Implementing Deep Learning for Atrial Fibrillation Diagnosis

Justine Mae S. Buenaventura   
College of Engineering, Architecture and Technology  
De La Salle University - DasmariñasCavite, Philippines  
bjs1388@dlsud.edu.ph   
  
Carl Louise D. Rostata   
College of Engineering, Architecture and Technology  
De La Salle University - DasmariñasCavite, Philippines  
rcd0891@dlsud.edu.ph  
  
Angel Ezekiel O. Anzures  
College of Engineering, Architecture and Technology  
De La Salle University - DasmariñasCavite, Philippines  
aao0986@dlsud.edu.ph

Dr. Gracita O. Topacio   
Internal Medicine  
ManilaMed – Medical Center ManilaManila, Philippines  
+63 999 192 3149   
  
Dr. Arvin R. Yumul   
Internal Medicine  
ManilaMed – Medical Center ManilaPampanga, Philippines  
+02 8523 8131  
  
Mark Montances  
Department of Electronics Engineering  
De La Salle University - DasmariñasCavite, Philippines  
momontances@dlsud.edu.ph

*Abstract*—Atrial fibrillation (AF) is a critical yet often underdiagnosed cardiac arrhythmia, particularly in low-resource environments where digital ECG systems and cardiology specialists may be inaccessible. This study presents a deep learning-based diagnostic tool designed to classify AF and normal sinus rhythm (NSR) from real-world, clinic-verified ECG images. Leveraging EfficientNet-B0, a convolutional neural network optimized for efficiency and accuracy, the model was trained on preprocessed images of 12-lead ECG printouts captured via mobile phone. Preprocessing included grayscale conversion, denoising, sharpening, and resizing, ensuring consistent input quality across variable capture conditions. The model achieved a 93% accuracy, balanced F1-scores of 0.93 for both classes, and an AUC of 0.98 on a validation dataset. Beyond validation, the model was integrated into a web-based interface and tested on newly acquired ECG images under varied real-world conditions, maintaining strong classification performance with stable confidence. By combining accessible imaging workflows with robust deep learning inference, this system offers a practical and scalable approach to AF detection in clinics without access to advanced diagnostic infrastructure.

Keywords—atrial fibrillation, ECG, deep learning, image classification, EfficientNet, web development

# Introduction

Atrial fibrillation (AF) is the most prevalent type of sustained cardiac arrhythmia, known to increase the risk of stroke, heart failure, and other life-threatening complications when left undetected. Despite its clinical significance, AF often remains undiagnosed, especially in community-level clinics and remote areas lacking access to advanced diagnostic tools or trained cardiologists. Conventional AF detection relies on digital ECG data interpreted by specialists or signal-based algorithms—resources that may be impractical or unavailable in many primary care settings.

This study addresses that gap by introducing an image-based deep learning solution capable of diagnosing AF from photographs of printed 12-lead ECG strips. These images, often captured via mobile phones, reflect the realities of front-line healthcare: inconsistent lighting, varied paper quality, and limited equipment. Instead of depending on raw digital signal inputs, this approach uses convolutional neural networks (CNNs) trained on preprocessed ECG images, offering a more accessible and cost-effective alternative.

Our system builds upon the EfficientNet-B0 architecture, chosen for its strong balance of accuracy and computational efficiency. It was trained using real ECG data collected from a clinical setting and integrated into a web-based interface designed for real-time, point-of-care use. This paper outlines the full process—from dataset preparation and model development to deployment and real-world evaluation—demonstrating how deep learning can be practically applied to support AF screening in under-resourced environments.

# Methodology

## Dataset Acquisition and Preprocessing

The dataset used in this study was collected from Dr. Topacio's clinic in Manila, where patients undergoing ECG examinations were screened under the supervision of a licensed physician. Only ECGs classified as either atrial fibrillation (AF) or normal sinus rhythm (NSR) were included, and each was manually validated by a medical expert to ensure diagnostic accuracy. The final set consisted of clinically verified 12-lead ECG printouts that were captured using mobile phone cameras under natural lighting conditions, representing the practical constraints of low-resource clinics.

Given the variability introduced by photographic capture—such as shadows, folds in the paper, inconsistent angles, and image noise—a robust preprocessing pipeline was designed to standardize the dataset while preserving relevant waveform information. Each image was first converted to grayscale to eliminate color inconsistencies and reduce computational complexity. Denoising techniques were applied using adaptive filters to remove visual noise without distorting key signal features. Sharpening algorithms were then used to enhance the contrast and visibility of ECG waveforms, making important components like P-waves and R-peaks more distinguishable.

To further emphasize the structure of the waveforms, morphological operations such as erosion and dilation were employed. These helped reinforce the line quality of the traces while suppressing background text and markings commonly found in ECG paper reports. Images were then padded to maintain a square aspect ratio and resized to 224×224 pixels—dimensions aligned with the input requirements of the EfficientNet-B0 model architecture.

This preprocessing strategy ensured that despite inconsistencies in capture conditions, each ECG image fed into the model had consistent spatial formatting and enhanced waveform clarity. By bridging the gap between uncontrolled real-world image capture and deep learning requirements, this pipeline enabled the model to generalize well and maintain high classification performance across varied input conditions.

## Model Development and Training

To develop the final model that achieved the best diagnostic performance, EfficientNet-B0 was selected due to its proven capability in handling image classification tasks with high accuracy and computational efficiency. The architecture’s compound scaling and depth-wise separable convolutions made it well-suited for the lightweight deployment constraints of a real-time web-based environment.

The training was conducted exclusively on a fully preprocessed dataset composed of 12-lead ECG images sourced from a clinical setting. Each image was converted to grayscale and resized to 224×224 pixels. Preprocessing steps included denoising, sharpening, and morphological operations to enhance the visual clarity of the ECG waveforms. These techniques were critical in enabling the model to learn consistent and relevant spatial features despite potential variability in printed ECG formats.

For the final training phase, the model was trained from scratch using the Adam optimizer with an initial learning rate of 0.0005 and a batch size suitable for stable convergence in Google Colab’s GPU environment. A learning rate scheduler was used to reduce the rate once performance plateaued. To further improve model generalization and calibration, label smoothing was introduced, helping the model make more measured predictions, especially on visually ambiguous samples.

The model was evaluated using the macro F1-score to ensure balanced performance across the two classes—AF and NSR. Early stopping was applied based on this metric to prevent overfitting. No class imbalance correction was necessary, as the final dataset was carefully curated to ensure equal representation of AF and NSR cases.

The final model checkpoint was selected based on validation performance, ensuring the most reliable and stable predictions under real-world testing. This final version was integrated directly into the deployed web application and demonstrated consistent accuracy, responsiveness, and reliability during live inference.

## System Integration and Development

The diagnostic platform was developed as a full-stack web system using Laravel and deployed on a LEMP (Linux, Nginx, MySQL, PHP) server, designed to function efficiently in real-world clinical workflows. The frontend was built to be mobile-responsive and intuitive, allowing users to upload ECG images directly through file input or live capture via device camera. The layout prioritized simplicity while offering a detailed display of diagnostic results.

Upon image submission, the Laravel backend handles file validation—ensuring correct formats, resolution, and safe size limits—before invoking a Python-based inference pipeline using shell execution. The backend communicates with the trained EfficientNet-B0 model, which processes the image through the same preprocessing steps used during training, ensuring consistency.

Inference results include the predicted class (AF or NSR), confidence score, and a Grad-CAM-based heatmap for interpretability. These outputs are returned in real-time and rendered on a custom results page showing the input ECG, diagnosis summary, and corresponding visualization overlays.

The system includes an admin panel with user authentication and activity logs for tracking uploaded files and test history. The backend supports modular scaling, enabling future additions of multi-class classification models or API-based mobile integration.

During testing and deployment, attention was also given to error handling, network latency, and user education prompts, ensuring a smooth experience from capture to result. This system was optimized not only for technical performance but also for clinical practicality.

# Results

## Performance Evaluation

The final EfficientNet-B0 model underwent comprehensive evaluation using a validation dataset composed of preprocessed ECG images unseen during training. The model achieved an overall accuracy of 93%, with an F1-score of 0.93 for both AF and NSR classes, demonstrating its balanced sensitivity and specificity. The area under the receiver operating characteristic curve (AUC) reached 0.98, indicating the model’s strong discriminative capability between the two classes.

Additional performance metrics, including precision and recall, were closely aligned across both categories, supporting the conclusion that the model was neither overfitting nor favoring one class over another. A confusion matrix generated from validation results further revealed a low number of false positives and false negatives, solidifying the model’s reliability in capturing the subtle variations in ECG waveform patterns associated with AF and NSR.

The evaluation phase confirmed that the model could consistently and confidently recognize complex rhythm characteristics, even without relying on raw signal data, establishing it as a viable solution for image-based ECG classification.

## Real-World Testing Insights

To complement the validation metrics, the system was deployed for real-world testing using 20 printed ECG strips from Dr. Topacio’s clinic—10 labeled as AF and 10 as NSR. These strips were captured via mobile phones under varying environmental conditions that mimicked actual field use: ideal lighting with stable angles, shadowed or close-up captures, and distant or skewed perspectives.

In controlled conditions, the model's predictions aligned perfectly with the clinical labels, producing confidence scores exceeding 90%. However, when image quality was compromised due to shadows, reflections, or off-angle positioning, confidence scores were noticeably reduced, often ranging from 65% to 80%. Importantly, the classification accuracy remained robust, but the reduction in confidence signaled the model’s cautious approach when visual clarity was diminished.

One key behavioral observation was that in highly ambiguous cases—such as ECGs with waveform distortion from wrinkled paper or poor focus—the model leaned toward predicting NSR. This conservative decision tendency reflects a bias toward safety, reducing the risk of unnecessary AF alarms in uncertain conditions.

Real-world testing affirmed the model's usability and resilience outside of ideal scenarios. It highlighted the importance of user training and proper image capture protocols, which could be integrated into future system updates through feedback mechanisms or automated quality checks prior to inference.

# Discussion

This study highlights the practical application of deep learning in detecting atrial fibrillation using real-world ECG image inputs. By focusing on printed 12-lead ECGs captured via mobile devices, the system addresses a pressing clinical need—offering diagnostic support in areas without access to advanced digital ECG infrastructure or cardiology expertise.

The model’s consistent performance across varied environmental conditions demonstrates the strength of its design. The preprocessing techniques ensured robustness to common issues like image noise, shadows, and misalignments, while the use of label smoothing and confidence calibration contributed to a more cautious and realistic diagnostic output. These factors collectively made the model well-suited for real-world deployment, especially in frontline healthcare settings where decision support tools can make a critical difference.

Integration into a web-based interface not only enhanced accessibility but also offered interpretability features such as Grad-CAM overlays, which help clinicians understand the model's focus areas. However, the system's tendency to default conservatively toward NSR in ambiguous cases, while clinically safer, points to areas for improvement—such as integrating uncertainty estimation or recommending repeat capture when confidence is low.

Challenges remain, particularly in ensuring consistent image quality across devices and users. Expanding training data to include more varied arrhythmias, integrating real-time image quality checks, and enhancing feedback mechanisms will be essential next steps.

# Conclusion

This research successfully developed and deployed a lightweight, accurate, and accessible deep learning model for detecting atrial fibrillation from printed ECG images. By leveraging real clinic-sourced data and optimizing for real-world usability, the system achieved high diagnostic performance, with 93% accuracy and balanced class-wise F1-scores.

The deployment through a Laravel-based web application ensures that healthcare workers—particularly in community clinics or underserved areas—can easily use the system without specialized hardware or software. The model's real-time predictions, interpretability features, and responsive design offer a scalable foundation for AI-driven cardiac screening.

Future work will aim to expand the model’s diagnostic capabilities to include multiple arrhythmias, improve capture consistency via guided feedback, and implement adaptive learning mechanisms to allow for continuous improvement based on field data. Ultimately, this project underscores the transformative potential of combining clinical collaboration with AI-driven innovation.

# Acknowledgment

The authors would like to express their sincere appreciation to De La Salle University – Dasmariñas for cultivating an academic environment that fosters innovation, research, and collaboration. We are also deeply grateful to our professor, Ms. Kathleen Villanueva, whose dedicated mentorship, expert guidance, and unwavering support were vital throughout the course of this project. Her encouragement and insight helped shape this work into a meaningful contribution to the field of AI-assisted healthcare.

##### References

[1] Y. Baek et al., "A deep learning algorithm for detecting AF during sinus rhythm," Scientific Reports, vol. 11, 12818, 2021.

[2] A. Król-Józaga, "Visual transformation of ECGs for CNN classification," Computing in Cardiology, vol. 49, 2022.

[3] R. Ribeiro et al., "Automatic diagnosis of ECGs at scale," Nature Communications, vol. 11, no. 1, p. 1760, 2020.

[4] J.W. Hughes et al., "Performance of CNNs versus physicians in ECG interpretation," JAMA Cardiology, vol. 6, no. 11, pp. 1285–1293, 2021.

[5] R. Ao and G. He, "Image-based deep learning for ECG interpretation," Frontiers in Artificial Intelligence, vol. 5, 2023.

[6] D. Patwari et al., "A Ramanujan-based CNN for AF detection," IEEE Journal of Biomedical and Health Informatics, 2024.

[7] S. Hong et al., "Opportunistic detection of atrial fibrillation from ECGs by a deep learning algorithm," Nature Medicine, vol. 26, no. 1, pp. 81–88, 2020.

[8] J. Lin et al., "AI-assisted ECG analysis for AF diagnosis," IEEE Access, vol. 9, pp. 98765–98773, 2021.

[9] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proc. Int. Conf. Mach. Learn., 2019, pp. 6105–6114.

[10] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. Int. Conf. Learn. Represent., 2015.

[11] C. Szegedy et al., "Rethinking the inception architecture for computer vision," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 2818–2826.

[12] G. Li et al., "Classification of ECG signals based on 1D convolutional neural network and attention mechanism," IEEE Access, vol. 9, pp. 34349–34357, 2021.

[13] F. S. Yen et al., "Application of convolutional neural network in the diagnosis of arrhythmia based on ECG signals," Sensors, vol. 20, no. 18, p. 5219, 2020.

[14] M. Abedi et al., "Interpretation of 12-lead ECGs using deep convolutional neural networks," Journal of Electrocardiology, vol. 58, pp. 66–70, 2020.

[15] Z. Zhang et al., "ECG image classification using deep CNNs with enhanced spatial resolution," Biomedical Signal Processing and Control, vol. 72, 2022.

[16] B. Yildirim, "A novel wavelet sequence-based on deep bidirectional LSTM network model for ECG signal classification," Computers in Biology and Medicine, vol. 96, pp. 189–202, 2018.

[17] J. Hannun et al., "Cardiologist-level arrhythmia detection and classification in ambulatory ECG using a deep neural network," Nature Medicine, vol. 25, pp. 65–69, 2019.

[18] A. M. Ayyad and A. B. Nassif, "A deep learning framework for detecting arrhythmia using 2D ECG images," IEEE Access, vol. 10, pp. 87438–87450, 2022.

[19] Y. LeCun et al., "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[20] J. Esteva et al., "A guide to deep learning in healthcare," Nature Medicine, vol. 25, no. 1, pp. 24–29, 2019.