

场景文本检测：全卷积网络（FCN），非极大值抑制（NMS）

FCN 网络结构：

(1) 特征提取

特征提取，得到不同尺度的特征图

(2) 特征合并

逐层合并，从特征提取网络的顶部特征按照相应的规则向下进行合并

(3) 输出层

按照文本形状（旋转框 RBOX、水平框 QUAD）输出，得到相应损失函数

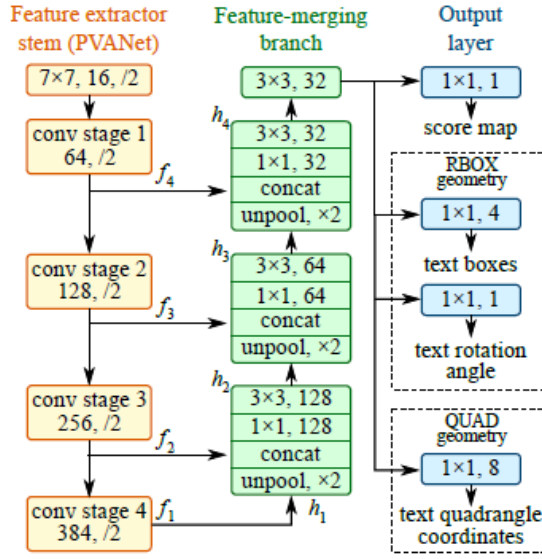


Figure 3. Structure of our text detection FCN.

NMS 结构

对输出集合结合阈值选择是否合并，合并之后进行 NMS 操作。

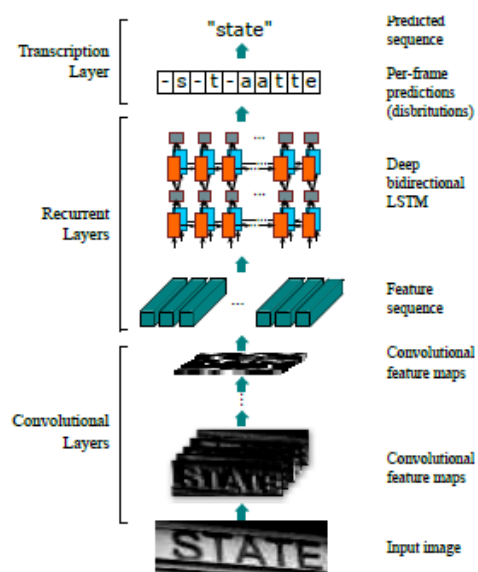
Algorithm 1 Locality-Aware NMS

```

1: function NMSLOCALITY(geometries)
2:    $S \leftarrow \emptyset, p \leftarrow \emptyset$ 
3:   for  $g \in \text{geometries}$  in row first order do
4:     if  $p \neq \emptyset \wedge \text{SHOULDMERGE}(g, p)$  then
5:        $p \leftarrow \text{WEIGHTEDMERGE}(g, p)$ 
6:     else
7:       if  $p \neq \emptyset$  then
8:          $S \leftarrow S \cup \{p\}$ 
9:       end if
10:       $p \leftarrow g$ 
11:    end if
12:  end for
13:  if  $p \neq \emptyset$  then
14:     $S \leftarrow S \cup \{p\}$ 
15:  end if
16:  return STANDARDNMS( $S$ )
17: end function

```

卷积循环神经网络 (CRNN)



CRNN 网络结构

(1) 卷积层

从输入图像上提取特征序列，主要有 CNN 模型中卷积层和最大池化层构成。

(2) 循环层

深度双向循环神经网络，通过 LSTM 基本网络单元实现。

(3) 转录层，

将每帧预测转换为标签序列（无词典转录，有词典转录）