(Re-) Imag(in)ing Prices Trend

Jingwen Jiang, Bryan Kelly, Dacheng Xiu

Hong Kong University of Science and Technology

May 2, 2022

Presented by: Alvin Lo, Ning Liu, Tony Tan, Ziqing Guo

Outline

Introduction

INTRODUCTION

Main Question: Can we predict stock price? If so, how to do it?

- Some research say yes, some research say no.
- This paper: YES, by using Convolutional Neural Network (CNN).
- Not surprised that we can. We can analyse data and build Mathematical model.
- But CNN? Image?

CONTRIBUTION

Contribution:

- Develop a CNN model to predict stock return.
- Importance: image can be helpful too (New Thing).

Image of the historical price plot as input!

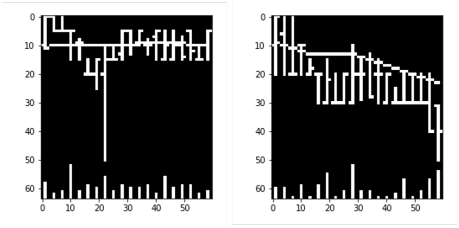

Figure 1: Historical price plot as input image.

WHY USE CNN/IMAGE?

Three reasons to consider using image/CNN:

1. Automated signal generation by CNN
2. Image representations of data allows the model to focus on relational attributes of the data.
3. Imaging process transform historical data to a comparable scale.

OUR REPORT

There are two main tasks of our report:

- Replicate the result from the paper.
- Extend the model to test new implications.

Model

MODEL BASIC

- The model use CNN to predict stock price.
- Three different models: 5-days/20-days/60-days.
- Increasing complexity when we increase the prediction period.

Input of the model: Time series plot of the stock price. Our

Task: Focus on the 20-days Model, replicate and extend!

ARCHITECTURE OF CNN

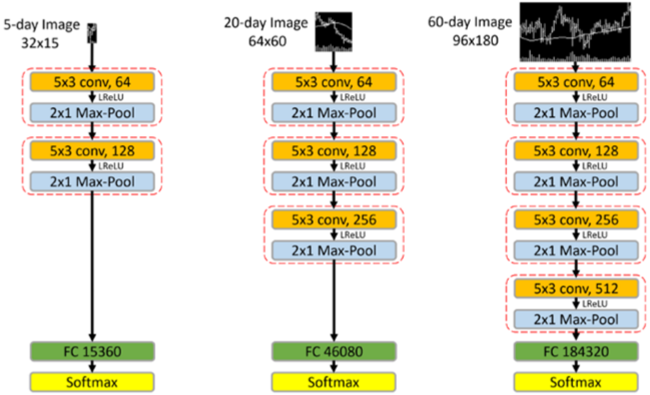

Figure 2: The building blocks of the CNN model.

More details

Replication

REPLICATION SUMMARY

- Model accuracy: 54% (Validation) and 52% (test)
- Other performance statistics also match the finding by the original paper [**Table 1**].
- CNN returns cannot be explained by traditional predictors [**Table 2**].

More details of the traditional predictors

REPLICATION RESULT - PORTFOLIO RETURN

	Equally-Weighed		Value-Weighted	
Portfolio	Ret	Sharpe Ratio	Ret	Sharpe Ratio
1	−0.68%	-0.04	6.38%	0.38
2	5.08%	0.26	4.31%	0.26
3	8.55%	0.44	5.30%	0.32
4	9.20%	0.47	7.82%	0.49
5	10.72%	0.56	6.36%	0.39
6	12.36%	0.64	7.49%	0.47
7	12.20%	0.64	7.80%	0.48
8	13.85%	0.74	7.78%	0.52
9	15.48%	0.81	6.28%	0.39
10	19.96%	0.91	10.07%	0.64
H-L	**20.34%****	2.05	**3.69%***	0.35
p-value		0.00		0.06
turnover		174%		181%

Table 1: We divide the portfolio into 10 monthly deciles. Each has different probability of positive return (decile 1 has the lowest, while decile 10 has the highest). The H-L portfolio represents the portfolio constructed by long decile 10 and short decile 1 combination.

REGRESSION RESULT - UNEXPLAINABLE FACTOR

	Equally-Weighted			Value-Weighted		
	Fama-French Five Factor	Momentum/ Reversal	LowVol	Fama-French Five Factor	Momentum/ Reversal	LowVol
const (alpha)	**0.02***	**0.02***	**0.02***	**0.00***	**0.00***	**0.00***
	(8.92)	(9.04)	(8.87)	(2.34)	(1.89)	(1.49)
CMA	-0.06			-0.20		
	(-0.49)			(-1.56)		
HML	0.08			0.02		
	(0.98)			(0.20)		
Mkt-RF	-0.19*			-0.18*		
	(-3.78)			(-3.25)		
RMW	0.15			-0.03		
	(1.69)			(-0.35)		
SMB	0.09			-0.06		
	(1.34)			(-0.79)		
Momentum		0.07*			0.08*	
		(1.86)			(2.19)	
Reversal		-0.08*			-0.20*	
		(-1.66)			(-3.84)	
LowVol			0.00*			-0.01
			(0.15)			(0.82)

Table 2: The result shows that under the baseline CNN specification, the model predicted returns cannot be explained by using traditional models. However, the CNN model does not capture much additional information for the value-weighted portfolio.

Extension

EXTENSION 1: GRAD-CAM

Issue: We have the model to predict ups/downs, can we interpret the result?

- Grad-Cam can be useful.
- Provides insight: when will the model predicts positive/negative?

EXTENSION 1: GRAD-CAM

New insights from the Grad-Cam:

1. Model activated most heavily on days with high volatility. For low-volatility days, CNN may be activated at moving the average line.
2. Model pays attention to open/close prices.
3. High-volume days are activated more intensively (more important).
4. For "UP" image, more attention on bottom regions, uniform attention for "DOWN" images.

▸ heatmaps

EXTENSION 2: ROBUSTNESS

We check whether the model result is robust to model setting changes, including:

1. Number of filters
2. Number of layers
3. Dropout
4. Batch normalization
5. Initializer
6. Activation function
7. Filter size
8. Dilation/Stride

EXTENSION 2: ROBUSTNESS

	Variations	Acc.		Sharpe Ratio	
		V	T	EW	VW
Baseline		**0.538**	**0.522**	**1.961**	**0.472**
Filter(64)	32	0.536	0.523	1.981	0.283
	128	0.540	0.524	1.904	0.486
Layer(3)	2	0.530	0.529	1.674	0.318
	4	0.539	0.520	2.029	0.372
Dropout(0.50)	0.00	0.515	0.523	1.644	0.097
	0.25	0.534	0.520	1.720	0.169
	0.75	0.540	0.520	1.904	0.063
BN(yes)	no	0.543	0.520	1.806	0.457
Xavier(yes)	no	0.527	0.528	1.943	0.379
Activation (LReLU)	ReLU	0.533	0.527	1.653	0.230
Max-pool Size (2×1)	(2×2)	0.534	0.523	1.678	0.471
Filter Size (5×3)	(3×3)	0.528	0.51	1.599	0.266
	(7×3)	0.532	0.515	1.854	0.307
Dilation/Stride(2,1)/(3,1)	(2,1)/(1,1)	0.527	0.501	1.759	0.155
	(1,1)/(3,1)	0.542	0.526	2.069	0.614
	(1,1)/(1,1)	0.518	0.492	1.707	0.557

Table 3: Table shows robustness check of the model. The model is very robust to model setting changes, in terms of the accuracy and the predicted Sharpe ratio.

EXTENSION 3: REGRESSION

Can we do more than predicting ups/downs?

- Imagine two scenarios:
  1. Model predicts the return will up by 4 percent.
  2. Model predicts the return will up by 0.1 percent.
- Will you make the same investment decision in these two cases?

Model that can predict actual return is more beneficial from the investment perspective.

More details

EXTENSION 3: REGRESSION RESULT

		Return Regression	**Sharpe Regression**	**Baseline Classification**
Loss	Validation	0.08	0.08	-
	Test	0.08	0.08	-
R-Square	Validation	0.01%	1.14%	-
	Test	−0.10%	−0.17%	-
Correlation	Pearson	0.05	0.04	0.04
	Spearman	0.07	0.06	0.06
SR	Equal-Weight	1.98	2.13	2.05
	Value-Weight	0.61	0.47	0.35
Vol	Equal-Weight	13.94%	10.84%	9.92%
	Value-Weight	12.68%	10.84%	9.92%
Max Draw-Down/Vol	Equal-Weight	**0.34**	0.40	0.64
	Value-Weight	**1.92**	3.03	3.05

Table 4: The table summarises the performance of using regression to predict the stock market return. In the regression result, the Sharpe Ratio result has similar performance as that in the Baseline classification, while the Max Draw-Down improves.

EXTENSION 3: REGRESSION

- Regression result shows some investment performance improvement.
- We also conduct other analysis using regression based model:
    - Extracting explainable components using traditional predictors (Fama-French Five Factors).
    - Using the Sharpe ratio as the label.
    - Using a **partially-connected network**.
- Result are robust to model setting changes.

More details

CONCLUSION

This paper:

- Explores if CNN can be useful in stock return predictions.
- Point out the benefits of using images to predict the stock market.

What we do:

- Replicate the model.
- Explores numerous extensions:
  1. GradCam
  2. Classification $\Rightarrow$ Regression
  3. Robustness

Future research:

- Improving the CNN regression model in portfolio design.

Thank You!

Appendix

THE OHLC CHART

Figure 3: Description of the OHLC chart. The paper uses the format displayed on the left (Down OLHC): the high and low price in the day is represented by the upper and lower end of the bar, while the opening and the closing price are represented by the small horizontal line on the left and right of the bar.

Go Back

DETAIL DESCRIPTION OF THE BASELINE ARCHITECTURE

Here is the detail description of the model building blocks:

Components	**Details**
Input Channels	1
Number of layers	3
Filters in each layers	64, 128, 256
Filter size	(5,3)
Dilation	(2,1)
Strike	(3,1)
Parameter initialization	Xavier Initializer
FCC Dimension	40680

Table 5: Summary of the model building blocks.

Go Back

MODEL DESCRIPTION

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1           [-1, 64, 65, 60]           1,024
         LeakyReLU-2           [-1, 64, 65, 60]               0
         MaxPool2d-3           [-1, 64, 32, 60]               0
       BatchNorm2d-4           [-1, 64, 32, 60]             128
            Conv2d-5          [-1, 128, 15, 60]         123,008
         LeakyReLU-6          [-1, 128, 15, 60]               0
         MaxPool2d-7           [-1, 128, 7, 60]               0
       BatchNorm2d-8           [-1, 128, 7, 60]             256
            Conv2d-9           [-1, 256, 7, 60]         491,776
        LeakyReLU-10           [-1, 256, 7, 60]               0
        MaxPool2d-11           [-1, 256, 3, 60]               0
      BatchNorm2d-12           [-1, 256, 3, 60]             512
          Flatten-13                [-1, 46080]               0
          Dropout-14                [-1, 46080]               0
           Linear-15                    [-1, 2]          92,162
================================================================
Total params: 708,866
Trainable params: 708,866
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.01
Forward/backward pass size (MB): 11.31
Params size (MB): 2.70
Estimated Total Size (MB): 14.03
----------------------------------------------------------------
```

Figure 4: Model summary

DETAILS OF REGRESSION RESULT

- We try two different labels:
  1. N-days ahead return
  2. N-days ahead Sharpe ratio
- Loss function: Mean Square Error
- We applied normalization to the data.
- To ensure the model stability, we also test the result with different learning rate.

Go Back

TRADITIONAL PREDICTORS

- Our goal is to find out the component of the stock return predicted by the CNN model which is not explained by using traditional stock return predictor.
- Method: **Regression and find out whether the constant is significant.**
- In JBD, they use five traditional return predictors: Mean return, market risk-free rate, momentum, short-term reversal, weekly short-term reversal.
- Due to limitation of data availability, we use:
  - Fama-French Five Factor Model
  - Momentum/Reversal
  - Low volatility model

Go Back

GRAD-CAM HEATMAP "UP"

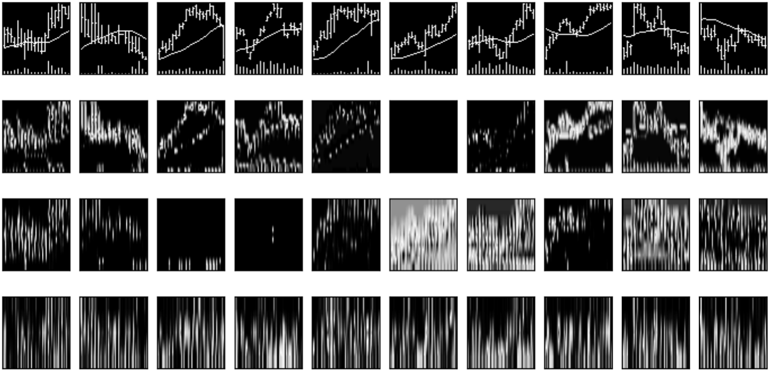

Figure 5: Images receiving "Up" classification.

Go Back

GRAD-CAM HEATMAP "DOWN"

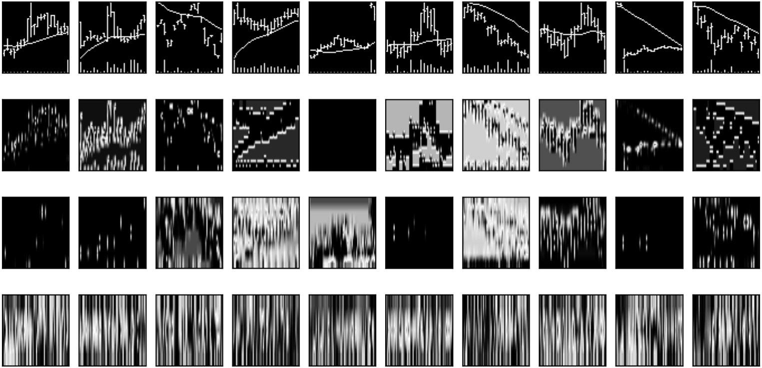

Figure 6: Images receiving "Down" classification.

Go Back

REGRESSION - SHARPE RATIO

	Value-Weighted			Equally-Weighted		
	Fama-French Five Factor	Momentum/ Reversal	LowVol	Fama-French Five Factor	Momentum/ Reversal	LowVol
const	0.01*	0.00*	0.00*	0.02*	0.02*	0.02*
	(2.79)	(2.21)	(2.13)	(9.79)	(9.49)	(9.37)
CMA	-0.12			0.04		
	(-0.84)			(0.30)		
HML	-0.02			0.10		
	(-0.26)			(1.18)		
Mkt-RF	-0.28*			-0.28*		
	(-4.79)			(-5.24)		
RMW	0.01			0.02		
	(0.14)			(0.21)		
SMB	0.10			0.10		
	(1.21)			(1.38)		
Momentum		0.18*			0.13*	
		(4.41)			(2.28)	
Reversal		-0.08*			-0.11*	
		(-1.37)			(-1.93)	
LowVol			0.00			-0.05*
			(-0.06)			(-2.41)

Table 6: Regression on the Sharpe ratio (* indicates statistical significance at 5% level). The result shows that the constant term is significant, meaning that using the regression specification can detect meaningful alpha.

Go Back

REGRESSION - ALPHA SEARCH

	Value-Weighted			Equally-Weighted		
	Fama-French Five Factor	Momentum/ Reversal	LowVol	Fama-French Five Factor	Momentum/ Reversal	LowVol
const	**0.01***	**0.01***	**0.01***	**0.02***	**0.02***	**0.02***
	(3.34)	(3.30)	(2.87)	(8.72)	(8.96)	(8.90)
CMA	-0.08			0.03		
	(-0.50)			(0.21)		
HML	0.09			0.10		
	(0.82)			(0.89)		
Mkt-RF	-0.30*			-0.30*		
	(-4.57)			(-4.14)		
RMW	-0.04			0.08		
	(-0.34)			(0.66)		
SMB	0.11			0.14		
	(1.23)			(1.47)		
Momentum		0.19*			0.13*	
		(4.40)			(2.50)	
Reversal		-0.25*			-0.17*	
		(-4.12)			(-2.48)	
LowVol			-0.01			-0.10*
			(-0.56)			(-3.93)

Table 7: Loadings on Academic Factors, Regression on Returns (* indicates statistical significance at 5% level). The significance of the constant term indicates that the return generated from the CNN regression model is orthogonal to the known factors in the literature, implying the additional benefits of considering the image-based model relative to using the traditional model.

REGRESSION - PARTIALLY CONNECTED

		Baseline Classification	Regression: Sharpe	Regression: Sharpe + Partially Connected	Regression: Returns	Regression: Returns + Partially Connected
Loss	Validation	0.69	0.08	0.08	0.08	0.08
	Test	0.70	0.08	0.08	0.08	0.08
R-square	Validation	-	1.14%	1.02%	0.01%	0.88%
	Test	-	−0.17%	0.34%	−0.10%	0.31%
Correlation	Pearson	0.04	0.04	0.04	0.05	0.05
	Spearman	0.06	0.06	0.07	0.07	0.08
Sharpe Ratio	Equal-Weight	2.05	2.13	1.91	1.98	1.82
	Value-Weight	0.35	0.47	0.60	0.61	0.71
Vol	Equal-Weight	9.92%	10.84%	13.10%	13.94%	16.00%
	Value-Weight	10.64%	11.36%	11.88%	12.68%	15.40%
Max Draw-Down/Vol	Equal-Weight	0.64	0.40	0.44	0.34	0.41
	Value-Weight	3.05	3.03	2.53	1.92	1.28

Table 8: Performance of Regression + Sparse Linear Layers. This table shows that the Sharpe ratio and maximum draw-downs under value-weighted scheme are improved with partially-connected layers.