



SPRNet: Automatic Fetal Standard Plane Recognition Network for Ultrasound Images

Jiajun Liang¹, Rian Huang¹, Peiyao Kong¹, Shengli Li²,
Tianfu Wang^{1(✉)}, and Baiying Lei^{1(✉)}

¹ School of Biomedical Engineering, National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Shenzhen University, Shenzhen, China

{tfwang, leiby}@szu.edu.cn

² Department of Ultrasound, Affiliated Shenzhen Maternal and Child Healthcare Hospital of Nanfang Medical University, Shenzhen, China

Abstract. Fetal standard plane recognition is a crucial clinical part of prenatal diagnosis. However, it is also a sophisticated, subjective, and highly empirical process. Thus, there is a huge demand for proposing an effective and precise automatic method to help experienced as well as inexperienced doctors to complete this process, efficiently. In order to satisfy this clinical need, we propose an automatic fetal standard plane recognition network called SPRNet. Specifically, we adopt DenseNet as the basic network of SPRNet and implement data-based partial transfer learning on it by weight-sharing strategy. We then train our network with a task dataset (fetal ultrasound images) and a transferring dataset (placenta ultrasound images) so that our network can discover and learn the potential relationship between these two datasets to improve the performance and avoid overfitting. Finally, we achieve automatic fetal standard plane recognition by utilizing the feature extracted from SPRNet. The experimental results indicate that our network can attain an accuracy of 99.00% and perform better than conventional networks.

Keywords: Fetal standard plane recognition · Data-based partial transfer learning · Fetal ultrasound images · Placenta ultrasound images

1 Introduction

Prenatal diagnosis is an effective examination to assess the growth of fetuses and it is also helpful to reduce birth defect rate and neonatal mortality. Due to the advantages of the non-invasion, no radiation and low cost, ultrasonography plays an important role in prenatal diagnosis, nowadays. This ultrasonography method can be generally divided into five steps: ultrasound images scanning, standard planes recognition, structural observation, parameter measurement, and diagnosis. Among these steps, standard plane recognition is the key part of the process, as the standard planes are the foundation of parameter measurement and directly reveal the congenital anomaly of fetus [1].

Currently, the recognition of standard planes mainly depends on artificial examination. Slight differences exist between standard planes and non-standard planes and an example is shown in Fig. 1. This high similarity between the planes make it hard for sonographers to effectively distinguish the planes and increases likelihood for misdiagnosis, especially when they are working in a high-workload environment. In addition, underdeveloped areas are lacking of experienced prenatal diagnosis doctors. This is detrimental to decline the birth defect rate and neonatal mortality. Therefore, it is great significance to propose an effective and automatic method to help experienced as well as inexperienced sonographers to efficiently distinguish fetal standard planes from non-standard planes.

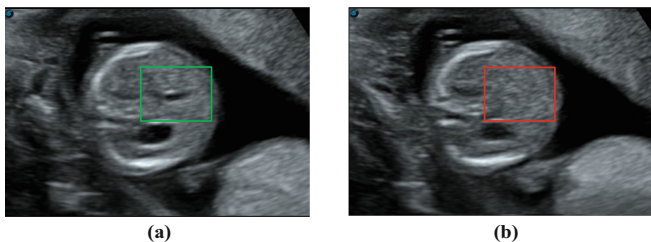


Fig. 1. Abdominal standard plane (a) and non-standard plane (b) appear in two adjacent frames of an ultrasound video. The green and red boxes show the nuanced difference between these two images. (Color figure online)

Recently, the-state-of-the-art deep learning based method, convolutional neural network (CNN) and its variants like VGG [2], ResNet [3] and SeNet [4], showed high performance in different image classification tasks. It can also provide a new insight for researchers to realize automatic fetal standard planes recognition. Accordingly, many works also have been devoted into this area [5, 6]. Although some of these works addressed the automatic recognition of certain fetal standard planes, their frameworks have limitations in the generalization ability and accuracy. Recently, Kong et al. [11] and Cai et al. [12] address these issues by higher performance network and multi-task learning respectively. Inspired by their works, we propose an automatic fetal standard plane recognition network called SPRNet. Specifically, the proposed SPRNet is based on DenseNet architecture [7], which could maximize the use of features and outperform other different deep neural network architectures. However, it still suffers from the problem of overfitting. Inspired by work in Wang et al. [8], we propose a transfer learning method called data-based partial transfer learning to alleviate overfitting and adopt a placenta ultrasound image dataset as the transferring dataset. After preprocessing, the features extracted from SPRNet are used to classify input images into corresponding categories by Softmax layer. The experimental results indicate that, with the transfer learning method we proposed, our network can utilize the potential relationships between two different datasets to improve classification performance, show higher generalization ability, and outperform other conventional networks.

2 Methodology

The overall architecture of the proposed SPRNet is shown in Fig. 2. The principles of the method used in this network are demonstrated as follow.

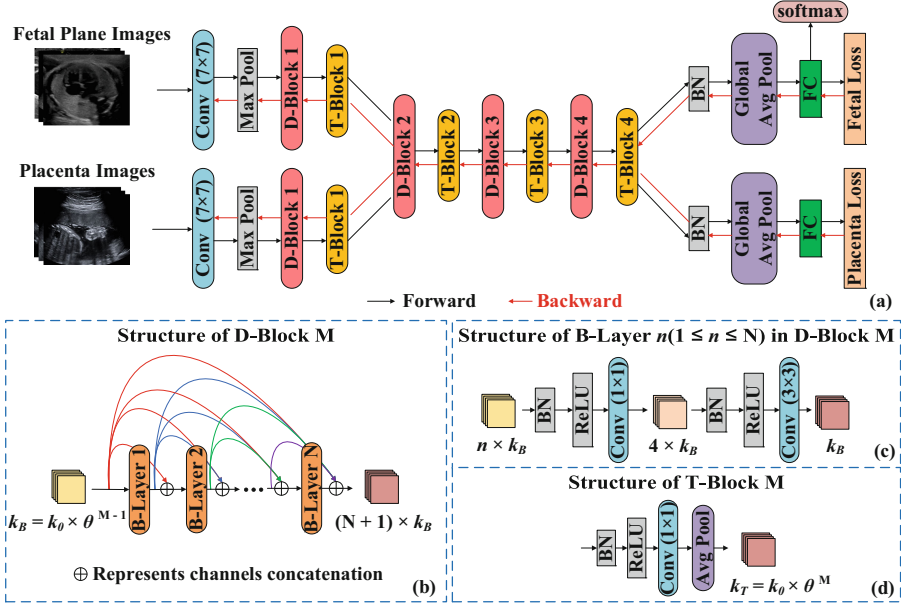


Fig. 2. Overall structure of the proposed method. (a) the architecture of SPRNet; (b), (c) and (d) represent the basic modules in our network. k_0 is the number of initial feature channels, k_B and k_T denote the number of feature channels, M represents the index of D-Block and T-Block, N is the amount of B-Layer in D-Block M , n ($1 \leq n \leq N$) refers to the index of B-Layer in D-Block M , and θ ($0 < \theta \leq 1$) denotes the channel decay coefficient.

2.1 Data Processing

The datasets used to train SPRNet is composed of a task dataset and a transferring dataset. The task dataset is constituted by fetal plane images in ultrasound and contains seven categories: 4 channel chamber (4CH), abdomen, brain, axial face (AF), coronal face (CF), sagittal face (SF) and others. Others are a collection of all non-standard planes, and the rest categories are collections of corresponding structures' standard planes. The transferring dataset is a collection of placenta ultrasound images, and it is divided into four grades: grade 0 to 3, according to Grannum standards [9].

The sizes of task and transferring datasets are summarized in Tables 1 and 2, respectively. Due to the limited number of cases and the difficulties of data annotation, the size of transferring dataset is significantly smaller than task dataset. This problem may impose an adverse effect on the performance of SPRNet. Therefore, we extend the training set of transferring dataset by cropping original images (700×500 pixels) into

448×448 pixels from top left corner to bottom right corner with different strides for different grades. The horizontal cropping strides for grade 0 to grade 3 are 126 pixels, 84 pixels, 63 pixels, and 21 pixels, respectively, and the longitudinal cropping stride for all images is 26 pixels. Eventually, we get more than 1,000 images in each category of transferring dataset and solve the problem of unbalanced data.

Table 1. The size of every category of training and testing in task dataset.

Fetal	4CH	Abdomen	Brain	AF	CF	SF	Others
Training set	1927	1349	1472	966	1542	1376	9208
Testing set	481	337	367	241	385	343	2301

Table 2. The size of every grade of original and extended training and testing in transferring dataset.

Placenta	Grade 0	Grade 1	Grade 2	Grade 3
Original training set	149	108	68	28
Extended training set	1341	1296	1020	1092
Testing set	37	26	16	7

2.2 Basic Modules

We adopt the B-Layer (Bottleneck Layer), D-Block (Dense Block) and T-Block (Transition Block) in DenseNet as the basic modules of our network.

D-Block is an intensive connection mechanism. It connects each layer to the previous layers in the same block and reuse the features extracted from previous layers by concatenation. The advantage of this connection strategy is that, it protects the information while reusing them and allows gradient to propagate from deep layers to shallow layers more easily. With this structure, D-Block performs better than the residual block in ResNet with less parameters and alleviates the problems of gradient vanishing and model degradation. B-Layer is the basic unit of D-Block, which is used for extracting information. T-Block, which is an interlayer between two D-Blocks, is mainly used for reducing the number of parameters.

In our SPRNet, there are 4 D-Block and 4 T-Block, and we set $k_0 = 32$ and $\theta = 0.5$. From D-Block 1 to 4, N is 6, 16, 24 and 24, respectively.

2.3 Data-Based Partial Transfer Learning

Transfer learning is used to utilize the knowledge learned from transferring dataset to improve the performance of CNN in task dataset and it was proved to be effective to augment the generalization ability of CNN by Yosinski et al. [10]. The conventional methods of transfer learning are based on transferring the weights of a pre-trained model to a new model as initial weights and then fine-tuning the new model. Although this method can boost the generation ability of network, it ignores the relationship between task dataset and transferring dataset during the fine-tuning process and still

suffers from the problem of overfitting when dataset size is limited. Wang et al. proposed a novel transfer learning method called data-based transfer learning [8]. In this method, networks for different datasets are integrated into a general network by weight-sharing strategy, but they still possess their own fully connected layer and loss function to finish their own task. With this structure, the general network is able to extract and learn the potential relationship between task dataset and transferring dataset, prevent network from overfitting to any one dataset and perform better generalization ability than conventional transfer learning methods.

Conventional transfer learning method usually adopt natural images like ImageNet [13] as transferring data, but for transfer learning in medical area, there is huge difference between medical images and natural images, such as morphological difference and acquisition method, which may bring some adverse effects. So, in order to avoid these disadvantages, we also try to adopt the placenta ultrasound images as transferring data. Although there is still huge morphological difference between placenta ultrasound images and fetal plane ultrasound images, we believe that the different medical images which are collected by the same method have some common features which can be used for transfer learning.

When we apply data-based transfer learning into our network, we discover that the performance of our network declines. The reason is that, unlike the datasets used in Wang's et al. work [8], which are selected and closely related, there are huge morphological differences between our task dataset and transferring dataset, and these differences make it difficult for shallow layers, which are prone to extract morphological information like textures and corner point, to extract common features from task dataset and transferring dataset. To settle this problem, we do not apply weight-sharing strategy into shallow layers and only use deep layers to extract the common features hidden in the task and transferring datasets. Therefore, our network can avoid the performance decline while the task dataset and transferring dataset are not closely related to each other. We call this method as data-based partial transfer learning.

3 Experiments and Results

3.1 Experiment Design

We design a control experiment which uses three different networks (DenseNet-145, DenseNet-145-global-transfer and SPRNet) to finish two tasks (fetal standard plane recognition and placenta maturity grading), respectively, to demonstrate the improvement of SPRNet. DenseNet-145 is a densely connected convolutional networks with 145 convolutional layers. DenseNet-145-global-transfer is a network where weight-sharing strategy is applied on every convolutional layer. SPRNet is our proposed and it also includes 145 convolutional layers.

We randomly divide both datasets into 80% for training set and 20% as testing set, and data processing is applied to the training set.

The experiments are implemented using Python via Tensorflow and runs at a 32 GBs-RAM computer with a GeForce GTX 1080 Ti GPU. Accuracy (ACC),

sensitivity (SEN), specificity (SPE), and F1-Score (F1) are adopted to evaluate the performance of the networks.

3.2 Results

As shown in Table 3, benefitting from data-based partial transfer learning, SPRNet outperforms other methods in fetal standard plane recognition as well as placenta maturity grading. However, for the DenseNet-145-global-transfer, there is a performance degradation in both tasks, if we regard DenseNet-145 as the benchmark. This degradation of performance mainly caused by the huge morphological differences between task dataset and transferring dataset. As shown in Fig. 3(b) and (c), the features extracted from the separated shallow layers in SPRNet mainly contain morphological information, such as corner point and texture, and there are huge differences between the features extracted from different datasets, which will result in an intense antagonism in shallow weight-sharing convolutional layers. For the data-based global transfer learning, this intense antagonism is too strong for it to find a proper point to learn common information from both datasets and yields performance degradation. On the contrary, data-based partial transfer learning, which cancels the weight-sharing on shallow layers, can effectively weaken this strong antagonism, ensuring the network will not be impaired. Furthermore, the improvement of SPRNet in placenta maturity grading task suggest that data-based partial transfer learning can effectively prevent the overfitting problem which is caused by limited data and improve the performance of network by extracting common features from task dataset and transferring dataset.

Table 3. The performance of the proposed SPRNet against other networks.

Task	Network	ACC %	SEN %	SPE %	F1 %
Fetal standard plane recognition	DenseNet-145	98.86	96.04	99.34	95.24
	DenseNet-145-global-transfer	98.03	93.11	98.85	92.25
	SPRNet	99.00	96.51	99.41	95.58
Placenta maturity grading	DenseNet-145	92.44	84.88	94.96	85.13
	DenseNet-145-global-transfer	91.27	82.55	94.18	83.02
	SPRNet	94.76	89.53	96.51	90.85

Table 4 shows the SPRNet’s recognition results of different fetal planes, indicating that SPRNet achieves the best results in this task. The confusion matrix shown in Fig. 4 reveals the specific recognition result of SPRNet and prove the effectiveness of the proposed method.

To further explain the effectiveness of SPRNet, we implement feature visualization by t-SNE. Specifically, we reshape all the input test images and the feature extracted from SPRNet to a matrix, respectively, in which every row represents an images or features, and then demonstrate the distribution of these two matrixes by the t-SNE

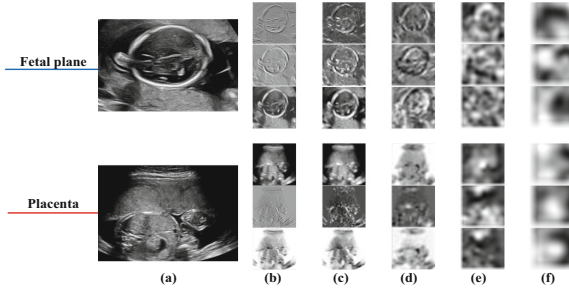


Fig. 3. Feature maps extracted from SPRNet. (a) is the original input image. (b) is the output of the first convolutional layer. (c)–(f) refer to the feature extracted from D-Block 1 to 4, respectively.

Table 4. Recognition results

Classes	ACC %	SEN %	SPE %	F1 %
4CH	99.51	97.54	99.75	97.74
Abdomen	99.00	95.60	99.27	93.54
Brain	99.17	98.92	99.20	95.21
AF	99.37	99.58	99.36	94.50
CF	99.42	93.81	99.97	96.67
SF	99.04	96.56	99.22	93.87
Others	97.48	96.16	98.89	97.53

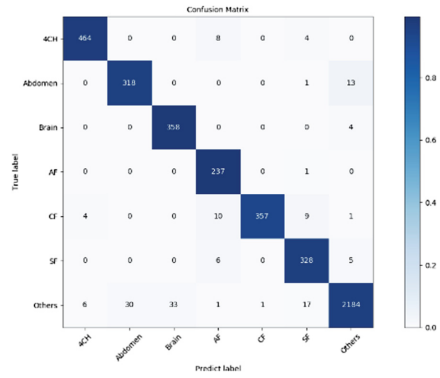


Fig. 4. Confusion matrix of SPRNet.

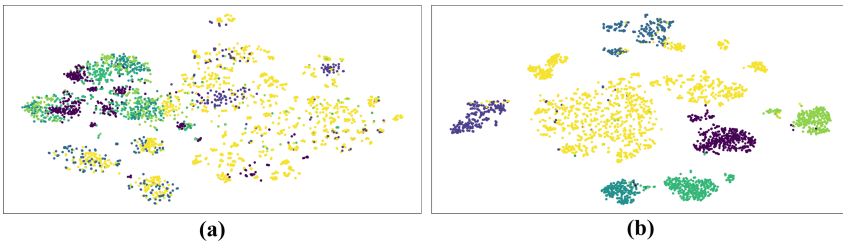


Fig. 5. Feature visualization via t-SNE. (a) the distribution of the input test images; (b) the distribution of the features extracted from SPRNet.

function provided in the sklearn. Different colors refer to different categories of fetal ultrasound planes. As shown in Fig. 5, the distribution of the input test images is unordered, showing that the distinction of standard fetal planes and non-standard planes is unobtrusive. On the contrary, after feature extraction of SPRNet, corresponding categories are grouped together, and the distribution of features becomes separable. This result further proves the effectiveness of the proposed network.

4 Conclusion

In this paper, we propose an effective fetal standard plane recognition network, which adopts D-Block and T-Block as the basic module and introduces data-based partial transfer learning. The experimental results demonstrate that SPRNet is accurate and effective, and the data-based partial transfer learning brings a considerable improvement to our network. In the future, we will expand our dataset to realize standard plane recognition on more fetal structures and try to apply automatic parameter measurement and structure localization to our method.

References

1. Li, J., Wang, Y., et al.: Automatic fetal head circumference measurement in ultrasound using random forest and fast ellipse fitting. *IEEE J. Biomed. Health Inf.* **22**, 215–223 (2018)
2. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014)
3. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778 (2016)
4. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141 (2018)
5. Cai, Y., Sharma, H., Chatelain, P., Noble, J.: SonoEyeNet: standardized fetal ultrasound plane detection informed by eye tracking. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 1475–1478 (2018)
6. Baumgartner, C.F., Kamnitsas, K., Matthew, J., Fletcher, T.P., Smith, S., Koch, L.M.: SonoNet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound. *IEEE Trans. Med. Imaging* **36**, 2204–2215 (2017)
7. Huang, G., Liu, Z., Van Der Maaten, L., et al.: Densely connected convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708 (2017)
8. Wang, L., Wang, Z., Qiao, Y., et al.: Transferring deep object and scene representations for event recognition in still images. *Int. J. Comput. Vis.* **126**(2–4), 390–409 (2018)
9. Grannum, P.A., Berkowitz, R.L., Hobbins, J.C.: The ultrasonic changes in the maturing placenta and their relation to fetal pulmonic maturity. *Am. J. Obstet. Gynecol.* **133**(8), 915 (1979)
10. Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks?. In: *Advances in Neural Information Processing Systems*, pp. 3320–3328 (2014)

11. Kong, P., Ni, D., Chen, S., Li, S., Wang, T., Lei, B.: Automatic and efficient standard plane recognition in fetal ultrasound images via multi-scale dense networks. In: Melbourne, A., et al. (eds.) PIPPI/DATRA -2018. LNCS, vol. 11076, pp. 160–168. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00807-9_16
12. Cai, Y., Sharma, H., Chatelain, P., Noble, J.A.: Multi-task SonoEyeNet: detection of fetal standardized planes assisted by generated sonographer attention maps. In: Frangi, Alejandro F., Schnabel, Julia A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11070, pp. 871–879. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00928-1_98
13. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255 (2009)