBIR system really impact decisions of physicians in a clinical environment? In the 26th IEEE International Symposium on Computer-Based Medical Systems, IEEE, Porto, Portugal: 41-46. https://doi.org/10.1109/CBMS.2013.6627762

Pourghassem H and Ghassemian H (2008). Content-based medical image classification using a new hierarchical merging scheme. Computerized Medical Imaging and Graphics. 32(8): 651-661. https://doi.org/10.1016/j.compmedimag.2008.07.006 **PMid:18789648**

Qayyum A, Anwar SM, Awais M, and Majid M (2017). Medical image retrieval using deep convolutional neural network. Neurocomputing, 266: 8-20. https://doi.org/10.1016/j.neucom.2017.05.025

Quattoni A, Collins M, and Darrell T (2008). Transfer learning for image classification with sparse prototype representations. In the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Anchorage, USA: 1-8. https://doi.org/10.1109/CVPR.2008.4587637

Quellec G, Lamard M, Cazuguel G, Cochener B, and Roux C (2011). Fast wavelet-based image characterization for highly adaptive

image retrieval. IEEE Transactions on Image Processing, 21(4): 1613-1623.
https://doi.org/10.1109/TIP.2011.2180915 **PMid:22194244**

Rahman MM, Antani SK, and Thoma GR (2011). A learning-based similarity fusion and filtering approach for biomedical image retrieval using SVM classification and relevance feedback. IEEE Transactions on Information Technology in Biomedicine, 15(4): 640-646.
https://doi.org/10.1109/TITB.2011.2151258
PMid:21690019

Rajam F and Valli S (2013). A survey on content based image retrieval. Life Science Journal, 10(2): 2475-2487.

Rui Y, Huang TS, and Chang SF (1999). Image retrieval: Current techniques, promising directions, and open issues. Journal of Visual Communication and Image Representation, 10(1): 39-62. https://doi.org/10.1006/jvci.1999.0413

Rundo L, Tangherloni A, Galimberti S, Cazzaniga P, Woitek R, Sala E, and Mauri G (2019). HaraliCU: GPU-powered Haralick feature extraction on medical images exploiting the full dynamics of gray-scale levels. In the 15th International Conference on Parallel Computing Technologies, Springer, Almaty, Kazakhstan: 304-318.
https://doi.org/10.1007/978-3-030-25636-4_24

Safavian SR and Landgrebe D (1991). A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man, and Cybernetics, 21(3): 660-674.
https://doi.org/10.1109/21.97458

Shirkhorshidi AS, Aghabozorgi S, and Wah TY (2015). A comparison study on similarity and dissimilarity measures in clustering continuous data. PLOS ONE, 10(12): e0144059.
https://doi.org/10.1371/journal.pone.0144059
PMid:26658987 PMCid:PMC4686108

Sidney S (1988). Nonparametric statistics for the behavioral sciences. McGraw-Hill, New York, USA.

Simonyan K and Zisserman A (2014). Very deep convolutional networks for large-scale image recognition. In the International Conference on Learning Representations, San Diego, USA: 1-14.

Smeulders AW, Worring M, Santini S, Gupta A, and Jain R (2000). Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12): 1349-1380.
https://doi.org/10.1109/34.895972

Sundararajan SK, Sankaragomathi B, and Priya DS (2019). Deep belief CNN feature representation based content based image retrieval for medical images. Journal of Medical Systems, 43: 174.
https://doi.org/10.1007/s10916-019-1305-6
PMid:31069547

Suykens JA and Vandewalle J (1999). Least squares support vector machine classifiers. Neural Processing Letters. 9(3): 293-300.
https://doi.org/10.1023/A:1018628609742

Szegedy C, Ioffe S, Vanhoucke V, and Alemi AA (2017). Inception-v4, inception-ResNet and the impact of residual connections on learning. In the 31st AAAI conference on artificial intelligence, AAAI Press, San Francisco, USA: 4278-4284.
https://doi.org/10.1609/aaai.v31i1.11231

Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, and Rabinovich A (2015). Going deeper with convolutions. In the 28th IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Boston, USA: 1-9.
https://doi.org/10.1109/CVPR.2015.7298594

Szegedy C, Vanhoucke V, Ioffe S, Shlens J, and Wojna Z (2016). Rethinking the inception architecture for computer vision. In the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Las Vegas, USA: 2818-2826.
https://doi.org/10.1109/CVPR.2016.308

Tamura H and Yokoya N (1984). Image database systems: A survey. Pattern Recognition, 17(1): 29-43.
https://doi.org/10.1016/0031-3203(84)90033-5

Tian Y and Fu S (2020). A descriptive framework for the field of deep learning applications in medical images. Knowledge-Based Systems, 210: 106445.
https://doi.org/10.1016/j.knosys.2020.106445

Tsai HY, Zhang H, Hung CL, and Min G (2017). GPU-accelerated features extraction from magnetic resonance images. IEEE Access, 5: 22634-22646.
https://doi.org/10.1109/ACCESS.2017.2756624

Ucar F and Korkmaz D (2020). COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. Medical Hypotheses, 140: 109761.
https://doi.org/10.1016/j.mehy.2020.109761
PMid:32344309 PMCid:PMC7179515

Veltkamp RC and Tanase M (2002). A survey of content-based image retrieval systems. In: Marques O and Furht B (Eds.), Content-based image and video retrieval: 47-101. Volume 21, Springer, Boston, USA.
https://doi.org/10.1007/978-1-4615-0987-5_5

Wan J, Wang D, Hoi SCH, Wu P, Zhu J, Zhang Y, and Li J (2014). Deep learning for content-based image retrieval: A comprehensive study. In the 22nd ACM International Conference on Multimedia, Association for Computing Machinery, Orlando, USA: 157-166.
https://doi.org/10.1145/2647868.2654948
PMid:24955297 PMCid:PMC4050312

Wang XF, Huang DS, Du JX, Xu H, and Heutte L (2008). Classification of plant leaf images with complicated background. Applied Mathematics and Computation. 205(2): 916-926. https://doi.org/10.1016/j.amc.2008.05.108

Weszka JS, Dyer CR, and Rosenfeld A (1976). A comparative study of texture measures for terrain classification. IEEE Transactions on Systems, Man, and Cybernetics, SMC-6(4): 269-285. https://doi.org/10.1109/TSMC.1976.5408777

Win KY, Maneerat N, Hamamoto K, and Sreng S (2020). Hybrid learning of hand-crafted and deep-activated features using particle swarm optimization and optimized support vector machine for tuberculosis screening. Applied Sciences, 10(17): 5749. https://doi.org/10.3390/app10175749

Yadav SS and Jadhav SM (2019). Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data, 6(1): 1-18.
https://doi.org/10.1186/s40537-019-0276-2

Yeung M, Sala E, Schönlieb CB, and Rundo L (2021). Focus U-Net: A novel dual attention-gated CNN for polyp segmentation during colonoscopy. Computers in Biology and Medicine, 137: 104815.
https://doi.org/10.1016/j.compbiomed.2021.104815
PMid:34507156 PMCid:PMC8505797

Zhang F, Song Y, Cai W, Hauptmann AG, Liu S, Pujol S, and Chen M (2016). Dictionary pruning with visual word significance for medical image retrieval. Neurocomputing, 177: 75-88.
https://doi.org/10.1016/j.neucom.2015.11.008
PMid:27688597 PMCid:PMC5036942

Zhang K, Zuo W, Chen Y, Meng D, and Zhang L (2017). Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Transactions on Image Processing, 26(7): 3142-3155.
https://doi.org/10.1109/TIP.2017.2662206 **PMid:28166495**

Zhang Z, Li W, and Li B (2009). An improving technique of color histogram in segmentation-based image retrieval. In the 5th International Conference on Information Assurance and Security, IEEE, Xi'an, China, 2: 381-384.
https://doi.org/10.1109/IAS.2009.156