Time-Series Forecasting of Physical and Mental Fatigue Based on Wearable Data

JUN NI (JENNY) DU, University of Toronto, Canada

1 INTRODUCTION

Fatigue is a broad, multifactorial concept that is generally defined by the feelings of reduced physical and mental energy levels. It can strongly impact a patient's health-related quality of life (HRQOL), and has been a primary focus in various treatments [1]. The cause of fatigue can be categorized into pathological and non-pathological types. Pathological fatigue can be attributed to medical conditions like anemia, heart disease, cancer, and other chronic illnesses [3], while non-pathological fatigue can occur in healthy individuals under a variety of circumstances and environmental factors, making it harder to track. Fatigue can also be categorized as physical or mental. Physical fatigue typically occurs in response to extensive muscle activity, such as exercise, and often recovers quickly with rest. In contrast, mental fatigue tends to accumulate over time due to prolonged exposure to stress, anxiety, or excessive stimulation [7].

Existing approaches for studying fatigue often rely solely on patient-reported outcomes and single performance tests performed in a laboratory setting, such as tracking eye movements in a car-driving simulation to assess driver fatigue[8] or utilizing clinical technologies like mechanomyography, electromyography, and ultrasounds to detect muscle fatigue at rest and after specific exercise activities [2]. These approaches exhibit limitations in capturing the complete spectrum of factors that contribute to fatigue, including physiological, psychological, and contextual elements. Furthermore, they face challenges such as data sparsity (caused by the limitation of collecting data only in laboratory settings), limited reproducibility, and sub-optimal compliance (in attending lab sessions). An emerging approach to improving fatigue assessment is the use of wearable sensors to collect continuous and objective data over extended periods in real-life environments. These sensor data, complemented by patient-reported outcomes, are driving a new trend in mobile health research [5].

In a pilot study conducted by Luo et al. [2020], the authors explored the relationship between self-reported non-pathological physical and mental fatigue and multisensor wearable data in healthy subject as they went about their daily routines. They collected sensor data from 28 healthy adult participants, using an armband-sized wearable device that was continuously worn on each participant's non-dominant arm for one or more week-long periods, along with a daily questionnaire on physical and mental fatigue levels. The authors found that, for both physical and mental fatigue, combining multiple sensor parameters resulted in a stronger correlation with the fatigue labels compared to using individual sensor data. Moreover, clustering analysis identified distinct top predictive features for physical fatigue compared to mental fatigue, implying differences in the underlying mechanisms of these two types of fatigue.

Luo et al.'s study [2020] was among the first to analyze multimodal data sources related to physical activity, vital signs, and other physiological parameters in the context of their relationship to self-reported non-pathological physical and mental fatigue in real-world settings. This study offers a glimpse into the potential of a machine learning-driven framework connecting multisensor wearable data with patient-reported outcomes, potentially enhancing the understanding of HRQOL factors in clinical trials and daily medical practice. However, the flip-side of its novelty lies in the limitations of the model's performance, which also serves as the underlying motivation for this project. The best models in the pilot study achieved only 70% weighted accuracy for physical fatigue and 65% weighted accuracy for mental fatigue. The authors also simplified each of the fatigue measures into binary values (fatigue/no fatigue), despite the original questionnaire response containing fatigue scores across five levels (never; sometimes; regularly; often; always) and a numerical score ranging from 1 to 10. Furthermore, each participant's daily sensor data and questionnaire response are treated as independent samples, disregarding the time-series nature of the data which could reveal any potential long-term patterns. These long-term patterns could prove especially critical in the prediction of mental fatigue, since accumulated stress or over-stimulation may be reflected in the chronic changes that are not observable within the data from just one day. Therefore, we propose a project to enhance the models presented by Luo et al. [2020]. Specifically, we aim to develop a time-series forecasting algorithm to predict physical and mental fatigue levels using continuous data from wearable sensors collected over multiple days. In addition, we wish to explore alternative approaches for improving the model's accuracy, while eliminating some of the simplifications that were applied to the original models.

2 RESEARCH QUESTION

Building upon the pilot study, the core research question of this project is as follows: Can the predictive performance of fatigue prediction models be significantly improved by integrating a time-series forecasting algorithm that utilizes historical data from each participant, such as all sensor data collected over the past 7 days?

In addition to the primary research question, our project aims to investigate and potentially enhance several other aspects of the models. Firstly, the pilot study employed simplified, binarized fatigue labels to reduce intra- and inter-rater variability. We intend to examine whether representing these labels in their original format, as presented in the questionnaire, through the use of ordinal encoding can result in improved performance. Secondly, conditional on project timeline constraints, we plan to explore more advanced machine learning models to evaluate the predictive capabilities of sensor data in fatigue prediction.

3 DATASET DESCRIPTION

The dataset for this project has been published by the authors [6] and is accessible at the following link: https://zenodo.org/record/4266157. It includes data from 28 healthy adult participants aged 26 to 55 years (average age: 42 years; 11/16/1 female/male/unknown gender), with a total of 973 recording days. The recorded data consists of continuous sensor data on physical activity, vital signs, and other physiological parameters at 1-minute intervals, as well as responses to the daily questionnaires on fatigue. The sensor data are collected using an armband-sized clinical-grade multisensor wearable, Everion (Biovotion AG, Switzerland), worn continuously on each participant's non-dominant arm. The Everion device combines a 3-axis accelerometer, barometer, galvanic skin response electrode, and temperature and photo sensors. The daily fatigue questionnaire is delivered using a mobile app, SymTrack (Gastric GmbH, Switzerland). A complete list of parameters recorded by the wearable device and the questionnaire questions is presented in Table 1.

This dataset is highly suitable for our project, given that the motivation of our research problem is built upon the pilot study [6] where this data is collected. The sensor data is gathered from a clinical-grade medical device that is both discreet and highly portable, making it a prime example of mobile health technology. Moreover, several of the sensors used to produce this dataset, such as the accelerometer, barometer, and skin temperature sensor, are already present in current consumer-facing smartwatches and fitness trackers, albeit they might have a somewhat reduced degree of accuracy [4]. We can anticipate that advancements in technology will likely make it feasible to replicate this study using only consumer-grade wearables in the near future.

Nevertheless, several limitations persist within the dataset. These include the presence of missing data, which can be particularly troublesome for time-series forecasting that relies on complete data, as well as the inherent intra- and inter-rater variability in the self-reported fatigue levels. Additionally, the dataset's relatively small size, consisting of only 28 adult participants who are all healthy, raises concerns about the generalizability of the findings. While Luo et al. [2020] addressed some of the shortcomings in their pilot study, such as using a recurrent neural network-based algorithm to impute missing data and binarizing the fatigue levels to reduce variability, it is essential to acknowledge that these potential limitations might have an inherent impact on the results.

Table 1. List of parameters measured by the wearable device (Everion) and the daily questionnaire (SymTrack)

Sensor	Parameter Name	Description
Accelerometer	ActivityClass	Type of physical activity: 0 = undefined, 1 = resting, 9 = other, 10 = biking, 11 = running, 12 = walking
	ActivityCounts	Intensity of motion (counts per minute)
	Steps	Number of steps
	EnergyExpenditure	Amount of energy a person uses to complete all regular bodily functions (calories/min)
Photoplethysmography	HR	Heart rate (beats per minute)
	HRV	Heart rate variability; indicates the beat to beat variations (milliseconds)
	RESP	Respiration rate (breaths per minute)
	BloodPerfusion	Percentage change in blood volume in local tissue resulting from a heartbeat
	BloodPulseWave	Amplitude of pulse wave generated during the contraction of the heart
Temperature	SkinTemperature	Skin temperature ($^{\circ}C$)
Galvanic Skin Response	GalvanicSkinResponse	Changes in the electrical conductivity of the skin in response to sweat gland activity; a measure of emotional arousal (kOhm)
Barometer	Barometer	Atmospheric pressure and altitude (mbar)

Measure	Questionnaire Question	Possible Answers
Physical fatigue score (PhF)	"Physically, today how often did you feel exhausted?"	never; sometimes; regularly; often; always
Mental fatigue score (MF)	"Mentally, today how often did you feel exhausted?"	(same as above)
Visual analogue scale score (VAS)	"Describe fatigue on a scale of 1–10, where 1 means you don't feel tired at all and 10 means the worst tiredness you can imagine."	a number from 1 to 10
Indicator of relative perception (RelP)	"Are you feeling better, worse, or the same as yesterday?"	better; worse; same
Sports	"Did you do sport today?"	yes; no

REFERENCES

[1] Lauren S. Aaronson, Cynthia S. Teel, Virginia Cassmeyer, Geri B. Neuberger, Leonie Pallikkathayil, Janet Pierce, Allan N. Press, Phoebe D. Williams, and Anita Wingate. 1999. Defining and Measuring Fatigue. *Image: the Journal of Nursing Scholarship* 31, 1 (March 1999), 45–50. https://doi.org/10.1111/j.1547-5069.1999.tb00420.x

[2] Mohamed R. Al-Mulla, Francisco Sepulveda, and Martin Colley. 2011. A Review of Non-Invasive Techniques to Detect and Predict Localised Muscle Fatigue. *Sensors* 11, 4 (March 2011), 3545–3594. https://doi.org/10.3390/s110403545

[3] Josef Finsterer and Sinda Zarrouk Mahjoub. 2013. Fatigue in Healthy and Diseased Individuals. *American Journal of Hospice and Palliative Medicine®* 31, 5 (July 2013), 562–575. https://doi.org/10.1177/1049909113494748

[4] Marjolein E Haveman, Mathilde C van Rossum, Roswita M E Vaseur, Claire van der Riet, Richte C L Schuurmann, Hermie J Hermens, Jean-Paul P M de Vries, and Monique Tabak. 2022. Continuous Monitoring of Vital Signs With Wearable Sensors During Daily Life Activities: Validation Study. *JMIR Formative Research* 6, 1 (Jan. 2022), e30863. https://doi.org/10.2196/30863

[5] Xiao Li, Jessilyn Dunn, Denis Salins, Gao Zhou, Wenyu Zhou, Sophia Miryam Schüssler-Fiorenza Rose, Dalia Perelman, Elizabeth Colbert, Ryan Runge, Shannon Rego, Ria Sonecha, Somalee Datta, Tracey McLaughlin, and Michael P. Snyder. 2017. Digital Health: Tracking Physiomes and Activity Using Wearable Biosensors Reveals Useful Health-Related Information. *PLOS Biology* 15, 1 (Jan. 2017), e2001402. https://doi.org/10.1371/journal.pbio.2001402

[6] Hongyu Luo, Pierre-Alexandre Lee, Ieuan Clay, Martin Jaggi, and Valeria De Luca. 2020. Assessment of Fatigue Using Wearable Sensors: A Pilot Study. *Digital Biomarkers* 4, Suppl. 1 (Nov. 2020), 59–72. https://doi.org/10.1159/000512166

[7] Kei Mizuno, Masaaki Tanaka, Kouzi Yamaguti, Osami Kajimoto, Hirohiko Kuratsune, and Yasuyoshi Watanabe. 2011. Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity. *Behavioral and Brain Functions* 7, 1 (2011), 17. https://doi.org/10.1186/1744-9081-7-17

[8] Junli Xu, Jianliang Min, and Jianfeng Hu. 2018. Real-time eye tracking for the assessment of driver fatigue. *Healthcare Technology Letters* 5, 2 (Jan. 2018), 54–58. https://doi.org/10.1049/htl.2017.0020