Faculty of Computers
& Artificial
Intelligence

Benha University

An Efficient Information Retrieval Model for Heterogeneous Database

A thesis presented for the degree of Master in Information Systems
Department of Information Systems
Faculty of Computers and Artificial Intelligence
Benha University

By

Ibrahem Mohamed Ibrahem Afifi

Demonstrator in Information Systems Department
Faculty of Computers and Artificial Intelligence - Benha University

Supervised By

Prof. Mohamed Abdelfatah Awad
Professor of Information Systems and Vice Dean for Education and Student Affairs
Faculty of Computers and Artificial Intelligence
Benha University

Dr. Metwally Rashad Metwally
Assistant Professor of Computer Science
Faculty of Computers and Artificial Intelligence
Benha University

Benha – 2023

Abstract

Retrieving and managing Content-Based Medical Images Retrieval (CBMIR) are considered more important now, especially with increasing in medical imaging and expanding the medical image database. Also, these systems allow for the benefit of medical images in having a better grasp on and deeper insights into the causes and treatment of different diseases, not only for diagnostic purposes. CBMIR, therefore, played an important role in the field of image processing and the extraction of low-level features such as color histograms, edges, texture, and shape, as well as similarity measures for comparison and retrieval of medical images. The majority of the methods already used in CBMIR enhance the retrieval of a medical image and disease diagnosis by reducing the issue of the semantic gap between low visual and high semantic levels. To overcome these problems, there is a critical need for an efficient and accurate content-based medical image retrieval method. This thesis proposes an efficient method of Retrieval based on Query Expansion (RbQE) for the retrieval of Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and histopathological images. RbQE is based on expanding the features of querying and exploiting the pre-trained learning models AlexNet, VGGNets, and ResNets to extract compact, deep, and high-level features from medical images. There are two searching procedures in RbQE: a rapid search and a final search. In the rapid search, the original query is expanded by retrieving the top-ranked images from each class and is used to reformulate the query by calculating the mean values for deep

features of the top-ranked images, resulting in a new query for each class. In the final search, the new query that is most similar to the original query will be used for retrieval from the database. The performance of the proposed method has been compared to state-of-the-art methods on five publicly available standard datasets, namely, TCIA-CT, EXACT09-CT, NEMA-CT, OASIS-MRI, and KIMIA Path960. Experimental results show that the proposed method exceeds the compared methods by 0.84%, 4.86%, 1.24%, 14.34%, and 1.96% in average retrieval precision (ARP) on the retrieval of the top ten, for TCIA-CT, EXACT09-CT, NEMA-CT, OASIS-MRI, and KIMIA Path960, respectively. But for KIMIA Path960, the ARPs for the top five, fifteen, and twenty exceed the compared methods by 2.14%, 6.74%, and 9.42%, respectively.

Acknowledgments

All gratitude and thankfulness to ALLAH for guiding and aiding me to bring this work out to light. Also, it is impossible to give sufficient thanks to the people who gave help and advice during the writing of this thesis. However, I would like to single out both my supervisors for their valuable discussions, enriching collaborations, and their vital comments on the contents of this thesis. Without their kind help and generous encouragement, this work would not have seen the light.

I am deeply indebted to Dr. *Metwally Rashad Metwally* for his kindness and emotional support throughout the period it took to prepare this work in its final form. I am particularly grateful for his genuine concern and his prompt replies to all the questions I addressed to him, for giving so generously of his time in revising this work and, in the process, pointing out some relevant and interesting ideas.

I also wish to express my deep gratitude to Professor Dr. *Mohamed Abdelfatah Awad* for his guidance and moral support during the making of this work and also for his insightful remarks, fruitful suggestions and constructive criticism.

I also wish to express my deep gratitude to the staff of the department of information systems, faculty of computers and artificial intelligence, benha university for their guidance and moral support during the making of this work. Finally my deep appreciation and gratitude to my family for their kind help, their continuous encouragement, and support to keep going through this work.

List of Publications

- **Paper 1:** Metwally Rashad, Ibrahim Afifi, Mohamed Abdelfattah. **RbQE: An Efficient Method for Content-Based Medical Image Retrieval Based on Query Expansion**. J Digit Imaging (2023). https://doi.org/10.1007/s10278-022-00769-7.

- **Paper 2:** Metwally Rashad, Ibrahim Afifi, Mohamed Abdelfattah. **Content-based Medical Image Retrieval based on Deep Features Expansion**. In: 2022 5th International Conference on Computing and Informatics (ICCI). IEEE; 2022. Available from: https://doi.org/10.1109%2Ficci54321.2022.9756114.

- **Paper 3:** Metwally Rashad, Ibrahim Afifi, Mohamed Abdelfattah. **Effective of Modern Techniques on Content-Based Medical Image Retrieval: A Survey**. International Journal of Computer Science and Mobile Computing. 2022 mar;11(3):56-77. https://doi.org/10.47760/ijcsmc.2022.v11i03.008.

Contents

List of Figures

List of Tables

List of Abbreviations

Table 1: Table of Abbreviations

Abbreviation	**Referenced Terms**
CBMIR	Content-based Medical Image Retrieval
RbQE	Retrieval based on Query Expansion
IR	Information Retrieval
VIR	Visual Information Retrieval
MIR	Medical Image Retrieval
RDBMS	Relational Database Management System
TBMIR	Text / Tag Based Medical Image Retrieval
ED	Euclidean Distance
HVS	Human Visual System
CNN	Convolutional Neural Network
SIFT	Scale Invariant Feature Transform
SURF	Speed Up Robust Features
BoW	Bag of Words
BBF	Best Bin First
NNDRS	Nearest Neighbor Distance Ration Scoring
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
TCIA	The Cancer Imaging Archive
DICOM	Digital Imaging and Communications in Medicine
EXACT09	Extracting Airways from CT 2009
NEMA	National Electrical Manufacturers Association
OASIS	Open Access Series
DCNN	Deep Convolutional Neural Network
FC	Fully Connected
NQE	New Query Expansion
FNQE	Final New Query Expansion
QDF	Query with Deep Features
ResNets	Residual Networks
ARP	Average Precision Retrieval
ARR	Average Retrieval Rate
SDAE	Stacked Denoising Autoencoders
BBF	Best Bin First
ED	Euclidean Distance
CS	Cosine Similarity
MD	Manhattan Distance

1. Introduction

Chapter 1

Introduction

In this chapter, we introduce the background of Information Retrieval (IR) and Content-Based Medical Image Retrieval (CBMIR), the thesis objectives and main contributions, and the thesis organization.

1.1 Background and Problem Statement

The term Information Retrieval (IR) is the science of searching for information in a document, searching for documents themselves, and also searching for the metadata. This term was coined in 1952, and since 1961 it has gained popularity in the research community. Also, one may describe such a system as storing and retrieving information. In the past, information retrieval has meant textual information retrieval, but the above definition still holds when applied to visual information retrieval (VIR). However, there is a distinction between the type of information and the nature of the retrieval of text and visual objects [68]. Textual information is linear, while images are bi-dimensional. Generally, there are two approaches to solving the VIR problem. They are based on the form of visual information: attribute-based and feature-based methods. Attribute-based methods rely on traditional textual information retrieval and Relational Database Management System (RDBMS) methods.

While feature-based solutions concentrate on visual features such as color, texture, and shape of images that would be indexed according to these features [126]. Medical Image Retrieval (MIR) is the way for searching and retrieving relevant medical images from large medical databases. There is a three types of retrieval for medical images as follow:

- Text / Tag Based Medical Image Retrieval (TBMIR) is based on the textual description or annotation of images.
- Content-Based Medical Image Retrieval (CBMIR) is based on the image features like color, texture, and shape or any other features being derived from the image itself.
- Hybrid Approach both textual query, and visual features are combined (also known as fusion) to obtain the desired output medical image [49].

CBMIR is an image search technology that aims to find images most comparable to a certain query, where the search for images is based on their features, whether high-level or low-level. The main difference between low and high-level features lies in the fact that the first one is characteristics extracted from an image, such as colors, edges, and textures. These features are frequently unique to a single image or video and have little relevance on their own. In contrast, the second one is extracted from low-level features, denotes more semantically meaningful concepts, and contains more complicated details about an image's or video's topic. Items, scenarios, and interactions are instances of high-level features. CBMIR's success depends mainly on the selected features where the images are represented as features with a high dimension [70]. The similarity between the query and the stored images in the datasets is measured using distance metrics such as Euclidean Distance (ED) and Cosine Similarity (CS). The encoding of images in terms of features

and the selection of an algorithm for measuring similarity are thus the most significant components of CBMIR systems. Despite the fact that several researchers have extensively researched these topics [128].

The CBMIR main architecture is shown in Fig. 1.1, which is composed of two fundamental steps: the feature extraction step (offline phase) and the similarity measurement computation step (online phase) [93, 122, 104]. There are many similarity measurements that can be used in (online phase), let X = (x_1, x_2,..., x_n) and Y = (y_1, y_2,..., y_n), two feature vectors with n dimension, the similarity distance is computed as follows:

- Euclidean Distance (ED) [26]

$$ED\left(X,Y\right)=\sqrt{\sum_{i=1}^{n}\left(x_i-y_i\right)^2} \tag{Ch1-1}$$

- Manhattan Distance (MD) [25]

$$MD\left(X,Y\right)=\sum_{i=1}^{n}\left|x_i-y_i\right| \tag{Ch1-2}$$

- Cosine Distance [69]

$$Cosine\left(X,Y\right)=\frac{x\cdot y}{|x||y|} \tag{Ch1-3}$$

- Minkowski Distance [41]

$$Minkowski\left(X,Y\right)=\left(\sum_{i=1}^{n}\left|x_i-y_i\right|^p\right)^{\frac{1}{p}} \tag{Ch1-4}$$

- Jaccard Distance [59]

$$Jaccard\,(X,Y) = \frac{|X \cap Y|}{|X \cup Y|} \qquad \text{(Ch1-5)}$$

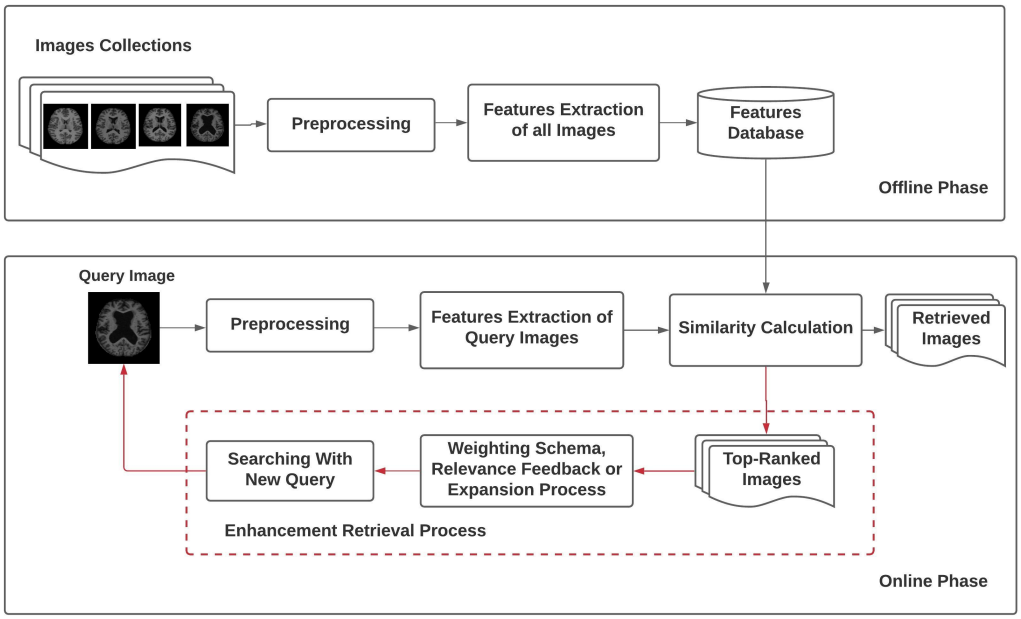

Figure 1.1: The CBMIR Main Architecture.

The CBMIR may also provide an effective means of supplementing diagnostics and treatment of various diseases, as well as an efficient tool for managing massive amounts of data [133]. The implementation of an efficient framework for medical image retrieval is therefore important to help clinicians access such large datasets. Many algorithms for automated medical image analysis and retrieval have been proposed in the literature to encourage the creation and management of these massive datasets of medical images [79, 53, 97, 100, 101].

Different descriptors at a global level have been created over the last few decades to represent images. For example in features of shape, color [50], and texture-based

features [75]. Eakins [35] categorizes image features into three levels. first, the spatial location of image elements and colors, as well as texture and shape, are considered from level one features. Second, the features that are derived or logical features are considered to be at level two, which implies a degree of inference with regard to the identities of the objects in the image. Third, the features at the final level are recognized as abstract attributes, and they involve complex inferences about the significance of objects belonging to the image.

Shape and color are considered significant features in images, where various visual indications are given by the shape, such as characteristics of curve, contour, and surface, while powerful visual cues are provided by color and dominate color imaging. The features of texture spatially organized the pixels in an image. These features need standard tools to analyze them, like Gabor filters [35], Fourier transforms, and wavelets. In addition, there are some local descriptors that have been developed like Scale Invariant Feature Transform (SIFT) [73], Speed Up Robust Features (SURF) [14], and model of Bag of Words (BoW) [137, 111, 135]. The semantic gap problem was tried to address in various studies [132, 76, 7], but the features used in these studies cannot fully able address this problem. The SIFT features are used in [132] as a representation of images of the training and query, and the performance of the system was tested by using two classifiers, where the classification and indexing were done by algorithms of the nearest neighbor and KD tree with the best bin first (BBF).

A modified voting approach called "Nearest Neighbor Distance Ration Scoring" (NNDRS) is aggregating the scores of candidate images, and the aggregate scores are then sorted into descending order. After that, the images with top rank are retrieved. In [76], a general image CBMIR system is introduced with three features for representing the images in the feature space. Color and texture features are

considered low-level, while the structure of the binary tree is utilized to capture higher levels. The Human Visual System (HVS) is utilized for the features of color, while wavelet decomposition is employed for the features of texture.

Many upgrades to the CBMIR system have been created to improve the effectiveness and retrieval performance, which can be at the stage of pre-processed or extracted features [109, 57]. Texture-based features are well accepted and popular among researchers worldwide, according to the extensive medical image retrieval literature [103, 112, 58, 71, 80]. However, medical imaging becomes more sophisticated over time as it attempts to gather as much information about the patient's anatomy as possible.

The fast rise of digital, multimedia, and storage systems in recent years has resulted in massive image and multimedia warehouses. The advances in digital storage and information retrieval benefit clinical and diagnostic studies. Diagnostic and investigative imaging systems in hospitals generate a massive amount of images, which contributes to the expansion of medical image archives. Also, the retrieval of medical images based on textual details like tags and manual annotations is poor in productivity since they require manpower, clinical skills, and time. Medical image retrieval systems, which can retrieve and identify the images automatically based on the features coming from the images themselves, are so critical. The most difficult problem in CBMIR systems remains the reduction of the semantic gap. This gap occurs between the visual input of the HVS and the imaging system when information is lost during the process of representing the image from high-level semantics to low-level comprehension features [61].

As a result, developing a powerful CBMIR system based solely on texture is insufficient. It is, therefore, necessary for the hour to build a system for the multidimensional retrieval of medical images that will combine multidimensional informa-

tion, for example, texture, edge, and shape. It is considered the fundamental component of any CBMIR system that compares an image to an image in the dataset to determine how similar they are and to find matching pairings to the image [4, 5]. Traditional methods rely on low-level extraction by assessing their colors, textures, forms, and spatial structure from medical imagery. These are all low-level features that frequently do not accurately reflect semantic notions in images. Using these features for retrieval usually yields unsatisfactory results.

1.2 Thesis Objectives

Our objective in this thesis is to develop an efficient CBMIR technique that is able to achieve high accuracy in the retrieval of medical images that have different modalities with less retrieval time, reduce the semantic gap problem, and represent the medical image with a low-dimensional feature vector.

The semantic gap will be reduced by either incorporating domain-specific knowledge or employing machine learning techniques to develop intelligent systems capable of acting like the HVS.

Machine learning technology has evolved dramatically, and deep learning systems have been a breakthrough, where deep learning includes numerous algorithms of machine learning for modeling abstractions of high-level data [15] by employing a deep structure that contains multiple nonlinear transformations. Deep learning is modeled after the human brain [128], whose architecture is complex and processes information through transformation composed of multiple layers. Thus, deep learning led to learning the complex features from raw images by a machine (deep networks) without using handcrafted features, this enables us to automatically get feature representations by using several abstract levels to automatically learn fea-

tures by exploring deep structures.

The use of pre-trained Convolutional Neural Network (CNN) model features has lately achieved superior performance and flexibility than classical descriptors in common image retrieval applications due to the quick advancement of deep learning. (e.g., image retrieval or object recognition). Rich image semantic information is provided by this feature, which is crucial for improving the precision of image retrieval. Also, recent research has shown that deep learning systems can effectively implement image and video classification [8, 140, 55], visual tracking [134], speech recognition [48], and natural language processing [142].

1.3 Thesis Contributions

The main contributions of our thesis are summarized as follows:

1. Presented an overview on modern CBMIR methods after searching and studying the topic thoroughly as shown in Chapter 3.

2. Proposed an efficient medical image retrieval method (RbQE), which expands the query in a new automated way as illustrated in Chapter 4.

3. Use the pre-trained deep convolutional neural networks (AlexNet, VGGNets, and ResNets) as feature extractors that describe and represent medical images to obtain complex and high-level features that have the best ability to withstand external interferences such as changes in lighting, noise, rotation, and blurred images as shown in Chapter 5.

4. Extensive tests have been carried out to compare the performance of the proposed method (RbQE with DCNN) with the existing and modern methods,

and it has been demonstrated that the proposed method exceeds all these methods in retrieving medical images as shown in Chapter 5.

5. Tested the proposed method on five different modalities of standard and publicly available datasets: TCIA-CT, EXACT09-CT, NEMA-CT, OASIS-MRI, and KIMIA Path960.

1.4 Thesis Organization

We describe the various medical image datasets used in our experimental evaluations, hardware and software issues in Chapter 2. In Chapter 3, we define an overview of modern CBMIR methods. In Chapter 4, we will show an efficient method for content-based medical image retrieval based on query expansion (RbQE). In Chapter 5, we present the experimental evaluation of the RbQE method. Finally, the conclusion and future work are presented in Chapter 6.

2. Datasets of Medical Images

Chapter 2

Datasets of Medical Images

The evaluation of CBMIR methods greatly depends on the accuracy and time of the retrieval for the standard datasets, which have different modalities. As our purpose is to develop a method that achieves high accuracy in less time, we have compared the proposed method with modern methods on five publicly available image datasets with different modalities.

2.1 Medical Image Datasets

There are many medical images in different modalities, and we have used five publicly available standard image datasets in different modalities, such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and histopathological images, for testing purposes.

2.1.1 TCIA-CT Dataset

The cancer image archive (TCIA), which was compiled by the authors in [32] and contains a huge number of cancer images [24]. This dataset contains 604 CT images in DICOM format with series number 1.3.6.1.4.1.9328.50.4.2 of study instance UID

1.3.6.1.4.1.9328.50.4.1 for subject 1.3.6.1.4.1.9328.50.4.0001. The TCIA-CT collection is divided into eight classes, contains 75, 50, 58, 140, 70, 92, 78, and 41 images respectively. The TCIA-CT dataset samples for each class can be viewed as shown Fig. 2.1.

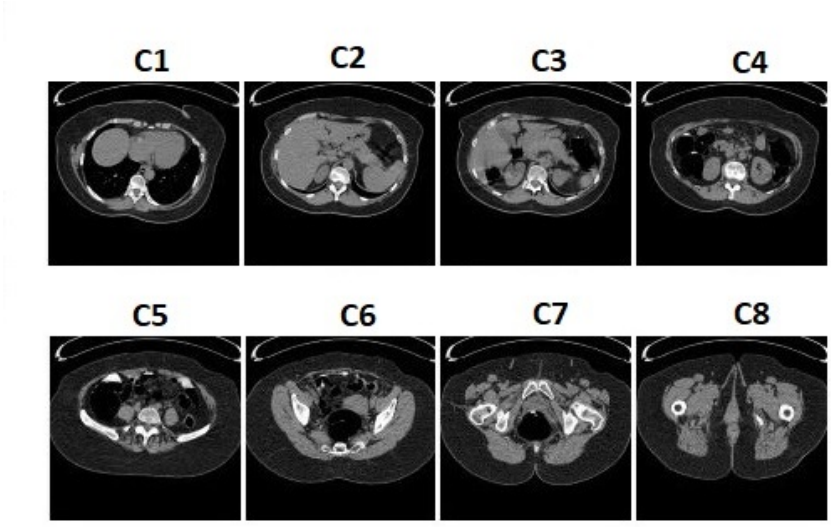

Figure 2.1: Sample images from each class of TCIA-CT dataset.

2.1.2 EXACT09-CT Dataset

The Extracting Airways from CT 2009 (EXACT09) is composed of chest CT scans [72]. In order that we can assess our suggested features, as indicated in [32], we picked a CASE 23 test set, a subset for the EXACT09-CT dataset. In this dataset, there are 675 CT images, organized into 19 classes as outlined in [32]. 36, 23, 30, 30, 50, 42, 20, 45, 50, 24, 28, 24, 35, 40, 50, 35, 30, 28, and 55 are the image numbers for each class, as shown in Fig. 2.2.

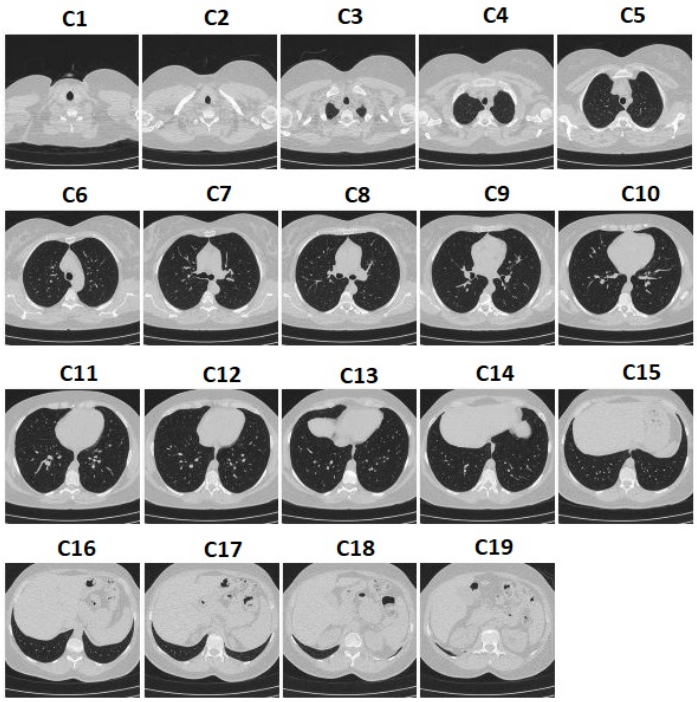

Figure 2.2: Sample images from each class of EXACT09-CT dataset.

2.1.3 NEMA-CT Dataset

The National Electrical Manufacturers Association [88], which comprises CT images of various sections of the human body. Following the experimental setup in [32], the NEMA-CT dataset was create by selecting 315 CT images from CT0001, CT0003, CT0057, CT0060, CT0080, and CT0083. Furthermore, those images are separated into 9 classes each comprising 36, 18, 36, 37, 41, 30, 23, 70, and 24 images representing different regions of the body displayed in Fig. 2.3.

2.1.4 OASIS-MRI Dataset

The Open Access Series (OASIS) with MRI [77], where the OASIS-MRI is considered a type of medical dataset created by the Image Studies Open Access Series (OASIS)

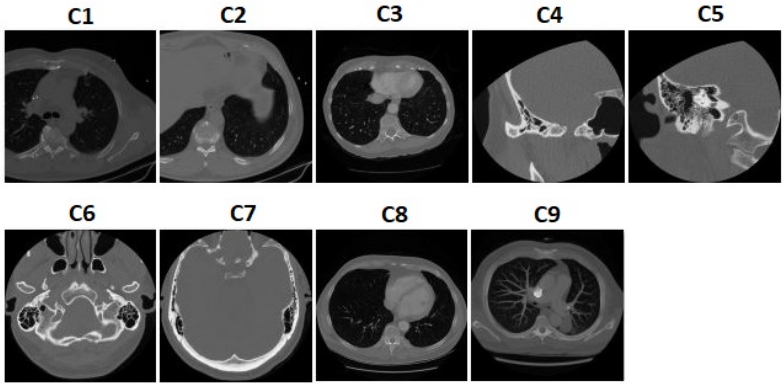

Figure 2.3: Sample images from each class of NEMA-CT dataset.

with MRI, within the original datasets, a sectional range of 421 topics between 18 and 96 years can be studied and analyzed. These images were classed into four classes by authors in [83], with 124, 102, 89, and 106 images in each class, illustrated in Fig. 2.4.

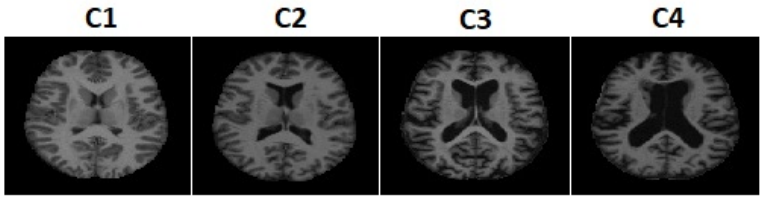

Figure 2.4: Sample images from each class of OASIS-MRI dataset.

2.1.5 KIMIA Path960 Dataset

The second is the KIMIA Path960 dataset [62] 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 intraclass

differences and certain similarities between different classes. The KIMIA Path960 dataset contains 960 Images, where there are 48 images in each class as shown in Fig. 2.5.

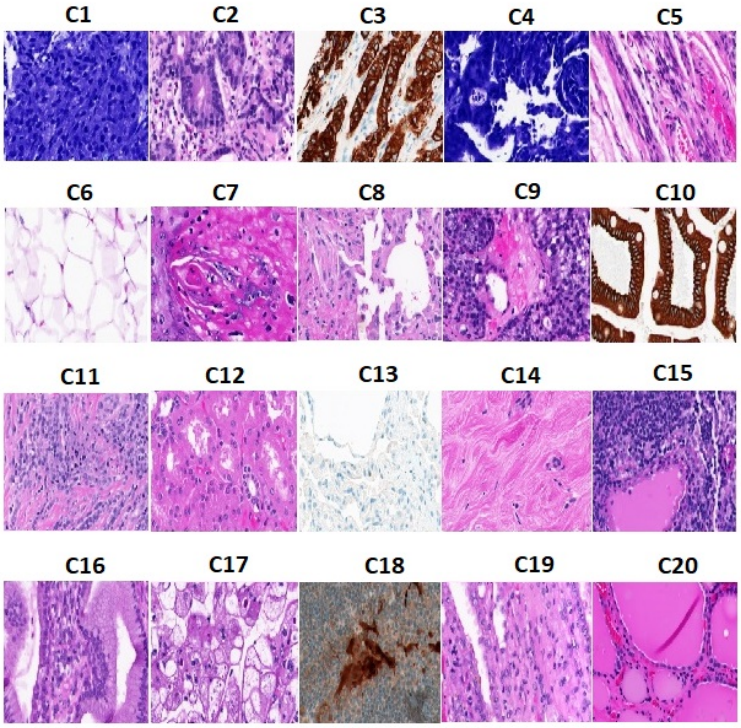

Figure 2.5: Image from each class of the KIMIA Path960 dataset.

2.2 Hardware and Software Issues

The development and some offline and online evaluations were performed on a computer, with Intel(R) Core(TM) i7-4510U CPU @2.00 GHz processor, 8 GB RAM, and 64-bit Windows 10 Enterprise LTSC operating system. Software development was done in spider using OpenCV and Matlab.

3. Modern CBMIR Methods: Literature Reviews

Chapter 3

Modern CBMIR Methods: Literature Reviews

In this chapter, we introduce and summarize how modern CBMIR methods applied to the mentioned datasets, mentioning the advantages and disadvantages of each method. Section 3.1 presents the majority of CBMIR methods that is based on local descriptors to extract low-level features for representing medical images. While Section 3.2 presents most of the CBMIR methods that depends on machine learning and deep learning descriptors for extracting high-level features for representing the medical images. In Table 3.1, we presented the summarization of the CBMIR methods in terms of the model, datasets, classification technique, feature vector length, extraction and retrieval time, and the accuracy of each model.

3.1 Methods of CBMIR based on Low-Level Features

In order to enhance the process of identifying dominant features, researchers aim to provide the best models and descriptors to increase the quality of medical im-

age retrieval. In [33] a new CT image retrieval image descriptor was suggested for a Local Wavelet (LWP) pattern. Firstly, the decomposition of the local wavelet takes place in the local pixel area to encode the relationship with the neighboring pixel. Secondly, the local wavelet was compared to the center pixel's transformed values for the encoding relationship between the center and neighboring pixels, and the LWP pattern for the center pixel was calculated. Thirdly, they generated the LWP for every pixel of the image and eventually used it as a feature vector to find the histogram. The LWP feature descriptor was tested using three CT medical image retrieval studies on three medical CT-image bases (NEMA-CT dataset [88], EXACT09-CT dataset [72], TCIA-CT datasets [32]), and compared LWP with LTP descriptors [119], LTCoP [84], LBP [90], and LMeP [84]. From the experiments, the LWP feature descriptor executes the current feature descriptors for each dataset, where the LWP is superior to almost every group within each dataset, and the LWP feature descriptor's time complexity decreases. The local neighbors that influence the LWP dimension are only taken into consideration as local neighbors. Experiments and research show that their LWP feature descriptor can be used more easily, and reliably for medical CT image diagnosis, but the high-level semantic information is not fully extracted from the image.

A simple method to use the texture features for retrieval of the medical images has been suggested in [65]. This method is considered more efficient and discriminative compared to recent methods such as HCSCs [66], where just 21% of the time is spent in HCSC scattering transform for the image filtering and partitioning operation. The images in this method are filtered with different filters from Gabor and Schmid and the filtered images are then divided into patches that do not overlap. Finally, they used the BoW model to represent the features of images. This method was applied to two benchmark CT datasets, EXACT09-CT, and TCIA-CT, where

the experiments showed their superiority compared to the recent methods. However, somewhat it did not reduce the semantic gap significantly too.

A new descriptor has been proposed in [30] for biomedical images retrieval and indexing called Local Bit Blane Decoding Pattern (LBDP), where for LBDP binary pattern creation, in each plane the LBDP transforms the local district, encoding the relationship between the center pixel intensity value and the values that are transformed. The structure of the LBDP process differs from the existing descriptors of the features belonging to the image. In their approach, the dimension was depending on the image bit depth and local neighbors count. So, three tests of the retrieval for the biomedical images were carried out to assess the improvements and efficiency of LBDP discriminating in terms of ARP, ARR, and F_{score}. Two datasets of CT images (Emphysema-CT [36], and NEMA-CT) and one MRI (OASIS-MRI) [77] repositories are tested to prove that LBDP is above these current modern descriptors. The time to retrieval with LBDP is considerably reduced while simultaneously increasing efficiency. Furthermore, the outline of the LBDP feature may also be employed in an invariant face recognition task. But, this approach includes a disadvantage where the system has relatively low output because the high level of semantical information is usually not defined as within the mind of the user.

In [52] the authors proposed a new extract approach that relies on the extraction using the Gray-level Co-occurrence Matrix (GLCM) [44] and the Local Pattern Descriptors (LPD) of texture features from medical images. To gain improved retrieval output on medical images in order to make quick decisions in clinical applications. The Local Mesh Vector Co-occurrence Pattern Descriptor is presented (LMVCoP). Through integrating the LPD with the GLCM the LMVCoP extracted textural features in the MRI brain images. LMVCoP has been compared to existing descriptors, such as the LTCoP, LMeP, LBDP, PVEP, and LVP [37], for performance appraisal

measures. Their outcomes are evaluated in terms of evaluation metric accuracy. The LMVCoP descriptor has been used on the MR (OASIS-MR) dataset, where 87.57% of ARP was achieved and 53.21% for ARR. The time used for extraction of features of LMVCoP descriptors is 04.29 s and for retrieval is also 0.21 s which is considered less. The results of the simulation and the experiments have shown that the LMVCoP exceeds the other state-of-the-art. Such improvements, however, have a disadvantage, where the dimension of the feature vector is very large, and it is highly inefficient to index and match processes.

One aspect of the success of any CBMIR method is its ability on the feature extraction that describes the high-level semantics in images. So, in [63] they suggested a novel method for biomedical image retrieval utilizing the Zernike Moments (ZMs) [121]. ZMs belong to the global descriptors group and are orthogonal moments. ZMs acquire gross image information that extractors of the local feature like LBP, LDEP [31], and others cannot obtain. This broad perspective is excellent for distinguishing medical images from various regions of the body precisely. Furthermore, their noise insensitivity makes them ideal for the retrieving of images that are real-life, which are commonly harmed by noise. In order to evaluate the method, they have applied their method using two image datasets, namely EXACT09-CT, and OASIS-MRI respectively. The experiment showed that when compared to other known features such as LBP, ULBP, or LDEP the methodology of ZMs-based methods has achieved high performance and other states of the art recently released methods like SS-3D-LTP, LBDP, LMePVEP [85] and CSLBCOP [124]. Despite these advantages, it has some drawbacks. Firstly, the outer circle mapping method has been used for computing ZMs. The use of outer circle mapping might result in inaccurate image retrieval. Secondly, the images in the classes of both the Emphysema-CT and OASIS-MRI datasets appear to overlap in the existence of noise, i.e. to fall into an-

other class, thereby causing incorrect retrieval. Thirdly, fewer features of the ZMs are used for classification purposes. Finally, in the Emphysema-CT and OASIS-MRI datasets, tiny images lead to poorer retrieval levels when they may be connected.

In [1] a new method for the enhancement and resolution of the major problems in [63, 94], and in the majority of modern methods, where Orthogonal Fourier-Mellin Moments (OFMMs) was suggested. OFMMs have several features, such as greater information packing power, small feature dimensions, improved image noise robustness, rotation invariance, and scaling, as a result, they are ideal for applications of retrieval biomedical images. When compared to other state-of-the-art methods like (Local Binary Pattern (LBP), Local Ternary Pattern (LTP), Local Quinary Patterns (LQP), etc..)., this approach is also incredibly efficient since it requires very few CPU seconds to obtain the top images that are matched to the query. They have applied their method on four various medical datasets namely the Emphysema-CT, NEMA-CT, OASIS-MR, and NEMA-MR [89]. The results show that the rate of retrieval from all current methods for noisy and non-noise images of each of the four medical test bases improved considerably by around 20% and 7% (average) respectively, but this method has disadvantages, where the major difference between multimodal datasets when working with a large dataset may not be very satisfactory.

A robust descriptor for features named the antithetic isomeric cluster pattern (ANTIC) was suggested for CBMIR and modification applications in [74]. The ANTIC is inspired by isomerism, which uses ANTI and clustering characteristics. It is able to extract information about the line and corner points in the local neighborhood and therefore provides a robust texture descriptor. In addition, only four isomeric antithetic patterns are needed for all directional information to be obtained. By reducing the length of the feature vector, this trait ensures robustness. Further-

more, the MANTIC, which combines multi-resolution information using Gaussian filters, also improves its performance. This descriptor has been applied to four benchmark datasets MESSIDOR, NEMA-CT, OASIS-MRI, and VIA/I-ELCAP, where the experiments showed their superiority compared to the recent methods. However, it did not reduce the semantic gap somewhat, as well as although the length of the feature vector was reduced to half, but still large somewhat.

In [115], an effective approach has suggested retrieving biomedical images using the global and local image features. The Local Binary Edge Directional Pattern (LDEBP) has been developed to extract local features, where information for each pixel in the image is collected from every direction, i.e. 0^0, 45^0, 90^0 and 135^0. Directional information shall be calculated based on the magnitude of the local signals code between the central pixel and its directional pixels. All spatial information is used for the evaluation of four edges for each pixel. Local and shape features of the image have been derived from the Lower Order Zernike Moments (LZMs) [114]. Once the shape and texture descriptors were integrated, the results were improved. Such results also demonstrated substantial progress when applied to benchmark datasets such as Emphysema-CT, OASIS-MRI, brain tumor [19], and compared to validated approaches such as LBDP, ZM, and LDEP. Yet, there is a downside to all these enhancements, where the time for extraction and retrieval is comparatively growing compared to some other approaches.

In [18] a new indexing and retrieval method for medical images was proposed as a kind of improved CBMIR retrieval. The Local Binary AND Pattern threshold (TLBAP), as well as the adjacent Local Average Differential Pattern (LANADP), were proposed as two new descriptor features. Every descriptor has been added to medical images, every descriptor generates a histogram and at the end, the two histograms are combined in order to create the last features vector. Various description

patterns, such as the Local Binary Pattern (LBP), are considered to be the binary pattern threshold for each pixel center in the image, but the TLBAP descriptor improves LBP by using the highest pixel intensity of neighboring pixels in calculating the threshold value. Then, they perform a logical AND operation between the LBP pattern and TLBP pattern to generate a TLBAP pattern. Furthermore, the other descriptor (LANADP) was suggested as a way to extract all dominant features and reduce the semantic gap by finding the relationship of adjacent pixels to their next neighbors in the diagonal, vertical, and horizontal directions. They have used three datasets of biomedical images, namely, NEMA-CT, OASIS-MRI, and VIA/ELCAP-CT, for testing the efficacy of their method in comparison with the other methods. All the experiment results of the proposed method were superior. However, all these improvements in retrieval efficiency have a downside, where the retrieval duration is significantly longer than the other methods.

A new 3D Local-Oriented Zigzag Fused Pattern descriptor (3D-LOZFP) was proposed in [45] for the purpose of retrieval of CT images. The three distinct 3D zig-zag patterns in four directions on a 3D plane are used by this descriptor to encode the link between the center and the neighboring pixels. In total, there have been 12 effective 3D zig-zag patterns. The 3D plane is created through a Gaussian filter bank which generates numerous, multi-scale, filtered images. A high-dimensional vector was produced by this descriptor. The number of features was therefore decreased by utilizing a quantization and fusion technique. This method was applied to two benchmark CT image datasets: the NEMA-CT dataset and the TCIA-CT dataset. The results of the experiments have shown that this method is highly efficient relative to other recent methods. Although this superiority, this method has some drawbacks, where it failed to extract high-level semantic information from the image and the feature extraction process is considered time-consuming.

The researchers have continued to improve the methods proposed in [98, 2] in which a new technique is applied which seeks to reduce semantic gaps and obtain the best possible results, Therefore, in [107] have been proposed new CBMIR system that focused on the spatial matching of visual words and a new similarity metric, the "Skip Similarity Index" (SSI), is efficiently utilized in computing visual word similarity spatial similarity. The results of the experiment revealed that it's most successful in extracting multimodal image datasets from various body organs with similar anatomical structures. The mean Average Precision (mAP) for the retrieval process of 69.70% was achieved in the proposed method. Nevertheless, the method did not achieve a high-precision average retrieval of 50 images on the Kvasir dataset as opposed to the method in [3] where 74.02% had been achieved, but [107] had 60.48%. Even when comparing the images in the whole dataset to establish similarity, the computational complexity of the suggested technique is substantial when compared to standard BoVW techniques.

More was made at [108], when a modern medical image retrieval technique was introduced to retrieve medical images using a topic location feature vector from a huge multimodal medical image dataset. In order to obtain a topic probability, the Guided Latent Dirichlet Allocation Model (GuidedLDA) is used, and the suggested location model is used to determine the location probability. The location model is depicted in images of each topic as the standard distribution of visual word locations. The location likelihood of the images was evaluated based on the closeness of the location of visual words in the established location models. The feature vector was created by combining the topic and location probabilities of visual words in the images. The significant position information in the topics has reduced the problem of the semantic gap, where the images were represented by a low-dimensional location feature vector. By predicting an image label using the topic-location feature vector,

search times have been decreased by avoiding a full search for images that are similar over the whole dataset. The location-weighted precision (wPrecision) was introduced and became a better way of evaluating the normal precision metric set of medical images. Experiments demonstrate that including spatial information in the topics improved the performance of CBMIR for a medical image when compared to other modern retrieval systems for medical images. Despite its accuracy, this approach has a disadvantage in that it requires a large amount of quality-labeled data for the generation of models, which is difficult to obtain.

In [56] the authors introduced a method of CBMIR, where this method is composed of two phases: enrollment and querying. Firstly, the Discrete Wavelet Transform (DWT) coefficients were calculated for each inbound image with four types of Haar [11, 17], Daubechies [22, 129, 139], Coiflet Wavelets [51], and the most suitable type of wavelet for tested and compared retrieval system for each inbound image. Then, from wavelet coefficients, the Block Truncation Codes (BTCs) have been extracted. The suggested method was also enhanced with various scanning techniques. Rasters, zigzags, Morton, and Hilbert scans were used to split the image into subblocks to match BTC. The derived codes are then stored as a dataset of the feature vectors. Secondly, the BTCs are extracted from the wavelet coefficients of the query image during the query process. Eight different distance metrics were used to measure similarity. In order to test their method, three medical image datasets were used, namely, 7500 CT brain images collected from a teaching hospital in Egypt, Kvasir, and VIA-ELCAP. The experimental results indicate that, in contrast with other methods, the method is very efficient. The proposed method also showed strong results in the extraction of Morton BTCs from the DB2 DWT analysis and the best estimation of similarities with the Manhattan distance. Given these benefits, but with some disadvantages, where the feature vector will be very

large, in [3] have better accuracy on the Kvasir dataset than this method and the high computational complexity where this method searches the images in the whole dataset to compute the similarity.

In [6], a Relevant Feedback Retrieval Method (RFRM) for CBMIR has been suggested. Here the feedback is based on voting values performed in the image repository by each class, where a group of color features and texture was extracted using well-known color moments and GLCM texture features extraction methods and eight common coefficients of similarity were used as a base for similarity measures. After a rapid search using a single query that is taken randomly, the images that have top rank are selected from the dataset to determine the most efficient coefficient of similarity for the final search process. This method was applied to benchmark datasets the Kvasir dataset, and the PH2 dataset, where the experiments proved the superiority for improving the retrieval effectiveness of related medical images. Although this superiority, this method has a drawback, where it requires labeled data that are not easily available.

Because histological images gave a clear perspective of illnesses, the pathologist thought them to be the sole way to diagnose a disease. Automatic identification and extraction of advantageous traits from histopathological images can aid in the diagnosis, whereas manual assessment of microscopic images requires subjective pathologist interpretation and takes time. Therefore, a new CBMIR technique was introduced in [117]. The scheme presented downsamples the image to different scales, in which LTP is used for each downsampled image for lower, upper, and then divides into concentrated patches. The scheme defines the images on a different scale. In order to achieve the high accuracy of the retrieval of histopathological imagery, they have applied power-law normalization and a Vector of Locally Aggregated Descriptors (VLAD) to get the histogram images, which have high accuracy.

The experiments have been conducted using the KIMIA Path960 dataset [62] of standard 20-class histopathological data, and the results demonstrated that their method achieves better retrieval accuracy than other techniques. However, this technique has a drawback, as it requires improvements in computation complexity and performance.

In order to distinguish between normal and abnormal lung CT images, in [138] the Local Diagonal Laplacian Pattern (LDLP) has been proposed as a novel low-dimensional descriptor with computational efficiency. In order to define the diagonal neighbor center pixel relation, LDLP uses a derivative method of the second order. This led to a feature vector low-dimensional with rich local structure information and thus the calculation costs were significantly reduced. First, the feature histograms for the chest CT image slices of the EMPHYSEMA dataset were identified. Then the statistical method ANOVA was also used for measuring and analyzing the distance between normal and emphysematous tissue features. In addition, a classification of four-class has been produced using the classification of a neural artificial network. In the end, the results of the experiment have been compared with prominent methods like LBP, Local Tetra Pattern (LTrP) [82], Local Diagonal Extrema Pattern (LDEP) [31] and have shown high accuracy compared with them. This method failed to fully extract the high-level semantic information.

3.2 Methods of CBMIR based on High-Level Features

Hand-crafted features can negatively influence medical image retrieval. So, we have found in [66] a unique feature for medical imaging, where a Histogram of Compressed Scattering Coefficients (HCSCs) is suggested. Their method works as fol-

lows: First, it was necessary to create invariant representations of its translation scattering transformation to a medical image. Second, they employed the BoW frame to drive the histogram as a feature vector after the compression process. Thirdly, to assess the features efficiency of the HSCSs, their method was applied using three CT image datasets, namely NEMA-CT, EXACT09-CT, and TCIA-CT respectively. The result of the experiments proved that their method has achieved high performance compared to many features, such as LWP, LMeP, SS-3D-LTP, [86], and LBP. Despite these advantages, it has some drawbacks. Firstly, the HCSC used partial image information as it allows only one path of projection for compression. Secondly, the time consumption arising from the derivation of the codebook used in the BoW model due to the clustering process affected the efficiency of the HCSC adversely.

Due to the restrictions in [66] which influence the results' accuracy. A novel and integrated scattering feature for the retrieval of medical images was thus proposed in [64]. The method proposed combines two types of compressed dissemination data from a variety of points of view, namely Data Concentration (DC) and Canonical Correlation Analysis (CCP), where the feature presented takes account of the relevant scattering data capable of high-level representations of the original medical image. For the validation of the method, two of the CT image datasets, EXACT09-CT, and TCIA-CT, have been used in their process. The experiment results demonstrated high performance compared to other current featured methods including LWP, HSCSs, LMeP, SS-3D-LTP, and LBP. Their system has proven high performance. Given these benefits, it has also some disadvantages because the application of the scattering transformation is often time-consuming.

A new method characterized by encoding the relationship among neighbors was suggested in [39], where the method takes into account multiple image dimensions,

in order to encode information on the local depth. It also searches for directional edges and retrieves local information in many directions. This method was applied to three benchmark datasets MESSIDOR (Retinal images) [27], OASIS-MRI, and VIA/I-ELCAP (CT images) [125] respectively. The result of the experiments proved the superiority of this method to retrieve images accurately compared to other recent methods, where the other methods considered the one-dimensional image information in most literature for encoding, resulting in less accuracy for retrieval. For example, LMeP encodes the neighboring neighbors' relationship, and LBP encodes the neighbor's relationship to the center pixel. Although it has this superiority, this method has some drawbacks, including a high-dimensional feature vector that leads to an increase in computational complexity and a failure to fully extract the high-level semantic information from the image.

In [94] the authors proposed a method that utilizes two forms of profound neural network techniques: a supervised model of learning for Convolutional Neural Networks, (CNN) and a non-supervised model of training, is Stacked with Denoising Autoencoders (SDAE) to represent the most discriminatory features of medical images. In addition to indexing all images to look for related images. For the first time, they employed Preferential Learning technology (PL) to learn and train a model reference image capable of generating a list of the ranked image of similarities from the dataset of the medical images. They have carried out a great trial of three datasets of different complexity levels (NEMA-CT image dataset, image dataset OASIS-MR, dataset TCIA-CT) to investigate whether the method proposed would provide reliable output over medical datasets of various formats and types of imagery. The method is compared with the descriptor for the local binary pattern in [90]. The results of the experiments have shown that this method makes all datasets with high accuracy in comparison to other modern technologies and reduces computing

complexity by utilizing the Preference Learning (PL) model as a supervised ranking algorithm. This method also has a disadvantage, where the deep neural networks need to be fine-tuned with large-scale biomedical images acquired to extract more professional features.

The majority of the conventional methods in CBMIR perform fairly poorly because they often do not define the user's high-level semantic data. Finally, the rapid development of deep learning has brought higher performance and versatility than traditional descriptors in standard image retrieval tasks to features obtained from pre-trained CNN models. In [21] a new CBMIR (CNN) and Hash coding framework has been built. The new framework uses a Siamese network that uses image pairs (similar and dissimilar) to generate images of the same class and uses a contrasting loss function and weight sharing to learn how to create images of the same class. The CNN is used to extract features from each of the network branches and reduces via hash-mapping the dimension of the features vectors. During training, a novel loss function has been created, which would make the feature vectors more included and allow real value outputs to approximate the required binary values. The trained network generates the compact binary hash code of the image query and is comparable with the hash code of the dataset images during the retrieval process. Two medical image datasets were used: the (VIA/I-ELCAP) and the TCIA-CT [120]. Experimental tests have proven that this method is better than the current solutions to CNN and hash. Also, in comparison with the conventional CNN methods in which the Siamese network is paired with the hash system, the solution is superior.

The authors in [95] wanted to address the limitations of current handcrafted methods of feature extraction. So, a new feature learning method was thus suggested to address these limitations and extract robust features from medical scans. The proposed feature learning method is separated into two phases: the first is an IR-Net

(Image Reconstruction Network), which is meant to reconstruct the supplied input image using a unique deep network encoder-decoder. The suggested IR-Net was utilized as a new encoder network to encode the image to a variety of features and to reconstruct the input image using a new network of decoders. They have shown that the encoded features represent the input image in a strong way if the input image is properly reconstructed. In the second phase, these features that are encoded are used in the retrieval of medical images. In three medical image datasets, they assessed the output of the method, namely OASIS-MR, ILD [29], and VIA/I-ELCAP CT. In comparison with the existing methods, the new methodology exceeded them with the accuracy of the results, but this method has some limitations; it is not viable for multimodal medical image retrieval because they developed a CBMIR system that retrieves medical images that have a disease close to the input medical image.

A new deep hashing method is suggested in [99], in which the operations of deep extraction, binary code learning, and deep hash function learning are conducted under supervision. The discreet restricted objective function in the learning of the hash code is particularly iteratively optimized so that the binary code may be solved without the requirement for relaxation. Semantic similarity is retained in the meantime by exploring supervision information in discreet optimization in order to apply a principle of graphical regularization to the neighborhood structure of the training data to conserve them. Furthermore, for re-ranking the images a new scheme to refine the measure of similarity by considering the Euclidean Distance among realistic descriptors and their category data between these images is proposed to increase the ranking of the returned medical images with a specific hamming distance. This method was applied to the pulmonary nodule image dataset (LIDC-IDRI) [10], where the experiments show its superiority compared to the recent methods. However, it requires a sufficient amount of excellent quality labeled data, which are not

readily available, to create models.

In [98] a great solution was presented to enhance the efficiency of CBMIR with the application of a deep learning system for the purpose of CBMIR by training a Deep Convolution Neural Network (DCNN). Introduced two methods for medical image retrieval; one is a class prediction by the trained network to predict the class of the query image and then search for images that are appropriated in that specified class. The second method is to not combine any information on the query image class and check for appropriate images in the entire dataset. The presented method reduced the semantic gap by learning discriminating features directly from images. The suggested method can achieve a mean average precision of 69% and average classification accuracy of 99.77% in the retrieval task. The most suitable method to obtain a multimodal medical images dataset for various body parts is the method provided, but this method has an adverse effect whereby fully trained DCNN-based methods demand a large collection of labeled medical images to achieve their full potential and the feature dimension is 4096, that is highly time-consuming.

In [3], they suggested a method for solving the CBMIR system problem based on handmade features, i.e. inefficient content modeling, high dimensional features, and extraction of unnecessary and less useful features. An effective method has been used to represent images utilizing outstanding convolutions features. They worked on AlexNet's initial convolutional layer of kernels, demonstrating how visual content may be produced successfully by collecting color and texture features. Two measurements, texture sensitivity, and color sensitivity were used to study the properties of these kernels. On the basis of these discoveries, three different clusters divided the convolutional feature space, where there are a number of kernels with similar features in each of these clusters. These individual kernel sets are utilized to extract texture and color features from the image and to add them to a single

feature map termed the Spatial Maximum Activator Map (SMAP). The features of these maps have been recorded in a histogram so that their spatial layout information is also obtained with a structured pooling technique without increasing the dimensions of the features. This allows them to extract various strengths from the features of texture and color while clustering the space for convolution. Furthermore, the Spatial Maximal Activator map (SMAP) approach enables them to choose the most discriminative features. The Kvasir dataset [96] comparison of the classification performance indicates that this method has the best precision of 75.4, also the comparison of the retrieval performance with existing feature extract on methods shows that this method has the best retrieval performance with Average precision retrieval of 74.02 @50 but this method has a drawback, where a huge collection of labeled medical images is required for training to reach its full potential.

In [2], a new method has been introduced to reduce the semantic gap issue and achieve high accuracy in the retrieval of multimodal medical images over existing state-of-the-art methods. The suggested solution was tested with the two standard datasets, namely, the Kvasir dataset, and the IRMA 2009 dataset [133]. It was an efficient way to pick a perfect subset of features from pre-trained convolutional CNN layers. They found that the chosen subset of features is much better than the entire set of features in large datasets of medical image retrieval. They have also suggested a highly efficient method for depicting such features as compact binary codes using Fast Fourier Transform (FFT) to reduce the search space and enhance medical image retrieval efficiency. The vector of the feature is patronized as a one-dimensional signal and transformed by FFT into the frequency domain. Then, they transformed the feature vector into bits on the basis of their selection, using a simple linear transformation, by selecting the necessary number of frequency components and they used the binary codes obtained to provide hash codes and allow efficient retrieval of

big data sets. This proposed hash code method was compared extensively with many state-of-the-art methods, several of which outperformed significantly. However, with very shortcodes, the proposed method was not successful, mainly because the Fourier spectrum was easy to convert to binary codes and also the method did not obtain a high average accuracy at a 50-image accuracy in the Kvasir data set compared with the approach in [3], where that approach reached 47%, but in [3] 74.02%. The performance was in addition low compared to Spherical Hashing (SpH) [47], Sensitive Hashing (DSH) [54], and Random-Rotation Principal Component Analysis (PCA-RR) [40]. However, efficient FFT-based code computation time makes them more suitable for use.

In [42], an efficient and scalable histopathological image retrieval method was suggested for the learning of binary representation using Densely Connected Multi-Magnification Hashing (DCMMH), where in contrast to earlier works, which focuses only on one magnifications level. So, this method cooperates with many magnification levels, to learn hashing functions based on CNNs. There is a reciprocal learning guidance model focusing on image data of high-low magnification pairs. To fully utilize cross-magnification information, a dense-connected architecture has been implemented. This method has been applied to two benchmark datasets BreakHis [113], and PLOSONE [9]. In comparison to previous hasher methods based on handmade features, top performances were achieved in these experiments. So, this method contributed to decreasing clinician manual effort in the diagnosis of the histopathological images. However, retrieval by cross-magnification is not available. Although the binary codes for histopathological images are employed by various magnification levels. Furthermore, computational complexity is increased.

Also, the histopathological image contains more information that can be used in the early detection of breast cancers, and it's important to learn the compact

representation of histopathological image retrieval. So, in [43], a new framework was presented for learning binary histopathologic image codes. This framework is multi-Magnification Correlation Hashing (MMCH), in which both low-magnification and high-magnification data are used to learn discriminatory features. In particular, they built a patch-link graph with local patches extracted randomly from the labeled high-mag and low-mag images and propagated the link between high-mag images and low-mag patches to assess the semantic similarity of local patches. Similarities are then kept on the patch link graph and global labels to learn the binary codes of local patches. Furthermore, $L_{2,1}$ hashing function constraints have been introduced to select the more informative features from the original visual representation, where $L_{2,1}$ constraints get better discrimination for the feature. They used the BreakHis dataset and the PLOSONE dataset to perform the experiments, and the experiments have shown that their method has improved accuracy through a series of cell-based and comprehensive approaches to histopathological image retrieval tasks. Despite these advances, this method has a drawback: distance computation time is $O(n^2)$, and the query is not as effective as holistic-based methods when dealing with high-resolution data.

A Deep Convolutional Hash (DCH) method has been proposed in [105] for the encoding of images in binary codes. There is an embedded LBE layer available on the proposed network, and it may be trained "point-wise." The joint objective optimization function has also been constructed to encourage the network to learn discriminating representations from label information and decrease the gap between embedded features of low dimensions and necessary binary values. Binary encoding for new images has been provided through network propagation and the quantification of the LBE layer output. This method has been applied to the histopathology image dataset that has been created by them, including lung cancer images

and skeletal muscle images. From the experiments, superior performances in the histopathology image dataset demonstrate the efficacy of this method. However, the optimization is currently eased through the use of the non-linear (tanh) saturation function, which might restrict the efficiency of the binary codes learned. In addition, noise from labels would have a detrimental effect on model learning and the ultimate diagnosis of disease because no noise is assumed by the method for image labels.

The researchers in [20] suggested a method for improving medical decision support systems, i.e. breast cancer diagnosis and including AL/RF in the medical imaging region, by utilizing the Medical Active leaRning and Retrieval (MARRow). In order to minimize many of the disadvantages in terms of efficiency and quality of the system, the commission introduced an AL strategy that was suitably incorporated into the CBIR core mechanism. Since their AL strategy chooses a few more informative images based not just on similarity, but also on the degree of uncertainty and diversity. These selected images will benefit more from literature than those from the same class. This strategy has achieved a high level of improvement over the most advanced methods, achieving a precision increase of up to 87.3%. MARRow also showed a clear and acceptable rate of growth over the course of the learning iterations. The time of the learning procedure is reduced by this method, as it eliminates the involvement of the expert in the analysis process (eliminating up to 88%) as the expert doesn't have to annotate (correct) the labels on all samples as needed by the work of literature. They adopt a more rigorous class approach (i.e., provide fewer misclassifications) as more information samples are chosen for learning. However, complexity is increased in this method by its intrinsic inter-class similarity which leads to the process of fine-grain annotations to make a more complicated difference between relevant and irrelevant pictures.

In [131] a similarity measurement method is suggested, which integrates deep features for a mammogram. The images are processed first in advance to extract low-level data, including content and location, but before extracting location features, the registration is done using the standard image. Then, the following are developed to extract the deep features that are regarded to be very effective: the Stacked Auto-Encoder Network (SAE), the Convolutional Neural Network (CNN), and the Deep Belief Network (DBN). The similarity of the contents and the deep similarities are computed using the ED among the query and the images in the dataset separately, however, for the similarity of the location, the intersection ratio to the mass regions is calculated. The similarity between content, location, and deep similarity is ultimately merged to form the image fusion similarity. This procedure has been applied to 740 cases of mammograms, which are 740 MLO mammograms [91] from northeast china and their corresponding diagnostic reports. The experiments demonstrated the superiority in enhancing the efficacy of medical image retrieval and reducing the issue of the semantic gap. Although this superiority, this method has a high dimensional feature vector that will lead to increasing the computational complexity.

Individual features of a multi-phase CT image to indicate the region that is abnormal from the liver and apply it in an accurate retrieval of liver CT images were suggested in [78]. Two methods for retrieval of liver lesions were suggested. In the first method, for every imaging phase, individual features were selected to enhance the lesion class separability. In the second method, an 87×1 element vector was presented to represent different regions of an abnormal region. The correlation graph distance [102], which represents the feature vector's nonlinear structure, was also utilized to calculate the distance between an input image and other dataset images. Their method has been applied to 411 CT liver data consisting of three phases

[130, 136]. From the experiments, has been demonstrated that the use of a manifold image distance measurement scheme enhances CBMIR 's discrimination efficiency. Furthermore, when compared to recent methods, a more complicated feature vector enhances the outcomes. and on average, the overall recall of their results with the proposed features vector was increased by 7.5%. However, this method needs additional features to replace the existing huge multi-phase data feature vector and needs to obtain a larger dataset and test the method's efficacy.

Due to their different fracture locations, the retrieval of bone images affected by Avascular Necrosis (AN) is difficult. So, an effective methodology for retrieving AN image using Deep Belief CNN feature representation has been suggested in [118]. At first, pre-processing takes place for the input dataset. In addition, throughout the pre-processing step, image noise was eradicated with the use of the Median Filter (MF) and resized. Then, the features were transmuted to binary codes after being represented by the (DB-CNN). Ultimately, the calculation of the similarity was determined using the modified Hamming distance and retrieving the images. This method has been applied to the (Femur, Humerus, and Knee) of the AN images dataset, where the experiments showed its superiority compared to the recent methods. However, it is limited to this dataset, and getting a large number of data for analysis is very difficult. So, the real data is very limited in this research.

The images in the current methods are first identified and then categorized and this resulted in increasing the period for image retrieval as the entire dataset must be searched for performing the retrieval. So, a Grey Wolf Optimization-Support Vector Machine (GWO-SVM) method was proposed in [16] which classifies the query image's class at first. In addition, only the query image dataset can organize the retrieval process. The GWO algorithm makes a distinctly improved value for the SVM classification for the solved and optimized parameters. Therefore, it is evi-

dent that the retrieval rate is high after the classification is performed relative to the current methods by defining the retrieval levels. This has been applied to 7641 medical CT images that include 27 classes of different body parts. Tests have shown superior accuracy classification with a 97% and improvement in the accuracy of retrieval compared to current methods. However, this method has a high dimensional feature vector which considers a drawback.

In [143] the authors proposed an effective indexing and retrieval framework using deep features. The proposed system aims to produce the most effective and least parameterized hash codes by using image features. For this reason, deep features are obtained from medical images using the CNN model. The length of the acquired raw deep feature vectors for an image is relatively inefficient for retrieval speed. Feature reduction methods are used for the most effective reduction of the length of the deep feature vector. The main reason for producing a reduced class-driven hash code with feature selection algorithms is the drawbacks of medical image datasets. These drawbacks prevent the CNN output from being used directly as hash code. The performance of the proposed method is tested on NEMA MRI and NEMA CT datasets.

In [81] the authors proposed a DCNN model developed for classification and image retrieval purposes. Deep learning techniques obtained deep features from different abstraction levels represented using multiple hidden layers. This ultimately improves retrieval performance. They developed a framework based on a multi-layered CNN to represent medical images as deep features which are used for retrieval purposes. The proposed network's architecture is similar to Alexnet. The extracted deep features are used to calculate the similarity index between the images using distance measures and based on the similarity index, retrieval is done. By using a data augmentation technique the model achieved better retrieval accuracy.

Table 3.1: Summarization of CBMIR methods in terms of the model, dataset, classification technique, feature vector length, extraction and retrieval time, and the accuracy of each model.

Paper	Model	Dataset	Classification Technique	Feature Length	Extraction Time in(s)	Retrieval Time in(s)	Accuracy (%) ARP	ARR	F-measure	mAP
[33]	LWP	NEMA-CT	Manually	256	98.25	0.45	95.32 @10	31.33 @10	47.16 @10	-
		EXACT09-CT		256	≈ 150	≈ 1	83.00 @10	24.87 @10	38.27 @10	-
		TCIA-CT		256	≈ 140	≈ 0.8	88.42 @10	13.09 @10	22.8 @10	-
[66]	HCSCs	NEMA-CT	Manually	200	384.1	-	98.33 @10	33.64 @10	50.13 @10	-
		EXACT09-CT		200	384.1	-	91.50 @10	28.83 @10	43.84 @10	-
		TCIA-CT		200	384.1	-	95.12 @10	14.52 @10	25.20 @10	-
[65]	Filtering+ Partitioning+BoW	EXACT09-CT	Manually	-	-	-	92.97 @10	29.31 @10	44.57 @10	-
		TCIA-CT		-	-	-	96.32 @10	14.67 @10	25.46 @10	-
[64]	$ST-DC_{avg}$	EXACT09-CT	Manually	200	384.1	-	92.09 @10	29.07 @10	44.19 @10	-
	$ST-DC_{max}$	EXACT09-CT		200	384.1	-	91.93 @10	28.98 @10	44.07 @10	-
	$ST-CCA_h$	EXACT09-CT		-	14.98	-	91.92 @10	28.90 @10	43.98 @10	-
	$ST-CCA_v$	EXACT09-CT		-	14.98	-	93.35 @10	29.44 @10	44.76 @10	-
	$ST-DC_{avg}$	TCIA-CT		200	384.1	-	95.80 @10	14.65 @10	25.41 @10	-
	$ST-DC_{max}$	TCIA-CT		200	384.1	-	95.71 @10	14.58 @10	25.30 @10	-
	$ST-CCA_h$	TCIA-CT		-	14.98	-	96.33 @10	14.68 @10	25.48 @10	-
	$ST-CCA_v$	TCIA-CT		-	14.98	-	96.45 @10	14.71 @10	25.52 @10	-
[30]	LBDP	Emphysema-CT	Manually	256	2.39	0.06	≈ 50@100	≈ 89.5 @100	≈ 64.2 @100	-
		NEMA-CT		256	606.20	0.43	99.55 @5	75.83 @50	67.64 @50	-
		OASIS-MRI		256	56.78	0.33	≈ 58 @10	≈ 24 @10	≈ 34 @10	-
[52]	LMVCoP	OASIS-MRI	Manually	-	4.29	0.21	87.57 @10	53.21 @10	61.04 @10	-
[39]	$(MD)^2$MaMEP	MESSIDOR	Manually	2560	-	-	56.93 @5	-	-	-
		OASIS MRI		2560	-	-	62.49 @10	-	-	-
		VIA/I-ELCAP		2560	-	-	93.36 @10	60.40 @10	73.35 @10	-
[94]	CNN+PL	NEMA-CT	Manually	4096	-	-	≈ 90 @50	≈ 69 @50	≈ 78 @50	-
	SDAE+PL	NEMA-CT		256	7.77	3.72	≈ 90 @50	≈ 69 @50	≈ 78 @50	-
	CNN+PL	OASIS-MRI		4096	-	-	68.8 @50	32.8 @50	44.4 @50	-
	SDAE+PL	OASIS-MRI		256	2.88	3.11	61.1 @50	28.9 @50	39.24 @50	-
	CNN+PL	TCIA-CT		4096	-	-	≈ 90 @50	≈ 58 @50	≈ 70 @50	-
	SDAE+PL	TCIA-CT		256	6.24	4.50	91 @50	59.1 @50	71.7 @50	-
[63]	ZMs	Emphysema-CT	Manually	12	0.85	0.24	44.88 @100	79.48 @100	57.37 @100	60.23 @100
		Noisy Emphysema-CT					43.90 @100	78.50 @100	56.39 @100	59.16 @100
		OASIS-MRI	Manually	12	16.85	1.37	42.81 @100	40.52 @100	41.63 @100	49.41 @100
		Noisy OASIS-MRI					42.39 @100	39.54 @100	41.50 @100	48.72 @100
[74]	ANTIC	MESSIDOR	Manually	2048	-	-	-	≈ 54 @10	-	-
		NEMA-CT		2048	-	-	96.79 @10	-	-	-
		OASIS MRI		2048	-	-	≈ 56.5 @10	-	-	-
		VIA/I-ELCAP		2048	-	-	≈ 91 @10	≈ 9 @10	≈ 16.38 @10	-

To be continued

Table 3.1 (continued)

Paper	Model	Dataset	Classification Technique	Feature Length	Extraction Time in(s)	Retrieval Time in(s)	Accuracy (%) APR	ARR	F-measure	mAP
[115]	LDEBP+ LZMs	Emphysema-CT	Manually	85	11.85	0.49	≈ 52 @100	≈ 87 @100	≈ 65 @100	-
		OASIS-MRI		85	23.51	1.95	71.08 @100	30.41 @100	42.6 @100	-
		Brain-Tumor		85	185.72	2.81		≈ 51 @100	-	-
[1]	OFFMs	Emphysema-CT	Manually	16	0.26	0.11	45.31 @100	80.34 @100	57.94 @100	62.35 @100
		Noisy Emphysema-CT					45.03 @100	79.82 @100	57.58 @100	62.39 @100
		NEMA-CT	Manually	16	65.42	1.46	68.03 @100	97.24 @100	80.05 @100	99.12 @100
		Noisy NEMA-CT					48.51 @100	74.12 @100	58.64 @100	76.52 @100
		OASIS-MRI	Manually	16	7.49	0.67	42.52 @100	40.17 @100	41.31 @100	51.09 @100
		Noisy OASIS-MRI					33.76 @100	30.39 @100	31.98 @100	35.75 @100
		NEMA-MR	Manually	16	10.14	1.17	80.30 @100	100.00 @100	89.07 @100	100.00 @100
		Noisy NEMA-MR					56.82 @100	76.32 @100	65.14 @100	78.82 @100
[18]	TLBAP + LANADP	NEMA-CT	Manually	512	-	0.1160	96.34 @10	11.76 @10	20.96 @10	-
		OASIS-MRI		512	-	0.0714	63.39 @10	58.11 @10	60.64 @10	-
		VIA/I-ELCAP		512	-	0.1760	43.80 @100	43.81 @100	41.44 @100	-
[21]	Siamese network+ hash coding	TCIA-CT [120]	Siamese network+ hash coding	48-bits	-	0.382952	-	-	-	≈ 80 @100
		VIA/I-ELCAP		48-bits	-	0.382952	85.7 @50	-	-	81.63 @100
[95]	IR-Net	OASIS-MRI	Manually	512	-	-	70.45 @10	-	-	-
		VIA/I-ELCAP		512	-	-	99.36 @10	99.36 @10	99.36 @10	-
		ILD		512	-	-	80.87 @10	-	-	-
[45]	3D-LOZFP	NEMA-CT	Manually	192	1.73	-	84.37 @30	69.30 @30	76.1 @30	-
		TCIA-CT		192	1.73	-	93.17 @30	63.36 @30	75.43 @30	-
[99]	deep hashing	LIDC-IDRI	Radiologists	64	-	-	-	-	-	63.51 @100
[98]	DCNN	Multimodal	DCNN	4096	-	-	71.50 @21	-	-	69 @21
[3]	AlexNet and SMAP	Kvasir	SVM	96	-	-	74.02 @50	-	-	-
[2]	CNN and FFT	Kvasir	-	128-bits	-	-	46.5 @50	-	-	-
		IRMA 2009		128-bits	-	-	77.00 @50	-	-	-
[107]	BoVW and SSI	Multimodal	BoVW	-	-	-	80.50 @21	-	-	69.70 @21
		Kvasir		-	-	-	60.48 @50	-	-	-
		IRMA 2009		-	-	-	83.9 @50	-	-	-
[108]	Topic and Location	Multimodal	Topic and Location	24	-	-	88.75 @21	-	-	86.74 @21
		IRMA 2009		57	-	-	97.5 @10	-	-	-

To be continued

Table 3.1 (continued)

Paper	Model	Dataset	Classification Technique	Feature Length	Extraction Time in(s)	Retrieval Time in(s)	Accuracy (%) APR	ARR	F-measure	mAP
[6]	RFRM	Kvasir	Medical Experts	18	-	-	85 @10	-	-	-
[56]	DWT+ BTCs	Kvasir	Manually	-	-	-	55.3 @50	-	-	-
		VIA/I-ELCAP	-	-	-	-	86.7 @50	-	-	-
[117]	LTP +VLAD	KIMIA Path960	Manually	-	-	-	95 @10	20.21 @10	33.33 @10	-
				-	-	-	84.65 @20	36.02 @20	50.71 @20	-
[42]	DCMMH-All (40-100-200-400)	BreakHis 40×	Manually	16-bit	-	-	-	-	-	95.41 @5
		BreakHis 100×		16-bit	-	-	-	-	-	93.12 @5
		BreakHis 200×		16-bit	-	-	-	-	-	96.55 @5
		BreakHis 400×		16-bit	-	-	-	-	-	95.14 @5
	DCMMH	PLOSONE 200×		64-bit	-	-	-	-	-	32.46 @5
[43]	MCCH-40-200	BreakHis 40×	Manually	16-bit	-	-	-	-	-	92.52 @5
	MCCH-100-200	BreakHis 100×		16-bit	-	-	-	-	-	91.52 @5
	MCCH-200-400	BreakHis 200×		16-bit	-	-	-	-	-	89.74 @5
	MCCH-200-400	BreakHis 400×		16-bit	-	-	-	-	-	88.76 @5
	MCCH	PLOSONE 100×		64-bit	-	-	-	-	-	55.76 @5
	MCCH	PLOSONE 200×		64-bit	-	-	-	-	-	54.19 @5
[105]	DCH	Skeletal Muscle+ lung cancer	Manually	64-bit	-	-	-	-	-	96 @1000
[20]	LBP-JD	I_1-VIENNA	K-NN K-means	108	-	-	$\approx$ 93.7 @30	-	-	97.0 @30
	LBP-L_2	I_2-MIAS		108	-	-	100 @30	-	-	99.7 @30
	Zernike-X_2	I_3-DDSM		36	-	-	$\approx$ 49.4 @30	-	-	53.3 @30
	Zernike-X_2	I_4-MIAS-DDSM		36	-	-	$\approx$ 49.1 @30	-	-	52.7 @30
	Daubechies -L_∞	I_5-VIENNA-DDSM		16	-	-	$\approx$ 61.6 @30	-	-	63.1 @30
[131]	SAE+CNN+DBN	MLO mammograms	Manually	-	-	-	74.5 @15	85 @15	79.4 @15	-
[78]	manifold learning	CT liver data	Manually	87	-	16	P=83.6	R=84.2	83.9	-
[118]	DB-CNN	Femur images	Manually	-	-	11	P=92.3	R=81.56	86.6	-
		Humerus images		-	-	9	P=87.38	R=80.23	83.65	-
		Knee images		-	-	7	P=80.5	R=79.9	80.2	-
[138]	LDLP	Emphysema-CT Slice	Manually	25-bits	-	-	P=84.21	R=96.67	90.01	-
[16]	GWO-SVM	CT scan	GWO-SVM	-	-	-	-	-	-	90

3.3 Literature Reviews Analysis

In this chapter, we introduced a summary of modern methods of CBMIR that depend on low-level and high-level features. We also presented the advantages and disadvantages of each method, and we found that most of the methods that depend on low-level features shared some disadvantages. First, we failed to extract the high-level semantic information from the image, which resulted in a semantic gap problem still being found. Second, it has high computational complexity when it searches the whole dataset to compute the similarity. Third, the feature vector is very large. For methods of the high-level features, we found that most of them require an adequate number of good quality labeled data to generate the models, which are not easily available in the medical domain. In Chapter 4, we will introduce the proposed a method that exploits the pre-trained DCNN models to solve most of these disadvantages.

4. Proposed RbQE Method

Chapter 4

Proposed RbQE Method

In this chapter, we introduce a method for content-based medical image retrieval (RbQE) using a new query expansion method. The query expansion method depends on the top-rank images in the expansion process, which improves the chances of obtaining the most semantically related output. The proposed method uses deep features (AlexNet, VGGNets, and ResNets) to increase the accuracy of retrieval with less time for CBMIR systems. The expansion method is employed for the first time on deep features (AlexNet, VGGNets, and ResNets).

4.1 Medical Image Retrieval Method using New Query Expansion (RbQE)

The proposed method (RbQE) has two important parts, presented in Fig. 4.1. The first part is a feature extractor, and the second is a method for matching and retrieval of medical images. Three deep feature extractors have been used in the proposed method. Based on pre-trained models, the deep feature extractor can extract compact and high-level features to represent all images in the medical dataset. There are two aims of using the deep neural network instead of raw pixels in the

analysis of medical images: the first is to extract invariant features, which are more robust against different interferences like noise and changes in the light that appear during the generation of the medical images. Second, there is no need for the deep feature extractor to be retrained, if trained offline using a huge image dataset with heterogeneous formats, even in the case of analyzing various types and formats of medical images. As a result, the used deep model is likely to dramatically increase computational efficiency and lower calculation costs in comparison to other retrieval systems that also use deep models.

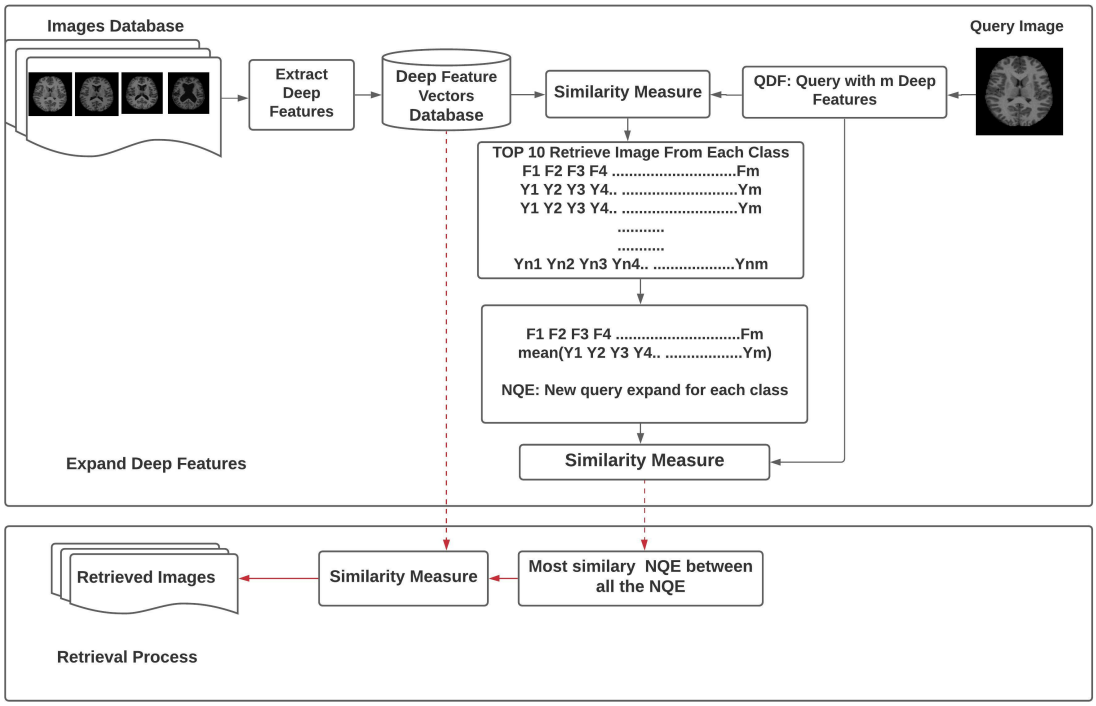

Figure 4.1: Illustration of the RbQE method.

In addition, the RbQE technique target improving the matching and retrieval in the CBMIR by expanding the deep features of the original query and the construction of a new query. The RbQE method is based on two search procedures: rapid and final. Firstly, a rapid search of the dataset is used to retrieve the top-ranked

images of the original query from each class of the dataset and then will form a New Query Expansion (NQE) for each class. Secondly, in the final search, the image that is most similar (NQE) to the query images is taken and used as the Final New Query Expansion (FNQE), which is one of the main benefits of our suggested method. Fig. 4.2 illustrates the proposed algorithm for the RbQE method. The next subsections provide more information on these feature extractors and query expansion methods.

Algorithm RbQE method

Input: Single random image form the dataset as query

Output: Top n similar images

1: **Start**
2: Extract deep features for all images in dataset
3: **for** $j = 1 : m$ **do** ▷ where m refer to all classes in dataset
4: Retrieve the top n similar images to query image
5: The values of the mean features are computed for the top n relevant images
6: Reformulate the NQE by using the values of the mean features
7: **end for**
8: end for
9: Retrieve the most NQE similar to query image and used it as FNQE
10: Use the FNQE to retrieve the top n similar images
11: end start

Figure 4.2: RbQE method algorithm.

4.1.1 Deep Feature Extraction

We used more robust and efficient deep features to extract more discriminatory and high-level features for medical images, thus minimizing interference problems. Deep learning has gained enormous popularity recently, with promising applications in a variety of areas [127]. The fundamental idea behind deep learning hasn't changed, despite the fact that numerous architectures have been proposed and put into practice: deep learning is a feature representation learning approach that concentrates on huge amounts of unprocessed image data and can use different levels of representation. This concept is stable in spite of several models of deep learning that have been suggested and implemented. Many levels of abstraction allow learning representations of data by computational models with many layers of processing L ($L > 1$), where after the input layer, each layer transforms the representation of the preceding layer into a more abstract representation, then you can obtain complex structures indirectly from large format imagery and ideally use them to create the original image or the image of the query after studying most of distinctive variations layer by layer.

For medical image retrieval, three types of supervised CNN learning models have been used as deep neural networks. CNN is a form of neural network that has been proposed to deal with images and obtain local features located in images. To work with high-resolution images, CNN has three properties: Firstly, each convolution kernel has a small function in depth that represents a feature that, despite its small size, can distinguish between different images. Secondly, the same deep feature may filter and obtain information from various locations in the input image because the same convolution kernel is employed by each convolution feature map. Finally, by subsampling from the convolutional layer to the pooling layer, the image's dimensionality is reduced and computing efficiency is increased. Fig. 4.3 shows a

previously trained Deep Convolutional Neural Network (DCNN) model (AlexNet); Fig. 4.4 shows the previously trained DCNN models (VGG-19); and ResNets, which have been trained offline in the ImageNet dataset [28] and contain millions of labeled images.

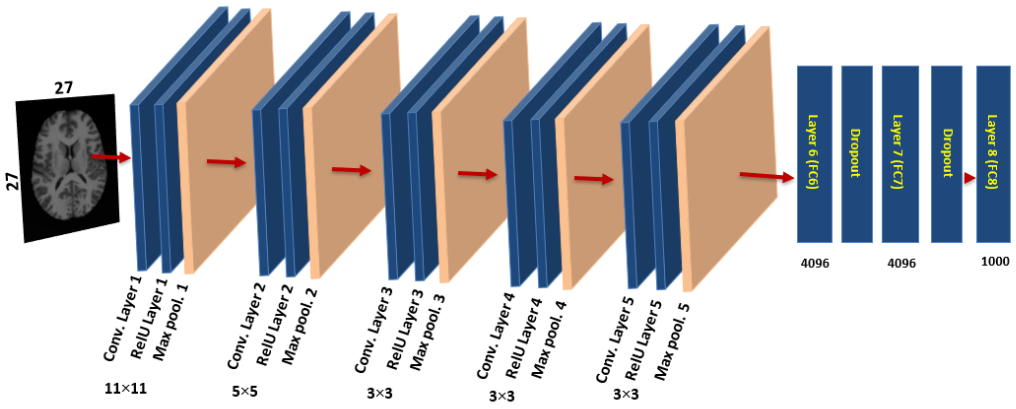

Figure 4.3: The pre-trained CNN (AlexNet) on ImageNet dataset.

Significantly deeper neural networks cannot be used in medical image processing. Since the small distinguishes among identical biomedical images with high-level features and most semantical are difficult to differentiate and with greater abstraction, the small disparity will go away. However, a small difference is particularly essential in biological images and may be applied precisely to discriminate biomedical images of several types, such as images from our research from the OASIS-MRI dataset used in our research. The AlexNet, as demonstrated in Fig. 4.3, is inspired by biological processes where the object is recognized from the low-level step by step to the semantic level, it is usually made up of four key components: Firstly, the convolutional layers, which are connected to a limited, mostly human visual system location by a convolutional kernel and considered the greatest highlight of AlexNet, Secondly, the activation functions are frequently followed by the convolutional layers, where the

ReLU (rectified linear unit) activation function is used to extract from the input signals the more complicated features. Thirdly, the dimensionality of the feature map is lessened by the pooling layers, while the convolutional layer sensitivity is decreased. Finally, at the conclusion of the AlexNet structure, the fully connected layer is combined to generate a feature vector, which provides the prediction result. By applying the backpropagation approach, the loss function between the prediction outcomes and ground truth is minimized using the AlexNet training procedure until the error loss is considerably reduced or a certain number of iterations have been completed. We have used the learned AlexNet as an extractor of biomedical frameworks, utilizing the fully connected layer-6. Our usage of completely connected layer-6 features, since various studies have shown that layer-6 features are more efficient than layer-7 features in biomedical image processing [12, 123, 106, 13, 23].

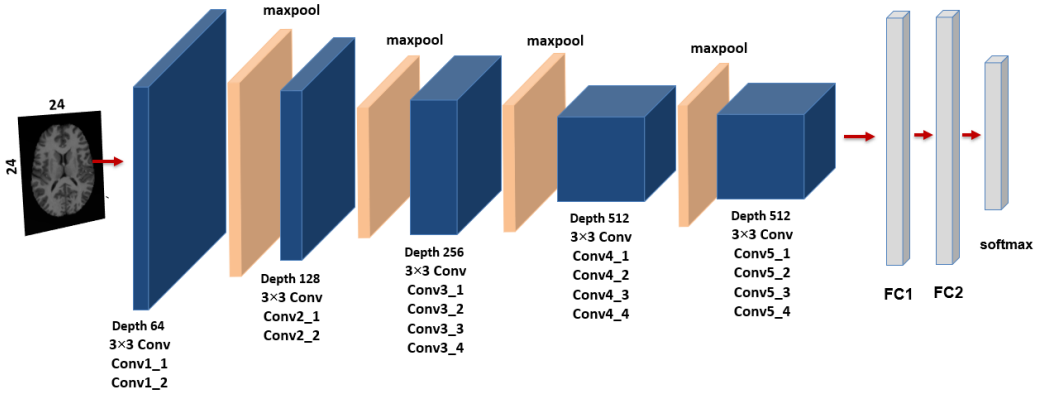

Figure 4.4: The pre-trained DCNN (VGG-19) on ImageNet dataset.

Convolutional networks with extremely deep layers that up to 19 layers (VGG-19) were employed as a feature extractor, where there are 16 convolutional layers and 3 FC (Fully Connected) layers, as shown in Fig. 4.4, where the number of channels is quite small, starting at 64 in the first layer and growing by a factor of two after each max-pooling layer until it reaches 512. In this network (VGG-

19), the image was transmitted through a stack of a convolutional layer that is a composite of filters with an extremely narrow receptive field 3×3 to gripe the notion of up/down, left/right, and center). The convolution stride was set to one pixel, and the spatial padding of convolution is 1 pixel 3×3 convolution layers. There are five max-pooling layers, each of which was conducted across a 2×2 pixel window with stride 2. Three Fully-Connected (FC) layers follow a stack of convolutional layers: the first two (FC1, FC2) have 4096 channels (features) apiece, while the third (FC3) has 1000 channels (features), and the soft-max layer is found in the final. Also, here we have utilized the fully connected layer-1 (FC1) of the VGG-19 as a feature vector extractor. There are also various studies in which we did show that FC1 features are more efficient than layer-2 (FC2) features in biomedical image processing, except in the TCIA-CT dataset, where FC2 features have achieved higher accuracy than FC1 features.

Residual networks (ResNets) [46] are utilized in this thesis 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, which 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 dataset [28] 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.

4.1.2 Query Expansion Method

As shown in Fig. 4.1, which describes the complete idea of the expansion of deep features for the original query and the reformulation of a new query for the final

search process, where after the rapid search using a Query with 4096 or 1000 Deep Features (QDF) of the original query to all deep features of images in the dataset, the RbQE technique employs the mean values of the deep feature values for images of top-ranked. From each class in the dataset, the top 10 similar images to the original query are retrieved, and the mean value of deep features for each of the top 10 is calculated. This process produces a number of NQEs equal to the number of classes in the dataset. After that, the most similar NQE to the original query will be taken as the final NQE (FNQE), and then the FNQE is used for the final search procedure. Table 4.1 provides a simple numerical example of building NQE, where the feature vector dimension for the image is 4096 for both descriptors AlexNet and VGGNets and 1000 for the ResNets descriptor.

Table 4.1: NQE based on mean values.

	F1	F2	F3	F4	F5	...	F4096
Img1	0.99	-15.05	-5.02	-41.11	-23.11	...	6.76
Img2	2.21	-12.71	-3.55	-43.14	-23.29	...	8.57
Img3	-1.29	-13.40	-7.85	-39.39	-28.45	...	8.29
Img4	-1.48	-11.07	-3.12	-37.64	-23.31	...	13.11
Img5	-6.69	-7.70	1.03	-40.86	-22.97	...	6.07
NQE	-1.252	-11.986	-3.702	-40.428	-24.226	...	8.56

5. Experimental Evaluation and Results

Chapter 5

Experimental Evaluation and Results

In this chapter, we introduce an analysis of the results of our experiments on five publicly available datasets with different modalities (TCIA-CT, EXACT09-CT, NEMA-CT, OASIS-MRI, and KIMIA Path960) and we present different CBMIR methods used for comparison with the proposed method RbQE.

5.1 Experimental Framework

The computational method that was used is presented in this section to compare the performance of the proposed method with other modern retrieval methods. The name and abbreviations of all methods used for comparison with the proposed method are presented in Table 5.1. We also show the similarity measurement that has been used.

5.1.1 Image Similarity Estimation

The similarity values in the proposed method are calculated with the Euclidean Distance (ED), which is used to calculate the similarity for both rapid and final searches. Let $\mathrm{NQE}_1 = (f_1, f_2, ..., f_n)$ and $\mathrm{NQE}_2 = (f_1^{'}, f_2^{'}, ..., f_n^{'})$, two feature vectors

Table 5.1: Name and Abbreviations of all methods used in the comparison.

S.No.	Abbreviation	Method Name	Reference	Year
1	LBP	Local Binary Pattern	[90]	1996
2	LTP	Local Ternary Pattern	[119]	2010
3	LDP	Local Derivative Pattern	[141]	2010
4	LTrP	Local Tetra Patterns	[82]	2012
5	AlexNet	Deep Convolutional Neural Networks	[60]	2012
6	LTCoP	Local Ternary Co-Occurrence Patterns	[84]	2013
7	LMeP	Local Mesh Patterns	[83]	2014
8	VGG-16	Visual Geometry Group	[110]	2014
9	LWP	Local Wavelet Pattern	[33]	2015
10	SS3D	Spherical Symmetric 3D Local Ternary Patterns	[87]	2015
11	MDLBP	Multichannel Decoded Local Binary Patterns	[34]	2016
12	ResNet	Residual Neural Network	[46]	2016
13	QWLD	Quaternionic Weber Local Descriptor	[67]	2017
14	HCSCs	Histogram of Compressed Scattering Coefficients	[66]	2017
15	ST-CCA [64]	Scattering Transform with Canonical Correlation Analysis	[64]	2018
16	MDMEP	Multi-dimensional multi-directional mask maximum edge pattern	[38]	2018
17	LTP+VLAD	Local Ternary Pattern and Vector of Locally Aggregated Descriptors	[117]	2019
18	IR-Net	Image Reconstruction Network	[95]	2020

with n dimension, the similarity is computed as follows:

$$ED\left(NQE_1, NQE_2\right) = \sqrt{\sum_{i=1}^{n}\left(f_i - f_i'\right)^2} \qquad \text{(Ch5-6)}$$

5.1.2 Performance Estimation

In the experiments, every image in the dataset is used as a query, and an image is only relevant if it falls under the same category as the query. The Average Retrieval Precision (ARP), Average Retrieval Rate (ARR), and F_{score} are three performance indicators used to assess each retrieval strategy.

$$precision[92] : P(q) = \frac{\text{Number of relevant images retrieved}}{\text{Number of images retrieved}} \qquad \text{(Ch5-7)}$$

$$recall[92] : R(q) = \frac{\text{Number of relevant images retrieved}}{\text{Number of relevant images in the dataset}} \quad \text{(Ch5-8)}$$

$$ARP(\%) = \frac{100}{\mid DB \mid} \sum_{i=1}^{|DB|} P(I_i) \quad \text{(Ch5-9)}$$

$$ARR(\%) = \frac{100}{\mid DB \mid} \sum_{i=1}^{|DB|} R(I_i) \quad \text{(Ch5-10)}$$

$$F_{score}(\%) = \frac{2 \times \text{ARP} \times \text{ARR}}{\text{ARP} + \text{ARR}} \quad \text{(Ch5-11)}$$

Where $\mid DB \mid$ indicates the count of all dataset images.

5.2 Experimental Results

This section includes various experiments that illustrate the efficiency of RbQE with ResNet-18, ResNet-50, ResNet-101, AlexNet, VGG-16, and VGG-19 as feature extractors and compare its results to those of the existing methods presented in Table 5.1. The RbQE method applied two different searching techniques: a rapid search for each dataset class using one query image selected from the dataset's image collection, where every image in the dataset is considered a query. Then, the final search is done using the final NQE (FNQE). Note that all searches are automated without user participation or suggestion, which is considered a strong point. The performance of the proposed method is compared to the modern methods, whether deep learning-based or not.

5.2.1 Retrieval Performance on TCIA-CT Dataset

The performance of the RbQE method is evaluated in the TCIA-CT dataset using six feature extractors: ResNet-18, ResNet-50, ResNet-101, AlexNet, VGG-16, and VGG-19. Table 5.2 shows the retrieval results in terms of ARP, ARR, and F_{score}. When compared to other methods, the suggested RbQE method using VGG-19 performs the best on the top 10 images and exceeds the RbQE with other feature extractors. The retrieval results of the proposed method are improved by 0.84%, 0.16%, 0.27% as compared with $ST - CCA_v$ in terms of ARP, ARR, and F_{score}. Fig. 5.1 exhibits the TCIA-CT dataset query outcomes of the RbQE method with VGG-19 features, which shows all the top 10 images in the same query image class.

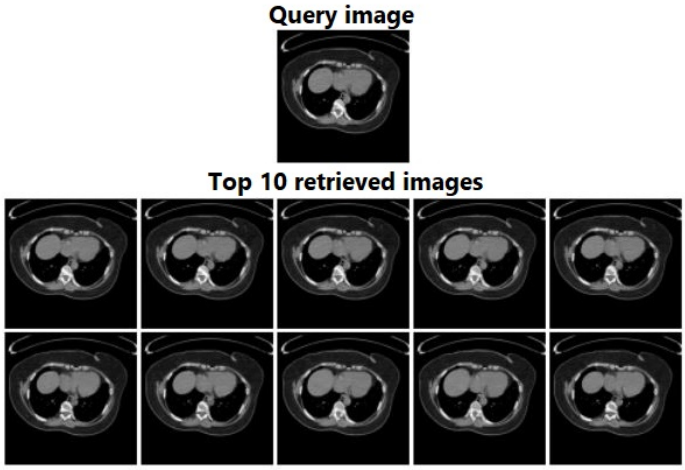

Figure 5.1: Retrieved images for a TCIA-CT dataset query using RbQE with VGG-19.

5.2.2 Retrieval Performance on EXACT09-CT Dataset

The comparison methods utilized in Subsection 5.2.1 are also taken into consideration here and are assessed using the same experimental parameters for the TCIA-CT dataset. The retrieval result of the RbQE with different feature extractors is

Table 5.2: Performance of different methods on TCIA-CT dataset with the top 10 matches considered.

Method	ARP	ARR	F_{score}
LBP [90]	66.91	9.74	17.00
LMeP [83]	73.71	10.77	18.79
LWP [33]	88.40	13.09	22.80
SS3D [87]	80.54	11.71	20.45
$HCSCs_h$ [66]	94.74	14.45	25.08
$HCSCs_v$ [66]	95.12	14.52	25.20
$ST-DC_{avg}$ [64]	95.80	14.65	25.41
$ST-DC_{max}$ [64]	95.71	14.58	25.30
$ST-CCA_h$ [64]	96.33	14.68	25.48
$ST-CCA_v$ [64]	96.45	14.71	25.52
RbQE with ResNet − 18	**96.84**	**14.78**	**25.65**
RbQE with ResNet − 50	**96.87**	**14.76**	**25.61**
RbQE with ResNet − 101	**96.04**	**14.66**	**25.43**
RbQE with VGG − 16	**96.92**	**14.80**	**25.68**
RbQE with VGG − 19	**97.29**	**14.87**	**25.79**
RbQE with AlexNet	**96.86**	**14.77**	**25.63**

shown in Table 5.3. In comparison to other methods, the features of the AlexNet descriptor with the RbQE method achieve the highest performance on the top 10 images and exceed the RbQE with other feature extractors, while all the descriptors with RbQE exceed the $ST-CCA_v$ method. The outcomes of the AlexNet with the RbQE method in relation to $ST-CCA_v$ in the ARP, ARR, and F_{score} ranges are improved by 4.86%, 1.64%, 2.47%. The results of the top 10 images obtained using the RbQE search technique with AlexNet features are shown in Fig. 5.2.

5.2.3 Retrieval Performance on NEMA-CT Dataset

We also use the NEMA-CT dataset to assess the performance of RbQE with different feature extractors and other modern methods. The proposed RbQE with VGG-19 features achieves the most satisfactory accuracy on the top 10 images and is superior

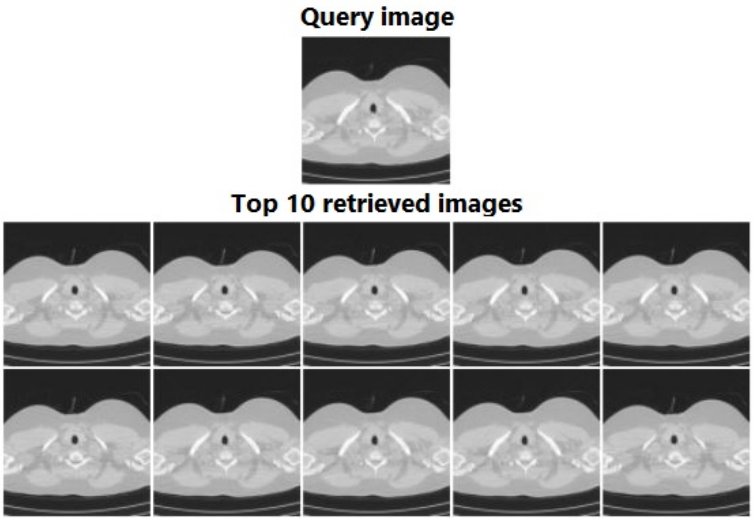

Figure 5.2: Retrieved images for an EXACT09-CT dataset query using RbQE with AlexNet.

Table 5.3: Performance of different methods on EXACT09-CT dataset with the top 10 matches considered.

Method	ARP	ARR	F_{score}
LBP [90]	65.03	19.51	30.02
LMeP [83]	63.23	18.91	29.11
LWP [33]	83.00	24.87	38.27
SS3D [87]	67.00	20.09	30.91
$HCSCs_h$ [66]	90.74	28.53	43.41
$HCSCs_v$ [66]	91.50	28.83	43.84
$ST - DC_{avg}$ [64]	92.09	29.07	44.19
$ST - DC_{max}$ [64]	91.93	28.98	44.07
$ST - CCA_h$ [64]	91.92	28.90	43.98
$ST - CCA_v$ [64]	93.35	29.44	44.76
RbQE with ResNet – 18	**96.06**	**30.33**	**46.1**
RbQE with ResNet – 50	**97.45**	**30.86**	**46.88**
RbQE with ResNet – 101	**96.79**	**30.53**	**46.42**
RbQE with VGG – 16	**98.16**	**31.06**	**47.19**
RbQE with VGG – 19	**96.99**	**30.63**	**46.56**
RbQE with AlexNet	**98.21**	**31.08**	**47.23**

to all other descriptors used by the RbQE. The retrieval results of the RbQE method with VGG-19 are improved by 1.24%, 0.18%, and 0.36% compared to the HCSCs method, as shown in Table 5.4. The result of the top 10 images for the query using

the RbQE method with VGG-19 features is shown in Fig. 5.3.

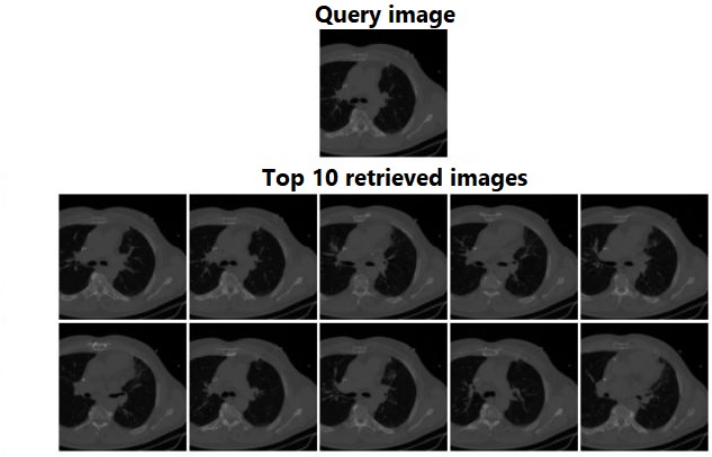

Figure 5.3: Retrieved images for a NEMA-CT dataset query using RbQE with VGG-19.

Table 5.4: Performance of different methods on NEMA-CT dataset with top 10 matches considered.

Method	ARP	ARR	F_{score}
LBP [90]	90.55	29.33	44.31
LTCoP [84]	92.15	30.31	45.62
LMeP [83]	93.09	30.62	46.08
LWP [33]	95.32	31.33	47.16
SS3D [87]	92.24	30.26	45.57
LTP [119]	92.00	30.23	45.51
LDP [141]	94.22	31.08	46.74
LTrP [82]	93.69	30.96	46.54
HCSCs [66]	98.33	33.64	50.13
RbQE with ResNet – 18	**99.05**	**33.54**	**50.12**
RbQE with ResNet – 50	**98.12**	**33.03**	**49.43**
RbQE with ResNet – 101	**98.58**	**33.24**	**49.71**
RbQE with VGG – 16	**99.14**	**33.56**	**50.15**
RbQE with VGG – 19	**99.57**	**33.82**	**50.49**
RbQE with AlexNet	**99.44**	**33.75**	**50.39**

5.2.4 Retrieval Performance on OASIS-MRI Dataset

The efficiency of the RbQE method with different feature extractors was also compared to IR-Net [95] in addition to some other modern methods. Table 5.5 presents the performance of the top 10 images in terms of ARP. In terms of average group-wise, the RbQE method with AlexNet, VGG-16, and VGG-19 features exceeds the IR-Net the RbQE with ResNet-18, ResNet-50, and ResNet-101 in terms of ARP, as shown in Table 5.6. On the top 10 images, the RbQE with AlexNet features performs with the highest level of accuracy compared to the RbQE with VGG-16, VGG-19, ResNet-18, ResNet-50, and ResNet-101. The results of the query using the RbQE method with AlexNet features are shown in Fig. 5.4.

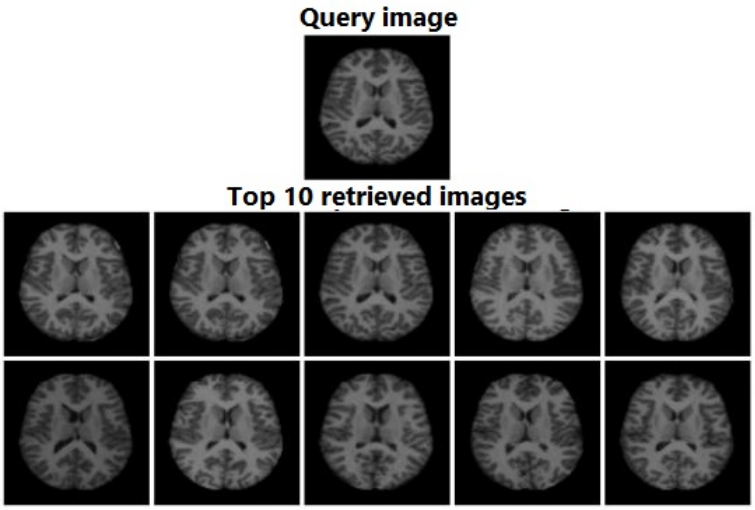

Figure 5.4: Retrieved images for an OASIS-MRI dataset query using RbQE with AlexNet.

5.2.5 Retrieval Performance on KIMIA Path960 Dataset

The Path960 KIMIA dataset is used in this experiment. We compared RbQE with different types of feature extractors against a modern medical image retrieval method

Table 5.5: Performance of different methods on OASIS dataset in terms of ARP for top 10 matches.

Method	Top1	Top2	Top3	Top4	Top5	Top6	Top7	Top8	Top9	Top10
LBP [90]	100	72.92	61.20	56.89	53.25	50.04	48.63	47.21	45.87	44.66
LTCoP [84]	100	73.87	62.79	57.66	54.49	51.58	49.37	47.57	46.93	45.82
SS3D [87]	100	68.53	56.77	51.54	47.89	45.68	43.77	42.37	41.44	40.45
LTP [119]	100	73.63	63.42	57.84	53.87	51.46	49.37	47.77	46.45	45.04
LTrP [82]	100	70.67	59.62	52.55	48.69	47.43	45.81	44.30	43.34	42.52
MDMEP [38]	100	81.47	73.87	70.19	67.84	66.71	65.52	63.9	63.37	62.49
AlexNet [60]	100	80.88	73.56	68.41	65.46	62.98	61.83	60.54	59.46	58.36
ResNet [46]	100	78.50	71.18	67.99	64.94	62.87	60.54	59.26	58.38	57.48
VGG-16 [110]	100	76.60	68.88	63.84	61.24	59.62	58.23	57.60	56.03	55.06
IR-Net [95]	100	83.37	78.62	76.37	74.68	73.44	72.21	71.62	70.89	70.45
RbQE with ResNet – 18	**100**	**100**	**99.92**	**97.27**	**91.07**	**87.69**	**83.68**	**81.26**	**78.33**	**76.77**
RbQE with ResNet – 50	**100**	**100**	**99.45**	**95.78**	**92.02**	**87.81**	**83.71**	**81.09**	**78.57**	**76.32**
RbQE with ResNet – 101	**100**	**100**	**99.13**	**94.00**	**88.84**	**84.56**	**80.59**	**78.27**	**76.43**	**74.73**
RbQE with VGG – 16	**100**	**100**	**99.92**	**98.22**	**94.54**	**91.05**	**87.72**	**84.68**	**81.13**	**79.88**
RbQE with VGG – 19	**100**	**100**	**99.92**	**98.28**	**95.01**	**91.61**	**88.77**	**85.69**	**83.27**	**80.83**
RbQE with AlexNet	**100**	**100**	**100**	**99.23**	**97.1**	**93.98**	**91.58**	**88.81**	**86.8**	**84.79**

Table 5.6: Performance of different methods on OASIS dataset in terms of ARP for group-wise.

Method	Group1	Group2	Group3	Group4	Average
LBP [90]	55.08	35.20	32.70	51.60	43.64
LTCoP [84]	50.08	41.08	34.16	55.19	45.12
SS3D [87]	44.19	39.02	35.73	41.42	40.08
LTP [119]	52.90	37.06	35.73	51.32	44.25
LTrP [82]	52.26	37.35	34.04	43.21	41.75
MDMEP [38]	69.52	50.59	48.31	77.64	62.49
AlexNet [60]	68.87	43.73	41.01	74.72	57.08
ResNet [46]	69.11	45.59	44.27	66.42	56.35
VGG-16 [110]	70.73	44.61	38.09	61.04	53.62
IR-Net [95]	77.10	55.29	57.19	88.40	69.49
RbQE with ResNet – 18	**87.18**	**65.39**	**69.33**	**81.79**	**75.92**
RbQE with ResNet – 50	**84.76**	**65.69**	**68.09**	**83.58**	**75.53**
RbQE with ResNet – 101	**83.87**	**57.35**	**72.02**	**83.02**	**74.07**
RbQE with VGG – 16	**91.21**	**79.41**	**60.79**	**83.11**	**78.63**
RbQE with VGG – 19	**91.45**	**80.88**	**62.25**	**83.96**	**79.64**
RbQE with AlexNet	**94.44**	**76.67**	**76.97**	**87.92**	**84.00**

on a recent histopathological image, LTP+VLAD [116] in addition to some other methods as explained in [116]. The retrieval performance is presented in Tables 5.7, 5.8, and 5.9 in terms of ARP, ARR, and F_{score} respectively. It is clear from the tables that RbQE with ResNet-50 is superior to methods on the top 5, 10, 15,

and 20 images but does not exceed RbQE with ResNet-101 on the top 10 images. The result of the retrieval for RbQE with ResNet-50 is improved by 2.14%, 1.94%, 6.74%, and 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. 5.5, 5.6, 5.7, and 5.8 present the query results of the top 5, 10, 15, and 20 images, respectively for the RbQE with ResNet-50 over the KIMIA Path960 dataset. It appears that the images retrieved for the query are from the same category as the query.

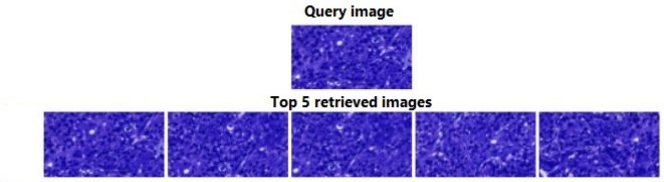

Figure 5.5: Retrieving the top 5 images for a KIMIA Path960 dataset query using RbQE with ResNet-50.

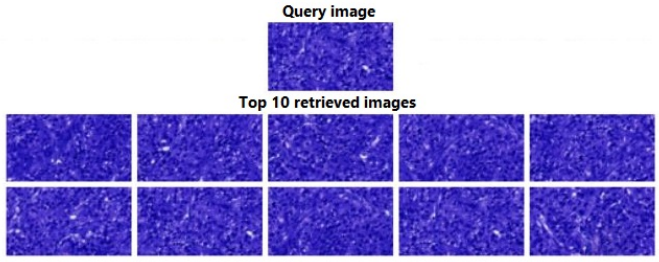

Figure 5.6: Retrieving the top 10 images for a KIMIA Path960 dataset query using RbQE with ResNet-50.

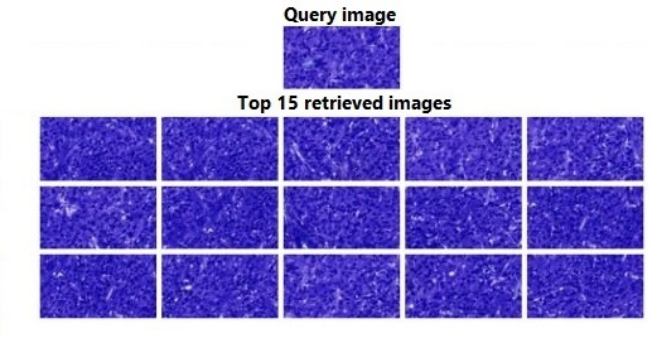

Figure 5.7: Retrieving the top 15 images for a KIMIA Path960 dataset query using RbQE with ResNet-50.

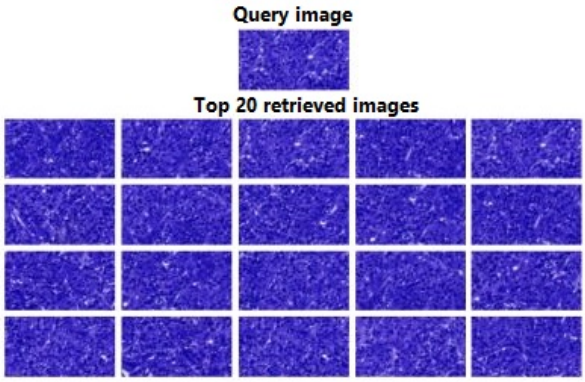

Figure 5.8: Retrieving the top 20 images for a KIMIA Path960 dataset query using RbQE with ResNet-50.

5.3 Time Complexity

The feature extraction average time, retrieval average time, and total CPU time in seconds are shown in Table 5.10 using the proposed RbQE with the different feature extractions methods (VGG-16, VGG-19, AlexNet, ResNet-18, ResNet-50, and ResNet-101) over each dataset (TCIA-CT, EXACT09-CT, NEMA-CT, OASIS-MRI, and KIMIA Path960). All the experiments were performed on a computer, with Intel(R) Core(TM) i7-4510U CPU @2.00 GHz processor, 8 GB RAM, and 64-bit

Table 5.7: ARP for various methods on the KIMIA Path960 dataset for the top 5, 10, 15, and 20 retrieved images.

Method	Top5	Top10	Top15	Top20
LBP [90]	88.74	84.05	80.50	75.89
HCSCs [66]	47.48	42.21	37.50	31.54
MDLBP [34]	86.87	81.71	76.27	70.61
QWLD [67]	41.44	37.46	30.89	23.34
AlexNet [60]	93.10	91.04	85.01	81.98
VGG-16 [110]	94.12	92.40	85.41	82.24
LTP+VLAD [117]	96.92	95.00	88.57	84.65
RbQE with ResNet – 18	**98.6**	**94.99**	**92.34**	**89.48**
RbQE with ResNet – 101	**99.04**	**96.96**	**94.69**	**92.89**
RbQE with ResNet – 50	**99.06**	**96.94**	**95.31**	**94.07**
RbQE with VGG – 16	**98.6**	**95.32**	**92.34**	**89.48**
RbQE with VGG – 19	**98.6**	**93.66**	**92.34**	**89.48**
RbQE with AleNet	**98.69**	**95.04**	**91.89**	**88.96**

Table 5.8: ARR for various methods on the KIMIA Path960 dataset for the top 5, 10, 15, and 20 retrieved images.

Method	Top5	Top10	Top15	Top20
LBP [90]	9.44	17.88	25.69	32.29
HCSCs [66]	5.05	8.98	11.97	13.42
MDLBP [34]	9.24	17.38	24.34	30.05
QWLD [67]	4.41	7.97	9.86	9.93
AlexNet [60]	9.90	19.37	27.13	34.88
VGG-16 [110]	10.01	19.66	27.26	34.99
LTP+VLAD [117]	10.31	20.21	28.27	36.02
RbQE with ResNet – 18	**10.49**	**20.21**	**29.47**	**38.08**
RbQE with ResNet – 101	**10.54**	**20.63**	**30.22**	**39.53**
RbQE with ResNet – 50	**10.54**	**20.63**	**30.42**	**40.03**
RbQE with VGG – 16	**10.49**	**20.28**	**29.47**	**38.08**
RbQE with VGG – 19	**10.48**	**19.93**	**29.47**	**38.08**
RbQE with AleNet	**10.49**	**20.22**	**29.33**	**87.36**

Windows 10 Enterprise LTSC operating system. The total CPU time of ResNet-18 is less than that of VGG-16, VGG-19, ResNet-18, ResNet-50, and ResNet-101. The retrieval times of VGG-16, VGG-19, and AlexNet are equal on the same dataset

Table 5.9: F_{score} for various methods on the KIMIA Path960 dataset for the top 5, 10, 15, and 20 retrieved images.

Method	Top5	Top10	Top15	Top20
LBP [90]	17.06	29.49	38.95	45.30
HCSCs [66]	9.13	14.81	18.15	18.83
MDLBP [34]	16.70	28.66	36.90	42.16
QWLD [67]	7.97	13.14	14.95	13.93
AlexNet [60]	17.89	31.94	41.13	48.94
VGG-16 [110]	18.24	19.66	32.42	49.09
LTP+VLAD [117]	18.64	33.32	42.86	50.54
RbQE with ResNet – 18	**18.96**	**33.33**	**44.68**	**53.42**
RbQE with ResNet – 101	**19.05**	**34.02**	**45.82**	**55.46**
RbQE with ResNet – 50	**19.05**	**34.01**	**46.12**	**56.16**
RbQE with VGG – 16	**18.96**	**33.45**	**44.68**	**53.42**
RbQE with VGG – 19	**18.96**	**32.86**	**44.68**	**53.42**
RbQE with AleNet	**18.98**	**33.35**	**44.26**	**53.11**

because they have the same dimension for the feature vector (4096). Also, the retrieval times of ResNet-18, ResNet-50, and ResNet-101 are equal on the same dataset because they have the same dimension of feature vector (1000).

Table 5.10: CPU elapse time (sec) for proposed RbQE with the different features extractions methods over all five test datasets.

Dataset	Feature Extraction Method	Feature Extraction Time (A) (sec)	Retrieval Time (B) (sec)	Total CPU time (A + B) (sec)
TCIA-CT	VGG-16	1.69	3.99	5.68
	VGG-19	1.84	3.99	5.83
	AlexNet	1.09	3.99	5.08
	ResNet-18	1.17	1.15	2.32
	ResNet-50	1.68	1.15	2.83
	ResNet-101	3.9	1.15	5.05
EXACT09-CT	VGG-16	1.69	4.93	6.62
	VGG-19	1.84	4.93	6.77
	AlexNet	1.09	4.93	6.02
	ResNet-18	1.17	1.6	2.77
	ResNet-50	1.68	1.6	3.28
	ResNet-101	3.9	1.6	5.5
NEMA-CT	VGG-16	1.69	2.62	4.31
	VGG-19	1.84	2.62	4.46
	AlexNet	1.09	2.62	3.71
	ResNet-18	1.17	0.75	1.92
	ResNet-50	1.68	0.75	2.43
	ResNet-101	3.9	0.75	4.65
OASIS-MRI	VGG-16	1.69	2.5	4.19
	VGG-19	1.84	2.5	4.34
	AlexNet	1.09	2.5	3.59
	ResNet-18	1.17	0.76	1.93
	ResNet-50	1.68	0.76	2.44
	ResNet-101	3.9	0.76	4.66
KIMIA Path960	VGG-16	1.69	2.5	4.19
	VGG-19	1.84	2.5	4.34
	AlexNet	1.09	2.5	3.59
	ResNet-18	1.17	2.08	3.25
	ResNet-50	1.68	2.08	3.76
	ResNet-101	3.9	2.08	5.98

6. Conclusion and Future Work

Chapter 6

Conclusion and Future Work

In this chapter, we first conclude what has been done in this thesis in Section 6.1. Section 6.2 presents our insights to improve the accuracy of a content-based medical image retrieval system.

6.1 Conclusions

An efficient method (RbQE) has been presented for the retrieval of CT, MRI, and histopathological images. RbQE method benefits from pre-trained deep convolutional neural networks (AlexNet, VGGNets, and ResNets) as extractors of the deep and high-level features. RbQE method relies on query expansion, where this expansion method is considered a fully automatic process that only requires the user to enter the query image without any involvement or feedback. RbQE method consists of two basic searching procedures: a rapid search and a final search. In the rapid search, the original query is expanded by retrieving the top-ranked images from each class and is used to reformulate the query by calculating the mean values for deep features of the top-ranked images, resulting in a new query for each class. In the final search, the new query that is most similar to the original query will be used for retrieval from the database. This method was tested on five different formats

of standard and publicly available datasets: TCIA-CT, EXACT09-CT, NEMA-CT, OASIS-MRI, and KIMIA Path960. In terms of ARP, ARR, and F_{score}. Experimental results show that the proposed method exceeds the compared methods by 0.84%, 4.86%, 1.24%, 14.34%, and 1.96% in average retrieval precision (ARP) on the retrieval of the top ten, for TCIA-CT, EXACT09-CT, NEMA-CT, OASIS-MRI, and KIMIA Path960, respectively. But for KIMIA Path960, the ARPs for the top five, fifteen, and twenty exceed the compared methods by 2.14%, 6.74%, and 9.42%, respectively.

6.2 Future Work

In the future, we can use new extractors like Stacked Denoising Autoencoders (SDAE) that can describe the medical image more accurately and help increase the accuracy of retrieval. Also, we can reduce the time of retrieval by using an indexing method to index all the images and find similarly referenced images. We will try to improve the dimension of feature vectors.

Bibliography

[1] Ashutosh Aggarwal, Suchita Sharma, Karamjeet Singh, Harpreet Singh, and Sunil Kumar. A new approach for effective retrieval and indexing of medical images. *Biomedical Signal Processing and Control*, 50:10–34, 2019.

[2] Jamil Ahmad, Khan Muhammad, and Sung Wook Baik. Medical image retrieval with compact binary codes generated in frequency domain using highly reactive convolutional features. *Journal of medical systems*, 42(2):1–19, 2018.

[3] Jamil Ahmad, Khan Muhammad, Mi Young Lee, and Sung Wook Baik. Endoscopic image classification and retrieval using clustered convolutional features. *Journal of medical systems*, 41(12):1–12, 2017.

[4] Jamil Ahmad, Muhammad Sajjad, Irfan Mehmood, Seungmin Rho, and Sung Wook Baik. Saliency-weighted graphs for efficient visual content description and their applications in real-time image retrieval systems. *Journal of Real-Time Image Processing*, 13(3):431–447, nov 2015.

[5] Jamil Ahmad, Muhammad Sajjad, Seungmin Rho, and Sung Wook Baik. Multi-scale local structure patterns histogram for describing visual contents in social image retrieval systems. *Multimedia Tools and Applications*, 75(20):12669–12692, mar 2016.

[6] Ali Ahmed. Implementing relevance feedback for content-based medical image retrieval. *IEEE Access*, 8:79969–79976, 2020.

[7] Alexandra Alfanindya, Noramiza Hashim, and Chikannan Eswaran. Content based image retrieval and classification using speeded-up robust features (surf) and grouped bag-of-visual-words (gbovw). In *2013 International Conference on Technology, Informatics, Management, Engineering and Environment*, pages 77–82. IEEE, 2013.

[8] Grigorios Antonellis, Andreas G Gavras, Marios Panagiotou, Bruce L Kutter, Gabriele Guerrini, Andrew C Sander, and Patrick J Fox. Shake table test of large-scale bridge columns supported on rocking shallow foundations. *Journal of Geotechnical and Geoenvironmental Engineering*, 141(5):04015009, 2015.

[9] Teresa Araújo, Guilherme Aresta, Eduardo Castro, José Rouco, Paulo Aguiar, Catarina Eloy, António Polónia, and Aurélio Campilho. Classification of breast cancer histology images using convolutional neural networks. *PloS one*, 12(6):e0177544, 2017.

[10] Samuel G Armato III, Geoffrey McLennan, Luc Bidaut, Michael F McNitt-Gray, Charles R Meyer, Anthony P Reeves, Binsheng Zhao, Denise R Aberle, Claudia I Henschke, Eric A Hoffman, et al. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. *Medical physics*, 38(2):915–931, 2011.

[11] Sangeeta Arora, Yadwinder S Brar, and Sheo Kumar. Haar wavelet transform for solution of image retrieval. *signal*, 2(2):1, 2001.

[12] Yaniv Bar, Idit Diamant, Lior Wolf, and Hayit Greenspan. Deep learning with non-medical training used for chest pathology identification. mar 2015.

[13] Yaniv Bar, Idit Diamant, Lior Wolf, Sivan Lieberman, Eli Konen, and Hayit Greenspan. Chest pathology identification using deep feature selection with non-medical training. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 6(3):259–263, may 2016.

[14] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). *Computer vision and image understanding*, 110(3):346–359, 2008.

[15] Yoshua Bengio, Aaron C Courville, and Pascal Vincent. Unsupervised feature learning and deep learning: A review and new perspectives. *CoRR, abs/1206.5538*, 1:2012, 2012.

[16] Renita D Benyl and Christopher C Seldev. Novel real time content based medical image retrieval scheme with gwo-svm. *Multimedia Tools and Applications*, 79(23-24):17227–17243, 2020.

[17] Kiran Ashok Bhandari and R Manthalkar Ramchandra. An innovative remote sensing image retrieval techniques based on haar wavelet-ltrp and anfis. *Procedia Computer Science*, 79:391–401, 2016.

[18] Ranjit Biswas, Sudipta Roy, and Debraj Purkayastha. An efficient content-based medical image indexing and retrieval using local texture feature descriptors. *International Journal of Multimedia Information Retrieval*, 8(4):217–231, 2019.

[19] Brain-Tumor Database. [online], 2017.

[20] Rafael S Bressan, Pedro H Bugatti, and Priscila TM Saito. Breast cancer diagnosis through active learning in content-based image retrieval. *Neurocomputing*, 357:1–10, 2019.

[21] Yiheng Cai, Yuanyuan Li, Changyan Qiu, Jie Ma, and Xurong Gao. Medical image retrieval based on convolutional neural network and supervised hashing. *IEEE Access*, 7:51877–51885, 2019.

[22] D Abraham Chandy, A Hepzibah Christinal, Alwyn John Theodore, and S Easter Selvan. Neighbourhood search feature selection method for content-based mammogram retrieval. *Medical & biological engineering & computing*, 55(3):493–505, 2017.

[23] Francesco Ciompi, Bartjan de Hoop, Sarah J. van Riel, Kaman Chung, Ernst Th. Scholten, Matthijs Oudkerk, Pim A. de Jong, Mathias Prokop, and Bram van Ginneken. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2d views and a convolutional neural network out-of-the-box. *Medical Image Analysis*, 26(1):195–202, dec 2015.

[24] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, Lawrence Tarbox, and Fred Prior. The cancer imaging archive (TCIA): Maintaining and operating a public information repository. *Journal of Digital Imaging*, 26(6):1045–1057, jul 2013.

[25] Susan Craw. *Manhattan Distance*, pages 639–639. Springer US, Boston, MA, 2010.

[26] Per-Erik Danielsson. Euclidean distance mapping. *Computer Graphics and image processing*, 14(3):227–248, 1980.

[27] Etienne Decencière, Xiwei Zhang, Guy Cazuguel, Bruno Lay, Béatrice Cochener, Caroline Trone, Philippe Gain, Richard Ordonez, Pascale Massin, Ali

Erginay, et al. Feedback on a publicly distributed image database: the messidor database. *Image Analysis & Stereology*, 33(3):231–234, 2014.

[28] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. jun 2009.

[29] Adrien Depeursinge, Alejandro Vargas, Alexandra Platon, Antoine Geissbuhler, Pierre-Alexandre Poletti, and Henning Müller. Building a reference multimedia database for interstitial lung diseases. *Computerized medical imaging and graphics*, 36(3):227–238, 2012.

[30] Shiv Ram Dubey, Satish Kumar Singh, and Rajat Kumar Singh. Local bit-plane decoded pattern: a novel feature descriptor for biomedical image retrieval. *IEEE Journal of Biomedical and Health Informatics*, 20(4):1139–1147, 2015.

[31] Shiv Ram Dubey, Satish Kumar Singh, and Rajat Kumar Singh. Local diagonal extrema pattern: a new and efficient feature descriptor for ct image retrieval. *IEEE Signal Processing Letters*, 22(9):1215–1219, 2015.

[32] Shiv Ram Dubey, Satish Kumar Singh, and Rajat Kumar Singh. Local wavelet pattern: A new feature descriptor for image retrieval in medical CT databases. *IEEE Transactions on Image Processing*, 24(12):5892–5903, dec 2015.

[33] Shiv Ram Dubey, Satish Kumar Singh, and Rajat Kumar Singh. Local wavelet pattern: a new feature descriptor for image retrieval in medical ct databases. *IEEE Transactions on Image Processing*, 24(12):5892–5903, 2015.

[34] Shiv Ram Dubey, Satish Kumar Singh, and Rajat Kumar Singh. Multichannel decoded local binary patterns for content-based image retrieval. *IEEE Transactions on Image Processing*, 25(9):4018–4032, sep 2016.

[35] John P Eakins. Towards intelligent image retrieval. *Pattern Recognition*, 35(1):3–14, 2002.

[36] Emphysema-CT Database. [online], 2010.

[37] Kuo-Chin Fan and Tsung-Yung Hung. A novel local pattern descriptor?local vector pattern in high-order derivative space for face recognition. *IEEE transactions on image processing*, 23(7):2877–2891, 2014.

[38] G. M. Galshetwar, L. M. Waghmare, A. B. Gonde, and S. Murala. Multi-dimensional multi-directional mask maximum edge pattern for bio-medical image retrieval. *International Journal of Multimedia Information Retrieval*, 7(4):231–239, jun 2018.

[39] GM Galshetwar, Laxman Madhavrao Waghmare, Anil Balaji Gonde, and Subrahmanyam Murala. Multi-dimensional multi-directional mask maximum edge pattern for bio-medical image retrieval. *International Journal of Multimedia Information Retrieval*, 7(4):231–239, 2018.

[40] Yunchao Gong, Svetlana Lazebnik, Albert Gordo, and Florent Perronnin. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. *IEEE transactions on pattern analysis and machine intelligence*, 35(12):2916–2929, 2012.

[41] Patrick JF Groenen, Rudolf Mathar, and Willem J Heiser. The majorization approach to multidimensional scaling for minkowski distances. *Journal of Classification*, 12:3–19, 1995.

[42] Yun Gu and Jie Yang. Densely-connected multi-magnification hashing for histopathological image retrieval. *IEEE journal of biomedical and health informatics*, 23(4):1683–1691, 2018.

[43] Yun Gu and Jie Yang. Multi-level magnification correlation hashing for scalable histopathological image retrieval. *Neurocomputing*, 351:134–145, 2019.

[44] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.

[45] Rakcinpha Hatibaruah, Vijay Kumar Nath, and Deepika Hazarika. Biomedical ct image retrieval using 3d local oriented zigzag fused pattern. In *2020 National Conference on Communications (NCC)*, pages 1–6. IEEE, 2020.

[46] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, jun 2016.

[47] Jae-Pil Heo, Youngwoon Lee, Junfeng He, Shih-Fu Chang, and Sung-Eui Yoon. Spherical hashing. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2957–2964. IEEE, 2012.

[48] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine*, 29(6):82–97, 2012.

[49] Kyung Hoon Hwang, Haejun Lee, and Duckjoo Choi. Medical image retrieval: Past and present. *Healthcare Informatics Research*, 18(1):3, 2012.

[50] Anil K Jain and Aditya Vailaya. Image retrieval using color and shape. *Pattern recognition*, 29(8):1233–1244, 1996.

[51] K Jaspreet and K Rajneet. Biomedical image denoising using symlet wavelet with wiener filter. *Int. J. Eng. Res. Appl*, 3:548–550, 2013.

[52] A Jenitta and R Samson Ravindran. Image retrieval based on local mesh vector co-occurrence pattern for medical diagnosis from mri brain images. *Journal of medical systems*, 41(10):1–10, 2017.

[53] Gnanasigamony Wiselin Jiji and Peter Savariraj Johnson Durai Raj. Content-based image retrieval in dermatology using intelligent technique. *IET Image Processing*, 9(4):306–317, 2015.

[54] Zhongming Jin, Cheng Li, Yue Lin, and Deng Cai. Density sensitive hashing. *IEEE transactions on cybernetics*, 44(8):1362–1371, 2013.

[55] Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 1725–1732, 2014.

[56] Hany Kasban and DH Salama. A robust medical image retrieval system based on wavelet optimization and adaptive block truncation coding. *Multimedia Tools and Applications*, 78(24):35211–35236, 2019.

[57] Palwinder Kaur and Rajesh Kumar Singh. A panoramic view of content-based medical image retrieval system. In *2020 International Conference on Intelligent Engineering and Management (ICIEM)*. IEEE, jun 2020.

[58] Manesh Kokare, B N Chatterji, and P K Biswas. A survey on current content based image retrieval methods. *IETE Journal of Research*, 48(3-4):261–271, may 2002.

[59] Sven Kosub. A note on the triangle inequality for the jaccard distance. *Pattern Recognition Letters*, 120:36–38, 2019.

[60] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[61] K Kranthi Kumar and T Venu Gopal. A novel approach to self order feature reweighting in cbir to reduce semantic gap using relevance feedback. In *2014 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2014]*, pages 1437–1442. IEEE, 2014.

[62] Meghana Dinesh Kumar, Morteza Babaie, Shujin Zhu, Shivam Kalra, and H. R. Tizhoosh. A comparative study of CNN, BoVW and LBP for classification of histopathological images. nov 2017.

[63] Yogesh Kumar, Ashutosh Aggarwal, Shailendra Tiwari, and Karamjeet Singh. An efficient and robust approach for biomedical image retrieval using zernike moments. *Biomedical Signal Processing and Control*, 39:459–473, 2018.

[64] Rushi Lan, Huadeng Wang, Si Zhong, Zhenbing Liu, and Xiaonan Luo. An integrated scattering feature with application to medical image retrieval. *Computers & Electrical Engineering*, 69:669–675, 2018.

[65] Rushi Lan, Si Zhong, Zhenbing Liu, Zhuo Shi, and Xiaonan Luo. A simple texture feature for retrieval of medical images. *Multimedia Tools and Applications*, 77(9):10853–10866, 2018.

[66] Rushi Lan and Yicong Zhou. Medical image retrieval via histogram of compressed scattering coefficients. *IEEE journal of biomedical and health informatics*, 21(5):1338–1346, 2016.

[67] Rushi Lan, Yicong Zhou, and Yuan Yan Tang. Quaternionic weber local descriptor of color images. *IEEE Transactions on Circuits and Systems for Video Technology*, 27(2):261–274, feb 2017.

[68] Michael S. Lew, editor. *Principles of Visual Information Retrieval*. Springer London, 2001.

[69] Baoli Li and Liping Han. Distance weighted cosine similarity measure for text classification. In *Intelligent Data Engineering and Automated Learning–IDEAL 2013: 14th International Conference, IDEAL 2013, Hefei, China, October 20-23, 2013. Proceedings 14*, pages 611–618. Springer, 2013.

[70] Ying Liu, Dengsheng Zhang, Guojun Lu, and Wei-Ying Ma. A survey of content-based image retrieval with high-level semantics. *Pattern recognition*, 40(1):262–282, 2007.

[71] Ying Liu, Dengsheng Zhang, Guojun Lu, and Wei-Ying Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1):262–282, jan 2007.

[72] Pechin Lo, Bram van Ginneken, Joseph M. Reinhardt, Tarunashree Yavarna, Pim A. de Jong, Benjamin Irving, Catalin Fetita, Margarete Ortner, Rômulo Pinho, Jan Sijbers, Marco Feuerstein, Anna Fabijanska, Christian Bauer, Reinhard Beichel, Carlos S. Mendoza, Rafael Wiemker, Jaesung Lee, Anthony P. Reeves, Silvia Born, Oliver Weinheimer, Eva M. van Rikxoort, Juerg Tschirren, Ken Mori, Benjamin Odry, David P. Naidich, Ieneke Hartmann, Eric A. Hoffman, Mathias Prokop, Jesper H. Pedersen, and Marleen de Bruijne. Extraction of airways from ct exact-09). *IEEE Transactions on Medical Imaging*, 31(11):2093–2107, nov 2012.

[73] David G Lowe. Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision*, volume 2, pages 1150–1157. Ieee, 1999.

[74] Murari Mandal, Mallika Chaudhary, Santosh Kumar Vipparthi, Subrahmanyam Murala, Anil Balaji Gonde, and Shyam Krishna Nagar. Antic: Antithetic isomeric cluster patterns for medical image retrieval and change detection. *IET Computer Vision*, 13(1):31–43, 2019.

[75] Bangalore S Manjunath and Wei-Ying Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on pattern analysis and machine intelligence*, 18(8):837–842, 1996.

[76] Zahra Mansoori and Mansour Jamzad. Content based image retrieval using the knowledge of texture, color and binary tree structure. In *2009 Canadian Conference on Electrical and Computer Engineering*, pages 999–1003. IEEE, 2009.

[77] Daniel S. Marcus, Anthony F. Fotenos, John G. Csernansky, John C. Morris, and Randy L. Buckner. Open access series of imaging studies: Longitudinal MRI data in nondemented and demented older adults. *Journal of Cognitive Neuroscience*, 22(12):2677–2684, dec 2010.

[78] Mansoureh Sadat Mirasadi and Amir Hossein Foruzan. Content-based medical image retrieval of ct images of liver lesions using manifold learning. *International Journal of Multimedia Information Retrieval*, 8(4):233–240, 2019.

[79] Maxim Mizotin, Jenny Benois-Pineau, Michèle Allard, and Gwenaelle Catheline. Feature-based brain mri retrieval for alzheimer disease diagnosis. In *2012*

19th IEEE International Conference on Image Processing, pages 1241–1244. IEEE, 2012.

[80] Henning Müller, Nicolas Michoux, David Bandon, and Antoine Geissbuhler. 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, feb 2004.

[81] Nilima B. Mohite and Anil B. Gonde. Deep features based medical image retrieval. *Multimedia Tools and Applications*, 81(8):11379–11392, feb 2022.

[82] Subrahmanyam Murala, RP Maheshwari, and R Balasubramanian. Local tetra patterns: a new feature descriptor for content-based image retrieval. *IEEE transactions on image processing*, 21(5):2874–2886, 2012.

[83] Subrahmanyam Murala and Q. M. Jonathan Wu. Local mesh patterns versus local binary patterns: Biomedical image indexing and retrieval. *IEEE Journal of Biomedical and Health Informatics*, 18(3):929–938, may 2014.

[84] Subrahmanyam Murala and QM Jonathan Wu. Local ternary co-occurrence patterns: a new feature descriptor for mri and ct image retrieval. *Neurocomputing*, 119:399–412, 2013.

[85] Subrahmanyam Murala and QM Jonathan Wu. Mri and ct image indexing and retrieval using local mesh peak valley edge patterns. *Signal processing: image communication*, 29(3):400–409, 2014.

[86] Subrahmanyam Murala and QM Jonathan Wu. Spherical symmetric 3d local ternary patterns for natural, texture and biomedical image indexing and retrieval. *Neurocomputing*, 149:1502–1514, 2015.

[87] Subrahmanyam Murala and Q.M. Jonathan Wu. Spherical symmetric 3d local ternary patterns for natural, texture and biomedical image indexing and retrieval. *Neurocomputing*, 149:1502–1514, feb 2015.

[88] NEMA-CT image database. [online], 2012.

[89] NEMA-MR Database. [online], 2012.

[90] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7):971–987, 2002.

[91] Paulo H Oliveira, Lucas C Scabora, Mirela T Cazzolato, Marcos VN Bedo, Agma JM Traina, and Caetano Traina-Jr. Mammoset: An enhanced dataset of mammograms. In *Proc. Satellite Events 32nd Brazilian Symp. Databases*, pages 256–266, 2017.

[92] David L Olson and Dursun Delen. *Advanced data mining techniques*. Springer Science & Business Media, 2008.

[93] Muhammad Owais, Muhammad Arsalan, Jiho Choi, and Kang Ryoung Park. Effective diagnosis and treatment through content-based medical image retrieval (CBMIR) by using artificial intelligence. *Journal of Clinical Medicine*, 8(4):462, apr 2019.

[94] Shuchao Pang, Mehmet A Orgun, and Zhezhou Yu. A novel biomedical image indexing and retrieval system via deep preference learning. *Computer methods and programs in biomedicine*, 158:53–69, 2018.

[95] Rohini Pinapatruni and C Shoba Bindu. Learning image representation from image reconstruction for a content-based medical image retrieval. *Signal, Image and Video Processing*, pages 1–8, 2020.

[96] Konstantin Pogorelov, Kristin Ranheim Randel, Carsten Griwodz, Sigrun Losada Eskeland, Thomas de Lange, Dag Johansen, Concetto Spampinato, Duc-Tien Dang-Nguyen, Mathias Lux, Peter Thelin Schmidt, et al. Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In *Proceedings of the 8th ACM on Multimedia Systems Conference*, pages 164–169, 2017.

[97] Marcelo Ponciano-Silva, Juliana P Souza, Pedro H Bugatti, Marcos VN Bedo, Daniel S Kaster, Rosana TV Braga, Angela D Bellucci, Paulo M Azevedo-Marques, Caetano Traina, and Agma JM Traina. Does a cbir system really impact decisions of physicians in a clinical environment? In *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pages 41–46. IEEE, 2013.

[98] Adnan Qayyum, Syed Muhammad Anwar, Muhammad Awais, and Muhammad Majid. Medical image retrieval using deep convolutional neural network. *Neurocomputing*, 266:8–20, 2017.

[99] Yongjun Qi, Junhua Gu, Yajuan Zhang, Gengshen Wu, and Feng Wang. Supervised deep semantics-preserving hashing for real-time pulmonary nodule image retrieval. *Journal of Real-Time Image Processing*, 17(6):1857–1868, 2020.

[100] Gwénolé Quellec, Mathieu Lamard, Guy Cazuguel, Béatrice Cochener, and Christian Roux. Fast wavelet-based image characterization for highly adaptive

image retrieval. *IEEE Transactions on Image Processing*, 21(4):1613–1623, 2011.

[101] Md Mahmudur Rahman, Sameer K Antani, and George R Thoma. 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, 2011.

[102] Sharmili Roy, Yanling Chi, Jimin Liu, Sudhakar K Venkatesh, and Michael S Brown. Three-dimensional spatiotemporal features for fast content-based retrieval of focal liver lesions. *IEEE Transactions on Biomedical Engineering*, 61(11):2768–2778, 2014.

[103] Yong Rui, Thomas S Huang, and Shih-Fu Chang. Image retrieval: Current techniques, promising directions, and open issues. *Journal of visual communication and image representation*, 10(1):39–62, 1999.

[104] Mahya Sadeghi, Parmit Chilana, Jordan Yap, Philipp Tschandl, and M. Stella Atkins. Using content-based image retrieval of dermoscopic images for interpretation and education: A pilot study. *Skin Research and Technology*, 26(4):503–512, dec 2019.

[105] Manish Sapkota, Xiaoshuang Shi, Fuyong Xing, and Lin Yang. Deep convolutional hashing for low-dimensional binary embedding of histopathological images. *IEEE journal of biomedical and health informatics*, 23(2):805–816, 2018.

[106] Pierre Sermanet, David Eigen, Xiang Zhang, Michaël Mathieu, Rob Fergus, and Yann LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*, 2013.

[107] P Shamna, VK Govindan, and KA Abdul Nazeer. Content-based medical image retrieval by spatial matching of visual words. *Journal of King Saud University-Computer and Information Sciences*, 2018.

[108] P Shamna, VK Govindan, and KA Abdul Nazeer. Content based medical image retrieval using topic and location model. *Journal of biomedical informatics*, 91:103112, 2019.

[109] Amita Shinde, Amol Rahulkar, and Chetankumar Patil. Content based medical image retrieval based on new efficient local neighborhood wavelet feature descriptor. *Biomedical Engineering Letters*, 9(3):387–394, may 2019.

[110] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[111] Josef Sivic, Bryan C Russell, Alexei A Efros, Andrew Zisserman, and William T Freeman. Discovering objects and their location in images. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 1, pages 370–377. IEEE, 2005.

[112] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

[113] Fabio A Spanhol, Luiz S Oliveira, Caroline Petitjean, and Laurent Heutte. A dataset for breast cancer histopathological image classification. *Ieee transactions on biomedical engineering*, 63(7):1455–1462, 2015.

[114] G Sucharitha and Ranjan K Senapati. Shape based image retrieval using lower order zernike moments. *International Journal of Electrical and Computer Engineering*, 7(3):1651, 2017.

[115] G Sucharitha and Ranjan Kumar Senapati. Biomedical image retrieval by using local directional edge binary patterns and zernike moments. *Multimedia Tools and Applications*, 79(3):1847–1864, 2020.

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

[117] Komal Nain Sukhia, Muhammad Mohsin Riaz, Abdul Ghafoor, Syed Sohaib Ali, and Naima Iltaf. Content-based histopathological image retrieval using multi-scale and multichannel decoder based ltp. *Biomedical Signal Processing and Control*, 54:101582, 2019.

[118] Senthil Kumar Sundararajan, B Sankaragomathi, and D Saravana Priya. Deep belief cnn feature representation based content based image retrieval for medical images. *Journal of medical systems*, 43(6):1–9, 2019.

[119] Xiaoyang Tan and Bill Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE transactions on image processing*, 19(6):1635–1650, 2010.

[120] TCIA-CT Database. [online], 2010.

[121] Michael Reed Teague. Image analysis via the general theory of moments. *Josa*, 70(8):920–930, 1980.

[122] P. Tschandl, G. Argenziano, M. Razmara, and J. Yap. Diagnostic accuracy of content-based dermatoscopic image retrieval with deep classification features. *British Journal of Dermatology*, 181(1):155–165, oct 2018.

[123] Bram van Ginneken, Arnaud A. A. Setio, Colin Jacobs, and Francesco Ciompi. Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. apr 2015.

[124] Manisha Verma and Balasubramanian Raman. Center symmetric local binary co-occurrence pattern for texture, face and bio-medical image retrieval. *Journal of Visual Communication and Image Representation*, 32:224–236, 2015.

[125] VIA/I-ELCAP Database. [online], 2003.

[126] Linh Viet Tran. *Efficient image retrieval with statistical color descriptors*. PhD thesis, Linköping University Electronic Press, 2003.

[127] Michael Wainberg, Daniele Merico, Andrew Delong, and Brendan J Frey. Deep learning in biomedicine. *Nature Biotechnology*, 36(9):829–838, oct 2018.

[128] Ji Wan, Dayong Wang, Steven Chu Hong Hoi, Pengcheng Wu, Jianke Zhu, Yongdong Zhang, and Jintao Li. Deep learning for content-based image retrieval: A comprehensive study. In *Proceedings of the 22nd ACM international conference on Multimedia*, pages 157–166, 2014.

[129] James Ze Wang, Gio Wiederhold, Oscar Firschein, and Sha Xin Wei. Content-based image indexing and searching using daubechies' wavelets. *International Journal on Digital Libraries*, 1(4):311–328, 1998.

[130] Jian Wang, Jing Li, Xian-Hua Han, Lanfen Lin, Hongjie Hu, Yingying Xu, Qingqing Chen, Yutaro Iwamoto, and Yen-Wei Chen. Tensor-based sparse representations of multi-phase medical images for classification of focal liver lesions. *Pattern Recognition Letters*, 130:207–215, 2020.

[131] Zhiqiong Wang, Junchang Xin, Yukun Huang, Ling Xu, Jie Ren, Hao Zhang, Wei Qian, Xia Zhang, and Jiren Liu. A similarity measure method fusing deep feature for mammogram retrieval. *Journal of X-ray science and technology*, 28(1):17–33, 2020.

[132] Xu Wangming, Wu Jin, Liu Xinhai, Zhu Lei, and Shi Gang. Application of image sift features to the context of cbir. In *2008 International Conference on Computer Science and Software Engineering*, volume 4, pages 552–555. IEEE, 2008.

[133] Petra Welter, Thomas M Deserno, Benedikt Fischer, Rolf W Günther, and Cord Spreckelsen. Towards case-based medical learning in radiological decision making using content-based image retrieval. *BMC medical informatics and decision making*, 11(1):1–16, 2011.

[134] Guoxing Wu, Wenjie Lu, Guangwei Gao, Chunxia Zhao, and Jiayin Liu. Regional deep learning model for visual tracking. *Neurocomputing*, 175:310–323, 2016.

[135] Lei Wu and Steven CH Hoi. Enhancing bag-of-words models with semantics-preserving metric learning. volume 18, pages 24–37. IEEE, 2011.

[136] Yingying Xu, Lanfen Lin, Hongjie Hu, Dan Wang, Wenchao Zhu, Jian Wang, Xian-Hua Han, and Yen-Wei Chen. Texture-specific bag of visual words model and spatial cone matching-based method for the retrieval of focal liver lesions using multiphase contrast-enhanced ct images. *International journal of computer assisted radiology and surgery*, 13(1):151–164, 2018.

[137] Jun Yang, Yu-Gang Jiang, Alexander G Hauptmann, and Chong-Wah Ngo. Evaluating bag-of-visual-words representations in scene classification. In *Pro-*

ceedings of the international workshop on Workshop on multimedia information retrieval, pages 197–206, 2007.

[138] Praveen Kumar Reddy Yelampalli, Jagadish Nayak, and Vilas Haridas Gaidhane. A novel binary feature descriptor to discriminate normal and abnormal chest ct images using dissimilarity measures. *Pattern Analysis and Applications*, 22(4):1517–1526, 2019.

[139] Ela Yildizer, Ali Metin Balci, Tamer N Jarada, and Reda Alhajj. Integrating wavelets with clustering and indexing for effective content-based image retrieval. *Knowledge-Based Systems*, 31:55–66, 2012.

[140] Matthew D Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *European conference on computer vision*, pages 818–833. Springer, 2014.

[141] Baochang Zhang, Yongsheng Gao, Sanqiang Zhao, and Jianzhuang Liu. Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor. *IEEE Transactions on Image Processing*, 19(2):533–544, feb 2010.

[142] Shusen Zhou, Qingcai Chen, and Xiaolong Wang. Active deep learning method for semi-supervised sentiment classification. *Neurocomputing*, 120:536–546, 2013.

[143] Şaban Öztürk. Class-driven content-based medical image retrieval using hash codes of deep features. *Biomedical Signal Processing and Control*, 68:102601, jul 2021.

ملخص الرسالة

يعتبر استرجاع وإدارة الصور الطبية اعتمادا علي محتواها (CBMIR) أكثر أهمية الآن ، خاصة مع زيادة التصوير الطبي وتوسيع قاعدة بيانات الصور الطبية. أيضًا ، تسمح هذه الأنظمة بالاستفادة من الصور الطبية في فهم أفضل ورؤى أعمق لأسباب وعلاج الأمراض المختلفة ، وليس فقط لأغراض التشخيص. لذلك ، لعب CBMIR دورًا هامًا في مجال معالجة الصور واستخراج الميزات منخفضة المستوى مثل الرسوم البيانية الملونة والحواف والملمس والشكل, وكذلك مقاييس التشابه لمقارنة واسترجاع الصور الطبية. تعزز غالبية طرق CBMIR المستخدمة بالفعل في استرجاع الصورة الطبية وتشخيص الأمراض عن طريق تقليل مشكلة الفجوة الدلالية بين المستويات البصرية المنخفضة والمستويات الدلالية العالية. لتحقيق كل هذه الأهداف ، هناك حاجة ماسة إلى طريقة ذات كفاءة ودقة في استرجاع الصور الطبية اعتمادا علي محتواها.

تقترح هذه الأطروحة (RbQE) طريقة فعالة لاسترجاع الصور الناتجة عن الأشعة المقطعية (CT) والأشعة بالرنين المغناطيسي (MRI) وصور الأنسجة. تعتمد RbQE على توسيع ميزات الاستعلام واستغلال نماذج التعلم المدربة مسبقًا AlexNet و VGGNets و ResNets لاستخراج الميزات المدمجة والعميقة والعالية المستوى من الصور الطبية. هناك إجراءان في RbQE للبحث: بحث سريع وبحث نهائي. في البحث السريع, يتم توسيع الاستعلام الأصلي عن طريق استرداد الصور ذات الترتيب الأعلى من كل فئة ويتم استخدامه لإعادة صياغة الاستعلام عن طريق حساب القيم المتوسطة للميزات العميقة للصور ذات الترتيب الأعلى, مما يؤدي إلى استعلام جديد لكل فئة. في البحث النهائي ، سيتم استخدام الاستعلام الجديد الأكثر تشابهًا مع الاستعلام الأصلي للاسترجاع من قاعدة البيانات.

تم اختبار أداء الطريقة المقترحة علي خمس مجموعات بيانات معتمدة متاحة للجمهور ، وهي TCIA-CT و EXACT09-CT و NEMA-CT و OASIS-MRI و KIMIA Path960 ومقارنتها بأحدث الطرق الموجودة. تظهر النتائج التجريبية أن الطريقة المقترحة تتجاوز الطرق المقارنة بنسبة 0.84٪ و 4.86٪ و 1.24٪ و 14.34٪ و 1.96٪ في متوسط دقة الاسترجاع (ARP) في استرجاع العشرة الأوائل ، لـ TCIA-CT و EXACT09-CT و NEMA-CT و OASIS-MRI و KIMIA Path960 ، على التوالي. ولكن بالنسبة لـ KIMIA Path960 ، فإن ARPs لأعلى خمسة وخمسة عشر وعشرين تتجاوز الطرق المقارنة بنسبة 2.14٪ و 6.74٪ و 9.42٪ على التوالي.

جامعـة بنهـــا

كلية الحاسبات والذكاء الاصطناعي

نموذج فعال لإسترجاع المعلومات لقواعد البيانات الغير متجانسة

رسالة

مقدمة للحصول على درجة الماجستير في نظم المعلومات

قسم نظم المعلومات

كلية الحاسبات والذكاء الاصطناعى

جامعة بنهـــا

مقدمة من الباحث
إبراهيم محمد إبراهيم عفيفي

معيد بقسم نظم المعلومات
كلية الحاسبات والذكاء الاصطناعى - جامعة بنهـــا

تحت إشراف

أ.د. محمد عبدالفتاح عوض

أستاذ نظم المعلومات ووكيل الكلية لشؤون التعليم والطلاب
- كلية الحاسبات والذكاء الاصطناعي- جامعة بنها

د. متولي رشاد متولي

مدرس علوم الحاسب
- كلية الحاسبات والذكاء الاصطناعي- جامعة بنها

بنهـــا - 2023