

An Insight Framework for Content based Medical Image Retrieval using Deep Convolutional Neural Network

HariPriya P, Porkodi R

Abstract: Content Based Image Retrieval system is essentially proven techniques for retrieving the DICOM based images. DICOM format is use to acquire, store and shared medical images. Typically, retrieval of image under query upon the medical image databases is performed by fusing the low-level and high-level descriptors along with DCNN features. In this paper, DICOM image meta data are extracted by using dicom function and create the local database for reserve the information. To accomplish the initial search and retrieve the images by using extraction of Semantic information. The DICOM tags are extracted from the DICOM images and use DCNN feature to build a feature vector database. Subsequently prediction is done based on the by leveraging the convolution layers based on the meta data along with DCNN image features. This paper attempts to implement pre filter method to all DICOM images which further decrease the searching no of images, searching time and ultimately gives fast processing. DCNN based prediction model was constructed and finest results are accomplished. Average accuracy of precision and recall up to 0.80 and 0.87 respectively is achieved based on precision value which would be suitable for high quality image retrieval based on semantic information confined in the image.

Keyword: CBMIR, DICOM, DCNN, Semantic Gap

I. INTRODUCTION

As the medical field and domains are emerging the medical imaging becomes the fundamental cornerstone. Images which are Multidimensional (3D or Time-Varying) are good examples if diverse range of modalities in the medical imaging. These help a lot in the areas such as diagnosis, teaching and research. In general with respect to images, Content Based Image Retrieval (CBIR) is done by properties such as shape, color, texture etc. , for non-medical images [3]. Hospital and Healthcare institutions predominantly rely upon the Computer Aided Diagnostics (CAD). The important information on the image is embedded as part of the Medical Image.

The data accompanied the image depends on the manual data entry or manual labeling process which many times are not efficient. This could also vary based on the experience and knowledge of medical imaging device operators. This would also pose as a challenge for the physician to give the right inference by doing a comparison of the images with the previous history of images.

The residues of the paper involved as following. Section 2 describes the content-based image retrieval, Section 3 presented their view of related works, Section 4 briefly discuss about concept of deep convolutional neural network, Section 5 provides the debate of the details about proposed methodology, Section 5 demonstrate the results and finally, Section 6 provides conclusion of this study.

II. CONTENT BASED MEDICAL IMAGE RETRIEVAL (CBMIR)

CBMIR is very good tool for manage this problem better. We can categories the CBMIR into 5 categories such as 2 Dimensional, 3 Dimensional, Use of Non-Image Data, and retrieval of Multi Modal images from varied datasets. [4].The essential ingredient of any CBMIR should be is give most accurate, faster response to image queries with ease.

Metadata describing the image defines the quality and quantitative aspect of the features of the image extracted [12]. As there is more research happening on reducing the semantic gap using proper annotation, there are research efforts required from this front. In paper [11] which discusses about automatic annotation of Liver CT data which would help in faster retrieval of Medical image. This has grown up to leveraging such technologies to annotating anatomical structures in the medical image.

DICOM provide meta information which can enable interoperability between different device vendors for easy interpretation of data which is residing in the images. It also helps universally accepted way of using it with different digital imaging devices such as MRI, CT, Ultrasonography, Nuclear medicine scanners etc. [1], As more trends towards lowering storage and computational cost and adoption big data and web emerges highly for Digital Imaging [2]. DICOM helps in implementations of CBMIR kind of systems to enable interoperability and scale for future needs.

Revised Manuscript Received on October 12, 2019.

HariPriya P*, Department of Computer Science, Bharathiar University, Tamil nadu, India, p.haripriya28@gmail.com

Porkodi R, Department of Computer Science, Bharathiar University, Tamil nadu, India, porkodi_r76@buc.edu.in

As DICOM images go beyond the visual representation by itself, it's very important to go deeper into the features aspect of the DICOM image. As every DICOM image having content and its features such as Color, Patterns, Texture, etc., extracting its features and processing them as semantic content is important for accurate retrieval and understanding the morphological structure of image as whole [12]. When organizations like Facebook[18] are using CNN extensively, its usage for DICOM and healthcare need not be substantiated enough.

Convolutional Neural networks could be based on the Neocognitron model which was proposed for simulating Human Visual System ability to recognize stimulus patterns [20]. CNN is class of deep networks and it is largely adopted in visual image related problems. CNN has some technical features of Visual cortex which outlined by YannLeCun in 1988 in this paper about back propagation [17].

III. RELATED WORKS

CBMIR has been implemented using various techniques. In this paper [5] Pilevar has discussed a method of extracting the regions of the medical image and decomposed them into effective regions. Based on similarity measurement between images using regional matches were able to attain 93% success rate with a 5000 images dataset.

Ramamurthy[6] discusses the implementation of CBMIR using hybrid methods through combining text features extraction and shape feature extraction at 4 different levels. The implementation was done with Matlab 7.2 with Image processing toolbox to achieve Precision and recall about 90% except that the query was done only against sample medical images from the toolbox.

Mohamed, O.[7] , a paper related CBIR by CNN and SVM has been discussed for fast retrieval. More than 6000 images has been used for training from the ImageNet and Caltech256 database [7]. The classifier has reached the accuracy of 95% with SVM which would enable to use less images for training with error rate up to 11.9%. Also identified that pre-processing of images helps to improve the accuracy. Though this paper discusses the adoption of CNN and SVM with better accuracy this is done on general database and not on DICOM image or draws any line of references to CBMIR.

Yu-An Chung & Wei-Hung Weng[8] conducted the experiments on the DR fundus image dataset with 26517 images after pre-processing the images at 224 x 224 pixel images. Deep SCNN has been used and performance evaluated with Mean Average Precision and Mean Reciprocal Rank with 0.6492 and 0.7737 respectively with some inconsistent trends. It also concluded to try and perform similar research on different medical datasets.

Ma, L., Liu, X.[9] proposed the Fused Context-Sensitive Similarity based on Content based Medical Image

retrieval focusing specifically on CT images especially for Lung related diseases[9]. This was done based on similarity between visual feature and classification information from the LISS dataset images. The Gaussian implementation of SVM gave accuracy up to 86.27%

Qayyum, A [10], developed the CBMIR model using DCNN technique for medical images. They are archived classification and retrieval accuracy is 99.77% and 69%. The implementation was done using Torch tool with 24 classes of Medical image dataset from cancer imaging archive. The number of images totaled to 7200 as 300 images for 24 classes and 80% of them were DICOM images.

IV. DEEP CONVOLUTIONAL NEURAL NETWORK (DCNN)

Deep learning technique adopted to achieve true Artificial Intelligence [16]. Though these can be computationally intensive advanced cloud based technologies and availability of GPUs has made this possible. DCNN has been used to classify the images and helps in performing the object identification with in image scene. Convolution by itself gives the combined integration of two functions as provided in the equation given below:

$$[x * y](t) = \int_0^t x(\tau)y(t - \tau) d\tau$$

$[x * y](t)$ represent the convolution of x and y. When it's considered over infinite image it represented as follows:

$$[x * y] = \int_{-\infty}^{\infty} x(\tau)y(t - \tau) d\tau$$

$$= \int_{-\infty}^{\infty} x(\tau)y(t - \tau) d\tau$$

The word “convolve” means to roll together. The convolutional neural networks consider images as 3D objects with Width, Height and Number of pixels along the dimension. These numbers of pixels would be in three layers for RGB respectively which are called as Channels. With respect to leveraging the Deep Convolutional Neural Network using Matlab it might need to load the image data from the source and explore the metadata from the image.

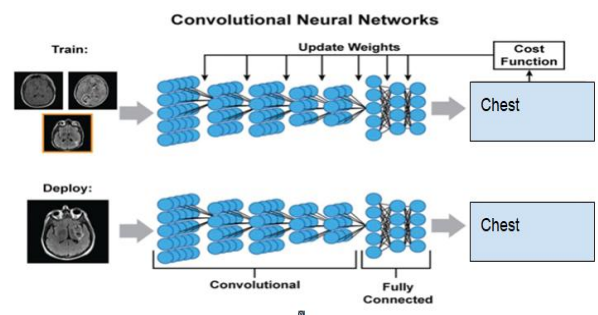


Fig. 1. DCNN Architecture



Figure 1 shows that the example reference image given above has been sourced based on “Deep Learning in Neuroradiology”[15] helps to understand this a simplified way. The DCNN models have three components apart from the Input and output layer namely convolutional, pooling and fully connected layers. The convolutional neural networks are generally similar to the Neural networks except that Neurons in the CNN are having the capability of learning the weights and biases. Each neuron in the CNN performs a dot product on the given input.

V. PROPOSED METHODOLOGY

In this proposed method, Deep CNN implementation to improve the retrieval efficiency of the medical images

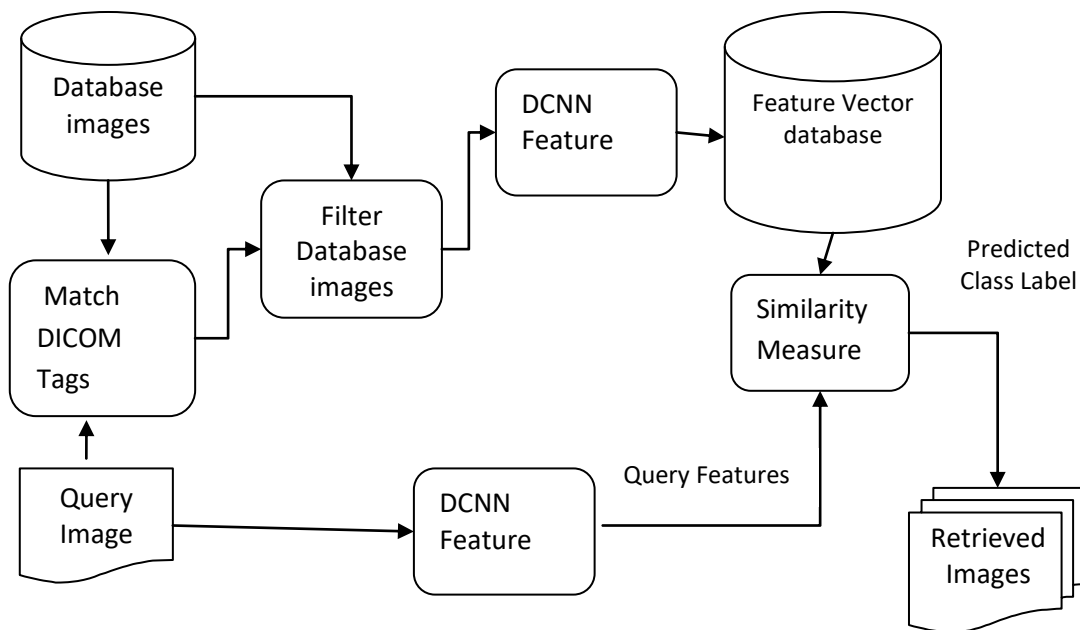


Fig. 2. Proposed Framework for CBMIR

Initially the all the images from the database image repository are extracted with the DICOM Meta tags. Then the features of the images are segregated and persisted in the feature vector database. These data are fed to the DCNN for developing the appropriate prediction of the class labels. Convolutional layer is applied to image and this layer performs convolutions with its neighboring pixels as small areas incrementally. The inputs query image is extracted with the Features which are to be performed for comparison using the existing DCNN prediction function which will retrieve the list of images which are closest match with higher accuracy. Finally similarity measures are calculated by Euclidian distance matrix function for image retrieval.

VI. DATA PRE-PROCESSING

When we need to build different layers through neural network models we need to consistent aspect ratio, so each class images are resized with 256 X 256 dimensions. To minimize the data processing the RGB related to intensities in the images, the images are converted to Grey scale. This helps to disregard lot of unnecessary information which needs to be processed. This could also help in processing time reduction. The metadata of the DICOM image extract using the Image

with multi modality. In Figure 2 shows the comprehensive illustration of the proposed methodology. Fundamental Deep CNN framework need to discover the kernel filter by produce concrete illustration of image. The images are categorized to class and it eventually frames into a multiclass classification issue. Normally, image classification consists of two important steps such as extraction of features and classification. Deep CNN absorbs with more convolutional features and classifier subsequently. Deep Learning acquires the low level, midlevel and abstract level features from images and it is circumstance for handcrafted features. For medical image retrieval, the classes are recognized based on query image and consequently the extracted features used in retrieval.

Processing Toolbox function dicominfo. This helps to extract Metadata features Such Modality, Body Part Examined, Patient Gender, and Patient Age using Dicom Matlab functions for all images. Subsequently labels are assigned to all the class of image based up on Meta data features. Figure 3 shows that the steps which are performed in pre-processing stage.

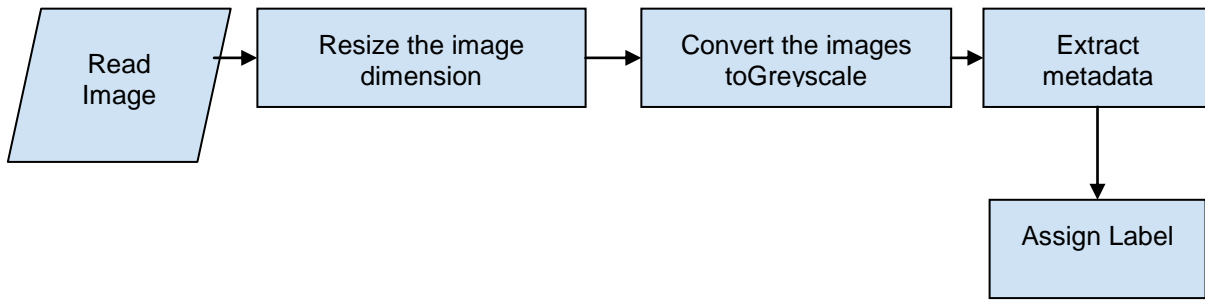


Fig. 3. Data Pre-processing Steps

VII. DCNN MODEL TRAINING

The training process and its output of classification is based on input image. The DCNN models contain the three important layers namely convolutional, pooling and fully connected layers.

Here, DCNN contains the 5 convolutional layer, 3 fully connected layers and max-pooling layer were used. This model the convolutional, Polling and Fully connected layers are represented in CL1, C12, C13, CL4 &CL5 , PL1, PL2 & PL3 and FCL1, FCL2 & FCL3 respectively. In CL1 consist of 96 filters of 11 x 11 filter dimension with stride value is 4..From the output of CL1 feed to a ReLU activation function after that output have forward toPL1 to reduced the dimension. Next, CL2 processed the output from PL1 with 256 filters of size is 5 x 5 with stride 1 and padding 2.The CL3 and CL4 contains 384 kernels with 3 x 3 filter dimension and stride and padding value is 1. In CL5 have 256 kernels with 3 x 3 filter dimension. PL2 is processed after CL2 convolve operation and PL 3 is used after CL5 with same dimension, stride and padding. The FCL1 and FCL2 contain 4096 neurons and FCL3 contain 1000 neurons. After construed the DCNN model was trained by SGD

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, triaing_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] ← calculaterehidden_delta1(delta1[l+1], W1[l], y1[l])
          endif
          triaing_W1[l] ←
np.dot(delta1[l+1][:,np.newaxis],np.transpose(h[l][:,np.newaxis]))
  
```

algorithm. The first step in an in the model training process is to load the data into your desired tool for analysis. Adopted the SGD (Stochastic Gradient Descent with Back propagation)[19] for training the models. The term “Stochastic” refers to the randomness involved in the training process. This would be used to update the weights to achieve better accuracy. This is being used to with goal to minimize error function.

The image classification task primarily consists of two steps one is using the score function and another one is the loss function. The score function is to map the raw image pixels and meta data to the class scores (linear), where in the loss function helps to ascertain the quality of a particular set of parameters based on the truth table. When network layers are increase then SGD helps Back propagation by updating the weights during each iteration. Gradient Descent method typically takes a random data from the parameter space and reduces it iteratively until the error E reaches a local minimum. If there are 50 training images the SGD adjusts the weight 50 times as part of the iteration [13] in the training pool During every iteration it determines the deepest descent and move in that direction.



```

        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

```

Similarity Measure

Euclidean distance is the most common metric for measuring the distance between two vectors in CBMIR approaches. Similarity images retrieved through obtained similarity feature for query image from tranined database image. The distance between vector a and b defined as:

$$d(\mathbf{q}, \mathbf{r}) = \sqrt{\sum_{i=1}^p (q_i - r_i)^2}$$

where q_i, r_i denote the image features of query and database image.

VIII. EXPERIMENTAL RESULTS

The dataset used for the experimentation of the CBMIR is obtained from TCIA (The Cancer Image Archive) open access medical portal [14]. The dataset primarily uses DICOM format of images in the archive. Data is available based on the collections based on the 3 Modalities - CT, MRI, and Ultrasound. 22 classes of images are used with 3 types of modality. In this paper, different set of 22 classes used for the purpose of research such as Brain, Breast, Liver, Stomach, Leg, Kidney, chest, Heart, Cervix, Uterus, Thyroid, Rectum, Prostate, Phantom, Pancreas, Ovary, Lung, Head neck, Head, Colon, Chest, Bladder. Totally 6600 images are used spanning 300 images for each class. These DICOM files were having the extension “.dcm”. In this research used training and test data set is 70% and 30% respectively. This image retrieved has been performed based on the Metadata and the DCNN features of the image itself. In figure 4 (a) shows the query input image as CT modality and Uterus of female of age 37 and figure4 (b) shows that the retrieved images for given input image. In figure 5 shows the performance matrix for image retrieval based on precision, recall, accuracy and F measure.

Performance Metrics

The precision outlines the accurately retrieve images from the relevant images based on the given query. Precision is the capability of the classifier to not label as positive a data sample that is negative. It is considers all the given inputs and provides how accurate the model based on the actual positives. This measure is very beneficial when the cost of false positive is high. This can be outlined as P (retrieved DICOM Images/relevant DICOM Images) based on the given query. Recall is the fraction of documents which are significant images from retrieved DICOM images which can be represented as R (retrieved DICOM Images/relevant DICOM Images). Recall is the capability of the classifier to find all positive samples from the given query. Recall is also termed as sensitivity, which is a proportion of true positives among all truly positive samples.

In summary, to determine the closest match for the given query image from the non-relevant images, precision is a fraction of correctly predicted queried DICOM image to all predicted DICOM images. Recall is a fraction of predicted DICOM image to the total number of true DICOM images. F measure is an important in medical image retrieval and calculated by using precision and recall values. The proposed framework is evaluated by following performance matrices for CBMIR.

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

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

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

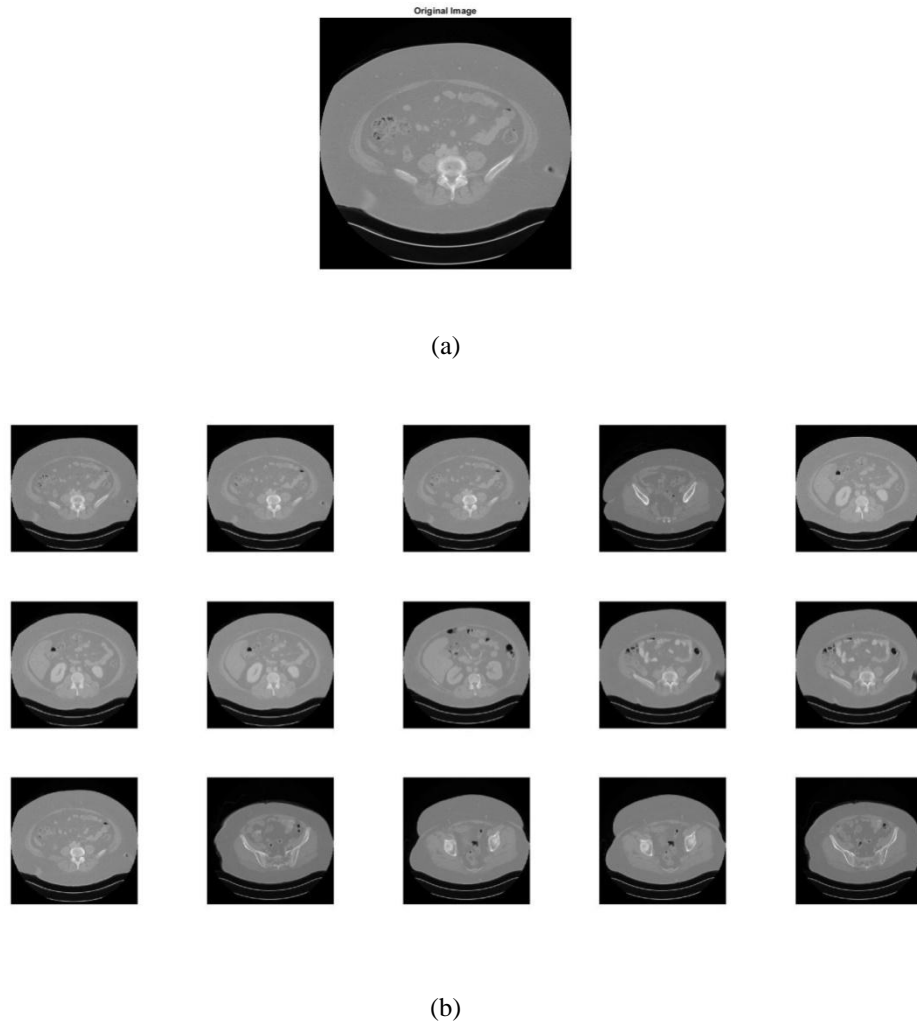


Fig. 4. Retrieval result (a) Query Image, (b) Retrieved Image

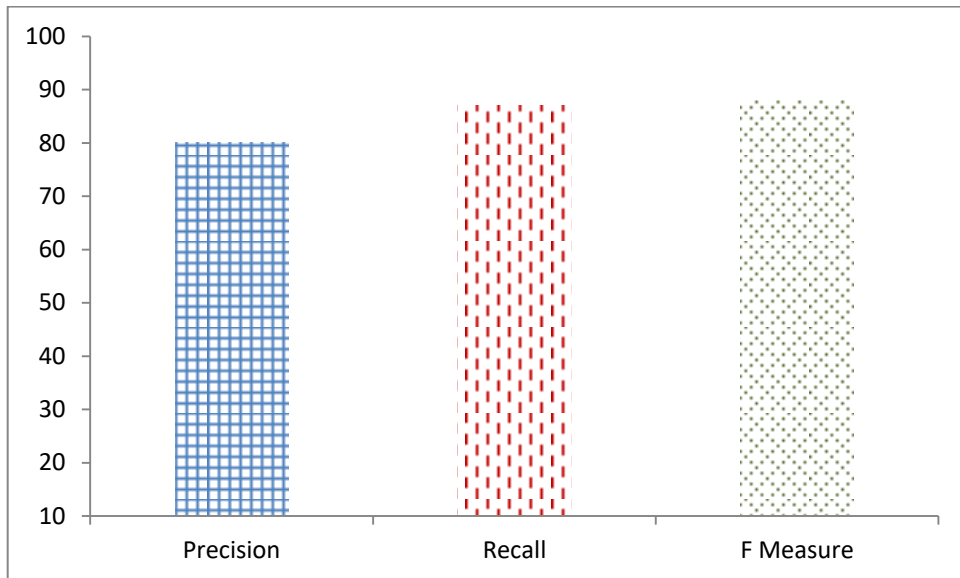


Fig. 5. Performance Metrics

Based on the given query through the DCNN with SGD approach we see Recall being significant. In general, F-Score would be less when classes which have fewer samples and high for classes with large samples. With more than 22 classes and 6600 images we see that F-Measure at 88. The result has fared better when compared with the previous outcomes as provided in the table 1. In searches that are expected to be semantic, we see an improvement in precision better than previous result where precision stands at 0.80.

| Method | Images | Modalities | Classes | Precision |
|---|--------|---|---------|-----------|
| DCNN with SGD (Proposed) | 6600 | CT, MRI, <u>UltraSound</u> (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

IX. CONCLUSION

In this paper, the proposed deep learning for CBMIR by using deep CNN. The medical image database were used the low level and high level semantic feature along with DCNN features are extracted. It is developed a deep CNN based classification system which can help in classification of DICOM using pre filtering by meta data. The proposed work decreases the semantic gap through learning discriminative features directly from the images and also reduced search space. In this system gives the better result compared to existing result and obtained the precision rate is 0.80 for multimodal image data for given query image. Futher, research approach can be improve the performance accuracy for big data and also acclimate the network for 3D volumetric images.

REFERENCES

- W. D. Bidgood, S. C. Horii, F. .Prior and D. E. V. Syckle, "Understanding and Using DICOM, the Data Interchange Standard for Biomedical Imaging," Journal of the American Medical Informatics Association, vol. 4, no. 3, pp. 199-212, 1997.
- M. & E. K. & O. M. & A. B. Ouhda, Content-Based Image Retrieval Using Convolutional Neural Networks, 2019.
- S. . Khan and S. . Khan, "An Efficient Content based Image Retrieval: CBIR," International Journal of Computer Applications, vol. 152, no. 6, pp. 33-37, 2016.
- A. . Kumar, J. . Kim, W. .Cai, M. J. Fulham, M. J. Fulham, D. .Feng and D. .Feng, "Content-Based Medical Image Retrieval: A Survey of Applications to Multidimensional and Multimodality Data," Journal of Digital Imaging, vol. 26, no. 6, pp. 1025-1039, 2013.
- Pilevar, Abdol Hamid. "CBMIR: Content-based Image Retrieval Algorithm for Medical Image Databases" Journal of medical signals and sensors vol. 1,1 (2011): 12-8.
- Ramamurthy, Byrav and K. R. Chandran. "CBMIR: Content Based Medical Image Retrieval Using Multilevel Hybrid Approach." (2015).
- Mohamed, O., Khalid, E. A., Mohammed, O., &Brahim, A. (2019). Content-Based Image Retrieval Using Convolutional Neural Networks. In J. Mizera-Pietraszko, P. Pichappan, & L.

- Mohamed (Eds.), Lecture Notes in Real-Time Intelligent Systems(pp. 463–476). Cham: Springer International Publishing.
- Chung, Y.-A., &Weng, W.-H. (2017). Learning Deep Representations of Medical Images using Siamese CNNs with Application to Content-Based Image Retrieval.CoRR, abs/1711.08490. Retrieved from <http://arxiv.org/abs/1711.08490>
- Ma, L., Liu, X., Gao, Y., Zhao, Y., Zhao, X., & Zhou, C. (2017). A new method of content based medical image retrieval and its applications to CT imaging sign retrieval. Journal of Biomedical Informatics, 66, 148–158. <https://doi.org/https://doi.org/10.1016/j.jbi.2017.01.002>
- Qayyum, A., Anwar, S. M., Awais, M., &Majid, M. (2017). Medical image retrieval using deep convolutional neural network.Neurocomputing. <https://doi.org/10.1016/j.neucom.2017.05.025>
- Kumar, A., Dyer, S., Kim, J., Li, C., Leong, P. H. W., Fulham, M., &Feng, D. (2016). Adapting content-based image retrieval techniques for the semantic annotation of medical images. Computerized Medical Imaging and Graphics, 49, 37–45. <https://doi.org/https://doi.org/10.1016/j.compmedimag.2016.01.001>
- Artificial Intelligence for Maximizing Content Based Image Retrieval, Zongmin Ma Northeastern University, China, Information Science Reference (an imprint of IGI Global) 701 E. Chocolate Avenue, Suite 200, Hershey PA 17033, 2009
- Kim, P. (2017). MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence(1st ed.). Berkeley, CA, USA: Apress.
- 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>
- G., Gong, E., Wintermark, M., Rubin, D., &Langlotz, C. P. (2018). Deep Learning in Neuroradiology.American Journal of Neuroradiology.<https://doi.org/10.3174/ajnr.A5543>
- Mattsson, N. (2016). Classification Performance of Convolutional Neural Networks.
- LeCun, Y. (1992). A theoretical framework for Back-Propagation. In P. Mehra& B. Wah (Eds.), *Artificial Neural Networks: concepts and theory*. Los Alamitos, CA: IEEE Computer Society Press.
- Güler, R. A., Neverova, N., & Kokkinos, I. (2018). DensePose: Dense Human Pose Estimation In The Wild. *CoRR*,

abs/1802.00434. Retrieved from <http://arxiv.org/abs/1802.00434>

19. Amari, S. (1993). Backpropagation and stochastic gradient descent method. *Neurocomputing*, 5, 185–196. [https://doi.org/10.1016/0925-2312\(93\)90006-O](https://doi.org/10.1016/0925-2312(93)90006-O)
20. Kuniyiko Fukushima (1980), Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, Volume 36, Issue 4, pp 193–202
21. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 770-778. doi: 10.1109/CVPR.2016.90
22. Pierre Sermanet, David Eigen, Xiang Zhang, Michael Mathieu, Rob Fergus, YannLeCun, "OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks", 2014 doi:<https://arxiv.org/abs/1312.6229>.

AUTHORS PROFILE



P. Haripriya, received MCA degree in Anna university and pursuing PhD in Bharathiar University. Her research interests include CBMIR, Medical Image Mining and Deep Learning.



Dr. R. Porkodi, received MCA degree and pursued Ph.D in Bharathiar University. She received UGC grant for her research study. She is the member of many academic bodies. She is the life member in computer society in India and member in IAENG and IACSIT. She acted as a committee member/ resource person/coordinator for various research conferences/events/Faculty development programmes. She published many articles in various reputed journals. Her research interests include Data mining, NLP, Image mining, CBMIR, Hyperspectral remote sensing and Bioinformatics