Q2 : CNN for Colorization a) Regression

(iv) **Regression vs Classification** : For Colourization as Regression, we usually Mean Square Error (MSE) for loss function.

Q: what does MSE encourage ? We are trying to convert grayscale to RGB – but as per RGB Luminance eqn

$$\text{Luminance} = 0.3\,R + 0.59\,G + 0.11\,B$$

where R, G, B are the pixel values in respective channels

This eqn is used in OpenCV's cvtColor() as well from RGB to gray conversion.

Now, for gray to RGB conversion, a single grayscale can have multiple possible colored version ie the colorization problem is multi-modal.

Therefore using MSE, the model chooses a desaturated color (more black & white, less "colorful")

This is because saturation,

$$S = \frac{\text{Max(RGB)} - \text{Min(RGB)}}{1 - |2L - 1|}$$

where L, Luminosity $= \frac{1}{2}(\text{Max(RGB)} + \text{Min(RGB)})$

To minimize the MSE value, as the network predicts R, G, B values of near-equal value, ie the picture output is more black & white, less vivid / brilliant color (ie. low Saturation)

This is why most colourization SOT models use classification instead of regression.