Graph Neural Blocks on Segmentation

Context

Semantic Segmentation is one of the key steps in computer vision applications.

(a) X-ray  (b) Agriculture  (c) Remote Sensing  (d) Autonomous Driving

Semantic segmentation problems:
- Low-resolution in output heatmaps (solved)
- Loss of spatial precision (remain)

(a) DUNet  (b) OCNet  (c) HRNet+OCR

Table: Loss of spatial precision, generally displayed on the segmented objects' boundaries

Quantitative Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Sky</th>
<th>Building</th>
<th>Road</th>
<th>Parking</th>
<th>Sky+Building</th>
<th>Vegetation</th>
<th>Pedestrian</th>
<th>Sign</th>
<th>Person</th>
<th>Cyclist</th>
<th>mIoU</th>
</tr>
</thead>
<tbody>
<tr>
<td>ParseNet</td>
<td>90.76</td>
<td>85.20</td>
<td>92.01</td>
<td>63.49</td>
<td>39.00</td>
<td>48.82</td>
<td>89.20</td>
<td>61.90</td>
<td>63.91</td>
<td>52.73</td>
<td>71.20</td>
</tr>
<tr>
<td>ESPNet</td>
<td>91.79</td>
<td>86.36</td>
<td>95.73</td>
<td>71.84</td>
<td>48.52</td>
<td>88.44</td>
<td>49.06</td>
<td>87.29</td>
<td>54.60</td>
<td>61.83</td>
<td>75.51</td>
</tr>
<tr>
<td>FC-DenseNet67</td>
<td>92.19</td>
<td>86.77</td>
<td>96.60</td>
<td>75.40</td>
<td>41.55</td>
<td>88.07</td>
<td>52.92</td>
<td>87.09</td>
<td>63.89</td>
<td>60.48</td>
<td>72.54</td>
</tr>
<tr>
<td>BiSeNet</td>
<td>91.64</td>
<td>87.42</td>
<td>96.48</td>
<td>75.41</td>
<td>44.05</td>
<td>89.07</td>
<td>39.50</td>
<td>89.34</td>
<td>58.63</td>
<td>66.81</td>
<td>72.86</td>
</tr>
<tr>
<td>ENet</td>
<td>91.61</td>
<td>87.48</td>
<td>96.44</td>
<td>75.34</td>
<td>44.44</td>
<td>89.23</td>
<td>43.14</td>
<td>89.24</td>
<td>56.04</td>
<td>65.13</td>
<td>73.26</td>
</tr>
<tr>
<td>ICNet</td>
<td>92.03</td>
<td>88.46</td>
<td>96.61</td>
<td>77.22</td>
<td>42.59</td>
<td>89.48</td>
<td>48.27</td>
<td>90.74</td>
<td>58.71</td>
<td>66.18</td>
<td>74.23</td>
</tr>
<tr>
<td>DeepLab v3</td>
<td>92.22</td>
<td>89.02</td>
<td>96.74</td>
<td>78.13</td>
<td>41.00</td>
<td>90.81</td>
<td>49.74</td>
<td>91.02</td>
<td>64.48</td>
<td>66.52</td>
<td>75.21</td>
</tr>
<tr>
<td>PSPNet</td>
<td>91.94</td>
<td>89.93</td>
<td>96.94</td>
<td>78.37</td>
<td>53.64</td>
<td>90.19</td>
<td>43.47</td>
<td>92.12</td>
<td>64.40</td>
<td>70.71</td>
<td>76.61</td>
</tr>
<tr>
<td>DA-Net</td>
<td>92.25</td>
<td>90.26</td>
<td>97.25</td>
<td>79.95</td>
<td>51.33</td>
<td>90.60</td>
<td>45.20</td>
<td>92.50</td>
<td>68.38</td>
<td>71.47</td>
<td>77.13</td>
</tr>
<tr>
<td>AdapNet++</td>
<td>93.07</td>
<td>89.46</td>
<td>96.06</td>
<td>80.03</td>
<td>49.46</td>
<td>90.58</td>
<td>52.10</td>
<td>92.22</td>
<td>66.26</td>
<td>72.88</td>
<td>77.61</td>
</tr>
<tr>
<td>CCNet</td>
<td>90.97</td>
<td>89.01</td>
<td>96.59</td>
<td>77.36</td>
<td>42.49</td>
<td>91.36</td>
<td>58.24</td>
<td>91.22</td>
<td>71.18</td>
<td>74.80</td>
<td>77.62</td>
</tr>
<tr>
<td>OCNet</td>
<td>92.72</td>
<td>90.73</td>
<td>97.39</td>
<td>80.80</td>
<td>54.58</td>
<td>90.86</td>
<td>45.60</td>
<td>92.64</td>
<td>67.35</td>
<td>71.81</td>
<td>77.86</td>
</tr>
<tr>
<td>DUNet</td>
<td>93.34</td>
<td>91.05</td>
<td>97.28</td>
<td>80.18</td>
<td>55.15</td>
<td>93.19</td>
<td>53.70</td>
<td>92.89</td>
<td>68.33</td>
<td>73.33</td>
<td>79.29</td>
</tr>
<tr>
<td>GNNBlock* (Ours)</td>
<td>93.10</td>
<td>90.86</td>
<td>96.96</td>
<td>79.13</td>
<td>48.61</td>
<td>90.82</td>
<td>59.05</td>
<td>92.58</td>
<td>72.86</td>
<td>74.96</td>
<td>79.11</td>
</tr>
<tr>
<td>HRNet</td>
<td>94.56</td>
<td>90.98</td>
<td>97.48</td>
<td>82.46</td>
<td>50.27</td>
<td>92.35</td>
<td>61.57</td>
<td>93.96</td>
<td>73.14</td>
<td>78.55</td>
<td>80.98</td>
</tr>
</tbody>
</table>

Table: IoU results on Cityscapes validation set for semantic segmentation, using 11 classes and with resize of 384 x 768.

Methodology

Graph convolutional network, creating and deleting edges and updating features values.

\[
H^{(l+1)} = \sigma_l [\mathcal{A}^{(l)} H^{(l)} W^{(l)}], \quad \tau [A^{(l)}] = [D^{(l)}]^{-\frac{1}{2}} [A^{(l)} + I] (D^{(l)})^{-\frac{1}{2}}, \quad D^{(l)} = D^{(l)} + I_{s_{\tau}} \]

GNNBlock

\[
X' = g(\alpha_l (A^{(l)} f(X) W^{(l)})) \]

Loss Functions

\[
\mathcal{L}_{cross_{-}ss} = \frac{1}{N_{i=1}} \sum \alpha_i \log P(s = s_i | X; \phi), \quad (5) \\
\mathcal{L}_{iou_{-}ss} = 1 - \frac{1}{N_{i=1}} \sum I_{s_i \cap \hat{s}_i} I_{s_i \cup \hat{s}_i} \quad (6) \\
\mathcal{L}_{ss} = \psi_1 \mathcal{L}_{cross_{-}ss} + \psi_2 \mathcal{L}_{iou_{-}ss} \quad (7)
\]

Qualitative Results

Table: Comparison results on validation set from Cityscapes dataset.