THE SIGNIFICANCE OF THE CONVOLUTIONAL DEEP LEARNING MODEL IN THE INTELLIGENT COLLABORATIVE CORRECTION OF ENGLISH WRITING

Hong Wang*

Department of Social Work, Formerly the Ministry of Civil Affairs Management Cadre Institute, Beijing College of Social Administration, Beijing, 102628, China

hongw_200466@163.com

Reception: 18/11/2022 **Acceptance**: 07/01/2023 **Publication**: 26/01/2023

Suggested citation:

W., Hong. (2023). **The significance of the convolutional deep learning model in the intelligent collaborative correction of English writing**. *3C Tecnología. Glosas de innovación aplicada a la pyme, 12(1)*, 142-157. https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

ABSTRACT

To ensure the normal operation of English composition grammar correction and avoid inaccurate detection caused by faults, it is of great significance to detect abnormal working conditions in time and diagnose them accurately. Aiming at the complexity of grammar correction, this paper proposes a PLSTM-CNN model for fault detection in the grammar correction process. The model effectively combines the global feature extraction ability of LSTM for time series data and the ability of the CNN model to extract local features, which reduces the loss of feature information and achieves a higher fault detection rate. A one-dimensional dense CNN is used as the main body of the CNN, and the LSTM network is sensitive to changes in sequence information to avoid model overfitting while building a deeper network. The maximum mutual information coefficient (MMIC) data preprocessing method is adopted to improve the local correlation of the data and improve the efficiency of the PLSTM-CNN model to detect faults from different initial conditions. The research results show that the parallel PLSTM-CNN has better prediction performance than the serial PLSTM-CNN, and its FDR and FPR are 90.5% and 0.051, respectively. It shows that the use of convolutional deep learning models for the prediction of writing grammar correction faults has strong application prospects.

KEYWORDS

PLSTM-CNN; Fault detection; English writing; grammar; Deep learning.

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

PAPER INDEX

1. INTRODUCTION

Writing ability is an important basis for measuring students' English learning and practical ability, and plays an significant role in promoting the overall development of language skills. In recent years, the rapid development of computers and networks has laid the foundation for the reform of college English writing [1]. Due to the limitations of the current technology of the automatic evaluation system, more feedback is given to students at the vocabulary level, but there are still many problems in the evaluation of syntax, text structure, logic, and coherence. Therefore, relying solely on machine correction and feedback is of limited help to students [2]. To solve this problem, manual intervention and feedback are required [3]. Therefore, the need for intelligent collaboration is extremely necessary.

The research on English marking begins with writing feedback. Research on writing feedback began in the 1950s. Before this, writing feedback was conducted by teachers. During this period, some scholars conducted comparative studies on teachers' correction of students' compositions and students' peer evaluation under the guidance of teachers and found that peer feedback was better than teacher feedback [4]. In the following decades, more and more researchers from China and abroad began to pay attention to the application of peer feedback in practical teaching and research, and their research results provided a lot of guidance for our teaching and research [5]. However, some problems and difficulties with peer feedback have also been found in some studies. For example, the correctness, fairness, and effectiveness of feedback are often questioned, and the operability in writing classrooms also needs to be verified. In the 1960s, Professor Ellis Page of Duke University in the United States developed the PEG automatic composition scoring system, and the automatic writing evaluation system (AWES) gradually developed [6]. For decades, especially with the development of artificial intelligence technology, the development of foreign writing automatic evaluation systems has made great progress, such as Criterion, My Access, and Writing Roadmap [7]. Then there is the study of translation. According to the basic principle of deep reinforcement learning algorithm, some researchers designed a neural machine translation model, introduced the evaluation mechanism to the level of the sentence to be translated, predicted the convergence of the translation, and used the deep reinforcement learning algorithm as a guiding strategy for translation. Optimize the word sequence of the translation target, integrate the monolingual corpus into the training of deep reinforcement learning, and alleviate the data-sparse problem of translation sentences. After experimental tests, it is found that this model can improve the overall performance of machine translation. Compared with other translation models, Whether it is Chinese-Korean or Korean-Chinese, the BLEU value has been significantly improved [8], but the performance in other aspects is poor. Considering the problems existing in traditional machine translation, some researchers have designed a neural machine translation model based on the basic principle of quality estimation, and used the quality estimation method to score the pseudo-parallel data generated by reverse translation, and use the data with higher scores as the basis for quality estimation. Design of CNN's English Machine Translation Minor Error Detection System [9]. Some scholars use the input of the

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

neural network to control the quality of the pseudo-parallel data generated by the reverse translation, and provide a rich training network as the output for the neural network model. After experimental tests, it is found that compared with the traditional model, the model has no effect on the forward translation or reverse translation. For translation, the BLEU value has been improved, but the function cannot meet the design requirements [10].

One-dimensional CNN [11] divides the input data along a single dimension without windowing operation, which is easier to train and has less computational complexity. Although 1D convolution is the current popular deep learning method, it still has some limitations. Although the CNN can extract the local features of the data, its ability to extract the global features of the data is weak. To obtain beneficial features such as global and local use of data at the same time, this paper proposes an LSTM-CNN structure based on a parallel structure, which combines the local features extracted by the CNN and the global features extracted by LSTM to make full use of the data. features to improve the accuracy of the model, thereby reducing the translation accident rate [12].

2. CORRELATION MODEL THEORY

2.1. 1D-CNN

A CNN is one of the widely studied deep learning algorithms, which has the characteristics of local connection, weight sharing, and downsampling [13]. The difference between 1D-CNN and classical CNN is the dimension of the convolution kernel, which has been widely used in time series feature extraction in recent years [14]. The one-dimensional convolution operation is shown in Figure 1, and its mathematical model is shown in formula (1):

$$H_i = f(H_{i-1} \otimes W_i + b_i) \tag{1}$$

Among them, Hi is the input feature quantity of the ith layer; Wi and bi represent the weight and corresponding bias of the convolution kernel of the ith layer respectively; f represents the activation function, here is the Relu activation function, which has a good nonlinear expression ability.

Pooling layers are also known as subsampling layers. The sub-sampling layer downsamples the feature map according to the rules and reduces the dimension of the convolutional feature to reduce the parameters and calculation amount inside the CNN, and at the same time suppress the network overfitting.

Suppose H^l_t is the jth feature map of the l-th sub-sampling layer, and its sampling process is as shown in formula (2), which represents the next sampling function. Each output feature map corresponds to its own multiplicative bias B^l_j and an additive bias b^l_j [15].

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

$$H_j^l = f\left(\beta_j^l down\left(H_j^{l-1}\right) + b_j^l\right) \tag{2}$$

2.2. DCNN

To alleviate the gradient explosion problem caused by the network depth, the dense CNN [16] (Dense CNN, DCNN) further connects each sub-layer based on the residual network structure, so that the output of each layer of the network is used as The input of one layer ensures maximum feature reuse, alleviates the gradient disappearance and gradient explosion problems caused by the increase in the number of network layers and makes the network information flow more smoothly.

Assuming that the number of network layers is, the DCNN contains a total of N/(N+1)/2 connections. Through the sequence x_0 passed through the convolutional layer, the input of the nth layer is the feature map of all previous layers, as shown in the following formula (3):

$$x_n = H([x_0, x_1, \cdots, x_{n-1}]) \tag{3}$$

Among them, $[x_0, x_1, \ldots, x_{n-1}]$ represents the feature map in the 0,1...,n-1 layer; Hl( ) represents the normalized linear correction unit, Relu activation function, pooling operation and volume Product equivalence transformation. The output of the convolutional layer and the pooling method formula is as follows:

$$\lambda_{output} = \left\lceil \frac{\lambda_{input}}{S} \right\rceil \tag{9}$$

$$\lambda_{output} = \left\lceil \frac{\lambda_{input} - F + 1}{S} \right\rceil \tag{10}$$

2.3. LSTM LONG SHORT-TERM MEMORY NETWORK

A Long Short-Term Memory Network (LSTM) is a temporal recurrent neural network that is an optimization of a Recurrent Neural Network (RNN). RNN is often used to analyze and predict time series, but it is used for short time series, and it is not suitable for long-distance and long period time series. LSTM perfectly solves the shortcomings of RNN. It significantly improves the model's ability to analyze and predict long sequences by designing hidden layers without changing the original model structure. The operation state of LSTM is almost linear, and the entire operation mode is also chain operation, and there will be no problems such as gradient expansion and disappearance in the RNN training process, which improves the prediction effect and accuracy. The LSTM model is mainly composed of a forgetting gate, input gate, and output gate. After the unit gate enters the forgetting gate, the forgetting gate is responsible for screening out the unit state that can be retained to the current moment at the previous moment, and the input gate is responsible for

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

screening out a certain number of cells. The input at the current moment is the unit state at the current moment, and the output gate is responsible for the output of the unit state [17].

Information is mainly selected through three gate structures: input gate, forget gate, and output gate. Taking the t-th sample as an example is the input data at the current moment. Equation (13)~(18) is the process performed by the LSTM unit [20]：

$$f_t = \sigma\left(w^f \cdot \left[h_{t-1}, x_t\right] + b^f\right) \tag{11}$$

$$i_t = \sigma\left(w^i \cdot \left[h_{t-1}, x_t\right] + b^i\right) \tag{12}$$

$$\tilde{C}_t = \tanh\left(w^c \cdot \left[h_{t-1}, x_t\right] + b^c\right) \tag{13}$$

Where w_f, w_i, $w_{c,}$ and w_o are the weights of the corresponding forgetting gate, input gate, and output gate respectively [18].

2.4. MODEL STRUCTURE DIAGRAM

According to the above model theory, a joint prediction and correction model of DCNN and LSTM is established, combined with the advantages of the LSTM algorithm, an LSTM-CNN hybrid model is formed to perform model training on the data set, and the short text of unknown category is predicted by the trained model. As shown in Fig.1 [19], this parallel network structure avoids feature loss to the greatest extent and preserves the global and local information of data features; the residual structure enhances CNN stability and reduces resource occupancy; batch normalization layer (Batch normalization, BN) to speed up network training and suppress network overfitting; the global average pooling layer on the left splices the dimensionality-reduced feature map with the network feature map on the right, and finally sends it to the classification layer [20].

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

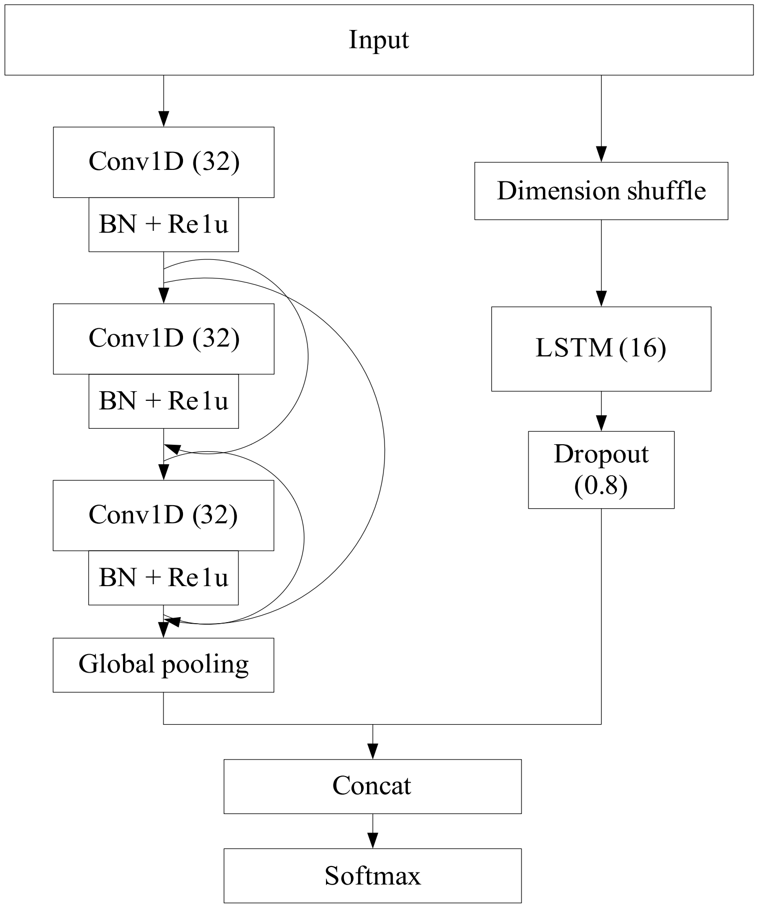

Figure 1. PLSTM-CNN network model

3. EXPERIMENT AND RESULT ANALYSIS

3.1. EVALUATION RESULTS

To show the fault checking results and evaluate the model performance, two metrics, Fault Diagnosis Rate (FDR) and False Positive Rate (FPR), were used to evaluate the model performance. FDR and FPR are defined in [21].

Among them, TP (True Positives) represents the number of instances that are positive classes and are predicted to be positive classes, FN (False Negative) represents the number of positive classes that are predicted to be negative, and FP (False Positives) that instances are negative classes that are predicted to be positive The number of classes, TN (True Negative) represents the number of instances that are predicted to be negative classes.

To verify the performance of the proposed method, the diagnostic results of the test data based on the PLSTM CNN model are compared with the results of the two-dimensional CNN model-based method and the LSTM model-based method on the test data, and the results are shown in the table., the average fault diagnosis rate based on the PLSTM-CNN model in the table is 91.4% [22].

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

Table 1. The comparison of fault results

Fault type	FDR			FPR		
	2D-CNN	LSTM	PLSTM-CNN	2D-CNN	LSTM	PLSTM-CNN
Normal	0.91	1.0	1.0	0.08	0.03	0
Fault1	1.0	1.0	1.0	0	0	0
Fault2	1.0	1.0	0.81	0	0	0
Fault3	0.48	1.0	0.92	0.24	0.14	0.07
Fault4	1.0	0.81	0.34	0	0	0
Fault5	1.0	0.92	1.0	0	0	0
Fault6	1.0	0.34	1.0	0	0	0
Fault7	1.0	1.0	1.0	0	0	0
Fault8	1.0	1.0	1.0	0.1	0	0
Fault9	1.0	1.0	1.0	0.58	0.21	0.17
Fault10	1.0	1.0	0.37	0.06	0.01	0.02
Fault11	0.81	1.0	0.33	0.03	0.02	0.01
Fault12	0.92	0.48	0.98	0.05	0.04	0.01
Fault13	0.34	1.0	0.92	0.16	0.05	0.04
Fault14	0.82	1.0	0.89	0.15	0.61	0.05
Fault15	0.96	1.0	1.0	0.79	0.61	0.51
Fault16	0.84	1.0	0.98	0.69	0.62	0.48
Fault17	0.09	1.0	1.0	0.05	0.02	0.02
Fault18	0.96	1.0	1.0	0	0	0
Fault19	1.0	0.48	1.0	0	0	0
Fault20	1.0	1.0	1.0	0	0	0
Average	0.86	0.91	0.88	0.14	0.11	0.07

We find that the classification accuracy of different faults varies widely. The detection rate of faults 3 and 9 is less than 90%, and the detection rate of faults 15 and 16 is less than 50%. In addition, the detection rates of the remaining 16 faults are all higher than 90%, of which the detection rate of faults 1, 2, 4, 5, 6, 7, 8, 17, 19, and 20 is 100%. Therefore, PLSTM-CNN can effectively isolate most of the faults, and only a few faults perform poorly. Through the analysis of the data in the table, it is found that the lower accuracy of faults 3, 9, 15, and 16 is due to the higher degree of confusion among them, which is consistent with the previous mutual information calculation results [23]. The PLSTM-CNN model proposed in this paper makes full use of the local and global features of the original data, and the fault detection rate and false-positive rate are better than 2D-CNN and LSTM models in faults 3, 9, 15, and 16 with a high degree of confusion. Because both fault 3 and fault 9 are related to the

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

change in the initial number of articles, this paper rearranges the variables according to the method of the maximum mutual information coefficient based on the initial number of articles, which further improves the fault detection rate of faults 3 and 9. The degree of confusion of faults 15 and 16 is too high, and there is no effective fault detection method [24].

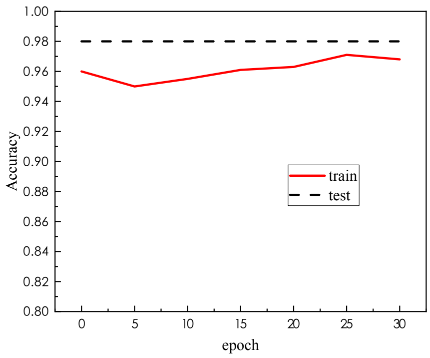

Figure 2. Recognition accuracy of learning rate self-enhancement algorithm

It can be seen from Fig.2 that when the number of iterations is 25, the accuracy rate is the highest, which is 97.3%, which shows that the use of PLSTM-CNN for English composition grammar detection has strong applicability. In addition, it can be seen from Figure 4 that as the iteration continues, the recognition accuracy of the DCNN model continues to rise. After reaching a certain level, the accuracy does not change and the network model converges. Therefore, the DCNN model used in this study has a good effect., it starts to converge after reaching a certain number of iterations. In this study, this algorithm is used to optimize DCNN with a certain preprocessing effect [25-27].

3.2. COMPARATIVE EXPERIMENT

To further illustrate the advantages of the parallel neural network structure, we did the following experiments on average fault detection rate, model training and testing time, and model stability under small sample conditions.

3.2.1. COMPARISON OF AVERAGE FAILURE DETECTION RATES

To further compare the fault detection capabilities of serial and parallel network models, based on the data set above, we designed a traditional serial network structure, as shown in Fig.3. The experimental results are shown in Fig.4. The average F1 scores on the PLSTM-CNN, tandem LSTM-CNN, LSTM, 1D-CNN, and 2D-CNN models are 92.13%, 89.54%, 84.08%, 84.80%, and 85.78%, respectively. This shows that parallel CLSTM-CNN has better fault detection performance than serial LSTM-CNN [28].

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

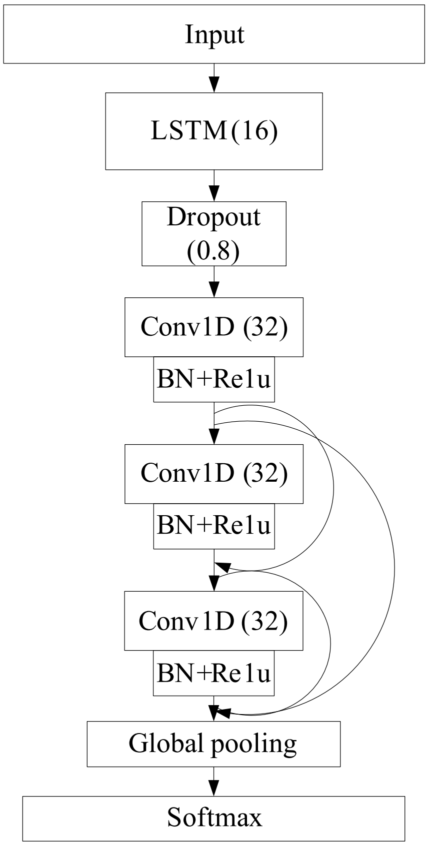

Figure 3. Serial LSTM-CNN network structure

The experimental results are shown in Fig.4. The average F1 scores on the PLSTM-CNN, tandem LSTM-CNN, LSTM, 1D-CNN, and 2D-CNN models are 92.13%, 89.54%, 84.08%, 84.80%, and 85.78%, respectively. This shows that parallel CLSTM-CNN has better fault detection performance than serial LSTM-CNN.

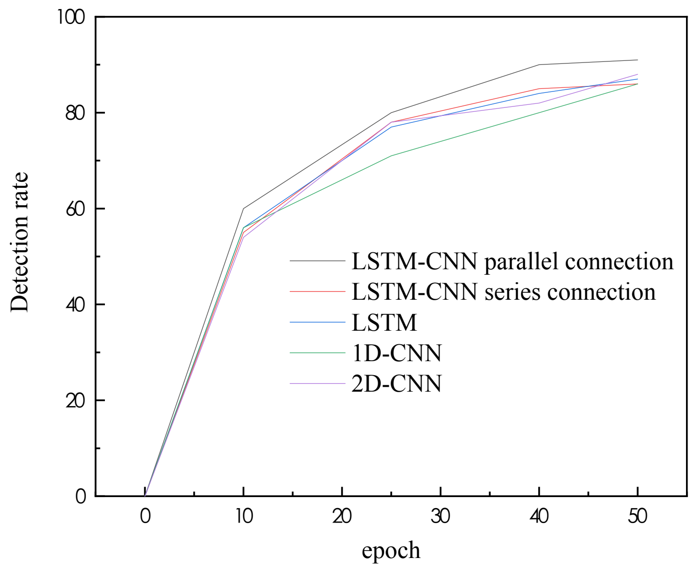

Figure 4. Mean failure detection rate

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

3.2.2. MODEL INFERENCE AND INFERENCE TIME COMPARISON

The PLSTM-CNN model takes 4.2 seconds to train for each epoch and 7 minutes to train for 100 epochs. The main reasons for its faster training speed are: considering the real-time nature of fault monitoring, the convolution layer in this paper is selected as one-dimensional convolution, which has fewer parameters. Under the conditions of the same network and hyperparameters, the calculation time is shorter, but Some accuracy will be lost.

Meanwhile, training deep 2D convolutional networks usually require special hardware devices, such as cloud computing or GPU acceleration. But 1D-CNN can be implemented on the CPU of ordinary computers, and its low computational requirements and compact structure are very suitable for real-time monitoring and low-cost applications.

Table 2. Comparison of training and inference time

Model	Training time for one epoch (s)	Reasoning time for one epoch (ms)
1D-CNN	2.54	10
2D-CNN	65	200
LSTM	3	12
PC LSTM-CNN	3.8	20
SC-LSTM-CNN	4.2	25

3.2.3. COMPARISON OF AVERAGE FAULT DETECTION RATES FOR SMALL SAMPLES

Considering that the actual English composition grammar detection fault samples are scarce, the experiment will reduce the number of each type of fault sample. Set the sampling time to 3 minutes, run for 10 hours under normal conditions, and collect 2000 normal samples. In the simulation of 20 kinds of faults, the simulator runs normally for 1 hour, then introduces the corresponding faults, and then continues to run for 1 hour. Thus, 1 hour of failure data (200 failure samples) was collected per simulation. The simulations for each failure type were repeated ten times with ten different initial states. The simulation platform collects a total of 4200 sample data, including 2000 normal samples and 2000 samples for each fault. Choose 70% of the data for training, 20% for testing, and 10% for validation. The experimental results are shown in Table 3:

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

Table 3. Average failure detection rate of small samples

Model	FDR	FPR
1D-CNN	83.4%	0.1
2D-CNN	78.5%	0.08
LSTM	84.6%	0.09
PC LSTM-CNN	84.8%	0.12
SC-LSTM-CNN	90.5%	0.051

The study found that SC-LSTM-CNN has a high FDR and a small value of FPR, which shows that the parallel LSTM-CNN has extremely high prediction accuracy and stability. Although PC-LSTM-CNN has a high FDR, its FPR value is high, indicating that its stability is poor and the prediction accuracy is average. The above results confirm that the parallel LSTM-CNN can still maintain high accuracy on small sample datasets, and its network structure is more stable than the serial network [29]. 2D-CNN requires a large number of training samples to guarantee the accuracy, while 1D-CNN still performs well on small-sample datasets[30-31].

4. CONCLUSION

This paper proposes a grammar writing check fault detection method based on the PLSTM-CNN network model. The model is constructed by LSTM, one-dimensional dense convolutional layer, one-dimensional global pooling layer, and Dropout layer, which can effectively extract the local and Global features; after data analysis and variable reordering based on the maximum information coefficient method, the data distribution is made more regular and easy to train. The study compares the fault detection results of PLSTM-CNN, tandem LSTM-CNN, LSTM, and 2D-CNN. The experimental results show that: (1) the fault detection accuracy and false positive rate of PLSTM-CNN are significantly better than other methods; (2) for the difficult-to-detect faults 3 and 9, the PLSTM-CNN model still performs well; (3) parallel Compared with the serial structure, the PLSTM-CNN structure has better accuracy and stability, and its FDR and FPR are 90.5% and 0.051, respectively.

5. CONFLICT OF INTEREST

The authors declared that there is no conflict of interest. REFERENCES

REFERENCES

(1) Tao, Y., Shi, H., Song, B., & Tan, S. (2020). **A Novel Dynamic Weight Principal Component Analysis Method and Hierarchical Monitoring Strategy for Process Fault Detection and Diagnosis**. *IEEE Transactions on Industrial Electronics*, (99), 1-1. https://doi.org/10.1109/TIE.2019.2942560

https://doi.org/10.17993/3ctecno.2023.v12n1e43.142-157

(2) A, C. W., A, X. W., Jz, A., Liang, Z. A., Xiao, B. A., Xin, N. B. Ehd, A. (2021). **Uncertainty Estimation for Stereo Matching Based on Evidential Deep Learning.** https://doi.org/10.1016/j.patcog.2021.108498

(3) Cai, W., Zhai, B., Liu, Y., Liu, R., & Ning, X. (2021). **Quadratic polynomial guided fuzzy C-means and dual attention mechanism for medical image segmentation**. *Displays*, 70, 102106. https://doi.org/10.1016/j.displa.2021.102106

(4) Ning, X., Duan, P., Li, W., & Zhang, S. (2020). **Real-time 3D face alignment using an encoder-decoder network with an efficient deconvolution layer**. *IEEE Signal Processing Letters*, 27, 1944-1948. https://doi.org/10.1109/LSP.2020.3032277

(5) Miao, J., Wang, Z., Ning, X., Xiao, N., Cai, W., & Liu, R. (2022). **Practical and secure multifactor authentication protocol for autonomous vehicles in 5G.** *Software: Practice and Experience*. https://doi.org/10.1002/SPE.3087

(6) Beruvides, G., Casta?O, F., Quiza, R., & Haber, R. E. (2016). **Surface roughness modeling and optimization of tungsten–copper alloys in micro-milling processes**. *Measurement*, 246-252. https://doi.org/10.1016/j.measurement.2016.03.002

(7) Chen, Y., Wang, L., Hu, J., & Ye, M. (2020). **Vision-Based Fall Event Detection in Complex Background Using Attention Guided Bi-directional LSTM**. https://doi.org/10.1109/ACCESS.2020.3021795

(8) Shan, W. (2022). **Digital streaming media distribution and transmission process optimisation based on adaptive recurrent neural network**. *Connection Science*, 34(1), 1169-1180. https://doi.org/10.1080/09540091.2022.2052264

(9) Yan, C., Pang, G., Bai, X., Liu, C., Xin, N., Gu, L., & Zhou, J. (2021). **Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss**. *IEEE Transactions on Multimedia*. https://doi.org/10.1109/TMM.2021.3069562

(10) Hu, X., Liu, T., Hao, X., & Lin, C. (2022). **Attention-based Conv-LSTM and Bi-LSTM networks for large-scale traffic speed prediction**. *The Journal of Supercomputing*, 1-24. https://doi.org/10.1007/s11227-022-04386-7

(11) Huang, Z., Wei, X., & Kai, Y. (2015). **Bidirectional LSTM-CRF Models for Sequence Tagging**. *Computer Science*. https://doi.org/10.48550/arXiv.1508.01991

(12) Jin, C., Shi, Z., Li, W., & Guo, Y. (2021). **Bidirectional LSTM-CRF Attention-based Model for Chinese Word Segmentation**. https://doi.org/10.48550/arXiv.2105.09681

(13) Ying, L., Nan, Z. Q., Ping, W. F., Kiang, C. T., Pang, L. K., Chang, Z. H. Nam, L. (2021). **Adaptive weights learning in CNN feature fusion for crime scene investigation image classification**. *Connection Science*. https://doi.org/10.1080/09540091.2021.1875987

(14) Li, D., & Lasenby, J. (2021). **Spatiotemporal Attention-Based Graph Convolution Network for Segment-Level Traffic Prediction**. *IEEE Transactions on Intelligent Transportation Systems*, (99), 1-9. https://doi.org/10.1109/TITS.2021.3078187

(15) Li, M., Liu, X., & Xiong, A. (2002). **Prediction of the mechanical properties of forged TC11 titanium alloy by ANN**. *Journal of Materials Processing Technology*, 121(1), 1-4. https://doi.org/10.1016/S0924-0136(01)01006-8

(16) Liu, Z., Zhou, W., & Li, H. (2019). **AB-LSTM: Attention-based Bidirectional LSTM Model for Scene Text Detection**. *ACM Transactions on Multimedia Computing Communications and Applications*, 15(4), 1-23. https://doi.org/10.1145/3356728

(17) Olave, M., Sagartzazu, X., Damian, J., & Serna, A. (2010). **Design of Four Contact-Point Slewing Bearing With a New Load Distribution Procedure to Account for Structural Stiffness**. *Journal of Mechanical Design*, 132(2), 021006. https://doi.org/10.1115/1.4000834

(18) Tang, D., Wei, F., Nan, Y., Ming, Z., & Bing, Q. (2014). **Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification**. *Paper presented at the Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics,* 1.

(19) Gao, Y., & Yu, D. (2020). **Total variation on horizontal visibility graph and its application to rolling bearing fault diagnosis**. *Mechanism and Machine Theory*, 147, 103768. https://doi.org/10.1016/j.mechmachtheory.2019.103768

(20) Nguyen, T. (2019). **Spatiotemporal Tile-based Attention-guided LSTMs for Traffic Video Prediction**. https://doi.org/10.48550/arXiv.1910.11030

(21) Sagnika, S., Mishra, B., & Meher, S. K. **An attention-based CNN-LSTM model for subjectivity detection in opinion-mining**. *Neural Computing and Applications*, 1-14. https://doi.org/10.1007/s00521-021-06328-5

(22) Shan, X., Wang, Y., Dong, M., & Xia, J. (2021). **Application Research and Analysis of Geographic Information System in Intelligent City Surveying and Mappinge**. *Journal of Physics: Conference Series*, 1881(4), 042071. https://doi.org/10.1088/1742-6596/1881/4/042071

(23) Shi, X., & Wang, B. (2021). **Application of New Surveying and Mapping Technology in the Construction of Smart City.** *E3S Web of Conferences*, 236, 04031. https://doi.org/10.1051/e3sconf/202123604031

(24) Shi, Z. L., Gong, Y., Cao, M., & Xiao, S. (2010). D**iscussion on the Application of Surveying and Mapping Technology in the Internet of Things Times**. *Modern Surveying and Mapping*, 65(4), 503-515. https://doi.org/10.1016/j.neuron.2010.01.035

(25) Andrejic, M., Bojovic, N., & Kilibarda, M. (2016). **A framework for measuring transport efficiency in distribution centers**. *Transport Policy*, 45(JAN.), 99-106. https://doi.org/10.1016/j.tranpol.2015.09.013

(26) Bergstrom, J. C., Braden, J. B., & Kolstad, C. D. (1991). **Measuring the demand for environmental quality**. *American Journal of Agricultural Economics*, 75(1), 244. https://doi.org/10.2307/1242975

(27) Brock, W. A., & Taylor, M. S. (2005). **Economic Growth and The Environment: A Review of Theory and Empirics**. *Handbook of Economic Growth*. https://doi.org/10.1016/S1574-0684(05)01028-2

(28) Wang, M., Zhou, J., Gao, J., Li, Z., & Li, E. (2020). **Milling Tool Wear Prediction Method Based on Deep Learning under Variable Working**

Conditions. *IEEE Access*, 99, 1-1. https://doi.org/10.1109/ACCESS.2020.3010378

(29) Amin, T., Khan, F., Ahmed, S., & Imtiaz, S. (2020). **A novel data-driven methodology for fault detection and dynamic risk assessment**. *The Canadian Journal of Chemical Engineering*. https://doi.org/10.1002/cjce.23760

(30) Horani M. O., Najeeb, M., y Saeed, A. (2021). **Model electric car with wireless charging using solar energy**. *3C Tecnología. Glosas de innovación aplicadas a la pyme*, 10(4), 89-101. https://doi.org/10.17993/3ctecno/2021.v10n4e40.89-101

(31) Chang Jingying,Lan Weibin & Lan Wenhao. (2021). **Higher education innovation and reform model based on hierarchical probit**. *Applied Mathematics and Nonlinear Sciences, 7*(1), 175-182. https://doi.org/10.2478/AMNS.2021.2.00154