Self-supervised difference detection for weakly supervised segmentation

Wataru Shimoda  Keiji Yanai  The University of Electro-Communications, Tokyo, Japan

Objective

Self-supervised difference detection (SSDD)

The definition of difference detection

Estimate difference between knowledge and advice

Difference region:

\[ M_{KA}^{d} = \begin{cases} 1 & \text{if } (m_{k}^{A} = m_{d}^{A}) \\ 0 & \text{if } (m_{k}^{A} \neq m_{d}^{A}) \end{cases} \]

Input

Knowledge

Difference detection

SSDD

Self-supervised supervision

Knowledge

Advice

Define the border of the good advice

The high score indicates good advice. But the threshold is ambiguous. It should be different in each sample.

Knowledge

Hard

Advice

Easy

w/o difference

Mask integration

The computation of the confidence scores

\[ w_{n} = d_{e}^{K} - d_{e}^{A} + bias_{n} \]

We integrate \( m^{K} \) and \( m^{A} \) using the confidence scores

\[ m_{p}^{K} = \begin{cases} m_{p}^{K} & \text{if } (w_{n} \geq 0) \\ 0 & \text{if } (w_{n} < 0) \end{cases} \]

We denote this integration process as SSDD module

\[ m^{D} = SSDD(e(x; \theta_{e}), m^{K}, m^{A}; \theta_{d}) \]

SSDD training/inference

- Training
  Train a difference detection model using the difference of the Knowledge and advice.

- Inference
  Integrate a pair of the mask using DD-net outputs.

Static region refinement [A]

Dynamic region refinement [B]

Knowledge: PSA\[1\]

Advice: PSA\[1\] + CRF

Knowledge: Pseudo labels
Advice: The outputs of the segmentation model

We attempt to adapt the re-training scheme\[2\] to SSDD module

References and source codes


Source codes: https://github.com/shimoda-uec/ssdd

How to improve pseudo labels?

Recent weakly-supervised segmentation methods generate pseudo labels in advance and train a segmentation model with them.

We can evaluate the pseudo labels in weakly supervised setting.

Most previous approaches based on heuristic knowledge. There are noise in the mapped results.

We denote the inputs of the mapping functions as knowledge.

We consider the supervision containing the noise as advice.

Problem:
- It is unclear which advice is useful.
- We want to detect good advice.
- Different opinions from the adviser are important.
- Predict the important advice in advance.
- Use the prediction for detecting good advice.

Key idea

We denote the inputs of the mapping functions as knowledge. We consider the supervision containing the noise as advice.

Input (A) ➔ Label (3) Input (B) ➔ Label (4) Input (C) ➔ Label (5) Input (D) ➔ Label (5)

Knowledge

Advice

Assumption

The number of the training sample is related to the degree of the difficulty of the inference in difference detection.

Good advice

Bad advice

Easily predictable difference ➔ good advice

If

If

Good advice

Bad advice

Easily predictable difference ➔ bad advice

w/o difference

Knowledge

Hard

Advice

Easy

w/o difference

The detailed results on PASCAL VOC 2012 val set.

Table

<table>
<thead>
<tr>
<th>Method</th>
<th>Val set</th>
<th>Test set</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCN-MIL (ICLR2015)</td>
<td>25.7</td>
<td>24.9</td>
</tr>
<tr>
<td>DCNN (CVPR2015)</td>
<td>36.1</td>
<td>35.6</td>
</tr>
<tr>
<td>EM-Adapt (ICCV2015)</td>
<td>38.2</td>
<td>39.6</td>
</tr>
<tr>
<td>DCNN (CVPR2015)</td>
<td>44.1</td>
<td>43.1</td>
</tr>
<tr>
<td>BPF (ICCV2016)</td>
<td>47.6</td>
<td>47.0</td>
</tr>
<tr>
<td>NLC (ECCV2016)</td>
<td>54.7</td>
<td>51.7</td>
</tr>
<tr>
<td>TPUR (CVPR2018)</td>
<td>52.8</td>
<td>53.7</td>
</tr>
<tr>
<td>PSL (CVPR2018)</td>
<td>61.7</td>
<td>63.7</td>
</tr>
<tr>
<td>SSDD (ours)</td>
<td>64.9</td>
<td>65.5</td>
</tr>
</tbody>
</table>

The detailed results on PASCAL VOC 2012 val set.

The comparison with WSS methods w/o additional supervision.

The comparison with WSS methods w/ additional supervision.

References and source codes


Source codes: https://github.com/shimoda-uec/ssdd