

ECG ARRHYTHMIA DIAGNOSIS USING AI



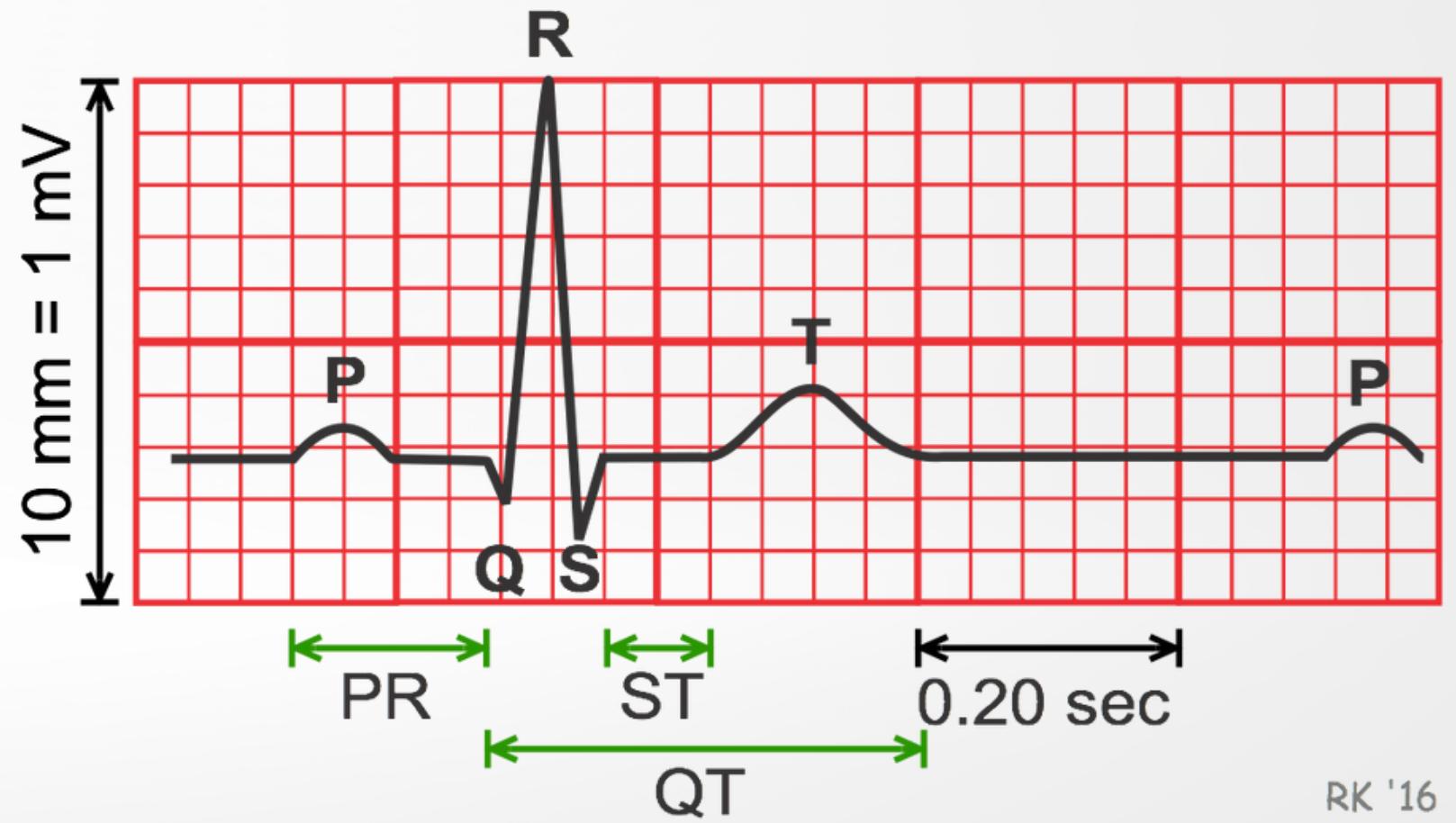
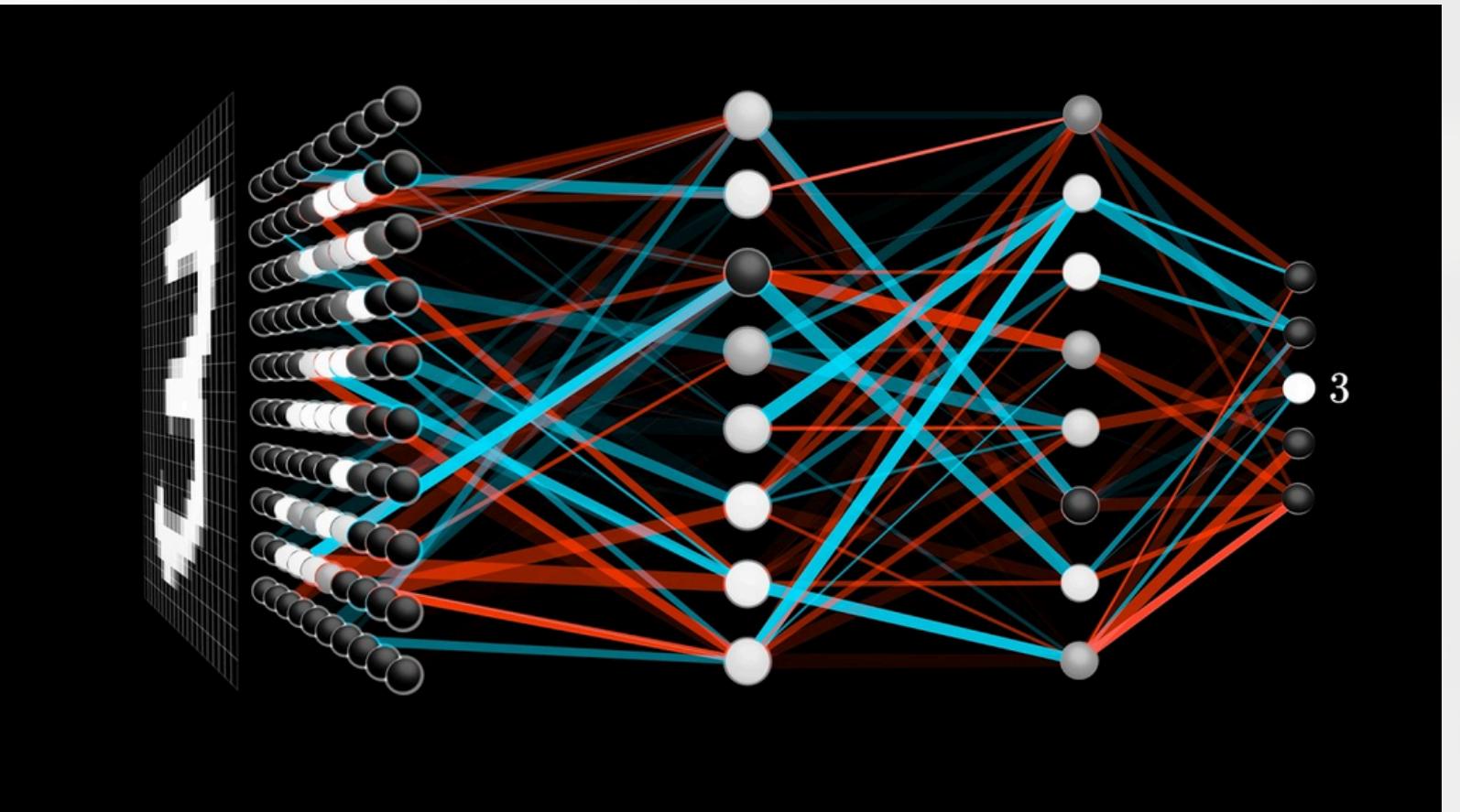
Department of AI
IIT Kharagpur

Ravindra Mina (25AI60R02)
Sarthak Dey (25AI60620)
Biswajit Das (25AI60616)

Guide:
Dr. Mahesh Mohan M.R.
Department of AI

PROBLEM STATEMENT

- This project focuses on developing an automated deep-learning system that classifies ECG heartbeat segments into six arrhythmia categories using the MIT-BIH Arrhythmia Database.
- By building and comparing three different architectures: SimpleCNN, BiLSTM, and ResNet1D-Lite.
- we aim to identify which model best captures the subtle morphological variations in heartbeat patterns and delivers clinically reliable accuracy.



WHY THE PROBLEM IS RELEVANT

Rising use of wearable ECG monitoring devices

Need for real-time anomaly detection

Deep learning's ability to detect subtle waveform patterns

1

2

3

- Cardiovascular diseases are responsible for the highest mortality worldwide.
- timely arrhythmia detection plays a crucial role in prevention and early intervention.
- Manual ECG interpretation can be time-intensive and prone to human error.
- Automated classification systems help support clinicians by providing quick, accurate, and scalable analysis.

EXISTING WORK LITERATURE REVIEW

- **Recent studies highlight that beat-level segmentation and R-peak alignment improve arrhythmia classification accuracy.**
- **Hybrid CNN-RNN models have shown enhanced robustness across varying patient conditions.**
- **Lightweight architectures have gained attention for achieving high performance with lower computational cost.**
- **Researchers have also emphasized the need for model interpretability in clinical ECG analysis.**
- **These findings support our use of efficient deep models and clean, well-segmented ECG beats.**

Previous research has explored a variety of approaches, from classical signal processing such as the Pan-Tompkins algorithm to more sophisticated deep learning models. CNN-based architectures have shown strong performance in detecting morphological patterns, while recurrent models like LSTM and BiLSTM excel at handling temporal information in ECG signals. Residual networks further improve feature extraction by enabling deeper architectures without degrading performance.

Our work builds on:

- Pan - Tompkins Algorithm
- CNNs for morphology
- BiLSTMs for temporal dependencies
- Lightweight ResNet structures for deeper feature learning

DATASET USED

The MIT-BIH Arrhythmia Database was chosen due to its wide adoption and high-quality expert annotations. It contains 48 signal records sampled at 360 Hz, from which individual beats were extracted using R-peak detection. Only six clinically significant labels were retained to simplify the classification problem while maintaining medical relevance. The dataset's structure and volume make it ideal for training robust deep learning models.

- Dataset: MIT-BIH Arrhythmia Database (48 subjects)
- Duration per record: 30-minute ECG recordings
- Sampling frequency: 360 Hz
- Leads used: Lead 0 (single-lead classification)
- Beat-level segmentation:
Heartbeats extracted using R-peak detection → ~100,000 individual ECG beats
- Input format to model:
Each beat is a 0.6-second segment
→ 216 samples
→ Shape: (1×216) (1D ECG signal)
- Classes used (6):
Normal (N), PVC (V), APC (A), LBBB (L), RBBB (R), Paced (/)
- Why this dataset?
 - Clinically annotated
 - High-quality ECG recordings
 - Large number of beat samples
 - Standard benchmark for ECG deep learning

METHODOLOGY: PREPROCESSING

ECG preprocessing involved applying a light band-pass filter to remove noise, followed by accurate R-peak detection. Each beat was segmented using a fixed 0.6-second window centered around the R-peak (72 samples before, 144 after). The resulting signals were then normalized to ensure uniform input behavior.

Purpose of preprocessing:

- Remove noise
- Standardize beat length
- Align morphological landmarks across samples

Raw recording length & sampling
Each record length: 30 minutes.
Compute seconds: $30 \text{ minutes} \times 60 \text{ seconds/minute} = 30 \times 60 = 1800 \text{ seconds}$.
Sampling frequency: $360 \text{ Hz} \rightarrow 360 \text{ samples per second}$.
Samples per record: $1800 \text{ seconds} \times 360 \text{ samples/second} = 1800 \times 360 = 648000 \text{ samples}$.
Heartbeats per recording (approx.)
Typical resting heart rate ≈ 70 beats per minute (example).
Beats in 30 minutes: $70 \text{ beats/minute} \times 30 \text{ minutes} = 70 \times 30 = 2100 \text{ beats per recording}$.
Total beats across 48 recordings
 $2100 \text{ beats/record} \times 48 \text{ records} = 2100 \times 48 = 100800 \text{ beats}$ (approximate total).
So even though there are only 48 recordings (patients), each recording contains $\sim 2,000$ beats, yielding $\sim 100k$ beats overall.

METHODOLOGY: PREPROCESSING

R-Peak Detection (Pan-Tompkins Algorithm)

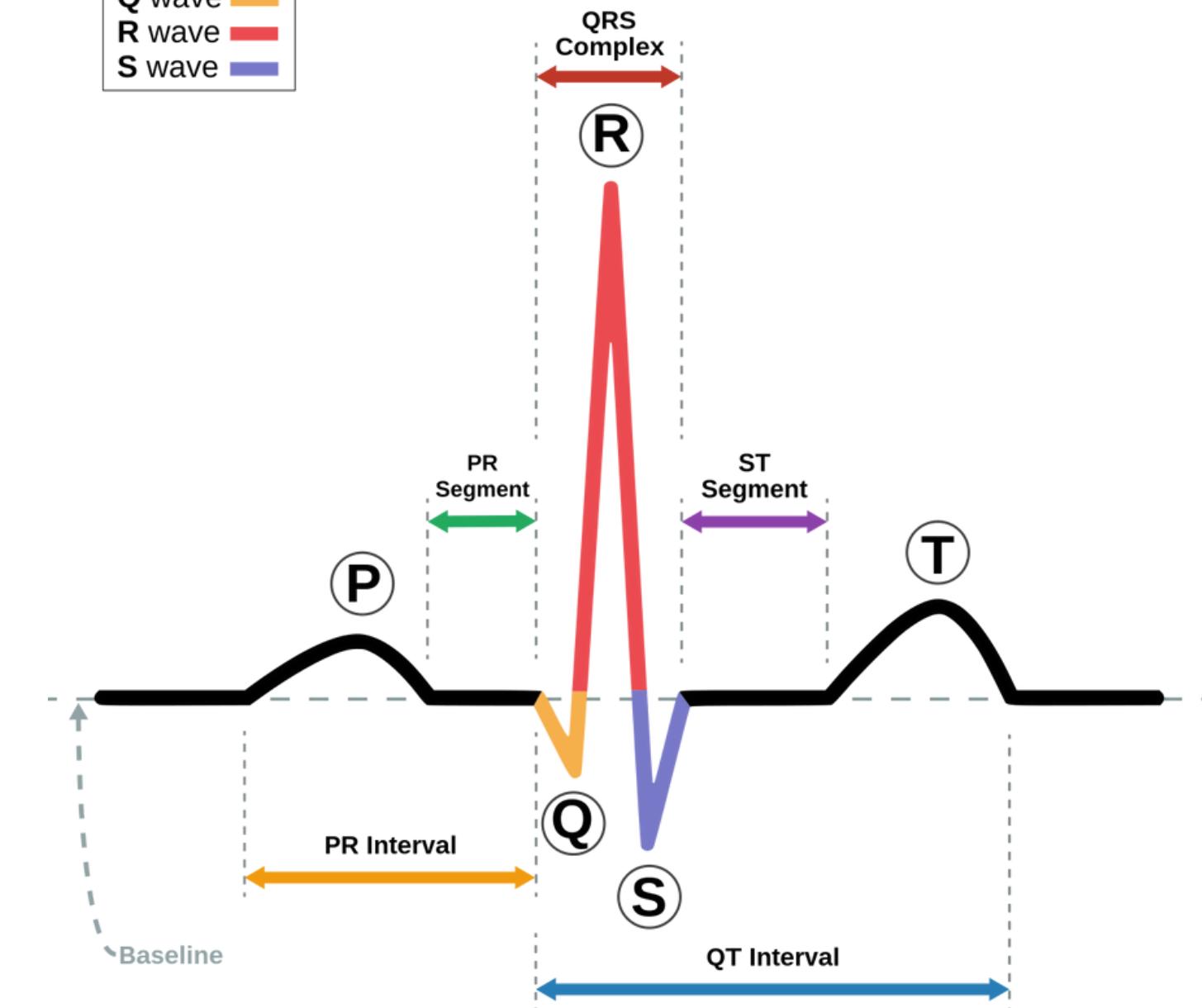
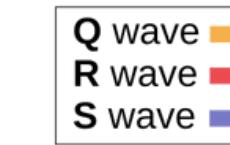
- Bandpass filter (5–15 Hz) isolates QRS energy
- Derivative highlights steep slopes
- Squaring amplifies QRS peaks
- Moving Window Integration (150 ms) forms a QRS energy envelope
- Adaptive threshold + 200 ms refractory detects true R-peaks
- Outputs ~2000 R-peaks per patient → ~100k beats total

Annotation Alignment (MIT-BIH Labels)

- MIT-BIH provides beat annotation sample locations
- Each annotation is matched to the nearest detected R-peak
- Ensures accurate beat labeling even with small detection shifts

Beat Segmentation (0.6-Second Window)

- Extract a fixed segment around each aligned R-peak:
 - 0.2 s before R (72 samples)
 - 0.4 s after R (144 samples)
- Total segment length: 216 samples (0.6 s)
- Final input to the model: 1×216 ECG beat



METHODOLOGY: PREPROCESSING

Filtering & cleaning

- Discard segments that:
 - Cross recording edges (incomplete windows),
 - Have wrong length after processing,
 - Belong to unwanted annotation symbols.
- Keep only the chosen classes (N, V, A, L, R, /), mapping symbols → integer IDs.

Final dataset

- After extraction and filtering you get roughly the same order as the beat count estimate (e.g., ~90k–110k segments depending on filtering and discarded edge cases).
- Example: if you lose 5–10% during filtering:
 $100800 \times 0.90 = 90720$ samples remain – still ample for training.

What the model sees

- Input: torch tensor of shape (batch_size, 1, 216)
- Label: integer in {0,...,5} corresponding to the six retained classes.

Train/val split

- Split by sample (not by full record) as implemented: e.g., VAL_SPLIT = 0.2 → 80% train / 20% val on the beat-level samples.

METHODOLOGY: TRAINING SETUP

Each model was trained by the following criteria

- 30 epochs
- Adam optimizer
- learning rate of 0.001 and
- weight decay for regularization.
- A batch size of 256 ensured training efficiency
- an 80-20 train-validation split was maintained across all experiments.
- The best-performing models were saved based on the highest macro-F1 score.

Metrics tracked:

- Train/Validation loss
- Validation accuracy
- Macro-F1 score for class balance evaluation

Class Distribution

Class Distribution

Normal (N) class dominates (~75,000 samples)

Minority classes (/ , A, L, R, V) have only 2,000–8,000 samples. Dataset is highly imbalanced

Why This Matters

Model may learn to predict Normal more often

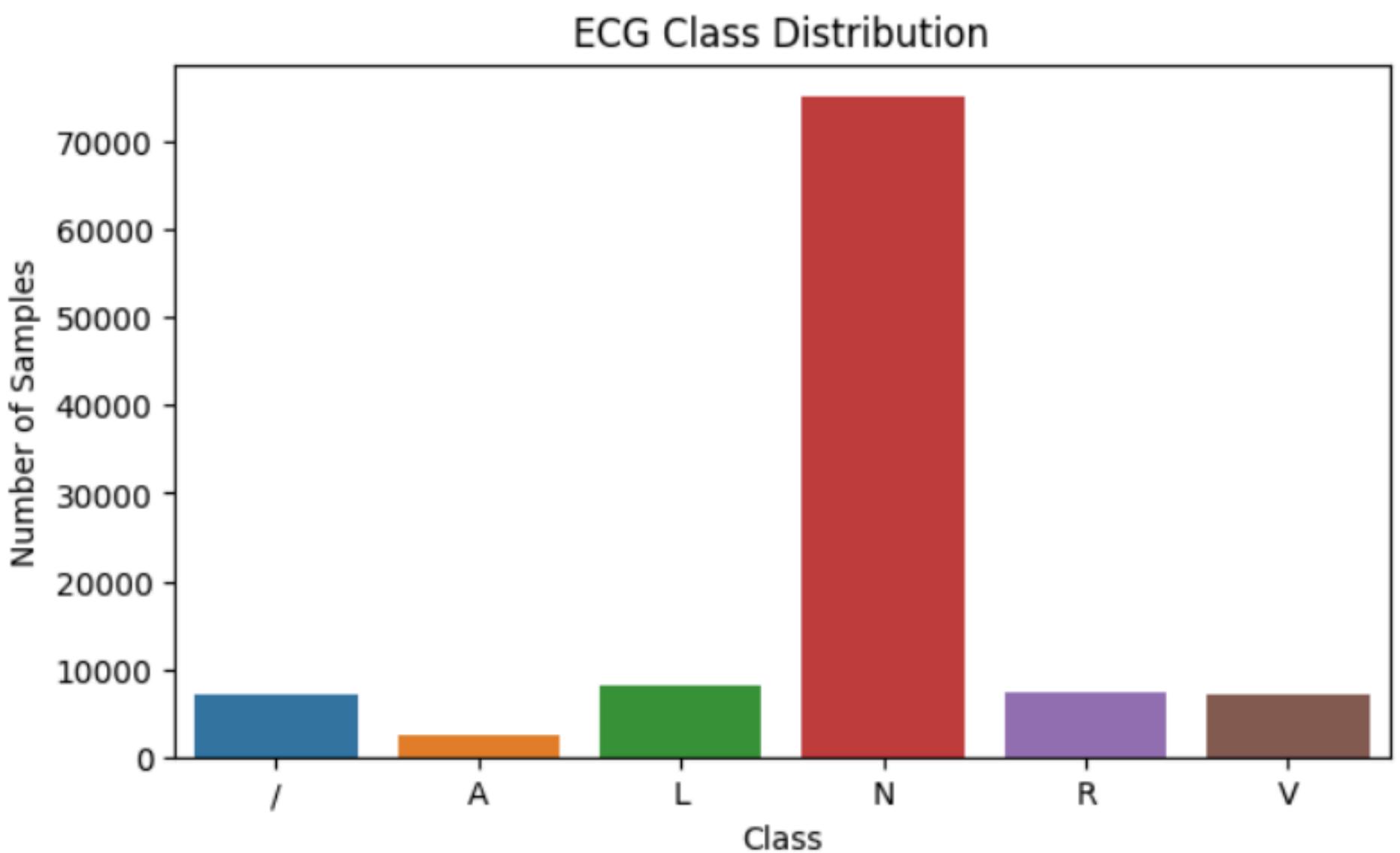
Accuracy becomes misleading (majority-biased)

Minority classes contribute less to gradient updates

Risk of lower recall for rare arrhythmias

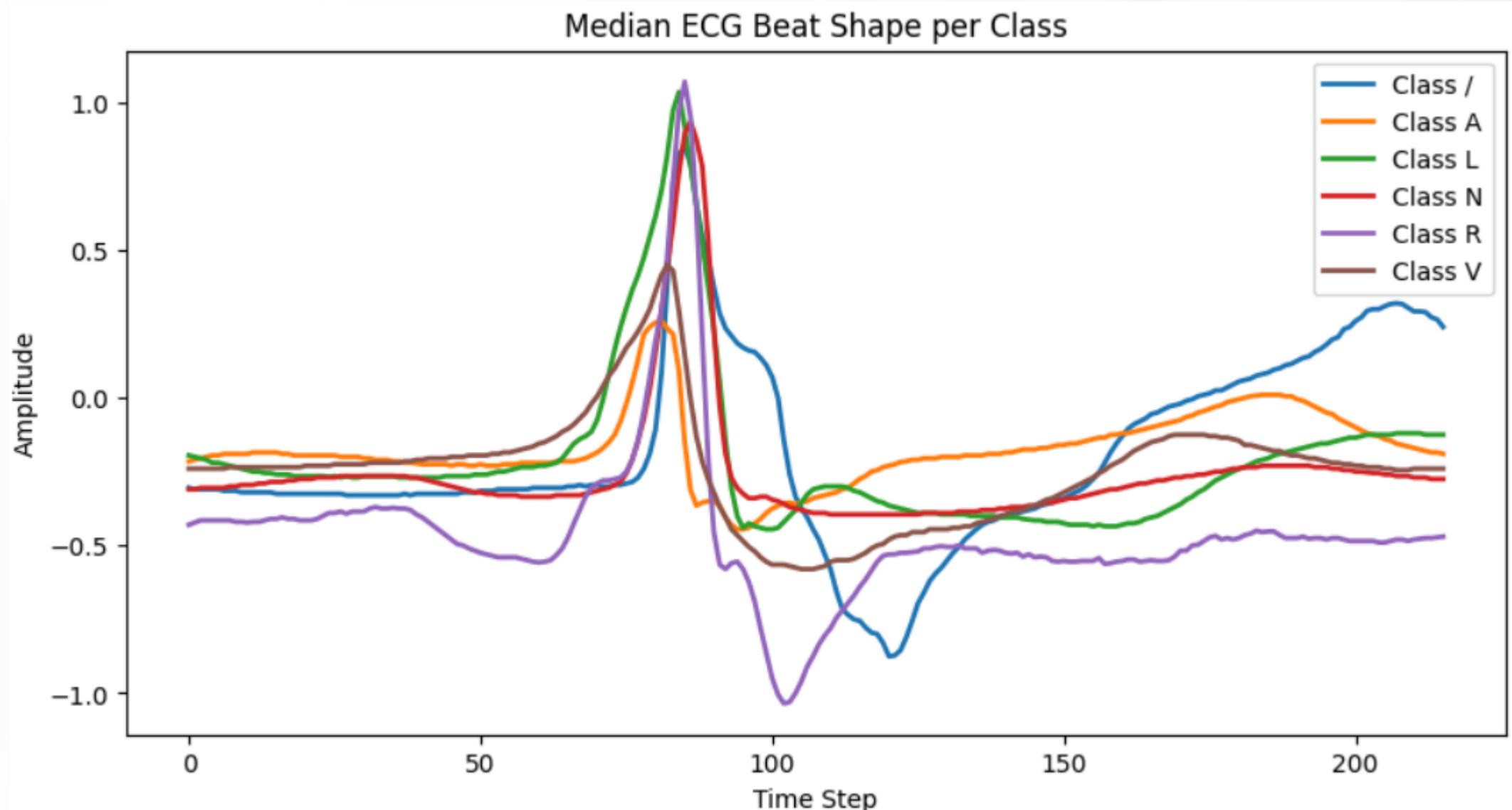
Why Our Model Still Performs Well

- ECG beat morphology is distinctive → easy to learn patterns
- Clean R-peak-centered segments improve feature clarity
- CNN/ResNet/BiLSTM models capture both shape & temporal cues
- High Macro-F1 scores (>0.94) indicate strong performance across all classes despite imbalance



ADDITIONAL VISUAL INSIGHTS

- The median beat shapes show clear morphological differences between classes, especially in the QRS complex.
- Normal (N) beats have a sharp, narrow R-peak, while arrhythmias like LBBB, RBBB, and PVC show widened or distorted waveforms.
- These consistent shape variations help the deep learning models distinguish between arrhythmia types effectively.



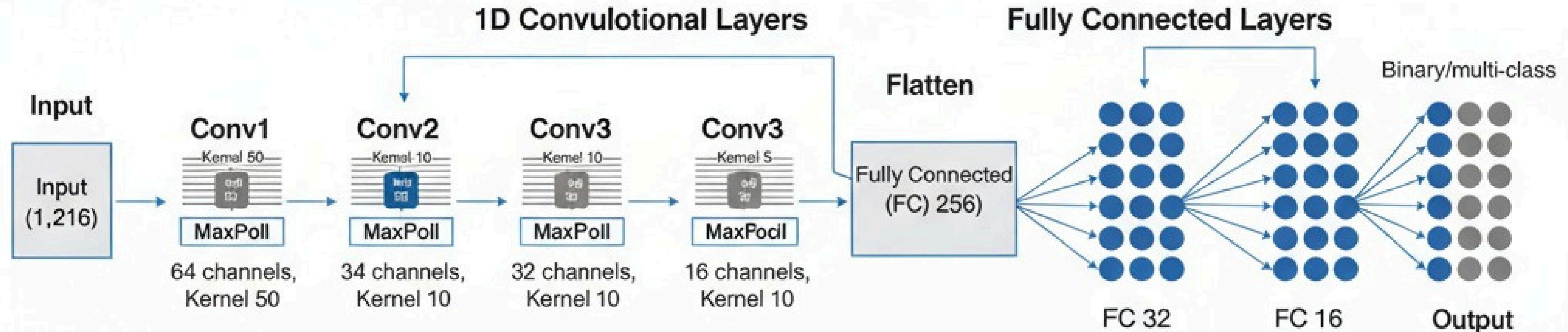
METHODOLOGY: MODEL ARCHITECTURES

Three deep learning architectures were implemented and compared. The SimpleCNN model extracted spatial features directly from signal morphology. The BiLSTM model captured temporal relationships inherent to ECG signals. The ResNet1D-Lite model leveraged residual connections, enabling efficient deeper feature extraction without vanishing gradients.

Why these models?

- CNN → local waveform patterns
- BiLSTM → temporal rhythm
- ResNet → deeper representation learning

RESULTS: SIMPLE CNN

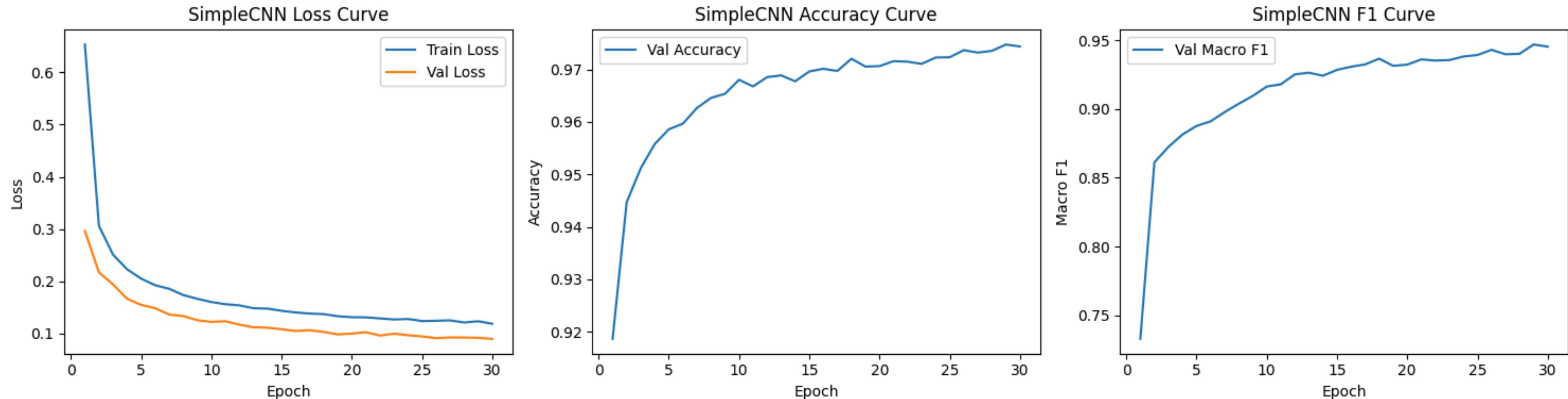


Deep learning first gained traction in ECG analysis through 1D CNN architectures. These models learn morphological features directly from raw beats, outperforming traditional handcrafted approaches. Studies like Acharya et al. and Kiranyaz et al. showed that CNNs effectively capture QRS shape, ST changes, and beat morphology.

Key Points:

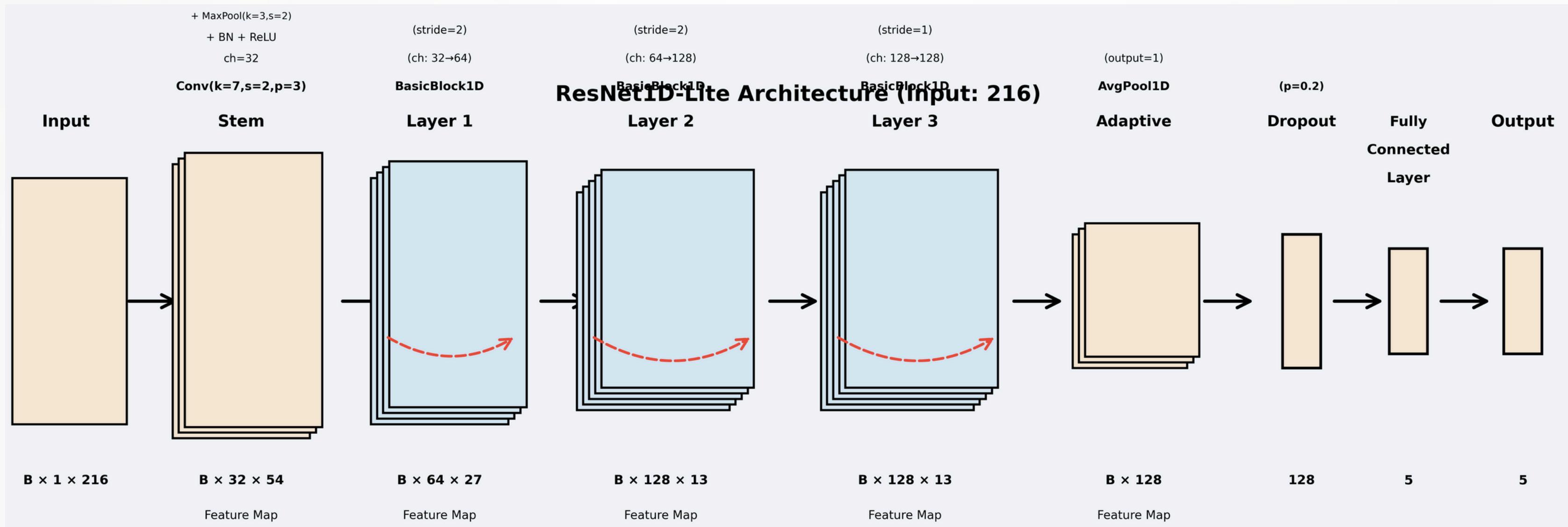
- CNNs excel at learning spatial/morphological ECG features
- Strong performance on MIT-BIH without handcrafted features
- Basis for our SimpleCNN, which follows a similar multi-layer convolution + pooling philosophy
- This gives the Accuracy of 95% and F1-Score of 0.94.

SimpleCNN Training Process



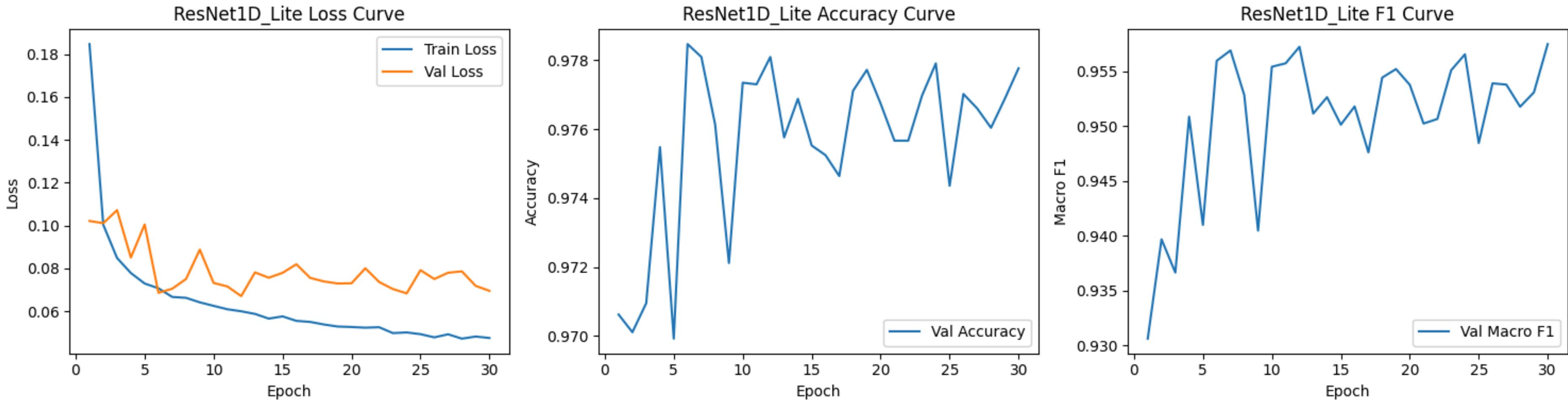
- Both training and validation loss consistently decrease, indicating that the model is learning effectively without major instability.
- Validation loss stays close to training loss, suggesting good generalization and no significant overfitting across training epochs.

RESNET1D-LITE



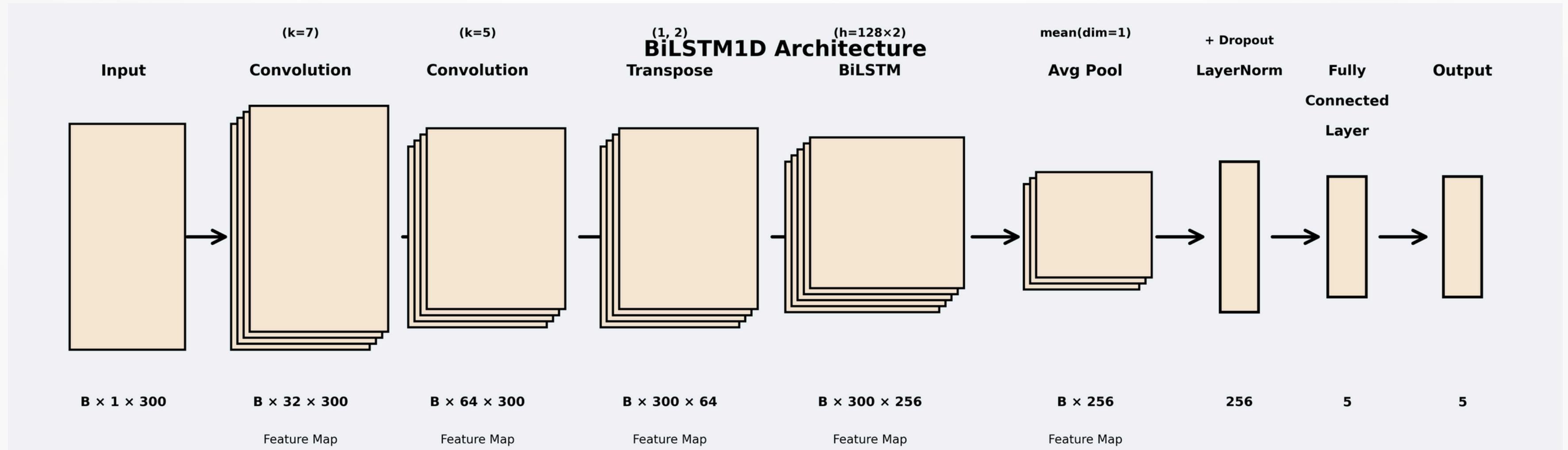
- The ResNet1D-Lite model achieved the strongest overall performance, with a validation accuracy of 97.85% and macro-F1 of 0.9559.
- Its confusion matrix demonstrated clear class boundaries, and the ROC curves showed near-perfect AUC for several classes (some reaching 1.0).
- These results indicate that residual learning enhances the network's ability to capture detailed ECG signatures.

RESULTS: RESNET1D-LITE



- Training loss steadily decreases and stabilizes, indicating that the model is learning robust features over time.
- Validation loss shows small fluctuations, which is typical for deep models but suggests some sensitivity to validation samples.
- No major gap between training and validation loss, meaning no strong overfitting, and overall the model generalizes reasonably well.

BiLSTM

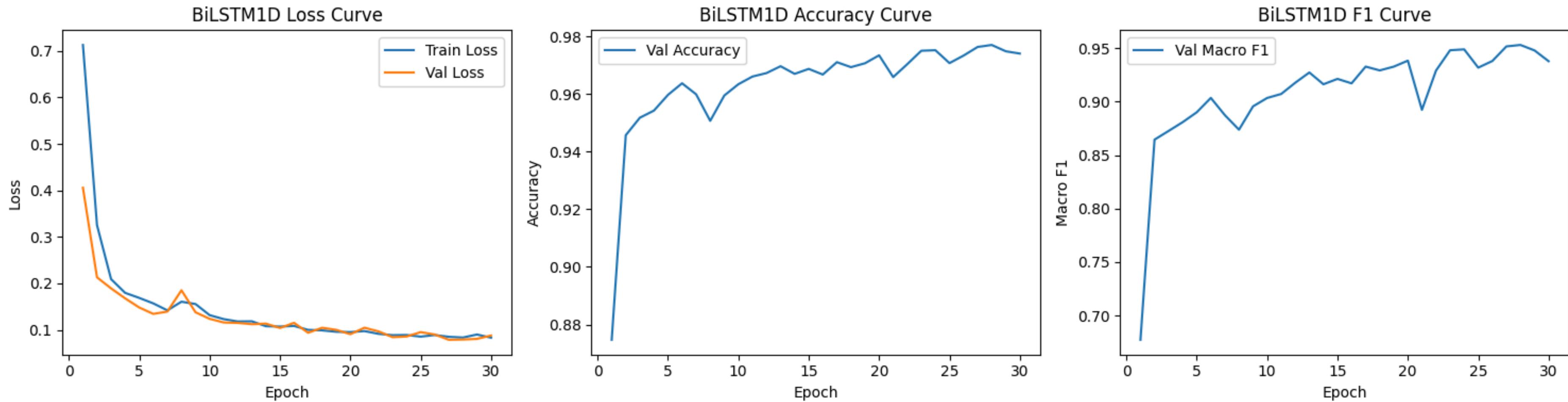


To capture temporal dependencies within heartbeat dynamics, researchers introduced LSTM and BiLSTM architectures. Works such as Faust et al. and Xia et al. demonstrated that bidirectional recurrence helps model subtle timing-based arrhythmias better than pure CNNs.

Key Points:

- LSTMs capture rhythm, timing, and sequential dependencies
- BiLSTM improves by integrating both forward and backward temporal context
- Motivation for our BiLSTM1D, combining conv front-end + bidirectional recurrence
- This Gives accuracy of 98% and Macro F1 score of 0.96

RESULT BiLSTM



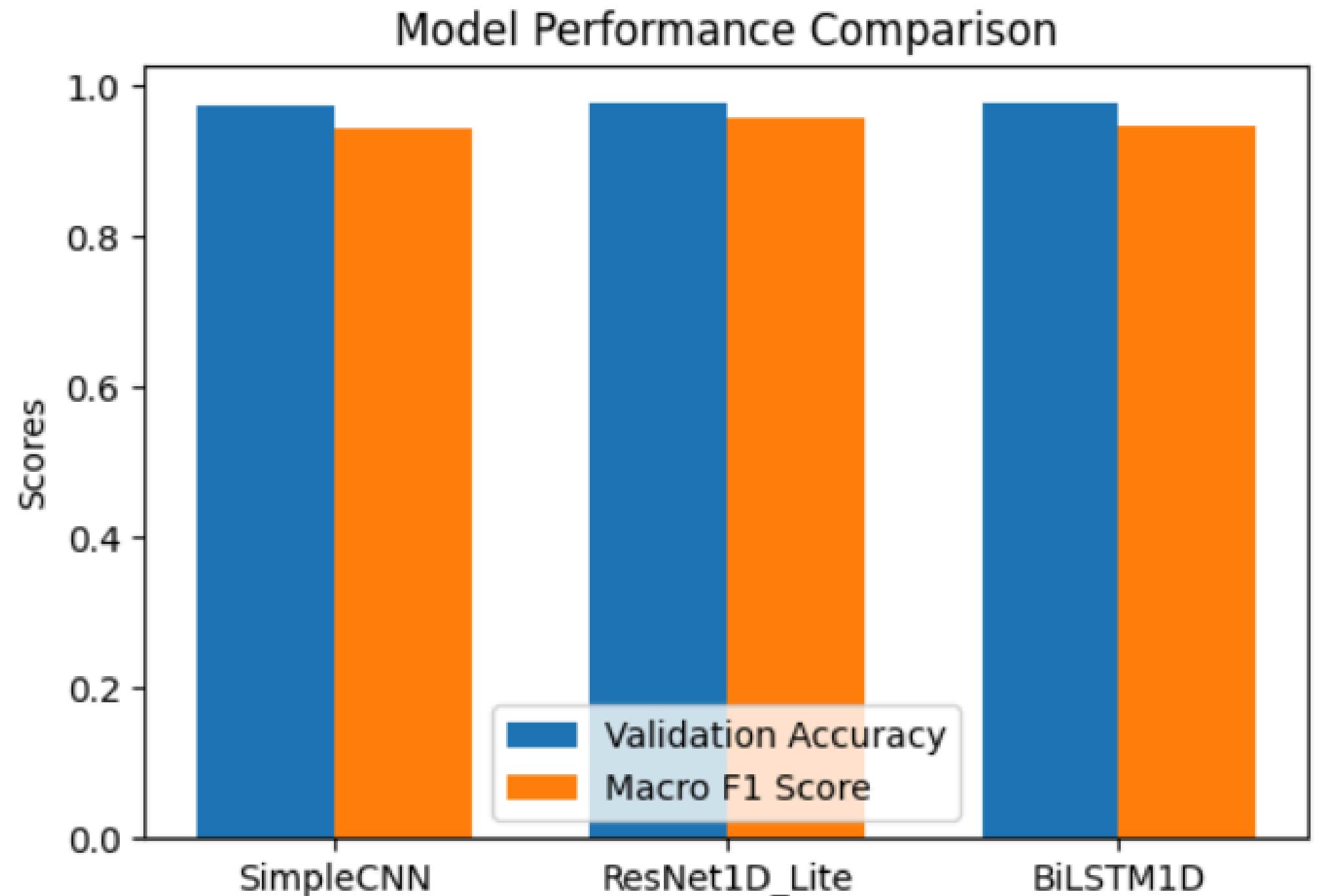
- Training and validation loss follow almost identical trajectories, showing excellent generalization and no signs of overfitting.
- Rapid loss reduction in the first few epochs, indicating that the BiLSTM quickly learns temporal ECG patterns.
- Stable and low loss after ~10 epochs, suggesting that the model converges effectively and maintains consistent performance.

MODEL COMPARISON SUMMARY

To compare the models directly, we evaluated their validation accuracy and macro-F1 scores:

Model	Validation Accuracy	Macro-F1 Score
SimpleCNN	97.38%	0.9449
ResNet1D_Lite	97.82% (Highest)	0.9579 (Best)
BiLSTM1D	97.60%	0.9487

The differences reveal that while all models are highly performant, ResNet1D-Lite provides a slight edge in both robustness and class-wise balance, making it the most reliable architecture in this experiment.



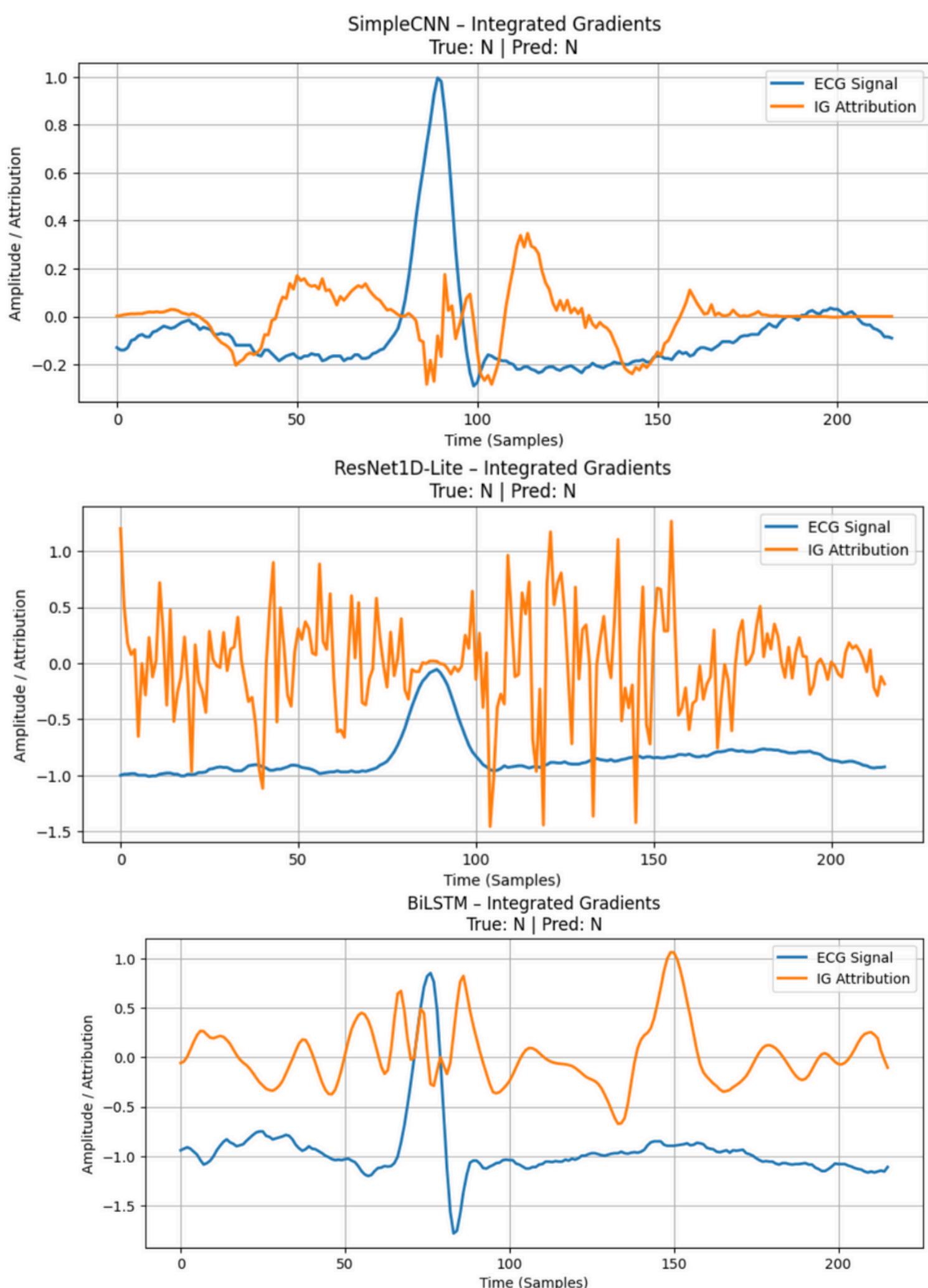
Explainability Using Integrated Gradients

What the Plots Show

- Each plot displays the original ECG beat and its Integrated Gradients (IG) attribution map.
- IG highlights the most influential time-points that led to the model's prediction.
- Higher attribution values indicate regions that strongly impacted the classification decision.

How Different Models Interpret the ECG

- SimpleCNN:
 - Focuses mainly on the QRS complex, capturing sharp and prominent morphological changes.
- ResNet1D-Lite:
 - Learns more detailed features and often highlights subtle waveform variations across QRS, ST, and T-wave regions.
- BiLSTM:
 - Attributes importance across a broader temporal range, reflecting its ability to use sequential dependencies before and after the R-peak.



FUTURE WORK

- **Integrate Attention Mechanisms**

Incorporate self-attention or Transformer-based layers to better capture long-range temporal dependencies in ECG signals.

- **Deploy as a Real-Time Monitoring System**

Optimize and convert the trained model into a lightweight mobile/edge application for continuous ECG monitoring.

- **Extend to Multi-Class & Multi-Lead ECG Analysis**

Evaluate model performance on more arrhythmia types and use 12-lead ECG data for improved diagnostic capability.



THANK YOU