

Title	Estimating perceived fatigue using machine learning and biomechanical features from wearable sensors
Authors	Qirtas, Malik Muhammad;Yasar, Merve Nur;Sica, Marco;Tedesco, Salvatore;Visentin, Andrea
Publication date	2025-07-03
Original Citation	Qirtas, M. M. , Yasar, M. N., Sica, M., Tedesco, S. and Visentin, A. (2025) 'Estimating perceived fatigue using machine learning and biomechanical features from wearable sensors', 2025 IEEE International Conference on Smart Computing (SMARTCOMP), Cork, Ireland, 16-19 June 2025, pp. 342-347. https://doi.org/10.1109/SMARTCOMP65954.2025.00083
Type of publication	Conference item;proceedings-article
Link to publisher's version	10.1109/smartcomp65954.2025.00083
Rights	© 2025, IEEE. For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. - https://creativecommons.org/licenses/by/4.0/
Download date	2026-10-30 15:00:26
Item downloaded from	https://hdl.handle.net/10468/18195

Estimating Perceived Fatigue Using Machine Learning and Biomechanical Features from Wearable Sensors *

Malik Muhammad Qirtas*, Merve Nur Yasar†, Marco Sica‡, Salvatore Tedesco‡, Andrea Visentin*

* SFI Insight Centre for Data Analytics, School of Computer Science and IT, University College Cork, Cork, Ireland
mqirtas@ucc.ie, andrea.visentin@ucc.ie

† Department of Coaching Education, Faculty of Sports Sciences, Sakarya University of Applied Sciences, Sakarya, Turkey
merv.nur.yasar@gmail.com

‡ Tyndall National Institute, University College Cork, Cork, Ireland
marco.sica@tyndall.ie, salvatore.tedesco@tyndall.ie

Abstract—Physical fatigue is a state of reduced physical ability caused by prolonged activity or repetitive tasks. It can affect performance in tasks that require effort, focus or precision and can lead to reduced strength and increased risk of injuries. Detecting physical fatigue is important for timely interventions and enhancing safety, efficiency and overall well-being in workplaces, sports and rehabilitation. In this study, we propose a robust and generalizable framework for fatigue detection using wearable sensor data, specifically using Inertial Measurement Unit and Electromyography sensors. A comprehensive set of biomechanical features was extracted from raw sensor data to capture both kinematic and neuromuscular aspects of fatigue progression. These features were evaluated across shoulder internal rotation and external rotation movements under different resistance levels. We trained and compared multiple regression models for fatigue estimation using subjective fatigue ratings based on the Borg Rating of Perceived Exertion scale and performed feature importance analysis to get model interpretability. The extracted feature set showed strong generalizability specifically for IR movements, as proved by leave one task out cross-validation, where models maintained robust performance across unseen movement-resistance task settings. This work highlights the potential of combining IMU and EMG data, along with biomechanical features extracted from these two sensor modalities for accurate and interpretable fatigue estimation. It opens the way for real-world applications in dynamic and diverse environments for effective fatigue estimation.

Index Terms—Wearable Sensors, Fatigue Estimation, Biomechanical Features, Inertial Measurement Unit, Electromyography, Machine Learning

I. INTRODUCTION

Physical fatigue is a complex physiological and psychological state that can significantly impact performance, safety, recovery and overall health in various domains like athletic training, sports, rehabilitation and workplace ergonomics [1]. It arises from repetitive or sustained movement patterns which cause biomechanical and neuromuscular changes. These changes lead to reduced muscle activation efficiency, altered movement patterns and increased joint instability. All of these factors can increase the risk of injuries and compromise task

performance. To reduce these effects, there is a critical need for accurate and real-time fatigue estimation systems that can monitor neuromuscular and biomechanical changes in muscles. Specifically, the shoulder joint plays an important role in a wide range of upper body movements which makes it essential for tasks requiring precision, strength and endurance in sports, rehabilitation and daily activities. For example, in overhead sports, 49% of athletes experience shoulder problems at some point in their career [2]. The high prevalence and incidence of shoulder pain in these athletes is thought to be largely due to the repetitive and excessive stress placed on the shoulder during overhead sports. Such repetitive movements are thought to create muscle fatigue thus increase the risk of shoulder injury by changing kinematics. Detecting arm and shoulder fatigue is particularly important as it helps prevent injuries and supports long term joint health in physically demanding tasks and athletic performance.

Fatigue assessment has traditionally relied on subjective self-reports, such as the Borg Rating of Perceived Exertion (RPE) scale [3]. However, due to their subjective nature, self-reported ratings alone may not fully capture the biomechanical and neuromuscular changes associated with fatigue, especially in real-time applications, and they can vary significantly across individuals. To address these limitations, recent approaches have explored combining subjective measures with objective sensor data for more robust fatigue estimation [4]. Wearable sensors such as Electromyography (EMG) and Inertial Measurement Units (IMUs) enable continuous monitoring of muscle activity and movement kinematics. EMG captures electrical activity in skeletal muscles, useful for analysing activation patterns and fatigue-related changes [5], while IMUs provide kinematic data such as acceleration and angular velocity to assess movement behaviour and efficiency during repetitive tasks [6].

Previous studies have shown that IMU and EMG are effective tools for monitoring movement and muscle activity for fatigue detection in domains such as sports, rehabilitation and industrial ergonomics [7], [8]. While fatigue estimation

has been well explored for lower limb activities like running and cycling [9], [10], there are not many studies related to shoulder and upper limb movements, despite their importance in overhead tasks and upper body sports. Some studies have investigated shoulder fatigue using IMU or EMG during specific actions such as throwing or walking, and a few have combined both sensor types to classify fatigue states during upper limb rehabilitation [4], [11], [12]. However, most of this work has focused on classification, which is less suitable for modeling fatigue as a continuous progressive process. Moreover, many studies rely on single-task setups which limits their generalizability across real world conditions.

This paper addresses these gaps by using regression-based modeling with interpretable biomechanical features derived from IMU and EMG data to evaluate across multiple movement-resistance combinations. We integrate subjective Borg ratings with objective features extracted from sensor data to develop a generalizable and robust framework for capturing fatigue progression across multiple movement-resistance combinations. By using machine learning, we combine data across multiple participants to generalise individual perceived fatigue to reduce their inherent variability and enhance the robustness of the fatigue estimation. Below are the main contributions:

- 1) Extracted a comprehensive set of biomechanical features from raw IMU and EMG data that can capture both kinematics and neuromuscular aspects of fatigue progression.
- 2) Trained and evaluated multiple regression models on shoulder internal rotation (IR) and external rotation (ER) movements across different resistance levels and their combinations to analyze how fatigue dynamics vary across task settings.
- 3) Conducted feature importance analysis to identify a set of top performing biomechanical features that serve as key indicators of fatigue across different task settings while enhancing model interpretability.
- 4) By using multi-task datasets, we showed that the extracted features are not overfitted to any specific setting. Instead, they prove strong generalizability across different tasks. This robust testing shows the robustness and adaptability of our feature set for real-world applications.

II. METHODS

A. Dataset

This study used a pre-collected dataset collected from 34 healthy participants at the Tyndall National Institute’s Wearable Laboratory, University College Cork, Ireland, between April and July 2023 [13]. The participants included 23 males and 11 females. All of these participants were physically active, having physical training at least three times per week and had no prior history of musculoskeletal injuries.

IMU and EMG sensors were applied to the dominant side of upper body to assess muscle activation and movement. EMG electrodes were placed on six muscles using placement guidelines from Cram’s Introduction to Surface Electromyography, and signals were recorded at 1000 Hz using a BTS

FREEEMG 300 system. IMU sensors (MTw Awinda, Xsens) were also attached to 6 body parts to capture acceleration and angular velocity at 100 Hz [Xsens system]. Maximal Voluntary Isometric Contraction (MVIC) was assessed using a push/pull dynamometer and resistance levels were set to 30–40%, 40–50%, and 50–60% MVIC. Participants performed shoulder internal and external rotations under these conditions in pseudo-randomised order, with 10-minute rest intervals, while maintaining consistent posture with 90° elbow flexion and straight wrists.

Fatigue levels were measured using the Borg RPE scale [3]. This scale ranges from 6, meaning “no exertion at all” to 20, meaning “maximum exertion”. Participants reported their perceived exertion before starting the exercises and every 10 seconds during the exercises. The exercises continued until participants reached maximal exertion, corresponding to an RPE level of 20. A metronome set at 40 beats per minute was used to ensure a consistent exercise pace.

B. Data Cleaning and Preprocessing

We applied filtering, segmentation and interpolation of subjective Borg RPE ratings in preprocessing. The dataset was divided according to the six exercises performed, specifically 30-40%, 40-50% and 50-60% of IR and ER, respectively. Data were processed separately for each movement-resistance task setting before being used for model training. The dataset and code are publicly available: the dataset can be accessed at ¹ and the code is available at ².

For IMU sensor signals, we first removed the mean value from each axis (X, Y, Z) of acceleration and gyroscope signal to eliminate DC offset. Then based on existing literature, a 4th-order low-pass Butterworth filter with a cutoff frequency of 20 Hz was applied to reduce high-frequency noise [14]. Fig. 1 shows the raw and filtered signals.

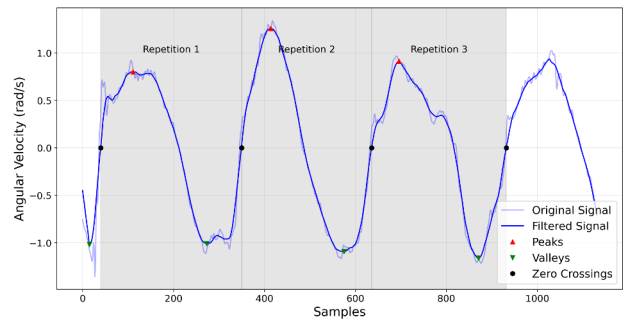


Fig. 1: Repetition-based segmentation for internal rotation with original and filtered signals shown.

For EMG signals, the mean was subtracted from the signal to eliminate DC offset. We then applied a 4th-order Butterworth band-pass filter with a low cut-off at 30 Hz and a high cut-off at 350 Hz to remove motion artifacts and minimize ECG contamination, following established methods [15].

¹<https://zenodo.org/record/8415066>

²<https://github.com/Qirtas/SensorFatigueEstimation>

To analyze individual repetitions of the exercises, the IMU and EMG data were segmented into repetitions. For IR, both peaks and valleys were identified, and a repetition was considered as the signal between two successive zero-crossing points located between a valley and a peak. Conversely, for ER, a repetition was considered as the signal between two consecutive zero-crossing points located between a peak and a valley. These segmentation points are illustrated in Fig. 1, where peaks, valleys and relevant zero crossings are marked on the filtered signal. Since the subjective Borg RPE scale ratings were collected at 10-second intervals, not every repetition had a directly recorded rating. To address this issue, missing values were estimated using linear interpolation between the recorded Borg RPE values.

C. Features Extraction

A key contribution of this study is the extraction of a diverse set of biomechanical features from raw IMU and EMG sensor data. We achieved this by identifying studies in sports science and technical domains that used IMU or EMG sensor data for various tasks such as movement analysis, fatigue estimation and activity recognition. These studies provided useful information about commonly used features for analyzing sensor data. Based on this review, we compiled a list of interpretable features and extracted a comprehensive set for our study. This set covered a wide range of biomechanical characteristics for both IMU and EMG data to ensure a robust representation of fatigue related changes. This features extraction step transforms raw sensor data into meaningful biomechanical indicators for fatigue estimation. These features can capture multiple aspects of movement dynamics and muscle activity and can provide a granular representation of fatigue related changes during repetitive movements. We computed features for each segmented repetition to ensure temporal alignment with the corresponding Borg RPE scale ratings.

For IMU data, we extracted features from both accelerometer and gyroscope signals along each axis (X, Y, Z) separately. To capture movement behaviour comprehensively, we extracted features from both the time and frequency domains. Time domain features focused on movement patterns and intensity. These included slope to measure the rate of change in movement, zero-crossing rate to identify movement transitions and the duration of extreme values to capture high intensity movements. Jerk, which is the time derivative of acceleration, was also computed to assess movement smoothness. Other movement related features such as peak velocities, time to peak velocity, cumulative displacement and range of motion were also extracted to capture movement efficiency. Angular measurements, including maximum and minimum angles, angular displacement ranges and root mean square of angular velocity, were also extracted.

Frequency domain features were extracted using the Fast Fourier Transform and Power Spectral Density. These features included dominant frequency and peak frequency components, energy distribution across frequency bands and a smoothness index, which was calculated as the ratio of low-frequency to

high-frequency power. To provide a more detailed analysis of the signals, both time and frequency features were further analyzed using statistical sub-features, such as mean, median, minimum, maximum, standard deviation, variance, skewness and kurtosis to describe the distribution and variability of the signals.

For EMG data, we extracted features related to muscle activation patterns and fatigue progression. Time domain features included Integrated EMG, which calculates total muscle activity, Root Mean Square and Mean Absolute Value to assess signal intensity and variance to measure signal consistency. Signal complexity was analyzed using waveform length. Since fatigue affects the frequency content of EMG signals, frequency domain features were also extracted to track changes in muscle activation over time. These features included Short-Time Fourier transform to analyze frequency variations, median and mean frequency calculations, Power Spectral Density measurements and frequency band ratios to detect shifts in muscle activity patterns as fatigue progressed. In addition to time and frequency domain features, we also extracted hybrid features that combined aspects of both. Zero-crossing rates were computed to capture both frequency changes and muscle activation patterns. Slope indicators and peak characteristics were extracted to track how muscle fatigue evolved over time. Finally, we extracted statistical sub-features, similar to those used for IMU data, to provide a more detailed representation of neuromuscular changes associated with fatigue.

A total of over 3,780 features were extracted by combining IMU and EMG data along with their statistical sub-features. This ensured a detailed understanding of biomechanical and neuromuscular indicators.

D. Models Training

The model training process was designed to systematically evaluate fatigue estimation across three key scenarios: (1) individual tasks, (2) combined task settings and (3) leave one task out. These scenarios were selected to rigorously assess model performance, generalizability and robustness under diverse scenarios.

For the first evaluation, models were trained and tested on data from individual movement-resistance task setting, which included 30-40%, 40-50% and 50-60% for both IR and ER. This approach isolates the effect of specific resistance levels and movement types on fatigue estimation. By training separate models for each task, we assessed whether models could learn task specific fatigue patterns without interference from other scenarios. These models also helped identify task settings where fatigue dynamics were easier or harder to predict.

To evaluate generalizability across tasks, models were trained on data combined from different movement-resistance combinations. These combinations involved grouping IR and ER settings with the same resistance levels (e.g., IR 30–40%, combined with ER 30–40%, IR 40–50% combined with ER 40–50%). In addition, models were trained on all IR data combined, all ER data combined and finally on the full dataset,

which included every movement and resistance level together. Combining data exposed models to diverse biomechanical variations just like in real-world scenarios where fatigue may occur across different tasks. These experiments also tested whether models could learn generic fatigue indicators that generalize beyond individual task settings.

To further evaluate generalizability and robustness, we used a leave one task out cross validation approach. For each target task setting (e.g., IR 30-40%), models were trained on all other task settings and tested only on the left out task setting. This approach simulated real-world situations where models encounter completely unseen movement-resistance combinations.

Given the large feature set, we used a two step feature selection process. First, we trained a RF model on the full feature set and selected the top 100 features with the highest importance scores for each task setting. Selecting the top 100 features provided a consistent method to evaluate the importance and interpretability of the overall feature. Then, using this reduced feature set, we trained three regression models including XGBoost, Support Vector Regression (SVR) and RF for each task setting. To ensure generalizability across individuals, we applied leave one group out cross-validation. In this method, each fold excluded all data from one participant, resulting in 34 folds. The hyperparameters for the models were tuned using RandomizedSearchCV. We normalized all features using StandardScaler and evaluated model performance using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

E. Features Importance Analysis

To identify the most important features for fatigue estimation, we performed a feature importance analysis. This analysis was performed separately for different task settings, including individual resistance levels such as 30–40% and 40–50% for both IR and ER movements. Additionally, we evaluated feature importance across combined task settings, including 30–40% and 40–50% together for IR and ER movements, as well as all movement-resistance levels combined. To evaluate the overall impact of each feature, we summed its raw importance scores across all task settings. This method captured the total contribution of each feature to fatigue estimation, regardless of the specific resistance task setting. It provided a clear and holistic view of feature relevance. This summation approach also reduced task specific fluctuations, which made the assessment more stable and reliable. Finally, we normalized the raw scores to have a direct comparison across different task settings.

III. RESULTS

A. Single Task Performance

Table I presents the performance of the models across different task settings. To establish a reference for model performance, we trained a simple baseline model for predicting the mean fatigue score. This baseline approach, which does

TABLE I: Performance Comparison of Models for Single Task Setting

Task Setting	Model	MAE	MSE	RMSE	MAPE (%)
IR 30-40%	XGBoost	2.1425	7.2359	2.6900	16.44
	SVR	2.2370	7.8724	2.8058	17.56
	RF	2.1654	7.3139	2.7044	16.49
ER 30-40%	XGBoost	1.6146	4.0337	2.0084	13.10
	SVR	1.6465	4.3009	2.0739	13.52
	RF	1.8498	5.0758	2.2530	15.29
IR 40-50%	XGBoost	2.0298	6.1874	2.4875	15.93
	SVR	1.9950	6.0588	2.4615	15.59
	RF	2.1672	6.8947	2.6258	17.19
ER 40-50%	XGBoost	1.3211	2.7399	1.6553	11.54
	SVR	1.4243	3.0671	1.7513	12.14
	RF	1.4457	3.1729	1.7813	12.80
IR 50-60%	XGBoost	1.8687	5.3924	2.3221	15.32
	SVR	1.9270	5.5617	2.3583	15.96
	RF	1.9754	5.6860	2.3845	16.33
ER 50-60%	XGBoost	1.2627	2.4838	1.5760	11.44
	SVR	1.3152	2.7341	1.6535	12.16
	RF	1.3321	2.7336	1.6534	12.13

IR: Internal, ER: External

MAE: Mean Absolute Error, MSE: Mean Squared Error, RMSE: Root Mean Squared Error, MAPE: Mean Absolute Percentage Error

RF: Random Forest, SVR: Support Vector Regressor

not account for sensor data or individual variations, resulted in an MAE of 3.2. This serves as a reference point to evaluate model performance, where any machine learning approach should ideally show a lower error to be considered effective. Our machine learning pipeline significantly outperformed this baseline, reducing MAE by up to 60% in some tasks.

The best performance was observed in ER 50-60% (XGBoost: MAE = 1.26, MAPE = 11.44%) and ER 40-50% (XGBoost: MAE = 1.32, MAPE = 11.54%). In contrast, the worst performance was seen in IR 30-40% (MAE range: 2.14–2.23).

XGBoost consistently outperformed SVR and RF across most tasks. The largest performance gap was observed in the ER, such as in the ER 30-40%, where XGBoost achieved an MAE of 1.61 compared to RF's MAE of 1.84. ER movements also showed better predictability than IR.

B. Combined Task Settings Performance

Table II presents the performance of the models trained on combined movement-resistance task settings. XGBoost achieved the lowest errors in all cases, with MAE values ranging from 1.99 to 2.09. The ER Only task setting resulted in the lowest MAE of 1.99 and an MAPE of 16.21%. The IR Only task setting showed slightly higher errors, with an MAE of 2.08 and MAPE of 15.92%. Models trained on a combination of IR and ER movements exhibited MAE values between 2.06 and 2.16, comparable to those trained on individual rotation tasks.

TABLE II: Performance Comparison of Models for Combined Task Settings

Task Setting	Model	MAE	MSE	RMSE	MAPE (%)
IR Only	XGBoost	2.0893	6.6937	2.5872	15.92
	SVR	2.1816	7.3650	2.7139	16.49
	RF	2.1035	6.7181	2.5919	15.99
ER Only	XGBoost	1.9902	5.8449	2.4176	16.21
	SVR	2.0058	5.9747	2.4443	16.13
	RF	2.0427	5.9775	2.4449	16.72
IR + ER	XGBoost	2.0615	6.6012	2.5693	16.17
	SVR	2.1051	6.8680	2.6207	16.50
	RF	2.1637	7.0088	2.6474	17.04

ER Only: ER 30-40% + ER 40-50% + ER 50-60%

IR: Internal Rotation, ER: External Rotation

IR Only: IR 30-40% + IR 40-50% + IR 50-60%

IR + ER: IR Only + ER Only

MAE: Mean Absolute Error, MSE: Mean Squared Error, RMSE: Root Mean Squared Error, MAPE: Mean Absolute Percentage Error

RF: Random Forest, SVR: Support Vector Regressor

TABLE III: Performance Comparison of Models for Leave One Task Out Cross Validation to Evaluate Model Generalizability Across Unseen Task Settings

Task Setting	Model	MAE	MSE	RMSE	MAPE (%)
IR 30-40%	XGBoost	2.1403	6.8562	2.6184	16.79
	SVR	2.0370	6.3582	2.5215	16.05
	RF	2.2154	7.2348	2.6898	17.54
ER 30-40%	XGBoost	2.0430	6.2164	2.4933	15.75
	SVR	1.9401	5.6016	2.3668	14.41
	RF	2.1253	6.5252	2.5544	16.28
IR 40-50%	XGBoost	1.6807	4.3028	2.0743	12.36
	SVR	1.7696	4.8248	2.1966	12.97
	RF	1.9048	5.0274	2.2422	14.16
ER 40-50%	XGBoost	1.9897	4.7745	2.2843	15.19
	SVR	2.2643	5.8847	2.7478	16.92
	RF	2.3018	7.5228	2.8194	17.15
IR 50-60%	XGBoost	1.8830	5.3922	2.3221	13.90
	SVR	1.9191	5.6913	2.3856	14.10
	RF	1.9255	5.5401	2.3537	14.17
ER 50-60%	XGBoost	2.5970	10.0869	3.1703	17.51
	SVR	2.5473	10.1630	3.1874	17.83
	RF	2.6853	10.8042	3.2879	17.94

IR: Internal Rotation, ER: External Rotation

MAE: Mean Absolute Error, MSE: Mean Squared Error, RMSE: Root Mean Squared Error, MAPE: Mean Absolute Percentage Error

RF: Random Forest, SVR: Support Vector Regressor

C. Generalizability Using Leave One Task Out Cross Validation

Table III summarizes the results of the leave one task out cross validation to assess model generalizability to unseen task settings. XGBoost achieved the lowest errors for most left out tasks, with MAE values ranging from 1.68 for IR 40-50% to

2.59 for ER 50-60%. The highest errors occurred when ER 50-60% was excluded, where XGBoost had an MAE of 2.59 and SVR had an MAE of 2.54.

IR movements showed lower error values compared to ER movements. When IR 30-40% was excluded, models recorded MAE values ranging from 2.03 (SVR) to 2.21 (RF). Excluding IR 40-50% resulted in lower errors, with XGBoost achieving an MAE of 1.68. ER movements showed higher variability in generalizability across different task settings. While model performance declined compared to single task and combined ones training, errors were still much lower than the baseline mean prediction MAE of 3.2.

D. Features Importance Analysis

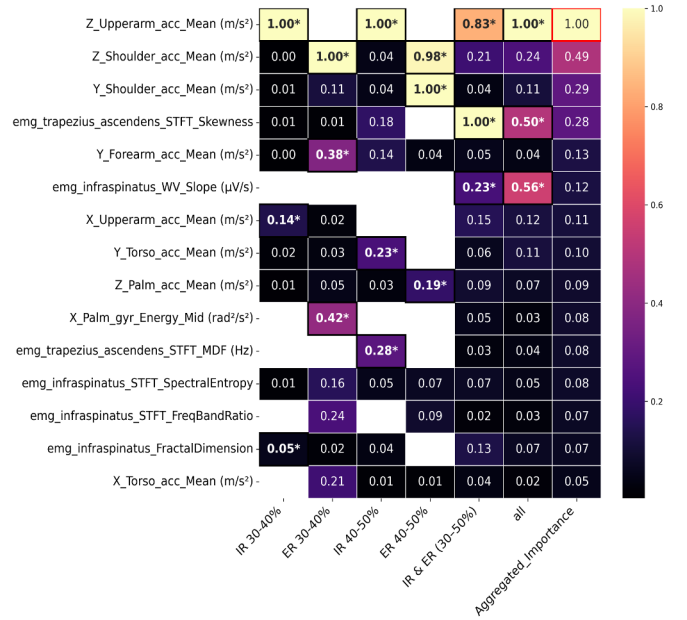


Fig. 2: Heatmap of the top 15 features by overall aggregated importance across movement-resistance task settings. Asterisks (*) indicate the top 3 most important features for each task setting. IR & ER (30-50%): Combined data from 30-40% and 40-50% resistance levels for both internal and external movements. all: Combined data from all resistance levels for both IR and ER movements. Importance scores have been normalized for comparability. STFT: Short-time Fourier Transform, WV: Wavelet, MDF: Median Frequency

We conducted a feature importance analysis to identify key predictors of fatigue using both IMU and EMG data across various movement-resistance task settings. The heatmap in Fig. 2 shows the top 15 features and their normalised importance scores, with lighter colours indicating higher contribution to fatigue estimation across individual and combined task settings. Top IMU features were average accelerations of the upper arm and shoulder, y-axis shoulder acceleration and palm rotational energy. From EMG, top features are the frequency distribution asymmetry of the trapezius muscle and the rate

of change in activation signals from the infraspinatus muscle which were extracted through time-frequency transformations such as STFT and wavelet analysis.

To further validate the relationship between the top acceleration features and fatigue, we performed correlation analysis and independent t-tests. Pearson's correlation showed a negative relationship between Borg RPE scores and both upper arm and shoulder acceleration, with upper arm acceleration showing stronger sensitivity to fatigue. Statistical testing confirmed these differences, with significantly lower acceleration values found in the high fatigue group compared to the low fatigue group ($p < 0.0001$ for upper arm; $p = 0.0028$ for shoulder).

IV. DISCUSSION AND CONCLUSION

This study presented the effectiveness of using wearable sensor data to estimate perceived fatigue which aligns with prior research on IMU and EMG sensors for tracking movement and neuromuscular responses [16]. Our framework achieved robust fatigue estimation across diverse movement resistance settings by combining subjective Borg scores with interpretable biomechanical features and applying machine learning across multiple participants. While predicting Borg scores is inherently challenging due to individual variability and subjectivity, our approach reduced this inconsistency through generalizable modeling.

Our results showed strong predictive performance, especially for internal rotation tasks. However, a significant drop in performance for the ER 50–60% task setting during leave-one-task-out validation shows a trade-off between generalizability and task specific accuracy. This may be due to limited data for ER movements and nuanced aspects that could not be effectively learned with such limited data.

The top features identified are consistent with existing literature, which associates fatigue with reduced acceleration, increased movement variability and compensatory strategies in joints [17]. EMG features such as trapezius signal skewness and infraspinatus wavelet slope are also consistent with established neuromuscular fatigue markers linked to changes in spectral content and motor unit recruitment [18].

From an applied perspective, our findings support the potential of sensor based fatigue estimation in real-time systems for sports, rehabilitation and occupational settings. The ability of our models to generalize across different task conditions indicates potential applicability in real-world scenarios with different movement variations.

Despite these contributions, the study has limitations. The data was collected in a controlled lab environment, which may not fully capture the variability encountered in real-world settings. The dataset was mainly on arm movements which highlights the need to test these extracted biomechanical features for other exercises and movement types. Future work should focus on expanding datasets to include a broader range of participants, tasks and environmental conditions as well as exploring deep learning architectures to capture temporal fatigue patterns.

ACKNOWLEDGMENT

This work was conducted with the financial support of Science Foundation Ireland under Grant Nos. 12/RC/2289-P2-INSIGHT2, 13/RC/2077-CONNECT, 21/RC/10303_P2-VISTAMILK2 and 16/RC/3918 which are co-funded under the European Regional Development Fund. This research was partially supported by the EU's Horizon Digital, Industry and Space program under grant agreement ID 101092989-DATAMITE. For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

REFERENCES

- [1] R. M. Enoka and J. Duchateau, "Translating fatigue to human performance," *Medicine and Science in Sports and Exercise*, vol. 48, no. 11, pp. 2228, 2016.
- [2] M. Leroux and F. Lagniaux, "Relationship between Shoulder Pain, Trunk and Lower Limb Pain in Overhead Athletes: A Systematic Review with Meta-analysis," *International Journal of Sports Physical Therapy*, vol. 19, no. 12, pp. 1509, 2024.
- [3] G. A. Borg, "Psychophysical bases of perceived exertion," *Medicine and Science in Sports and Exercise*, vol. 14, no. 5, pp. 377–381, 1982.
- [4] S. P. Nair, M. Sica, S. Tedesco, and A. Visentin, "Machine learning physical fatigue estimation approach based on IMU and EMG wearable sensors," *CEUR-WS.org*, 2024.
- [5] R. Merletti and P. J. Parker, "Electromyography: Physiology, Engineering, and Non-Invasive Applications," vol. 11. John Wiley & Sons, 2004.
- [6] H. J. Luinge and P. H. Veltink, "Measuring orientation of human body segments using miniature gyroscopes and accelerometers," *Medical and Biological Engineering and Computing*, vol. 43, pp. 273–282, 2005.
- [7] N. R. Adão Martins, S. Annaheim, C. M. Spengler, and R. M. Rossi, "Fatigue monitoring through wearables: A state-of-the-art review," *Frontiers in Physiology*, vol. 12, pp. 790292, 2021.
- [8] M. Rana and V. Mittal, "Wearable sensors for real-time kinematics analysis in sports: A review," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1187–1207, 2020.
- [9] T. Op De Beeck, W. Meert, K. Schütte, B. Vanwanseele, and J. Davis, "Fatigue prediction in outdoor runners via machine learning and sensor fusion," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 606–615, 2018.
- [10] A. Biró, A. I. Cuesta-Vargas, and L. Szilágyi, "AI-assisted fatigue and stamina control for performance sports on IMU-generated multivariate time series datasets," *Sensors*, vol. 24, no. 1, pp. 132, 2024.
- [11] Y. Ganesan, S. Gobee, and V. Durairajah, "Development of an upper limb exoskeleton for rehabilitation with feedback from EMG and IMU sensor," *Procedia Computer Science*, vol. 76, pp. 53–59, 2015.
- [12] J.-W. Seo and H.-S. Kim, "Biomechanical analysis in five bar linkage prototype machine of gait training and rehabilitation by IMU sensor and electromyography," *Sensors*, vol. 21, no. 5, pp. 1726, 2021.
- [13] M. N. Yasar, M. Sica, B. O'Flynn, S. Tedesco, and M. Menolotto, "A dataset for fatigue estimation during shoulder internal and external rotation movements using wearables," *Scientific Data*, vol. 11, no. 1, pp. 433, 2024.
- [14] F. Crenna, G. B. Rossi, and M. Berardengo, "Filtering biomechanical signals in movement analysis," *Sensors*, vol. 21, no. 13, p. 4580, 2021.
- [15] M. S. Redfern, R. E. Hughes, and D. B. Chaffin, "High-pass filtering to remove electrocardiographic interference from torso EMG recordings," *Clinical Biomechanics*, vol. 8, no. 1, pp. 44–48, 1993.
- [16] J. R. Cram, *Introduction to Surface Electromyography*, 1st ed., Aspen Publishers, Gaithersburg, MD, 1998.
- [17] D. H. Gates and J. B. Dingwell, "The effects of muscle fatigue and movement height on movement stability and variability," *Experimental Brain Research*, vol. 209, no. 4, pp. 525–536, 2011.
- [18] J. L. Dantas, T. V. Camata, M. A. O. C. Brunetto, A. C. Moraes, T. Abrão, and L. R. Altimari, "Fourier and wavelet spectral analysis of EMG signals in isometric and dynamic maximal effort exercise," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2010, pp. 5979–5982.