Name: Nandita Koppisetty
Net-ID: nxk174230

K-means clustering for Image Compression

Image compression ratios for different values of K:

1. Penguins.jpg

Original Image
Height x Width: 768 x 1024
Image Size: 777.835 KB

Image with K = 2
Height x Width: 768 x 1024
Input Image Size: 777.835 KB
Output Image size: 83.093 KB
Compression Ratio: 9.361017

Image with K = 5
Height x Width: 768 x 1024
Input Image Size: 777.835 KB
Output Image size: 111.138 KB
Compression Ratio: 6.998821

Image with K = 10
Height x Width: 768 x 1024
Input Image Size: 777.835 KB
Output Image size: 68.269 KB
Compression Ratio: 11.393677

Name: Nandita Koppisetty
Net-ID: nxk174230

Image with K = 15
Height x Width: 768 x 1024
Input Image Size: 777.835 KB
Output Image size: 123.615 KB
Compression Ratio: 6.292399

Image with K = 20
Height x Width: 768 x 1024
Input Image Size: 777.835 KB
Output Image size: 117.681 KB
Compression Ratio: 6.609690

2. Koala.jpg:

Original Image
Height x Width: 768 x 1024
Image Size: 780.831 KB

Image with K = 2
Height x Width: 768 x 1024
Input Image Size: 780.831 KB
Output Image size: 132.608 KB
Compression Ratio: 5.888265

Name: Nandita Koppisetty
Net-ID: nxk174230

Image with K = 5
Height x Width: 768 x 1024
Input Image Size: 780.831 KB
Output Image size: 170.457 KB
Compression Ratio: 4.580809

Image with K = 10
Height x Width: 768 x 1024
Input Image Size: 780.831 KB
Output Image size: 168.692 KB
Compression Ratio: 4.628737

Image with K = 15
Height x Width: 768 x 1024
Input Image Size: 780.831 KB
Output Image size: 156.292 KB
Compression Ratio: 4.995975

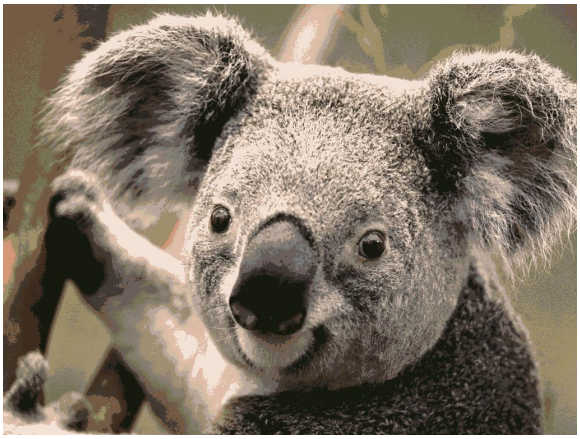

Image with K = 20
Height x Width: 768 x 1024
Input Image Size: 780.831 KB
Output Image size: 162.697 KB
Compression Ratio: 4.799296

Name: Nandita Koppisetty
Net-ID: nxk174230

Analysis:

There is a trade-off between quality of the image and the compression ratio.

For lower K values, the quality of the image is low as the number of colours representing the image is low and hence a lot of details in the image are compromised. Whereas higher K values produce a better image quality but takes longer time to run.

The goal here is to reconstruct the original image from the compressed representation without any loss of details rather than higher compression. Hence, for both the images the good values of K lie between 15 and 20.

For Penguins.jpg, 15 would be a good choice of K with compression ratio 6.292399 which is close to that of K=20. For Koala.jpg, 10 would be a good choice of K with compression ratio 4.628737 which is close to that of K=15 and K=20.