

整字预测准确率评估算法

1.采用连通域分析判断错误笔画

在评估模型预测效果时，采用连通域分析结合面积占比判定来区分局部错误（像素级）和全局错误（整笔画错误），从而更合理地计算预测准确率。

若预测笔画连通域与真实笔画差异过大（面积占比超阈值）→ 判定为整笔画错误，直接扣除该笔画所有像素的得分。

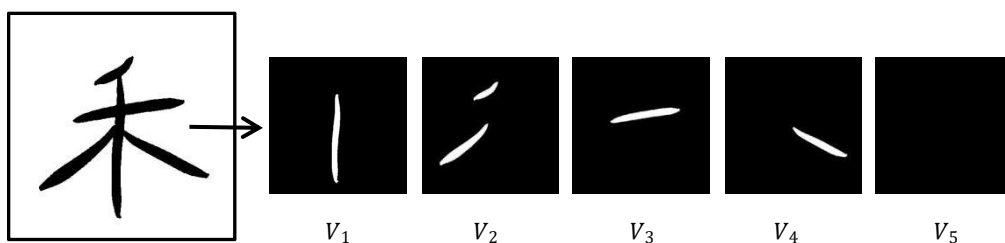
若预测笔画连通域与真实笔画差异较小（面积占比低于阈值）→ 则按像素级错误计算。

1.1 连通域提取：

对于每类笔画类别 $k \in \{1, 2, 3, 4, 5\}$

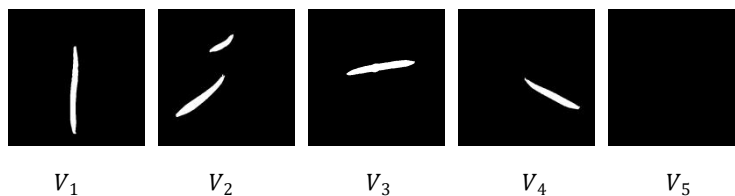
1.1.1 真实连通域提取

$$A_{true}^k = \sum_{i=1}^n C_i^k$$



1.1.2 预测连通域提取

$$\bar{A}_{prediction}^k = \sum_{i=1}^n \bar{C}_i^k$$



1.2 预测连通域匹配：

由于预测连通域与真实连通域都是二值化掩码，所以直接进行交运算，从而算出预测连通域与真实连通域的重合面积，与真实连通域面积相比进行有效性判别。

$$IoU^k = \frac{A_{true}^k \cap \bar{A}_{prediction}^k}{A_{true}^k}$$

1.3 有效笔画判定

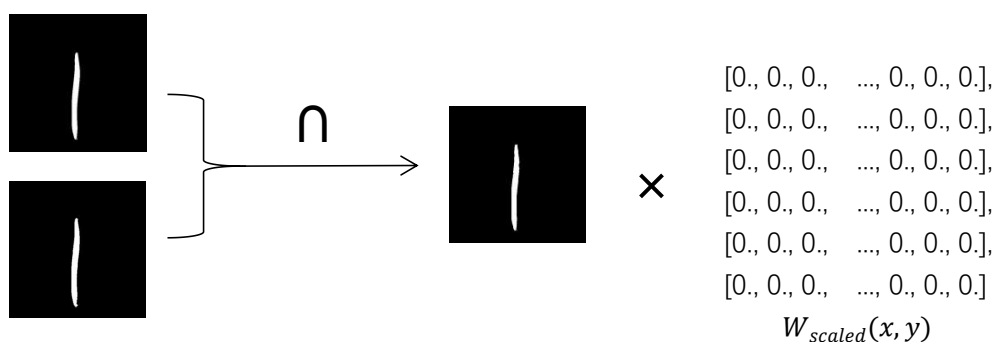
$$valid^k = \begin{cases} 1 & \text{if } IoU^k \geq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

2.计算每一类笔画(连通域)的准确率 L_k

对于第 k 类笔画，若其 $valid^k$ 为 0，则该类笔画整体预测结果判定为错误，其评估率为 0，若其 $valid^k$ 为 1，则该类笔画整体预测结果符合要求，按照像素权重计算评估率

$$L^k = 100 \times \frac{valid^k \times \sum_{(x,y) \in A_{true}^k} W_{scaled}(x,y) \times I(X_{x,y;k} = \bar{X}_{x,y;k})}{\sum_{(x,y) \in A_{true}^k} W_{scaled}(x,y)}$$

这里的 $I(*)$ 表示当真实图像与预测图像在 (x,y) 处值相等时，取 1，反之取 0。



3.计算每一类笔画的权重分配

基于笔画面积分配第 k 类笔画的权重，计算该类笔画面积 S_k ：

$$S_k = \sum_{i=1}^{500} \sum_{j=1}^{500} \Pi(X_{i,j,k} = 1)$$

这里的 $\Pi(*)$ 表示该像素值是否为 1，满足条件取 1，不满足取 0。

获取每一类笔画的面积后，计算各类权重 F_k ：

$$F_k = \frac{S_k}{\sum_{k=1}^5 S_k}$$

4.计算总准确率

$$G = \sum_{k=1}^5 F_k \times L_k$$