

基于特征融合网络的场景文字检测方法

蔡晨沁*, 吕品, 苏冰

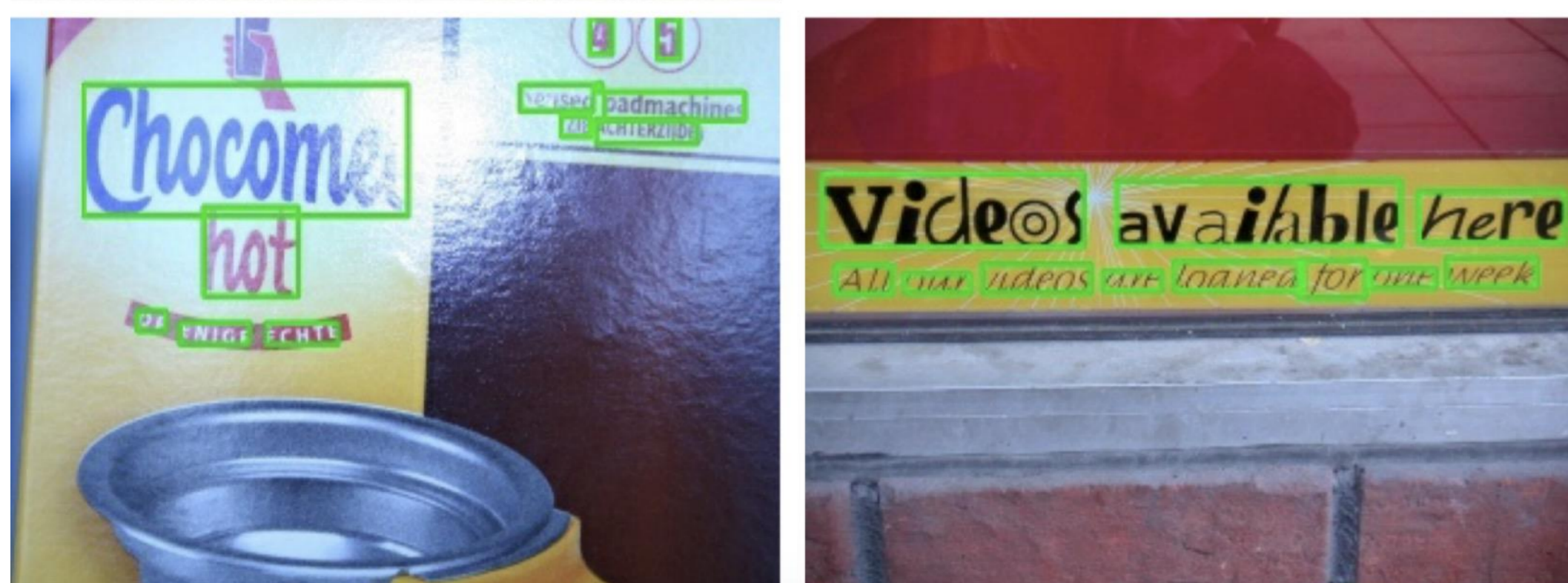
Feature fusion network for scene text detection[J].
international conference on image processing, 2018:2755-2759.

*通讯方式: 18810217852; chenqin2016@iscas.ac.cn

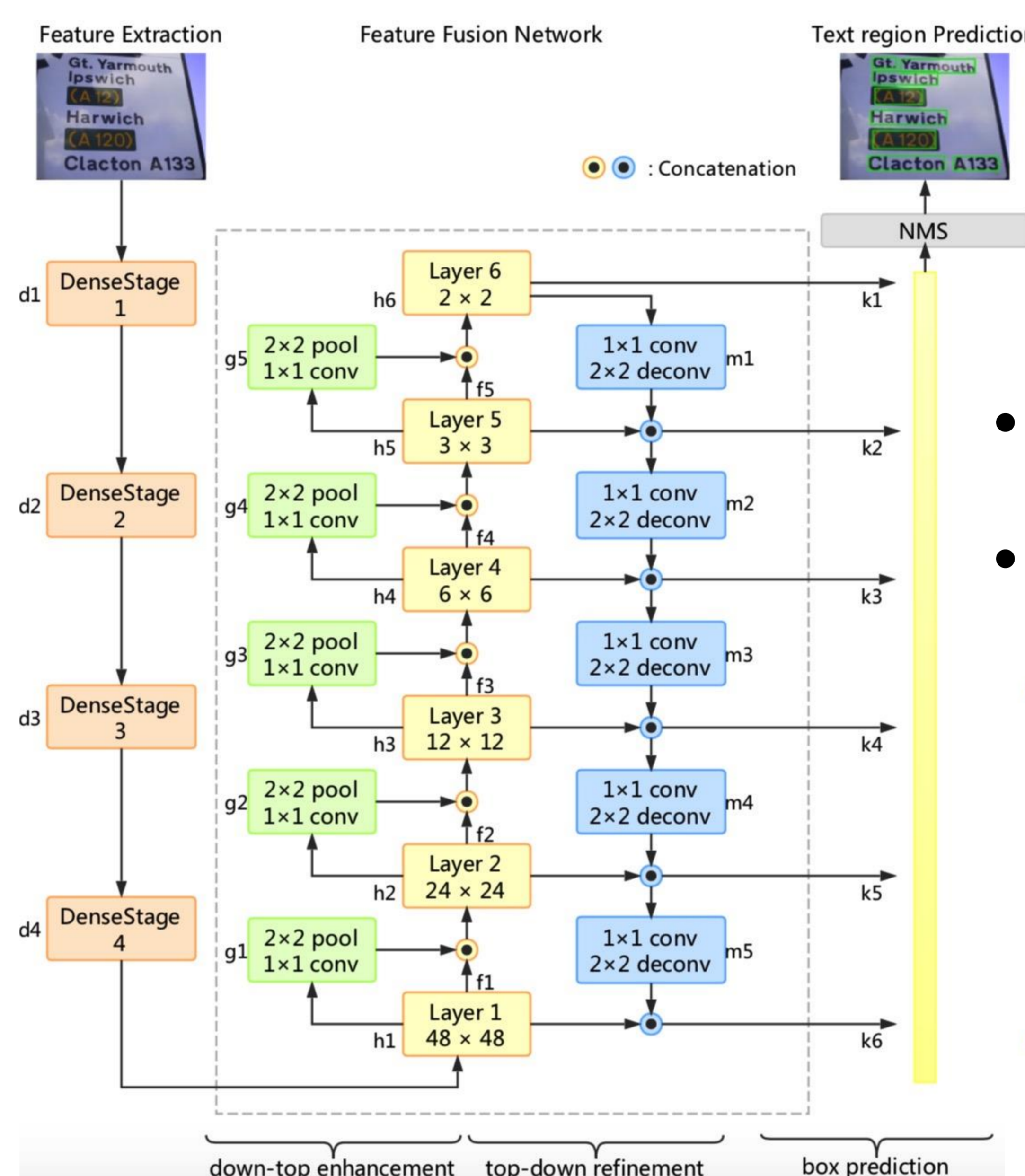
摘要

设计一个特征融合网络，双向连接深度神经网络中深层与浅层的网络，以检测多个尺度的文本区域，该模型支持端到端场景文本检测器，无需预先训练的模型即可检测各种比例的文本区域。

检测结果



特征融合网络



融合公式

- 由浅到深的特征增强阶段
- 由深到浅的特征修正阶段

① Down-top enhancement stage:

$$f_i = \text{Conv}_{3 \times 3}(\text{Conv}_{1 \times 1}(h_i)),$$

$$g_i = \text{Conv}_{1 \times 1}(\text{Pool}_{2 \times 2}(h_i)),$$

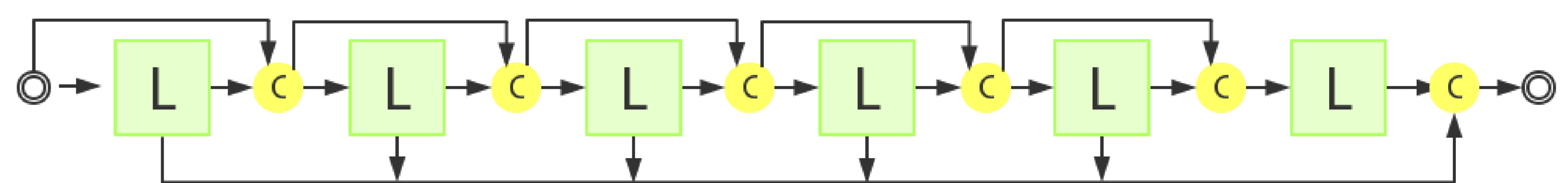
$$h_i = \begin{cases} d_4, & i = 1, \\ \text{Concat}(f_{i-1}, g_{i-1}), & 2 \leq i \leq 6, \end{cases}$$

② Top-down refinement phase:

$$m_j = \text{Deconv}_{2 \times 2}(\text{Conv}_{1 \times 1}(k_j)),$$

$$k_j = \begin{cases} h_6, & j = 1, \\ \text{Concat}(m_{j-1}, h_{7-j}), & 2 \leq j \leq 6, \end{cases}$$

改进的DenseNet



总结

双向融合的特征网络，同时考虑到文本区域的大小小于或大于特定比例的情况，通过防止不必要的检测（例如，针对一个文本区域的不同比例的多个框），在适当的层中检测特定大小的文本区域，提高准确率和召回率。