EVALUATION OF THE FETAL HEAD CIRCUMFERENCE FROM ULTRASOUND IMAGES DLA U-NET

MOHANA PRIYAG¹, DR.S.K.B.SANGEETHA²

¹Research Scholar, Department of Computer Science and Engineering, SRMIST, Vadapalani, Chennai, India
²Assistant Professor (Sr. G), Department of Computer Science and Engineering, SRMIST, Vadapalani, Chennai, India
E-mail: ¹mg8504@srmist.edu.in, ²sangeets8@srmist.edu.in

ABSTRACT

This study introduces the DLA U-Net deep technique architecture in order to segment images method, which is specifically created for autonomously segmenting fetal ultrasonography images and head circumference (HC) biological dimension. To improve segmentation performance and accuracy, this method applies the consideration technique and deep supervision approach to U-Net representations. The evaluation of the proposed U-Net deep learning architecture (DLA) was conducted with the HC18 evaluation set, consisting of 355 cases. Four indicators of performance, namely the HC Mean Absolute Error (MAE), Hausdorff distance (HD), Dice similarity coefficient (DSC), and HC difference (DF) were used to assess the method's performance. U-Net deep learning architecture achieved a DSC of 97.94%, DF of 0.09 2.49 mm, MAE of 1.78 1.69 mm, and HD of 1.30 0.79 mm, according to experimental results. Gestational age (GA) was resolved by means of crown-rump length (CRL) dimension as the reference. GA estimations showed a mean difference of 0.5 ± 4.2, 0.3 ± 4.8 and 2.4 ± 12.3 days. The proposed U-Net architecture demonstrated superior or comparable segmentation performance when contrasted with contemporary techniques published in the literature. Hence, the introduced DLA U-Net can be considered as a viable approach for the delineation of fetal ultrasonography images and head circumference (HC) biological measurement.

Keywords: Fetal ultrasound, DSC, Image segmentation, DLA, Head circumference, DF, Deep learning, U-Net

1. INTRODUCTION

Medical imaging is a robust procedure utilized for both medicinal and diagnostic reasons. Various image-processing approaches are employed, such as computed tomography (CT) scanners, ultrasound (US), medical radiography, and magnetic resonance imaging (MRI). The fact that Ultrasound doesn't emit radiation, affordable, and real-time makes it a popular choice for diagnosing pregnancies. The accepted method for determining gestational age and fetal development has been to use fetal biometry measurements taken during US scans. Secondly and final trimesters, in its entirety most crucial biometrics is head circumference (HC). Monitoring gestational age (GA) is essential for identifying prenatal hazard elements and developing a correct diagnosis plan with newborn issues. Because it is affordable, radiation-free, and allows for real-time measurement, ultrasonography is typically utilized to assess the growth of neonates and GA. Fetal head circumference (HC) measurements from ultrasonography images are one method used for estimating GA, particularly for fetuses in the secondly and final trimesters. Fetal (HC) measurements from ultrasound images are one method for estimating GA, particularly for fetuses during the secondly and final trimesters. During medical settings, HC is manually the context of ultrasonography images, calculated through establishing a stripe among the two locations after each main then minor elliptical axis. The circumference of an ellipse that serves as a representation of the HC is then determined using the two pieces of information.

With advancing gestational age, low liquor, anterior placenta position, assisted vaginal birth, and those with elevated fetal HC, the measurement inaccuracy increased. Specific clinical situations, such as monitoring pregnancies accompanied by
fetal growth limitation, possible variations in fetal head development, and labor outcome, may be significantly impacted by the measurement error [1]. This article [2] explains how to measure HC from US photos automatically. In order to localize the HC ROI, we created and studied a structure for acquiring knowledge that takes preceding information into account. The RF-based identification within the ROI classifiers was significantly facilitated by the preexisting information of GA and screening complexity. According toward the exploratory findings on 145 embryonic head pictures, using phase symmetry to locate the fetal skull's center line and match the HC ellipse with ElliFit for assessment surpasses more conventional approaches in relations of accurateness and effectiveness of the HC measurement.

In contrast, this study [3] to use convolutional neural networks (CNN) toward ration the head circumference precise, deprived of the need for manually labeled segmented images or hand-crafted features. The rationale behind this theory is that the CNN will eventually train itself to locate and recognize the contour of the head. Examined three loss functions and various regression CNN architectures. According to experimental findings, the MSE loss function and the deeper Reg-ResNet50 model outperformed each other. Results are encouraging because the top models produce error that is comparable to the variability in manual measurement; however, CNN-based regressors still have space for improvement, especially when compared to the precision of segmentation-based techniques.

For autonomous HC measurement, a multiscale light convolutional neural network is suggested in this research [4]. A light convolutional neural network's segmentation accuracy and HC assessments done on an ultrasound dataset of a fetus in various pregnancy trimesters are equivalent to those of deep convolutional neural networks, according to experimental results. They [5] performed preliminary analysis methods to eradicate noise from the RoI then choose the best pixels that indicate a fetal head offer an improved rapid elliptical fitting method based on ElliFit. The suggested system [6] consists of a regression CNN for precisely delineating the HC and a regional CNN head localization and centering approach. While we suggest a distance-based training technique to simulate regression CNN, the first CNN is trained using transfer learning.

In this research [17], we introduced Deeply supervised attention-gated (DAG) V-Net, a revolutionary deep learning technique, was suggested for automatically segmenting fetal ultrasonography images then measuring head circumference (HC). The consideration mechanism and the deep supervision approach were both introduced obsessed by the V-Net architecture using this technique. The DAG V-Net showed enhanced performance in accurately segmenting fetal ultrasound images and predicting head circumference measures after combining these two approaches. The multi-scale loss equation has been added by deep supervision. As a result [8] of the statistic that the normal range in fetal growth rises with gestational age, it was demonstrated that it is crucial to divide the outcomes into each trimester. The findings from each trimester were assessed independently for the first time under this system. The normal range in fetal growth rises with gestational age, it was demonstrated that it is crucial to divide the outcomes into each trimester. Probably the fetal skull readily seen in initial trimester, making computerized HC determination is a more challenging task. However, the GA can be approximated more precisely in the first trimester.
The technique works much better in the first and second trimesters than a medical investigator, although the 2D standard plane must still be obtained.

2. RELATED WORKS

In ultrasound (US) imaging, determining head circumference frequently entails a two-step technique, which is frequently used in the literature: Identification and localization of the baby head within the ultrasound image are the main objectives of this first step. This step is essential because it establishes the region of interest (ROI) for the measurements that will follow. The fetal head in the image is precisely located using a variety of methods, such as image processing and machine learning. The next stage is to carefully segment or outline the fetal head's outlines after the fetal head has been localized. This segmentation process aids in precisely defining the head's boundaries. To improve the initial localization and obtain a more accurate head circumference measurement, various segmentation algorithms, such as region-growing, active contours, or deep learning-based methods, may be utilized.

Researchers hope to increase the precision and dependability of head circumference measurements in ultrasound images, which are crucial for keeping track of baby growth and development throughout pregnancy. These techniques have important ramifications for determining the health and wellbeing of fetuses. In Ref [8], The initial phase entails using machine learning to find the fetal head, with Haar-like characteristics being used to train a random forest classifier; the second step entails measuring the HC using elliptical fitting and the Hough transform. The same approach is applied in Ref [2]. Alternative approaches like prediction and ellipse fitting, additionally rely on deep segmented frameworks Ref [16]. The same concept is expanded upon by the authors, who combine image division with ellipse tuning in a multi-task network Ref [4]. In Ref [18], conventional segmentation technique manually assigned labels are used to train U-Net, and segmentation outcomes are fitted to ellipses. When compared to modern methods reported in the literature, the suggested U-Net design showed better or equal segmentation performance.

3. METHODOLOGY

It originated from the traditional neural network with convolution and was utilized in 2015 for processing images used in biomedicine. A standard convolutional neural network is concerned with classifying images, with an input of an image and an outcome of a single label. However, in biomedical applications, it is necessary to identify both the presence of a disease and the location of the abnormality. UNet (Figure 3) is devoted to finding a solution to this problem. This technique effectively localizes and distinguishes borders by classifying each pixel, resulting in output and input images with identical size.

Figure 3: U-Net Architecture

It looks like a "U" at first. The method structure is symmetrical and is separated into two key sections: the procedure that contracts on the left, that is composed of the convolutional analysis in its entirety and the wide path on the right, and this is composed of transposed 2D convolutional layers. (for the time being, think of it as an up-sampling technique). Figure 4 describes the Flowchart for measuring head circumference (HC) and effectively segmenting prenatal ultrasound images by U-Net method.
4. **MATERIALS AND PROCEDURES**

A 12-month prospective cohort observational research was carried out. Fetal HC was assessed postpartum and fetal biometry, including HC, was conducted using sonography. Accuracy metrics and various influences on accuracy are examined. Figure 5 explains graph for learning set then the test set's foresee count.

**4.1 Augmentation**

Image augmentation is a method frequently used in visual analysis and deep studying expand the variety and volume of training data. The original images are exposed to a number of transformations and adjustments, producing fresh enriched copies of the data. The following are some frequently used images enhancement methods: Flipping the image horizontally or vertically allows you to imitate various orientations. Rotation: To introduce variations in object orientations, rotate the image by a predetermined angle, such as 90 degrees or random angles. Scaling and Resizing: You can resize an image to a particular dimension or adjust the image's size by scaling it up or down.

**4.2 Training Set**

A number of important hyperparameter settings were previously discussed in this study. The network parameters were then set, and the model underwent 25 epochs of training in each training session. The input size for model training was set to 768 512 pixels. It was shown that training on this scale increased the results' precision.

**4.3 Testing Set**

The fetal head circumference (HC) measurements were the main focus of the testing set of images utilized to assess the effectiveness of the U-Net model. The precision and efficiency of the U-Net model for calculating HC on unidentified information were evaluated using these images, which were only used for testing factors. The testing set was essential in determining how well the model could be generalized and the extent to which well it could estimate HC in different contexts.

**4.4 Ellipse fitting**

A common technique for determining the embryonic head circumference determined by ultrasonography images ellipse fitting. A general description of the elliptical fitting procedure for fetal HC measurement is given below: To improve the fetal head region's quality and conciseness, the ultrasound image is preprocessed. Edge detection, contrast enhancement, and denoising are merely a few possible processes that are utilized. Selection of a Region of Interest (ROI): In the preprocessed image, an area of interest enclosing the fetal head is found. This phase aids in concentrating the elliptical fitting procedure on the important region. Using methods like edge detection algorithms, the edges region within the fetal head are identified. The edges within the identified fetal head can be fitted with an ellipse using a variety of techniques. The least squares fitting method is a popular technique, where the ellipse's center coordinates, major and minor axes lengths, and rotation angle are optimized to reduce the distance between the ellipse and the edge points that are being identified. Calculating the Head Circumference: After obtaining the ellipse parameters, the head circumference may be determined using the formula for an ellipse's circumference.

\[ 2\pi\sqrt{(c^2 + d^2) / 2} \]

Where c and d are the major and minor axis' respective lengths.

For the purpose of fitting the ellipse, lone the pixels identified through the dynamic programming methodology that fell beyond the probability map's top fifth percentile generated by the pixel classifier was utilized into account. The strong possibility that these pixels came from the fetal skull served as the basis for this selection criterion. The elliptical fitting procedure concentrated on this subset of pixels with the goal of correctly capturing the shape and features of the infant skull.
Table 1: Together the learning set then the test set's foresee count.

<table>
<thead>
<tr>
<th>Trimester</th>
<th>Set for training</th>
<th>Set for evaluating</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>165</td>
<td>55</td>
</tr>
<tr>
<td>Second</td>
<td>693</td>
<td>233</td>
</tr>
<tr>
<td>Third</td>
<td>141</td>
<td>47</td>
</tr>
<tr>
<td>Total</td>
<td>999</td>
<td>335</td>
</tr>
</tbody>
</table>

5. EXPERIMENTAL RESULTS

5.1 The HC18 data resource

The HC18 set for training data resource (van den Heuvel et al., 2018b), which consists of 999 ultrasound pictures taken by various points throughout the pregnancy collectively with the corresponding head circumference. The average HC value is 1263.3 ± 264.4 pixels (174.465.2 mm), with values ranging the size of the region of interest (ROI) ranged from 439.1 pixels (equivalent to 44.3 mm) to 1786.5 pixels (equivalent to 346.4 mm).

5.2 Performance Metrics

5.2.1 Dice similarity coefficient (DSC)

The ground truth region, which is often manually annotated or expertly labeled, and the segmented region acquired from the algorithm or model are represented as binary masks. These masks encode the area of interest as white (foreground) pixels and the backdrop as black pixels. The intersection and union of the segmental and ground truth masks are used to calculate the DSC. The percentage of white pixels in both the ground truth mask and categorized mask is the intersection. Union is the overall sum of white pixels in the categorized mask as well as the ground truth mask. By dividing twice the intersection the DSC is determined as the sum of the pixels in the categorized mask with the ground truth mask.

\[
\text{DSC} = \frac{(\text{Segmented pixels} + \text{Ground truth pixels})}{(2 \times \text{Intersection})}
\]

The region that was segregated and the actual data region must perfectly match for the DSC to be at 1, which varies from 0 to 1. An increased DSC value suggests better segmentation accuracy since it shows a greater degree of overlap and similarity between the segmented and ground truth regions. As it provides a quantitative assessment of the extent to which the segmentation algorithm or model works in capturing the actual segmentation boundaries, the DSC is frequently employed as an evaluation metric for image segmentation tasks.

The Dice coefficient is equal to two times the connection of E and F separated by the total of the areas of E and F. It is also referred to as the Dice coefficient and the F1 score.

\[
\text{DSC} = \frac{2 \times |E \cap F|}{|E| + |F|} = \frac{2 \times TP}{2 \times TP + FP + FN}
\]

(TP=True Positives, FP=False Positives, FN=False Negatives)

5.2.2 Hausdorff distance (HD)

The Hausdorff distance (HD) is a metric frequently used in image analysis and segmentation activities to evaluate the similarity or disagreement between two collections of points or regions. The Hausdorff distance is prone to quantify the difference between the segmented region and the ground truth region when analyzing picture segmentation results. The ground truth region (usually manually annotated or expert-labeled) and the segmented region (obtained through the algorithm or model) are represented as sets of points or coordinates. The maximum distance between any argument in one set and its adjoining neighbor in the supplementary set is used to calculate the Hausdorff distance. The segmented region to the ground truth region and vice versa are both involved in this procedure. Forward Hausdorff distance (FHD) is the determined distance amongst any argument in the segmented region and its adjoining neighbor in the ground truth region. The greatest separation amongst each point in the segmented region and its closest neighbor in the ground truth region is called as the reverse Hausdorff distance (RHD).

The biggest difference between the onward and inverse Hausdorff distances is referred to as the Hausdorff distance.

\[
\text{HD} = \text{maximum}(\text{FHD}, \text{RHD})
\]

A measure of dissimilarity, the Hausdorff distance shows the greatest separation between the two sets of points. A decreasing Hausdorff distance means that the divided region and the region of the ground truth are more similar or correspond. A shorter Hausdorff distance indicates that the divided region accurately captures the ground truth region's boundaries and shape in picture segmentation evaluations. Therefore, the Hausdorff distance can be used as a useful statistic to measure how different or off the segmentation results are from the actual data in terms of spatial location and extent.
5.2.3 DF or HC difference
The head circumference difference is referred to as DF, also known as HC difference. The difference between the estimated Head Circumference (HC) value and the actual HC value is measured using this performance index. The absolute variance amongst the estimated HC and the actual HC is used to determine the DF or HC difference:

\[ DF = |\text{Estimated HC} - \text{Ground Truth HC}| \]

A smaller HC differential reflects greater HC estimation accuracy since the estimated and ground truth HC values are more closely aligned. By quantifying the absolute deviation between the estimated and actual HC values, the HC difference measure aids in assessing the accuracy and dependability of HC estimation methods or models.

5.2.4 Mean Absolute Error (MAE)
A different approach frequently utilized in loss function and assessment measure in regression projects is a mean absolute error (MAE). MAE is used to quantify the size of the forecasts' errors and offers a reliable indicator of the model's effectiveness. The following equation can be used to determine MAE:

\[ \text{MAE} = \frac{1}{n} \sum |y\_pred - y\_true| \]

where:
y\_pred signifies the predicted values.
y\_true signifies the ground truth (actual) values.
n is the sum of data points in the dataset.

MAE takes into account the absolute differences as opposed to DF, which squares the discrepancies between anticipated and actual values. Because of this, MAE offers a simpler interpretation of the average error and is less sensitive to outliers. By quantifying the absolute difference between the estimated and actual HC values, it is possible to evaluate the accuracy and dependability of HC estimation methods or models.

Table 2: The concert of various deep learning models in relations of segmentation and biometry was evaluated. HC Mean Absolute Error (MAE), Hausdorff distance (HD), Dice similarity coefficient (DSC), and HC difference (DF), and DLA deep learning architecture.

<table>
<thead>
<tr>
<th>Techniques</th>
<th>DSC (%)</th>
<th>HD (mm)</th>
<th>MAE (mm)</th>
<th>DF (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-Net</td>
<td>97.84 ± 1.33</td>
<td>1.20 ± 0.75</td>
<td>2.00 ± 1.85</td>
<td>0.93 ± 2.57</td>
</tr>
<tr>
<td>U-Net</td>
<td>97.70 ± 1.39</td>
<td>1.33 ± 0.84</td>
<td>2.34 ± 2.10</td>
<td>1.80 ± 2.55</td>
</tr>
<tr>
<td>DLA V-Net</td>
<td>97.85 ± 1.25</td>
<td>1.29 ± 0.74</td>
<td>1.89 ± 1.73</td>
<td>−0.38 ± 2.53</td>
</tr>
<tr>
<td>DLA U-Net</td>
<td>97.87 ± 1.27</td>
<td>1.30 ± 0.77</td>
<td>2.01 ± 1.81</td>
<td>−1.01 ± 2.55</td>
</tr>
</tbody>
</table>

The V-Net achieved an average DSC of 97.84%, with a standard deviation of 1.33%. The DSC procedures the overlay amongst the foreseen separation and the ground truth, with higher values demonstrating improved segmentation accuracy. The V-Net achieved an average HD of 1.20 mm, through a standard deviation of 0.75 mm. The HD quantifies the maximum distance comparison of the anticipated segmentation and the actual segmentation, with smaller values denoting better precision in segmentation and position. The V-Net achieved an average MAE of 2.00 mm, with a standard deviation of 1.85 mm. The MAE measures the absolute variance amongst the predictable and ground truth measurements, with smaller values indicating higher accuracy in biometric measurement estimation. The V-Net achieved an average DF of 0.93 mm, with a standard deviation of 2.57 mm. The DF measures the variance amongst the predictable and ground truth measurements, with smaller values indicating better accuracy in biometric measurement estimation. The U-Net model achieved impressive results in terms of segmentation performance, as indicated by the following evaluation metrics an average DSC of 97.70%, with a standard deviation of 1.39%; HD of 1.33 mm, with a standard deviation of 0.84 mm;
MAE of 2.34 mm, with a standard deviation of 2.10 mm and DF of 1.80 mm, with a standard deviation of 2.55 mm. The DLA V-Net model demonstrated remarkable segmentation performance DSC of 97.85 ± 1.25, HD of 1.29 ± 0.74, MAE of 1.89 ± 1.73 and DF of −0.38 ± 2.53. The DLA U-Net model achieved promising results in terms of segmentation performance, as indicated by the evaluation metrics. In terms of fetal head segmentation performance, the DLA U-Net model produced good results compared to the other techniques. The DLA V-Net model confirmed remarkable segmentation concert DSC of 97.87 ± 1.27, HD of 1.30 ± 0.77, MAE of 2.01 ± 1.81 and DF of −1.01 ± 2.55. These results highlight the effectiveness of the DLA U-Net model in accurately segmenting and estimating biometric measurements, such as head circumference. The high DSC, low HD, and small values for DF and MAE indicate the model's ability to capture precise and accurate delineation of regions of interest, leading to reliable biometric measurements.

<table>
<thead>
<tr>
<th>Performance metrics</th>
<th>Initial trimester</th>
<th>Second trimester</th>
<th>Final trimester</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF (mm)</td>
<td>0.30</td>
<td>−0.011</td>
<td>−0.30</td>
</tr>
<tr>
<td>MAE (mm)</td>
<td>1.42</td>
<td>1.718</td>
<td>2.50</td>
</tr>
<tr>
<td>HD (mm)</td>
<td>0.85</td>
<td>1.20</td>
<td>2.16</td>
</tr>
<tr>
<td>DSC (%)</td>
<td>96.84</td>
<td>98.15</td>
<td>98.11</td>
</tr>
</tbody>
</table>

Figure 8: Graph for Predictable technique was assessed in terms of segmentation and biometry.

In the first trimester, the performance metrics for the model were as follows are HC difference was measured to be 0.30 mm, indicating a small discrepancy amongst the predictable and ground truth HC values in the first trimester. The HC MAE was found to be 1.42 mm, which represents the average absolute variance amongst the predictable and ground truth HC values in the first trimester. The Hausdorff distance, a measure of dissimilarity, was calculated to be 0.85 mm, indicating a relatively close alignment between the segmented region and the ground truth in the first trimester. The DSC was determined to be 96.84%, signifying a high degree of overlap and similarity amongst the separated region and the ground truth in the initial trimester. A second trimester is when, the performance metrics for the model were DF (mm) −0.011, MAE (mm) 1.718, HD (mm) 1.2 and DSC (%) 98.15. The result of the third trimester is as follow as DF (mm) −0.30, MAE (mm) 2.50, HD (mm) 2.16 and DSC (%) 98.11.

6. FUTURE WORK

In the future, we want to fix problems with the model, like overfitting and narrow generalization. Reducing the number of layers in the computational framework is how we want to accomplish this, with the goal of improving its performance with the existing requirement for manual dataset extension. With this method, the model's segmentation skills will be improved, increasing its efficacy and efficiency for practical uses.

7. CONCLUSION

Despite this study, we created a deep learning technique called DLA U-Net for measuring head circumference (HC) and segmenting fetal heads from two-dimensional ultrasound images. In comparison to prior investigations, the empirical findings demonstrated that the U-Net deep learning architecture achieved better outcomes with a DSC of 97.94%, DF of 0.09 2.49 mm, MAE of 1.78 1.69 mm, and HD of 1.30 0.79 mm. This study acknowledges certain limitations need to be considered into account. Firstly, the numeral of cases involved in the learning is inadequate, which might impact the generalizability of the results. Next, the separation precision for the initial and second trimester of pregnancy is relatively inferior, indicating a potential area for improvement. Additionally, the final trimester of pregnancy shows larger discrepancies in terms of the absolute difference (AD). These limitations provide opportunities for future research to address these issues and enhance the methodology.

REFERENCES:


