Colorful Image Colorization

- LAB space
- Quantize ab output space: bins with grid size = 10, Q = 313 (Number of quantized ab pairs)
- For a given input X, predict color distribution $\hat{Z}$:

$$\hat{Z} = G(X) \text{ where } \hat{Z} \in [0,1]^{H \times W \times Q}$$

- From $\hat{Z}$ (a distribution) to $\hat{Y}$ (a point in ab space):
  - **Mode**: vibrant but strange details
  - **Mean**: desaturated color, similar to Euclidean loss
  - **Annealed mean**: $H(Z_{h,w}) = E[f_T(Z_{h,w})]$, $f_T(z) = \frac{e^{log(z)/T}}{\sum_q e^{log(z_q)/T}}$
- Ground truth color Y is converted to distribution Z using **soft encoding**:

  Find **5** nearest neighbours to Y in output space, weight them $\propto$ distance from Y using Gaussian kernel with $\sigma = 5$
- **Multinomial cross entropy loss**:

$$L_{cl}(\hat{Z}, Z) = -\sum_{h,w} v(Z_{h,w}) \sum_q Z_{h,w,q} log(\hat{Z}_{h,w,q})$$

  - $v(Z_{h,w})$ class rebalancing:
    - low ab values dominate natural images (grayish, due to clouds, pavement, dirt, walls, etc.)
    - Increase importance of rare colors:
      1. Estimate empirical probability distribution of colors in quantized ab space $p \in \Delta Q$.
      2. Smooth p to $\tilde{p}$ with Gaussian kernel G_σ, $\sigma = 5$.
      3. Mix $\tilde{p}$ with a uniform distribution $\frac{1}{Q}$ (tones down importance of rare colors slightly), then take reciprocal (rare colors importance > frequent colors): $w \propto ((1-\lambda)\tilde{p} + \lambda \times \frac{1}{Q})^{-1}$, $\lambda = .5$.
      4. Normalize w so that $E[w] = \sum_q \tilde{p}_q w_q = 1$