

Received April 5, 2020, accepted April 20, 2020, date of publication May 6, 2020, date of current version May 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2991845

Standard Plane Identification in Fetal Brain Ultrasound Scans Using a Differential Convolutional Neural Network

RUOWEI QU¹, GUIZHI XU¹, (Member, IEEE), CHUNXIA DING²,
WENYAN JIA³, AND MINGUI SUN³, (Senior Member, IEEE)

¹State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300130, China

²Ultrasound Department, Zhangjiakou Maternal and Child Health Hospital, Zhangjiakou 075000, China

³Laboratory for Computational Neuroscience, University of Pittsburgh, Pittsburgh, PA 15213, USA

Corresponding author: Guizhi Xu (gzxu@hebut.edu.cn)

This work was supported in part by the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, and in part by the National Natural Science Foundation of China under Grant 51677053, Grant 51737003, and Grant 51807048.

ABSTRACT Ultrasound scanning has become a highly recommended examination in prenatal diagnosis in many countries. The accurate identification of fetal brain ultrasound scans is crucial to accurate head measurement and brain lesion detection, such as the measurement of the biparietal diameter and the detection of hydrocephalus. In recent years, deep learning has made great progress in the field of image processing. However, there are two difficulties in the identification of fetal brain ultrasound standard planes (FBSPs). First, since the fetal brain tissue is not mature, the fetal brain tissue features are not easy to be detected. Second, because of the expensive collection costs, the amount of labeled image data is limited, which can cause over-fitting and decrease the identification precision. In this study, we proposed a differential convolutional neural network (differential-CNN) to automatically identify six fetal brain standard planes (FBSPs) from the non-standard planes. In this differential-CNN framework, the additional differential feature maps were derived from the feature maps in the original CNN using differential operators. The derivation process did not increase the number of convolution layers and parameters. Moreover, the differential convolution maps have the large advantage of analyzing the directional pattern of pixels and their neighborhoods using additional variation calculations. Therefore, the differential convolution maps would result in good identification performance and cost no extra computational burden. To test the performance of these algorithms, we constructed a dataset consisting of 30,000 2D ultrasound images from 155 fetal subjects ranging from 16 to 34 weeks. The experimental results showed that this method achieved an accuracy of 92.93%. Our work shows that the differential-CNN can be used to facilitate the implementation of the automated identification of FBSPs.

INDEX TERMS Convolutional neural network, ultrasound scan image, medical image processing, convolution techniques, differential operator.

I. INTRODUCTION

Ultrasound imaging technology has been applied to the prenatal observation and measurement of fetuses and diagnosis of fetal diseases for approximately 30 years [1], [2]. It is the most common and simple imaging method to understand the main anatomical structures of embryos and fetuses [3]–[6]. In the initial application stage, the prenatal ultrasound examination is limited to the purposed determination of pregnancy,

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Ayoub Khan ¹.

the survival of the fetus, the adequacy of the amniotic fluid, the condition of the placenta and so on. Currently, ultrasound imaging technology has become an indispensable imaging diagnostic tool for obstetrics-based medical departments [7]. Ultrasound images can not only observe and understand the fetal morphology and structure but also observe the fetal activity and behavior in the mother's uterus and the dynamic changes of the fetal blood flow in real time [8].

In prenatal examination, the brain development of the fetus is the most important. First, the measurement of the fetal brain on the standard planes can directly reflect the development

of the fetus [9] and fetal age [10]. For example, the biparietal diameter (BPD) is the length of the widest part between the left and right sides of the fetal head [11]. After five months of pregnancy, the BPD is basically relevant to the fetal age and fetal weight. If the BPD is larger than the transverse diameter of the pelvic outlet of the parturient, the parturient may have dystocia and other issues. Therefore, the BPD is an important index for doctors to make decisions for natural delivery or cesarean section. Second, in the fetal brain development process, doctors can observe brain abnormalities using standard planes, such as ventricular expansion, intracranial hematomas, hydrocephalus and subependymal cysts, which are usually the manifestations of intracranial hemorrhage, ischemic and hypoxic brain diseases. If doctors can observe the changes of these indicators during pregnancy, it will be far-reaching significant for determining the death of an immature fetus and the congenital malformation of a fetus, selecting the appropriate diagnosis and treatment plan in time, and even terminating a pregnancy.

In the diagnosis of fetal brain development, the accurate recognition of fetal brain ultrasound standard planes is the basis of developmental diagnosis and data measurement [12]. Taking the measurement of the BPD as an example, the standard plane for the BPD measurement is the horizontal transverse section of the thalamus, which passes through the anterior horn of the lateral ventricle, the posterior horn of the lateral ventricle, the septum pellucidum and the thalamus. When ultrasound waves pass through the complex anatomy of a fetus, doctors observe the echoes received by the ultrasound receivers. It is a very complex task to observe and guide the ultrasonic probe to the correct standard planes, which techniques need to be trained by years. Because of the complexity of this work, the acquisition of ultrasonic standard planes has low reproducibility and large operator differences. Even with a given standard plane, it is a very challenging task for clinicians to identify the relevant organizational structures, especially for inexperienced operators and ultrasound diagnosticians. Therefore, research on the automatic recognition of the fetal brain ultrasonic standard plane has very important and far-reaching significance.

There are several limitations in the diagnosis based on ultrasound imaging [5]. In the ultrasound scan, the diagnostic precision can be affected by the inappropriate choice of the standard scan section, which is dependent on the clinical experience of the doctor [13]. Moreover, the same captured scan image can be diagnosed as a different disease due to different professional skills related to the inspection criteria and quality control standards [14].

An ultrasound image is formed by echoes from human tissues [4]. There may be artifacts as a result of the image processing. Image artifacts originate from many subtle factors, such as the shaking of machines and equipment. When different instruments are used, the performance of the equipment itself and the adjustments of doctors will also affect the identification and acquisition of standard sections. These factors will not only cause minor errors in the diagnosis

and measurement but also reduce the efficiency of doctor's assessments [15].

In recent years, the automatic identification of fetal brain standard planes has been extensively discussed in the literature. Yeh *et al.* [16] used a gray level co-occurrence matrix and wavelet decomposition to extract features, and then used a support vector machine (SVM) to classify features. Lei *et al.* [17] explored the automatic recognition of fetal facial standard planes (FFSPs) using the Fisher Vector and SVM on a small sample dataset. The limitations of the above traditional methods are mainly reflected in the following three aspects: (1) manual feature extraction settings are usually based on humans' subjective experience; (2) due to artificial selection, the number of extracted features and the types of features are very limited; and (3) the above feature extraction method cannot be optimized in real time when data sets change.

However, with the development of deep learning algorithms, a new development stage of computer-aided diagnosis is emerging. Cao *et al.* evaluated the performance of several existing state-of-the-art deep learning methods for breast tumor detection [18]. Yu *et al.* continued their research by applying a deep convolutional neural network on a relatively large dataset and distinguished four statuses of the FFSPs [19]. Baumgartner *et al.* [20] proposed a novel detection and localization method based on a convolutional neural network (CNN) that can automatically detect 13 fetal standard views for freehand 2D ultrasound data and provide a localization of the fetal structures via a bounding box. Feng Dagan *et al.* [21] used two fusion CNNs and saliency maps. Lei *et al.* [22]–[25] explored the automatic recognition of fetal facial standard planes and breast segmentation using a deep learning method. These two CNNs learned from a network trained using natural images and achieved an automatic detection and measurement using VGG-Net [26]. Ricardo *et al.* improved the topological coherence strategy in an auto-encoder to augment the number of blood vessel images with noise [27].

Inspired by [28], we proposed a differential-CNN-based fetal brain standard planes (FBSPs) identification system to identify six classes of FBSPs and one class of non-standard planes. The feature maps in differential CNNs were generated with one single convolution feature map by applying predefined hyper-parameters and a differential operator. In this way, the differential-CNN used more differential feature maps to extract more details in the image without increasing the numbers of convolution layers and parameters. In addition, one more fixed filter is added to calculate the difference between the pixel and the pixel on the corresponding position of the adjacent layer. The relevant differential feature map and the relevant back propagation processing on the original algorithm to improve the FBSPs identification performance. Therefore, the proposed differential-CNN reduced the complexity of the convolutional network structures and meets the requirements of portal computing equipment.

The rest of the paper is organized as follows. Section II introduces the methodology of the proposed differential-CNN algorithm. The data collection and augmentation are given in Section III. The experiment results and analysis are given in Section IV. The conclusion and future works are given in the last section.

II. METHODOLOGY

A. CONVOLUTIONAL NEURAL NETWORK

Compared with traditional neural networks, the CNN provides a fast and convenient algorithm and performs well in target detection and classification [29]–[32]. In bionics, the CNN can successfully simulate mammalian visual cortex nerve operations [33]. Inspired by visual perception, the features are extracted using localized convolution. Then, the hidden layer of the CNN carries the spatial correlation information. Therefore, when applied in vision-related imaging problems, the CNN has significantly improved classification accuracy for many standard image databases, such as MNIST [34] and CIFAR 10 [35].

The basic structure of the CNN includes several convolutional layers, pooling layers and fully connected layers [36]–[38]. The task of the **convolution layer** is to detect the local connection features of the previous layer. The formula for computing a single output matrix is defined as follows:

$$A_j = f \left(\sum_{i=1}^N I_i * K_{i,j} + B_j \right) \quad (1)$$

where I is an input vector, and K is the corresponding convolution kernel with the size of $n \times n$ ($n < \text{input size}$). Then, all the convoluted matrices are added up and a bias value B_j is added to each element of the resulting matrix. f is a non-linear activation function working on each element of the previous matrix to produce the output matrix A . In this paper, the rectified linear function $f(x) = \max(0, x)$ (ReLU) is chosen as the **activation function** to improve both the learning speed and classification performance of the CNN [39]. The **pooling layer** that integrates the semantically similar features is used to reduce the resolution of the feature maps [40]. The **back propagation** can update the training weights and train the weights of all feature maps [41].

In the design of CNN models, different architectures can affect the training and testing performances [42]. A deep network can achieve better performance but it needs a longer calculation time; conversely, a shallow network can achieve high calculation efficiency but cause the underfitting problem. As a result, an appropriate network system can improve the performance of the identification system. Our differential-CNN model contains five convolutional layers and five average pooling layers between the convolutional layers. The architecture of the original CNN is shown in Table 1.

In Table 1, C represents the convolutional layers, P represents the pooling layers (the average pooling layers in this structure), and F represents the fully connected layers. The original feature map in the CNN is randomly generated.

TABLE 1. Architecture of the original CNN model.

Layer	Number of feature maps	Kernel size	Stride	Size of the feature maps
Input		11		1020×1020
C1	12	2	2	500×500×12
P1	1	5		250×250×12
C2	5	2	1	250×250×60
P2	1	6		125×125×60
C3	5	3	1	120×120×300
P3	1	3		40×40×300
C4	2	2	1	40×40×600
P4	1	3		20×20×600
C5	1	3	1	18×18×600
P5	1			6×6×600
F1				21600
F2				2520
F3				200

Because the model's task is to identify six classes of FBSPs images, the final fully connected layer has seven channels. The previous research [43] shows that a prominent reduction of the fully connected layer size in a CNN will not reduce the network performance. As a result, the proposed fully connected layer was greatly shortened to improve the calculation efficiency while retaining the network performance.

The final goal of the experiment is to compare the probability of FBSPs classification on the last output node. The most probable class is considered to be the final prediction class. The output is the final predictive classification.

B. DIFFERENTIAL CONVOLUTIONAL FEATURE MAP

Convolution is the key of a deep learning structure, which is realized by sliding several filters over the input images. It extracts the features from the input image by simulating human vision. Therefore, the more feature maps included in the feature extraction layers in the structure, the more features the classifier obtains.

In traditional convolution neural networks, feature maps are generated via random initialization or transferred knowledge. Compared with traditional CNNs, the feature maps in differential CNNs are generated using traditional convolution feature maps by applying pre-defined hyperparameters and a differential operator [25]. Differential convolution maps are used to analyze the directional patterns of pixels and their neighborhoods through an additional variation calculation. It is worth mentioning that, in mathematical differentiation, the sequence change is considered by calculating the difference between pixel activations. The predefined feature maps are shown in Fig. 1. Each feature map is used to compute the difference in one direction. Therefore, the additional feature maps containing differences along different directions are obtained. Different from [28], we add a fixed filter to the original algorithm to extract more features in the FBSPs identification task.

Because we add a fixed filter here, the feature maps are added correspondingly. Let the initial feature map generated

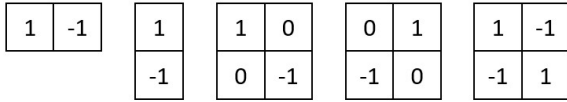


FIGURE 1. Predefined filters.

from the traditional convolution neural networks be g_1 and the resulting five feature maps using the differential operator be $g_2, g_3, g_4, g_5,$ and g_6 . The neurons in these maps are calculated using (2)-(6):

$$g_{2,i,j} = g_{1,i,j} - g_{1,i+1,j} \tag{2}$$

$$g_{3,i,j} = g_{1,i,j} - g_{1,i,j+1} \tag{3}$$

$$g_{4,i,j} = g_{1,i,j} - g_{1,i+1,j+1} \tag{4}$$

$$g_{5,i,j} = g_{1,i+1,j} - g_{1,i,j+1} \tag{5}$$

$$g_{6,i,j} = g_{1,i+1,j+1} - g_{1,i,j+1} \tag{6}$$

where i and j are the coordinates of the neurons in the convolutional feature maps. We assume that the size of g_1 is $M \times N$, and the sizes of $g_2, g_3, g_4, g_5,$ and g_6 are $(M - 1) \times N, M \times (N - 1), (M - 1) \times (N - 1), (M - 1) \times (N - 1)$ and $(M - 1) \times (N - 1)$, respectively. The calculation process of these feature maps is shown in Fig. 2.

After the first feature map is generated by the traditional convolution feature map, the differential convolution feature maps are calculated from the first feature map using differential operators. The differential convolution feature maps are used to detect the basic features of an image, such as edges and corners.

From the above derivation process, we can see that the differential CNN uses more differential feature maps to extract more details in the image without increasing the convolution layers. Therefore, the proposed differential CNN reduces the complexity of convolution network structures, thus reducing the computing requirements.

C. BACK PROPAGATION

The BP algorithm would be improved while the feature maps changed. If the network cannot identify the expected output value in the output layer, the sum of the error between the expected value and output value is taken as the objective function that is transmitted in the reverse direction. Then, the partial derivative of the objective function is calculated layer by layer. This partial derivative is the learning gradient. The CNN modifies the weights of the feature maps based on the gradient and learning rate. When the error decreases below the expected value, the training does stop.

Let the error transmitted to the first map be h_1 ; the errors transmitted to the generated extra maps be $h_2, h_3, h_4,$ and h_5 ; and the error matrix be E . Functions (7)-(11) show the error calculations for the relevant filter.

$$E_{i,j} = h_{1,i,j} - h_{2,i,j-1} + h_{2,i,j} - h_{3,i-1,j} + h_{3,i,j} - h_{4,i-1,j-1} + h_{4,i,j} - h_{5,i-1,j} + h_{5,i,j-1} - h_{6,i,j-1} + h_{6,i-1,j-1} \tag{7}$$

where $1 < i < M$ and $1 < j < N$. The $E_{i,j}$ in formula (7) represents the error for the neurons neither at the edges nor in the corners. It receives error feedback from all neighboring neurons.

$$E_{i,j} = \begin{cases} h_{1,i,j} + h_{2,i,j} + h_{3,i,j} + h_{4,i,j}, & i = 1, j = 1 \\ h_{1,i,j} - h_{2,i,j-1} + h_{3,i,j} + h_{5,i,j-1}, & i = 1, j = N \\ h_{1,i,j} + h_{2,i,j} - h_{3,i-1,j} - h_{5,i-1,j}, & i = M, j = 1 \\ h_{1,i,j} - h_{2,i,j-1} - h_{3,i-1,j} - h_{6,i-1,j-1}, & i = M, j = N \end{cases} \tag{8}$$

The $E_{i,j}$ in formula (8) represents the error of the neurons at the corners and edges. It receives error feedback from 3 neighboring neurons.

$$E_{i,j} = \begin{cases} h_{1,i,j} - h_{2,i,j-1} + h_{3,i,j} + h_{4,i,j} + h_{5,i,j-1}, & i = 1, 1 < j < N \\ h_{1,i,j} - h_{2,i,j-1} + h_{2,i,j} - h_{3,i-1,j} - h_{4,i-1,j-1} - h_{6,i-1,j}, & i = M, 1 < j < N \\ h_{1,i,j} + h_{2,i,j} + h_{3,i,j} - h_{1,i-1,j} + h_{4,i,j} - h_{5,i-j}, & 1 < i < M, j = 1 \\ h_{1,i,j} - h_{2,i,j-1} + h_{3,i,j} - h_{4,i-1,j} + h_{6,i,j-1}, & 1 < i < M, j = N \end{cases} \tag{9}$$

The $E_{i,j}$ in (9) represents the error of the neurons propagated to the edge neurons. It receives error feedback from 5 neighboring neurons.

III. DATA COLLECTION AND AUGMENTATION

A. DATA COLLECTION

After the approval from Institutional Review Board (IRB), we recorded a set of images corresponding to a complete pregnancy. With the permissions of the pregnant women, their ultrasound scan videos during the whole pregnancy were recorded and kept. All ultrasonic scan images were converted into gray images so that there was only one input channel for the CNN. Since the successive frames in the video change little, we down-sampled the image sequence. Even after this down-sampling, some images were still quite similar. Due to this persistent similarity among the images, doctors helped us to further manually delete similar images, blurry images and excessively dark images to keep the number of images from the same event to less than 10. In this way, we collected 19,142 fetal brain standard planes in six classes and one class of non-standard planes from 155 subjects. They were collected by a Hitachi ARIETTA 70 B-mode ultrasonic apparatus and the corresponding probe with frequency of 4-6 MHz.

In addition, all private information has been removed from the image. Furthermore, all the ground and detailed information (such as the Sylvian fissure (SF), lateral ventricle (LV) and thalamus (T)) were labeled in the images.

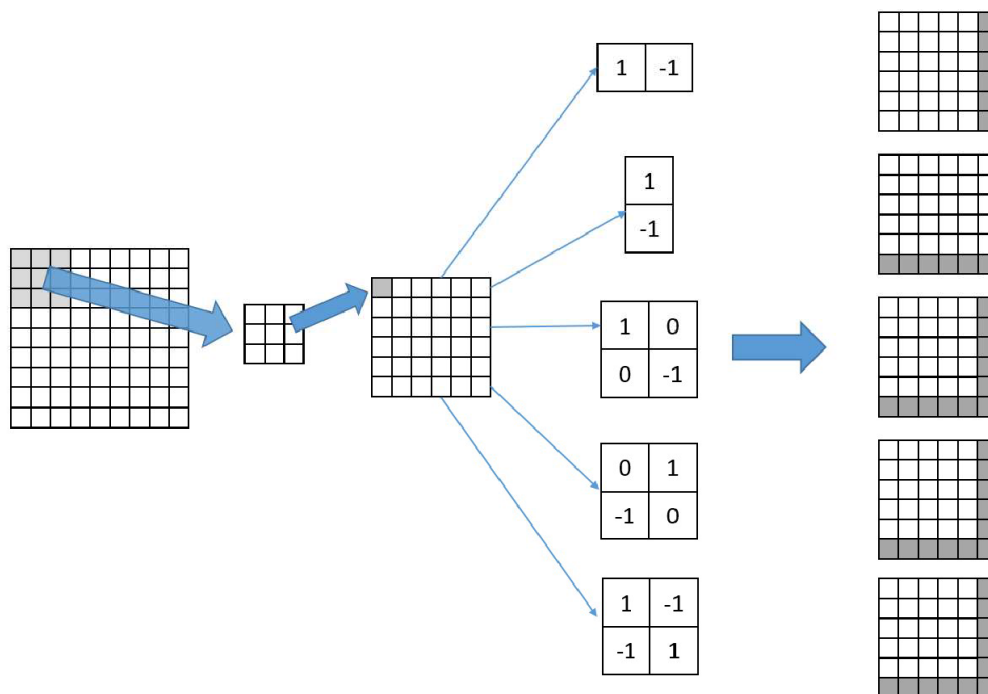


FIGURE 2. The calculation of the differential feature maps.

The six fetal brain standard planes, which are the horizontal transverse plane of thalamus, the horizontal transverse plane of lateral ventricle, the transverse plane of the cerebellum, the midsagittal plane, the paracentral sagittal plane, and the coronal plane of the anterior horn of the lateral ventricle, are shown in Fig. 3.

B. DATA AUGMENTATION

Although the CNN has a strong advantage in the representation of the learned features, the deep structure and the supervised learning may lead to overfitting when the amount of training data is limited, such as in many medical situations. When the dataset is small, an excessive number of parameters in a CNN would result in overfitting. The common features among different images may be ignored. As a result, the CNN's generalization capability will be weakened.

To avoid this problem, several data augmentation methods are proposed to prevent overfitting in this paper, which are shown in Fig. 4. They include the following: (A) Rotation: rotate the ultrasound image by a certain degree; (B) Reflection: change the direction of the ultrasound image's content; (C) Gaussian white noise: add white noise depending on a Gaussian distribution; (D) Flip: flip the ultrasound scans vertically or horizontally; (E) Zoom: zoom in or out on an ultrasound scan; (F) Cubic spline interpolation: reconstruct a new image using the cubic spline interpolation method; and (G) Bilinear interpolation: reconstruct a new image using bilinear interpolation.

After conducting all the pre-processing and data augmentation for the original dataset, a total of 30,000 1020×1020

fetal brain standard planes and non-standard planes formed FBSPs. In this FBSPs dataset, there are 4,000 images for each class of standard planes and 6,000 images for the non-standard planes. Here, we choose 5-fold validation as the training strategy. It means that each fold contains 800 images for each class of standard planes and 1200 images for the non-standard planes.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In this experiment, the differential-CNN uses the same parameters as the original CNN structure showed in Table 1. Additionally, they have the same numbers and positions for the convolutional layers and pooling layers. There are 5 convolutional layers in this differential-CNN. The number of feature maps in these convolutional layers is 48, 20, 20, 8, and 4. We set the size of the feature maps in first two convolutional layers as 2×2 and the size of the feature maps in the other convolutional layers as 3×3 . All the convolutional layers are followed by an averaged pooling layer. All the pooling filters are followed by the pooling layers in the CNN showed in Table 1.

In this section, we will carry out the numerical simulation using an HP Z640 Workstation with an Intel Xeon E5-2620v4 2.1 2133 8C CPU and an NVIDIA Quadro M4000 8GB GeForce GPU. The framework used to establish the CNN architecture is TensorFlow. Our system also used the mixed programming technology of MATLAB.

We conducted testing using our collected FBSPs image dataset to evaluate and analyze the various factors in our proposed CNN architecture. To evaluate the efficiency of our

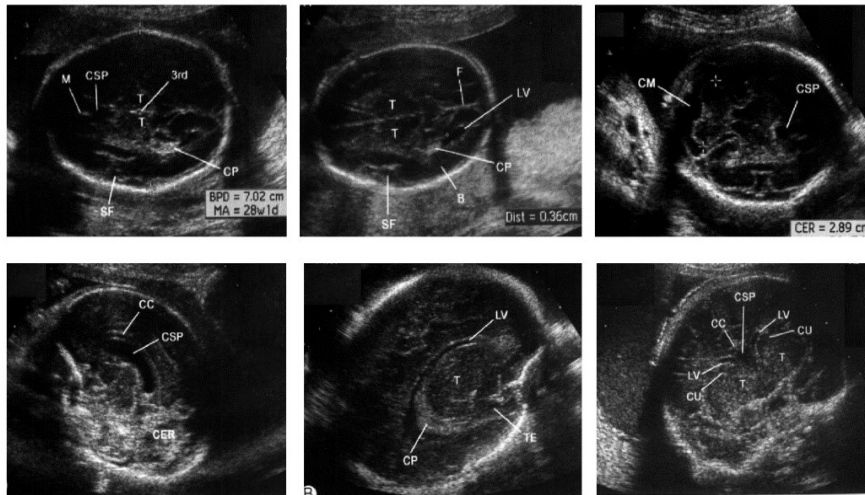


FIGURE 3. Samples of the six fetal brain standard planes.

proposed algorithm, four classical machine learning models were explored and several CNN systems with different network depths were trained for comparison purposes. To compare the differences between the various algorithms, we computed the follow statistical indexes: Accuracy (A), Precision (P), Recall (R) and F1-measure (F1) [44].

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$P = \frac{TP}{TP + FP} \quad (11)$$

$$R = \frac{TP}{TP + FN} \quad (12)$$

$$F_1 = \frac{2RP}{R + P} \quad (13)$$

where TP, TN, FP and FN denote the number of true positives, true negatives, false positives and false negatives, respectively. To evaluate the efficiency of the proposed algorithm, four classical machine learning models were explored and several CNN systems with different network depths were trained for comparison purposes. **K-means clustering** is a vector quantization algorithm. In machine learning, it divides n observations into k classes, each of which belongs to the classes closest to the mean [45], [46]. Compared to k -means clustering, the **support-vector machine (SVM)** is a supervised learning model. It maps the points in space to divide the samples of individual categories with as large of a gap as possible [47]. The **RCM** is a new detection method that is added to the detection program to improve the performance [48]. The identification performance of the **original CNN** listed in Table 1 is also compared in Table 2. Recently, **SonoNet** performed well in the classification of medical images [8]. In this paper, we discussed the performance of two SonoNet structures with different depths. In recent years, transfer learning has shown its advantage at preventing overfitting problems. In this paper, we compared two other popular structures, AlexNet and VGG 16, with the

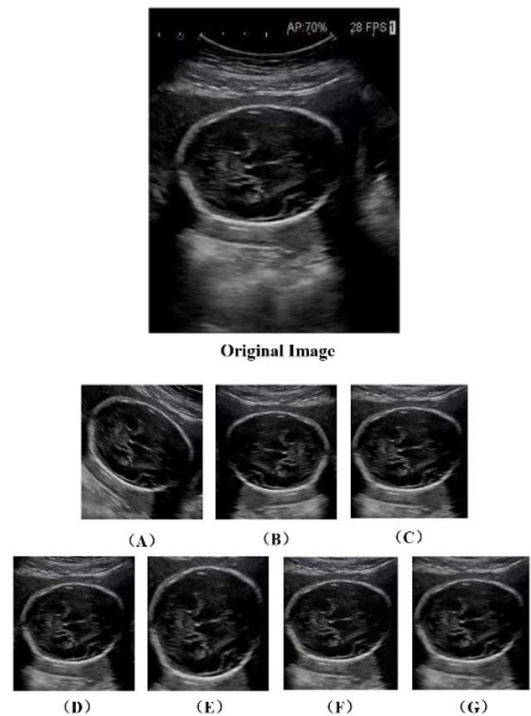


FIGURE 4. Samples showing the results of the seven transformation methods for the original image. (A-G) represent rotation, reflection, Gaussian white noise, flip, zoom, cubic spline interpolation and bilinear interpolation.

proposed differential-CNN. The detailed comparison results are shown in the Table 2. Moreover, to consider the hardware resource consumption of these networks for the identification task, the running times and the numbers of parameters of the deep networks are measured and listed.

It can be seen in Table 2 that the accuracy, precision, recall, and F1-score of the differential-CNN on the testing data achieved the best results of 93.11%, 92.62%, 92.39% and 93.53%, respectively, which significantly outperformed

TABLE 2. FBSPs identification results for various models.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC	Running time (min)	Number of parameters
(a) K-means clustering	74.62	77.30	75.64	80.56	0.792	35	/
(b) RCM	80.02	75.53	85.02	85.10	0.838	78	/
(c) SonoNet-32	85.23	86.48	84.98	85.85	0.877	26	6,556,338
(d) SonoNet-64	86.82	84.04	85.40	84.35	0.873	37	26,265,354
(e) CNN+SVM	84.37	84.11	85.96	83.55	0.863	92	16,157,435
(f) Original CNN	84.51	85.52	83.39	85.84	0.865	66	16,157,435
(g) Pre-trained AlexNet	87.58	88.35	89.93	88.17	0.879	25	54,542,338
(h) Pre-trained VGG16	90.91	89.52	91.39	90.27	0.917	32	119,554,050
(i) Differential-CNN	93.11	92.62	92.39	93.53	0.937	37	4,830,554

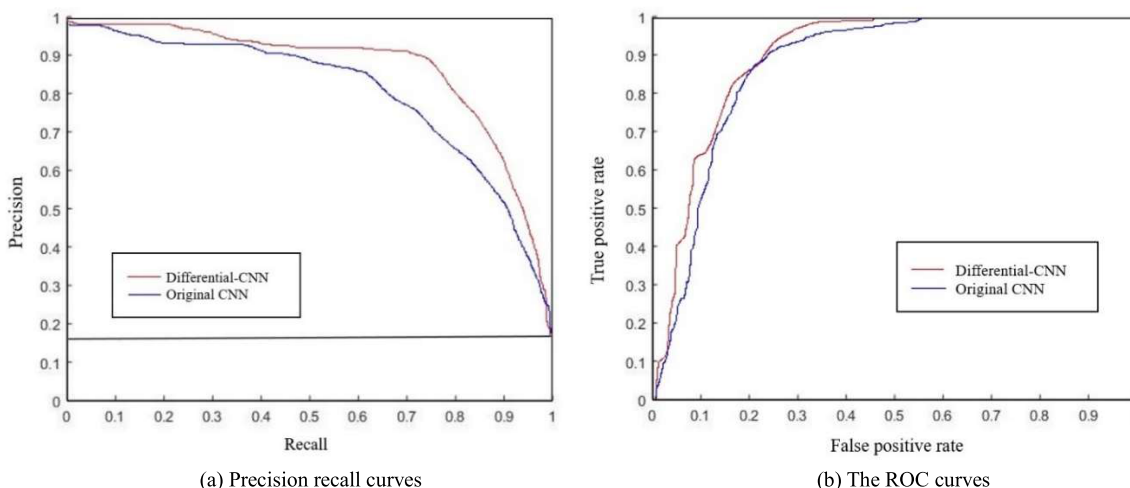


FIGURE 5. The precision and rates of different algorithms.

the other network architectures. The running time of our proposed differential-CNN is similar to those of the pre-trained neural networks, and our differential-CNN is significantly faster than the original CNN. In addition, the number of parameters of our differential-CNN is only 4,830,554, which is significantly smaller than those of the listed deep architectures. Compared with (f) in Table 2, the feature maps calculated from the original CNN using differential operators greatly reduce the number of parameters that need to be trained in the networks and the complexity of the models.

This shows that the differential-CNN has a distinct advantage over traditional CNNs in hardware resource consumption. In this way, our proposed differential-CNN can complete the FBSPs identification task with better performance consuming less memory than CNNs listed from models (c) to (h).

Because the original feature maps in the differential-CNN are convoluted by a pre-defined filter, the depth of a single convolution layer expands without increasing the number of convolution layers. Moreover, these differential feature maps record the differences in different directions, which improves the differential-CNN performance of detecting the basic structures of images such as edges and angles. Therefore, the identification accuracy is improved by applying the

differential-CNN. In addition, the differential feature maps are derived from the original feature map. The number of weights in the neural network is reduced. Thus, the calculation time of the differential-CNN is significantly shortened.

The precision-recall (PR) and receiver operating characteristic (ROC) curves shown in Fig. 5 are standard indexes used to evaluate the identification performance of learning models for a given fixed dataset. For the PR curve, when the shape of the ROC curve is more convex, it indicates that the algorithm represented by the ROC curve has a better recognition effect. The area under the curve (AUC) is defined as the ratio of the area under the ROC curve to the area of the whole square, which is positively related to identification performance. Since the ROC curves of different classification algorithms may intersect, the performances of the algorithms cannot be directly judged using the ROC curves. Therefore, the AUC value is often used as the index to judge the performances of algorithms. The area under the ROC curve (AUC) obtained by our proposed differential-CNN is 0.937.

For models (c) to (f), it can be seen that SonoNet-64 (d) and SonoNet-32 (c) performed similarly on the identification task. They performed quite well as CNN frameworks that do not transfer the initial parameters. Model (f) represents the algorithm that we introduced in Table 1. In model (e),

we replaced the last fully connected layer with an SVM classifier. As a result, the CNN can be used as a feature extractor.

V. CONCLUSION

Although deep CNNs have achieved remarkable success in medical image analysis in recent years, their relative low identification accuracy and expensive computational costs limit the application of CNNs in clinical practice [49]. In addition, overfitting is likely to occur when directly training CNNs using limited ultrasound scans. Meanwhile, the CNNs pre-trained on natural images are not very suitable for FBSPs analysis. In order to increase the identification accuracy of automated ultrasound scan machines and decrease the computational requirements, a novel differential-CNNs for FBSPs identification was proposed.

First, the feature maps in the differential-CNN are generated from an original CNN with differential operators. In this way, in the differential-CNN, more differential feature maps are adopted to extract more details from images without increasing the numbers of convolutional layers and parameters. Therefore, the proposed differential-CNN used in the identification achieves better performance than the other deep learning methods in the FBSPs identification problem. Moreover, the proposed differential-CNN consumes less memory, which means that it can be applied in portable devices with computational limitations and is beneficial to reducing production costs [50].

Second, the development of deep learning in the field of medicine is based on large numbers of medical images. An FBSPs dataset was collected by our team, which can be used to not only evaluate the performance of our proposed method but also provide a data reference for other researchers. Moreover, our approach uses samples that have been diagnosed by clinicians for evaluation. The experimental results demonstrated that the identification of FBSPs using our model agreed with the diagnoses of clinicians. Our proposed differential-CNN demonstrated the great prospects of applying deep learning in clinical applications.

In future works, we will try to improve the hyper-parameters of the differential filters such as the number of filters, the filter size and the initial values in filters to make the network converge faster. We will also try to improve the network architectures by adding a multi-channel classifier, which has been proven to effectively improve the performance of identification systems.

REFERENCES

- [1] T. L. A. van den Heuvel, H. Petros, S. Santini, C. L. de Korte, and B. van Ginneken, "Automated fetal head detection and circumference estimation from free-hand ultrasound sweeps using deep learning in resource-limited countries," *Ultrasound Med. Biol.*, vol. 45, no. 3, pp. 773–785, Mar. 2019.
- [2] B. Potočník, B. Cigale, and D. Zazula, "Computerized detection and recognition of follicles in ovarian ultrasound images: A review," *Med. Biol. Eng. Comput.*, vol. 50, no. 12, pp. 1201–1212, Dec. 2012.
- [3] G. N. Nguyen, L. H. Son, A. S. Ashour, and N. Dey, "A survey of the state-of-the-arts on neutrosophic sets in biomedical diagnoses," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 1, pp. 1–13, Jan. 2019.
- [4] X. Yan, T. Yang, and H. Luo, "Diagnostic value of power Doppler ultrasonography for Sirenomelia Sequence in prenatal," *J. Sichuan Univ.*, vol. 42, no. 6, pp. 870–873, 2011.
- [5] D. S. Kermany, M. Goldbaum, and W. Cai, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, Feb. 2018.
- [6] C. Dubourg, V. Bendavid, and L. Pasquier, "Holoprosencephaly," *Orphanet J. Rare Diseases*, vol. 2, no. 1, pp. 1–14, 2007.
- [7] M. I. Daoud, A. A. Atallah, F. Awwad, and M. Al-Najar, "Accurate and fully automatic segmentation of breast ultrasound images by combining image boundary and region information," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 718–721.
- [8] C. F. Baumgartner, K. Kamnitsas, J. Matthew, T. P. Fletcher, S. Smith, L. M. Knch, B. Kainz, and D. Rueckert, "SonoNet: Real-time detection and localisation of fetal standard scan planes in freehand ultrasound," *IEEE Trans. Med. Imag.*, vol. 36, no. 11, pp. 2204–2215, Nov. 2017.
- [9] L. Zhang, S. Chen, C. T. Chin, T. Wang, and S. Li, "Intelligent scanning: Automated standard plane selection and biometric measurement of early gestational sac in routine ultrasound examination," *Med. Phys.*, vol. 39, no. 8, pp. 5015–5027, 2012.
- [10] A. I. L. Namburete, R. V. Stebbing, B. Kemp, M. Yaqub, A. T. Papageorgiou, and J. A. Noble, "Learning-based prediction of gestational age from ultrasound images of the fetal brain," *Med. Image Anal.*, vol. 21, pp. 72–86, Apr. 2015.
- [11] N. H. Khan, E. Tegnander, J. M. Dreier, S. Eik-Nes, H. Torp, and G. Kiss, "Automatic measurement of biparietal diameter with a portable ultrasound device," in *Proc. IEEE Int. Ultrason. Symp.*, Sep. 2014, pp. 459–462.
- [12] R. W. Mann and M. Kobayashi, "Biparietal thinning: Accidental death by a fall from standing height," *J. Forensic Sci.*, vol. 62, no. 5, pp. 1406–1409, 2017.
- [13] C. Desai, J. Eledath, H. Sawhney, and M. Bansal, "De-correlating CNN features for generative classification," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Jan. 2015, pp. 428–435.
- [14] H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern Recognit.*, vol. 43, no. 1, pp. 299–317, Jan. 2010.
- [15] Z. Chen, N. He, Y. Huang, W. T. Qin, X. Liu, and L. Li, "Integration of a deep learning classifier with a random forest approach for predicting malonylation sites," *Genomics, Proteomics Bioinf.*, vol. 16, no. 6, pp. 451–459, Dec. 2018.
- [16] W.-C. Yeh, S.-W. Huang, and P.-C. Li, "Liver fibrosis grade classification with B-mode ultrasound," *Ultrasound Med. Biol.*, vol. 29, no. 9, pp. 1229–1235, Sep. 2003.
- [17] B. Lei, E.-L. Tan, S. Chen, L. Zhuo, S. Li, D. Ni, and T. Wang, "Automatic recognition of fetal facial standard plane in ultrasound image via Fisher vector," *PLoS ONE*, vol. 10, no. 5, 2015, Art. no. e0121838.
- [18] Z. Cao, L. Duan, and G. Yang, "Breast tumor detection in ultrasound images using deep learning," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*. Quebec City, QC, Canada: Springer, 2017, pp. 121–128.
- [19] Z. Yu, E. L. Tan, D. Ni, J. Qin, S. Chen, S. Li, B. Lei, and T. Wang, "A deep convolutional neural network-based framework for automatic fetal facial standard plane recognition," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 3, pp. 874–885, May 2018.
- [20] A. Moshfegh, A. Javadzadegan, M. Mohammadi, L. Ravipudi, S. Cheng, and R. Martins, "Development of an innovative technology to segment luminal borders of intravascular ultrasound image sequences in a fully automated manner," *Comput. Biol. Med.*, vol. 108, pp. 111–121, May 2019.
- [21] A. Kumar, P. Sridar, A. Quinton, R. K. Kumar, D. Feng, R. Nanan, and J. Kim, "Plane identification in fetal ultrasound images using saliency maps and convolutional neural networks," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 791–794.
- [22] L. Wu, J.-Z. Cheng, S. Li, B. Lei, T. Wang, and D. Ni, "FUIQA: Fetal ultrasound image quality assessment with deep convolutional networks," *IEEE Trans. Cybern.*, vol. 47, no. 5, pp. 1336–1349, May 2017.
- [23] H. Chen, L. Wu, Q. Dou, J. Qin, S. Li, J.-Z. Cheng, D. Ni, and P.-A. Heng, "Ultrasound standard plane detection using a composite neural network framework," *IEEE Trans. Cybern.*, vol. 47, no. 6, pp. 1576–1586, Jun. 2017.

- [24] C. Bian, R. Lee, Y. H. Chou, and J. Z. Cheng, "Boundary regularized convolutional neural network for layer parsing of breast anatomy in automated whole breast ultrasound," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2017, pp. 259–266.
- [25] B. Lei, S. Huang, R. Li, C. Bian, H. Li, Y. H. Chou, and J. Z. Cheng, "Segmentation of breast anatomy for automated whole breast ultrasound images with boundary regularized convolutional encoder–decoder network," *Neurocomputing*, vol. 321, pp. 178–186, Dec. 2018.
- [26] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Proc. Adv. Neural Inform. Process. Syst.*, 2014, pp. 3320–3328.
- [27] R. J. Araujo, J. S. Cardoso, and H. P. Oliveira, "A deep learning design for improving topology coherence in blood vessel segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2019, pp. 93–101.
- [28] M. Sarigül, B. M. Ozyildirim, and M. Avci, "Differential convolutional neural network," *Neural Netw.*, vol. 116, pp. 279–287, Aug. 2019.
- [29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
- [30] J. Yang, Z. Lei, and S. Z. Li, "Learn convolutional neural network for face anti-spoofing," *Comput. Sci.*, vol. 92, no. 18, pp. 373–384, 2014.
- [31] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
- [32] L. Shi, W. Liu, H. Zhang, Y. Xie, and D. Wang, "A survey of GPU-based medical image computing techniques," *Quantitative Imag. Med. Surv.*, vol. 2, no. 3, pp. 188–206, 2012.
- [33] H. D. Cheng, X. Cai, X. Chen, L. Hu, and X. Lou, "Computer-aided detection and classification of microcalcifications in mammograms: A survey," *Pattern Recognit.*, vol. 36, no. 12, pp. 2967–2991, Dec. 2003.
- [34] E. Kussul and T. Baidyk, "Improved method of handwritten digit recognition tested on MNIST database," *Image Vis. Comput.*, vol. 22, no. 12, pp. 971–981, Oct. 2004.
- [35] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [36] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
- [37] G. E. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-R. Mohamed, N. Jaitly, and A. Senior, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Oct. 2012.
- [38] A.-R. Mohamed, G. E. Dahl, and G. Hinton, "Acoustic modeling using deep belief networks," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 1, pp. 14–22, Jan. 2012.
- [39] Y. Matsuda, H. Hoashi, and K. Yanai, "Recognition of multiple-food images by detecting candidate regions," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2012, pp. 25–30.
- [40] J. Yang, F. Xie, H. Fan, Z. Jiang, and J. Liu, "Classification for dermoscopy images using convolutional neural networks based on region average pooling," *IEEE Access*, vol. 6, pp. 65130–65138, 2018.
- [41] Y. Jia, C. Huang, and T. Darrell, "Beyond spatial pyramids: Receptive field learning for pooled image features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3370–3377.
- [42] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [43] H. Chen, D. Ni, J. Qin, S. Li, X. Yang, T. Wang, and P. A. Heng, "Standard plane localization in fetal ultrasound via domain transferred deep neural networks," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 5, pp. 1627–1636, Sep. 2015.
- [44] H. G. Lewis and M. Brown, "A generalized confusion matrix for assessing area estimates from remotely sensed data," *Int. J. Remote Sens.*, vol. 22, no. 16, pp. 3223–3235, Jan. 2001.
- [45] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A K-Means clustering algorithm," *Appl. Statist.*, vol. 28, no. 1, pp. 100–108, 1979.
- [46] C. W. Chen, J. Luo, and K. J. Parker, "Image segmentation via adaptive K-mean clustering and knowledge-based morphological operations with biomedical applications," *IEEE Trans. Image Process.*, vol. 7, no. 12, pp. 1673–1683, Dec. 1998.
- [47] Y. Tan and J. Wang, "A support vector machine with a hybrid kernel and minimal vovnik-chervonenkis dimension," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 4, pp. 385–395, Apr. 2004.
- [48] D. Ni, X. Yang, X. Chen, C.-T. Chin, S. Chen, P. A. Heng, S. Li, J. Qin, and T. Wang, "Standard plane localization in ultrasound by radial component model and selective search," *Ultrasound Med. Biol.*, vol. 40, no. 11, pp. 2728–2742, Nov. 2014.
- [49] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5353–5360.
- [50] J. Tang and N. Suga, "Modulation of auditory processing by cortico-cortical feed-forward and feedback projections," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 21, pp. 7600–7605, May 2008.



RUOWEI QU received the B.S. degree from the School of Science, University of Science and Technology Beijing, Beijing, China, in 2012. She is currently pursuing the Ph. D. degree with the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin, China. Her research interests include medical image processing, pattern recognition, and machine learning.



GUIZHI XU (Member, IEEE) was born in 1962. She received the Ph.D. degree from the School of Electrical Engineering, Hebei University of Technology, Tianjin, China, in 2002.

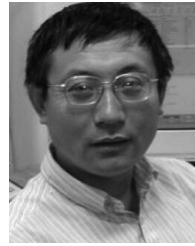
She is currently a Professor, the Doctoral Tutor, and the Dean of the School of Electrical Engineering, Hebei University of Technology. She is also the Head of the Key Subjects at the Provincial Level of Biomedical Engineering, a Professor of Meta-optics with the Hebei University of Technology, and the Head of the National Top-quality Course of Engineering Electromagnetic Field. She has published more than 90 academic articles retrieved by the SCI and EI and published three monographs. She has presided over one key project of the National Natural Science Foundation, three projects of the National Natural Science Foundation and one pre-research project of the Ministry of General Equipment, and completed two key projects of the National Natural Science Foundation in cooperation with Tsinghua University and the Fourth Military Medical University. She received the Hebei Science and Technology Outstanding Contribution Award, the Hebei Natural Science Second Prize and Third Prize, the Hebei Science and Technology Progress Second Prize and Third Prize, and the Hebei Excellent Teaching Achievement Second Prize. She achieved the honorary titles of the first famous teaching teacher in Hebei Province, an outstanding young and middle-aged expert in Hebei Province, an outstanding young and middle-aged backbone teacher in Hebei Province, an advanced individual in Hebei Province, and an outstanding Communist Party member in Hebei Province's education system.



CHUNXIA DING received the Bachelor of Medicine degree from the Zhangjiakou Medical College, in 2003. She has been involved in ultrasound diagnosis for 16 years and worked on prenatal screening for more than ten years. She is skilled in gynecology and obstetrics, the heart, blood vessels, the abdomen, superficial small organs, the neonatal brain and other parts of conventional diagnosis, prenatal screening, and fetal malformation diagnosis.



WENYAN JIA received the B.S. and M.S. degrees in biomedical engineering from Capital Medical University, Beijing, China, in 1998 and 2001, respectively, and the Ph.D. degree in biomedical engineering from Tsinghua University, Beijing, in 2005. She is currently a Research Assistant Professor with the Department of Neurosurgery, University of Pittsburgh, Pittsburgh, PA, USA. Her current research interests include biomedical signal processing and brain–computer interfacing.



MINGUI SUN (Senior Member, IEEE) received the B.S. degree in instrumental and industrial automation from the Shenyang Chemical Engineering Institute, Shenyang, China, in 1982, and the M.S. and Ph.D. degrees in electrical engineering from the University of Pittsburgh, Pittsburgh, PA, USA, in 1986 and 1989, respectively. He is currently a Professor of neurosurgery, electrical and computer engineering, and bioengineering with the University of Pittsburgh. His current research interests include advanced biomedical electronic devices, biomedical signal and image processing, sensors and transducers, and artificial neural networks.

• • •