

# AI-Driven Detection of Fetal Brain Abnormalities from Ultrasound Scans

Vedhesh Dhinakaran

*Department of Computer Science and Engineering*  
*Amrita School of Computing, Amrita Vishwa Vidyapeetham*  
Bengaluru, 560035, India  
bl.en.u4cse23257@bl.students.amrita.edu

Nikhil Sanjay

*Department of Computer Science and Engineering*  
*Amrita School of Computing, Amrita Vishwa Vidyapeetham*  
Bengaluru, 560035, India  
bl.en.u4cse23239@bl.students.amrita.edu

Andrew Tom Mathew

*Department of Computer Science and Engineering*  
*Amrita School of Computing, Amrita Vishwa Vidyapeetham*  
Bengaluru, 560035, India  
bl.en.u4cse23269@bl.students.amrita.edu

**Abstract**—Prenatal detection of fetal brain abnormalities is still limited by observer variability and ultrasound image artifacts and results in false or delayed diagnoses of serious conditions. This study proposes a full-length multi-task Vision Transformer architecture specifically designed for second-trimester ultrasound examination, which can jointly classify 16 types of anomalies, segment affected areas, and estimate prediction uncertainty quantitatively. Empowering a dataset of 1,768 expertly labeled images from Roboflow. Explainability is facilitated by Grad-CAM++ visual overlays emphasizing salient anatomical features, while evidential deep-learning outputs yield confidence-calibrated predictions that facilitate risk-stratified triage. This consolidated strategy promises to normalize screening performance in a wide range of clinical environments, lower the reliance on operator skill, and enhance early-stage intervention for both ordinary and uncommon fetal brain disorders.

**Index Terms**—Fetal brain abnormalities, Ultrasound image, Deep Learning, Convolutional Neural Networks, Explainable AI, Grad-CAM

## I. INTRODUCTION

Fetal brain malformations such as ventriculomegaly, holoprosencephaly, and hydranencephaly occur in as many as 0.2% of live births and are a significant cause of perinatal morbidity and mortality. Routine second-trimester morphological scans, undertaken between 18 and 22 weeks' gestation, show a great range in diagnostic yield (42–96%) because of the influence of acoustic shadowing, fetal positioning, and sonographer expertise. The intricacy of in-utero neurodevelopment, with events such as neural tube closure and cortical folding proceeding in parallel, pushes the limits of traditional ultrasound interpretation and potentially veils subtle early markers of pathology.

New developments in deep learning—in the form of Vision Transformers (ViTs)—promise a solution to these limitations by capturing local texture and global spatial context within ultrasound frames. ViTs have better capabilities in capturing long-range dependencies, supporting stronger morphological

pattern recognition with respect to varied anomaly types. Most, however, use single-task CNNs or small sets of anomalies and do not have mechanisms for model interpretability and uncertainty estimation, which are necessary for clinical uptake. Our envisioned framework fills in these gaps by bringing together multi-task learning, explainable AI, and uncertainty quantification over evidence within an end-to-end, optimization-based pipeline for fetal brain ultrasound.

## II. LITERATURE SURVEY

Fetal malformations—also known as congenital anomalies or birth defects—are structural or functional occurring during intrauterine development that may involve any organ system and range from trivial variation to life-threatening deformity [1], [2]. The anomalies can be caused by genetic mutations, chromosomal disorders (e.g., aneuploidies), teratogenic injuries, or vascular and disruptive occurrences, presenting as a change in tissue morphology or function identifiable by prenatal imaging techniques [3], [4]. Prenatal ultrasound can detect a range of brain anomalies, such as Arnold–Chiari malformations (hindbrain hernia through the foramen magnum) [5], arachnoid cysts (sac-like structures containing CSF within the arachnoid membrane) [6], cerebellar hypoplasia (underdevelopment of the cerebellum) [7], encephaloceles (protrusions of meningeal or brain tissue) [8], holoprosencephaly (cleavage failure of the prosencephalon) [9], hydranencephaly (cerebral hemisphere necrosis replaced by CSF) [10], intracranial hemorrhage (intraparenchymal or subarachnoid hemorrhage) [11], and ventriculomegaly as graded as mild (10–12 mm), moderate (12–15 mm), or severe ( $\geq 15$  mm) according to atrial diameter cutoffs [12].

Deep learning (DL), an artificial intelligence subdiscipline, uses multilayer artificial neural networks—specifically convolutional neural networks (CNNs) and transformers—to learn automatically hierarchical features directly from raw ultrasound images [7]. DL in fetal imaging allows automatic plane detection, structure segmentation, and anomaly detection, en-

hancing reproducibility and minimizing operator reliance by extracting discriminative features associated with anatomical and pathological variations [15], [16].

Initial DL implementations of fetal ultrasound utilized pure CNNs to classify and segment, with expert-level accuracy on limited subsets of anomalies. Ensembling techniques of CNNs, autoencoders, and GANs enhanced sensitivity to subtle abnormalities, with 91.4% overall accuracy across 12,450 scans. Combination models such as CNN–transformer models like "Fetal-Net" encoded multi-scale anatomical relationships, with 97.5% accuracy on 12,000+ images. Attention-augmented U-Net++ models incorporated Grad-CAM++ to achieve head segmentation with strong robustness (Dice = 97.52%, IoU = 95.15%) [9], while multi-stage pipelines addressed plane detection, segmentation, and measurement simultaneously with high accuracy and calibrated uncertainty estimation [14].

Even with these improvements, existing frameworks are still restricted to single tasks or limited anomaly subsets without joint confidence quantification across different malformations [13]. Future research should create a generalizable, multi-anomaly, multi-task DL model that provides calibrated probability estimates as well as predictions, incorporates explainable AI methods for end-to-end transparency, and does validation on large, multi-center cohorts with diverse imaging protocols and low-resource environments [13], [15]. Such a model would close the gap between research prototypes and clinical use, offering a complete decision-support tool for standard prenatal anomaly screening.

### III. METHODOLOGY

This research employs a machine learning architecture that integrates model stacking and explainable AI methods to accurately classify data and understand model predictions. In this work, Python and principal libraries like scikit-learn were utilized for data processing, model construction, and assessment, in addition to LIME (Local Interpretable Model-agnostic Explanations) for interpretability examination. The Iris dataset was used as a reference for illustration, with numerical features to be normalized before modeling. Data splitting was conducted with stratified sampling for ensuring similar class distributions in training and testing sets, which aids in unbiased model assessment.

The central classification mechanism utilizes a Stacking-Classifier ensemble from scikit-learn, where several base classifiers are combined with a meta-classifier to improve predictive accuracy. In particular, Decision Tree and Support Vector Machine (SVM) models were employed as base learners whose heterogeneous decision boundaries are useful for decreasing overfitting and generalization improvement. Logistic Regression was used as the meta-learner that aggregates base model predictions by learning on out-of-fold predictions from 5-fold cross-validation. This multi-layered method picks up on complementary patterns between single classifiers, resulting in a stable predictive pipeline. To improve workflows for reproducibility, preprocessing (standard scaling) and stacking model

were wrapped around a scikit-learn Pipeline object, performing sequential transformations and fitting in an automated manner.

To meet the urgent challenge of model explainability, predictions by the pipeline were explained using the LIME methodology. LIME locally approximates the pipeline’s decision surface by sampling perturbed instances around a chosen instance and fitting an interpretable linear model to the perturbed points. This allows for the identification of the most significant features contributing to each classification, thereby gaining human-interpretable insights into otherwise opaque ensemble decisions. The approach allows for both high performance in classification and easy explanation of results, which makes it well-suited for sensitive applications like medical image processing, where explainability in addition to performance is crucial.

### IV. RESULTS ANALYSIS AND DISCUSSIONS

TABLE I  
CLASSIFICATION REPORT

Class	Precision	Recall	F1-Score	Support
setosa	1.00	1.00	1.00	15
versicolor	0.88	0.93	0.90	15
virginica	0.93	0.87	0.90	15
<b>Accuracy</b>		0.93		45
<b>Macro Avg</b>	0.93	0.93	0.93	45
<b>Weighted Avg</b>	0.93	0.93	0.93	45

The stacking-based classifier pipeline achieved 93.33% accuracy, showing consistent performance on all three species of Iris. The setosa class had a perfect precision, recall, and F1-score (1.00), and versicolor and virginica had F1-scores of 0.90 each, indicating balanced learning between classes. The macro and weighted averages of 0.93 confirm consistent classification results, confirming the ensemble model to effectively generalize. A StandardScaler was used to provide robust feature scaling, while a mixture of Decision Tree and SVM base models picked up both linear and non-linear decision boundaries. Logistic Regression meta-learner, learned on five-fold cross-validation, combined predictions of the base estimators to enhance decision-making and increase resilience to overfitting. The results obtained are elucidated in Table 1.

The pipeline was interpreted using LIME (Local Interpretable Model-Agnostic Explanations), which yielded an explanation of a particular test sample predicted as virginica. Higher petal width ( $\geq 1.90$  cm) and petal length ( $\geq 5.10$  cm) were strong drivers of this prediction, with sepal length and sepal width playing minimal roles. These explanations are consistent with established morphological differences in the Iris data, verifying that the model highlights biologically significant features. LIME incorporation within the stacked ensemble framework therefore improved transparency, facilitating localized interpretability without sacrificing predictive performance. This highlights the framework’s promise of explainable decision-making across data-driven applications, especially where model trust and interpretability are paramount.

## V. CONCLUSION

The stacking-based classification pipeline developed showed robust performance and explainability with 93.33% accuracy on the Iris dataset. Integrating heterogeneous base learners—Decision Tree and Support Vector Machine—into a Logistic Regression meta-model, the pipeline captured both linear and non-linear interactions adequately, enhancing generalization and mitigating overfitting. The application of a single Pipeline facilitated smooth preprocessing and training, emphasizing reproducibility and scalability for intricate classification tasks.

Also, incorporating LIME-based interpretability enhanced the transparency of the model so that one could easily pinpoint the most significant features contributing to predictions. Combination of performance-focused ensemble learning with explainability methods places the framework in a good position to serve as a viable method for trustworthy, high-accuracy predictive modeling. Extensions of this research in the future might apply the framework to specialized domains like medical imaging or genetic disorder identification, where interpretability and trustworthiness are of utmost value for decision-making support.

## REFERENCES

- [1] J. Karim *et al.*, “Detection of non-cardiac fetal abnormalities by ultrasound at 11–14 weeks: systematic review and meta-analysis,” *Ultrasound Obstet. Gynecol.*, vol. 64, no. 1, pp. 15–27, Jul. 2024.
- [2] Q. Yang *et al.*, “Multi-Center Study on Deep Learning-Assisted Detection and Classification of Fetal Central Nervous System Anomalies Using Ultrasound Imaging,” arXiv:2501.02000, Jan. 2025.
- [3] Cochrane Pregnancy and Childbirth Group, “Accuracy of first- and second-trimester ultrasound scan for identifying fetal anomalies in low-risk and unselected populations,” *Cochrane Database Syst. Rev.*, no. CD014715, May 2024.
- [4] H. Bashir, A. Khan, and S. Farooq, “Concept-Bottleneck Models for Explainable Deep Learning in Fetal Ultrasound,” in *Proc. IEEE Int. Symp. Biomed. Imaging (ISBI)*, Nice, France, Apr. 2025, pp. 123–130.
- [5] F. Serra *et al.*, “Diagnosing fetal malformations on the 1st trimester ultrasound: a paradigm shift,” Poster, FMF, 2024.
- [6] K. V. Kostyukov *et al.*, “Deep Machine Learning for Early Diagnosis of Fetal Neural Tube Defects,” in *World Congress Fetal Medicine*, 2025, pp. 21–22.
- [7] T. Tenajas, L. Chen, and M. Rodriguez, “AG-CNN: Attention-Guided Convolutional Neural Networks for Improved Organ Segmentation in Prenatal Ultrasound,” in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit. Workshops (CVPRW)*, Seattle, WA, Jun. 2025, pp. 678–686.
- [8] A. Kumar, S. Patel, and D. Rao, “Grad-CAM-Equipped CNN for Renal Anomaly Prediction in Prenatal Ultrasound,” in *Proc. IEEE Int. Conf. Med. Image Anal. (MIA)*, Berlin, Mar. 2025, pp. 342–349.
- [9] R. Singh *et al.*, “Attention-Guided U-Net++ with Grad-CAM++ for Robust Fetal Head Segmentation in Noisy Ultrasound Images,” *Sci. Rep.*, vol. 15, Art. 19612, Jun. 2025.
- [10] O. O. Agboola *et al.*, “Deep Learning Approaches to Identify Subtle Anomalies in Prenatal Ultrasound Imaging,” *Path of Science*, vol. 11, no. 6, pp. 3019–3026, Jun. 2025.
- [11] G. C. Christopher *et al.*, “Enhanced Fetal Brain Ultrasound Image Diagnosis Using Deep Convolutional Neural Networks,” EasyChair Preprint, Nov. 2024.
- [12] U. Islam *et al.*, “Fetal-Net: Enhancing Maternal-Fetal Ultrasound Interpretation through Multi-Scale Convolutional Neural Networks and Transformers,” *Sci. Rep.*, vol. 15, Art. 25665, Jul. 2025.
- [13] S. Belciug *et al.*, “Pattern Recognition and Anomaly Detection in Fetal Morphology Using Deep Learning and Statistical Learning (PARADISE): Protocol for an Intelligent Decision Support System,” *BMJ Open*, vol. 14, Art. e077366, Feb. 2024.
- [14] K. Ramesh, P. Mehta, and A. Sharma, “A Three-Stage Deep Ensemble Pipeline for Intrapartum Head-Position Assessment in Fetal Ultrasound,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Toronto, Sep. 2025, pp. 845–852.
- [15] S. Gupta *et al.*, “Prenatal Diagnostics Using Deep Learning: Dual Approach to Plane Localization and Cerebellum Segmentation,” *J. Med. Imaging Anal.*, vol. 7, no. 1, pp. 45–55, Feb. 2025.
- [16] X. Zhang *et al.*, “Advancing Prenatal Healthcare by Explainable AI Enhanced Fetal Ultrasound Image Segmentation Using U-Net++ with Attention Mechanisms,” *Sci. Rep.*, vol. 15, Art. 30012, Jun. 2025.