Stage-Wise Neural Architecture Search

Artur Jordão, Maiko Lie, Fernando Akio and William Robson Schwartz
Smart Sense Laboratory, Computer Science Department
Federal University of Minas Gerais, Brazil
Introduction

• Neural Architecture Search (NAS)
  – Current NAS strategies analyze a large set of possible candidate architectures

• Motivated by this, we propose a NAS approach to efficiently design accurate and low-cost convolutional architectures
Proposed NAS
Problem Definition

- Modern architectures are composed of stages
  - Each stage consists of \( b \) modules
Problem Definition

• Proposed NAS
Proposed Approach

Input Architecture

Importance 0.87

Importance 0.72

Importance 0.92

Temporary Architecture

Importance 0.93

Importance 0.68

Importance 0.94

Candidate Architecture 1
Experiments
Comparison with human-designed architectures

- CIFAR-10
  - * indicates human-designed architectures

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Depth</th>
<th>Param. ↓ (Million)</th>
<th>FLOP↓ (Million)</th>
<th>Accuracy↑</th>
</tr>
</thead>
<tbody>
<tr>
<td>ResNet44*</td>
<td>44</td>
<td>0.66</td>
<td>97</td>
<td>92.83</td>
</tr>
<tr>
<td>Ours (it=1)</td>
<td>43</td>
<td>0.60</td>
<td>92</td>
<td>93.38</td>
</tr>
<tr>
<td>ResNet56*</td>
<td>56</td>
<td>0.86</td>
<td>125</td>
<td>93.03</td>
</tr>
<tr>
<td>Ours (it=3)</td>
<td>59</td>
<td>0.69</td>
<td>130</td>
<td>93.36</td>
</tr>
<tr>
<td>ResNet110*</td>
<td>110</td>
<td>1.7</td>
<td>253</td>
<td>93.57</td>
</tr>
<tr>
<td>Ours (i=5)</td>
<td>67</td>
<td>0.88</td>
<td>149</td>
<td>94.27</td>
</tr>
</tbody>
</table>
Comparison with state-of-the-art NAS

- CIFAR-10

<table>
<thead>
<tr>
<th>Model</th>
<th>Evaluated Models</th>
<th>GPUs ↓</th>
<th>Param. ↓ (Million)</th>
<th>Accuracy ↑</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zoph et al. [2018]</td>
<td>20,000</td>
<td>800</td>
<td>2.5</td>
<td>94.51</td>
</tr>
<tr>
<td>Real et al. [2017]</td>
<td>1,000</td>
<td>250</td>
<td>5.4</td>
<td>94.60</td>
</tr>
<tr>
<td>Dong and Yang [2019]</td>
<td>240</td>
<td>1</td>
<td>2.6</td>
<td>96.25</td>
</tr>
<tr>
<td>Jin et al. [2019]</td>
<td>60</td>
<td>1</td>
<td>---</td>
<td>88.56</td>
</tr>
<tr>
<td>Ours (it=5)</td>
<td>11</td>
<td>1</td>
<td>2.3</td>
<td>94.74</td>
</tr>
</tbody>
</table>

Zoph et al. (2018). *Learning transferable architectures for scalable image recognition*. In CVPR.
Real et al. (2017). *Large-scale evolution of image classifiers*. In ICML.
Dong and Yang et al. (2019). *Searching for a robust neural architecture in four GPU hours*. In CVPR.
Jin et al. (2019). *Auto-keras: An efficient neural architecture search system*. In SIGKDD.
Comparison with state-of-the-art NAS

- CIFAR-10

Zoph et al. (2018). *Learning transferable architectures for scalable image recognition*. In CVPR.
Baker et al. (2017). *Designing neural network architectures using reinforcement learning*. In ICML.
Real et al. (2017). Large-scale evolution of image classifiers. In ICLR.
Brock et al. (2018). *SMASH: one-shot model architecture search through hypernetworks*. In ICLR.
Conclusions

• We demonstrate that it is possible to design high-performance convolutional architectures by inserting layers based on their importance.

• Compared to NAS strategies, our method is extremely more efficient, as it evaluates one order of magnitude fewer models and discovers architectures on par with the state of the art.

• Code: github.com/arturjordao/StageWiseArchitectureSearch
Acknowledgments