AUTOMATIC DETECTION OF LOCAL FETAL BRAIN STRUCTURES IN ULTRASOUND IMAGES

M. Yaqub[1], R. Napolitano [2], C. Ioannou[2], A. T. Papageorghiou[2], J. A. Noble[1]
[1] Institute of Biomedical Engineering, [2] Nuffield Dept of Obstetrics & Gynaecology
University of Oxford, UK

ABSTRACT

Although ultrasound imaging is the main modality to study fetal growth, the analysis of such images is an under-studied area of research especially using 3D ultrasound. In this paper we propose an automatic technique to locate four local fetal brain structures in 3D ultrasound images. Clinically, the localization of such structures in 3D is hard and subjective. However, structure localization is critical to detect some brain abnormalities and to do fetal biometry. The technique we propose is based on a discriminative model (Random Forests) which is gaining a lot of interest recently. The novelty of this work lies in 1) the implicit integration of image features and the relative position of structures in the fetal brain within the technique and 2) the application itself since it is very challenging. We report promising first results which are consistent with published literature on visual detection of fetal brain structures, and suggest that automated analysis of 3D fetal neurosonography may be possible in the future.

Index Terms— Ultrasound Imaging, Random Decision Forests, Object Detection, Fetal Brain, Fetal neuro-sonography.

1. INTRODUCTION

Ultrasound (US) imaging is radiation free, inexpensive and real-time, and is the main imaging modality used in monitoring pregnancy. Although fetal MRI is safe, and can give images of higher anatomical definition (after motion correction), it is not approved for use below 20 week gestation and is not cost-effective or practical for general abnormality screening. 2D US is the current modality of choice for fetal growth assessment in the clinic. However, there is a lot of interest in determining the merits of using 3D US for this purpose as 3D acquisition is quick and the quality of a 3D scan is now approximately that of 2D scans. The downside is that currently it takes some skill to find the best 2D viewing planes in a 3D scan, thus negating the acquisition time benefits from the perspective of workflow analysis.

Underpinning finding a good viewing plane is finding good anatomical structures that define that imaging plane. This topic is the subject of this paper where we focus on doing this for the fetal brain (fetal neuro-sonography).

Identification of fetal brain structures is an essential part of basic prenatal screening for abnormalities [1]. It is also necessary for accurate estimation of fetal head circumference, which is used in dating pregnancies beyond 14 weeks. Head circumference is measured on either the trans-thalamic or trans-ventricular planes [2], Figure 1. These planes are located by a clinician via visually identifying several important structures [2], e.g., cavum septum pellucidum, posterior ventricle, etc. Finding such structures is not straightforward and highly subjective since the visibility of fetal brain structures are related to many conditions like the US machine, maternal tissue, orientation of the fetus (this changes the attenuation of signal and as a result the visibility of some structures), sonographer experience and many others. Therefore, automatic detection of fetal brain structures is important because it is the first and most important step toward accurate fetal biometry and abnormality detection.

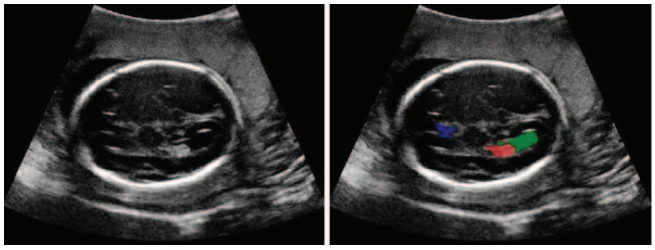


Figure 1. Original image that represents a trans-ventricular plane (left) and the same image with manual delineation of three structures superimposed. *Green* is the choroid plexus, *red* is the posterior ventricle cavity and *blue* is the cavum septum pellucidum.

In the literature, there are only a handful of papers published on detecting fetal structures. In [3], the authors used a method based on the probabilistic boosting tree (PBT) algorithm to detect large (or global) fetal structures (abdomen, femur, head) in 2D US automatically. The detection was used to define regions-of-interest for basic fetal biometry and for quite large structures. In [4] the fetal stomach and umbilical vein were automatically detected in 2D fetal abdominal US to define whether the imaging plane was correctly acquired or not. However, both these works only look at 2D US images (though [4] has used their approach to identify the best plane in a 3D scan (2.5D detection)).

In this paper, we propose a novel and efficient automatic method to detect local 3D fetal brain structures in 3D US. The method is based on Random Forests (RF), a general machine learning framework gaining popularity in biomedical image analysis (and for analysis of US images which are one of the most challenging modalities to work with) because of its power to embed simple constraints on image interpretation. The novelty of this paper is in the design of the solution based on this framework and the way that we incorporate image and relative positions features. To our knowledge this is the first paper considering detection of local 3D fetal structures in 3D US images. It is also important to stress that we target to locate subtle structures and not to perform global organ/structure detection as in for instance [3].

The use of the RF technique for medical image segmentation and detection is relatively new. Lempitsky et al. [5] used the standard binary RF to automatically delineate the myocardium in 3D echocardiography, later extending the work to the temporal (2D+T) case [6]. Yi et al. [7] segmented three brain tissues from 20 MRI volumes using the traditional RF technique. Geremia et al. [8] also applied the traditional RF technique to segment multi-channel brain MR images for analysis of multiple sclerosis. Yaqub et al. [9] proposed a modification to the traditional RF segmentation technique validated on MRI brain structures and the fetal femur in 3D US. Criminisi et al. [10-12] have extensively looked at object detection using regression forests to automatically detect 3D bounding boxes around several organs in CT and MRI volumes. The purpose of this paper is not to compare the RF technique with other techniques since this has been done before [13-14] but to demonstrate its utility for the task of local fetal structure detection for the first time.

2. METHOD

Our fetal structure detection solution is designed using the Random Forests framework. A classifier is built to classify voxels to one of five classes; *background*, *Choroid Plexus (CP), Posterior Ventricle Cavity (PVC), Cavum Septum Pellucidum (CSP) and Cerebellum (CER)*.

2.1. Dataset

40 fetal US brain volumes were used in our experiments. Fetal gestational age was 19-24 week gestation. The only assumption made during 3D acquisition was that the 3D acquisition was roughly performed on the trans-thalamic plane which is a common scanning convention in 3D fetal brain acquisition. Four structures (CP, PVC, CSP and CER) were manually segmented on 20 volumes by one experienced sonographer while the other 20 volumes have no manual segmentation. The manual segmentation was performed on only one hemisphere (distal) because brain structures are usually highly attenuated on the proximal hemisphere. See Figure 1.

2.2. Training

A classifier was trained using the RF technique [5, 7-9, 13, 15]. The RF classifier is a set of multiple random decision trees. Training requires a set of training examples (e.g., voxels and their class label). A random training set from the original training set is created for each tree (bagging) [16]. Each internal node in every tree is a classifier that consists of a feature and a threshold. Weak features are usually used in such techniques [5, 7-11, 13, 15]. To create an internal node in a tree, a small subset of features is randomly selected and the best feature and its best threshold chosen as a classifier on that node. Maximizing the gain of information is used to find the best randomly selected feature/threshold. A split of the training examples between two branches happens after finding the best classifier. The process continues recursively at each branch node until reaching a leaf node. Leaves are created when either the maximum allowed depth is reached or no gain of information is achieved if a split happens. Associated with each leaf is a histogram of the training examples that reached it.

In this paper, we used Haar, cuboid, binary and unary features as in [9, 15]. We employed absolute and relative position features which are important to 1) associate structures to each others, e.g., the posterior ventricle cavity appears next to the choroid plexus in many slices, and 2) relate the position of the structure to the trans-thalamic plane. The size and displacement of these features were used depending on the physical size of the structures and their relative positions.

Training was performed on 10 volumes that were randomly selected. We experimentally found that the classification error saturated when the number of trees = 9 and overfitting started to happen if the tree depth exceeded 10 so we used a maximum tree depth = 10. Because the images have high variability in appearance and to avoid overfitting, the number of randomly selected features was set low compared to some other published work. We used 20 random features to train a node.

2.3. Classification and Detection

The classification (testing) was performed on 10 separate volumes. The classification of a voxel from a test volume proceeded by traversing it over all trees starting from their roots until it reaches a leaf node. The mean posterior at the leaves from all trees is assigned to the test voxel. A nice property of the RF technique is that it gives a probability of a test example to belong to each class label (in our case there are 5 classes). We utilized this soft assignment to locate the structure of interest. First, a mean probability for each voxel were found within a 3D neighborhood. This was done for each class label separately. Then structures were detected using one of two strategies. In the first approach, a structure was detected by finding its highest probable location in 3D. This was simply done by finding the voxel with the highest probability within a region (neighborhood). The second approach finds the voxel of the maximum

probability for each class on every 2D axial slice (mimicking the process a radiologist would use of scrolling through 2D scans). The selected voxel probability is used as a certainty measure of the structure existence and location. A probability threshold was used to define if a structure exists or not. In this way every 2D slice of the fetal brain can be tested for the existence and location of each structure.

2.4. Validation Methodology

Quantitative and qualitative validation was done. We used the 20 volumes where no manual segmentation was performed in the qualitative evaluation. The structures in these volumes were automatically located using the proposed technique and an expert graded the result by eye as "Correctly localized" and "Not correctly localized". It is important to note that the expert who did the manual segmentation is different than the one who did the grading of the correctly localized structures. This was done to avoid reader bias.

3. RESULTS

3.1. Qualitative Analysis

Figure 2 presents some visual detection results on several 2D slices from different 3D volumes. We show in Figure 2 good (top & middle rows) and bad (bottom row) detection results. Table 1 shows the percentage of correctly localized structures using the proposed technique compared to the expert "by-eye" localization.

3.2. Quantitative Analysis

The quantitative validation was performed on 10 3D manually segmented volumes from the test set. Table 2 shows the 3D detection accuracy of the four brain structures.

Although we show in Table 2 that we achieve an excellent detection in 3D, detection of structures on each 2D slice is clinically important. To address this, we compared the detection of structures in all brain slices with their manual localization. It is important to note that this is a 2D detection of structures but from a 3D classification. Table 3 shows the detection accuracy. In Table 3, we show true negative, false negative, false positive, true positive, precision, recall and accuracy. True negative (TN) represents the number of slices the technique decides that the structure does not exist where it actually does not exist. False negative (FN) represents the number of slices the technique decides that the structure does not exist where it actually does exist. False positive (FP) represents the number of slices the technique decides that the structure does exist where it actually does not exist. True positive (TP) represents the number of slices the technique decides that the structure does exist where it actually exists. These four measures were found by thresholding the probabilities. Precision, recall and accuracy were calculated as

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

Balancing the precision and recall can be achieved by changing the value of the threshold and finding the best threshold. This is also an application dependent choice. Results in Table 3 are reported for the case where the maximum accuracy is achieved.

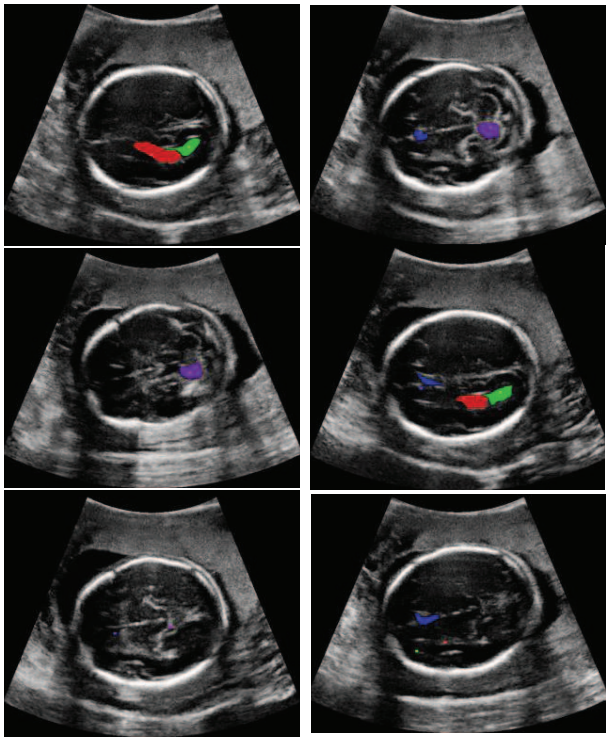


Figure 2. Automatic detection on 2D slices from 3D volumes. The top two rows shows four slices of good detection of the structures while the bottom row shows two slices of incorrect detection of structures. In all images, the colors represent the manual delineation of structures while the small brighter color square is the automatically detected position of a structure. Color codes are *green* for CP, *red* for PVC, *blue* for CSP and *purple* for CER.

Table 1. Correctly localized brain structures. Comparison between the automatic detection and the expert detection by a naked eye.

CP	PVC	CSP	CER
95%	75%	85%	90%

Table 2. 3D detection accuracy of the brain structures. Comparison between the automatic detection and the manual delineation.

CP	PVC	CSP	CER
100%	80%	90%	90%

Table 3. 2D detection from 3D classification. Comparison between the automatic detection and the manual delineation on each 2D slice from the 3D volumes.

	CP	PVC	CSP	CER
True Negative	733	784	813	830
False Positive	32	44	48	47
False Negative	39	45	33	34
True Positive	196	127	106	89
Precision	86.0%	74.3%	68.8%	65.4%
Recall	83.4%	73.8%	76.3%	72.4%
Accuracy	92.9%	91.1%	91.9%	91.9%

4. DISCUSSION AND FUTURE WORK

In this paper we presented an automatic detection technique for four important fetal brain structures in 3D US which is a challenging image modality to analyze. The technique is based on the Random Forests classifier.

Note that the accuracy on detecting the CP is better than detecting the other three structures, see Table 2 & Table 3. However, it is important to discuss why the accuracy in Table 3 is much higher than the precision and recall. Notice in equation (3) that the accuracy is calculated as a measure of detection but taking into consideration the true negatives as well. Since the four structures only appear on a small number of slices compared to the total number of slices in any fetal brain, the technique succeeded in ignoring such slices where none of the four structures appear. Therefore, the true negatives are much higher than the other three measures (TP, FP and FN) and this is the main reason why accuracy is higher than precision and recall.

There are several opportunities for further improvement of this work from both a technical and a clinical point of view. For instance, on the technical side, adding more structures to the model would constrain the method to look for more structures and relate them to each others, e.g., skull and falx. Furthermore, training and testing on rigidly aligned images might provide more accurate localization although it requires more preprocessing.

ACKNOWLEDGEMENT

This work was supported by provision of equipment and funding from Philips Healthcare (Andover, Massachusetts). The authors would like to thank all the mothers who participated in the project.

REFERENCES

[1] L. J. Salomon, Z. Alfirevic, V. Berghella *et al.*, "Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan," *Ultrasound in Obstetrics & Gynecology,* vol. 37, no. 1, pp. 116-126, 2011.

[2] "Sonographic examination of the fetal central nervous system: guidelines for performing the 'basic examination' and the 'fetal neurosonogram'," *Ultrasound in Obstetrics and Gynecology,* vol. 29, no. 1, pp. 109-116, 2007.

[3] G. Carneiro, B. Georgescu, S. Good *et al.*, "Detection and Measurement of Fetal Anatomies from Ultrasound Images using a Constrained Probabilistic Boosting Tree," *Medical Imaging, IEEE Transactions on,* vol. 27, no. 9, pp. 1342-1355, 2008.

[4] A. P. Bahbibi Rahmatullah, and J. Alison Noble, "Automated Selection of Standardized Planes From Ultrasound Volume," in Machine Learning in Medical Imaging, Toronto, Canada, 2011.

[5] V. Lempitsky, M. Verhoek, J. A. Noble *et al.*, "Random Forest Classification for Automatic Delineation of Myocardium in Real-Time 3D Echocardiography," in Proceedings of the 5th International Conference on Functional Imaging and Modeling of the Heart, Nice, France, 2009.

[6] M. Verhoek, M. Yaqub, J. McManigle *et al.*, "Learning Optical Flow Propagation Strategies Using Random Forests for Fast Segmentation in Dynamic 2D & 3D Echocardiography," in Machine Learning in Medical Imaging, Toronto, Canada, 2011.

[7] Z. Yi, A. Criminisi, J. Shotton *et al.*, "Discriminative, semantic segmentation of brain tissue in MR images," in Medical Image Computing and Computer Assisted Intervention, 2009, pp. 558-65.

[8] E. Geremia, B. H. Menze, O. Clatz *et al.*, "Spatial Decision Forests for MS Lesion Segmentation in Multi-Channel MR Images.," in Medical Image Computing and Computer Assisted Intervention, Beijing, 2010.

[9] M. Yaqub, K. Javaid, C. Cooper *et al.*, "Efficient Volumetric Segmentation using 3D Fast-Weighted Random Forests " in Medical Image Computing and Computer Assisted Intervention-MLMI, Toronto, Canada, 2011.

[10] A. Criminisi, J. Shotton, D. Robinson *et al.*, "Regression Forests for Efficient Anatomy Detection and Localization in CT Studies," in MICCAI workshop in Medical Computer Vision 2010: Recognition Techniques and Applications in Medical Imaging, 2010.

[11] A. Criminisi, J. Shotton, and S. Bucciarelli, "Decision Forests with Long-Range Spatial Context for Organ Localization in CT Volumes," in MICCAI workshop on Probabilistic Models for Medical Image Analysis (MICCAI-PMMIA), London, 2009.

[12] O. Pauly, B. Glocker, A. Criminisi *et al.*, "Fast Multiple Organs Detection and Localization in Whole-Body MR Dixon Sequences," in Medical Image Computing and Computer Assisted Intervention, Toronto, 2011.

[13] L. Breiman, "Random Forests," *Machine Learning,* vol. 45, no. 1, pp. 5-32, 2001.

[14] R. Caruana, and A. Niculescu-Mizil, "An Empirical Comparison of Supervised Learning Algorithms," in ICML, 2006.

[15] M. Yaqub, P. Mahon, K. Javaid *et al.*, "Weighted Voting in 3D Random Forest Segmentation," in Medical Image Understanding and Analysis, Coventry, UK, 2010.

[16] L. Breiman, "Bagging predictors," *Machine Learning,* vol. 24, no. 2, pp. 123-140, 1996.