*Segmentation is the process by which a digital image is partitioned into various subgroups (of pixels) called Image Objects, which can reduce the complexity of the image, and thus analysing the image becomes simpler.* image segmentation algorithms to split and group a certain set of pixels together from the image. By doing so, we are actually assigning labels to pixels and the pixels with the same label fall under a category where they have some or the other thing common in them.
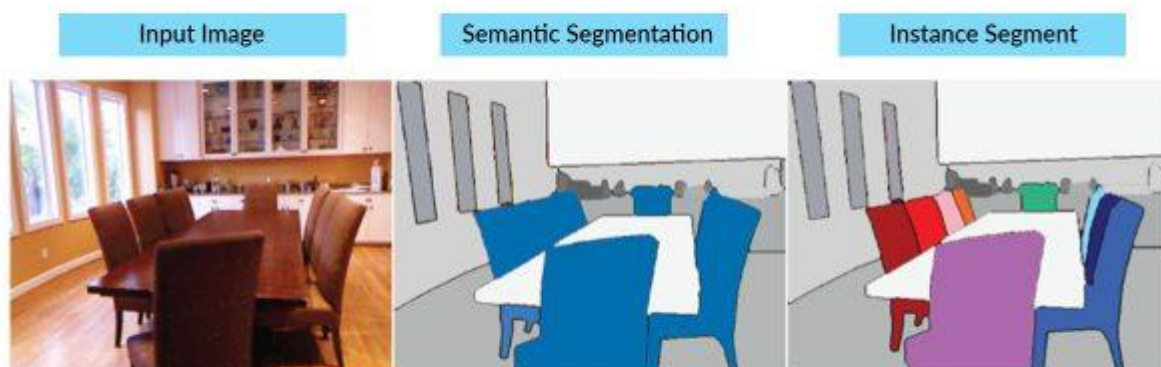
Using these labels, we can specify boundaries, draw lines, and separate the most required objects in an image from the rest of the not-so-important ones. In the below example, from a main image on the left, we try to get the major components, e.g. chair, table etc. and hence all the chairs are colored uniformly. In the next tab, we have detected instances, which talk about individual objects, and hence the all the chairs have different colors.

This is how different methods of segmentation of images work in varying degrees of complexity and yield different levels of outputs.

Types of Image Segmentation
1. The Approach
Whenever one tries to take a bird's eye view of the Image Segmentation tasks, one gets to observe a crucial process that happens here – object identification. Any simple to complex application areas, everything is based out of object detection.

And as we discussed earlier, detection is made possible because the image segmentation algorithms try to – if we put it in lay man's terms –

collect similar pixels together and separate out dissimilar pixels. This is done by following two approaches based on the image properties:

1.1. Similarity Detection (Region Approach)
This fundamental approach relies on detecting similar pixels in an image – based on a threshold, region growing, region spreading, and region merging. Machine learning algorithms like clustering relies on this approach of similarity detection on an unknown set of features, so does classification, which detects similarity based on a pre-defined (known) set of features.

1.2. Discontinuity Detection (Boundary Approach)
This is a stark opposite of similarity detection approach where the algorithm rather searches for discontinuity. Image Segmentation Algorithms like Edge Detection, Point Detection, Line Detection follows this approach – where edges get detected based on various metrics of discontinuity like intensity etc.

Image Source: scikit-image

2. The Types of Techniques
Based on the two approaches, there are various forms of techniques that are applied in the design of the Image Segmentation Algorithms. These techniques are employed based on the type of image that needs to be processed and analysed and they can be classified into three broader categories as below:

2.1 Structural Segmentation Techniques
These sets of algorithms require us to firstly, know the structural information about the image under the scanner. This can include the pixels, pixel density, distributions, histograms, color distribution etc. Second, we need to have the structural information about the region that we are about

to fetch from the image – this section deals with identifying our target area, which is highly specific to the business problem that we are trying to solve. Similarity based approach will be followed in these sets of algorithms.

2.2 Stochastic Segmentation Techniques

In these group of algorithms, the primary information that is required for them is to know the discrete pixel values of the full image, rather than pointing out the structure of the required portion of the image. This proves to be advantageous in the case of a larger group of images, where a high degree of uncertainty exists in terms of the required object within an object. ANN and Machine Learning based algorithms that use k-means etc. make use of this approach.

2.3 Hybrid Techniques

As the name suggests, these algorithms for image segmentation make use of a combination of structural method and stochastic methods i.e., use both the structural information of a region as well as the discrete pixel information of the image.

Image segmentation Techniques

Based on the image segmentation approaches and the type of processing that is needed to be incorporated to attain a goal, we have the following techniques for image segmentation.

1. Threshold Method
2. Edge Based Segmentation
3. Region Based Segmentation
4. Clustering Based Segmentation
5. Watershed Based Method
6. Artificial Neural Network Based Segmentation

Techniques	Description	Advantages	Disadvantages
Thresholding Method	Focuses on finding peak values based on the histogram of the image	Doesn't require complicated pre-processing, simple	Many details can get omitted, threshold errors are common
Edge Based Method	based on discontinuity detection unlike	Good for images having better contrast between objects.	Not suitable for noisy images
Region-Based	based on partitioning an image into	Works really well for images with a considerate amount of noise,	Time and memory consuming
Traditional Segmentation	Divides image into k number of homogenous, mutually	Proven methods, reinforced with fuzzy logic and more useful for real-time application.	Determining cost function for minimization can be
Watershed Method	based on topological interpretation of image	segments obtained are more stable, detected boundaries are	Gradient calculation for ridges is complex.

1. Threshold Method

This is perhaps the <mark>most basic and yet powerful technique to identify the required objects in an image.</mark> Based on the intensity, the pixels in an image get divided by comparing the pixel's intensity with a threshold value. The threshold method proves to be advantageous when the objects in the image in question are assumed to be having <mark>more/LESS intensity than the background (and unwanted components) of the image.</mark>
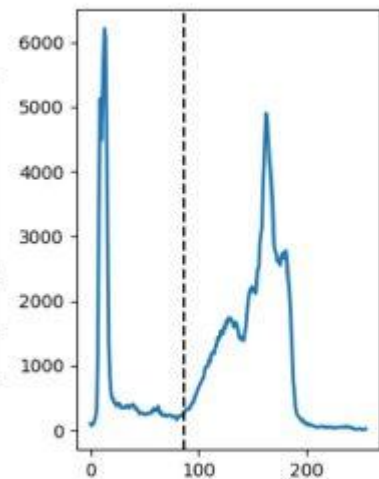
At its simpler level, the threshold value T is considered to be a constant. But that approach may be futile considering the amount of noise (unwanted information) that the image contains. So, we can either keep it constant or change it dynamically based on the image properties and thus obtain better results. Based on that, thresholding is of the following types:

1.1 Simple Thresholding

This technique replaces the pixels in an image with either black or white. If the intensity of a pixel ($I_{i,j}$) at position (i,j) is less than the threshold (T), then we replace that with black and if it is more, then we replace that pixel with white. This is a binary approach to thresholding.

1.2 Otsu's Binarization

In global thresholding, we had used an arbitrary value for threshold value and it remains a constant. The major question here is, how can we define

and determine the correctness of the selected threshold? A simpler but rather inept method is to trial and see the error.

But, on the contrary, let us take an image whose histogram has two peaks (bimodal image), one for the background and one for the foreground. According to Otsu binarization, for that image, we can approximately take a value in the middle of those peaks as the threshold value. So in simply put, it automatically calculates a threshold value from image histogram for a bimodal image.

The disadvantage here, however, is for images that are not bimodal, the image histogram has multiple peaks, or one of the classes (peaks) present has high variance.

However, Otsu's Binarization is widely used in document scans, removing unwanted colors from a document, pattern recognition etc.

1.3 Adaptive Thresholding
A global value as threshold value may not be good in all the conditions where an image has different background and foreground lighting conditions in different actionable areas. We need an adaptive approach that can change the threshold for various components of the image. In this, the algorithm divides the image into various smaller portions and calculates the threshold for those portions of the image.

Hence, we obtain different thresholds for different regions of the same image. This in turn gives us better results for images with varying illumination. The algorithm can automatically calculate the threshold value. The threshold value can be the mean of neighborhood area or it can be the weighted sum of neighborhood values where weights are a Gaussian window (a window function to define regions).

2. Edge Based Segmentation

Edge detection is the process of locating edges in an image which is a very important step towards understanding image features. It is believed that edges consist of meaningful features and contains significant information. It significantly reduces the size of the image that will be processed and filters out information that may be regarded as less relevant, preserving and focusing solely on the important structural properties of an image for a business problem.

Edge-based segmentation algorithms work to detect edges in an image, based on various discontinuities in grey level, colour, texture, brightness, saturation, contrast etc. To further enhance the results, supplementary processing steps must follow to concatenate all the edges into edge chains that correspond better with borders in the image.

Image Source: researchgate.net

Edge detection algorithms fall primarily into two categories – Gradient based methods and Gray Histograms. Basic edge detection operators like sobel operator, canny, Robert's variable etc are used in these algorithms.

These operators aid in detecting the edge discontinuities and hence mark the edge boundaries. The end goal is to reach at least a partial segmentation using this process, where we group all the local edges into a new binary image where only edge chains that match the required existing objects or image parts are present.

3. Region Based Segmentation

The region based segmentation methods involve the algorithm creating segments by dividing the image into various components having similar characteristics. These components, simply put, are nothing but a set of pixels. Region-based image segmentation techniques initially search for some seed points – either smaller parts or considerably bigger chunks in the input image.

Next, certain approaches are employed, either to add more pixels to the seed points or further diminish or shrink the seed point to smaller segments and merge with other smaller seed points. Hence, there are two basic techniques based on this method.

3.1 Region Growing

It's a bottom to up method where we begin with a smaller set of pixel and start accumulating or iteratively merging it based on certain pre-determined similarity constraints. Region growth algorithm starts with choosing an arbitrary seed pixel in the image and compare it with its neighboring pixels.

If there is a match or similarity in neighboring pixels, then they are added to the initial seed pixel, thus increasing the size of the region. When we reach the saturation and hereby, the growth of that region cannot proceed further, the algorithm now chooses another seed pixel, which necessarily does not belong to any region(s) that currently exists and start the process again.

Region growing methods often achieve effective Segmentation that corresponds well to the observed edges. But sometimes, when the algorithm lets a region grow completely before trying other seeds, that usually biases the segmentation in favour of the regions which are segmented first. To counter this effect, most of the algorithms begin with the user inputs of similarities first, no single region is allowed to dominate and grow completely and multiple regions are allowed to grow simultaneously.

Region growth, also a pixel based algorithm like thresholding but the major difference is thresholding extracts a large region based out of similar pixels, from anywhere in the image whereas region-growth extracts only the

adjacent pixels. Region growing techniques are preferable for noisy images, where it is highly difficult to detect the edges.

3.2 Region Splitting and Merging

The splitting and merging based segmentation methods use two basic techniques done together in conjunction – region splitting and region merging – for segmenting an image. Splitting involves iteratively dividing an image into regions having similar characteristics and merging employs combining the adjacent regions that are somewhat similar to each other.

A region split, unlike the region growth, considers the entire input image as the area of business interest. Then, it would try matching a known set of parameters or pre-defined similarity constraints and picks up all the pixel areas matching the criteria. This is a divide and conquers method as opposed to the region growth algorithm.

Now, the above process is just one half of the process, after performing the split process, we will have many similarly marked regions scattered all across the image pixels, meaning, the final segmentation will contain scattered clusters of neighbouring regions that have identical or similar properties. To complete the process, we need to perform merging, which after each split which compares adjacent regions, and if required, based on similarity degrees, it merges them. Such algorithms are called split-merge algorithms.

4. Clustering Based Segmentation Methods

Clustering algorithms are unsupervised algorithms, unlike Classification algorithms, where the user has no pre-defined set of features, classes, or groups. Clustering algorithms help in fetching the underlying, hidden information from the data like, structures, clusters, and groupings that are usually unknown from a heuristic point of view.

The clustering based techniques segment the image into clusters or disjoint groups of pixels with similar characteristics. By the virtue of basic Data Clustering properties, the data elements get split into clusters such that elements in same cluster are more similar to each other as compared to other clusters. Some of the more efficient clustering algorithms such as k-means, improved k means, fuzzy c-mean (FCM) and improved fuzzy c mean algorithm (IFCM) are being widely used in the clustering based approaches proposed.

K means clustering is a chosen and popular method because of its simplicity and computational efficiency. The Improved K-means algorithm can minimize the number of iterations usually involved in a k-means

algorithm. FCM algorithm allows data points, (pixels in our case) to belong to multiple classes with varying degrees of membership. The slower processing time of an FCM is overcome by improved FCM.
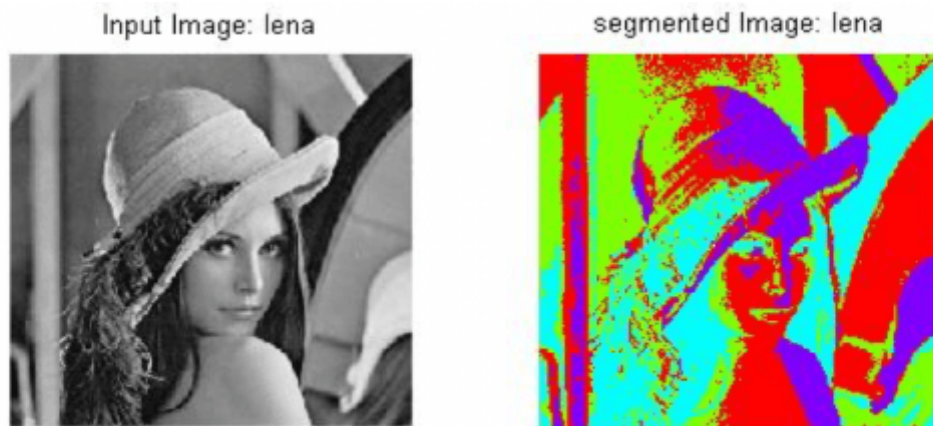

Image Source: researchgate.net

A massive value add of clustering based ML algorithms is that we can measure the quality of the segments that get generated by using several statistical parameters such as: Silhouette Coefficient, rand index (RI) etc.

4.1 k-means clustering

K-means is one of the simplest unsupervised learning algorithms which can address the clustering problems, in general. The process follows a simple and easy way to classify a given image through a certain number of clusters which are fixed apriori. The algorithm actually starts at this point where the image space is divided into k pixels, representing k group centroids. Now, each of the objects is then assigned to the group based on its distance from the cluster. When all the pixels are assigned to all the clusters, the centroids now move and are reassigned. These steps repeat until the centroids can no longer shift.

At the convergence of this algorithm, we have areas within the image, segmented into "K" groups where the constituent pixels show some levels of similarity.

4.2 Fuzzy C Means

k-means, as discussed in the previous section, allows for dividing and grouping together the pixels in an image that have certain degrees of similarity. One of the striking features in k-means is that the groups and their members are completely mutually exclusive. A Fuzzy C Means clustering technique allows the data points, in our case, the pixels to be clustered in more than one cluster. In other words, a group of pixels can

belong to more than one cluster or group but they can have varying levels of associativity per group. The FCM algorithm has an optimization function associated with it and the convergence of the algorithm depends on the minimization of this function.

At the convergence of this algorithm, we have areas within the image, segmented into "C" groups where the constituent pixels inside a group show some levels of similarity, and also they will have a certain degree of association with other groups as well.

5. Watershed Based Methods

Watershed is a ridge approach, also a region-based method, which follows the concept of topological interpretation. We consider the analogy of geographic landscape with ridges and valleys for various components of an image. The slope and elevation of the said topography are distinctly quantified by the gray values of the respective pixels – called the gradient magnitude. Based on this 3D representation which is usually followed for Earth landscapes, the watershed transform decomposes an image into regions that are called "catchment basins". For each local minimum, a catchment basin comprises all pixels whose path of steepest descent of gray values terminates at this minimum.
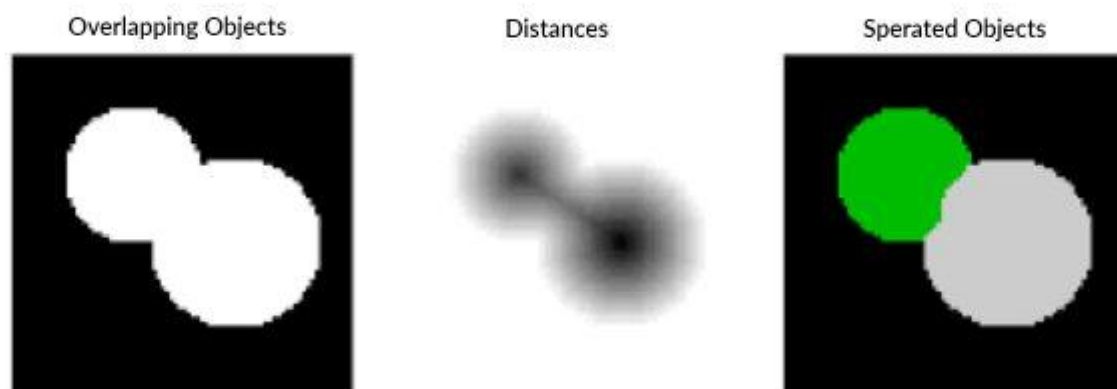
Image Source: scikit-image

In a simple way of understanding, the algorithm considers the pixels as a "local topography" (elevation), often initializing itself from user-defined markers. Then, the algorithm defines something called "basins" which are

the minima points and hence, basins are flooded from the markers until basins meet on watershed lines. The watersheds that are so formed here, they separate basins from each other. Hence the picture gets decomposed because we have pixels assigned to each such region or watershed.

6. Artificial Neural Network Based Segmentation Method

The approach of using Image Segmentation using neural networks is often referred to as Image Recognition. It uses AI to automatically process and identify the components of an image like objects, faces, text, hand-written text etc. Convolutional Neural Networks are specifically used for this process because of their design to identify and process high-definition image data.

An image, based on the approach used, is considered either as a set of vectors (colour annotated polygons) or a raster (a table of pixels with numerical values for colors). The vector or raster is turned into simpler components that represent the constituent physical objects and features in an image. Computer vision systems can logically analyze these constructs, by extracting the most important sections, and then by organizing data through feature extraction algorithms and classification algorithms.
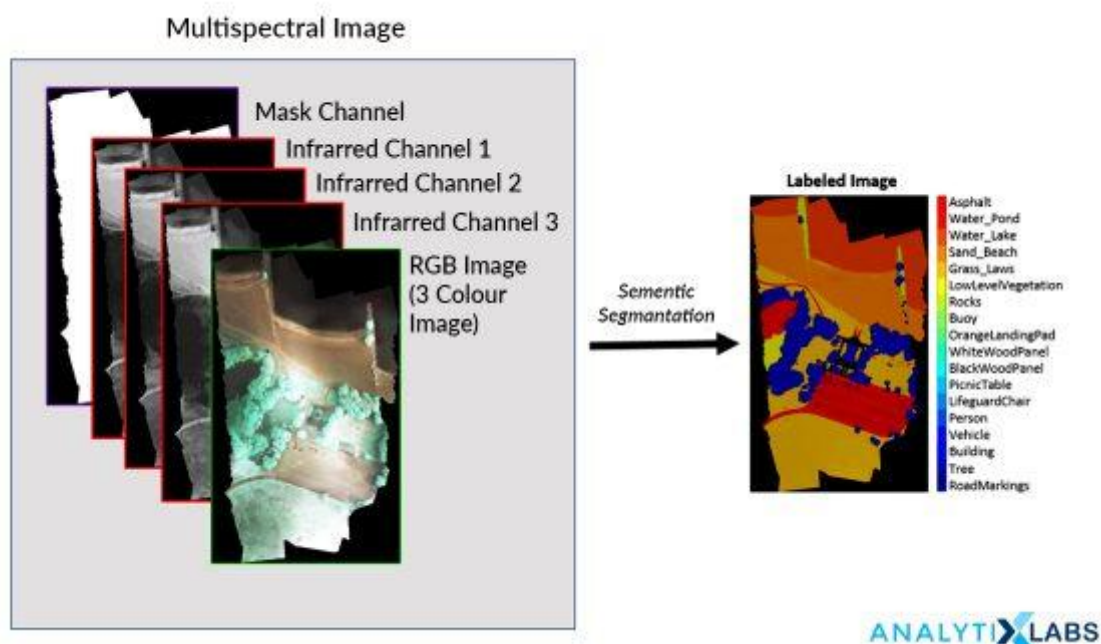

Image Source: mathworks.com

These algorithms are widely used in a variety of industries and applications. E-commerce industry uses it for providing relevant products to

users for their search requirements and browsing history. The manufacturing industry uses it for anomaly detection, detecting damaged objects, ensuring worker safety etc. Image Recognition is famously used in education and training for visually impaired, speech impaired students. Although Neural Nets are time consuming when it comes to training the data, the end results have been very promising and the application of these has been highly successful.