POLYNET
A PURSUIT OF STRUCTURAL DIVERSITY WITHIN A NETWORK

Dahua Lin, on behalf of the CU-DeepLink team
Single Network Accuracy

- ResNet-152: 5.71% (single-crop), 4.49% (multi-crop)
- Inception-ResNet: 4.90% (single-crop), 3.70% (multi-crop)
- PolyNet: 4.25% (single-crop), 3.45% (multi-crop)
Going Deeper?

Error Rate vs. # Layers for ResNet

- Error Rate decreases with increasing # Layers.
- There is a point of diminishing return after a certain number of layers.

Graph shows:
- Error Rate on the y-axis (ranging from 4 to 7).
- # Layers on the x-axis (ranging from 50 to 500).
- The line indicates a trend that decreases gradually, with a notable plateau after a certain layer count.
Going Wider?

Wide ResNet

Cost vs. Widening factor (k)
Dimensions to explore

- **Depth**
  
  Diminishing return &
  Increased training difficulty

- **Width**
  
  Quadratic growth in both
  computational cost &
  memory demand.

- Any other dimensions to explore?
Clues from the History

What do they have in common?

Inception
A combination of complementary paths — the most successful design of CNN modules.

ResNet
Veit et al. showed that a ResNet is an exponential ensemble of relatively shallow paths.

Ensemble
Ensemble usually gives you a considerable gain no matter how powerful individual models are.
PolyInception

\[ y = (I + F + G \circ F)(x) \]
PolyNet

<table>
<thead>
<tr>
<th></th>
<th>#layers</th>
<th>param (MB)</th>
<th>ms/iter</th>
<th>single-crop error</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR-v2 (5-10-5)</td>
<td>132</td>
<td>135</td>
<td>880</td>
<td>5.05</td>
</tr>
<tr>
<td>IR-v2 (10-20-10)</td>
<td>242</td>
<td>237</td>
<td>1380</td>
<td>4.83</td>
</tr>
<tr>
<td>IR-v2 (20-56-20)</td>
<td>655</td>
<td>531</td>
<td>1957</td>
<td>4.50</td>
</tr>
<tr>
<td>PolyNet</td>
<td>537</td>
<td>365</td>
<td>1792</td>
<td>4.25</td>
</tr>
</tbody>
</table>

![Diagram of PolyNet](image)
Overview of CLS Results

- multi-crop (PolyNet G5) 4.25%
- multi-crop (selective pooling) 3.45%
- ensemble (weighted comb) 2.93%

This is only the first step …
Parrots

A new deep learning framework developed by us (from scratch)

**Efficient**
Highly efficient scheduling & optimal memory reuse

**Scalable**
Multi-node & multi-GPU support. Scalable to 64 GPUs and more …

**Extensible**
Highly extensible modular design based on a novel notion of VM

Will be open sourced …
Thank You

CU-DeepLink
Xingcheng Zhang
Zhizhong Li
Yang Shuo
Yuanjun Xiong
Yubin Deng
Xiaoxiao Li
Kai Chen
Yingrui Wang
Chen Huang
Tong Xiao
Wanshen Feng
Xinyu Pan
Yunxiang Ge
Hang Song
Yujun Shen
Boyang Deng
Ruohui Wang

Shengen Yan
Wenzhi Liu
Chen Change Loy
Dahua Lin

Multimedia Lab, CUHK
SenseTime. Inc