



Introduction – COMP4423 Computer Vision

Xiaoyong Wei (魏驍勇)

x1wei@polyu.edu.hk



The Instructor

- > Visiting Professor, Dept. of Comp., PolyU, HK
- > Visiting Professor, Tampere University, Finland
- > Adjunct Professor, Peng Cheng Lab
- > Professor, Department Head, Sichuan University
- > Postdoctoral Fellow, UC Berkeley
- > PhD, City University of Hong Kong
- > Founding Member of VIREO
- > Senior Member of IEEE



Department of Computing
電子計算學系



香港城市大學
City University
of Hong Kong



VIREO



IEEE



NEWS WEBSITE OF THE YEAR
The Telegraph

Subscribe now
Free for one month

Log in

Coronavirus News Politics Sport Business Money Opinion Tech Life Style Travel Culture

See all News

News
Chinese lecturer to use facial-recognition technology to check boredom levels among his students

f share t Share S Save

digitaltrends

EMERGING TECH

A professor in China is using facial recognition to gauge student interest

By Luke Dormehl
September 13, 2016

就问你怕不怕！川大老师研发出无人机，
让新闻离你更近

2018-05-25 07:30:40 来源：华西都市报

关注新华网
微信
微博
Qzone

The Telegraph Sep. 13, 2016

Digital Trends Sep. 13, 2016

XinHua News May 25, 2018

央視新聞
<https://tw.weibo.com/cctvxinwen>

2013年1月28日 19:04

【川大教授又來了！逆天面部識別“刷臉”點名】上大學，不免會翹幾節課，同宿舍的哥們姐們也總會幫忙喊到，但是，遇到“刷臉”簽到，只能愛莫能助了。近日，四川大學計算機學院副教授魏驍勇發明了一套面部識別系統，上課從不點名，直接“刷臉”。學生們紛紛感慨，“技術宅”的老師惹不起！

澎湃新闻

精选 视频 时事 财经 澎湃号 思想 生活 问题 订阅 下载APP

用改装无人机“点名”的川大教授：表情识别技术已申专利

四川大学 澄清公众号 2018-05-27 15:02

QUARTZ

ALL EYES ON YOU

A Chinese professor is using facial recognition to gauge how bored his students are

CCTV Jan. 13, 2013

The Paper May 27, 2018

QUARTZ Sep. 13, 2016

中青網
中国新闻网
105万 文章 631亿 总阅读
查看TA的文章>
评论 0
分享 微信分享 新浪微博 QQ空间 复制链接

川大魏晓勇：“无人机点名”不是为了点名
2018-05-26 13:56

演讲中的魏晓勇。四川大学供图

2018年06月13日 星期三 收藏本站 中青在线 www.cyol.com
往期回顾 新闻列表 返回目录 < 上一期 下一期 > 中青报系
传奇 让科学飞
中国青年报·中青在线记者 王景烁 来源：中国青年报 (2018年06月13日 10 版)
2018年06月13日 10 版

深视新闻 SZTVNEWS “安心”抗疫“圳”在行动 深圳多举措筑牢抗疫心理防线
18°~23° Shenzhen takes measures for psychological health of medical workers and residents
深圳大学

China News May 26, 2018

China Youth Daily Jun. 13, 2018

Shenzhen TV Mar. 2, 2020

中国青年网 Youth.cn 新闻 重页 >> 社会 >> 正文
“劈砖教授”发明“刷女”软件 获取信息助联谊
发稿时间：2013-06-04 09:03:00 来源：华西都市报 中国青年网

中国新闻网 首页 → 新闻中心 → 文化新闻
字号：大 中 小
“劈砖教授”获菠萝科学奖 劈砖功夫完胜诺奖得主
2013年04月09日 13:28 来源：天府早报 参与互动(0)

我们都应该知道，喜欢并传播科学的人都有三个阶段：见自己，见天地，见众生。
魏老师已经到了见众生的层次，劈砖的境界在于念念不忘必有回响。魏老师教会我们如何科学地成为一代宗师。”——魏晓勇获奖的颁奖词

网易首页 > 教育频道 > 考研 > 最新资讯 > 正文
“劈砖教授”走红网络 大学不应只有高精尖
2012-10-25 17:39:02 来源：新华网

走红后首堂课，魏晓勇特别声明：教学还要继续，自己讲的不是功夫。用“劈砖”讲
务实，大学不应只有高、精、尖。

10月24日，因为劈砖太多，魏晓勇右手已经红肿。

China Youth Net Jun. 04, 2013

“Pineapple Science Prize”
2013

Awarded by Nobel Prize
laureate

Office Hours

- > Every Wednesday before the lectures 13:30 - 15:30 at PQ829
- > Appointments via x1wei@polyu.edu.hk or by phone 2766-7258

Languages

- >English (Professional)
- >Mandarin (Native)
- >Cantonese (Pretend to be native)

Teaching Assistants

- > Xulu Zhang, compxulu.zhang@connect.polyu.hk
- > Changmeng Zheng, csczheng@comp.polyu.edu.hk
- > Bo Liu, bokelvin.liu@connect.polyu.hk

What is Computer Vision?

I have no idea ...









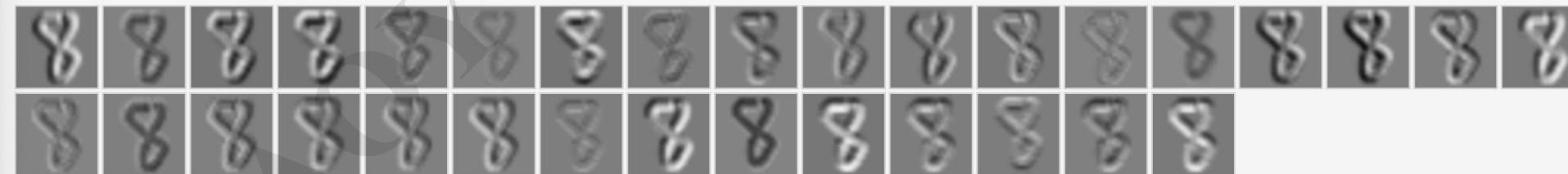
But let's play with it!

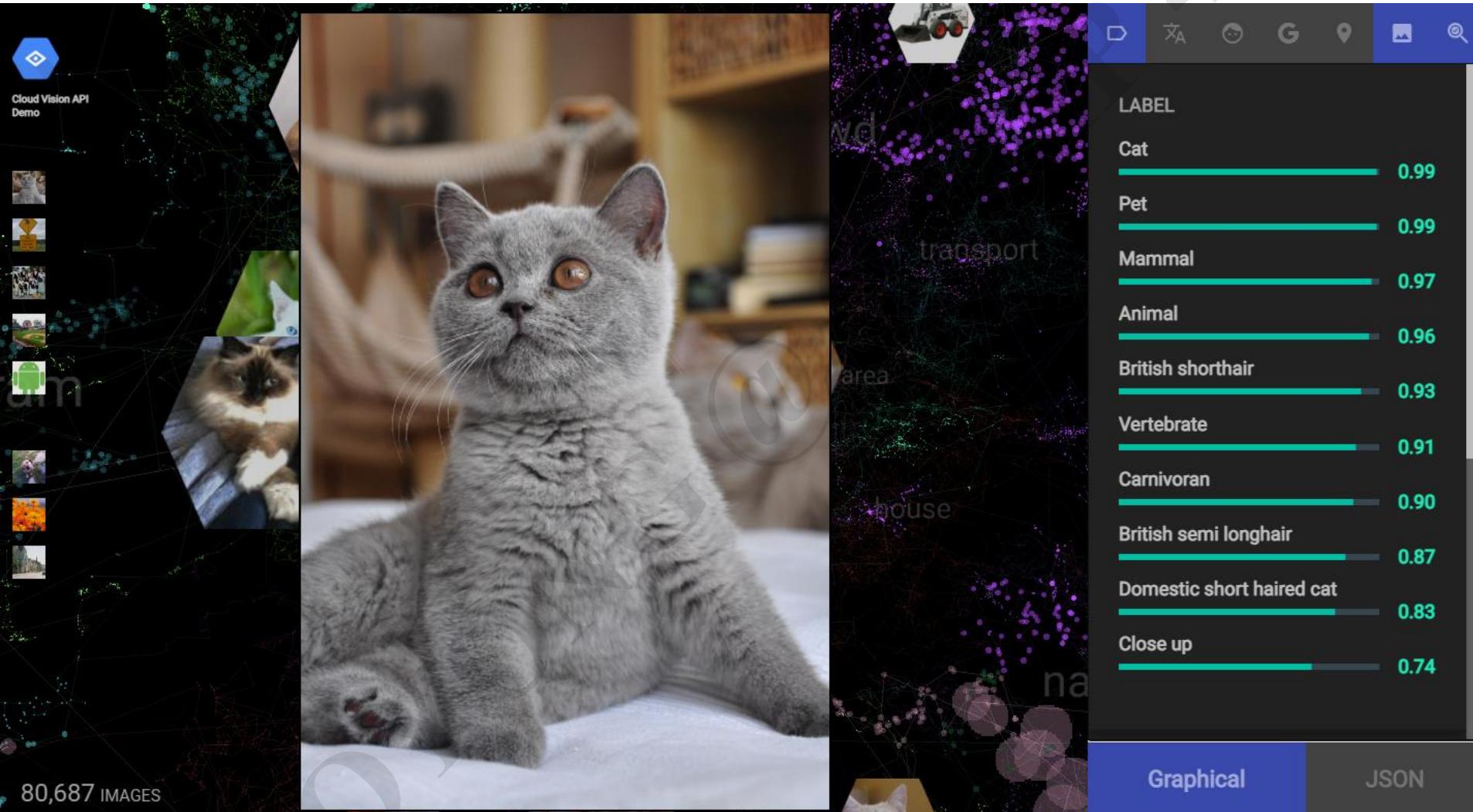
Basic Convnet for MNIST



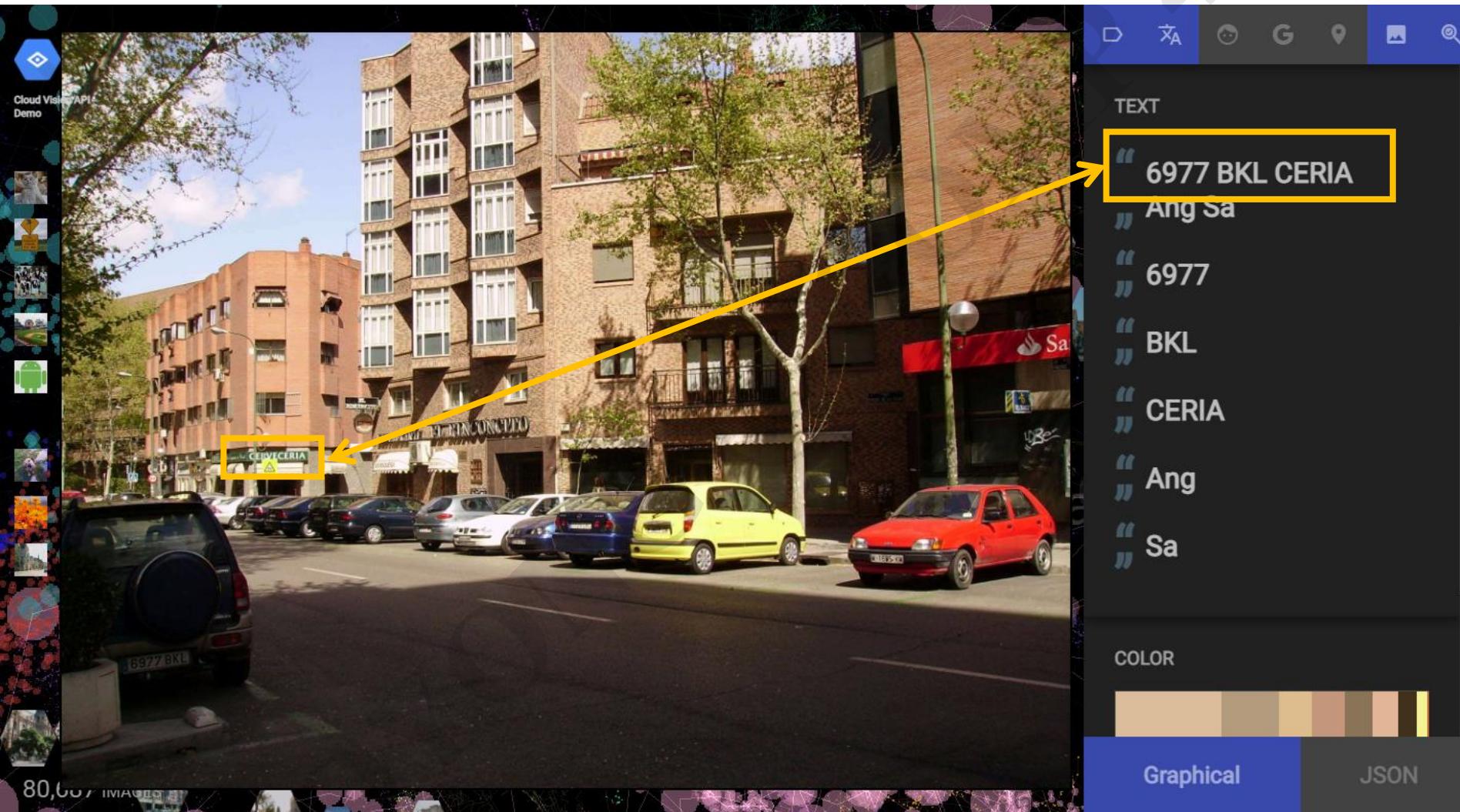
Conv2D

32 3x3 filters, padding valid, 1x1 strides





http://vision-explorer.reactive.ai/#/galaxy?_k=8d4jqd



<http://vision-explorer.reactive.ai/#/galaxy? k=8d4jqd>







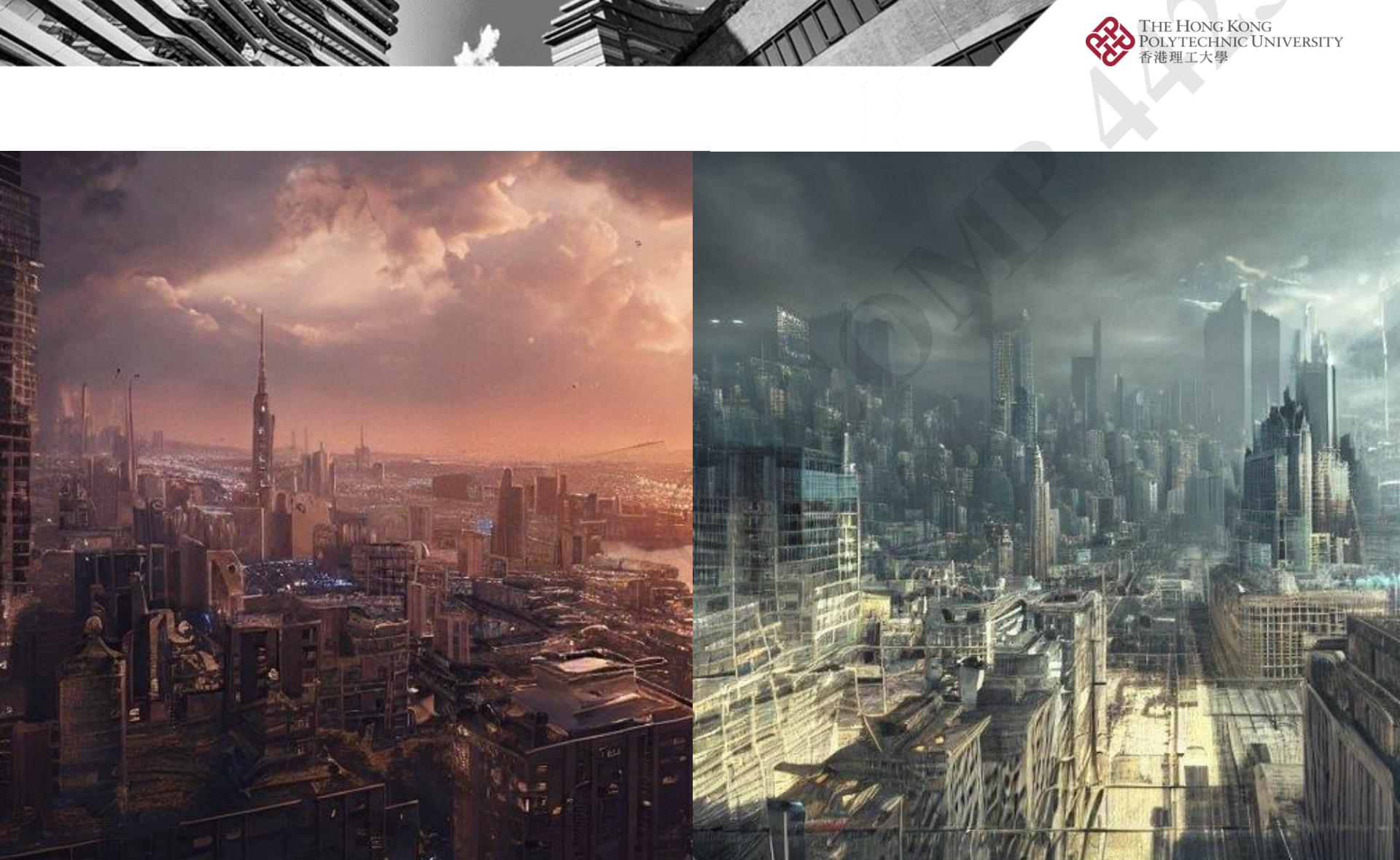
<https://thispersondoesnotexist.com/>

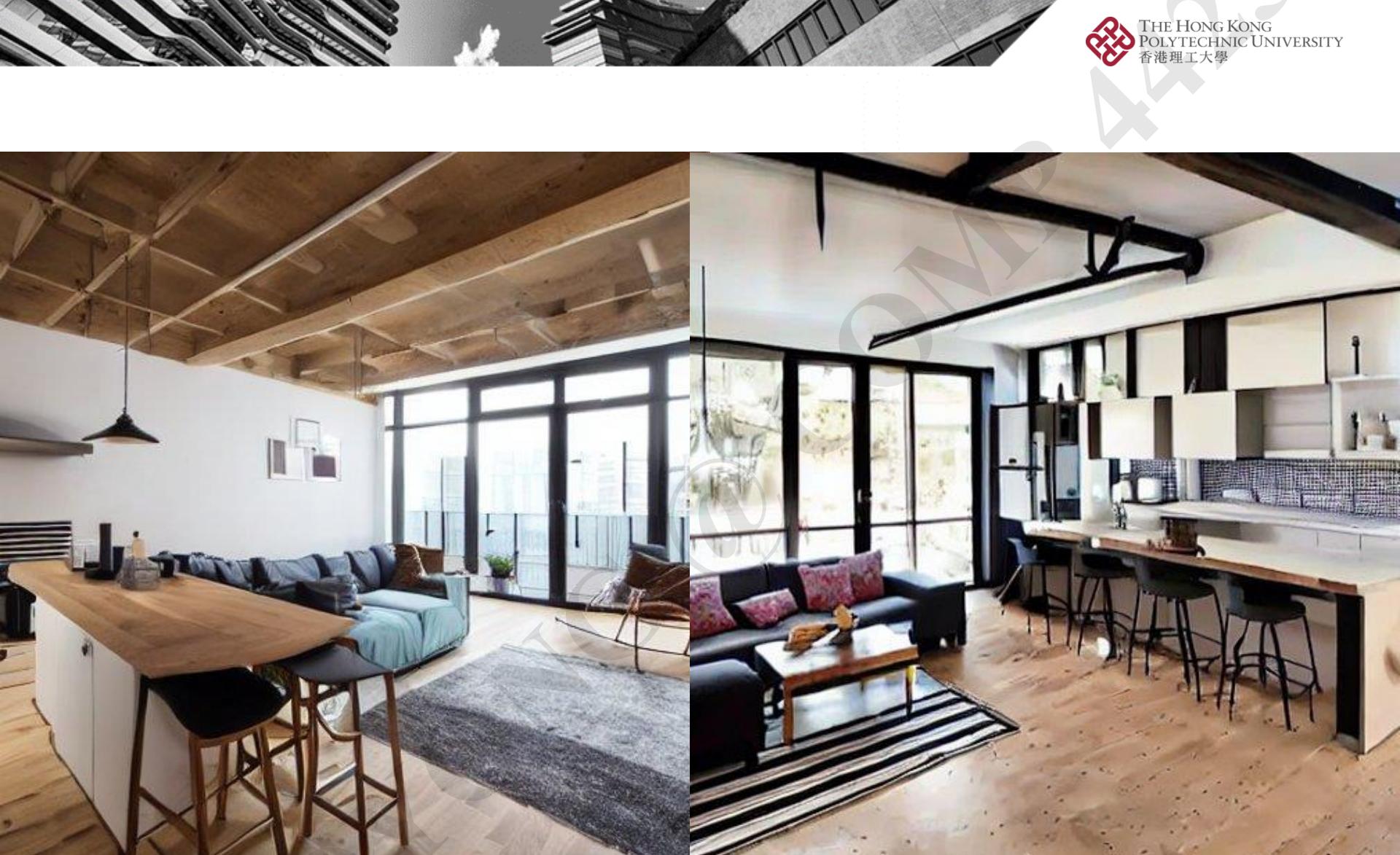


<https://carpedm20.github.io/faces/>











<https://stablediffusionweb.com/>

Try It

interior design, frank lloyd wright house cave with forest canopy, dark wood, streaks of light, light fog, living room :: bubbletech –test –ar 9:16

cyberpunk techwear streetwear look and clothes, we can see them from feet to head, highly detailed and intricate, golden ratio, beautiful bright colors, hypermaximalist, futuristic, cyberpunk setting, luxury, elite, cinematic, techwear fashion, Errolson Hugh, Sacai, Nike ACG, Yohji Yamamoto, Y3, ACRNYM, matte painting –w 2176 – h 3840 –iw 1

woman, warrior, detailed illustration, digital art, overdetailed art, concept art, full character, character concept, long hair, full body shot, highly saturated colors, fantasy character, detailed illustration, hd, 4k, digital art, overdetailed art, concept art, Dan Mumford, Krzysztof Maziarz, trending on artstation

a full page design of spaceship engine, black and bronze paper, intricate, highly detailed, epic, infographic, marginalia –ar 9:16

<https://stablediffusionweb.com/>

AI Art at Christie's Sells for \$432,500



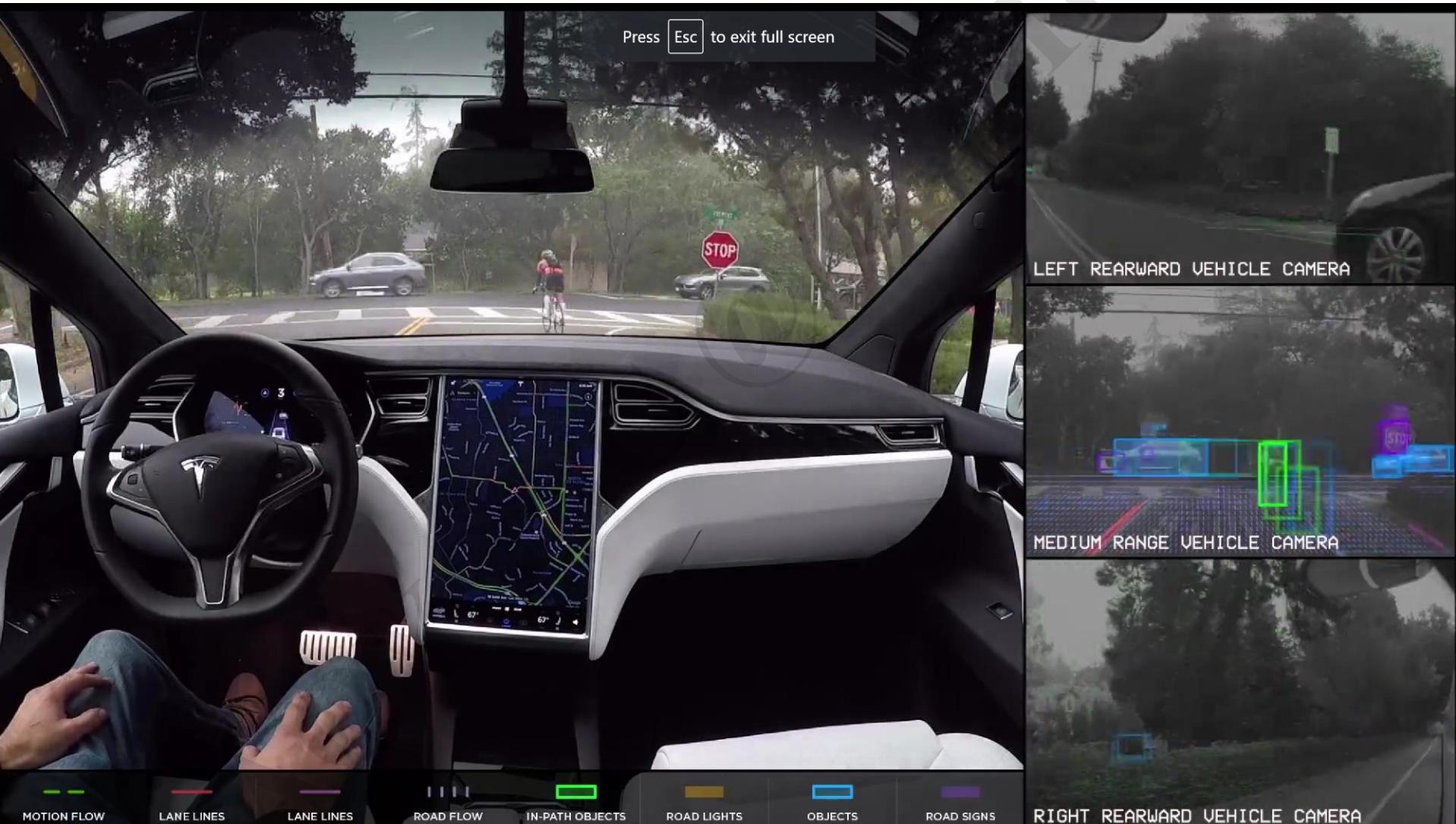
“Edmond de Belamy, from La Famille de Belamy,” by the French art collective Obvious, was sold on Thursday at Christie’s New York. Christie’s

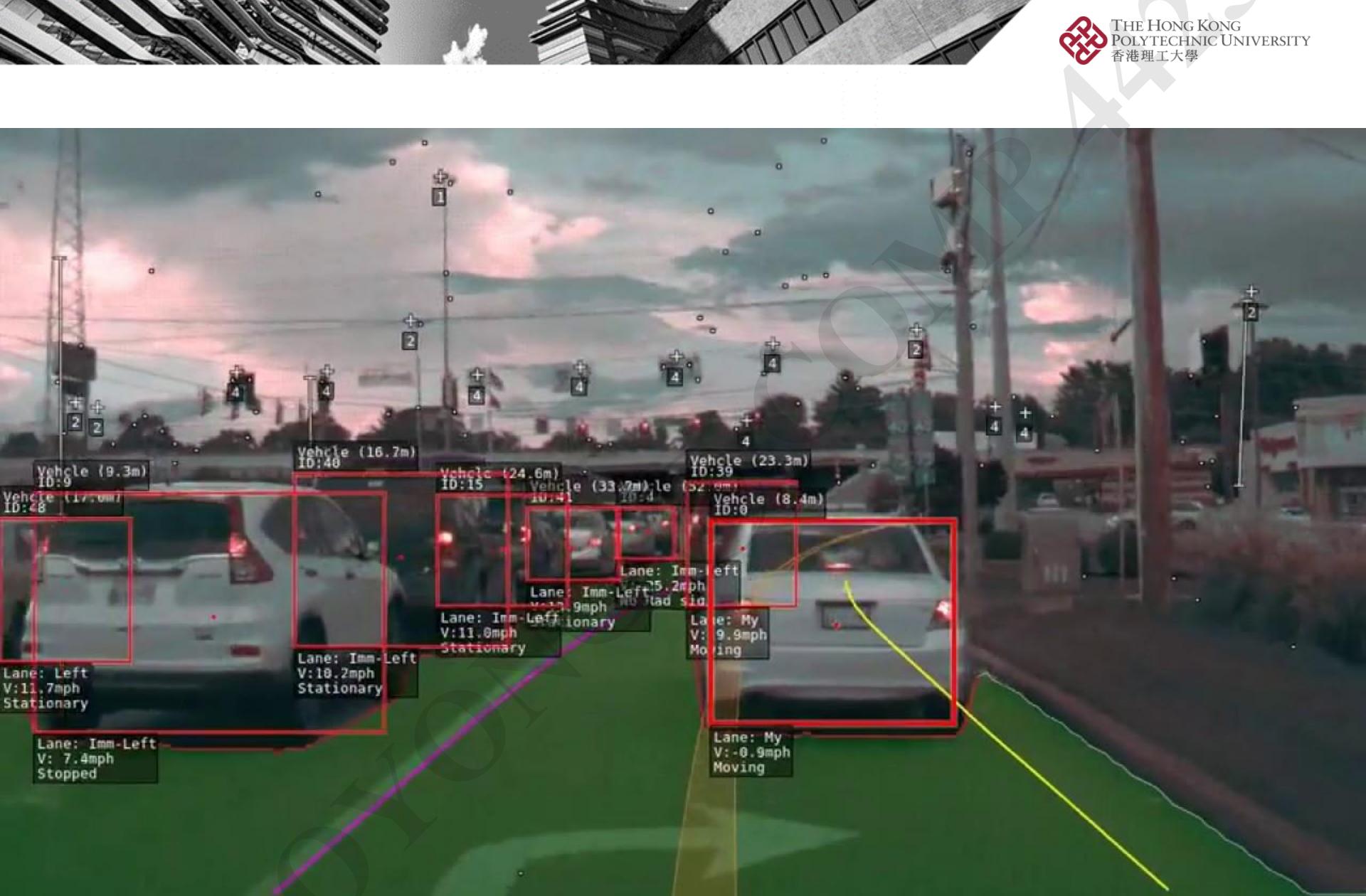
<https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html>

Applications

<https://www.tesla.com/autopilot>

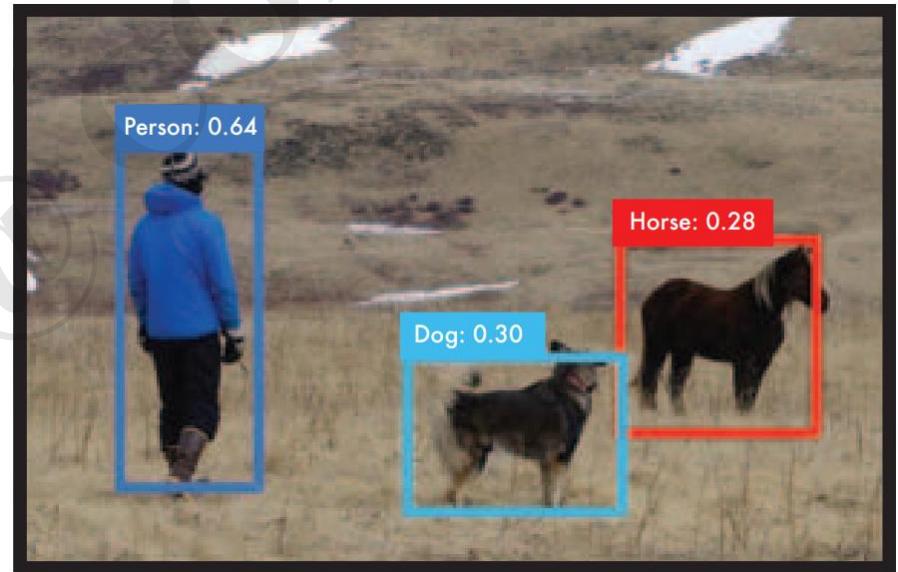
https://en.wikipedia.org/wiki/Self-driving_car





Applications

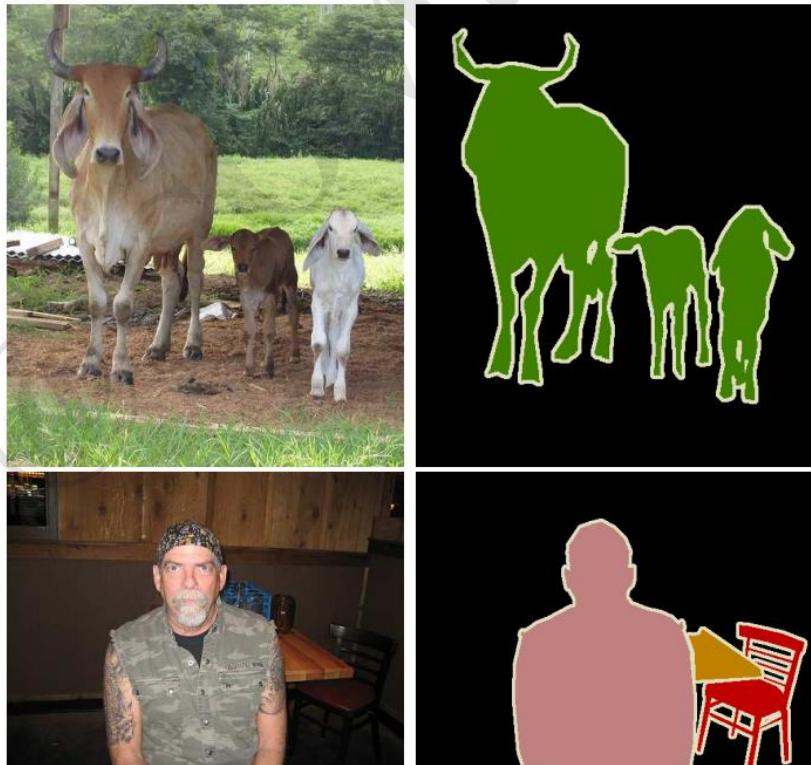
The task of **Object Detection** is to find out all the objects of interest in the image and determine their categories and positions. Object detection is always the most challenging problem in the field of computer vision because of the different appearance, shape and posture of objects and the interference of illumination and occlusion.



Redmon J , Divvala S , Girshick R , et al. You Only Look Once: Unified, Real-Time Object Detection[C]// Computer Vision & Pattern Recognition. IEEE, 2016.

Applications

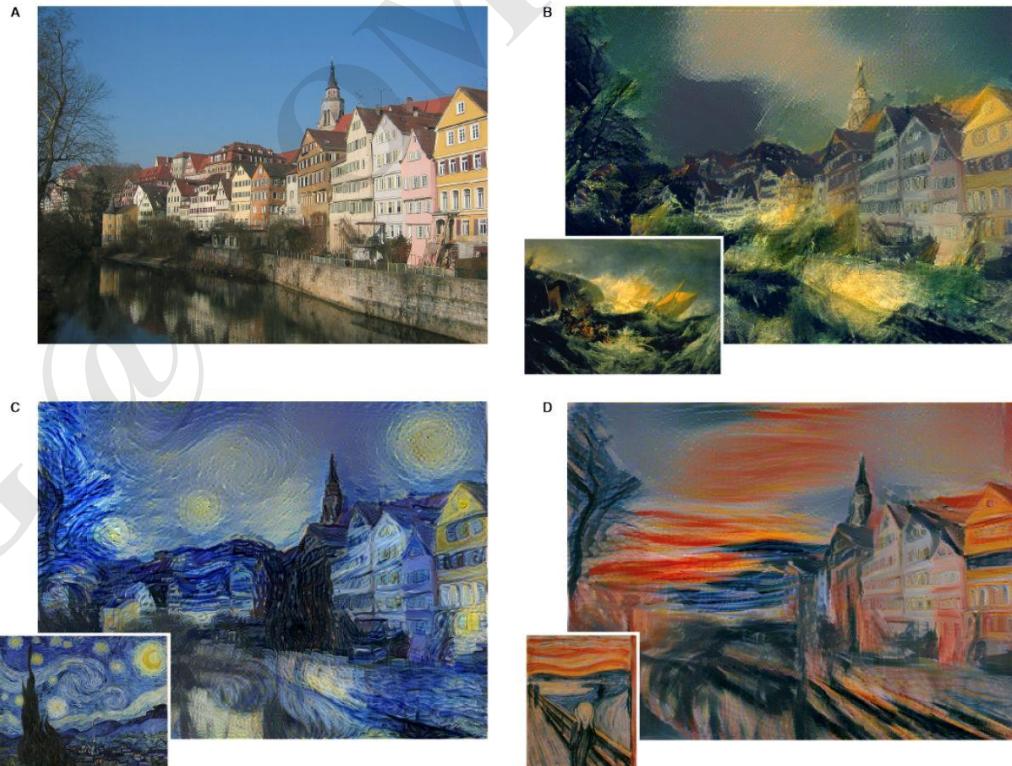
Semantic Segmentation, or image segmentation, is the task of clustering parts of an image together which belong to the same object.



M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. IJCV, 88(2):303–338, 2010.

Applications

Style Transfer refers to applying the style of one domain or several images to other domains or images. For example, abstract styles are applied to realistic images.



Gatys L A , Ecker A S , Bethge M . A Neural Algorithm of Artistic Style[J]. Journal of Vision, 2015.



Driving



Surveillance



Medicine



Services



Smart Building



Entertainment

ENGINE - 96%

FORKLIFT - 98%

FORKLIFT - 95%

TOOLS - 97%

TOOLS - 93%
GEARBOX - 99%

TOOLS - 94%

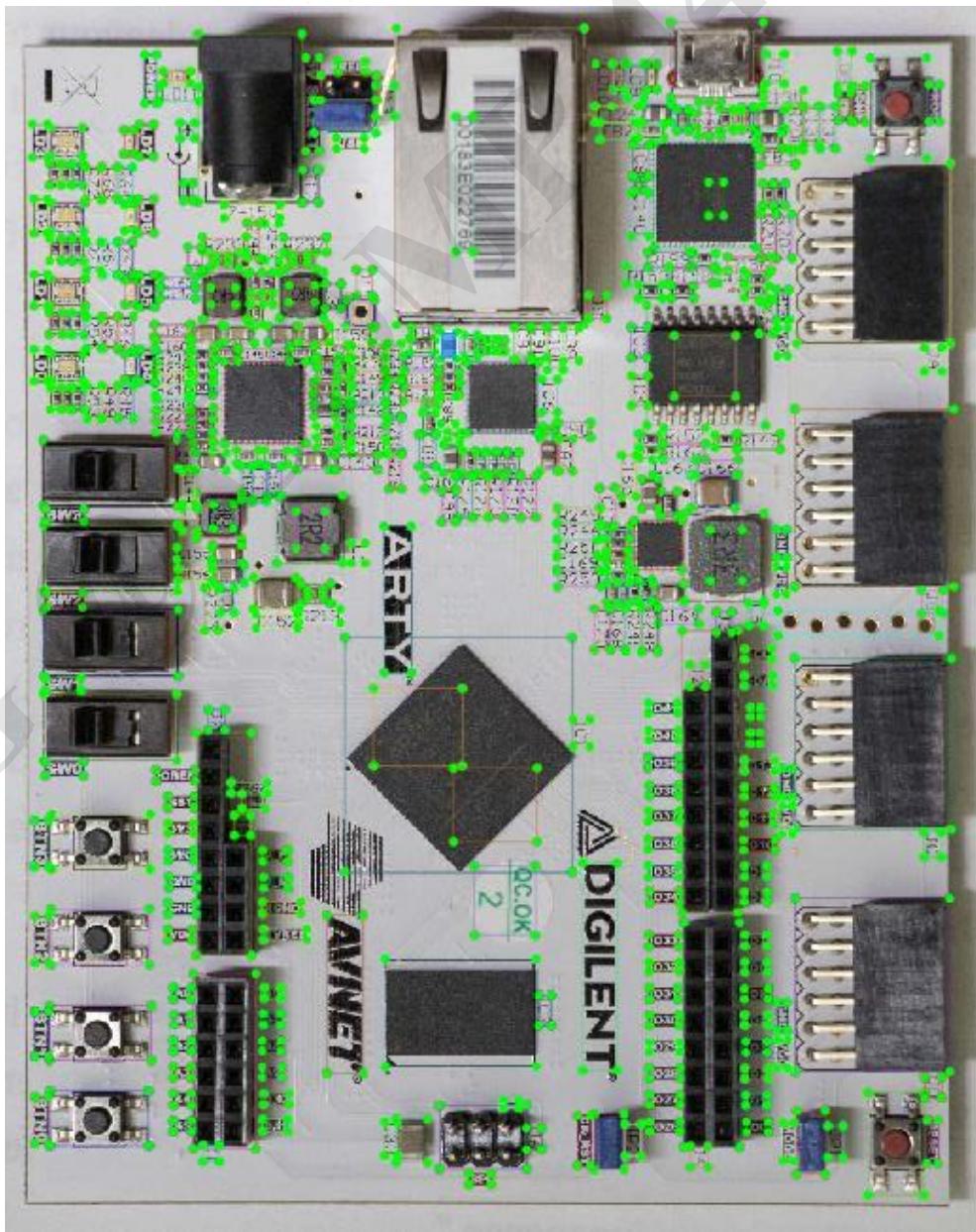
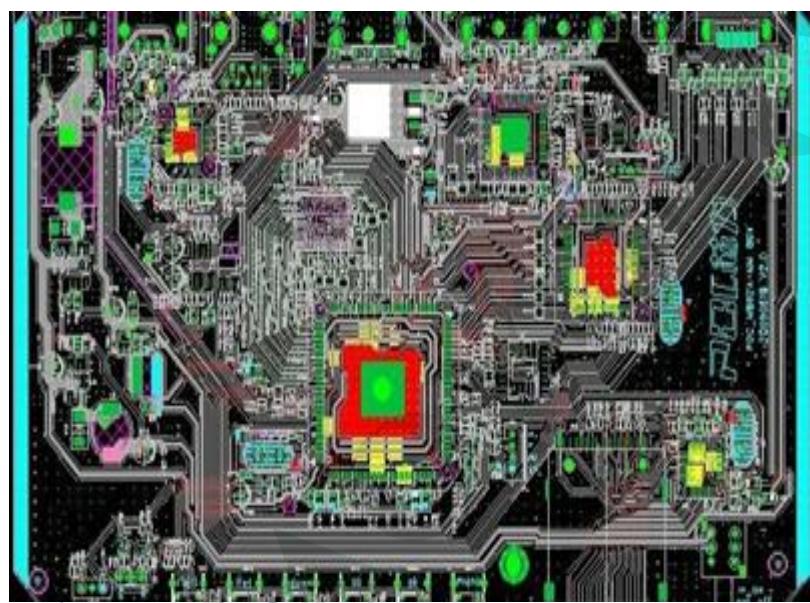
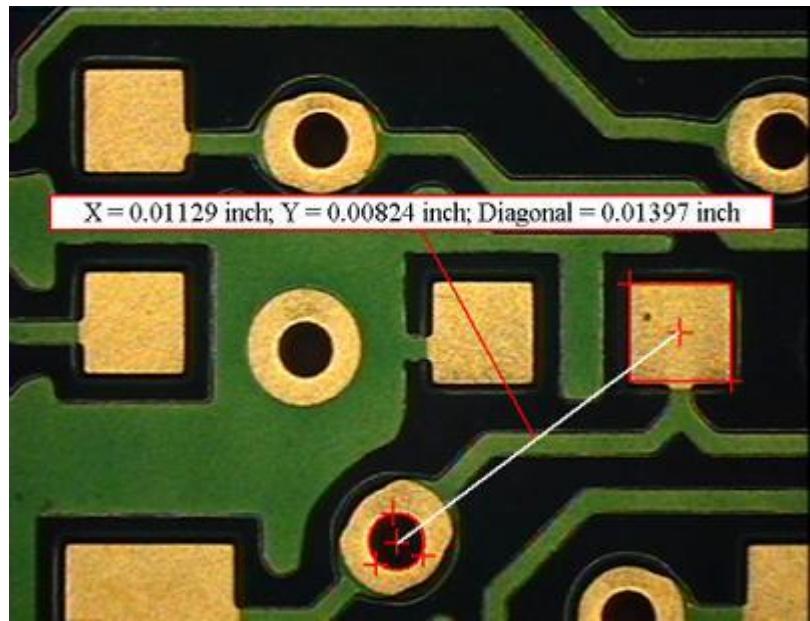
PERSON - 99%

PERSON - 94%

PERSON - 98%
PERSON - 97%

ENGINE - 97%





What is Computer Vision?

A brief history – Digital Images



Russell's infant son

Work on computer image processing started much earlier, in **1957**, at the American National Institute for Standards and Technology, where a group of engineers, led by **Russell Kirsch**, succeeded in making the **first ever digital scan of an image**. The image they used was of Russell's infant son, which became so famous that the *Life* magazine included it in its article about 100 images that changed the world. The original image is being kept in the Portland Art Museum.

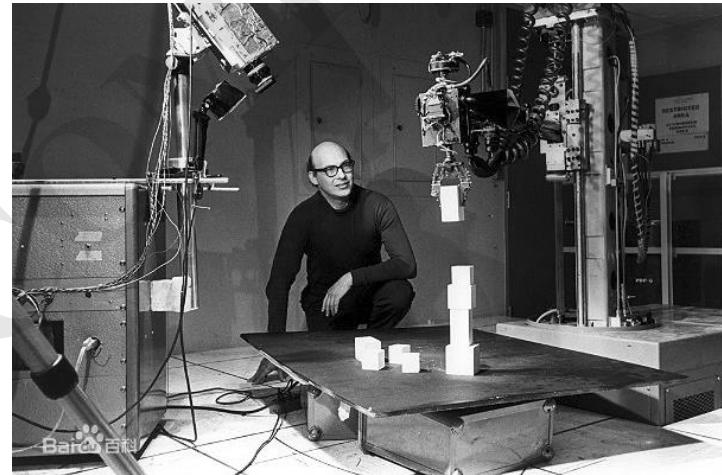
<https://www.megatrend.com/en/the-history-of-computer-vision-and-the-evolution-of-autonomous-vehicles/>

A brief history – 1960s-1980s

It was believed by some of the early pioneers of artificial intelligence and robotics (at places like MIT, Stanford, and CMU) that solving the “visual input” problem would be an easy step along the path to solving more difficult problems such as higher level reasoning and planning.

In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”.

Marvin Minsky proposed that framework was the basis of visual perception. For example, here is a framework for cars as follows:

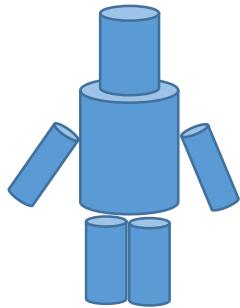


Marvin Minsky

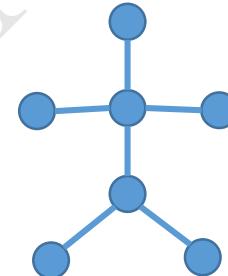
Name: Car
Super-class: Vehicle
Sub-class: Van, Jeep
Number of wheels: Unknown
Length: Unknown

A brief history – 1960s-1980s

How can we recognize and represent objects?



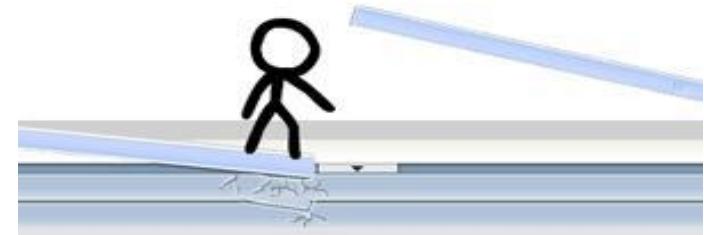
generalized cylinder
Brooks & Binford, 1979



pictorial structure
Fischler & Elschlager, 1973



Minecraft



Matchstick Man

A brief history – 1990s-2000s



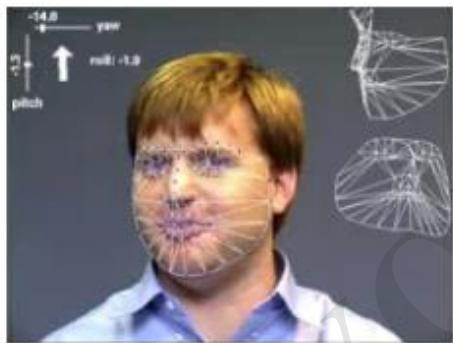
(a)



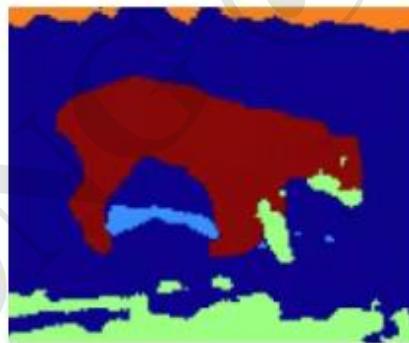
(b)



(c)



(d)



(e)



(f)

(a) Factorization-based structure from motion (Tomasi and Kanade, 1992) ; (b) Dense stereo matching (Boykov, Veksler and Zabih, 2001) ; (c) Multi-view reconstruction (Seitz and Dyer, 1999) ; (d) Face tracking (Matthews, Xiao and Baker, 2007) ; (e) Image segmentation (Belongie et al., 2002) ; (f) Face reorganization (Turk and Pentland, 1991).

A brief history – 1990s-2000s



(a)



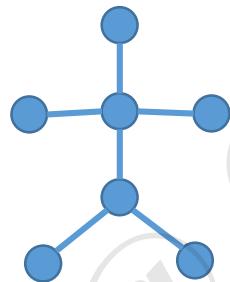
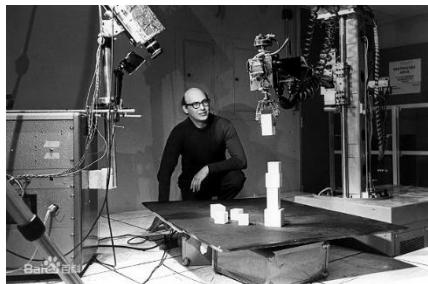
(b)



(c)

(a) Image stitching: merging different views (Szeliski and Shum 1997) ; (b) Exposure bracketing: merging different exposures; (c) Morphing: blending between two photographs (Gomes et al. 1999).

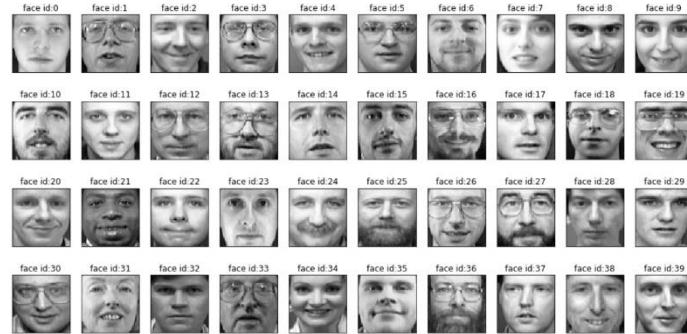
A brief history – 1960s-2000s



Observations on Small Samples

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9

MINIST Handwritten Digits

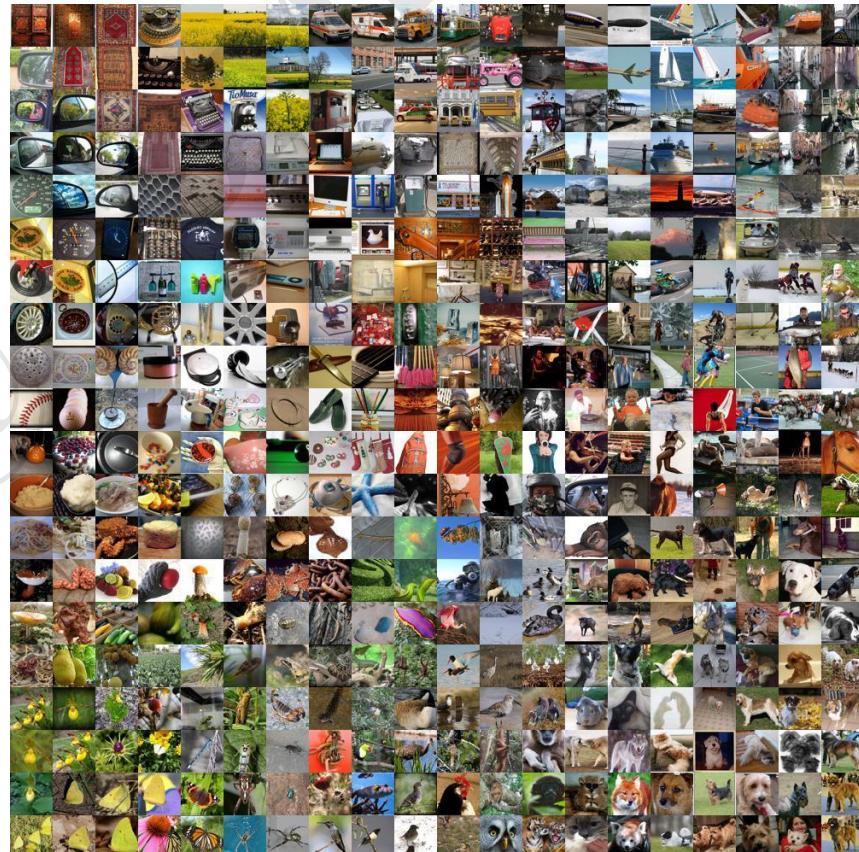


Olivetti Face Dataset, AT&T

A brief history – 2010

ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

1,000 object classes
1,431,167 images



<https://cs.stanford.edu/people/karpathy/cnnembed/>

Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25: 1097-1105.

A brief history – 2013

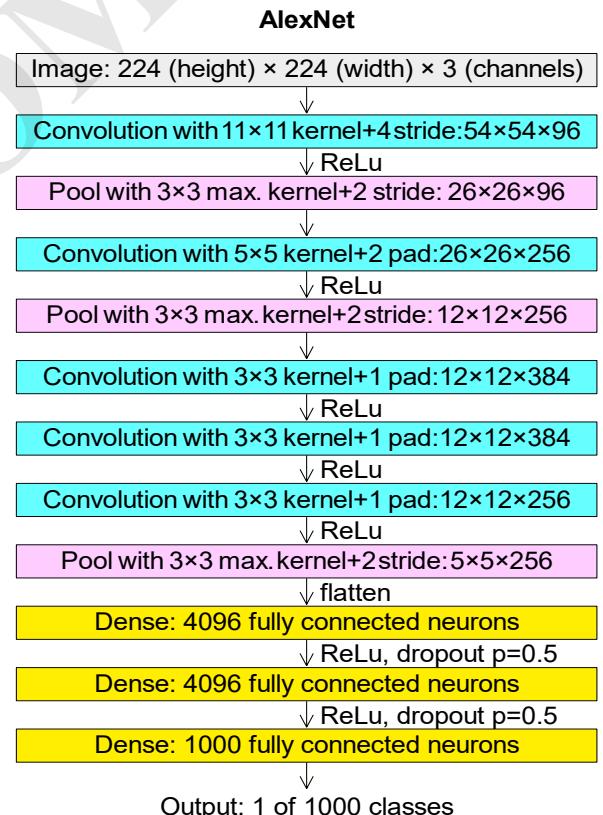
AlexNet

Winner in ILSVRC 2012

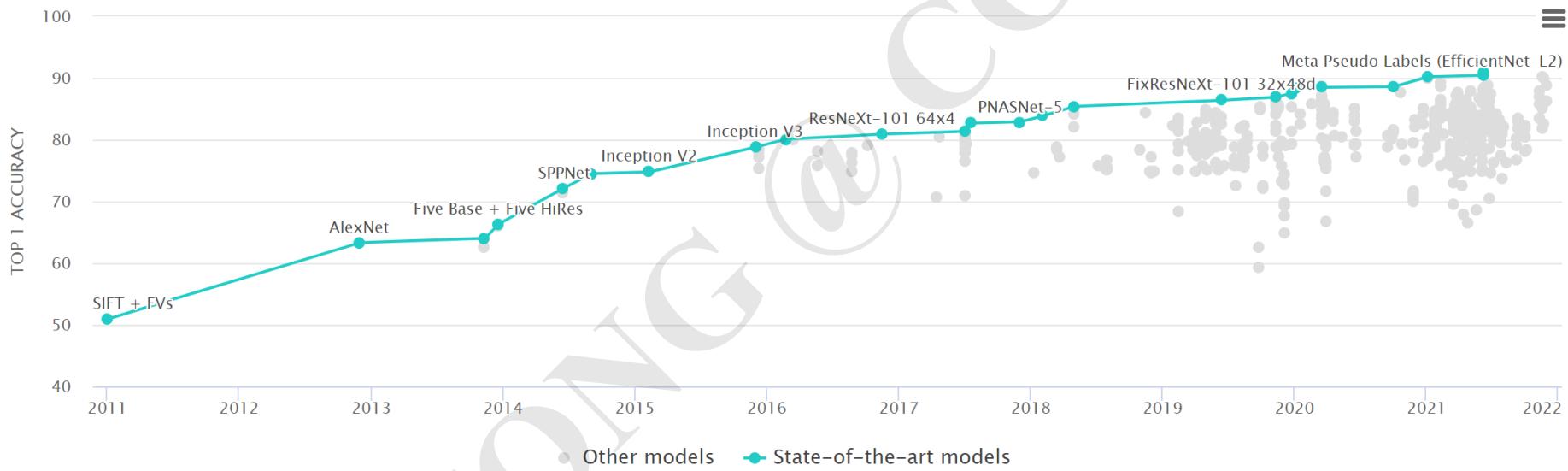
Top-5 error rate (rate of not finding the true label of a given image among its top 5 predictions) of 15.3%. The next best result trailed far behind (26.2%).

<https://en.wikipedia.org/wiki/AlexNet>

Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25: 1097-1105.



A brief history – 2010s-2020s



What is Computer Vision?

Keywords

>Recognition, Classification,
Segmentation, Detection, Image
Processing, Generation,
Reconstruction, Semantics, Machine
Learning, Models, Deep Learning ...

Learning CV by playing it

In-class Social Network Mining

Do you like attendance-check?

No!!!

Why don't we make it fun?

An Interesting Way for Attendance-Check

- >Take photos in class
- >Publish them on course homepage
- >Online check (by students) when downloading course materials

An Interesting Way for Attendance-Check



Photos taken from class

An Interesting Way for Attendance-Check



Image Stitching

An Interesting Way for Attendance-Check



Face Detection

ucopy.net - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Inbox (1,298) - 1200@... 优酷-中国第一视频网站,提供... ucop... Balancing Thread Based Na... 百度搜索_attendance check 我的首页 优酷视频空间-聊聊... Google

ucopy.net

Most Visited Getting Started Latest Headlines

ID	Portrait	Student(s)
0		
1		
2		
3		

Attendance Check

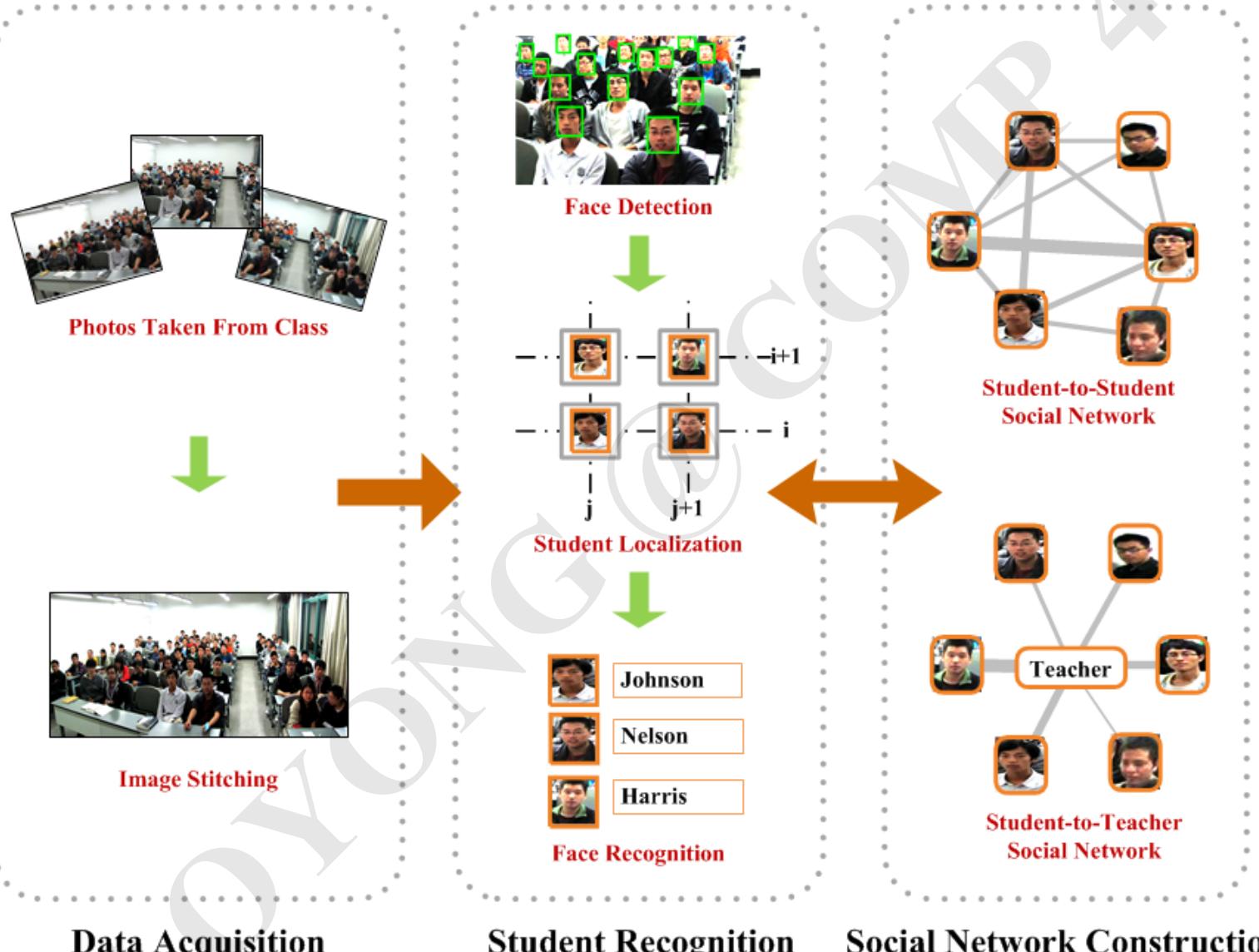
It's me! (0) Not me! (0)
It's me! (1) Not me! (1)
It's me! (2) Not me! (2)
It's me! (3) Not me! (3)

Administration File Manager New Page Edit This Page Page Settings

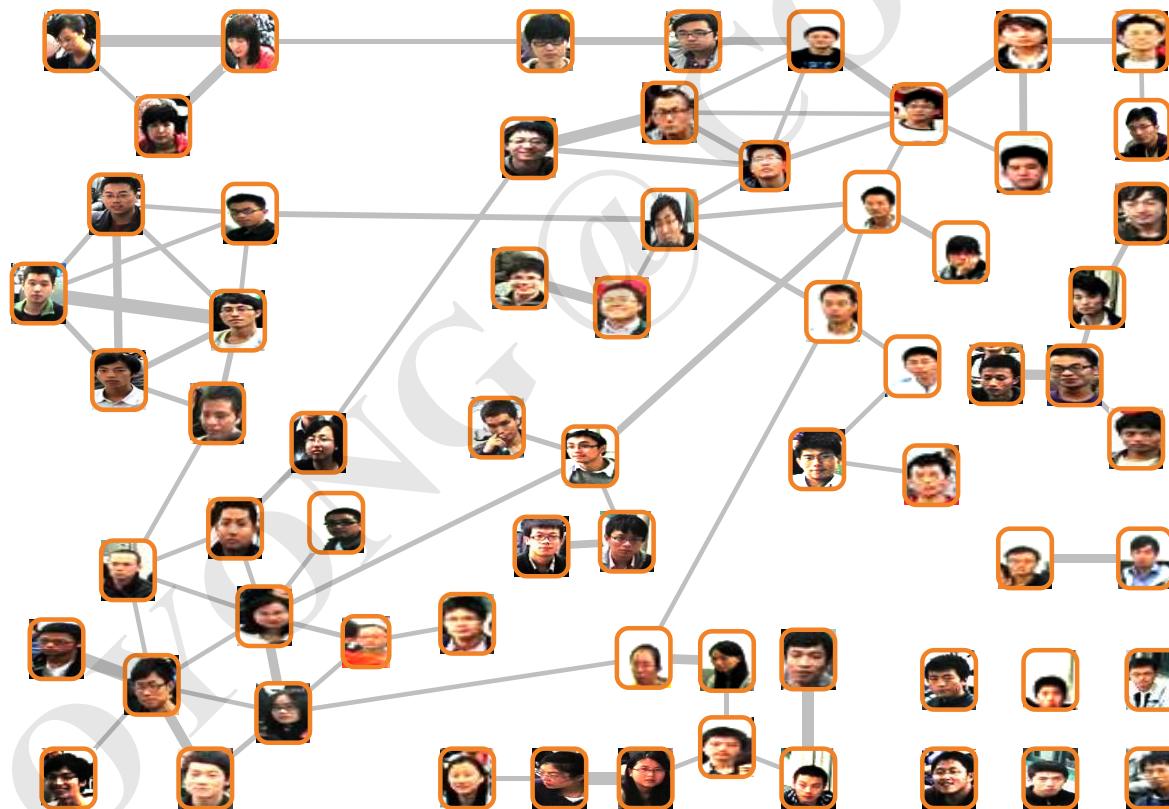
Online Check

Take a step further

- >Integrate with facial recognition
- >Conduct in-class social network analysis

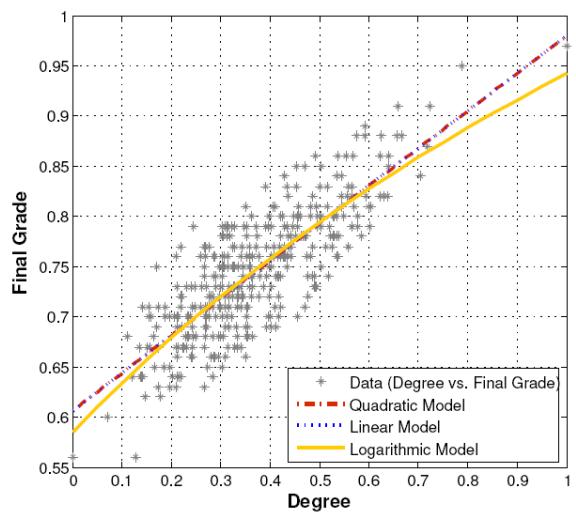
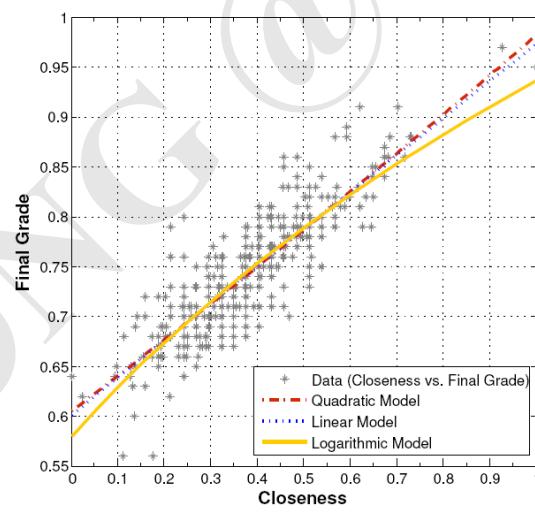
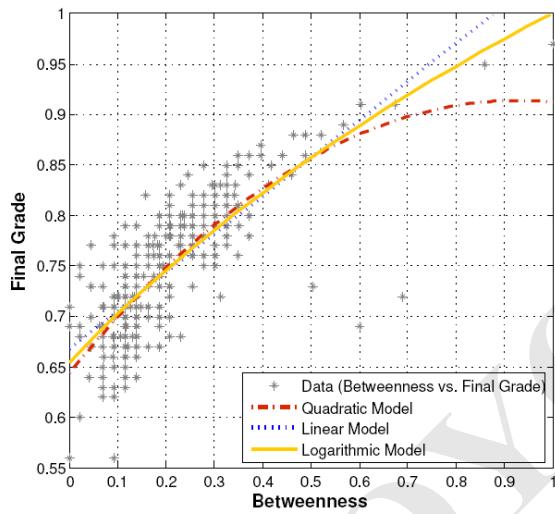


Social Network Construction & Pedagogical Analysis



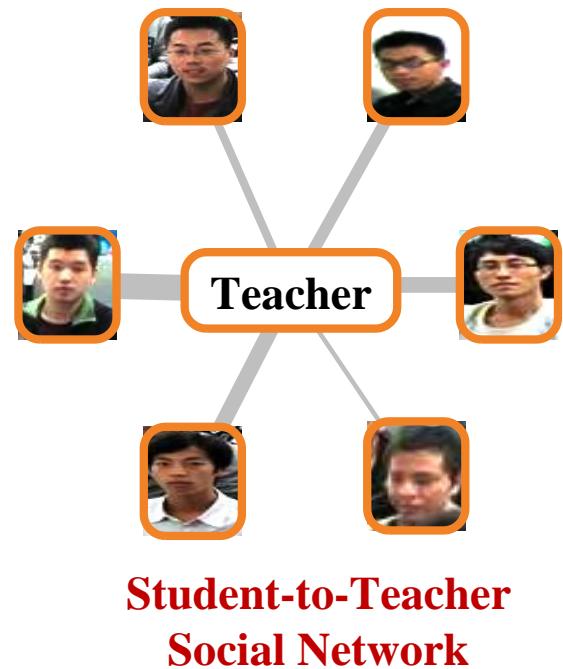
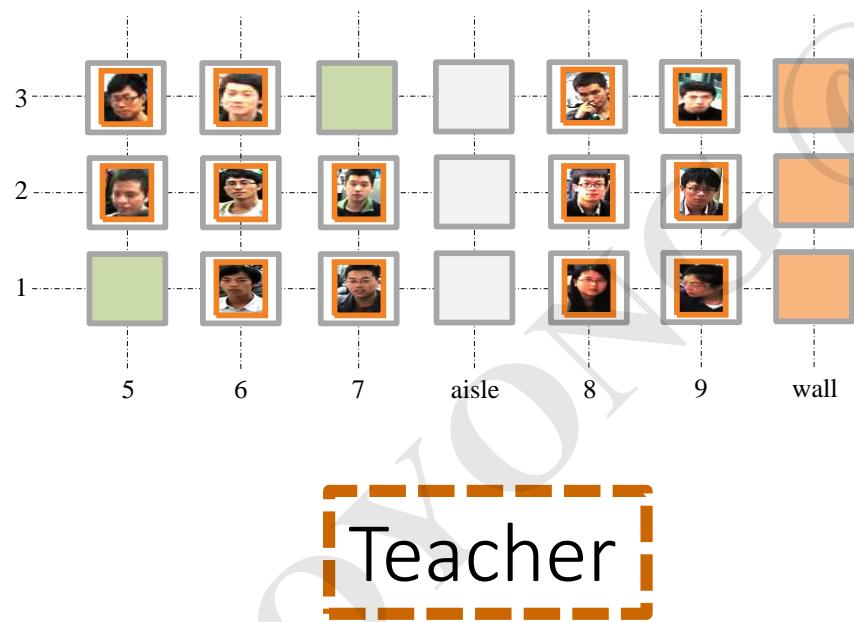
Social Network Construction & Pedagogical Analysis

- > Student-to-Student network construction
 - Experimental Results



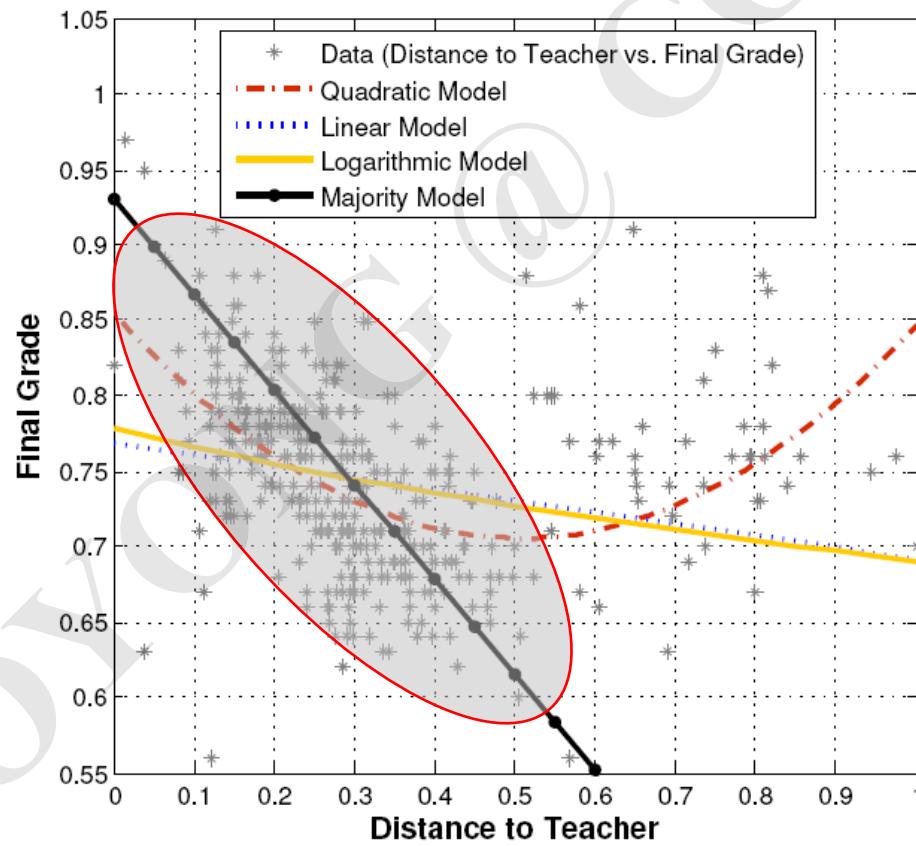
Social Network Construction & Pedagogical Analysis

> Student-to-Teacher network construction



Social Network Construction & Pedagogical Analysis

> Student-to-Teacher network construction



Conclusions

- > Facial recognition in extreme conditions can be improved by social inference
- > In-class social network is useful for pedagogical analysis to reach objective and quantitative conclusions

Conclusions

Attendance-check can be fun



2012-10-10: 116 students



More steps further

Multimedia Computing Lab (McLab)
Sichuan University (cswei@scu.edu.cn)



封面 封面视频

—亿万年轻人的生活方式 —

秒拍

Why don't we make it fun when
you have to do it in your classes?

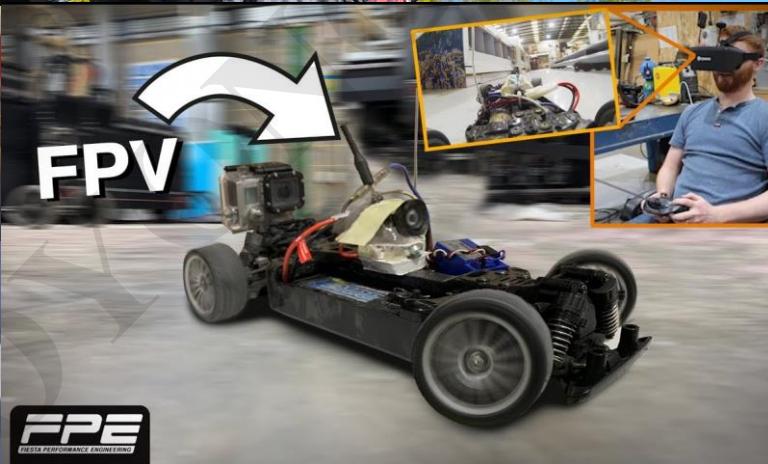
不如我们课上来点**有趣的吧**

A step we can take together

First Person View (FPV)



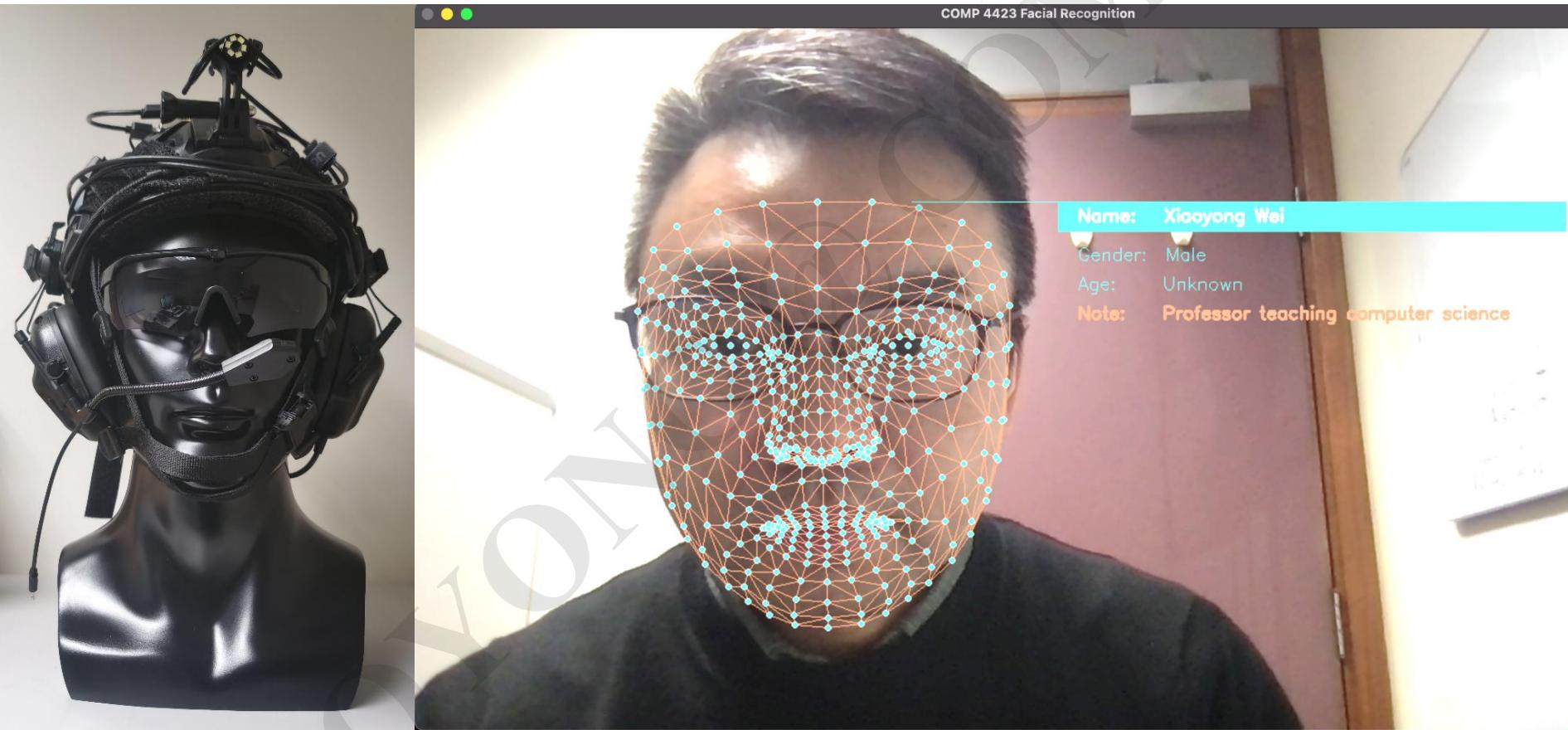
First Person View (FPV)



First Person View (FPV)



FPV Recognition @ 4423



Possible Applications

- >Facial Recognition
- >Celebrity Lookalike



Du Yiheng (left) is known for his remarkable resemblance to Andy Lau (right).



Data Sharing Agreement

- > A Survey is going to be announced on Blackboard for your willingness of sharing
- > Data will be used **ONLY** for assignments and demos. Data will be used **ONLY** in this class. We will **delete** all data at the end of this semester.
- > Data is shared only among group members. The larger your group is, the more data you have. You can join a CLASS group to share with everyone in that group. Or you can form a PRIVATE group with your friends
- > You can share data with/without masks

Road Map (tentative)

- >L01: Introduction
- >L02: Image Processing 1
- >L03: Image Processing 2
- >L04: Feature Extraction
- >L05: Image Retrieval
- >L06: Image Classification 1
- >L07: Image Classification 2

Road Map (tentative)

- >L08: Advanced Image Retrieval
- >L09: Advanced Image Classification
- >L10: 3D Representation and Reconstruction
- >L11: Advanced topics & Course projects
- >L12: Advanced topics & Course projects
- >L13: Advanced topics & Course projects
- >Final Exam

Course websites

>PolyU Blackboard learn.polyu.edu.hk

- Announcements
- Slides
- Assignments
- Projects

Timetable

> Lectures at R1205

- Wed. 15:30-17:20
- Week 1-13

> Tutorials

- **TUT001:** Mon. 12:30-13:20 at **PQ604C**
- **TUT002:** Mon. 11:30-12:20 at **PQ604C**
- Week 2-13

Assessments

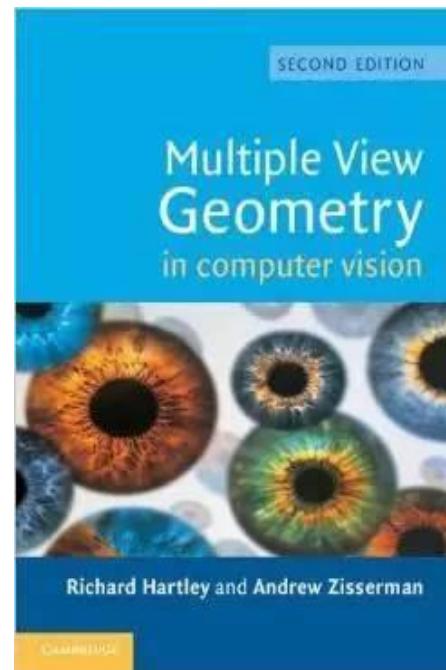
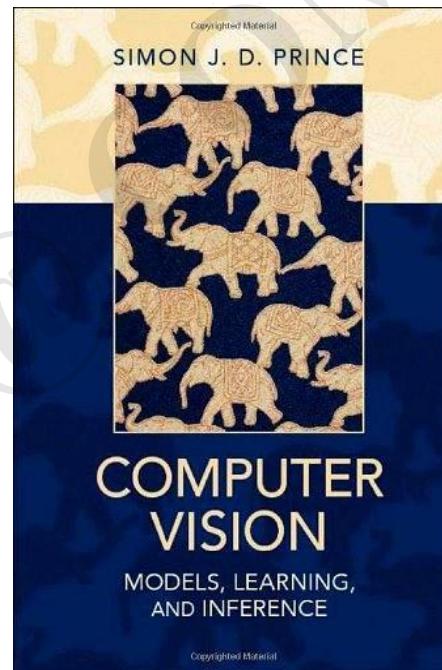
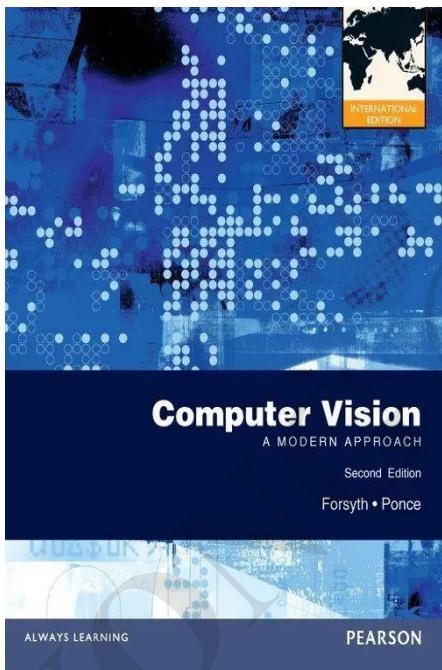
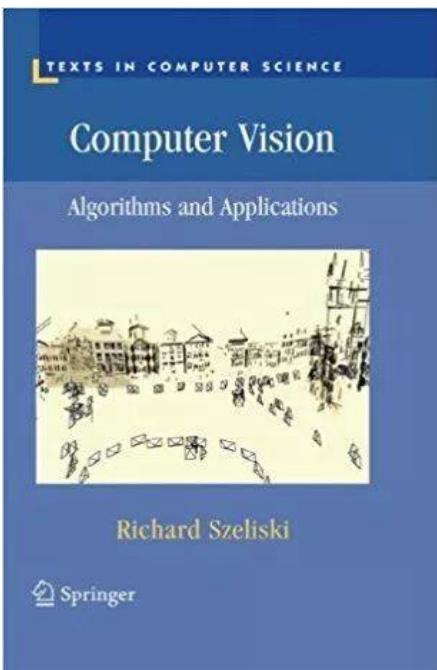
> Continuous assessments 55%

- 2 Assignments 30% (15% each, late submissions acceptable up to 3 days with 10% deduction per day)
- 1 Project 25%

> Final Examination 45% (open book)

> **Zero tolerance to cheating and plagiarisms**

References



Szeliski R. Computer Vision: Algorithms and Applications[J]. 2011.

Forsyth D, Ponce J. Computer vision: A modern approach[M]. Prentice hall, 2011.

Prince S J D. Computer vision: models, learning, and inference[M]. Cambridge University Press, 2012.

Andrew A M. Multiple view geometry in computer vision[J]. Kybernetes, 2001.

Thank you!

