Recurrent GlimpseNet

CS231N Final Presentation

Isaac Supeene and Kaiyue Wang
Spring Quarter 2022
Contents

- Problem Statement & Related Work
- Summary of Findings
- Model Experiments
- Model Inspections
Related Work

GlimpseNet - Hang et al.
- Previous CS231N project
- Trained the glimpse generator on ground truth salience labels

MultiGlimpse Network
- Recurrent architecture
- End-to-end training

Tan et al. Illustration of Recurrent Downsampled Attention
Related Work

**GlimpseNet** - Hang et al.
- Previous CS231N project
- Trained the glimpse generator on ground truth salience labels

**MultiGlimpse Network**
- Recurrent architecture
- End-to-end training
- **Slow to train**

*Tan et al.*, Illustration of Recurrent Downsampled Attention
Related Work
Contents

- Problem Statement & Related Work
- Summary of Findings
- Model Experiments
- Model Inspections
Findings

- Our experiments replicated the training instability of the Localization Network reported by Tan et al.
- Model inspections show that the gradient signal to the Localization Network is extremely unstable.
- Model experiments show no benefit to increasing the capacity of the Localization Network, consistent with the finding that its learning signal is unreliable.
Contents

- Problem Statement & Related Work
- Summary of Findings
- Model Experiments
- Model Inspections
Model Experiments

- Even very significant increases in Localization Network capacity do not improve glimpse quality.
- Substantially increasing the image scaling factor does not encourage the network to learn useful glimpses.
- Minor performance improvement by using 2d positional encoding to describe glimpse location.

<table>
<thead>
<tr>
<th>Method</th>
<th>ImageNet 100 Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>baseline</td>
<td>83.99</td>
</tr>
<tr>
<td>2d-spatial-clue</td>
<td>84.16</td>
</tr>
<tr>
<td>2d-spatial-loc</td>
<td>83.81</td>
</tr>
<tr>
<td>2d-spatial-loc-s7</td>
<td>62.00</td>
</tr>
<tr>
<td>2d-spatial-loc-s7-no-aux</td>
<td>67.45</td>
</tr>
<tr>
<td>multihead-attention</td>
<td>83.72</td>
</tr>
<tr>
<td>pretrained</td>
<td>85.14</td>
</tr>
<tr>
<td>pretrained-frozen</td>
<td>84.75</td>
</tr>
</tbody>
</table>
Contents

▪ Problem Statement & Related Work
▪ Summary of Findings
▪ Model Experiments
▪ Model Inspections
Model Inspections

- Glimpse parameter sweeps clearly demonstrate the unstable gradient signal to the Localization Network.
Recurrent GlimpseNet

CS231N Final Presentation

Isaac Supeene and Kaiyue Wang
Spring Quarter 2022