CHB-MIT

Feature Representation and Dimensionality Reduction

Seizures

- **198 clinical seizures** annotated by experts
- However, in one case (*chb 12*), some seizure recordings **were changed** from bipolar to unipolar montage (these are **excluded** – 13 in total)
- Channel consistency across all recordings is ensured; this means only **18 channels are used**
- To ensure a **balanced dataset** 3 seizure files (*for now at least*) were **excluded** to make a 4-sec segment **60 sec after annotated seizure ending**
- Right now, this yields 182 interictal and 182 ictal segments

Feature Representation

- ”A High-Performance Seizure Detection Algorithm based on Discrete Wavelet Transform (DWT) and EEG” – Chen et al., 2017
- 159 citations and ”Scopus” peer-reviewed

- The paper investigates the **six frequency sub-bands** for EEG with **different wavelet families** and selects **different statistical features** of the sub-bands **for machine learning**
- ML model is a **binary SVM classfier** and evaluation was done with **“leave-one-out” cross-validation** and evaluation metrics used were primarily **accuracy, sensitivity and specificity**
- According to Chen et al., the result was best when using the wavelet family **(Coiflets; coif3)**, and seven statistical features (**Max, Min, Mean, STD, Skewness, Energy and Normalized STD**).
- **Accuracy**: 92.30%, **Sensitivity**: 91.71%, **Specificity**: 92.89%

Feature Representation

- **182 ictal segments** of 4 seconds selected at **the middle of annotated seizure start and end**

- **182 interictal segments** of 4 seconds selected **60 seconds after annotated seizure end**

- Segments are **DWT decomposed according** to the results of **Chen et al**. and the **seven statistical features** are extracted yielding the following data shape for one segment**: [1, 18, 42]** corresponding to [segment, channels, features]

t-SNE: t-Distributed Stochastic Neighbor Embedding

- t-SNE is an **unsupervised, non-linear technique** used for data exploration and visualizing high-dimensional data
- In contrast to Principal Component Analysis (PCA) which was developed in 1933, **t-SNE was developed in 2008 by Van Der Maaten and Hinton**
- PCA seeks to maximize variance and preserve large pairwise distances and t-SNE differs by **preserving only small pairwise distances or local similarities**.

- **t-SNE calculates a similarity measure between pairs in both the high dimensional space and the low dimensional space**. It then tries to optimize these two similarity measures using a cost function

t-SNE: How The Algorithm Works

- **Step 1:** Measure similarities between data points in the high-dimensional space and center a Gaussian distribution over each point. Points are then renormalized and this will yield a set of probabilities P_{ij} for all points. Probabilities are proportional to the similarities.
- **Step 2:** Similar to **Step 1**, however, instead a student t-distribution with one degree of freedom is used to give a second set of probabilities Q_{ij} in the low dimensional space.
- **Step 3:** The set of probabilities from the low-dimensional space, Q_{ij}, should reflect those of the high-dimensional space P_{ij} . Therefore, the Kullback-Lieber divergence is used to measure the difference between probability distributions of the two-dimensional spaces. And finally, gradient descent is used to minimize the KL cost function

t-SNE: Hyperparameters

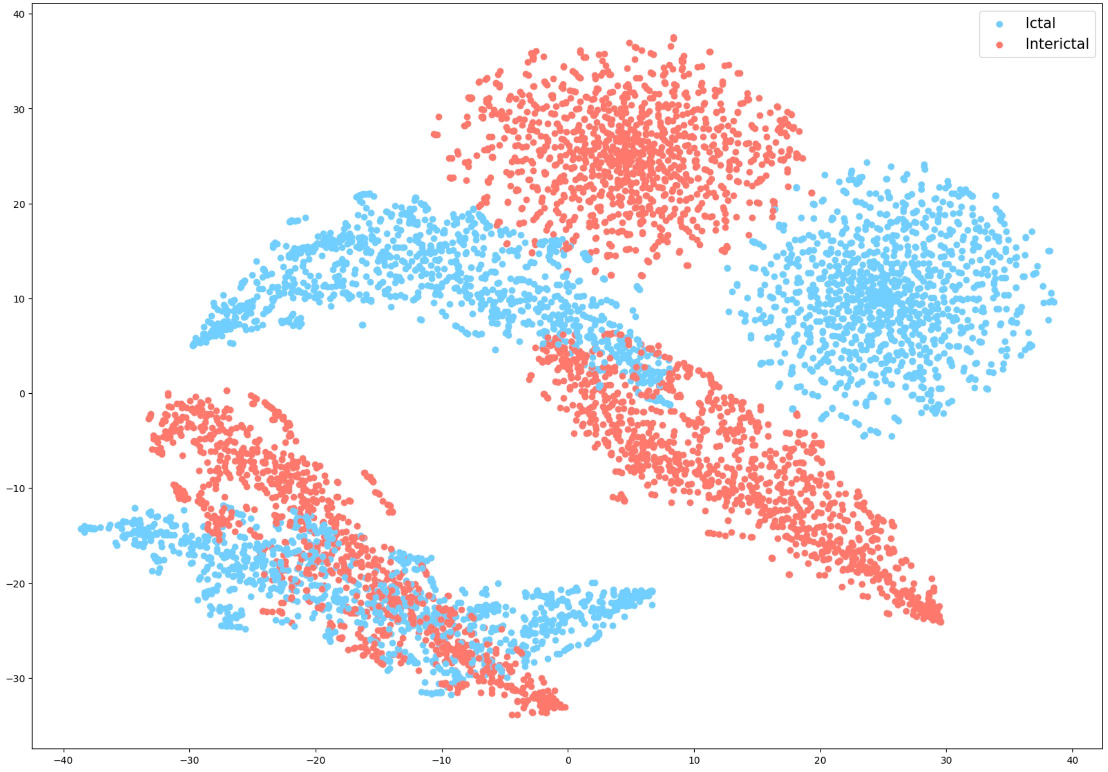
- **Perplexity: Influences the variance of the Gaussian distribution** and essentially the number of nearest neighbors (*according to Van Der Maaten and Hinton, normal range for perplexity is between 5 and 50*)

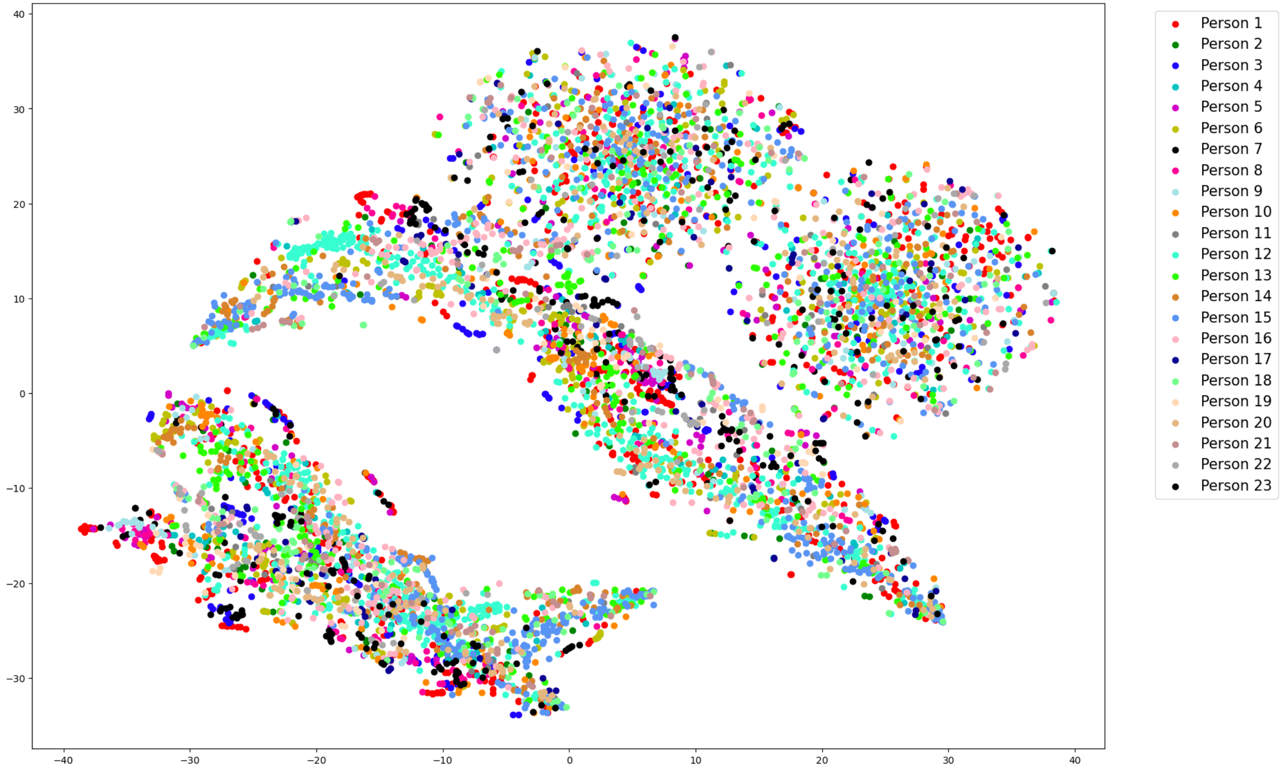
- **Iterations:** Number of iterations needed for tSNE to converge; **the more iterations the better**, however, typically not feasible to have 10.000 for big data

- **Number of Components:** Just like PCA, you would like to have a few amounts of components to explain most of the data and to visualize the data in 2D or 3D.

t-SNE: Results

Perplexity: 50, Iterations: 8



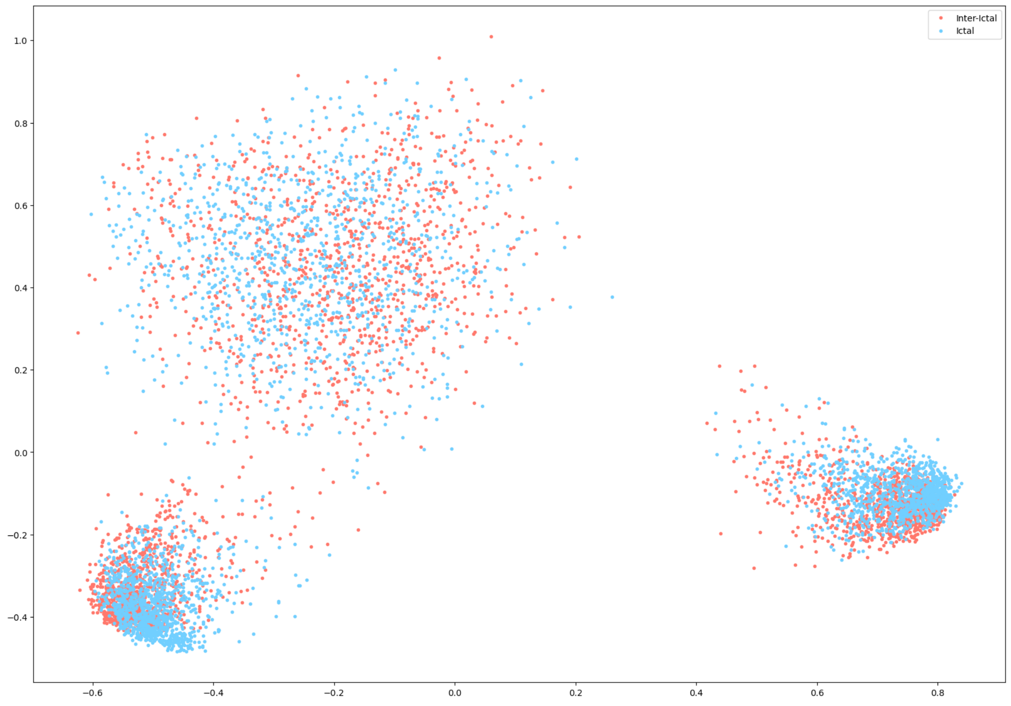


Labeled by ictal and inter-ictal state

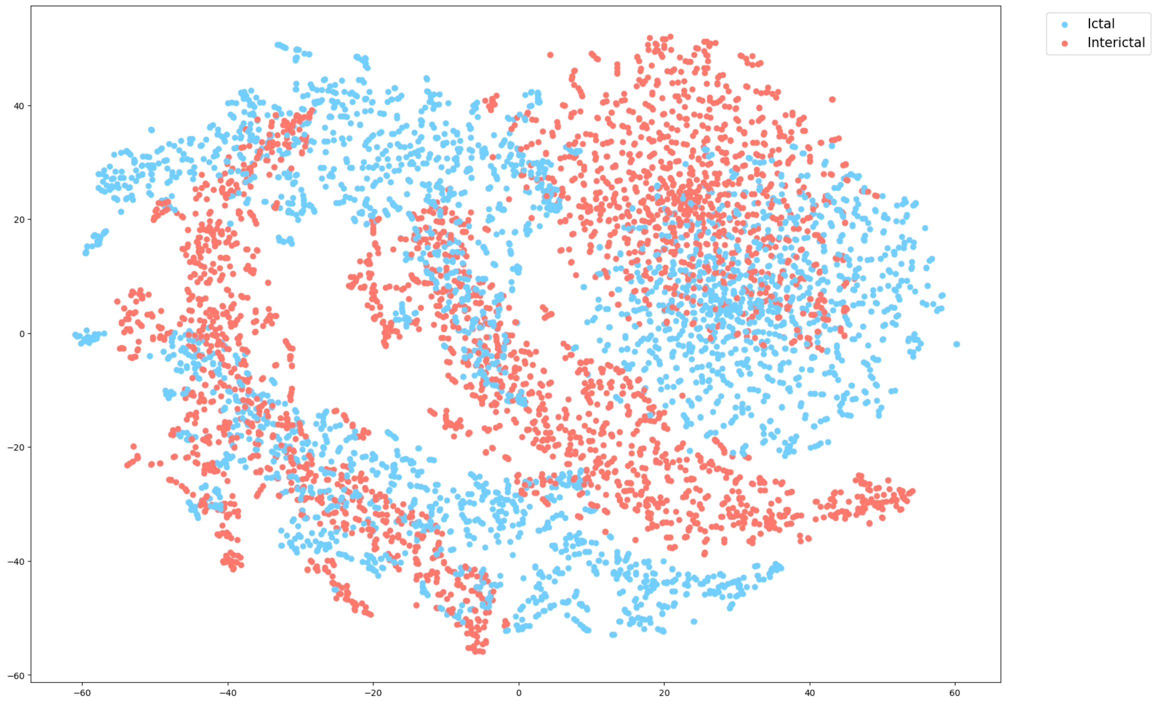
Labeled by patients

… and PCA + lower perplexities

PCA



t-SNE
Perplexity: 15, Iterations: 8



What Next?

- Machine Learning on t-SNE embeddings based on Chen et al.

- SVM or Random Decision Forest

- Hoping results are not too far from those obtained in the paper…

- Write report… ☺