CLASSIFICATION OF DEFECTED SPINE AND SEGMENTATION USING DEEP LEARNING

[1]**Mrs.G.ValarmathiM.E,(Ph.D),**

Department of ECE, Sri Sairam Institue of Technology, Chennai-44

[2]**Prof.Dr.B.Ashraf Ahmed MD[RD],**

HoD Department of Radiology,
Melmaruvathur Adhiparasakthi Institute of Medical Sciences & Research

[3]**S.Padmapriya, M.ShafanaAasmi, M.N.Varshini**

Department of ECE, Sri Sairam Institute of Technology, chennai-44

Abstract

This paper presents an efficient method to delineate the degenerated portion of the spinal cord from magnetic resonance images (MRI) of the patients. One of the major issue in human body is back pain. To diagnose the problems in lumbar region of spine we use deep learning neural networks using Convolutional Neural Networks(CNN) algorithm which can take in an input image, assign importance to various aspects/objects in the image and be able to differentiate one from the other .By using this CNN algorithm we get a meticulous result.. Spine is the principal transmission pathway for neural signals between the brain and the rest of the body. The primary purpose of this paper is to compare various methods used for segmentation of spine. Matlab has been used to perform spine segmentation and classification.Our approach produces better segmentation results than other existing methods.

NTRODUCTION

The recognition and segmentation of anatomical structures play critical roles in the quantitative interpretation of CT images, and, conventionally,the image processing is accomplished through human interpretation and manual annotations by expert radiologists. However, human interpretation is often qualitative and subjective, with relatively high intra- and inter-observer variability. Manual annotations also have limited reproducibility and are very time-consuming. Therefore, automatic image segmentation would offer improved efficiency and reliability, as wellas reducing the burden on radiologists[1].The human spine typically consists of 24 vertebrae (7 cervical:C1-C7,12thoracic:T1-T12,5lumbar:L1-L5)with in-between situated intervertebral disks, aligned in a double-S shaped curve[2]. We therefore use the disks as high level features (parts) for localizing the spine column and individual vertebrae. There are several pathologies that may significantly affect the appearance of the spine such as fracture, neoplasm, deformity (e.g. scoliosis) and degeneration. Also the number of vertebrae may differ from 24, e.g. by lumbalization of the cranial sacral segment into an L6.

Fully automatic image segmentation, which transfers the physical image signal to a useful abstraction, is a crucial prerequisite for computer-based image analysis of 3D CT cases[1].

Recently, machine learning-based methods have gained more and more interest in the medical image analysis community. Most of these methods are based on ensemble learning principles that can aggregate predictions of multiple classifiers and demonstrate superior performance in various challenging medical image analysis problems [3]. Zhan et al. [4] presented a hierarchical strategy and local articulated model to detect vertebrae and discs from 3D MR images.

More recently, with the advance of deep learning techniques [5,6,7], many researchers have proposed deep learning based methods for automatic localization and segmentation of vertebrae from CT images. For example, Chen et al. [8] proposed a method for automatically locating and identifying vertebrae in 3D CT volumes by exploiting high level feature representation with deep convolutional neural networks(CNN).

EXISTING WORK:

Many researchers have proposed methods for the diagnosis of certain vertebral column abnormalities[10]. Bounds et al.[11] utilized a neural network for diagnosis of back pain and sciatica. Sciatica might be caused by lumbar discherniation as well as many other reasons. They have three groups of doctors to perform diagnosis as their validation mechanism. They achieved better accuracy than the doctors in the diagnosis. However, the lack of data prohibited themfrom full validation of their system. The spine provides a natural patient-specific coordinate system, where individual vertebrae serve as anatomical landmarks[9]. These can be used, for instance, for semantically guided inspection tools, linking of radiological reports with corresponding image regions, or for robust initialization of image registration.

Two learning paradigms in machine learning, namely supervised and unsupervised learning, have always been a popular subject of research, comparison and analysis by researchers.In supervised learning, the algorithm is trained to map the input variables to the output variables using pairs of known input and output values called training data set. The resulting algorithm, which can manifest as a mapping function, a decision tree or a neural network, can then be tested for performance using another set of known input and output values, called test data set. Both training and test data sets consist of Ground Truth Data that are developed by either manually assigning labels to the input data, or collected by taking measurements from real world experiments.

The process of how the ground truth data is obtained depends on the task of which the machine learning is set to do. Furthermore, due to the nature by which they are obtained, ground truth data can have some degree of inaccuracy[12].In[15] the training data set was collected from 22 spinal MRI datasets. Each MRI dataset was acquired from a 1.5 T Siemens MRI system with spin-echo scanning sequence and 2-Dsagittal-view acquisition.

In these datasets, the repetition time (TR) parameter values are between 4000 ms and 4510 ms, and the echo time (TE) ranges from 67 ms to 104 ms.We manually labeled 1398 vertebra regions from the sagittal images in these datasets and crop the vertebra regions with the minimal bounding rectangles to obtain the positive samples.The first set of negative training images are randomly clipped from the nonvertebral regions in the training spinal MRI datasets. Then, during the bootstrap learning, the current classifiers are applied to send the false positive patterns, which are collected to form the new negative data set in the next training stage. These positive and negative data sets are given to the feature extraction procedure to extract features for AdaBoost learning.

There also exist a learning-based[16], unified random forest regression and classification framework totackle the problems of fully automatic localization and segmentation of VBsfroma3DCT image ora3DT2-weighted TSEMR image. More specifically, in the first step, the localization ofa3DVB in a given image is solved with random forest regression where we aggregate votes from a set of randomly sampled 3D image patches to get a probability map. The resultant probability map is then further regularized by HMM to eliminate potential ambiguity caused by the neighboring VBs.The output from the first step allows us to define a region of interest (ROI) for the segmentation step, where we use random forest classification to estimate the likelihood of a voxel in the ROI being foreground or background.

PROPOSED SYSTEM:

Our neural network shares a similar architecture to other patch-based deep neural networks[13]. Neural networks are complex models, which try tomimic the way the human brain develops classification rules. A neural net consists of many different layers of neurons, with each layer receiving inputs from previous layers, and passing outputs to further layers.

It includes an Image Input Layer, a cascade of Convolution Layers that will produce multiscale classification features and a Fully Connected Layer for classification of these features.The goal of using deep learning is to perform end to end learning. CNNs capture better representation of data and hence we don’t need to do feature engineering.

The main three parts of layer are input layer,hidden layer,output layer. Convolutional Neural Networks(CNN) is a type of neural network.Pretrained alex-net is an example of a CNN. CNNs eliminate the need for manual feature extraction—the features are learned directly by the CNN.CNNs produce state-of-the-art recognition results.CNNs can be retrained for new recognition tasks, enabling you to build on pre-existing networks.

CNN architecture is designed for application where the input has an inherent two-dimensional structure,like an image.The layers used are

- input layer
- convolutional layer
- pooling layer
- ReLu layer

- fully connected layer
- softmax layer
- output layer

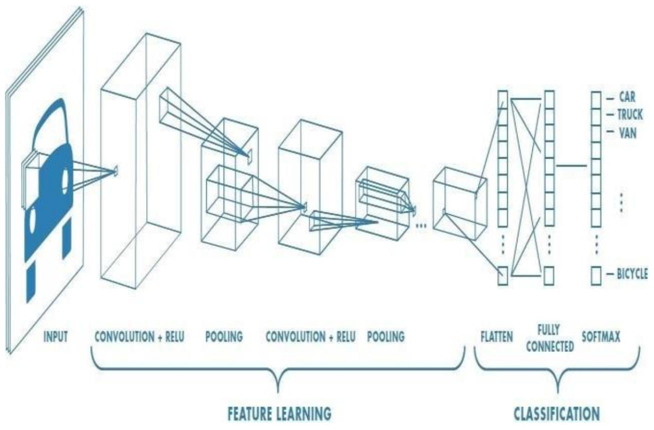

Fig.1 Layers of the neural network

The layers themselves comeout with many parameters known as weights. The weights determine how the layers behave when the data is passed through them.The values of weights are determined by training the network on known data. Transfer Learning take layers from a network trained on a large data set and fine-tune on a new data set.Components needed for transfering learning are network layer, training layer and algorithm options.The ReLufunction is used as activation function.The fixed configurations are number of hidden layers, number of nodes per hidden layer, activation function.

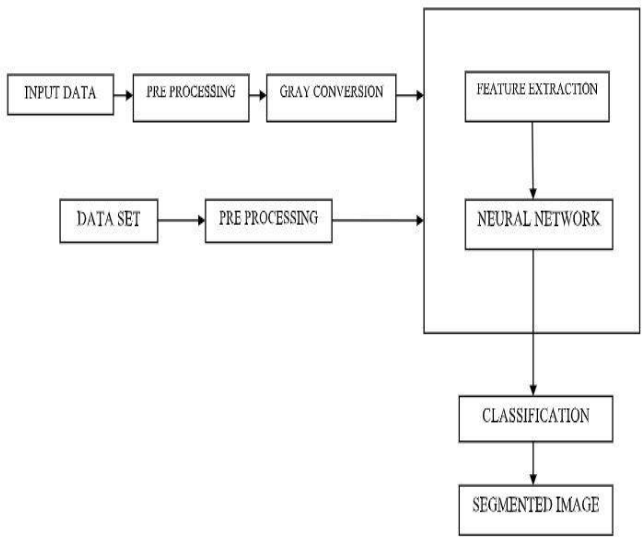

Fig.2 Block diagram

The output of the Image Input Layer is fed to a cascade of Convolution Layers. In our network, each convolution layer has a fixed number of equal-sized kernels. The kernel size chosen is relatively small since the input of the network is already a small subset of the entire image. Each kernel corresponds to each classification feature trained within the layer. Each of the output of these layers is calculated as the dot product of the input and the kernel. As with other deep neural network architectures, we opt to use Rectified Linear Unit (ReLU) activation function [14]. This decision is based on the function's advantages over others, such as Sigmoid function, that include faster training speed due to a reduced likelihood of vanishing gradient. ReLU activation function also has been proven to train networks with sparser, and hence considered better, weightsrepresentation.

Rectified linear unit (ReLU) allows for faster and more effective training by mapping negative values to zero and maintaining positive values. This is sometimes referred to as activation, because only the activated features are carried forward into the next layer.

CNN compares the images piece by piece.Then it looks for are called features. By finding rough feature matches,in roughly the same position in two images,CNN gets a lot better at seeing similarity than whole-image matching schemes.The convolutional layer line up the feature and the image.Then multiply each image pixel by the corresponding feature pixel and add them up.Then divide by total number of pixels in the feature.

After each convolutional layer, there is a Rectified linear unit (ReLU) activation function. Activation function decides, whether a neuron should be activated or not by calculating weighted sum and further adding bias with it. The purpose of theactivation function is to introduce non-linearity into the output of a neuron. ReLu is less computationally expensive than tanh and sigmoid because it involves simpler mathematical operations.

At a time only a few neurons are activated making the network sparse making it efficient and easyfor computation. This layer is used to remove all the negative value(change to 0).This is done to avoid the values from summing up to zero.The advantage of using ReLu layer are Computationally efficient (allows the network to converge very quickly) and Non-linear(although it looks like a linear function, ReLU has a derivative function and allows for backpropagation).

The max pooling has a stride of two and halves the spatial dimensionality[13]. Maximum pooling, or max pooling, is a pooling operation that calculates the maximum, or largest, value in each patch of each feature map. The results are down sampled or pooled feature maps that highlight the most present feature in the patch, not the average presence of the feature in the case of average pooling.

Pooling layer is used to reduce the size of the array matrix. Fully connected layer puts filtered shrinked images into single list. This is the layer where actual classification happens. FullyConnected layers in a neural networks are those layers where all the inputs from one layer are connected to every activation unit of the next layer. In most popular machine learning models, the last few layers are fully connected layers which compiles the data extracted by previous layers to form the final output. It is the secondmost time consuming layer second to Convolution Layer. The input to the fully connected layer is the output from the final Pooling or Convolutional Layer, which is flattened and then fed into thefully connected layer.

Feature extraction is an easy and fast way to use the power of deep learning without investing time and effort into training a full network. Because it only requires a single pass over the training images, it is especially useful if you do not have a GPU. You extract learned image features using a pretrained network, and then use those features to train a classifier.
Pooling simplifies the output by performing nonlinear downsampling, reducing the numberparameters that the network needs to learn. Softmax layer is the final layer used in neural network. It must have same number of nodes as the output layer. Softmax is useful is because it converts the output of the last layer in your neural network into what is essentially a probability distribution.The advantages of using softmax layer are they can Able to handle multiple classes and Useful for output neurons. Backpropogation is algorithm used to train neural network.Batch normalization is used to increase the stability of a neural network.

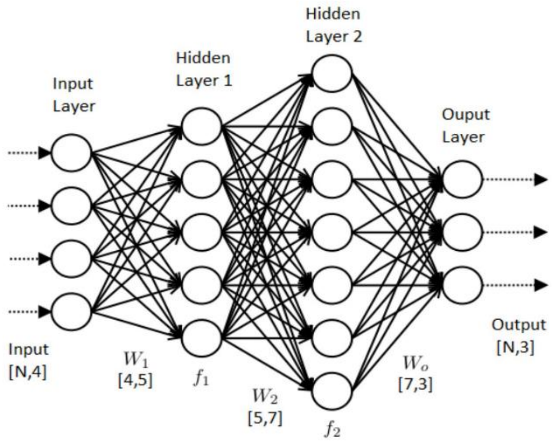

Fig.3 Deep learning neural network

We have used MatLab to write code for training and testing as this is a simple to use code of Convolution Neural Network -a deep learning tool.60% of dataset is used for training and 40% of the images where used for testing. We trained the network using normal and abnormal images of spine. This paper mainly focuses on the lumbar region of the spine.We useddeep learning as theyhave extremely high accuracy in image recognition applications.The advantages of using convolutional neural network are: they can capture/are able to learn relevant features from an image at different levels similar to a human being. So CNN was applied for better results.

Additionally, automatically predicting the patient overall size could replace the current scale search step and reduce testing times to only a few seconds. Further investigation will also be carriedout with respect to highly pathological cases of spine such as high-grade scoliosis.

RESULT:

1. The first result obtained by the method we developed is classification of herniated spine by the processed image input.Fig.4 represents the reference of the obtained result.

2. The other additional method is to calculate the distance between the discs. The physician may select the required boundary for calculating the distance in the processed image.

3. Fully segmented spine image is obtained from the processed input image.

4. The segmented image is precisely accurate for the physicians for further diagnosis.

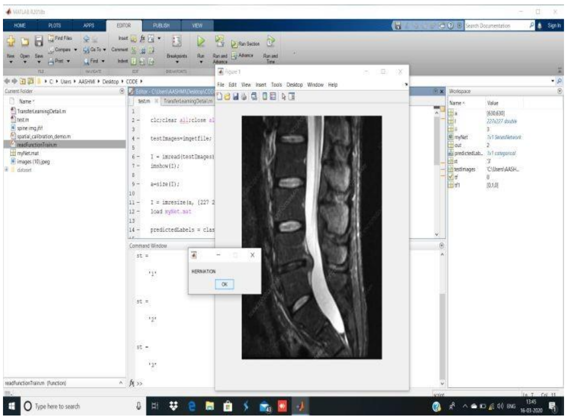

Fig 4: Classified result among the Testing data

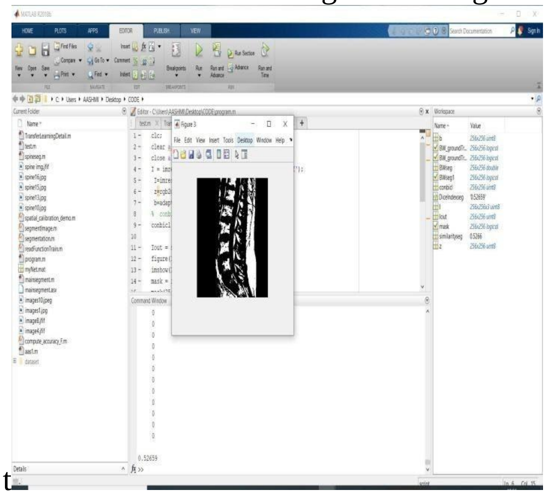

Fig.5 Segmented image

Fig.6 width between two discs

CONCLUSION:

We validated our model using 33 abnormal clinical MRI cases. Each case contains at least one herniated disc. We consider the clinical report of the radiologist as our gold standard for herniation condition of each disc. We perform a cross-validation experiment on the 33 cases by leaving 13 cases for testing each round. The overall herniation detection accuracy is around 90%. We also reported specificity of 91% and sensitivity of 94%. A semi-automatic detection approach for the localization and identification of vertebrae in generic MRI is implemented. The algorithm does not make any assumptions on the input images and can deal with highly cropped scans and partially visible spines. In the future, increasing the amount of training data, in particular, for the lumbar region would produce an increase in accuracy across the entire spine.

Automatically locating and segmenting human organs from medical images constitute the first primary step for an intelligent medical image analysis system. In this paper, we presented a semi-automatic vertebra detection and segmentation algorithm for spinal MR images.

REFERENCES:

1.X. Zhou, R. Takayama, S Wang, T Hara &H Fujita,"Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method", Medicalphysics,44(10),pp.5221-5233,2017.

2.S. Schmidt, J. Kappes , M. Bergtholdt, V. Pekar, S. Dries, D. Bystrov, and C. Schnoerr, "Spine detection and labeling using a parts-based graphical model," in Information Processing in Medical Imaging, 2007, vol. 4584, pp. 122–133.

3.R. Janssens, G. Zeng,& G. Zheng,"Fully Automatic Segmentation of Lumbar Vertebrae from CT Images Using Cascaded 3D Fully Convolutional Networks.",IEEE 15th International Symposium on Bio medical Imaging(ISBI2018),pp.893-897,2018.

4.Y.Zhan,D.Maneesh,M.Harder,and X.Zhou,"Robustmr spine detection using hierarchical learning and local articulated model.," in Proceedings of MICCAI 2012, 2012, pp. 141–148.

5.Jonathan Long, Evan Shelhamer, and Trevor Darrell, "Fully convolutional networks for semantic segmentation,"in Proceedings of the IEEE Conference on Computer VisionandPatternRecognition,2015,pp.3431–3440.

6.Olaf Ronneberger, Philipp Fischer, and ThomasBrox, "U-net: Convolutional networks for biomedicalimage segmentation," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015, pp. 234–241.

7.O. Cicek, A. Abdulkadir, S. Lienkamp, T. Brox, andO. Ronneberger, "3d u-net: Learning dense volumetric segmentation from sparse annotation," inMICCAI2016,vol.LNCS9901,pp.424–432.Springer, 2016.

8.H. Chen, C. Shen, J. Qin, and et al., "Automatic localization and identification of vertebrae in spine ct via ajoint learning model with deep neural network," in Proceedings of MICCAI 2015, 2015, pp. 515–522.

9.B. Glocker, J. Feulner, and et al., "Automatic localization and identification of vertebrae in arbitrary
field-ofviewct scans," in Proceedings of MICCAI 2012, 2012, pp. 590–598.

10.R. S. Alomari, J. J. Corso, V. Chaudhary, and G. Dhillon, "Automatic diagnosis of lumbar disc herniation with shape and appearance features from MRI," in SPIE Medical Imaging, 2010, p. 76241A--76241A.

11.Bounds, D. G., Lloyd, P., and et al, "A multilayer perceptron network for the diagnosis of low back pain," 2, 481–489 (July 1988).

12.F. Natalia et al., ''Development of ground truth data for automatic lumbar spine MRI image segmentation,'' in Proc. IEEE 20th Int. Conf. High Perform. Comput. Commun.; IEEE 16th Int. Conf. Smart City; IEEE 4th Int. Conf. Data Sci. Syst., Jun. 2018, pp.1449–1454.

13.A. S. Al-Kafri et al., "Segmentation of Lumbar Spine MRI Images for Stenosis Detection using Patch-based Pixel Classification Neural Network," in IEEE Congress on Evolutionary Computation, 2018,p. toappear.

14.V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in Proceedings of the 27th International Conference on International Conference on Machine Learning, 2010, pp. 807–814.

15.Huang, S.H., Chu, Y.H., Lai, S.H., Novak, C.L.:Learning-based Vertebra Detection and Iterative Normalized-Cut Segmentation for Spinal MRI. IEEE TMI 28(10), 1595–1605 (2009).

16.C. Chu, D.L. Belavy, G. Armbrecht, M. Bansmann, D.Felsenberg,andG.Zheng, "Fully automatic localization and segmentation of 3d vertebral bodies from ct/mr images via a learning-based method," PLOS ONE, vol. 10, no. 11, pp. e0143327,2015.