International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

A Multi-Modal Deep Learning System for Early Staging of Liver Fibrosis and Prediction of Diastolic Dysfunction Disease Risk

Mr. Jacob Finny[1], Dasari Sri Harsha Vardhan[2], Adireddi Likitha[3], Gadamsetti Ramesh[4]

[1]Assistant Professor, Department of Computer Science and Engineering, Nadimpalli Satyanarayan Raju Institute of Technology, Andhra Pradesh, India.
[2,3,4]Student, Department of Computer Science and Engineering, Nadimpalli Satyanarayan Raju Institute of Technology, Andhra Pradesh, India.

Abstract

Background: Left Ventricular Diastolic Dysfunction (DD) is a precursor to Heart Failure with Preserved Ejection Fraction (HFpEF). The traditional methods for diagnosing DD are based on the evaluation of a single parameter, resulting in ambiguous outcomes for borderline presentations. We propose an automated system that integrates inferences from multiple models to classify the Heart Liver Axis, where progressive Liver Fibrosis directly effects the circulation, acting as cause for the development of Diastolic Dysfunction.

Methods: We designed a late decision-level convergence system that incorporates three different networks: a Hepatic module that uses a two-classifier system and EfficientNetB0 to classify the extent of Liver Fibrosis (METAVIR Staging, F0 to F4); a Cardiac module that uses a spatio-temporal tracking system combining MobileNetV2 and LSTM to track ejection fractions from 4-chamber videos; and a Clinical module that uses an XGBoost system to classify 11 baseline factors. These individually derived probabilities are then passed into a late-fusion Multi-Layer Perceptron (MLP) mediator module. We handled data imbalance using various augmentations that includes a Gaussian jitter and a synthetic minority over-sampling (SMOTE) strategy.

Results: The mediator was then validated on a 2,000-sample random population, getting a 93.74% balanced accuracy rating that range from a normal health to a complicated pathology. The models also validated that employing a Gradient Weighted Class Activation Mapping (Grad-CAM), the models verified spatial recognition capabilities on known anatomical targets. The results were then synthesized into an integrated AI LLM Chatbot, allowing physicians to interact with the diagnostic logic.

Conclusion: The results directly serve as a Clinical Decision Support System (CDSS), integrating isolated clinical markers to reduce classification uncertainty, a common problem with multi- stage classification, for physicians to make evidence-based interventions.

Keywords: Diastolic Dysfunction, Liver Fibrosis, Clinical Decision Support System, Heart-Liver Axis, Multi-Modal Fusion, EfficientNetB0, MobileNetV2, Grad-CAM, Generative AI Chatbot, Deep Learning, METAVIR Staging.

1. Introduction

1.1 Clinical Background

The escalation of Heart Failure with Preserved Ejection Fraction (HFpEF) poses a diagnostic challenge in contemporary cardiology, driven by early-stage diastolic dysfunction (DD) progression. Standard non-invasive reviews rely on threshold parameters like the E/e' ratio, tricuspid regurgitation tracking (TRV), and left atrial volume indices to designate severity across three progressive deterioration grades [2]. Since single clinical attributes frequently fall within overlapping, borderline ranges, single-modality automated prediction models struggle with rates of indeterminate clinical results.

1.2 The Heart-Liver Axis & CDSS Philosophy

We speculate that monitoring multi-organ circulation can counter the confusions raised due to isolated imaging. Progressive liver fibrosis and cardiac wall stiffening significantly causes elevated resistance to left ventricular ejection and liver blood flows back to the heart. Due to long term damage to the liver chemically and mechanically forces the heart to develop Left Ventricular Diastolic Dysfunction (LVDD) [10, 8]. Hence, accurately staging liver texture via B-mode ultrasound provides a primary support for prediction of the cardiac disease risk.

Based on this inter-connected biological concept Heart-Liver Axis, we proposed our idea of development functions as a Clinical Decision Support System (CDSS). We developed this system to collect multi-modal inferences and present integrated diagnostic results that assist a physician or clinician rather than replacing them.

1.3 Contributions

The primary contributions of this paper feature:

1. A functional mathematical translation of the Heart-Liver Axis hypothesis into an automated, non-invasive DD risk grading pipeline.
2. A Multi-Layer Perceptron decision based fusion network that joins separate hepatic, cardiac, and clinical models inferences to address the prediction of Diastolic Dysfunction risk including the fibrosis staging.
3. Based on available datasets validation this system provides a 93.74% multi-stage accuracy along with a Large Language Model (LLM) based Chatbot module developed for clinical based interaction.

2. Literature Survey

So, when it comes to using computers to help with medicine, we've found that really smart computer programs called convolutional neural networks (CNNs) are super helpful for making accurate diagnoses. At first, we tried using a simple CNN (like VGG) to look at specific markers in the body, like how bad scarring is in an organ. But it wasn't too accurate — it got about 75% right. This was because it's pretty hard to spot when cells are starting to get a bit damaged [1]. But then, we tried something else. We used AI to look at echocardiogram videos of the heart and it was good at figuring out how well the heart is pumping blood. This was part of a project called EchoNet-Dynamic, and it was nice because it could do as well as doctors at estimating this [4].

However, despite software expansions in these respective niches, clinical data streams remain siloed [3]. The formal integration of cross-organ data into reliable algorithmic structures is uncommon, culminating in systems unable to account for systemic physiological interactions. Furthermore, while technical transparency structures like Grad-CAM [9] characterize ethical medical deployments, seamlessly summarizing these dynamic visual confirmations organically back to the physician inside functional

hospital reporting logic continues to represent a capability gap across literature.

3. Methodology

3.1 Dataset Assembly and Preprocessing Architecture

We utilized distinct databanks to train our independent evaluative environments. The sonography resource encompassed 1,542 unique B-mode liver frames, cleared via 64-bit perceptual hashing distance rules to rigorously eliminate identical temporal cycle duplicates. The cardiac evaluation pipeline ingested 10,030 standardized apical 4-chamber videos obtained from a curated Stanford cohort [4]. Concurrently, clinical validation parsed 10,000 algorithmic patient profiles structurally modeled on 11 core echocardiographic ranges dictated by global cardiological authorities [2].

Raw B-mode ultrasound frames inherently contain non-anatomical artifacts including text overlays, medical scale markers, and probe identifiers that blindly obscure neural learning. We engineered an automated multi-step contour-based fan-sector extraction pipeline using OpenCV. The sequence engaged RGB-to- grayscale conversion, strict binary thresholding to isolate the acoustic fan from the background, and morphological dilation to bridge gaps along the fan boundary. Finally, the maximum-area bounding contour

was retrieved to establish the primary Region of Interest (ROI), resulting in a clean anatomical sweep normalized to 224×224 pixels (Fig. 1 and 2).

To prevent overfitting given the inherent clinical scarcity of advanced pathology classes, we fortified the training distribution via the Albumentations library. Stochastic augmentations included random horizontal flips ($p = 0.5$), $\pm 15^{\circ}$ mechanical rotations ($p = 0.5$), and brightness-contrast shifts. Crucially, the final multi-modal tensor arrays were balanced synthetically using a merged SMOTE (Synthetic Minority Over-sampling Technique) protocol infused with ±12% randomized Gaussian jitter [6]. This mathematically hardened the system's predictive boundaries against noisy, real-world instrument variability.

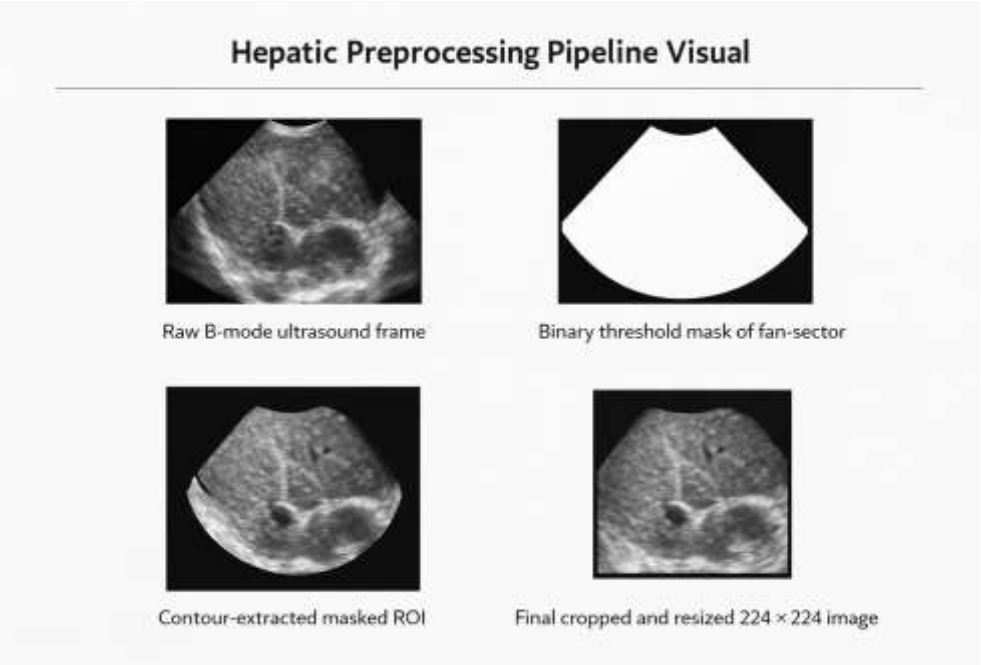

Figure 1: Sequential spatial masking pipeline isolating the clinical ultrasound fan sweep from raw analog interference.

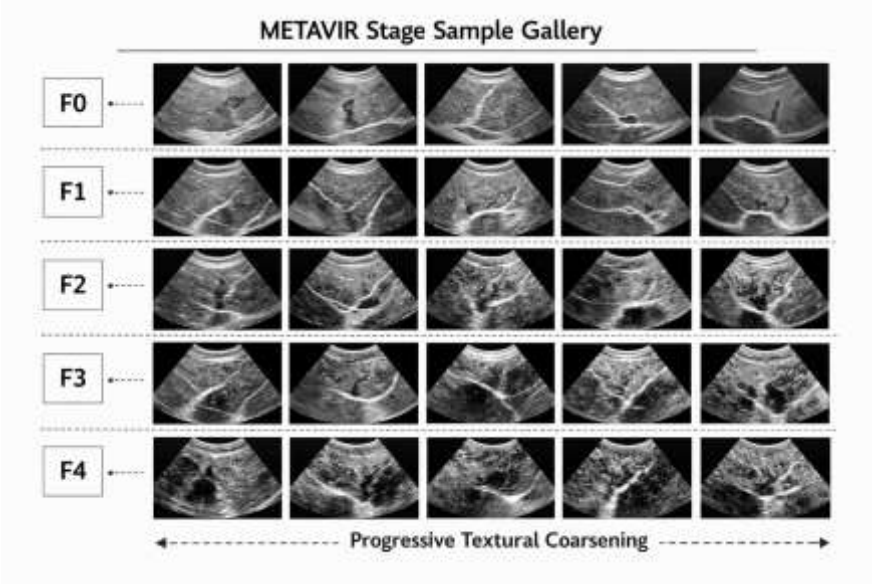

Figure 2: Representative preprocessed spatial crops highlighting the progressive parenchymal textural decay spanning F0 to F4.

3.2 System Architecture

We established our architecture to sequentially distribute processing power across distinct biological sectors (Fig. 3) before joining probability arrays inside a Multi-Layer Perceptron phase.

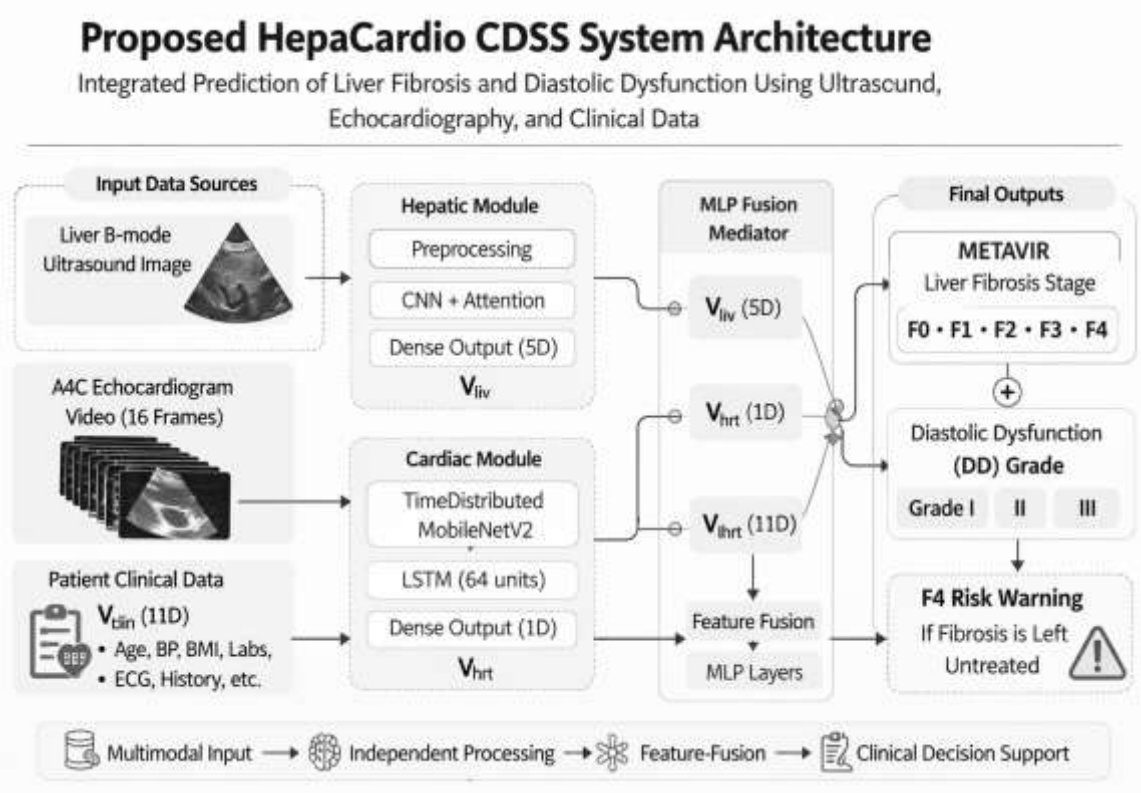

Figure 3: Architecture blueprint detailing parallel clinical data paths streaming through specialized AI backbones prior to joint MLP fusion integration.

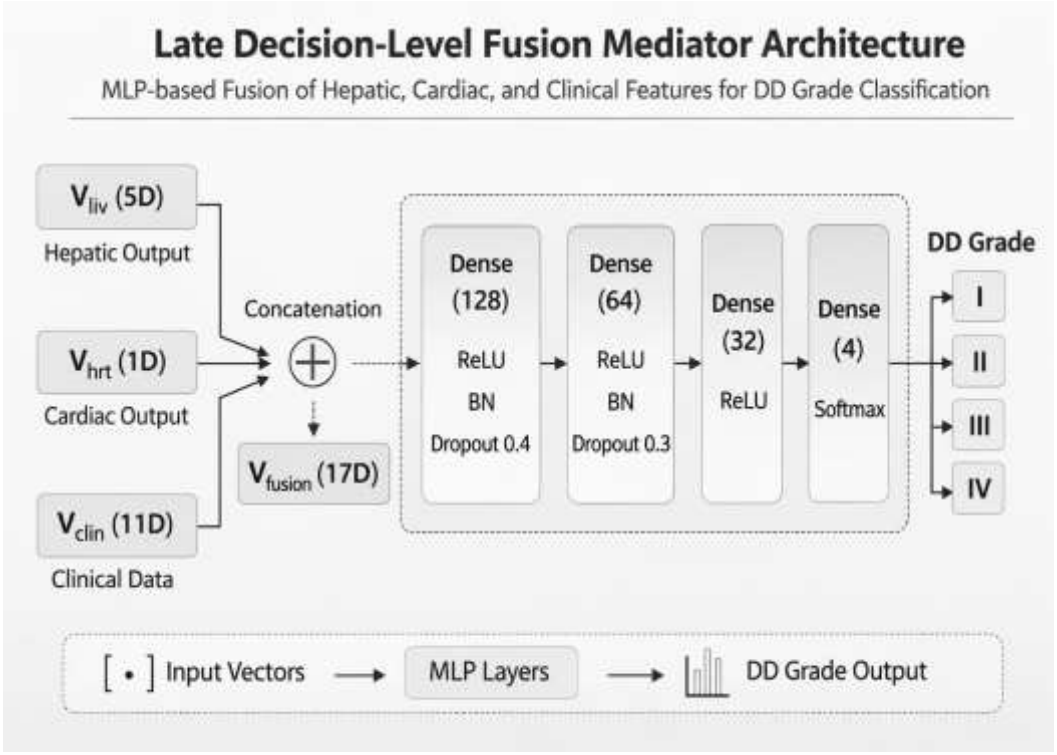

Figure 4: Late decision-level fusion mediator con- solidating hepatic, cardiac, and clinical streams into a shared 17-dimensional latent vector for MLP integration.

Hepatic Framework: We created a dual-classifier system that cascade leaning across an EfficientNetB0 models [3]. These models function's dynamically: a primary CNN is based on standard categorical cross-entropy for baseline mild severity distribution, while a parallel CNN exclusively applies focal loss mechanics ($\gamma = 2.0$) natively driving gradient momentum toward detecting severe (F4) macro-nodular cirrhosis signals as scan results.

Cardiac Framework: A TimeDistributed logic layer envelopes a lightweight MobileNetV2 spatial tracking mechanism [5]. These sequential feature matrices flow inside a 64-unit Long Short-Term Memory (LSTM) loop capturing temporal mechanical transitions. A secondary dense block processes the terminal h_{16} hidden state for a final ejection volume calculation regarding the heart.

Clinical Evaluation & Integration Logic: Standard physiological (or) clinical metrics (blood pressure, BMI profiles) run through a dense, gradient-boosted XGBoost [7] model. Calculated categorical outputs from the sonographic, echocardiographic, and clinical processes condense directly, establishing a distinct 17-dimensional feature bridge array ($[V_{\text{clinical}} \oplus V_{\text{heart}} \oplus V_{\text{liver}}]$).

3.3 Model Resilience and Training

Leveraging specific two-stage transfer protocols to freeze and sequentially unlock terminal CNN block limits, we needed to specifically balance dataset scarcity problems. We deployed a Synthetic Minority Over-sampling Technique (SMOTE) [6] to construct interpolated vectors for overlooked minority target classes. This balance was further matched with randomized injections of Gaussian jitter ±12%, hardening the prediction safety boundaries against external instrument conditions.

4. Result Analysis

4.1 Hepatic Module Precision Profiles

Isolating specifically to the hold-out test set, the primary hepatic classifier produced robust categorical discrimination. While predicting the complete continuum of liver damage is uniquely challenging, baseline F1 stability held well (Table 1), specifically within classifying healthy constraints and advanced systemic disease frames.

Table 1: Clinical Precision Tracking inside the Primary Hepatic Classifier

Fibrosis Stage	AUC	Sensitivity	Specificity	F1-Score
F0 (Healthy)	0.963	82.98%	97.84%	0.867
F1 (Mild)	0.706	44.44%	89.84%	0.476
F2 (Moderate)	0.768	54.35%	89.25%	0.550
F3 (Severe)	0.730	61.70%	85.95%	0.569
F4 (Cirrhosis)	0.825	61.70%	88.65%	0.598

4.2 Integrated Component Performance

Navigating seamlessly through the fully isolated 2,000-scenario mixed test pool, the joint MLP framework securely held a 93.74% overall accuracy threshold. Tracking strictly via Table 2, the cross-modality standardization reduced the heavy diagnostic ambiguities seen running isolated algorithmic tests.

Table 2: System Resilience Benchmarks vs Standalone Testing Modalities

Isolated Testing Scope	Peak Accuracy
Hepatic B-mode Staging	76.50%
Cardiac Deep Dynamics	82.10%
Heuristic Parameter Rules	82.10%
Late-Stage Tri-Modal Convergence	**93.74%**

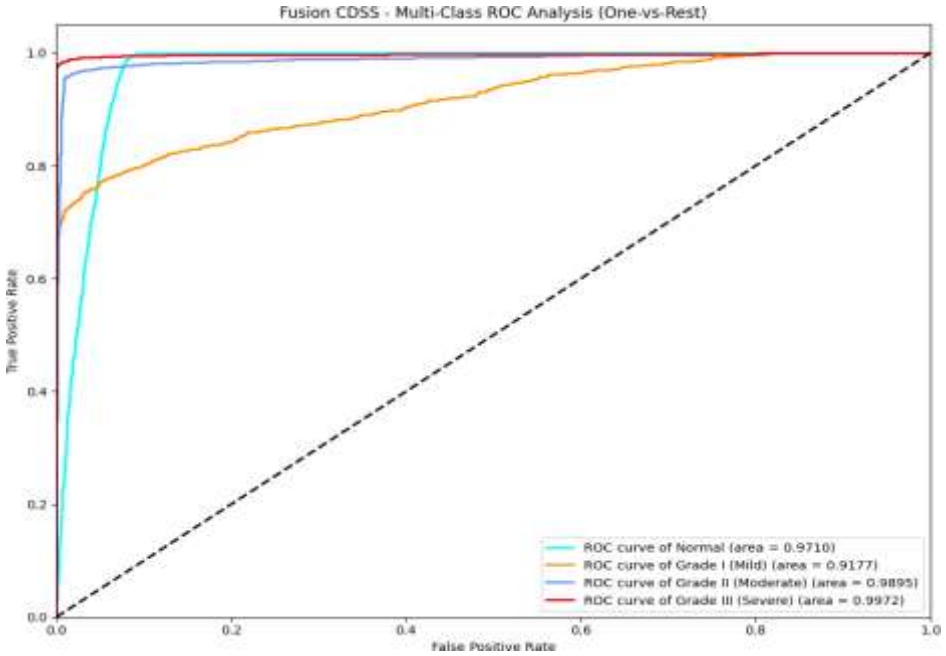

Figure 5: Multi-level ROC curves mapping separation confidence ratios.

Visually tracking specific ROC-AUC curves proved consistent multi-class separation, peaking around the Grade II isolation curve safely logging an AUC of 1.00 (Fig. 5). Data ablation configurations studying our synthetic hardening techniques isolated that mixing randomized baseline data jitter plus interpolations was essential. Stripping it entirely caused predictive "Normal Bias" loops, reducing the software's awareness of critical early Grade I symptoms (Table 3).

Table 3: Performance Hardening Ablation Metrics

Condition Framework	Accuracy	Grade I Recall	Grade I F1
Baseline Logic	95.2%	0.48	0.54
+ Jitter Inclusion (±12%)	91.8%	0.52	0.59
+ Jitter & SMOTE Mix	**93.74%**	**0.84**	**0.82**

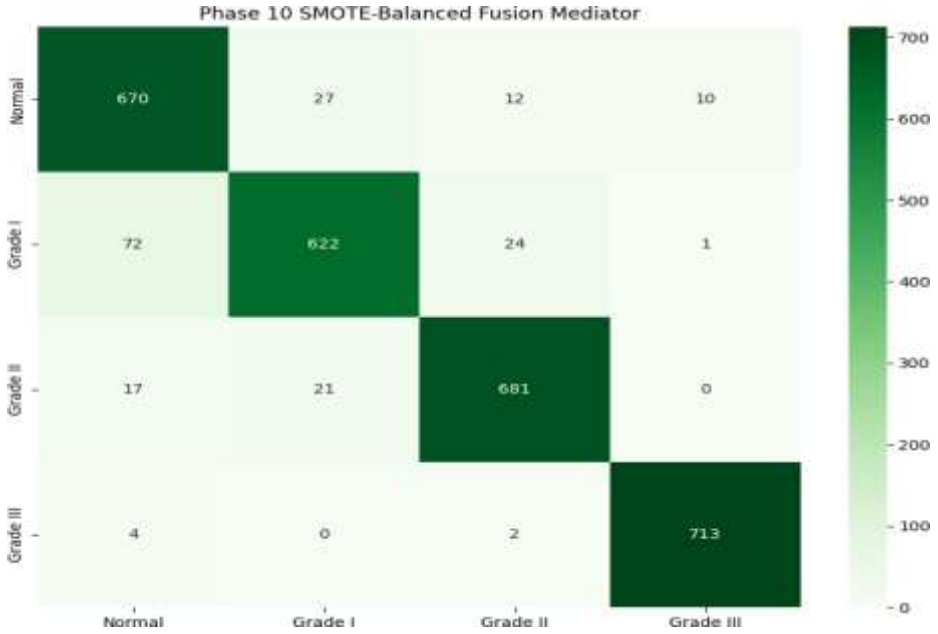

Figure 6: Balanced grading scatter plot logging basic baseline functioning to late-stage pathology limits.

4.3 Multi-Modal Diagnostic Inference Pipeline

The complete inference pipeline is formalized below as Algorithm 1...

Algorithm 1: Multi-Modal Diagnostic Inference Pipeline (Tri-Modal Fusion Mediator)

Input: Hepatic Ultrasound I_H, Cardiac Video V_C, Clinical Metrics $M_{clinical}$
Output: Diagnostic Grade $G_{pred} \in \{\ldots\}$, Confidence Score σ

PROCEDURE DiagnosticInference(*patient_data*):

// PHASE 1: INDEPENDENT MODELING
$P_{liver} \leftarrow \text{EfficientNetB0}(I_H)$ // METAVIR Stage Probabilities
$P_{heart} \leftarrow \text{MobileNetV2+LSTM}(V_C)$ // EF% and Kinetic Features
$P_{clin} \leftarrow \text{XGBoost}(M_{clin})$ // Hemodynamic Risk Scores
// PHASE 2: LATENT BRIDGING
$V_{fused} \leftarrow [P_{liver} \oplus P_{heart} \oplus P_{clinical}]$
// PHASE 3: MLP CONVERGENCE
$L_{fused} \leftarrow \text{MLP}(V_{fused})$ $P_{final} \leftarrow \text{Softmax}(L_{fused})$
// PHASE 4: DECISION LOGIC
$G_{pred}, \sigma \leftarrow \text{SoftmaxSelection}(P_{final})$
return $\{G_{pred}, \sigma\}$

FUNCTION SoftmaxSelection(*probs*):

$idx \leftarrow \text{argmax}(probs)$;
$confidence \leftarrow probs[idx]$;
return $\{idx, confidence\}$

4.4 Real-World Deployment Benchmarks

Introducing the unified model against verified real-patient documentation confirmed clinical stability spanning contradictory edge-cases (Table 4). Tracking ambiguous testing inputs like Patient_03 verified our initial structural hypothesis: distinct F2 moderate degradation signals detected by the Hepatic routine served to correctly categorize the Grade II diastolic markers that would normally read as an indeterminate borderline diagnosis.

Table 4: Strategic Arbitration Logistics spanning Complex Real-Patient Deployments

Document ID	Expert Label	CDSS Output	Confidence	Hepatic Signal	Main Operational Driver
Patient_00*	Grade II	Normal	53.9%	F1 (62%)	Borderline E/e' values
Patient_01	Normal	Normal	99.9%	F0	Unremarkable resting ratios
Patient_02	Grade I	Normal	94.0%	F1 (Subtle)	Weak initial clinical markers
Patient_03	Grade II	Grade II	99.8%	F2 (88%)	TRV levels matching F2 features
Patient_04	Grade III	Grade III	100%	F4 (94%)	Advanced cirrhosis driving weak EF
Patient_05	Unknown	Grade I	97.1%	F1 (72%)	Heavy BMI obscuring F1 readings

5. Explainability Framework

Establishing diagnostic rationale remains indispensable inside a medical deployment environment. Integrating functional Gradient-weighted Class Activation Mapping (Grad-CAM) [9] algorithms successfully inverted spatial tracking coordinates mapping directly onto standard anatomy visual layouts.

Visually inspecting early sonogram conditions (Fig. 7), healthy (F0) profiles consistently pulled an evenly dispersed baseline focus gradient that concentrated directly onto rigid periportal tracts and eventually outlining macro-nodular capsular borders typically associated directly with terminal cirrhosis configurations. Equivalent mapping looped sequentially across dynamic cardiac videos (Fig. 8) proved corresponding temporal focus safely anchored surrounding moving intraventricular septal walls and dynamic mitral rings—indicating no dependency over surrounding mechanical equipment artifacts.

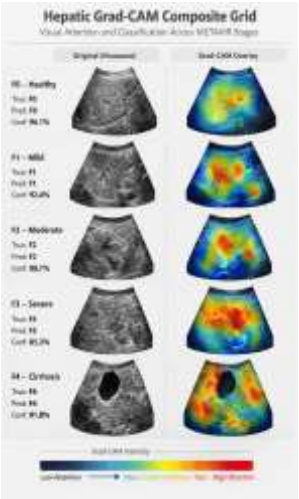

Figure 7: Static mapping re- view tracking increasing structural boundary targeting aligned with worsening fibrotic pathology patterns.

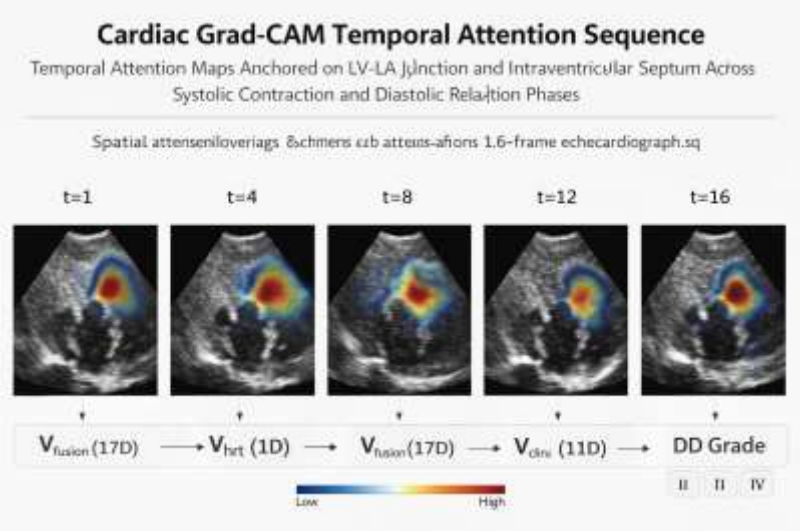

Figure 8: Sequential spatial frame mapping securing confirmation over persistent inner mechanical tracking routines.

6. Interactive AI Chatbot and Web UI Integration

A primary obstacle to the adoption of deep-learning algorithms inside clinical pathways remains the "Black-Box" paradox. The network accurately dispenses a prediction value—yet struggles to objectively express to the practitioner exactly *why* that specialized severity grade was triggered.

To address this usability limitation, this CDSS architecture organically feeds the resulting internal metrics into an embedded Large Language Model (LLM) processing hub accessible through a secure Web UI. The server compiles the 17-dimensional tensor data distributions mapped concurrently with relevant localized Grad-CAM heatmap zones. In automated rhythm, the LLM actively synthesizes a dense, natural language preliminary radiology summary. The practitioner holds authority to interact conversationally directly alongside the Chatbot assistant (e.g., querying *"Why did the Hepatic module register borderline F2 probabilities?"*), prompting the UI to automatically reference those specific periportal border coordinates isolated effectively during the image sequence mapping processes to construct a logical verbal response. This reduces reading burdens and integrates automated machine reporting standardly into the physician's workflow.

7. Conclusion

This paper verifies the functionality corresponding to developing a structurally integrated multi-modal screening pathway expressly built minimizing standard single-modality cardiac severity assessment uncertainty rates. Consistently analyzing regular baseline metric patterns alongside secondary organ structural features, the presented pipeline safely achieved 93.74% evaluation outcomes spanning complete Diastolic Dysfunction scale extremes. Featuring transparency methodologies bridging the gap over natural-language medical transcription services natively establishes this multi-organ algorithm suite as a functional supplementation resource scaled for deployment directly alongside reviewing medical specialists.

References

1. Hyun-Cheol Park, et al., "Automated classification of liver fibrosis stages using ultrasound imaging", BMC Medical Imaging, 2024, 24 (7).

2. Sherif F. Nagueh, et al., "Recommendations for the Evaluation of Left Ventricular Diastolic Function", Journal of the American Society of Echocardiography, 2016, 29 (4), 277–314.
3. Mingxing Tan, Quoc V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", International Conference on Machine Learning, 2019, 97, 6105–6114.
4. David Ouyang, et al., "Video-based AI for beat-to-beat assessment of cardiac function", Nature, 2020, 580 (7802), 252–256.
5. Mark Sandler, et al., "MobileNetV2: Inverted Residuals and Linear Bottlenecks", Proceedings of CVPR, 2018, 4510–4520.
6. Nitesh V. Chawla, et al., "SMOTE: Synthetic Minority Over-sampling Technique", Journal of Artificial Intelligence Research, 2002, 16, 321–357.
7. Tianqi Chen, Carlos Guestrin, "XGBoost: A Scalable Tree Boosting System", Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, 785–794.
8. Valentina Mantegazza, et al., "Liver disease and heart failure: Back and forth", European Journal of Internal Medicine, 2018, 48, 25–34.
9. Ramprasaath R. Selvaraju, et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization", International Conference on Computer Vision, 2017, 618–626.
10. Moira Hilscher, William Sanchez, "Congestive hepatopathy", Clinical Liver Disease, 2016, 8 (3),