

# 1 Normalization 常用归一化公式

## 1.1 Instance Normalization

实例归一化

对于一个输入特征  $\mathbf{x}$  (通常是一个多通道的特征图), `InstanceNorm` 的计算过程如下:

$$\hat{x}_{i,c} = \frac{x_{i,c} - \mu_c}{\sigma_c + \varepsilon} \times \gamma_c + \beta_c \quad (1)$$

其中:

- $\mathbf{x}(i, c)$  是通道  $c$  处的特征值。
- $\mu_c$  是该通道的均值:

$$\mu_c = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W x_{i,c,h,w} \quad (2)$$

- $\sigma_c$  是该通道的标准差:

$$\sigma_c = \sqrt{\frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W (x_{i,c,h,w} - \mu_c)^2} \quad (3)$$

- $\gamma_c$  和  $\beta_c$  是可学习的缩放和平移参数。
- $\varepsilon$  是一个小的常数, 用于防止除零错误。

`InstanceNorm` 的归一化计算是针对每个样本的每个通道单独进行的, 与 `BatchNorm` 不同, 它不依赖于 `mini-batch` 维度。