Channel Transformation Enhanced Deep Convolutional Neural Network enforced Image Retrieval Mechanism for Medical Image Applications

Senthil Kumar Sundararajan, B. Sankaragomathi, D. Saravana Priya

Abstract--- The diversified utilization of digital imaging data in the medical domain has in turn increased the size of the medical image repository. This increase in the size of the repository imposes huge challenges during the process of querying and handling huge databases will lead to the requirement of Content Based Medical Image Retrieval Systems(CBMIR). In this paper, a Channel Transformation Enhanced Deep Convolutional Neural Network-based Image Retrieval Mechanism (CTEDCNN-IRM) is proposed for handling the issue of semantic gap that prevails between human perceived high level semantic information and imaging devices' captured low level visual information in medical imaging applications. The experimental results of the proposed CTEDCNN-IRM confirmed a mean classification accuracy and mean precision rate of 99.83% and 0.78 in the process of image retrieval. This proposed CTEDCNN-IRM is also determined to be well suited and applicable to the processing of multimodal medical images that relates to different body organs.

I. INTRODUCTION

In the recent decade, consistent and dynamic development of digital computers, use of multimedia and incorporation of storage systems emerged in huge multimedia and image content repositories [1]. The development of huge multimedia and image content repositories has innovated its use in the diagnosis and clinical studies of medical applications [2]. Moreover, the hospitals produce large degree of imaging data in medical image databases, since they are equipped with well equipped diagnosis and investigation-based imaging capabilities [3]. Hence, the need for developing a potential large degree of imaging data in medical image databases is essential for helping the clinicians towards the effective handling of large datasets [4] Further, a diversified number of algorithms were contributed in the literature for automated investigation of medical images in order to produce, store and manage large degree of imaging data in medical image databases [5].

In specific, a ContentBased Medical Image Retrieval Systems (CBMIR) is considered to be the predominant approach for effective diagnosis and treatment of diseases through the process of producing, storing and managing large degree of imaging data in medical image databases [6]. This CBMIR system utilizes the features such as texture, color and shape or other relevant features that could be derived from the considered medical image during the process of searching and retrieval. Further, CBMIR systems' potential depends on the kind of selected features considered for medical image searching and retrieval from large medical image databases of hospitals. This CBMIR system initially represents the features using a high dimensional feature space, which then uses a distance metric like Euclidean distance for identifying the closeness between the query image and the medical image stored in the database [7]. Hence, the image data representation highlighted in terms of similarity and features metric is considered to be an indispensable factor during the implementation of any CBMIR system. In spite of many research contributions in the design of an effective CBMIR system formulated based on feature image data representation and similarity metric, the issue of semantic gap is the predominant hurdle that needs to be resolved in the process of medical image searching and retrieval [8]. This semantic gap is the deviation that exists between human perceived high level semantic information and imaging devices' captured low level visual information in medical imaging applications. Furthermore, deep learning is one of the breakthrough in the potential development of machine learning research. This deep learning approaches utilize a diversified number of machine learning algorithms in order to model high degree of data abstractions that incorporates deep architectures for handling multiple non-linear characteristic transformations [9].Thus, deep learning approaches are considered to be predominant in reducing the semantic gap, which is more common in medical image retrieval process. In addition, hybridized CNN and LSTM deep neural network architectures with enhanced channel transformation is determined to be significant in ensuring optimal medical image retrieval process [10].

In this paper, CTEDCNN-IRM is proposed for effective medical image retrieval process by incorporating the merits of CNN and LSTM with updated channel transformation. This proposed CTEDCNN-IRM aims at facilitating high degree feature representations that correspond to different body organs and diversified imaging modalities.

Manuscript received February 01, 2019

Senthil Kumar Sundararajan, Research Scholar, Department of Computer Science, Bharathiar University, Coimbatore, Tamil Nadu, India. (e-mail: senthilsskhp@gmail.com)

Dr.B. Sankaragomathi, Professor, Department of Electronics and Instrumentation, National Engineering College, Kovilpatti, Tamil Nadu, India. (e-mail: Sankaragomathi.paramasivan@gmail.com)

Dr.D. Saravana Priya, Department of Computer Science Engineering, PA College of Engineering and Technology, Pollachi, Tamil Nadu, India. (e-mail: dspriyapacet@gmail.com)

The simulation experiments of the proposed CTEDCNN-IRM is conducted using the evaluation metrics of classification accuracy, prediction, recall and F-score for identifying the predominant in medical image retrieval process.

II. RELEVENT WORK

This section presents the works of the existing ContentBased Medical Image Retrieval systems(CBMIR) contributed in the recent years with its merits and limitations.

Initially, a large Guassian-inspired hierarchical and repulsive balloon method-based medical image retrieval approach was contributed to ensure maximized classification accuracy in the medical field [11]. This Guassian-inspiredmedical image retrieval approach was proposed for discriminating between two kinds of lung cancers that corresponds to the category of squamous carcinoma and adenocarcinoma. This Guassian-inspiredmedical image retrieval approach was also determined to be significant in handling even a half-million of cells that are derived from the dataset considered for training.

Finally, a deep learning method using CNN was proposed for effective training and classification of medical image retrieval in medical applications [20]. This CNN-based deep learning method used an inter-modal data set that possessed five modalities and twenty four classes in order to achieve effective training and classification process in the retrieval process.The class-based forecasting utilization in this CNN-based Deep Learning Method aided in enabling better accuracy having 99.77% with mean precision rate of 0.69 under the task of medical image retrieval process.

III. THE PROPOSED CHANNEL TRANSFORMATION ENHANCED DEEP CONVOLUTIONAL NEURAL NETWORK-BASED IMAGE RETRIEVAL MECHANISM

This proposed Channel Transformation Enhanced Deep Convolutional Neural Network-based Image Retrieval Mechanism (CTEDCNN-IRM) is proposed as a classification inspired framework that is capable of retrieving similar kind of images from the considered medical database considered for exploration. The underlying Channel Transformation Enhanced Deep Convolutional Neural Network model focuses on learning the filter kernel by generating a greater potential description of the data existing in each layer. The underlying Channel Transformation Enhanced Deep Convolutional Neural Network (CTEDCNN) comprises of three layers that correspond to fully connected layers, pooling layers and convolutionlayers .The convolution layer plays the task of generating feature maps through the process of convolving the input feature maps with the kernel. In the fully connected layers, each and every neuron consists of maximum number of connections on par with the compared convolution layers utilized in the improvised Deep Convolutional Neural Network model. This fully connected layers acts as the intermediate layer that separates feature extractor layer with the classifier layer. Further, the pooling layers in the CTEDCNN are designed for scaling down the number of feature maps that are generated by the convolution layers. This pooling layerpresent in the CTEDCNN facilitates the reduction of feature maps through the estimation of local maxima in the neighboring region. This pooling layer is also significant in ensuring translational invariance that in turn minimizes the neuron count is being processed in the subsequent connected layers. In this CTEDCNN, the output of the layer is traditionally considered as a special and separate layer with the model extracting input samples from the input layer.

3.1 Classification phase of CTEDCNN

In this initial phase of CTEDCNN, the method of supervised learning scheme is utilized for the purpose of classifying the medical images by training the incorporated Channel Transformation Enhanced Deep Convolutional Neural Network. In the current phase, the medical images are partitioned into diversified classes based on the information associated with the organ or the body part considered from the dataset. The core objective of the phase focuses on the classification of each medical image into classes as 2D images are considered for investigation. This classification process aids in the essential formulation and transformation that converts it into a problem of multi-class image classification. This classification process incorporates feature extraction and classification module for effective ContentBased Medical Image Retrieval (CBMIR). The proposed CTEDCNN possess the significance of learning the hierarchy of a deep convolution classifier and features for constructing an end-to-end learning framework based on the data derived from the training medical images. This proposed CTEDCNN incorporates the learning of abstract, mid-level and low-level features from the medical images as it is contrary to the option of considering domain specific consideration under the context of handcrafted features used for investigation. Thus, the proposed CTEDCNN is capable in identifying the class associated with the image being queried in a more effective manner, such that the image retrieval work is facilitated through the learned features.

The proposed CTEDCNN possess a training model that comprises of eight layers that includes five convolution layers (CVL1, CV2, CV3,CV4 and CV5) and three fully connected layers (FCL1, FCL2 and FCL3) as portrayed in Figure 1. The proposed CTEDCNN considers gray scale images of 224x224 dimensions as input in contrast to the model that utilizes a reduced number of kernels as presented in [21]. The input gray scale image is filtered by CVL1 of size 11 x 11 with 64 kernels is converted into an equivalent stride of 4 pixels. In this context, stride refers to the distance determined between the neighborhood neuron centers related to the corresponding domain existing in the kernel map. The output of CVL1 is then fed into a series of a nonlinear and spatial max pooling for consolidating neighboring neurons.The importance of nonlinearity is imposed through the incorporation of the rectified linear unit, which is applicable to the outputs of five convolution layers and three fully connected layers.

This rectified linear unit aids in eliminating gradient problems that in turn helpin increasing the potential of getting trained several times compared to its equivalent hyperbolic tangent units. The output of CVL1 is considered as input to CVL2 for filtering it into size of 5 x 5 with 192 kernels with the help of rectified linear unit and spatial max pooling layer. Further, the output of CVL2 is considered as input to CVL3 for filtering it into size of 5 x 5 with 384 kernels with the help of rectified linear unit and spatial max pooling layer. Furthermore, the last two convolution layers CVL4 and CVL5 possess size of 3 x 3 with 256 kernels respectively. The three fully connected layers (FCL1, FCL2 and FCL3) equally consist of 4096 neurons for the potential training process. Finally, probability distributions associated with each class are determined through the application of Softmax function that comprises of 24 outputs over the last fully connected layers FCL3. Hence, the probability distributions associated with each class are of size 1 x 24 with each vector element related to the class of medical image dataset.

A pooling layer is used subsequently after the first, second and fifth convolution layer for the effective classification process. This incorporated pooling layer having the grid of pooling entities that are spaced at a distance of 2 pixels with each pixel emphasizing the significance of the neighboring neurons. These grids of pooling entities are also centered at the proximity of 3 x 3 considered to the median of the pooling unit. However, the complexity in over fitting the model under training under overlapping pooling operation is high. Thus, the benefits of dropout regularization layer pertaining to the first and second FCLs are incorporated for eliminating the issue of over fitting based on probability $1-\delta$ with δ corresponding to the remaining probability of utilized neurons. Furthermore, the possible connections entering into and out of the dropped neurons are discarded during the training period, since it does not contribute any potential task in the forward and backward pass of the training process.

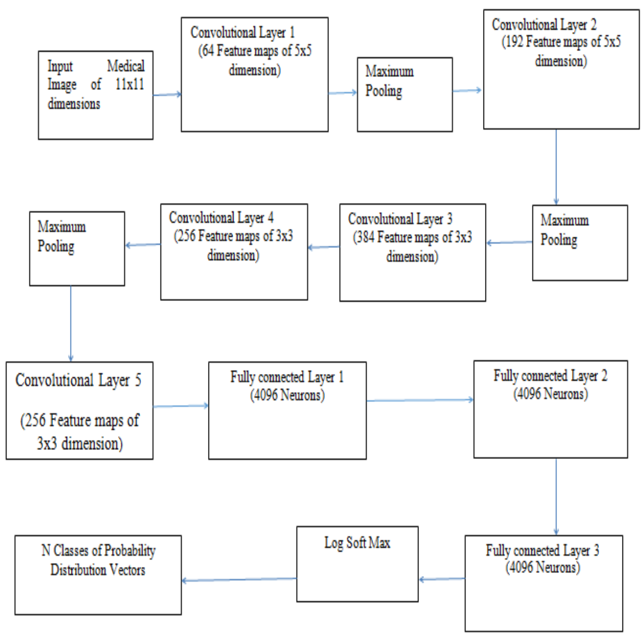


Figure 1: The architecture of the proposed CTEDCNN approach used for CBMIR

3.2 Training phase of the proposed CTEDCNN

There are two type of phases exists ie training phase,evaluation phase.When the training phase is completed the neurons which are dropped out are selected once again with the orginal weights for the next level training stage, In the next phase ie evaluation phase all the selection neurons are used once again with dropping them.. However, to balance the expected values of the neurons in the test phase to that of the training phase a factor of keeping probability ‘p’ weights them. Finally, the output of FCL3 is fed to input of log Softmax with four thousand nighty six inputs and and twenty four outputs.

For a N dimensional vector, log softmax imposes a softmax function to extend the valuesof the vector in the range of [0,1] and its addition gives a value of 1. By accomplishing this it allows probability distribution to each and every class.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The predominance of the proposed CTEDCNN is investigated using Torch7, which is the widely popular learning tool utilized for implementing and training the deep learning architecture used in contentbased medical image retrieval. The simulation of the proposed CTEDCNN is conducted using Dell Inpiron Laptop with Linux Version 14.04 with Intel Core of i3 CPU and 6 GB RAM with the considered 2.4GHz of clock speed. The proposed CTEDCNN is investigated using retrieval and classification results by utilizing medical images collected from the medical databases that are publicly available. The investigation of the proposed CTEDCNN is performed over the classes that are derived pertaining to the body organs such as lymph nodes and lung organs. The data set used for determining the performance of the proposed CTEDCNN consists of 24 classes, out of which 22 classes are extracted from diversified public medical images. The complete dataset consists of 7200 medical images, in which a random sample of 300 images is considered for determining the significance of the proposed CTEDCNN scheme towards effective image retrieval. In the process of testing and training nearly 30% and 70% of the images are used for testing and training from the images derived from the dataset that are partitioned into classes that have unique characteristics. Thus, 2160 and 5040 number of images are considered as the testing and training set from the collection of 7200 medical images for predominant investigation. Further, the testing and training set images are mutually exclusive in nature with TIF format images. Furthermore, the input medical images that pertain to each and every class used for analysis are resized to 256x256 size images. In addition, the classes of the input medical images are labeled using numerical values for facilitating effective supervised learning process.

Then, the investigation of the suggested CTEDCNN is conducted using Precision and recall value with and without class prediction for estimating the quality of medical image retrieval.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Figure 2 and 3 highlights the proposed CTEDCNN scheme is predominant in terms of Precision and recall value investigated with and without class prediction, since feature representations derived from the three entirely connected layers are superior in this medical image retrieval scheme.

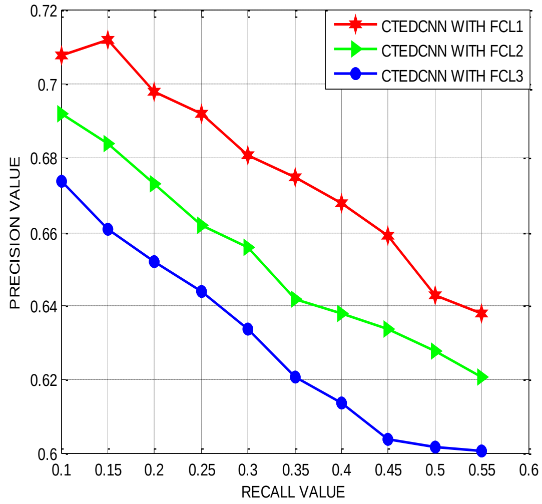


Figure 2: Proposed CTEDCNN-Precision and recall value under class prediction

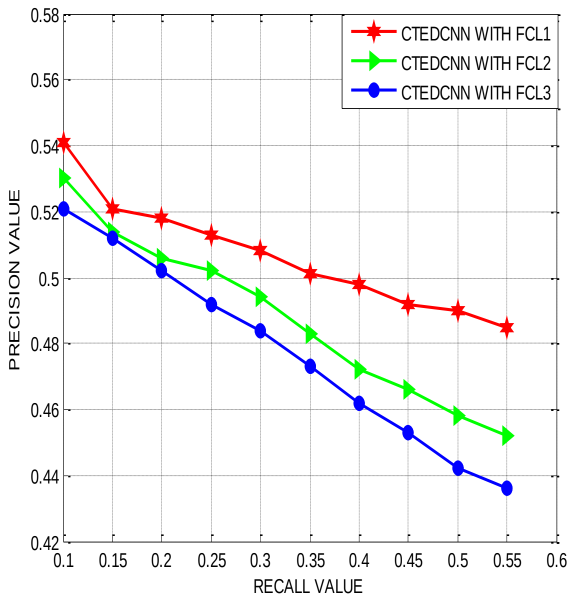


Figure 3: Proposed CTEDCNN-Precision and recall value without class prediction

The precision and recall value under FCL1 incorporated in the proposedCTEDCNN scheme is determined to be enhanced in an average by 16%, 13% and 11% excellent to the precision and recall value under the incorporation of FCL2 and FCL3.

Further, the significance of the proposed CTEDCNN is conducted using Average precision value(AP), the average recall value (AR), accuracy and F1 measureis illustrated using Table 1.

Furthermore, the potential of the proposed CTEDCNN is conducted using mean precision Average precision value (AP), the average recall value (AR), accuracy and F1 measure is illustrated using Table 2.

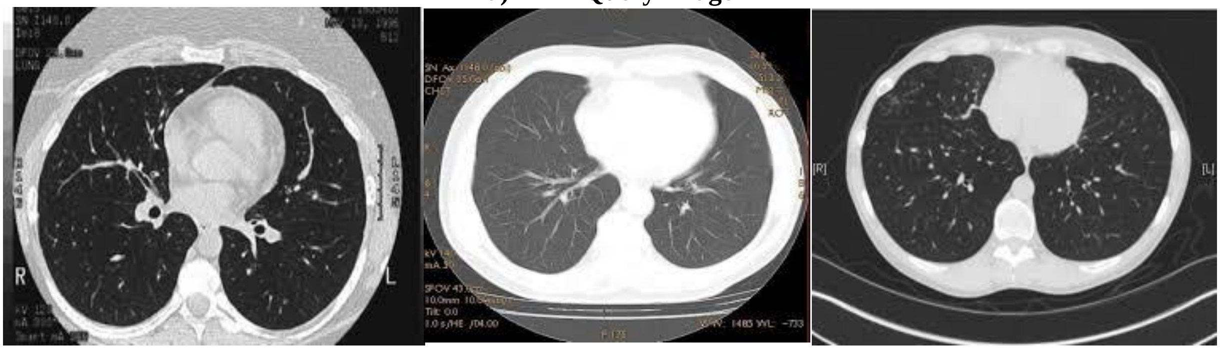
Table 1: Average recall (AR), Average precision (AP), accuracy and F1 measure of the proposed CTEDCNN-IRM with benchmarked deep learning-based state of art schemes

Scheme	**Images**	**Training set of Images**	**Testing set of Images**	**Classes**	**Accuracy**	**AP** (In %)	**AR** (In %)	**F1** (In %)
Proposed CTEDCNN-IRM	7200	5040	2160	24	99.82	99.86	99.79	99.86
CNN-based Deep Learning Method [20]	7200	5040	2160	24	99.76	99.81	99.77	99.75
Spatial Matching of Visual Words for medical image retrieval [19]	6400	5760	640	28	99.12	99.65	98.34	98.41
Optimized Multidimensional Spectral Hashing-based Indexing [18]	6400	5760	640	28	98.56	99.57	97.28	97.84

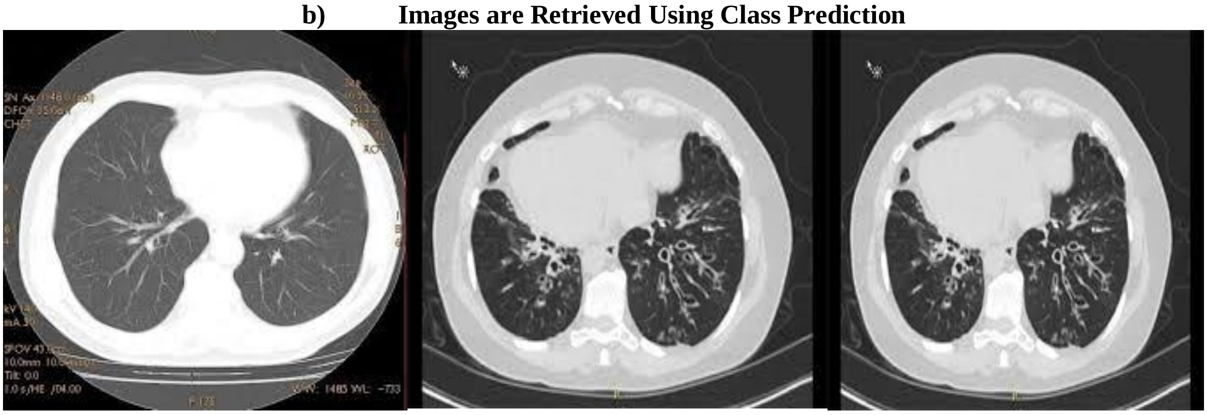
Table 2: Mean precision investigated with the proposed CTEDCNN-IRM with benchmarked deep learning - Established medical image retrieval approaches

Scheme	**Images**	**Training set of Images**	**Testing Set of Images**	**Classes**	Modalities	**Mean Precision**
Proposed CTEDCNN-IRM	7200	5040	2160	24	OPT , PET, PT, CT, MR	0.82with predictionclass 0.77 prediction without class
CNN-based Deep Learning Method [20]	7200	5040	2160	24	OPT , PET, PT , CT, MR,	0.69with prediction class 0.56prediction without class
Spatial Matching of Visual Words for medical image retrieval [19]	6400	5760	640	28	MR, CT, PT, PET, OPT	0.67with prediction class 0.52prediction without class
Optimized Multidimensional Spectral Hashing-based Indexing [18]	6400	5760	640	28	MR, CT, PT, PET, OPT	0.64withprediction class 0.46prediction without class

a) Query Image

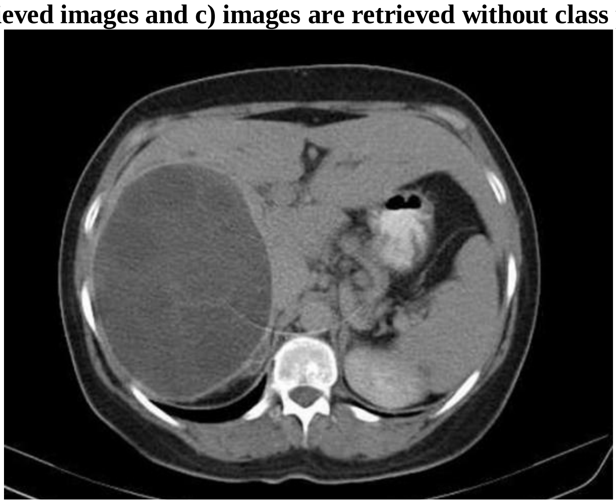


b) Images are Retrieved Using Class Prediction



c) Images are Retrivied without Using Class Prediction

Figure 4: Proposed CTEDCNN-IRM-based retrieval results for lymph image: a) Query image, b) Class prediction-based retrieved images and c) images are retrieved without class prediction



a) Query Image

b) Image Retrived Using Class Predicition

c) Image are Retrived Using with out Class Prediction

Figure 5: Proposed CTEDCNN-IRM-based retrieval results for liver image: a) Query retrievied image, b) Class prediction-based retrieved images and c) Images are retrivied beyond the class prediction

In addition, Figure 4 and 5 exemplars the potential of the proposed CTEDCNN-IRM scheme quantified in terms of effective medical image retrieval process using lymph and liver data set. The mean precision of the proposed CTEDCNN-IRM scheme is determined to be 0.82 with class prediction and 0.77 without class prediction, which is a considerable improvement over the compared CNN-based Deep Learning Method [20] that possess mean precision of 0.69 with class prediction and 0.56 without class prediction.

V. CONCLUSIONS

This proposed CTEDCNN-IRM was presented based on the enhancement of channel transformation over CNN for effective medical image retrieval process with the view the resolve the impacts of semantic gap. The mean classification accuracy of the proposed CTEDCNN-IRM was determined to be 99.82% during the training the network using 24 diversified classes of medical images. The utilization of final three FCLs is another predominant improvement in the proposed CTEDCNN-IRM, since it has ensured competent feature extraction process for efficient image retrieval task. The average precision and recall value of the recommended CTEDCNN-IRM was estimated to be 0.75 and 0.82 under class prediction of multimodal medical image retrieval process. In the near future, it is also planned to formulate a hybrid deep neural network that integrates the characteristics of CNN and LSTM for ensuring efficient image retrieval through effective training process

REFERENCES

1. Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., & Mougiakakou, S. (2016). Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. *IEEE Transactions on Medical Imaging*, *35*(5), 1207-1216.
2. Zhang, F., Song, Y., Cai, W., Hauptmann, A. G., Liu, S., Pujol, S., … Chen, M. (2016). Dictionary pruning with visual word significance for medical image retrieval. *Neurocomputing*, *177*(1), 75-88.
3. Quellec, G., Lamard, M., Cazuguel, G., Cochener, B., & Roux, C. (2012). Fast Wavelet-Based Image Characterization for Highly Adaptive Image Retrieval. *IEEE Transactions on Image Processing*, *21*(4), 1613-1623.
4. Rahman, M., Antani, S., & Thoma, G. (2011). A Learning-Based Similarity Fusion and Filtering Approach for Biomedical Image Retrieval Using SVM Classification and Relevance Feedback. *IEEE Transactions on Information Technology in Biomedicine*, *15*(4), 640-646.
5. Jiji, G. W., & Durai Raj, P. S. (2015). Content-based image retrieval in dermatology using intelligent technique. *IET Image Processing*, *9*(4), 306-317.
6. Cai, W., Zhang, F., Song, Y., Liu, S., Wen, L., Eberl, S., … Feng, D. (2014). Automated feedback extraction for medical imaging retrieval. *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, *1*(1), 56-68.
7. Saritha, R. R., Paul, V., & Kumar, P. G. (2018). Content based image retrieval using deep learning process. *Cluster Computing*, *1*(1), 67-78.
8. Wu, L., & Hoi, S. C. (2011). Enhancing Bag-of-Words Models with Semantics-Preserving Metric Learning. *IEEE Multimedia*, *18*(1), 24-37.
9. Shin, H., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., … Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics

and Transfer Learning. *IEEE Transactions on Medical Imaging*, *35*(5), 1285-1298.

10. Yogapriya, J., Saravanabhavan, C., Asokan, R., Vennila, I., Preethi, P., & Nithya, B. (2018). A Study of Image Retrieval System Based on Feature Extraction, Selection, Classification and Similarity Measurements. *Journal of Medical Imaging and Health Informatics*, *8*(3), 479-484.
11. Markonis, D., Schaer, R., & Müller, H. (2015). Evaluating multimodal relevance feedback techniques for medical image retrieval. *Information Retrieval Journal*, *19*(1-2), 100-112.
12. Seetharaman, K., & Sathiamoorthy, S. (2016). A unified learning framework for content based medical image retrieval using a statistical model. *Journal of King Saud University - Computer and Information Sciences*, *28*(1), 110-124.
13. Zhang, X., Liu, W., Dundar, M., Badve, S., & Zhang, S. (2015). Towards Large-Scale Histopathological Image Analysis: Hashing-Based Image Retrieval. *IEEE Transactions on Medical Imaging*, *34*(2), 496-506.
14. Li, P., & Ren, P. (2017). Partial Random Spherical Hashing for Large-Scale Image Retrieval. *2017 IEEE Third International Conference on Multimedia Big Data (BigMM)*, *1*(2), 13-25.
15. Mengting, L., & Jun, L. (2017). Deep hashing for large-scale image retrieval. *2017 36th Chinese Control Conference (CCC)*, *1*(1), 45-54.
16. Zhang, J., & Peng, Y. (2017). SSDH: Semi-supervised Deep Hashing for Large Scale Image Retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, *1*(1), 1-1.
17. Balasundaram, R., & Sudha, G. F. (2017). An efficient and reduced memory indexing approach based on priority rank spectral hashing for multibiometric database. *International Journal of Biometrics*, *9*(2), 113-122.
18. Revathi, B., & Sudha, G. (2018). Retrieval performance analysis of multibiometric database using optimized multidimensional spectral hashing based indexing. *Journal of King Saud University - Computer and Information Sciences*, *1*(1), 67-78.
19. Shamna, P., Govindan, V., & Abdul Nazeer, K. (2018). Content-based medical image retrieval by spatial matching of visual words. *Journal of King Saud University - Computer and Information Sciences*, *1*(2), 34-48.
20. Qayyum, A., Anwar, S. M., Awais, M., & Majid, M. (2017). Medical image retrieval using the deep convolutional neural network. *Neurocomputing*, *266*(1), 8-20.