Project : Software Competencies

Filière : GIIADS

Année universitaire : 2025–2026

– Version 2025.1 –

CardioECHO : Precision Cardiology through Artificial Intelligence

---

Présenté par :
Azami Hassani Adnane

Encadré par :
Mr. Hosni Mohammed

Résumé

This report details the development of **CardioEcho**, an end-to-end system for the automated assessment of cardiac function. The project begins with a comprehensive statistical analysis of the EchoNet-Dynamic dataset (10,030 studies) to characterize key physiological distributions. Building on this data, a hybrid deep learning pipeline was developed, utilizing **DeepLabV3** for semantic segmentation and 3D spatiotemporal networks (**MC3** and **R(2+1)D**) for the prediction of Ejection Fraction (EF) and End-Systolic Volume (ESV). These models were integrated into a standalone desktop application built with PyQt6. The resulting software features a "human-in-the-loop" workflow, enabling clinicians to visualize AI predictions, manually correct segmentation masks, and generate standardized clinical reports, thereby bridging the gap between deep learning research and practical clinical deployment.

Chapitre 1

Introduction and Dataset Analysis

1.1 Introduction

The **EchoNet-Dynamic dataset** is a publicly available collection of echocardiogram videos designed to advance machine learning research in cardiac function assessment. This analysis focuses on understanding the dataset's structure, variable distributions, and data quality to inform subsequent modeling efforts.

1.1.1 Project Objectives

— Characterize the distribution of cardiac function measurements
— Assess data quality and completeness
— Analyze relationships between cardiac variables
— Examine temporal tracking data structure
— Validate dataset consistency across train/validation/test splits

1.2 Dataset Description

The EchoNet-Dynamic dataset consists of three primary components :

1.2.1 FileList Dataset

Contains summary measurements and metadata for each echocardiogram video :

— **Total Records** : 10,030 echocardiogram studies corresponding to the videos
— **Key Variables** : Ejection Fraction (EF), End-Systolic Volume (ESV), End-Diastolic Volume (EDV), such as :

$$\mathrm{EF} = \frac{\mathrm{EDV} - \mathrm{ESV}}{\mathrm{EDV}} \times 100\%$$

— **Video Properties** : Frame dimensions, frame rate (FPS), number of frames
— **Data Splits** : Training (7,465), Validation (1,288), Test (1,277)

1.2.2 VolumeTracings Dataset

Contains frame-by-frame bounding box coordinates for cardiac region tracking :
- **Coordinate Variables** : X1, Y1, X2, Y2 (bounding box corners)
- **Temporal Variable** : Frame number
- **Expected Structure** : 42 measurements per video file
- **Purpose** : Temporal tracking of cardiac regions throughout cardiac cycles

1.2.3 AVI files

Short, unannottated videos of A4C **(Apical 4 chamber)** view of 10030 patients.

1.3 Variable Definitions

1.3.1 Cardiac Function Measurements

Variable	Description	Units
EF	Ejection Fraction	Percentage (%)
ESV	End-Systolic Volume	Milliliters (mL)
EDV	End-Diastolic Volume	Milliliters (mL)
Stroke Volume	EDV - ESV	Milliliters (mL)

Table 1.1 – Cardiac Function Variables

1.3.2 Video Characteristics

Variable	Description	Units
FPS	Frames Per Second	Hz
NumberOfFrames	Total video frames	Count
Vid_duration	Calculated video length	Seconds
FrameHeight	Video frame height	Pixels
FrameWidth	Video frame width	Pixels

Table 1.2 – Video Property Variables

1.4 Statistical Analysis Results

1.4.1 Cardiac Function Distribution

Ejection Fraction (EF)

- **Mean** : 55.7% ± 12.4%

- **Range** : 6.9% - 97.0%
- **Distribution** : Normal-like with slight left skew

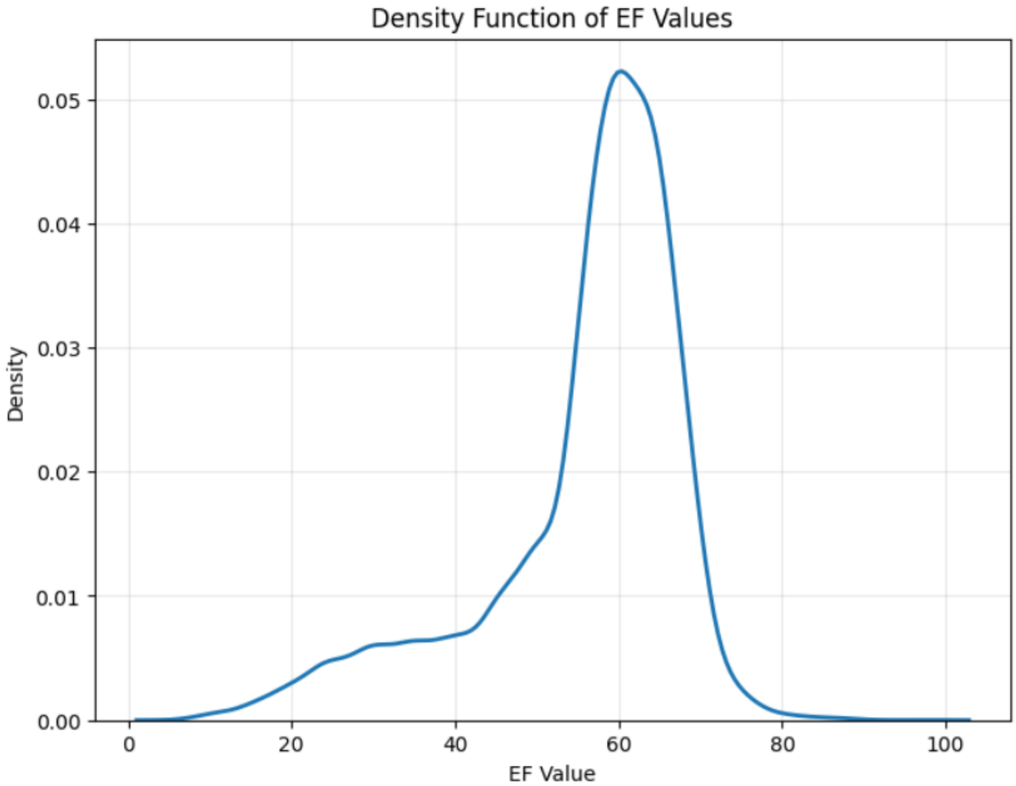


FIGURE 1.1 – Density function of EF values

- **Clinical Categories** :
  - **Upper** (EF $\geq$ 70%) : 781 cases (7.8%)
  - **Normal** (50% $<$ EF $<$ 70%) : 3,249 cases (32.4%)
  - **Reduced** (EF $\leq$ 50%) : 6,000 cases (59.8%)

Ventricular Volumes

Variable	Mean	Std	Min	Max	Median
ESV (mL)	43.4	35.8	4.4	612.5	33.6
EDV (mL)	91.3	45.7	12.6	695.0	82.1
Stroke Volume (mL)	47.9	18.8	5.3	178.8	45.8

TABLE 1.3 – Ventricular Volume Statistics

1.4.2 Video Characteristics

- **Video Duration** : Mean 3.46s $\pm$ 1.13s (Range : 0.56s - 20.04s)
- **Frame Rate** : Predominantly 50 FPS (median), Range : 18-138 FPS
- **Frame Dimensions** : Primarily 112×112 pixels (standard resolution)
- **Number of Frames** : Mean 176.5 $\pm$ 57.9 frames

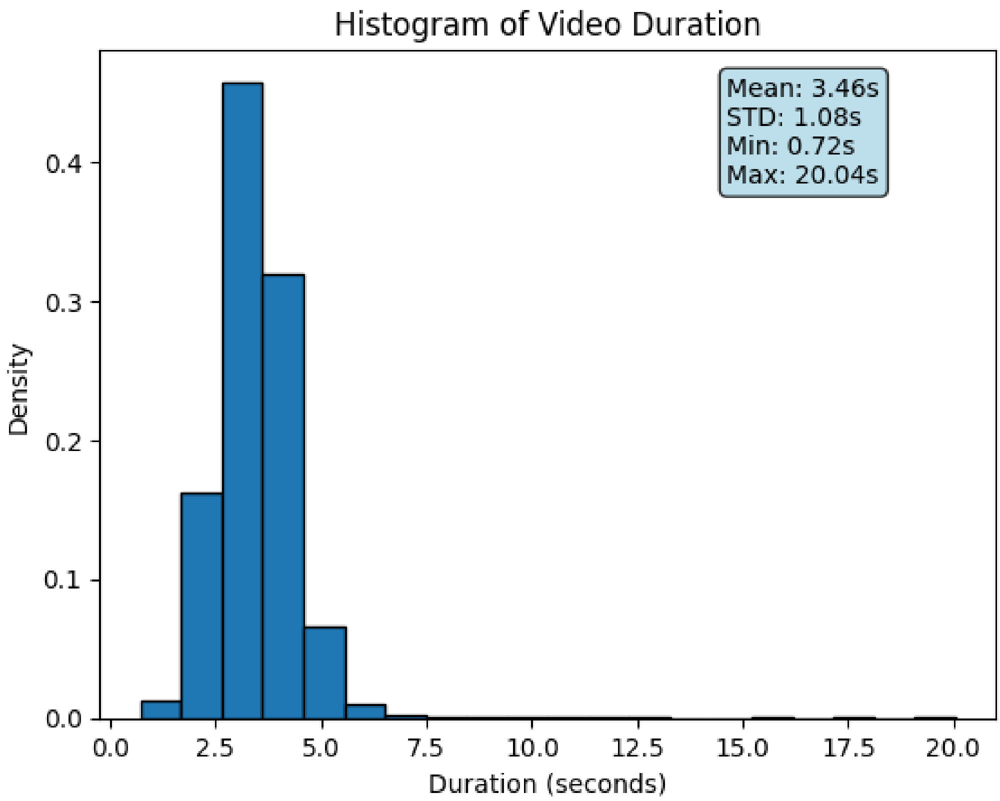


FIGURE 1.2 – AVI Duration Histogram

1.4.3 Correlation Analysis

Strong correlations ($|r| > 0.5$) identified :

- **EDV - Stroke Volume** : $r = 0.772$ (expected physiological relationship)
- **ESV - EF** : $r = -0.845$ (inverse relationship as expected)
- **EDV - EF** : $r = -0.532$ (moderate inverse relationship)

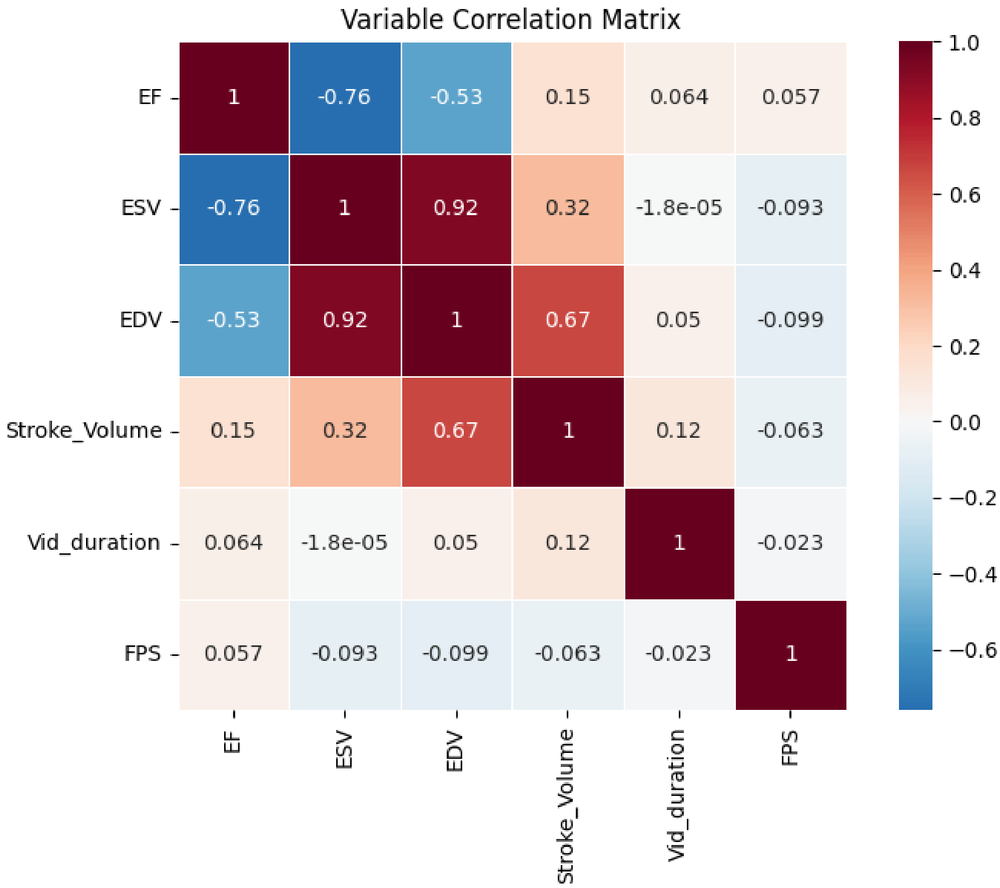


FIGURE 1.3 – Feature Heatmap

1.5 Volume Tracings Analysis

1.5.1 Data Structure Assessment

- **Expected Pattern** : 42 coordinate measurements per video
- **Compliance Rate** : 99.06% (93 out of 10025 files have more or less instances (ranging from (34-252) instances))
- **Coordinate Ranges** :
  - X coordinates : Span image width (112 pixels typical)
  - Y coordinates : Span image height (112 pixels typical)

1.6 Data Quality Assessment

1.6.1 Dataset Completeness

Quality Metric	Count	Percentage
Total Studies	10,030	100.0%
Missing EF Values	0	0.0%
Missing Volume Data	0	0.0%
EF Outliers (<30% or >80%)	1,183	11.8%

Table 1.4 – Data Quality Summary

1.6.2 Dataset Split Balance

The train/validation/test splits show consistent statistical properties :
— **Training Set** : 74.4% (7,465 studies)
— **Validation Set** : 12.8% (1,288 studies)
— **Test Set** : 12.7% (1,277 studies)
— **Statistical Balance** : No significant differences in mean EF, ESV, or EDV across splits

1.7 Key Findings

1.7.1 Dataset Characteristics

1. The dataset contains a **representative distribution** of cardiac function, with a higher prevalence of reduced ejection fraction cases
2. Video characteristics are **standardized** (primarily 112×112 resolution, ∼50 FPS)
3. **Strong physiological correlations** confirm data validity
4. **Balanced train/validation/test splits** ensure robust model evaluation capability

1.7.2 Clinical Relevance

1. EF distribution reflects **clinical populations** with cardiac disease
2. Volume measurements span **normal to severely impaired** cardiac function
3. Temporal tracking data enables **dynamic cardiac analysis**
4. Dataset suitable for **supervised learning** of cardiac function assessment

1.8 Recommendations

1.8.1 Data Preprocessing

— Standardize video durations and frame rates if needed

— Validate coordinate tracking completeness (42 measurements per file)
— Consider outlier handling for extreme EF values

1.8.2 Modeling Considerations

— Leverage strong ESV-EF correlation for validation
— Utilize temporal coordinate data for region-of-interest focusing
— Consider class imbalance in EF categories for classification tasks
— Maintain dataset split integrity for fair evaluation

1.9 Conclusion

The EchoNet-Dynamic dataset provides a **comprehensive foundation** for cardiac function analysis research. The descriptive analysis reveals well-structured, complete data with physiologically consistent relationships between variables. The dataset's balance across clinical conditions and standardized video properties make it suitable for developing **robust machine learning models** for automated cardiac function assessment.

The temporal coordinate tracking data adds valuable **spatial-temporal context**, enabling sophisticated analysis of cardiac motion patterns. Data quality is high with minimal missing values and consistent statistical properties across dataset splits.

Technical Notes

Analysis Software : Python with pandas, matplotlib, seaborn, and numpy libraries
Statistical Methods : Descriptive statistics, correlation analysis, distribution analysis
Dataset Version : As provided in CSV format (FileList.csv, VolumeTracings.csv)

Chapitre 2

Feasability Analysis

2.1 Project Objectives

1. **Image & Video Processing** – Develop algorithms to analyze both still images and moving sequences of the heart.
2. **Cardiac Structure Detection** – Automatically segment chambers, valves, and myocardial walls.
3. **Functional Measurement Extraction** – Derive key parameters such as ejection fraction, chamber volumes, wall thickness, valve motion, and blood flow patterns.
4. **Data Structuring** – Convert extracted measurements into standardized formats (DICOM, HL7 FHIR) for integration with hospital systems.
5. **Validation & Reliability** – Ensure clinical accuracy through comparison with expert cardiologist evaluations and benchmark datasets.
6. **User Platform** – Provide a clinician-friendly interface to upload echocardiographic images or videos and obtain structured digital reports.

Deliverables

— AI algorithms for both image and video-based cardiac analysis.
— A validated dataset of annotated echocardiogram images and cine loops.
— A prototype application for automated measurement and reporting.
— Performance validation reports against expert cardiologist assessments.

Feasibility Assessment and Proposed Solutions

Task	**Feasibility**	**Proposed Solution / Notes**
Image & Video Processing	High	Use OpenCV/PyAV for preprocessing frames and videos ; normalize and resize ; data available in EchoNet-Dynamic.
Cardiac Structure Detection	Medium-High	Segment chambers, valves, and myocardial walls using U-Net / Mask R-CNN ; feasible for LV segmentation with EchoNet-Dynamic, other chambers may require additional annotation.
Functional Measurement Extraction	Medium	Ejection fraction, EDV, ESV : feasible using pretrained EchoNet-Dynamic models.
Wall Thickness	Low	Requires epicardial and endocardial border annotations ; EchoNet-Dynamic does NOT have this data ; solution : manual annotation of 500–1000 studies or acquire CAMUS dataset.
Valve Motion	Low	Temporal action recognition (3D CNN or CNN+LSTM) with valve landmark detection ; EchoNet-Dynamic does NOT have frame-level valve state labels ; requires new annotations or hospital data.
Blood Flow Patterns	Very Low	Doppler signal processing and classification ; EchoNet-Dynamic does NOT have Doppler images ; must acquire hospital/commercial datasets for training.
Data Structuring & Integration	High	Convert extracted measurements to DICOM / HL7 FHIR ; independent of data source ; can start once extraction pipeline is ready.

Table 2.1 – Feasibility and Proposed Solutions for Cardiac Analysis Tasks

Chapitre 3

Methodology and Model Development

3.1 Overview of the Pipeline

The core processing pipeline of the **CardioEcho** system is designed to transform raw echocardiogram videos into actionable clinical metrics. The pipeline consists of four deep learning components operating sequentially and in parallel :

1. **Semantic Segmentation** : Frame-by-frame delineation of the Left Ventricle (LV).
2. **End-Systolic Volume (ESV) Prediction** : A spatiotemporal regression model estimating the minimum LV volume.
3. **Ejection Fraction (EF) Prediction** : A video-based regression model estimating global cardiac function.
4. **End-Diastolic Volume (EDV) Derivation** : A mathematical derivation combining EF and ESV predictions.

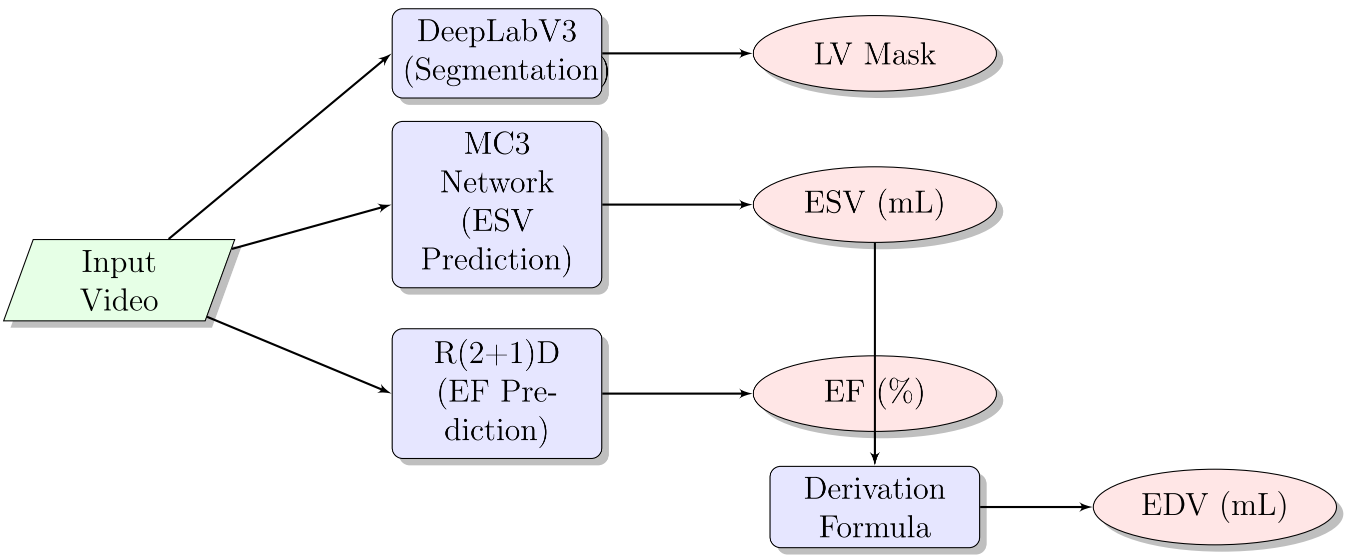


Figure 3.1 – High-level architecture of the CardioEcho processing pipeline showing the parallel flow of segmentation and regression tasks.

3.2 Data Preprocessing

Prior to model training, the EchoNet-Dynamic dataset underwent rigorous preprocessing to ensure consistency across inputs.

3.2.1 Video Standardization

— **Frame Extraction** : Videos were sampled into fixed-length clips (typically 32 frames) to capture at least one full cardiac cycle.
— **Spatial Resizing** : Frames were resized to 112×112 pixels to reduce computational cost while preserving anatomical detail.
— **Normalization** : Pixel intensities were normalized (min–max scaling or mean–standard deviation normalization) to stabilize training.

3.3 Deep Learning Models

3.3.1 Left Ventricle Segmentation (DeepLabV3)

To visualize cardiac wall motion and compute 2D ventricular areas, a semantic segmentation approach was employed.

— **Architecture** : DeepLabV3 with a ResNet-50 backbone.
— **Objective** : Pixel-wise classification into *Left Ventricle* or *Background*.
— **Rationale** : The Atrous Spatial Pyramid Pooling (ASPP) module captures multi-scale contextual information, which is crucial for separating the LV boundary from surrounding tissue with similar echogenicity.
— **Training** : A combination of Dice Loss and Cross-Entropy Loss was used to mitigate class imbalance.

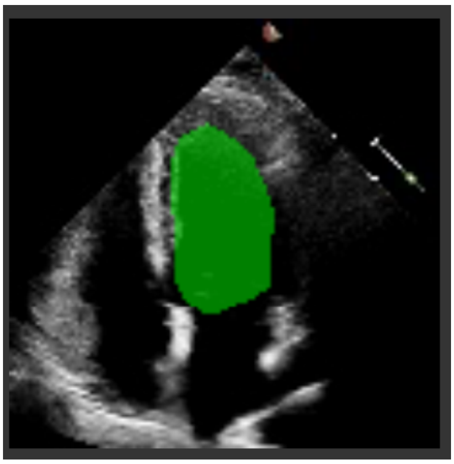


FIGURE 3.2 – Example of Left Ventricle segmentation mask overlay (green) on an A4C view.

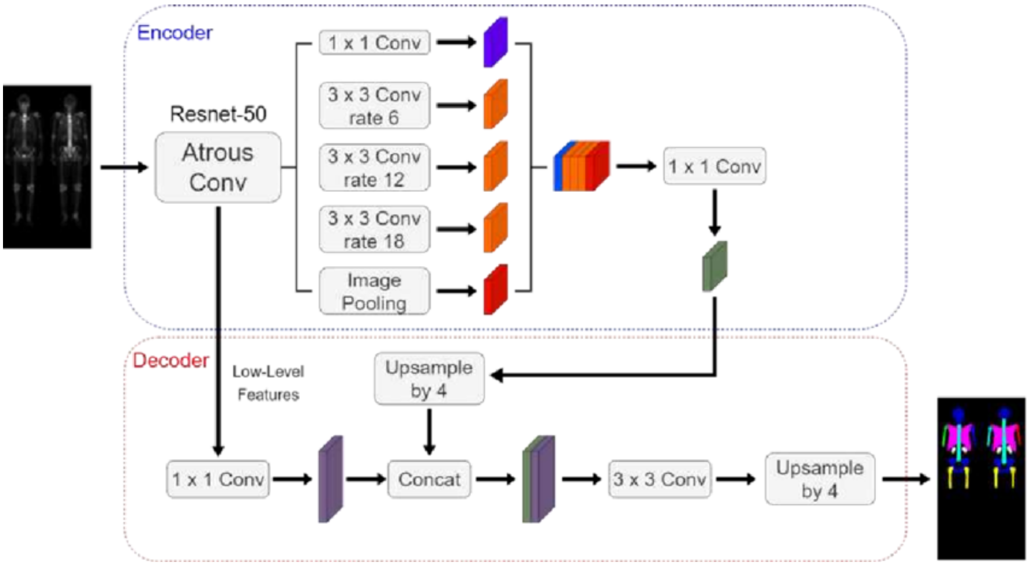


FIGURE 3.3 – DeepLabV3 with a ResNet-50 backbone

3.3.2 End-Systolic Volume Prediction (MC3)

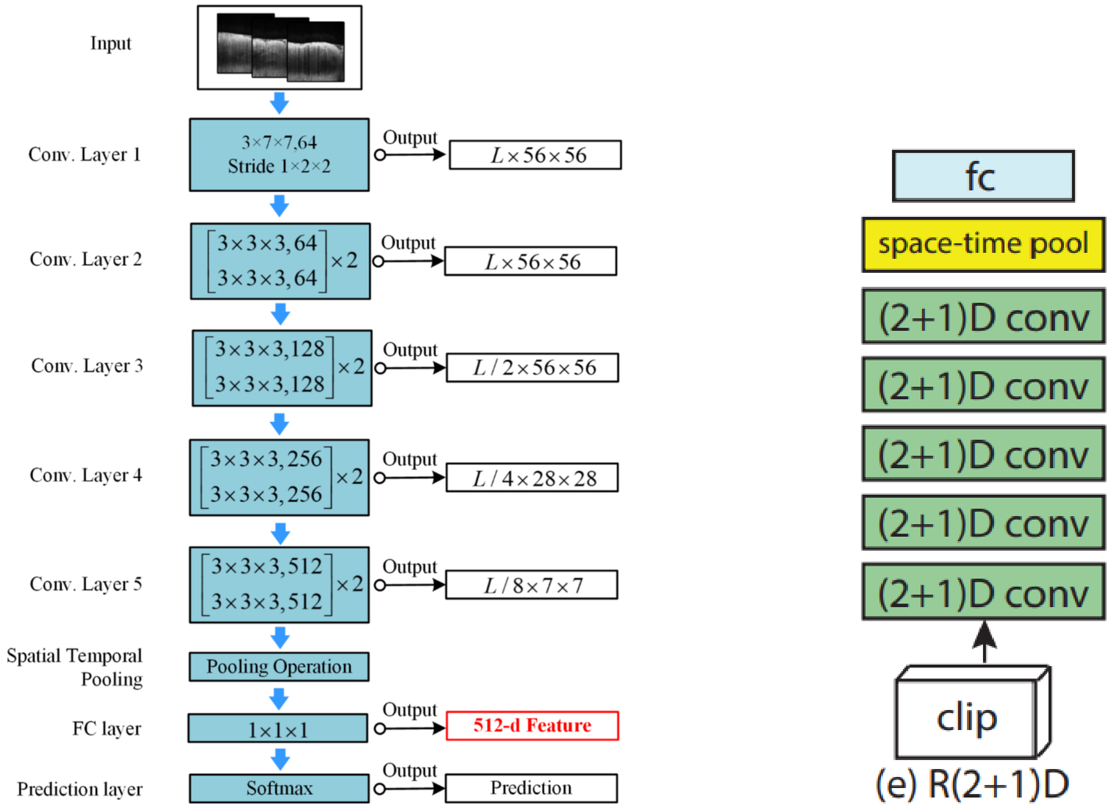
Estimating ventricular volume from 2D echocardiographic video requires spatiotemporal understanding.

— **Architecture** : MC3 (Mixed Convolutional 3D) network.
— **Mechanism** : Factorization of 3D convolutions into 2D spatial and 1D temporal operations, reducing model complexity while preserving temporal modeling.
— **Input** : Video tensor of shape (C, T, H, W).
— **Output** : Scalar End-Systolic Volume (ESV) in milliliters (mL).

3.3.3 Ejection Fraction Prediction (R(2+1)D)

Ejection Fraction is a dynamic cardiac metric requiring robust motion analysis.

— **Architecture** : R(2+1)D-18.
— **Mechanism** : Decomposed 3D convolutions, pre-trained on large-scale video datasets (e.g., Kinetics-400) and fine-tuned on EchoNet-Dynamic.
— **Output** : Scalar Ejection Fraction (EF) expressed as a percentage.



(a) mc3 architecture (b) ResNets with (2+1)D convolutions

FIGURE 3.4 – Regression architecture

3.4 Derived Metrics

While ESV and EF are directly predicted by neural networks, End-Diastolic Volume (EDV) is derived mathematically to enforce physiological consistency. Based on the clinical definition of EF :

$$EF = \frac{EDV - ESV}{EDV} \times 100$$

Rearranging yields :

$$EDV = \frac{ESV}{1 - \frac{EF}{100}}$$

This formulation guarantees coherence between the reported metrics, preventing physiologically implausible outputs (e.g., $EDV < ESV$).

Chapitre 4

Application Development

4.1 The CardioECHO Desktop Application

To translate research models into a practical clinical tool, a standalone desktop application named **CardioECHO** was developed. The application uses **PyQt6** for the graphical user interface and integrates PyTorch-based inference.

4.1.1 System Architecture

The application follows a modular architecture to ensure responsiveness and stability :

— **Frontend (GUI)** : Manages user interaction, video visualization, and plotting.
— **Worker Threads** : All inference tasks run in background `QThreads` to prevent UI blocking.
— **Backend Logic** : Handles model loading, frame extraction using OpenCV, and cardiac cycle signal processing.

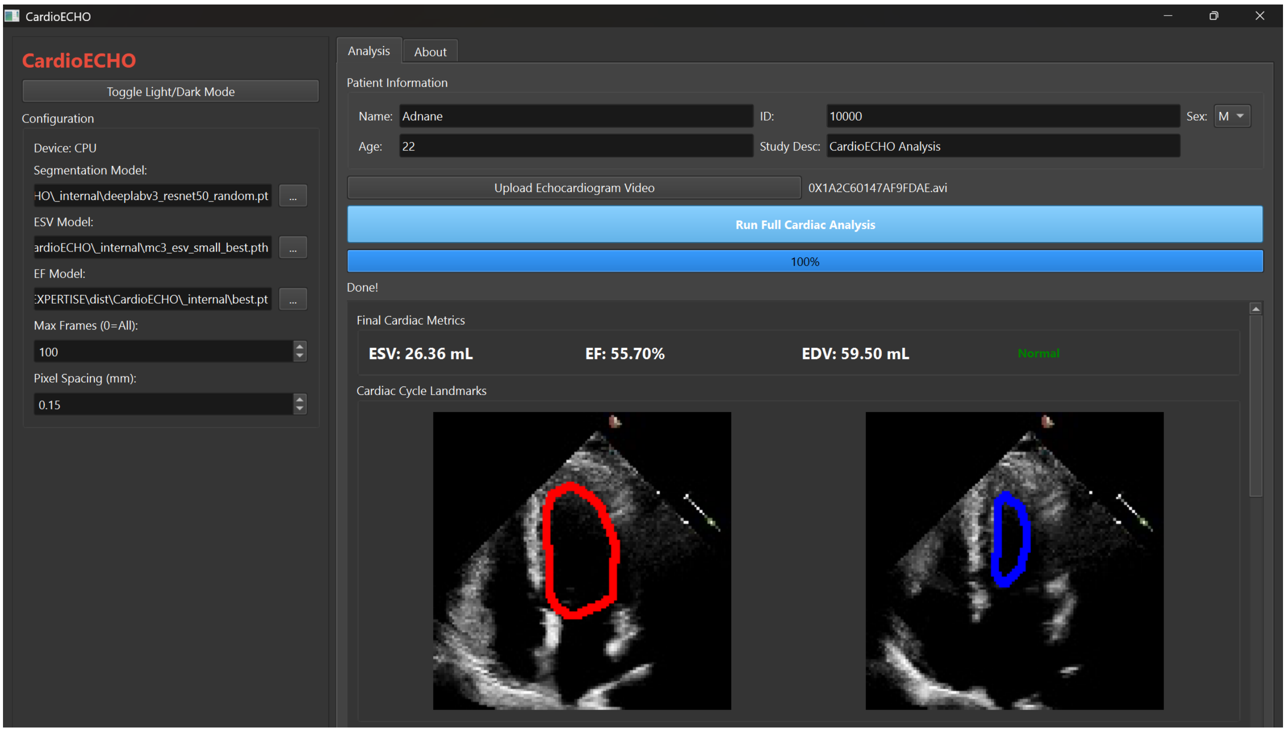


FIGURE 4.1 – Screenshot of the CardioECHO main interface showing the analysis dashboard.

4.2 Key Features

4.2.1 Automated Analysis Pipeline

Users can upload standard echocardiogram videos (AVI/MP4). The system automatically :

1. Segments the LV in each frame.
2. Identifies End-Diastole (ED) and End-Systole (ES) based on maximal and minimal LV areas.
3. Predicts EF, ESV, and EDV.
4. Generates a temporal LV area curve.

4.2.2 Interactive Frame Browser and Correction

To accommodate occasional model errors, a human-in-the-loop mechanism is provided :

- **Frame Navigation** : Slider-based frame-by-frame inspection.
- **Mask Editing** : Manual segmentation correction using brush tools.
- **Recalculation** : Updated metrics and plots after manual edits.

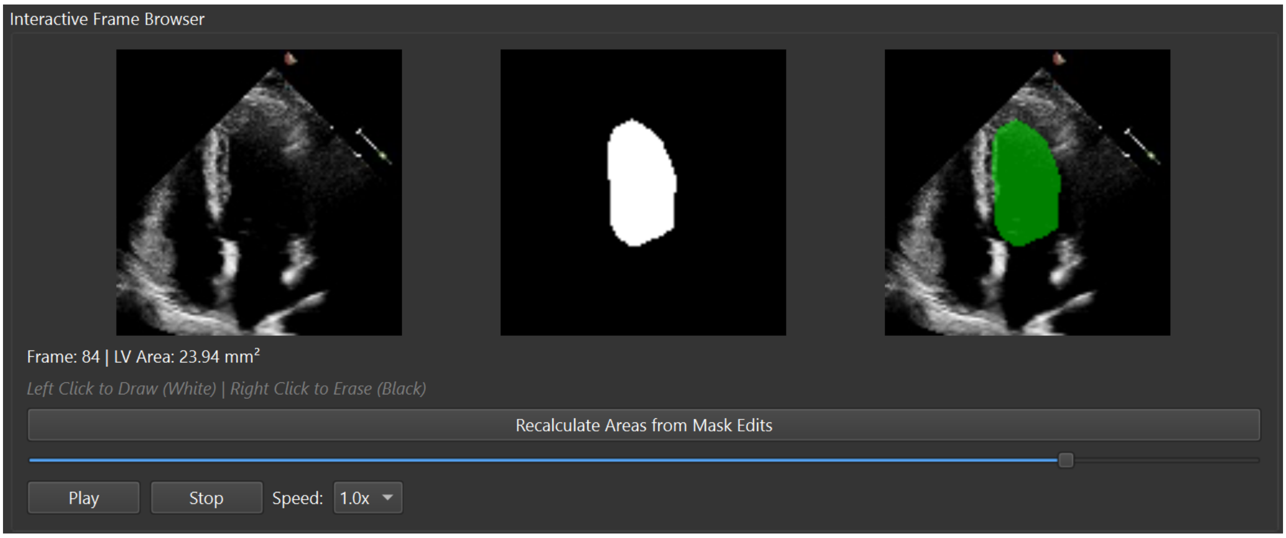


FIGURE 4.2 – Interactive frame browser enabling manual segmentation correction.

4.2.3 Clinical Reporting and Export

The application supports clinical workflow integration through :

— **PDF Reports** : Structured reports including patient data, metrics, LV area curves, and ED/ES keyframes.

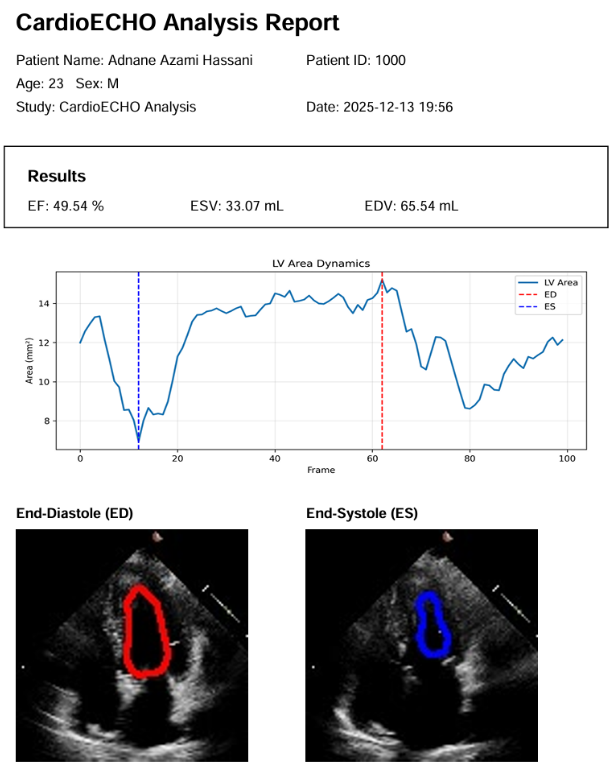


FIGURE 4.3 – PDF EXPORT

— **DICOM Export** : Conversion of processed videos to DICOM format with AI-derived metrics embedded in private tags (group `0x0011`).

FIGURE 4.4 – DICOM EXPORT

4.3 Performance Optimization

For deployment on standard medical workstations, the application was compiled into a standalone executable using **PyInstaller**. This process involved :

- Bundling PyTorch and OpenCV dependencies.
- Embedding trained model weights (`.pt`, `.pth`) into the executable.
- Implementing robust runtime resource path handling.

Chapitre 5

Conclusion and Future Work

5.1 Summary of Achievements

This project successfully bridges the gap between medical AI research and real-world deployment. Leveraging the EchoNet-Dynamic dataset, robust deep learning models were developed for segmentation and functional cardiac assessment. These models were integrated into **CardioECHO**, a clinician-oriented application that automates measurements while preserving expert oversight.

5.2 Future Directions

Future work will address the limitations identified in Chapter 2 :

— **Multi-View Analysis** : Extension to Parasternal Long Axis (PLAX) views.
— **Real-Time Processing** : Model optimization (e.g., TensorRT) for live inference.
— **Longitudinal Tracking** : Integration of a local patient database for follow-up analysis.