Fast Candlestick Patterns Detection with Limited Training Samples Using RGB Gramian Angular Field and YOLO-LITE-V1

Background and Problems

- Candlestick chart is a visualizable tokenization of price time series.
- Potential in high-frequency trading, as automation is very important.
- Many traders come up with new candlestick pattern indicators and need automatic pipeline to test their performance.
- Insufficient data
- Tidous labeling
- Numeric value based rather than graph based

Datasets

- Input data:
  1. Time series price data
  2. Candlestick pattern label
  3. Real candlestick chart (background changed)
- Datasets include currency and stock patterns, 15 to 50k data samples in total
- Start_t, End_t, Label
- Yahoo Finance and Td Ameritrade

Problem Statement

- Goal:
  A pipeline that can train on a small training set and yield an object detection model to detect candlestick patterns graphically
- Input data:
  1. Time series price data
  2. Candlestick pattern label
  3. Pictures of real candlestick charts (background changed)
- Core Model
  YOLO-LITE for object detection
- Simple CNN with dropout/peeling for data augmentation
- Metrics
  mAP for object detection
  ROC_AUC for simple CNN

Method

- Time Series -> polar coordinate -> Gramian Angular Field
- Original time series
- Polar mapping
- Gramian Angular Field
- YOLO-LITE -- Augmented timeseries with labels = SimpleCNN
- Chart's GAF without RGB
- Model Structure
- mAP > 0.4 = 0.481
- ROC_AUC = 0.71

Conclusions

While there still exist cases that YOLO-LITE model labels far-off pictures for reasons that I cannot understand, the overall performance of both bounding box precision and classification precision is already reasonably good with a mAP > 0.4 at 0.481.

Part of the error is due to imperfection in manual labelled test sets. My design of RGB Gramian Angular Field, data ARIMA-CNN-based data augmentation, and model choice of YOLO-LITE contributes to the performance to a great extent.

If more time, computational resource, and people are available in future, we will increase the number of precisions of testing labels and explore more completed setup for object detection model.