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

CAPES UFMG CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico senselab

Introduction

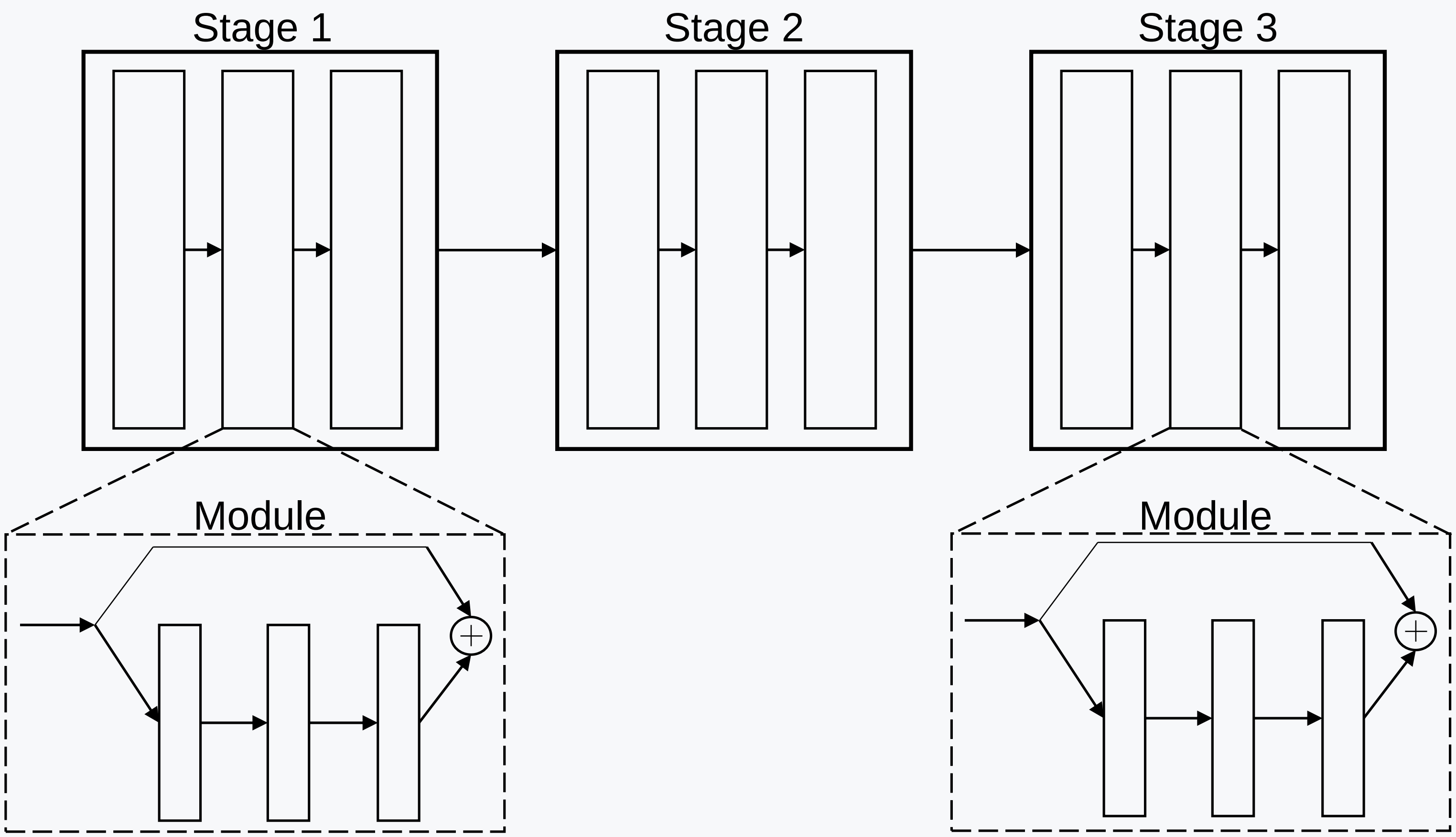
- Neural Architecture Search (NAS)
  - Current NAS strategies analyze a large set of possible candidate architectures

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

Proposed NAS

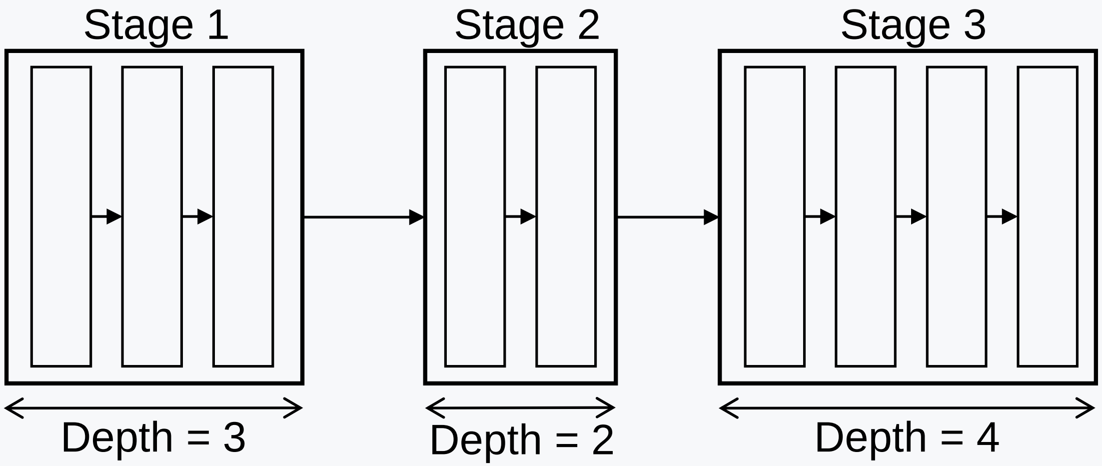
Problem Definition

- Modern architectures are composed of stages
  - Each stage consists of b modules



Problem Definition

- Proposed NAS



Proposed Approach

Experiments

Comparison with human-designed architectures

- CIFAR-10
  - * indicates human-designed architectures

Architecture	Depth	Param. ↓ (Million)	FLOP↓ (Million)	Accuracy↑
ResNet44*	44	0.66	97	92.83
Ours (it=1)	**43**	**0.60**	**92**	**93.38**
ResNet56*	**56**	0.86	**125**	93.03
Ours (it=3)	59	**0.69**	130	**93.36**
ResNet110*	110	1.7	253	93.57
Ours (i=5)	**67**	**0.88**	**149**	**94.27**

Comparison with state-of-the-art NAS

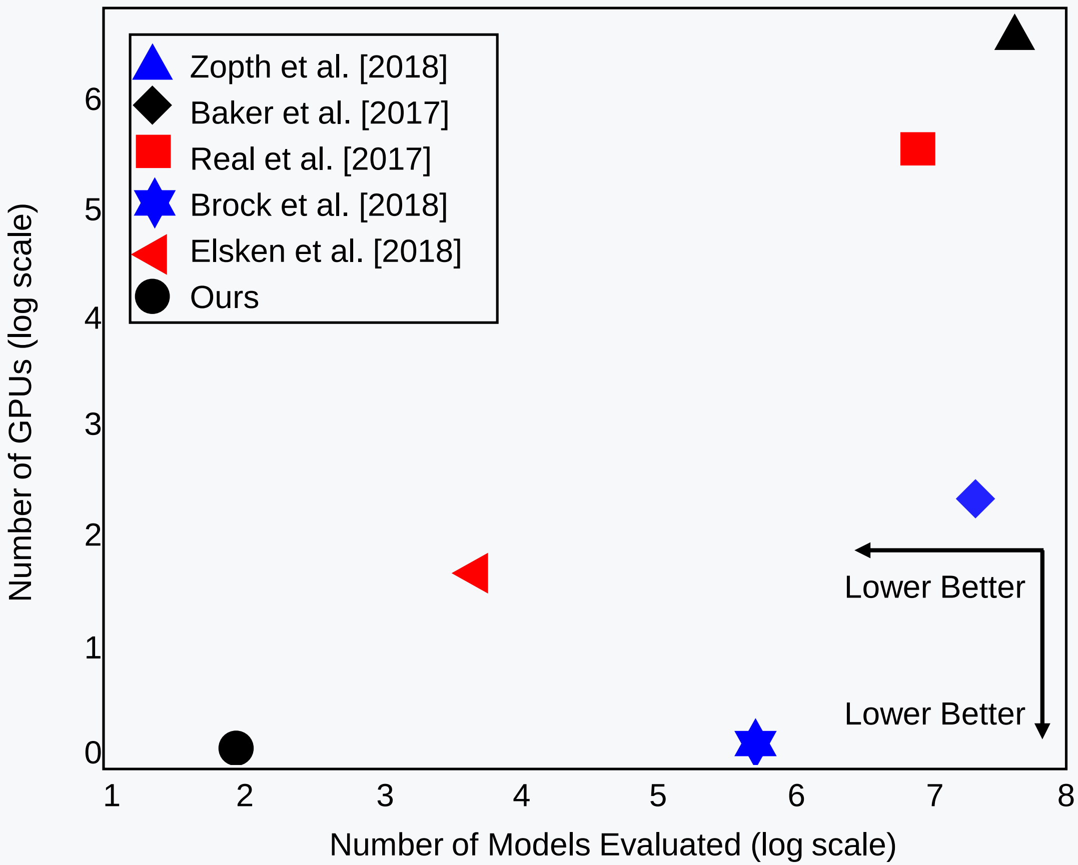
- CIFAR-10

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

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.
Elsken et al. (2018). Simple and efficient architecture search for convolutional neural networks. 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

UFMG

CNPq

Conselho Nacional de Desenvolvimento Científico e Tecnológico

CAPES

senselab