Welcome to the Computer Vision Course
UCS 532

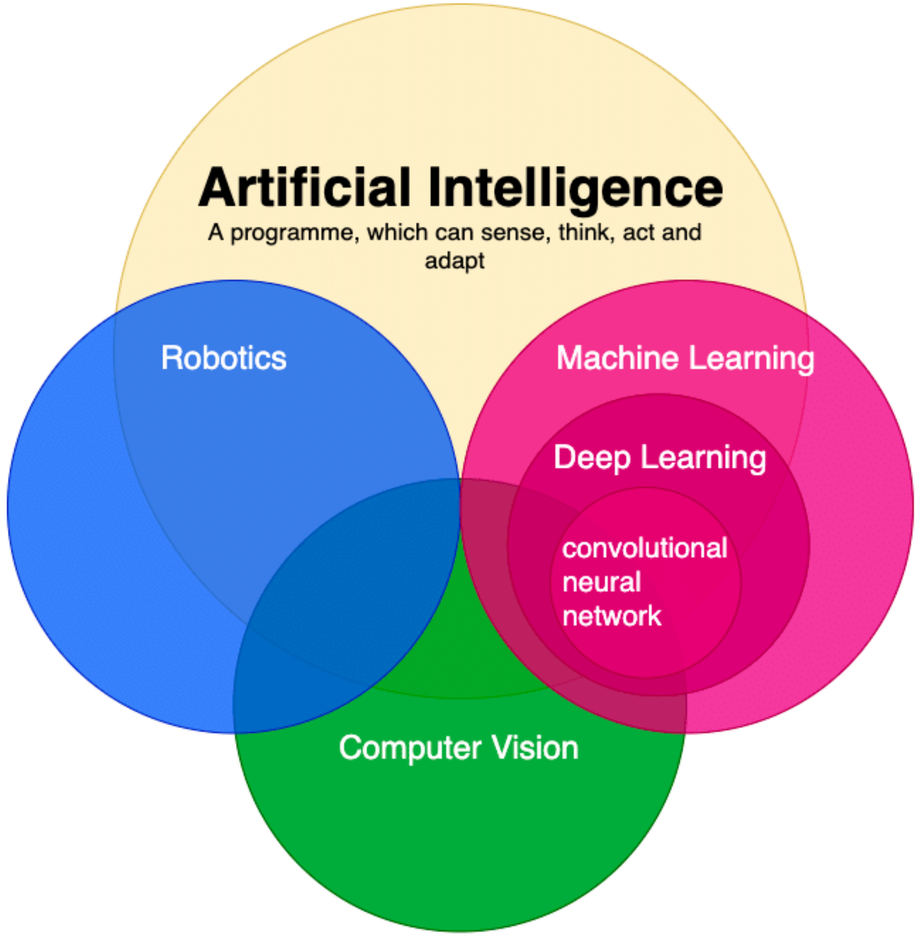
Course Instructor: Dr. Shailendra Tiwari,
Email: shailendra@thapar.edu

So what exactly is Computer Vision?

"An interdisciplinary field that aims to enable computers to gain ***understanding of what is being seen in*** *images and videos."*

Source - Terminator 2: Judgement Day (1991)

Is Computer Vision Artificial Intelligence?



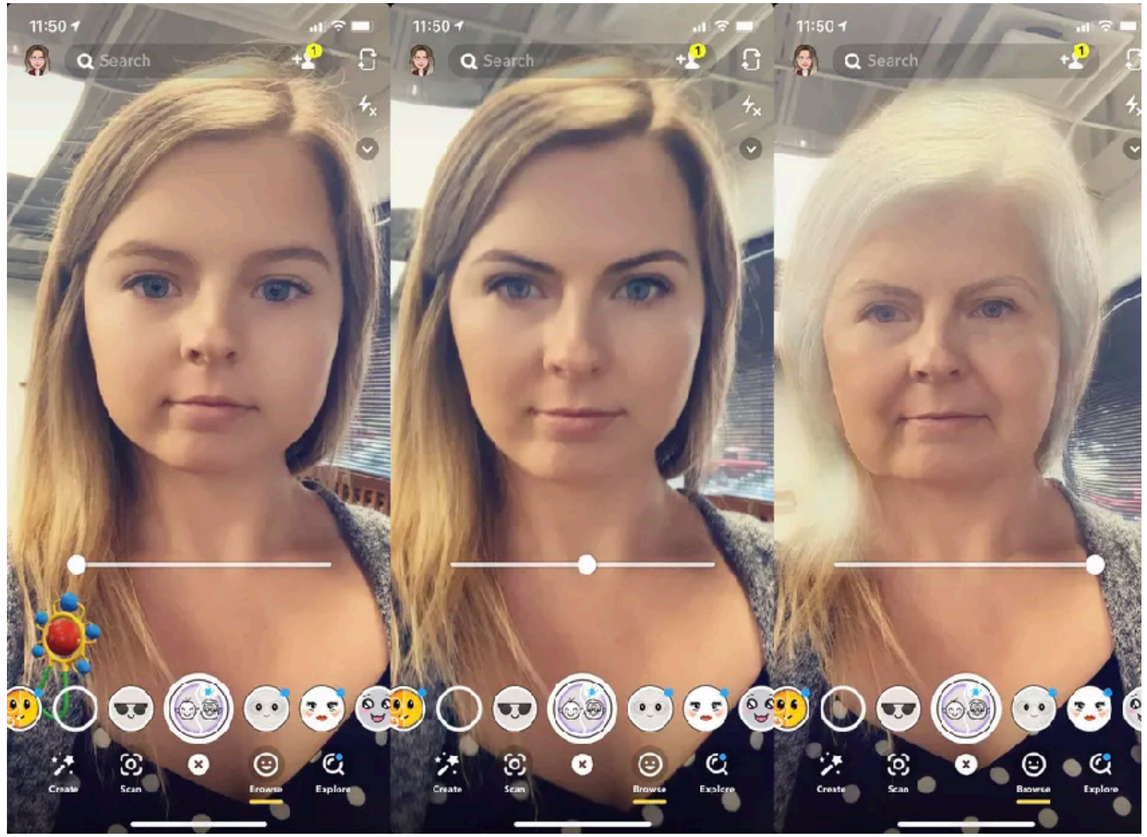
Relation between Artificial Intelligence, Machine Learning and Deep Learning,Computer Vision.
https://www.researchgate.net/figure/Relation-between-Artificial-Intelligence-Machine-Learning-and-Deep-Learning-Computer_fig1_342978934

Computer Vision is an amalgamation of many fields

So what can Computer Vision do?

You might be familiar with these…

- **Snapchat and Instagram filters**
- Optical Character Recognition (OCR)
- Licence Plate Reading
- Self-driving cars
- Sporting Analysis
- Facial Recognition

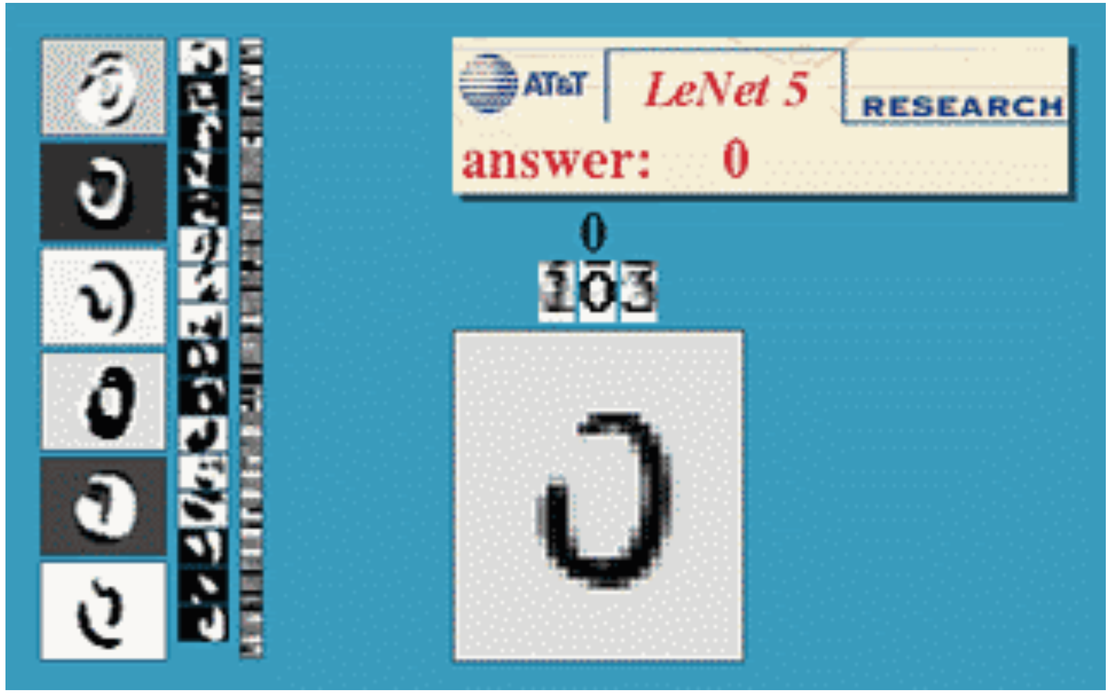


Source - Cnet - Snapchat's Time Machine AR lens creepily shows what you'll look like old

So what can Computer Vision do?

You might be familiar with these…

- Snapchat and Instagram filters
- **Optical Character Recognition (OCR)**
- Licence Plate Reading
- Self-driving cars
- Sporting Analysis
- Facial Recognition

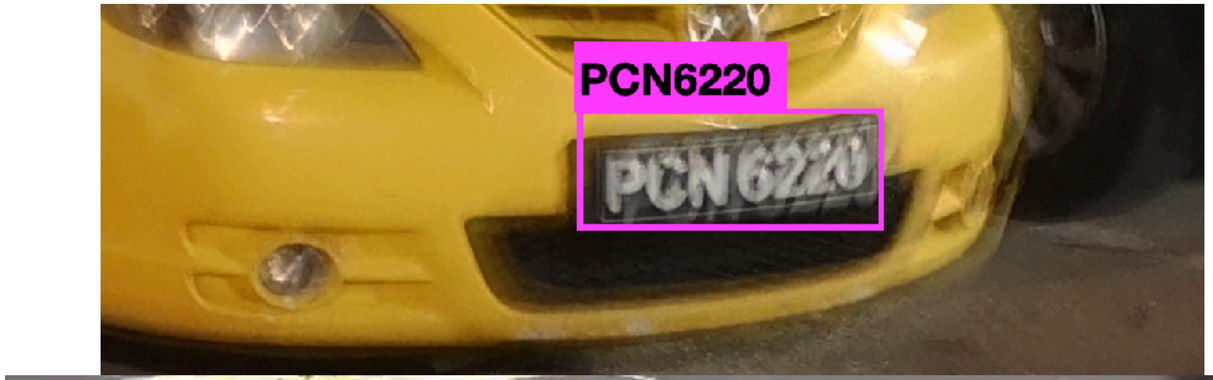


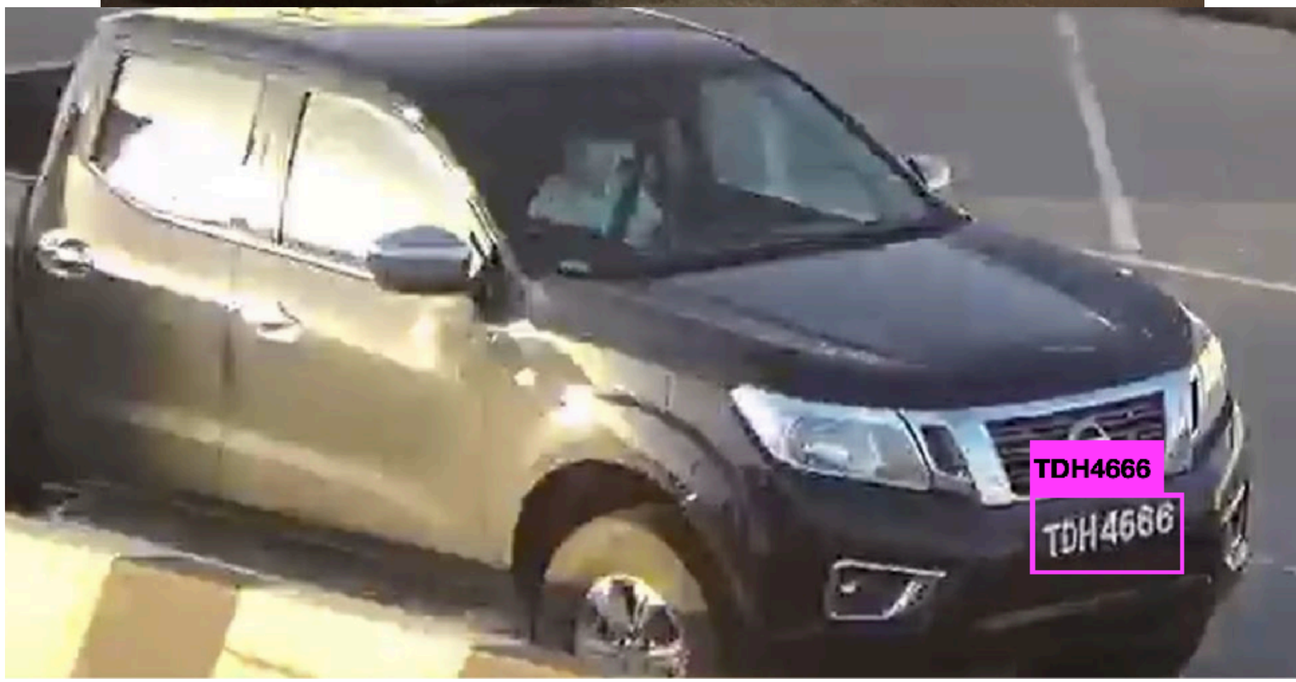
Source -AT&T's LeNet OCR for Handwritten Digits

So what can Computer Vision do?

You might be familiar with these…

- Snapchat and Instagram filters
- Optical Character Recognition (OCR)
- **Licence Plate Reading**
- Self-driving cars
- Sporting Analysis
- Facial Recognition





So what can Computer Vision do?

You might be familiar with these…

- Snapchat and Instagram filters
- Optical Character Recognition (OCR)
- Licence Plate Reading
- **Self-driving cars**
- Sporting Analysis
- Facial Recognition

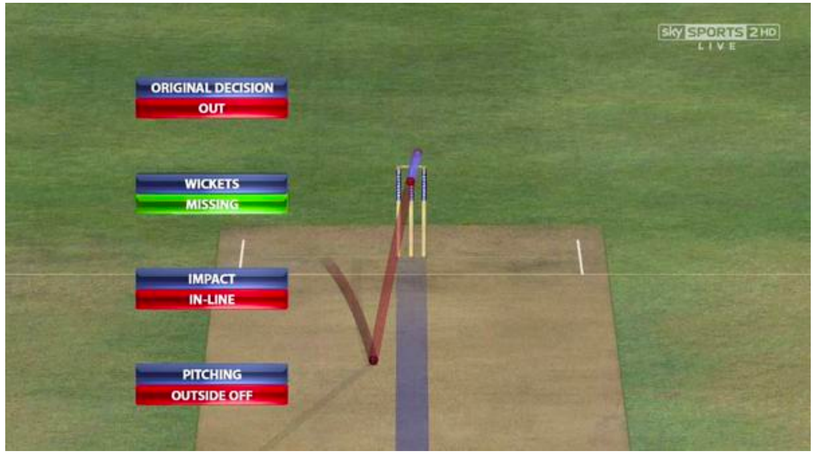
Source Inc. - The Tech That Powers Your Self-Driving Car Might Be Built Using People Playing Games on Their Phones

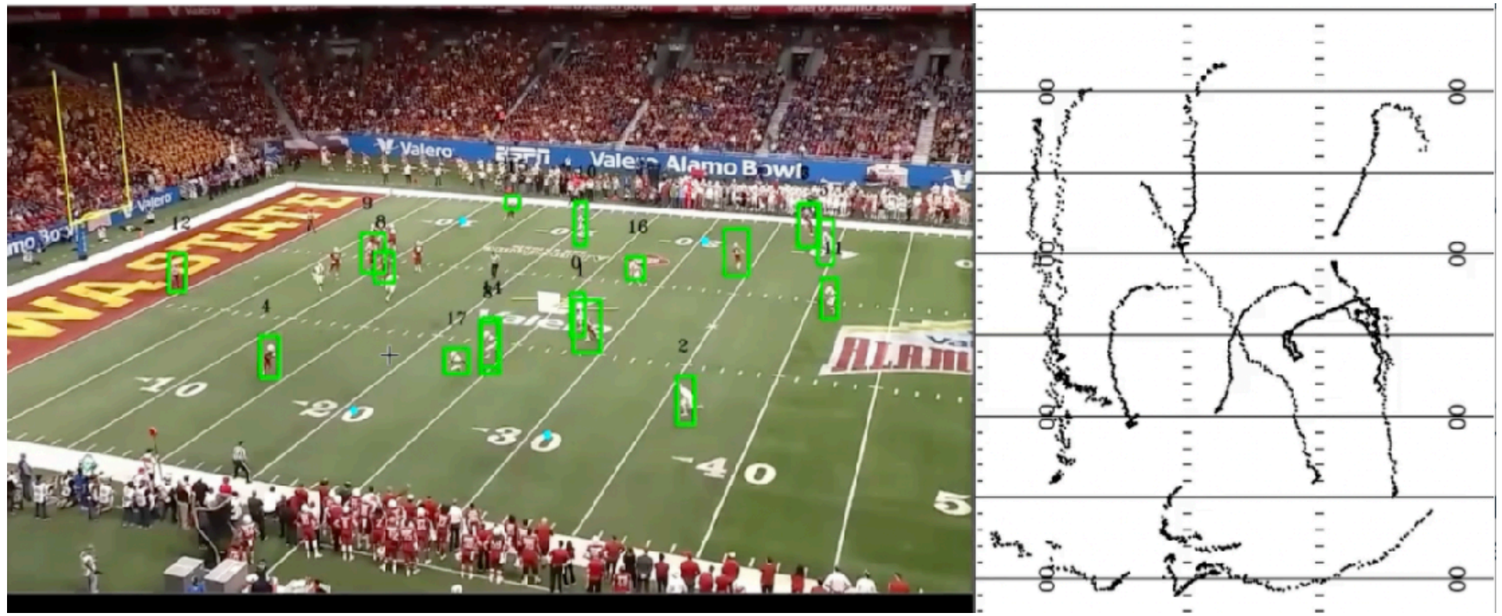
So what can Computer Vision do?

You might be familiar with these…

- Snapchat and Instagram filters
- Optical Character Recognition (OCR)
- Licence Plate Reading
- Self-driving cars
- **Sporting Analysis**
- Facial Recognition

HawkEye in Cricket



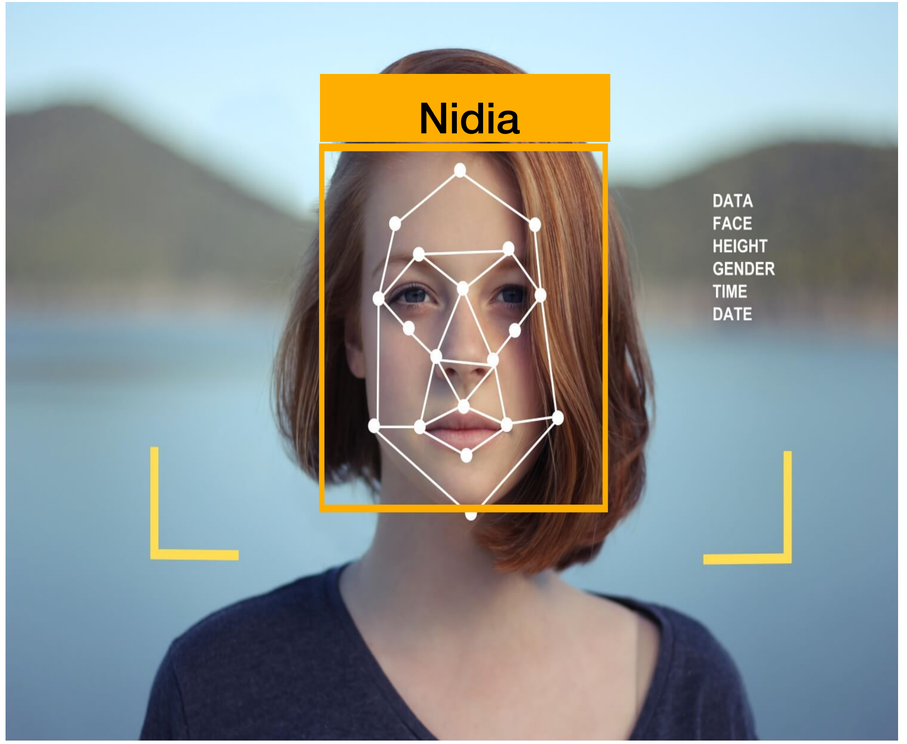


Source -Roboflow - AI Coach

So what can Computer Vision do?

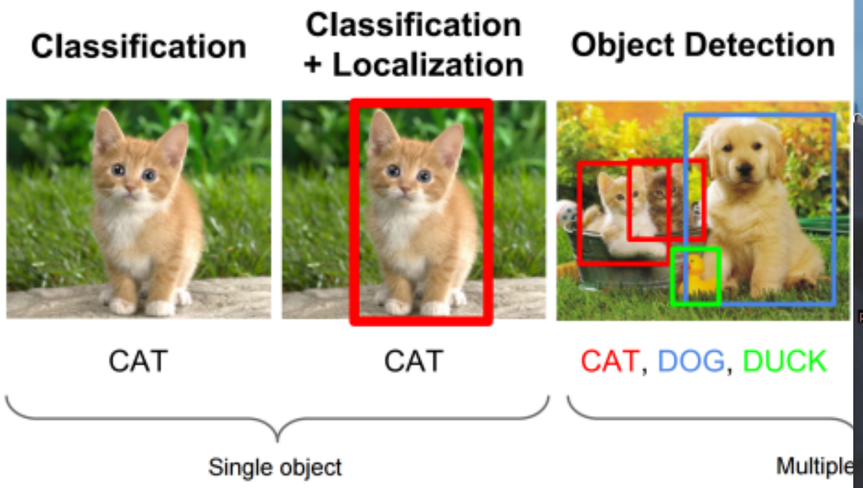
You might be familiar with these…

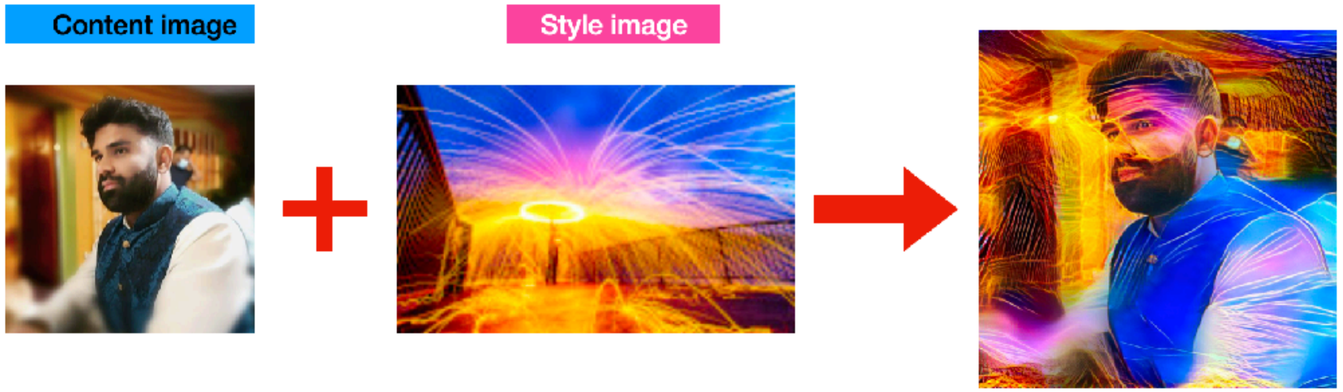
- Snapchat and Instagram filters
- Optical Character Recognition (OCR)
- Licence Plate Reading
- Self-driving cars
- Sporting Analysis
- **Facial Recognition**



But there’s so much more!

- **Image Recognition**
- **Object Detection**
- **Segmentation**
- **AI Art**
- Image Similarity
- Deep Fakes
- Body Pose Detection
- Image Generation





But there’s so much more!

- Image Recognition
- Object Detection
- Segmentation
- AI Art
- **Image Similarity**
- **Deep Fakes**
- Body Pose Detection
- Image Generation

But there’s so much more!

- Image Recognition
- Object Detection
- Segmentation
- AI Art
- Image Similarity
- Deep Fakes
- **Body Pose Detection**
- **Image Generation**

Computer Vision Applications are endless!

Electric Scooter ID
Gas Leak Detection
Document Digitization
Plant Phenotyping
Flare Stack Monitoring
Resume Parsing
Augmented Reality
Weed Detection
Microscopy
Bean Counting
Garbage Cleanup
Drone Video Analysis
Conveyer Belt Debris
Traffic Counter
Pothole Identification
Soccer Player Tracker

Steelyard Throughput
Security Cam Analysis
Self Driving Cars
Fish Measuring
Remote Tech Support
Tennis Line Tracking
Know Your Customer
Endangered Species Tracking
Inventory Management
Hard Hat Detection
Pest Identification
OCR Math
Basketball Shot Tracking
Logo Identification
Satellite Imagery
Traffic Cone Finder

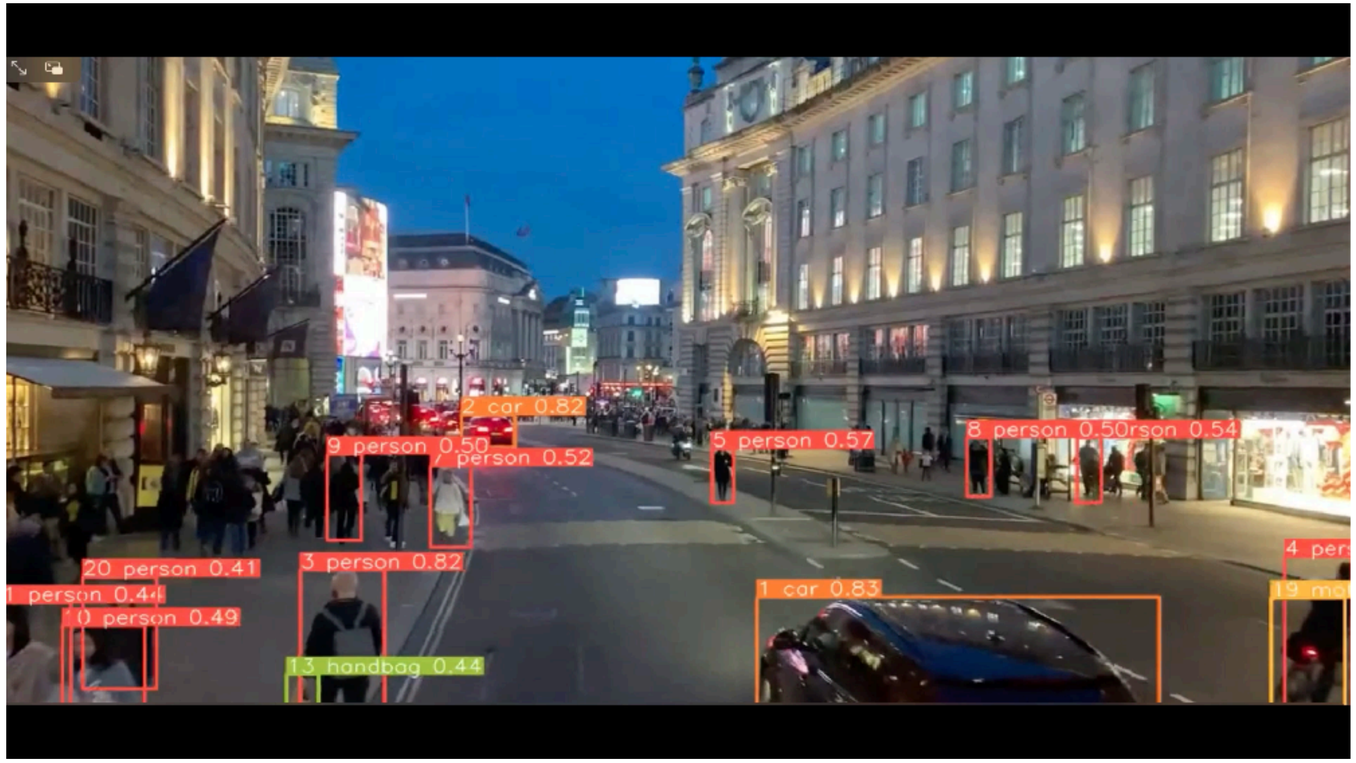
Airplane Maintenance
Tumor Detection
D&D Dice Counter
Plant Disease Finder
X-Ray Analysis
Roof Damage Estimator
City Bus Tracking
Board Game Helpers
Dental Cavity Detection
Drought Tracking
Hog Confinements
Sushi Identifier
Oil Storage Estimator
Car Wheel Finder
License Plate Reader
Exercise Counter

roboflow

Why do this course?

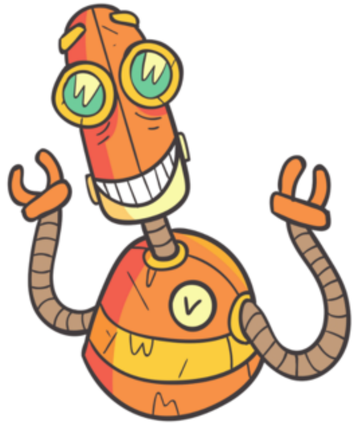
What are you going to learn exactly?

- Foundation in **Classical Computer** Vision with **OpenCV**
- **Deep Learning** applied to Computer Vision
  - PyTorch
  - TensorFlow Keras



Requirements?

- **Internet connection** (preferably at least 2mb) and web browser
- **High school level math** knowledge (preferred, not required)
- **Basic programming** knowledge is **not required** but preferred (especially if it's Python)
- **Enthusiasm** about Artificial Intelligence!



High Level Overview

OpenCV - Classical Computer Vision Outline

1.Image Manipulations
2.Segmentation
3.Feature Extraction
4.Object Detection with Haar Cascade Classifiers
5.Image analysis and Transformations
6.Background Removal
7.Tracking - Optical Flow and MeanShift
8.Face Swapping
9.OCR and text detection
10.QR and Barcode reading
11.YOLOv3 & SSDs in OpenCV
12.Neural Style Transfer
13.Colorising Black and White Photos
14.Computation Photography (noise removal, in-painting)
15.Facial Recognition
16.Working with Video and Video Stream

High Level Overview

Deep Learning Outline

1.Introduction to Deep Learning in Computer Vision
2.Deep Learning Basics with PyTorch and TensorFlow Keras
3.Comparing Libraries and Analysing Performance
4.Regularisation and Overfitting
5.Understanding What CNN's See
6.Designing CNNs and Overview of Modern CNN Architectures
7.PyTorch Lightning
8.Transfer Learning
9.Google DeepDream
10.Neural Style Transfer
11.Autoencoders
12.Generative Adversarial Networks
13.Siamese Networks
14.Facial Recognition
15.Object Detection
16.Deep Segmentation
17.Tracking with Deep Learning
18.Vision Transformers
19.Depth Estimation
20.Video Classification
21.Point Cloud Classification and Segmentation
22.3D Image Classification
23.OCR Readers
24.Image Captioning
25.Computer Vision APIs

Course Overview

This is a BIG course! Let’s get started!

So how do we do Computer Vision?

What makes it possible?

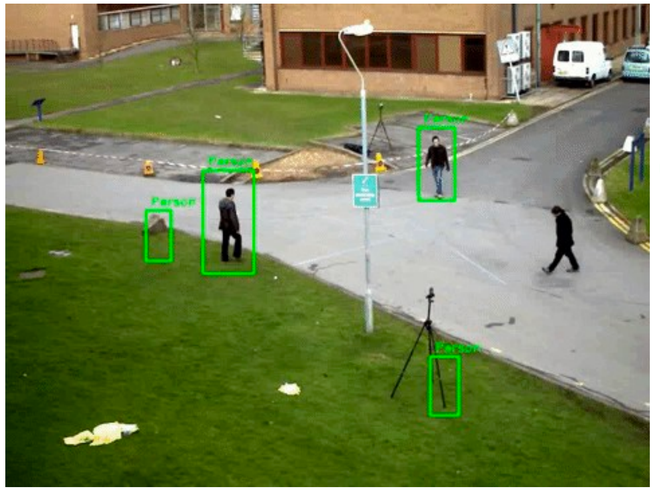
- We need tools! Namely a programming language
- Many exist such as:
  - Matlab
  - C++ & Java
  - **Python**

What makes Python so great for Computer Vision?

- Easy to learn and accessible
- Language of Artificial Intelligence
- The best and easiest to use Libraries, such as:
  - **OpenCV** for Classical Computer Vision
  - **PyTorch** (Facebook) and **TensorFlow with Keras** (Google)

Classical Computer Vision?

- What is meant by **Classical Computer Vision?**
- It encompasses Computer Vision algorithms that **do not** involve Machine Learning
- Before the advent of Machine Learning and Deep Learning, Computer Vision was a deeply explored field and many useful algorithms were developed for things like **feature extraction, OCR, Segmentation and simple transformations.**
- **OpenCV** is the Classical Computer Vision library of choice!



OpenCV

Deep Learning Computer Vision

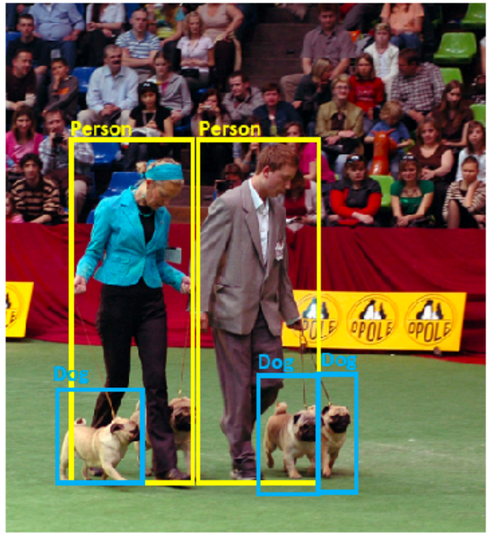
- Deep Learning was used in Computer Vision since the 1990s, however due to the computational requirements and intricate design, it remained on the sidelines for decades.
- Until the the mid 2010s…which brought **two important building blocks** together.
  - **Mature Deep Learning libraries** (TensorFlow, Keras, Theano, Caffe)
  - **Accessible GPU processing** (NVIDIA's CUDA)

Deep Learning Examples

Image Classification

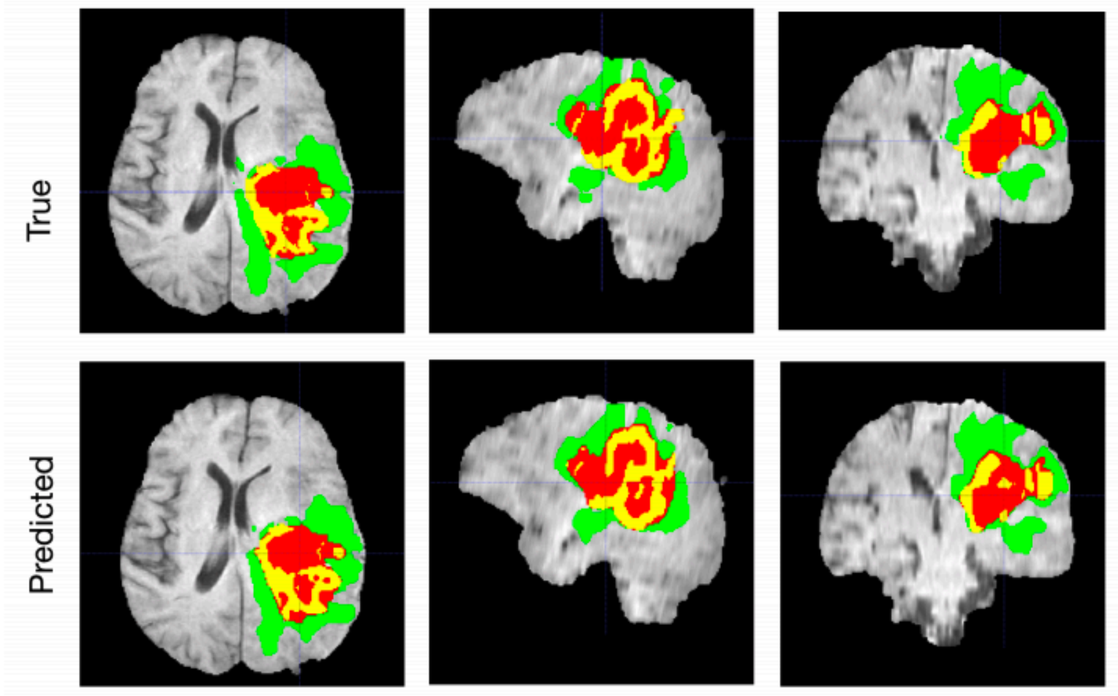
{Dog}

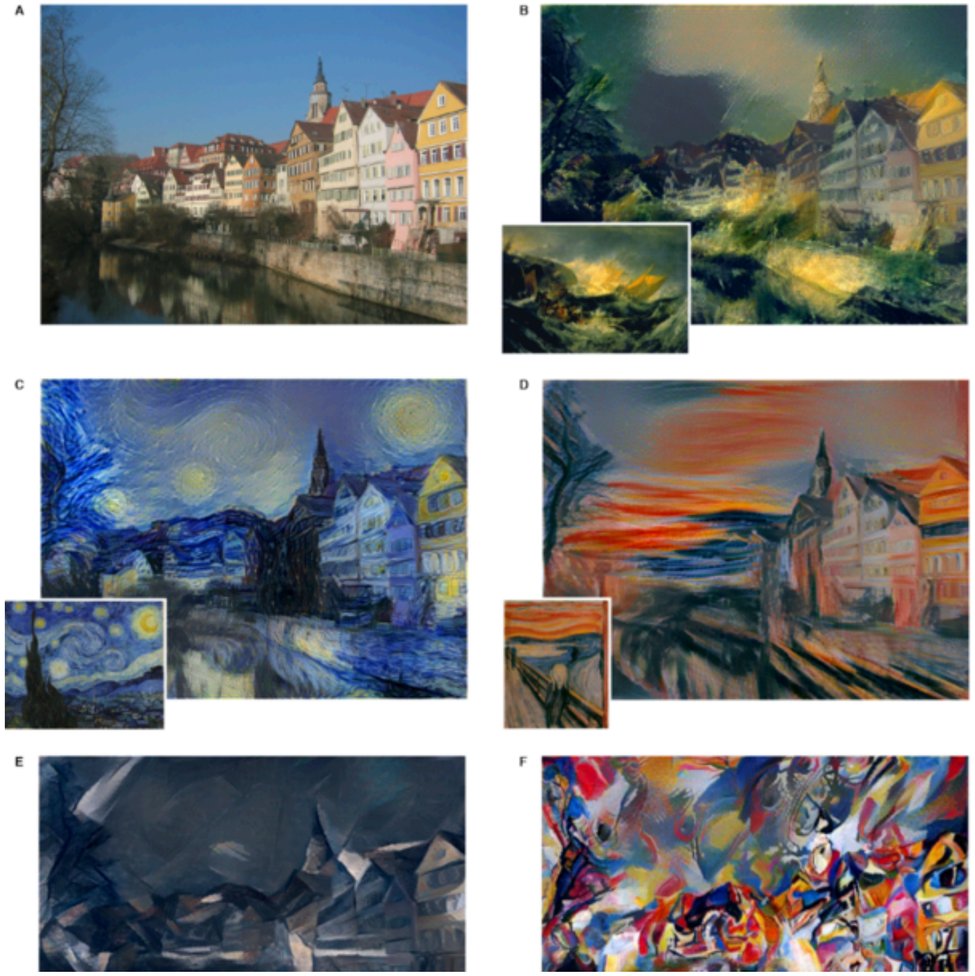
Object Detection



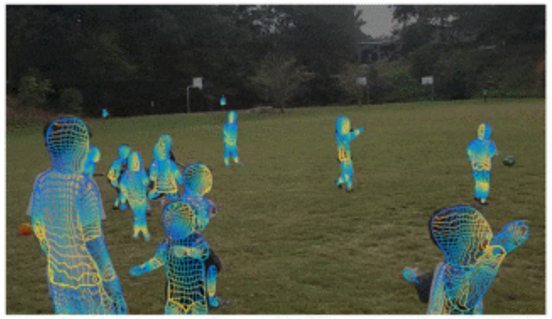
{Dog, Dog, Dog, Person, Person}

Deep Learning Examples





Deep Learning Examples



Deep Learning CV vs Classical CV

Deep Learning	Classical Computer Vision
Adapts to new images well (assuming it’s similar to the data it was trained on)	Small changes can have big negative impacts
Requires Models to be trained	Doesn’t require training and can be used once coded
Model weights learn to adapt to varying image conditions	Relies on hardcode features and parameters
Requires GPU hardware (most times)	Can be run on CPU

Our OpenCV Outline

1. Getting Started with OpeNCV
2. Grey-scaling Images
3. Color Spaces (HSV & RGB)
4. Drawing on Images
5. Transformations - Translations and Rotations
6. Scaling and re-sizing and Cropping
7. Arithmetic and Bitwise Operations
8. Convolutions, Blurring and Sharpening
9. Thresholding & Binarization
10. Dilation, Erosion and Edge Detection
11. Contours - Drawing and Hierarchy and Modes
12. Moments, Matching and Sorting Contours
13. Line, Circle, Blob Detection
14. Counting Circles, Ellipses and Finding Waldo
15. Finding Corners
16. Face and Eye Detection with HAAR Cascade Classifiers
17. Vehicle & Pedestrian Detection
18. Perspective Transforms
19. Histograms and K-means clustering for finding dominant colours
20. Comparing Images with MSE and Structural Similarity

Our OpenCV Outline

21. Filtering Colors
22. Watershed Algorithm marker-based image segmentation
23. Background and Foreground Subtraction
24. Motion tracking using Mean Shift and CAM-Shift
25. Optical Flow Object Tracking
26. Simple Object Tracking by Colour
27. Facial Landmarks Detection with Dlib
28. Face Swapping with Dlib
29. Tilt Shift Effect
30. Grabcut Algorithm for Background Removal
31. OCR with PyTesseract and EasyOCR
32. Barcode and QR generation and reading
33. YOLOv3 in OpenCV
34. Neural Style Transfer with OpenCV
35. SSDs in OpenCV
36. Colorise Black and White Photos
37. Repair Damaged Photos with Inpainting
38. Add and remove Noise, Fix Contrast with Histogram Equalisation
39. Detect Blur in Images
40. Facial Recognition

Next...

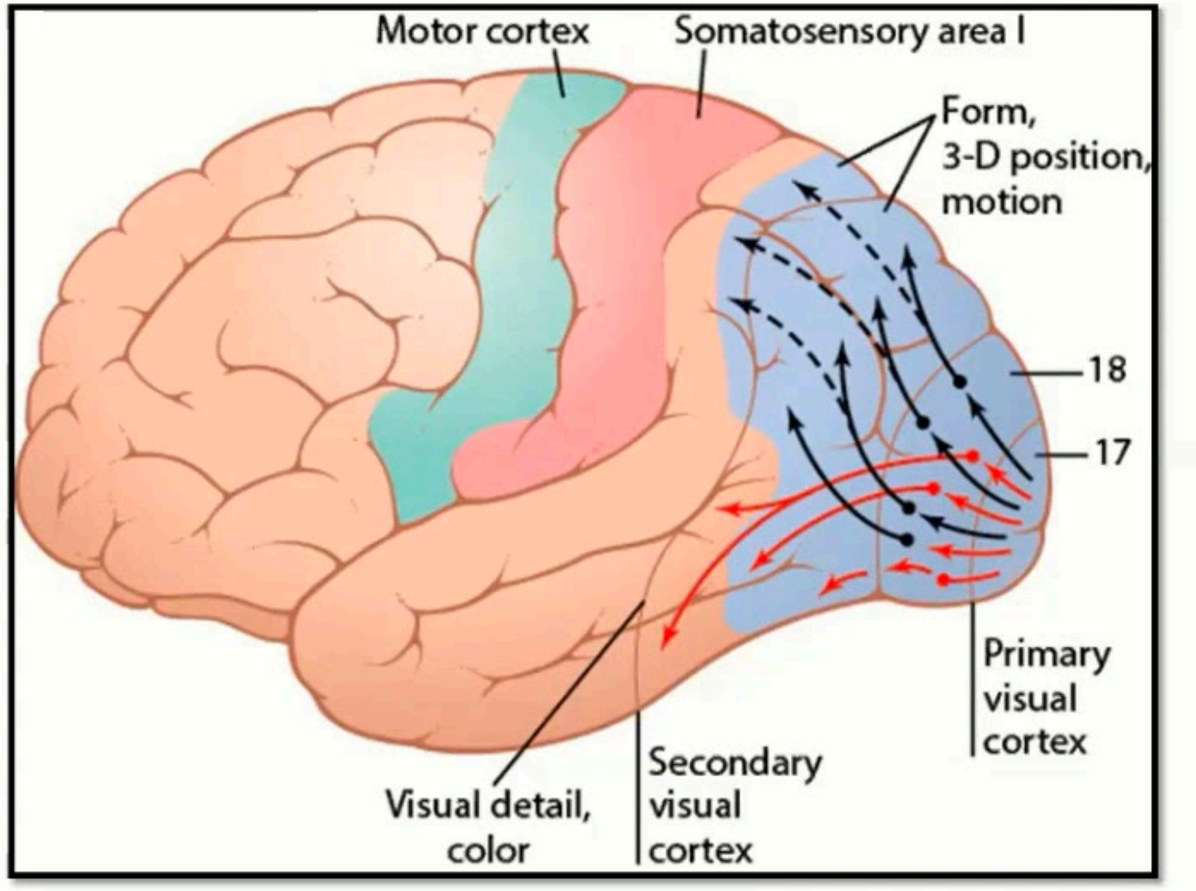
Introduction to Computer Vision

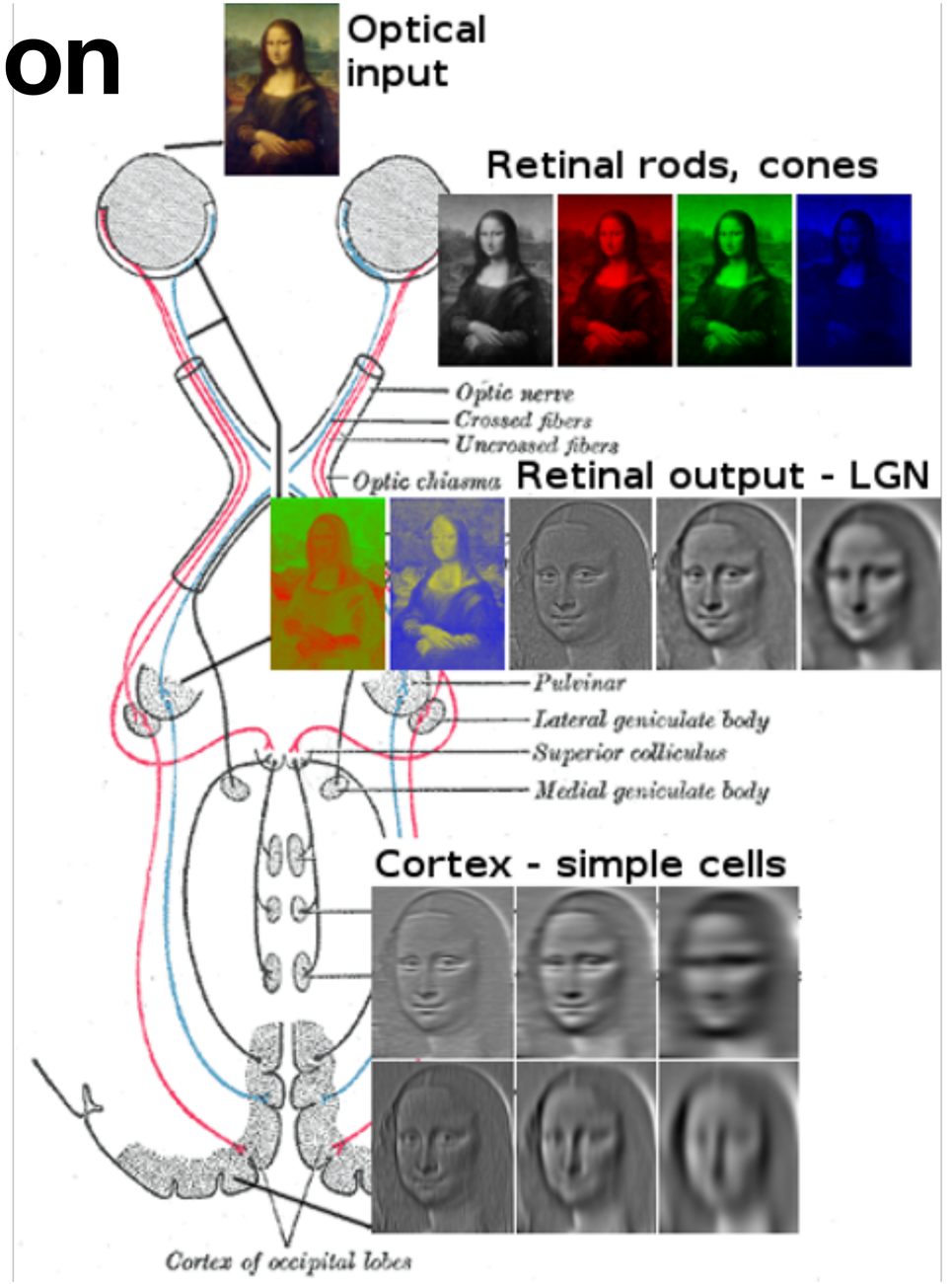
What Makes Computer Vision Hard?

Nothing worth having comes easy

Our Brains are amazing at Vision

The **visual cortex** (located in the occipital lobe) is the primary cortical region of the brain that receives, integrates, and processes visual information relayed from the retinas





Can Artificial Intelligence Come Close?

The AI Dream, a machine that can see and understand better than humans.

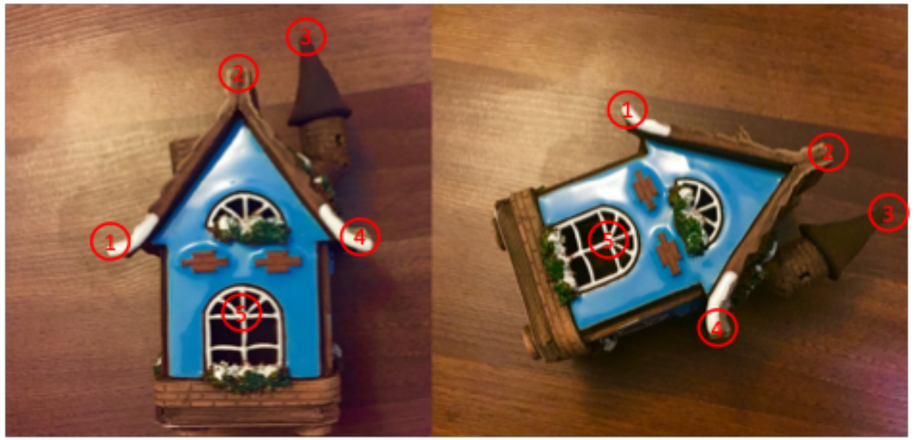
Source - Terminator 2: Judgement Data (1991)

What Makes Computer Vision Hard?

Camera and Senor Limitations

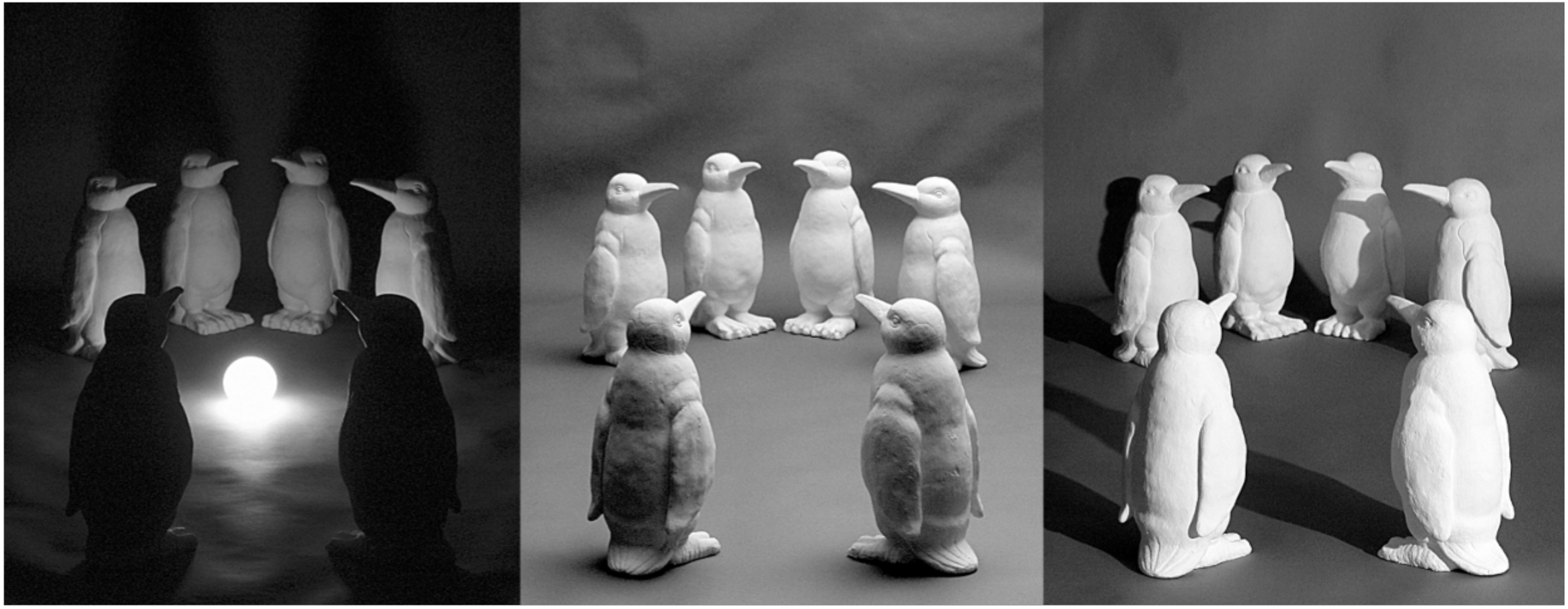
What Makes Computer Vision Hard?

Viewpoint Variations



What Makes Computer Vision Hard?

Changing Lighting Conditions

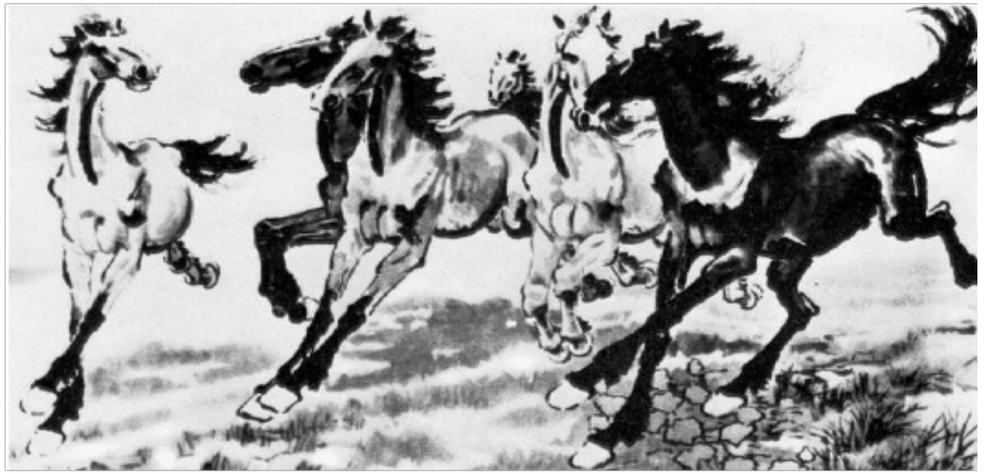


What Makes Computer Vision Hard?

Scaling Issues

What Makes Computer Vision Hard?

Non-rigid Deformations



What Makes Computer Vision Hard?

Occlusion

What Makes Computer Vision Hard?

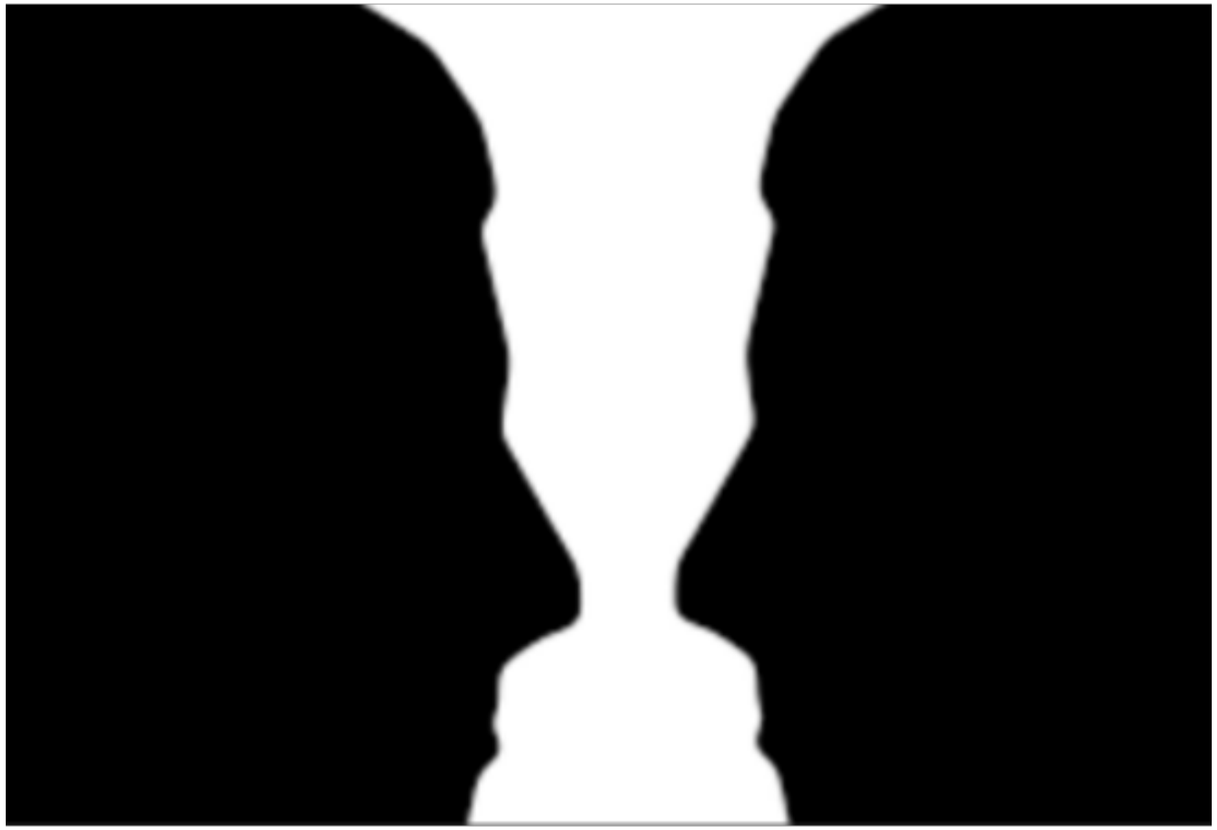
Clutter

What Makes Computer Vision Hard?

Object Class Variation

What Makes Computer Vision Hard?

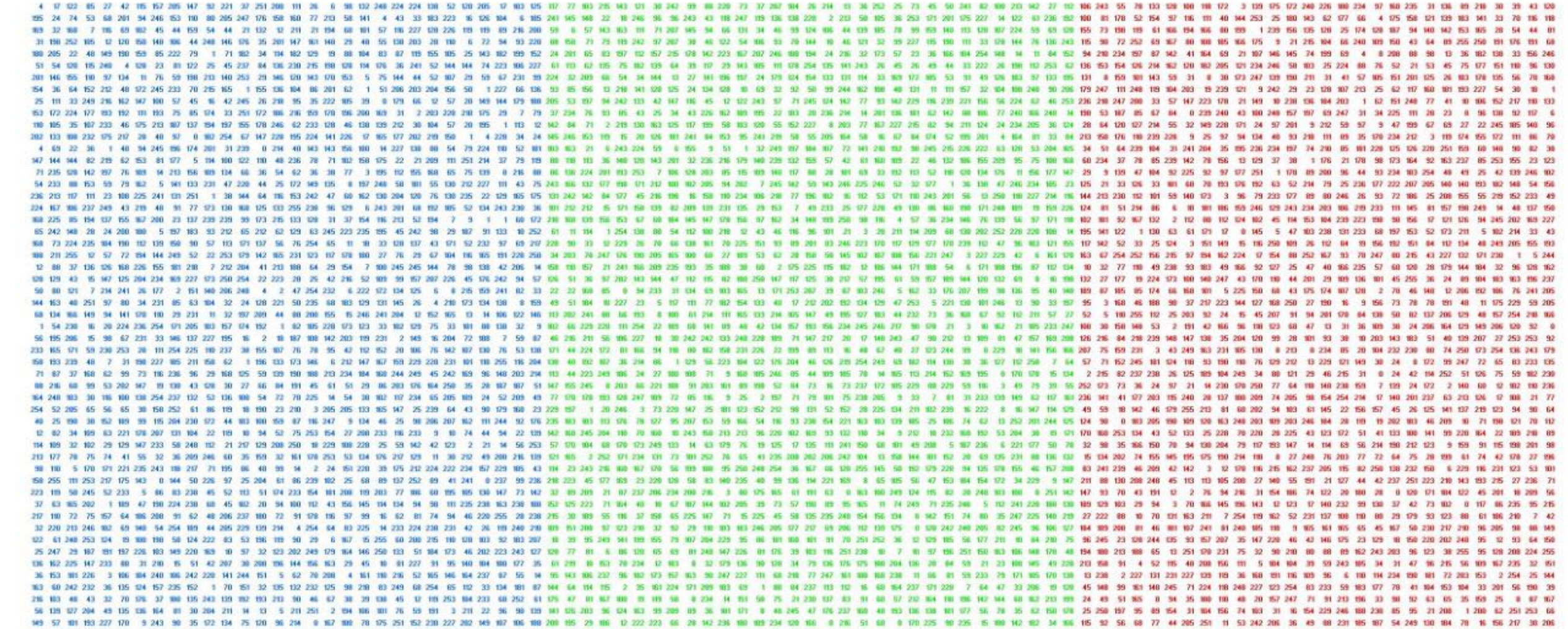
Ambiguous Optical Illusions



What Are Images?

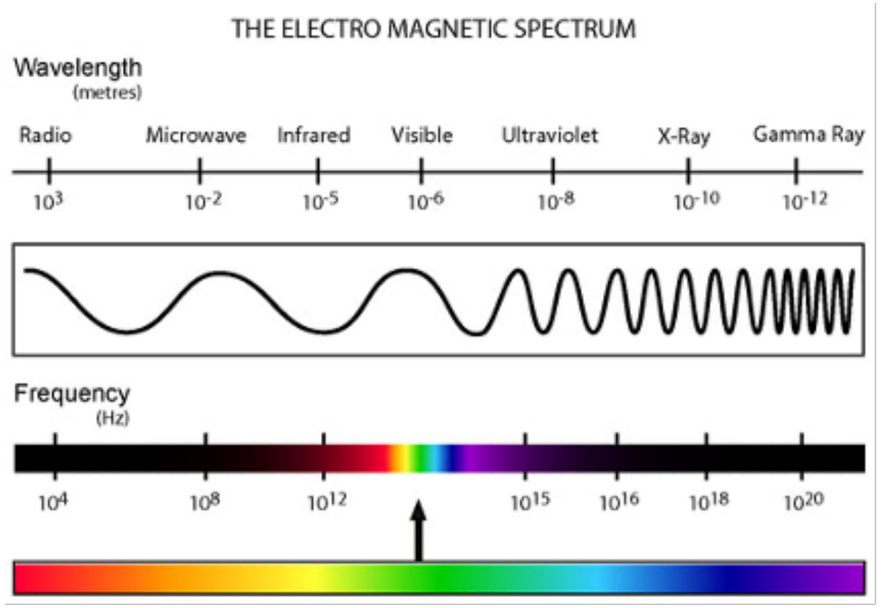
A foundation in understanding digital images

How do Computers ‘see’ Images?



What are Images?

- Images are 2-D representations of the **visible light spectrum**



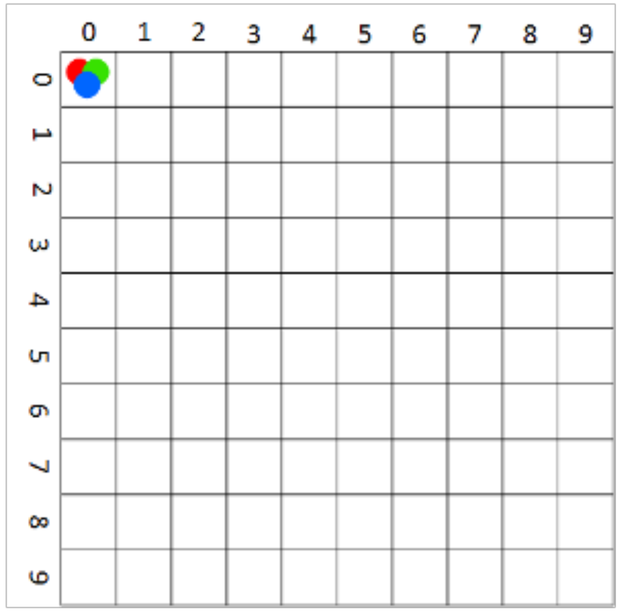
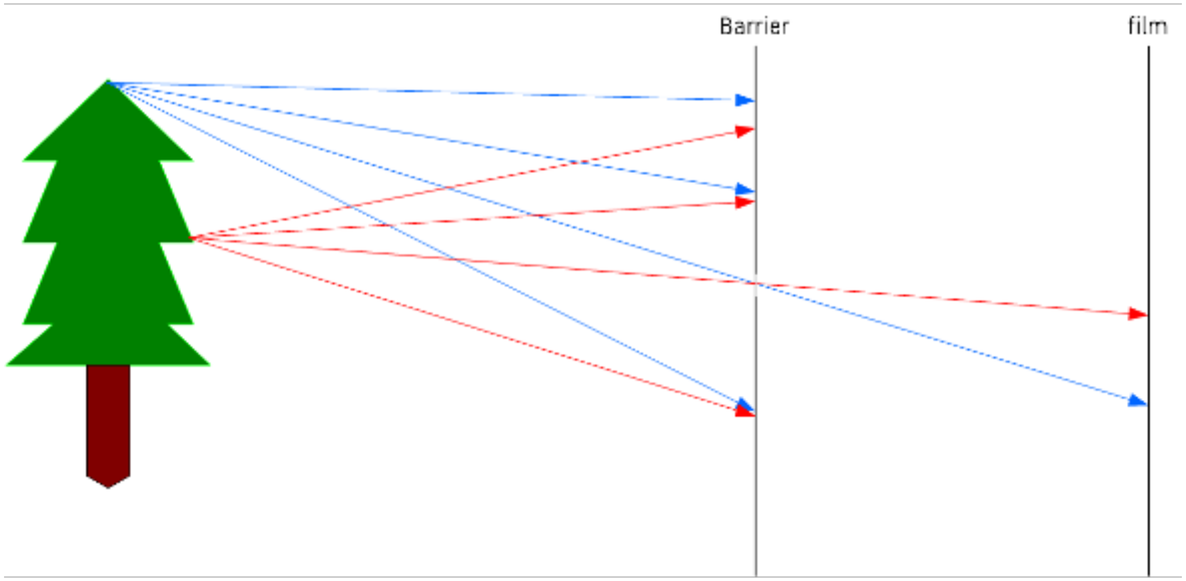


Image Creation

- Digital images are created either by using software on computer (think clip art or logos) or by digital cameras (or scanners).

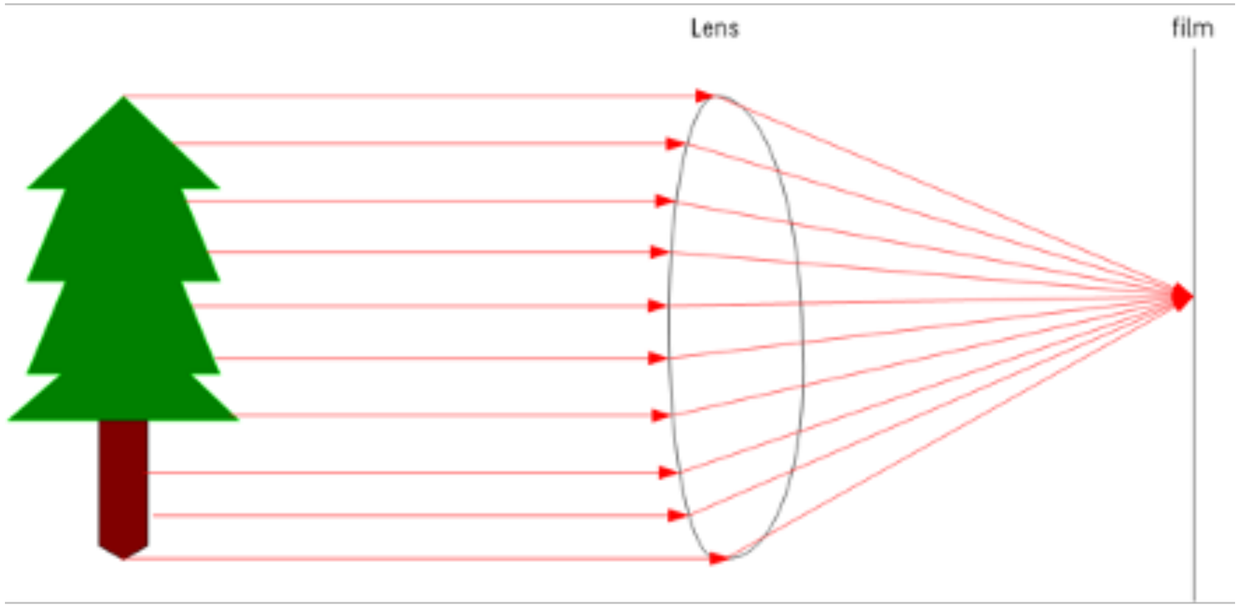


- The pinhole camera model is the basis of how we form an image using a photosensitive sensor

Modern Cameras

Our eyes as well as cameras use an **adaptive lens** to control many aspects of image formation such as:

- **Aperture Size**
  - Controls the amount of light allowed through (f-stops in cameras)
  - Depth of Field (Bokeh)
- **Lens width** - Adjusts focus distance (near or far)



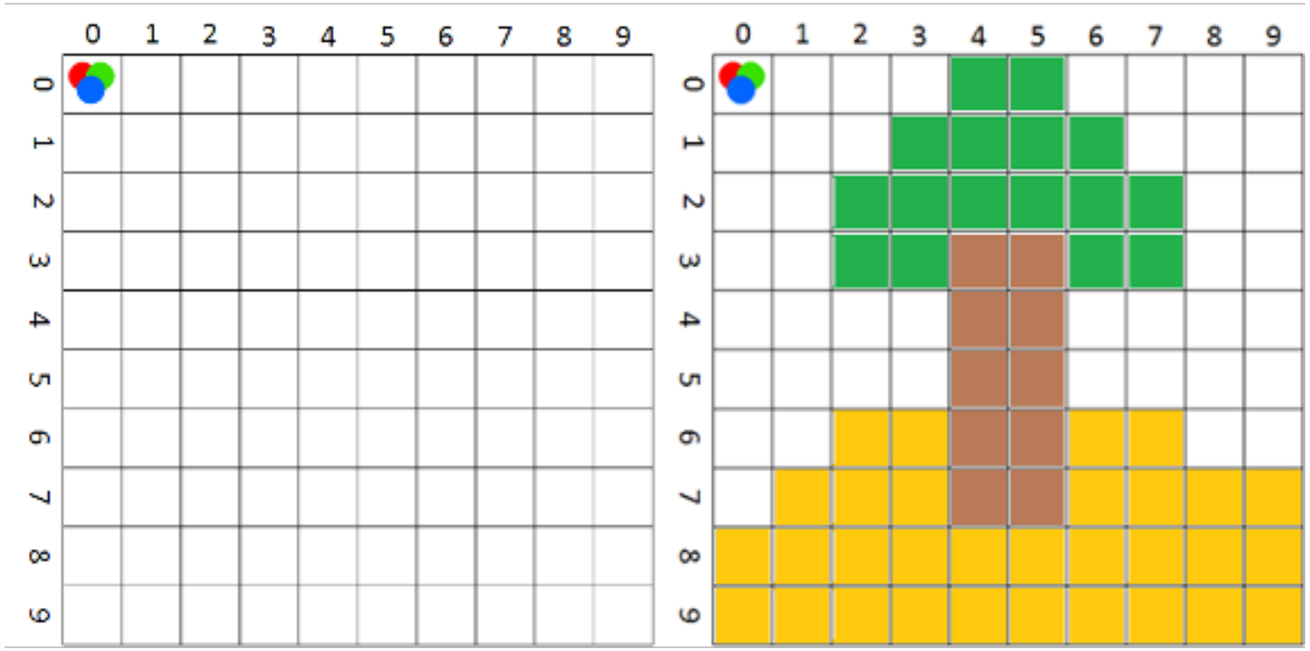
Digital Images

- Computers store images in a variety of formats, however they all involve mixing some of the following such as: colours, hues, brightness or saturation.

- The one most commonly used in Computer Vision is **RGB (Red, Green, Blue)**

- Each pixel is a combination of the brightness of blue, green and red, ranging from 0 to 255 (brightest)

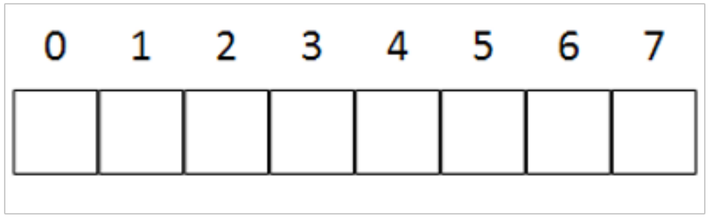
- Yellow is represented as:
  - Red – 255
  - Green – 255
  - Blue - 0



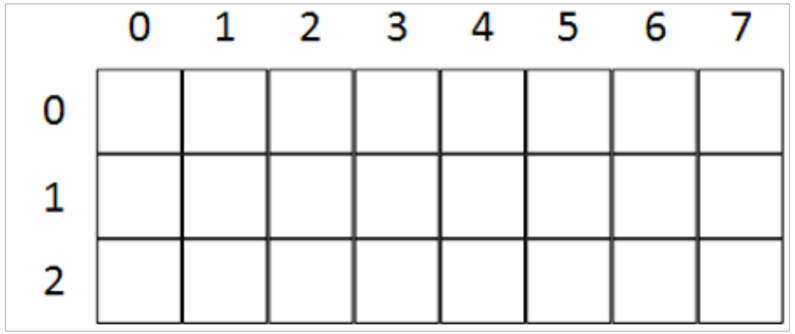
Digital Images Format

Images are stored in Multi-Dimensional Arrays

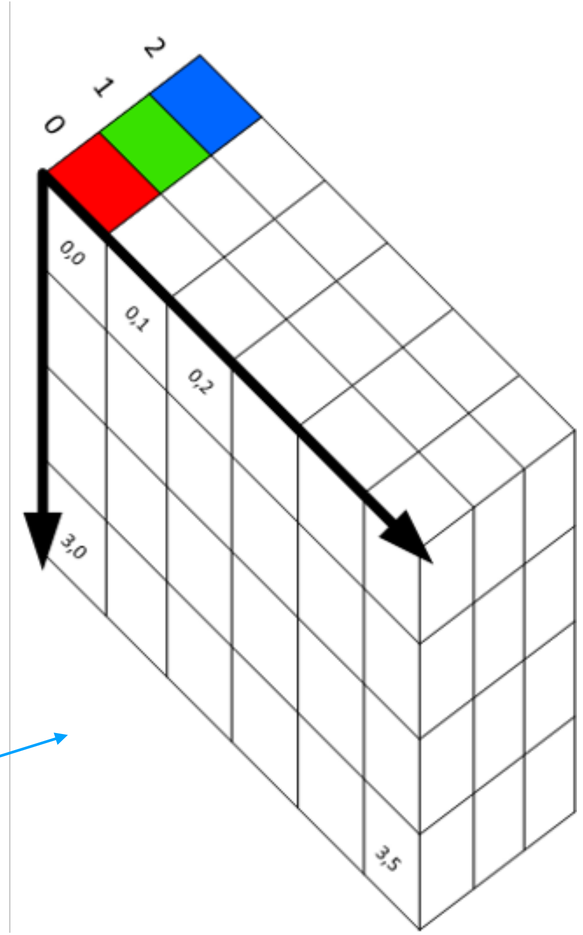
- A 1-Dimensional array looks like this



- A 2-Dimensional array looks like this



- A **3-Dimensional** array looks like this



A Grayscale Image

Sometimes referred to as black and white image

- Info is stored in a 2-Dim array

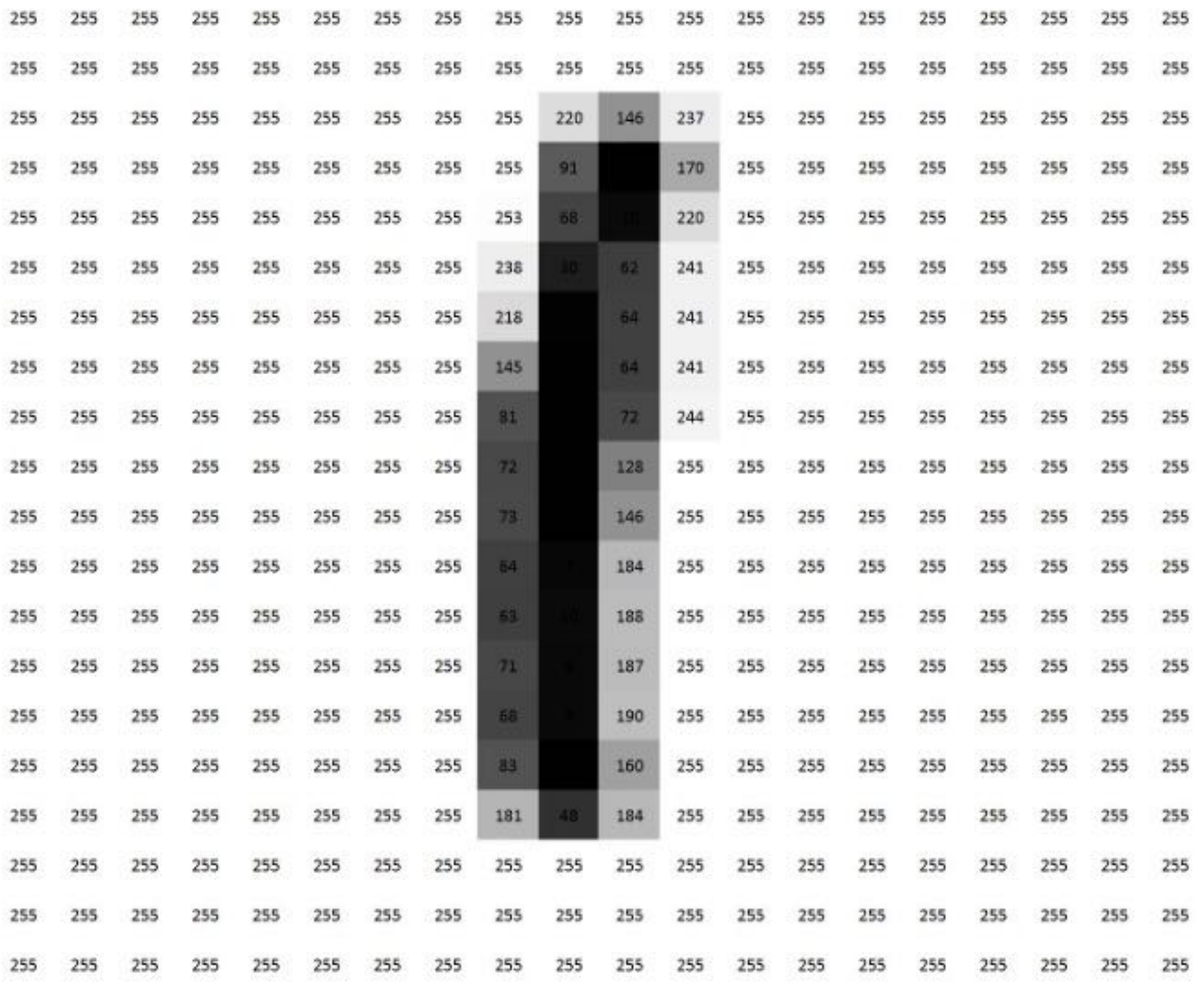


Image File Formats