MULTIMODAL DEEP LEARNING MODELS FOR EARLY DETECTION OF ALZHEIMER'S DISEASE STAGE

Summary

FCI-Bioinformatics
Graduation Project

Problem Diagnosis ***AD*** disease (Classification problem)

Objective Distinguish between 3-stages of ***AD***

- normal (*CN*)
- Mild cognitive impairment (*MCI*)
- *AD*

Datasets from **ADNI** (ADNI1, ADNI2, ADNI GO) databases (adni.loni.usc.edu) contains

- MRI (**ADNI**)
- Clinical (**EHR**) data (ADNI1, ADNI2, ADNI GO)
- Genetic (**SNP**) data (ADNI1, ADNI2, ADNI GO)

	CN	MCI	AD
Clinical Data	598	699	707
Imaging	132	104	266
Genetic	245	338	226

Figure 1. Number of samples for each data modality per class of ***AD.***

Methodology

Make Classification Model for all combinations of each data modality, Model for

- **EHR** Data to classify (*CN* | *MRI* | *AD*)
- **MRI** Data to classify (*CN* | *AD*)
- **SNP** Data to classify (*CN* | *MRI* & *AD* *as one class*)
- **EHR** + **SNP** + **MRI** Data to classify (*CN* | *MRI* | *AD*)
- **EHR** + **SNP** Data to classify (*CN* | *MRI* | *AD*)
- **EHR** + **MRI** Data to classify (*CN* | *MRI* | *AD*)
- **SNP** + **MRI** Data to classify (*CN* | *MRI* & *AD* *as one class*)

And use **shallow** &**Traditional** algorithms | **DL** algorithm as classifiers (comparing between two types). *Figure 2 for performance*

Algorithms

K-mean for clustering **SNP & EHR,** to show difference between associations in *intermediate features* and associations in *raw features*.

For Extracting Features

- **3D-convolutional neural networks (CNNs)** from **MRI**

- **Stacked denoising auto-encoders** from **SNP & EHR**

As Classifiers

- **Shallow Models (SM)**
    - simple **ANN** with 1-2 hidden layers.
    - Traditional (support vector machines **SVM**, decision trees **DT**, random forests **RF**, and k-nearest neighbors **KNN**)
- **DL** complex **ANN** with serval hidden layers.

In Case one of three data modalities is missing, mask it as zeros when using **DL**.

Results

DL Models outperform **shallow models** except in **SNP** + **MRI** data modality, and in **EHR** only **DT** and **RF** are somewhat similar to **DL**.

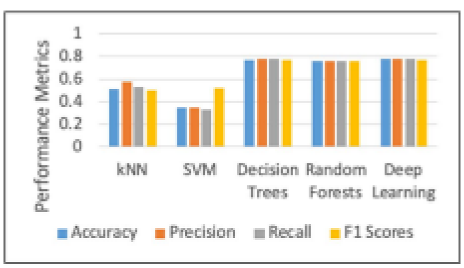
DL methods become better in multi-modalities data than single data-modality.

DL methods are capable of extracting top performing features.

To evaluate Models performance, using **Accuracy**, **Precision**, **Recall** and **meanF1** matrices, and separate Dataset is into

- 90% internal (training) and make ***tenfold cross-validation***
- 10% external (testing)

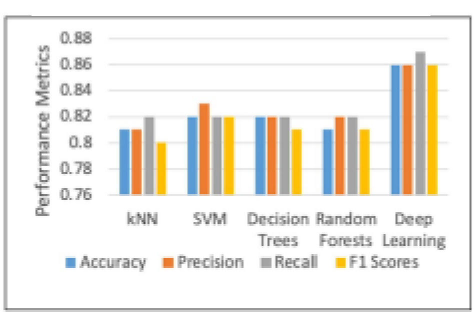
Internal Cross-Validation test



External test DL			
Accuracy	**Precision**	**Recall**	**F1 Score**
0.76%	0.76%	0.77%	0.76%

a. EHR Model

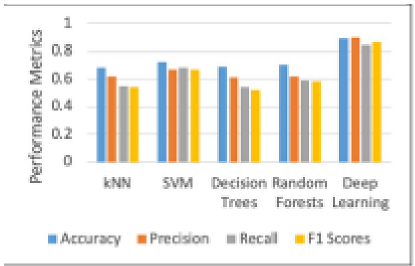
Internal Cross-Validation test



External test DL			
Accuracy	**Precision**	**Recall**	**F1 Score**
0.84%	0.83%	0.83%	0.83%

b. MRI Model

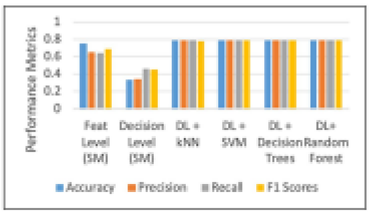
Internal Cross-Validation test



External test DL			
Accuracy	**Precision**	**Recall**	**F1 Score**
0.66%	0.66%	0.57%	0.53%

c. SNPs Model

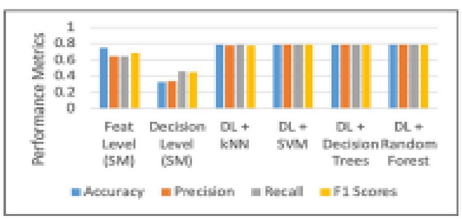
Internal Cross-Validation test



External test DL			
Accuracy	**Precision**	**Recall**	**F1 Score**
0.78%	0.77%	0.78%	0.78%

d. HER + **MRI** + **SNPs** Model

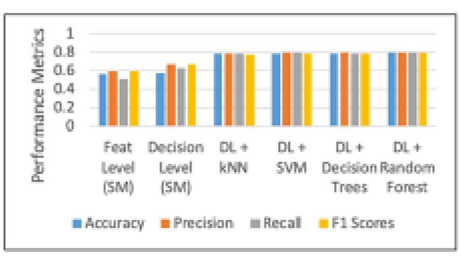
Internal Cross-Validation test



External test DL			
Accuracy	**Precision**	**Recall**	**F1 Score**
0.78%	0.78%	0.79%	0.78%

e. HER + **SNPs** Model

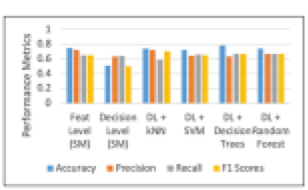
Internal Cross-Validation test



External test DL			
Accuracy	**Precision**	**Recall**	**F1 Score**
0.77%	0.76%	0.77%	0.77%

f. HER + **MRI** Model

Internal Cross-Validation test



External test SM			
Accuracy	**Precision**	**Recall**	**F1 Score**
0.63%	0.62%	0.57%	0.56%

g. SNPs + **MRI** Model

Figure 2. Performance measures

Limitations **DL** Models can't perform well in **small** Dataset such as **SNPs** + **MRI** combination 220 *samples*.

Summary

J.V., contributed to the study design, the pre-processing, data analysis for the EHR data, the combination of the three data modalities, and the writing of the manuscript**.**

L.T., contributed to the pre-processing and analysis of the SNP data, the writing of the manuscript.

H.H., contributed to the image processing pipeline and writing of the results pertaining to image processing.

Prof. M.D.W., contributed to the study design, result evaluation, and extensive relevant and the revision of the manuscript.

Preprocessing

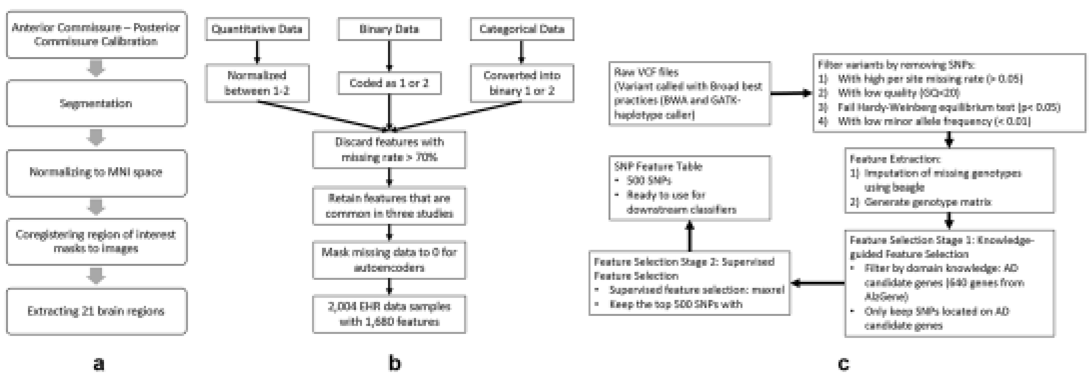


Figure 3. Preprocessing for multi-modality data,

a. MRI data, **b. HER** data and **c. SNPs**

Feature Extraction

Input EHR and **SNP** with noise

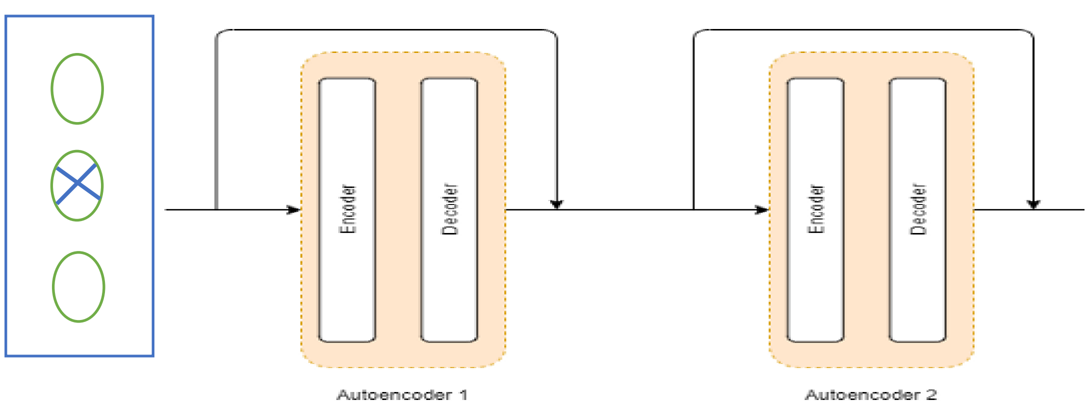


Figure 4. extract features from **EHR and SNP** using **stacked denoising auto-encoder network**

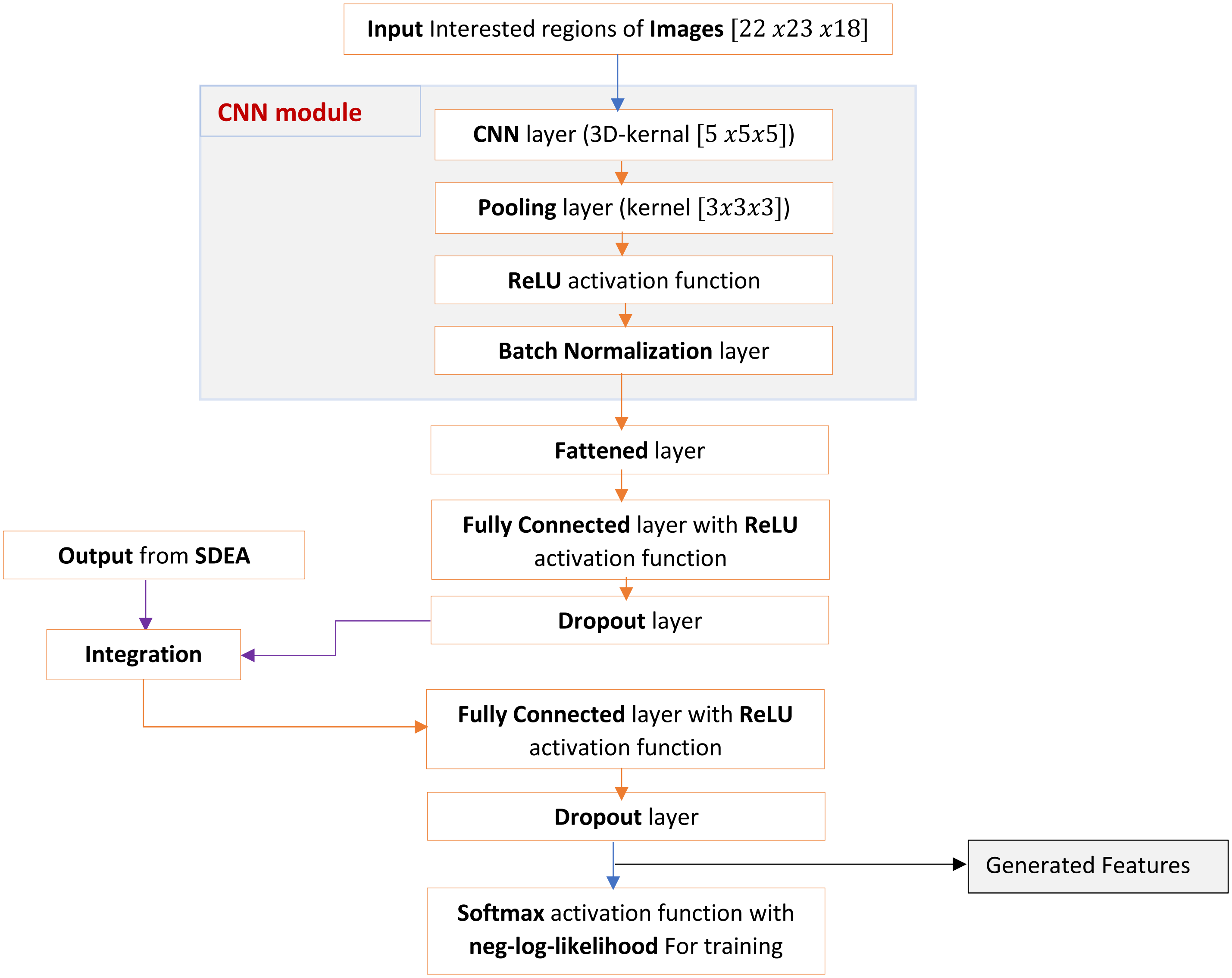


Figure 5. extract features from Images using **3D- Convolutional Neural Network Architecture (CNN).**

Integration with training.

References

Reference	Data Modeling	Methods	Dataset	Objective
[Multimodal deep learning models for early detection of Alzheimer's disease stage | Scientific Reports (nature.com)]	**MRI-SNPs-HER**	**3D CNN – SDAE – SM – DT – RF - KNN**	**ADNI**	Distinguish between 3-stages of ***AD (CN, MRI, AD)***