NAS-Bench-101: Towards Reproducible Neural Architecture Search

Chris Ying[*1], Aaron Klein[*2],
Esteban Real[1], Eric Christiansen[1], Kevin Murphy[1], Frank Hutter[2]

[1]Google Brain, [2]University of Freiburg
[*]equal contribution

ICML 2019

Motivation

Neural architecture search (NAS) methods are notoriously difficult to reproduce and compare:

1. **Different search spaces** and **training procedures**
    - Implicit biases imposed by search space and training, different NAS methods optimized for different setups
    - Cannot separate benefit of NAS from the careful design of the search space and training procedures

Motivation

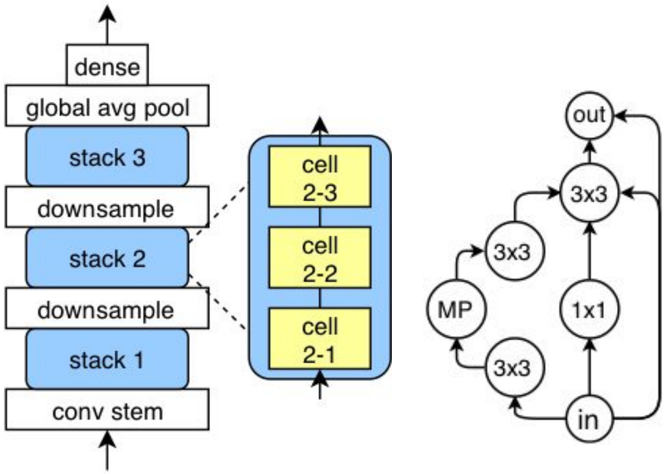
Neural architecture search (NAS) methods are notoriously difficult to reproduce and compare:

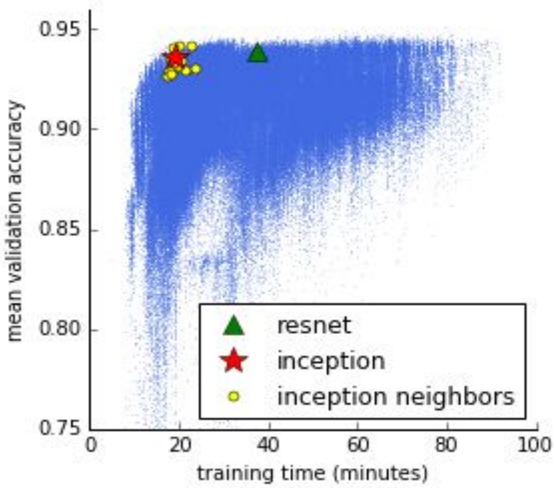
1. **Different search spaces** and **training procedures**
    - Implicit biases imposed by search space and training, different NAS methods optimized for different setups
    - Cannot separate benefit of NAS from the careful design of the search space and training procedures
2. **Compute cost** limits number of trials and makes methods inaccessible to most researchers

NAS-Bench-101

- General search space of directed acyclic graphs for cell-based NAS methods
- Exhaustively trained & evaluated all models on CIFAR-10 to create a queryable dataset

~423K <u>unique</u> cells
* 4 epoch budgets
* 3 repeats
= **~5M <u>total</u> models** trained





NAS-Bench-101

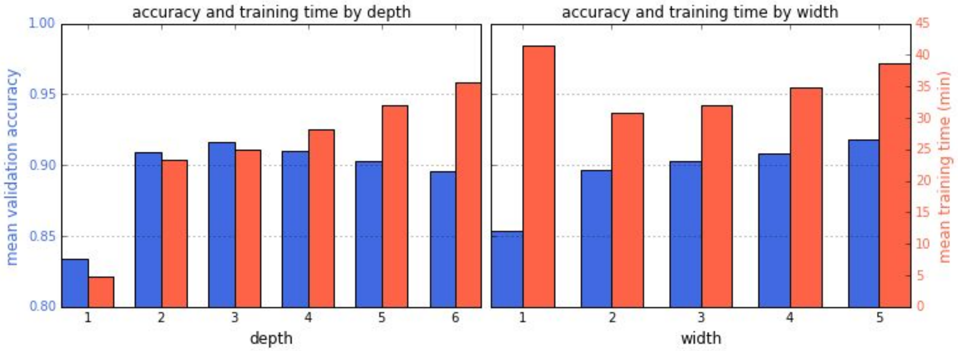
Enables:

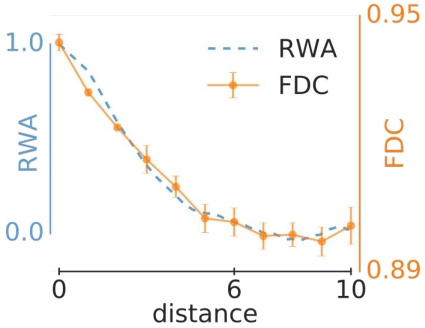
1) Studying the landscape of a neural architecture search space as a discrete optimization space

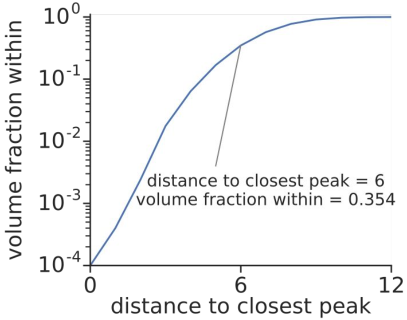
2) Efficient benchmarking of NAS methods by separating the process of **searching for models** (cheap) from **evaluating the models** (expensive)

Aggregate Analysis of Search Space

- Search space exhibits *locality*: similar architectures often have similar performance
- Randomly selecting top model is extremely unlikely, but many models within short edit-distance away

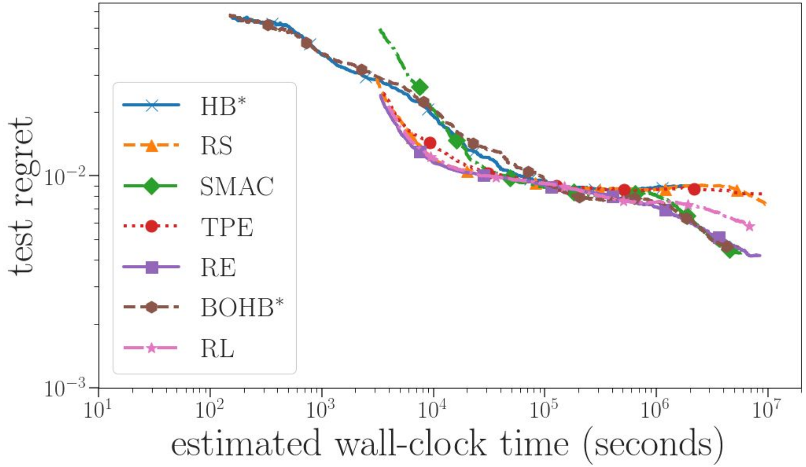


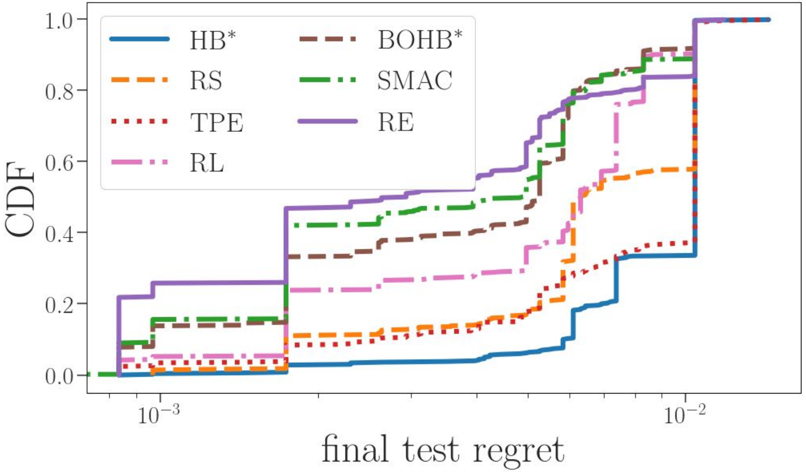




Benchmarking

- Querying dataset enables running entire NAS experiments in seconds
- Can investigate the robustness of NAS methods across random repeats
- Results suggest that conclusions may generalize to larger spaces





Pacific Ballroom Poster #12

Dataset and code available at:
https://github.com/google-research/nasbench