

HW #14

2017029970

우원진

사용한 Library

```
import cv2
import numpy as np
from sklearn.cluster import MeanShift, KMeans
from sklearn.datasets import make_blobs
import csv
```

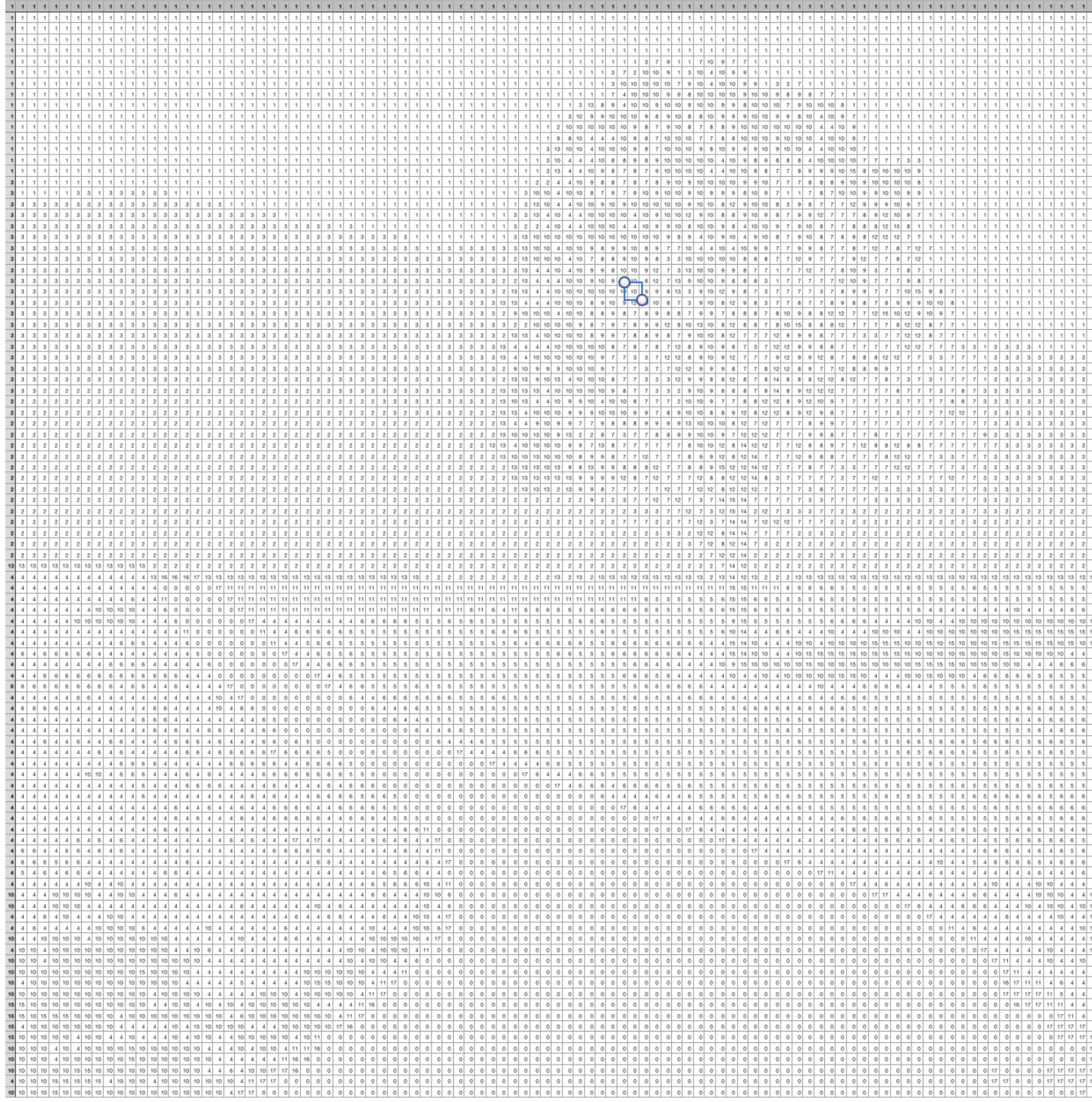
1. Mean Shift Cluster

Mean Shift Cluster를 구현하기 위해 sklearn 라이브러리의 MeanShift를 사용하여 모델링 했습니다.

처음 사진의 크기가 350 x 500 이었는데 이를 Clustering 하기 위해서는 엄청난 시간이 소요되어 100x100 Image로 변경했습니다. 이를 LAB로 변경했습니다.

또 100 x 100 으로 했어도 코드를 실행할때마다 모델링을 하는데에는 무리가 있어서, 모델링을 한번 하여 그 모델링의 결과로 나온 label의 값을 index 순서대로 csv 파일에 담았습니다.

Mean Shift Cluster를
사용해 군집화 되어
Labeling 되었기 때문에
index 별로의 Label 값이
저장 되어있습니다.
크게 보니까 그림이
되어있네요...
자세히 보면 모두
숫자입니다!!(Label)
3개의 사진 모두 이
방법을 사용했습니다.



BandWidth(h)

처음에 h의 크기를 어느 정도로 잡아야 할지 감이 안와서 먼저 2로 잡았습니다.

그랬더니 몇 천개의 Label로 구분 되었습니다.

그로 인해 h의 크기를 많이 크게하여 10으로 했더니 약 40개 정도의 label이 나왔습니다.

이로인해 h를 다시 15로 바꿨고 label 은 18개의 cluster가 나왔습니다.

이 방법으로 2번째 사진은 h를 15로 하여 16개의 cluster를 만들었고, 3번째 사진은 h를 10으로 하여 20개의 cluster를 만들었습니다.

```
clustering = MeanShift(bandwidth=15).fit(X) # label 18개
```

2. K-Means Clustering

K-Mean Cluster를 구현하기 위해 sklearn 라이브러리의 KMeans 를 사용하여 모델링 했습니다.

사진은 위에서 만든 사진을 사용했습니다.

Mean Shift Clustering에서 18개의 군집으로 만들었기 때문에 K-Means에서 군집의 갯수를 18개로 설정하여 모델링 했습니다.

2번째 사진은 16개, 3번째 사진은 20개로 모델링 했습니다.

```
kmeans_model = KMeans(n_clusters=18)
```


K-Means Clustering도
코드 실행시마다
모델링을 하기에는
부담이 크기 때문에
csv파일에 각각
index마다의 label을
저장했습니다.
(근데 생각보다 K-Means
Clustering은 속도가
빨랐습니다.)

결과 : 모든 결과는 LAB에서 다시 RGB로 변환하여 표현했습니다.

Original Image



Mean-Shift Clustered Image



K-Means Clustered Image



결과 2

Original Image



Mean-Shift Clustered Image



K-Means Clustered Image



결과 3

Original Image



Mean-Shift Clustered Image



K-Means Clustered Image



3. Comparing

Mean Shift Clustering은 군집에서 가장확률적으로 높은 값이 대표값이 되고, K-Means Clustering은 군집의 평균값이 대표값이 된다.

Mean Shift Clustering으로 만든 Image는 색 표현을 할 때 대표되는 색상 들로만 표현했기 때문에 당연히 색상이 단순해 졌지만 discontinuity 한 부분들은 대부분 잘 유지가 된다.

K-Means Clustering으로 만든 Image도 색 표현을 할 때 대표되는 색 들로만 표현했기때문에 역시 색상이 단순해지고, discontinuity한 부분들이 약간 뭉개지는 느낌 이있다.

Mean Shift Clustering이 조금 더 original에 가까워서 좋다고 생각되지만 모델링에 오래 걸리고, K-Means Clustering은 모델링 속도가 빠르기 때문에 예민한 결과값을 바랄때는 Mean Shift Clustering을, 속도가 중요할 때는 K-Means Clustering을 사용하는것이 좋다고 생각된다.