Semantic Graph Convolutional Networks for 3D Human Pose Regression

Long Zhao¹ Xi Peng² Yu Tian¹ Mubbasir Kapadia¹ Dimitris Metaxas¹
¹ Rutgers University ² Binghamton University

Highlights
- We study the problem of learning Graph Convolutional Networks (GCNs) for regression.
- We propose Semantic Graph Convolutional Networks (SemGCNs) to address two limitations of GCNs.
- We apply SemGCNs to 3D human pose regression: Both 2D and 3D poses are able to be naturally represented by a canonical skeleton, i.e., a graph.
- The proposed SemGCN outperforms state of the art while using 90% fewer parameters.

Semantic Graph Convolutional Networks
- **Semantic Graph Convolutional Layers** learn local relationships between nodes.
- **Non-local Blocks** capture global and long-range relationships among nodes.
- Our proposed **Semantic Graph Convolutional Networks** interleave Semantic Graph Convolutional Layers and Non-local Blocks to capture local and global semantic relations of nodes in a graph.

3D Human Pose Regression
Motivation: We argue that image content is able to offer important cues for solving ambiguous cases of 3D human poses. Therefore, we treat image content as an additional constraint for the human skeleton.
- The pre-trained **2D pose estimation network** encodes the perceptual features of the input image.
- **Perceptual Feature Pooling** is proposed to extract intermediate features from the 2D pose detector.
- Perceptual features are concatenated with the 2D coordinates and fed into the proposed SemGCN.

Results
- **Ablation Study:** Training Curves and Testing Errors
- **Evaluation on 3D Human Pose Regression**

<table>
<thead>
<tr>
<th>Methods</th>
<th># of params</th>
<th>MPIPE (GT)</th>
<th>Methods</th>
<th>MPIPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>iGCN / GAT</td>
<td>0.16M</td>
<td>82.9</td>
<td>Martinez et al. (ICCV' 17)</td>
<td>62.9</td>
</tr>
<tr>
<td>ST-GCN</td>
<td>0.27M</td>
<td>57.4</td>
<td>Yang et al. (CVPR' 18)</td>
<td>58.6</td>
</tr>
<tr>
<td>Martinez et al.</td>
<td>4.2M</td>
<td>45.5</td>
<td>Ours (HG)</td>
<td>60.8</td>
</tr>
<tr>
<td>Ours</td>
<td>0.43M</td>
<td>43.8</td>
<td>Ours (RN w/ FF)</td>
<td>57.6</td>
</tr>
</tbody>
</table>

- **Qualitative Results**

![Image of 3D Human Pose Regression Diagram](image-url)