Images Segmentation

Mohamed Metwalli, Mariam Saeed, Nada Hassan

1. Downloading the data

First, we've downloaded the dataset then uploaded it to a drive folder to ease the access during tuning and testing.

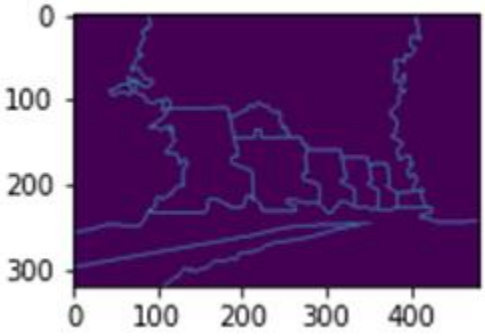
2. Visualizing the data

The function that visualizes the images take two parameters:

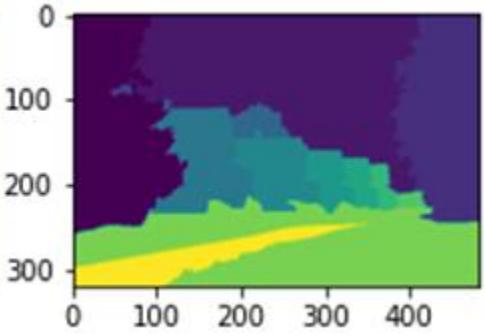
- Name of the image
- Category of the image (train - test)

Then it shows the available ground-truths in two formats as segment map and edges map

Example



Edges



segments

3. Segmentation using K-means

First, we ran our own implementation of K-means on 200 images from the training set to tune the parameter K between {3,5,7,9,11} and we evaluated the segmentation results for every M available ground-truth for image I using f-measure and conditional entropy that we implemented as well.
We found that K=3 is the average K among all images producing acceptable measures.
Then, we used K=3 on 50 images of the testing set to evaluate the performance.

Average of 200 images of train set:

k	3	5	7	9	11
f_Score	0.5320985	0.4526	0.3920	0.34791	0.31384697
con_entropy	0.5507327	0.4996	0.4734	0.45371	0.43856059

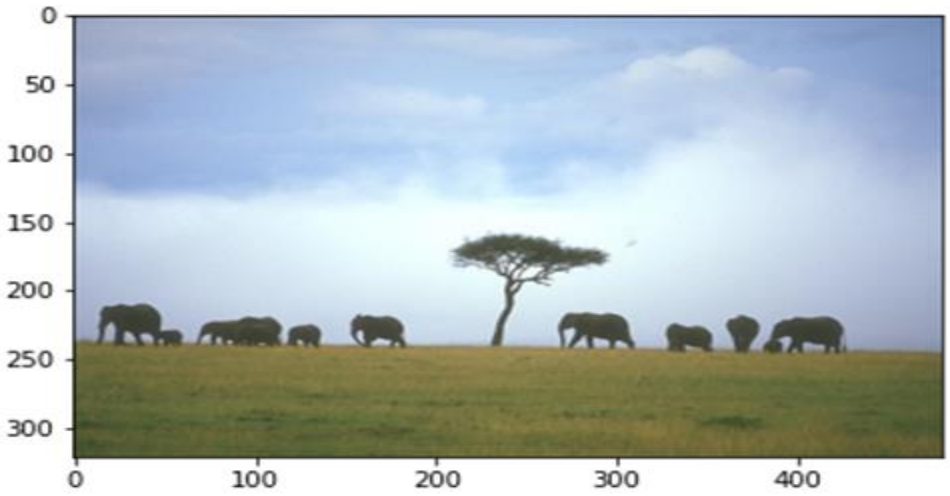
Average of 50 images of test set:

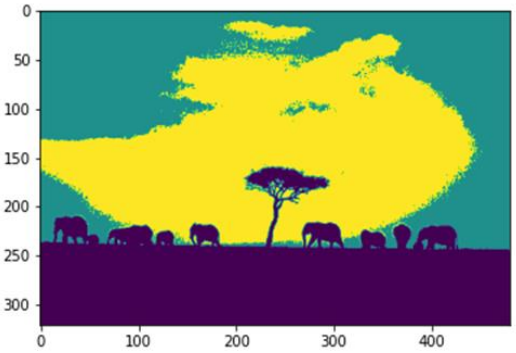
F_Score	Entropy
0.4993944032424352	0.6062242045276335

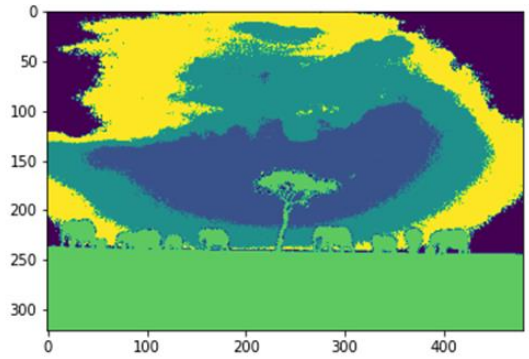
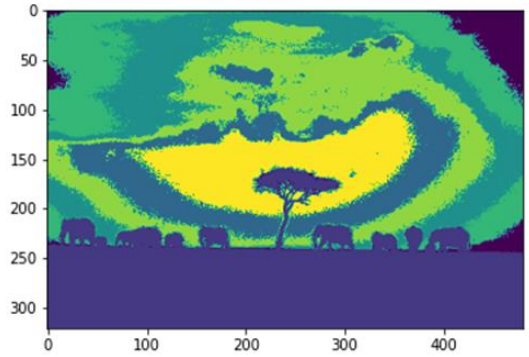
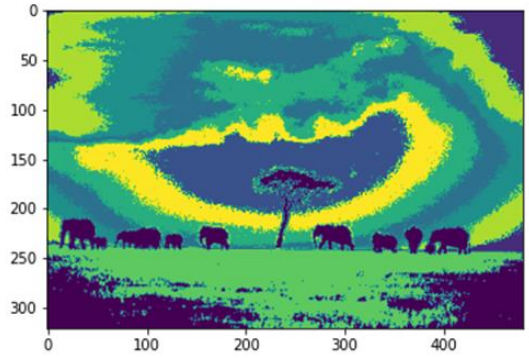
Success vs Failures in train set

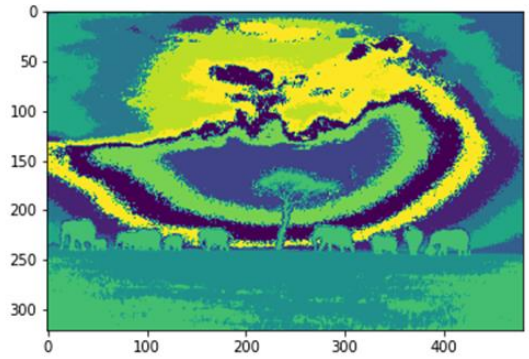
Success cases

Original image

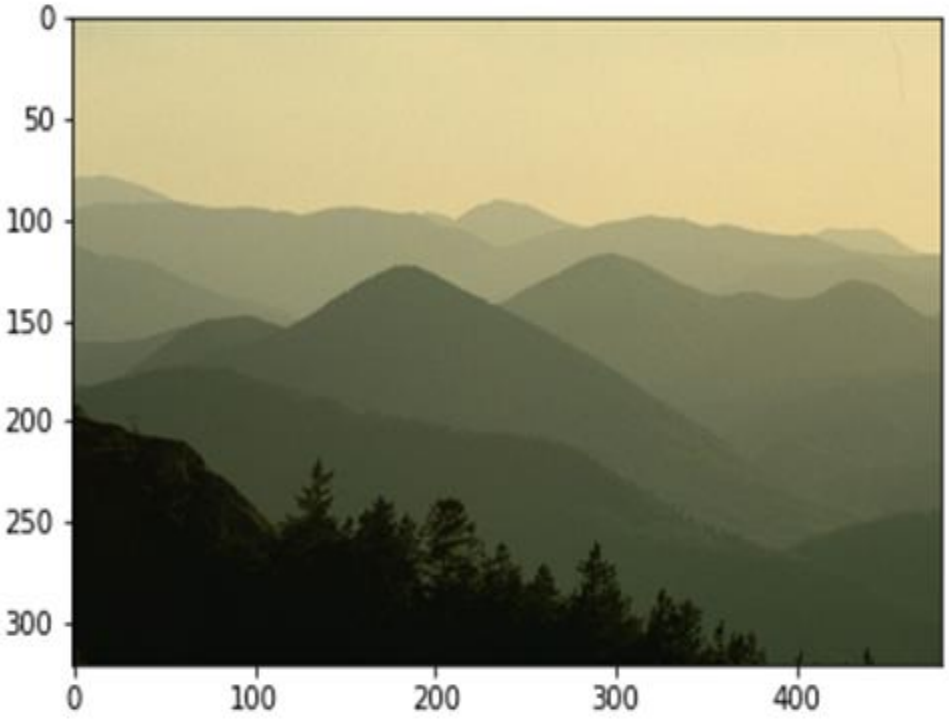


k	image	f_score	Entropy
3	 Picture is similar to ground truth	0.75434	0.1545395

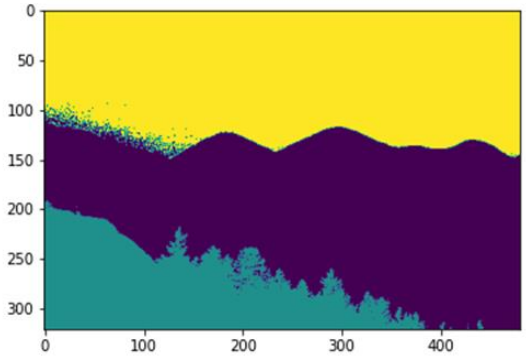
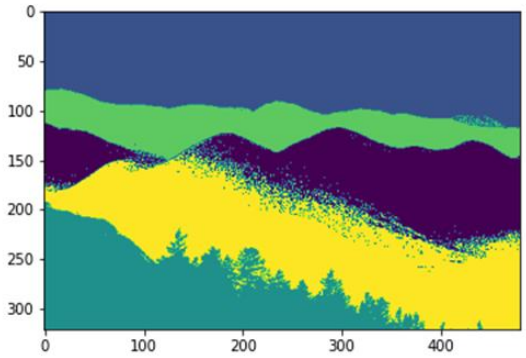
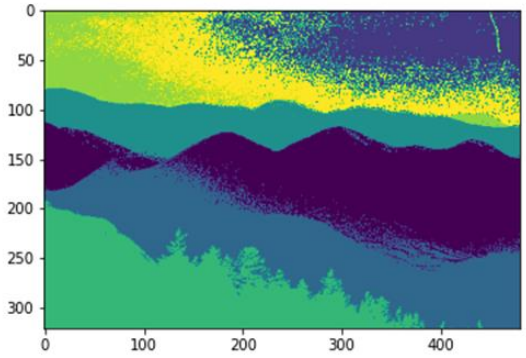
5		**0.499608**	**0.11946**
7		**0.37157**	**0.1475**
9		**0.34092**	**0.1376**

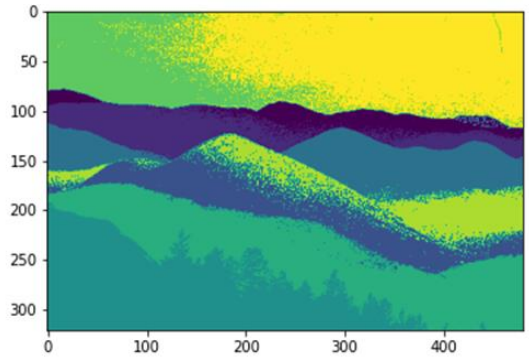
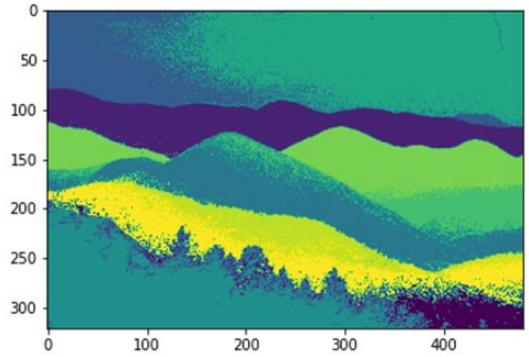
11		**0.28817**	**0.1326**

Original image



k	image	f_score	Entropy

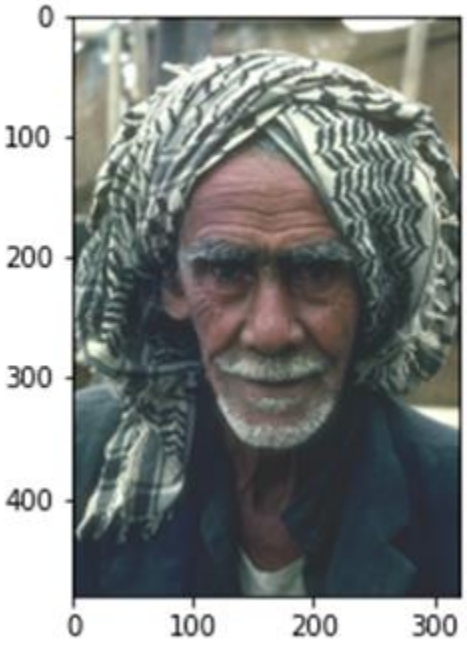
3	 Picture is similar to groundtruth	**0.81844**	**0.3496**
5		**0.832**	**0.1876**
7		**0.6675**	**0.178**

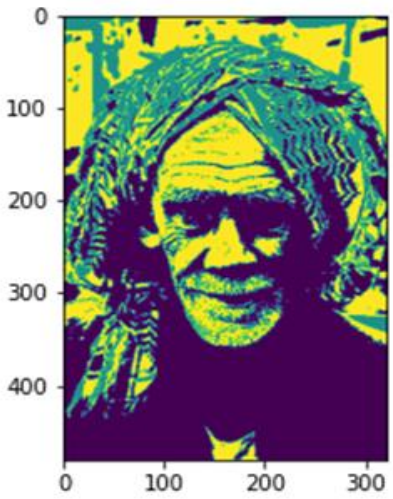
9		**0.658**	**0.1136**
11		**0.568**	**0.1123**

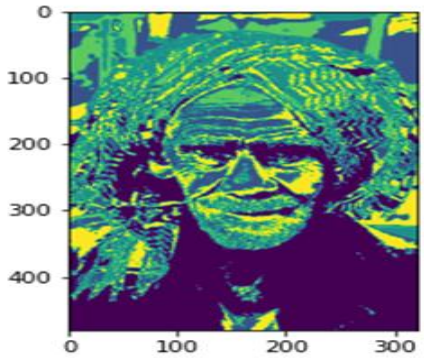
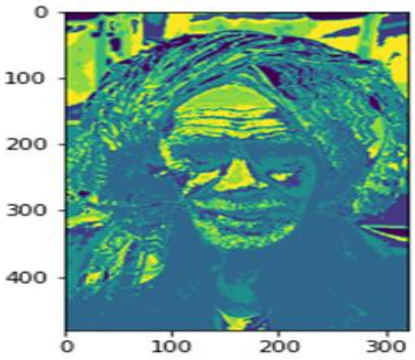
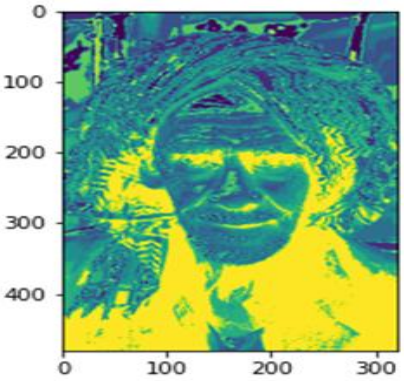
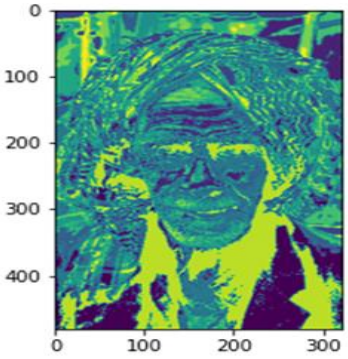
- When the image has few details, the clustering result matches the ground truth segments
- When we increase k and there's a few colors and shades in the image (less than k) the clustering mis matches the groundtruth

failure cases

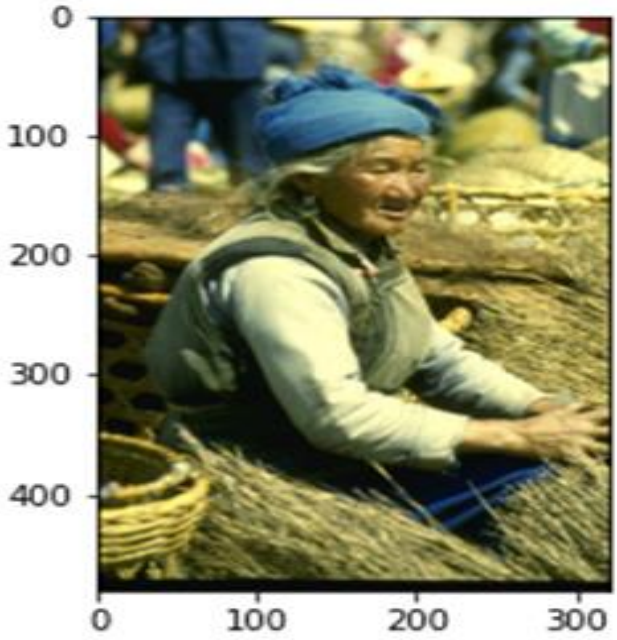
Original image:



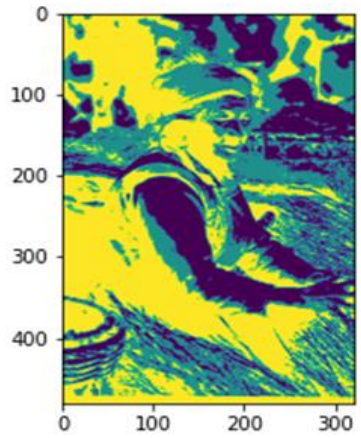
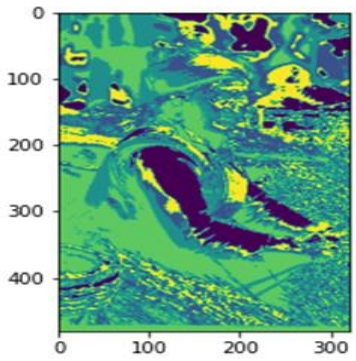
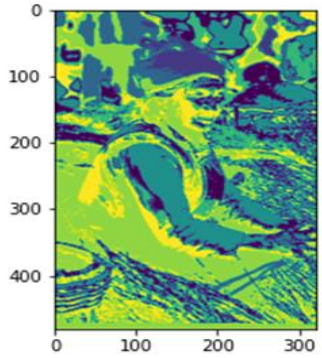
k	image	f_score	Entropy
3		0.35316	0.73001

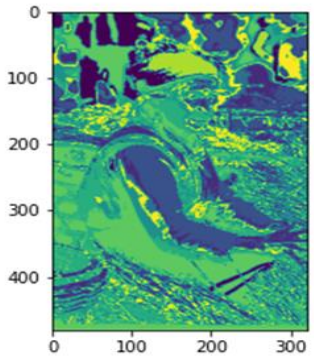
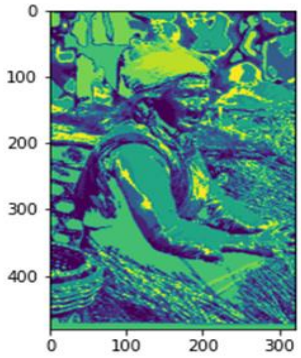
5		**0.30482**	**0.64041**
7		**0.27184**	**0.62241**
9		**0.23538**	**0.61784**
11		**0.23094**	**0.59487**

Original image:



k	image	f_score	Entropy

3		**0.32637**	**1.0142**
5		**0.27918**	**0.98883**
7		**0.3278**	**0.92757**

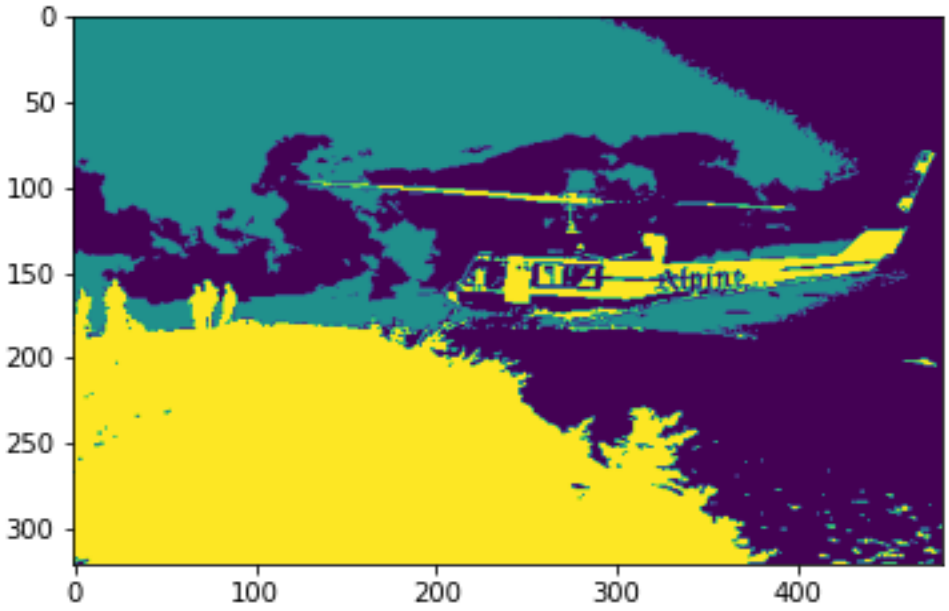
9		**0.2907**	**0.91967**
11		**0.26802**	**0.89448**

Bad results produce less f-score although it shows more details but as these details are not shown in the ground-truth so the pixels are mismatched. F-score decreases as K increases as it shows more details.

Success vs Failures in test set :

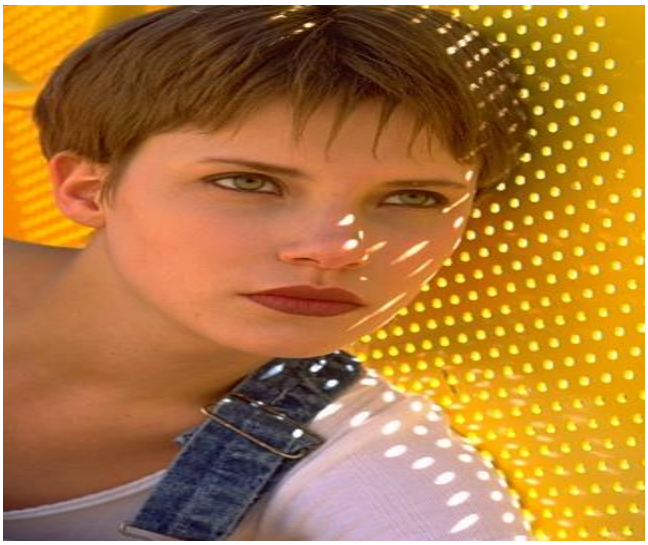
Success

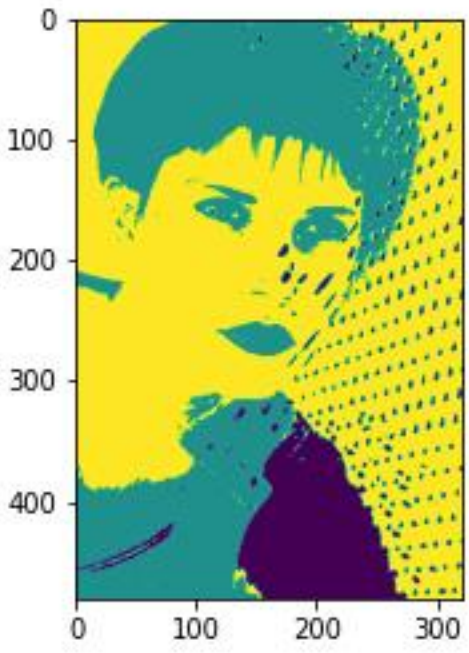
Original image:



Average fscore: 0.717742984794211 conditional_entropy: 0.435613431357418

Original image:

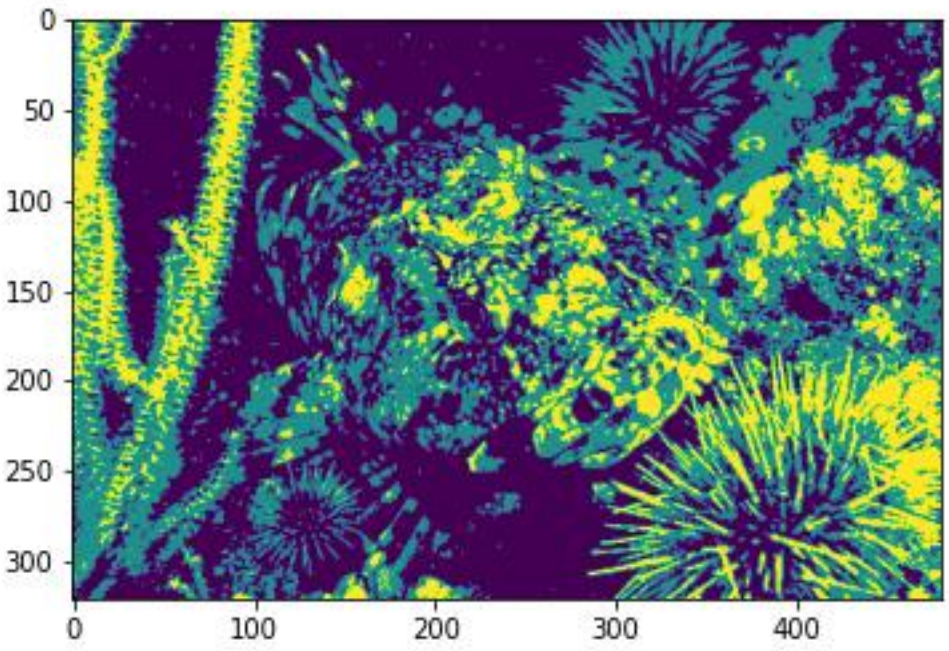




Average fscore: 0.6548447753658493 conditional_entropy: 0.5115402955520272

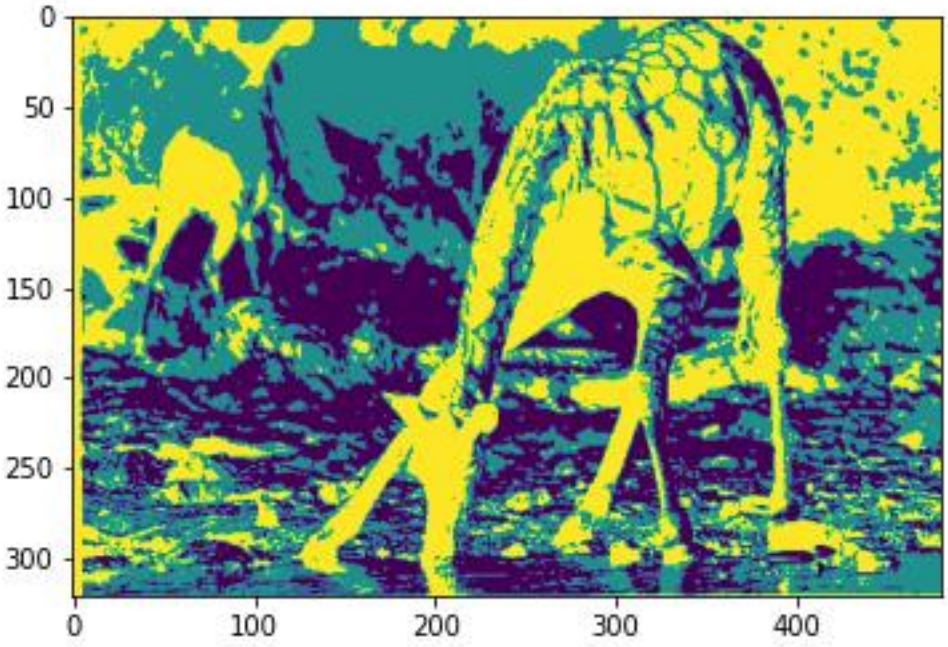
Failures:

Original image:



Average fscore: 0.32404107526209136 conditional_entropy: 0.8072440733834156

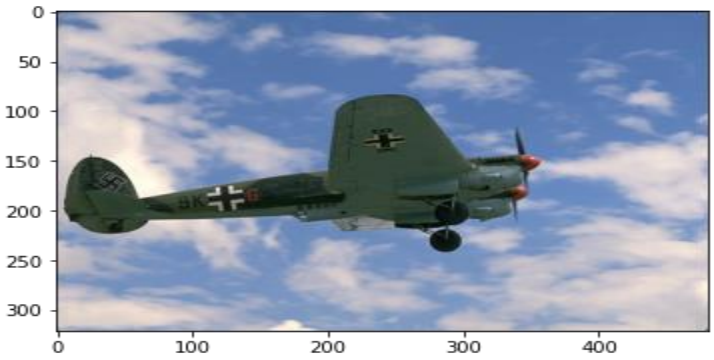
Original image:

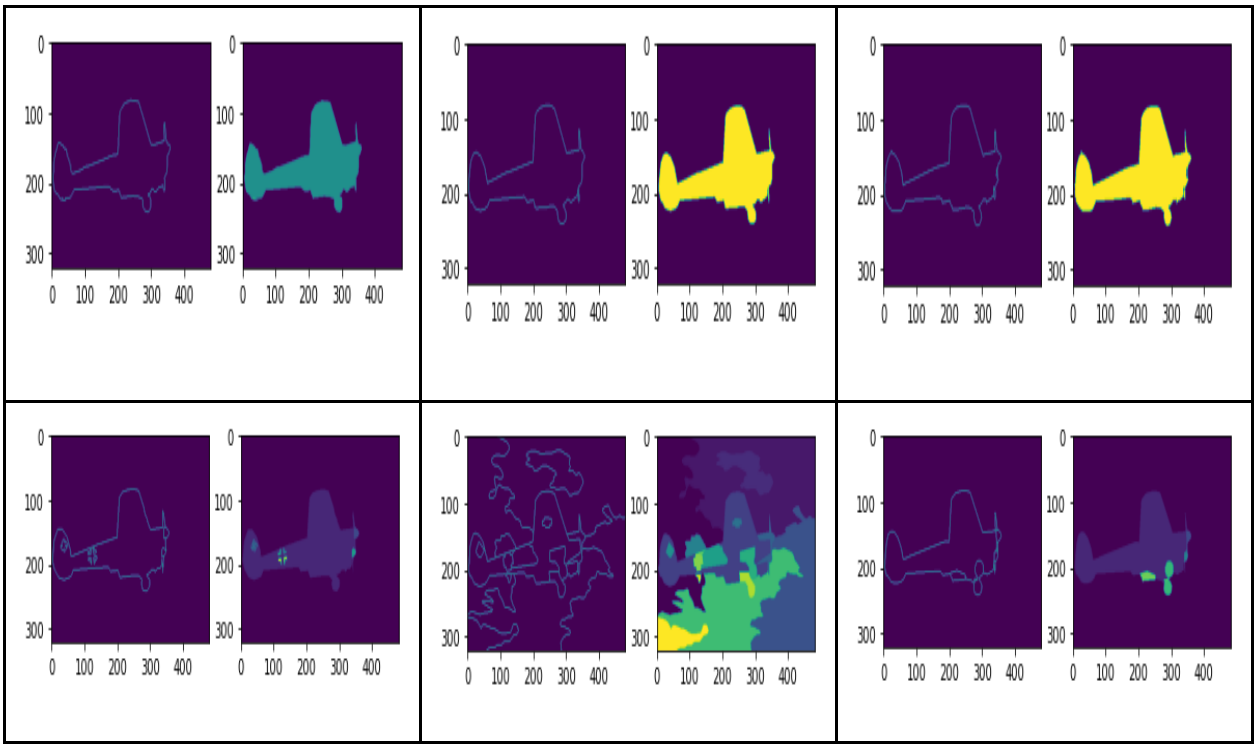


Average fscore: 0.3699420133273374 conditional_entropy: 0.8516124277308098

4. Big picture

Image 3036 :





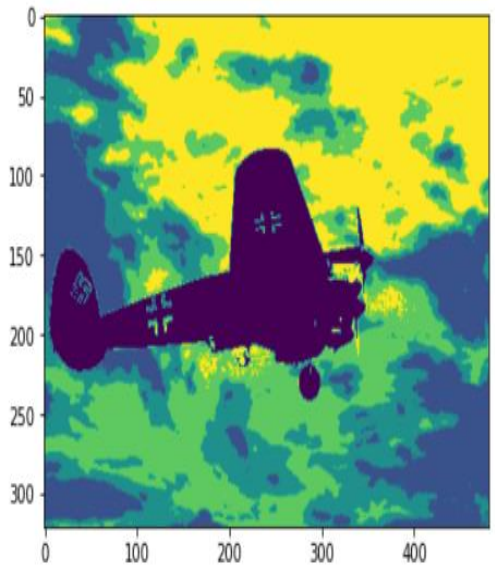
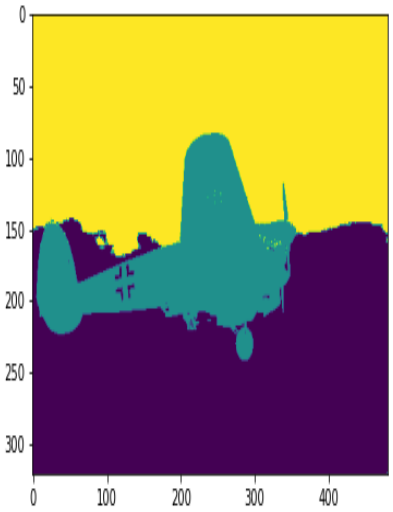
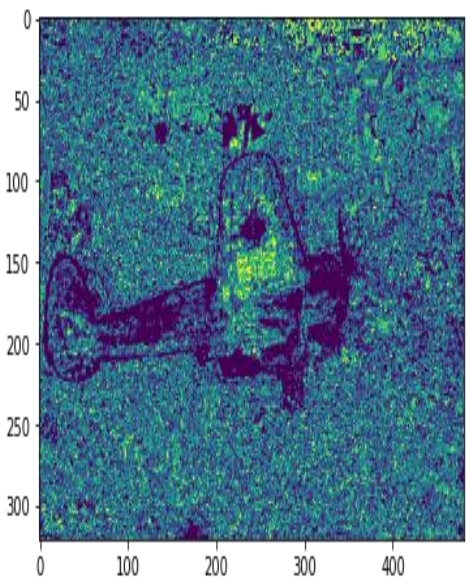
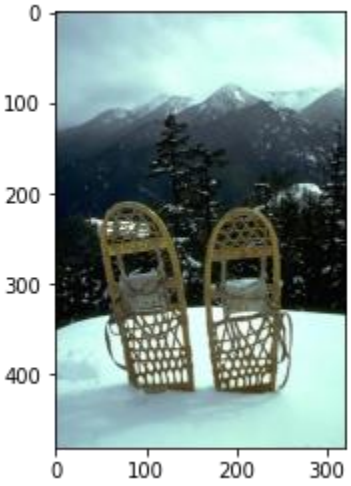
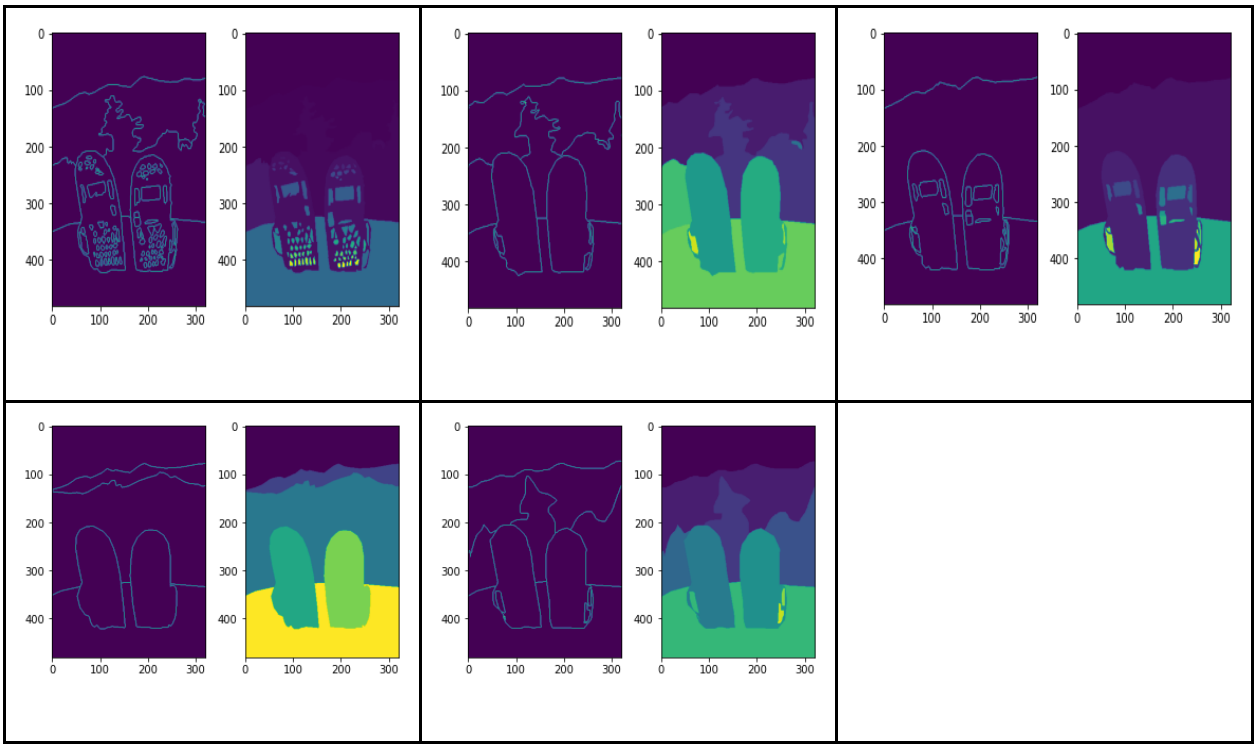
k-means	Spatial layout	Normalized cut
		
Average F-measure : 0.5280139532293684 Average entropy : 0.11225977295102128	Average fscore: 0.7571137485935723 conditional_entropy: 0.10741786308526106	Average F-measure : 0.2968065822313833 Average entropy : 0.2971368953481392

Image 2018 :





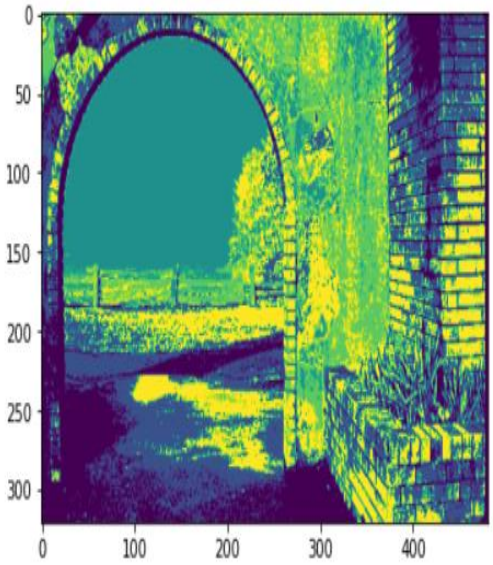
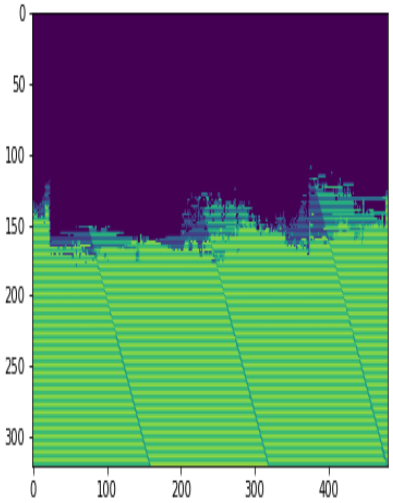
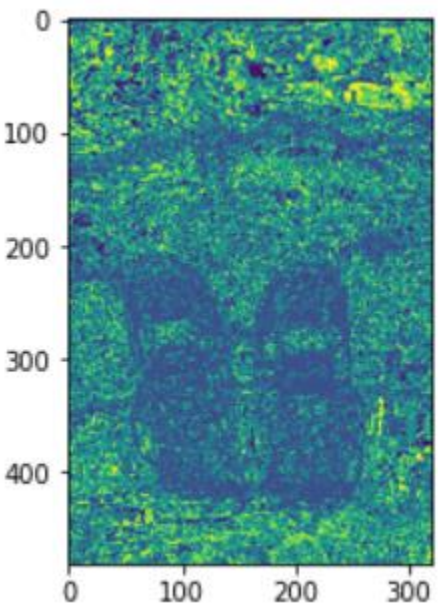
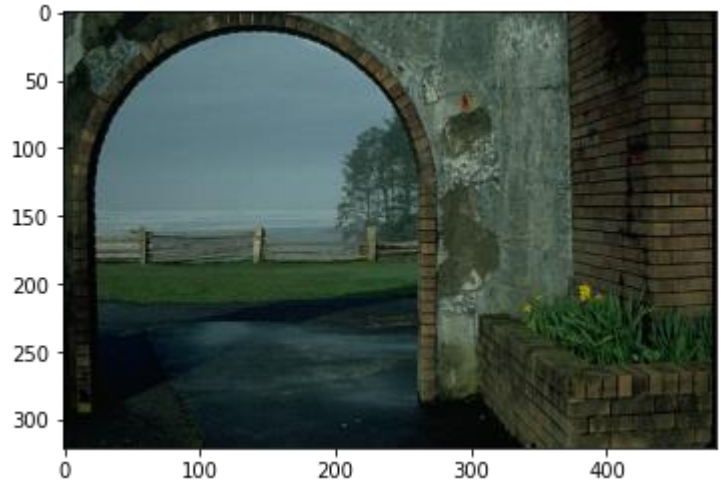
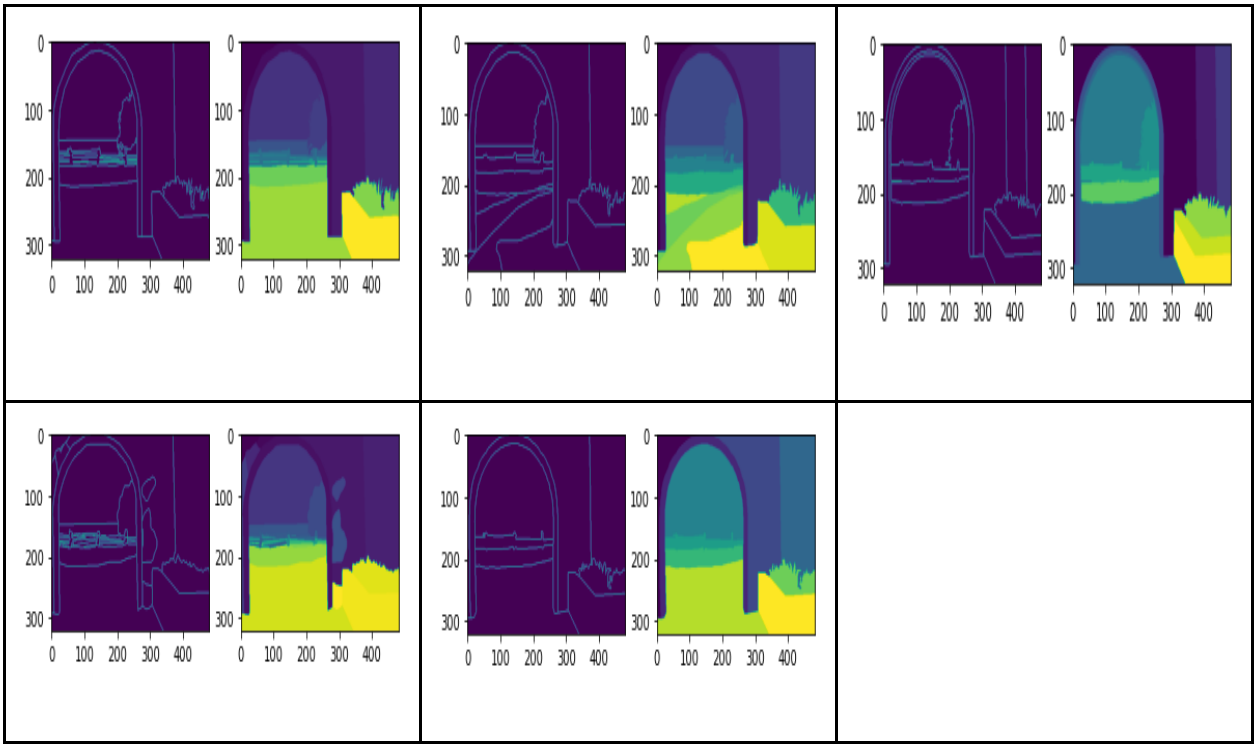
k-means	Spatial layout	Normalized cut
		
Average F-measure : 0.4999559861917925 Average entropy : 0.7210260105041936	Average fscore: 0.41369176639882516 conditional_entropy: 0.8347727414910888	Average F-measure : 0.2916552763745857 Average entropy : 0.30299593211363557

Image 5096 :





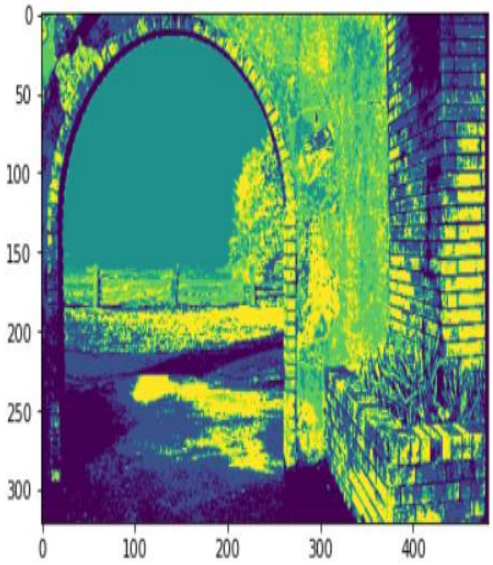
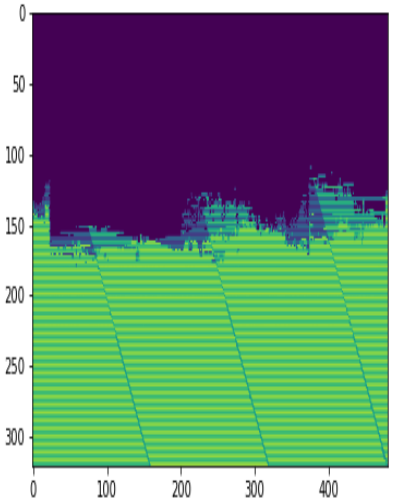
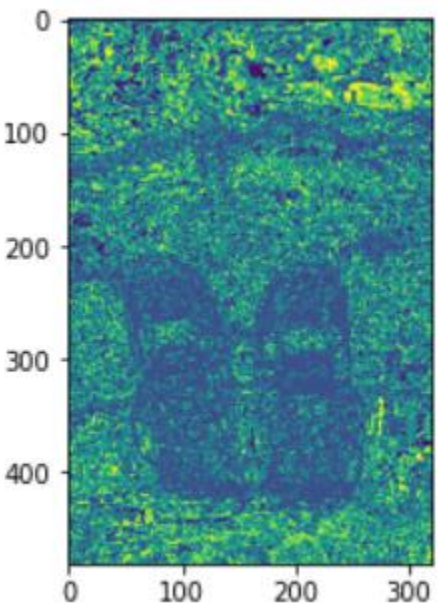
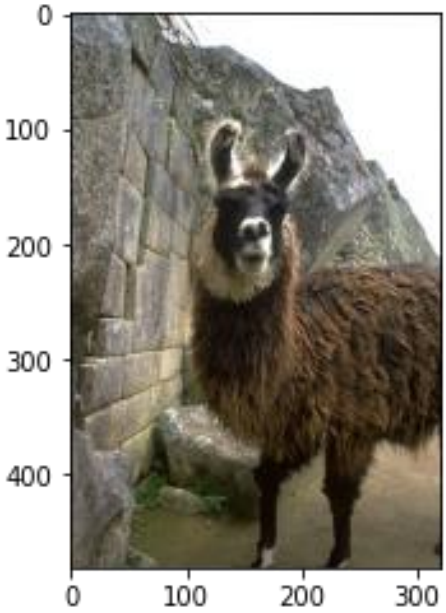
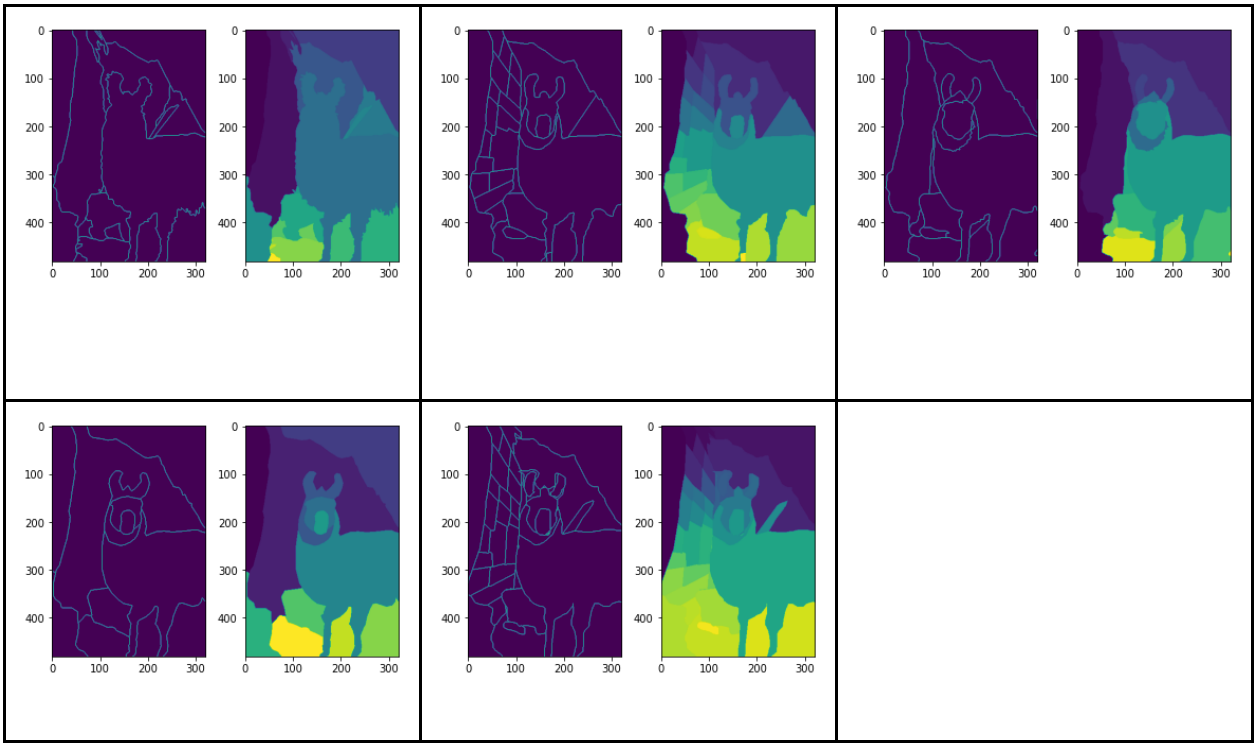
k-means	Spatial layout	Normalized cut
		
Average F-measure : 0.4999559861917925 Average entropy : 0.7210260105041936	Average fscore: 0.41369176639882516 conditional_entropy: 0.8347727414910888	Average F-measure : 0.2916552763745857 Average entropy : 0.30299593211363557

Image 6046 :





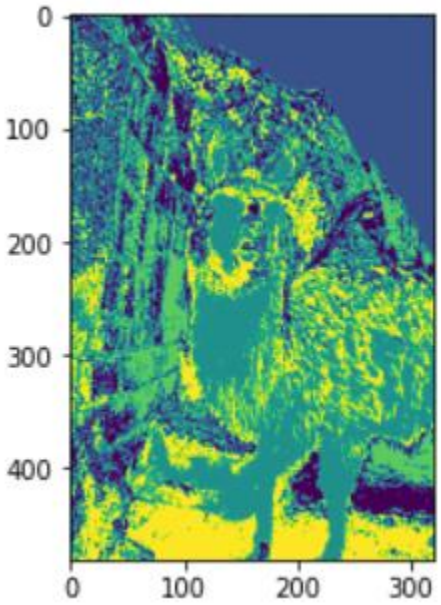
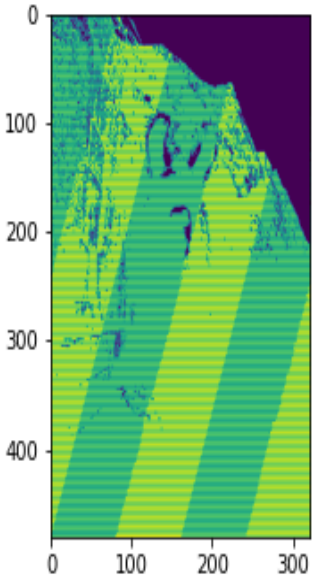
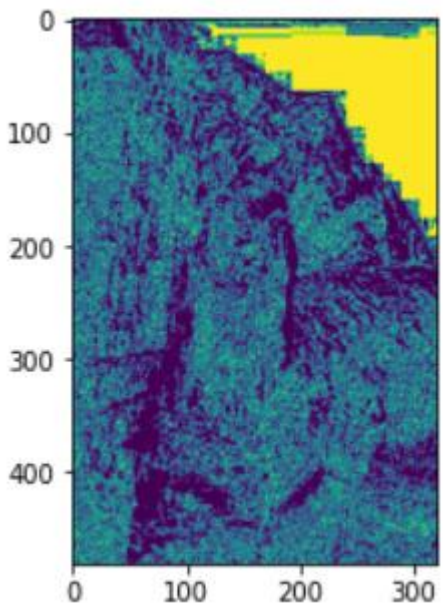
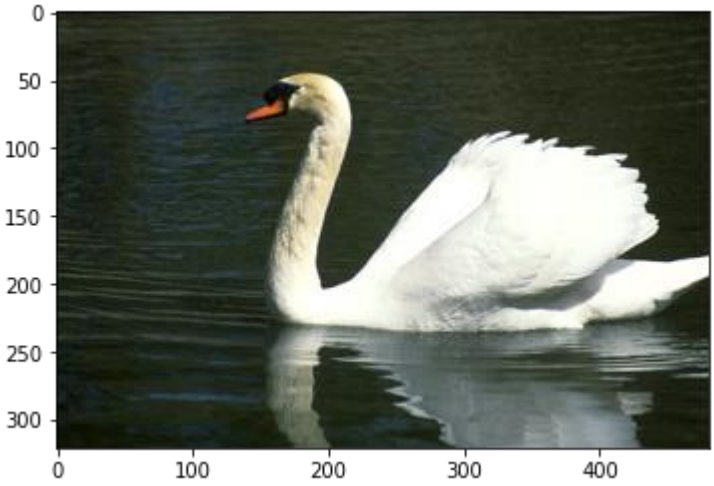
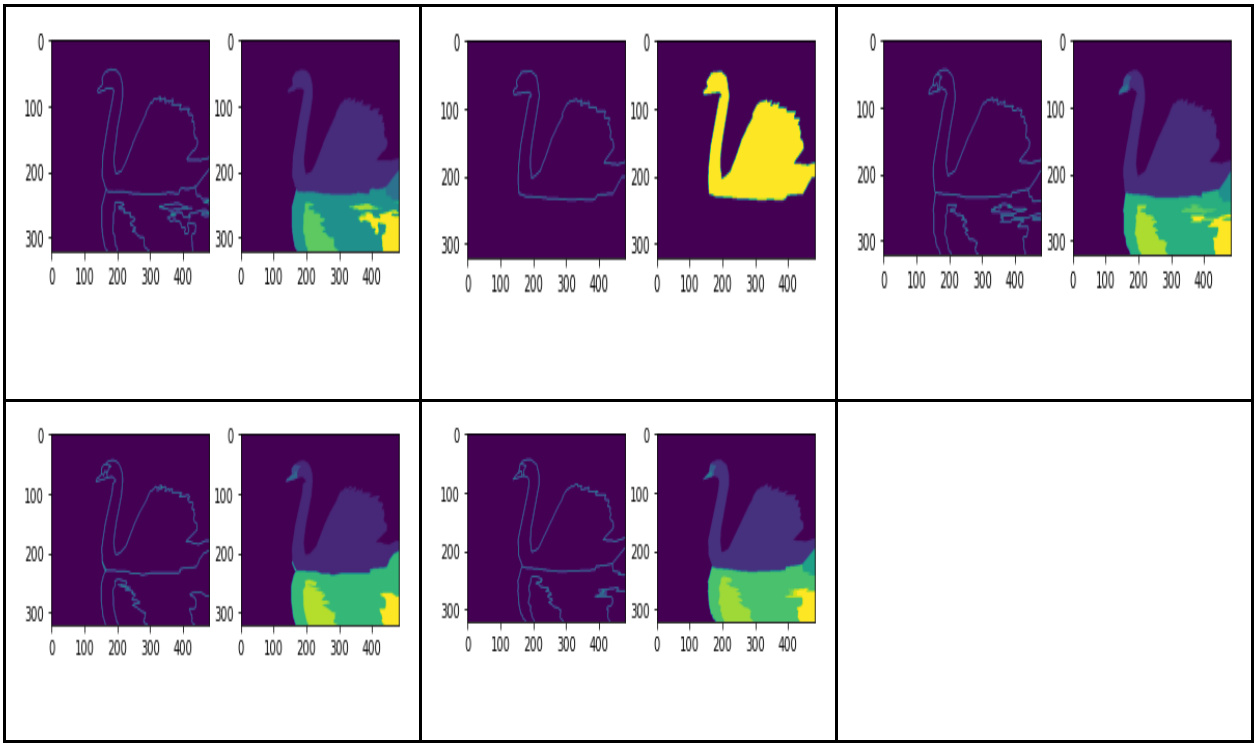
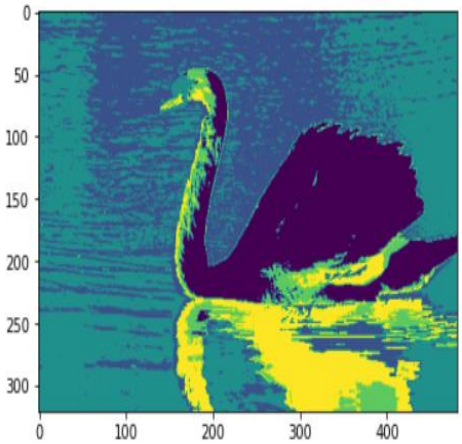
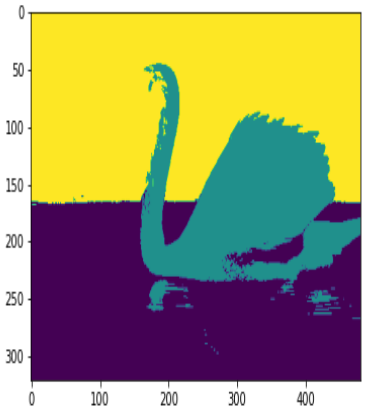
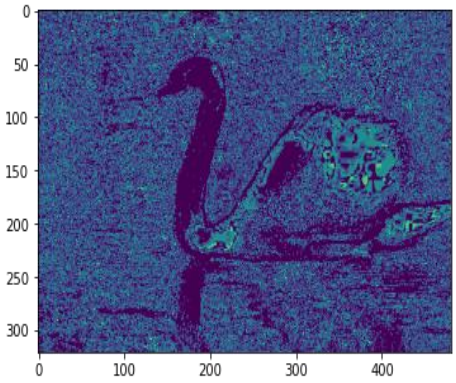
k-means	Spatial layout	Normalized cut
		
Average F-measure : 0.5143757188104001 Average entropy : 0.7765816442616964	Average fscore: 0.5323195681807157 conditional_entropy: 0.9126612337133657	Average F-measure : 0.35793727000133196 Average entropy : 0.9449552940322565

Image 8068 :





k-means	Spatial layout	Normalized cut
		
Average F-measure : 0.5860721264508238 Average entropy : 0.18679724156168204	Average F-measure : 0.707828782675141 Average entropy : 0.21553555891904114	Average F-measure : 0.26428553749368944 Average entropy : 0.4348731322841834

5. Extra

a. Modifying dataset

- To modify the feature vector of the previous runs, we've added two features x and y that represent the position of each pixel.

- Now the data sets contains 321x481(width x height) rows and 6 columns (R G B X Y)

b. Contrasting results

- Next, we ran the tuning again on 200 images of the dataset to observe the difference between the two approaches.
- Adding the new features reduced the overfitting of clustering in the previous part but that made the pictures became unclear as objects of pictures are non-convex shapes so k-means fail to segment it
- We found that k = is the average k between all images in the dataset
- We used that k on the first 50 images of the test set

Average of 200 images of train set:

k	3	5	7	9	11
f_Score	0.515708	0.4284	0.3765	0.33648	0.3034156

con_entropy	0.55007	0.5008	0.4647	0.43870	0.4221187

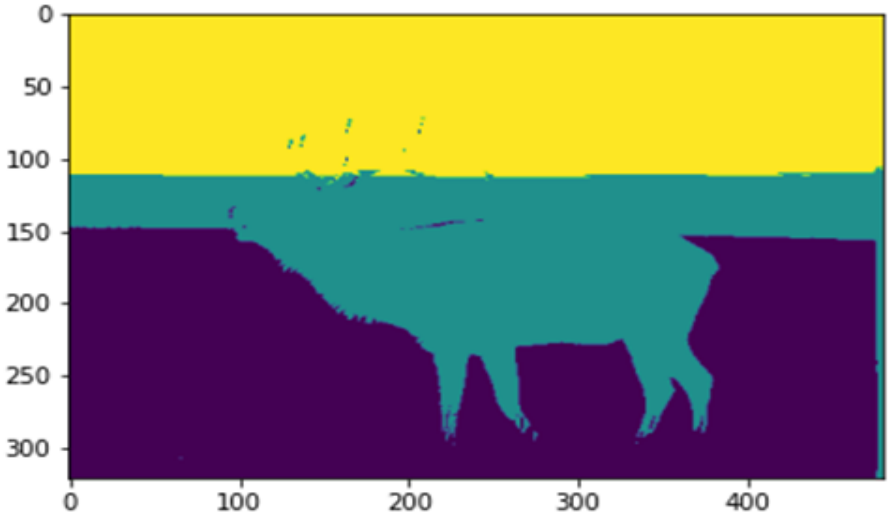
Average of 50 images of test set on k = 3:

f_score	con_entropy
0.49779	0.59976

Success vs Failures in train set when k = 3 :

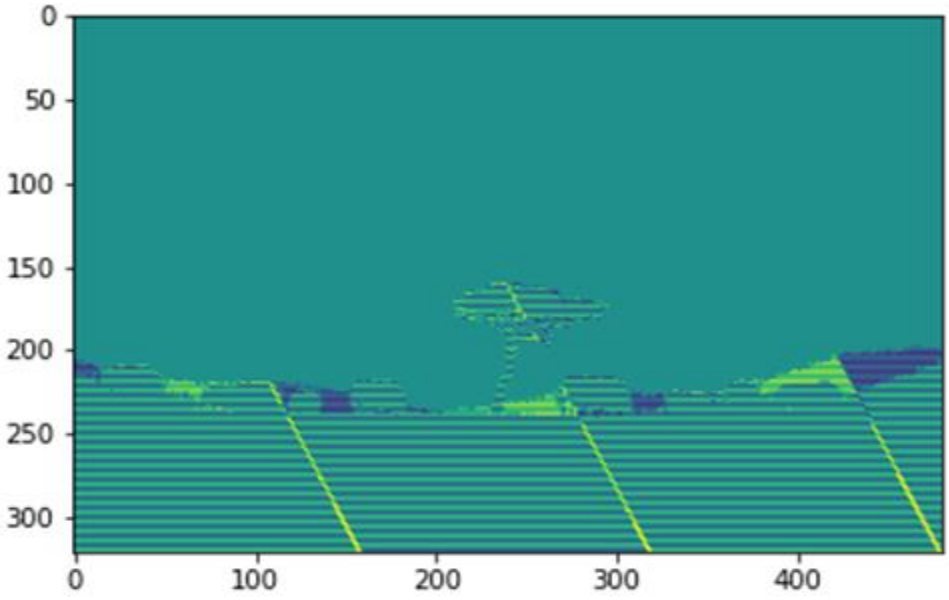
Success

Original image:



Average fscore: 0.8284478025116421 conditional_entropy: 0.18147965853216558

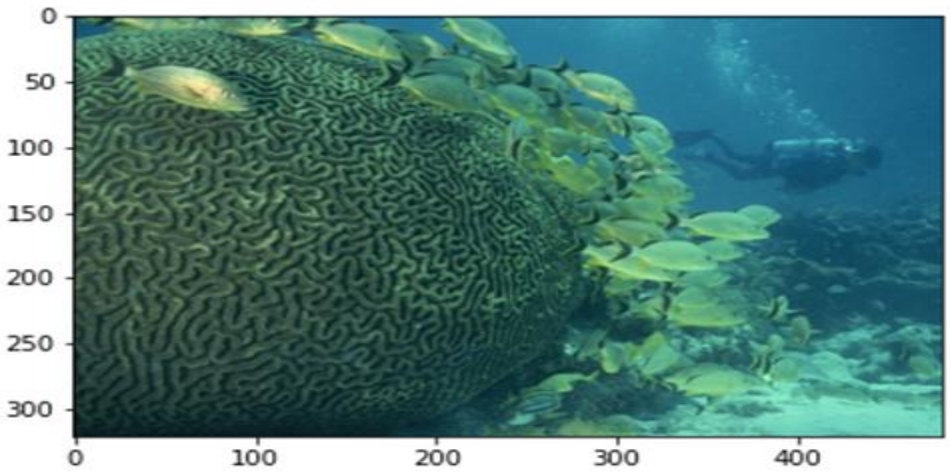
Original image

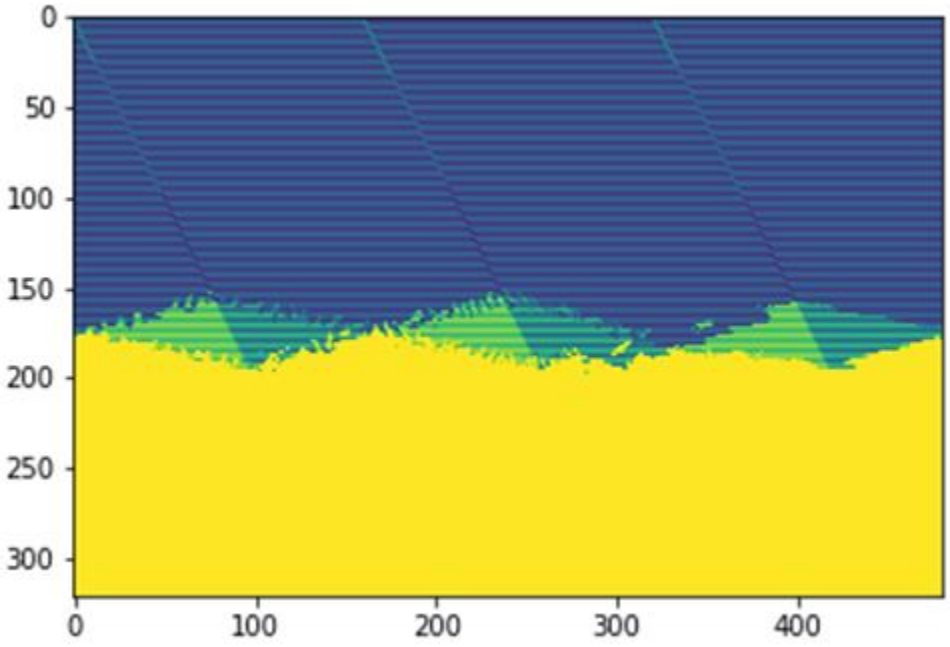


Average fscore: 0.7174326553490545 conditional_entropy: 0.18831131721885586

Failures

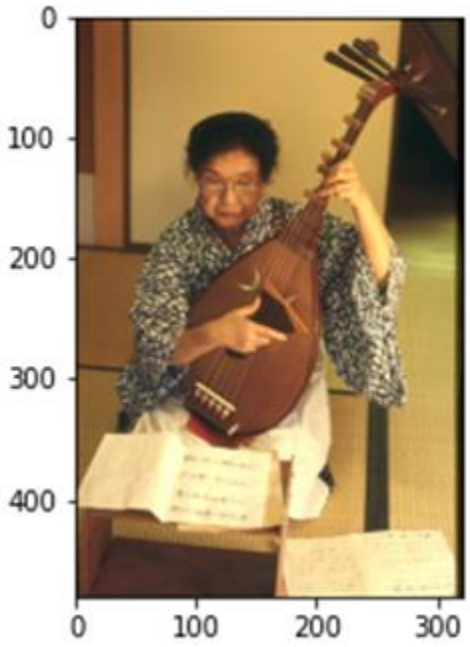
Original image:

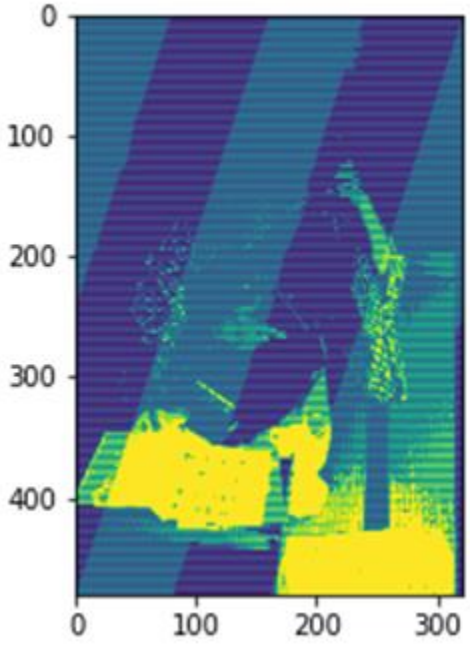




Average fscore: 0.3781043622079655 conditional_entropy: 0.600936571684321

Original image:





Average fscore: 0.33549314737219327 conditional_entropy: 1.0947403121001686