Localization in the Crowd with Topological Constraints
Shahira Aboussamra, Minh Hoai, Dimitris Samaras, Chao Chen
Stony Brook University, USA

Introduction

Crowd Localization Problem: Finding the location of each person in a crowded scene.

Ground truth: a single dot on each head.

Challenges in Crowd Localization
1. Perspective, occlusion, and cluttering.
2. The features of dots are not specific.
3. Difficult to prevent spatial semantic errors.

Contributions:
1. Overcome these challenges by introducing topological constraints in the training phase.
2. Propose persistence loss to enforce topological constraints.
3. Achieve high quality localization that is useful for crowd counting and spatial analysis.

Method: TopoCount

• Formulate crowd localization as a structured prediction problem.
• Each component in the binary prediction represents one dot.

Topological Constraint for Crowd Localization
Within any local patch, the number of connected components in the prediction equals to the number of ground truth dots.

Persistence Loss $L_{Pers}$

- To enforce topological constraints.
- Consider likelihood map as a terrain function $f_i$.
- Each mode $f$ corresponds to a possible dot prediction.
- Persistence Loss captures all modes and chooses to suppress or enhance.

Persistence Constraint $\delta$

- A patch $\delta$ with $c$ ground truth dots:
  - Persistence loss reinforces the total saliency of the top $\alpha$ modes of $f$ and suppresses the saliency of the rest.

$\text{Saliency/Persistence of a mode } m_i = f(m_i) - f(x) = \sum_{m \in M_i} \text{Pers}(m) + \sum_{m \notin M_i} \text{Pers}(m)$

Training Loss $L = L_{DICE} + \lambda_{pers} L_{Pers}$

Model Architecture

Persistence Loss $L_{Pers}$

- To enforce topological constraints.
- Consider likelihood map as a terrain function $f_i$.
- Each mode $f$ corresponds to a possible dot prediction.
- Persistence Loss captures all modes and chooses to suppress or enhance.

Evaluation

1. Localized Counting

Grid Average Mean Absolute Error (GMAE) divide the image into $d^3$ non-overlapping cells and computes the mean MAE over all grid cells.

<table>
<thead>
<tr>
<th>Method</th>
<th>ShanghaiTech A</th>
<th>ShanghaiTech B</th>
<th>SCI (2020)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>6.2</td>
<td>5.8</td>
<td>5.4</td>
</tr>
<tr>
<td>RMSE</td>
<td>8.6</td>
<td>7.0</td>
<td>6.5</td>
</tr>
<tr>
<td>F1/Precision</td>
<td>58.7 / 32.5 / 132.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dice</td>
<td>52.5 / 55.8 / 49.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOU</td>
<td>57.0 / 56.6 / 54.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TopoCount</td>
<td>68.1 / 63.1 / 60.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Ablation Study: Loss function

Compare training with and without persistence loss.
- Dice loss only better than binary cross entropy (BCE) loss only.
- Dice loss + Persistence loss gives lower error.

3. Dot Matching Accuracy

Compare precision, recall, and F-score on NWPU localization challenge.

Integration with Density-map Counting

Input Image

Baseline density map count err = -323

TopoCount density map count err = -47

code: https://github.com/TopoXLab/TopoCount
email: shahira.aboussamra@stonybrook.edu