

멀티미디어설계 HW5 보고서

19010692 이준호

보고서 순서

1. Define 정의 + 상대경로 설정
2. 환경 변수에 따른 결과
 - 클러스터 개수
 - 차원 (Dimension) - 4, 16, 64
3. 최종 평가

1. Define 정의 + 상대 경로

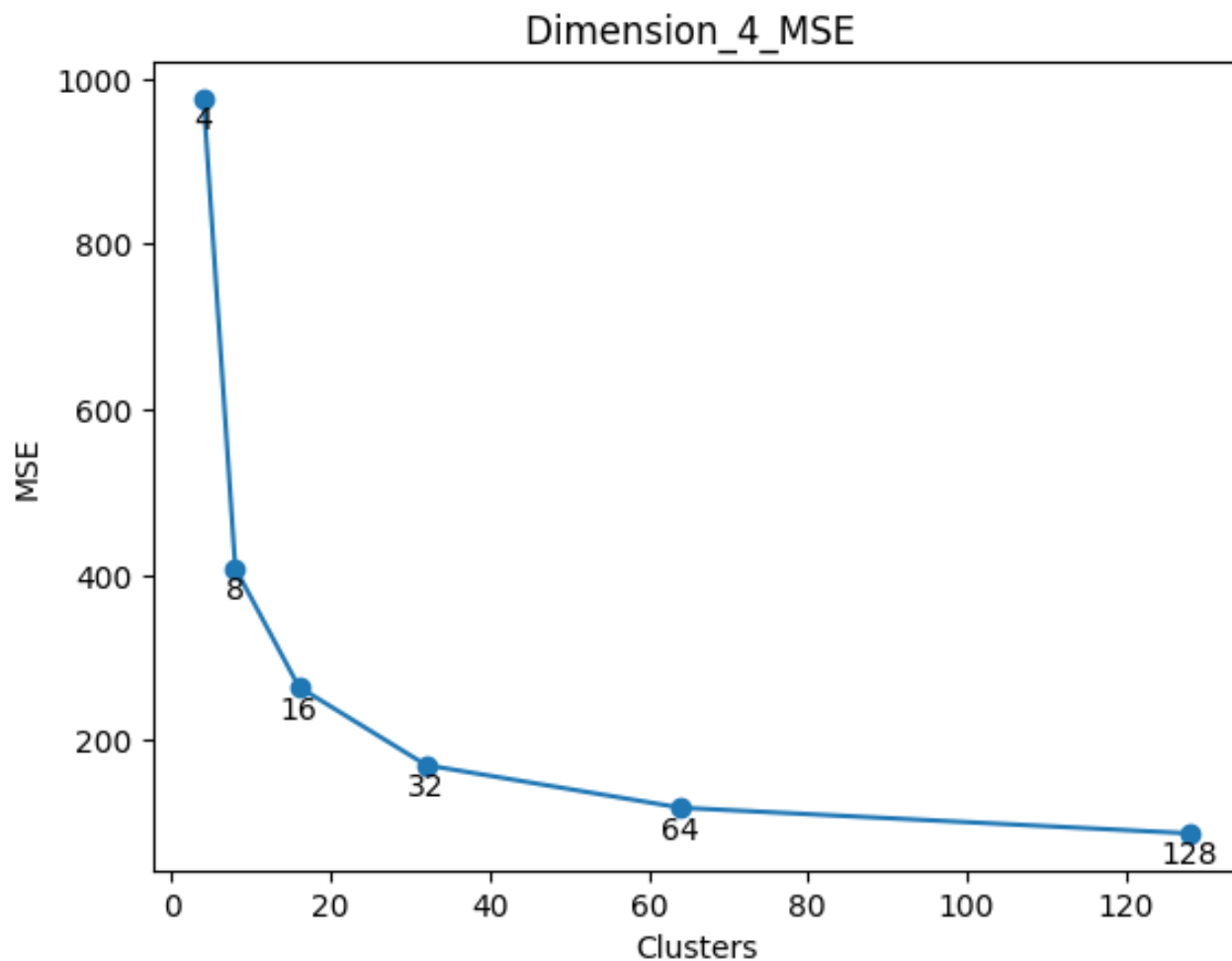
```
from google.colab import drive
drive.mount('/content/drive')
PATH = '/content/drive/MyDrive/Colab Notebooks/test_image/'
```

구글 코랩을 이용해 프로젝트를 진행해 Colab Notebooks 안에 **test_image** 라는 폴더에서 모든 프로젝트를 진행
test_image 폴더 안에 training directory에 있는 19장 영상 저장 후 프로젝트 진행

```
# 전역 변수 설정
TOTAL_IMAGES = 19
WIDTH, HEIGHT = 512, 512
K = [4, 16, 64]
N = [4, 8, 16, 32, 64, 128]
divided_images = [] # K차원 이미지 저장
```

과제에서 주어진 조건을 전역 변수로 설정

2. 환경 변수에 따른 결과 - 4차원

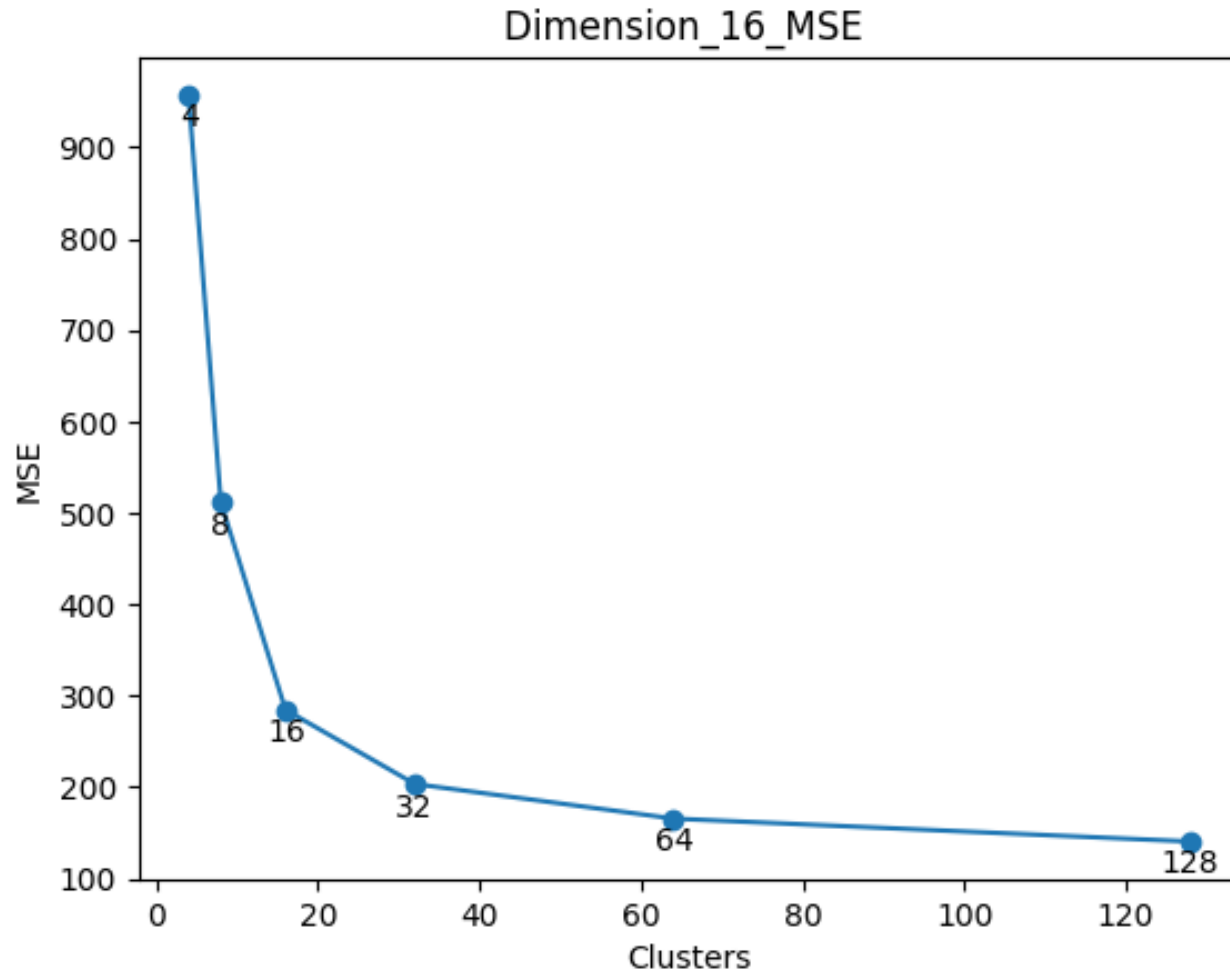


클러스터의 개수가 증가할수록 눈에 띄게 MSE가 감소되는 것을 확인할 수 있다.

클러스터의 개수가 증가함에 따라 클러스터 중심 간의 거리가 줄어들어 전체적인 MSE가 감소되는 것을 확인했다.

4차원에서는 이미지 조각이 너무 많아 128개의 클러스터를 진행할 때 가장 많은 시간이 걸렸다.

2. 환경 변수에 따른 결과 - 16차원



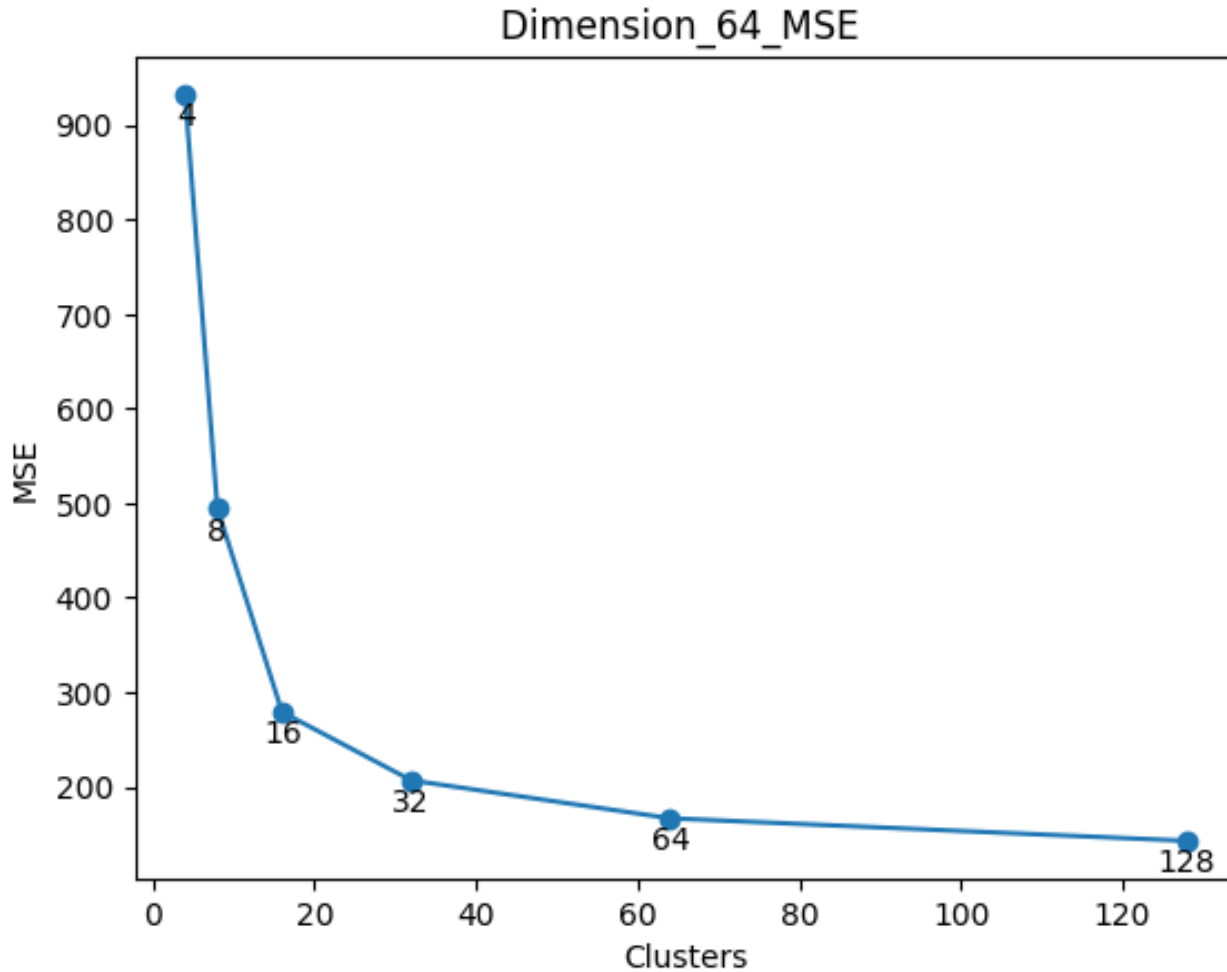
앞선 결과와 동일하게 클러스터의 개수가 증가함에 따라 MSE 값이 감소하는 경향을 보인다.

4차원에 비해 16차원으로 나눈 이미지는 더 넓은 영역으로 계산하다 보니 MSE값이 전체적으로 증가했다.

같은 조건에서 차원을 늘리기보단 줄이는 쪽이 더 정밀한 결과를 가져온다고 볼 수 있다.

하지만 4차원에 비해 전체적인 실행 시간이 적게 걸렸다.

2. 환경 변수에 따른 결과 - 64차원



해당 결과도 동일하게 클러스터의 개수가 증가할수록 MSE 값은 감소했다.

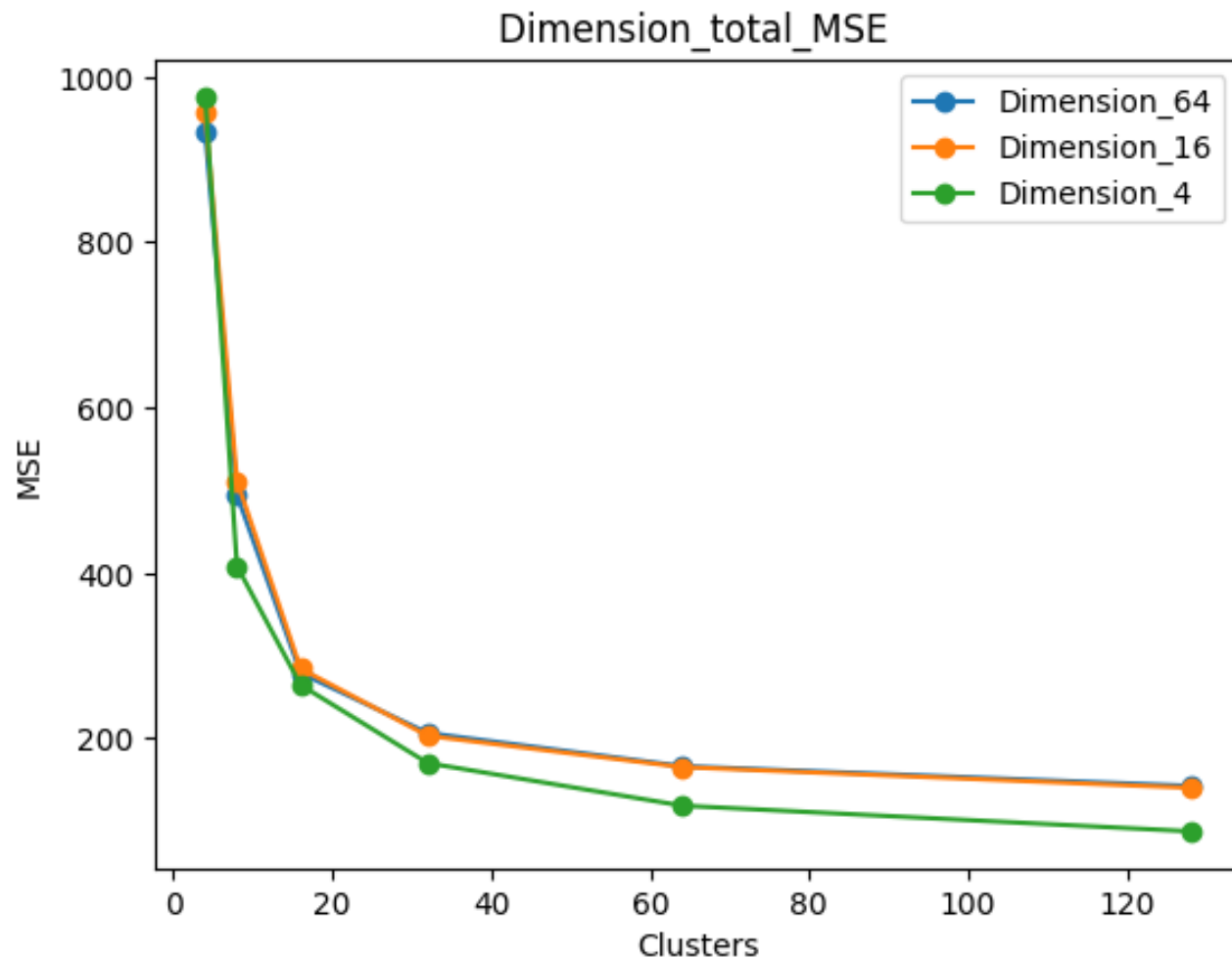
16차원의 MSE는 4차원의 MSE와 큰 차이가 나는데 비해 64차원의 MSE는 16차원의 MSE와 크게 차이가 나지 않았다.

이는 차원이 증가해도 각 차원 내 데이터의 변동성이 비슷하게 유지되었다고 볼 수 있다.

3. 최종 평가

MSE	K = 4	K = 16	K = 64
N = 4	976.0034	957.1981	932.4804
N = 8	408.3491	511.3271	494.5252
N = 16	264.7076	284.4688	279.0424
N = 32	170.245	203.2751	206.648
N = 64	118.7588	165.415	166.5509
N = 128	87.71	140.3218	142.6558

3. 최종 평가



클러스터의 개수가 적을 때 64차원과 16차원의 MSE는 큰 차이가 없는 반면, 4차원과는 차이가 보인다.

하지만 클러스터의 개수가 커짐에 따라 확실한 MSE 결과의 차이를 확인할 수 있다.

가장 정밀한 MSE 결과는 4차원의 128개의 클러스터를 사용할 때 **87.71**의 결과가 나왔다.

3. 최종 평가

차원이 높아질수록, 즉 데이터의 특성이 많아질수록 공간이 기하급수적으로 커져 데이터 간의 거리가 늘어난다.

이로 인해 차원과 MSE의 결과가 비례하여 증가한다고 생각한다.

또한 클러스터의 개수가 증가함에 따라, 각 데이터 포인트는 자신과 더 가까운 클러스터 중심을 가질 확률이 높아진다.

이는 MSE를 감소시키는 주요 원인으로 모든 시행에서 클러스터의 개수가 커짐에 따라 MSE가 작아지는 결과를 보였다.

최종적으로 4차원과 클러스터 128개에서 **87.71**로 가장 낮은 MSE 결과를 보였다.