Friedrich-Alexander-Universität

Integration of Structural and Diffusion Metrics in Fetal MRI: A Model to Support Clinical Evaluation

Master's Thesis in Medical Engineering

submitted by

Javier Munoz Trujillo
born 27.05.1995 in Bogota, Colombia

Written at

Smart Imaging Lab
Institutes of Radiology and Neuroradiology
Friedrich-Alexander-Universität Erlangen-Nürnberg.

Advisors: Prof. Doc. Jana Hutter,
Ph.D. Candidate Nyvenn Castro,

Started: 01.11.2024
Finished: 17.04.2024

Declaration of Originality

I, Javier Munoz Trujillo, student registration number: 23187330, hereby confirm that I completed the submitted work independently and without the unauthorized assistance of third parties and without the use of undisclosed and, in particular, unauthorized aids. This work has not been previously submitted in its current form or in a similar form to any other examination authorities and has not been accepted as part of an examination by any other examination authority.

Where the wording has been taken from other people's work or ideas, this has been properly acknowledged and referenced. This also applies to drawings, sketches, diagrams and sources from the Internet.

In particular, I am aware that the use of artificial intelligence is forbidden unless its use an aid has been expressly permitted by the examiner. This applies in particular to chatbots (especially ChatGPT) and such programs in general that can complete the tasks of the examination or parts thereof on my behalf.

Furthermore, I am aware that working with others in one room or by means of social media represents the unauthorized assistance of third parties within the above meaning, if group work is not expressly permitted. Each exchange of information with others during the examination, with the exception of examiners and invigilators, about the structure or contents of the examination or any other information such as sources is not permitted. The same applies to attempts to do so.

Any infringements of the above rules constitute fraud or attempted fraud and shall lead to the examination being graded "fail" („nicht bestanden").

Place, date	Signature

Abstract

This thesis presents a unified framework that integrates multiple Magnetic Resonance Imaging (MRI) contrasts with advanced machine learning techniques to support the clinical evaluation of fetal conditions, specifically ventriculomegaly and pre-eclampsia, after 23 weeks of gestation. By combining structural T2-weighted imaging and functional T2* relaxation, the proposed approach offers a comprehensive tissue characterization that overcomes the limitations of single-modality assessments. Structural brain volumetric measurements were leveraged to accurately predict gestational age, with both Linear Regression and Random Forest models achieving an R^2 of 0.958. For anomaly detection, supervised classifiers, notably an XGBoost model, attained an accuracy of 98.1% in ventriculomegaly detection, while a Neural Network classifier utilizing placental radiomic features achieved 99.7% accuracy for pre-eclampsia prediction.

A significant innovation of this work is the employment of a Vector Quantization Variational Auto-encoder (VQ-VAE) to generate interpretable brain heatmaps that highlight latent structural irregularities, thereby enhancing clinical decision-making. In addition, the application of Shapley Additive Explanations (SHAP) provided transparent insights into feature contributions, enabling a clear interpretation of the decision process behind the automated predictions.

While the results are promising, limitations such as sample size disparities and potential overfitting in underrepresented cases highlight the need for further validation through multi-center studies and longitudinal data integration. Future work will aim to expand the dataset, incorporate additional imaging modalities, and develop a user-friendly clinical decision support tool to facilitate real-world adoption, especially in resource-limited settings.

Overall, this thesis demonstrates a significant advancement in automated fetal diagnostics by merging multi-modal MRI data with state-of-the-art machine learning to deliver a robust, interpretable, and clinically viable framework, thereby paving the way for improved maternal–fetal healthcare.

Contents

Chapter 1

Introduction

In recent decades, medical imaging has led to significant advances in the field of obstetrics, allowing healthcare professionals to diagnose and monitor fetal development more accurately during pregnancy [Rec+24]. Among the available imaging techniques, MRI has become a valuable complement to conventional ultrasound, providing deeper insights into fetal growth and potential complications [Lev01]. While ultrasound plays a crucial role in diagnostic procedures such as fetal echocardiography [Leu21], it has limitations in detecting all fetal abnormalities in other organs, with a persistent risk of false positives and false negatives despite ongoing technological progress [Sun]. A key advantage of fetal MRI over ultrasound lies in its ability to generate high-resolution multi-planar images with a larger field of view, superior tissue contrasts and volumetric measurements, critical consideration in prenatal diagnostics [Pat+22].

The superior soft tissue contrast provided by MRI represents a significant advance in fetal imaging [OCo+12]. In contrast to ultrasound, which can be limited by maternal body habitus, fetal position, or reduced amniotic fluid, MRI overcomes these challenges to provide a detailed and consistent visualization of fetal anatomy [Roe+07]. This enhanced capability is particularly valuable for the examination of complex structures such as the developing brain, where subtle abnormalities might otherwise go undetected [OCo+12]. The ability of technology to differentiate between various types of tissue enables clinicians to identify and characterize structural anomalies with remarkable precision [Nov+21].

Furthermore, fetal MRI has established itself as an indispensable tool in cases where there is not enough information obtained by ultrasound. In instances where ultrasound findings are inconclusive or suggest complex congenital anomalies, advanced MRI capabilities can provide the detailed anatomical information necessary for accurate diagnosis and appropriate

treatment planning [Lev01]. This is particularly evident in placental abnormalities where three-dimensional visualization and tissue characterization are essential for a comprehensive assessment [Nov+21].

Monitoring of fetal development represents a cornerstone of modern prenatal care, with implications that extend far beyond the gestational period. This approach aligns closely with the concept of Developmental Origins of Health and Disease (DOHaD), which emphasizes how environmental factors during early life development, including the fetal period, can influence health outcomes throughout an individual's lifespan [Lac19]. Early detection of developmental abnormalities or deviations from expected growth patterns can fundamentally alter the trajectory of prenatal and postnatal care [Li+23]. When healthcare professionals can identify potential complications during fetal development, they have the invaluable opportunity to develop comprehensive management strategies before birth [Sou+24]. This approach enables medical teams to gather the necessary specialists, plan appropriate interventions, and prepare both healthcare facilities and families for post-delivery challenges, significantly improving neonatal outcomes across a spectrum of conditions, from cardiac abnormalities to neurological complications [Li+23].

With all this, modern obstetrics faces growing challenges due to an increase in complicated pregnancies, driven by societal changes and shifts in maternal health [Sou+24]. The increasing prevalence of advanced maternal age in High Income Countries (HICs) or the incidence of early pregnancy in low- and middle-income countries (LMICs) introduces additional risk factors that require close monitoring during pregnancy [Sou+24]. Moreover, the increasing prevalence of chronic conditions such as diabetes or pre eclampsia among expectant mothers further complicates fetal development, facing more intensive care and specialized approaches [Ala+23].

Despite significant advances in medical imaging technology, the field of fetal diagnostics continues to face substantial challenges that impact clinical practice and patient care [Sit+15], healthcare systems worldwide struggle not only with fundamental access to medical equipment, but also with severe shortages of trained personnel, resulting in a lack of primary attention to the pregnant population [Hil+25].

According to a UNICEF's study presented at the AlignMNH Opening Forum (20-21 April 2021) [Cou21], In 88 LMICs surveyed since 2010: 75% of pregnant women received four or more prenatal care contacts, 86% delivered in a health facility and 87% of mother-baby pairs received a postnatal visit within two days of delivery. However, only 57% of

women-baby received all three interventions: four prenatal care visits, institutional delivery care, and postnatal care visit within two days.

This project aims to integrate the functionalities of MRI with Machine Learning (ML) algorithms to facilitate the access of healthcare professionals to comprehensive information regarding fetal health. While ML can be applied to various aspects of fetal monitoring, this project primary focuses on two cases, ventriculomegaly and pre-eclampsia, due to their high clinical relevance and the potential for MRI to offer a holistic view through different functional contrasts.

Cases of ventriculomegaly have not been extensively reported in the literature regarding LMICs. However, the incidence of congenital hydrocephalus, which is often associated with ventriculomegaly, has been significantly higher. Specifically, the incidence rate is approximately 123 per 100,000 births, compared to 79 per 100,000 births in HICs [Dew+19].

Pre-eclampsia is the second leading cause of avoidable maternal deaths, with approximately 16% of all maternal deaths attributable to this cause [May+24] and more than 99% of pre-eclampsia-related maternal deaths occur in LMICs according to The European Foundation for the Care of Newborn Infants (EFCNI). [EFC24]

A comprehensive system has been developed to automate biometric measurements and enhance Gestational Age (GA) prediction. As part of this, an algorithm extracts and quantifies key brain parameters. Additionally, a predictive model has been designed to detect deviations in brain and placental development in fetuses from 23 weeks of gestation onward.

MRI offers protocols with different parameters that results in different functional imaging contrasts such as diffusion-weighted T2 *-relaxation, Apparent Diffusion Coefficient (ADC), and Intravoxel Incoherent Motion (IVIM). These techniques enable the characterization of tissue microstructure, oxygenation, and perfusion, offering new opportunities for early detection of developmental abnormalities [Cro+24] . Recent studies have refined acquisition protocols and post-processing techniques, improving the reliability and reproducibility of these biomarkers across different GA [Shi+20].

In the detection of ventriculomegaly and pre-eclampsia, novel algorithms based on diffusion-weighted T2* mapping and texture analysis have been leveraged. All these approaches integrate artificial intelligence and ML techniques to enhance diagnostic accuracy and provide deeper insights into fetal and placental health.

For ventriculomegaly detection, researchers have developed an automated workflow using a deep learning model to measure lateral ventricle diameter in fetal brain MRI scans

as explained in the article from the Rush Medical College in Chicago [Vah+23]. This UNet-based model segments the fetal brain into seven tissue categories and performs automatic measurements at the level of the thalamus and choroid plexus.

In the realm of pre-eclampsia detection, researchers have explored the use of diffusion-derived vascular density (DDVD), a new diffusion-weighted MRI-based technique, to assess placental micro-perfusion as exposed by a group of Mayo Clinic [Al +23]. In addition, the UT Southwestern Medical Center has investigated texture analysis to identify changes in placental heterogeneity using Haralick metrics or voxel-wise geometric analysis. These methods are correlated with GA and can distinguish between normal and diseased placental tissue. [Do+19].

The proposal model leverages the integration of diverse MRI contrasts, combining structural imaging with functional techniques. With this comprehensive approach, the model aims to generate classifications of ventriculomegaly and pre-eclampsia in conjunction with probabilistic maps that offer highlights of deviations from developmental patterns or normal clinical cases. To support classification and prediction tasks across multiple conditions, this framework incorporates a diverse set of machine learning models tailored to the structure and complexity of fetal MRI data. Random Forest and XGBoost were selected for their strong performance with high-dimensional structured datasets and their ability to model non-linear relationships while addressing class imbalance [Bre01; Che+16]. K-Nearest Neighbors (KNN) provides a simple, interpretable approach that is particularly effective in capturing localized spatial variations, making it suitable for medical imaging applications [Cov+67]. Additionally, Bayesian Neural Networks (BNNs) were integrated for their capacity to model predictive uncertainty, an important feature in clinical decision-making where confidence in classification is critical [Nga+24].

To translate these technical capabilities into clinically meaningful tools, the probabilistic map is created based on a VQ-VAE. This deep learning architecture offers a lower-dimensional latent space while preserving essential developmental features in high-resolution images.

In addition to volumetric and texture-based metrics, a Gyrification Index (GI) experiment was explored to incorporate cortical folding characteristics as potential biomarkers of neurodevelopment [Zil+88]. However, this approach was not integrated into the final pipeline due to inconsistency in the results and sensitivity to segmentation quality.

The significance of this project extends to technological assistance in global maternal-fetal healthcare, addressing access gaps and providing a potential guide for healthcare

professionals. Many healthcare facilities in LMICs, while equipped with MRI capabilities, lack the specialized expertise required for complex fetal image interpretation. By enhancing, for instance, the capabilities of teleradiology, this tool could support radiologists and clinicians in analyzing fetal imaging remotely, facilitating earlier interventions in settings where specialist consultation might otherwise be delayed or unavailable.

In the future, this project could incorporate the implementation of early warning systems in real time by incorporating automated alerts during the acquisition and interpretation of MRI. With this, the risk of missed diagnoses that could impact maternal-fetal outcomes can be significantly reduced. These real-time notification systems serve as a safety net, alerting potential abnormalities that could be subtle or easily overlooked during routine evaluation. This capability is particularly valuable in high-volume clinical settings, where radiologists face increasing workloads and time pressures. The system's ability to provide immediate feedback during image acquisition also offers the opportunity to optimize scanning protocols in real-time, ensuring the capture of diagnostically valuable information while the patient is still in the scanner.

Furthermore, the methodological framework developed in this research has the potential for expansion beyond the detection of ventriculomegaly or preeclamsia. Future advancements could integrate this model to identify and monitor Fetal Grouth Restruction (FGR), congenital heart defect (CHD), and abnormalities in cortical folding patterns, which are key indicators of neurodevelopmental disorders. Expanding its application would enhance fetal imaging capabilities, providing a more comprehensive tool for early diagnosis, risk assessment, and intervention planning across a broader range of conditions.

Chapter 2

Methodology

This project aims to develop and validate a ML-based algorithm to assess fetal conditions using MRI. To ensure tight evaluation and increase the potential for immediate clinical translation, this thesis focuses as initial applications for this framework on the detection of ventriculomegaly and pre-eclampsia after 23 weeks of GA.

To achieve this, a quantitative approach was adopted, involving the development and training of various ML models. This approach allows for statistical validation of its precision in identifying deviations from normal fetal development, which is key to addressing the clinical challenges associated with these conditions, particularly in limited resource environments.

The project utilizes a large pre-existing dataset of fetal MRI scans acquired on a clinical 0.55 T scanner (MAGNETOM Free.Max, Siemens Healthcare, Erlangen, Germany), with a 9 element spine coil integrated into the patient table in conjunction with a 6 element flexible coil (BioMatrix Contour Coil, Siemens Healthcare, Erlangen, Germany) conducted as part of three ethically approved prospective single-center studies. These studies contain both healthy cases and cases with ventriculomegaly. Furthermore, data was gathered from an additional study with pre-eclampsia cases [Hal+24] using a clinical 1.5T Philips Ingenia scanner with a combined 24-channel posterior and torso (dStream) coil. For this project, only T2* relaxation and T2-weighted images were utilized from all described studies.

All computational experiments presented, were executed on the TinyGPU cluster hosted by the National High Performance Computing Center (NHR@FAU). Specifically, the A100 partition was utilized, which provides access to NVIDIA A100 Tensor Core GPUs with 80 GB of memory, high-throughput NVMe storage, and AMD EPYC 7742.[]

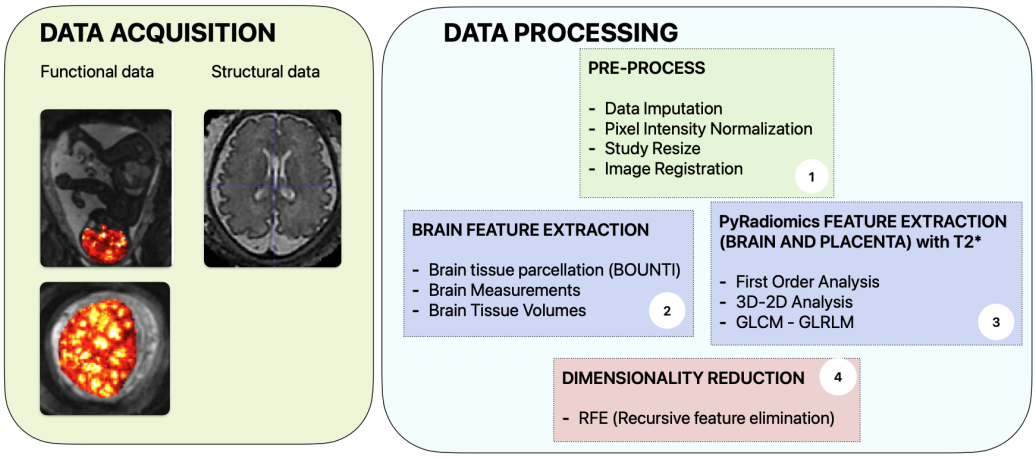


Figure 2.1: Overview of data acquisition and processing pipeline. The data acquisition step includes functional and structural MRI scans, while the data processing pipeline involves 1) pre-processing, 2-3) feature extraction, and 4) dimensionality reduction for machine learning model development.

2.1 Pre-processing

As first (illustrated in Figure 2.1) a pre-processing pipeline was implemented to ensure data quality and consistency. This process included data imputation, pixel normalization, study resizing, and image registration. The specific steps are described in detail below.

Initially, all studies were manually inspected, and only those with complete, high-quality brain parcellation and spatial continuity were selected. For instance, the left image in Figure 2.2 represents a poorly structured segmentation with imprecise boundary delineation and artifacts, indicating reduced data, for example, due to increased fetal motion. In contrast, the right image demonstrates a well-defined segmentation with clear anatomical structures, representing high-quality data suitable for further analysis.

To standardize the intensity range specifically across T2-weighted images, the Min-Max normalization method was applied in SimpleITK, ensuring that the pixel intensity values fall within the range $0 \leq \text{Pi} \leq 1$ [25c].

Image resizing was performed using SimpleITK, which provides interpolation techniques designed for medical imaging while preserving critical metadata, such as orientation, spacing, and direction, despite varying voxel sizes [25d].

For registration, the framework itk-Elastik [25b] was employed. This tool allows for accurate alignment of MRI scans by optimizing transformations that correct for motion artifacts and anatomical differences between subjects. This step is crucial for ensuring

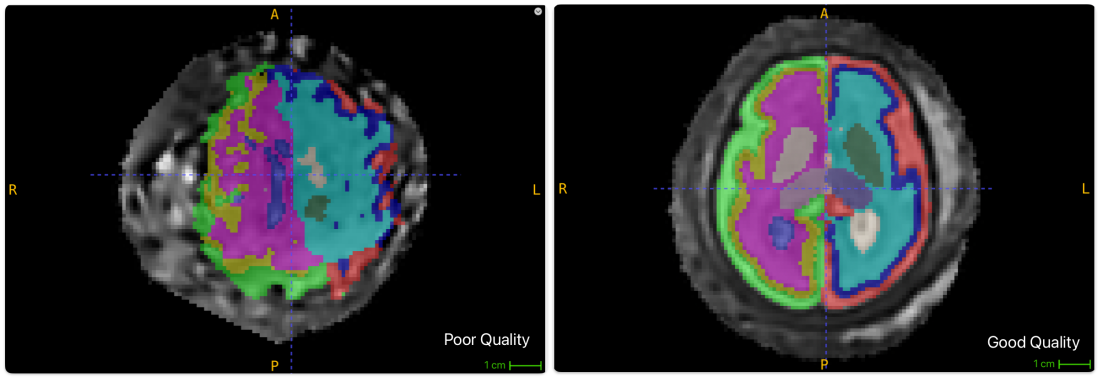


Figure 2.2: Comparison of segmentation quality. The left image shows low-quality segmentation with artifacts and unclear boundaries, while the right image presents a high-quality segmentation with well-defined anatomical structures.

that subsequent analyses are performed on properly aligned data, reducing variability and enhancing model performance.

Brain feature extraction is the second step (Figure 2.3), focusing on quantitative biomarkers relevant to ventriculomegaly, pre-eclampsia and their strong correlation with the prediction of GA.

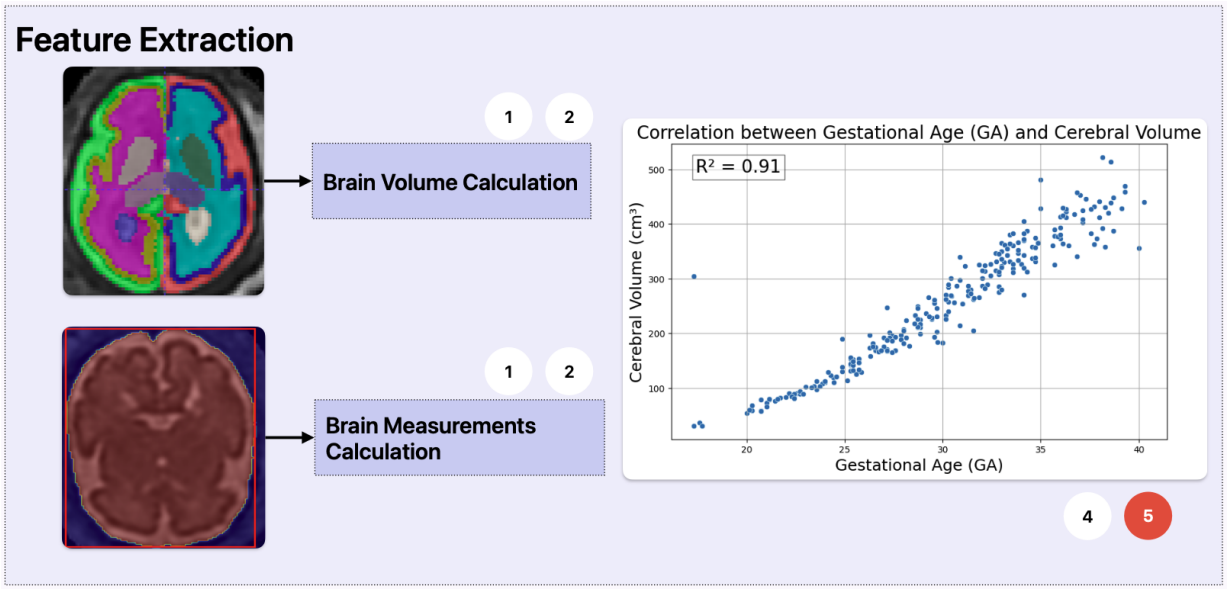


Figure 2.3: Brain Feature Extraction Method. This step involves calculating brain volume and other structural measurements from MRI scans. The extracted features are then analyzed for their correlation with gestational age (GA), as shown in the scatter plot ($R^2 = 0.91$), highlighting the predictive value of cerebral volume in fetal development assessment.

A deep learning-based segmentation pipeline [25e], built upon the BOUNTI framework, was employed for brain parcellation. This segmentation resulted in tissue masks that allowed for the computation of volumetric measurements across different brain regions, which are essential for quantitatively characterizing fetal brain development and identifying potential structural abnormalities. Accurate parcellation is crucial because subtle deviations in regional brain volumes can indicate pathologies or developmental delays that may otherwise remain undetected [Li+24]. In Table A.1, all 19 labels representing distinct anatomical regions are listed. Once the volume of each specific brain region was determined, further brain tissue volumes were derived using established methodologies as outlined in [Nap+20] and [Bin+21]. These additional derived volumes include composite metrics such as total brain volume, ventricular volume ratios, and normalized cortical volumes. These metrics correspond to Step 5 and provide deeper insights into possible abnormalities within the brain structure as suggested by [Gro+06] study.

2.2 GA Prediction

As illustrated in Step 6 of the methodological pipeline (Figure 2.4), the prediction of GA was performed using a feature selection process based on correlation analysis. Specifically, features that exhibit a Pearson correlation coefficient greater than 0.90 with gestational age were selected from previously extracted brain volumetric measurements (Step 5). These features demonstrated strong linear or non-linear associations with fetal development milestones and were therefore considered highly informative for prediction tasks.

Two machine learning models were trained and evaluated: Linear Regression and Random Forest. The rationale for selecting these models lies in their complementary strengths. Linear Regression offers interpretability and computational efficiency, enabling a straightforward assessment of how each individual feature influences GA [Mol24]. In contrast, the Random Forest model is capable of capturing more complex, non-linear relationships between features while providing robustness to outliers and overfitting [Aur+12].

Both models were trained on the selected features using a 70:30 train-validation split. Performance metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2, were calculated and analyzed in the Results chapter (Section 3.1). These metrics provide a comprehensive assessment of the prediction accuracy and generalizability of each model.

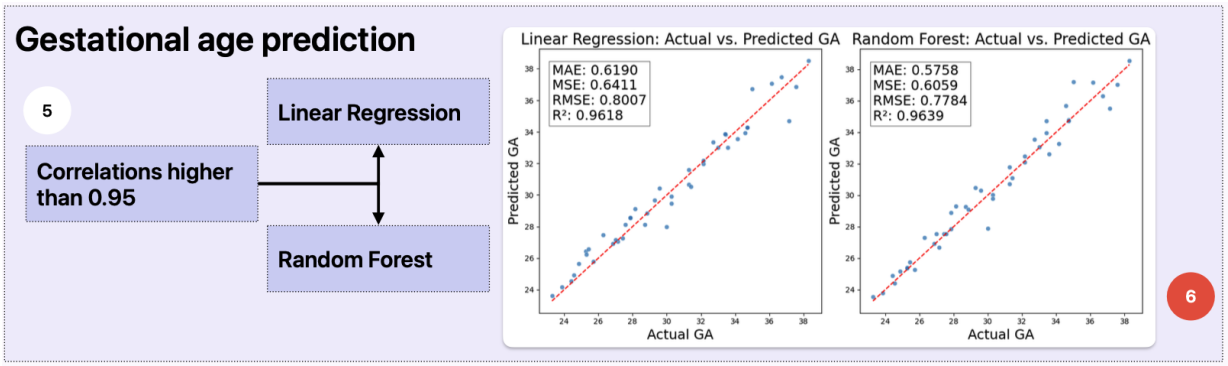


Figure 2.4: GA prediction model. Step 6 involves training and evaluating both a Linear Regression and a Random Forest model using brain features strongly correlated with gestational age

In addition to standard regression metrics, a residual distribution analysis was performed to evaluate the spread and symmetry of prediction errors across the gestational age range. This method supports the identification of potential biases, such as systematic under- or over-estimation, and is a common diagnostic tool for assessing regression model assumptions and generalizability [Jam+21].

2.3 VQ-VAE and Brain Heatmap

To improve brain analysis, a VQ-VAE [Pin+22] was developed to detect potential abnormalities in heatmaps, marked as Step 7 and depicted in Figure 2.5. The VQ-VAE is a variation of the Variational Auto-encoder (VAE) architecture that incorporates vector quantization, enabling discrete latent representations. Unlike traditional architectures, which use continuous latent spaces, this approach learns a discrete codebook that helps capture meaningful structural features while reducing reconstruction artifacts. To enhance training performance, the dataset was divided into one-week gestational age bins, enabling the model to learn age-specific anatomical patterns more effectively and various hyperparameters were tested, including the number of embeddings, batch size, and latent dimensionality. The model was trained using the Adam optimizer with a learning rate of 0.001 and a β value of 0.25, favoring convergence without overfitting.

After training, the model was used to reconstruct brain scans, and abnormalities were estimated by comparing the original and reconstructed volumes. Specifically, each scan was normalized and passed through the trained VQ-VAE. The absolute difference between the original and reconstructed scans was computed, highlighting regions where the model failed

to accurately reconstruct the input. These differences indicate potentially abnormal regions. To reduce noise, a threshold of 0.3 was applied and values below this threshold were set to zero, resulting in a focused heatmap of high-probability abnormal areas.

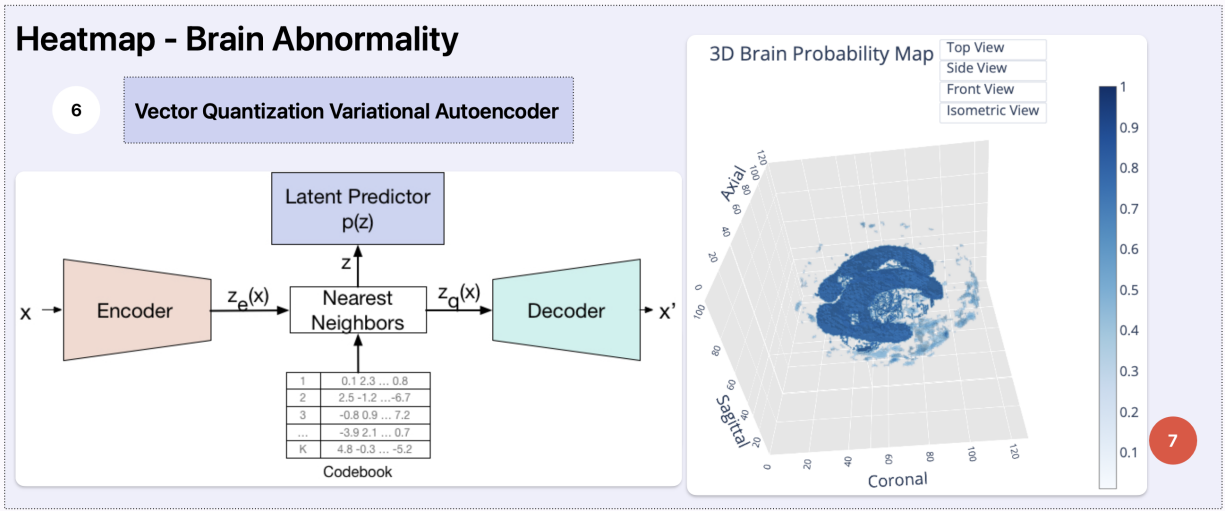


Figure 2.5: Brain Heatmap Model. The diagram illustrates the structure of the VQ-VAE, emphasizing the intermediate codebook step. The heatmap (right) represents the model's probability output for abnormality detection, with darker blue indicating regions of higher abnormality probability and lighter (white) areas signifying regions of low or no probability. The image illustrates a brain identified as a positive case of ventriculomegaly.

2.4 Ventriculomegaly classification

In Step 8 of the methodological pipeline (Figure 2.6), the classification of ventriculomegaly was estimated using four machine learning models, each selected for their specific advantages in handling high-dimensional, clinical imaging-derived data.

The dataset, acquired on the MAGNETOM Free.Max scanner (Siemens Healthcare, Erlangen, Germany) as described earlier, originally included 18 confirmed positive cases and 108 negative cases. To address class imbalance while maintaining statistical robustness, a subset of 53 negative cases was selected, resulting in a more balanced dataset suitable for model training. A 70:30 train-validation split was applied to ensure reliable performance evaluation.

The models implemented in this step were configured as follows: the Random Forest classifier was set with 100 estimators and a balanced class weight to mitigate potential class imbalance. The K-Nearest Neighbors (KNN) model used $k = 5$ to classify instances

based on feature-space similarity. The XGBoost classifier was configured with a softprob output function and a multi-logarithmic (mlogloss) loss objective, enabling probabilistic predictions and iterative improvement of misclassified cases. Finally, the Bayesian Neural Network (BNN) architecture incorporated ReLU-activated layers and a dropout rate of 0.3, and was optimized using the Adam algorithm with categorical cross-entropy as the loss function.

Each model's performance was assessed using standard classification metrics: accuracy, precision, recall, and F1 score. The results are further detailed in the chapter (Section 3.3).

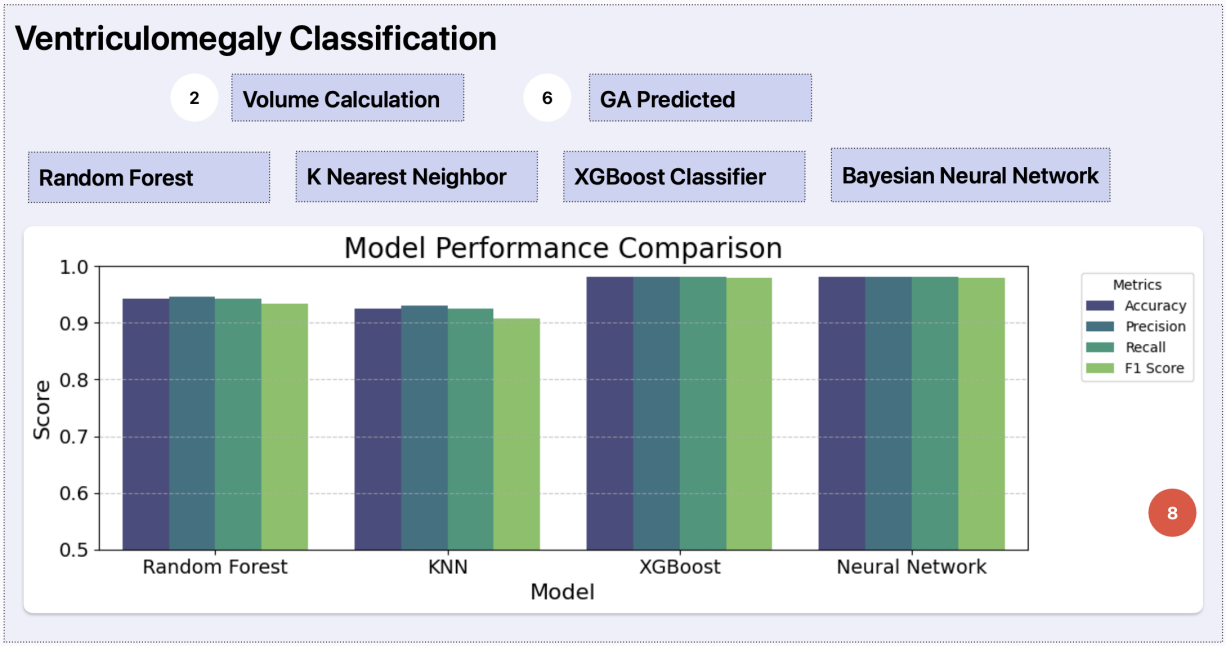


Figure 2.6: Ventriculomegaly classification. Step 8 includes training of four machine learning models on structural brain features to classify ventriculomegaly presence.

To further enhance model interpretability, SHAP analysis was conducted on the best model to assess the relative importance of input features. This approach assigns each feature a contribution value toward the final model prediction, supporting transparency and consistency in decision-making.

2.5 Pre-eclampsia Likelihood Estimation

As illustrated in Step 9 of the methodological framework (Figure 2.7), the estimation of pre-eclampsia likelihood was conducted using the same four machine learning models

described in the ventriculomegaly section. This analysis was extended across two anatomical regions—brain and placenta—in order to determine which provided superior predictive performance based on T2* mapping-derived features.

The dataset used for this analysis was derived from a prospective cohort study described by [Hal+24], where pregnant women were recruited from antenatal clinics and categorized into three cohorts:

- 291 controls with no hypertensive complications (Cohort 0)
- 82 confirmed cases of pre-eclampsia (Cohort 1)
- 27 cases of chronic hypertension (Cohort 2)

Participants who were initially enrolled as controls but later developed unrelated pregnancy complications were retrospectively excluded to preserve cohort integrity.

From the placental and brain T2* images, a total of 107 radiomic features were extracted [25a] (Step 3 in the pipeline, Figure 2.1), encompassing three main descriptor categories:

- First-order intensity statistics, including mean, standard deviation, skewness, kurtosis, and entropy, characterized the global distribution of voxel intensities and provided insights into perfusion and oxygenation heterogeneity [Do+19].
- Shape-based features, such as volume, surface area, compactness, and sphericity, captured geometric attributes of the segmented placenta, which are known to reflect vascular development abnormalities in pre-eclamptic conditions [Hal+24].
- Second-order texture metrics derived from the Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), and Gray Level Dependence Matrix (GLDM) quantified local spatial heterogeneity and microstructural disorganization within the tissue [Al +23].

All models were trained and validated using a 70:30 split and evaluated through standard classification metrics including accuracy, precision, recall, and F1 score, with performance results reported in Section 3.4.

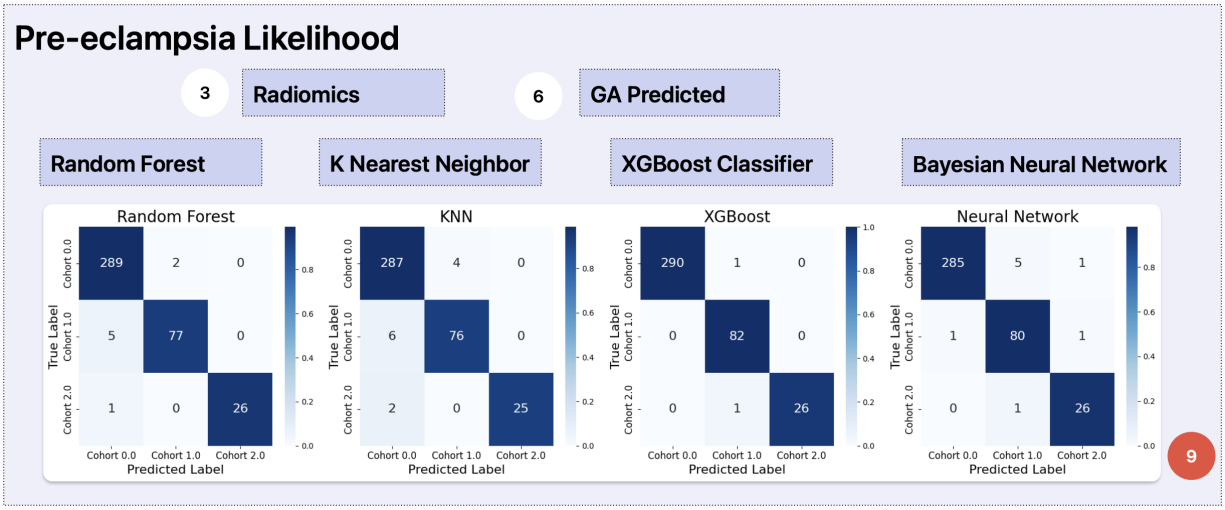


Figure 2.7: Pre-eclampsia model likelihood. Step 9 includes model training on brain and placental radiomic features extracted from T2* MRI to assess pre-eclampsia likelihood.

2.6 Clinical Integration via Web Interface

To facilitate the clinical translation of the developed fetal MRI analysis pipeline, a dynamic and interactive web-based interface was implemented using the Dash framework [Par+24]. The application was designed to visualize and report patient-specific data and model predictions in a single format.

The interface also loads an interactive plot that allows users to assess individual brain metrics, in relation to gestational age norms via regression-based outlier detection. Studies identified with potential ventriculomegaly are flagged based on deviations in lateral ventricle volumes or asymmetry indices.

2.7 GI Exploratory Experiment

A GI exploratory experiment was conducted, the analysis began by simplifying the brain segmentation into three main tissue types: cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM). A binary mask of gray matter was extracted and processed using morphological operations to isolate the pial surface. The surface mesh was then reconstructed via the marching cubes algorithm, yielding two key 3D surface representations:

- The **pial surface**: outlining the outer boundary of the cortical gray matter.

- The **exposed (smoothed) surface**: estimated through dilation and smoothing to represent a convex approximation of the cortex.

Surface areas were computed from both meshes using the spacing information from the original MRI volume. The global Gyrification Index was then defined as:

$$\text{GI} = \frac{\text{Pial Surface Area}}{\text{Exposed Surface Area}}$$

To explore regional variability, a vertex-wise estimation was implemented. For each exposed surface vertex, a spherical neighborhood was defined using a Euclidean radius, and a local GI value was computed by measuring the ratio of corresponding surface patches. This produced a continuous GI heatmap that could be projected onto the cortical mesh to visualize local folding intensity.

Chapter 3

Results

This chapter presents the outcomes of the methodology as outlined, maintaining alignment with the structured analytical pipeline. The results are organized to reflect each step in the proposed workflow, beginning with the prediction GA, followed by the reconstruction and evaluation of the heat map VQ-VAE, the estimation of ventriculomegaly, and concludes with the pre-eclampsia likelihood. Each section includes quantitative performance metrics.

3.1 Gestational Age Prediction

As described in Section 2.2, brain features with a Pearson correlation coefficient greater than 0.95 with gestational age were selected for model training. The features derived from key volumetric measurements such as Biparietal diameter or Ventricle/parenchyma ratio, are listed in Annex Table A.2, with the complete correlation matrix available in Figure A.2.

The performance of the Linear Regression model and the Random Forest is summarized in Table 3.1, corresponding to Step 6 of the methodological pipeline (Figure 2.4).

Table 3.1: Performance comparison of GA prediction models. Best results per metric are highlighted in bold.

Model	MAE	MSE	RMSE	R^2
Linear Regression	**0.606**	**0.725**	**0.851**	**0.958**
Random Forest	0.613	0.748	0.865	**0.958**

Both models achieved the same R^2 scores of 0.95, indicating a strong linear association between predicted and actual gestational ages. The Linear Regression model reported

slightly lower error values across all metrics, with a MAE of 0.606, MSE of 0.725, and RMSE of 0.851 with an uncertainty prediction of 1.4 weeks, compared to the Random Forest model, which reported a MAE of 0.613, MSE of 0.748, and RMSE of 0.865. Figure 3.1 illustrates the predictions of both models across the full gestational age range.

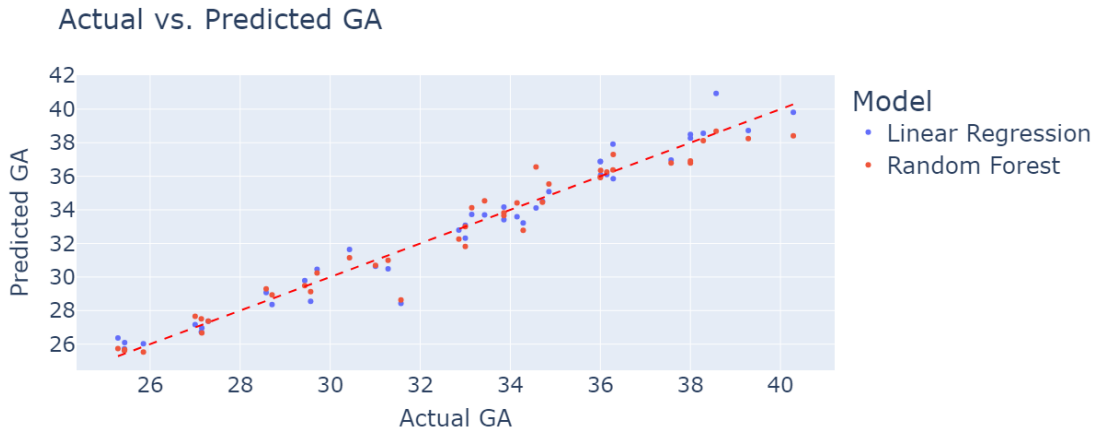


Figure 3.1: Models prediction across all GA cases

As part of the model inspection, the residual distribution analysis (Figure 3.2) provides additional information on the prediction behavior. The Linear Regression model shows a bell-shaped distribution centered near zero, with a skewness value of 0.61. The Random Forest model also displays a distribution centered near zero, with a slightly higher skewness of 0.77.

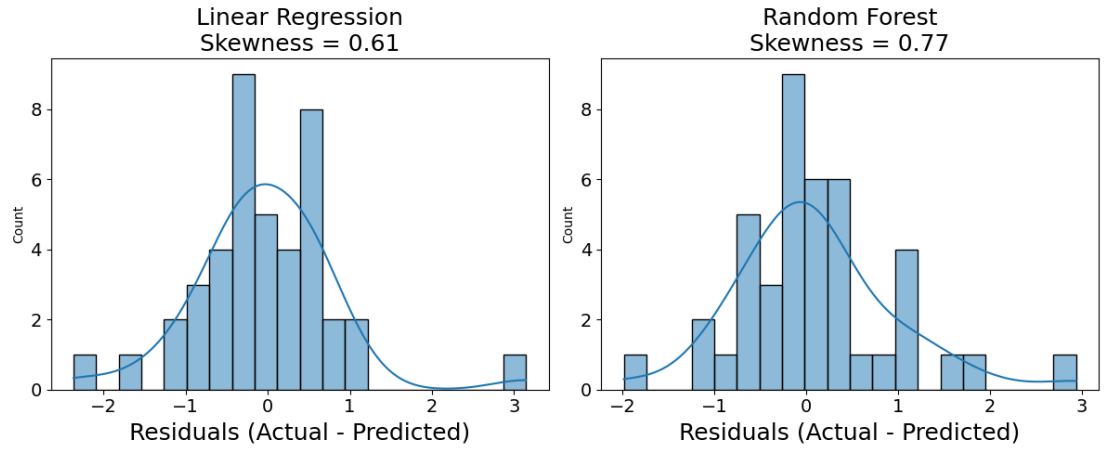


Figure 3.2: Residual distributions in GA prediction

3.2 VQ-VAE and Brain Heatmap Analysis

Aligned with Step 7 of the methodological pipeline, the VQ-VAE was implemented to generate brain heatmaps for the detection of potential structural abnormalities. This approach enables unsupervised learning of latent representations by compressing input data into discrete codebook embeddings, thereby facilitating the identification of deviations from expected anatomical patterns throughout gestation.

Multiple hyperparameter configurations were tested, varying the number of embeddings, batch size, and latent dimensionality, to identify the optimal setup. Figure 3.3 displays the performance of the top three configurations based on reconstruction quality and structural similarity.

Figure 3.3: Performance comparison of top VQ-VAE configurations

The configuration using **256 embeddings**, a **batch size of 50**, and a **latent dimension of 32** yielded the most balanced and robust results. As shown in Table 3.2, this setting achieved the lowest MSE (0.054), highest Structural Similarity Index Measure (SSIM) (0.254), and best DICE similarity coefficient (0.282).

Table 3.2: Performance comparison of top VQ-VAE configurations **(embeddings + batch size + latent dim)**

Metric	256+10+16	256+10+64	256+50+32
Last MSE	0.060	0.058	**0.054**
Last VQ-VAE Loss	0.065	0.069	0.080
Last Reconstruction Loss	0.060	0.056	**0.054**
Last Loss	0.126	0.128	0.135
Average SSIM	0.222	0.201	**0.254**
DICE	0.276	0.275	**0.282**

3.3 Ventriculomegaly Classification

Following Step 8 of the analytical pipeline, ventriculomegaly classification was performed using four supervised machine and the results are through the confusion matrices in Figure 3.4 and the quantitative metrics reported in Table 3.3.

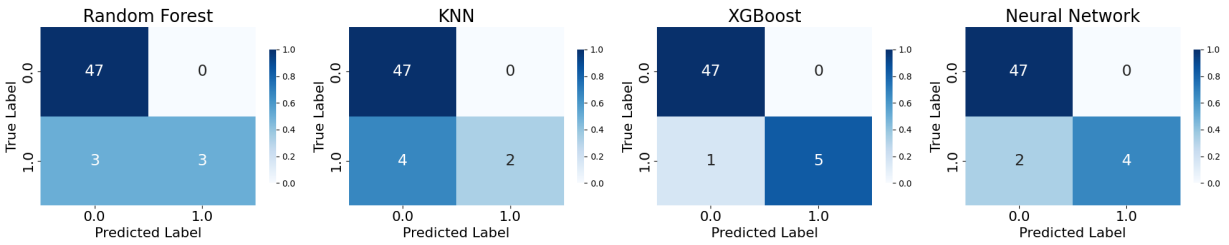


Figure 3.4: Confusion matrix for ventriculomegaly classification

Table 3.3: Ventriculomegaly Classification Models Performance Comparison

Model	Accuracy	Precision	Recall	F1 Score
Random Forest	0.962	0.963	0.962	0.958
KNN	0.924	0.930	0.924	0.907
XGBoost	**0.981**	**0.981**	**0.981**	**0.980**
Neural Network	0.962	0.963	0.962	0.958

Among the evaluated models, **XGBoost demonstrated the highest overall performance**, achieving an accuracy of 0.981 and an F1 score of 0.980. The model also achieved an ROC-AUC score of 0.997.

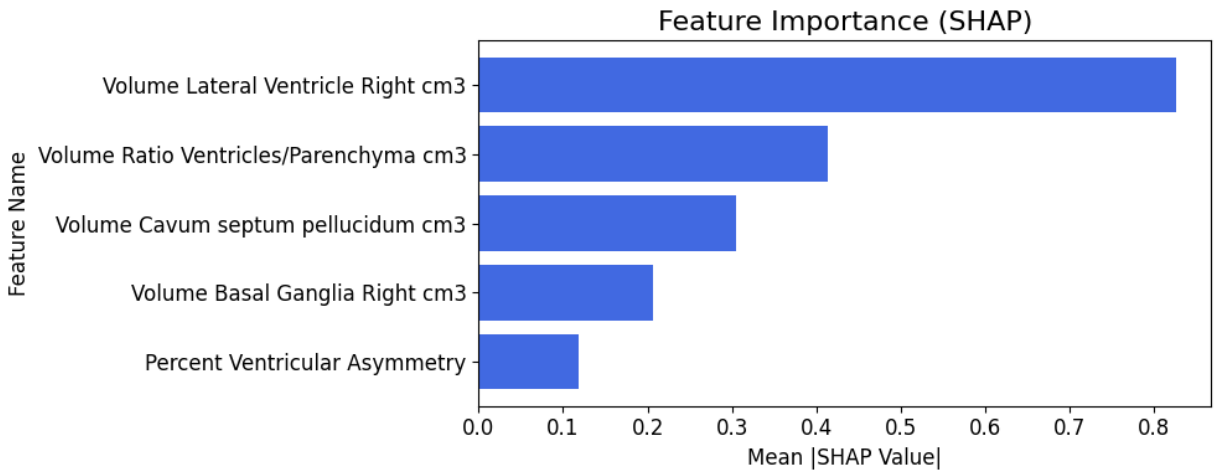


Figure 3.5: Feature Importance in XGBoost Model Prediction for Ventriculomegaly

As illustrated in Figure 3.5, the most influential predictor was the **volume of the right lateral ventricle**, with a mean SHAP value exceeding 0.8. The second most significant feature was the **volume ratio between ventricles and parenchyma**,

Other contributing features included the volume of the cavum septum pellucidum, basal ganglia (right), and percent ventricular asymmetry.

3.4 Pre-eclampsia Likelihood Estimation

Aligned with Step 9 of the analytical framework, pre-eclampsia likelihood estimation was performed using radiomic features extracted from two distinct anatomical regions: the fetal brain and the placenta.

Pre-eclampsia likelihood Based on Brain Images

Initial classification models were trained using the brain T2* tissue analysis extracted from Radiomics [25a]. Their performance is illustrated in the confusion matrix (Figure 3.6) and detailed in Table 3.4.

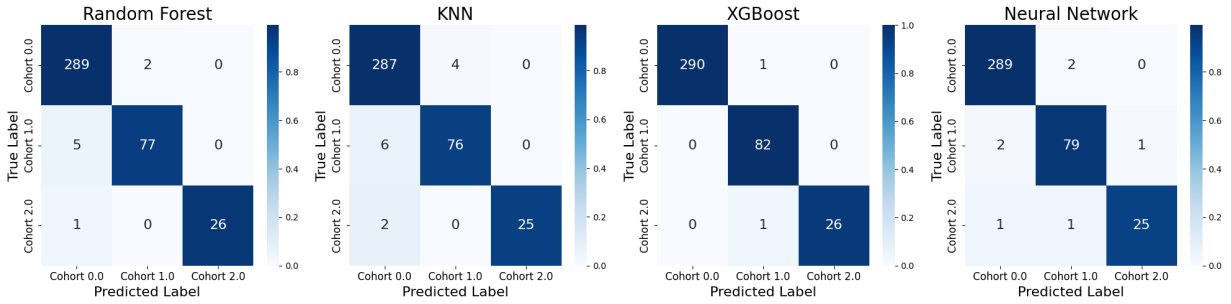


Figure 3.6: Confusion matrix for pre-eclampsia likelihood model output based on brain T2* mapping features from Radiomics

Table 3.4: Pre-eclampsia Likelihood Models Performance Comparison based on brain T2* mapping

Model	**Accuracy**	**Precision**	**Recall**	**F1 Score**
Random Forest	0.980	0.980	0.980	0.979
KNN	0.965	0.964	0.965	0.964
XGBoost	**0.995**	**0.995**	**0.995**	**0.994**
Neural Network	0.972	0.972	0.972	0.972

The XGBoost model once again achieved the best results across all evaluated metrics, with an accuracy, precision, and recall of 0.995. It produced one single false positive and one single false negative throughout the entire test set of 400 cases.

Pre-eclampsia likelihood Based on Placenta Images

The same machine learning pipeline was applied to radiomic features extracted from placental T2*. This parallel analysis enabled a direct comparison with brain results to determine which anatomical region yielded more reliable predictive signals. Classification results are illustrated in Figure 3.7, with corresponding performance metrics detailed in Table 3.5.

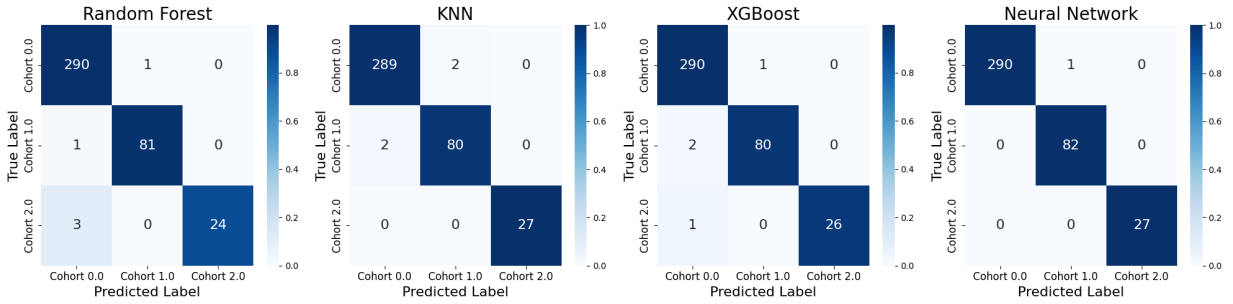


Figure 3.7: Confusion matrix for pre-eclampsia likelihood model based on placental T2* mapping

Table 3.5: Pre-eclampsia Likelihood Models Performance Comparison based on placental T2* mapping

Model	**Accuracy**	**Precision**	**Recall**	**F1 Score**
Random Forest	0.987	0.987	0.987	0.987
KNN	0.990	0.990	0.990	0.990
XGBoost	0.990	0.990	0.990	0.990
Neural Network	**0.997**	**0.997**	**0.997**	**0.997**

Among all classifiers, the **Neural Network model** achieved the highest performance, correctly identifying nearly all cases with only one false negative recorded. With accuracy, precision, recall, and F1 scores of 0.997, it outperformed all models trained on brain features.

SHAP Analysis by Cohort

To further interpret the model's decisions, SHAP waterfall analyses were conducted across all three cohorts using the Bayesian Neural Network model. These analyses quantify the contribution of individual radiomic features to each prediction, offering a transparent view of the model's internal decision-making process.

Cohort 0: Negative Cases For the control group (Cohort 0), the most impactful predictor was **gestational age**, with a SHAP value of +0.74, as shown in Figure 3.8. Additional contributions came from texture-related metrics such as the Informational Measure of Correlation and Zone Entropy. All these radiomic markers collectively added meaningful support to the model by capturing normal structural and textural placental characteristics.

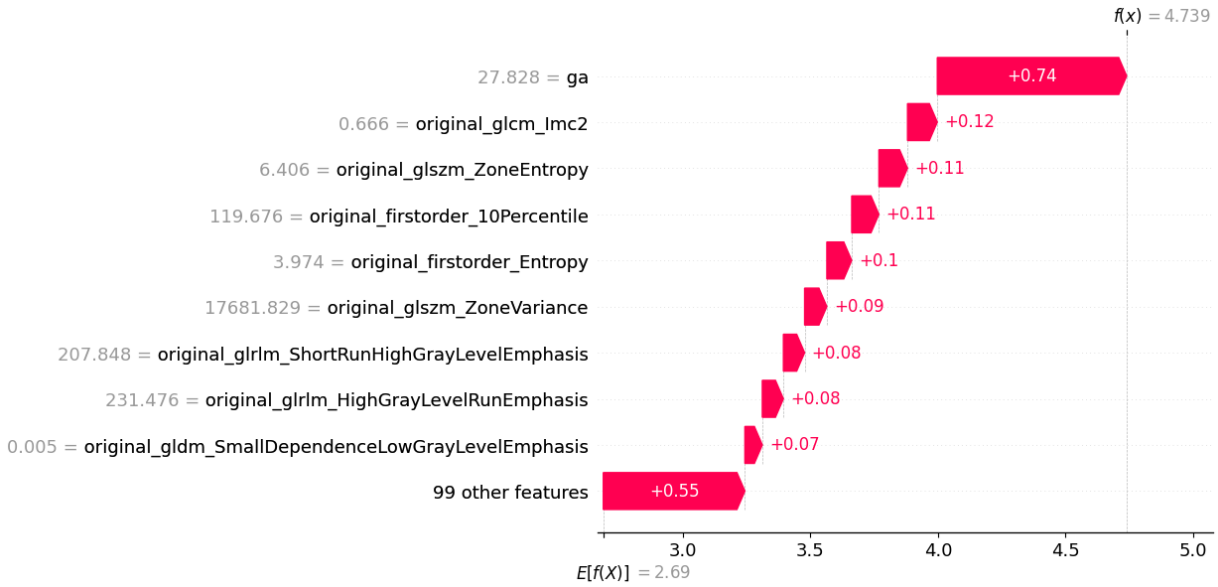


Figure 3.8: Waterfall SHAP for Cohort 0 (negative pre-eclampsia cases). Gray values represent baseline radiomic features from placental T2* mapping; red bars indicate individual contributions to the final model output.

Cohort 1: Pre-eclampsia Cases In pre-eclamptic cases (Cohort 1), gestational age remained the dominant feature, with a substantially increased SHAP contribution of +1.78, as shown in Figure 3.9. Other highly relevant features included the 10th percentile intensity and Small-Dependent Low Gray Level Emphasis.

Cohort 2: Chronic Hypertension For cases of chronic hypertension (Cohort 2), as shown in Figure 3.10, the top feature was Run Entropy (+1.28), followed by Large Dependence High Gray Level Emphasis (+0.51). These metrics point to distinct textural signatures compared to both control and pre-eclamptic cases.

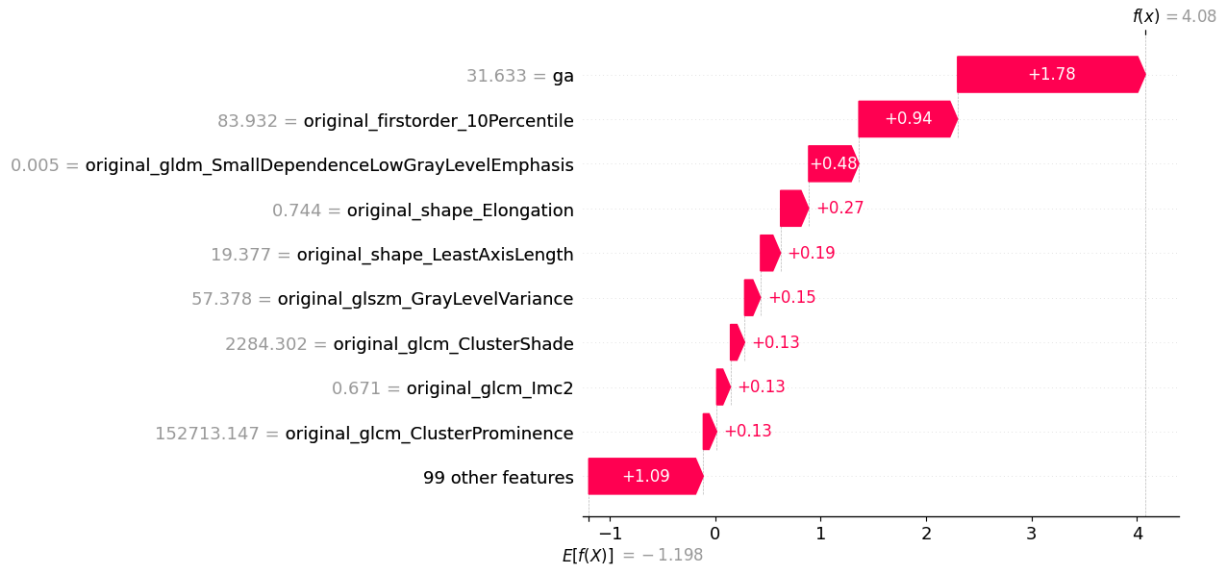


Figure 3.9: Waterfall SHAP for Cohort 1 (positive pre-eclampsia cases). Gray values represent baseline radiomic features from placental T2* mapping; red bars indicate individual contributions to the final model output.

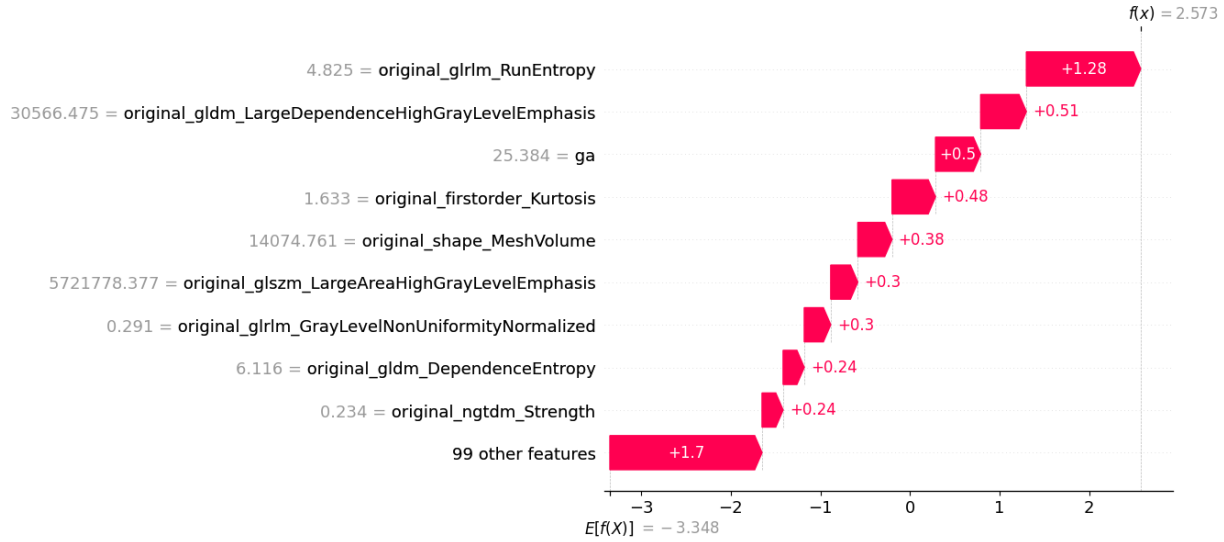
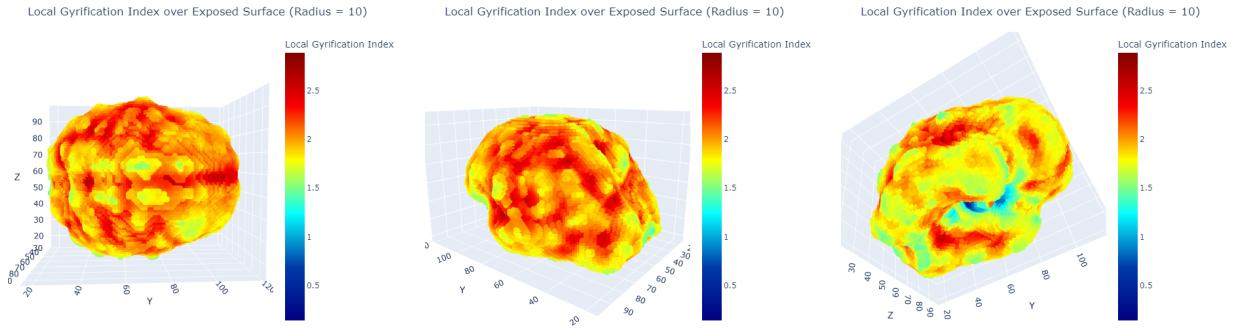


Figure 3.10: Waterfall SHAP for Cohort 2 (chronic hypertension cases). Gray values represent baseline radiomic features from placental T2* mapping; red bars indicate individual contributions to the final model output.

3.5 GI Exploratory Experiment



(a) Frontal view (b) Lateral view (c) Oblique view

Figure 3.11: Local Gyrification Index visualized from three different angles. Warmer colors represent regions of higher cortical folding complexity.

Vertexwise GI was calculated and visualized on the cortical surface. As shown in Figure 3.11, the local GI values ranged approximately between 0.5 and 2.5. Regions with higher folding complexity, indicated by warmer colors (red and orange), were primarily localized in the central and lateral portions of the cortex. In contrast, lower GI values (cooler colors) appeared in areas more convex or less folded, such as the medial and inferior surfaces. The 3D projections from multiple angles provided a comprehensive view of regional variability in cortical folding patterns, confirming the utility of this method for capturing local morphological differences.

3.6 Clinical Integration via Web Interface

To streamline the use of these findings in clinical settings, a user-friendly graphical interface was developed to visualize individual patient results. As illustrated in Figure 3.12, the interface includes automatic segmentation, gestational biomarker estimations, heatmaps, and SHAP-based feature interpretations all in an HTML format suitable for deployment in diagnostic workflows.

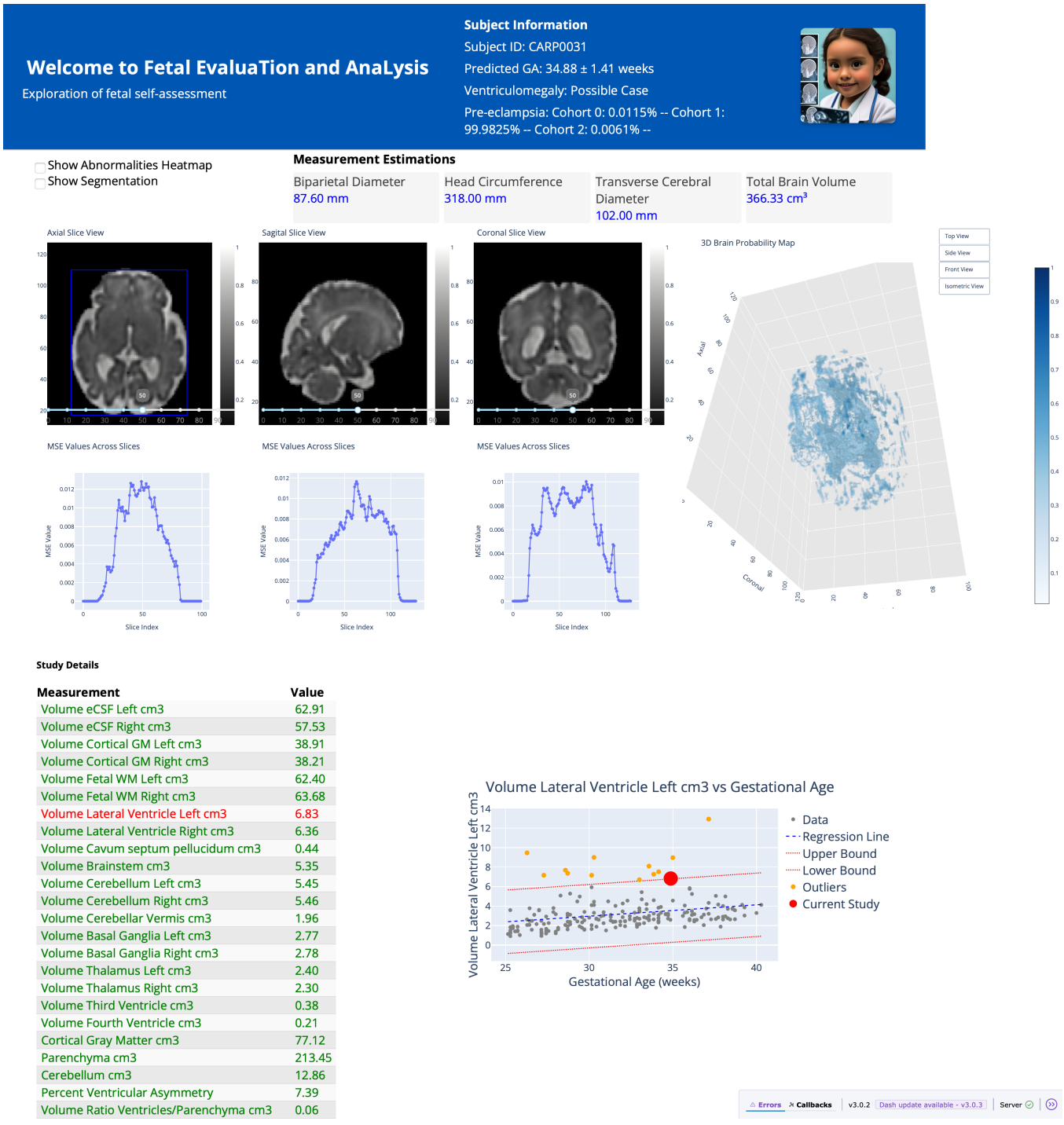


Figure 3.12: HTML output interface. Includes subject basic information, brain probability heatmap, brain measurement available and measurement bounds.

Chapter 4

Discussion

The primary objective of this thesis was to integrate multiple MRI contrasts and advanced machine learning algorithms to facilitate the automated assessment of fetal conditions after 23 weeks of gestation, illustrated exemplarily with ventriculomegaly and pre-eclampsia. The results from this study provide compelling evidence that both conventional regression models and more complex supervised classifiers can be harnessed to capture key developmental and pathological patterns in fetal MRI.

4.1 Gestational Age Prediction

The performance of the Linear Regression and Random Forest models in predicting gestational age (GA) was highly encouraging, with both models achieving an R^2 score of 0.958. This strong correlation demonstrates that structural brain features, such as cerebellum and parenchyma volume (as illustrated in Annex Figure A.2), serve as reliable indicators of fetal developmental progress. The residual plot (Figure 3.2) reveals a systematic tendency for both models to slightly underestimate gestational age in a subset of cases. This consistent bias suggests that while the overall predictive signal is strong, the models may benefit from calibration techniques, such as isotonic regression or Platt scaling, that can adjust predictions to better align with true values, particularly in regions of the input space where errors are more frequent.

Looking ahead, as more data becomes available—especially covering a wider and more balanced gestational age range—it would be worthwhile to explore more expressive model architectures. For example, transformer-based models, which have shown success in other

biomedical domains, could provide a data-efficient alternative for capturing subtle and complex patterns in structural brain features.

4.2 VQ-VAE Reconstruction and Heatmap Analysis

The evaluation of the VQ-VAE model indicated that the configuration with 256 embeddings, a batch size of 50, and a latent dimensionality of 32 achieved the most favorable performance across multiple metrics (Table 3.2). Despite this optimal setting, the overall Structural Similarity Index Measure (SSIM) of 0.254 and DICE coefficient of 0.282 were relatively low. Such modest values can be attributed to the inherent complexity of fetal brain anatomy, variability across gestational ages, and the unsupervised nature of the model, which lacks ground truth spatial supervision. Nonetheless, the quality of the reconstructions generates informative heatmaps that highlight structural deviations in anomalous cases, supporting the use of the VQ-VAE primarily as an anomaly detection tool rather than for precise segmentation.

4.3 Anomaly Detection for Ventriculomegaly

For the detection of ventriculomegaly, the anomaly detection task yielded promising results with the XGBoost classifier, which achieved an accuracy of 98.1% and an F1 score of 0.980. The analysis identified lateral ventricle volume and the ratio between ventricles and parenchyma as critical markers, consistent with previous literature indicating that disproportionate ventricular enlargement is a primary indicator of ventriculomegaly [Gro+06]. Furthermore, features such as the volume of the cavum septum pellucidum, right basal ganglia volume, and percent ventricular asymmetry also contributed significantly, underscoring the model's ability to integrate diverse anatomical cues into a cohesive predictive framework.

4.4 Pre-eclampsia Detection and SHAP Analysis

For pre-eclampsia detection, two parallel pipelines based on brain T2* (Table 3.4) and placental T2* (Table 3.5) imaging were evaluated. The brain-based models, with the XGBoost classifier, demonstrated robust classification performance. However, the placenta-

based models slightly outperformed them, with the Bayesian Neural Network achieving the highest overall accuracy.

This difference in performance suggests that placental T2* imaging may offer a more discriminative and sensitive signature for detecting pre-eclampsia, aligning with prior findings that highlight the placenta's central role in the disease [Ton+19]. The complex microstructural and perfusion characteristics of the placental tissue appear to yield richer radiomic biomarkers—such as textural and intensity-based features—that are more strongly associated with early pathological changes. Furthermore, the success of the Bayesian Neural Network highlights the advantages of probabilistic modeling in this context, as it not only improves predictive performance but also provides calibrated uncertainty estimates, which are particularly valuable in clinical decision-making.

SHAP Analysis by Cohort

Cohort 0: Negative Cases For the control group, the most impactful predictor was **gestational age**, with a SHAP value of +0.74 (Figure 3.8). Additionally, texture-related metrics such as the **Informational Measure of Correlation** and **Zone Matrix Entropy** contributed to a negative classification. The Informational Measure of Correlation quantifies the consistency of texture patterns, where higher values indicate a homogeneous and uniform tissue structure. In conjunction, Zone Matrix Entropy reflects the distribution and variability of pixel intensities, with moderate values suggesting balanced heterogeneity showing well-perfused tissue.

Cohort 1: Pre-eclampsia Cases In pre-eclamptic cases, **gestational age** remains a dominant factor but with a substantially higher SHAP value of +1.78 (Figure 3.9), indicating that deviations in placental features become more pronounced with advancing gestation in these cases. Two additional key predictors in this cohort are the **10th Percentile Intensity** and **Small Dependence Low Gray Level Emphasis**. The 10th Percentile Intensity captures the lower end of the pixel intensity distribution, which is indicative of hypoxic or under-perfused regions, while the Small Dependence Low Gray Level Emphasis reflects fine-grained textural disruptions and potential early calcific changes. These factors contribute to a radiomic profile that is markedly different from healthy placental tissue.

Cohort 2: Chronic Hypertension For cases of chronic hypertension, the SHAP analysis revealed a unique textural signature. The top contributor was **Run Entropy** from the

Gray Level Run Length Matrix (GLRLM), with a SHAP value of +1.28 (Figure 3.10). Run Entropy measures the irregularity and diversity in the distribution of consecutive pixels with the same intensity, with higher values indicating significant texture variation often associated with impaired perfusion. In addition, **Large Dependence High Gray Level Emphasis** from the Gray Level Dependence Matrix (GLDM) contributed with a SHAP value of +0.51. This feature highlights clusters of high-intensity pixels which may reflect compensatory tissue remodeling or focal fibrosis as a response to chronic hypertensive stress.

However, it is important to note that the gestational age range in cohort 2 is imbalanced, with most samples clustering around an average of 25 weeks. This limited variation may reduce the representativeness of this group, potentially affecting the generalizability of these observed textural signatures.

4.5 GI Exploratory Experiment

The global GI was computed as the ratio between the pial surface area and the exposed (smoothed) surface area, and a vertex-wise local GI heatmap was generated to visualize regional cortical folding intensity. However, inconsistencies in the results and high sensitivity to segmentation and meshing artifacts—particularly in early gestational weeks—limited the reliability of the GI measurements. Moreover, the absence of an established reference algorithm or validated protocol to cross-correlate the GI outputs with clinical benchmarks further impeded its integration into the final pipeline. It is likely that the current method of approximating cortical neighborhoods using a spherical model requires further refinement, and a closer analysis in collaboration with clinical experts is warranted to ensure that the measured GI accurately reflects cortical folding patterns.

4.6 Future Directions

Although the results of this study are promising, several limitations and methodological challenges must be addressed before the proposed framework can be directly implemented in a clinical setting. A key obstacle is the disparity in data volume among diagnostic categories, particularly the comparatively small number of pre-eclampsia and ventriculomegaly cases. In addition, the chronic hypertension cohort (Cohort 2) exhibited an imbalanced gestational age range, with most samples clustering around 25 weeks. While techniques such as

resampling and stratified validation were deployed to mitigate these issues, the risk of overfitting remains, thereby emphasizing the need for larger and more balanced datasets.

To further enhance the current framework and build on the promising results of this study, several future directions are proposed. First, the use of larger, multi-center datasets is critical. Collaborative data collection and external validation across institutions could address current limitations such as class imbalance and reduce overfitting, ultimately improving the robustness and generalizability of the models across diverse populations and imaging protocols.

Second, future studies should explore advanced network architectures and novel algorithms. Emerging deep learning models such as Vision Transformers and hybrid convolutional-transformer approaches offer enhanced feature extraction capabilities and could be particularly beneficial in handling high-dimensional imaging data. Moreover, probabilistic frameworks that incorporate real-time uncertainty estimation could support clinical decision-making by highlighting predictions with low confidence. Exploring anomaly detection algorithms may also facilitate the identification of atypical cases and the inclusion of additional pregnancy-related pathologies beyond ventriculomegaly and pre-eclampsia. For example, intrauterine growth restriction (IUGR), a condition with both placental and fetal brain implications, would be a clinically meaningful extension. Including IUGR could help validate the general applicability of the radiomics pipeline and address a wider spectrum of fetal health outcomes.

Third, the inclusion of additional imaging contrasts may offer complementary information to improve disease detection. While this study focused on T2-weighted and T2* imaging, incorporating other MRI modalities such as diffusion-weighted imaging with apparent diffusion coefficient (ADC) mapping or intravoxel incoherent motion (IVIM) metrics may yield further insight into microstructural and perfusion characteristics. T1-weighted imaging and T1 mapping also provide opportunities to detect subtle tissue changes, particularly those related to hemorrhage or tissue integrity, which could support more comprehensive characterization of fetal and placental conditions.

In addition, expert cross-validation and interdisciplinary collaboration will be essential for advancing the clinical relevance of these methods. Given the limited number of studies applying radiomics to placental tissue, input from clinical specialists—including radiologists, obstetricians, and perinatologists—can provide valuable feedback on model interpretability and clinical integration.

Finally, the development of a teleradiology-integrated interface would support real-world deployment. A web-based platform designed to work with existing telemedicine infrastructure could help democratize access to expert-level fetal MRI interpretation, particularly in low-resource settings where radiological expertise is scarce. This challenge is not only about data or models—limited access to both equipment and trained personnel remains a major barrier to equitable care and must not be underestimated. A striking example comes from Peru, where Siemens Healthineers recently donated a high-end MRI scanner to a regional hospital to support maternal-fetal health initiatives [Sie24]. While this donation represents a significant step forward, the full impact of such initiatives depends on concurrent investment in local training and remote diagnostic support. In this context, a teleradiology platform could serve as a practical bridge, offering centralized analysis and expert guidance that help translate cutting-edge imaging technologies into clinical benefits, even in settings where local resources are limited.

Chapter 5

Conclusions

The present thesis developed and evaluated a comprehensive framework leveraging fetal MRI data and machine learning algorithms to detect ventriculomegaly and pre-eclampsia—two clinically significant prenatal conditions. By integrating structural and functional MRI contrasts (T2-weighted and T2* sequences) with both ensemble and deep learning models, the research demonstrates significant potential for enhancing fetal health assessments. The demonstrated efficacy in identifying both ventriculomegaly and pre-eclampsia underscores the capacity of advanced machine learning methods to streamline diagnostic workflows, enabling earlier and more targeted interventions. Future iterations, incorporating larger datasets and more sophisticated deep learning architectures, could evolve into a broadly applicable platform for prenatal diagnostics—valuable in both highly specialized tertiary centers and resource-limited settings worldwide.

Key Contributions

- **GA Prediction and Brain Biometrics:** Brain volumetric features strongly correlated with GA (Section 3.1). Both Linear Regression and Random Forest models achieved high R^2 scores (0.958), effectively predicting GA within MAE of just over 0.6 weeks. These results illustrate that standard volumetric measurements, particularly from the cerebellum and parenchymal regions, can offer an efficient proximation for fetal maturation
- **VQ-VAE Heatmaps for Abnormality Detection:** An unsupervised anomaly detection method based on VQ-VAE was introduced (Section 3.2). Although Structural Similarity Index Measure (SSIM) values were modest, the generated heatmaps

nonetheless highlighted anatomical deviations in cases of ventriculomegaly, demonstrating the capacity of unsupervised learning to uncover subtle pathological features in a data-driven manner

- **Accurate Ventriculomegaly Classification:** Four supervised classification models (Random Forest, KNN, XGBoost, and a Bayesian Neural Network) were trained on structural brain features to detect ventriculomegaly (Section 3.3). XGBoost produced the best results (accuracy of 0.981 and F1 score of 0.980) with a notably high ROC-AUC score of 0.997. SHAP analysis identified lateral ventricle volumes and ventricle-to-parenchyma ratios as key predictive features, aligning with established clinical knowledge

- **Pre-eclampsia Detection:** Parallel classification pipelines trained on brain versus placental radiomics (Section 3.4) showed that placental T2*-based features more robustly distinguish pre-eclampsia than brain-based features. In particular, the Neural Network classifier reached an accuracy and F1 score of 0.997. SHAP analysis revealed several distinct texture signatures indicative of microvascular or perfusion abnormalities within the placenta, highlighting the placental tissue as a critical imaging target for this condition

- **Practical Clinical Integration:** The final pipeline was paired with an HTML-based interface designed to display automated segmentation maps, gestational biomarker evaluations, and heatmap overlays (Figure 3.12). This streamlined presentation facilitates clinician interpretation, supports rapid decision-making, and helps bridge specialized radiological expertise gaps, an especially relevant feature for LMICs

Limitations

Despite the promising outcomes, certain methodological constraints must be acknowledged:

- **Class Imbalance:** The relatively small number of confirmed cases of ventriculomegaly and pre-eclampsia, as well as an uneven distribution across gestational ages (particularly in the chronic hypertension cohort), limited the diversity of training samples. While data augmentation and stratified splitting partially mitigated bias, expanded datasets remain crucial for more robust and generalized models.

- **Multi-Center Heterogeneity:** Data were primarily obtained from single-center sources—plus an additional single site for pre-eclampsia scans—making it uncertain how well the model would perform on data collected under different scanning protocols or using different MRI vendors.

- **VQ-VAE Reconstruction Precision:** Although heatmaps offered valuable qualitative insights for identifying structural deviations, the moderate SSIM and DICE values underscore the complexities of unsupervised fetal brain modeling. Enhancing spatial alignment and refining training hyperparameters may yield further improvements.

Future Directions

- **Data Expansion and Multi-Center Collaboration:** Acquiring larger, more diverse fetal MRI datasets will be pivotal for refining model reliability, especially for rare conditions. Collaborations across multiple hospitals or imaging centers could standardize protocols and augment sample sizes, leading to improved generalizability.

- **Integration of Additional MRI Contrasts:** Incorporating perfusion or diffusion-based modalities (e.g., IVIM,ADC) may yield richer microstructural and vascular biomarkers, potentially enhancing early detection of fetal conditions.

- **Advanced Deep Learning Architectures:** Future models could adopt self-supervised or transformer-based methods to better capture 3D context from multi-parametric scans, enabling detection of more nuanced structural variants. Similarly, Bayesian Neural Networks and uncertainty quantification measures can help clinicians gauge the confidence of each prediction.

- **Real-Time Monitoring and Early Warning Systems:** As MRI technology moves toward faster acquisition protocols, real-time anomaly detection could alert clinicians during scanning, enabling immediate protocol adjustments and potentially earlier interventions.

- **Teleradiology and Resource-Constrained Settings:** Emphasizing lightweight, interpretable interfaces will aid in scaling these methods to LMICs. Tools that can operate effectively even with limited computing resources (e.g., smaller GPU clusters or cloud-based solutions) will expand the reach of advanced fetal imaging diagnostics.

In summary, the work presented here contributes to bridging the gap between prenatal imaging technology and clinical utility. By offering accurate gestational biomarkers, robust classification of high-risk conditions, and interpretable decision support, this framework supports the overarching goal of improved maternal-fetal care on a global scale.

Appendix A

Appendix for additional large figures or tables

Table A.1: Brain Segmentation Areas

Label	**Anatomical area**
1	Volume eCSF Left
2	Volume eCSF Right
3	Volume Cortical GM Left
4	Volume Cortical GM Right
5	Volume Fetal WM Left
6	Volume Fetal WM Right
7	Volume Lateral Ventricle Left
8	Volume Lateral Ventricle Right
9	Volume Cavum Septum Pellucidum
10	Volume Brainstem
11	Volume Cerebellum Left
12	Volume Cerebellum Right
13	Volume Cerebellar Vermis
14	Volume Basal Ganglia Left
15	Volume Basal Ganglia Right
16	Volume Thalamus Left
17	Volume Thalamus Right
18	Volume Third Ventricle
19	Volume Fourth Ventricle

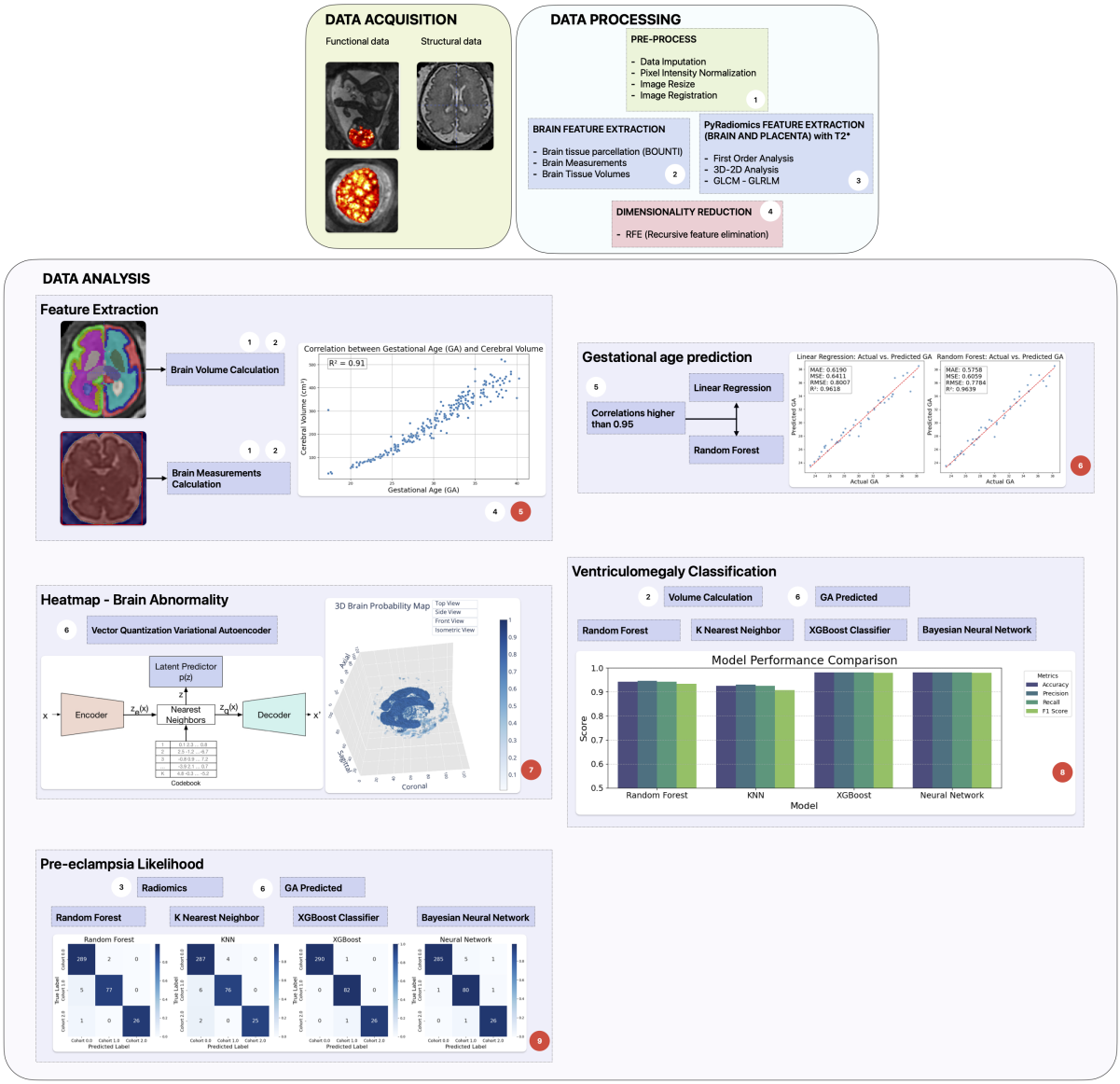


Figure A.1: Project Methods

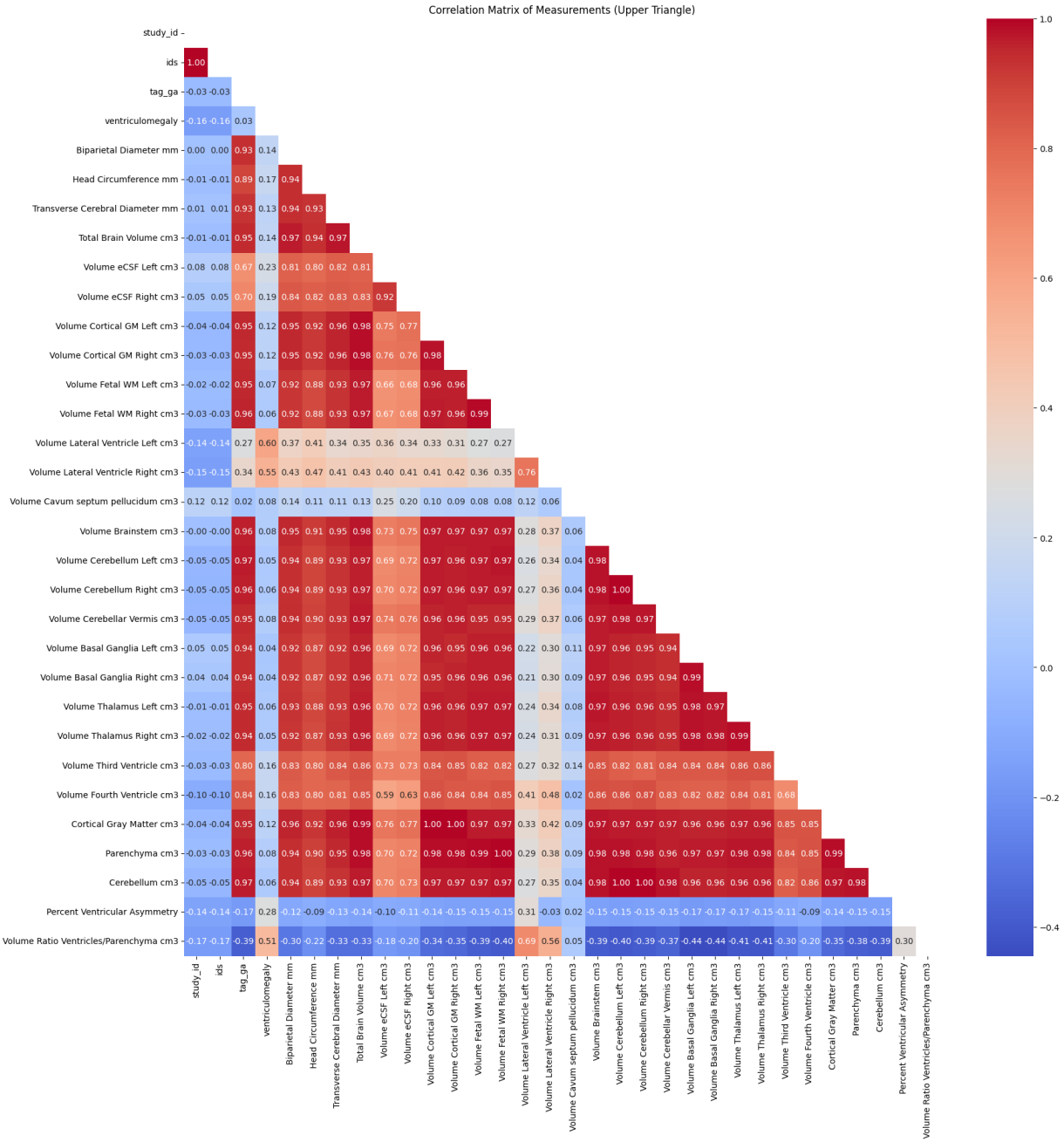


Figure A.2: Correlation matrix of measurements

Table A.2: Brain Measurements and Their Calculations

Measurement	**Calculation Method**
Biparietal Diameter (mm)	Calculated as the difference between the leftmost and rightmost x-coordinates of the brain edges, multiplied by the voxel spacing in the x-direction.
Head Circumference (mm)	Computed by summing the perimeter pixels of the segmented head boundary and multiplying by the voxel spacing in the corresponding direction.
Transverse Cerebral Diameter (mm)	Derived from the difference between the topmost and bottommost y-coordinates of the brain edges, multiplied by the voxel spacing in the y-direction.
Total Brain Volume (cm^3)	Computed as the sum of all labeled brain voxels multiplied by the voxel volume and converted to cubic centimeters.
Cortical Gray Matter (cm^3)	Sum of the left and right cortical gray matter volumes extracted from the segmentation mask.
Parenchyma (cm^3)	Sum of cortical gray matter, fetal white matter, basal ganglia, and thalamus volumes.
Cerebellum (cm^3)	Sum of the left cerebellum, right cerebellum, and cerebellar vermis volumes extracted from segmentation.
Percent Ventricular Asymmetry	Computed as $\left(\frac{\text{larger lateral ventricle volume}}{\text{smaller lateral ventricle volume}} - 1\right) \times 100$.
Volume Ratio (Ventricles/Parenchyma)	Sum of lateral, third, and fourth ventricle volumes divided by the parenchyma volume.

List of Abbreviations

MRI Magnetic Resonance Imaging

ADC Apparent Diffusion Coefficient

IVIM Intravoxel Incoherent Motion

VAE Variational Auto-encoder

VQ-VAE Vector Quantization Variational Auto-encoder

FGR Fetal Grouth Restruction

LMICs low- and middle-income countries

AI Artificial Intelligence

HICs High Income Countries

DOHaD Developmental Origins of Health and Disease

DDVD diffusion-derived vascular density

GA Gestational Age

ML Machine Learning

MAE Mean Absolute Error

MSE Mean Squared Error

RMSE Root Mean Squared Error

GI Gyrification Index

GLDM Gray Level Dependence Matrix

SSIM Structural Similarity Index Measure

GLRLM Gray Level Run Length Matrix

SHAP Shapley Additive Explanations

CHD congenital heart defect

List of Figures

List of Tables

Bibliography

[] TinyGPU - NHR@FAU HPC documentation. URL: https://doc.nhr.fau.de/clusters/tinygpu/ (visited on 04/15/2025) (cited on p. 7).

[25a] AIM-harvard/pyradiomics, April 19, 2025. URL: https://github.com/AIM-Harvard/pyradiomics (visited on 04/22/2025). original-date: 2015-09-15T15:11:28Z (cited on pp. 14, 21).

[25b] InsightSoftwareConsortium/ITKElastix, February 14, 2025. URL: https://github.com/InsightSoftwareConsortium/ITKElastix (visited on 02/18/2025). original-date: 2019-09-10T02:54:40Z (cited on p. 8).

[25c] SimpleITK: itk::simple::NormalizeImageFilter class reference, February 14, 2025. URL: https://simpleitk.org/doxygen/v2_3/html/classitk_1_1simple_1_1NormalizeImageFilter.html (visited on 02/19/2025) (cited on p. 8).

[25d] SimpleITK: itk::simple::ResampleImageFilter class reference, February 14, 2025. URL: https://simpleitk.org/doxygen/v2_4/html/classitk_1_1simple_1_1ResampleImageFilter.html (visited on 02/19/2025) (cited on p. 8).

[25e] SVRTK/auto-proc-svrtk, February 18, 2025. URL: https://github.com/SVRTK/auto-proc-svrtk (visited on 02/28/2025). original-date: 2023-09-09T08:19:29Z (cited on p. 10).

[Al +23] M. Al Saeedi and L. O. Lerman. The use of magnetic resonance imaging to detect placental microvascular rarefaction that heralds preeclampsia. *Hypertension (Dallas, Tex. : 1979)*, 80(8):1668–1670, August 2023. ISSN: 0194-911X. DOI: 10.1161/HYPERTENSIONAHA.123.21324. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10372807/ (visited on 02/10/2025). Number: 8 (cited on pp. 4, 14).

[Ala+23] N. Alakonda, N. Patil, R. Yaliwal, A. Biradar, S. Shiragur, S. Kori, and S. Mathapati. A cross-sectional study to evaluate the impact of placental location on maternal and fetal outcomes. *Cureus*, June 12, 2023. ISSN: 2168-8184. DOI: 10.7759/cureus.40291. URL: https://www.cureus.com/articles/145090-a-cross-sectional-study-to-evaluate-the-impact-of-placental-location-on-maternal-and-fetal-outcomes (visited on 10/01/2024) (cited on p. 2).

[Aur+12] L. Auret and C. Aldrich. Interpretation of nonlinear relationships between process variables by use of random forests. *Minerals Engineering*, 35:27–42, August 2012. ISSN: 08926875. DOI: 10.1016/j.mineng.2012.05.008. URL: https://linkinghub.elsevier.com/retrieve/pii/S0892687512001987 (visited on 04/14/2025) (cited on p. 10).

[Bin+21] C. Binder, J. Buchmayer, A. Thajer, V. Giordano, V. Schmidbauer, K. Harreiter, K. Klebermass-Schrehof, A. Berger, and K. Goeral. Association between fat-free mass and brain size in extremely preterm infants. *Nutrients*, 13(12):4205, November 24, 2021. ISSN: 2072-6643. DOI: 10.3390/nu13124205. URL: https://www.mdpi.com/2072-6643/13/12/4205 (visited on 10/08/2024). Number: 12 (cited on p. 10).

[Bre01] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, October 1, 2001. ISSN: 1573-0565. DOI: 10.1023/A:1010933404324. URL: https://doi.org/10.1023/A:1010933404324 (visited on 03/04/2025). Number: 1 (cited on p. 4).

[Che+16] T. Chen and C. Guestrin. XGBoost: a scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA. Association for Computing Machinery, August 13, 2016. ISBN: 978-1-4503-4232-2. DOI: 10.1145/2939672.2939785. URL: https://dl.acm.org/doi/10.1145/2939672.2939785 (visited on 03/04/2025) (cited on p. 4).

[Cou21] Countdown to 2030 for Women's, Children's and Adolescent's Health. Maternal and newborn health in low- and middle-income countries. 2021, AlignMNH, April 21, 2021, page 19. (Visited on 01/26/2025). Issue: 2021 (cited on p. 2).

[Cov+67] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, January 1967. ISSN: 1557-9654. DOI: 10.1109/TIT.1967.1053964. URL: https://ieeexplore.ieee.org/

document/1053964 (visited on 03/04/2025). Number: 1 Conference Name: IEEE Transactions on Information Theory (cited on p. 4).

[Cro+24] D. Cromb, P. J. Slator, M. Hall, A. Price, D. C. Alexander, S. J. Counsell, and J. Hutter. Advanced magnetic resonance imaging detects altered placental development in pregnancies affected by congenital heart disease. *Scientific Reports*, 14(1):12357, May 29, 2024. ISSN: 2045-2322. DOI: 10.1038/s41598-024-63087-8. URL: https://www.nature.com/articles/s41598-024-63087-8 (visited on 09/24/2024). Number: 1 (cited on p. 3).

[Dew+19] M. C. Dewan, A. Rattani, R. Mekary, L. J. Glancz, I. Yunusa, R. E. Baticulon, G. Fieggen, J. C. Wellons, K. B. Park, and B. C. Warf. Global hydrocephalus epidemiology and incidence: systematic review and meta-analysis. *Journal of Neurosurgery*, 130(4):1065–1079, April 1, 2019. ISSN: 1933-0693. DOI: 10.3171/2017.10.JNS17439. Number: 4 (cited on p. 3).

[Do+19] Q. N. Do, M. A. Lewis, A. J. Madhuranthakam, Y. Xi, A. A. Bailey, R. E. Lenkinski, and D. M. Twickler. Texture analysis of magnetic resonance images of the human placenta throughout gestation: a feasibility study. *PLOS ONE*, 14(1):e0211060, January 22, 2019. G. E. Woloschak, editor. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0211060. URL: https://dx.plos.org/10.1371/journal.pone.0211060 (visited on 09/27/2024). Number: 1 (cited on pp. 4, 14).

[EFC24] EFCNI. Estimating pregnant women's' risk of pre-eclampsia using artificial intelligence – with a focus on expectant mothers around the world. EFCNI. 2024. URL: https://www.efcni.org/news/estimate-risk-of-pre-eclampsia-using-ai/ (visited on 01/26/2025) (cited on p. 3).

[Gro+06] R. Grossman, C. Hoffman, Y. Mardor, and A. Biegon. Quantitative MRI measurements of human fetal brain development in utero. *NeuroImage*, 33(2):463–470, November 2006. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2006.07.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S1053811906007816 (visited on 11/28/2024). Number: 2 (cited on pp. 10, 30).

[Hal+24] M. Hall, A. de Marvao, R. Schweitzer, D. Cromb, K. Colford, P. Jandu, D. P. O'Regan, A. Ho, A. Price, L. C. Chappell, M. A. Rutherford, L. Story, P. Lamata, and J. Hutter. Preeclampsia associated differences in the placenta, fetal brain, and maternal heart can be demonstrated antenatally: an observational cohort

study using MRI. *Hypertension (Dallas, Tex.: 1979)*, 81(4):836–847, April 2024. ISSN: 1524-4563. DOI: 10.1161/HYPERTENSIONAHA.123.22442. Number: 4 (cited on pp. 7, 14).

[Hil+25] B. S. Hilabi, S. A. Alghamdi, and M. Almanaa. Impact of magnetic resonance imaging on healthcare in low- and middle-income countries. *Cureus*, 15(4):e37698, January 26, 2025. ISSN: 2168-8184. DOI: 10.7759/cureus.37698. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10112545/ (visited on 01/26/2025). Number: 4 (cited on p. 2).

[Jam+21] G. James, D. Witten, T. Hastie, and R. Tibshirani. *An introduction to statistical learning: with applications in R*. Springer texts in statistics. Springer, New York, NY, second edition edition, 2021. 607 pages. ISBN: 978-1-07-161420-4 978-1-07-161417-4 (cited on p. 11).

[Lac19] S. Lacagnina. The developmental origins of health and disease (DOHaD). *American Journal of Lifestyle Medicine*, 14(1):47, October 11, 2019. DOI: 10.1177/1559827619879694. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC6933571/ (visited on 02/05/2025). Number: 1 (cited on p. 2).

[Leu21] K.-Y. Leung. Applications of advanced ultrasound technology in obstetrics. *Diagnostics*, 11(7):1217, July 6, 2021. ISSN: 2075-4418. DOI: 10.3390/diagnostics11071217. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8306830/ (visited on 02/10/2025). Number: 7 (cited on p. 1).

[Lev01] D. Levine. Ultrasound versus magnetic resonance imaging in fetal evaluation: *Topics in Magnetic Resonance Imaging*, 12(1):25–38, February 2001. ISSN: 0899-3459. DOI: 10.1097/00002142-200102000-00004. URL: http://journals.lww.com/00002142-200102000-00004 (visited on 12/21/2024). Number: 1 (cited on pp. 1, 2).

[Li+23] S. Li, Q. Yang, S. Niu, and Y. Liu. Effectiveness of remote fetal monitoring on maternal-fetal outcomes: systematic review and meta-analysis. *JMIR mHealth and uHealth*, 11:e41508, February 22, 2023. ISSN: 2291-5222. DOI: 10.2196/41508. URL: https://mhealth.jmir.org/2023/1/e41508 (visited on 12/21/2024) (cited on p. 2).

[Li+24] C. Li, S. Yu, and Y. Cui. Individual brain parcellation: review of methods, validations and applications, July 1, 2024. DOI: 10.48550/arXiv.2407.00984.

arXiv: 2407.00984[q-bio]. URL: http://arxiv.org/abs/2407.00984 (visited on 04/14/2025) (cited on p. 10).

[May+24] J. Mayrink and Z. S. N. Reis. Pre-eclampsia in low and middle-income settings: what are the barriers to improving perinatal outcomes and evidence-based recommendations? *International Journal of Gynaecology and Obstetrics: The Official Organ of the International Federation of Gynaecology and Obstetrics*, 164(1):33–39, January 2024. ISSN: 1879-3479. DOI: 10.1002/ijgo.14913. Number: 1 (cited on p. 3).

[Mol24] C. Molnar. 6 linear regression – interpretable machine learning. Github. April 22, 2024. URL: https://christophm.github.io/interpretable-ml-book/limo.html (visited on 04/14/2025) (cited on p. 10).

[Nap+20] R. Napolitano, M. Molloholli, V. Donadono, E. O. Ohuma, S. Z. Wanyonyi, B. Kemp, M. K. Yaqub, S. Ash, F. C. Barros, M. Carvalho, Y. A. Jaffer, J. A. Noble, M. Oberto, M. Purwar, R. Pang, L. Cheikh Ismail, A. Lambert, M. G. Gravett, L. J. Salomon, Z. A. Bhutta, S. H. Kennedy, J. Villar, A. T. Papageorghiou, and International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). International standards for fetal brain structures based on serial ultrasound measurements from fetal growth longitudinal study of INTERGROWTH-21st project. *Ultrasound in Obstetrics & Gynecology*, 56(3):359–370, 2020. ISSN: 1469-0705. DOI: 10.1002/uog.21990. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/uog.21990 (visited on 03/18/2025). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/uog.21990 (cited on p. 10).

[Nga+24] L. Ngartera, M. A. Issaka, and S. Nadarajah. Application of bayesian neural networks in healthcare: three case studies. *Machine Learning and Knowledge Extraction*, 6(4):2639–2658, December 2024. ISSN: 2504-4990. DOI: 10.3390/make6040127. URL: https://www.mdpi.com/2504-4990/6/4/127 (visited on 03/04/2025). Number: 4 Publisher: Multidisciplinary Digital Publishing Institute (cited on p. 4).

[Nov+21] M. I. Novis, A. P. C. Moura, A. D. P. F. Watanabe, L. C. L. E. Pereira, G. Warmbrand, and G. D'Ippolito. Placental magnetic resonance imaging: normal appearance, anatomical variations, and pathological findings. *Radiologia Brasileira*, 54(2):123–129, April 2021. ISSN: 1678-7099. DOI: 10.1590/0100-

3984.2020.0010. URL: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-39842021000200123&tlng=en (visited on 09/24/2024). Number: 2 (cited on pp. 1, 2).

[OCo+12] S. C. O'Connor, V. J. Rooks, and A. B. Smith. Magnetic resonance imaging of the fetal central nervous system, head, neck, and chest. *Seminars in Ultrasound, CT and MRI*, 33(1):86–101, February 2012. ISSN: 08872171. DOI: 10.1053/j.sult.2011.10.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S0887217111001375 (visited on 12/21/2024). Number: 1 (cited on p. 1).

[Par+24] C. Parmer, P. Duval, and A. Johnson. A data and analytics web app framework for python, no JavaScript required. Version 2.18.2, November 2024. DOI: 10.5281/zenodo.14182630. URL: https://github.com/plotly/dash (visited on 04/14/2025) (cited on p. 15).

[Pat+22] V. Patil, M. Tafheem, S. Melkundi, and S. Masimade. A comparative study of fetal MRI and antenatal ultrasound in prenatal diagnosis with histopathological correlation. *Journal of Fetal Medicine*, 09(3):83–90, December 2022. ISSN: 2348-1153, 2348-8859. DOI: 10.1007/s40556-022-00347-5. URL: http://www.thieme-connect.de/DOI/DOI?10.1007/s40556-022-00347-5 (visited on 02/05/2025). Number: 03/04 (cited on p. 1).

[Pin+22] W. H. Pinaya, P.-D. Tudosiu, R. Gray, G. Rees, P. Nachev, S. Ourselin, and M. J. Cardoso. Unsupervised brain imaging 3d anomaly detection and segmentation with transformers. *Medical Image Analysis*, 79:102475, July 2022. ISSN: 13618415. DOI: 10.1016/j.media.2022.102475. URL: https://linkinghub.elsevier.com/retrieve/pii/S1361841522001220 (visited on 12/13/2024) (cited on p. 11).

[Rec+24] F. Recker, U. Gembruch, and B. Strizek. Clinical ultrasound applications in obstetrics and gynecology in the year 2024. *Journal of Clinical Medicine*, 13(5):1244, February 22, 2024. ISSN: 2077-0383. DOI: 10.3390/jcm13051244. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10931841/ (visited on 12/26/2024). Number: 5 (cited on p. 1).

[Roe+07] N. M. Roelfsema, W. C. J. Hop, L. N. A. Van Adrichem, and J. W. Wladimiroff. Craniofacial variability index *in utero* : a three-dimensional ultrasound study. *Ultrasound in Obstetrics & Gynecology*, 29(3):258–264, March 2007. ISSN: 0960-7692, 1469-0705. DOI: 10.1002/uog.3904. URL: https://obgyn.onlinelibrary.

wiley.com/doi/10.1002/uog.3904 (visited on 12/21/2024). Number: 3 (cited on p. 1).

[Shi+20] Y. Shi, Y. Xue, C. Chen, K. Lin, and Z. Zhou. Association of gestational age with MRI-based biometrics of brain development in fetuses. *BMC Medical Imaging*, 20(1):125, December 2020. ISSN: 1471-2342. DOI: 10.1186/s12880-020-00525-9. URL: https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880-020-00525-9 (visited on 02/10/2025). Number: 1 (cited on p. 3).

[Sie24] H. Siemens. Future prize 2023 winners donate MRIScanner toPeruHospital. December 13, 2024. URL: https://www.siemens-healthineers.com/press/releases/peru-donation (visited on 04/22/2025) (cited on p. 34).

[Sit+15] N. A. Sitkin, D. Ozgediz, P. Donkor, and D. L. Farmer. Congenital anomalies in low- and middle-income countries: the unborn child of global surgery. *World Journal of Surgery*, 39(1):36–40, January 2015. ISSN: 0364-2313, 1432-2323. DOI: 10.1007/s00268-014-2714-9. URL: https://onlinelibrary.wiley.com/doi/10.1007/s00268-014-2714-9 (visited on 12/23/2024). Number: 1 (cited on p. 2).

[Sou+24] J. P. Souza, L. T. Day, A. C. Rezende-Gomes, J. Zhang, R. Mori, A. Baguiya, K. Jayaratne, A. Osoti, J. P. Vogel, O. Campbell, K. Y. Mugerwa, P. Lumbiganon, Ö. Tunçalp, J. Cresswell, L. Say, A. C. Moran, and O. T. Oladapo. A global analysis of the determinants of maternal health and transitions in maternal mortality. *The Lancet Global Health*, 12(2):e306–e316, February 2024. ISSN: 2214109X. DOI: 10.1016/S2214-109X(23)00468-0. URL: https://linkinghub.elsevier.com/retrieve/pii/S2214109X23004680 (visited on 12/23/2024). Number: 2 (cited on p. 2).

[Sun] Y. Sun. Ultrasound in obstetrics and gynecology: benefits, limitations and future developments. 09(2). Number: 2 (cited on p. 1).

[Ton+19] W. Tong and D. A. Giussani. Preeclampsia link to gestational hypoxia. *Journal of developmental origins of health and disease*, 10(3):322–333, June 2019. ISSN: 2040-1744. DOI: 10.1017/S204017441900014X. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6658216/ (visited on 04/10/2025) (cited on p. 31).

[Vah+23] F. Vahedifard, H. A. Ai, M. P. Supanich, K. K. Marathu, X. Liu, M. Kocak, S. M. Ansari, M. Akyuz, J. O. Adepoju, S. Adler, and S. Byrd. Automatic ventriculomegaly detection in fetal brain MRI: a step-by-step deep

learning model for novel 2d-3d linear measurements. *Diagnostics*, 13(14):2355, July 13, 2023. ISSN: 2075-4418. DOI: 10.3390/diagnostics13142355. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10378043/ (visited on 02/10/2025). Number: 14 (cited on p. 4).

[Zil+88] K. Zilles, E. Armstrong, A. Schleicher, and H.-J. Kretschmann. The human pattern of gyrification in the cerebral cortex. *Anatomy and Embryology*, 179(2):173–179, November 1988. ISSN: 0340-2061, 1432-0568. DOI: 10.1007/BF00304699. URL: http://link.springer.com/10.1007/BF00304699 (visited on 11/09/2024). Number: 2 (cited on p. 4).