# Uncovering the topology of the temporal region in Alzheimer's disease.

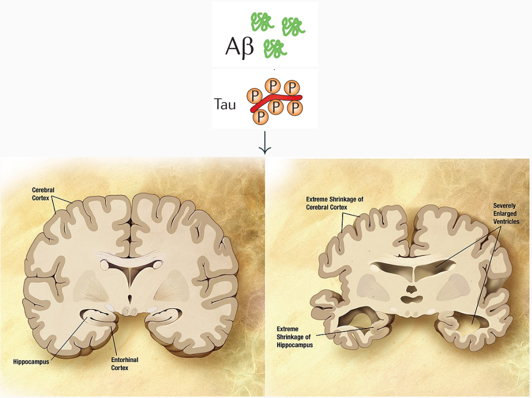
Philip Hartout

November 23, 2020

# Alzheimer’s disease

**Alzheimer’s disease:**

- Nearly 40 million people live with AD
- Cost in US alone $ 2 trillion by 2030
- Among leading causes of death in EU




**Topology:**

- Concerned with “properties of a geometric object that are preserved under **continuous deformations**, such as stretching, twisting, crumpling and bending, but not tearing or gluing.”
- Recently, *persistent homology* has emerged to quantify (differences in) the shape of data.
- **How about applying persistent homology to quantify changes in shape due to Alzheimer’s disease?**

# Topology in AD - Research Avenues

1. Diagnosis (classification)
2. Subtype identification
3. Progression & forecasting

$\rightarrow$ Some findings in these directions will be presented today

# Analysis setting

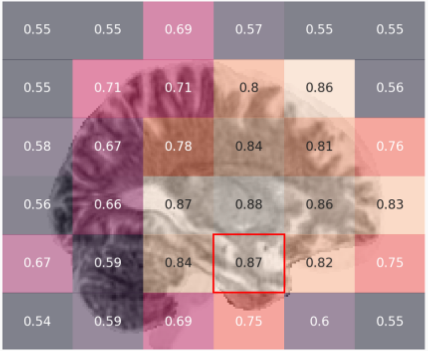



**Figure 2:** AUCPRC on each patch, achieved using a model described in earlier work. Chosen patch for analyses is boxed in red (patch with highest accuracy).

# I - Diagnosis prediction

- Computed the cubical filtration to obtain persistence image for each patch
- Simple CNN to classify AD/CN patients.
- Use same partition as NeurIPS submission, train ANN 3 times on each partition
- Export misclassified samples

# I - Persistent homology

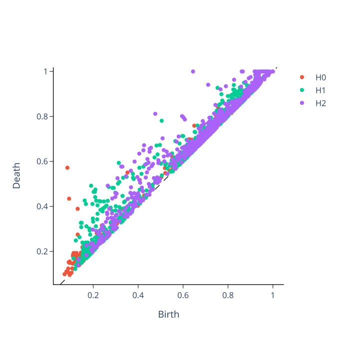



**Figure 3:** CN

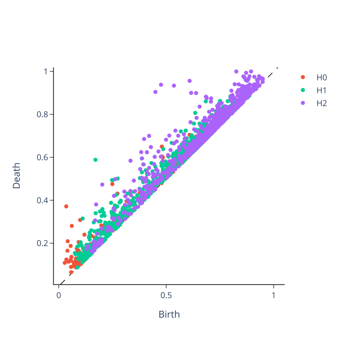



**Figure 4:** MCI

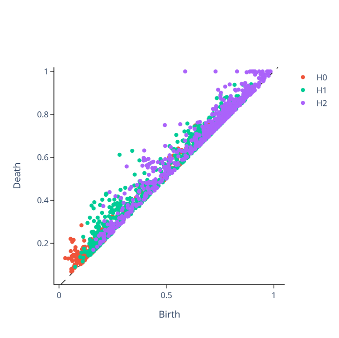



**Figure 5:** AD

# I - Diagnosis prediction - Persistence Images

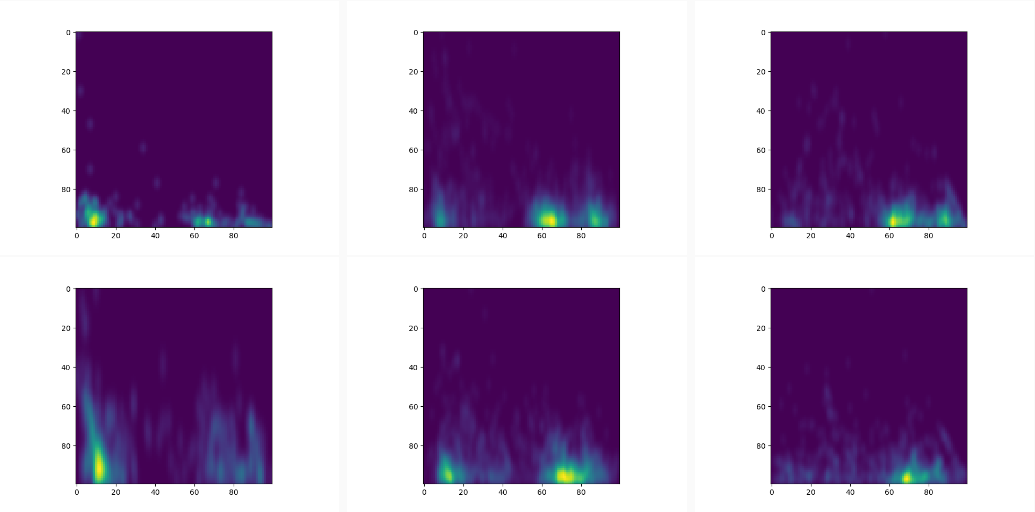
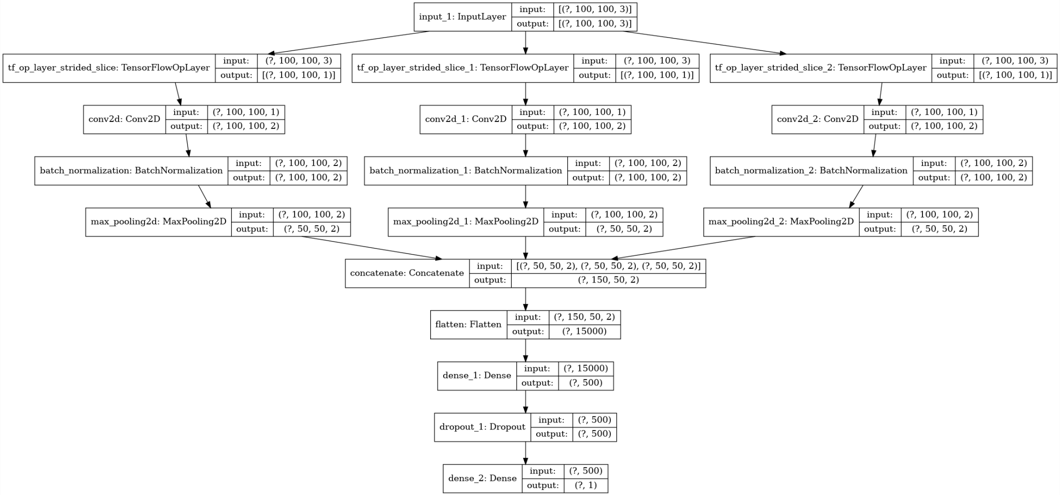



**Figure 6:** Columns: hom. dimension (0,1,2); Row: Diagnostic category (CN top, AD bottom).

# I - Diagnosis prediction - Network architecture



# I - Diagnosis prediction - Performance

| Performance metric | Model trained on PIs from patches |
| --- | --- |
| Training accuracy | $0.81 \pm 0.01$ |
| Validation accuracy | $0.78 \pm 0.03$ |
| Precision | $0.81 \pm 0.04$ |
| Recall | $0.77 \pm 0.03$ |
| AUC | $0.85 \pm 0.03$ |

**Table 1:** Performance metrics of the model.

$\rightarrow$ using a normal CPU, training takes 1 minute!

## II - Distance analysis

**Question:** How topologically heterogenous is the data?

- Compute a persistence landscape (allows for statistical analysis)
- Compute a median persistence landscape with one layer for each category (CN, MCI, AD)
- Compute $L^1$ norm from median landscape for each image within category

Average persistence landscape.

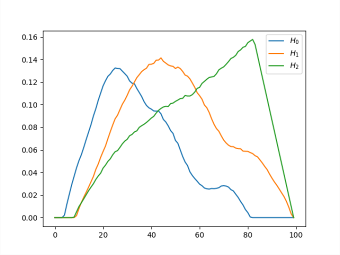



**Figure 7:** CN

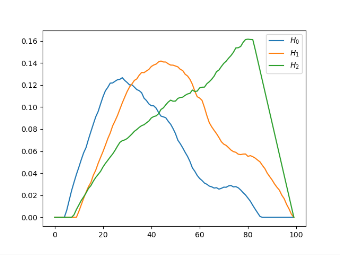



**Figure 8:** MCI

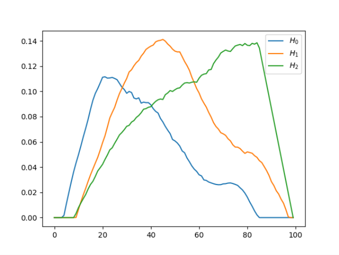



**Figure 9:** AD

# II - Distance analysis

**Question:** How topologically heterogenous is the data?

| | Mean | Median | Standard deviation | Q3 | Max | Skewness |
|---|---|---|---|---|---|---|
| CN $H_0$ | 2.16 | 2.00 | 0.78 | 2.50 | 7.41 | 1.78 |
| CN $H_1$ | 2.61 | 2.27 | 1.17 | 2.93 | 9.47 | 1.92 |
| CN $H_2$ | 2.38 | 2.23 | 0.88 | 2.79 | 7.19 | 1.39 |
| MCI $H_0$ | 2.24 | 2.04 | 0.82 | 2.55 | 6.21 | 1.71 |
| MCI $H_1$ | 2.57 | 2.19 | 1.29 | 2.80 | **11.87** | **2.57** |
| MCI $H_2$ | 2.40 | 2.27 | 0.83 | 2.82 | 6.55 | 1.18 |
| AD $H_0$ | 2.40 | 2.18 | 0.96 | 2.77 | 7.77 | 1.97 |
| AD $H_1$ | 2.47 | 2.13 | 1.15 | 2.77 | **9.28** | **2.10** |
| AD $H_2$ | 2.36 | 2.20 | 0.80 | 2.75 | 8.39 | 1.64 |

**Table 2:** Summary statistics of the distribution of distances

# III - Distance analysis - outline

**Question:** Among the patients who deteriorate, do we see higher average pairwise distances compared to patients who don't deteriorate?

- The data is a longitudinal dataset (multiple timepoints are available for each patient)
- Some patients deteriorate (transition from CN$\rightarrow$MCI or from MCI$\rightarrow$AD)
- Compute pairwise distance between patients ($L^1$ PL distance, Wasserstein distance and bottleneck distance), and average for each patient.

# III - Distance analysis - examples

Example of $L^!$ norm between PLs of a deteriorating patient.

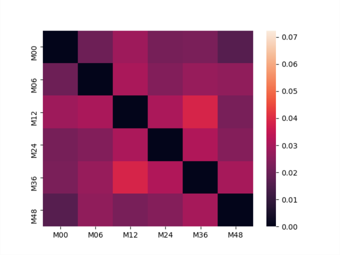
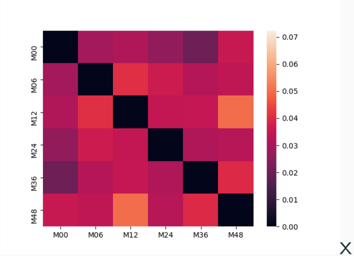



**Figure 10:** $H_0$




X

**Figure 11:** $H_1$

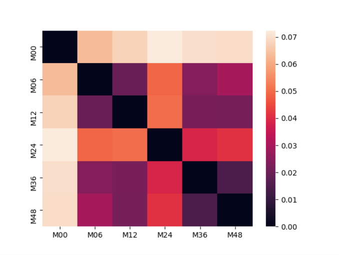



**Figure 12:** $H_2$

# III - Distance analysis - further examples

Example of $L^1$ norm between PLs of a subject who does *not* deteriorate.

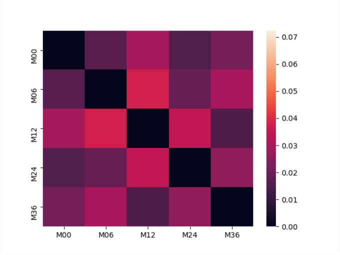



**Figure 13:** $H_0$

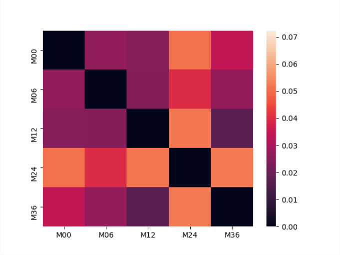



**Figure 14:** $H_1$

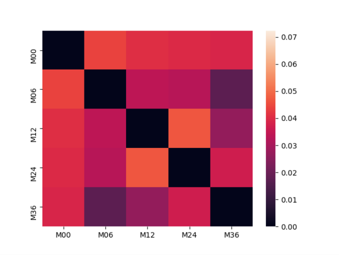



**Figure 15:** $H_2$

# III - Distance analysis - population-level results

Persistence landscape $L^1$ norm

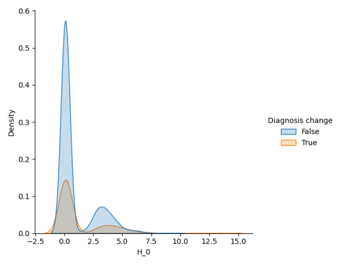



**Figure 16:** $H_0$

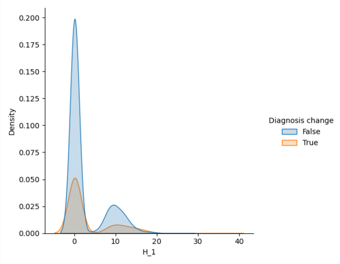



**Figure 17:** $H_1$

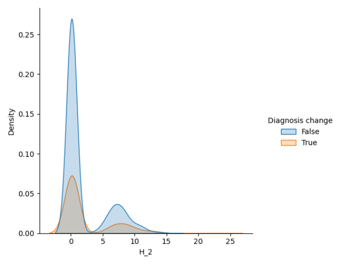



**Figure 18:** $H_2$

# III - Distance analysis - population-level results

Bottleneck $L^1$ norm

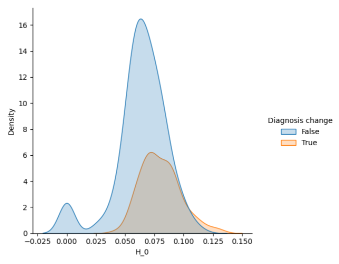



**Figure 19:** $H_0$

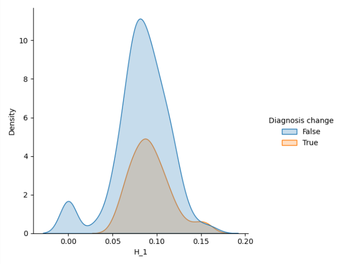



**Figure 20:** $H_1$

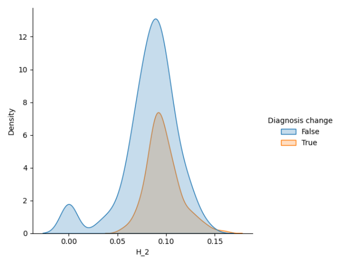



**Figure 21:** $H_2$

# III - Distance analysis - population-level results

Wasserstein distance

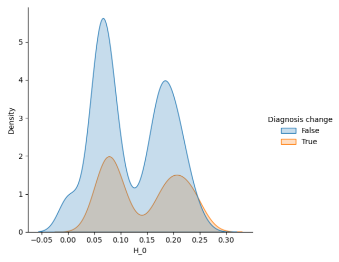



**Figure 22:** $H_0$

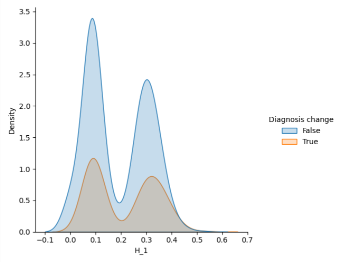



**Figure 23:** $H_1$

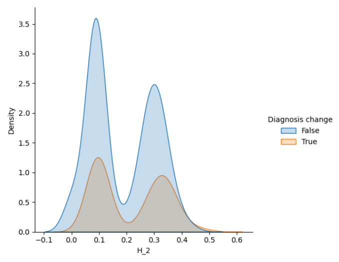



**Figure 24:** $H_2$

# Limitations & Outlook

1. Diagnosis
   - Performance can be enhanced by taking local features from other patches and passing them through convolutions. Especially useful when dealing with subtypes of Alzheimer's which show atrophy in other, larger brain regions.
   - Neural architecture could be made a little deeper with additional convolutions, but was not the sole aim of the lab rotation
   - Problem could also be made more useful (and challenging) by looking at diagnosing prodromal forms of AD.
2. Distances
   - In general, very *coarse* analysis (does not pick out subtypes or use individual features to differentiate between images), sensitive to noise.
   - Deep image clustering could be applied for subtype identification & better heterogeneity investigation.
   - Could learn better (topology-based) embeddings to better distinguish between people who progress versus those who do not.