Action Recognition Experiments

By Kevin Yu
Hardware Setup

Ubuntu 14.04
i7-6700HQ, 8 core
16Gb RAM
CPU only
NVIDIA Tesla K80 GPU
Which experiment to pick

- Two-stream convolutional networks for action recognition in videos?
- Large-scale video classification with convolutional neural networks?
- Learning Spatiotemporal Features with 3D Convolutional Networks?
Installation Guides

Main Website:

https://github.com/facebook/C3D

Installation Instructions:

https://gist.github.com/arundasan91/b432cb011d1c45b65222d0fac5f9232c
http://vra.github.io/2016/03/03/c3d-use/
## Legitimacy of C3D

<table>
<thead>
<tr>
<th>Dataset Task</th>
<th>Sport1M action recognition</th>
<th>UCF101 action recognition</th>
<th>ASLAN action similarity labeling</th>
<th>YUPENN scene classification</th>
<th>UMD scene classification</th>
<th>Object object recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>[29]</td>
<td>[39][25]</td>
<td>[31]</td>
<td>[9]</td>
<td>[9]</td>
<td>[32]</td>
</tr>
<tr>
<td>Result</td>
<td>90.8</td>
<td>75.8 (89.1)</td>
<td>68.7</td>
<td>96.2</td>
<td>77.7</td>
<td>12.0</td>
</tr>
<tr>
<td>C3D</td>
<td>85.2</td>
<td>85.2 (90.4)</td>
<td>78.3</td>
<td>98.1</td>
<td>87.7</td>
<td>22.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Number of Nets</th>
<th>Clip hit@1</th>
<th>Video hit@1</th>
<th>Video hit@5</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeepVideo’s Single-Frame + Multires [18]</td>
<td>3 nets</td>
<td>42.4</td>
<td>60.0</td>
<td>78.5</td>
</tr>
<tr>
<td>DeepVideo’s Slow Fusion [18]</td>
<td>1 net</td>
<td>41.9</td>
<td>60.9</td>
<td>80.2</td>
</tr>
<tr>
<td>Convolution pooling on 120-frame clips [29]</td>
<td>3 net</td>
<td>70.8*</td>
<td>72.4</td>
<td>90.8</td>
</tr>
<tr>
<td>C3D (trained from scratch)</td>
<td>1 net</td>
<td>44.9</td>
<td>60.0</td>
<td>84.4</td>
</tr>
<tr>
<td>C3D (fine-tuned from I380K pre-trained model)</td>
<td>1 net</td>
<td>46.1</td>
<td>61.1</td>
<td>85.2</td>
</tr>
</tbody>
</table>
Example Videos

1. Badminton: 1.00
2. Seepak Takraw: 0.00

1. Pehlwani: 1.00
2. Wrestling: 0.00
Action Classification

Tested 150 clips
Action Classification Successful

True: Still Rings
Predicted: Still Rings

True: BandMarching
Predicted: BandMarching

True: Punch
Predicted: Punch
Action Classification Unsuccessful

True: Kayaking
Predicted: Surfing

True: PlayingSitar
Predicted: ShavingBeard

True: HammerThrow
Predicted: PlayingSitar