

# SS-EPA：基于补丁语义亲和力增强的单阶段弱监督语义分割方法

姜景杰<sup>1)</sup> 周肖桐<sup>1)</sup> 李轩慧<sup>1)</sup> 张轩豪<sup>1)</sup> 杨钧茗<sup>1)</sup>

<sup>1)</sup>(南京信息工程大学 计算机学院、网络空间安全学院 南京 中国 210044)

**摘要** 图像级弱监督语义分割 (Image Level Weakly Supervised Semantic Segmentation, Image Level WSSS) 通常利用类激活图 (Class Activation Map, CAM) 来生成伪标签 (Pseudo Label)。先前WSSS方法无论使用CNN还是Transformer框架, 多数都采用多阶段方法, 需要分阶段地训练模型和采取不同的训练策略, 多个阶段间的复杂交互较为繁琐。且先前的方法通常直接通过ViT中的语义亲和力优化CAM, 对计算资源要求较高, 还可能会给CAM带来错误和误导。本文提出了一种名为SS-EPA (Single Stage WSSS with Enhanced Patch Affinity) 的单阶段WSSS方法, 集成了端到端式的多头自注意力CAM优化方法, 利用ViT中的补丁语义亲和力 (Patch Affinity) 信息, 对从补丁令牌生成的初始CAM执行优化。为了解决语义亲和力中噪声、错误和注意力图过于庞大的问题, 本文提出了头平均注意力融合增强模块 (Head Average Attention Fusion, HAAF)。HAAF通过对ViT中不同头的注意力权重执行平均操作来聚合语义信息, 去除头重复关注、包含无效信息的冗余问题, 提取更加精简有效的信息, 极大减少计算资源占用。在Pascal VOC 2012数据集上的实验表明, 本文所提方法可以显著优化生成的CAM和伪标签, 语义分割性能在验证集和测试集上分别达到了72.4%和73.3%。相较先前的单阶段方法, SS-EPA误分类的概率更小, 且有更加完整和准确的对象边界, 充分验证了本文方法的有效性。

**关键词** 计算机视觉; 语义分割; 弱监督学习; Transformer; 语义亲和力

Title \* (中英文题目一致) 字体为4号Times New Roman,加粗\* Title

NAME Name-Name<sup>1)</sup> NAME Name<sup>2)</sup> NAME Name-Name<sup>3)</sup> \*字体为5号Times new Roman\*Name

<sup>1)</sup>(Department of \*\*\*\*, University, City ZipCode, China) \*字体为6号Times new Roman\* Depart.Correspond

<sup>2)</sup>(Department of \*\*\*\*, University, City ZipCode)\*中国不写国家名\*

<sup>3)</sup>(Department of \*\*\*\*, University, City ZipCode, country)\*外国写国家名\*

**Abstract** (500英文单词, 内容包含中文摘要的内容).字体为Times new Roman,字号5号\* Abstract

Do not modify the amount of space before and after the artworks. One- or two-column format artworks are preferred. and Tables, create a new break line and paste the resized artworks where desired. Do not modify the amount of space before and after the artworks. One- or two-column format artworks are preferred. All Schemes, Equations, Figures, and Tables should be mentioned in the text consecutively and numbered with Arabic numerals, and appear below where they are mentioned for the first time in the main text. To insert Schemes, Equations, Figures, and Tables, create a new break line and paste the resized artworks where desired. Do not modify the amount of space before and after the artworks. One- or two-column format artworks are preferred. Do not modify the amount of space before and after the artworks. One- or two-column format artworks are preferred. and Tables, create a new break line and paste the resized artworks where desired. Do not modify the amount of space before and after the artworks. One- or two-column format artworks are preferred. All Schemes, Equations, Figures, and Tables should be mentioned in the text consecutively and numbered with Arabic numerals, and appear below where they are mentioned for the first time in the main text.

**Keywords** 中文关键字与英文关键字对应且一致, 不要用英文缩写); key word; key word; key word\*  
\*字体为5号Times new Roman \* Key words

## 1 一级标题\*字体为4号黑体\*标题1

### 对投稿的基本要求:

(1) 研究性论文主体应包括引言(重点论述研究的科学问题、意义、解决思路、价值、贡献等)、相关工作(为与引言部分独立的一个章节)、主要成果论述、关键实现技术、验证(对比实验或理论证明)、结论(结束语)等内容;系统实现或实验应有关键点的详细论述,以便读者能够重复实现论文所述成果。实验应有具体的实验环境设置、全面细致的数据对比分析。[1]

(2) 综述应包括引言、问题与挑战、研究现状分析、未来研究方向、结论等内容。以分析、对比为主,避免堆砌文献或一般性介绍、叙述。

(3) 定理证明、公式推导、大篇幅的数学论述、原始数据,放到论文最后的附录中。

### 稿件提交时的基本要求:

(1) 本模板中要求的各项内容正确齐全,无遗漏;

(2) 语句通顺,无中文、英文语法错误,易于阅读理解,符号使用正确,图、表清晰无误;

(3) 在学术、技术上,论文内容正确无误,各项内容确定。

### 1.1 二级标题\*字体为5号黑体\*标题2

#### 1.1.1 三级标题\*字体为5号宋体\*标题3

\*正文部分,字体为5号宋体\*正文文字

正文文字要求语句通顺,无语法错误,结构合理,条理清楚,不影响审稿人、读者阅读理解全文内容。以下几类问题请作者们特别注意:

- 1) 文章题目应明确反映文章的思想和方法;文字流畅,表述清楚;
- 2) 中文文字、英文表达无语法错误;
- 3) 公式中无符号、表达式的疏漏,没有同一个符号表示两种意思的情况;
- 4) 数学中使用的符号、函数名用斜体;
- 5) 使用的量符合法定计量单位标准;
- 6) 矢量为黑体,标量为白体;
- 7) 变量或表示变化的量用斜体;
- 8) 图表规范,量、线、序无误,位置

正确(图表必须在正文中有所表述后出现,即...如图1所示)(注意纵、横坐标应有坐标名称和刻度值)。

9) 列出的参考文献必须在文中按顺序引用,即参考文献顺序与引用顺序一致,各项信息齐全(格式见参考文献部分);

10) 首次出现的缩写需写明全称,首次出现的符号需作出解释。

11) 图的图例说明、坐标说明全部用中文或量符号。

12) 图应为矢量图。

13) 表中表头文字采用中文。

14) 公式尺寸:

标准: 10.5磅

下标/上标: 5.8磅

次下标/上标: 4.5磅

符号: 16磅

次符号: 10.5磅

15) 组合单位采用标准格式,如:“pJ/bit/m<sup>4</sup>”应为“pJ/(bit·m<sup>4</sup>)”

**定理1.** \*\*\*\*\*.\*定理内容.\*

【“定义”、“假设”、“公理”、“引理”等的排版格式与此相同,详细定理证明、公式可放在附录中】

证明.\*证明过程.\*【“例x”等的排版格式相同】

证毕.

#### 示例图片

(请插入当时做图时的矢量版 如有当时的文件,例如 Visio,origin,matlab, smartdraw,Execl,powerpoint 等各种软件作的图,图字用6号宋体,外文Times new roman, **图中文字尽量用翻译成中文**)

如插入图为截图,必将原图文件如\*.vsd,\*.opj, \*.fig\*.sdr,\*.eps,\*.emf, \*.wmf,\*.ps 等后缀名随修改稿压缩后传过来排版。

图X 图片说明\*字体为小5号,图片应为黑白图,图中的子图要有子图说明\*

**过程X.** 过程名称

\*《计算机学报》的方法过程描述字体为小5号宋体,IF、THEN等伪代码关键词全部用大写字母,变量和

**\*示例表格\***    **\*第1行为表头,表头要有内容\***

附录X.

本文的成果是解决大课题中的哪一部分, 如果涉及 863\973 以及其项目、基金、研究计划, 注意这些项目的英文名称应书写正确。