

---

# Welcome to Deep Learning Online Bootcamp

## Day 12 - Computer Vision & Open CV



Democratizing Data Science Learning

# Learning Objectives

---

**How do Computers  
see images?**

**Computer Vision**

**OpenCV**



---

When we see this image, we can say, without even thinking, that this is an image of an adorable dog. Even an 8-year-old kid would not have any problem identifying the dog in the image.



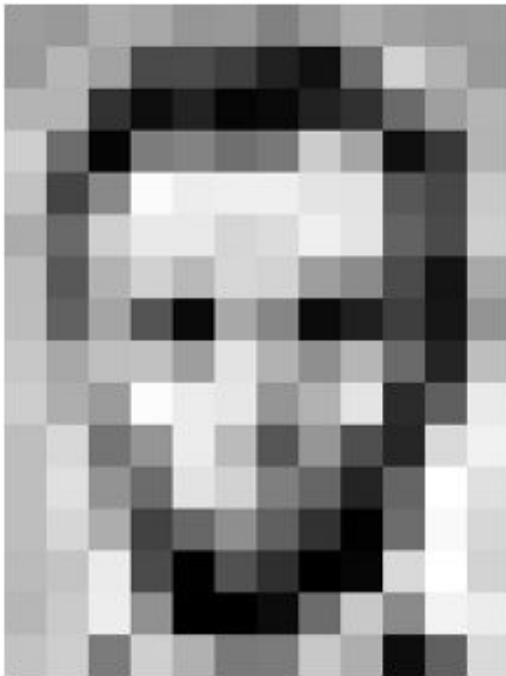
---

Have you ever wondered how a computer sees  
the same image?



# Pixel

A computer sees an image as 0s and 1s (binary digits).  
**Pixel** is the smallest unit in an image.



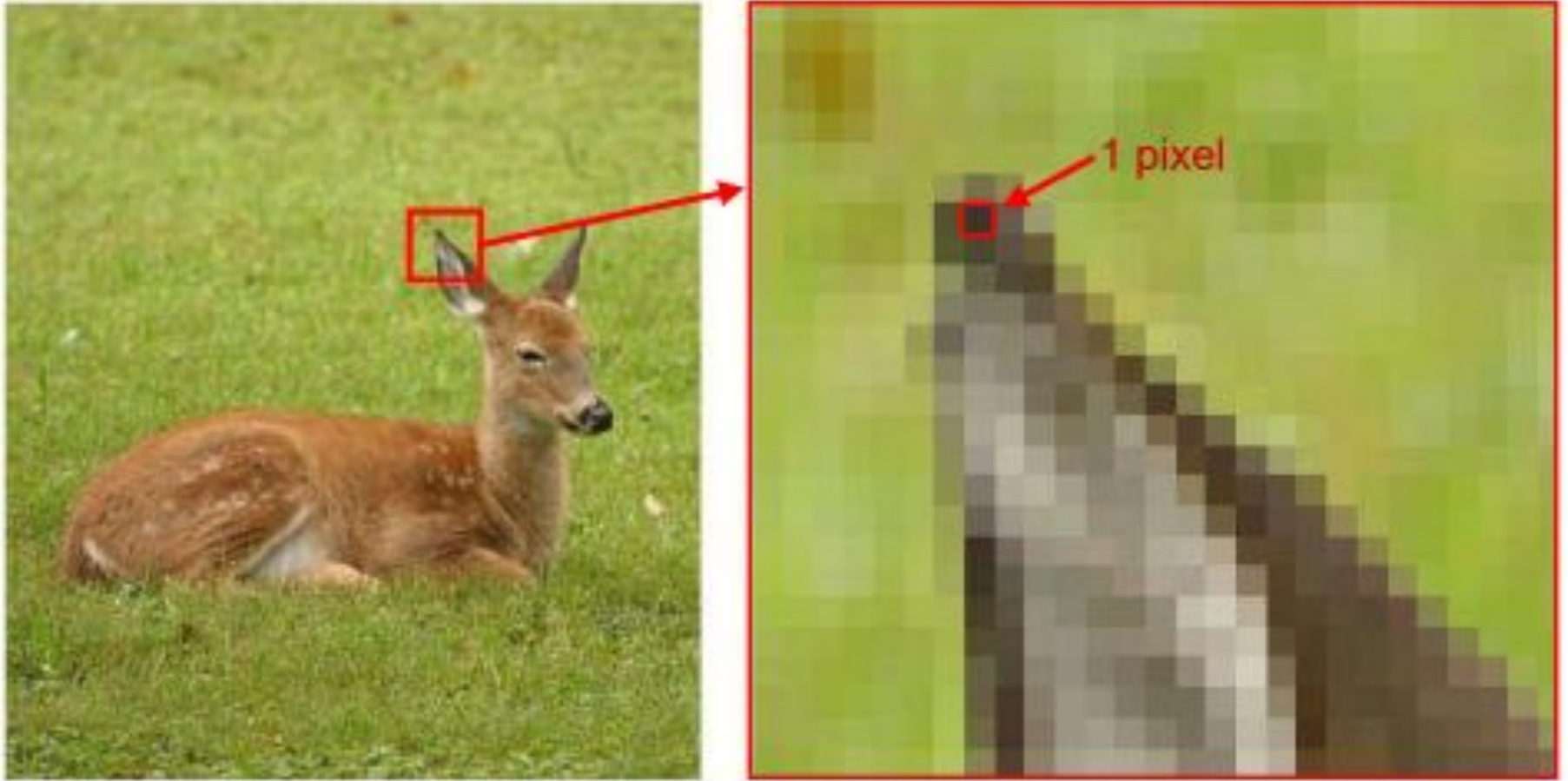
157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	83	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	6	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	85	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	60	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	83	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	6	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	85	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	60	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Pixel data diagram. At left, image of Lincoln; at center, the pixels labeled with numbers from 0–255, representing their brightness; and at right, these numbers by themselves. What is this? 0-255, it is RGB color combination values, we will learn about it soon.

# Pixel

---



Ear has of the animal has been zoomed to show the pixel of an image

# Channels

---

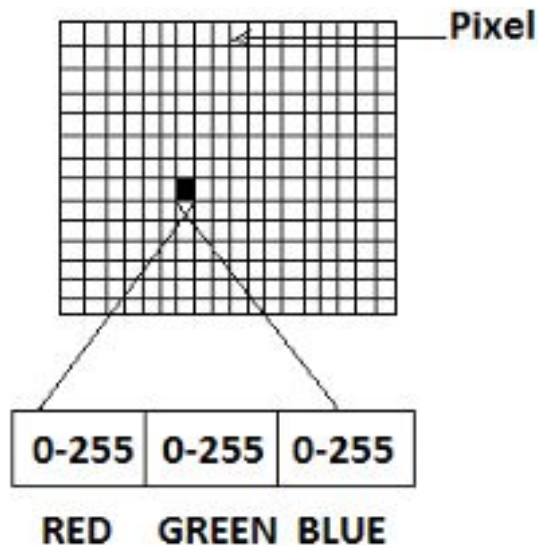
Computers are actually unable to recognize or look at images the way we humans would. So, we would have to find a way to convert these images into numbers. There are two common ways to do this when it comes to Image Processing:

1. **Greyscale** - A range of grey shades from white to black.  
When using the Greyscale, the computer assigns each pixel a value (in numbers) based on its level of darkness.
2. **RGB Values** - A combination of red, green, and blue.  
Once a colour is given a RGB Value, the computer extracts that value from each pixel and puts the results in an array.



# Channels






Each pixel contains a different number of channels. If it is a **grayscale image**, it has **only one channel**, whereas if it is a **coloured image**, it contains **three channels: red, green and blue**.



## How to create colors with RGB?

Combine parts of the three primary colors **red**, **green** and **blue**.

Each of the primary colors can have a value in the range from 0 to 255.

					
R:	255	0	0	0	255
G:	0	255	0	0	255
B:	0	0	255	0	255

As shown in the above representation of a digital coloured image, each channel of each pixel has a value between 0 and 255.



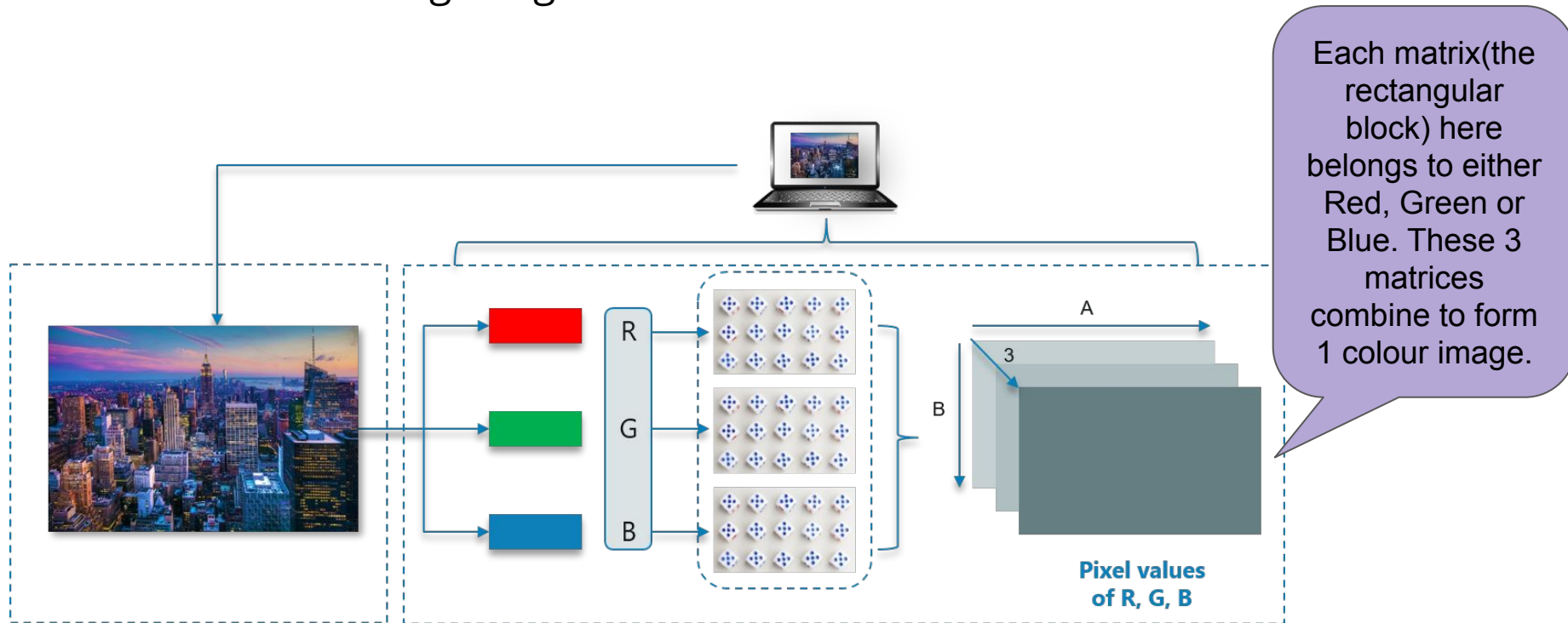


# Understanding Colour Images

For any color image in RGB format, there are 3 primary channels – Red, green and blue. How it works is pretty simple.

A matrix is formed for every primary color and later these matrices combine to provide a Pixel value for the individual R, G, B colors.

Consider the following image:



# Understanding Colour Images

---

**Note:** Images with 4, 5 or more channels also exist. An image can actually have up to 56 channels.

For example, images used in sophisticated graphics editing use four channels: three channels for RGB plus an extra alpha or "a" channel.

Also, a CMYK image has four channels: cyan, magenta, yellow, and key (black).

When we say 3 channels, we are considering images of the RGB type.

# Image Representation

---

Images are usually represented as **Height x Width x #Channels** where **#Channels is 3 for RGB images** and **1 for grayscale images**.

Sometimes you see **Width x Height x #Channels**, but the third dimension is the “channels.”

The size of the image in the previous slide can be calculated as  
**B x A x 3.**

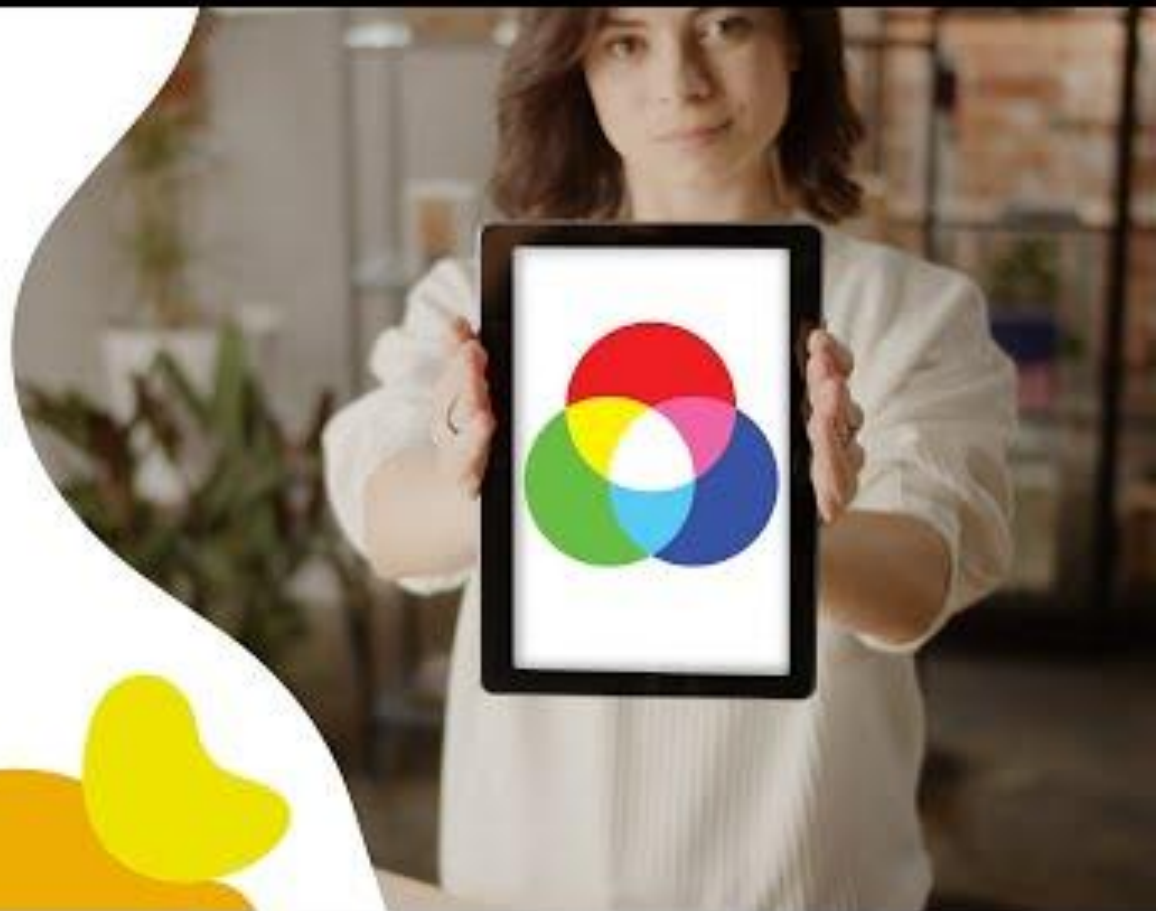
Can you now understand why?

(# means number of something)



TIPS

What is  
RGB?



---

The next question is, how can we say that the given image contains a picture of a dog?



---

Just being able to read the image is of no use if it cannot understand what it means, or if it cannot describe what the image is about, and what it contains.  
**This is where machine learning comes in.**

A machine (or a computer) can be taught how to understand an image and say what the image contains. This is an example of machine learning teaching a computer to understand and describe an image. It is similar to how we teach kids to identify different alphabets or differentiate between an apple and a banana by showing examples of each case. This is exactly how a computer learns to identify objects in an image.



---

Like humans have different skills and one of the skills is to identify an object in an image (a dog in the above image), computers have machine learning models, which can be thought of as a skill, to perform the same task. As humans need to be trained to perform a particular skill, computers need to train the machine learning model as well.

In both cases, training happens by examples. Similar to how a kid is taught to identify an apple, a machine learning model can be taught how to identify an apple in an image by giving several example images that contain an apple. From these example images, the model learning features of an apple, like its shape and colour. Now when a new image of an apple is presented to this computer with this model, it can use what it had learned about apples earlier and identify that this new image also contains apple.



# How we teach computers to understand pictures

---

[In this amazing TED Talk](#), **Fei Fei Li** - the co-director of the Stanford Institute for Human-Centered Artificial Intelligence talks about the [ImageNet Dataset](#) - *a large dataset of over 14 million images, designed by academics intended for computer vision research* that she created with fellow researchers.

The video dates back to 5 years ago but is still relevant today. ImageNet has played an important role in the advancement of computer vision. It is useful for many computer vision applications and is one of the most popular datasets for Computer Vision till date.

She also talks about the Convolutional Neural Networks and the crucial role they are playing in this domain. We will be learning more about CNN's in [tomorrow's session](#).





# Computer Vision

# What is Computer Vision?

---

Computer vision, often abbreviated as CV is a field of artificial intelligence that trains computers to interpret and understand the visual world.

Breaking it down into its constituent words,  
Computer Vision = giving vision to the computers.

Using digital images from cameras and videos and deep learning models, machines can accurately identify and classify objects — and then react to what they “see.”

Computer Vision is just one of the many applications of Deep Learning.



# What is Computer Vision?

---

The goal of computer vision is to understand the content of digital images. Typically, this involves developing methods that attempt to reproduce the capability of human vision.



# How ML and DL revolutionised Computer Vision

---

Machine learning helped solve many problems that were historically challenging for classical software development tools and approaches. For instance, years ago, machine learning engineers were able to create a software that could predict breast cancer survival windows better than human experts. However building the features of the software required the efforts of dozens of engineers and breast cancer experts and took a lot of time develop.

Deep learning provided a fundamentally different approach to doing machine learning. Deep learning relies on neural networks, a general-purpose function that can solve any problem representable through examples. When you provide a neural network with many labeled examples of a specific kind of data, it'll be able to extract common patterns between those examples and transform it into a mathematical equation that will help classify future pieces of information.



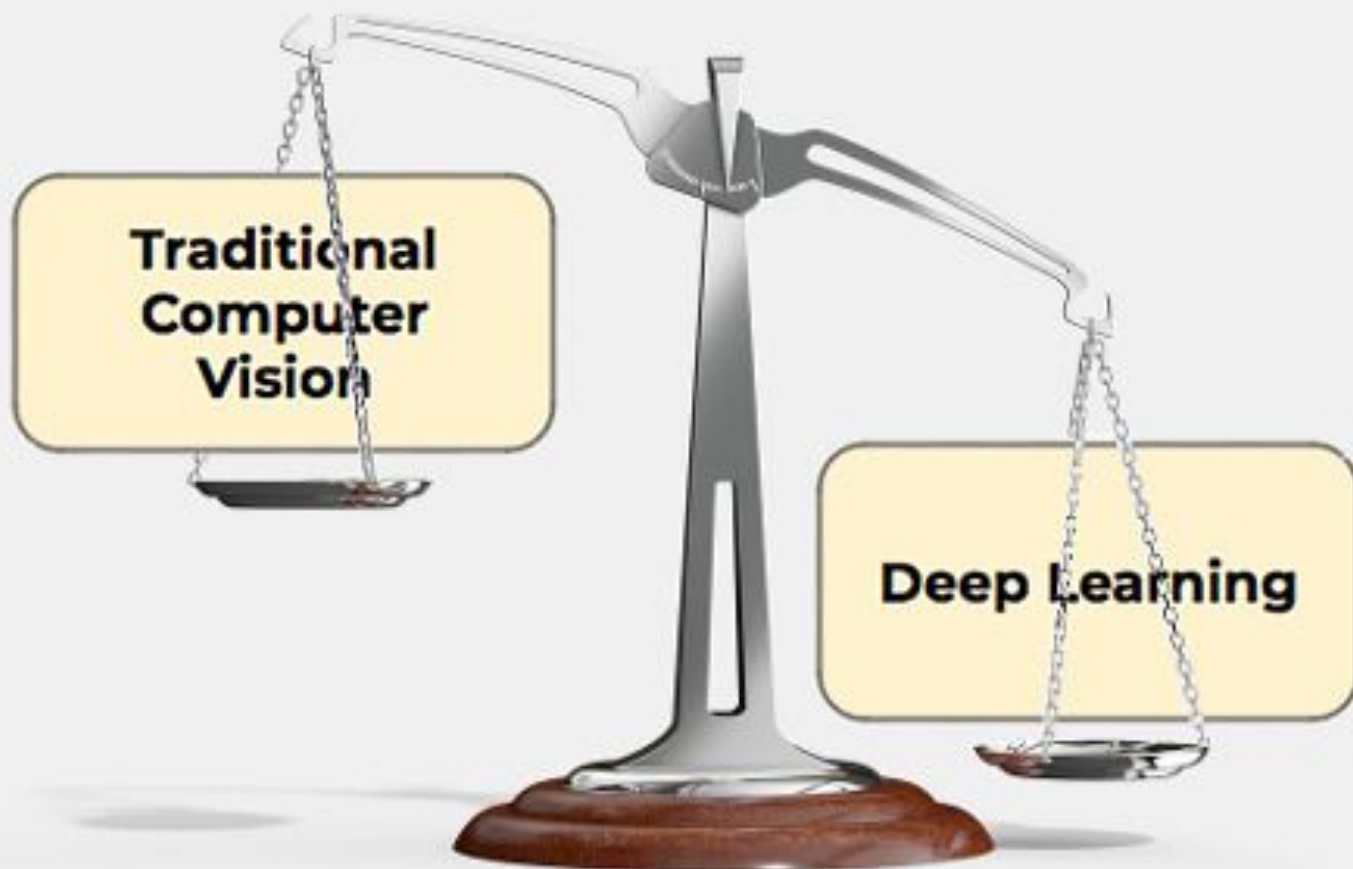


Image source: zbigatron

# DL - The star of Computer Vision

---

Deep learning is a very effective method to do computer vision. In most cases, creating a good deep learning algorithm comes down to gathering a large amount of labeled training data and tuning the parameters such as the type and number of layers of neural networks and training epochs. Compared to previous types of machine learning, deep learning is both easier and faster to develop and deploy.

Most of current computer vision applications such as cancer detection, self-driving cars and facial recognition make use of deep learning. Deep learning and deep neural networks have moved from the conceptual realm into practical applications thanks to availability and advances in hardware and cloud computing resources.



---

# Applications Of Computer Vision



Democratizing Data Science Learning

# CV In Self-Driving(Autonomous) Cars

Computer vision enables self-driving cars to **make sense of their surroundings**. Cameras capture video from different angles around the car and feed it to computer vision software, which then processes the images in real-time to find the extremities of roads, read traffic signs, detect other cars, objects and pedestrians. The self-driving car can then steer its way on streets and highways, avoid hitting obstacles, and (hopefully) safely drive its passengers to their destination.

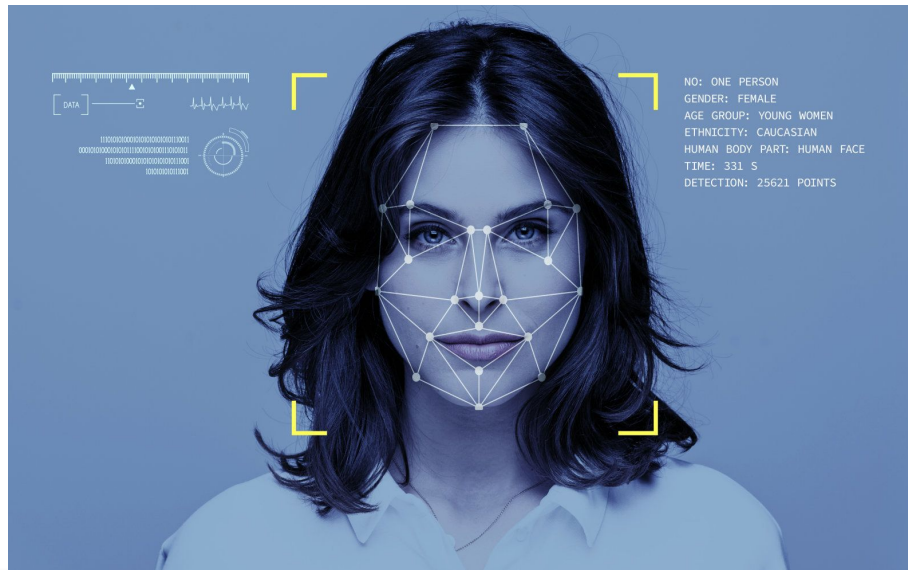


Democratizing Data Science Learning



# CV In Facial Recognition

Computer vision also plays an important role in facial recognition applications, the technology that enables computers to match images of people's faces to their identities. Computer vision algorithms detect facial features in images and compare them with databases of face profiles. Consumer devices use facial recognition to authenticate the identities of their owners. Social media apps use facial recognition to detect and tag users. Law enforcement agencies also rely on facial recognition technology to identify criminals in video feeds.



Democratizing Data Science Learning

# CV In Healthcare

Computer vision has also been an important part of advances in health-tech. Computer vision algorithms can help automate tasks such as detecting cancerous moles in skin images or finding symptoms in x-ray and MRI scans.

## Benign



## Malignant



---

# Computer Vision Tasks

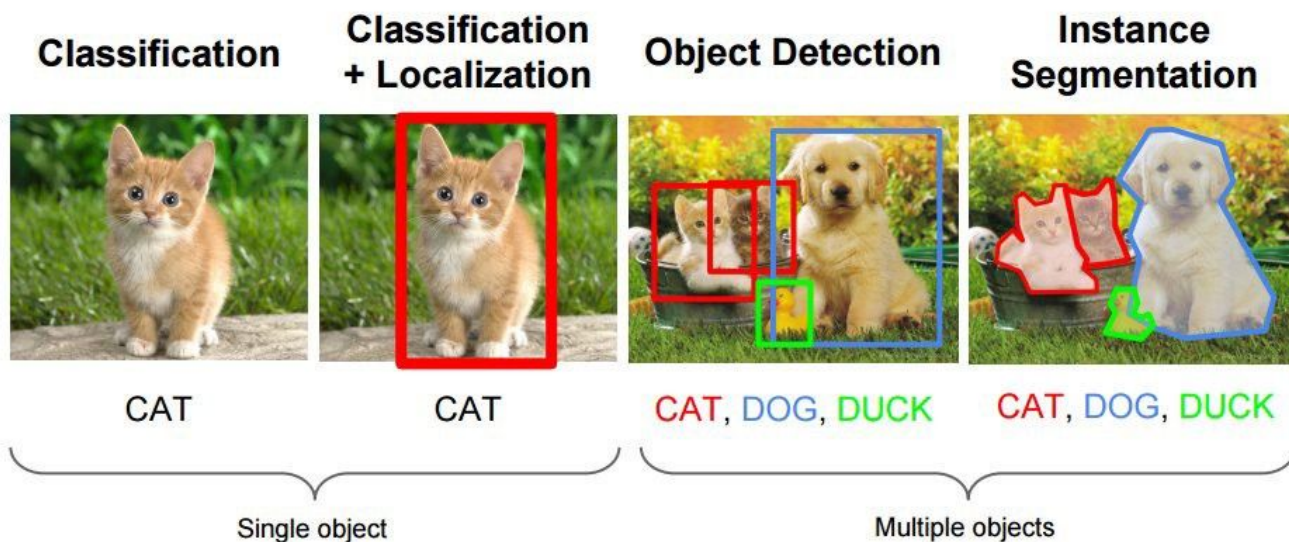


Democratizing Data Science Learning

# CV Tasks

Many popular computer vision applications involve trying to recognize things in photographs; for example:

- **Object Classification:** What broad category of object is in this photograph?
- **Object Identification:** Which type of a given object is in this photograph?
- **Object Verification:** Is the object in the photograph?
- **Object Detection:** Where are the objects in the photograph?
- **Object Landmark Detection:** What are the key points for the object in the photograph?
- **Object Segmentation:** What pixels belong to the object in the image?
- **Object Recognition:** What objects are in this photograph and where are they?



# CV Tasks

---

Outside of just recognition, other methods of analysis include:

- **Video motion analysis** uses computer vision to estimate the velocity of objects in a video, or the camera itself.
- In **image segmentation**, algorithms partition images into multiple sets of views.
- **Scene reconstruction** creates a 3D model of a scene inputted through images or video.
- In **image restoration**, noise such as blurring is removed from photos using Machine Learning based filters.

Any other application that involves understanding pixels through software can safely be labeled as computer vision.

Stating these tasks tackled by Computer Vision is not to overwhelm you but to introduce you to the immense capabilities Computer Vision holds.



# **Getting started with the basics of OpenCV and image processing**

# What is OpenCV?



OpenCV (Open Source Computer Vision) is a library with functions that mainly aim at real-time computer vision. OpenCV supports Deep Learning frameworks like Caffe, Tensorflow, Torch/PyTorch.



# Role of OpenCV in a deep learning computer vision project

---

- **OpenCV is not used** to train the neural networks—you should do that with a framework like TensorFlow or PyTorch, and then export the model to run in OpenCV.
- **OpenCV is used** to take a trained neural network model, prepare and preprocess images for it, apply it to the images and output results. You can also use it to combine neural networks with other computer vision algorithms available in OpenCV.





# OpenCV Applications

---

- Detecting and recognizing faces
- Identifying objects
- Classifying human actions in videos
- Tracking camera movements
- Tracking moving objects
- Extracting 3D models of objects
- Producing 3D point clouds from stereo cameras
- Stitching images together to produce an image of a scene
- Finding similar images from an image database
- Removing red eyes from images
- Following eye movements
- Recognizing scenery adding markers to enable augmented reality (AR)

And many more.



# Colour images in OpenCV

---

- In OpenCV color images in the RGB (Red, Green, Blue) color space have a 3-tuple associated with each pixel: (B, G, R) .
- Notice the **ordering is BGR rather than RGB**. This is because when OpenCV was first being developed many years ago the standard was BGR ordering. Over the years, the standard has now become RGB but OpenCV still maintains this “legacy” BGR ordering to ensure no existing code breaks.



# Installing OpenCV

---

Before you can start learning OpenCV you first need to install the OpenCV library on your system.

- If you are working on Jupyter notebook in your device, you can install OpenCV with:

```
pip install opencv-python
```

- Colab users don't need to worry about this. OpenCV is installed by default in Google Colab.



# Basic OpenCV Operations in practice

---

- The following notebook elaborates on the most commonly used commands in OpenCV:  
[https://github.com/dphi-official/Deep\\_Learning\\_Bootcamp/blob/master/OpenCV/DL\\_Day12\\_OpenCV.ipynb](https://github.com/dphi-official/Deep_Learning_Bootcamp/blob/master/OpenCV/DL_Day12_OpenCV.ipynb)



# Slide Download Link

---

- You can download the slides here:

[https://docs.google.com/presentation/d/1h0Vw2h\\_HISqP4Di9DCSk-A\\_wwlgGtQf-j4ZBJInSuJ0/edit?usp=sharing](https://docs.google.com/presentation/d/1h0Vw2h_HISqP4Di9DCSk-A_wwlgGtQf-j4ZBJInSuJ0/edit?usp=sharing)



# References

---

- <https://towardsdatascience.com/how-does-computer-understand-images-c1566d4537bf>
- <https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e>
- <https://www.pyimagesearch.com/2018/07/19/opencv-tutorial-a-guide-to-learn-opencv/>
- <https://missinglink.ai/guides/computer-vision/opencv-deep-learning/>
- <https://www.edureka.co/blog/python-opencv-tutorial/>



---

That's it for the day. Thank you!

Feel free to post any queries on the  
Discuss forum or #help channel on Slack

