Application of Improved Deep Neural Network in Complex Chemical Soft Measurement

Xilong Ding

Weifang University of Science and Technology, Shandong 262700, China
xilongding23189@126.com

To improve the application effect of deep neural network in complex chemical soft measurement. After Kalman filtering equations were introduced, the algorithms of extended Kalman filter, volumetric Kalman filter and square root volumetric Kalman filter were optimized. After comparing the original algorithm with the optimization algorithm, we found that the design algorithm had better performance in all aspects. In this paper, the comparison between the original algorithm and the optimized algorithm is proved to be effective and the purpose of the algorithm optimization is achieved.

1. Introduction

In the modern chemical field, chemical soft measurement will be used to measure the results of some of the work. However, the accuracy of traditional measurement techniques does not meet expectations. Therefore, deep neural networks are used in modern technology. At present, neural network has become an important soft measurement modeling tool, and the common technology is recursive neural network. As a dynamic neural network, this technology has been successfully applied in data-driven soft measurement modeling. It mainly uses the Kalman filter to update the connection weights between neurons in each layer of the neural network as the state of the filter. It replaces the traditional training algorithm of the traditional RNN network, effectively improves the prediction accuracy of the network and has been successfully applied to time series prediction. However, in practical applications, this technology still does not reach the expected level, hence further improvements are needed.

For the purpose of improvement, based on the linear Kalman filter, this paper presents an optimization algorithm for extended Kalman filter, volumetric Kalman filter and square root volumetric Kalman filter. Then the algorithm is applied in the training of RNN. Through the application of the results, it shows that the improved method in this paper has achieved satisfactory results in many aspects in the case of soft-sensing modeling of complex chemical processes.

2. Literature review

In recent years, there are many researches and applications of deep neural network, and its application in complex chemical soft sensing is also paid certain attention to. A lot of foreign scholars explored the deep neural network from a variety of aspects. Narayanan and Wang pointed out that although deep neural network (DNN) acoustic models were known to be inherently noise robust, especially with matched training and testing data, the use of speech separation as a frontend and for deriving alternative feature representations was shown to improve performance in challenging environments. In addition, they first of all presented a supervised speech separation system that significantly improved automatic speech recognition (ASR) performance in realistic noise conditions. Then, they proposed a framework that unified separation and acoustic modeling through joint adaptive training (Narayanan and Wang, 2015). Kandaswamy and so on proposed a novel feature transference approach, especially when the source and the target problems were drawn from different distributions. They applied deep neural networks to transfer either low or middle or higher-layer features for a machine trained in either unsupervised or supervised way. Applying this feature transference approach on Convolutional Neural Network and Stacked Denoising Autoencoder on four different

datasets, they achieved lower classification error rate with significant reduction in computation time with lower-layer features trained in supervised way and higher-layer features trained in unsupervised way for classifying images of uppercase and lowercase letters dataset (Kandaswamy, 2014). Prusa and Khoshgoftaar proposed a new method of creating character-level representations of text to reduce the computational costs associated with training a deep convolutional neural network. They pointed out that the method of character embedding greatly reduced training time and memory use, while significantly improving classification performance. Additionally, they suggested that the proposed embedding could be used with padded convolutional layers to enable the use of current convolutional network architectures, while still facilitating faster training and higher performance than the previous approach for learning from character-level text (Prusa and Khoshgoftaar, 2017). Except for the researches done by foreign scholars, some scholars in domestic also made great effort to the exploration of deep neural network. They mainly discussed the deep neural network from the concept, application, characteristics and so on aspects. Hu and other scholars improved mispronunciation detection performance with a Deep Neural Network (DNN) trained acoustic model and transferred learning-based Logistic Regression (LR) classifiers. The acoustic model trained by the conventional GMM-HMM based approach was refined by the DNN training with enhanced discrimination (Hu et al., 2015). Liu and other scholars put forward a cluster-based senone selection method to speed up the computation of deep neural networks (DNN) at the decoding time of automatic speech recognition (ASR) systems. In DNN-based acoustic models, the large number of senones at the output layer was one of the main causes that led to the high computation complexity of DNNs. The senone selection strategy was derived by clustering acoustic features according to their transformed representations at the top hidden layer of the DNN acoustic model. Experimental results showed that the average number of DNN parameters used for computation could be reduced by 22% and the overall speed of the recognition process could be accelerated by 13% without significant performance degradation after using the proposed method (Liu et al., 2017). Tian and so on adopted an improved deep neural network method to predict the surface appearance parameters. In order to meet the high accuracy requirements for the prediction results of the contact model, a novel surface appearance prediction model was established utilizing a regularized deep belief network. The Bayesian regularization strategy was used to reduce the network weights during unsupervised training, which could effectively restrain the contribution of unimportant neurons. This allowed limiting the occurrence of overfitting, and the layer-by-layer training was performed for each hidden layer based on a continuous transfer function. Meanwhile, the surface appearance parameters of the joint interface could be obtained by plugging arbitrary machining parameters into the training model (Tian et al., 2016). Ling and other researchers presented a method of using deep neural networks to learn a model for the Reynolds stress anisotropy tensor from high-fidelity simulation data. A novel neural network architecture was proposed which used a multiplicative layer with an invariant tensor basis to embed Galilean invariance into the predicted anisotropy tensor. It was demonstrated that this neural network architecture provided improved prediction accuracy compared with a generic neural network architecture that did not embed this invariance property. The Reynolds stress anisotropy predictions of this invariant neural network were propagated through to the velocity field for two test cases. For the test cases, significant improvement versus baseline RANS linear eddy viscosity and nonlinear eddy viscosity model was demonstrated (Ling et al., 2016). To address the challenging problem of vector quantization (VQ) for high dimensional vector using large coding bits, Jiang and others proposed a novel deep neural network (DNN) based VQ method. This method used a k-means based vector quantizer as an encoder and a DNN as a decoder. The decoder was initialized by the decoder network of deep auto-encoder, fed with the codes provided by the k-means based vector quantizer, and trained to minimize the coding error of VQ system. The experiments on speech spectrogram coding showed that, compared with the k-means based method and a recently introduced DNN-based method, the proposed method significantly reduced the coding error. Furthermore, in the experiments of coding multi-frame speech spectrogram, the proposed method achieved about 11% relative gain over the k-means based method in terms of segmental signal to noise ratio (SegSNR) (Jiang et al., 2017).

In a word, as a novel speech recognition software works by imitating the way of thinking of human brain, deep neural network is widely applied, and the speech recognition speed of the software is faster and the recognition accuracy is higher. To sum up, the above studies enabled use have a clearer understanding of deep neural network and know the difference between deep neural network and other neural networks, so that we can better apply it in various fields, but it still needs to be further explored.

3. Methods

3.1 Recursive neural network based on extended kalman filtering

The Kalman filter equation contains two aspects, namely the time update equation and the measurement update equation. The time update equation mainly forwards the current state and the error covariance

estimate to obtain the prior estimate at the next moment, while the measured update equation can obtain updated posteriors based on the measured values and a priori estimate. It is also possible to regard the time update equation as a prediction equation, and to regard the measurement update equation as a check equation. In fact, the Kalman filtering algorithm is similar to an iterative "prediction-correction" algorithm that solves numerical problems, which is as shown in Figure 1.
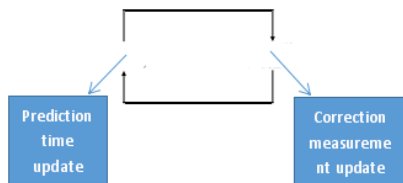

Figure 1: continuous Kalman filtering cycle

After completing a time update and measurement update, the a priori estimate at the previous moment will be used for the prediction of the a priori estimate at the next moment to repeat. Compared to Wiener filtering, the recursive nature of Kalman filtering increases its feasibility in the actual implementation process. And as long as the current measurement data is obtained, you can calculate the filter value at the current moment without storing a large amount of data, which is easy to computer real-time processing. However, the linear Kalman filter can only deal with the state estimation problem of the linear system. For most of the non-linear system state estimation problems that exist in real life, this method seems to be ineffective.

3.2 Extended kalman filter algorithm

The EKF learning algorithm is a suboptimal state estimation technique, and it is a generalized form of the Kalman filter applied to nonlinear systems. The method uses Taylor series to discrete local linearization of continuous nonlinear system equations, and calculates the estimated value based on the partial derivative of the process and measurement equations, which can solve the problem that linear Kalman filter can not effectively settle problems of the nonlinear system.

3.3 Full-connection recursive neural network based on extended kalman filter algorithm

According to the traditional neural network training algorithm, the essence of the network training process is the process of adjusting the weights, namely adjusting the connection weights between the various layers of the network through a given input-output sample set. Therefore, neural network training can be viewed as a nonlinear state estimation problem.

The basic idea of the FCRNN method based on the EKF algorithm is to treat the network training problem as a dynamic parameter estimation problem of a nonlinear dynamic system, That is, the vector of connection weights among the neurons in the FCRNN is used as the state of the filter. As the time sequence k continues to update the connection weight parameters of the network, it makes the mean square error between the output of the actual output of the network gradually decrease with the increase of the timing, thus achieving to improve the training accuracy.

3.4 Recursive neural network based on square root volumetric kalman filtering

CKF is a nonlinear filter used to solve the problem of high-dimensional state estimation. Its core idea is the Spherical-Radial criterion, that is, by using the third-order Spherical-Radial volume rule to generate volume points and using the mean and covariance of the posterior distribution of the approximation point of the volume point, multivariable integral numerical solution in high-dimensional nonlinear filtering becomes possible.

The reasons that affect the accuracy of the CKF algorithm:

1) System model error

The premise of Kalman filtering is that the assumptions are linear, the noise obeys Gaussian white noise with a normal distribution, and it requires a known and accurate system mathematical model. In real systems, the noise error is usually large, which can lead to one-step prediction of the state is not accurate enough.

2) Filter Initial Value Selection

Although CKF can avoid some of the deficiencies of the EKF method, it has unique advantages for linear filtering, but in the actual application process, the choice of initial value has a great influence on its

performance. CKF is a nonlinear suboptimal Gaussian filtering method that approximates the posterior mean and covariance based on the determined volume point in the Bayesian filtering framework. Theoretically, the Spherical-Radial volume criterion has proved to be able to approximate at least the posterior mean and covariance of any nonlinear system state with a second-order Taylor precision. Therefore, the accuracy of CKF is significantly improved compared to EKF. In addition, CKF does not require the Jacobians matrix in the EKF algorithm, which effectively reduces the computational complexity of the algorithm, but the numerical stability of the CKF algorithm needs to be improved.

3.5 Square root volumetric kalman filter algorithm

In the process of propagating a volume point, the standard CKF algorithm has error effects due to arithmetic operations performed on a computer with limited precision, and two basic characteristics of the error covariance matrix, namely positive definiteness and symmetry, are often lost, and the loss of their positive definiteness may cause the operation of the CKF algorithm to be terminated. In the CKF time updating and measurement updating process at each time point, the square root operation, the inverse operation, and the rounding error of the matrix involved in the CKF time update and measurement updating process will abolish the characteristics of the covariance matrix. Moreover, some nonlinear filtering problems may also be numerically ill-conditioned, making the covariance matrix possibly non-definite, resulting in unstable or even non-convergent algorithms. To solve the above problem, CKF's square-root improvement algorithm, SCKF, was first proposed. SCKF essentially propagates the square root factor of prediction and posterior error covariance, avoids square root operations of the matrix, and improves the numerical stability of the algorithm due to the condition number. In addition, SCKF also maintains the symmetry and positive definiteness of the covariance.

4. Results and analysis

Soft measurement of C4 concentration at the bottom of debutanol:
The debutanizer is an essential part of the desulfurization and naphtha separation unit in the refining process of a refinery. The naphtha separation unit is mainly used to separate several kinds of alkanes from naphtha, which mainly includes a desulfurization tower, a depropanizer tower, a debutanizer tower and other equipment.
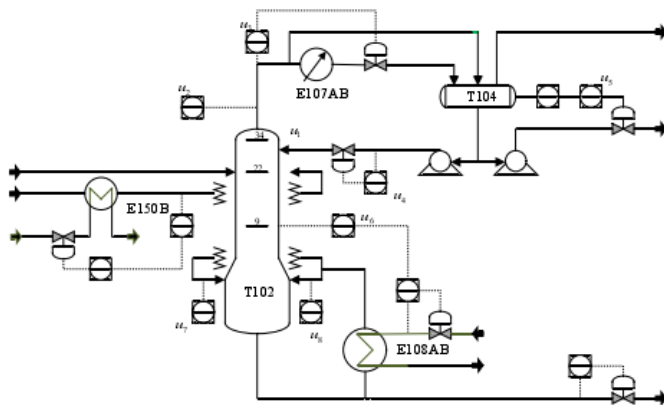

Figure 2: Operation flow of debutane tower

The main role of the debutanizer is to separate C5 from C3 and C4. During the separation process, the equipment removes C3 and most of the C4 from the top and removes C5 and residual C4 at the bottom. Since C5 is an important substance for the production of gasoline, if a small amount of C4 is added to C5 will affect the quality of gasoline, the concentration of C4 contained in the C5 separated from the bottom of the tower is required to be minimized to ensure the control of gasoline quality.
Figure. 3 and Figure. 4 show the comparison results of the network estimation output and the actual output of the soft measurement model for C4 concentration estimation based on the FCRNN-EKF and the SRN-SCKF method respectively in the test data set. From Figure. 3 and Figure. 4, it can be seen that the prediction results of the two methods are satisfactory. Among them, the SRN-SCKF method has better prediction results.
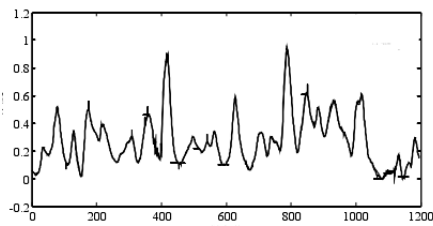
Figure 3: Results of C4 concentration estimation on test set based on FCRNN-EKF method
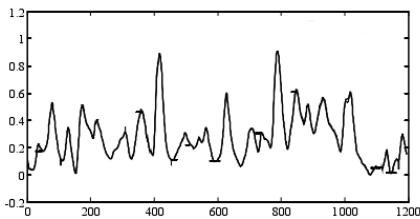

Figure 4: Results of C4 concentration estimation on test set based on SRN-SCKF method

Figure. 5 shows the MSE of the training data set as a function of the number of iterations when training the network based on different methods. As can be seen from Figure. 5, the SRN-SCKF method can achieve higher accuracy with faster convergence.
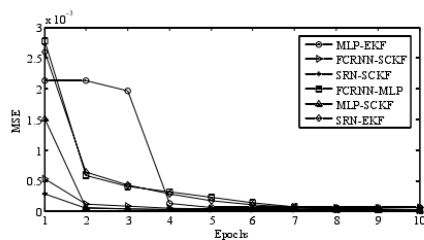

Figure 5: Comparison of MSE curves for estimating C4 concentration based on different methods in training process

On the test data set, the specific performance indicators for C4 concentration estimation by different soft measurement methods are given in Table 1. From Table 1, we can see that compared with the traditional RNN training algorithm, the Kalman filtering algorithm is used to train the neural network faster and the convergence accuracy is higher. Given that the EKF algorithm and the SCKF learning algorithm are also used, the accuracy of the RNN is higher than that of the feed-forward neural network, and the accuracy of the SRN is equal to that of the FCRNN. The prediction accuracy of the SRN-SCKF method is higher than that of SRN-EKF, validating its effectiveness.

Table 1: Results of performance evaluation for C4 concentration on different sets of test sets

different methods	correlation coefficient	MSE	epochs
FCRNN-BPTT	0.9857	$9.19×10^{-4}$	300
SRN-BPTT	0.9861	$9.86×10^{-4}$	300
FCRNN-RTRL	0.9903	$6.72×10^{-4}$	300
SRN-RTRL	0.9903	$5.06×10^{-4}$	300
MLP-EKF	0.9957	$3.52×10^{-4}$	10
FCRNN-EKF	0.9988	$2.08×10^{-4}$	10
SRN-EKF	0.9982	$1.41×10^{-4}$	10
MLP-SCKF	0.9991	$7.82×10^{-4}$	10
FCRNN-SCKF	0.9991	$6.10×10^{-4}$	10
SRN-SCKF	0.9997	$2.70×10^{-4}$	10

In addition, the method of this paper is also compared with the theoretical results. Among them, the 16-12-1 structure MLP method is used, and Li ESN network is adopted; off-line and online RLS learning algorithm is employed, and the literature uses Bayesian network. Consequently, the estimation accuracy of this method is obviously better than that of the literature, and is slightly better than the result of the pair.

5. Conclusion

Firstly, in this paper, the algorithms of extended Kalman filter, volumetric Kalman filter and square root volumetric Kalman filter are optimized on the basis of linear Kalman filtering. Afterwards, three optimization results of full-connection recursive neural network based on extended Kalman filter algorithm, recursive neural network based on square root volumetric Kalman filter, and square root volumetric Kalman filter algorithm are proposed. Then it analyzes the three results, and in the practical application of the soft-sensing C4 concentration at the bottom of the debutanizer, a comparison chart of the network estimation output and the actual output of the soft measurement model for C4 concentration estimation based on the FCRNN-EKF and the SRN-SCKF method established in the test data set is presented. The comparison results show that the prediction results of the two methods are satisfactory, but the prediction methods used in this paper are better. The curve of the MSE of the training data set changing with the number of iterations is given when training the network based on different methods. The comparison results show that the convergence speed of the method used in this paper is fast, indicating that the algorithm has a good optimization effect.

Reference

Hu W., Qian Y., Soong F.K., Wang Y., 2015, Improved mispronunciation detection with deep neural network trained acoustic models and transfer learning based logistic regression classifiers, Speech Communication, 67, 154-166, DOI: 10.1016/j.specom.2014.12.008

Jiang W., Liu P., Wen F., 2017, An improved vector quantization method using deep neural network. AEU - International Journal of Electronics and Communications, 72, 178-183, DOI: 10.1016/j.aeue.2016.12.002

Kandaswamy C., Silva L.M., Alexandre L.A., Santos J.M., Sá J.M.D., 2014, Improving deep neural network performance by reusing features trained with transductive transference, ICANN, 8681, 265-272, DOI: 10.1007/978-3-319-11179-7_34

Ling J., Andrew K., Jeremy T., 2016, Reynolds averaged turbulence modelling using deep neural networks with embedded invariance, Journal of Fluid Mechanics, 807, 155-166, DOI: 10.1017/jfm.2016.615

Liu J.H., Ling Z.H., Wei S., Hu G.P., Dai L.R., 2017, Improving the decoding efficiency of deep neural network acoustic models by cluster-based senone selection, Journal of Signal Processing Systems, (2), 1-13, DOI: 10.1007/s11265-017-1288-9

Narayanan A., Wang D.L., 2015, Improving robustness of deep neural network acoustic models via speech separation and joint adaptive training, Audio Speech & Language Processing, 23(1), 92, DOI: 10.1109/taslp.2014.2372314

Prusa J.D., Khoshgoftaar T.M., 2017, Improving deep neural network design with new text data representations, Journal of Big Data, 4(1), 7, DOI: 10.1186/s40537-017-0065-8

Tian Y., Liu Z., Cai L., Pan G., 2016, The contact mode of a joint interface based on improved deep neural networks and its application in vibration analysis, Journal of Vibroengineering, 18(3), 1388-1405, DOI: 10.21595/jve.2016.16373