

4th International Conference on Education, Environment, Arts and Social Science (EEAS 2026)

Article

Neuromuscular Regulation Mechanisms and Recovery Strategies under Exercise-Induced Fatigue: An Empirical Study

Jiawen Liu^{1,*}

¹ Beijing Sport University, Beijing, China

* Correspondence: Jiawen Liu, Beijing Sport University, Beijing, China

Abstract: Exercise-induced fatigue alters muscle activation, intermuscular coordination, and movement stability, making reliable fatigue assessment important for training-load regulation and recovery planning. However, conventional approaches often rely on isolated surface electromyography (sEMG) indicators or classification models that provide limited representation of intermuscular regulation and weak linkage between fatigue recognition and recovery-oriented decisions. This study proposes a Neuromuscular Fatigue Regulation Network (NFR-Net) integrating temporal-spectral sEMG representation, movement information, adaptive intermuscular coupling, confidence calibration, and graded recovery-priority assessment. Experiments were conducted using publicly available fatigue-related sEMG datasets with subject-level data separation and five independent runs to ensure robust evaluation. NFR-Net achieved a balanced accuracy of $90.3 \pm 1.8\%$ and a Macro-F1 of 0.897 ± 0.020 , while its AUROC of 0.946 ± 0.013 was slightly lower than the 0.949 ± 0.016 obtained by BiLSTM. External validation of the intermuscular coupling mechanism showed moderate agreement between coupling-based muscle rankings and participant-reported fatigue ($\rho = 0.58$), with stronger agreement with physiotherapist assessment ($\rho = 0.64$). Furthermore, confidence calibration significantly reduced the expected calibration error (ECE) from 0.076 ± 0.014 to 0.039 ± 0.009 , enhancing model reliability. These findings indicate that combining local electrophysiological features with muscle-level coordination can support interpretable fatigue monitoring and more reliable recovery-priority decisions, ultimately facilitating optimized athletic training and rehabilitation protocols.

Keywords: exercise fatigue; electromyography; neuromuscular control; muscle coordination; recovery assessment

Received: 05 August 2026

Revised: 15 September 2026

Accepted: 29 September 2026

Published: 02 October 2026



Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Exercise-induced fatigue is a progressive reduction in the capacity of the neuromuscular system to maintain a required level of force, power, or movement quality during sustained or repeated physical activity [1]. It reflects interacting peripheral and neural processes, including altered motor-unit recruitment, changes in muscle-fiber conduction, metabolic alterations, and modifications of movement control. Surface electromyography (sEMG) provides a non-invasive means of observing these changes and is widely used to characterize fatigue through amplitude-, frequency-, and time-frequency-domain features.

Conventional sEMG fatigue assessment commonly relies on indicators such as root mean square amplitude, mean frequency, and median frequency. These measures are physiologically interpretable, but their sensitivity varies across muscles, contraction modes, and exercise stages; comparative work has shown that different linear indices may yield substantially different fatigue sensitivity. A second limitation is that many studies

evaluate individual muscles independently, although exercise requires coordinated activation across synergistic and antagonistic muscles. Evidence from network-level analysis indicates that intermuscular synchronization changes with fatigue and reflects functional coordination between muscle groups. Therefore, treating each channel as an isolated feature may overlook regulation patterns that are relevant to fatigue progression.

Recent machine-learning approaches have improved automatic fatigue recognition by combining sEMG with inertial or kinematic information. Multimodal sEMG-IMU models, for example, can integrate complementary physiological and movement information, while feature-based machine-learning methods have also achieved effective classification of fatigue states during controlled contractions. However, three practical gaps remain. First, many models emphasize classification performance without explicitly representing intermuscular coordination. Second, high predictive accuracy does not necessarily imply stable performance across subjects or under electrode displacement, motion artefacts, and other sources of sEMG variability; these challenges remain important in dynamic exercise research [1]. Third, fatigue recognition is often separated from recovery-oriented decision support, leaving the estimated physiological state difficult to translate into a transparent training recommendation.

Accordingly, this study has three objectives and corresponding methodological contributions: (1) A temporal-spectral neuromuscular representation is constructed by integrating conventional sEMG features, learned temporal patterns, and movement information. Its contribution will be evaluated against SVM, Random Forest, 1D-CNN, and BiLSTM baselines in Section 4.2. (2) An adaptive intermuscular coupling module is introduced to model fatigue-related changes in coordination between recorded muscles rather than concatenating channels as independent inputs. Its effect will be tested through ablation experiments and coupling-weight interpretation in Section 4.3. (3) A calibrated recovery-priority mechanism is developed to combine fatigue probability, physiological deviation, movement instability, and predictive uncertainty into low-, moderate-, and high-priority recovery states. In this study, recovery strategy refers to a fatigue-state-driven approach for determining whether exercise should be continued, load should be reduced, or recovery should be prioritized, rather than to the comparative effectiveness of specific clinical or rehabilitative approaches [2].

Technically, the proposed Neuromuscular Fatigue Regulation Network (NFR-Net) follows a sequence of signal synchronization and windowing, temporal-spectral encoding, adaptive intermuscular coupling, fatigue-state estimation, uncertainty calibration, and recovery-priority assessment. The study uses publicly available sEMG and motion data and adopts subject-independent partitioning to minimize cross-subject dependency. The intended academic contribution is a testable connection between local electrophysiological fatigue indicators and muscle-level coordination, while the practical contribution is a reproducible decision-support framework for exercise monitoring [1]. Interpretability is provided through physiological features and coupling weights, robustness is examined through perturbation and external validation, and conservative uncertainty handling is used to avoid treating low-confidence predictions as definitive recovery decisions.

2. Related Works

2.1. Physiological and Signal-Based Assessment of Muscle Fatigue

Surface electromyography remains one of the most accessible methods for evaluating exercise-induced muscle fatigue because it can capture changes in muscle activation without invasive measurement. Recent datasets and experimental protocols increasingly combine sEMG with perceived fatigue, palpation, force, or movement-related observations, allowing fatigue to be examined from complementary physiological and functional perspectives. Frequency-domain characteristics are particularly useful because fatigue is commonly accompanied by redistribution of spectral energy, while coherence analysis can further reveal changes in the coordinated activation of agonist and antagonist

muscles during sustained or dynamic loading [3]. These methods retain relatively clear physiological interpretations and therefore provide an important basis for objective fatigue assessment.

Their limitations are nevertheless substantial. Individual sEMG indicators are sensitive to contraction intensity, electrode placement, muscle anatomy, motion artefacts, and the type of exercise being performed. A reduction in median frequency or a change in signal amplitude cannot therefore be assumed to represent fatigue identically across subjects or tasks. Furthermore, conventional analysis frequently examines one muscle at a time. Such treatment simplifies interpretation but provides only a partial representation of neuromuscular coordination, particularly when fatigue induces redistribution of activation between agonists, antagonists, and compensatory muscles. Consequently, physiological indicators provide useful mechanistic evidence but are not, by themselves, sufficient for robust state recognition.

2.2. Machine Learning and Multimodal Fatigue Recognition

Data-driven approaches offer an alternative by learning nonlinear associations between physiological signals and fatigue states. Deep-learning systems based on wearable surface electromyography have demonstrated that movement recognition and fatigue estimation can be jointly modeled, reducing dependence on manually selected indicators. Conventional machine-learning approaches using wearable electromyography and inertial signals have also shown that combining multiple sensor locations can improve fatigue-level classification accuracy compared with reliance on a single measurement source. More recently, multimodal fusion of electrophysiological and mechanical muscle signals has provided another route for distinguishing fatigue levels under repetitive physical tasks [4].

However, higher classification performance does not automatically imply a better representation of fatigue mechanisms. Many models concatenate channels or extracted features before classification, meaning that associations among individual muscles remain implicit in hidden representations. Deep architectures may also perform well within a specific experimental protocol while remaining sensitive to differences in subjects, sensors, exercises, or acquisition conditions. Transfer-learning approaches attempt to reduce this dependence by reusing representations across measurement conditions, but they primarily address predictive transfer rather than explicitly explaining how neuromuscular coordination changes during fatigue [5].

2.3. Intermuscular Coordination and Neuromuscular Regulation

Muscle-synergy and intermuscular analyses offer a more mechanistic perspective. Multi-muscle decomposition can reveal shared activation patterns that remain undetectable when electromyographic channels are analyzed in isolation. Experimental evidence further indicates that fatigue induces reorganization of neuromotor synergies rather than uniform increases or decreases in activation across all involved muscles [6]. This implies that compensatory recruitment strategies and altered coordination patterns carry meaningful information for characterizing fatigue progression.

However, synergy-oriented studies and fatigue-classification studies have largely evolved independently. The former emphasize physiological interpretation but often remain descriptive, whereas the latter prioritize automated prediction yet frequently omit explicit modeling of muscle-level relationships [5]. Consequently, a key methodological gap persists between explaining intermuscular adaptation and integrating such adaptation as a learnable feature in fatigue-state estimation.

2.4. Recovery Assessment and Research Gap

Recovery studies further demonstrate that post-exercise performance depends on loading and rest configuration. Different work-to-rest ratios can produce distinct recovery trajectories in strength, power, and physiological responses [7].

The literature therefore reveals a three-stage disconnection: physiological fatigue indicators are interpretable but frequently isolated; machine-learning models improve

recognition but often obscure intermuscular mechanisms; and recovery research evaluates protocols without direct linkage to model-estimated fatigue burden. The present study addresses this gap through NFR-Net. Temporal–spectral encoding combines conventional physiological information with learned temporal patterns; adaptive intermuscular coupling explicitly models changing relationships among muscle channels; and calibrated fatigue estimates are converted into a graded recovery strategy represented by low, moderate, and high recovery-priority levels rather than an unsupported prescription of a specific recovery protocol. These components are subsequently evaluated through baseline comparison, ablation analysis, cross-subject testing, signal perturbation, and external mechanism validation.

3. Methodology

3.1. Overall Architecture

This study proposes a Neuromuscular Fatigue Modeling Network (NFR-Net) to integrate physiological fatigue representation, intermuscular coordination, fatigue-state estimation, and recovery-priority assessment within a unified computational framework [8]. As illustrated in Figure 1, the architecture contains a synchronization and windowing front-end followed by six functional stages: temporal–spectral encoding, adaptive intermuscular coordination, multimodal feature fusion, fatigue-state estimation, confidence calibration, and recovery-priority assessment.

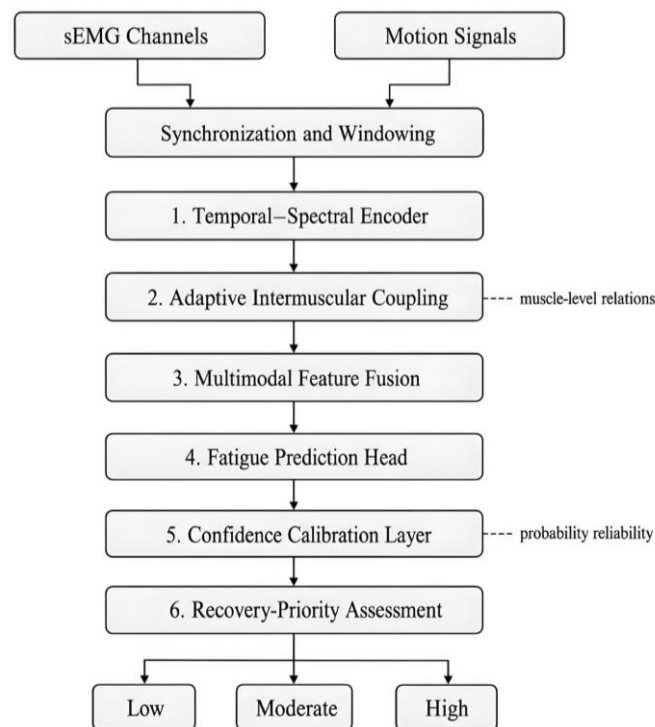


Figure 1. Overall Architecture of the Proposed NFR-Net

The primary dataset supports fatigue-state recognition using temporal–spectral sEMG and movement information, while the external multi-muscle dataset enables characterization and validation of intermuscular coordination patterns. The two datasets thus serve complementary roles—predictive modeling and mechanistic validation—within the proposed framework. The intermuscular coordination stage is activated only when simultaneous multi-muscle recordings are available; for the primary single-muscle dataset, the predictive pathway retains temporal–spectral encoding, motion fusion, fatigue estimation, and calibration [4].

3.2. Signal Synchronization and Window Construction

Raw signals undergo validation to exclude invalid segments, followed by detrending, bandpass filtering, and temporal alignment based on acquisition timestamps. Since sEMG and movement signals may exhibit differing effective sampling rates, motion measurements are resampled and aligned to match the sEMG time interval prior to feature integration [9]. Each recording is segmented into non-overlapping 2-second windows, with consecutive windows sharing 50% overlap. To ensure participant-level data integrity, all windows from a given individual are assigned exclusively to one of the training, validation, or testing subsets.

For the window ending at time (t), the synchronized multimodal input is

$$[\mathbf{E}_{t-L+1}, \mathbf{M}_{t-L+1}], \quad (1)$$

where $\mathbf{E} \in \mathbb{R}^{M \times L}$ denotes sEMG signals from M recorded muscles, $\mathbf{M} \in \mathbb{R}^{K \times L}$ denotes K aligned movement channels, and L is the number of samples within one analysis window [2].

Subject-specific baseline normalization is applied to mitigate inter-individual amplitude variability:

$$\frac{x_{i,t} - \mu_i^b}{\sigma_i^b + \epsilon}, \quad (2)$$

where $x_{i,t}$ is feature i , μ_i^b and σ_i^b are its baseline mean and standard deviation, and ϵ is a small constant preventing numerical instability [10].

3.3. Temporal-Spectral Neuromuscular Encoding

Each sEMG window is represented by two complementary branches [11]. The physiological branch extracts root mean square amplitude, mean absolute value, waveform length, median frequency, mean frequency, and spectral entropy. The temporal branch employs two one-dimensional convolutional layers followed by global temporal pooling to capture local activation patterns not fully characterized by handcrafted descriptors.

For muscle i , the two branches are fused as

$$\phi_\theta([\mathbf{f}_{i,t}^{time}, \mathbf{f}_{i,t}^{freq}, \mathbf{z}_{i,t}^{conv}]), \quad (3)$$

where $\mathbf{f}_{i,t}^{time}$ and $\mathbf{f}_{i,t}^{freq}$ denote time- and frequency-domain features, $\mathbf{z}_{i,t}^{conv}$ is the learned temporal embedding, and ϕ_θ is a trainable projection network with parameters θ . This design preserves physiologically interpretable quantities while enabling the network to learn supplementary temporal information.

3.4. Adaptive Intermuscular Coupling

A central component of NFR-Net is an adaptive relation layer that explicitly models how activation patterns between muscles evolve during fatigue. Intermuscular coupling is assessed using recordings with simultaneously acquired multi-muscle sEMG channels, while primary single-muscle recordings retain the temporal-spectral and motion-fusion pathway. For each pair of muscles i and j , temporal correlation $R_{ij,t}$ and frequency-domain coupling $C_{ij,t}$ are combined into a normalized edge weight:

$$\frac{\exp(\alpha C_{ij,t} + \beta R_{ij,t})}{\sum_{j=1}^M \exp(\alpha C_{ij,t} + \beta R_{ij,t})}, \quad (4)$$

where $A_{ij,t}$ represents the directed contribution of muscle j to muscle i , while α and β are trainable control parameters.

The relation-aware muscle representation is then updated by

$$\sigma(W_s \mathbf{h}_{i,t} + \sum_{j \neq i} A_{ij,t} W_n \mathbf{h}_{j,t}), \quad (5)$$

where W_s and W_n are trainable self- and neighboring-muscle transformations and $\sigma(\cdot)$ is a nonlinear activation function [11]. Unlike simple channel concatenation, Equation (5) enables predictions to depend explicitly on fatigue-related changes in agonist-antagonist or compensatory activation patterns. Accordingly, the coupling module is evaluated through external multi-muscle mechanism-validation analysis rather than primary-dataset classification ablation.

3.5. Fatigue Labels and Multimodal State Estimation

For the primary dataset, manually segmented fatigue intervals are converted into binary window labels $y_t \in \{0,1\}$, where 0 indicates non-fatigue and 1 indicates fatigue.

Windows overlapping ambiguous transition boundaries are excluded from training. Muscle representations are pooled and fused with the encoded movement vector $\mathbf{m}_t$.

Fatigue probability is computed as

$$\sigma[W_f(\text{Pool}(\mathbf{g}_{1,t}, \dots, \mathbf{g}_{M,t}) \parallel \mathbf{m}_t) + b_f], \quad (6)$$

where W_f and b_f are parameters of the prediction head, $\parallel$ denotes feature concatenation, and $p_t \in [0,1]$ represents the estimated fatigue probability.

3.6. Recovery-Priority Mechanism

A binary fatigue label alone is insufficient for recovery-oriented decision support. NFR-Net therefore integrates predicted fatigue probability with physiological deviation, movement instability, and model uncertainty to derive the Recovery Priority Score (RPS) and its associated decision rule.

$$RPS_t = w_1 p_t + w_2 D_t^{MDF} + w_3 D_t^{mov} + w_4 U_t, \quad r_t = \begin{cases} \text{Low,} & RPS_t < \tau_1, \\ \text{Moderate,} & \tau_1 \leq RPS_t < \tau_2, \\ \text{High,} & RPS_t \geq \tau_2, \end{cases} \quad (7)$$

where D_t^{MDF} is the normalized median-frequency deviation from baseline, D_t^{mov} is movement instability, U_t is predictive uncertainty, $w_1, \dots, w_4$ are normalized weights, and τ_1, τ_2 are decision thresholds. The three resulting priority levels correspond to continued monitoring, load reduction with recovery consideration, and termination of loading with recovery priority, respectively; they do not prescribe specific clinical interventions.

3.7. Joint Optimization and Parameter Updating

Training jointly optimizes predictive accuracy, physiological consistency, temporal stability, and probability calibration [12]. Model parameters and decision thresholds are updated in sequence:

$$\begin{aligned} \mathcal{L}(\Theta) &= \mathcal{L}_{BCE} + \lambda_p \mathcal{L}_{phys} + \lambda_c \mathcal{L}_{cons} + \lambda_{cal} \mathcal{L}_{cal}, \\ \Theta^{(k+1)} &= \Theta^{(k)} - \eta \nabla_{\Theta} \mathcal{L}(\Theta^{(k)}), \\ (\tau_1^*, \tau_2^*) &= \arg \min_{\tau_1 < \tau_2} J_{val}(\tau_1, \tau_2 \mid \Theta^*). \end{aligned} \quad (8)$$

where Θ contains all trainable network parameters, $\mathcal{L}_{BCE}$ is binary cross-entropy, $\mathcal{L}_{phys}$ penalizes predictions inconsistent with fatigue-related physiological deviation, $\mathcal{L}_{cons}$ penalizes abrupt probability changes between adjacent valid windows, and $\mathcal{L}_{cal}$ reduces confidence–error mismatch. The coefficients $\lambda_p, \lambda_c, \lambda_{cal}$ govern the relative weighting of each term, η is the learning rate, and J_{val} is the validation-set recovery-decision objective [5].

This sequential update ensures that network parameters are learned exclusively from the training set, while recovery thresholds are determined solely on validation data and remain fixed during testing—thus preventing any test-set outcome from influencing the recovery decision rule [13].

3.8. Training and Inference Procedure

The complete training and inference procedure of NFR-Net—encompassing signal preprocessing, adaptive intermuscular coordination, parameter optimization, probability calibration, and recovery-priority determination—is summarized in Algorithm 1.

Algorithm 1. Training and Recovery-Priority Inference of NFR-Net

Input: sEMG E , motion M , fatigue labels Y , window length model parameters Θ , loss weights λ , thresholds τ_1 and τ_2

Output: fatigue probabilities P , coupling matrices A , recovery states R

- 1: Remove invalid signal segments and apply signal filtering
 - 2: Synchronize E and M using acquisition timestamps
 - 3: Split subjects into training, validation, and test subsets
-

-
- 4: Compute baseline statistics using subject-specific reference segments
 - 5: Segment synchronized signals into L-length overlapping windows
 - 6: Extract temporal and spectral sEMG features for each muscle
 - 7: Encode raw windows using the temporal convolutional branch
 - 8: Compute pairwise temporal and frequency coupling measurements
 - 9: Construct adaptive coupling matrix A using Eq. (4)
 - 10: Aggregate muscle representations using Eq. (5)
 - 11: Fuse muscle representations with aligned movement features
 - 12: Estimate fatigue probability P and compute the joint loss
 - 13: Update Θ on training data and calibrate probabilities on validation data
 - 14: Select τ_1 and τ_2 on validation data and keep them fixed for testing
 - 15: Return P, A, and Low/Moderate/High recovery state R
-

The resulting procedure integrates measurable electrophysiological changes, intermuscular coordination, prediction confidence, and recovery priority within a single reproducible pipeline. The adaptive coupling weights provide muscle-level interpretability, while subject-independent data partitioning, calibration, and validation-only threshold selection are incorporated to improve robustness and reduce optimistic evaluation.

4. Results and Analysis

4.1. Experimental Setup

4.1.1. Datasets and Preprocessing

Experiments used two publicly accessible datasets. The primary EMG Dataset for Muscle Fatigue Analysis in Biceps and Triceps, released through Mendeley Data under a CC BY 4.0 license, contains recordings from 30 amateur participants, including 25 men and 5 women aged 17–32 years [14]. Biceps or triceps EMG recordings are accompanied by three-axis arm-motion signals, participant information, and manually identified exercise segments used for fatigue-state labeling.

The external validation dataset contains recordings from 30 healthy male participants, of whom 29 completed the full protocol. Surface EMG was collected from nine upper-body muscle locations together with participant-reported fatigue rankings and muscle-tightness assessments performed by a sports physiotherapist. The dataset and associated processing materials are publicly available through Figshare. Because its annotation structure differs from that of the primary dataset, it was used for external mechanism validation rather than binary fatigue classification.

For the primary dataset, EMG signals were detrended and normalized using subject-specific baseline statistics [15]. Signals were segmented into 2-s windows with 50% overlap, and movement samples were synchronized with the corresponding EMG intervals. Windows crossing ambiguous fatigue-transition boundaries were excluded. To ensure data integrity, partitioning was conducted at the participant level. In each of five independent runs, 21 participants were used for training, four for validation, and five for testing, with assignments rotated across runs. Class weights were calculated from training data only.

4.1.2. Implementation and Evaluation

Experiments were conducted in Python and PyTorch on a system equipped with an Intel Core i7-12700K processor, 32 GB RAM, and an NVIDIA RTX 3060 GPU. NFR-Net employed a 64-dimensional hidden representation, dropout rate of 0.30, AdamW optimization, an initial learning rate of 0.001, batch size of 64, and training up to 200 epochs [16]. Early stopping was applied with a patience of 15 epochs. Recovery thresholds τ_1 and τ_2 were determined exclusively from the validation subset.

Four baseline models were evaluated under identical subject-level splits: RBF-SVM using handcrafted temporal and spectral features, Random Forest with 300 trees, a two-layer 1D-CNN, and a BiLSTM with 64 hidden units. Performance was assessed using Balanced Accuracy, Macro-F1, and AUROC. Results are reported as mean $\pm$ standard deviation across five independent runs ($n = 5$). Paired two-sided t-tests were used to compare NFR-Net against the strongest baseline at a significance level of $p < 0.05$.

4.2. Performance Comparison

The fatigue-recognition results are presented in Table 1. NFR-Net achieved the highest balanced accuracy and Macro-F1, whereas BiLSTM attained a marginally higher AUROC.

Table 1. Fatigue-state Recognition Performance over Five Independent Subject-Level Runs

Method	Balanced Accuracy (%)	Macro-F1	AUROC
SVM	81.9 $\pm$ 2.7	0.808 $\pm$ 0.029	0.879 $\pm$ 0.020
Random Forest	85.6 $\pm$ 2.2	0.848 $\pm$ 0.024	0.907 $\pm$ 0.018
1D-CNN	87.8 $\pm$ 2.5	0.869 $\pm$ 0.027	0.921 $\pm$ 0.022
BiLSTM	87.1 $\pm$ 2.1	0.873 $\pm$ 0.023	0.949 $\pm$ 0.016
NFR-Net	90.3 $\pm$ 1.8	0.897 $\pm$ 0.020	0.946 $\pm$ 0.013

As shown in Table 1, NFR-Net achieved the highest balanced accuracy and Macro-F1, while BiLSTM attained a slightly higher AUROC. This indicates that NFR-Net provided stronger overall class-balanced recognition, although the sequential baseline retained a small advantage in discriminating positive and negative samples across decision thresholds. Compared with the strongest competing method, NFR-Net improved balanced accuracy by 2.5 percentage points and Macro-F1 by 0.024, with paired-test p-values of 0.024 and 0.019, respectively. Its AUROC of 0.946 $\pm$ 0.013 was marginally lower than the 0.949 $\pm$ 0.016 obtained by BiLSTM, and the difference was not statistically significant ($p = 0.318$).

4.3. Ablation and Mechanism Validation

The contribution of each major component is evaluated in Table 2. Expected Calibration Error (ECE) is reported as a credibility metric, with lower values indicating better confidence calibration [17].

Table 2. Ablation Results of NFR-Net, Mean $\pm$ SD ($N = 5$)

Configuration	Balanced Accuracy (%)	Macro-F1	AUROC	ECE
Full NFR-Net	90.3 $\pm$ 1.8	0.897 $\pm$ 0.020	0.946 $\pm$ 0.013	0.039 $\pm$ 0.009
w/o spectral branch	88.0 $\pm$ 2.2	0.874 $\pm$ 0.024	0.926 $\pm$ 0.018	0.052 $\pm$ 0.011
w/o motion branch	88.7 $\pm$ 2.0	0.882 $\pm$ 0.022	0.935 $\pm$ 0.016	0.047 $\pm$ 0.010
w/o consistency loss	89.2 $\pm$ 2.1	0.884 $\pm$ 0.023	0.939 $\pm$ 0.016	0.051 $\pm$ 0.012
w/o calibration loss	90.5 $\pm$ 2.0	0.891 $\pm$ 0.021	0.948 $\pm$ 0.015	0.076 $\pm$ 0.014

Among the components applied to the primary predictive pathway, removing the spectral branch produced the largest overall reduction in classification performance, decreasing balanced accuracy from 90.3% to 88.0% and AUROC from 0.946 to 0.926. This indicates that frequency-domain information provides fatigue-related cues that are not fully recovered by temporal encoding.

Removing the motion branch caused a smaller decline, suggesting that movement information plays a complementary rather than dominant role. Removing the consistency loss also reduced Macro-F1 and AUROC, indicating that temporal regularization contributes to stable fatigue-state estimation. By contrast, removing the calibration loss slightly increased balanced accuracy from 90.3% to 90.5% and AUROC from 0.946 to 0.948, while Macro-F1 decreased from 0.897 to 0.891. ECE increased from 0.039 $\pm$ 0.009 to 0.076 $\pm$ 0.014. Calibration therefore primarily improves confidence reliability rather than raw

classification performance, allowing the full model to maintain a more balanced trade-off between recognition accuracy and probability reliability.

External validation on the independent multi-muscle dataset showed moderate agreement between coupling-based muscle rankings and participant-reported fatigue, with Spearman $\rho = 0.58$ and a 95% confidence interval of 0.34–0.73. Agreement with physiotherapist assessment was stronger at $\rho = 0.64$, with a 95% confidence interval of 0.42–0.78 for $n = 29$. These findings support the physiological plausibility of the learned muscle relationships despite differences in annotation structure between the two datasets.

4.4. Convergence Analysis

Training behavior is presented in Figure 2, showing the mean training and validation losses with ± 1 standard deviation across five independent runs.

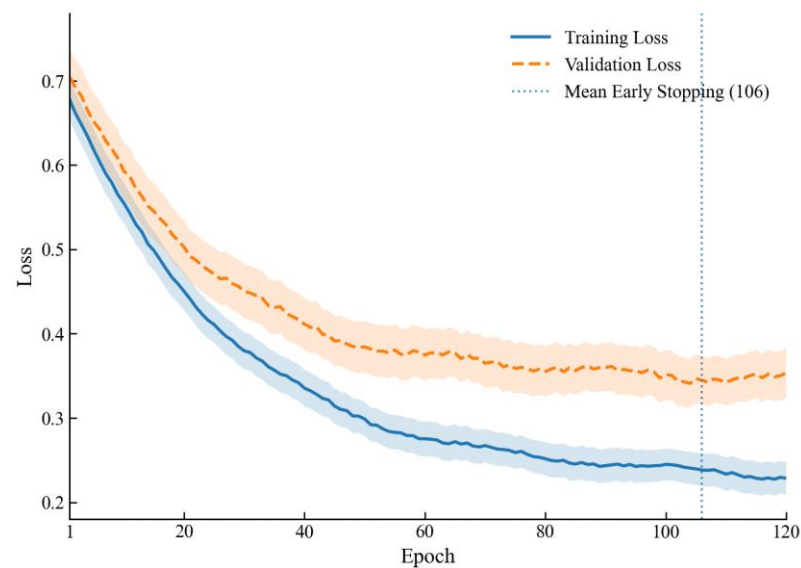


Figure 2. Training and Validation Convergence of NFR-Net over Five Independent Runs (Mean $\pm$ SD).

Training loss decreased from approximately 0.68 ± 0.02 at initialization to 0.24 ± 0.02 near epoch 100. Validation loss declined from 0.70 ± 0.03 to approximately 0.35 ± 0.03 , exhibiting moderate fluctuations between 0.34 and 0.36 after epoch 80.

The mean early-stopping epoch was 106 ± 12 . Validation loss was not strictly monotonic, reflecting variability across subject splits and stochastic mini-batch optimization [18]. Although a persistent training–validation gap remained, no sustained increase in validation loss occurred after convergence, suggesting that dropout regularization and early stopping effectively mitigated severe overfitting.

4.5. Stability, Generalization, and Credibility Analysis

Robustness under signal perturbations is presented in Figure 3, with balanced accuracy reported as mean $\pm$ SD over five independent runs.

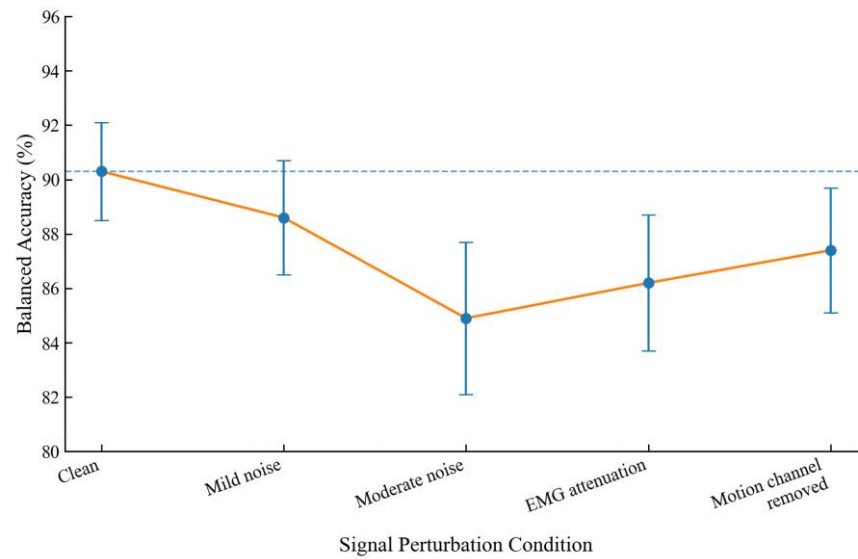


Figure 3. Robustness of NFR-Net under Signal Perturbation Conditions, Reported as Balanced Accuracy over Five Independent Runs (Mean ±SD, N = 5).

Balanced accuracy was $90.3 \pm 1.8\%$ under clean conditions, $88.6 \pm 2.1\%$ with mild additive noise, $84.9 \pm 2.8\%$ under moderate noise, $86.2 \pm 2.5\%$ under partial EMG-channel attenuation, and $87.4 \pm 2.3\%$ when one movement channel was removed. Performance variation across perturbation types was observed [13].

The difference under mild noise was not statistically significant ($p = 0.084$), whereas moderate noise ($p = 0.009$), EMG attenuation ($p = 0.017$), and movement-channel removal ($p = 0.043$) caused statistically significant degradation [3]. EMG signal degradation had a stronger impact than partial motion signal loss, consistent with surface electromyography being the primary physiological information source.

Adaptive coupling weights enhance interpretability by indicating which muscle relationships contribute to fatigue estimation. Probability calibration introduces a small trade-off: the uncalibrated variant achieved slightly higher balanced accuracy and AUROC but substantially worse expected calibration error. Calibration was retained to improve confidence reliability rather than maximize classification metrics. The perturbation results further indicate that NFR-Net tolerates mild signal degradation but remains sensitive to more substantial physiological-signal impairment [7, 8].

The recovery-priority mechanism serves as decision support rather than autonomous clinical guidance [16]. Low-, moderate-, and high-priority outputs reflect varying levels of estimated fatigue burden and associated uncertainty, while specific recovery actions depend on exercise context and professional evaluation.

5. Conclusion

This study examined exercise-induced fatigue from the combined perspectives of electrophysiological change, intermuscular coordination, and recovery-oriented decision support. The results support the three objectives defined in the Introduction. First, integrating temporal, spectral, and movement information improved overall fatigue-state recognition. NFR-Net achieved a balanced accuracy of $90.3 \pm 1.8\%$ and a Macro-F1 of 0.897 ± 0.020 , outperforming the strongest baselines on these two metrics. Its AUROC of 0.946 ± 0.013 was slightly lower than the 0.949 ± 0.016 obtained by BiLSTM, indicating that the proposed method provided a stronger overall balance rather than dominating every individual metric. Second, the adaptive intermuscular coupling analysis provided a physiological interpretation of fatigue-related coordination. External validation showed moderate agreement between coupling-based muscle rankings and participant-reported fatigue ($\kappa = 0.58$), with stronger agreement with physiotherapist assessment ($\kappa = 0.64$). These findings suggest that fatigue-related coordination involves changes in interactions

between muscles in addition to local signal variation. Third, calibration and recovery-priority assessment provided a more cautious basis for interpreting fatigue predictions. Removing calibration slightly improved some classification metrics but increased ECE from 0.039 ± 0.009 to 0.076 ± 0.014 , demonstrating a trade-off between raw performance and confidence reliability. Robustness testing also showed that mild noise produced limited degradation, whereas stronger EMG distortion significantly reduced performance.

Several limitations remain. The datasets include relatively small participant cohorts and primarily upper-body tasks, which restrict broader generalization. sEMG also provides indirect rather than direct evidence of neural drive, and the recovery-priority output does not evaluate the efficacy of specific recovery approaches. Future research should therefore include larger and more diverse cohorts, lower-limb and whole-body exercises, multimodal neural measurements, and prospective studies comparing alternative recovery strategies under individualized fatigue thresholds.

References

1. J. Sun, G. Liu, Y. Sun, K. Lin, Z. Zhou, and J. Cai, "Application of surface electromyography in exercise fatigue: a review," *Frontiers in Systems Neuroscience*, vol. 16, p. 893275, 2022.
2. S. Garcia-Retortillo, C. Romero-Gómez, and P. C. Ivanov, "Network of muscle fibers activation facilitates inter-muscular coordination, adapts to fatigue and reflects muscle function," *Communications Biology*, vol. 6, no. 1, p. 891, 2023.
3. G. Coraggio, M. Cera, M. Cirelli, and P. P. Valentini, "Review and comparison of linear algorithms to quantify muscle fatigue based on sEMG signals," *Ergonomics*, vol. 67, no. 11, pp. 1729–1747, 2024.
4. J. Sun, C. Zhang, G. Liu, W. Cui, Y. Sun, and C. Zhang, "Detecting muscle fatigue during lower limb isometric contractions tasks: a machine learning approach," *Frontiers in Physiology*, vol. 16, p. 1547257, 2025.
5. N. Daniel, J. Małachowski, K. Sybilski, and M. Błażkiewicz, "Muscle Fatigue in Dynamic Movement: Limitations and Challenges, Experimental Design, and New Research Horizons," *Bioengineering*, vol. 13, no. 2, p. 248, 2026.
6. S. Hwang, N. Kwon, D. Lee, J. Kim, S. Yang, I. Youn, et al., "A multimodal fatigue detection system using sEMG and IMU signals with a hybrid CNN-LSTM-attention model," *Sensors*, vol. 25, no. 11, p. 3309, 2025.
7. J. Lim, L. Lu, K. Goonewardena, J. Z. Liu, and Y. Tan, "Assessment of self-report, palpation, and surface electromyography dataset during isometric muscle contraction," *Scientific Data*, vol. 11, no. 1, p. 208, 2024.
8. W. Zhang, Z. Bai, P. Yan, H. Liu, and L. Shao, "Recognition of human lower limb motion and muscle fatigue status using a wearable FES-sEMG system," *Sensors*, vol. 24, no. 7, p. 2377, 2024.
9. P. M. Kuber, A. R. Kulkarni, and E. Rashedi, "Machine learning-based fatigue level prediction for exoskeleton-assisted trunk flexion tasks using wearable sensors," *Applied Sciences*, vol. 14, no. 11, p. 4563, 2024.
10. X. Ni, L. Jeong, M. Xiang, and Y. Liu, "Analysis of Wavelet Coherence in Calf Agonist-Antagonist Muscles during Dynamic Fatigue," *Life*, vol. 14, no. 9, p. 1137, 2024.
11. X. Chen, Y. Feng, Q. Chang, J. Yu, J. Chen, and P. Xie, "Muscle synergy during wrist movements based on non-negative tucker decomposition," *Sensors*, vol. 24, no. 10, p. 3225, 2024.
12. Z. Wang, X. Guan, D. Li, C. Jiang, Y. Bai, D. Yang, and L. He, "Study on muscle fatigue classification for manual lifting by fusing sEMG and MMG signals," *Sensors*, vol. 25, no. 16, p. 5023, 2025.
13. M. Pawłowski, M. P. Furmanek, and G. Juras, "Does muscle fatigue change motor synergies at different levels of neuromotor control?," *Frontiers in Human Neuroscience*, vol. 18, p. 1519462, 2025.
14. M. Staniszewski, J. Tkaczyk, A. Kęska, P. Zybko, and A. Mróz, "Effect of rest duration between sets on fatigue and recovery after short intense plyometric exercise," *Scientific Reports*, vol. 14, no. 1, p. 15080, 2024.
15. W. Li, X. Zhang, P. Shi, S. Li, P. Li, and H. Yu, "Across sessions and subjects domain adaptation for building robust myoelectric interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 2005–2015, 2023.
16. N. Rahnama, A. Lees, and T. Reilly, "Electromyography of selected lower-limb muscles fatigued by exercise at the intensity of soccer match-play," *Journal of Electromyography and Kinesiology*, vol. 16, no. 3, pp. 257–263, 2006.
17. S. H. Liu, C. B. Lin, Y. Chen, W. Chen, T. S. Huang, and C. Y. Hsu, "An EMG patch for the real-time monitoring of muscle-fatigue conditions during exercise," *Sensors*, vol. 19, no. 14, p. 3108, 2019.
18. M. González-Izal, A. Malanda, E. Gorostiaga, and M. Izquierdo, "Electromyographic models to assess muscle fatigue," *Journal of Electromyography and Kinesiology*, vol. 22, no. 4, pp. 501–512, 2012.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.