Self-adversarial Multi-scale Contrastive Learning for Semantic Segmentation of Thermal Facial Images

Jitesh Joshi, Nadia Bianchi-Berthouze, Youngjun Cho*

Department of Computer Science, University College London, UK

Motivation: Challenges in Identifying Regions of Interest

Thermal facial images exhibit:
- Less salient facial features.
- Physiology driven changing spatio-temporal patterns.
- Pose variations.
- Occlusions such as hairs, eyeglasses.
- Changing spatio-temporal patterns due to varying ambient thermal conditions.

Proposed Framework: Self-adversarial Multi-scale Contrastive Learning (SAM-CL)

This work introduces SAM-CL framework to address below challenges in training existing segmentation networks for thermal images:
- Limited size of datasets.
- Unavailability of thermal images acquired in real-world settings such as:
  - Unavailability of pretrained weights that are essential for applying existing state-of-the-art (SOTA) learning strategies, such as supervised contrastive learning.

SAM-CL framework consists of:
- Thermal image augmentation (TiAug) module to generate representations of unconstrained Thermal settings by applying domain-specific transformations to the thermal images acquired in controlled settings.
- Self-adversarial multi-scale contrastive loss function, that utilizes the adversarial attacks made by the TiAug module to achieve intra-class proximity and inter-class separation of the feature representations.

Proposed SAM-CL Loss Function

- Predicted segmentation mask (Y) or logits, ground-truth mask (X*) and class-swapped mask ($Y'$) are passed through the auxiliary network and down-convolved.
- Triplet loss is applied at each layer, resulting in multi-scale contrastive loss function.

Project web-page: https://github.com/PhysiologicAILab
* Corresponding author: youngjun.cho@ucl.ac.uk

Raw Text Content

Quantitative Performance Analysis

<table>
<thead>
<tr>
<th>Segmentation Network</th>
<th>Learning Strategy (Loss Function)</th>
<th>mIoU (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blind segmentation (BCE)</td>
<td>78.69</td>
<td></td>
</tr>
<tr>
<td>Blind segmentation (BCE)</td>
<td>80.11</td>
<td></td>
</tr>
<tr>
<td>SAM (DeepSL)</td>
<td>80.96</td>
<td></td>
</tr>
<tr>
<td>SAM (DeepSL)</td>
<td>76.97</td>
<td></td>
</tr>
<tr>
<td>SAM (DeepSL)</td>
<td>81.33</td>
<td></td>
</tr>
<tr>
<td>SAM-CL (Ours)</td>
<td>81.58</td>
<td></td>
</tr>
<tr>
<td>SAM-CL (Ours)</td>
<td>82.45</td>
<td></td>
</tr>
<tr>
<td>UBComfort</td>
<td>78.69</td>
<td></td>
</tr>
<tr>
<td>UBComfort</td>
<td>80.11</td>
<td></td>
</tr>
<tr>
<td>UBComfort</td>
<td>81.33</td>
<td></td>
</tr>
<tr>
<td>UBComfort</td>
<td>82.45</td>
<td></td>
</tr>
</tbody>
</table>

Ablation Study

- Consistent performance gains are observed for all the segmentation networks when thermal images are augmented using the TiAug module and the SAM-CL loss function is used for optimization.

Qualitative Performance Analysis

- Different Datasets
- Different thermal images
- Different thermal ambience

Code and demo are available at: https://github.com/PhysiologicAILab/SAM-CL

SAM-CL framework is needed only for training the segmentation network, resulting in no computational overhead during inference.