Transfer Learning for ECG Classification Using ResNet

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

***Abstract*—This paper evaluates the use of transfer learning for classifying Electrocardiogram (ECG) images using a pre-trained ResNet model. Two approaches were studied: (1) using the model as a fixed feature extractor and (2) fine-tuning the network. The fixed feature extractor achieved approximately 66% accuracy, while fine-tuning improved accuracy to 73.5%. Results demonstrate that fine-tuning better adapts deep features to ECG-specific patterns, though at a higher computational cost.**

***Index Terms*—transfer learning, ECG, ResNet, fine-tuning, feature extraction**

I. Introduction

Electrocardiograms (ECG) are widely used in diagnosing cardiovascular conditions. Deep learning has shown great promise in ECG analysis, but training deep networks from scratch is computationally expensive and data-intensive. Transfer learning offers an efficient solution by adapting pre-trained convolutional neural networks (CNNs) to new tasks. This paper compares two strategies for ECG classification: a fixed feature extractor and fine-tuning.

II. Methodology

A. Dataset

The dataset consisted of 65,600 (after removing missing images) training and 18,136 testing ECG images across two balanced classes.

B. Model and Training Setup

A pre-trained ResNet model was used.

- Optimizer: Adam, learning rate 0.001
- Loss function: Cross-Entropy
- Epochs: 5

C. Experiments

1) **Fixed Feature Extractor**: Backbone layers were frozen; only the final fully connected layer was trained.
2) **Fine-Tuning**: Entire ResNet was updated during training.

III. Results

A. Fixed Feature Extractor

The fixed feature extractor froze the pre-trained ResNet backbone and trained only the final fully connected layer. Training accuracy plateaued around 66–70%, with test accuracy of 66.6%. Precision, recall, and F1-scores were balanced but modest.

Training Logs:

```
[Fixed Feature] Epoch 1 | Loss: 0.5481 | Acc: 0.652
[Fixed Feature] Epoch 2 | Loss: 0.5376 | Acc: 0.662
[Fixed Feature] Epoch 3 | Loss: 0.5061 | Acc: 0.678
[Fixed Feature] Epoch 4 | Loss: 0.4712 | Acc: 0.694
[Fixed Feature] Epoch 5 | Loss: 0.4543 | Acc: 0.699
```

Classification Report (Testing):

	precision	recall	f1-score	support
0	0.6673	0.6673	0.6673	502
1	0.6647	0.6647	0.6647	498
accuracy			0.6660	1000
macro avg	0.6660	0.6660	0.6660	1000
weighted avg	0.6660	0.6660	0.6660	1000

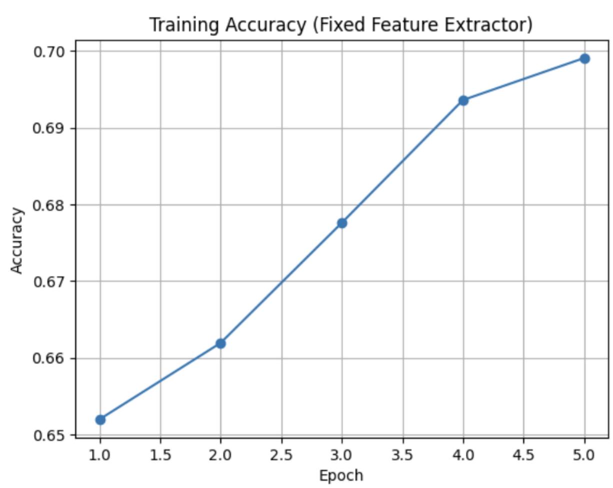

Fig. 1. Training accuracy curve for Fixed Feature Extractor.

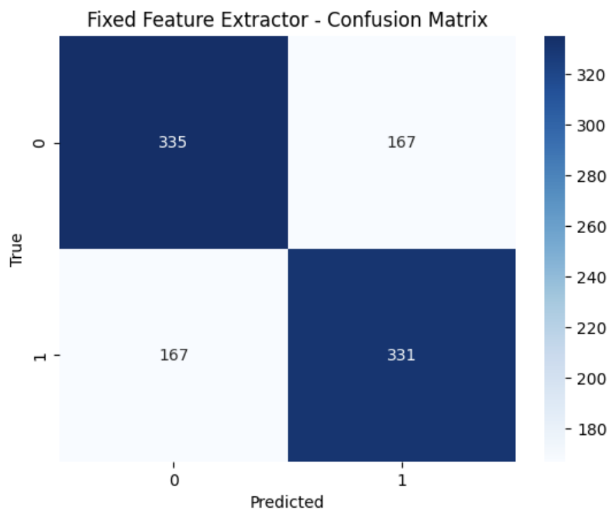

Fig. 2. Confusion matrix for Fixed Feature Extractor.

B. Fine-Tuning

Fine-tuning updated the entire ResNet during training, allowing adaptation to ECG-specific features. Training accuracy improved steadily, and the model achieved a test accuracy of 73.5%. Classification metrics also improved compared to the fixed feature extractor.

Training Logs:

```
[Fine-Tuning] Epoch 1 | Loss: 1.0266 | Acc: 0.710
[Fine-Tuning] Epoch 2 | Loss: 0.5174 | Acc: 0.720
[Fine-Tuning] Epoch 3 | Loss: 0.3344 | Acc: 0.727
[Fine-Tuning] Epoch 4 | Loss: 0.2982 | Acc: 0.735
[Fine-Tuning] Epoch 5 | Loss: 0.2468 | Acc: 0.740
```

Classification Report (Testing):

	precision	recall	f1-score	support
0	0.7327	0.7400	0.7363	500
1	0.7374	0.7300	0.7337	500
accuracy			0.7350	1000
macro avg	0.7350	0.7350	0.7350	1000
weighted avg	0.7350	0.7350	0.7350	1000

IV. Discussion

Fine-tuning consistently outperformed fixed feature extraction. Freezing the backbone limited adaptation to ECG-specific patterns, leading to lower accuracy. Fine-tuning allowed better representation learning but required more training time and carried a risk of overfitting. Hyperparameter tuning, particularly the learning rate, was critical to stable performance.

When training on an Apple Silicon M3 chip, several practical challenges were encountered. Increasing the number of workers for data loading sometimes caused issues with NumPy not recognizing certain libraries, leading to errors during preprocessing. Additionally, while the chip accelerates computation, training and testing times still increased significantly with larger batch sizes or deeper models, requiring careful balancing between performance and resource utilization.

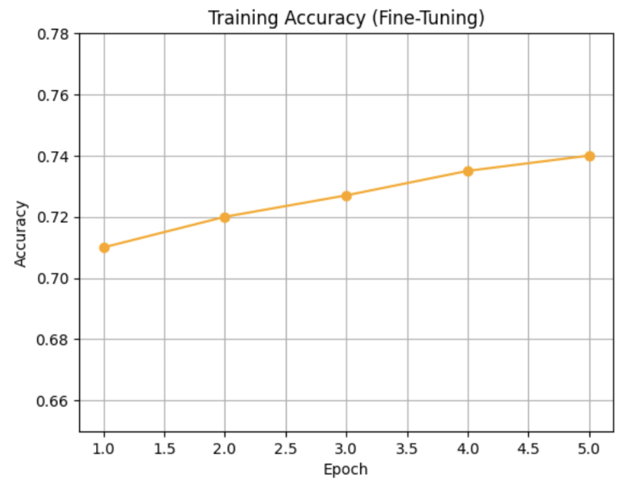

Fig. 3. Training accuracy curve for Fine-Tuning.

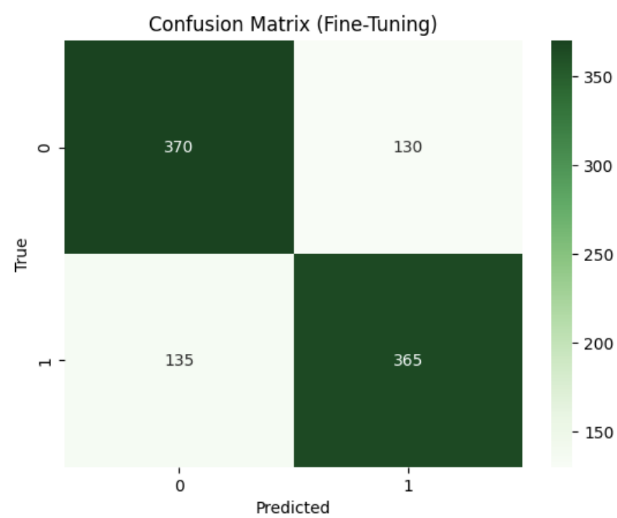

Fig. 4. Confusion matrix for Fine-Tuning.

V. Conclusion

Transfer learning significantly improved ECG classification performance. The fixed feature extractor achieved modest results (66%), while fine-tuning improved accuracy to 73.5%. Fine-tuning proved to be more effective but computationally expensive.

Acknowledgment

The author thanks Dr. Nima Karimian and the University of South Florida for providing guidance, valuable datasets, and all resources necessary for completing this work.