Automated measure of fetal head circumference by means of Distance Field based Mask R-CNN

Daniele Delli Rocili
Student of computer engineering and automation master's degree at Università Politecnica delle Marche

Alberto Pierini
Student of computer engineering and automation master's degree at Università Politecnica delle Marche

Emanuele Frontoni
Computer Vision teacher at Università Politecnica delle Marche

Sara Moccia
PhD at Università Politecnica delle Marche

Maria Chiara Fiorentino
M.Eng. at Università Politecnica delle Marche

***Abstract*—This article presents a method for the automatic detection and measurement of fetal head circumference (HC) through 2D ultrasound images. This measure can be used to estimate the gestational age of the fetus and to identify growth problems. Our approach propose the use of a Deep Convolutional Neural Network for the detection and regression of the HC. Being ours an edge detection problem, we used a distance field approach to identify the contours of the head. Our system was trained on 999 ultrasound images and validated on independent test set of 335 images from all trimesters. The method used in this approach involve fine tuning of a modified version of the Mask RCNN that we called Distance Field Mask RCNN (DF Mask RCNN), where the Mask Branch has been replaced with a Regression Branch. A threshold is then applied on each Distance Field in order to extract the Binary Mask. We then used an Ellipse Fitting method on the Binary Mask to estimate the ellipse that approximates the head circumference. To train the Distance Field Mask RCNN we used regression masks generated from the skeletonized masks with a Gaussian distribution mounted above. Our approach has achieved a Dice of 96.83±1.72%, a Difference of 0.22±2.96 mm, an Absolute Difference of 2.10±2.10 mm and an Hausdorff Distance of 1.80±0.96 mm. The results show good performance and are comparable to the state of the art, a sign of the fact that our approach can be pursued for identifying the contours of the head of a fetus.**

I. Introduction

Ultrasound imaging is widely used in prenatal screening in most of the world due to it's relative safety, low cost, noninvasive nature, real-time display, operator comfort and operator experience. Trough the ultrasounds a series of parameters are estimated to identify the state of health and the gestational age of the fetus. One of them is the Head Circumference [1]. The manual procedure for estimating the circumference involves an expert who manually annotates the perimeter of the fetal head by drawing the ellipse that best suits the circumference of the head. This operation can lead to problems like inter and intra-observer variability due to subjective experience and can be very time consuming. The aim our work is to propose an automatic method to estimate the fetal head circumference.This will allow to reduce the errors introduced by the specialist during the measurement and to reduce the execution time of the operation. This technology can be used where experts are lacking, especially in developing countries.

This article is structured as follows: in Section II there is a description of the techniques developed in the state of the art. In Section III material and methods used in our approach are presented, with focus on the Distance Field Mask RCNN. Section IV presents the experimental protocol. Section V presents in detail the results obtained in the test phase. In Section VI there are discussions about test results. The conclusions are presented in Section VII and the biographical references are reported in Section VIII.

II. Related Work

The methods present in literature can be divided into three categories: Model based, Machine Learning and Deep Learning. For Model based methods Perez-Gonzalez et al. [2] propose a method that incorporate texture map, morphological operations and active contours together with optimal ellipse detection. Rajinikanth et al. [3] use a pre-processing with Jaya algorithm,Otsu thresholding and Chan-Vase with Level-set for segmentation. These model based methods use thresholds and therefore cannot be generalized.

For Machine Learning methods Li et al. [4] propose a random forest classifier for ROI Detection, head circumference generation using phase symmetry feature map and ElliFit. Van den Hevel et al. [5] use a computer aided detection of the fetal head circumference consisting in two steps: first, Haar-like features were computed from ultrasound images to train a random forest classifier to locate the fetal skull, then, HC was extracted using Hough transform, dynamic programming and an ellipse fit. One of the limitations of these methods is the use of handcrafted features that are not problem specific.

For Deep Learning methods Sobhaninia et al. [6] [7] propose two methods based on LinkNet, the first uses a unique network for segmentation and regression of ellipse parameters, but no comparison between ellipse regression and an ellipse fitting algorithm is presented, the second uses a Mini-LinkNet that does not extract deep feature, both of these methods have a multi scale input. Irene et al. [8] propose a method based on YOLO for ROI detection and use DoGell for ellipse detection, in this method the CNN is only used to locate the ROI. Zhang et al. [9] propose regression

CNNs for direct HC measure, but the error obtained with this system is doubled if compared to segmentation based approaches. Desai et al. [10], Kim et al. [11], Aji et al. [12] propose different segmentation methods based on the UNet. Segmentation-based approaches may have difficulty segmenting objects with great intensity variation such as ultrasound.

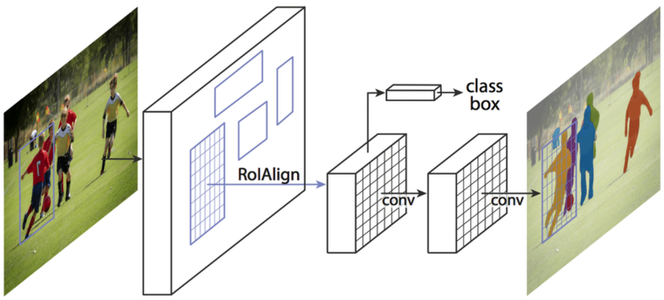

Fig. 1. Mask RCNN Architecture

III. Material and Methods

Section III.A describe the Mask RCNN [13], the instance segmentation network that we later modified according to our needs. Section III.B presents the change we made to the mask branch which became a regression branch. Section III.C describe the used Ellipse Fitting algorithm [14].

A. Mask-RCNN

Mask RCNN [13] extends the Faster RCNN [15] by adding a branch for predicting a segmentation mask on each Region of Interest (RoI), in parallel with the existing branch for classification and bounding box regression. The mask branch is a small Fully Convolutional Network (FCN) applied to each RoI, predicting a segmentation mask in a pixel-to-pixel manner. The general architecture is shown in Fig. 1.
in detail, a pre-processing is applied to each image by subtracting the mean, performing a rescale and padding. The backbone used is a ResNet101 with Feature Pyramid Network (ResNet101-FPN), this network has 4 layers and for each of these a feature map is extracted. At every layer the feature maps size is reduced by half and number of feature maps are doubled. To generate final feature maps, we use an approach called top-bottom pathway. We start from the smallest feature map and work our way down to bigger ones, by upscale operations. For each feature map a set of Anchor Box is generated, those and the feature maps of the four levels are given as input to the Region Proposal Network (RPN). The Loss function used in RPN is the sum of Binary Cross Entropy (classification loss) and the sum of all smooth-L1 losses (bounding box regression loss) for all foreground anchors. Each of the feature maps generated is passed through a 3 x 3 convolution layer. The resulting output is passed to two branches, one pertaining to foreground/background scores and other to bounding box regressors. Then the top K (default K for train=12000, test=6000) anchor are selected based on their foreground score and the bounding boxes are modified according to the regression coefficients obtained from RPN. The bounding boxes in which either of coordinate lies outside the image are removed and is applied a pruning based on Non Maximum Suppression with a default threshold of 0.7. All the anchor boxes are now grouped together and are selected the top N (default N=2000) proposals based on their corresponding foreground score. An FPN-ROI mapping is now applied, it associates a specific feature map of the FPN with an ROI based on its size. Via the ROI Align operator is generated a uniform P x P matrix for all the ROIs, quantization is not used and the ROI feature maps are obtained through bilinear interpolation. The output for each ROI is reshaped to be passed through two fully connected layers, one for classification and one for bounding box regression coefficients. The Loss function used in classification branch is Cross Entropy and in Bounding Box Regression branch is smooth L1. The ROI Align output is passed trought the Mask Branch, a small FCN that predict a binary mask for each class independently, without competition among classes, and rely on the network's ROI classification to predict the category. The Loss function used in Mask Branch is Binary Cross Entropy.

B. Regression branch

In the method we propose, the mask branch has been changed to a regression branch. The Loss function on the mask branch, that was Binary Cross Entropy (BCE), was replaced with Mean Squared Error (MSE).
For sample i given prediction p_i and target t_i we define:

$$BCE_i = \sum_j t_{i,j} \log p_{i,j}$$

$$MSE = \frac{1}{N} \sum_{i=1}^{n} (t_i - p_i)^2$$

The inputs of the new network are the ultrasound images and the regression masks of the circumference. The output of the new regression branch is a Distance Field (DF), a map where each pixel is labeled with the distance from the nearest boundary pixel. In our case we obtain the HC Distance Field.

C. Ellipse fitting

Once obtained the distance fields we apply a threshold (TH) to obtain the points of the fetal head distance field with a confidence value greater than the threshold value. From the obtained binary mask we extract the contours and apply the ElliFit [14] algorithm to estimate the ellipse parameters $\Theta = (x_c, y_c, a, b, \theta_c)$, where x_c and y_c are the center coordinates, a and b are the axes and θ_c is the rotation angle. This method models the non-linear problem of ellipse fitting as a combination of two operators, with one being linear, numerically stable, and easily invertible, while the other being non-linear but unique and easily invertible operator.
The steps of the method are illustrated in Fig. 2.

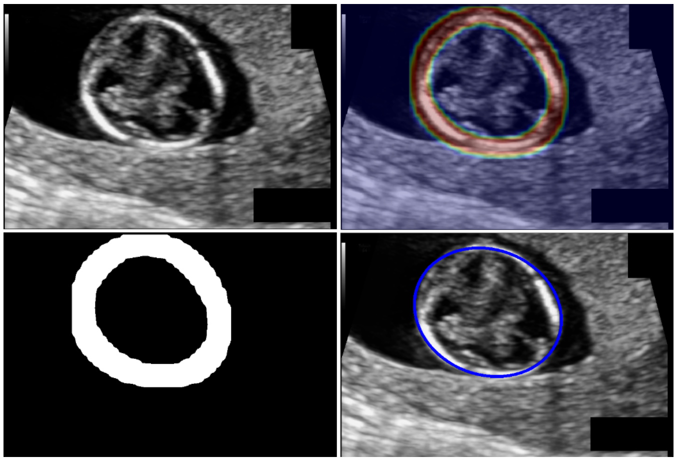

Fig. 2. In the upper left there is the US image, in the upper right there is the Distance Field, at the bottom left there is the Binary Mask and at the bottom right there is the estimated ellipse

IV. Experimental Protocol

In this section, dataset, training settings and performance metrics are presented

A. Dataset

To train our network we used the HC18 Grand Challenge dataset [5][16], which contains 999 ultrasound images and their annotations masks for training a 335 ultrasound images for testing, both dataset are related to first, second and third pregnancy trimesters. The ultrasound images were acquired from 551 pregnant women who received a routine ultrasound screening exam between May 2014 and May 2015. Only fetuses that did not exhibit any growth abnormalities were included in this study [5]. The size of each 2D ultrasound image and annotation was 800 by 540 pixels. Ultrasound images datasets present several challenges like head size variability, artifacts, attenuation, shadows, speckle noise, missing boundaries and low signal-to-noise ratio.

A stratified split on the training set was performed to divide the dataset in two parts. An 80% for training and the remaining 20% for validation.

The original annotation masks have been one-pixel skeletonized. For each point of the ellipse perimeter was calculated the normal line and on this was drawn a 100 pixel length segment with the same angular coefficient and centered on the point of the ellipse. Above these segments, was mounted a Gaussian distribution (Fig. 3). We used these regression mask to train our Distance Field Mask RCNN.

Data augmentation on the fly was performed on the training set. Ultrasound images and regression masks were scaled, translated, rotated and sheared.

B. Training setting

For the training of the Distance Field Mask RCNN we used a ResNet101-FPN backbone and as optimizer a Mini Batch Gradient Descent with momentum set to 0.9. The learning rate has been set to 0.001, the Non Maximum Suppression

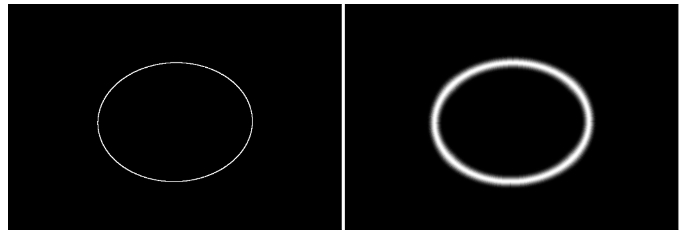

Fig. 3. On the left there is a skeletonized mask, on the right a regression mask

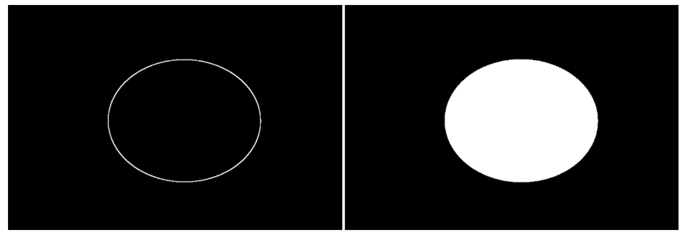

Fig. 4. On the left there is a skeletonized mask, on the right there is a full mask

threshold to 0.7, the batch size to 4. A resize to 512x512 was performed and starting from the COCO [17] weights for the Mask RCNN we trained for 10 epochs the heads layers of the network and for other 12 epochs all the layers.

C. Performance Metrics

To evaluate the performance of the Distance Field Mask-RCNN we used the HC18 Grand Challenge metrics:

- Dice Coefficient (DSC):

$$DSC = \frac{2 * |Area_p \cap Area_{gt}|}{|Area_p| + |Area_{gt}|}$$

  where the $Area_p$ is the area of the predicted ellipse and the $Area_{gt}$ is the ground truth area.

- Difference (DF):

$$DF = HC_p - HC_{gt}$$

  where HC_p is the predicted head circumference and HC_{gt} is the groundtruth head circumference.

- Absolute Difference (ADF):

$$ADF = |HC_p - HC_{gt}|$$

- Hausdorff Distance (HD):

$$HD(P, G) = max(h(P, G), h(G, P))$$

  where $P = p_1, ..., p_q$ are the pixels from predicted ellipse and $G = g_1, ..., g_k$ are the pixels from ground truth ellipse, given:

$$h(P, G) = \max_{p \in P} \min_{g \in G} \|p - g\|$$

TH	DC	DF (mm)	ADF (mm)	HD (mm)
0.9	97.30±1.22	-2.88±3.04	3.29±2.59	1.53±0.84
0.8	97.29±1.27	-2.47±3.01	3.00±2.48	1.52±0.82
0.6	97.25±1.39	-1.65±2.94	2.50±2.27	1.54±0.82
0.2	96.83±1.73	0.22±2.96	2.10±2.10	1.80±0.96

TABLE I
COMPARISON OF RESULTS BASED ON THRESHOLD APPLIED TO DISTANCE FIELD

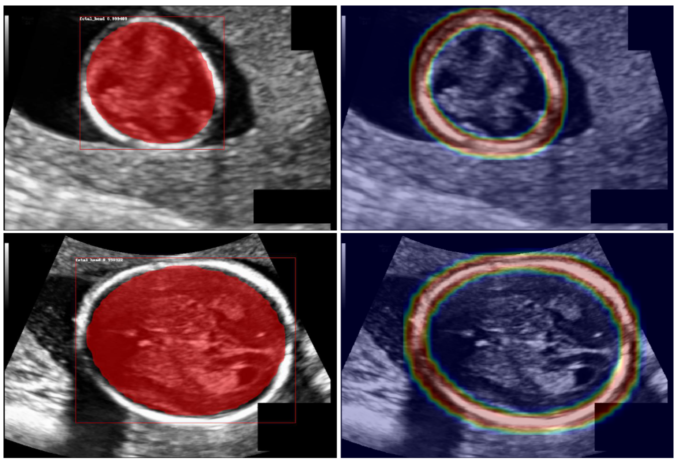

Fig. 5. On the left column the Binary Masks obtained with the original Mask RCNN, on the right column the Distance Field obtained with the DF Mask RCNN on the same US

V. RESULTS

We compared the results obtained with the Distance Field Mask RCNN with those of the original Mask RCNN, which was trained with full binary masks (Fig. 4) of the fetal heads. Some results are showed in Fig. 5.
The best results for the original Mask RCNN were obtained with a ReNet101-FPN backbone, with the learning rate set to 0.001 and the batch size to 4. We trained starting with coco weights for 95 epoch, 20 for the heads layer and 75 for all layers. The best results for the Distance Field Mask RCNN were obtained with TH=0.2 on the Distance Field. The results obtained with different threshold are shown in Table 1.

As we can see in Table 2 the error obtained with binary masks and the original Mask RCNN is greater than the one obtained with the Distance Field Mask RCNN.

Considering the good results obtained with the Distance Field Mask RCNN we have compared ours with those present in the state of the art. As shown in Table 2 our results are comparable with those obtained with other methods.
Table 3 shows the results divided by trimester of gestation.

Method	DC	DF (mm)	ADF (mm)	HD (mm)
DF Mask RCNN (ours)	96.83±1.73	0.22±2.97	2.10±2.10	1.80±0.97
Mask RCNN	89.25±1.59	-18.11±7.56	18.13±7.52	3.85±1.58
Rajinikanth et al. 2019	94.27±−	-	-	-
Van den Hevel et al. 2018	97.0±2.8	0.6±4.3	2.8±3.3	2.0±1.6
Desai et al., 2020	97.33±1.41	-	-	1.58±0.97
Sobhaninia et al., 2019	96.84±2.89	1.13±2.69	2.12±1.87	1.72±1.39
Sobhaninia et al., 2020	92.46±−	1.19±−	2.22±−	3.40±−
Zhang et al., 2020	-	-	4.52±4.27	-
Aji et al., 2019	-	41.75±−	-	-

TABLE II
COMPARISON OF OUR DISTANCE FIELD MASK RCNN AGAINST ORIGINAL MASK RCNN AND STATE OF THE ART

Trimester	DC	DF (mm)	ADF (mm)	HD (mm)
1	95.17±2.97	1.39±2.72	1.96±2.33	1.35±0.91
2	97.16±1.15	0.13±2.56	1.83±1.79	1.69±0.83
3	97.16±0.85	-0.68±4.44	3.61±2.62	2.88±0.90

TABLE III
COMPARISON OF ALL TRIMESTER USING DISTANCE FIELD MASK RCNN

VI. DISCUSSION

Having to face an edge detection task, we have chosen to use a distance field approach as we think it is a good method for identifying the contours. In fact, using the original Mask RCNN we observe an underestimation of the segmentation of the head, a sign that the network is not able to correctly recognize the contours of the fetal head. This is due to the large variation in intensity between the inside and the contour of the head. Instead, a distance field approach allows us to exploit this difference in intensity to directly regress the contour of the head.
The Distance Field approach has brought good results with a mean absolute difference of 2.10±2.10 mm. We observe that the Dice is the lower in the first trimester, we assume that it depends on the fact that the fetal skull is not clearly visible in ultrasounds related to this one. Regarding the mean absolute difference and the Hausdorff distance of the third trimester, we observe major errors due to the spread of the error caused by the grater size of the head.
Future improvements of the proposed methodology may include a larger dataset for training and adding convolutional layers to directly regress the ellipse parameters from the network, eliminating the errors caused by threshold and ElliFit algorithm and transform the estimate of the ellipse into a learning process.

VII. CONCLUSIONS

A two-step framework was exposed in this paper, focused on extracting the perimeter of the head circumference of a fetus from an ultrasound image.
The first step includes a Mask RCNN with a Regression Branch used to obtain the Distance Field of the fetal head to which a threshold is applied to obtain a binary mask.

From this we extract the contours to be passed to the ellipse fitting algorithm in order to obtain the ellipse parameters.
This method was evaluated on a large independent test set of 335 ultrasound images that included data of all trimesters and shows similar or superior results to the other system published in literature.

REFERENCES

[1] R. UM, F. RA, and C. JA, "Prenatal imaging: ultrasonography and magnetic resonance imaging." *Obstet Gynecol 2008;112(1):145–57.*

[2] J. Perez-Gonzalez, J. B. Muñoz, M. R. Porras, F. Arámbula-Cosío, and V. Medina-Bañuelos1, "Automatic fetal head measurements from ultrasound images using optimal ellipse detection and texture maps," *Latin American Congress on Biomedical Engineering*, 2014.

[3] V. Rajinikanth, N. Deyb, R. Kumarc, J. Panneerselvamd, V. Rajinikantha, N. D. R. Kumar, J. Panneerselvam, and N. S. M. Raja, "Fetal head periphery extraction from ultrasound image using jaya fetal head periphery extraction from ultrasound image using jaya algorithm and chan-vese segmentation," *Procedia Computer Science 152 66-73*, 2019.

[4] J. Li, Y. Wang, B. Lei, J.-Z. Cheng, J. Qin, T. Wang, S. Li, , and D. Ni, "Automatic fetal head circumference measurement in ultrasound using random forest and fast ellipse fitting," *IEEE Journal of Biomedical and Health Informatics*, 2016.

[5] van den Heuvel TLA, de Bruijn D., de Korte C.L., and G. B., "Automated measurement of fetal head circumference using 2d ultrasound images," *PLoS ONE 13(8): e0200412*, 2018.

[6] Z. Sobhaninia, S. Rafiei, A. Emami, N. Karimi, K. Najarian, S. Samavi, and S. Soroushmehr, "Fetal ultrasound image segmentation for measuring biometric parameters using multi-task deep learning," *2019 IEEE 6545-6548*.

[7] Z. Sobhaninia, A. Emami, N. Karimi, and S. Samavi, "Localization of fetal head in ultrasound images by multiscale view and deep neural networks," *2020 IEEE*.

[8] K. Irene, A. Y. P., H. Haidi, N. Faza, and W. Chandra, "Fetal head and abdomen measurement using convolutional neural network, hough transform, and difference of gaussian revolved along elliptical path (dogell) algorithm," 2019.

[9] J. Zhang, C. Petitjean, and P. L. S. Ainouz, "Direct estimation of fetal head circumference from ultrasound images based on regression cnn," 2020.

[10] A. Desai, R. Chauhan, and J. Sivaswamy, "Image segmentation using hybrid representations," *arXiv:2004.07071v1*, 2020.

[11] H. P. K. et al., "Automatic evaluation of fetal head biometry from ultrasound images using machine learning," *Physiol. Meas. 40 065009*, 2019.

[12] A. C. P., F. M. H., Sardjono, and T. A., "Automatic measurement of fetal head circumference from 2-dimensional ultrasound," *2019 IEEE*.

[13] K. He, G. G. Piotr, Dollàr, and R. Girshick, "Mask r-cnn," *arXiv:1703.06870*, 2018.

[14] D. K. Prasad, M. K. Leung, and C. Quek, "Ellifit: An unconstrained, non-iterative, least squares based geometric ellipse fitting method," *2012 Elsevier Ltd.*

[15] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *arXiv:1506.01497*, 2015.

[16] T. L. A. van den Heuvel, D. de Bruijn, C. L. de Korte, and B. van Ginneken, "Automated measurement of fetal head circumference using 2d ultrasound images [data set]," *Zenono*, 2018.

[17] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollar, "Microsoft coco: Common objects in context," *arXiv:1405.0312 (cs)*, 2014.