**DAY 07**

**26.06.2023**

**Image recognition**

Image Recognition is the task of identifying images and categorizing them in one of several predefined distinct classes. So, image recognition software and apps can define what’s depicted in a picture and distinguish one object from another.

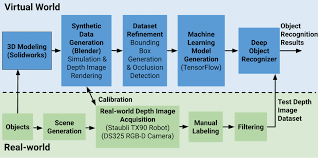
The field of study aimed at enabling machines with this ability is called **computer vision**. Being one of the computer visions (CV) tasks, image classification serves as the foundation for solving different CV problems, including:

***Image classification with localization*** – placing an image in a given class and drawing a bounding box around an object to show where it’s located in an image.

***Object detection***– categorizing multiple different objects in the image and showing the location of each of them with bounding boxes. So, it’s a variation of the image classification with localization tasks for numerous objects.

***Object (semantic) segmentation*** – identifying specific pixels belonging to each object in an image instead of drawing bounding boxes around each object as in object detection.

***Instance segmentation****–* differentiating multiple objects (instances) belonging to the same class (each person in a group).



***Reference***

<https://www.altexsoft.com/blog/image-recognition-neural-networks-use-cases/>

<https://machinelearningmastery.com/object-recognition-with-deep-learning/>

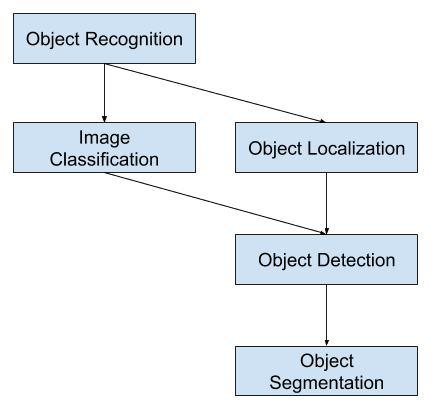
<https://viso.ai/computer-vision/image-recognition/>

<https://www.researchgate.net/publication/304294528_Image_recognition_based_on_deep_learning>

<https://www.analyticsvidhya.com/blog/2021/06/image-processing-using-cnn-a-beginners-guide/>

<https://www.sciencedirect.com/topics/engineering/image-recognition>

<https://medium.com/intro-to-artificial-intelligence/simple-image-classification-using-deep-learning-deep-learning-series-2-5e5b89e97926>



***Image recognition***

**YOLO: Real – Time Object Detection**

You Only Look Once (YOLO) is a state-of-the-art, real-time object detection algorithm introduced in 2015 by[Joseph Redmon](https://arxiv.org/search/cs?searchtype=author&query=Redmon%2C+J),[Santosh Divvala](https://arxiv.org/search/cs?searchtype=author&query=Divvala%2C+S),[Ross Girshick](https://arxiv.org/search/cs?searchtype=author&query=Girshick%2C+R), and[Ali Farhadi](https://arxiv.org/search/cs?searchtype=author&query=Farhadi%2C+A) in their famous research paper “[**You Only Look Once: Unified, Real-Time Object Detection**](https://arxiv.org/abs/1506.02640)”.

It frames the object detection problem as a regression problem instead of a classification task by spatially separating bounding boxes and associating probabilities to each of the detected images using a single convolutional neural network (CNN).

***YOLO Popular for Object Detection***

* Speed
* Detection accuracy
* Good generalization
* Open-source

***YOLO Architecture***

YOLO architecture is similar to[**GoogleNet**](https://arxiv.org/pdf/1409.4842.pdf)**.** It has overall 24 convolutional layers, four max-pooling layers, and two fully connected layers.

The architecture works as follows:

* Resizes the input image into 448x448 before going through the convolutional network.
* A 1x1 convolution is first applied to reduce the number of channels, which is then followed by a 3x3 convolution to generate a cuboidal output.
* The activation function under the hood is ReLU, except for the final layer, which uses a linear activation function.
* Some additional techniques, such as batch normalization and dropout, respectively regularize the model and prevent it from overfitting.

***Reference***

https://pjreddie.com/yolo/

https://www.datacamp.com/blog/yolo-object-detection-explained/