EE655: Computer Vision & Deep Learning

Lecture 01

Koteswar Rao Jerripothula, PhD
Department of Electrical Engineering
IIT Kanpur

Lecture Outline

- **Course Overview** 
- Guidelines & Course Policies
- Introduction to Computer Vision & Deep Learning

EE 604 Team

Instructor

Dr. Koteswar Rao Jerripothula
Assistant Professor
Department of Electrical Engineering
Indian Institute of Technology Kanpur

Email: kotesrj@iitk.ac.in
Tel: +91 512 259 2483
Add: #302, ACES Building, IIT Kanpur
Office Hours: WTh 11:30-12:30

ORCiD: 0000-0002-3507-3731
Vidwan-ID: 153811

TAs

Sai Lakshmi Yejjipurapu	lakshmi24@iitk.ac.in
Suryansh Singh01	Ssingh24@iitk.ac.in
Kuldeep Chaudhary	Kuldeepc24@iitk.ac.in
Kuldeep Kumar	kuldeepk21@iitk.ac.in

Logistics

Course Schedule: Tuesdays and Fridays, 2:00 to 3:15 p.m. (L02)

Office Hours: Fridays, 3:15 to 4:30 p.m.

- *@ my office, ACES 302*
- *Take appointment before coming*

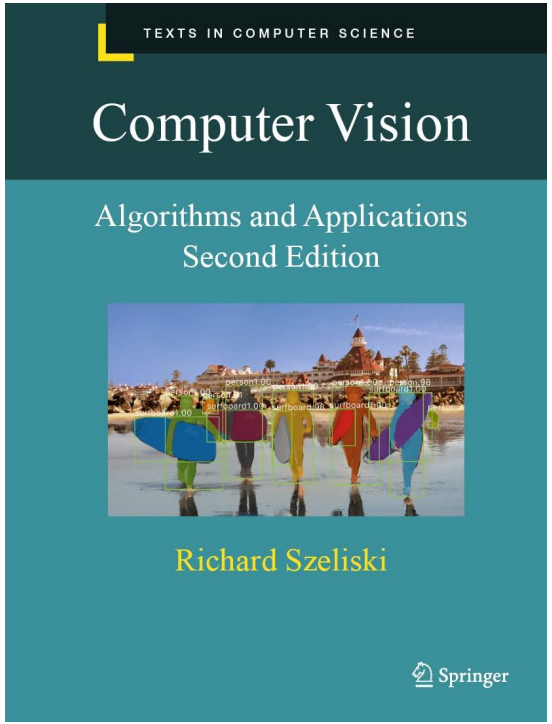
Course Objectives

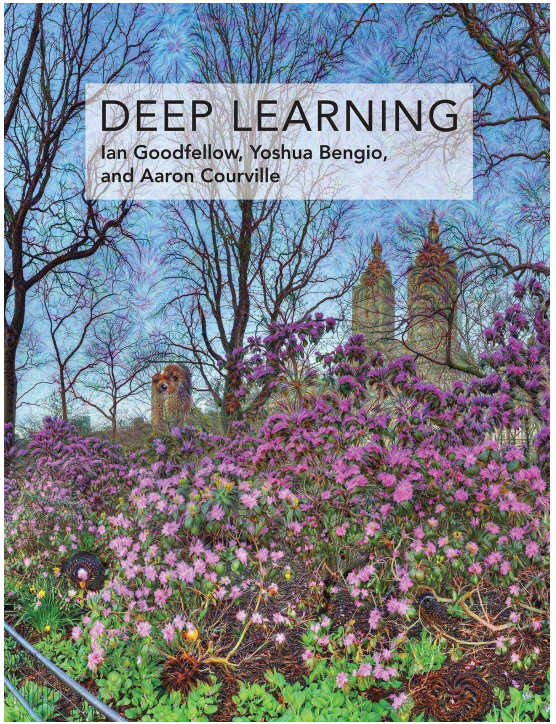
At the end of the course, the student should be able to:

- **Learn** the basics of Computer Vision & Deep Learning
- **Understand** the wide applicability of different Computer Vision & Deep Learning Techniques
- **Develop** an insight towards which method to use when
- **Apply** learned techniques to build a real-world system

Contents

Primary References





Other References

Internet & Research Papers from Top Venues:

Top publications

Categories > Engineering & Computer Science > **Computer Vision & Pattern Recognition**

	Publication	h5-index	h5-median
1.	IEEE/CVF Conference on Computer Vision and Pattern Recognition	440	689
2.	IEEE/CVF International Conference on Computer Vision	291	484
3.	European Conference on Computer Vision	206	306
4.	IEEE Transactions on Pattern Analysis and Machine Intelligence	196	348
5.	IEEE Transactions on Image Processing	150	202

Top publications

Categories > Engineering & Computer Science > **Artificial Intelligence**

	Publication	h5-index	h5-median
1.	Neural Information Processing Systems	337	614
2.	International Conference on Learning Representations	304	584
3.	International Conference on Machine Learning	268	424
4.	AAAI Conference on Artificial Intelligence	220	341

Evaluation

Components	Weightage	Remarks
Homeworks (10)	10%	Simple questions; once a week; binary marking
Quizzes (2)	20%	MCQ, T/F, FITB type question; announced & closed-book
Assignments (2)	20%	You will have one week; late submissions are not allowed.
Course Project	20%	Presentation (5%) + Implementation (10%) + Write-up (5%)
End-sem Exam	30%	Theoretical & Numerical questions; closed-book

NOTE: No mid-sem exam

Grading Policy

Score = max (marks% + bonus, percentile)

Grade	Cutoff Score
A*	99
A	90
B+	81
B	72
C+	63
C	54
D+	45
D	36

NOTE 1: Minimum 10% students will get A*/A

NOTE 2: “marks% + bonus” must be at least 36 to pass

NOTE 3: Upto 3% bonus marks for submitting a research paper based on your course project.

Travel Support: https://www.iitk.ac.in/doaa/data/travel-support-national-conference-Revised.pdf

Lecture Outline

- Course Overview
- **Guidelines & Course Policies** 
- Introduction to Computer Vision & Deep Learning

Guidelines

- The easiest way to reach the instructor is via email.
- Email regarding the course should start with [EE655] in the subject line.
- Homeworks, Assignments & Course Project should be submitted at Hello IITK.
- Quizzes & End-sem exam will be conducted offline, and will be closed-book.
- Maximum team-size in the course project: 4
- A list of projects will be shared soon.
- Assignments will be entirely programming-based.

OR

Course Policies

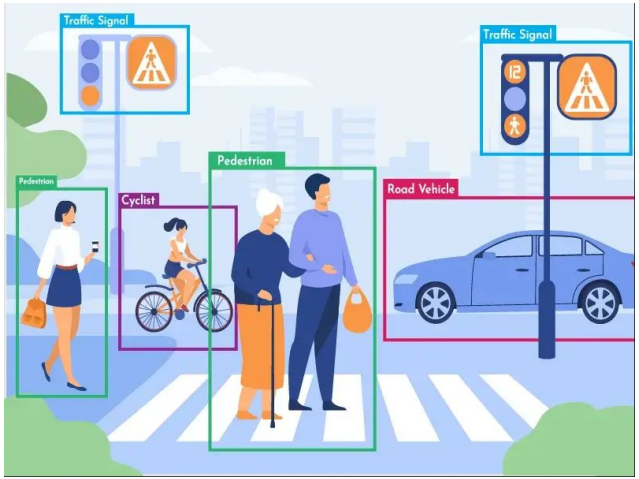
- It’s expected that you will maintain academic honesty in every form.
- Please familiarize yourself with the general rules and laws of plagiarism.
- If plagiarism is detected, you'll be given straight zero marks in the entire evaluation component concerned. That is,
  - If it’s detected in any homework, you get 0 for entire 10% (weightage of homeworks).
  - If it’s detected in any assignment, you get 0 for entire 20%
  - If it’s detected in any quiz, you get 0 for entire 20%
  - If it’s detected in any part of the course project, you get 0 for entire 20%
  - If it’s detected in any question of the end-sem exam, you get 0 for entire 30%
- Attendance will not be taken, but it’s important to attend lectures to do well in the homeworks’ component, and the course at large.

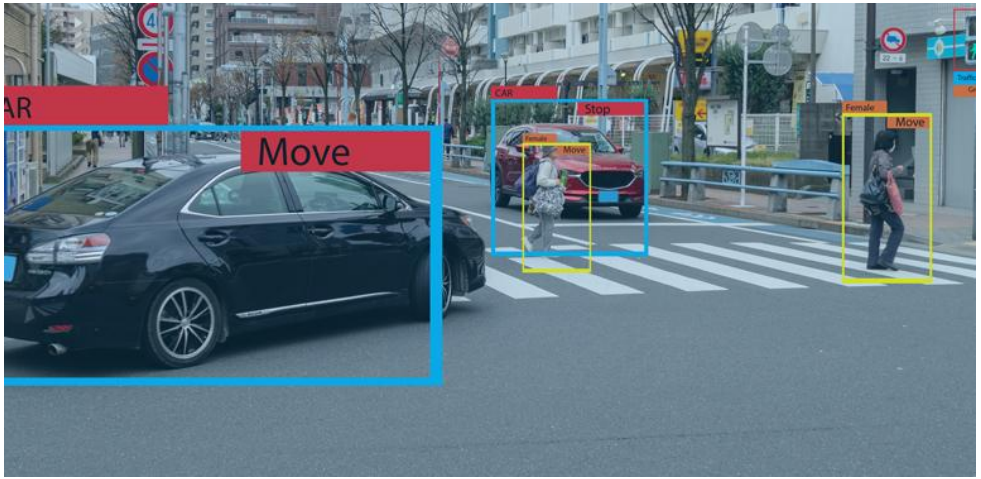
Lecture Outline

- Course Overview
- Guidelines & Course Policies
- **Introduction to Computer Vision & Deep Learning** 

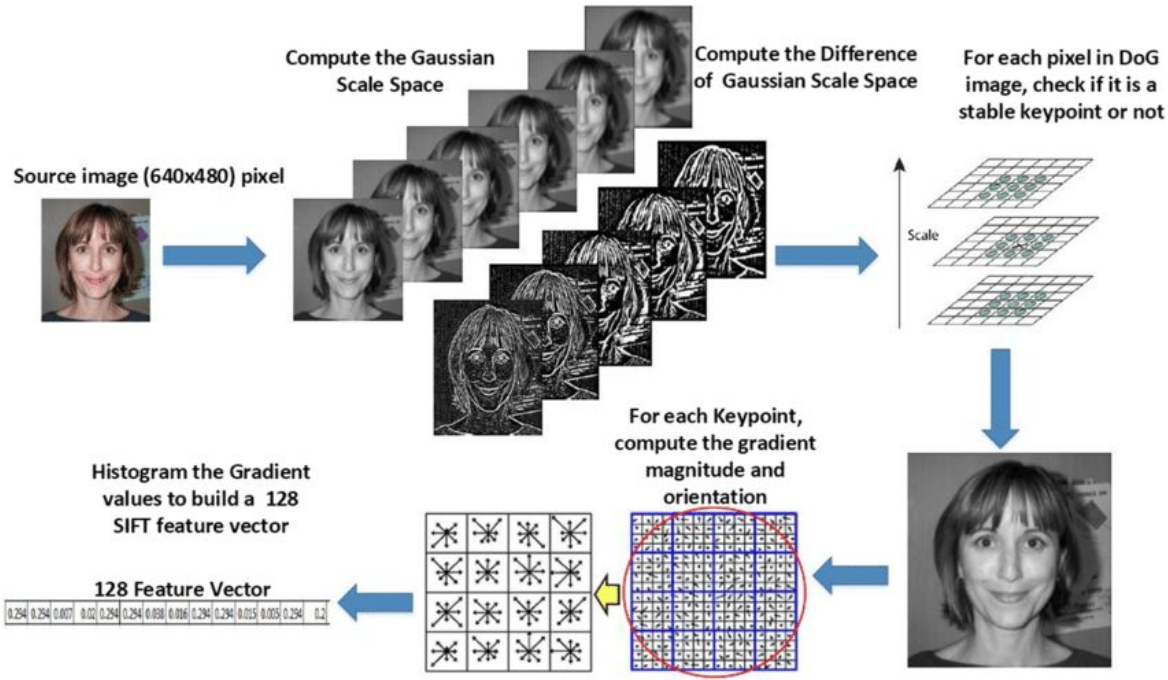
Computer Vision

Computer vision is a sub-field of AI that focuses on enabling computers to identify and understand objects and people in images and videos



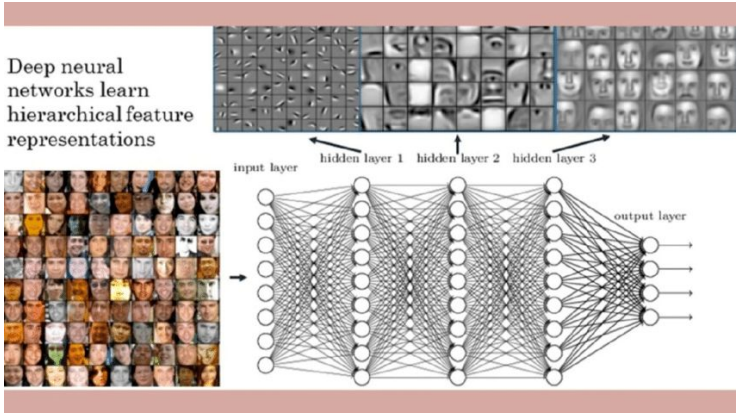


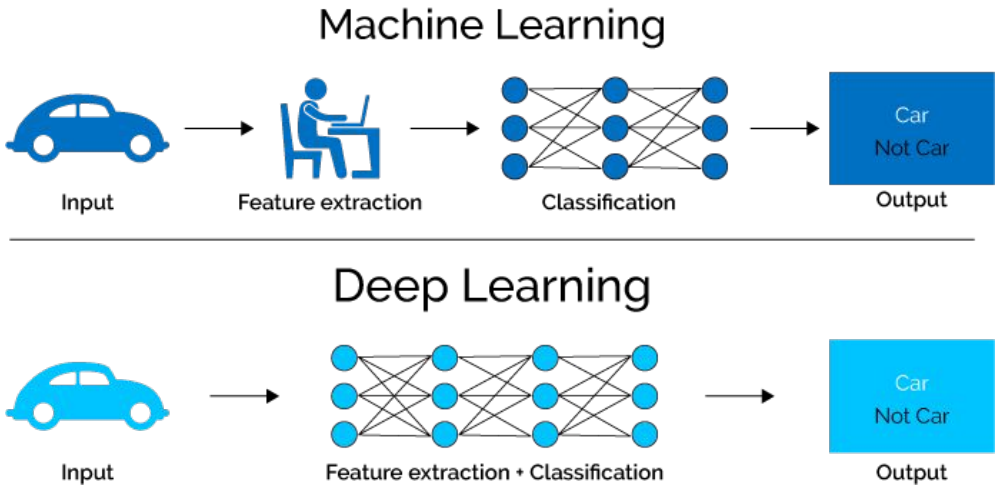
Feature Extraction



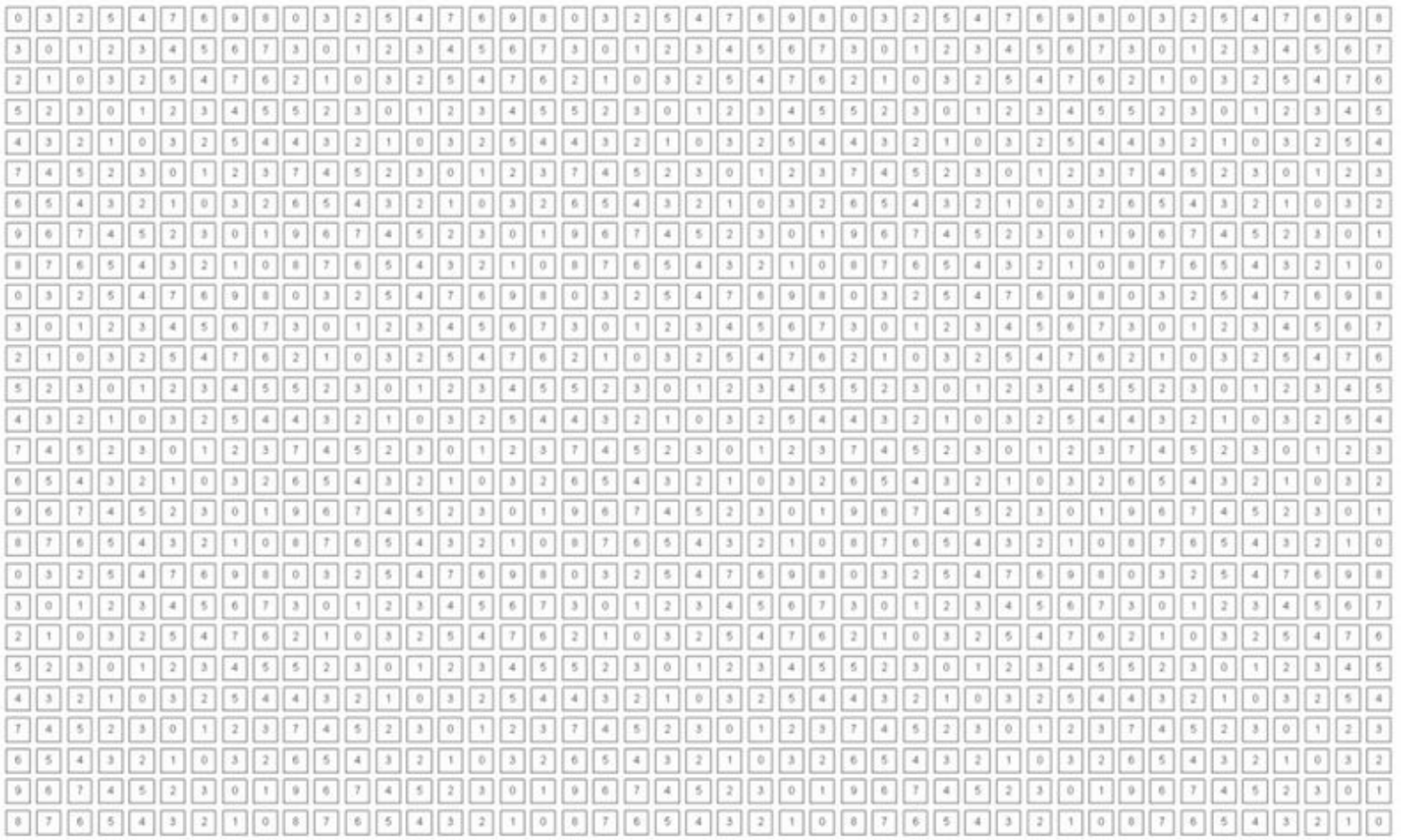
An ML model can be trained using such a feature vector to identify the gender

Deep Learning





What a person sees



What a computer sees

True for even CV Algorithm outputs

CV Applications

Automated visual inspection

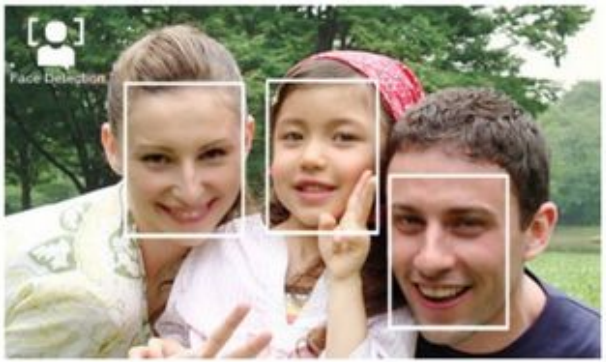
Smart Shopping Carts

Toshiba Tech IS-910T 2013

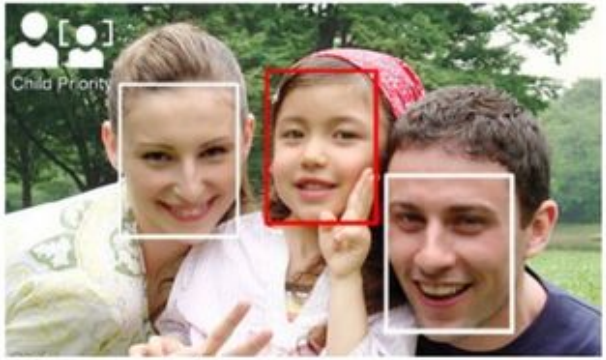
DataLogic LaneHawk LH4000 2012

Face detection

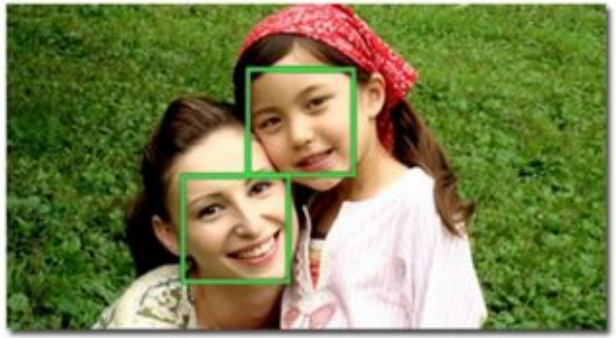
Multiple applications are possible once the face is detected



Sony Cyber-shot



Age recognition



Smile recognition

Face ID

Identifying plants

Computer Vision also tries provide machines/computers some superhuman capabilities like seeing in dark

BMW 5 series

BMW night vision

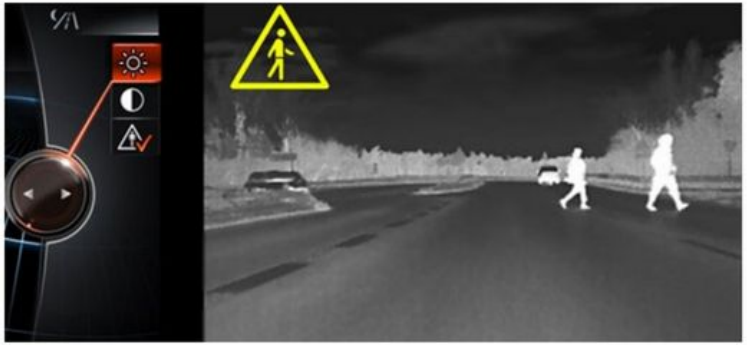
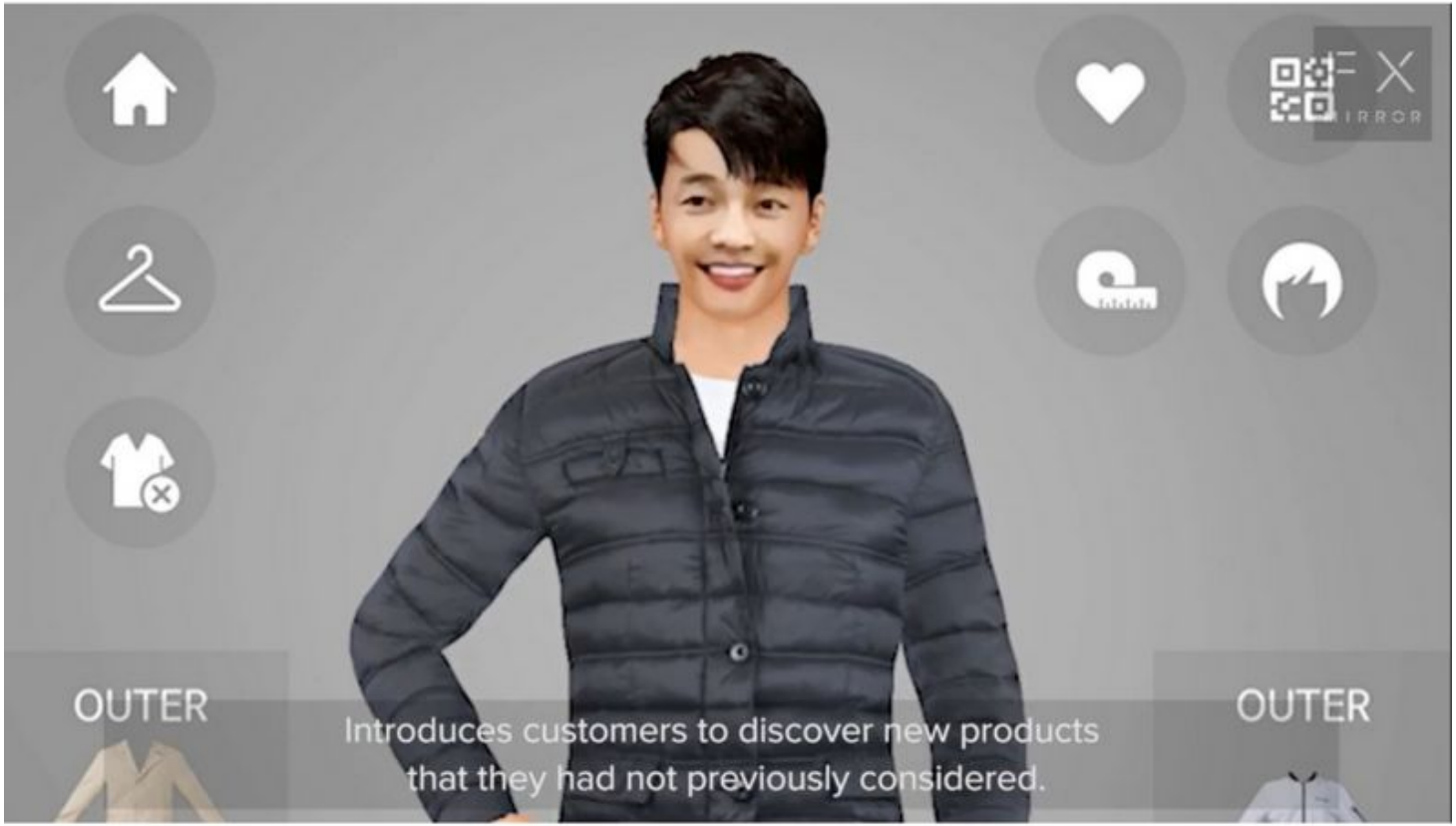


Image stitching

Virtual Fitting



Deep Fake

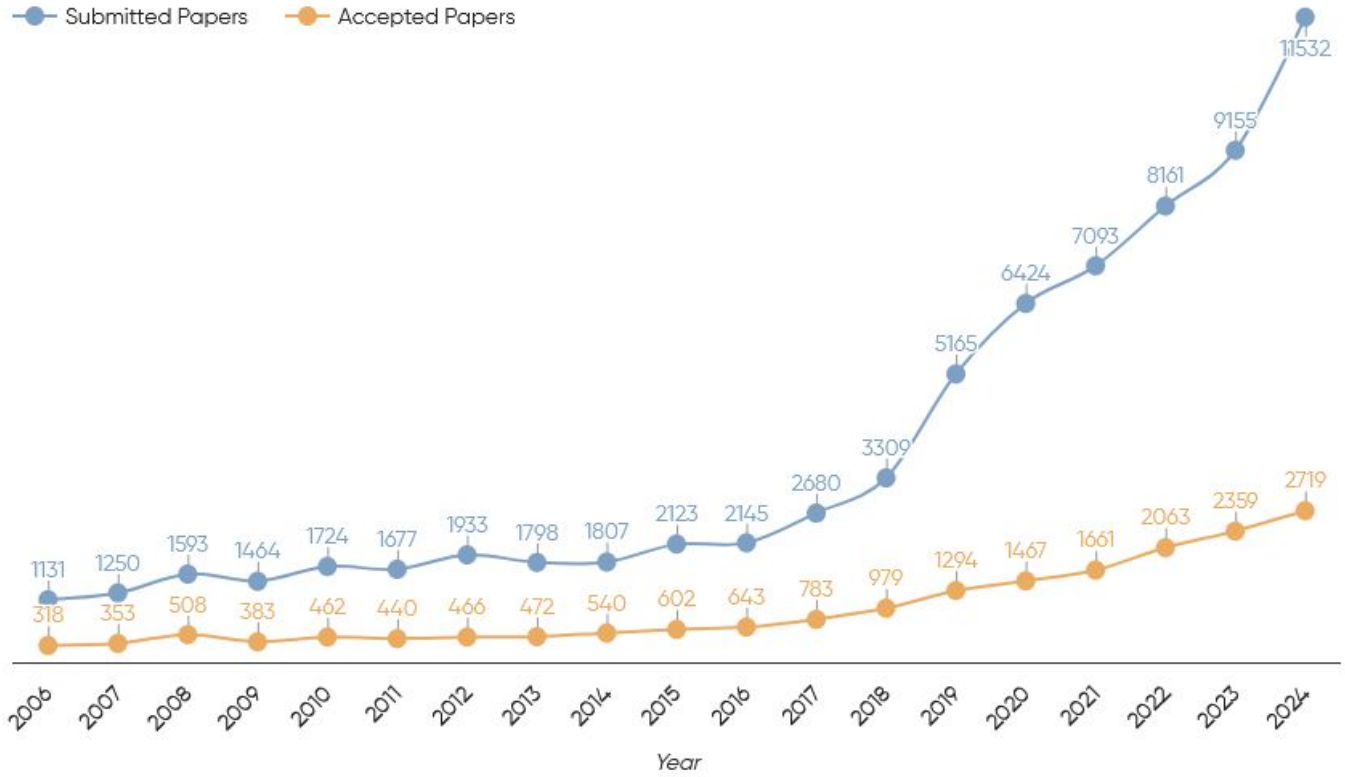
It’s a good time to do computer vision

Industry aggressively hiring CV graduates, or even students!

(~~strong~~ dominant industrial presence at conferences for recruitment)

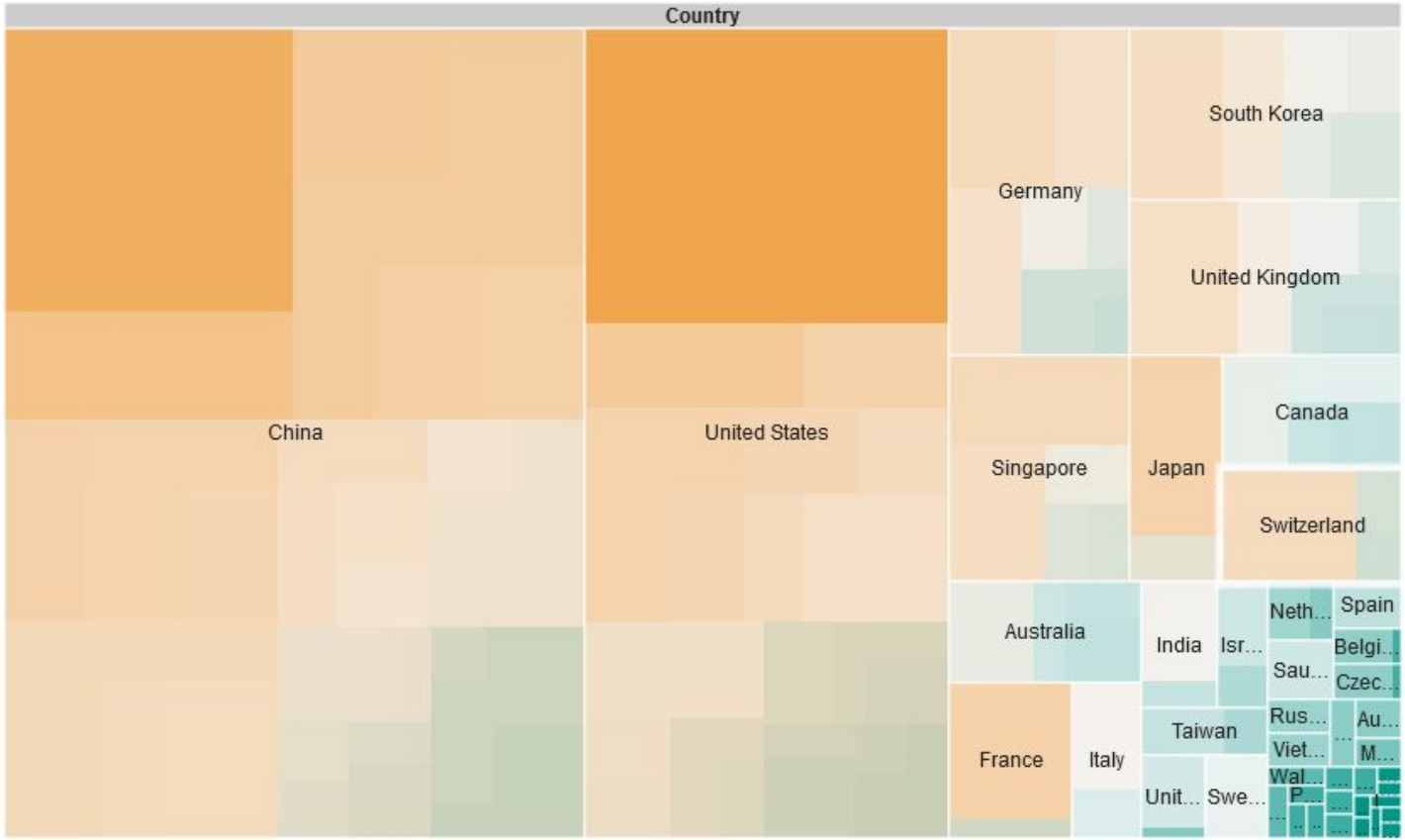
CVPR is Growing Faster than Ever

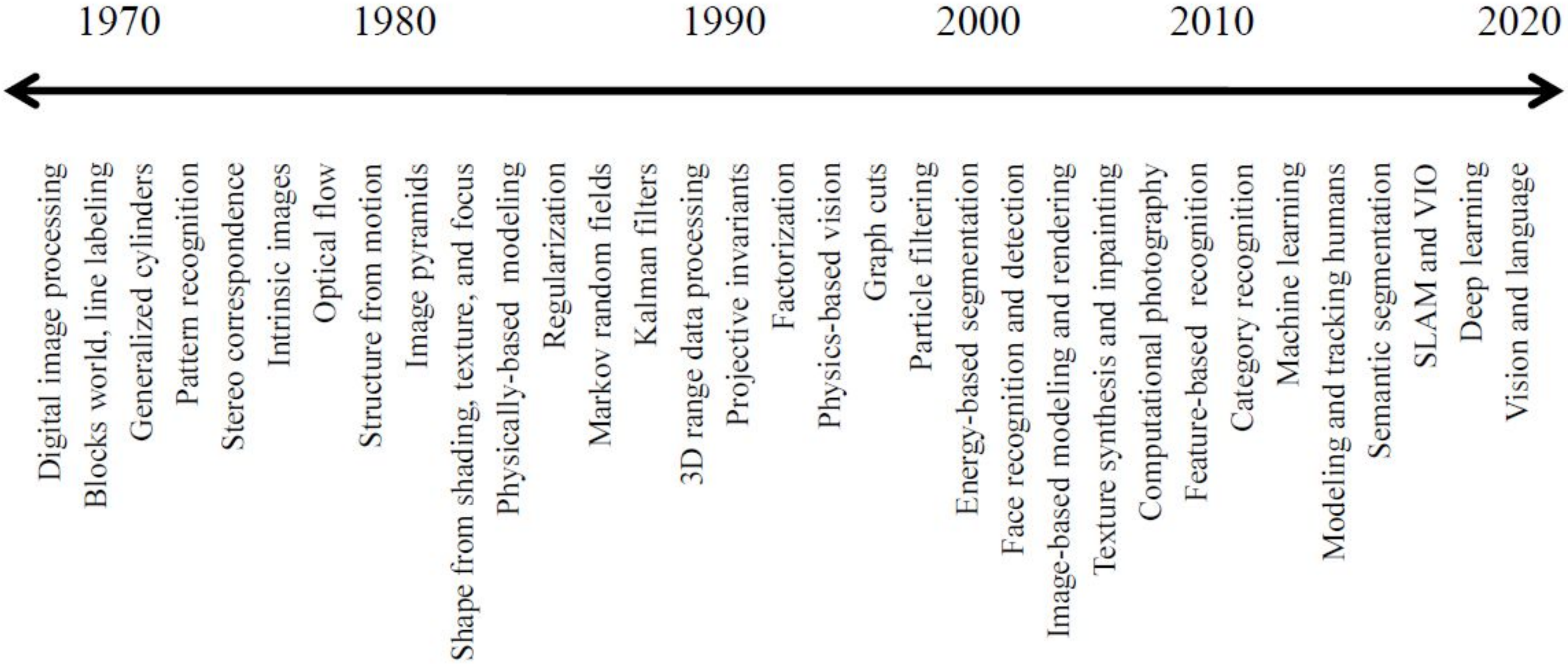
At this year's CVPR, there were 11 532 submissions out of which 2 719 were accepted to the main conference (23.6% acceptance rate). Looking at the historical data we see that not only is CVPR still growing, it is in fact growing faster than ever — compared to last year there was 25+% increase in paper submissions, which is more than last two years combined.



ICCV & ECCV are other two top CV conferences

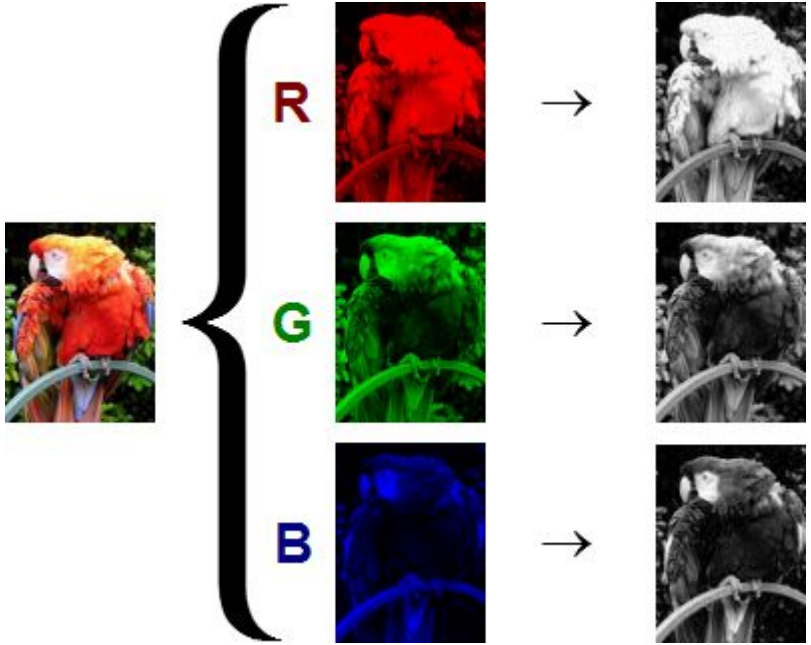
Country-wise participation





A rough timeline of some of the most active topics of research in computer vision.

Image Formation



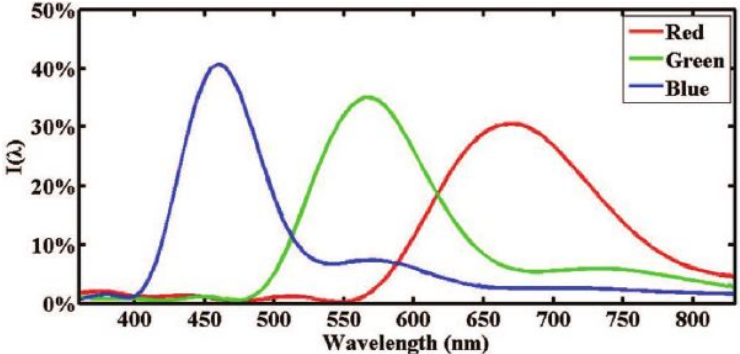
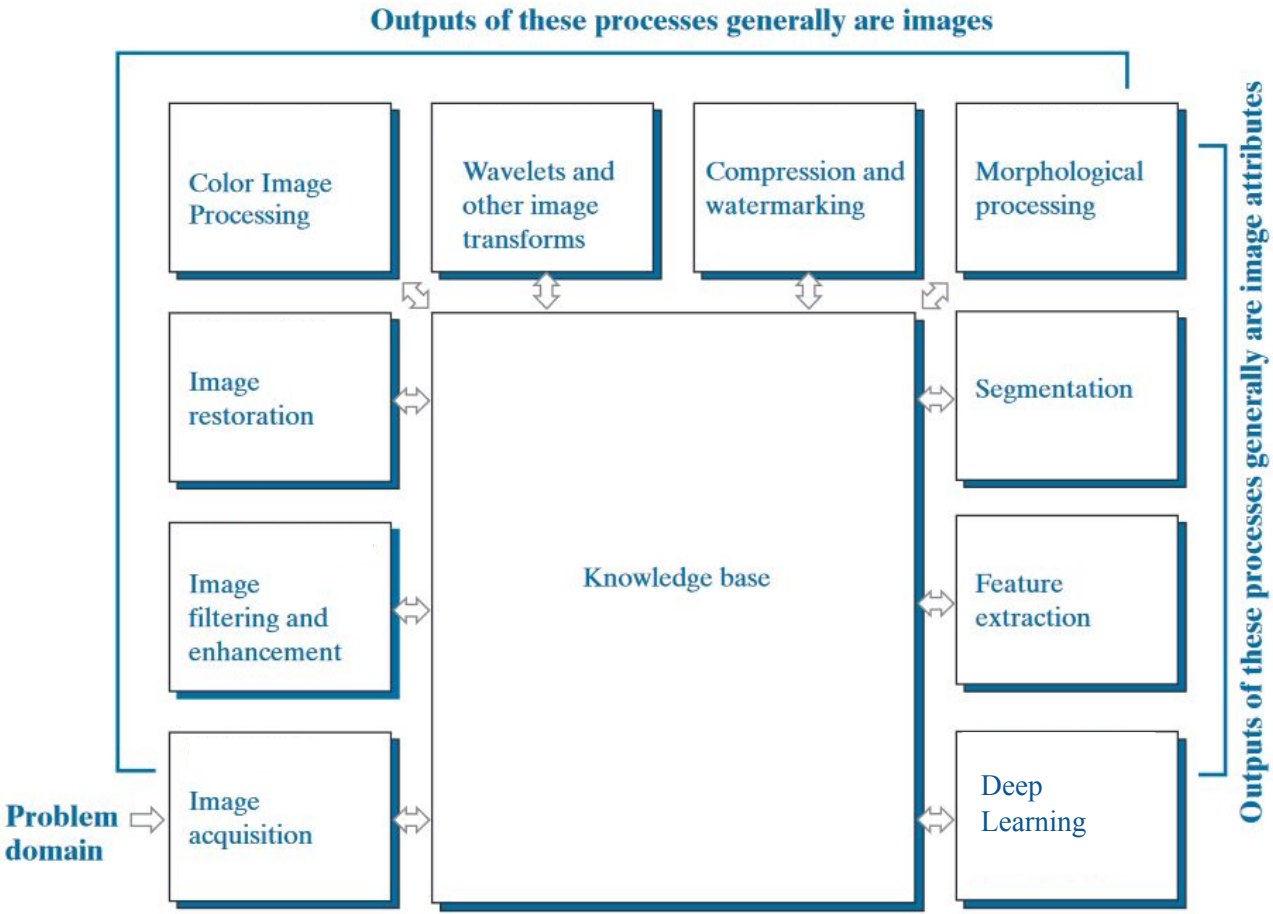
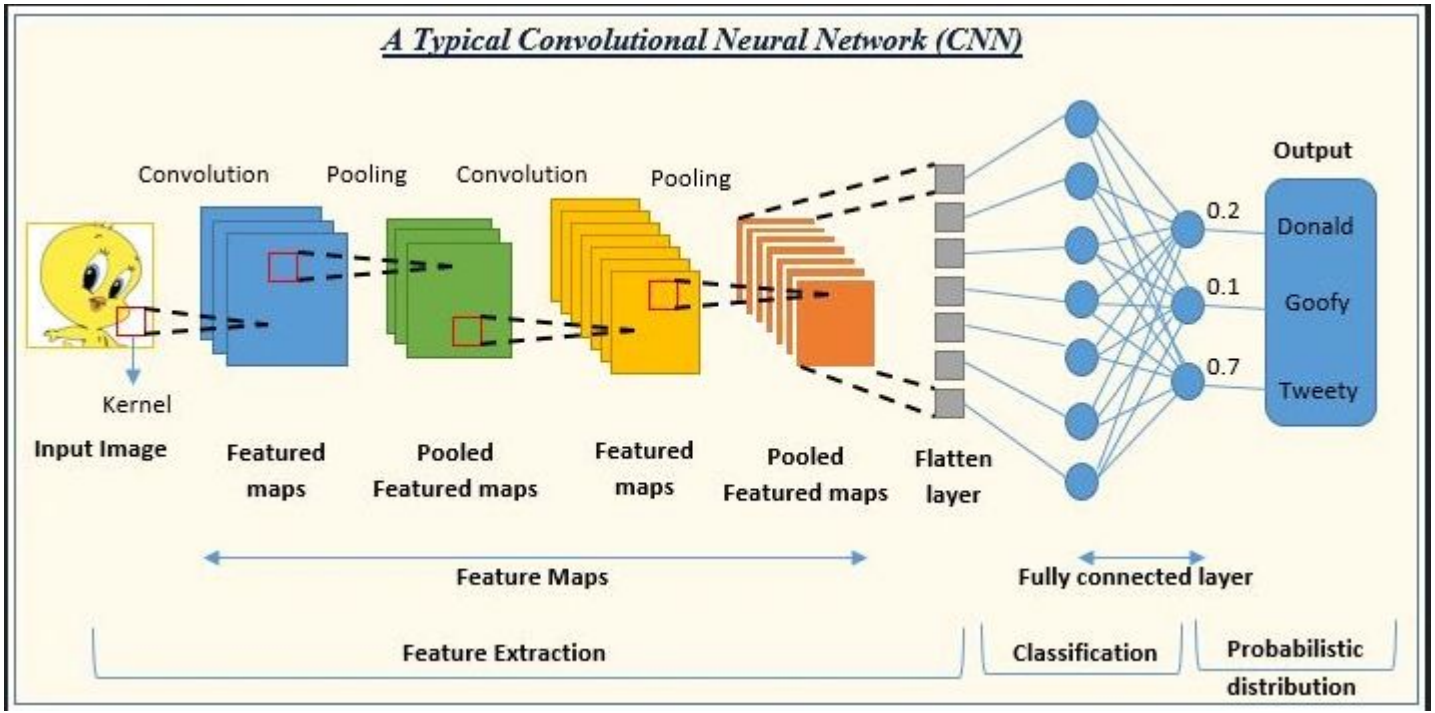


Image Processing

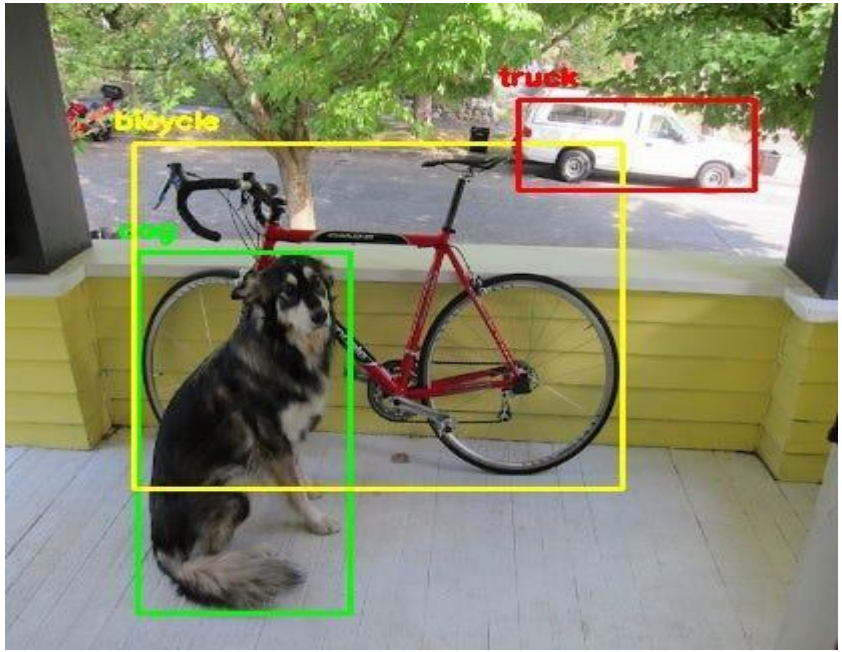


CNNs

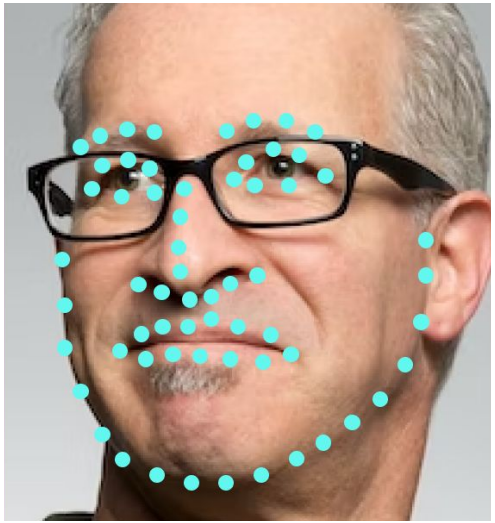

A Typical Convolutional Neural Network (CNN)
Convolution
Pooling
Convolution
Pooling
Output
0.2
0.1
0.7
Donald
Goofy
Tweety
Kernel
Input Image
Featured maps
Pooled Featured maps
Featured maps
Pooled Featured maps
Flatten layer
Feature Maps
Fully connected layer
Feature Extraction
Classification
Probabilistic distribution

Semantic Segmentation

Object Detection

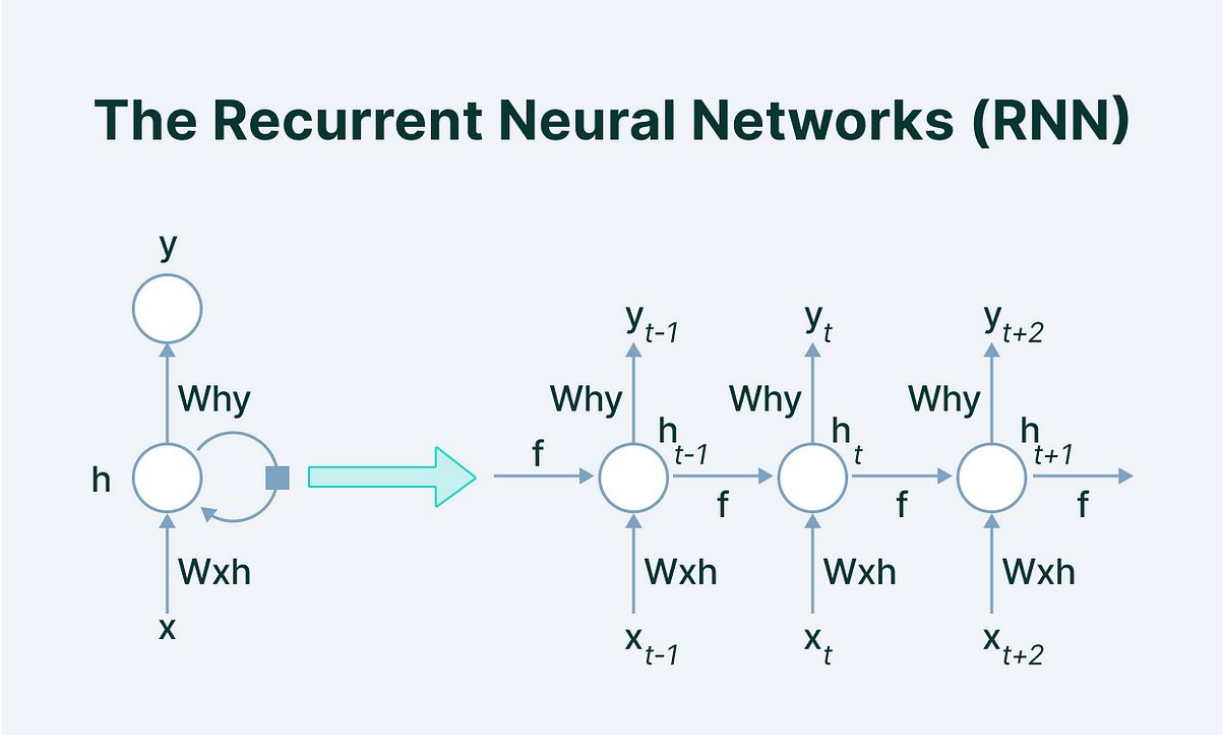


Landmark Detection & Pose Estimation



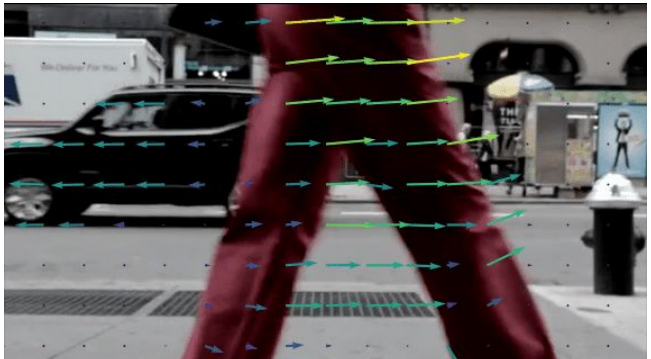
RNNs

Useful for handling sequential information (for e.g., videos)



Video Understanding & Motion Estimation

Class	Score
Swing Dancing	0.52
Salsa Dancing	0.39
Holding hands	0.03
Walking	0.01
...	...

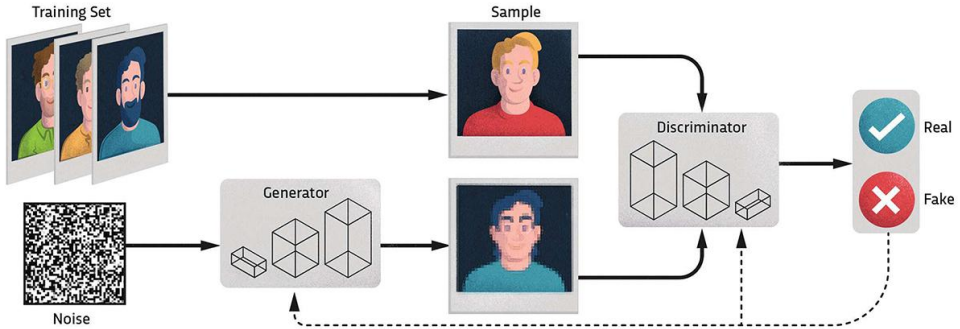


GANs (Generative adversarial networks)

Generator tries to generate fake images

Discriminator tries to classify whether the image is real or fake

Output of the discriminator provides the feedback to the generator how successful it has been in fooling the discriminator



Depth Estimation & 3D Reconstruction