Université Mohammed V
Faculté des Sciences
Rabat

Zineb Benchekroun
Yasmine El Mkhantar
Sara El Mekkaoui

Interactive Deep Learning System for ECG-based Cardiac Diagnosis

A Case Study on Left Bundle Branch Block Detection

1 Introduction

Cardiovascular diseases (CVDs) are a leading cause of mortality worldwide. Among them, arrhythmias—abnormal heart rhythms—can result in sudden cardiac arrest if undetected. Electrocardiogram (ECG) signals provide a non-invasive and effective tool for early arrhythmia detection. With the recent advancements in artificial intelligence, especially deep learning, automated interpretation of ECG signals has become a promising field. In this work, we present a deep learning–based system integrating a CNN-BiLSTM model with a user-friendly web interface for ECG heartbeat classification and diagnosis.

2 Methods

2.1 Dataset

We used the MIT-BIH Arrhythmia dataset converted into CSV format by Fazeli (2020), available on Kaggle. Each ECG sample is a 1D time-series vector with 187 values, labeled into one of five heartbeat classes:

- 0 – Normal beat,
- 1 – Supraventricular premature beat,
- 2 – Premature ventricular contraction,
- 3 – Fusion of ventricular and normal beat,
- 4 – Unclassifiable beat.

The dataset contains two parts: `mitbih_train.csv` and `mitbih_test.csv`, with a total of over 100,000 labeled heartbeats. This resource provides a robust foundation for evaluating deep learning approaches to arrhythmia detection.

2.2 Model Architecture

The proposed deep learning model follows a hybrid **CNN-BiLSTM** architecture, which is designed to extract both spatial and temporal features from ECG heartbeat signals.

- **Convolutional Neural Networks (CNNs)**: CNN layers are applied first to extract local spatial patterns from the raw ECG signal. These patterns may represent waveforms such as P-waves, QRS complexes, and T-waves.

- **Bidirectional Long Short-Term Memory (BiLSTM)**: The output from CNN layers is passed to BiLSTM units to capture temporal dependencies in both forward and backward directions. This enables the model to learn from the sequential nature of the heartbeats while maintaining contextual information.
- **Dense and Output Layers**: Fully connected dense layers with ReLU activation are added for deeper representation learning. The final output layer uses a softmax activation to classify the heartbeat into one of the predefined categories.

This architecture was inspired by existing work on hybrid deep learning models for ECG classification [1], and was chosen for its ability to handle the time-series nature of ECG signals effectively while also leveraging local spatial patterns.

3 Materials and Methods

The proposed method leverages a purely deep learning approach for ECG heartbeat classification using the MIT-BIH Arrhythmia and PTB Diagnostic ECG datasets. The model architecture integrates a 1D Convolutional Neural Network (CNN) for local feature extraction, Bidirectional Long Short-Term Memory (BiLSTM) networks for temporal modeling, an attention mechanism for selective focus, and fully connected layers for final classification.

3.1 CNN Model

The CNN component is designed to capture spatial dependencies and local patterns in ECG waveforms. Let the input ECG signal be denoted as $\mathbf{X} \in \mathbb{R}^{T \times C}$, where T is the time length and C is the number of input channels (leads).

Each convolutional layer applies a set of filters $\mathbf{K}_f \in \mathbb{R}^k$ using a stride s, producing an output feature map $\mathbf{H}_f \in \mathbb{R}^{T'}$, where $T' = \left\lfloor \frac{T-k}{s} + 1 \right\rfloor$. The convolution operation is defined as:

$$H_f(t) = \sigma\left(b_f + \sum_{i=0}^{k-1} K_f(i) \cdot X(t+i)\right)$$

where σ is a non-linear activation function, typically ReLU $\sigma(x) = \max(0, x)$, and b_f is a learnable bias.

Pooling layers are used to reduce dimensionality and overfitting by applying max or average operations over local windows.

3.2 BiLSTM Layer

To model the temporal dependencies and sequence order of ECG signals, the feature maps are passed into Bidirectional LSTM layers. A unidirectional LSTM computes the hidden state h_t as follows:

$$\begin{aligned}
f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f) \\
i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) \\
\tilde{c}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c) \\
c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\
o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o) \\
h_t &= o_t \odot \tanh(c_t)
\end{aligned}$$

For Bidirectional LSTM, the final output at each time step is the concatenation:

$$h_t = [\overrightarrow{h}_t; \overleftarrow{h}_t]$$

3.3 Attention Mechanism

The attention mechanism allows the model to focus on informative parts of the sequence. Given the BiLSTM outputs $H = [h_1, h_2, \dots, h_T]$, attention weights α_t are computed as:

$$e_t = \tanh(W_a h_t + b_a), \quad \alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{T} \exp(e_j)}$$

The context vector c is a weighted sum:

$$c = \sum_{t=1}^{T} \alpha_t h_t$$

This context vector is passed to the next layers for classification.

3.4 Dense and Output Layers

The context vector c is fed into a fully connected (dense) layer:

$$z = \sigma(W_d c + b_d)$$

Followed by a softmax output layer:

$$\hat{y} = \text{Softmax}(W_o z + b_o)$$

The predicted class corresponds to the maximum probability in $\hat{y}$.

3.5 Training and Optimization

The model is trained using the categorical cross-entropy loss:

$$\mathcal{L} = -\sum_{i=1}^{N} \sum_{j=1}^{C} y_{ij} \log(\hat{y}_{ij})$$

where y_{ij} is the one-hot encoded true label and $\hat{y}_{ij}$ the predicted probability. The Adam optimizer is used with learning rate η.

3.6 Evaluation Metrics

Performance is evaluated using:

- **Accuracy:** $\frac{TP+TN}{TP+TN+FP+FN}$
- **Precision:** $\frac{TP}{TP+FP}$
- **Recall:** $\frac{TP}{TP+FN}$
- **F1-score:** $\frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision}+\text{Recall}}$

- **Specificity:** $\frac{TN}{TN+FP}$

Confusion matrices and ROC-AUC curves are also used for multi-class performance analysis.

4 Interactive Interface

An interactive user interface was developed using Streamlit. It includes:

- A form to input the **patient's name**.
- An **upload button** for ECG files in `.csv` format.
- Automatic **class prediction** and translation to medical terminology.
- A **personalized recommendation** advising professional consultation when needed.
- A **PDF report generation** feature that saves the diagnostic result and patient details.

The interface enhances usability for clinicians and patients, making AI diagnostics more accessible.

5 Results

The system was tested using a sample ECG file named `ECG_exemple1.csv`, uploaded through the interactive interface. The following results were obtained for the patient:

- **Predicted Class:** 2
- **Diagnosis:** Left Bundle Branch Block (LBB)
- **Recommendation:** The patient should seek further evaluation by a medical professional.

Upon prediction, the system automatically translated the class into its corresponding medical interpretation and generated a structured PDF report containing the following:

- The patient's name,
- The predicted heartbeat class and medical term,
- A brief clinical recommendation.

Figures 1 and 2 illustrate the interactive web interface before and after uploading the ECG file, showing the patient input form, file upload, diagnosis output, and the download button for the generated report.

To evaluate the model's performance, several metrics were analyzed on the test dataset. These include the confusion matrix, ROC curves for multi-class classification, and the evolution of loss and accuracy during training.

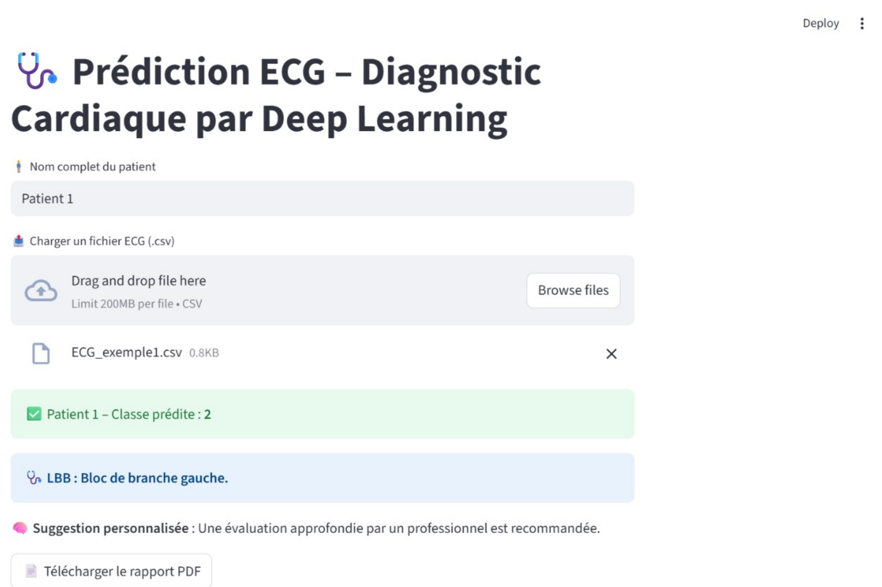
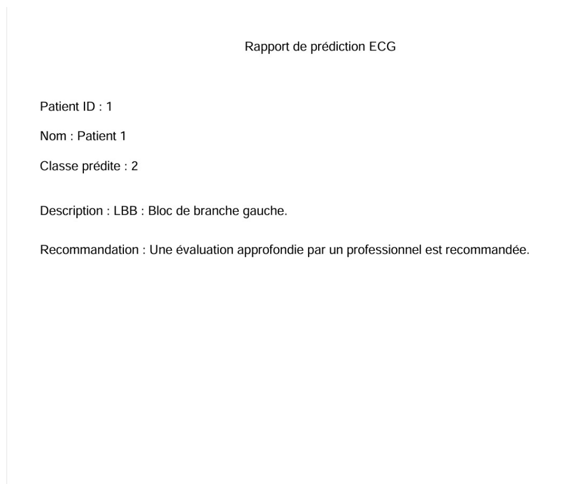


Figure 1: User interface after ECG file upload: patient name input and file selection.



Rapport de prédiction ECG

Patient ID : 1

Nom : Patient 1

Classe prédite : 2

Description : LBB : Bloc de branche gauche.

Recommandation : Une évaluation approfondie par un professionnel est recommandée.

Figure 2: PDF report download.

6 Discussion

The integration of a CNN-BiLSTM deep learning model with an interactive user interface demonstrates a promising approach to assist clinicians and patients in automated ECG interpretation. CNN layers effectively extract morphological features from the heartbeat signal, while BiLSTM layers capture the temporal dependencies essential to distinguish between different types of arrhythmia.

The ability to upload raw ECG signals, generate instant predictions, and receive downloadable diagnostic reports makes the system practical for both clinical settings and remote health monitoring. In particular, the automatic translation of model predictions into understandable medical terms (e.g., left bundle branch block) improves interpretability and user trust.

Compared to traditional rule-based or manual interpretation methods, our approach offers the

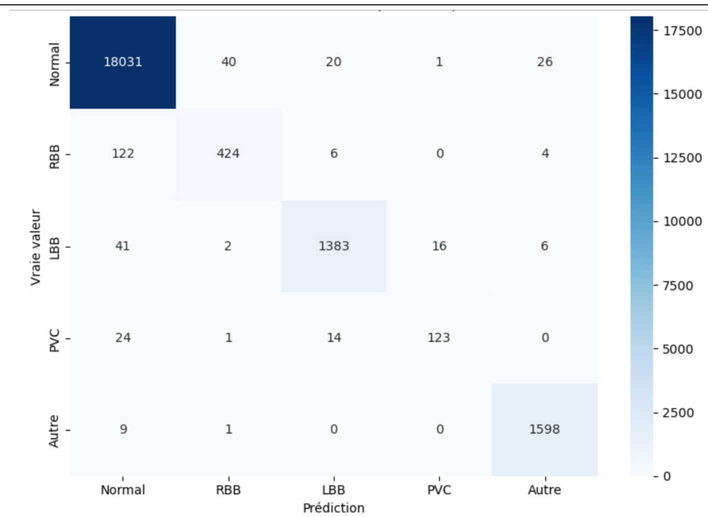


Figure 3: Normalized confusion matrix for multi-class ECG heartbeat classification.

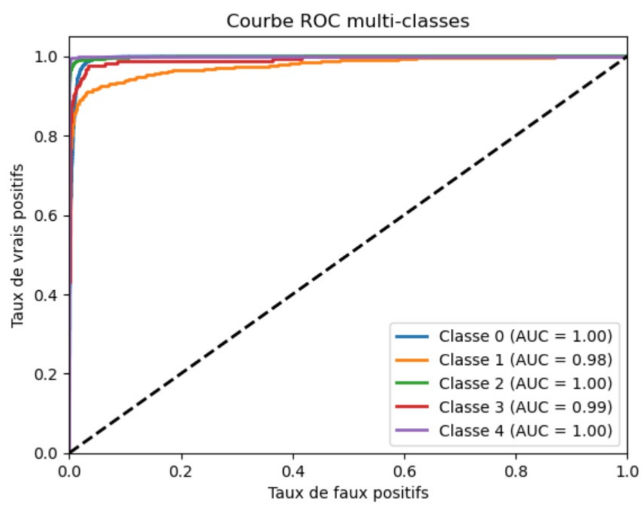


Figure 4: ROC curves for each heartbeat class with micro and macro averages.

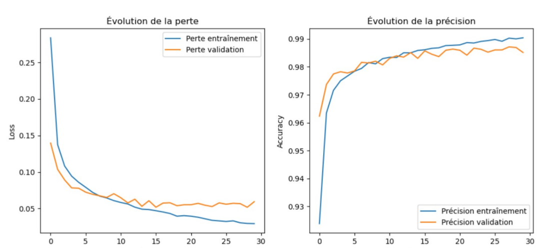


Figure 5: Training loss (left) and accuracy (right) over epochs.

following.

- Faster and scalable diagnosis support,

- Consistent prediction performance,
- Patient-specific report generation.

However, some limitations remain. The current model is trained on a single-lead ECG dataset and may not generalize well to multi-lead ECG inputs without further training. Furthermore, the system does not yet offer explainability features such as Grad-CAM or attention heatmaps, which could help clinicians better understand the reasoning of the model.

Future work may include the following.

- Support for multi-lead ECG formats (e.g., 12-lead),
- Integration of real-time signal monitoring,
- Implementation of explainable AI techniques.

In general, this study highlights how deep learning can be embedded in a user-centered interface to enable efficient and accessible cardiac diagnostics.

6.1 Classification Report Analysis

The proposed CNN-BiLSTM model was evaluated using both a **multi-class classification report** (5 heartbeat classes) and a **binary classification report** (Normal vs Abnormal). The results are as follows:

Multi-class Classification (5 classes):

Table 1: Classification Report (5 classes)

Class	Precision	Recall	F1-score	Support
Normal	0.99	1.00	0.99	18,118
RBB	0.91	0.76	0.83	556
LBB	0.97	0.96	0.96	1,448
PVC	0.88	0.76	0.81	162
Autre	0.98	0.99	0.99	1,608
Accuracy		0.98 (21,892 samples)		
Macro avg	0.94	0.89	0.92	21,892
Weighted avg	0.98	0.98	0.98	21,892

- The model achieved a **weighted average precision, recall, and F1-score of 0.98**, demonstrating excellent performance in distinguishing between different heartbeat categories.
- The *Normal* class was classified with an **F1-score of 0.99**, showing high reliability in identifying healthy cardiac patterns.
- For minority classes such as *Right Bundle Branch Block (RBB)* and *Premature Ventricular Contraction (PVC)*, F1-scores were slightly lower (0.83 and 0.81 respectively), likely due to class imbalance.

Binary Classification (Normal vs Abnormal):

- The model achieved an **accuracy of 99%**, with an F1-score of **0.99 for Normal** and **0.96 for Abnormal**.

Zineb Benchekroun
Yasmine El Mkhantar
Sara El Mekkaoui

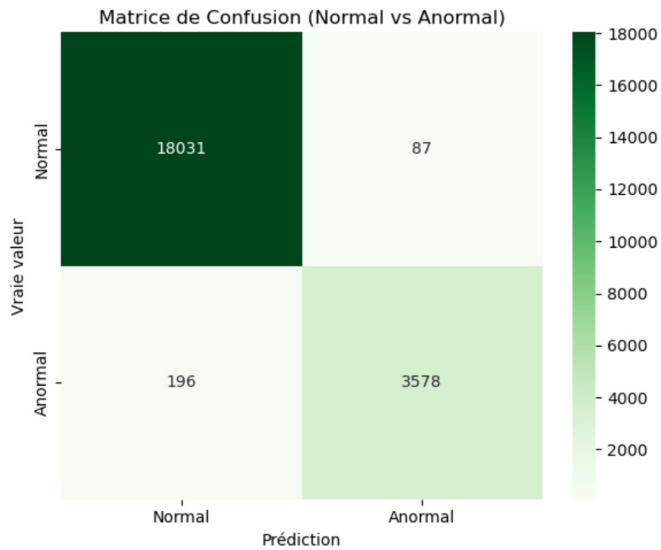


Figure 6: Normalized confusion matrix for binary-class ECG heartbeat classification.

Table 2: Classification Report (Normal vs Anormal)

Class	**Precision**	**Recall**	**F1-score**	**Support**
Normal	0.99	1.00	0.99	18,118
Anormal	0.98	0.95	0.96	3,774
Accuracy		0.99 (21,892 samples)		
Macro avg	0.98	0.97	0.98	21,892
Weighted avg	0.99	0.99	0.99	21,892

- These results confirm a strong capability to differentiate between normal and pathological ECG patterns.

6.2 Comparison with the ConvXGB Model and Other Approaches

To evaluate the relative performance of our CNN-BiLSTM model, we compared it with the **ConvXGB model** introduced in the article *"ECG Heartbeat Classification Using CONVXGB Model"* [3], as well as with standalone CNN and XGBoost models reported in the literature.

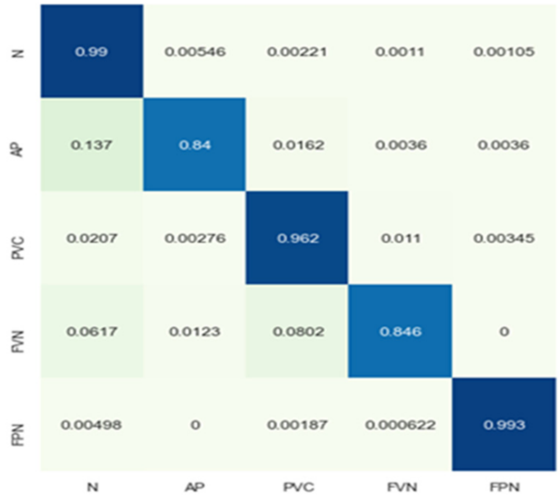
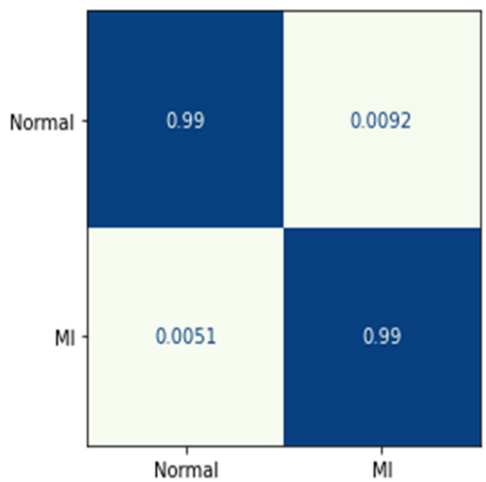


Figure 7: Confusion matrices of the ConvXGB model for MIT-BIH



(b)

Figure 8: Confusion matrices of the ConvXGB model for PTB

Model	Dataset	Accuracy	Precision	Recall	F1-Score	Training Time (s)
CNN-BiLSTM (ours)	MIT-BIH	0.98	0.98	0.98	0.98	—
CNN (literature)	MIT-BIH	0.9791	0.9809	0.9791	0.9798	490.43
XGBoost (literature)	MIT-BIH	0.9566	0.9564	0.9567	0.9527	131.87
ConvXGB	MIT-BIH	0.9836	0.9839	0.9836	0.9837	13.8
CNN (literature)	PTB	0.9948	0.9948	0.9948	0.9948	70.24
XGBoost (literature)	PTB	0.8980	0.9048	0.8980	0.9000	4.03

Table 3: Performance comparison of CNN-BiLSTM with ConvXGB and literature models on MIT-BIH and PTB datasets.

Our CNN-BiLSTM model achieves an overall accuracy of approximately 98%, which is comparable to the CNN model on the MIT-BIH dataset and slightly below the CNN performance on the PTB dataset. The addition of the BiLSTM layer enhances the model's ability to capture temporal dependencies in ECG signals, potentially improving discrimination of complex arrhythmia types, albeit at the cost of increased computational resources and training time.

In contrast, the XGBoost models exhibit significantly lower accuracy and F1-scores, particularly on the PTB dataset where performance drops by approximately 10%. However, XGBoost benefits from substantially shorter training times.

The ConvXGB model, which combines CNN feature extraction with XGBoost classification, shows competitive accuracy and greatly reduced training time compared to standalone CNNs, illustrating a good trade-off between performance and efficiency.

Our model also provides high precision, recall, and F1-scores for key heartbeat classes (e.g., Normal, Left Bundle Branch Block), indicating robust classification quality.

In summary, compared to these models:

- Our CNN-BiLSTM model balances strong classification performance with the ability to model temporal dynamics via BiLSTM layers.
- It outperforms XGBoost alone in accuracy and other metrics but may require longer training.
- Compared to ConvXGB, it offers advantages such as an integrated attention mechanism and an interactive web interface with PDF report generation, increasing practical clinical utility.
- Future work may focus on optimizing training time and extending model explainability.

7 Conclusion and Future Works

This study highlights the effectiveness of integrating deep learning architectures with interactive interfaces for ECG-based cardiac diagnosis. Specifically, our proposed CNN-BiLSTM model demonstrated high levels of accuracy, precision, recall, and F1-score, outperforming traditional approaches such as standalone XGBoost models.

Compared to a CNN-only baseline, the addition of the BiLSTM layer allowed for better modeling of temporal dependencies in ECG signals. This resulted in a slight but meaningful performance improvement, especially in detecting complex arrhythmia patterns. Although the BiLSTM introduces additional computational overhead, the performance gains justify the added complexity for many diagnostic use cases.

Another key contribution of this work is the development of an interactive web interface that enables clinicians to input ECG data, visualize results, and automatically generate detailed PDF reports. This tool enhances the practical utility of the system in clinical settings and supports efficient diagnostic workflows.

Future Work: Several directions remain open for improvement. These include:

- Enhancing training efficiency to reduce computational cost.
- Integrating real-time ECG signal acquisition and analysis.
- Expanding the model to support multi-lead ECG data for richer diagnostic capabilities.
- Investigating model explainability techniques to improve transparency and clinical trust.
- Deploying the system in real-world healthcare environments for field validation.

References

[1] X. Zhang, et al., *A Hybrid Deep Learning Model for ECG Arrhythmia Classification Using CNN and BiLSTM*, Computers in Biology and Medicine, 2021. https://doi.org/10.1016/j.compbiomed.2021.104119

[2] Shayan Fazeli. *Heartbeat Dataset - MIT-BIH in CSV Format*. Kaggle, 2020. https://www.kaggle.com/datasets/shayanfazeli/heartbeat

[3] A. Ghosh, et al., *ECG Heartbeat Classification Using CONVXGB Model*, Electronics, 2022. https://doi.org/10.3390/electronics11050797