Computer Image Processing

<u>Lecture 11</u>

Image segmentation

Image segmentation

- **Image segmentation** is a prime domain of computer vision backed by a huge amount of research involving both image processing-based algorithms and learning-based techniques.
- In conjunction with being one of the most important domains in computer vision, Image Segmentation is also one of the oldest problem statements researchers pondered upon, with first works involving primitive region growing techniques and optimization approaches developed as early as 1970-72.

Image segmentation

1) What is Image Segmentation?
2) Types of Image Segmentation tasks
3) Traditional Image Segmentation techniques
4) Deep Learning-based Image Segmentation
5) Image Segmentation applications

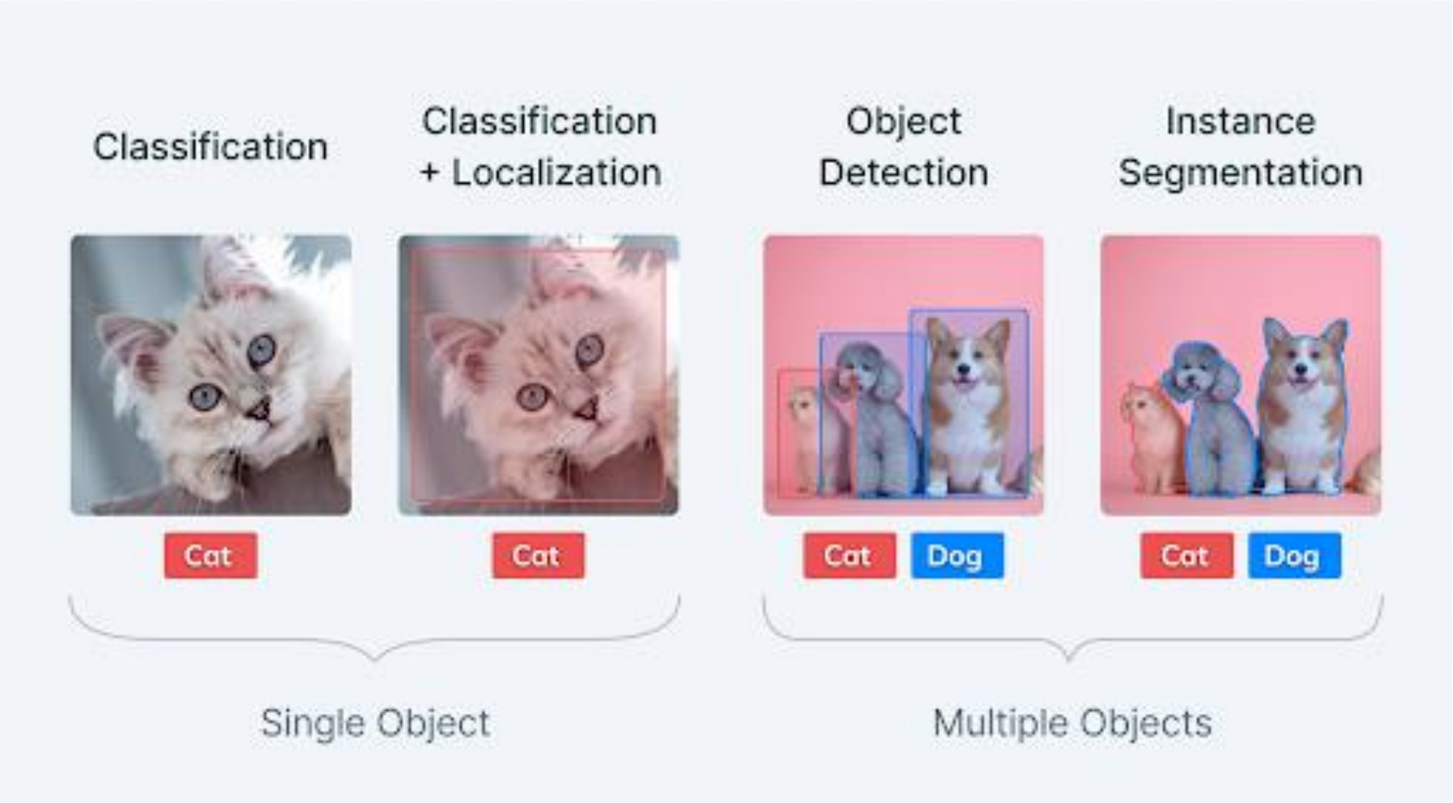
What is Image Segmentation?

Image segmentation is a sub-domain of computer vision and digital image processing which aims at grouping similar regions or segments of an image under their respective class labels.

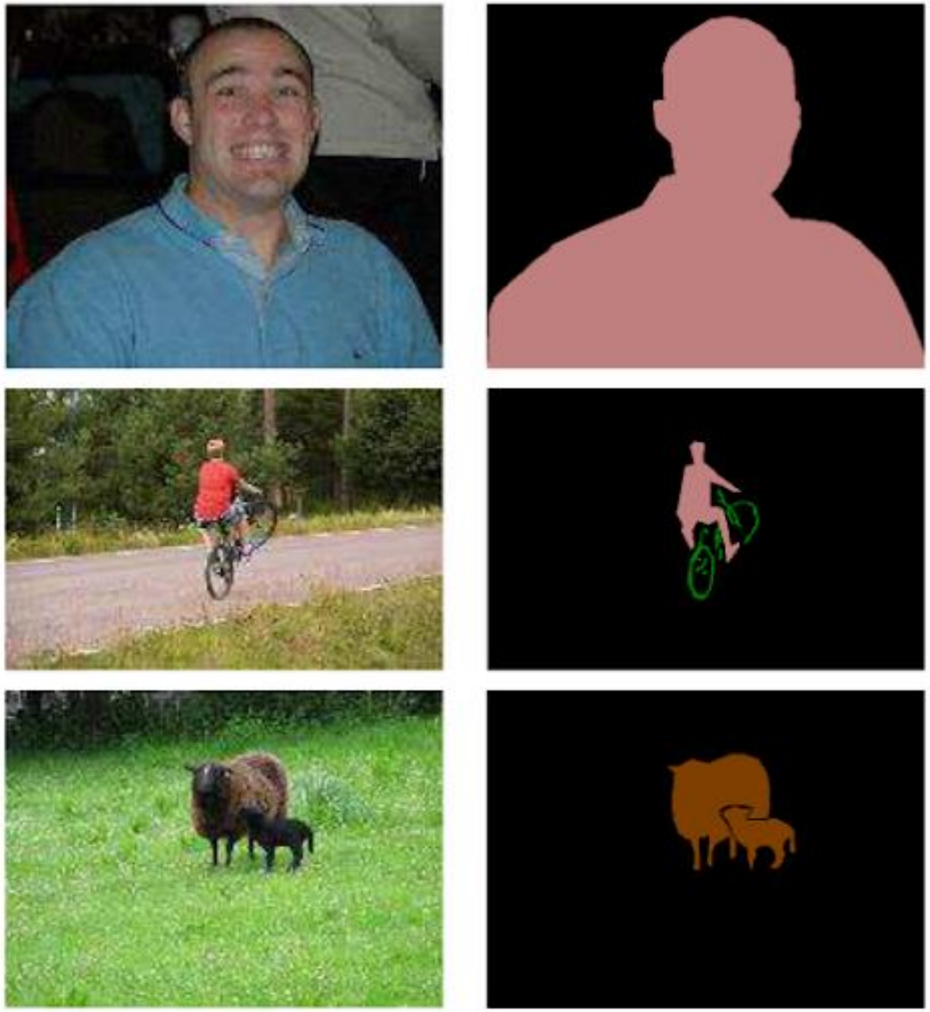
Since the entire process is digital, a representation of the analog image in the form of pixels is available, making the task of forming segments equivalent to that of grouping pixels.

Image segmentation is an extension of image classification where, in addition to classification, we perform localization. Image segmentation thus is a superset of image classification with the model pinpointing where a corresponding object is present by outlining the object's boundary.

Image segmentation



A typical segment map looks something like this



Types of Image Segmentation tasks

Image segmentation tasks can be classified into three groups based on the amount and type of information they convey.

While **semantic segmentation segments** out a broad boundary of objects belonging to a particular class, instance segmentation provides a segment map for each object it views in the image, without any idea of the class the object belongs to.

Panoptic segmentation is by far the most informative, being the conjugation of instance and semantic segmentation tasks. Panoptic segmentation gives us the segment maps of all the objects of any particular class present in the image.

Semantic segmentation

- **Semantic segmentation** refers to the classification of pixels in an image into semantic classes. Pixels belonging to a particular class are simply classified to that class with no other information or context taken into consideration.

- As might be expected, it is a poorly defined problem statement when there are closely grouped multiple instances of the same class in the image. An image of a crowd in a street would have a semantic segmentation model predict the entire crowd region as belonging to the "pedestrian" class, thus providing very little in-depth detail or information on the image.

Instance segmentation

Instance segmentation models classify pixels into categories on the basis of "instances" rather than classes.

An instance segmentation algorithm has no idea of the class a classified region belongs to but can segregate overlapping or very similar object regions on the basis of their boundaries.

If the same image of a crowd we talked about before is fed to an instance segmentation model, the model would be able to segregate each person from the crowd as well as the surrounding objects (ideally), but would not be able to predict what each region/object is an instance of.

Panoptic segmentation

- **Panoptic segmentation**, the most recently developed segmentation task, can be expressed as the combination of semantic segmentation and instance segmentation where each instance of an object in the image is segregated and the object's identity is predicted.

- Panoptic segmentation algorithms find large-scale applicability in popular tasks like self-driving cars where a huge amount of information about the immediate surroundings must be captured with the help of a stream of images.

Semantic Segmentation vs. Instance Segmentation vs. Panoptic Segmentation

(a) Image
(b) Semantic Segmentation
(c) Instance Segmentation
(d) Panoptic Segmentation

Traditional Image Segmentation techniques

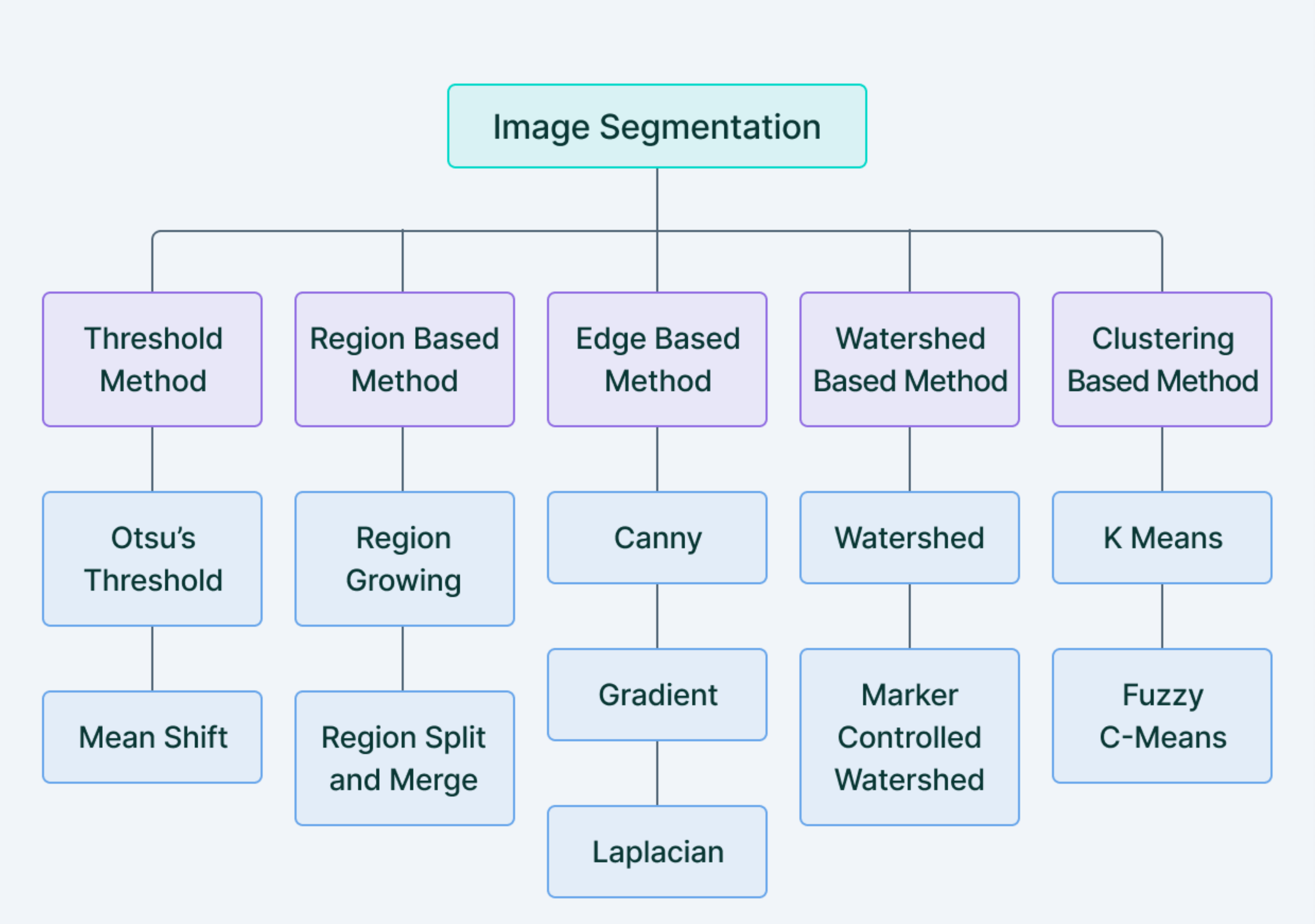
- Image segmentation originally started from Digital Image Processing coupled with optimization algorithms. These primitive algorithms made use of methods like region growing and snakes algorithm where they set up initial regions and the algorithm compared pixel values to gain an idea of the segment map.

- These methods took a local view of the features in an image and focused on local differences and gradients in pixels.

- Algorithms that took a global view of the input image came much later on with methods like adaptive thresholding, Otsu's algorithm, and clustering algorithms being proposed amongst classical image processing methods.

.

Image Segmentation

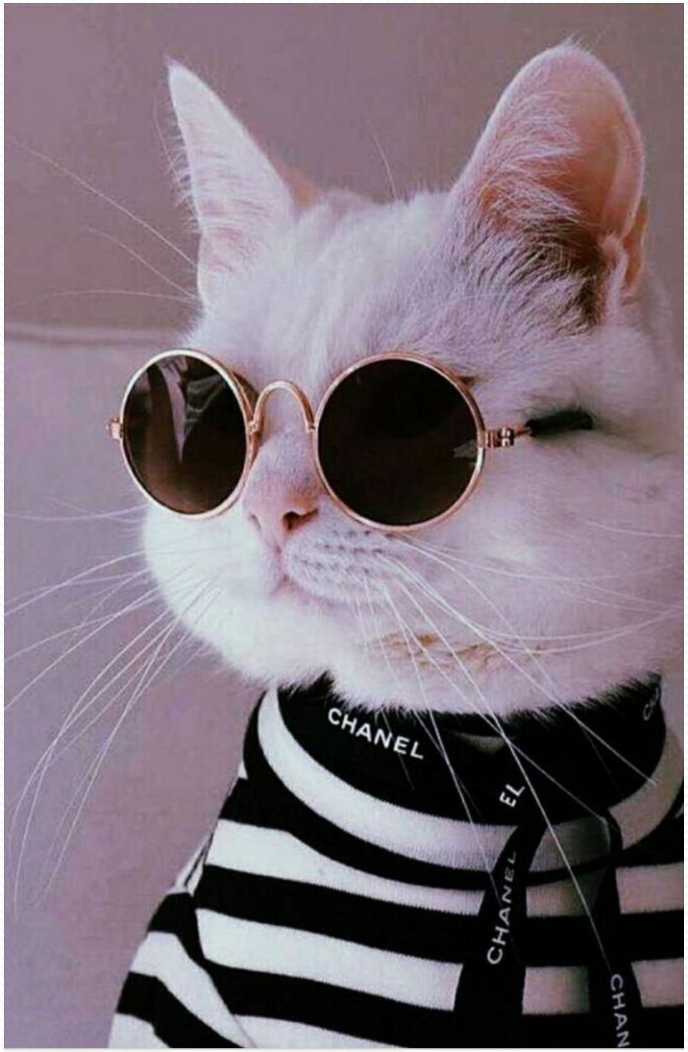


Thresholding

Thresholding is one of the easiest methods of image segmentation where a threshold is set for dividing pixels into two classes. Pixels that have values greater than the threshold value are set to 1 while pixels with values lesser than the threshold value are set to 0.

The image is thus converted into a binary map, resulting in the process often termed binarization. Image thresholding is very useful in case the difference in pixel values between the two target classes is very high, and it is easy to choose an average value as the threshold.

Thresholding is often used for image binarization so that further algorithms like contour detection and identification that work only on binary images can be used.



Source image

Image after thresholding

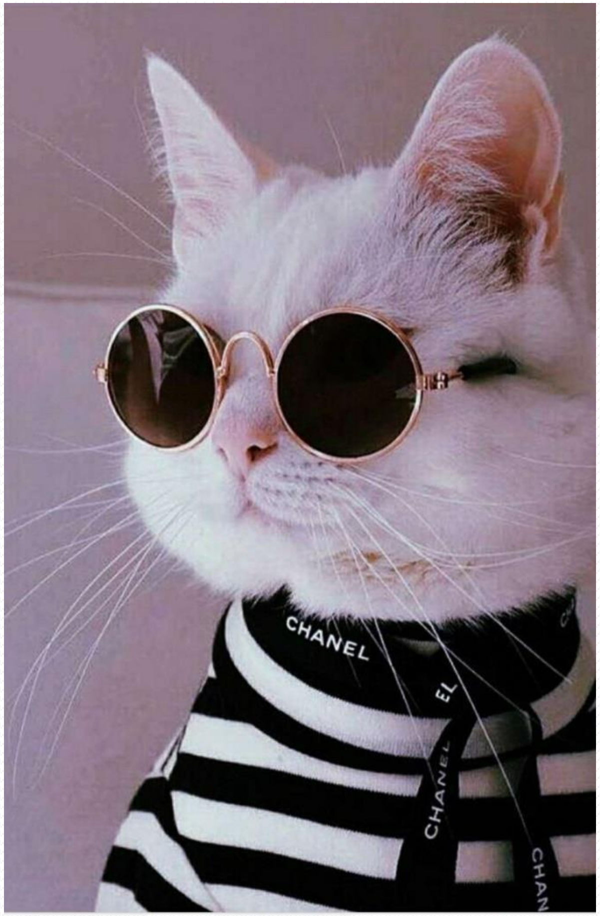
Region-Based Segmentation

- **Region-based segmentation** algorithms work by looking for similarities between adjacent pixels and grouping them under a common class.

- Typically, the segmentation procedure starts with some pixels set as seed pixels, and the algorithm works by detecting the immediate boundaries of the seed pixels and classifying them as similar or dissimilar.

- The immediate neighbors are then treated as seeds and the steps are repeated till the entire image is segmented. An example of a similar algorithm is the popular watershed algorithm for segmentation that works by starting from the local maxima of the euclidean distance map and grows under the constraint that no two seeds can be classified as belonging to the same region or segment map.

Region-Based Segmentation



Source image

Image after watershed segmentation

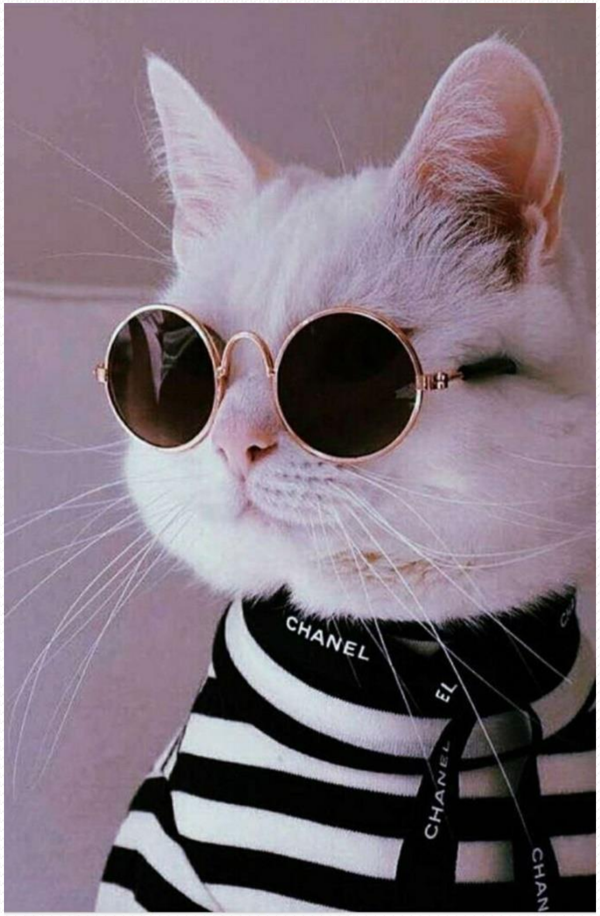
Edge Segmentation

Edge segmentation, also called edge detection, is the task of detecting edges in images.

From a segmentation-based viewpoint, we can say that edge detection corresponds to classifying which pixels in an image are edge pixels and singling out those edge pixels under a separate class correspondingly.

Edge detection is generally performed by using special filters that give us edges of the image upon convolution. These filters are calculated by dedicated algorithms that work on estimating image gradients in the x and y coordinates of the spatial plane.

An example of edge detection using the Canny edge detection algorithm, one of the most popular edge detection algorithms is shown below.



Source image

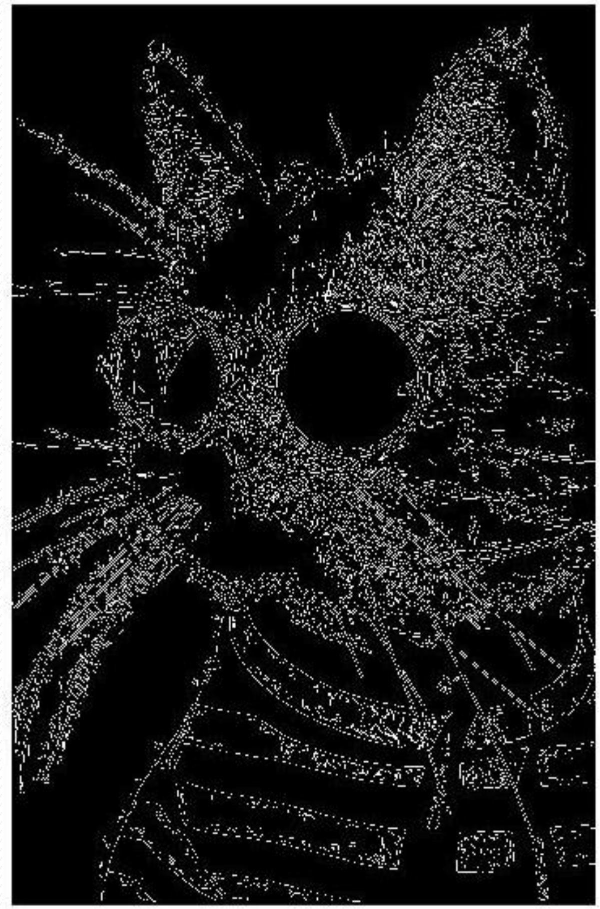


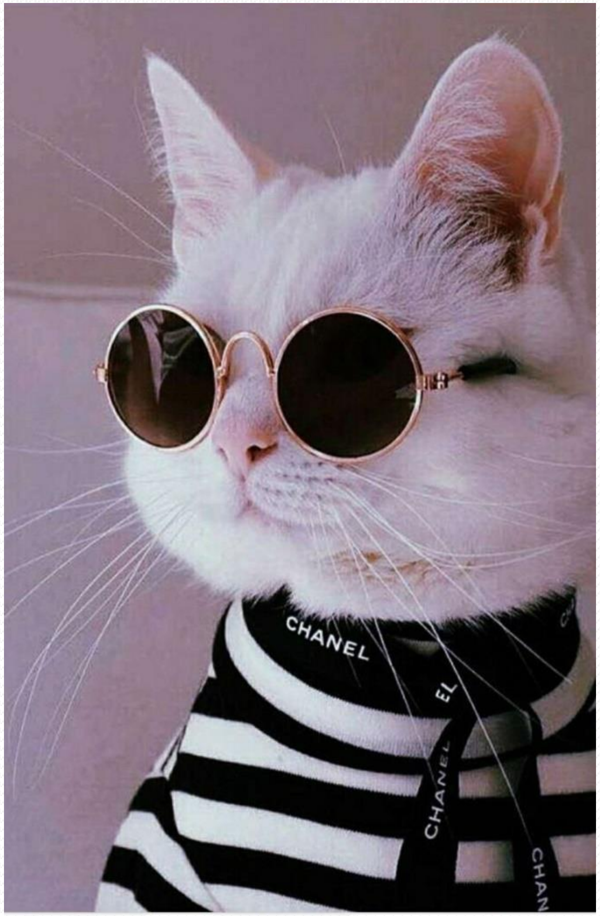
Image after edge detection using the Canny

Clustering-based Segmentation

Modern segmentation procedures that depend on image processing techniques generally make use of clustering algorithms for segmentation.

Clustering algorithms perform better than their counterparts and can provide reasonably good segments in a small amount of time. Popular algorithms like the K-means clustering algorithms are unsupervised algorithms that work by clustering pixels with common attributes together as belonging to a particular segment.

K-means clustering, in particular, takes all the pixels into consideration and clusters them into “k” classes. Differing from region-growing methods, clustering-based methods do not need a seed point to start segmenting from.

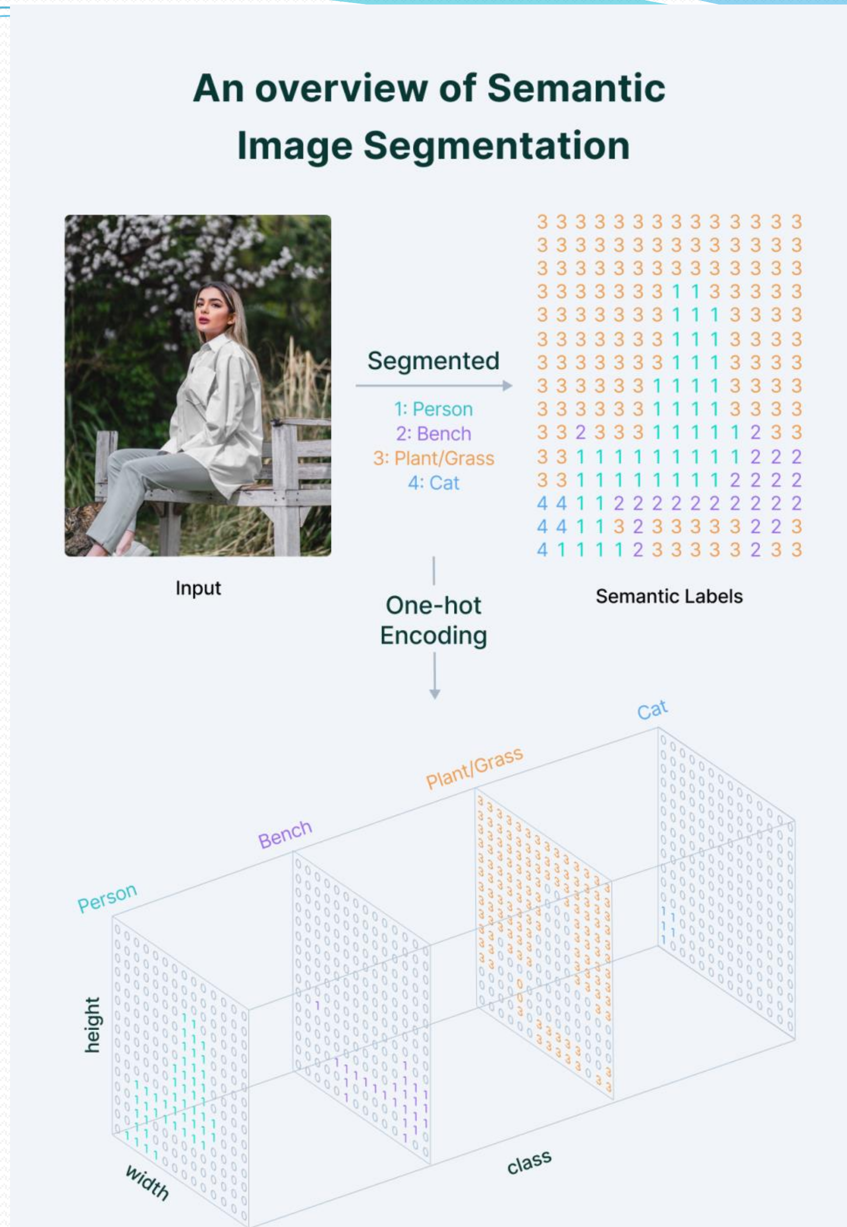


Source image

After segmentation by the method of k-means

Deep Learning-based methods

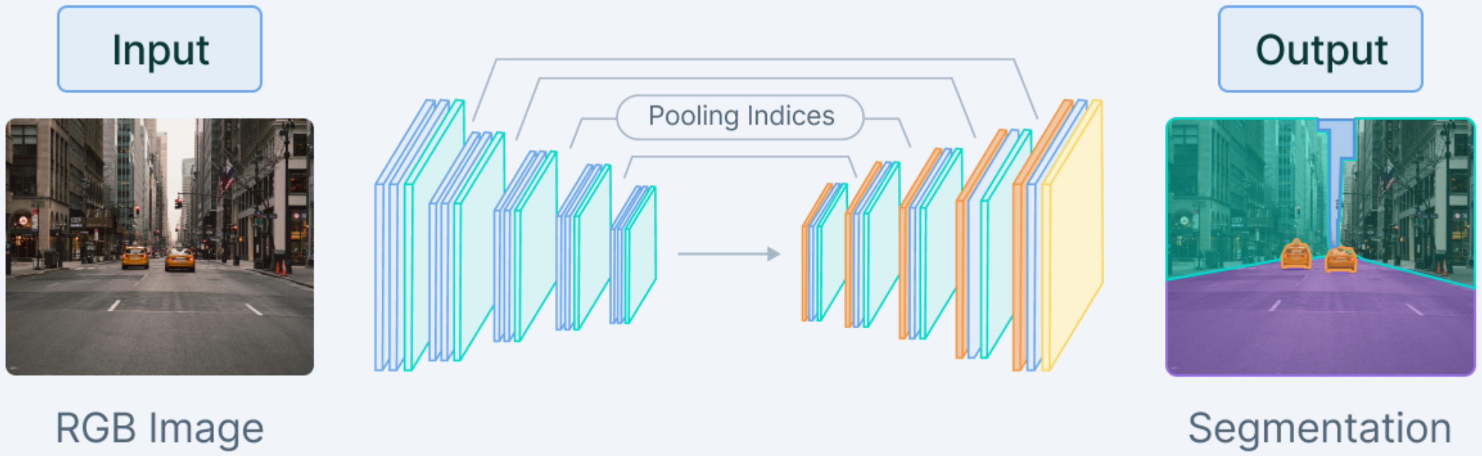
- **Semantic segmentation** models provide segment maps as outputs corresponding to the inputs they are fed.
- These segment maps are often n-channeled with n being the number of classes the model is supposed to segment. Each of these n-channels is binary in nature with object locations being "filled" with ones and empty regions consisting of zeros. The ground truth map is a single channel integer array the same size as the input and has a range of "n", with each segment "filled" with the index value of the corresponding classes (classes are indexed from 0 to n-1).


An overview of Semantic Image Segmentation
Input
Segmented
1: Person
2: Bench
3: Plant/Grass
4: Cat
3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 1 1 3 3 3 3 3
3 3 3 3 3 3 3 1 1 1 3 3 3 3
3 3 3 3 3 3 3 1 1 1 3 3 3 3
3 3 3 3 3 3 3 1 1 1 3 3 3 3
3 3 3 3 3 3 1 1 1 1 3 3 3 3
3 3 3 3 3 3 1 1 1 1 3 3 3 3
3 3 2 3 3 3 1 1 1 1 1 2 3 3
3 3 1 1 1 1 1 1 1 1 1 2 2 2
3 3 1 1 1 1 1 1 1 1 2 2 2 2
4 4 1 1 2 2 2 2 2 2 2 2 2 2
4 4 1 1 3 2 3 3 3 3 3 2 2 3
4 1 1 1 1 2 3 3 3 3 3 2 3 3
Semantic Labels
One-hot Encoding
Person
Bench
Plant/Grass
Cat
height
width
class

Convolutional Encoder-Decoder Architecture

- Encoder decoder architectures for semantic segmentation became popular with the onset of works like SegNet (by Badrinarayanan *et. a.*) in 2015.
- SegNet proposes the use of a combination of convolutional and downsampling blocks to squeeze information into a bottleneck and form a representation of the input. The decoder then reconstructs input information to form a segment map highlighting regions on the input and grouping them under their classes.
- Finally, the decoder has a sigmoid activation at the end that squeezes the output in the range (0,1).

Convolutional encoder-decoder



Applications of Image Segmentation

- Image segmentation is an important step in artificial vision. Machines need to divide visual data into segments for segment-specific processing to take place.
- Image segmentation thus finds its way in prominent fields like Robotics, Medical Imaging, Autonomous Vehicles, and Intelligent Video Analytics.
- Apart from these applications, Image segmentation is also used by satellites on aerial imagery for segmenting out roads, buildings, and trees.
- Here are a few of the most popular real-world use cases of image segmentation.

Robotics (Machine Vision)

- Image segmentation aids machine perception and locomotion by pointing out objects in their path of motion, enabling them to change paths effectively and understand the context of their environment.
- Apart from locomotion, segmentation of images helps machines segregate the objects they are working with and enables them to interact with real-world objects using only vision as a reference. This allows the machine to be useful almost anywhere without much constraint.

1. Instance segmentation for robotic grasping
2. Recycling object picking
3. Autonomous navigation and SLAM

Medical imaging

- Medical Imaging is an important domain of computer vision that focuses on the diagnosis of diseases from visual data, both in the form of simple visual data and biomedical scans.

- Segmentation forms an important role in medical imaging as it helps doctors identify possible malignant features in images in a fast and accurate manner.

- Using image segmentation, diagnosis of diseases can not only be speeded up but can also be made cheaper, thereby benefiting thousands across the globe.

- X-Ray segmentation
- CT scan organ segmentation
- Dental instance segmentation
- Digital pathology cell segmentation
- Surgical video annotation