AIP Assignment 4

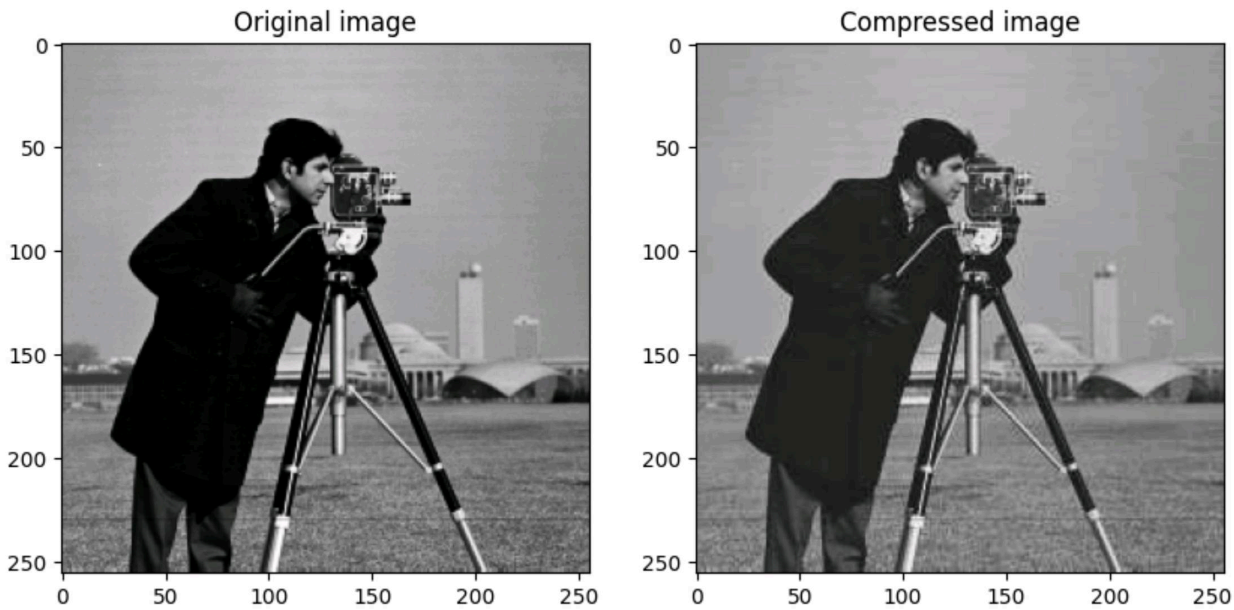
Name : B Srinath Achary

Sr No: 21441

Problem 1

Question 1 :

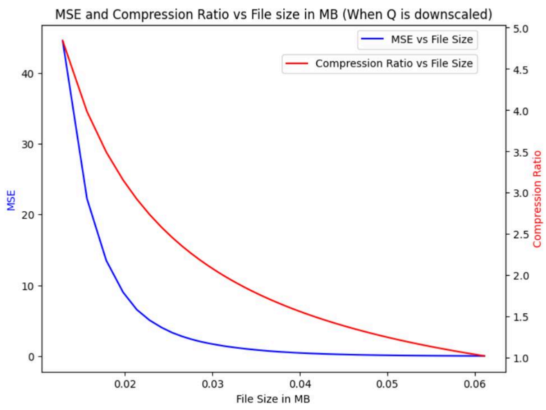
Results:

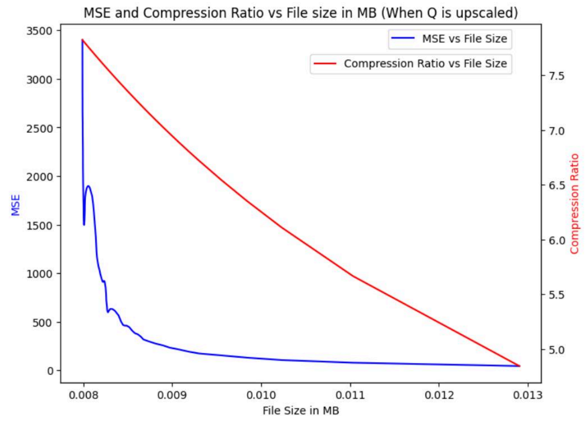


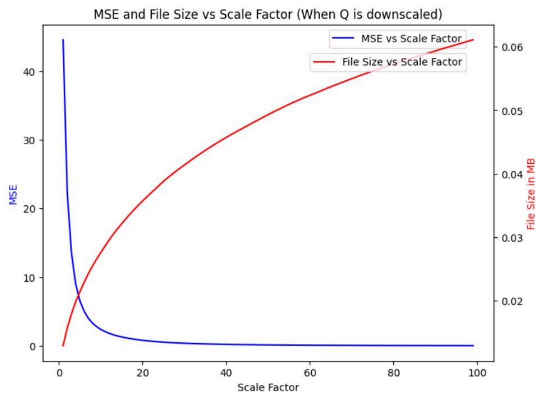
- Input Size : 64 KB
- Output Size : 13.21 KB
- Compression Ratio : 4.84
- MSE : 44.54

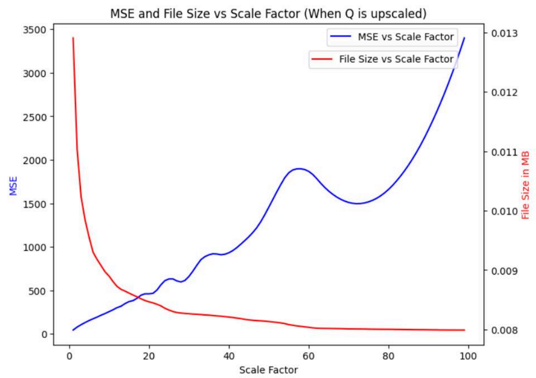
Question 2 :

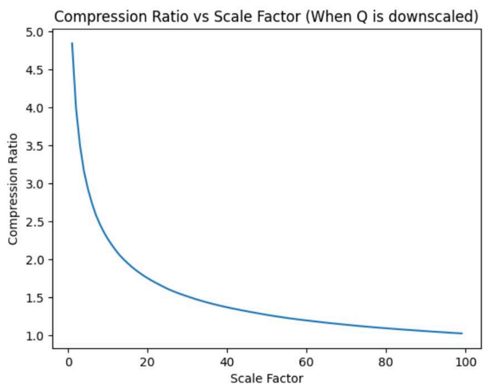
Results :

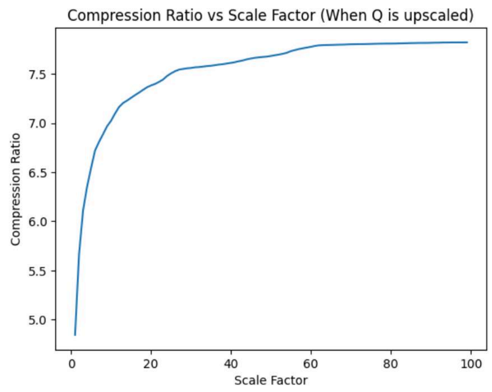












Analysis :

- Upon examining Figure 1 and Figure 2, we can observe that there is an inverse relationship between Mean Squared Error (MSE) and File Size in both cases, where the quantization matrix Q is either up-scaled or down-scaled. This implies that as the file size decreases, the MSE between the original and reconstructed images increases. Additionally, we can see that there is also an inverse relationship between Compression Ratio and File Size. This behavior can be explained by the fact that by up-scaling the quantization matrix, leads to a larger number of DCT coefficients being quantized to zero or very small values, resulting in a sparser bitstream. As a result, the file size decreases, but at the same time, there is a loss of information due to quantization, which leads to an increase in MSE between the original and reconstructed images. Conversely, by down-scaling the quantization matrix, leads to a smaller number of DCT coefficients being quantized to zero or very small values, resulting in a denser bitstream. Consequently, the file size increases, but the reconstructed image has a lower MSE compared to the quantized version since there is less loss of information due to quantization. Overall, these observations highlight the trade-off between compression ratio and image quality in JPEG encoding, where decreasing the file size through quantization leads to a loss of information and increase in distortion.

- Based on figures 3 and 4, it is observed that when the quantization matrix Q is downscaled, the mean squared error (MSE) decreases as the scaling factor increases. However, the file size increases with an increase in the scaling factor, indicating that more bits are required to represent the compressed image. This can be explained by the fact that a smaller quantization matrix results in less quantization error, thereby leading to lower MSE. However, a smaller quantization matrix also implies that more coefficients need to be encoded, leading to a larger file size. On the other hand, when the quantization matrix Q is upscaled, we observe an opposite trend, where the MSE increases with an increase in the scaling factor, while the file size decreases. This can be explained by the fact that a larger quantization matrix results in higher quantization error, thereby leading to higher MSE. However, a larger quantization matrix also implies that fewer coefficients need to be encoded, leading to a smaller file size.

- The results from Figure 5 and 6 show that for a down-scaled Q, the compression ratio decreases with an increase in the scale factor. On the other hand, when Q is upscaled, we observe an opposite trend. The compression ratio is defined as the ratio of the input image in bits to the size of the output file in bits. When the Q matrix is downscaled, it leads to a higher level of quantization, which results in larger quantized indices and, thus, larger bitstream sizes. This, in turn, reduces the compression ratio. However, when the Q matrix is upscaled, it leads to lower quantization levels, resulting in smaller quantized indices and, thus, smaller bitstream sizes, leading to a higher compression ratio. Hence, we observe an inverse relationship between the compression ratio and the scale factor of the Q matrix.

Problem 2

Given 2 uniformly distributed continuous independent sources X_1 & X_2 both with mean zero and variances 5 & 10 respectively

Then $R_i = \log_2 K_i$ for $i \in \{1,2\}$

Constraint: $R_1 + R_2 = 3$.

Now we have $E[(X-\hat{X})] = \frac{\sigma^2}{2^{2R}}$

So we have to minimize : $\frac{\sigma_1^2}{2^{2R_1}} + \frac{\sigma_2^2}{2^{2R_2}}$ s.t $R_1 + R_2 = 3$. & $R_1, R_2 \in \mathbb{Z}^+$

Here $\sigma_1^2 = 5$, $\sigma_2^2 = 10$.

Now possible pairs (R_1, R_2): $(1,2)$ & $(2,1)$.

For $(R_1, R_2) = (1,2)$ MSE : $\boxed{1.875}$

For $(R_1, R_2) = (2,1)$ MSE : $\boxed{2.8125}$

Best pair : $\boxed{R_1 = 1, R_2 = 2}$

Allocation of bits :

From the above calculation,

For X_1 we allocate 1 bit & for X_2 we allocate 2 bits.

Now $\sigma_1^2 = \frac{(x_1^{max} - x_1^{min})^2}{12}$

$\Rightarrow x_1^{max} - x_1^{min} = 2\sqrt{15}$ $\Rightarrow$ $x_1^{max} = \sqrt{15}$, $x_1^{min} = -\sqrt{15}$.

And we have $\frac{x_1^{max} + x_1^{min}}{2} = 0$

Now $K_1 = 2^{R_1} \Rightarrow K_1 = 2.$

$$\Delta_1 = \frac{x_1^{max} - x_1^{min}}{K_1} = \sqrt{15}$$

Now $v_0 = -\sqrt{15}, \quad v_1 = 0, \quad v_2 = \sqrt{15}.$

And $\hat{x}_0 = \frac{-\sqrt{15}}{2} \quad \& \quad \hat{x}_1 = \frac{\sqrt{15}}{2}.$

So we can encode $\hat{x}_0 \rightarrow 0$ & $\hat{x}_1 \rightarrow 1.$

For X_2:

$$\sigma_2^2 = \frac{(x_2^{max} - x_2^{min})^2}{12}$$

$x_2^{max} = \sqrt{30}$

$x_2^{min} = -\sqrt{30}$

$\Rightarrow \quad x_2^{max} - x_2^{min} = 2\sqrt{30}.$

Now $K_2 = 2^{R_2} \Rightarrow K_2 = 2^2 = 4.$

$$\Delta_2 = \frac{x_2^{max} - x_2^{max}}{K_2} = \frac{2\sqrt{30}}{4} = \frac{\sqrt{30}}{2}$$

Now $v_0 = -\sqrt{30}, \quad v_1 = -\frac{\sqrt{30}}{2}, \quad v_2 = 0, \quad v_3 = \frac{\sqrt{30}}{2}, \quad v_4 = \sqrt{30}$

$\hat{x}_0 = -\frac{3}{4}\sqrt{30}, \quad \hat{x}_1 = \frac{-\sqrt{30}}{4}, \quad \hat{x}_2 = \frac{\sqrt{30}}{4}, \quad \hat{x}_3 = \frac{3}{4}\sqrt{30}$

So we can encode

$\hat{x}_0 \rightarrow 00$

$\hat{x}_1 \rightarrow 01$

$\hat{x}_2 \rightarrow 10$

$\hat{x}_3 \rightarrow 11$

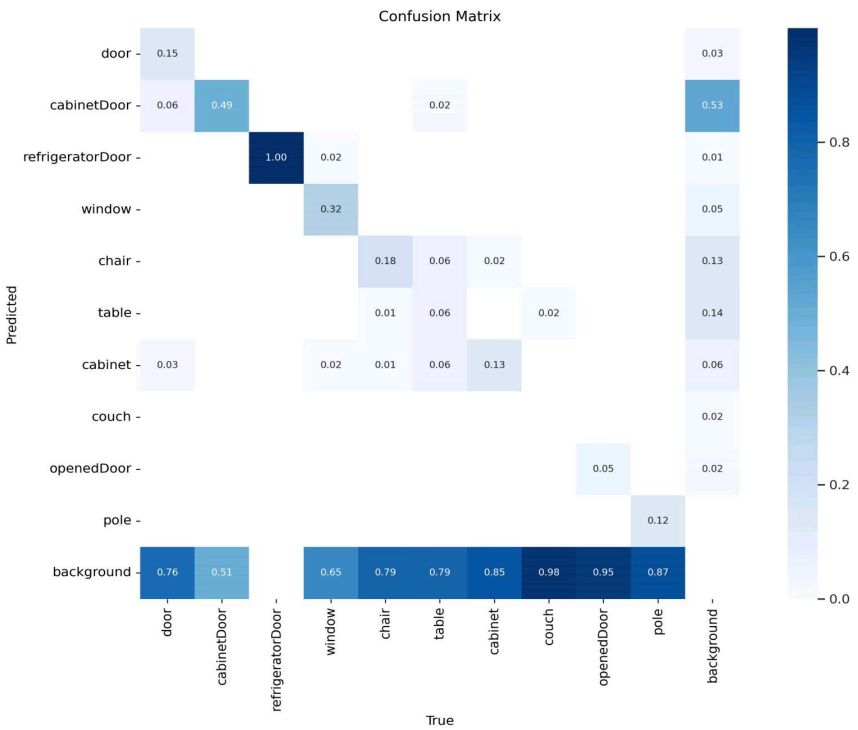
Problem 3

Results :

Metric Table:

Class	Images	Instances	P	R	mAP50	mAP50-95:
all	107	550	0.372	0.26	0.263	0.108
door	107	34	0.481	0.176	0.226	0.0976
cabinetDoor	107	179	0.459	0.464	0.393	0.134
refrigeratorDoor	107	2	0.31	1	0.995	0.338
window	107	63	0.635	0.331	0.398	0.22
chair	107	87	0.364	0.184	0.196	0.0751
table	107	47	0.0913	0.0851	0.0323	0.0103
cabinet	107	52	0.287	0.135	0.111	0.0653
couch	107	58	0	0	0.0334	0.00828
openedDoor	107	20	0.262	0.1	0.0545	0.0198
pole	107	8	0.836	0.125	0.191	0.112

Confusion Matrix:



True Labels :

Predicted Labels :

Analysis :

These results show the performance of YOLOv5 object detection model on different classes of objects in terms of precision (P), recall (R), mean average precision (mAP) with IoU threshold of 0.5 (mAP50), and mean average precision with IoU thresholds ranging from 0.5 to 0.95 (mAP50-95). The model was evaluated on a total of 107 images.

- Looking at the results, we can see that the overall performance of the model is not very impressive, with a low precision and a relatively low recall for most of the classes. This means that while the model was able to detect some objects, it also produced many false positives and missed many true positives.

- Breaking down the results by class, we can see that the model performed best on the "refrigeratorDoor" class, with a perfect recall of 1 and a high precision of 0.31. This means that the model was able to detect all instances of refrigerator doors in the images, and most of the detections were correct.

- On the other hand, the model performed poorly on the "table" and "couch" classes, with very low precision and recall values. This indicates that the model struggled to detect these objects in the images, and most of the detections it made were false positives.

- From the confusion matrix we can see that for almost all classes (except refrigerator Door) model is categorizing most of the objects as background. One possible reason could be that the dataset used to train the model has a lot of images with a cluttered background, making it difficult for the model to distinguish between the object and the background.