Multi-modal detection of Alzheimer's Disease

Sahasra Kovvuru, Gautam Ganesh & Krish Chhugani

What is Alzheimer’s Disease?

Alzheimer's disease is a progressive brain disorder that gradually damages memory, thinking ability, and behavior. It is the most common cause of dementia, especially in older adults.

- Alzheimer’s disease is the most common form of dementia, accounting for roughly 80% of all dementia cases.
- Early detection improves treatment and patient care.
- Combining MRI and clinical data can improve prediction accuracy compared to single-modality systems.
- Existing Alzheimer’s detection systems often rely only on MRI data and use computationally expensive models, limiting practical clinical deployment.

Fun Fact: A new case of dementia arises somewhere in the world every 3 seconds

20
Number of years you could have the disease before symptoms appear

6th
Alzheimer’s disease is the sixth leading cause of death in the U.S.

1/10
Number of people over 65 who have Alzheimer’s

Potential Applications

- Can be used for early Alzheimer’s screening in hospitals and diagnostic centers using MRI scans and clinical data.
- Helps doctors and radiologists by providing AI-assisted diagnosis and faster patient analysis.

Potential Impact

- Early detection can improve treatment planning and patient care.
- The lightweight ResNet-18 model makes the system more practical for real-world healthcare applications.

PAPER-1

Methodology:

- Uses transfer learning in 3D CNNs, which allows the transfer of knowledge from 2D image datasets (ImageNet) to a 3D image dataset.

Observations:

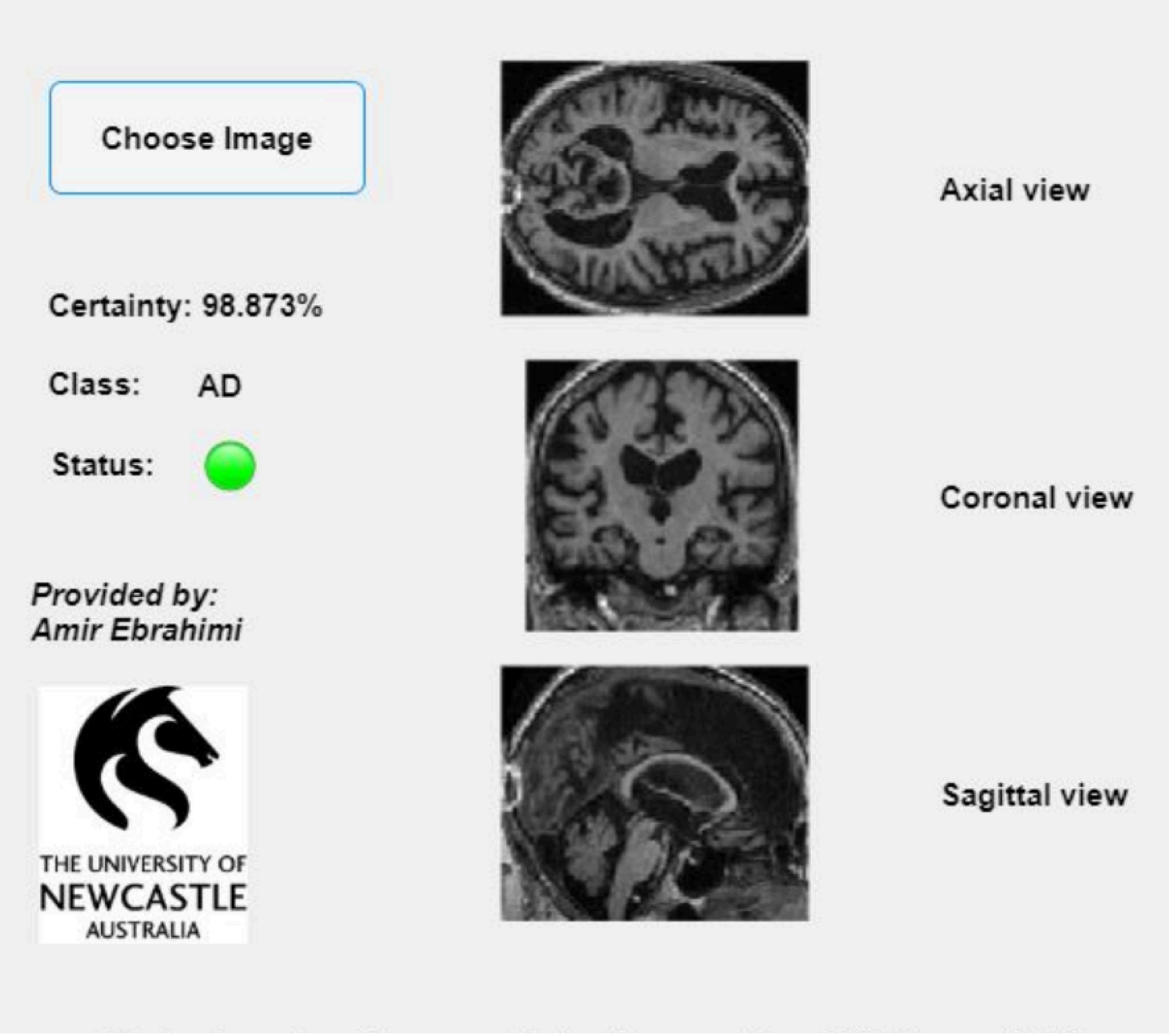
- To build 3D ResNet-18, 2D filters of 2D ResNet-18 were extended in the third dimension to have 3D filters.
- Then, the entire MRIs were used for training 3D ResNet-18 to make one decision per person.

Performance Metrics:

- Accuracy: 96.88% , Sensitivity: 100%, Specificity: 93.75% on ADNI dataset.

Analysis:

- Introducing transfer learning to a 3D CNN improves an AD detection system's accuracy



Reference:
Ebrahimi, A., Luo, S., & Chiong, R.
"Introducing Transfer Learning to 3D ResNet-18 for Alzheimer's Disease Detection on MRI Images" 2020 International Conference on Image and Vision Computing New Zealand (IVCNZ).

PAPER-2

Methodology:

- Can model the longitudinal analysis using RNN from the imaging data at various time points.

Observations:

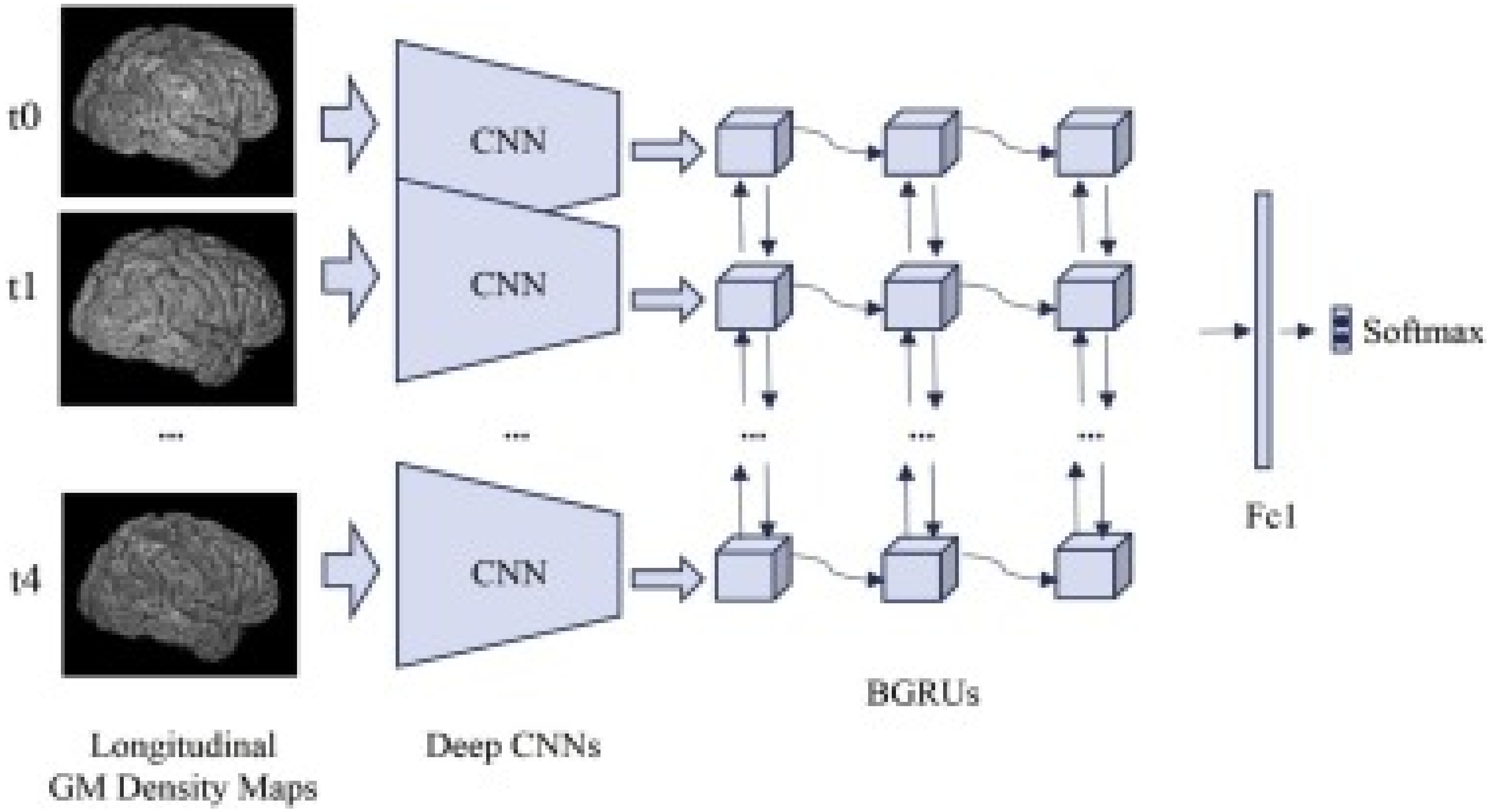
- Jointly learns the spatial and longitudinal features and disease classifier, which can achieve optimal performance

Performance Metrics:

- Classification accuracy of 91.33% for AD vs. NC

Analysis:

- RNN's are used for longitudinal analysis for the image data at various time points
- longitudinal analysis of sequential data yield better results for progression of AD.



Reference:
Cui, R., & Liu, M.
"RNN-based longitudinal analysis for diagnosis of Alzheimer's disease"
Computerized Medical Imaging and Graphics, 2019.

PAPER-3

Methodology:

- this paper introduces DeepALZNET, a dual-pathway computational framework designed to enhance AD prediction by offering two independent processing pathways

Observations:

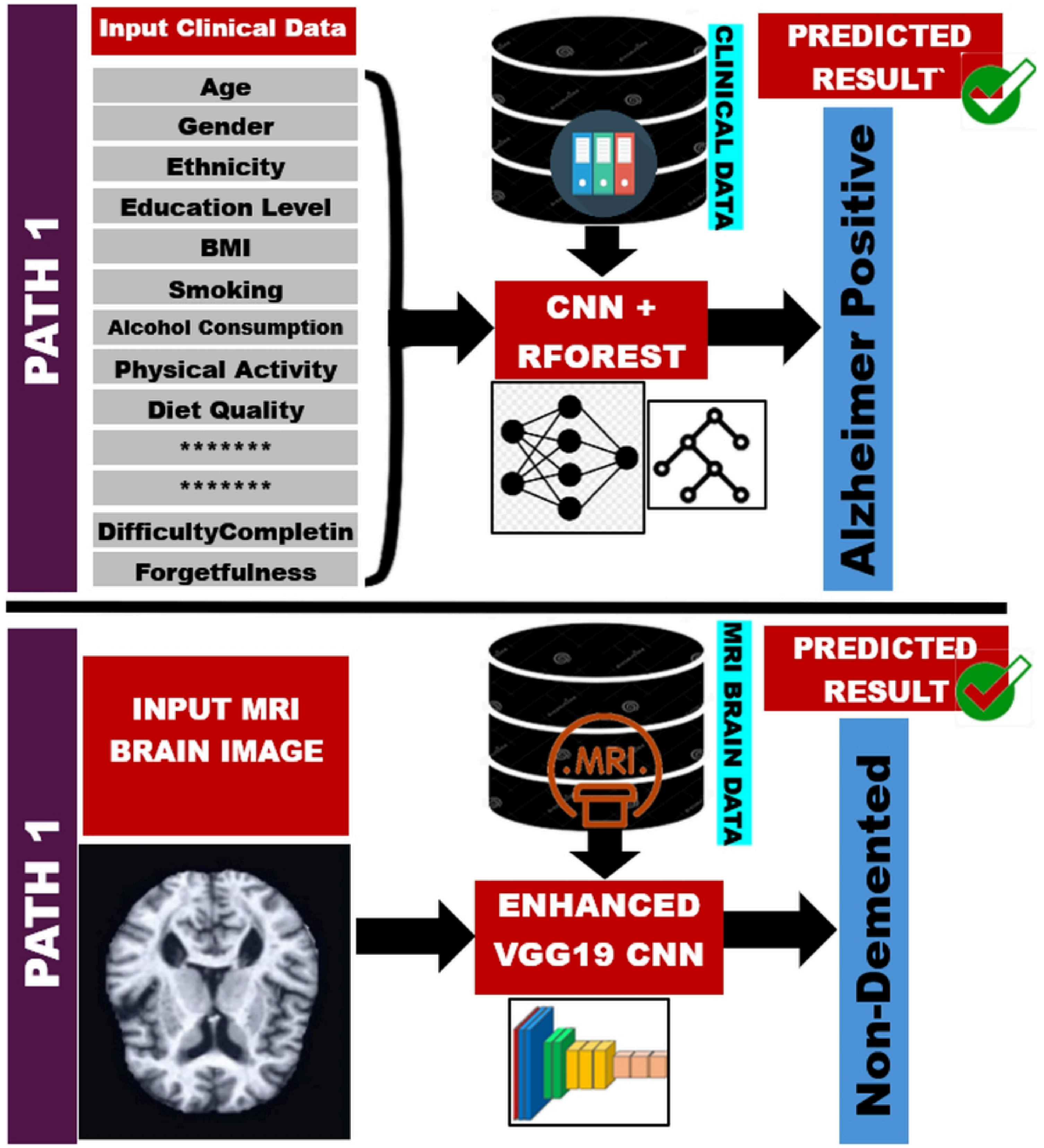
- one path for structured clinical data and another for unstructured brain MRI scans.
- combines 1-D CNN + VGG19 for transfer learning

Performance Metrics:

- both pathways: accuracy >90%

Analysis:

- Improved feature learning through multimodal fusion.
- Achieved better classification performance than single-modal approaches.



Reference:
Bekhet, S., Saad, N., & Farag, S.
"Alzheimer disease predicting from clinical and MRI data using DeepALZNET dual pathwa
Scientific Reports, 2025.

Research Gap:

From the literature survey, it is observed that:

- Many studies rely only on MRI image features.
- Clinical indicators such as MMSE, CDR, Age, and Education are often underutilized.
- Advanced architectures such as 3D CNNs increase computational complexity and training cost.

Proposed work

→ Our work extends these approaches by combining lightweight ResNet-18 image features with clinical data using multimodal learning.

The proposed system aims to:

- Improve classification performance through multimodal learning,
- Reduce computational complexity using a lightweight residual network,
- Develop a more practical and efficient framework for real-world clinical screening applications.

Dataset

→ Procured from Washington University Medicine

→ OASIS-2 Longitudinal MRI Dataset, the dataset is publicly available and anonymized to protect patient privacy.

→ Min visits: 1, Max visits: 5

Property	Value
Subjects	150
Age Range	60 – 96
Total MRI Sessions	373
MRI Type	T1-weighted*
Scans per session	3–4

Category	Count
Non-Demented	72
Demented	64
Converted to Demented	14

Dataset

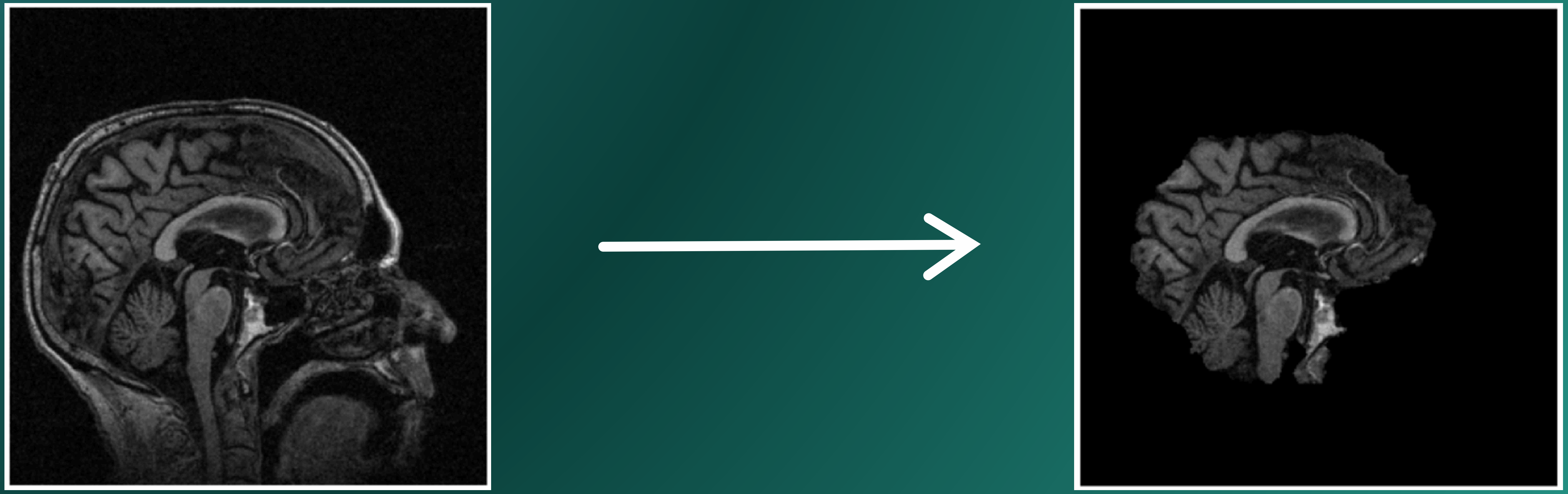
→ Clinical Features:

- Age
- Gender
- Education
- SES
- Time Gap(MR delay)
- MMSE score
- CDR score
- Brain volume

```
CLINICAL_COLS = [
    'MMSE',
    'eTIV',
    'nWBV',
    'ASF'
]

STATIC_COLS = [
    'EDUC',
    'SES',
    'M/F_M'
]

FEATURE_COLS = (
    CLINICAL_COLS +
    STATIC_COLS
)

TARGET_COL = 'CDR'
```

Preprocessing

→ For preprocessing the MRI images, we used DeepBET for brain extraction.



Preprocessing

→ Preprocessing of clinical features

→ converted MR delay from days to years → imputed missing values → standardized: eTIV, nWBV, ASF

before:

Subject ID	MRI ID	Group	Visit	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
OAS2_0001	OAS2_0001_MR1	Nondemented	1	0	M	R	87	14	2	27	0	1987	0.696	0.883
OAS2_0001	OAS2_0001_MR2	Nondemented	2	457	M	R	88	14	2	30	0	2004	0.681	0.876
OAS2_0002	OAS2_0002_MR1	Demented	1	0	M	R	75	12		23	0.5	1678	0.736	1.046
OAS2_0002	OAS2_0002_MR2	Demented	2	560	M	R	76	12		28	0.5	1738	0.713	1.010
OAS2_0002	OAS2_0002_MR3	Demented	3	1895	M	R	80	12		22	0.5	1698	0.701	1.034
OAS2_0004	OAS2_0004_MR1	Nondemented	1	0	F	R	88	18	3	28	0	1215	0.710	1.444
OAS2_0004	OAS2_0004_MR2	Nondemented	2	538	F	R	90	18	3	27	0	1200	0.718	1.462
OAS2_0005	OAS2_0005_MR1	Nondemented	1	0	M	R	80	12	4	28	0	1689	0.712	1.039
OAS2_0005	OAS2_0005_MR2	Nondemented	2	1010	M	R	83	12	4	29	0.5	1701	0.711	1.032
OAS2_0005	OAS2_0005_MR3	Nondemented	3	1603	M	R	85	12	4	30	0	1699	0.705	1.033
OAS2_0007	OAS2_0007_MR1	Demented	1	0	M	R	71	16		28	0.5	1357	0.748	1.293
OAS2_0007	OAS2_0007_MR3	Demented	3	518	M	R	73	16		27	1	1365	0.727	1.286

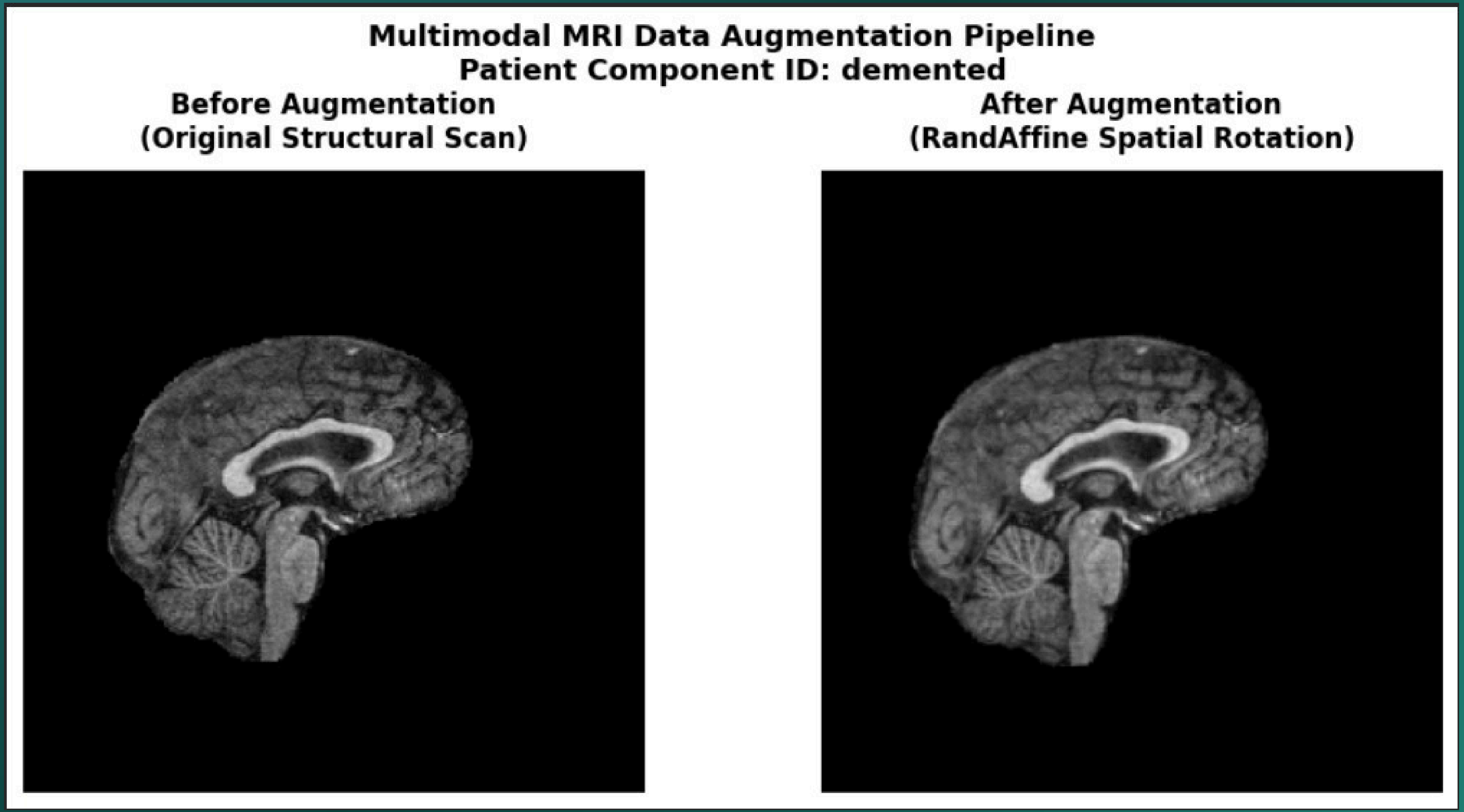
after:

Subject ID	MRI ID	Group	Visit	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF	Clean_ID	Severity
OAS2_000	OAS2_000	Nondeme	1	0	0	R	87	14	2	27	0	2.833595	-0.90182	-2.26232	OAS2_000	0
OAS2_000	OAS2_000	Nondeme	2	1.251198	0	R	88	14	2	30	0	2.935525	-1.30741	-2.31961	OAS2_000	0
OAS2_000	OAS2_000	Dementec	1	0	0	R	75	12	2	23	0.5	1.081119	0.182804	-1.08578	OAS2_000	1
OAS2_000	OAS2_000	Dementec	2	1.533196	0	R	76	12	2	28	0.5	1.418413	-0.43551	-1.34469	OAS2_000	1
OAS2_000	OAS2_000	Dementec	3	5.188227	0	R	80	12	2	22	0.5	1.192666	-0.76351	-1.17341	OAS2_000	1
OAS2_000	OAS2_000	Nondeme	1	0	1	R	88	18	3	28	0	-1.55084	-0.54039	1.802488	OAS2_000	0
OAS2_000	OAS2_000	Nondeme	2	1.472964	1	R	90	18	3	27	0	-1.63742	-0.30578	1.935319	OAS2_000	0
OAS2_000	OAS2_000	Nondeme	1	0	0	R	80	12	4	28	0	1.139618	-0.48673	-1.13203	OAS2_000	0
OAS2_000	OAS2_000	Nondeme	2	2.765229	0	R	83	12	4	29	0.5	1.208652	-0.49881	-1.1858	OAS2_000	1
OAS2_000	OAS2_000	Nondeme	3	4.388775	0	R	85	12	4	30	0	1.200386	-0.65985	-1.1794	OAS2_000	0
OAS2_000	OAS2_000	Dementec	1	0	0	R	71	16	2	28	0.5	-0.74356	0.499563	0.707073	OAS2_000	1
OAS2_000	OAS2_000	Dementec	3	1.418207	0	R	73	16	2	27	1	-0.7028	-0.06804	0.657769	OAS2_000	2
OAS2_000	OAS2_000	Dementec	4	3.507187	0	R	75	16	2	27	1	-0.6595	-0.52847	0.606051	OAS2_000	2

Data Augmentation:

- Performed small rotations of 0.05 radians.
- Did small magnifications
- Performed with certain probability for each training example

Data Augmentation:



Feature Extraction:

- We used the CNN ResNet-18 for feature extraction of our MRI images.
- We got 512 features for each MRI scan
- We applied PCA on this to reduce it 128 features.

```
Epoch 1, Avg Loss: 0.8011
Epoch 2, Avg Loss: 0.4086
Epoch 3, Avg Loss: 0.3076
Epoch 4, Avg Loss: 0.2125
Epoch 5, Avg Loss: 0.2528
Epoch 6, Avg Loss: 0.2355
Epoch 7, Avg Loss: 0.1485
Epoch 8, Avg Loss: 0.1363
Epoch 9, Avg Loss: 0.1510
Epoch 10, Avg Loss: 0.1515
Epoch 11, Avg Loss: 0.0817
Epoch 12, Avg Loss: 0.0701
Epoch 13, Avg Loss: 0.0844
Epoch 14, Avg Loss: 0.0357
Epoch 15, Avg Loss: 0.0632
```

```
Scans per class:
     non_demented: 692
         demented: 542
        converted: 134
```

```
Train patients: 120
Test patients : 30
Train scans: 1093
Test scans : 275
```

```
torch.Size([1, 512])
```

Method 1: ResNet-18 + T-LSTM

- We did feature extraction using ResNet-18 and split the features into Train + Test (80/20 split)
- Created temporal sequences of the visits and trained the train features on T-LSTM model.
- Tested the test features on T-LSTM model.
- Regression model, predicts continuous CDR scores

Method 1: Performance Metrics

```
TEST RESULTS

MAE  : 0.2265
RMSE : 0.2946
R²   : 0.0917
```

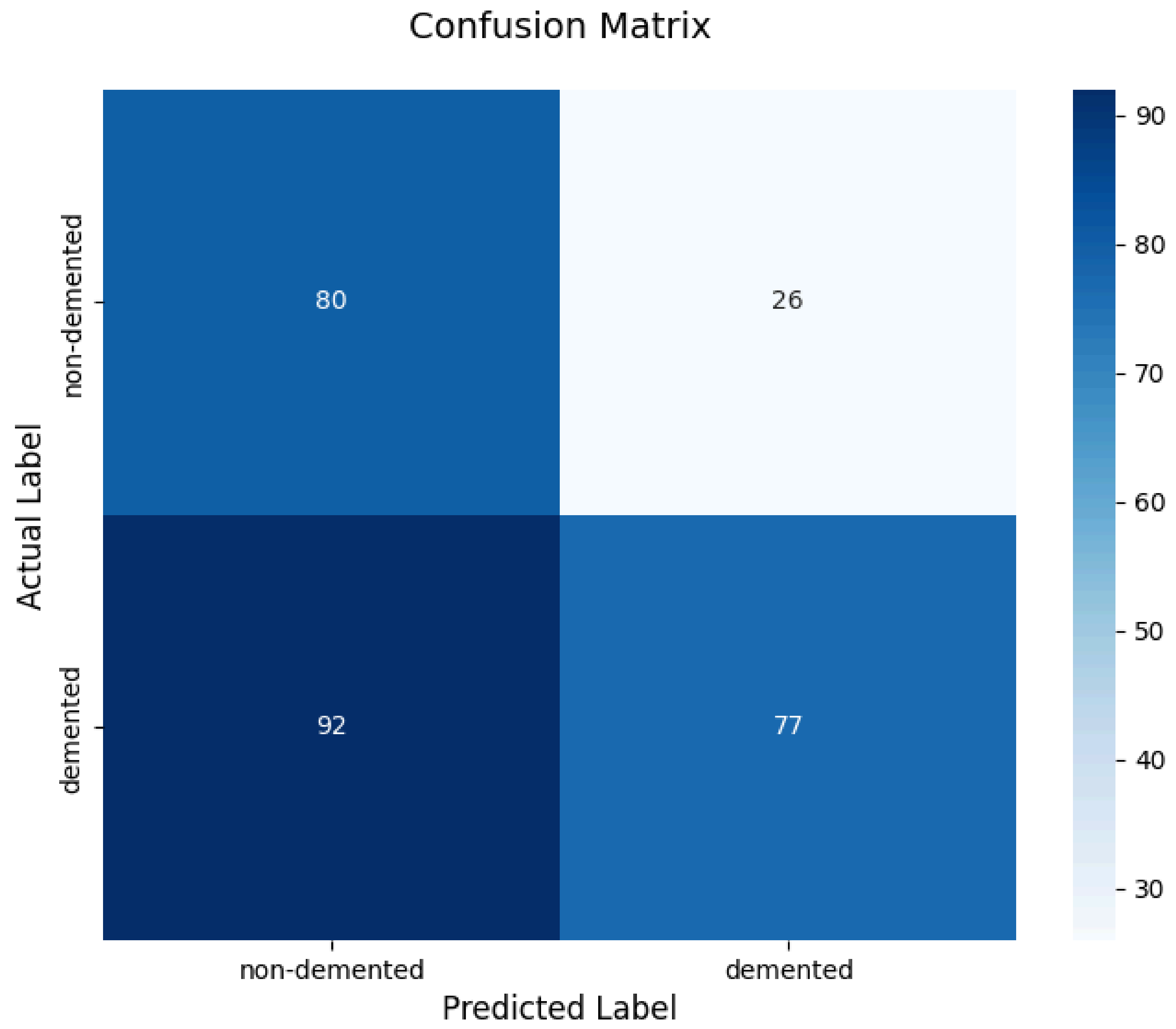
Method 1: Challenges

- Model gave bad results, R^2 value was extremely low.
- Temporal sequences created was padded with zeroes due to lack of data.
- Which might have led to bad predictions hence, bad model.

Method 2: ResNet-18

- We unfroze last 2 hidden layers of ResNet 18 to extract 512 features.
- We used 5-fold cross validation for testing.
- Changed ResNet 18's classification layer to classify the 512 features into 2 categories:
  (Demented and Non-Demented)

Method 2: Performance Metrics



Method 2: Performance Metrics

Classification Report:

	precision	recall	f1-score	support
0	0.47	0.75	0.58	106
1	0.75	0.46	0.57	169
accuracy			0.57	275
macro avg	0.61	0.61	0.57	275
weighted avg	0.64	0.57	0.57	275

0 → Demented

1 → Non-Demented

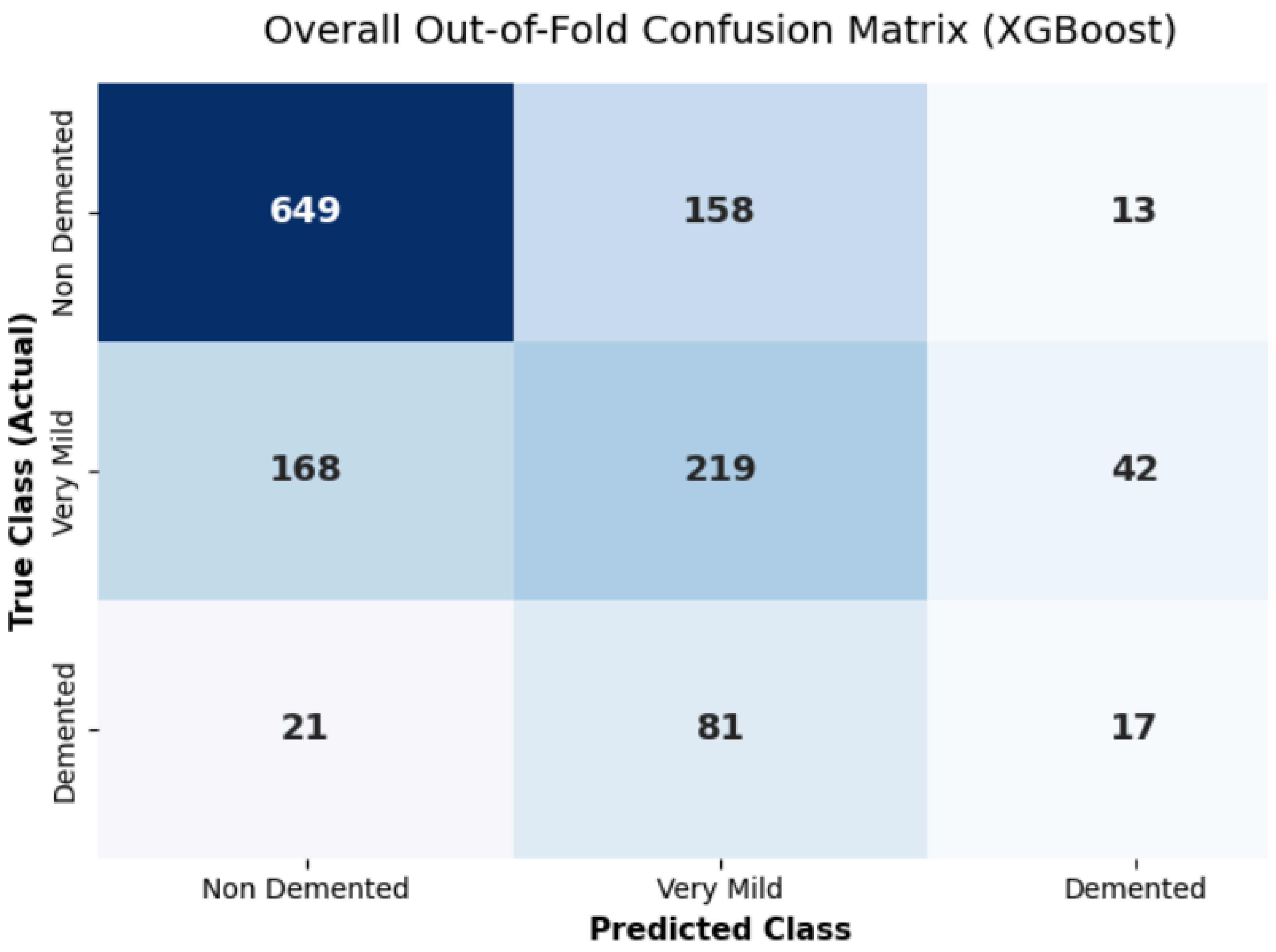
Method 2: Challenges

- Heavily biased toward predicting demented as we have 92 False Positives.
- Low accuracy and precision, misdiagnoses healthy patients too frequently.
- The cause may be that the extracted features do not provide enough information to confidently separate dementia from normal aging

Method 3: ResNet-18 + XGBoost

- We did feature extraction using ResNet-18 and K-Fold Group Stratified CV
- We also did augmentation on the MRI images, small rotations and zoom.
- Got 64-dimensional feautres in the end and concatenated with clinical data
- Combined features classified using XGBoost

Method 3: ResNet–18 + XGBoost



Method 3: ResNet-18 + XGBoost

	precision	recall	f1-score	support
Non Demented	0.7745	0.7915	0.7829	820
Very Mild	0.4782	0.5105	0.4938	429
Demented	0.2361	0.1429	0.1780	119
accuracy			0.6469	1368
macro avg	0.4962	0.4816	0.4849	1368
weighted avg	0.6347	0.6469	0.6396	1368

```
XGBoost Accuracy:
0.7073170731707317
```

BEST MODEL PERFORMANCE

Overall Challenges

- Limited Longitudinal Data: Few MRI visits per patient
- Class Imbalance: Way more Non-Demented samples than Demented
- High-Dimensional MRI Data: Heavy preprocessing and computational cost
- Multimodal Fusion Complexity: Combining MRI and clinical features effectively
- Model Generalization: Risk of overfitting because of minimal data

Deployment:

- The model can be integrated into a software platform where MRI scans and patient clinical information are uploaded for automated analysis and classification.

- It can support researchers, healthcare professionals, and collaborative medical studies.

Challenges while scaling:

- Collecting large and diverse MRI datasets while maintaining patient privacy.
- High computational and storage requirements for medical imaging data.
- Ensuring consistent model accuracy across different hospitals and MRI machines.

Conclusion & future work

Conclusion

- Multimodal learning improved Alzheimer’s classification.
- ResNet-18 + XGBoost achieved the best performance out of the 3 models we used.
- Clinical + MRI features (method 3) performed better than MRI-only methods (method 2).

Future Work

- Use larger datasets
- Improve minority class prediction
- Explore 3D CNNs
- Real-time hospital deployment

Thank you!


Man with Alzheimer's forgets he had Alzheimer's, remembers everything
That's not how Alzheimer's works
made with mematic