

Transfer Learning Strategies for Leveraging Multi-Session and Multi-Subject Datasets to Improve sEMG Recognition

N. Sicard

Dept. of Electrical and Computer Engineering
Université du Québec à Trois-Rivières
Trois-Rivières, CANADA
nicolas.sicard@uqtr.ca

F. Nougrou

Dept. of Electrical and Computer Engineering
Université du Québec à Trois-Rivières
Trois-Rivières, CANADA
Francois.Nougrou@uqtr.ca

Abstract—Building movement detection models based on surface electromyography (sEMG) requires extensive data collection, which is time-consuming due to inter-session and inter-subject variability in the EMG signals. In fact, sensor placement can differ across sessions, and muscle physiology and user experience vary between subjects. The model-based approach of transfer learning offers a solution by transferring knowledge from an already acquired dataset into a model and deriving this pre-trained model to achieve good performance for a new session or a new subject with fewer data. However, how can this approach effectively exploit datasets from multiple sessions and multiple subjects? Compatibility issues arise between some pre-trained models and data from different sessions or subjects, affecting prediction accuracy. This study addresses these challenges and proposes strategies for combining and selecting transfer learning models from diverse datasets that significantly improve the prediction of eight forearm gestures with minimal training data. The proposed strategies produce promising results, with the combination strategy exceeding 90% precision and others between 85 and 90%.

Keywords—EMG, gesture recognition, transfer learning strategies, inter-session, inter-subject.

I. INTRODUCTION

Surface electro-myographic (sEMG) data presents high variability in the case of inter-session and inter-subject signals. The signals depend on the sensor placement relative to the arm and contact with the subject's skin and the individual muscle physiology that differs between subjects. This high variability makes training models for applications using these signals very long and straining on the subject. The use of transfer learning (TL) approach has gained traction as an extra step in hand recognition based on sEMG signals to add depth to the process[1-10]. Model-based transfer learning leverages knowledge from a pre-existing model trained from an initial dataset, which is then adjusted to achieve good performance for a new session or a new subject with less data. This improves the ease of model training, because the model only needs fine-tuning. This approach has been used within a variety of different models in sEMG recognition such as simple Neural Networks (NN) [1], Convolutional Neural

Networks (CNN) [2-7], CNN combined with Long Short-Term Memory (LSTM) [8], Domain Generalization (DG) [9], and TL model ensemble decisional trees [10]. These studies show that recognition methods without transfer learning (TL) achieve high precision, and the TL improvement is often marginal but with fewer learning data[1-10].

However, in the presence of several datasets from multiple sessions and multiple subjects, the strategies for exploiting these datasets using the TL approach are not trivial. Indeed, new data to be interpreted may not be compatible with some pre-trained models and this may affect the performance of the TL approach. This study highlights these compatibility challenges and explores strategies for combining and selecting multiple TL models from multiple datasets to improve motion prediction based on sEMG signals.

To do this, we conducted preliminary experiments in which EMG data from 8 sensors placed around the forearm were recorded on two subjects over several sessions with different electrode positions to detect 8 hand gestures. All these data enabled us (i) to evaluate the behavior of the pre-trained transfer learning models under different inter-session and inter-subject detection conditions, (ii) to observe the variability in the performance of these models and (iii) to propose transfer learning strategies for effectively exploiting multiple datasets in order to improve the gesture prediction while using a reduced number of training sEMG data.

II. MATERIALS AND METHODS

A. Data Acquisition and Features Extraction

The sEMG data is recorded using the MyoArmband® (Thalmic Labs, Canada), which consists of a bracelet with

TABLE I: Recorded hand gestures

Label	Image	Name	Label	Image	Name
M1		Rest	M5		Wrist ulnar flexion
M2		Hand closed	M6		Wrist extension
M3		Hand opened	M7		Wrist radial flexion
M4		Wrist flexion and fingers extension	M8		Thumb, index and little extension

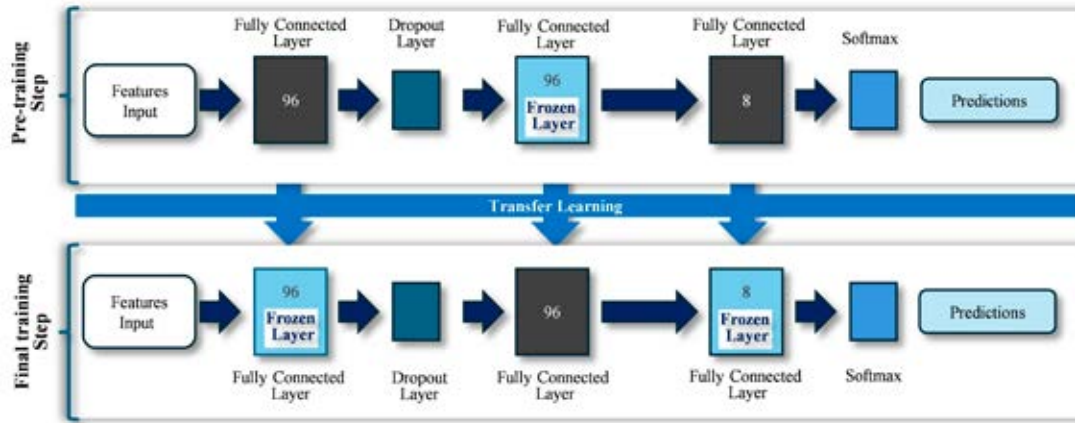


Fig. 1: Illustration of the training of the pre-trained model, 8 electrodes that measure muscle activity around the forearm at a sampling rate of 200 Hz. Eight hand and wrist gestures (Table I) were recorded for training and testing using a custom Matlab/Python platform for target setting, sensor connection, and data collection. Two subjects were involved in this study to conduct a preliminary investigation on transfer learning based on multi-subject and multi-session datasets. Subject 1 completed three data collection sessions and Subject 2 completed four sessions. The bracelet is rotated after each session. Rotating the bracelet changes the electrode placement on the arm, simulating the variations that can occur between usage sessions of a given subject. Finally, a data collection session involves recording the sequences of each hand gesture three times, each lasting 30 seconds. A total of 48 size features were extracted from the EMG channels within a 100ms window, with a 50ms overlap: Mean Absolute Value (MAV), Modified MAV (MMAV), MAV Type 2 (MMAV2), Root Mean Square (RMS), Waveform Length (WL), and Mean Absolute Value Slope (MAVSLP) [11].

B. Training phases of the model-based transfer learning

1) *Pre-training step*: As shown in Fig. 1, the initial model is a neural network with three fully connected (FC) layers (96, 96, and 8 neurons), dropout, softmax, and cross-entropy. The second FC layer is frozen in the pre-training step and will only learn in the final training step. The purpose of freezing that layer is to teach the first and third layers to classify the position independent from the second layer. This pre-training step uses all data from the subject and session with a high number of epochs, allowing the model to capture data variability and

the knowledge transfer and the training of the final model specialize in hand gesture classification. Pre-training was done for all positions and both participants to obtain a model for each training condition.

2) *Final training step*: In model-based transfer learning, the pre-trained model serves as the initial state. The first and third FC layers are frozen with pre-trained weights, while the second FC layer is active (see Figure 1). The training dataset differs in bracelet position and/or subject from the pre-training dataset. These datasets can be considered as belonging to two different domains. The model is retrained to map features and predict based on the input domain. The goal is to train the second FC layer to classify hand gestures despite domain changes, using minimal data. The resulting model is called the "TL model." The final training used all pre-trained models to create TL models for each domain and test inter-session and inter-subject TL.

C. Proposed combination and Selection strategies

With pre-trained models from multiple sessions and subjects, compatibility with new domain data varies. Some TL models perform well, others match no TL. Therefore, with multiple datasets, the strategies for using the TL approach are not trivial. Two strategies are explored for better prediction: combining and pre-selecting TL models (see Figure 2).

1) *Combination strategy*: Figure 2 (a) presents the proposed combination strategy based on the average of soft labels. For each sample, this approach averages the soft labels from all TL models and selects the gesture with the highest score. This technique uses maximum information during testing to ensure accuracy without selecting TL models. However, it requires

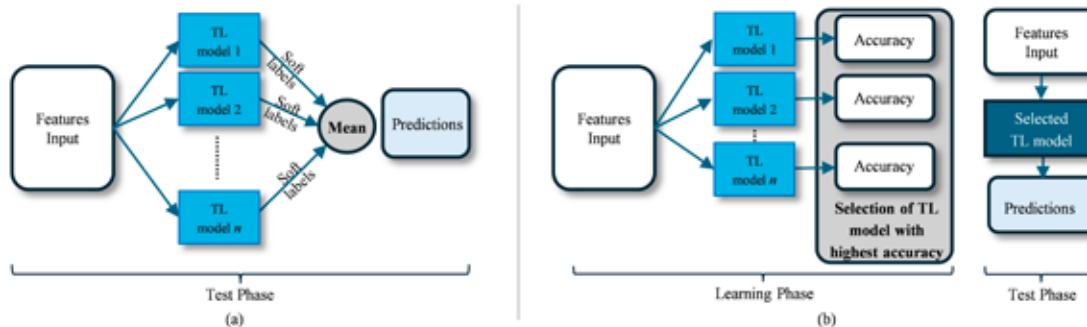


Fig. 2: Illustration of the proposed (a) combination strategy based on soft label average and the (b) selection strategy.

processing all TL models, with complexity increasing as their number grows.

2) *Selection strategy*: As shown in Figure 2 (b), during the learning phase, model selection considers all trained TL models and evaluates their accuracy using data that were not used for training. TL model accuracies are compared, and the highest-performing model is selected, which has the advantage of reducing computational complexity of the testing phase.

3) *Selection & combination strategy*: This approach combines the proposed selection and combination strategies, offering a compromise between computational efficiency and the amount of information used. Here, the two highest-performing models are first selected and then combined. This solution slightly increases the system's complexity but also enhances the amount of information leveraged for prediction. Additionally, it is possible to increase the number of TL models selected for averaging, further improving predictive accuracy at the cost of added complexity.

III. RESULTS

A. Behavior of TL Models

As shown in Figure 3, the proposed strategies were evaluated using 5 to 25 training samples per gesture. All models aim to detect the 8 gestures from Table I using EMG data from a domain different from pre-training. The datasets used are detailed in Table II. Several observations can be made from Figure 3. First, the TL models, pre-trained on data from different domains, exhibit distinct behaviors. One model performs similarly to the model without TL (Model 6), most of them achieve better performance, and one model stands out significantly (Model 5). This indicates that TL models behave differently due to varying levels of compatibility with the new domain data. This variability in TL model behavior highlights the importance of developing strategies to effectively leverage multiple TL models. The model without TL underperforms compared to the TL models, and even with just 5 training

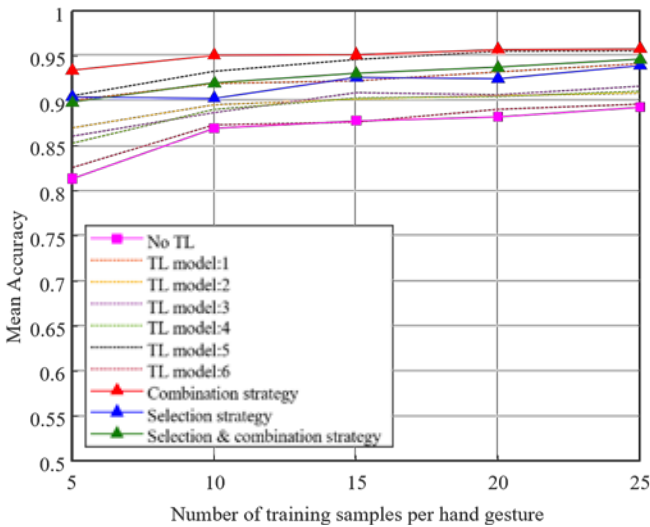


Fig. 3: Precision performances for available TL models and proposed strategies with varying training sample sizes.

TABLE II: TL data configurations used in Figure 3 for given participants and electrode position

TL Model	Pre-trained (Pos., Part.)	TL (Pos., Part.)
TL Model 1	Pos. 1, Part. 1	Pos. 2, Part. 2
TL Model 2	Pos. 2, Part. 1	Pos. 2, Part. 2
TL Model 3	Pos. 3, Part. 1	Pos. 2, Part. 2
TL Model 4	Pos. 1, Part. 2	Pos. 2, Part. 2
TL Model 5	Pos. 3, Part. 2	Pos. 2, Part. 2
TL Model 6	Pos. 4, Part. 2	Pos. 2, Part. 2

samples per movement, the models can achieve over 90% accuracy to detect all gestures. In this scenario, the strategy combining all 6 models shows strong potential, reaching nearly 95% accuracy with just 5 samples per gesture and improving slightly with more data. The model selection strategy also delivers satisfactory results, the accuracy at 5 samples per movement is of 90%. Finally combining both strategies, which requires only two TL models, balances performance and complexity, showing accuracies between the other two strategies.

B. Evaluation of the combination and selection strategies

Accuracy results are presented in Figure 4 across three scenarios for a given subject x : inter-session (Cond.Ax), separated inter-subject for each of the 2 subjects (Cond.Bx), and inter-subject and inter-session combined (Cond.Cx). Subject 1 is inexperienced, while Subject 2 is experienced in performing the gestures to acquire forearm EMG signals. To highlight the benefits of the TL approach with minimal training data, all simulations used only 5 samples per movement. The average precision, recall and F1-Score are also evaluated in Table III. Finally, Table IV presents the accuracy for each individual hand gesture. All results are averages over 20 cross-validation repetitions using all available TL model conditions. In each scenario, a reference model was trained from scratch using the same training and testing data as the TL models. The reference model for Subject 1 has an average accuracy of 70%, while Subject 2 reached 77%. As shown in Figure 4, across all tested training conditions, the proposed strategies consistently outperformed the reference model by approximately 10%.

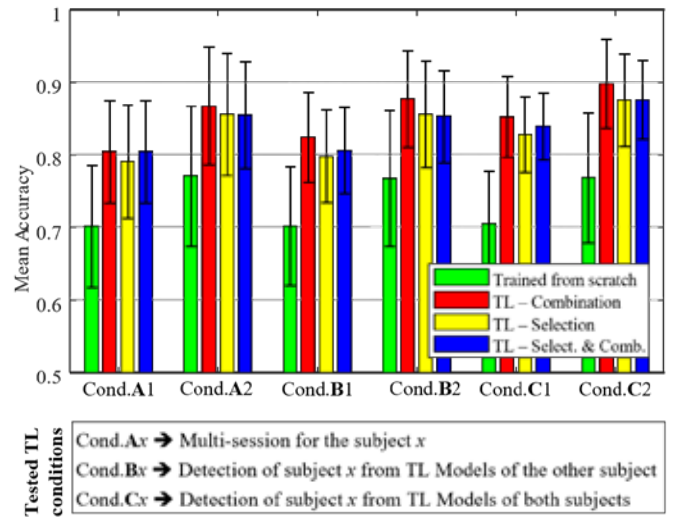


Fig. 4: Average accuracies of the model without TL and the proposed TL strategies in different TL conditions.

TABLE III: Performance metrics for CondC1 and CondC2 over 20 repetitions

Strategy	Precision	F1-Score	Recall
<i>CondC1</i>			
Scratch	69.36 ± 7.50	66.78 ± 8.81	70.96 ± 7.76
Selection	78.99 ± 6.58	76.85 ± 7.21	80.25 ± 6.18
Select. & Comb.	80.62 ± 5.80	78.73 ± 6.82	82.46 ± 4.88
Combination	83.79 ± 6.33	82.21 ± 7.61	84.70 ± 5.54
<i>CondC2</i>			
Scratch	80.08 ± 6.65	78.88 ± 6.91	82.57 ± 5.98
Selection	88.82 ± 5.77	88.16 ± 6.31	90.27 ± 5.32
Select. & Comb.	88.45 ± 5.12	87.70 ± 5.58	90.03 ± 4.56
Combination	92.03 ± 4.77	91.66 ± 5.07	93.21 ± 4.11

For both subjects, the proposed strategies show similar behaviors. The combination strategy shown in Figure 2 consistently provides the best performance under all conditions and training sample datasets. In the case of the inter-session with subject 1, the combination strategy and the selection & combination strategy give identical results, as only two TL models are available, meaning that both approaches use the same models. In most cases, the selection and combination strategy (section II.C.3) achieves the second best performance, and the selection strategy (section II.C.2) performs slightly worse. Thus, as previously mentioned, the combination strategy is the approach that makes the best use of the available TL models, although it requires all TL models during the testing phase. Selection-based strategies are slightly less effective but require fewer models for prediction.

When considering Cond. Ax and Cond. Bx, the proposed strategies effectively address the challenges of classification scenarios with minimal training data by leveraging the information of the available datasets. Note that Subject 2, being more experienced, consistently outperforms Subject 1. However, the results from Cond.B1 and Cond.B2 demonstrate effective knowledge transfer between subjects. For each subject, Cond.Cx yields the best results, where the strategies use information from both inter-session and inter-subject data. The proposed strategies perform better with more information, indicating their effectiveness and great potential for achieving higher performance with fewer training data.

Table III gives a better perspective of the performances of the proposed strategies in CondC1 and CondC2. The bold values are the highest for each performance metric in each condition. For each metric of each condition, the combination strategy has the highest performance. For Subject 1, the combination strategy achieves over 80% average accuracy across all metrics, while Subject 2 exceeds 90%. Notably, models trained from scratch perform significantly worse.

Table IV details the results per movement for CondC2. It shows the effectiveness of the proposed methods. By leveraging multiple datasets, the detection of movements that are more difficult for the reference model improves.

TABLE IV: Accuracy separated by class using CondC2 over 20 repetitions

Strategy	M1	M2	M3	M4	M5	M6	M7	M8
Scratch	79.88	90.10	79.20	78.35	90.65	85.60	66.31	83.35
Selection	98.41	93.72	88.65	87.45	95.96	91.89	79.04	96.58
Select. & Comb.	96.13	91.05	86.42	81.96	94.47	91.85	80.33	98.91
Combination	98.63	95.36	89.20	87.89	98.02	94.99	83.52	98.31

IV. CONCLUSION

The goal of this preliminary study is to propose transfer learning (TL) strategies to exploit multi-session and multi-subject datasets to improve the prediction of eight hand gestures using reduced surface electromyography (sEMG) data. Three TL strategies are presented and proposed based on combination of TL models, selection of TL models, and selection & combination of TL models. The combination strategy is the most complex but most efficient approach to leverage the information from the available TL models. Selection-based strategies are slightly less effective but are significantly less complex. These three approaches outperform a model that doesn't go through TL, and only necessitates minimal training data due to its utilization of the information from the existing pre-trained models. Finally, it has been shown that the more these strategies leverage information, the better they perform. Thus, future work will involve a more in-depth evaluation of these proposals with a more complex recognition structures and a larger number of subjects and inter-session data.

ACKNOWLEDGMENT

Research supported by Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

- [1] C. Andronache et al., "Towards extending real-time EMG-based gesture recognition system," 43rd International Conference on Telecommunications and Signal Processing (TSP), 2020, pp. 301-304.
- [2] U. Côté-Allard et al., "Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning," in IEEE Trans. on Neural Systems and Rehabilitation Engineering, vol. 27, no. 4, pp. 760-771, April 2019.
- [3] R. Deepa et al., "Revolutionizing Hand Gesture Recognition: A Transfer Learning Approach using Surface Electromyography and Convolutional Neural Networks," Conference on Augmented Intelligence and Sustainable Systems (ICAIS), 2023, pp. 70-75.
- [4] J. Lu, P. Liang et al., "EffiE: Efficient Convolutional Neural Network for Real-Time EMG Pattern Recognition System on Edge Devices," 11th International IEEE/EMBS Conference on Neural Engineering (NER), Baltimore, MD, USA, 2023, pp. 1-5.
- [5] S. S. Roy et al., "Hand Movement Recognition Using Cross Spectrum Image Analysis of EMG Signals-A Deep Learning Approach," Conference on Emerging Trends on Sustainable Technology and Engineering Applications (NCETSTE), 2020, pp. 1-5.
- [6] H. Shi, X. Jiang, C. Dai and W. Chen, "EMG-based Multi-User Hand Gesture Classification via Unsupervised Transfer Learning Using Unknown Calibration Gestures," in IEEE Trans. on Neural Systems and Rehabilitation Engineering, vol. 32, pp. 1119-1131, 2024.
- [7] C. Tang, W. Yi, J. M. Kim and L. G. Occhipinti, "EMG-Based Human Motion Analysis: A Novel Approach Using Towel Electrodes and Transfer Learning," IEEE SENSORS, 2023, pp. 1-4.
- [8] X. Chen et al., "Hand Gesture Recognition based on Surface Electromyography using Convolutional Neural Network with Transfer Learning Method," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 4, pp. 1292-1304, April 2021.
- [9] Y. Ye, Y. He, T. Pan, Q. Dong, J. Yuan and W. Zhou, "Cross-subject EMG hand gesture recognition based on dynamic domain generalization," 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2023, pp. 1-4.
- [10] D. Wu, J. Yang and M. Sawan, "Transfer Learning on Electromyography (EMG) Tasks: Approaches and Beyond," in IEEE Trans. on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 3015-3034, 2023.
- [11] A. Phinyomark et al., "Evaluation of EMG feature extraction for hand movement recognition based on Euclidean distance and standard deviation," in Proc. ECTI-CON, May 2010, pp. 856-860.