Archives available at journals.mriindia.com

International Journal on Advanced Electrical and Computer Engineering

ISSN: 2349-9338

Volume 15 Issue 01s, 2026

Two-Stage Machine Learning Pipeline for Fetal Head Analysis

[1]Shivanand S. Gornale, [2*]Priyanka Kamat, [3]Rashmi Siddalingappa, [4]Khang Wen Goh, [5]Kefang Li

[1 2]Department of Computer Science, School of Mathematics and Computing Sciences, Rani Channamma University, Belagavi, Karnataka, India.

[3]York St John University, London campus, Clove Crescent, E14 2BA, UK.

[4]Faculty of Data Science and Information Technology, INTI International University,71800 Nilai, Malaysia.

[5]Faculty of Applied Sciences, Macao Polytechnic University, R. de Luís Gonzaga Gomes, Macao, China.

E-mails: [1]shivanand1971@rcub.ac.in ,[2]priyankakamat@rcub.ac.in, [3]r.siddalingappa@yorksj.ac.uk, [4]kefengl@mpu.edu.mo, [5)]khangwen.goh@newinti.edu.my.

Corresponding author: [2]priyankakamat@rcub.ac.in

Peer Review Information

Submission: 05 Dec 2025

Revision: 25 Dec 2025

Acceptance: 10 Jan 2026

Keywords

Fetal Ultrasound Imaging, Head Circumference Measurement, Machine Learning Methods, Automated Segmentation Processes, Trimester-Based Classification.

Abstract

Background: The accurate measurement of fetal head circumference (HC) and the precise estimation of gestational age are vital components of prenatal assessment. However, manual techniques are often labor-intensive and subject to variability among different observers. Automated methods have the potential to enhance measurement consistency and reduce the clinical workload. **Purpose:** This study presents an integrated two-stage machine learning pipeline aimed at automating fetal head segmentation and trimester-specific gestational age classification using ultrasound data. **Methods**: The proposed framework is composed of two stages. In the first stage, fetal head segmentation is conducted using a Random Forest classifier that leverages handcrafted pixel-level features. In the second stage, features related to shape, intensity, and texture are extracted from the segmented region and classified using a top-3 ensemble of LightGBM, Support Vector Machine with a Radial Basis Function (RBF) kernel, and XGBoost, with adjustments for class imbalance and feature optimization. **Results:** The segmentation framework achieved a mean accuracy of 89.02%. The trimester classification model achieved an overall accuracy of 85.62%, a precision of 86.57%, a recall of 85.62%, and an F1-score of 85.59%. **Conclusion:** The proposed two-stage pipeline effectively achieves precise segmentation of the fetal head and accurate trimester classification, highlighting its considerable potential for integration into clinical decision-support systems, especially in healthcare environments with constrained resources.

1. Introduction

Ultrasound imaging is commonly used in prenatal care because it is safe, non-invasive, and enables real-time visualization of the fetus. Biometric measurements are used to track fetal growth and development, to identify developmental issues, and to calculate gestational age[1, 2]. Intrauterine growth restriction (IUGR), microcephaly, macrocephaly, and other conditions are commonly diagnosed

using fetal head circumference (HC). It is an important hallmark of brain growth [3]. It is necessary to precisely segment the fetal head boundary to measure HC.

Deep learning methods have shown excellent accuracy in fetal head segmentation, particularly U-Net-based designs. Despite their success, these methods' usage in real-time and resource-constrained clinical settings is restricted due to their requirement for substantial processing power and sizable annotated datasets[4, 5]. Interpretable and computationally efficient alternatives are offered by traditional machine learning methods. However, complex ultrasound patterns, border ambiguities, and class differences sometimes present challenges[6, 7]. Capturing multi-scale contextual information remains a major challenge in obtaining accurate fetal head delineation.

This paper suggests a two-step approach for fetal head circumference analysis utilizing ultrasound images in order to solve these problems. To increase boundary accuracy, the first step uses a random forest-based segmentation model that has been improved with contrast-limited adaptive histogram equalization, bilateral filtering, and ellipse-constrained post-processing[8–10]. In the second stage, quantitative shape, intensity, and texture features are extracted from the segmented fetal head region and used for classification. Class imbalance is handled using synthetic sampling, followed by ensemble feature selection and robust feature normalization [11–13]. A top-3 ensemble classifier, combining LightGBM, SVM with an RBF kernel, and XGBoost, is employed to classify images into the first, second, and third trimesters.

1.1 Key Contributions

The main contributions of this study are as follows:

1. A new two-stage framework is proposed for fetal ultrasound analysis, integrating segmentation and classification in a unified pipeline.
2. An ellipse-guided random forest segmentation strategy is introduced to enhance anatomical accuracy and robustness.
3. A robust fetal trimester classification pipeline is proposed using outlier-resistant feature scaling and multiple machine learning models to improve stability and generalization on medical ultrasound features.
4. Voting and stacking ensembles with top-model selection are used to improve classification performance. Standard evaluation metrics confirm the clinical reliability of the proposed approach.

2. Related Work

Recent research has applied machine learning and deep learning techniques to fetal ultrasound imaging to support automated diagnosis and biometric analysis. Early studies relied on handcrafted feature extraction combined with traditional classifiers, while more recent work has focused on convolutional neural networks for segmentation, classification, and gestational age estimation[9, 14]. Although deep learning methods often achieve high accuracy, they typically require large annotated datasets and high computational resources, motivating the exploration of hybrid and feature-based approaches.

Micucci et al. [6] reviewed machine learning (ML) applications in ultrasound imaging for medical diagnosis, categorizing studies by tasks such as detection, segmentation, and classification. Traditional ML pipelines using handcrafted features like local binary patterns and histograms of oriented gradients with support vector machines have shown effectiveness. Deep learning (DL) methods, especially those using convolutional neural networks (CNNs), have been applied to fetal gestational age estimation, improving diagnostic accuracy and reducing analysis time. Carneiro et al. [15] proposed an automated system for detecting and measuring fetal anatomical structures, including biparietal diameter, head and abdominal circumferences, femur and humerus lengths, and crown-rump length, using probabilistic boosting tree classifiers. Rawat et al. [16] highlighted the importance of early detection of fetal abnormalities and noted that noise in ultrasound images can reduce accuracy. Their approach used artificial neural networks with Levenberg–Marquardt backpropagation and gradient vector flow segmentation to enhance performance. Gornale et al. [3] introduced trimester-aware, multi-stage DL pipelines for fetal biometric analysis using custom datasets and the HC18 benchmark dataset. Their framework combined preprocessing, segmentation, and classification with multiple architectures and ensemble strategies, using hybrid loss functions to improve accuracy. Lee et al. [14] developed a CNN-based framework for gestational age estimation directly from ultrasound images, removing the need for manual biometric measurements and showing superior performance, especially in late gestation. Gornale et al. [17] used UNet for precise segmentation of head, abdominal, and femur measurements and evaluated transfer learning models such as ResNet, VGG16,

MobileNet, and DenseNet. Kumar et al. [1] reviewed noise reduction techniques in medical imaging, focusing on statistical (wavelet, Fourier) and probability-based (Bayesian, maximum likelihood) methods. Zhang et al. [18] compared segmentation-based and segmentation-free approaches for fetal head circumference estimation, analyzing the computational efficiency and accuracy of CNNs and regression forests. Alzubaidi et al. [4] developed an end-to-end framework integrating segmentation, biometric measurement, and gestational age estimation using UNet, DeepLabv3, and EfficientNetB0, validated on the HC18 dataset. Yousefpour et al. [8] reviewed ML and DL methods for fetal ultrasound, emphasizing classification, object detection, and segmentation for anomaly detection, and identified the need for greater robustness and clinical applicability. Kim et al. [7] proposed DL pipelines for automated fetal biometric measurements, including head circumference and biparietal diameter, by combining plane validation, boundary detection, and measurement refinement to achieve accurate localization despite imaging artifacts.

Over the past decade, researchers have explored various analytical methods and abnormalities in fetal imaging using diverse image-processing techniques. Accurate image representation and description are fundamental steps in these analyses. Numerous studies have made significant contributions to this field, and a summary of recent research is presented in Table 1.

Table 1. Summary of Recent Work

Authors	Dataset	Methodology	Results
Alzubaidi M., et al. (2022) [4]	HC18 Grand Challenge dataset	Ensemble Transfer Learning	MIoU=98.53%
Gornale S., et al. (2024) [17]	Created Own Dataset	U-Net	Accuracy= 99.86%
Govindarajan, M. P., et al.[2]	HC18 Grand Challenge dataset	Unsupervised CAD-based method	DSC=97.5%
Torres HR, et al. (2022)[19]	HC-18 dataset	Regression convolutional neural network (RCNN)	DSC=97.97%
Gornale S., et al. (2025)[3]	HC-18 dataset & Own dataset	Ensemble approach	DSC=98.74%
Wang J., et al. (2021) [20]	HC-18 dataset	Ellipse-guided multi-task network	DSC=97.97%
Zeng Y., et al. (2021) [21]	HC-18 dataset	Deeply supervised attention-gated (DAG) V-Net	DSC=97.93%
Li P., et al. (2020) [22]	HC-18 dataset	Scale attention pyramid deep neural network (SAPNet)	DSC=97.94%
Fiorentino MC., et al. (2021)[23]	HC-18 dataset	Region-proposal CNN for head localization	DSC=97.75%
Al-Bander., et al. (2019)[24]	HC-18 dataset	R-CNN and FCN	DSC=97.73%

Trimester-based fetal head circumference (HC) categorization has a substantial research gap. The inability of existing methods to capture HC changes over the course of trimesters is the reason why they are unable to predict fetal growth patterns [3, 10]. To address this, the current study proposes a conventional machine learning approach for trimester-specific HC classification. The recommended solution makes use of machine learning and robust data management techniques to increase the accuracy of HC predictions, as will be discussed in the next sections.

3. Materials and Methods

A two-step methodology for the automatic analysis of fetal heads from ultrasound images is presented in this research. The framework is proposed to segment the fetal head first, then classify the trimester. The primary goal is to ensure that trimester classification is based on the anatomically significant aspects of the fetal head. The fetal head region is identified in the first step using a pixel-wise segmentation technique. Handcrafted characteristics that define intensity, spatial location, edge details, and local texture are used to build an ML

classifier [11, 25]. In the following step, trimester classification is conducted using features extracted solely from the segmented fetal head region. These features capture the shape, intensity distribution, and texture patterns that change with gestational age. This two-step approach enhances both robustness and interpretability by clearly distinguishing anatomical localization from gestational age classification [26]. Figure 1 illustrates a schematic overview of the proposed framework.

3.1 Dataset Description

The dataset included 1,334 fetal head ultrasound images. A CSV file containing head circumference (HC) measurements and pixel sizes was included in the training set, which comprised 999 images. The test set included 335 images with only pixel size data. The Department of Obstetrics at Radboud University Medical Center in Nijmegen, Netherlands, took images of 551 pregnant women in the first, second, and third trimesters. The Voluson E8 and Voluson 730 ultrasound machines (General Electric, Boston, MA, USA) were used to acquire them. The study was authorized by the local ethics council (CMO Arnhem-Nijmegen), and all examinations were carried out by qualified sonographers in accordance with the Declaration of Helsinki. Depending on the size of the fetus, each photograph had a resolution of 800 × 540 pixels with pixel sizes varying from 0.052 to 0.326 mm. A sonographer manually marked each image's skull region with an ellipse [27].

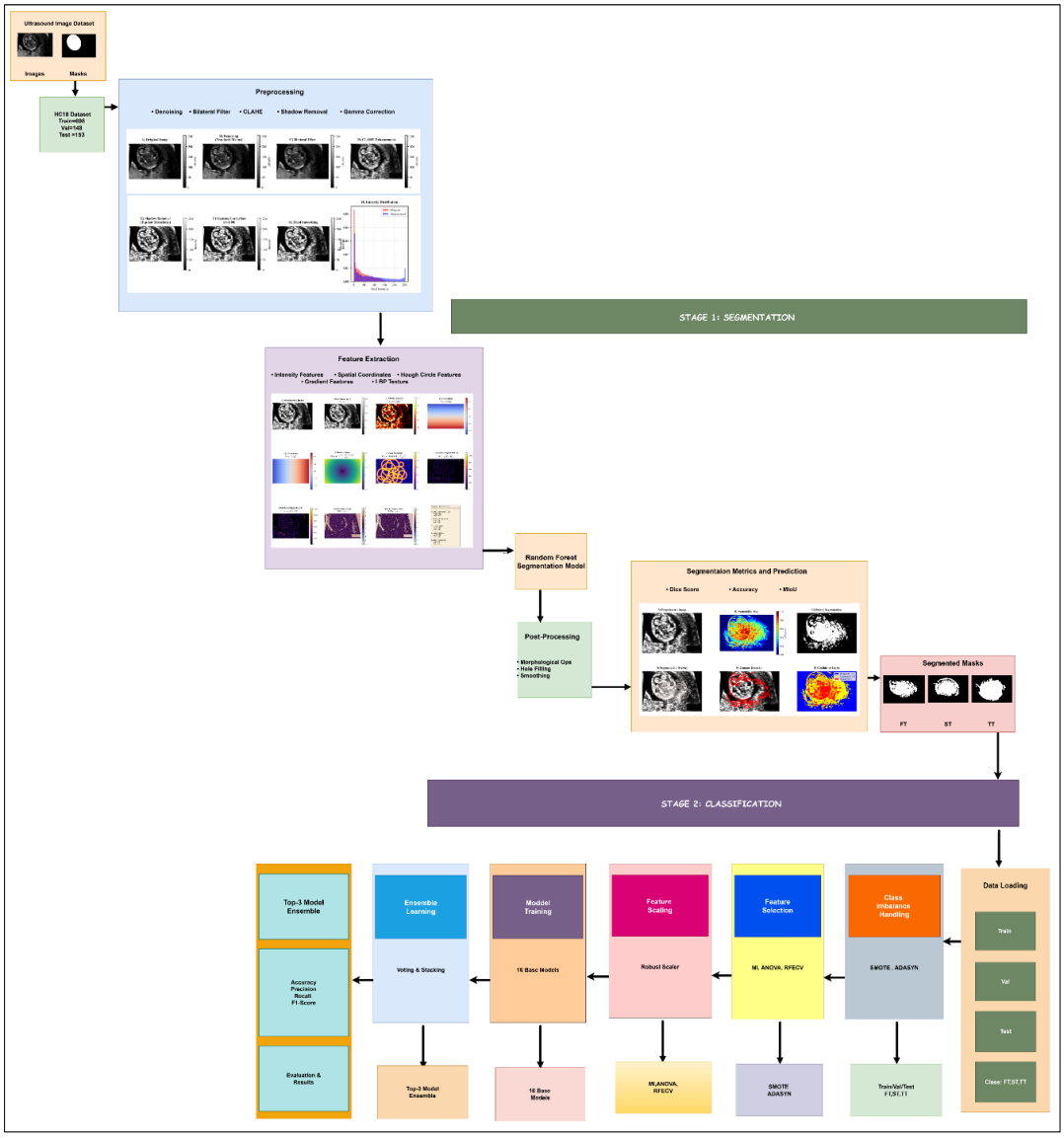

Figure 1. Proposed Two-Stage Pipeline Architecture

3.2 Data Preparation

The dataset consisted of grayscale fetal ultrasound images acquired during routine prenatal examinations. The images are categorized into three gestational age groups: First Trimester (FT, 1–13 weeks), Second Trimester (ST, 14–27 weeks), and Third Trimester (TT, 28–40 weeks)[28]. Each image is associated with a binary ground-truth mask delineating the fetal head circumference. Figure. 2 presents representative ultrasound images and their corresponding annotations (masks), arranged according to trimester-wise categories.

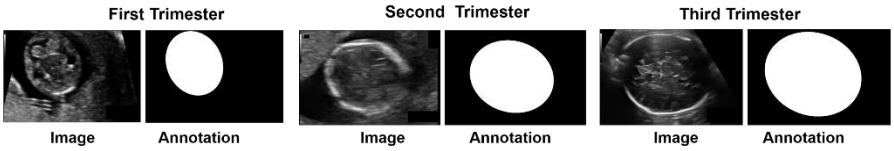

FIGURE. 2. Trimester-wise fetal ultrasound images with ground-truth annotations.

3.3 Preprocessing

Ultrasound images typically suffer from speckle noise, low contrast, and acoustic shading. To reduce these effects, a fixed preprocessing pipeline is applied to all images[29].

Let $I \in \mathbb{R}^{H\times W}$ denote a grayscale ultrasound image of height H and width W. The proposed framework maps the input image I to a binary segmentation mask. M and an associated trimester label y. Non-local means (NLM) denoising is applied to suppress speckle noise while preserving important edges. For a pixel p, the denoised intensity $I_d(p)$ is computed as using Equation (1).

$$I_d(p) = \sum_{q\in\Omega} w(p,q)I(q),\ \sum_{q\in\Omega} w(p,q) = 1 \quad (1)$$

In Equation (1), Ω denotes a local search window centered on the pixel p, and $w(p,q)$ represents similarity-based weights between pixels p and q. Next, bilateral filtering is applied to smooth the homogeneous regions while preserving the edges. The filtered intensity $I_b(p)$ is given by Equation (2).

$$I_b(p) = \frac{1}{W_p}\sum_{q\in\Omega} \exp\left(-\frac{\| p-q \|^2}{2\sigma_s^2}\right) exp\left(-\frac{|I(p)-I(q)^2}{2\sigma_r^2}\right) I(q) \quad (2)$$

Here, $\sigma_s$and σ_r control the spatial and intensity similarity, respectively, and W_p is a normalization factor.

Local contrast enhancement is performed using contrast-limited adaptive histogram equalization (CLAHE)[30]. Shadow artifacts are reduced using a white top-hat morphological transformation, as defined in Equation (3).

$$I_{\mathrm{TH}} = I - (I \circ B) \quad (3)$$

In Equation (3), ∘ denotes morphological opening, and B is a disk-shaped structuring element.

Finally, adaptive gamma correction is applied to normalize the image brightness. The gamma-corrected intensity $I_\gamma(p)$is expressed in Equation (4).

$$I_\gamma(p) = I(p)^\gamma \quad (4)$$

Where the gamma value γ is selected based on the mean image intensity to ensure consistent intensity distribution across the images.

The visual effect of each preprocessing step is illustrated in Figure 3. (A) shows the original ultrasound image. (B) and (C) demonstrate speckle noise reduction and edge-preserving smoothing achieved using non-local means denoising and bilateral filtering, respectively. (D) illustrates local contrast enhancement using CLAHE, while (E) highlights the reduction of acoustic shadow artifacts through top-hat transformation. (F) presents the gamma-corrected image, and (G) shows the final preprocessed image used for subsequent segmentation. (H) compares the intensity distributions before and after preprocessing, confirming improved contrast uniformity.

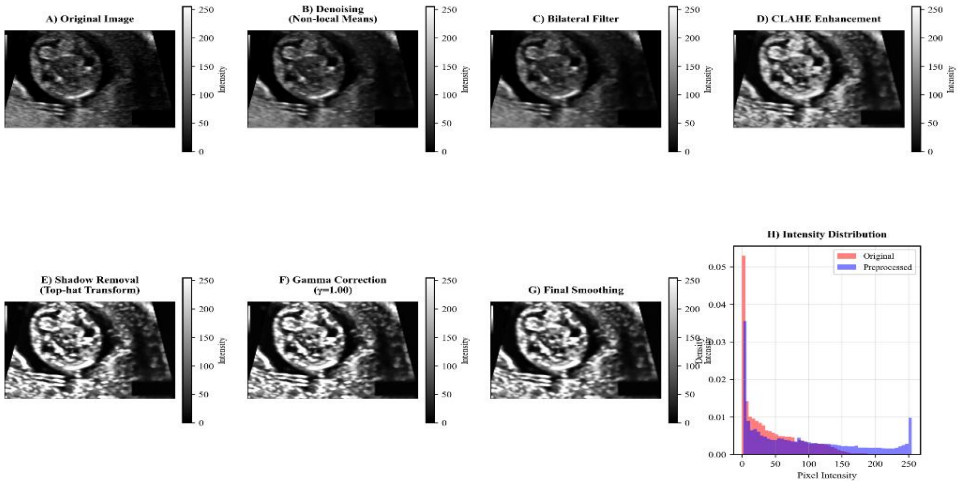

Figure 3. Illustration of the Ultrasound Image Preprocessing Pipeline

3.4 Fetal Head Segmentation

3.4.1 Feature Extraction for Segmentation

In this step, pixel-level features are extracted from each pre-processed image to characterize the appearance of the fetal head. Each pixel is represented using 11 pixel-level features. The intensity features include the grayscale pixel value and its squared value to capture nonlinear intensity relationships[31, 32]. The spatial features encoded the normalized horizontal and vertical distances from the image center, as well as the normalized radial distance. These features provide contextual information regarding the expected location of the fetal head. is defined in Equation (5).

$$f_x = \frac{x-x_c}{W}, f_y = \frac{y-y_c}{H} \quad (5)$$

Here, (x_c, y_c) represents the image's center.

The normalized radial distance feature is computed using Equation (6), as follows:

$$f_r = \frac{\sqrt{(x-x_c)^2+(y-y_c)^2}}{\sqrt{H^2+W^2}} \quad (6)$$

Edge information is captured using the Sobel gradient magnitude computed at multiple scales. It is calculated as follows (Equation (7)):

$$\mathrm{G}(p) = \sqrt{G_x(p)^2 + G_y(p)^2} \quad (7)$$

Texture information is extracted using rotation-invariant uniform Local Binary Patterns (LBP), which describe local texture variations around each pixel. The LBP code at pixel p is defined by Equation (8).

$$\mathrm{LBP}_{P,R} = \sum_{i=0}^{P-1} s(g_i - g_c)2^i, s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (8)$$

In Equation (8), g_c denotes the gray value of the central pixel, and g_i denotes the gray values of the neighboring pixels.

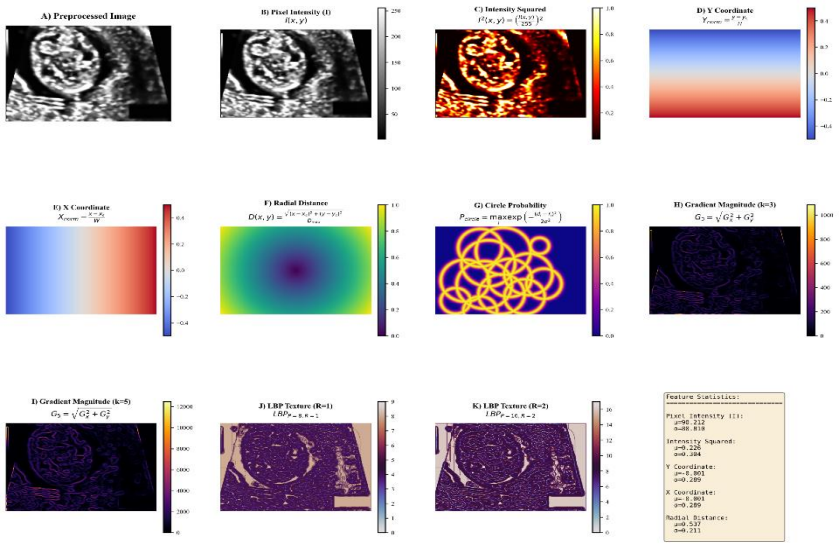

Figure 4. Intermediate Feature Maps Extracted from the Preprocessed Ultrasound Image for Fetal Head Segmentation.

Figure. 4 illustrates the intermediate feature representations extracted from the preprocessed ultrasound images before segmentation and classification. In addition to the original preprocessed image (A), pixel-level intensity transformations (B–C) are used to enhance contrast-related information. Spatial coordinate maps in the horizontal and vertical directions (D–E) explicitly encode positional priors, while the normalized radial distance map (F) captures the geometric relationship of pixels with respect to the image center. A probabilistic circular prior (G) highlights regions likely to correspond to the fetal head boundary. Edge-aware features are extracted using multi-scale gradient magnitude operators (H–I), and local texture characteristics are modeled using local binary pattern (LBP) descriptors at different radii (J–K). Together, these complementary intensity, spatial, geometric, edge, and texture features form a robust representation for accurate fetal head segmentation and trimester-wise classification.

3.4.2 Random Forest–Based Segmentation

Fetal head segmentation is formulated as a pixel-wise binary classification problem, in which each pixel is classified as either the fetal head or background. A Random Forest classifier is trained using the extracted pixel-level features[33].

To address the class imbalance and improve the contour accuracy, pixels are sampled from three regions: boundary pixels near the annotated head circumference, interior head pixels, and exterior background pixels[30]. Boundary pixels are assigned higher sampling weights to emphasize accurate head boundary detection. Each pixel label $\hat{y}_p$ is predicted using the Random Forest classifier as shown in Equation (9).

$$\hat{y}_p = f(\mathbf{f}_p), \hat{y}_p \in \{0,1\} \quad (9)$$

Here, $f(\cdot)$ represents the learned Random Forest model, where label 1 corresponds to the fetal head pixels and label 0 to background pixels.

3.4.3 Ellipse-Constrained Post-Processing

The fetal head typically exhibits an approximately elliptical shape in the ultrasound cross-sections. To enforce anatomical plausibility, the segmentation outputs are refined using geometry-based post-processing[19, 20].

First, small isolated regions are removed, and the largest connected component is retained for further analysis. An ellipse is then fitted to the segmented region using geometric properties. The general equation for the fitted ellipse is given by Equation (10).

$$\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1 \quad (10)$$

In Equation (11), (x_0, y_0) denotes the ellipse center, while a and b are the semi-major and semi-minor axes, respectively.

If the fitted ellipse closely matched the segmented region, the ellipse and original mask are blended to regularize the shape. Morphological closing and opening operations are subsequently applied to smooth the boundaries and fill small holes.

The final segmentation mask M_{final} is obtained by blending the fitted ellipse mask E with the original segmentation M as expressed in Equation (11).

$$M_{\text{final}} = 0.7E + 0.3M \quad (11)$$

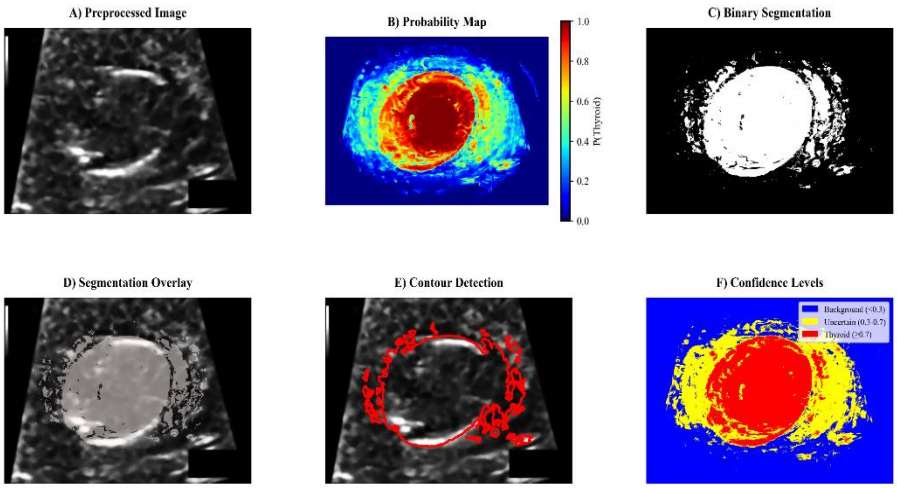

Figure. 5 Illustrates the Random Forest–Based Fetal Head Segmentation.

This post-processing step, to improve segmentation consistency while preserving anatomical variation, uses a Random Forest–based approach (Figure 5). The preprocessed ultrasound image (A) is used to generate a pixel-wise probability map (B), representing the likelihood of each pixel belonging to the fetal head region. Thresholding the probability map yielded a binary segmentation (C). The resulting mask is overlaid on the original image to assess spatial agreement (D). Contour detection is then applied to delineate the fetal head boundary for shape analysis (E). Finally, confidence-level visualization identified regions of high certainty and areas of boundary uncertainty (F).

3.5 Trimester Classification

This stage presents a structured multistage classification framework for automated trimester identification (FT, ST, and TT) based on quantitative features extracted from fetal ultrasound images[3, 10]. Class-imbalance handling, robust feature selection, feature normalization, and ensemble learning techniques are all included in the suggested pipeline to improve classification accuracy and generalization capacity. In the next subsections, the fetal ultrasound analysis and categorization procedures are described in detail.

3.5.1 Data Loading and Class Imbalance Handling

The fetal ultrasound dataset is first divided into subsets for training, validation, and testing. The test set is used for objective performance evaluation, the training set is utilized for model learning, and the validation set assists with hyperparameter tuning and model selection. The three trimester classes that are the focus of the classification task are FT, ST, and TT. In addition to reducing information loss, our methodical data-splitting technique guarantees accurate and repeatable performance evaluation. Class imbalance, in which the number of samples varies significantly between trimesters, is a prevalent problem with medical ultrasonography datasets. Classifiers may perform worse in underrepresented trimesters due to this imbalance, which can also bias them against the majority classes. In order to address this problem, only the training dataset is subjected to synthetic sampling approaches. By interpolating between preexisting minority class samples and their closest neighbors, the Synthetic Minority Over-Sampling Technique (SMOTE) creates new synthetic samples[30, 32]. A synthetic sample is generated as expressed in Equation (12)

$$x_{new} = x_i + \lambda(x_{nn} - x_i), \lambda \in (0,1) \quad (12)$$

where $x_i$is a minority class sample and x_{nn}denotes one of its k-nearest neighbors within the same class.

Further to improve learning in regions with higher classification difficulty, Adaptive Synthetic Sampling (ADASYN) is employed. ADASYN adaptively generates synthetic samples based on the local data distribution by assigning higher weights to minority samples that are surrounded by majority-class neighbors. The number of synthetic samples generated for each minority class instance is proportional to the local data distribution complexity, as defined in Equation (13).

$$G_i = \frac{r_i}{\sum_{j=1}^{N_m} r_j} \times G \quad (13)$$

where r_i represents the ratio of majority-class neighbors in the local neighborhood of the sample i, $N_m$is the number of minority samples, and G is the total number of synthetic samples to be generated.

By integrating SMOTE and ADASYN, a balanced training dataset is obtained while preserving the intrinsic structure of the feature space [32, 34]. This balanced representation enabled the classification models to learn trimester-specific fetal growth patterns more effectively and improved the generalization performance across all trimester classes.

3.5.2 Feature Selection

Following class imbalance correction, feature selection is performed to reduce the dimensionality and retain only the most discriminative features for fetal trimester classification. Because ultrasound-derived fetal head features may contain redundancy and noise, an ensemble feature selection strategy is adopted to ensure the robustness and stability of the selected feature subset. This process improves the model interpretability and reduces the risk of overfitting.

Let, $\mathbf{X} = [X_1, X_2, \dots, X_d]$

denote the feature set extracted from the segmented fetal head region, including shape, intensity, and texture descriptors, and let.

$$Y \in \{FT, ST, TT\}$$

Represents the corresponding trimester class labels.

3.5.3 Mutual Information-Based Feature Ranking

Mutual Information (MI) is employed to quantify the statistical dependency between each feature. $X_j$and the trimester class label Y. MI measures the amount of information a given feature

provides about trimester membership and is defined as shown in Equation (14).

$$MI(X_j;Y) = \sum_{x\in X_j} \sum_{y\in Y} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right) \quad (14)$$

Features with higher MI values are considered more informative for distinguishing between FT, ST, and TT and are ranked in accordance.

3.5.4 Analysis of Variance (ANOVA)

To further assess feature discriminability, an Analysis of Variance (ANOVA) is applied to identify features that exhibited statistically significant differences across trimesters[34]. For each feature X_j, the ANOVA F-statistic is computed as defined in Equation (15).

$$F_j = \frac{\text{Between-trimester variance}}{\text{Within-trimester variance}} \quad (15)$$

Features with higher F-scores indicate stronger separation among the three trimester classes and are retained for further analyses.

3.5.5 Recursive Feature Elimination with Cross-Validation (RFECV)

Recursive Feature Elimination with Cross-Validation (RFECV) is then employed to iteratively refine the feature subset. At each iteration, a classifier is trained using the current feature set, and feature importance scores are computed[34]. The least informative features are progressively removed, while cross-validation is used to identify the optimal number of features that maximizes the classification performance, as shown in Equation (16).

$$\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} \setminus \left\{ arg_j^{min} \left| w_j^{(k)} \right| \right\} \quad (16)$$

This procedure ensures the selection of a compact and highly informative feature subset.

3.5.5 Final Feature Subset Selection

The final feature set is obtained by combining the outputs of MI ranking, ANOVA filtering, and RFECV optimization as defined in Equation (17).

$$\mathbf{X}_{final} = \mathbf{X}_{MI} \cap \mathbf{X}_{ANOVA} \cap \mathbf{X}_{RFECV} \quad (17)$$

Only features that are consistently identified as informative across multiple selection criteria are retained [33, 35]. Reliable trimester classification based on fetal ultrasound data is ensured by this ensemble feature selection technique, which also increases robustness and improves generalization.

3.5.6 Feature Scaling and Multi-Model Training

To ensure numerical stability and consistent feature contribution, all selected features are normalized using Robust Scaling, which is perfect for medical ultrasonography data with non-Gaussian distributions and potential outliers. This scaling method reduces the impact of extreme numbers without compromising the data structure [9, 37]. 16 different machine learning classifiers are trained separately using the scaled feature set. Training several models at once improves the classification framework's overall discriminative performance for identifying the prenatal trimester and boosts model diversity by capturing a range of learning patterns and decision limits.

3.5.7 Ensemble Learning and Performance Evaluation

Ensemble learning strategies are employed to enhance robustness and reduce individual model bias. Voting ensembles combine predictions from multiple classifiers using majority or weighted voting, as defined in Equation (18).

$$\hat{y} = \arg\max_c \sum_{i=1}^{N} w_i \cdot \mathbb{I}(h_i(x) = c) \quad (18)$$

where h_i represents the i-th base classifier, w_i is its weight, and $\mathbb{I}(\cdot)$ is the indicator function. Additionally, stacking ensembles are employed, where predictions from the base models are supplied to a meta-learner that found the optimal collection of classifiers [5]. Based on their validation findings, the top three models are selected to form the final ensemble. Figure 6. Predicting Fetal Trimesters using the Top-3 Ensemble Classification Pipeline. The diagram displays the recommended ensemble architecture for automated fetal trimester categorization (FT, ST, TT) based on features gathered from the segmented fetal head area[9, 10]. It combines XGBoost classifiers, SVM with RBF kernel, and LightGBM. The pipeline begins with segmentation-derived standardized quantitative features, which are produced by utilizing a vote aggregator (majority or weighted voting) to combine the output from the three classifiers [11, 35]. An ensemble approach leverages the corresponding characteristics of different classifiers to improve robustness and classification accuracy.

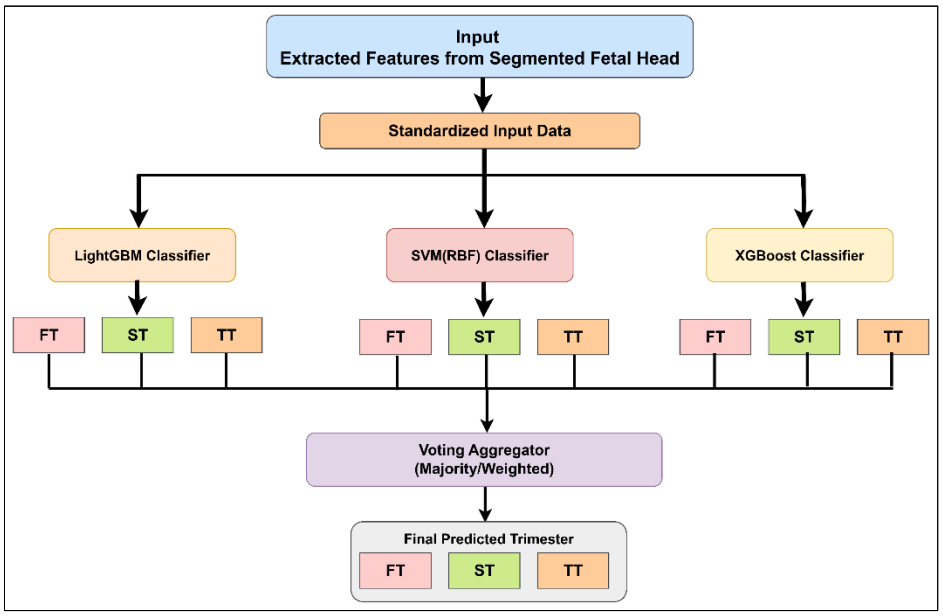

Figure 6. Top-3 Ensemble Framework for Fetal Trimester Classification.

An integrated method for automatic fetal trimester identification utilizing ultrasound images is offered by the suggested two-stage framework. By combining pixel-wise fetal head segmentation with a powerful feature-driven classification method, the model successfully identified gestational age-dependent anatomical variations. Combining multi-model ensemble learning, ensemble feature selection, and class-imbalance handling improves prediction accuracy and generalization [3]. All things considered, the proposed model offers a reliable and clinically meaningful approach for classifying trimesters in fetal ultrasound investigations.

4. Results and Discussion

4.1 Ablation Study

4.1.1 Segmentation Performance Analysis

In the segmentation stage, there are 999 fetal ultrasound images utilized. To preserve class balance, the dataset is split up by trimester. Training uses 70% of the images (698), validation uses 15% (148), and testing uses the remaining 15% (153). To assess the impact of various segmentation models, an ablation study is carried out. The segmentation performance of several models in terms of accuracy, mIoU, and dice score on the test set is displayed in Table 2. The K-means approach is found to perform the worst, suggesting that it is not very good at capturing intricate fetal head boundaries. Supervised models, such as SVM and XGBoost, are shown to achieve higher accuracy and Dice scores. The Random Forest model is found to achieve the best performance, with an accuracy of 89.27%, an mIoU of 78.12%, and a Dice score of 81.98%. These results indicate that Random Forest is selected for the subsequent stages of the proposed framework.

Table 2. Segmentation performance results (%) of different models evaluated on the test dataset.

Model	Accuracy	Dice Score	MIoU
K-means	69.51	57.26	50.92
SVM	81.23	79.50	6844
XG-Boost	82.88	79.89	68.87
Random Forest	**89.27**	**81.98**	**78.12**

Fetal head segmentation is performed using a Random Forest classifier on handcrafted pixel-level features, with ellipse-constrained post-processing to ensure anatomically consistent boundaries. Although the Dice score (~82%) is lower than state-of-the-art deep learning methods, this choice is justified due to the limited size of the HC18 dataset, which can cause overfitting in deep models. Random Forest reliably leverages handcrafted features while

producing consistent fetal head contours. The resulting segmentation provides a robust basis for extracting shape, intensity, and texture features for trimester-based classification.

4.1.2 Feature Selection Analysis

The feature selection analysis using Mutual Information and ANOVA F-test demonstrates that a limited subset of features contributes most of the discriminative power for trimester-based classification. Figure 7 illustrates the feature ranking, score distribution, and cumulative importance obtained using Mutual Information and ANOVA F-test methods. The ranking results show that the top features achieve higher relevance scores, while the remaining features contribute marginally. The score distribution and cumulative curves indicate that approximately 80–90% of the total information is captured using a reduced number of features. This confirms that dimensionality reduction is feasible without significant loss of information. The consistency observed between Mutual Information and ANOVA F-test further validates the robustness of the selected features. These selected features are subsequently used in the classification models, contributing to improved performance and better generalization.

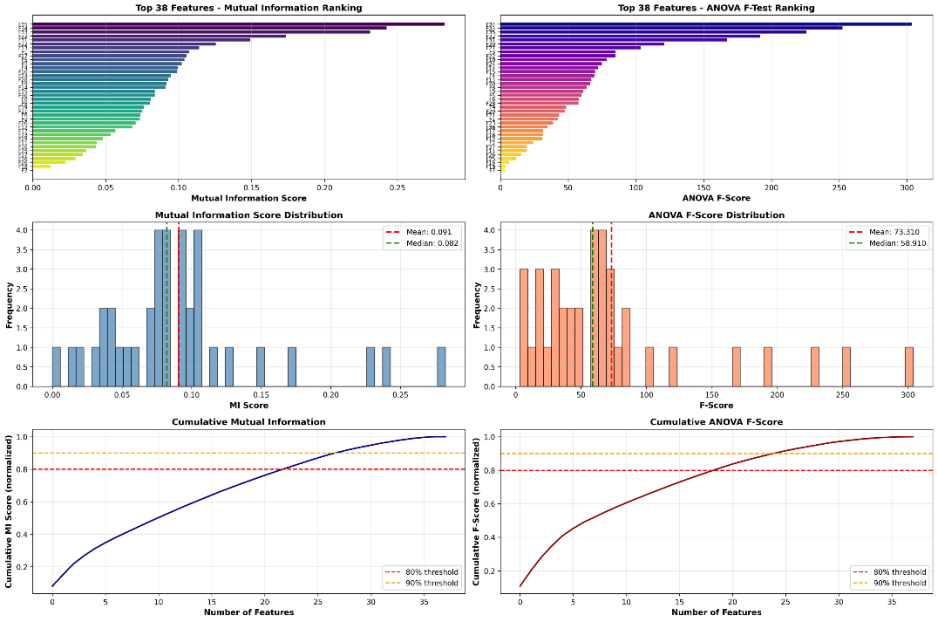

Figure 7. Mutual Information and Anova F-Test–Based Feature Ranking, Score Distribution, And Cumulative Importance Analysis for Feature Selection.

4.1.3 Classification Feature Extraction

The segmentation study demonstrates that the Random Forest method provides the most reliable performance among all evaluated models. Based on these results, the segmentation outputs generated by the Random Forest model are used as input for the subsequent classification stage. The dataset consists of 698 training images, 148 validation images, and 153 test images. The classification models are trained using features extracted from the segmented fetal head regions. Recursive Feature Elimination with Cross-Validation (RFECV) is employed to analyze the impact of feature selection on classification performance during the model development stage. Stratified 5-fold cross-validation is used to preserve class distribution across folds, and classification accuracy is reported as the mean and standard deviation over the validation folds. Table 3 presents the RFECV-based cross-validated performance comparison of individual and ensemble classifiers, while Figure. 8 provides a visual comparison of their corresponding cross-validation accuracies.

Classical linear models like QDA and LDA exhibit moderate accuracy, as seen by Table 3 and Figure. 8, demonstrating a limited capacity to describe complicated feature interactions. On the other hand, because they can capture non-linear relationships within the chosen feature space, tree-based and kernel-based models perform better. By combining complementary learners, ensemble techniques further improve robustness. With a cross-validated accuracy of 94.39% ± 0.81%, the Stacking Ensemble outperforms all other examined approaches, followed by SVM (RBF) and XGBoost. These findings demonstrate that ensemble learning significantly increases classification stability during training and that RFECV successfully discovers discriminative features.

Table 3. RFECV-based cross-validated accuracy of individual and ensemble classifiers.

Model	CV-Accuracy
QDA	0.8490 (±0.0229)
LDA	0.8765 (±0.0116)
Extra Trees	0.9183 (±0.0108)
SVM (Linear)	0.8970 (±0.0190)
Logistic Regression	0.8958 (±0.0131)
Random Forest	0.9195 (±0.0197)
Deep Neural Network	0.9208 (±0.0110)
Balanced Random Forest	0.9089 (±0.0126)
CatBoost	0.9201 (±0.0126)
Gradient Boosting	0.9183 (±0.0158)
Voting Ensemble	0.9314 (±0.0080)
XGBoost	0.9251 (±0.0148)
SVM (RBF)	0.9295 (±0.0160)
Weighted Voting	0.9282 (±0.0130)
Stacking Ensemble	0.9439 (±0.0081)
LightGBM	0.9257 (±0.0180)

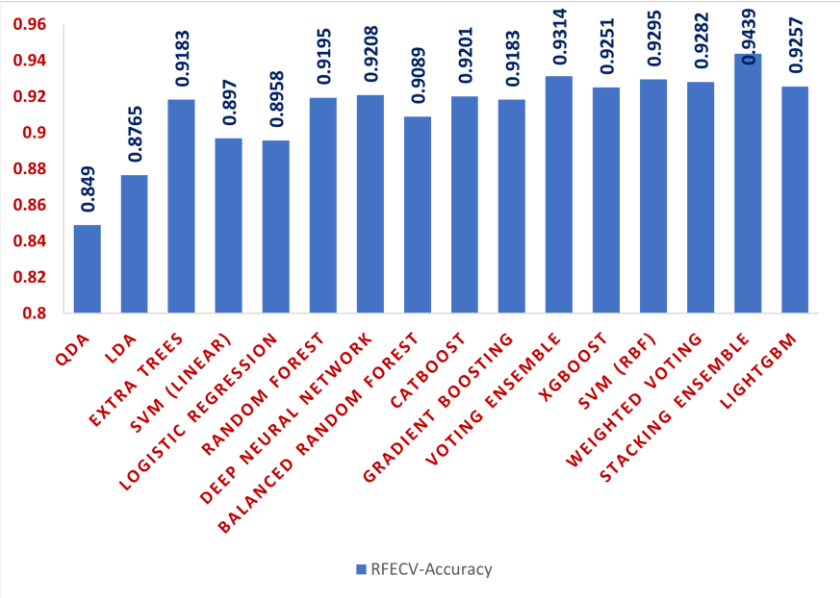

Figure 8. Cross-Validated Accuracy Comparison of Classifiers Using Rfecv.

4.1.4 Classification Model Comparison

An ablation study is conducted to analyse the contribution of different classification models using features obtained from the segmentation stage. Figure 9(a) presents the accuracy comparison of all evaluated classification models, while Figure 9(b) illustrates the multi-metric comparison of the top-performing models using accuracy, precision, recall, and F1-score. The corresponding quantitative results are summarized in Table 4. Linear models, such as QDA and LDA, achieve moderate performance, indicating limited effectiveness in capturing complex feature relationships. Individual machine learning models, including Random Forest, SVM, and gradient boosting methods, show improved results due to their ability to model non-linear patterns. Ensemble-based approaches further enhance classification performance by combining multiple learners. Among all evaluated models, the Top-3 Ensemble achieves the highest performance, with an accuracy of 85.62% and an F1-score of 85.90. These results demonstrate that ensemble learning provides a significant contribution to the robustness and accuracy of the proposed classification framework.

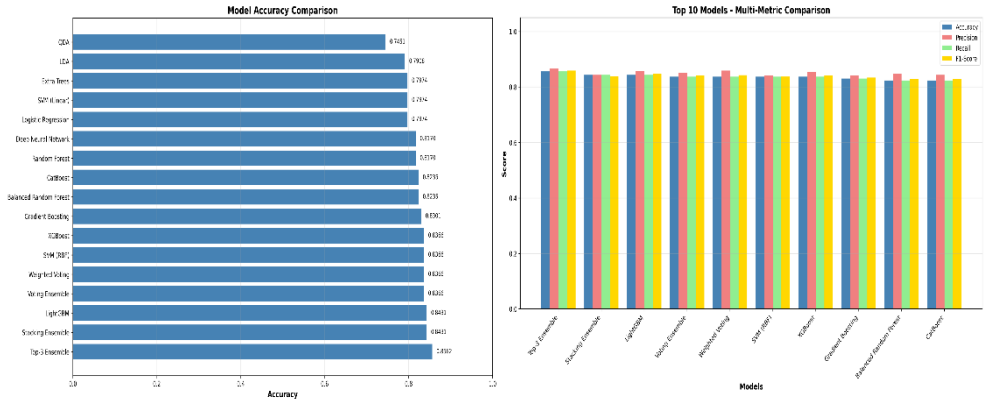

Figure 9. (A) Accuracy of Individual and Ensemble Classifiers. (B) Multi-Metric Performance of Top Models for Trimester Classification on the Test Set.

Table 4. Ablation Results (%) of Individual and Ensemble Models for Trimester Classification.

Model	Accuracy	Precision	Recall	F1-Score
QDA	74.51	81.84	74.51	76.12
LDA	79.08	82.46	79.08	79.92
Extra Trees	79.74	80.58	79.74	80.02
SVM (Linear)	79.74	84.59	79.74	80.52
Logistic Regression	79.74	85.19	79.74	80.68
Random Forest	81.7	82.36	81.7	81.92
Deep Neural Network	81.7	82.73	81.7	82.03
Balanced Random Forest	82.35	84.77	82.35	82.95
CatBoost	82.35	84.34	82.35	82.9
Gradient Boosting	83.01	84.15	83.01	83.4
Voting Ensemble	83.66	85.1	83.66	84.08
XGBoost	83.66	85.39	83.66	84.16
SVM (RBF)	83.66	84.16	83.66	83.75
Weighted Voting	83.66	85.92	83.66	84.27
Stacking Ensemble	84.31	84.32	84.31	83.88
LightGBM	84.31	85.8	84.31	84.76
Proposed: Top-3 Ensemble	**85.62**	**86.57**	**85.62**	**85.9**

4.2 Classification Results

The classification performance of the selected models is evaluated using accuracy, precision, recall, and F1-score. Table 5 shows the performance comparison of XGBoost, SVM (RBF), LightGBM, and the proposed Top-3 Ensemble model on the test set. The individual classifiers demonstrate strong and comparable results, with LightGBM achieving slightly higher performance among single models. The proposed Top-3 Ensemble, which integrates LightGBM, SVM (RBF), and XGBoost, achieves the best overall performance, with an accuracy of 85.62% and an F1-score of 85.90%. These findings indicate that ensemble learning improves classification robustness and enhances trimester-based prediction accuracy.

Table 5. Performance Results (%) of Individual Models and the Top-3 Ensemble for Trimester Classification.

Model	Accuracy	Precision	Recall	F1-Score
XGBoost	83.66	85.39	83.66	84.16
SVM (RBF)	83.66	84.16	83.66	83.75
LightGBM	84.31	85.8	84.31	84.76
Proposed: Top-3 Ensemble (LightGBM+ SVM (RBF)+ XGBoost)	**85.62**	**86.57**	**85.62**	**85.9**

Figure 10 illustrates the performance evaluation of the proposed Top-3 ensemble model for fetal trimester classification through two complementary analyses: (a) the confusion matrix and (b) Receiver Operating Characteristic (ROC) curves with corresponding Area Under the Curve (AUC) values. The first trimester (FT), second trimester (ST), and third trimester (TT) classes' properly and incorrectly identified cases are highlighted in the confusion matrix, which offers a class-wise assessment of prediction accuracy. The suggested classification framework's robustness, dependability, and clinical relevance are validated by high AUC values and excellent agreement in the confusion matrix.

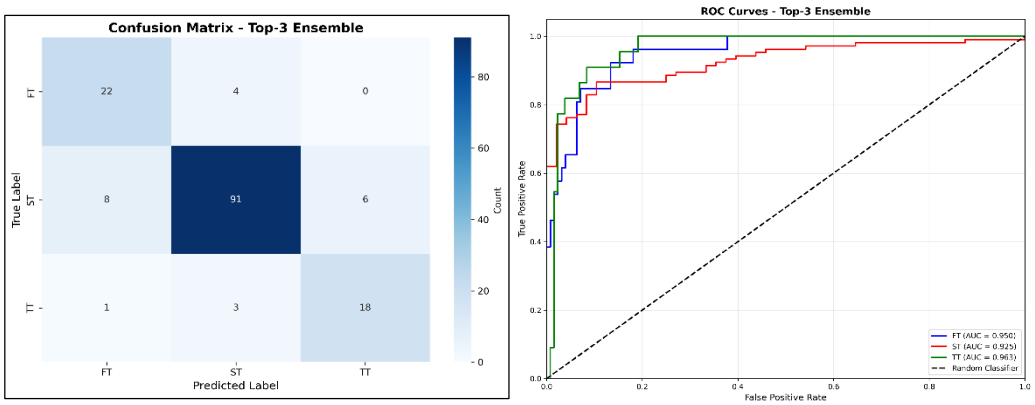

Figure 10. Confusion Matrix (a) And ROC Curves with AUC (b) For the Top-3 Ensemble in Fetal Trimester Classification.

4.3 Comparative Analysis

On the HC18 dataset, Table 6 contrasts the suggested Top-3 ensemble method with new trimester-based fetal ultrasound categorization techniques. Using MobileNet and a multi-input DenseNet121 model, the authors reported accuracy rates of 79.42% and 86.68%, respectively. DenseNet121 takes significant processing resources and extensive deep learning training, although achieving somewhat greater accuracy. The suggested ensemble strategy, on the other hand, uses handmade features taken from segmented fetal head regions to attain a competitive accuracy of 85.62%. This indicates that the suggested approach offers a good trade-off between computational efficiency, interpretability, and classification performance. The findings demonstrate that reliable feature-based ensemble models can be useful substitutes for intricate deep learning architectures, especially in clinical contexts with constrained data or processing power.

Table 6. Accuracy Comparison of Fetal Trimester Classification Methods on the HC18 Dataset.

Methodology	Dataset	Results
MobileNet [10]	HC18 dataset	79.42%
DenseNet121 (Multi-Input Classification)[3]	HC18 dataset	86.68%
Proposed Method: Top-3 Ensemble (LightGBM+ SVM (RBF)+ XGBoost)	**HC18 dataset**	**85.62%**

4.4 Discussion

The proposed two-stage framework effectively performs fetal head segmentation and trimester classification from ultrasound images. In segmentation, it achieves competitive accuracy compared with deep learning methods such as U-Net, DAG V-Net, and SAP-Net (DSC 97.5–98.74%) while requiring lower computational resources and offering better interpretability [3, 10]. While regression-based methods estimate gestational age in days, they are highly sensitive to annotation errors and inter-observer variability, particularly in ultrasound imaging. In contrast, trimester-based classification groups gestational age into clinically meaningful categories, reducing the impact of labeling noise and improving model robustness. This approach also simplifies interpretation and integration into clinical workflows, providing reliable information for decision-making without requiring precise day-level estimates [14, 35, 36].

Trimester classification on the HC18 dataset, MobileNet, and DenseNet121 achieve 79.42% and 86.68% accuracy, respectively, whereas the proposed Top-3 ensemble model attains 85.62%, demonstrating that feature-based ensemble learning can reliably distinguish first, second, and third trimesters with efficiency suitable for clinical deployment. Although deep learning–based segmentation methods generally achieve higher accuracy, they require large annotated datasets and significant computational resources. In contrast, the proposed Random Forest–based segmentation provides reliable performance on limited data with lower computational cost, at the expense of a moderate reduction in segmentation accuracy, which is acceptable as segmentation primarily serves as a preprocessing step for classification[11, 12].Overall, the framework provides a practical balance of segmentation precision, classification performance, and interpretability, highlighting the potential of hybrid feature-based ensemble approaches as robust and efficient alternatives to computationally intensive deep learning models in prenatal assessment. These results suggest that the proposed method can support clinicians by providing an automated, consistent, and reliable tool for fetal gestational stage evaluation.

4.5 Limitations

The HC18 fetal ultrasound dataset, which is obtained utilizing certain imaging settings and equipment, is used to assess the suggested framework. The results may therefore not apply to ultrasound images from other devices or medical environments. Frequent ultrasound artifacts and image noise can affect the Random Forest classifier and manually created features utilized in the fetal head segmentation technique. Moreover, the classification stage uses specified features from the divided fetal head area rather than explicitly modeling fetal growth over time.

4.5.1 Future Multi-Center Validation and Device Variability

The proposed framework demonstrates robust performance on the HC18 dataset; future work will focus on validating the method across multi-center datasets acquired using different ultrasound devices and imaging protocols. Such an evaluation will help assess the generalizability of the model under variations in scanner hardware, acquisition settings, and population characteristics. Incorporating data from diverse clinical centers will further strengthen the reliability of the proposed framework for real-world deployment.

5. Conclusion and Future Work

This paper presents a reliable two-stage approach for automated fetal trimester identification from ultrasound images by combining ensemble-based classification with fetal head segmentation. In the first phase, the fetal head region is carefully defined using a Random Forest-based pixel-wise segmentation technique, which enables the reliable extraction of anatomically important characteristics. In the second stage, a comprehensive classification pipeline, which includes class imbalance handling, ensemble feature selection, feature normalization, and a Top-3 ensemble model (LightGBM, SVM with RBF kernel, and XGBoost), effectively captures gestational age-dependent differences across FT, ST, and TT classes. The confusion matrix and ROC-AUC analysis support the experimental results, which show strong classification accuracy and robustness.

The clinical perception, the proposed framework can assist obstetricians with prenatal screening and gestational age assessment by offering an objective, fully automated technique for determining the fetal trimester. By reducing reliance on manual measures and operator competence, the system has the potential to improve workflow efficiency and consistency in clinical decision-making, particularly in settings with limited resources. Accurately determining the trimester can help with early discovery of abnormal fetal growth patterns and timely clinical interventions.

In order to improve generality, future research will concentrate on expanding the study to larger and more diverse multi-center ultrasound datasets. Combining handcrafted features with deep learning-based feature representations may enhance discriminative performance even more. Regression-based gestational age estimate and longitudinal analysis should be studied in order to provide continuous age prediction rather than discrete trimester categorization. The therapeutic utility of the proposed strategy would be further enhanced by these modifications.

Acknowledgment

The authors thank KSTEPS, the DST, and the Government of Karnataka for their financial help and for giving them a Ph.D. fellowship, which made this research possible. We also want to thank Dr. Vanita B. Metgud and Dr. Satwik B. Metgud from Metgud Hospital - Advanced Laparoscopy Centre and IVF in Belagavi, Karnataka, India, for giving us the ultrasound fetal images and sharing their knowledge of fetal development, which was very helpful to this study.

References

1. Kumar, R.R., Priyadarshi, R.: Denoising and segmentation in medical image analysis: A comprehensive review on machine learning and deep learning approaches. Multimed Tools Appl. 84, 10817–10875 (2025). https://doi.org/10.1007/s11042-024-19313-6.
2. Govindarajan, M.P., Bharathi, S.S.K.: Automatic Fetal Segmentation Designed on Computer-Aided Detection with Ultrasound Images. Computers, Materials and Continua. 81, 2967–2986 (2024). https://doi.org/10.32604/cmc.2024.055536.
3. Gornale, S.S., Kamat, P.C., Hiremath, P.S., Siddalingappa, R.: A Hybrid Ensemble of Denoising Autoencoders and Deep Learning Models for Fetal Image Analysis. Cureus Journal of Computer Science. (2025). https://doi.org/10.7759/s44389-025-09506-x.
4. Alzubaidi, M., Agus, M., Shah, U., Makhlouf, M., Alyafei, K., Househ, M.: Ensemble Transfer Learning for Fetal Head Analysis: From Segmentation to Gestational Age and Weight Prediction. Diagnostics. 12, (2022). https://doi.org/10.3390/diagnostics12092229.
5. Faska, Z., Khrissi, L., Haddouch, K., El Akkad, N.: A robust and consistent stack generalized ensemble-learning framework for image segmentation. Journal of Engineering and Applied Science. 70, (2023). https://doi.org/10.1186/s44147-023-00226-4.
6. Sujan Hiregundagal Gopal Rao. (2023). Safety and Security Co-Design in Automotive Semiconductor Systems: Challenges and Future Directions. International Journal of Intelligent Systems and Applications in Engineering, 12(4s), 830–834.
7. Kim, H.P., Lee, S.M., Kwon, J.-Y., Park, Y., Kim, K.C., Seo, J.K.: Automatic evaluation of fetal head biometry from ultrasound images using machine learning. (2018). https://doi.org/10.1088/1361-6579/ab21ac.
8. Yousefpour Shahrivar, R., Karami, F., Karami, E.: Enhancing Fetal Anomaly Detection in Ultrasonography Images: A Review of Machine Learning-Based Approaches, (2023). https://doi.org/10.3390/biomimetics8070519.
9. Zhang, H.W., Wang, Y.R., Hu, B., Song, B., Wen, Z.J., Su, L., Chen, X.M., Wang, X., Zhou, P., Zhong, X.M., Pang, H.W., Wang, Y.H.: Using machine learning to develop a stacking ensemble learning model for the CT radiomics classification of brain metastases. Sci Rep. 14, (2024). https://doi.org/10.1038/s41598-024-80210-x.
10. Gornale, S.S., Kamat, P., Hiremath, P.S., Kumar, S., Goh, K.W.: Automated Segmentation and Trimester-Based Classification of Fetal Head Circumference in Ultrasound Images Using Deep Learning Techniques. Intern J Pattern Recognit Artif Intell. 0, null. https://doi.org/10.1142/S021800142552038X.
11. Mukadam, S.B., Patil, H.Y.: Machine Learning and Computer Vision Based Methods for Cancer Classification: A Systematic Review, (2024). https://doi.org/10.1007/s11831-024-10065-y.
12. Attallah, O., Gadelkarim, H., Sharkas, M.A.: Detecting and Classifying Fetal Brain Abnormalities Using Machine Learning Techniques. Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018. 1371–1376 (2019). https://doi.org/10.1109/ICMLA.2018.00223.
13. Regmi, B., Shah, C.: Classification Methods Based on Machine Learning for the Analysis of Fetal Health Data. (2023).
14. Lee, L.H., Bradburn, E., Craik, R., Yaqub, M., Norris, S.A., Ismail, L.C., Ohuma, E.O., Barros, F.C., Lambert, A., Carvalho, M., Jaffer, Y.A., Gravett, M., Purwar, M., Wu, Q., Bertino, E., Munim, S., Min, A.M., Bhutta, Z., Villar, J., Kennedy, S.H., Noble, J.A., Papageorghiou, A.T.: Machine learning for accurate estimation of fetal gestational age based on ultrasound images. NPJ Digit Med. 6, (2023). https://doi.org/10.1038/s41746-023-00774-2.
15. Carneiro, G., Georgescu, B., Good, S., Comaniciu, D.: Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree. IEEE Trans Med Imaging. 27, 1342–1355 (2008). https://doi.org/10.1109/TMI.2008.928917.
16. Rawat, V., Jain, A., Shrimali, V., Raghuvanshi, S.: Performance Analysis of Different Learning Algorithms of Feed Forward Neural Network Regarding Fetal Abnormality Detection. Springer International Publishing (2018). https://doi.org/10.1007/978-3-319-99810-7_6.
17. Gornale, S., Kamat, P., Siddalingappa, R., Kumar, S.: Deep Learning Techniques for a Comprehensive Analysis of Fetal Biometric Parameters Across Trimesters. Transactions on Machine Learning and Artificial Intelligence. 12, 18–45 (2024). https://doi.org/10.14738/tecs.123.16985.
18. Zhang, J., Petitjean, C., Ainouz, S.: Segmentation-Based vs. Regression-Based Biomarker Estimation: A Case Study of Fetus Head Circumference Assessment from Ultrasound Images. J Imaging. 8, (2022). https://doi.org/10.3390/jimaging8020023.
19. Torres, H.R., Oliveira, B., Morais, P.R., Fritze, A., Birdir, C., Rüdiger, M., Fonseca, J.C., Vilaça, J.L.: Fetal head circumference delineation using convolutional neural networks with registration-based ellipse fitting. In: Proc.SPIE. p.

120323L (2022). https://doi.org/10.1117/12.2611150.
20. Wang, J., Fang, Z., Yao, S., Yang, F.: Ellipse guided multi-task network for fetal head circumference measurement. Biomed Signal Process Control. 82, 104535 (2023). https://doi.org/https://doi.org/10.1016/j.bspc.2022.104535.
21. Zeng, Y., Tsui, P.H., Wu, W., Zhou, Z., Wu, S.: Fetal Ultrasound Image Segmentation for Automatic Head Circumference Biometry Using Deeply Supervised Attention-Gated V-Net. J Digit Imaging. 34, 134–148 (2021). https://doi.org/10.1007/s10278-020-00410-5.
22. Li, P., Zhao, H., Liu, P., Cao, F.: Automated measurement network for accurate segmentation and parameter modification in fetal head ultrasound images. Med Biol Eng Comput. 58, 2879–2892 (2020). https://doi.org/10.1007/s11517-020-02242-5.
23. Fiorentino, M.C., Moccia, S., Capparuccini, M., Giamberini, S., Frontoni, E.: A regression framework to head-circumference delineation from US fetal images. Comput Methods Programs Biomed. 198, (2021). https://doi.org/10.1016/j.cmpb.2020.105771.
24. Al-Bander, B., Alzahrani, T., Alzahrani, S., Williams, B.M., Zheng, Y.: Improving Fetal Head Contour Detection by Object Localisation with Deep Learning BT - Medical Image Understanding and Analysis. Presented at the (2020).
25. Fiorentino, M.C., Villani, F.P., Di Cosmo, M., Frontoni, E., Moccia, S.: A Review on Deep-Learning Algorithms for Fetal Ultrasound-Image Analysis. (2022). https://doi.org/10.1016/j.media.2022.102629.
26. Włodarczyk, T., Płotka, S., Szczepański, T., Rokita, P., Sochacki-Wójcicka, N., Wójcicki, J., Lipa, M., Trzciński, T.: Machine learning methods for preterm birth prediction: A review. Electronics (Switzerland). 10, 1–24 (2021). https://doi.org/10.3390/electronics10050586.
27. van den Heuvel, T.L.A., de Bruijn, D., de Korte, C.L., van Ginneken, B.: Automated measurement of fetal head circumference using 2D ultrasound images. PLoS One. 13, (2018). https://doi.org/10.1371/journal.pone.0200412.
28. Kiserud, T., Piaggio, G., Carroli, G., Widmer, M., Carvalho, J., Neerup Jensen, L.: The World Health Organization Fetal Growth Charts: A Multinational Longitudinal Study of Ultrasound Biometric Measurements and Estimated Fetal Weight. PLoS Med. 14, (2017). https://doi.org/10.1371/journal.pmed.1002220.
29. Gornale, S.S., Patravali, P.U., Hiremath, P.S.: Osteoarthritis Detection in Knee Radiographic Images Using Multiresolution Wavelet Filters. In: Communications in Computer and Information Science. pp. 36–49. Springer Science and Business Media Deutschland GmbH (2021). https://doi.org/10.1007/978-981-16-0493-5_4.
30. Liu, X., Guo, L., Wang, H., Guo, J., Yang, S., Duan, L.: Research on imbalance machine learning methods for MR T1 WI soft tissue sarcoma data. BMC Med Imaging. 22, (2022). https://doi.org/10.1186/s12880-022-00876-5.
31. Ganie, S.M., Pramanik, P.K.D., Zhao, Z.: Ensemble learning with explainable AI for improved heart disease prediction based on multiple datasets. Sci Rep. 15, (2025). https://doi.org/10.1038/s41598-025-97547-6.
32. Chhillar, I., Singh, A.: An improved soft voting-based machine learning technique to detect breast cancer utilizing effective feature selection and SMOTE-ENN class balancing. Discover Artificial Intelligence. 5, (2025). https://doi.org/10.1007/s44163-025-00224-w.
33. Kuhle, S., Maguire, B., Zhang, H., Hamilton, D., Allen, A.C., Joseph, K.S., Allen, V.M.: Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: A retrospective cohort study. BMC Pregnancy Childbirth. 18, 1–9 (2018). https://doi.org/10.1186/s12884-018-1971-2.
34. Epifano, J.R., Tripathi, A., Silvestri, A., Yu, A., Lee, H., Ramachandran, R.P.: A Comparison of Feature Selection Techniques for First-day Mortality Prediction in the ICU Ghulam Rasool.
35. Fung, R., Villar, J., Dashti, A., Cheikh Ismail, L., Staines-Urias, E., Ohuma, E.O., Salomon, L.J., Victora, C.G., Barros, F.C., Lambert, A., Carvalho, M., Jaffer, Y.A.: Achieving accurate estimates of fetal gestational age and personalised predictions of fetal growth based on data from an international prospective cohort study: a population-based machine learning study. (2020).
36. Naz, S., Noorani, S., Jaffar Zaidi, S.A., Rahman, A.R., Sattar, S., Das, J.K., Hoodbhoy, Z.: Use of artificial intelligence for gestational age estimation: a systematic review and meta-analysis, (2025). https://doi.org/10.3389/fgwh.2025.1447579.
37. Savakar, D.G., Telsang, D., Kannur, A.: Artificial Neural Network for Identification and Classification of Natural Body Marks. In: Lecture Notes in Electrical Engineering. pp. 541–555. Springer Science and Business Media Deutschland GmbH (2021). https://doi.org/10.1007/978-981-33-4909-4_41.