Basic Deep Learning in Computer Vision

Day 1

Morning: Introduction to Computer Vision and Image Libraries

Afternoon: Image Preprocessing and Augmentation

[Course Materials]

https://bit.ly/bdlcv_2021

Getting to know you

Step 1: Go to the following url

https://pollev.com/jimmygoh831

Introduction of Trainer

Mr Hew Ka Kian

Name
Hew Ka Kian

Telegram
@

Email
hew_ka_kian@rp.edu.sg

Mr Jimmy Goh

Name
Jimmy Goh

Telegram
@jimmygohrp

Email
jimmy_goh@rp.edu.sg

Programme

	Morning	**Afternoon**
Day 1	• Computer Vision • Image Libraries • Activity 1: Getting Started with Libraries	• Image Preprocessing • Image Augmentation • Activity 2: Image preprocessing • Activity 3: Image Augmentation
Day 2	• Basic of Neural Network • Activity 4: Building NN with Python • Introduction to Keras • Activity 5: Building NN with Keras	• Image Convolution • Convolution Neural Network (CNN) • Activity 6: Create and use CNN • Quiz

Prerequisites

- Programming Language
  - Python
- Tools / System used
  - Google’s Colab
    - https://colab.research.google.com/

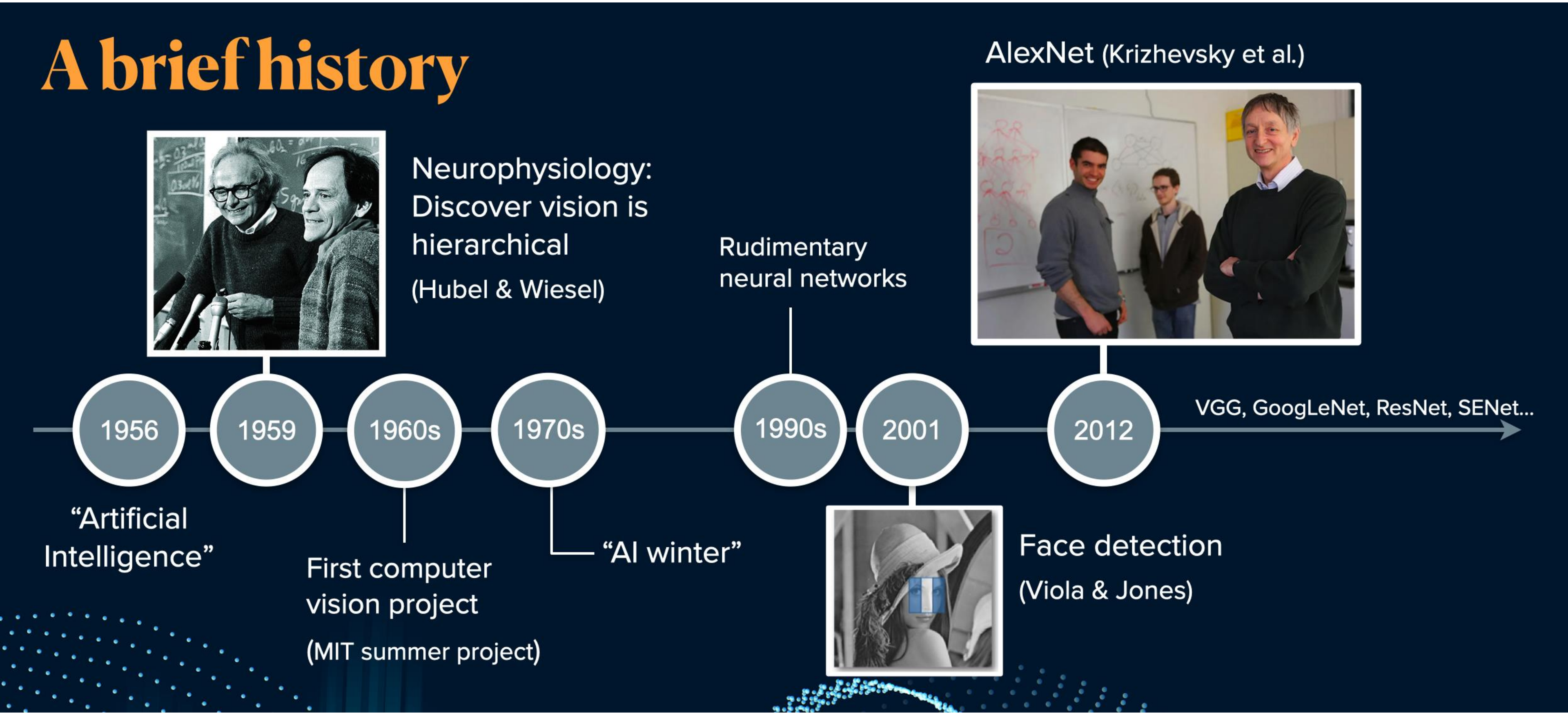
What is Computer Vision?

Definition of CV

- Computer vision is a field of computer science that works on enabling computers to **see**, **identify** and **process** images in the same way that human vision does, and then **provide appropriate output**.
  - Reference: https://www.techopedia.com/definition/32309/computer-vision

- Computer vision tasks include methods for **acquiring**, **processing**, **analyzing** and **understanding** digital images, and extraction of high-dimensional data from the real world in order to **produce numerical or symbolic information**, e.g., in the forms of decisions.
  - Reference: https://en.wikipedia.org/wiki/Computer_vision

Brief History of CV



A brief history

1956 — "Artificial Intelligence"

1959 — Neurophysiology: Discover vision is hierarchical (Hubel & Wiesel)

1960s — First computer vision project (MIT summer project)

1970s — "AI winter"

1990s — Rudimentary neural networks

2001 — Face detection (Viola & Jones)

2012 — AlexNet (Krizhevsky et al.)

VGG, GoogLeNet, ResNet, SENet...

Ref: https://katiehuang1221.medium.com/an-introduction-to-computer-vision-131826e2b512

Components of CV

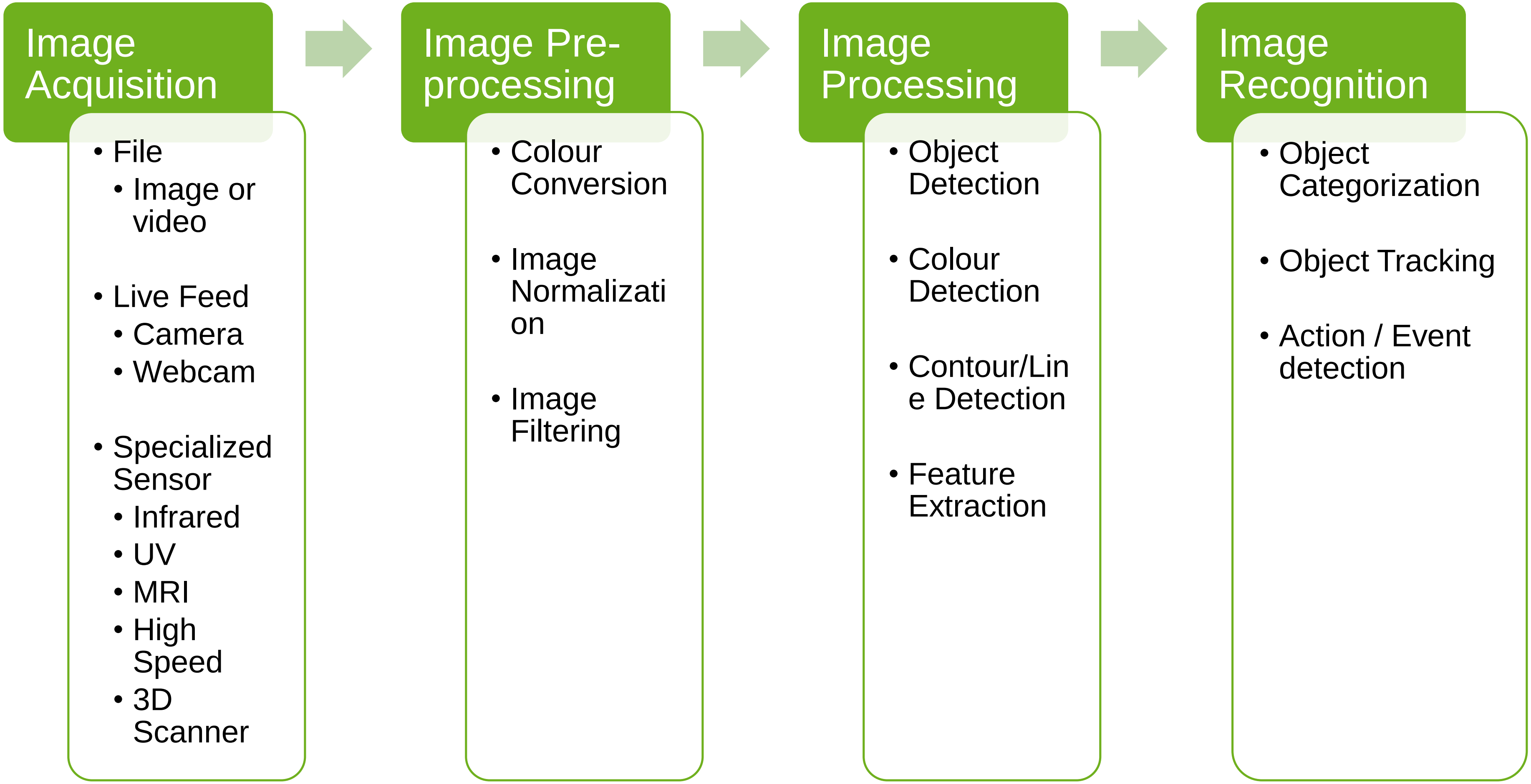


Image Acquisition

- File
  - Image or video
- Live Feed
  - Camera
  - Webcam
- Specialized Sensor
  - Infrared
  - UV
  - MRI
  - High Speed
  - 3D Scanner

Image Pre-processing

- Colour Conversion
- Image Normalization
- Image Filtering

Image Processing

- Object Detection
- Colour Detection
- Contour/Line Detection
- Feature Extraction

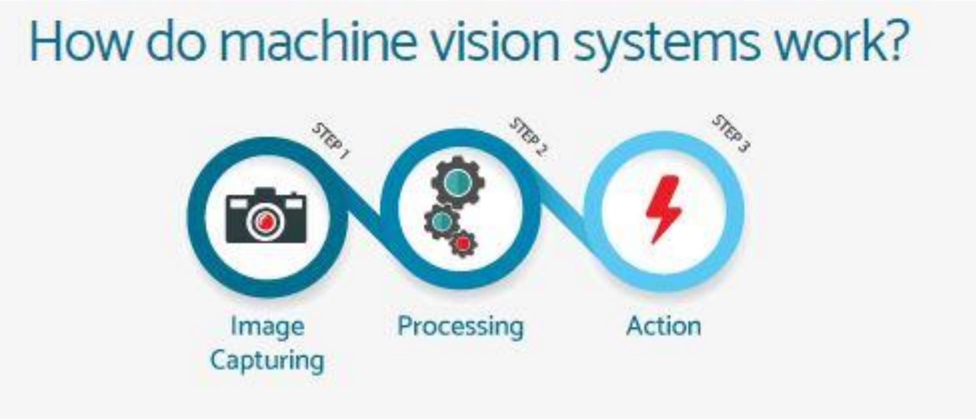
Image Recognition

- Object Categorization
- Object Tracking
- Action / Event detection

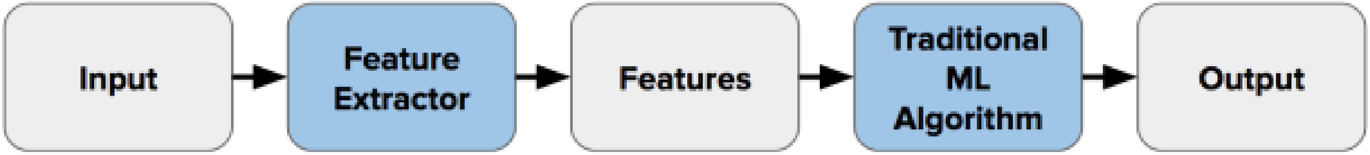
CV vs Image Processing (optional)

https://www.youtube.com/watch?v=9-8Js62wzQs

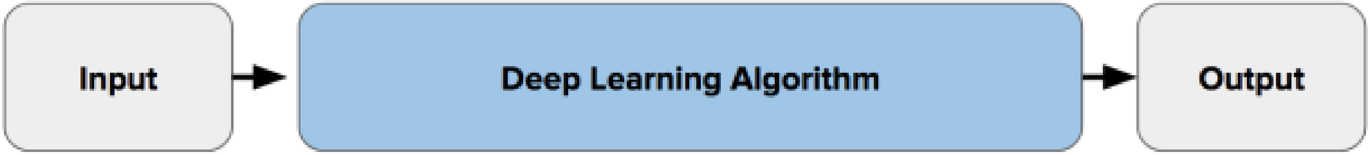
CV Workflow



Ref: https://research.aimultiple.com/machine-vision/



Traditional Machine Learning Flow

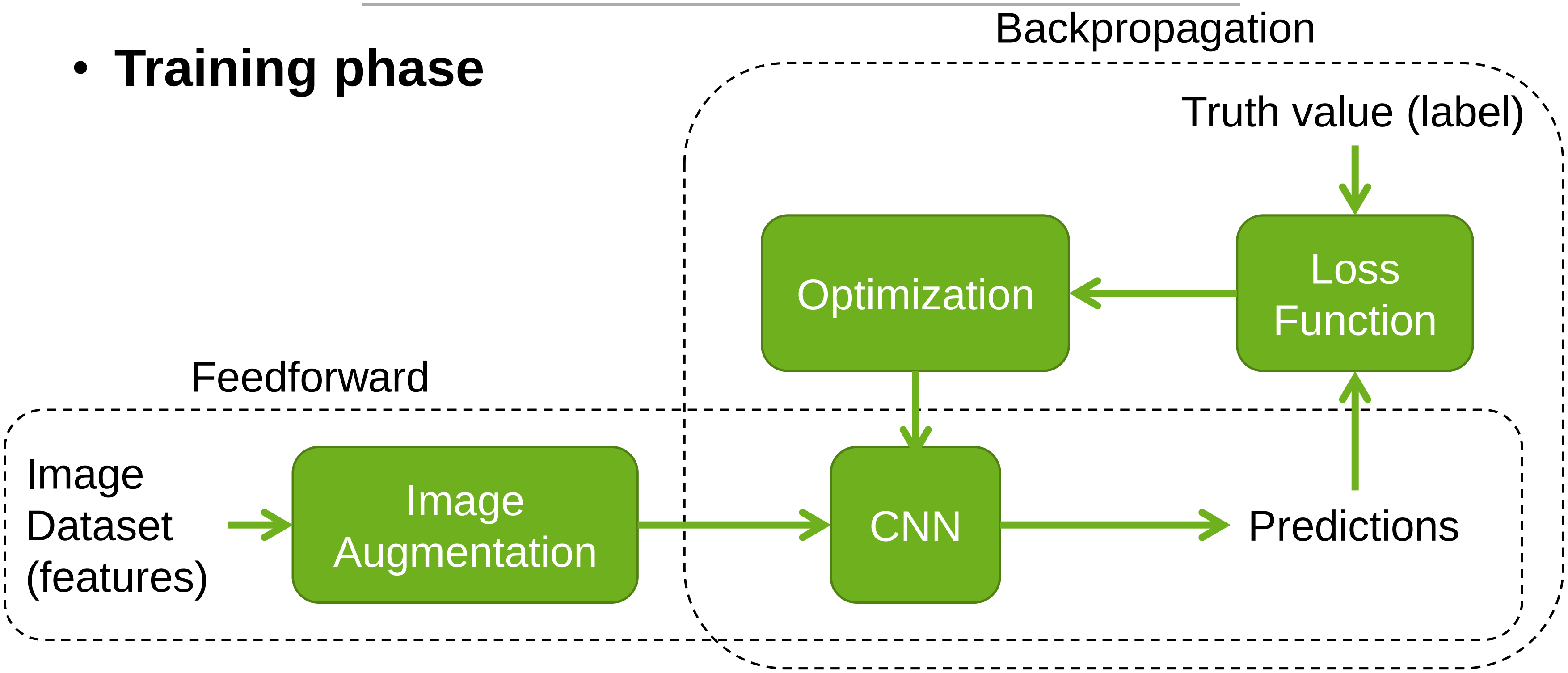


Deep Learning Flow

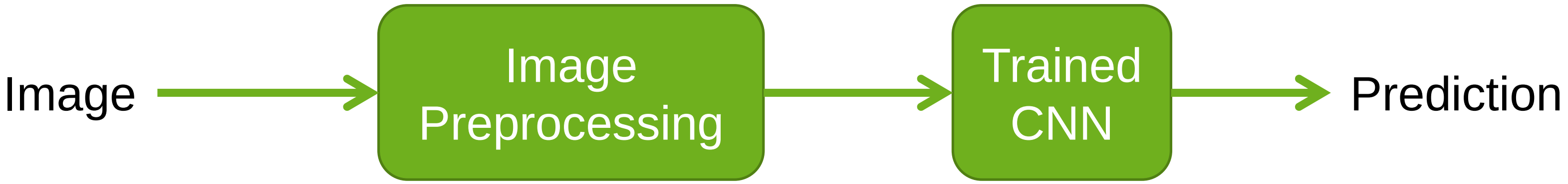
Ref: https://towardsdatascience.com/roadmap-to-computer-vision-79106beb8be4

CV Deep Learning Workflow

- **Training phase**



- **Inference Phase**



CV Tasks

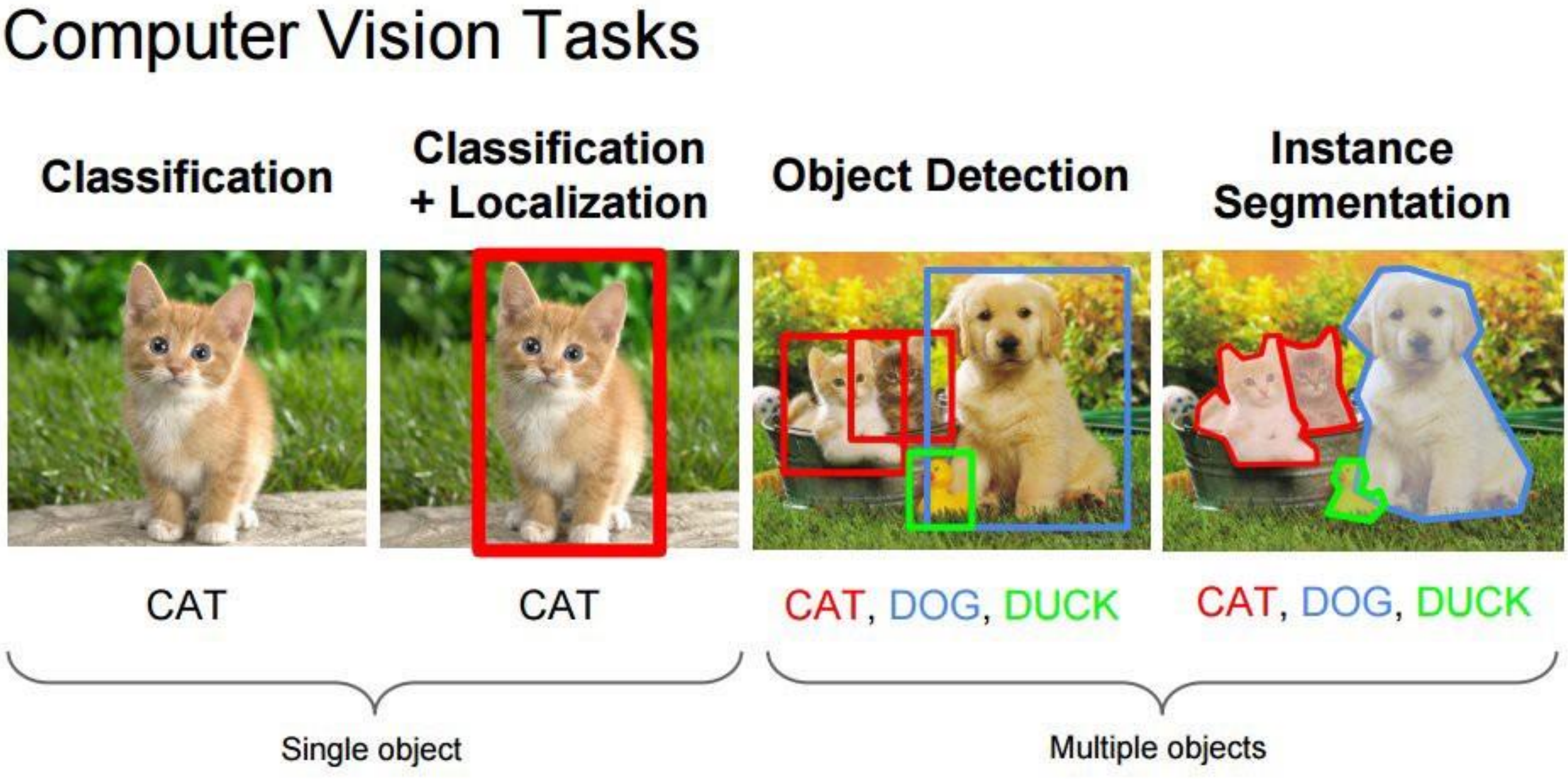
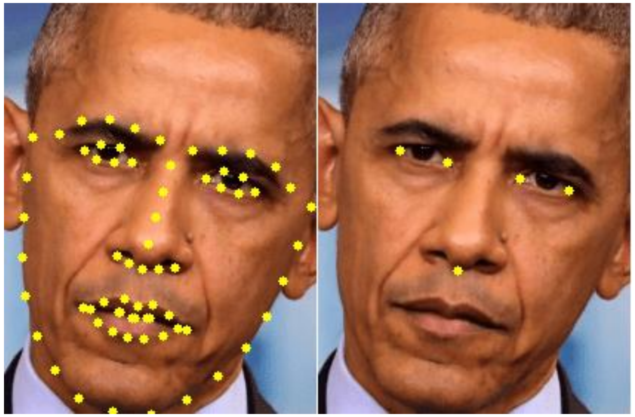


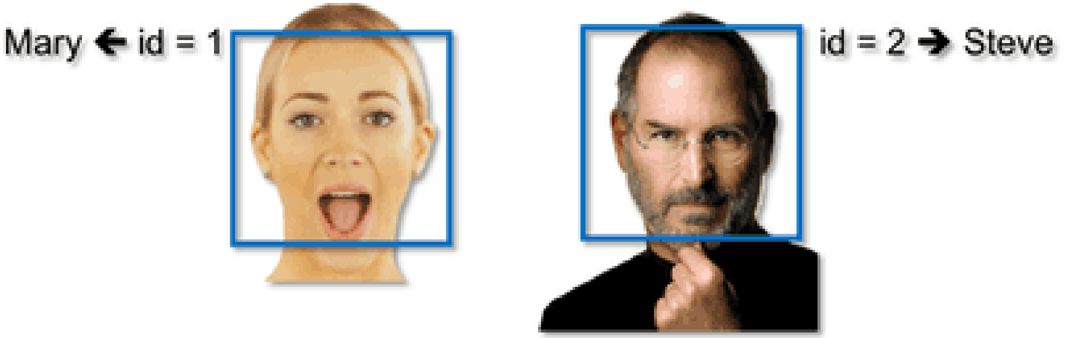
Image: https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e

CV Tasks

- **Object Landmark – hand, face or pose**



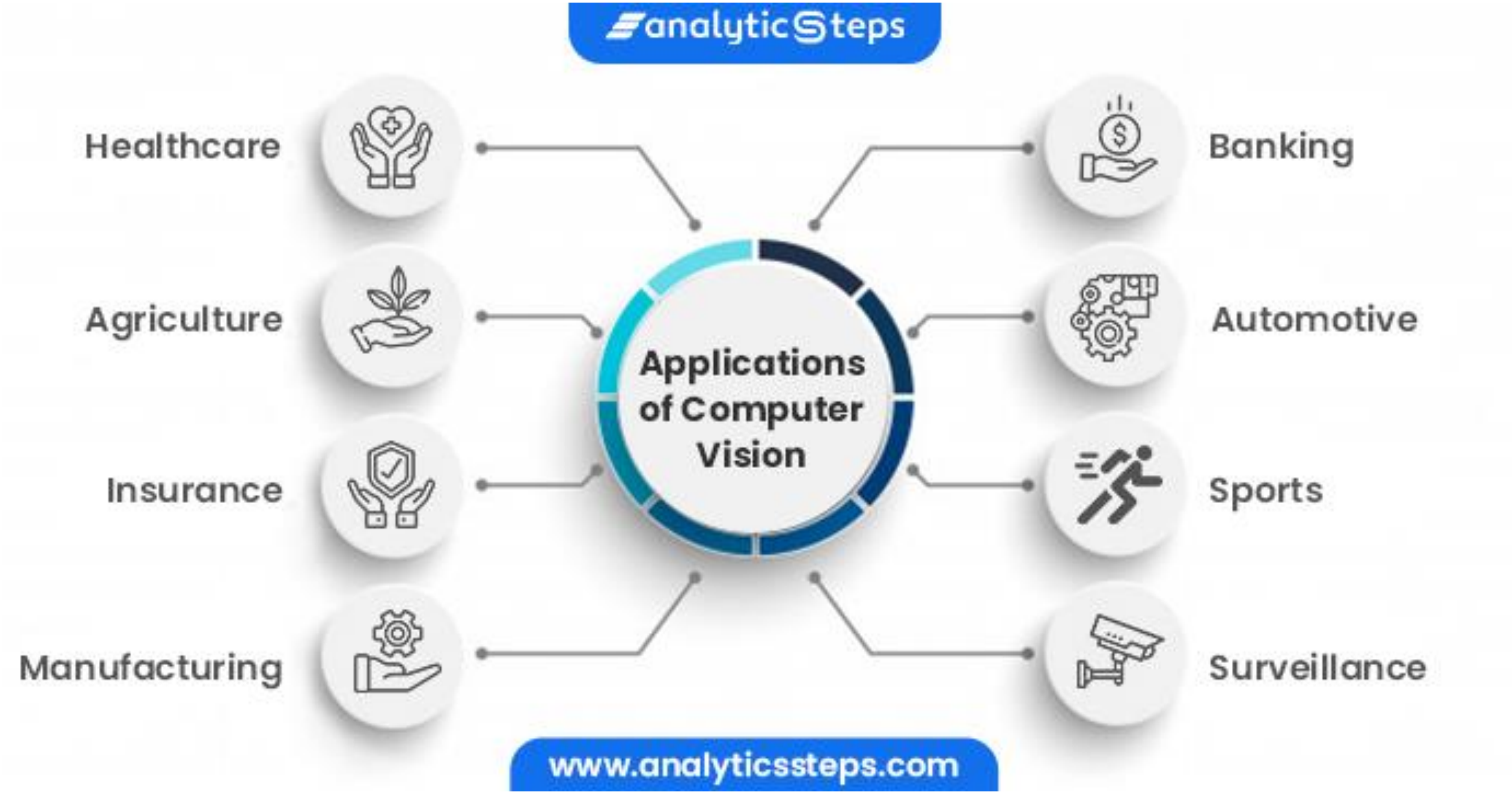
- **Facial Recognition**



CV Applications

- **Image Recognition**
  - Image Classification
  - Image Localization
  - Object Detection
  - Image Segmentation
  - Face Recognition
- **Motion Analysis**
  - Object Tracking
  - Event Alert
- **Scene Reconstruction**
  - Image Stitching
  - 3D Modelling
  - Image Captioning
- **Image Restoration**
  - Motion Blur
  - Sensor Noise

CV Applications



Ref: https://www.analyticssteps.com/blogs/8-popular-computer-vision-applications

CV Applications (optional)

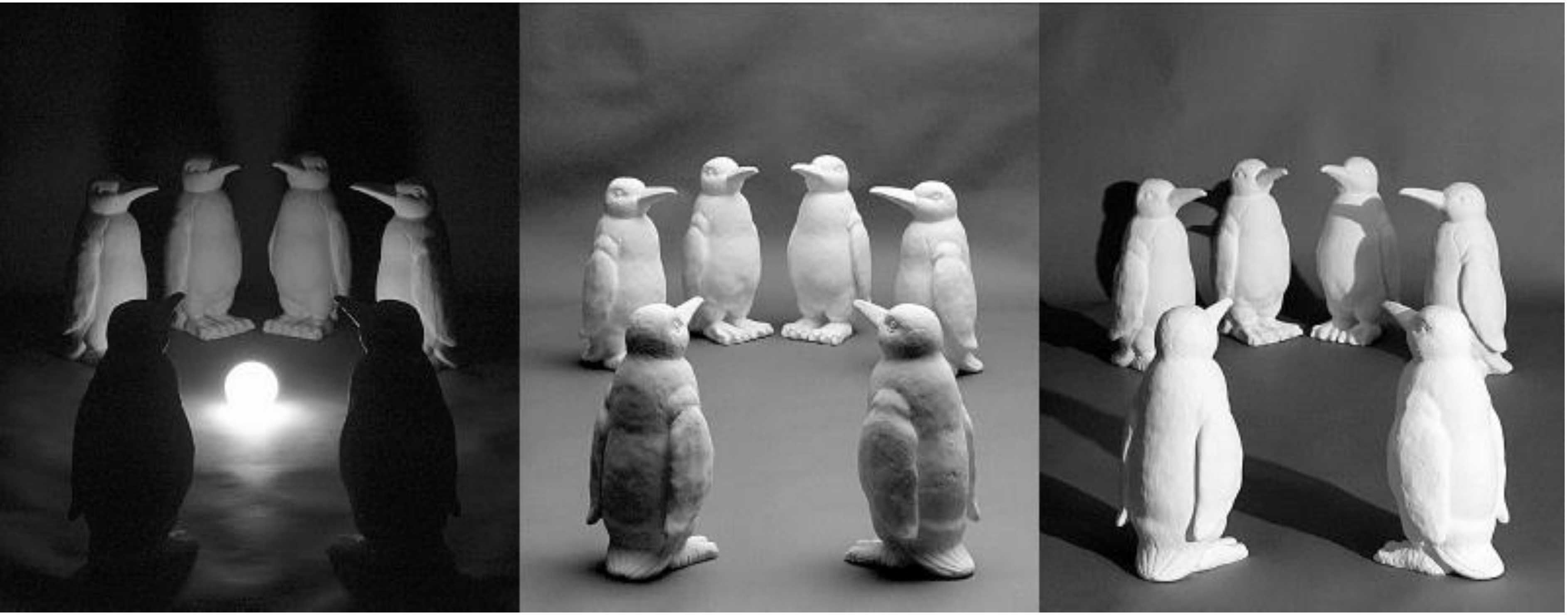
https://www.youtube.com/watch?v=aDpnaxPAmtU

Challengers of CV

- **Viewpoint Variation**

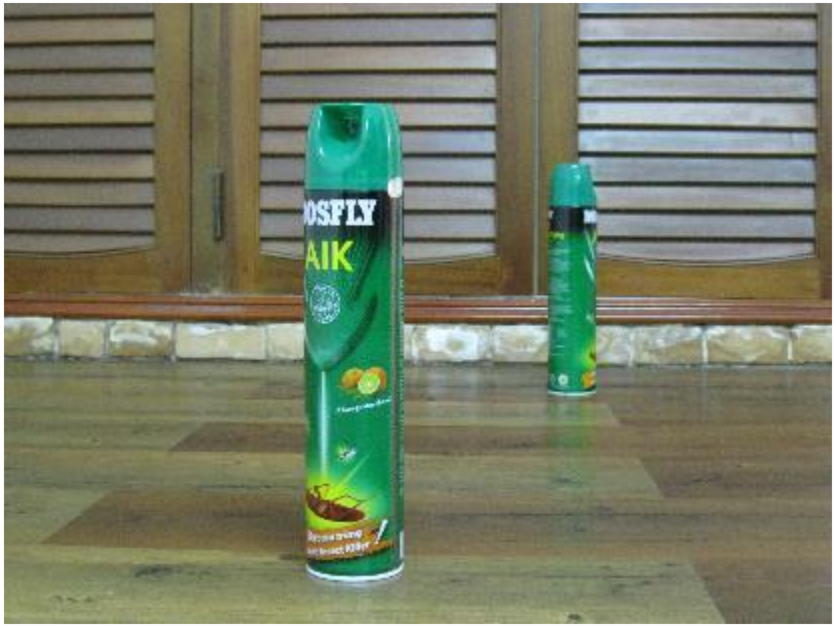
Challengers of CV

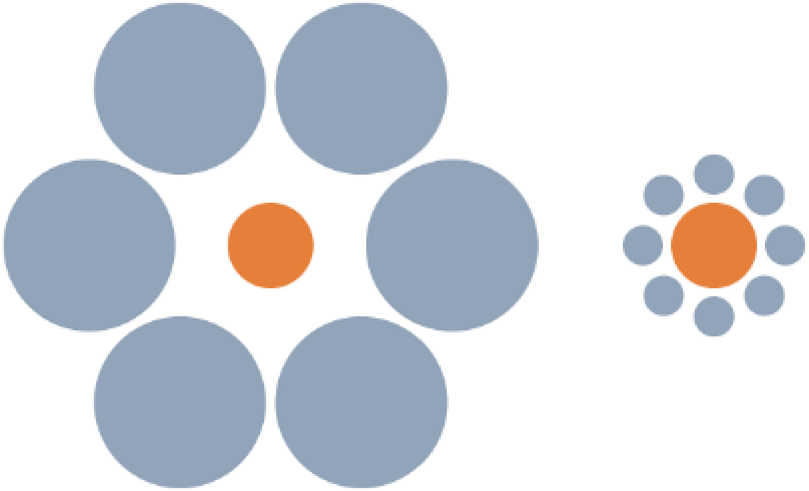
- **Illumination**



Challengers of CV

- **Scale**



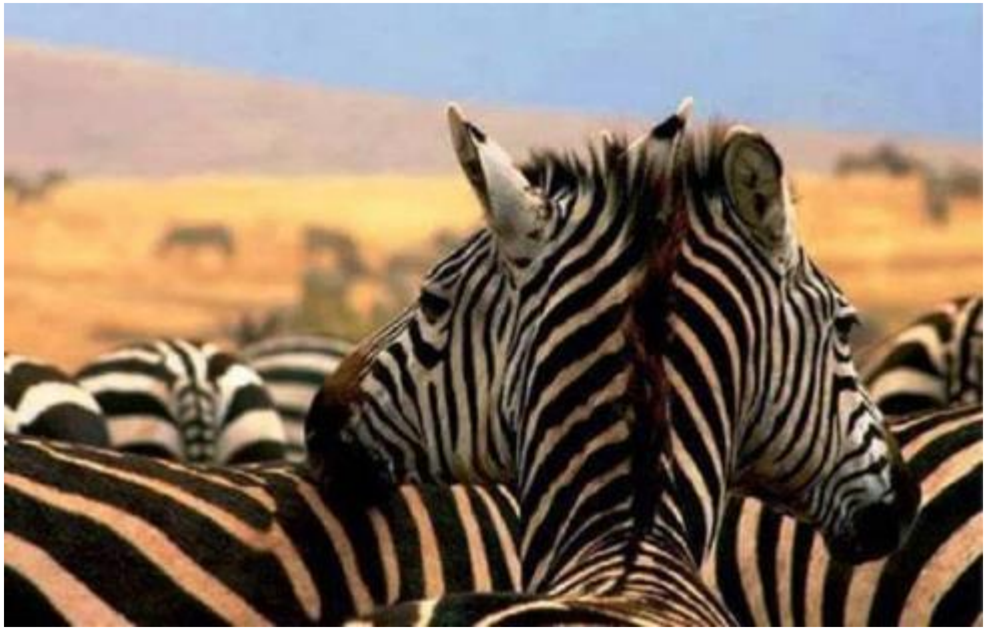


Challengers of CV

- **Deformation**

Challengers of CV

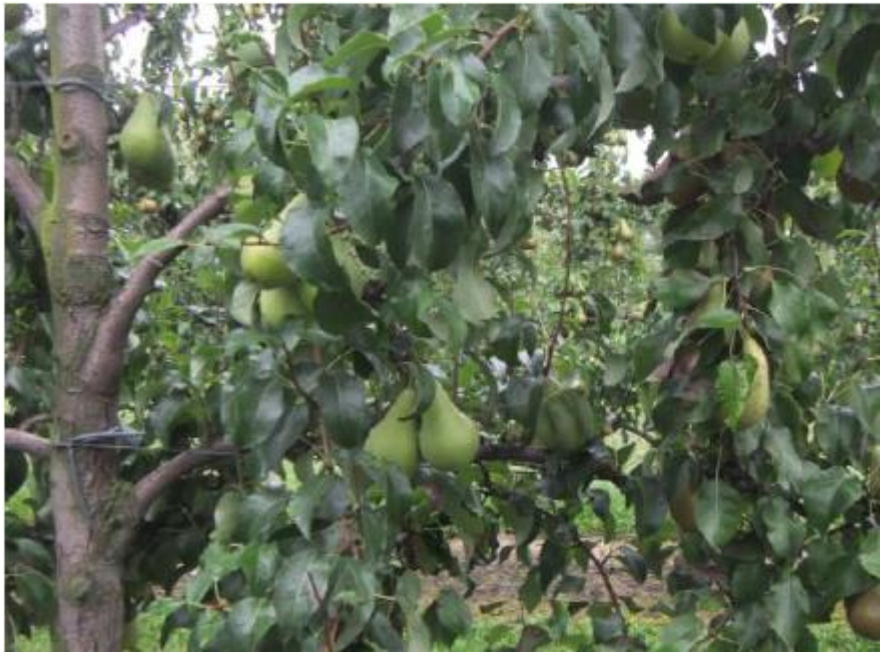
- **Occlusion**

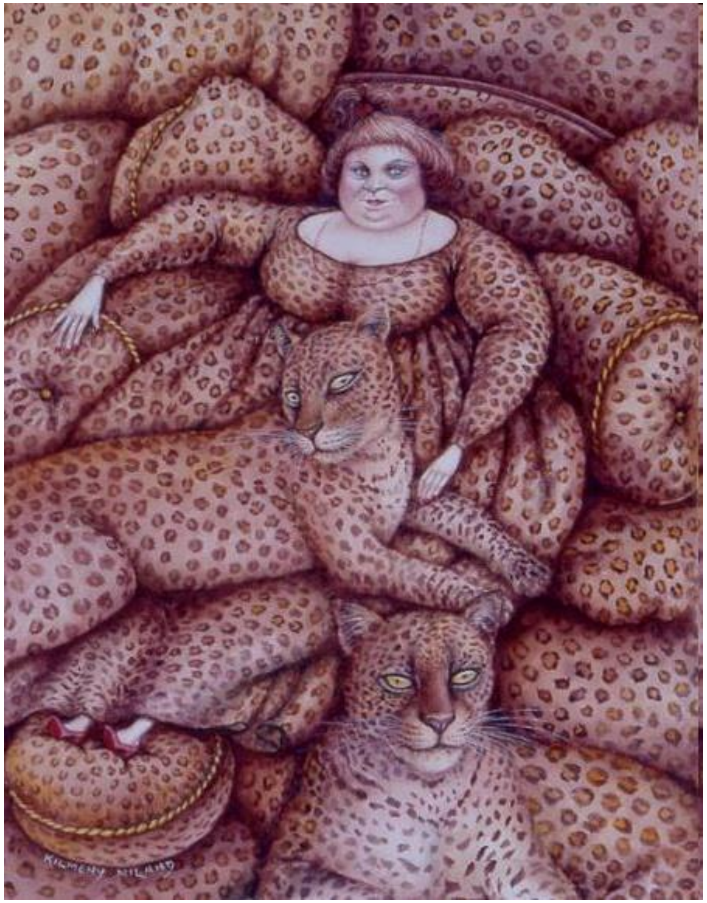


T-Rex Illusion

Challengers of CV

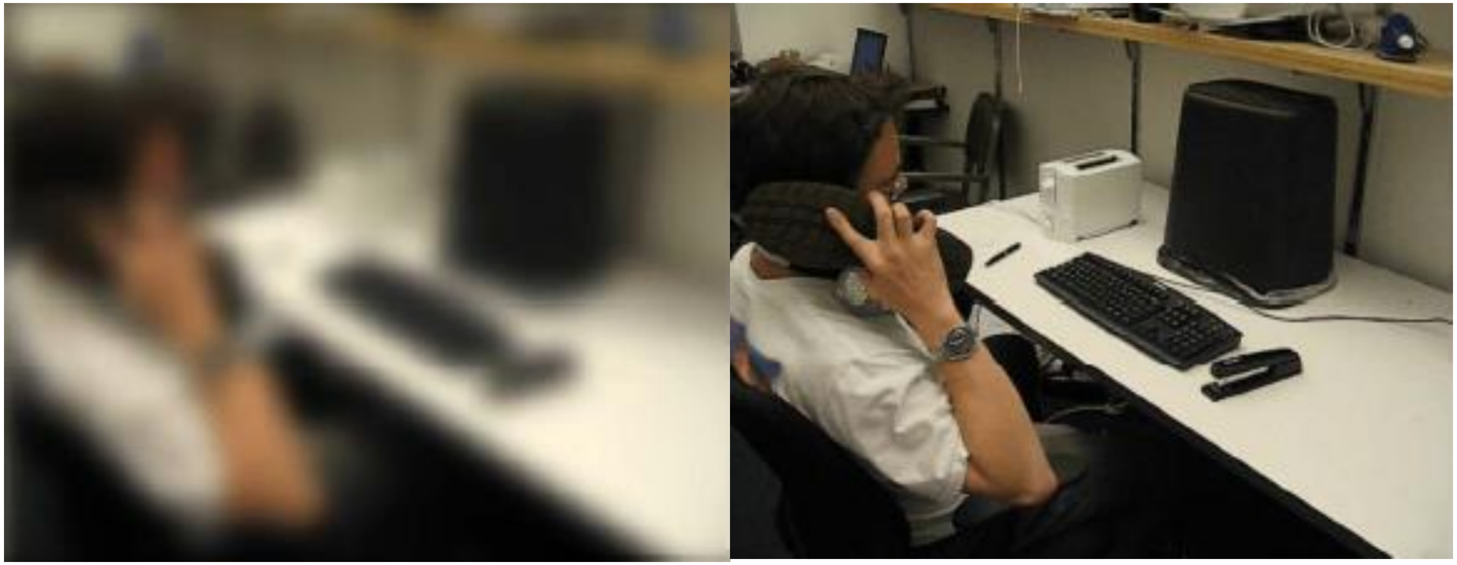
- **Background Clutter**





Challengers of CV

- **Local ambiguity**



Challengers of CV

- **Motion**

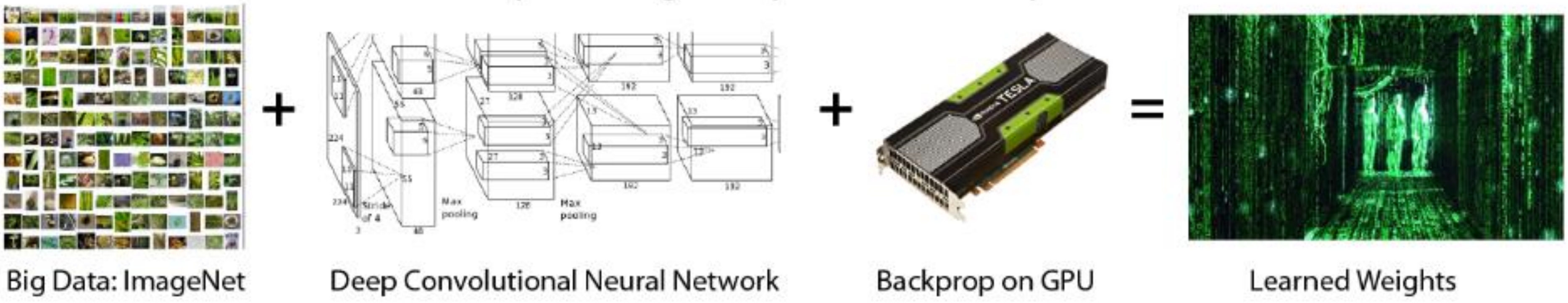
Challengers of CV

- **Intra-class variation**

Success of CV

- **Large dataset:** images are readily available on the internet
- **Good algorithm:** Deep Convolution Neural Network for better and more accurate CV tasks
- **Fast computing**: huge reduction of computing time using GPU (Graphical Processing Unit)

The Deep Learning "Computer Vision Recipe"



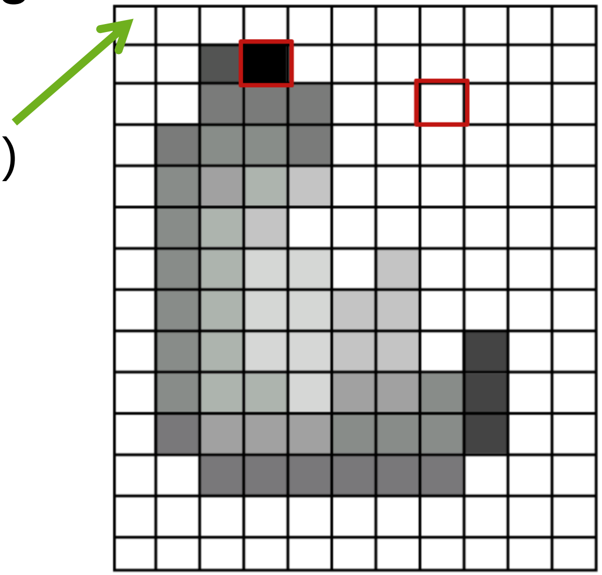
15 Mins Break

Image Basics

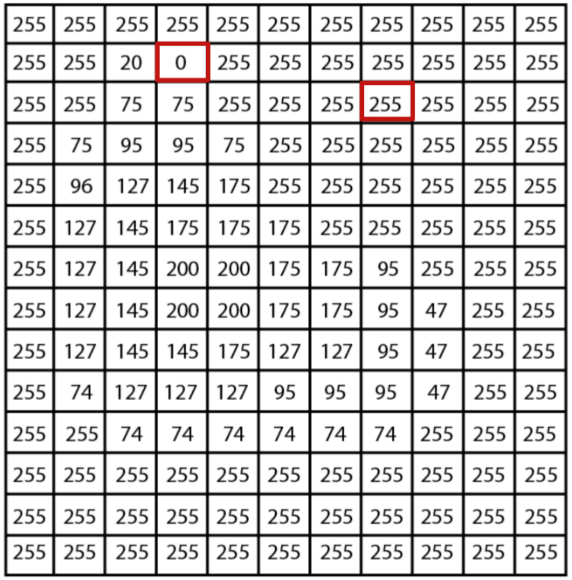
Image representation

- **Pixel (picture element) represents one point in an image. It is the smallest unit within an image.**
- **The intensity of an pixel is represented by one value (grey scale) or multiple values (colour)**
- **A 2-D collections (array) of these value(s) represents an image**

Origin (0, 0)



=



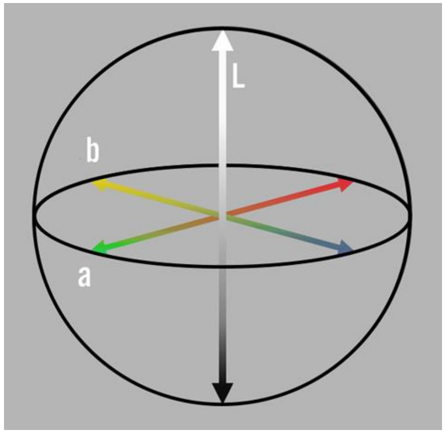
255	255	255	255	255	255	255	255	255	255	255
255	255	20	0	255	255	255	255	255	255	255
255	255	75	75	255	255	255	255	255	255	255
255	75	95	95	75	255	255	255	255	255	255
255	96	127	145	175	255	255	255	255	255	255
255	127	145	175	175	175	255	255	255	255	255
255	127	145	200	200	175	175	95	255	255	255
255	127	145	200	200	175	175	95	47	255	255
255	127	145	145	175	127	127	95	47	255	255
255	74	127	127	127	95	95	95	47	255	255
255	255	74	74	74	74	74	74	255	255	255
255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255

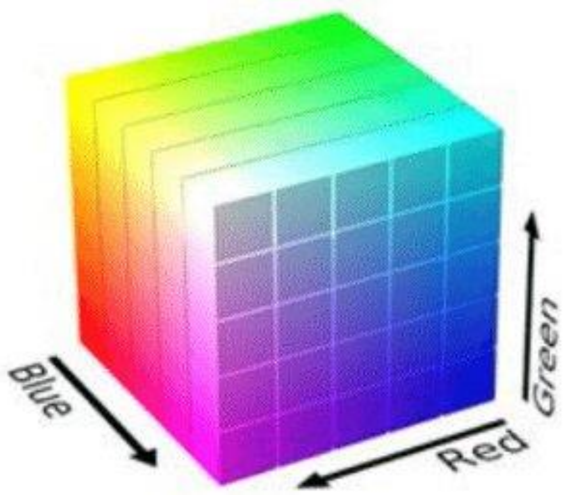
Image taken from:
https://edtech.engineering.utoronto.ca/object/2d-image-digital-representation

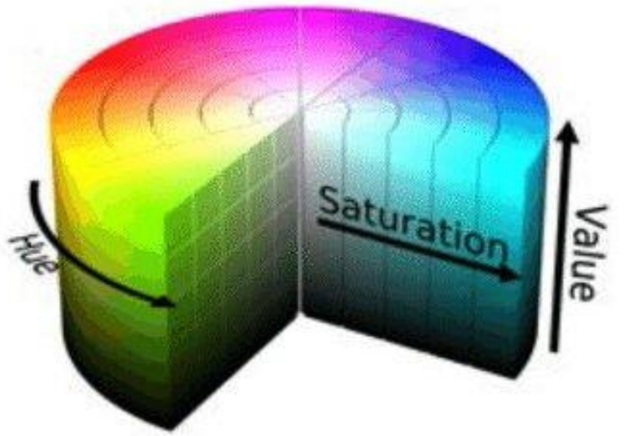
0 = black; 255 = white

Image Modes/channels

- **Colour representations using multiple values**
  - RGB – Red, Green Blue
  - HSV – Hew, Saturation Value
  - BGR – Blue, Green, Red
  - CMYK – Cyan, Magenta, Yellow, Black
  - LAB - Lightness, Red/Green coordinate(A), Yellow/Blue coordinate(B)
  - YCbCr – Lumination (Y), Chroma-Blue, Chroma-Red







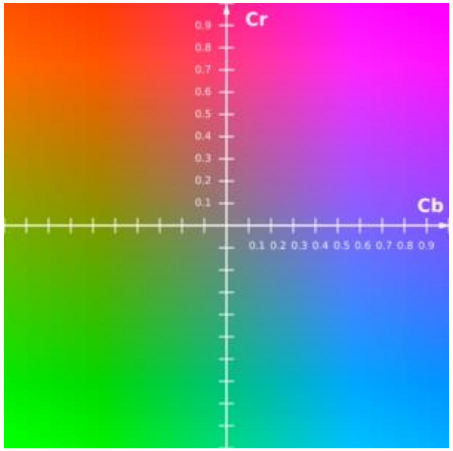


Image Compression

- **Storing the pixel values without any compression (RAW format) will result in a very large filesize.**
  - An colour image of 800 by 600 pixels requires (800x600x3) 1.44MB.
- **Image compression greatly reduced the size required to store the image**
- **Most common compression**
  - **JPEG** (Joint Photographic Experts Group): It is used for digital images, especially for those images which are composed of digital photography.
  - **PNG** (Portable Network Graphics): These files are commonly used to store graphics for web images. PNG was developed to enhance the non-registered replacement for Graphics Interchange Format.
  - **GIF** (Graphics Interchange Format): It is a file format for storing graphical images up to 256 colors. PNG is based on a lossless compression method, which makes higher quality output. PNG was created as a more powerful option to the GIF file format.
  - **TIFF/ TIF** (Tagged Image File): These files can be saved in a collection of color formats and many forms of compression. TIFF file is used to maintain image integrity and clarity. It is often used for professional photography.

Pro and Cons of each format (Optional)

https://www.youtube.com/watch?v=M247l1LktG8

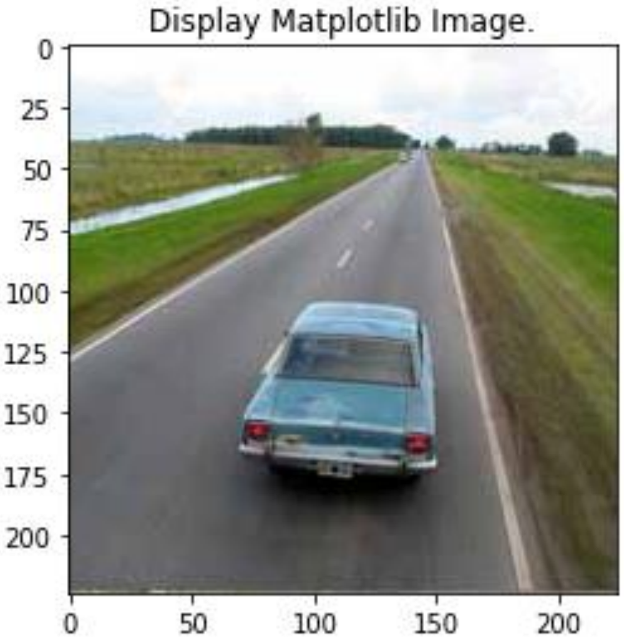
Python Image Libraries

- **Python Image Library (pil / pillow)**
  - Default image library that comes with python
- **OpenCV (https://opencv.org/about/)**
  - OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products.
  - Most popular image library
- **Imutil (https://github.com/jrosebr1/imutils)**
  - A series of convenience functions to make basic image processing functions such as translation, rotation, resizing, skeletonization, and displaying Matplotlib images easier with OpenCV and ***both*** Python 2.7 and Python 3.
- **Matplotlib (https://matplotlib.org/)**
  - Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python.
  - Use together with Jupyter Notebook (Co-lab) to display image within the notebook.
  - Note: All libraries provide image display feature, however Matplotlib is the only library that is integrated into notebook.

Activity 1 – Getting Started on Libraries

Car-224.jpg

PIL
OpenCV
Matplotlib
Imutils



Exercises:

- Work with more images concurrently.
- Create an array (list) of images.

Step 1:
Watch and listen to the instructor's demonstration

15 mins

Step 2:
Work through the activities

Individual Activity

20 mins

LUNCH BREAK

60 mins Lunch Break

Lunch break xx:xx – yy:yy

Image Pre-processing

Image Pre-processing

- **Image pre-processing is traditionally mean the enhancement of an image for analysis and processing.**
- **In the context of Deep Learning and CNN, image pre-processing prepares the images to be used in Deep Learning or CNN**
- **The image pre-processing includes**
  - Image resizing
  - Image mode/channel conversion
  - Image normalization

Image resizing

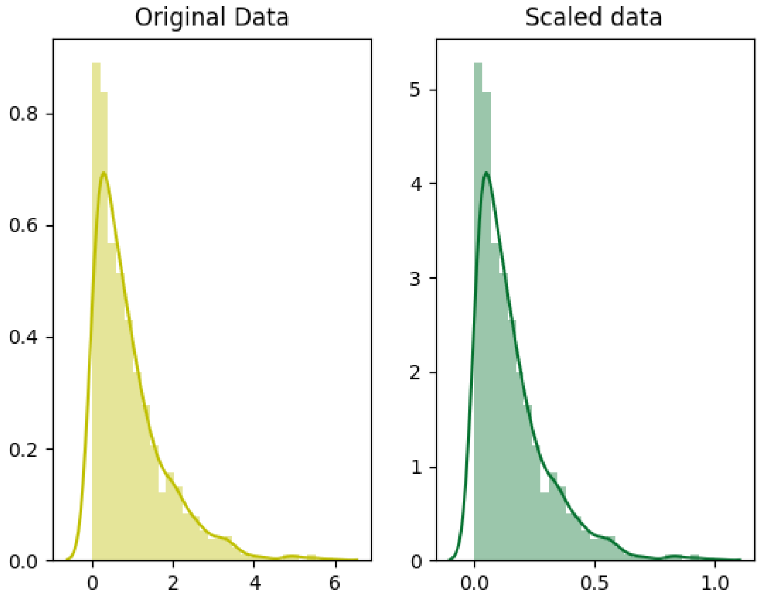
- **As the number of nodes of the input layer is pre-defined and constant, the number of pixels (image size) inputting into the CNN is also constant.**
- **Solutions:**
  - Resize the image to the correct dimensions
  - Crop the image with the correct dimensions to obtain a ‘subset’ of the image
  - Slicing (cutting) a large image into smaller images with correct dimensions

Image mode/channel Conversion

- **There are several different colour modes. Different colour mode is useful for different purpose of the CV tasks. Example:**
  - Greyscale is better for shape or contour detection regardless of colour.
  - HSV / LAB is better for colour detection.
  - RGB/BGR is better for complex image analysis as a whole.
- **Most libraries load in RGB or BGR format, image colour conversion is required if a different format is used.**

Image Normalization

- **Normalization is the process of ‘fitting’ the data into a range of 0 to 1 while maintaining the histogram of the data.**



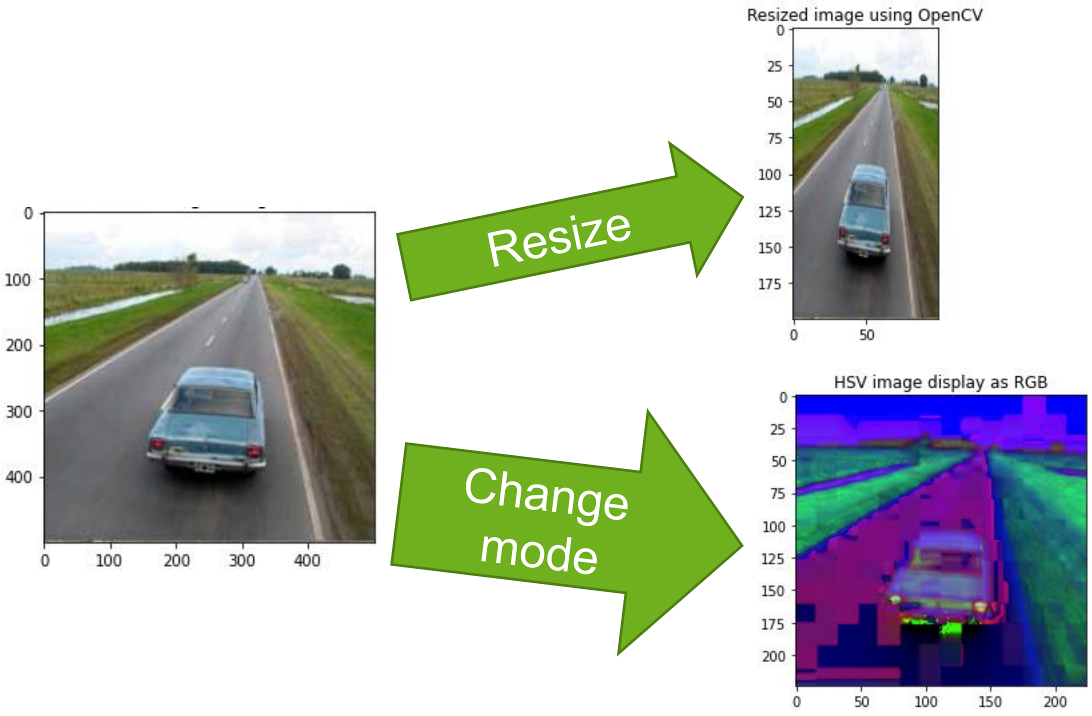
Normalization Formula

$$X_{normalized} = \frac{(X - X_{minimum})}{(X_{maximum} - X_{minimum})}$$

- **Since minimum value for pixel is 0 and maximum value is 255. The normalization formula for image is simplified to:**

$$Pixel_{normalized} = \frac{Pixel}{255}$$

Activity 2 – Image Preprocessing



Exercises:

- Play around the image preprocessing functions with different parameters and values

Step 1:
Watch and listen to the instructor's demonstration

15 mins

Step 2:
Work through the activities

Individual Activity

20 mins

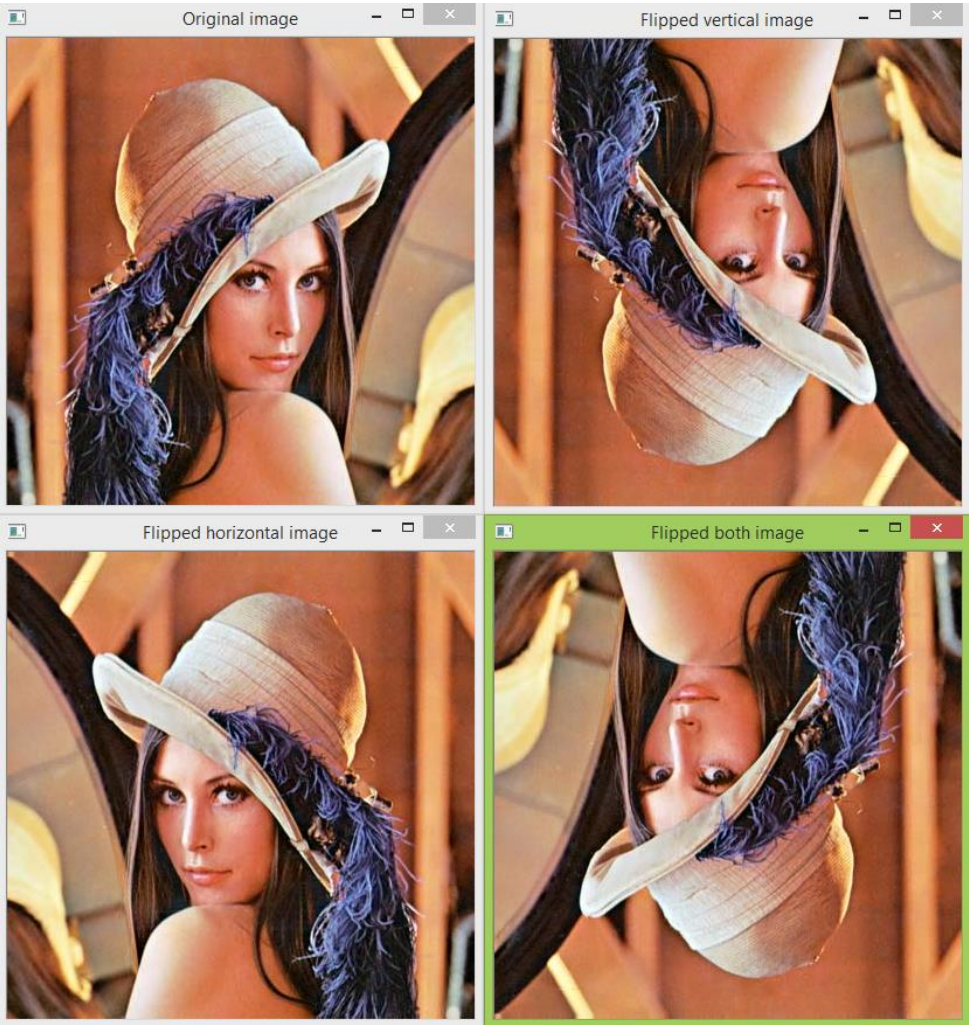
15 Mins Break

Image Augmentation

Image Augmentation

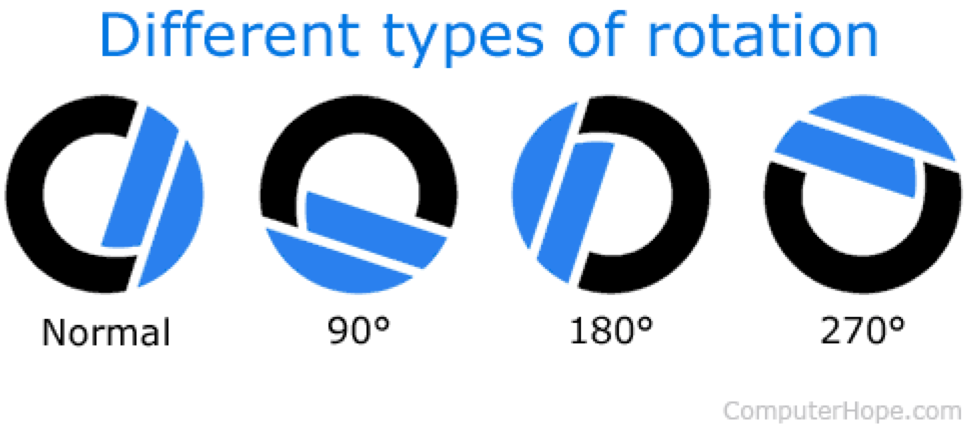
- **Training of the CNN requires large amount of images**
- **More images can be generated through applying some primitive image pre-processing methods on the existing images.**
- **This process is known as Image Augmentation**
- **Augmentation the can be performed**
  - Flipping
  - Rotating
  - Zooming
  - Shifting
  - Channel shifting (change colour)

Image Flipping

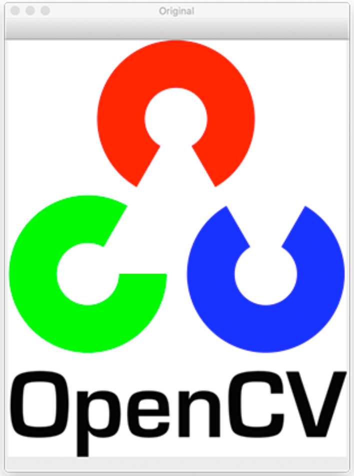


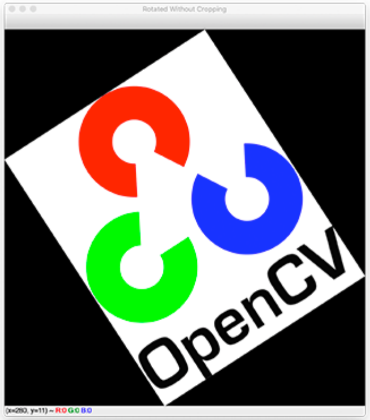
https://techtutorialsx.com/2019/04/21/python-opencv-flipping-an-image/

Image Rotating



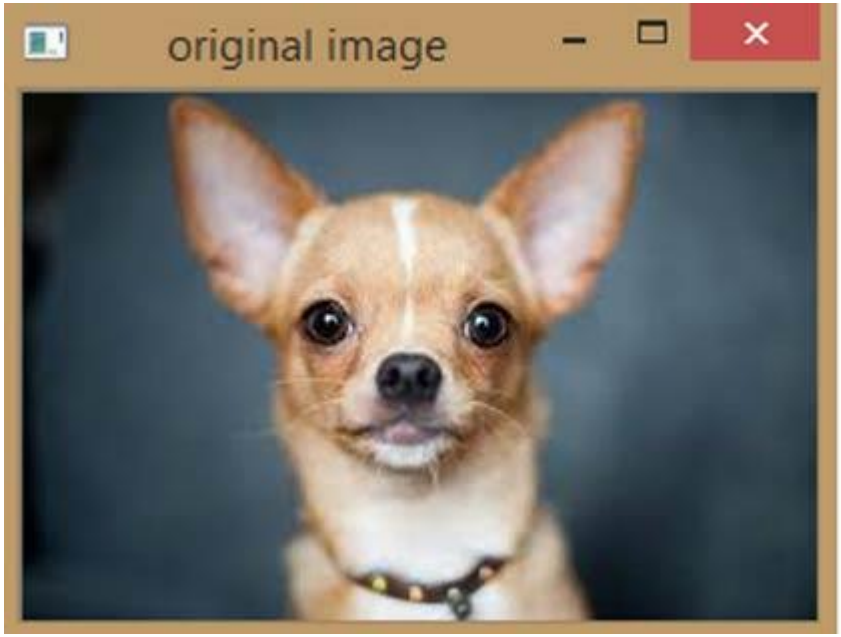
https://www.computerhope.com/issues/ch000990.htm

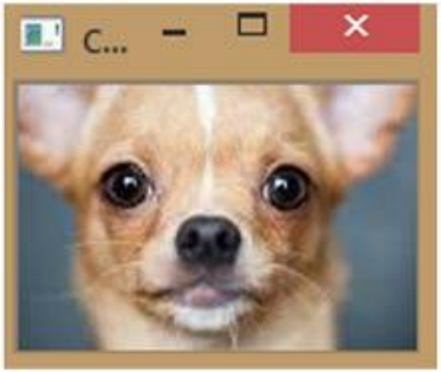




https://www.pyimagesearch.com/2021/01/20/opencv-rotate-image/

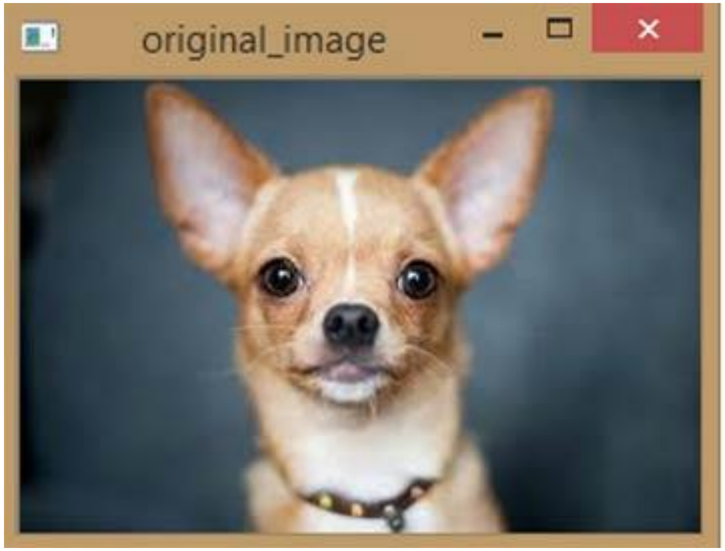
Image Zooming / Cropping

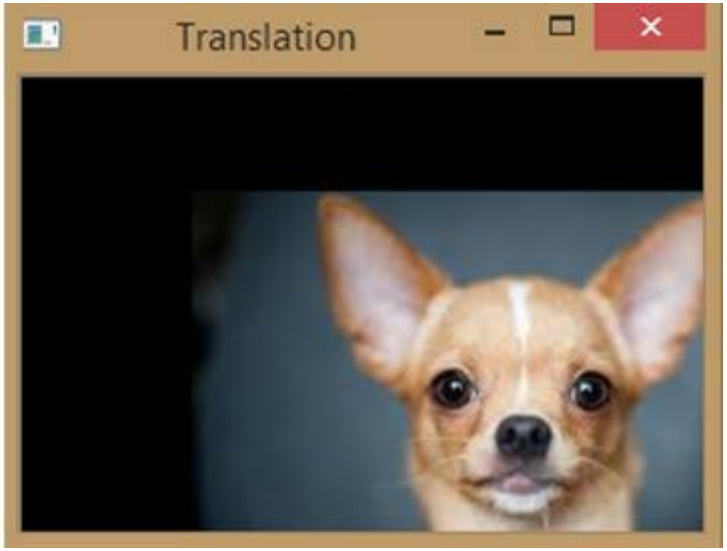




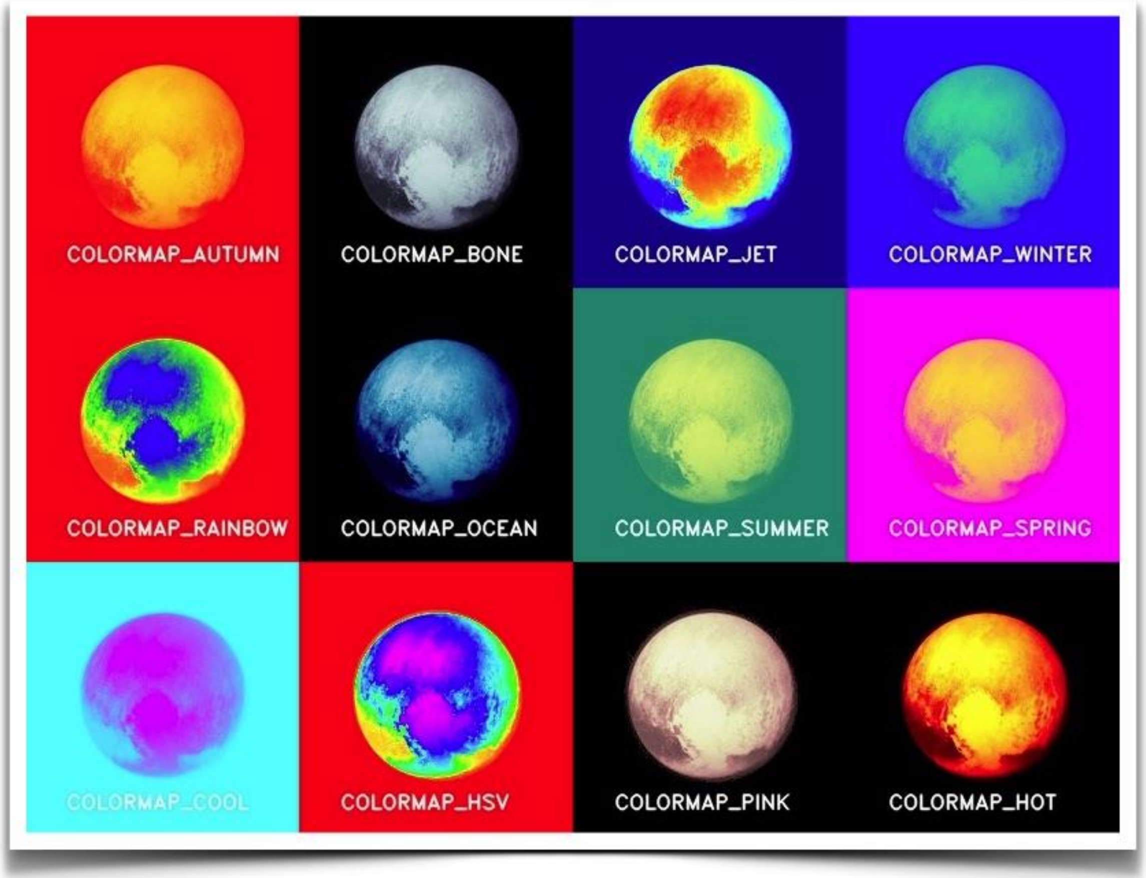
https://circuitdigest.com/tutorial/image-manipulation-in-python-opencv-part1

Image Shifting (Translation)





Channel Shifting (Colour Change)



https://learnopencv.com/applycolormap-for-pseudocoloring-in-opencv-c-python/

Activity 3 – Image Augmentation

Exercises:

- Play around with the parameters of the image augmentation functions.

Step 1:
Watch and listen to the instructor's demonstration

15 mins

Step 2:
Work through the activities

Individual Activity

20 mins

Thank you