Program Digest, 2019 IEEE International Ultrasonics Symposium (IUS)
Glasgow, Scotland, October 6-9, 2019

Deep-Learning based Identification of Frames Containing Foetal Gender Region During Early Second Trimester Ultrasound Scanning

Lakra, Praveen Paul
Department of Engineering Design
Indian Institute of Technology Madras
Chennai, India
p.paullakra@gmail.com

Kumar, Ashish
Department of Engineering Design
Indian Institute of Technology Madras
Chennai, India
ashishkumar191295@gmail.com

Mohanram, Narayanaswamy
Institute of Obstetrics and Gynecology
Madras Medical College
Chennai, India
nmohanram@gmail.com

Krishnamurthi, Ganapathy
Department of Engineering Design
Indian Institute of Technology Madras
Chennai, India
gankrish@iitm.ac.in

Thittai, Arun Kumar
Department of Applied Mechanics
Indian Institute of Technology Madras
Chennai, India
akthittai@iitm.ac.in

***Abstract*—Sex-selective abortions, based on ultrasound foetal scanning of pregnant women, is a rampant problem in developing countries like India. Although prenatal gender screening has been outlawed, poor enforcement of it has abetted a steady rise in their number. Currently, the B-scans are displayed on the screen in real-time and the operator can make out the gender from the displayed image. Therefore, enforcement of the secrecy has not been effective and has only led to limitations on the usage of the technology. Hence, it may be useful to develop a method that can identify frames containing gender-indicative features in real-time and block it out automatically from the display, thus, preventing unauthorized viewing. In this work, deep learning-based techniques have been explored to detect images containing the gender-defining features among the entire set of images in cine-loop, with a conforming accuracy averaging above 80%.**

***Index Terms*—foetal ultrasound, foetal gender, neural networks, imaging, sex-selective abortion, frame filtering**

I. INTRODUCTION

Census data of the previous few decades indicate that sex-selective abortion is increasing in many parts of the world [1] [2]. Two of the worlds most populous countries, China and India, are among the top four countries with the highest number of cases of female infanticides, with male to female ratio being 1.15 and 1.12 at birth, respectively. Although prenatal gender screening has been outlawed, poor enforcement of it has abetted a steady rise in their number. For example, there is a legal act in India prohibiting pre-natal gender diagnosis known as the Pre-Conception and Pre-Natal Diagnostic Techniques Act (hereafter referred to as PCPNDT) of 1994. It was amended in 2003 to include the usage of ultrasound and sales of such machines only to registered bodies. PCPNDT also has provisions to bar the communication of gender of the foetus to the pregnant woman or her relatives through any means. Currently, the onus lies on the operator to not reveal the gender details from the ultrasound scan. However, enforcement of this has been challenging because the B-scans are displayed on the screen in real-time and the operators can make out the gender from the displayed image. Therefore, the focus of the work reported in this paper is on developing a method that can prevent unauthorized viewing of the ultrasound frames that contain foetal gender region during real-time scanning. This may allow for better enforcement of acts like e.g., PCPNDT act in India.

Obstetric ultrasonography is the use of medical ultrasound in pregnancy. The procedure is a standard part of prenatal care in many countries, as it can provide a variety of information about the health of the mother, the timing and progress of the pregnancy, and the health and development of the embryo or foetus. The International Society of Ultrasound in Obstetrics and Gynaecology (ISUOG) recommends routine obstetric ultrasounds between 18 weeks' and 22 weeks' gestational age (the anatomy scan) in order to confirm pregnancy timing, to measure the growth of the foetus, recognize any abnormalities quickly in pregnancy, and assess multiple pregnancies (twins, etc) [3]. The sex of the foetus may be discerned by ultrasound as early as 11 weeks' gestation. The accuracy is relatively imprecise when attempted early [4] [5] [6]. After 13 weeks' gestation, high accuracy of between 99% and 100% is possible if the foetus does not display intersex external characteristics [7]. Table I shown below highlights the accuracy of determining the gender of the foetus at different gestational ages. It is important to note that the accuracies obtained in determining the gender of the foetus from the ultrasound images were done by trained professionals.

Deep learning models have shown great results in the domain of image classification in the recent past. In the case of natural images, deep neural networks have achieved accuracies similar to the human performance [8]. However, the use of deep learning in the domain of medical images is still an upcoming area of research. In this work, the use of deep learning approach was explored to identify a frame containing gender indicative features so that it may potentially blocked from the display thus preventing unauthorized viewing.

TABLE I: SEX DETERMINATION ACCURACY REPORTED IN LITERATURE

Gestational Age	Kings College Hospital Medical School [5]	Taipei City Hospital & Li Shin Hospital [6]
11 weeks	70.3%	71.9%
12 weeks	98.7%	92%
13 weeks	100%	98.3%

II. Methodology

A. Data Pre-processing

Ultrasound B-mode foetal scan data were obtained from 50 women, 12 to 20 weeks into their pregnancy. This pregnancy time frame was chosen because the gender of the foetus becomes visible in ultrasound scans during this period and abortion in India is legal until 20 weeks of pregnancy. A minimum of one cine loop of ultrasound image frames was obtained from each patient. Every possible view was accounted for while obtaining the scans. The scans predominantly contain the top view of the foetus where the visibility of gender region is not hindered by other parts of the foetus like legs, hands and umbilical cord. Data obtained from the scans were in DICOM 3D format as generated by the machines. Data were then converted and saved in the standard neuroimaging format NIfTI [9]. After conversion, the images were annotated with the help of an expert physician practitioner. This served as the ground truth for the network architecture that was used. After annotation, this image data was resized to 230x230 for training purposes.

B. Model Architecture and Training

As observed from the data statistics reported in Table II, there was a heavy imbalance among the number of frames that contained gender-defining features compared to the frames that did not contain such features. Training a deep neural network with such a high imbalance among the different classes can lead to the model getting biased towards one class, often giving poor results on images of the other class. Hence, to remove this imbalance in the data classes, additional copies of the less represented class were created (images containing the gender-defining region), such that its proportion is equal to the other class (images that do not contain gender region). Data augmentation was done by horizontal flipping or rotation, whilst maintaining the aspect ratio of the images.

This work evaluated the performance of many state-of-the-art deep learning architectures and finally adapted a residual network architecture. Residual networks came into prominence after being declared the winning algorithm for the ImageNet classification challenge (ILSVRC 2015) [10]. Unlike a traditional neural network where each layer feeds the next layer, in residual networks each layer not only feeds the next layer but also the layer that is 2-3 jumps away using what is called as skip connections or residual connections as shown in figure 1. The layers in this residual network try to learn the difference of the true output and the input to the layer. The model architecture created is a residual network with 23 convolutional layers, as shown in Table III, to address the twin problems of underfitting and overfitting when the commonly used ResNet-18 and ResNet-34 architecture are used, respectively.

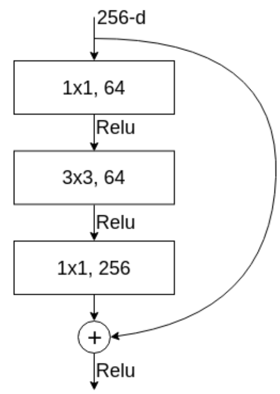


Fig. 1: A schematic of one residual block used in ResNet is shown.

The optimizer used for the network is AdaMax, instead of Stochastic Gradient Descent (SGD) with Nesterov momentum, because AdaMax is better at picking up sparse information because of the $1/\sqrt{t}$ decay term for the learning rate. Further, AdaMax is better than Adam as it uses L_∞ norm that helps in removing initialization bias and reducing the influence of gradient noise in learning rate update, which are issues with Adam that uses L2-norm. [11]. A batch size of 16 was chosen as bigger sizes perform poorly when it comes to generalizing the learning features and smaller sizes were unable to converge to a minima within the number of iterations specified in the training. The loss function chosen to evaluate the model was binary cross-entropy because this loss function can deal with binary classification problems. As a result, the output ranges between 0 and 1, portraying the probability of a frame containing gender indicative features or not. Training parameters are listed in Table IV. Five-fold cross-validation was performed on the network to also check for over-fitting.

III. Results and Discussion

The network was trained on randomly captured videos, which were not specifically taken to identify gender or the gender region of the foetus, as a routine obstetric scan. The data split was strongly biased towards images containing no gender region, but the network was still able to do a satisfactory job with a high recall rate. Images containing gender region only comprised 27.89% of the total number of

TABLE II: DATA STATISTICS

Fold	Images containing gender region	Images devoid of gender region
1	1250	2888
2	1010	3348
3	1681	2614
4	757	3238
5	1329	3497
Total	6027	15585

TABLE III: MODEL ARCHITECTURE

Number of layers	23 layer
Convolution 1	7x7, 64, stride 2
Convolution 2	$\left\{\begin{array}{l} 1\times 1, 64 \\ 3\times 3, 64 \\ 1\times 1, 256 \end{array}\right.$ x 1
Convolution 3	$\left\{\begin{array}{l} 1\times 1, 128 \\ 3\times 3, 128 \\ 1\times 1, 512 \end{array}\right.$ x 2
Convolution 4	$\left\{\begin{array}{l} 1\times 1, 256 \\ 3\times 3, 256 \\ 1\times 1, 1024 \end{array}\right.$ x 2
Convolution 5	$\left\{\begin{array}{l} 1\times 1, 512 \\ 3\times 3, 512 \\ 1\times 1, 2048 \end{array}\right.$ x 2
	max pooling, 1-d fully connected, sigmoid

TABLE IV: TRAINING PARAMETERS

Optimizer	AdaMax
Loss function	Binary cross-entropy
Batch Size	16
Number of GPUs used	1

images. After performing the cross-validation, the accuracy of the network was found to be 83.56% on average. The true positive rate (recall) of the network was 82.73% and positive predictive value (precision) was 70.37% on average. The F1 score (harmonic mean of recall and precision) comes out to nearly 69.57. Refer to Table V and Table VI for the results of the training.

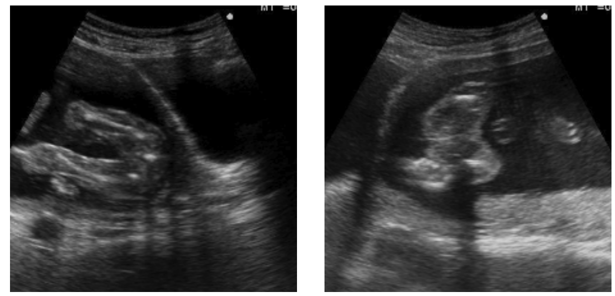
TABLE V: FIVE FOLD CROSS VALIDATION TRUTH MATRIX

Fold	True Positive	False Positive	False Negative	True Negative
1	904	471	346	2417
2	652	348	358	3000
3	1180	280	501	2334
4	465	177	292	3061
5	1035	497	294	3000

Figure 2(a) shows an example image in which gender region is visible to the human eye and the network was able to flag the frame. Figure 2(b) shows an image that has no gender region visible but the network confused the umbilical cord to a female gender determining feature and flagged the frame. Figure 3(a) shows an image that has a gender region. The neural network was unable to make out the gender region possibly because it might have considered the anatomy as apart of the background, due to low contrast in the image. Figure 3(b) shows another example image that has gender region visible, which the network was unable to detect. The network probably mistook the gender of the foetus as a separate part because of the lack of visible connection between the organ and the body.

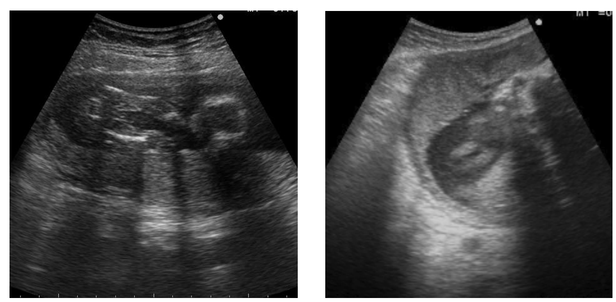
TABLE VI: FIVE FOLD CROSS VALIDATION RESULTANT SCORES

Fold	Validation Accuracy	True Positive Rate (Recall)	Positive Predictive Value (Precision)	F1 Score
1	80.27%	72.35%	65.76%	68.90
2	83.79%	64.59%	65.16%	64.87
3	81.83%	70.21%	80.84%	75.15
4	88.28%	61.47%	72.50%	66.53
5	83.63%	77.93%	67.59%	72.39



(a) (b)

Fig. 2: Example images of foetal scans where; (a) contains the gender region that was correctly detected by the network, a true positive; and (b) does not contain the gender region that was incorrectly detected by the network, a false positive.



(a) (b)

Fig. 3: Example images of foetal scans where; (a) contains the gender region that was incorrectly detected by the network, a false negative; and (b) contains the gender region that was incorrectly detected by the network, another false negative.

IV. Conclusion

This paper reported the evaluation of using neural networks to identify and flagging the ultrasound frames containing gender-indicative features of a foetus from real-time B-mode images. The results demonstrate that it is feasible to use a trained network to automatically detect an ultrasound image frame that may contain gender-indicative features. Importantly, the network learnt to capture the necessary features in the noisy ultrasound image with no provision for noise removal from the input fed to the network.

V. Acknowledgement

The authors would like to thank the National Health Mission - Tamil Nadu, for providing necessary approvals for data collection.

References

[1] Population Division of Department of Economic and Social Affairs, *World Population Prospects 2019: Highlights*, ser. World Population Prospects. United Nations, 2019. [Online]. Available: https://www.un.org/en/development/desa/population/index.asp

[2] Census of India Ministry of Home Affairs Government of India. (2017) Sample registration system statistical reports. [Online]. Available: http://censusindia.gov.in/vital_statistics/SRS_Statistical_Report.html

[3] L. J. Salomon, Z. Alfirevic, V. Berghella, C. Bilardo, E. Hernandez-Andrade, S. L. Johnsen, K. Kalache, K.-Y. Leung, G. Malinger, H. Munoz, F. Prefumo, A. Toi, W. Lee, and on behalf of the ISUOG Clinical Standards Committee, "Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan," *Ultrasound in Obstetrics & Gynecology*, vol. 37, no. 1, pp. 116–126, 2011. [Online]. Available: https://obgyn.onlinelibrary.wiley.com/doi/abs/10.1002/uog.8831

[4] E. Merz, *Ultrasound in Obstetrics and Gynecology, Volume 1 Obstetrics*. Thieme, 2011. [Online]. Available: https://books.google.co.in/books?id=gVx-L9_NYU0C

[5] Z. Efrat, O. O. Akinfenwa, and K. H. Nicolaides, "First-trimester determination of fetal gender by ultrasound," *Ultrasound in Obstetrics & Gynecology*, vol. 13, no. 5, pp. 305–307, 1999. [Online]. Available: https://obgyn.onlinelibrary.wiley.com/doi/abs/10.1046/j.1469-0705.1999.13050305.x

[6] C. HSIAO, H. WANG, C. HSIEH, and J. HSU, "Fetal gender screening by ultrasound at 11 to 13+6 weeks," *Acta Obstetricia et Gynecologica Scandinavica*, vol. 87, no. 1, pp. 8–13, 2008. [Online]. Available: https://obgyn.onlinelibrary.wiley.com/doi/abs/10.1080/00016340701571905

[7] B. L. Whitworth, M and C. Mullan, "Ultrasound for fetal assessment in early pregnancy," *Cochrane Database of Systematic Reviews*, no. 7, 2015. [Online]. Available: https://doi.org//10.1002/14651858.CD007058.pub3

[8] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, and et al., "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, p. 211252, Apr 2015. [Online]. Available: http://dx.doi.org/10.1007/s11263-015-0816-y

[9] P. A. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[11] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv e-prints*, p. arXiv:1412.6980, Dec 2014.