

Leveraging transfer learning with convolutional neural networks for cardiovascular disease detection

Jerry Chen¹, Yue Chen²

¹Loveless Academic Magnet Program High School, Montgomery, Alabama

²Auburn University at Montgomery, Department of Mathematics, Montgomery, Alabama

SUMMARY

Cardiovascular diseases (CVDs) are the top global cause of death. Early diagnosis of CVDs could mitigate their severity. A common method for diagnosis is electrocardiogram (ECG) signal analysis; however, it is time-consuming and inaccurate, even for expert physicians. To address these limitations, deep learning models offer a promising alternative for automated, high-accuracy analysis. This study aimed to determine whether parameter efficiency or a residual network architecture is more impactful in creating an accurate deep learning model. We hypothesized that residual networks would achieve superior accuracy in ECG-based CVD detection compared to other non-residual network architectures when leveraging transfer learning on small medical datasets. This advantage was attributed to the ability of residual networks to mitigate the vanishing gradient problem and preserve low-level signal details, whereas other non-residual network architectures instead prioritize parameter efficiency. To test this hypothesis, we trained EfficientNetV2-S and ResNet-50 on the same open-source dataset of 928 ECG images obtained from Ch. Pervaiz Elahi Institute of Cardiology in Multan, Pakistan to detect CVDs and potential signs of CVDs. The two models were evaluated on a test set. The findings supported our hypothesis, as the ResNet-50 model outperformed the EfficientNetV2-S model on all metrics. We concluded that architectural factors such as residual connections outweigh parameter efficiency in ECG analysis. Deep learning residual network architectures that leverage transfer learning could offer an efficient alternative to manual analysis.

INTRODUCTION

Cardiovascular diseases (CVDs) are the primary cause of death globally, killing around 17.9 million people annually (1, 2). As a result, improved methods for diagnosis are crucial, and early diagnosis is key to saving lives (2, 3). If CVDs are detected early, medical professionals can intervene in a timely manner to prevent a life-threatening situation (2). Currently, electrocardiograms (ECGs) are a commonly used non-invasive method to measure heart activity and spot potential warning signs of CVDs (4). However, manual ECG analysis is both time-consuming and challenging for even the most expert physicians, as it requires careful wave-by-wave measurement and complex pattern recognition (5).

To solve this problem, many researchers have begun to look towards artificial intelligence (AI) for solutions (3, 6). AI can extract information from datasets, analyze complex relationships in ECG signals, diagnose diseases

accurately, and shorten data processing time, saving medical professionals a significant amount of time and allowing for improved clinical analysis and diagnosis (3, 6, 7). Specifically, the convolutional neural network (CNN) is a revolutionary technology in computer vision, allowing machines to see the images provided to them as input (8). Unlike typical fully connected neural networks, which treat each input pixel as a separate node, a CNN uses specialized layers, such as convolutional and pooling layers, to process spatial hierarchies of features. This architecture enables the network to automatically and adaptively learn spatial feature hierarchies directly from pixel data. Consequently, CNNs have considerably advanced image processing in clinical backgrounds, such as improving x-ray diagnosis and showing potential in ECG data analysis (7-10). CNNs could fundamentally change how physicians detect CVDs and analyze ECG images.

However, creating a CNN from scratch is resource-intensive and time-consuming, which is often an obstacle to researchers with limited resources. Transfer learning overcomes this obstacle by reusing pre-trained models and retraining them on smaller datasets (9). Transfer learning has also seen great success in the medical field. Pre-trained models like AlexNet, ResNet, VGGNet, and GoogleNet were able to understand medical images, requiring fewer images to train compared to a CNN trained from scratch (11). In fact, models that leveraged transfer learning generally outperformed CNNs trained from scratch (12). For example, on the PTB-XL ECG dataset, which contains 21,799 clinical 12-lead ECGs from 18,869 patients of 10 second length, transfer learning led to models achieving higher accuracy with faster convergence and training speeds (12).

However, different pre-trained models could lead to varying results. Thus, we set out to investigate whether the presence of residual connections provides a greater performance benefit in ECG image analysis than a design focused on parameter efficiency. In this paper, we used EfficientNetV2-S and ResNet-50 as our two pre-trained models because they achieved excellent performance on the ImageNet dataset through vastly different model architectures (13, 14). ResNet-50 consists of residual nets with a depth of 50 layers, yet ResNet-50 is less complex than the Visual Geometry Group (VGG) pre-trained models, a class of models known for their depth and use in early deep learning breakthroughs, despite ResNet-50 having more layers (14, 18). Residual networks (ResNets) are a specific type of CNN that use skip connections, also known as residual connections, to jump over some layers. This architecture helps mitigate the vanishing gradient problem, a common issue in deep networks where

gradients become too small for effective learning, which can lead to feature degradation—the loss of important low-level details as information passes through many layers. In addition, ResNet-50 has shown excellent results in many medical applications, highlighting its vast potential in the medical field (15).

On the other hand, EfficientNetV2-S attained top-tier results on the ImageNet dataset by using a design approach focused on faster training and higher efficiency through smart model search and scaling (13). We hypothesized that ResNet-50 would outperform EfficientNetV2-S in ECG image classification due to its residual connections, which mitigate feature degradation in small datasets by preserving low-level signal details (14, 15). As the model with the residual network architecture (ResNet-50) displayed the best performance across all metrics, our results supported the hypothesis, suggesting that residual network architectural factors outweigh parameter efficiency on ECG image classification.

RESULTS

For both models, the same dataset was used to allow for comparisons. The dataset was downloaded from an open-

source dataset of ECG records from the Ch. Pervaiz Elahi Institute of Cardiology in Multan, Pakistan (16). The dataset was composed of 928 ECG images from 928 different patients that were categorized into 4 classes (the four possible ECG classifications): Myocardial Infarction (MI), History of Myocardial Infarction, Abnormal Heartbeat, and Normal Heartbeat; the classes had 239 images, 172 images, 233 images, and 284 images, respectively with representative images shown (Figure 1) (16, 17).

To evaluate whether residual network architecture or parameter efficiency is more critical for ECG image classification, we compared ResNet-50 (residual) and EfficientNetV2-S (parameter-efficient) across 10 independent trials. In each trial, both models were trained from their ImageNet-initialized weights on our dataset of ECG images. Model performance was measured by evaluating the final trained model on the identical, “hidden” test set for each trial, using classification accuracy, precision, recall, and F1-score as the evaluation metrics. Multiple trials were conducted to ensure procedure’s replicability and the results’ reliability. For the 10 trials conducted on each model, the residual neural network model (ResNet-50) had a higher score on all

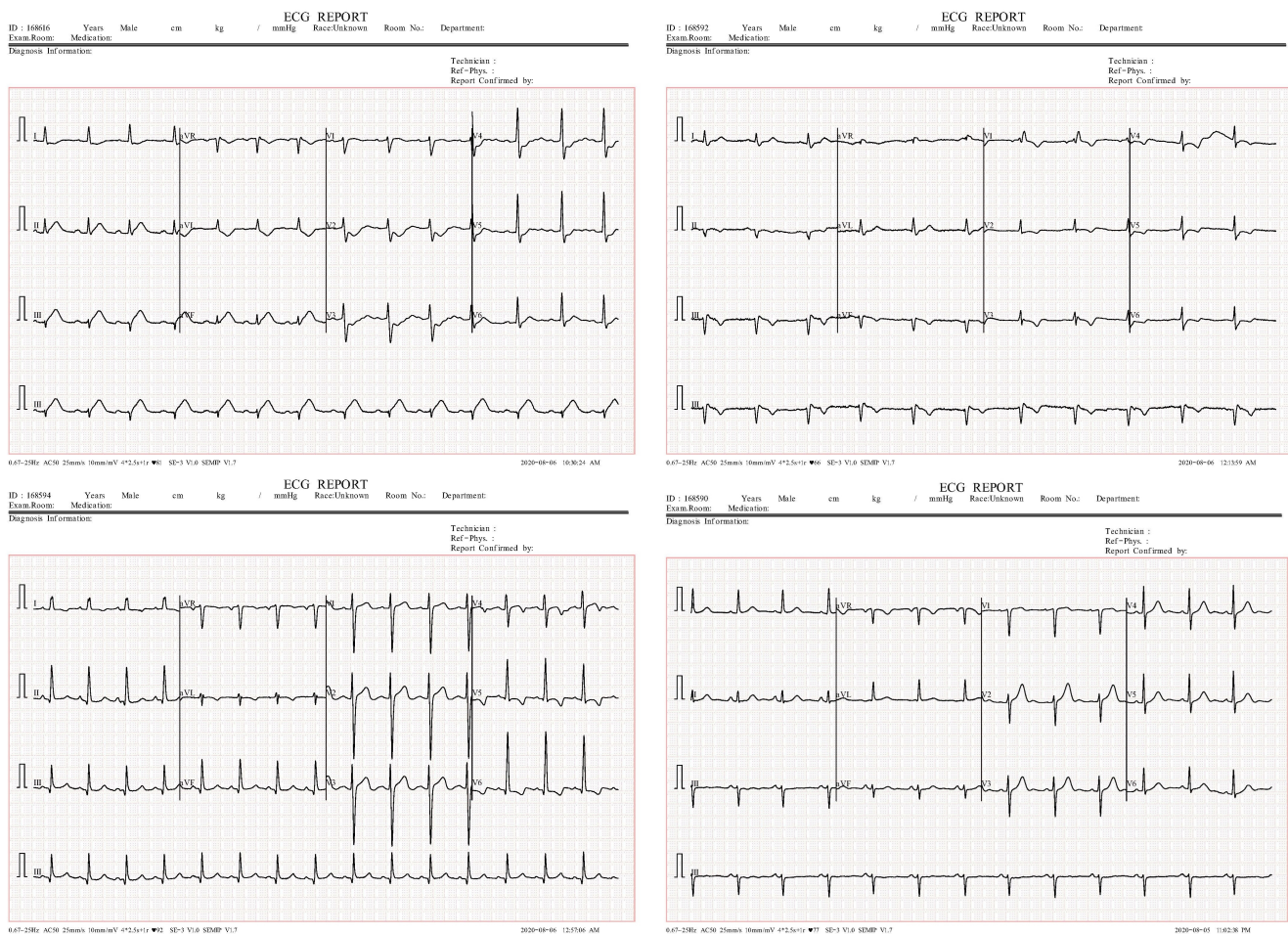


Figure 1. Sample images in the dataset. The four images represent the four possible diagnoses the models can return. Top Left: Myocardial Infarction. Top Right: History of Myocardial Infarction. Bottom Left: Abnormal Heartbeat. Bottom Right: Normal Heart Activity. Each ECG report is from a different patient.

the performance metrics than the non-residual, parameter-efficient neural network model (EfficientNetV2-S) (Figure 2). ResNet-50 achieved a mean accuracy of 94.76%, which was 7.03% higher than EfficientNetV2-S's mean accuracy of 87.73%. Similarly, ResNet-50's mean precision, recall, and F1-score were 6.72%, 7.51%, and 7.45% higher, respectively. Furthermore, the residual network architecture's performance displayed less variability, as evidenced by its narrower confidence intervals (Figure 2).

In order to test for statistical significance, we performed Bonferroni-corrected paired t-tests across all metrics (accuracy, precision, recall, F1-score). We found that the residual network architecture model (ResNet-50) was statistically significantly superior to the parameter-efficiency based neural network (EfficientNetV2-S) in ECG classification across accuracy, precision, recall, and F1-score (all $p < 0.0001$) (Figure 2). Thus, we have convincing statistical evidence to support our hypothesis that residual network architectures outweigh parameter efficiency in ECG image diagnosis.

DISCUSSION

In a clinical setting, if potential signs of CVDs were present and detected, physicians could take actions that would minimize the risk of a potentially life-threatening event. This could lead to significant health benefits as well as lower costs in the long run. An automated system that can accurately use ECG images to detect CVDs can significantly boost the accuracy of CVD diagnosis, allowing CVDs to be detected earlier. In this study, we compared the results of leveraging transfer learning from EfficientNetV2-S and ResNet-50 to detect CVDs and the potential signs of CVDs. Our findings show that ResNet-50 demonstrated superior performance across all metrics when compared EfficientNetV2-S (all $p < 0.0001$), suggesting that a residual network architecture is superior in ECG classification. The superior performance of residual network architectures can be attributed to their residual learning framework, which allows them to capture

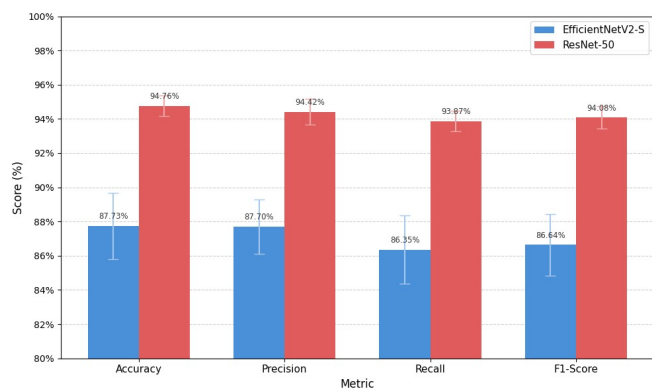


Figure 2. Performance differences between ResNet-50 and EfficientNetV2-S on ECG image classification. Bars represent mean accuracy, precision, recall, and F1-score for EfficientNetV2-S and ResNet-50 across 10 independent test trials. Models were trained on 741 ECG images and tested on the remaining 187 images. Precision, recall, and F1-score reflect macro-averages across all classes. Error bars represent 98.75% confidence intervals. All differences were statistically significant under a Bonferroni-adjusted alpha level of 0.0125 (all $p < 0.0001$).

spatially relevant features in medical imaging (14, 15). ResNet-50's shortcut connections enabled stable training of deeper layers, capturing spatially localized features essential for distinguishing myocardial infarction (MI) and MI warning signs from normal heartbeats. EfficientNetV2-S's neural architecture search framework has a higher emphasis on parameter efficiency and training speed and could not effectively capture all the complex patterns in the small dataset of ECG images (13).

This project shows the efficacy of leveraging transfer learning, specifically from residual networks, to detect CVDs and possible signs of CVDs. We selected ResNet-50 and EfficientNetV2-S for this comparison because they represent two of the leading but philosophically different approaches to modern CNN design: one prioritizing architectural innovation to enable stable, very deep networks (ResNet), and the other prioritizing parameter efficiency and training speed (EfficientNet) (13, 14). However, the dataset is quite limited, containing only 928 images from the Ch. Pervaiz Elahi Institute of Cardiology in Multan, Pakistan (16). The results could be biased towards a certain geographic region in Pakistan and may not be generalized to the entire population of the world. Furthermore, with only four classes, there may be many clinical situations that the model has not learned to classify. Future studies could build upon the existing dataset, adding ECG images from patients all across the globe. The new dataset should include a diverse set of images to allow the model to become more adaptive. A dataset that contains ECG images from patients across the globe would allow for the model's results to become generalized. Also, in this project, we only explored two CNN architectures. More CNN architectures could be explored to see which is the most suitable for ECG image analysis.

Overall, our results suggest that the architectural advantage from residual connections outweighs the benefits of parameter efficiency when it comes to detecting CVDs from ECG images. A model created by leveraging transfer learning from residual networks can be used in clinical settings to improve CVD detection and diagnosis. Thus, our findings indicate that leveraging transfer learning from residual networks alongside medical professionals is a highly promising approach for CVD detection and diagnosis, warranting further investigation.

MATERIALS AND METHODS

Dataset

The dataset used in this study was an open-source dataset of 928 images of ECG records from the Ch. Pervaiz Elahi Institute of Cardiology in Multan, Pakistan (16). This dataset had roughly equal class sizes, thus reducing the risk of creating a model that is biased towards a certain majority class; this is essential for building an effective deep learning model (17). Out of the 928 images, 187 images (approximately 20%) were set aside for testing, and the other 741 images (approximately 80%) were used for model development. Furthermore, out of the 741 images used for model development, 148 images (approximately 20% of the development data) were randomly chosen to validate the model to tune hyperparameters during training. The remaining 593 images (approximately 80% of the development data) were used as the training set.

Pre-trained Models

Transfer learning from two pre-trained neural networks, EfficientNetV2-S and ResNet-50, was leveraged because of the small size of the dataset and the ability of transfer learning to preserve the feature extraction abilities of the pre-trained models to improve training performance (9). EfficientNetV2-S is a state-of-the-art image classification model trained on the ImageNet dataset (13). It consists of 252 layers, has approximately 22 million parameters, and exhibits high parameter efficiency and a fast training speed (13). ResNet-50 is a residual network trained on the ImageNet dataset, and it is made up of 50 layers and approximately 25.6 million parameters (14). It has demonstrated excellent performance in several types of image classification applications, including many in the medical field (15).

Model Architecture, Image Preprocessing, and Training Procedure

Both models were set to train for up to 100 epochs with an initial learning rate of 0.0001, using the Adam optimizer. To prevent overfitting, early stopping was implemented by monitoring validation loss. Finally, the test set was used to evaluate the performance of the trained models. All ECG images were resized to 512 pixels × 1024 pixels to save computational resources. A size of 512 pixels × 1024 pixels was selected to maintain the original aspect ratio of the images. Cross-validation was applied to make sure the model could generalize to unseen data in order to prevent overfitting. All training and testing images were loaded into a dataset generator using the Keras library in Python. Furthermore, all labels were one-hot encoded to transform categorical data into binary columns so that machine learning algorithms could effectively interpret the information. Integer encoding was avoided in an attempt to prevent the model from assuming a natural ordering between categories, which could result in poor and/or unexpected results.

Two pre-trained models were considered to be the baseline model of the overall neural network: ResNet-50 and EfficientNetV2-S. Transfer learning was leveraged in that pre-trained weights on both ResNet-50 and EfficientNetV2-S were preserved by freezing all layers in the two pre-trained networks to maintain their feature extraction abilities. A global pooling layer was added to the end of the pre-trained models to reduce the number of spatial features to improve computational efficiency and reduce model complexity to prevent overfitting. A dense layer (a layer that connects every neuron - the basic computational unit that processes the input data - of the previous layer to every neuron of the current layer) of 1024 neurons and a dense layer of 512 neurons were added, as it is a common configuration used in medical neural networks that allows the neural network to learn complex features without overfitting (19). Both dense layers had a 20% dropout rate to prevent the model from overfitting. A softmax activation layer with 4 neurons was appended to the end of the model to output the probabilities for each of the 4 classes. Validation loss was tracked, and a patience of 5 epochs was used for early stopping to prevent the model from overfitting. With early stopping enabled, training could be stopped before the models are trained for 100 epochs.

Evaluation and Statistical Analysis

The models were then evaluated on the test set using the metrics accuracy, precision, recall, and F1-score. These metrics were calculated using the equations below:

$$Accuracy = \frac{Correct\ Predictions}{Total\ Predictions} \quad \dots(Equation\ 1)$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad \dots(Equation\ 2)$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad \dots(Equation\ 3)$$

$$F1-Score = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad \dots(Equation\ 4)$$

The above process was then repeated another 9 times for a total of 10 trials. Bonferroni-corrected paired t-tests were then conducted, with an adjusted α -level of 0.0125 taken as significant.

Platform and Computational Resources

All experiments were conducted on Kaggle. Kaggle's NVIDIA P100 GPU was used to train the models. Keras' pre-trained EfficientNetV2-S and ResNet-50 models were used to download the models' architectures and ImageNet weights. All needed Python libraries were imported into the Kaggle environment for use.

Received: June 18, 2025

Accepted: October 29, 2025

Published: May 25, 2026

REFERENCES

1. "Cardiovascular Diseases." *World Health Organization*. www.who.int/health-topics/cardiovascular-diseases#tab=tab_1. Accessed 17 Mar. 2025.
2. Oude Wolcherink, Martijn J., et al. "Health Economic Research Assessing the Value of Early Detection of Cardiovascular Disease: A Systematic Review." *PharmacoEconomics*, vol. 41, no. 10, 16 June 2023, pp. 1183–1203, <https://doi.org/10.1007/s40273-023-01287-2>.
3. Almansouri, Naiela E., et al. "Early Diagnosis of Cardiovascular Diseases in the Era of Artificial Intelligence: An In-Depth Review." *Curēus*, 9 Mar. 2024, <https://doi.org/10.7759/cureus.55869>.
4. Sattar, Yasar and Lovely Chhabra. "Electrocardiogram." *National Library of Medicine, StatPearls Publishing*, 2023, www.ncbi.nlm.nih.gov/books/NBK549803/.
5. Cook, David A., et al. "Accuracy of Physicians' Electrocardiogram Interpretations: A Systematic Review and Meta-Analysis." *JAMA Internal Medicine*, vol. 180, no. 11, 1 Nov. 2020, pp. 1461–1471, <https://doi.org/10.1001/>

- [jamainternmed.2020.3989](https://doi.org/10.59720/25-139).
6. Martínez-Sellés, Manuel and Manuel Marina-Breyse. "Current and Future Use of Artificial Intelligence in Electrocardiography." *Journal of Cardiovascular Development and Disease*, vol. 10, no. 4, 17 Apr. 2023, pp. 175–175, <https://doi.org/10.3390/jcdd10040175>.
 7. Xiong, Zhaohan, et al. "Robust ECG Signal Classification for the Detection of Atrial Fibrillation Using Novel Neural Networks." *Computing in Cardiology Conference (CinC)*, 14 Sept. 2017, <https://doi.org/10.22489/cinc.2017.066-138>.
 8. Kourounis, Georgios, et al. "Computer Image Analysis with Artificial Intelligence: A Practical Introduction to Convolutional Neural Networks for Medical Professionals." *Postgraduate Medical Journal*, 4 Oct. 2023, <https://doi.org/10.1093/postmj/qgad095>.
 9. Salehi, Ahmad W., et al. "A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope." *Sustainability*, vol. 15, no. 7, 1 Jan. 2023, p. 5930, <https://doi.org/10.3390/su15075930>.
 10. Yen, Chih-Ta and Chia-Yu Tsao. "Lightweight Convolutional Neural Network for Chest X-Ray Images Classification." *Scientific Reports*, vol. 14, no. 1, 30 Nov. 2024, <https://doi.org/10.1038/s41598-024-80826-z>.
 11. Kora, Padmavathi, et al. "Transfer Learning Techniques for Medical Image Analysis: A Review." *Biocybernetics and Biomedical Engineering*, vol. 42, no. 1, Jan. 2022, pp. 79–107, <https://doi.org/10.1016/j.bbe.2021.11.004>.
 12. Nguyen, Cuong V and Cuong D Do. "Transfer Learning in ECG Diagnosis: Is It Effective?" *ArXiv.org*, 2024, <https://doi.org/10.48550/arXiv.2402.02021>.
 13. Tan, Mingxing and Quoc V. Le. "EfficientNetV2: Smaller Models and Faster Training." *International Conference on Machine Learning 2021*, <https://doi.org/10.48550/arXiv.2104.00298>.
 14. He, Kaiming, et al. "Deep Residual Learning for Image Recognition." *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 10 Dec. 2015, <https://doi.org/10.48550/arXiv.1512.03385>.
 15. Rath, Arati, et al. "ResNet50-based Deep Learning model for accurate brain tumor detection in MRI scans." *Next Research*, vol. 2, no. 1, Mar. 2025, <https://doi.org/10.1016/j.nexres.2024.100104>.
 16. Khan, Ali Haider and Muzammil Hussain. "ECG Images dataset of Cardiac Patients." *Mendeley Data*, vol. 2, 19 Mar. 2021, <https://doi.org/10.17632/qwbz3fsgp8.2>.
 17. Sayin, Irem, et al. "CNN Based Detection of Cardiovascular Diseases from ECG Images." *ArXiv.org*, 29 August 2024, <https://doi.org/10.48550/arXiv.2408.16800>.
 18. Simonyan, Karen and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *International Conference on Learning Representations*, 4 September 2014, <https://doi.org/10.48550/arXiv.1409.1556>.
 19. Raina, Diya, et al. "Convolved neural network and transfer learning algorithm for improved brain tumor classifications in MRI." *Journal of Knowledge Learning and Science*, vol. 3, no. 24, Dec. 2024, <https://doi.org/10.60087/jkfst.v3.n4.p200>.
 20. Kingma, Diederik P. and Jimmy Ba. "Adam: A method

for stochastic optimization." *International Conference on Learning Representations*, 22 December 2014, <https://doi.org/10.48550/arXiv.1412.6980>.

Copyright: © 2026 Jerry Chen and Yue Chen. All JEI articles are distributed under the Creative Commons Attribution Noncommercial No Derivatives 4.0 International License. This means that you are free to share, copy, redistribute, remix, transform, or build upon the material for any purpose, provided that you credit the original author and source, include a link to the license, indicate any changes that were made, and make no representation that JEI or the original author(s) endorse you or your use of the work. The full details of the license are available at <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>.

Appendix

Link to GitHub repository for code:

<https://github.com/chenhaoyun0415/Transfer-Learning-ECG-Analysis>