Healthcare Technology Innovation Centre

Collaborate · Innovate · Impact

— A joint initiative of —

Indian Institute of Technology Madras

Dept. of Biotechnology
Ministry of Science & Technology

Interpreting Deep Neural Networks for Single-Lead ECG Arrhythmia Classification

Sricharan Vijayarangan, Balamurali Murugesan, Vignesh R, Preejith SP

Jayaraj Joseph and Mohansankar Sivaprakasam

Motivation

- **Deep Learning methods for arrhythmia detection**
  - Scales automated systems
  - Removes requirement for expert rules
  - Augments doctor’s ability
- **Limitations**
  - Black box
  - Unreliable
- **Requirement**
  - Correlation b/w model outputs and ECG input samples
  - Comparing visualizations with medical literature

Contribution

- Novel adaptation of CNN saliency visualization to 1D ECG signals
- Extension of the LSTM visualization procedure for ECG signals
- Rigorous analysis of the saliency maps
- Draw comparisons to traditional diagnosis as highlighted in medical literature

Problem formulation

Dataset → $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), ..., (x^{(m)}, y^{(m)})\}$

Input ECG Signal → $x^{(i)}$

Labels → $y^{(i)} \in \{0, 1, ...7\}$

CNN → $z_1^{(i)} = F_1(x^{(i)}; \theta_1)$

LSTM --> $z_2^{(i)} = F_2(x^{(i)}; \theta_2)$

FC → $z_3^{(i)} = F_3(z_1^{(i)} || z_2^{(i)}; \theta_3)$

Softmax → $p(z_3^{(i)})$

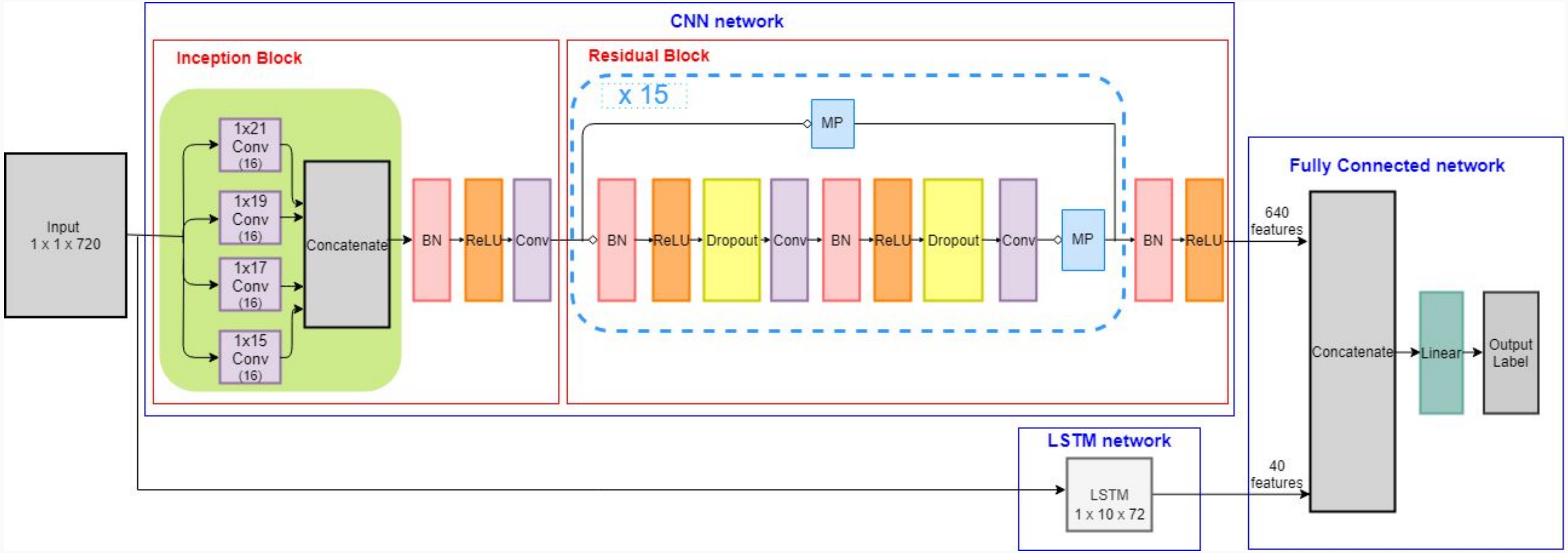
Rationale behind Architecture Choice

Three popular DL architectures in literature were compared for this specific 8 class classification problem.

Model	Precision	Recall	F1-Score	Accuracy
Hannun *et al.* [2]	0.93	0.93	0.93	0.93
Zihlmann *et al.* [3]	0.94	0.94	0.94	0.94
Murugesan *et al.* [4]	**0.98**	**0.97**	**0.97**	**0.97**

Murugesan *et al.*'s model is clearly the best for this classification task.
Thus, it was chosen for the interpretability task.

CONT.



Architecture Design

Image Credits: https://ieeexplore.ieee.org/abstract/document/8438739

Dataset Description

Rhythm Types	MITDB	LTAFDB	LTDB	Total
N	75013	10756	517402	603171
PVC	7121	1318	5137	13576
PAC	2542	14914	-	17456
AFIB	102	7241	-	7343
SVTA	22	3265	-	3287
SBR	-	11323	-	11323
LBBB	6580	-	-	6580
RBBB	5400	-	-	5400

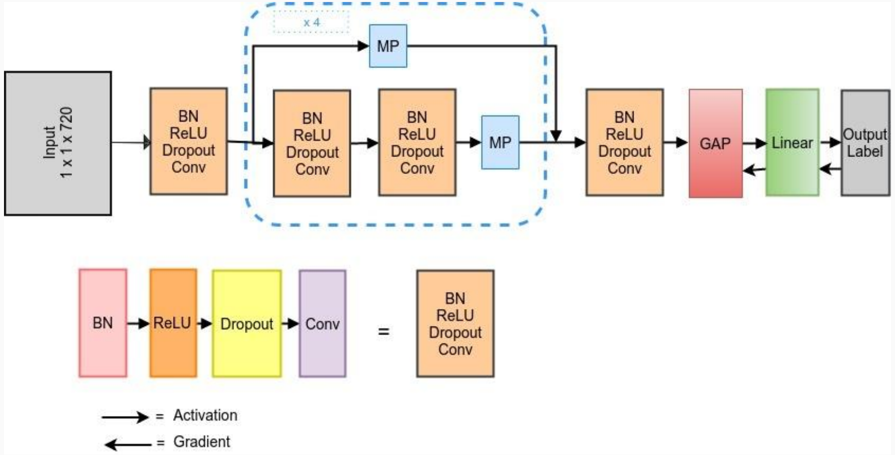
Visualization of CNN

Activation of unit k → $f_k(x)$

Weight for class C corresponding to Unit ‘k’ → w_k^c

I/p to Softmax → $\sum_k w_k^c \sum_x f_k(x)$

CAM → $M_c(x) = \sum_k w_k^c f_k(x)$



- CAM of vector length 48 is obtained.
- Upsampled to 720

LSTM Visualization

LSTM visualization network (ψ)

Input to the LSTM network (ϕ)

Input ECG signal $(\mathbf{x}_{1:T})$

Mask $(\mathbf{m}_{1:T})$

Weights of the saliency term λ_1

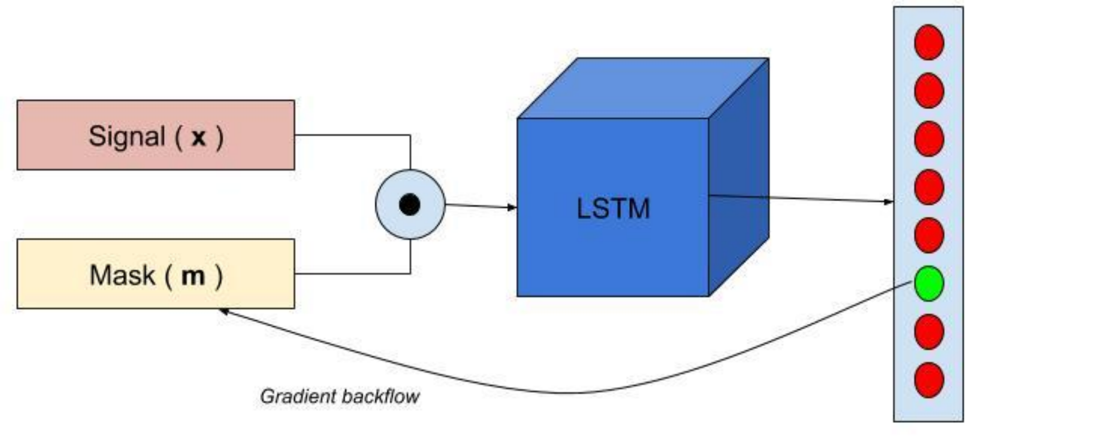
Weights of the smoothing term λ_2

$$J = \underset{m_{1:T}}{argmin}\ \lambda_1 ||\mathbf{1} - \mathbf{m}_{1:T}||_1 + \lambda_2 \sum_{t=1}^{T-1} |\mathbf{m}_{t+1} - \mathbf{m}_t| + s_c(\psi(\phi(\mathbf{x}_{1:T}; \mathbf{m}_{1:T})))$$

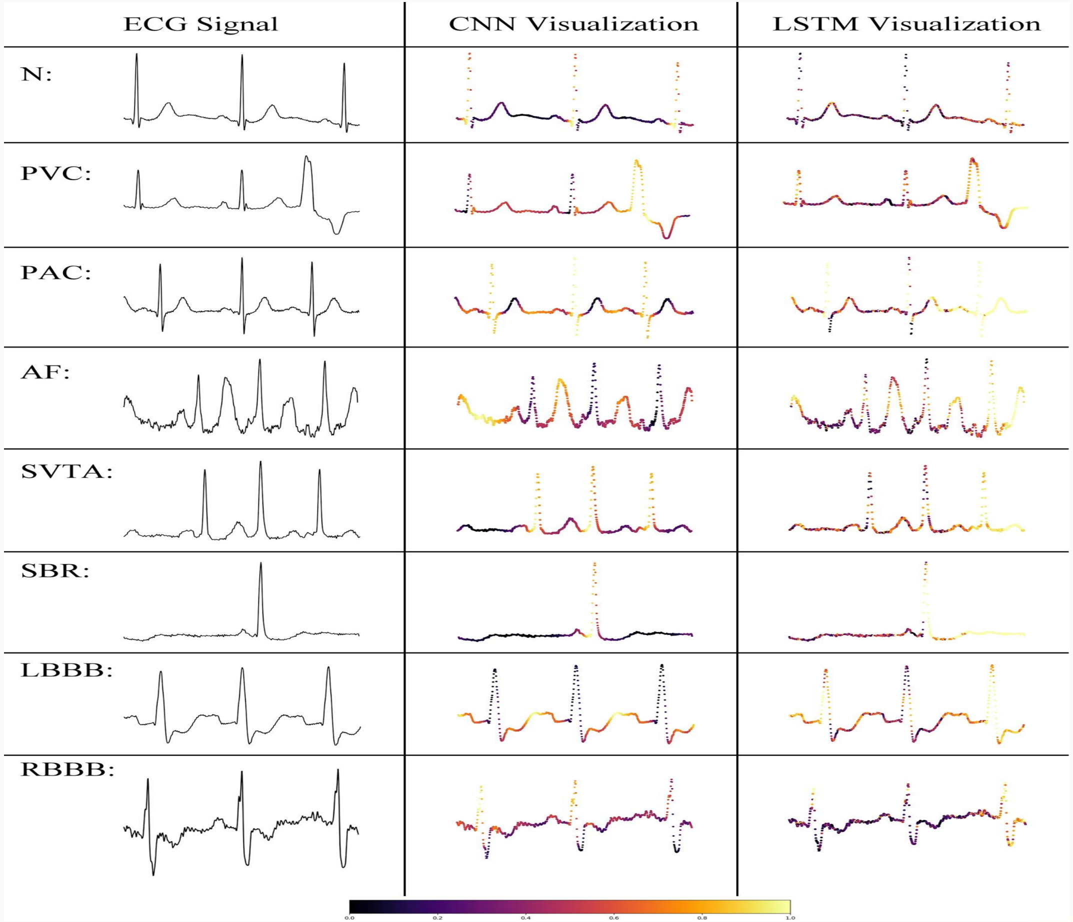
$$\phi(\mathbf{x}_{1:T}; \mathbf{m}_{1:T}) = \mathbf{m}_{1:T} \odot \mathbf{x}_{1:T} + k(\mathbf{1} - \mathbf{m}_{1:T})$$

CONT

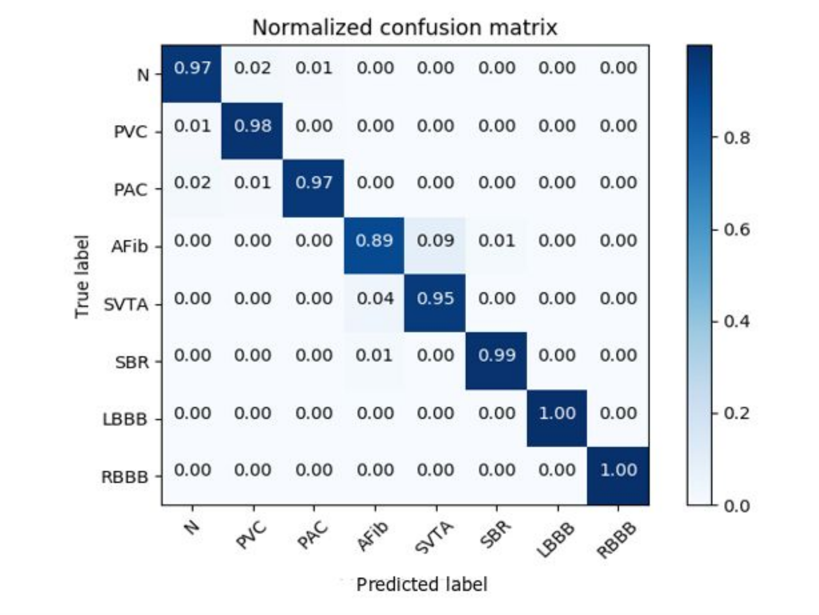
- Initially, $m_{1:T} = 0$
- λ_1, λ_2 and learning rate are set to 1, 0.001 and 0.001 respectively.
- Gradient update is done for 500 iterations.

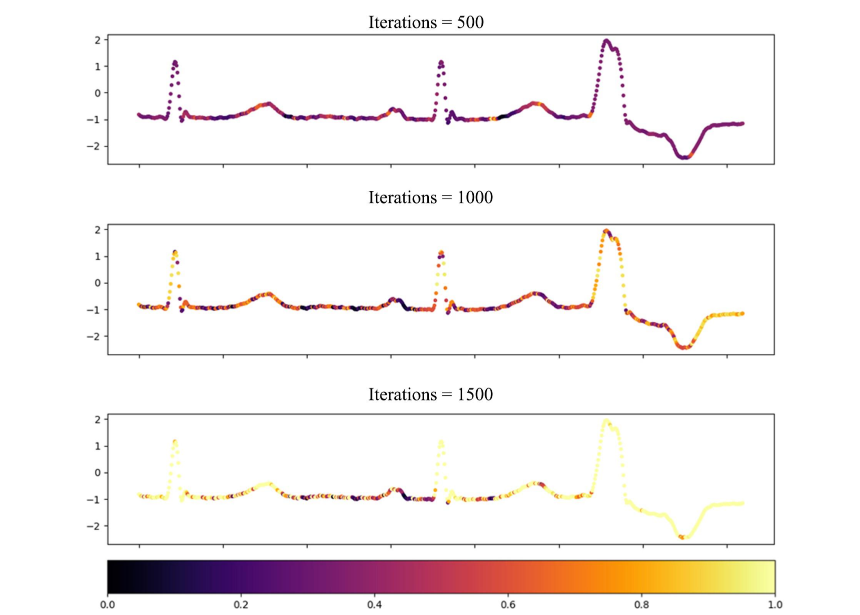


ECG Visualization



CONT.





Confusion Matrix of Predictions

Interpretability w.r.t epochs no

Conclusion and Future scope

- A novel adaptation of visualization techniques of CNN and LSTM for ECG signals was proposed.
- Visualizations were observed to line up with the clinical literature in ECG interpretation
- Extension to other arrhythmia classes
- Extension to entire arrhythmia records
- Exploring Explainability

Healthcare Technology Innovation Centre
Collaborate · Innovate · Impact

— A joint initiative of —

Indian Institute of Technology Madras

Dept. of Biotechnology
Ministry of Science & Technology

Thank you

Contact - sricharanv@htic.iitm.ac.in