

Introduction to Computer Vision Applications

Dr. Halit Orenbas



Dr. Halit Orenbas

Engineering Manager at Huawei

Contents

Part 1: CV101

1. What is CV?
2. What a CV Engineer does?
3. What are the start points?
4. What is the research direction?

Part 2: CV in Huawei

1. Video Advertisement Placement
2. Optical Character Recognition
3. Environment, Health and Safety Project
4. Search by Photo & AI Shopping
5. Network Compression

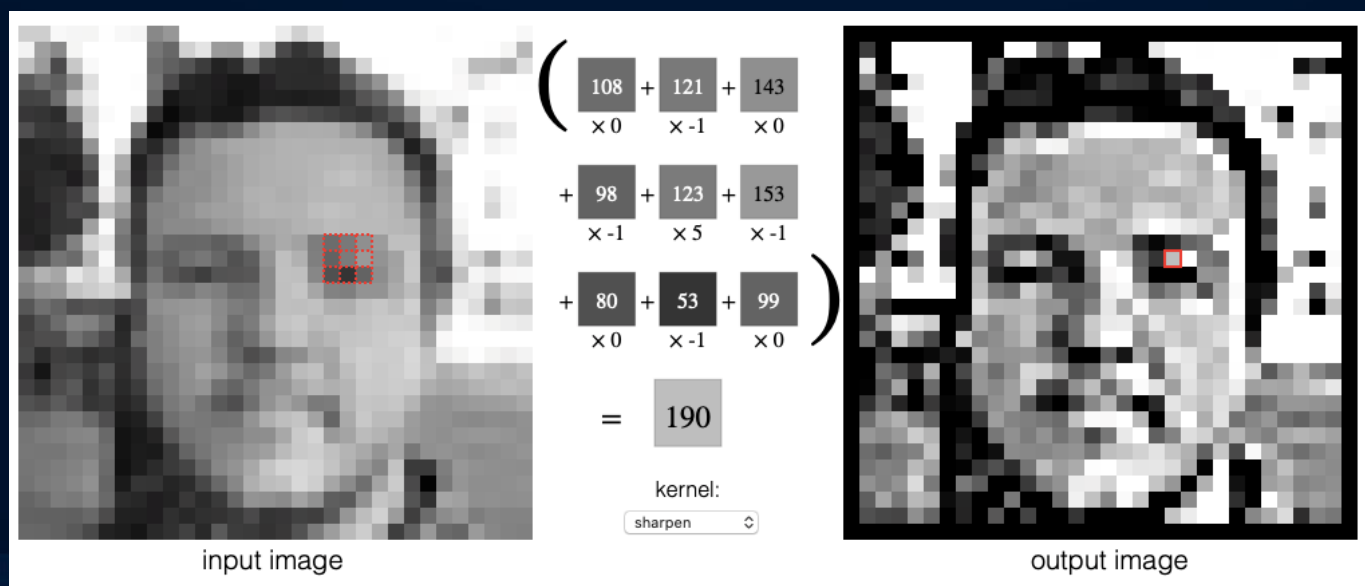
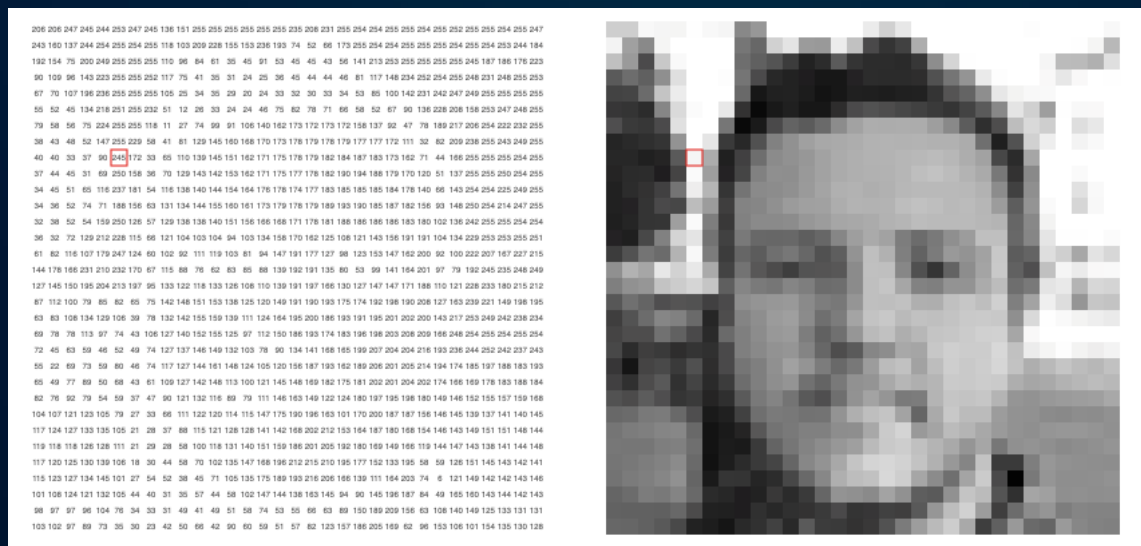


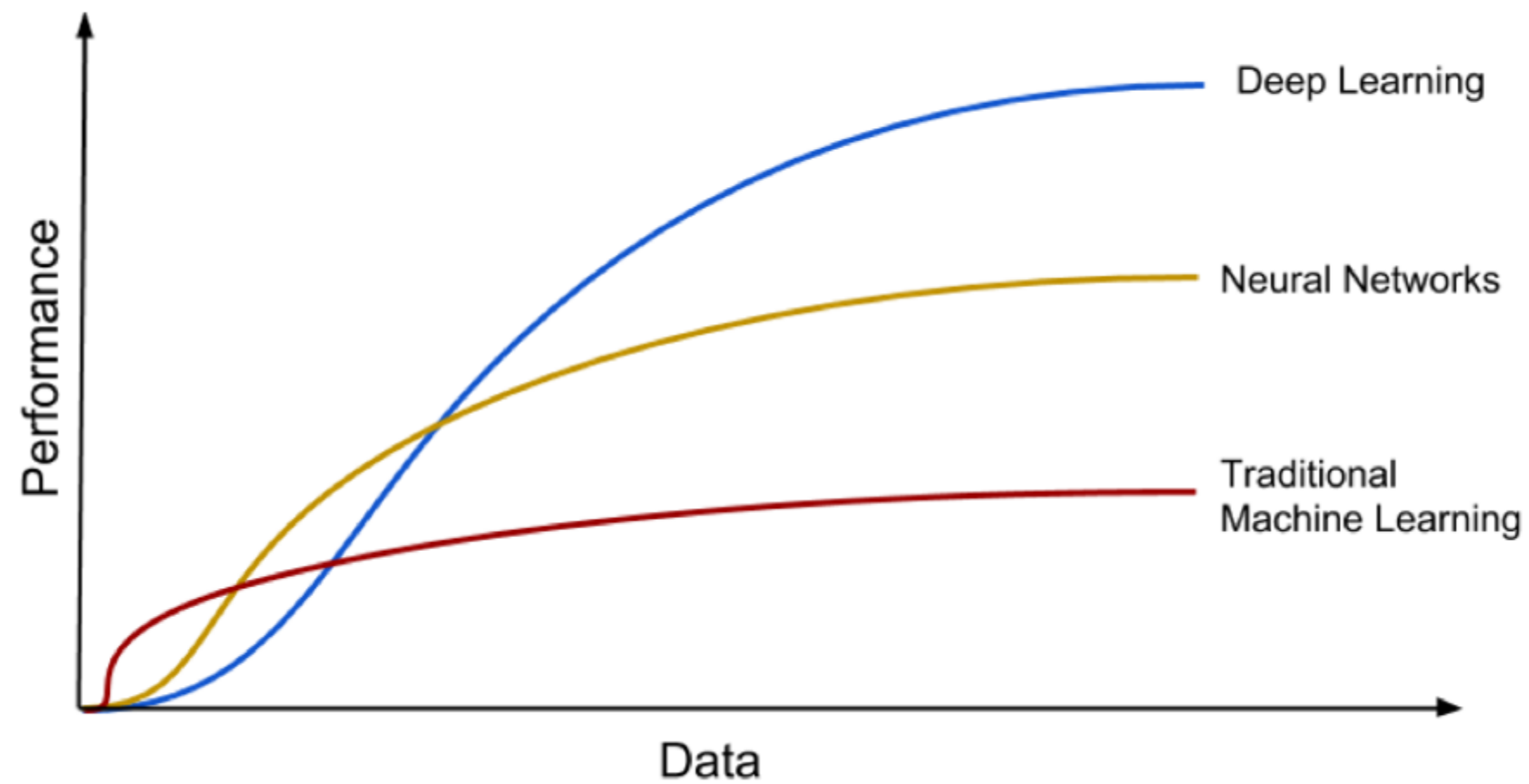
Part 01

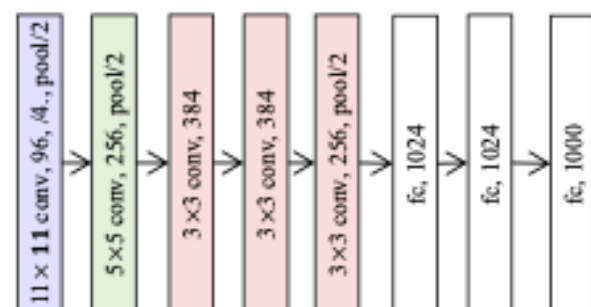
A Gentle Introduction to Computer Vision

1. What is CV?
2. What a CV Engineer does?
3. What are the start points?
4. What is the next?

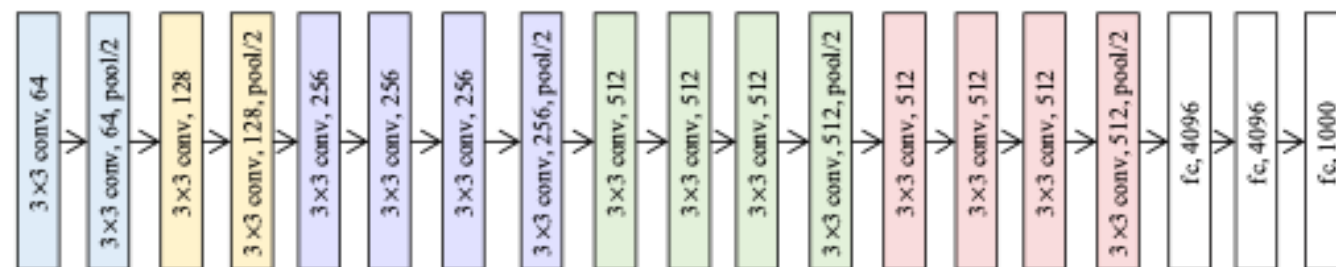
What is Computer Vision ?



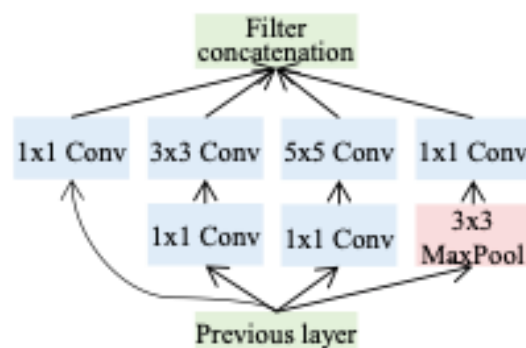




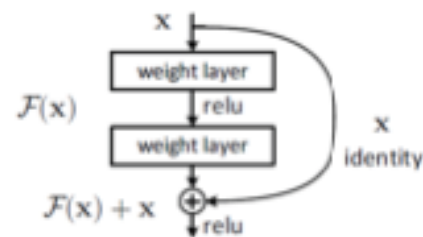
(a) Alexnet



(b) VGGNet



(c) GoogleNet

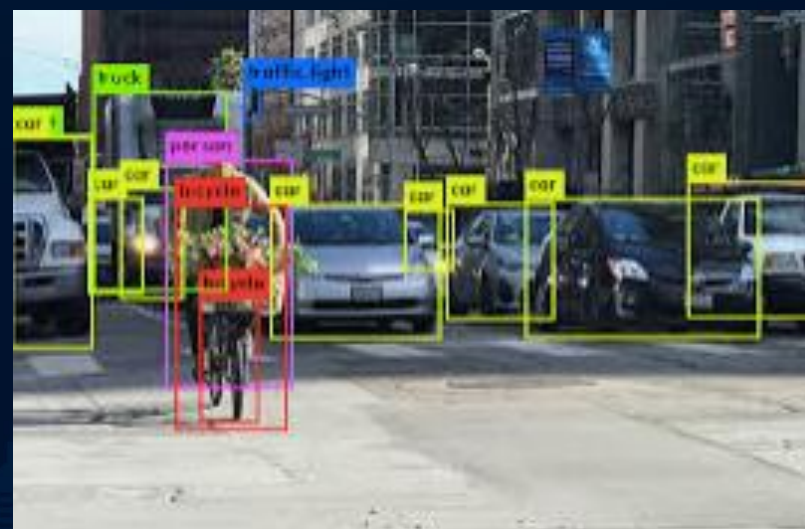
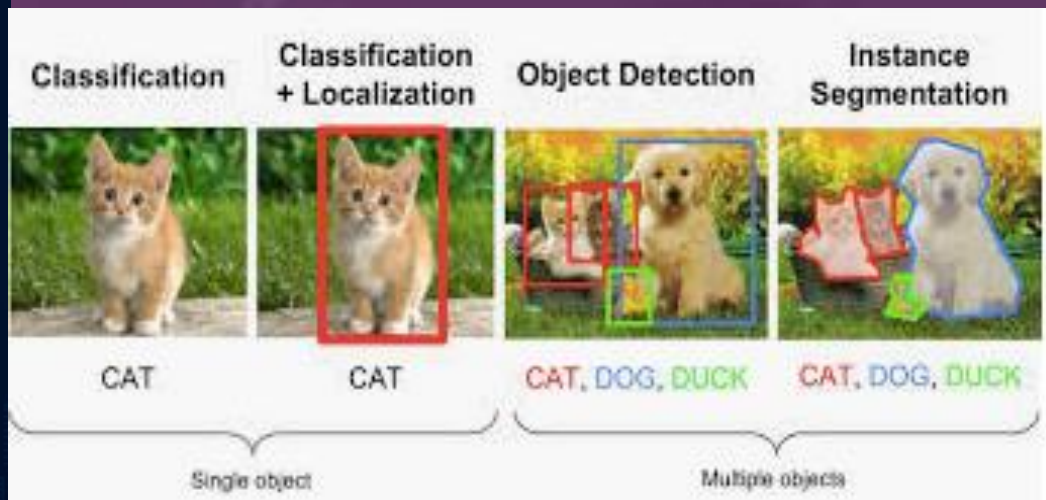


(d) ResNet



(e) SENet

What a CV Engineer does ?



Start Point

- Python
- Machine Learning
- Neural Networks
- Deep Learning

Research Directions

- Few data
- Unlabeled data
- Model sizes
- Inference times

Suggestions

- Conferences: CVPR, ECCV, ICLR, ICCV
- PyTorch:
<https://pytorch.org/tutorials/>
- Open source:
<https://github.com>
- Papers and codes:
<https://paperswithcode.com>
- State of the art (SOTA):
<https://paperswithcode.com/sota>



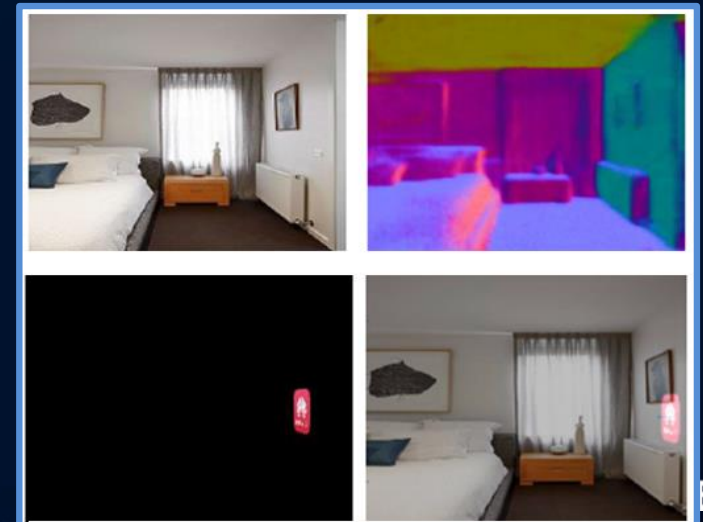
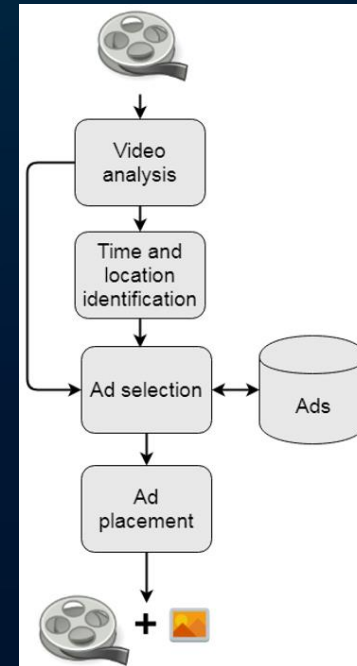
Part 02

Computer Vision in Huawei


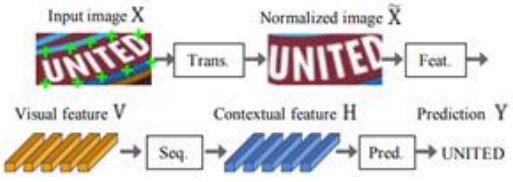


1. Video Advertisement Placement
2. Optical Character Recognition
3. Environment, Health and Safety Project
4. Search by Photo & AI Shopping
5. Network Compression

Video Advertisement Placement

- Industrial **research** project supported by The Scientific and Technological Research Council of Turkey (TUBITAK)
- Purpose
 - Choose the **best location and time frame in the video**
 - Select the **most relevant advertisement** to the video contents
 - Place and keep the advertisement on the selected surface by **3D transformations** and **handle the occlusions**.
- Methods
 - Video content analysis
 - Object detection, semantic segmentation, surface normal estimation, rectangle/line detection, visual tracking, room layout estimation, feature extraction, etc.
 - Spatial & temporal geometric analysis
 - Word embedding based advertisement matching



Optical Character Recognition

TEXT DETECTION	TEXT RECOGNITION	TEXT SPOTTING	SYNTHETIC DATA GENERATION
<p>Detection of arbitrary shaped text from real-world images</p> 	<p>Recognition of arbitrary shaped text content from real-world images</p> 	<p>Simultaneous detection and recognition of text in arbitrary shapes more efficiently and</p> 	<p>Synthetic text data generation for training of complex models</p> 

STRUCTURED TEXTUAL DATA EXTRACTION

Extraction of text from images in a structured format (line by line, paragraph by paragraph)

Welcome to the Tcl Developer Xchange!
Join the many thousands of software developers who are already more productive with help from the **Tcl** programming language and the **Tk** graphical user interface toolkit.

Tcl (Tool Command Language) is a very powerful but easy to learn dynamic programming language, suitable for a very wide range of use, including web and desktop applications, networking, administration, testing and many more. Open source and business-friendly. Tcl is a mature yet evolving language that is truly cross platform, easily deployed and highly extensible.

Tk is a graphical user interface toolkit that takes developing desktop applications to a higher level than conventional approaches. Tk is the standard GUI not only for Tcl, but for many other dynamic languages, and can produce rich, native applications that run unchanged across Windows, Mac OS X, Linux and more.



Welcome to the Tcl Developer Xchange!
Join the many thousands of software developers who are already more productive with help from the **Tcl** programming language and the **Tk** graphical user interface toolkit.

Tcl (Tool Command Language) is a very powerful but easy to learn dynamic programming language, suitable for a very wide range of use, including web and desktop applications, networking, administration, testing and many more. Open source and business-friendly. Tcl is a mature yet evolving language that is truly cross platform, easily deployed and highly extensible.

Tk is a graphical user interface toolkit that takes developing desktop applications to a higher level than conventional approaches. Tk is the standard GUI not only for Tcl, but for many other dynamic languages, and can produce rich, native applications that run unchanged across Windows, Mac OS X, Linux and more.



```
1 Welcome to the Tcl Developer Xchange!
2 Welcome to the Tcl Developer Xchange!
3 Join the many thousands of software developers who are already more productive with help from the Tcl
4 programming language and the Tk graphical user interface toolkit
5 Tcl (Tool Command Language) is a very powerful but easy to learn dynamic programming language suitable for a
6 very wide range of uses, including web and desktop applications, networking, administration, testing and many
7 more. Open source and business-friendly. Tcl is a mature yet evolving language that is truly cross platform,
8 easily deployed and highly extensible.
9 Tk is a graphical user interface toolkit that takes developing desktop applications to a higher level than
10 conventional approaches. Tk is the standard GUI not only for Tcl, but for many other dynamic languages and
11 can produce rich, native applications that run unchanged across Windows, Mac OS X, Linux and more.
```


Environment, Health and Safety Project

- 3 Main areas

- Human detection (area intrusion)
 - Safety equipment check
 - Face + mask detection



- Correct mask

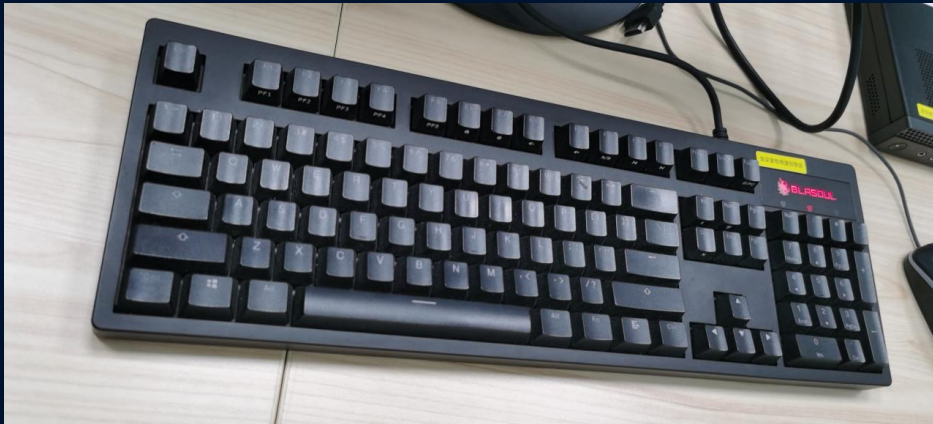


- Incorrect mask



Search by Photo - AI Shopping

- Global image feature extraction
- Dataset cleaning with autoencoder based outlier detection
- Feature based classification
- Recommendation and ranking of query images
- Links to purchase recommended items



Query Image



Recommended Image

Network Compression

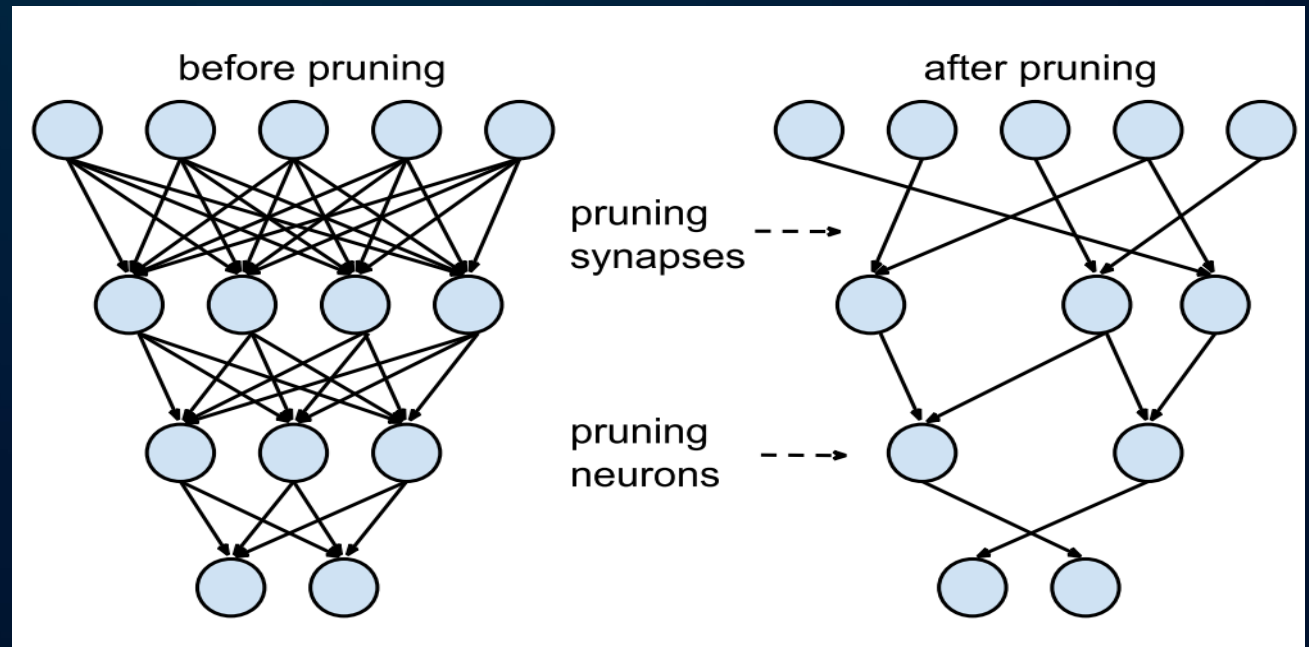
- Weight Pruning

- Filter Pruning

Faster inference time, model size reduction, lighter networks

- Knowledge Distillation

Increasing AI model accuracy via teacher and student model structure



Computer Vision Application Areas Training

BTK Academy Computer Vision Application
Areas Training:



Thank you.

Bring digital to every person, home and organization for a fully connected, intelligent world.

**Copyright©2018 Huawei Technologies Co., Ltd.
All Rights Reserved.**

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

