Indian Institute of Information Technology, Vadodara

Government Engineering College, Sector-28, Gandhinagar, Gujarat - 382028

B.Tech. Project Report

on

Study of Deep Learning Architecture for Hippocampus Segmentation

submitted by

Dilip Puri
(201351014)

under the supervision of

Dr. Ashish Phophalia[1]

Abstract

This report provides an overview of Deep Learning Architecture on Medical Imaging. Early diagnosis of any disease plays a major role in improving treatment possibilities and increases the survival rate of the patient [8]. The brain image segmentation is a crucial part of diagnosis so we can find the status of illness [5] like a brain tumor, Alzheimers disease (AD). Segmentation of Hippocampus is an important task in the treatment of AD, as it is one of the part first affected by the illness. A depletion in hippocampal volume can be used as a marker for AD treatment [5]. The aim is to classify each voxel in an MR Image as non-hippocampus and hippocampus. In this study, we are going to segment the Hippocampus from MR Image of a brain and produce a ground truth value for Hippocampus region in the image using the technique based on Convolutional Neural Network Auto-Encoder.

Keywords - MRI, Image Processing, Image Segmentation, Deep Learning, Neural Network.

1 Introduction

Machine learning is applied in the various field of the medical imaging such as computer-aided diagnosis, image segmentation, image registration, image fusion, image-guided therapy, image annotation, and image database retrieval [15]. Machine learning methods can be applied and can help in early detection of AD [19].

Humans have two hippocampi, shaped like seahorses, as shown in Figure 1. The hippocampus is a component of human brains responsible for committing short-term episodic and declarative memory into long-term memory, as well as navigation [17]. Segmentation of Hippocampus is an important task in the diagnosis of AD, as it is one of the part first affected by the disease. A depletion in hippocampal volume can be used as a marker for AD diagnosis [5].

Nowadays, a large number of MR images generated in clinical routine and manual segmentation of hippocampal (hippocampus, amygdala, etc.) is tedious, time-consuming, susceptible to human errors and even non-reproducible and expensive. The construction of accurate, robust, and reliable segmentation techniques for the automatic extraction of anatomical structures is becoming a major challenge in quantitative MR analysis [6].

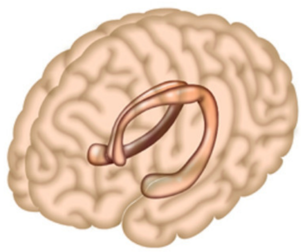

Figure 1: Left-Right Hippocampus

In this study we present the results of applying deep architecture to the hippocampus segmentation problem, and comparing their classification performances.

2 Literature Review

In recent years, comprehensive studies have been conducted to develop various image segmentation methods [8]. Many deep learning methods are proposed for the efficient processing and objective assessment for large MRI image data [8]. Since 1990, artificial neural networks have come to be used as a different approach for image segmentation [1]. The mostly used being Kohonen and Hopfield ANNs [1]. For particular hippocampus segmentation, many deep convolutional network based methods proposed, among these methods, SegNet[†] [3] and U-Net[‡] [12] deep architecture are preformed state-of-the-art performance. Originally, SegNet is an encoder-decoder deep architecture unsupervised learning model [3]. U-Net is U-shaped deep architecture based on the fully connected convolutional network. U-net architecture won the ISBI cell tracking challenge 2015 [12].

[1] ashish_p@iiitvadodara.ac.in

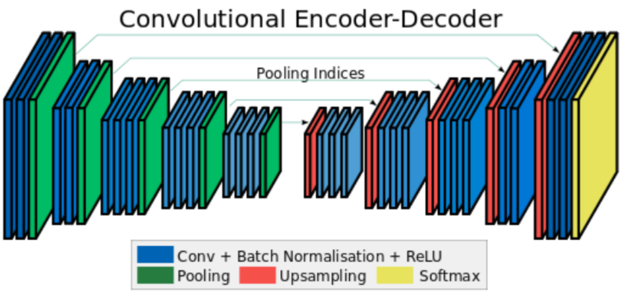

Figure 2: SegNet† architecture

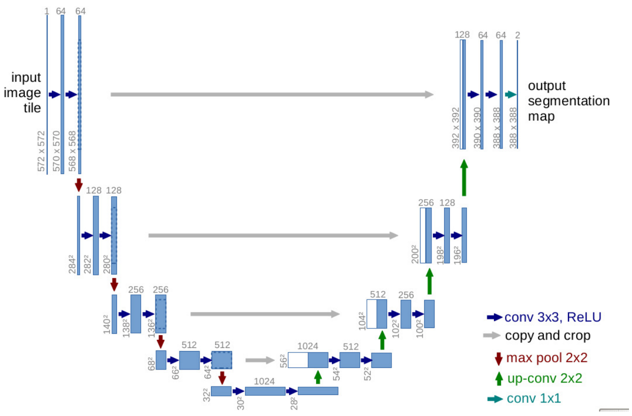

Figure 3: U-net‡ architecture

Various kind of deep architecture are proposed Autoendcoder, ImageNet [9], VGG-net [16], U-net [12], SegNet [3] etc. In many of these applications, methods based on deep learning has go beyond the previous state-of-the-art performance.

3 Deep Learning

Deep learning methods are a set of algorithms in machine learning, which learn different levels of presentation and characterization that help to understand data [4]. Higher-level abstractions are determined from lower-level ones, so more complex functions can be learned. In particular, if a deep architecture can represent a very complex function, the same function could require a remarkably broad architecture if the depth of this architecture is made more shallow [4] [2].

3.1 Artificial Neural Network

An Artificial Neural Network(ANN) is an information-processing paradigm that is motivated by the way biological nervous systems, such as the brain, process information. It is composed of a vast number of highly interconnected processing elements (neurons) working simultaneously to solve specific problems. ANN is like human brain it will learn by examples. In a typical neural network, nodes are arranged in layers, with the first layer being the input layer, and the last layer being the output layer. The input nodes are special arranged in the fashion that their outputs are simply the value of the corresponding features in the input vector.

3.2 Convolution Neural Network(CNN)

The CNN pioneered by Kunihiko Fukushima, Geoff Hinton, and Yann LeCun back in the 1970s and '80s [11]. Just like brain consists of billions of highly connected neurons a basic operating unit in Neural Network is a neuron like a node. CNNs are especially powerful in object recognition and localization in natural images [11]. Medical image analysis groups across the world are quickly entering the field and applying CNNs and other deep learning methodologies to a wide class of applications [7]. CNNs have applied to medical image processing as long ago as 1996 in work of Sahiner [14]. It takes input from other nodes and sends output to others. Moreover, these hundreds of thousands or even millions of nodes are organized in hierarchical layers, also similar to the human brain. Typical CNN model consists of bunch of layers(convolution, sampling, fully-connected, etc.).

4 Methods

For this study, we have 3-D MR image as input and output will be corresponding ground truth image for the hippocampus region in the brain.

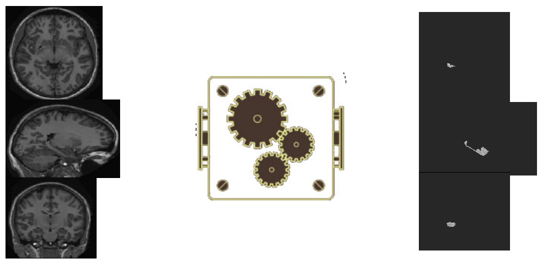

Figure 5: Input as MR Image(top: axial, middle: sagittal, bottom: coronal) and Output as Ground Truth

There are many methods to do such task, here we explore convolutional neural network architectures for patch-based segmentation on OASIS and ADNI dataset [13]. For both dataset pre-processing and post-processing are identical. We have 23 images in OASIS and 25 images in ADNI set with ground truth. For training, validation, and testing, 60% of the images are used as the training set, 20% as the validation set, and 20% as the testing set.

4.1 Dataset

For any training, the most basic requirement is good training examples. Choosing dataset to train our model is crucial. For this we've chose two different datasets - Open Access Series of Imaging Studies(OASIS)[2] dataset and Alzheimer's Disease Neuroimaging Initiative(ADNI)[3] [18]. In the OASIS dataset, the original images are obtained from and corresponding ground truth images provided by Visual Computing Lab[4].

4.2 Pre-processing

Before we begin segmenting an image, we first crop it down to a rectangular matrix so that we can perform masking in normalized coordinates. In the both cases of

[2]http://www.oasis-brains.org/
[3]http://adni.loni.usc.edu/
[4]https://vcl.iti.gr/hippocampus-segmentation/

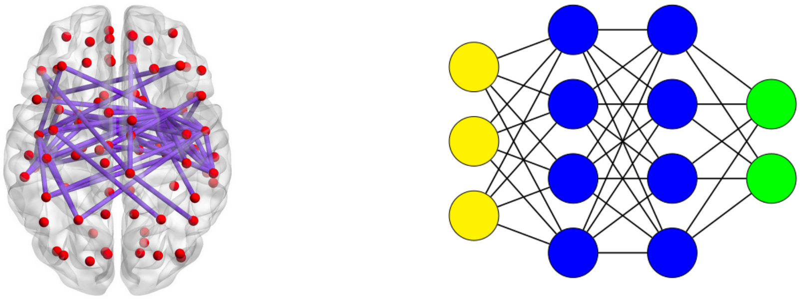

Figure 4: Connection in Human Brain & Artificial Neural Network

the ADNI dataset and OASIS dataset, all images are already in the same orientation, so no rotation is required. From going through all images in the training set, it is determined that the hippocampi are always in the region $(0.42 < x < 0.81, 0.30 < y < 0.67, 0.22 < z < 0.80)$, relative to each dimension of the matrix of respective brains. The new size of region is expanded by 0.03 on each side, and use $(0.39 < x < 0.84, 0.27 < y < 0.70, 0.19 < z < 0.83)$ as the mask [10]. Outside of this region, all voxels are automatically classified as non-hippocampus. All training patches are drawn from within the mask.

4.3 Training

After pre-processing data matrix produced. For each 3-D image, 2-D slices extracted. Each 2-D slice matrix is reshaped in a 1-D array, and that is the row of a data matrix. This data matrix separated into training and testing data and input for a neural network.

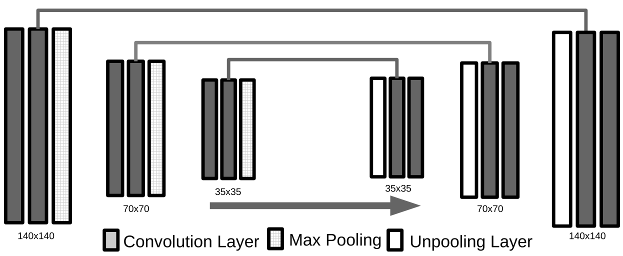

Figure 6: Basic Model

4.3.1 Convolution Layer

This is the very first layer of the model. The input of this layer is an image matrix. This layer extract features from given input and maintains the spatial relationship among the image pixels. It convolves (slide pixel by pixel called stride) the mask or filter or kernel through the whole image and produces the output matrix. It computes element-wise multiplications then multiplications are summed, and a single number is generated[5].

4.3.2 ReLU function

The Rectified Linear Unit (ReLU) is a simple activation function which apply through each element of the matrix along the spatial dimension (hight, width). The basic function is $\mathbf{f(x) = max(0, x)}$ thresholding at *zero*. It preserves spatial dimensions of input. ReLU is used after each convolution layer. Instead of using ReLU we use other function such as **tanh** or **sigmoid** but in most cases ReLU is better[6].

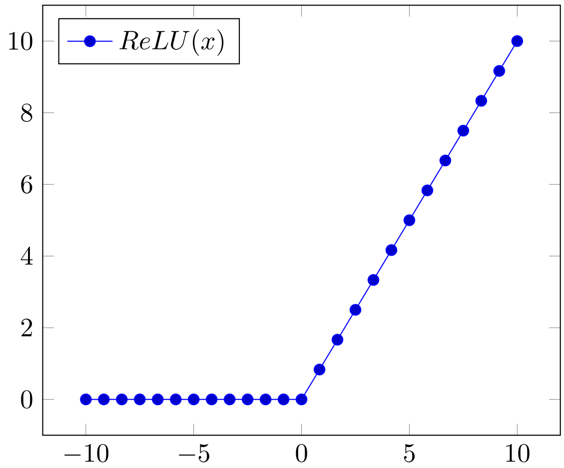

Figure 7: ReLu:=$\mathbf{max(0, x)}$

4.3.3 Max Pooling

Along width and height of given input matrix, it operates down-sampling, and the pool mask convolves through a dimension of the matrix and pools the maximum element for each position. It keeps important features and drop the less importance and reduces the size of the input matrix. There are many other types of pooling such as average, sum, etc. also known as spatial pooling[5].

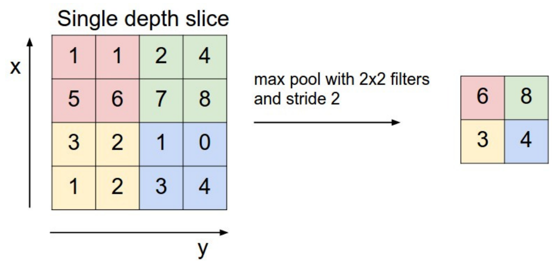

Figure 8: Max Pooling

4.3.4 Unpooling or Upsampling

This is the layer which is used in decoding part basically, it the reverse operation of Max pooling. In the studied model Figure 6, encoding part contains 6 convolution layers and 3 max-pooling layers and 3 unpooling layers and 6 convolution layers in the decoding part. After each convolution layer, ReLU is used. Using 2x2 mask (window) with 2 strides max pooling is performed.

[5] http://cs231n.github.io/convolutional-networks/
[6] https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

5 Results

All the computation done on nVidia GTX GPU, with Python and Keras neural network library using Theano's CUDA implementation as backend. In order to evaluate the proposed architecture, we have experimented the our model on the OASIS and ADNI dataset. In the OASIS dataset, 140x140 2162 2-D images are extracted from the 23 3-D images and similarly, 2350 2-D are extracted from 25 3-D images from ADNI dataset. We have iterated experiment for 5 epochs and 10 epochs for above configuration and achieved the testing accuracy of 86.65% with training accuracy 97.06%. The following table gives complete patch-wise and sample-wise results.

Sr	X	Y	Z	δ	X Accuracy	Y Accuracy
1	13	5	5	5	97.11	86.23
2	13	5	5	10	97.23	87.47
3	15	5	5	5	96.43	85.39
4	15	5	5	10	96.93	86.21
5	28	10	10	5	97.27	87.11
6	28	10	10	10	97.41	87.51

Table 1: X: #Training Images, Y: #Testing Images, Z: #Validation Images, δ: #Iterations; here - #: Times, sr no. 1, 2 used OASIS dataset, and 3, 4 used ADNI dataset and in the 5, 6 both datasets are used(13 images from OASIS and 15 images from ADNI)

A comparative study of method presented in this report and some of the other methods is given in below table. The dataset used by all methods is provided by OASIS and ADNI.

Sr	Method	Samples	Accuracy
1	Graph Cut	23	86%
2	Gradient Distribution	23	87%
3	Patch Based LF	80	88%
4	Patch Based CNN	423	91.12%

Table 2: LF: Label Fusion

Graph cut, Gradient Distribution, Patch Based CNN used OASIS dataset and in the Patch Based Label fusion ADNI dataset is used.

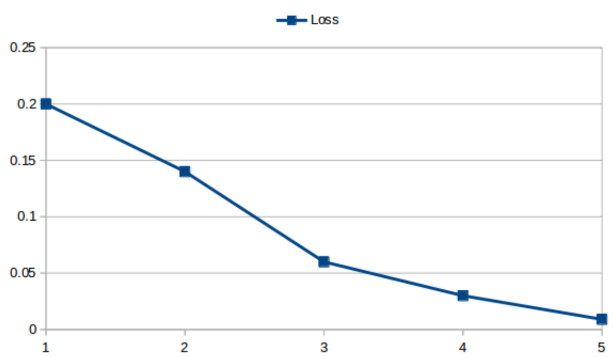

Figure 9: Loss Function x-axis: no. of epochs and y-axis: loss %

6 Conclusion and Future work

Neural-network-based techniques are used successfully for the segmentation process. The methodology was tested on the OASIS and ADNI datasets using Auto-Encoder convolutional neural network. In the future we can achieve state-of-the-art and even surpass state-of-the-art. In the future we will make our model more robust, less time consuming and extend to different datasets.

Acknowledgement

I would like to express my special thanks of gratitude to my mentor Dr. Ashish Phophalia who helped me in doing a lot of Research, and providing the resources required for the faster computation, and I came to know about so many new things. I am thankful to my teammate Mr. Kushal Jangid for his immense help to complete this project. At last, I would also like to thank classmates who helped me a lot in finishing this project.

References

[1] Catalin Amza. A review on neural network-based image segmentation techniques. *De Montfort University, Mechanical and Manufacturing Engg., The Gateway Leicester, LE1 9BH, United Kingdom*, pages 1–23, 2012.

[2] Jimmy Ba and Rich Caruana. Do deep nets really need to be deep? In *Advances in neural information processing systems*, pages 2654–2662, 2014.

[3] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv:1511.00561*, 2015.

[4] Yoshua Bengio et al. Learning deep architectures for ai. *Foundations and trends® in Machine Learning*, 2(1):1–127, 2009.

[5] Owen T Carmichael, Howard A Aizenstein, Simon W Davis, James T Becker, Paul M Thompson, Carolyn Cidis Meltzer, and Yanxi Liu. Atlas-based hippocampus segmentation in alzheimer's disease and mild cognitive impairment. *Neuroimage*, 27(4):979–990, 2005.

[6] Ivana Despotović, Bart Goossens, and Wilfried Philips. Mri segmentation of the human brain: challenges, methods, and applications. *Computational and mathematical methods in medicine*, 2015, 2015.

[7] Hayit Greenspan, Bram van Ginneken, and Ronald M Summers. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5):1153–1159, 2016.

[8] Ali Işın, Cem Direkoğlu, and Melike Şah. Review of mri-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, 102:317–324, 2016.

[9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[10] Matthew Lai. Deep learning for medical image segmentation. *arXiv preprint arXiv:1505.02000*, 2015.

[11] Fei Fei Li. How we're teaching computers to understand pictures, February 2015.

[12] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer, 2015.

[13] Holger R Roth, Le Lu, Ari Seff, Kevin M Cherry, Joanne Hoffman, Shijun Wang, Jiamin Liu, Evrim Turkbey, and Ronald M Summers. A new 2.5 d representation for lymph node detection using random sets of deep convolutional neural network observations. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 520–527. Springer, 2014.

[14] Berkman Sahiner, Heang-Ping Chan, Nicholas Petrick, Datong Wei, Mark A Helvie, Dorit D Adler, and Mitchell M Goodsitt. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE transactions on Medical Imaging*, 15(5):598–610, 1996.

[15] P Seebock. Deep learning in medical image analysis. *Master's thesis, Vienna University of Technology, Faculty of Informatics*, 2015.

[16] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[17] Endel Tulving and Hans J Markowitsch. Episodic and declarative memory: role of the hippocampus. *Hippocampus*, 8(3):198–204, 1998.

[18] Bradley T Wyman, Danielle J Harvey, Karen Crawford, Matt A Bernstein, Owen Carmichael, Patricia E Cole, Paul K Crane, Charles DeCarli, Nick C Fox, Jeffrey L Gunter, et al. Standardization of analysis sets for reporting results from adni mri data. *Alzheimer's & Dementia*, 9(3):332–337, 2013.

[19] Pingkun Yan, Kenji Suzuki, Fei Wang, and Dinggang Shen. Machine learning in medical imaging, 2013.