CPE42

IMPLEMENTING DEEP LEARNING FOR ATRIAL FIBRILLATION DIAGNOSIS

Presented by :

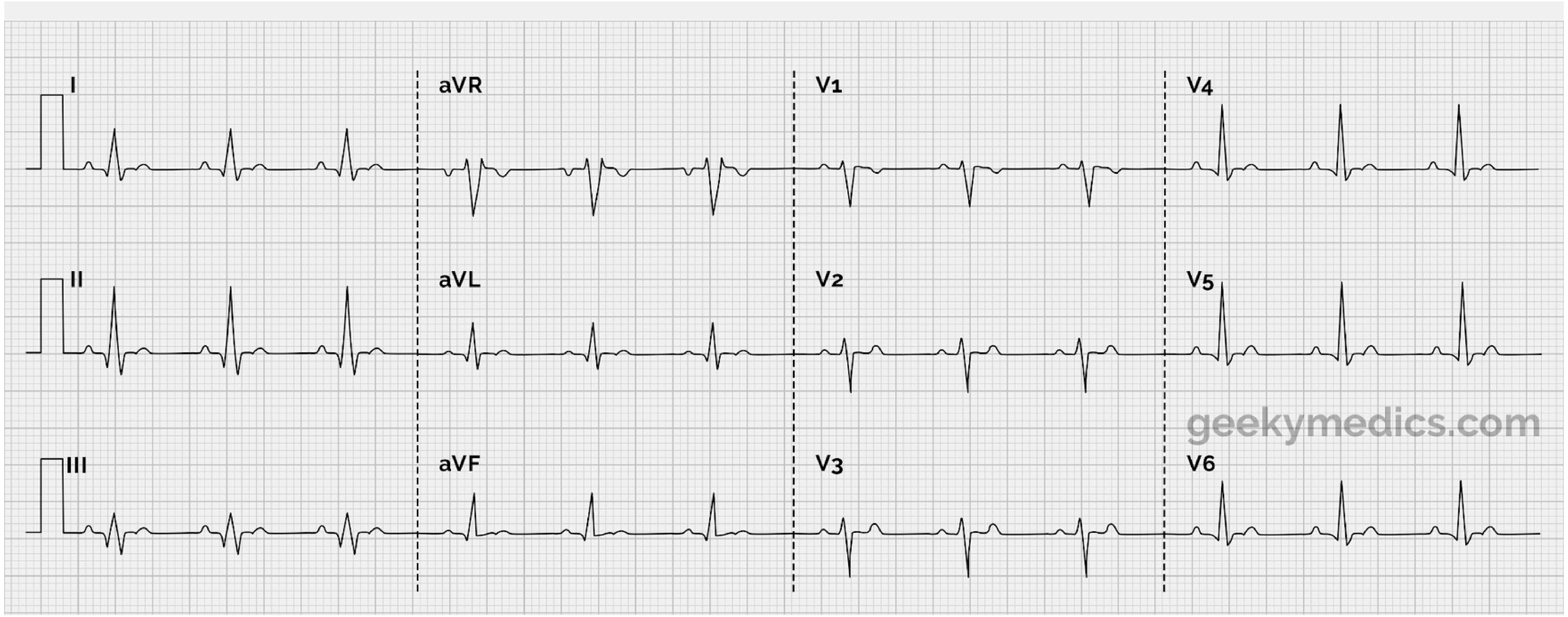
BUENAVENTURA, JUSTINE

ROSTATA, CARL

ANZURES, ANGEL

RESEARCH OBJECTIVES

1. **To collect a valid and comprehensive dataset** from online datasets, clinics, or hospitals for developing the model.
2. To build a suitable **machine learning model** for analyzing 12-lead ECG digital image data using deep learning architecture.
3. To develop a **web application** that integrates web security features and a deployed final deep learning model for diagnosing Atrial Fibrillation, facilitating data input, and model inference.
4. To **test and validate the accuracy** of the model in clinical application with the help of a cardiologist, specifically in diagnosing Atrial Fibrillation.



SCOPE

- Develops a deep learning model to detect Atrial Fibrillation (AF) from 12-lead ECG images.
- Uses data solely from Dr. Gracita Topacio’s clinic.
- Includes model training, preprocessing, and deployment in a web application.

LIMITATIONS

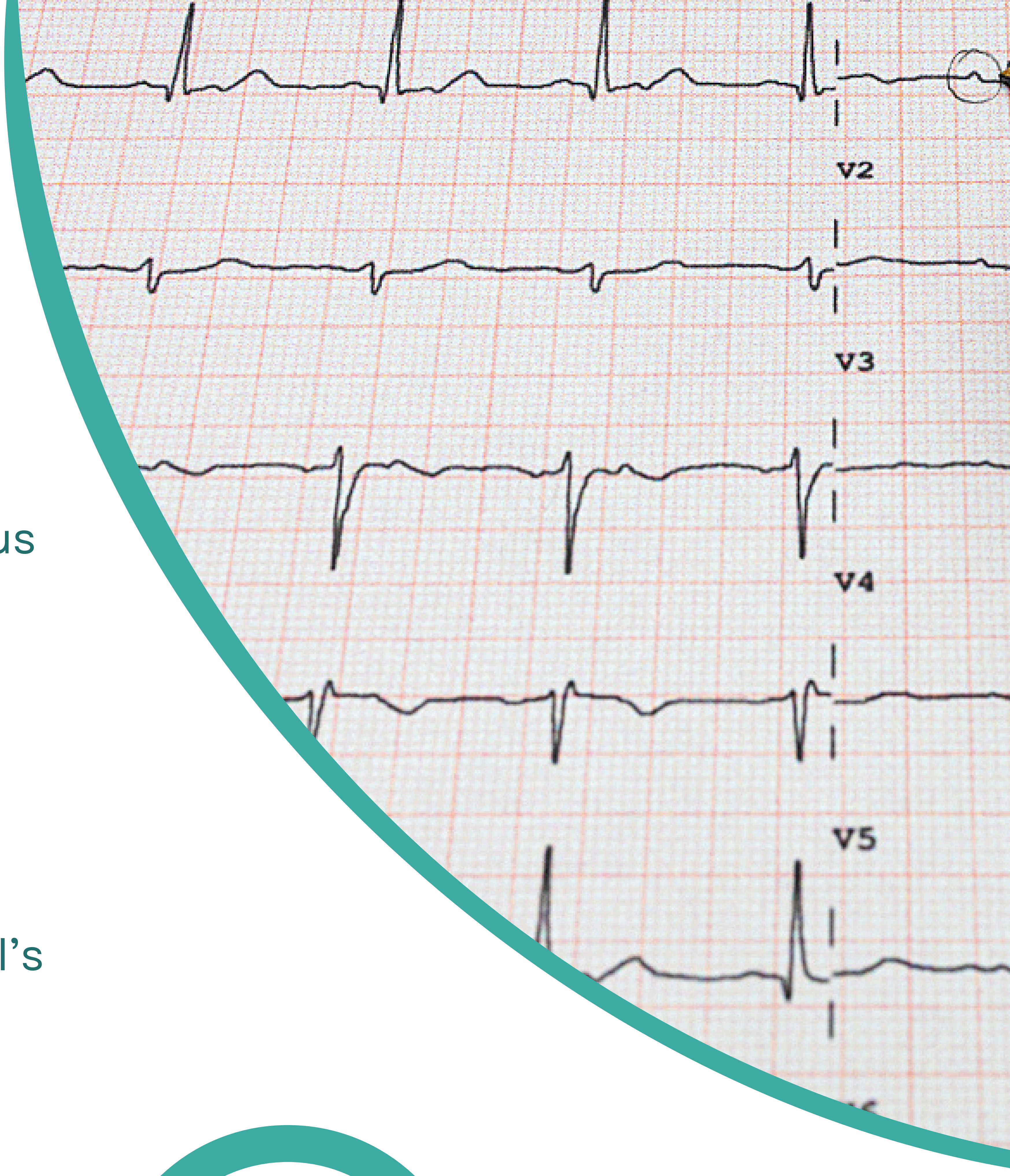
- Limited dataset from a single source may affect generalizability.
- Focuses only on AF, excluding other heart conditions.
- Does not explore alternative predictive methods or clinical treatments.

BENEFICIARIES

- Non-Cardiologist Medical Experts
- Health Practitioners
- Medical Students
- Future Researchers

RESEARCH INSTRUMENT

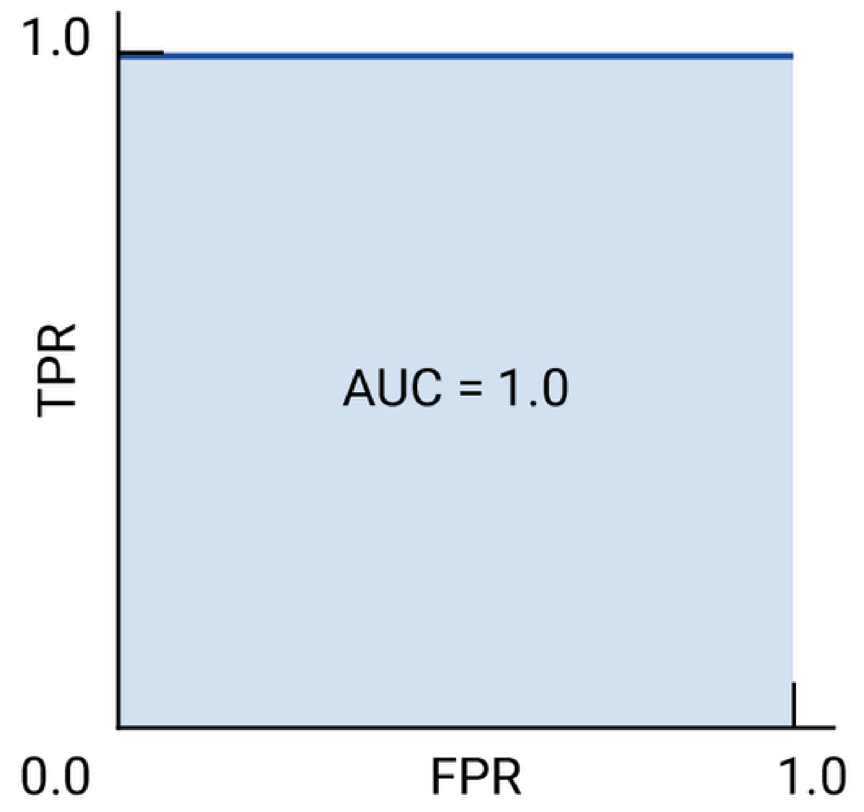
- **ECG Images:**
  - **Source:** ECG printouts at Dr. Gracita Topacio's clinic.
  - **Classes:** Atrial Fibrillation and normal sinus rhythm.
  - Captured using a mobile phone camera
- **Deep Learning Model:**
  - EfficientNetB0
- **Web Application:**
  - Web application secured through Laravel's built in security features.

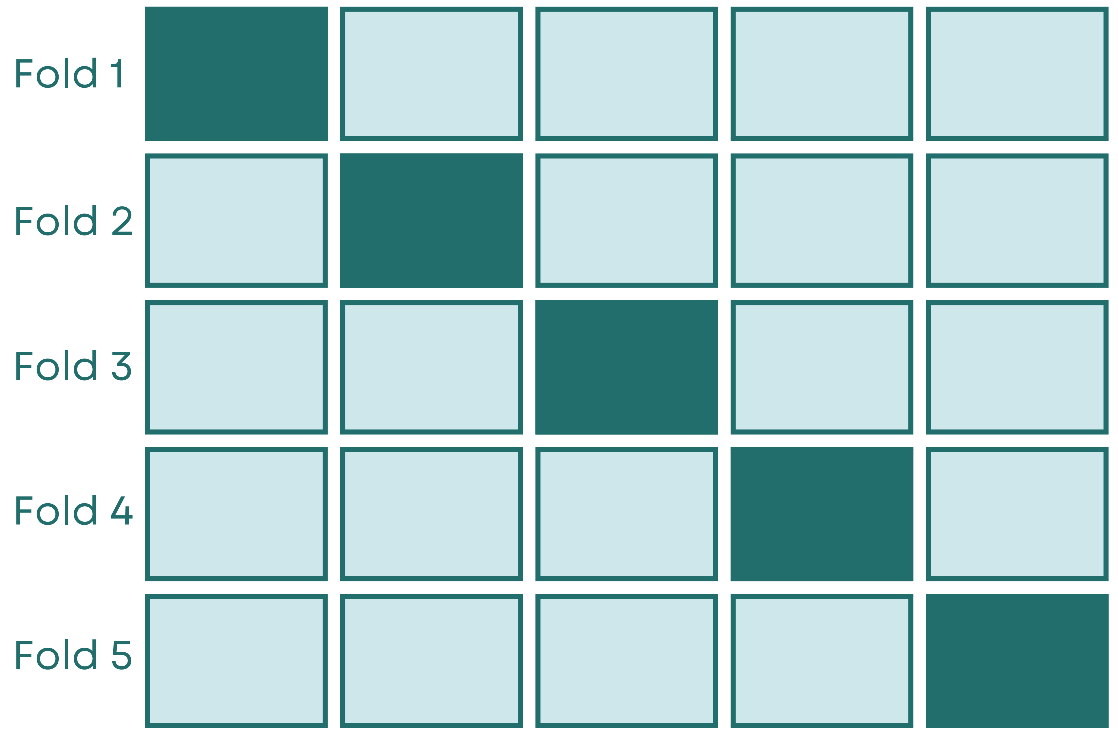


TESTING METHODOLOGY

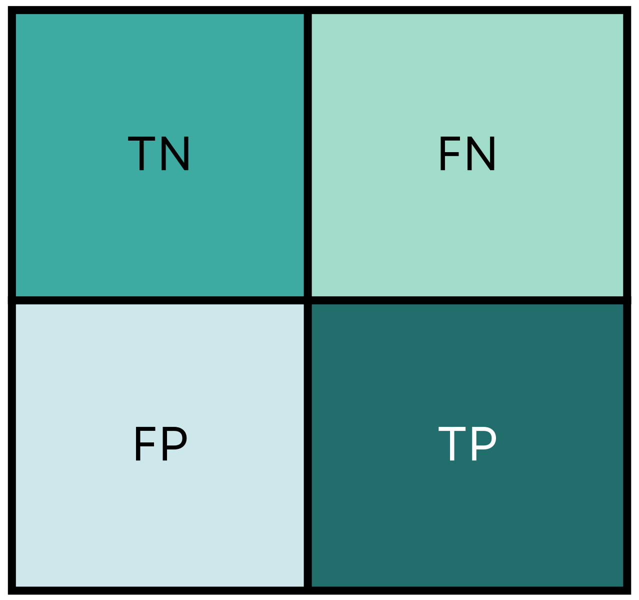
- Cross validation
- Performance Metrics
  - Accuracy
  - F1 Score
  - ROC-AUC
  - Confusion Matrix)
- Clinical Validation
- Out of Sample Testing

ROC-AUC



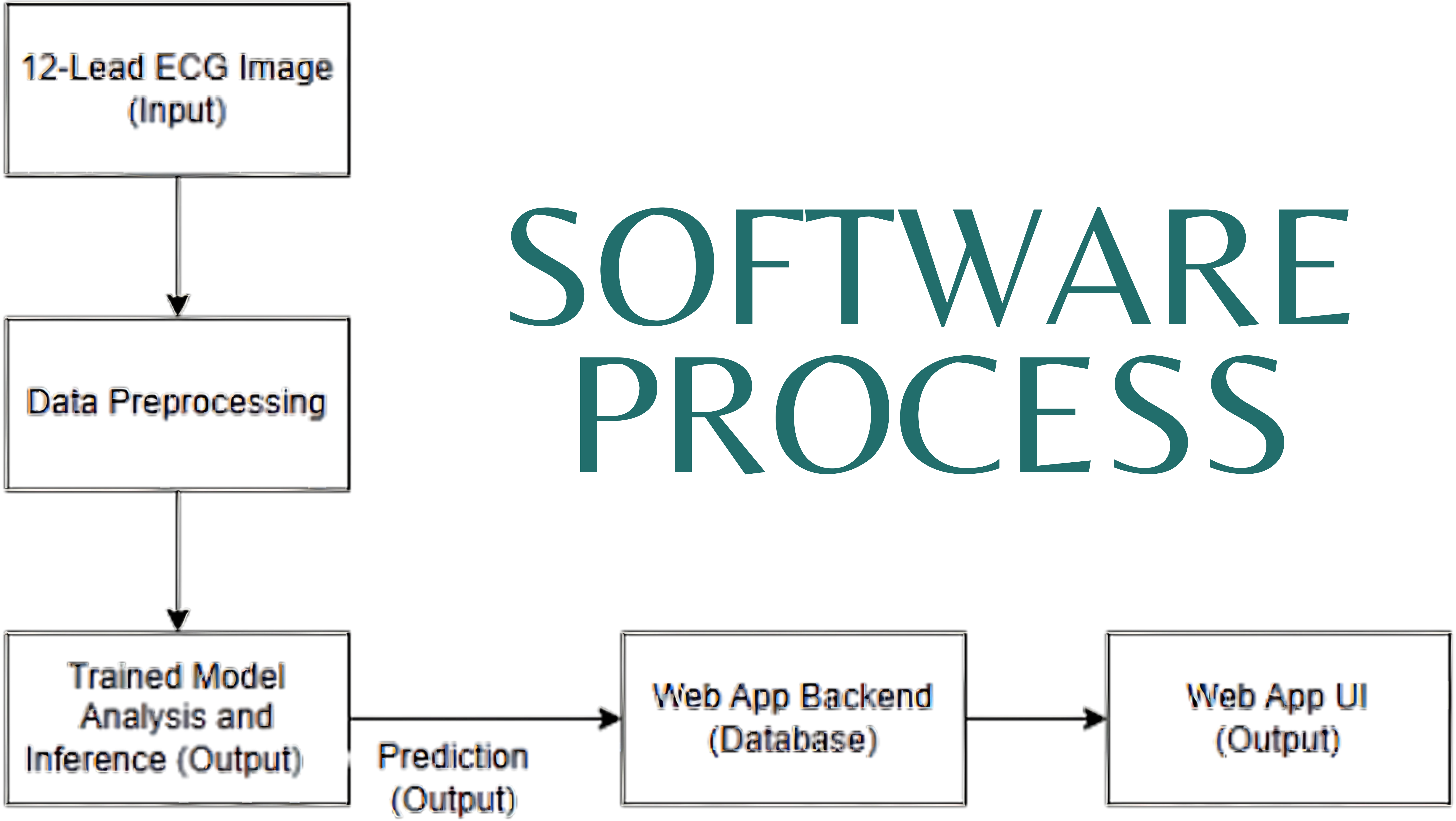


Cross validation

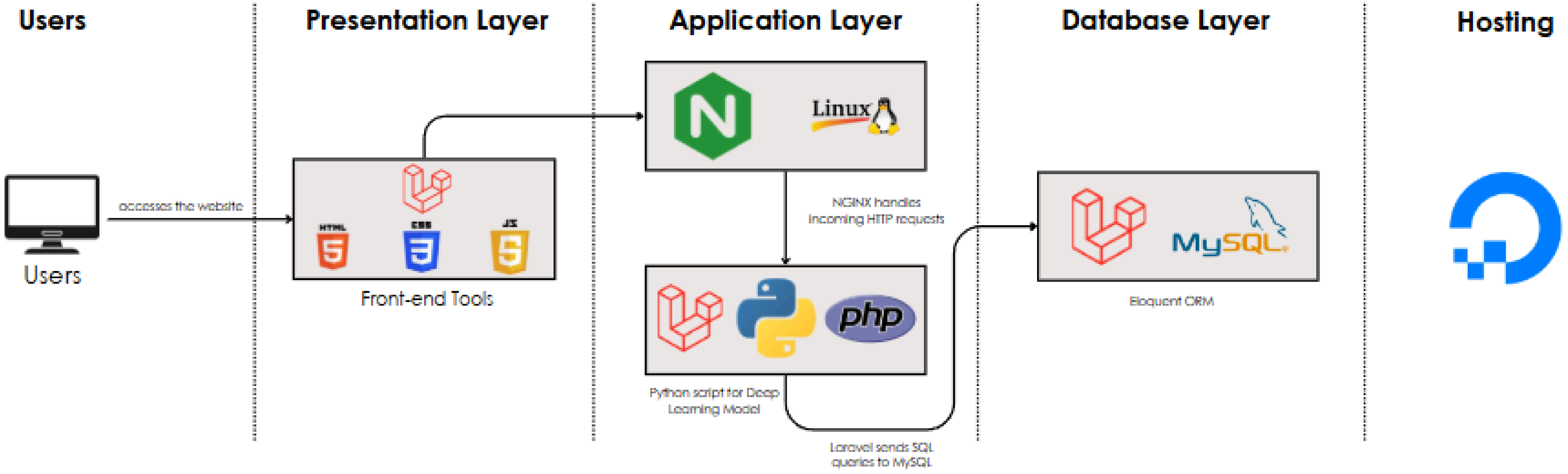


Confusion Matrix

SOFTWARE PROCESS



Three-tier Web Application Diagram



SECURITY

- **CSRF Protection -** Prevents unauthorized commands from being sent from a user's browser.
- **SQL Injection Protection** - Prevents malicious SQL code from being executed in your database.
- **Password Hashing** - Converts password into a one-way, fixed-length string of characters.
- **XSS Protection** (Cross-Site Scripting)- Prevents browsers from interpreting user inputs
- **Authentication System** - Middleware ensures users are logged in.

SUMMARY OF PANEL REVIEWS

ENGR. KATHLEEN VILLANUEVA

- Make a deeper study on the characteristics of heart sounds
- Include real time classification and interpretation
- Include at least two VHD
- Use the best algorithm from RRL
- Revise Chapter 3 and 4
- Consult with doctors

ENGR. RON ANTHONY DELOS SANTOS

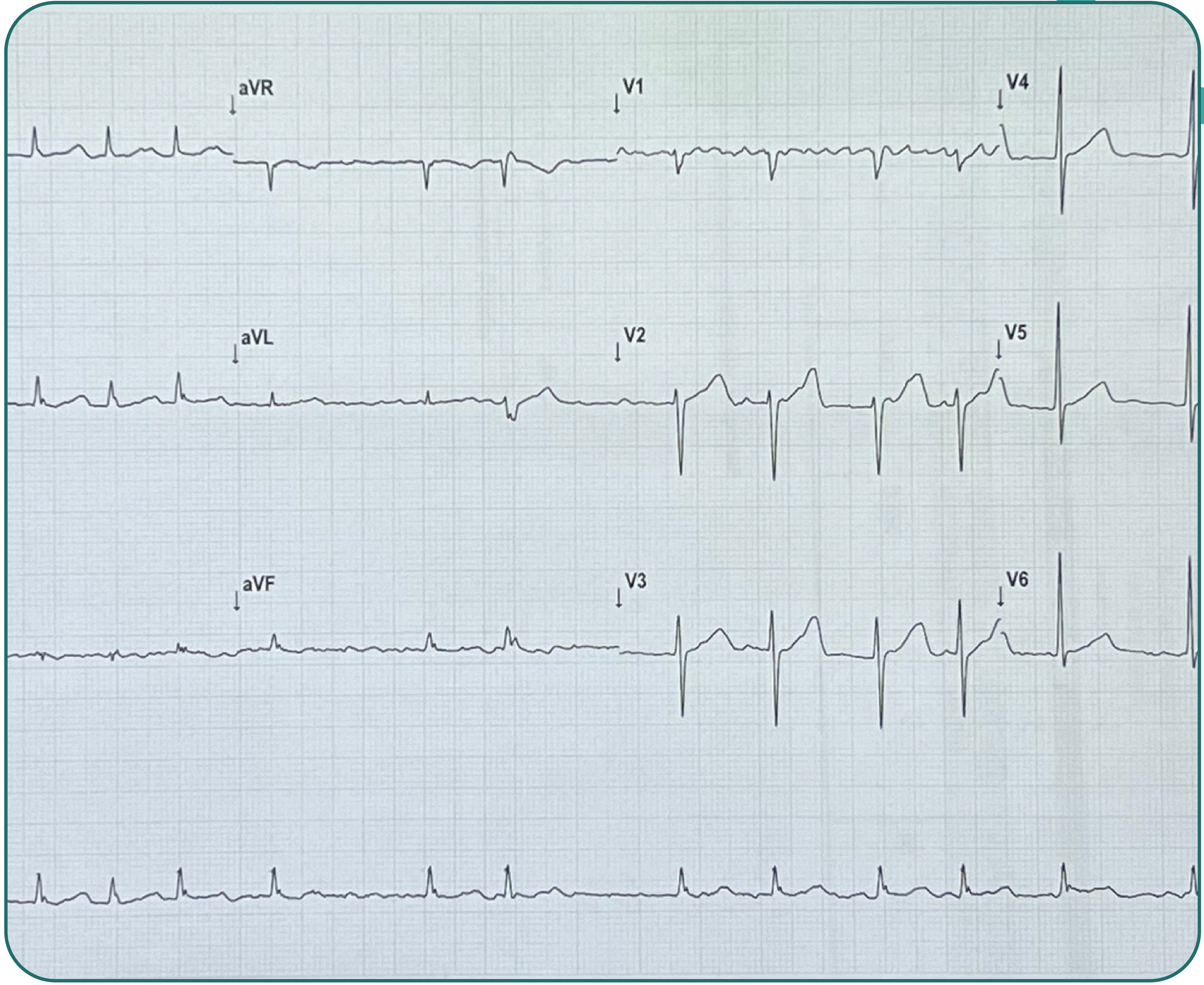
- Re-write your background of the study and statement of the problem.
- Rewrite your objective #3 with detailed feature of web application
- Include a database for patient record, data records, etc.
- Rewrite your obj 4 - accuracy of diagnosis
- Revise your title
- Add obj on security.

ENGR. MICHAEL OLIVO

- Create web application for heart valve disease.
- Include module for patient, doctor
- Payment module in methodology
- Generate reports
- Use Reactjs for Web application
- Use Chartjs for charts

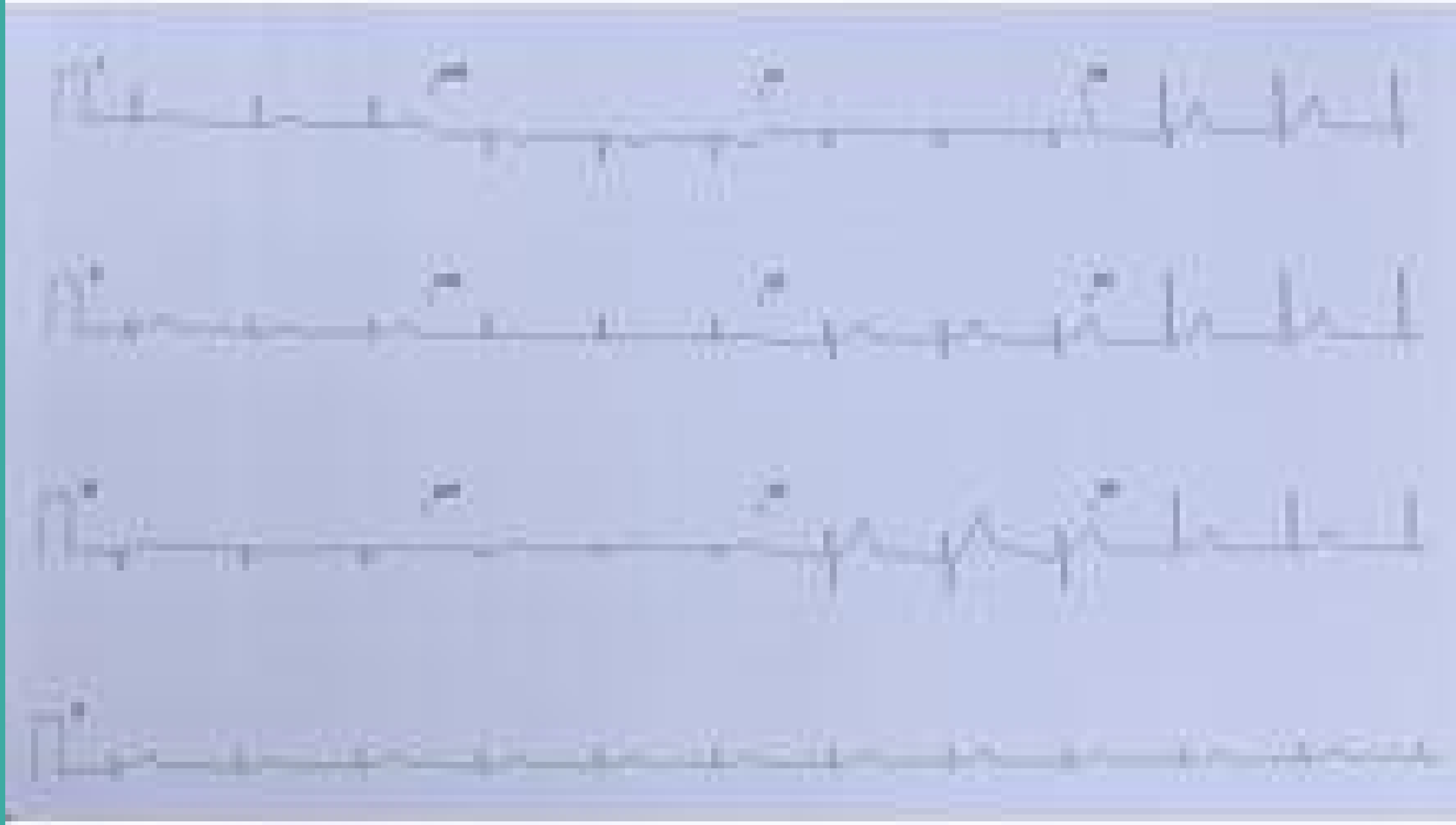
DATASET

- ECG Images:
  - Atrial Fibrillation
  - Normal Sinus Rhythm
- Sourced from Dr. Topacio's clinic in Manila.
- Augmentations:
  - Random rotation
  - Brightness and shadow adjustments
  - Scaling
  - Horizontal flips.



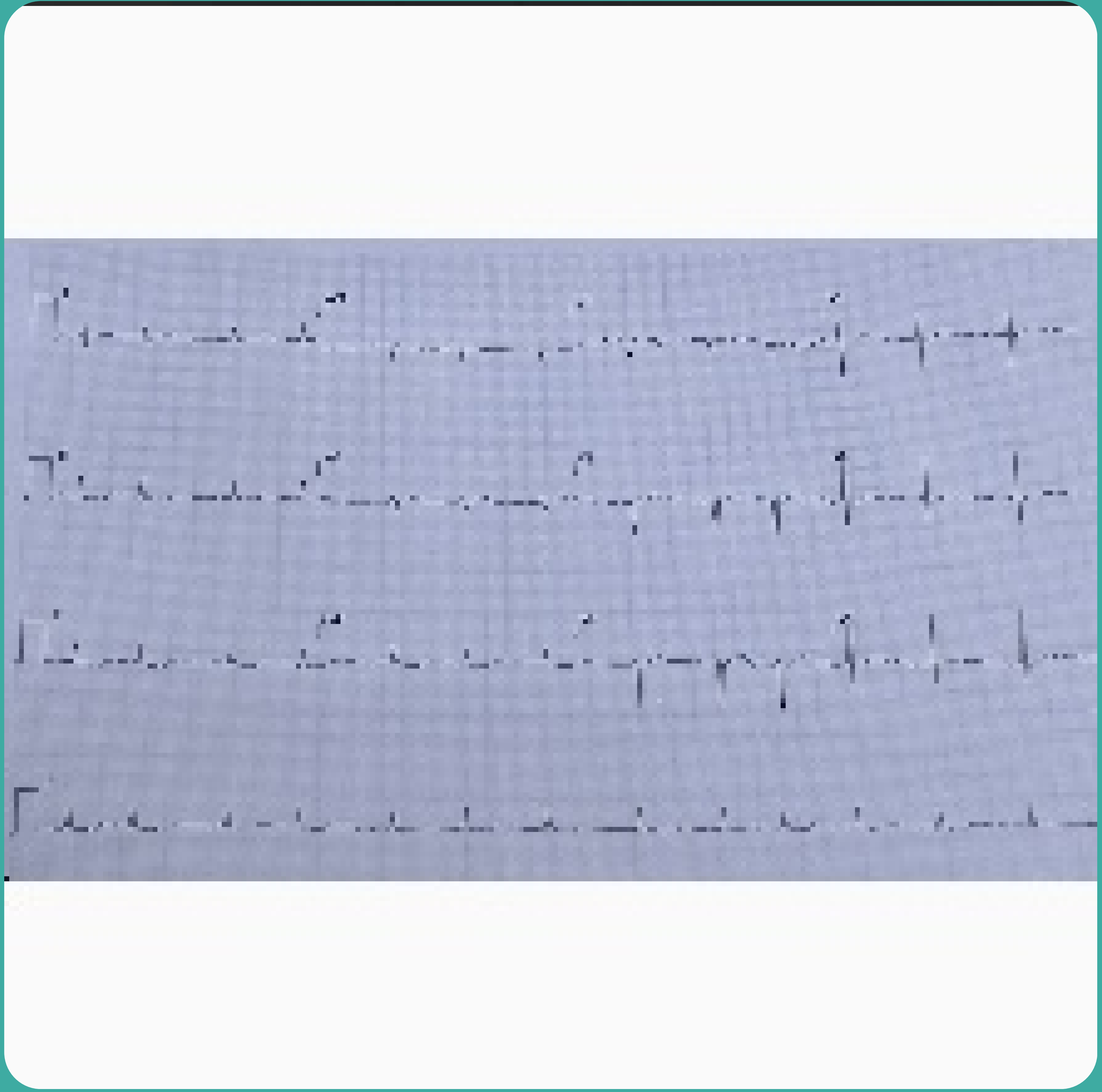
DATASET 1

- 73 Original Images
- 365 Augmented
- No preprocessing
- Padded to squares
- resized to 224x224



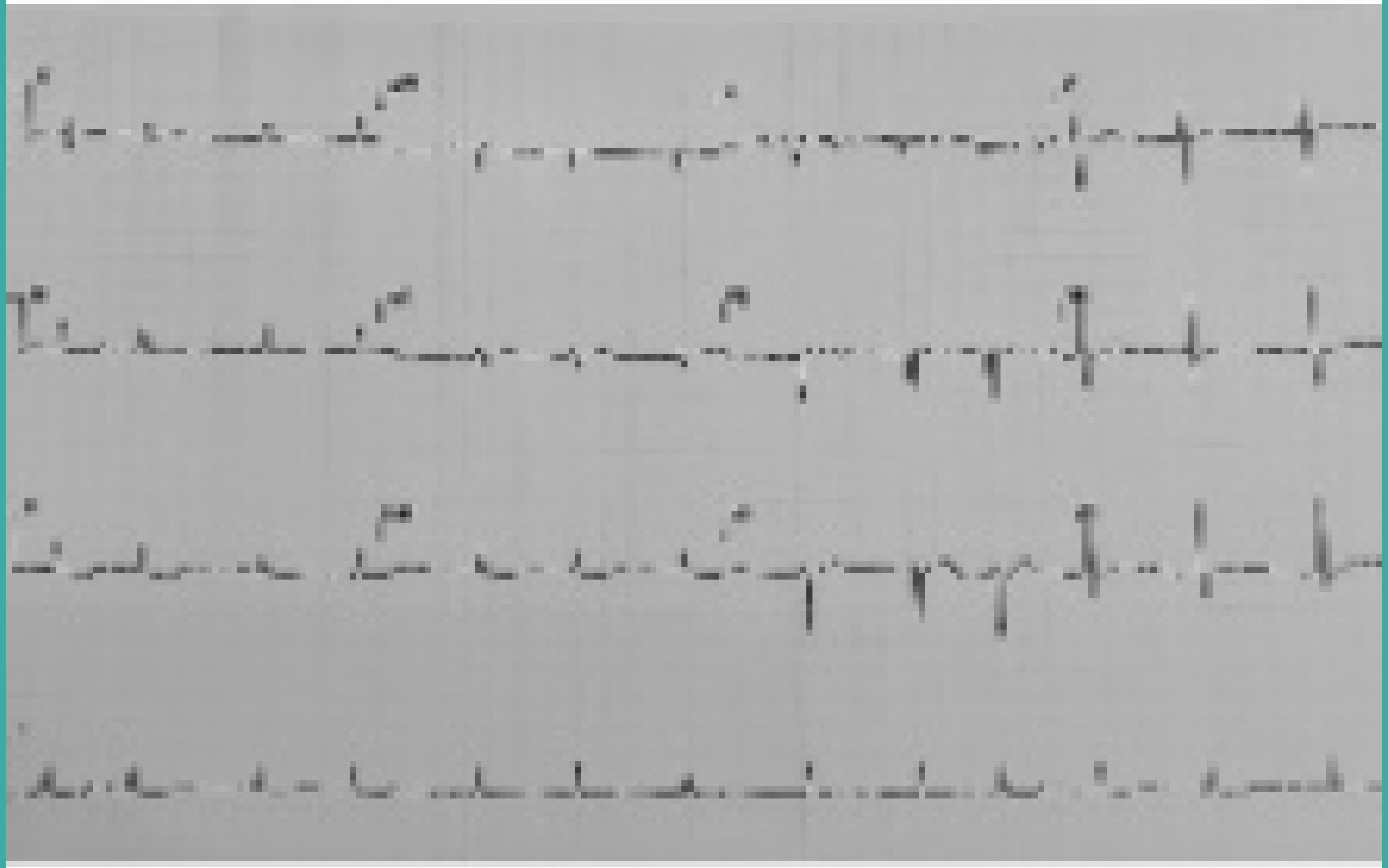
DATASET 2

- 121 Original Images
- 1936 Augmented
- No preprocessing
- Padded to squares
- resized to 224x224



DATASET 3

- Dataset 2 With Preprocessing
    - Grayscale conversion
    - Noise reduction via fastNlMeansDenoising
    - Morphological erosion
    - Padding to square dimensions



MODEL EVALUATION

Metric	ResNet18	VGG16	MobileNetV2	EfficientNet-B0	InceptionV3
Accuracy	89.44%	45.96%	94.41%	97.52%	97.52%
F1-Score	89.44%	62.98%	94.67%	97.33%	97.47%
ROC-AUC	97.34%	49.67%	98.78%	99.92%	99.57%
Sensitivity	97.30%	100.0%	95.95%	98.65%	96.25%
Specificity	82.76%	0.00%	97.70%	96.55%	98.77%

TOP 3

EFFICIENTNETB0

INCEPTIONV3

MOBILENETV2

MODEL EVALUATION

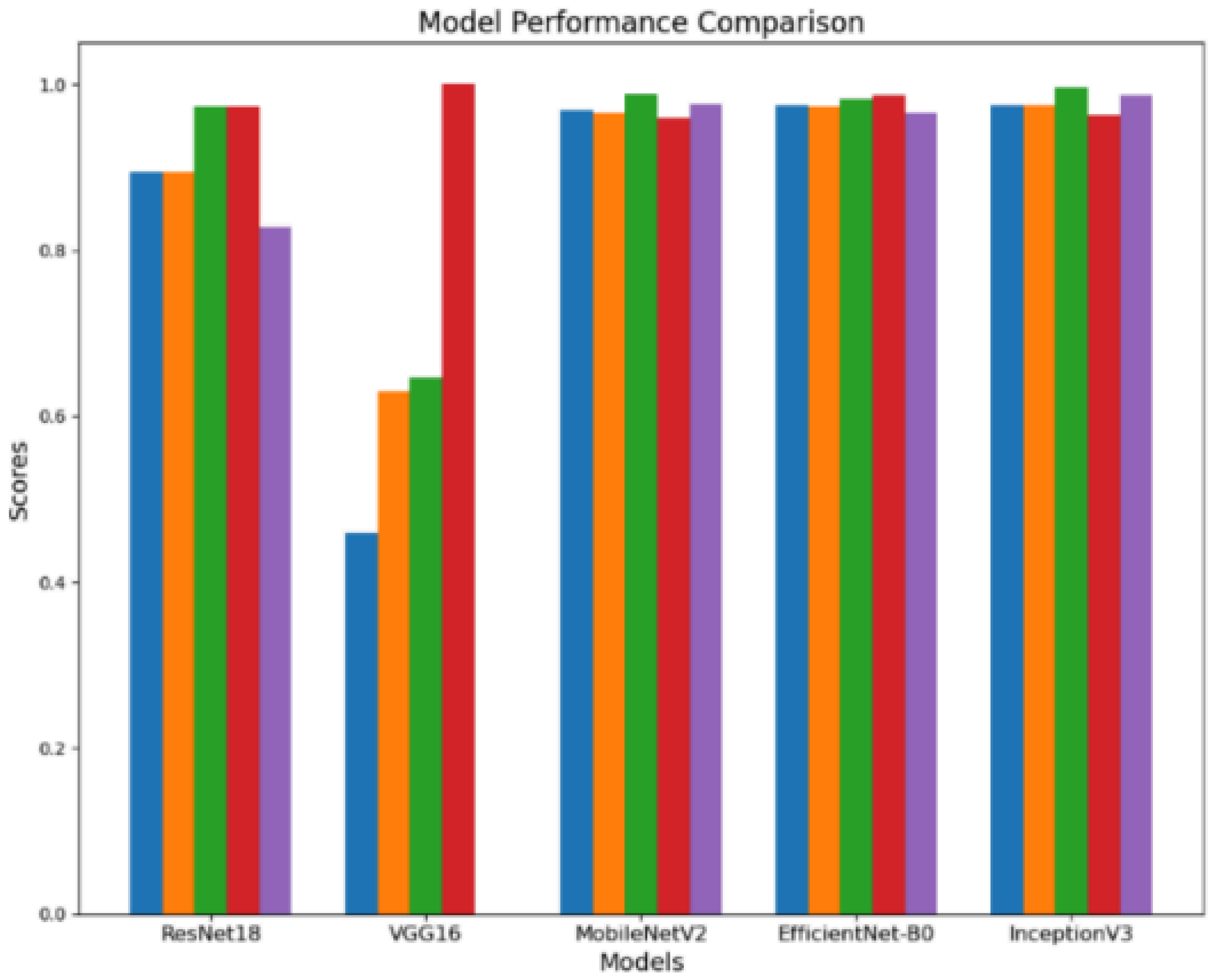
Accuracy

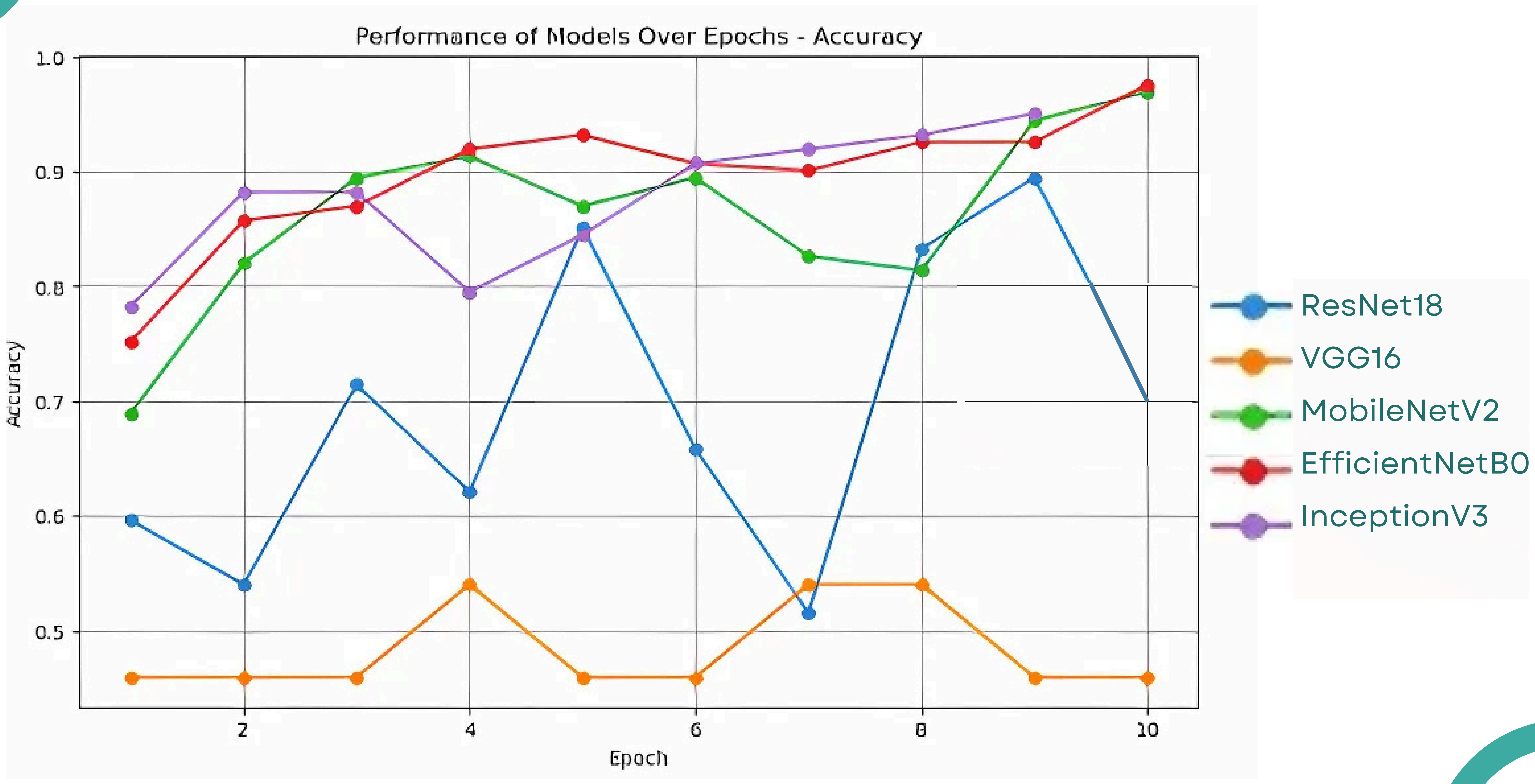
F1 Score

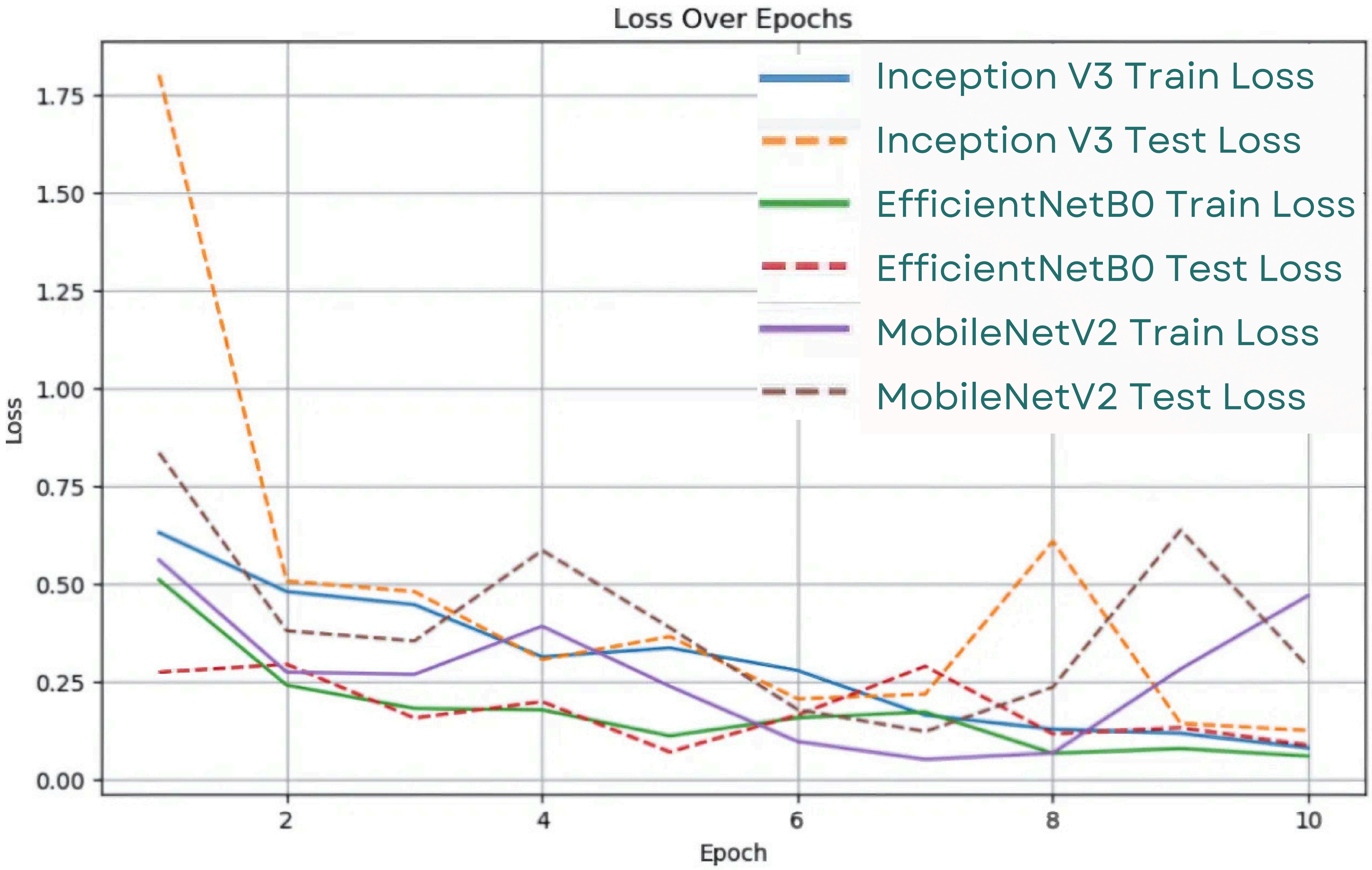
ROC-AUC

Sensitivity

Specificity



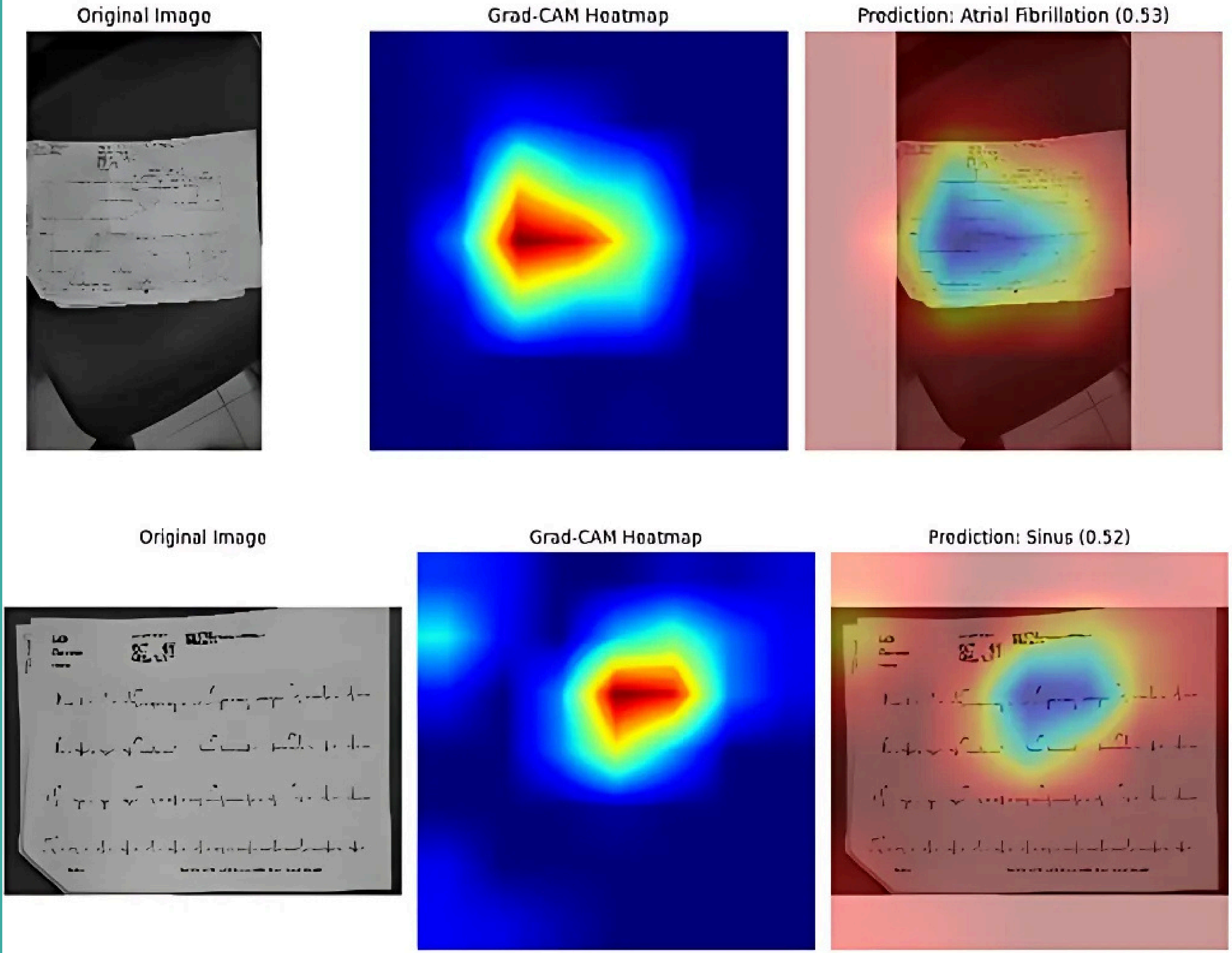

Performance of Models Over Epochs - Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
Accuracy
2
4
6
8
10
Epoch
ResNet18
VGG16
MobileNetV2
EfficientNetB0
InceptionV3


Loss Over Epochs
Inception V3 Train Loss
Inception V3 Test Loss
EfficientNetB0 Train Loss
EfficientNetB0 Test Loss
MobileNetV2 Train Loss
MobileNetV2 Test Loss
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Loss
2
4
6
8
10
Epoch

FIRST TRAINING ATTEMPT

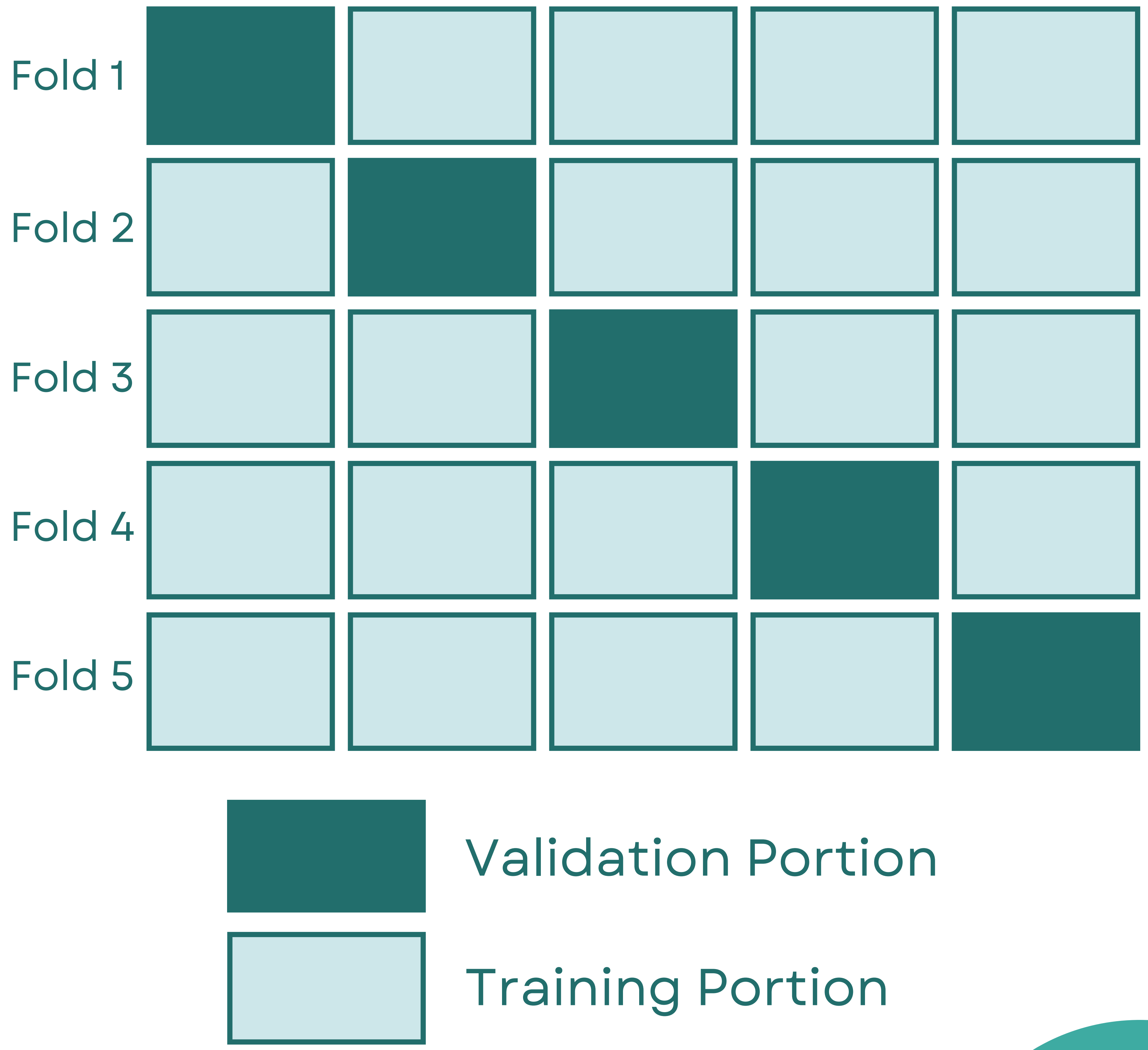
	Initial K Fold Model	Retraining With Dataset 2	Retraining With Dataset 3
Dataset	Dataset 1	Dataset 2	Dataset 3
Model	EfficientNetB0 with ImagenetV1 weights	Fold 3	Retrained Fold 3
Loss Function	Cross Entropy Loss	Cross Entropy Loss	Focal Loss
Optimizer	Adam (Learning Rate: 1e-4)	Adam (Learning Rate: 1e-5)	Adam (Learning Rate: 1e-5)
Frozen Layers	No	Blocks 0 to 3	No
Output	Best Fold: Fold 3	Retrained Fold 3.1	Retrained Fold 3.2
Tested on	Dataset 1	Dataset 2	Dataset 3
Performance	Accuracy: 89.26% F1 Score: 89.96%	Accuracy: 94% F1 Score: 88%	Accuracy: 95% F1 Score: 89.60%

FIRST ATTEMPT MODEL TESTING

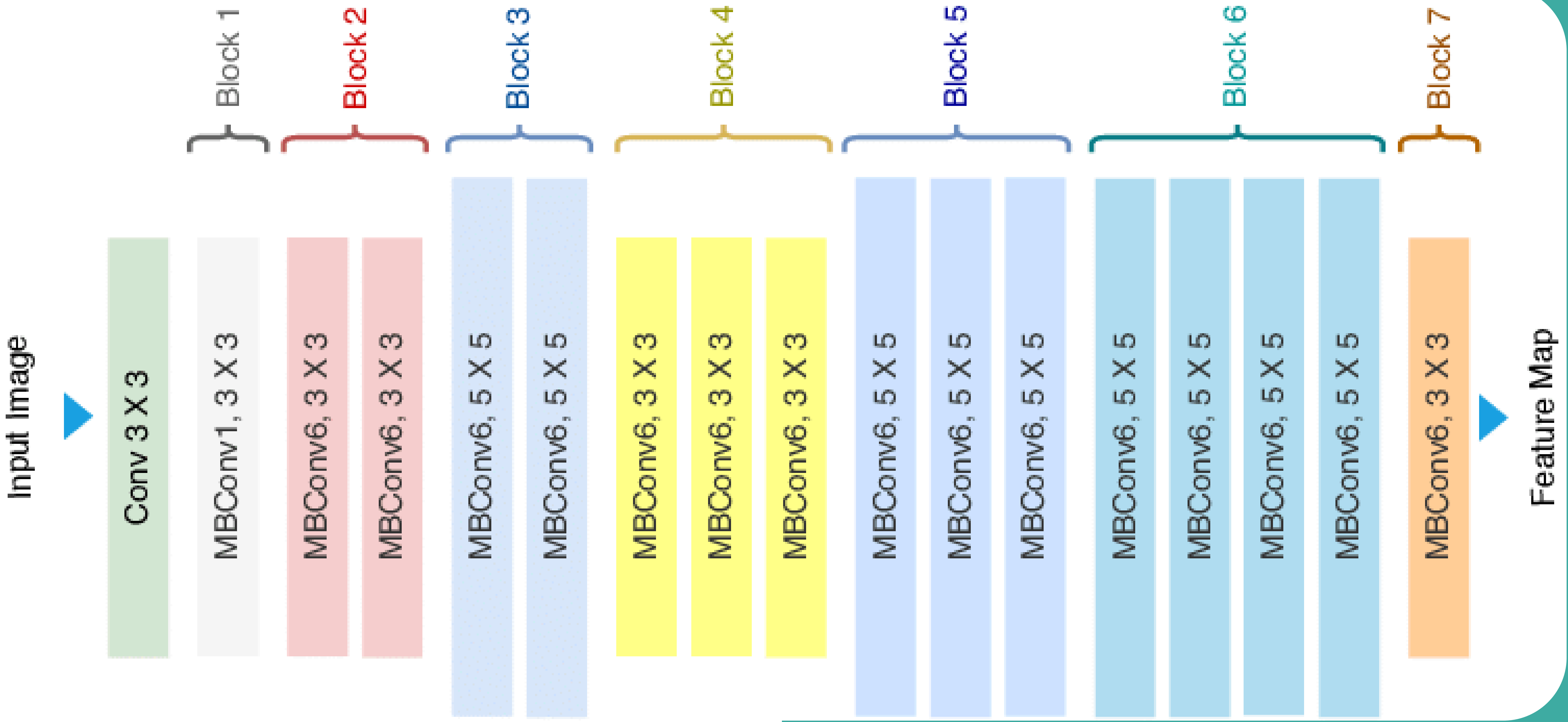


IDENTIFIED INNEFICIENT TRAINING METHOD

- When using KFold during training, Fold models are trained on only a portion of the data.
- Deploying a fold model reduces generalization and may cause bias or overfitting.
- KFold cross validation should only be used in testing the model in the dataset.

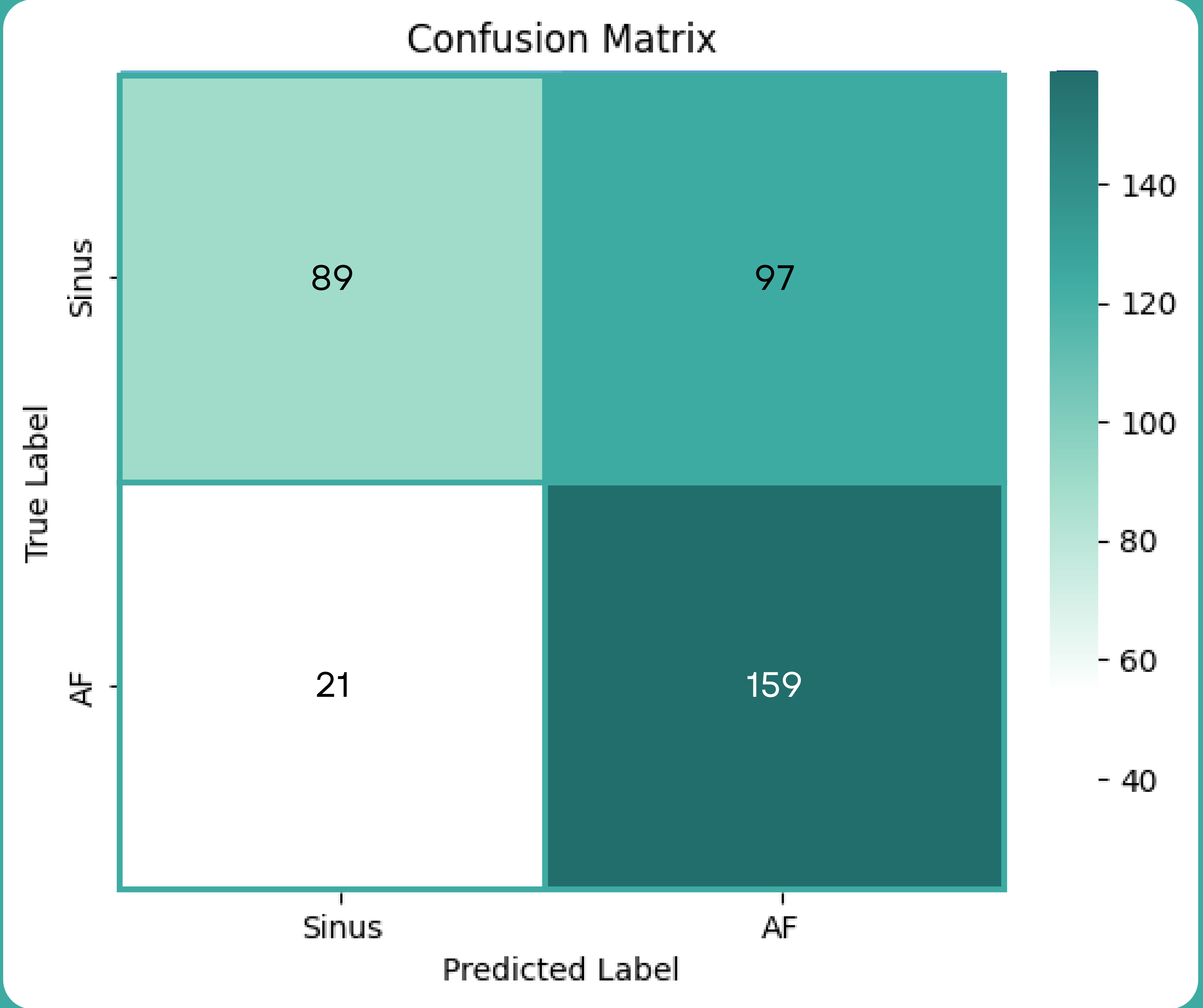


EFFICIENTNET BLOCKS



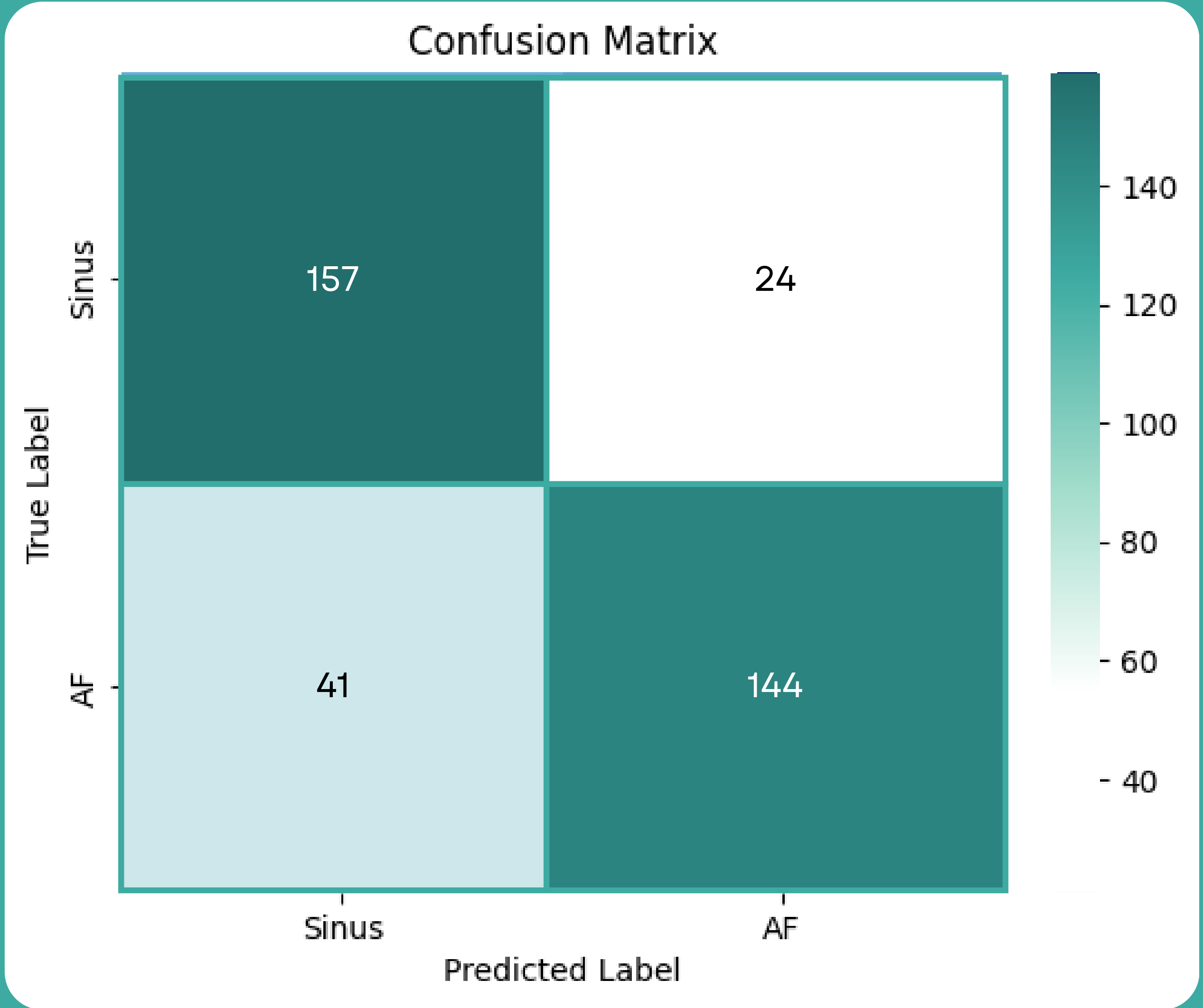
FINAL MODEL TRAINING

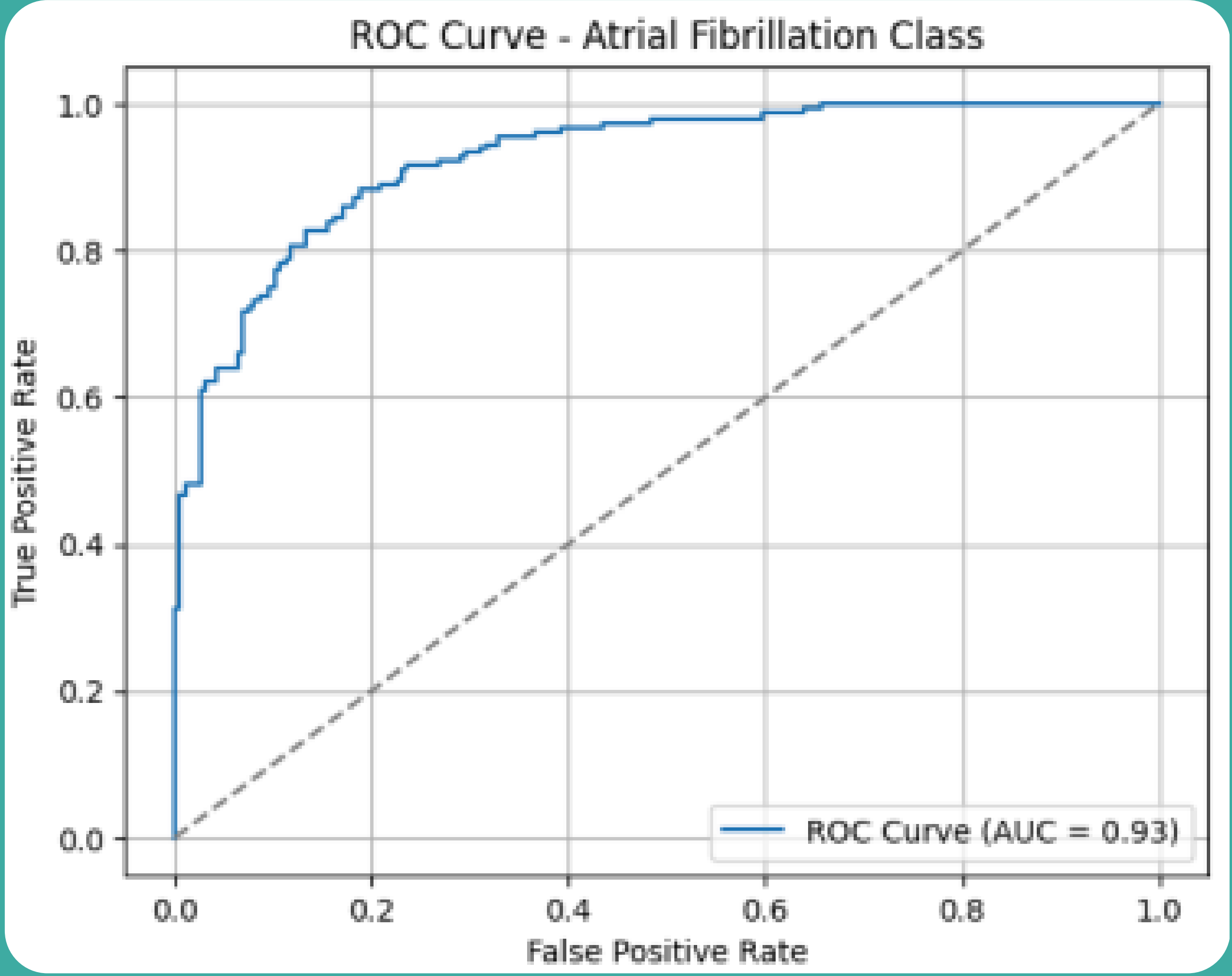
Phase 1	**Initial Full Dataset Model**
Dataset	Dataset 1
Model	EfficientNetB0 with ImagenetV1 weights
Loss Function	Cross Entropy Loss
Optimizer	Adam (Learning Rate: 1e-4)
Frozen Layers	No
Output	Initial Full Model
Tested on	Dataset 1
Performance	Accuracy: 68% F1 Score: 66.5%



FINAL MODEL TRAINING

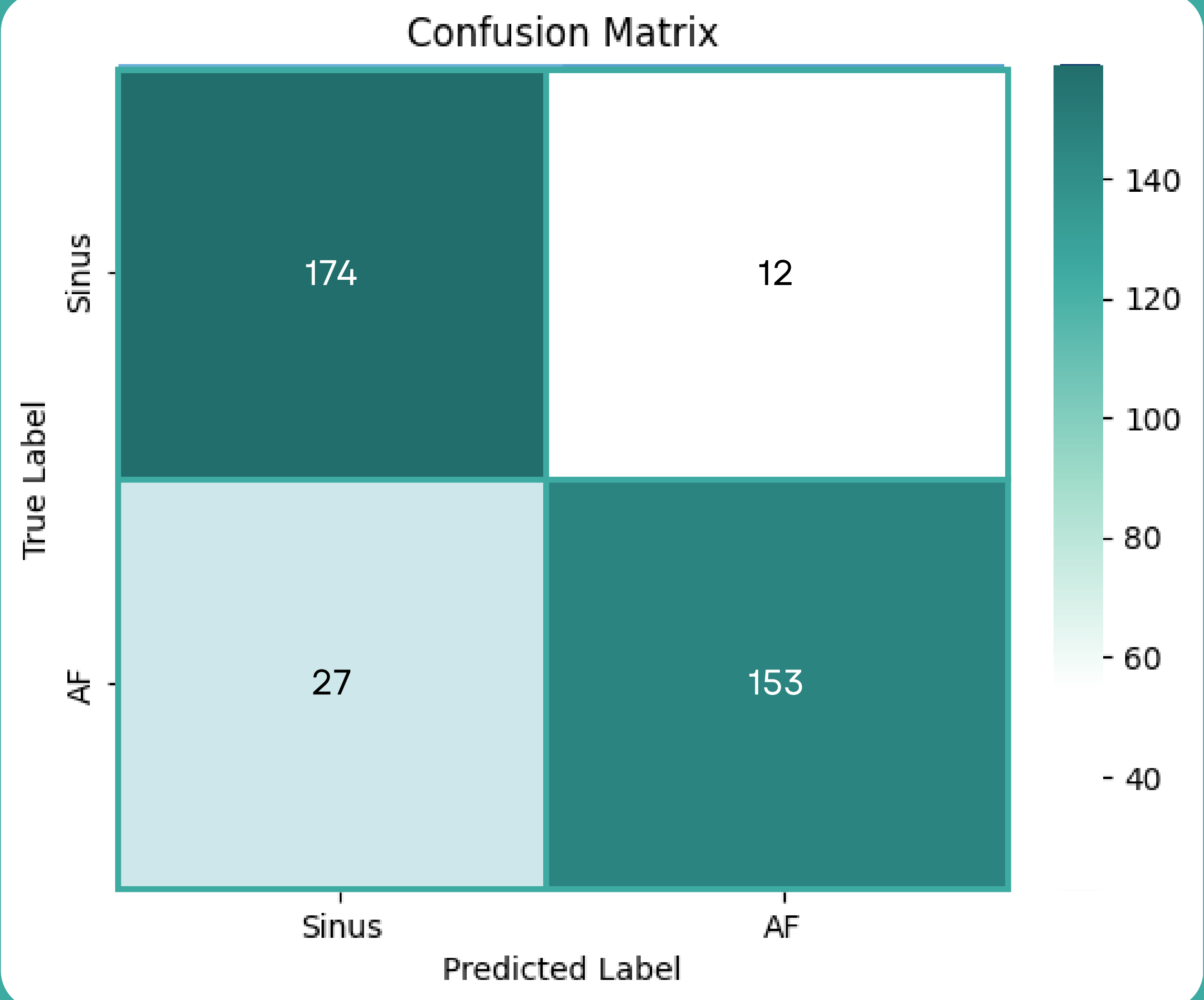
Phase 2	**Retraining with Dataset 3**
Dataset	Dataset 3
Model	Initial Full Model
Loss Function	Focal Loss (Sinus)
Optimizer	Adam (Learning Rate: 1e-5)
Frozen Layers	Block 0-3
Output	Retrained Full Model
Tested on	Dataset 3
Performance	Accuracy: 82% F1 Score: 82%

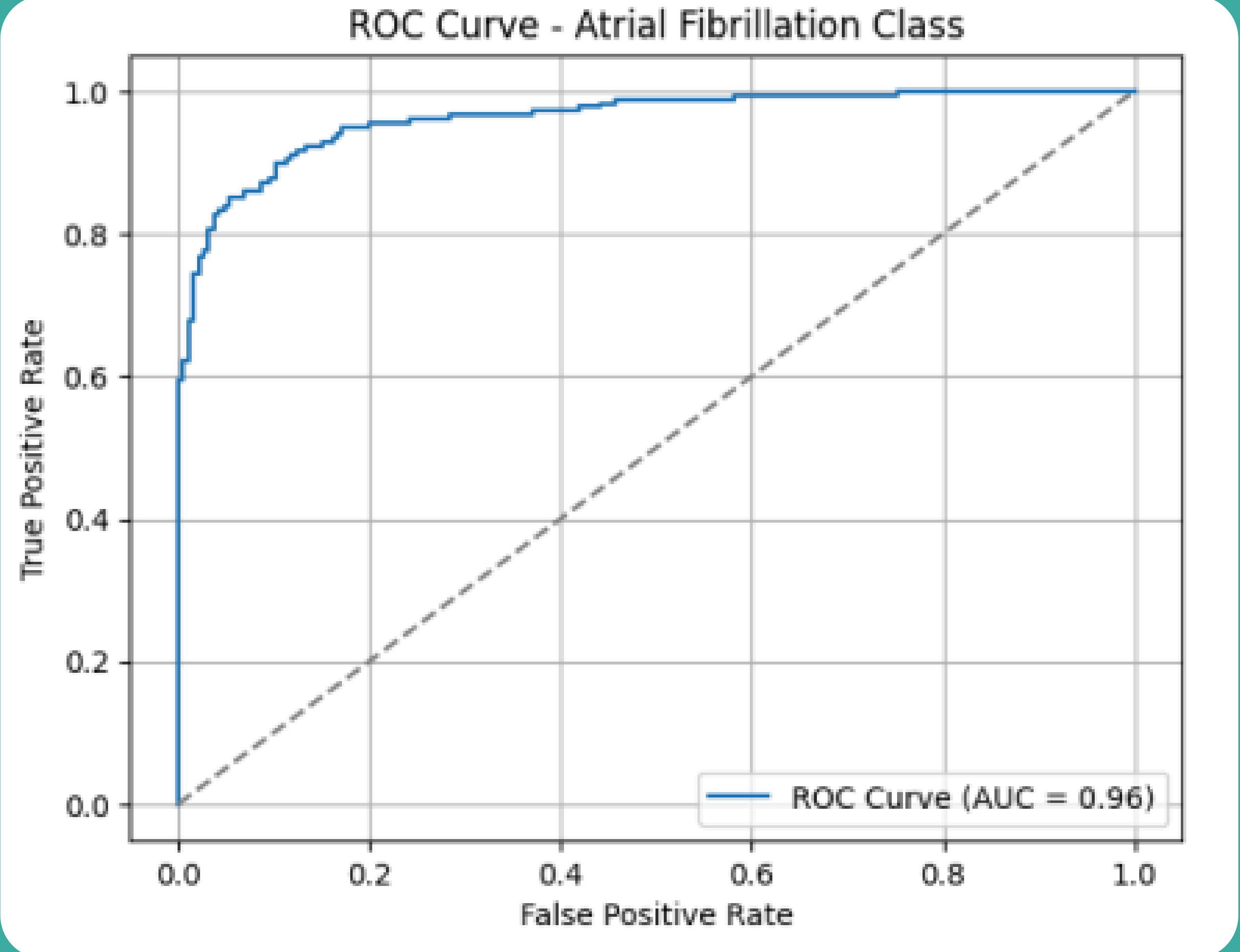




FINAL MODEL TRAINING

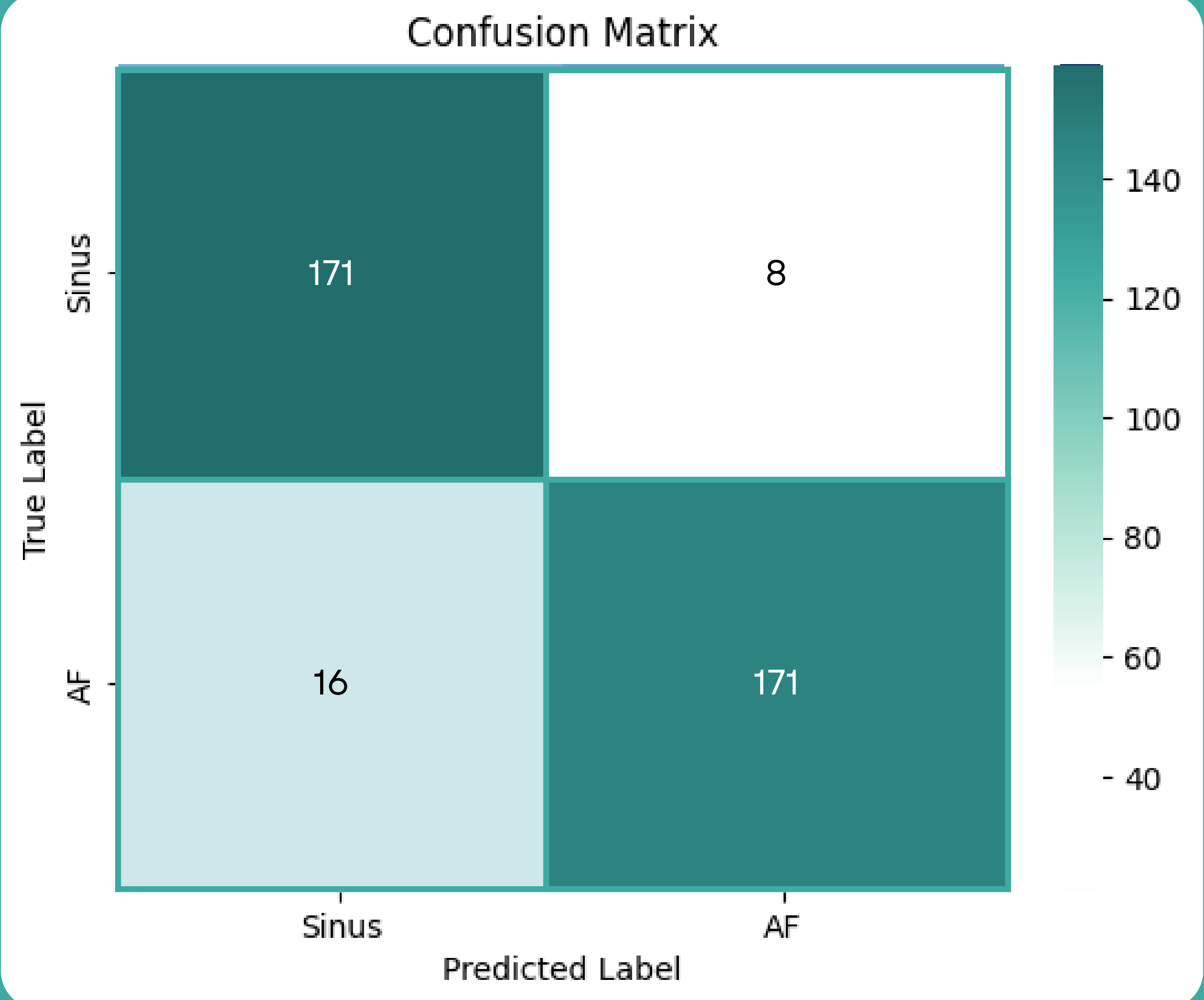
Phase 3	**Fine-Tuning Retrained Model**
Dataset	Dataset 3
Model	Retrained Full Model
Loss Function	Focal Loss (Sinus)
Optimizer	Adam (Layer wise LR) ○ 5e-6 backbone ○ 1e-4 classifier head
Frozen Layers	Block 0-2
Output	Fine Tuned Full Model
Tested on	Dataset 3
Performance	Accuracy: 89% F1 Score: 89%

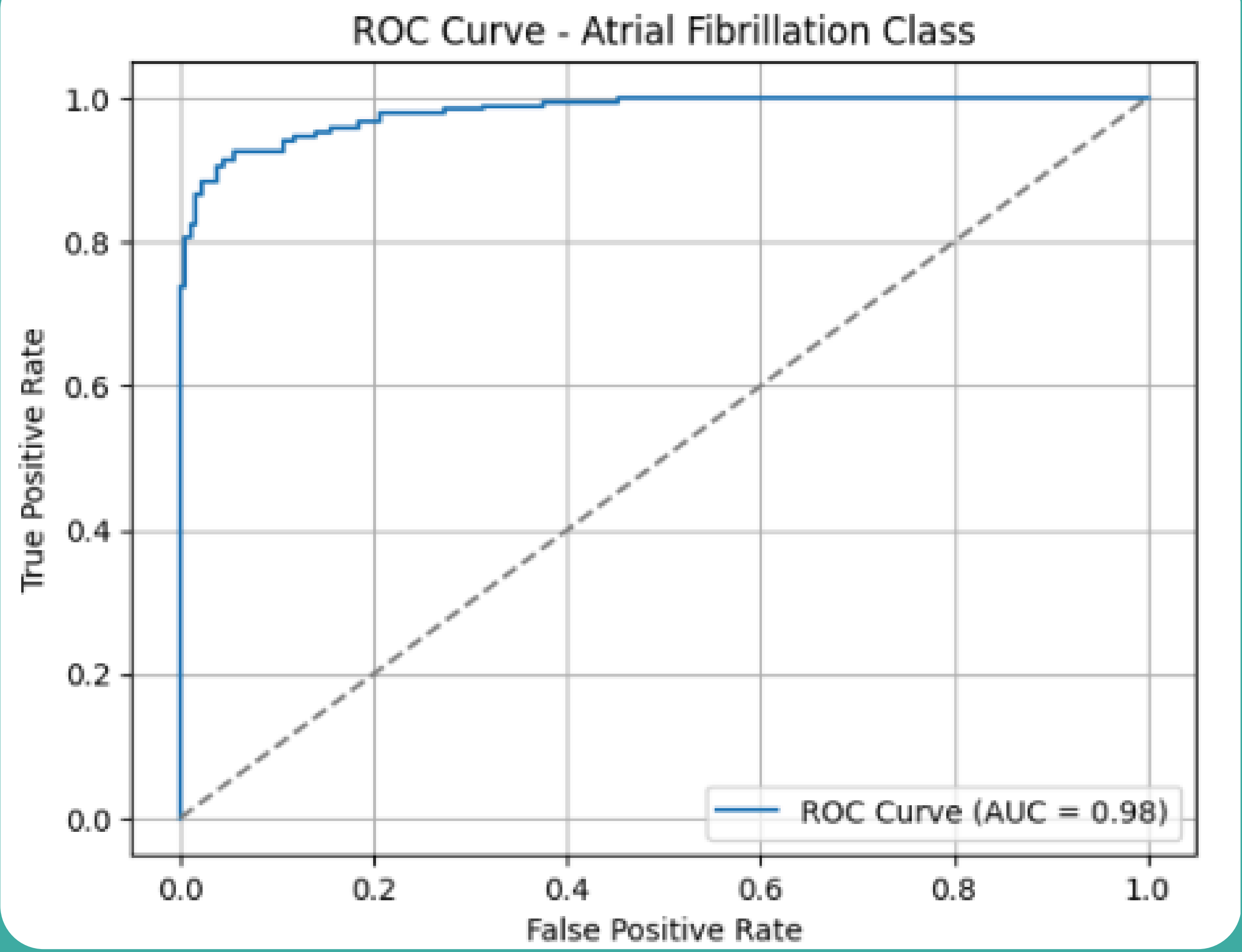




FINAL MODEL TRAINING

Phase 4	**Final Fine-Tuned Model**
Dataset	Dataset 3 Training
Model	Fine Tuned Full Model
Loss Function	Cross Entropy Loss with Label Smoothing
Optimizer	Adam (Learning Rate: 1e-6)
Frozen Layers	Block 0-2
Output	Final Full Model
Tested on	Dataset 3 Validation
Performance	Accuracy: 93% F1 Score: 93%





MODEL DEVEELOPMENT TIMELINE

START HERE

Model Evaluation

TOP 3

EfficientnetB0

InceptionV3

MobileNetV2

Initial KFold Training

Accuracy: 89.26%

F1 Score: 89.96%

Retraining Kfold with Dataset 2

Accuracy: 94%

F1 Score: 88%

Retraining Kfold with Dataset 2

Accuracy: 95%

F1 Score: 89.60%

Identified Inefficient Training Method

START HERE

Initial Full Model Training

Accuracy: 68%

F1 Score: 66.5%

Retrained Model

Accuracy: 82%

F1 Score: 82%

Fine Tuned Model

Accuracy: 89%

F1 Score: 89%

Final Deployed Model

Accuracy: 93%

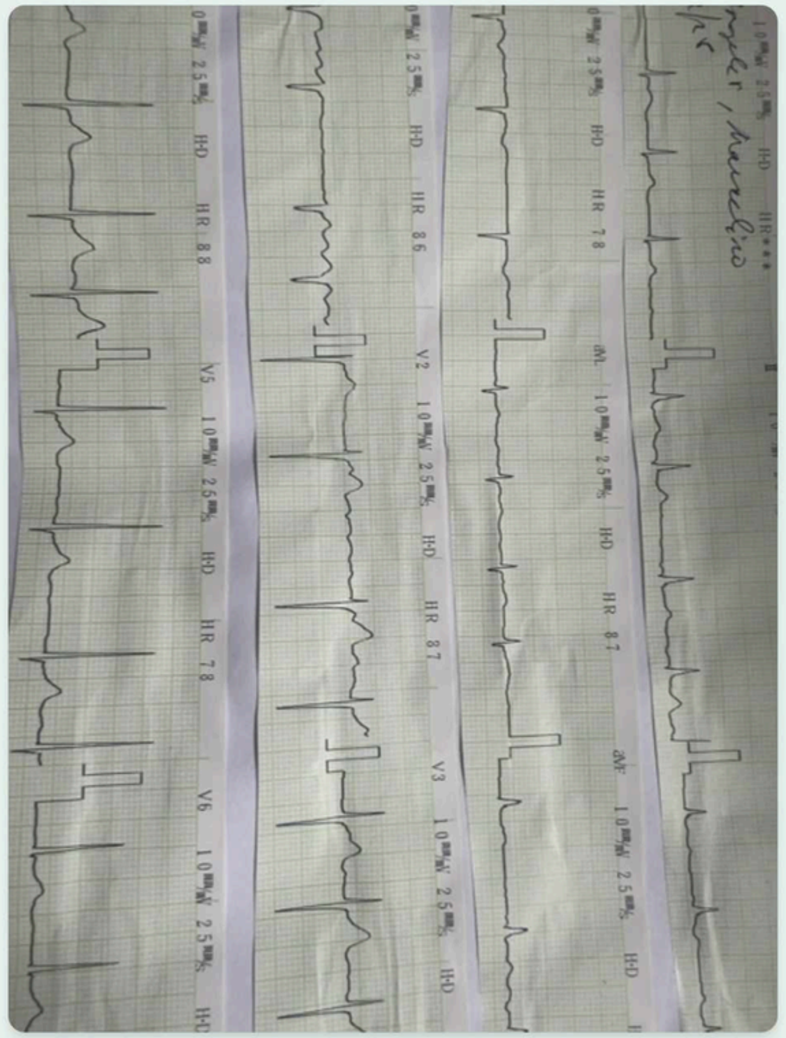
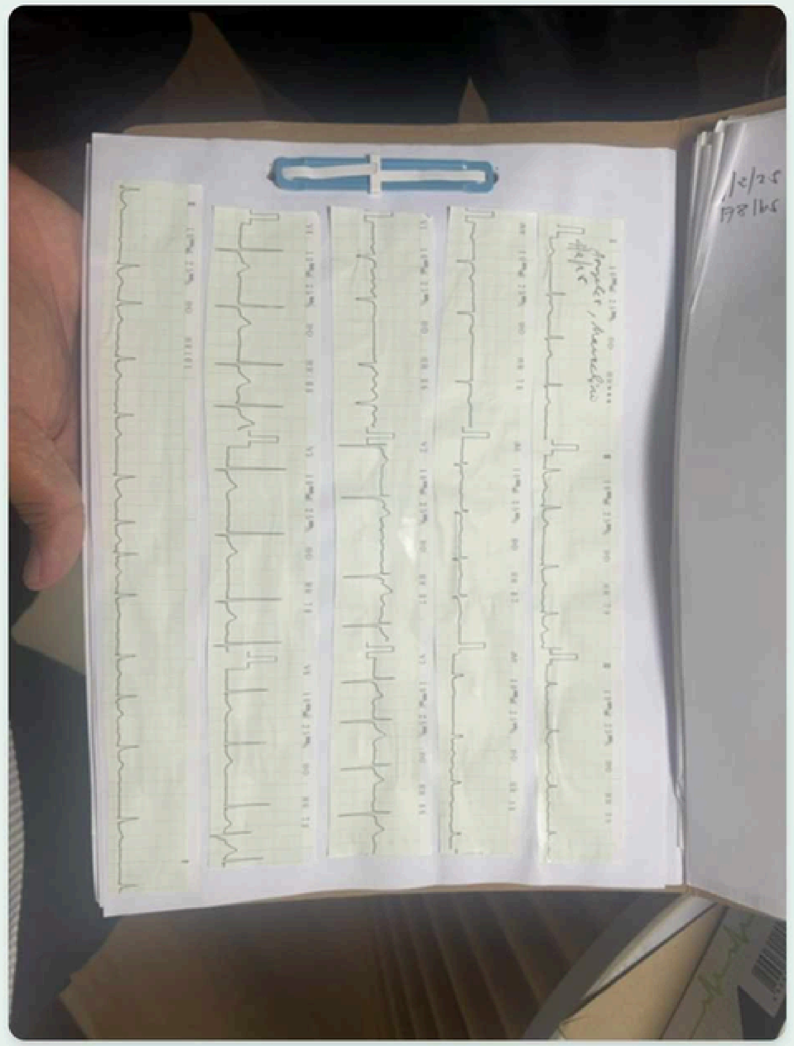
F1 Score: 93%

REAL WORLD TESTING

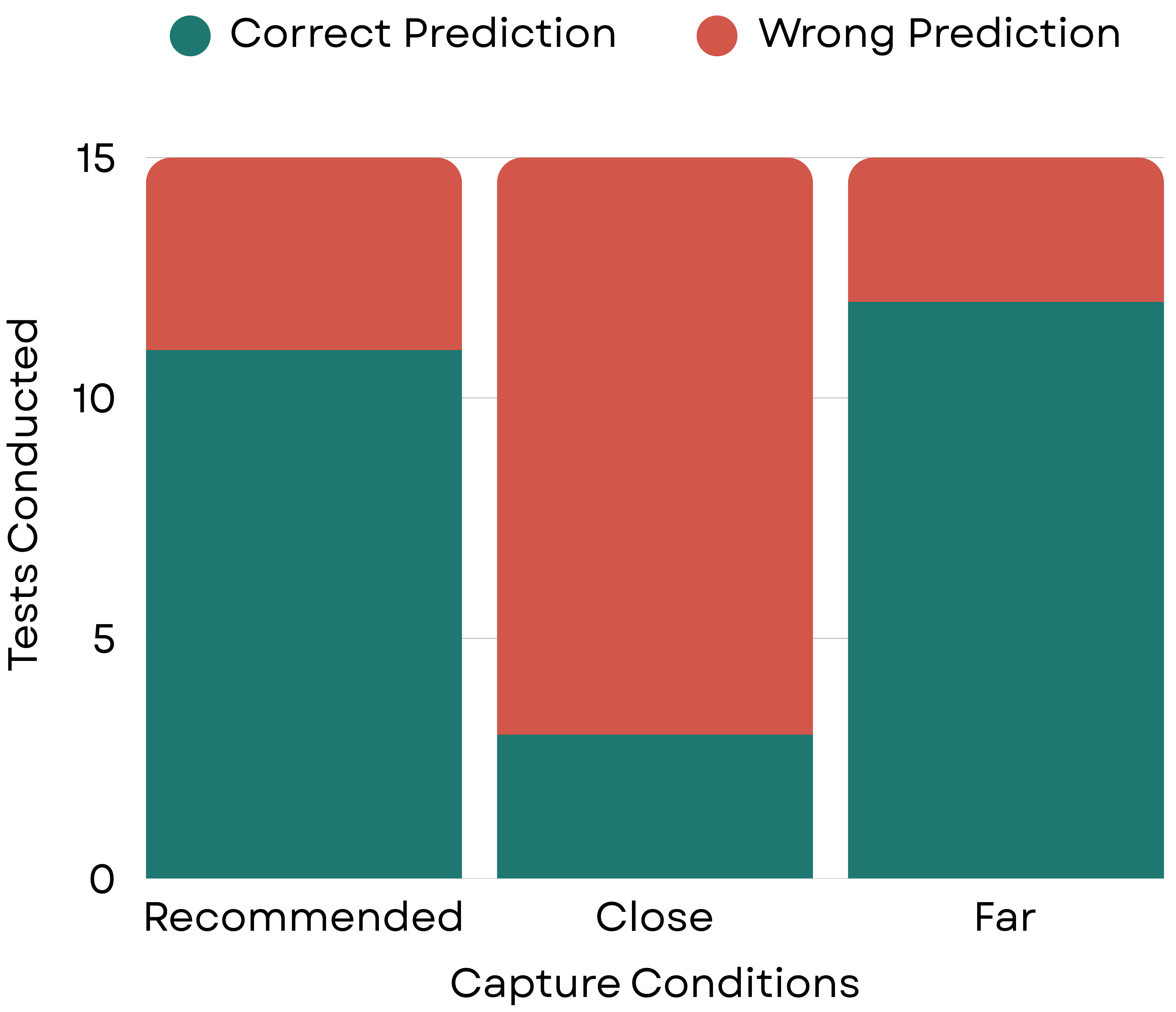
- Tested on **Actual ECG Images.** These images were **NOT included in our dataset.**
- **Supervised and validated by Dr. Gracita Topacio** for clinical accuracy.
- Tested on **15 AF** and **15 Normal Sinus** ECG images.

Limitations Observed

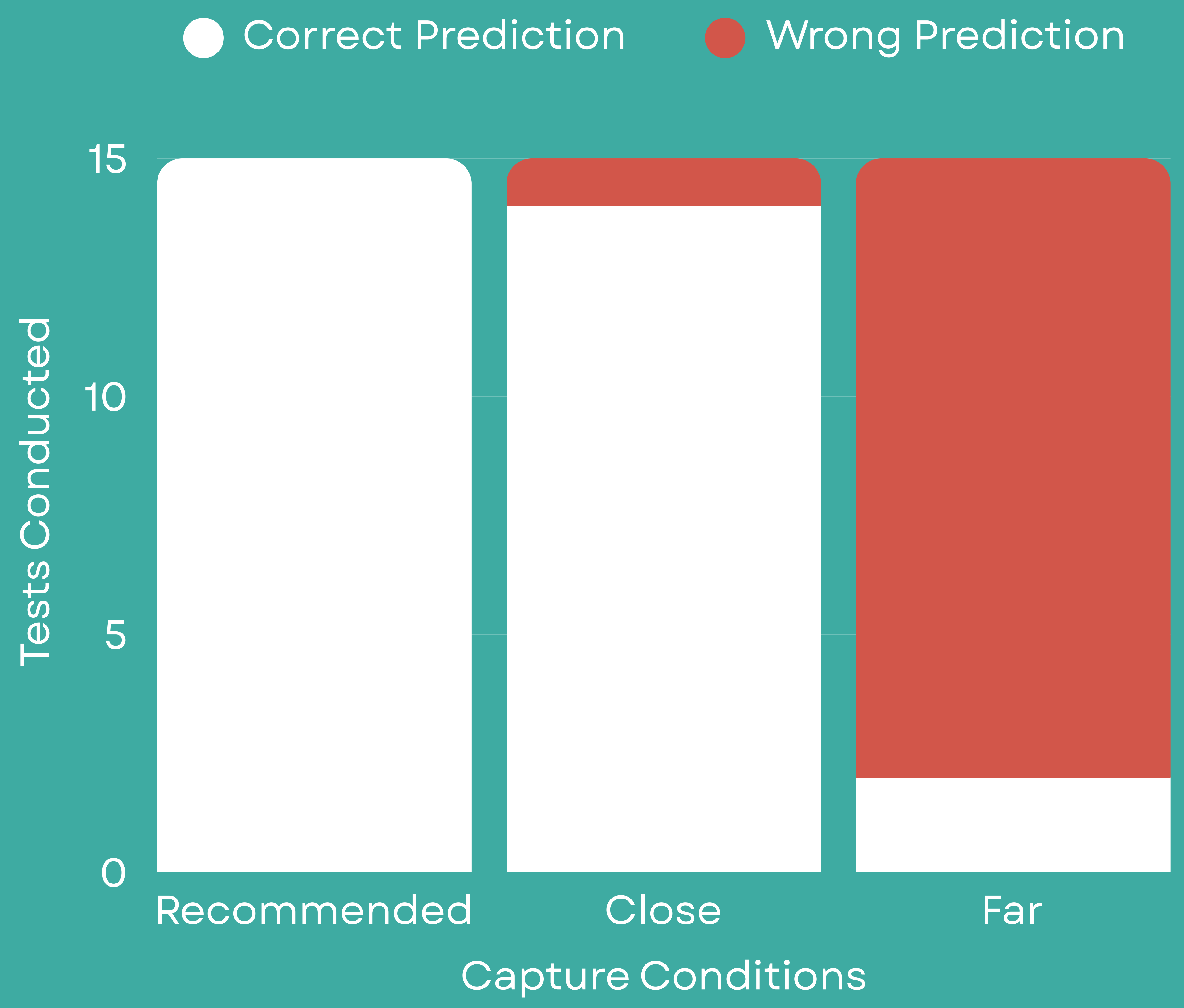
- Accuracy drops under poor lighting, shadows, or distance.
- Confidence scores ranges from 50–60% due to label smoothing.
- Consistent performance requires proper image capture.

	Recommended	Too Close/with Shadow	Too Far
Capture Condition	Proper lighting, full ECG view, moderate distance	Partial ECG, shadow interference	Lower resolution, background noise
Sample	Test # 115 **Atrial Fibrillation** Confidence: 51.10% Is this correct? ✓ 	Test # 114 **Sinus Rhythm** Confidence: 58.02% Is this correct? ✗ 	Test # 113 **Atrial Fibrillation** Confidence: 56.18% Is this correct? ✓ 

AF TEST RESULTS



SINUS TEST RESULTS



DEPLOYED MODEL TESTING

April 14, 2025

April 21, 2025

DEPLOYED MODEL TESTING

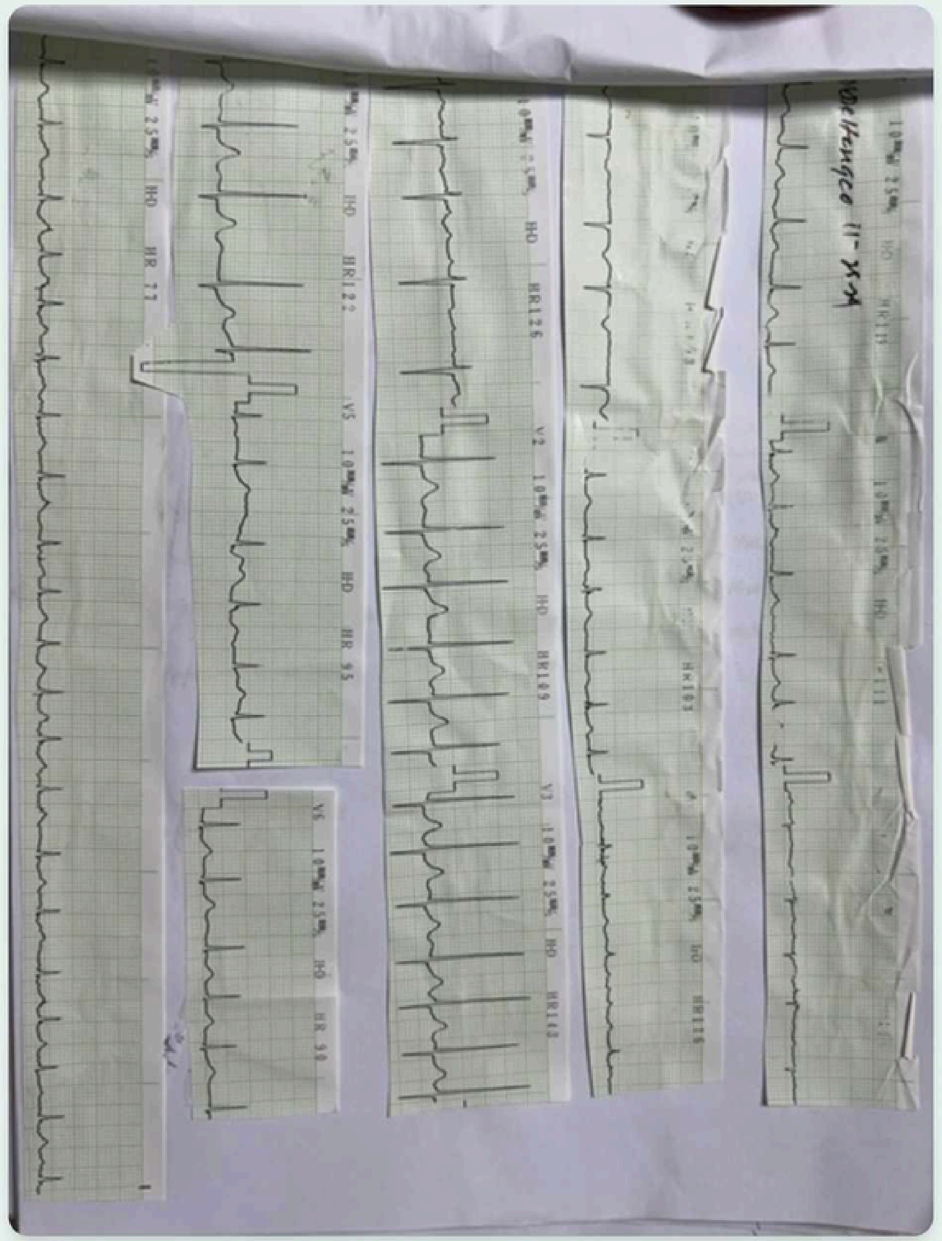
DEPLOYED MODEL TESTING

Test # 155

Atrial Fibrillation

Confidence: 52.20%

Is this correct?

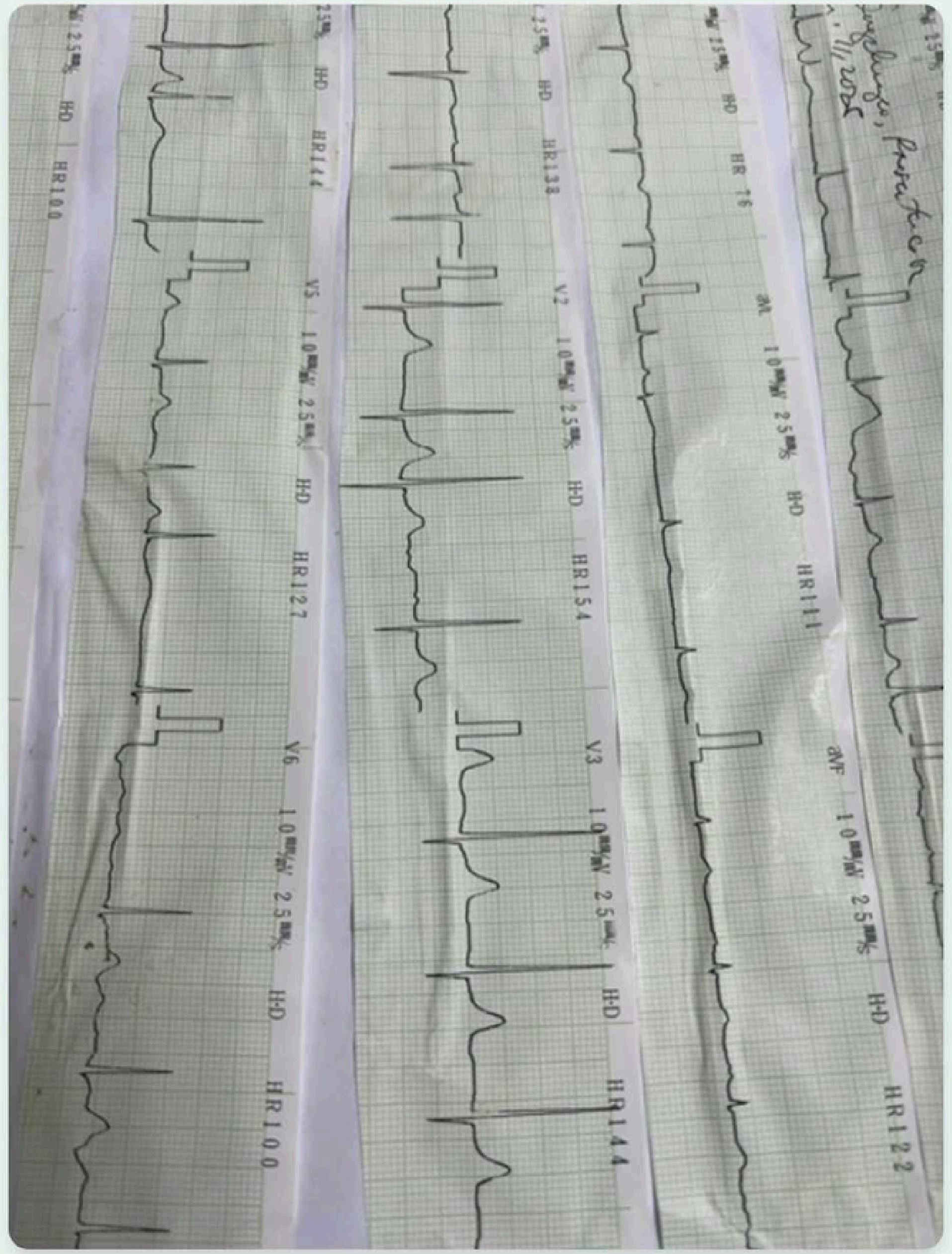


Test # 154

Atrial Fibrillation

Confidence: 50.44%

Is this correct?

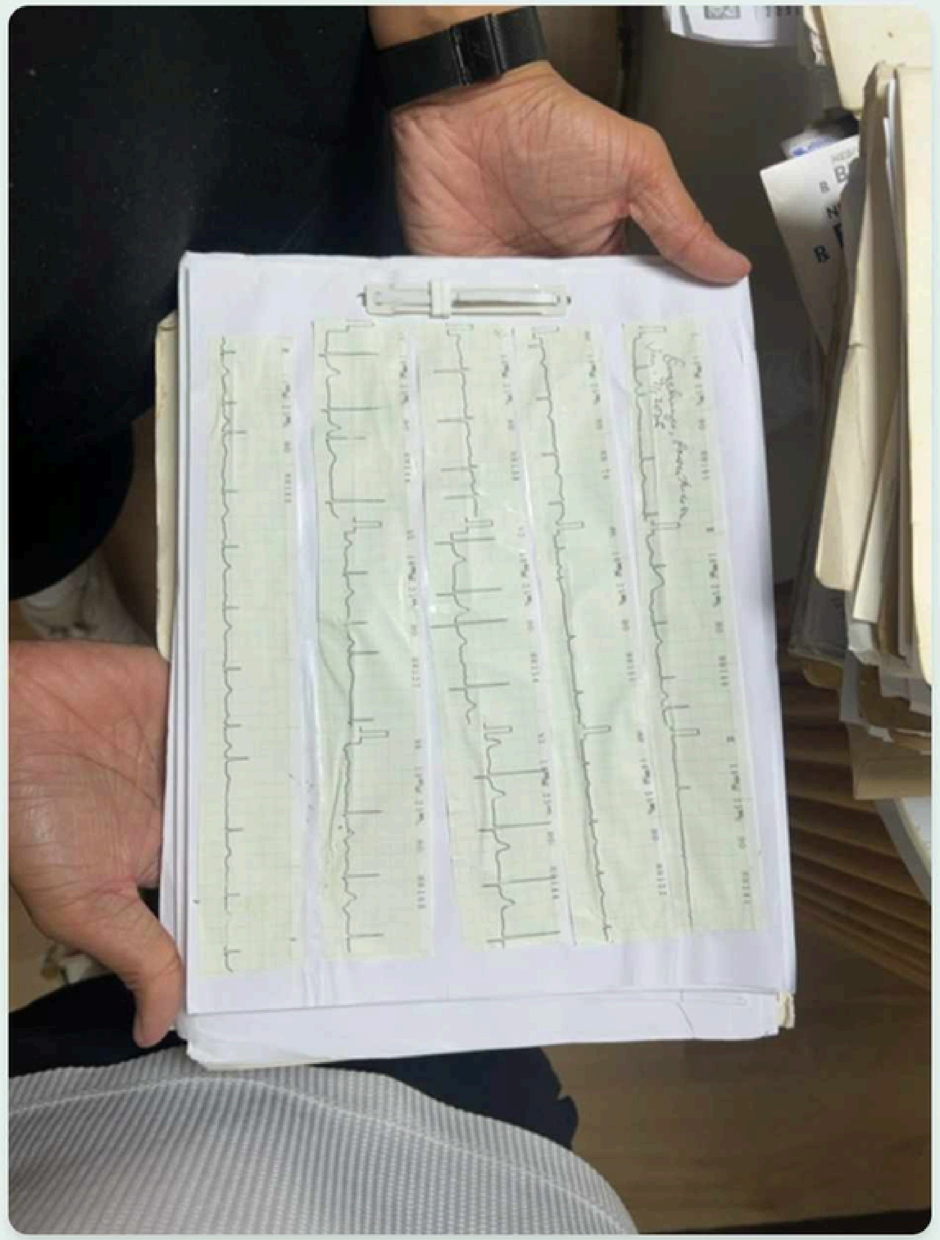


Test # 153

Atrial Fibrillation

Confidence: 64.34%

Is this correct?

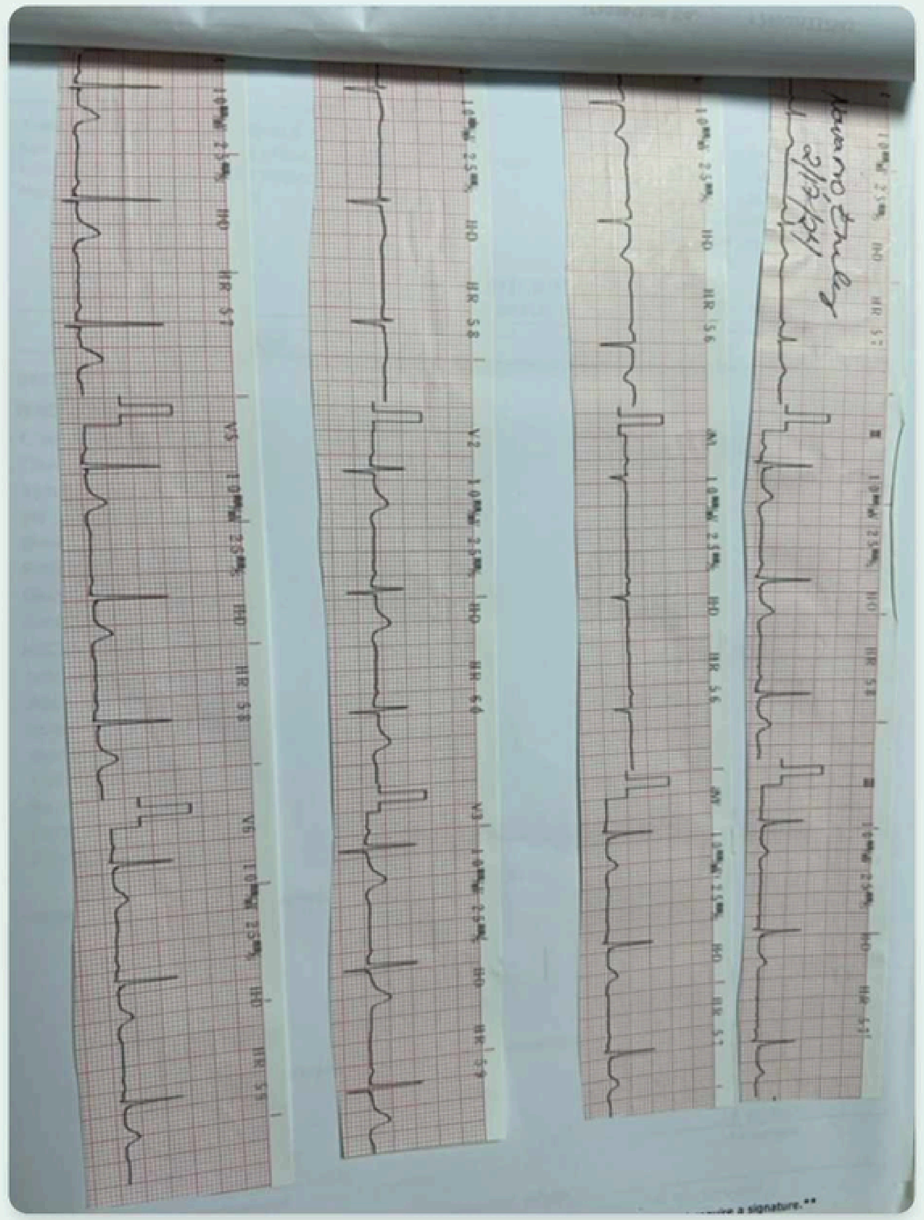


Test # 149

Sinus Rhythm

Confidence: 54.62%

Is this correct?

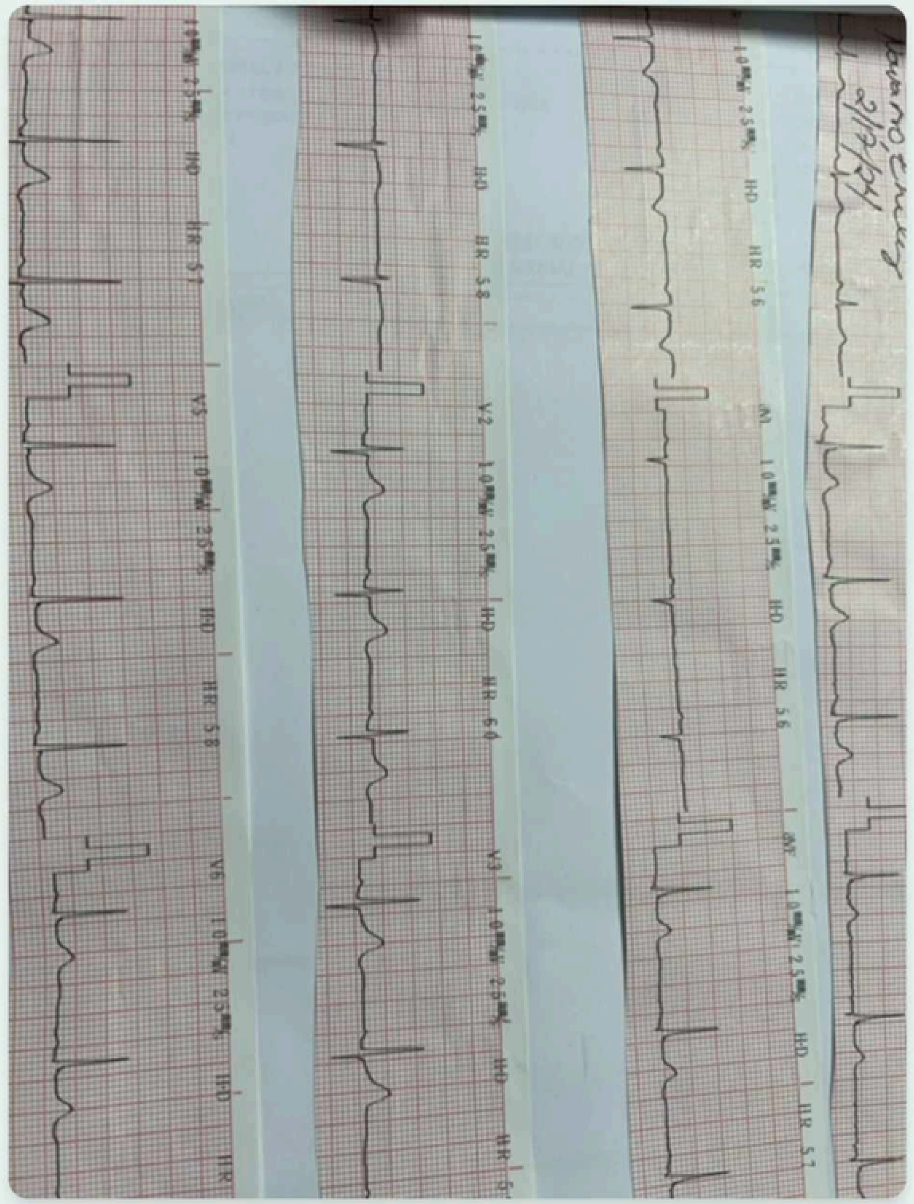


Test # 151

Sinus Rhythm

Confidence: 53.77%

Is this correct?

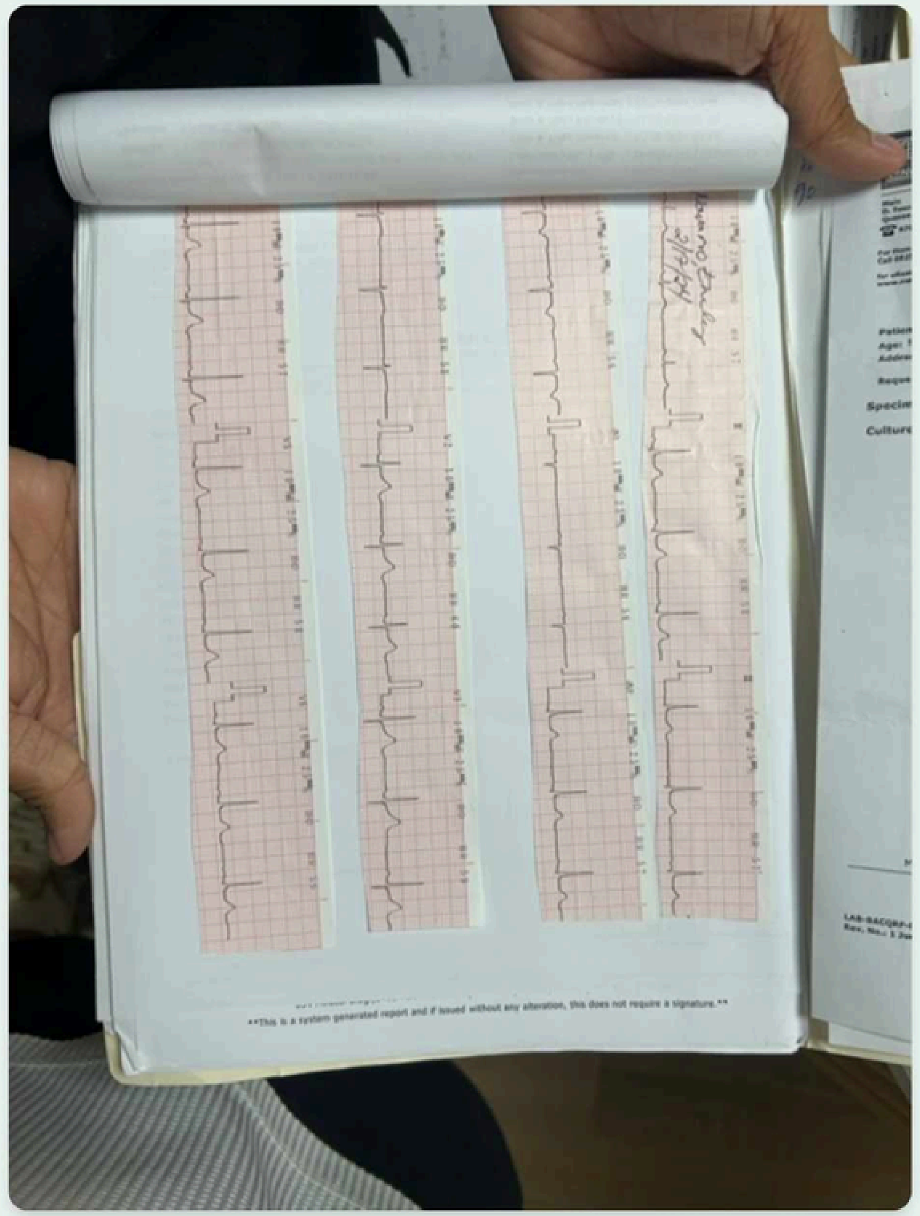


Test # 150

Atrial Fibrillation

Confidence: 58.37%

Is this correct?



CONCLUSION

- A **valid and clinically relevant ECG dataset was collected from Dr. Gracita Topacio's clinic in Manila** and medically verified by the attending cardiologist, ensuring real-world representation for model development.
- A **deep learning model was developed using preprocessed and augmented ECG dataset**. The final fine-tuned model achieved **93% accuracy, balanced F1-scores of 0.93, and an AUC of 0.98** on validation set.
- The **final model was deployed on the AFsense web platform,** secured through Laravel's built-in security features. The website **handles ECG image input and performs model inference for Atrial Fibrillation diagnosis.**
- **Clinical testing, supervised and validated by Dr. Gracita Topacio**, showed that the model correctly classified **all 15 Sinus cases** and **11 of 15 AF cases under ideal conditions**. Performance declined with poor-quality images, revealing practical limitations in uncontrolled environments.

RECOMMENDATION

Train Models Specifically on Lead II

- Lead II provides clearer rhythm representation; focusing on it may improve AF detection and model interpretability.

Integrate Image Quality Checks into the Web App

- Add features like lighting detection, alignment prompts, and clarity validation to ensure better ECG image inputs and reduce errors.

Expand and Diversify the Dataset

- Collaborate with more clinics and utilize public ECG repositories to improve generalizability and reduce data bias.

Use Stronger Backbone Models and Dynamic Input Handling

- Test advanced architectures (e.g., EfficientNetV2, Vision Transformers) and pooling strategies that support high-resolution, variable-size ECGs.

Enable Mobile and Offline Inference

- Optimize the model for mobile devices with offline functionality to support diagnostics in rural or low-connectivity areas.

CPE42

THESIS DEFENDED!

We are now open for your questions

Presented by :

BUENAVENTURA, JUSTINE

ROSTATA, CARL

ANZURES, ANGEL