

UDC 612.743:004.942

# Improving EMG Signal Classification with Transfer Learning Under Low-Data and Cross-Subject Conditions

*Kolomiets B. Yu., Karpliuk Ye. S.*

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

E-mail: [kolomiets.bohdan.yuriyovych@gmail.com](mailto:kolomiets.bohdan.yuriyovych@gmail.com)

Surface electromyography is a non-invasive method used for monitoring muscle activity and is widely applied in rehabilitation, prosthetics, assistive robotics, and human–computer interaction. However, its practical use often remains limited by large differences between individuals and the effort required to train models for each new user. This study explores whether transfer learning can help address these challenges when using deep learning to classify hand and wrist gestures. The experiments use a dataset that includes eleven gestures, each repeated eight times by 22 healthy participants. Three training approaches are evaluated: (i) training and testing on the same subject (intra-subject), (ii) training on some subjects and testing on a new one (inter-subject), and (iii) transfer learning with and without resetting the fully connected output layer of the convolutional neural network. All models are evaluated using a leave-one-out cross-validation strategy across both subjects and repetitions. Results show that both transfer learning methods outperform the other two approaches in terms of classification accuracy. The best performance is observed when the fully connected layer is reset before fine-tuning (F1-score = 0.907,  $\sigma = 0.074$ ). Wilcoxon signed-rank statistical tests confirm that these improvements are statistically significant, even when only a few repetitions are used for fine-tuning. In fact, using transfer learning with just four repetitions instead of eight achieves accuracy comparable to training from scratch on all eight repetitions. These findings suggest that fine-tuning pre-trained models can significantly reduce the effort needed to adapt EMG-based systems to new users, providing a practical and effective approach for developing user-friendly interfaces suited to assistive and rehabilitation applications.

*Keywords:* surface electromyography (sEMG); gesture recognition; transfer learning; deep learning; convolutional neural networks (CNN); inter-subject variability; model generalization; fine-tuning strategies; subject adaptation; myoelectric control; EMG-based interface; cross-validation; biomedical signal processing; rehabilitation technologies; low-effort calibration

DOI: [10.64915/RADAP.2026.103.61-68](https://doi.org/10.64915/RADAP.2026.103.61-68)

## Introduction

Surface electromyography (sEMG) is a non-invasive technique for monitoring muscle activity and is widely applied in rehabilitation, prosthetics, assistive robotics, and human–computer interaction [1]. It supports gesture recognition, prosthetic control, and assessment of neuromuscular recovery [2].

However, the utility of sEMG is limited by physiological and technical factors, particularly electrode placement. Minor variations in electrode position can lead to substantial differences in signal amplitude and frequency content due to anatomical variability [3]. Additionally, signal quality is affected by skin impedance, motion artifacts, and muscle fatigue [4].

Classical signal processing methods, including time- and frequency-domain analysis, wavelet transforms, and handcrafted features, have been widely used for EMG interpretation [5–7]. These techniques are com-

putationally efficient and interpretable but struggle to capture the complex spatiotemporal patterns of multichannel or dynamic sEMG data. Their generalization across subjects or sessions is also limited [8]. Deep learning, especially convolutional neural networks (CNNs), has advanced EMG processing by enabling automatic feature learning from raw signals. CNNs outperform traditional machine learning models like SVMs or decision trees in classification accuracy [9]. However, they require large, annotated datasets to generalize effectively, posing challenges in biomedical contexts where data collection is costly and time-consuming [10]. Transfer learning has emerged as a solution, enabling pre-trained models to adapt to smaller, task-specific datasets. Strategies include fine-tuning the entire model, freezing earlier layers, or employing domain adaptation [11, 12]. Their effectiveness varies depending on the dataset and model architecture.

Recent advances in wearable electronics, such as flexible sensors, edge AI modules, and low-power microcontrollers, are making it feasible to deploy deep learning models for real-time sEMG analysis [13]. Bringing these technologies together supports the creation of intelligent, adaptive systems for physical activity monitoring and neurorehabilitation.

## 1 Statement of the Problem

Despite the success of CNNs in EMG gesture classification, their performance declines significantly when applied across different subjects without retraining. This inter-subject variability limits the practicality of deploying EMG-based human-machine interfaces without per-user calibration [11, 14].

Transfer learning offers a promising way to address this issue by adapting models trained on one subject group to new individuals. However, its practical implementation, such as whether to fine-tune all layers or use fixed feature extractors, has not been carefully studied for sEMG tasks [15].

Furthermore, most studies do not systematically compare transfer learning strategies using standardized datasets or well-defined evaluation protocols. As a result, there is little guidance on which methods offer the best accuracy and stability, particularly when the amount of data collected from each subject is limited.

This work aims to fill this gap by experimentally comparing two transfer learning strategies on a real-world sEMG dataset: (1) fine-tuning with the reset of the fully connected (FC) output layer and (2) fine-tuning without resetting the FC layer. We assess classification accuracy and model variance on the 3DC dataset [16], and investigate whether these methods can reduce the amount of required subject-specific data while maintaining or improving performance.

## 2 Dataset

The *3DC Dataset* dataset comprises sEMG recordings collected from 22 able-bodied participants performing eleven distinct hand and wrist gestures. Each subject completed eight repetitions of a predefined gesture sequence, resulting in a substantial volume of labeled gesture data.

The 3DC Dataset is introduced in the work of Côté-Allard et al. [16], where the authors designed a low-cost, 3D-printed, wireless myoelectric armband called as the 3DC Armband. This device has 10 dry sEMG electrodes and a 9-axis inertial measurement unit (IMU), with a sampling rate of 1000 Hz. The data acquisition protocol involved placing the 3DC Armband on the dominant forearm of each participant, alternating the armband’s position relative to the elbow between participants to simulate real-world variability in wearability.

Subjects perform a sequence of eleven gestures, including wrist flexion, wrist extension, and various hand shapes. Each gesture lasts 5 seconds. They repeat the entire sequence eight times, taking a 5-minute rest between the fourth and fifth repetitions.

The original study uses diverse electrode placement protocols to simulate real-world variability, resulting in substantial inter-subject signal differences. On the one hand, this variability challenges the generalization capability of learned models. On the other hand, it makes the dataset a valuable benchmark for evaluating inter-subject generalization and transfer learning strategies in sEMG-based gesture classification.

## 3 Methods

This research is conducted using the LibEMG Python library [17], which provides a unified framework for data loading, preprocessing, training, and evaluation of machine learning models in electromyographic control systems.

Several studies have shown that feeding raw sEMG signals directly into deep learning models, particularly CNNs, is an effective approach for gesture classification. In their early review, Oskoei and Hu [20] outlined the challenges of handcrafted feature extraction and discussed the potential of neural networks to bypass manual processing by learning directly from raw signals. Later, Rehman et al. [19] demonstrated that deep learning models trained on raw multiday sEMG data can outperform traditional methods that rely on engineered features, especially in terms of robustness across recording sessions. Similarly, Côté-Allard et al. [11] showed that a CNN trained on raw sEMG windows achieves high classification accuracy and reduced the need for complex preprocessing steps, making the system more practical for real-time applications.

Based on these findings, this study adopts a raw-input strategy. This approach eliminates the need for manual feature engineering and allows the model to learn directly from raw signal patterns, which is particularly beneficial when handling variability in electrode placement and signal characteristics.

Following existing best practices [18] and prior works using the chosen dataset [16, 17], raw sEMG signals are split into 200-sample overlapping windows with a step size of 100. Each window forms a  $10 \times 200$  matrix that serves as input to the CNN.

### CNN Model Implementation

The CNN architecture used in this study builds on implementations from previous works [16, 17], as shown in Fig. 1. The network processes raw sEMG inputs of shape  $10 \times 200$  (channels  $\times$  samples). All layer weights are initialized using the Glorot (Xavier) uniform method, with biases set to zero. Training employs the Adam optimizer with an initial learning rate

of  $10^{-3}$ , adjusted over time using a cosine annealing schedule. Cross-entropy serves as the loss function, and early stopping is applied with a patience of four epochs and a tolerance threshold of 0.03. Each training session runs for up to 50 epochs, although convergence typically occurs within 15. Random seeds for PyTorch, NumPy, and Python’s built-in generators are explicitly set to ensure reproducibility. The exact implementation of the CNN is available in the publicly accessible repository [21].

## Evaluated training approaches

All experiments in this study follow the same cross-validation protocol to ensure a fair comparison. Specifically, a leave-one-out cross-validation (LOO-CV) strategy is applied to the gesture sequence repetitions of each subject. In each fold, we select one repetition for testing, one for validation during training, and use the remaining repetitions for model training.

Three training approaches are explored. The first, referred to as *Intra-subject training*, trains and tests the model using data from the same individual. Each repetition is assigned to either training, validation, or testing according to the LOO-CV protocol.

The second approach, called *Inter-subject training*, trains the model on data from 21 participants and evaluates its performance on the one subject excluded from training. The same repetition-based LOO-CV strategy is applied to each subject’s data.

The third approach, called *Transfer learning*, fine-tunes a model pre-trained on data from 21 subjects

using data from the left-out subject. This approach uses the same cross-validation setup described earlier. We evaluate two fine-tuning strategies: one that resets the FC output layer before fine-tuning — *Fine-tuning with FC reset*, and one that keeps the FC layer unchanged, referred to as *Fine-tuning without FC reset*. In both strategies, the convolutional layers remain trainable during fine-tuning.

To further examine whether transfer learning with FC reset continues to outperform training from scratch under limited subject-specific data, this study conducts additional experiments using fewer training repetitions. Instead of all eight available repetitions, six, four, and three are used. The same LOO-CV protocol is applied in each case to ensure consistency.

When all eight repetitions are used, the cross-validation process generates 56 folds per subject ( $8 \times 7 = 56$ ), resulting in a total of 1232 folds across 22 subjects. Using six repetitions yields 30 folds per subject ( $6 \times 5 = 30$ ), totaling 660 folds. With four repetitions, we obtain 12 folds per subject ( $4 \times 3 = 12$ ), giving 264 folds overall. Finally, three repetitions generates six folds per subject ( $3 \times 2 = 6$ ), resulting in 132 total folds.

The F1-score serves as the primary metric for evaluating classification accuracy in this study. To assess whether the observed performance differences are statistically significant, Wilcoxon signed-rank tests are applied to the distributions of F1-scores obtained from the folds of each experiment.

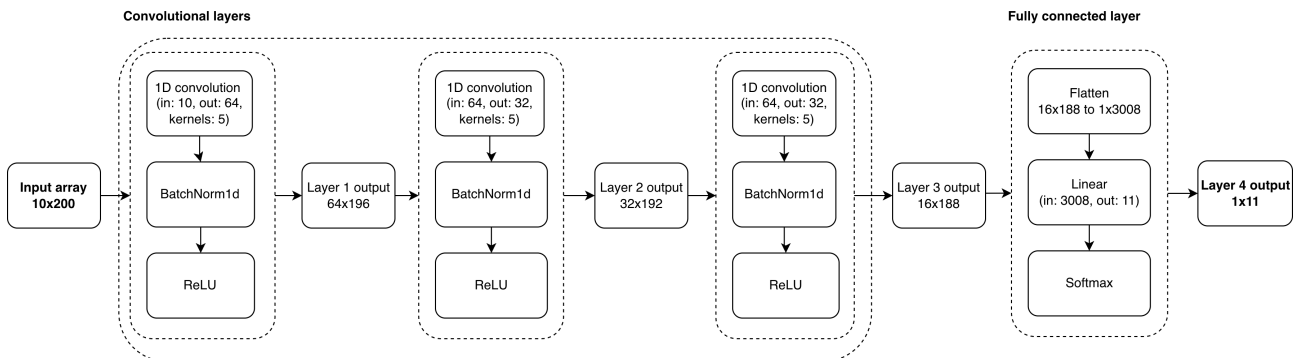


Fig. 1. Diagram of the CNN architecture used in this study. It consists of three sequential 1D convolutional layers followed by batch normalization and ReLU activations, ending with a FC layer mapping to 11 gesture classes.

## 4 Results and discussion

### Compare training approaches

Figure 2 presents box-and-whisker plots summarizing the F1-score distributions for inter-subject generalization, intra-subject training, and two transfer learning strategies: fine-tuning with and without resetting the FC layer. All eight available repetitions are used for training in those experiments. Consequently, distributions are based on  $1232(8 \times 7 \times 22)$  cross-validation folds per training approach across all 22 subjects.

The inter-subject approach yields a mean F1-score of 0.382 ( $\sigma = 0.149$ ), which confirms the significant limitations in cross-subject generalization caused by variability in electrode placement and anatomical differences. In contrast, intra-subject training, where the model trains and tests on data from the same participant, achieves a much higher mean F1-score of 0.869 ( $\sigma = 0.089$ ).

Fine-tuning pre-trained models using the “without FC reset” transfer learning strategy further improves performance to 0.896 ( $\sigma = 0.071$ ). However, the “with FC reset” strategy achieves the best result,

reaching a mean F1-score of 0.907 ( $\sigma = 0.074$ ), which suggests that resetting the CNN’s head enables better adaptation to new subjects. Both transfer learning strategies outperform the intra-subject training approach, reaffirming the benefit of leveraging pre-trained models.

### Evaluate the effect of reduced subject-specific data

To assess whether transfer learning reduces the need for subject-specific data, we conduct additional experiments using fewer training repetitions (6, 4, and 3). As shown in Fig. 3, the transfer learning approach with FC layer reset consistently achieves higher F1-scores across all scenarios. Reducing the number of repetitions to three leads to a noticeable decline in accuracy, while using four repetitions causes only a modest drop. This setup effectively cuts the required subject effort in half compared to using all eight repetitions, without significantly compromising classification accuracy.

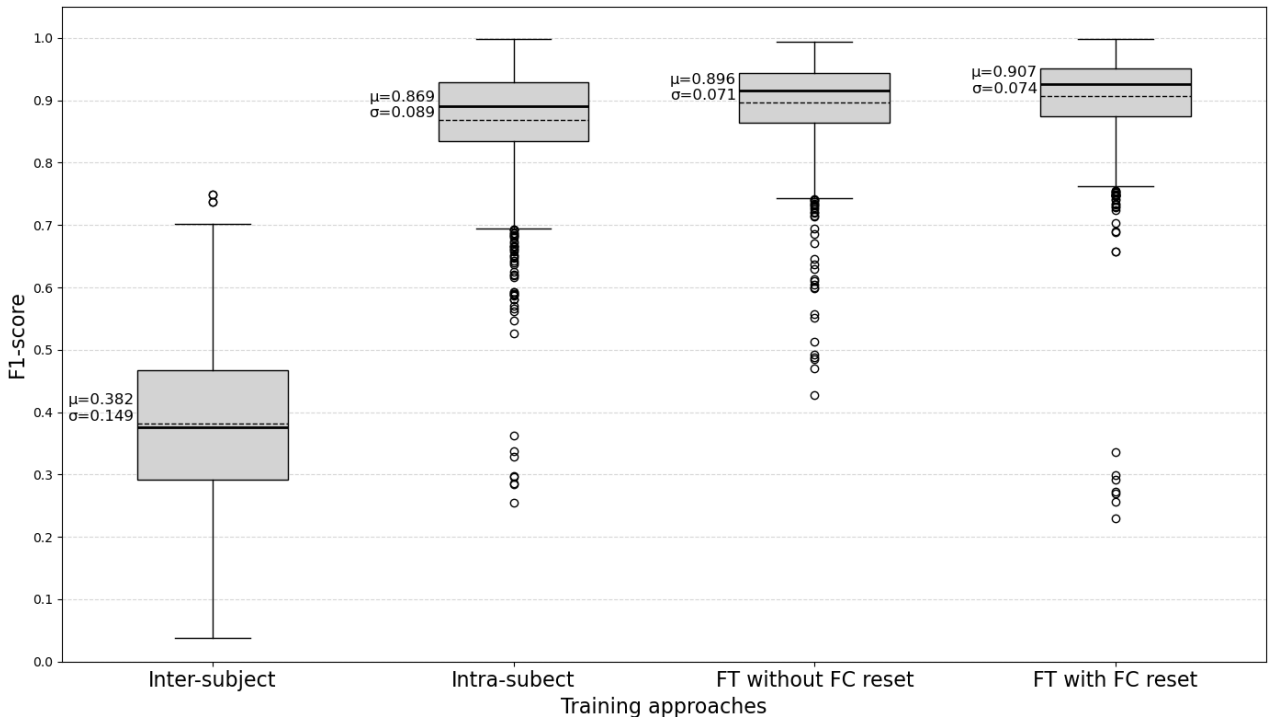


Fig. 2. Box-and-whisker plots of F1-score distributions across training approaches: “Inter-subject”, “Intra-subject” training approaches, and two fine-tuning (FT) strategies in transfer learning (TL) approach: with and without FC layer reset. Boxes represent interquartile range (IQR), lines show medians, and dashed lines indicate means. Outliers are shown as individual points.

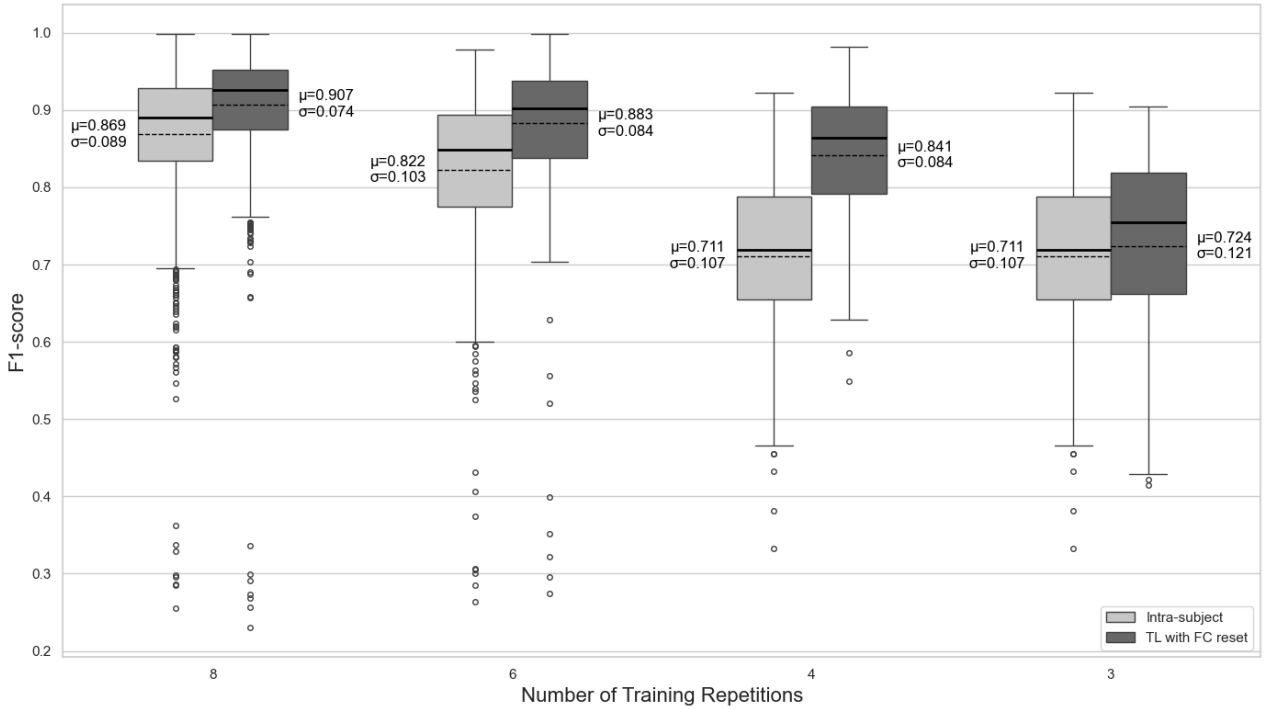


Fig. 3. Box-and-whisker plots showing the distribution of F1-scores across different numbers of training repetitions (6, 4, 2, 1) for “Intra-subject” and “Transfer learning” (TL) approach with FC layer reset. Note that repetitions numbers here represent a total number of repetitions consisting of eleven gestures a subject would need to execute to achieve shown accuracy. Each box shows the interquartile range and median; dashed lines denote the mean. Outliers are plotted as individual points.

### Statistical significance of transfer learning improvements

To assess whether the observed differences in model performance are statistically meaningful, Wilcoxon signed-rank tests are applied to compare subject-wise F1-score means ( $\mu$ ) and standard deviations ( $\sigma$ ). Table 1 summarizes the results for all training approaches and repetition settings evaluated in this study.

When using all eight repetitions for training, transfer learning without FC reset yields a statistically significant improvement in mean F1-score over intra-subject training ( $p = 7.87 \times 10^{-6}$ ), although it does not significantly improve model stability ( $p = 0.0829$ ). In contrast, transfer learning with FC reset results in both a significant increase in classification accuracy ( $p = 2.38 \times 10^{-7}$ ) and a notable reduction in F1-score variability ( $p = 0.0016$ ), indicating more consistent performance across cross-validation folds. Moreover, when directly comparing the two transfer learning strategies, resetting the FC layer leads to significantly higher mean F1-scores ( $p = 3.46 \times 10^{-4}$ ), though the difference in standard deviations remains statistically insignificant ( $p = 0.118$ ).

Fine-tuning with FC reset consistently maintains its advantage across all tested repetition counts. Even when we reduce the training data to six, four, or three repetitions, the mean F1-score stays significantly higher than that of intra-subject models trained under the same constraints. In each scenario, Wilcoxon tests confirm not only improved accuracy but also significantly lower variability in model performance (all  $p < 0.05$ ), even under data-limited conditions.

The results of all training experiments, as well as the Python scripts used for statistical evaluation, are available in the publicly accessible repository [21].

Overall, these findings confirm that transfer learning strategies offer statistically robust improvements in terms of gesture classification accuracy. Among the evaluated strategies, fine-tuning with FC reset provides statistically better accuracy and stability, regardless of how much subject-specific data is available. The ability to retain these advantages under reduced-data conditions supports its suitability for real-world deployment, especially in real-time or wearable applications where minimizing calibration effort is essential.

Tabl. 1 Wilcoxon signed-rank test results comparing subject-wise F1-score means ( $\mu$ ) and standard deviations ( $\sigma$ ) across training strategies and training set sizes.  $H_0$  represents null hypothesis, while  $H_1$  – alternative hypothesis. TL+FC: Transfer Learning with FC reset; TL-FC: Transfer Learning without FC reset; Intra: Intra-subject training. The number of training repetitions used in each experiment is denoted by  $r$ . Bold horizontal lines separate hypothesis groups. Statistically significant results ( $p < 0.05$ ) indicate a meaningful difference in accuracy or model stability.

Hypothesis	p-value
$H_0 : \mu (F1_{TL-FC, r=8}) = \mu (F1_{Intra, r=8})$ $H_1 : \mu (F1_{TL-FC, r=8}) > \mu (F1_{Intra, r=8})$	$p = 7.87 \times 10^{-6} < 0.05$
$H_0 : \sigma (F1_{TL-FC, r=8}) = \sigma (F1_{Intra, r=8})$ $H_1 : \sigma (F1_{TL-FC, r=8}) < \sigma (F1_{Intra, r=8})$	$p = 0.0829 > 0.05$
$H_0 : \mu (F1_{TL+FC, r=8}) = \mu (F1_{Intra, r=8})$ $H_1 : \mu (F1_{TL+FC, r=8}) > \mu (F1_{Intra, r=8})$	$p = 2.38 \times 10^{-7} < 0.05$
$H_0 : \sigma (F1_{TL+FC, r=8}) = \sigma (F1_{Intra, r=8})$ $H_1 : \sigma (F1_{TL+FC, r=8}) < \sigma (F1_{Intra, r=8})$	$p = 0.0016 < 0.05$
$H_0 : \mu (F1_{TL+FC, r=8}) = \mu (F1_{TL-FC, r=8})$ $H_1 : \mu (F1_{TL+FC, r=8}) > \mu (F1_{TL-FC, r=8})$	$p = 3.46 \times 10^{-4} < 0.05$
$H_0 : \sigma (F1_{TL+FC, r=8}) = \sigma (F1_{TL-FC, r=8})$ $H_1 : \sigma (F1_{TL+FC, r=8}) < \sigma (F1_{TL-FC, r=8})$	$p = 0.118 > 0.05$
$H_0 : \mu (F1_{TL+FC, r=6}) = \mu (F1_{Intra, r=6})$ $H_1 : \mu (F1_{TL+FC, r=6}) > \mu (F1_{Intra, r=6})$	$p = 2 \times 10^{-5} < 0.05$
$H_0 : \sigma (F1_{TL+FC, r=6}) = \sigma (F1_{Intra, r=6})$ $H_1 : \sigma (F1_{TL+FC, r=6}) < \sigma (F1_{Intra, r=6})$	$p = 0.033 < 0.05$
$H_0 : \mu (F1_{TL+FC, r=4}) = \mu (F1_{Intra, r=4})$ $H_1 : \mu (F1_{TL+FC, r=4}) > \mu (F1_{Intra, r=4})$	$p = 2 \times 10^{-5} < 0.05$
$H_0 : \sigma (F1_{TL+FC, r=4}) = \sigma (F1_{Intra, r=4})$ $H_1 : \sigma (F1_{TL+FC, r=4}) < \sigma (F1_{Intra, r=4})$	$p = 0.0018 < 0.05$
$H_0 : \mu (F1_{TL+FC, r=3}) = \mu (F1_{Intra, r=3})$ $H_1 : \mu (F1_{TL+FC, r=3}) > \mu (F1_{Intra, r=3})$	$p = 2 \times 10^{-5} < 0.05$
$H_0 : \sigma (F1_{TL+FC, r=3}) = \sigma (F1_{Intra, r=3})$ $H_1 : \sigma (F1_{TL+FC, r=3}) < \sigma (F1_{Intra, r=3})$	$p = 0.0473 < 0.05$

## Limitations, Comparison to Related Work, and Future Directions

While our experiments clearly show that transfer learning improves model accuracy and stability, there are some limitations to consider.

First, the CNN model we used is relatively simple. We chose this to keep training fast and make our results easier to compare with past research [16, 17]. However, more advanced models like temporal convolutional networks [2] or CNN-LSTM combinations [9] may be able to improve accuracy further.

Second, the 3DC dataset is designed to simulate real-world changes in electrode placement, which supports testing model generalization. However, this variability also introduces noise into the data, potentially lowering accuracy compared to more controlled datasets.

Third, this study conducts all experiments offline. To determine whether the models perform effectively in real-world use, future research should evaluate them under real-time conditions, such as with wearable systems or interactive applications [13].

Compared to earlier works such as Côté-Allard et al. [11], which demonstrate the benefits of transfer learning for sEMG, our study goes further by comparing two fine-tuning strategies and evaluating their statistical significance. While other studies [12, 15] explore transfer learning under electrode shift conditions, few measure how much subject-specific data can be reduced. This is the key contribution of our work.

In the future, it would be useful to test more complex models, add domain adaptation techniques, and use fewer labeled samples. Testing the models on small devices like microcontrollers would also show how well they work in real applications. Public benchmarks on larger and more varied datasets would also help to

confirm how well these results generalize. The source code and evaluation results for this study available publicly [21] can help other researchers build on this work.

## Conclusion

This study investigated the effectiveness of transfer learning for sEMG gesture classification using CNN. By comparing intra-subject training, inter-subject generalization, and two transfer learning strategies (with and without resetting the FC layer), we demonstrate that transfer learning significantly enhances classification performance and model stability.

Our experiments show that fine-tuning pre-trained models, especially when resetting the FC layer, not only improves F1-scores but also reduces standard deviation across cross-validation folds, indicating more consistent performance. Notably, this benefit remains statistically significant even if to reduce the number of repetitions used for subject-specific fine-tuning, suggesting that transfer learning can substantially lower the effort required for user calibration.

In addition, Wilcoxon signed-rank test results confirm the superiority of fine-tuned models under various data constraints. These findings emphasize the practicality of transfer learning in real-world applications where data availability remains limited.

Nevertheless, the study faces several limitations. The CNN architecture remains relatively simple and might fail to capture more complex spatiotemporal dependencies in the sEMG signal. Additionally, the evaluation takes place in an offline setting and relies on a single dataset (3DC), which, despite its realistic variability, may not generalize to all sEMG acquisition systems or application environments.

Future works could explore deeper and hybrid models, real-time implementations, and domain adaptation techniques to further reduce calibration overhead. Expanding evaluation to diverse datasets and hardware platforms will be critical to validating these results at scale.

In summary, transfer learning, particularly fine-tuning with FC reset, proves to be an effective strategy for subject adaptation in sEMG-based gesture recognition. It enhances both performance and model stability while reducing the amount of subject-specific training data required.

## References

- [1] Al-Ayyad, M., Owida, H. A., De Fazio, R., Al-Naami, B., & Visconti, P. (2023). Electromyography Monitoring Systems in Rehabilitation: A Review of Clinical Applications, Wearable Devices and Signal Acquisition Methodologies. *Electronics*, Vol. 12, Iss. 7, 1520, doi:10.3390/electronics12071520.
- [2] Xie, P., Xu, M., Shen, T., Chen, J., Jiang, G., Xiao, J., & Chen, X. (2024). A channel-fused gated temporal convolutional network for EMG-based gesture recognition. *Biomedical Signal Processing and Control*, Vol. 95, Part A, 106408, doi:10.1016/j.bspc.2024.106408.
- [3] Merletti, R. & Farina, D. (2020). Tutorial. Surface EMG detection, conditioning and pre-processing: Best practices. *Journal of Electromyography and Kinesiology*, Vol. 54, 102440, doi:10.1016/j.jelekin.2020.102440.
- [4] Where should I place my Electromyography (EMG) electrodes? *Plux Biosignals*. (2022). Accessed: 25 May 2025.
- [5] Chowdhury, R. H., Reaz, M. B. I., Ali, M. A. B. M., Bakar, A. A. A., Chellappan, K., & Chang, T. G. (2013). Surface Electromyography Signal Processing and Classification Techniques. *Sensors*, Vol. 13, Iss. 9, pp. 12431–12466, doi:10.3390/s130912431.
- [6] Ojha, A. (2023). An Introduction to Electromyography Signal Processing and Machine Learning for Pattern Recognition: A Brief Overview. *Extsv. Rev.*, Vol. 3, No. 1, pp. 24–37, doi:10.21467/exr.3.1.8382.
- [7] Moctar, S. M. S.-E., Rida, I., & Boudaoud, S. (2024). Comprehensive Review of Feature Extraction Techniques for sEMG Signal Classification: From Handcrafted Features to Deep Learning Approaches. *IRBM*, Vol. 45, Iss. 6, 100866, doi:10.1016/j.irbm.2024.100866.
- [8] Phinyomark, A., Campbell, E., & Scheme, E. (2019). Surface Electromyography (EMG) Signal Processing, Classification, and Practical Considerations. In: G. Naik, ed., *Biomedical Signal Processing*, Singapore: Springer, pp. 3–29, doi:10.1007/978-981-13-9097-5\_1.
- [9] Faust, O., Hagiwara, Y., Hong, T.J., Lih, O.S. & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine*, Vol. 161, pp. 1–13. doi:10.1016/j.cmpb.2018.04.005.
- [10] Wu, D., Yang, J., & Sawan, M. (2022). Transfer Learning on Electromyography (EMG) Tasks: Approaches and Beyond. *arXiv*, doi:10.48550/arXiv.2210.06295.
- [11] Côté-Allard, U., Fall, C. L., Campeau-Lecours, A., et al. (2019). Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 27, Iss. 4, pp. 760–771, doi:10.1109/TNSRE.2019.2896269.
- [12] Ameri, A., Akhaee, M. A., Scheme, E. & Englehart, K. B. (2020). A Deep Transfer Learning Approach to Reducing the Effect of Electrode Shift in EMG Pattern Recognition-Based Control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 28, Iss. 2, pp. 370–379, doi:10.1109/TNSRE.2019.296218.
- [13] Shen, S., Gu, K., Chen, X.-R., Lv, C.-X., & Wang, R.-C. (2020). Gesture Recognition Through sEMG with Wearable Device Based on Deep Learning. *Mobile Networks and Applications*, Vol. 25, pp. 2447–2458, doi:10.1007/s11036-020-01590-8.
- [14] Geng, Y., Zhou, P. & Li, G. (2012). Toward attenuating the impact of arm positions on electromyography pattern-recognition based motion classification in transradial amputees. *Journal of NeuroEngineering and Rehabilitation*, Vol. 9, Art. 74, doi:10.1186/1743-0003-9-74.
- [15] Lehmler, S. J., Saif-ur-Rehman, M., Glasmachers, T. & Iosifidis, I. (2021). Deep Transfer-Learning for patient specific model re-calibration: Application to sEMG-Classification. *arXiv*, doi:10.48550/arXiv.2112.15019.

- [16] Côté-Allard U., Gagnon-Turcotte G., Laviolette F., Gosselin B. (2019). A Low-Cost, Wireless, 3-D-Printed Custom Armband for sEMG Hand Gesture Recognition. *Sensors*, Vol. 19, Iss. 12, 2811, doi:10.3390/s19122811.
- [17] Eddy, E., Campbell, E., Phinyomark, A., Bateman, S., Scheme, E. (2023). LibEMG: An Open Source Library to Facilitate the Exploration of Myoelectric Control. *IEEE Access*, Vol. 11, pp. 87380-87397, doi:10.1109/ACCESS.2023.3304544.
- [18] Smith, L. H., Hargrove, L. J., Lock, B. A. and Kuiken, T. A. (2011). Determining the Optimal Window Length for Pattern Recognition-Based Myoelectric Control: Balancing the Competing Effects of Classification Error and Controller Delay. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 19, Iss. 2, pp. 186-192, doi:10.1109/TNSRE.2010.2100828.
- [19] Rehman, M. Z. U., Waris, A., Gilani, S. O., Jochumsen, M., Niazi, I. K., et al. (2018). Multiday EMG-Based Classification of Hand Motions With Deep Learning Techniques. *Sensors*, Vol. 18, Iss. 8, 2497, doi:10.3390/s18082497.
- [20] Oskoei, M. A. and Hu, H. (2007). Myoelectric control systems—A survey. *Biomedical Signal Processing and Control*, Vol. 2, Iss. 4, pp. 275–294, doi:10.1016/j.bspc.2007.07.009.
- [21] Kolomiets, B. Yu. (2025). \*Code used for experiments presented in this article\* [online]. *GitHub*, accessed May, 2025.

## Покращення класифікації ЕМГ-сигналів за допомогою трансферного навчання в умовах обмежених даних та варіативності між користувачами

Коломієць Б. Ю., Карпліук Є. С.

Поверхнева електроміографія (ЕМГ) — це неінвазивний метод моніторингу м'язової активності, який широко застосовується в реабілітації, протезуванні, асистивній робототехніці та системах взаємодії людини з комп'ютером. Однак її практичне застосування часто обмежується значною варіативністю

між користувачами та необхідністю налаштування моделей під кожного нового користувача. У цьому дослідженні розглядається застосування трансферного навчання до задачі класифікації жестів кисті та зап'ястя на основі методів глибокого навчання. Експерименти проводилися на датасеті, що містить одинадцять жестів, кожен з яких повторено вісім разів 22 здоровими учасниками. Було протестовано три підходи до навчання: (i) навчання та тестування на даних одного учасника (внутрішньоособове навчання), (ii) навчання на даних декількох учасників і тестування на новому (міжособове узагальнення), (iii) трансферне навчання з донавчанням, із скиданням та без скидання повнозв'язного вихідного шару згортової нейронної мережі. Усі моделі оцінювались за допомогою стратегії перехресної перевірки з виключенням одного елемента (leave-one-out) як по учасниках, так і по повтореннях.

Результати показали, що обидві стратегії трансферного навчання дали вищу точність класифікації порівняно з іншими двома підходами. Найкращого результату вдалося досягти при скиданні FC шару перед донавчанням (F1-оцінка = 0.907,  $\sigma = 0.074$ ). Статистичні тести Вілкоксона підтвердили, що ці покращення є статистично значущими навіть за використання лише кількох повторень для донавчання. Ба більше, трансферне навчання з чотирма повтореннями забезпечило подібну точність, як і навчання з нуля на восьми повтореннях.

Отримані результати свідчать, що донавчання попередньо натренованих моделей може суттєво зменшити зусилля, необхідні для адаптації ЕМГ-систем до нових користувачів, і є практичним та ефективним підходом для створення інтерфейсів, орієнтованих на допоміжне та реабілітаційне застосування.

*Ключові слова:* поверхнева електроміографія (ЕМГ); розпізнавання рухів; трансферне навчання; глибоке навчання; згортові нейронні мережі; міжіндивідуальна варіативність; узагальнення моделей; стратегії донавчання; персоналізація моделей; міоелектричне керування; ЕМГ-інтерфейс; перехресна перевірка (cross-validation); обробка біомедичних сигналів; реабілітаційні технології; спрощене калібрування