



Hash Code Generation using Deep Feature Selection Guided Siamese Network for Content-Based Medical Image Retrieval

Saban OZTURK* *Amasya University, Electrical and Electronics Department, 05100, Amasya, Turkey*

Highlights

- An effective and short hash code technique is provided for content-based medical image retrieval.
- Manhattan distance based Siamese network activation functions are changed to hyperbolic tanh.
- Pre-hash code approach is introduced.
- Short-length hash code generation with the iterative downsampling approach is introduced.
- Higher performance is obtained compared to current state-of-the-art methods in the literature.

Article Info

*Received: 28 Mar 2020**Accepted: 04 Dec 2020*

Keywords

*Cbmir
Cnn
Retrieval
Hashing*

Abstract

It is very pleasing for human health that medical knowledge has increased and the technological infrastructure improves medical systems. The widespread use of medical imaging devices has been instrumental in saving lives by allowing early diagnosis of many diseases. These medical images are stored in large databases for many purposes. These datasets are used when a suspicious diagnostic case is encountered or to gain experience for inexperienced radiologists. To fulfill these tasks, images similar to one query image are searched from within the large dataset. Accuracy and speed are vital for this process, which is called content-based image retrieval (CBIR). In the literature, the best way to perform a CBIR system is by using hash codes. This study provides an effective hash code generation method based on feature selection-based downsampling of deep features extracted from medical images. Firstly, pre-hash codes of 256-bit length for each image are generated using a pairwise siamese network architecture that works based on the similarity of two images. Having a pre-hash code between -1 and 1 makes it very easy to generate hash code in hashing algorithms. For this reason, all activation functions of the proposed convolutional neural network (CNN) architecture are selected as hyperbolic tanh. Finally, neighborhood component analysis (NCA) feature selection methods are used to convert pre-hash code to binary hash code. This also downsamples the hash code length to 32-bit, 64-bit, or 96-bit levels. The performance of the proposed method is evaluated using NEMA MRI and NEMA CT datasets.

1. INTRODUCTION

Over the years, the increase in scientific knowledge and the rapid spread of technological developments in every field raise human life standards and expectations. Human health is at the top of these areas. Many parameters such as changes in various economic parameters, the life expectancy of individuals and increase in quality of life, and rising standards allow rapid developments in the medical field [1]. Medical devices and medical imaging devices play an essential role in early diagnosis and treatment [2]. Recently, the widespread use of medical imaging devices brings some challenges. These medical images need to be stored, and the storage problem is increasing day by day due to a large number of images. Problems with accessing these images are another challenge. Radiologists and students who want to gain experience are trained using these images. In doubtful situations, some radiologists may need to examine similar images. In these cases, CBIR systems, which are long-standing research topics for artificial intelligence society, become very important [3]. A code is determined for each image in CBIR systems, and every image is stored with this code in the dataset. When searching similar images to a query, search processing takes

place according to the distance between these codes. This process has two advantages; 1- storage capacity, 2-retrieving speed [4].

The nearest neighbor (NN) search approach may be preferred when searching images or points with similar features to a large dataset. However, using the exact NN is quite costly, and this process takes quite a long time. Approximate NN (ANN) is preferred in the literature for this process for relaxation of exact NN [5]. The most popular and attractive ANN approach is hashing. According to the hashing approach, a binary code consisting of few parameters is generated for each image. In this way, the storage cost is minimized. Hamming distance is used to access images using binary hash codes. In this way, the querying speed is maximized. There are two important components for generating a hash code; robust features and effective downsampling. Many computer-aided diagnosis (CAD) techniques are presented in the literature to perform these operations automatically. These studies can be divided into two parts in terms of feature extraction: hand-crafted features and deep features. In terms of downsampling, it can be divided into internal downsampling and external downsampling. Hand-crafted features represent manually specified features for each dataset, while deep features represent automatically extracted features by CNN architectures. While the internal downsampling refers to the hash code produced directly by the main algorithm, the external downsampling refers to the hash code produced by an extra algorithm.

One of the essential components of the performance of CBIR studies is feature extraction. The task of feature extraction is a field that has been studied for a long time and developed with the development of artificial intelligence (AI) methods. In the early days, when feature extraction algorithms were applied to image processing problems, quite basic features were used (e.g., edge detection, grey level information, color information, shape information, etc.) [6]. These low-level features had great difficulties in obtaining spatial information. For this reason, text-based image retrieval studies were more popular before the development of better hand-crafted feature extraction algorithms. The speed of widespread use of imaging devices and the excessive information contained in the images made CBIR systems attract more attention. In this period, texture features and wavelet-based techniques were used quite popularly for general image problems [7]. In CBIR studies of this period, Fourier [8], curvelet [9], Gaussian [10], combined texture features [11], radon features [12] were used to extract features from images. Since features are significant in CBIR methods, the most powerful algorithms of the period are generally used by CBIR researchers. However, the mentioned features cannot maintain their performance when the dataset changes or the dataset variation increases. For this reason, these methods have not been very helpful in narrowing the 'semantic gap' between the codes produced by algorithms and linguistic variables produced by humans [13]. For this purpose, hand-crafted feature extraction methods have continued to be recommended and strengthened. More advanced hand-crafted feature extraction methods are recommended before deep features examine the image from multiple angles. Sensitivity to dataset variations has decreased thanks to this multi-specific approach. Besides, a serious performance increase was experienced. The most popular hand-crafted feature extraction algorithms of the last period are gray level co-occurrence matrix (GLCM) [14], local binary patterns (LBP) [15], speeded up robust features (SURF) [16], scale-invariant feature transform (SIFT), and histogram of oriented gradients (HOG) [17]. The contribution of the mentioned hand-crafted algorithms to the CBIR domain is quite large, but it falls short of generating high-level features to close the 'semantic gap' in most cases.

CNN architectures automatically learn low, middle, and high-level features of the image and process them for various tasks. Studies in the literature prove that it has higher performance than hand-crafted methods [18]. Recent CBIR studies generally use CNN-based feature extraction. CNN architectures help reduce the 'semantic gap'. Retrieval studies using CNN-based methods are frequently carried out in medical fields such as content-based mammographic image retrieval [19], diabetic retinopathy [20], hepatobiliary images [21], and lung cancer [22]. Although different CNN architectures are presented for different tasks in the literature, more than one CNN architecture is offered for the same task. Some studies select a ready-made CNN architecture from literature according to the problem being studied. Some studies suggest a novel CNN architecture for CBIR. Linear CNN models [3], siamese CNN models [23], and residual CNN architecture [24] are among the most commonly used CBIR CNN models.

This study provides an effective hash code generation method that can work with unlabeled or labeled data for CBMIR tasks. The proposed method forces all parameters within the CNN architecture to be -1 or 1 to support hash code generation. For this purpose, all activation functions are used as a hyperbolic tanh after convolution layers. Hash code generation is applied to pre-hash codes generated by the CNN-F. These codes are called pre-hash codes because they are converted into hash codes by downsampling only and can be used as hash codes in their original form. Pre-hash codes in the output layer contain 256 parameters. With feature selection algorithms, this number of parameters can be downsampled to 32, 64, or 96-bit. Contributions of the proposed method can be listed as follows;

- Thanks to the proposed framework flexible siamese network architecture, hash codes can be generated using unlabeled and labeled examples,
- The hyperbolic tanh relaxation used as the activation layer provides a binary hash code with less effort.
- Feature selection based downsampling process allows the hash code length to change without interfering with the CNN architecture.
- The performance of the proposed method is compared with current state-of-the-art methods using medical datasets. It is achieved better than other algorithms in 32-bit, 64-bit, and 96-bit precision experiments.

The rest of this paper is organized as follows; Section 2 provides a detailed analysis of the proposed framework. It is analyzed in parts in this section. Section 3 includes dataset details, experimental settings and results, and a discussion of the results. Finally, the conclusion is presented in Section 4.

2. Material and Methods

2.1. Overview of the Proposed Hashing Method

Medical images contain information about organs, diseases, and disease stages. If the classification process is preferred instead of using CBMIR, it raises the possibility of encountering some problems. First, medical images often examine a disease or its stage. Since the area covered by this disease in the image can be quite small, healthy organs and diseased organs are classified in the same class due to the classification process. On the other hand, information about the class of the disease is also assigned to the same class in the classification process. But these small details are very important for radiologists. In the CBMIR, the inspection process does not take place over the dominant object in the image. It can distinguish between an organ with diseased tissue and an organ with healthy tissue. Also, if the same disease is in other organs, it can also retrieve these organs. In these respects, CBMIR systems differ from classification systems.

Nowadays, considering the increasing number of medical images dataset, three tasks are expected from a CBMIR system; low storage cost, fast retrieving speed, and high precision. Increasing the hash code length is one of the simplest solutions to achieve high performance. However, in this case, the retrieving time extends, and the storage capacity increases. The simplest solution for a faster retrieval is to keep the hash code length very short. In this way, the need for storage is reduced. However, retrieval performance decreases. The balance between these two parameters is very important.

The proposed method can work with both labeled and unlabeled samples. It only requires similarity knowledge. Nowadays, it is very difficult to reach labeled medical datasets, so it is imperative to provide algorithmic solutions [25]. It can be seen as the first contribution of this proposed framework thanks to the Siamese network architecture. The second important contribution of the proposed method is that the trainable CNN parameters are forced to be -1 or 1. In this way, it is very easy to convert the output vector to binary. Since this process is quite difficult for the CNN architecture, the relaxation technique is used. Hyperbolic tanh is added after each learnable layer (convolution layers and fully connected layers). The CNN architecture used in Siamese architecture is CNN-F [26]. It is chosen for its simple structure, speed, and high performance. Hyperbolic tanh activation is not the only change made in the CNN-F architecture, it is also used as a new activation function in FCL layers. Two FCLs are added, containing 1024 and 256 parameters, respectively. The last layer of the proposed CNN-F architecture with 256 parameters is called

the pre-hash layer. This layer can be used as a hash code by applying a soft threshold. But, 256-bit hash code requires quite a long time for real-time applications. For this reason, the parameters in this layer are downsampled with the proposed NCA-based feature selection method. The most important advantage of the downsampling part is that it is very easy to obtain hash codes of different lengths. In addition, there is no need to retrain the CNN architecture to change the hash code length.

The proposed framework is shown in Figure 1. CNN-F architecture and details, siamese network details, and feature selection used in the proposed framework are explained in detail in the following sections.

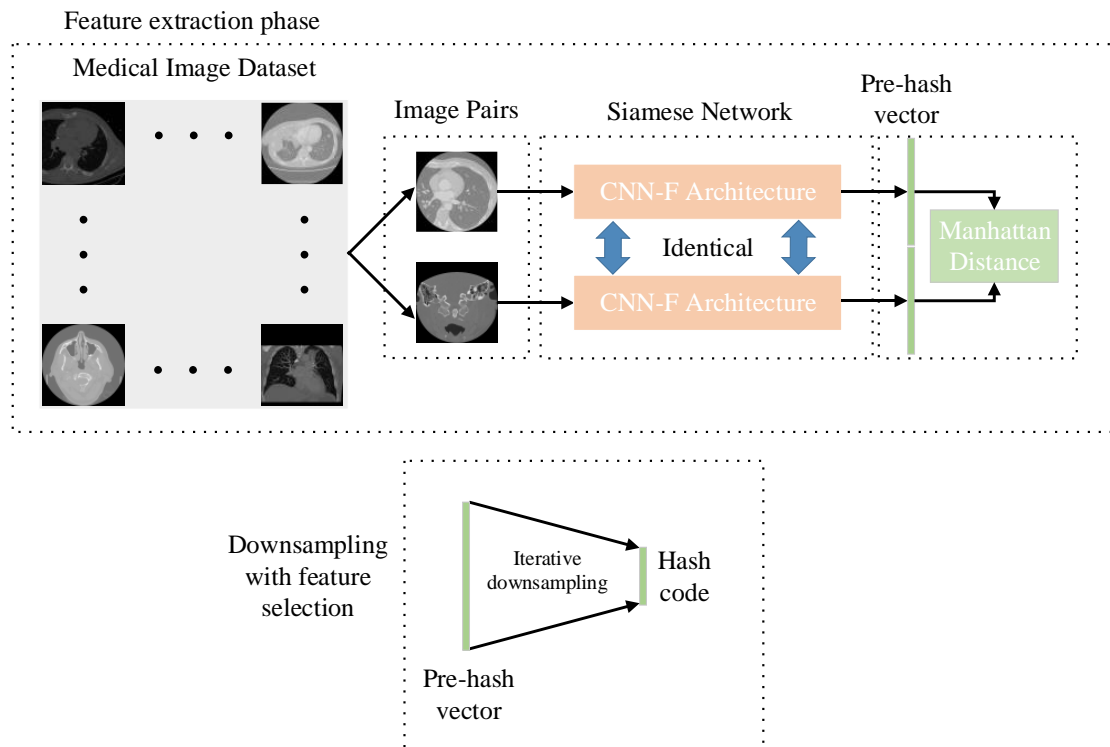


Figure 1. Proposed hashing framework

2.2. Siamese Network Details and Modified CNN-F Architecture

The proposed method is trained in a pairwise fashion using the Siamese network. In other words, two images are simultaneously applied to the siamese network input. The only information about these images is that they are similar or different. This similarity may be related to organs or related to the disease. The Siamese network includes two CNN architectures for pairwise manner training. These two CNN architectures are identical. The Siamese network structure compares feature vectors obtained by these two CNN architectures from different images. Manhattan distance is used to calculate distance, as in Equation (1).

$$d(y, t) = \|y - t\|_1 = \sum_{i=1}^n |y_i - t_i| \quad (1)$$

where y represents the output vector of the first CNN-F (pre-hash code), t represents the output vector of the second CNN-F (pre-hash code), n is the length of the vectors. CNN-F trainable parameters in the siamese structure are updated according to the distance. Some modifications are made to the CNN-F architectures inside the Siamese structure. Since these architectures are used to generate hash code, their parameters and outputs must conform to hash codes. Therefore, its parameters are expected to be -1 or 1. It is quite costly to train CNN architecture with exact -1 and exact 1. For this reason, relaxation is applied. By adding hyperbolic tanh to each layer, the values are between [-1, 1], as shown in Figure (2).

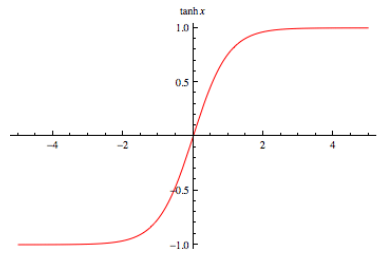


Figure 2. Hyperbolic tanh function

Numerical values in the pre-hash vector can be easily converted to -1 or 1 with a simple thresholding application with the help of the relaxation process. Because, as seen in Figure 2, most of the parameters of the network are very close to -1 and 1. The second change made in CNN-F architecture is the replacement of the fully connected layer (FCL). The new FCL contains two layers of 1024 and 256 parameters, respectively. The pre-hash vector is the most important layer in this section. This vector convert to shorter or longer codes easily. Since 32-bit or 64-bit hash codes are often used in the literature, this vector is downsampled to the specified numbers. 256 is chosen for the least loss in downsampling operation. The output of a portion of the CNN-F architecture consisting of only one convolution, one activation, and one pooling layer is calculated using Equation 2.

$$f(o_i) = \maxpool_{n \times n} (\sigma(w \otimes I + b)). \quad (2)$$

where o_i indicates the current layers' output, \maxpool is the maximum pooling, n represents kernel size of pooling operation, σ indicates hyperbolic tanh activation, w represents convolution kernels, \otimes indicates 2D convolution operator, I is the input matrix of the current layer, and b indicates bias. As shown in Figure 3, a hyperbolic tanh is used after each convolution layer and after each layer in FCL. Figure 3(a) shows the hyperbolic tanh function used after the convolution layers, Figure 3(b) shows the hyperbolic tanh function used after the fully connected layers.

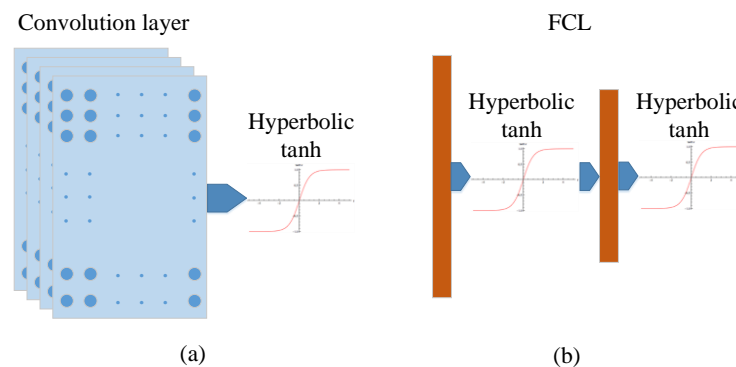


Figure 3. Adding hyperbolic tanh to CNN-F for relaxation, a) for convolution layer, b) for FCL

2.3. Deep Feature Selection Based Downsampling

The 256-bit length pre-hash code generated by the Siamese network is not effective for CBMIR. Some bits of 256-bit long pre-hash code are so insignificant that they are the same for all images. Some bits are specific to only one or two images, and they are meaningless for the whole of the dataset. It is necessary to remove these components that cause storage load and time loss from the hash code. However, hash code values should not be reconstructed for this process. This section aims to identify meaningless bits in the pre-hash code and reduce the pre-hash code length by eliminating unimportant bits. Besides, the meaning of the hash code should remain the same. In the literature, methods such as principal component analysis (PCA), independent component analysis (ICA), reconstruction independent component analysis (RICA), and reliefF are actively used in the feature reduction task [27]. However, these methods change the domain of the features while performing the downsampling task. For the classification task, this does not cause any problems, in some cases it even contributes positively to the classification performance. However, for hash

codes in CBIR systems, this change can cause big problems. For this reason, an additional transformation is required after this process. Thus, while the computational load increases, the loss from the original space increases at each step.

This study focuses on a method that can only select the most meaningful ones without changing the hash code bits. Neighborhood component analysis (NCA) can successfully perform feature selection for both the classification task and the regression task. NCA's most advantageous property for hash code oversampling is its learning-based algorithm and freedom to determine the distance. There is no need to change the space, as it tries to reduce the distance using linear transformations. The second important contribution of this study is updating the NCA to enable feature selection in CBIR tasks. Consider a pre-hash code with n -bits, $S=\{x_i\}$, $i=1,2,\dots,n$. Significance values are calculated by randomly selecting reference points from S . Every bit in S has the possibility to be selected. This probability can be calculated using $P(\text{Ref}(x)=x_j/S)$. This random selection process is carried out according to the distance function (d_w) in Equation (3).

$$d_w(x_i, x_j) = \sum_{r=1}^p w_r^2 |x_{ir} - x_{jr}| \quad (3)$$

where d_w represents distance function, w_r represents feature weights, x_j indicates picked point from S . In this case, the probability of reference point x is calculated as in Equation (4).

$$P(\text{Ref}(x) = x_j | S) = \frac{k(d_w(x, x_j))}{\sum_{j=1}^n k(d_w(x, x_j))} \quad (4)$$

where $k(z)=\exp(-z/\sigma)$. Now consider the leave-one-out application of this randomized retrieval model, the probability of point x_j for reference point x_i is calculated using Equation (5).

$$p_{ij} = P(\text{Ref}(x_i) = x_j | S^{-i}) = \frac{k(d_w(x_i, x_j))}{\sum_{j=1, j \neq i}^n k(d_w(x_i, x_j))} \quad (5)$$

The loss is determined by comparing the obtained downsampled hash code with the pre-hash code representation ratio. The NCA ranks all bits in order of importance in the hashing task. The location of the most important 32, 64, or 96 bits, is learned. With these bits, new 32-bit, 64-bit, and 96-bit hash codes are generated.

2.4. Binarization of Pre-Hash Codes

The proposed framework is designed to pull the hash code values to exact -1, and 1. However, since this process is quite difficult, relaxation is applied. For this reason, values that are very close to 1 and very close to -1 are found in the hash code. A simple binarization is sufficient to convert these values to 1 and -1. Equation 6 is used to convert hash bits to 1 and -1

$$x = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{otherwise} \end{cases} \quad (6)$$

3. Experiments and Experimental Results

In this section, the performance of the proposed hashing method is evaluated. For this purpose, three medical datasets used in this study are introduced. Appropriate evaluation parameters should be selected so that the performance of the proposed method can be evaluated fairly and compared with the state-of-the-

art method in the literature. The selection and calculation of evaluation parameters are presented in this section. Then, the effect of contributions in the proposed framework on performance is examined in detail in the ablation study sub-section. Finally, experimental results and comparisons with methods in the literature are discussed.

The proposed method is trained on a computer with Intel Core i7-7700K CPU (4.2 GHz), 32 GB DDR4 RAM, and NVIDIA GeForce GTX 1080 graphic card.

3.1. Dataset

The retrieval performance of the proposed method is performed using three known medical datasets. NEMA CT¹, NEMA MRI¹, and IRMA [28] datasets are used for training, evaluation, and testing. NEMA CT and NEMA MRI datasets are formed DICOM format images provided by National Electrical Manufacturers Association. 663 images from the NEMA CT dataset are selected for this study. The number of organ classes represented by these images is ten; 'CT0001', 'CT0002', 'CT0015', 'CT0021', 'CT0054', 'CT0057', 'CT0060', 'CT0070', 'CT0080', and 'CT0083'. All images in this data are resized to 128x128. Table 1 contains details on NEMA CT image selection. Figure 4 shows some sample images of the NEMA CT dataset.

Table 1. Detail of NEMA CT Image Selection

Name of dataset	Total number of images in class	Total number of images used for NEMA CT experiments
CT0001	54	50
CT0002	461	80
CT0015	88	80
CT0021	66	66
CT0054	52	52
CT0057	104	70
CT0060	75	60
CT0070	112	80
CT0080	253	80
CT0083	69	45



Figure 4. Sample images from NEMA CT dataset

In the experiments related to NEMA MRI data, 775 MRI images are used. These images are selected from the NEMA MRI dataset randomly. The NEMA MRI codes of these images, which represent eleven organ classes in total, are as follows; 'ABDFATDY', 'BIGHEART', 'BRAINSAG', 'BRFLAIR', 'CARANSRC',

¹ <ftp://medical.nema.org/medical/Dicom/Multiframe/>

'DYNAMAMMO', 'ILIACSUB', 'KNEE3D', 'LIVERAN', 'VHMFLEGS', 'VOLBRAIN'. Table 2 contains details on NEMA MRI image selection. Figure 5 shows some sample images of the NEMA MRI dataset.

Table 2. Detail of NEMA MRI Image Selection

Name of dataset	Total number of images in class	Total number of images used for NEMA MRI experiments
ABDFATDY	112	80
BIGHEART	252	70
BRAINSAG	60	44
BRFLAIR	136	64
CARANSRC	91	65
DYNAMAMMO	168	80
ILIACSUB	304	85
KNEE3D	100	65
LIVERAN	72	54
VHMFLEGS	308	88
VOLBRAIN	200	80

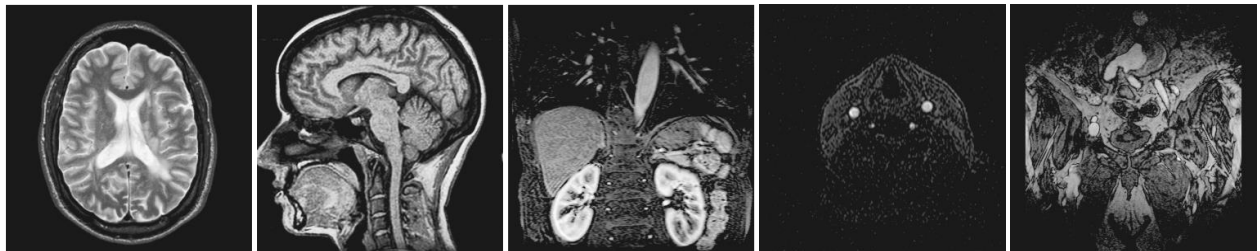


Figure 5. Sample images from NEMA MRI dataset

IRMA dataset was created during routine radiology at the Department of Diagnostic Radiology, Aachen University of Technology (RWTH), Aachen, Germany. X-Ray images in this dataset contain 57 different organ images. The total number of training images is 12677, and the total number of test images is 1733. All images in the dataset are rescaled to 128x128 pixel dimensions. Some example images from the IRMA dataset are shown in Figure 6.



Figure 6. Sample images from IRMA dataset

3.2. Retrieval Performance Evaluation Measures

The same parameters used in the literature are used to evaluate the performance of the proposed method fairly. In the literature, average precision, average recall, and F-score parameters are often used to evaluate

retrieval performance. Precision, average precision, recall, average recall, and F-score parameters are calculated as in Equations (7)-(11), respectively

$$P(I_q) = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}} \quad (7)$$

$$ARP = \frac{1}{n} \sum_{i=1}^n P(I_i) \quad (8)$$

$$R(I_q) = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in the database}} \quad (9)$$

$$ARR = \frac{1}{n} \sum_{i=1}^n R(I_i) \quad (10)$$

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

where P indicates precision, R indicates recall, ARP represents average precision, ARR represents average recall, I_q is query image, n represents the number of images in the dataset.

3.3. Ablation Study

The contributions offered by the proposed framework are examined as experimental in this section. For this purpose, the performance of the hyperbolic tanh activation function is compared with the ReLU activation function traditionally found in each CNN architecture. Figure 7 (a) shows the training curve of the ReLU activation function. Figure 7 (b) shows the training curve performed with hyperbolic tanh. The conditions are kept the same for the two trainings. Accordingly, the initial learning rate of CNN architectures is 0.0001, minibatch size is 96, gradient decay is 0.9, and the margin value is 0.3. Contrastive loss is used for both experiments. The output vector obtained is converted as -1 and 1, according to Equation 6. But in this part, Equation 6 has to be updated for the ReLU function. In ReLU experiments, the threshold value is chosen as the average of the current vector. Training curves show that the proposed architecture produces pre-hash codes that are more suitable for generating hash code. While the standard deviation in Figure 7(a) is quite high, Figure 7(b) has a more stable training curve. When the training process is completed, the loss of the network trained with ReLU is almost twice the loss of the network trained with hyperbolic tanh.

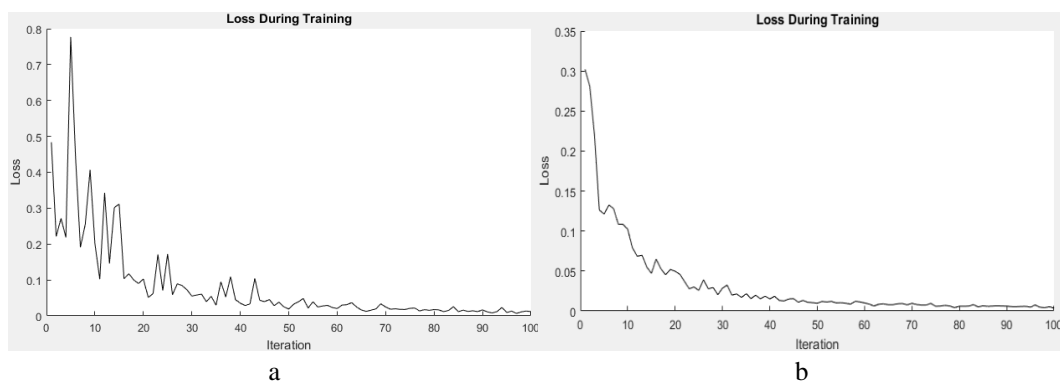


Figure 7. Comparison of ReLU and hyperbolic tanh, a) Training with ReLU, b) Training with hyperbolic tanh

Hyperbolic tanh makes a great contribution to hash code learning by bringing parameters closer to -1 and 1. On the other hand, it is a possibility that it produces an exact -1 or exact 1. If exact -1 and exact 1 values are used, then the training process will be quite difficult. The relaxation effect can be seen in Figure 8. Figure 8 (a) shows the training of the net with the relaxation approach, and Figure 8 (b) shows the training of the net without using relaxation (training with exact -1 and exact 1). The training process without relaxation is very unstable, and the network takes a long time to learn the parameters correctly.

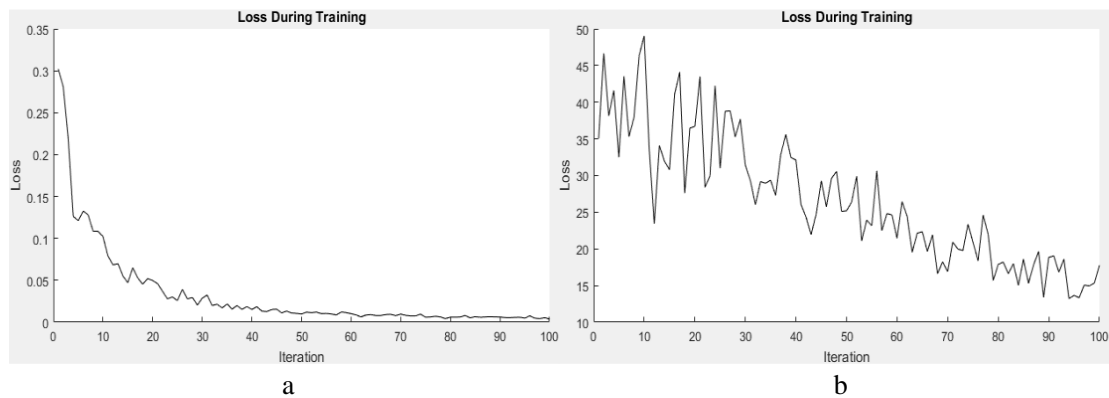


Figure 8. Relaxation effect, a) Training with relaxation, b) Training without relaxation

In retrieval studies, distance is used to determine similar images in the dataset with a query image. For this reason, determining the most appropriate distance approach is beneficial for performance. While Euclidean distance is generally preferred for problems such as classification and regression, hamming distance is generally used for retrieval. But recent research shows that Manhattan distance can well express the difference between two binary variables, especially when using relaxed space [29]. For this reason, CNN-F architecture is trained according to Manhattan distance.

Conversion of pre-hash codes to hash codes is carried out by feature selection based oversampling. For this purpose, a feature selection method based on NCA is offered in this study. Using the NCA, the locations of the most effective features are determined. Features at specified locations are used for the hash code. The NCA transaction is updated according to the length of the hash code. Instead of using a single NCA, sequential NCA is used in this section. Thus, the downsampling process is done iteratively. For this reason, it is called iterative downsampling. The performance difference between iterative downsampling and hard downsampling is shown in Table 3. Accordingly, iterative downsampling increases precision performance by more than 2%.

Table 3. Comparison of Iterative Downsampling and Hard Downsampling Performance using NEMA CT dataset

Method	Average Precision (%)		
	32-bit	64-bit	96-bit
Hard Downsampling	91.01	96.27	97.82
Iterative Downsampling	98.77	99.41	99.92

3.4. Experimental Results

The proposed method's performance is tested in this section using NEMA CT, NEMA MRI, and IRMA medical datasets. 32-bit, 64-bit, and 96-bit long hash codes are produced in order to effectively evaluate the retrieval performance of the proposed method. Thus, the performance of the proposed method for hash codes of different lengths can be evaluated. On the other hand, average precision is very important to evaluate the retrieval performance. In order to generalize the retrieval performance of the proposed method, 10, 20, 30, 40, and 50 answer images are retrieved in response to a query image. Figure 9 shows ARP curves for NEMA CT, NEMA MRI, and IRMA data of the proposed method. 32-bit, 64-bit, and 96-bit long hash codes are presented in the same frame for each dataset. As expected, the higher the hash code length, the higher the retrieval performance. However, this situation negatively affects response time. Also, Figure 9 reveals that precision performance decreases as the number of retrieved images increases.

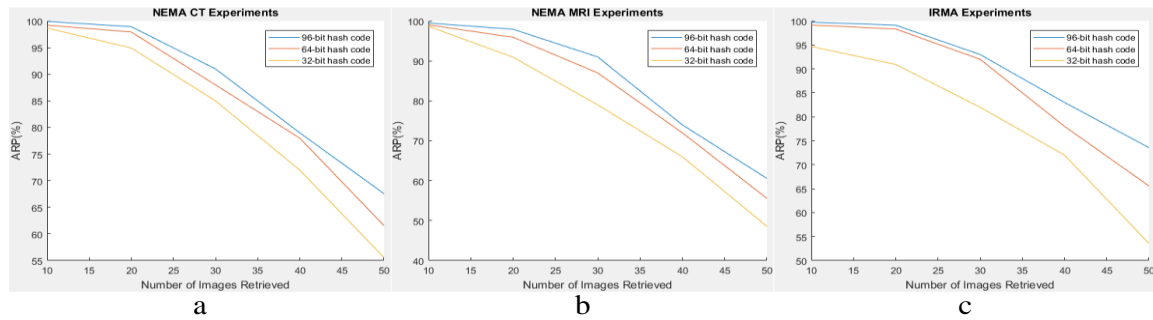


Figure 9. Comparison of performance in respect of a) ARP curves for NEMA CT, b) ARP curves for NEMA MRI, c) ARP curves for IRMA

Table 4 shows the performance of the proposed framework for experiments performed using NEMA CT and NEMA MRI data. Table 4 also shows the performance of other methods in the literature on NEMA CT and NEMA MRI datasets. Accordingly, the proposed method has a higher performance than other methods in both datasets. An almost 10% increase in performance is offered compared to the current methods in the literature. Besides, the fast modifiable hash code length of the proposed method provides a great advantage. In the experiments in Table 4, 64-bit is used as the hash code length. The number of images retrieved is chosen as 25. The structure of all techniques is preserved for comparison, and the same conditions are met.

Table 4. Comparison of the proposed framework with other methods

Method	NEMA CT			NEMA MRI		
	ARP (%)	ARR (%)	F-score (%)	ARP (%)	ARR (%)	F-score (%)
DLTerQEP	66.91	44.78	53.65	62.45	50.29	55.71
LBDP	71.04	49.41	58.28	67.77	53.05	59.51
LMePTerP	72.37	55.65	62.91	68.33	55.65	61.34
LBP	75.10	58.19	65.57	75.08	61.71	67.74
LWP	75.73	61.43	67.83	73.41	58.97	65.40
OFMMs	77.28	62.40	69.04	75.18	60.40	66.98
LBDISP	80.49	64.16	71.40	77.82	62.52	69.33
DBWP	84.11	68.75	75.66	83.73	66.73	74.27
Proposed framework	99.81	74.44	85.27	99.40	67.39	80.32

Table 5 shows the retrieval performance of the proposed method on the IRMA dataset. It presents the performances of other methods used IRMA dataset for the retrieval task. It appears that methods involving CNN produce higher performance. In addition, the bag of visual words (BoVW) method draws attention among the hand-crafted methods. As a result, the proposed method surpasses both hand-crafted feature extraction methods and other CNN-based methods.

Table 5. IRMA dataset retrieval performance comparison of the proposed framework with other methods (the methods with * are reported in [3])

Method	Feature Extraction Approach of Method	P@20
SC with Pyramid*	Hand-crafted	84.29
Camlica et al.*	Hand-crafted	86.8
GSC with Pyramid*	Hand-crafted	91.07
Shamna et al.*	BoVW	97.22
Öztürk's method [3]	CNN	99.12
Proposed method	CNN	99.15

4. CONCLUSION

In this study, an effective hash code generation method is presented for medical datasets containing a large number of samples. The proposed method is trained with a siamese network structure to work with labeled and unlabeled medical datasets. Siamese network structure includes CNN-F architecture with hyperbolic tanh activation. The added hyperbolic tanh activation function forces the trainable CNN parameters to be close to -1 and 1. In this way, the values produced by the CNN-F architecture can be easily converted to hash code. CNN-F architecture produces a 256-bit vector called the pre-hash code. Pre-hash code can be used alone as a hash code, but it is not effective. The most important advantage of the pre-hash code approach proposed in this study is that it can change the hash code length quite simply. There is no need to retrain the CNN architecture to generate a hash code of different lengths. The feature selection based downsampling method has been proposed to make this code effective. The proposed feature selection based downsampling method makes use of NCA. The NCA selects the most important 32, 64, or 96 bits for hashing. Final hash codes are created using these bits. The proposed method's retrieval performance is compared with current state-of-the-art methods in the literature, and its performance is better than other methods.

CONFLICTS OF INTEREST

No conflict of interest was declared by the author.

REFERENCES

- [1] Nour, M., Sindi, H., Abozinadah, E., Öztürk, Ş., Polat, K., "A healthcare evaluation system based on automated weighted indicators with cross-indicators based learning approach in terms of energy management and cybersecurity", *International Journal of Medical Informatics*, 144: 104300, (2020).
- [2] Font, M. M., "Clinical applications of nuclear medicine in the diagnosis and evaluation of musculoskeletal sports injuries", *Revista Española de Medicina Nuclear e Imagen Molecular (English Edition)*, 39(2): 112-34, (2020).
- [3] Öztürk, Ş., "Stacked auto-encoder based tagging with deep features for content-based medical image retrieval", *Expert Systems with Applications*, 161: 113693, (2020).
- [4] Alsmadi, M. K., "Content-Based Image Retrieval Using Color, Shape and Texture Descriptors and Features", *Arabian Journal for Science and Engineering*, 45: 3317–3330, (2020).
- [5] Jiang, Q. Y., Cui, X., Li, W. J., "Deep discrete supervised hashing", *IEEE Transactions on Image Processing*, 27(12): 5996-6009, (2018).
- [6] Jianhua, X., Adali, T., Yue, W., "Segmentation of magnetic resonance brain image: integrating region growing and edge detection", *Proceedings International Conference on Image Processing*, 544-547, (1995).
- [7] Choi, H., Baraniuk, R. G., "Multiscale image segmentation using wavelet-domain hidden Markov models", *IEEE Transactions on Image Processing*, 10(9): 1309-1321, (2001).
- [8] Zhang, G., Ma, Z. M., Tong, Q., He, Y., Zhao, T., "Shape Feature Extraction Using Fourier Descriptors with Brightness in Content-Based Medical Image Retrieval", *2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 71-4, (2008).
- [9] Chandra, P. N. R. L. C., Prasad, P. S., Kumar, M. V., Santosh, D. H. H., "Image retrieval with rotation invariance", *2011 3rd International Conference on Electronics Computer Technology*, 194-198, (2011).

- [10] Jai-Andaloussi, S., Lamard, M., Cazuguel, G., Tairi, H., Meknassi, M., Cochener, B., "Content based Medical Image Retrieval: use of Generalized Gaussian Density to model BEMD's IMF", World Congress on Medical Physics and Biomedical Engineering, 1249-1252, (2009).
- [11] Ramamurthy, B., Chandran, K. R., Meenakshi, V. R., Shilpa, V., "CBMIR: Content Based Medical Image Retrieval System Using Texture and Intensity for Dental Images", Eco-friendly Computing and Communication Systems, 125-134, (2012).
- [12] Babaie, M., Tizhoosh, H. R., Khatami, A., Shiri, M. E., "Local radon descriptors for image search", 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), 1-5, (2017).
- [13] Öztürk, Ş., "Image Inpainting based Compact Hash Code Learning using Modified U-Net", In 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT2020), 1-5, (2020).
- [14] Beura, S., Majhi, B., Dash, R., "Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer", Neurocomputing, 154: 1-14, (2015).
- [15] Banerji, S., Sinha, A., Liu, C., "A New Bag of Words LBP (BoWL) Descriptor for Scene Image Classification", Computer Analysis of Images and Patterns, 490-497, (2013).
- [16] Hadjiiski, L. M., Tourassi, G. D., Sadek, I., Sidibé, D., Meriaudeau, F., "Automatic discrimination of color retinal images using the bag of words approach", Medical Imaging 2015: Computer-Aided Diagnosis, 9414, (2015).
- [17] Vetrivel, A., Gerke, M., Kerle, N., Vosselman, G., "Identification of Structurally Damaged Areas in Airborne Oblique Images Using a Visual-Bag-of-Words Approach", Remote Sensing, 8: 1-22, (2016).
- [18] Yamashita, R., Nishio, M., Do, R. K. G., Togashi, K., "Convolutional neural networks: an overview and application in radiology", Insights into imaging, 9(4): 611-629, (2018).
- [19] Bressan, R. S., Alves, D. H. A., Valerio, L. M., Bugatti, P. H., Saito, P. T. M., "DOCToR: The Role of Deep Features in Content-Based Mammographic Image Retrieval", 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), 158-163, (2018).
- [20] Bootwala, A., Breininger, K., Maier, A., Christlein, V., "Assistive Diagnosis in Ophthalmology Using Deep Learning-Based Image Retrieval", Bildverarbeitung für die Medizin, 144-149, (2020).
- [21] Kumar, M., Singh, K. M., "Content based medical image retrieval system (CBMIRS) to diagnose hepatobiliary images", In International Conference on Next Generation Computing Technologies, 663-676, (2017).
- [22] Cai, Y., Li, Y., Qiu, C., Ma, J., Gao, X., "Medical image retrieval based on convolutional neural network and supervised hashing", IEEE Access. 7: 51877-51885, (2019).
- [23] Chung, Y. A., Weng, W. H., "Learning deep representations of medical images using siamese CNNs with application to content-based image retrieval", arXiv preprint arXiv: 1711.08490, (2017).
- [24] Ayyachamy, S., Alex, V., Khened, M., Krishnamurthi, G., "Medical image retrieval using Resnet-18", In Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications, 1095410, (2019).

- [25] Latif, A., Rasheed, A., Sajid, U., Ahmed, J., Ali, N., Ratyal, N. I., “Content-Based Image Retrieval and Feature Extraction: A Comprehensive Review”, *Mathematical Problems in Engineering*, 1-21, (2019).
- [26] Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A., “Return of the devil in the details: Delving deep into convolutional nets”, *arXiv preprint arXiv: 1405.3531*, (2014).
- [27] Öztürk, Ş., Özkaya, U., Barstuğan, M., “Classification of Coronavirus (COVID-19) from X-ray and CT images using shrunken features”, *International Journal of Imaging Systems and Technology*. (2020).
- [28] Lehmann, T. M., Güld, M. O., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Wein, B. B., “Content-based image retrieval in medical applications”, *Methods of information in medicine*. 43(04): 354-361, (2004).
- [29] Cao, Y., Qi, H., Gui, J., Li, K., Tang, Y. Y., Kwok, J. T. Y., “Learning to Hash with a Dimension Analysis-based Quantizer for Image Retrieval”, *IEEE Transactions on Multimedia*, (2020).