(IJAER) 2019, Vol. No. 18, Issue No. II, August **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

# DEEP LEARNING FOR CONTENT BASED MEDICAL IMAGE RETRIEVAL

**Sheela M. G.**
*Lecturer in Electronics,*
*Women's Polytechnic College, Ernakulam, Kerala.*

## ABSTRACT

*The size of medical image repositories is increasingly growing with the widespread use of digital imaging images in hospitals. The management and querying of these broad databases is also difficult to accomplish content based medical image retrieval (CBMIR) systems. The semantic distance between low visual information obtained by imaging devices and high semanthemistic information interpreted by humans is a significant challenge in CBMIR systems. In terms of feature representations that fully characterize high-level details, the efficiency of such systems is more important. In this paper, we propose a structure for profound training for CBMIR by the use of the Deep Convolutional Neural Network (CNN). The network is trained with an intermodal dataset comprising twenty four classes and five modalities. Medical images are retrieved using the learned characteristics and classification outcomes. Best results are obtained for retrieval by using class predictions. The recovery work achieves an average grade accuracy of 99.77% and an average accuracy of 0.69. It is the safest way to acquire multimodal diagnostic images for various body organs.*

***Keywords:*** *Deep Learning; Content Based Medical Image Retrieval (CBMIR); Similarity Metric; CBMIR*

## INTRODUCTION

The rapid development in the last few years has led to vast images and multimedia content repositories for various computers, multimedia and storage systems. The advances in digital storage and information delivery are also useful in clinical and diagnostic studies. Diagnostic and survey imagery hospitals generate a large number of imagery data and thereby contribute to a significant increase in the production of medical image collections. Therefore, the implementation of an efficient framework for recovery of medical images is important to help clinicians browse these large datasets. Many literary algorithms have been suggested for automated analysis of medical images to promote the creation and management of such broad medical image databases[1-5]. A medical image retrieval system based on content (CBMIR) can be used to complement diagnosis and treatment of different diseases and also to process vast volumes of data effectively.

Content based image retrieval (CBIR) is A technology for a computer vision that enables the search of images related to broad bases of data. This research is focused on the image characteristics such as color, texture and form, or some other image characteristics. The efficiency of a CBIR device depends primarily on these functions [6]. The images are shown in a high-dimensional function space first of all. Then the image similarity in the database with that in a query image is measured using distance metrics such as the Euclidian distance in the feature space. Therefore, the most important components of CBIR systems are image data representation with regards to features and the option of similitude steps. Many scientists have researched these areas extensively [7], but the most challenging problem remains the "semantic gap" reduction in CBIR systems.

# RELATED WORK

Existing research work is briefly discussed in this section.

## Content Based Image Retrieval (CBIR)

A general CBIR framework block diagram is shown in the Fig. 1. The CBIR is based on images extracted from, or derived from image contenue from, the broad databases [8]. The first phase is offline, the second is online. There are normally two phases in every CBIR system. Features are extracted in the offline process from large sets of images (used to train the system) to construct a database of local features. This process normally takes time and depends on how many training images the machine is used to train. The same functions are extracted in the online process from the query picture and distance metrics are measured from the query picture features and from the image database functions for similarity calculation. These images are then displayed to the user in terms of results of high similarity or low distance. In both steps, the process used for pre- and extraction of features is the same.

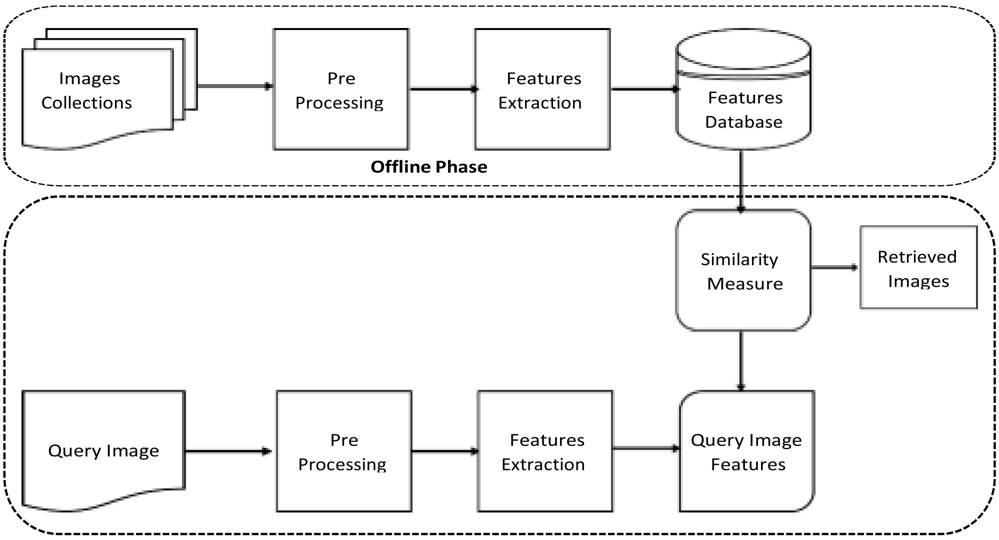



Fig. 1. Block diagram of a generic CBIR system.

The topic of semantic gaps was explored with local characteristics recently by machine learning techniques, some of which were based on learning havoc or compact codes [9-11]. A new way to minimize semantine gaps has been provided by advanced machine learning techniques and deep learning has provided an opportunity to bridge these gap by learning visual characteristics straight from images with no hand-crafted characteristics.

### Content Based Medical Image Retrieval (CBMIR)

As the scale of medical image collections grows rapidly with the widespread distribution of Picture Archiving and Communications Systems (PACS) in hospitals. Therefore, the implementation of an efficient system for medical image recovery is important to handle such large medical databases. In addition to this database management mission, a specialized method of CBMIR lets physicians decide objectively on a specific disease or lesions. Using similar photos and case reports, doctors will determine better the stage and diagnosis of the disease of the patient [1]. Global features extraction systems did not provide compact feature representations for medical images, because clinically positive information is highly located in small image regions [8]. The [1] method for the diagnosis of Alzheimer's disease based on a Bag of Visual Words (BoVWs) based on the use of SIFT features was proposed for a magnetic resonance image MRI. As a feature vector for the picture corresponding they proposed the Laguerre Circular Harmonic Functions Coefficients (LG-CHF). Global features extraction systems did not provide compact feature representations for medical images, because clinically positive information is highly located in small image regions [8]. The [1] method for the diagnosis of Alzheimer's disease based on a Bag of Visual Words (BoVWs) based on the use of SIFT features was proposed for a magnetic resonance image MRI. As a feature vector for the picture corresponding they proposed the Laguerre Circular Harmonic Functions Coefficients (LG-CHF). A regression function was used to optimize recovery efficiency. In [5], a biomedical image retrieval classification-driven supervised method was suggested. The image filtering and similitude fusion were used to predict the type of image query as the basis and a multi-class support vector machine (SVM). Thus, the search area in wide database for similitude calculation is reduced by deleting the irrelevant images.

### Deep Learning

Deep learning is a topic of machine learning that uses a variety of algorithms to model high-level abstractions in data using a profound architecture that has several layers of transformation with linear and nonlinear transformation functions. Deep learning began in 1965, but only recently made substantial progress, with enhanced computational capacity, nonlinearities that allow deeper networking and more efficient ways to initialize deep networking [12, 13]. Deep learning is based on artificial neural networks that seek to emulate the workings of the human brain. Neural feed networks with several hidden layers are excellent examples of deep architecture models. The 1980 common standard back propagation algorithm is still an important way to train neural networks [8]. For hand-written digit recognition, a standard back propagation algorithm was used for training multilayer neural networks. The image classification challenge of the ILSVRC-2012 was won by the Deep Convolutional Neural Network (DCNN) presented in the ILSVRC.

The use of deep learning techniques in the field of medicine was seen successfully from recent research. In [14], an interstitial lung disease classification neuronal network dependent framework was presented (ILDs). They had 7 classes, six of which consisted of ILD models and a balanced tissue class. In characterizing lungs patterns, they achieve 85.5%

classification accuracy. In [15] a Boltzmann machine-based approach to lung compute tomography (CT), which combines generative and discriminatory representation learning, has been implemented for convolutionary classification. Two methods were addressed to two separate data sets: one for classification of the lung texture and another for identification of airways. [16] CNN multi-scale was applied by classifying voxel into brain tissue groups to automatically segment the MR pictures. Depending on the patch size, the network was trained on several image patches with different kernel sizes. In [17], the body organ recognition was introduced as a two-stage deep learning system. The CNN was trained in local patches in the first stage to extract discriminatory and insightful patches from training samples. The second phase consisted of 12 classes of CT which MR 2D slices, and included extracted biased patches for the classification assignment. domain

## METHODOLOGY

In this approach, the implementation of the deep CNN system to boost medical image retrieval efficiency in multiple modes. The full description of the proposed technique is presented in Figure 2. The kernel filter must be discovered by generating a concrete image picture in fundamental Deep CNN system. The images are classified and are finally classified into a multi-class problem. Two essential steps are usually the classification of images like the extraction and classification of features. Deep CNN absorbs and classifies subsequently with more convolutionary functions. Deep Learning takes on tiny, intermediate and abstract image features and for handcrafted features. It is a situation. For medical picture retrieval, the classes are recognized based on the query picture and hence the features extracted used in the retrieval process.

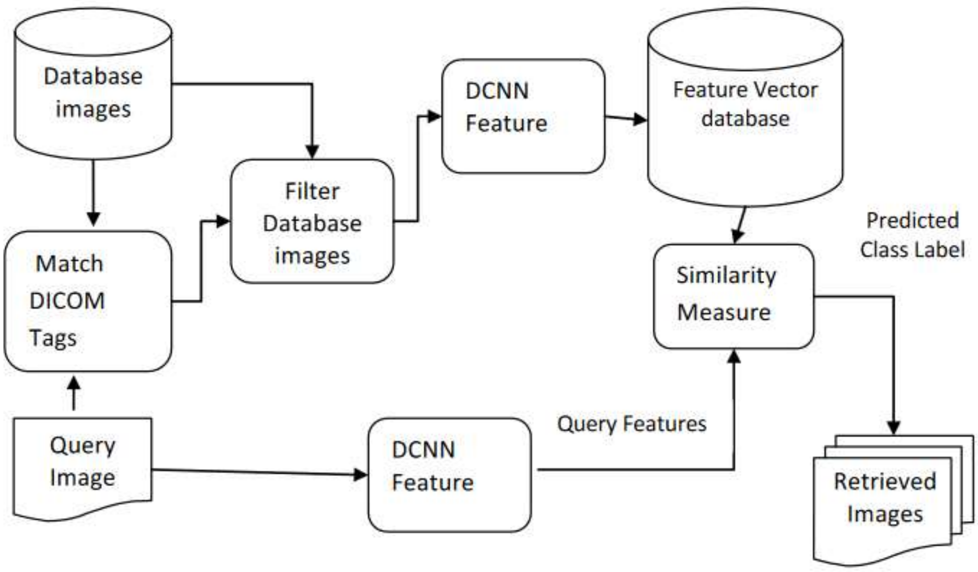



Fig. 2. Proposed Framework for CBMIR

At first, the DICOM Meta Tags are used to retrieve all images from the database image repository. In the functional vector database, the images will then remain separate and persistent. These data are given to the DCNN to establish the related class mark prediction. Convolutionary layer is added to the image and this layer eventually transforms into small areas with its neighboring pixels. The image query inputs is extracted using the functions to be compared using the current DCNN forecasting, so that a list of images that fit closely with higher precision is found. The Euclidian Distance Matrix Function for Retrieval of Image finally tests similitude.

## DATA PRE-PROCESSING

For the construction by neural network models of different layers, we need a coherent aspect ratio, so that every image is resized to 256 X256 dimensions in each class. The images are transformed to gray to eliminate data processing in connection with intensities in images. This tends to miss much of the need to process knowledge. This can also help to reduce processing time. The DICOM picture extract metadata with the Dicom details feature Image Processing Toolbox. This helps to eliminate metadata from all images by using Dicom Matlab functions, such as modality, body part examined, gender in the patient and patient age. Late marks based on Meta data attributes are allocated to the entire image class. The steps performed in pre-processing stage are shown in Figure 3.

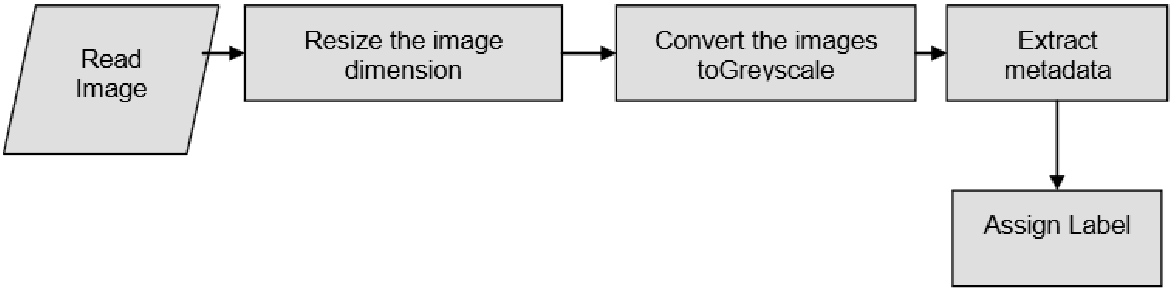



Fig. 3. Data Pre-processing Steps

## THE DCNN MODEL ARCHITECTURE

The 5convolutionary layer DCNN contains here, 3 fully linked layers were used, and max-pool. This model is presented in CL 1, Cl2, Cl3, CL4 &CL5 , PL1, PL2 & PL3 and FCL1, FCL2 & FCL3 respectively in the transforming, polling, and fully linked layers. In CL1 consist of 96 filters with an average size 11 x 11. After that output, from a CL1 feed to a ReLU activation feature, the size is reduced to PL1. Next, CL2 has processed the PL1 output with 256 size filter 5 x 5 stage 1 and 2 padding. The CL3 and the CL4 contain 384 kernels with 3 x 3 size filter and the stride value is 1. 256 kernels in CL5 have 3 x 3 filter dimensions.. PL2 is processed with the same dimension, stride and padding following CL2 and PL 3 is used after CL5. There are 4096 neurons in FCL1 and FCL2 and there are 1000 neurons in FCL3. The SGD algorithm was trained after construction of the DCNN model. The first step in the model training is to load the data into your desired analysis tool. Adopted the SGD (Stochastic

Gradient Descent with Back propagation) [19] for training the models. The word "Stochastic" refers to the randomness involved in the training process. This will be used to update the weights to achieve better accuracy. This is being used to with aim to minimize error function.

**Pseudo Code:**

```
FUNCTION TRAIN_NN_SGD(nn_structure1, P, Q, iteration_number=3000, alpha1=0.25,
lamb1=0.000):
   W1, b1 ← weight intilization(nn_structure1)
   count ← 0
   m1 ← length(Q1)
   average_cost_function ← []
   OUTPUT 'SGD for {} iterations'
   While count<iteration_number:
      if count%50 = 0:
         OUTPUT 'iteration {} of {}'.format(count, iteration_number)
      endif
      triaing_W1, trianing_b1 ← initilation_trianing_values(nn_structure1)
      average_cost ← 0
      for i ← 1 tolength(Q1)
         Delta1 ← {}
          x, y← feedforward(P[i, :], W1, b1)
         for l ← 0 to length(nn_structure1)
            if l = length(nn_structure1):
               delta1[l] <- calculateoutlayer_delta1(Q[i,:], x1[l], y1[l])
               average_cost += np.linalg.norm((y1[i,:]-x1[l]))
            else
               if l > 1:
                  delta1[l] ← calculatehidden_delta1(delta1[l+1], W1[l], y1[l])
               endif
               trianing_W1[l]                                              ←
np.dot(delta1[l+1][:,np.newaxis],np.transpose(h[l][:,np.newaxis]))
```

```
               trianing_b1[l] ← delta1[l+1]
              endif
          endfor
      endfor
         for l ← 0 to length(nn_structure1) - 1
              W1[l] += -alpha * (trianing_W1[l] + lamb * W1[l])
              b1[l] += -alpha * (trianing_b1[l])
          endfor
     endfor
      average_cost ← 1.0/m * average_cost
      average_cost_function.append(average_cost)
      count += 1
  endwhile
  return W1, b1, average_cost_function
```

(IJAER) 2019, Vol. No. 18, Issue No. II, August **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

**Similarity Measure**

The most common metric for measuring the distance between CBMIR approaches is the euclidean distance. Similarity images obtained from qualified database image by obtaining the similarity attribute for query image. The distance from vector a to b is:

$$\mathbf{d(q,r)} = \sqrt{\sum_{i=1}^{p}(q_i - r_i)}$$

Where qi,ri denotes the query and image database characteristics.

## RESULT

TCIA (The Cancer Image Archive) provides the dataset used in the experimentation of the CBMIR[19]. In the archive, the dataset mainly uses DICOM image format. Data is accessible from the 3 Modalities-CT, MRI and Ultrasound sets. 22 image groups of 3 modality forms are used. In this paper there are 22 separate classes, including brain, breast, liver, stomach, leg, kidneys, poitrine, core, cervix, uterus, thyroid, rectum, prostate, fantasy, pancreas, ovary, lung, head neck, head, colon and chest for study. 6600 pictures covering 300 pictures for each class are included. The DICOM files had a.dcm extension. The training and test data sets used in this study are 70% and 30%. This image was taken based on the image's metadata and the DCNN features. The question input picture is shown in figure 4(a) and uterus of female ages 37 and 4(b) indicates that the pics retrieving for the specific input picture have been shown. The output matrix for picture recovery is shown in Figure 5 based on precise calculation, recall and accuracy.

The accuracy of the images can be correctly retrieved depending on the query. Precision is the classifier's ability not to mark a negative sample as positive. All of the inputs are tested and the model based on the real positive values is correct. This measure is rather advantageous if the costs of false positives are high. This can be represented as P (DICOM images / DICOM images) based on the question given. Remember is the fraction of documents that are essential images from the DICOM images retrieved that can be interpreted as R (DICOM Images retrieved / valid DICOM images). Recall is the classifier's ability to find all positive samples from the question given. Remembering is often referred to as sensitivity, which is one aspect of the true positive samples.

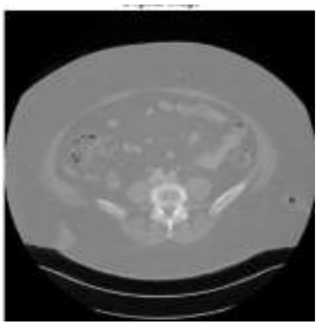
In summary, the accuracy of a fraction of the correctly predicted queried DICOM images to all predicted DICOM images to decide the closest match for a given query image from non-relevant images. Note, the total number of true DiCOM images is a fraction of the expected DICOM image. F calculation is critical in the recovery of medical images and measured using accuracy and retrieval values. Following efficiency matrices for CBMIR will test the proposed system.

(IJAER) 2019, Vol. No. 18, Issue No. II, August **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

$$Recall = \frac{TruePostive}{TruePositive + FalseNegative}$$

$$FMeasure = 2 * \frac{Precision * Recall}{Precision + Recall}$$



(a) Original Image

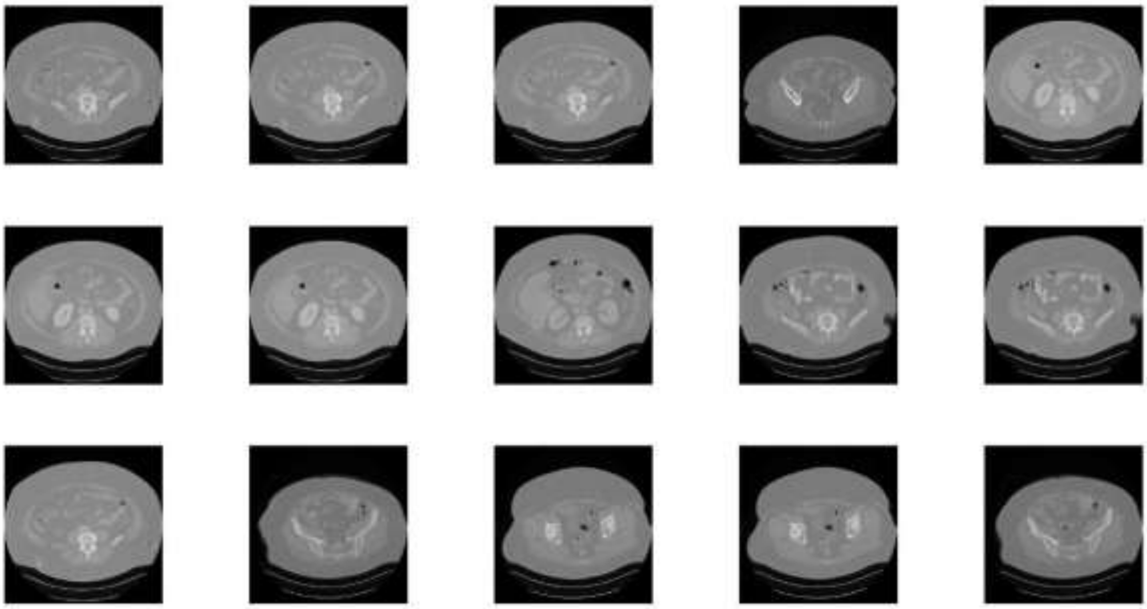


Fig. 4. Retrieval result (a) Query Image, (b) Retrieved Image

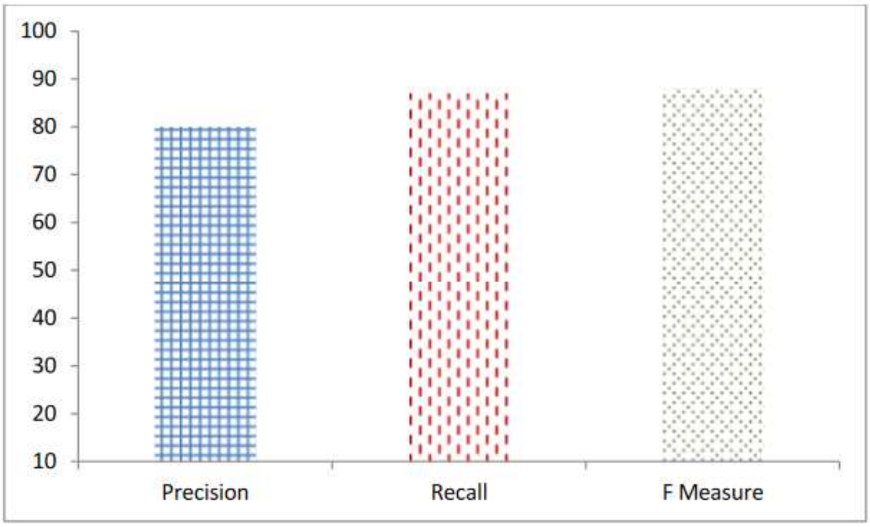



Fig. 5. Performance Metrics

We see Recall as significant based on the question via DCNN with SGD approach. Generally, if classes with less samples and strong for classes with big ones, the F-Score will be less. We see this F-Measure at 88 with over 22 classes and 6,600 images. In comparison with the previous results as given in Table 1, the results have been better. In searches expected to be semantic, we can see a better accuracy than previous results when accuracy is 0, 80.

| Method | Images | Modalities | Classes | Precision |
|---|---|---|---|---|
| DCNN with SGD (Proposed) | 6600 | CT, MRI, Ultrasound (Multi-Modality) | 22 | 80 |
| DCNN Trained on whole images[10] | 7200 | MR,CT,PT,PET,OPT(Multi-Modality) | 24 | 69 |
| LBP, SVM, and auto encoder[21] | 14410 | X-ray (Single Modality) | 57 | 86 |
| Auto regressive model and binary tree based SVM[22] | 6400 | Multi-Modality | 83 | 57 |

Table 1. Comparison of retrieval accuracy of the proposed DCNN

## CONCLUSION:

In this article, deep CNN learning is proposed for CBMIR. A low level and high semanthropic functionality along with DCNN features were extracted from the medical image database. A deep CNN classification system has been developed that can help with DICOM classification with pre-filtering by Meta data. The work proposed reduces the semantine distance by learning descriptive characteristics directly from the images and reduces search space. The precise rate for multimodal image data in this device is 0.80 for the given query image, compared to the

(IJAER) 2019, Vol. No. 18, Issue No. II, August **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

current performance. In addition, analysis methods will boost the accuracy of big data and acclimatize the 3D volumetric image network.

## REFERENCES

1. M. Mizotin, J. Benois-Pineau, M. Allard, and G. Catheline, "Feature-based brain MRI retrieval for Alzheimer disease diagnosis," in 2012 19th IEEE International Conference on Image Processing, 2012, pp. 1241-1244.
2. G. W. Jiji and P. S. J. D. Raj, “Content-based image retrieval in dermatology using intelligent technique,” IET Image Processing, vol. 9, no. 4, pp. 306-317, 2015.
3. M. Ponciano-Silva et al., "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, 2013, pp. 41-46.
4. G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, and C. Roux, “Fast wavelet-based image characterization for highly adaptive image retrieval,” IEEE Transactions on Image Processing, vol. 21, no. 4, pp. 1613-1623, 2012.
5. M. M. Rahman, S. K. Antani, and G. 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, vol. 15, no. 4, pp. 640-646, 2011.
6. Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, “A survey of content-based image retrieval with high-level semantics,” Pattern recognition, vol. 40, no. 1, pp. 262-282, 2007.
7. J. Wan et al., "Deep learning for content-based image retrieval: A comprehensive study," in Proceedings of the 22nd ACM international conference on Multimedia, 2014, pp. 157-166.
8. K. K. Kumar and T. V. Gopal, "A novel approach to self order feature reweighting in CBIR to reduce semantic gap using Relevance Feedback," in Circuit, Power and Computing Technologies (ICCPCT), 2014 International Conference on, 2014, pp. 1437-1442
9. L. Zhang, Y. Zhang, X. Gu, J. Tang, and Q. Tian, “Scalable similarity search with topology preserving hashing,” IEEE Transactions on Image Processing, vol. 23, no. 7, pp. 3025-3039, 2014.
10. M. Norouzi, D. J. Fleet, and R. R. Salakhutdinov, "Hamming distance metric learning," in Advances in neural information processing systems, 2012, pp. 1061-1069.
11. H. Jegou et al., “Aggregating local image descriptors into compact codes,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 9, pp. 1704-1716, 2012
12. L. Deng and D. Yu, “Deep Learning,” Signal Processing, vol. 7, pp. 3-4, 2014.
13. A. G. e. Ivakhnenko and V. G. Lapa, Cybernetic predicting devices, CCM Information Corporation, 1965.
14. M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, “Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional

Neural Network," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1207-1216, 2016.

15. G. van Tulder and M. de Bruijne, "Combining Generative and Discriminative Representation Learning for Lung CT Analysis With Convolutional Restricted Boltzmann Machines," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1262-1272, 2016.
16. P. Moeskops et al., "Automatic segmentation of MR brain images with a convolutional neural network," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1252-1261, 2016.
17. Z. Yan et al., "Multi-Instance Deep Learning: Discover Discriminative Local Anatomies for Bodypart Recognition," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1332-1343, 2016.
18. H.-C. Shin et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1285-1298, 2016.
19. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Prior, F. (2013). The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. Journal of Digital Imaging, 26(6), 1045–1057. https://doi.org/10.1007/s10278-013-9622-7