MULTI CENTER BRAIN IMAGING CLASSIFICATION USING A NOVEL 3D CNN APPROACH

Seminar Report

Submitted in partial fulfillment of the requirements for the award of degree of

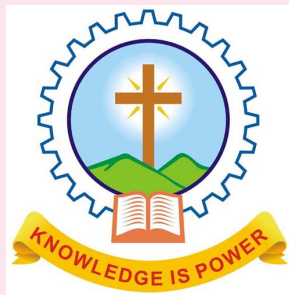**BACHELOR OF TECHNOLOGY**

In

COMPUTER SCIENCE AND ENGINEERING

of

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Submitted By

ROHITH R NAIR

Department of Computer Science & Engineering
Mar Athanasius College Of Engineering Kothamangalam

MULTI CENTER BRAIN IMAGING CLASSIFICATION USING A NOVEL 3D CNN APPROACH

Seminar Report

*Submitted in partial fulfillment of the requirements for
the award of degree of*
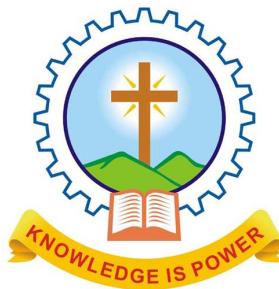
BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

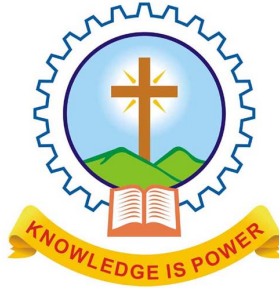*of*

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Submitted By

ROHITH R NAIR

Department of Computer Science & Engineering
Mar Athanasius College Of Engineering Kothamangalam

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
MAR ATHANASIUS COLLEGE OF ENGINEERING
KOTHAMANGALAM

CERTIFICATE

This is to certify that the report entitled **Multi center Brain Imaging Classification Using A Novel 3D CNN Approach** *submitted by* **Mr. ROHITH R NAIR**, *Reg.No.***MAC15CS050** *towards partial fulfillment of the requirement for the award of Degree of Bachelor of Technology in Computer science and Engineering from APJ Abdul Kalam Technological University for December 2018 is a bonafide record of the seminar carried out by him under our supervision and guidance.*

....
Prof. Joby George
Faculty Guide

....
Prof. Neethu Subash
Faculty Guide

....
Dr. Surekha Mariam Varghese
Head of the Department

Date:

Dept. Seal

ACKNOWLEDGEMENT

ABSTRACT

Using a Novel 3D CNN Approach with the development of brain imaging technology, increasing amounts of magnetic resonance imaging (MRI) data are being acquired. Deep learning technology has shown the powerful ability to solve the classification problem while it has not been widely used in brain imaging classification. The proposed novel 3D deep adding neural network to classify samples from the largest datasets in the brain imaging field collected from centers. The proposed method utilizes multiple convolutional layers to extract gradient information in different orientation. Which combines spatial information at two scales via the adding operation. The cross-site classification results prove that the proposed method is robust when training on a dataset and testing on another dataset. This is the first work to classify neuroimaging data on such a large scale from multiple centers. With its improved performance in classification and transferable program codes, the proposed method can potentially be used in intelligent medical treatment strategies and clinical practices.

Contents

List of Figures

LIST OF ABBREVIATION

3DDANet	3D Deep Adiing Neural Network
LSTM	Long Sort Term Memory
CNN	Convolutional Neural Network
ADHD	Attention Deficit Hyperactivity Disorder
MRI	Magnetic Resonance Imaging
GM	Grey Matter
MNI	Montral Neurological Institute
CSF	Cerebrospinal Fluid
WM	White Matter

Introduction

Brain imaging data,including magnetic resonance imaging (MRI) data, have been widely used in neural science and clinical applications. However, with the development of brain science and neuroimaging technology, increasing amounts of neuroimaging data are being acquired, and the number of samples is unprecedented. Traditional statistical analysis methods based on single sites and small samples are becoming problematic for addressing big data from multiple centers. Large individual differences and dif- ferences between centers will make the previous methods ineffective.

Based on the development of the deep learning algo- rithm, some researchers have introduced models into the neuroimaging analysis field, with some model architectures being based on three-dimensional (3D) convolutional neural networks (CNNs) Recently, deep learning has underwent unprecedented development and soared in popularity. Numerous companies and research groups have successively published all kinds of deep learning models to solve image classification problems in rather large datasets (such as ImageNet, which has over 100,000 images). These models (such as AlexNet,VGGNet,GoogleInceptionNet,ResNet, and GAN and so on) have been very successful at image classification. Some models, such as long sort term memory (LSTM), can effectively address video recognition and time series classifications.

We think that deep learning will have many more successes in the near future because it requires very little engineering by hand, so it can easily take advantage of increases in the amount of available computation data.

Deeper neural networks are more difficult to train. We present a residual learning framework to ease the training of networks that are substantially deeper than those used previously. We explicitly reformulate the layers as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions. We provide comprehensive empirical evidence showing that these residual networks are easier to optimize, and can gain accuracy from considerably increased depth.

In this paper, we use the subjects gender label as the predicting label because gender

is the golden standard in the neuroimaging field and does not include subjective factors. The proposed method was tested with a standard 5-fold cross-validation strategy on over 6000 MRI samples of healthy people collected from more than 61 sites in six datasets. All MRI data were preprocessed in a uniform pipeline to maximally eradicate differences between scanners. Furthermore, we also use deconvolutional computation to show differences between groups to demonstrate the effectiveness of the proposed method

Existing Methods

Recently, deep learning has underwent unprecedented de- velopment and soared in popularity. Numerous com- panies and research groups have successively published all kinds of deep learning models to solve image classifica- tion problems in rather large datasets (such as ImageNet [8], which has over 100,000 images). These models (such as AlexNet,VGGNet, GoogleInceptionNet, ResNet, and GAN and so on) have been very successful at image classification. Some models, such as long sort term memory (LSTM), can effectively address video recognition and time series classifications. Further- more, some deep learning approaches have been used in medical practice. However, all the models are based on two-dimensional (2D) images or 2D images with time series. Because the human brain has a special spatial structure that is strongly related to brain functions and activities, ignoring the spatial structure of the human brain is not advisable.

Dolz et al investigated a 3D full CNN for subcortical brain structure segmentation in MRI. They tested the model on the ABIDE dataset, which includes 1112 subjects, and yielded segmentations that were highly consistent with a standard atlas-based approach. Kawahara et al. proposed using CNNs for brain networks to predict clinical neurodevelopmental outcomes, while Wachinger et al. proposed a deep CNN that included three con- volutional layers with pooling, batch normalization (BN), non-linearities and fully connected layers with dropout for segmenting neuroanatomy. Cole et al utilized a CNN to predict brain age using raw T1 MRI data on a large dataset of healthy adults (2001 subjects). Chen et al proposed using VoxResNet, comprising 25 layers, for segmenting key brain tissues from 3D magnetic resonance images, and this approach was awarded first place in the challenge out of 37 competitors. Ktena et al proposed to learn a graph similarity metric using a siamese graph convolutional neural network (s-GCN) to classify neuroimaging on brain connec- tivity networks. They tested the proposed framework on the ABIDE database and the UK Biobank, and this approach performed better than traditional methods. Zou et al proposed a 3D CNN classification approach to automatic

3

diagnose attention deficit hyperactivity disorder (ADHD) and achieved the accuracy of 69.15%, which outperformed reported classifiers in the literature.

3D deep learning approaches were focused on brain or neuroanatomy segmentation and brain maturity analyses, and few 3D CNN models for multi-center MRI data classifications exist on a large scale

The CNN was first proposed by Lecun et al in 1998 and has been very actively investigated, especially recently. As mentioned in the Introduction , several different network architectures have been proposed, which have demonstrated state-of-the-art performances in many computer vision and speech recognition tasks. We hypothesize that a CNN can provide the appropriate architecture to infer imag- ing features from processed brain MRI scans to predict a subjects gender. The 2D residual neural network (ResNet), which eases network training, was proposed by Kaiming He et al in 2015, and they won 1st place in the ILSVRC 2015 classification task. Some architectures based on the 3D CNN model for predicting brain maturity, predicting neurodevelopment, Alzheimers disease classifications, brain segmentation and skull stripping. With the inspiration from ResNet and previous 3D CNN architectures, we propose the 3D DANet model for neuroimaging classification

we selected some typical traditional classifica- tion methods and deep learning approach architectures. In our previous study , we proposed a 3D descriptor (3D weighted histogram of gradient orientation, 3D WHGO) and selected the PCA+SVM approach, a typical multivariable pattern analysis (MVPA) approach, as the comparison model. In this paper, we also compared our proposed method with the 3D WHGO method and the PCA+SVM approach. For the deep learning approach, we compare our method with the typical 3D CNN and 3D ResNet.

Proposed Method

Traditional 3D deep learning approaches were focused on brain or neuroanatomy segmentation and brain maturity analyses, and few 3D CNN models for multi-center MRI data classifications exist on a large scale. The proposed method using a 3D deep learning model, termed 3D deep adding neural network (3D DANet), to solve the human brain MRI classification problem.

This model contains two parts; the first part includes one 3D convolutional layer and a 3D max pooling layer, while the second part includes four 3D convolutional layers and a 3D max pooling layer. Then, the two 3D pooling layers are added together as the input for the final full connection layer. Finally, we use softmax as the classifier to generate predictive labels.In the proposed solution, we use the subjects gender label as the predicting label because gender is the golden standard in the neuroimaging field and does not include subjective factors. Furthermore, we also use deconvolutional computation to show differences between groups to demonstrate the effectiveness of the proposed method.

3.1 Datasets

T1-weighted MRI scans comprised all the neuroimaging data used in the study, and all the scans were from six subsets collected from more than sixty sites or scanners. Details regarding the participants in the specific samples and the respective acquisition parameters used are outlined below. We used gender as the predicting label, and the females and males included in each dataset are also outlined below.
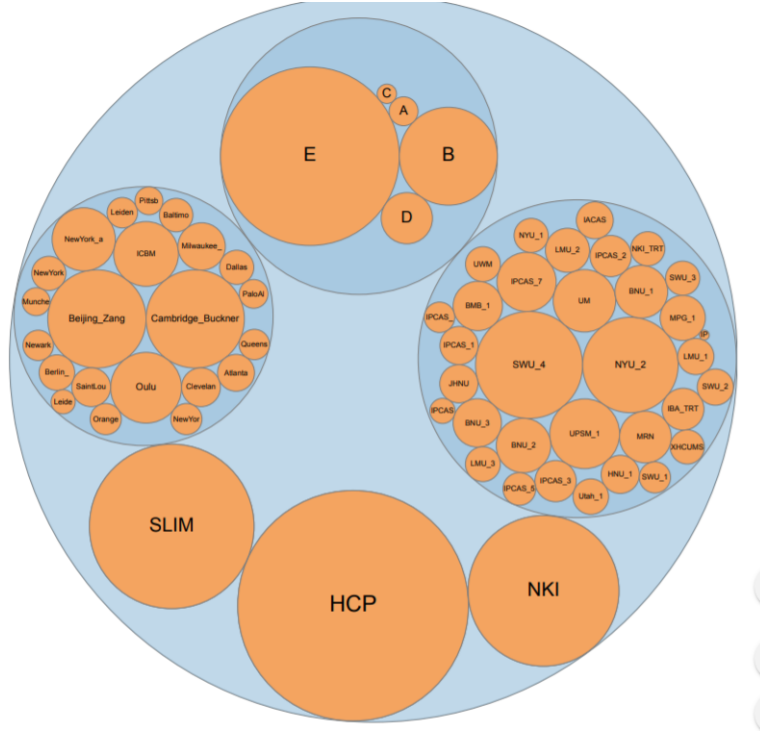
Fig. 3.1: Venn diagram of the collected data.

- 1000 Functional Connectomes Project

  The 1000 Functional Connectomes Project (FCP) publishes the unrestricted public release of 1200 functional MRI (Rs-fMRI) datasets independently collected from 33 sites. All datasets were generously donated by the principal investigators from the member sites to provide the broader imaging community complete access to a large-scale functional imaging dataset. The age and sex of the participants as well as the imaging center information are provided for each of the datasets. As we used gender as the predicting label to test the proposed model, some of centers that did not provide gender information were removed.

- Brain Genomics Superstruct Project

  The goal of the Brain Genomics Superstruct Project (GSP) is to enable large-scale exploration of the links between brain function, behavior, and ultimately, genetic variation. MRI data from 1570 unique subjects are included in five sub-datasets, with each dataset

containing high-resolution anatomical T1-weighted structural MRI scans. Imaging data, provided in the NIfTI format, were obtained from 905 female and 665 male subjects.

- Human Connectome Project

  The Human Connectome Project (HCP) is an ambitious 5- year effort to characterize brain connectivity and function and their variability in healthy adults to study a population of 1200 subjects including T1- and T2-weighted MRI structural neuroimaging data. The HCP samples are described in further detail in [23]. All the HCP samples used herein were down- loaded from ConnectomDB (https://db.humanconnectome. org) [24]. At the time of this manuscript submission, the HCP had released T1-weighted MRI data from 1090 subjects, which included 592 females and 498 males. We selected all the raw T1-weighted MRI data as the input for our preprocessing pipeline.

- Nathan Kline Institute

  The enhanced Nathan Kline Institute-Rockland Sample is deeply phenotyped, community-ascertained, lifespan sample comprising advanced neuroimaging and genetics data. We collected 461 healthy participants from the NKI-RS database, which included 257 females and 204 males. All approvals and procedures for collection and data sharing were approved by the NKI institutional review board, and each participant were provided written informed consent. MRI data were acquired on a 3.0 T SIMENS Trio scanner. For each subject, high-resolution T1-weighted images were acquired using the magnetization-prepared rapid gradient echo (MPRAGE) sequence.

- Consortium for Reliability and Reproducibility

  The Consortium for Reliability and Reproducibility aims to create an open science resource for the imaging community that facilitates the assessment of test-retest reliability and reproducibility for functional and structural connectomics. To accomplish this, they aggre- gated resting state fMRI (Rs-fMRI) and diffusion imaging data from laboratories worldwide, which are shared via the International Neuroimaging Data-sharing Initiative (INDI). The CoRR dataset includes 18 centers and 33 sites, 32 of which are currently available for download, comprising 1629 subjects and 3357 anatomical scans

- Southwest University Longitudinal Imaging Multimodal

  Southwest University Longitudinal Imaging Multimodal brain data comprise a long-term test-retest sample of young, healthy adults in Southwest China. The SLIM dataset aims to build a large sample with a long-term longitudinal design and a narrow age-span for the assessment of test-retest reliability and reproducibility of brain-behavior correlations as well as for the development of novel causal insights into these correlational findings. The SLIM dataset includes brain sMRI scans that provide a set of structural, diffusion and Rs-fMRI images acquired from 555 subjects (308 females and 247 males)

- Summary

  The compositions of the datasets used herein are shown. Some repeated subjects (18 female subjects from the CoRR database were repeated in the NKI database; 115 female and 118 male subjects were repeated in the SLIM database), subjects missing gender information, or data that were not subjected to the data preprocessing pipeline due to pool imaging quality were removed. Therefore, the final dataset used herein totalled 6008 subjects, including 3292 females and 2716 males.Figure shows the details for each datasets.

Dataset	Total number	Female/Male	Number of sites
HCP	1090	592/498	-
FCP	1058	596/462	22
GSP	1570	905/665	5
NKI	461	257/204	-
SLIM	555	308/247	-
CoRR	1274	634/640	31
Total	6008	3292/2716	>61

Fig. 3.2: Subjects information for each subset used herein.

3.2   Neuroimaging data preprocessing

To minimize the influence of different centers machines and scanning parameters, all T1-MRI data from all datasets were preprocessed to generate normalized brain volume maps via the same preprocessing pipeline. In this study, we used the previously outlined protocol to generate volumetric maps as the original input. All the structural images were preprocessed using SPM12 software and the CAT12 package. All the T1 images were first normalized to the standard Montreal Neurological Institute (MNI) space, and the images were segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) partitions using the tissue prior free segmentation routine of the Segment Data toolbox in CAT12. Then, using the options Display One Slice For All Images and Check Sample Homogeneity Using Covariance in Check Data Quality, we evaluated the segment and normalization quality. To retain the brain edge information and maintain the local differences, we did not utilize the smoothing step, which is a typical step in preprocessing MRI data. The detailed flowchart for structural MRI preprocessing is shown below
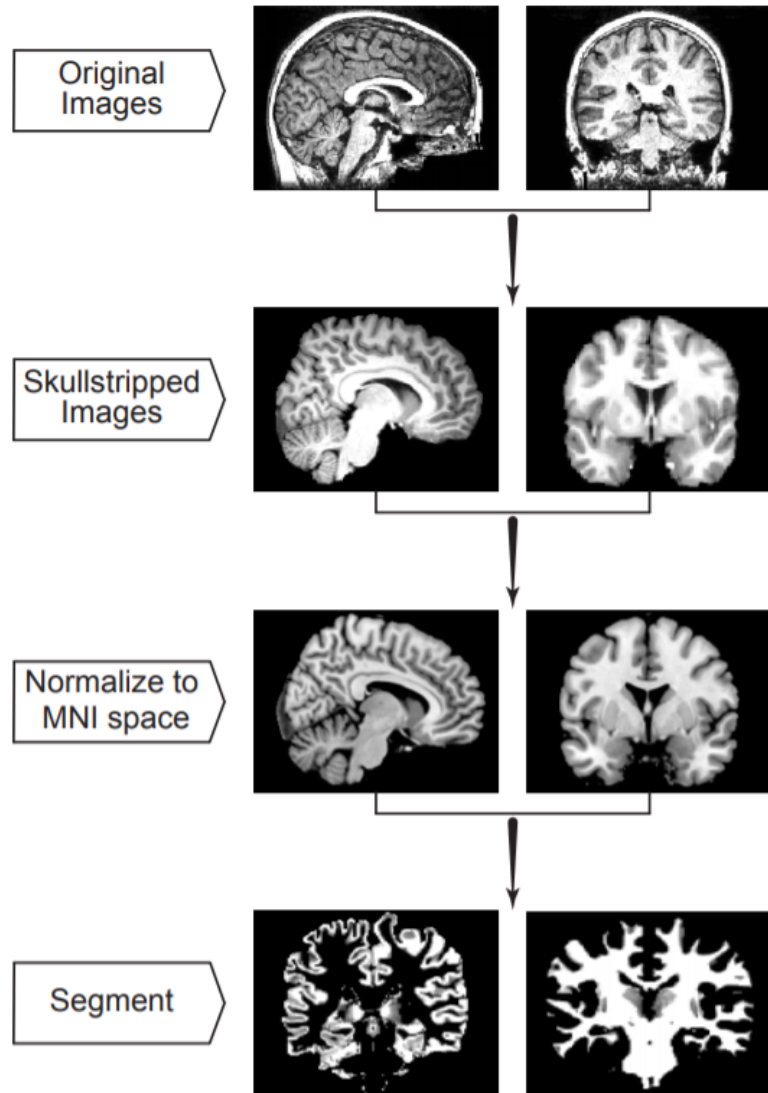
Fig. 3.3: Structural MRI preprocessing

3.3   3D deep adding neural network

The CNN was first proposed by Lecun et al in 1998 and has been very actively investi-
gated, especially recently. As mentioned in the Introduction , several different network archi-
tectures have been proposed, which have demonstrated state-of-the-art performances in many
computer vision and speech recognition tasks. We hypothesize that a CNN can provide the
appropriate architecture to infer imag- ing features from processed brain MRI scans to predict

a subjects gender. The 2D residual neural network (ResNet), which eases network training, was proposed by Kaiming He et al in 2015, and they won 1st place in the ILSVRC 2015 classification task. Some architectures based on the 3D CNN model for predicting brain maturity, predicting neurodevelopment, Alzheimers disease classifications , brain segmentation and skull stripping. With the inspiration from ResNet and previous 3D CNN architectures, we propose the 3D DANet model for neuroimaging classification.

Our proposed 3D DANet architecture used MRI volumes of a specific size (z x h x w) as inputs. Because we used CAT 12 as the preprocessing toolbox, the MRI resolutions of all subjects were normalized into 1.5 mm x 1.5 mm x 1.5 mm . Thus, in our application, the size of the input GM data was 121 x 145 x 121 . The predicted output was a single scalar representing the biological gender.

Fig. 3.4: Illustration of the proposed 3D DANet

A schematic illustration of the 3D DANet architecture, which includes two parts, is given. The bottom side contains two layers: a 3D convolutional layer (with a kernel of 3 x 3 x 3 and a stride of 2), a 3D BN layer, a rectified linear unit (ReLU), and finally, a 3D max

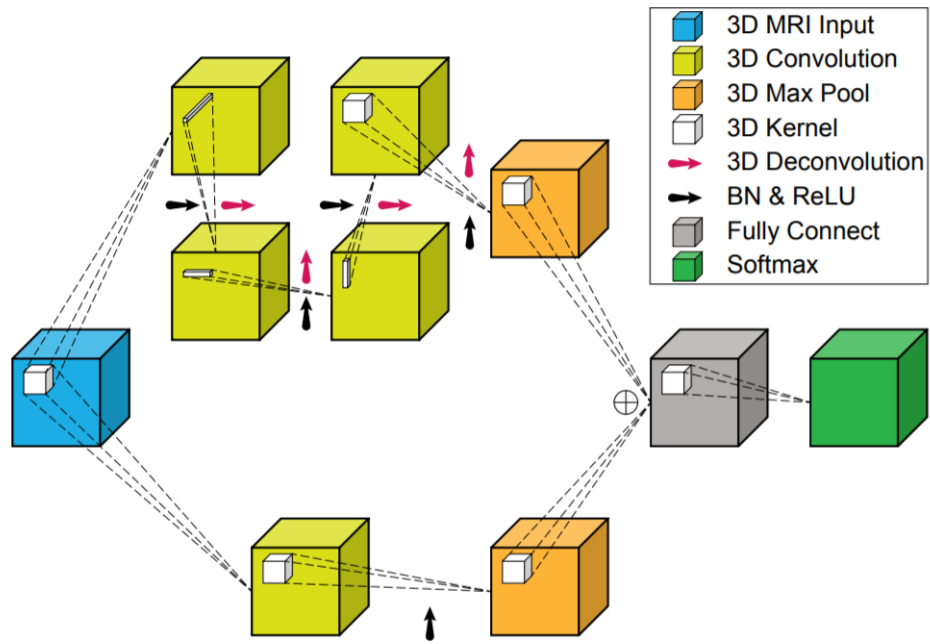pooling layer (with a stride of 1). The top side contains five layers: a 3D convolutional layer (with a kernel of 3 x 1 x 1 , a stride of 1 and a 3D deconvolutional layer), a 3D BN layer, a ReLU, a 3D convolutional layer (with a kernel of 1 x 3 x 1 , a stride of 1 and 3D deconvolutional computation), a 3D BN layer, a ReLU, a 3D convolutional layer (with a kernel of 1 x 1 x 3 , a stride of 1 and 3D deconvolutional computation), a 3D BN layer, a ReLU, a 3D convolutional layer (with a kernel of 3 x 3 x 3 , a stride of 1 and 3D deconvolutional computation), a 3D BN layer, a ReLU, and finally, a 3D max pooling layer (with a stride of 2). The 3D convolutional layer indicates that a convolutional calculation was performed with a 3D kernel in the 3D spatial space. The kernel size of the 3D deconvolutional computation was the same as that used for the 3D convolutional layer. The BN represents the layer prior to activation, and the ReLU represents the method of activation. This function can be formulated as

$$f\left(x_i\right) = \max\left(0, x_i\right)$$

The number of feature channels was set to eight for all the layers. Like in standard CNNs, a fully connected layer was added at the end of the network to encode semantic information. Finally, the neurons in the final layer (the classification layer) were grouped into m = C feature maps, where C denotes the number of classes (C typically equal 2 herein). To avoid overfitting, a dropout approach was used after the fully connected layer. The sparse softmax cross-entropy function was selected as the loss function, and the Adam algorithm was used as the optimizer. The probability score of the class from 1 to C was computed as follows:

$$p_c = \frac{\exp\left(y_L^c\right)}{\sum_{c'=1}^{C} \exp\left(y_L^{c'}\right)}$$

All the results refer to the experiments in which the models were initialized with random truncated normally distributed parameters. Training the 3D DANet architectures with only GM

input using a graphics processing unit (GPU, Nvidia GTX 1080) sped up the training. All the programs were written in TensorFlow, an open source software library, for numerical computation using data flow graphs with support for deep learning algorithms and GPU computing.

we evaluate the performance of the proposed method, est the classification accuracy of the proposed method with two validation strategies,

3.4 Performance evaluation

Cross-validation was used in our experiments, to overcome the problem of overfitting and to make the predictions more general.In detail we have made reference to a 5 fold crossvalidation and leave one subset out cross validation.. The performance of the selected machine learning classification techniques was evaluated in terms of generalization rate,sensitivity, specificity and ROC area by using the following measurements

- True Positive (TP): number of males correctly predicted

- True Negative (TN): the number of females correctly predicted

- False Positive (FP): the number of females classified as male

- False Negative (FN): the number of males classified as female

- Sensitivity (SS): the proportion of male subjects predicted correctly

- Specificity (SC): the proportion of female subjects predicted correctly

- Generalization rate (GR): overall accuracy of all subjects predicted correctly

The generalization rate (accuracy), that is the percentage of correctly classified instances, is defined as:

$$generalization\ rate = (TP + TN)/(TP + TN + FP + FN) \tag{Equ:3.1}$$

while the sensitivity and the specificity, that represent respectively the tests ability to detect positive results or the identification of negative results, are defined as:

$$Specificity = TN/(TN + FP) \qquad \text{(Equ:3.2)}$$

$$Sensitivity = TP/(TP + FN) \qquad \text{(Equ:3.3)}$$

3.4.1 5 Fold cross validation

The classifier design can be viewed as a task of a learning model of the relationship between features and class labels in training sets. If the classifier can predict the classes of new samples satisfactorily, we can trust that it truly captures the relationship between features and classes. In this study, a standard 5-fold cross-validation strategy was used to estimate the performance of the proposed 3D DANet architecture, where all the subjects were separated into five folds randomly. One fold was used as the test set, while the others were used to train the model, and the test was repeated 5 times to ensure that every subject was tested. To remove the influence of the random separation and consider the computational time, we repeated the 5-fold cross-validation strategy five times. Based on the limitation of GPU memory and the sizes of the input images, we choose to utilize a batch size of five.

To remove the influence of the random separation of the datasets from the 5-fold strategy, we calculated the SS, SC and GR of the collected datasets 5 times. The mean accuracy and standard deviation results for SS, SC, GR on each and the total dataset are shown. We obtained an accuracy greater than 92.5% for the total dataset, which is the highest recently reported value. The the table shows that the standard deviation is very small, indicating that the influence of random separation is negligible and that the model is robust for changing data.

Dataset	Evaluation		
	SS	SC	GR
HCP	0.8882 ± 0.0091	0.9780 ± 0.0061	0.9370 ± 0.0043
FCP	0.8997 ± 0.0201	0.9172 ± 0.0161	0.9096 ± 0.0043
GSP	0.9454 ± 0.0157	0.9724 ± 0.0057	0.9609 ± 0.0035
NKI	0.9379 ± 0.0186	0.9105 ± 0.0206	0.9212 ± 0.0055
SLIM	0.9419 ± 0.0102	0.9524 ± 0.0131	0.9478 ± 0.0065
CoRR	0.8880 ± 0.0099	0.8769 ± 0.0182	0.8825 ± 0.0043
Total	0.9125 ± 0.0128	0.9383 ± 0.0098	0.9267 ± 0.0017

Fig. 3.5: Performance of the proposed method with a standard 5-fold strategy

3.4.2 Leave one subset out cross validation

To evaluate the performance of cross-site classification, we also tested the proposed approach using the leave one subset out cross-validation (LOSO-CV) strategy, which means that the model was tested with one subset, while training was performed using the other subsets. Because six subsets were used herein, the LOSO-CV strategy was repeated six times such that all the subsets were tested in the proposed approach. The results are shown

Dataset	Evaluation		
	SS	SC	GR
HCP	0.9598	0.8412	0.8954
FCP	0.8983	0.8842	0.8828
GSP	0.9459	0.9160	0.9046
NKI	0.8971	0.9182	0.9050
SLIM	0.9474	0.9156	0.9073
CoRR	0.8891	0.8691	0.8791
Total	0.9234	0.8879	0.9039

Fig. 3.6: Detected regions of interests.

3.4.3 Regions of interests

To explore the ability of 3D DANet to extract features related to the gender label, we also used deconvolution computations to show the corresponding regions of interests (ROIs), and the results are shown. From the figure, we can see that the sparsity of the detected ROIs increased when the depth of the neural network increased, which corresponds with our knowledge. When the network was shallower, the features learned from the network were rough and comprised nearly the entire brain.

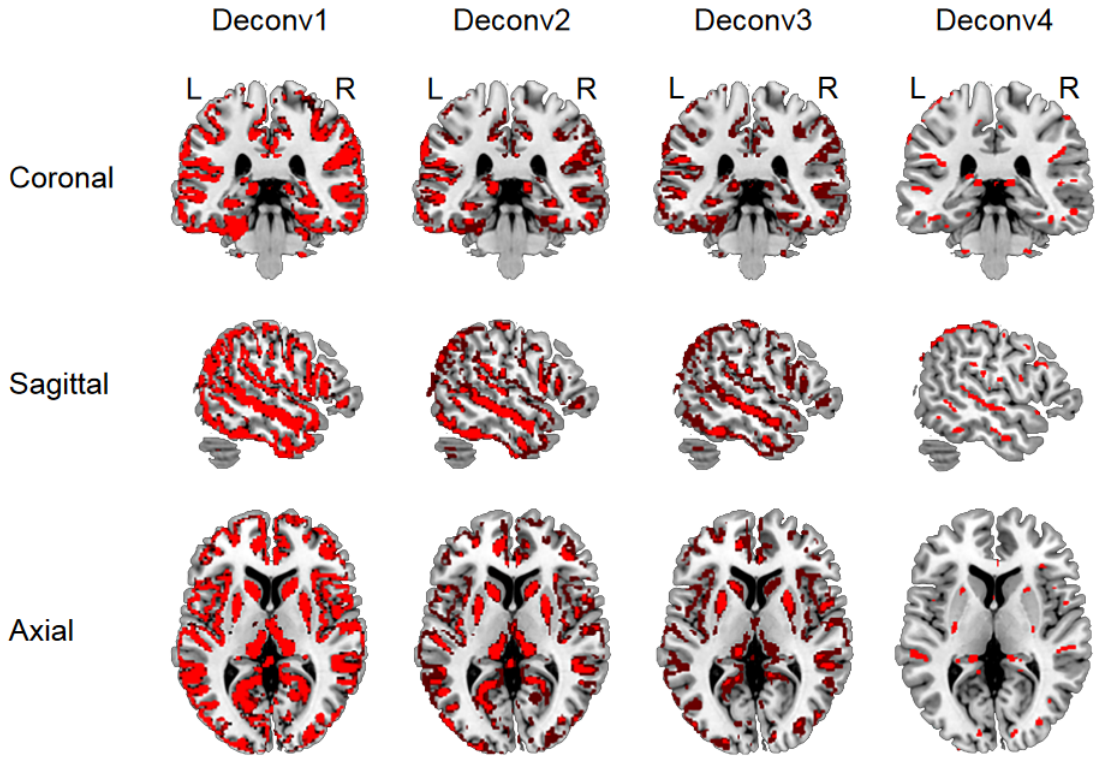
Fig. 3.7: Deconvolutional mapping of the proposed method. The four rows correspond to the four convolutional layers with different sized kernels. The four columns represent the three orientation views of the ROIs. L, left; R, right.

As the network deepened, the features became sparse and some specific regions were detected.used the AAL template to locate the detected regions, which are listed

Regions	Regions
Cerebelum	Vermis
Temporal_Inf_L/R	Temporal_Mid_L/R
Temporal_Sup_R	ParaHippocompal_L/R
Frontal_Mid_L/R	Frontal_Sup_L/R
Cingulum_Post_L	Cingulum_Ant_L/R
Insula_L	Putamen_R
Precuneus_L/R	Cuneus_L/R
Occipital_Sup_L/R	Occipital_Mid_L/R
Angular_L/R	Calcarine_L/R
Supp_Motor_Area_L/R	Cingulum_Post_L
Parietal_Inf_L/R	Parietal_Sup_L/R
Frontal_Med_Orb_L/R	Frontal_Inf_Tri_L/R
Rolandic_Oper_R	Frontal_Sup_Medial_L

Fig. 3.8: Computational results of the proposed method using the LOSO-CV strategy. Dataset indicates the subset that was removed for testing, and Total indicates the statistical results from all the subset

We used a standard template without a gender label to obtain the deconvolutional results, which are equivalent to the features learned from the neural network. The spatial ROIs shown in demonstrate that the features learned from the neural network are located throughout almost entire brain. However, as the network deepens, the features become sparser, which corresponds with our knowledge. The features from the first convolutional layer are rough and comprise nearly the entire brain. However, in the fourth convolutional layer, some special regions to group differences are elucidated. Figure shows that some features are located on the edge of the brain, indicating that the shape of the human brain varies, especially between females and males.

table shows that some regions corresponding to group differences have been previously reported, such as the vermis , occipital lobe, cerebellum, cingulum cortex , precuneus and rolandic regions are reported previously. Zhang et al. utilized functional connectivity to elucidate differences between gender, demonstrating that these differences do existed. Similar results from other previous studies demonstrated that the proposed method can potentially be

applied in clinical practices and to reflect the special regions of some brain diseases.

3.5   Comparing our model with other methods

To compare the performance of our proposed model to other methods, we selected some typical traditional classification methods and deep learning approach architectures. In our previous study, we proposed a 3D descriptor (3D weighted histogram of gradient orientation, 3D WHGO) and selected the PCA+SVM approach, a typical multivariable pattern analysis (MVPA) approach, as the comparison model. In this paper, we also compared our proposed method with the 3D WHGO method and the PCA+SVM approach. For the deep learning approach, we compare our method with the typical 3D CNN and 3D ResNet.Itillustrates the comparisons of the selected models.

We performed the same evaluation for each method on the total dataset and on each subset. The first column depicts the results of the 3D CNN model, and the second column shows those of the 3D ResNet approach, both of which are deep learning approaches whose number of layers correspond with the best performance. The third column depicts the results of our previous work, and the fourth column shows the results of the PCA+SVM approach, both of which are traditional methods selected for comparison with our method.shows that the standard deviation randomly separating the datasets is sufficiently small, indicating that the influence of random separation can be ignored. Therefore, to save computational time, we tested the other models with the 5-fold strategy only once.

The comparison results of our proposed method with the other methods are shown in Figure. The average accuracies of each subset and the total dataset are shown as the area of the circle. The proposed method was the most accurate among all the datasets, which further proves its discriminative power to classify neuroimaging data.
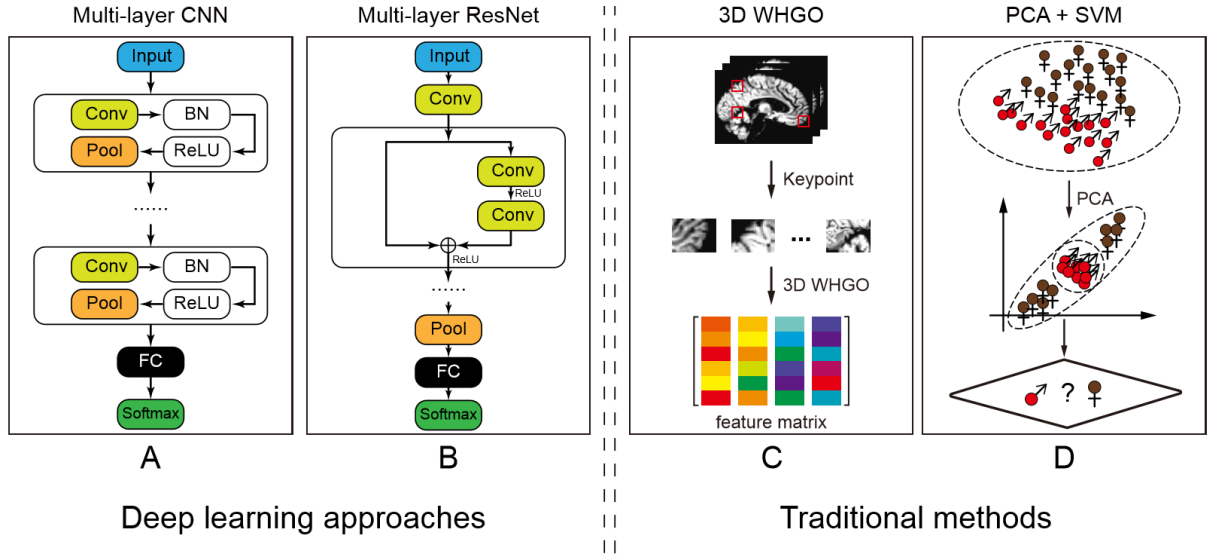
Fig. 3.9: Illustration of some models selected for comparison with our proposed model. Deep learning approaches are shown on the left. The first approach illustrates a multiple layers 3D CNN, and the second depicts 3D ResNet model with multiple layers. Traditional methods are shown on the right. The first method is our previous 3D descriptor (3D WHGO), and the second represents PCA+SVM, a typical MVPA approach..

With the increasing number of neuroimaging data samples, traditional analysis methods begin to lose efficacy. Some traditional statistical methods, such as principal component analysis (PCA), independent component analysis (ICA), and classifications based on feature description and feature selection approaches, such as scale invariant feature transform (SIFT), 3D descriptors (e.g., 3D WHGO), support vector machine (SVM) and MVPA, are incapable of handling large datasets. Because traditional methods must load all the training data to generate a model for classification, the cost of memory becomes prohibitive for completing the computation as the number of samples increases.

Dataset	Evaluation	Methods			
		3D CNN	3D ResNet	3D WHGO	PCA+SVM
HCP	SS	0.9719	0.9578	0.8112	0.7711
	SC	0.8767	0.9172	0.9037	0.8041
	GR	0.9202	0.9358	0.8615	0.7889
FCP	SS	0.9351	0.8918	0.6527	0.6818
	SC	0.8507	0.8926	0.9094	0.8389
	GR	0.8875	0.8922	0.7977	0.7703
GSP	SS	0.9609	0.9564	0.7624	0.7879
	SC	0.9238	0.9591	0.9348	0.8950
	GR	0.9395	0.9579	0.8618	0.8497
NKI	SS	0.9509	0.9363	0.7246	0.7647
	SC	0.8638	0.8482	0.9183	0.8482
	GR	0.9024	0.8872	0.8329	0.8113
SLIM	SS	0.9676	0.9393	0.8907	0.8381
	SC	0.9221	0.9448	0.9221	0.8571
	GR	0.9423	0.9423	0.9081	0.8486
CoRR	SS	0.9250	0.8891	0.7859	0.7500
	SC	0.7918	0.8533	0.8359	0.7965
	GR	0.8587	0.8713	0.8108	0.7732
Total	SS	0.9499	0.9267	0.7673	0.7607
	SC	0.8218	0.9092	0.9031	0.8423
	GR	0.9071	0.9171	0.8417	0.8054

Fig. 3.10: Comparing results with those of other methods.

In contrast to traditional methods, the deep learning approach has a natural advantage to address big data. The approach of partitioning large training datasets into small batches and using iterations to decrease loss-of-function by speeding up the GPU easily satisfies the need for computational speed. By performing iterations numerous times, the training dataset is utilized much more often in the deep learning approach much than in traditional methods, which is why deep learning performs better than traditional methods.
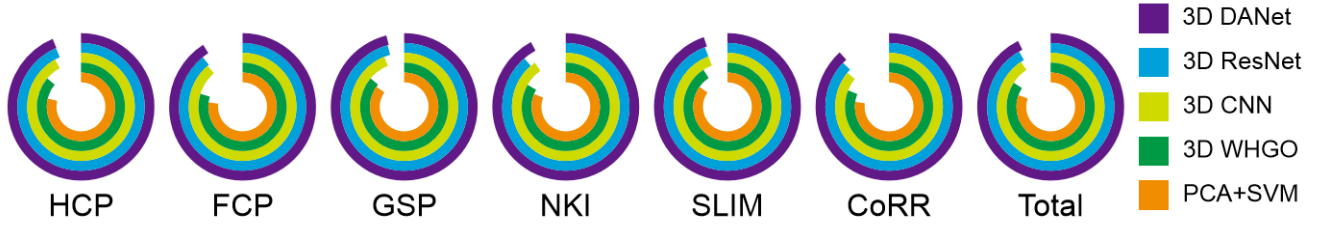
Fig. 3.11: Comparison of our proposed method with other methods. The different colors indicate the various methods, and the circle area indicates the accuracy; larger areas correspond with higher accuracies

3.6 Dept of the neural network

Previously, researchers have proposed numerous types of deep neural network architectures to classify millions of images with thousands of classes. Researchers have developed networks to comprise more than one hundred layers to obtain a satisfactory result. The previous classification was made to solve the multiple classes problem, and each class contained thousands of pictures. By contrast, in the neuroimaging classification field, the number of samples dose not reach the level of natural scenes, and the number of neuroimaging data classes is also fewer. Furthermore, the neuroimaging dimension is much larger than the number of subjects. Due to the neuroimaging characteristics, especially the two-class classification problem described herein, the neural network should not be deeper. A shallower neural network may grasp the differences between classes, while deeper neural networks may learn the pattern uncorrelated with class labels. Based on these considerations, we herein proposed a shallower deep adding neural network to address neuroimaging classifications.

He et al demonstrated that when deeper networks begin to converge, a degradation problem is exposed because the accuracy becomes saturated (unsurprisingly) and then rapidly degrades as the network depth increases. We believe that for the two-class neuroimaging classification problem, an effective neural network should not be deeper but rather more shallow, which is why the proposed 3D DANet has only one adding operation.

3.7 Kernal selection and scale combination

We choose four different sizes of convolution kernel sizes because the four kernels can yield gradient information in different orientations. The first three kernels, whose sizes were 3 x 1 x 1, 1 x 3 x 1 and 1 x 1 x 3, can extract gradient information in three different orientations, while the final kernel (3 x 3 x 3) can extract gradient information in a solid cube. The adding operation can combine the information from the different scales.

On the top branch, 3D convolution operations are completed in the spatial space with a stride of 1, corresponding to the original 3D imaging scale. For the bottom branch, the 3D convolution operation is performed with a stride of 2 in the spatial domain, which is a smaller scale. Using two pooling layers, the results under the two scales can be altered to equal sizes and then added. Thus, the two scales are combined via the adding operation. Our previous work demonstrated that combining the information from different scales is effective for improving the classification performance.

3.8 Transfer ability

We also tested the cross-site classification of the proposed method. Herein, we extended the standard k-fold crossvalidation strategy and used the LOSO-CV strategy, rate in table is shown. For the validation of each time, one center subset was left out, and the other subsets were used as the training datasets; the test subset thus had no overlap with the training datasets. All the accuracies remained high, which demonstrated that the proposed method is transferable and can effectively handle multicenter classifications. Reasonably, all the accuracies obtained using the LOSO-CV strategy were slightly lower than those obtained using the 5-fold cross-validation strategy because the distributions of the test subsets differed from those of the training datasets

The code for our proposed model was written using TensorFlow, an open source library that is well suited for mobile deep learning, and the application thus removes the platform limitation. In the future, we plan to build a system in which the model is trained in the cloud based on big data and test the practical data on the clinical diagnosis made on mobile terminal.

This may lead to significant developments in artificial intelligence (AI) and a clinical revolution.

3.9 Big data from multiple centers

To the best of our knowledge, our work is the first to classify neuroimaging data on such a large scale and multicenter data with such a high accuracy (over 92.5%). Furthermore, this is the first time that the deep learning approach has benn used to solve the multicenter classification problem. With the development of neuroimaging technology, increasing amounts of MRI data are being acquired, creating large differences between scanners exist and varying image qualities. These difficulties are problematic for traditional classification methods based on single sites and small sample sizes. Our method utilizes deep learning technology to address the big data classification problem on multicenter datasets and yields satisfactory results, demonstrating that the proposed approach can handle multicenter classifications. We believe that analyses based on big data from multiple centers will increase in the future.

Conclusion

With the development of neuroimaging technology, increas- ing amounts of MRI data are being acquired. Herein, we proposed a 3D deep adding neural network (3D DANet) to address the brain imaging classification problem on multi- center datasets. Our method utilizes multiple convolutional kernels of varying sizes to extract gradient information in different orientations and combines spatial features at two scales via the adding operation. A standard 5-fold cross- validation strategy was used to test the performance of the proposed method. We collected a large dataset compris- ing 6008 healthy subjects from more than sixty scanners at six centers. A high accuracy (over 92%) in distinguish females from males was observed, which demonstrates that the proposed method can handle big data classification from multiple centers. Compared with some traditional methods and architectures of the deep learning approach, the proposed method performed better, highlighting its advantages for addressing two-class neuroimaging data classifications. With its improved performance for neuroimaging classifications and transferable program codes, the proposed method can potentially be used for intelligent medical treatment strategies and clinical practices based on mobile terminal

The tests have been carried out over the overall dataset and over three different subset where we only have considered the selected features by features selection meth- ods. The results have shown that the best accuracy in brain classification is achieved using the 3D deep addiing neural network

REFERENCES

[1] F. Ge, X. Li, M. J. Razavi, H. Chen, T. Zhang, S. Zhang, L. Guo, X. Hu, X. Wang, and T. Liu,, "Denser growing fiber connections induce 3-hinge gyral folding", *IEEE Computer*, Cerebral Cortex, pp. 112, 2017.

[2] A. C. Altamura, E. Maggioni, T. Dhanoa, V. Ciappolino, R. A. Paoli, L. Cremaschi, C. Prunas, G. Orsenigo, E. Caletti, C. M. Cinnante, F. M. Triulzi, B. DellOsso, L. Yatham, and P. Brambilla,"The impact of psychosis on brain anatomy in bipolar disorder: A structural mri study,,"]*IEEE Wireless Commun.,* Journal of Affective Disorders, 2017.

[3] Y.-K. Kim and K.-S. Na "Application of machine learning classification for structural brain mri in mood disorders: Critical review from a clinical perspective "*IT Professional,*Progress in Neuropsychopharmacology and Biological Psychiatry, vol. 80, pp. 7180, 2018

[4] C. Tas, H. Mogulkoc, G. Eryilmaz, I. Gogcegoz-Gul, T. T. Erguzel, B. Metin, and N. K. Tarhan,, "Discriminating schizophrenia and schizoobsessive disorder: a structural mri study combing vbm and machine learning methods,," *Neural Computing and Application, vol. 29, pp. 377 387. J.,* 2018.

[5] Y. LeCun, Y. Bengio, and G. Hinton,, "Deep learning, " *Nature, vol. 521, pp. 436444, .,* 2017.

[6] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sanchez,, "A survey on deep learning in medical image analysis,, " *Medical Image Analysis,. ,* vol. 42, 2017, pp. 6088.

[7] Q. Zhang, L. T. Yang, Z. Chen, and P. Li,, "MA survey on deep learning for big data,, "*Information Fusion, vol. 42, pp. 146157, 2018.,* vol. 42, pp. 146157, 2018..

[8] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei,., "Imagenet large scale visual recognition challenge,, " *International Journal of Computer Vision, vol. 115, pp. 211252, 2015.*

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks,, " *Communications of the ACM, *, vol. 6, pp. 8490, 2017..

[10] K. Simonyan and A. Zisserman,, "Very deep convolutional networks for large-scale image recognition " *in International Conference on Learning Representations 2015.* 2015

[11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich,., "Going deeper with convolutions, "*n 2015 IEEE Conference on Computer Vision and Pattern Recognition,* 2015, pp. 19..

[12] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun,, "Dermatologist-level classification of skin cancer with deep neural networks " *Nature, vol. 542, 2017 ,* vol. 542, 2017 .

[13] J. Dolz, C. Desrosiers, and I. B. Ayed,, "3d fully convolutional networks for subcortical segmentation in mri: A large-scale study, " *NeuroImage* 2017