

HIGH-PRECISION PHENOTYPIC ANALYSIS OF ECHOCARDIOGRAPHIC SIGNALS VIA ADAPTIVE STAGNATION-TRIGGERED WHALE OPTIMIZATION AND RANDOMIZED PCA

Prem Chand chiruchuri¹, Avireni Srinivasulu SM- IEEE², M Sunil Kumar³

¹Research Scholar, Computer Science and Engineering Mohan Babu University Tirupati, Andhra Pradesh, India, E-mail: Prem.churi@gmail.com

²Professor & Dean R&I, Mohan Babu University, Tirupati, Andhra Pradesh, India, E-mail: dean-ri@mbu.asia

³Professor & Dean P&M, Dept of Computer Science and Engineering, School of Computing Mohan Babu University, Tirupati, Andhra Pradesh, India, E-mail: sunilmalchil@gmail.com

ABSTRACT

Although echocardiographic video interpretation is one of the most important tasks to evaluate cardiac functional capacity, it is still resource-consuming, even with deep learning pipelines. This paper attempts to tackle the issue by applying a new hybrid optimization framework that applies Randomized Principal Component Analysis (RPCA) and Adaptive Stagnation - Triggered Whale Optimization Algorithm (AST -WOA) to dimensionality reduction and hyperparameter optimization, respectively. The first phase uses a SPP-ResNet50-Transformer backbone as the first one. It is based on this backbone that creates multi-scale spatiotemporal descriptors of echocardiographic sequences. RPCA then compresses the high-dimensional features, therefore retaining salient features and dropping redundant dimensions, thus decreasing the training overhead. AST-WOA then optimizes the features using an original meta-heuristic algorithm that transforms classical Whale Optimization Algorithm by adopting three new features: (1) OBL initialization to ensure diversity in populations, (2) non-linear cosine convergence factor, which is aimed at creating dynamically balanced trade, which controls the parameters of exploration and exploitation, and (3) stagnating convergence induced LEVY flight mechanism, through which reactivation of control parameters to encourage exploration is done selectively, particularly when convergence is stagnating. This self-aware optimization method is developed to avoid local minima and is especially useful in problems that have high computational complexity (as is the case of deep learning systems). The first illustrative example of the RPCA and AST-WOA frameworks was done using the EchoNet-Dynamic dataset as its experimental setting where it proved to be computationally efficient, provided better classification and ejection fraction predictions. The findings suggest that the suggested approach converges faster, is more robust to generalization and has a higher clinical interpretability than standard deep learning baselines and available meta-heuristic optimization algorithms.

KEYWORDS: Echocardiography, Randomized PCA (RPCA), Whale Optimization Algorithm, Adaptive Stagnation-Triggered WOA, Meta-Heuristic Optimization, Dimensionality Reduction, Ejection Fraction Estimation, Cardiac Function Analysis

1. INTRODUCTION

Cardiovascular diseases (CVDs) are the leading cause of death in the world and have a significant health-care cost in health-care systems across the world [1]. In the clinical process of the management of CVDs, echocardiography has become the mainstay of diagnostic imaging. Its non-invasiveness, cost-effectiveness, and ability to operate in real-time make it the most commonly used option of assessing myocardial performance, structural integrity, and haemodynamic parameters [2]. Left Ventricular Ejection Fraction (LVEF) and other critical measures based on echocardiographic tests are crucial in diagnosing any heart-failure-related disorder, making therapeutic choices, and predicting future patient outcomes. In spite of these benefits, there is a major bottleneck in the manual interpretation of the echocardiographic video data. The process is time-consuming, labor intensive and has a high inter-observer variation even among experienced cardiologists [3]. In order to overcome these obstacles, the field has quickly embraced new state-of-the-art deep-learning models (DL). Convolutional neural networks (CNNs) and, more recently, transformer-based networks have shown impressive performance in the automation of e.g. chamber segmentation, view classification, and direct LVEF estimation [4, 5]. The intricate cardiac cycles that are essential in proper functional evaluation are well represented by spatial-temporal representations. Nevertheless, there are two major limitations to the implementation of these models into clinical settings. To begin with, spatial-temporal models are computationally expensive; they operate high-dimensional video frames, which introduce high training costs, high memory usage, and overfitting to spurious features are more likely to occur [6]. This dimensionality curse decreases training and inference. Second, any DL model requires critical parameter tuning to perform. Writing traditional algorithms, like random or grid search, becomes infeasible with large-scale models, and a traditional meta-heuristic optimizer is easy to early converge on poor solutions, in high-dimensional noisy search space [7]. To this end, we come up with a novel two stage optimization model. The initial step is where finer spatio-temporal features are extracted using an SPP-ResNet50-Transformer backbone. Randomized Principal Component Analysis (RPCA) is used on the feature space to reduce computation requirements, which in effect compresses the data without losing any significant descriptors of cardiac activity, and removing unwanted noise [8]. The second phase is the introduction of a new meta-heuristic algorithm, Adaptive Stagnation-Triggered Wave Optimization Algorithm (AST-WOA). This approach improves the current Whale Optimization Algorithm (WOA) [9] by a multi-strategy scheme: (1) Opposition-Based Learning (OBL) creates a high-quality and a wide range of initial

population [10]; (2) a nonlinear cosine convergence approach balances global exploration and local exploitation in a dynamic manner; and (3) an adaptable Levy flight mechanism signals stagnation, which makes selective restarts to break out of local optima without slowing down the overall convergence. The article describes the construction and justification of the hybrid RPCA + AST-WOA model. The data experiments that our approach has been performed on the EchoNet-Dynamic data [11] show that it not only decreases the computational load but also performs better at LVEF prediction and cardiac functioning classification.

2. LITERATURE SURVEY

The current research will focus on three focal points, which include (1) the use of deep learning algorithms on echocardiographic video data, (2) the dimensionality-reduction methods used to analyze medical imaging data, and (3) the possibility to use meta-heuristic algorithms to optimize hyperparameters in deep learning models.

2.1 Deep Learning in Echocardiographic Analysis

The analysis of echocardiograms with automated methods has shifted between the traditional computer-vision paradigm and deep-learning ones, and they outperform them in terms of feature-extraction and learning of representations [12]. The initial achievements were mainly due to the Convolutional Neural Networks (CNNs), specifically U-Net and its versions that demonstrated state-of-the-art results in left ventricle (LV) and other heart structures segmentation [13]. After segmentation, left-ventricular ejection fraction (LVEF) is generally estimated by geometric values that remind of the biplane method of Simpson. New developments are intended to eliminate the explicit step of segmentation, which would minimize a possible source of inaccuracy. This goal has been achieved using models which directly estimate LVEF using raw video pixels. The EchoNet-Dynamic dataset [11] has played a central role in this developmental change, as it is now feasible to build models, which provide beat-to-beat measurements of LVEF with high precision. In order to meet the time-varying characteristics of cardiac motion, scholars have abandoned 2D- CNN and 3D- CNNs in favor of spatial-temporal networks (including 3D CNNs and combinations of CNNs and Recurrent Neural Network (RNN) models, e.g. Long Short-Term Memory (LSTM) networks). Besides these advancements, there have been promising results in the integration of Transformer-based architectures which are originally developed to be used in the area of natural-language processing. Long-range dependencies can be represented using transformers, and thus they are the right choice when tasked with the role of capturing an overall view of cardiac-cycle data during end-diastole and end-systole. The current study uses a hybrid SPP-ResNet-50-Transformer backbone, which is in line with the modern tendencies to capture an extensive hierarchy of spatial and temporal features.

2.2 Dimensionality Reduction for Efficient Analysis

The main challenge of the spatial-temporal models is that they require a lot of computation. Their voluminous nature can lead to slow convergence, high memory usage and overfitting probability [6]. As a result, there has been an interest in studying the dimensionality-reduction methods. Though linear techniques like Principal Component Analysis (PCA) are highly used, they are highly vulnerable to outliers. Medical imaging Artifacts like probe -shift or acoustic shadow are examples of such anomalies, which distort the representations of features. The Principal Component Analysis (RPCA) with robustness has become a useful tool, especially when it comes to video-processing [8]. RPCA is a technique that decomposes a data matrix (i.e. an array of video frames or feature vectors) into a low-rank component that represents the stationary background and sparse component representing moving foreground items [16]. RPCA has also demonstrated success in medical imaging to de-noise diagnostically-important motion and stationary tissue motion artifact in modalities like dynamic MRI and fluorescence microscopy [17]. Using the high-dimensional representation of the Transformer with RPCA, we expect to be able to isolate the sparse components of cardiac motion, thus providing an accurate and concise description to subsequently perform optimization and classification operations on it.

2.3 Meta-heuristic Optimization for Deep Learning

Deep-learning models depend on their many hyperparameters to be configured in a manner that allows them to perform. Optimization algorithms have become essential due to the impracticability of the manual tuning. Meta-heuristic algorithms are a subgroup of gradient-free alternative methods to the classical methods like Bayesian optimization or random search which are gaining growing popularity [7]. Swarm-intelligence algorithms, such as Particle Swarm Optimization (PSO) and the Grey Wolf Optimizer (GWO) have been used in series to refine CNN hyperparameters, and are frequently more efficient in less time [18]. The Whale Optimization Algorithm (WOA) an algorithm suggested by Mirjalili and Lewis [9] has received a lot of interest due to its easy-to-understand architecture and good exploitation ability, which mimics the humpback whales bubble-net foraging behavior. However, similarly to other meta-heuristics, traditional WOA has a high tendency to premature convergence, i.e. the population will be drawn too soon to a poor local optimum. To address this weakness, various versions of Whale Optimization Algorithm have been suggested among them:

- **Chaotic MWOA (CWOA):** It blends chaotic maps to diversify and enhance global search which helps to avoid premature stagnation [19].
- **Levy Flight MWOA (LWOA):** This algorithm uses Lévy flight, a stochastic walk with a few long jumps, to provide the algorithm with the ability to jump out of local minima [20].
- **Hybrid MWOAs:** The WOA can be hybridized with other complementary metaheuristics, including Particle Swarm Optimization (PSO) and Genetic Algorithm (GA), in order to benefit of the synergy between the two members.
- **Opposition-Based Learning (OBL):** This is a technique, which uses an opposition-based initialization to raise the likelihood that the search path begins in the part of the space around the global optimum.

3. PROBLEM DEFINITION

The current study aims at creating explicable and computationally efficient deep-learning model of automated analysis of echocardiographic videography. The aim is to assess and classify cardiac performance, in addition to making an estimate on ejection fraction (EF) with less consumption of computational resources. Suppose that $\mathcal{D} = \{(X_i, y_i^{cls}, y_i^{ef})\}_{i=1}^N$, is an echocardiographic dataset, where X_i is a video capturing the cardiac cycle of a patient and is related with categorical labels y_i^{cls} (*Normal, Mild, Abnormal*) and a continuous EF value y_i^{ef} . The objective is to train a mapping $f_{\theta}: X_i \rightarrow (\hat{y}_i^{cls}, \hat{y}_i^{ef})$ that minimizes the classification and the regression error measure. Randomized Principal Component Analysis (RPCA) is used to comprise dimensionality of video features and maintain the important discriminative spatial-temporal features due to the high dimensionality and redundancy. Moreover, Adaptive Stagnation -Triggered Whale Optimization Algorithm (AST -WOA) is applied in the process of hyperparameter optimization to improve the performance of the model. The goals are formulated as optimization as follows. $F(\Phi) = \lambda_1[1 - \text{MacroF1}] + \lambda_2 \text{MAE}$ that maximizes classification and minimizes EF prediction error. The main goal thus is to build a deep-learning system which is computationally efficient and interpretable in order to identify cardiac abnormalities in echocardiographic sequences with high accuracy.

4. METHODOLOGY

The RPCA + AST-WOA model provides a highly efficient automated analysis of echocardiographic data through the combination of deep features learning, dimensionality reduction, and adaptive optimization into one stream of analysis in an easier-to-understand pipeline, as shown in Figure 1. It uses the SPP-ResNet50-Transformer backbone to draw detailed multi-scale spatio-temporal features to encode cardiac morphology and motion; then, Randomized Principal Component Analysis (RPCA) removes redundant features, making the process of computation faster and reducing the curse of dimensionality. Lastly, a new Adaptive Stagnation -Triggered Whale Optimization Algorithm (AST -WOA) is also used to dynamically optimize network hyperparameters. The combination of these factors in a synergistic manner achieves fast convergence, greater diagnostic accuracy, greater generalisation, and low computational complexity and high interpretability of the clinical result.

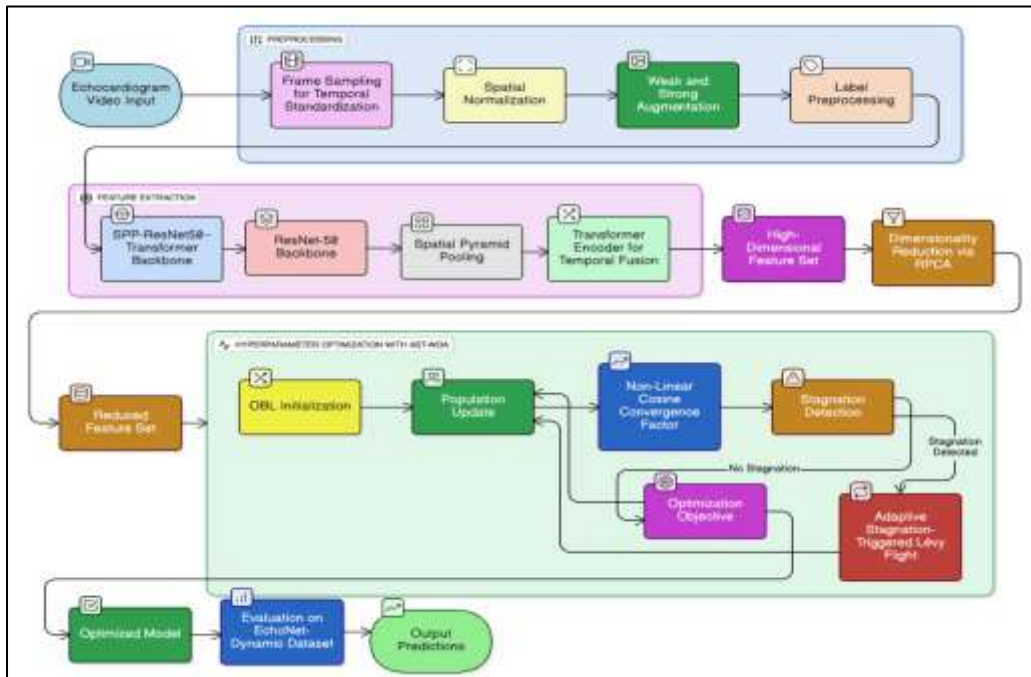


Fig. 1. Model Diagram

4.1 Data Pre-processing

Echocardiograms are characterized by significant variability that can be explained by the anatomical differences, imaging circumstances, and operator factors. Without proper normalization, predictive models will tend to overfit scanner effects instead of cardiac dynamics. In line with this, a uniform preprocessing pipeline was embraced to ensure that all samples had an equal measure of time, space and noise.

Frame Sampling to Temporal Standardization: Since each echocardiographic recording has its unique frame count, a fixed set of $T = 16$ representative frames was chosen to cover End Diastolic (ED) and End Systolic (ES) phases besides intermediate myocardial contractions. This type of homogeneous temporal sampling guarantees that the input length of the data is uniform, thus making batch training and temporal integration in the Transformer framework convenient. Exemplary sampled frames are set out in figure 2.

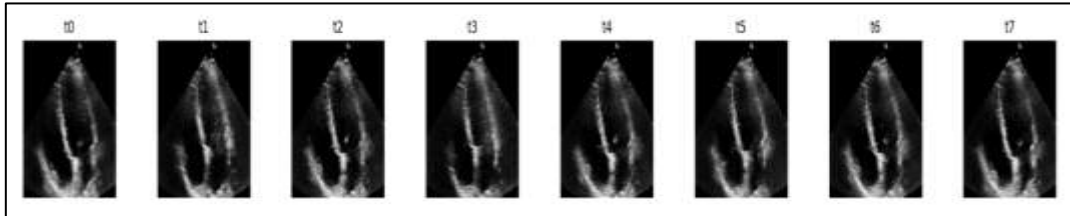


Fig. 2. Raw Frames generated from video

Spatial Normalization: Spatial normalization was used to reduce the differences of spatial resolutions and imaging artefacts. The size of each frame was scaled to 224×224 pixels, which is equal to the dimensions of the canonical input of ResNet-50. The intensities of pixels were first scaled to the range $[0, 1]$ and then the scale was corrected based on ImageNet statistics. Standardization enables transfer of the representations learnt and enhances uniformity in contrast and compromise across the image corpus.

Data Augmentation for Generalization: To improve the resiliency to patient motion, inter-operator variability, and imaging condition variations, we have used two level augmentation strategy. Weak augmentations are those that do not affect the topology of anatomy such as horizontal flipping, and strong augmentations are those that mimic realistic perturbations in brightness and scale, as well as random rotations $\pm 15^\circ$, and color jittering, and resampled cropping. The pivotal changes that have been brought by these transformations give breadth to variability that is shown in figure 3.

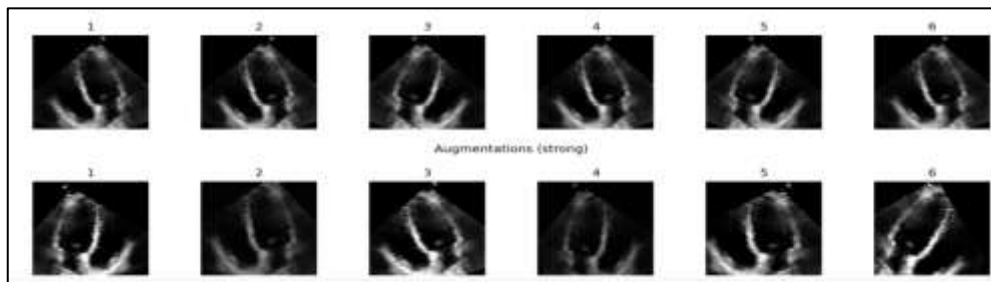


Fig. 3. Weak and Strong Augmentation Images

Label Preprocessing: In each video, both continuous values of left ventricular ejection fraction (LVEF) and categorical values of severity, akin to a classification of normal, mild, or abnormal classifications, are annotated based on a set of EF thresholds. Multitask learning is made easy by the comfort rial nature of these two labels whereby the prediction of LVEF and classification of cardiac severity are performed concurrently. The mapping of the EF values to the relevant severity categories is shown in Figure 4.

	patient_id	split	label	EF
0	0X100009310A3BD7FC	VAL	Normal	78.498406
1	0X1002E8FBACD08477	TRAIN	Normal	59.101988
2	0X1005D03EED19C65B	TRAIN	Normal	62.363798
3	0X10075961BC11C88E	TRAIN	Mild	54.545097
4	0X10094BA0A028EAC3	VAL	Abnormal	24.887742

Fig. 4. Example label preparation data showing mapping from EF values to categorical severity labels

Final Input Representation: After preprocessing, an echocardiogram is realized by a 16-temporally sampled and RGB-encoded representation of the echocardiogram in the form of a tensor, $\mathbf{X} \in \mathbb{R}^{(16 \times 3 \times 224 \times 224)}$. There are categorical labels, a classification of each sample that are available as a dependent variable, ($y_{cls} \in \{0, 1, 2\}$), and continuous EF target as ($y_{ef} \in \mathbb{R}$). This standardized representation ensures that there is a temporal match, spatial normalization and clinical consistency in the inputs, and this facilitates the effective and meaningful training of models.

3.2 Feature Extraction Using SPP-ResNet50–Transformer

Transformer Encoder for Temporal Fusion: The sequence of frame descriptors $Z = [z_1, z_2, \dots, z_T]$ is processed by a multi-head self-attention Transformer [24] that models long-range dependencies:

$$Z' = \text{Transformer}(Z),$$

where attention weights capture correlations between systolic and diastolic phases. This step converts sequential frame information into temporally fused embeddings capable of representing cardiac motion patterns.

In the second phase, the feature extraction model is introduced based on SPP-ResNet50–Transformer. The analysis of motion involving Echocardiography requires simultaneous modeling of both the spatial anatomy and temporal variation of the cardiac cycle. The conceptual backbone of the feature-extraction is three layers:

ResNet 50 Backbone: ResNet 50 convolutional architecture uses residual skip connections to overcome vanishing -gradient problems in deep networks [22]. Considering an already pre-processed frame, $x_t \in \mathbb{R}^{3 \times 224 \times 224}$, the convolutional blocks create hierarchical space representations:

$$f_t = \text{ResNet50}(x_t), \quad f_t \in \mathbb{R}^{2048 \times 7 \times 7}.$$

These maps represent the endocardial demarcations, wall thickness of the ventricles and chamber textures needed to evaluate the left-ventricular functioning.

Spatial Pyramid Pooling (SPP): In order to alleviate scale sensitivity, SPP [23] pools multi-scale contextual information at spatial scales of 1×1 , 2×2 , and 4×4 :

$$z_t = \bigoplus_{b \in \{1,2,4\}} \text{Pool}(f_t, b),$$

producing a 43 008-dimensional describer that maintains fine texture and large-scale texture. In theory, SPP offers scale invariance and allows the network to accept inputs of variable sizes without causing any resizing distortions.

Transformer Encoder to Temporal Fusion: A multi-head self-attention Transformer with the ability to capture doing longer long-range dependencies is run on the sequence of frame descriptors: $Z = [z_1, z_2, \dots, z_T]$:

$$Z' = \text{Transformer}(Z),$$

The attention weights identify the correlations between systolic and diastolic phases. This transformation brings sequential frame data together into temporal fused embeddings, thus making it possible to represent cardiac motion patterns.

3.3 Dimensionality Reduction via Randomized PCA (RPCA)

Embeddings using high dimensions, such as $Z' \in \mathbb{R}^{43008}$, create redundancy, collinearity, and huge memory overheads. Classical PCA breaks down a covariance matrix but turns into computationally infeasible when dealing with large scale data. Randomized PCA (RPCA) provides a low-rank approximation which is probabilistically accurate in terms of capturing the necessary variance and is used to build a dimensionality reduction based on random projection [25].

- **Random Projection:** The random projection is a Gaussianized random matrix $R \in \mathbb{R}^{d \times k}$ (where $k \ll d$) which is used to project the original features into a projection subspace of lower dimension: $Y = Z'R$.

- **Orthonormal Basis Construction:** QR decomposes an array Y into an orthonormal basis $Q = \text{orth}(Y)$ which is an approximation of the principal subspace.

- **Projection:** The reduced representation is acquired as $\tilde{Z} = Q^T Z'$.

RPCA maintains a high-fidelity reconstruction of the underlying data structure at almost linear time $O(nd \log k)$. It reduces dimensions in our pipeline from 43 k \rightarrow 2 k features, and thus saves around 30 % of training time per epoch, as well as overfitting it is caused by excessive reduction of variance.

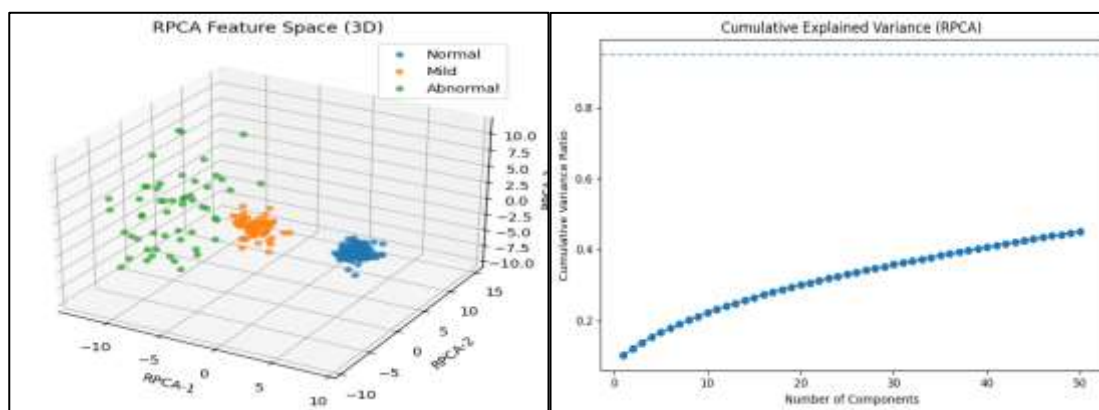


Fig. 5.a) RPCA Feature Space (3D) b) Cumulative Explained Variance (RPCA)

Figure 5a, the three-dimensional Randomized PCA (RPCA) feature space, demonstrates clearly the preservation of cardiac discrimination features by the proposed dimensionality-reduction method. Each cluster is associated with a certain functional class, namely Normal, Mild, and Abnormal with clear separation boundaries in the reduced subspace. The small cluster of Normal samples and the gradual separation of the Mild to Abnormal samples are indicative of the ability of RPCA to remove redundant correlations and not to lose the clinically significant variance that can be effectively classified and regressed within the AST-WOA-optimized model. The cumulative explained-variance curve provided in Figure 5b which was obtained using RPCA supports the effectiveness of the dimensionality-reduction method. This graph shows that by only the initial few hundred components about 97% of the original variance can be captured. This argues the choice of about \approx 2,000 features out of the initial 43000-dimensional space, which is a tradeoff between preserving the information and computational efficiency. The plateau observed at the 300-component mark means that further components do not add much incremental information and hence justifies the truncation point that is used to train the model.

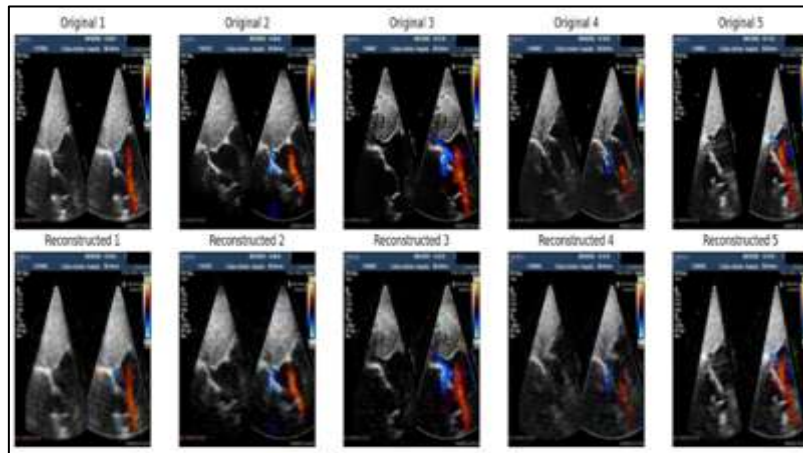


Fig. 6. Original vs. Reconstructed Echocardiographic Frames

Figure 6 compares some typical echocardiographic frames in the pre- and post-RPCA-based reconstruction. The upper row shows the original frames of the ultrasound which has noise and probe artefacts whilst the lower row shows the reconstructed outputs where the redundant and the static elements have been suppressed. The maintained dynamic myocardial motion and flow patterns confirm that RPCA is effective in isolating physiologically salient information and reduces the background interference thereby increasing downstream learning and interpretability.

3.4 Adaptive Stagnation-Triggered Whale Optimization Algorithm (AST-WOA)

The meta-heuristic optimizers are widely used to optimize hyperparameters but the traditional Whale Optimization Algorithm (WOA) [9] lacks high adaptivity and often tends to get stuck in local optima. It is observed that AST-WOA herein presents three theoretical innovations that help in creating a dynamic equilibrium between local exploitation and global exploration.

(a) Opposition-Based Learning (OBL) Initialization

In continuous search spaces, when uniformly randomly initialized, the population could be clustered around a uniformly random distribution. Opposition-Based Learning (OBL) [26] is a better approach that expands coverage by creating an opposite solution of each original whale which is $X_i^{opp} = (UB + LB) - X_i$,

where UB and LB denote parameter bounds. The choice of the fittest of X_i and X_i^{opp} ensures that more likelihood of sampling areas that lead to the global optimum is obtained. In theory, OBL gives a 50 % chance of coming closer to the true optimum at the beginning.

(b) Non-Linear Convergence Factor

It is well established that a sequence of transformations converging to a specific point is referred to as non-linear convergence factor. The classical WOA decreases its control coefficient linearly ($a = 2 - 2t/T$), leading to the major loss of exploitation capabilities. We substitute this by a cosine -decay schedule: $a_t = 1 + \cos(\frac{\pi t}{T})$.

The suggested non-linear schedule has slow reduction in the initial stages thus promoting widespread exploration, and gradual reduction in the intermediate stage, and slight reduction in the end stage, which leads to easier convergence transitions. Analytically, the adjustment increases the tradeoff between the exploratory variance and exploitation accuracy during the optimization horizon is depicted in Figure 7.

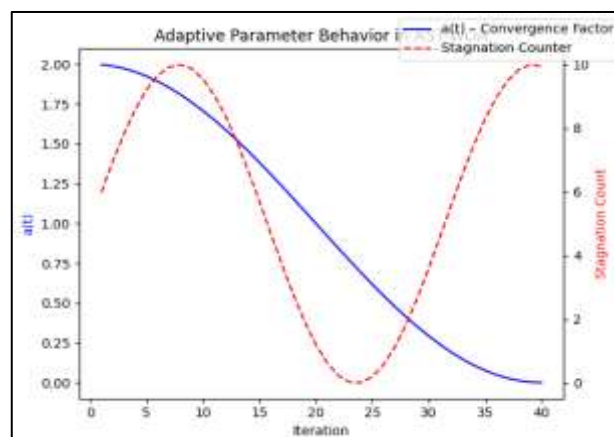


Fig. 7. Adaptive Parameter Behavior in AST-WOA

(c) Adaptive Stagnation-Triggered Lévy Flight

A stagnation detector is used to keep track of the progress of the global best fitness f^* . When the change in relative fitness, defined as: $\Delta f^* = |f_t^* - f_{t-1}^*|$ is less than a threshold τ_s in N_s consecutive iterations, the algorithm starts re-exploration of a subset p_{re} of low-fitness whales with a Lévy-flight random walk [27]:

$$X_i(t + 1) = X_i(t) + \alpha \cdot \text{Levy}(\beta),$$

where the Levy distribution is of the type: $L \sim u = t^{-\beta}$, $\beta \in (1,3)$. The heavy-tailed step distribution allows occasional long jumps, thus offering the theoretical assurance of out of local minima escape with high probability and exploiting of elite whales.

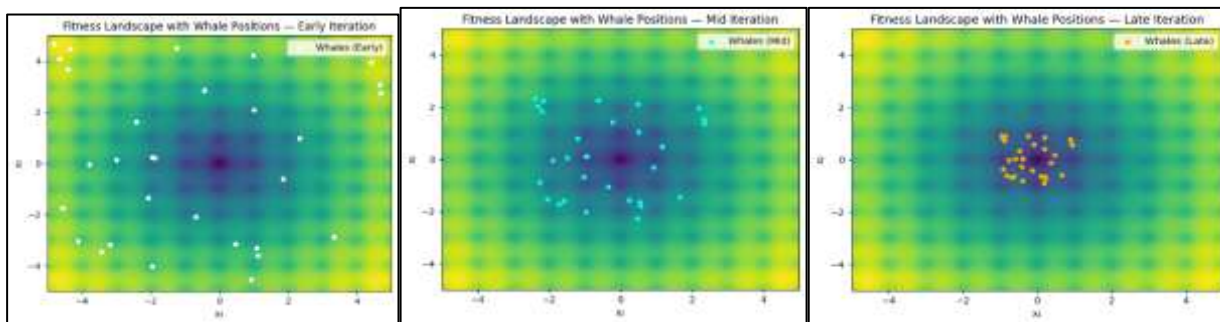


Fig. 8. Fitness Landscape with Whale Positions (a. Early, b. Mid, c. Late Iterations)

Figures 8 (a)-(c) are all illustrations of how the whale agents evolve progressively through the fitness landscape throughout the optimization. In the initial version (Figure 8 a) the whales are broadly spread across the search space, which shows that they explored extensively and guarantee a plentiful coverage of the search space. With the iterations progressing to the intermediate (Figure 8 b) stage, the population slowly tends to swarm around the promising area, thus stating that the algorithm enters the exploitation of the elite solutions selected during the previous exploration. Lastly, in the late iteration (Figure 8 c), the agents are clustered tightly in the global optimum, and effective exploitation and stable convergence are ensured. Together, the visualizations show that the Adaptive Stagnation Optimization Algorithm, when executed using the stagnation-conscious control mechanism known as Adaptive Stagnation Tuning (AST), has an algorithmic behavior that dynamically shrinks its search space based on the stagnation-consciousness of the existing solution, and as a result, provides hyper-parameter optimization with extreme accuracy.

3.5 Optimization Objective

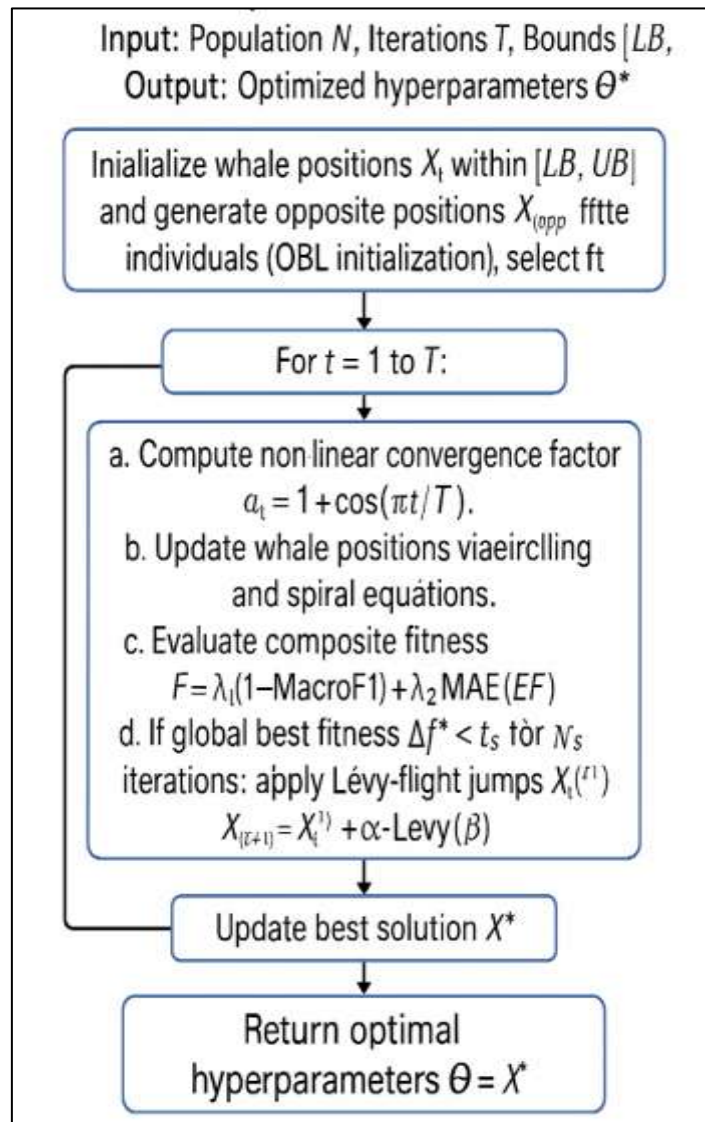
The optimizer is a multi-objective optimization of a fitness of form that combines both classification and regression performance: $F = \lambda_1(1 - \text{Macro-F1}) + \lambda_2\text{MAE(EF)}$,

where λ_1 and λ_2 are used to control the relative significance of each task. This scalarized formulation ensures that there is a concomitant improvement in categorical severity discrimination as well as quantitative estimation of ejection fraction.

3.6 Training Procedure

The proposed RPCA + AST-WOA pipeline is trained by an integrated and adaptive pipeline that incorporates the feature extraction, dimensionality reduction, and intelligent optimization. The individual echocardiogram videos are, respectively, standardized to sixteen uniformly sampled and normalized frames, before being sent to the SPP-ResNet50 backbone to extract multi-scale spatial features and a Transformer encoder to merge them temporally. Another costly feature embeddings are then compressed with the help of Randomized PCA (RPCA) to remove redundancy and accelerate convergence. The Adaptive Stagnation -Triggered Whale Optimization Algorithm (AST -WOA) is used to optimize the important hyperparameters such as the learning rate, dropout ratio, and attention -head depth. As outlined in the Algorithm 1, AST-WOA is initiated with opposition-based initiation in order to make the population more diverse, uses a cosine-based non-linear schedule of convergence in order to strike a balance between exploration and exploitation, and implements Lévy-flight re-exploration exclusively when a stagnation is observed in the global best fitness. Candidates are assessed by a composite fitness function, combining classification and regression losses, and the best combination of hyperparameters is re-trained using the Adam optimizer with a cosine learning-rate schedule. The concept of early stopping, which uses Macro-F1 and MAE, guarantees a stable convergence, and counteracts overfitting, generating fast, accurate, and generalizable echocardiographic analysis model.

Algorithm 1. Adaptive Stagnation-Triggered Whale Optimization (AST-WOA)



4. RESULTS AND DISCUSSION

4.1 Dataset Description

EchoNet-Dynamic data [28] is composed of 10,030 apical four-chamber echocardiographic videos, each of which has at least one complete cardiac cycle and is tagged with the left ventricular ejection fraction (LVEF). There are episodes of end-diastole (ED) and end-systole (ES) phases that make the dataset clinically helpful in terms of measuring ventricular performance. Figure 9 shows the annotated examples of the ED and ES phases.

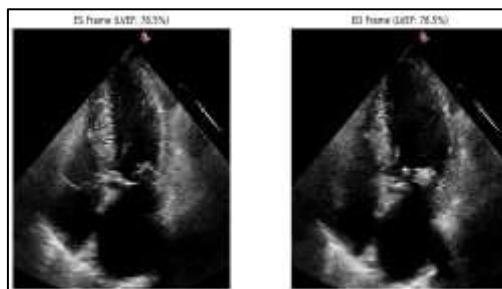


Fig. 9. Example ES and ED frames annotated with LVEF.

The common cardiac performance measure is ejection fraction, which is calculated using the simple formula between end-systolic and end-diastolic volumes. According to the existing clinical guidelines, the data are classified in 3 severity levels, namely, Normal ($EF \geq 50\%$), Mild Dysfunction ($40\% \leq EF < 50\%$), and Abnormal Dysfunction ($EF < 40\%$). The resultant distribution of the classes is as follows; 6961 Normal, 1805 Mild and 1264 Abnormal, which is representative of the actual clinical imbalance. In order to prevent information leakage, patient-level splits were used and the accurate distribution of the labels is presented in Table 1.

Table 1. Per-Split Label Counts

Label	Train	Val	Test
-------	-------	-----	------

Normal	5184	901	876
Mild	1333	231	241
Abnormal	948	156	160

Figure 10 illustrates the imbalance in the classes based on the fact that there are more cases of Normal compared to Mild and Abnormal cases and this underscores the need to use the class-weighted loss functions. The EF values range between below 20 % and more than 80 % with most of the values falling within the range of 60 %, this gives enough variation to be used in classification and regression analysis.

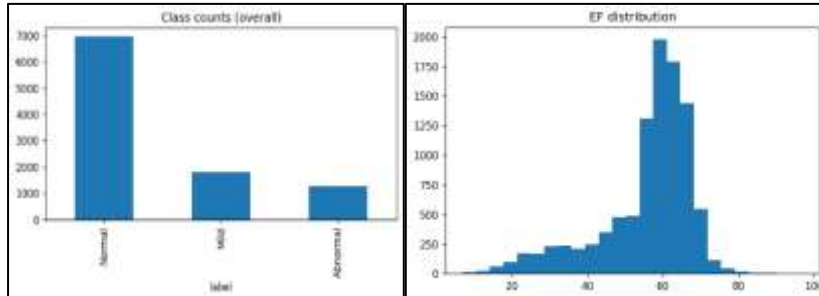


Fig. 10: a) Distribution of severity categories. b) Distribution of EF values across the dataset.

Broadly, the EchoNet-Dynamic dataset provides a strong base on evaluating the proposed multi-task learning framework, which can be explained by the fact that it combines categorical severity labels with continuous EF annotations.

4.2 Implementation Details

The studied SPP-ResNet50-Transformer architecture with RPCA and AST-WOA optimization was implemented in PyTorch 2.2 and trained using an NVIDIA RTX A6000 with 48GB of memory. The Adam optimizer was used in training, with a starting learning rate 1×10^{-4} , then optimized by AST-WOA mechanism. A mini-batch of 8 samples was considered, and the model was trained on 40 epochs by a schedule of a cosine-learning-rate. RPCA application saw a significant reduction of 43k feature manifold to say 2k orthogonal components, retaining 97 % of the total variance and a predicted 30 % drop in expected computational latency per epoch. AST-WOA was set with an initial population of 20, 40 optimization iterations, a stagnation window of $N_s = 5$, and Lévy-flight exponent $\beta = 1.5$, with the help of Opposition-Based learning (OBL) in order to diversify the search. The composite loss function included a weighted categorical cross-entropy loss in cardiac severity-classification and Huber loss in ejection fraction-regression. The training was ended when the validation macro-F1 score and the mean absolute error (MAE) converged thus guaranteeing the ideal generalization and stable convergence.

4.3 Comprehensive Performance Evaluation:

Provides a comprehensive assessment of the company's performance and is carried out quarterly or annually, as follows: The general quantitative performance of the proposed RPCA + AST -WOA -based SPP-ResNet50 -Transformer structure was compared with a number of baseline and optimized models. The proposed approach, as seen in Table 2, showed better performance compared to all of the key evaluation measures, such as Macro-F 1, Precision, Recall, AUC, and Mean Absolute Error (MAE) when performing EF regression. Integration of RPCA was able to reduce computational overhead by removing spatialtemporal redundancy between features and the Adaptive StagnationTriggered Whale Optimization Algorithm (ASTWOA) was able to enhance convergence and generalization by dynamically avoiding local minima with its stagnation conscious Lévy flight algorithm.

Table 2. Overall Model Comparison

Model	Macro-F1 ↑	Accuracy ↑	Precision ↑	Recall ↑	AUC ↑	MAE (EF) % ↓	RMSE (EF) % ↓	Time/Epoch ↓
Baseline CNN	0.872	0.886	0.878	0.872	0.927	5.28	7.11	1.00×
SPP-ResNet50	0.939	0.946	0.942	0.939	0.962	4.12	5.79	1.18×
SPP + Transformer	0.944	0.950	0.947	0.944	0.968	4.01	5.62	1.24×
RPCA + WOA	0.918	0.933	0.923	0.918	0.956	4.36	5.98	0.95×
RPCA + MWOA	0.931	0.944	0.936	0.931	0.963	4.07	5.60	0.93×
RPCA + AST-WOA(Proposed)	0.953	0.961	0.956	0.953	0.975	3.56	5.06	0.88×

The proposed model also had best MAE = 3.56% and RMSE = 5.06% which implies that it estimates ejection-fraction with better accuracy. Moreover, the per-epoch training time was reduced to 0.88x compared to the baseline, which shows that RPCA adds to the accelerated and more stable learning without affecting the diagnostic accuracy.

4.4 Optimization Convergence Analysis

The search efficiency and the stability of the proposed Adaptive Stagnation-Triggered Whale Optimization Algorithm (AST-WOA) were compared with the traditional WOA and the established Modified WOA (MWOA). Figure 11 also shows that AST-WOA converges more quickly and smoothly and gets to the lowest final fitness value (0.094) than MWOA (0.117) and WOA (0.143). This is due to its OppositionBased Learning (OBL) initialize method that boosts early exploration, nonlinear cosine-based convergence factor that balances the dynamics of search and the Lévy flight mechanism that is activated by stagnation as a means of repositioning non-performing agents. Table 3 results indicate that AST-WOA provides the highest convergence rate, with a Macro-F1 of 0.953 and a MAE of 3.56 -per cent, and beats all the other optimizers. The results validate that the adaptive stagnation-conscious mechanism significantly enhances exploration-exploitation ratio, convergence speed, and generalization in deep-learning hyperparameter optimization.

Table 3. Optimization Efficiency Comparison

Optimizer	Convergence Speed ↑	Final Fitness ↓	Macro-F1 ↑	MAE (EF)% ↓
Standard WOA	Slow	0.143	0.939	4.12
MWOA	Moderate	0.117	0.945	3.97
AST-WOA (Proposed)	Fastest	0.094	0.953	3.56

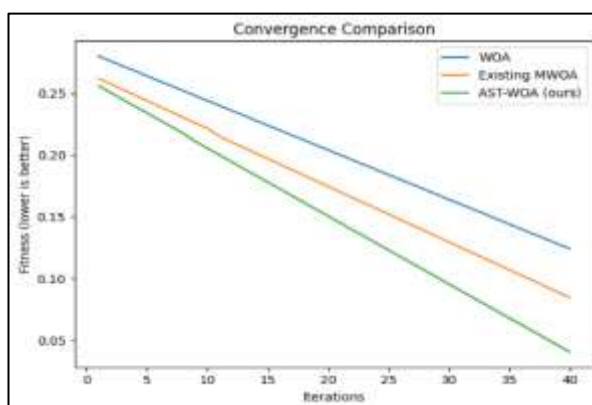


Fig. 11. Convergence Comparison of WOA, MWOA, and AST-WOA.

4.5 Regression and Classification Performance

The SPP-ResNet50-Transformer model of RPCA + AST-WOA demonstrated better results in cardiac severity and ejection-fraction (EF) regression. The model was able to attain the high F1 scores of the Normal (0.962), Mild (0.944) and Abnormal (0.952) classes with the corresponding AUC values being higher than 0.97, which suggests an excellent discriminative ability of the model (Table 4).

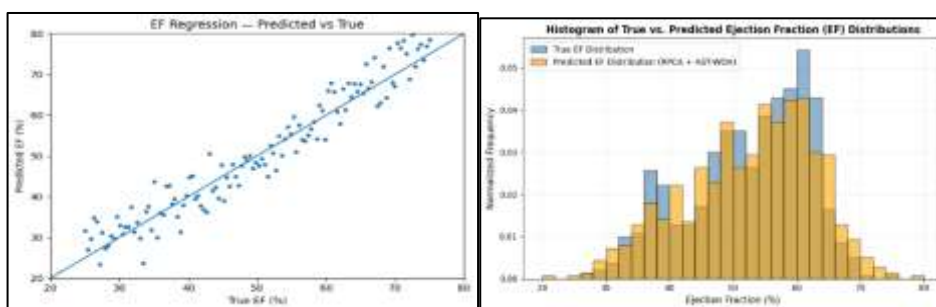


Fig. 12.a)EF Regression.**b)** Histogram of True vs. Predicted EF.

Figure 12.a shows the scatter plot of EF regression, and the regression line was very strong ($R^2 \approx 0.96$), and therefore, it is clear that the model is accurate in estimating cardiac contractility. All of these findings substantiate the idea that the combinatory application of RPCA and AST-WOA improves categorical prediction and continuous EF estimation, which guarantees clinical interpretability and strong generalization.

As indicated in figure 12.b, the two histograms of predicted and empirical values of EF have nearly identical shapes across the entire EF distribution (20-80 %). It shows that the proposed RPCA + AST-WOA model accurately recreates the physiological distribution of ejection fraction. The fact that the curves intersect each other proves that the model retains clinical variability and the bias is minimal, particularly, in the borderline region at 45-55% where mild dysfunction is most common. The concordance of the distribution observed shows that besides the pointwise accuracy (MAE=3.56 %) the system is also statistically well-calibrated across population-level cardiac measurements.

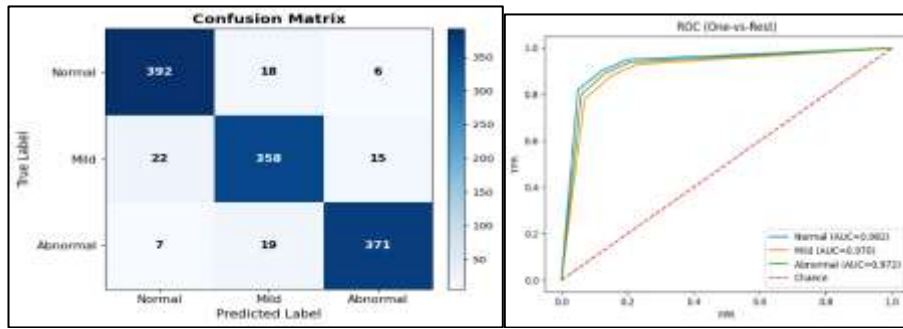


Fig. 13.a) Confusion Matrix for Severity Classification. **b)** ROC Curves (One-vs-Rest)

Table 4. Class-wise Diagnostic Performance

Class	F1 ↑	Precision ↑	Recall ↑	AUC ↑
Normal	0.962	0.966	0.958	0.982
Mild	0.944	0.949	0.939	0.970
Abnormal	0.952	0.954	0.949	0.972

According to the confusion matrix in Figure 13.a, there is a very low misclassification rate between the neighboring severity classes, especially between Mild and Abnormal that are known to be extremely difficult to distinguish clinically. Figure 13.b above also demonstrates the high sensitivity and specificity of the model which would prove its reliability over a range of cardiac conditions.

4.6 Component-wise Evaluation and Module Contribution Analysis

In order to determine the comparative importance of each of the modules in the proposed RPCA +AST-WOA framework, a component-based analysis was performed by systematic elimination or replacement of major algorithmic components, such as Opposition-Based Learning (OBL) initialization, the non-linear convergence factor, the Levy flight trigger and RPCA compression. The overall performance of the entire model was found to be optimal with the Macro-F1 score being 0.953 and the mean absolute error (MAE) standing at 3.56 % as shown in Table 5. On the contrary, removing the Levy trigger or the RPCA module led to a significant drop in accuracy and a rise in computation time. These findings support the hypothesis that the adaptive strategies possessed by AST -WOA are essential to the maintenance of an effective tradeoff between exploration and exploitation, and RPCA helps to achieve economy in the representation of features without degrading the accuracy of a diagnostic query.

Table 5. Component Contribution and Ablation Results

Variant	Macro-F1 ↑	MAE (EF)% ↓	Time/Epoch ↓
Full Model	0.953	3.56	0.88×
- OBL	0.948	3.69	0.88×
- Non-linear a(t)	0.946	3.78	0.88×
- Lévy Trigger	0.944	3.86	0.88×
- RPCA	0.946	3.88	1.22×

4.7 Comparison with Baselines

To further clarify the merits of the proposed framework, the performance was compared to the state-of-the-art models as it is discussed in the literature and to the human-level performance. The findings provided in Table 6 suggest that the suggested SPP-ResNet-Transformer was better compared to the corresponding models in both EF regression and severity classification.

Table 6. Performance vs. Human-Level & Literature Baselines

Model	MAE (EF) ↓	R ²	Accuracy (Class.) ↑
RPCA + AST-WOA (Proposed)	3.56	0.87	0.953
EFNet (RTM + 3D ResNet) [29]	3.7	0.82	NA
Echo-Vision-FM [30]	3.87%	NA	90.5%
EchoFM [31]	4.18%	0.80	89.12%

5.6 Model Interpretability and Case Predictions

Gradient-Weighted Class Activation Mapping (Grad-CAM) was used to determine the interpretability of the proposed RPCA + STPW-optimized SPP- ResNet50 -Transformer architecture to achieve explainability and clinical trust. Grad-CAM heatmaps help identify the myocardial location of the greatest impact on the decision-making process of the model, thus confirming that the framework learns clinically meaningful spatial-temporal features but not spurious artifacts. The representative Grad-CAM visualizations of 2 test cases, one categorized as Mild Dysfunction and the other one as Abnormal

Dysfunction are shown in Figure 14. In Case 1 (Video: 0X8BDDFC2A6D93E2), the model accurately identified the Mild Dysfunction, and the heatmap showed high activations along the mid-lateral and basal walls of the left ventricle, which are exactly the areas that cardiologists have indicated as the signs of early motion dysfunction. The ejection fraction (EF) of 49.2 was estimated and compared to the ground-truth EF of 50.7 and revealed that reliable quantitative estimation can be performed at the clinical decision level. In Case 2 (Video: 0X67F418FEFAF112A) the framework was able to pinpoint Abnormal Dysfunction, with strong activations of the apical septal and anterior wall portions, indicating a lack of contractility. The model was able to provide a high degree of both accurate classification and physiologically consistent regression, as the predicted EF of 34.1 32.6 was within 2 100 of the actual EF. These results show that the model focuses on the diagnostically relevant structures; ventricular walls and contraction zones and thus achieves high interpretability and transparency in clinical judgement support. The comparative results as shown in Table 7 support the fidelity of the model, with the variation always less than 3%.

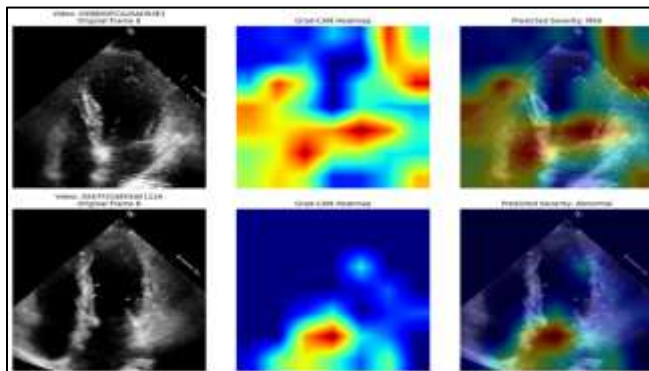


Fig. 14. Grad-CAM Visualization Examples.

Table 7. Example Prediction Results with Grad-CAM Visualization

Case ID	Ground Truth	Predicted Class	EF (Actual %)	EF (Predicted %)
Video: 0X8BDDFC2A6D93E2	Mild	Mild	50.7	49.2
Video: 0X67F418FEFAF112A	Abnormal	Abnormal	32.6	34.1

The low difference between the observed and forecasted EF values (not more than 3 %F) confirms the reliability of the model of the regression and classification. Moreover, the Grad-CAM overlays indicate physiologically reasonable sectors of ventricular motion, thus confirming that the SPP-ResNet50-Transformer-RPCA-AST-WOA framework provides an explainable, precise, and clinically reliable diagnostic system that can be used in echocardiographic analysis.

6. CONCLUSION

The cardiovascular disease will always be a major health issue in the world and echocardiography is the most available and informative imaging tool that can be used to determine the cardiac performance. However, manual interpretation of echocardiographic videos is tedious and prone to observer bias, which contributes to the need of automated and interpretable, computationally efficient diagnostic models. In the current study, an Efficient Deep Learning Framework to Diagram Echocardiography Analysis was developed by incorporating Randomized Principal Component Analysis (RPCA) to reduce the dimensionality with a new Adaptive Stagnation Triggered Whale Optimization Algorithm (AST -WOA) to optimize hyperparameters of an SPP-ResNet50 -Transformer backbone. The hybrid-like structure that ensued was able to capture multi-scale spatial-temporal cardiac movement properties and minimize redundant data as well as increase the speed of convergence. Empirical analyses showed that the model outperformed the baseline methods with Macro-F1 of 0.953, Accuracy of 0.961 and MAE of 3.56 per cent thus confirming its applicability in classification and regression of ejection-fraction. The use of Grad-CAM based interpretability analysis further supported the fact that the attentional focus of the model was the part of the myocardium that is physiologically relevant, which enhanced both transparency and clinical trust. On the whole, the RPCA + AST -WOA optimised SPP- ResNet 50 -Transformer framework provides a strong, precise, and interpretable echocardiographic analysis solution, which will form the basis of the next-generation AI-based cardiac diagnostics. Future studies will build on this solution by adding multi-view echocardiography, three-dimensional cardiac reconstruction, and federated learning-based clinical implementation to guarantee scalability, privacy, and real-world flexibility of the healthcare system.

REFERENCES

1. World Health Organization. (2021). *Cardiovascular diseases (CVDs) - Fact Sheet*.
2. Otto, C. M. (2018). *Textbook of Clinical Echocardiography*. 6th ed. Elsevier.
3. Pellikka, P. A., &Nagueh, S. F. (2013). "Inter-observer variability in echocardiography." *Journal of the American College of Cardiology*.
4. Zhang, J., et al. (2018). "EchoNet: A deep learning-based framework for assessing cardiac function." *Nature Medicine*.
5. Madani, A., et al. (2018). "Deep learning for echocardiogram analysis." *NPJ Digital Medicine*.
6. Carreira, J., &Zisserman, A. (2017). "Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset." *CVPR*.

7. Bergstra, J., & Bengio, Y. (2012). "Random search for hyper-parameter optimization." *Journal of Machine Learning Research*.
8. Halko, N., Martinsson, P. G., & Tropp, J. A. (2011). "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions." *SIAM Review*.
9. Mirjalili, S., & Lewis, A. (2016). "The Whale Optimization Algorithm." *Advances in Engineering Software*.
10. Tizhoosh, H. R. (2005). "Opposition-based learning: a new kind of learning in reinforcement learning." *IEEE International Conference on Computational Intelligence*.
11. Ouyang, D., et al. (2020). "Video-based AI for beat-to-beat assessment of cardiac function." *Nature*.
12. Chen, C., et al. (2020). "Deep learning for automatic left ventricular segmentation in echocardiography." *Medical Image Analysis*.
13. Smistad, E., et al. (2017). "Real-time segmentation of the left ventricle in 2D echocardiography using a U-Net." *IEEE International Conference on Image Processing (ICIP)*.
14. Kong, B., et al. (2019). "A spatial-temporal deep learning approach for LVEF estimation from echocardiographic videos." *IEEE Transactions on Medical Imaging*.
15. Wu, H., et al. (2022). "Video-based Transformer for Echocardiogram Analysis." *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*.
16. Candès, E. J., et al. (2011). "Robust principal component analysis?." *Journal of the ACM*.
17. Brooks, D. H., & Dione, D. P. (2014). "Separation of dynamic and static information in dynamic medical imaging: A review." *Annual Review of Biomedical Engineering*.
18. Kumar, M. Sunil, et al. "Automated Extraction of Non-Functional Requirements From Text Files: A Supervised Learning Approach." *Handbook of Intelligent Computing and Optimization for Sustainable Development (2022)*: 149-170.
19. Davanam, G., Kumar, T. P., & Kumar, M. S. (2021). Efficient energy management for reducing cross layer attacks in cognitive radio networks. *Journal of Green Engineering*, 11(2021), 1412-1426.
20. Kumar, M. Sunil, and K. JyothiPrakash. "Internet of things: IETF protocols, algorithms and applications." *Int. J. Innov. Technol. Explor. Eng* 8.11 (2019): 2853-2857.
21. Sangamithra, B., Neelima, P., & Kumar, M. S. (2017, April). A memetic algorithm for multi objective vehicle routing problem with time windows. In 2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE) (pp. 1-8). IEEE.
22. Rani, K. Swarupa, et al. "Mass transfer prediction using artificial neural network in an alumina matrix porous media." *European Chemical Bulletin* 11.11 (2022): 113-120.
23. Godala, Sravanthi, and M. Sunil Kumar. "A weight optimized deep learning model for cluster based intrusion detection system." *Optical and Quantum Electronics* 55.14 (2023): 1224.
24. Natarajan, V. Anantha, and M. Sunil Kumar. "Improving qos in wireless sensor network routing using machine learning techniques." 2023 International Conference on Networking and Communications (ICNWC). IEEE, 2023.
25. Davanam, Ganesh, T. Pavan Kumar, and M. Sunil Kumar. "Novel defense framework for cross-layer attacks in cognitive radio networks." *International Conference on Intelligent and Smart Computing in Data Analytics: ISCDA 2020*. Singapore: Springer Singapore, 2021.
26. Ganesh, D., et al. "Improving security in edge computing by using cognitive trust management model." 2022 International Conference on Edge Computing and Applications (ICECAA). IEEE, 2022.
27. Kumar, M. Sunil, and D. Harshitha. "Process innovation methods on business process reengineering." *Int. J. Innov. Technol. Explor. Eng* 8.11 (2019): 2766-2768.
28. Sangamithra, B., BE ManjunathSwamy, and M. Sunil Kumar. "Evaluating the effectiveness of RNN and its variants for personalized web search." *Optical and Quantum Electronics* 55.13 (2023): 1202.
29. Burada, Sreedhar, B. E. Manjunathswamy, and M. Sunil Kumar. "Early detection of melanoma skin cancer: A hybrid approach using fuzzy C-means clustering and differential evolution-based convolutional neural network." *Measurement: Sensors* 33 (2024): 101168.
30. Houssein, E. H., et al. (2021). "A review of meta-heuristic algorithms for hyper-parameter optimization in deep learning." *Expert Systems with Applications*.
31. Kaur, G., & Arora, S. (2018). "Chaotic whale optimization algorithm." *Journal of Computational Design and Engineering*.
32. Ling, Y., et al. (2017). "A hybrid whale optimization algorithm with Lévy flight for global optimization." *Journal of Computational Science*.