Research Article
Assessing Dry Weight of Hemodialysis Patients via Sparse Laplacian Regularized RVFL Neural Network with L_{2,1}-Norm

Xiaoyi Guo,1 Wei Zhou,1 Qun Lu,2 Aiyan Du,1 Yinghua Cai,2 and Yijie Ding3

1Hemodialysis Center, The Affiliated Wuxi People’s Hospital of Nanjing Medical University, 214000 Wuxi, China
2Nursing Department, The Affiliated Wuxi People’s Hospital of Nanjing Medical University, 214000 Wuxi, China
3School of Electronic and Information Engineering, Suzhou University of Science and Technology, 215009 Suzhou, China

Correspondence should be addressed to Wei Zhou; 285403434@qq.com, Yinghua Cai; 179098331@qq.com, and Yijie Ding; wuxi_dyj@163.com

Received 1 January 2021; Revised 21 January 2021; Accepted 25 January 2021; Published 5 February 2021

1. Introduction

Fluid overload in patients with chronic renal failure is closely related to poor cardiovascular outcomes [1, 2]. Maintenance of hemodialysis (HD) is the main method for patients with renal failure [3]. However, the accurate assessment of body water volume is still a concern [4]. At present, dry weight has been used as an important indicator to assess the homeostasis of fluids in hemodialysis patients. Medical staff can use the patient’s dry weight to estimate the amount of water needed for dialysis during hemodialysis. The conventional clinical-based dry weight assessment method is time-consuming and labor-intensive [1]. There are already some methods based on bioelectrical impedance analysis (BIA) [5] to determine dry weight, including body composition monitor (BCM) [6] and lung ultrasound (LUS). However, all the above methods require special instruments and professional technicians to complete. Medical staff can use some clinical data to build predictive models [7] to accurately assess dry weight. Currently, machine learning (ML) or deep learning has solved many common clinical problems in medicine, such as brain diseases [8–10], cancer analysis, and diabetes.

Some scholars have used artificial neural networks (ANN) to predict the total water volume of hemodialysis patients and have obtained better results than conventional clinical calculation equations [11]. In addition, deep learning methods are also emerging in clinical diagnosis, including pixel-based convolutional neural networks to diagnose skin cancer [12]. In the biological field, microbiology analysis [13], CircRNAs [14], microRNAs, and cancer association prediction [15–17], lncRNA-miRNA association prediction, O-GlcNAcylation site prediction [18], DNA methylation site [19–21], protein remote homology [22], function prediction
of proteins [23–29], electron transport proteins [30], breast cancer [31], cell-specific replication [32], osteoporosis diagnoses [33], and drug complex network analysis [34–38].

In our previous research, a Multiple Kernel Support Vector Regression (MKSVR) [39] predictor was proposed to assess the dry weight and obtain good predictive performance. Inspired by the previous work and baseline Random Vector Functional Link (RVFL) network [40], we propose a new dry weight assessment model, called Sparse Laplacian regularized RVFL neural network with L1,1-norm (SLapRVFL), which considers the topological relationship between samples and more sparse connections between the input layer and the hidden layer.

2. Materials and Methods

2.1. Materials. This work collects demographic and anthropometric data and bioimpedance spectroscopy (BIS) from historical data (2018-9 to 2019-9) from Wuxi people’s hospital and the northern Jiangsu people’s hospital. This study has been approved by the ethics committees of the hospitals (Nos. KYLLKS201813 and 2018KY-001). The collected patient data meet the following requirements: age greater than 18 years; ESRD for more than three months and maintenance hemodialysis [41]; no heart failure, no metal implants, no pregnancy, no disability, no infection, and no edema and other diseases; and hemodialysis treatment 3 times a week, 4 hours each time. Finally, we obtain a data set of 476 hemodialysis patients. DW is the normal body weight after clinical diabetes. DW is obtained by a clinician under strict clinical supervision using a clinical scoring system (using trial and error method) [42, 43].

We choose 7 features, including age, gender (binary feature), systolic blood pressure (SBP), diastolic blood pressure (DBP), body mass index (BMI), heart rate (HR), and years of dialysis (YD) to build our predictive model. Table 1 shows the information of the data set. BMI is measured before hemodialysis treatment.

2.2. Methods. The baseline RVFL was proposed for regression or classification. The schematic diagram of RVFL is shown in Figure 1. The basic information of the patient is put into the RVFL neural network model for processing, and the predicted dry weight is the output.

Suppose, there are \( N \) training samples with \( \{x_i, y_i\}, i = 1, 2, \cdots, N \). The output value is \( y_i \in \mathbb{R}^{1 \times c} \) and the input data is \( x_i \in \mathbb{R}^{1 \times d} \). \( d \) denotes the dimension of \( x_i \). As per Figure 1, RVFL randomly initializes all weights and deviations between the hidden layer and the input layer. These parameters are fixed during the training process and do not need to be tuned. There are connections between the output layer, input layer, and hidden layer. This part of the weight needs to be obtained by training RVFL. The output layer of RVFL is connected to both the input layer and the hidden layer, so as to ensure the nonlinear and linear relationships between the input and the output. The RVFL network with \( P \) hidden nodes are formulated as

\[
H \beta = Y, \tag{1}
\]

\[
H = [H_1 \; H_2], \tag{2}
\]

\[
H_1 = \begin{bmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{Nd} \end{bmatrix}_{N \times d}, \tag{3}
\]

\[
H_2 = \begin{bmatrix} G(a_1 x_1 + b_1) & \cdots & G(a_p x_1 + b_p) \\ \vdots & \ddots & \vdots \\ G(a_1 x_N + b_1) & \cdots & G(a_p x_N + b_p) \end{bmatrix}_{N \times p}, \tag{4}
\]

\[
\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_P^T \end{bmatrix}_{(d+p) \times C}, \tag{5}
\]

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
<th>( r^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>54.17 ± 14.22</td>
<td>-0.2341</td>
</tr>
<tr>
<td>Gender (males/females)</td>
<td>312/164</td>
<td>-0.4489</td>
</tr>
<tr>
<td>BMI</td>
<td>22.96 ± 2.95</td>
<td>0.9558</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>150.64 ± 29.36</td>
<td>-0.1739</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
<td>88.32 ± 19.56</td>
<td>-0.1249</td>
</tr>
<tr>
<td>Heart rate (times/min)</td>
<td>73.41 ± 8.92</td>
<td>0.1862</td>
</tr>
<tr>
<td>Years of dialysis (years)</td>
<td>5.97 ± 3.22</td>
<td>-0.1069</td>
</tr>
</tbody>
</table>

\*Denotes that each feature correlated with dry weight using Pearson correlation coefficient (\( r \)).

Figure 1: Schematic of our proposed method.

where \( \beta \) denotes the output weight matrix; \( H \) is the concatenated matrix, which combines the output of the hidden layer and the input layer; and \( Y \) denotes the label matrix. \( H \) and \( \beta \) can be represented as
In Equation (4), $a_j$ and $b_j$ are the weights and bias of the hidden and input layers. $C$ and $P$ are numbers of output and hidden layer nodes. In general, the activation function is a Gaussian function: $g(x) = e^{-x^2}$. The activation function has a nonlinear approximation effect. To consider the potential linear relationship between the input data and the output value, RVFL adds a direct connection weight between the input layer and the output layer. Therefore, RVFL is a model that contains both linear and nonlinear approximations to improve prediction performance. For optimal $\beta$, the RVFL can be formulated as a regularized least-squares:

$$
\beta^* = \text{arg min} \frac{1}{2} \|H\beta - Y\|^2 + \frac{\lambda}{2} \|\beta\|^2,
$$

where $\lambda$ is the parameter of regularization term. The solution of Equation (6) can be found by setting its gradient to 0:

$$
\beta^* = (H^T H + \lambda I)^{-1} H^T Y,
$$

where $I$ denotes the identity matrix. However, the RVFL network did not consider the topological relationship between samples. For the output node, it must be connected to both the input and the hidden layer.

In order to further improve the robustness of RVFL, we propose Sparse Laplacian regularized RVFL neural network with $L_{2,1}$ norm (SLapRVFL). The objective function is

$$
\beta^* = \text{arg min} \frac{1}{2} \|H\beta - Y\|^2 + \frac{\lambda_1}{2} \text{Tr}(H\beta)^T LH\beta + \frac{\lambda_2}{2} \|\beta\|^2_{2,1},
$$

where $L \in R^{N \times N}$ denotes the Laplacian matrix. $\lambda_1$ and $\lambda_2$ are the coefficients of Laplacian regularization and $L_{2,1}$-norm term, respectively. Laplacian regularization is used to indicate the potential manifold between samples. It can better describe the topological association between samples to improve the generalization ability of the model. Since the third term of $\|\beta\|^2_{2,1}$ is not diversified, we convert Equation (8) to

$$
\beta^* = \text{arg min} \frac{1}{2} \|H\beta - Y\|^2 + \frac{\lambda_1}{2} \text{Tr}(H\beta)^T LH\beta + \frac{\lambda_2}{2} \text{Tr}(\beta^T G\beta),
$$

where $G \in R^{(d+P) \times (d+P)}$ denotes a diagonal matrix whose ith-diagonal element

$$
G_{ii} = \frac{1}{2} \|\beta_i\|^2, \quad i = 1, 2, \cdots, (d + P).
$$

We take the derivative of the formula Equation (10) as

$$
H^T (H\beta - Y) + \lambda_1 H^T LH\beta + \lambda_2 G\beta = 0,
$$

$$
H^T H\beta + \lambda_1 H^T LH\beta + \lambda_2 G\beta = H^T Y,
$$

**Figure 2:** The RMSE under different numbers of hidden layer nodes (SLapRVFL network).
\((H^TH + \lambda_1 H^TLH + \lambda_2 G)\beta = H^TY,\)
\(\beta = (H^TH + \lambda_1 H^TLH + \lambda_2 G)^{-1}H^TY.\)

We use the baseline RVFL solution with Equation (7) as the initial \(\beta^0\). In addition, the Laplacian matrix can be calculate as
\[L = D^{-1/2}\Delta D^{-1/2},\]
\[\Delta = D - S,\]
where \(D\) is diagonal matrix, \(D_{ii} = \sum_{j=1}^{N} S_{ij}\). Similarity matrix \(S\) is built by Radial Basis Function (RBF):
\[S_{ij} = \exp \left(-\gamma ||x_i - x_j||^2\right).\]

The process of SLapRVFL is list in Algorithm 1.

3. Results

We test our model on the benchmark data set and obtain the optimal parameters of the predictor through cross-validation. The SLapRVFL network is compared to other machine learning-based models. In addition, the body composition monitor (BCM) device (Fresenius Medical Care, Baden Humboldt, Germany) is also compared with the SLapRVFL network.

3.1. Evaluation Measurements. The 10-fold cross-validation (10-CV) is employed to evaluate the robustness of methods. Root Mean Square Error (RMSE), \(R^2\) square, correlation coefficient (\(R\)), Bland–Altman analysis, and Empirical Cumulative Distribution Plot (ECDP) [44] are all used in our study. To evaluate the agreement of two different methods, the Bland–Altman analysis usually can obtain whether the two
Table 2: Comparison on existing methods via 10-fold cross-validation.

<table>
<thead>
<tr>
<th>Method</th>
<th>$R$</th>
<th>$R$ squared</th>
<th>RMSE</th>
<th>Empirical cumulative distribution plot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Highest value</td>
</tr>
<tr>
<td>BCM*</td>
<td>0.9473</td>
<td>0.9137</td>
<td>1.9694</td>
<td>3.2235</td>
</tr>
<tr>
<td>LR*</td>
<td>0.9403</td>
<td>0.9308</td>
<td>1.4335</td>
<td>4.2524</td>
</tr>
<tr>
<td>ANN (BP)*</td>
<td>0.9398</td>
<td>0.9295</td>
<td>1.4794</td>
<td>7.3661</td>
</tr>
<tr>
<td>MKRR*</td>
<td>0.9399</td>
<td>0.9289</td>
<td>1.5015</td>
<td>4.9227</td>
</tr>
<tr>
<td>MKSVR*</td>
<td>0.9412</td>
<td>0.9321</td>
<td>1.3817</td>
<td>4.3962</td>
</tr>
<tr>
<td>RVFL</td>
<td>0.9389</td>
<td>0.9300</td>
<td>1.3828</td>
<td>6.7004</td>
</tr>
<tr>
<td>SLapRVFL (our method)</td>
<td>0.9632</td>
<td>0.9501</td>
<td>1.3136</td>
<td>3.1940</td>
</tr>
</tbody>
</table>

*The results are from previous work on MKSVR [39].

Figure 5: Folded empirical cumulative distribution plot between different methods.

Table 3: Bland–Altman plot analysis for different models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Differences with DW (%)</th>
<th>Limits of agreement (%)</th>
<th>Number (ratio) of outside agreement interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCM*</td>
<td>Mean: -1.8232, SD: 2.7466, 95% confidence interval: -2.0706 to -1.5759</td>
<td>Lower limit: -7.2066, Upper limit: 3.5601</td>
<td>Number: 30/476 (6.30%)</td>
</tr>
<tr>
<td>LR*</td>
<td>Mean: 0.0002, SD: 2.4269, 95% confidence interval: -0.2184 to 0.2187</td>
<td>Lower limit: -4.7566, Upper limit: 4.7569</td>
<td>Number: 21/476 (4.41%)</td>
</tr>
<tr>
<td>ANN (BP)*</td>
<td>Mean: 0.1152, SD: 2.5139, 95% confidence interval: -0.1112 to 0.3416</td>
<td>Lower limit: -4.8119, Upper limit: 5.0424</td>
<td>Number: 22/476 (4.62%)</td>
</tr>
<tr>
<td>MKRR*</td>
<td>Mean: -0.0801, SD: 2.5007, 95% confidence interval: -0.3053 to 0.1451</td>
<td>Lower limit: -4.9814, Upper limit: 4.8212</td>
<td>Number: 23/476 (4.83%)</td>
</tr>
<tr>
<td>MKSVR*</td>
<td>Mean: -0.2638, SD: 2.3372, 95% confidence interval: -0.4743 to -0.05329</td>
<td>Lower limit: -4.8446, Upper limit: 4.3171</td>
<td>Number: 22/476 (4.62%)</td>
</tr>
<tr>
<td>SLapRVFL (our method)</td>
<td>Mean: 0.0867, SD: 2.2202, 95% confidence interval: -0.1133 to 0.2866</td>
<td>Lower limit: -4.2650, Upper limit: 4.4383</td>
<td>Number: 20/476 (4.20%)</td>
</tr>
</tbody>
</table>

*The results are from previous work on MKSVR [39].
Figure 6: Bland–Altman plot analysis.
methods can be substituted for each other (equivalence). Evaluating the agreement of the two methods can answer the question, "Can these two methods replace each other?"

3.2. Selection of Optimal Parameters. To get the optimal parameters of the predictive method, we obtain them through a grid search method. The parameters that need to be determined include the numbers of hidden layer nodes \( P \), maximum iterations, and coefficients of \( \lambda_1 \) and \( \lambda_2 \). For the numbers of hidden layer nodes \( P \), we fix the iterations, \( \lambda_1 \) and \( \lambda_2 \). Setting the maximum number as 50, \( \lambda_1 = 1 \) and \( \lambda_2 = 1 \). The value of \( P \) is from 10 to 140 with step of 10. The results are shown in Figure 2. From 10 to 100, the more neurons in the hidden layer, the lower the RMSE. Since then, RMSE has gradually increased. So, we get the lower RMSE under \( P = 100 \).

Next, \( P = 100 \), \( \lambda_1 = 1 \), and \( \lambda_2 = 1 \). We gradually increase the number of iterations from 1 to 100 (shown in Figure 3). After the number of iterations reaches 10, the RMSE value drops to a minimum and slightly oscillates within a certain value. In our study, maximum number of iterations is 10.

Then, we use the better number of hidden layer nodes and iterations to search for the best \( \lambda_1 \) and \( \lambda_2 \). The search range of parameters is from \( 2^{-5} \) to \( 2^{10} \) (with step of \( 2^{0.5} \)). Figure 4 shows the results of different parameters. When \( \lambda_1 \) and \( \lambda_2 \) are \( 2^{-3} \) and \( 2^{-2.5} \), RMSE is the lowest.

3.3. Comparison to Other Predictive Models and BCM. To evaluate our model, SLapRVFL is compared with our previous work of Multiple Kernel Support Vector Regression (MKSVR) [39], Multikernel Ridge Regression (MKRR), Linear Regression (LR), Artificial Neural Network based on Back Propagation algorithm (ANN with BP), and BCM measuring instrument. Clinical dry weight is our reference standard (also the regression target value of the prediction model). The comparisons are listed in Table 2, which shows that SLapRVFL achieves best performance of RMSE (1.3136). Although the ECDP median value (peak) of MKSVR (0.0082) is more closed to zero, Figure 5 shows that SLapRVFL has the least bias and much less tails than MKSVR (smaller width). The RMSE of BCM is 1.9694, which is larger than SLapRVFL.

3.4. Bland–Altman Analysis. Bland–Altman plot is a useful tool to evaluate the agreement between predictive methods and clinical DW. In Table 3 and Figure 6, SLapRVFL, MKSVR, LR, ANN (BP), MKRR, and BCM are analyzed via Bland–Altman difference plot. SLapRVFL achieves the smallest range of 95% confidence interval (-0.1133 to 0.2866) and standard deviation (2.2202). In addition, the number (ratio) of outside agreement interval for predictive models is all less than 24 (5%) predictive samples. These results of models are clinically acceptable. SLapRVFL achieves least number (20) of the outside agreement interval in Table 3. As shown in Figure 6, two red horizontal dotted lines (upper and lower) denote the upper and lower limits of the 95% agreement limit, respectively. The middle blue solid line is the average value of the difference (between measurement methods and clinical DW). While one measurement method and clinical method can be considered as a better agreement, they can be substituted for each other (equivalence). If 95% of the points of the data set are in the agreement range, the measurement method (predictive model) is clinically acceptable. The results of the evaluation show that SLapRVFL can help clinicians assess DW with low cost.

4. Discussion

Due to the limitations of clinical and BCM measurement (more time and cost), this study uses a machine learning method to assess the dry weight of hemodialysis patients. Based on the basic RVFL, we propose a sparse Laplace regularized RVFL network (SLapRVFL) model. SLapRVFL is compared not only with other machine learning methods (such as LR, MKRR, ANN with BP, and MKSVR) but also with BCM equipment (commonly used in hospitals). The RMSE and Bland–Altman analysis of the model are better than the BCM instrument. It is proven that the predictive model driven by data can provide reference for clinical dry weight assessment.

BCM requires the patient’s information on weight (before hemodialysis) and height. It is a portable, inexpensive, and noninvasive technology that has been used to measure DW [45, 46]. For the Bland–Altman analysis, SLapRVFL achieves the least number (20) of outside agreement interval. However, BCM has 30/476 (6.30%) points (ratio) of the outside agreement interval. Obviously, our method has better agreement with the clinical method.

5. Conclusions

To further improve the robustness of RVFL, we introduce sparse Laplacian regular term with \( L_{2,1} \)-norm. In the training process, the graph topology information and the sparse weight matrix (output) are employed to improve the robustness of the RVFL. In fact, our work provides a new idea for assessing patients’ dry weight. Not only that, in the fields of biology [47–57], pharmacy [58], and medicine [12, 59, 60], machine learning methods have helped solve many analysis tasks. In future research, we will consider collecting more samples, introducing more patient personal information, and building a predictor based on a deep learning model to more accurately assess the dry weight of hemodialysis patients.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Ethical Approval

This study had been approved by the ethics committee of the hospital (ethical approval Nos. KYLLKS201813 and 2018KY-001). The experimental protocol was established, according to the ethical guidelines of the Helsinki Declaration, and was approved by the Human Ethics Committee (Wuxi
People’s Hospital Ethics Committee and Northern Jiangsu People’s Hospital Ethics Committee).

Consent
Written informed consent for publication was obtained from all participants.

Conflicts of Interest
The authors declare that they have no conflict of interest.

Authors’ Contributions
Xiaoyi Guo and Wei Zhou are joint first authors.

Acknowledgments
The authors give their thanks to the Hemodialysis Center of Wuxi People’s Hospital and Northern Jiangsu People’s Hospital for collecting the data in this study. This work is supported by a grant from the National Natural Science Foundation of China (NSFC 61902271, 61772362, and 61972280) and the Natural Science Research of Jiangsu Higher Education Institutions of China (19KJB520014).

References


