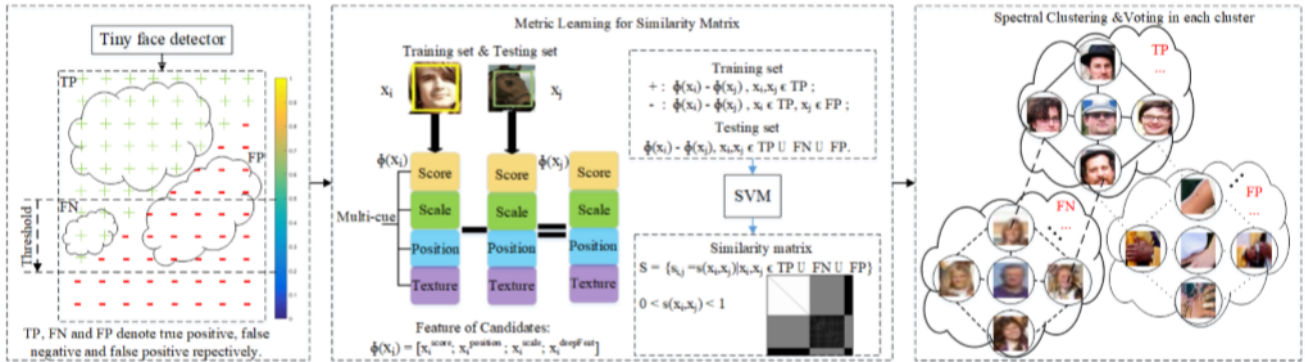


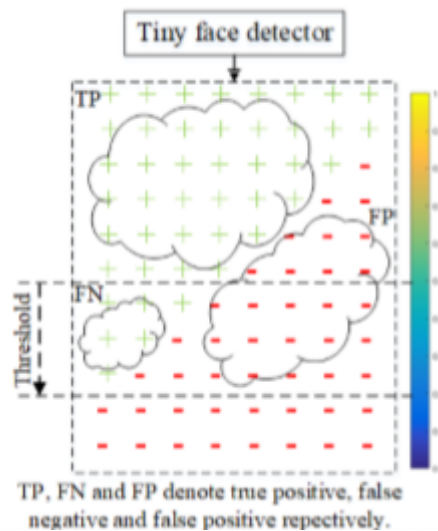
BEYOND CONTEXT: EXPLORING SEMANTIC SIMILARITY FOR TINY FACE DETECTION ---- 读后感

本文从图片中选出的后选区之间的上下问联系进行分析，所以本文的实验是基于其他已有基础的候选框选取方法进行。而本文的创新为对所有的后选区进行聚类。



利用线性SVM计算相似度

在微小人脸检测器中获取的特征值 $\phi(x_i)$ ，(分类得分，标量，位置，纹理等)，然后将这些后选区在训练阶段按照降序排列。然后假定排名前10%的是人脸区域，后10%是非人脸区域。

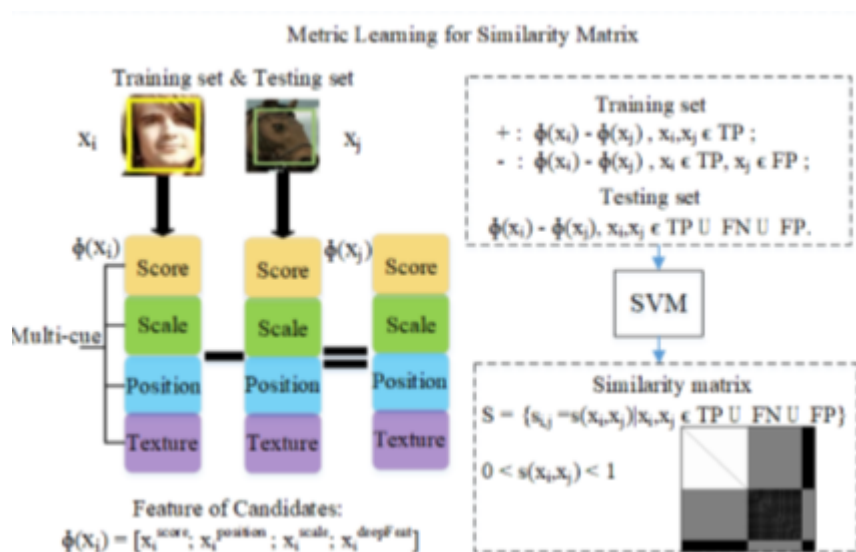


输入训练集如下

$$\{(x'_{11}, y'_{11}), (x'_{12}, y'_{12}), \dots, (x'_{nn}, y'_{nn})\}, \text{ 其中 } x'_{ij} = \phi(x_i) - \phi(x_j)$$

$$y'_{ij} = \begin{cases} 1 & x_i, x_j \in X_{TOP10\%} \\ 0 & x_i \in X_{TOP10\%}, x_j \in X_{BOTTOM10\%} \end{cases}$$

当测试阶段，将 x'_i 输入到SVM分类器中，得到 x_i 与 x_j 之间的相似度得分 $s_i[0,1]$ ，从而构建相似度矩阵S



基于光谱的分割图像聚类

现在我们已经获取了一个后选区的相似度矩阵，由此可以构建一个全连接无向图 $G=(V, E)$, $v_i \in V$ 表示后选区 x_i , $e_{ij} \in E$ 表示 x_i 和 x_j 的相似度。所以如果要获取一个聚类，那么应该满足：

- 不同聚类点之间的权重较低(每个后选区与其他类中的后选区相似度较低)
- 同一个聚类点之间的权重较高(同一个类型的后选区的相似度应该较高)

于是对于一个聚类，有如下公式可以表达：

$$cut(A_1, A_2, \dots, A_K) = \frac{1}{2} \sum_{i=1}^k W(A_i, \vec{A}_i)$$

$$A_i \subset V, A_i \cap A_j = \emptyset, A_1 \cup A_2 \cup \dots \cup A_k = V \quad W(A_i, \vec{A}_i) = \sum_{m \in A_i, n \in A_j} exp(-S_{mn}/2\delta^2)$$

为了防止聚类个数的不均衡情况，用 $|A|$ 表示 A 中定点的个数，来确保分类 $\{A_1, A_2, \dots, A_k\}$ 的聚类大小是均衡的。因此可将上述公式改为：

$$cut(A_1, A_2, \dots, A_K) = \frac{1}{2} \sum_{i=1}^k \frac{W(A_i, \vec{A}_i)}{|A_i|}$$

又根据参考文献 Ulrike Von Luxburg, "A tutorial on spectral clustering," Statistics and computing, 2007.

$$\arg \min_H cut(A_1, A_2, \dots, A_k) = \arg \min_H Tr(H^T L H)$$

其中 L 是 Laplacian matrix (图的度矩阵减去邻接矩阵) 且 $H^T H = E$, $H = \{h_1, h_2, \dots, h_k\}$

$$h_{i,j} = \begin{cases} \frac{1}{\sqrt{|A_j|}} & v_i \in A_j \\ 0 & otherwise \end{cases} \quad i = 1, 2, \dots, N, j = 1, 2, \dots, k$$

通过选择包含 L 的第 k 个特征向量的矩阵利用 K-means 算法得到结果，然后就获取了每一个类是否是人脸还是非人脸。

个人总结

利用每个后选区的分类进行聚类来对那些可能判断失误的后选区进行补救，即对FP的后选区进行进一步的筛选，减少FP的区域，同时可能导致FN和TG的后选区错误选择。即是对后选区进行聚类，然后根据聚类中的后选区的人脸可能综合判断人脸。