GROUP 2

(Re)Imag(in)ing Price Trends

EMANUELE SALA 3137553
FABIO GIOVANNI STEFANA 3121555
LUCA SOLERI 3122497

CONTENT

THE IDEA

(Re-)Imag(in)ing Price Trends*

Jingwen Jiang
Department of Computer Science
University of Chicago

Bryan Kelly
Yale University, AQR Capital
Management, and NBER

Dacheng Xiu
Booth School of Business
University of Chicago

Abstract

We reconsider the idea of trend-based predictability using methods that flexibly *learn* price patterns that are most predictive of future returns, rather than testing hypothesized or pre-specified patterns (e.g., momentum and reversal). Our raw predictor data are images—stock-level price charts—from which we elicit the price patterns that best predict returns using machine learning image analysis methods. The predictive patterns we identify are largely distinct from trend signals commonly analyzed in the literature, give more accurate return predictions, translate into more profitable investment strategies, and are robust to a battery of specification variations. They also appear context-independent: Predictive patterns estimated at short time scales (e.g., daily data) give similarly strong predictions when applied at longer time scales (e.g., monthly), and patterns learned from US stocks predict equally well in international markets.

Predicting future returns based on past prices is hard

PC-LAB 4!!!

THE IDEA

Predicting future returns based on past prices is hard

THE IDEA

Predicting future returns based on past prices is hard

Traditional empirical finance uses standard methods to detect patterns like **momentum** and **reversal** in price data.

THE IDEA

Predicting future returns based on past prices is hard

Traditional empirical finance uses standard methods to detect patterns like **momentum** and **reversal** in price data.

HOWEVER...

Prices may hold subtle, **complex patterns** missed by standard finance methods. This paper reevaluates price trends using **machine learning**.

DATA

Kelly et al.

Daily data for all stocks in the **NYSE**, **AMEX** and **NASDAQ**.
Training set is comprised of data from 1993 to 2000.
While the Out of Sample test set 2001 to 2019.

Our study

Daily stock data for all **S&P 500** companies.
Training set from Jan 1st 2010 to Dec 31st 2019.
Out of sample test set from Jan 1st 2021 to Nov 24th 2024.

We decided to skip 2020 entirely as it’s a year where significant price shifts are mainly due to exogenous components.

THE METHODOLOGY

Step 1

Create images to represent price trends. They used **OHLC charts** with the addition of **volume** and a **moving average**

Step 2

Train a **CNN** to extract features and patterns that help us predict future stock returns based on price chart images, followed by a **classification head** to estimate the probability of positive returns.

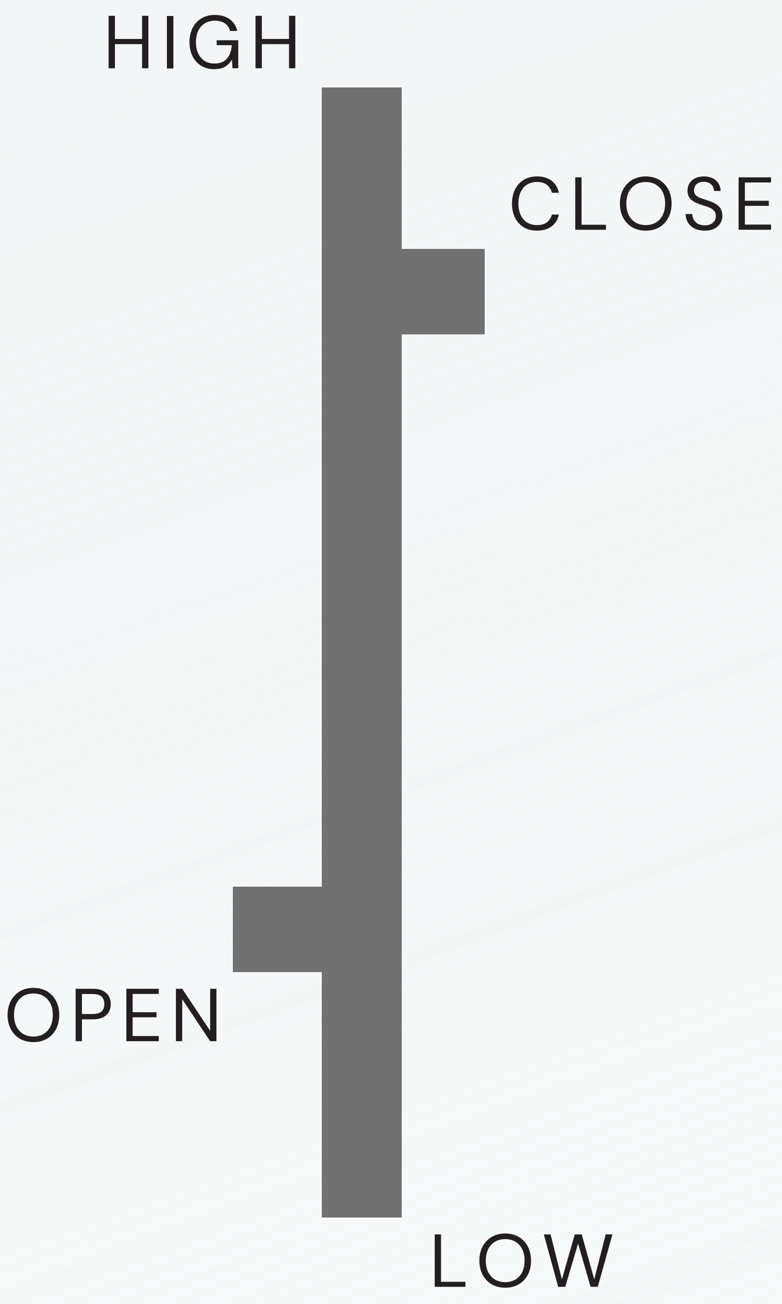
Step 3

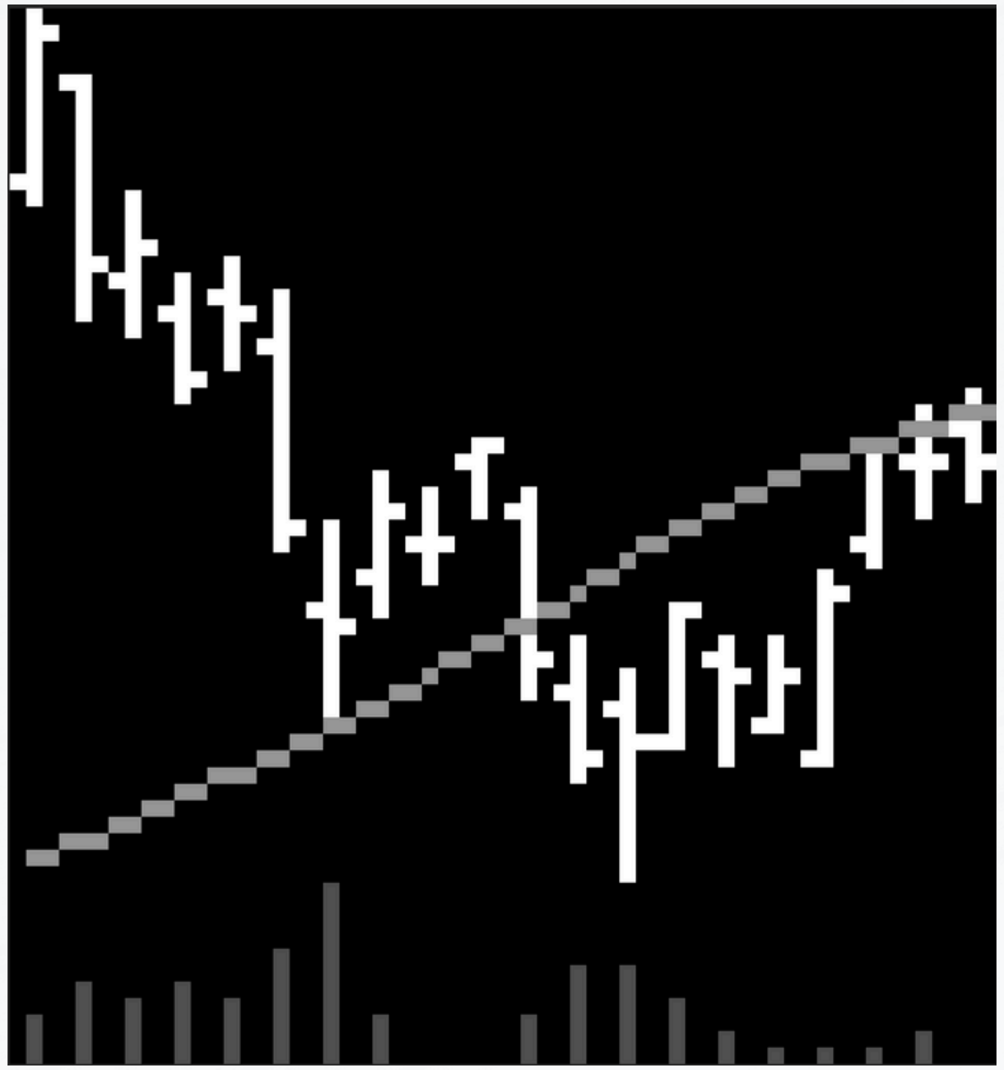
Create **portfolios** based on the CNN decisions, implementing long-short and long only strategies

THE METHODOLOGY

Step 1

Create images to represent price trends. They used **OHLC charts** with the addition of **volume** and a **moving average**





THE METHODOLOGY

Step 1

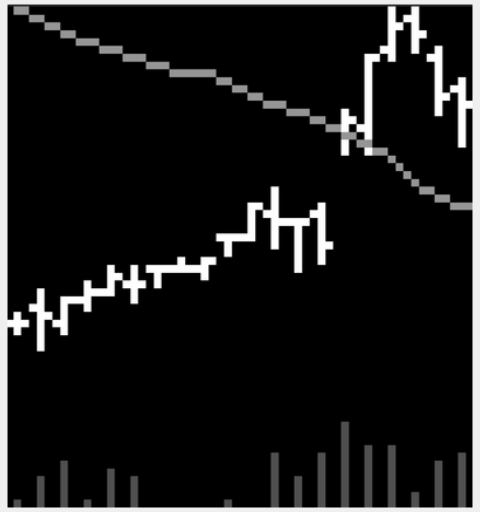
Create images to represent price trends. They used **OHLC charts** with the addition of **volume** and a **moving average**

AAPL

5-day

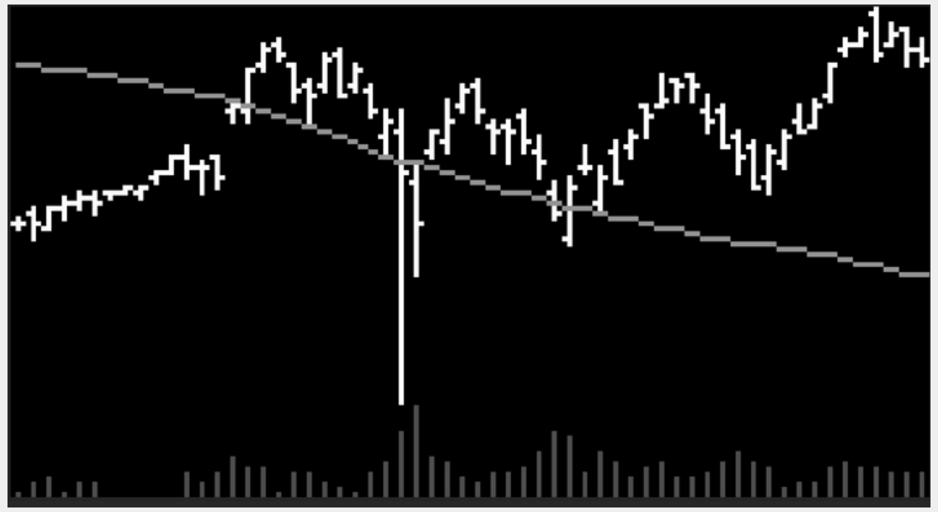
32x15

20-day



64x60

60-day



96x180

(All charts starting from March 31st 2010)

THE METHODOLOGY

Step 2

Train a **CNN** to extract features and patterns that help us predict future stock returns based on price chart images, followed by a **classification head** to estimate the probability of positive returns.

5-day

32x15

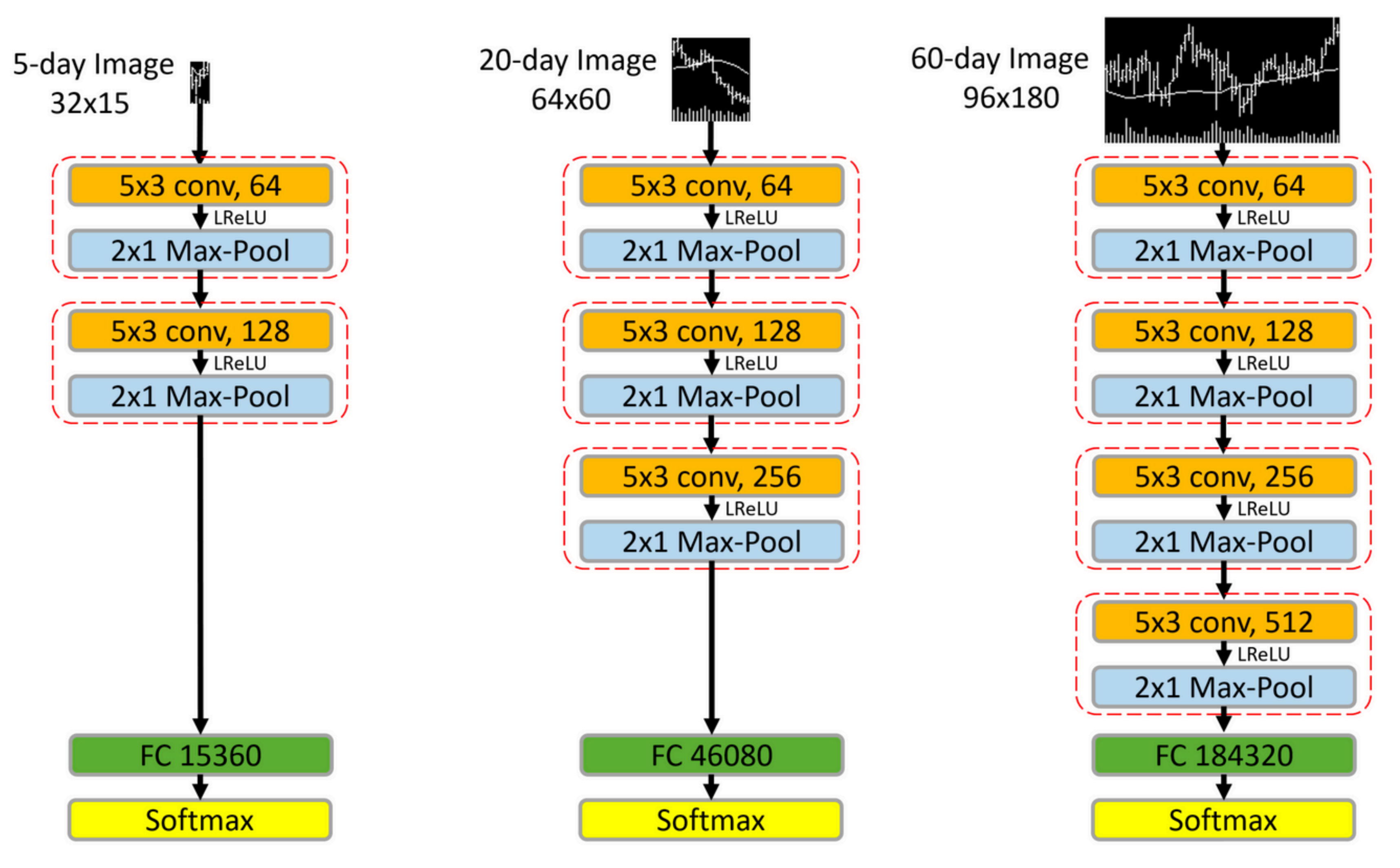
Each image is assigned a **binary label:**

- **True**, if in the next time period there is a **positive return**
- **False**, if in the next time period there is a **negative return**

THE METHODOLOGY

Step 2

Train a **CNN** to extract features and patterns that help us predict future stock returns based on price chart images, followed by a **classification head** to estimate the probability of positive returns.

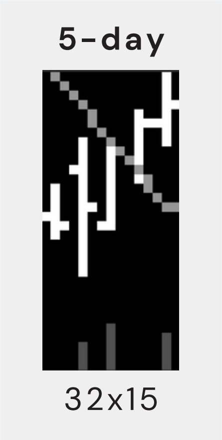


THE METHODOLOGY

Step 3

Create portfolios based on the CNN decisions, implementing long-short and long only strategies

Assuming we are using 5 day images...



t

Calculate the probability of a positive return for all stocks

t+5

t+10

THE METHODOLOGY

Step 3

Create portfolios based on the CNN decisions, implementing long-short and long only strategies

Buy or sell an equal amount of each stock

Buy the top decile and Short the bottom decile

t

Calculate the probability of a positive return for all stocks

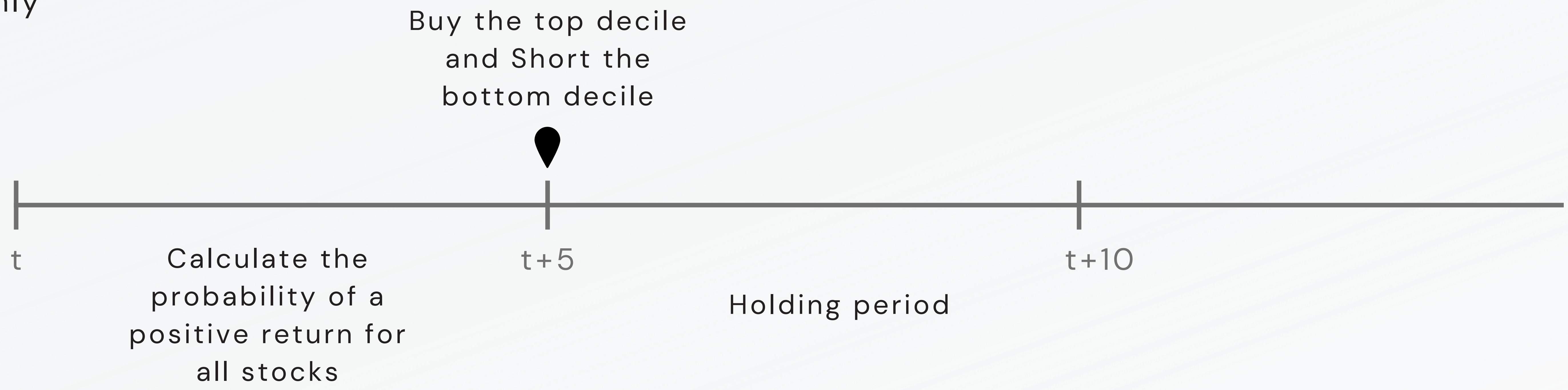
t+5

t+10

THE METHODOLOGY

Step 3

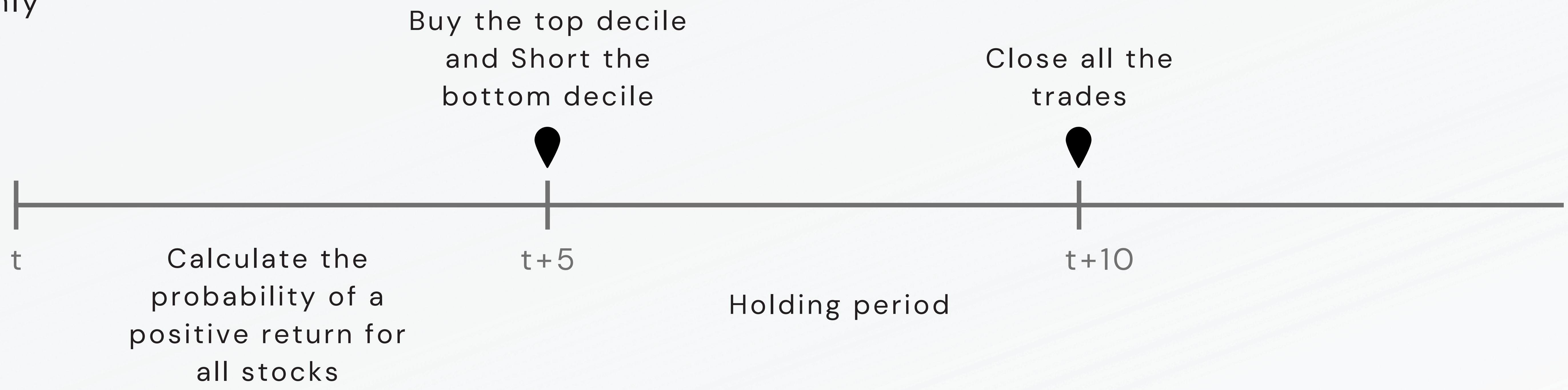
Create portfolios based on the CNN decisions, implementing long-short and long only strategies



THE METHODOLOGY

Step 3

Create portfolios based on the CNN decisions, implementing long-short and long only strategies

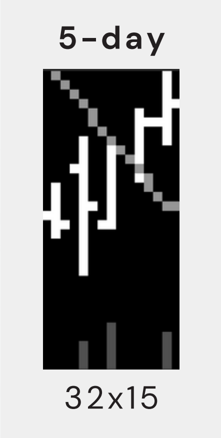


THE METHODOLOGY

Step 3

Create portfolios based on the CNN decisions, implementing long-short and long only strategies

Rinse and repeat...



Buy the top decile and Short the bottom decile

t

t+5

t+10

Calculate the probability of a positive return for all stocks

OUR RESULTS

Long-only

Model	Overall Return	Avg Yearly Return
5 day	65.16%	13.98%
20 day	34.95	8.13%
60 day	28.07%	6.67%

Long-short

Model	Overall Return	Avg Yearly Return
5 day	5.29%	1.35%
20 day	-8.50%	-2.29%
60 day	-8.17%	-2.20%

OUR RESULTS

Long-only

Model	Overall Return	Avg Yearly Return
5 day	65.16%	13.98%
20 day	34.95	8.13%
60 day	28.07%	6.67%

Long-short

Model	Overall Return	Avg Yearly Return
5 day	5.29%	1.35%
20 day	-8.50%	-2.29%
60 day	-8.17%	-2.20%

04

OUR RESULTS

Long-only

Model	Overall Return	Avg Yearly Return
5 day	65.16%	13.98%

Model	Overall Return	Avg Yearly Return
Uniform	54.71%	12.06%

Did we beat a uniform portfolio?

OUR RESULTS

Long-only

Model	Overall Return	Avg Yearly Return
5 day	65.16%	13.98%

Did we beat a random portfolio?

Model	Overall Return	Avg Yearly Return
Uniform	54.71%	12.06%
Random	53.12%	11.76%

OUR RESULTS

Long-only

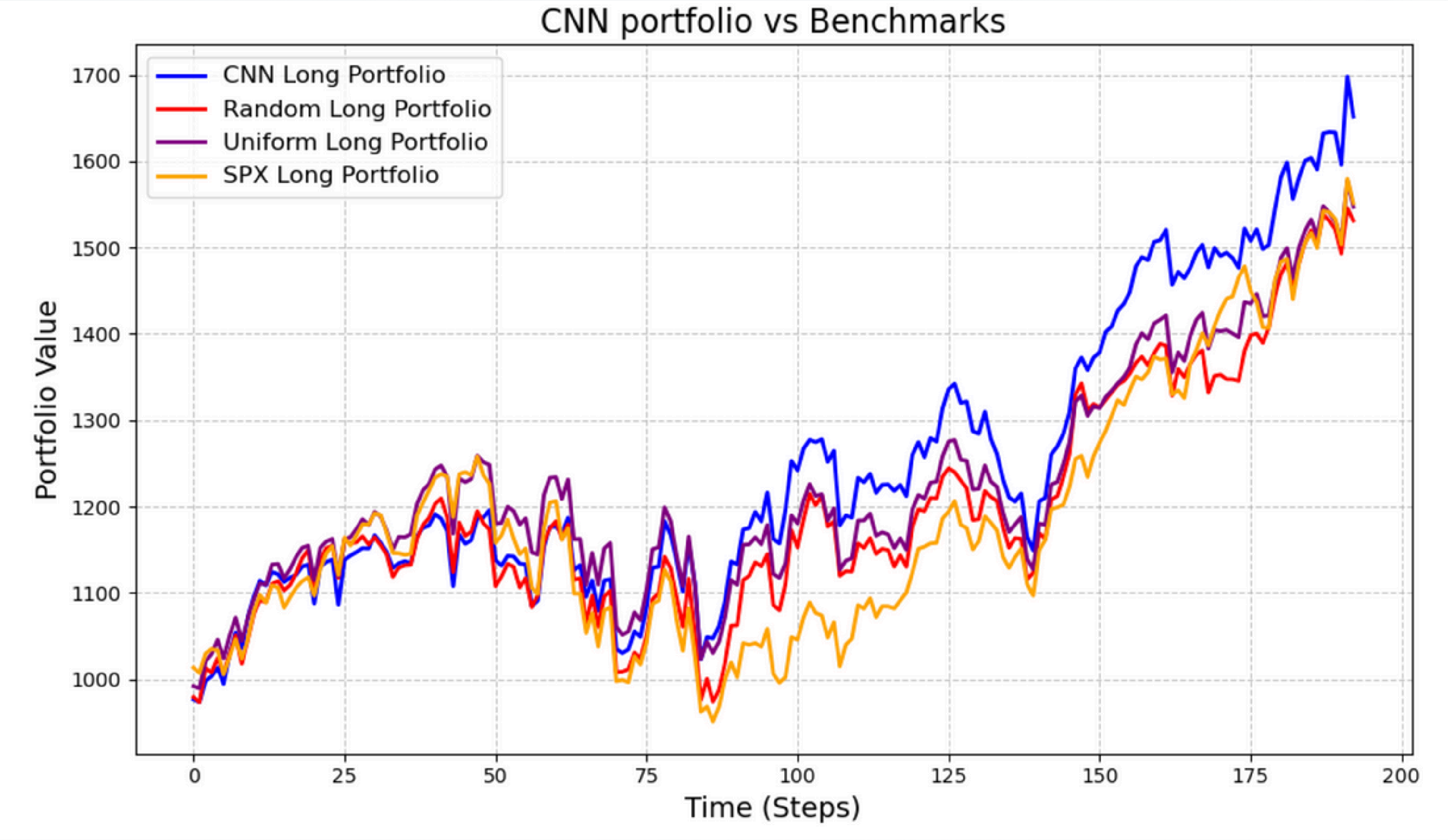
Model	Overall Return	Avg Yearly Return
5 day	65.16%	13.98%

Did we beat the S&P 500?

Model	Overall Return	Avg Yearly Return
Uniform	54.71%	12.06%
Random	53.12%	11.76%
S&P 500	55.13%	12.14%

OUR RESULTS

Long-only		
Model	Overall Return	Avg Yearly Return
5 day	65.16%	13.98%



OUR RESULTS

Long-only

Model	Overall Return	Avg Yearly Return
5 day	65.16%	13.98%

TRANSACTION COSTS

KELLY'S RESULTS

Kelly et al. explore a variety of portfolio constructions, going beyond our simple approach. They experiment with predicting different time steps using varying image inputs, such as predicting **20-day returns with 5-day** images, and construct CNN portfolios using **value-based weighting** between the stocks selected by the CNN. While there are areas of agreement between their findings and ours, significant differences also emerge, highlighting contrasting outcomes in portfolio performance and predictive patterns.

Main Common Findings

1. The CNN methodology yields promising results.
2. The best-performing strategies use 5-day images.

Interpretation:
This suggests that predictive patterns are particularly effective for **weekly trading strategies**.
Additionally, providing the model with **more information does not always improve performance,** indicating that simpler, focused inputs can often be more impactful in capturing actionable patterns.

KELLY'S RESULTS

Long-Short vs Long-only

In Kelly et al., **long-short** portfolios outperform long-only ones, while in our case, long-only portfolios excel, and long-short portfolios often underperform, sometimes yielding negative returns.
Possible explanations include:

1. Our CNN may have learned **only long patterns**, struggling with short ones.
2. The **S&P 500** dataset may lack sufficient examples to train the CNN for effective short trades.

Huge sharpe ratios

Our best performer

Long Only	Us	Kelly
I5/R5	0.7	1.8
Long Short		
I5/R5	-	7.2
I20/R5	-	6.8
I60/R5	-	4.9

I5/R5 --> uses 5 day **Image**, to predict 5 day **Return**

Insights from Kelly et al. We Didn’t Investigate

Key insight

Transfer learning extended CNN methodology beyond U.S. markets to 26 other countries with impressive performance.

Findings

- CNN trained on U.S. data outperformed country-specific CNNs in international markets.
- CNN trained on international data performed slightly better than U.S.-trained CNN when re-applied to U.S. markets.

Implications

- Patterns learned from U.S. stock data are generalizable across global markets.
- Leveraging global financial data enhances predictive accuracy, even in the originating market.

WHAT COULD WE IMPROVE AND HOW COULD WE EXPAND

Improvements

To improve our replication, we could have used the same data as Kelly et al., but due to time and computational constraints, we chose not to. Additionally, our goal was not to perfectly replicate their study, but to **add our own perspective to the research**. We could have also explored all the different portfolio combinations that Kelly et al. tested and compared the results, but we decided to focus on a more streamlined approach to see **how the methodology could perform in a simplified context.**

Expansion

We have several ideas for expanding this paper. First, we would like to **apply the methodology to different assets**, broadening the scope of the analysis.

Additionally, we try **different chart**, such as **candlestick charts with color coding,** to provide more structure.

Furthermore, we are interested in testing a **non-image-based approach** to determine whether the success of the strategy is primarily due to the use of images or the fact that images allow us to incorporate more data, beyond just price, into the model.

THANKS FOR WATCHING

GROUP 2