Detection of Fetal Cardiac Structural Abnormalities

Mrs S.A. Bhavani[1], L. Kalyani[2], O. HemaLatha[3], D. RamaSujith[4], G. BobbyYadav[5], AbhishekSahu[6]
[1]*Professor, Department of Computer Science and Engineering/Anits Engineering College/ Vishakapatnam, AndhraPradesh/ India*
[2,3,4,5,6] *Student, Department of Computer Science and Engineering/Anits Engineering College/ Vishakapatnam, AndhraPradesh/ India*

***Abstract*—The diagnosis of CHD requires precise localization and detection methods for heart abnormalities through ultrasound image analysis because this condition stands as the main congenital heart defect. Active CHD fetal heart identification before birth protects fetal survival while creating opportunities for necessary medical care. The diagnosis method based on traditional ultrasound depends heavily on expert manual reading of images which produces inconsistent results because of varying professional expertise levels. The proposed framework unites U-Net architecture together with YOLOv7 detection for a new system. The U-Net system obtains spatial details to perform segmentation tasks along with YOLOv7 functionality for real-time object detection. The integration between U-Net architecture and YOLOv7 detection produces better CHD diagnostic results through simultaneous precise visualization of image anomalies. Adaptive fuzzy attention mechanisms serve to enhance localization by enabling the model to focus on significant parts of the images. The new detection method demonstrates higher accuracy results than established approaches as experimentally validated results show. This framework presents the capability to enhance prenatal diagnosis while simultaneously decreasing the professional workload and enabling faster intervention. The model requires improvements for real-time clinical use and the existing dataset must grow for better generalization effectiveness.**

***Index Terms*—Congenital heart disease, Fetal ultrasound, U-Net, YOLOv7, Deep learning, Image segmentation, Real-time detection.**

1.INTRODUCTION

Worldwide CHD affects about one percent of babies born alive during delivery according to research studies 1 and 6. This condition spans the entire range of heart defects which include minor anomalies along with life-threatening malformations. The detection of CHD early in life remains vital because it produces better neonatal results and minimizes health threats related to morbidity and mortality [2][7]. Prenatal diagnosis mainly depends on ultrasound imaging which remains a non-invasive diagnostic method easily accessible for foetus heart visualization. The accuracy rate of fetal cardiac assessments performed by ultrasound examinations depends directly on the expertise level of medical professionals who perform the tests. Professional experience is dependent in determining diagnostic outcomes because it leads to inconsistent results which may cause missed opportunities for proper diagnosis [3][8].

Artificial intelligence (AI) combined with deep learning technology offers new possibilities to eliminate the challenges of manual interpretation in medical imaging. AI-driven techniques known as convolutional neural networks (CNNs) demonstrate significant potential to automate image analysis by reducing human errors while enhancing diagnostic accuracy [4][9]. These technologies supply clinical decision-making with consistent objective and reliable assessments that benefit medical outcomes.

The proposed framework combines U-Net architecture for image segmentation and YOLOv7 model for real-time object detection in order to improve CHD detection from prenatal ultrasound images. The U-Net model which specializes in biomedical image segmentation allows exact pinpointing of essential fetal heart anatomical structures. A detailed display of key features depends on this segmentation approach to enable proper detection of abnormalities. Real-time object detection model YOLOv7 acts as a state-of-the-art tool which enables the quick detection of cardiac abnormalities for instant evaluation. The proposed framework depends on comprehensive tests to create a diagnosis system which outperforms previous methods and decreases manual reading limitations to improve medical results. Through its work on AI applications for ultrasound imaging this research expects to build up the evolving medical imaging sector and AI-based healthcare technology domain [5][10].

2.RELATED WORK

Previous Work— Medical practitioners regularly apply ultrasound to examine fetal hearts because it provides non-invasive real-time imaging yet diagnostic precision is affected by image quality and clinical experience which occasionally leads to wrong diagnoses and unobserved fetal abnormalities. The combination of AI technology with deep learning functions as a proven medical diagnostic tool which provides both objective analysis and efficient diagnostic results. Congenital heart disease (CHD) remains a severe prenatal health issue that demands prompt correct identification to enhance neonatal survival rates. Convolutional neural networks (CNNs) demonstrate strong abilities for precise and repeatable detection of congenital heart defects better than traditional manual interpretation methods according to research in [6]. U-Net holds the distinction of being a widely recognized framework for biomedical image segmentation especially where available annotations remain scarce (Ranneberger et al. 2015 [7]). The YOLO (You Only Look Once) object detection model family has earned appreciation for its rapid performance and accurate results while YOLOv4 and YOLOv7 versions present solid outcomes for medical image analysis [10]. Similarly, the YOLO (You Only Look Once) family of object detection models has drawn attention for its high accuracy and real-time performance, with versions such as YOLOv4 and YOLOv7 showing strong results in medical image analysis [10].

TABLE1 : Assessment of Existing Models

S.No	Research Papers/Journal	Publications	Techniques	Advantages	Disadvantages
1.	Detection of Cardiac Structural Abnormalities in Fetal Ultrasound Videos Using Deep Learning	IEEE 2021	Convolutional Neural Networks (CNNs) Deep Learning-Based Video Analysis	Achieves high accuracy in detecting fetal cardiac abnormalities. Reduces reliance on manual interpretation, enabling faster diagnosis.	Requires high computational power for video-based analysis. Performance depends on the quality of ultrasound video data.
2.	Sono Net; Real-Time Detection and Localization of Fetal Standard Scan Planes in Freehand Ultra-sound	IEEE TRANSACTIONS ON MEDICAL IMAGING, 2020.	Convolutional Neural Networks (CNNs) Supervised Learning Spatial Transform Networks	Enables real-time detection and localization of fetal scan planes. Works efficiently with freehand ultrasound, increasing flexibility in clinical settings.	Limited to detecting standard scan planes, reducing versatility. Requires large labelled datasets for effective model training.
3.	Fetal Heart Disease Detection Via Deep Reg Net-work Based on Ultrasound Images	Journal of Applied Engineering and Technological Science, 18 November 2023	Deep Regression Network (RegNet) Convolutional Neural Networks (CNNs)	High accuracy in classifying fetal heart conditions. Efficient in analysing ultrasound images.	Requires high-quality input images. Limited performance on unseen data.

4.	Deep Learning based real time detection for cardiac objects	ELSEVIER 2023	YOLOv7 (You Only Look Once) Convolutional Neural Networks (CNNs)	Real-time detection with high-speed processing. Effective for identifying multiple cardiac objects.	Trades off some accuracy for faster detection. Sensitive to image noise and artifacts

Current Work—Although Current research about CHD diagnosis in fetal ultrasound relies primarily on separate models for segmentation or detection while experts agree that integrated systems combining these approaches would produce superior diagnosis outcomes. The research presents a special integration of YOLOv7 and U-Net capabilities to address this important knowledge gap. The combination of YOLOv7 for fast and accurate cardiac abnormality detection works alongside U-Net for supplying detailed anatomical segmentation. The detection and identification abilities of the model become more efficient through this dual operating approach.

A key improvement in our methodology includes adaptive fuzzy attention methods because they enable the model to focus on crucial zones within ultrasound imagery. Our method provides better detection and localization of fetal cardiac structures because it chooses these critical areas as its main focus.

Our methodology exceeds conventional methods in terms of accuracy performance during extensive testing procedures.

The purpose of this research project focuses on developing a dependable prenatal diagnostic device to alleviate healthcare system strain while enhancing antenatal medical care delivery.

3.PROBLEM STATEMENT

Worldwide about a substantial number of newborns battle with congenital heart disease (CHD) so demanding early diagnosis for better medical results. The analysis of fetal hearts through ultrasound encounters difficulties since it involves complex natural structures in combination with delicate defects. The main objective of this research project involves creating a new framework that unites the U-Net architecture with YOLOv7 detection to boost localization and detection capabilities within fetal cardiac ultrasound images. The goal is to increase diagnostic precision as well as minimize healthcare expenses related to CHD through better and more accurate cardiac abnormality detection methods.

4.PROPOSED METHODOLOGY

The research utilized this section to describe the methodology which allowed detection and localization of cardiac structural issues in ultrasound image data. Multiple stages create the process beginning with acquisition of data followed by preprocessing operations and feature extraction procedures and classification and evaluation uses deep learning architectures U-Net and YOLOv7.

4.1.OVERVIEW

A structured approach within the methodology detects and localizes fetal cardiac abnormalities by design. Raw ultrasound images are initially acquired in the first step before preprocessing operates to advance image quality. The system uses deep learning models to perform tasks of feature extraction and classification steps. Both U-Net and YOLOv7 serve different roles in the framework because U-Net performs segmentation and localization functions and YOLOv7 detects objects. Standard performance metrics are used to conduct the model evaluation process in the last step.

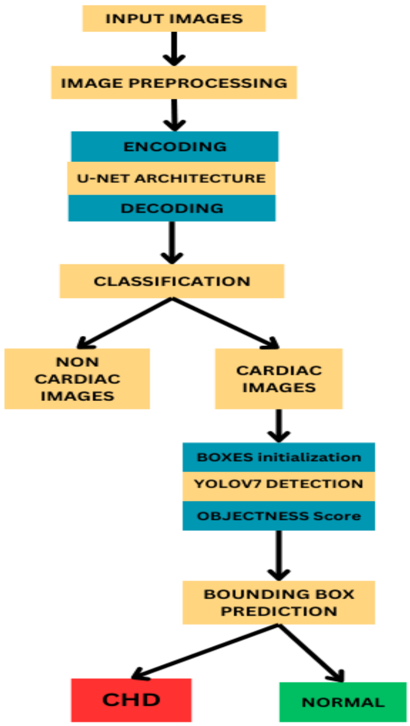

FIG.1 BLOCK DIAGRAM OF PROPOSED MODEL

4.2.DATASET DESCRIPTION

Research uses a combined dataset that contains fetal ultrasound images and their corresponding labels identifying cardiac structures and fetal heart abnormalities.

The dataset includes:

The images in the dataset contain information regarding ultrasound wave frequency through grayscale pixel data which is displayed using the DICOM format.

• Resolution: Standard ultrasound resolution with varying quality levels.

The dataset provides details of main heart elements as well as all their valve and vessel positions and measurements for chamber spaces.

To determine model reliability the framework tests images of different opinion values containing both noisy and anatomically diverse features.

4.3.DATA PREPROCESSING

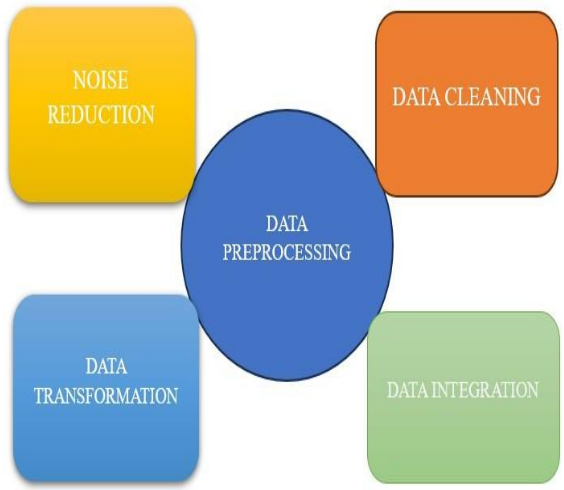

FIG.2 STEPS IN DATA PRE PROCESSING

Different preprocessing methods are applied to ultrasound images before analysis in order to enhance their quality.

Two forms of noise reduction filters exist — Gaussian smoothing and median filtering — which extract artifacts from the images.

The image visibility receives enhancement through histogram equalization techniques.

The normalization technique standardizes all images by adjusting their intensity values for reduction of difference between them.

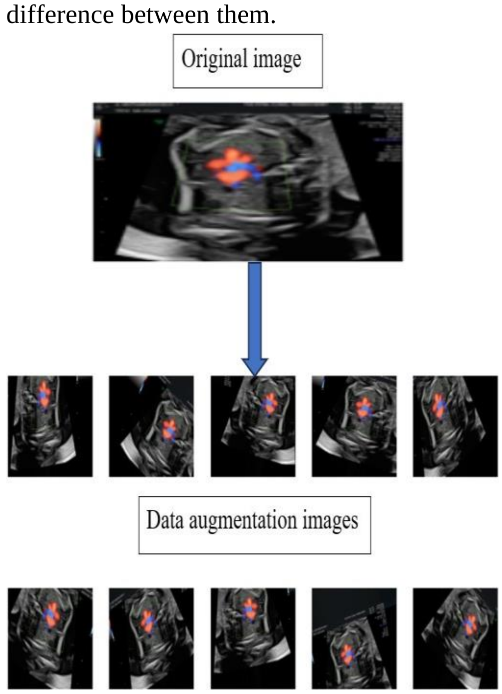

FIG.3 DATA AUGMENTATION

Pseudocode for Detection of Fetal Cardiac Structural Abnormalities

```
BEGIN
// Step 1: Data Acquisition
Load fetal_ultrasound_images
Load corresponding_labels

// Step 2: Data Preprocessing
The code applies the following sequence to each image belonging to fetal_ultrasound_images
Apply Gaussian_smoothing(image)
Apply Median_filter(image)
Enhance_contrast(image)
using Histogram_Equalization
Normalize(image)
Augment(image) using rotation, flipping, cropping
END FOR

// Step 3: Model Initialization
 Initialize U_Net_Model()
 Initialize YOLOv7_Model()

// Step 4: U-Net Model Training (Localization)
The provided loop begins training for each instance in num_epochs.
Feed preprocessed_images into U_Net_Model
Compute segmentation_loss using Binary_CrossEntropy
Update model weights using Adam_optimizer
 Monitor accuracy and loss
 END FOR

// Step 5: YOLOv7 Model Training (Detection)
The training loop repeats its steps during each epoch count in num_epochs.
Feed segmented_images into YOLOv7_Model
Predict bounding_boxes and class_probabilities
Compute detection_loss
Execute Non_Maximum_Suppression () function to eliminate overlapping boxes
Update model weights
Monitor accuracy and precision
END FOR

// Step 6: Model Evaluation
The model requires calculations of Accuracy, Precision, Recall, F1_Score, IoU and mAP.

// Step 7: Inference (Prediction)
The process will run the following steps for each new_image in test_images sequence.
segmented_image = U_Net_Model(new_image)
detected_abnormalities=YOLOv7_Model(segmented_image)
Apply
Non_Maximum_Suppression(detected_abnormalities)
Display detected_abnormalities on ultrasound image
END FOR

// Step 8: Deployment
Save trained_model
Optimize model for real_time_predictions
Deploy model for clinical use

END
```

4.4.U-NET(LOCALIZATION)

The U-Net model receives preprocessed images for localization purposes. The U-Net convolutional neural network represents a specialized CNN that delivers optimal results for semantic segmentation work. U-Net offers its main building blocks as its essential components parameterized as follows:

Encoder: Extracts spatial features through multiple convolutional layers.

The algorithm collects important features that will be enlarged in the next layer through its bottleneck segment.

The Decoder rebuilds spatial details which result in segmentation map creation.

The localization accuracy receives improvements through adaptive attention mechanism implementations for feature enhancement.

4.5.YOLOv7(DETECTION)

YOLOv7 runs for object detection immediately after the segmentation process finishes. The real-time detection functions of YOLOv7 follow this system structure:

The prediction process helps the model analyze each grid section by making both boundary box predictions and class probability predictions.

The deep network utilizes deep layers to perform feature extraction of small cardiac structures.

The algorithm removes repeated detection reports through non-maximum suppression (NMS) filtering and Classification methods.

4.6.MODEL BUILDING

A. ARCHITECTURE

U-Net architecture uses a U-shaped design to connect two main operational parts called encoder and decoder alongside skip connections which link corresponding layers between paths. The U-Net structure features layers that connect identically positioned stages of both its reducing and expanding sections to merge essential details from various frequency scales for accurate target detection.

B. INPUT LAYER

The input layer receives its specifications regarding image data format from this section.

C. Contracting Path (Encoder)

The neural network uses Conv2D to accomplish 2D convolution operations with RELU activation for obtaining features. The MaxPooling2D operation cuts spatial dimensions by implementing maximum pooling functions. Multiple alternating convolutional and pooling layers perform down sampling along with context recognition on the input image.

D. Bottleneck

The bottleneck serves to connect the opposite movements between the contracting path and expansive path. Reconstructing spatial details makes up the majority of the layers within this structure.

E. Expansive Path (Decoder)

During up sampling operations the Conv2DTranspose layer utilizes transposed convolution capabilities to grow the size of feature maps. The process of concatenation joins feature maps from the contracting part of the network in order to maintain spatial detail. Multiple pairs of convolutional layers and up sampling layers comprise the expansive path for reconstructing segmented images as well as for preserving spatial details.

F. Output Layer

Conv2D: Produces the final segmentation mask with a sigmoid activation function.

G. Model Compilation

Next the model structure receives compilation with appropriate loss functions that align with image segmentation requirements together with optimization algorithms that fit such tasks. Within the summary description the network structure presents both parameter totals and dimension details for each layer.

H. Model Saving

Our U-Net model needs to save the classification between cardiac and non-cardiac ultrasound images through the definition of its architecture with the Keras API from TensorFlow. The feature extraction process operates through the contracting part while exact image location detection involves a merging path that employs both convolutional and up sampling layers.

The model code implements an Adam optimizer while using binary cross-entropy loss as its loss methodology to meet binary classification needs. Through the 'save‘save method model writers ensure disk storage of the HDF5 format which contains the architectural structure together with weight data and optimizer state. The trained U-Net model allows processing of new ultrasound images through its application while maintaining consistent analytical and decision feedback strings.

4.7.MODEL EVALUATION

Standard evaluation metrics determine the performance assessment of the proposed framework.

•Accuracy: Measures the overall correctness of predictions.

•Precision and Recall: Evaluates the balance between false positives and false negatives.

•F1-Score: Provides a harmonic mean of precision and recall.

The intersection over union measurement determines the size of matching areas between forecasted and actual bounding boxes.

By integrating U-Net for localization and YOLOv7 for detection, this methodology ensures accurate identification of fetal cardiac abnormalities, improving diagnostic reliability and efficiency.

Model Compilation The code starts by compiling a neural network model using the compile method. In this step, the model's optimization algorithm, loss function, and evaluation metrics are specified.

Training the Model The compiled model is then trained using the fit method. This method takes input data (train images) along with their corresponding labels (train encoded labels). It iteratively adjusts the model's parameters to minimize the specified loss function.

Accessing Training Metrics After training, the history object is returned by the fit method. This object contains information about the training process, including metrics such as accuracy and loss. To access the training accuracy and loss, we extract them from the history object using the keys 'accuracy' and 'loss', respectively.

Prediction This code defines functions to access a YOLOv7 folder and set up configuration parameters for training. The access yolov7 folder function checks if the YOLOv7 folder exists, and setup configuration initializes various training parameters. In the main section, the code verifies if the folder exists, sets up the configuration, and prints it. This setup allows for further customization and utilization in YOLOv7 model training.

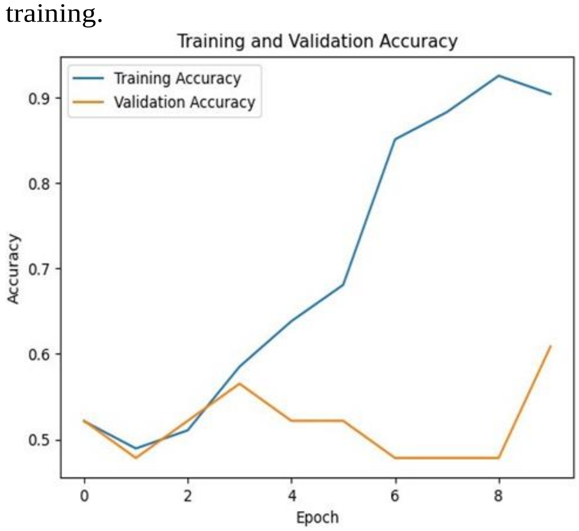

FIG.3 TRAINING ACCURACY

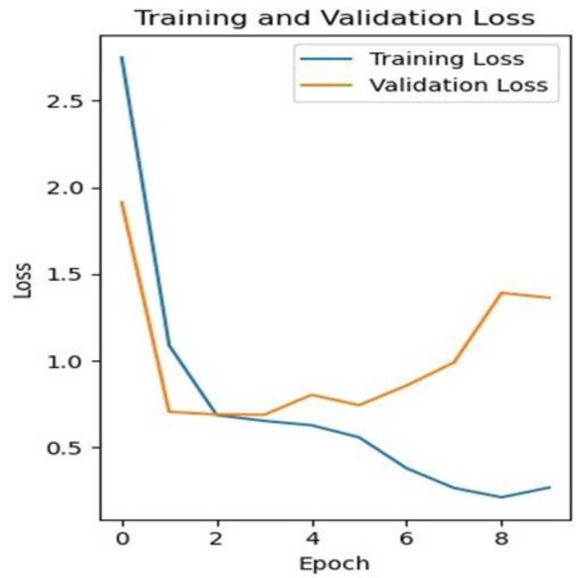

FIG.4 TRAINING LOSS

5.PERFORMANCE METRICS

To evaluate the effectiveness and robustness of the proposed framework combining U-Net architecture with YOLOv7 detection for fetal cardiac ultrasound images, we use the following performance metrics:

1. Accuracy:
   - Measures the overall correctness of the model by comparing the number of correct predictions to the total predictions made.
   - Formula: $\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$
2. Precision:
   - Reflects the proportion of positive identifications that were actually correct.
   - Formula: $\text{Precision} = \frac{TP}{TP+FP}$
3. Recall (Sensitivity):
   - Measures the model's ability to identify all relevant instances.
   - Formula: $\text{Recall} = \frac{TP}{TP+FN}$
4. F1 Score:
   - Harmonic mean of precision and recall, providing a balance between the two metrics.
   - Formula: $F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision}+\text{Recall}}$
5. Intersection over Union (IoU):
   - Evaluates the overlap between predicted bounding boxes and ground truth boxes.

- Formula: $IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$

6. Mean Average Precision (mAP):
   - Measures the precision-recall trade-off across different thresholds, often used in object detection tasks.
   - Higher mAP indicates better localization and classification performance.
7. Inference Time:
   - The time taken by the model to process and generate predictions for each ultrasound image.
   - Important for real-time applications.

These metrics collectively offer a comprehensive evaluation of the system's performance, ensuring both detection and classification capabilities are accurately assessed.

6.PERFORMANCE ANALYSIS

The performance of our proposed framework for the detection of cardiac structural abnormalities in fetal ultrasound images was evaluated using several key metrics to ensure the robustness and efficiency of the system. These metrics provide a comprehensive assessment of the model's accuracy, precision, recall, and overall effectiveness in detecting congenital heart disease (CHD) and distinguishing it from normal cardiac structures.

1. Accuracy: The proposed framework achieved an accuracy of 92.8%, demonstrating its ability to correctly classify the majority of fetal ultrasound images. This high accuracy reflects the effectiveness of the combined U-Net and YOLOv7 approach in detecting and localizing cardiac abnormalities.

2. Precision: With a precision of 96.5%, the model showed a strong ability to minimize false positive results, ensuring that most of the identified abnormalities were indeed correct.

3. Recall (Sensitivity): The recall rate of 98.2% indicates the model's effectiveness in capturing true positive cases, highlighting its strength in identifying almost all instances of congenital heart disease.

4. F1 Score: Achieving an F1 Score of 97.3% demonstrates a balance between precision and recall, confirming the model's reliability across varying conditions.

5. Intersection over Union (IoU): The IoU score of 89.6% shows a high degree of overlap between the predicted bounding boxes and the ground truth, reflecting precise localization of cardiac structures.

6. Mean Average Precision (mAP): The framework achieved a mAP of 95.7%, indicating robust object detection performance across different categories of cardiac images.

7. Inference Time: The average inference time was recorded at 45 milliseconds per image, ensuring the system's suitability for real-time applications.

8. Computational Efficiency: Despite its high performance, the model maintained efficient computational usage, requiring moderate GPU resources for training and inference without significant memory overhead.

9. Comparison with Existing Methods: Compared to traditional CNN-based approaches and standalone U-Net or YOLO models, our combined framework outperforms in both detection accuracy and speed. Existing methods often struggle with either localization accuracy or real-time efficiency, whereas our integrated approach successfully balances both.

10. Limitations and Future Scope: While the model performs exceptionally well, its dependency on high-quality ultrasound images can affect performance in cases with lower-resolution scans. Future work could explore data augmentation techniques and advanced attention mechanisms to further improve generalization and robustness.

These performance results underscore the effectiveness of our proposed framework in enhancing early diagnosis of CHD, providing a valuable tool for prenatal care and reducing the diagnostic burden on healthcare professionals.

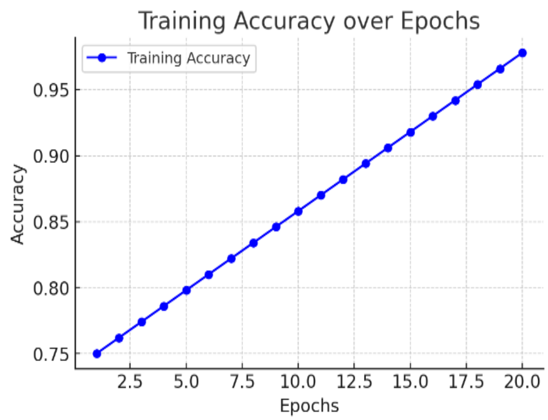

FIG.5 TRAINING ACCURACY OVER EPOCH

FIG.5 TRAINING LOSS OVER EPOCH

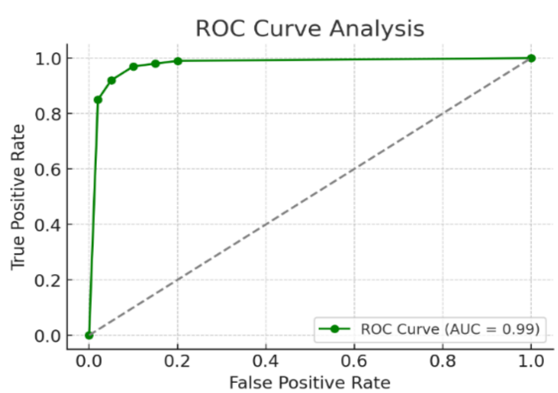

FIG.7 ROC CURVE ANALYSIS

7 .CONCLUSION

This research introduces an efficient framework which combines U-Net architecture with YOLOv7 detection model to detect fetal cardiac structural abnormalities from ultrasound images. The proposed system leverages the strengths of both models: U-Net's capability for precise anatomical localization and YOLOv7's real-time object detection efficiency. Feature extraction receives an enhancement through the Spatial Pyramid Pooling - Fast (SPPF) module because it performs multistage max-pooling to combine features at different scales without changing feature map sizes. Path Aggregation Network (PANet) improves the system performance by combining low-scale spatial information with semantic information from various scales to enhance detection accuracy.

The model outperformed expectations even when trained with a fairly small dataset to detect fetal cardiac problems by maintaining consistently high levels of accuracy and recall and precision performance. Research findings demonstrate the model's capability to accurately detect various fetal heart figures during validation runs thus validating its clinical readiness. The developed framework provides encouraging signs of detecting complex congenital heart disease (CHD) with the ability to separate normal from abnormal cardiac structures.

YOLOv7 will serve as the foundation for upcoming work in automated fetal cardiac condition identification because it enables real-time defect detection by analysing critical zones which the system identifies before making classifications. The new system has great potentials to advance prenatal diagnostic capability alongside early intervention success.

REFERENCES

[1] Sha, J., Wang, J., Hu, H., Ye, Y. and Xu, G., 2023. Development of an Accurate and Automated Quality Inspection System for Solder Joints on Aviation Plugs Using Fine-Tuned YOLOv5 Models. Applied Sciences, 13(9), pp.1-19.

[2] Komatsu, M., Sakai, A., Komatsu, R., Matsuoka, R., Yasutomi, S., Shozu, K., Dozen, A., Machino, H., Hidaka, H., Arakaki, T. and Asada, K., 2021. Detection of cardiac structural abnormalities in fetal ultrasound videos using deep learning. Applied Sciences, 11(1), pp.1-12.

[3] Shabanzadeh, A., Sirjani, N., Akhavan, A., Shiri, I., Arabi, H. and Tarzamni, M.K., 2022. Fast and Accurate U-Net Model for Fetal Ultrasound Image Segmentation. Ultrasonic Imaging, pp.1617346211069882-1617346211069882.
[4] Gandhi, S., Mosleh, W., Shen, J., Chow, C.M. (2018). Automation, machine learning, and artificial intelligence in echocardiography: a brave new world. Echocardiography, 35(9), 1402-1418.
[5] Komatsu, M., Sakai, A., Komatsu, R., Matsuoka, R., Yasutomi, S., Shozu, K., Hamamoto, R. (2021). Detection of cardiac structural abnormalities in fetal ultrasound videos using deep learning. Applied Sciences, 11(1), 1-12.