Preview poster

Back to overview

Abstract Title:

Deep Learning model for fetal ultrasound abdominal plane quality improvement

Authorblock: S. Slimani[1], S. Hounka[2], A. Tlemçani[3], M. Akiki[3], D. LAOUDIYI[1], A. Lamrissi[1], A. Bouzyane[3], S. Hanane[4], S. Bouhya[1]; [1]CASABLANCA/MA, [2]Rabat/MA, [3]Casablanca/MA, [4]Oujda/MA

Authors:
1. Dr. Saad Slimani [1] (presenter)
2. Mr Salaheddine Hounka [2]
3. Dr. Abed Tlemçani [3]
4. Dr. Mustapha Akiki [4]
5. Mrs. DALALE LAOUDIYI [1]
6. Pr. Amine Lamrissi [1]
7. Pr. Amal Bouzyane [3]
8. Pr. Saadi Hanane [5]
9. Pr. Said Bouhya [1]

Disclosures:
Dr. Saad Slimani: Nothing to disclose
Mr Salaheddine Hounka: Nothing to disclose
Dr. Abed Tlemçani: Nothing to disclose
Dr. Mustapha Akiki: Nothing to disclose
Mrs. DALALE LAOUDIYI: Nothing to disclose
Pr. Amine Lamrissi: Nothing to disclose
Pr. Amal Bouzyane: Nothing to disclose
Pr. Saadi Hanane: Nothing to disclose
Pr. Said Bouhya: Nothing to disclose

Keywords: Artificial Intelligence, Obstetrics (Pregnancy / birth / postnatal period), Ultrasound, Computer Applications-Detection, diagnosis, Obstetrics

Purpose or Learning Objective:
To develop a deep learning (DL) model able to classify fetal ultrasound abdominal planes according to quality criteria described by the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG)[1] as a stepping stone to fetal biometry workflow automation.

Methods or Background:
Quality assessment of fetal ultrasound (US) standard biometry planes is crucial in ensuring adequate measurements to estimate fetal growth[2].

Abdominal circumference (AC) is one of the major biometric parameters measured from 14 weeks onwards, that reflects fetal growth and wellbeing along with head circumference (HC), biparietal diameter (BIP), and femur length (FL)[3].

AC measurement is essential because it can by itself be a diagnostic indicator of fetal growth restriction (FGR), the world's leading cause of perinatal mortality and morbidity[4].

Indeed, to come up with a definition of early and late FGR, a recent Delphi survey[5] acknowledged AC < 3rd centile as a solitary diagnostic parameter along with Estimated fetal weight (EFW) < 3rd centile.

Correct biometric parameters measurement is paramount to ensure early diagnosis of FGR. However, studies have shown inter-observers differences in AC measurement up to 8.8%[6]. Adding to this variability in caliper placement is the variability in terms of quality assessment and choice of the right abdominal plane leading to costly quality audits[7].

With the advent of performant DL models applied to medical imaging, attempts have been made to automate the two tasks that compose biometric measurement workflow: plane detection (image classification) and caliper placement (fetal structure's segmentation)[8].

Human-level performances were achieved in fetal US plane classification, with models able to distinguish between standard US planes that showcase very different anatomical structures: fetal brain, abdomen, and femur, for example[9]. However, few attempts have been made to classify fetal US images involving subtle differences, also called fine-grained image classification problems.

Researchers have tried to tackle this issue by applying an acceptance check approach to abdominal plane image classification with a few features as input, such as the presence of the umbilical vein or the stomach bubble reaching accuracies ranging from 0.80 to 0.87 [10,11]. However, these studies did not consider the entirety of the abdomen plane quality criteria as described by Salomon et al.[12], which are: symmetrical plane, plane showing stomach bubble, plane showing portal sinus, kidneys not visible.

In this work, we conducted a retrospective study in three fetal ultrasound centers in Morocco. The data were collected from the PACS systems of each facility and merged with the open dataset in Xavier P. Burgos-Artizzu et al. publication[9].

Before handling the data, all images were deidentified using the RSNA Dicom anonymizer tool[13]. Preprocessing steps included converting the images from Dicom to PNG, resizing the images to the 224 x 224 format, and cropping of non-relevant anatomical structures (AF and uterine walls). Furthermore, data augmentation techniques were used, including random flip both vertical and horizontal, random rotation of 10 degrees, and random contrast enhancement.

Fig 2: Abdominal image before cropping.

Preprocessing steps included converting the images from Dicom to PNG, resizing the images to the 224 x 224 format, and cropping of non-relevant anatomical structures (AF and uterine walls).

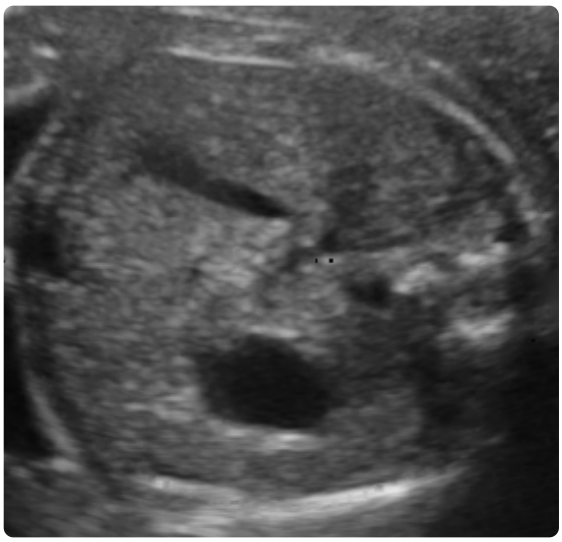

Fig 1: Abdominal image after cropping.

Furthermore, data augmentation techniques were used, including random flip both vertical and horizontal, random rotation of 10 degrees, and random contrast enhancement.

Two radiology residents with fetal ultrasound training annotated each image according to the four ISUOG quality criteria mentioned above using a customized version of the open-access image annotation tool Label Studio. Before annotating the images, the two residents held two training sessions with a fetal US expert (>10 years' experience). The final dataset consisted of 3135 annotated images and was randomly split into training and test sets with a 3/4 ratio.

The problem was treated as a multilabel classification task where the model assigned one or more of four classes to each image. Our best model was a finetuned ResNet152V2 pre-trained architecture[14], where 60% of layers were frozen.

Results or Findings:

The precision of the final model was 0.8 for symmetry assessment, 0.88 for stomach bubble visualization, 0.81 for portal sinus visualization, 0.8 for kidneys visibility, leading to a total accuracy of 0.83.

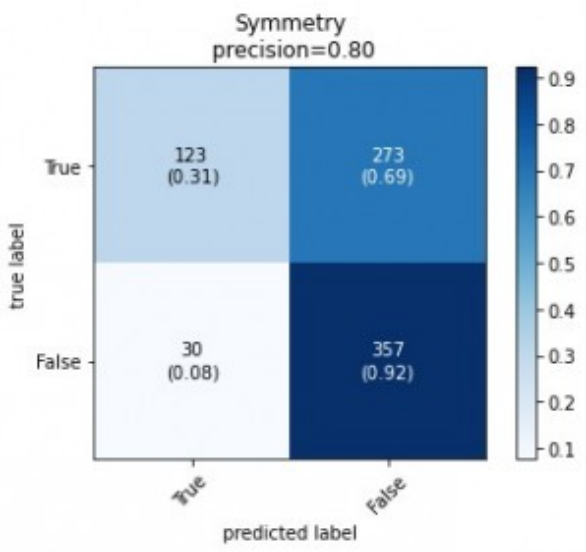

Fig 4: Confusion matrix of the model's performance on symmetry assessment. Matrix rows show the true class, labeled by our radiology residents.

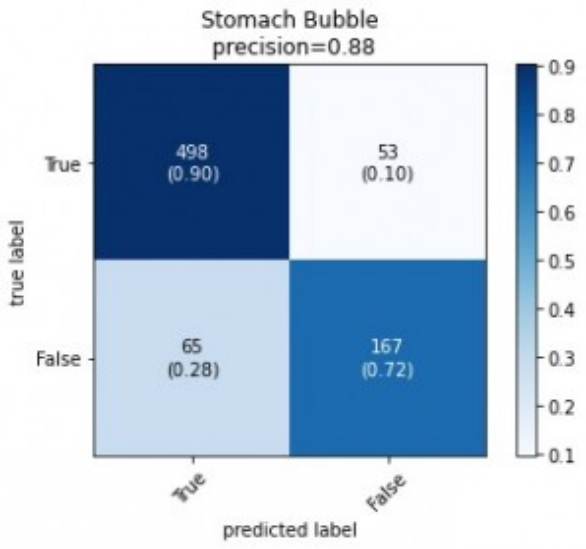

Fig 5: Confusion matrix of the model's performance on stomach bubble visibility assessment. Matrix rows show the true class, labeled by our radiology residents.

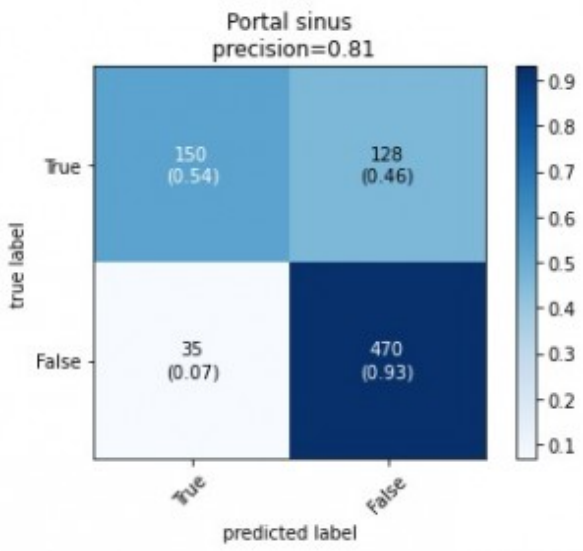

Fig 6: Confusion matrix of the model's performance on portal sinus visibility assessment. Matrix rows show the true class, labeled by our radiology residents.

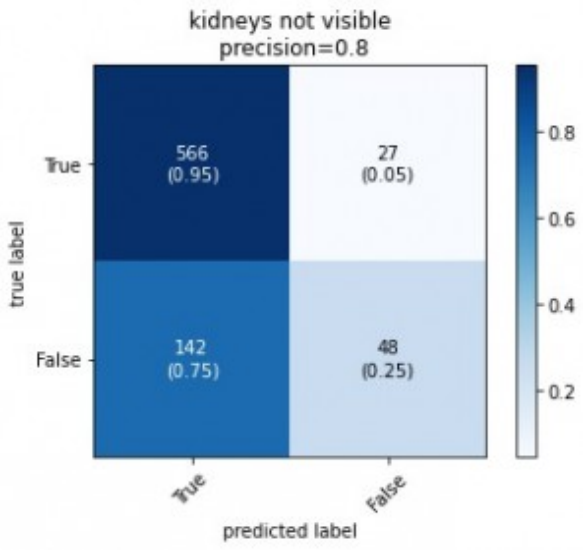

Fig 3: Confusion matrix of the model's performance on kidney visibility assessment. Matrix rows show the true class, labeled by our radiology residents.

Conclusion:

We developed a DL model with encouraging results in the task of abdominal images quality assessment.It performed worst on the symmetry assessment task because the abstract nature of such a feature depends on many factors such as abdomen visibility, the integrity of the contours, and shape. The comparable performance on portal sinus and kidney visualization might be due to the relatively small size of the PS, particularly in poorly zoomed images and the scarcity of examples that contained visible kidneys.Including all the criteria from the ISUOG guidelines is critical to ensure image conformity

and allow future models autonomy in fetal biometry assessment. With improved results, such a tool combined with a segmentation model could automate part of the routine fetal US screening scan and improve FGR diagnosis in stranded resources environments.

References:

1. Salomon, L. J. *et al.* ISUOG Practice Guidelines: ultrasound assessment of fetal biometry and growth. *Ultrasound in Obstetrics & Gynecology* **53**, 715–723 (2019).
2. Salomon, L. J. *et al.* Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound in Obstetrics & Gynecology* **37**, 116–126 (2011).
3. O'Gorman, N. & Salomon, L. J. Fetal biometry to assess the size and growth of the fetus. *Best Practice & Research Clinical Obstetrics & Gynaecology* **49**, 3–15 (2018).
4. Colella, M., Frérot, A., Novais, A. R. B. & Baud, O. Neonatal and Long-Term Consequences of Fetal Growth Restriction. *Current Pediatric Reviews* **14**, 212 (2018).
5. Gordijn, S. J. *et al.* Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound in Obstetrics & Gynecology* **48**, 333–339 (2016).
6. Sarris, I. *et al.* Intra- and interobserver variability in fetal ultrasound measurements. *Ultrasound in Obstetrics & Gynecology* **39**, 266–273 (2012).
7. Cavallaro, A. *et al.* Quality control of ultrasound for fetal biometry: results from the INTERGROWTH-21st Project. *Ultrasound in Obstetrics & Gynecology* **52**, 332–339 (2018).
8. Drukker, L., Noble, J. A. & Papageorghiou, A. T. Introduction to artificial intelligence in ultrasound imaging in obstetrics and gynecology. *Ultrasound in Obstetrics & Gynecology* **56**, 498–505 (2020).
9. Burgos-Artizzu, X. P. *et al.* Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. *Scientific Reports* **10**, 10200 (2020).
10. Jang, J. *et al.* Automatic Estimation of Fetal Abdominal Circumference from Ultrasound Images. *IEEE Journal of Biomedical and Health Informatics* (2017) doi:10.1109/JBHI.2017.2776116.
11. Kim, B. *et al.* Machine-learning-based automatic identification of fetal abdominal circumference from ultrasound images. *Physiological Measurement* **39**, 105007 (2018).
12. Salomon, L. J. *et al.* Feasibility and reproducibility of an image-scoring method for quality control of fetal biometry in the second trimester. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* **27**, 34–40 (2006).
13. The DicomAnonymizerTool - MircWiki. http://mircwiki.rsna.org/index.php?title=The_DicomAnonymizerTool.
14. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* **2016-December**, 770–778 (2015).

Status:
SUBMITTED

Images

Fig 2: Abdominal image before cropping.

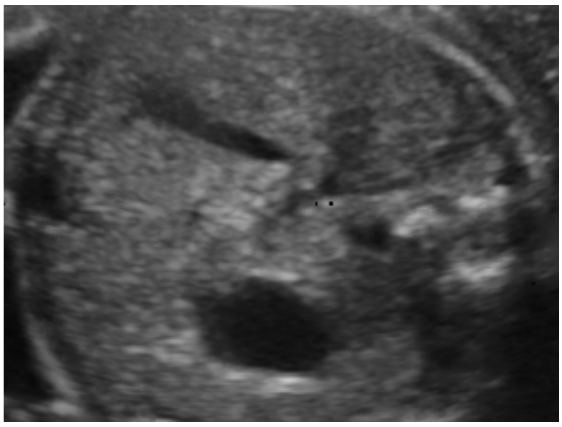

Fig 1: Abdominal image after cropping.

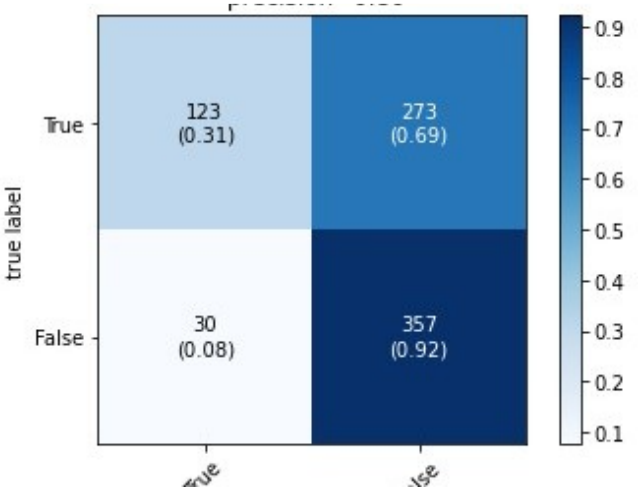

Fig 4: Confusion matrix of the model's performance on symmetry asse...

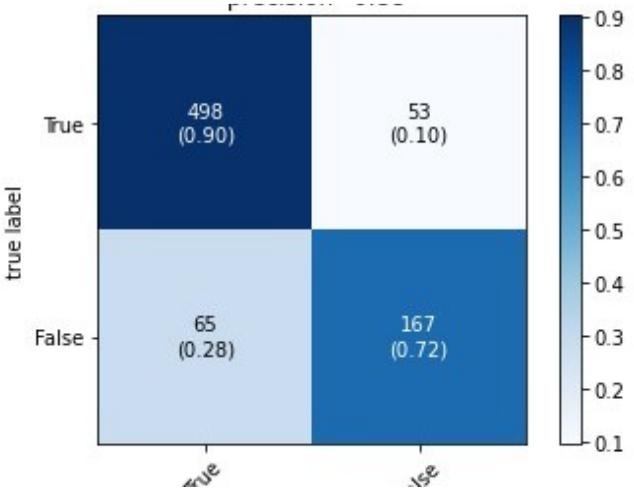

Fig 5: Confusion matrix of the model's performance on stomach bubbl...

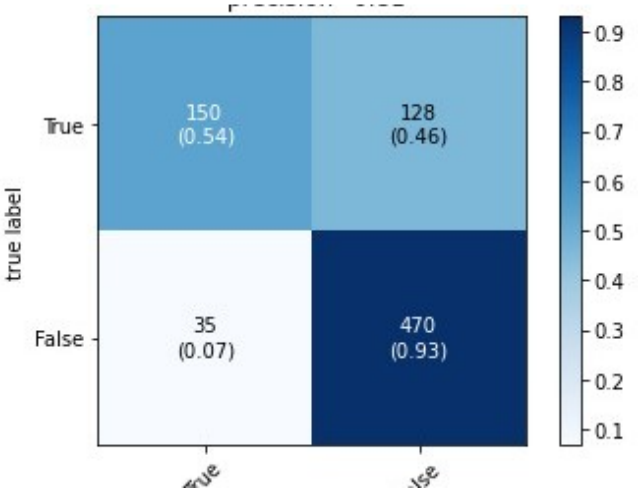

Fig 6: Confusion matrix of the model's performance on portal sinus ...

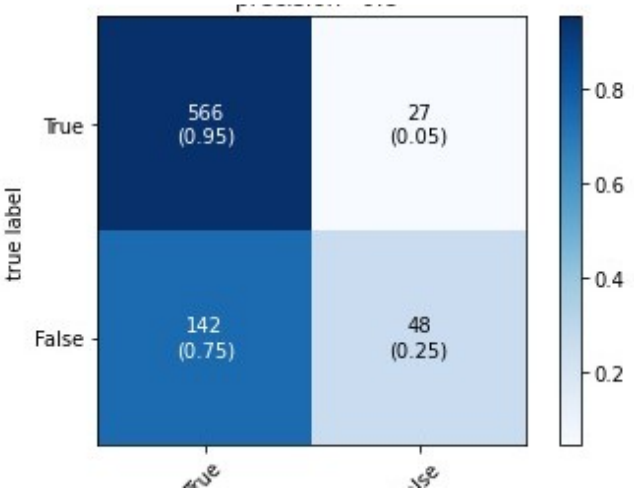

Fig 3: Confusion matrix of the model's performance on kidney visibi...