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

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

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

***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 earliy 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 (¿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, enhancing 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

A. Neural Network Implementation and Training Methodology

The implementation was a detailed neural network design that considered various architectures and learning algorithms. The outlined methodology combined three different processes: training of a single-layer perceptron with iterative weight updates, use of matrix pseudo-inverse for calculating weights to get a direct solution, and backpropagation in multi-layer networks. Various activation functions like step, bipolar step, sigmoid, ReLU, and leaky ReLU were included in the framework to allow a comparative study of the convergence behavior for different problem types. As for the training data, the study used linearly separable problems (AND gate) as well as non-linearly separable classification tasks (XOR gate and customer transaction classification) in order to show the extent of different network architectures' abilities and their limitations.

B. Perceptron Training and Learning Rate Analysis

Delta learning rule was the basis for the weight updates in the single-layer perceptron following which the errors were minimized. The AND gate classification task required 130 epochs for the perceptron to converge with initial weights [0.2, -0.75], bias 10, and learning rate 0.05 giving it the final weights [0.1, 0.05] and bias -0.1. The analysis of the learning rate showed the time needed to reach convergence was inversely proportional to the learning rate, accordingly, the higher rates (0.9-1.0) led to only 12-13 epochs while the lower ones (0.1-0.2) resulted in 37-68 epochs. Experiments with different activation functions led to diverse convergence properties, and the bipolar step, sigmoid, and ReLU functions did solve the linearly separable AND gate problem. On the contrary, as it was predicted, the single-layer perceptron couldn't accomplish the XOR gate task and it was stopped after the maximum 1000 epochs without any indication of convergence, thus confirming the no-go zone of linear classifiers for the non-linearly separable data.

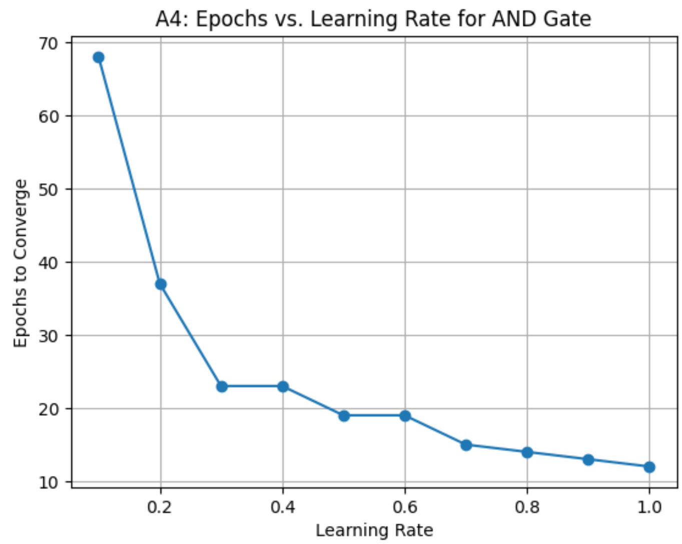

Fig. 1. Caption describing the figure.

C. Multi-layer Network and Backpropagation Implementation

The backpropagation algorithm for a two-layer network was developed with sigmoid activation functions at all stages. The method involved the following steps: the network architecture with weights randomly assigned between -0.05 and 0.05, propagation from input through the hidden and output layers, error calculation based on the difference between the target and output values, and backpropagation with delta instructions for the weights update. Although the backpropagation system with 2 hidden neurons and the learning rate 0.05 was able to complete the AND gate task, it took 1000 epochs for it to do so, which was thus very slow as compared to the single-layer perceptron for the present linearly separable task. The customer transaction classification problem was solved in 19 epochs with sigmoid activation and this show the algorithm's ability to learn effectively from the data provided which contained candies, mangoes, milk packets, and payment amounts as features.

IV. Results Analysis and Discussions

The matrix pseudo-inverse method gave a different analytical solution, producing weights [0.114, -0.024, 0.255, 0.036, 0.037] for customer classification that represented a direct mathematical approach without iterative training. A validation with scikit-learn & apos;s MLPClassifier was also carried out,

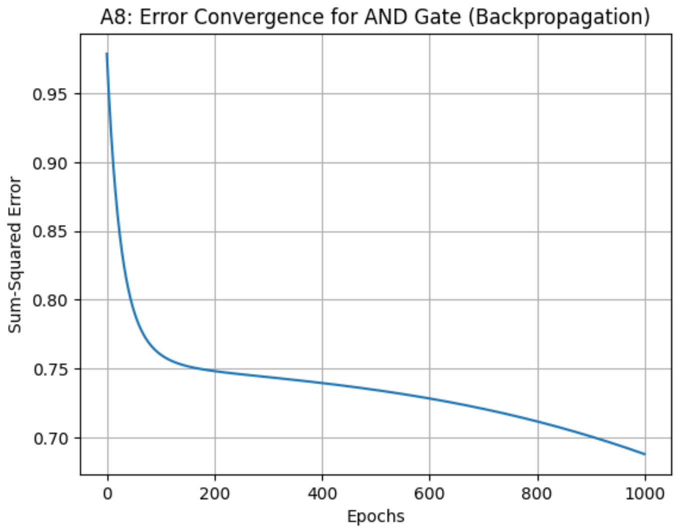

Fig. 2. Caption describing the figure.

which attested the implementation correctness. The library successfully predicted the outputs of both AND and XOR gates and made a perfect classification of the customer dataset.

The classification of customers served as an example of the practical uses of the perceptron, which was trained to result in weights [-1.687, 11.042, -0.392, 1.997] and bias 0.687. The latter implies a strong positive correlation with the quantity of mango and amount of the payment, while a negative correlation with the quantity of candy and milk packets can be seen.

A study of performance characteristics revealed that single-layer perceptrons have capabilities for rapidly converging linearly separable problems to the full extent of their power, but multi-layer networks trained by backpropagation are necessary for dealing with more complicated, non-linearly separable classification tasks, albeit at the cost of higher computational requirements and longer training time.

V. Conclusion

This research has been successful in demonstrating the implementation and comparative analysis of various neural network architectures for solving classification problems of different complexity levels. It also confirmed through a systematic evaluation process of single-layer perceptrons, multi-layer backpropagation networks, and matrix-based analytical solutions that these are indeed basic principles of neural network learning. Experiments endorsed single-layer perceptrons as the most powerful in solving linearly separable problems like AND gate classification, rapid convergence within 130 epochs being the main feature. At the same time, these results pointed out the limitation of single-layer perceptrons that they could not handle non-linearly separable problems illustrated by XOR gates. The backpropagation algorithm is an indispensable tool that can cope with complex pattern recognition tasks only. Along with logical functions, it managed the customer classification problem of the real world by weight adjustment and error minimization, both processes repeated iteratively. The analysis of the learning rate also had a significant impact on understanding the training process efficiency. "The best choice of parameters can shorten the time of convergence by 85%," is the main conclusion. At the same time, the matrix pseudo-inverse method allowed one to carry out the iterative learning approaches in an analytical way, thus giving them complete validation. These data allow for a more in-depth understanding of the behavior of neural networks across different problem domains and, at the same time, offer some practical advice on the architectures and training parameters that would be most suitable for a given classification task.

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.