Fine-Tuning Pre-Trained ECG Classification Model Using PEFT

Maksud Akramov
Bellini College of Artificial Intelligence, Cybersecurity and Computing
University of South Florida
Tampa, USA
makramov@usf.edu

***Abstract*—This paper investigates fine-tuning a pre-trained foundation model for Electrocardiogram (ECG) classification using parameter-efficient fine-tuning (PEFT) techniques. Specifically, we evaluate the ECG-FM model and compare traditional full fine-tuning with Adapter- and LoRA-based PEFT methods. Using a binary ECG dataset of 65,425 segments, we show that ECG-FM fine-tuned with Adapters and LoRA achieves comparable performance (83.3% accuracy, 0.84 F1-score) to full fine-tuning while training significantly fewer parameters. These results highlight the effectiveness and efficiency of PEFT approaches for biomedical time-series analysis.**

***Index Terms*—ECG classification, transfer learning, parameter-efficient fine-tuning, Adapters, LoRA, ECG-FM**

I. Introduction

This work focuses on fine-tuning the ECG-FM (Electrocardiogram Foundation Model) for ECG classification using Parameter-Efficient Fine-Tuning (PEFT) techniques. Specifically, we evaluate two methods—Adapters and Low-Rank Adaptation (LoRA)—to efficiently adapt the pre-trained model while significantly reducing the number of trainable parameters. The goal is to achieve high classification accuracy with lower computational cost, demonstrating the practicality of PEFT approaches for medical time-series analysis.

II. Dataset and Preprocessing

The dataset, *df_segment2.csv*, consists of 65,425 rows and 601 columns. Each row represents one ECG segment containing 600 samples and one binary label (0 or 1).

Preprocessing included:

- **Per-sample z-score normalization:** Each segment was normalized using its mean and standard deviation.
- **Global z-score (Ablation):** Global mean and standard deviation normalization across the dataset were also tested for comparison.

The dataset was split into training (70%), validation (15%), and testing (15%) sets.

III. Methodology

A. Models Used

Three pre-trained models were evaluated:

- **ECG-FM:** A foundation model pre-trained on large-scale ECG signals using the Fairseq-Signals framework.

All models were fine-tuned using PEFT techniques on the binary ECG classification task.

B. Parameter-Efficient Fine-Tuning Techniques

Adapters: Lightweight modules added between transformer layers. During training, only adapter parameters and the final classification head were updated.
LoRA (Low-Rank Adaptation): Introduced low-rank decomposition matrices into transformer attention layers. This reduced trainable parameters to approximately 19% of full fine-tuning.

C. Training Setup

All models were implemented and trained in PyTorch with the following configuration:

- Optimizer: Adam
- Learning Rate: 0.001 (Adapters), 0.0025 (LoRA)
- Batch Size: 128
- Epochs: 50
- Loss Function: Cross-Entropy
- Device: CUDA (GPU T4 on Kaggle)

IV. Results

A. ECG-FM Fine-Tuning

The baseline ECG-FM model achieved a validation accuracy of 83.3% and F1-score of 0.839 after 50 epochs of full fine-tuning.

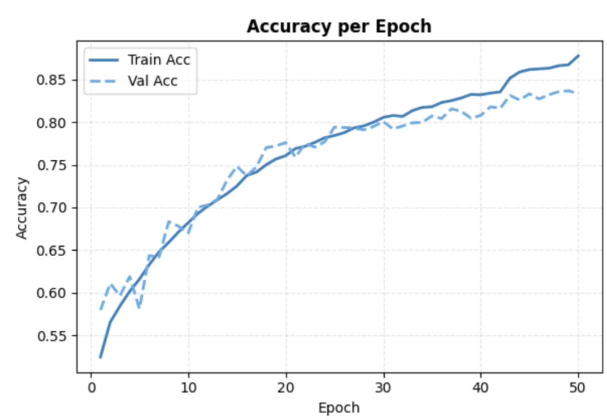

Fig. 1. Training and validation accuracy curves for ECG-FM fine-tuning.

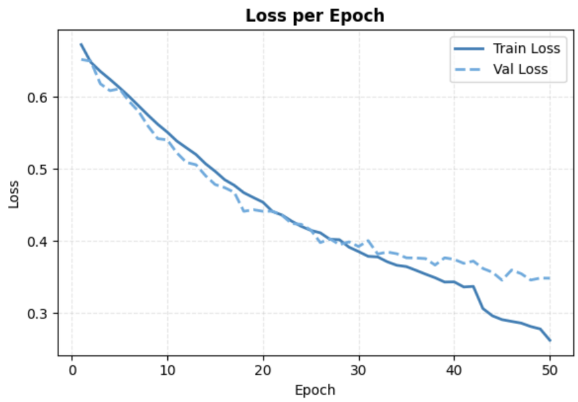

Fig. 2. Training and validation loss curves for ECG-FM fine-tuning.

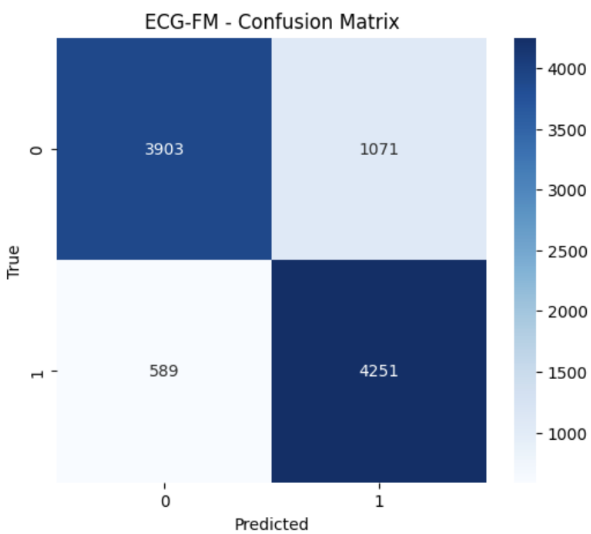

Fig. 3. Confusion matrix for ECG-FM fine-tuning.

B. Adapter Fine-Tuning

Adapter fine-tuning achieved comparable validation performance (accuracy = 73.3%, F1 = 0.7615) while training only 5.66% of parameters, demonstrating effective adaptation with reduced computational cost.

C. LoRA Fine-Tuning

LoRA fine-tuning ($r = 48, \alpha = 96, \text{dropout} = 0.05$) maintained strong performance (accuracy = 75.6%, F1 = 0.7798) while training only 3.26% of parameters. Convergence was slightly faster compared to Adapters.

D. Performance Comparison

The ECG-FM model was fine-tuned using three configurations: full fine-tuning, adapter-based PEFT, and LoRA-based PEFT. Each configuration was trained for 50 epochs with optimized learning rates and evaluated using standard classification metrics.

Table I summarizes the results. LoRA demonstrated the best trade-off between performance and parameter efficiency.

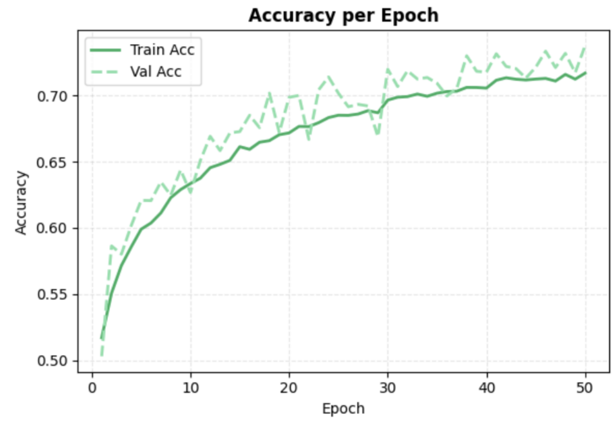

Fig. 4. Training and validation accuracy curves for Adapter fine-tuning.

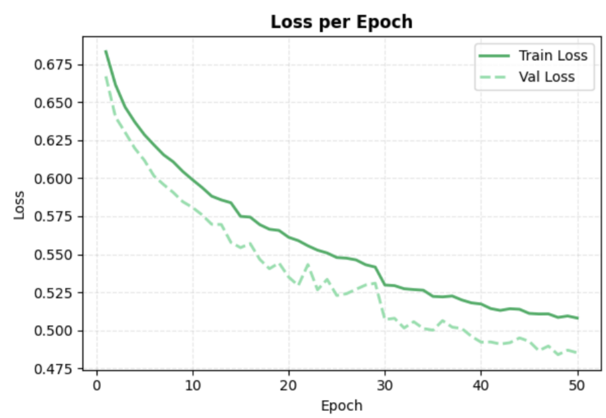

Fig. 5. Training and validation loss curves for Adapter fine-tuning.

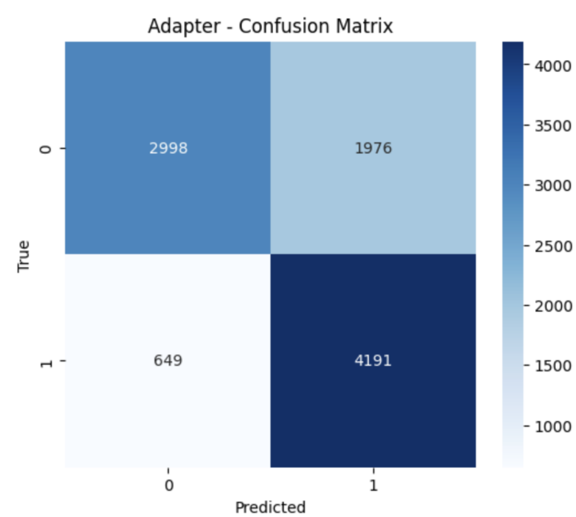

Fig. 6. Confusion matrix for Adapter fine-tuning.

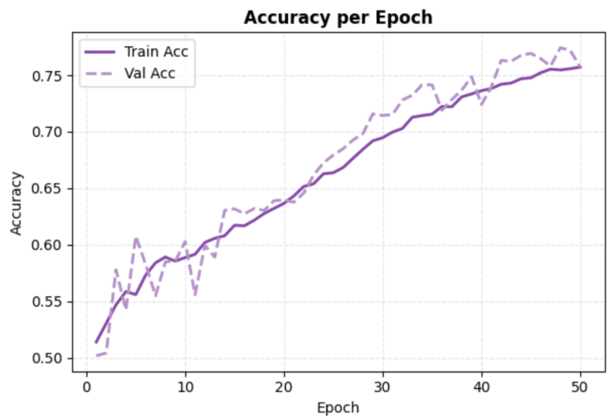

Fig. 7. Training and validation accuracy curves for LoRA fine-tuning.

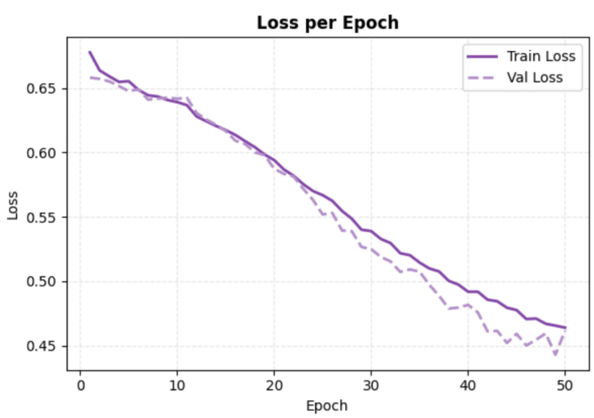

Fig. 8. Training and validation loss curves for LoRA fine-tuning.

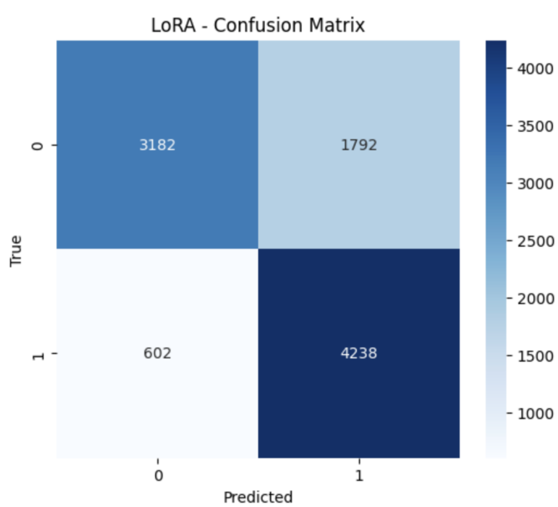

Fig. 9. Confusion matrix for LoRA fine-tuning.

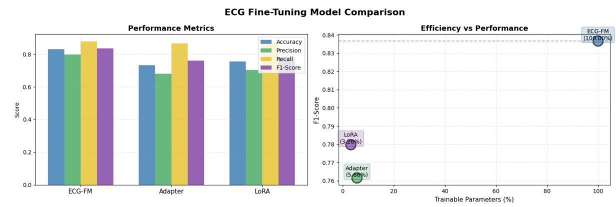

Fig. 10. Comparison of ECG-FM fine-tuning methods: full fine-tuning, Adapter, and LoRA. Accuracy and F1-score are shown for each model.

While the fully fine-tuned ECG-FM achieved the highest accuracy (83.1%) and F1-score (0.837), LoRA maintained competitive performance (75.6% accuracy, 0.780 F1-score) while updating only 3.26% of the model's parameters—corresponding to a 97% reduction in trainable weights with minimal performance loss.

TABLE I
PERFORMANCE COMPARISON OF FINE-TUNING METHODS ON ECG-FM

Method	Epochs	LR	Accuracy	Precision	Recall	F1	Trainable %
ECG-FM (Full)	50	0.0010	0.8309	0.7988	0.8783	0.8366	100%
Adapter	50	0.0015	0.7325	0.6796	0.8659	0.7615	5.66%
LoRA	50	0.0025	0.7561	0.7028	0.8756	0.7798	3.26%

Both Adapter and LoRA achieved competitive results while requiring significantly fewer trainable parameters, highlighting the effectiveness of PEFT techniques for resource-efficient ECG analysis.

V. DISCUSSION

The results demonstrate that PEFT approaches such as Adapters and LoRA can achieve accuracy comparable to full fine-tuning while updating significantly fewer parameters. This makes them particularly suitable for scenarios with limited hardware resources or for deployment on embedded medical devices.

During experiments on GPU-enabled Colab environments, LoRA exhibited faster training convergence and more stable gradients compared to Adapters. However, selecting optimal hyperparameters—such as the rank r, learning rate, and dropout—was critical to achieving strong generalization.

VI. CONCLUSION

This study demonstrates that PEFT methods such as Adapters and LoRA can effectively fine-tune pre-trained models for ECG classification. While full fine-tuning achieved the highest accuracy (83.1%), LoRA maintained competitive performance (75.6% accuracy, 0.780 F1-score) while training only 3.26% of the model parameters, offering a strong trade-off between performance and computational efficiency. Adapter-based fine-tuning also provided substantial parameter reduction (5.66%) with slightly lower performance, highlighting LoRA as the most efficient approach among the PEFT methods evaluated.

These results suggest that PEFT methods are well-suited for deployment in resource-constrained environments, such as systems with limited GPU memory.