

Image Compression using K-means clustering in R

This was a task issued by Sunayu as an evaluation of approach, style, ability, and performance. The R script will compress an image successfully by removing some information without losing the overall detail. Thus, the image will require less memory for storage, interpretation, and it still remains meaningful and informative.

Problem Statement

Image compression is a type of data compression applied to digital images, to reduce their cost for storage or transmission. Storage, while becoming cheaper and easier to obtain, remains a key factor for operations and processes of an organization. Image quality continues to get better and more detailed, therefore, requiring additional storage space than before. Find a way to compress an image to help combat this issue.

Data set

You may choose any image that is professionally reasonable to demonstrate a successful compression of an image using K-means clustering in R.

Compression Algorithm

Here we will compress the image into compressed images with clusters of 2, 4, 8, 16, 32, 64, and 128.

This compression algorithm caused the original image chosen (Satellite.png) to go from 6.5MB to:

- 2 clusters: 262KB
- 4 clusters: 459KB
- 8 clusters: 647KB
- 16 clusters: 1.05MB
- 32 clusters: 1.54MB
- 64 clusters: 1.94MB
- 128 clusters: 2.57MB