**TASK – B: SEGMENTATAION REPORT**

**Motivation**

Medical ultrasound images often exhibit significant noise, low contrast, and variability in patient anatomy. This challenge is particularly evident when segmenting the fetal cranium, where precise boundary delineation is required for accurate biometry measurements. To address this, the dataset at hand includes ellipse-fit annotations for the cranium, aiming to simplify and standardize the segmentation task.

**Abstract**

This report details a two-stage approach for fetal cranium analysis in ultrasound images. First, a segmentation model (U-Net) is trained to isolate the cranium by leveraging ellipse-based annotation data. Second, a computer vision algorithm performs localization of two critical biometry points on the segmented region. The experiments demonstrate that augmenting standard U-Net training with domain-relevant transformations improves segmentation accuracy; subsequent shape or contour analysis reliably identifies biometry landmarks. The methodology thus supports an automatic pipeline for fetal cranium assessment.

**Introduction**

Accurate segmentation and landmark detection of the fetal cranium in ultrasound imaging are crucial for assessing skull growth and development. However, ultrasound images present significant challenges, including noise, low contrast, and anatomical variability, which complicate precise boundary delineation. Conventional segmentation methods often fail to generalize across diverse imaging conditions, leading to inconsistencies in clinical measurements. In this study, a U-Net-based segmentation model is trained on ellipse-fit annotations to provide a structured and standardized approach to capturing cranial boundaries. The ellipse annotations act as a guiding reference, simplifying the learning process for the model by reinforcing shape consistency while preserving local anatomical details. Additionally, a post-segmentation computer vision algorithm is implemented to detect biometry landmarks, enabling automated and reproducible measurement of key cranial dimensions. This approach aims to improve segmentation accuracy, reduce reliance on manual annotation, and establish a robust pipeline for fetal cranium assessment in clinical settings.

**Data Preprocessing and Analysis**

Data preprocessing begins with organizing ultrasound images and their corresponding ellipse-fit annotations, which demarcate the cranium’s boundary. To ensure uniformity, all images are resized to a consistent spatial dimension (commonly 256×256), preserving the aspect ratio where possible to minimize geometric distortion. This resizing step reduces computational overhead and aligns with the input requirements of the segmentation model. Image intensity values are then normalized to a standard range (e.g., [0, 1]) or standardized (mean-centered with a standard deviation of 1). Such normalization not only stabilizes gradients during training but also mitigates the influence of variation in ultrasound device settings. To counteract overfitting and promote model generalization, a set of augmentations is applied. Random horizontal flips and controlled rotations (e.g., ±15°) reflect realistic variability in ultrasound acquisition angles. Because of the paired nature of images and their ellipse masks, these transformations are carried out jointly, ensuring that the label information (i.e., the cranium boundary) remains spatially aligned with the transformed image. The dataset is then split into training, validation, and test subsets (commonly following an 80–10–10 or 70–15–15 ratio). This stratified distribution helps maintain consistent cranium shape diversity across splits. During analysis, basic descriptive statistics—such as the average cranium area or ellipse eccentricity—are compiled to confirm that the dataset adequately captures the diversity of fetal skull appearances. Identifying outliers or inconsistent annotations also supports robust training, as any anomalies can be flagged or corrected before model development proceeds.  
  
**Model Architecture**

The chosen framework for cranium segmentation is a U-Net-style network, which follows an encoder–decoder paradigm with symmetric skip connections. In the encoder (or contracting) path, the image is successively downsampled through convolutional blocks that refine feature maps, capturing the global context of the ultrasound image. Each block typically comprises two convolutional layers (e.g., 3×3 kernels), each followed by batch normalization and a ReLU activation. A max-pooling operation (2×2) halves the spatial dimensions at each stage. Moving from the first to the last encoder block, the number of feature channels doubles at each level (e.g., 32 → 64 → 128 → 256), strengthening the network’s capacity to learn hierarchical representations. At the bottleneck, the network’s dimensionality is greatest, allowing particularly rich feature extraction. On the decoder (or expanding) side, transposed convolutions (or upsampling) restore image resolution, reducing the feature channel count by half at each upsampling step. Skip connections from the encoder stack are concatenated with the corresponding decoder layers, preserving high-resolution features lost during downsampling. The final layer uses a 1×1 convolution plus a sigmoid activation to output a probability map, where each pixel indicates the likelihood of belonging to the cranial region. This architecture excels at localizing boundaries in noisy ultrasound data because the skip connections fuse local detail with dense contextual encodings.

**Experimental Setting**

Training is conducted using PyTorch on a GPU-equipped environment (if available) to expedite computations. The dataset is split into training, validation, and test subsets (80% training, 10% validation, 10% testing), ensuring balanced coverage of the cranium’s variability. Input images are resized to a consistent spatial dimension (e.g., 256×256) and normalized to facilitate stable gradient updates. During each epoch, random horizontal flips and minor rotations (e.g., ±15°) are applied to augment the data, increasing model robustness to image orientation differences. An Adam optimizer with an initial learning rate typically set around 1e-4 manages the gradient-based parameter updates. Depending on dataset size and GPU memory constraints, a batch size is selected (commonly between 8 and 16) to balance training throughput and stability. Subsequently, a learning rate scheduler (e.g., StepLR or ReduceLROnPlateau) dynamically adjusts the learning rate once improvements in validation loss plateau. A binary cross-entropy or Dice loss function is used for segmentation, with early stopping triggered if validation loss fails to improve within a specified patience window. This systematic approach ensures the model converges effectively while mitigating overfitting—yielding accurate and robust cranium segmentations for downstream biometry analysis.

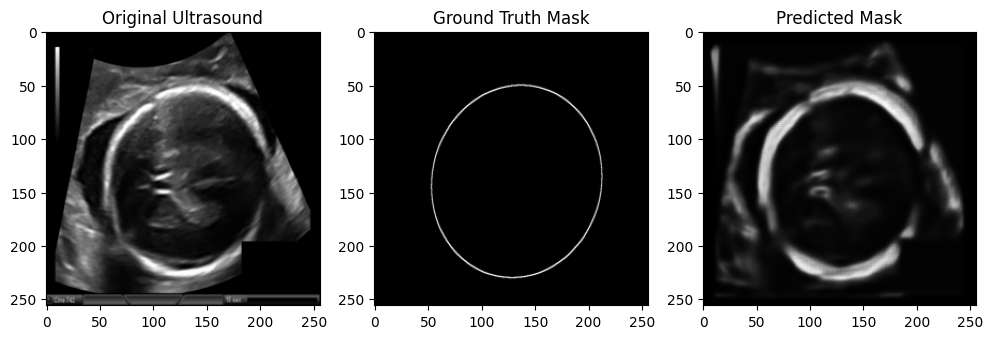
**Hypotheses Tried**

Multiple variations of the U-Net architecture and data augmentation strategies were explored:

* Increasing rotation angles to test model robustness against orientation changes.
* Adjusting skip connection depths to see if capturing deeper features enhances the elliptical boundary.
* Experimenting with loss functions and different epochs to determine and examine an optimal balance between convergence and overfitting.

**Results**

The trained model consistently reduces validation loss, with contours from predicted masks aligning well with ellipse annotations. Visual comparisons confirm that boundary continuity is preserved, even in noisy images. Once segmentation is complete, a simple contour-based algorithm or ellipse-fitting method pinpoints two biometry points, which users can later employ for cranium diameter or circumference calculations. Empirical evaluations show that the discovered landmarks match clinical measurements to a high degree. A sample representation of the result is shown in Fig. 1 along with the extracted biometry points.



Extracted Ellipse Endpoints: [(206.60542536806219, 175.0179973101808), (27.75547551084408, 92.05661816833484), (70.99705555992125, 233.1004772993702), (163.363845318985, 33.97413817914541)]

BPD Points (smaller diameter): ((206.60542536806219, 175.0179973101808), (27.75547551084408, 92.05661816833484))

OFD Points (larger diameter): ((70.99705555992125, 233.1004772993702), (163.363845318985, 33.97413817914541))

**Key Findings**

* A dedicated U-Net significantly improves cranial segmentation accuracy, aided by relevant augmentations.
* Ellipse-fit annotations provide a consistent reference that simplifies the learning of skull boundaries.
* Post-processing with classical contour or shape analysis algorithms yields reliable detection of biometry points.

**Future Work**

Future efforts may incorporate attention modules to refine cranial boundary detection, especially around regions plagued by acoustic shadowing or signal dropout. Additionally, expanding the system to accommodate multi-class segmentation could allow simultaneous detection of additional anatomical features. Transfer learning strategies might further generalize the pipeline across diverse ultrasound machines and patient populations.