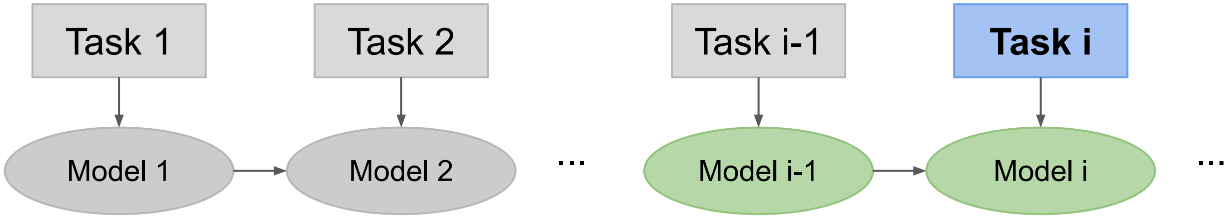
Learn to Grow: A Continual Structure Learning Framework for Overcoming Catastrophic Forgetting

Xilai Li[1*], Yingbo Zhou[2*], Tianfu Wu[1], Richard Socher[2], and Caiming Xiong[2]

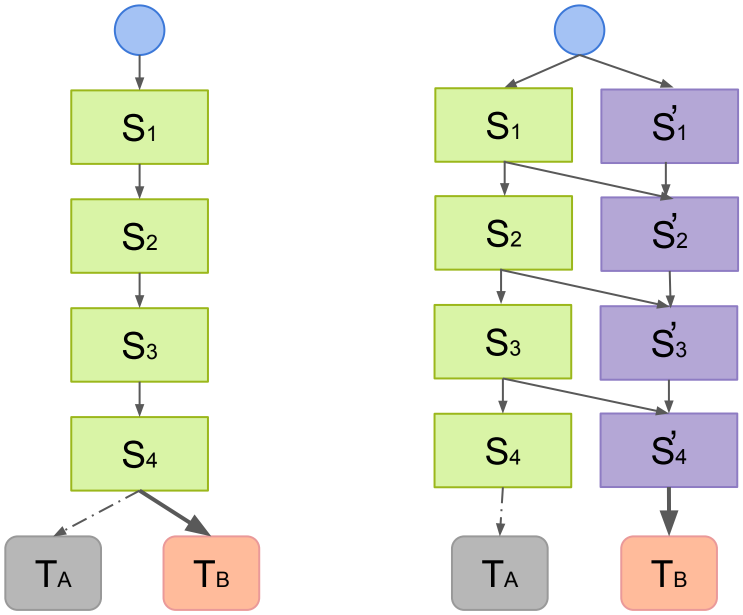
North Carolina State University[1], Salesforce Research[2]

Continual Learning



- Continual Learning is letting model learn multiple tasks sequentially
- Suffers from Catastrophic Forgetting

Current Approaches

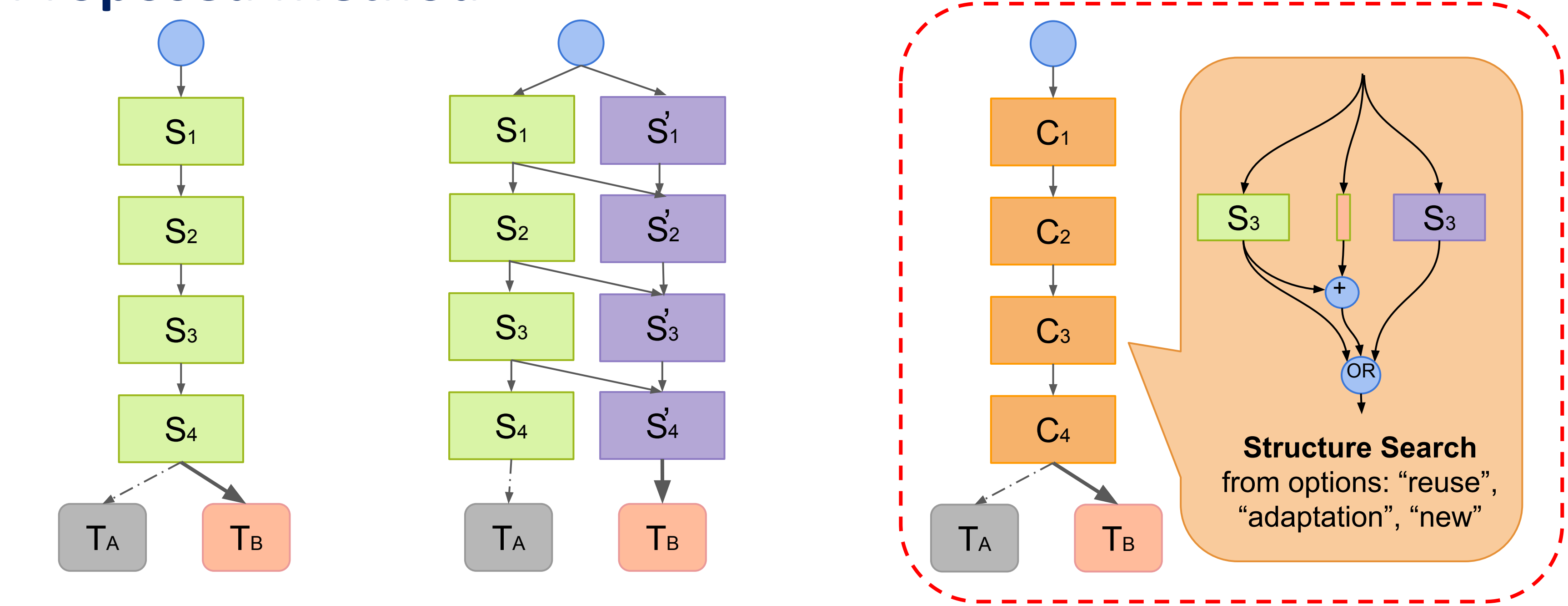


Fixed Structure (**Regularization based method**, e.g. EWC, Kirkpatrick et. al., 2016)

Uniformly Growth for new task (Progressive Nets, Rusu et. al., 2016)

- **Fixed structure: Will finally limited by the capacity**
- **Manually growing is sub-optimal**

Proposed Method



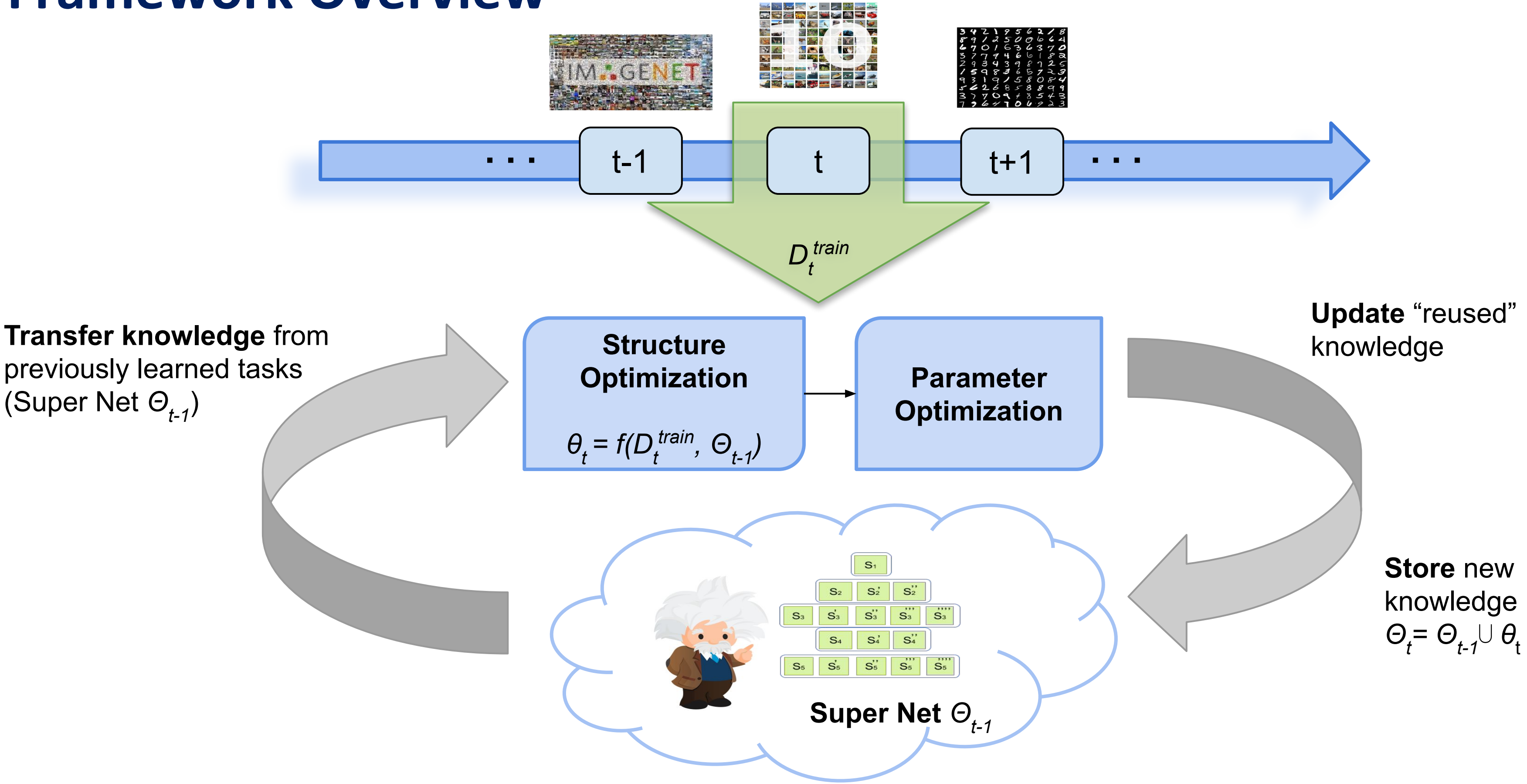
Fixed Structure (**Regularization based method**, e.g. EWC, Kirkpatrick et. al., 2016)

Uniformly Growth for new task (Progressive Nets, Rusu et. al., 2016)

Learn-to-Grow

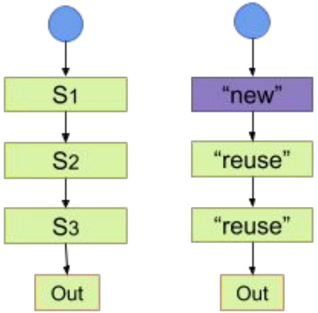
Input | Reused weight | New weight | Adapter | Task specific layer (prev) | Task specific layer (current)

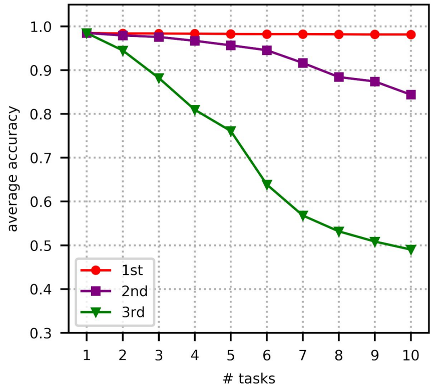
Framework Overview



Structures found by Learn-to-Grow Are Sensible

- Experiments on Permuted MNIST

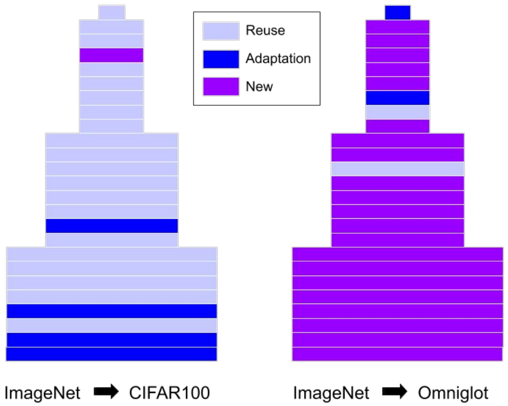




- The structure optimization results in "**new**" on the **first layer** and "reuse" for the rest.

- Ablations experiments validates the search results.

- Experiments on Visual Domain Decathlon (VDD)

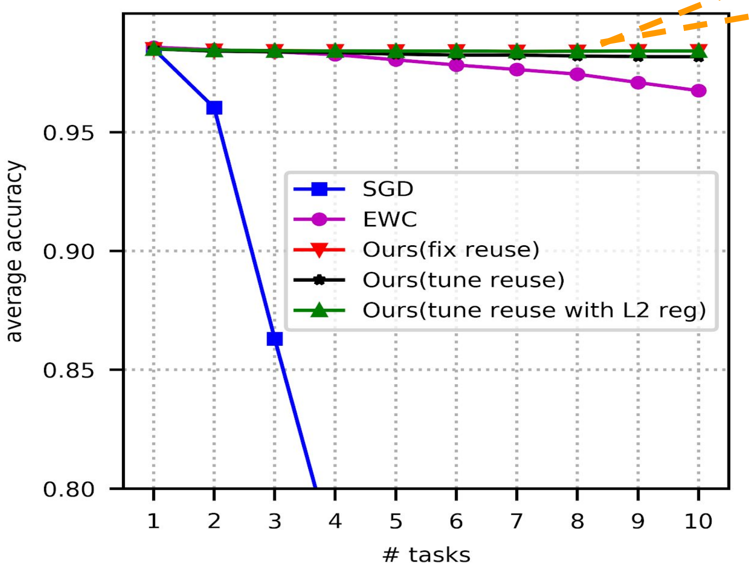


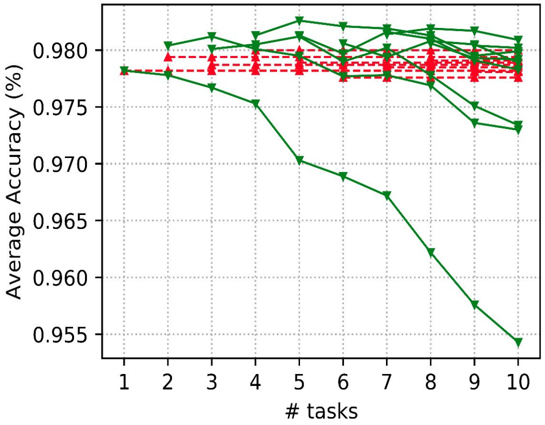
Qualitative analysis on the Searched Structure on Task 2 (Task 1: ImageNet)

➢ Learned structure is **sensible**
  - Similar tasks tends to share more structure and parameters
  - Distant tasks share less

Are Forgetting Alleviated in Learn-to-Grow?

- Experiments on Permuted MNIST



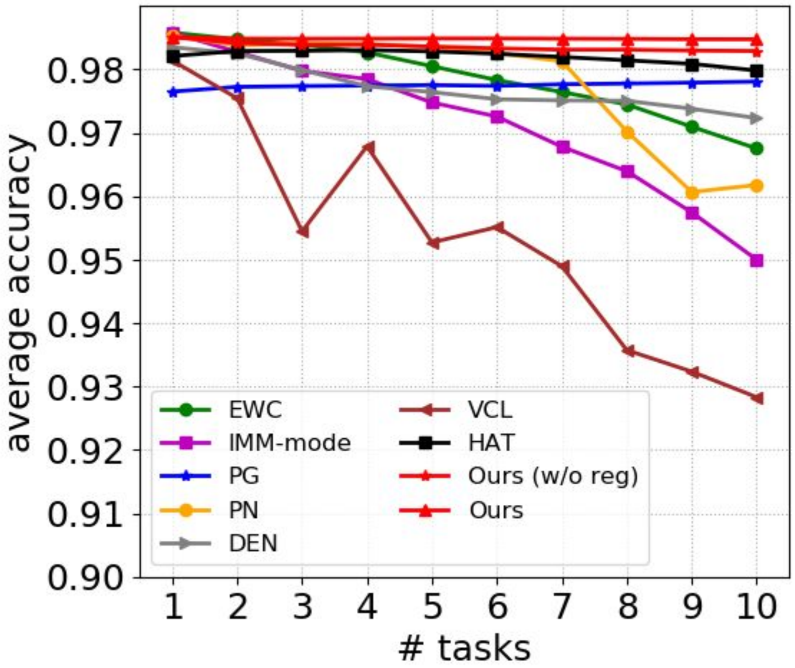


Comparison between "**tune reuse**" and "**fix reuse**"

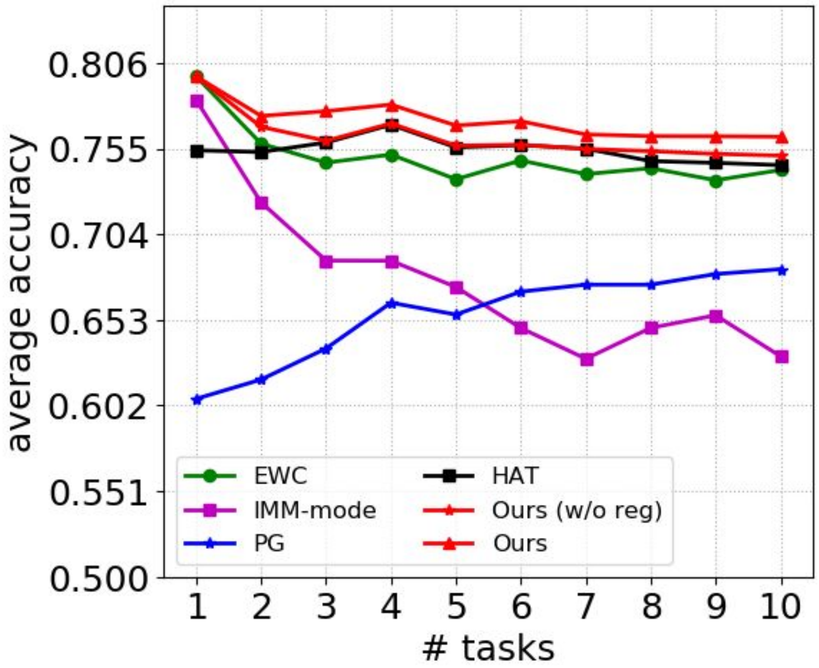
- The "tune" higher than "fix" at certain task indicates "**positive forward transfer**"
- The "tune" curve "goes up" means "**positive backward transfer**"

Comparison with Other Approaches

Permuted-MNIST



Split-CIFAR100



Come to see our poster!

Today 6:30 – 9:00 PM @ Pacific Ballroom **#89**

NC STATE UNIVERSITY Electrical & Computer Engineering

salesforce research