dynAmiC: Dynamic Domain Adaptation with Efficient Coreset Selection

Anonymous Author(s)
Affiliation
Address
email

Abstract

Building subject-independent models for electrocardiogram (ECG)-based emotion recognition is challenging due to substantial inter-subject variability and the high cost of utilizing large volumes of unlabeled data. While prior domain adaptation (DA) methods have mitigated distribution shifts, they typically rely on static alignment strategies and overlook data efficiency. In this work, we propose dynAmiC domain adaptation, a semi-supervised framework that dynamically balances Maximum Mean Discrepancy (MMD) and Local Structure Discriminative (LSD) losses to achieve effective global–local alignment. Furthermore, we introduce a coreset selection strategy that leverages only 1% of the unlabeled target data while delivering performance comparable to using the entire dataset. Extensive experiments on the DREAMER and WESAD benchmarks, evaluated under leave-one-subject-out cross-validation, demonstrate that our approach consistently outperforms state-of-the-art baselines. These findings highlight dynAmiC domain adaptation as a robust and data-efficient pathway toward practical, calibration-free affective computing systems.

1 Introduction

Emotion recognition (ER) is becoming an essential component of human–machine interaction, particularly in applications that demand real-time responses, such as monitoring a patient's mental health status [1]. Emotion recognition from physiological signals like electrocardiogram (ECG) is a promising yet challenging area of research. A primary obstacle is domain shift, a phenomenon where a model's performance degrades when applied to data from different subjects due to individual-specific physiological variations. This is a critical issue in cross-subject emotion recognition, as models trained on one population fail to generalize to new, unseen individuals. Traditional domain adaptation (DA) methods attempt to align marginal distributions, which can be counterproductive as it risks conflating the discriminative features of different emotion categories. We argue that a more effective strategy is to align domains at the level of specific emotion categories. By treating each emotion as a distinct subdomain, we can ensure that feature distributions are aligned for semantically similar data points, preserving the inter-class boundaries crucial for classification. Motivated by this insight, we propose a novel Dynamic Domain Adaptation (DDA) framework. Our method dynamically transitions from a global alignment phase, where it minimizes the Maximum Mean Discrepancy (MMD), to a fine-grained phase focused on minimizing the Local Subdomain Discrepancy (LSD). This progressive approach allows the model to first learn a robust, shared representation before fine-tuning it with a class-conditional alignment objective. As true labels are scarce in the target domain, our framework leverages pseudo-labels to estimate LSD, while also supporting a semi-supervised setting by incorporating a small number of true target labels.

A significant challenge in semi-supervised DA is the computational burden associated with processing large volumes of unlabeled data. To address this, we introduce an innovative coreset selection strategy

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

that reduces the amount of unlabeled data used for adaptation. Our method selects a small, highly informative subset of the target pool prior to the main adaptation stage. We design a hybrid scoring mechanism to select this coreset, combining three key signals for each data point: a diversity score (Euclidean distance to the feature-center), a difficulty proxy (negative similarity to the center), and a local-density term (mean distance to k-nearest neighbors). This approach ensures that the selected coreset is a representative and challenging subset of the data, maximizing the information gain while minimizing computational overhead.

The key contributions of this work are as follows: $i)$ We used a dynamic domain adaptation framework that progressively aligns source and target domains at the subdomain level, which is particularly suited for cross-subject emotion recognition for ECG. $ii)$ We integrate an efficient coreset selection strategy based on a novel hybrid scoring function to reduce computational costs and the reliance on large unlabeled datasets. $iii)$ We validate our framework on the two benchmark datasets, i.e., DREAMER and WESAD, demonstrating state-of-the-art performance .

2 Related Work

Emotion recognition from physiological signals, particularly EEG and ECG, has attracted growing interest due to applications in affective computing, healthcare, and human–computer interaction. The *DREAMER* database [6] established an important benchmark with multimodal EEG–ECG recordings and self-assessed emotion labels, demonstrating that wearable devices can achieve performance comparable to medical-grade systems. To address *inter-subject variability*, domain adaptation (DA) has been widely explored, especially in EEG-based studies [3]. However, most methods focus on global distribution alignment, overlooking class-specific discrepancies that weaken discriminative features. Dynamic domain adaptation approaches have since emerged to jointly minimize global and local shifts for improved cross-subject generalization. Compared with EEG, *ECG* offers a robust modality for emotion recognition due to its direct link to autonomic activity. Deep models such as DFF-STM [13] and ensemble RNNs [17] capture temporal and spatial ECG patterns, while attention mechanisms further highlight informative regions. Other strategies investigate segmentation [9] and multimodal fusion [18, 19] for richer representations. Despite these advances, ECG-based methods still struggle with *data scarcity* and *poor cross-dataset generalization*. Recently, unsupervised domain adaptation (UDA) and self-supervised learning (SSL) have been proposed to reduce reliance on labeled data and improve robustness [8, 12]. Transformer-based architectures and multimodal fusion have achieved strong results, but often require large-scale unlabeled datasets and lack mechanisms to efficiently exploit them.

In this work, we address these challenges by proposing *dynAmiC domain adaptation*, which dynamically balances MMD and LSD losses for effective global–local alignment, and introduces a novel coreset selection strategy that achieves comparable performance using only 1% of the unlabeled target data.

3 Methodology

3.1 Framework

We propose a semi-supervised domain adaptation pipeline for subject-independent ECG-based emotion recognition. The model combines a lightweight 1D-ResNet feature extractor with a shallow two-layer perceptron classifier. Given an input segment $x \in \mathbb{R}^{1\times T}$, the network outputs a 64-dimensional embedding $f = \phi(x)$ and logits $z = \psi(f) \in \mathbb{R}^{C}$. The backbone consists of three residual blocks with convolution, batch normalization, ReLU activations, and shortcut connections, followed by pooling and a linear projection to the embedding space.

Training couples supervised source-domain classification with an alignment loss to reduce distribution shift between source and target subjects. To minimize labeling and computation, we introduce a *coreset* selection strategy that adapts using only 1% of unlabeled target samples.

Evaluation follows a leave-one-subject-out (LOSO) protocol: in each fold, one subject is treated as the target and the remaining 22 as sources. Target data is split into 25% few-shot supervision, 25% unlabeled data for adaptation, and 50% held-out for testing. Pseudo-label alignment uses only

Hyperparameters	Values
Optimizer	Adam
LR Scheduler	Cosine Annealing
Patience (Early stopping)	10
Epochs	100
ECG Segment length	10 sec (50% ovelap)
Learning rate	10^{-4}
Weight decay	10^{5}
Sigma for Gaussian kernel (σ)	10
Batch-size	128

Table 1: Hyperparameters of the proposed model

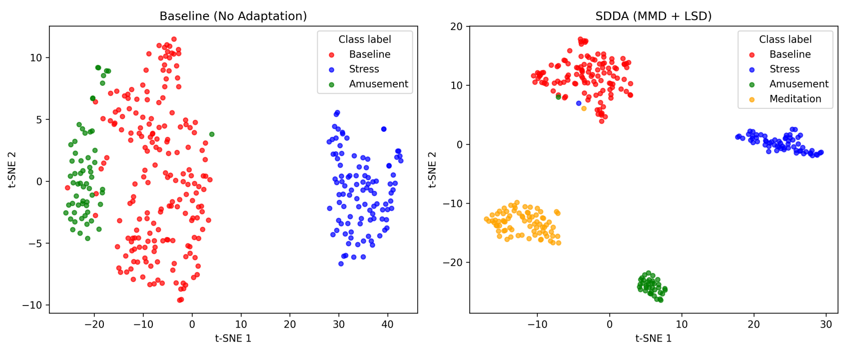

Figure 1: t-SNE plot representing the feature space for baseline (no DDA) and with DDA. *(It can be seen that after applying DDA approach, model is able to clearly distinguish all four clusters for four emotional states).*

confident predictions ($\tau = 0.9$). Performance is reported per subject using accuracy and macro-F1, and averaged across all 23 folds.

4 Experimental setup

4.1 Dataset and Preprocessing

This work utilizes the DREAMER dataset, a multimodal database comprising EEG and ECG signals recorded from 23 participants exposed to audio–visual stimuli in the form of film clips. In total, 18 clips were employed, with two clips selected for each of the nine target emotional categories: *amusement, excitement, happiness, calmness, anger, disgust, fear, sadness, and surprise*. EEG and ECG signals were acquired wirelessly using portable devices, namely the Emotiv EPOC EEG headset and the SHIMMER ECG sensor. The ECG recordings were sampled at 256 Hz, with the sensor supporting both $RA \rightarrow LL$ and $LA \rightarrow LL$ lead configurations. In this work, we considered only the ECG modality, and specifically utilized the $RA \rightarrow LL$ lead for feature extraction.

The preprocessing pipeline consisted of three main steps. (a) Filtering: ECG signals were bandpass filtered between 0.5 and 40 Hz to suppress baseline wander and high-frequency noise. (b) Normalization: Trial-wise z-score normalization was applied to mitigate inter-subject variability. (c) Segmentation: The normalized signals were divided into 10-second windows with a 50% overlap to generate fixed-length input segments. For labeling, the three emotional dimensions i.e., valence, arousal, and dominance - were binarized in accordance with prior literature: scores greater than 3 were categorized as *high*, whereas scores lower than 3 were categorized as *low*.

5 Results and Conclusion

This study addresses the challenge of inter-subject variability in ECG-based emotion recognition by reducing the discrepancy between source and target domains from a subdomain perspective. We used Dynamic Domain Adaptation (DDA) framework that goes beyond global distribution alignment and explicitly considers local divergences across emotion categories. By dynamically aligning both global and category-specific distributions, DDA facilitates more fine-grained knowledge transfer

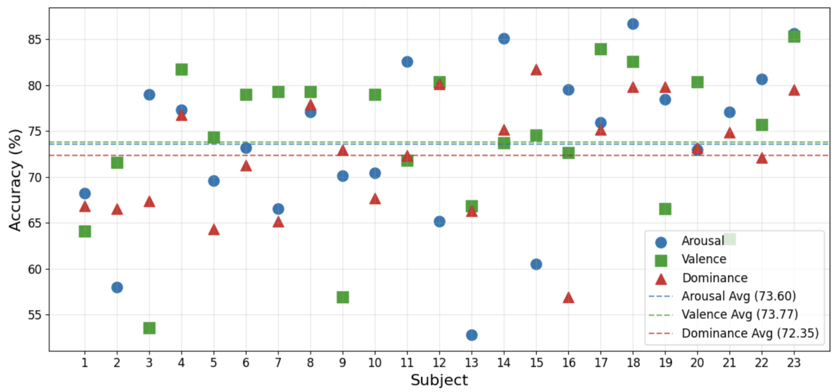

Figure 2: Classification accuracy for dominance, arousal and valence across each of the 23 subjects evaluated in a LOSO fashion. *(The average performance for each of the three classes is shown by the dotted lines).*

Table 2: Comparisons of proposed method with existing works in terms of accuracy and F1-score (mean(std.))on the WESAD dataset for 2-class *(stress Vs non-stress)*, 3-class *(Baseline, amusement and stress)*, and 4-class *(Baseline, amusement, meditation and stress)*; $SD : Subject - dependent; SI : Subject - independent$

Study	Method	Approach	2-class		3-class		4-class	
			Acc. (%)	F1 (%)	Acc. (%)	F1 (%)	Acc. (%)	F1 (%)
Sarkar et al. [12]	Self- Supervised	SD	–	–	–	–	99.4(0.004)	98.3(0.007)
Schmidt et al. [14]	LDA	SI	85.44	81.31	66.29	56.03	–	–
Li et al. [15]		SI	–	–	67.65 (13.48)	43.05 (17.20)	–	–
Abd et al.[16]	MLP	SD			90.2	90.00	–	–
Our Approach (MMD+LSD)		SI	89.12 (8.91)	87.10 (12.01)	73.77 (8.51)	69.78 (11.97)	86.5 (7.9)	84.64 (9.4)
Our Approach (MMD)		SI	90.14 (5.67)	88.35 (6.44)	72.93 (8.58)	67.72 (12.85)	72.60 (8.10)	66.99 (10.87)
Our Approach (LSD)		SI	86.4 (8.95)	84.50 (12.42)	73.14 (8.85)	68.72 (11.61)	71.45 (7.80)	65.57 (9.85)

between domains. Extensive experiments show that this strategy consistently improves generalization to unseen subjects, yielding substantial gains over state-of-the-art domain adaptation methods. These results highlight the importance of fine-grained alignment in mitigating domain shift and demonstrate the effectiveness of DDA in advancing robust cross-subject emotion recognition.

References

[1] Regan L Mandryk, Kori M Inkpen, and Thomas W Calvert, “Using psychophysiological techniques to measure user experience with entertainment technologies,” Behaviour & information technology, vol.25, no. 2, pp. 141–158, 2006.

[2] T. D. T. Mai, D. -Q. Vu, H. N. Phuong and T. -N. Phung, "EEG-SKDNet: A Self-Knowledge Distillation Model With Scaled Weights for Emotion Recognition From EEG Signals," in IEEE Access, vol. 13, pp. 140239-140250, 2025, doi: 10.1109/ACCESS.2025.3594671.

[3] Z. Li, E. Zhu, M. Jin, C. Fan, H. He, T. Cai, and J. Li. Dynamic domain adaptation for class-aware cross-subject and cross-session EEG emotion recognition. IEEE Journal of Biomedical and Health Informatics,

Table 3: Comparisons of proposed method with existing works on the DREAMER dataset for Arousal, Valence and Dominance $SD : Subject - dependent; SI : Subject - independent$

Study	Method	Approach	Arousal		Valence		Dominance	
			Acc. (%)	F1 (%)	Acc. (%)	F1 (%)	Acc. (%)	F1 (%)
Hammad et al. [4]		SD	96.34	94.74	97.56	97.30	–	–
Katsigiannis et al. [6]	SVM	SD	62.00	58.00	62.00	53.00	–	–
Vaezi et al. [5]		SD	94.00	–	95.00	–	–	–
Fang et al.[7]		SD	98.80	94.00	95.10	98.30	–	–
He et al. [8]	UDA+ SVM	SI	71.00	65.00	72.00	69.00	–	–
Our Approach (MMD+LSD)		SI	73.60 (8.91)	68.16 (12.01)	73.77 (8.51)	69.78 (11.97)	72.35 (6.35)	66.87 (7.86)
Our Approach (MMD)		SI	73.88 (9.90)	68.68 (12.49)	72.93 (8.58)	67.72 (12.85)	72.60 (8.10)	66.99 (10.87)
Our Approach (LSD)		SI	73.78 (8.95)	67.58 (12.42)	73.14 (8.85)	68.72 (11.61)	71.45 (7.80)	65.57 (9.85)

26(12):5964–5973, 2022.

[4] Hammad, D. S. and Monkaresi, H. (2022). ECG-based emotion detection via parallel-extraction of temporal and spatial features using convolutional neural network. Traitement du Signal, 39(1):43–57.

[5] M. Vaezi, M. Nasri, F. Azimifar, and M. Mosleh. Hybrid attention-based deep learning network for emotion recognition by ECG signal. Majlesi Journal of Electrical Engineering, 19(2), June 2025.

[6] Katsigiannis, S. and Ramzan, N. (2017). DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. IEEE Journal of Biomedical and Health Informatics, 22(1):98–107.

[7] A. Fang, F. Pan, W. Yu, L. Yang, and P. He. ECG-based emotion recognition using random convolutional kernel method. Biomedical Signal Processing and Control, 91:105907, 2024.

[8] He, W., Ye, Y., Li, Y., Pan, T., and Lu, L. (2021). Online cross-subject emotion recognition from ECG via unsupervised domain adaptation. In Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pages 1001–1005. IEEE.

[9] Govarthan, Praveen Kumar, et al. "Emotion classification using electrocardiogram and machine learning: A study on the effect of windowing techniques." Expert Systems with Applications 254 (2024): 124371.

[10] Li, J. and Washington, P. (2024). A comparison of personalized and generalized approaches to emotion recognition using consumer wearable devices: Machine learning study. JMIR AI, 3(1):e52171.

[11] T. Fan, S. Qiu, Z. Wang, H. Zhao, J. Jiang, Y. Wang, et al. A new deep convolutional neural network incorporating attentional mechanisms for ECG emotion recognition. Computers in Biology and Medicine, 159:106938, 2023.

[12] P. Sarkar and A. Etemad. Self-supervised ECG representation learning for emotion recognition. IEEE Transactions on Affective Computing, 13(3):1541–1554, 2020.

[13] Xiao, S., Qiu, X., Tang, C., & Huang, Z. (2023, June). A spatial-temporal ecg emotion recognition model based on dynamic feature fusion. In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1-5). IEEE.

[14] Schmidt, P., Reiss, A., Duerichen, R., Marberger, C., & Van Laerhoven, K. (2018, October). Introducing wesad, a multimodal dataset for wearable stress and affect detection. In Proceedings of the 20th ACM international conference on multimodal interaction (pp. 400-408).

[15] Li, J., & Washington, P. (2024). A comparison of personalized and generalized approaches to emotion recognition using consumer wearable devices: Machine learning study. JMIR AI, 3(1), e52171.

[16] Abd Al Aleem, M., Mubarak, R., Salem, N. M., & Sadek, I. (2024, October). A Deep Learning Approach Using WESAD Data for Multi-Class Classification with Wearable Sensors. In 2024 6th Novel Intelligent and Leading Emerging Sciences Conference (NILES) (pp. 194-197). IEEE.

[17] Nandini, Durgesh, et al. "An ensemble deep learning framework for emotion recognition through wearable devices multi-modal physiological signals." Scientific Reports 15.1 (2025): 17263.

[18] Ali, Kamran, and Charles E. Hughes. "A unified transformer-based network for multimodal emotion recognition." arXiv preprint arXiv:2308.14160 (2023).

[19] Vazquez-Rodriguez, Juan, et al. "Transformer-based self-supervised learning for emotion recognition." 2022 26th international conference on pattern recognition (ICPR). IEEE, 2022.