

CNN-Based Candlestick Patterns Recognition and Stock Trend Prediction

Team 16



Introduction

Candlestick Charts

- Widely used by stock traders & investors to gauge market sentiment
- Offering insights into market psychology and momentum



Introduction

Caveat of manual interpretation

- Time Consuming (time is money)
- Human Bias (same chart, different interpretation)

ML / AI empowering us to explore automated methods to extract meaningful signals



Stock Price Prediction

- **Fundamental** Analysis vs. **Technical** Analysis
- Stock trend prediction is important on two levels:
 - a. for individual market participants: maximizing profits or minimizing losses
 - b. For the market as a whole: drive price movements towards fair value by triggering buy or sell activities



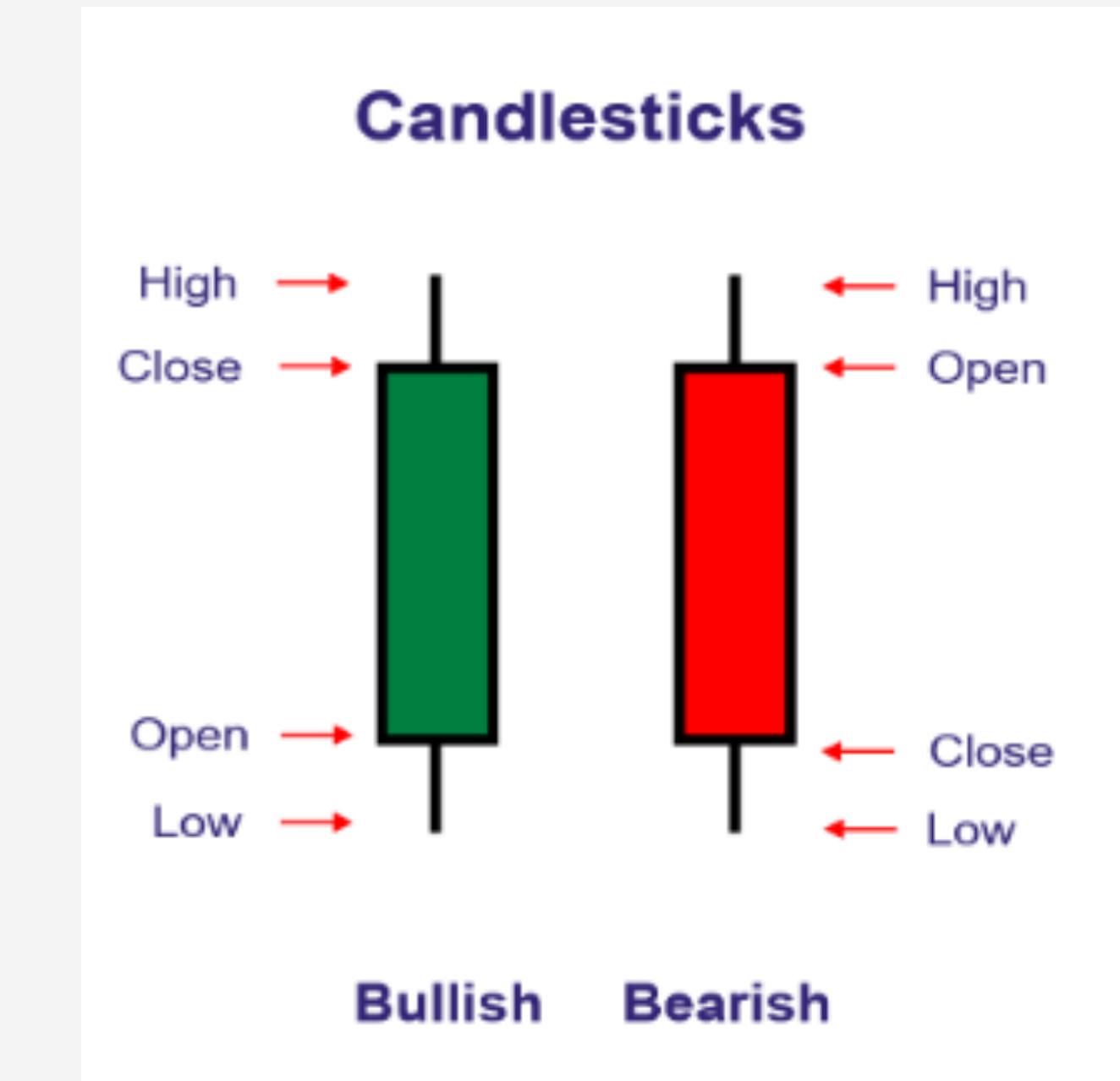
Technical Analysis

- technical analysis charts are applied to discover potential trading opportunities
- technical patterns are developed from historical price charts, enabling investors to make predictions of future price movements and in turn make trading profits

Candlestick Charts

4 Critical Points -

1. Opening Price
2. Closing Price
3. Highest Price (high) within a specific time frame
4. Lowest Price (low) within a specific time frame



Source: <https://www.strike.money/technical-analysis/types-of-candlesticks-patterns>

Versatile Patterns of Candlestick Charts

Candlestick Patterns Cheat Sheet for Cryptomarkets

Bullish			Bearish			Neutral
Reversal		Continuation	Reversal		Continuation	
Hammer	Inverted Hammer	Bullish Three Line Strike	Hanging Man	Shooting Star	Bearish Three Line Strike	Doji
Bullish Engulfing	Tweezer Bottom	Rising Three Methods	Bearish Engulfing	Tweezer Top	Falling Three Methods	Gravestone Doji
Morning Star	Three Stars in the South	Bullish Mat Hold	Evening Star	Advance Block	Bearish Mat Hold	Dragonfly

Benefits of Applying CNN to Stock Trend Analysis

1. CNNs can
 - a. extract meaningful features from sequential data
 - b. learn hierarchical representation
2. CNN can learn to detect patterns at different levels
(Low-shapes, Medium-combinations, High-other indicators)
3. CNNs are highly parameter efficient

What is Our Data?

- Historical stock price and trading volume data of companies listed on S&P 500 index from 2010-01-01 to 2019-12-31 (from yahoo finance)
- Converted to Candlestick Charts using Matplotlib



Source: <https://finance.yahoo.com/quote/%5EGSPC/>

	Date	Open	High	Low	Close	Volume	ticker
0	2010-01-04 00:00:00-05:00	40.869485	41.046558	40.662898	40.835056	3640265	MMM
1	2010-01-05 00:00:00-05:00	40.726839	40.938345	40.185783	40.579281	3405012	MMM
2	2010-01-06 00:00:00-05:00	41.258047	41.612192	41.076055	41.154751	6301126	MMM
3	2010-01-07 00:00:00-05:00	40.982593	41.199018	40.392351	41.184261	5346240	MMM
4	2010-01-08 00:00:00-05:00	41.164598	41.474476	40.972768	41.474476	4073337	MMM

Supplementary Data:

- Sector in which the company operates.
- Date when the company went public and added to the S&P 500 index.

	Date	Open	High	Low	Close	Volume	ticker
0	2010-01-04 00:00:00-05:00	40.869485	41.046558	40.662898	40.835056	3640265	MMM
1	2010-01-05 00:00:00-05:00	40.726839	40.938345	40.185783	40.579281	3405012	MMM
2	2010-01-06 00:00:00-05:00	41.258047	41.612192	41.076055	41.154751	6301126	MMM
3	2010-01-07 00:00:00-05:00	40.982593	41.199018	40.392351	41.184261	5346240	MMM
4	2010-01-08 00:00:00-05:00	41.164598	41.474476	40.972768	41.474476	4073337	MMM

- Dimension of the dataframe: **1,178,469 rows × 7 columns**
- 500 companies (tickers) * 10 years * 250 trading days each year ~ **1,250,000 rows**
- Reasons for missing data:
 - Some companies were not public until later in the 2010s.
 - Some companies quit the index in the middle of the observation period.
 - Other data recording issues.

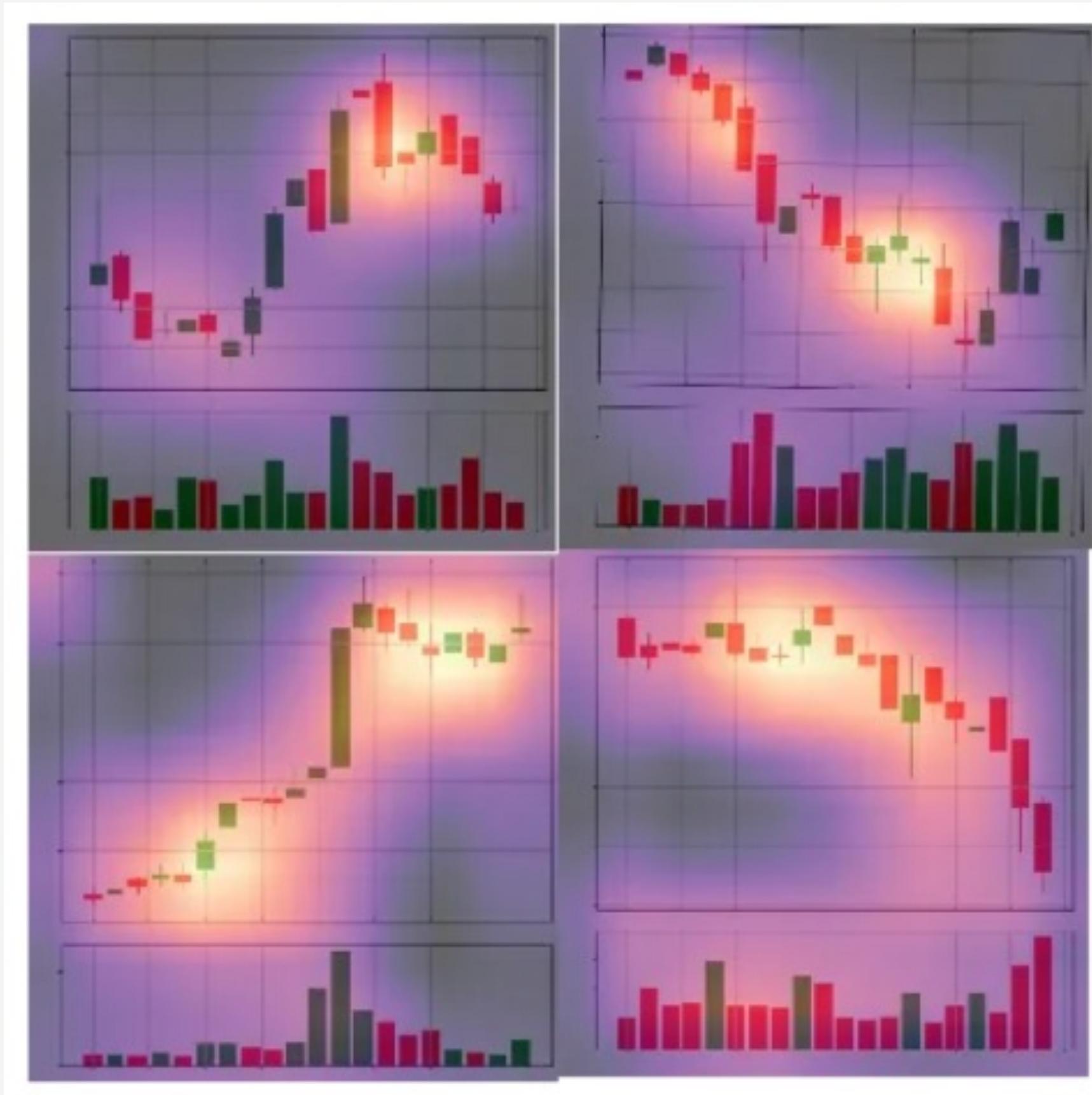
- 20-day window
- Target picture resolution: $448 * 336$ (4:3) with RGB
- Target picture size: < 50KB



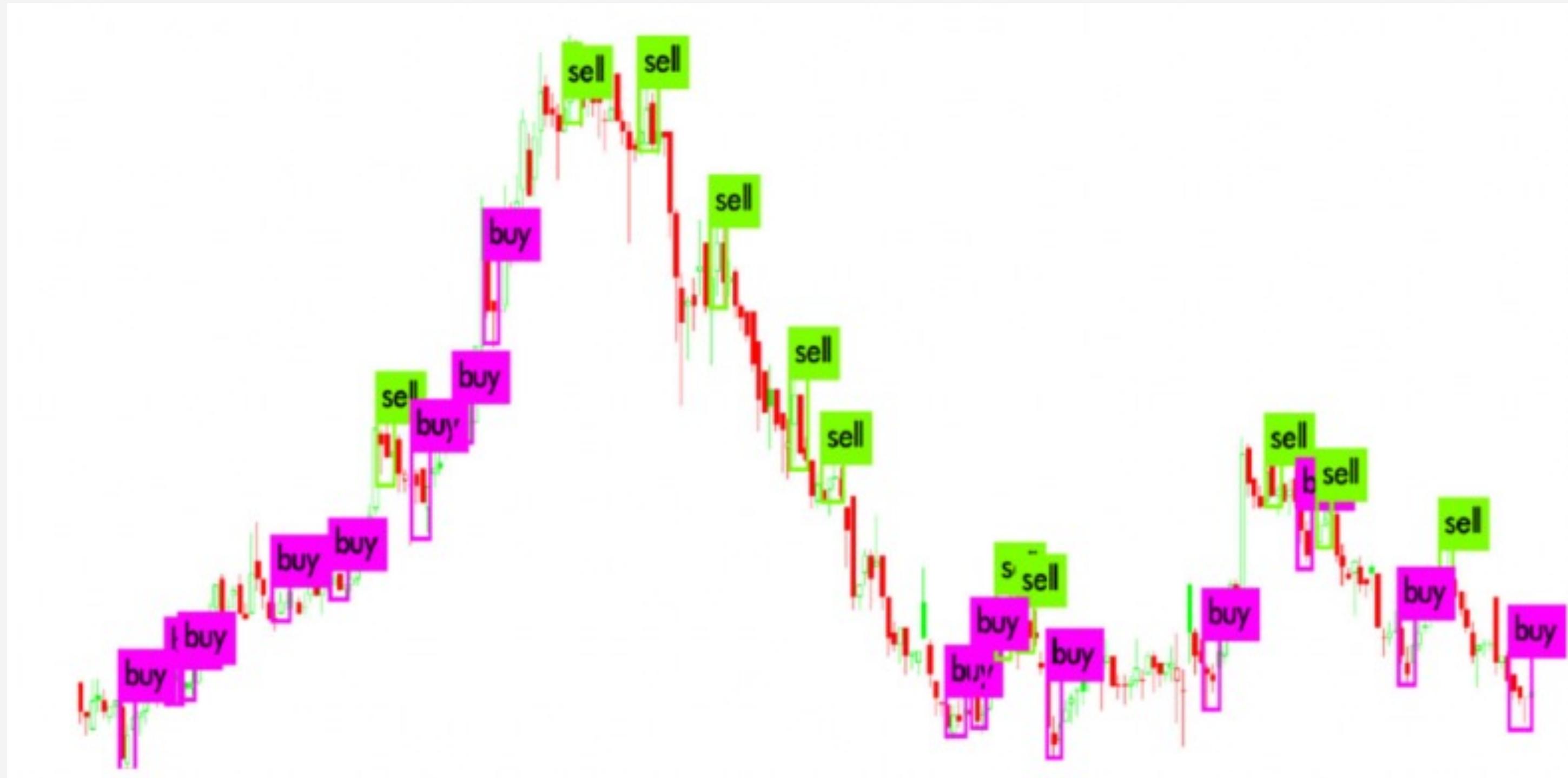
- If 5-day close price > 5-day open * 1.01: **Buy** (0)
- If 5-day close price < 5-day open * 0.99: **Sell** (1)
- Else: **Hold** (2)
- The resulting proportion of 3 target labels is approximately balanced ($\frac{1}{3}$ each)
- The annualized return of 1.01 is 65%, which is a reasonable return that affects the investor's buy-sell-hold decision.



Extracted Patterns (EDA)

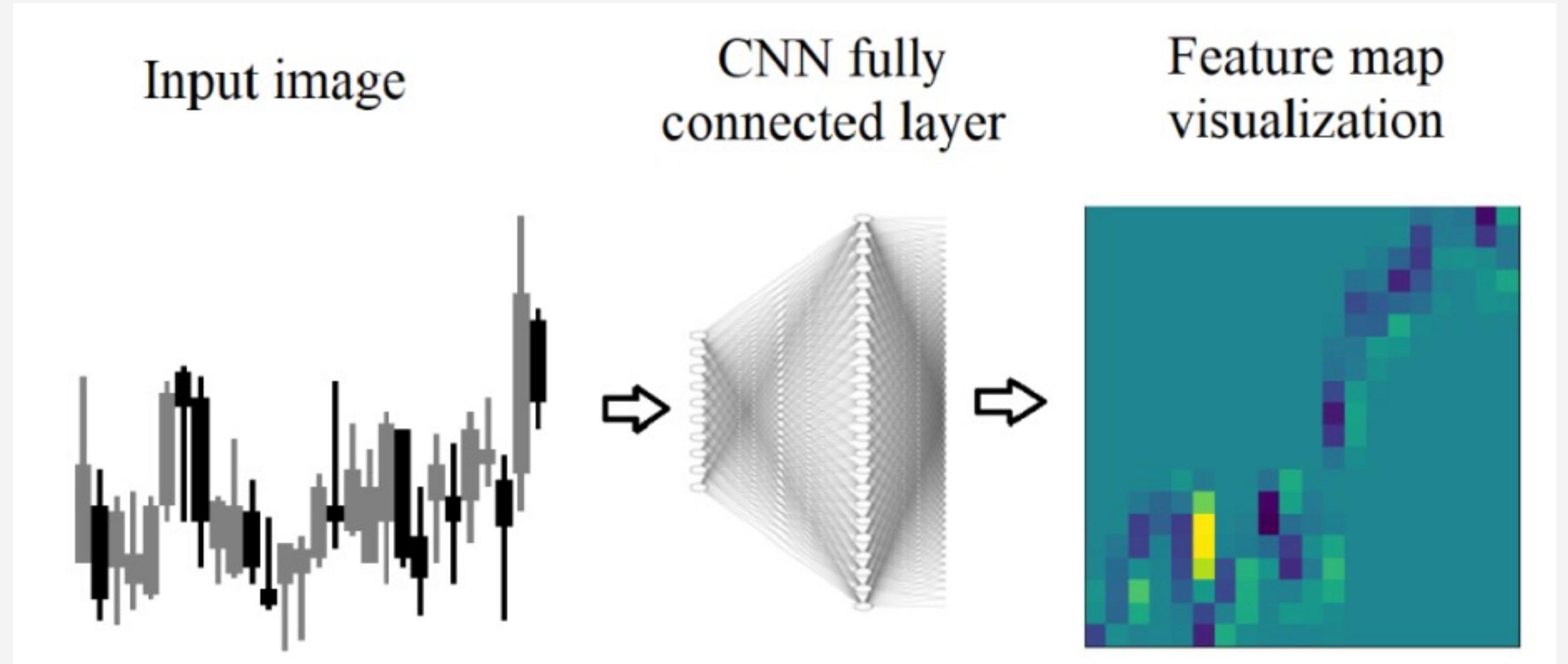


Extracted Patterns (EDA)



Source: [ps://towardsdatascience.com/identifying-candlestick-patterns-using-deep-learning-b7d706726874](https://towardsdatascience.com/identifying-candlestick-patterns-using-deep-learning-b7d706726874)

Extracted Patterns (EDA)



Source: https://www.researchgate.net/publication/326764676_Stock_Chart_Pattern_recognition_with_Deep_Learning

Data Loader

- Resize the shape of the photo to 224 * 224
- Scale the color value by 255
- Batch size: 32
- Handling unreadable images



0_A46.png



0_A446.png



0_A571.png



0_A1346.png



0_A1396.png



0_A1421.png



0_A1846.png



0_A1871.png



0_A1971.png

Challenges of the Experiments

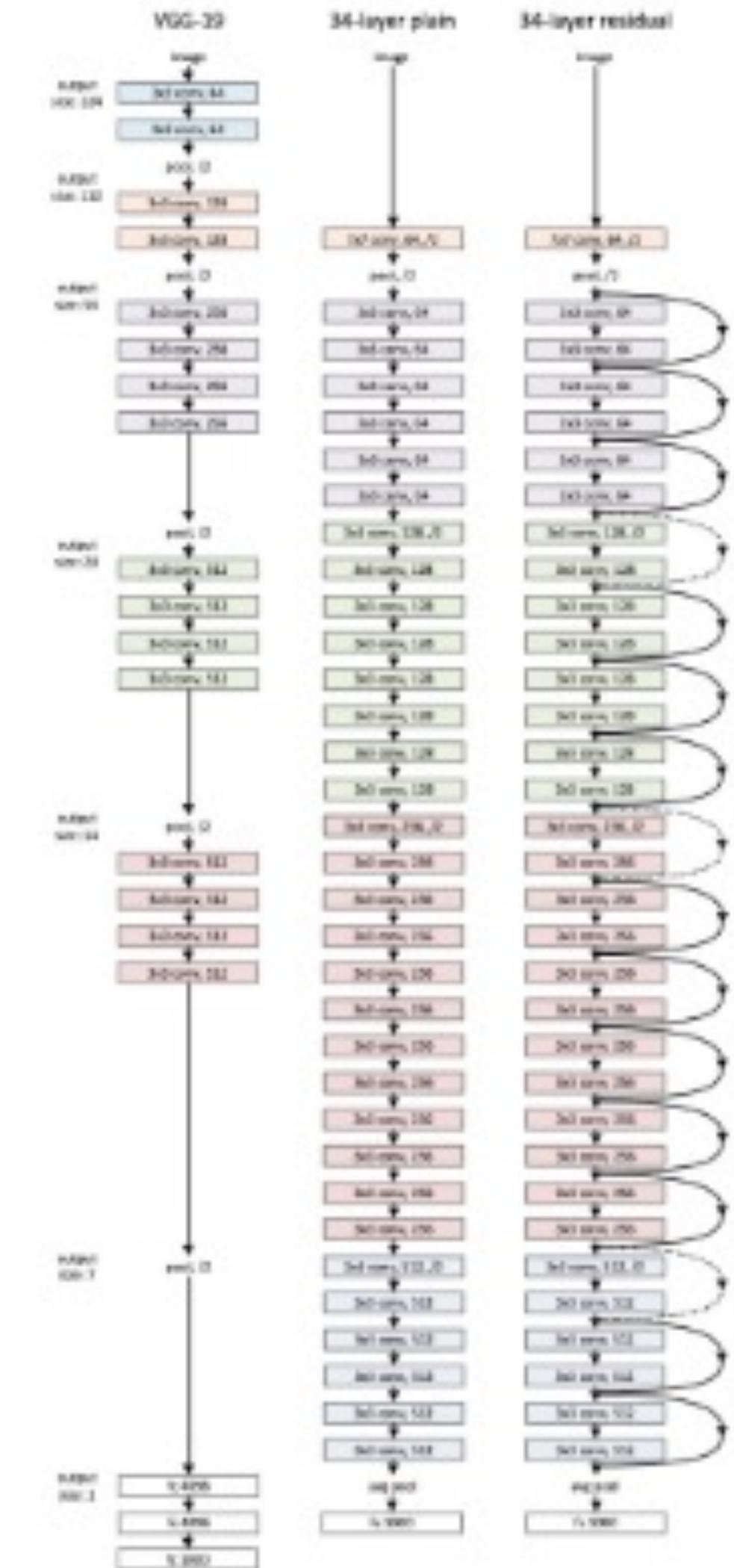
- Data size & sample challenge
- Computational power challenge



Model 1 - ResNet

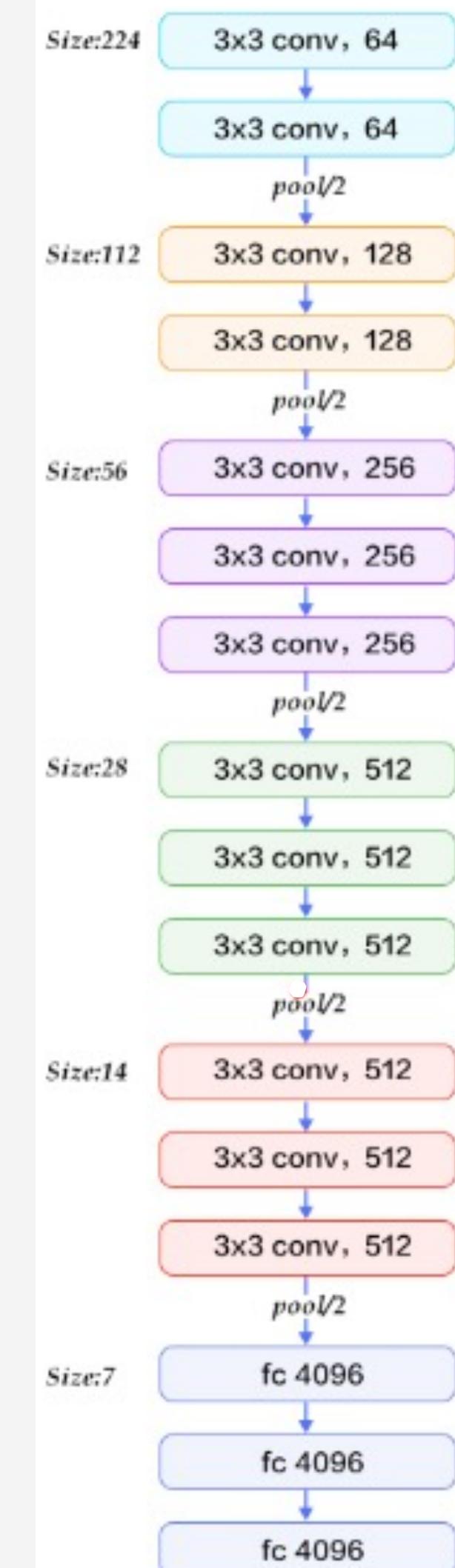
Pre-Trained Model

- Base model : ResNet50, a pre-trained model
- Layers : Dense layer , global average 2D layer , softmax layer
- Freezing layer
- Compilation
- Training



Model 2 - VGG16

- Base model : VGG16 pretrained model
- Custom Layers : same as the ResNet model
- Freezing Layers : VGG16
- Compilation & Training the same way as the ResNet model



Model 3 - CNN Model

Input
Conv2D-32 ReLU
max-pooling
Conv2D-48 ReLU
max-pooling
Dropout
Conv2D-64 ReLU
max-pooling
Conv2D-96 ReLU
max-pooling
Dropout
Flatten
Dense-256
Dropout
Dense-2

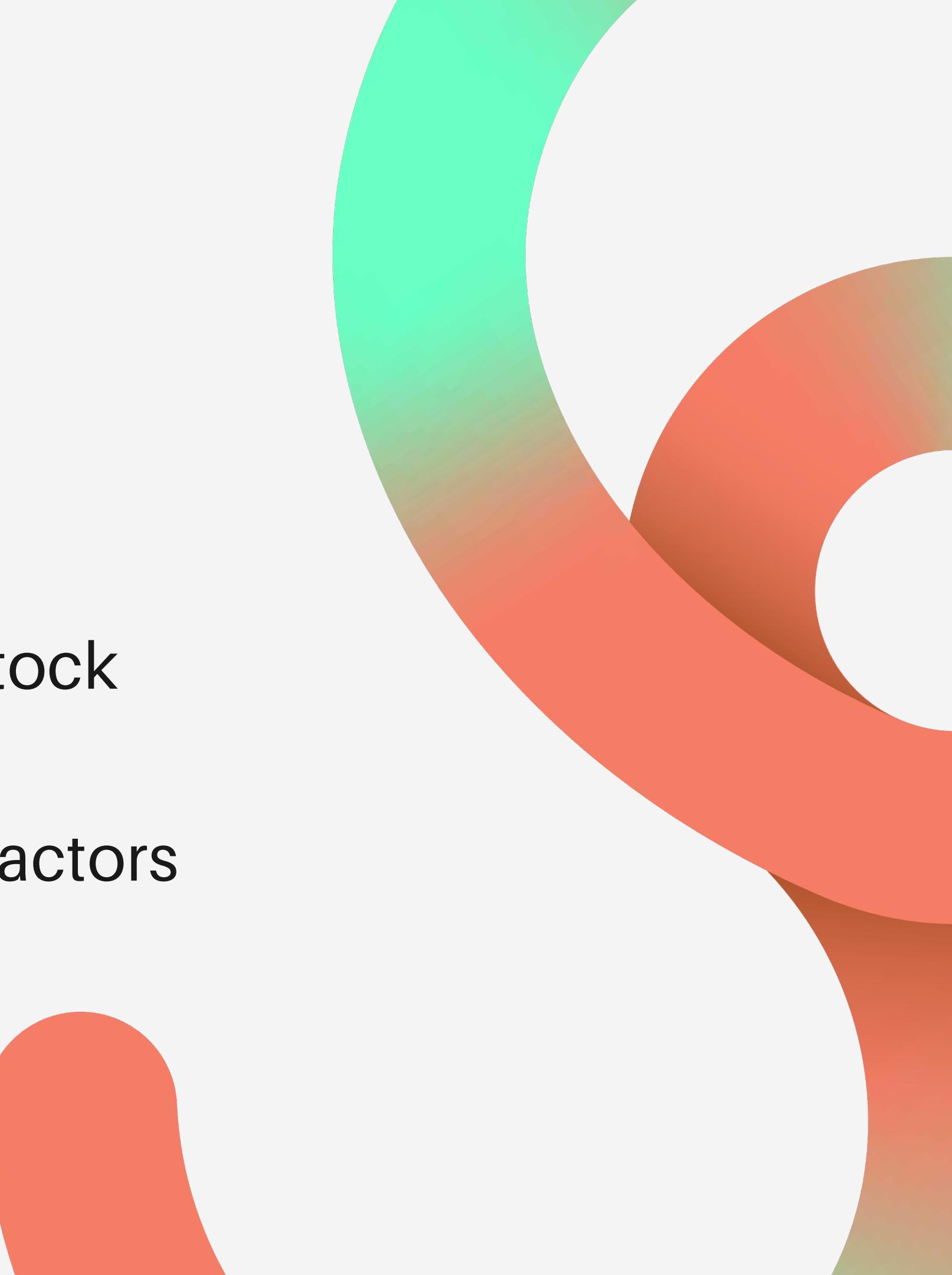


Result

Model	Class	Precision	Recall	F1-Score
Customized CNN	Buy (0)	0.73	0.65	0.69
	Sell (1)	0.41	0.44	0.42
	Hold (2)	0.47	0.45	0.46
VGG 16	Buy (0)	0.68	0.73	0.7
	Sell (1)	0.5	0.43	0.46
	Hold (2)	0.56	0.61	0.58
ResNet 50	Buy (0)	0.79	0.75	0.77
	Sell (1)	0.63	0.7	0.66
	Hold (2)	0.58	0.55	0.56

Limitations

- 1 Data's accessibility
- 2 Computation power
- 3 Application to different sectors / stock markets
- 4 Stock price's randomness / macro factors
- 5 Efficient Market Hypothesis



Future Scope

- 1 Algorithm Trading
- 2 Risk Management
- 3 Sentiment Analysis Integration
- 4 Financial Advisory Services



Reference

- [1]<https://towardsdatascience.com/identifying-candlestick-patterns-using-deep-learning-b7d706726874>
- [2]<https://www.kaggle.com/code/farshidhossain68/candle-stick-full-project-accuracy-0-9542#Image-Load>
- [3]https://www.researchgate.net/publication/326764676_Stock_Chart_Pattern_recognition_with_Deep_Learning
- [4]<https://arxiv.org/pdf/1903.12258.pdf>
- [5]<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9092995>
- [6]<https://cs231n.stanford.edu/reports/2022/pdfs/157.pdf>
- [7]https://www.researchgate.net/publication/376685209_Enhancing_Financial_Chart_Analysis_Advanced_Detection_of_Candlestick_Patterns_Using_Deep_Learning_Models_for_Mastering_Trend_Recognition
- [8]<https://www.strike.money/technical-analysis/types-of-candlesticks-patterns>



Q&A