1. Cover Form

Insert the title here

■ Authors: Jungsun Im, Subin An, Soochahn Lee

■ Corresponding Author: Soochahn Lee

■ Affiliation: Department of Electronics Engineering, Kookmin University

■ Please check your area of study below.

⃞ Broadcast content – voice/audio ⃞ Broadcast content - video/image ⃞ Content management and application

⃞ Broadcasting system ⃞ Broadcasting system - modulation and transmission

⃞ Broadcasting communication convergence ⃞ Broadcasting industry technique

■ Paper classification

⃞ Experimental paper: A paper that implements the content of the paper's proposal and presents its experimental results.

⃞ Production paper: A paper that introduces the results and effects of producing systems, devices, performances, etc. by synthesizing various existing technologies.

⃞ Standard paper: A paper introducing the standard's architecture, use cases, requirements, standard specifications, etc.

⃞ Survey paper: A paper that summarizes and classifies the research contents of existing research or technologies.

⃞ Policy paper: A paper on policy research and proposals for industrial development and innovation in the field of broadcasting media technology.

■ Key words (maximum of five key words in English)

weakly supervised semantic segmentation, affinity enhancement, prototype exploration, self-supervised learning, image-specifc

■ Address including zip code : 02707

■ Phone number: cell phone number: 01024395865

■ E-mail address: ijs\_1016@kookmin.ac.kr

■ Please indicate the type of paper below.

⃞ Special Paper (Please specify the title.)

⃞ Special Paper(Broadcast and Media Technology(영문특집호))

⃞ Regular Paper

⃞ Letter Paper (short, four-page paper; it can be submitted on-line as a J2 Letter Paper.)

2. Requirements for Submitting a Paper

■ All authors of a paper must register for membership and pay the membership fee.

Membership site:

<http://acoms.atit.co.kr:8090/ksbembc/index.jsp?publisher_cd=ksbembc&cid=&cid_year=2013&cid_seq=A&lang=kor>

■ Account number for payment: Citibank Korea, 124-50340-241 (The Korea Institute of Broadcast and Media Engineers)

3. Submission of Paper

■All authors must comply with the instructions regarding how to submit a paper.

- The instructions are provided in our Online Submission System.

- All pages should be in one column and numbered.

- Figures and tables should be included in the body of the paper; their titles and contents should be written in English.

- To comply with the Korean Federation of Science and Technology Societies, the references should be written in English according to the instructions on how to submit a paper. The titles of Korean journals should be translated into English.

- The summary should be approximately 500 Korean characters in order to register it with Scopus. It should not be longer or shorter than the designated length.

■ Before uploading the paper, please check the spelling and word spacing.

- The file to be submitted should be in earlier versions of Hangul 2010 or in Word 2010.

**Author Checklist**

**Title of the paper: AFFINITY ENHANCED IMAGE-SPECIFIC PROTOTYPES FOR**

**WEAKLY SUPERVISED SEMANTIC SEGMENTATION**

**Before you submit your manuscript, please review the following items and mark applicable items with ✔.**

1. General details of preparation

☑ The manuscript includes the original copy of the manuscript, a checklist, an agreement for copyright transfer, and an author introduction.

☑ The manuscript is prepared in Korean or English, and English terms are used when necessary (provided that paper titles, author names, institutions to which authors belong, abstracts, and explanations of figures and tables in English accompany those in Korean).

☑ The manuscript begins with a cover page followed by a title page, an abstract in Korean, an abstract in English, the body (including necessary figures and tables), an appendix, and references in order of precedence.

☑ The body begins with an introduction and ends with conclusions.

☑ The manuscript is prepared using a PC with a Hangeul word processor in a single column, and the length of the manuscript is less than 12 pages (A4 size) including figures and tables (around 1,000 letters per page, double spaced).

☑ In manuscripts in English, only the first letter in each sentence is capitalized (except for proper nouns), and the remaining letters are written in lowercase.

2. Cover page

☑ On the cover page of the manuscript, the article title (including the title in English) and detailed area of the study are provided. Moreover, the submitter’s address, postal code, phone number, and e-mail address must be written.

☑ In the case of names in English, given names are written first followed by family names.

3. Title page

☑ The first page (the page after the cover page) is a title page, and the names of the authors or the organization to which the authors belong are not written; only the title (including the title in English) is written.

4. Abstract and keywords

☑ Both an abstract in Korean (less than 500 letters) and an abstract in English (less than 200 words) are included in the paper.

☑ Five or fewer keywords are written below the abstract in English.

5. Body

☑ Page numbers are included in the paper. Numbers for chapters are written in Roman numerals (Ⅰ, Ⅱ, Ⅲ, Ⅳ, …), and numbers for sections are written in Arabic numerals (1, 2, 3, …).

☑ Figures and tables are presented on the pages in which they are mentioned if possible; however, if they must be presented on a separate page, they are presented at the beginning of the next page. The titles of figures are written below the figures, and the titles of tables are written above the tables in both Korean and English.

☑ With regard to figures or photos, a sufficiently high resolution is maintained so that they can be identified.

6. References

☑ References are written in the order in which they are cited in the body in the manner shown below for both those in Korean and in English, and journal and book names are italicized.

A. Journal: author name, article title, journal name, volume, issue, page, month, year (year followed by month in the case of references in Korean)

B. Book: author name, book name, publishing company name, location, page, year

☑ All references with English descriptions (not the whole paper, just title, author names, etc.) are written in English, and references without English descriptions are written in both Korean and English by translating the Korean into English.

☑ The numbers of references cited in the body shall be written in superscript and in square brackets after cited sentences[1].

7. Ethics

☑ The author(s) of this article has (have) checked the Korean Institute of Broadcast and Engineers’ Internal Regulations for Paper Ethics and will observe the ethics regulations.

Insert the Korean Title here

AFFINITY ENHANCED IMAGE-SPECIFIC PROTOTYPES FOR

WEAKLY SUPERVISED SEMANTIC SEGMENTATION

요 약

Insert the Korean abstract here.

Abstract

Many weakly supervised semantic segmentation methods rely on the pixel-level features extracted from networks trained for image classification. These features can be used to create class activation maps for semantic scores, define pixel affinity as feature similarities, and construct per-class feature prototypes based on feature clustering. This paper proposes a method that enhances previous works by

incorporating affinity-based refinement into the generation of image-specific per-class prototypes, resulting in significantly improved representative strength. These prototypes then lead to improved pseudo-labels, ultimately improving segmentations. Experimental results show significant improvements compared to baseline methods, and are on par with recent state-of-the-art methods. The code is available at <https://github.com/IJS1016/AE_SIPE>

Keywords (insert a maximum of five keywords in English)

weakly supervised semantic segmentation, affinity enhancement, prototype exploration, self-supervised learning, image-specifc

I. Introduction

The goal of weakly supervised semantic segmentation (WSSS) is to learn how to generate pixel-level labels from limited supervision, usually in the form of image-level class labels [2]. The introduction of the Class Activation Map (CAM) [3] was a significant advancement towards achieving this goal, as it provides a means of generating pixel-level per-class scores based on image classification. However, it has been observed that meaningful CAM scores are often only assigned to a selective number of the most discriminative pixels, leading to limitations in directly using CAM as a segmentation solution.

Nonetheless, CAM proves to be a highly efficient technique for utilizing image-level annotations to make pixel-level predictions. It has frequently served as a base upon which multiple methods have been proposed to enhance and optimize the acquisition of pixel-level class probabilities.

One approach is to erase [5] or suppress [6] the more discriminative regions, further mine discriminative pixels. Another approach is to assign the limited discriminative regions as seeds and expand them into full segmentation labels using conventional region growing algorithms [7, 8], based on the similarities of local pixel values. Further methods extended this approach by incorporating pixel adaptive refinement [4], random walks on semantic features [9], or multitask inference of displacement and class boundary [9, 10].

Many recent methods are based on self-supervised learning. A contrastive learning framework, with positive image pairs defined by pairing an image with its linear transform and negative pairs of different images, were applied in [11, 12].

Another approach uses network features to create a per-class feature prototype-based alternative score map, providing supervision to guide the network towards generating consistent features with pixel affinities and image-level class labels. [1].

Combining these methods with others has shown benefits, as seen in recent works [11, 12, 1]. The improved CAM-like score maps generated by these methods are used to enhance pixel affinities and generate pseudo-labels [13, 10], which are used to train a fully supervised semantic segmentation network [14].

In this paper, we propose a method to incorporate pixel-adaptive mask refinement (PAMR) [4] so that pixel affinity is maximized when generating score maps within the self-supervised image-specific prototype exploration (SIPE) method [1]. Experimental results demonstrate that our proposed method provides substantial improvements over the baseline method SIPE. We also propose additional modifications that further improve quantitative results.

II. PROPOSED METHOD

1. Framework

(Enter the text)

\* The text within a figure should be written in English.

Figure 1. Enter the figure title (no period is needed).

Table 1. Enter the table title (no period is needed).

|  |  |  |  |
| --- | --- | --- | --- |
| Content in English | Content in English | Content in English | Content in English |
| Content in English |  |  |  |
| Content in English |  |  |  |
| Content in English |  |  |  |

References

Articles

[1] Names of authors, Title, Title of journal, Volume, Issue, Page, Month, Year (in a Korean document, “Year” is followed by “Month”)

Books

[2]Name(s) of author(s), Title, Name of publisher, Location, Page, Year (Titles of journals and books should be italicized.)

\* References must be written in English.