IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 30, 2019, accepted June 22, 2019, date of publication July 2, 2019, date of current version July 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2925803

Fetal Weight Estimation via Ultrasound Using Machine Learning

MIAO FENG[1], LI WAN[2], ZHI LI[1], LINBO QING[1], (Member, IEEE), AND XIAORONG QI[2]
[1]College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China
[2]Development and Related Disease of Women and Children Key Laboratory of Sichuan Province, Key Laboratory of Birth Defects and Related Diseases of Women and Children, Ministry of Education, Department of Gynecology and Obstetrics, West China Second Hospital, Sichuan University, Chengdu 610041, China

Corresponding authors: Linbo Qing (qing_lb@scu.edu.cn) and Xiaorong Qi (qixiaorong11@163.com)

This work was supported in part by the Key Research and Development Project of Science and Technology Department of Sichuan Province under Grant 2019YFG0192.

ABSTRACT Accurate fetal weight estimation is important for both fetuses and their mothers. The low birth weight (LBW, birth weight < 2500 g) and high birth weight (HBW, birth weight ≥ 4000 g) fetuses and their mothers are linked to both short and long-term health outcomes such as high perinatal mortality rate, various complications, and chronic disease in life. Because of the imbalanced small data sets and body size heterogeneities between different fetal weight groups, it is difficult for the commonly used regression formulas to get a satisfying performance, especially for the HBW and LBW fetuses. The aim of this paper is to propose a machine learning solution to improve fetal weight estimation accuracy and to help the clinicians identify potential risks before delivery. A clinical data set of 7875 singleton fetuses were analyzed. The synthetic minority over-sampling technique (SMOTE) was employed to solve the imbalanced learning problem. Then, the support vector machine (SVM) algorithm was utilized for fetal weight classification. Finally, the deep belief network (DBN) was employed to estimate the fetal weight based on different ultrasound parameters. The estimation result of the proposed model showed a mean absolute percent error (MAPE) of $6.09 \pm 5.06\%$ and mean absolute error (MAE) of 198.55 ± 158.63g. It demonstrated that our model outperformed the commonly-used regression formulas, especially for the HBW and LBW fetuses.

INDEX TERMS Deep belief network, fetal weight estimation, synthetic minority over-sampling technique, support vector machine, ultrasound.

I. INTRODUCTION

Fetal weight is an essential factor to predict the short and long-term health consequences [1]. According to birth weight (BW), the neonates are defined by the World Health Organization (WHO) as three groups, namely low birth weight (LBW, BW < 2500g), normal birth weight (NBW, 2500g $\leq$ BW < 4000g) and high birth weight (HBW, BW ≥ 4000 g) which is also called macrosomia [2]. Low birth weight is connected with fetal and neonatal mortality and inhibited growth, it can also cause long-term diseases in their childhood, such as mental retardation and learning disabilities [3], [4]. Macrosomia can cause perinatal asphyxia and death, moreover, for maternities, the risk of caesarean section, prolonged labor, abnormal haemorrhage, and perineal trauma increases [5], [6]. In the long term, macrosomia is more likely to be associated with obesity, diabetes, and heart disease [4]. Therefore, it is significant to estimate fetal weight accurately during pregnancy and identify low birth weight fetuses or macrosomia correctly. Once the risk has been identified, the maternal or neonatal morbidity and mortality can be reduced by taking appropriate clinical decisions and precautions [7].

A recognized method for fetal weight estimation is ultrasound measurement since it is non-invasive, non-hazardous, and relatively accurate [8], [9]. Various regression formulas based on different combinations of ultrasound parameters have been introduced. Dudley [10] evaluated regression formulas delivered from 11 different methods and claimed that there was no consistently superior formula for ultrasound fetal weight estimation. All the regression formulas have a problem, which is that they perform well among NBW

The associate editor coordinating the review of this manuscript and approving it for publication was Yonghong Peng.

fetuses, but is likely to be less accurate when applied to the entire fetal weight ranges [11]. One particular reason is the imbalanced sample size, the real clinical data set consists primarily of NBW fetuses with just a small part of HBW fetuses and LBW fetuses, therefore, using a regression formula to estimate the birth weight of an HBW or LBW fetus is much more difficult than estimating the birth weight of an NBW fetus. Previous studies have reported that except for ultrasound parameters, there are multiple variables related to fetal birth weight, such as fetal sex, maternal age, height, large gestational weight gain, and gestational diabetes [12], [13]. It's hard for a simple traditional regression formula to reflect the complex multi-dimensional and nonlinear relation between all these variables and the fetal weight. Besides, most of the fixed regression formulas are derived from a typical clinical population, so it may be less accurate to apply these formulas to other populations [10].

Recently, artificial neural networks (ANN) has been applied by many researches to predict fetal weight to overcome the problems of traditional regression methods [14], [15], [16]. Comparison of estimated fetal weight (EFW) accuracy showed that these ANN models both significantly outperformed the commonly-used regression formulas. However, the case numbers of these methods were small, neural networks are more suitable for finding patterns in large sample data, and enlarging the neonatal cases may improve the performance of these ANN models. Although the mean accuracy of all fetal weight ranges has improved significantly, the estimation is still less accurate when body weight is below 2500g and above 4000g due to the imbalanced data sets. What's more, it is revealed that the fetal density, morphologic feature or body configurations are different among the three fetal weight groups [8], [11]. However, most of the methods were designed for the entire fetal weight ranges by using a single ANN model which ignore those differences between each group. Thus it is necessary to classify fetuses into different groups and determine different significant variables for each group. All these ANN models used the back propagation (BP) network as the learning algorithm to train the ANN. Nevertheless, the BP network suffers from an uncontrolled convergence speed and local optima, it also needs large numbers of tag data.

Deep belief network (DBN) [17] is a kind of deep learning model which is composed of multiple layers of restricted boltzmann machines (RBMs). DBN has a greedy layer-wise unsupervised pre-training process as well as a top-down fine-tuning procedure for optimizing the model's performance. The training process of DBN is faster than ANN, because the RBM is trained by just comparing the divergence. Besides, the pretraining procedure helps to find latent variables behind the data, which can be regarded as weights initialization of a BP network. Because of these advantages, DBN can avoid the problems that BP network have [18], [19].

In this study, we proposed a classification-based birth weight prediction model, which is built upon DBN networks. We collected 7875 singleton fetuses from West China

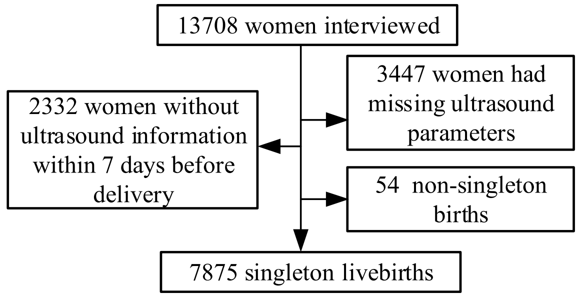

FIGURE 1. Flow chart of patient selection.

Second University Hospital, the synthetic minority oversampling technique (SMOTE) [20] was utilized to enlarge the training data of LBW and HBW fetuses to overcome the imbalanced learning problem. Then, the support vector machine (SVM) algorithm was utilized to classify the fetuses into two groups: BW < 4000g (LBW, NBW) and BW ≥ 4000g (HBW). Finally, these two groups were trained in two different DBN models respectively where each DBN model had different significant input parameters. The experimental results demonstrate that the proposed method outperforms the regression formulas and is an effective way to estimate the fetal birth weight. This paper also provides a possible method for predicting fetal weight in gestation period. We hope that the proposed method can help clinicians to assess fetal growth before delivery and provide valuable information in delivery management and clinical decision-making.

II. METHODS

A. DATA

This study was a retrospective review of delivery records in West China Second University Hospital between January 2016 and December 2017. As one of the medical centers for women and children in China. The patients are from all over China with different provinces and ethnic groups which can fully reflect the diversity of patients, thus making the data from this hospital represent the general population of China. All the fetuses were examined by ultrasound which was performed by skilled obstetric residents within seven days prior to delivery [11]. As shown in Fig. 1, the women who had missing ultrasound parameters or without ultrasound information within seven days of delivery were excluded, twins or multiple births were also removed. In total, 7875 women with singleton fetus were analyzed in our study.

Table 1 shows the maternal and fetal characteristics of the 7875 singleton neonates. There were 7200 (91.43%) NBW fetuses, 485 (6.16%) HBW fetuses, and 190 (2.41%) LBW fetuses. The age of these women was between 18 and 48 years old, with an average age of 30.81 $\pm$ 3.97 years. The mean gestational age at delivery was 39.50 $\pm$ 0.79 weeks (36-40weeks), the average actual birth weight was 3331.59 $\pm$ 409.18g (930-5120g).

In this study, the following six parameters before delivery were adopted in constructing the estimation models,

TABLE 1. Maternal and fetal characteristics of neonates (*n=7875*).

Maternal and fetal characteristics	Values
Maternal ages(years), mean ± SD	30.81 ± 3.97
Gestation at delivery (weeks), mean ± SD	39.50 ± 0.79
Actual birth weight (g), mean ± SD	3331.6 ± 409.2
2500 ≤ BW < 4000g, number	7200
BW ≥ 4000g, number	485
BW < 2500g, number	190

SD = standard deviation; BW = birth weight.

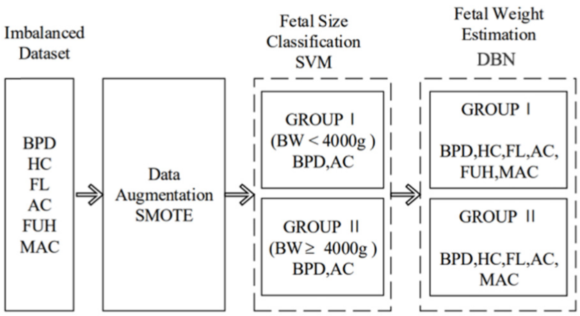

FIGURE 2. Architecture of the birth weight estimation model.

i.e. four ultrasonographic parameters including biparietal diameter (BPD), head circumference (HC), femur length (FL), abdominal circumference (AC), and two maternal parameters, namely maternal fundal height (FUH) and maternal abdominal circumference (MAC).

The architecture of the proposed birth weight estimation model is illustrated in Fig. 2, before fetal size classification, we employed SMOTE algorithm to solve the imbalanced learning problem. The fetuses were classified into group I (BW < 4000g) and group II (BW ≥ 4000g), then two DBN models were employed to estimate the birth weight of group I and group II separately. In this paper, we classify all the fetal samples into two categories instead of three because the case number of LBW samples are too small to achieve a satisfying classification performance, meanwhile we mainly focus on improving the prediction accuracy of NBW and HBW fetuses, so we treat the LBW samples and the NBW samples as the same class.

B. SMOTE-BASED DATA AUGMENTATION

As shown in Table 1, the sample numbers for HBW and LBW are insufficient, resulting in an imbalanced learning problem. Machine learning networks typically expect large amounts of balanced data sets. Therefore, when applied to complex data sets with imbalanced class distributions, these networks tend to provide inaccurate performance and cannot properly represent the distributive characteristics of the data [21].

SMOTE algorithm is an excellent synthetic data augmentation algorithm which has been applied in various areas, it can prevent overfitting problems compared with those methods that simply replicate the minority samples [20]. The illustration of SMOTE is shown in Fig. 3.

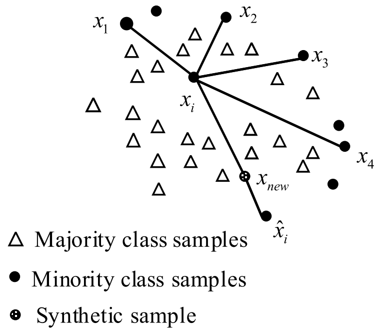

FIGURE 3. Illustration of synthetic minority over-sampling technique.

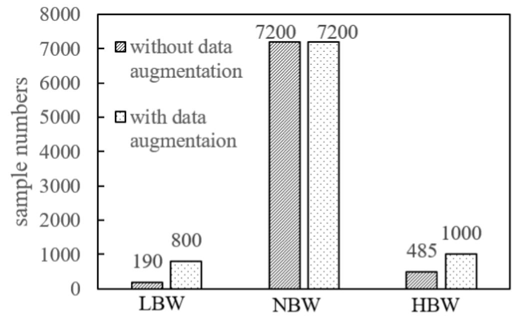

FIGURE 4. Distribution comparison of data with and without SMOTE-based data augmentation.

The steps of SMOTE are as follows:

(1) We define the given training data set as S, the set S_{min} is the minority class examples ($S_{min} \in S$), for each minority class sample $x_i \in S_{min}$, obtain its k-nearest neighbors within the minority class.

(2) Choosing neighbors from the k-nearest neighbors randomly based on the required over-sampling amount, assuming that the selected neighbor is $\hat{x}_i \in S_{min}$.

(3) Synthetic samples are generated according to (1): multiply the difference between the sample x_i and its nearest neighbor $\hat{x}_i$ by a random number η ($0 < \eta < 1$), then add it to the sample x_i to create a new synthetic sample x_{new}.

$$x_{new} = x_i + \eta(\hat{x}_i - x_i) \tag{1}$$

In this paper, the numbers of the nearest neighbors for each minority class sample was set to 5. The data distribution before and after the data augmentation is shown in Fig. 4, there were 9000 samples covering 7200(80%) NBW fetuses, 1000 (11.11%) HBW fetuses, and 800(8.89%) LBW fetuses after data augmentation. All these augmented data were only put into the training set, then we divided the other data without data augmentation into training and testing set randomly. Thus there is only real clinical data in the testing set. As a result, 6900 fetuses (76.66%) were included as the training set while 2100 fetuses (23.34%) made up the testing set.

C. SVM-BASED FETAL SIZE CLASSIFICATION

The body configurations of different fetal size may be different and therefore, different groups of fetal weight may

TABLE 2. Pearson correlation coefficient analysis for three fetal weight ranges and all parameters.

Fetal weight ranges	Pearson correlation coefficient					
	BPD	HC	FL	AC	MAC	FUH
LBW	0.706**	0.757**	0.785**	0.838**	0.081	0.168
NBW	0.480**	0.496**	0.523**	0.690**	0.333**	0.387**
HBW	0.199**	0.194**	0.185**	0.305**	0.158**	0

**, the parameter with a significance of $p < 0.01$; LBW = low birth weight (BW < 2500g); NBW = normal birth weight (2500g $\leq$ BW $<$ 4000g); HBW = high birth weight (BW $\geq$ 4000g).

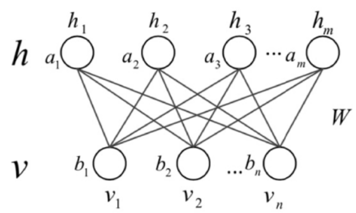

FIGURE 5. RBM structure.

be affected by different parameters. As illustrated in Fig. 2, in order to reduce the effect of body size heterogeneities between different fetal weight groups, SVM algorithm was used to classify fetal size before fetal weight regression. The fetuses were classified into two groups, group I (BW< 4000g) and group II (BW $\geq$ 4000g).

To determine which parameters were highly correlated with different fetal birth weight ranges, we utilized the Pearson correlation coefficient analysis to investigate the most significant parameters for each fetal weight range. The Statistical Product and Service Solutions (SPSS 22.0) was employed to perform this statistical analysis. The Pearson correlation coefficients between different fetal weight ranges and parameters are shown in Table 2. It shows that for HBW range, BPD and AC have a higher correlation coefficient with the actual birth weight than the other parameters, so BPD and AC are selected as the classification features for SVM.

D. DBN-BASED FETAL WEIGHT ESTIMATION

In our study, DBN was used to build the fetal birth weight estimation model, according to the classification result of SVM model, two DBN models with different input parameters were designed to predict the fetal birth weight. DBN is a multilayer structure consists of a series of individual RBMs [22], [23]. The RBM is a bipartite connectivity graph, as shown in Fig. 5, v and h represent the visible layer and the hidden layer of RBM respectively [24].

For one particular RBM, the energy of the joint configuration of (v, h) is defined as:

$$p(v, h\,|\theta) \propto \exp(-E(v, h\,|\theta)) = \exp(\sum_i b_i v_i + \sum_j a_j h_j + \sum_i \sum_j w_{ij} v_i h_j) \quad (2)$$

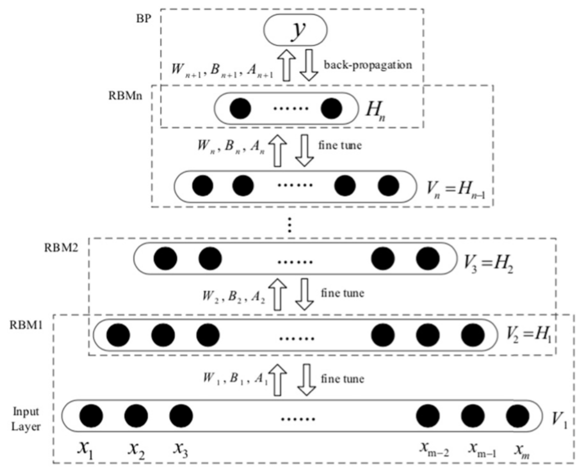

FIGURE 6. Architecture of the DBN model.

where $\theta = (w, b, a)$ is the model parameter set, w_{ij} is the weight between v_i and h_j, b_i and a_j are the bias for the layer v and layer h, respectively.

The activation probability of the j*th* hidden unit is:

$$p(h_j\,|v, \theta) = \frac{1}{1 + \exp(-\sum_i w_{ij} v_i - a_j)} \quad (3)$$

The activation probability of the i*th* visible unit is:

$$p(v_i\,|h, \theta) = \frac{1}{1 + \exp(-\sum_j w_{ij} h_j - b_i)} \quad (4)$$

where the activation function is sigmoid equation.

The gradient of log-likelihood function was used to optimize the θ, it can be described as:

$$\frac{\partial \log p(v\,|\theta)}{\partial w_{ij}} = \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model} \quad (5)$$

$$\frac{\partial \log p(v\,|\theta)}{\partial a_j} = \langle h_j \rangle_{data} - \langle h_j \rangle_{model} \quad (6)$$

$$\frac{\partial \log p(v\,|\theta)}{\partial b_i} = \langle v_i \rangle_{data} - \langle v_i \rangle_{model} \quad (7)$$

where $\langle\cdot\rangle_{data}$ and $\langle\cdot\rangle_{model}$ represent the expectations with respect to the training data distribution and the model distribution respectively. The $\langle\cdot\rangle_{data}$ can be obtained by calculating the conditional probability distributions from training set, while computing the $\langle\cdot\rangle_{model}$ is intractable, as a simple and efficient solution, the contrastive divergence (CD) was used to approximate $\langle\cdot\rangle_{model}$ [25].

The architecture of the DBN model is shown in Fig. 6. The DBN model consists of n RBM layers and a BP output layer, in our study, n is set to be 2. In the input layer, we use a sequence $V_1 = \{x_1, x_2, \ldots, x_m\}$ to represent the input parameters, where x_i represents each parameter and m is the number of the input parameters. As shown in Table 2, all parameters are significantly correlated with NBW fetuses ($p < 0.01$) while BPD, HC, FL, AC, MAC are significantly

TABLE 3. The result of the proposed birth weight estimation model (in testing set, *n= 2100*).

Fetal weight ranges	MAPE±SD (%)	MAE±SD (g)
ABW	6.09±5.06	198.55±158.63
LBW	11.09±10.38	245.94±174.20
NBW	5.65±4.49	179.70±147.79
HBW	5.85±5.04	244.84±217.70

ABW = all birth weight; LBW = low birth weight (BW<2500g); NBW = normal birth weight (2500g ≤BW<4000g); HBW = high birth weight (BW ≥4000g); MAPE = mean absolute percent error; MAE = mean absolute error; SD = standard deviation.

correlated with HBW fetuses ($p < 0.01$). So for Group I (BW < 4000g), BPD, HC, FL, AC, FUH, and MAC are selected as the input parameters, while the input parameters of Group II (BW ≥ 4000g) are BPD, HC, FL, AC, and MAC. In each RBM layer, V_i and H_i is the visible layer and hidden layer of this RBM layer respectively, and H_i is also regarded as the visible layer of the next RBM layer. There is a BP network on the top of the RBMs, which is regarded as the output layer, y represents the predicted fetal birth weight. W_i, B_i, and A_i are the connection weights and bias between two layers.

There are two steps in the training process of a DBN model:

Step I : pre training

In this step, each RBM layer is trained separately with the greedy unsupervised algorithm.

Step II: fine-tuning

In this step, a supervised BP network is utilized to fine tune the whole model, the input of this BP network is H_n, which is the output vector of the last RBM layer.

We evaluated the performance of the model by using mean absolute error (MAE) and mean absolute percent error (MAPE), the equations are represented as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{8}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i - \hat{y}_i\right|}{y_i} \times 100\% \tag{9}$$

where n is the number of fetuses, y_i and $\hat{y}_i$ are the actual birth weight and estimated birth weight of the i^{th} fetus.

III. RESULTS

A. PERFORMANCE COMPARISON OF IMBALANCED AND BALANCED DATA SETS

Table 3 shows the estimation performance of the proposed model, the MAPE and MAE for the fetuses of all the fetal weight ranges are 6.09± 5.06% and 198.55 ±158.63g. Table 4 provides the result of the prediction model without SMOTE based data augmentation. It can be observed that the prediction performance of minority groups HBW and LBW has been improved after data augmentation in the proposed model, especially for the HBW group, there is a 1.17% decline in MAPE and 49.02g decline in MAE respectively. Thus, it can be concluded that using the balanced data set to predict the fetal weight is more effective than just using the original imbalanced data set.

TABLE 4. The result of estimation model without smote-based data augmentation (in testing set, *n= 2100*).

Fetal weight ranges	MAPE±SD (%)	MAE±SD (g)
ABW	6.21±4.23	202.69±141.96
LBW	11.72±9.21	260.82±201.00
NBW	5.76±4.23	183.30±133.85
HBW	7.02±4.33	293.86±168.10

TABLE 5. The result of estimation model in the unclassified data sets (in testing set, *n= 2100*).

Fetal weight ranges	MAPE±SD (%)	MAE±SD (g)
ABW	6.17±5.17	201.22±155.33
LBW	11.89±11.01	248.94±184.56
NBW	5.76±4.50	188.08±144.32
HBW	7.10±4.77	297.07±206.54

B. PERFORMANCE COMPARISON OF CLASSIFIED AND UNCLASSIFIED DATA SETS

To investigate the effectiveness of SVM based classification in the proposed model, we validated the performance of the estimation model in the unclassified data set. Table 5 illustrates the result, compared with the result shown in Table 3, it was demonstrated that using SVM algorithm to classify the fetuses into different fetal size groups can improve the prediction performance, especially for HBW fetuses.

C. ACCURACY COMPARISON OF DIFFERENT FETAL WEIGHT ESTIMATION METHODS

We compared the fetal weight estimation accuracy of the proposed model with that of the commonly used regression formulas shown in Table 6, including the formulas of Hadlock *et al.* (1985) [26], Shepard *et al.* (1982) [27], Hsieh *et al.* (1987) [28], and Woo *et al.* (1985) [29]. We calculated the MAPE and MAE for each regression formula method in the same testing set (n=2100) as our proposed method. The result is illustrated in Table 7, compared with the accuracy of regression formulas, the proposed method increased the predictive accuracy MAPE for 0.25% to 21.01% and MAE for 12.49g to 681.24g. Fig. 7 and Fig. 8 illustrate that the proposed approach is superior to regression formulas, it can reach a statistically higher degree of accuracy.

For the regression formulas, the fetal weight estimation accuracy of different fetal birth weight ranges is shown in Table 8. As compared with the result of the proposed model shown in Table 3, it is demonstrated that our model outperforms the regression formulas among NBW and HBW fetuses, while for LBW fetuses, the estimation results of Hadlock1 (1985) and Hadlock2 (1985) are slightly better than those of the proposed model.

TABLE 6. Five published regression formulas for fetal weight estimation.

References	Formulas
Hadlock1 (1985)	$\log_{10}(BW)=1.326-0.00326\times AC\times FL+0.0107\times HC+0.0438\times AC+0.158\times FL$
Hadlock2 (1985)	$\log_{10}(BW)=1.3596-0.00386\times AC\times FL+0.0064\times HC+0.00061\times BPD\times AC+0.0424\times AC+0.174\times FL$
Shepard (1982)	$\log_{10}(BW)=-1.7492+0.166\times BPD+0.046\times AC-2.646\times AC\times BPD/1000$
Hsieh (1987)	$\log_{10}(BW)=2.7193+1.745\times 0.001\times BPD^2\times FL-7.6742\times 0.0001\times AC\times BPD^2-0.1432\times FL+9.4962\times 0.001\times AC\times BPD$
Woo (1985)	$\log_{10}(BW)=1.14+0.16\times BPD+0.05\times AC-2.8\times 0.001\times BPD\times AC+0.04\times FL-4.9\times 0.0001\times FL\times AC$

TABLE 7. The estimation results of regression formula methods (in testing set, *n= 2100*).

Methods	MAPE±SD (%)	MAE±SD (g)
Hadlock1 (1985)	6.34±4.76	211.90±163.40
Hadlock2 (1985)	6.42±4.93	211.04±159.65
Shepard (1982)	7.42±5.83	240.00±179.13
Hsieh (1987)	7.29±5.68	235.64±174.30
Woo (1985)	27.10±10.45	879.79±301.31

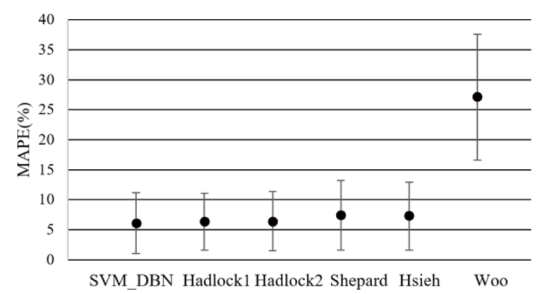

FIGURE 7. The mean absolute percent error (MAPE) of the proposed model as compared to other methods.

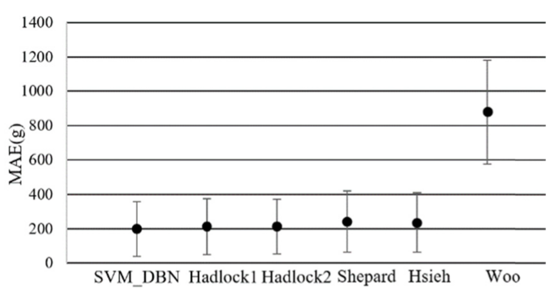

FIGURE 8. The mean absolute error (MAE) of the proposed model as compared to other methods.

IV. DISCUSSION

Most of the EFW models were designed for all fetal weight ranges by using a single formula or prediction model, it may have a high estimation accuracy among NBW group, but the accuracy of HBW and LBW was still not satisfactory. We considered one of the reasons was that the sample size of the minority weight ranges was small, in our study, the data augmentation algorithm SMOTE was used to solve this problem. Bernstein and Catalano claimed that for the HBW fetuses, their soft tissue mass in the limbs increased a lot, but the measurements of AC, HC, and FL may do not account for this, which would cause fetal weight underestimation [30]. Therefore, we hypothesized that when the fetal weight grows to more than 4000g, the growth trend is no longer reflected in the changes of ultrasound parameters, but in the changes of the maternal parameters such as FUH and MAC. So the FUH and MAC before delivery were adopted as the input parameters of the DBN models. Chuang *et al.* [8] also hypothesized other possible reasons, such as the density, the fetal morphologic feature might be different between each birth weight groups. In order to reduce the influence of body size heterogeneities between different fetal weight groups, SVM algorithm was utilized to classify fetuses into two groups: birth weight below 4000g and over or equal to 4000g, we analyzed these two groups and found the most significant parameters for each group. Then they were trained in two different DBN models with the corresponding input parameters. We compared the prediction results of the classified and unclassified models. As the estimation accuracy shown in Table 3 and Table 5, it is demonstrated that classify the fetuses into different groups and predict the birth weight using different significant parameters result in higher accuracy observably.

In our study, we investigated the effectiveness of SMOTE-based data augmentation, it is shown in Table 4 that if the original imbalanced dataset was utilized directly as the training set of the prediction model, the minority groups HBW and LBW fetuses would be seriously biased to the majority group NBW fetuses. As a result, the prediction accuracy of HBW and LBW fetuses were relatively low. After data augmentation based on SMOTE, a balanced training dataset was reconstructed, the performance of HBW groups had realized a significant increase. For the LBW group, there was a slight rise on the performance, because the sample size of this group was too small to get a satisfying improvement. It also proved that the increase of the sample number of HBW and LBW fetuses can improve the estimation performance, future study is suggested to collect more actual HBW and LBW fetuses to improve the prediction accuracy.

From Table 3 and Table 8, we found that the proposed method outperformed the regression formulas almost in all fetal ranges. There were also a lot of EFW researches based on ANN model. Wu *et al.* [16], Chuang *et al.* [8], and Cheng *et al.* [14] both adopted ANN to estimate birth weight, they all claimed that ANN model outperformed the regression formulas, however, their researches were all based on small study groups with a fetal case of 109, 1353 and 2127 respectively. Wu *et al.* [16], Chuang *et al.* [8] also claimed that the EFW was less accurate for LBW and HBW fetuses. In our study, there were a total of 7875 cases which was significantly larger than the above studies. The sufficient

TABLE 8. Estimation results of the regression formulas for different fetal weight ranges (in testing set, *n*= *2100*).

Fetal weight ranges	Hadlock1 (1985)	Hadlock2 (1985)	Shepard (1982)	Hsieh (1987)
LBW- MAPE±SD (%)	10.19±7.09	10.91±7.44	12.38±8.39	12.13 ± 8.13
LBW- MAE±SD (g)	230.49±163.95	247.38±173.23	282.05±196.26	273.64±187.14
NBW- MAPE±SD (%)	5.81±4.31	6.01±4.57	7.24±5.61	7.13±5.47
NBW- MAE±SD (g)	190.87±141.22	195.27±144.16	233.32±173.06	230.11±169.15
HBW- MAPE±SD (%)	9.26±5.22	7.92±5.05	6.63±4.99	6.34±4.80
HBW- MAE±SD (g)	391.38±225.98	334.82±217.65	279.51±212.74	267.34±204.96

balanced sample cases are more conducive to the DBN to analyze the complex relationship between input and output, resulting in higher accuracy. In addition, DBN can solve the problem of uncontrolled convergence speed and local optima of ANN.

There are several limitations of the study that should be recognized. First, those women who gave birth to twins or multiple births were excluded, and it may prevent us from fully understanding the birth weight of both singleton, twins, and multiple births. Future study may take these patients into consideration. Secondly, there are various parameters we can get from an ultrasound measurement, such as occipito-frontal diameter (OFD), femur length (FL), head circumference (HC), biparietal diameter (BPD), abdominal circumference (AC), and fetal gender. In our study, we selected BDP, HC, FL, AC, FUH, and MAC as the input parameters of the proposed model according to the expert's clinical experience and Pearson correlation analysis. However, studies showed that different ultrasound parameters were associated with birth weight at different degrees, and different combinations of parameters may affect the prediction results. Cheng *et al*. [14] cross-validate the significance among ultrasound parameters by using Spearman correlation analysis, it showed that compared with other factors, AC is most correlated with birth weight. In addition, several maternal characteristics, such as pre-existing or gestational diabetes, prolonged gestation, maternal age, significant weight gain during pregnancy, and body mass index are also linked to fetal weight [4], [31]. Future studies are warranted to improve the accuracy of EFW by adding more effective parameters and the rational combination of maternal and fetal parameters is the key to establish a better estimation model. Thirdly, the proposed model improved the estimation accuracy at the extreme weight ranges (HBW, LBW) due to the balanced data sets produced by SMOTE-based data augmentation. However, there are also drawbacks in SMOTE algorithm, such as variance and over generalization, the variant of SMOTE such as Borderline-SMOTE can be utilized to improve the estimation performance [32]. Besides, these augmented samples cannot replace the real neonates. To solve this problem, further studies are suggested to collect more data from real HBW and LBW fetuses born in hospitals. Finally, this work was a retrospective study from a single center, in the further work, a multicenter prospective study is planned to expand the sample size to test the accuracy of the proposed model.

V. CONCLUSION

In this study, we proposed a novel fetal weight estimation model which combined SVM based classification with DBN to improve the performance of EFW in all fetal weight ranges, we also solved the imbalanced learning problem by utilizing SMOTE based data augmentation. It was demonstrated from the result that the proposed model outperformed the regression formulas. Our study revealed that DBN is a promising approach for fetal weight estimation, it also proved that classify fetuses into different groups and predict birth weight using different significant parameters are effective. We believe that the proposed method may help clinicians to assess fetal growth during the pregnancy and provide valuable information in delivery management and clinical decision-making. With respect to the future work, we intend to collect more new-born fetuses of all fetal weight ranges from multiple centers, explore additional effective parameters and optimize the structure of DBN to improve the prediction accuracy.

APPENDIX

Abbreviations and full names:

BW: Birth weight
LBW: Low birth weight (birth weight $<$ 2500g)
NBW: Normal birth weight (2500g$\leq$ birth weight $<$ 4000g)
HBW: High birth weight (birth weight $\geq$4000g)
SMOTE: Synthetic minority over-sampling technique
SVM: Support vector machine
DBN: Deep belief network
MAPE: Mean absolute percent error
MAE: Mean absolute error
WHO: the World Health Organization
ANN: Artificial neural networks
EFW: Estimated fetal weight
BP: Back propagation
RBMs: Restricted boltzmann machines
BPD: Biparietal diameter
HC: Head circumference
FL: Femur length
AC: Abdominal circumference
FUH: Maternal fundal height
MAC: Maternal abdominal circumference
CD: Contrastive divergence
OFD: Occipito-frontal diameter

ACKNOWLEDGMENT

The authors would like to thank Dr. Xiaorong Qi and Mrs. Li Wan at the Department of Obstetrics and Gynecology of West China Second University Hospital for clinical data organization. They would like to also thank all doctors for the ultrasound measurements.

REFERENCES

[1] S. Badshah, L. Mason, K. McKelvie, R. Payne, and P. J. Lisboa, ''Risk factors for low birthweight in the public-hospitals at Peshawar, NWFP-Pakistan,'' *BMC Public Health*, vol. 8, p. 197, Jun. 2008.

[2] J. Berard, P. Dufour, D. Vinatier, D. Subtil, S. Vanderstichéle, J. C. Monnier, and F. Puech, ''Fetal macrosomia: Risk factors and outcome. A study of the outcome concerning 100 cases >4500 G,'' *Eur. J. Obstetrics Gynecology Reproductive Biol.*, vol. 77, no. 3, pp. 51–59, Mar. 1998.

[3] F. Chiarotti, A. M. Castignani, M. Puopolo, F. Menniti-Ippolito, and E. D. Minniti, ''Effects of socio-environmental factors on neurocognitive performance in premature or low-birth weight preschoolers,'' *Annali Dell'Istituto Superiore Di Sanita*, vol. 37, no. 4, pp. 553–559, 2001.

[4] H. G. Mengesha, A. D. Wuneh, B. Weldearegawi, and D. L. Selvakumar, ''Low birth weight and macrosomia in Tigray, Northern Ethiopia: Who are the mothers at risk?'' *BMC Pediatrics*, vol. 17, no. 1, p. 144, Jun. 2017.

[5] K. Haram, J. Pirhonen, and P. Bergsjo, ''Suspected big baby: A difficult clinical problem in obstetrics,'' *Acta Obstetricia et Gynecologica Scandinavica*, vol. 81, no. 3, pp. 185–194, Mar. 2002.

[6] T. Henriksen, ''The macrosomic fetus: A challenge in current obstetrics,'' *Acta Obstetricia et Gynecologica Scandinavica*, vol. 87, no. 2, pp. 134–145, 2008.

[7] A. M. Phillips, A. B. Galdamez, S. T. Ounpraseuth, and E. F. Magann, ''Estimate of fetal weight by ultrasound within two weeks of delivery in the detection of fetal macrosomia,'' *Austral. New Zealand J. Obstetrics Gynaecology*, vol. 54, no. 3, pp. 441–444, Oct. 2014.

[8] L. Chuang, J. Y. Hwang, C. H. Chang, C. H. Yu, and F. M. Chang, ''Ultrasound estimation of fetal weight with the use of computerized artificial neural network model,'' *Ultrasound Med. Biol.*, vol. 28, no. 3, pp. 991–996, Aug. 2002.

[9] R. M. Farmer, A. L. Medearis, G. I. Hirata, and L. D. Platt, ''The use of a neural network for the ultrasonographic estimation of fetal weight in the macrosomic fetus,'' *Amer. J. Obstetrics Gynecology*, vol. 166, no. 3, pp. 1467–1472, May 1992.

[10] N. J. Dudley, ''A systematic review of the ultrasound estimation of fetal weight,'' *Official J. Int. Soc. Ultrasound Obstetrics Gynecology*, vol. 25, no. 3, pp. 80–89, Jan. 2005.

[11] R. E. Sabbagha, J. M. Dmin, R. K. Tamura, and S. A. Hungerford, ''Estimation of birth weight by use of ultrasonographic formulas targeted to large-, appropriate-, and small-for-gestational-age fetuses,'' *Amer. J. Obstetrics Gynecology*, vol. 160, no. 3, pp. 854–860, Apr. 1989.

[12] A. Koyanagi, J. Zhang, and A. Dagvadorj, ''Macrosomia in 23 developing countries: An analysis of a multicountry, facility-based, cross-sectional survey,'' *Lancet*, vol. 381, no. 3, pp. 476–483, Feb. 2013.

[13] M. Oliver, G. Mcnally, and L. Leader, ''Accuracy of sonographic prediction of birth weight,'' *Austral. New Zealand J. Obstetrics Gynaecology*, vol. 53, no. 3, pp. 584–588, Dec. 2013.

[14] Y. C. Cheng, G.-L. Yan, Y. H. Chiu, F.-M. Chang, C.-H. Chang, and K.-C. Chung, ''Efficient fetal size classification combined with artificial neural network for estimation of fetal weight,'' *Taiwanese J. Obstetrics Gynecology*, vol. 51, no. 3, pp. 545–553, Dec. 2012.

[15] H. Mohammadi, M. Jahanseir, Z. Allahmoradi, and F. Samiee, ''Ultrasound estimation of fetal weight by artificial neural network in normal pregnancies,'' in *Proc. Int. Conf. Meas. Control Eng.*, 2010, pp. 575–579.

[16] J. Wu, T. Yang, J. Lin, H. Luo, and D. Li, ''Estimation of fetal weight on the basis of neural network,'' *J. Biomed. Eng.*, vol. 22, no. 5, pp. 922–925, Oct. 2005.

[17] G. E. Hinton, S. Osindero, and Y.-W. Teh, ''A fast learning algorithm for deep belief nets,'' *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.

[18] Q.-S. Tan, W. Huang, and Q. Li, ''An intrusion detection method based on DBN in ad hoc networks,'' in *Proc. Int. Conf. Wireless Commun. Sensor Netw. (WCSN)*, 2015, pp. 477–485.

[19] C. Shang, F. Yang, D. Huang, and W. Lyu, ''Data-driven soft sensor development based on deep learning technique,'' *J. Process Control*, vol. 24, no. 3, pp. 223–233, 2014.

[20] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, ''SMOTE: Synthetic minority over-sampling technique,'' *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.

[21] H. He and E. A. Garcia, ''Learning from imbalanced data,'' *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.

[22] G. E. Hinton and R. R. Salakhutdinov, ''Reducing the dimensionality of data with neural networks,'' *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[23] Y. Bengio, ''Learning deep architectures for AI,'' *Found. Trends Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.

[24] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, ''A learning algorithm for Boltzmann machines,'' *Cogn. Sci.*, vol. 9, no. 1, pp. 147–169, Jan./Mar. 1988.

[25] G. E. Hinton, ''Training products of experts by minimizing contrastive divergence,'' *Neural Comput.*, vol. 14, no. 8, pp. 1771–1800, 2002.

[26] F. P. Hadlock, R. B. Harrist, R. S. Sharman, R. L. Deter, and S. K. Park, ''Estimation of fetal weight with the use of head, body, and femur measurements—A prospective study,'' *Amer. J. Obstetrics Gynecology*, vol. 151, no. 3, pp. 333–337, Feb. 1985.

[27] M. J. Shepard, V. A. Richards, R. L. Berkowitz, S. L. Warsof, and J. C. Hobbins, ''An evaluation of two equations for predicting fetal weight by ultrasound,'' *Amer. J. Obstetrics Gynecology*, vol. 142, no. 3, pp. 47–54, Jan. 1982.

[28] F. J. Hsieh, F. M. Chang, H. C. Huang, C. C. Lu, T. M. Ko, and H. Y. Chen, ''Computer-assisted analysis for prediction of fetal weight by ultrasound-comparison of biparietal diameter (BPD), abdominal circumference (AC) and femur length (FL),'' *Taiwan Yi Xue Hui Za Zhi*, vol. 86, no. 3, pp. 957–964, Sep. 1987.

[29] J. S. Woo, C. W. Wan, and K. M. Cho, ''Computer-assisted evaluation of ultrasonic fetal weight prediction using multiple regression equations with and without the fetal femur length,'' *J. Ultrasound Med.*, vol. 4, no. 2, pp. 65–67, Feb. 1985.

[30] I. M. Bernstein and P. M. Catalano, ''Influence of fetal fat on the ultrasound estimation of fetal weight in diabetic mothers,'' *Obstetrics Gynecology*, vol. 79, no. 3, pp. 561–563, Apr. 1992.

[31] T. Wondie, D. Jara, and M. Ayana, ''Factors associated with macrosomia among neonates delivered at Debre Markos referral hospital, Northwest Ethiopia, 2014: A case control study,'' *J. Diabetes Metabolism*, vol. 5, no. 12, p. 468, Jan. 2014.

[32] B. X. Wang and N. Japkowicz, ''Imbalanced data set learning with synthetic samples,'' in *Proc. IRIS Mach. Learn. Workshop*, Jun. 2004.

MIAO FENG received the B.S. degree in electronic information engineering from Northwest A&F University, Yangling, China, in 2013. She is currently pursuing the M.S. degree with the College of Electronics and Information Engineering, Sichuan University, Chengdu, China.

Her research interests include medical data analysis, machine learning, and deep learning.

LI WAN is currently the Head Nurse with the West China Second University Hospital of Sichuan University.

ZHI LI received the B.S. degree in electronics engineering from Shenyang Aerospace University, Shenyang, China, in 1997, and the M.S. degree in pattern recognition and intelligent systems and the Ph.D. degree in applied mathematics from Sichuan University, Chengdu, China, in 2000 and 2004, respectively.

He joined the Autonomous Computing Laboratory, University of Arizona, USA, in 2013. He is currently a Professor with Sichuan University. His research interests include medical data analysis technologies, biomedical signal processing, compressive sensing, and spectrum sensing.

LINBO QING (M'16) received the B.S. degree in electronic information science and technology and the Ph.D. degree in communication and information system from Sichuan University, China, in 2003 and 2008, respectively.

He is currently an Associate Professor with the College of Electronics and Information Engineering with Sichuan University. His main research interests include image processing, visual computing, data mining, and video coding and transmission.

XIAORONG QI received the Ph.D. degree in obstetrics and gynecology, in 2011.

She is currently an Associate Professor with the West China Second University Hospital of Sichuan University. Her main research interests include perinatal medicine and gynecologic oncology.