Zoom Pattern Signatures for Fetal Ultrasound Structures

Mohammad Alsharid[1,3*], Robail Yasrab[1], Lior Drukker[2,4], and Aris T. Papageorghiou[2], J. Alison Noble[1]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK
[2] Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK
[3] Department of Computer Science, Khalifa University of Science and Technology, Abu Dhabi, UAE
[4] Rabin Medical Center, Tel-Aviv University Faculty of Medicine, Tel Aviv, Israel

Abstract. During a fetal ultrasound scan, a sonographer will zoom in and zoom out as they attempt to get clearer images of the anatomical structures of interest. This paper explores how to use this zoom information which is an under-utilised piece of information that is extractable from fetal ultrasound images. We explore associating zooming patterns to specific structures. The presence of such patterns would indicate that each individual anatomical structure has a unique signature associated with it, thereby allowing for classification of fetal ultrasound clips without directly reading the actual fetal ultrasound images in a convolutional neural network.

Keywords: Fetal Ultrasound · Zoom · Magnification.

1 Introduction

Magnifying an image to get a better view of a structure is part and parcel of the fetal ultrasound (US) scanning process. This fact is especially true when attempting to capture biometric measurements of interest, such as the crown rump length (CRL) and the nuchal translucency (NT). During a fetal ultrasound scan, a sonographer will zoom in and zoom out as they attempt to get clearer images of the anatomical structures of interest. In this work, we discuss how to use this zoom information which is an under-utilised piece of information. Zoom information is extractable from fetal ultrasound frames. We explore associating zooming patterns to specific structures. The presence of such patterns would indicate that each individual anatomical structure has a unique zoom pattern signature associated with it, thereby allowing for classification of fetal ultrasound clips without directly reading the actual fetal ultrasound images into a convolutional neural network. Magnification matters because structures can be viewed more or or

---

* Corresponding author: mohammad.alsharid@ku.ac.ae

less easily depending on how zoomed in or out an image is. It stands to reason then that consistent zooming patterns may exist when attempting to find the standard planes of these structures.

Some datasets come with specific labels that have to do with magnification settings, e.g. datasets on histopathological breast cancer images with different magnifications levels [15]. In this, work we use the PULSE dataset [9]. The PULSE dataset does not have labels specifically for magnification or zoom; however, in a previous work of ours [1] that has been accepted and peer-reviewed but is yet to be published, we demonstrated how to extract, from ultrasound images, labels that we called RQZ (Reverse Quasi-Zoom) values that are inversely correlated to the zoom values.

$$RQZ \propto \frac{1}{zoom} \tag{1}$$

RQZ values range from 0 to 4, where 0 is most zoomed in and 4 is most zoomed out. RQZ values are inexpensive to obtain, relying only on reading specific pixel values in a raw US image. We also focused on proving experimentally that a classifier can be trained to distinguish between the different zoom levels even when including images of different anatomical structures, rescaling them, and cropping them to make them all consistently have a fan-shape. That paper determined that zoom is meaningful information. US images at different RQZ values are clearly distinguishable. In this work, we mean to take this further and show that we can discover that specific zooming patterns can be attributed to different anatomical structures in the first trimester fetal ultrasound scanning.

We were motivated to work on this paper in an effort to actualise the untapped potential of zoom information. In low-resource and low-compute settings, being able to circumvent processing entire videos and make classifications possible without such processing could potentially be time and resource efficient. Classifying a 301 long sequence of integers is less demanding than classifying a video clip that consists of 301 frames where each frame has a dimension of 224x224 pixels.

1.1   Related Work

Use of Zoom or Magnification Histopathology is the most prominent subdivision within medical imaging that utilizes images at varying magnification levels. For instance, D'Amato et al. [4] employ magnification as a key technique in the classification of histopathological images through their multi-scale approach. Others attempt to identify the optimal magnification level for histopathological images, in order to determine the magnification level at which the best performance can be obtained when training convolutional neural networks to detect breast cancer in histopathological images [17, 2]. There are other ways to consider the zoom level of an image. For example, predicting whether or not the entire chest silhouette area is visible within the US fan-shaped area of the image [6].
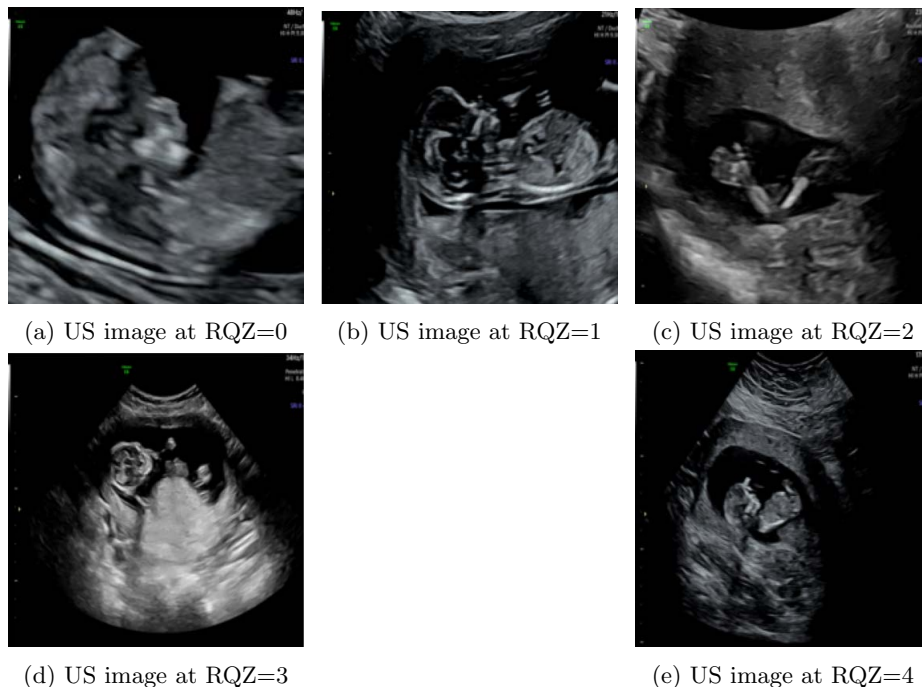
(a) US image at RQZ=0   (b) US image at RQZ=1   (c) US image at RQZ=2

(d) US image at RQZ=3   (e) US image at RQZ=4

Fig. 1: Collection of US images at different RQZ values

Studying Sonographer or Operator Behaviour In this work, we first attempt to associate zooming pattern behaviours to the two most important fetal views in the first trimester, crown rump length and nuchal translucency, proving how a sonographer manipulates the image content during a scan can tell us what anatomical structure is being viewed. We focus on these two views specifically, because in these views, biometric measures are taken once a clear standard plane depicting the CRL or NT views appears. This implies that a certain view needs to be achieved before taking the measurements. It stands to reason that sonographers will attempt to obtain a clearer, easier image to work with for their measurements. This finetuning phase prior to reaching the standard plane is of particular interest. The period of the finetuning can vary. Yasrab et. al [21, 20] operate on the basis that it is 3 seconds (or 90 frames worth of content). Droste et. al [7] analyse probe motion behaviour in order to predict the probe action taken when conducting a fetal US scan until before a frame worthy of being considered standard frame quality is found. Droste et. al look into 10 seconds (300 frames worth of content) before the freeze frame. Teng et. al [19], among other things, attempt to study how the eye gaze tracking patterns are different for an expert and a novice. In their work, they look at a 100 frames before the standard frame.

1.2   Contributions

Fundamentally, we prove that different anatomical views can be distinguished from one another solely based on how a sonographer used the zooming functionality of the ultrasound machine without relying on the fetal ultrasound anatomical content to predict the anatomical class.

2   Methods

Initially, we considered using a window of 2700 frames (or 90 seconds) for the fine tuning phase; however, that proved too long. 90 seconds prior to an NT measurement frame could also cover the fine tuning process and then the acquisition and measurement of a CRL frame. Hence, we shortened the window to 300 frames, as shown in Fig. 2. We had also thought to include the entirety of
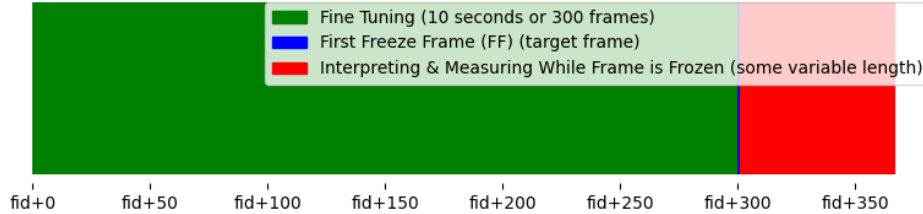

Fig. 2: The first freeze frame is akin to the point at which a sonographer has decided to make a measurement. The sequences, on which we train the 1-D CNN on, end at that first freeze frame. During fine tuning, the sonographer is still attempting to obtain a better view of some anatomical structure. After the approporiate structure or view is identified, the sonographer freezes the frame and proceeds to perform the expected biometry measurements.

freeze frames sequence (so specifying the target frame to be the end of the freeze frames sequence); however since the actual image content is unchanging during said time, it would not have told us much about the changing zooming pattern. It could give us, however, insight into the time spent performing measurements. Our work involves six operators, reducing operator bias.

2.1   Modelling and Analysis

Model Architecture The architecture of the model used was a 1-D convolutional neural network [13, 11, 14]. You can view the architectural breakdown of the model in Table 1.

Model Training Process The cross entropy loss was used was the Adam optimisation algorithm [12]. The model was trained for 1000 epochs.

$$CrossEntropy(y, y^*) = - \left(y \log(y^*) + (1 - y) \log(1 - y^*) \right) \qquad (2)$$

Table 1: Model architecture breakdown for the 1-D CNN. [13, 11, 14]

Layer	Type	Output Channels	Kernel Size	Stride	Padding
1	Conv1d	32	3	1	1
2	MaxPool1d	-	2	2	0
3	Conv1d	64	3	1	1
4	MaxPool1d	-	2	2	0
5	Flatten	-	-	-	-
6	Linear	128	-	-	-
7	Linear	2 (number of classes)	-	-	-

2.2 Evaluation Metrics

Accuracy was calculated as follows:

$$accuracy = \frac{correct}{total} \times 100 \tag{3}$$

where *correct* is the number of correct predictions, and *total* is the total number of predictions made. Other evaluation metrics we have used include precision, recall, and the F1-score. All of which can be found in Table 2.

3 Experiments

In the PULSE study [9, 8, 5], Drukker et al. acquired data from full-length fetal ultrasound scans, which included ultrasound scan videos. Drukker et al. made use of a state-of-the-art Voluson E8 version BT18 ultrasound machines from General Electric Healthcare, based in Zipf, Austria [10]. These machines were equipped with both standard curvilinear probes (C2-9-D, C1-5-D) and 3D/4D probes (RAB6-D) to conduct the ultrasound examinations featured in the PULSE study.

In our previous work [1], we determined that zoom is meaningful by doing various tests with our RQZ classifier which is based on the efficientnet-b0 model [18]. The RQZ classifier was able to achieve 99.42% accuracy in RQZ value classification, demonstrating near complete success at predicting an ultrasound image's RQZ level (even with attempts to manipulate an image post-acquisition by rescaling and cropping images).

In this work, we look at the target frames, and the 300 frames before each target frame, this covers the 10 seconds of finetuning before a standard frame. Target frame is the first freeze frame in the sequence of freeze frames when a standard plane is likely to be be found. It is important to note that these freeze frames are obtained from the process introduced by Yasrab et. al [21, 20]. They had trained a model to identify the frames that are standard-frame worthy. The raw dataset they used had originally come from the work of Drukker et. al [9], but Yasrab et. al [21, 20] used optical character recognition (OCR) to determine when during the scan did the sonographer save the present image. The sonographers' intention to save the image was considered to be an indication that

the standard frame had been obtained and a suitable frame at which measurement can take place had been reached. With these labels, Yasrab et. al. trained a model that could then be used to identify the standard frame quality freeze frames automatically. We use the predictions of their trained model. The first of such frames in a sequence that have been automatically labelled is considered to be our target frame.

4   Results and Discussion

An accuracy of 71.43% is promising for this initial endeavour and shows that there is indeed a distinct nature to the way that CRL and NT views are approached by sonographers as Figs. 4 and 5 show. In Fig. 4, the image content, for both structures, starts at a more zoomed out state but becomes relatively more zoomed in the further into the clip we go. This behaviour makes sense because as sonographers draw closer to the specific ultrasound image content they are trying to acquire and measure, we expect them to zoom in to get a better view of it. We also notice that for CRL the RQZ value is higher than that of NT throughout the average clip. This makes sense, since with NT, the sonographer is trying to focus on observing the fluid in the nuchal fold at the back of the fetus's neck. On the other hand, with CRL images, the sonographer is attempting to measure the entirety of the fetus from its crown to its rump, and so, as one would expect CRL images are on average more zoomed out than NT images, where the focus is on the back of the neck.

We believe that the results are not spurious, primarily from the fact that intuitively the results make strong clinical sense and are clinically explainable. The CRL measurement requires the entire fetus to be visible on the screen before a measurement is made (therefore zoomed out), while NT requires a specific part behind the fetus's head to be focused on for a measurement (therefore zoomed in to fill 75% of the screen according to the guidelines) [3]. At this age, the CRL is roughly 65 ± 19mm, while NT is usually around 1.1 to 3.0 mm, so to view them well, one would expect different zoom levels for each. [16] We discuss the main reason that have prevented a higher score from currently being achieved. Primarily, the data; only 121 sequences were used for training and validation, and 21 sequences were used for testing.

Table 2: Classification report for the anatomical classes, CRL and NT.

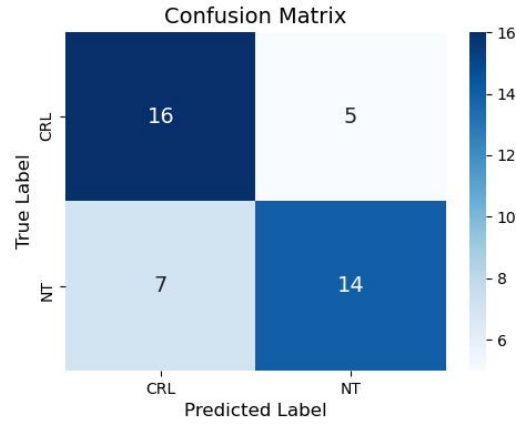|	Precision	Recall	F1-Score	Support
CRL	0.70	0.76	0.73	21
NT	0.74	0.67	0.70	21
Accuracy			0.71	42
Macro Avg	0.72	0.71	0.71	42
Weighted Avg	0.72	0.71	0.71	42

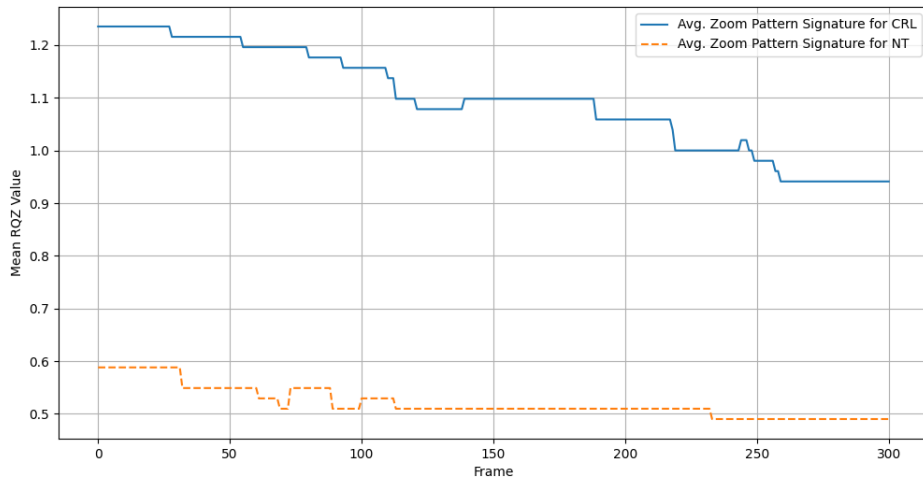Fig. 3: Confusion matrix of the zoom signature prediction task.

Fig. 4: Comparison of the average zoom pattern signatures. The sequences in the training set are used for this plot. This figure serves as qualitative evidence on how CRL and NT differ in terms of zoom.
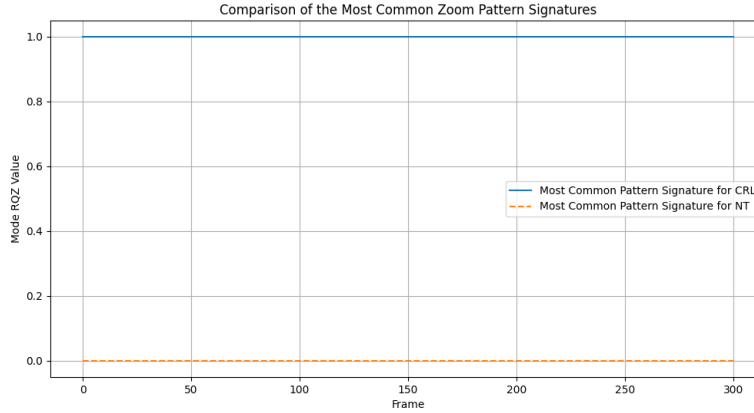
Fig. 5: The Most Common Pattern Signature for each structure. The Most Common RQZ Value at a given frame is determined by the mode at each index in the sequence.

5   Conclusion

In this paper, we show that CRL and NT images exhibit distinct zooming pattern signatures that allow them to be distinguished from each other. One possible use of the proposed approach is faster classification, rather than requiring the direct classification of a huge number of video clips. We can sidestep that by attempting to classify the video by observing the zoom, or magnification, values and how they change to determine the structure being classified.

In the future, we will first retrain the RQZ classifier on the fine-tuning frames before using the RQZ classifier's predictions in our analysis of zoom signature patterns. In this current work, we have considered clips of the two primary biometry measurements of the first trimester being acquired (CRL and NT). In the future, we will expand to other investigating the zoom pattern signature of other clips of interest, and how zooming behaviour may change with changes in the subject's BMI and the positioning of the fetus. There are additional biometry measurements in the second trimester whose zoom pattern signatures we would also be interested in. Investigating the zoom pattern signatures of other gestational ages could be worthwhile as well. In this work, we have put particular focus on 300 frames (so 10 seconds) before the target frame as the fine-tuning phase. We could explore different size windows for the fine-tuning including 90 frames and 100 frames in the future.

Acknowledgements

Disclosure of Interests

The authors have no financial conflicts of interest to disclose related to this work.

References

1. Alsharid, M., Yasrab, R., Drukker, L., Papageorghiou, A.T., Noble, J.A.: Zoom is meaningful: Discerning ultrasound images' zoom levels. In: Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI) (2024), to be published
2. Ashtaiwi, A.: Optimal histopathological magnification factors for deep learning-based breast cancer prediction. Applied System Innovation **5**(5), 87 (2022)
3. Bilardo, C., Chaoui, R., Hyett, J., Kagan, K., Karim, J., Papageorghiou, A., Poon, L., Salomon, L., Syngelaki, A., Nicolaides, K.: Isuog practice guidelines (updated): performance of 11–14-week ultrasound scan. Ultrasound in Obstetrics and Gynecology **61**(1) (2023)
4. D'Amato, M., Szostak, P., Torben-Nielsen, B.: A comparison between single-and multi-scale approaches for classification of histopathology images. Frontiers in Public Health **10**, 892658 (2022)
5. Department of Engineering Science, University of Oxford: Pulse (2019), https://www.eng.ox.ac.uk/pulse/, available at https://www.eng.ox.ac.uk/pulse/
6. Dong, J., Liu, S., Liao, Y., Wen, H., Lei, B., Li, S., Wang, T.: A generic quality control framework for fetal ultrasound cardiac four-chamber planes. IEEE journal of biomedical and health informatics **24**(4), 931–942 (2019)
7. Droste, R., Drukker, L., Papageorghiou, A.T., Noble, J.A.: Automatic probe movement guidance for freehand obstetric ultrasound. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part III 23. pp. 583–592. Springer (2020)
8. Drukker, L., Sharma, H., Karim, J., Droste, R., Noble, J., Papageorghiou, A.: Clinical workflow of sonographers performing fetal anomaly ultrasound scans: deep-learning-based analysis. Ultrasound in Obstetrics & Gynecology **60**(6), 759–765 (2022)
9. Drukker, L., Sharma, H., Droste, R., Alsharid, M., Chatelain, P., Noble, J.A., Papageorghiou, A.T.: Transforming obstetric ultrasound into data science using eye tracking, voice recording, transducer motion and ultrasound video. Scientific Reports **11**(1), 14109 (2021)
10. GE Healthcare: Voluson e8 (2023), available at https://www.gehealthcare.co.uk/products/ultrasound/voluson/voluson-e8
11. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016), http://www.deeplearningbook.org

12. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

13. Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, D.J.: 1d convolutional neural networks and applications: A survey. Mechanical systems and signal processing **151**, 107398 (2021)

14. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. nature **521**(7553), 436–444 (2015)

15. Mayouf, M.S., Dupin de Saint-Cyr, F.: Curriculum incremental deep learning on breakhis dataset. In: Proceedings of the 2022 8th International Conference on Computer Technology Applications. pp. 35–41 (2022)

16. Nicolaides, K.: Screening for chromosomal defects (2003)

17. Patra, A., Behera, S.K., Barpanda, N.K., Sethy, P.K.: Effect of microscopy magnification towards grading of breast invasive carcinoma: An experimental analysis on deep learning and traditional machine learning methods. Ingénierie des Systèmes d'Information **27**(4) (2022)

18. Tan, M., Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. pp. 6105–6114. PMLR (2019)

19. Teng, C., Drukker, L., Papageorghiou, A.T., Noble, J.A.: Skill, or style? classification of fetal sonography eye-tracking data. In: Annual Conference on Neural Information Processing Systems. pp. 184–198. PMLR (2023)

20. Yasrab, R., Alsharid, M., Sarker, M.M.K., Zhao, H., Papageorghiou, A.T., Noble, J.A.: Automated description and workflow analysis of fetal echocardiography in first-trimester ultrasound video scans. In: 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2023)

21. Yasrab, R., Fu, Z., Zhao, H., Lee, L.H., Sharma, H., Drukker, L., Papageorgiou, A.T., Noble, J.A.: A machine learning method for automated description and workflow analysis of first trimester ultrasound scans. IEEE Transactions on Medical Imaging (2022)