Multi-tasks Deep Learning Model for classifying MRI images of AD/MCI Patients

S.Sambath Kumar[1], Dr M. Nandhini[2],

[1]Research scholar, [2]Assistant Professor
[1,2]Department of Computer Science, Pondicherry University, India
[1]sambathkumars06@gmail.com
[2]mnandhini2005@yahoo.com

Abstract-The accurate diagnosis of Alzheimer's diseases (AD) and prodromal stage like Mild Cognitive Impairment (MCI) play a vital role in preventing progression of Alzheimer diseases and mild cognitive impairment. In view of that, a multi-kernel classifier model with noninvasive imaging technique for AD/MCI patients is proposed in this paper. The proposed model includes four techniques, such as PCA, Stability Selection, Multitask deep models with dropout and AD/MCI diagnosis with kernel SVM leads into the deep learning framework. Also, the proposed approach is evaluated with real-world ADNI datasets (Alzheimer's diseases Neuroimaging Initiative) and its results are analyzed.

Keyword - Alzheimer's diseases (AD), Mild Cognitive Impairment (MCI), Deep Learning, Magnetic Resonance Imaging (MRI), early diagnosis and Support Vector Machine.

I. INTRODUCTION

Machine Learning, a type of artificial intelligence (AI) provides automatic program execution[1]. In real time, it focuses on dynamic updating of programs when exposed to new data. ML is classified into three categories; supervised learning, unsupervised learning, and reinforcement learning. Supervised learning is trained learning model whereas unsupervised learning draws inferences from the data provided. Reinforcement Learning considers the output from previous experience to take an accurate decision. Machine learning approach and most existing pattern classification method have been growing in research to the diagnosis of diseases. Deep Learning, common to all the above-mentioned three methods, extract a pattern from any given data such as text, images, audio, and video in unsupervised learning.

In pattern classification method, each point coordinated in the image is denoted as picture elements or pixels. The strength at every pixel is represented by an integer (0, 255) say 0 for black color and 255 for white color. To classify an image, methods are used to compare it with neighbourhood stored models known as patterns to choose the best match. The comparison of images is directly performed by template matching. Every template consists of pixels or images of known patterns. To find the best match each template is compared with a new image[2]. These patterns are processed using image modalities. Image modalities are classified into two categories: single image modality, MRI or PET only, and multi-image modality, a combination of MRI, PET, fMRI, EEG signals, CT brain CSF, and FDG.

Siqi Liu et al.2015[3]proposed a novel framework using stacked autoencoders (SAEs) and softmax logistic regressor merged with multi-kernel support vector machine (MKSVM) for multiclass AD diagnosis with deep learning architecture. The achieved accuracy was 91.67% with this method. Andre Santos Ribeiro et al.2015[4] introduced new Multimodal Imaging Brain Connectivity Analysis (MIBCA) tool, which they used for processing fMRI image and PET scans. This toolbox is able to process aMRI from high resolution scanned images, fMRI image from Blood Oxygen Level Dependent (BOLD) and dMRI from DTI.

The above works motivated us to develop an effective model with multi-kernel support vector machine (MKSVM) for multiclass AD to improve the accuracy levels with low false positive and false negative values.

In order to improve the accuracy level, a framework of the multi-kernel classifier for AD diagnosis with data from MRI scans is proposed with the hierarchy steps as follows. The 3-D MRI image scans are preprocessed and their feature was further extracted. PCA, Stability selection, deep learning and dropout techniques are applied in addition to MoCA, BNT, and CFT achieving better accuracy with single image modalities. Finally, trained feature representation is used for AD/MCI categorization by multi-kernel support vector machine (MKSVM).

The organization of this paper is presented as follows. Section 2 focuses on related works on deep learning neural model of ADNI datasets. Multi-kernel classifier approach is proposed in section 3. Section 4 presents experimental implementations, dataset used and results. Section 5 concludes the paper.

II. RELATED WORKS

In this section, some of the related works with wide range of approaches and techniques are presented with different datasets that achieve different accuracy levels.

Xiaohong W.Gao et al,[5] introduced machine learning approach(CNN) for finding the Alzheimer's disease or normal ageing or lesion by using 2D and 3D scans (282 subjects: AD- 51, lesions- 118, normal healthy subjects -117). These scanned images have a larger thickness and the depth direction of 5mm. The accuracy value of 86.8% was achieved.
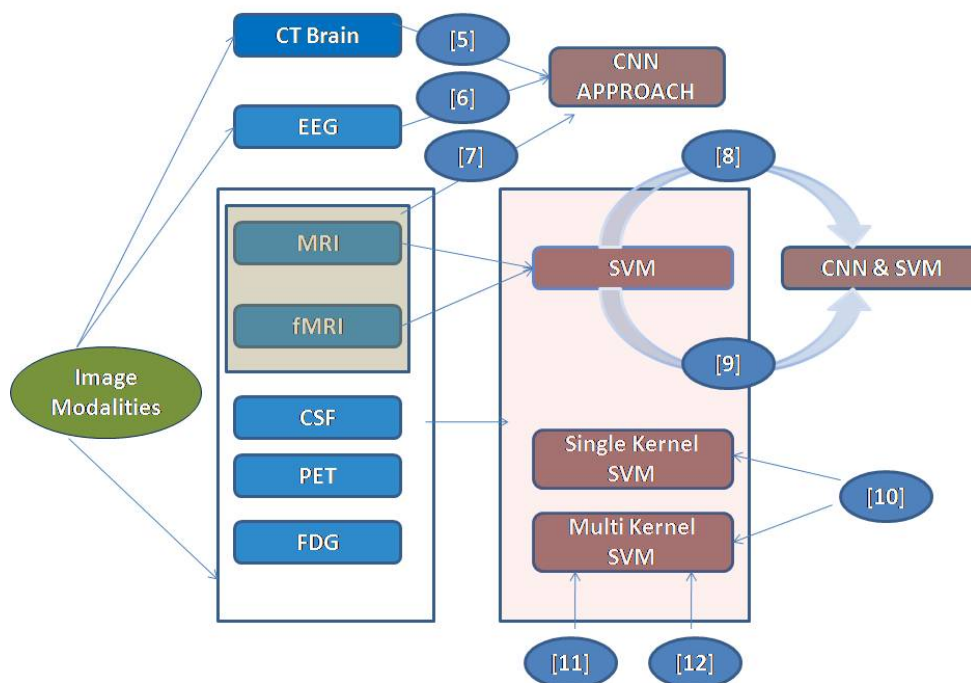

Fig.1.Classification based on multi-model image modalities

Francesco Carlo Morabito et al,[6]discussed the use of different biomarkers yielding promising accurate outcome and various image modalities giving a different view of the function of brain images for diagnosing AD/MCI using CNN approach (119 dementia patients: 63 AD subjects, 56 MCI subjects). For example, EEG signals classify EEG pattern of AD from the prodromal version of dementia. CNN approach attained an accuracy level of 95%.

Only few research studies focus on multimodal neuroimaging for AD/MCI classification. For example,

Ehsan Hosseini-Asl et al,[7] introduced structural MRI scans using deep learning approach (3D CNN). The proposed method involved 3D ACNN classifier approach; enabled to find the AD using structural MRI in an accurate manner (210 subjects of ADNI).The performance of ACNN classifier approach reached an accuracy level of 95%.

Saman Sarraf et al,[8], discussed state functional magnetic resonance imaging (rs-fMRI)data and fMRI data to recognize the AD disease during a clinical scan. The proposed method involves SVM and CNN approach (ADNI - 24 female and 19 male). The accuracy value of 96.85% was achieved.

Saman Sarraf et al,[9] introduced machine learning approach (CNN)with SVM to classify Alzheimer's disease from a normal healthy subject using single image modalities(ADNI - age group > 75). The accuracy value of 98.84% accomplished.

Sigi Liu et al[10], introduced stacked autoencoder (SAE) and softmax layer for diagnosing AD and MCI in the early stages. The proposed method incorporates optimized graph cut algorithm, single kernel SVM (SKSVM) and multi-kernel SVM (MKSVM). It is a semi-supervised method, which allows unlimited use of unlabeled data samples (ADNI- from 311 subjects: 65 AD, 67 cMCI, 102 ncMCI, 77 normal controls). SAE and softmax layer concluded with an accuracy value of 87.76%.

Chen Zu et al.,[11] Chen developed two main components (i.e.) multimodal classification and multitask feature selection for AD/MCI patients. Multi-kernel SVM method is proposed to classify the multimodality image data for final classification (ADNI 202 participants: 51 AD, 99 MCI, and 52 normal Control). Multi-kernel SVM method shows an accuracy value of 95.95%.

Heung-Il Suk et al[12], introduced deep learning approach stacked auto-encoder(SAE) for AD/MCI diagnosis. The SAE model permits us to find the best possible parameter in adjustment with the exact sample data (ADNI (51 AD patients, 99 MCI52, HC subjects). This method involved SAE learned feature representation in brain diseases and its achieved maximum accuracy based on the classification. The accuracy value of 98.8% was achieved.

The above-discussed work focuses on using machine learning approach and deep learning approach (CNN, SVM&CNN, SKSVM, and MKSVM) with single image modalities and multi-image modalities. The classifier approach has some cons, larger datasets and increased memory cost, data redundancy, the privacy of original data, scalability, efficiency and increase memory cost. To overcome this disadvantage the following approaches are proposed. PCA, Ridge regression technique, lasso method, Elastic Net in Stability selection, deep learning and dropout in addition to MoCA, BNT and CFT achieving better accuracy.

III. DESIGN OF PROPOSED FRAMEWORK

In this section method of this proposed work is discussed. Principal Component Analysis (PCA) is used for reducing the dimensionality and feature extraction from the preprocessed data. Stability selection is used to identify and preserve commonly found dominant features. Ridge and Lasso are an eminent technique instability selection, where a myriad number of features have been generated. The Elastic Net method is particularly used when a large number of correlations are present between the feature variables. In deep learning, the preferred features are processed by dropout techniques. Dropout is used to enhance the generalization capability of the model. And finally, support vector machine (SVM) is used to classify AD/MCI[15].

In support of the techniques used, complementary clinical scores are provided to achieve better accuracy in diagnostic. The complementary clinical scores are generated by clinical experts. There is no single test that proves a person has Alzheimer's. There are 5 additional clinical scores added namely Minimum Mental State Examination (MMSE)[16], the ADAS-CogIRT[17][18], BNT, MoCA and CFT[19]. When the first and second of these methodologies are incorporated simultaneously it requires fewer data sets and minimal trial durations. The information from the score is related to identify the AD diagnosis. The main concept behind this approach is the deep learning structure as a multi-learning framework (MTL).
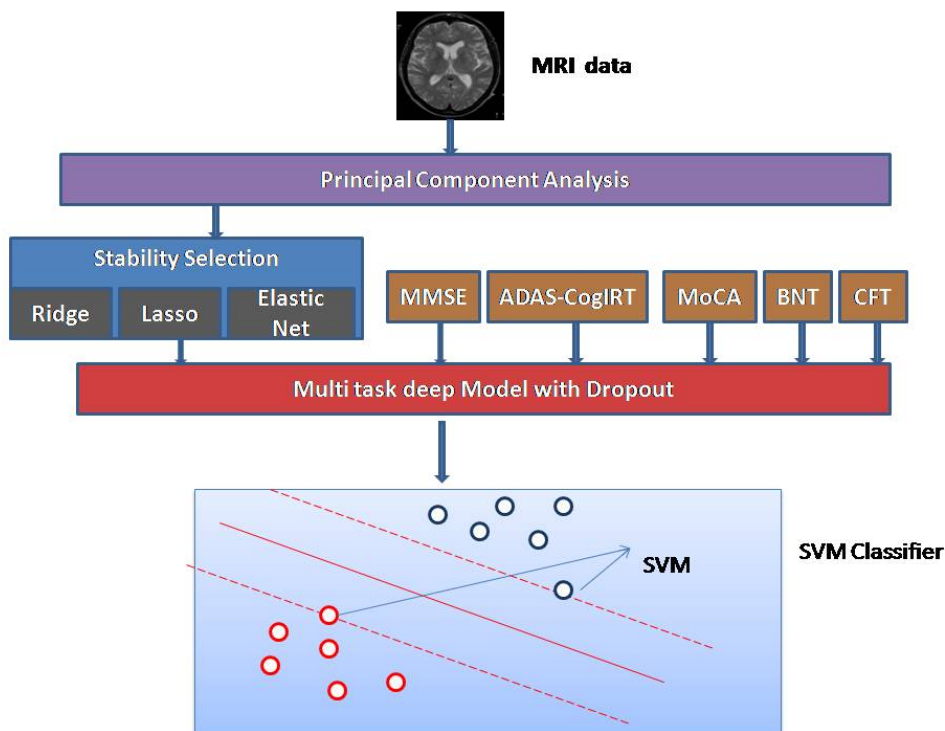

Fig.2. The design of proposed multitasks deep learning model.

The proposed work consists of multiple structures: PCA, Stability selection, unsupervised feature learning, multi-deep learning, and kernel SVM, as shown in Fig 2. In the following section, each of this components is explained in detail.

A. Principal Component Analysis

Principle Component Analysis is an efficient technique to process the medical images to find the exact diseases in the medical field. The main objective of the PCA is to rotate the sample data multiple times to align the maximum variance with the reference coordinate. For each rotation, the next maximum variance is recognized for feature selection. PCA derives uncorrelated variables from a set of features. It actually converts and changes the direction of the variables in the best manner where the required components can be seen clearly. The biggest variance is PC-1 among the original datasets, which has been used much in transformation such as predicting, redundancy detection and removal extraction of required variables and compression of data.

B. Stability Selection

Initially, PCA is applied to reduce the dimensions. Secondly, Stability method is applied for feature selection method. Some of the stability selection methods widely accepted is ridge regression, Lasso tool[20], and Elastic Net. Ridge regression technique is applied to find covariance value of the given data. It starts with the smallest value among the PC and proceeds toward the diagonal elements of the matrix. One of the main goals of using ridge regression technique is to select an exact value. The feature values obtained in ridge regression is fed to lasso tool for further extraction[21]

The lasso method will identify the top level feature for AD/MCI diagnosis for example (a1, a2, a3, a4, a5) as well as try to reduce the cost of function for feature selection. Then lasso picks the distinct features (PCs) in PCA, $s = [s1, s2, s3....st]^t$. The idea behind using this technique is to improve the selection procedure method multiple times based on the data. The above lasso method is repeated for 50 times. Then we applied the Elastic Net technique to solve the regularization and high dimensional data. Lasso will select only one covariance but Elastic Net will select multiple covariances of the variable.

C. Multi-task Learning

Multi-task Learning is mainly used for improving the performance and to learn simultaneously from the appropriate given information. The baseline assessment is applied to identify the exact AD/MCI patients namely, MMSE, ADAS-CogIRT, BNT, MoCA and CFT. These five different tasks perform collaboratively for the final outcome.

→Minimum Mental State Examination (MMSE) is one of the eminent clinical tests to assess the AD patients. ADAS-Cog is low sensitive to find out the AD diseases. To achieve high sensitivity ADAS-CogIRT (Alzheimer's diseases Cog methodology with IRT modeling) has been implemented.

→Boston Naming Test involves handing over line drawing containing an outline of the objects for the patients to recognize. These objects are of the frequently encountered type in the day to day life. The score from the test is taken as an input.

→CFT is a test of calculation of semantic memory (i.e. mother tongue and verbal fluency). The participant will be asked to identify the exemplars from the given semantic category.

→MoCA assessment is designed as a fast screening tool for mild cognitive impairment. MoCA involves a number of tests to assess the different type of cognitive domains (i.e. attention and concentration, language, memory, visuoconstructional skills, executive function, calculation, orientation, and thinking.)

D. Adaptive Dropout Technique:

The Dropout technique is very effective and simple way to characterize the deep neural network model. It is a very successful technique to train the deep learning model. It can process thousands or millions of parameter in dropout technique within minutes. It finds some weight coadaptation by haphazardly dropping out some smaller unit in the model during the training period [24]. Adaptation dropout technique is applied in this proposal to improve the AD/MCI diagnosis patients.

E. Support Vector Machine:

SVM plays a vital role in machine learning. Supervised learning model and associated learning algorithm studies data used for regression analysis and classification[25]. This SVM technique cannot be used in deep learning. Here the baseline method is only available for further selection with SVM steps. Linear kernel method will not improve the classification accuracy because the scalar method is in dual form. All the experiments were carried out with three hidden layer model.

IV. EXPERIMENTAL ANALYSIS

A. Dataset

The Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset is used for training and testing the proposed multi-task deep model.

The Subjects are classified into three main categories. Such as MRI scans from 51 AD patients, 52 normal healthy control (NC) as well as 99 MCI patients [43 MCI patients who changed over to AD (MCI.C) 56 MCI Patients not converted into AD with 18 months (MCI.NC)]. The detailed information acquiring from ADNI data set is used for processing of AD/MCI diseases to improve the patients with a better life. MRI image modalities are sensed using 1.5 T scanners. The MRI scans were downloaded from the ADNI site(adni.loni.usc.edu) with the help of Digital Imaging and Communication in Medicine(DICOM).This dataset contains five additional clinical scores MMSE, ADAS-CogIRT, **BNT, MoCA, CFT** for each and every patient. The process of image processing is applied 3-D MRI images with skull stripping, cerebellum removal, anterior commissure correction, and spatially normalization.

B. Experiments

The proposed work is to identify the AD/MCI patients based on their brain image dataset. To identify AD/MCI diseases choose a normal healthy image to differentiate AD/MCI patients with other normal patients. This can be done by four different combinations i.e. (i) AD/normal healthy control (AD/NC), (ii) MCI subject/ healthy control (MCI/NC), (iii) MCI patients versus AD patients (MCI/AD), and (iv) MCI non-converted versus MCI converted (MCI.NC/MCI.C). In the first combination, one AD brain image and one normal brain image is loaded from the set of images. Later the clinical exports will provide the clinical scores to the loaded images. Based on these clinical scores and brain images, the variation between these two images is calculated by using four techniques one by one. Initially, PCA is applied and the output of PCA is given as input for stability selection. In stability selection, the covariance value of the given data is identified. Later in multitasking, the clinical scores are added and dropout technique is used to remove the smaller units. Finally, the SVM classifier is used to classify those images. The same can be followed for the other three combinations and a comparative table is drawn to show this results.

C. Comparison method

The proposed framework is compared with baseline method. Baseline method has same components as in proposed work excluding deep learning concepts and evaluated with more than two image modalities for diagnosing AD/MCI patients. The accuracy of proposed work is tested with single image modalities and is compared with individual and baseline components as shown in Table.1.

D. Results

Table 1 demonstrates the overall experience of the proposed system in the framework. Our proposed method performed each component such as PCA, Stability Selection, and Drop out technique with higher accuracy. The high-level accuracy was achieved better than baseline method.

It is found that in the conversion diagnosis (MCI.C versus MCI.NC), the PCA component has, in a small degree downgraded the proposed method (i.e. from 58.8 % to 57.9 %). However, it still holds a significantly better position than the baseline method (58.8% versus 50.6%).

In comparison to all the components, it is highly evident that "dropout" has the most important in deep learning. This is because without "dropout," deep learning did not have much impact on the baseline method (67.65 % versus 69.42 % in terms of average accuracy.).

TABLE 1. Result accuracy of proposed and traditional methods.

Task	Proposed	PCA	Dropout	SS	Multitask	Baseline
AD vs HC	91.4	83.2	85.6	92	87.2	86.4
MCI vs HC	77.4	72.8	74.9	81.8	76.3	72.1
AD vs MCI	71.8	70.1	64.6	69.3	70.2	61.5
MCI.C vs MCI.NC	57.9	58.8	52.6	56.8	58.1	50.6
Average accuracy(%)	74.625	71.225	69.425	74.975	72.95	67.65

The proposed method consists of four components: PCA, Dropout, Stability Selection, and Multitask. All the above components except Dropout technique comes under Machine Learning while the later comes under Deep Learning. The Baseline method denotes without deep learning component. The result of AD versus HC equals 91.4%, MCI versus HC equals 77.4%, AD versus MCI equals 71.8 and lastly and MCI.C versus MCI.NC equals 57.9%.

V. CONCLUSION

In recent years, Deep learning is an emerging area in the healthcare industry. An improved model with multi-kernel support vector machine (MKSVM) for multitasking AD is proposed to increase the accuracy levels with high true positive and true negative values. The proposed model used four techniques, Principal Component Analysis (PCA), Stability Selection (SS), Dropout and Multitask deep learning with SVM. Good accuracy level is achieved with the help of introduced dropout technique in deep learning using MRI image modalities for the subject of AD, MCI, and HC. The future work is extended by evaluating multimodal images using multi-kernel classifier approach.

REFERENCES

[1]   A. Heydarzadegan, Y. Nemati, and M. Moradi, "Evaluation of Machine Learning Algorithms in Artificial Intelligence," vol. 4, no. 5, pp. 278–286, 2015.
[2]   P. B. Christos Davatzikosa, "Prediction of MCI to AD conversion, via MRI, CSF biomarkers, pattern classification* Christos," Biophys. Chem., vol. 257, no. 5, pp. 2432–2437, 2011.
[3]   S. Liu et al., "Multimodal Neuroimaging Feature Learning for Multiclass Diagnosis of Alzheimer's Disease," IEEE Trans. Biomed. Eng., vol. 62, no. 4, pp. 1132–1140, 2015.
[4]   A. S. Ribeiro, L. M. Lacerda, N. André, and H. A. Ferreira, "Multimodal imaging brain connectivity analysis (MIBCA) toolbox : preliminary application to Alzheimer ' s disease," EJNMMI Phys., vol. 1, no. Suppl 1, p. A61, 2014.
[5]   X. W. Gao, "A deep learning based approach to Classification of CT brain images," pp. 13–16, 2016.
[6]   F. C. Morabito, C. Ieracitano, L. Bonanno, A. Bramanti, S. De, and I. C. N. Bonino-pulejo, "Deep Convolutional Neural Networks for Classification of Mild Cognitive Impaired and Alzheimer ' s Di sease Patients from Scalp EEG Recordings," 2016.
[7]   E. Hosseini-asl, R. Keynton, and A. El-baz, "ALZHEIMER'S DISEASE DIAGNOSTICS BY ADAPTATION OF 3D CONVOLUTIONAL NETWORK Electrical and Computer Engineering Department, University of Louisville , Louisville , KY, USA .," no. 502.
[8]   S. Sarraf and G. Tofighi, "Classification of Alzheimer's Disease Using fMRI Data and Deep Learning Convolutional Neural Networks," pp. 8–12.
[9]   S. Sarraf and G. Tofighi, "Classification of Alzheimer's Disease Structural MRI Data by Deep Learning Convolutional Neural Networks," pp. 1–14.
[10]   S. Liu et al., "EARLY DIAGNOSIS OF ALZHEIMER'S D ISEASE WITH DEEP LEARNING BMIT Research Group , School of IT , University of Sydney , Australia Surgical Planning Lab , Brigham & Women ' s Hospital , Harvard Medical School , Boston , USA," pp. 1015–1018, 2014.
[11]   C. Zu, B. Jie, M. Liu, S. Chen, and D. Shen, "Label-aligned Multi-task feature learning for Multimodal classification of Alzheimer ' s disease and mild cognitive impairment," Brain Imaging Behav., pp. 1148–1159, 2016.
[12]   H. S. S. Lee and D. Shen, "Latent feature representation with stacked auto-encoder for AD / MCI diagnosis," pp. 841–859, 2015.
[13]   N. Meinshausen and B. Peter, "Stability Selection", vol. 72, pp. 417–473, 2009.
[14]   R. Tibshirani, "Regression shrinkage and selection via the lasso.pdf," vol. 58, no. 1, pp. 267–288, 1996.
[15]   S. Rastegar, M. S. Baghshah, H. R. Rabiee, and S. M. Shojaee, "MDL-CW : A Multimodal Deep Learning Framework with Cross Weights," pp. 2601–2609, 2016.
[16]   Verna C. Pangman. .Jeff Sloan, "An examination of psychometric properties of the Mini-Mental State Examination and the Standardized Mini-Mental State Examination: Implications for clinical practice," vol. 13, no. 4, pp. 209–213.
[17]   E. Kolibas, V. Korinkova, V. Novotny, K. Vajdickova, and D. Hunakova, "ADAS-cog (Alzheimer ' s Disease Assessment Scale — cognitive subscale) — validation of the Slovak version ADAS-cog (Alzheimer ' s Disease Assessment Scale — cognitive subscale) — validizácia slovenskej verzie Abstract," vol. 101, pp. 598–602, 2000.
[18]   S. Ueckert, E. L. Plan, and K. Ito, "Improving cognitive testing with IRT and optimal design," 2013.
[19]   P. Manual, "Alzheimer's Disease Neuroimaging Initiative 2 (ADNI2)."
[20]   A. JAIN, "A Complete Tutorial on Ridge and Lasso Regression in Python," 2016.
[21]   "ridge-regression-lasso-and-elastic-net," 2014. [Online]. https://www.slideshare.net/ShangxuanZhang/ridge-regression-lasso-and-elastic-net.
[22]   H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," vol. 67, pp. 301–320, 2005.
[23]   G. Hinton, "Dropout : A Simple Way to Prevent Neural Networks from Overfitting," vol. 15, pp. 1929–1958, 2014.
[24]   F. Li, L. Tran, K. Thung, S. Ji, D. Shen, and J. Li, "A Robust Deep Model for Improved Classification of AD / MCI Patients," vol. 19, no. 5, pp. 1610–1616, 2015.
[25]   C. Cortes and V. Vapnik, "Support-Vector Networks," vol. 297, pp. 273–297, 1995.
[26]   D. Zhang, Y. Wang, L. Zhou, H. Yuan, and D. Shen, "NeuroImage Multimodal classi fi cation of Alzheimer â€TM s disease and mild cognitive impairment," vol. 55, pp. 856–867, 2011.
[27]   H. Suk and D. Shen, "Deep Learning-Based Feature Representation for AD / MCI Classification," pp. 583–590, 2013.