

# Activity-Aware Deep Cognitive Fatigue Assessment using Wearables

Mohammad Arif Ul Alam

**Abstract**—Cognitive fatigue is a common problem among workers which has become an increasing global problem. While existing multi-modal wearable sensors-aided automatic cognitive fatigue monitoring tools have focused on physical and physiological sensors (ECG, PPG, Actigraphy) analytic on specific group of people (say gamers, athletes, construction workers), activity-awareness is utmost importance due to its different responses on physiology in different person. In this paper, we propose a novel framework, Activity-Aware Recurrent Neural Network (*AcRoNN*), that can generalize individual activity recognition and improve cognitive fatigue estimation significantly. We evaluate and compare our proposed method with state-of-art methods using one real-time collected dataset from 5 individuals and another publicly available dataset from 27 individuals achieving max. 19% improvement over the baseline model.

## I. INTRODUCTION

Cognitive fatigue is a syndrome conceptualized as resulting from chronic workplace stress that has not been successfully managed [1]. Although, cognitive fatigue is not a clinical condition which can occur in any workplace or home environment where there is stress, it is recognized by the World Health Organization (WHO) as a syndrome [1]. In short term, cognitive fatigue may cause sleeping disturbances, anxiety, irritability and hormonal disturbances and in long run, this may result more severe impacts on health safety such as cardiovascular, gastrointestinal and neuropsychological disorders [2].

Current frameworks for cognitive fatigue estimation are mostly self-reported questionnaire based [18], which is impossible to generate continuous fatigue report by avoiding recall bias [18]. Recent advancement of wearable physical and physiological sensor technologies enable accurate estimation of cognitive fatigue related partial outcomes such as stress, anxiety, sleep quality etc, which provides ultimate opportunity to researchers to estimate cognitive fatigue continuously [15], [16] that includes actigraphy [17], heart rate (HR) [17], Electrocardiography (ECG) and Electromyography (EMG) sensors along with traditional and deep machine learning techniques. Combining accelerometer with ECG is a successful attempt as well before [3] which proposed to use deep learning frameworks (LSTM with Consistency Self-Attention, LSTM-CSA) but suffers with the lack of adaptability across diverse population.

Due to the dissimilarities among different individual group's responses on cognitive fatigue in terms of physical and physiological contexts, current wearable-based fatigue estimation research is constrained in group specific cognitive

fatigue estimation [19]. As per many clinical psychologists and mental health researchers, cognitive fatigue estimation should be more personalized, rather than generalized on specific group of people to keep mental healthcare systems sustainable for future generations [20].

While, the needs of building personalized cognitive fatigue estimation tool, we have the following key question: *can we develop personalized cognitive fatigue assessment tool considering their activity as a domain invariant feature and fatigue as personalized response on each activity?* As we know, autonomic nervous system (ANS) restrains the body's major physiological activities including the heart rate (HR) and gland secretion or electrodermal activity (EDA) [21]. However, these responses are contaminated with physical activity artifacts significantly [22]. The central **hypothesis** of the this paper is: *each performed activity context generates similar artifacts on same activity over diverse population, thus, we can align similar activities (activity-awareness) as person invariant feature and its physiological responses as personalized fatigue feature.* For example, in Fig. 1, we illustrate the EDA responses on two different activities: (i) steady hand (ii) waving hand over two different physiological states: (i) stress and (ii) no stress. The Fig. 1 clearly shows that same activity (waving hands) has similar EDA response patterns (but different amplitude) due to similar artifacts which signifies our hypothesis.

In this paper, we develop a novel Activity-Aware Recurrent Neural Network (*AcRoNN*) model and utilize it to design a personalized cognitive fatigue assessment framework, that provides the following **key contributions**

- We develop a novel Activity-Aware Recurrent Neural Network (*AcRoNN*) framework that is able to exploit contextual cues present in any event from actigraphy sensor and then assess cognitive fatigue from physiological (EDA and HR) sensor signal using a deep recurrent neural network.
- Apply *AcRoNN* on three datasets and evaluate the capability of *AcRoNN* framework to improve cognitive fatigue assessment.

## II. ACTIVITY-AWARE RECURRENT NEURAL NETWORK

Fig. 3 shows the overall schematic diagram of our context-aware cognitive fatigue assessment framework. In this framework, recognized activity information is intertwined with the cognitive fatigue assessment architecture. We take (stage 1) cognitive fatigue assessment scoring in terms of gestural activity, and then postural activity information, and finally we (stage 2) re-evaluate the cognitive fatigue assessment scoring based on activity relationships it has learned from stage 1.

Department of Computer Science, University of Massachusetts Lowell  
mohammadarifullalam@uml.edu

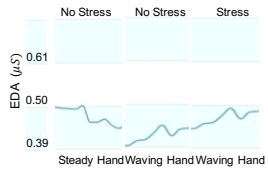


Fig. 1. Electrodermal activity (EDA) responses (micro Siemens  $\mu S$ ) on 2 seconds of 2 dominant hand gestures: steady hand and waving hand twice during two physiological states: stress and no stress of the same subject. Note that, the subject was in sitting position.

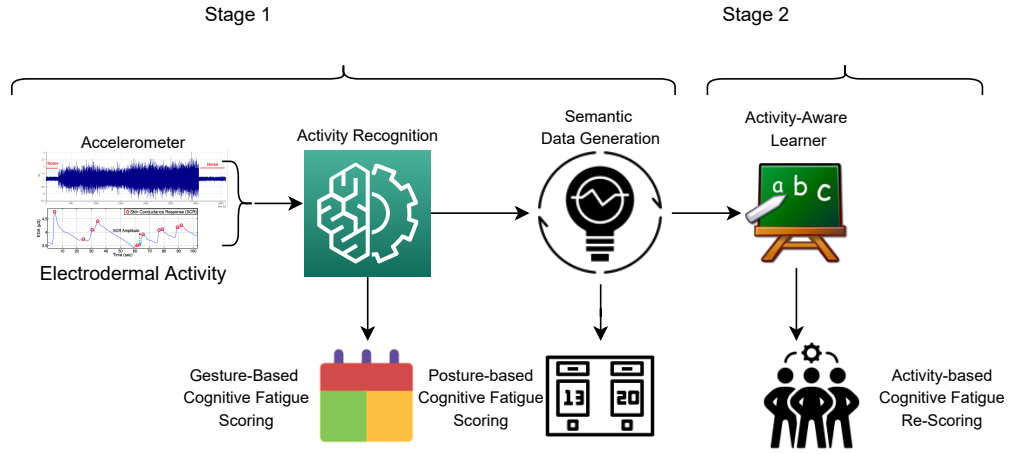


Fig. 2. High-level schematic diagram of *AcRoNN* architecture that consists two stages of learning

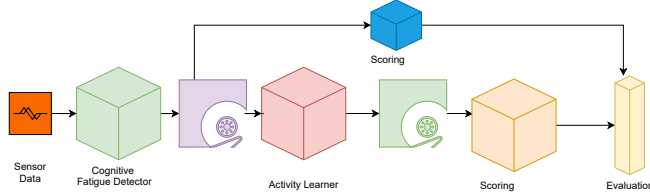


Fig. 3. Proposed Activity-Scoring and LSTM-CSA (Consistency Self-Attention) based two stage *AcRoNN* model

#### A. Stage One: Fatigue Detection and Feature Mapping

This stage involves feature extraction, activity recognition and activity-based score mapping for fatigue detection.

1) *Wearable Sensor Signal Processing*: Wearable sensors can be two types: physical and physiological. Physical sensors (accelerometer, gyroscope etc.) signal values change over the movements of the sensor devices. Physiological sensors change over physiological condition of body such as EDA changes over stress and PPG changes over heart rate. However, physical movements also impose noises on physiological sensor signals which is called *motion artifacts*.

**Physiological Signal Processing**: A continuous and discrete decomposition of EDA, and time and frequency domain analytics of EDA signal are investigated before to extract relevant physiological features which were contaminated with noises and motion artifacts [25]. [26] denoised and classified EDA from cognitive load and stress with accuracy higher than 80%. Though motion artifacts removal techniques such as exponential smoothing and low-pass filters provide significant improvement in filtering EDA signals, wavelet transforms offer more sophisticated refinement for any kind of physiological sensors such as electroencephalogram, electrocardiogram [24], and PPG [27]. [28] proposed a stationary wavelet transform (SWT) based motion artifacts removal technique. ‘cvxEDA’ proposed a convex optimization technique considering EDA as a mixture of white gaussian noise, tonic and phasic components where white gaussian noise includes motion artifacts and external noises [24]. We combine SWT and ‘cvxEDA’ together to remove noises and motion artifacts from EDA signal. We used Periodic Moving Average Filter (PMAF) to reduce

noise. Then, we generated 33 heart rate variability (HRV) features from PPG (as per [3]) and 12 statistical features from EDA (as per [21]) signals with a 10-seconds window.

**Accelerometer Signal Processing**: We used Bai et. al. proposed accelerometer signal processing method [3]. We used ActiLife tool [29], and calculated the Actigraphy counts (from accelerometer) every 5 seconds, and detect the non-wear time (for invalid data removal). Within every 10-seconds window, based on Actigraphy counts we further extracted 8 statistical features, i.e., mean, median, standard deviation, variance, minimum value, maximum value, skewness and kurtosis for further processing.

**Multimodal Feature Sequence Construction**: After preprocessing and feature engineering, the original segment can be transformed into a D-dimensional sequence  $X = \{x_t \in \mathcal{R}^D\}_{t=1}^T$  where  $T$  is the sequence length (i.e., the number of windows/epochs within a segment), and  $x_t = feat_{acc} \cup feat_{eda} \cup feat_{hrv}$  where  $feat_{acc}$ ,  $feat_{eda}$ ,  $feat_{hrv}$  represent accelerometer, EDA and HR features extracted above. Since, each of the extracted features were in 10-seconds window, the concatenated input feature  $x_t$  has a dimension of 53 (33 + 12 + 8).

2) *Activity Recognition Module*: We develop a two step multi-label activity recognition framework which consists of two LSTM with Consistency Self-Attention (LSTM-CSA) [23] models, (1) gestural activity recognition and (2) postural activity recognition. Both of the LSTM-CSA models are independent from each other, trained and tested separately using hand gesture and postural activity labels respectively using the input accelerometer features ( $feat_{acc}$ ) and their corresponding labels. For all LSTM-CSA models, we used the following regularization term as  $\Gamma(\alpha) = T \sum_t |\alpha_t - \alpha_{t-1}|$  where  $T$ ,  $\Gamma(\alpha)$  tend to penalize heavily with a larger contextual scores to maintain its global consistency.

3) *Class Contextual Feature Maps*: We develop a contextual feature mapping for each cognitive fatigue label as

$$cfm_c(x, y) = \sum_{cls(b)=c} hm_b(x, y)$$

$$CFM_c = \frac{cfm_c}{max(cf m_c)} \quad (1)$$

The output of the contextual feature map layer is a  $(H_c \times W_c \times C)$  tensor, where  $H_c$  and  $W_c$  are the segment dimensions, and  $C$  is the number of classes. We set  $H_c = 23$  while we have already defined  $W_c = 53$ .

4) *Contextual Scoring*: We designed a scoring function to measure the contextual relevance of cognitive fatigue detection in relation to multiple-label's presence in a window [30]. Scores are computed using the contextual feature maps generated in each stage of our pipeline, and are used as the ranking score in AP calculations to measure whether contextual learner confirms or refutes detections passed to it, based on learned semantic relationships. The scoring process is designed as a new network layer, and appended to the end of each stage in our pipeline. We defined two contextual-scoring method as per [30].

- **Individual Contextual Scoring**: We calculate gestural and postural-based cognitive fatigue score as

$$Score_1(b) = \frac{\sum FM_b(x, y)}{2\sigma_{bx}^2 \times 2\sigma_{by}^2} \quad (2)$$

where  $FM_b$  represents the activity bounding box (start and end of an activity) related relevance score and  $b$  represents each activity type i.e. gestural or postural activity as per [30]. We have two types of individual contextual scoring in our framework, gestural-based cognitive fatigue scoring and postural-based cognitive fatigue scoring (Fig. 3).

- **Cumulative Contextual Scoring**: In this scoring method, we add both gestural and postural activity based cognitive fatigue scoring together for producing final Activity-based cognitive fatigue re-scoring which can be defined as follows

$$Score_2(b) = \frac{\sum CFM_b(x, y)}{2\sigma_{bx}^2 \times 2\sigma_{by}^2} \quad (3)$$

where  $CFM_b$  represents the cumulative activity bounding box (start and end of an activity) related relevance score and  $c$  represents cumulative activity type i.e., either gestural or postural activity-based cognitive fatigue scoring or re-scoring [30].

### B. Stage Two: Activity-Aware Cognitive Fatigue Learner

The second stage is an LSTM with Consistency Self-Attention (LSTM-CSA) model that is trained to learn semantic relationships using the cumulative contextual score mapping generated by the primary cognitive fatigue detector using the Equation 1.

## III. EXPERIMENTAL EVALUATION

In this section, we aim to evaluate our proposed Activity-Aware Recurrent (*AcRoNN*) performance towards developing a personalized cognitive fatigue assessment system using wearables without any target labels.

### A. Datasets

**A1: Activity Recognition Dataset**: It is our previously collected data consists of hand gestural (8-hand gestures) and postural (#4) activities [22], [31] using Empatica E4 watch.

**D1: Gamer's Fatigue Dataset**: We recruited 5 student video games players (age ranges from 19-25) for 7 days who stayed up during a 22 hour shift every alternative day (4 days each) to simulate cognitive fatigue while wearing Empatica E4 watch [4]. Empatica E4 watch consists of accelerometer (ACC), electrodermal activity (EDA), photoplethysmography (PPG) and skin temperature (TEMP) sensors. During the data collection (including non-gaming days), participants were asked to measure their sleepiness based on the 'Stanford Sleepiness Scale' (SSS) [5], [8] (ranges 1-7 representing active to extremely sleepy) and the 'Sleep-2-peak' score [6] (ranges 1-7 representing active to extremely sleepy) using Sleep2Peak Android App [7], [9].

**D2: Healthy Adults Fatigue Dataset**: We have used publicly available health adults fatigue dataset [10]. Data from 28 healthy individuals (26–55 years of age, average age 42 years, 41/51% female/male), of which 17 enrolled up to 2 days after returning from long-haul flights with 3–7 time zone differences and hence were recovering from jet lag, from 1 to 219 consecutive days (973 days) were collected. Objective data were collected using a multisensor wearable device, Everion (Biovotion AG, Switzerland [11]), in conjunction with a mobile app, SymTrack (Gastric GmbH, Switzerland), to deliver a daily fatigue questionnaire. Volunteers were asked to continuously wear the Everion device around their non-dominant arm over a 1-week period. The device combines a 3-axis accelerometer, barometer, galvanic skin response electrode, and temperature and photo sensors. Dataset tracked a total of 12 parameters at 1-Hz temporal resolution on physical activity and physiology. Volunteers were instructed to complete a 4-item daily questionnaire in the evening to capture their subjective assessment of fatigue, adapted from the Fatigue Assessment Scale [12] and Visual Analogue Scale to evaluate fatigue severity [13]: (1) Physical fatigue score (PhF), (2) Mental fatigue score (MF), (3) Visual analogue scale score (VAS), and, (4) Indicator of relative perception (RelP) (see [10] for more details).

### B. Pre-Processing

We excluded days where more than 80% of the samples were missing to ensure an acceptable performance of downstream analysis. Missing samples were due to subjects not wearing the device (e.g., during charging) or low-quality segments (e.g., loss of skin contact). This filtering step led to a total of 5 subjects and 821 hours of data annotated (Stanford Sleepiness Scale and Sleep-2-peak) labels with continuous Empatica E4 sensor data (excluding 1 hour recharging sessions). Finally, we imputed missing data gaps using the state-of-the-art unidirectional uncorrelated recurrent imputation model [14].

### C. Baseline Algorithm development

We re-implemented latest cognitive fatigue estimation framework [3]. In Bai et. al. [3], authors generated highest

accuracy of cognitive fatigue assessment using LSTM-CSA model. Although, Bai et. al. provided cognitive fatigue detection method based on PPG (ECG) and Accelerometer sensor signal processing, we implemented the following baselines from Bai et. al. method as follows.

- **B1:** In this framework, we considered original Bai et. al. approach i.e. using only ECG and Actigraphy features (total number of features  $41 = 33+8$ ) and followed Bai et. al. to re-implement original baseline paper’s result without any alteration.
- **B2:** In this framework, we considered our 53 features that include Actigraphy, PPG and EDA features and applied Bai et. al. proposed LSTM-CSA model for cognitive fatigue assessment.
- **B3:** In this framework, we combined our activity recognizer produced hand gesture and posture detections, and applied Bai et. al. considered 41 features (Actigraphy and ECG) and used our proposed *AcRoNN* framework.
- **AcRoNN:** We combine everything together (proposed).

#### D. Results and Comparisons

Table I shows details of our experimental results and comparisons with our different baseline models. We can easily identify that our model *AcRoNN* outperforms all of the baseline models significantly in both of our collected datasets and already available datasets. Also, we can firmly say that, our *AcRoNN* model outperforms baseline significantly even though we chose to use baseline proposed sensors (Actigraphy and ECG) related features only.

TABLE I

*AcRoNN* PERFORMANCE COMPARISONS ACROSS BASELINE METHODS

Data D1	B1 (D1)	B2 (D1)	B3 (D1)	AcRoNN (D1)
Precision	69.65 ± .1	70.45 ± .1	79.58 ± .1	83.87 ± .2
Recall	68.64 ± .1	71.76 ± .1	80.4 ± .2	82.45 ± .1
F1	69.23 ± .1	74.56 ± .2	79.45 ± .1	83.45 ± .1
Data D2	B1 (D2)	B2 (D2)	B3 (D2)	AcRoNN (D2)
Precision	65.36 ± .4	66.45 ± .1	72.56 ± .1	76.76 ± .2
Recall	66.42 ± .1	66.45 ± .1	73.45 ± .1	76.34pm.2
F1	67.35 ± .1	67.34 ± .1	74.47 ± .1	77.47 ± .2

#### IV. CONCLUSION

To develop an automated cognitive fatigue assessment system, we introduced a new pipeline from data collection, data preprocessing, feature engineering, attention based LSTM and a novel context-aware LSTM model flow. To our best knowledge, *AcRoNN* is the best cognitive fatigue detection model in the existing literature which can be extended to any other physiological health assessment with proper study design and data collection. Our efficient two-step feature map scoring method provides a new concept in context-aware activity and health monitoring research area that can be utilized to provide appropriate care to patients with dementia, asthma, post-traumatic stress disorder and so on.

#### REFERENCES

- [1] Burn-out an "occupational phenomenon": International Classification of Diseases, 28 May 2019 Departmental news, <https://www.who.int/news/item/28-05-2019-burn-out-an-occupational-phenomenon-international-classification-of-diseases>
- [2] Walker.et al. Circadian rhythm disruption and mental health. *Transl Psychiatry* 10, 28 (2020).
- [3] Yang Bai et. al. Fatigue assessment using ECG and actigraphy sensors. ISWC 2020
- [4] Empatica E4 Watch: <https://www.empatica.com/research/e4/>
- [5] Hoddes E. (1972). "The development and use of the stanford sleepiness scale (SSS)". *Psychophysiology*. 9 (150).
- [6] Brunet JF et. al. Validation of sleep-2-Peak: A smartphone application that can detect fatigue-related changes in reaction times during sleep deprivation. *Behav Res Methods*. 2017 Aug
- [7] <https://sleep-2-peak.com>
- [8] "Management of Excessive Daytime Sleepiness Reviewed". *Medscape*. Retrieved 2021-04-07
- [9] Brunet, JF. et al. Validation of sleep-2-Peak: A smartphone application that can detect fatigue-related changes in reaction times during sleep deprivation. *Behav Res* 2017.
- [10] Luo H., et. al. (2020), Assessment of Fatigue Using Wearable Sensors: A Pilot Study. *Digit Biomark*.
- [11] <https://www.biovotion.com/everion/>
- [12] Michielsen HJ et. al. Psychometric qualities of a brief self-rated fatigue measure: The Fatigue Assessment Scale. *J Psychosom Res*. 2003 Apr.
- [13] Lee KA., et. al. Validity and reliability of a scale to assess fatigue. *Psychiatry Res*. 1991 Mar;36(3):291–8.
- [14] Cao W et. al. Brits: Bidirectional recurrent imputation for time series. In *Advances in Neural Information Processing Systems*. 2018.
- [15] Nils Y. Hammerla et. al. PD Disease State Assessment in Naturalistic Environments Using Deep Learning. *AAAI*, 2015.
- [16] Bethany Little, et. al. Deep learning-based automated speech detection as a marker of social functioning in late-life depression. *Psychological Medicine* (01 2020), 1–10.
- [17] Zahra Sedighi Maman et. al. A data-driven approach to modeling physical fatigue in the workplace using wearable sensors. *Applied Ergonomics* 65 (03 2017).
- [18] Alaa Althubaiti, *Information bias in health research: Definition, pitfalls, and adjustment methods*. *Multidisciplinary Healthcare* 2016.
- [19] Vourvopoulos, A et. al.. EEG correlates of video game experience and user profile in motor-imagery-based brain-computer interaction. *Vis Comput* 33, 533–546 (2017).
- [20] McMorris T et. al. Cognitive fatigue effects on physical performance: A systematic review and meta-analysis. *Physiol Behav*. 2018, May.
- [21] Mohammad Arif Ul Alam, Nirmalya Roy, Sarah Holmes, Aryya Gangopadhyay, Elizabeth Galik, Automated Functional and Behavioral Health Assessment of Older Adults with Dementia, CHASE 2016.
- [22] Mohammad Arif Ul Alam, Sarah Holmes, Nirmalya Roy, A. Gangopadhyay, E. Galik, AutoCogniSys: Iot-Assisted Cognitive Health Assessment of Older Adults, arXiv:2003.07492, 17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous 2020).
- [23] Ming Zeng et. al.. 2018. Understanding and improving recurrent networks for human activity recognition by continuous attention. ISWC 2018.
- [24] A. Greco,et. al., Combining Electrodermal Activity and Speech Analysis towards a more Accurate Emotion Recognition System, EMBC, 2019
- [25] M. Alam, N. Roy, S. Holmes, A. Gangopadhyay, E. Galik, Automated Functional and Behavioral Health Assessment of Older Adults with Dementia, CHASE, 2016.
- [26] C. Setz et. al. Discriminating stress from cognitive load using a wearable EDA device., *IEEE transactions on information technology in biomedicine*, vol. 14, no. 2, Mar. 2010.
- [27] C. Lee et. al.. Reduction of motion artifacts from photoplethysmographic recordings using a wavelet denoising approach. *IEEE EMBS Asian-Pacific Conference on Biomedical Engineering*, 2003
- [28] W. Chen et. al. Wavelet-based motion artifact removal for electrodermal activity. *EMBC* 2015
- [29] Weblink.: [n.d.]. ActiLife. <https://www.actigraphcorp.com/actilife/>
- [30] K. Bardool et. al. A Systematic Analysis of a Context Aware Deep Learning Architecture for Object Detection. *BNAIC* 2019
- [31] Mohammad Arif Ul Alam, AI-Fairness Towards Activity Recognition of Older Adults, *MobiQuitous* 2020