



## Near infrared spectroscopy quantification based on Bi-LSTM and transfer learning for new scenarios

Ailing Tan <sup>a</sup>, Yunxin Wang <sup>a,\*</sup>, Yong Zhao <sup>b</sup>, Bolin Wang <sup>a</sup>, Xiaohang Li <sup>a</sup>, Alan X. Wang <sup>c</sup>

<sup>a</sup> School of Information and Science Engineering, Yanshan University, The Key Laboratory for Special Fiber and Fiber Sensor of Hebei Province, Qinhuangdao 066004, China

<sup>b</sup> School of Electrical Engineering, Yanshan University, The Key Laboratory of Measurement Technology and Instrumentation of Hebei Province, Qinhuangdao 066004, China

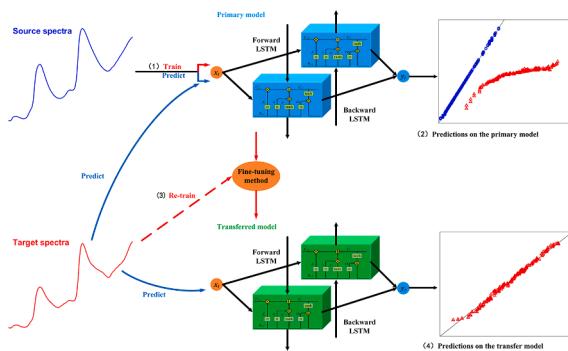
<sup>c</sup> Department of Electrical and Computer Engineering, Baylor University, Waco, TX 76706, USA



### HIGHLIGHTS

- A deep transfer learning methodology based on Bi-LSTM is proposed for the first time to address the NIR model transfer issue for samples produced in different conditions.
- Bi-LSTM modeling outperformed 1D-CNN, PLS and ELM for source manure and  $\gamma$ -PGA cases.
- Transfer learning by fine-tuning Bi-LSTM layers and re-training FC layers performed even better than by fixing Bi-LSTM layers and only re-training FC layers.
- A Bi-LSTM-fine-tuning enabled NIRS technology is beneficial to improve the adaptability of NIR spectroscopy in various new application scenarios.

### GRAPHICAL ABSTRACT



### ARTICLE INFO

#### Keywords:

Near infrared spectroscopy  
Bi-directional Long Short-Term Memory  
Transfer learning  
Fine-tuning  
Manure  
 $\gamma$ -PGA

### ABSTRACT

This study proposed a deep transfer learning methodology based on an improved Bi-directional Long Short-Term Memory (Bi-LSTM) network for the first time to address the near infrared spectroscopy (NIR) model transfer issue between samples. We tested its effectiveness on two datasets of manure and polyglutamic acid ( $\gamma$ -PGA) solution, respectively. First, the optimal primary Bi-LSTM networks for cattle manure and the first batch of  $\gamma$ -PGA were developed by ablation experiments and both proved to outperform one-dimensional convolutional neural network (1D-CNN), Partial Least Square (PLS) and Extreme Learning Machine (ELM) models. Then, two types of transfer learning approaches were carried out to determine model transferability to non-homologous samples. For poultry manure and the second batch of  $\gamma$ -PGA, the obtained predicting results verified that the second approach of fine-tuning Bi-LSTM layers and re-training FC layers transcended the first approach of fixing Bi-LSTM layers and only re-training FC layers by reducing the  $RMSEP_{target}$  of 23.4275% and 50.7343%, respectively. Finally, comparisons with fine-tuning 1D-CNN and other traditional model transfer methods further identified the superiority of the proposed methodology with exceeding accuracy and smaller variation, which decreased  $RMSEP_{target}$  of poultry manure and the second batch of  $\gamma$ -PGA of 7.2832% and 48.1256%, 67.1117%

\* Corresponding author at: School of Information and Science Engineering, Yanshan University, Qinhuangdao, Hebei Province, China.  
E-mail address: [james\\_wang2020@163.com](mailto:james_wang2020@163.com) (Y. Wang).

and 80.6924% when compared to that acquired by fine-tuning 1D-CNN, Tradaboost-ELM and CCA-PLS which were the best of five traditional methods, respectively. The study demonstrates the potential of the Fine-tuning-Bi-LSTM enabled NIR technology to be used as a simple, cost effective and reliable detection tool for a wide range of applications under various new scenarios.

## 1. Introduction

The combination of Near Infrared spectroscopy (NIRS) and chemometrics has been widely applied for quantitative detection of complex chemical substances [1,2]. However, a major practical challenge is the generality of NIRS, in which the developed predictive model only works well for spectra that has the same features as the original calibration set, and usually severely degrades the quantification accuracy when new spectral features are introduced by different backgrounds. For example, it often leads to unreliable predictions when the model is directly used for NIR spectra acquired by a different type of spectrometer due to the difference in photodetector, illumination source, spectral range, and resolution [3,4], or directly used for samples in a new scenario such as from a different source or production batch, because the differences may emerge local shifts in peaks or variations in absorption intensities [5,6]. Model transfer aims to solve the problem of mismatch between new spectra and the established NIR model. In the field of chemometrics, many studies have focused on the model transfer between instruments [7–9], whereas little research has dealt samples produced in different conditions [10]. A quantitative NIRS model with high transferability between samples is crucial for many practical engineering applications.

There are two traditional model transfer approaches that are frequently undertaken: The first approach is to correct the target spectra based on spectral standardization to correlate the primary spectra that was used in the development of the model. Various spectral standardization methods, including Direct Standardization (DS) [11], Piecewise Direct Standardization (PDS) [12], and Canonical Correlation Analysis (CCA) have been studied by calculating transfer matrices between target and primary spectra [13]. The other approach is to improve the model itself by enhancing the algorithm or using additional models. The most typical method is Slope Bias Correction (SBC) which corrects the univariate slope and biases of the target calibration set based on the source linear model to revise the predicted values [14]. These traditional approaches are usually used for model calibration between different spectrometers [15–17]. Furthermore, corrections by PDS, CCA and SBC are only suitable for linear multivariate models, such as partial least square (PLS) regression analysis [18,19], they are not desirable when there exist non-linear relationships between NIR spectra.

In machine learning, a similar concept to model transfer is called transfer learning (TL), which aims to transfer the obtained knowledge of source area to target area to make the transferred model more robust [20,21]. Among which, Transfer Component Analysis (TCA) and Tradaboost are two effective approaches. TCA is popular in supervised transfer learning, the source and target data are both mapped into a high-dimensional regenerative Hilbert space, in which the maximum mean discrepancy(MMD) is used to minimize the distribution difference between source and target data [22,23]. TCA is simple to implement, but it requires more calculation time for high dimensional spectral data. Tradaboost is especially available for migrations of regression issue by increasing the weight of samples in target area that has high spectral similarity with the source area [24,25]. Subsequently, the transfer performance becomes limited when the characteristics of target spectra differ greatly from source spectra.

Deep learning (DL) algorithms have revealed spectroscopy analysis to an even higher level [26,27]. One-dimensional convolutional neural network (1D-CNN) is one of the most widely used one-dimensional spectral analysis method [28–30]. In some recent studies, researchers extracted low-order to high-order features of spectral signals, such as NIR [31], Raman [32], Magnetic Resonance (MR) [33] and Terahertz

(THz) spectroscopy [34], by arranging the convolution layers in different ways. For regression tasks, several fully connected layers (FCs) are usually added to the end of the convolutional layer to form a regressor [35,36]. Long- and short-term memory (LSTM) is a special type of recurrent neural network (RNN) which is designed to learn long-distance dependencies and can overcome the gradient extinction and explosion problems of RNN [37,38]. For example, LSTM deep network has been used for MR spectroscopy detection of pseudo brain tumors [39], Raman spectroscopy screening of hepatitis B [40], THz spectrum detection of internal defects in glass fiber reinforced polymer [41] and NIRS detection of cumin and fennel [42].etc. However, Like the classical chemometrics methods, DL models must also be updated or adapted to capture new variabilities induced by new spectrometers or different samples, which can be accomplished by deep transfer learning(DT) for neural network layers [43,44]. DT is prevalent in computer vision and natural language processing, which aims to utilize the already acquired knowledge to avoid the time-consuming work of re-labeling samples and the expensive task of training large network from scratch [45,46]. DT algorithms basically contain two main parts: retrieval blocks including but not limited to the convolutional framework that performs feature extraction and combination, and a mapping block that is usually fully connected layers (FCs) to map extracted features to target variable [43,44]. For certain computer vision transfer tasks, the pre-trained weights of the underlying network are taken as the initial values in the new task, herein, the transferring process has a better “starting line” because these pre-trained weights reflect the extracted features of source data [47,48]. Besides, there are evidence that by initializing and updating the weights in the mapping block (FC layers), target features can be learned to adapt to the new data even though the new prediction data may be different from the original data [49,50].

In the field of NIRS, most research on deep network model and migration are based on 1D-CNN, while LSTM is relatively rare. LSTM is a good candidate for NIRS since spectral data that arranged by wavelength or frequency magnitude behaves like time-frequency sequence [42]. Inspired by the correlations of long bands in spectrum mined by LSTM and the deep transfer learning of CNN, we proposed a quantitative NIRS method with high transferability enabled by an improved bidirectional LSTM (Bi-LSTM) architecture and fine-tuning network approach. We proved the effectiveness of our method based on two cases: openly shared NIR dataset of fresh cattle manure and poultry manure [51], and experimentally acquired NIR dataset of two batches of polyglutamic acid ( $\gamma$ -PGA) solution products. Fig. 1 is the flowchart that summarizes our work of NIRS modeling and transferring. We first developed an improved neural network based on Bi-LSTM and FCs to build the NIR primary model using source data. Then, based on two types of fine-tuning approaches, we re-calibrated the primary networks using the target data to obtain the transferred model. We further tested the predicting accuracy and transferability of the proposed methodology by comparing with 1D-CNN and traditional model transfer algorithms respectively. To the best of knowledge, this is the first attempt to combine LSTM and DT for NIRS quantitative analysis between different samples, which provides a new approach for promoting the transfer of NIRS quantitative modeling for new scenarios.

## 2. Materials and methods

### 2.1. Data sets

#### 2.1.1. Fresh cattle manure and poultry manure data set

Near-infrared spectroscopy and chemical measurements of fresh cattle manure and poultry manure are provided by the French National Institute for Agri-Food and Environment website (<https://doi.org/10.15454/JIGO8R>). Manure samples were from livestock production areas in mainland France and Reunion Island: Broilers were fed with grains (wheat, maize, barley), oilseeds, vitamin, and minerals, dairy and feedlot cattle were fed with hay, sugarcane residue and concentrated feed. The dry matter (DM) content in fresh cattle manure ranges from 11.2550 to 46.6120 % (wet weight), and the DM content in fresh poultry manure ranges from 27.0700 to 82.4800 %, therefore, the cattle manure and poultry manure were non-homologous fertilizers with different range of DM contents. The NIR spectra of 196 fresh cattle manure samples and 136 fresh poultry manure samples were acquired by 3 spectrometers in 3 different laboratories (2 Foss XDS, 1 Buchi NIRflex). In this study, DM content was chosen as the property of interest, the dataset of 196 cattle manure were used as source data, and 136 poultry manure were used as target data.

#### 2.1.2. Polyglutamic acid ( $\gamma$ -PGA) solutions data set

Polyglutamic acid ( $\gamma$ -PGA) solution is a kind of biological fertilizer in which the  $\gamma$ -PGA composition has the effect of improving the rooting rate of seedlings, transplanting survival rate and fertilizer utilization rate. The  $\gamma$ -PGA solutions in this work came from two batches of products produced by Qinhuangdao Dongnuo Biotechnology Co., Ltd., Hebei Province, China, and 30 samples were taken from each batch. During the production of two batches, the fermentation substrate and bacteria were different, hence, the samples of two batches were non-homologous. The concentrations of  $\gamma$ -PGA composition in two batches were both 50 g/L, we diluted two batches of original solutions (50 g/L) respectively to

obtain 300 samples with concentration values of 50 g/L, 38.7 g/L, 25 g/L, 18.75 g/L and 12.5 g/L. The NIR spectra of 300 samples were collected by a MPA Fourier Transform Near Infrared Spectrometer (Bruker Optics Inc., Germany) equipped with TE-InGaAs detector and Opus 6.5 analysis software. The spectral wavelength scanning range was 12493–4000 cm<sup>-1</sup>, the scanning resolution was set at 4 cm<sup>-1</sup>, each sample was measured for 3 times and averaged. The dataset of the first batch of 150  $\gamma$ -PGA solutions were used as source data (measured in November 2021), and the dataset of the second batch of 150  $\gamma$ -PGA solutions were used as target data (measured in January 2022).

### 2.2. Bi-LSTM network architecture

Bi-LSTM is a neural network composed of a forward LSTM and a backward LSTM, which splices hidden layer state variables of the forward and backward neural networks. Under this mechanism, Bi-LSTM neural network can fully extract features according to the bi-directional information of NIR characteristic spectral bands. Its hidden layer state architecture is shown in Fig. 2.

In Fig. 2, the structure of forward-propagated LSTM and back-propagated LSTM are symmetrical, and  $f_t$  is the output of the “forget gate” of LSTM which is calculated as follows:

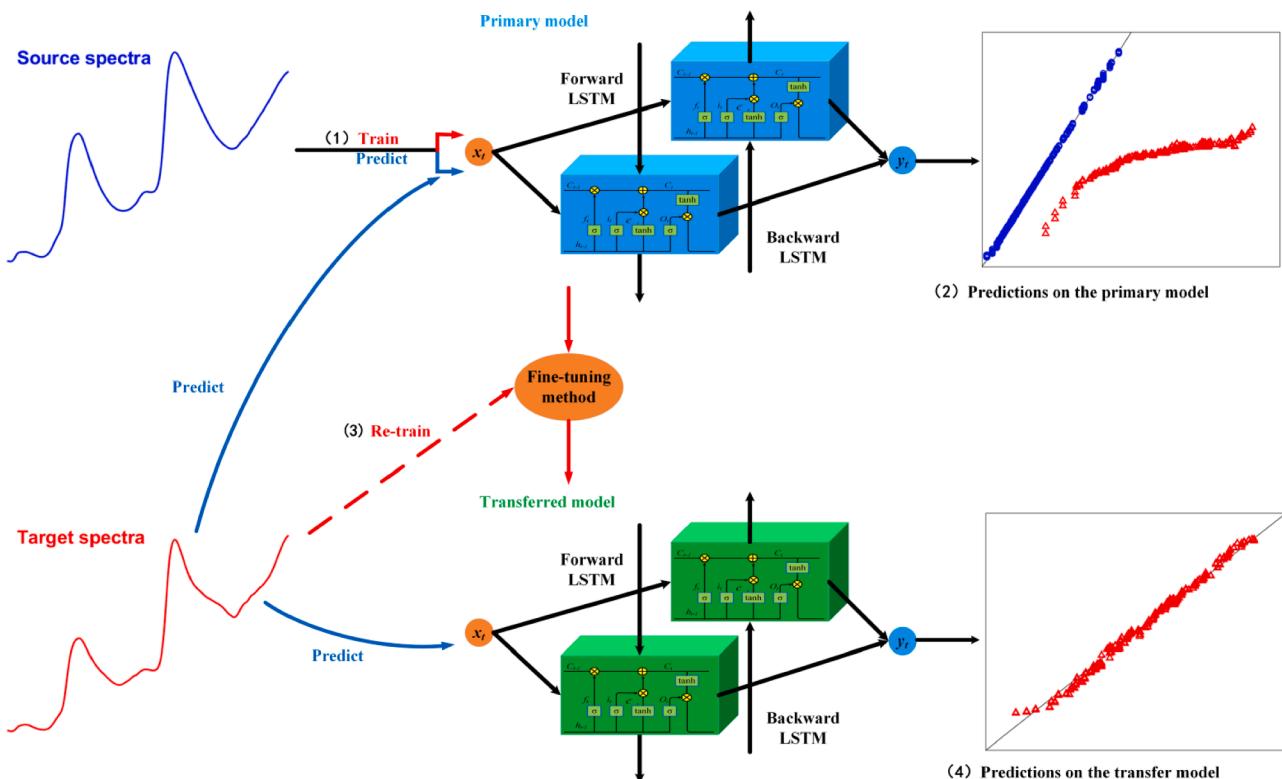
$$f_t = \text{sigmoid}(y_1) \quad (1)$$

$$y_1 = w_f * h_{t-1} + u_f * x_t + b_f \quad (2)$$

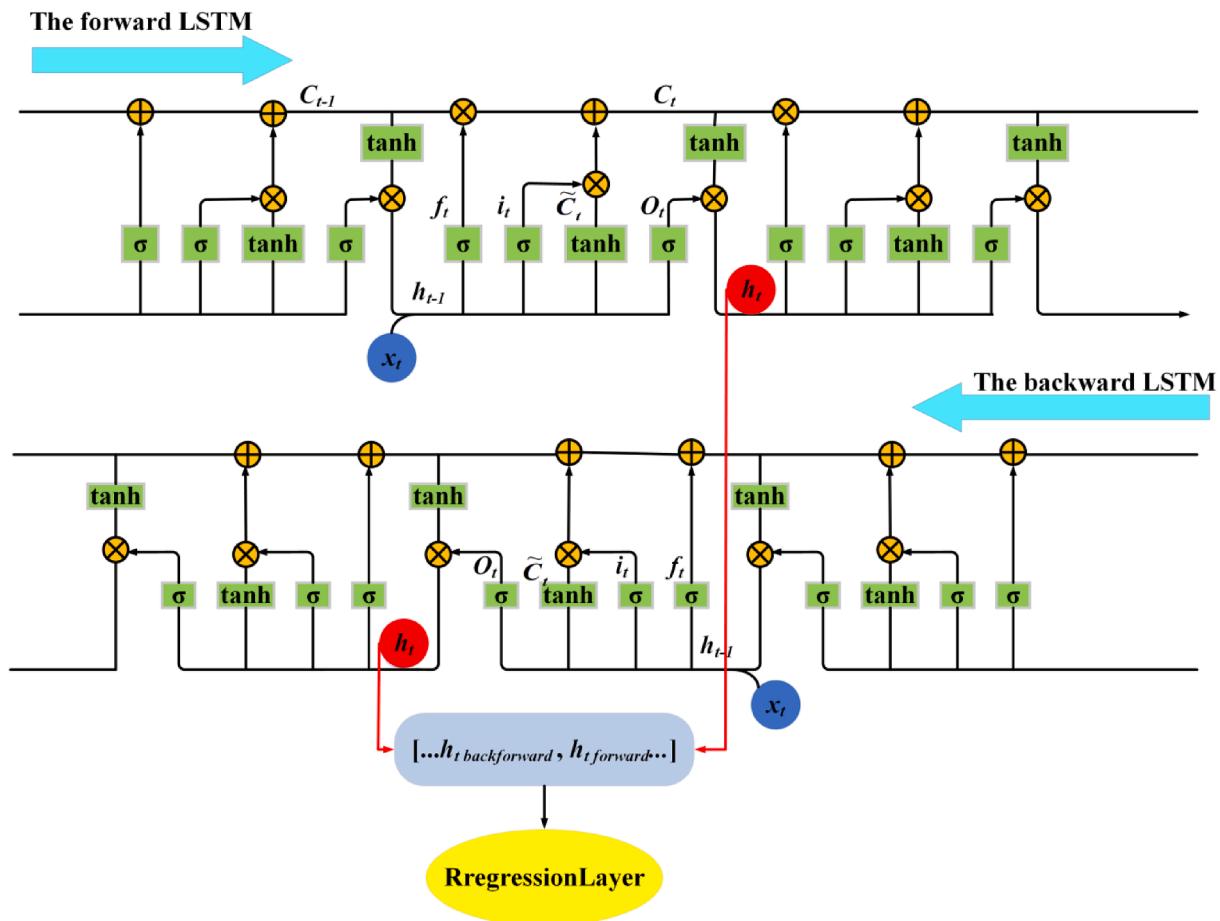
where  $\text{sigmoid}(y_1) = \frac{1}{1+e^{-y_1}}$ ;  $h_{t-1}$ 、 $x_t$  are the hidden layer state of the  $t-1$ -th wavelength and the spectral data of the  $t$ -th wavelength;  $w_f$ 、 $u_f$ 、 $b_f$  are the hidden layer weights, input weights and biases of the forget gate, respectively.

$i_t$  is the output of the “memory gate” of LSTM which is expressed as formula (3):

$$i_t = \text{sigmoid}(w_i * h_{t-1} + u_i * x_t + b_i) \quad (3)$$



**Fig. 1.** Flowchart: (1)Train the primary Bi-LSTM network; (2)Test the primary model with source & target spectra; (3)Fine-tuning the primary network; (4)Test the transferred model with target spectra.



**Fig. 2.** The architecture of Bi-LSTM hidden layer state.

where  $w_i$ ,  $u_i$ ,  $b_i$  are the hidden layer weights, input weights and biases of the memory gate, respectively.

$\tilde{C}_t$  is the temporary neuron state, and is calculated as follows:

$$\tilde{C}_t = \tanh(y2) \quad (4)$$

$$y2 = w_c * h_{t-1} + u_c * x_t + b_c \quad (5)$$

where  $\tanh(y2) = \frac{e^{y2} - e^{-y2}}{e^{y2} + e^{-y2}}$ ;  $w_c$ ,  $u_c$ ,  $b_c$  are the hidden layer weights, input weights and biases of temporary neurons, respectively.

$C_t$  is the neuron state of LSTM at the  $t$ -th wavelength point, which not only reflects the neuron partial spectral information of LSTM forgets and inherits at the  $t-1$ -th wavelength, but also represents the temporary neuron state retained at the  $t$ -th wavelength. The formula is as follow:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

$O_t$  is the output of the "output gate" of LSTM at the  $t$ -th wavelength, and is calculated based on formula (7):

$$O_t = \text{sigmoid}(w_o * h_{t-1} + u_o * x_t + b_o) \quad (7)$$

where  $w_o$ ,  $u_o$ ,  $b_o$  are the hidden layer weights, input weights and biases of the output gate, respectively.

$h_t$  is the hidden layer state of LSTM at the  $t$ -th wavelength, which represents the hidden information obtained by LSTM at the  $t$ -th wavelength point, and is expressed by formula (8):

$$h_t = O_t * \tanh(C_t) \quad (8)$$

All hidden layer states of the forward LSTM neural network  $\{h_0, h_1, \dots, h_{t-1}\}$  are spliced with all hidden layer states of the backward LSTM neural network  $\{h'_0, h'_1, \dots, h'_{t-1}\}$  to obtain a hidden layer state sequence

$\{(h_0, h'_0), (h_1, h'_1), \dots, (h_{t-1}, h'_{t-1})\}$  with twice the number of wavelength, which is used to represent the correlation features extracted by Bi-LSTM for NIR spectra.

In this study, the fusion method of FCs and Bi-LSTM by element addition is designed to improve the performance of the network. The advantages of this feature fusion method lie in that FCs not only assist to extract spectral features, but also avoid redundancy of network parameters and save computation time. In addition, FCs also serve as important carriers for updating and storing the parameters of model transferring, which is conducive to maintain a strong transferability of the source neural network. To further explore how to combine FCs and Bi-LSTM to establish the optimal source primary network, we designed the ablation experiment, as shown in Fig. 3. The backbone network is composed of three Bi-LSTM and FCs in series, and FC sections are added to the backbone network by seven ways. The calibration set of source spectra were used to train seven types of improved networks and the validation set and prediction set were used to test its effectiveness respectively to determine the optimal source primary model. The ablation experiment has been implemented on both manure dataset and  $\gamma$ -PGA dataset.

### 2.3. Fine-tuning approaches for transfer learning

Fine-tuning is the basic method of deep network transfer learning. It adjusts the parameters of the trained neural network with new data input which can reduce the distribution difference between source data and target data to adapt to new prediction tasks. In computer vision, there are evidence that the trained convolutional block (CNN) doesn't need retraining when used on new data [33,34]. Whereas, in spectra, re-

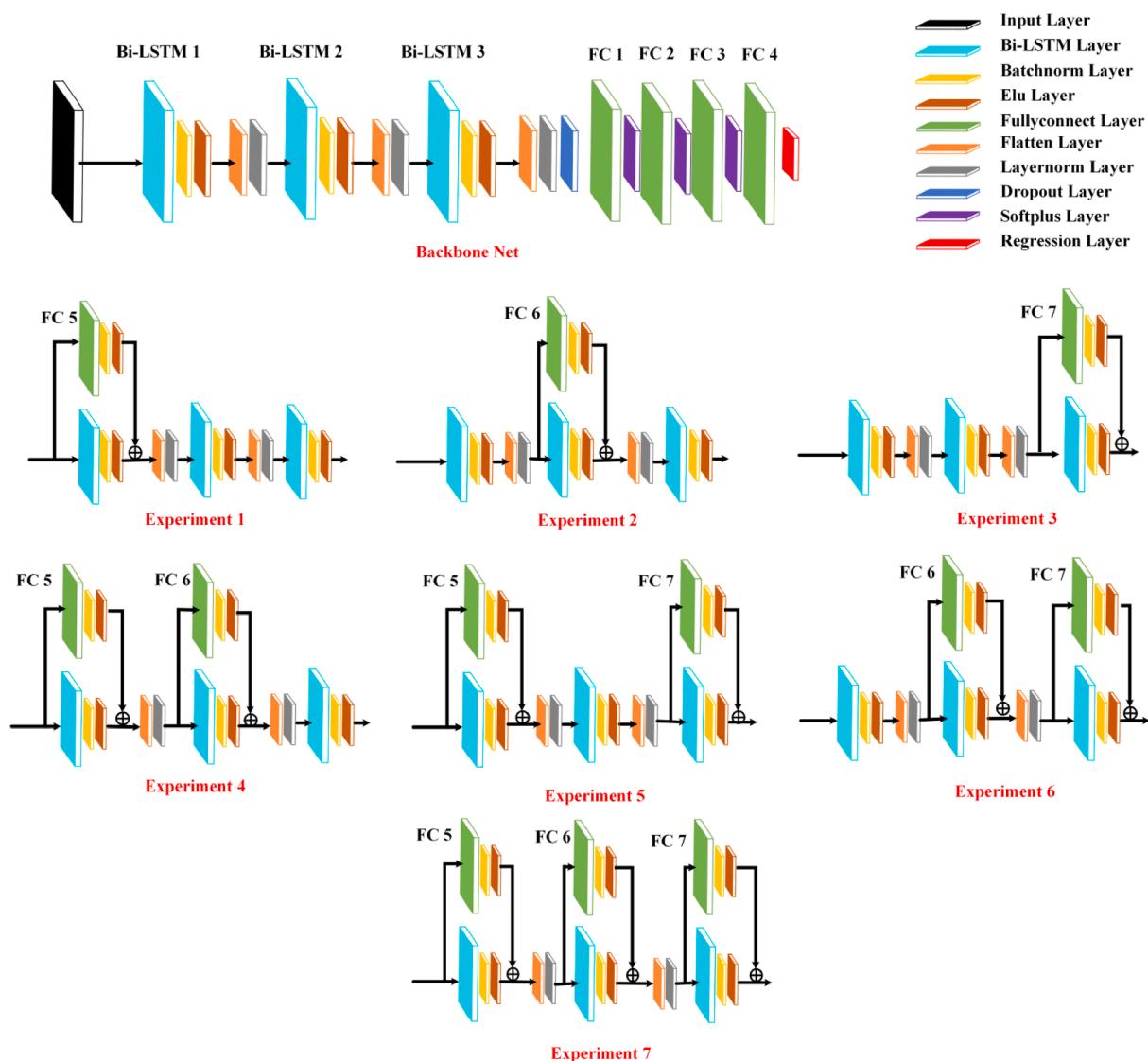


Fig. 3. The architecture of the backbone Bi-LSTM network and the ablation experiment design.

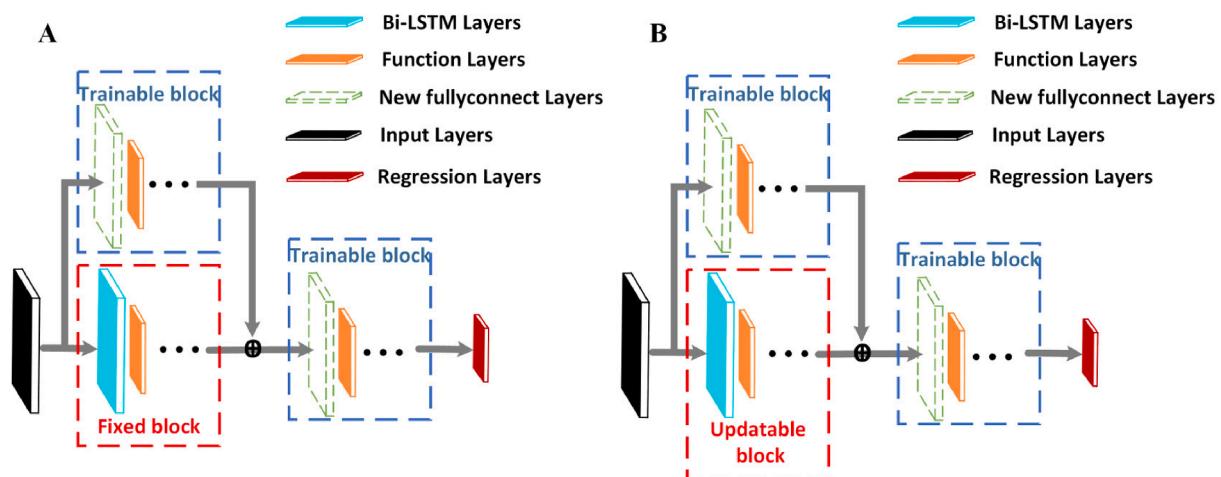


Fig. 4. The framework of two types of fine-tuning the improved Bi-LSTM network: **A.** Fixing the parameters of the feature extraction module; **B.** Updating the parameters of the feature extraction module and the FCs.

training the retrieval block might provide some extra insights. Herein, in this study, two approaches of fine-tuning the primary network were considered to verify the NIRS model transfer tasks of manure and  $\gamma$ -PGA cases, and further to determine which fine-tuning approach performs better in predicting the property values of the target area.

For the first fine-tuning network approach shown in Fig. 4A, the parameters of Bi-LSTM layers and function layers of the trained primary network are fixed while the FCs are initialized, and the parameters of the FCs are re-trained by the target dataset.

For the second fine-tuning network approach shown in Fig. 4B, the FCs of the trained primary model are initialized, and the parameters of Bi-LSTM layers, function layers and FCs are all updated by re-training the primary network with the target dataset. The predicting capacities of the transferred models based on these two fine-tuning approaches for the target NIR spectra will both be discussed.

#### 2.4. Network parameters settings and implementation

The hidden layer parameters of Bi-LSTM and FCs in Fig. 3 and Fig. 4 for manure case and  $\gamma$ -PGA case are both shown in Table 1. The maximum number of epochs, the batch size, the initial global learning rate, and the momentum parameter were set as 300, 256, 0.001 and 0.9, respectively, after each 30 epochs of training, the global learning rate will be decayed by 90 %.

To ensure the strong robustness of the source primary model and the transferred model, root mean square error (RMSE) and L2 norm regularization were used in the training process to minimize the square sum loss to prevent overfitting. The loss function is expressed by formula (9):

$$\text{Loss} = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - \hat{y}_n)^2} + \lambda \|\omega\| \quad (9)$$

Where  $y_n$  and  $\hat{y}_n$  are the true value and the predicted value respectively,  $N$  is the number of samples in the calibration set,  $\lambda$  is the regularization coefficient and  $\omega$  is the weight matrix.

The Adam optimizer was used to find the minimum value of the loss function during the training process, and the early stop condition was set as: when the loss value of the validation set is greater than or equal to the minimum value for 50 times, the iterative training of the network stops. Besides, the dropout mechanism was added, during the training of each epoch, the output of Bi-LSTM module at the end of network was randomly deactivated by 10 % to prevent overfitting.

All data preprocessing, neural network design, training processes and transfer learning simulation were implemented on Deep Network Designer toolkit of Matlab 2021b. The CPU is Intel (R) core (TM) i5-10400 CPU @ 2.90ghz, and the GPU is NVIDIA Geforce RTX 2060.

### 3. Results and discussion

#### 3.1. Spectral preprocessing

Fig. 5A shows the NIR spectra of cattle manure and poultry manure. It is obvious that there are spectral line shifts may arise from a range of

**Table 1**  
Hyper-parameters of training network settings for manure and  $\gamma$ -PGA cases.

| Hyper-parameters                | manure           | $\gamma$ -PGA    |
|---------------------------------|------------------|------------------|
| Hidden number of Bi-LSTM 1 to 3 | 1024, 512, 256   | 2048, 1024, 512  |
| Hidden number of FC 1 to 4      | 300, 300, 300, 1 | 300, 300, 300, 1 |
| Hidden number of FC 5 to 7      | 1024, 512, 256   | 2048, 1024, 512  |
| Max epochs                      | 300              | 300              |
| Batch size                      | 256              | 256              |
| Initial learning rate           | 0.001            | 0.001            |
| momentum parameter              | 0.9              | 0.9              |
| Learning rate decay             | 0.1              | 0.1              |
| Learning rate decay period      | 30               | 30               |
| Regularization coefficient      | 0.0001           | 0.0001           |

factors such as differences in spectrometer instruments, samples, and measurement conditions. Linear mapping method was first used to expand the source and target NIRS data by 10 times: ten slope seeds were set as 0.95, 0.96, 0.97, 0.98, 0.99, 1.01, 1.02, 1.03, 1.04, 1.05, respectively, the standard deviation of  $\text{data}_{\text{source}}$  and  $\text{data}_{\text{target}}$  were calculated as  $\text{std}_1$  and  $\text{std}_2$ . The linear transformation of spectral data in source area and target area under 10 slope seeds are based on formula (10) and (11):

$$\text{source}_{\text{amplify}}(i) = \text{seed}(i) * \text{data}_{\text{source}} + 0.5 * \text{std}_1, i \in [1, 10] \quad (10)$$

$$\text{target}_{\text{amplify}}(i) = \text{seed}(i) * \text{data}_{\text{target}} + 0.5 * \text{std}_2, i \in [1, 10] \quad (11)$$

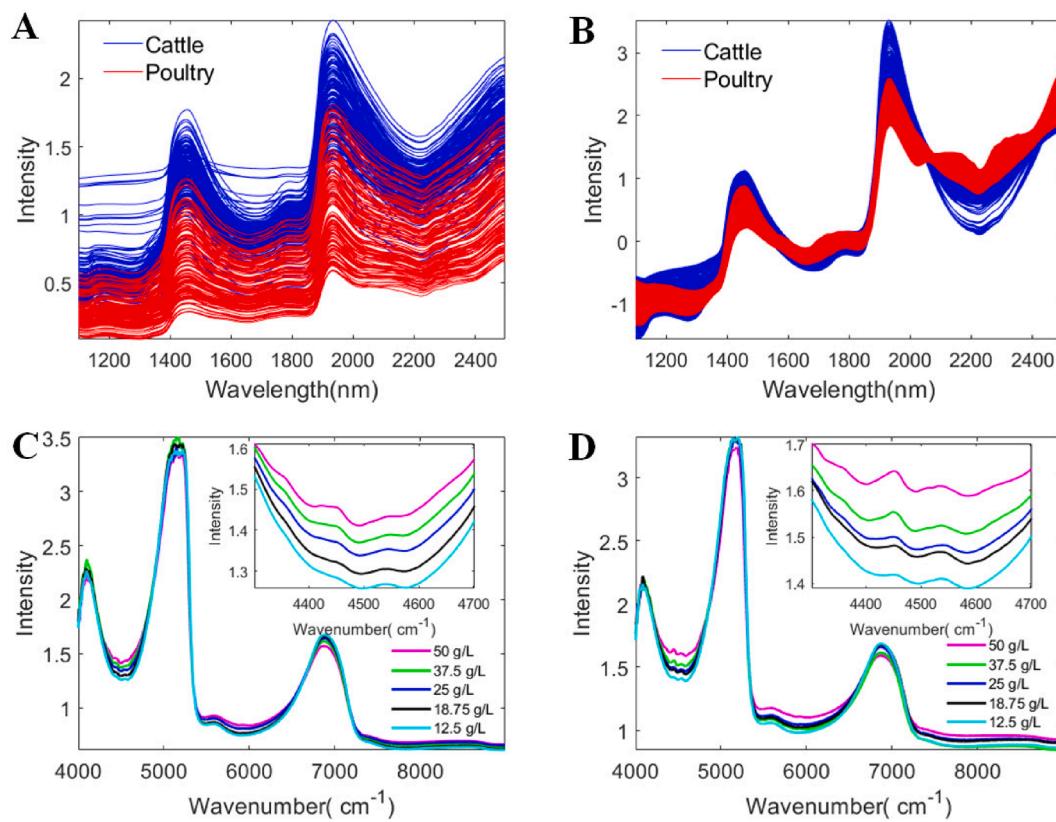
Standard normal variate transformation (SNV) was then carried out on all NIR spectra to eliminate the shifts, the spectra after SNV pre-treatment are shown in Fig. 5B. It is easy to find that the spectral baseline shifts are significantly reduced, and the NIR spectral characteristics of cattle and poultry manure are more prominent. Finally, the spectral data of cattle and poultry manure were divided into calibration set, validation set, and prediction set in a ratio of 3:1:1 by sample set partitioning based on joint x-y distances (SPXY) algorithm.

Fig. 5C and D shows the averaged NIR spectra of the source and target  $\gamma$ -PGA solutions of 5 different concentrations, respectively, the spectral range of 9000–4000 cm<sup>-1</sup> were selected. As can be seen from zoom-in views in the range of 4300 cm<sup>-1</sup> to 4700 cm<sup>-1</sup> that NIR absorption intensity decreases regularly with the decrease of  $\gamma$ -PGA concentrations. Again, NIR spectra of two batches were first expanded to 1960 and 1360 respectively by the same data amplification method of linear mapping. After that, variation mode decomposition (VMD) was performed and the mean NIR spectra of two batches after VMD pre-processing are show in Figure S1 and Figure S2 in Supporting Information(SI). It can be found that VMD algorithm had obvious effect on eliminating high-frequency noise in the range of 4000–4200 cm<sup>-1</sup> and 5050–5250 cm<sup>-1</sup>, and the shape and position of the characteristic absorption peaks of  $\gamma$ -PGA in the spectral range of 4300 cm<sup>-1</sup>–4700 cm<sup>-1</sup> are well preserved. In the end, SPXY algorithm was used to divide the spectra of two batches into calibration set, verification set, and prediction set according to the ratio of 3:1:1, respectively.

#### 3.2. Primary Bi-LSTM networks for cattle manure and the first batch of $\gamma$ -PGA solution

Based on the ablation experiment designed in Fig. 3, quantitative models of DM content, and the  $\gamma$ -PGA concentration based on backbone network and seven improved Bi-LSTM networks were trained with the calibration set of cattle manure and the first batch of  $\gamma$ -PGA, respectively. For manure case, the loss curves of the backbone network and each improved network in the ablation experiment are shown in Fig. 6. As can be seen from Fig. 6, both the loss curves of calibration set (red) and validation set (black) oscillated and converged, and the interval between the red and black loss curves is very small, indicating that the models have strong generalization ability, and the networks are not over-fitting. In addition, the positions indicated by blue arrows in the zoom-in views represent the validation loss values reached the minimum, namely the training met the early stopping condition and would stop after the network continued to train for 50 epochs. For  $\gamma$ -PGA case, the validation loss curves are shown in Figure S3 in SI.

The source calibration, validation and prediction sets, and the target prediction set of two datasets were used to verify the performance of the acquired source models in ablation experiment, respectively. The predicting evaluation index of two datasets are both presented in Table 2. For manure case, it can be seen from Table 2 that each model has good robustness since the accuracies of the validation and prediction sets are all very close. Furthermore, the  $\text{RMSEP}_{\text{source}}$  value of backbone network is 0.6833, and after adding FC sections fusion with backbone network by seven ways, the  $\text{RMSEP}_{\text{source}}$  values of six improved networks have all been reduced except the second fusion method, which proves that the combination of FCs and Bi-LSTM can improve the quantitative



**Fig. 5.** Original and preprocessed NIR spectra: A. Original spectra of cattle and poultry manure; B. Spectra of cattle and poultry manure after preprocessing; C. Average original spectra of source  $\gamma$ -PGA solution (the first batch); D. Average original spectra of target  $\gamma$ -PGA solution (the second batch).

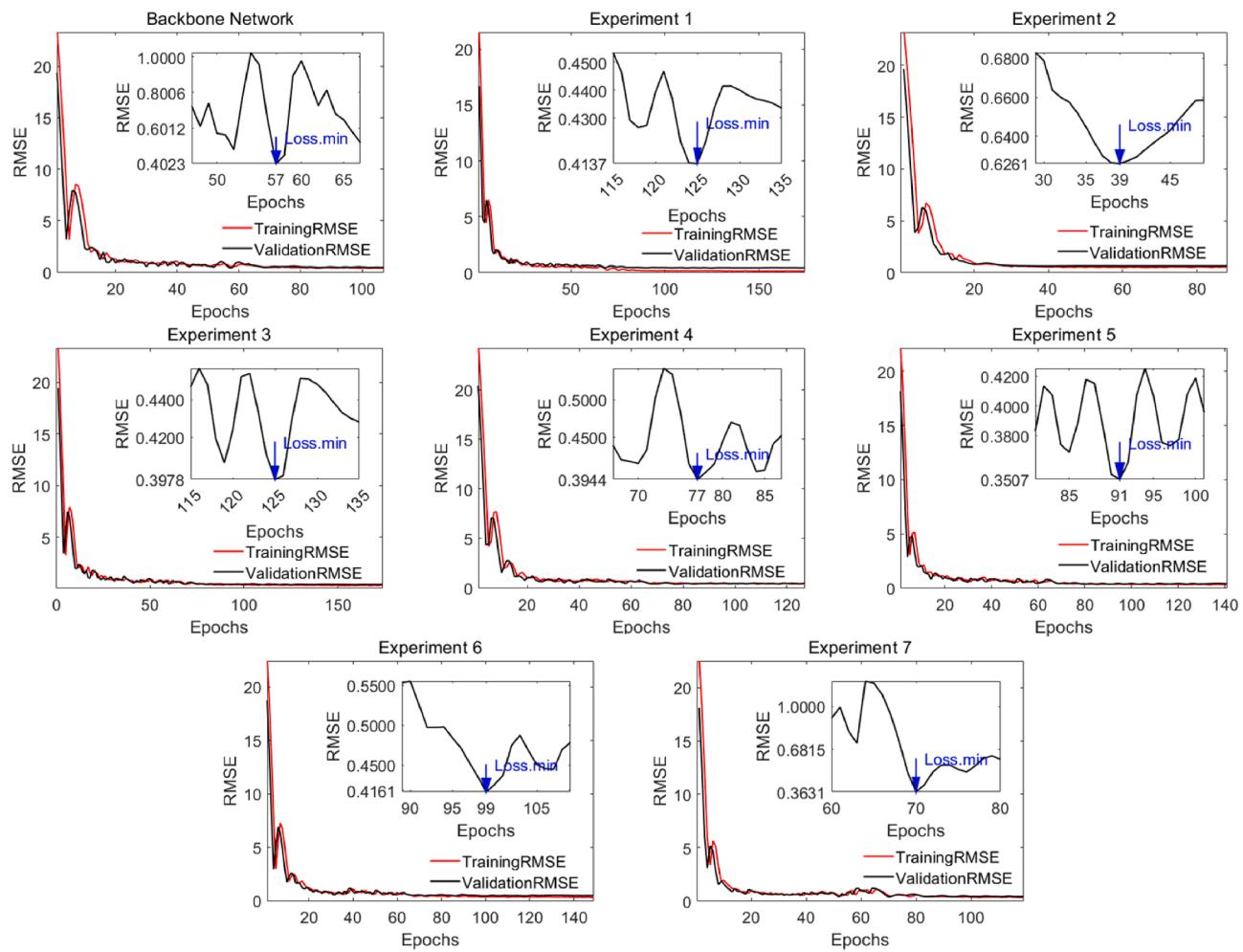
prediction accuracy. Among them, the seventh improved Bi-LSTM network has achieved the most excellent DM content quantification results, which the  $RMSEP_{source}$  value is as low as 0.4276. Therefore, the DM content quantitative model constructed based on the seventh improved network was the optimal primary model. However, for the target prediction set (poultry manure, number of 272), the primary model without re-calibration performed very poorly: the  $R^2_{Ptarget}$  value is only 0.8612, and  $RMSEP_{target}$  value is high to 31.4137 (Table 2), as also can be demonstrated from Fig. 7, the coordinate points composed by the predicted and the true values of the source prediction set (cross in blue) are densely distributed on the  $Y = X$  line, while the coordinate points of the target prediction set (circled in red) are distributed in the lower right area far away from the  $Y = X$  line. Furthermore, the true values of DM content of poultry manure range from 27.0700 to 82.4800 %, while the predicted values range was from 16.8747 to 32.4502 %, which are still within the DM content range of cattle manure (11.2550 to 46.6120 %).

And likewise, for  $\gamma$ -PGA case, as can be seen from the evaluation index shown in Table 2 and the scatter plots shown in Fig. 8, it can also be concluded that the network models in ablation experiment are accurate and stable for predicting source spectra (the first batch product). Among which, the  $RMSEP_{source}$  values of the first and the fourth improved networks are 0.1929 and 0.1926 respectively, which are extremely close, and the first network structure is more concise. Therefore,  $\gamma$ -PGA concentration quantitative model constructed based on the first improved network model was the optimal primary model. Similarly, the source primary model does not match the target spectra (the second batch product) as indicated by the  $R^2_{Ptarget}$  and  $RMSEP_{target}$  values shown in Table 2 and scatter plots of the target prediction set (circled in red) shown in Fig. 8. Therefore, it is still unable to overcome the spectral characteristics differences between the two batches of products, even though the differences appear to be small as shown in Fig. 5C and D. The results prove that the established primary Bi-LSTM

networks cannot be directly used for predicting the target spectra and confirm the requirement for the model update. Accordingly, we next conducted transfer learning for primary networks based on the two types of fine-tuning approaches described in section 2.3.

### 3.3. Transfer learning for poultry manure and the second batch of $\gamma$ -PGA solution

Herein, we used target calibration set of two datasets to re-train the optimal primary networks based on the two types of fine-tuning approaches shown in Fig. 4, respectively. The hyper-parameters of training the transferring network were consistent with those of training the primary network, the loss value curves of re-training network for manure case and  $\gamma$ -PGA case are shown in Figure S4 and Figure S5 in SI, respectively. the curves of calibration and validation converged, and the gap was small, indicating that the re-training has not been over-fitted. Table 3 gives the prediction accuracy indicators of the target calibration and prediction set on the transferred models obtained by two fine-tuning approaches for two cases. Based on fine-tuning approach 1, the  $R^2_{Ptarget}$  and  $RMSEP_{target}$  values of the target prediction set of poultry manure are 0.9943 and 1.0906, values of the second batch of  $\gamma$ -PGA are 0.9990 and 0.4494, respectively. Comparing with those obtained on the primary model shown in Table 2 of 0.8612 and 31.4137, 0.8853 and 8.9876, respectively, the determination coefficients both have been greatly improved and the root mean square errors have been significantly reduced. The results show that Bi-LSTM layers stored the extracted source spectral features and were no longer updated, only the parameters of the initialized FCs layers were re-trained by the target calibration set, and the non-linear relationship between the spectra and the property was adaptatively adjusted according to the characteristics of the target spectra, which can prove that FCs act as a “firewall” in the process of transfer learning and ensure the effectiveness and accuracy of



**Fig. 6.** Loss curves of source improved Bi-LSTM networks of DM content in ablation experiment.

**Table 2**

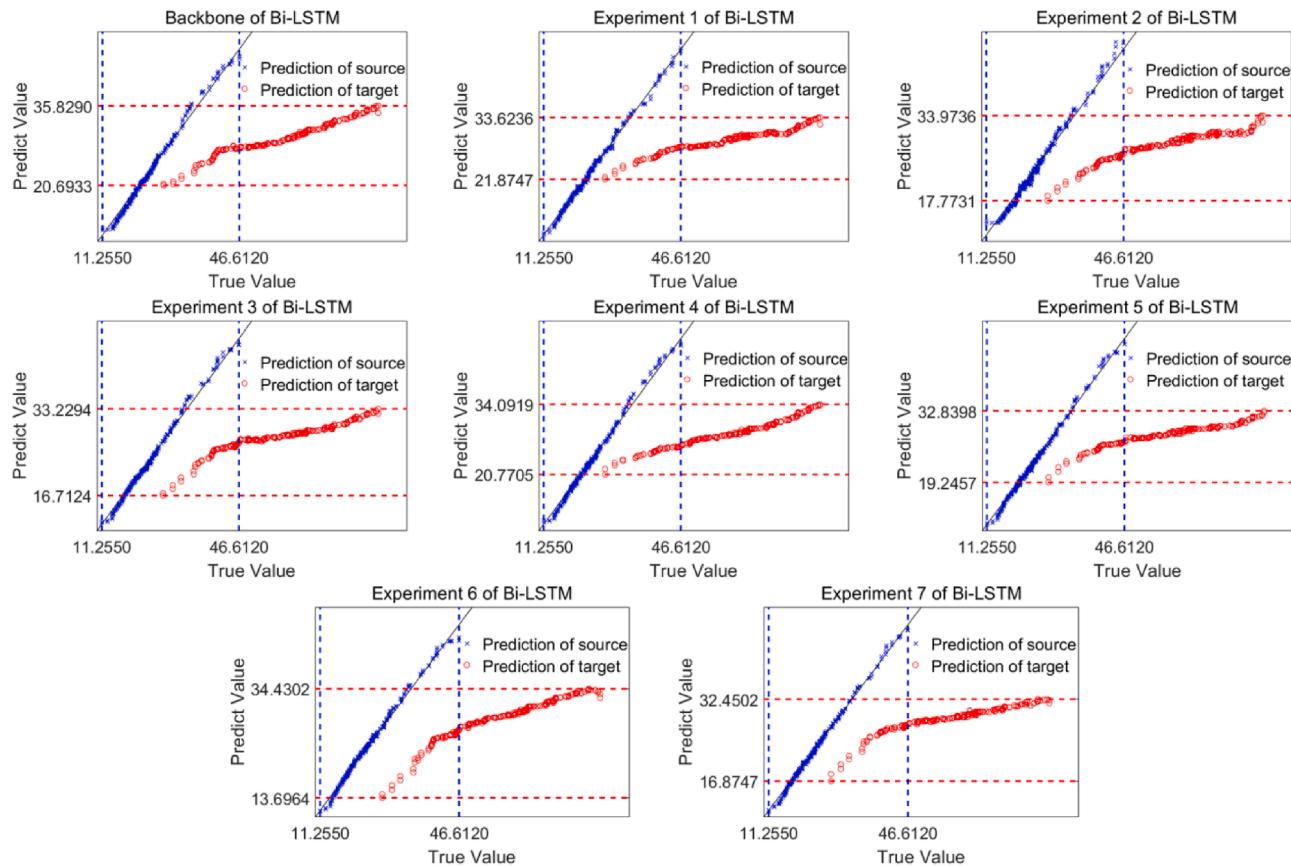
Evaluation index based on the Bi-LSTM models in ablation experiment.

| Sample        | experiment | $R^2_{Csource}$ | $RMSEC_{source}$ | $R^2_{Vsource}$ | $RMSEV_{source}$ | $R^2_{Psource}$ | $RMSEP_{source}$ | $R^2_{Ptarget}$ | $RMSEP_{target}$ |
|---------------|------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|
| Manure        | Backbone   | 0.9994          | 0.3776           | 0.9956          | 0.6813           | 0.9956          | 0.6833           | 0.9602          | 30.2639          |
|               | 1          | 0.9999          | 0.0511           | 0.9958          | 0.4626           | 0.9985          | 0.4661           | 0.9435          | 31.1956          |
|               | 2          | 0.9977          | 0.4113           | 0.9932          | 0.8253           | 0.9928          | 0.8443           | 0.8971          | 31.7568          |
|               | 3          | 0.9998          | 0.2615           | 0.9962          | 0.5701           | 0.9963          | 0.5769           | 0.8833          | 32.1189          |
|               | 4          | 0.9994          | 0.4307           | 0.9966          | 0.5830           | 0.9966          | 0.5836           | 0.9751          | 31.7811          |
|               | 5          | 0.9998          | 0.2951           | 0.9972          | 0.5273           | 0.9972          | 0.5412           | 0.9169          | 31.9693          |
|               | 6          | 0.9997          | 0.2492           | 0.9954          | 0.4740           | 0.9953          | 0.4688           | 0.9116          | 30.8511          |
| $\gamma$ -PGA | 7          | 0.9997          | 0.2243           | <b>0.9976</b>   | <b>0.4201</b>    | <b>0.9976</b>   | <b>0.4276</b>    | <b>0.8612</b>   | <b>31.4137</b>   |
|               | Backbone   | 1               | 0.0333           | 0.9997          | 0.3844           | 0.9996          | 0.3857           | 0.9182          | 6.9003           |
|               | 1          | 1               | 0.0216           | <b>1</b>        | <b>0.1003</b>    | <b>0.9998</b>   | <b>0.1929</b>    | <b>0.8853</b>   | <b>8.9876</b>    |
|               | 2          | 1               | 0.0506           | 0.9998          | 0.2979           | 0.9997          | 0.2998           | 0.9144          | 7.6822           |
|               | 3          | 1               | 0.0363           | 0.9998          | 0.2271           | 0.9993          | 0.3764           | 0.9213          | 8.0601           |
|               | 4          | 1               | 0.0210           | 0.9999          | 0.1470           | 0.9998          | 0.1926           | 0.9235          | 7.3297           |
|               | 5          | 1               | 0.0510           | 0.9998          | 0.1886           | 0.9994          | 0.3306           | 0.9167          | 8.0953           |
|               | 6          | 1               | 0.0966           | 0.9998          | 0.2443           | 0.9990          | 0.4363           | 0.9682          | 5.6035           |
|               | 7          | 1               | 0.0745           | 0.9997          | 0.2362           | 0.9993          | 0.3785           | 0.9599          | 7.0889           |

the migration network.

More importantly, as shown in Table 3, a satisfactory predicting results has been achieved by performing fine-tuning approach 2. Specifically, the  $R^2_{Ptarget}$  value for poultry manure was further increased from 0.9943 to 0.9967 and the  $RMSEP_{target}$  value was further decreased from 1.0906 to 0.8351, a reduction of  $RMSEP_{target}$  by 23.4275 %; and the  $R^2_{Ptarget}$  value for the second batch of  $\gamma$ -PGA was further improved from 0.9990 to 0.9999, and  $RMSEP_{target}$  value was further decreased from

0.4494 to 0.2214, an predicting error reduction by 50.7343 %. It is also easy to conclude from the scatter plots presented in Fig. 9A for poultry manure and B for the second batch of  $\gamma$ -PGA. Obviously, for poultry manure, the points formed by the predicted DM content values and the true values has notably shifted from scattering in the far lower area of the Y = X line (circled in red in Fig. 7) to almost all points were closely distributed on the Y = X line (Fig. 9A), and the points composed by predicted  $\gamma$ -PGA concentration values and the true values has been



**Fig. 7.** Scatter plots of the DM content of source and target on the Bi-LSTM models in ablation experiments.

corrected from scattering in the upper left corner of  $Y = X$  line (red circled in Fig. 8) to densely concentrated on  $Y = X$  line (Fig. 9B).

The results in this section illustrate that the second transfer approach by updating the weights of Bi-LSTM layers with the target calibration set and then re-training the initialized FCs layers performed even better than the first approach of fixing Bi-LSTM layers and only re-training FC layers, which is beneficial for reducing the feature differences between source and target spectra and fully learn the underlying characteristics of the target spectra. At the same time, FCs layers also ensure the transferability of the network, therefore, the prediction performance of the migration model can be remarkably improved.

### 3.4. Comparisons with one dimensional convolutional neural network

1D-CNN, has recently been applied in the field of near-infrared technology. The essential difference between 1D-CNN and LSTM is that 1D-CNN has no memory function and can only recognize part of spectral features, while LSTM can correlate the forward and backward spectral features. In this section, we aim to explore the superiority of LSTM network compared to 1D-CNN network for NIRS quantification and transferring for new scenarios.

#### 3.4.1. Comparisons of source primary models

For FCs and CNN fusion, the FC ablation experiment was also carried out according to Fig. 3 design. To control variables, the number of filters in the convolutional layer was set as the same with the number of hidden units in Bi-LSTM, which ensured that the network structure and number of parameters of 1D-CNN and Bi-LSTM were consistent. In addition, the same hyper-parameters and anti-overfitting mechanism were set according to section 2.4, and the loss curves of CNN backbone network and improved networks were shown in Figure.S6 and Figure.S7 in SI.

All 1D-CNN models were tested by source and target validation set

and prediction set, and the evaluation index of CNN models in the ablation experiment were recorded in Table 4. For manure case, it can be found in Table 4 that the  $RMSEP_{source}$  value of CNN backbone network is 1.6121, while the  $RMSEP_{source}$  values of all improved models decrease after adding FC sections, which proves that fusion of FC and feature extraction layer (Bi-LSTM or 1D-CNN) does enhance the learning ability of the network and improve the prediction accuracy. Among them, the quantitative model based on the seventh improved CNN structure has achieved the most accurate prediction with  $RMSEP_{source}$  value of 1.3164, which is 18.3425 % lower than that of the backbone network. However, this prediction result is still far inferior to Bi-LSTM network with a  $RMSEP_{source}$  value of 0.4276, which is 67.5175 % lower. For  $\gamma$ -PGA case, it can also be found in Table 4 that the seventh experiment has obtained the optimal model with the lowest  $RMSEP_{source}$  value of 0.7975. However, the  $RMSEP_{source}$  value of the optimal Bi-LSTM network is 0.1926, which is 75.8495 % lower than that of the CNN network.

Moreover, Fig. 10 A and B present the scatter plots of DM content and  $\gamma$ -PGA concentration based on the optimal improved CNN network, respectively. When compared with the scatter plots of the source prediction set (cross in blue) based on the improved Bi-LSTM network in Fig. 7, it is obvious that most points in Fig. 10 are relatively scattered on both sides of  $Y = X$  line, which also proved the superiority of Bi-LSTM network. Scatter plots of DM content and  $\gamma$ -PGA concentration based on other CNN networks of ablation experiment are shown in Figure.S8 and Figure.S9 in SI, respectively.

#### 3.4.2. Comparisons of fine-tuned transferred models

In this section, the calibration sets of two datasets were used to retrain the best improved CNN primary network model respectively according to the two fine-tuning approaches described in Fig. 4. The loss curves of the migrating neural network are shown in Figure.S8 and Figure.S9. Each transferred 1D-CNN model was tested with target

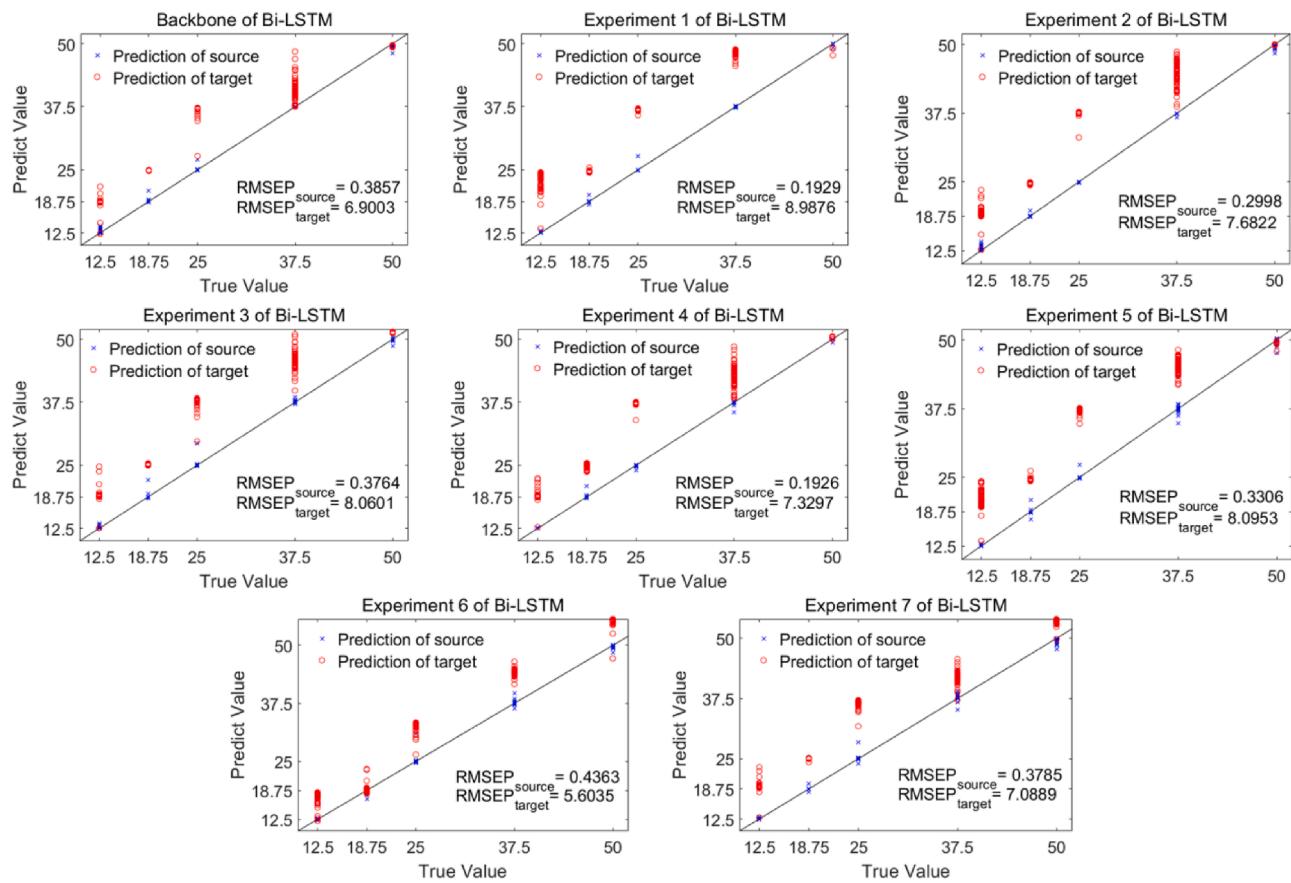


Fig. 8. Scatter plots of source and target  $\gamma$ -PGA concentrations on primary models in ablation experiments.

Table 3

Prediction accuracy on the transferred models obtained by two fine-tuning approaches.

| Sample | approach               | $R^2_{Ctarget}$ | $RMSEC_{target}$ | $R^2_{Ptarget}$ | $RMSEP_{target}$ |
|--------|------------------------|-----------------|------------------|-----------------|------------------|
| Manure | Fine-tuning approach 1 | 0.9972          | 0.7384           | 0.9943          | 1.0906           |
|        | Fine-tuning approach 2 | 0.9991          | 0.5098           | 0.9967          | 0.8351           |
|        | Fine-tuning approach 1 | 0.9999          | 0.1093           | 0.9990          | 0.4494           |
|        | Fine-tuning approach 2 | 1               | 0.0747           | 0.9999          | 0.2214           |

spectra, and the evaluation index of the transferred model were recorded in Table 5. From the comparisons with the results shown in Table 4, it is not difficult to find that two type of fine-tuning 1D-CNN approaches have both repaired the prediction of DM content and  $\gamma$ -PGA concentration for target spectra, which reduced the  $RMSEP_{target}$  values from 30.8854 and 7.0506 (Table 4) to 1.2536, 0.9007 and 0.4536, 0.4268, respectively. The results prove once again that FC acts as a “firewall” for neural network migration to enable model transfer between samples. Moreover, obviously, the accuracy of the transferred model based on fine-tuning approach 2 is better than that based on approach 1.

Furthermore, the  $RMSEP_{target}$  values obtained by fine-tuning Bi-LSTM network were 0.8351 and 0.2214 for manure and  $\gamma$ -PGA respectively (in Table 3), which reduced by 7.2832 % and 48.1256 % compared to 0.9007 and 0.4268 that obtained by fine-tuning 1D-CNN network. The comparisons results illustrate that the transferability of Bi-LSTM

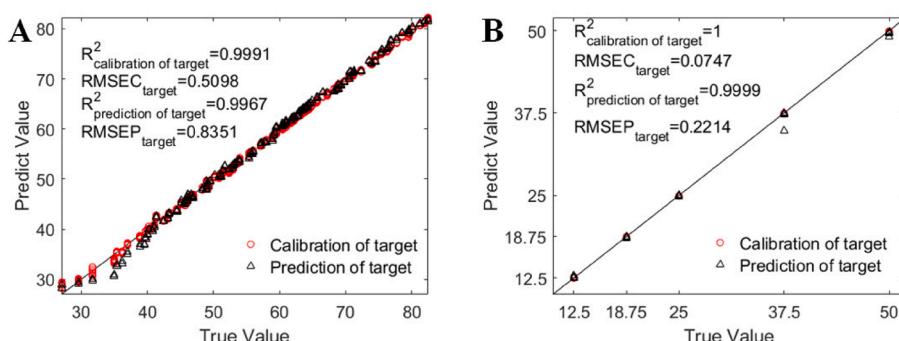
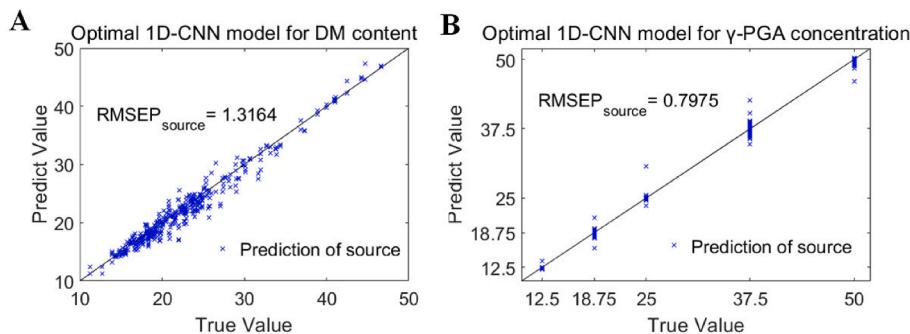


Fig. 9. Scatter plots of the transferred models based on fine-tuning approach 2 for two cases: A. DM content of poultry manure.; B.  $\gamma$ -PGA concentrations of the second batch.

**Table 4**

Evaluation index based on the primary 1D-CNN models in ablation experiment.

| Sample | experiment | $R^2_{\text{source}}$ | $RMSE_{\text{source}}$ | $R^2_{\text{source}}$ | $RMSEV_{\text{source}}$ | $R^2_{\text{source}}$ | $RMSE_{\text{source}}$ | $R^2_{\text{target}}$ | $RMSE_{\text{target}}$ |
|--------|------------|-----------------------|------------------------|-----------------------|-------------------------|-----------------------|------------------------|-----------------------|------------------------|
| Manure | Backbone   | 0.9645                | 1.2096                 | 0.9453                | 1.6141                  | 0.9458                | 1.6121                 | 0.9602                | 29.4879                |
|        | 1          | 0.9669                | 1.1759                 | 0.9527                | 1.5193                  | 0.9528                | 1.5204                 | 0.9238                | 29.0739                |
|        | 2          | 0.9666                | 1.2108                 | 0.9508                | 1.5324                  | 0.9510                | 1.5381                 | 0.9239                | 30.9670                |
|        | 3          | 0.9649                | 1.2133                 | 0.9519                | 1.5228                  | 0.9519                | 1.5228                 | 0.9151                | 19.3664                |
|        | 4          | 0.9692                | 1.1584                 | 0.9580                | 1.3728                  | 0.9579                | 1.3808                 | 0.9025                | 31.4241                |
|        | 5          | 0.9633                | 1.2490                 | 0.9647                | 1.3222                  | 0.9646                | 1.3269                 | 0.9223                | 31.2763                |
|        | 6          | 0.9654                | 1.1978                 | 0.9591                | 1.3921                  | 0.9595                | 1.3910                 | 0.8943                | 31.4580                |
|        | 7          | 0.9671                | 1.1954                 | 0.9630                | 1.3065                  | 0.9628                | 1.3164                 | 0.9162                | 30.8854                |
|        | Backbone   | 0.9989                | 0.5122                 | 0.9969                | 0.7644                  | 0.9955                | 0.9037                 | 0.9664                | 9.5569                 |
|        | 1          | 0.9992                | 0.4471                 | 0.9976                | 0.6801                  | 0.9962                | 0.8570                 | 0.9942                | 7.6934                 |
|        | 2          | 0.9991                | 0.4817                 | 0.9973                | 0.7181                  | 0.9961                | 0.8803                 | 0.9882                | 7.6892                 |
|        | 3          | 0.9991                | 0.4744                 | 0.9973                | 0.7127                  | 0.9958                | 0.8851                 | 0.9815                | 7.6123                 |
|        | 4          | 0.9993                | 0.3925                 | 0.9976                | 0.6689                  | 0.9965                | 0.8136                 | 0.9906                | 7.5753                 |
|        | 5          | 0.9992                | 0.4313                 | 0.9973                | 0.6960                  | 0.9961                | 0.8513                 | 0.9805                | 8.3294                 |
|        | 6          | 0.9993                | 0.4477                 | 0.9975                | 0.7080                  | 0.9962                | 0.8796                 | 0.9919                | 7.6049                 |
|        | 7          | 0.9994                | 0.3937                 | 0.9979                | 0.6382                  | 0.9968                | 0.7975                 | 0.9936                | 7.0506                 |

**Fig. 10.** The scatter plots of DM content and  $\gamma$ -PGA concentration based on the optimal improved CNN network.**Table 5**

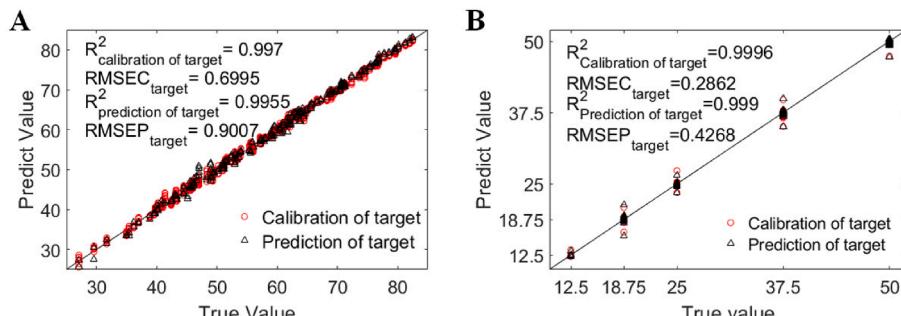
Prediction accuracy on the transferred 1D-CNN models obtained by two fine-tuning approaches.

| Sample        | approach               | $R^2_{\text{target}}$ | $RMSE_{\text{target}}$ | $R^2_{\text{target}}$ | $RMSE_{\text{target}}$ |
|---------------|------------------------|-----------------------|------------------------|-----------------------|------------------------|
| Manure        | Fine-tuning approach 1 | 0.9921                | 1.1486                 | 0.9921                | 1.2536                 |
|               | Fine-tuning approach 2 | 0.9970                | 0.6995                 | 0.9955                | <b>0.9007</b>          |
| $\gamma$ -PGA | Fine-tuning approach 1 | 0.9996                | 0.2965                 | 0.9989                | 0.4536                 |
|               | Fine-tuning approach 2 | 0.9996                | 0.2862                 | 0.9990                | <b>0.4268</b>          |

network is superior to 1D-CNN network, which can also be identified by the scatter plots shown in Fig. 9 and Fig. 11, it is obvious that more scatter points are located farther from the  $Y = X$  line in Fig. 11. The result can be explained that the prediction accuracy of the primary model based on Bi-LSTM was better than that based on 1D-CNN, namely the transfer can be adjusted at a higher “starting point”.

### 3.5. Comparisons with traditional model transfer methods

To further evaluate the superiority of the proposed NIRS model transfer based on the deep neural network of Bi-LSTM and FCs, we constructed Partial Least Square (PLS) and Extreme Learning Machine (ELM) primary models and implemented primary model transfer based on spectra-based migration methods (PDS and CCA), result-based migration method (SBC), feature-based migration method (TCA), and samples-based migration method (Tradaboost), respectively. Each

**Fig. 11.** Scatter plots of the transferred 1D-CNN models based on fine-tuning approach 2 for two cases: A. DM content of poultry manure.; B.  $\gamma$ -PGA concentrations of the second batch.

algorithm was performed both on the manure dataset and the  $\gamma$ -PGA dataset.

### 3.5.1. Comparisons of different NIR primary models

The primary models for cattle manure and the first batch of  $\gamma$ -PGA solutions based on PLS, ELM were built to compare with the improved Bi-LSTM network. For the prediction set of both the source and target datasets of the manure and  $\gamma$ -PGA samples, the detailed values of the determination coefficient  $R_p^2$  and the root mean square error  $RMSEP$  based on the primary models are compared in Table 6. Moreover, Fig. 12 A to C respectively presents the scatter plot for the prediction set of the cattle manure built by PLS, ELM and Bi-LSTM, whilst Fig. 12 D to F show the scatter plots for the prediction set of the first batch of  $\gamma$ -PGA built by the above three algorithms. The following conclusions can be derived:

- (1) The coordinate points formed by the predicted and the true value in Fig. 12 C (DM content) and in F ( $\gamma$ -PGA concentration), which were both obtained by the proposed Bi-LSTM network, all were uniformly concentrated on the  $Y = X$  line, while the points in A and B (DM content built by PLS and ELM), D and E ( $\gamma$ -PGA concentration built by PLS and ELM) are relatively discrete. Besides, as can be derived from Table 6, for the prediction set of cattle manure, the proposed Bi-LSTM network increased the  $R_{source}^2$  by 12.4944 % and 25.8642 % when compared with PLS and ELM, and decreased the  $RMSEP_{source}$  by 80.1973 % and 85.3617 %, respectively; for the prediction set of the first batch of  $\gamma$ -PGA solutions, the proposed Bi-LSTM increased the  $R_{source}^2$  by 1.4202 % and 1.8645 % when compared with PLS and ELM, and decreased the  $RMSEP_{source}$  by 88.0720 % and 89.5243 %, respectively. Deep learning based on Bi-LSTM network for NIR quantification has greatly outperformed classical PLS and ELM regression algorithms.
- (2) As can be derived from Table 6, based on the three primary models built by PLS, ELM and Bi-LSTM, for the prediction set of the target data of two cases, the  $RMSEP_{target}$  values of poultry manure were 8.9465, 16.1508 and 31.4137, respectively, which were 3.1432, 4.5290 and 72.4652 times larger than  $RMSEP_{source}$  of cattle manure; The  $RMSEP_{target}$  values of the second batch of  $\gamma$ -PGA were 9.4821, 6.5875, 8.9876, respectively, which were

**Table 6**

Evaluation index of the primary models built by PLS, ELM and improved Bi-LSTM for manure and  $\gamma$ -PGA.

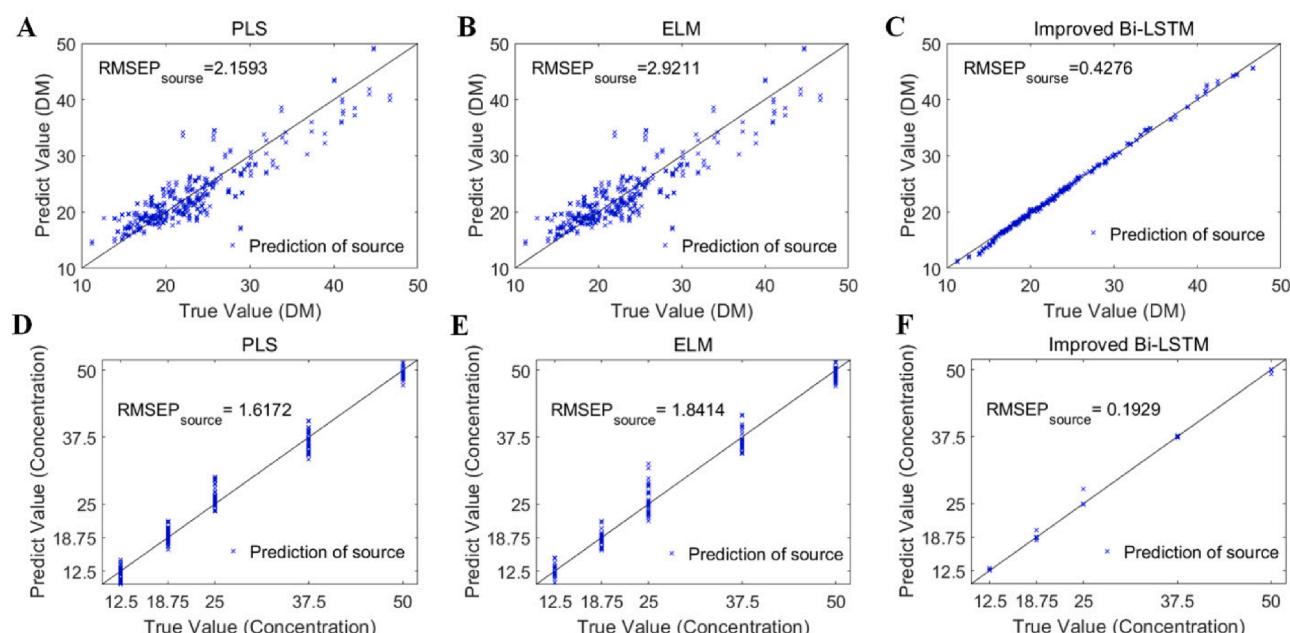
| Samples       | algorithm        | $R_{source}^2$ | $RMSEP_{source}$ | $R_{target}^2$ | $RMSEP_{target}$ |
|---------------|------------------|----------------|------------------|----------------|------------------|
| Manure        | PLS              | 0.8868         | <b>2.1593</b>    | 0.7967         | <b>8.9465</b>    |
|               | ELM              | 0.7926         | <b>2.9211</b>    | 0.5564         | <b>16.1508</b>   |
|               | Improved Bi-LSTM | 0.9976         | <b>0.4276</b>    | 0.8612         | <b>31.4137</b>   |
| $\gamma$ -PGA | PLS              | 0.9858         | <b>1.6172</b>    | 0.9880         | <b>9.4821</b>    |
|               | ELM              | 0.9815         | <b>1.8414</b>    | 0.9589         | <b>6.5875</b>    |
|               | Improved Bi-LSTM | 0.9998         | <b>0.1929</b>    | 0.8853         | <b>8.9876</b>    |

4.8633, 2.5774 and 45.5920 times larger than  $RMSEP_{source}$  of the first batch. The results illustrate that when a quantitative NIRS model is used for a new scenario, no matter it is built by a machine learning or by a deep learning algorithm, model transfer or transfer learning is inevitable to be implemented.

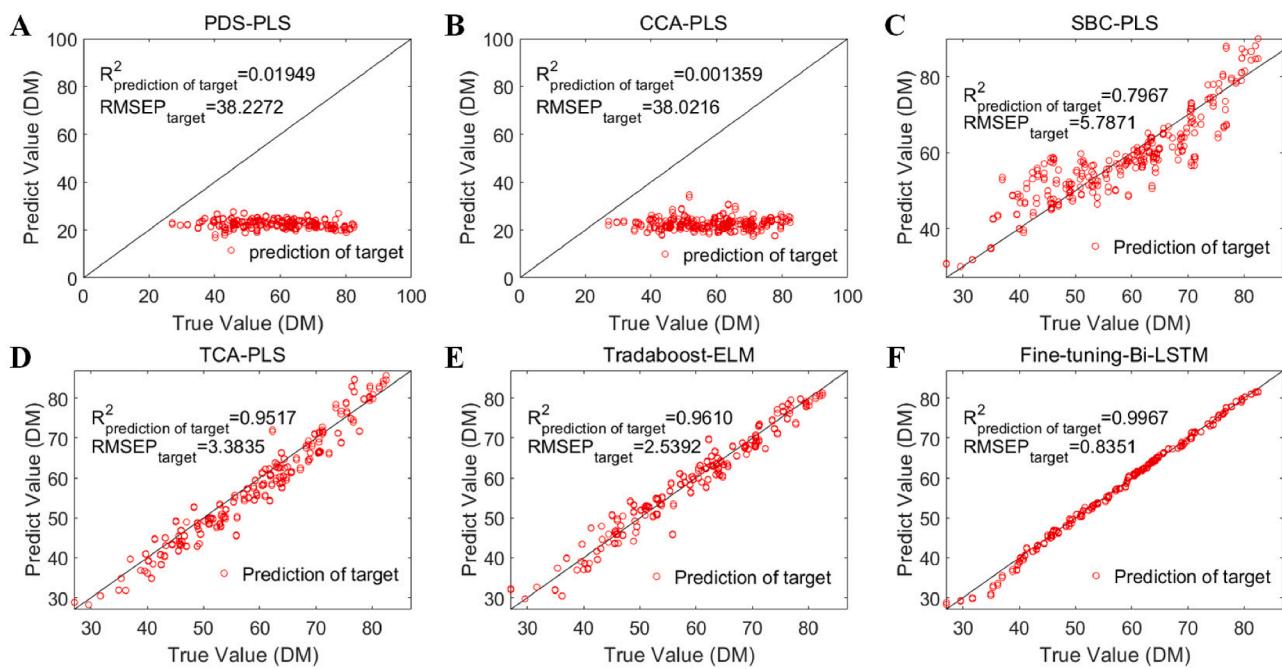
### 3.5.2. Comparisons of traditional model transfer methods and transfer learning

According to the conclusions in sections 3.3, the transferred models for the two cases (manure and  $\gamma$ -PGA) based on fine-tuning approach 2 were both more superior to approach 1, therefore, we only carry out comparisons between fine-tuning approach 2 with traditional model transfer methods.

Firstly, the transferred models for poultry manure based on PDS-PLS, CCA-PLS, SBC-PLS, TCA-PLS, Tradaboost-ELM and Fine-tuning-Bi-LSTM algorithms are compared. The detailed values of  $R_{target}^2$  and  $RMSEP_{target}$  of six migration models are shown in Table 7. Interestingly, it can be found that both PDS and CCA transferred the primary PLS model to a negative migration prediction for the DM content in poultry manure, because the  $RMSEP_{target}$  values are as high as 38.2272 and 38.0216, respectively, which are both even much higher than 8.9465 that obtained by the primary PLS model without recalibration. We can also see it from the scatter plots in Fig. 13 A and B that the predicted and true values hardly exhibit linear distribution, indicating that PDS and CCA both have no correction effect on the PLS model in this context. It can be explained that the source nonlinear polynomial coefficients fitted by PLS



**Fig. 12.** Comparisons of scatter plots based on different primary models: A. PLS of cattle manure; B. ELM of cattle manure; C. Improved Bi-LSTM of cattle manure; D. PLS of  $\gamma$ -PGA; E. ELM of  $\gamma$ -PGA; F. Improved Bi-LSTM of  $\gamma$ -PGA.



**Fig. 13.** Scatter plots of six different transferred models for DM contents of poultry manure: A. PDS-PLS; B. CCA-PDS; C. SBC-PLS; D. TCA-PLS; E. Tradaboost-ELM l; F. fine-tuning-Bi-LSTM.

algorithm is determined by the NIR spectra of cattle manure and the DM content, although PDS and CCA can narrow the characteristic differences by calculating the transfer matrix between the NIR spectra of cattle and poultry manure, however, the spectra-based transfer algorithms cannot update and adjust the polynomial coefficients of the primary PLS model, which led to larger errors and negative migration prediction.

On the other hand, as can be derived from Table 7, among the other 3 traditional methods, the Tradaboost-ELM obtained the highest prediction accuracy with the  $R^2_{\text{target}}$  value of 0.9610 and the  $\text{RMSEP}_{\text{target}}$  value of 2.5392, which can be interpreted that Tradaboost can increase the data weights in the target area that are helpful to the DM content regression task and decrease the data weights that are useless by iterative optimization continuously. But the proposed fine-tuning-Bi-LSTM in this study transcended Tradaboost-ELM with  $R^2_{\text{target}}$  value of 0.9967, which increased by 3.7149 %, and  $\text{RMSEP}_{\text{target}}$  value of 0.8351, which decreased by 67.1117 %. Besides, it can also be demonstrated by the points of the scatter plot in Fig. 13 F (fine-tuning-Bi-LSTM) which has a much higher degree of aggregating on the  $Y = X$  line than the points in Fig. 13 E (Tradaboost-ELM). The comparative experimental results prove that by fine-tuning the parameters of Bi-LSTM layers and re-training the initialized FC layers, the feature difference between target

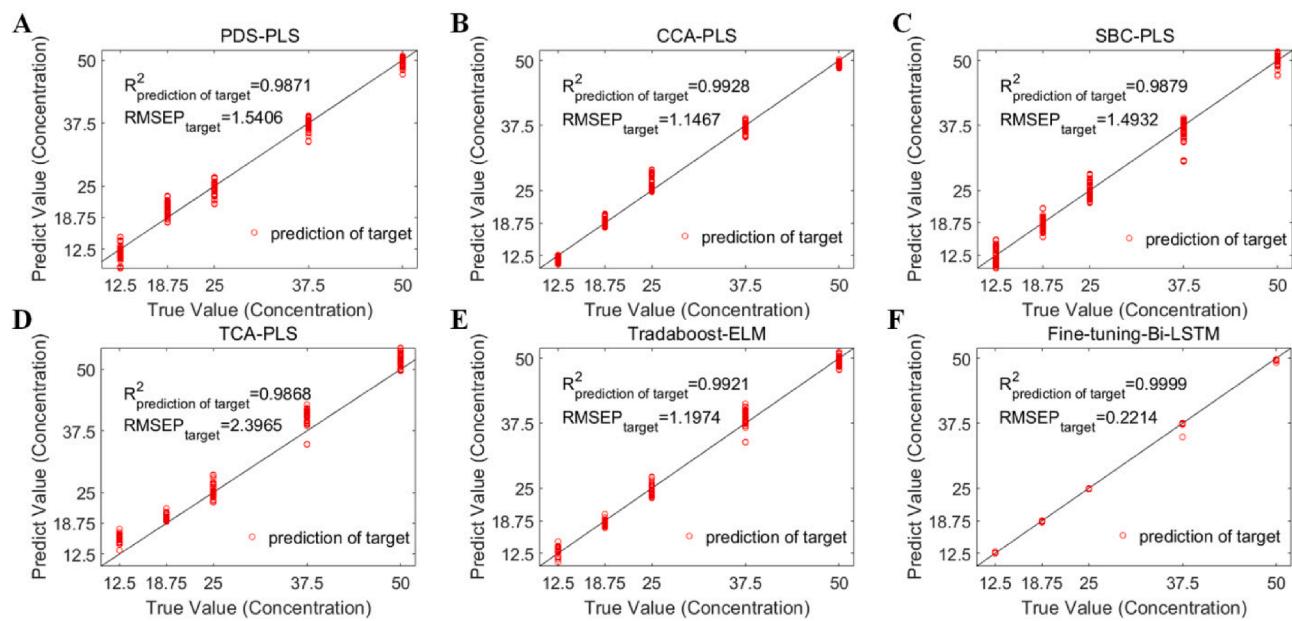
and source spectra can be more fully explored, and the quantitative relationship between the underlying characteristics of target spectra and the new prediction property can be re-constructed. This advantage is particularly prominent when NIRS model is used for a new scenario where the predicting property range is quite different.

Secondly, six transferred models for predicting the  $\gamma$ -PGA concentrations of the second batch of products were compared. From the comparisons of the evaluation index acquired by the primary models in Table 6 and by the transferred models in Table 7, the  $R^2_{\text{target}}$  values were all greatly improved, and the  $\text{RMSEP}_{\text{target}}$  values were all appreciably reduced. Specifically, after PDS, CCA, SBC and TCA re-calibrated the primary PLS model, the  $\text{RMSEP}_{\text{target}}$  values decreased from 9.4821 to 1.5406, 1.1467, 1.4932 and 2.3965, respectively, whilst the  $\text{RMSEP}_{\text{target}}$  value decreased from 6.5875 to 1.1974 after Tradaboost re-calibrated the primary ELM model. However, the proposed fine-tuning-Bi-LSTM in this study, too, like performed in the manure scenario, again has achieved the best performance for different batch of  $\gamma$ -PGA samples which reduced  $\text{RMSEP}_{\text{target}}$  value from 8.9876 to 0.2214. It reduced  $\text{RMSEP}_{\text{target}}$  by 80.6924 % when compared with the results of CCA-PLS method which has obtained the highest prediction accuracy among five traditional methods. Similarly, comparisons between the scatter plots of the six transferred models shown in Fig. 14 A to F respectively

**Table 7**

Prediction accuracy of poultry manure and the second batch of  $\gamma$ -PGA based on different transferred models.

| Samples       | algorithm         | $R^2_{\text{Ctarget}}$ | $\text{RMSE}_{\text{Ctarget}}$ | $R^2_{\text{Vtarget}}$ | $\text{RMSE}_{\text{Vtarget}}$ | $R^2_{\text{Ptarget}}$ | $\text{RMSE}_{\text{Ptarget}}$ |
|---------------|-------------------|------------------------|--------------------------------|------------------------|--------------------------------|------------------------|--------------------------------|
| Manure        | PDS-PLS           | 0.0153                 | 38.2004                        | 0.0009                 | 38.2057                        | 0.0195                 | 38.2272                        |
|               | CCA-PLS           | 0.0092                 | 38.3465                        | 0.0088                 | 38.3039                        | 0.0014                 | 38.0216                        |
|               | SBC-PLS           | 0.7851                 | 5.9492                         | 0.7793                 | 6.0314                         | 0.7967                 | 5.7871                         |
|               | TCA-PLS           | 0.9648                 | 2.4908                         | 0.9492                 | 3.5994                         | 0.9517                 | 3.3835                         |
|               | Tradaboost-ELM    | 0.9611                 | 2.5342                         | 0.9610                 | 2.5403                         | 0.9610                 | 2.5392                         |
| $\gamma$ -PGA | Fine-tune-Bi-LSTM | 0.9991                 | 0.5098                         | 0.9967                 | 0.8378                         | 0.9967                 | 0.8351                         |
|               | PDS-PLS           | 0.9882                 | 1.4777                         | 0.9871                 | 1.5430                         | 0.9871                 | 1.5406                         |
|               | CCA-PLS           | 0.9931                 | 1.1169                         | 0.9928                 | 1.1456                         | 0.9928                 | 1.1467                         |
|               | SBC-PLS           | 0.9888                 | 1.4217                         | 0.9885                 | 1.4507                         | 0.9880                 | 1.4932                         |
|               | TCA-PLS           | 0.9722                 | 2.2432                         | 0.9872                 | 2.2555                         | 0.9867                 | 2.3965                         |
|               | Tradaboost-ELM    | 0.9910                 | 1.2918                         | 0.9911                 | 1.2731                         | 0.9921                 | 1.1974                         |
|               | Fine-tune-Bi-LSTM | 1                      | 0.0747                         | 0.9999                 | 0.2171                         | 0.9999                 | 0.2214                         |



**Fig. 14.** Scatter plots of six different transferred models for  $\gamma$ -PGA concentration of the second batch: A. PDS-PLS; B. CCA-PLS; C. SBC-PLS; D. TCA-PLS; E. Tradaboost-ELM; F. fine-tuning-Bi-LSTM.

also indicate that the deep transfer learning proposed in this study has obviously outperformed the traditional model transfer methods.

Particularly, there is an obvious reversal of CCA-PLS method performed in  $\gamma$ -PGA case, unlike the negative migration occurred in manure case, the corrected  $RMSEP_{target}$  of CCA transferring method is 1.1467, only second to 0.2214 of the fine-tuning-Bi-LSTM. And this is because the concentration gradient of the two batches were the same (while the DM content range in cattle and poultry manure were different), CCA conducted the transfer matrix to convert target  $\gamma$ -PGA spectra to source  $\gamma$ -PGA spectra and combined with the inherited polynomial coefficient between the source spectra features and the concentration, a more accurate prediction for target  $\gamma$ -PGA was obtained.

#### 4. Conclusion

When a well-built NIRS quantitative model is applied to samples produced in different conditions, model transfer is required to re-calibrate the primary model to adapt the target prediction task. Modern deep learning and transfer learning algorithms have started to gain interest in the chemometrics domain as a powerful tool for NIR analysis. In this work, we proposed for the first time a methodology based on an improved Bi-LSTM neural network and transfer learning by two types of fine-tuning approaches to deal with NIRS quantification and model transfer for two kinds of non-homologous samples. The feasibility of the deep neural network and its transferability in our research has great promotion effect on NIRS quantitative analysis and transfer for various scenarios. The findings and contributions can be summarized as follows:

- (1) Ablation experiments of FCs addition to backbone networks of Bi-LSTM and 1D-CNN were implemented to determine the optimal primary model for DM content of the fresh cattle manure and  $\gamma$ -PGA concentration of the first batch, respectively. The obtained  $RMSEP_{source}$  value of cattle manure was 0.4276, which was 67.5175 %, 80.1973 % and 85.3617 % lower than those of optimal CNN, PLS and ELM models; and the obtained  $RMSEP_{source}$  value of the first batch of  $\gamma$ -PGA was 0.1929, which was 75.8119 %, 88.0720 % and 89.5243 % lower than those of optimal CNN, PLS and ELM models. The outstanding prediction accuracy of both the primary Bi-LSTM models demonstrated the feasibility of

the proposed neural network and made the next migration process stand at a higher “starting line”.

- (2) Two types of fine-tuning primary network approaches were both conducted on the poultry manure and the second batch of  $\gamma$ -PGA. The transfer learning results both indicated that the second transfer approach by fine-tuning Bi-LSTM layers and re-training FC layers performed even better than the first approach by fixing Bi-LSTM layers and only re-training FC layers, which reduced the  $RMSEP_{target}$  by 23.4275 % and 50.7343 %, respectively.
- (3) The effectiveness and superiority of the proposed Fine-tuning-Bi-LSTM deep transfer learning methodology were verified by comparing with fine-tuning-1D-CNN and five traditional methods of PDS-PLS, CCA-PLS, SBC-PLS, TCA-PLS, Tradaboost-ELM. The  $RMSEP_{target}$  value of poultry manure obtained by fine-tuning-Bi-LSTM was 0.8351, which was decreased by 7.2832 % and 67.1117 % compared to fine-tuning-1D-CNN and Tradaboost-ELM (the best one of 5 traditional methods); The obtained  $RMSEP_{target}$  value of the second batch of  $\gamma$ -PGA was 0.2214, which was reduced by 48.1256 % and 80.6924 % compared to fine-tuning-1D-CNN and CCA-PLS (the best one of 5 traditional methods).

Besides, future research can be extended to study the quantitative model transfer of different compositions (such as dry matter and nitrogen content) between non-homologous samples based on the proposed Fine-tuning-Bi-LSTM methodology, which is beneficial to make NIR spectroscopy suitable for more complex scenarios.

#### CRediT authorship contribution statement

**Ailing Tan:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Yunxin Wang:** Methodology, Software, Investigation, Formal analysis, Data curation, Writing – original draft, Visualization. **Yong Zhao:** Conceptualization, Methodology, Investigation, Resources, Supervision, Project administration. **Bolin Wang:** Investigation. **Xiaohang Li:** Investigation. **Alan X. Wang:** Supervision.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgment

The authors would like to acknowledge the support from the Hebei Provincial Natural Science Foundation under grant No. C2020203010, the Hebei province Science and Technology Supporting Program under grant No. 19975704D United States Foundation N/a and the National Natural Science Foundation of China under Grant No. 62073280.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.saa.2022.121759>.

## References

- [1] P. Mishra, I. Herrmann, M. Angileri, Improved prediction of potassium and nitrogen in dried bell pepper leaves with visible and near-infrared spectroscopy utilising wavelength selection techniques, *Talanta* 225 (2021), 121971, <https://doi.org/10.1016/j.talanta.2020.121971>.
- [2] X. Miao, Y. Miao, H. Gong, S.U. Tao, Z. Chen, J. Wang, Y. Chen, Y. Chen, NIR spectroscopy coupled with chemometric algorithms for the prediction of cadmium content in rice samples, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 257 (2021), 119700, <https://doi.org/10.1016/j.saa.2021.119700>.
- [3] P. Mishra, R. Nikzad-Langerodi, F. Marini, J.M. Roger, A. Biancolillo, D. N. Rutledge, S. Lohumi, Are standard sample measurements still needed to transfer multivariate calibration models between near-infrared spectrometers? The answer is not always, *TrAC, Trends Anal. Chem.* 143 (2021) 116331.
- [4] P. Mishra, R. Klont, T. Verkleij, S. Wisse, Translating near-infrared spectroscopy from laboratory to commercial slaughterhouse: Existing challenges and solutions, *Infrared Phys. Technol.* 119 (2021), 103918, <https://doi.org/10.1016/j.infrared.2021.103918>.
- [5] P. Yuan-Yuan, C. O'Donnell, J.T. Tobin, N. O'Shea, Review of near-infrared spectroscopy as a process analytical technology for real-time product monitoring in dairy processing, *Int. Dairy J.* 103 (2020), 104623, <https://doi.org/10.1016/j.idairy.2019.104623>.
- [6] N.L. Velez, J.K. Drennen, C.A. Anderson, Challenges, opportunities and recent advances in near infrared spectroscopy applications for monitoring blend uniformity in the continuous manufacturing of solid oral dosage forms, *Int. J. Pharm.* 615 (2022), 121462, <https://doi.org/10.1016/j.ijpharm.2022.121462>.
- [7] C. McVey, T.F. McGrath, S.A. Haughey, C.T. Elliott, A rapid food chain approach for authenticity screening: The development, validation and transferability of a chemometric model using two handheld near infrared spectroscopy (NIRS) devices, *Talanta* 222 (2021), 121533, <https://doi.org/10.1016/j.talanta.2020.121533>.
- [8] L. Li, W. Huang, Z. Wang, S. Liu, X. He, S. Fan, Calibration transfer between developed portable Vis/NIR devices for detection of soluble solids contents in apple, *Postharvest Biol. Technol.* 183 (2022), 111720, <https://doi.org/10.1016/j.postharvbio.2021.111720>.
- [9] P.H. Ciza, P.-Y. Sacre, C. Waffo, T.M. Kimbeni, B. Masereel, P.h. Hubert, E. Ziemons, R.D. Marini, Comparison of several strategies for the deployment of a multivariate regression model on several handheld NIR instruments. Application to the quality control of medicines, *J. Pharm. Biomed. Anal.* 215 (2022) 114755.
- [10] Xiao Wang, Dan-Zhuo Mao, Yong-Jian Yang, Calibration transfer between modelled and commercial pharmaceutical tablet for API quantification using backscattering NIR, Raman and transmission Raman spectroscopy (TRS). 194 (2021) 113766. *Journal of Pharmaceutical and Biomedical Analysis*. <https://doi.org/10.1016/j.jpba.2020.113766>.
- [11] V. Panchuk, D. Kirsanov, E. Oleneva, V. Semenov, A. Legin, Calibration transfer between different analytical methods, *Talanta* 170 (2017) 457–463, <https://doi.org/10.1016/j.talanta.2017.04.039>.
- [12] E. Bouveresse, D.L. Massart, Standardisation of near-infrared spectrometric instruments: A review, *Vib. Spectrosc.* 11 (1996) 3–15, [https://doi.org/10.1016/0924-2031\(95\)00055-0](https://doi.org/10.1016/0924-2031(95)00055-0).
- [13] W. Fan, Y. Liang, D. Yuan, J. Wang, Calibration model transfer for near-infrared spectra based on canonical correlation analysis, *Anal. Chim. Acta* 623 (2008) 22–29, <https://doi.org/10.1016/j.jaca.2008.05.072>.
- [14] S. Fan, J. Li, Y. Xia, X. Tian, Z. Guo, W. Huang, Long-term evaluation of soluble solids content of apples with biological variability by using near-infrared spectroscopy and calibration transfer method, *Postharvest Biol. Technol.* 151 (2019) 79–87.
- [15] D. Galvan, E. Bona, D. Borsato, E. Danieli, M.H. Montazzoli Killner, Montazzoli Killner, Mário Henrique, Calibration Transfer of Partial Least Squares Regression Models between Desktop Nuclear Magnetic Resonance Spectrometers, *Anal. Chem.* 92 (19) (2020) 12809–12816.
- [16] B. Jun, L.i. Xin, F. Wei, Z. Ji-Heng, W. Cheng-Wei, Calibration transfer of near-infrared spectroscopy by canonical correlation analysis coupled with wavelet transform, *The Analyst*. 142 (2017) 2229–2238, <https://doi.org/10.1039/c7an00280g>.
- [17] L. Salguero-Chaparro, B. Palagos, F. Peña-Rodríguez, J.M. Roger, Calibration transfer of intact olive NIR spectra between a pre-dispersive instrument and a portable spectrometer, *Comput. Electron. Agric.* 96 (2013) 202–208, <https://doi.org/10.1016/j.compag.2013.05.007>.
- [18] Z. Xin, S. Jun, T. Yan, C. Quansheng, W. Xiaohong, H. Yingying, A deep learning based regression method on hyperspectral data for rapid prediction of cadmium residue in lettuce leaves, *Chemometrics Intelligent Laboratory Systems*. 200 (2020) 103996.
- [19] Y. Xinxie, L. Huanda, W. Di, Development of deep learning method for predicting firmness and soluble solid content of postharvest Korla fragrant pear using Vis/NIR hyperspectral reflectance imaging, *Postharvest Biol. Technol.* 141 (2018) 39–49, <https://doi.org/10.1016/j.postharvbio.2018.02.013>.
- [20] Pan Sino Jialin, Yang Qiang, A Survey on Transfer Learning, *IEEE Trans. Knowledge and Data Engineering*. 22 (10) (2010) 1345–1359.
- [21] Z. Fuzhen, Q. Zhiyuan, D. Keyu, X. Dongbo, Z. Yongchun, Z. Hengshu, X. Hui, H. Qing, A Comprehensive Survey on Transfer Learning, *Proc. IEEE* 109 (2021) 43–76, <https://doi.org/10.1109/JPROC.2020.3004555>.
- [22] S.J. Pan, I.W. Tsang, J.T. Kwok, Q. Yang, et al., Domain Adaptation via Transfer Component Analysis, *IEEE Trans. Neural Networks* 22 (2) (2011) 199–210.
- [23] Z. Qiu, S. Zhao, X. Feng, Y. He, Transfer learning method for plastic pollution evaluation in soil using NIR sensor, *Sci. Total Environ.* 740 (2020), 140118, <https://doi.org/10.1016/j.scitotenv.2020.140118>.
- [24] Jiang Shuhui, Mao Haiyi, Ding Zhengming, Fu Yun, Deep Decision Tree Transfer Boosting, *IEEE Transactions on Neural Networks and Learning Systems*. 31 (2020) 383–395. Doi:10.1109/TNNLS.2019.2901273.
- [25] Y. Yan, J. Huang, S. Liu, J. Zhu, S. Liang, Cross target attributes and sample types quantitative analysis modeling of near-infrared spectroscopy based on instance transfer learning, *Measurement* 117 (2021), 109340, <https://doi.org/10.1016/j.measurement.2021.109340>.
- [26] J.A. Martins, R. Guerra, R. Pires, M.D. Antunes, T. Panagopoulos, A. Brázio, A. M. Afonso, L. Silva, M.R. Lucas, A.M. Cavaco, SpectraNet-53: A deep residual learning architecture for predicting soluble solids content with VIS-NIR spectroscopy, *Comput. Electron. Agric.* 197 (2022), 106945, <https://doi.org/10.1016/j.compag.2022.106945>.
- [27] L. Zhou, L. Tan, C. Zhang, N. Zhao, Y. He, Z. Qiu, A portable NIR-system for mixture powdery food analysis using deep learning, *LWT*. 153 (2022), 112456, <https://doi.org/10.1016/j.lwt.2021.112456>.
- [28] Mozaffari Mohammad Hamed, Tay Li-Lin, A Review of 1D Convolutional Neural Networks toward Unknown Substance Identification in Portable Raman Spectrometer. 2020. <https://arxiv.org/abs/2006.10575v1>.
- [29] J. Yang, X. Jinfan, X. Zhang, W. Chiyu, T. Lin, Y. Ying, Deep learning for vibrational spectral analysis: Recent progress and a practical guide, *Anal. Chim. Acta* 1081 (2019) 6–17, <https://doi.org/10.1016/j.aca.2019.06.012>.
- [30] Malek, Salim, Melgani, Farid, Bazi, Yakoub, One-dimensional convolutional neural networks for spectroscopic signal regression, *Journal of Chemometrics*, 32 (2018) e2977. <https://doi.org/10.1002/cem.2977>.
- [31] S.S.N. Chakravartula, R. Moscetti, G. Bedini, M. Nardella, R. Massantini, Use of convolutional neural network (CNN) combined with FT-NIR spectroscopy to predict food adulteration: A case study on coffee, *Food Control* 135 (2022), 108816, <https://doi.org/10.1016/j.foodcont.2022.108816>.
- [32] W. Xijun, S. Gao, Y. Niu, Z. Zhao, R. Ma, X. Baoran, H. Liu, Y. Zhang, Quantitative analysis of blended corn-olive oil based on Raman spectroscopy and one-dimensional convolutional neural network, *Food Chem.* 385 (2022), 132655, <https://doi.org/10.1016/j.foodchem.2022.132655>.
- [33] X. Hou, G. Wang, X. Wang, X. Ge, Y. Fan, S. Nie, Convolutional neural network based approach for classification of edible oils using low-field nuclear magnetic resonance, *J. Food Compos. Anal.* 92 (2020), 103566, <https://doi.org/10.1016/j.jfca.2020.103566>.
- [34] J. Liang, L. Xingxing, T. Chang, H.-L. Cui, Deep learning aided quantitative analysis of anti-tuberculosis fixed-dose combinatorial formulation by terahertz spectroscopy, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 269 (2020), 120746, <https://doi.org/10.1016/j.saa.2021.120746>.
- [35] H. Chen, A. Chen, X. Lili, H. Xie, H. Qiao, Q. Lin, K. Cai, A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources, *Agric. Water Manag.* 240 (2020), 106303, <https://doi.org/10.1016/j.agwat.2020.106303>.
- [36] W. Xijun, X. Baoran, R. Ma, Y. Niu, S. Gao, H. Liu, Y. Zhang, Identification and quantification of adulterated honey by Raman spectroscopy combined with convolutional neural network and chemometrics, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 274 (2022), 121133, <https://doi.org/10.1016/j.saa.2022.121133>.
- [37] X. Zhang, J. Yang, T. Lin, Y. Ying, Food and agro-product quality evaluation based on spectroscopy and deep learning: A review, *Trends Food Sci. Technol.* 112 (2021) 431–441, <https://doi.org/10.1016/j.tifs.2021.04.008>.

- [38] Simranjit Singh, Singara Singh Kasana, Estimation of soil properties from the EU spectral library using long short-term memory networks, *Geoderma Regional*. 18 (2019) e00233. <https://doi.org/10.1016/j.geodrs.2019.e00233>.
- [39] E. Dandil, S. Karaca, Detection of pseudo brain tumors via stacked LSTM neural networks using MR spectroscopy signals, *Biocybernetics and Biomedical Engineering*. 41 (2021) 173–195, <https://doi.org/10.1016/j.bbe.2020.12.003>.
- [40] X. Wang, S. Tian, Y. Long, X. Lv, Z. Zhang, Rapid screening of hepatitis B using Raman spectroscopy and long short-term memory neural network, *Lasers Med Sci*. 35 (2020) 1791–1799, <https://doi.org/10.1007/s10103-020-03003-4>.
- [41] Q. Wang, Q. Liu, R. Xia, P. Zhang, H. Zhou, B. Zhao, G. Li, Automatic defect prediction in glass fiber reinforced polymer based on THz-TDS signal analysis with neural networks, *Infrared Phys. Technol.* 115 (2021), 103679, <https://doi.org/10.1016/j.infrared.2021.103679>.
- [42] C. Chen, B. Yang, R. Si, C. Chen, F. Chen, R. Gao, Y. Li, J. Tang, X. Lv, Fast detection of cumin and fennel using NIR spectroscopy combined with deep learning algorithms, *Optik*. 242 (2021), 167080, <https://doi.org/10.1016/j.jleo.2021.167080>.
- [43] P. Mishra, D. Passos, Deep calibration transfer: Transferring deep learning models between infrared spectroscopy instruments, *Infrared Phys. Technol.* 117 (2021), 103863, <https://doi.org/10.1016/j.infrared.2021.103863>.
- [44] P. Mishra, D. Passos, Realizing transfer learning for updating deep learning models of spectral data to be used in new scenarios, *Chemometrics and Intelligent Laboratory Systems*. 212 (2021), 104283, <https://doi.org/10.1016/j.chemolab.2021.104283>.
- [45] Z.N.K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, L. Jianfeng, Brain tumor classification for MR images using transfer learning and fine-tuning, *Comput. Med. Imaging Graph.* 75 (2019) 34–46, <https://doi.org/10.1016/j.compmedimag.2019.05.001>.
- [46] A. Abdalla, H. Cen, L. Wan, R. Rashid, H. Weng, W. Zhou, Y. He, Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure, *Comput. Electron. Agric.* 167 (2019), 105091, <https://doi.org/10.1016/j.compag.2019.105091>.
- [47] Z. Zhou, J.Y. Shin, S.R. Gurudu, M.B. Gotway, J. Liang, Active, continual fine tuning of convolutional neural networks for reducing annotation efforts, *Med. Image Anal.* 71 (2021), 101997, <https://doi.org/10.1016/j.media.2021.101997>.
- [48] M. Hamed Mozaffari, W.-S. Lee, Won Sook Lee, Domain adaptation for ultrasound tongue contour extraction using transfer learning: A deep learning approach, *The Journal of the Acoustical Society of America*. 146 (5) (2019) EL431–EL437.
- [49] Liang Huan, Fu Wenlong, Yi Fengji, A Survey of Recent Advances in Transfer Learning, 2019 IEEE 19th International Conference on Communication Technology (ICCT). (2019) 1516–1523. Doi:10.1109/ICCT46805.2019.8947072.
- [50] N. Shuteng, L. Yongxin, W. Jian, S. Houbing, A Decade Survey of Transfer Learning (2010–2020), *IEEE Trans. Artificial Intelligence*. 1 (2020) 151–166, <https://doi.org/10.1109/TAL.2021.3054609>.
- [51] F. Gogé, L. Thuriès, Y. Fouad, N. Damay, F. Davrieux, G. Moussard, C. Le Roux, S. Trupin-Maudemain, M. Valé, T. Morvan, Dataset of chemical and near-infrared spectroscopy measurements of fresh and dried poultry and cattle manure, *Data in Brief*. 34 (2021), 106647, <https://doi.org/10.1016/j.dib.2020.106647>.