Vinit Ramesh Gore (M20CS064), Vikanksh Nath (MT19AI024)

KMeans

Description

The KMeans algorithm is a popular clustering algorithm. It finds application in bioimage segmentation for clustering different organs out of a given image.

Algorithm KMeans(X: datapoints, k: number of clusters):   // X is a matrix of image properties like intensities or texture. Each data point is an intensity vector having 3 dimensions corresponding to RGB values.

Step 1: Initialize k centroids randomly from the given dataset.

Step 2: Run Step 3 and Step 4 until I number of iterations are completed.

Step 3: Allocate each data point to the nearest centroid in order to form clusters.

Step 4: Find average of data points of each cluster and assign them as new centroids.

Results

(i) Left Input Image; Right Output of k-means image segmentation with k =4
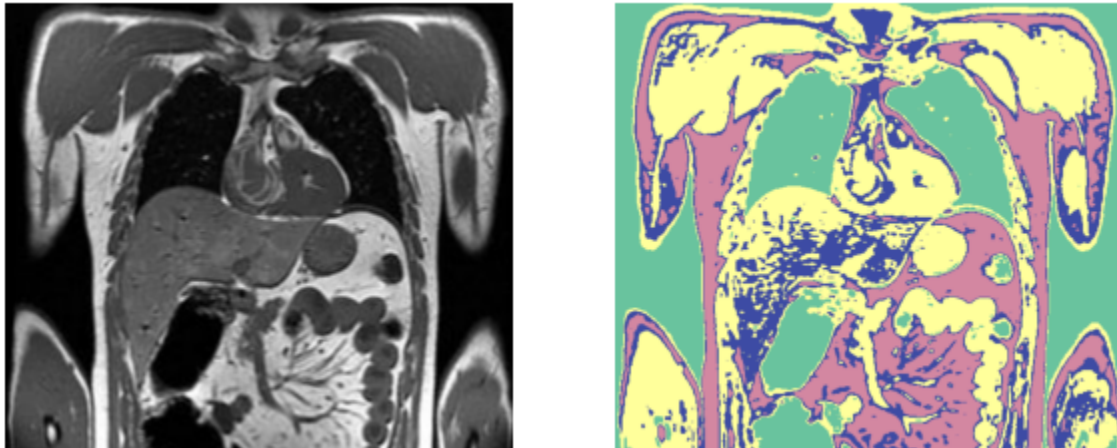
Each pixel having the same color belongs to the same cluster, i.e. all the pixels (image regions) having green color belong to the same cluster. It shows that k-means clustering is clustering together lungs with the background of the image, as in original image lungs and background both regions have similar pixel values (i.e. black).

Vinit Ramesh Gore (M20CS064), Vikanksh Nath (MT19AI024)

(ii) Left Input Image; Right Output of k-means image segmentation with k = 6
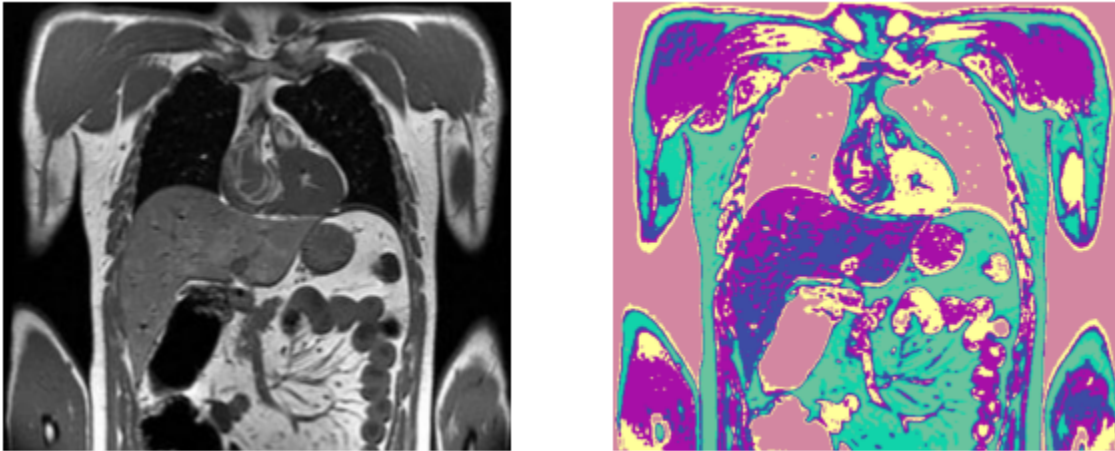
As seen from the output, k-means cluster the two different organs which is undesirable, e.g. see the violet region has a portion from the lever as well as the muscles from the arm clustered together.

(iii) Left Input Image; Right Output of k-means image segmentation with k = 8
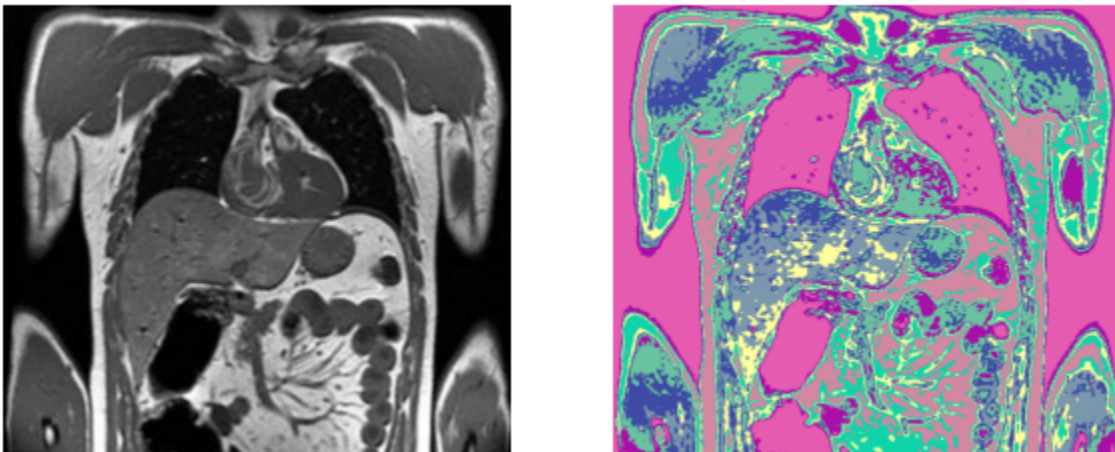
Simple Linear Iterative Clustering (SLIC)

Description

The SLIC algorithm is an improvement over KMeans algorithm. It can group visually similar pixels by minimizing the cost of computation. In the case of KMeans, the complete image area is searched for getting better results. However, in SLIC farther pixels are allotted more cost and therefore they are less considered. The algorithm is the same as KMeans. Here in Step 3, the allocation is performed using an updated criterion:

Vinit Ramesh Gore (M20CS064), Vikanksh Nath (MT19AI024)

$$\sum_{i=1}^{k} \sum_{u \in C_i} (||P(u) - P(\mu_i)||_2)^2 + \lambda(||u - \mu_i||_2)^2$$

Here *P(x)* is the image property as discussed in KMeans. *u* is the spatial location of the data point in the image. For every point *u* in cluster *Cᵢ*, search is based on the image property as well as its spatial location.

So, in SLIC the input data matrix contains two information: intensity values of pixels and spatial location i.e. coordinates of pixels weighted by a parameter λ. In our implementation we added the weighted pixel coordinates to each pixel value, resulting in a 5 dimensional vector for each pixel. Now the traditional k-means is applied on the 5 dimensional data points for each pixel in the image.

Results

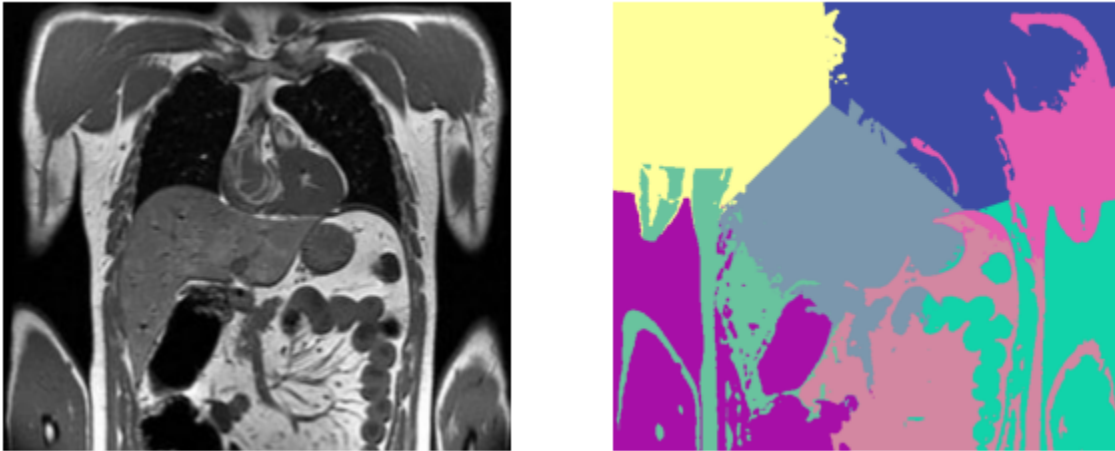(i) Left Input Image; Right Output of k-means image segmentation with k = 8, λ= 0.25
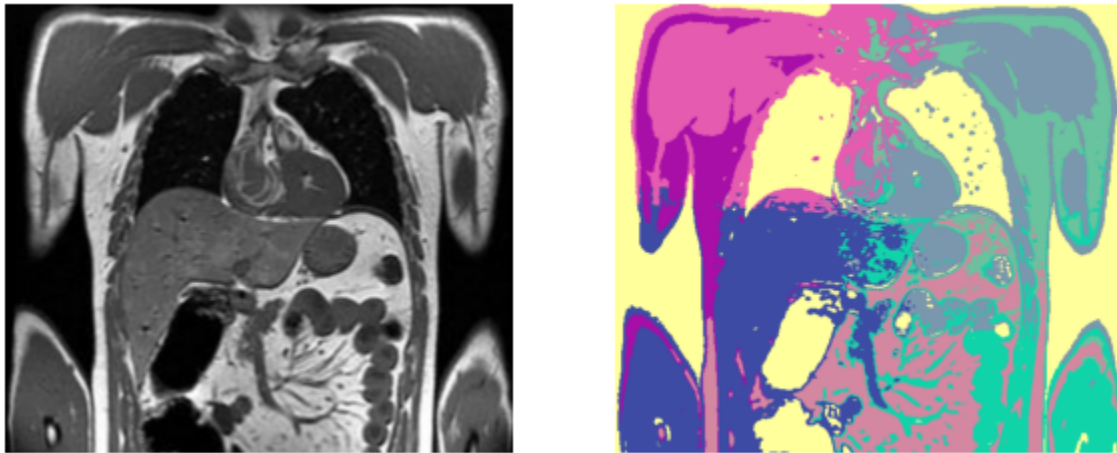
(i) Left Input Image; Right Output of k-means image segmentation with k = 8, λ= 0
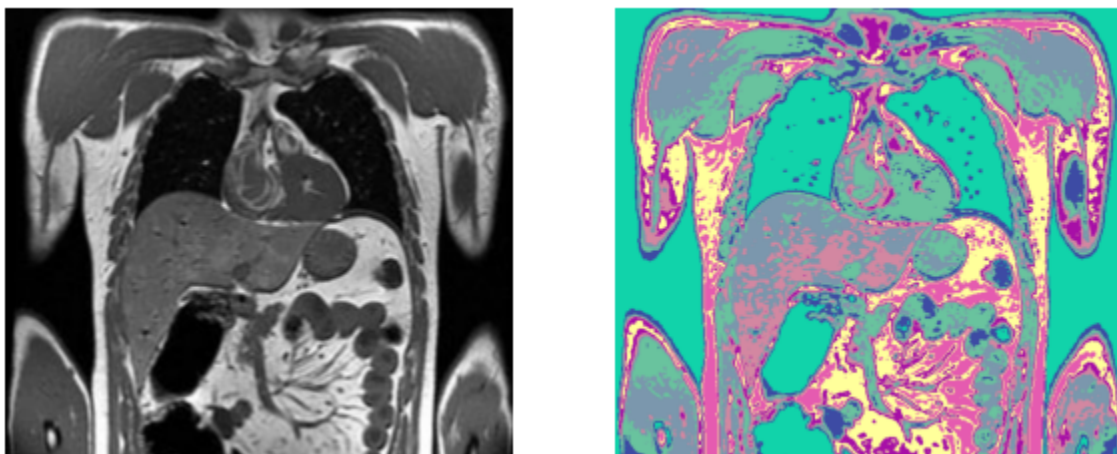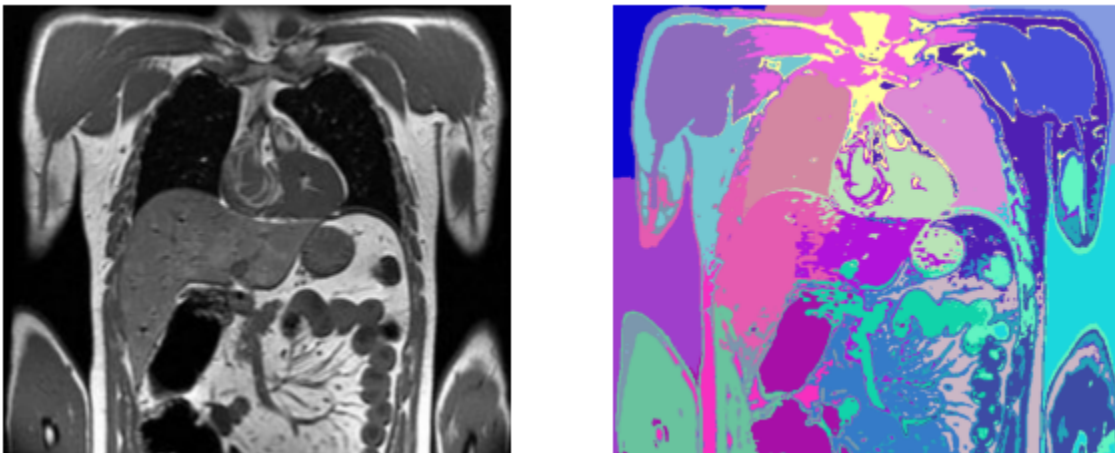

Output equivalent to KMeans result with λ= 0.

(ii) Left Input Image; Right Output of k-means image segmentation with k = 8, λ = 1

Output gives equal weightage to intensities and spatial coordinates.

(i) Left Input Image; Right Output of k-means image segmentation with k = 25, λ = 0.25

The spatial grouping in SLIC algorithm is observed in the results. The clusters are localized and do not extend beyond a certain area. The algorithm works great for images with small localized organs.

Spectral Clustering

Description

Both KMeans and SLIC form linear divisions around clusters. Spectral Clustering forms non-linear divisions. It uses graph cuts to implement the same.

Vinit Ramesh Gore (M20CS064), Vikanksh Nath (MT19AI024)

Input: Adjacency matrix W of a graph having image pixels as nodes and similarity values between pixels as edges, k: number of clusters.

Step 1: Calculate graph Laplacian matrix L = D - W where D is a diagonal matrix having degree of each node present on the diagonal.

Step 2: Find the EigenValue Decomposition of L as l (list of eigenvalues), V (matrix having columns as eigenvectors). Here eigenValues are arranged in ascending order and corresponding eigenVectors are placed in V.
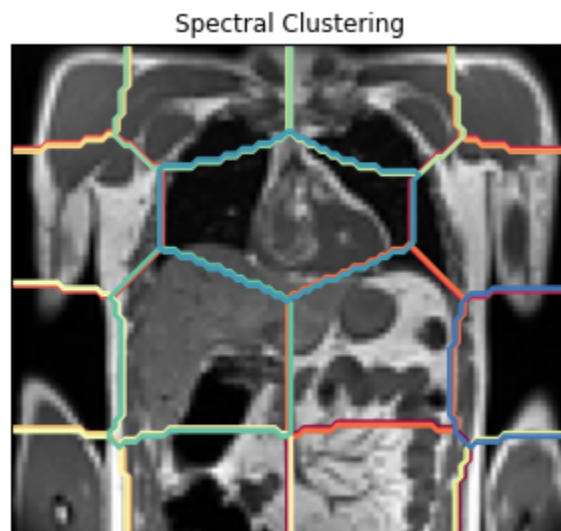
Step 3: Assign the first k columns of V to a new matrix H and apply k-means on H for k clusters. H has the shape (n,k) n: total number of pixels.

Output: The labels generated by KMeans.

Results

(a)   (b)   (c)
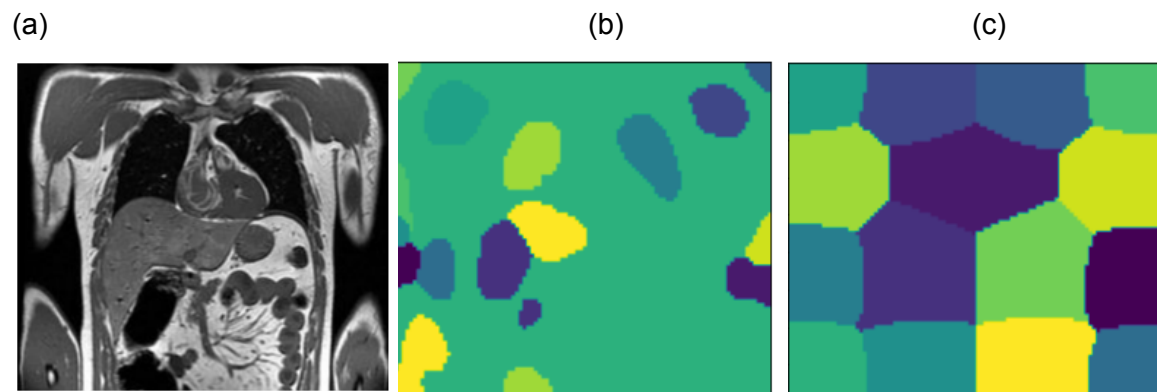
(a b c) a: Input Image, b: Output of self-built Spectral clustering algorithm where the organs were detected very grossly (eg. lungs) c: Output of in-built Spectral clustering algorithm.

Overlapping of obtained clusters with input image