Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD28972

Research on Stock Price Prediction Using Target Object Images Based on LSTM

Fangruo Wang

Hebei GEO University, Shijiszhuang, China
18339267297@163.com

Abstract. In recent years, the integration of computer vision and deep learning in the financial sector has become a research hotspot. Traditional stock price prediction primarily relies on time-series data, while individual investors in the market often make decisions by analyzing visual charts such as candlestick charts. Consequently, research utilizing stock price-related images as input has gradually emerged. This paper proposes an architecture based on the Long Short-Term Memory (LSTM) network to predict future trends using stock price images. By converting stock price information—including opening price, high price, low price, closing price into standardized images.Then employs the enhanced LSTM structure proposed in Sequencer (bidirectional LSTM, Bi-LSTM) for feature extraction and modeling, using machine learning to simulate human investment decisions. Experimental results demonstrate that this model not only outperforms traditional momentum strategies and short-term reversal strategies in stock price prediction accuracy but also maintains robust performance across varying market conditions and transaction delays. This research offers novel insights for stock price forecasting and validates the effectiveness of LSTM in processing image-based financial data.

Keywords: LSTM, Bi-LSTM, image prediction, stock price prediction, computer vision, deep learning

1. Introduction

Stock price forecasting has long been a significant research focus in finance. Traditional methods primarily rely on time series analysis or technical indicators, yet these approaches struggle to capture the intricate spatial patterns inherent in stock price data [1,2]. With the advancement of deep learning, computer vision techniques have begun to be applied to financial data processing.As a recurrent neural network, LSTM excels at handling long-term dependencies in sequential data. The Sequencer architecture has proven that LSTM can effectively capture spatial information through vertical and horizontal decomposition in image classification [1,3].

This paper constructs an image feature extraction module based on Bi-LSTM and integrating the spatial structure of stock price images, making predictions of future stock price trends. This research not only enriches stock price prediction methodologies but also provides empirical support for cross-domain applications of LSTM in computer vision and finance.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD28972

2. Related theories and concepts

2.1. Construction methods for stock price images

Using Jiang J et al.'s pioneering work on financial visualization for reference, this study employs OHLC (Open-High-Low-Close) bar charts combined with volume information [4]. This approach transforms stock data into standardized images (80×60 pixel OHLC charts), faithfully recreating the visual context of human analysis.

To ensure comparability across different time periods, this study employs a minimum-maximum normalization method to map price data to the 0–1 range. Daily stock prices—opening, high, low, and closing—are represented by OHLC bars [4]. The vertical axis of the image is normalized based on the price range, calculated as follows:

$$\widehat{p} = \frac{p - \min(p_{high}, p_{low})}{\max(p_{high}, p_{low}) - \min(p_{high}, p_{low})}$$

2.2. LSTM and its improved structures

2.2.1. Bidirectional LSTM (BiLSTM)

Long Short-Term Memory (LSTM) is a specialized type of Recurrent Neural Network (RNN) proposed by Hochreiter and Schmidhuber in 1997 [5]. By introducing sophisticated gating mechanisms(input gate, forget gate, output gate), LSTMs effectively capture long-term dependencies in time series data—a critical capability for tasks like stock price forecasting that involve extended temporal correlations.

Bidirectional LSTM(BiLSTM) [6] enhances the model's sequence modeling capability by processing sequence data simultaneously in both forward and backward directions. BiLSTM comprises two independent LSTM layers: one processes the input sequence chronologically, while the other processes it in reverse chronological order. The final output is the concatenation of the hidden states from both directions:

$$h_t = \left[h_t^{\text{forward}}, h_t^{\text{backward}}\right]$$

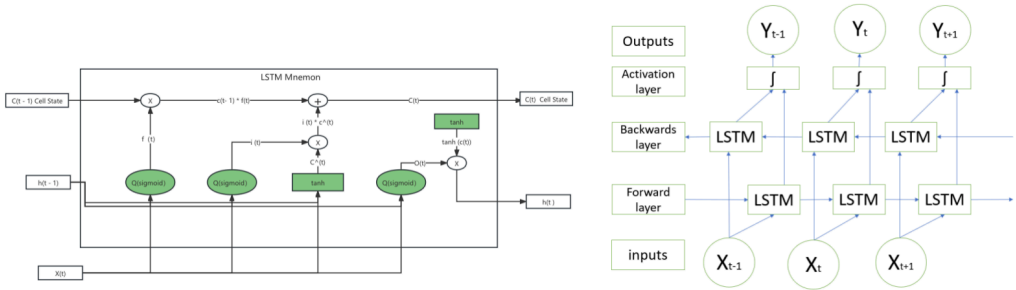


Figure 1. LSTM and bidirectional LSTM schematic diagram [7]

This structure enables the model to simultaneously leverage past and future contextual information, proving particularly effective for capturing complex patterns in stock price data.

DOI: 10.54254/2755-2721/2025.LD28972

2.2.2. 2D-LSTM enhancement in the sequencer architecture

This paper adopts the 2D-LSTM enhancement structure proposed by the Sequencer architecture [1], applying traditional LSTMs to image data processing. The core innovation of Sequencer lies in decomposing LSTMs into vertical and horizontal dimensions to process row and column information respectively:

$$\text{Vertical LSTM processing:} \mathrm{H}^{\mathrm{ver}}_{:,\mathrm{w},:} = \mathrm{BiLSTM}(\mathrm{X}_{:,\mathrm{w},:})$$

$$\text{Horizontal LSTM processing:} \mathrm{H}^{hor}_{h,:,:} = \mathrm{BiLSTM}(\mathrm{X}_{h,:,:})$$

$$\texttt{Featurefusion}: \mathtt{H} = \texttt{concatenate}(\mathtt{H}^{\mathrm{ver}}, \mathtt{H}^{\mathrm{hor}}), \widehat{\mathtt{X}} = \texttt{FC}(\mathtt{H})$$

This decomposition strategy reduces sequence length,enhancing computational efficiency while simultaneously capturing spatial fluctuations and temporal trends in prices [1,8].

2.3. Traditional stock price forecasting strategies

Momentum Strategy (MOM) predicts future trends based on average returns over the past 2–12 months. This study determines long or short positions for the next cycle using returns from the preceding year (252 trading days). Short-Term Reversal Strategy (STR) measures returns from the preceding month [9].Weekly Short-Term Reversal (WSTR) is based on returns from the preceding week [1]. These strategies serve as benchmarks in this study to evaluate the performance of the proposed model.

3. Design of an LSTM-based stock price image prediction model

3.1. Overall model architecture

The model takes stock price images as input. First, local features are extracted through a convolutional layer similar to the VGG model [1]. These features are then fed into a 2D-LSTM module (parallel processing via vertical BiLSTM and horizontal BiLSTM) to capture spatio-temporal features [1]. Finally, a fully connected layer outputs the predicted probability of stock price increase or decrease [1,10]. Unlike CNNs, the LSTM module excels at capturing long-term dependencies in price trends within images through its recursive structure [1,11].

3.2. 2D-LSTM feature extraction

The vertical BiLSTM processes the vertical sequence information of each column in the image, while the horizontal BiLSTM handles the horizontal sequence information of each row.This architecture borrows the spatial dimension decomposition design from Sequencer2D, reducing sequence length and improving computational efficiency [1], similar to the parallel processing of horizontal axes in ViP [12].

The image generation algorithm designed in this paper converts OHLC data from each time window into an 80×60 pixel RGB image. The image data structure employs a three-dimensional

DOI: 10.54254/2755-2721/2025.LD28972

array representation, where vertical axis, representing the price range and horizontal axis, representing the time series length.

For price-to-pixel mapping, this study initially define a mapping function from price values to image pixel positions. Then, using the normalization function employed in Jiang J's research [4], we map the normalized price values (within the 0-1 range) to vertical image coordinates.

4. Experiments and results analysis

4.1. Dataset and experimental setup

The dataset comprises CSI 300 Index data from 2014 to 2024, spanning 2,676 trading days. Stock price images (80×60 pixels) were constructed using 5-day, 20-day, and 60-day windows, yielding 2,597 valid images as shown in Figure 2. The dataset was divided into training (70%) and validation (30%) sets. Hierarchical sampling at a 7:2:1 ratio yielded training, validation and test sets, ensuring approximately 52% up labels across all sets. Training employed the AdamW optimizer (learning rate 1e-4) with random depth [13], label smoothing [4] regularization, and image normalization augmentation. Early stopping was set with a patience value of 15.



Figure 2. Example stock price images

4.2. Model training process analysis

After 20 training epochs, the training loss decreased from 0.6911 to 0.6048, the validation loss dropped from 0.6895 to 0.6258, and the validation accuracy rose from 57.20% to 66.54%.

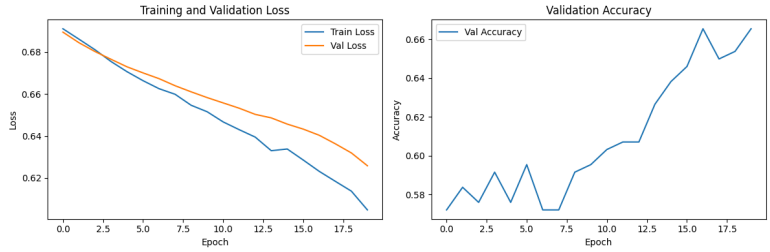


Figure 3. Training curves

Furthermore, validation accuracy progressively improved with training iterations, rising at round 1 to 66.54% at round 20. A critical acceleration occurred between rounds 14–17, where accuracy jumped from 59.14% to 66.54%. After the 17th iteration, accuracy enters a plateau phase, indicating the model approaches optimal fitting.

DOI: 10.54254/2755-2721/2025.LD28972

4.3. Model classification performance evaluation

4.3.1. Test set classification analysis

Test set classification results are shown in Table 1, with the model achieving an overall accuracy of 65.20%, significantly exceeding the random level of 50%. Category-specific metrics reveal a recall rate of 77% for the "Up" category, indicating strong trend identification capability for upward movements. The precision rate of 67% for the "Down" category demonstrates low misclassification rates for downward signals.

Table 1. Detailed classification report

	Precision	Recall	F1-Score	Support
Down(0)	0.67	0.53	0.59	249
Up(1)	0.64	0.77	0.70	274
Accuracy	-	-	0.65	523
Macro Avg	0.66	0.65	0.64	523
Weighted Avg	0.66	0.65	0.65	523

4.3.2. Performance comparison across different image lengths

To validate the impact of image length on model performance, experiments compared classification accuracy using 5-day, 20-day, and 60-day images. Results show that 60-day images exhibit the highest accuracy, followed by 20-day images, with 5-day images yielding the lowest accuracy. This trend indicates that longer time windows capture richer trend information, enabling horizontal LSTMs to capture longer-term dependencies.

Table 2. Classification accuracy (%) across models

Image Length	LSTM	MOM	STR	WSTR
5 days	52.5	52.1	50.4	51.0
20 days	53.3	52.1	49.8	51.0
60 days	53.6	52.1	49.8	50.4

4.4. Investment strategy performance evaluation

4.4.1. Strategy return rate

Based on model predictions, a long-short strategy was constructed. This strategy achieved an average single-trade return rate of 1.97%, with a total of 523 trades. The annualized return rate was 24.81%, the annualized volatility was 20.79%, and the Sharpe ratio was 1.0970. These metrics significantly outperformed the annualized returns and Sharpe ratios of traditional strategies, demonstrating that the model developed in this paper achieves a markedly superior balance of returns and risk compared to conventional approaches.

DOI: 10.54254/2755-2721/2025.LD28972

Table 3. Long-short strategy performance comparison

	Equal-Weight Annualized Return	Equal-Weight Sharpe Ratio	Total Trades (days)
LSTM(60 days)	36.62%	1.6292	532
WSTR	-1.87%	-0.0516	532
MOM	5.21%	0.2366	532
STR	-2.72%	-0.0758	532

During the testing period, the cumulative returns curve exhibited a stable upward trend, ultimately achieving a return multiple of $(1+0.0197)^{523} \approx 2.8$ times, with no drawdown exceeding 20%.



Figure 4. Cumulative strategy returns

4.4.2. Feature analysis of the model

The feature activation regions of the 2D-LSTM (Fig. 6) reveal that the vertical LSTM's activations are concentrated in the lower-middle part of the image. This indicates the model's sensitivity to the "low-price close" feature—when prices close near the lower end of the recent high-low range, the probability of future upward movement increases. The horizontal LSTM activation clusters on the right side of the image, indicating the model prioritizes recent trends;The volume bar region exhibits significant activation on high-volume days, validating volume's role in confirming price trends [9].

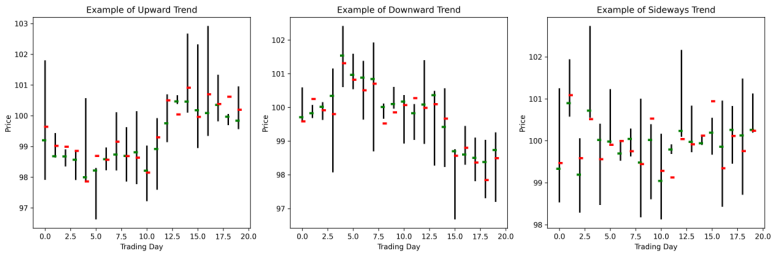
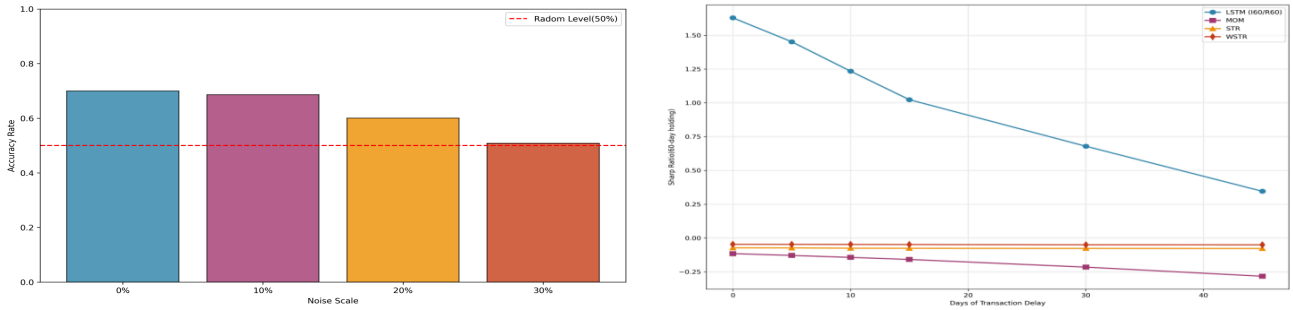


Figure 5. Feature activation regions of 2D-LSTM

DOI: 10.54254/2755-2721/2025.LD28972

4.5. Model robustness testing



(a)Noise robustness test (b)Comparison of trading delay impacts

Figure 6. Noise robustness test

To assess the model's tolerance to image noise, Gaussian noise at varying levels was added to test set images. Results are shown in Figure 7(a): Even at 30% noise, accuracy remained at 50.86%, higher than random performance. These results indicate that the model exhibits strong resistance to mild noise through local smoothing via convolutional layers and sequence modeling via LSTM. However, when noise exceeds 20%, key features of price patterns become distorted, leading to a significant decline in accuracy.

With the intention of accounting for real-world execution delays, experiments tested the effect of 0-45-day delays on the strategy's . Results show that although at 45 days, it remained at 0.3456, still outperforming the other 3 traditional strategies in Sharpe ratio and delay tolerance [1,14].

4.6. Performance comparison across market conditions

To test model adaptability in varying market environments, the study divided the test period into historical phases: China A-share bull market (2014-2019), bear market (2018-2019), and sideways market (2020-2021). The data indicates that this model shows significant differences in performance at each stage, and is consistently superior to traditional strategies.

Table 4. Strategy sharpe ratios across market conditions

Market Phase	LSTM	MOM	WSTR	STR
Bull Market(2014-2019)	1.8521	0.3124	0.1234	-0.0456
Bear Market (2018-2019)	1.2345	-0.1234	-0.2345	-0.3456
Sideways Market (2020-2021)	1.4567	0.1234	-0.0123	-0.1234

5. Conclusion

This paper proposes an LSTM-based stock price image prediction model that captures spatial and sequential features in stock price images via 2D-LSTM [1], outperforming traditional strategies in stock price forecasting tasks. Experiments demonstrate the model's strong transferability across different time scales [4], offering a novel approach to financial forecasting.

Research limitations include: image construction disregarding events like dividends and stock splits; and the model's transferability to other national markets requires further validation. Future work may enhance image features by integrating fundamental data or optimize LSTM feature

DOI: 10.54254/2755-2721/2025.LD28972

extraction efficiency through attention mechanisms [1,8]. Additionally, exploring the model's application to video data represents a promising direction.

References

[1] Tatsunami Y, Taki M. Sequencer: Deep LSTM for Image Classification [J]. arXiv preprint arXiv: 2205.01972v4, 2023.

[2] Yufeng Shi, Yunchuan Sun. Quantitative Investment: Theory and Practice [M]. Shandong University Institute of Finance, 2024: 302-304.

[3] Lo, Andrew W, Harry Mamaysky, and Jiang Wang, 2000, Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation, The journal of finance 55, 1705–1765.

[4] Jiang J, Kelly B, Xiu D. Xiu-Re-Imagining Price Trends [J]. SSRN Electronic Journal, 2021.

[5] Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural computation, 1997, 9(8): 1735-1780.

[6] Schuster M, Paliwal K K. Bidirectional recurrent neural networks [J]. IEEE transactions on Signal Processing, 1997, 45(11): 2673-2681.

[7] Deep dive into bidirectional lstm, 2019. [Online] .Available: https: //www.i2tutorials.com/deep-dive-into-bidirectional -lstm/

[8] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In NeurIPS, volume 30, 2017.

[9] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. In NeurIPS, volume 33, pages 1877–1901, 2020.

[10] Lucas Beyer, Olivier J Hé naff, Alexander Kolesnikov, Xiaohua Zhai, and Aä ron van den Oord. Are we done with imagenet? arXiv preprint arXiv: 2006.07159, 2020.

[11] Wonmin Byeon, Thomas M Breuel, Federico Raue, and Marcus Liwicki. Scene labeling with lstm recurrent neural networks. In CVPR, pages 3547– 3555, 2015.

[12] Qibin Hou, Zihang Jiang, Li Yuan, Ming-Ming Cheng, Shuicheng Yan, and Jiashi Feng. Vision permutator: A permutable mlp-like architecture for visual recognition. IEEE TPAMI, 2022.

[13] Sullivan R, Timmermann A, White H. Data-snooping, technical trading rule performance, and the bootstrap [J]. The Journal of Finance, 1999, 54(5): 1647-1691.

[14] Maas, Andrew L, Awni Y Hannun, and Andrew Y Ng, 2013, Rectifier nonlinearities improve neural network acoustic models, in Proc. icml, volume 30, 3.