MACHINE-LEARNING ASSIGNMENT 3

HARIPRIYAA U MANIAN(hum160030)

2.K Means Clustering on Images:

1)Koala.jpg

Original Image:

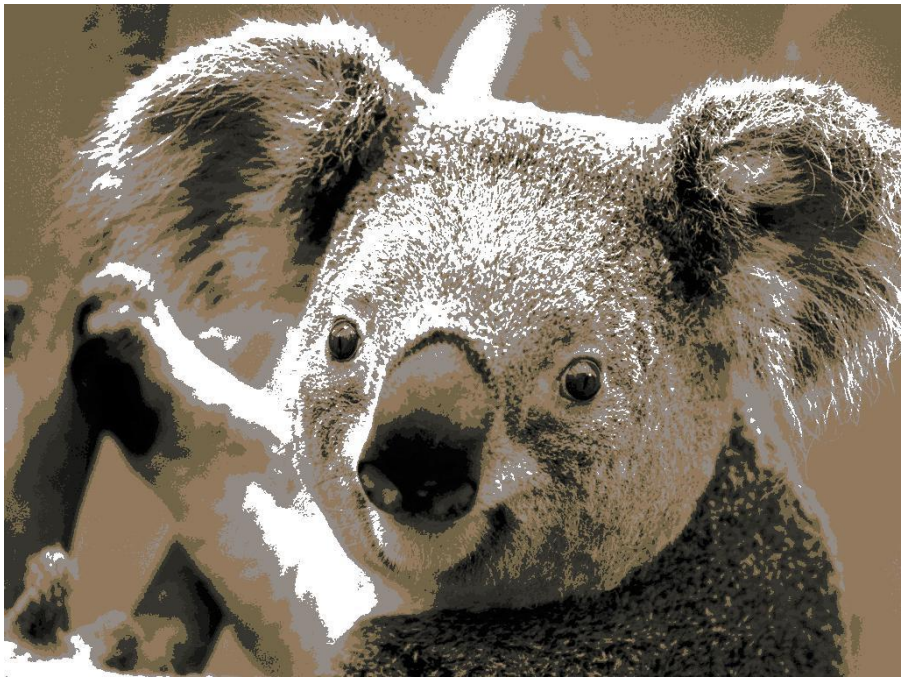
For k=2

Compression Ratio = 762kb/191kb = 3.9895

For k=5

Compression Ratio = 762kb/196kb =3.8877

For k=10

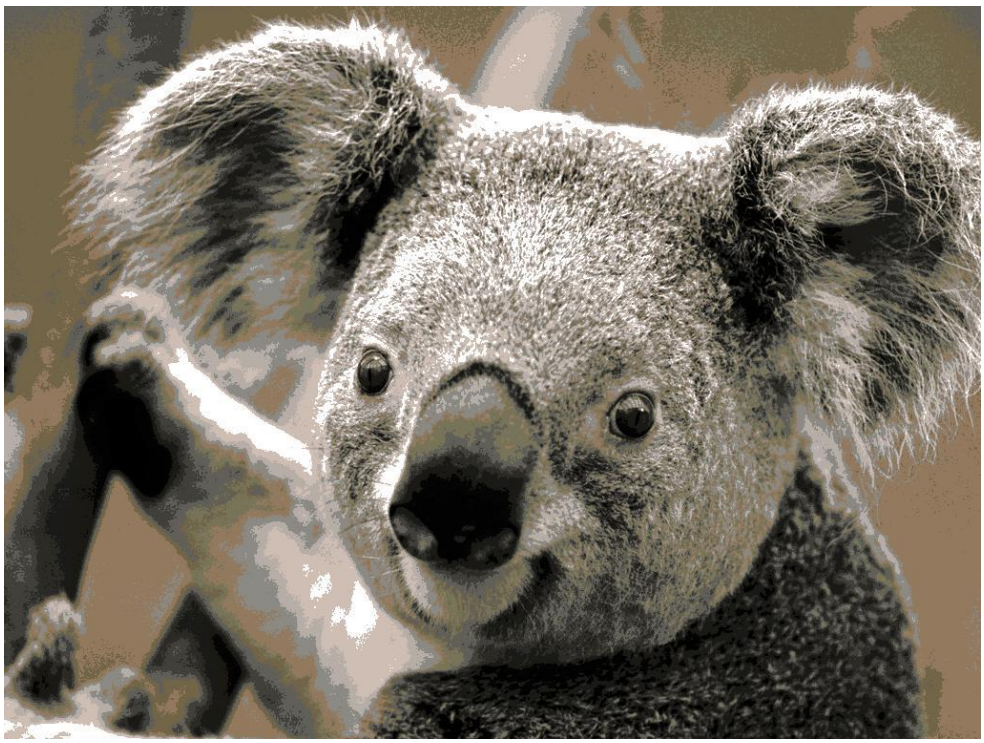


Compression Ratio = 762kb/180kb =4.2333

For k=15

Compression Ratio = 762kb/160kb =4.7625

For k=20

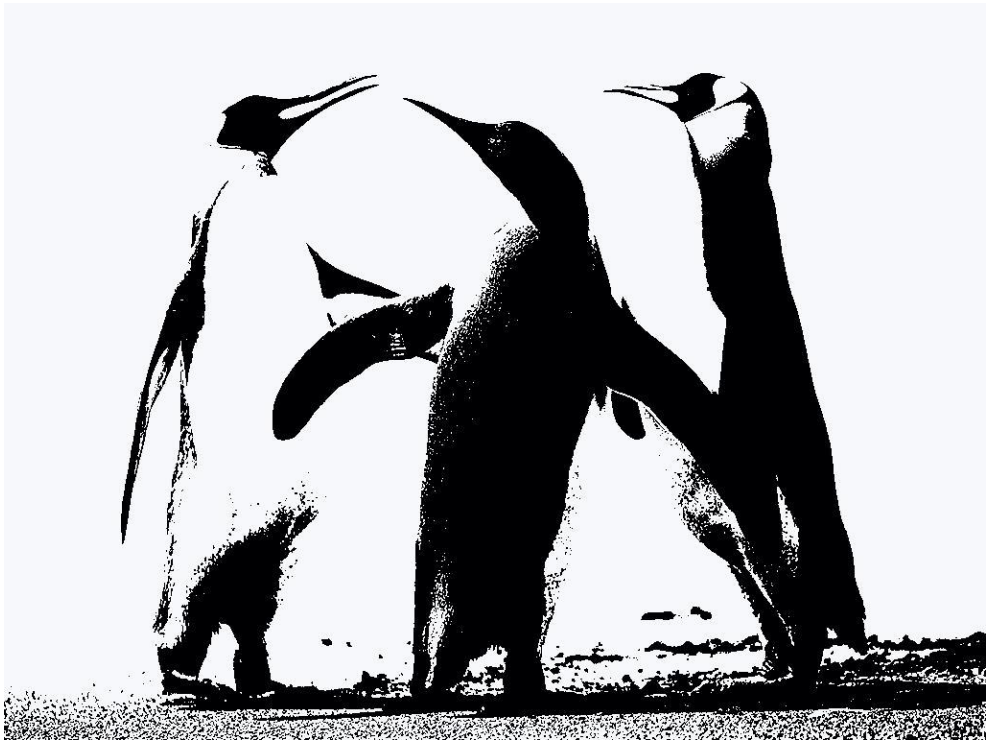


Compression Ratio = 762kb/162kb =4.7037

2) Penguins.jpg

Original Image

For k=2



Compression Ratio = 759kb/98.8kb =7.6821

For k=5

Compression Ratio = 759kb/129kb =5.8837

For k=10

Compression Ratio = 759kb/114kb =6.6578

For k=15

Compression Ratio = 759kb/113kb =6.7168

For k=20

Compression Ratio = 759kb/110kb =6.9

Average compression rate:

Koala:

Average compression ratio is 4.31534 and it varies by 0.2 for different k-values

Penguins:

Average compression ratio is 6.7680 and it varies by 0.4 for different k-values

Trade-off between image quality and degree of compression:

We get higher quality of images with the higher degree of compression

For the first image Koala—good k-value seems to be 20 since the quality seems to be high with twenty different clusters and image resembles the original image as shown

For Penguin.jpg, for a k-value of fifty, the image is better and also the degree of compression is high