[UCI]I804:24011-200000267368

August 2019
Master's Degree Thesis

Brain MRI segmentation and Alzheimer's disease classification using CNN

Graduate School of Chosun University

Department of Information and Communication Engineering

Bijen Khagi

Brain MRI segmentation and Alzheimer's disease classification using CNN

CNN 및 그 유도체를 이용한 뇌 MRI 분할 및 알츠하이머 병 분류

June 8, 2019

Graduate School of Chosun University

Department of Information and Communication Engineering

Bijen Khagi

Brain MRI segmentation and Alzheimer's disease classification using CNN

Advisor: Prof. Goo-Rak Kwon

This thesis is submitted to Chosun University in partial fulfillment of the requirements for a Master's degree

June 2019

Graduate School of Chosun University
Department of Information and Communication Engineering

Bijen Khagi

This is to certify that the master's thesis of

Bijen Khagi

has been approved by examining committee for the thesis requirement for the master's degree in engineering.

Committee Chairperson	**Prof. Jae-Young Pyun**	____________
Committee Member	**Prof. Bum-Shik Lee**	____________
Committee Member	**Prof. Goo-Rak Kwon**	____________

August 2019

Graduate School of Chosun University

Table of contents

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Brain MRI segmentation and Alzheimer's disease classification using CNN

Bijen Khagi

Advisor: Prof. Goo-Rak Kwon

Dept. of Information and Communication Engineering

Chosun University

Brain MRI is an important bio-marker for identifying neurodegenerative diseases like Alzheimer, Dementia etc. Based on the information associated with MRI, medical diagnosis is performed. Here the aim is to develop computer aided diagnosis (CAD) of MRI, so that the proper classification is carried out to assist final diagnosis. Similarly, the segmentation of brain into Gray, White and CSF parts is equally important for brain related diagnosis. Hence the goal is to develop such kind of CAD using image trained Convolution neural network (CNN) and its other pre-trained architectures.

The presented approach is to use deep neural network for segmenting the MRI images of heterogeneously distributed pixels into a definite class allocating a label to each pixel. This enables the application of the segmentation process on preprocessed MRI images, to train the network that can be used to segment other test images. Since labels are considered expensive resources in supervised training, fewer training images and training labels were processed to obtain optimal accuracy based segmentation CNN. In order to validate the performance of the proposed

idea, testing is conducted on other test images (available in the same database) that are excluded in the training; the obtained result is of decent visual quality in terms of segmentation and temperately comparable to the ground truth image. The average computed Dice similarity index for the test images is almost 0.8, whereas the intersection over union (IoU) based Jaccard similarity measure is approximately 0.6, which is better compared to some other methods.

The performance result of classification of Normal controls (NC) and Alzheimer's disease (AD) using pretrained model (trained on natural images) is presented along with its result in medical image classification via scratch trained model; trained from available medical MRI images, in order to have a comparative analysis. Shallow tuning and fine tuning of pretrained model (AlexNet, GoogLeNet, and Resnet50) in a bunch of layers were performed in order to find the impact of each section of layers in classification result. 28 NC and 28 AD patients were used for classification, selecting 30 important slices from each patient. Once all the slices are collected, each model was trained, validated and tested in ratio of 6:2:2 on random selection basis. The resulting testing results are reported and analyzed so, that the final CNN model was built with minimal number of layers for optimal performance.

한 글 요 약

CNN 및 그 유도체를 이용한 뇌 MRI 분할 및 알츠하이머 병 분류

비젠 카기

지도교수 : 권구락

정보통신공학과

조선대학교

뇌 MRI 는 알츠하이머 (Alzheimer), 치매 (Dementia)와 같은 신경 퇴행성 질환을 확인을 위한 중요한 바이오 마커이다. MRI 와 관련된 정보를 기반으로 의료 진단이 수행된다. 이 연구의 목표는 MRI 의 CAD(Computer Aided Diagnosis)를 개발하여 적절한 분류가 이루어 지도록 하는 것이다. 마찬가지로, 뇌를 진단하는 것에 있어서, 회백질, 백색질 및 뇌척수액으로 세분화하는 것은 중요하다. 따라서 본 연구에서는 영상 학습이 된 Convolution neural network (CNN) 와 기타 사전 훈련 된 아키텍처를 사용하여 뇌 진단 분류가 가능한 CAD 를 개발하는 것을 목표로 한다.

제안하는 접근 방식은 이질적으로 분포 된 픽셀의 MRI 이미지를 각 픽셀에 라벨을 할당하는 명확한 클래스로 세분화하기 위해 심층 신경망을 사용하는 것입니다. 이 방법은 전처리 된 MRI 이미지에 대한 영상 분할 프로세스의 적용을 가능하게 하여 다른 테스트 이미지를 분할하는 데 사용할 수 있는 네트워크를 학습한다. 기존의 학습된 데이터 리소스는 값이 비싸기 때문에 CNN 분할을 기반으로 최적의 정확도를 가지는 적은 수의 학습 영상과 라벨을 사용한다. 제안된 아이디어의 성능을 검증하기 위해 학습에서 제외된 다른 테스트 영상에 대해 테스트를 수행한다. 얻어진 결과는 실지 검증 이미지와 세분화 측면에서 높은 성능을 가진다.

테스트 이미지의 평균 계산된 Dice 유사도 지수는 0.8 에 가까우며, Jaccard 유사성 측정 기반의 유니온(IoU)는 0.6 이며 다른 방법과 비교했을 때 더 좋은 성능을 보인다.

사전 훈련된 모델을 사용하여 정상 대조군과 알츠하이머 병을 분류한 결과는 스크래치 훈련 모델을 통한 의료 영상 분류 결과와 함께 제시된다. 비교 분석을 하기 위해 사용 가능한 의료용 MRI 영상에서 훈련을 진행한다. 분류 결과에서 레이어의 각 섹션의 영향을 찾기 위해 일련의 레이이어에서 사전 조정된 모델(AlexNet, GoogLeNet, Resnet50)을 약간 수정하여 실행한다. 28 명의 NC 환자와 28 명의 AD 환자를 분류하기 위한 학습에 사용하여 각 환자에게서 약 30 개의 주요한 슬라이스를 선택한다. 모든 슬라이스가 수집되면 각 모델의 무작위 추출 기준에 따라서 6:2:2 의 비율로 훈련되고 이를 검증하기 위한 테스트가 진행된다. 최종 테스트 결과를 통해 최적화된 성능을 가진 최소한의 레이어를 가진 CNN 모델이 제작된다.

1. Introduction

Image processing has been successfully used in various fields for object prediction and detection. From finding defective items in factory, to weather prediction from satellite, it has been an essential tool and subject of study for many scientist and researchers. Similarly in the field of medicine, image processing has proven to be an efficient, potential tool for detection and diagnosis of Tumor [1, 2], Brain Lesion [3, 4], Alzheimer's disease [5–9] and cancer [10]. The practice of using machine learning approach in classification is an ongoing trend. With the development of various imaging technologies like MRI (Magnetic Imaging Resonance), PET (Positron-emission tomography), Computed Tomography (CT) scan in medical examination, there have been lots of efforts to process, simulate and interpret the result for the purpose of Computer Aided Diagnosis (CAD) that will be of vital importance for medical professionals.

MRI of brain needs an impeccable analysis to investigate all its structure and pattern. This analysis may be a sharp visual analysis by an experienced medical professional or by a CAD system that can help to predict, what may be the recent condition. Similarly, on the basis of various information, and technique a system can be designed to detect whether a patient is prone to Alzheimer's disease or not. And this task of detection of abnormalities at an initial stage from brain MRI is a major challenge in the field of neurosciences. Brain MRI segmentation is significant in several clinical applications and influences the outcome of the entire analysis that depends on its results of anatomical and structural regions. For instance, MRI segmentation is repeatedly used for imagining different brain structures, analyzing brain development in infants, calculating delineating lesions volume, and various image guided intrusions, surgical preparation and finally tissue segmentation.

In general image segmentation is the idea of presenting and partitioning image content into meaningful parts and patterns also called image-attributes. Deep neural networks (DNNs) have been extremely effective in segmenting outdoor scenes with high complication, dissimilar patterns, variable texture, and wide pixel range. In the present study, this model is used for segmenting MRI images of the brain, which are moderately simpler than natural outdoor scenes which have complex attributes and wider order of class. The precise segmentation of a two dimensional image has always been a puzzling task, and various approaches have been proposed for better result including manual and automatic, supervised, semi-supervised and unsupervised, standalone and cascaded neural network-based techniques [11]. Similarly, convolutional neural networks (CNNs) have been operational in machine learning and functional on various industrial, medical, and commercial fields particularly on imaging techniques. Conventional artificial neural networks (ANNs) have also been used in medical image processing, particularly in MRI image segmentation and Alzheimer's disease classification [12, 13].

Besides segmentation, classification is equally important for detecting healthy normal control (NC) from Alzheimer's disease (AD) affected patient MRI. Here comes the true contribution of image processing in the field of medical imaging, as a new CNN model is developed trained on those two types of MRI, so that network once trained, can be used to classify a test MRI image to determine its class i.e. either AD or NC. Here, the CNN is used as a generic feature extractor itself to build the classification model which will be discussed in detail in Chapter 2.

This thesis relates the implementation of deep learning scenario in two fields, first being the anatomical segmentation of brain MRI into its tissue content and second being the medical diagnosis of AD affected MRI to distinguish with NC.

1.1 Overview and motivation

Working in the field of medical imaging system, especially MRI, substantial work has been performed for AD diagnosis and also segmentation but the ideas are mainly conventional and use hand crafted features as primary features. Conventional algorithm like k-means, Expectation Maximization, bag-of-words and edge detection transforms like DWT, DTCWT, Slantlet etc. with supportive modification and enhancement are used as feature extractor or features itself [14–18]. But with the growth of deep learning techniques and its successful implementation in natural image classification in primary benchmark dataset being like ImageNet [19], Cifar-10[20], Oxford Flower [21] etc. have achieved the state of art performance, outperforming all conventional algorithms. Besides, deep learning is successfully implemented for cancer classification, lesion classification and much more medical diagnosis. With numerous research ongoing and paper being published in brain MRI, I have used deep learning with some proposed architecture and method to solve the problem domain i.e. brain tissue segmentation and detection of healthy and AD MRI. Especially I am inspired by the work of Vijay Badrinarayanan et al. [22] for segmentation who implemented semantic segmentation using SegNet in outdoor scene classification and the work of A. Krizhevsky in image classification for ImageNet dataset motivated in classification. The networks trained using AlexNet [23], GoogLeNet [24], Resnet50 [25] are part of the experiment.

1.2 Objectives

Currently, deep learning techniques are being widely used in every field for developing decision making system based on Artificial intelligence (AI) of Computer Aided System (CAS). The practice of using deep learning approach in classification is an ongoing trend. Convolutional Neural Network (CNN) originally designed for object detection finds its use in image classification, segmentation, pattern recognition etc. It is one of the initial deep learning techniques that have been developing as an important tool for machine vision and AI due to its autonomous working nature. CNN was successfully engaged in larger database with lowest error rate in 2012 by A. Krizhevsky et al. [23] in ImageNet database for classification of 1000 image types (class). Later various variants and advancement of CNN were proposed by different researcher for object recognition and image classification like Resnet50 [25], GoogLeNet [24] and R-CNN [26].

1.3 Contribution

This thesis highlights the use of CNN along with other supportive modification of CNN, for Brain MRI segmentation and Alzheimer's affected MRI classification task. I present here the experimental result of various CNN models that have been implemented, modified and customized for the above mentioned task. Basically, a CNN architecture model is proposed for brain MRI segmentation into its three constituent's tissue i.e. gray Matter (GM), cerebrospinal fluid (CSF) and white matter (WM).

Two separate CNN models are proposed for segmentation and classification task, former being motivated by pixel-label based segmentation and latter being motivated by object recognition algorithms. Similarly, segmentation architecture is inspired by SegNet model [22], and classification architecture is by AlexNet [23]. The reason behind idealizing these models is because SegNet was trained using fewer training material alike in present case with less distinct classes and it was the global success of AlexNet in ImageNet which has been a prevalent choice for many researcher.

For classification, we will investigate various CNN models along with the proposed scratch trained CNN. Then test the classification of two different types of MR images i.e. AD and NC using various tuning techniques of all models, transferring weights from them at different learning rates for different layers, to find out the best CNN model and their impact due to architectural modifications. On the other hand I will use simplified SegNet architecture based CNN with memory based encoder-decoder network for segmentation of three tissue types. The obtained result is convincing and supportive towards use of CNN, and can be aimed towards development of medical CAD system.

1.4 Thesis Layout

This thesis presents the cumulative work performed during my master degree course. It is organized as follow, Chapter 1 Introduction, presents the motivation behind my work and defines the problem domain from general point of view where I have clarified about my contribution in research. Chapter 2 describes the subjective matter of study and provides summary of how CNN can used for segmentation and classification of images. Chapter three provides the review of similar work done by various researchers in most concise and informative way as much as possible to my extent. The basics of CNN along with major mathematical operation performed are reviewed and explained briefly in Chapter 4. Besides the pre-trained CNN models used for transfer learning and tuning CNN to the problem domain are also explained in this chapter. To avoid confusion, proposed model for segmentation and classification process involved are separately included in different section. The fifth chapter explains the dataset employed separately for MRI segmentation and classification of AD and NC, along with related statistics. The complete process of converting raw Nifti MRI files to two dimensional images is explained here. Chapter six describes the proposed experiment along with performance parameter to evaluate the result. Segmentation and classification results have been thoroughly explained, discussed and compared with other algorithms in respective sections. Finally, I conclude the thesis and present the summary in the final chapter.

2. Background

2.1 Brain MRI and Alzheimer's disease

MRI is a magnetic-field gradient based neuroimaging biomarker technique that provides anatomic and physiological information for diagnosis [27] of different parts of body including the brain. It uses strong magnetic field and radio-wave to generate higher-quality picture of the structure and volume of the brain. The high quality and greater contrast image of the anatomical structures along with functional images of various organs helps the medical professionals to obtain maximum data and information without any physical operation of participant [28].

Alzheimer's disease is a neurodegenerative diseases, that affects the functional and structural parts associated with brain. It is one of the most familiar forms of dementia that develops problems with memory, behavior, thinking and other intellectual abilities disturbing personal and socio-economic aspect as well.

2.2 Convolutional neural network for supervised network training

Conventionally CNNs contain many convolutional layers that transform their input with convolution filters initialized differently with various size and stride of a small extent that runs over each image to pass the extracted feature vector to the succeeding layers. CNN is a supervised training phenomenon as it requires user defined target values generally called as label, ground truth or true value. On the basis of the error between the predicted value and target value, the loss function performs the iterative-training for different epochs using backward propagation until the parameters of the all layers participating in training remains constant or almost constant with minimum error between predicted value and target value. Here it is to be noticed that, training a CNN is directly affected by the number of training materials and the quality of label or ground truth. The network performs accordingly how it is trained hence called supervised network. I will be performing training at different layers using pre-trained and self-trained network on this study. But firstly before going detail into the work, it will be helpful if we go through some major layer wise mathematical operation used inside the CNN network.

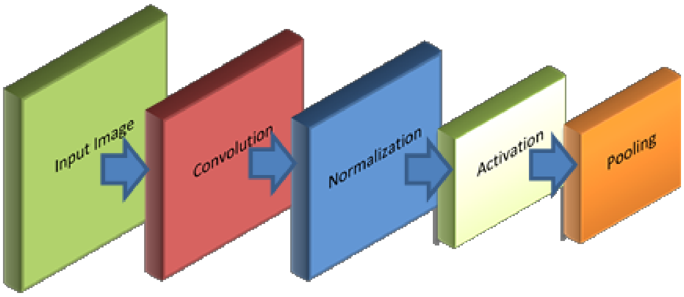


Figure 2-1. Convolutional layer architecture in CNN

(1) Convolutional layer: It consists of two dimensional filters (kernels) of specified size that runs across the input signal (image). Mathematically kernel is the matrix to be operated with the input signal and step or stride controls how much the filter convolves across the input signal. The convolution operation of the input signals with the kernel follows equation.

$$y_l = \sum_n^{N-1} x_n w_{l-n} \quad (2.1)$$

The convolution operation is as follows: where x_n is signal input for layer l , w_l is its filter weight, and N is the number of elements in x. The output vector is y_l. The subscripts denote the n^{th} element of the vector.

The output of the convolution is a reduced version of input image known as the feature map or feature vector. Here, one important consideration is the initial constituent of filter also known as filter-weight, which is normally a random value however different initialization techniques have been proposed to enhance the convergence process of network.

(2) Pooling layer: The feature vector or feature map obtained from convolution are bulkier in dimension due to lager number of filters used hence pooling operation is performed to select a representative feature map. This layer is also known as the down-sampling layer as it reduces the size of output neurons from the convolutional layer which may cause computational workload and over fitting. Various types of pooling operation may be average pooling, max-pooling, min-pooling which selects the average, maximum and minimum value from the selected pool size filter respectively.

(3) Activation layer: It a common practice to uses various activation functions to transform the feature between each layer so that the convolution process gets smoother and faster without losing important information. Mostly used is:

(a) Rectified linear activation unit (ReLu): Rectifier linear units are used to impart non-linearity to the network structure and select non-negative number as activated features as shown in equation [29].

$$f(x) = \max(0 \colon x) \tag{2.2}$$

As the equation suggests it misses the negative weights to maintain a range of [0, *x*] but a slightly different ReLu called leaky rectifier linear unit (LeakyRelu) [30] proved better than the original ReLu itself. This may be due to its properties which add nonlinearity and sparsity in the network structure, resulting network robustness to minor changes such as noise in the input. Equation (2.3) shows the LeakyRelu function.

$$f(x) = \begin{cases} x \; if \; x > 0 \\ 0.01x \; otherwise \end{cases} \tag{2.3}$$

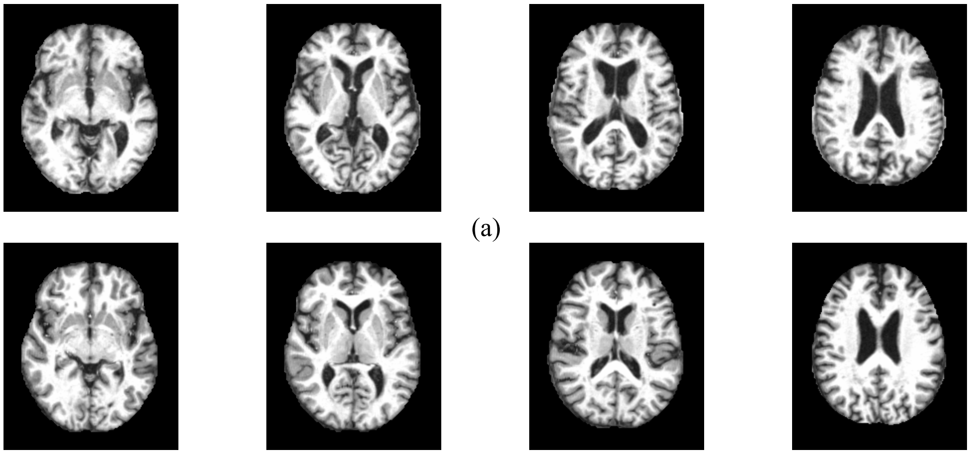
(4) Softmax: This function is a class-based prediction function that computes the probability distribution (PD) of the k output classes. Hence the final classification layer uses softmax function to predict the final class of the input MRI image. For $i=1, 2 \ldots k$ number of classes with input feature vector x_i, the i^{th} probability score p_i is,

$$p_i = \frac{e^{x_i}}{\sum_1^k e^{x_k}} \tag{2.4}$$

Equation (2.4) suggests p_i being value between 0 to 1; hence the i^{th} class with maximum probability score wins the race.

2.3 Brain MR image segmentation and AD classification

Brain MRI segmentation is significant in several clinical applications and influences the outcome of the intact analysis because various processing operations depends on accurate segmentation of anatomical and structural regions. As discussed earlier, MRI segmentation plays a vital role in imagining different brain structures, delineating lesions, analyzing brain development of infant that has immense role in clinical study of brain. In MRI, the obtained pictures of tissues are heterogeneously concerted in terms of intensity owing to the bias field and the partial volume effect that reflects the tissue content of the brain, i.e. GM, WM and CSF that needs to be accurately segmented and partitioned. Recent study has revealed that, there is a minor increase in CSF content in brain ventricles and sulci, along with substantial decrease in gray matter content and brain volume in AD patients [57]. The segmented tissue content reveals the volume of each type, and as AD is a neurodegenerative disease, the shrinking brain volume may alarm the case of possible diagnosis of brain atrophy that may cause dementia and finally AD. This kind of approach are used to develop structural brain-mapping and volume based feature extraction tools like Freesurfer, FSL etc. [58] which are actually based on gray-white matter segmentation and consequently detect AD. This fact suggests how segmentation helps in AD identification and classification.



(a)

(b)

Figure 2-2. (a) Typical AD patient MRI scan presenting 71st, 80th, 90th and 100th slice respectively Figure 2-2. (b) Typical NC patient MRI scan presenting 71st,80th,90th and 100th slice respectively

Existing methods [11] used extensively by many researchers and found excellent in case of MRI segmentation have a certain way of feature extraction and criteria's like thresholding, contours, clustering etc. But on the other hand deep neural network are now proving to be better, highly computational for large data because of encoder-decoder based network built in CNN architecture. The features are automatically investigated from low level features like edge, blob, line etc. to high level features like color, shape, detail etc. in a hierarchical manner by each layer. The activation layer like ReLu, Leaky ReLu, sigmoid helps to make those features more clear and computable. Hence, we can easily get the segmentation result, the only concern here is to design the network appropriately and train it properly as it requires large amount of ground truth for fine convergence.

3. Related works

Y. Zhang et al. [31] proposed singular value decomposition (SVD) algorithm for brain MRI segments feature extraction and classification of Alzheimer's disease (AD), Mild Cognitive Impairment (MCI) and Normal Control (NC using Principal Component Analysis (PCA) as feature reduction technique. They finally used 22 representative reduced features to be classified by Kernel SVM-Decision Tree. Chaplot et al. [32] proposed a two-dimensional Discrete Wavelet Transform (2D-DWT) with Daubechies wavelet decomposition to obtain the approximation coefficient as well as utilized a self-organization map and Support Vector Machine (SVM). Slantlet transform was used by Maitra et al. [18], which is an improvised version of DWT that they used for intelligent MRI classification system. El-Dahshan et al. [33] extracted all the coefficients using the multi-resolution decomposition of a DWT so that the features were reduced in smaller dimension using PCA. They achieved 97-98% accuracy using K-Nearest Neighbor (KNN) and feed forward back propagation artificial neural network (FP-ANN) respectively. Zhang et al. [34] accomplished 100% success rate by using a feed-forward back propagation neural network and Scaled Chaotic Artificial Bee Colony (SCABC) to classify normal MRI images from abnormal ones. Similarly, Jha et al. [35] used Dual-Tree Complex Wavelet Transform (DTCWT), PCA with Feed-Forward Neural Network (FNN) and achieved more than 90% accuracy to classify AD and NC MRI. Recently, Syrine Neffati et al. [36] used Downsized Kernel Principal Component Analysis (DKPCA) and multiclass Support Vector Machine (MSVM) for AD MRI detection and obtained accuracy up-to 92.5% using kernel MSVM.

Lo et al. [37] applied artificial CNN in medical image analysis in 1995 for lung nodule detection. CNN first successful real-world application was performed by LeCun et al. [38] with 'LeNet' for hand-written digit recognition using gradient based learning. Consequently various convolution based network with different feature extraction and optimization techniques bloomed and erupted. Regarding CNN implication in medical field, MRI and CT scan imaging has been successfully tested. Nima Tajbaksh et al. [39] tested CNN in medical Images for polyp detection and Pulmonary embolism detection, where they highlighted pretrained or fined tuned CNN performed as good as scratch trained CNN and suggested layer wise tuning for practical performance. Similarly, Hoo-Chang Shin et al. [40] tested CNN architecture for Lymph-Node detection and Interstitial Lung disease Classification, where they also tested pretrained CNN network (AlexNet, GoogLeNet and CIFAR-10 trained [41]) and also used transfer learning technique from this CNN. Transfer Learning in CNN can be done in principally two ways, either the weights of all CNN structure gets cascaded to another CNN layer with classification Layer output or simply using "off-the-shelf CNN features" [42] where CNN acts as a generic feature extractor to be evaluated further. I will present a comparative analysis of using the former idea and later perform shallow or fine tuning in the models to classify MR images.

Similarly it is used in MRI segmentation as well. Zhang et al. [43] used 2D patches of weighted and fractional anisotropic MR image for 6-8 month old infants as CNN training images, for segmentation of brain tissue viz. WM, GM and CSF. De Brébisson et al. [44] introduced multiple parallel 2D patches network based on multiple image patches and kernel sizes to support CNN based MRI classification Moeskops et al. [45] presented multi scale CNN for the automatic segmentation of anatomical brain T1-T2 MRI of young and ageing adults. The multi-scale method improved accurate segmentation details and spatial consistency for different tissue classes of brain MRI. This method allowed the CNN to learn multi-scale features that estimate both intensity and spatial characteristics which they used it for 134 regions segmentation as in MICCAI challenge of atlas labeling.

Technically, both the segmentation and classification have similar operation inside CNN, but it's the building architecture and ground truth that makes them different. For segmentation, a converging-diverging network based on encoder-decoder connected sequentially and parallel to transfer address of each maxpool layer is proposed. The network is supervised by pre-segmented image during training with loss-function being the difference of whole image. However, for classification the architecture is converging with only sequential connection. The network is trained by categorical image labels with the loss function as the difference of target values.

4. CNN and its architecture

Conventionally CNNs contain many convolutional layers that transform their input with convolution filters initialized differently with various size and stride of a small extent that runs over each image to pass the extracted feature vector to the succeeding layers.

4.1 Part I: MRI Segmentation

4.1.1 CNN models

a. AlexNet

CNN implementation in image classification and computer vision was lime lighted with the work of Alex Krizhevsky [23] in the ImageNet LSVRC-2012 contest. They introduced 8 layer based architecture with 5 convolutional and 3 fully connected layers based learnable CNN. These convolutional layers were followed by a maxpool layer and normalization layer. Softmax classifier was used for classification trained on basis of cross-entropy loss. This network later became popular as ‘AlexNet’.

The network consist around sixty million parameters and 650,000 neurons. They used rectified linear units (ReLu) $f(x) = max\ (0; x)$ for down-sampling features which could train faster than tradition non-linear activation functions like $f(x) = tanh(x)$ or $f(x) = (1 + e^{-x})^{-1}$. It also prevented the network from over-fitting on the ImageNet database which was mainly influenced by other layer called ‘dropout’. Dropout [46] technique sets the output of each hidden neuron having probability 0.5 to 0 so that, the neurons which are actually “dropped out” in this way do not subsidize to the forward pass and hence do not participate in back propagation. Although weights are shared but every time an input is presented, the neural network sample finds a changed architecture. Besides, data augmentation was performed by image translations and horizontal reflections. This reduces over-fitting by expanding the dataset and also brings variability in training material.

This architecture achieved top-1 and top-5 error rates of 37.5% and 17.0% in ImageNet LSVRC-2010 and later in ILSVRC-2012 accomplished a winning top-5 test error rate of 15.3%, leaving behind 26.2% error achieved by the second-best entry.

b. ResNet

In 2015 He et al. [25] proposed a residual learning framework that could the train the CNNs easily with residual learning framework substantially deeper layer than those used previously. In contrary to plain network, residual network insert shortcut connections which turn the network into its counterpart residual version, called ResNet blocks in their architecture. Rather than learning a function, the residual block solitary learns the residual and hence pre-conditioned towards learning mappings in each layer that are close to the identity function. This helped so that, deeper models could be trained effectively. Ultimately this architecture won the ImageNet challenge in 2015 with 3.57% error on the ImageNet test set, besides, they also obtained 28% improvement on the COCO object detection dataset.

ResNet architecture was tested with 18, 34, 50, 101 and 152 layers. Some of them are:

i) 50-layer ResNet: It replaces each 2-layer block previously used in 18 and 34 layers blocks with 3-layer bottleneck block, each of size 1×1, 3×3 and 1×1. This model has 3.8 billion FLOPs.

ii) 101-layer and 152-layer ResNet: 101- layer and 152-layer ResNet were constructed similarly as 50 layered ResNet using 3-layer blocks, however the numbers of blocks in each layers were different and more in 4th convolutional block. Interestingly it is stated that, although the depth is considerably increased, the 152-layer ResNet (11.3 billion FLOPs) still has reduced complication than VGG-16/19 nets (15.3/19.6 billion FLOPs).

c. GoogLeNet

Szegedy et al. [24] presented Inception architecture based network named GoogLeNet, also referred as Inception, blocks. Here a module replaces the mapping defined in conventional neural network weight update equation as (4.1) with a set of convolutions of different sizes like small kernels, this allows a similar function to be represented with less parameters.

$$X_k^l = \varphi(W_k^{l-1} \times X^{l-1} + b_k^{l-1}) \tag{4.1}$$

Here, at each l^{th} convolutional layer, the input image of fixed size is convolved with a set of k kernels of weights referred as $W_k^l = \{W_1, W_2, W_3 \ldots., W_k\}$ and added biases $b_k^l = \{b_1, b_2, \ldots., b_k\}$. X_k^l represents the obtained feature vector, applied after non-linear operation φ which is mostly a ReLu operation as,

$$f(x) = max(0; x) \tag{4.2}$$

It has deeper and wider network than previous design with computational cost being same. The optimization of architecture was based on Hebbian principle and multi-scale processing. GoogLeNet is 22 layers deep when only layers with parameters are counted but including pooling it is 27 layers. The complete number of layers i.e. independent building blocks) used for the design of the network is about 100. The use of average pooling before the classifier is for "network on network" [47], with an additional linear layer. Convolutional blocks from one layer to other is connected using 'DeptConcat' block .The linear layer enables the networks to easily adapt to other label sets. It was found that a move from fully connected layers to average pooling upgraded the top-1 accuracy by about 0.6%; though the use of dropout remained essential even after eliminating the fully connected layers.

d. SegNet Layer

The SegNet layer is a deep full CNN architecture adapted for semantic segmentation that was proposed by Vijay Badrinarayanan et al. [22]. Basically, the proposed semantic segmentation approach was used for outdoor, indoor, and road scenes mostly for larger number of classes. It was originally designed for scene understanding applications. Hence, it needs to be efficient in terms of memory, operation, and computational time. The network is also considerably smaller in terms of the number of trainable parameters than other competing architectures, and can be used in training end-to-end pixel-label classes using stochastic gradient descent (SDG) optimization and the cross-entropy loss function for back-propagation.

The encoder used in SegNet is very identical to the convolutional layers in VGG16 [48]. The fully connected layers of VGG16 have been removed in SegNet, and hence the encoder network is considerably reduced and easier to train compared to other recent architectures [22, 49, 50, 51, 52]. The most important constituent of SegNet is the encoder–decoder network, which consists of a hierarchy of down-sampling encoders matching each upsampling decoder with associated feature vectors cycling inside them.

4.1.2 Proposed CNN

CNN has always been an important tool in machine learning; by using various types of neural networks, systematic training and testing of image on the basis of pixel labels can be performed. The encoder network used here consist of convolution layers of 64 filters, each of size 3×3, manually zero-padded, followed by batch normalization and ReLu activation unit and repeatedly followed by same convolution, batch normalization and ReLu for proper down-sampling and robust feature extraction. Decoder network also follows the same layers of convolution, normalization and activation but firstly unpool feature from the maxpool layer connected parallel with corresponding layer of encoder network.

Table 1. Proposed CNN for brain MRI segmentation

S.N	Layer Name	Type	Description
1	"Image Input"	Image	208 × 1761 images with "zero center" normalization
2	"encoder1_conv1"	Convolution	64 3 × 3 × 1 convolutions with stride [1 1] and padding [1 1 1 1]
3	"encoder1_bn_1"	Batch normalization	Batch normalization with 64 channels
4	"encoder1_relu_1"	ReLU	ReLU
5	"encoder1_conv2"	Convolution	64 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
6	"encoder1_bn_2"	Batch normalization	Batch normalization with 64 channels
7	"encoder1_relu_2"	ReLU	ReLU
8	"encoder1_maxpool"	Max pooling	2 × 2 max pooling with stride [2 2] and padding [0 0 0 0]
9	"encoder2_conv1"	Convolution	64 3× 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
10	"encoder2_bn_1"	Batch normalization	Batch normalization with 64 channels
11	"encoder2_relu_1"	ReLU	ReLU
12	"encoder2_conv2"	Convolution	64 3× 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
13	"encoder2_bn_2"	Batch normalization	Batch normalization with 64 channels
14	"encoder2_relu_2"	ReLU	ReLU
15	"encoder2_maxpool"	Max pooling	2 × 2 max pooling with stride [2 2] and padding [0 0 0 0]
16	"decoder2_unpool"	Max unpooling	Max unpooling

17	"decoder2_conv2"	Convolution	64 3× 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
18	"decoder2_bn_2"	Batch normalization	Batch normalization with 64 channels
19	"decoder2_relu_2"	ReLU	ReLU
20	"decoder2_conv1"	Convolution	64 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
21	"decoder2_bn_1"	Batch normalization	Batch normalization with 64 channels
22	"decoder2_relu_1"	ReLU	ReLU
23	"decoder1_unpool"	Max unpooling	Max unpooling
24	"decoder1_conv2"	Convolution	64 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
25	"decoder1_bn_2"	Batch normalization	Batch normalization with 64 channels
26	"decoder1_relu_2"	ReLU	ReLU
27	"decoder1_conv1"	Convolution	4 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
28	"decoder1_bn_1"	Batch normalization	Batch normalization with 4 channels
29	"decoder1_relu_1"	ReLU	ReLU
30	"softmax"	Softmax	Softmax layer for classification
31	"Pixel_classify"	Pixel classification layer	Class weighted cross-entropy loss with "Background", "CSF," "GM," and "WM" classes

Proposed CNN has an encoder network and a matching decoder network, which is followed by a final pixel-based classification layer. This architecture is shown in Table 1. To simplify the architecture, two encoder and two decoder networks have been employed: encoder1 is mapped to decoder1, and encoder2 is mapped to decoder2. encoder1 consists in hierarchical order encoder1_conv1, encoder1_bn_1, encoder1_relu_1, and encoder1_maxpool_1, whereas dencoder1 consists in hierarchical order of dencoder1_unpool_1, dencoder1_conv1, dencoder1_bn_1, and dencoder1_relu_1. encoder2 and decoder2 are similarly structured. Here, encoder1 is followed by encoder2, and dencoder2 is followed by dencoder1, as shown in 4-1. The first 13 layers, after the input layer, act as an encoder network that implements the convolution with 64 filter banks of size 3 × 3 to find the sets of features along with batch normalization in a mini-batch set of 8 images. ReLU is the activation function f(x) = max (0, x), which is used to eliminate negative values. Thereafter, the max pooling layer with a 2 × 2 window and stride size 2 (non-overlapping window) is performed, so that the resulting output is down-sampled by a factor of 2. Multiple layers of down-sampling are used to achieve more translation invariance and robust pixel classification. Likewise, the decoder in the decoder network up-samples the input layer feature map(s) un-pooling the learnt max pooling indices with the location of maximum feature values from the corresponding encoder feature maps. It is followed by the convolution and batch normalization layers to generate dense features that are equal in size to the input image. The details of the simplified architecture are tabulated in Table 1.

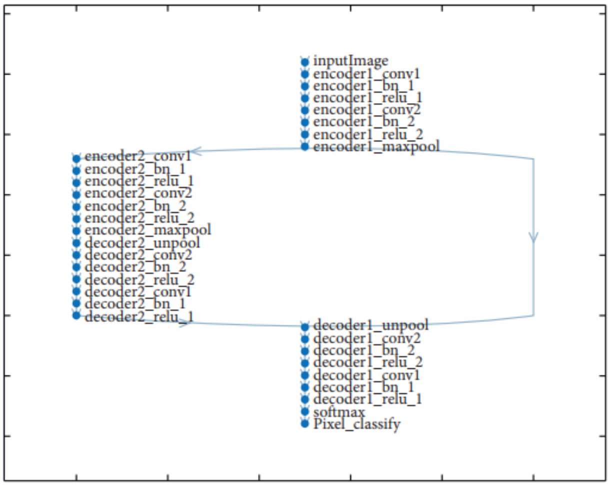


Figure 4-1. Schematic representation of 31 layers and 34 connections used in proposed CNN Network

4.2 Part II: MRI classification

4.2.1 Feature representation

Extracting features as mentioned above is a tedious process required with precise knowledge on respective method and also method are quite specific in nature like brain volume extraction needs one tool and cancer detection needs other. However with the introduction of deep neural layers, this extraction has been quite universally accepted, as they work irrespective of nature of work and extract the most possible features to represent the image. Using this proposed architecture I once trained the network using training sets and hence trained network was used to extract features from training set and testing set separated randomly in 7:3 ratios of all images.

To understand the weight update process [39] in CNN we can assume, if W_l^t denotes the corresponding weights of convolution kernel in l^{th} convolutional layer at *t* iteration, and then the weights in the next iteration are updated as follow:

$$W_l^{t+1} = W_l^t + V_l^{t+1} \tag{4.3}$$

where V_l^{t+1} is calculated as

$$V_l^{t+1} = m \times V_l^t - \gamma^t \times \alpha_l \frac{dL}{dW_l} \tag{4.4}$$

α_l is the learning rate of l^{th} layer, m is the momentum due to the previous weight update in the current iteration, and γ is the scheduling rate that decreases learning rate at the completion of each epoch.

$$L = -\frac{1}{N}\sum_i^N ln(p(Y^i/X^i)) \tag{4.5}$$

Here, L denotes the cost function (updated on the basis of stochastic gradient descent algorithm using back propagation for minimizing cost function) over a mini-batch of size N. Here, X^i is the i^{th} training image with respective labels or class-category denoted by Y^i and $p(Y^i/X^i)$ is the probability of X^i belonging to Y^i.

4.2.2 Scratch trained vs. fine/shallow tuned model

If we set $\alpha_l = 0$ for $1{:}\,l$ layers as shown in (4.4) then depending upon the value of l, those entire layer from 1:l is not updated in their weight and hence the weights are transferred as it is there in the final version of trained model. For AlexNet fine-tuned model, learning rate was setup as α_l=0 up to 17th layer i.e. up to first fully connected layer (note: there are 23 layers in AlexNet) whereas for the proposed scratch trained CNN α_l=0.0001 throughout the whole network. This approach of transferring weights from a trained network to other untrained network or classifier model is generally considered as transfer learning.

Pre-trained CNN model available in caffenet [9] library was used for the experiment. First the layer of each model was studied properly and the layers were selected for tuning. In Q1' tuning all the layers except the first quartile participates in training i.e. the weights from first layer to 1st quartile were immediately transferred from pre-trained model to the new tuning model, similarly in Q2' tuning the weights from first layer to 2nd quartile remains constant even after training process and so on for Q3' and Q4' tuning. The overall training accuracy along with validation accuracy and loss for each Q3' training is shown in Figure 6-2 a. to d. Scratch training from all the images was performed in CNN as proposed in Table 2. Similarly, Figure 6-3 shows the comparative bar-diagram of all the process tabulated in Table 6. We performed AD and NC classification task of OASIS MRI scans using various CNN models available and tested the result using Q1', Q2' and Q3' tuning.

Table 2. Proposed CNN layer for Scratch Training for MRI classification

S.N	Layer	Description
1	Image Input	227x227x3 images with 'zerocenter' normalization
2	Convolution	64 5x5x3 convolutions with stride [1 1] and padding [1 1 1 1]
3	Batch Normalization	Batch normalization with 64 channels
4	ReLU	ReLU
5	Max Pooling	2x2 max pooling with stride [2 2] and padding [0 0 0 0]
6	Convolution	32 5x5x64 convolutions with stride [1 1] and padding [1 1 1 1]
7	Batch Normalization	Batch normalization with 32 channels
8	ReLU	ReLU
9	Max Pooling	2x2 max pooling with stride [2 2] and padding [0 0 0 0]
10	Convolution	32 5x5x32 convolutions with stride [1 1] and padding [1 1 1 1]
11	Batch Normalization	Batch normalization with 32 channels
12	ReLU	ReLU
13	Max Pooling	2x2 max pooling with stride [2 2] and padding [0 0 0 0]
14	Convolution	64 5x5x32 convolutions with stride [1 1] and padding [1 1 1 1]
15	Batch Normalization	Batch normalization with 64 channels
16	ReLU	ReLU
17	Max Pooling	2x2 max pooling with stride [2 2] and padding [0 0 0 0]
18	Fully Connected	512 fully connected layer
19	ReLU	ReLU
20	Fully Connected	2 fully connected layer
21	Softmax	softmax
22	Classification	Cross-entropy with classes 'Alzheimer' and 'Normal'

5. Dataset and methodology

5.1 Data I: Segmentation dataset

The dataset consisted of different types of MRI scans with raw, processed as well segmented 3D raw files or analyze format file (.img, .hdr). Mid cross-sectional averaged and co-registered central slice scan images were used that were obtained in the native acquisition space resampled to 1 mm isotropic voxels [13] from 82 subjects (cross-sectional MRI brain scans of dimensions 208 × 176 × 160). Window OS supported software package named 'MRIcon' was used to visualize MRI slices and obtain PNG (Portable Network Graphics) format images from each mid cross-sectional MRI of size 208 × 176 pixels, which represents a single MRI subject. For simplicity we selected MRIs from ID OAS1_0001_MR1 to OAS1_0080_MR1, so finally seventy-six images were exported as training images. The training images were skull stripped, and the corresponding segmented images (the image is already segmented into four parts) of each training images were used as training labels or ground truth. Later, the trained network was used to segment the test MRI images, and the result was compared with the ground truth segmentation. Hence Seventy-six images selected from ID OAS1_0001_MR1 to OAS1_0080_MR1 (excluding four missing MRI) for training was increased by four times performing augmentation i.e. reflection in x-axis, y-axis, rotation in clockwise and anticlockwise 10 degrees. Later testing was performed in six images from ID OAS1_0081_MR1 to OAS1_0087_MR1 for excluding OAS1_0082_MR1 for performance evaluation.

All experiments were simulated in Matlab R2017b on an i3 4160, 4 GB RAM windows desktop. To lessen computation time, the neural network was trained by a single GeForce GTX 1050 Ti GPU using parallel computing.

5.2 Data II: Classification dataset

28 subjects from Healthy and 28 subjects from Alzheimer's patients were selected for study. OASIS provides two types of data: cross-sectional and longitudinal MRI data. In this study, cross-sectional MRIs data is used because my aim to develop an automatic system for detecting AD, which would not require longitudinal data that had been gathered from AD patients over long periods of time. Each subjects MRI axial slice were selected from 71st to 100th slice i.e. 30 slices from every patients. In total 840 images belonging to each class. Representative slice for a typical NC and AD patient MRI is shown in Figure 1. MRIcon [53] software is used as slice extraction tool. Here, we have used limited subjects to each class because, the structure belonging to same classes are almost identical so large number of subject included may cause redundancy and bulkiness in dimension.

All of these experiments were performed using Matlab R2017b, in i3 4160 CPU and NIVIDIA GeForce GTX 1050 Ti 4GB GPU with 4GB RAM windows desktop using Neural Network Toolbox™[54], Statistics and Machine Learning Toolbox™[55], ImageNet pre-trained AlexNet Model from MatConvNet: CNNs[56].

6. Experimental result and discussion

6.1 Result I: Segmentation result

6.1.1 Parameters

To assess the performance of the method, the Dice similarity index (DSI), the Jaccard coefficient (JC), and the mean square error (MSE) of each test image was calculated with reference to the ground truth image available in the same database as discussed in Chapter 5. For comparison, each image was converted into a label image as that of ground truth.

a. Dice and Jaccard Similarity Index

From the experiment, it can be clearly seen that results of high visual quality were obtained, with almost 80% Dice similarity index in each test image.

The Dice similarity coefficient of two sets m and n is defined as

$$Dice(m,n) = 2 \times |intersection(m,n)|/(|m| + |n|) \tag{6.1}$$

where $|m|$ represents the cardinality of the set m and $|n|$ represents the cardinality of the set n.

Similarly, the Jaccard similarity coefficient is defined as

$$Jaccard(m.n) = |intersection(m,n)|/|union(m,n)| \tag{6.2}$$

b. Mean Square Error

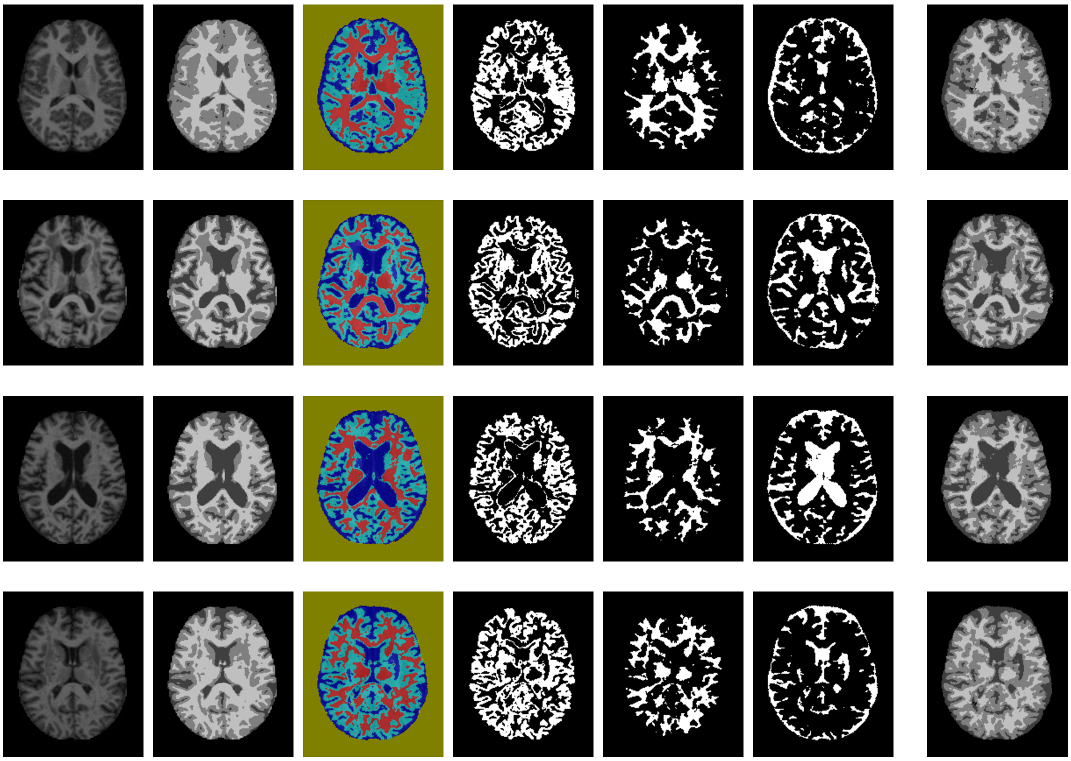
Mean square error (MSE) is defined as

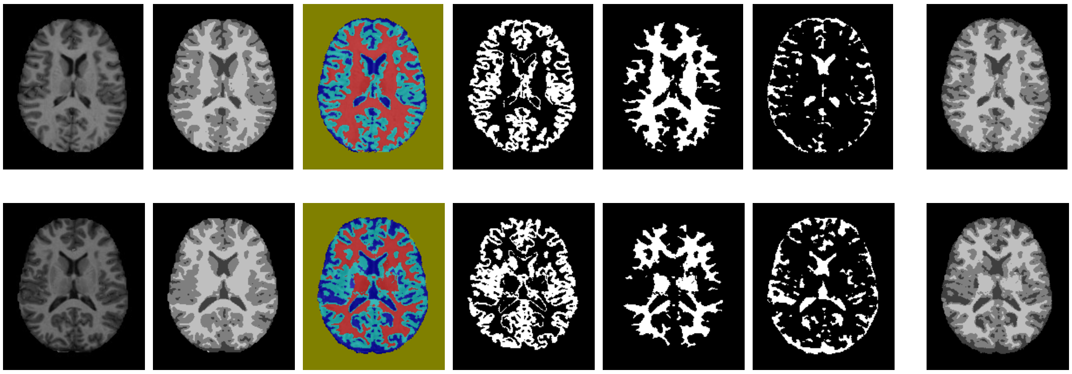
$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (I - I') \tag{6.3}$$

where I and I' stand for the pixel intensity value for the ground truth reference image of size M×N and the simulated image pixel value of the same image size, i.e., M×N, respectively. Both the DSI and the JC index are central parameters for determining how closely the images I and I' are related, and IoU is used for determining how closely they are spatially matched, with no wrong mapping. Similarly, *MSE* was calculated to authenticate the similarity index and the resemblance of the simulated result I' to the ground truth I with minimum loss of information.

c. Training and Testing Accuracy

The overall training accuracy was 91.47 with mean global accuracy 0.91, mean accuracy 0.88248, mean IoU 0.88248, and WeightedIoU 0.84. The intersection over union (IoU) for the best predicted image was approximately 0.8477, whereas IoU for the worst predicted image was approximately 0.625.





(a) Original image

(b) Ground truth

(c) Segmented Image by proposed method (color)

(d) GM part (Binary image)

(e) WM part (Binary image)

(f) CSF part (Binary image)

(g) Segmented Image in gray (c)

Figure 6-1. Original image and ground truth image presented along with other images, as a result of proposed segmentation

6.1.2 Discussion

The results obtained appear satisfactory and visually distinguishable. Figure 6-1 (a)–(g) shows the results of the experiment. The first column (a) shows the extracted MRI test images obtained from the OASIS database, which are skull stripped cross-sectional T1 images; the next column (b) contains the segmented or ground truth image of test image in column (a). The third column (c) shows the main results, which are segmented using the proposed method, i.e., segmentation based on pixel label. The remaining three columns (d), (e), and (f) show the extracted binary image as a classification result of (c). The segmented image (c) is represented by gray level intensity in (g), which is compared with the ground truth to evaluate the Dice similarity coefficient (DSC) of each class, namely GM, WM and CSF. Table 4 presents the performance parameter for each image presented in row (a) of the original image in Figure 6-1.

The computed mean DSC was approximately 80% (highest 84% and lowest 71%) among 6 test images. To compare this result, similar previous approaches for brain image segmentation are presented in Table 3. Zhang et al. 2015 [43] used a patch-wise CNN for private data of 10 healthy infants, and Nie et al. 2016 [49] used semantic approach for the same type of data. Proposed approach was superior to those by de Brébisson et al. 2015 [44] and Moeskops et al. 2016 [45] in terms of DSC, but the dataset used here is OASIS mid cross-sectional T1 MRI 2D images instead of MICCAI 2012 atlas.

Table 3- Comparison of deep learning approaches for brain structure segmentation

Authors	CNN Style	Dimension	Accuracy	Data
Zhang et al. 2015 [43]	Patch-wise	2D	DSC 83.5% (CSF), 85.2% (GM), 86.4% (WM)	Private data (10 healthy infants)
Nie et al. 2016 [49]	Semantic-pixel wise	2D	DSC 85.5% (CSF), 87.3% (GM), 88.7% (WM)	Private data (10 healthy infants)
de Brebisson et al. 2015 [44]	Patch-wise	2D/3D	Overall DSC 72.5% +/- 16.3%	MICCAI 2012-multi-atlas labeling
Moeskops et al. 2016 [45]	Patch-wise	2D/3D	Overall DSC 73.53%	MICCAI 2012-multi-atlas labeling
Proposed Method	Pixel-label Semantic	2D	DSC 72.2% (CSF), 74.6% (GM),81.9% (WM)	OASIS cross-sectional MRI

Table 4. Comparison of performance parameters for each result image, Figure 6-1 column (g), with respective ground truth image, Figure 6-1 column (b)

Test image ID	Parameter	CSF Part	Gray Part	White Part	Mean Value
OAS1_0081_MR1	Dice Similarity	0.54	0.75	0.85	0.71
	Jaccard Similarity	0.37	0.59	0.74	0.57
	Mean Square Error	-	-	-	29.47
OAS1_0083_MR1	Dice Similarity	0.84	0.75	0.79	0.80
	Jaccard Similarity	0.73	0.60	0.66	0.66
	Mean Square Error	-	-	-	19.32
OAS1_0084_MR1	Dice Similarity	0.85	0.71	0.78	0.78
	Jaccard Similarity	0.74	0.55	0.64	0.64
	Mean Square Error	-	-	-	25.02
OAS1_0085_MR1	Dice Similarity	0.72	0.67	0.73	0.71
	Jaccard Similarity	0.56	0.51	0.57	0.55
	Mean Square Error	-	-	-	32.52
OAS1_0086_MR1	Dice Similarity	0.74	0.85	0.92	0.84
	Jaccard Similarity	0.59	0.74	0.85	0.73
	Mean Square Error	-	-	-	9.52
OAS1_0087_MR1	Dice Similarity	0.64	0.75	0.85	0.74
	Jaccard Similarity	0.47	0.60	0.74	0.60
	Mean Square Error	-	-	-	27.58

6.2. Result II: Classification performance

6.2.1 Parameters

a. Training validation and testing accuracy

For performance analysis of the experiment conducted for classification, we trained, validated and tested the result in the ration of 6:2:2 on random selection basis. Then resulting testing results are reported and analyzed. Training accuracy determines how well the training images has well converged to form a stable network. The final weights induced in each filter can accurately determines the trainee set with almost 98-99% accuracy, whereas on the other hand validation set tests the accuracy after each epoch and confirms if the training is under fitting or over fitting the new(validation) data. Higher validation suggests over fitting and lower suggests under fitting. Once the training is completed and validation and training goes on parallel, we can test the new unknown data i.e. test set. This test set is completely new and unknown labels, which is predicted by the network. It was found testing accuracy to be highest in Q1' tuning i.e. less the weights transferred from pretrained to untrained, more is the accuracy. Whereas the scratch trained network could classify the test set with the highest 98.51% accuracy. Detail of testing accuracy for all models and tuning process is shown in Table 6 along with Cohen-Kappa and training time.

b. Cohen-Kappa value

In very simple term Cohen's Kappa (CK) measures the agreement between the two classes. So, higher the value more will be possibility of perfect prediction. CK value to 1 means a perfect prediction i.e. the predicted result perfectly matches the ground truth label. Detail is shown in Table 5.

Table 5. Cohen-Kappa calculation method

	Predicted class		
True class	TP	FN	po = the relative observed agreement among prediction vs. true class.
	FP	TN	pe = the hypothetical probability of chance agreement
	po=TP+TN		$pe = p1.p1' \times p2.p2'$
	p1=TP+FN		$k = \frac{po - pe}{1 - pe}$
	p2=FP+TN		Cohen kappa $= k$
	p1'=TP+FP		
	p2'=FN+TN		

c. Training time

It denotes the time required from the beginning of epoch 1 to the end of training until the final epoch is met or the convergence is achieved, whichever is the faster. The denoted time is GPU time here as GPU is used for training and feature extraction.

Figure 6-2 (a). Alexnet training Q3' @Initial Learning rate= 10^{-4}, Learning drops by 50% every epoch, Minibatch=8, epoch 5

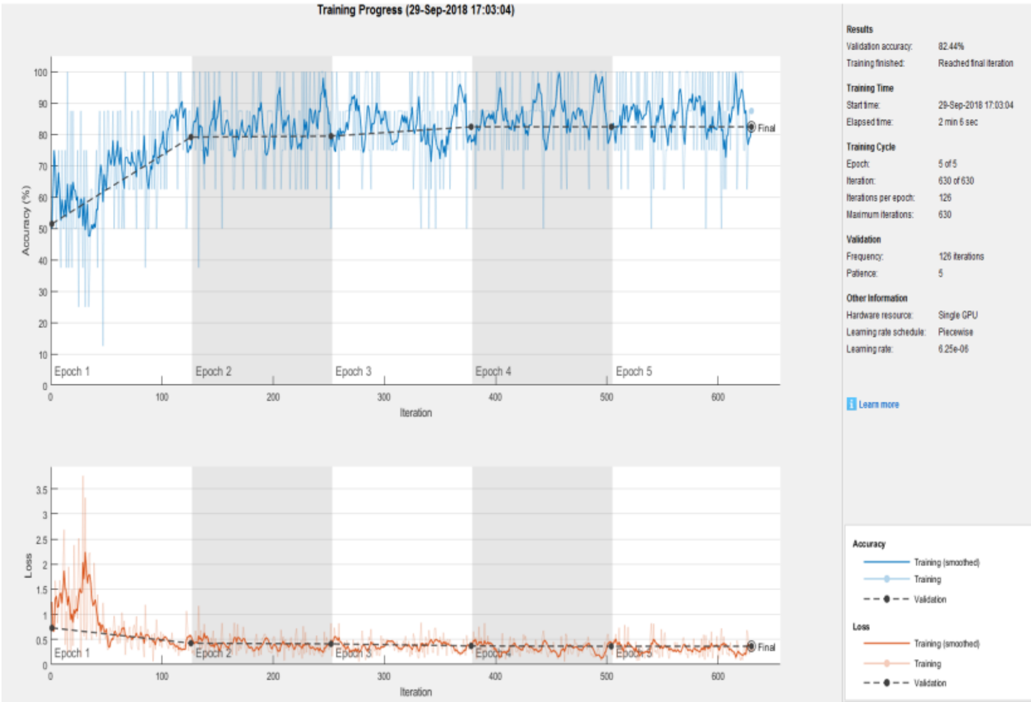


Figure 6-2(b). Googlenet training Q3' @Initial Learning rate= 10^{-4}, Learning drops by 50% every epoch, Minibatch=8, epoch 5

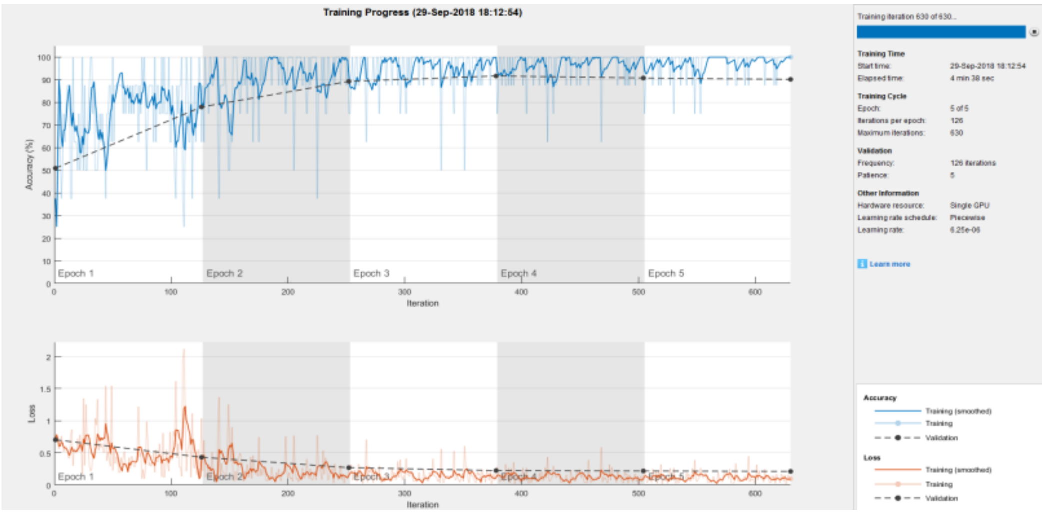


Figure 6-2(c). Resnet50 training Q3' @Initial Learning rate= 10^{-4}, Learning drops by 50% every epoch, Minibatch=8, epoch 5

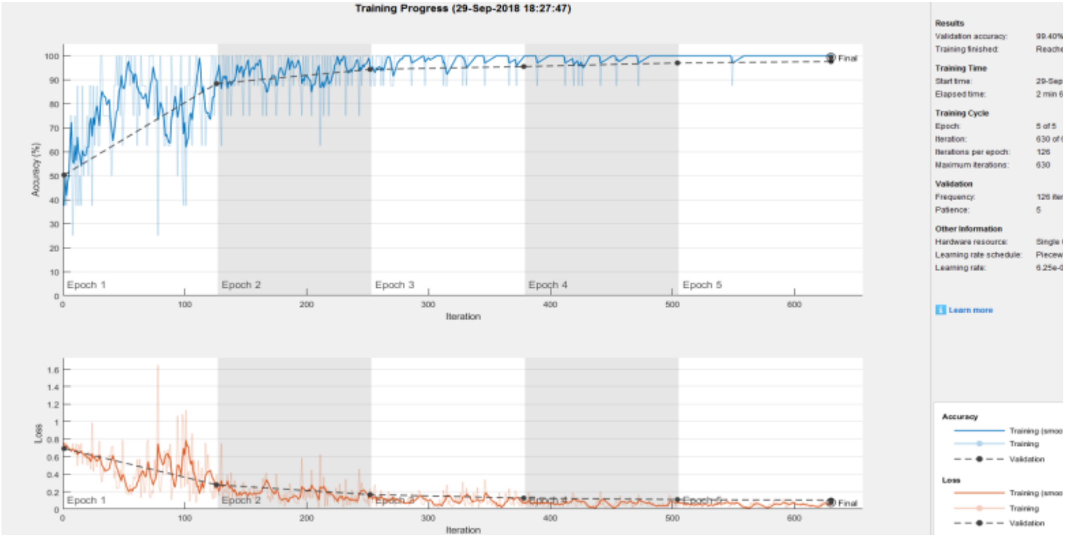


Figure 6-2(d). Scratch training @Initial Learning rate = 10^{-4}, Learning drops by 50% every epoch, Minibatch=8, epoch 5

6.2.2 Discussion

Three pre-trained models and one scratch trained model were tested on dataset of NC and AD patients MRI images. First model AlexNet has 25 layers, Google Net has 144 layers and Resnet50 has 177 layers (counting each unit as a single layer). They were tuned in different ways as discussed above and the obtained result is presented in Table 6. The result shows that the performance is better when majorities of layers are tuned i.e. during Q1' tuning but at the same time, the training period for Q1' is comparatively higher than Q2', Q3' and Q4' in all cases. On the basis of this result we have designed 22 layers CNN model similar in architecture with that of AlexNet [1] but instead of using cross-channel normalization, batch normalization was implemented and the size of filters were reduced eventually making it only 22 layered CNN. As from Figure 6-2(d), we can see that the scratch trained CNN model training and validation process is smoother and faster than other pre-trained model training process.

Table 6. 0.6:0.2:0.2=Training: Validation: Testing @Initial Learning rate= 10-4, Learning rate drops by 50% every epoch, Minibatch=8, Maxepoch 5, Maximum value for Accuracy and Cohen-Kappa value is 1.

Models	Total unit	Tuning process	Layers weights transferred	Accuracy	Cohen-Kappa value	Training Time (min:sec) for 5 epoch
AlexNet[1]	25 (last 3 layers being FCL,SM,CL)	Q1' Tuning	1:5	0.9464	0.8929	1:25
		Q2' Tuning	1:11	0.9315	0.8631	1:17
		Q3' Tuning	1:16	0.9196	0.8393	1:11
		Q4' Tuning	1:22	0.8482	0.6964	0:51
GoogLeNet [3]	144 (last 3 layers being FCL,SM,CL)	Q1' Tuning	1:33	0.8899	0.7798	3:21
		Q2' Tuning	1:62	0.8839	0.7679	3:13
		Q3' Tuning	1:104	0.8274	0.6548	2:06
		Q4' Tuning	1:141	0.7917	0.5833	1:40
Resnet50 [2]	177 (last 3 layers being FCL,SM,CL)	Q1' Tuning	1:43	0.9435	0.8869	6:59
		Q2' Tuning	1:91	0.9375	0.875	6:35
		Q3' Tuning	1:134	0.9256	0.8512	4:38
		Q4' Tuning	1:174	0.9167	0.8333	2:54
Scratch Trained	22	-	-	0.9851	0.9702	2:06

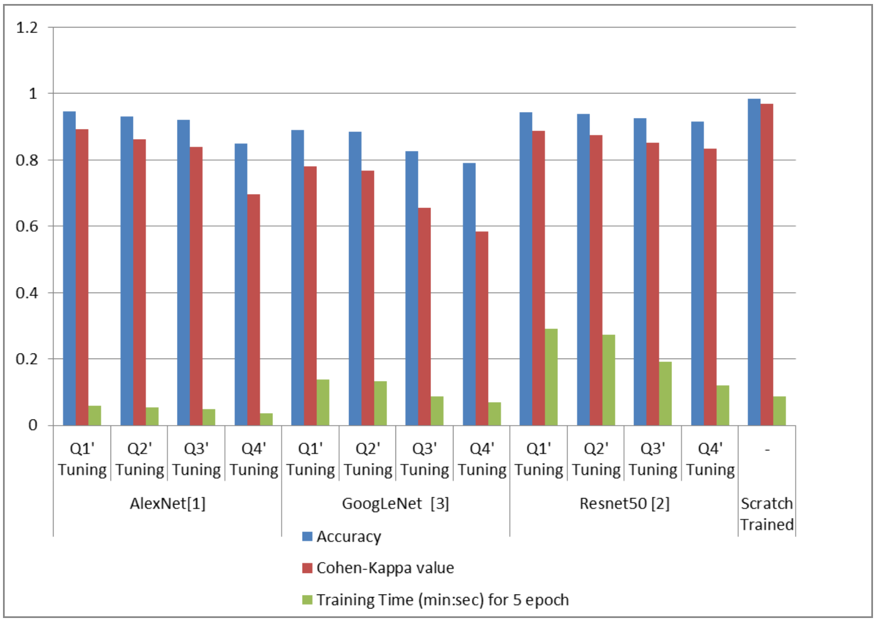


Figure 6-3. Bar diagram for result of experiment and test accuracy, Cohen-Kappa value and training time compared for all the process as in Table (6)

7. Conclusion

In conclusion, deep learning technique was successfully implemented for MRI segmentation and classification with convincing results. Specifically, in segmentation, closely related brain MRI images could be segmented on pixel-label basis using encoder-decoder network architecture like SegNet layer generally used in semantic segmentation of outdoor scene which suggests that, with necessary modification and simplified architecture, deep neural network can be effective in medical MRI segmentation as like natural outdoor images.

Also the CNN models result was studied on classification on medical MRI images and also a modest version of CNN was proposed. This model is a like a closed box system where, the training, validation and testing is performed on the same sets of data in random manner. So proper investigation needs to done for the generalization of model. From this experiment, I have been able to conclude following points:

- Increasing the depth of learning model doesn't not always result in good performance
- Training time increases with increase in number of layers to tune.
- Smaller and optimal model can be designed for optimal performance on the basis of application
- Non-medical images trained model can be used for transferring weights for training Medical Image classification model

Hence the CNN model implication for classification between healthy and Alzheimer's patients based on important MRI images using transfer learning and raw training from scratches have been successfully tested.

Acknowledgement

First and foremost, I would like to express my sincere gratitude to my advisor, Prof. Goo-Rak Kwon for his encouragement, support, guidance and valuable knowledge shared throughout the development of this thesis. Under his guidance, the exposure I have gained during my research would be a valuable treasure of my life.

At the same time I would like to appreciate the support of all my professors, seniors, juniors, lab mates, colleagues and friends who helped me in various ways during my study in Chosun.

And above all it's the blessing of my parents and love of my family without which I can achieve nothing in my life.

Bijen

REFERENCES

[1] L. Kapoor, S. Thakur, "A survey on brain tumor detection using image processing techniques," In: 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence; Noida, 12-13 Jan 2017, IEEE.

[2] N. N. Gopal, M. Karnan, "Diagnose brain tumor through MRI using image processing clustering algorithms such as Fuzzy C Means along with intelligent optimization techniques," In: 2010 IEEE International Conference on Computational Intelligence and Computing Research; Coimbatore, 28-29 Dec. 2010, IEEE.

[3] M. Kamber, R. Shinghal, D. L. Collins, G. S. Francis, A. C. Evans, "Model based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images," IEEE Transactions in Medical Imaging. 1995; 14(3):442–453.

[4] A. Zijdenbos, R. Forghani, A. Evans, "Automatic quantification of MS lesions in 3D MRI brain data sets: Validation of INSECT," Medical Image Computing and Computer Assisted Intervention —MICCAI'98, Lecture Notes in Computer Science, 1998; 439-448.

[5] Y. Zhang, S. Wang, "Detection of Alzheimer's disease by displacement field and machine learning," PeerJ. 2015.

[6] Y. Zhang, S. Wang, G. Ji, Z. Dong, "An MR brain images classifier system via particle swarm optimization and kernel support vector machine," The Scientific World Journal. 2013.

[7] S. Alam, G.R. Kwon, "Alzheimer disease classification using KPCA, LDA, and multi-kernel learning SVM," Int J Imaging Syst Technol. 2017; 27(2):133–143.

[8] Y. Ding, C. Zhang, T. Lan, Z. Qin, X. Zhang, W. Wang, "Classification of Alzheimer's disease based on the combination of morphometric feature and texture feature," In: 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); Washington DC, 9-12 Nov. 2015, IEEE.

[9] Y. Zhang, S. Wang, Z. Dong, "Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel uspport vector machine decision tree," Progress In Electromagnetics Research. 2014; 144:171-184.

[10] K. Sirinukunwattana, S. E. A. Raza, Y. W. Tsang, D. R. J. Snead, I. A. Cree, N. M. Rajpoot, "Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images," IEEE Transactions on Medical Imaging. 2016; 35(5):1196-1206.

[11] I. Despotović, B. Goossens, and W. Philips, "MRI Segmentation of the Human Brain: Challenges, Methods, and Applications," Computational and Mathematical Methods in Medicine, vol. 2015, Article ID 450341, 23 pages, https://doi.org/10.1155/2015/450341, 2015.

[12] M. Tabaton, P. Odetti, S. Cammarata et al., "Artificial neural networks identify the predictive values of risk factors on the conversion of amnestic mild cognitive Impairment," Journal of Alzheimer's disease. 2010; 19(3):1035–1040.

[13] D. Jha, J. I. Kim, and G. R. Kwon, "Diagnosis of Alzheimer's Disease Using Dual-Tree Complex Wavelet Transform, PCA, and Feed-Forward Neural Network," Journal of Healthcare Engineering, vol. 2017, Article ID 9060124, 13 pages, doi:10.1155/2017/9060124, 2017.

[14] J. B. MacQueen (1967), "Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability," Berkeley, University of California Press, 1:281-297

[15] Dempster, Arthur P, Nan M. Laird, and Donald B.Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," Paper presented at the Royal Statistical Society at a meeting organized by the Research Section, December 8, 1976.

[16] Yang, J., Jiang, Y. G., Hauptmann, A. G., & Ngo, C. W. (2007, September), "Evaluating bag-of-visual-words representations in scene classification," In Proceedings of the international workshop on Workshop on multimedia information retrieval (pp. 197-206). ACM.

[17] S. Alam, G.R. Kwon, J. I. Kim, and C. S. Park, " Twin SVM-based classification of Alzheimer's disease using complex dual-tree wavelet principal coefficients and LDA," Journal of healthcare engineering, 2017.

[18] M. Maitra, A. Chatterjee, "A Slantlet transform based intelligent system for magnetic resonance brain image classification," Biomedical Signal Processing and Control. 2006; 1(4):299-306.

[19] J. Deng, W. Dong, R. Socher, L. J. Li, L. & Fei-Fei, “ImageNet: A large-scale hierarchical image database”, In 2009 IEEE conference on computer vision and pattern recognition, June 2009 (pp. 248-255), IEEE.
[20] A. Krizhevsky, V. Nair and G. Hinton, “The CIFAR-10 dataset,” online: http://www. cs. toronto. edu/kriz/cifar. html, 4.
[21] X. Xia, C. Xu, B. Nan, “Inception-v3 for flower classification. In2017 2nd International Conference on Image”, Vision and Computing (ICIVC) 2017 Jun 2 (pp. 783-787), IEEE.
[22] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation,” IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017; 39(12):2481–2495.
[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks”, Proc. NIPS, 2012; 1097–1105.
[24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2015; 1–9.
[25] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition” arXiv:1512.03385, 2015.
[26] Girshick, Ross, “Fast r-cnn,” In Proceedings of the IEEE international conference on computer vision, pp. 1440-1448. 2015.
[27] Y. Zhang, S. Wang, G. Ji, and Z. Dong “Exponential wavelet iterative shrinkage thresholding algorithm with random shift for compressed sensing magnetic resonance imaging,” IEEJ Transactions on Electrical and Electronic Engineering. 2015; 10(1):116-117.
[28] Z. Dong, Y. Zhang, F. Liu, Y. Duan, A. Kangarlu, B. S. Peterson, “Improving the spectral resolution and spectral fitting of 1H MRSI data from human calf muscle by the spread technique,” NMR in Biomedicine. 2014; 27(11):1325-1332.
[29] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016. http://www.deeplearningbook.org.
[30] K. He, X. Zhang, S. Ren, J. Sun, “Delving Deep into Rectifiers: Surpassing Human level Performance on Image Net Classification,” 2015; 1026–103

[31] Y. Zhang, S. Wang, Z. Dong, “Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel uspport vector machine decision tree. Progress,” Electromagnetics Research. 2014; 144:171-184.
[32] S. Chaplot, L. M. Patnaik, N. R. Jagannathan, “Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network,” Biomedical Signal Processing and Control. 2006; 1(1):86-92.
[33] El-Dahshan, E. S. A., Hosny, T., & Salem, A. B. M, “Hybrid intelligent techniques for MRI brain images classification,” Digital Signal Processing, 2010; 20(2), 433-441
[34] Y. Zhang, L. Wu, S. Wang, “Magnetic resonance brain image classification by an improved artificial bee colony algorithm,” Progress In Electromagnetics Research. 2011; 116:65-79.
[35] D. Jha, J. I. Kim, G. R. Kwon, “Diagnosis of Alzheimer’s Disease Using Dual-Tree Complex Wavelet Transform, PCA, and Feed-Forward Neural Network,” Journal of Healthcare Engineering. 2017.
[36] Neffati, S., Ben Abdellafou, K., Jaffel, I., Taouali, O., & Bouzrara, K, “An improved machine learning technique based on downsized KPCA for Alzheimer's disease classification,” International Journal of Imaging Systems and Technology. 2018.
[37] Lo, S.-C., Lou, S.-L., Lin, J.-S., Freedman, M.T., Chien, M.V., Mun, S.K, “Artificial convolution neural network techniques and applications for lung nodule detection,” IEEE Trans. Med. Imaging 14.1995; 711–718.
[38] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P, “Gradient-based learning applied to document recognition,” Proc. IEEE 86, 1998; 2278–2324.
[39] N. Tajbakhsh, J. Shin, S. Gurudu, C. Kendall, M. Gotway, and J. Liang, “Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?” IEEE Transactions on Medical Imaging. 2016; 35(5):1299-1312.
[40] H. Shin, H. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, and J. Yao, “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics, and Transfer Learning,” IEEE Transactions on Medical Imaging. 2016; 35(5):1285-1298.

[41] H. R. Roth, L. Lu, J. Liu, J. Yao, A.Seff, K. Cherry, L. Kim, and Ronald M. Summers, "Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation," IEEE Transactions on Medical Imaging. 2016; 35(5):1170-1181.

[42] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: An astounding baseline for recognition," Proc.IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, pp. 512-519, 2014.

[43] W. Zhang et al., "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," NeuroImage. 2015; 108:214–224.

[44] A. de Brébisson and G. Montana, "Deep neural networks for anatomical brain segmentation," in CVPR Bioimage Comput. Workshop, 2015.

[45] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," IEEE Trans. Med. Imaging. 2016; 35(5):1252–1261.

[46] G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R.R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors". arXiv preprint arXiv:1207.0580, 2012

[47] M. Lin, Q. Chen, and S. Yan, "Network in network".

[48] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[49] D. Nie, L. Wang, Y. Gao and D. Sken, "Fully convolutional networks for multi-modality isointense infant brain image segmentation," 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, pp. 1342-1345, 2016.

[49] C. Liang-Chieh, G. Papandreou, I. Kokkinos, K. Murphy, and A.Yuille, "Semantic image Segmentation with deep convolutional nets and fully connected CRFs," in Proc. Int. Conf. Learn. Representations, 2015.

[50] R. Socher, C. C. Lin, C. Manning, and A. Y. Ng, "Parsing natural scenes and natural language with recursive neural networks," in ICML, pp. 129–136, 2011.

[51] C. Liang-Chieh, G. Papandreou, I. Kokkinos, K. Murphy, and A.Yuille, “Semantic image segmentation with deep convolutional nets and fully connected crfs,” in ICLR, 2015.

[52] S. Hong, H. Noh, and B. Han, “Decoupled deep neural network for semi supervised semantic segmentation,” NIPS. 2015; 1495–1503.

[53]https://www.nitrc.org/projects/mricron

[54]https://www.mathworks.com/products/-neuralnetwork.html?s_tid=FX_PR_info visited on 7th April 2018

[55]https://www.mathworks.com/products/statistics.html visited on 7th April 2018

[56]https://www.vlfeat.org/matconvnet/models/beta16/imagenet-caffe-alex.mat visited on 7th April 2018

[57] J. L. Tanabe, D. Amend, N. Schuff, V. DiSclafani, F. Ezekiel, D. Norman, G. Fein, and M. W. Weiner, “Tissue segmentation of the brain in Alzheimer disease”, American Journal of Neuroradiology. 1997; 18(1):115-123.

[58]https://surfer.nmr.mgh.harvard.edu/fswiki/FreeSurferWiki visited 4th June 2019

LIST OF PUBLICATIONS

Bijen Khagi, Goo Rak Kwon, Ramesh Lama: *Comparative analysis of Alzheimer's disease classification by CDR level using CNN, feature selection, and machine-learning techniques*. International Journal of Imaging Systems and Technology 03/2019; DOI:10.1002/ima.22316

Bijen Khagi, Goo-Rak Kwon: *Pixel-Label-Based Segmentation of Cross-Sectional Brain MRI Using Simplified SegNet Architecture-Based CNN*. Journal of Healthcare Engineering 10/2018; 2018:1-8., DOI:10.1155/2018/3640705

Bijen Khagi, Goo-Rak Kwon: *Brain MRI Segmentation using Simple Linear Iterative Clustering Based Superpixel approach on conventional Clustering algorithms*.

Bijen Khagi, Goo-Rak Kwon: *VBM based Alzheimer's disease classification from Region of Interest of T1 MRI with supportive Gaussian Smoothing and Bayesian regularized Neural Network.* Computer Methods and Programs in Biomedicine of Elsevier (under-review)

Bijen Khagi, Goo-Rak Kwon: *CNN Models Performance Analysis on MRI images of OASIS dataset for distinction between Healthy and Alzheimer patient.* IEIE Transactions on Smart Processing and computing (under-review)

Conference Proceedings

Bijen Khagi, Goo-Rak Kwon: *DCT-convolution based s-MRI detection of Alzheimer diseases patient*. 31th Workshop on Image Processing and Image Understanding (IPIU 2019), Maison Grad Jeju; 02/2019

Bijen Khagi, Chung Ghiu Lee, Goo-Rak Kwon: *Alzheimer's disease Classification from Brain MRI based on transfer learning from CNN*. 2018 11th Biomedical Engineering International Conference (BMEiCON), Chiang Mai, Thailand, IEEE Xplore: 14 January 2019; 01/2019, DOI:10.1109/BMEiCON.2018.8609974

Bijen Khagi, Goo-Rak Kwon: *Alzheimer's disease detection from Brain MRI using Feature extraction from CNN network and classification by KNN*. Proceedings of the 2018 Year of Korea Information and Communications Society Conference Summer General; 06/2018