Training of Deep Bidirectional RNNs for Hand Motion Filtering via Multimodal Data Fusion

Soroosh Shahtalebi†, S. Farokh Atashzar††, Rajni V. Patel‡, and Arash Mohammadi†

†Concordia Institute for Information System Engineering, Concordia University, Montreal, QC, Canada
†† Tandon School of Engineering, New York University (NYU), USA.
‡Electrical and Computer Engineering, University of Western Ontario, London, ON, Canada

7^{th} IEEE Global Conference on Signal and Information Processing (GlobalSIP'19)

Symposium on Advanced Bio-Signal Processing and Machine Learning for Assistive and Neuro-Rehabilitation Systems

Intelligent Signal and Information Processing (I-SIP) Lab

November 14, 2019

World Aging Trend

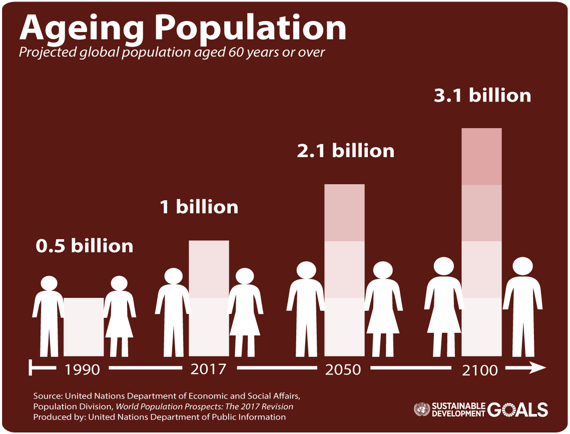


Figure: image taken from: United Nations Department of Economic and Social Affairs, Population Division, *World Population Prospects: The 2017 Revision*

Effects of Aging on Society

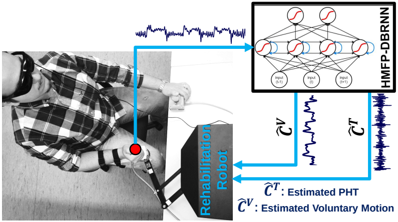
Aging:

- significant increase of the number of seniors over the age of 65,
- prevalent occurrence of age-related neurological disorders such as Parkinson's Disease (PD), Essential Tremor (ET),
- prevalent occurrence of their common motor symptoms such as Pathological Hand Tremor.

Tremor:

- a non-volitional and pseudo-rhythmic movement,
- affects coordination, targeting, and speed of movements in the individuals,
- reduces the ability of individuals to perform the activities of daily living (ADLs),
- affects the quality of life for patients.

Motivating Application: Rehabilitation and Assistive Technologies



(a) Block-diagram of an Augmented Haptic Rehabilitation (AHR) system, where tremor extraction is required to develop a safe haptics-enabled robotic rehabilitation system.

(b) Image taken from: https://keysoftwareservices.co.in/google-presents-smart-spoon-2/ Link to Video

Importance of Tremor Estimation

Clinical: The severity and characteristics of hand tremor are considered as a clinically-viable measure to

- assess the progression of the disease,
- tune the dosage and parameters of therapies, such as Botulinum toxin injection therapy,
- more accurate differential diagnosis of diseases.

Rehabilitative and Assistive Technologies:

- high accuracy in tremor estimation and minimum phase lag are the imperative requirements for the system to deliver the expected degree of performance and safety.

HMFP-DBRNN Architecture

HMFP-DBRNN:

- a data-driven framework based on deep bidirectional recurrent neural networks to extract pathological hand tremor,
- learns the behavior of tremor and voluntary movements through several training examples and provides a means for on-line and off-line estimation/extraction of tremor,
- an assumption-free framework and does not require any fine tuning of the parameters for different subjects,
- takes advantage of a devised training mechanism which addresses both unavailability of ground truth for collected action tremor signals, and the need for providing predictions on the voluntary motion, in a myopic fashion.

Network Architecture

The forward propagation of data in a vanilla RNN is formulated as

$$\boldsymbol{h}(k) = \mathrm{ReLU}\big(\boldsymbol{b} + \boldsymbol{W}\boldsymbol{h}(k-1) + \boldsymbol{U}\boldsymbol{m}(k_1 : k)\big), \quad (1)$$

$$\text{and } \hat{\boldsymbol{y}}(k) = \mathrm{softmax}\big(\boldsymbol{c} + \boldsymbol{V}\boldsymbol{h}(k)\big). \quad (2)$$

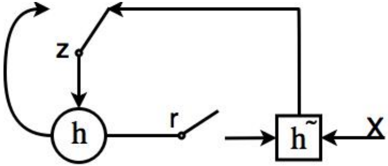
- $\boldsymbol{m}(k_1 : k) = [m(k_1), \ldots, m(k)]^T$ is the input sequence to the network constructed from the hand motion from time $(k_1 < k)$ to time k.
- $\boldsymbol{h}(k)$ represents the hidden states' sequence.
- $\boldsymbol{b}$ denotes the bias vector for the input nodes.
- $\boldsymbol{W}$ is the weight matrix for hidden-to-hidden connections.
- $\boldsymbol{c}$ models the bias vector for the output nodes.
- $\boldsymbol{V}$ denotes the weight matrix for hidden-to-output connections.
- ReLu$(\cdot)$ denotes the Rectified Linear Unit (ReLu) activation function.

The HMFP-DBRNN framework has a bidirectional architecture and employs Gated Recurrent Units (GRU) cells.

Gated Recurrent Units (GRU)

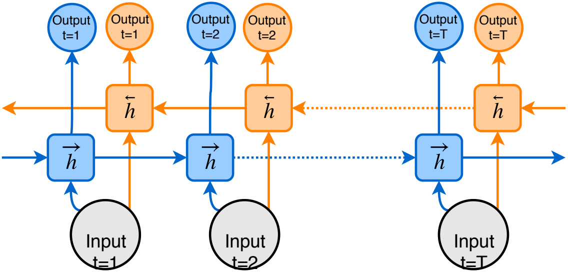
- GRU cells are capable of capturing the dependencies present within different time scales.
- This benefit comes from utilization of two internal gates, i.e., "*update gate*" and "*reset gate*".
- GRU cell is formulated as

$$\begin{aligned}
\boldsymbol{r} &= \sigma\big(\boldsymbol{U}_r\boldsymbol{m}(k_1:k) + \boldsymbol{W}_r\boldsymbol{h}(k-1)\big), && (3)\\
z &= \sigma\big(\boldsymbol{U}_z\boldsymbol{m}(k_1:k) + \boldsymbol{W}_z\boldsymbol{h}(k-1)\big), && (4)\\
\tilde{\boldsymbol{h}}(k) &= \mathrm{ReLU}\big(\boldsymbol{U}\boldsymbol{m}(k_1:k) + \boldsymbol{W}(\boldsymbol{r}\odot\boldsymbol{h}(k-1))\big), && (5)\\
\boldsymbol{h}(k) &= z\boldsymbol{h}(k-1) + (1-z)\tilde{\boldsymbol{h}}(k). && (6)
\end{aligned}$$



Bidirectional Recurrent Neural Network

- Bidirectional architecture provides a processing tool for both on-line and off-line (tuning) tasks.
- In vanilla RNNs, the cells which are analyzing the initial samples of the input sequence do not provide an accurate output, and bidirectional architecture can address this issue.



Motivation for Data Fusion

- Due to the data-hungry nature of deep neural networks, and unavailability of large datasets in medical fields, the application of deep learning methods may seem to be still limited.
- Neural networks trained over shallow datasets do not generalize well, and overfitting of the model over the studied phenomenon is always a possibility.
- In this work, we investigate the feasibility of combining two different multimodal datasets, collected under two different conditions with two different experimental setups, in order to train a tremor extraction neural network.
- The data fusion strategy is taken to improve the generalization of the model.

Employed Datasets

Motus Dataset:

- single channel recordings of patients with hand tremor,
- recorded with a bi-axial gyroscope, which is mounted on dorsum of hand,
- available online, courtesy of Motus Bioengineering Inc., Benicia, CA,
- sampling frequency of the signals is 100 Hz,
- the angular velocity of the movements is recorded,
- 5 sets of rest tremor and 5 sets of action tremor recordings are available from 10 patients.

Smartphone Dataset:

- tremor recordings of 10 patients with PD,
- recorded with the built-in tri-axial accelerometer of a smartphone (iPhone 5s) by placing it on the dorsum of hand,
- acceleration of hand motion is recorded in 3 axis,
- sampling rate is 100 Hz.

Data Fusion Strategy

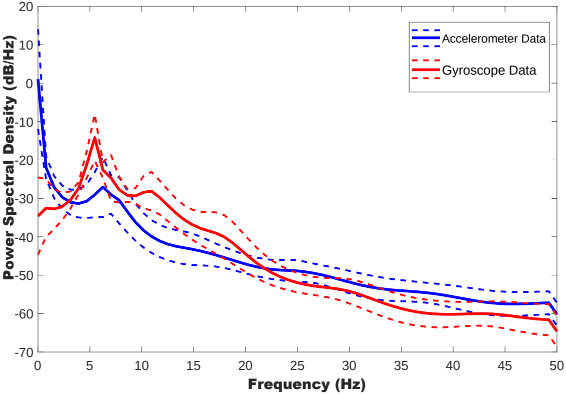


Figure: Representation of the Power Spectral Density (PSD) for Motus and Smartphone datasets. The mean of the PSDs along with its standard deviation lines are plotted for the two groups.

Training Mechanism

- We use real rest-tremor data combined with generated voluntary components to produce input-output pairs to train and validate the network.
- For the voluntary part, a sinusoidal waveform with random amplitude, frequency and phase is generated based on the following three uniform distributions $\sim \mathcal{U}(0, 0.25)$, $\sim \mathcal{U}(0, 3)$ and $\sim \mathcal{U}(0, \pi)$, respectively.
- To address the urge for predicting the voluntary component, at least one sample ahead of time, we form segments of length $N_s + 1$ samples for $\boldsymbol{c}^T$, $\boldsymbol{c}_a^V$ and $\boldsymbol{m}_a$. Then we provide $\boldsymbol{m}_a(1 : N_s)$ as the input to the network, and $\boldsymbol{c}_a^V(2 : N_s + 1)$ as the target of the network.

Simulation Results

Table: Quantitative testing of HMFP-DBRNN.

Scheme	MSE	NRMSE	PRF
Motus	0.001	0.0632	0.004
Smartphone	0.0019	0.0872	0.0076
Motus+Smartphone	0.0022	0.0938	0.0088

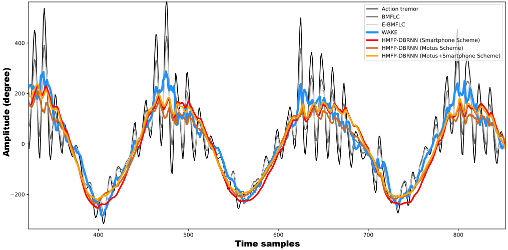


Figure: Performance of HMFP-DBRNN trained via three different schemes over the action tremor recordings.

Simulation Results

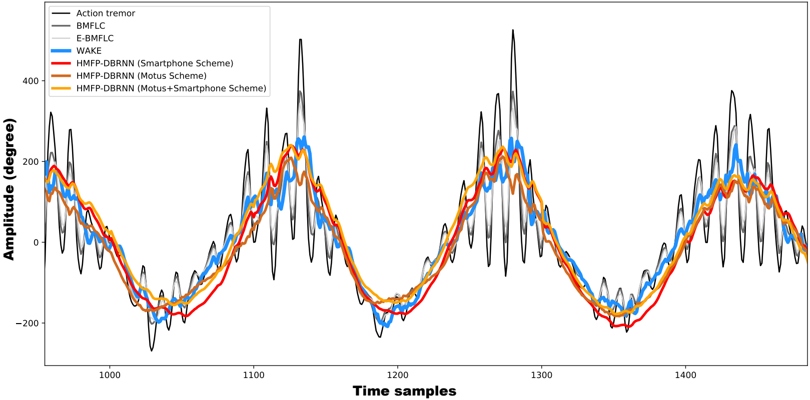


Figure: Performance of HMFP-DBRNN trained via three different schemes over the action tremor recordings.

Conclusion

- In this work, we investigated the idea of training a neural network by fusing multimodal datasets, which have recorded the same phenomenon (i.e., PHT) but with different devices.
- The multimodal fusion of datasets is evaluated based on our recently proposed HMFP-DBRNN framework, which offers the state-of-the-art results in the field of tremor extraction.
- As the results suggest, fusing the datasets, under certain conditions and at the cost of slightly higher estimation error, grants the network an acceptable degree of generalization over both of the datasets.

References

1. S. Shahtalebi, S. F. Atashzar, R. V. Patel, and A. Mohammadi, "Training of Deep Bidirectional RNNs for Hand Motion Filtering via Multimodal Data Fusion," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2019.

2. S. Shahtalebi, S. F. Atashzar, R. V. Patel, and A. Mohammadi, "HMFP-DBRNN: Real-time Hand Motion Filtering and Prediction via Deep Bidirectional RNN," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1061–1068, 2019.

3. S. Shahtalebi, S. F. Atashzar, R. V. Patel, and A. Mohammadi, "Wake: Wavelet Decomposition Coupled with Adaptive Kalman Filtering for Pathological Tremor Extraction," *Biomedical Signal Processing and Control*, vol. 48, pp. 179–188, 2019.

4. S. Shahtalebi, A. Mohammadi, S. F. Atashzar, and R. V. Patel, "A Multi-rate and Auto-adjustable Wavelet Decomposition Framework for Pathological Hand Tremor Extraction," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 432–436, 2017.

5. S. F. Atashzar, M. Shahbazi, O. Samotus, M. Tavakoli, M. S. Jog, R. V. Patel, "Characterization of Upper-limb Pathological Tremors: Application to Design of an Augmented Haptic Rehabilitation System", in *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, pp. 888-903, 2016.

Thanks for your attention!