Progressive Neural Architecture Search

Zhang, Liqiang

May 19,2018

1 SMBO

The author propose a new method for learning the structure of convolutional neural networks that is more e?cient than recent state-of-the-art methods based on reinforcement learning and evolutionary algorithms. The approach uses a sequential model-based optimization (SMBO) strategy, in which we search for structures in order of increasing complexity [1], while simultaneously learning a surrogate model to guide the search through structure space. This structures achieve state of the art classi?cation accuracies on CIFAR-10 and ImageNet.

The approach that the author propose t is able to learn a CNN which matches previous state of the art in terms of accuracy, while requiring 5 times fewer model evaluations during the architecture search.

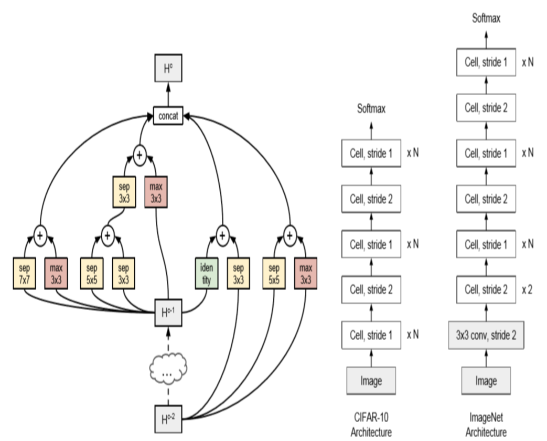

Figure 1: Left: The best cell structure found by our Progressive Neural Architecture Search, consisting of 5 blocks. Right: We employ a similar strategy as [2] when constructing CNNs from cells on CIFAR-10 and ImageNet. Note that we learn a single cell type instead of distinguishing between Normal and Reduction cell.

2 Architecture Search Space

To evaluate a cell, it have to convert it into a CNN. To do this, the author stack a prede?ned number of copies of the basic cell (with the same structure, but untied weights), using either stride 1 or stride 2, as shown in Fig. 1 (right). The number of stride-1 cells between stride-2 cells is then adjusted accordingly with up to N number of repeats. At the top of the network, we use global average pooling, followed by a softmax classi?cation layer. They then train the stacked model on the relevant dataset.

The overall CNN construction process is identical to [3], except the author only use one cell type (we do not distinguish between Normal and Reduction cells, but instead emulate a Reduction cell by using a Normal cell with stride 2), and the cell search space is slightly smaller (since we use fewer operators and combiners).

3 Search E?ciency

The results are shown in Table. 1. We see that PNAS is up to 5 times faster in terms of the number of models it trains and evaluates.

B	Top	Accuracy	PNAS	NAS	Speedup(models)	Speedup(ex.)
5	1	0.9183	1160	5808	5.0	8.2
5	5	0.9161	1160	4100	3.5	6.8
5	25	0.9136	1160	3654	3.2	6.4

Table 1: Relative e?ciency of PNAS (using MLP-ensemble predictor) and NAS under the same search space.

References

[1] Alex Krizhevsky. Learning multiple layers of features from tiny images. 2009.

[2] Kenneth O Stanley and Risto Miikkulainen. Evolving neural networks through augmenting topologies. *Evolutionary Computation*, 10(2):99–127, 2014.

[3] Ronald J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3-4):229–256, 1992.