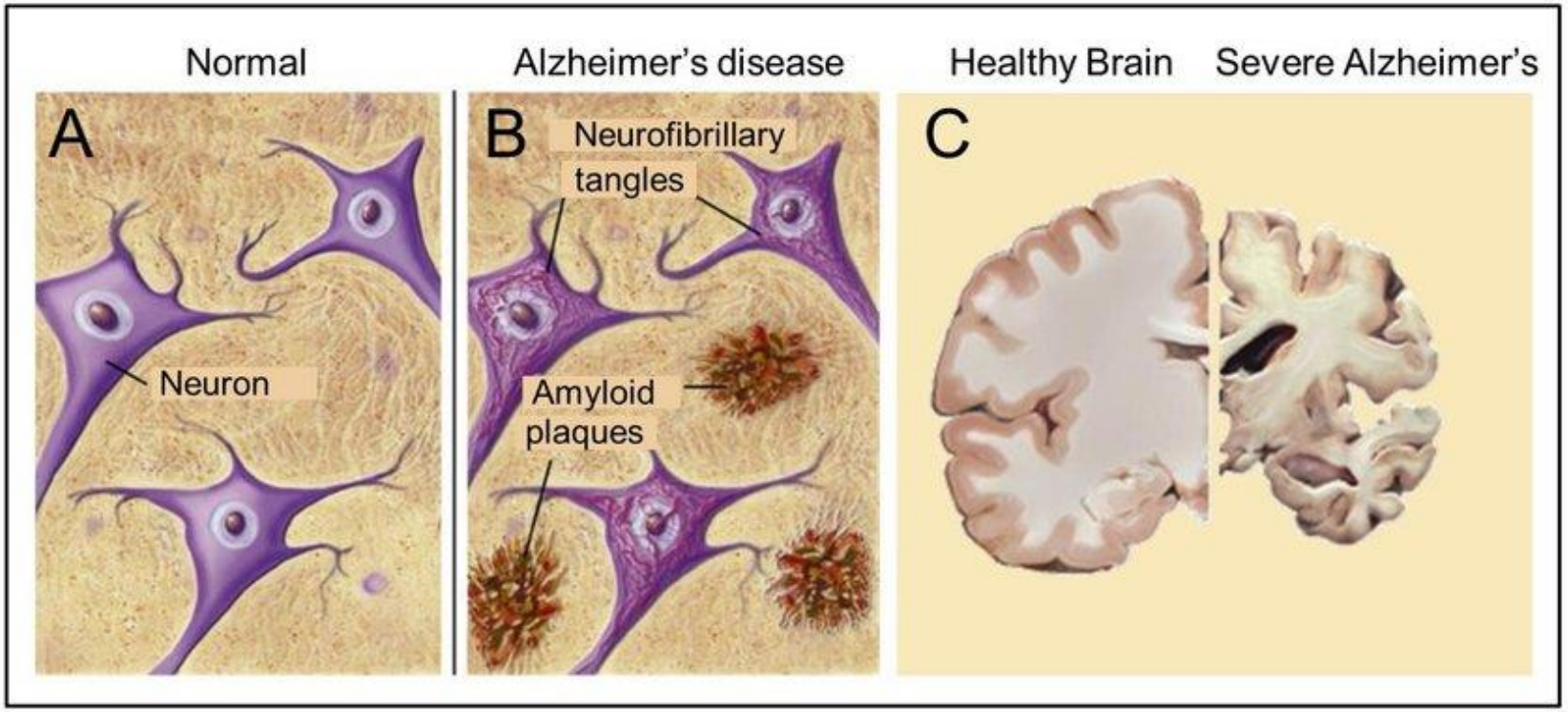
Multi-Modal Deep Learning Architectures for Detecting Alzheimer's Disease Stages

Mehmet Saygin Seyfioglu
10/25/2021

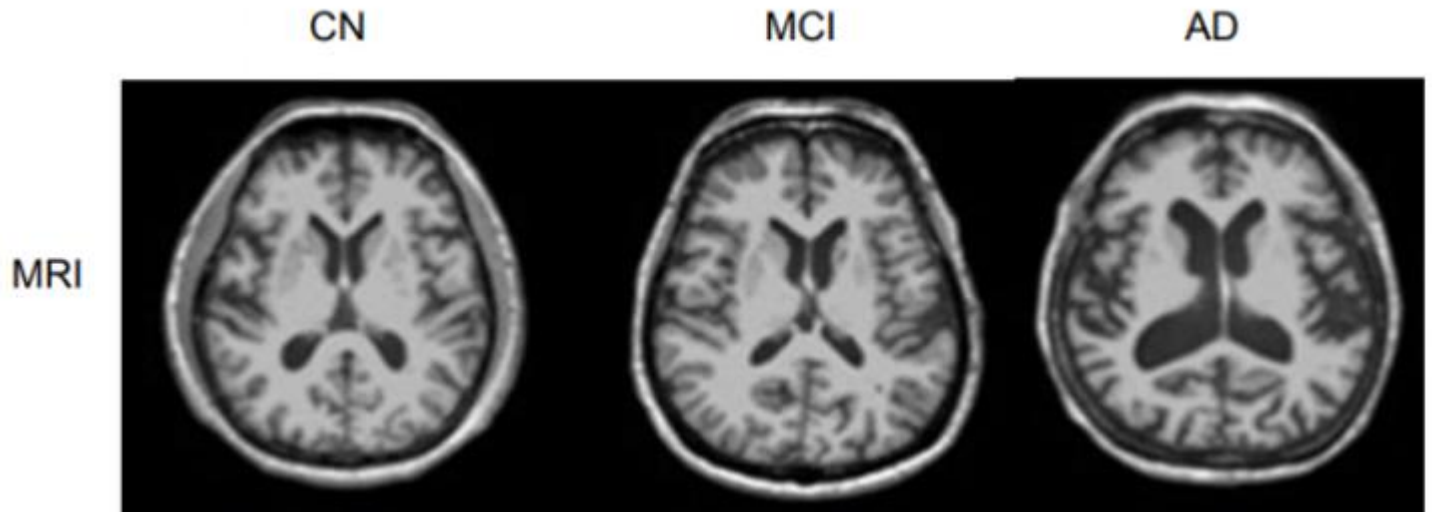
UNIVERSITY of WASHINGTON

Alzheimer's Disease (AD)

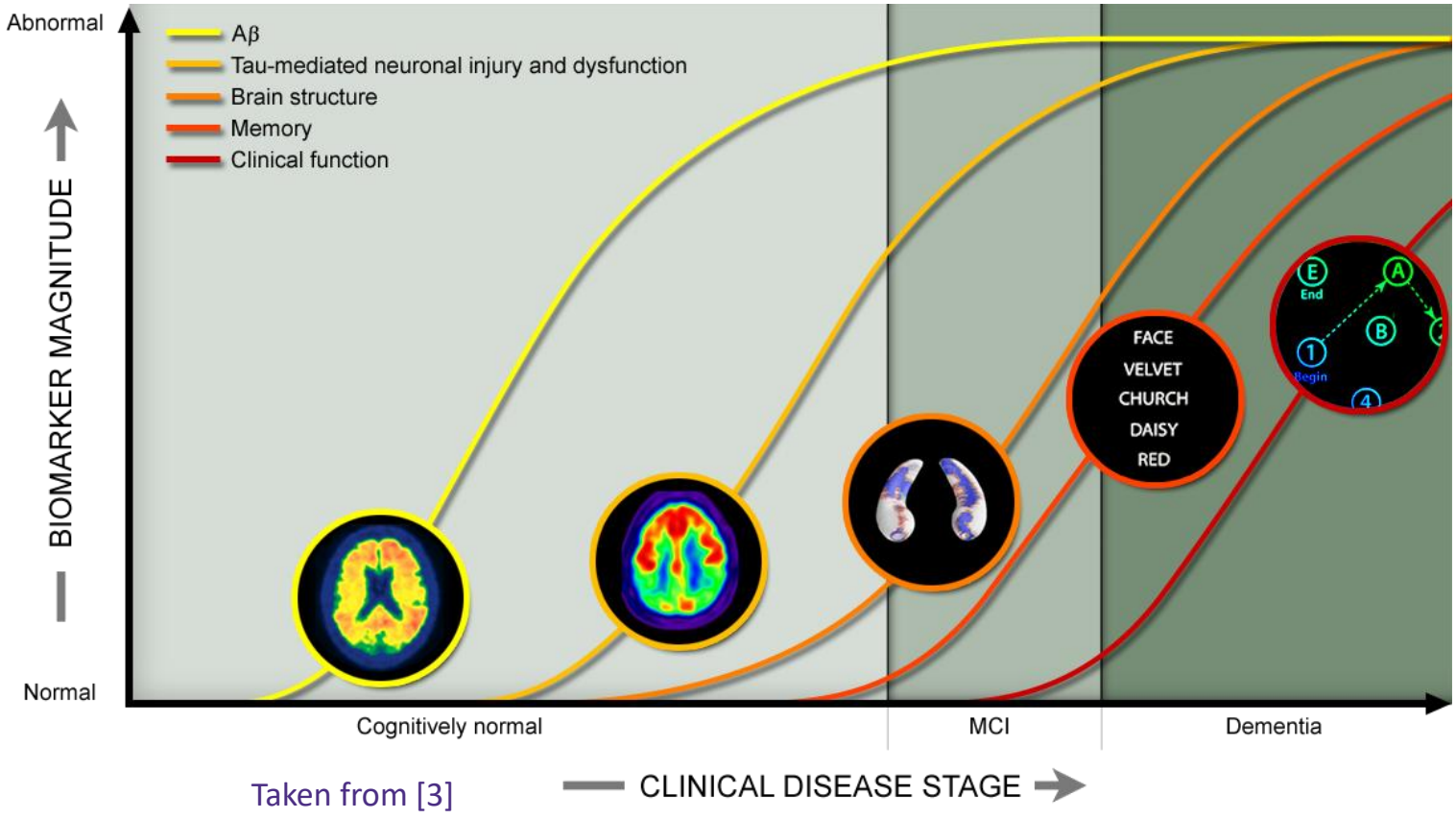


Taken from [1]

Alzheimer's Disease (Cont'd)



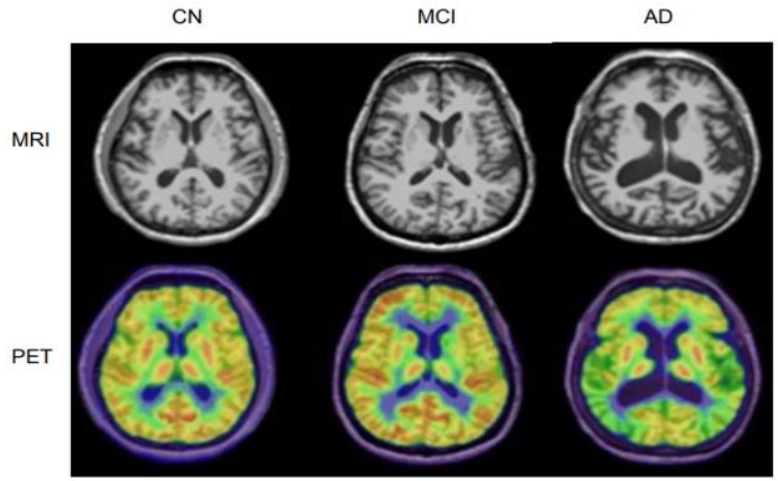
Disease progression



Taken from [3]

Biomarkers to Detect the Alzheimer's Disease

- Magnetic Resonance Imaging (MRI) provides anatomical brain structure information
- Positron Emission Tomography (PET) provides functional brain information



Taken from [2]

- Single Nucleotide Polymorphism (SNP) provides congenital disease risk information
- There are also Spinal Taps that measure the amount of amyloid and tau and Cognitive Tests measure memory and cognition
  - Hard to detect MCI!

Motivation and Problem

- **Detect the MCI stage to slow down the AD progression!**
  - **Exploiting complementary information from MRI, PET and SNP data could enhance prediction performance on AD stage detection**
    - **Humans are not good at analyzing higher order correlations between biomarkers**
      - **Suitable task for Deep Neural Networks (DNNs)!**
- **However, it is not straightforward to use multi modal data in training DNNs.**
  - **Hard to generalize using a small amount of training sample data**
  - **Not straightforward to exploit complementary information between biomarkers**

Data

- **The Alzheimer's Disease Neuroimaging Initiative (ADNI)**
  - **A global longitudinal initiative consists cognitive, imaging, biochemical and genetic biomarkers**
- **Labels for CN, MCI and AD**
- **MRI**
  - **Structural MRI, 3D T1 Scans**
- **PET**
  - **Fludeoxyglucose (FDG) Pet**
- **Genetic**
  - **Genome-wide Association Studies (GWAS) Data including 620k SNPs overall**

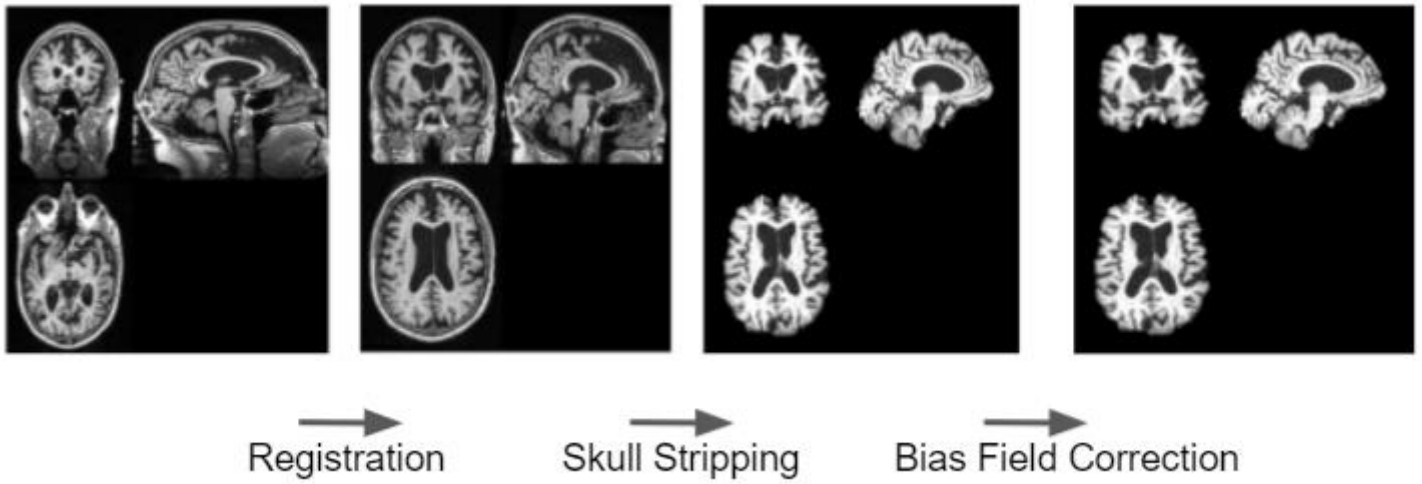
Modality	MRI	PET	SNP	MRI & PET	MRI & SNP	PET&SNP	ALL
# of Patients	805	360	867	360	737	360	360

- **From 360 Patients, we have 1296 visits for all 3 modalities distributed roughly balanced between the 3 classes.**

Preprocessing

- MRI Preprocessing
- PET Preprocessing
- SNP Preprocessing

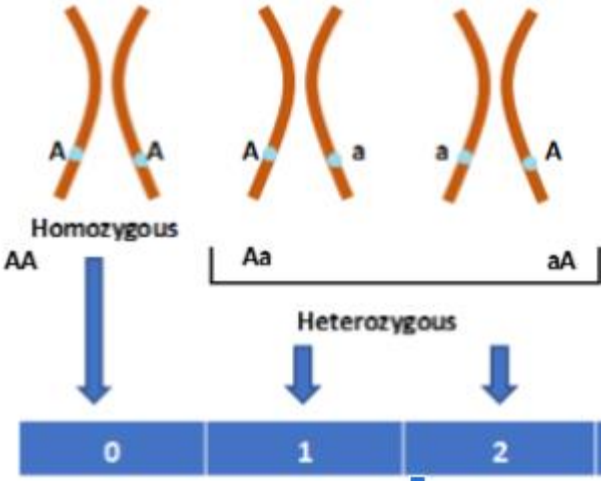
MRI and Pet Preprocessing Pipeline



The final image is 180x180x180

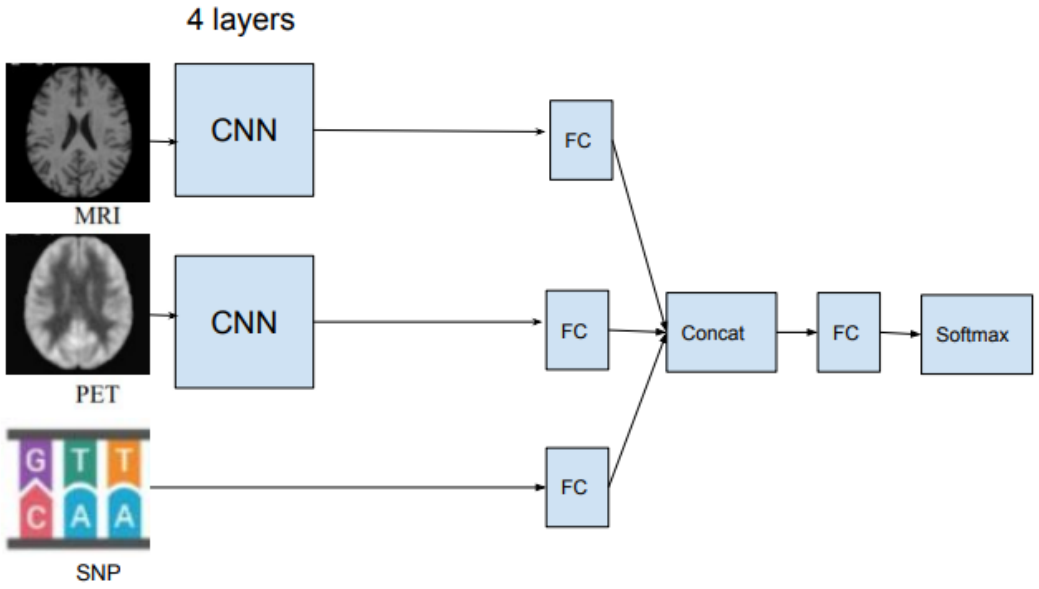
SNP Preprocessing Pipeline

- We first performed SNP quality control filtering.
- The International Genomics of Alzheimer's Project (IGAP) is then used to obtain the AD-related genes SNPs.
- Then the Genotype Matrix is generated by mapping genotypes to numbers based on allele combinations.

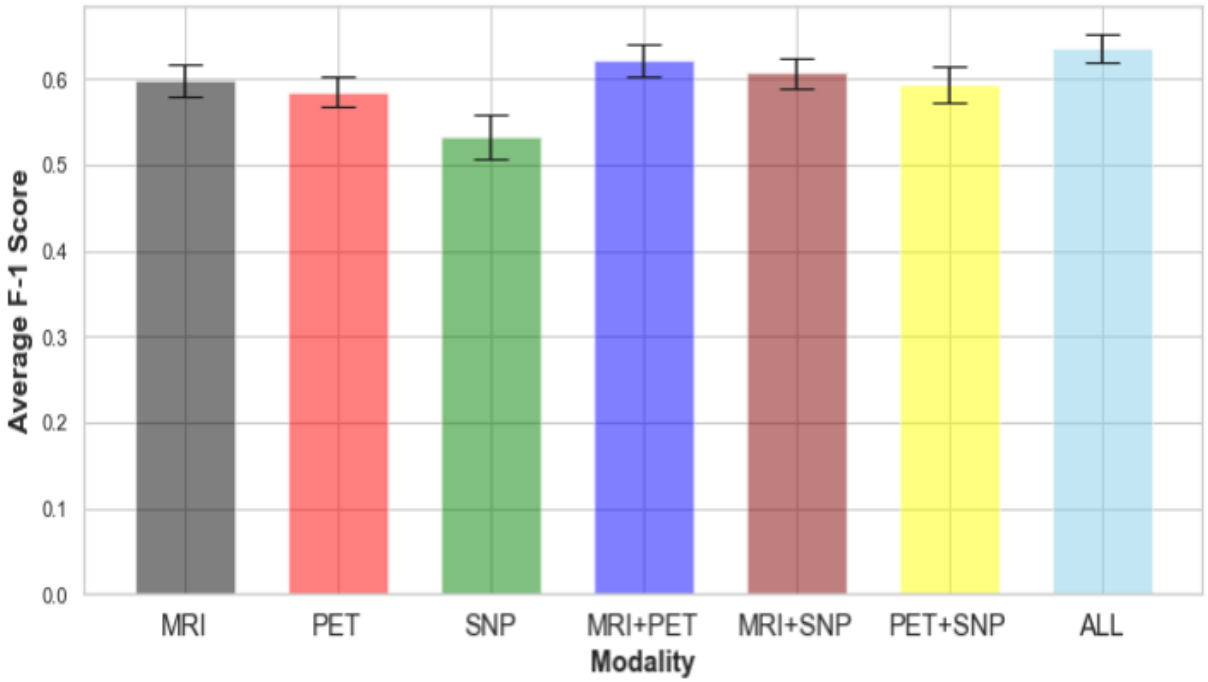


Baseline Model

- **To show that multi modality networks outperform the single modality approach, we built a simple baseline network that uses four layers of 3D CNNs.**



Impact of Modality Combinations

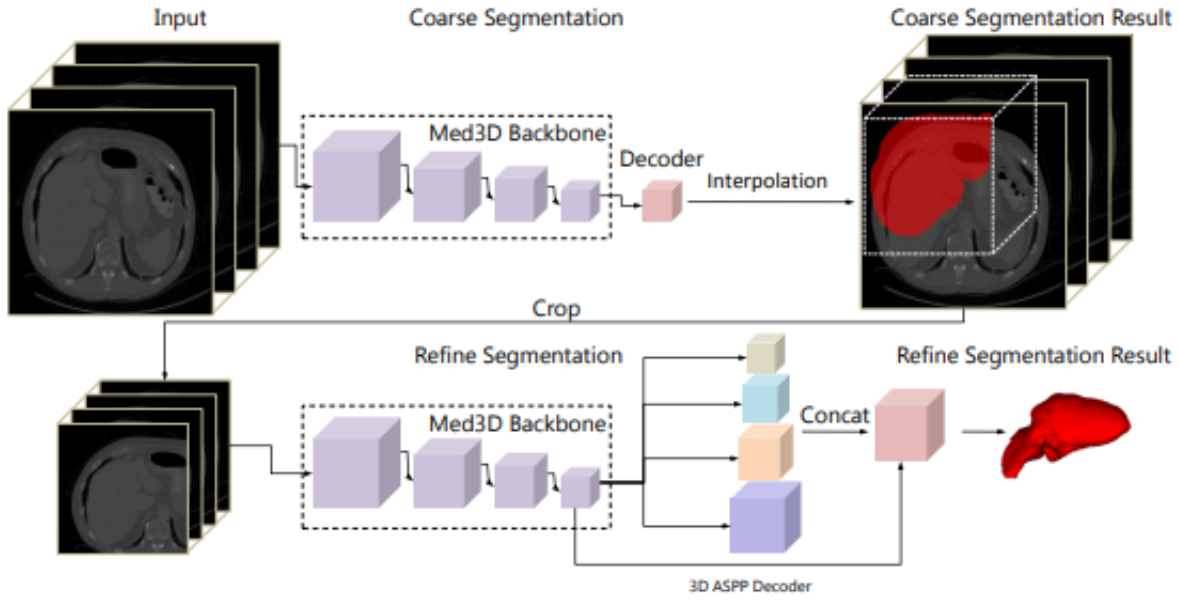


How to Improve

- **Baseline shows that multi modality approach works!**
- **However, the baseline model is quite simplistic because it does not**
  - **Exploit the complementary nature between the modalities.**
  - **Uses a very shallow model**

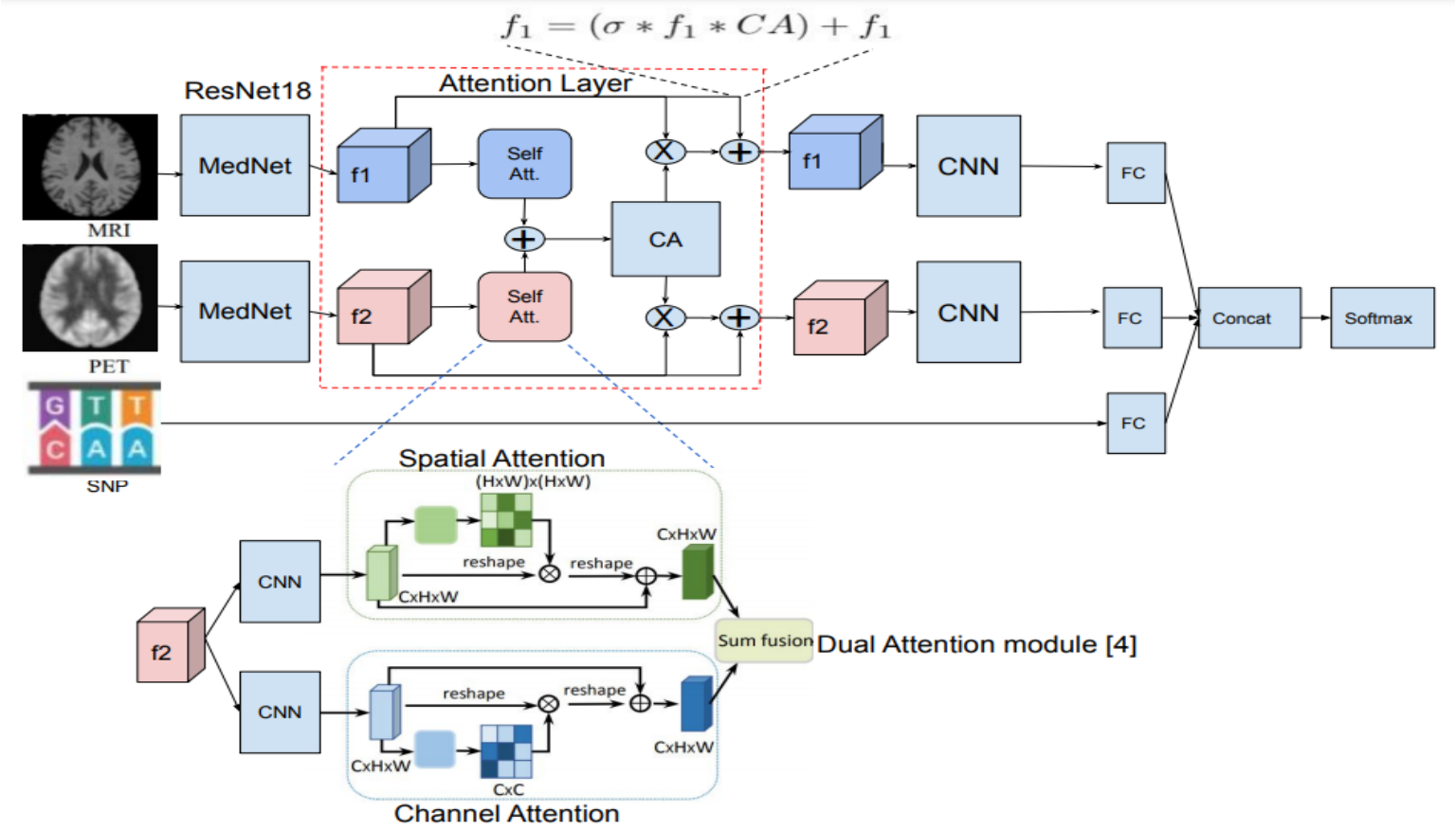
Transfer Learning from MedNet

- **The volume of training samples dramatically affects the performance of deep models, and this is especially the case for 3D CNNs.**
- **We added their convolutional backbone to our model and freeze all layers except for the final convolutional layer**

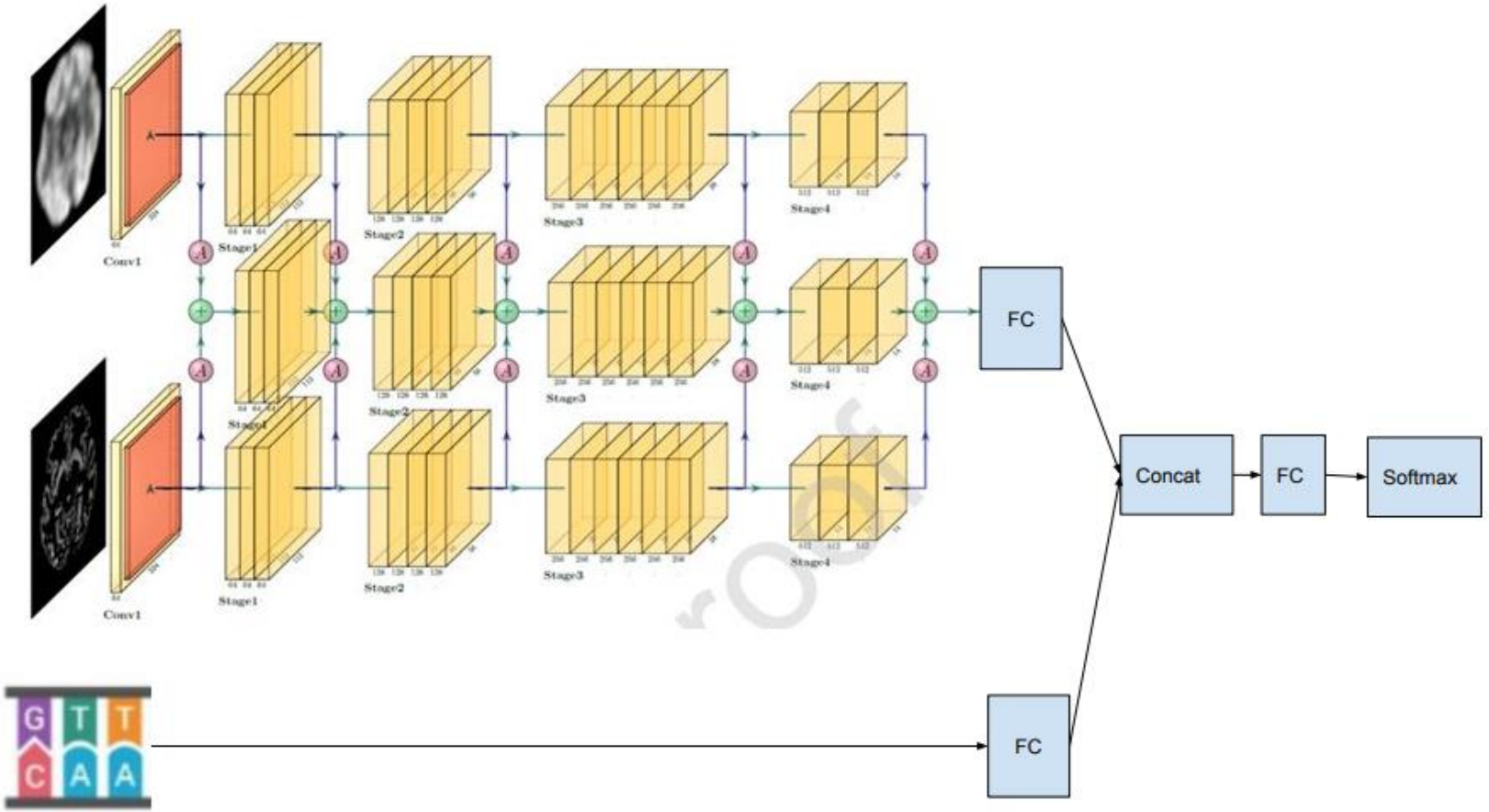


Taken from [6]

Transfer Learning and Complementary Attention



Comprehensive Baseline [7]



Results

Models	Class	Precision	Recall	F1-score
Joint Baseline	**CN**	0.686	0.649	0.667
	MCI	0.483	0.452	0.467
	AD	0.714	0.806	0.758
Zhang el al. [7]	**CN**	0.667	0.684	0.693
	MCI	0.600	0.469	0.526
	AD	0.700	**0.875**	0.778
Ours	**CN**	**0.735**	**0.694**	**0.714**
	MCI	**0.633**	**0.613**	**0.623**
	AD	**0.750**	0.828	**0.787**

Individual Contributions of CA and Transfer Learning

Model	Baseline	Transfer Learning	Attention	Transfer Learning and Attention
Average F1 Score	0.630	0.693	0.669	0.708

Discussion

- **We conclude that using multi-modality data in detecting AD is beneficial even on the simplest baselines.**
  - **MRI seems to be the most "representative" modality**
  - **SNP is the least important modality**
- **We outperformed both simple and comprehensive baselines, especially in detecting MCI.**
  - **Transfer learning and Complementary Attention modules both help greatly!**

References

[1] De Loof, Arnold, and Liliane Schoofs. "Alzheimer's disease: Is a dysfunctional mevalonate biosynthetic pathway the master-inducer of deleterious changes in cell physiology?." *OBM Neurobiology* 3.4 (2019).

[2] T. Thientunyakit, C. Sethanandha, W. Muangpaisan, O. Chawalparit, K. Arunrungvichian, T. Siriprapa, Y. Vichianin, S. Kamal, C. Suppasilp, T. Thongpraparn et al., "Relationships between amyloid levels, glucose metabolism, morphologic changes in the brain and clinical status of patients with alzheimer's disease," Annals of nuclear medicine, pp. 1– 12, 2020.

[3] S. Chen, K. Ma, and Y. Zheng, "Med3d: Transfer learning for 3d medical image analysis," arXiv preprint arXiv:1904.00625, 2019.

[4] Fu, Jun, et al. "Dual attention network for scene segmentation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.

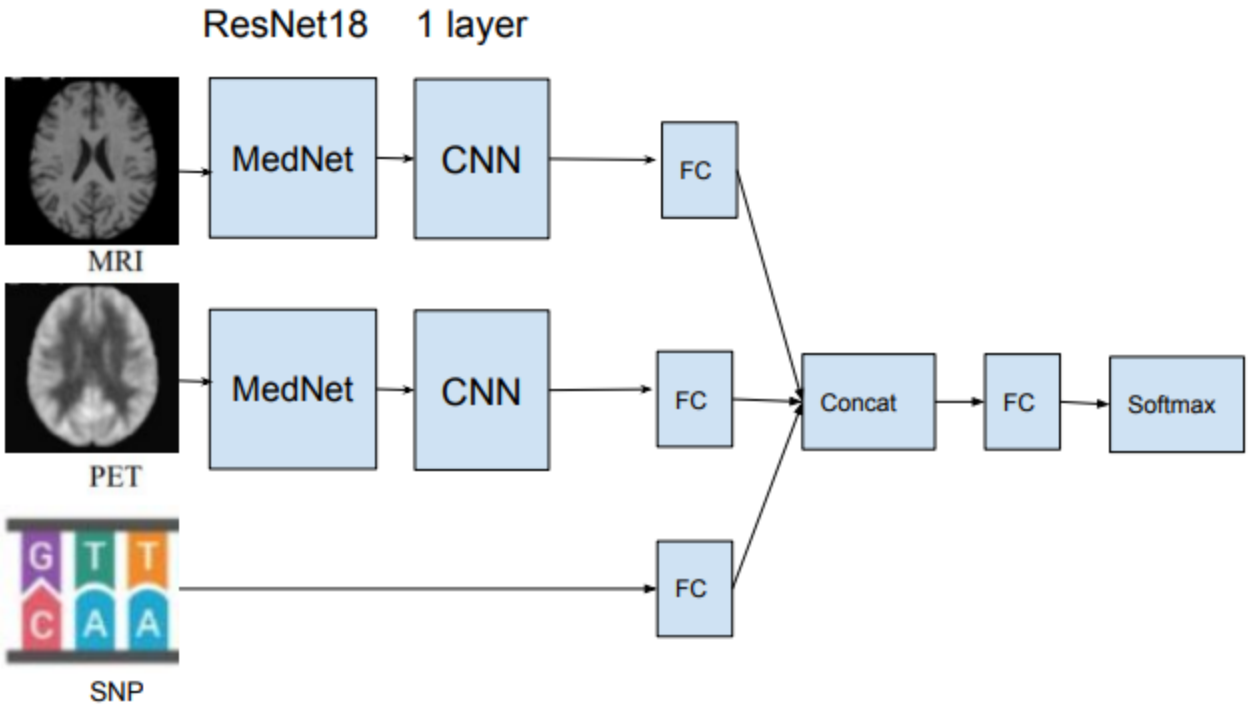
[5] T. Zhang and M. Shi, "Multi-modal neuroimaging feature fusion for diagnosis of alzheimer's disease," Journal of Neuroscience Methods, vol. 341, p. 108795, 2020

[6] Chen, Sihong, Kai Ma, and Yefeng Zheng. "Med3d: Transfer learning for 3d medical image analysis." arXiv preprint arXiv:1904.00625 (2019)

[7] T. Zhang and M. Shi, "Multi-modal neuroimaging feature fusion for diagnosis of alzheimer's disease," Journal of Neuroscience Methods, vol. 341, p. 108795, 2020.

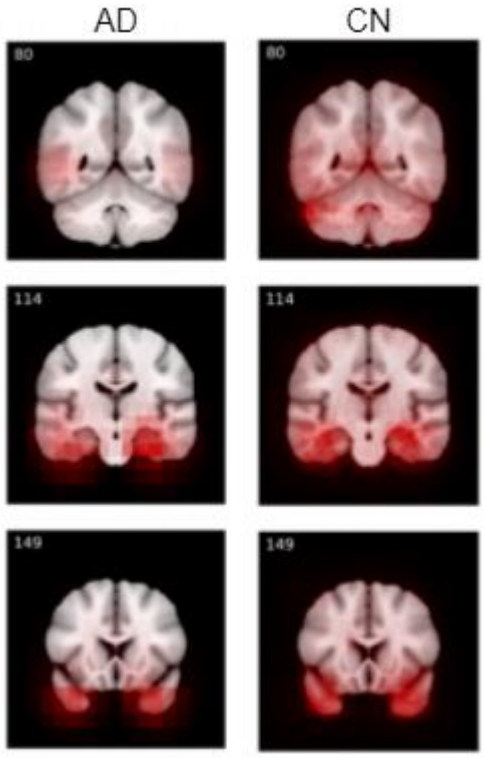
Thank you for your attention!

Supplemental Slides - Transfer Learning from MedNet

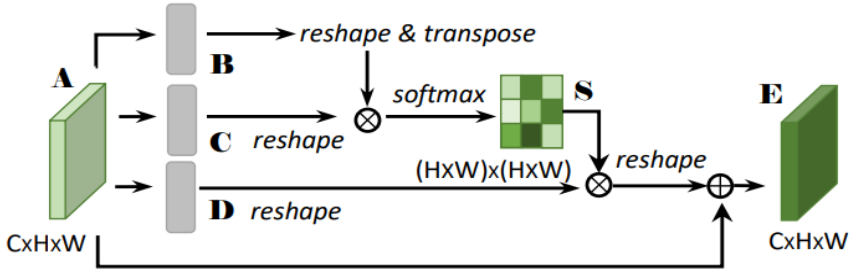


Supplemental Slides - Visualization

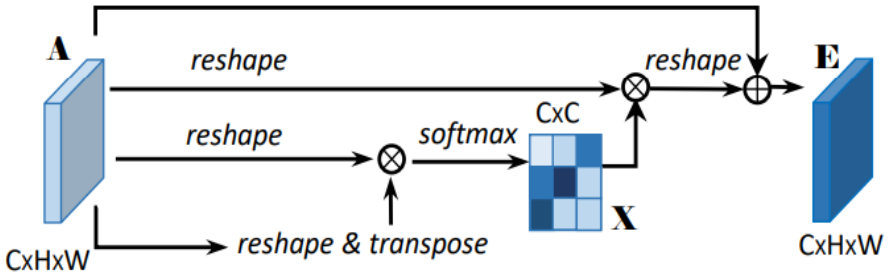
- **Qualitatively analysis is done over the model, found out that model focusing on inferior and middle temporal gyrus of the brain, which are known to be involved in AD.**



Supplemental Slides, Self-Attention

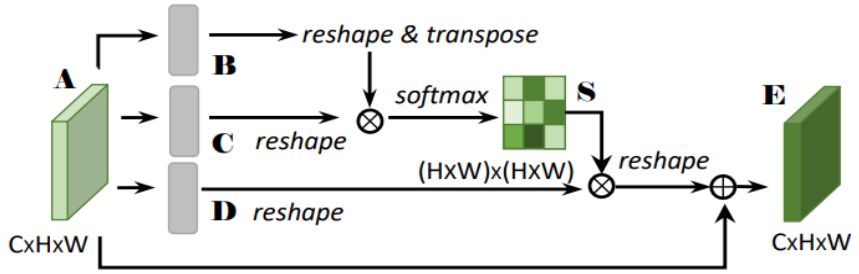


Spatial Attention [4]



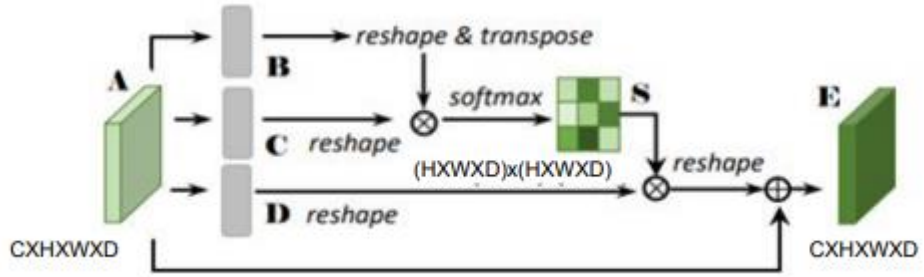
Channel Attention [4]

Supplemental Slides, Self-Attention for 2d Case



- Given A $\in \mathbb{R}^{CXHXW}$, B and C are obtained by feeding A into convolutional layers:
- $B, C \in \mathbb{R}^{CXHXW}$. Then, they reshaped in $\mathbb{R}^{CXN}$, where N = H x W (n of pixels)
- Then Transpose(C) x B is fed in the Softmax function to obtain: S $\in \mathbb{R}^{NXN}$ as
- $s_{ji} = \frac{exp(B_i \cdot C_j)}{\sum_{i=1}^{N} exp(B_i \cdot C_j)}$, where s_{ji} measures the i^{th} position's impact on j^{th} position.
- Lastly, D is obtained in the same way and reshaped into $\mathbb{R}^{CXN}$ and a D x Transpose(S) is calculated and reshaped in $\mathbb{R}^{CXHXW}$.
- Finally, it multiplied with a scale parameter to obtain the final output:
- $E_j = \alpha \sum_{i=1}^{N} (s_{ij} D_i) + A_j$
- Resulting feature E is a weighted sum of the features across all positions and original features. Therefore, it encodes the global contextual view.

Supplemental Slides, Self-Attention for 3d Case



- **On 3d case, only the reshaping part changes, rest stays the same!**

Supplemental Slides, Statistical Tests

- We used Friedman test to validate that our results are statistically significant on Baseline model to measure the effect of different modalities
- We also calculated p-value on the final models on average recall rate