The International Journal of Engineering and Science (IJES)
|| Volume || 9 || Issue || 03 || Series I || Pages || PP 36-43| 2020 ||
ISSN (e): 2319-1813 ISSN (p): 20-24-1805

A Portfolio Trading System of Digital Currencies

Xudong Sun[1], Liguo Weng[1], Min Xia[1]
[1]Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing University of Information Science and Technology, Nanjing, China
Corresponding Author:Min Xia

--ABSTRACT---
Portfolio management is an effective means of investment, but it is a very difficult task to make a proper portfolio management in virtual currency market. A portfolio management system based on deep-reinforcement learning is proposed to conduct more profitable portfolios in virtual currency market.The prices of virtual currency are driven by many factors(such as investors' confidence, emotions,regulations), but the traditional methods can't effectively deal with historical data. To solve this problem, a separable convolution is proposed. In a virtual currency market, the price rise only occurs in a flash. Experimental results in virtual currency market show that our model can get more returns. Besides, the higher Sharpe ratio indicates our method is more resistant to long-term risks.
KEYWORDS:*Portfolio ; Deep-reinforcement learning ; Convolution*

Date of Submission: 01-03-2020 Date of Acceptance: 16-03-2020

I. INTRODUCTION

Portfolio management is a dynamic strategy-decision process, which determines how,when and what to invest. The core of portfolio is resource allocation, determininghow to allocate funds and time. Investors shall focus more on project selection of portfoliosas a result, it demands investors should balance their funds and assets.In order to maximize profits, investors will continue renewing asset flows.During the decision process, all projects will be evaluated, chosen and optimized respectively.Then capital will be distributed to more beneficial projects.

The initial portfolio models composed of complete mathematical algorithmsstarted to emerge in 1960s. These idealized mathematical models are now called traditionalor classical methods, which included BCG, the General Electric, and Shell directional policy matrix etc[1].These methods utilized techniques such as linear programming, dynamic programming,and integer programming. Under the constraints of fixed capital, the objective of portfolio wasto optimize the expected return using existing and new asset items.Although these methods were quite attractive from the perspective of theory,these idealized methods could not be applied in practical markets.The main reasons of this phenomena were dependency on market informationand great uncertainty of market environment. Besides, through practitioners' practical use,they found it hard to comprehend and utilize these methods due to their complexity.Although portfolio management is very difficult, it plays a key rolefinancial problems, and hence a great number of portfolioapplications were introduced, for example, Fagereng[2]listed our portfolio choice in real life, Harvey [3] proposed to apply portfolio in insurance and Bertrand [4] optimized a Bayesian decision theoretic framework in portfolio. These methods realized optimalinvestment with the use of combinatorial optimization, which leads to huge computation. In additon,traditional methods were unable to effectively extract features because of their low efficiency in model optimization. In recent years, with the developmentof artificial intelligence technology [5,6], deep learningtheory has been more and more widely used in the field of datafeature extraction [7], and reinforcement learning theory isconsidered as an effective method of investment decision-making[8,9]. Reinforcement learning(RL) is a dynamic learning of an optimal policy and isused to solve decision making problems in a wide range of fields in natural sciences and engineering[10,11].

For some financial and economic issues, reinforcement learning is more suitable [9,12],like option pricing [13],and multi-period portfolio optimization [14,15],which utilized policy search to learn to trade. In handling some special issuessuch as risk management, reinforcement learning would provide better solutions in[9]and Garcia [16] proposed a new way to evaluate risk values based on deep learning.In the case of turbulent market conditions, investors are usually prone to make reckless investment behaviors [17].To relieve this dilemma, the adaptive markets hypothesis appeared in Khuntia's work [18], which may be realized by reinforcement learning.Some methods have been proposed to apply in portfolio management. The Best Stock (Best) is a benchmark widely used in portfolio selection, whose trading strategy is to invest in assets that have the best returns in the past [19].The Uniform Constant Rebalanced Portfolio(CRP) is a more challenging benchmark, which distributed the investor's funds equally to each asset in each round [20]. Inspired by the mean reversion principle of financial algorithm and the confidence weighting techniques in machine learning, the Confidence Weighted Mean Reversion(CWMR) built a model on the basis of the Gauss distribution and updated the model according to the mean reversion trading principle [21]. Unlike traditional trend tracking methods, the Passive Aggressive Mean Reversion(PAMR) built models on the mean regression relationship of financial markets [22].However, the above methods only improved the optimization strategy, but ignored the processing of historical data. In the virtual currency market, the prices of various assets change rapidly. These methods are difficult to accurately predict the time when assets rise. The recent method with deep-reinforcemnt learning [23]was designed based on CNN [24].Similarly, their methods failed to process historical data effectively.

In this article, we propose a deep-reinforcement-learning network that acts as a moreeffective approach in managing portfolios in virtual currency market, and our main contributions are threefold:

(i) First, we utilize the decision tree to preprocess the historical data.Different from other algorithms using capacity of computation of computers, the proposedmethod can extract more important features as inputs;

(ii) Second, a novel separable convolution and a modified residul module are proposed to extract historical data features. Our separable convolutions are designed to process each input independently helping cope with the rapidly changing market. In addition, our improved residual module is more sensitive to the time of price rise;

(iii) Third, our portfolio management system is applied to the virtual currency market in an end-to-end way. The proposed system directly conducts transactions through optimization strategies instead of predicting the prices of virtual currency. Unlike Our system does not predict the price trend, but directly optimizes the total value of the portfolio.

II. METHODS

Before specifying our proposed networks, we first review the nature of reinforcement learning.Shown in Fig 1, traditional reinforcement learning includes two key elements -- agent and environment,between which RL keeps learning through trials and errors. Despite uncertainty of environment,the agent seeks to achieve its goal. The studying progress of reinforcement learning involvesinteractions between agent and environment. In other words, agent takes actions to explore environment,and then future states of environment will be available to agent as feedbacks.

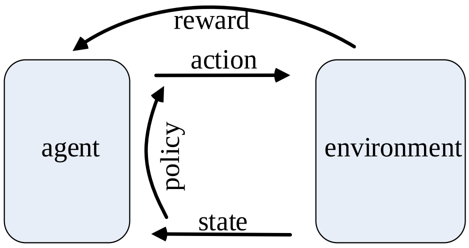


Fig 1. Reinforcement learning

The policy and a reward signal are also main elements of a reinforcement learning system.Roughly speaking, the policy is a mapping function, because it actually determines action ofagent according to current state. In deep-reinforcement learning, the policy is made up ofdeep neural networks where features of environment states will be extracted.In every time step, environment will return a reward signal to agent. Reward signalsare fundamental to adjust policy nets, because when the agent gets feedbacks fromenvironment of low rewards, it will change its policy. Agent's direct goal is to maximize rewards [10].We apply our method to virtual currency market, so the environment is a simulation of virtual currency market.The relevant state is the price of each currency in each time step, and the corresponding action is conductedby reinforcement learning networks to buy or sell certain currency. The essence of this policy is neuralnetworks using deep learning techniques about which we will give detailed descriptions in subsequent section.Reward signals are practical values of total assets and accumulation of rewards is the final portfolio value in the long run.

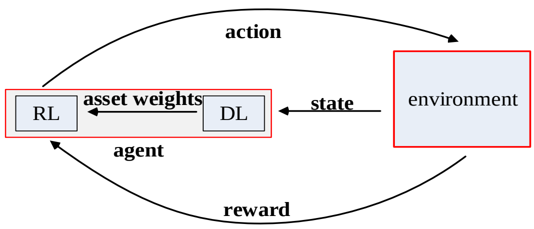


Fig 2. An overview of our proposed system

As illustrated in Fig 2, state of market price is fed into theagent. The price data is composed of m kinds of virtual currency. Tospecifically explain the agent of this paper, we split agentnetworks into two parts which are deep learning networks andreinforcement learning networks. Deep learning networks processstate information to predict asset weights in next time step.Reinforcement learning networks will take action according to asset weights. In next time step, the reward signalwill feedback to reinforcement learning networks. Namely, thepractical value of total assets is available to reinforcementlearning networks as a feedback. In the meanwhile, the present stateof environment is also transit to deep learning networks. This isthe operation of the entire networks.

We shall now illustrate deep learning networks and reinforcement learning networks with details.

Deep learning networks

Assuming that the portfolio is composed of m kinds of assets andthat the number of total time steps is t_f. We define that V_tis composed of close price of all assets in the t-th time step andthat stands for the close price of the i-th asset in the t-th timestep. Similarly, hV_t and lV_t respectively represents high pricevector and low price vector.

We set a constant BTC to 1 and the fund reallocation doesn'tinvolve BTC. The prices of all virtual currency are related to BTC'svalue, so BTC can be viewed as a price metric to evaluate the valueof each asset. The initial fund is 1 BTC and reinforcement learningwill allocate 1 BTC fund to other assets in the following time step.

As illustrated in Fig.2, the state of environment is transited todeep-learning networks.Close price, high price and low price constitute input data X_t.

The aim of using deep learning networks is to extract state features to get W_t --a weight vector of each asset. This process can be represented by a mapping function F(x):

$$W\_t = F(X\_t) \quad (1)$$

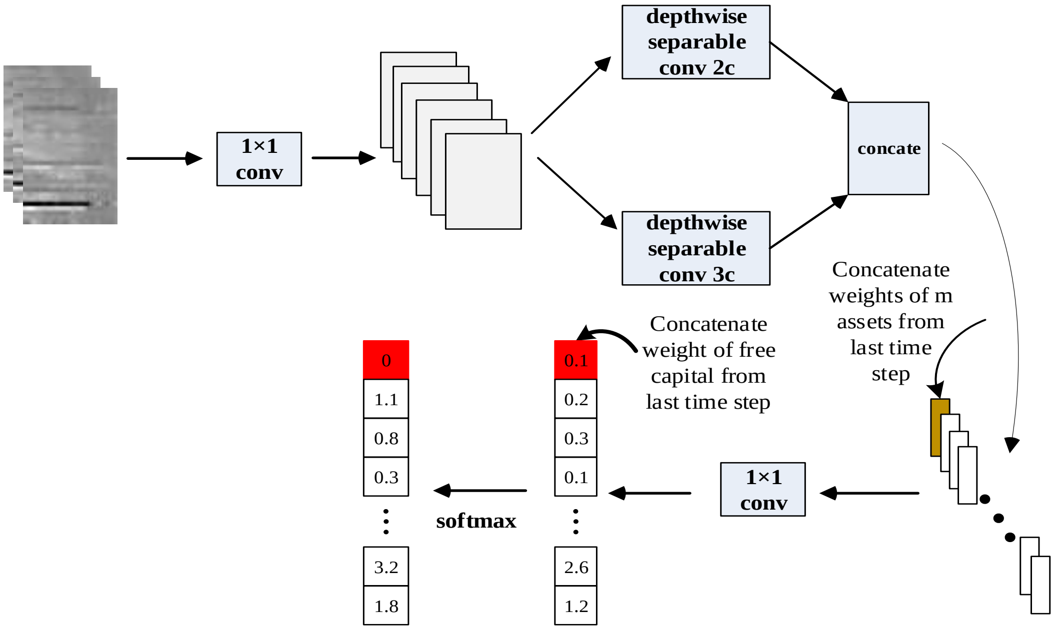


Fig 3. Separable networks

The specific architecture of deep learning networks is illustrated in Fig 3. We begin with our observation of great performance when applying depthwise separable convolutional networks. The results motivated us to proposea unique depth separable convolutional network. As illustrated in Fig.4 and Fig.5,depth-wise separable convolution is a network structure. In Fig.4, separable convolutions perform on each 2channels and we call these two channels as operated channels. This kind of networks is performed onfeature map channels. The input X_t has 6 channels after computation of 1×1 convolutional network. Shown in Fig.4,a 2×2 convolution performs computation on individual feature maps.Due to three different 2×2 convolutions, it realizes the goal to separate feature maps.This construction can achieve better performance while increasing few parameters. Afterwards,it just concatenates features by channel. Unlike Fig 4, Fig 5 presents another structureof separable networks which perform 2×2 convolutionson each 3 channels. Thus, operated channels in Fig.5 is 3.These two slightly different structures enable the networks to obtain richer and better features.

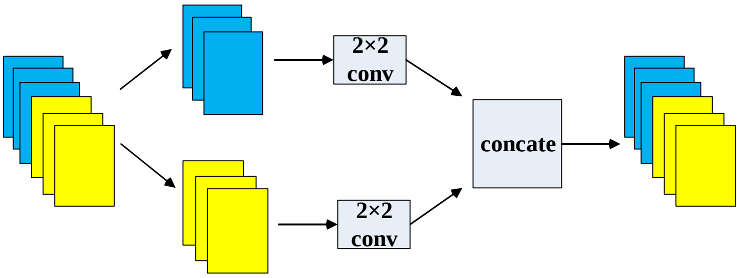


Fig 4. Separable convolutions on two channel

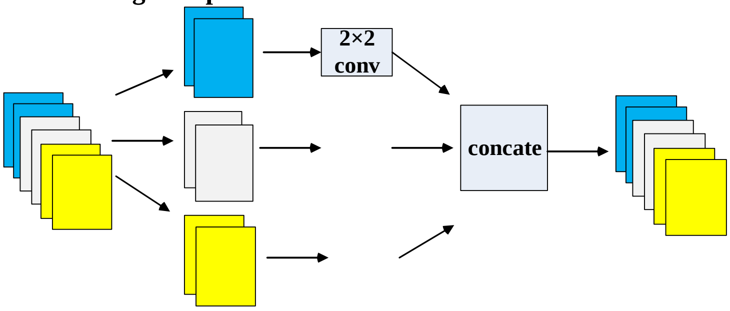


Fig 5. Separable convolutions on three channels

Reinforcement learning networks

Reinforcement learning aims at finding appropriate actions between virtual currency market and the agent.In our proposed reinforcement learning networks, environment is the entire data of virtual currency marketand the data consists of high price, low price, open price, close price, and volume etc.Action is conducted by reinforcement learning networks and to sell or to buy relevant assets isdetermined by their prices. Specifically speaking, we sample data of virtual currency every 30 minutes and we define 30 minutes as a time step. At the end of each time step, reinforcement learning networks will reallocatefunds to all assets according to each asset price.

Selection of features

The environment is just the marketdata and that market data is composed of close price,high price, low price, open price and volume. But market data also includes much redundant information.It's infructuous to deal with useless information. We first pay our attention to choose appropriate features of data.We adopt XGBoost to evaluate importance of all features. XGBoost is designed based on GradientBoosting Decision Tree(GBDT) [25]. Compared to GBDT, XGBoost can perform multi-kernel operation parallelly.XGBoost conducts feature evaluation through increasing trees and each tree in XGBoost is interrelated.

The selection of features is determined by the importance of features. XGBoostcalculates scores of feature importance through decision trees. The features ofvirtual currency market are high price, low price, open price, close price andvolume. We choose 20 virtual currency as market data which constitute our portfolio.

We respectively calculate scores of importance in each currency and then just sum them upto get score of the whole portfolio. The feature scores in each currency are listed inTable1, we can figure out thatclose price, high price and low price contribute most to the performanceof our networks. So we will treat these three features as input in the following experiments.

currency	Close(%)	High(%)	Low(%)	Open(%)	Volume(%)
xmr	66.4	10.9	17.4	2.5	2.8
eth	66.1	9.5	12.6	5.9	5.9
zec	39.8	24.7	15.8	10.5	9.2
fct	63.9	9.5	18.9	7.7	0
rusdt	83.3	2.5	9.2	5.0	0
etc	66.6	12.3	19.3	1.8	0
rep	62.8	15.9	17.1	2.5	1.7
xrp	100	0	0	0	0
dash	52.7	18.6	21.0	7.0	0.7
maid	86.7	2.5	10.0	0.8	0
ste	82.8	3.1	12.5	0	1.6
lsk	88.6	5.7	5.7	0	0
ltc	59.4	16.0	14.4	9.8	0.4
gam	100	0	0	0	0
xem	83.7	7.0	7.0	2.3	0
nav	100	0	0	0	0

bts	100	0	0	0	0
dog	42.7	22.3	20.9	9.7	4.4
sc	100	0	0	0	0
xcp	76.3	23.7	0	0	0

Table 1. Results of feature importance in each currency

III. EXPERIMENTS AND DISCUSSION

Portfolio planning problems include various practical restrictions such as delay of transactions, disturbance outside the market and limited transaction time. Unlike traditional financial products, the most notable advantages of virtual currency are as follows:

(i)virtual currency can be traded without delay at any time in a day;

(ii)virtual currency market has low correlation with other assets, so external disturbance won't affect the market.

Experiments	Training time	Testing time
validation	2014/02/01 to 2015/12/04	2015/12/05 to 2016/02/01
Back test1	2014/08/01 to 2016/06/04	2016/06/04 to 2016/08/01
Back test2	2015/02/01 to 2016/12/04	2016/12/05 to 2017/02/01
Back test3	2015/08/01 to 2017/06/04	2017/06/04 to 2017/08/01
Back test4	2016/02/01 to 2017/12/04	2017/12/05 to 2018/02/01

Table 2. Details of time

So, the experimental data is from real virtual currency market and the historical data is accessed from Poloniex's official dataset. We set the number of assets m to 20.We choose these 20 assets because trading volumes of them are bigger than other assets.In view of the goodness of a hypothetical strategy, there is no good measure other than back test.In common, back test stands for an assessment of historical performance of a suggested trading strategy.We set four back test experiments to evaluate our method. The first three sets of experiments were testedin the case of a bull market. In addition, the last set of experiment was conducted in a bear market.To adjust hyper parameters, we also set a cross validation experiment. In each experiment we split 92% ofthem as training data and the rest is test data. Details of time range are listed in Table 3.

An optimal portfolio is to measure assets with respect to both returns and risks. Markowitzdefined such a portfolio as an efficient portfolio, that is, it offers the highest level ofreturn for a given level of risk and the lowest level of risk for a given level of return.

In terms of return, we define the average return as follows:

$$R_{ave} = \frac{\sum_i R_i}{m} \quad (2)$$

where m stands for the number of assets (in our method we set m equal to 20) and R_i means the return of i-th asset.

Sharpe ratio is a standard index to measure portfolio performance, which is definedas expected return in excess of the risk-free rate over the portfolio standard deviation [26,27].If Sharpe ratio is positive, the average growth rate of asset exceeds the risk-free rate.

$$SR = \frac{E(\mathrm{R}) - \mathrm{R_f}}{\sigma(\mathrm{R})} \quad (3)$$

where E(R) is the average of returns and R_f is risk-free return. σ(R) is standard deviation of returns.

Another risk indicator is the maximum drawdown. Maximum drawdown is used to assessthe relative risk of two adjacent time periods. The maximum drawdown measures thelargest decline in portfolio value. A low maximum drawdown is preferred as this indicates losses from investment are small [28].The definition of maximum drawdown is formulated as follows:

$$MD = \frac{R_t - R_{t+1}}{R_t} \quad (4)$$

where R_t stands for return of t-th time period and Rt+1stands for return of t+1-th time period.

In the following subsections, we present experimental results.

Experiments of separable convolution strategy

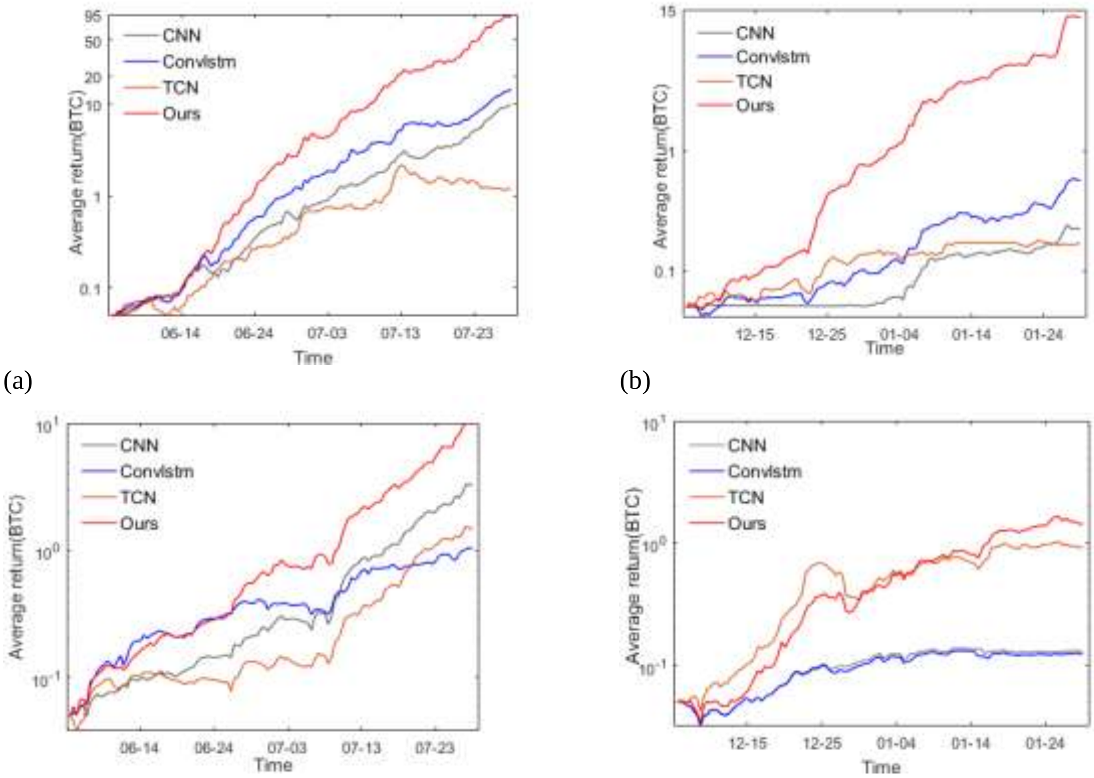
Separable convolutions function as primal feature extraction in our method.Separable convolutions are conducted on channel-wise features,so the number of operated channels is a key component.Experimental data is cross validation data in Table 2.The training time range is from 2014/02/01$ to $2015/12/04 and the range oftesting data is from 2015/12/05 to 2016/02/01. Results are listed in Table 3.From the perspective of average return,the combination of two-channel and three-channel structure performs better than other structures. Sharpe ratio indicates long-term risk of each structure.Obviously, combination of two-channel and three-channel is more resistant to long-term risk.While maximum drawdown shows that combination of two-channel and three-channel is hard todeal with short-term risk. For the purpose of maximizing average return, we consider thecombination of two-channel and three-channel as the best choice.

Channel	SR	MD (%)	average return (BTC)
only 2c	0.112	17.0	4.713
only 3c	0.091	20.3	1.778
2c+3c	0.126	21.9	12.754

Table 3. Experiments on separable convolution strategy

Experiments on different deep-learning networks

We now compare our method with different deep-learning networks,which are respectively based on CNN, convlstm and TCN. The trainingtime and testing time of three backtests are listed in Table 2.



(a) (b) (c) (d)

Fig 6. Results of backtests

Experimental results are shown in Fig 6 and fianl evaluation indicies are listed from Table 4 to Table 7. The three networks CNN, convlstm and TCN are all profitable in themarket as illustrated in the above tables and figures. But theirchangeable results show that these three methods are not robust to the market and they can't precisely predict the market changing trends. In different backtest experiments,the average return may vary dramatically. By contrast, our designed networks process the three-channel inputs independently by our separable convolution, and detects the critical moment of price rising with the modified residual module. As its high Sharpe ratio shows, our method can achieve a high return on venture investment and this verifies that our approach is helpful to catch the critical moments. In good marketconditions like backtest 1, our networks achieve amazing gains. It can get an average return of 93.817 and its Sharp ratio is the highest of the four methods.In a relatively stable market like backtest 2 and backtest 3, our approach is very steady. In the two backtests, average returns of our method are 13.375 and 10.584, which is much higher than that of other methods. In addition, our modified residual module not only seizes the opportunity of price rising, but also reduces the long-term risk to a certain degree. When it comes to the maximum drawdown, our method is unsatisfactory. Its high rateof maximum drawdown means that our method is unstable in a specific period.While in the bad market conditions like backtest 4,our designed networks still generates considerable revenue. However, the other three methods are not as profitable as they used to be.Actually, our proposed networks are constructed based on TCN.Compared to TCN, our method not only increases the average returnand Sharpe ratio but also controls maximum drawdown to a certain degree.

algo	SR	MD(%)	average return(BTC)
CNN	0.110	35.8	10.108
convlstm	0.020	16.5	0.079
TCN	0.063	51.9	1.482
ours	0.137	26.0	93.817

Table 4.Backtest1

algo	SR	MD(%)	average return(BTC)
CNN	0.058	17.8	0.215
convlstm	0.020	18.6	0.601
TCN	0.036	30.7	0.0169
ours	0.123	23.8	13.375

Table 5.Backtest 2

algo	SR	MD(%)	average return(BTC)
CNN	0.090	27.0	3.442
convlstm	-0.001	46.2	0.041
TCN	0.076	37.9	1.590
ours	0.111	30.1	10.584

Table 6 Backtest 3

algo	SR	MD(%)	average return(BTC)
CNN	0.041	36.6	0.122
convlstm	0.04	38.7	0.126
TCN	0.058	52.8	0.982
ours	0.063	38.51.492	

Table 7.Backtest 4

IV. CONCLUSION

Our networks are designed as an optimization of virtual currency portfolio.Our proposed method combines deep-learning networks and reinforcement-learning networks.Its purpose is to purchase or sell relevant currency to increase total assets. Besides,our method not only increases asset value but also controls risk. As shown in the above experiments,our method can increase average return to a large degree and is resistant to long-term risk.Experimental results indicate that our approach can be used for long-term investment in virtualcurrency market. However, the short-term risk is hard to avoid and intense market volatilityusually happen in virtual currency market. So further research is required to improve itsstability in short-term investment. On the other hand, traditional mathematical methods haveguiding significance to portfolio management, therefore we need to find a way to combine traditional methods with deep-reinforcement learning.

REFERENCE

[1]. Jackson B. Decision methods for selecting a portfolio of R&D projects. Research Management, 1983, 26(5): 21-26.
[2]. Fagereng A, Gottlieb C, Guiso L. Asset market participation and portfolio choice over the life cycle. The Journal of Finance, 2017, 72(2): 705-750.
[3]. Harvey C R, Liechty J C, Liechty M W, et al. Portfolio Insurance: The Extreme Value Approach Applied to the CPPI Method. Extreme Events in Finance: A Handbook of Extreme Value Theory and Its Applications, 2016: 465.
[4]. Bertrand P, Prigent J L. Portfolio selection with higher moments. Quant. Financ. 2010, 10(5): 469-485.
[5]. Xia M, Zhang C, Weng L, Liu J, and Wang Y, Robot path planning based on multi-objective optimization with local search, Journal of Intelligent and Fuzzy Systems, 2018, vol.35, 1755-1764.
[6]. Xia M, Shen M, Wang J, et al. Anti-spurious-state neural network using nonlinear outer product and dynamic synapses. Neural Network World, 2016, 26(4): 377.
[7]. Long W, Lu Z, Cui L. Deep learning-based feature engineering for stock price movement prediction. Knowledge-Based Syst. 2019, 164: 163-173.
[8]. Wang Y, Geng S, Gao H. A proactive decision support method based on deep reinforcement learning and state partition. Knowledge-Based Syst. 2018, 143: 248-258.
[9]. Deng Y, Bao F, Kong Y, et al. Deep direct reinforcement learning for financial signal representation and trading[J]. IEEE transactions on neural networks and learning systems, 2016, 28(3): 653-664.
[10]. Sutton R S, Barto A G.Reinforcement learning: An introduction. MIT press, 2018:1-7.
[11]. Bertsekas D P, Yu H.Q-learning and enhanced policy iteration in discounted dynamic programming. Math. Oper. Res. 2012, 37(1): 66-94.
[12]. Liu W, Zhuang P, Liang H, et al. Distributed Economic Dispatch in Microgrids Based on Cooperative Reinforcement Learning. IEEE transactions on neural networks and learning systems, 2018, 29(6): 2192-2203.
[13]. Maestre R, Duque J, Rubio A, et al. Reinforcement Learning for Fair Dynamic Pricing. Proceedings of SAI Intelligent Systems Conference.Springer, Cham, 2018: 120-135.
[14]. Brandt M W, Goyal A, Santa-Clara P, et al. A simulation approach to dynamic portfolio choice with an application to learning about return predictability. The Review of Financial Studies, 2005, 18(3):831-873.
[15]. Heaton J B, Polson N G, Witte J H. Deep learning for finance: deep portfolios. Appl. Stoch. Models. Bus. Ind. 2017, 33(1): 3-12.
[16]. Garcia G M, Carsteanu A A, Clempner J B. Continuous-Time Reinforcement Learning Approach For Portfolio Management With Time Penalization. Expert Syst. Appl. 2019.

[17]. Prashanth L A, Jie C, Fu M, et al. Cumulative prospect theory meets reinforcement learning: Prediction and control. International Conference on Machine Learning. 2016: 1406-1415.

[18]. Khuntia S, Pattanayak J K. Adaptive market hypothesis and evolving predictability of bitcoin. Econ. Lett. 2018, 167: 26-28.

[19]. Borodin A, El-Yaniv R, Gogan V. Can we learn to beat the best stock[C]//Advances in Neural Information Processing Systems. 2004: 345-352.

[20]. Kelly Jr J L. A new interpretation of information rate[M]//The Kelly Capital Growth Investment Criterion: Theory and Practice. 2011: 25-34.

[21]. Li B, Hoi S C H, Zhao P, et al. Confidence weighted mean reversion strategy for online portfolio selection[J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2013, 7(1): 1-38.

[22]. Li B, Zhao P, Hoi S C H, et al. PAMR: Passive aggressive mean reversion strategy for portfolio selection[J]. Machine learning, 2012, 87(2): 221-258.

[23]. Jiang Z, Liang J. Cryptocurrency portfolio management with deep reinforcement learning. 2017 Intelligent Systems Conference(IntelliSys). IEEE, 2017: 905-913.

[24]. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[25]. Chang Z, Hua Y , Yijie D , et al. Multi-scale encoding of amino acid sequences for predicting protein interactions using gradient boosting decision tree. PLOS ONE, 2017, 12(8):e0181426.

[26]. Jorion P. International portfolio diversification with estimation risk. Journal of Business, 1985: 259-278.

[27]. Sharpe W F. The sharpe ratio. J. Portf. Manage. 1994, 21(1): 49-58.

[28]. Magdon-Ismail M, Atiya A F. Maximum drawdown. 2004.

Min Xia, Et.Al "A Portfolio Trading System of Digital Currencies " *The International Journal of Engineering and Science (IJES)*, 9(03) (2020): 36-43.