

Research Article

A Stock Trend Forecast Algorithm Based on Deep Neural Networks

Yingying Yan and Daguang Yang 

Business School of Northeast Normal University, Changchun, Jilin 130000, China

Correspondence should be addressed to Daguang Yang; yangdg797@nenu.edu.cn

Received 26 April 2021; Accepted 17 June 2021; Published 5 July 2021

Academic Editor: Shah Nazir

Copyright © 2021 Yingying Yan and Daguang Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a recognized complex dynamic system, the stock market has many influencing factors, such as nonstationarity, nonlinearity, high noise, and long memory. It is difficult to explain it simply through mathematical models. Therefore, the analysis and prediction of the stock market have been a very challenging job since long time. Therefore, this paper adopts an encoder-decoder model of attention mechanism, adding attention mechanism from two aspects of feature and time. Both encoder and decoder use LSTM neural network. This method solves two problems in time series prediction; the first problem is that multiple input features have different degrees of influence on the target sequence, the feature attention mechanism is used to deal with this problem, and the weights of different input features can be obtained. A more robust feature association relationship is obtained; the second problem is that the data before and after the sequence have a strong time correlation. The time attention mechanism is used to deal with this problem, and the weights at different time points can be obtained to obtain more robustness and good timing dependencies. The simulation and experimental results show that the introduction of the attention mechanism can obtain lower forecast errors, which proves the effectiveness of the model in dealing with stock forecasting problems.

1. Introduction

The high yield of the stock market has attracted the majority of investors [1], making stock investment [2, 3] become one of the most common ways of investment and financial management. In order to avoid the high risk that comes with high returns, investors are in tireless pursuit of accurate analysis and prediction of the stock market. The amount of trading data generated by daily stock trading is considered to reflect the actual situation of the market, which is often used by investors to analyze and forecast the market. However, because the stock market is affected by many factors such as market supply and demand, exchange rate, company operating conditions, policy changes, and market interest rates, stock prices show the characteristics of fluctuation, which makes the analysis and prediction of stock prices [4–6] face great difficulties.

As a recognized complex dynamic system, the stock market has many influencing factors, such as nonstationarity, nonlinearity, high noise, and long memory. It is

difficult to explain it simply through mathematical models. Therefore, the analysis and prediction of the stock market have been a very challenging job since very long time. Random walk theory [7] believes that the fluctuation of stock prices is completely random, and there is no rule to follow. However, many researchers have found a certain rule in the fluctuation of stock prices, which shows that the stock market does have its own specific rules of operation, which lays the foundation for stock price prediction. Traditional stock market forecasting methods include fundamental analysis, technical analysis, multiple regression, and autoregressive moving translation method (ARIMA) [8]. The above methods are usually used for the simulation analysis of stationary or linear time series. The upper limit of the data introduced is small, and the data need preprocessing such as difference to smooth the nonstationary series. Therefore, traditional analysis and forecasting methods have certain discomforts in stock market forecasting [9–12].

The advent of the big data era and the continued advancement of artificial intelligence technology in recent

years [13–16], which has been gradually applied to various fields, brings profound changes to various industries. More and more experts and scholars begin to study the neural network [17–20] and apply it to the prediction of stock and other time series. Neural network has strong self-learning ability, can deal with massive data, and can make more accurate prediction for unstable and nonlinear time series. A long time ago, some researchers used BP neural network to study stock prediction related problems and achieved good prediction results. However, due to the disadvantage that BP neural network does not contain the concept of time sequence when dealing with problems, the prediction effect is still limited. However, the cyclic neural network solves this problem, it has the concept of time series, and it has better performance for stock prediction. However, the cyclic neural network cannot solve the problem of long-term dependence; so on the basis of its structural improvement, LSTM neural network is proposed. LSTM neural network has a great advantage in dealing with time series prediction and further improves the performance of stock prediction. In order to predict the stock price of the next day, the attention mechanism is added into the LSTM neural network and further optimized. It is hoped that we can further improve the handling ability of stock price prediction [21] and obtain better forecast results to meet the needs of current investors.

Following are the main innovations points of this paper:

- (1) This paper proposes an encoder-decoder model of attention mechanism and adds attention mechanism from two aspects of feature and time.
- (2) This paper uses the time attention mechanism to deal with this problem, which can get the weights of different time points and at the same time obtain the more robust timing dependence.
- (3) The simulation experiment results show that the introduction of the attention mechanism can obtain lower prediction errors, which proves the effectiveness of the model in dealing with stock prediction problems.

The organization of the paper is as follows. Section 2 discusses the related work to the proposed research. Methodology of the paper is given in Section 3 with details of the work done in the proposed research. Experiments and results are given in Section 4. The paper is concluded in Section 5.

2. Related Work

Some scholars used neural network technology to study financial time series in the 1990s and predicted the daily rate of return of IBM stocks. However, due to the gradient explosion problem of the traditional BP neural network [22, 23], the result will converge to a local minimum. With the advent of the big data era and the widespread application of deep learning, many scholars have also tried to apply the newly proposed recurrent neural network (RNN) model and its improved model LSTM model in financial research. In

particular, LSTM has become the first choice of many researchers due to its unique potential in time series modeling.

Siami and Namin [24] comparatively studied the two financial time series analysis methods, ARIMA and LSTM, and the results show that the LSTM model is 85% more accurate than the ARIMA model. Skehin et al. [25] used wavelet analysis to denoise the stock price data of the five groups listed on Nasdaq: Facebook, Apple, Netflix, Alphabet, and Amazon and established ARIMA models. Compared with the LSTM model, the prediction performance of these two models is compared. Xiong et al. [26] selected Google domestic trends, which can represent macroeconomic factors and public psychology, as input indexes, and used the neural network of long- and short-term memory to study the impact of the above indexes on the volatility of the S & P 500 index from 2004 to 2015. Fischer and Krauss [27] applied the LSTM network to predict the out-of-sample changes of the constituent stocks of the S & P 500 index from 1992 to 2015. The prediction results show that the daily return rate of 0.46% can be obtained by using the model, and the annualized Sharpe rate is 5.8.

In addition, Baek and Kim [28] proposed the MOD AUG NET framework, which consists of overfitting prevention LSTM module and prediction LSTM module, to build the model. The validity of the model is evaluated by using two different representative stock market data (S & P 500 and KOSPI200). The results show that the test error of MOD-AUGNET C is lower than that of LSTM prediction model alone. Yao et al. [29] used the LSTM network to build a short-term stock price change model and tested it by experimenting with some stocks randomly selected from the 300 Shanghai and Shenzhen stocks. Experiments show that the accuracy, recall rate, and critical error of LSTM are better than random prediction.

The application of the deep learning method to prediction [30] can improve the prediction accuracy of the model, but it does not mean that deep learning method has an inevitable relationship with high accuracy. Compared with traditional methods, this method has many unique advantages in the analysis and prediction of financial time series. However, when facing different practical financial problems, the specific characteristics of different financial problems should be considered to select the application.

3. Methodology

3.1. Attention Mechanism. Attention mechanism was originally used in machine translation and has now become an important method in the field of neural networks. In the field of artificial intelligence, the attention mechanism has become an important part of its structure and has a large number of applications in many fields such as natural language processing, time series prediction, speech, and computer. The attention mechanism is very similar to the human observation mechanism of external things. When we observe external things, we usually only pay attention to a certain part of them first, and everyone has different concerns. For example, when we observe a person, the first thing we notice may be his face, height, body shape, clothing, and

so on. After obtaining the information of each part, you can arrange and combine them and finally get the overall feeling of the person. The attention mechanism generally consists of two parts of tasks, determining which inputs require more attention and extracting features of key parts to obtain important information. Therefore, the attention mechanism can give different weights to each part of the input to filter the main features, which is why the attention mechanism is so widely used.

When processing the input, the calculation of the attention value is the most critical of all operations. Calculating the attention value usually requires two major operations. The first is to use all the inputs to calculate the attention weight and then calculate the weighted average of all inputs based on it.

First, record all the inputs as a whole as X and then select a query variable as q . The function of q is to find and select part of the information in the whole X . Here, a soft attention mechanism is used to select all inputs, give more attention to important inputs, and give less attention to less important inputs.

There are a total of N inputs, q is the query variable, and s is the attention calculation function, and then the attention weight of the i -th input can be expressed as follows:

$$a_i = \text{softmax}(s(x_i, q)) = \frac{\exp(s(x_i, q))}{\sum_{j=1}^N \exp(s(x_j, q))}. \quad (1)$$

There are many kinds of attention calculation functions. The following four methods are introduced:

$$\begin{aligned} s(x_i, q) &= v^T \tanh(Wx_i + Uq), \\ s(x_i, q) &= x_i^T q, \\ s(x_i, q) &= \frac{x_i^T q}{\sqrt{d}}, \\ s(x_i, q) &= x_i^T Wq. \end{aligned} \quad (2)$$

These four methods are called additive method, dot product method, scaled dot product method, and bilinear method where W , U , and v are the network parameters that need to be learned and d is the dimension of the input information.

After getting the attention weight, the attention value can be calculated by weighted average:

$$\text{att}(X\Delta q) = \sum_{i=1}^N a_i x_i. \quad (3)$$

3.2. Encoder-Decoder. The encoder-decoder model is a method to realize the problem of mapping the input sequence to the output sequence. It consists of two parts: an encoder and a decoder. Encoders and decoders can use neural networks such as CNN, RNN, LSTM, and GRU as shown in Figure 1.

The main purpose of the encoder is to extract features from the input time series data. The encoder extracts the information and encodes it into an intermediate vector as

the input of the decoder. In the decoder, the intermediate vector is decoded and combined with the input data at the current moment to predict the sequence data at the next moment.

In the encoder-decoder model, the intermediate vector corresponding to each encoder is the same, which means that each feature vector in the input sequence has the same effect on each feature vector in the output sequence. This will cause two problems. For a sequence with a huge amount of information, the intermediate vector cannot fully display it, and the information also has a first-come-first-served relationship; that is, the newly added information will dilute the information before it. This phenomenon is particularly prominent in long time series. This makes the decoder unable to obtain sufficient information from the input sequence when decoding, and the accuracy of the prediction will also be affected.

To solve the above problems, some scholars have proposed to use the attention mechanism to optimize the encoder-decoder model. There is not only one intermediate vector in the model. The increase in intermediate vectors enables the model to optimize the attention mechanism. The selection of the input sequence is realized, and the final output is obtained by learning according to the selection information. The model structure is shown in Figure 2.

First, the input sequence X is passed to the recurrent neural network of the encoder to calculate the hidden state as follows:

$$s_t = f_{\text{RNN}}(s_{t-1}, x_t). \quad (4)$$

After calculating the hidden state at each moment, they need to be synthesized and stored in the form of an intermediate vector as follows:

$$c = q(s_1, s_2, \dots, s_T). \quad (5)$$

The decoding process is actually the inverse operation of the encoding process, mainly using the intermediate vector and the output y_1 at the previous moment to predict y_2 to be output. The recurrent neural network of the encoder first uses the predicted output at the previous time, the hidden state at the previous time, and the intermediate vector to calculate the hidden state at the current time as follows:

$$h_t = f_{\text{RNN}}(y_{t-1}, h_{t-1}, c). \quad (6)$$

The predicted output at the previous moment, the hidden state at the current moment, and the intermediate vector are passed to a multilayer perceptron g , and the predicted output at the current moment is calculated as follows:

$$y_t = g(y_{t-1}, h_t, c). \quad (7)$$

3.3. LSTM Based on Attention Mechanism. Time series forecasting mainly faces two major problems. The first is that the target sequence in the time sequence is affected by multiple input feature sequences. Therefore, the influence of multiple input features on the target sequence changes over time, and the degree of influence of different input features is

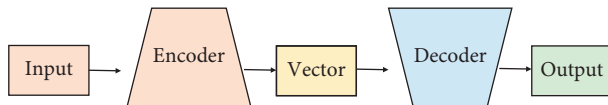


FIGURE 1: Encoder-decoder model.

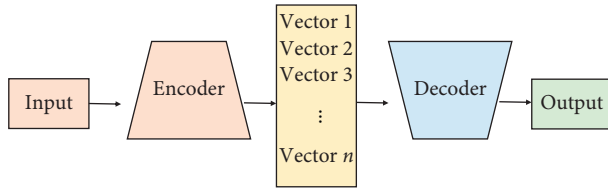


FIGURE 2: Encoder-decoder model of attention mechanism.

different. The second is the time correlation of the time series. The data before and after each series have a strong time correlation. Within a certain period of time, there is a strong interaction between the series. These time variables have different effects on the target sequence. When using traditional models for analysis, the impact of this interaction will be ignored, resulting in insufficient generalization capabilities of the model.

To solve the abovementioned problems in time series prediction, the attention mechanism can be added to the encoder and decoder parts of the encoder-decoder model. The feature attention mechanism is introduced into the encoder to calculate the attention weight of the input feature at the current moment. This weight indicates the importance of the input feature to the current target task. The sum of the attention weights of all input features is equal to 1, which completes the enhancement of key information and dilutes the information of general importance. The original input features are weighted according to the size of the attention weight, and the intermediate vector representing the updated information is obtained. A time attention mechanism is introduced into the decoder, and a neural network is used to comprehensively process the time sequence of the intermediate vector and input information at the current moment, thereby obtaining the predicted output at the next moment.

4. Experiments and Results

4.1. Hyperparameter Settings. The LSTM neural network model based on attention mechanism proposed in the previous section is referred to as ATT-LSTM model. In the prediction research of stock index in this section, the input characteristics of the LSTM model and ATT-LSTM model are the closing price, opening price, maximum price, minimum price, rising/falling price, and trading volume of the stock index, and the output is the predicted closing price of the next trading day. The number of neurons in the input and output layers is 6 and 1, respectively. In this section, the research on predicting the closing price of individual stocks is also added. In addition to the six features used in predicting the stock index above, the input features of the model are also added: the turnover ratio of individual stocks, volume ratio, price-earnings ratio, price-to-book ratio,

price-to-sales ratio, and total market value. The number of neurons in the input layer and output layer is 12 and 1, respectively. In the study of stock index and individual stock, the number of neurons in the hidden layer of both models is 128. ReLU function is used for activation function, MSE is used for loss function, and the Adam algorithm is used for the model learning algorithm. In order to verify the influence of time step on the predicted results, the LSTM model and ATT-LSTM encoder and decoder used LSTM time step which are set to groups for comparative analysis.

4.2. Experimental Results. The forecast results of the closing price of the Shanghai and Shenzhen 300 Index in this paper are fitted into a graph as shown in Figure 3 (the abscissa is time, the ordinate is stock index prices, the blue curve is the predicted value, and the yellow curve is the true value). Observing the curve of the fitted graph and comparing the predicted result with the actual result, we can find the following.

- (1) The predicted value of the forecast model for the closing price of the Shanghai and Shenzhen 300 Index is basically consistent with the true value in the overall trend. The predicted value is very close to the actual value, and the degree of deviation is low. However, it can be clearly seen from Figure 3 that during the period of the 2015 stock market crash, although the predicted value of the model is basically the same as the real value in trend, there is a large gap between the predicted value and the actual value. Considering that China's stock market is a developing policy market, the trading mechanism is immature and easily affected by policies, so the model has a certain degree of failure.
- (2) There is a certain lag between the predicted value of the model and the actual value, which is manifested as the right deviation of the overall true value curve of the predicted value curve. The predicted value in the rising interval is mostly lower than the true value, and the predicted value in the falling interval is mostly higher than the true value. Since the prediction of the model is usually based on past historical data, there will inevitably be a lag in the prediction results whether the lagging line affects the forecast results and the elimination of lagging still needs to be studied in depth.

In addition, the model trained with 160 stocks of the CSI 300 is recorded as M1, the model trained with 16 stocks of the banking category is recorded as M2, and the model trained with 14 stocks of the securities category is recorded as M3. It can be seen from Figures 4 and 5 that the M1 and M3 models have converged after about 100 iterations. From Figure 6, it can be seen that the M2 model converges after about 80 iterations. In the subsequent iterative training process, the loss function value of the verification set begins to oscillate, indicating that the model has overfitting. Therefore, when testing the model with the test set, the checkpoint model with the minimum loss function is selected.

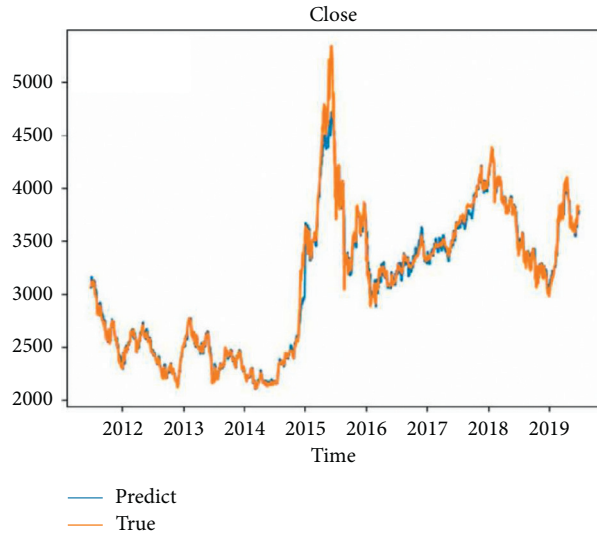


FIGURE 3: 2011–2019 Shanghai and Shenzhen 300 index forecast result and actual value fitting curve.

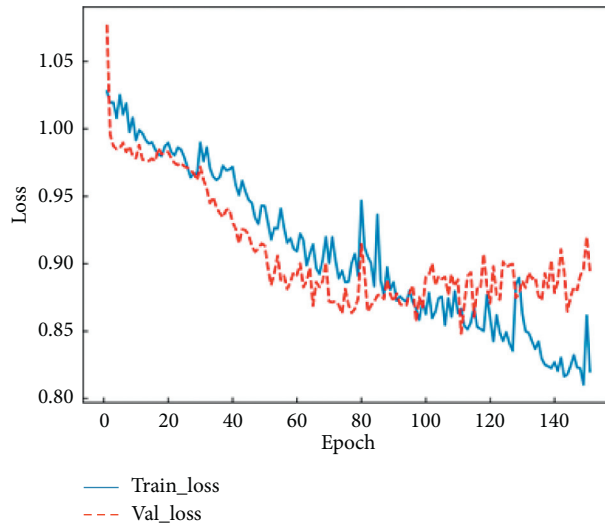


FIGURE 4: M1 model loss function.

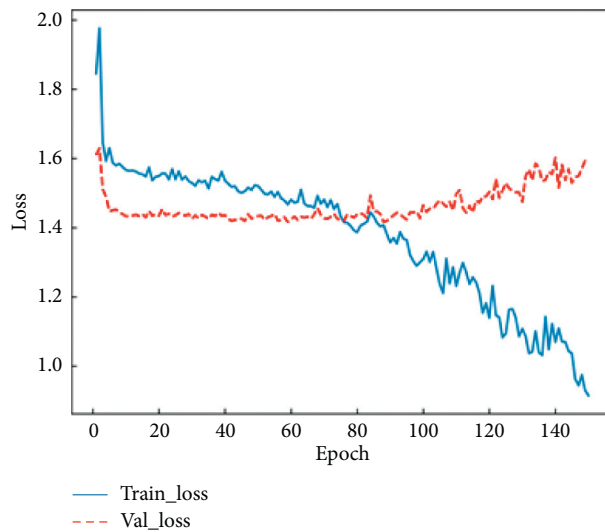


FIGURE 5: M2 model loss function.

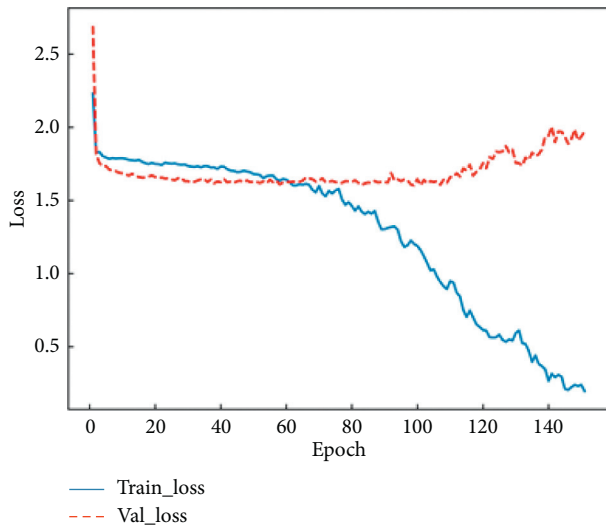


FIGURE 6: M3 model loss function.

5. Conclusion

In this paper, we adopt an encoder-decoder model of attention mechanism, adding attention mechanism from both features and time. Both encoder and decoder use LSTM neural network. This method solves two problems in time series prediction. The first problem is that multiple input features have different degrees of influence on the target sequence. The feature attention mechanism is used to deal with this problem, and the weights of different input features can be obtained. A more robust feature association relationship is obtained; the second problem is that the data before and after the sequence have a strong time correlation. The time attention mechanism is used to deal with this problem, and the weights at different time points can be obtained to obtain more robustness and good timing dependencies. The simulation experiment results show that the introduction of the attention mechanism can obtain lower forecast errors, which proves the effectiveness of the model in dealing with stock forecasting problems.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] S. Tuysuz, "Conditional correlations between stock index, investment grade yield, high yield and commodities (gold and oil) during stable and crisis periods," *International Journal of Economics and Finance*, vol. 5, no. 9, pp. 28–44, 2013.
- [2] Y. Yuniningsih, S. Widodo, and M. B. N. Wajdi, "An analysis of decision making in the stock investment," *Economic Times: Journal of Economic and Islamic Law*, vol. 8, no. 2, pp. 122–128, 2017.

- [3] L. Kengatharan and N. Kengatharan, "The influence of behavioral factors in making investment decisions and performance: study on investors of colombo stock exchange, Sri Lanka," *Asian Journal of Finance & Accounting*, vol. 6, no. 1, p. 1, 2014.
- [4] L. Zhang, F. Wang, B. Xu, W. Chi, Q. Wang, and T. Sun, "Prediction of stock prices based on LM-BP neural network and the estimation of overfitting point by RDCI," *Neural Computing & Applications*, vol. 30, no. 5, pp. 1425–1444, 2018.
- [5] R. Batra and S. M. Daudpota, "Integrating stocktwits with sentiment analysis for better prediction of stock price movement," in *Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1–5, IEEE, Sukkur, Pakistan, March 2018.
- [6] S. C. Nayak, B. B. Misra, and H. S. Behera, "ACFLN: artificial chemical functional link network for prediction of stock market index," *Evolving Systems*, vol. 10, no. 4, pp. 567–592, 2019.
- [7] S. Rehman, I. U. Chhapra, M. Kashif, and R. Rehan, "Are stock prices a random walk? an empirical evidence of asian stock markets," *ETIKONOMI*, vol. 17, no. 2, pp. 237–252, 2018.
- [8] M. Afeef, A. Ihsan, and H. Zada, "Forecasting stock prices through univariate ARIMA modeling," *NUML International Journal of Business & Management*, vol. 13, no. 2, pp. 130–143, 2018.
- [9] S. Carta, A. Ferreira, A. S. Podda, R. D. Reforgiato, and A. Sanna, "Multi-DQN: an ensemble of deep Q-learning agents for stock market forecasting," *Expert Systems with Applications*, vol. 164, Article ID 113820, 2021.
- [10] R. Efendi, N. Arbaiy, and M. M. Deris, "A new procedure in stock market forecasting based on fuzzy random auto-regression time series model," *Information Sciences*, vol. 441, pp. 113–132, 2018.
- [11] A. H. Bukhari, M. A. Z. Raja, M. Sulaiman, S. Islam, M. Shoaib, and P. Kumam, "Fractional neuro-sequential ARFIMA-LSTM for financial market forecasting," *IEEE Access*, vol. 8, Article ID 71326, 2020.
- [12] S. P. Chatzis, V. Siakoulis, A. Petropoulos, E. Stavroulakis, and N. Vlachogiannakis, "Forecasting stock market crisis events using deep and statistical machine learning techniques," *Expert Systems with Applications*, vol. 112, pp. 353–371, 2018.
- [13] C. Yan, G. Pang, X. Bai, J. Zhou, and L. Gu, "Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss," *IEEE Transactions on Multimedia*, 2021, In press.
- [14] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2020, In press.
- [15] X. Ning, Y. Wang, W. Tian, L. Liu, and W. Cai, "A biomimetic covering learning method based on principle of homology continuity," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 9–16, 2021.
- [16] Z. Huang, P. Zhang, R. Liu, and D. Li, "Immature apple detection method based on improved Yolov3," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 9–13, 2021.
- [17] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, In press.
- [18] J. Zhang, Y. Liu, H. Liu, and J. Wang, "Learning local-global multiple correlation filters for robust visual tracking with kalman filter redetection," *Sensors*, vol. 21, no. 4, p. 1129, 2021.

- [19] Z. Chu, M. Hu, and X. Chen, "Robotic grasp detection using a novel two-stage approach," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 19–29, 2021.
- [20] X. Ning, X. Wang, S. Xu et al., "A review of research on co-training," *Concurrency and Computation: Practice and Experience*, 2021, In press.
- [21] J. M. T. Wu, Z. Li, N. Herencsar, B. Vo, and J. C. W. Lin, "A graph-based CNN-LSTM stock price prediction algorithm with leading indicators," *Multimedia Systems*, pp. 1–20, 2021, In press.
- [22] L. Huang, G. Xie, W. Zhao, Y. Gu, and Y. Huang, "Regional logistics demand forecasting: a BP neural network approach," *Complex & Intelligent Systems*, pp. 1–16, 2021, In press.
- [23] G. Zhao, Y. Zhang, Y. Shi, H. Lan, and Q. Yang, "The application of bp neural networks to analysis the national vulnerability," *Computers, Materials & Continua*, vol. 58, no. 2, pp. 421–436, 2019.
- [24] N. S. Siami and A. S. Namin, "Forecasting economics and financial time series: ARIMA vs. LSTM," 2018, <http://arxiv.org/abs/1803.06386>.
- [25] T. Skehin, M. Crane, and M. Bezbradica, "Day ahead forecasting of FAANG stocks using ARIMA, LSTM networks and wavelets," in *Proceedings of the 26th AIAI Irish Conference on Artificial Intelligence and Cognitive Science*, Dublin, Ireland, December 2018.
- [26] R. Xiong, E. P. Nichols, and Y. Shen, "Deep learning stock volatility with google domestic trends," 2015, <http://arxiv.org/abs/1512.04916>.
- [27] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, 2018.
- [28] Y. Baek and H. Y. Kim, "ModAugNet: a new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module," *Expert Systems with Applications*, vol. 113, pp. 457–480, 2018.
- [29] S. Yao, L. Luo, and H. Peng, "High-frequency stock trend forecast using LSTM model," in *Proceedings of the 2018 13th International Conference on Computer Science & Education (ICCSE)*, pp. 1–4, IEEE, Colombo, Sri Lanka, 2018, August.
- [30] W. Sun, P. Zhang, Z. Wang, and D. Li, "Prediction of cardiovascular diseases based on machine learning," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 30–35, 2021.