Heart Disease Classification Using EfficientNetB5 with Three-Dimensional Scaled Electrocardiogram Images

Clinton Mwangi[1], David Warutumo[2] , Paul Bett[3], Mary Kariuki[4]

[1 2 3 4] Center for Data Science and Artificial Intelligence, Nyeri, Kenya

Abstract

In this study, we applied transfer learning to improve the classification of ECG images into various heart disease categories. We utilized a pre-trained convolutional neural network, EfficientNetB5 which is a convolutional neural network architecture that employs a compound scaling method to uniformly scale depth, width, and resolution using fixed coefficients. Unlike traditional methods that independently adjust these dimensions, EfficientNet scales all three dimensions simultaneously with a consistent set of scaling factors. We modified the architecture of the EfficientNetB5 model to accommodate our specific classification task of eleven heart disease categories. This adaptation involved replacing the final classification layer of the pre-trained model with a new output layer tailored for eleven distinct classes. We then fine-tuned the modified model on our dataset of 9,185 12-lead generated ECG images from PTB-XL database, which included 4,005 distortion-free images, 2,050 images with text distortions, and 3,130 images with wrinkled distortions. This fine-tuning process involved training the remaining layers of the network while leveraging the pre-trained weights to enhance feature extraction and accelerate convergence. The model performed with an f measure of 0.84 and an official f measure of 0.23.

1. Introduction

The electrocardiogram (ECG) is a critical tool in the diagnosis and management of cardiovascular diseases (CVDs) [1]. Despite advancements in digital ECG devices, physical or paper ECGs remain prevalent, particularly in regions with limited access to advanced medical technology [2]. The George B. Moody PhysioNet Challenge 2024 underscores the importance of developing algorithms that can digitize and classify ECGs from images or paper printouts, aiming to bridge the gap in cardiac care accessibility [3]. This study focuses on leveraging transfer learning to enhance the classification of ECG images into distinct heart disease categories. By utilizing a pre-trained convolutional neural network, EfficientNet B5, known for its superior image classification capabilities, we aim to improve the accuracy and efficiency of ECG-based diagnoses. EfficientNet employs a compound scaling method to uniformly scale depth, width, and resolution, which is crucial for handling the diverse and complex nature of ECG images.

2. Related Works

Transfer learning has become a vital technique for overcoming the challenges of training deep learning models from scratch. By fine-tuning a pre-trained model on a smaller, task-specific dataset, this method effectively utilizes the features learned from a large, often unrelated dataset, reducing the need for extensive labeled data and speeding up training [4] [5]. EfficientNet, a CNN architecture introduced by Tan and Le in 2019, has gained popularity for transfer learning in image classification tasks due to its compound scaling method, which optimally adjusts the network's depth, width, and resolution [3] [6]. This efficient scaling leads to better performance with fewer parameters and lower computational costs compared to traditional CNNs. EfficientNet's architecture has proven particularly effective in medical image classification, including ECG image classification, where it can be customized by replacing the final layer with a new output layer suited to the specific categories of heart disease. This fine-tuning not only improves diagnostic accuracy but also accelerates the training process [7][8][9].

The integration of EfficientNet and transfer learning in ECG image classification represents a significant advancement in the field of medical diagnostics. By leveraging pre-trained CNN models, researchers can enhance the accuracy and efficiency of ECG classification, contributing to more effective and timely diagnosis of heart diseases. Future research should continue to explore the optimization of these models, focusing on overcoming the current limitations and further improving diagnostic performance[10].

3. Methods

3.1. Challenge Data

In this study, we used dataset from publicly available PTB-XL database which consists of 21,799 12-lead ECG recordings, with ECG images [2]. We used ECG image generator to generate the ECG images for the classification task [11].

The data included 9185 ECG generated images which comprised of all the 11 classes namely: (Acute MI) acute myocardial infarction, (AFIB/AFL) atrial fibrillation or atrial flutter, (BRADY) bradycardia, (CD) conduction disturbances, (HYP) hypertrophy, (NORM) normal ECG, (Old MI) old myocardial infarction, (PAC premature atrial complex, (PVC) premature ventricular complex, (STTC) ST/T changes, and (TACHY) tachycardia. Figure 1 shows the distribution of the labels in our dataset. The generated ECG images could have more than one class where a patient could be having more than one cardiac health condition and this calls for multi-label classification.

To diversify the dataset and to avoid over fitting the model, different kind of distortions was introduced in the images as shown in Figure 3, specifically wrinkled and creases (Figure 4) , text distortions (Figure 5). Figure 2 depicts how different distortions were distributed in the dataset.

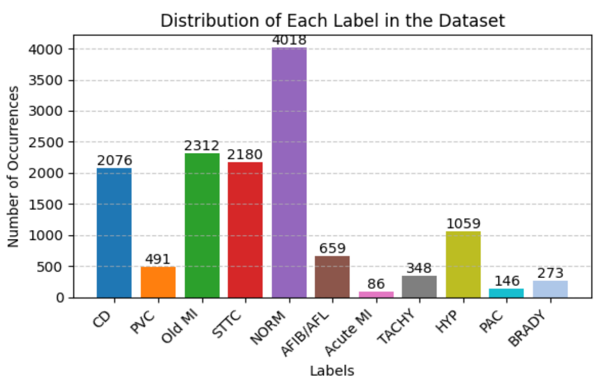

Figure 1. Distribution of labels in generated images data.

3.2. Proposed Model

Deep learning techniques has been used to classify images and have been having tremendous results. In this study, we used EfficientNetB5 pretrained model and fine tuned to classify ECG images. EfficientNetB5 models scales depth, width and resolution equally and progressively. The stem of this model include input layer, normalisation, zero padding, scaling, batch normalisation Conv2D, activation, and finally final layer block. The EfficientNetB5 model consist of this stem model and another

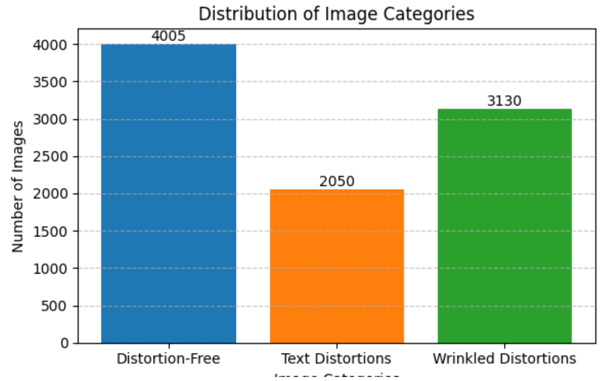

Figure 2. Distribution of distortions in the data.

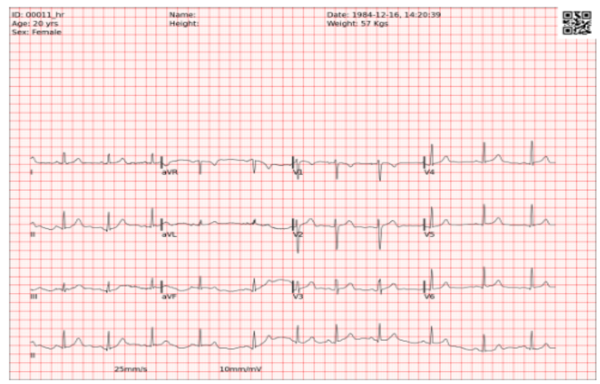

Figure 3. Generated ECG image without distortions.

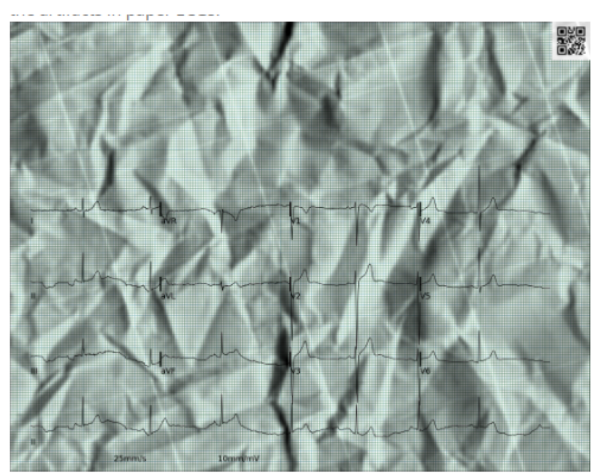

Figure 4. Generated ECG image with wrinkled distortions.

5 modules.The data is first split into 80% training and 20% testing. The training dataset is preprocessed by batch normalization and then passed to EfficientNetB5 sequential model. The model is trained and validation for the model done. Figure 6 shows the flowchart of the proposed architecture used in this study.

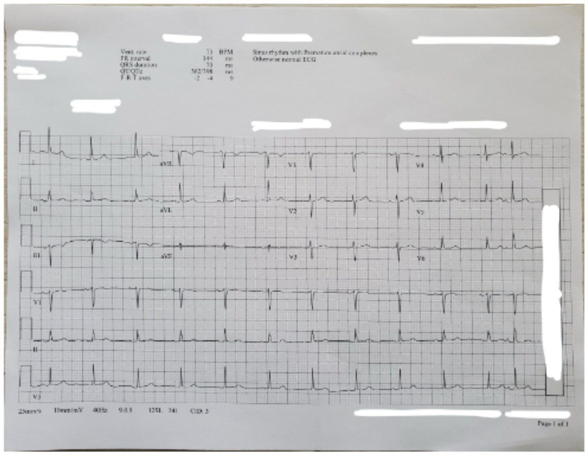

Figure 5. Generated ECG image with text distortions.

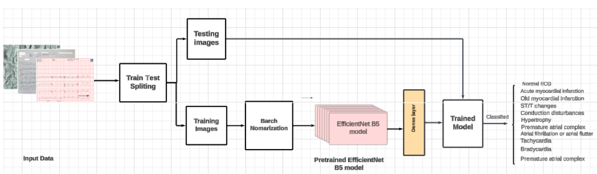

Figure 6. Proposed workflow for classification of ECG images.

4. Results and Discussion

The study utilized the Adam optimizer with a learning rate of 1e-4, a weight decay of 1e-5, and a scheduler with a step size of 10 and gamma value of 0.5. Training was conducted over 10 epochs with a batch size of 4. The model achieved an unofficial F-measure of 0.64, indicating moderate performance in balancing precision and recall during the training phase.

However, the official PhysioNet score was significantly lower, at 0.23. This discrepancy could be due to several factors, including the characteristics of the dataset. The data likely exhibited high diversity, leading to challenges in generalization. In such cases, the model may have overfitted to specific patterns in the training data, resulting in a good performance metric (unofficial F-measure) during internal validation but poor generalization to the unseen test data (as reflected in the PhysioNet score).

Figure **??** illustrates the loss function across epochs, which could provide insights into the model's convergence behavior. A decreasing loss indicates that the model is learning, but if the loss plateaus or fluctuates, it might suggest difficulties in further improvement. Figure **??**, which shows the Precision-Recall and Receiver Operating Characteristic (ROC) curves, offers additional evaluation metrics.

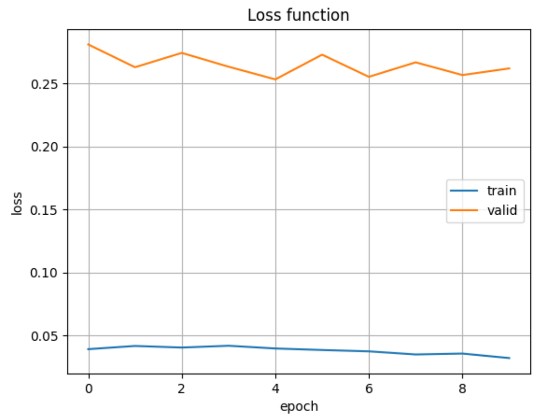

Figure 7. Loss function at each epoch.

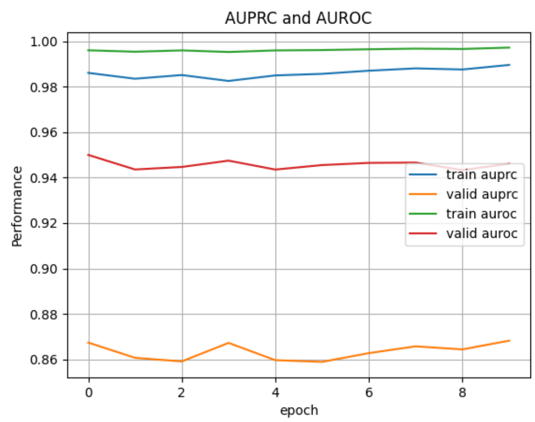

Figure 8. Plot showing the area under Precision Recall and Receiver Operating Characteristic curve.

5. Conclusions

This study utilized transfer learning with the EfficientNetB5 model to classify ECG images into heart disease categories. Despite achieving a moderate unofficial F-measure, the lower official PhysioNet score revealed challenges in generalization, likely due to dataset diversity. Future work should focus on addressing overfitting and enhancing model robustness, possibly through advanced regularization techniques or by expanding the dataset. Our results contribute to ECG image classification and offer insights for improving diagnostic tools in resource-limited settings.

Acknowledgments

We would like to express our sincere appreciation to the Center for Data Science and Artificial Intelligence, under the stewardship of Prof. Ciira, for their support and sponsorship of this study. We also extend our heartfelt gratitude to the organizers of the George B. Moody PhysioNet Challenge 2024 for providing this opportunity to contribute to the Computing in Cardiology (CinC) conference.

References

[1] PhysioNet. Physionet challenge 2024. `https://physionet.org/challenge/2024/`, 2024. Accessed: 2024-08-28.

[2] Wagner P, et al. Ptb-xl, a large publicly available electrocardiography dataset. Scientific Data 2020;7(1):1–15.

[3] Tan M, Le QV. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th International Conference on Machine Learning (ICML). 2019; 6105–6114.

[4] Pan SJ, Yang Q. A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering 2010; 22(10):1345–1359.

[5] Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? In Advances in Neural Information Processing Systems (NeurIPS). 2014; 3320–3328.

[6] Tan M, Le Q. Efficientnetv2: Smaller models and faster training. In Proceedings of the 38th International Conference on Machine Learning (ICML). 2020; 7754–7764.

[7] Adaei M, Zhang X. Efficientnet-based architectures for ecg arrhythmia classification. In Proceedings of the 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2020; 1301–1308.

[8] Hossain MS, Kim KJ. Deep learning-based ecg classification and heart disease prediction: A survey. Journal of Healthcare Engineering 2020;2020:8837412.

[9] Gupta S, Maji S. Transfer learning with efficientnet for medical image classification: A comparative analysis. IEEE Access 2021;9:104237–104249.

[10] Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 2018;172(5):1122–1131.

[11] Shivashankara KK, Deepanshi, Shervedani AM, Reyna MA, Clifford GD, Sameni R. Ecg-image-kit: A synthetic image generation toolbox to facilitate deep learning-based electrocardiogram digitization. Physiological Measurement May 2024;45(5):055019.

Address for correspondence:

Clinton Mwangi Kuya
Private Bag - 10143, Dedan Kimathi
clinton.mwangi@dkut.ac.ke

David Warutumo
Private Bag - 10143, Dedan Kimathi
david.warutumo@dkut.ac.ke

Paul Bett
Private Bag - 10143, Dedan Kimathi
bett.paul@dkut.ac.ke

Mary Kariuki
Private Bag - 10143, Dedan Kimathi
kariuki.mary20@students.dkut.ac.ke