Connectome Informed Attention — Predicting Tau Spreading for Alzheimer's Disease

Niklas Bühler, Mohamed Said Derbel, Andres Zapata

Medical Background and Motivation

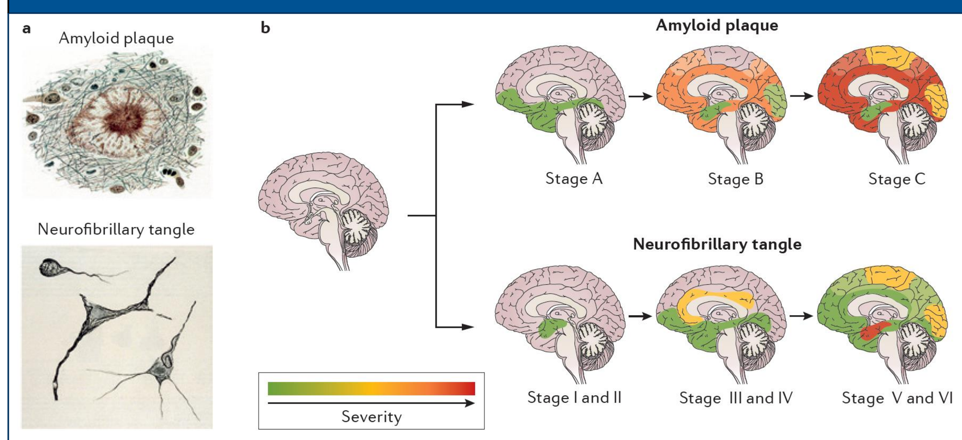


Fig 1: Depiction of Alzheimer's pathology taken from [2].

- Alzheimer's disease is characterized by two phenomena [2]:
  - Accumulation insoluble **amyloid-b plaques**
  - Aggregation of **tau protein** in tangles
- Tau protein has a higher correlation with cognitive degeneration

Tau spreading behavior [3]:
- Spreads 'prion-like' by diffusion and neural firing rates
- Tau concentration spreads from 'hubs' towards other 'hubs' to which they are connected

Goal

Predict next session tau accumulation values based on subject history and brain connectivity data

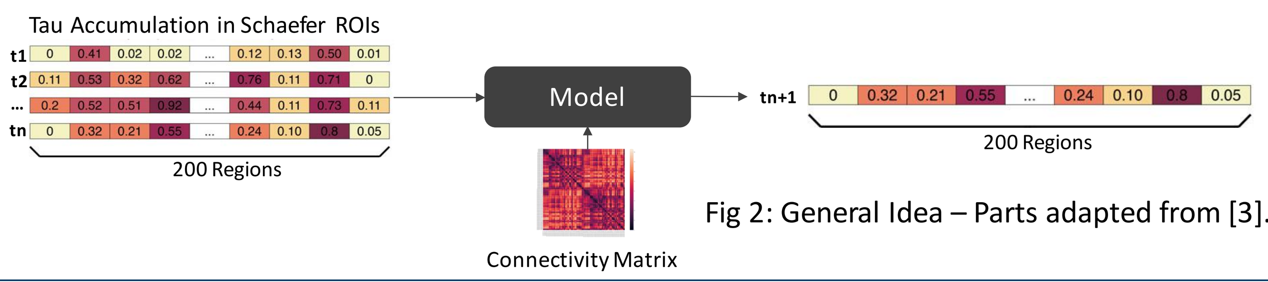


Fig 2: General Idea – Parts adapted from [3].

Dataset

1. **Dataset Generation**
   Out of the 728 provided sequences we generated 27,988 sequences
2. **Dataset Split**
   Patient-level Stratification based on diagnosis and input sequence length

	# Patients	# Sequences
Train	538	20,676
Validation	115	3,443
Test	116	3,869

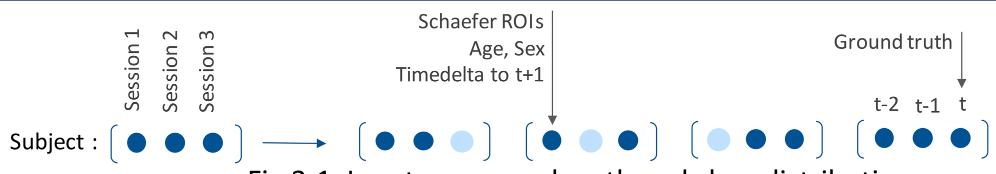


Fig 3.1: Input sequence length and class distributions.

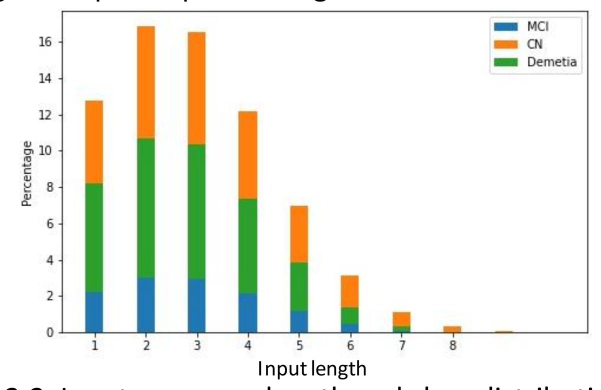


Fig 3.2: Input sequence length and class distributions.

Method

Baseline Model
- Our baseline attention model consists of a tradition Transformer encoder structure where its outputs are collapsed and reduced to 200 through a linear layer
- We extend the baseline model to incorporate the connectivity through different mechanism:

Early Fusion Transformer Encoder
Connectome Embedding Layer:
- Initialized with normalized functional connectivity data
- Weights are frozen
- In parallel for every entry in sequence
- Concatenation of original input and embedding is fed into the Transformer Encoder

Other Models
- **Late Fusion Transformer Encoder**: Analogous to Early Fusion Transformer Encoder
- **Connectome-Initialized Attention**: Query matrix of attention mechanism was initialized to apply functional connectivity relation to the input
- **Dual-Encoder**: pretraining a multi attention head to predict next session and then using the head as an extra encoder layer
- **Triformer**: Three transformer encoders, acting on the raw input, the connectome embedding and the concatenated outputs
- **Connectome-Head**: Utilizing a traditional encoder with the addition of one single frozen head that has its parameters fixed as the connectivity matrix

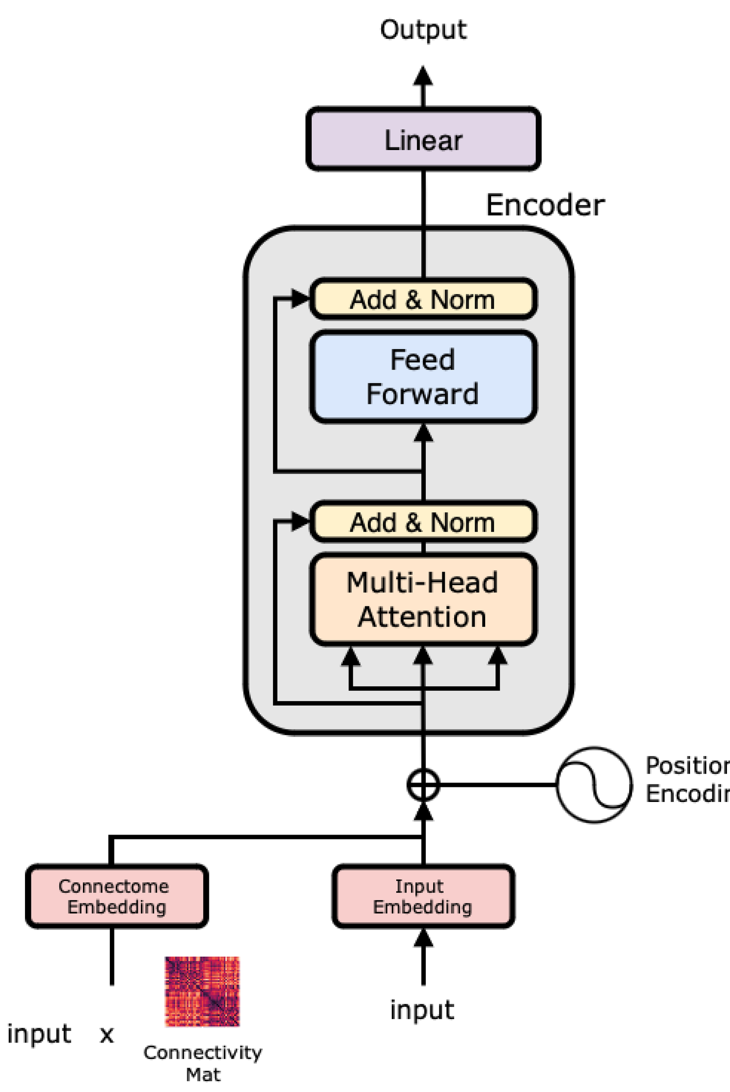


Fig 4. Early Fusion Transformer Encoder.

Model Performance

	Test Loss	Test Accuracy
MLP	0.036	0.898
LSTM	0.0297	0.939
Transformer	0.03215	0.9482
Early Fusion Transformer	**0.0282**	**0.9529**
Late Fusion Transformer	0.0441	0.9120
Connectome-Initialized Attention	0.0306	0.9445
Dual-Encoder	0.035	0.91
Triformer	0.0312	0.9498
Connectome-Head	0.0319	0.9438

Significance of Connectivity Information

Tau Protein tangles seem to have a higher correlation and temporal contingency with cognitive degeneration → Connectivity of brain regions can be useful for modeling the spreading behavior of tau

Functional Connectivity [3]:
Can be encoded as a static matrix computed from resting state fMRI data from multiple subjects. In total the creation of the connectivity matrix has three steps:
1. Preprocessing
2. Thresholding
3. Inversion

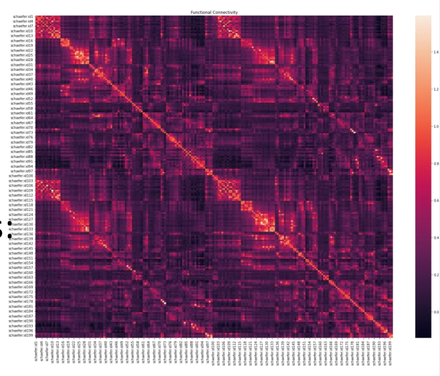
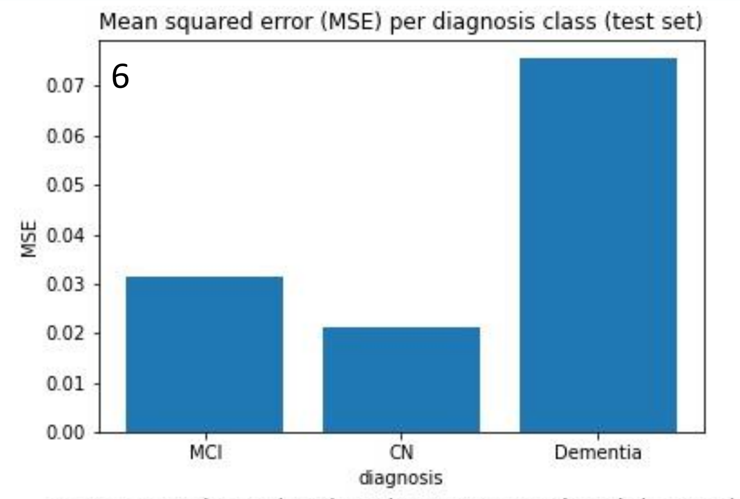


Fig 5: Functional Connectivity Matrix.

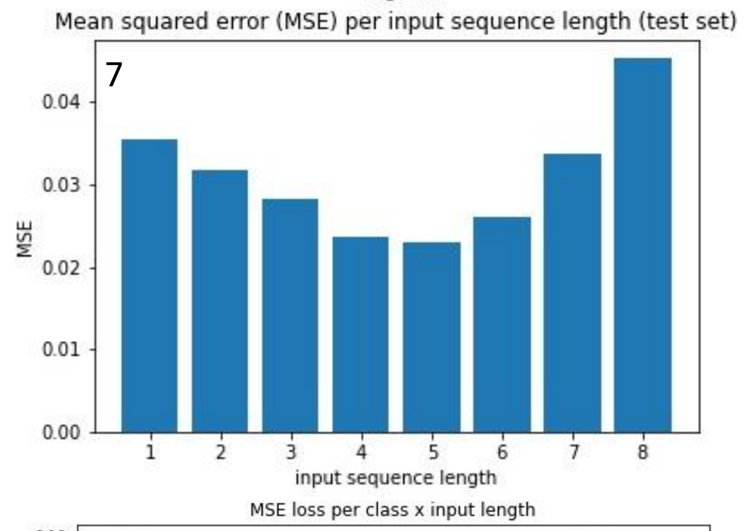
We demonstrated the significance of the positive effects of the connectivity matrix in sequence learning by using **Welch's t-test** as a statistical test

Results

We observed the following results on the test set:
- The model shows better performance on the CN and MCI classes than with the subjects with dementia
- The difference in performance is mainly due to class imbalance



- The model performs best when dealing with input sequences of length 5 and performs worse when given shorter or longer input sequences
- Possible reasons for this might be the that samples with low sequence length contain less information and there aren't a lot of samples with long sequence length in the dataset
- Another reason might be the increased time delta across session for shorter sequences, induced by the data augmentation process



- When stratifying for both factors, sequence length and target class, it can be seen that the test error decreases steadily for the MCI class and only increases for long sequence lengths for the CN and Dementia classes

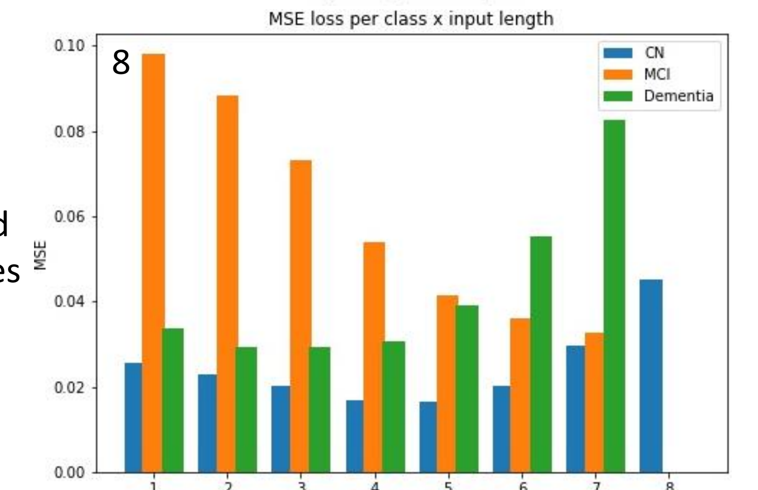


Fig 6, 7: MSE Loss | Fig 8: MSE loss per class and input len.

Visualization of Results

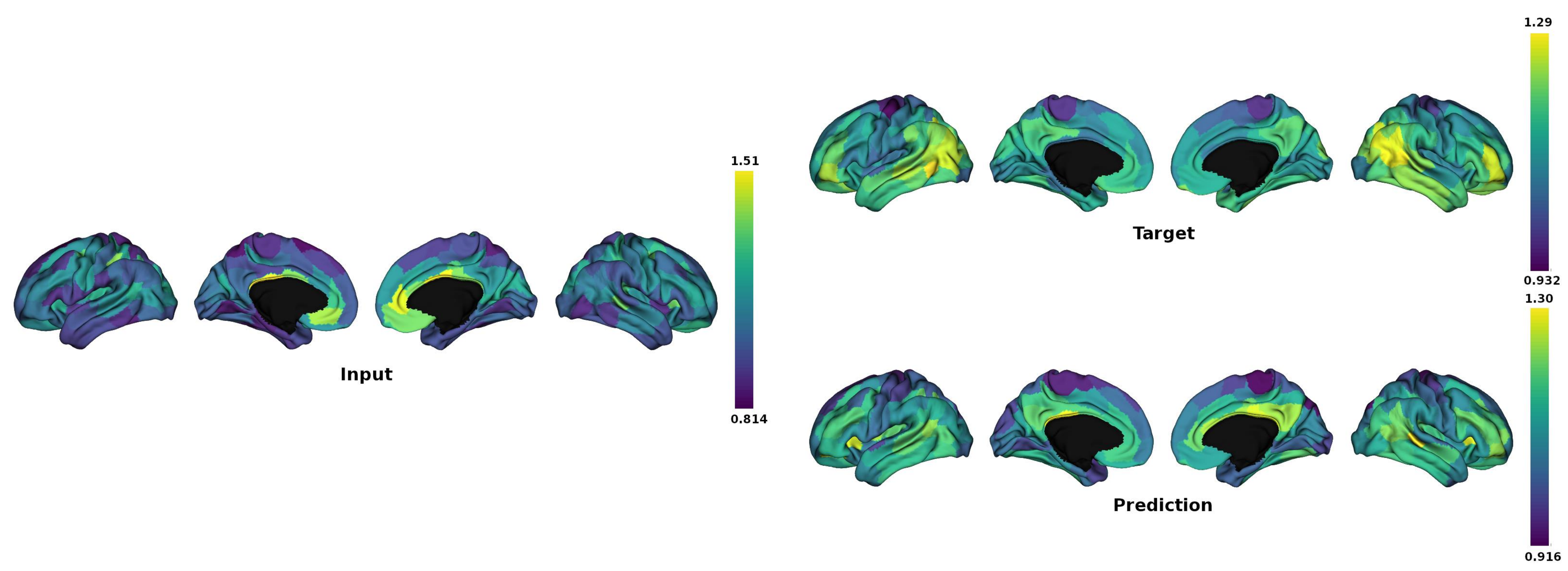


Fig 9: Visualization of input, prediction (Early Fusion Transformer model) and target values, for the CN class.

Main Findings & Future work

Findings:
- Our models were able to predict tau spreading across the Schaefer200 regions in the brain.
- The Transformer architecture is a suitable choice for this task
- Integration of functional connectivity information increased the models prediction capabilities significantly

Future work:
- Approaches to mitigate the class imbalance are to be employed, this can be achieved through loss weighting or robust optimization

References

[1] Masters, C., Bateman, R., Blennow, K. *et al.* Alzheimer's disease. *Nat Rev Dis Primers* **1**, 15056 (2015). https://doi.org/10.1038/nrdp.2015.56
[2] Takeda, S. (2019). Tau propagation as a diagnostic and therapeutic target for dementia: Potentials and unanswered questions. Frontiers in Neuroscience, 13, 1274
[3] Franzmeier N, Dewenter A, Frontzkowski L, et al. Patient-centered connectivity-based prediction of tau pathology spread in Alzheimer's disease. Sci Adv. 2020 Nov doi: 10.1126/sciadv.abd1327. PMID: 33246962; PMCID: PMC7695466.

Klinikum rechts der Isar
Technische Universität München