

# CNN- Above Image

12/21 Bgg

# Agenda

- Object Detection
- 3D CNN - Convolution + Pooling
  - A quick example
  - Why need Convolutional Neural Network(CNN)
  - Applications
- KKT example (17.11)

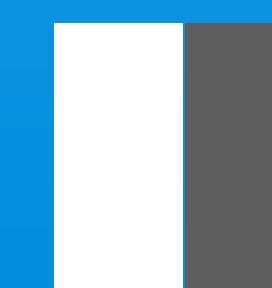
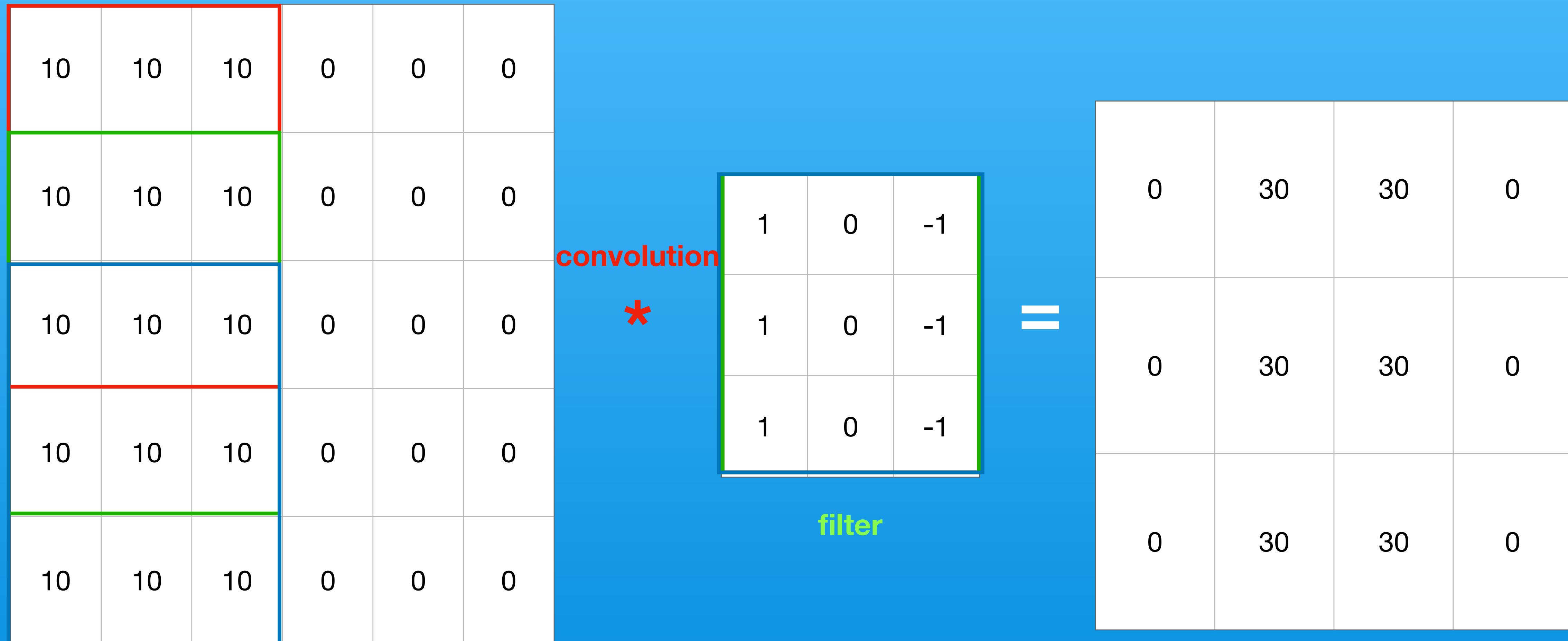
# Agenda

- **Object Detection**
- 3D CNN - Convolution + Pooling
  - A quick example
  - Why need Convolutional Neural Network(CNN)
  - Applications
- KKT example (17.11)

或

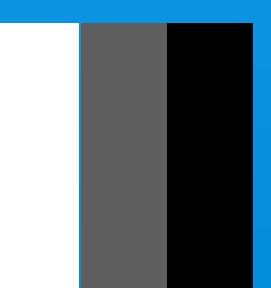


# Vertical edge detection



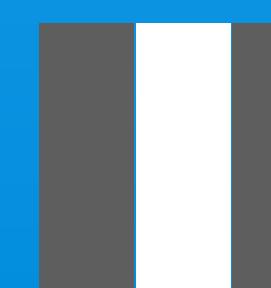
$5 \times 6$

\*



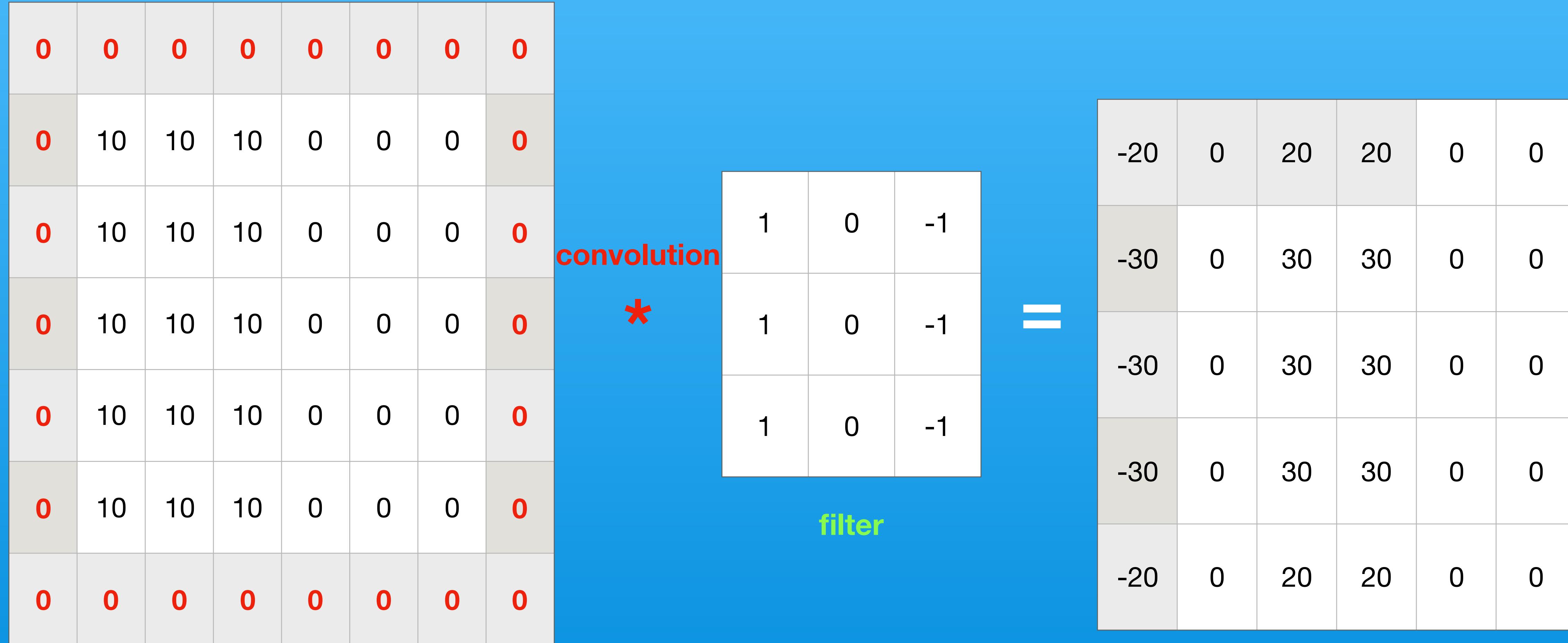
$3 \times 3$

=



$3 \times 4$

# Padding



$5 \times 6$

\*

$3 \times 3$

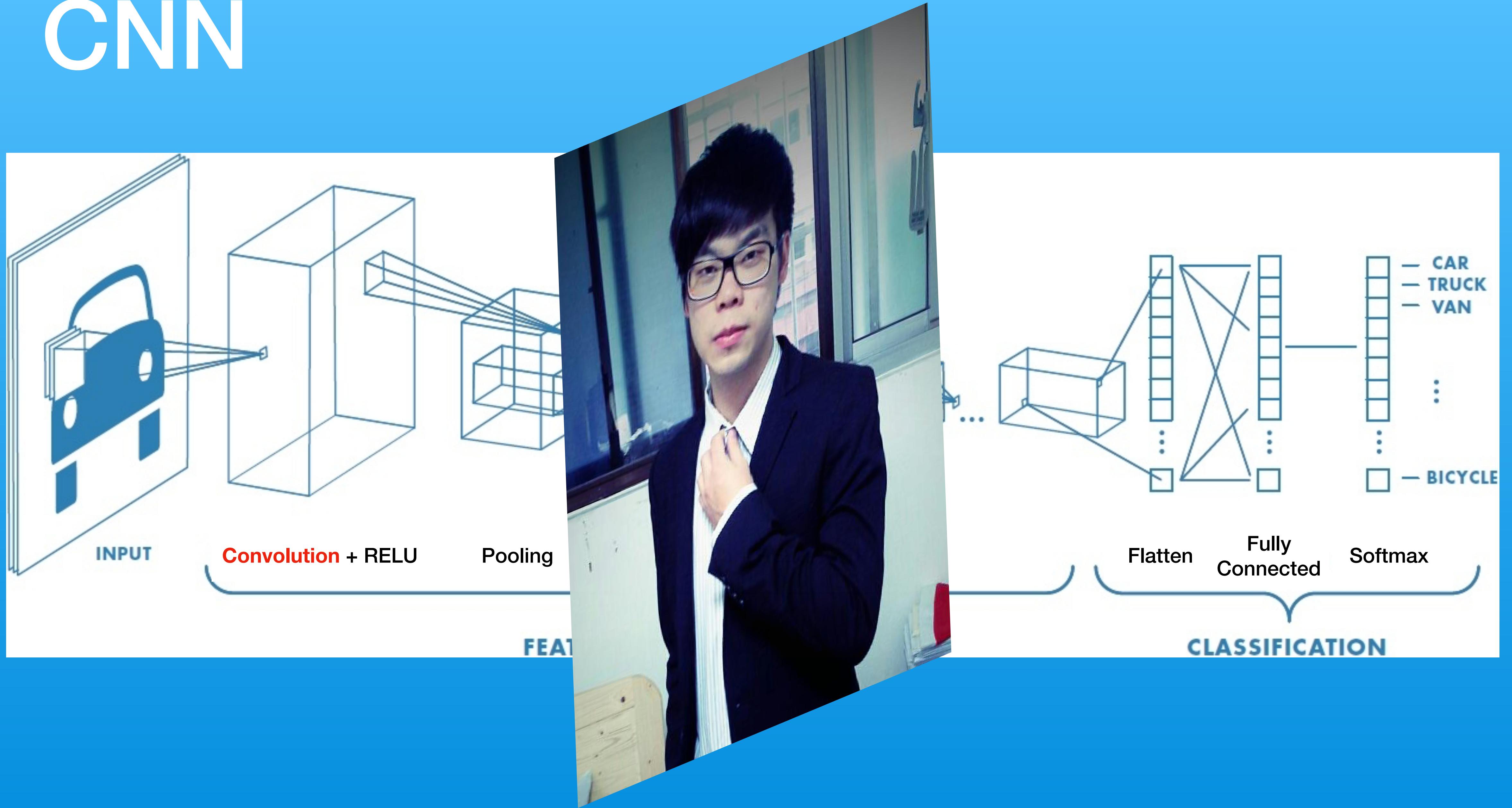
=

$5 \times 6$

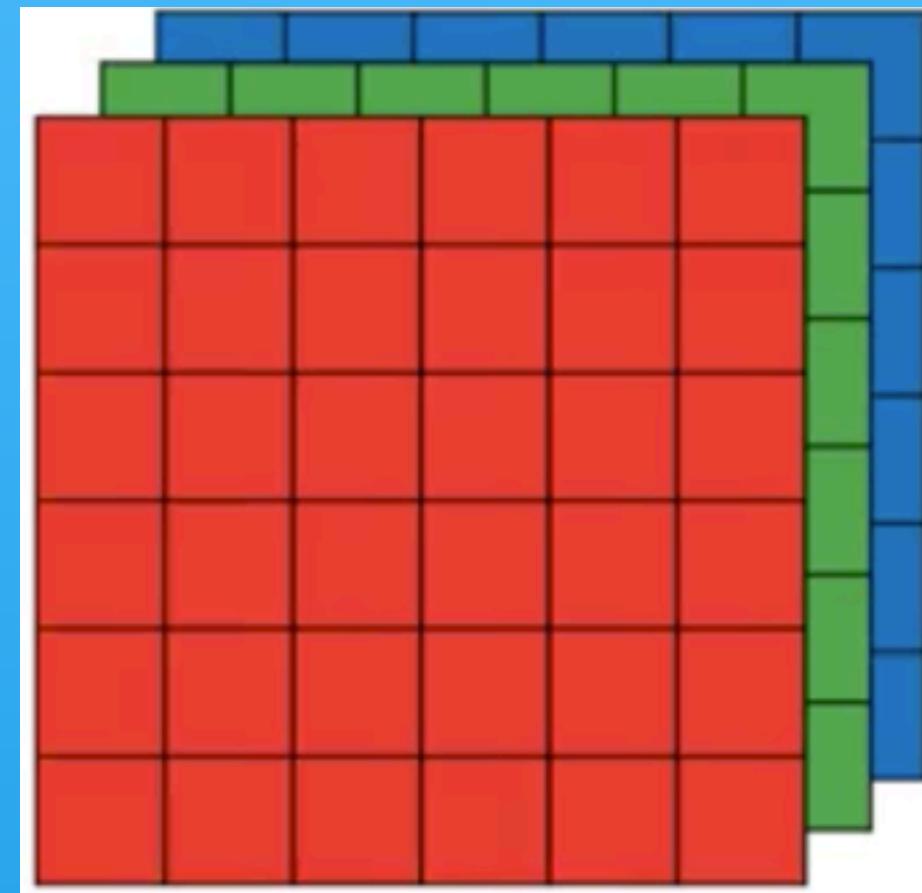
# Agenda

- Object Detection
- **3D CNN - Convolution + Pooling**
  - A quick example
  - Why need Convolutional Neural Network(CNN)
  - Applications
- KKT example (17.11)

# CNN



# Convolution

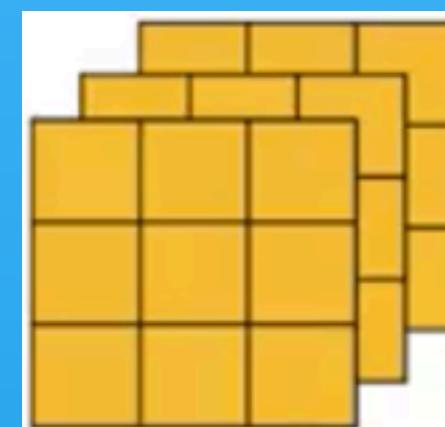
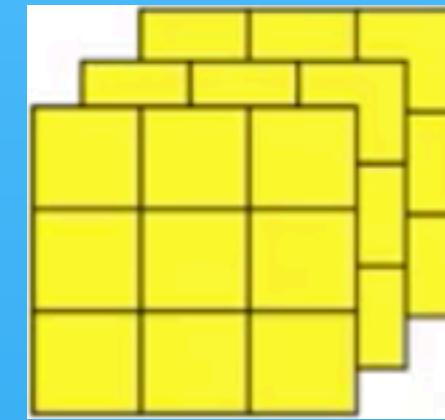


$6 \times 6 \times 3$

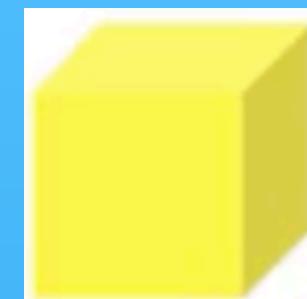
convolution

\*

$3 \times 3 \times 3$

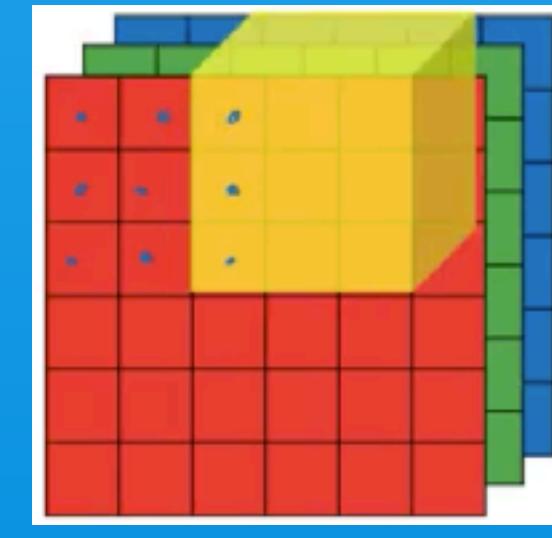
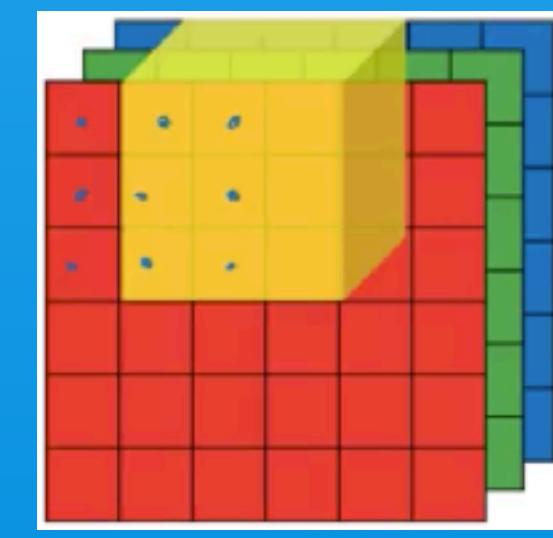
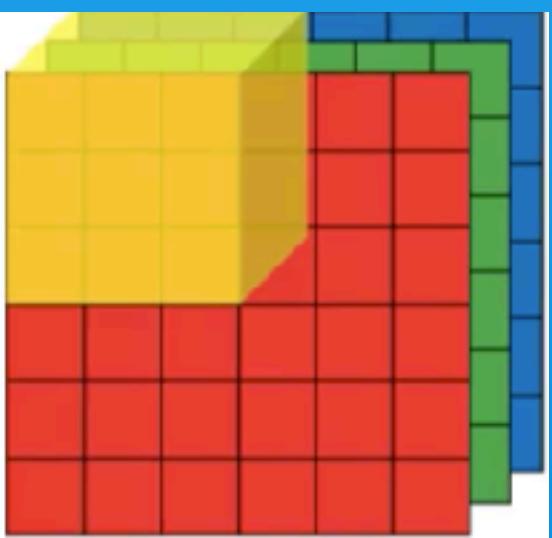
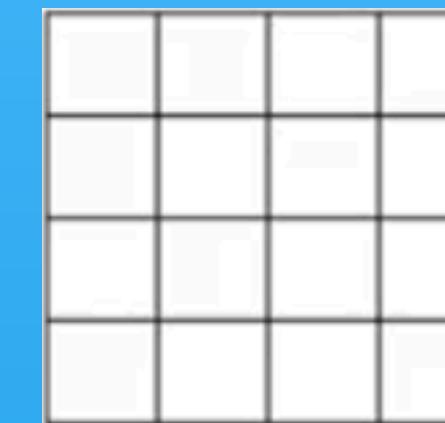


stride = 1  
filter = 2



=

$4 \times 4 \times 2$

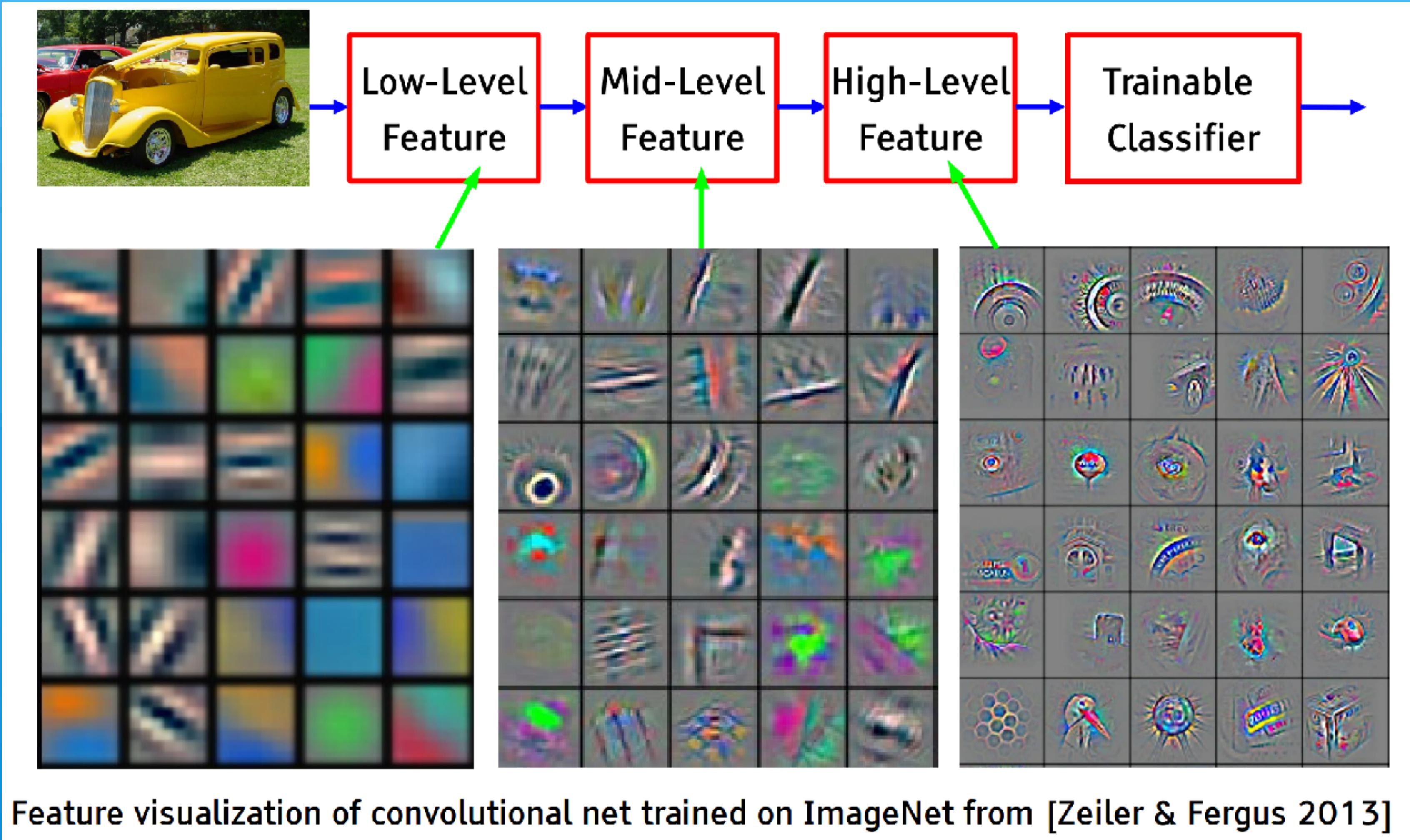


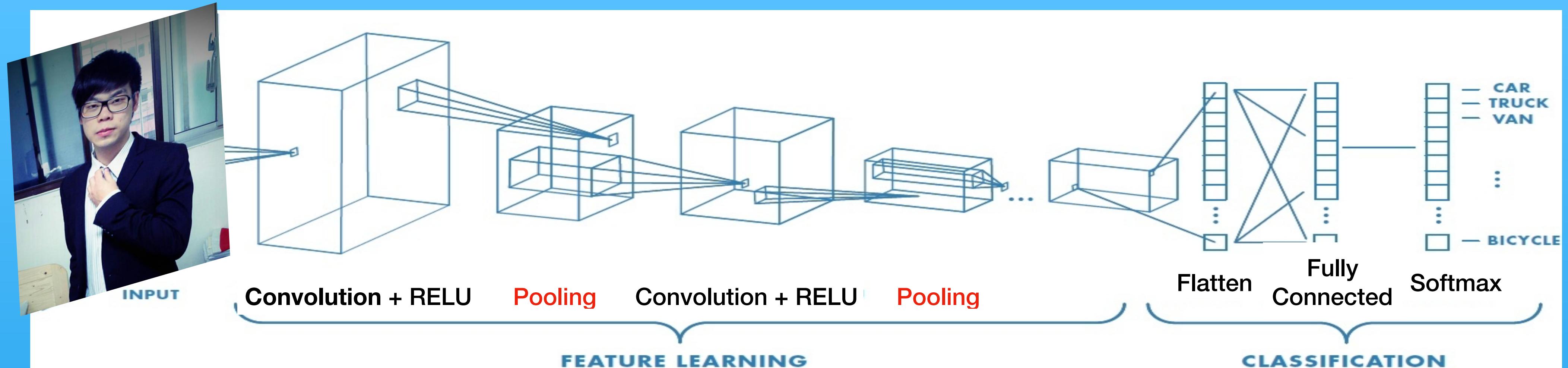
Stride = 魔鬼的步伐

Filter = 濾波器

# 訓練CNN = 訓練Filter

由簡入繁的特徵檢測過程





## Max Pooling

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

convolution  
\*

1	0	-1
1	0	-1
1	0	-1

filter

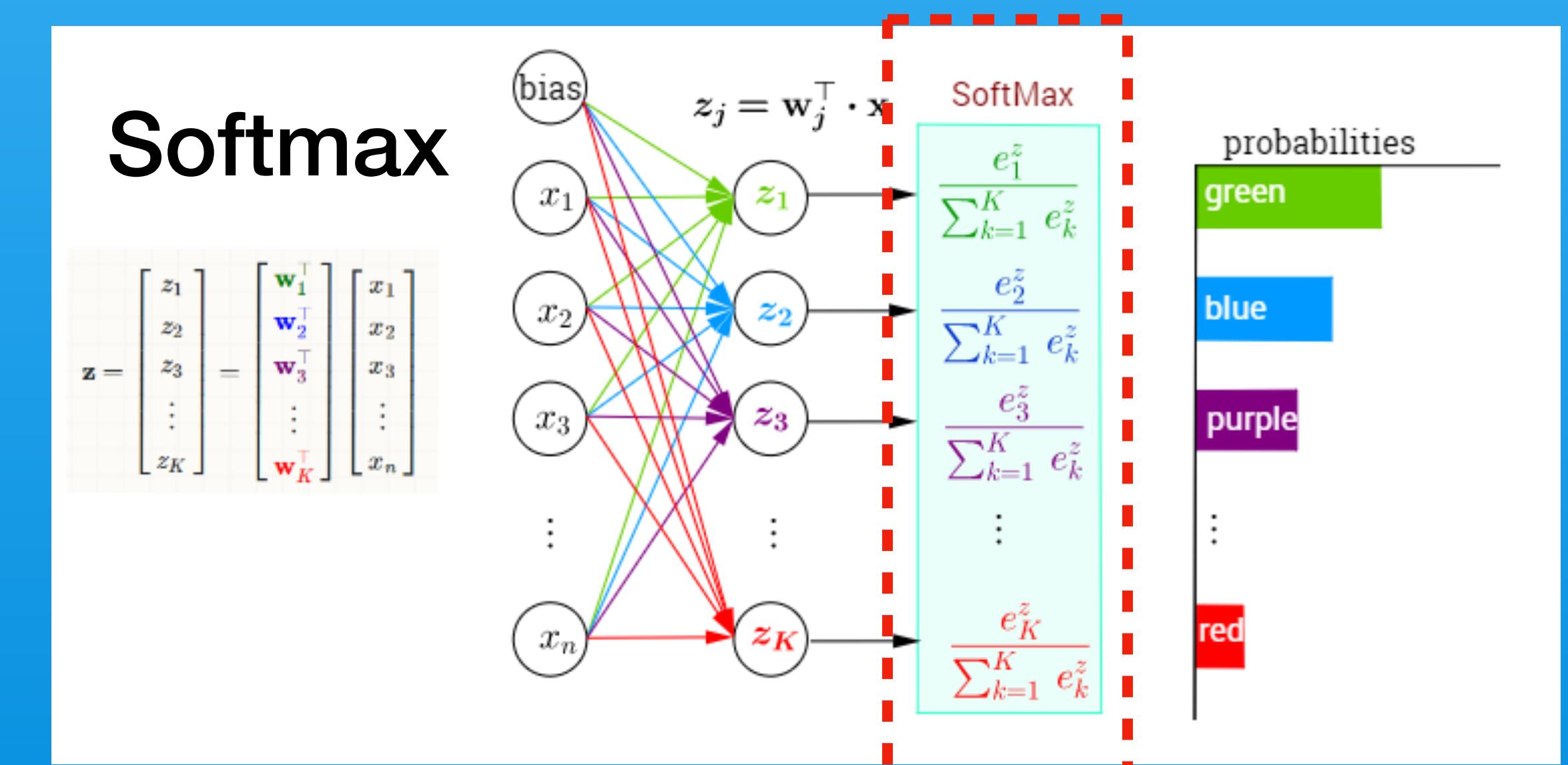
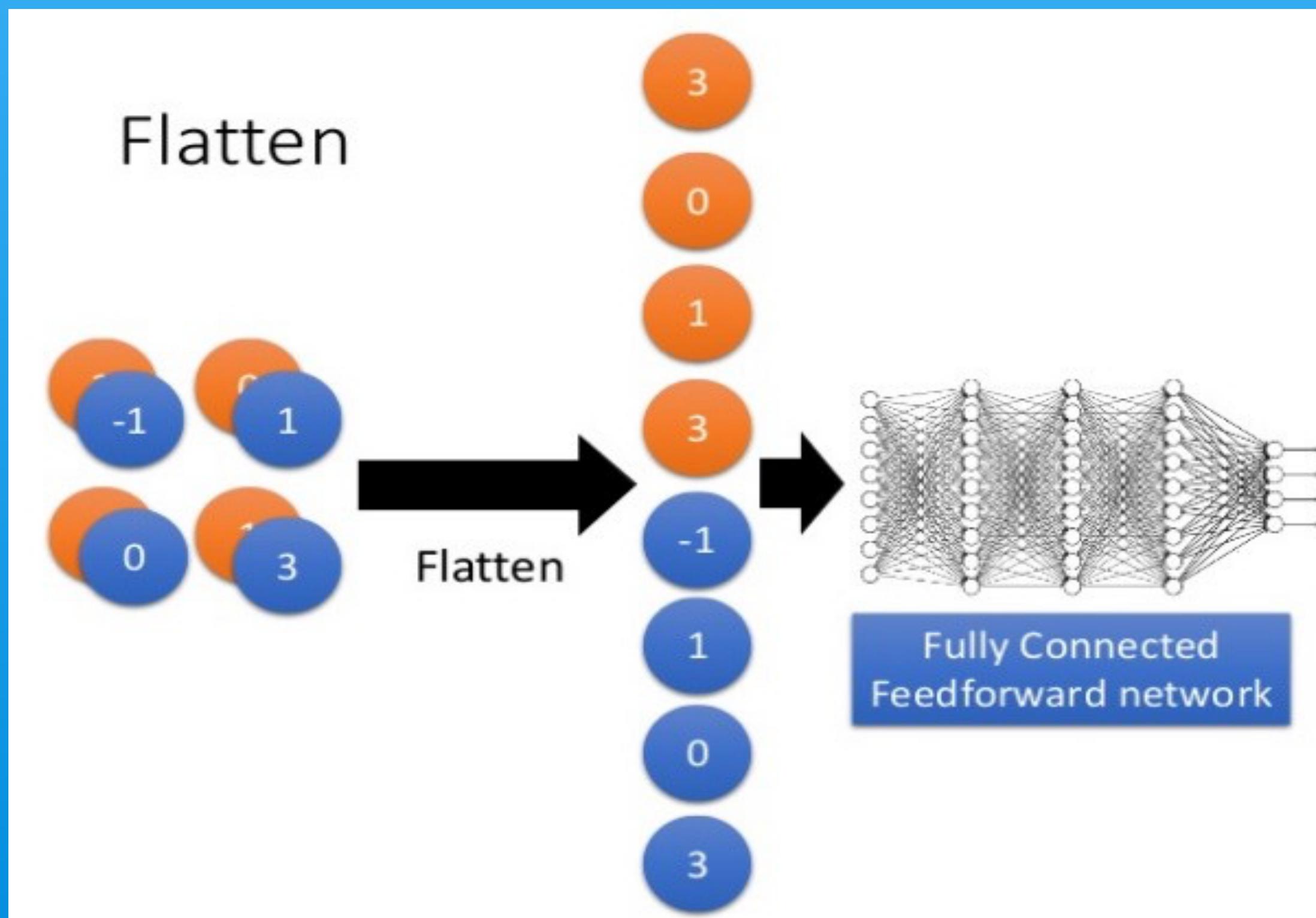
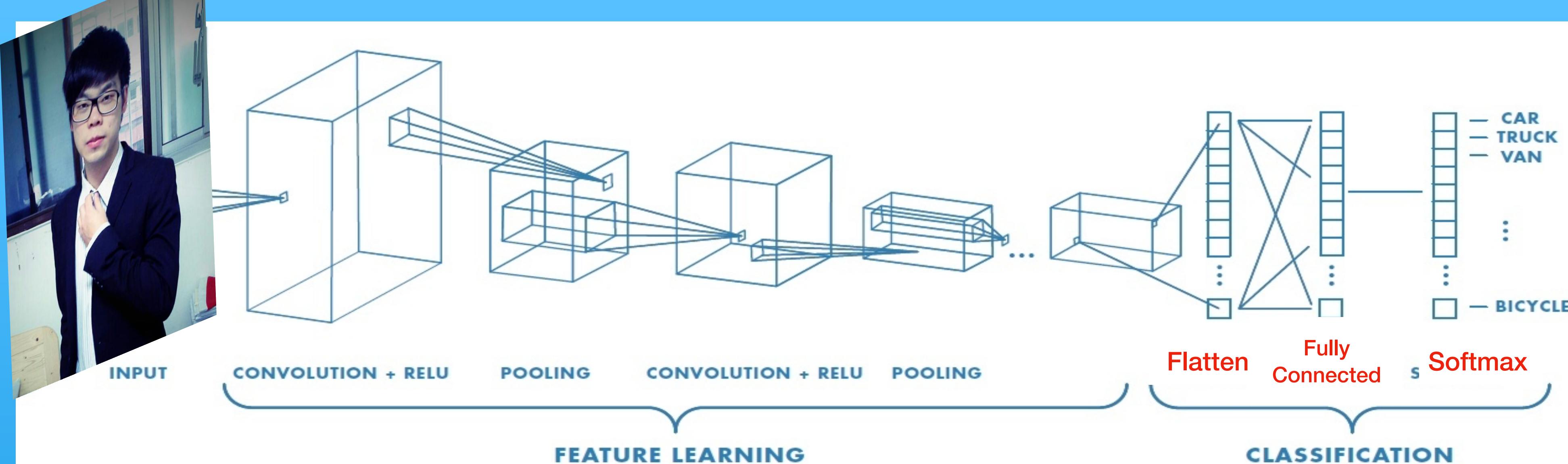
0	30	30	0
0	30	30	0
0	30	30	0
0	30	30	0

=

30	30
30	30

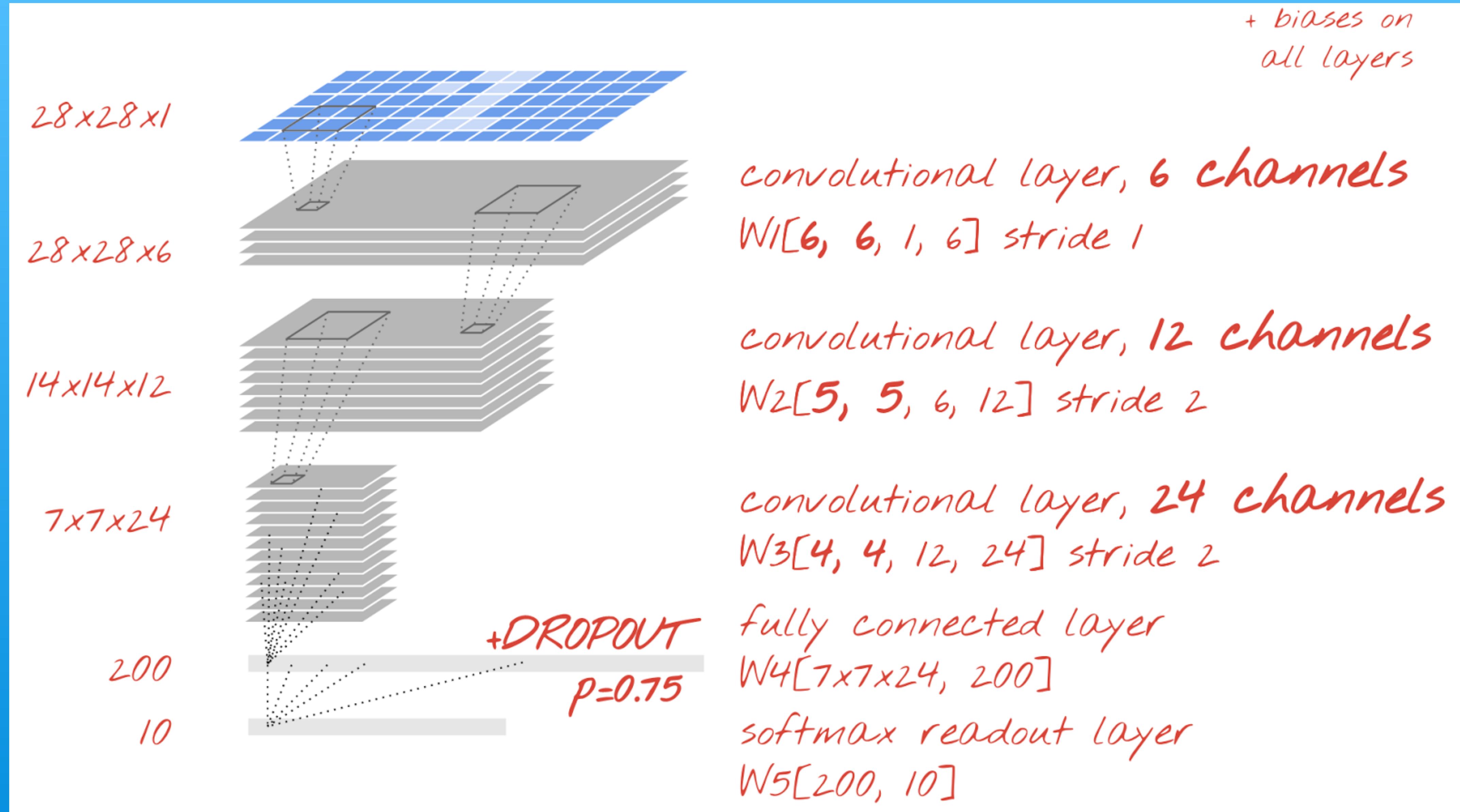
## Average Pooling

15	15
15	15



# Agenda

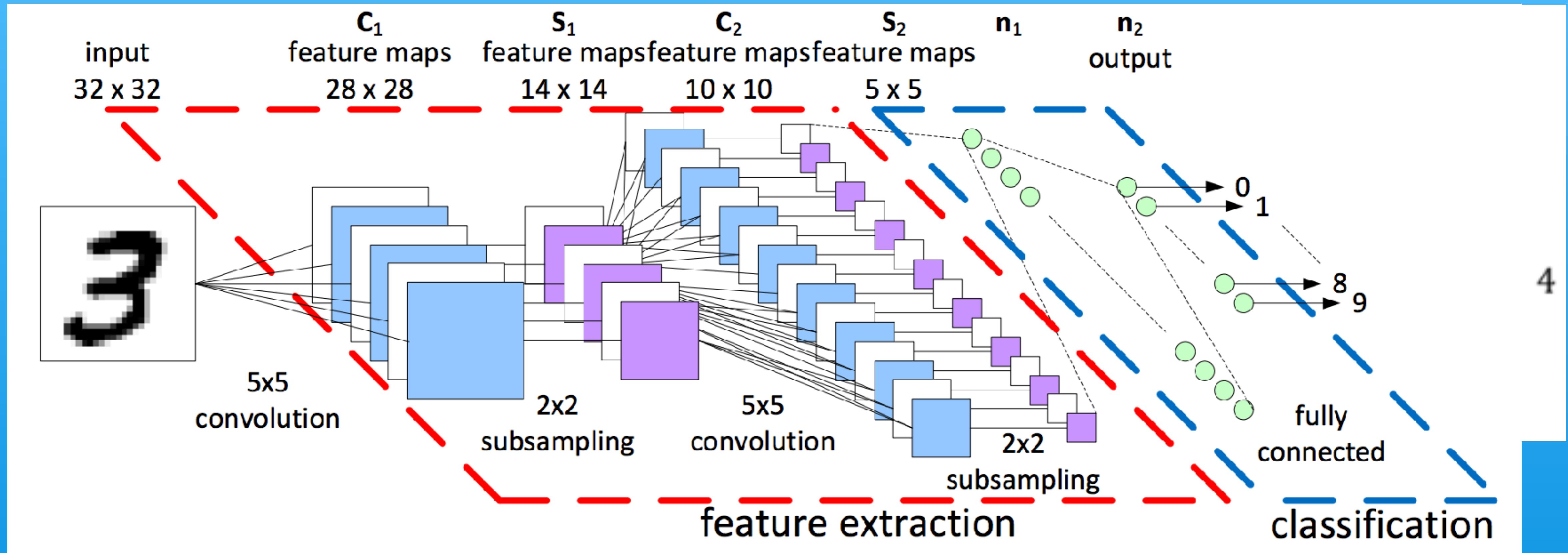
- Object Detection
- 3D CNN - Convolution + Pooling
  - A quick example
  - Why need Convolutional Neural Network(CNN)
  - Applications
- KKT example (17.11)



# Agenda

- Object Detection
- 3D CNN - Convolution + Pooling
  - A quick example
  - **Why need Convolutional Neural Network(CNN)**
  - Applications
- KKT example (17.11)

# Why need CNN ?



1. CNN保留圖像的空間排列
2. 取得局部圖像圖作為輸入特徵

THE CHAINSMOKERS & COLDPLAY

Applications

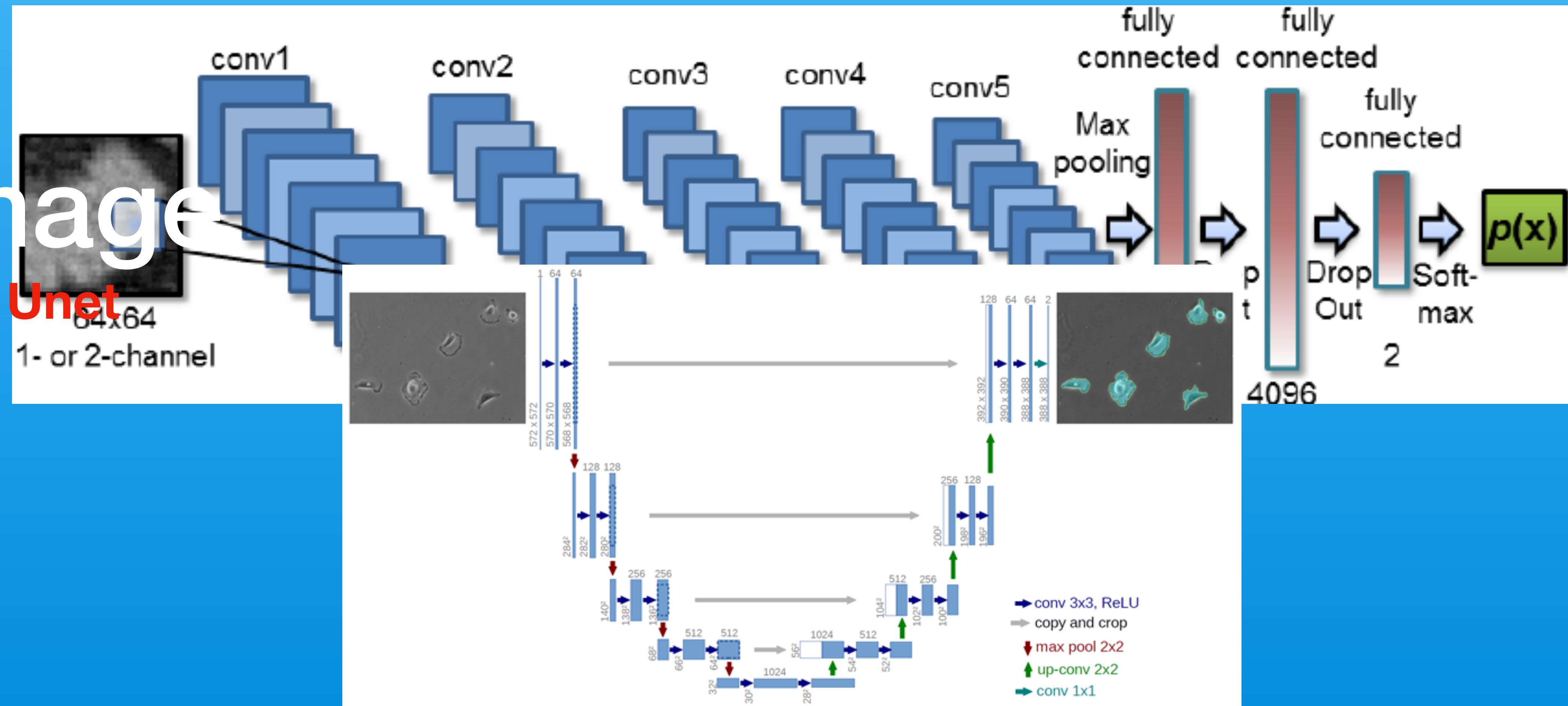
SOMETHING  
JUST  
LIKE THIS



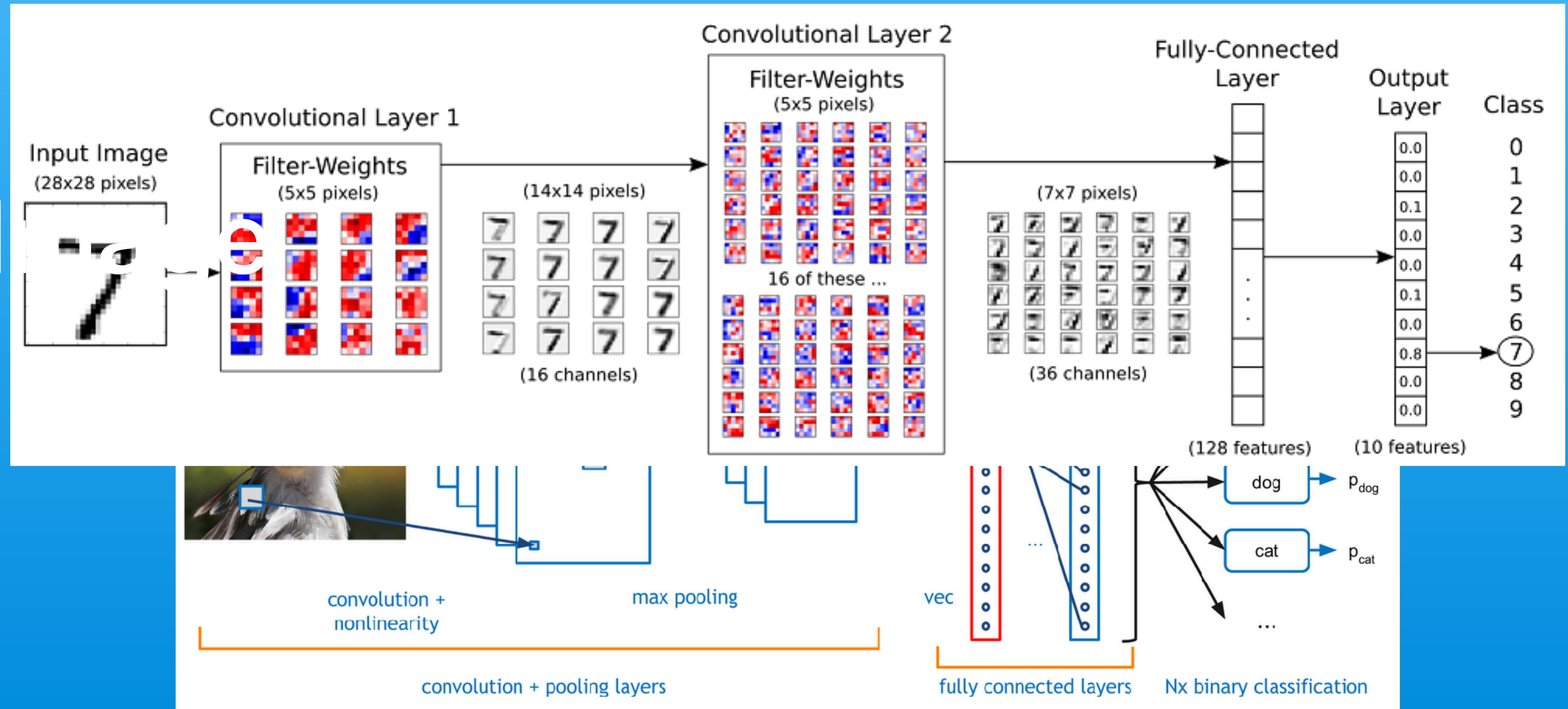
# Image

## Alexnet

Image  
Unet



# Image



Image

Image

Image

Image

Image

Image

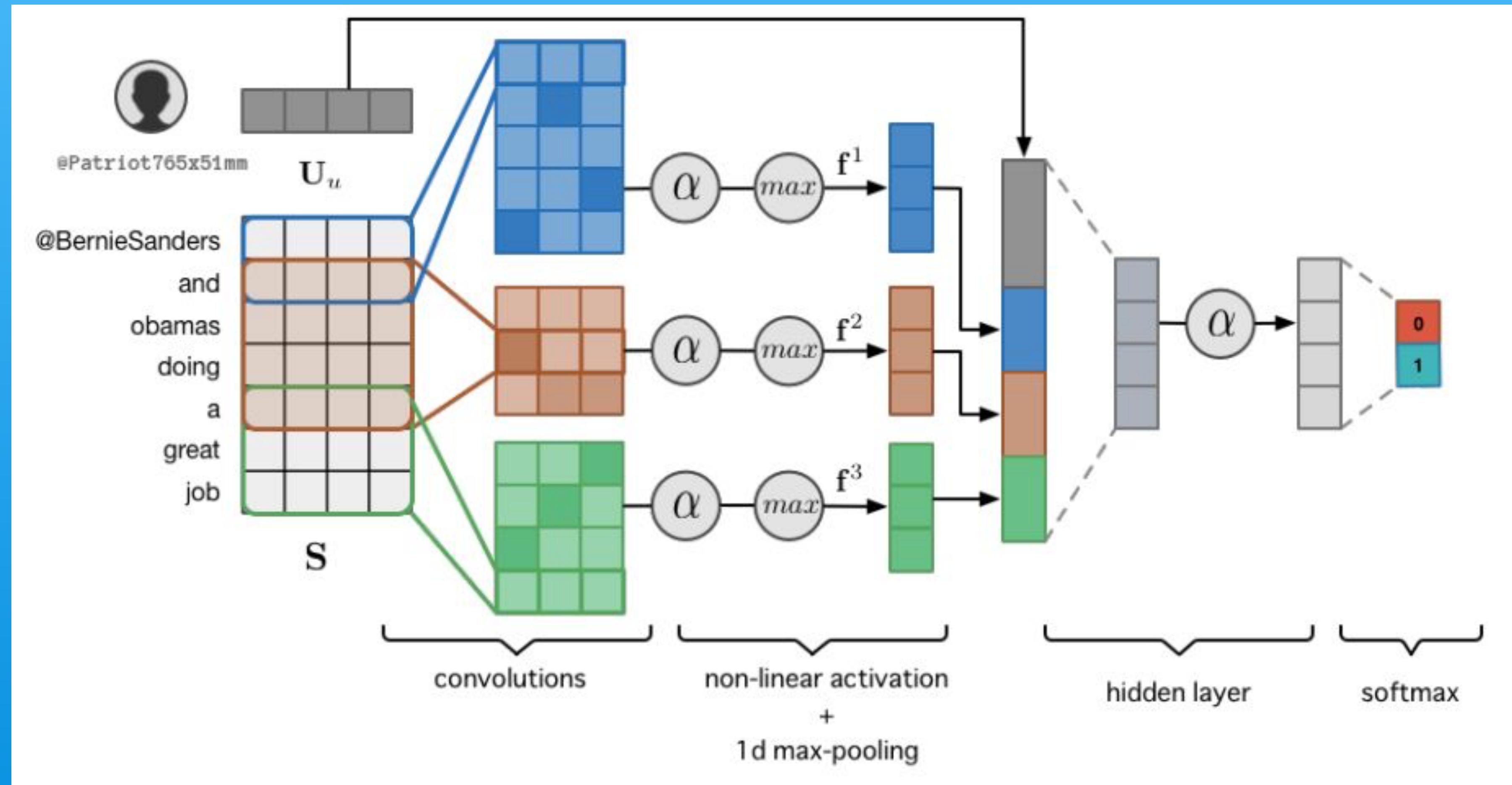
Image

Image

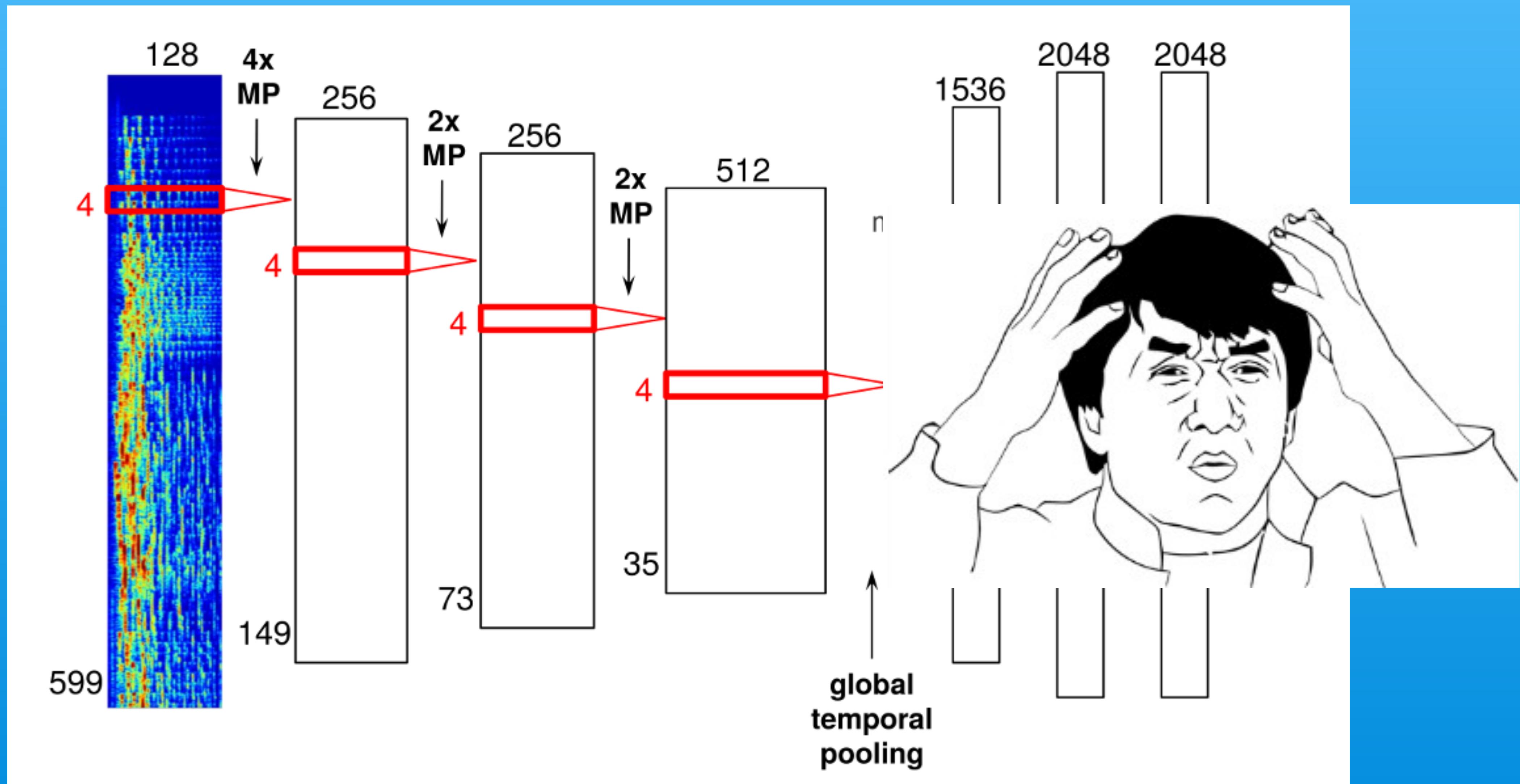


mage

# NLP



# Spotify recommender system



# Agenda

- Object Detection
- 3D CNN - Convolution + Pooling
  - A quick example
  - Why need Convolutional Neural Network(CNN)
  - Applications
- **KKTV example (17.11)**

	<b>user_id</b>	<b>device_id</b>	<b>session_id</b>	<b>title_id</b>	<b>event_time</b>	<b>played_duration</b>	<b>action_trigger</b>	<b>platform</b>	<b>episode_number</b>	<b>series_total_episodes_count</b>	<b>ernet_connection_type</b>	<b>is_trailer</b>
0	0	0	0	0	2017-06-08 14:54:30.325	2	seek	iOS	12	16	wifi	False
1	0	0	0	0	2017-06-08 14:54:34.857	2	seek	iOS	12	16	wifi	False
2	0	0	0	0	2017-06-08 14:54:36.730	1	seek	iOS	12	16	wifi	False
3	0	0	0	0	2017-06-08 14:54:38.976	1	seek	iOS	12	16	wifi	False
4	0	0	0	0	2017-06-08 14:54:40.373	1	seek	iOS	12	16	wifi	False
5	0	0	0	0	2017-06-08 14:54:41.779	1	seek	iOS	12	16	wifi	False
6	0	0	0	0	2017-06-08 14:54:46.615	3	seek	iOS	12	16	wifi	False
7	0	0	0	0	2017-06-08 14:57:26.866	159	pause	iOS	12	16	wifi	False
8	0	0	0	0	2017-06-08 15:18:25.334	1248	pause	iOS	12	16	wifi	False
9	0	0	0	0	2017-06-08 15:19:52.571	3	seek	iOS	12	16	wifi	False
10	0	0	0	0	2017-06-08 15:19:56.522	1	seek	iOS	12	16	wifi	False
11	0	0	0	0	2017-06-08 15:19:59.010	1	seek	iOS	12	16	wifi	False
12	0	0	0	0	2017-06-08 15:20:03.755	3	seek	iOS	12	16	wifi	False
13	0	0	0	0	2017-06-08 15:20:05.473	1	seek	iOS	12	16	wifi	False

# 預測8/14-8/21使用行為

衡量方式 : AUROC

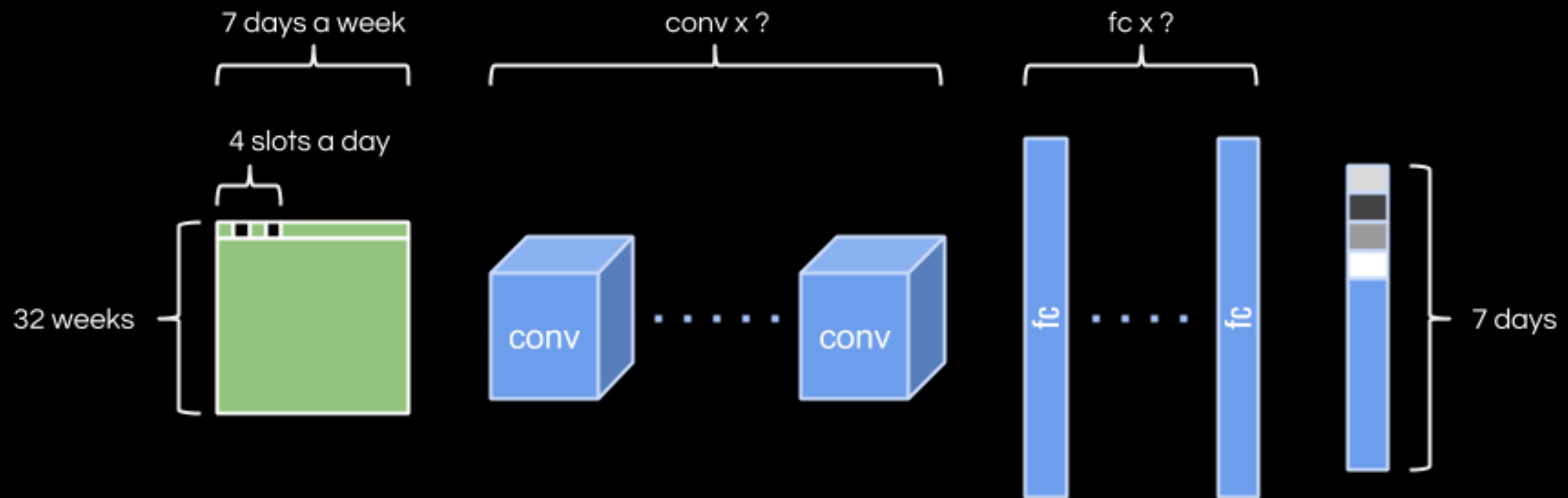
User	Time slot 1	Time slot 2	Time slot 3	Time slot 4	.....	Time slot 28
0	1	1	0	1	.....	1
1	0	0	1	0	.....	0
2	1	0	0	0	.....	0
...	0	1	0	1	.....	1

time\_slot\_0 : 2017/08/14-01:00:00~2017/08/14-09:00:00

time\_slot\_1 : 2017/08/14-09:00:00~2017/08/14-17:00:00

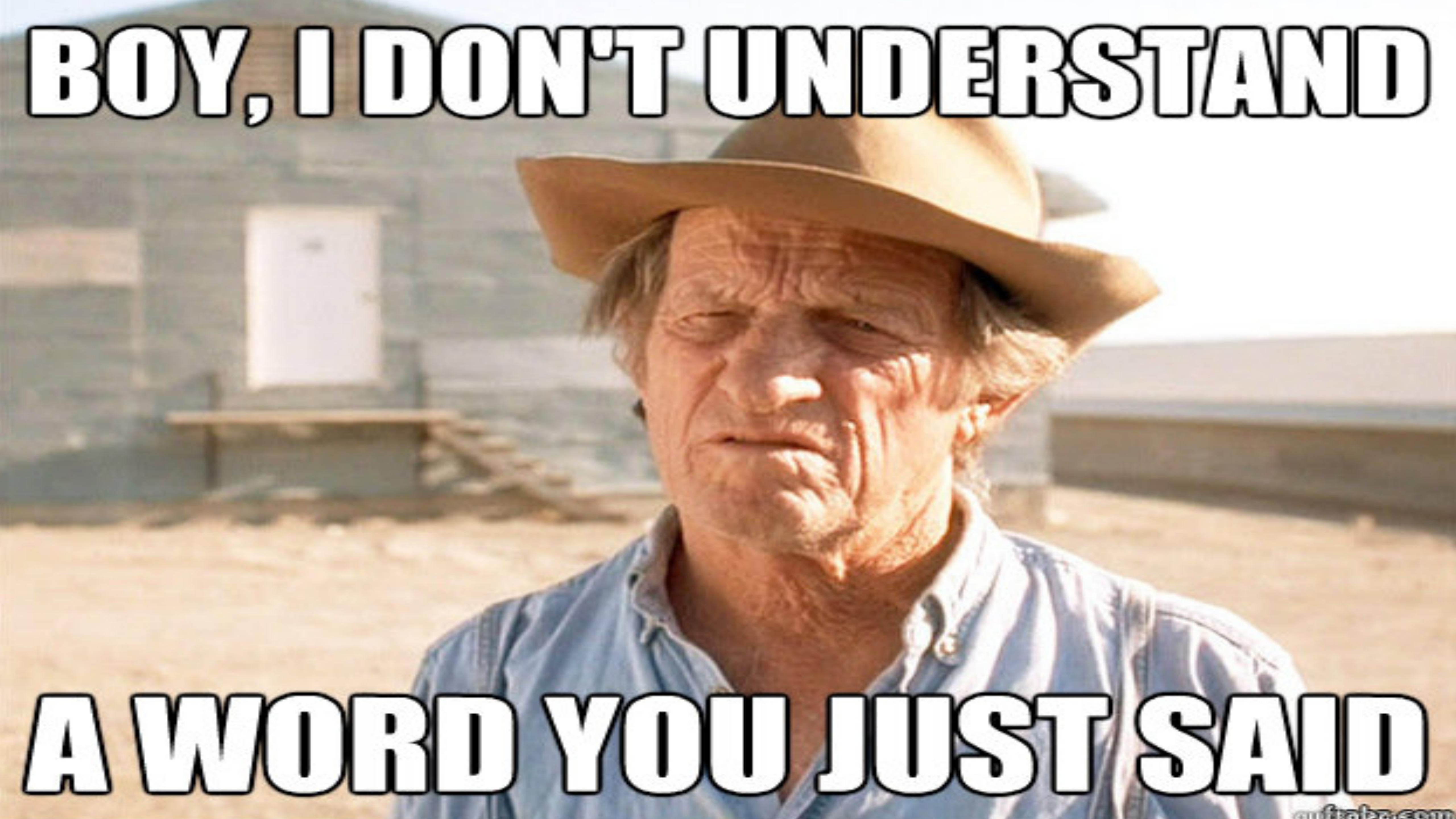
time\_slot\_2 : 2017/08/14-17:00:00~2017/08/14-21:00:00

time\_slot\_3 : 2017/08/14-21:00:00~2017/08/15-01:00:00



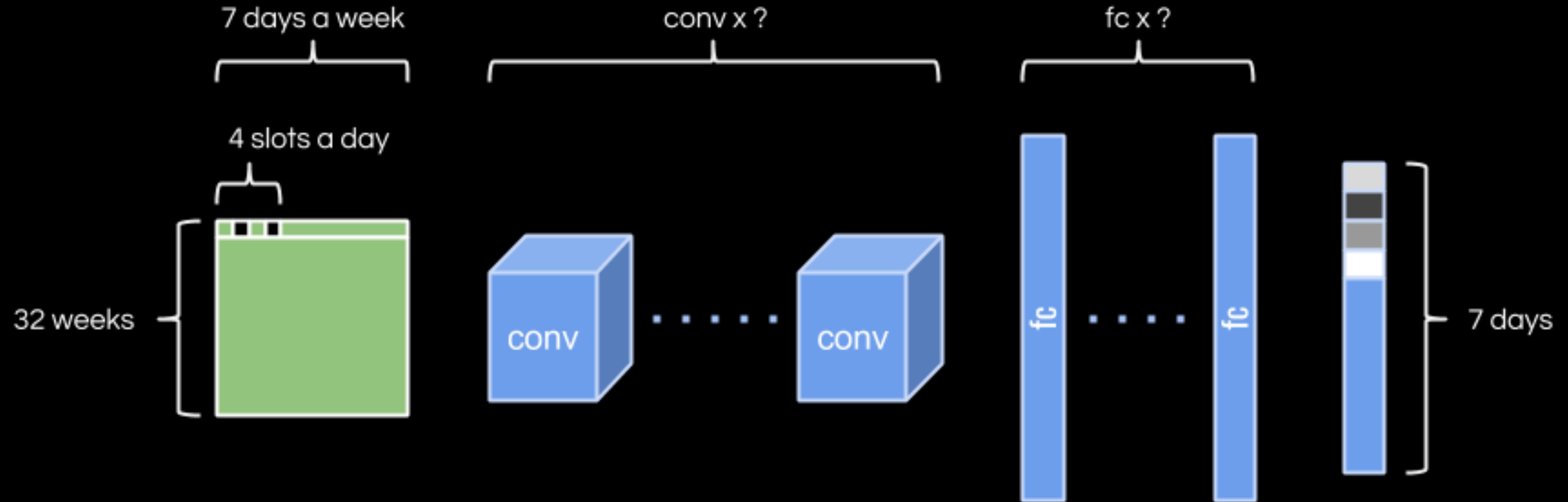
AUROC Top 10 - 0.88

**BOY, I DON'T UNDERSTAND**



**A WORD YOU JUST SAID**

# 思考



32 week : 1/2/01:00- 8/14/01:00

7 days/ week \ 4 slots/ day -> 28 slot/week

32 week : 1/2 0100- 8/14 0100

7 days/ week、4 slots/ day → 28 slot/week

	100	101	102	103	110	111	112	113	120	121	...	3242	3243	3250	3251	3252	3253	3260	3261	3262	3263	time_slot_25	time_slot_26	time_slot_27
user_id	5 rows x 896 columns																							
45707	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0	0
45708	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0	0
45709	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	7	0	0	0	...	0	0
45710	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	...	1	1
45711	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0	0	0	0

1/50 筆資料Training 即有 83% AUROC (Top 20)

# 問題

8000萬車資小票  
爆  
破  
py處理

怎麼辦？



你沒救了 我懶得治療了

憇解

Jupyter \*5(3+2), 分兩次處理儲存成h5  
再開第6隻 jupyter 匯入已完成資料整併

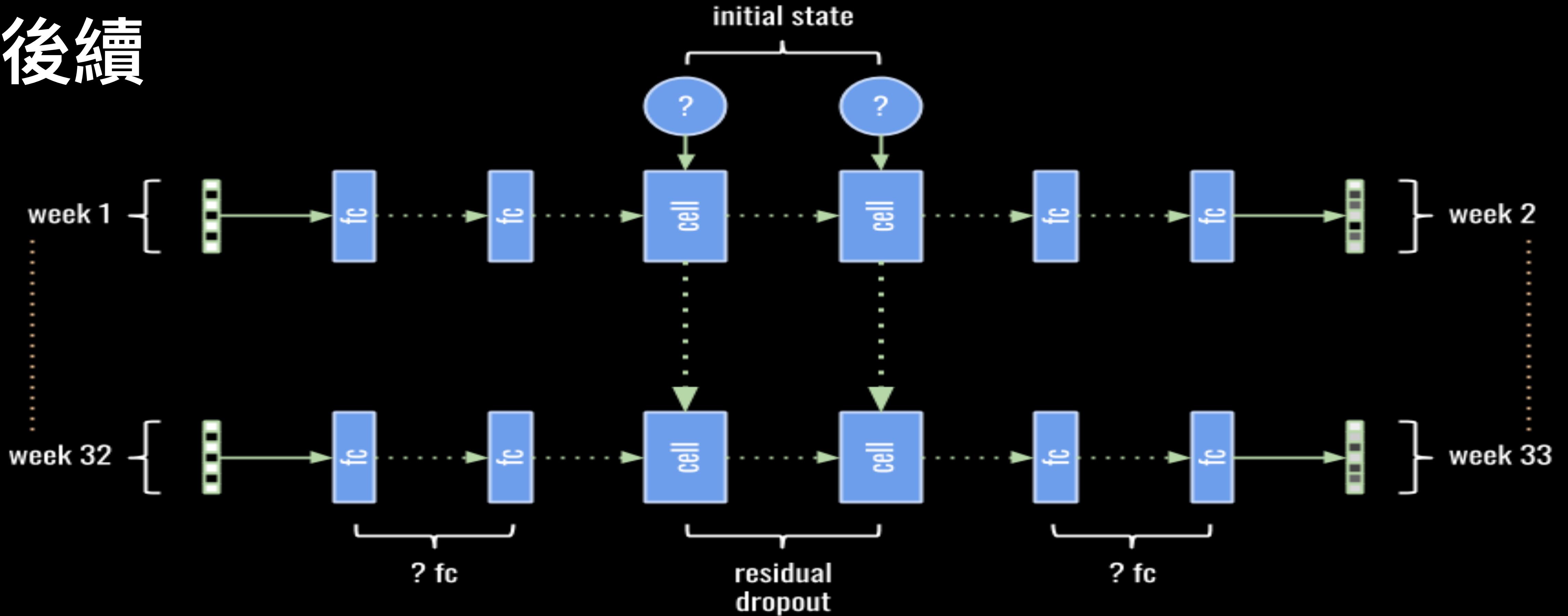
正解

跪著看Roger整理資料

趴著聽Alex闡述分散式結構

Spark hen 重要

# 後續



1. 時間資料  $\rightarrow$  RNN，AUROC 上升

2. 特徵加入行為資料，platform、connection\_type、hot\_titles...

3. 深度學習很強，但第一名用Xgboost

特徵hen重要！

# BIG THINGS



# HAVE SMALL BEGINNINGS

# Reference

- <https://www.coursera.org/learn/convolutional-neural-networks/lecture/ctQZz/convolutions-over-volume>
- <http://benanne.github.io/2014/08/05/spotify-cnns.html>
- <https://medium.com/@kstseng/kktv-data-game-17-11-1st-place-solution-96b3d62c594c>
- <https://hackmd.io/OwJgrAZgRhDMAmBaBYBsiAssDG3EE59sJFVptgJt4Qp8og==>
- <https://medium.com/kkstream/kktv-data-game-17-11-benchmarks-97fffc46fa23>
- <https://github.com/bgg11117/Python/blob/master/kaggle%20practice/Kktv%20CNN%20sample.ipynb>
- <https://github.com/b96705008/kktv>