Stress and Affect Detection on Resource-Constrained Devices

Scalable Deep Learning for Time-Series Data

Abhijith Ragav, **Nanda H Krishna**, Naveen Narayanan, Kevin Thelly
and Vineeth Vijayaraghavan

Contents

Overview

- Psychological stress in humans has constantly been rising, over **74%** of people interviewed in [1] were unable to cope with stress.
- Chronic stress has fatal consequences, including **cancer**, **heart disease** and **suicide**.
- **Early detection of stress** is key, possible through **continuous monitoring** using wearables and physiological sensor data.
- Data must be kept **private**, and **real-time** affective state classification must be **quick**, **accurate** and **efficient**.
- **On-device** classification eliminates need for a remote server, as well as problems of privacy and latency.
- We thus discuss **scalable deep learning models** for stress and affect detection on resource-constrained devices.

Datasets

Note

We focus on stress and affect detection from physiological modalities, which are great indicators of stress and can be recorded on a wearable device.

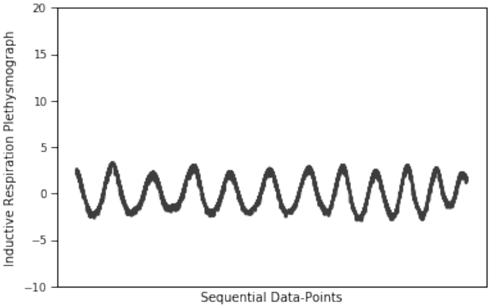
1. **WESAD** – Wearable Stress and Affect Detection Dataset [2]
   - 3 affective states – **Neutral**, **Stress**, **Amusement**.
   - Physiological and motion sensor data from chest-worn and wrist-worn devices.
   - We use **chest-worn device** data, due to **better classification performance** as shown in [2].
   - ECG, EEG, EMG, EDA, Temperature, Respiration, Accelerometer.
   - Benchmark – **87.18%** ANN [3].

Datasets

2. **SWELL-KW** – SWELL Knowledge Work Dataset [4]
   - 3 conditions – **Neutral (N)**, **Interruption (I)**, **Time-Pressure (T)**.
   - **Binary classification** of N versus I&T.
   - Various modalities, physiological modalities are HRV and SC.
   - Benchmark – **64.10%** SVM [5].
3. **DREAMER** – Database for Emotion Recognition [6]
   - 3 affective states – **Arousal**, **Dominance**, **Valence** – values range from 1 to 5.
   - **Multi-class classification** for each.
   - EEG and ECG signals.
   - Benchmark – **61.84%** for Arousal and Dominance, **62.32%** for Valence, SVM [6].

Data Analysis

- Respiration, Neutral State, WESAD

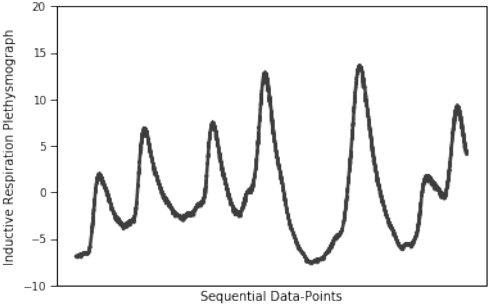


Note

In the neutral state, values lie between -5 and 5. We may understand values in this range to be representative of this class.

Data Analysis

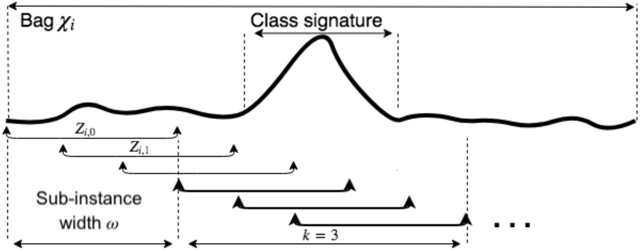
- Respiration, Stress State, WESAD



Note

It is evident that quite a few values lie between -5 and 5. Fewer values lie outside this range, and are representative of stress.

Multiple Instance Learning



- **Class signature** – Points which best represent the class.
- **Noise** – Points which do not truly represent the class.
- Class signature is **infrequent** compared to noise.
- MIL is a technique employed to **choose the best data points** to train the model. We use the technique illustrated in [7].

Multiple Instance Learning

- The dataset is split into **bags** of equal sizes. Class labels are bag-level.
- Every consecutive set of ω points in a bag constitutes a **sub-instance**.
- We choose the best k **consecutive sub-instances** to train the model.
- Helps in minimizing learning from noise, **maximizing learning from class signature**.

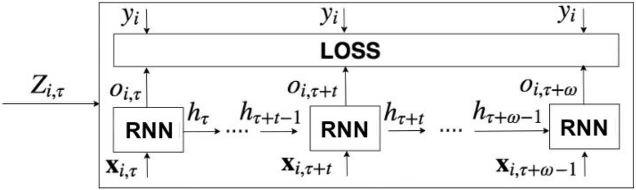
Note

This exploits the fact that the class signature is present in consecutive data points, which is a property of such time-series data.

Early Stopping

- LSTMs iterate through **all data-points** before a prediction is made.
- Computationally expensive, **increases inference time**.
- With Early Stopping, the output of **every step** of the LSTM is piped to a classifier.
- If a prediction is made with a probability greater than a **threshold** $\hat{p}$, the LSTM **delivers the prediction** without iterating through the rest of the steps.
- It is efficient to predict the probability after every k steps.
- Leads to greatly reduced inference times on resource-constrained devices.

EMI-RNN



- **Joint architecture** incorporating MIL and Early Stopping.
- MIL reduces the number of sub-instances learned from, Early Stopping delivers early predictions.
- Great reduction in inference time, **overall benefit greater than sum of parts** [7].
- We employ the EMI-RNN for stress and affect detection, with parameters chosen carefully for each dataset.

Choice of EMI-RNN Parameters

- Bag size (T)

$$B = \left\{ b : \frac{\nu_s}{m} \leq b \leq \frac{\nu_s}{n}, b \in \mathbb{N} \right\}$$

$$T = \arg\max_{b} f(b), b \in B$$

- Sub-instance width (ω)

$$W = \left\{ \omega_s : \frac{T}{p} \leq \omega_s \leq \frac{T}{q}, \omega_s \in \mathbb{N} \right\}$$

$$\omega = \arg\max_{\omega_s} g(\omega_s), \omega_s \in W$$

- Parameters chosen for each dataset

Dataset	T	ω	k
WESAD	175	88	2
SWELL-KW	20	10	3
DREAMER	128	48	2

Models Used

- We use **EMI-LSTM**, **EMI-GRU** and **EMI-FastGRNN**.
- LSTMs and GRUs are popular RNN architectures, perform well on sequential data.
- FastGRNN [8] out-performs unitary RNNs with respect to **training time**, **prediction cost** and **accuracy**.
- It learns a classifier of **very low model size and inference time**, but near **state-of-the-art accuracy**.
- These models are compared with the current state-of-the-art models for each dataset with respect to accuracy and inference time.

Evaluation Metrics

- We use **accuracy** and **inference time** to evaluate the performance of our models.
- Additionally, we introduce the metrics η_{comp} or **computational savings** and η_{ES} or **early savings**.

$$\eta_{\text{comp}} = \frac{I_B \text{ - } I_M}{I_B} \times 100$$

$$\eta_{\text{ES}} = \left(1 - \frac{1}{n}\sum_{i=1}^{n} \frac{T_{pred_i}}{T}\right) \times 100$$

- Accuracy and inference time are given greater weightage in choosing the best model.

Results

1. WESAD
   - **EMI-FastGRNN** performs best.
   - Accuracy **97.55%** and inference time **12.8 ms**.
   - Inference time **13x lesser** than the benchmark.
   - Leads to 92% computational saving, 96.07% early saving.
2. SWELL-KW
   - **EMI-FastGRNN** performs best.
   - Accuracy **87.87%** and inference time **3 ms**.
   - Inference time **35x lesser** than the benchmark.
   - Leads to 97.10% computational saving, 29.55% early saving.
3. DREAMER
   - **EMI-LSTM** performs best.
   - Accuracy for Arousal, Dominance and Valence were **67.75%**, **68.87%** and **65.60%**.
   - Inference time **49.2 ms**, **8.5x** lesser than the benchmark.
   - Leads to 87.88% computational saving, 20.52% early saving.

Conclusion

- EMI models **outperformed** the benchmark models with respect to accuracy and inference time.
- On average, there was an absolute **increase of 10%** in overall accuracy and **18x reduction** in inference times compared to the benchmark models.
- Average computational savings across the 3 datasets was 95.39%.
- The EMI-FastGRNN outperformed the EMI-LSTM and EMI-GRU for the WESAD and SWELL-KW datasets.
- EMI models are suitable for use on resource-constrained devices to make **real-time** and **on-device** predictions, with high accuracy and low inference time.

References

[1] Mental Health Foundation (2018)
Mental health statistics: stress.
https://www.mentalhealth.org.uk/statistics/mental-health-statistics-stress

[2] Schmidt, P., Reiss, A., Duerichen, Marberger, C. and Van Laerhoven, K. (2018)
Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection.
ICMI '18

[3] Marković, D., Vujičić, D., Stojić, D., Jovanović, Ž., Pešović, U. and Ranđić, S. (2019)
Monitoring System Based on IoT Sensor Data with Complex Event Processing and Artificial Neural Networks for Patients Stress Detection.
18th International Symposium INFOTEH-JAHORINA (INFOTEH)

References

[4] Koldijk, S., Sappelli, M., Verberne, S., Neerincx, M. A. and Kraaij, W. (2014)
The SWELL Knowledge Work Dataset for Stress and User Modeling Research.
ICMI '14

[5] Koldijk, S., Neerincx, M. A. and Kraaij, W. (2018)
Detecting Work Stress in Offices by Combining Unobtrusive Sensors.
IEEE Transactions on Affective Computing, 9 (2)

[6] Katsigiannis, S. and Ramzan, N. (2018)
DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals From Wireless Low-cost Off-the-Shelf Devices.
IEEE Journal of Biomedical and Health Informatics, 22 (1)

References

[7] Dennis, D. K., Pabbaraju, C., Simhadri, H. V., and Jain, P. (2018)
Multiple instance learning for efficient sequential data classification on resource-constrained devices
NeurIPS '18

[8] Kusupati, A., Singh, M., Bhatia, K., Kumar, A., Jain, P. and Varma, M. (2018)
FastGRNN: A Fast, Accurate, Stable and Tiny Kilobyte Sized Gated Recurrent Neural Network.
NeurIPS '18