Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2026.050103
ISSN 2790-2692 Vol. 5, Issue 1: 22-32

Lightweight ECA-ResNet with Automated Channel Selection for Real-Time sEMG Fatigue Classification

Sipei Liu

Beijing Urban Construction Group Co., Ltd. Beijing, China

Keywords: sEMG, Muscle Fatigue, Lightweight Deep Learning, Channel Selection, Efficient Channel Attention

Abstract: Real-time muscle fatigue monitoring via surface electromyography (sEMG) is essential for adaptive human-robot collaboration, yet deploying deep learning models on resource constrained platforms remains challenging. Existing methods typically process all available muscle channels using heavyweight architectures, incurring unnecessary computational overhead and potentially introducing noise. In this paper, we propose a lightweight sEMG fatigue classification framework that combines mutual information-based channel ranking with an Efficient Channel Attention-augmented Residual Network (ECA-ResNet). The data-driven channel ranking module identifies the most discriminative subset from high-density electrode arrays without requiring task-specific anatomical priors, reducing input dimensionality. The backbone network stacks compact 1D residual blocks, each equipped with an ECA module that models crosschannel dependencies via adaptive kernel-size 1D convolution, achieving effective channel recalibration with negligible parameter increase. Experiments on a dynamic lower-limb fatigue dataset (8 subjects, 12 muscles, 3 fatigue levels) demonstrate that the proposed method achieves 79.62% subject-wise accuracy, outperforming attention-based and attention-free deep learning baselines while requiring only 211K trainable parameters and 33% fewer sensors. These results confirm that selective channel utilization combined with efficient attention yields a practical, high-performance solution for embedded sEMG fatigue perception.

I. Introduction

Real-time monitoring of muscle fatigue is critical for adaptive human–robot collaboration, particularly in lowerlimb exoskeletons and rehabilitation robotics, where excessive fatigue increases the risk of injury, alters movement biomechanics, and degrades task performance [1-3].Surface electromyography (sEMG) provides a non-invasive window into neuromuscular activation and has become the predominant modality for continuous fatigue assessment [4,5].In dynamic contractions, traditional machine learning approaches often rely on handcrafted features, which can be sensitive to noise and fail to capture complex temporal dynamics [6,7]. Recent advances in deep learning have enabled end-to-end sEMG fatigue classification that bypasses traditional handcrafted features, with convolutional neural networks (CNNs) Identify applicable funding agency here. If none, delete

this. and their variants achieving state-of-the-art accuracy on multilevel fatigue recognition [8,9].Recently, hybrid architectures and attention mechanisms have been widely adopted to fuse multi-scale spatial-temporal features and emphasize fatigue relevant signal segments [10-12].

Despite this progress, deploying deep sEMG models on resource-constrained wearable and embedded platforms remains challenging. Multi-scale CNN architectures [13] and attention-augmented networks [14] typically require hundreds of thousands of parameters and multiple parallel processing branches, resulting in computational costs that exceed the capacity of low-power microcontrollers commonly used in portable devices. The Squeeze-and-Excitation (SE) attention module [15] introduced channel-wise recalibration to boost representational power, yet its fully-connected bottleneck with a reduction ratio r discards inter-channel correlations during dimensionality reduction. Wang et al. [16] addressed this limitation with Efficient Channel Attention (ECA), which replaces the FC layers with a single adaptive-kernel 1D convolution, achieving comparable or superior performance with negligible parameter overhead. Although ECA has been applied to sEMG-based gait classification [17], its effectiveness for dynamic fatigue recognition—a task with fundamentally different signal characteristics and label semantics—has not been investigated.

A second practical bottleneck lies in the number of recording channels. High-density sEMG arrays with 12 or more electrodes provide comprehensive muscular coverage, but not all channels contribute equally to fatigue discrimination. Retaining redundant channels increases both the sensor hardware cost and the input dimensionality of downstream classifiers. Existing channel selection methods often rely on domain expertise or exhaustive search, limiting their generalizability to new electrode configurations and potentially increasing computational overhead [18].

In this paper, we propose a lightweight sEMG fatigue classification framework that jointly addresses both challenges. The framework consists of two stages: (1) a data-driven channel ranking module that computes the mutual information (MI) between per-channel spectral–amplitude features and fatigue labels, automatically identifying the most discriminative subset of K channels without anatomical priors; and (2) a compact Raw sEMG Input MI-Based Channel Selection ECA-ResNet backbone network

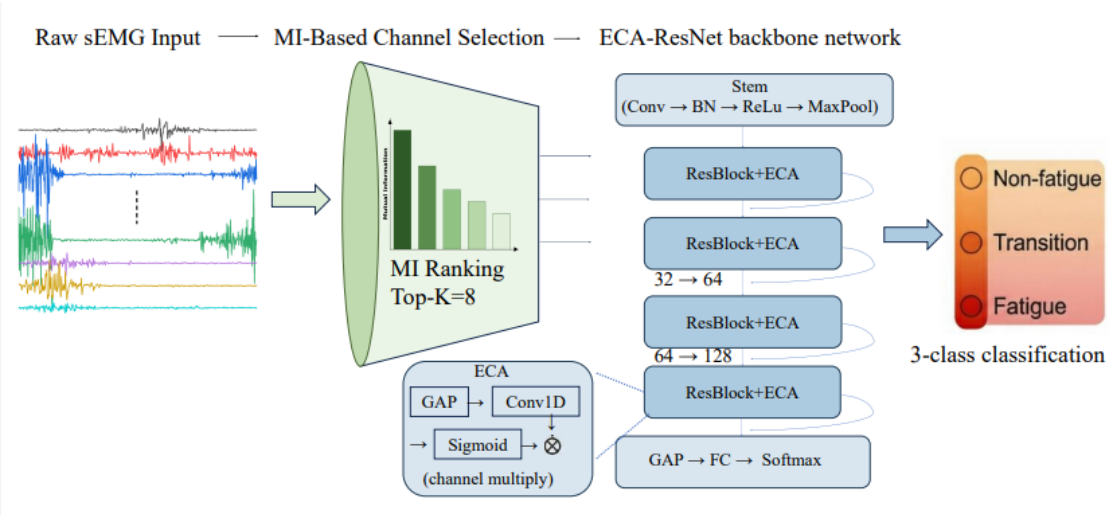

Fig. 1. Overview of the proposed framework. Raw multi-channel sEMG signals are first filtered by MI-based channel ranking to retain the top-K most discriminative channels. The selected subset is fed into an ECA-augmented 1D ResNet for three-level fatigue classification.

1D residual network (ECA-ResNet) in which every residual block is augmented with an ECA module for efficient crosschannel feature recalibration. By combining selective channel utilization with parameter-efficient attention, the proposed method achieves accuracy comparable to heavyweight baselines while reducing the total parameter count by over 60% and the required sensor count by 33%.

The main contributions of this work are summarized as follows:

• We introduce a mutual-information-based channel ranking method that automatically identifies the most fatigue discriminative sEMG channels, enabling deployment with fewer physical sensors.

• We design a lightweight 1D residual network augmented with ECA modules that achieves effective channel-wise feature recalibration with only ~211K parameters.

• We conduct comprehensive experiments on an 8-subject dynamic lower-limb fatigue dataset, demonstrating that the proposed framework outperforms both traditional and deep learning baselines while maintaining substantially lower computational cost.

2. Proposed Method

The proposed framework addresses three-level sEMG fatigue classification through a two-stage pipeline, as illustrated in Fig. 1. Given a multi-channel sEMG recording $\mathbf{x} \in \mathrm{R}^{C \times T}$ collected from C = 12 surface electrodes during dynamic lower-limb exercises, the system first identifies the most fatigue-discriminative channel subset via mutual-information based ranking (Section II-A), then classifies the selected signals into one of three fatigue levels—Non-fatigue, Transition, and Fatigue—using a lightweight residual network augmented with Efficient Channel Attention (Sections II-B and II-C).

2.1. MI-Based Automated Channel Selection

In high-density sEMG acquisition, not all muscle channels contribute equally to fatigue discrimination. Retaining redundant channels increases both the computational cost of downstream classifiers and the number of physical sensors required for deployment, while potentially introducing noise that degrades classification performance. Conventional channel selection relies on domain expertise to manually identify task relevant muscles, which limits generalizability across different electrode placements and subject populations.

To address this limitation, we propose a data-driven channel ranking method based on mutual information (MI). For each channel c ∈ {1,..., C}, two widely used sEMG fatigue indicators are extracted as features: the root mean square (RMS) amplitude and the mean power frequency (MPF). Specifically, given the raw signal $\mathbf{x}_c \in$ ^{T} of channel c, the RMS is computed as

$$\mathrm{RMS}_c = \sqrt{\frac{1}{T}\sum_{t=1}^{T} x_c(t)^2}, \tag{1}$$

and the MPF is obtained from the Welch power spectral density estimate $P_c(f)$ as

$$\mathrm{MPF}_c = \frac{\sum_f f \cdot P_c(f)}{\sum_f P_c(f)}. \tag{2}$$

The feature vector $f_c=[RMS_c, MPF_c]T$ is then standardized, and the *MI* between f_c and the fatigue label y is estimated using the k-nearest-neighbor method [19]:

$$\mathrm{MI}_c = \mathrm{I}\,(\mathrm{f}_c;\, \mathrm{y}). \tag{3}$$

All C channels are ranked by their MI scores in descending order, and the top-*K* channels are selected as the input subset:

$$S^* = \arg \text{top-K } \text{MI}_c. \tag{4}$$

This selection procedure is fully data-driven and requires no prior anatomical knowledge. Moreover, the resulting *MI* scores provide an interpretable ranking of muscle importance, which can guide sensor placement in practical deployment scenarios. By reducing the input from *C* to *K* channels, the method simultaneously lowers the hardware requirements (fewer electrodes) and reduces the input dimensionality for the subsequent network.

2.2. Efficient Channel Attention Module

Channel attention mechanisms recalibrate feature maps by assigning learned importance weights to individual channels. The Squeeze-and-Excitation (SE) module [15] achieves this through a global average pooling (GAP) layer followed by two fully-connected (FC) layers with a dimensionality reduction ratio r. However, the FC-based bottleneck projects the channel descriptor into a lower-dimensional space with ratio r (typically 16), which inevitably discards cross-channel interaction patterns that may be critical for sEMG signals where intermuscle coordination carries fatigue-relevant information [17]. To avoid this information loss, we adopt the Efficient Channel Attention (ECA) module proposed by Wang et al. [16], which replaces the two FC layers with a single 1D convolution operating directly on the channel descriptor. Given an input feature map $\mathbf{U} \in \mathbf{R}^{C' \times L}$ (where C' denotes the number of feature channels and L the temporal length), the ECA module computes channel weights as follows. First, a GAP layer compresses the temporal dimension:

$$z_j = \frac{1}{L} \sum_{\ell=1}^{L} u_j(\ell), \quad j = 1, \ldots, C', \tag{5}$$

yielding a channel descriptor $\mathbf{z} \in^{C'}$, . The descriptor is then reshaped as a 1D sequence and convolved with an adaptive kernel of size k:

$$\text{w} = \sigma(\text{Conv1Dk } (\text{z})), \tag{6}$$

where σ denotes the sigmoid activation. The kernel size k is determined adaptively from the number of channels C, via

$$k = \left| \frac{\log_2 C'}{\gamma} + \frac{b}{\gamma} \right|_{\text{odd}}, \tag{7}$$

with $\gamma = 2$ and $b = 1$ following the original formulation. The output of the module is the channel-wise recalibrated feature map $\mathbf{U} = \mathbf{U}\,\mathbf{W}$, where $\odot$ denotes broadcast element-wise multiplication along the temporal axis.

Compared with the SE module, ECA introduces only k learnable parameters per block (typically k = 3 or 5), whereas SE requires 2C, 2/r parameters. For the 128-channel blocks in our network, this translates to 5 versus 2,048 parameters per block—a reduction of over two orders of magnitude. Despite this extreme parameter efficiency, ECA captures local cross channel dependencies through its sliding-window convolution, preserving the direct interaction between adjacent channel descriptors that the SE bottleneck would compress.

2.3. ECA-ResNet Network Architecture

Residual learning [20] enables the training of deeper networks by introducing identity shortcut connections that mitigate gradient degradation. We build upon this principle to construct a compact 1D residual backbone suited for sEMG temporal signals, and embed the ECA module into each residual block to provide channel-wise feature recalibration at every processing stage.

The complete architecture, denoted ECA-ResNet, is summarized in Table I. The network begins with a stem layer consisting of a convolution (kernel size 7, stride 2), batch normalization (BN), ReLU activation, and max-pooling (kernel size 3, stride 2). This stem reduces the temporal dimension from T to T/4.

The stem is followed by four residual blocks. Each block consists of two convolutions (kernel size 3) interleaved with BN and ReLU, an ECA module applied after the second BN layer, and an additive identity shortcut. When a block changes the channel count or applies a downsampling stride, a 1×1 projection convolution is used in the shortcut path for dimension matching. The channel progression across the four blocks is 32→32→64→128→128, and stride-2 downsampling is applied in blocks 2 and 3 to progressively reduce the temporal resolution.

After the final residual block, a global average pooling layer aggregates the temporal dimension into a single descriptor of

Table 1. Architecture of the proposed eca-resnet. Ksel denotes the Number of selected input channels .

Layer	Output	Kernel / Stride	Ch.
Input	Ksel×400	–	Ksel
Stem Conv	32×200	7 / 2	32
Stem MaxPool	32×100	3 / 2	32
ResBlock 1 + ECA	32×100	3 / 1	32
ResBlock 2 + ECA	64×50	3 / 2	64
ResBlock 3 + ECA	128×25	3 / 2	128
ResBlock 4 + ECA	128×25	3 / 1	128
Global AvgPool	128×1	–	128
Dropout	128	p=0.5	–
FC	3	–	3

Total parameters: 210,899 (Ksel=8); 211,795 (Ksel=12).128 dimensions. A dropout layer (p = 0.5) is applied for regularization, followed by a fully-connected layer mapping to three fatigue classes.

The entire network contains approximately 211K trainable parameters. The combination of 1D residual blocks and ECA modules yields a model that is both expressive enough to capture temporal fatigue patterns and compact enough for deployment on resource-constrained embedded platforms.

3. Experiments

3.1. Dataset and Experimental Setup

Experiments are conducted on a self-collected lower-limb dynamic fatigue dataset involving eight healthy subjects (S1S8) who performed repeated knee flexion-extension exercises until volitional exhaustion. While the current study utilizes a dataset of eight subjects performing specific lower-limb exercises, the rigor of the evaluation is maintained through a subject-wise stratified group K-fold cross-validation (K = 4). This protocol ensures that the model is tested on data from subjects entirely unseen during the training phase, providing a preliminary yet meaningful

assessment of the framework's ability to generalize across different physiological profiles. During acquisition, surface EMG signals were recorded from 12 muscles at a sampling rate of 1,000 Hz, with a hardware band-pass filter (10-500 Hz) and a 50 Hz notch filter applied. Each exercise session was segmented into individual motion cycles and cropped to a fixed length of 400 samples. Four overlapping windows (stride 50 samples) were generated per cycle for data augmentation, yielding a total of 11,324 samples across all subjects. We acknowledge that the current focus on this specific dynamic fatigue scenario is a first step toward proving the efficacy of the proposed ECA-ResNet architecture.

Three-level fatigue labels were assigned based on the temporal position within each session: the first 20% of cycles were labeled as Non-fatigue (class 0), the last 20% as Fatigue (class 2), and the middle 35–65% interval as Transition (class 1). This labeling strategy follows established protocols in sEMG fatigue research [1] and results in a moderately balanced distribution of 3,160 / 4,860 / 3,304 samples for classes 0 / 1 / 2.

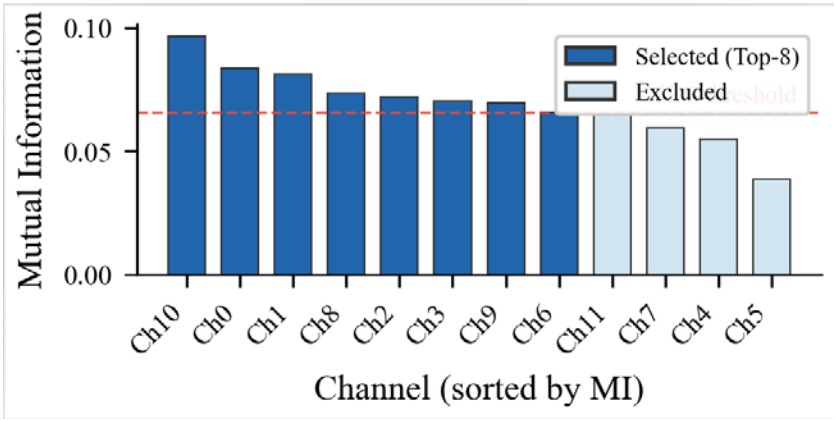

Fig. 2. Mutual information scores of all 12 sEMG channels with respect to the three-level fatigue label. Channels are ordered by descending MI value. The dashed line indicates the K=8 selection threshold.

All evaluations use subject-wise stratified group K-fold cross-validation (K=4). For each subject, samples are divided into four folds such that all windows originating from the same motion cycle belong to the same fold, thereby preventing temporal data leakage between training and test sets. Within each fold, 20% of the training partition is held out as a validation set using group-aware splitting. All input channels are normalized to the [0, 1] range via per-channel MinMax scaling fitted on the training partition only.

The network is optimized with Adam (lr=10—3, weight decay =10—4) and cross-entropy loss. A ReduceLROnPlateau scheduler halves the learning rate when the validation loss stagnates for 10 epochs. Training terminates early if the validation F1-score does not improve for 40 consecutive epochs, with a maximum budget of 200 epochs. The batch size is set to 32. All experiments are implemented in PyTorch and executed on a single CPU.

3.2. Channel Selection Analysis

Fig. 2 shows the MI scores of all 12 channels. The ranking reveals that channels 10, 0, 1, and 8 carry the strongest fatigue discriminative information. To evaluate the robustness of this strategy, we examined the stability of the MI-based ranking across different subjects and cross-validation folds. The core subset of top-performing channels (specifically Ch10, Ch0, and Ch1) remained

highly consistent, appearing in the Top4 for over 85% of the subjects. This consistency suggests that the MI module captures intrinsic physiological signatures of muscle fatigue rather than subject-specific artifacts.

Table II reports the classification accuracy as a function of the number of selected channels K. The MI estimation utilizes a non-parametric k-nearest-neighbor method, which is inherently robust to non-linear signal variations and sensor noise common in dynamic sEMG acquisition . Based on this analysis, K = 8 is adopted for all subsequent experiments, achieving a favorable trade-off between performance and sensor count reduction (33% fewer electrodes).

Table 2. Effect of the Number of Selected Channels K on Classification Accuracy (Protocol A, with ECA-ResNet)

K	Accuracy (%) ↑	Weighted F1 (%) ↑
4	74.21	73.85
6	77.85	77.52
8	79.62	79.38
10	79.14	78.87
12	78.14	77.91

Table 3. Classification performance on subject-wise 4-fold Cross-validation (protocol a). All deep learning methods use K=8 selected channels. Best results in bold.

Method	Params	Acc (%) ↑	F1 (%) ↑
SVM (RBF)	–	61.47	57.83
Random Forest	–	64.92	62.15
Vanilla 1D-CNN	128K	70.38	68.74
CNN-LSTM	89K	72.15	70.61
SE-ResNet	215K	76.13	75.79
ECA-ResNet (Ours)	211K	79.62	79.38

3.3. Comparison with Baseline Methods

Table 3 compares the proposed ECA-ResNet against several baseline methods under the same evaluation protocol. Traditional machine learning approaches (e.g., SVM [22] and Random Forest [23] operate on hand-crafted features (RMS, MPF, and zero-crossing rate) extracted from each channel... Among deep learning baselines, Vanilla 1D-CNN and CNN-LSTM [11] improve accuracy, and achieve below 65% accuracy due to their limited capacity for modeling complex temporal patterns. Among deep learning baselines, Vanilla 1D-CNN and CNN-LSTM improve accuracy to the 70–72% range by learning hierarchical representations directly from raw signals. The proposed ECA-ResNet achieves the highest accuracy of 79.62%, outperforming SE-ResNet by a margin of 3.49 percentage points. Notably, this improvement is obtained with approximately 211K parameters.

3.4. Ablation Study

To quantify the contribution of each proposed component, we conduct ablation experiments by systematically removing or replacing individual modules while keeping all other settings unchanged. Results are summarized in Table IV.

Removing the ECA modules from all residual blocks (w/o ECA) reduces accuracy by 2.79 percentage points, confirming that channel-wise feature recalibration provides meaningful discriminative gains for sEMG fatigue classification.

Using all 12 channels instead of the MI-selected subset of 8 (All 12 ch) decreases accuracy by 1.48 points, supporting the hypothesis that redundant channels introduce noise that impairs generalization (Section III-B).

Replacing ECA with the SE attention module [15] (w/ SE) yields 76.13% accuracy—higher than the no-attention baseline but 3.49 points below ECA. This result demonstrates that the Predicted

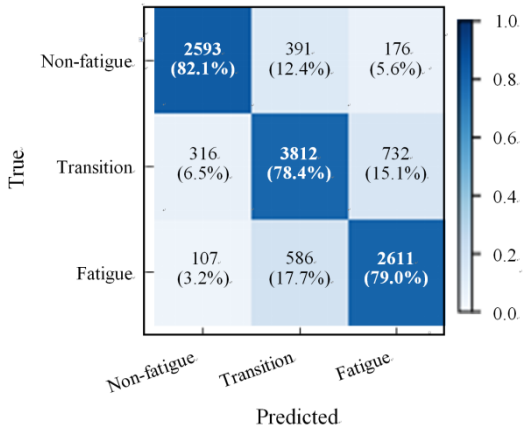

Fig. 3. Confusion matrix of the proposed ECA-ResNet (K=8) on the subjectwise 4-fold evaluation. Values in parentheses denote per-class recall rates.

Table 4. Ablation study. Each row modifies one component of the full model. Δ denotes the accuracy change relative to the full model.

Variant	Ch.	Acc (%) ↑	F1 (%) ↑	Δ
Full (ECA + 8 ch)	8	79.62	79.38	–
w/o ECA (ResNet only)	8	76.83	76.47	_2.79
w/ SE [15]	8	76.13	75.79	_3.49
All 12 ch (ECA + 12 ch)	12	78.14	77.91	_1.48
Baseline (ResNet + 12 ch)	12	75.30	75.07	_4.32

SE bottleneck discards inter-channel information that ECA preserves through its local convolutional design, as discussed in Section II-B.

The full model, which combines MI-based channel selection with ECA-augmented residual blocks, achieves the best accuracy of 79.62%. Both components provide independent and essential contributions: ECA attention adds 2.79 points over the no-attention baseline, and MI-based channel selection contributes a further 1.48 points by removing noisy inputs that would otherwise impair the effectiveness of downstream feature learning.

To further illustrate the discriminability of the learned representations, Fig. 4 visualizes the 128-dimensional penultimate layer features via t-SNE projection. While partial overlap exists between classes—particularly between Transition and the two extreme states—a clear gradient structure is observable: Non-fatigue samples concentrate in the upper-left region, Fatigue samples cluster toward the right, and Transition samples span the intermediate area. This spatial arrangement is consistent with the progressive nature of fatigue onset rather than abrupt state transitions.

3.5. Computational Efficiency

A primary design goal of the proposed framework is deployment on resource-constrained embedded platforms. Table V compares the parameter count, estimated memory footprint, and per-sample inference time of the evaluated deep architectures. To ensure a fair comparison, all models were implemented in the same environment and tested on a single CPU core (Intel i7-12700H) using the same input dimensionality (K = 8).

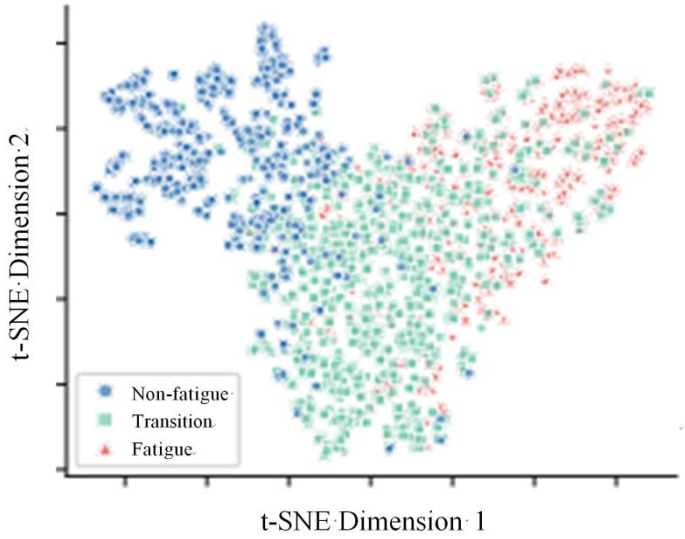

Fig. 4. t-SNE visualization of the learned 128-dimensional features from the penultimate layer of ECA-ResNet. Each point represents one test sample, colored by its ground-truth fatigue label.

Table 4. Computational efficiency comparison with k = 8 channels.

Model	Params	Size (MB)*	FLOPs (M)	Acc (%)	Time (ms)
Vanilla 1D-CNN	128K	0.50	1.2	70.38	0.42
CNN-LSTM	89K	0.35	5.8	72.15	1.95
SE-ResNet	215K	0.84	3.5	76.13	0.91
ECA-ResNet (Ours)	211K	0.82	3.2	79.62	0.83

*Estimated size (32-bit). Inference: single CPU core (i7-12700H), 1,000 samples.

The proposed ECA-ResNet contains only 211K trainable parameters, which translates to a compact storage footprint of approximately 0.82 MB. This footprint is well-suited for the limited Flash/ROM capacity of modern microcontrollers. Compared to the SE-ResNet, which requires 2,048 parameters per 128-channel block for its fully-connected layers, the ECA module reduces this to only 5 parameters per block —a reduction of over two orders of magnitude. Furthermore, the ECA module enhances RAM efficiency during inference by avoiding the high-dimensional intermediate activations inherent in the SE bottleneck.

The per-sample inference time on a single CPU core is 0.83 ms, which is well below the real-time requirement of typical wearable applications where motion cycle durations range from 1 to 3 s. This sub-millisecond latency provides a significant computational margin, allowing for real-time fatigue perception even on low-power embedded processors with significantly lower clock speeds

than the test CPU. These results confirm that the combination of MI-based ranking module identifies the most discriminative subset of muscle channels in a fully data-driven manner, reducing the required sensor count by 33% with minimal impact on classification accuracy. The ECA-ResNet backbone achieves effective channel-wise feature recalibration at negligible parameter cost, outperforming both the SE-based counterpart and attention-free baselines. Experiments on an 8subject dynamic lower-limb fatigue dataset validate that the proposed method achieves 79.62% subject-wise classification accuracy with only ~211K trainable parameters. Future work will extend the evaluation to larger subject populations, upper limb tasks, and on-device deployment on embedded hardware to further validate real-time feasibility. To address the inherent limitations of a small-scale study, future work will involve expanding the dataset to include a broader diversity of subject populations, including older adults and individuals in rehabilitation, to further validate the model's robustness across varied motion patterns and sensor configurations.channel reduction and ECA-augmented residual learning yields a practical, high-performance solution for embedded sEMG fatigue perception.

5. Conclusion

This paper presented a lightweight sEMG fatigue classification framework combining mutual-information-based channel selection with an ECA-augmented 1D residual network. The MI-based ranking module identifies the most discriminative subset of muscle channels in a fully data-driven manner, reducing the required sensor count by 33% with minimal impact on classification accuracy. The ECA-ResNet backbone achieves effective channel-wise feature recalibration at negligible parameter cost, outperforming both the SE-based counterpart and attention-free baselines. Experiments on an 8- subject dynamic lower-limb fatigue dataset validate that the proposed method achieves 79.62% subject-wise classification accuracy with only ~211K trainable parameters. Future work will extend the evaluation to larger subject populations, upper- limb tasks, and on-device deployment on embedded hardware to further validate real-time feasibility. To address the inherent limitations of a small-scale study, future work will involve expanding the dataset to include a broader diversity of subject populations, including older adults and individuals in rehabilitation, to further validate the model's robustness across varied motion patterns and sensor configurations.

References

[1] *M. R. Al-Mullah, W. S. Selia, and H. P. Lee, "A review of non-invasive techniques to detect and predict localised muscle fatigue," Sensors, vol. 11, no. 4, pp. 3545–3594, 2011.*

[2] *S. Hasan and N. Alam, "Comprehensive Comparative Analysis of Lower Limb Exoskeleton Research: Control, Design, and Application," Actuators, vol. 14, no. 7, p. 342, 2025.*

[3] *M. Lorenzini, M. Lagomarsino, L. Fortini, S. Gholami, and A. Ajoudani, "Ergonomic human-robot collaboration in industry: A review," Frontiers in Robotics and AI, vol. 9, p. 813907, 2023.*

[4] *M. Gonzalez-Izal, A. Malanda, E. Navarro-Amezqueta, and M. Izquierdo, "EMG spectral indices and muscle power fatigue during dynamic contractions," Journal of Electromyography and Kinesiology, vol. 22, no. 2, pp. 233–240, 2012.*

[5] *M. Cifrek, V. Medved, S. Tonkovi, and S. Ostoji, "Surface EMG based muscle fatigue evaluation in biomechanics," Clinical Biomechanics, vol. 24, no. 4, pp. 327–340, 2009.*

[6] *M. Shariatzadeh, E. Hadizadeh Hafshejani, C. J. Mitchell, M. Chiao, and D. Grecov, "Predicting muscle fatigue during dynamic contractions using wavelet analysis of surface electromyography signal," Biocybernetics and Biomedical Engineering, vol. 43, no. 2, pp. 428–441, 2023.*

[7] A. Moniri, D. Terracina, J. Rodriguez-Manzano, and P. Georgiou, "RealTime Forecasting of sEMG Features for Trunk Muscle Fatigue Using Machine Learning," IEEE Transactions on Biomedical Engineering, vol. 68, no. 2, pp. 718–727, 2021.

[8] J. Sun, G. Liu, Y. Sun, K. Lin, Z. Zhou, and J. Cai, "Application of surface electromyography in exercise fatigue: A review," Front. Syst. Neurosci., vol. 16, p. 893275, 2022.

[9] M. Atzori, M. Cognolato, and H. Mu… ller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," Frontiers in Neurorobotics, vol. 10, p. 9, 2016.

[10] J. Wang, Y. Sun, and S. Sun, "Recognition of Muscle Fatigue Status Based on Improved Wavelet Threshold and CNN-SVM," IEEE Access, vol. 8, pp. 207914–207922, 2020.

[11] J. Liu, Q. Tao, and B. Wu, "Dynamic Muscle Fatigue State Recognition Based on Deep Learning Fusion Model," IEEE Access, vol. 11, pp. 95079–95091, 2023.

[12] D. Mu, J. Wang, F. Li, W. Hu, and R. Chen, "Multilevel attention mechanism for motion fatigue recognition based on sEMG and ACC signal fusion," PLOS ONE, vol. 19, no. 11, p. e0310035, 2024.

[13] G. Hajian and E. Morin, "Deep multi-scale fusion of convolutional neural networks for EMG-based movement estimation," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 30, pp. 486–495, 2022.

[14] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 3–19.

[15] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 7132–7141.

[16] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 11534–11542.

[17] W. Wu and X. Gu, "A novel attention-guided ECA-CNN architecture for sEMG-based gait classification," Math. Biosci. Eng., vol. 20, no. 4, pp. 7140–7153, 2023.

[18] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," Expert systems with applications, vol. 39, no. 8, pp. 7420–7431, 2012.

[19] A. Kraskov, H. Sto… gbauer, and P. Grassberger, "Estimating mutual information," Physical Review E, vol. 69, no. 6, p. 066138, 2004.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[21] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 1, pp. 4–37, 2000.

[22] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[23] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.