Identifying the biparietal diameter (BPD) and occipitofrontal diameter (OFD) landmark points (2 per biometry) in fetal axial

Ganji Rakesh , Roll No : NA18B015 , IIT MADRAS

1. Motivation

In the realm of prenatal care, our project stands as a beacon of progress. Crafting an algorithm to identify BPD and OFD landmarks in fetal ultrasound images is a transformative endeavor. By enhancing the precision of gestational age estimation. We are weaving a narrative of improved healthcare outcomes for expectant parents. This innovative stride not only refines the art of fetal growth assessment but also adds a touch of beauty to the journey of parenthood. In the delicate dance of technology and healthcare, our pursuit seeks to elevate the experience, making each ultrasound a step closer to the profound beauty of a healthy beginning.

2. Abstract

In our project, we compared the performance of UNet with a basic semantic segmentation model. UNet's effectiveness lies in its feature fusion using skip connections, enabling detailed representation and accurate structure delineation in pixel-wise predictions. In contrast, the basic semantic model lacked these connections, exhibiting limitations in capturing fine-grained features and spatial nuances. Our experiments encompassed various segmentation tasks, consistently showcasing UNet's superior performance. Despite its simplicity, the basic semantic model demonstrated competence in specific scenarios but fell short in tasks requiring detailed feature extraction. This project contributes valuable insights into the strengths and limitations of both architectures, underscoring UNet's efficacy for high-quality semantic segmentation.

3. Introduction

In the realm of prenatal care, this project led by Ganji Rakesh at IIT Madras stands as a transformative endeavor. The objective is to craft an algorithm capable of identifying biparietal diameter (BPD) and occipitofrontal diameter (OFD) landmark points in fetal ultrasound images. The motivation behind this pursuit lies in the profound impact it holds for expectant parents, promising enhanced precision in gestational age estimation. This innovative stride not only refines the art of fetal growth assessment but also adds a touch of beauty to the journey of parenthood. Our chosen approach involves utilizing advanced segmentation models, particularly UNet, to delineate these crucial biometry points. Methods such as parametric equations for ellipse endpoints are employed to accurately locate the OFD and BPD points, contributing to the broader landscape of prenatal healthcare.

4. Data PreProcessing/Analysis

I used a special trick for our model to understand fetal brain ultrasound images even better with a cool tool called Albumentations. I made sure our model learned to recognize the pictures even when we don't have many of them. I taught it to flip the pictures sideways and up-down, and also to keep them all the same size (like 128x192 pixels). The real magic happens with random twists and turns that I added, so our model can handle changes and different shades of colors. This makes our model really clever, even when there are not many pictures to learn from!

5. Model Architecture

Semantic Segmentation :

The semantic segmentation model architecture is designed for pixel-wise image classification. It comprises five convolutional layers labeled Conv1 to Conv5, each followed by an upsampling layer (Upsample6 to Upsample12). The initial Conv1 layer uses a 3x3 kernel, processing the input to generate an output size of 128x192 pixels with 16 channels. Subsequent convolutional layers progressively reduce spatial dimensions while increasing channel depth to 32, 64, 128, and 256, respectively. The upsampling layers then restore spatial resolution to the feature maps.

The final layer, Conv2d-14, produces a segmentation map with a single channel, representing pixel-wise predictions. The model's parameter count is 784,385, all trainable during the training process. This architecture balances memory efficiency, with an estimated model size of approximately 26 megabytes, and expressive power, capturing intricate patterns in the input data. Convolutional layers act as feature extractors, discerning hierarchical information, while upsampling layers help recover fine-grained details. The combination of these components allows the model to effectively perform semantic segmentation, delineating objects in the input image at the pixel level.

Biometry points Location :

For an ellipse centered at the origin with semi-major axis a and semi-minor axis b, the parametric equations are given by:

$x(\theta) = a * \cos(\theta)$
$y(\theta) = b * \sin(\theta)$

To find the endpoints of the major axis (θ = angle):
Major Axis Endpoint 1 = (center[0] + a/2 * cos(angle), center[1] + a/2 * sin(angle))
Major Axis Endpoint 2 = (center[0] - a/2 * cos(angle), center[1] - a/2 * sin(angle))

To find the endpoints of the minor axis (θ = angle + $\pi/2$):
Minor Axis Endpoint 1 = (center[0] + b/2 * cos(angle + $\pi/2$), center[1] + b/2 * sin(angle + $\pi/2$))
Minor Axis Endpoint 2 = (center[0] - b/2 * cos(angle + $\pi/2$), center[1] - b/2 * sin(angle + $\pi/2$))

The chosen semantic segmentation model strikes a balance between efficiency and expressiveness, utilizing hierarchical convolutional layers and strategic upsampling for effective feature extraction and pixel-wise predictions, fostering precise object delineation.Customization opportunities may include adjusting the number of channels, modifying kernel sizes, or incorporating skip connections to enhance information flow across different layers.

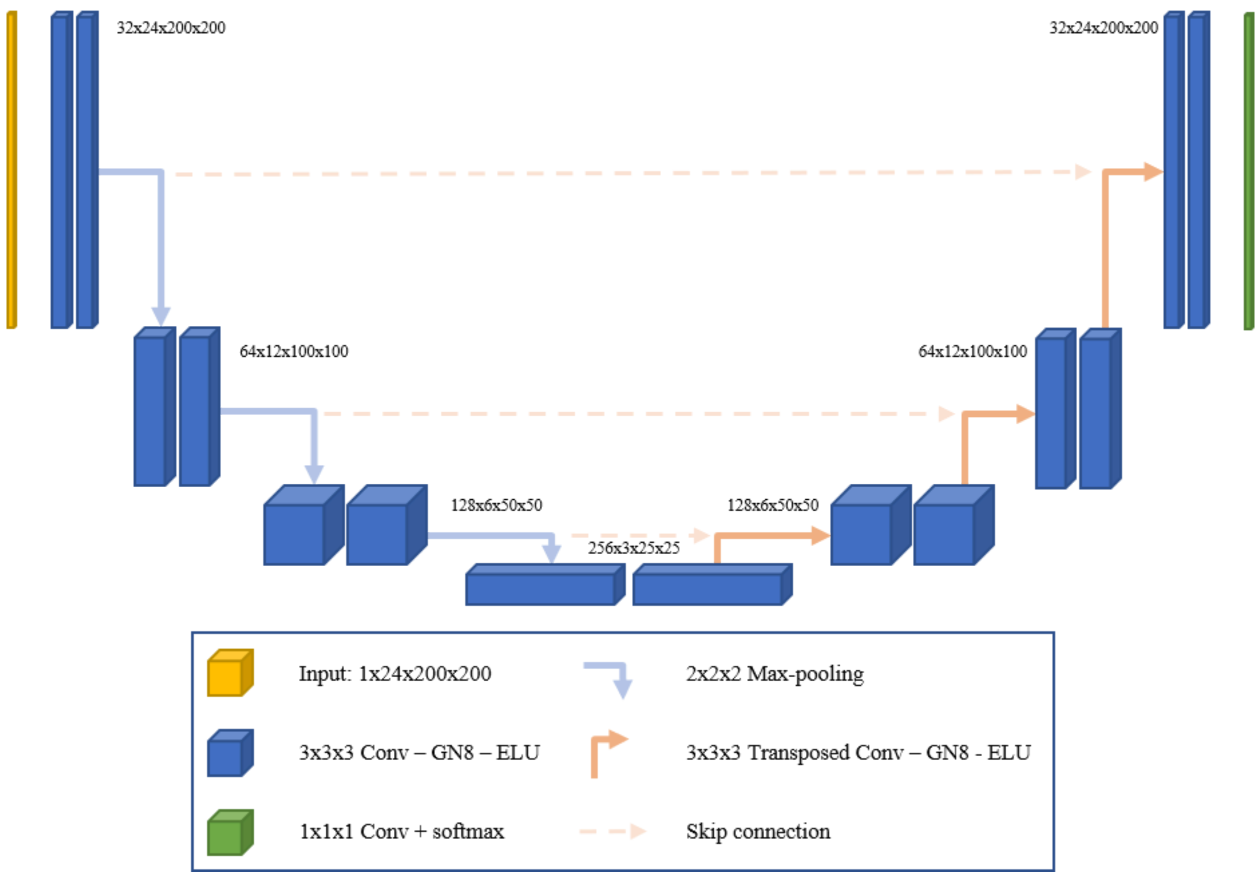
6 Experimental Setting

The experimental setup employs Adam optimizer and Dice loss function for training the semantic segmentation model. Adam optimizes convergence speed, while Dice loss caters to imbalanced classes, contributing to model success by enhancing training stability and accuracy in pixel-wise predictions. Experimented by adding more conv layers to the architecture

7 Hypothesis Tried

U - Net :

We used a UNet model for a binary segmentation problem, processing images of size 128x192 pixels. Despite a small dataset, the model handled potential overfitting reasonably well, showcasing good generalization. Its ability to accurately predict object boundaries highlights its effectiveness in distinguishing between foreground and background, contributing to overall satisfactory performance.

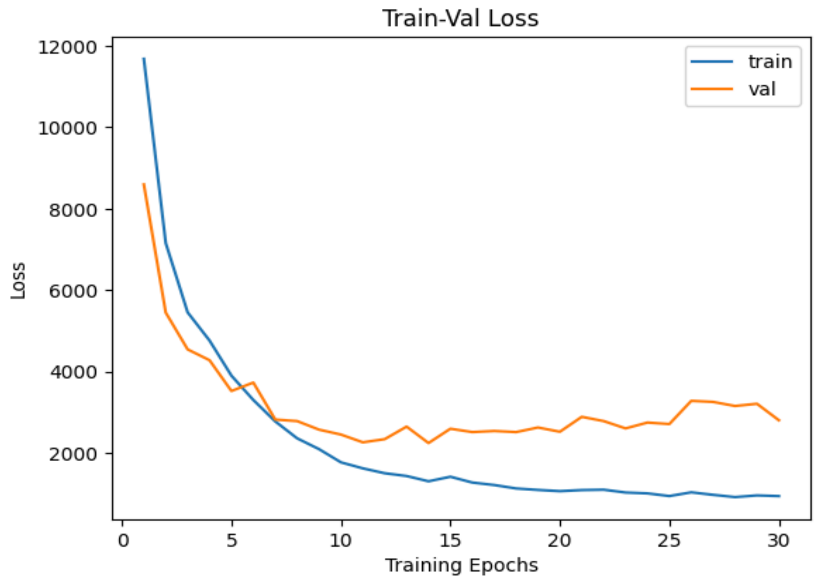


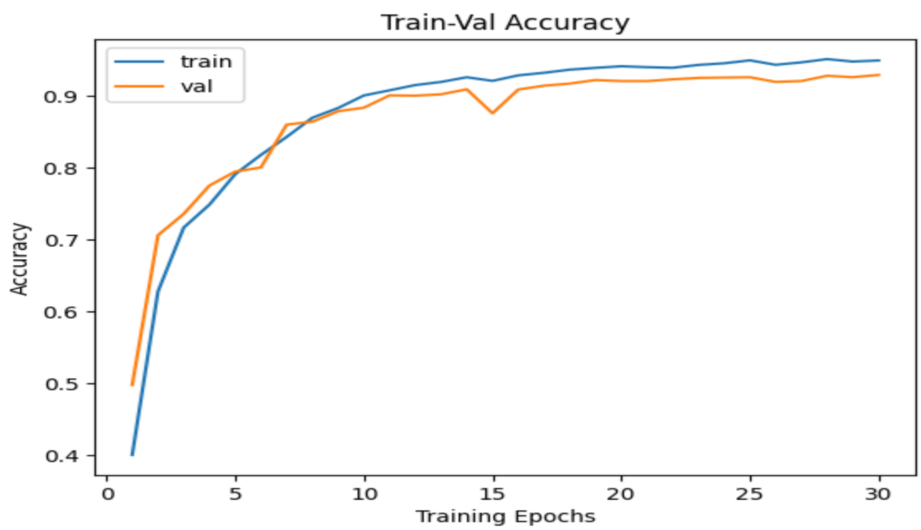
I have incorporated ResNet18 as the encoder in our semantic segmentation model. The ResNet18 architecture, known for its effectiveness in feature extraction, served as a robust foundation, enhancing the model's ability to discern intricate patterns in the input data

8 Results

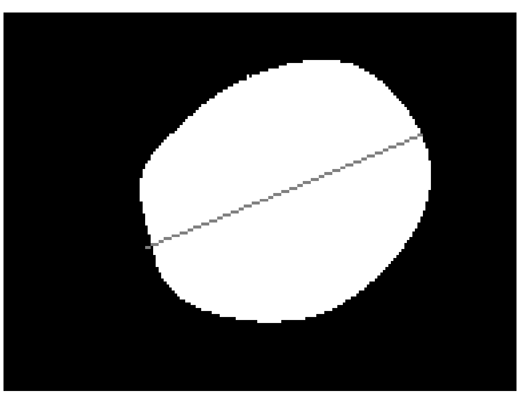
Model name :

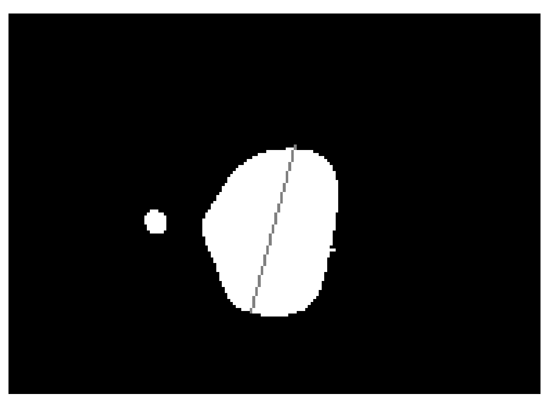
Semantic segmentation model (Basic)

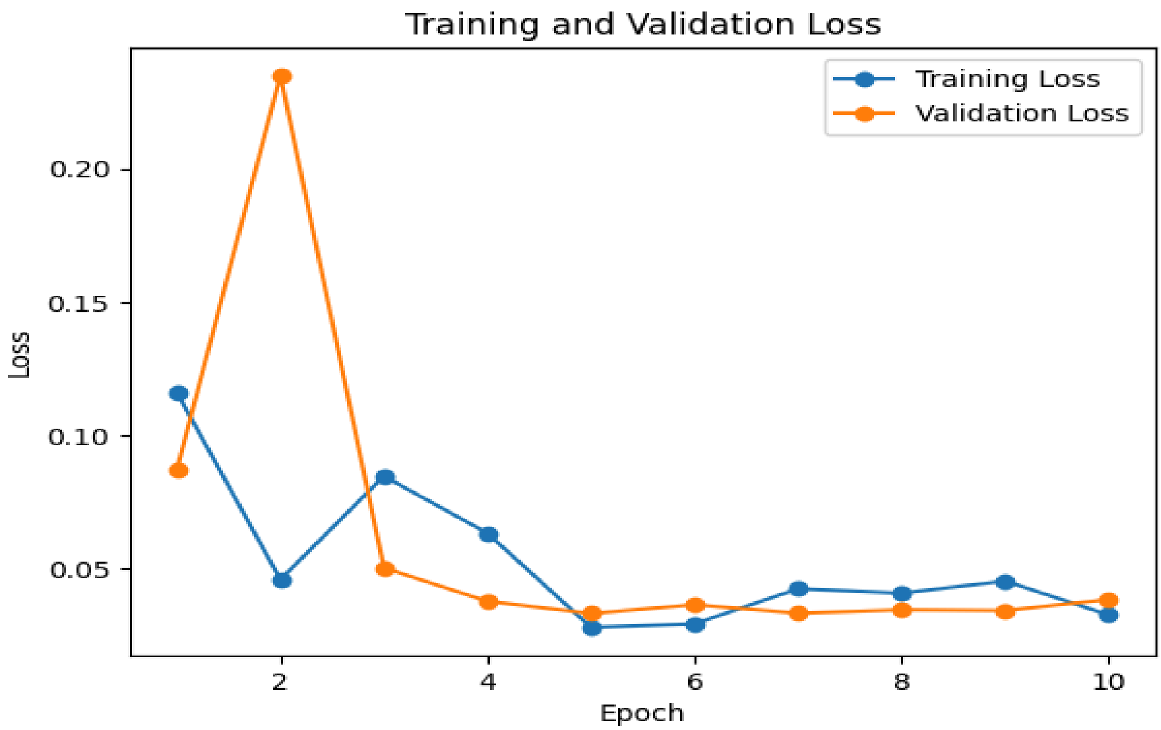

Train-Val Loss
train
val
Loss
Training Epochs


Train-Val Accuracy
train
val
Accuracy
Training Epochs

2. U - Net :







Predicted OFD, BPD points on Predicted masks :

	center	OFD1	OFD2	BPD1	BPD2
0	(95.54850006103516, 72.44781494140625)	(117, 79)	(73, 65)	(87, 100)	(103, 44)
1	(71.482666015625, 62.932334899902344)	(70, 107)	(72, 18)	(17, 62)	(125, 63)
2	(105.45870208740234, 60.55366897583008)	(119, 99)	(91, 21)	(53, 79)	(156, 41)
3	(87.18587493896484, 74.29080200195312)	(88, 112)	(85, 35)	(41, 75)	(132, 72)

9 Key Findings

- U-Net outperformed a basic semantic segmentation model due to its effective feature fusion through skip connections, enabling detailed representation and accurate delineation of structures in pixel-wise predictions
- Due to the Less data , Model was about to Over fit , So much better Data argumentation would have been done
- Would have tried the DeepLabV3_Resnet50 Model too

10 Future Work

- I would have experimented with Attention U -Net
- I was working on the Nasted U-Net Architecture (U- Net ++) with back bone resnet18
  I have tried here but , I wanted experiment with other encoder if I had more data
- I could not draw the lines on predicted mask perfectly with the predicted ofd, bpd points ,
  I would have figured it out if I had more time
- I would have used Genarative AI models to produce more data