UVIC
UNIVERSITAT DE VIC
UNIVERSITAT CENTRAL
DE CATALUNYA

PRELIMINARY RESULTS ON THE GENERATION OF ARTIFICIAL HANDWRITING DATA USING A DECOMPOSITION-RECOMBINATION STRATEGY

José Fernando Adrán Otero[1,2], Oscar Solans Caballer[1,2], Pere Marti-Puig[2], Zhe Sun[3], Toshihisa Tanaka[4,5], Jordi Solé-Casals[2,6*]

[1] Faculty of Computer Science, Multimedia and Telecommunications, Open University of Catalonia, Catalonia
[2] Data and Signals Processing Research Group, University of Vic-Central University of Catalonia, Catalonia
[3] Image Processing Research Team, RIKEN Center for Advanced Photonics RIKEN, Japan
[4] Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Japan
[5] Tensor Learning Team, RIKEN Center for Advanced Intelligence Project, Japan
[6] Department of Psychiatry, University of Cambridge, Cambridge CB2 3EB, U.K.

Singapore - ICASSP 2022

Overview

1 – Introduction

2 – Previous works

3 – Data augmentation using EMD

4 – Experimental results

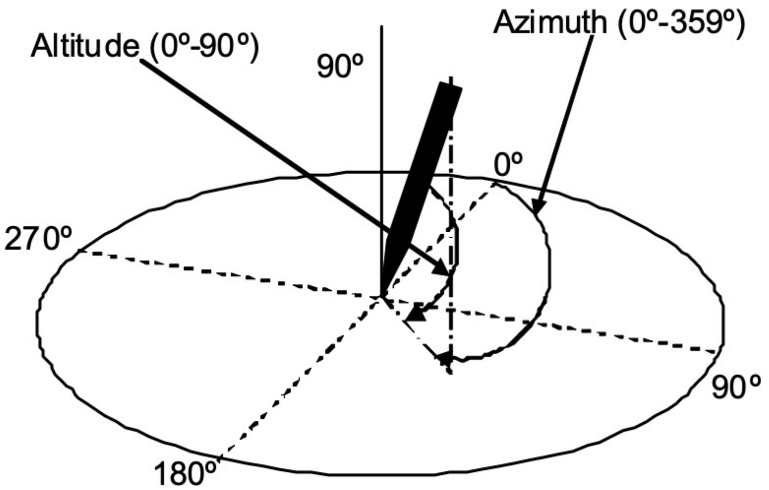
5 – Discussion and conclusions

Overview

1. Introduction

- Essential tremor (ET) is a disorder of the nervous system that causes involuntary and rhythmic movements that generally appears in people over 65 years of age, but it can appear earlier, from the age of 40
- Although it can be reflected in any part of the body, this tremor appears most frequently in the hands, and especially when the subject performs simple everyday tasks, such as eating, drinking or tying shoelaces.
- Different studies have focused on the automatic analysis of handwritten data to determine the patterns that allow an accurate diagnosis of this disease.
- Habitually, few subjects are available to properly train ML/DL models. Therefore, artificial data is needed to obtain useful models.

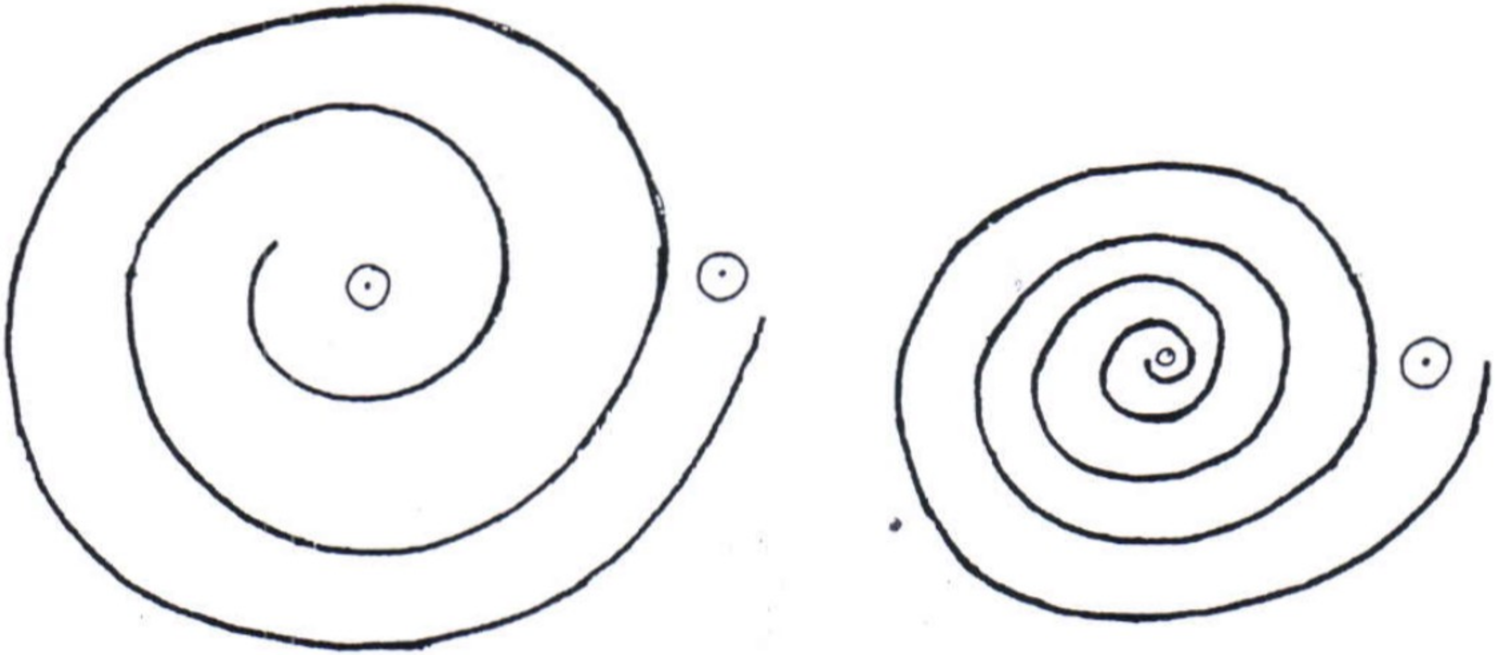
1. Introduction



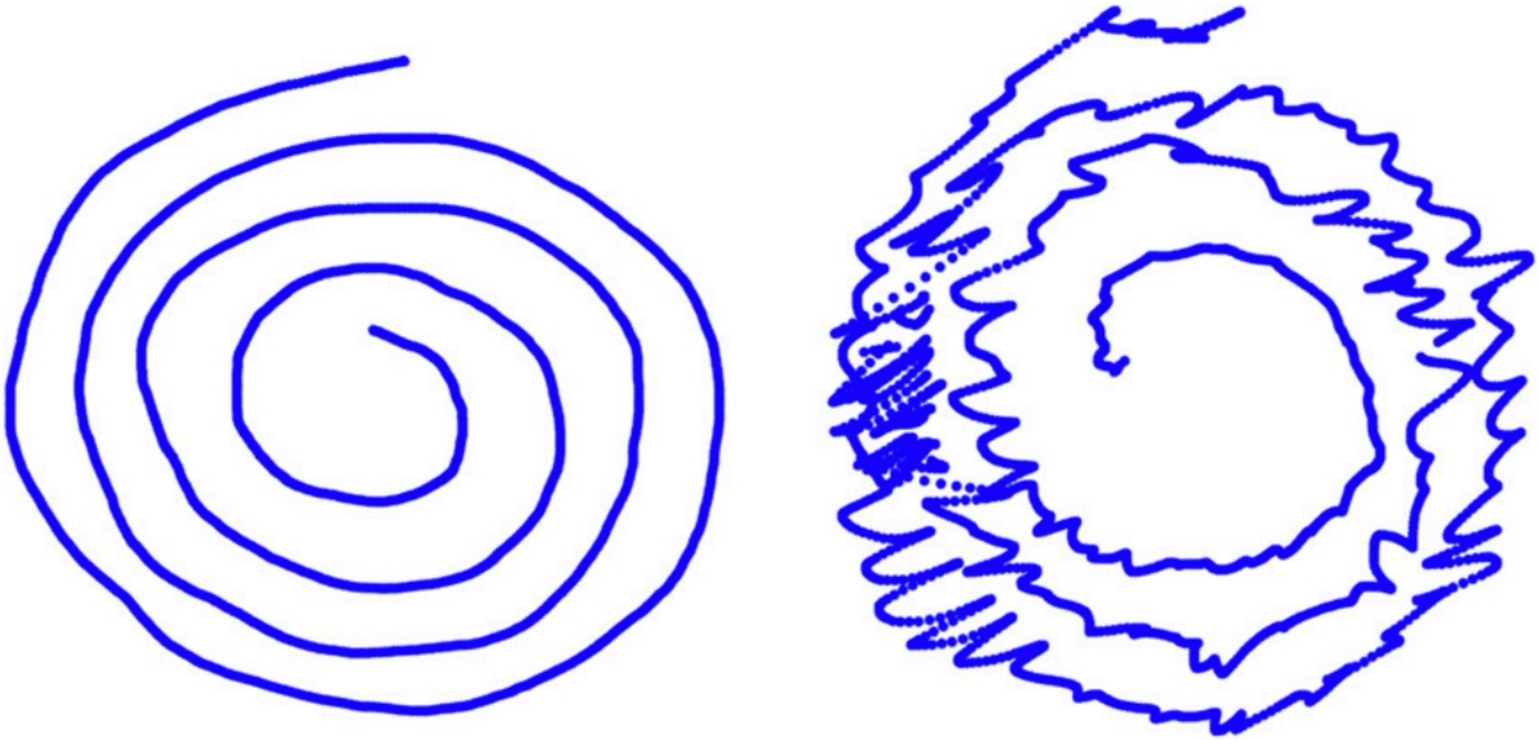
The digitizing tablet capture the spatial coordinates (x, y), the applied pressure of the pen, the azimuth (angle between the pen and the horizontal plane), and the altitude (angle between the pen and the vertical axis)

1. Introduction

Fahn, Stanley, Eduardo Tolosa, and Concepción Marín. "Clinical rating scale for tremor." *Parkinson's disease and movement disorders* 2 (1993): 271-280.



1. Introduction



Overview

2. Previous works

- Traditional methods rely on the calculation of many features and using feature selection methods to obtain a classification model.

- These features can be:
  - Time related features
  - Spatial related features
  - Pressure related features
  - Dynamic features
  - Frequency features
  - Nonlinear features such as Entropy, fractal dimension, etc.

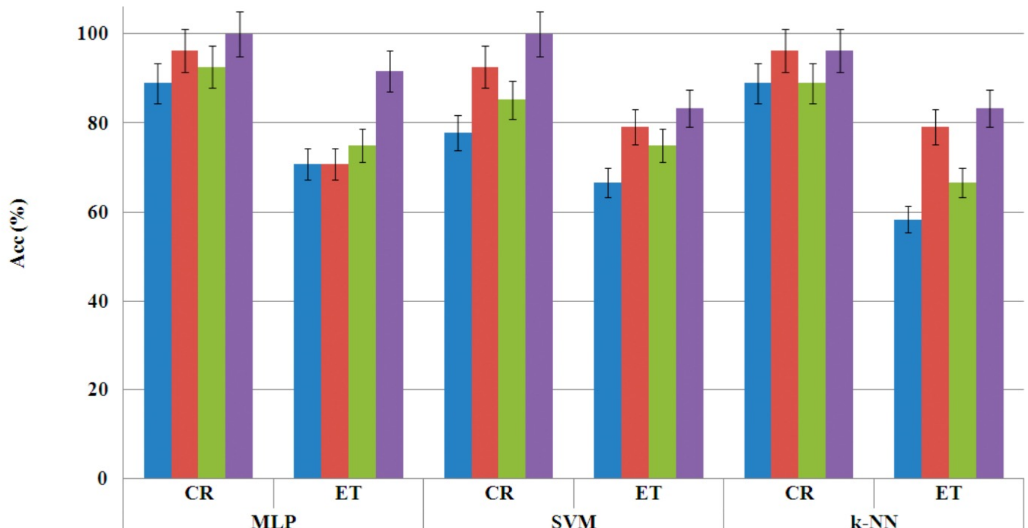
2. Previous works

Temporal Features	Descriptor
Sample entropy (SENT)	$m = 3, r = 0.2$
Mean absolute value (MAV)	$\frac{1}{N}\sum_{i=1}^{N}\lvert X_i\rvert$
Variance (VAR)	$\frac{1}{N-1}\sum_{i=1}^{N}\lvert X_i-\mu\rvert^2$
Root mean square (RMS)	$\sum_{i=1}^{N}\frac{1}{N}X_i^2$
Log detector (LOG)	$e^{\frac{1}{N}\sum_{i=1}^{N}log(\lvert X_i\rvert)}$
Waveform length (WL)	$\sum_{i=1}^{N-1}\lvert X_{i+1}-X_i\rvert$
Standard deviation (STD)	$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\lvert X_i-\mu\rvert^2}$
Difference Absolute standard deviation (AAC)	$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(X_{i+1}-X_i)^2}$
Fractal dimension (FD)	Higuchi's algorithms with $m = 5$
Maximum fractal length (MFL)	$log(\sum_{i=1}^{N-1}\lvert X_{i+1}-X_i\rvert)$
Myopulse percentage rate (MYO)	Percentage of time where the signal is bigger than two times the mean
Integrated EMG (IEMG)	$\sum_{i=1}^{N}\lvert X_i\rvert$
Simple square EMG (SSI)	$\sum_{i=1}^{N}X_i^2$
Zero crossing (ZC)	The number of times in which the signal crosses its mean
Slope sign change (SSC)	The number of times in which the slope of the sign changes
Wilson amplitude (WAMP)	$\sum_{i=1}^{N-1}\lvert X_i - X_{i+1}\rvert > \epsilon$ where ϵ is the mean of the signal
Autoregressive coefficients (AR, 4 coefficients)	AR parameter estimation via Yule-Walker method

Frequency Features	Descriptor
Main peak amplitude (Pmax)	Maximum peak
Main peak frequency (Fmax)	Frequency of the max peak
Mean power (MP)	$\frac{1}{N}\sum_{i=1}^{N}\lvert P_i\rvert$
Total power (TP)	$\sum_{i=1}^{N}P_i$
Mean frequency (MNF)	Estimates the mean normalized frequency of the power spectrum
Median frequency (MDF)	Estimates the median normalized frequency of the power spectrum
Standard deviation (STD)	$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\lvert P_i-\mu\rvert^2}$
1st spectral moment (SM1)	Spectral moments
2nd spectral moment (SM2)	Spectral moments
3rd spectral moment (SM3)	Spectral moments
Kurtosis (KUR)	Kurtosis of the power spectrum
Skewness (SKW)	Skewness of the power spectrum
Autocorrelation (Auto, 3 coefficients)	3 firsts coefficients of the autocorrelation

2. Previous works

López-de-Ipiña, Karmele, et al. "Selection of entropy-based features for automatic analysis of essential tremor." *Entropy* 18.5 (2016): 184.



	MLP CR	MLP ET	SVM CR	SVM ET	k-NN CR	k-NN ET
LF	88.89	70.83	77.78	66.67	88.89	58.33
SLF	96.30	70.83	92.59	79.17	96.30	79.17
PE-m7t7	92.59	75.00	85.19	75.00	88.89	66.67
SPE-m7t7	100.00	91.67	100.00	83.33	96.30	83.33

Figure 12. Accuracy (%) of classes for the paradigms for the references and the best options. SPE-m7t7 improves not only the global rates, but also the class rate even for the less powerful model: *k*-NN.

After a hard optimization process, the best results are obtained using a selection of linear features + Permutation Entropy, obtaining a final accuracy of 96%, using 78 features

2. Previous works

Lopez-de-Ipina, Karmele, et al. "Automatic analysis of archimedes' spiral for characterization of genetic essential tremor based on Shannon's entropy and fractal dimension." *Entropy* 20.7 (2018): 531.

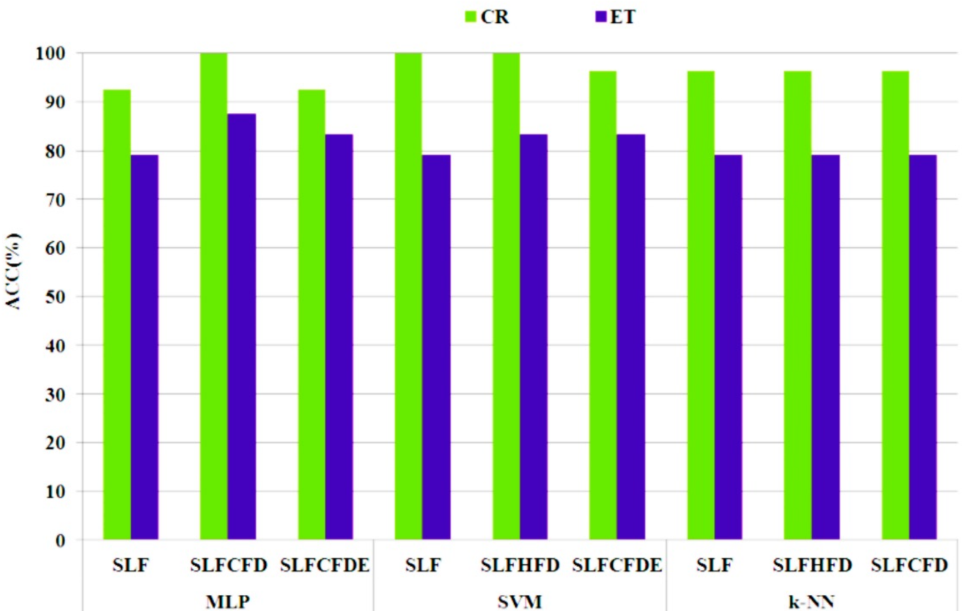
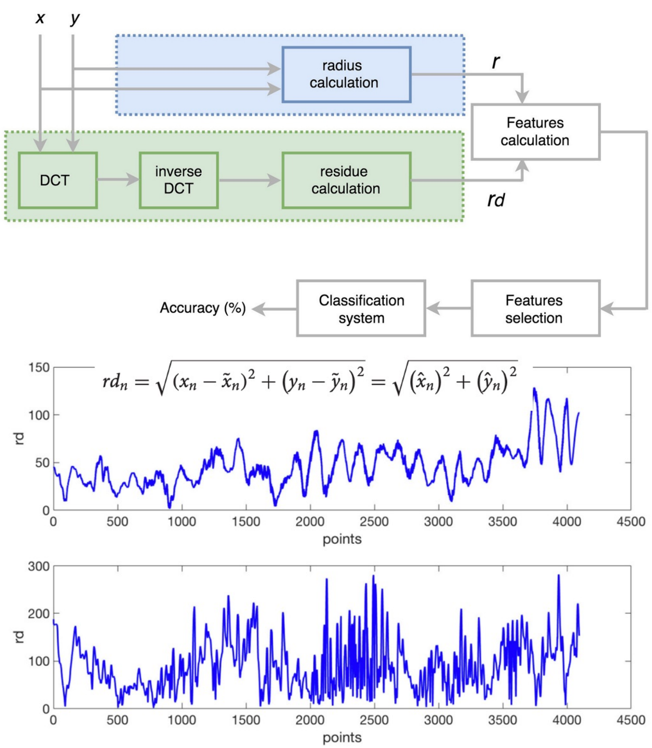


Figure 12. Accuracy (%) for each group (CR and ET) and classification algorithm when considering the best performing sets of features.

After a hard optimization process, the best results are obtained using SLFCFD with MLP, obtaining a final accuracy of 94.11% using 77 features

2. Previous works

Solé-Casals, Jordi, et al. "Discrete cosine transform for the analysis of essential tremor." *Frontiers in physiology* 9 (2019): 1947.



After a hard optimization process, the best results for the radius method is 85.71% and the best results for the residue method is 95.92% Combining both methods, the best result is 97.96% combining both methods and 7 features (5 from the residue method +2 from the radius method)

Overview

3. Data augmentation using EMD

- EMD (Empirical Mode Decomposition): is based on a data-driven method that allows decomposing the different time signals into a finite number of the so-called Intrinsic Mode Functions (IMFs).
- Each IMF represents a non-linear oscillation of the original signal
- The number of IMFs depends on the structure of the signal.
- All the IMFs have to fulfill the following two conditions:
    - The number of maximums has to be the same as the number of zero crossings, or at least they have to differ by only one.
    - For any sample, the mean value between the envelope of the local maxima and that of the local minima must be zero.
- A signal is completely restored by adding all its IMFs and the final residue:

$$x(t) = \sum_{k=1}^{n} IMF_k(t) + r_n(t)$$

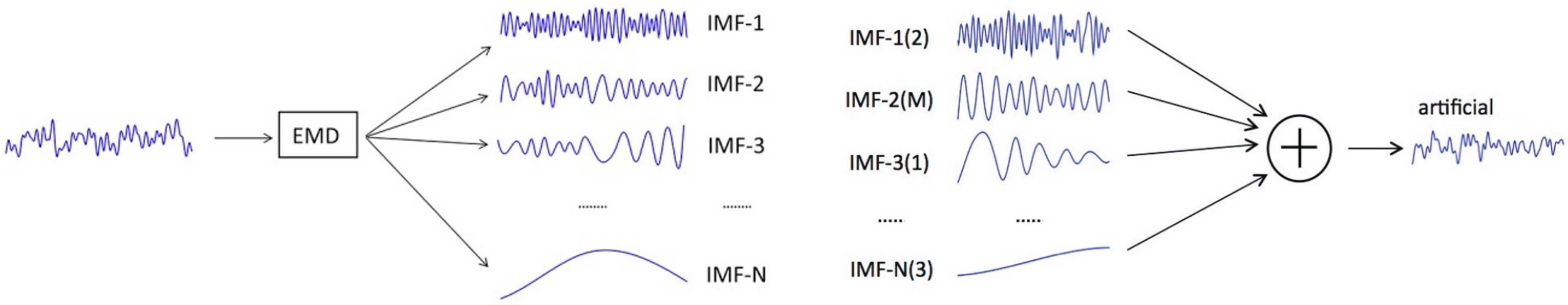
3. Data augmentation using EMD

The iterative process to obtain the IMFs from a signal $x(t)$ is as follows:

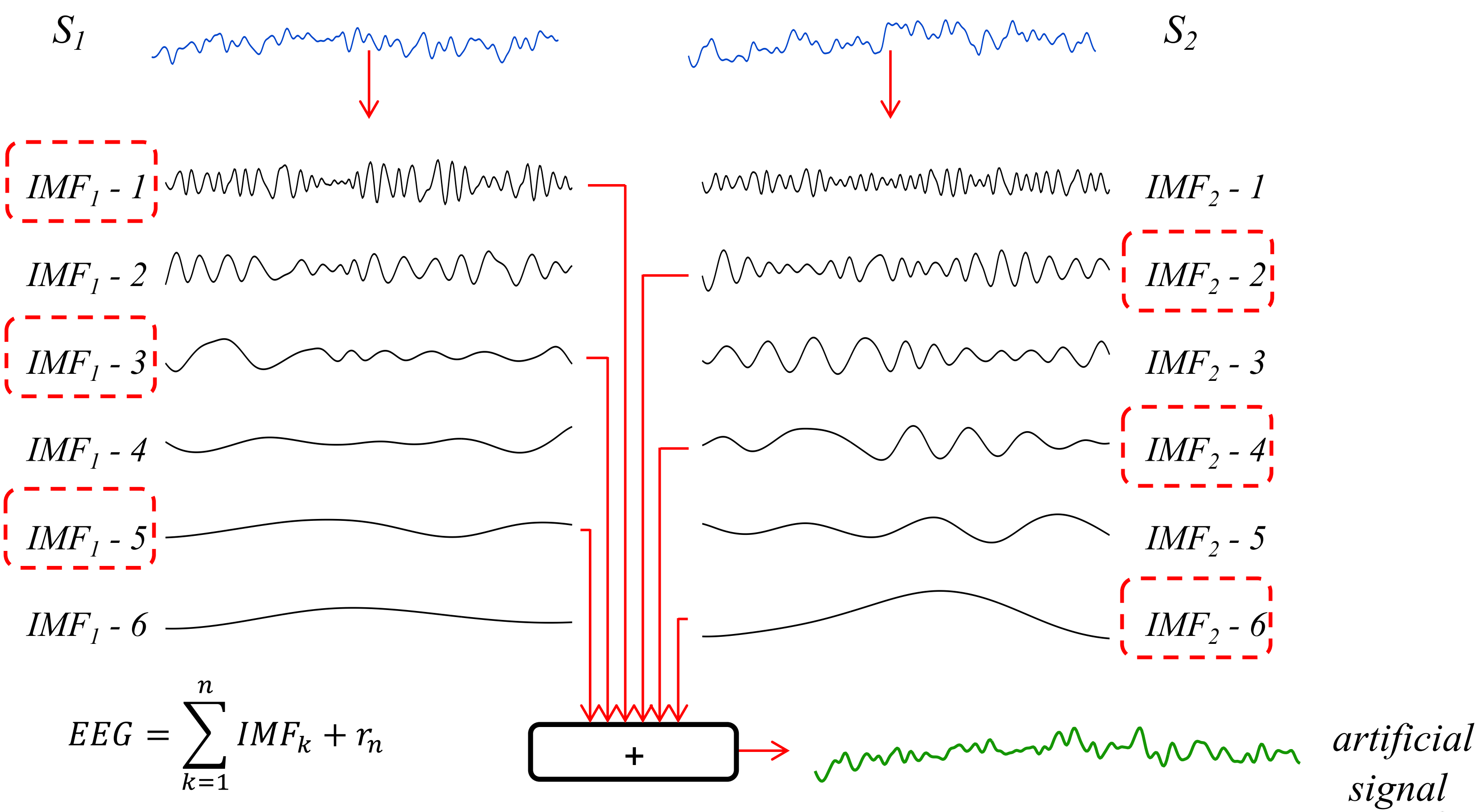
1. Set $s(t) = r_{i-1}(t)$. Initially, $i = 1$ and $r_0(t) = x(t)$.
2. Detect the local maxima and the local minima of $s(t)$.
3. Interpolate all local maxima & minima to generate the upper & lower envelopes.
4. Obtain the local mean $m(t)$ by averaging the upper and lower envelopes.
5. Get a candidate for IMF by subtracting the local mean $m(t)$ from the signal:
$$h(t) = s(t) - m(t).$$
6. If $h(t)$ does not satisfy the IMF's conditions, begin a new loop from step 2, setting:
$$s(t) = h(t).$$
7. Otherwise, $h(t)$ is defined as an IMF: $IMF_i(t) = h(t)$.
8. Set $r_i(t) = r_{i-1}(t) - IMF_i(t)$.
9. If $r_i(t)$ is a monotonic function or does not have enough extrema to calculate the upper and lower envelopes, then $IMF_i(t)$ is the last IMF function of $x(t)$ and the decomposition ends.
10. Otherwise, set $s(t) = r_i(t)$ and start a new loop from step 2 in order to obtain $IMF_{i+1}(t)$.

3. Data augmentation using EMD

- Idea: if one (or more) of these IMFs is replaced by another IMF from another previously decomposed signal, using the above formula, a different signal could be obtained.
- This new signal will maintain the important characteristics of the original ones because it has been generated only by mixing parts of the signals.



3. Data augmentation using EMD



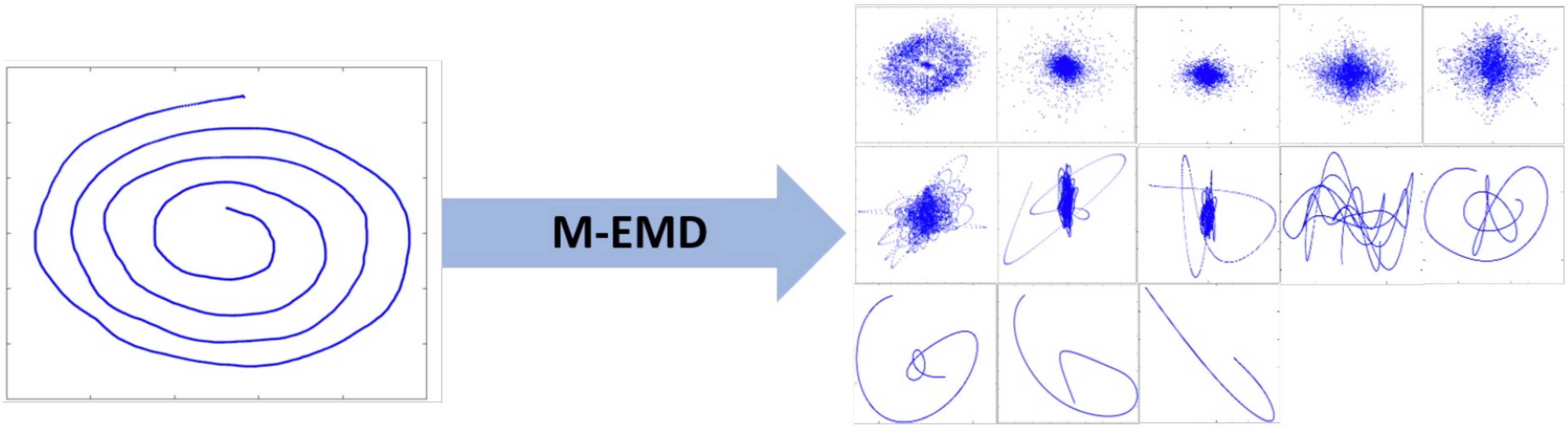
$$EEG = \sum_{k=1}^{n} IMF_k + r_n$$

3. Data augmentation using EMD

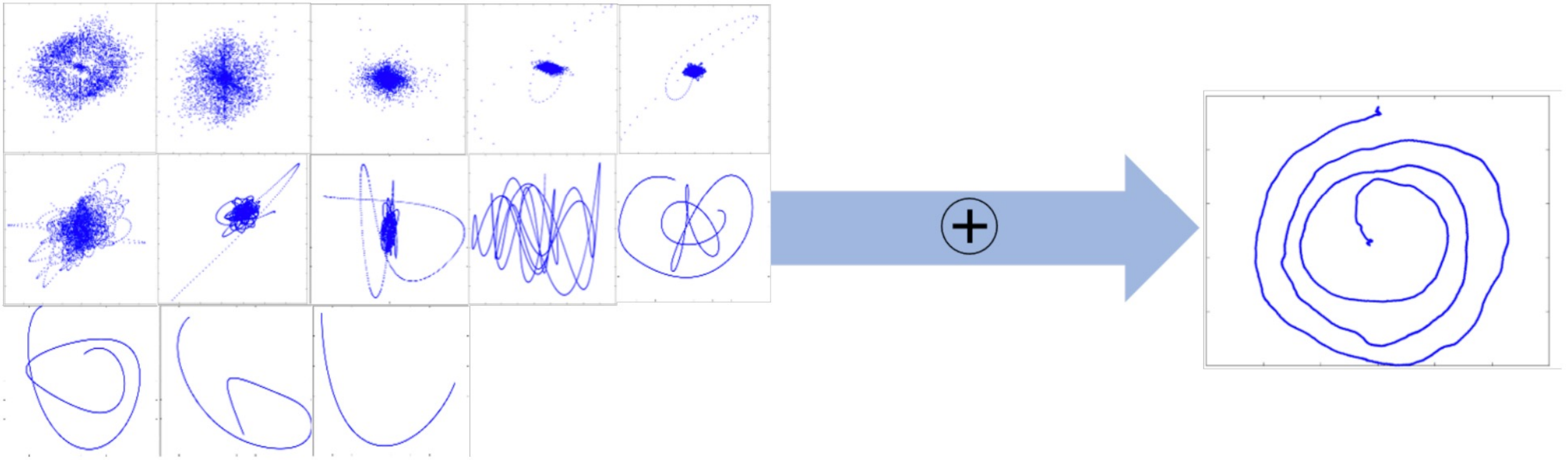
In this work, we propose the use an LSTM deep learning method.

- This is where the need to use a data augmentation method arises because the LSTM model must be adjusted accordingly.
- We will only use *x* and *y* coordinates from the digitizing tablet, following the previous work.
- We will use mEMD to simultaneously decompose the *x* and *y* time series.

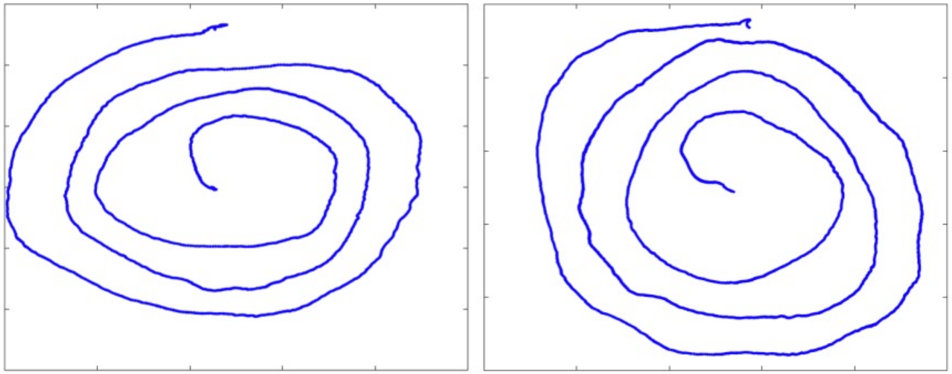
3. Data augmentation using EMD



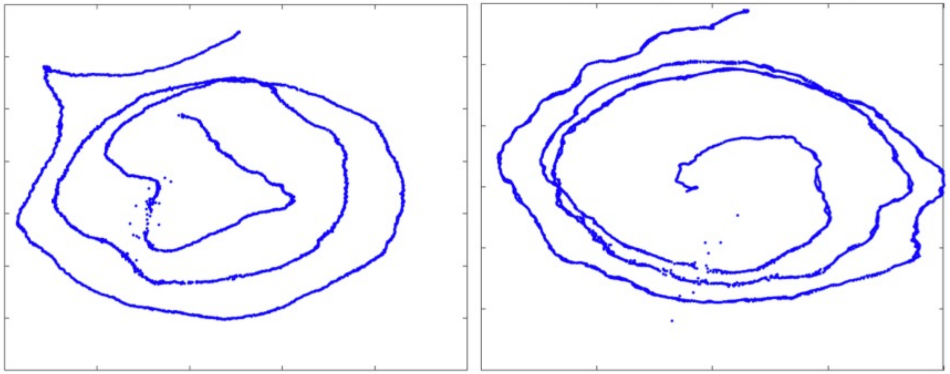
3. Data augmentation using EMD



3. Data augmentation using EMD



Artificial samples from CT group



Artificial samples from ET group

Overview

1 – Introduction

2 – Previous works

3 – Data augmentation using EMD

4 – Experimental results

5 – Discussion and conclusions

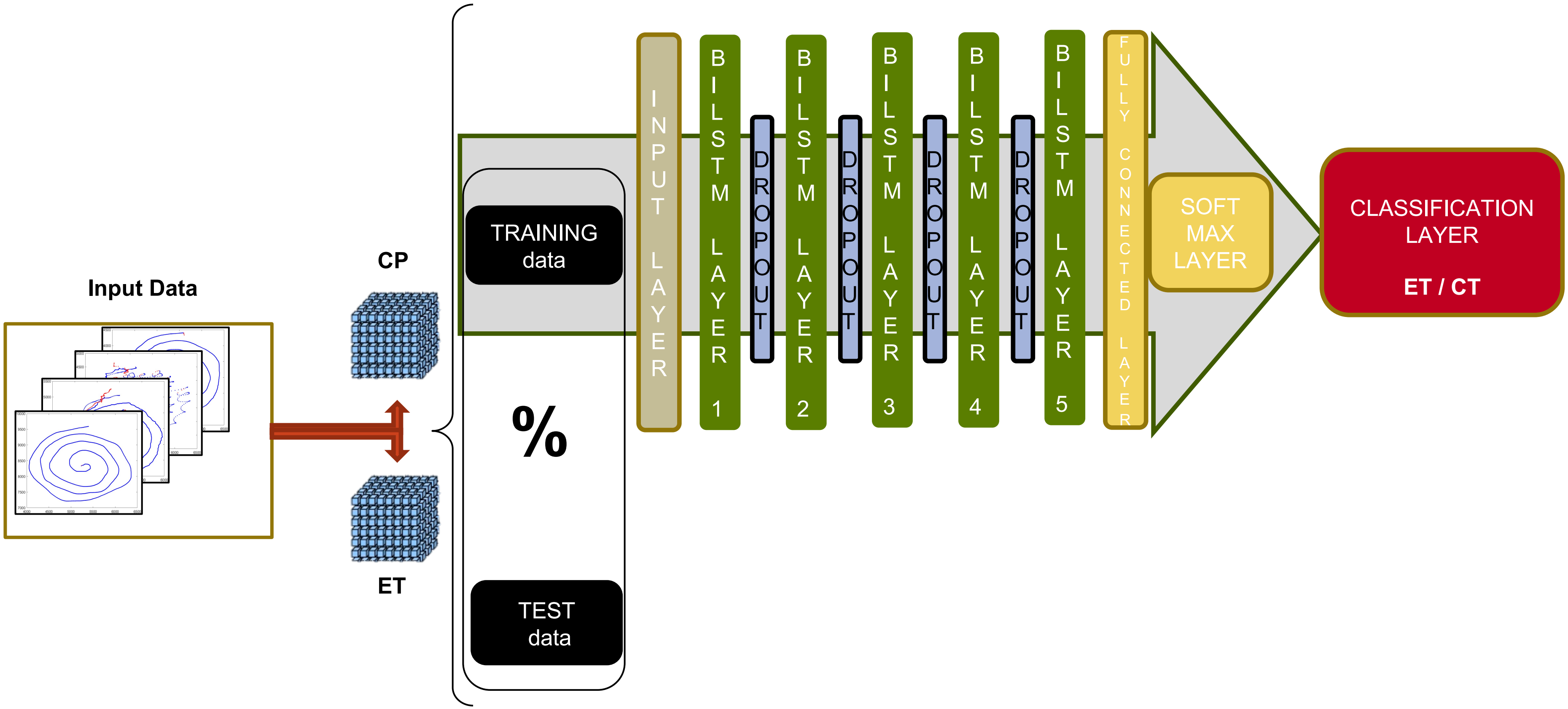
4. Experimental results

- The database used in this work is the subset BIODARWO from the BIODARW dataset.
- It consists of 51 samples: 24 samples of the essential tremor (ET) group and 27 samples of the control group (CT).
- The data available in this database was obtained from a Wacom 4 digitizing tablet at a sampling frequency of 100 Hz.
- Among all the possible features collected by the tablet (X coordinate, Y coordinate, time stamp, azimuth angle, angular angle and pressure), only the X and Y coordinates will be used to derive the classification model.

4. Experimental results

- For the experiment, 50% of subjects were used to train the LSTM model and the remaining 50% were used for testing it.
- The LSTM model was trained in three different scenarios:
  - Without artificial data
  - With 50 new artificial samples per group + the original training samples
  - With 100 new artificial samples per group + the original training samples
- Each experiment was repeated 10 times, with a random selection of subjects used for the training step.
- In each case, the artificial samples were generated using the IMFs of the selected training samples.

4. Experimental results



4. Experimental results

$$AC = \frac{TP+TN}{TP+TN+FP+FN}$$

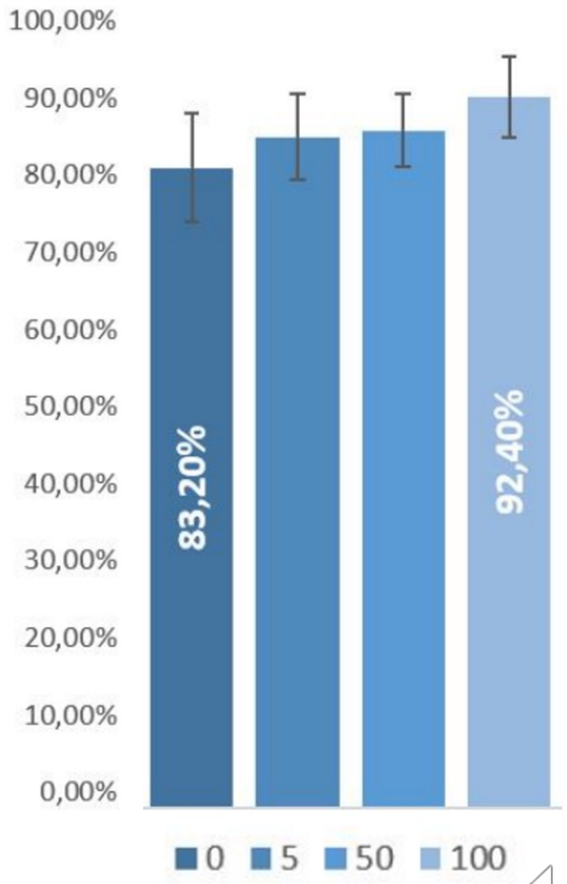
$$SE = \frac{TP}{TP+FN}$$

$$SP = \frac{TN}{TN+FP}$$

where TP, TN, FP, and FN are, respectively, the true positive, true negative, false positive and false negative values of the confusion matrix.

AS		AC	SE	SP
	0	83.2%	68.2%	**95.0%**
	50	88,0%	85,4%	90,0%
	100	**92,4%**	**93,6%**	91,4%

Accuracy, Sensitivity and Specificity of the testing diagnosis with respect to the number of artificial samples added (AS)



Overview

5. Discussion and conclusions

- As we can see, artificial samples always allow for increased accuracy when training and testing the model.
- The best accuracy is obtained with 100 artificial samples per group, with an improvement in accuracy of more than 9% (from 83.20% to 92.40%).
- The sensitivity is also improved in these cases, from 68.18% (no artificial samples) to 93.64%, but the specificity slightly decreases from 95% (no artificial samples) to 91.43%
- However, sensitivity and specificity are more balanced when using artificial data.

5. Discussion and conclusions

We can see that proposed method is powerful

- Only using the X and Y coordinates it achieves 92.40% accuracy
- This result is similar to the first two system using 77 and 84 linear and non-linear features, respectively, extracted from all the available variables generated by the digitizer tablet.

But the result is lower than the 97.96% reported in the third system:

- This system obtained this result after selecting the best 5 features over a set of 35 pre-computed linear and non-linear features (in time and frequency domains)
- This best result was obtained after combining the two different strategies (residual method and radius method)

Finally, note that we have the advantage of not needing to compute features in our LSTM model.

Thank you very much!

Grup de Recerca TDS
Tractament de dades i senyals
UVIC·UCC

UNIVERSITY OF CAMBRIDGE

Universitat Oberta de Catalunya

国立大学法人 東京農工大学
Tokyo University of Agriculture and Technology

RIKEN