Network Design and Some Thoughts About Trading by DNN

James Yijie Lu

jameslu.net (james.lu@magnumwm.com)

HKUST & AQUMON (Magnum Research Ltd.)

04/11/2021

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Outline of This Talk

Outline

- **Something about me**
  - I am a PhD in computer science, not from financial math background XD

- **Why deep learning (DL) is overpowered in quite a few areas?**
  - How we solve a problem in the pre-DL age?
  - How is the same kind of problem solved by DL?
    - *What has been changed?*
  - What drives the evolution of network design?
    - *A door has been opened for more sophisticated problems*
    - *Solving fundamental obstacles in deep networks*

Outline

- **How about for trading?**
  - Godlike? —*in terms of model performance*
  - Doubts?
    - *Super hungry about data*
    - *High cost*
      - Training a good model takes much more time and experience
      - Needs many GPUs and tons of electricity
    - *Hmmm… something is still missing?*

Outline

- **Action! rather than dreaming**
  - What we do in AQUMON
  - How AQUMON can help achieve your goal?

Something About Me

- **Multimedia research (2010 – 2017)**
  - Video content search and analysis
    - *Concept recognition*
    - *Event detection*
    - *Video summarization/presentation/search*
  - One year of internship in MSRA (’12 – ’13)
    - *Recommender system by click-through data*
  - A few Rank 1st in NIST TRECVID workshop/competition (’14 – ’17)
    - *Zero-example multimedia event detection*
    - *Ad-hoc video search*

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Something About Me

jameslu.net

AQUMON Investment Management

Something About Me

- **Attracted by trading (2018—)**
  - Happy to choose the “hard” mode
  - Challenging but rewarding

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

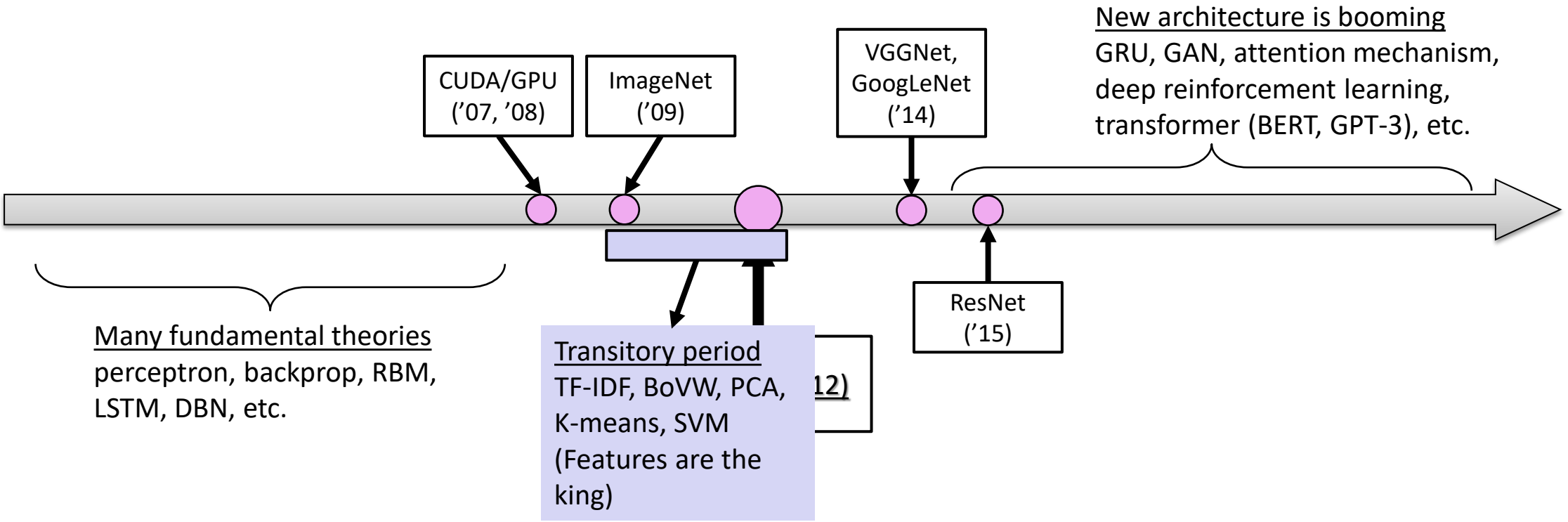
AQUMON | Investment Management

Why Deep Learning is Overpowered?

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON | Investment Management

Pre-DL & DL age



* To know more, you can go to this fancy page:
https://machinelearningknowledge.ai/brief-history-of-deep-learning/

Fathers of DL Revolution

2018 ACM A.M. TURING AWARD

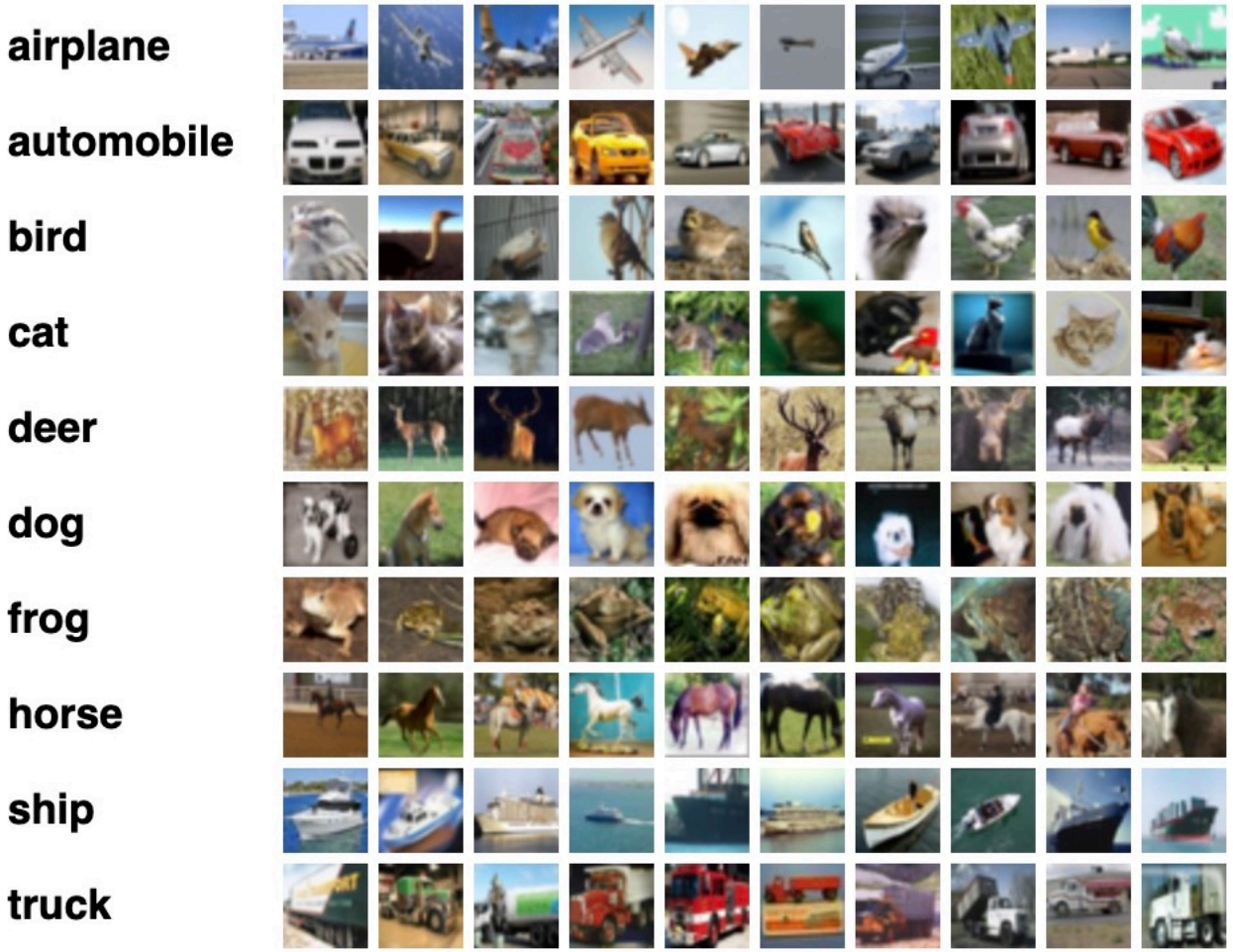
Honorable mention?

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

A problem-solution in pre-DL age

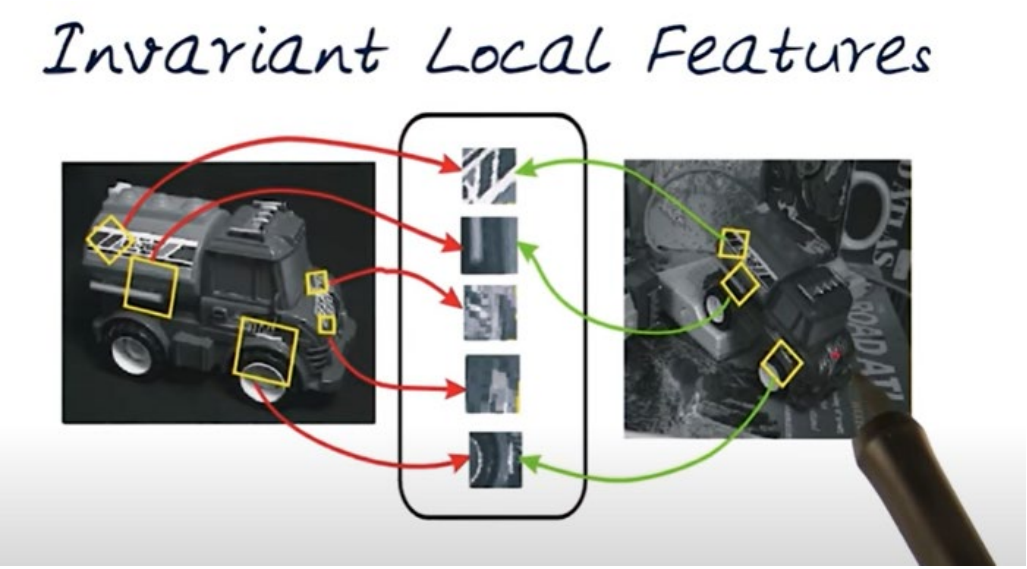
- **An example problem – image classification**

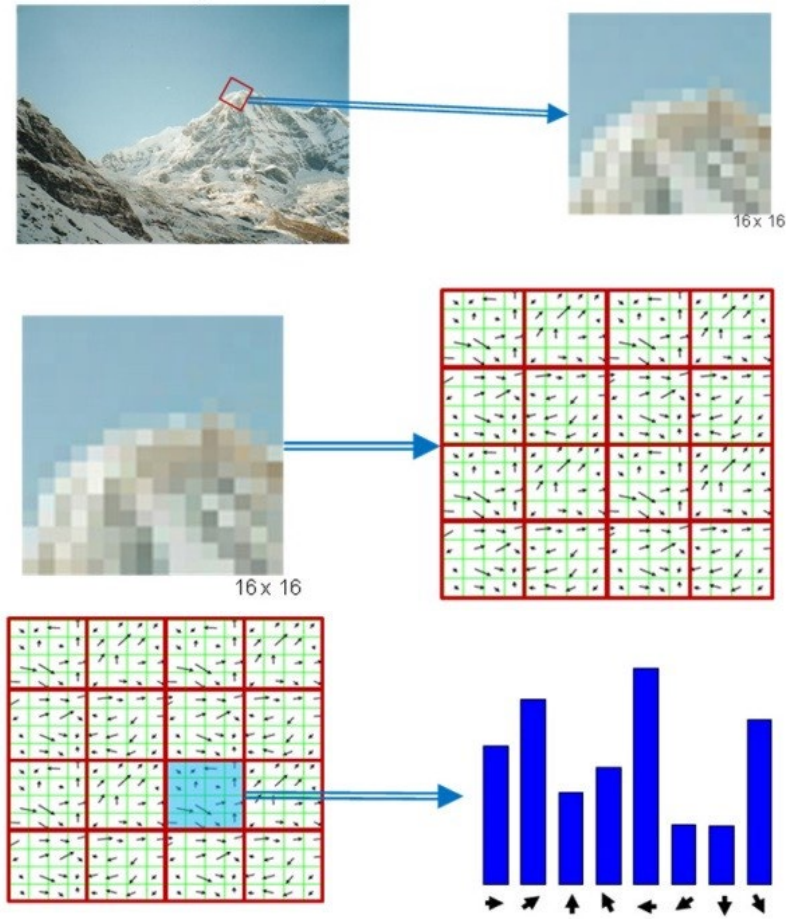


How we approach to solve this in the good old days?

A problem-solution in pre-DL age

- **Human-crafted low-level features** Extract SIFT features





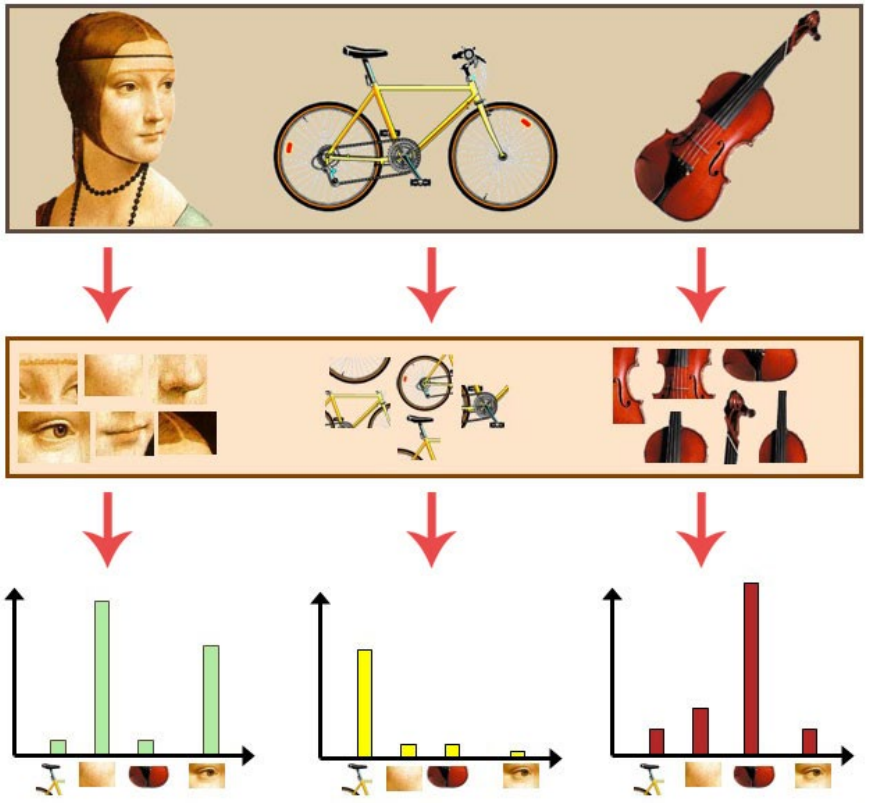
香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

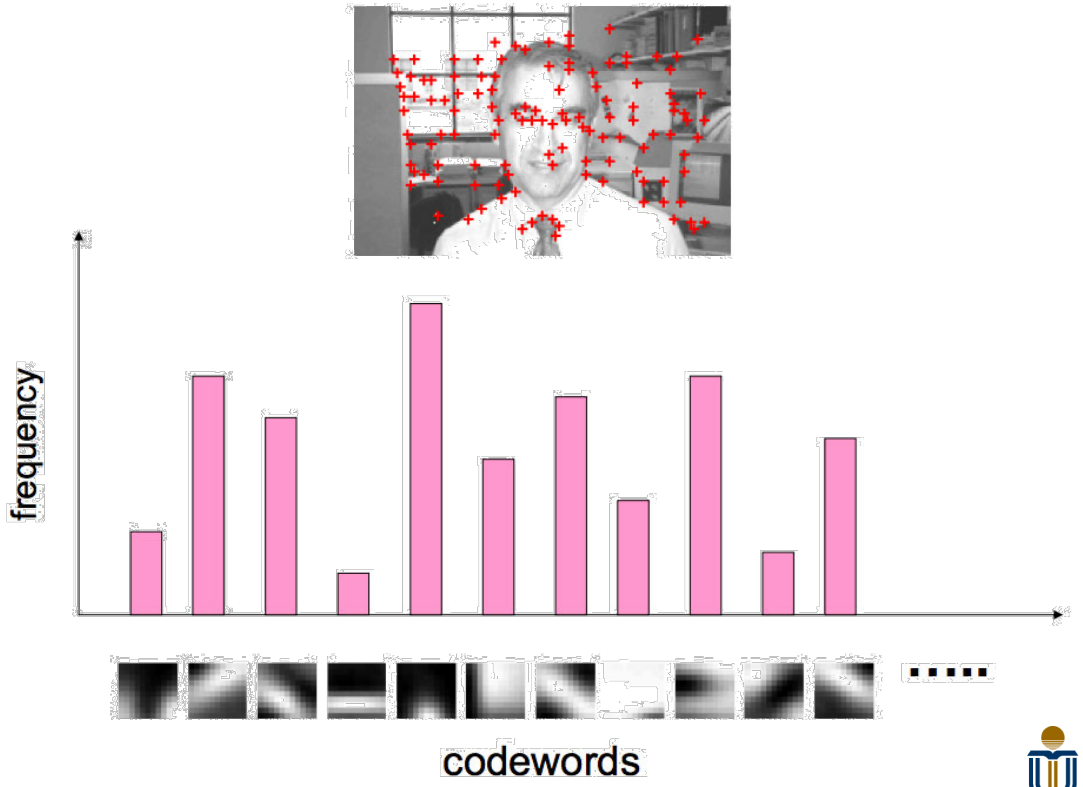
A problem-solution in pre-DL age

- **Bag-of-Visual-Word (a.k.a. Bag-of-Features) pipeline**

What we conceptually want:



The real implementation looks like:



A problem-solution in pre-DL age

- **Attribute models (2009)**

'is 3D Boxy'
'is Vert Cylinder'
'has Window'
'has Row Wind'
X'has Headlight'

'has Hand'
'has Arm'
X'has Screen'
'has Plastic'
'is Shiny'

'has Head'
'has Hair'
'has Face'
X'hasSaddle'
'has Skin'

'has Head'
'has Torso'
'has Arm'
'has Leg'
X'has Wood'

'has Head'
'has Ear'
'has Snout'
'has Nose'
'has Mouth'

'has Head'
'has Ear'
'has Snout'
'has Mouth'
'has Leg'

X'has Furniture Back'
X'has Horn'
X'has Screen'
'has Plastic'
'is Shiny'

' is 3D Boxy'
'has Wheel'
'has Window
'is Round'
' 'has Torso'

'has Tail'
'has Snout'
'has Leg'
X'has Text'
X'has Plastic'

'has Head'
'has Ear'
'has Snout'
'has Leg'
'has Cloth'

'is Horizontal Cylinder'
X'has Beak'
X'has Wing'
X'has Side mirror'
'has Metal'

'has Head'
'has Snout'
'has Horn'
'has Torso'
X'has Arm'

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

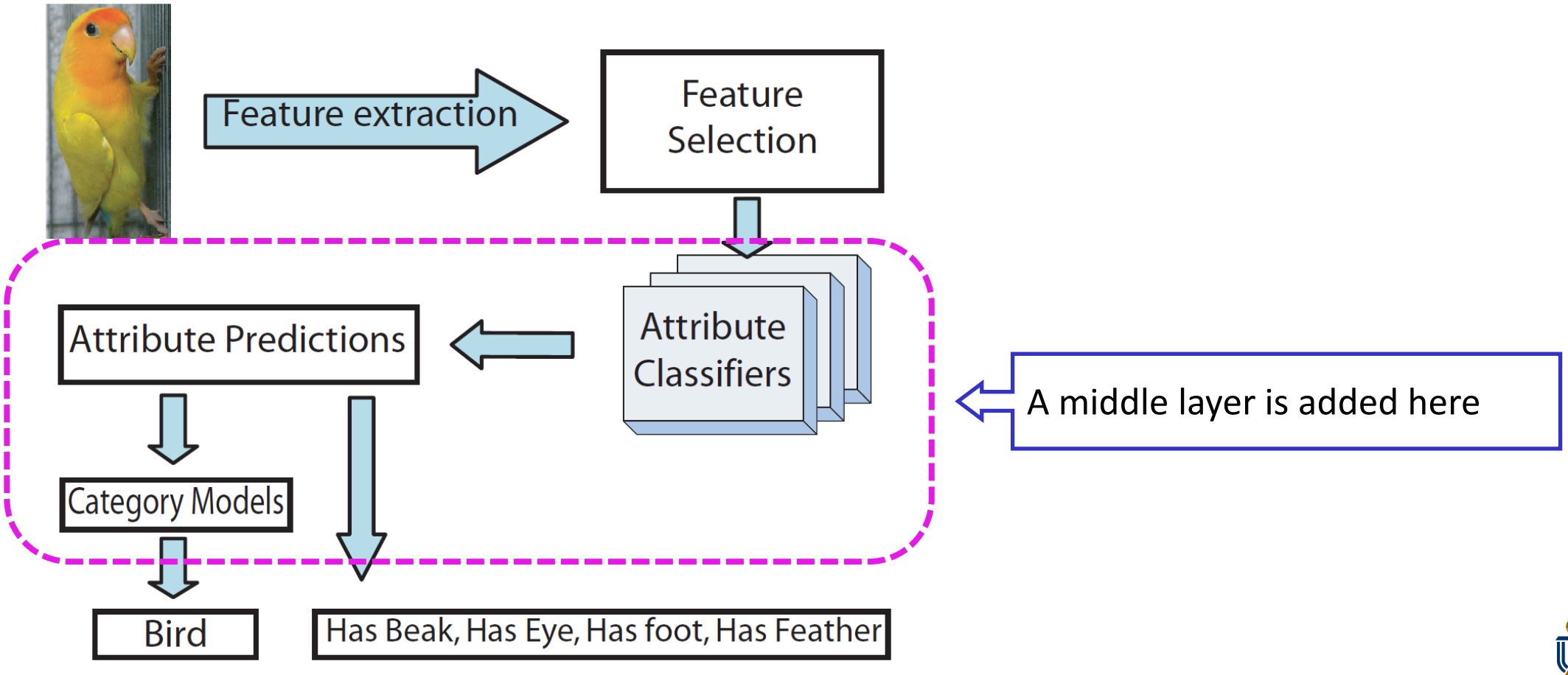
Semantic Concepts

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

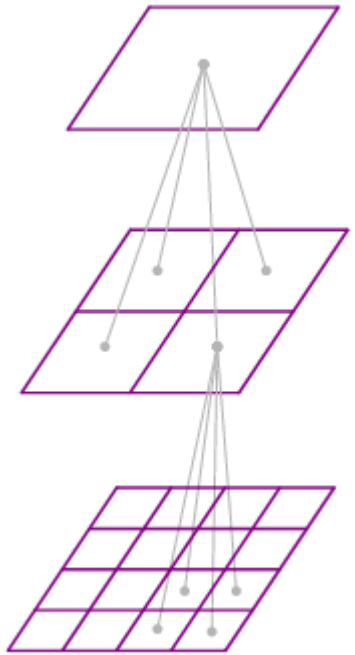
A problem-solution in pre-DL age

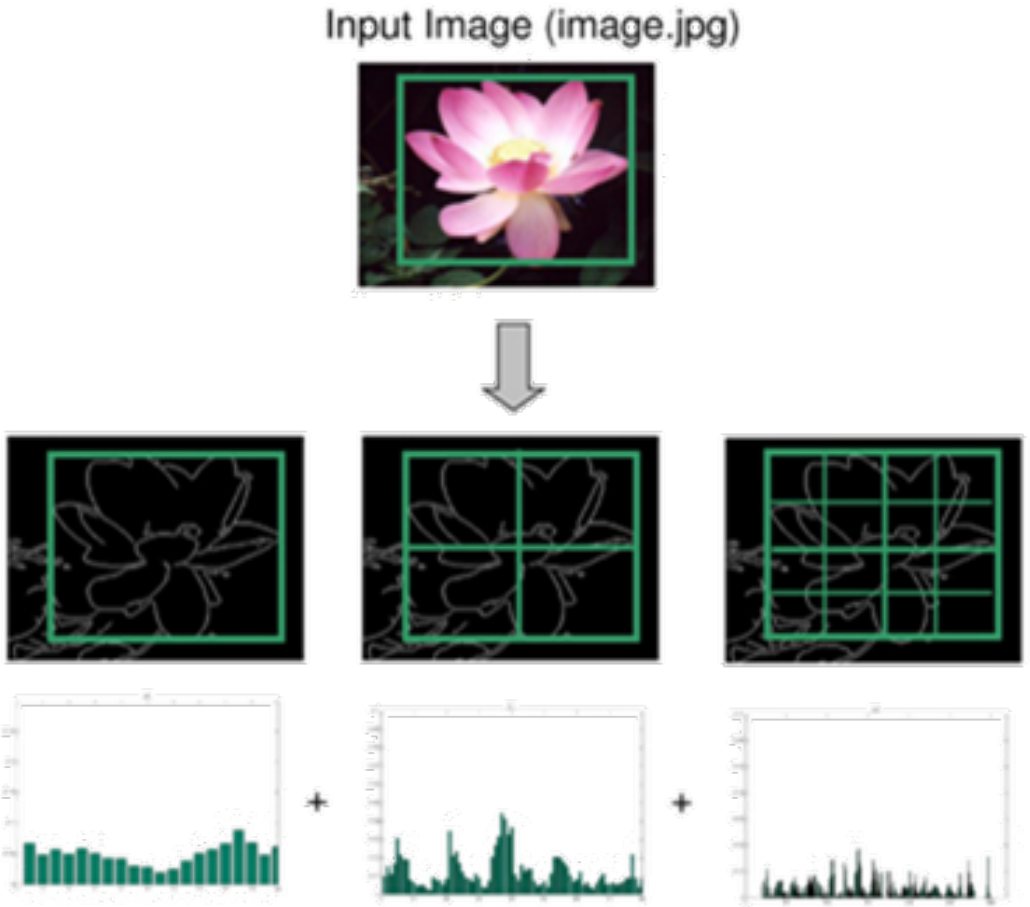
- **Attribute models (2009)**



A problem-solution in pre-DL age

- **Spatial pyramid**





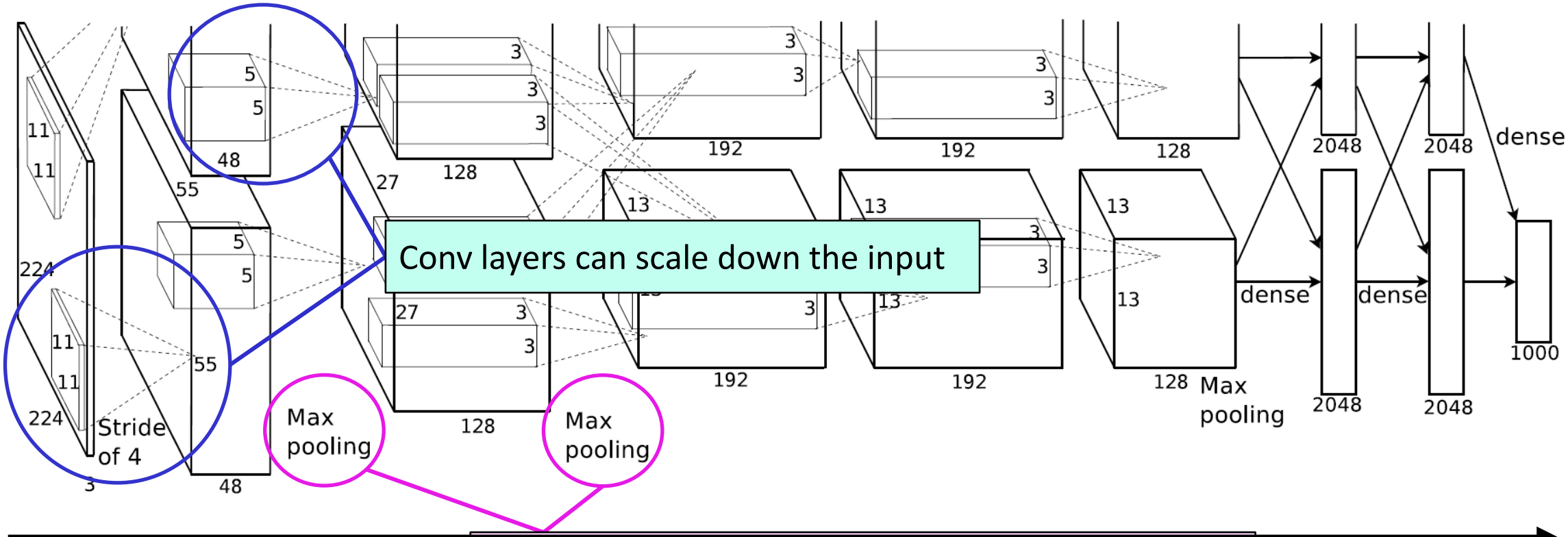
The Struggles

- **Difficult to scale up**
  - High dimensional features + millions of samples
    - *ImageNet ILSVRC Challenge (since 2010)*
      - 1000 object categories, 1.2 mil training images

    → *linear SVM or other linear models*
  - Computational complexity (spatial pyramid is luxury)
  - Nobody thinks this pipeline is ever suitable for live prediction by a PC/mobile
- **Supervised learning**
  - Too much efforts in human labeling

What has been changed by DL?

- **Recall AlexNet**

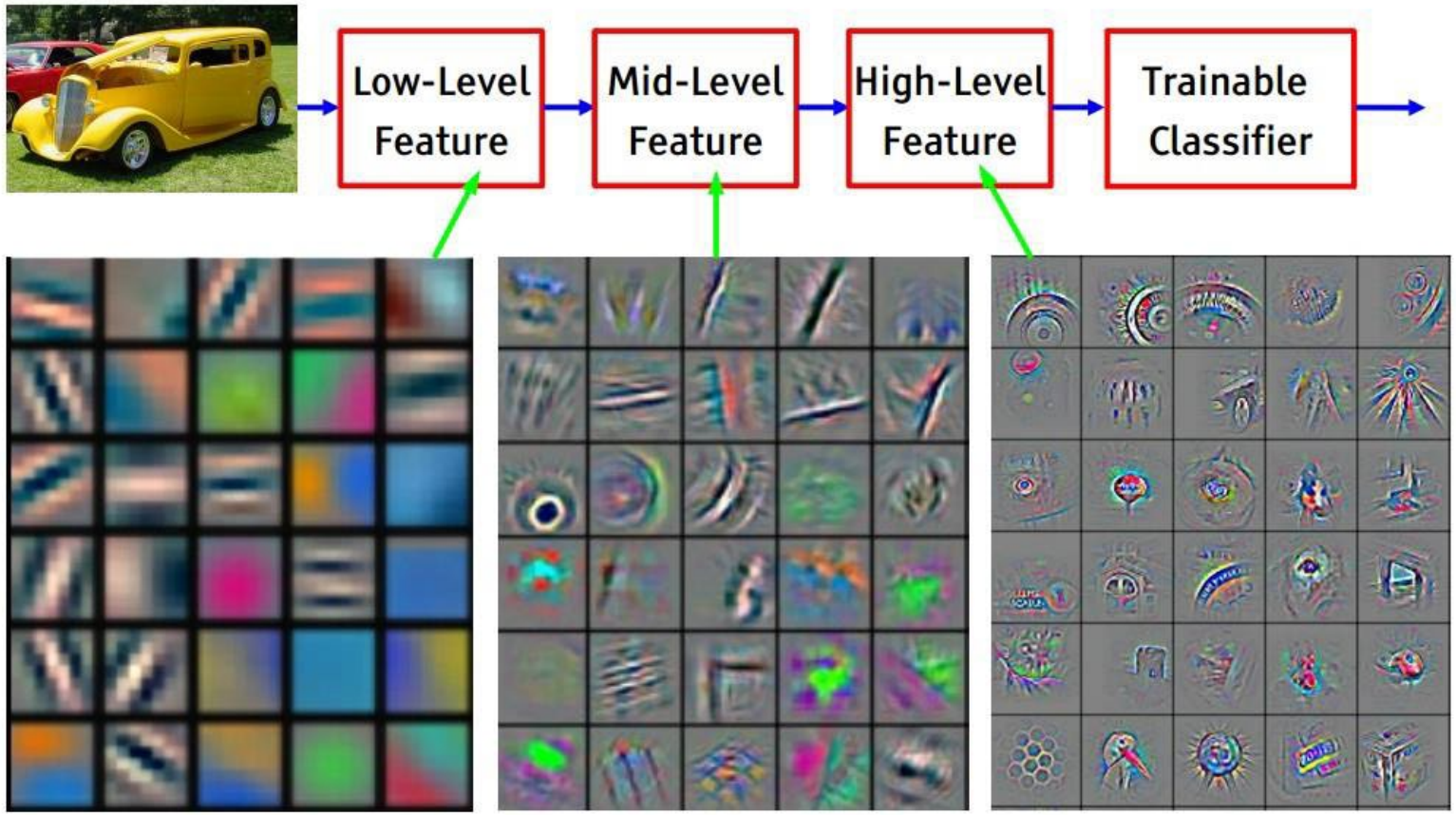


香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY

AQUMON Investment Management

What has been changed by DL?

- **Recall AlexNet**



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

What has been changed by DL?

- **What's more**
  - The learned features are *generalized* so well so that they can easily be used to recognize unseen concepts/classes (knowledge transfer)

query image | nearest neighbors in "code" space (Pool5, 4096D)

What has been changed by DL?

- **To summarize**
  - DCNN *automatically learns* the mid-level representations (~attributes) with various granularity
  - The granularity and mapping are “smoother” while human-defined attributes (the old approach) are “harsher”
  - The feature learning process (by conv layers) and the classification process (by FC layers) are gracefully combined, allowing the gradient to *pass through the whole network from the raw input to the objective (image classification in this case),* a process called ***end-to-end learning***. It is not possible to achieve this by old methods.

Why Deep is OP?

- **Above, imo, are the primary reasons**
- **Nowadays, we learn that**
  - DL can even beat human performance in some problems, e.g., image classification
  - The features learned by DL models from sufficient training data show overwhelming generalizability compared to human-crafted features

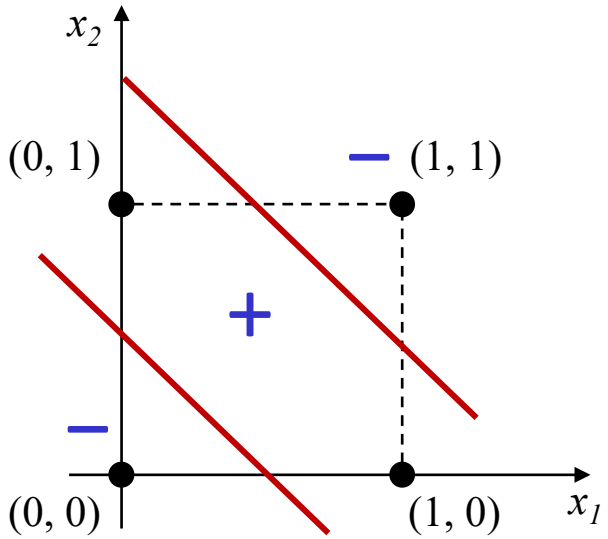
Why Deep is OP?

- **Native nonlinearity**
  - Compared to an SVM with kernel tricks, deep networks are computationally efficient by GPU parallelization and can scale up to extremely big data (as long as you have many GPUs)

- **The power of stacking**
  - This leads to the development of deeper and deeper network designs
  - It makes possible the popular pre-training/fine-tuning pipeline

Why Deep is OP?

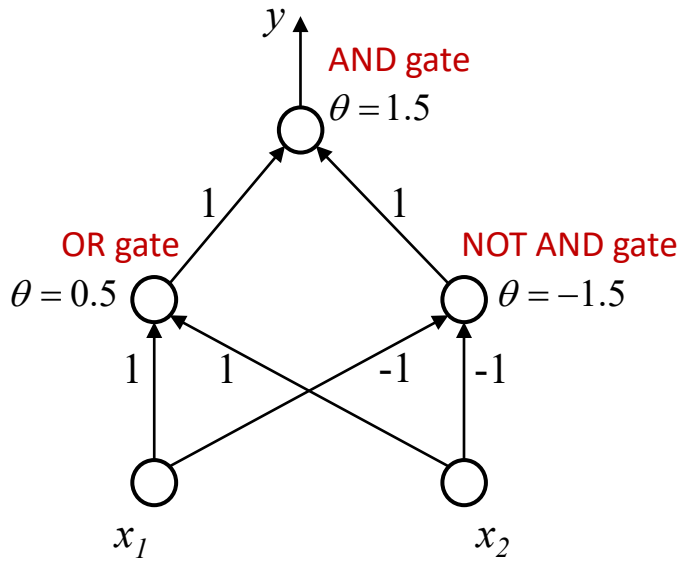
- **The power of stacking**



x_1 xor x_2

Nonlinearity is important

$$y = \mathrm{sgn}(\boldsymbol{\omega}^T \boldsymbol{x} - \boldsymbol{\theta})$$



not linearly separable

香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

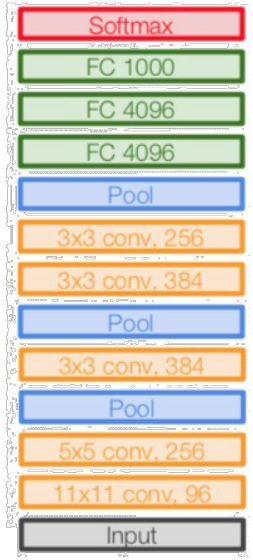
AQUMON Investment Management

Going Deeper

- **ImageNet ILSVRC Challenge (2010—2017)**
  - 1000 object categories
  - 1.2 million training images, human selected and labelled
  - 50k images as validation set
  - Test set: 150k images, blind test
  - Solve classic problems:
    - *Classification*
    - *Classification with localization (bounding box)*

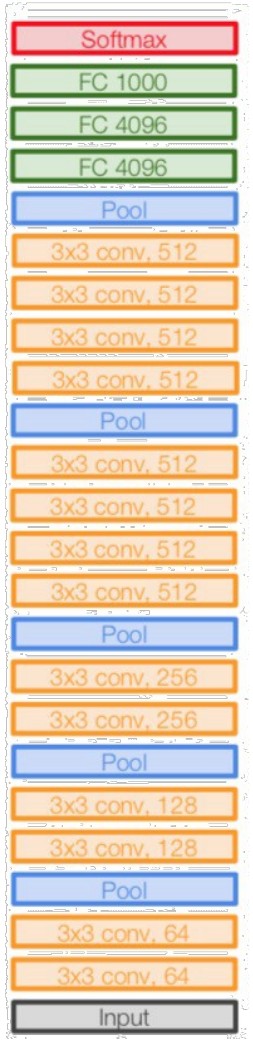
Going Deeper

- **VGGNet – ILSVRC ’14**
  - 16 and 19 layers
  - Simple stacking of 3x3 conv layers
    → *More elegant*
  - Image sizes (/2) are only reduced by pooling layers
    → *More elegant*
  - FC layers for classification are kept the same
  - # of parameters doubled compared to AlexNet
  - Huge resource usage (memory, FLOPs)



AlexNet

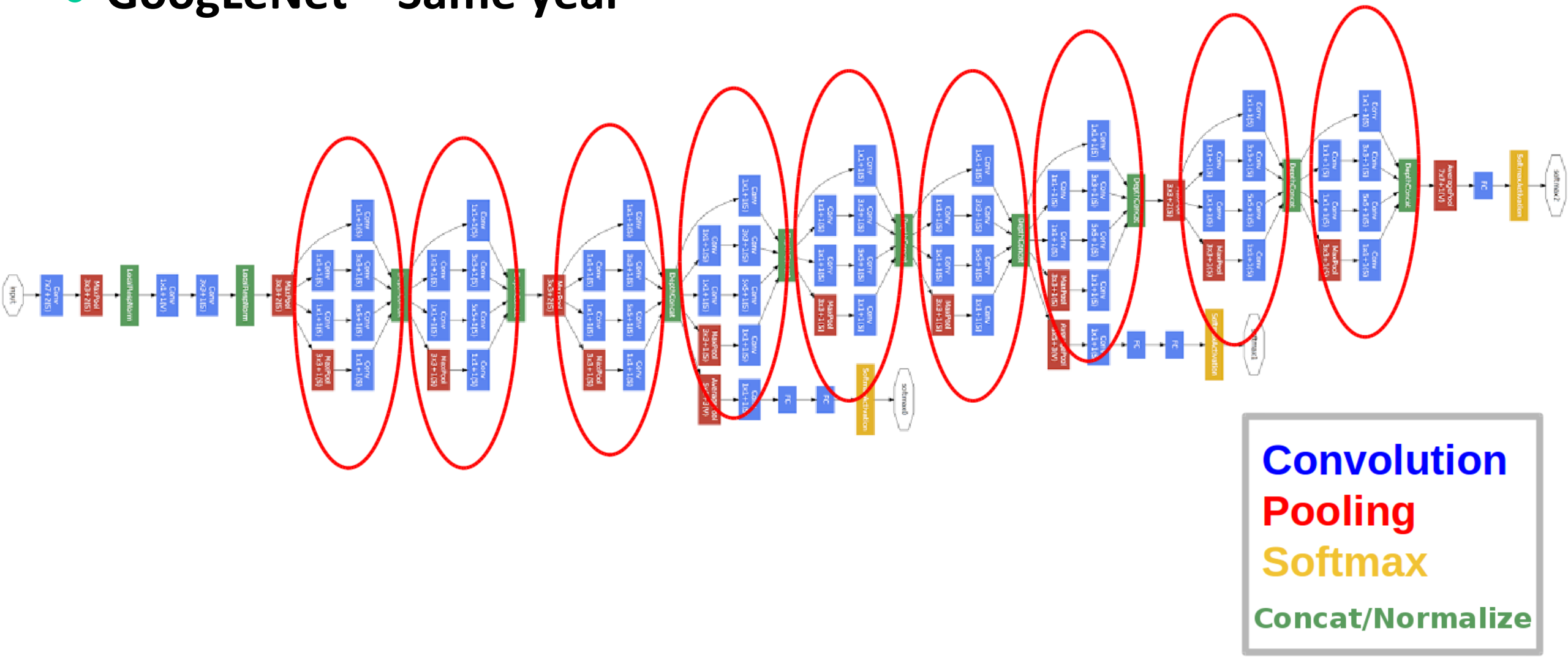
VGG16



VGG19

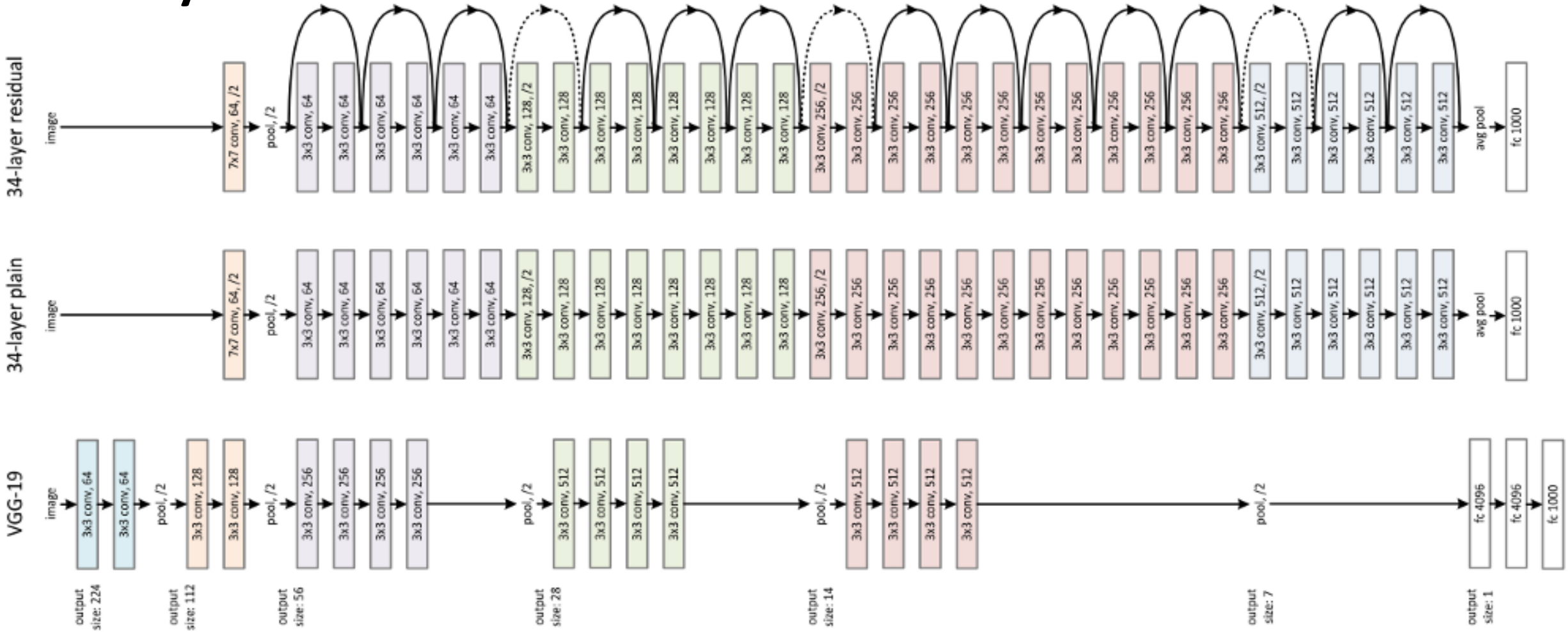
Going Deeper

- **GoogLeNet – Same year**



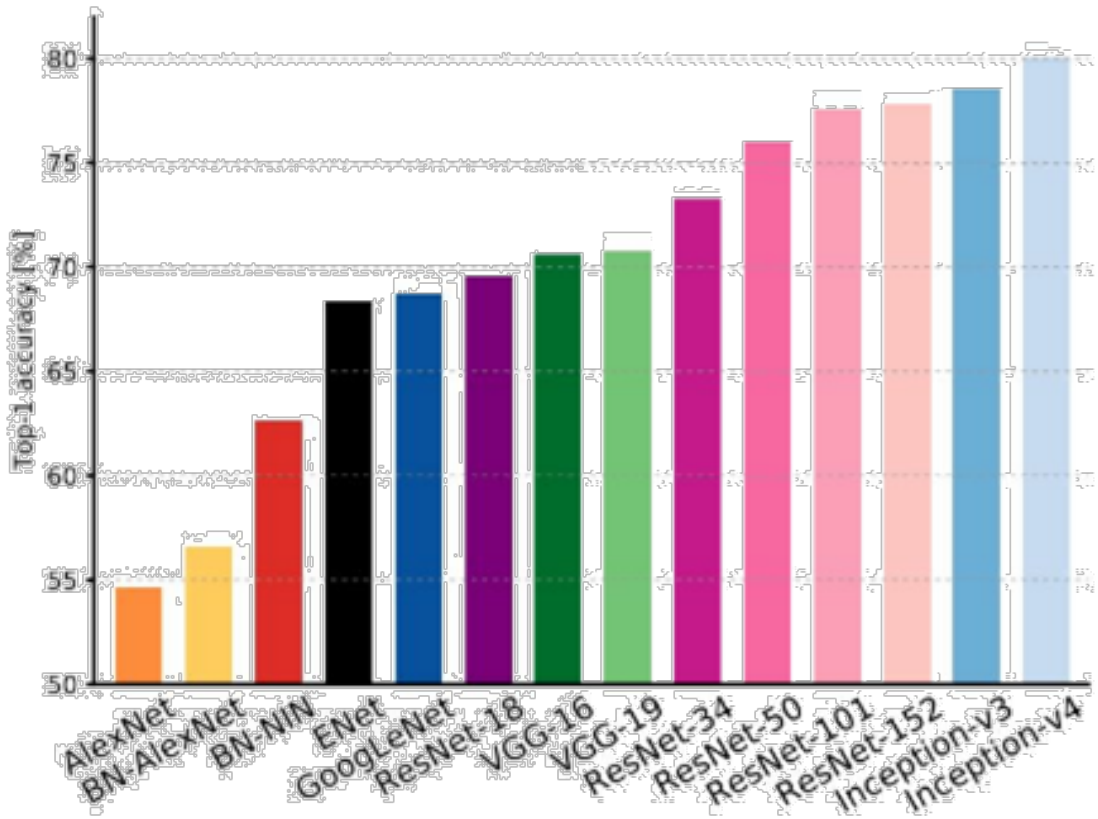
Going Deeper

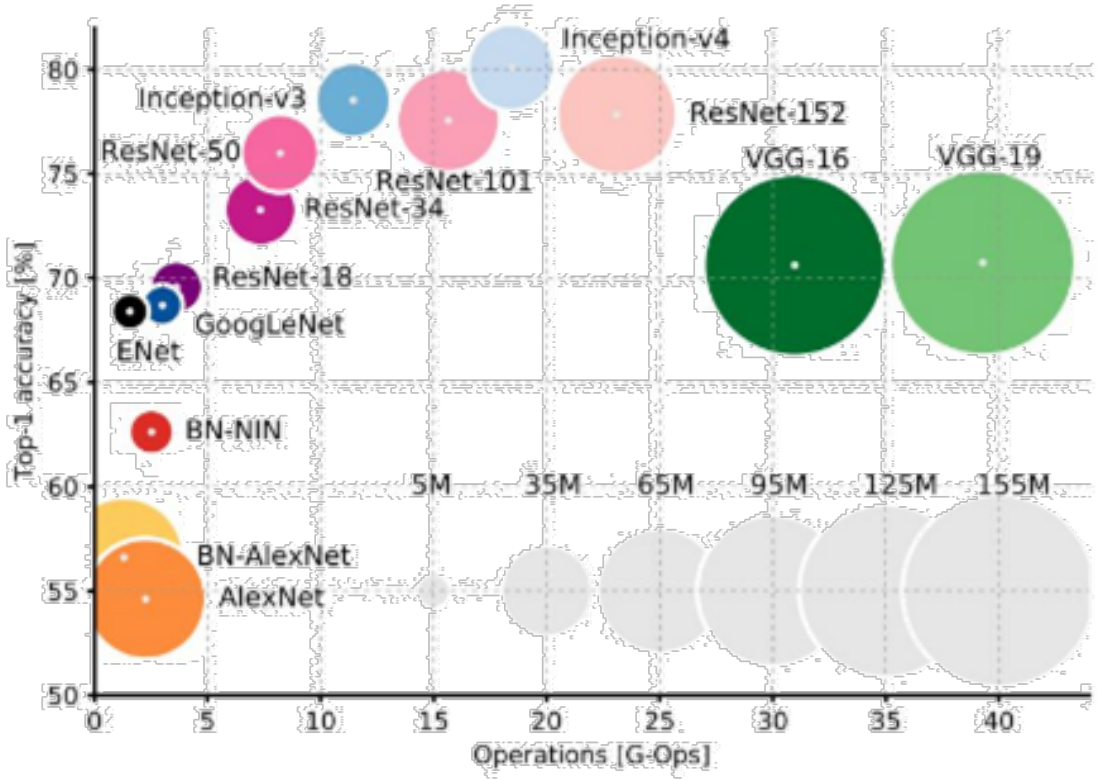
- **34-layer ResNet**



Going Deeper

- **Comparison**





An Analysis of Deep Neural Network Models for Practical Applications, 2017.

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Why Deep Learning is OP?

It's Highly Versatile

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON | Investment Management

Reinforcement Learning

- **Recently, we have seen remarkable success in the combination of**
  - Deep learning
  - Reinforcement learning

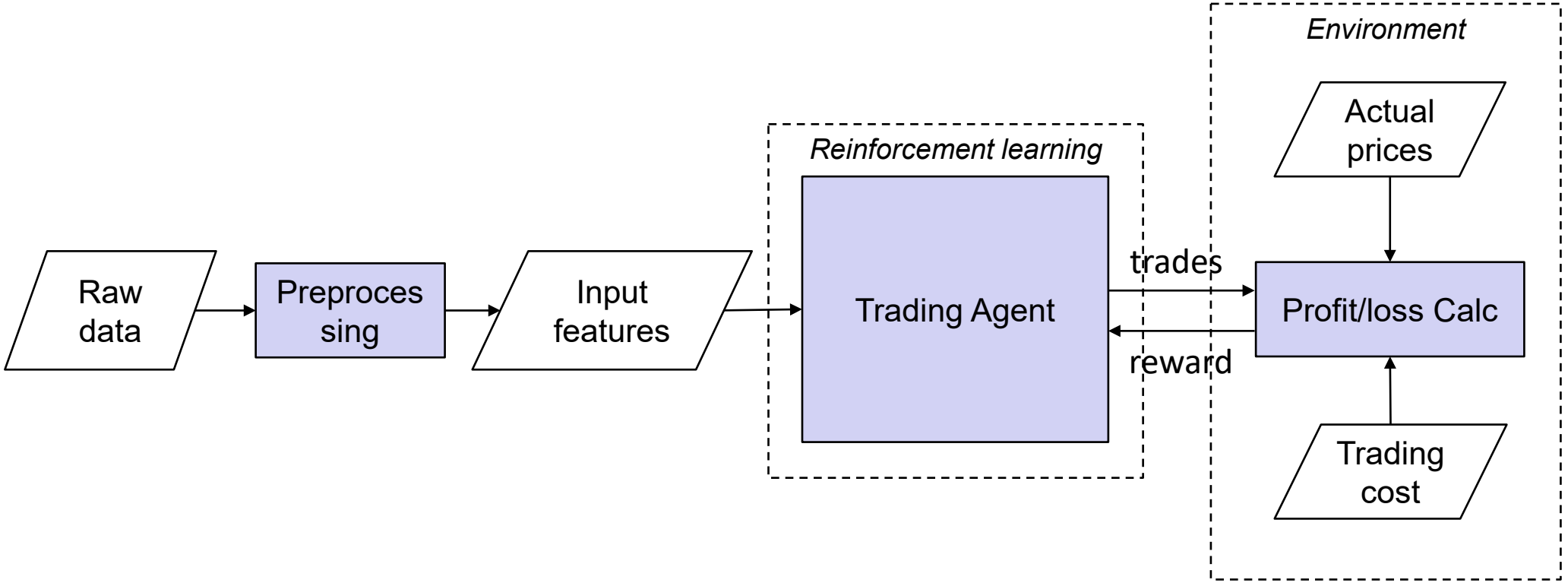
2015-2017

2017-2019

2014-2020

Reinforcement Learning

- **General Framework**



香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

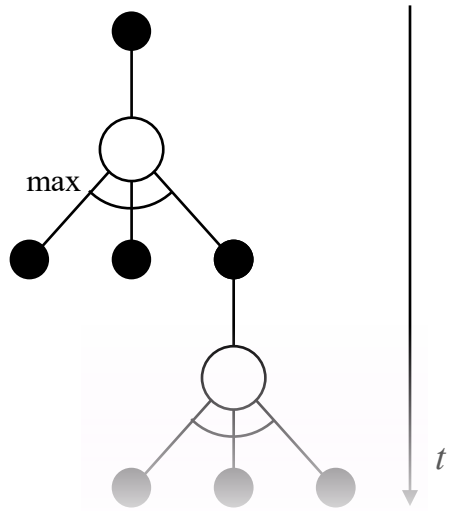
AQUMON Investment Management

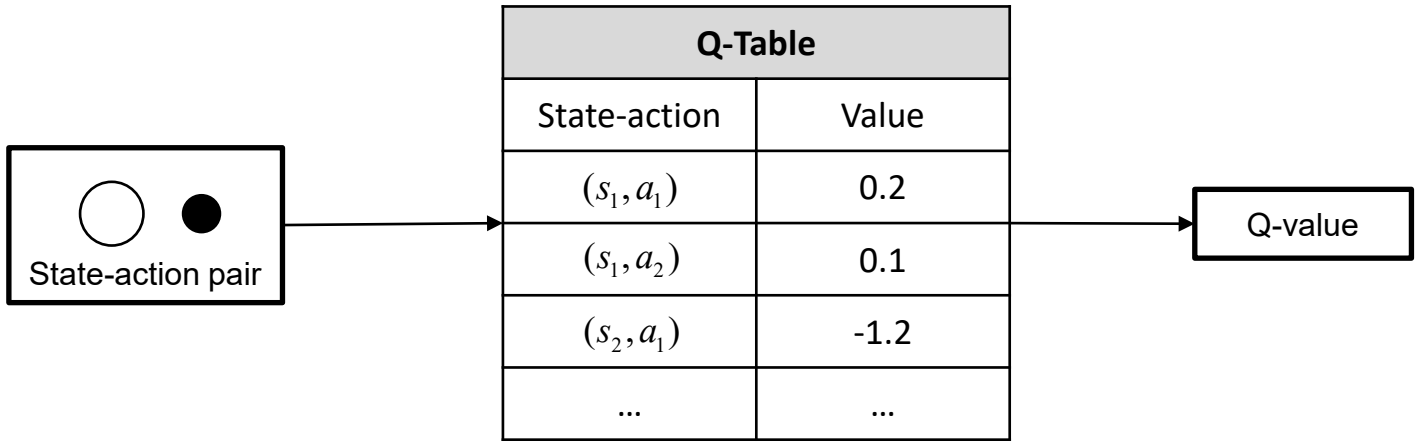
Reinforcement Learning

- **Q-Table**

$$Q(s,a) = r(s,a) + \gamma \max_{a'} Q(s',a')$$

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \max_{a} Q(S_{t+1}, a) - Q(S_t, A_t) \right]$$



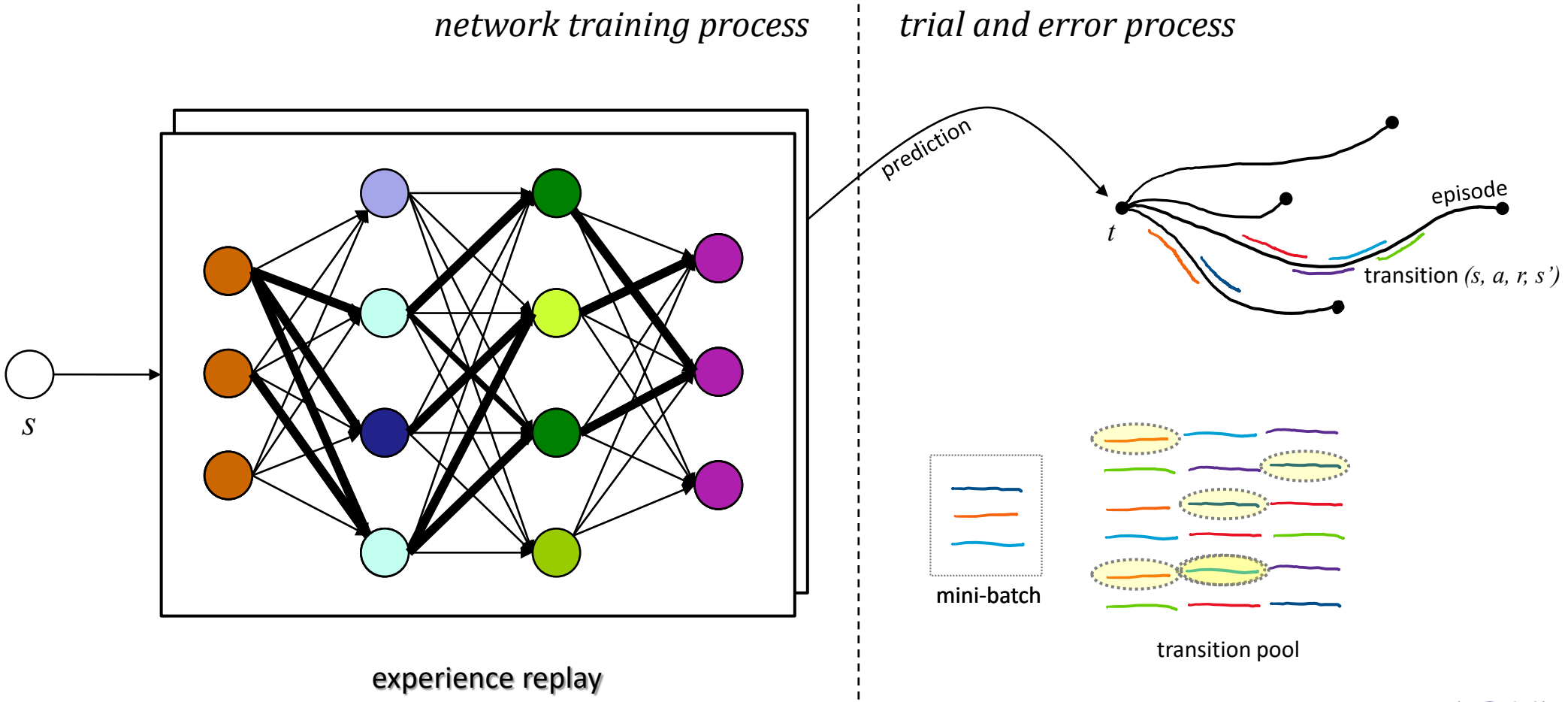


香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Reinforcement Learning

- **Q-Network Training**



香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

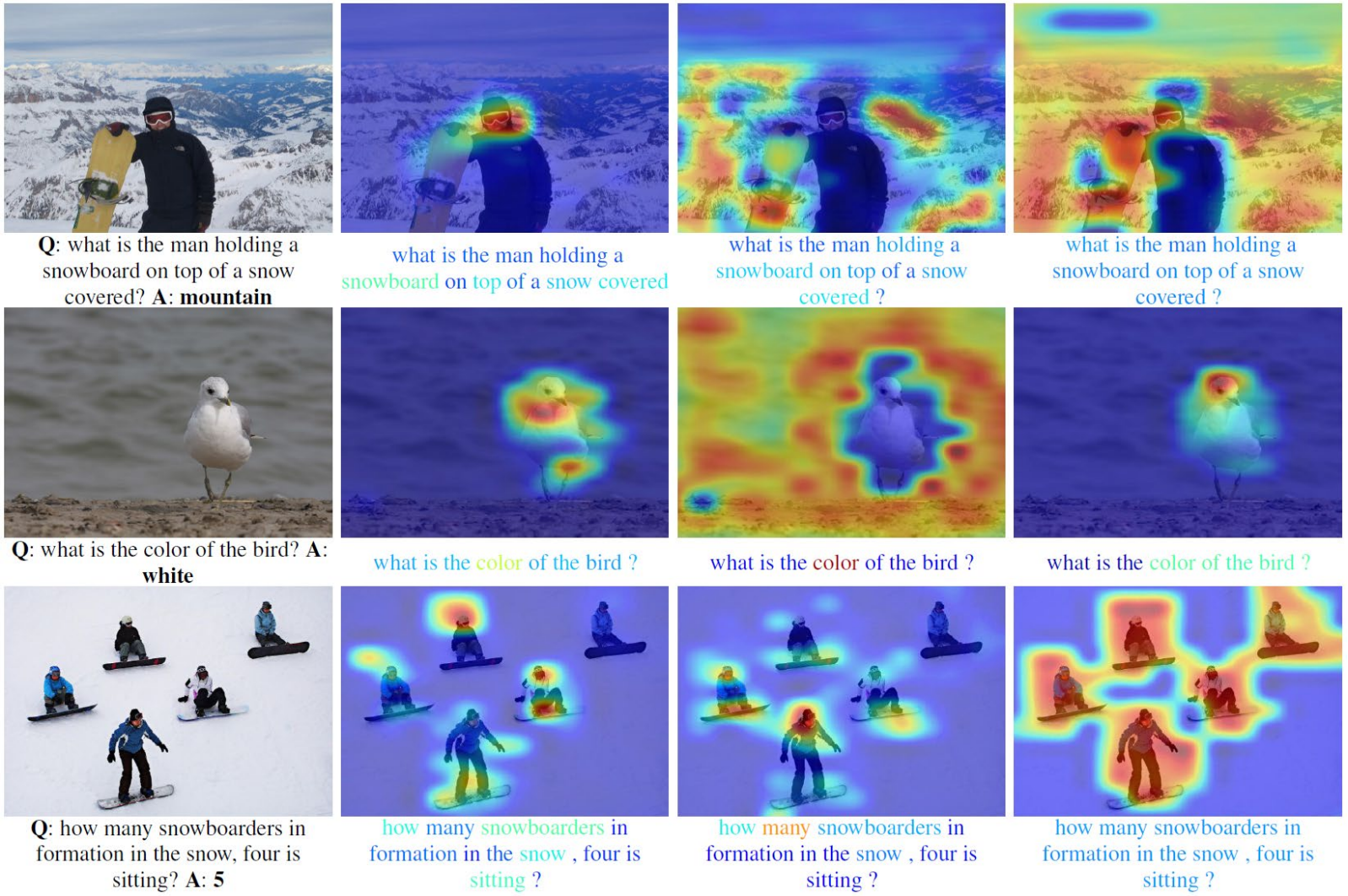
Reinforcement Learning

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Visual Question Answering

- **LSTM + Attention**



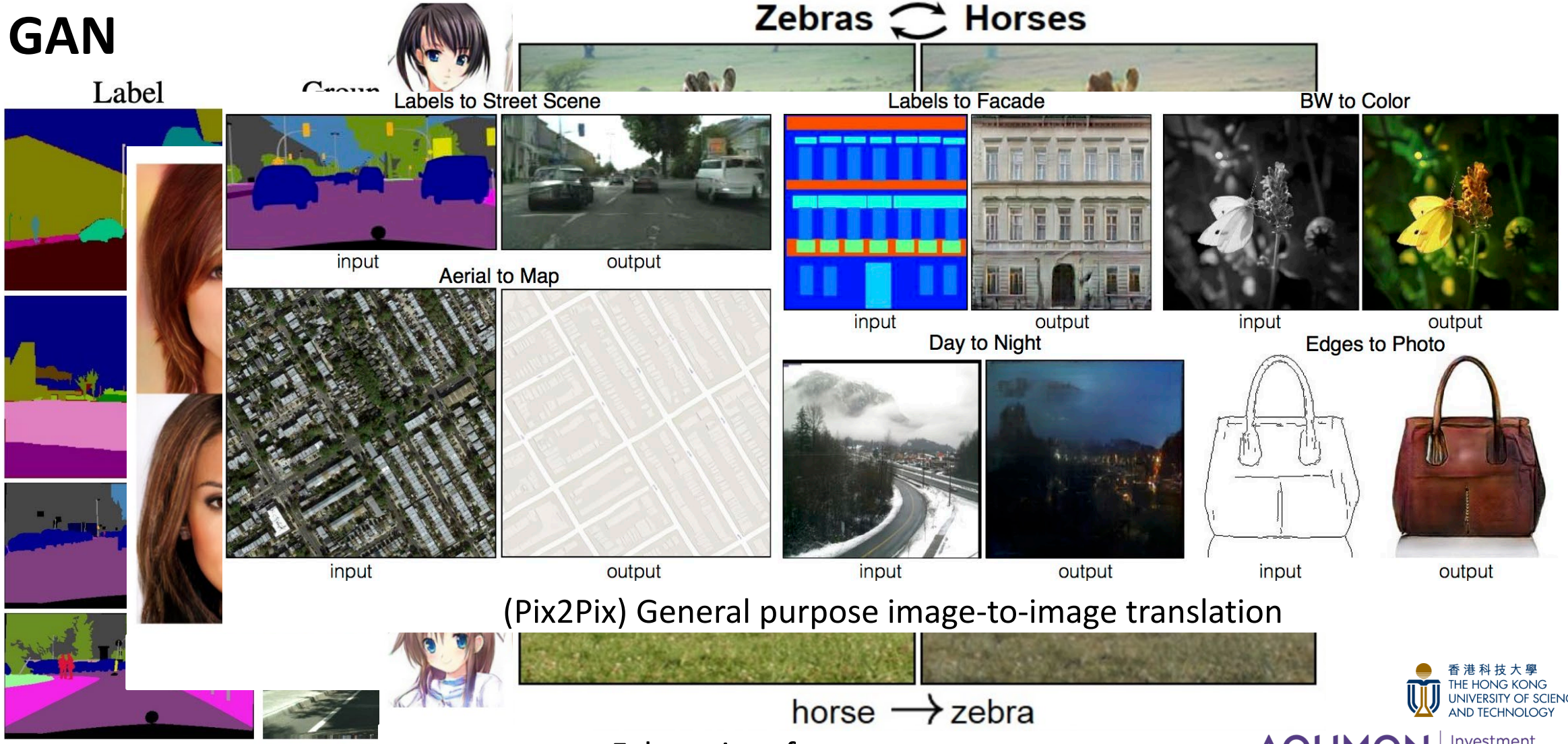
[1] Hierarchical Question-Image Co-Attention for Visual Question Answering. NIPS 2016.

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Image Creation | Styling | Transformation

- **GAN**



香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Why Deep Learning is OP?

- **To sum up**
  - Superior performance against classic ML models
  - Good generalizability that resembles human-level recognition
  - Building blocks like LEGO
    - *End-to-end learning*
    - *Unique pipeline of pre-training and fine-tuning*
  - Model capacity
    - *Inborn nonlinearity*
    - *Able to be stacked up*
  - Scalable to very big data
    - *GPU parallelization*
    - *Prediction can be very fast*

However, good things often come with a cost…

More Kind Suggestions

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON | Investment Management

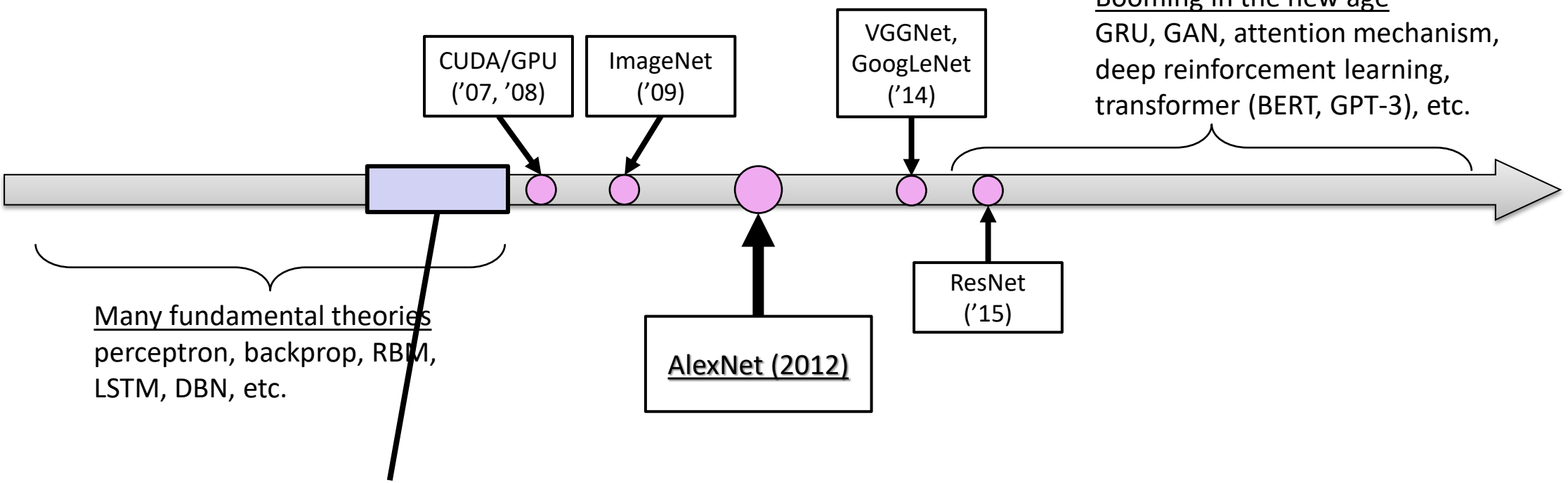
Practice Practice Practice!

- **Deep learning has always been the thing where *theory meets practice*, and often to the extreme:**
  - Theory is when you know everything but nothing works
  - Practice is when everything works but no one knows why
  - We feel obligated to combine the theory and practice:
    - ~~*Nothing works and no one knows why*~~

Practice Practice Practice!

- **Deep learning has always been the thing where *theory meets practice*, and often to the extreme**
  - Don’t treat it as a black box
  - Begin with the simplest model, gradually add up the complexity
  - Expect it’s not working at the beginning

Recall Pre-DL age



香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY

AQUMON Investment Management

Practice Practice Practice!

- **Deep learning has always been the thing where *theory meets practice*, and often to the extreme**
  - Don’t treat it as a black box
  - Begin with the simplest model, gradually add up the complexity
  - Expect it’s not working at the beginning
  - If it is still not working, be a good “copycat”—start with well-documented existing dataset first
  - Expect it takes much longer for you to figure it out compared to classic ML algorithms
  - Discuss with others

Big Data?

- **In AQUMON, we have a great amount of data:**
  - Subscribed
    - *Bloomberg, Refinitiv, MorningStar (for global equities and ETF)*
    - *Wind, Choice (for CN equities)*
  - TB levels of data maintained internally
    - *Multiple assets: stocks, futures, forex, crypto, company fundamentals, analyst opinions*
    - *Various frequency: tick, minutely, daily, level-2*
  - These are not available in personal collections or even at university

Discussion

How About DL for Trading?

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

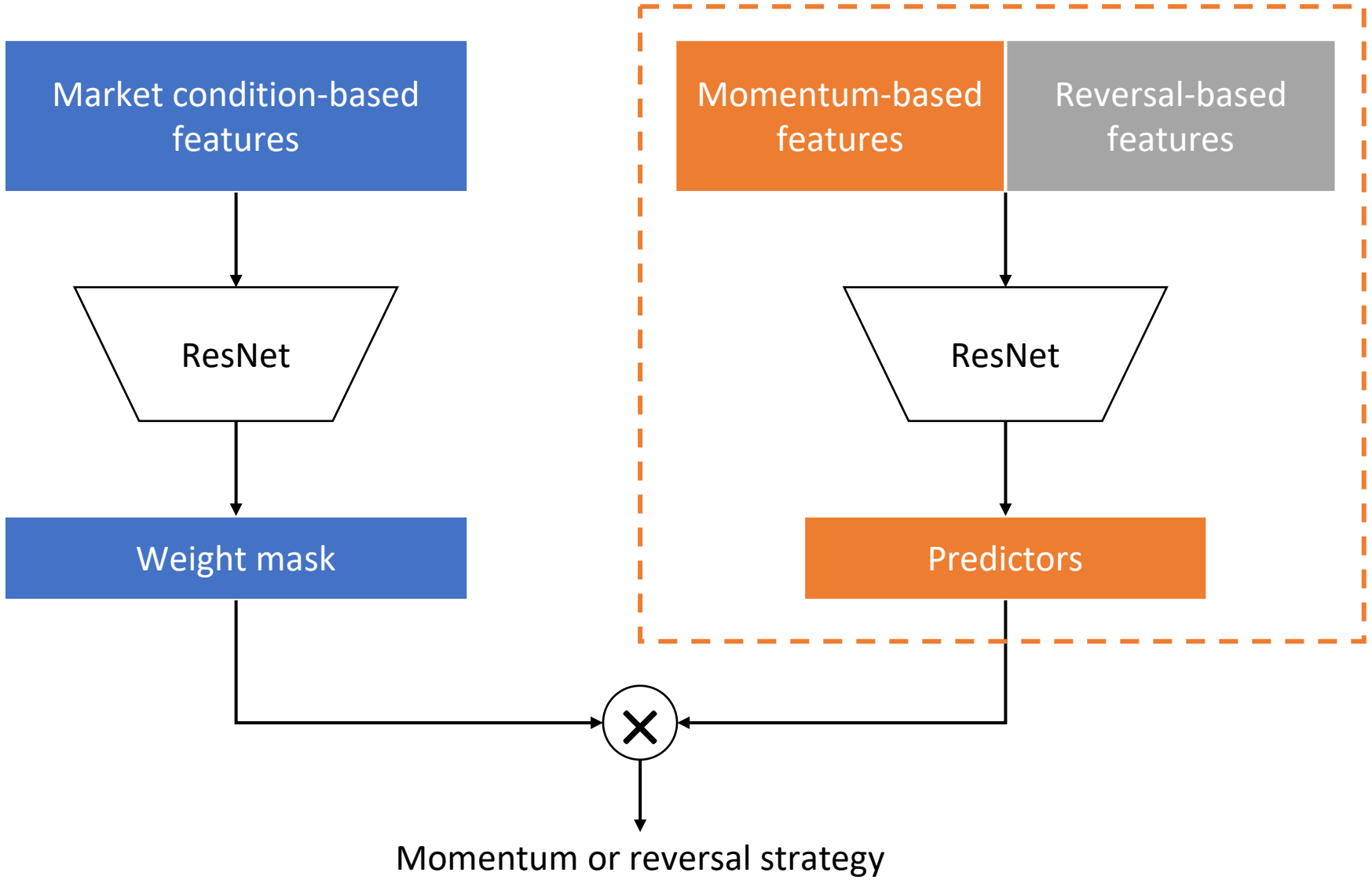
AQUMON Investment Management

DL for Trading

- **Industrywide, we care more about**
  - return that can eventually be *realized*
  - or *boost* the return that can eventually be realized :)

- **Here, we introduce some of the works we are doing**
  - that covers a wide range of spectrum from research to engineering

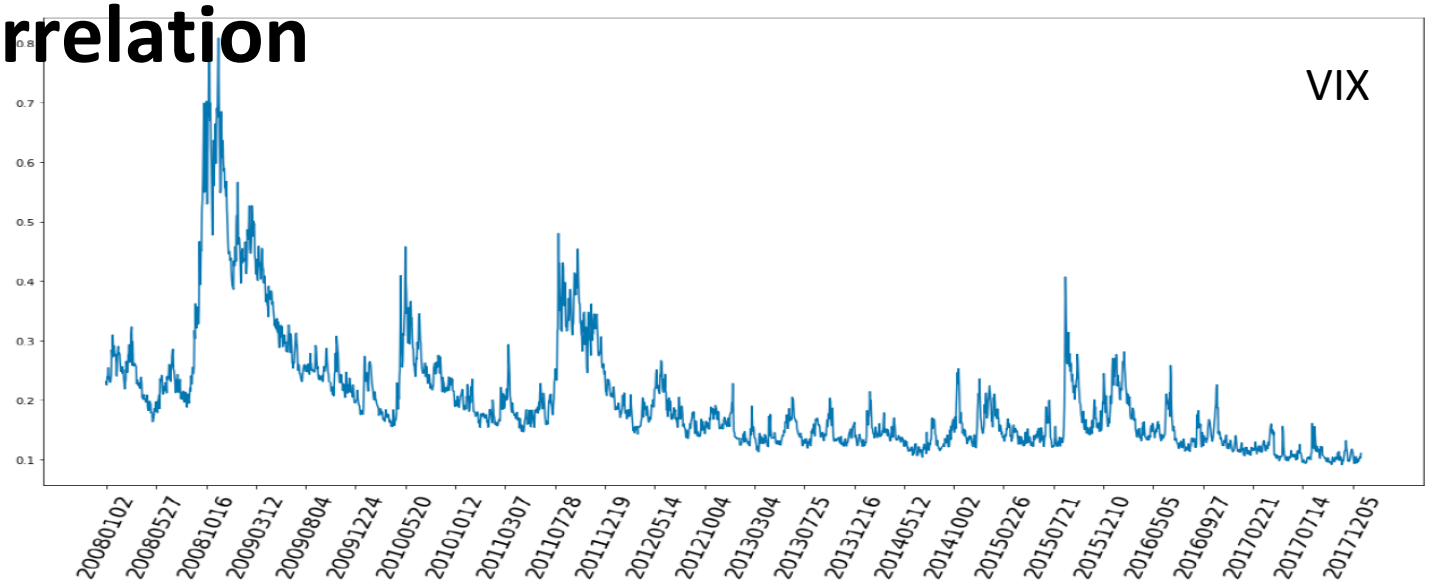
Momentum/Reversal Switch

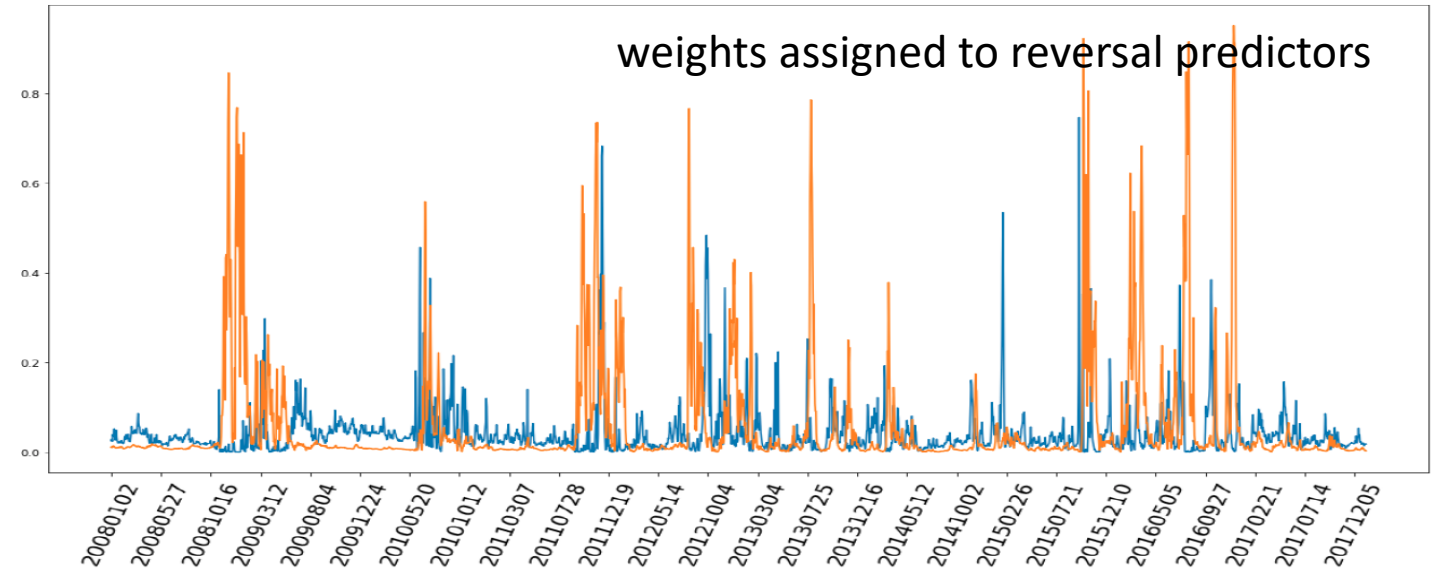


香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY
AQUMON Investment Management

Momentum/Reversal Switch

- **Positive correlation**



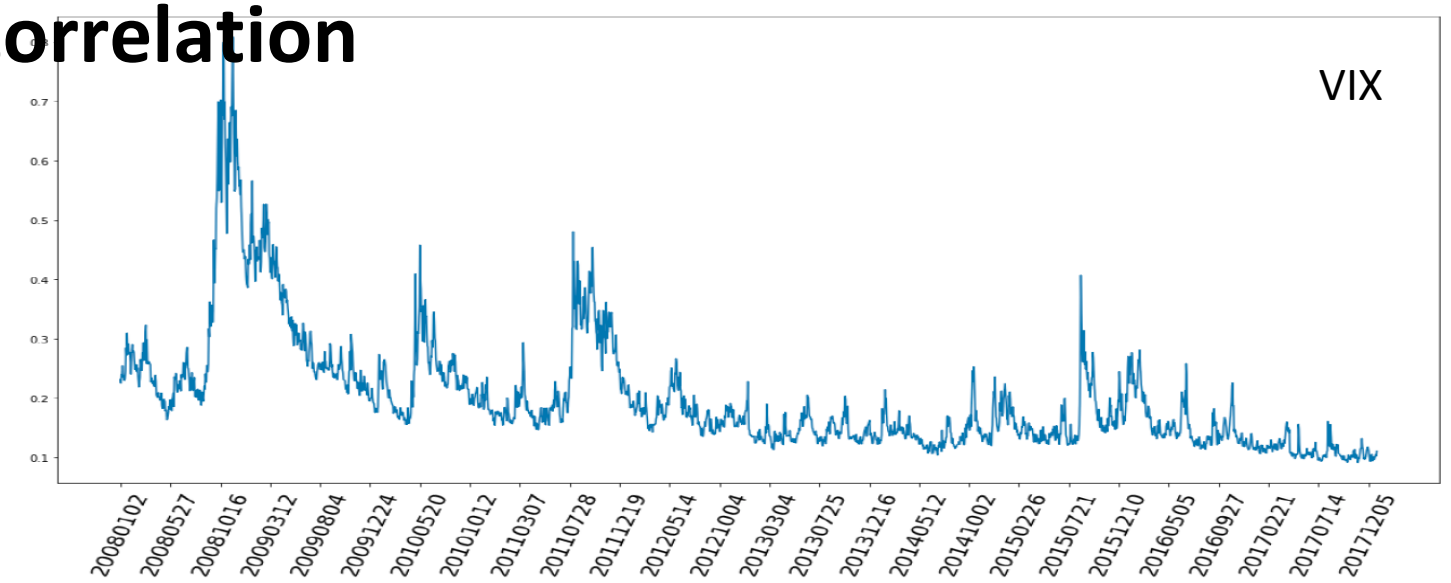


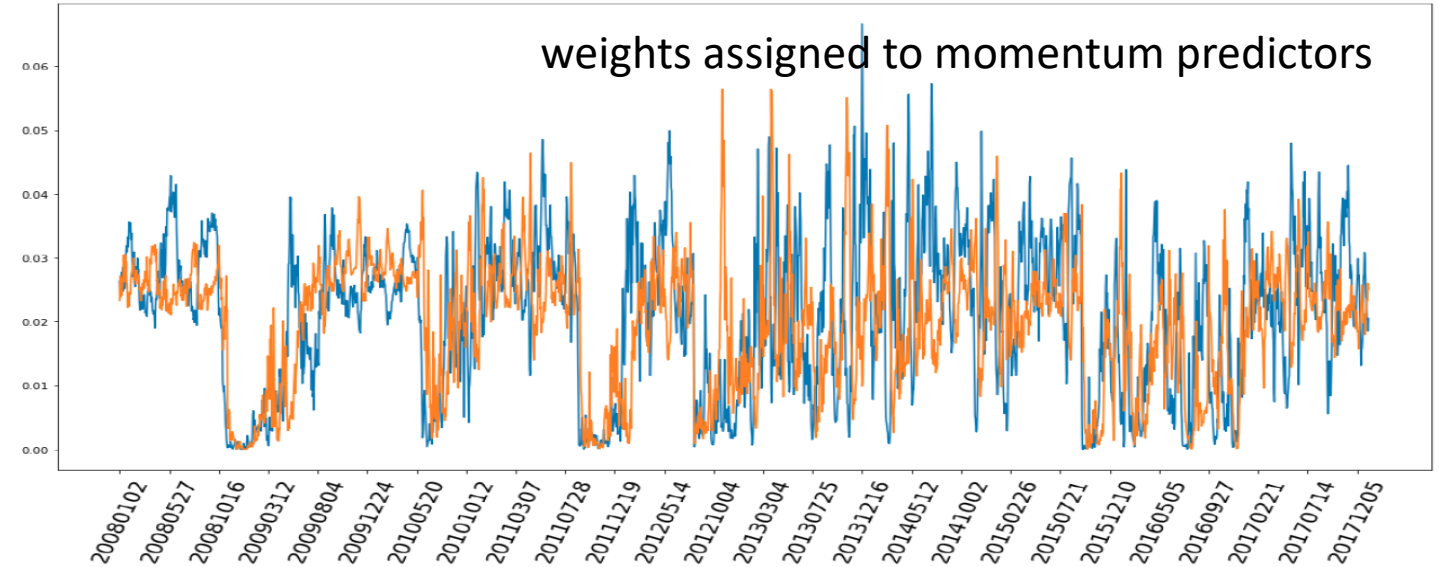
香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Momentum/Reversal Switch

- **Negative correlation**





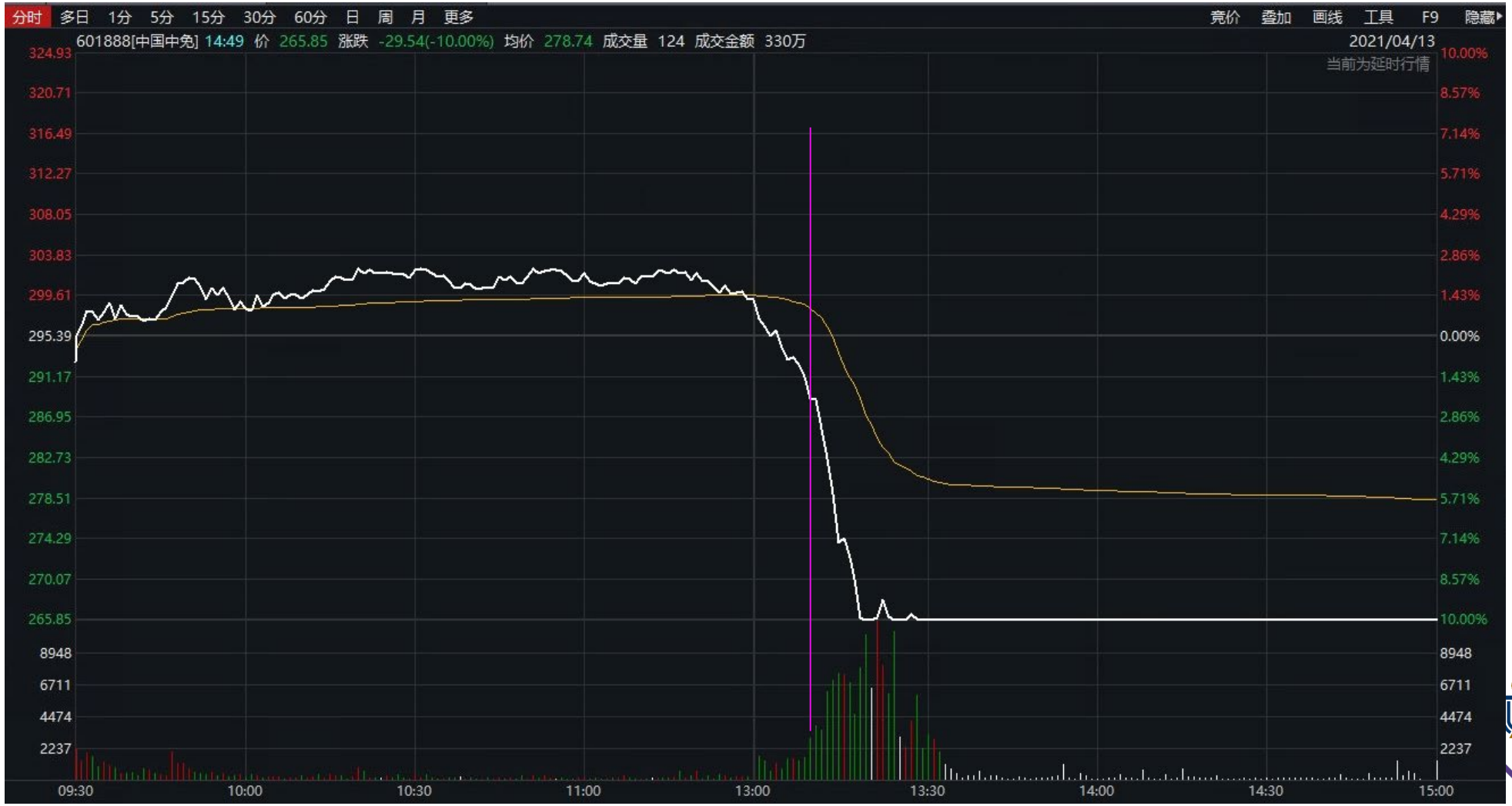
香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Volume Surge

- **Observation**

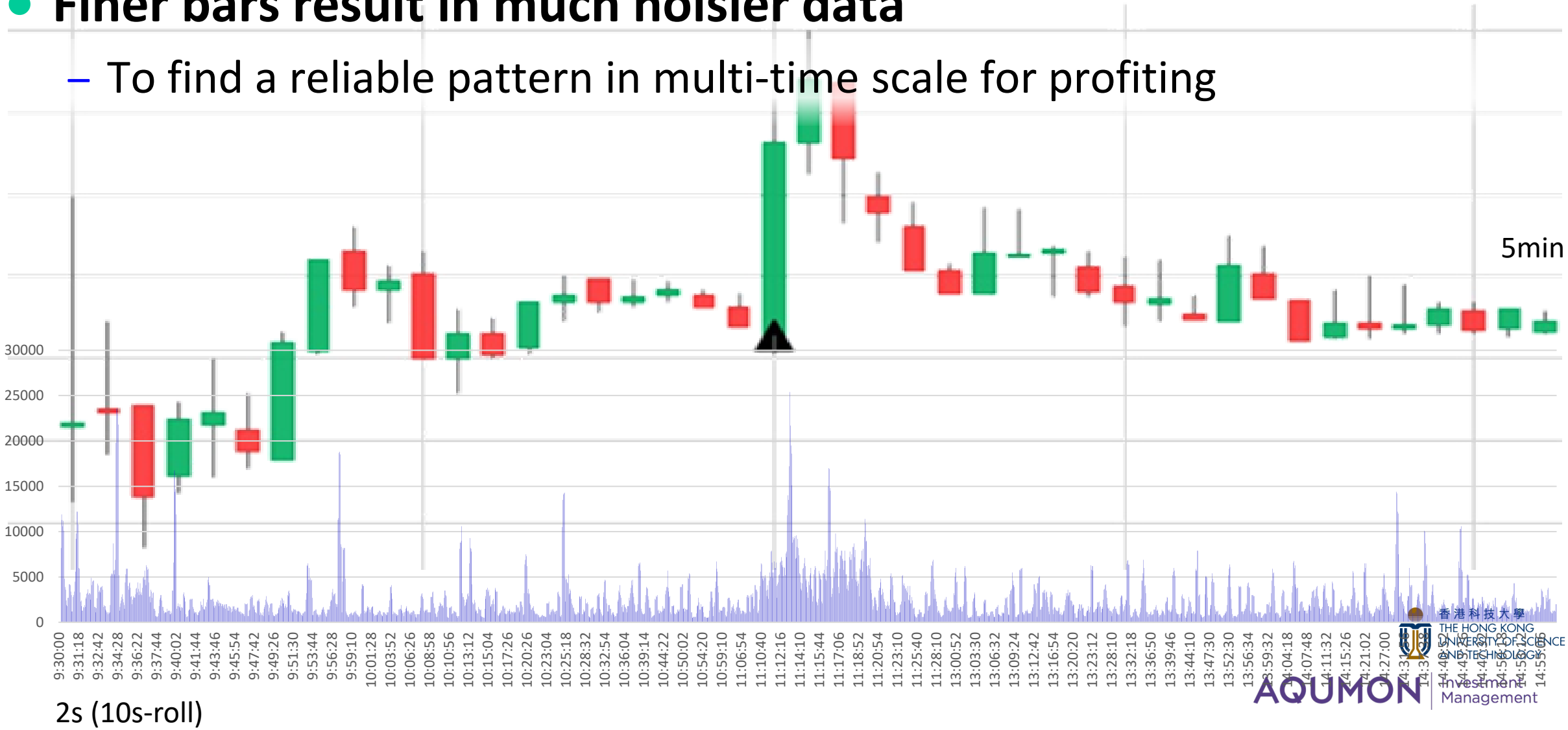
1min



香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY
Investment Management

Volume Surge

- **Finer bars result in much noisier data**
  - To find a reliable pattern in multi-time scale for profiting



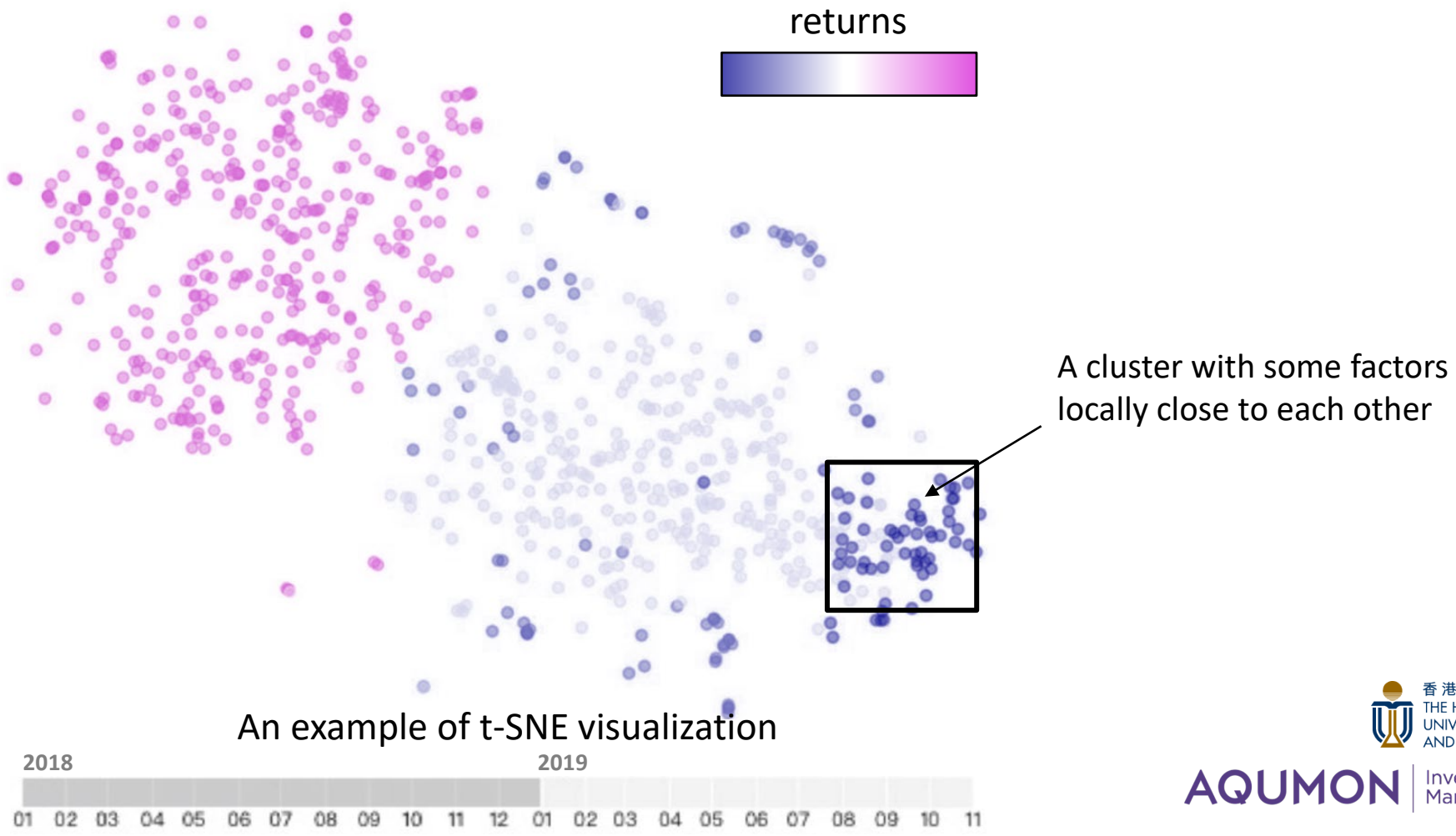
2s (10s-roll)

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Visualization

- **Strategy Visualization for Analysis**



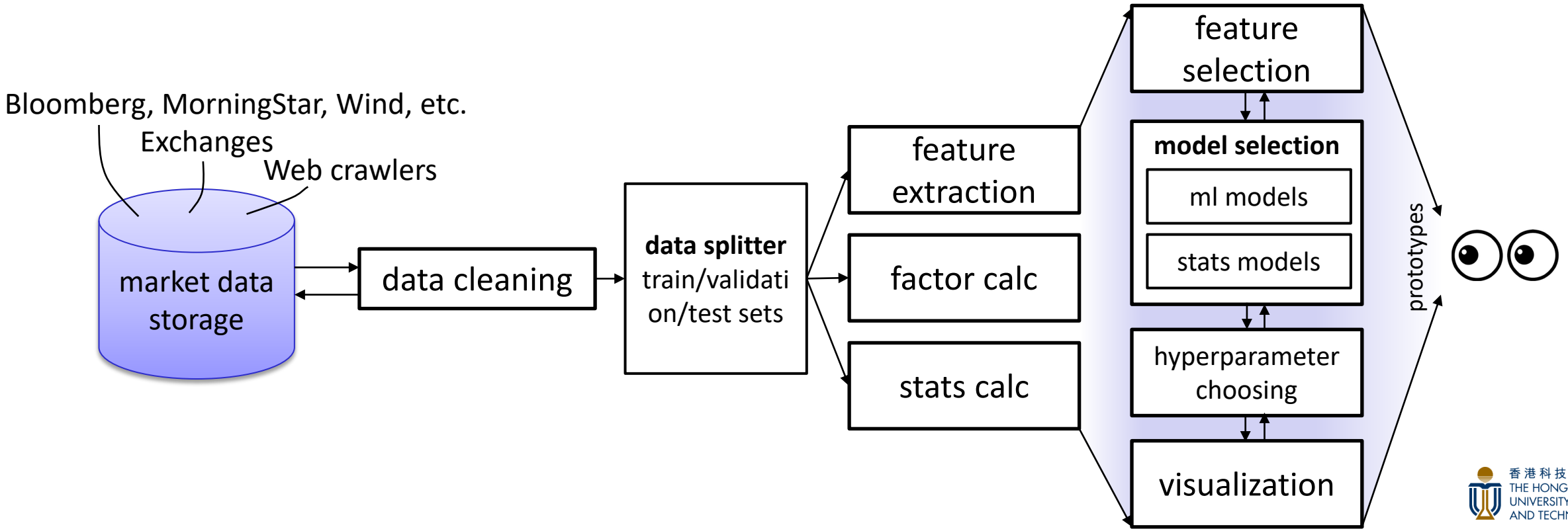
An example of t-SNE visualization

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

AutoML

- **Goal**
  - To reduce tedious and repetitive jobs for data science research



香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Action Rather than Dreaming

- **I do believe DL has a promising future in quant trading as long as we overcome a few obstacles like**
  - have sufficient data
  - avoid forward looking and overfitting
  - have satisfactory equipment
  - time input: practice and dedication

We as a Team

- **AQUMON**
  - Root in research
    - *3x PhD in Mathematics*
    - *2x PhD in Physics*
    - *1x PhD in Computer Science*
    - *1x PhD in Electronic Engineering*
    - *1x PhD in Finance*
    - *10x Master in Financial Engineering, Computer Science*
  - AUM: 1 billion RMB currently
  - We encourage **group sharing and believe in collaboration**

Interested? Contact our HR Ling: career@aqumon.com

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON Investment Management

Thank You

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AQUMON | Investment Management

Key References

[1] SIFT features (IJCV'04): https://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf

[2] ImageNet LSVRC: https://www.image-net.org/challenges/LSVRC/

[3] AlexNet (NIPS'12): https://proceedings.neurips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[4] VGGNet (ICLR'15): https://arxiv.org/abs/1409.1556

[5] GoogLeNet (CVPR'15): https://arxiv.org/abs/1409.4842

[6] ResNet (CVPR'16): https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf

[7] Deep Reinforcement Learning (Nature'15): https://deepmind.com/research/publications/2019/human-level-control-through-deep-reinforcement-learning