Medical Image Retrieval Using Convolution Neural Networks

N Kalaivani¹, D Devi², Dr. S Sophia³, N Manimaran⁴, P Kowsalya⁵
¹,²Assistant Professor, ³Professor, Sri Krishna College of Engineering and Technology, Coimbatore.
⁴Assistant Professor, SACS MAVMM Engineering College, Madurai.
⁵Assistant Professor, Sri Venkateswara College of Engineering and Technology, Chittoor.
Corresponding author mail Id: kalaivani@skcet.ac.in

Abstract: In this paper, medical image retrieval is done in an effective manner by using convolution neural network (CNN). This proposed system locates reference tags and classes the DICOM images using image processing techniques and retrieval of images are done by using GUI panel. First, the deep learning technique is used to extract powerful features of the image for tag description. Conversely, this technique performs tag matching directly by passing suitable parameters which recognizes the classes as queries. The comparison features are able to capture the general form of the input image and its class based on image tags. Here, we have used a collection of 22 classes of database to demonstrate the efficiency of our system. The experimental results show the classification of images by deep learning algorithm which is used to gain the rate of retrieval accuracy by using MATLAB.

Keywords: Medical image retrieval, DICOM images, Deep Learning.

1. Introduction

The development of computerized medical imaging uses Picture Archiving and Communication Systems (PACS) such as CT, MRI and X-ray has led to the need for efficient classification and visual data search in hospitals worldwide. The use of standards such as digital imaging and Medical Communication (DICOM) enables an easy and reliable exchange of information in the medical environment. DICOM Structured Reporting (SR) is a standardized report encoding concept. This allows further data extraction and automatic interpretation. Due to the advantages that DICOM posses, the integration of the healthcare company (IHE) uses DICOM SR as a number of integration profiles. DICOM SR has become the standard format for the exchange of clinical CAD results. An effective method of deep learning along with Deep Convolution network and image processing techniques are used to handle the DICOM images.

CNN has many convolution layers and fully connected layers. For the activation function, the standard way to activate the output of a neuron is by hyperbolic tangent function or Logistic function. These nonlinear activation functions may lead to saturation and cause vanishing gradient problem. These saturating nonlinear functions show a much slower convergence than non-saturating nonlinear functions, such as max(0, x) for training optimization with gradient descent. If the input x is greater than zero, the function returns x itself, otherwise the function returns 0. The nonlinear function is referred to as Rectified Linear Unit (ReLUs). Deep CNNs with ReLUs is trained several times faster than their equivalents with hyperbolic tangent units, which is crucial for training large models in large data sets. In view of the output of convolution layers, we add it to a range of fully connected layers.
The output of the last fully connected layer is regarded as a softmax classifier input to further transform the feature vector to probabilities. Thus, deep learning consists of three layers of input, hidden and output layers. In this system, input layer reads the file and detects the directory. The convolutionary layer (hidden layer) applies a mask or a filter to the input and the dot products of the values in the filter and the input is calculated. The convolution layer contains many filters and it produces different activation maps and proceeds to the respective fields such as rigid and non-rigid transforms of input images. Rigid transforms are pixel-oriented processes such as gray scaling, edge detection, image enhancement, image quantization and gray to RGB conversion. All layers can be called as hidden layer except the input and output layers. These hidden layers are used to store the different types of data for pooling layer purposes. Although hidden layers can be used with the high number of neurons, each pixel is relatively variable. pooling layer will down sample along the spatial dimensions. The depth dimension remains unchanged. The pooling layer accepts and produces the size volume of the image. The output layer connects all layers together to drive the output such as the plot and the numerical data driver.

2. Literature survey
In the recent past, the use of Neural Networks for data processing has been very expensive. The Deep Convolution Neural Network improves the geometric invariance of compact CNN descriptors [1]. A content-based image recovery system for gray images, RGB color space, YCbCr color space, images are grouped into various clusters and measured at the precision-recall crossover point [2,5]. In the field of retrieval, 3D matching is done by a 3D shape feature method of learning to extract high-level shape features which are insensitive to geometric shape deformations and uses a deep autoencoder to deform the invariable shape features as the hidden layer in this network [3]. A bidirectional model of learning representation is used to represent cross-modal text image retrieval. Since deep learning is used in Image and natural language processing, it is effective to use the deep neural networks to extract details from both text and image [4] which are trained with user generated contents data directly without any labeling effort. Although it deals with a common representation over multimodal inputs, training samples are used in the absence of certain modalities. A structured classification model is used for both visual and click functions in distance learning (DML). In particular, images and their results are first collected in order to form the training set. These images collect multimodal features which includes click and visual features. A method of sketch-based image recovery through image-assisted inter domain learning is used. DNN model is used to study the differential characteristics by using the generalized boundary. DCN with hash codes is used to learn a medical image. Therefore, a hash layer is added to the network which is used to represent the image information as binary hash codes. A deep network images are compressed to a bit stream and then trained to extract image features as a binary vector which are combined with the network for the CBIR task. Unsupervised-learning-based method using HNN (Hopfield Neural Network) is modeled to increase the retrieval efficiency of human visual memory and reduce the semantic gap.

3. Proposed system
In the proposed system, deep convolution neural network (DCNN) consists of several convolution layers, pooling layers and fully connected network. Here, 8 layers are used for retrieving the DCM files. The output of convolution layer is feature maps produced by convolution of the filter mask with the input images. Pooling layer is used to decimate the samples in the activation maps and also reduces the count of neurons and also gives translational invariance. CNN contains a feature extractor and a classifier.
The deep learning structure also learns about the retrieval function with 8 layers of neural network and image processing techniques to increase the speed and efficiency of image analysis. The dataset consists of 22 classes of different parts of the body. Each part of the body consists of 300 DICOM images which are trained to collect from the DCM files.

4 Data set and simulation results

4.1 Data set

Database is trained using 22 classes of body parts of DCM tag files. The 22 classes consists of brain, breast, bladder, cervix, chest, colon, esophagus, head, head neck, heart, kidney, leg, liver, lung, ovary, pancreas, phantom, prostate, rectum, stomach, thyroid, and uterus. Each class contains 300 DICOM images as queries. The tag file is stored in the confusion matrix form. The preprocessing process is done for the model, which contains two sets of functions.

i. Training data set for model running.

ii. Testing data set for model evaluation.

4.2 Experimental results

An image is given to the system and the DICOM image matches the tag by passing the parameter for recognizing classes as queries. Comparison features can capture the general form of the query and its classes based on image’s reference tags. Then, the similar images for the given feature are extracted and the image is retrieved in the MATLAB GUI panel.
Fig. 3. Original image of the DICOM images

Fig. 4. Enhanced image

Fig. 5. Enhanced image histogram
Fig. 6. RGB segmented image

Fig. 7. Image analysis data

Fig. 8. Edge detection
5. Conclusion
In this proposed model, the deep convolution neural network (CNN) is trained and tested using the DICOM images of 22 classes of body parts. The given Query image is predicted by the trained network and then relevant images are retrieved from the database. In the traditional neural networks, feature maps of the images are given as the input. If some features are missed in the input, then the image retrieval is not efficient.
But the proposed method improves the performance by extracting the features from the images by the convolution neural network itself. The depth of feature extraction is decided by the convolution neural network. So all the required features are extracted and given to the next layer which will improve the image retrieval rate. Increasing the data set and performing image retrieval for 3D applications are the future extension of this work.

References

[1] Aaron Chadha and Yiannis Andreopoulos” Voronoi based Compact Image Descriptors: Efficient Region-of-Interest Retrieval” IEEE Transaction on multimedia volume 2017
[2] Pukale D D S G Bhirud and V D Katkar “Content-based Image Retrieval using Deep Convolution Neural Network” IEEE International Conference on Computer and Communications 2017
[3] JinXie Guoxian Dai Fan Zhu Edward KWong and Yi Fang” Deep Shape: Deep-Learned Shape Descriptor for 3D Shape Retrieval” IEEE Transaction on Pattern Analysis and Machine Intelligence 2016
[4] Yonghao He Shiming Xiang Cuicui Kang Jian Wang and Chunhong Pan” Cross-Modal Retrieval via Deep and Bidirectional Representation Learning” IEEE Transactions on Multimedia NO 2017
[5] Balamurali R and Chandrasekar A 2019 Multiple parameter algorithm approach for adult image identification Cluster Computing 22(5) pp11909-11917