

Received December 6, 2021, accepted January 19, 2022, date of publication January 26, 2022, date of current version February 4, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3146371

Identifying Financial Market Trend Reversal Behavior With Structures of Price Activities Based on Deep Learning Methods

CHERN-BIN JU  AND AN-PIN CHEN

Institute of Information Management, National Yang Ming Chiao Tung University, Hsinchu 300093, Taiwan

Corresponding author: Chern-Bin Ju (jc.mg06@nycu.edu.tw)

ABSTRACT In recent years, with the popularity of artificial intelligence (AI) applications, financial market forecasting based on deep learning models has gotten more attention in academia and industry. According to statistics, a long short-term memory network (LSTM) is the first choice for deep learning to deal with the financial time series predicting problems due to its internal memory that can process incoming input with the previous state. However, most of the research used the raw financial time series data composed of opening, closing, highest, lowest price, and transactional time as learning features to feed into models. This study proposed a novel approach focusing on appropriate structures that can organize the price activities as a bell shape curve and identify the greed and fear in the market with deep convolutional neural networks. In addition, this study also considered the influence of time causality on patterns. It is difficult to accurately catch the rapid changes in the market, especially the causal link between various patterns and trend reversal behavior while using only static features extracted by AI models. We designed disparate methods while generating images that could keep the time-variant features on structures and be extracted by convolutional neural networks. Each image was labeled as the upward trend to downward trend, downward trend to upward trend, and trend did not change according to reversals of overall price direction. Finally, to evaluate the availability of the trained model, two-stage experiments were carried out. The first stage of the experiment mainly evaluated the accuracy and profitability of trades following the models, and the second stage considered practical trading rules such as the stop-loss mechanism. The results of the first and second stages showed the proposed models had better profitability and extremely well-matched classification capabilities when compared with the state-of-the-art deep learning models.


INDEX TERMS Algorithmic trading, convolutional neural network, deep learning, financial time series forecasting, market profile, price structure, time causality, trend reversals.

I. INTRODUCTION

Financial time series forecasting is undoubtedly one of the top topics that researchers in academia and industry have desired to solve for a long time. During the past decade, many studies tried to apply machine learning methods to discover the behavior of the financial market. Recently, the emergence of deep learning models has improved research results that significantly outperform previous models.

Deep learning is one of the popular research fields in artificial intelligence (AI), and it has been popularly applied

to many applications in our life. For instance, recurrent neural networks (RNN) and long short-term memory (LSTM) networks are used for speech recognition because they can consider not only current input but previous states at the same time. This feature makes them perform well for sequence and time series data. According to statistics [1], LSTM networks are the first choice to deal with the financial market forecasting problems by learning the raw time series data composed of opening, closing, the highest, and lowest price during a time as features. However, there are many different participants with diverse timeframes in the market, such as long-term investors, short-term traders, day traders, and scalpers. The market is influenced by their wide variety of timeframes

The associate editor coordinating the review of this manuscript and approving it for publication was Chien-Ming Chen .

and motivations such as speculating, hedging trading, and arbitrage, resulting in lots of noise in market price. These excessive noises may be too sensitive to learn the correct features to discover the trend behavior from the raw time series data. To cite an instance, raising up the price can be caused by a new buying business, or it may be carried out by an old buying business such as short-covering order.

Another common deep learning method is convolutional neural networks (CNN). It is widely used in computer vision because of its powerful ability to extract important features from images. With the powerful image classification ability of CNN, Sezer and Ozbayoglu [2] proposed CNN-BI using a 2-D CNN directly input closing price chart images without introducing any other time series data. This is the first attempt in the literature to adopt CNN to train a trading model. However, the financial market is a dynamic environment, and it is affected by not only a snapshot of patterns also its development characteristic with time. For example, a downward trend auction occurred last week, but a responsive buying just showed up three days ago. It would be classified the whole activities into a downward trend and omit the responsive buying, but it should start to reverse to an upward trend in the near time. Therefore, exploring the influence of time causality on features has become a very important topic.

A novel approach for analyzing the behavior of dynamic financial markets was proposed in this paper, which focused on appropriate structures of price activities that can identify the trend reversals with deep CNNs. In addition, this study also considered the time-variant extracted features. A series of extracted patterns with different occurred time could imply the behavior of financial market trend reversal. Equation (1) illustrated the idea of our perspective: y_t is the trend reversal result predicted by the models, and it is the function of extracted patterns and occurred time t . If the observed time period is from t to $t-n$ and there are m extracted patterns, the (1) can be expressed as the interaction of a series of patterns that occurred at time t_i during the observed period from t to $t-n$.

$$\begin{aligned} y_t &= f(\text{patterns}, t) \\ &= f(\{pattern_0, pattern_1, \dots, pattern_m\}, \{t, \dots, t-n\}) \\ &= f(\{pattern_0^t, pattern_0^{t-1}, \dots, pattern_m^{t-n}\}) \end{aligned} \quad (1)$$

Fig. 1 demonstrated the process of the proposed approach. First, to generate the images be able to utilize the power of the deep CNN, we adopted market profile theory [3], [4] to convert financial time series data into proper structures represented with text market profiles. In order to come up with a representation containing the characters of the time, we used different linear interpolation to convert the Time Price Opportunities (TPO) in profiles into greyscale values so that the generated images can maintain the time-variant features in structures. Next, images were grouped with the sliding window method and labeled with “downward trend to upward trend,” “upward trend to downward trend,” and “trend did not change” based on the current trend in the

grouped image and whether the trend will reverse. Finally, training the model with CNNs and its predicted results were evaluated using out-of-sample data.

The experimental results showed that the proposed method significantly outperformed the model trained by state-of-the-art benchmark models with raw financial time series data in terms of profitability after two stages of different trading mechanism analyses. Furthermore, the two models have extremely well-matched classification capabilities for identifying the trend reversal behavior. Therefore, the proposed approach is an exemplary method and could effectively discover the reversal behavior in the financial market, providing investors a reference for better rewards.

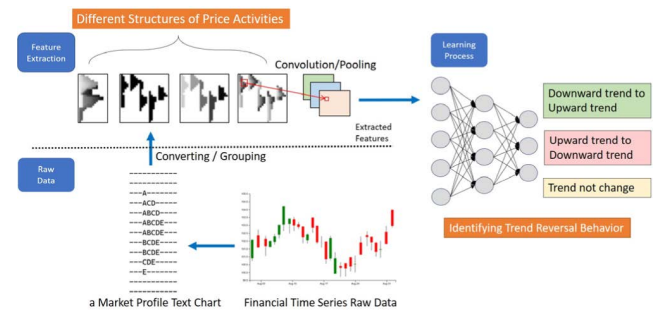


FIGURE 1. The process of the proposed approach.

The contributions of this study are briefly summarized as follows. (I) Constructing a deep CNN model to forecast the trend reversals of the financial market by converting appropriate structures of price activities from market profile theory as learning features. (II) In order to emphasize the influence of time causality on patterns, disparate methods were designed to convert structure shapes to images with different grayscale values so that the time-variant features could be extracted from structures instead of obtaining static features by CNNs. (III) This was probably the first attempt using market profile theory with deep learning methods in financial market behavior forecasting to the best of our knowledge. (IV) To evaluate the better-converted structure style, two types of formats to compare performance were arranged. (V) An available model for identifying the reversals based on overall price direction was proposed.

II. RELATED TECHNIQUES AND WORKS

A. MACHINE LEARNING AND DEEP LEARNING

AI has recently become one of the most important research and application topics. Machine learning is a branch of AI. It is an intelligent way to learn from data, train models that can make decisions automatically or as a decision support system for humans. Nowadays, machine learning has been widely used in many fields, such as image recognition, video analysis, object detection, financial time series prediction, weather forecasting, marketing, and medical image research [5], [6]. Moreover, the emergence of deep learning has led to another research with results that significantly outperform

previous models. The CNN is a type of deep learning network widely used in our lives, especially in the computer vision field, because of its powerful recognition ability. The simple architecture of the CNNs is LeNet, proposed by Yann [7] for digit recognition. Beyond image classification and recognition problems, CNN can also help us with financial time series prediction.

CNN consists of convolutional layers, pooling layers, and multiple fully-connected layers neural networks with many weights and biases that need to be trained with data. Unlike a traditional neural network, CNN shares the weights among neurons to reduce the number of parameters and avoid falling into an overfitting problem [8].

The convolution process convolves the original image with a specific filter which randomly generates several patterns. Its role is to obtain the boundary between some features and objects in the picture so that the input picture can be judged according to these extracted characters in the future. The pooling process is not only to reduce the number of features and parameters but also to reduce the computational complexity. A common pooling method is max pooling, as it picks the maximum value in the matrix. It retains all important feature information and improves the efficiency of the CNN. The fully-connected layers connect the features after the previously mentioned processes and the classifying labels. The whole process of CNN is illustrated in Fig. 2.

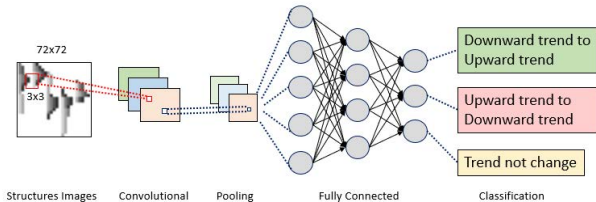


FIGURE 2. Illustration of the CNN architecture.

In literature, CNN is not preferred for time series data analytics. According to statistics, LSTM networks are the first choice to deal with the forecasting problems with OCHL datasets due to their memory cell connection [9], [10]. However, this study transformed financial time series data into market structure shapes with price activities and market profile theory. Through the different transforming processes, thousands of images were generated, and they could be regarded as feature graphs that were 2-D vectors full of the trend reversal characteristics at a specific time. Therefore, the financial time series forecasting problem is implicitly converted into an image classification problem. CNN had the potential to extract the characteristics from the structure shape images.

Krizhevsky *et al.* [11] proposed AlexNet, a trained CNN model, achieved a top-5 error rate of 15.3% that successfully classified 1.2 million images into 1000 different classes in the ImageNet datasets. The accuracy is beyond the past research results. Karpathy *et al.* [12] provided an empirical evaluation of CNNs on large-scale video classification.

They used one-million YouTube videos belonging to 487 classes. The results showed significant performance improvements compared to the baselines. The results of both [11] and [12] indicate that the CNN is suitable for image recognition and classification.

Sezer *et al.* [14] proposed CNN-TA using a 2-D CNN based on image processing properties. By converting 15 different technical indicators financial time series into 2-D images. They used Dow Jones 30 stock prices and ETFs as data. The results indicated that the built model CNN-TA performed well against the Buy & Hold, SMA, MLP, and even LSTM models. The CNN-TA gained average 6% more than the LSTM model in annualized return. According to the literature, CNN has good image recognition and classification capabilities. CNN-TA is a novel application that uses technical indicators to generate images to predict stock prices. However, the selection of technical indicators is handcrafted, and it is less meaningful to combine the calculated values of technical indicators into a picture. In this study, we directly used the transactional price of the raw data to perform structure shape conversion. In addition, through market profile theory, the transformed structure was meaningful due to the interaction between time and price. The next section will discuss the market profile structure shape.

B. MARKET PROFILE THEORY

Market profile is a text chart proposed by Steidlmayer [3] in 1984 and published on CBOT in 1985. Market profile mainly uses the information generated by the market and presents it in a bell curve. The bell curve concept is derived from the normal distribution in statistics. Steidlmayer believes that most of the trading results after a period of trading time should be concentrated in the center of the transactional price range. The center area is called the value area, composed of 70% trades, and the remaining extreme area is called excess price. The excess price above value area is recognized as too expensive to buy, while the excess price below value area is considered too cheap so that no one is willing to sell. Both up and bottom excess include 15% trades under normal circumstances. Fig. 3 illustrates the appearance of market profile.

By keeping tracking the transactional market data comprised of time and price, market profile could reveal the direction of the market trend once the evidence of controlled timeframe is uncovered. The basic building blocks of the market profile are called TPO. Every defined time period is designated by a letter. For example, the first time period is denoted as "A," and the second is denoted as "B." The defined time period can be a day of a week, an hour of a day, or any time segment. A market profile usually consists of at least five time periods. If a certain price is traded during a given time period, the corresponding letter is marked next to the price. Fig. 4 shows the complete profile composed of five-time periods (from A to E). Suppose the current market is dominated by long-term buyers, a higher probability of market prices rising. On the contrary, if the market is dominated

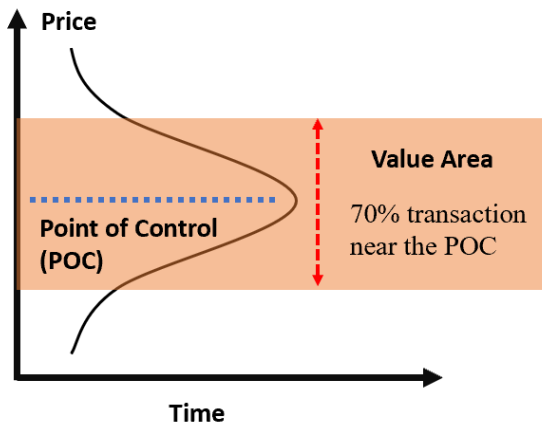


FIGURE 3. The bell-shaped curve of market profile.

by long-term sellers, there is a higher chance of falling, occurring a downward trend in the market. The remaining behavior that market trend is non-trend if it is analyzed that there are no long-term buyers and sellers in the current market, the future price may fluctuate within a narrow price interval.

Steidlmayer extracts some special patterns from the market profile to help traders understand which way the market is trying to go. There are three important patterns: (a) Pointer of control, POC (b) Value area, VA (c) Tail. POC indicates the most concentrated price range of transactions after a period of trading time, as illustrated in Fig. 4. The transactional price has been traded among the most time periods, and it means that every participant in the market agrees with the price, so that market price is balanced around the price. VA expands the concept of POC, making the balanced price be a range of price. Just like counting the height of students of the same race on campus, it would show a distribution similar to the normal distribution. The price is accepted by the market participants should be within a certain range above and below the POC. The market profile uses the concept of one standard deviation in statistics to define the range. It refers to approximately 70 percent of the transaction range near the point of control as VA. Most transactions will be dealt with in the VA if no major political and economic events or news may change the participants' perception of market value.

The third pattern is tail. The emergence of the tail is because short-term traders with greed and fear overreact to some event or news, resulting in an imbalance of supply and demand between short-term buyers and sellers. As the price moves away from the value area, long-term traders such as larger institutions and commercial participants in the marketplace would take advantage of the price. A tail occurs while the long-term participants act quickly, making the price rotate back into VA. In market profile, a tail appears when there are at least two single prints on the bottom of the profile or just the opposite, at the top of the profile, as shown in Fig. 4. Therefore, a tail is considered an important key: buying tails indicate the end of a downward auction, whereas selling tails

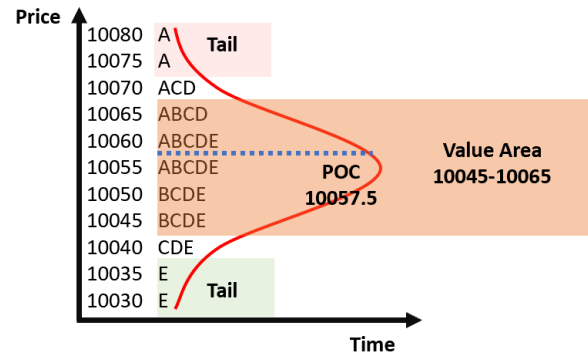


FIGURE 4. The appearance of market profile.

indicate the end of an upward auction. Fig. 5 demonstrates an example of the trend reversal from a downward trend to an upward trend. In addition, a buying tail is also called a support area, and a selling tail is also called a pressure area. Short-term traders would use support and pressure to trade, for example, buy when the price is close to the support area, and sell when the price is close to the pressure area. The greater the range between pressure and support, the more active market transactions because diverse prices can be traded to meet the needs of market participants.

The three patterns mentioned above form more meaningful structures of price activities than the raw time series data or calculated values of technical indicators. Undoubtedly, more than three patterns in the market profile structure. With the help of the market profile, market participants can understand the financial market trend reversal behavior.

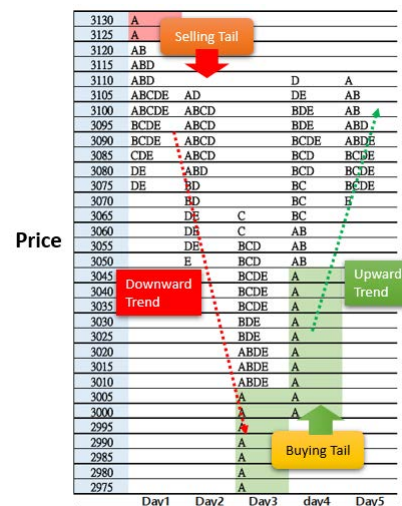


FIGURE 5. Trend reversals example with buying tail and selling tail.

Lin *et al.* [14] used the market profile theory and neural network to build the model that conducted empirical experiments on intraday trading. They adopted values of POC, VA, price range, and tail as the input and fed into fully connected neural networks for forecasting a short-term market trend.

Their experiments showed that the levels of accuracy and profitability among different short-term trading periods were significantly better than the random trading strategy model. The results confirmed the effectiveness of market profile theory. Chen *et al.* [15] proposed qualitative and quantitative methods to compute a market profile indicator by implementing the market profile theory on neural network architecture. The experimental results showed the quantitative market profile indicators had better trend-predicting ability in the long-term forecasting period. Moreover, they also suggested that the combination of long-term and short-term changes in the market can enhance forecasting performance and profitability.

The literature mentioned above showed that the market profile theory is an effective method to help discover a financial market trend. In this paper, we use the market profile to convert time series data into meaningful structure images. In addition to the POC, VA, Tail, and other patterns proposed by Steidlmayer, the powerful feature extraction capabilities of the CNNs are used to discover various features that can identify the key trend reversal behavior.

C. RECURRENT NEURAL NETWORK

The RNN is another type of deep learning network that has been widely used in nature language processing and speech recognition. Such as language translation application and sequential data processing [16], [17] because of its internal memory that can process incoming input with the previous state. The RNN model architecture is composed of different non-independent hidden layers so that it can take the current and previous input data into account at the same time. Therefore, RNN would perform well when dealing with sequential data. Fig. 6 illustrates the two basic structures of RNN. An Elman network is a three-layer network with an additional cell that connects to a hidden layer. The cell will pass the previous value of the hidden layer. Jordan networks are similar to Elman networks but the additional cell connecting to the output layer. Two structures of networks can maintain a short memory and perform sequence-prediction tasks [18], [19]. Equation (2) and (3) shows the forward pass of the Jordan network illustrated in Fig. 6. $\mathbf{a}^{(t)}$: the hidden layer vector; σ_a , σ_y : activation function; \mathbf{W}_i , \mathbf{W}_h , \mathbf{W}_o : weight matrices to be learned; \mathbf{b}_a , \mathbf{b}_y : bias vector to be learned; $\mathbf{y}^{(t)}$: output vector of the network.

$$\mathbf{a}^{(t)} = \sigma_a \left(\mathbf{W}_i \mathbf{x}^{(t)} + \mathbf{W}_h \mathbf{y}^{(t-1)} + \mathbf{b}_a \right) \quad (2)$$

$$\mathbf{y}^{(t)} = \sigma_y \left(\mathbf{W}_o \mathbf{a}^{(t)} + \mathbf{b}_y \right) \quad (3)$$

LSTM are one of the variants of the RNN. LSTM can avoid vanishing gradient problems when the RNN model updates the weights by controlling what is added and removed from memory in the hidden layers [20]. This is conducted by using a combination of three gates: (a) a forget gate, (b) an input gate, and (c) an output gate expressed in Fig. 7. Equation (4)-(9) shows the forward pass of the LSTM unit with the three gates. $\mathbf{X}^{(t)}$: input data vector; $\mathbf{f}^{(t)}$: forget gate's

activation vector; $\mathbf{i}^{(t)}$: input gate's activation vector; $\mathbf{o}^{(t)}$: output gate's activation vector; $\mathbf{g}^{(t)}$: new memory be added to memory cell state activation vector; $\mathbf{c}^{(t)}$: memory cell state vector; $\mathbf{h}^{(t)}$: output vector of LSTM unit; σ : sigmoid activation function; \mathbf{W} , \mathbf{U} : weight matrices to be learned; \tanh : hyperbolic tangent activation function; \mathbf{b} : bias vector to be learned; $*$: Hadamard product.

$$\mathbf{f}^{(t)} = \sigma \left(\mathbf{x}^{(t)} \mathbf{W}_f + \mathbf{h}^{(t-1)} \mathbf{U}_f + \mathbf{b}_f \right) \quad (4)$$

$$\mathbf{i}^{(t)} = \sigma \left(\mathbf{x}^{(t)} \mathbf{W}_i + \mathbf{h}^{(t-1)} \mathbf{U}_i + \mathbf{b}_i \right) \quad (5)$$

$$\mathbf{o}^{(t)} = \sigma \left(\mathbf{x}^{(t)} \mathbf{W}_o + \mathbf{h}^{(t-1)} \mathbf{U}_o + \mathbf{b}_o \right) \quad (6)$$

$$\mathbf{g}^{(t)} = \tanh \left(\mathbf{x}^{(t)} \mathbf{W}_g + \mathbf{h}^{(t-1)} \mathbf{U}_g + \mathbf{b}_g \right) \quad (7)$$

$$\mathbf{c}^{(t)} = \mathbf{f}^{(t)} * \mathbf{C}^{(t-1)} + \mathbf{g}^{(t)} * \mathbf{i}^{(t)} \quad (8)$$

$$\mathbf{h}^{(t)} = \mathbf{o}^{(t)} * \mathbf{c}^{(t)} \quad (9)$$

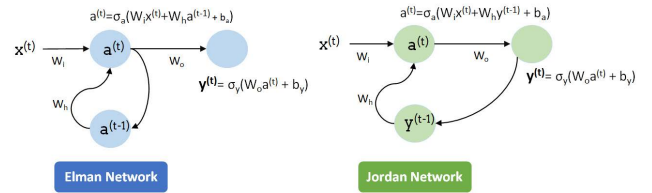


FIGURE 6. Two structures of recurrent neural networks.

In literature, most deep learning research used LSTM to solve the financial time series forecasting problems with the OCHLV datasets consisting of the opening price, highest price, lowest price, closing price, and trade volume of a specific timeframe. Chen *et al.* [21] used LSTM to predict China stock return by transforming 30-days transactional data with ten learning features composed by OCHLV of SSE index and stocks into a sequence. The results showed that LSTM is better than the random prediction model. Hiransha *et al.* [22] compared MLP, LSTM, RNN, and CNN with the stock price. They fed OCHLV, turnover, and number of trades to those models. After trained, the models were used for predicting the stock price of five stocks from NSE and NYSE. They concluded that deep learning models outperform the ARIMA model, the traditional method used for time series forecasting. Khare *et al.* [23] employed LSTM for intraday price prediction and proved the postulates of Inefficient Market Hypotheses. Zhang and Tan [24] implemented DeepStockRanker, an LSTM-based model for predicting the future return ranking of stocks by employing 11 values of technical indicators. By selecting the highly-ranked stocks predicted by the model, the results showed LSTM model approach significantly outperforms state-of-the-art techniques such as SVR and RBM. Based on a survey of previous works, LSTM is one of the common models used for solving financial time series forecasting problems. In addition, Selvin *et al.* [25] used RNN, LSTM, and CNN for predicting the price of three stocks of the National

Stock Exchange of India. Their result showed that CNN gave more accurate results than the other two models. Recently, Livieris *et al.* [26] proposed a model named CNN-LSTM. They conducted experiments for the accurate prediction of gold price and movement. The result showed CNN-LSTM model was against the LSTM model, too. Therefore, in this paper, the LSTM, CNN, and CNN-LSTM methods would be employed as our control groups models and compared with our proposed method.

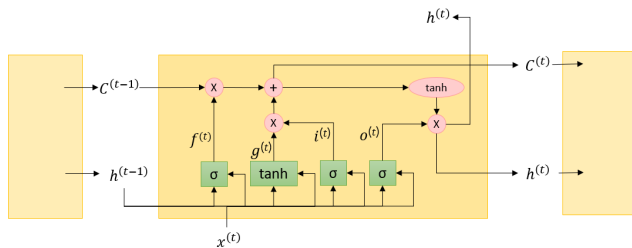


FIGURE 7. The structures of the LSTM unit.

III. PROPOSED METHOD

A. RESEARCH DESIGN

For the research model in this paper, a novel model that utilized structures of price activities to discover the trend reversal behavior was proposed. In addition, considering the influence of time causality on features was the key to solving the financial time series forecasting problems. This paper arranged a series of experiments to probe the importance of the structures and time. The first stage of experiments tried to validate the significance of proposed structures. In this stage, a preliminary experiment was conducted to see which way of presenting recording methods was proper for generating structures. There were two experimental groups to evaluate the effectiveness, including the single structure and multiple structures. The single structure organized all price activities in a large market profile for analyzing the trend after twenty-five days. By contrast, multiple structures recorded price activities in a structure every five days, so there was a total of five shapes of the structure. After the preliminary experiment, different occurred time was also considered by presenting the time as a feature within a shape or between shapes. Different pixel grayscale values were used to express the time causality when generating the structures' images. According to the transaction date of OCHL or the sequence number of structures in the images, disparate linear interpolation was performed for calculating the pixel grayscale values ranging from 0 to 255. Fig. 8 and Fig. 9 demonstrated the corresponding grayscale values after calculating based on twenty-five-day OCHL data in experiments.

The experimental groups were designed in accordance with the experiments mentioned above and divided into four groups:

Experimental Group A (EGA): twenty-five-day OCHL data was converted into a single shape of the grayscale structure in an image. The sample image of EGA was illustrated

in Fig. 10 (a), and its grayscale values in the structure varied with the transaction date of OCHL, as demonstrated in Fig. 8.

Experimental Group B (EGB): twenty-five-day OCHL data was converted into five shapes of the grayscale structure in an image. Each structure recorded five days of price activities. The sample image of EGB was illustrated in Fig. 10 (b). The grayscale value of these shapes was all black as an experiment regardless of the influence of time causality in the experimental groups, as shown in Fig. 9.

Experimental Group C (EGC): twenty-five-day OCHL data was converted into five shapes of the grayscale structure in an image. Each structure recorded five days of price activities. The sample image of EGC was illustrated in Fig. 10 (c), and its grayscale values of shapes varied with the sequence of structures to emphasize the influence of time causality on features between five shapes, as shown in Fig. 9.

Experimental Group D (EGD): twenty-five-day OCHL data was converted into five shapes of the grayscale structure in an image. Each structure recorded five days of price activities. The sample image of EGD was shown in Fig. 10 (d). The grayscale values of shapes varied with the transaction date of OCHL within each structure to stress the influence of time causality in every shape, as demonstrated in Fig. 9.

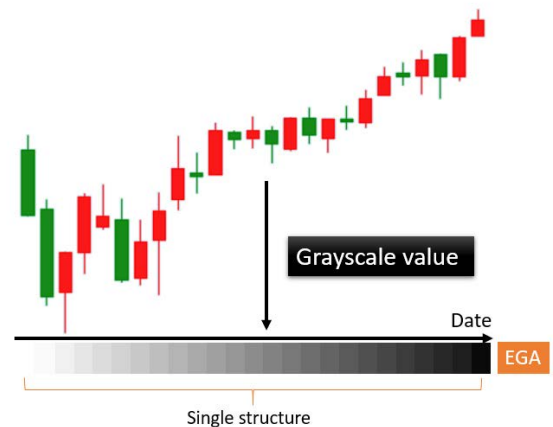


FIGURE 8. The corresponding grayscale values after calculating based on twenty-five-day OCHL data in EGA.

B. DATA COLLECTION

The research goal of this paper was to discover the trend behavior by market structures of price activities. In order to validate the proposed approach, we collected the futures data of the S&P 500 since its huge amount of transaction volume and great facilitation. The data source was the daily transaction data provided by barchart.com. The sample of data was shown in Fig. 11, including the date, the opening price of the day, the highest price of the day, the lowest price of the day, and the closing price of the day. The five fields were called raw data and composed for a candlestick bar which is the least element in financial time series analysis.

The experimental period of samples was collected from January 3, 2000, to December 31, 2020. A total of 5291 data

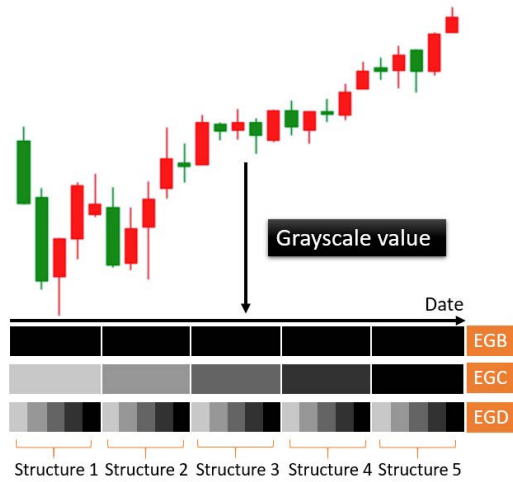


FIGURE 9. The corresponding grayscale values after calculating based on twenty-five-day OCHL data in EGB, EGC, and EGD.

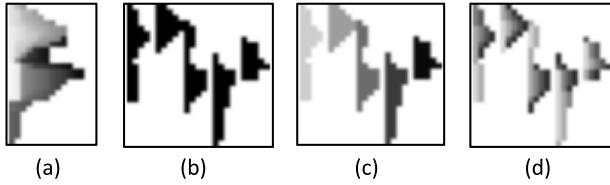


FIGURE 10. Illustration of input image generating results. (a) EGA: single structure with the different grayscale values calculated by the date, (b) EGB: multiple structures with the same grayscale value, (c) EGC: multiple structures with the different grayscale values between shapes calculated by the sequence of structures, (d) EGD: multiple structures with different grayscale values in shapes calculated by the date within a structure.

was obtained. The data was divided into training data and testing data to verify the effectiveness of the training network. The experiment used the ideal ratio of 80-20 for split, according to Kearns [27]. The training period was from January 3, 2000, to December 31, 2016. In comparison, the testing period was from the beginning of 2017 to the end transactional day of 2020.

Date	Opening	Closing	Highest	Lowest	Volume
2020/1/2	3237	3259	3261.75	3234.25	1360588
2020/1/3	3261	3235.375	3263.5	3206.75	1682453
2020/1/6	3220.25	3243.625	3249.5	3208.75	1460300
2020/1/7	3243.5	3235.25	3254.5	3226	1228094
2020/1/8	3231.75	3260.25	3267.75	3181	2223039
2020/1/9	3261.25	3275.875	3276.75	3257.75	1224853

FIGURE 11. The sample of collected raw data (OCHLV).

C. CALCULATE THE STRUCTURES OF PRICE ACTIVITIES

In order to generate the structures of price activities, market profile theory was adopted to convert the raw time series data to text profile according to the time, opening, closing, highest and lowest price. Algorithm 1 described the process

of generating the structure from raw data. In this paper, the time interval was set to one day, and the price interval was set to five points. The basic element in the market profile is called TPO, which is a set of letters from A to Z. Every alphabet represents a defined time. After inputting our collected S&P daily transactional data, the structure can be generated based on a defined time interval and a price interval.

Algorithm 1 GenerateStructureFromRawData()

Input Financial time series data (OCHL)

Output Text Market Profile Structure

```

1: define timeInterval = 1 day
2: define priceInterval = 5 point
3: while(dataRowCount < numberOfData) do
4:   MP = createMPDictionary<price,ListTPO>()
5:   TPO = MapDataTimeToTPO(data.Time)
6:   for p=data.Highest; p>=data.Lowest;
       p=p-priceInterval
7:     if (!MP.containsKey(p))
8:       MP.addKey(p)
9:     if (!MP[p] contains TPO)
10:      MP[p].push(TPO)
11:    dataRowCount++
12: return output

```

D. LABELING METHOD

In the labeling process, each input image is labeled to represent the trend reversal behavior. Let c_t be the closing price of a day, and c_{t-10} and c_{t-5} be the closing price ten days and five days ago. To identify the trend reverses from current time t to 5 days later time $t+5$ for models. Three labels y_t are defined in (10). The value of 0 means the trend reverses from an upward trend to a downward trend, and the value of 2 means the trend reverses from a downward trend to an upward trend. The value of 1 means that the trend does not change.

$$y_t = \begin{cases} 0, & \text{if } c_{t-5} > c_{t-10} \text{ and } c_t > c_{t-5} \text{ and } c_{t+5} < c_t \\ 2, & \text{if } c_{t-5} < c_{t-10} \text{ and } c_t < c_{t-5} \text{ and } c_{t+5} > c_t \\ 1, & \text{otherwise} \end{cases} \quad (10)$$

After learning the trend reversal behavior, corresponding actions including 'Buy action,' 'Sell action,' and 'No action' are taken by the determination of trend reversal behavior for evaluating the financial performance. A buy action was taken when the trend would transit from downward trend to upward trend. A sell action was taken while the trend would transit from upward trend to downward trend. In the remaining situation, no action is taken since there is no clear evidence to change the trend.

E. SLIDING WINDOW

We adopted a sliding window method while converting futures daily data of S&P 500 to grayscale images of

the structure. Therefore, every twenty-five-day OCHL daily data was included in an image. For example, the input image i contained OCHL data from t_{i-24} to t_i , and the next image started from t_{i-23} to t_{i+1} . Fig. 12 illustrated this sliding windows approach.

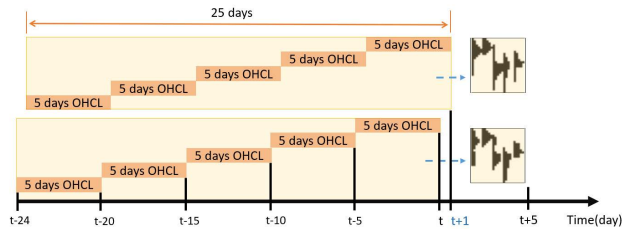


FIGURE 12. The sliding window method of converting data.

F. CNN NETWORKS

In the learning phase, each of the experimental groups was conducted by the CNN, which is successfully applied for image classification. The networks structure in this study consisted of eight layers, including input layers, three convolutional layers, two max pooling layers, fully connected layers, and an output layer. The combination of eight layers was similar to the LeNet CNN structures, the successful network in handwritten digits classification. We had tried the deeper networks such as ResNet or DenseNet, but the performance was not as good as shallow networks due to the overfitting problem in the training datasets. It may cause by the number of our input grayscale image data which was converted from the financial time series. Although twenty years of daily transactional data was collected and transformed with the sliding window method, the total amount of input data was roughly 5,300 images. The complete network structures were illustrated in Table 1.

TABLE 1. CNNs structure.

Layer	Shape	Parameters
Convolution	72, 72, 64	1664
Max pooling	35, 35, 64	0
Convolution	35, 35, 128	204928
Max pooling	11, 11, 128	0
Convolution	11, 11, 256	295168
Flatten	30976	0
Dense	256	7930112
Dense	3	771

G. CONTROL GROUPS

According to the related works mentioned in Part II, three state-of-the-art benchmark models were employed as the control groups. Control Group A (CGA) used the LSTM networks as the training model; Control Group B (CGB) employed CNN; Control Group C (CGC) applied CNN-LSTM networks. In order to compare with the experimental

groups, the input data was the financial time series data with the same range as experimental groups. Every record consisted of twenty-five-day transactional data. In addition, according to the recommendation of Chen and Hsu [28], applying the first-order value can increase the learning effect of neural networks. The input data fed into the LSTM and CNN-LSTM networks were processed by first-order differential and transformed the value between 0 and 1. The CNN model was fed with raw time series data as an ablation experiment and compared with the proposed model.

H. RESEARCH RESTRICTION AND TRADING STRATEGY

There were some concerns and restrictions in this study illustrated as follows: First of all, the transaction cost was calculated in this research. The futures contract of ES transaction cost was set to 2 ticks. Secondly, the slippage of every transaction did not consider, and transferring positions to the next delivery contract was also skipped. Besides, the initial capital was set to \$11,000 US dollars which was the required initial margin by exchanger CME. Last but not least, to complete the whole transaction based on all model predictions, a margin call occurring when the value of our margin account drops and fails to meet the account's maintenance margin requirement was not considered.

IV. RESULTS AND ANALYSIS

This study consisted of two stages of experiments, which contained four experimental groups and three control groups. All of them were conducted on S&P 500 mini futures. In the learning phase, twenty-five-day transactional data from the NYSE exchanger was fed into the input layer of each network in every experimental group and control group.

To evaluate the proposed methods in this paper, the following evaluation measures were adopted throughout this section: accuracy, precision, total profit, average profit per transaction (APpT), annualized return (AR), and profit to loss ratio (PL). All formulas for measures were illustrated in (11)-(17). The first two measures represent the classifier performance of the proposed model. Accuracy expresses whether the trained classifier distinguishes among the trend reversal classes, 'downward trend to upward trend,' 'upward trend to downward trend,' and 'trend not change.' In addition, precision shows the profitable ability to do buy and sell action after model prediction. The other assessment measures evaluated the financial performance of trained models by implementing a simulation environment to examine the empirical trading results in the financial market based on the model predictions. Algorithm 2 summarized all processes of this study.

$$\text{Accuracy} = \frac{\text{Number of Labels classfied Correctly}}{\text{Number of Labels Need to Classify}} \quad (11)$$

$$\text{Precision} = \frac{\text{Number of Profitable Trade}}{\text{Number of Entering Trade}} \quad (12)$$

$$\begin{aligned} \text{Total Profit (\%)} \\ &= \frac{\text{Final Value of Capital} - \text{Initial Value of Capital}}{\text{Initial Value of Capital}} \\ &\quad \times 100 \end{aligned} \quad (13)$$

$$\begin{aligned} \text{Total Profit (USD)} \\ &= \text{Final Value of Capital} - \text{Initial Value of Capital} \end{aligned} \quad (14)$$

$$\begin{aligned} \text{APpT} \\ &= \frac{\text{Total Profit(USD)}}{\text{Number of Entering Trade}} \end{aligned} \quad (15)$$

$$\begin{aligned} \text{AR} \\ &= \left(1 + \frac{\text{Total Profit(\%)}}{100}\right)^{\frac{1}{\text{number of years}}} - 1 \end{aligned} \quad (16)$$

$$\begin{aligned} \text{PL} \\ &= \frac{\text{Average Profit of Profitable Trade}}{\text{Average Loss of Loss Trade}} \end{aligned} \quad (17)$$

In the first stage, a preliminary experiment was carried out. The first two experimental groups (EGA, EGB) investigated the effectiveness of two types of generated structures based on market profile theory. All transactional data was merged into a single structure in EGA. In contrast with the merged structure, EGB split data into five structures. The results in Table 2 recommend that separating data into numerous structures be the better formation. As illustrated in Table 2, EGB outperformed EGA, which is considered a single structure. The average annualized return for EGB was over 42%, whereas EGA average annualized return was 1%. Although the two experimental groups performed equally in accuracy, there was a significant difference between the two models in financial performance comparison. Therefore, models in EGC and EGD were based on multiple structures.

After the preliminary experiment, proposed models were compared with the state-of-the-art models trained with time series data mentioned in the literature. The first stage was designed for investigating model prediction: while the model suggested ‘downward trend to upward trend,’ acquire a long ES futures position and sell it five days later. Otherwise, short a position when model forecasted ‘upward trend to downward trend’ and bought covering after the forecast number of days (five days). Moreover, two rules were considered in the second stage to make the experimental result more practical. In the second stage, it considered general rules while trading: taking maximum risk and trend continuation into account. Neither long nor short positions were closed when the model predicted the same trend reversal behavior as the current position. Furthermore, long and short positions should set the maximum stop-loss to prevent the models from constantly predicting the reversing to upward trend when the market continues to fall. The results and analysis of the two stages were illustrated in the following subsection A and B.

A. STAGE 1 - INVESTIGATING MODEL PERFORMANCE

The first stage evaluated the classifier performance. The results based on predictions of four models

(EGB, EGC, EGD, CGA, CGB, CGC) showed in Table 3. All experimental and control groups’ accuracy rates were close to 60%. Although the CGA trained by the LSTM network performed best with an accuracy rate of 64.78%, the accuracy of the proposed experimental group, EGB, and EGD, is about 64%. The accuracy of the two groups is almost the same. Besides, the precision of models among experimental and control groups was approximately the same except for CGC with 52%. In the financial performance evaluation, the average annualized return of the experimental groups was higher than 40%. It was much higher than all models in control groups. The profit to loss ratio (P/L) of all the experimental groups was higher than the control groups in Table 3, indicating that proposed models could obtain more rewards while profiting but smaller losses when they failed. The results showed that the classifier trained with proposed structures did identify the moment that trend reversed from downward trend to upward trend and upward trend to downward trend.

Algorithm 2 ALLProcess()

Input Financial time series data (OCHL)

```

1: Phase Get TextMP
2: dataset = GenerateStructureFromRawData()
3: if(createMultipleMP)
4:   groupDataset = multipleMP(5)
5: else
6:   groupDataset = singleMP(5)
7: Phase Data Labeling
8:   calculatingLabel(groupDataset)
9: Phase Image Creation
10:  images = converToGrayImages(groupDataset)
11:  convertImages = covertGrayImagesByTimeCausality
    (images,experimentName)
12: Phase Train Model
13:  trainData = splitData(startTime, 2016)
14:  testData = splitData(2017, endTime)
15:  model = createCNN()
16:  model.train(trainData)
17: Phase Model Performance Evaluation
18:  model.test(testData)
19: Phase Financial Performance Evaluation
20:  realEnv = createRealEnv(risk)
21:  predictions = model.predict(testData)
22:  realEnv.evaluatePerformance(predictions,stageName)

```

B. STAGE 2 – PRACTICAL PERFORMANCE

As mentioned earlier, the second stage conducted more practical experiments with two general trading rules. First, a position should be kept rather than closed when the signal of predicting the trend reversal did not change. Second, to avoid taking huge risks, a stop-loss order has to be set while acquiring an initial long and short position. The stop-loss is a commonly-used mechanism to handle it in

TABLE 2. Result of the preliminary experiment of the first stage.

Experimental Group	Model Performance		Financial Performance					
	Accuracy (%)	Precision (%)	Total Profit (%)	Total Profit (USD)	Number of Transactions	Average Profit per Transaction (USD)	annualized return (%)	P/L
EGA	61.83	39.92	4.61	507	75	6	1	1.56
EGB	63.89	42.96	304.69	33516	83	401	42	1.93

TABLE 3. Results of the experiment of the first stage.

Experimental Group	Model Performance		Financial Performance					
	Accuracy (%)	Precision (%)	Total Profit (%)	Total Profit (USD)	Number of Transactions	Average Profit per Transaction (USD)	annualized return (%)	P/L
EGB	63.89	42.96	304.69	33516	83	401.8	42	1.93
EGC	58.94	42.44	293.99	32339	92	351.5	41	1.89
EGD	63.38	45.5	305.12	33563	96	348.6	42	1.72
CGA	64.99	46.8	7.03	772.92	53	14.58	2	1.17
CGB	54.69	45.24	-176.27	-19390	80	-242.3	NaN	1.05
CGC	60.22	52.6	100.91	11100	37	300	19	1.13

TABLE 4. Results of experiment of the second stage (2% risk).

Experimental Group	Model Performance		Financial Performance				
	Precision (%)	Total Profit (%)	Total Profit (USD)	Number of Transactions	Average Profit per Transaction (USD)	annualized return (%)	P/L
EGB	44.6	320.5	35252	60	589.7	43.2	1.74
EGC	47.0	413.7	45504	61	740.9	50.5	1.70
EGD	46.0	232.9	25616	70	367.3	35.1	1.49
CGA	49.3	-66.7	-7340	39	-188.2	-24.1	0.94
CGB	48.4	-262.8	-28907	45	-649.6	NaN	0.79
CGC	54.2	57.2	6295	32	196.7	12	0.99

TABLE 5. Results of experiment of the second stage (5% risk).

Experimental Group	Model Performance		Financial Performance				
	Precision (%)	Total Profit (%)	Total Profit (USD)	Number of Transactions	Average Profit per Transaction (USD)	annualized return (%)	P/L
EGB	47.7	190.0	20900	51	409.0	30.5	1.38
EGC	51.9	334.4	36785	51	721.7	44.4	1.35
EGD	52.3	142.5	15677	59	266.6	24.8	1.08
CGA	51.9	143.8	15818	37	427.5	25.0	1.18
CGB	56.4	-50.9	-5594	39	-143.4	-16.3	0.79
CGC	58.0	144.8	15933	32	497.9	25.1	1.00

financial trading [29], [30]. It keeps away from suffering huge losses during a trade if the model predicts trend reversal behavior is continuously identical with the current position. Table 4 and Table 5 showed the results after setting a 2% and a 5% stop-loss order. When Table 4 and Table 5 were analyzed, it was observed that experimental group models performed well under any risk allowance. The average annualized return of the EGB model was over 30.5%, the average annualized return of the EGC model was over 41%, and the average annualized return of the EGD model was at least 24.8%. Each profit to loss ratio of experimental groups was

over 1.0 showing the profitable ability while considering different preferred risk values. In contrast with the experimental groups, all the control groups performed worse than the first stage while taking only 2% risk. It is not intuitive to see that the CGA with higher precision than experimental groups, but the total profit is negative. The model based on the correct prediction to do the transaction should profit, not lose money. However, not only precision but the expected return was the key to making the profit. It was possible to lose money with high precision if the model failed to manage the losses well. Fig. 13 demonstrated trading distributions of

trading return for EGC and CGA at the second stage of the experiment set a 2% stop-loss, showing that CGA suffered a greater number of large losses when it failed to detect the market reversals. In addition, EGB made more huge profits when it was correct. As increasing the risk allowance, control groups performed better. When allowed risk was 5%, the average annualized return of the CGA model was 25%, and the average annualized return of CGC was 25.1%. It showed that LSTM and CNN-LSTM network models trained by time series data might be profitable when high risk is allowed. In contrast with most control groups, the models trained with proposed structures performed well under the 2% risk allowance. The 2% risk setting achieved better financial performance, demonstrating that the models had outstanding ability to identify trend reversals. Lower stop loss provides a greater number of transactions so that more opportunities for making profits. In addition, the precision of taking 2% risk was slightly lower than the 5% risk preference. Taking a higher risk can bear the latency of trend reversals. However, it may cause a huge loss if the model mistakes the reversal of the trend and decreases trading chances. CGB was the worst of all the models. Although its precision was not bad, it failed to manage the expected return since the poor value of profit to loss ratio.

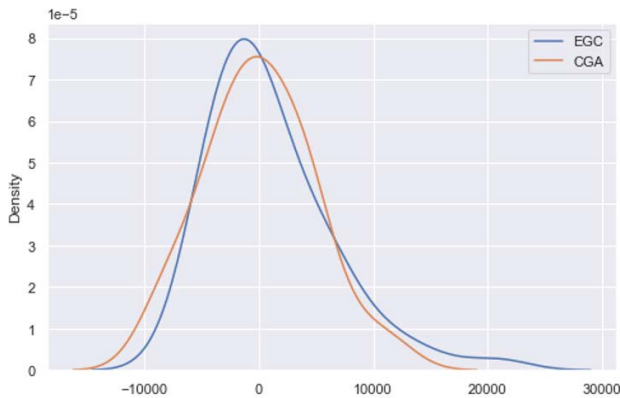


FIGURE 13. The distributions of trading return for EGB and CGA at the second stage of the experiment set a 2% stop-loss.

Fig. 14 analyzed experimental and control groups' financial performance details among the first and second stages with different risk preferences. The overall performance of EGC was the best and stable. The average profit per transaction exceeded \$600, which was better than EGB. It demonstrated the importance of the influence of time causality. The different grayscale values among structures provide the order of extracted patterns that can help the model discover the trend reversal behavior.

C. STATISTICAL SIGNIFICANCE TEST

This study investigated the novel approach as follows: (a) To investigate whether CNNs could extract the trend reversal behavior from proposed structures of price activities. (b) To prove the proposed models can make a profit in practice

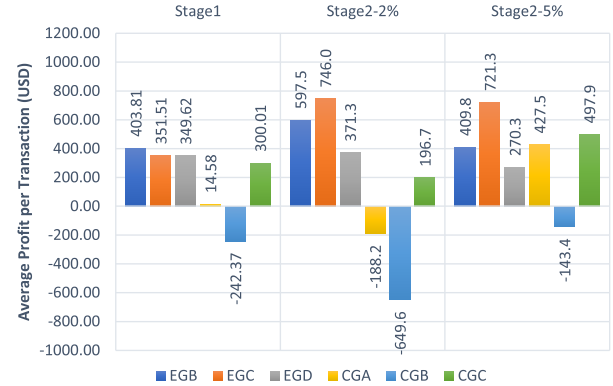


FIGURE 14. The overall financial performance of experiments.

and perform better than common models with sequence data (c) To explore the models that consider time causality is better than the models without time causality.

In this section, a statistical significance test was used to examine and compare the financial performance between each experimental and control group. At first, F-test was applied to test the population variances of financial performance among our models to select the right formula to carry out the T-test. The statistical significance test results of proposed models compared with models for control groups were tabulated in Table 6 and Table 7. It was worth noting that the annualized return of CGB performed not well. Since the base of (16) was negative, the t-test was performed only at the second stage of the experiment, which set a 5% stop-loss. The null hypothesis was expressed as:

$$H_0 = \mu_b \geq \mu_a \quad (18)$$

Table 6 showed the first result that the null hypothesis was rejected at the preliminary experiment of the first stage, which meant splitting data into five structures was better than merging into a single structure. In addition, all the remaining experimental groups indicated the trading performance of the proposed models was significantly better than the control groups. It showed that the proposed model did detect the trend switching behaviors and take the corresponding right actions in the financial market. However, the test result between EGB, EGC, and EGD accepted the null hypothesis, expressing that considering the time causality with structures has no significant difference at the first stage. It indicated that converting the time series data into structure images based on the market profile theory fed into CNNs could be a good model to identify the trend reversal.

The second stage compared the performance of each model in a more realistic context of the transactional environment. Table 7 showed the T-test results under the different proportions of stop-loss. Under the 2% stop loss setting, all null hypothesis was rejected, proposed models had better profitability than the control groups. As increased to a 5% stop-loss setting, the profitability of EGC was still significantly better than all the control groups. The results in Table 3-7

TABLE 6. T-Test results of annualized return of first stage.

Performance metrics	P-Value	T-Test results
EGB-EGA (preliminary experiment)	***	Reject H_0
EGB-CGA	***	Reject H_0
EGB-CGC	***	Reject H_0
EGC-CGA	***	Reject H_0
EGC-CGC	***	Reject H_0
EGD-CGA	***	Reject H_0
EGD-CGC	***	Reject H_0
EGD-EBG	0.4935	Do not reject H_0
EGD-EGC	0.347	Do not reject H_0
EBG-EGC	0.330	Do not reject H_0

TABLE 7. T-Test results of annualized return of second stage.

Risk allowance	Performance metrics	P-Value	T-Test results
2%	EGB-CGA	***	Reject H_0
	EGB-CGC	***	Reject H_0
	EGC-CGA	***	Reject H_0
	EGC-CGC	***	Reject H_0
	EGD-CGA	***	Reject H_0
	EGD-CGC	***	Reject H_0
	EGC-EBG	**	Reject H_0
	EGC-EGD	***	Reject H_0
	EGB-EGD	**	Reject H_0
	EGB-CGA	*	Reject H_0
5%	EGB-CGB	***	Reject H_0
	EGC-CGC	0.053	Do not reject H_0
	EGC-CGA	***	Reject H_0
	EGC-CGB	***	Reject H_0
	EGC-CGC	***	Reject H_0
	EGD-CGA	0.482	Do not reject H_0
	EGD-CGB	***	Reject H_0
	EGD-CGC	0.469	Do not reject H_0
	EGC-EBG	***	Reject H_0
	EGC-EGD	***	Reject H_0
	EGB-EGD	0.056	Do not reject H_0

presented the performance of the EGC against the state-of-the-art deep learning models.

V. CONCLUSION

In this research, we proposed a deep learning method that utilized a CNN with structures of price activities as input to identify trend reversal behavior. First, we adopted market profile theory to generate the structure to convert the financial time series data to a text market profile. Second, using disparate linear interpolation to convert TPO in profile into grayscale values, so that structure images were created. At first, single structure and multiple structures images were fed into CNN networks to learn the feature of trend change and compared. The multiple structures were superior to the single structure so that we focused on multiple structures. In addition, we also considered the patterns with time causality that need to integrate time with structure images by modifying the linear interpolation range to emphasize the importance of time.

To evaluate the model performance, S&P 500 daily mini futures data was collected as our financial time series data to carry out the simulated trades based on model prediction.

A long position would be held while the trend changed from a downward trend to an upward trend. On the other hand, a short position would be acquired when the model predicted the trend changed from an upward trend to a downward trend. If the trend did not change, no transaction was made. Ultimately, two stages of different trading mechanisms were designed to analyze both the model performance and the financial performance among the models. For the financial performance, the first stage results showed that all the experimental groups that extracted the features from multiple structures were significantly better than the control groups that utilized the state-of-the-art deep learning models, including LSTM networks, CNN, and CNN-LSTM networks which learning from time series data. Our proposed model has similar classification capabilities to those networks for the model performance. More complicated but close to real-life trades were carried out at the second stage. We considered holding positions longer and setting a stop-loss. As a result, almost all the experimental groups were performed very well against the control groups when allowing 2% risk. In addition, the model fed with structures integrated with the characters of time emphasizing the occurred time between shapes performed best among experimental groups in financial performance evaluation, indicating the importance of the influence of time causality on features. In the future, we will try to add the characters of the volume or other technical analysis methods into the deep learning model. Furthermore, ensemble learning with the traditional machine learning method is also considered to improve the model's accuracy. In the age of AI, our proposed model that could identify trend reversals will help investors avoid the dilemma of chasing highs and killing lows and increase the efficiency of their capital utilization.

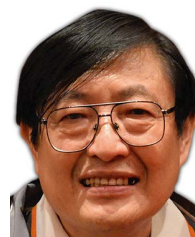
REFERENCES

- [1] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning : A systematic literature review: 2005–2019," *Appl. Soft Comput.*, vol. 90, May 2020, Art. no. 106181, doi: 10.1016/j.asoc.2020.106181.
- [2] O. B. Sezer and A. M. Ozbayoglu, "Financial trading model with stock bar chart image time series with deep convolutional neural networks," *Intelligent Automation and Soft Computing*, vol. 26, no. 2, pp. 323–334, 2020, doi: 10.31209/2018.100000065.
- [3] J. P. Steidlmyer, *Markets and Market Logic*. Chicago, IL, USA: Porcupine Press, 1984.
- [4] J. F. Dalton, E. T. Jones, and R. B. Dalton, *Mind Over Markets: Power Trading with Market Generated Information*, 2nd ed. Greenville, SC, USA: Traders Press, 1999.
- [5] R. L. Galvez, A. A. Bandala, E. P. Dadios, R. R. P. Vicerra, and J. M. Z. Maningo, "Object detection using convolutional neural networks," in *Proc. TENCON IEEE Region Conf.*, Oct. 2018, pp. 2023–2027.
- [6] E. Hoseinzade and S. Haratizadeh, "CNNpred: CNN-based stock market prediction using a diverse set of variables," *Expert Syst. Appl.*, vol. 129, pp. 273–285, Sep. 2019.
- [7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.
- [8] B. Hu, Z. Lu, H. Li, and Q. Chen, "Financial trading model with stock bar chart image time series with deep convolutional neural networks," *Intell. Automat. Soft Comput.*, vol. 26, no. 2, pp. 323–334, 2020, doi: 10.31209/2018.100000065.
- [9] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

- [10] H. Maqsood, I. Mehmood, M. Maqsood, M. Yasir, S. Afzal, F. Aadil, M. M. Selim, and K. Muhammad, "A local and global event sentiment based efficient stock exchange forecasting using deep learning," *Int. J. Inf. Manage.*, vol. 50, pp. 432–451, Feb. 2020.
- [11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 25. Stateline, NV, USA, Dec. 2012, pp. 1097–1105.
- [12] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1725–1732.
- [13] O. B. Sezer and A. M. Ozbayoglu, "Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach," *Appl. Soft Comput.*, vol. 70, pp. 525–538, Sep. 2018.
- [14] C.-C. Lin, C.-S. Chen, and A.-P. Chen, "Using intelligent computing and data stream mining for behavioral finance associated with market profile and financial physics," *Appl. Soft Comput.*, vol. 68, pp. 756–764, Jul. 2018, doi: [10.1016/j.asoc.2017.08.008](https://doi.org/10.1016/j.asoc.2017.08.008).
- [15] C.-C. Chen, Y.-C. Kuo, C.-H. Huang, and A.-P. Chen, "Applying market profile theory to forecast Taiwan index futures market," *Expert Syst. Appl.*, vol. 41, no. 10, pp. 4617–4624, Aug. 2014.
- [16] D. Li and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, 2014.
- [17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, Feb. 2015.
- [18] J. L. Elman, "Finding structure in time," *Cognit. Sci.*, vol. 14, no. 2, pp. 179–211, Mar. 1990.
- [19] M. I. Jordan, "Serial order: A parallel distributed processing approach," in *Advances in Psychology*, vol. 121. 1997, pp. 471–495, doi: [10.1016/S0166-4115\(97\)80111-2](https://doi.org/10.1016/S0166-4115(97)80111-2).
- [20] M. Sundermeyer, R. Schlüter, and H. Ney, "LSTM neural networks for language modeling," in *Proc. Interspeech*, Sep. 2012, pp. 1–4.
- [21] K. Chen, Y. Zhou, and F. Dai, "A LSTM-based method for stock returns prediction: A case study of China stock market," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Oct. 2015, pp. 2823–2824.
- [22] M. Hiransha, E. A. Gopalakrishnan, and V. K. Menon, "NSE stock market prediction using deep-learning models," *Proc. Comput. Sci.*, vol. 132, pp. 1351–1362, Jan. 2018.
- [23] K. Khare, O. Darekar, P. Gupta, and V. Z. Attar, "Short term stock price prediction using deep learning," in *Proc. 2nd IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT)*, May 2017, pp. 482–486.
- [24] X. Zhang and Y. Tan, "Deep stock ranker: A LSTM neural network model for stock selection," in *Proc. Int. Conf. Data Mining Big Data*. Cham, Switzerland: Springer, 2018, pp. 614–623.
- [25] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2017, pp. 1643–1647.
- [26] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN-LSTM model for gold price time-series forecasting," *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17351–17360, 2020, doi: [10.1007/s00521-020-04867-x](https://doi.org/10.1007/s00521-020-04867-x).
- [27] M. Kearns, "A bound on the error of cross validation using the 43 approximation and estimation rates, with consequences for the training-test split," in *Proc. Adv. Neural Inf. Process. Syst.*, 1996, pp. 183–189.
- [28] A.-P. Chen and Y.-C. Hsu, "Dynamic physical behavior analysis for financial trading decision support [application notes]," *IEEE Comput. Intell. Mag.*, vol. 5, no. 4, pp. 19–23, Nov. 2010.
- [29] J. Dhaene and M. J. Goovaerts, "Dependency of risks and stop-loss order," *ASTIN Bull.*, vol. 26, no. 2, pp. 201–212, Nov. 1996.
- [30] M.-E. Wu, J.-H. Syu, J. C.-W. Lin, and J.-M. Ho, "Effective fuzzy system for qualifying the characteristics of stocks by random trading," *IEEE Trans. Fuzzy Syst.*, early access, Aug. 18, 2021, doi: [10.1109/TFUZZ.2021.3105192](https://doi.org/10.1109/TFUZZ.2021.3105192).



CHURN-BIN JU received the B.S. degree from the Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan, in 2011, and the M.S. degree from the Institute of Information Management, National Chiao Tung University, Hsinchu, in 2013. He is currently pursuing the Ph.D. degree with the Institute of Information Management, National Yang Ming Chiao Tung University, Hsinchu. He also works as a Senior Researcher at Chunghwa Telecom Laboratories, Taiwan. His research interests include machine and deep learning, automatic trading systems in the financial market, the Internet of Things, big data, and data analytics.



AN-PIN CHEN received the B.S. degree from the Department of Agricultural Engineering, National Taiwan University, Taiwan, and the M.S. and Ph.D. degrees in industrial engineering from the University of Southern California. He is currently a Professor with the Institute of Information Management, National Yang Ming Chiao Tung University, Taiwan. His current research interests include decision support systems, data mining, financial engineering, knowledge management, and computational intelligence applications in the area of finance. He was elected to a fellow of the Chinese Society for Management of Technology, in 2020.

...