DARTS with Skip Prevention and Flexible Cell Stacking

Torben Eims[1], Jan Malte Töpperwien[1], Peer Duensing[1]

[1]Leibniz University Hannover

1 Summary

- DARTS-informed BOHB
- Cell search space with blocks from template
- Skip prevention in DARTS
  ↳ Scheduled surrogate skips
  ↳ Skip-based early termination
- DARTS instabilities don't transfer to our search space

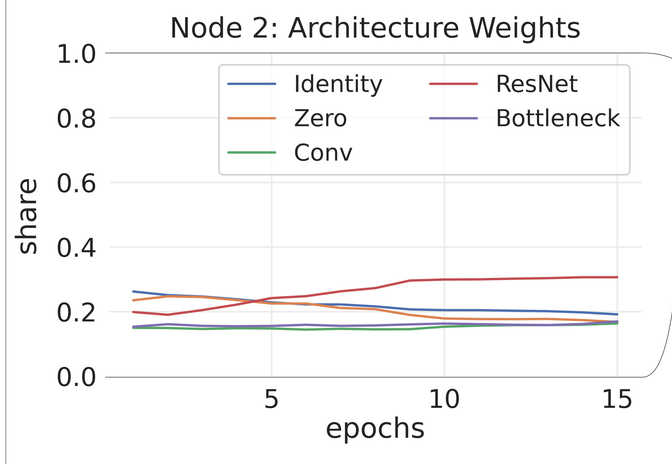
2 Motivation & Problem Setting

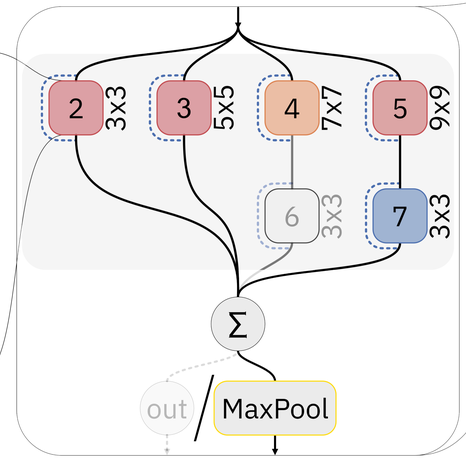
- Inform BOHB by sampling from DARTS
- DARTS selects too many skips
  ↳ DARTS- [Chu, X. et al. (2021)]: adds scheduled surrogate skips to combat vanishing gradients
  ↳ ZeroLess DARTS [Fayyazifar, N. et al. (2023)]: too many skips -> terminate early

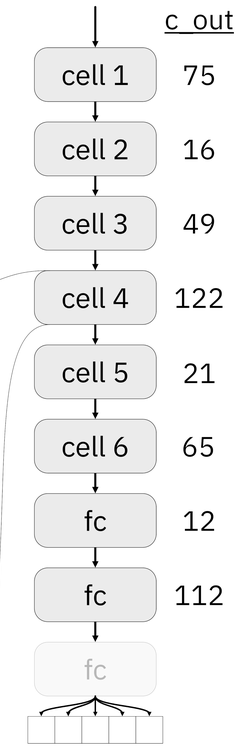
3 Approach

1. DARTS
- Scheduled surrogate skip (DARTS-)
  ▷ Normalized sum as combination
- Early termination (ZeroLess DARTS)
  ▷ Based on skip confidence threshold
- Enable zero operation selection

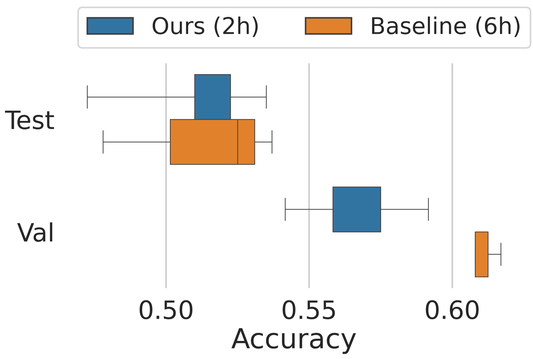
2. BOHB
- Focal loss / class weighting for tackling class imbalance
  ▷ Select cells sampled from DARTS
  ▷ Scheduled surrogate skip during training
- Learning rate scheduling

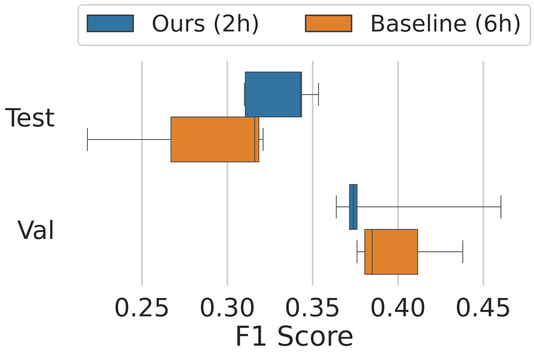




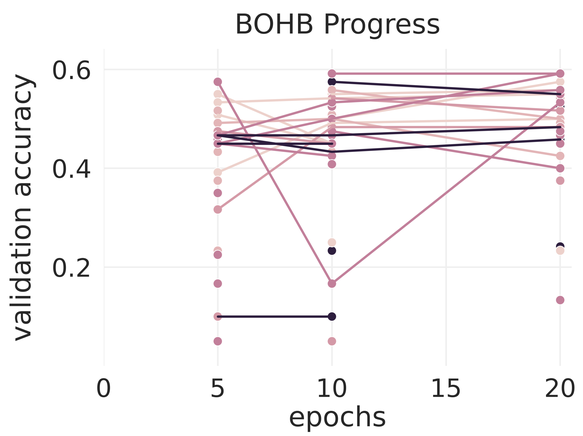


4 Key Insights





- Cannot reproduce DARTS instabilities
  ↳ Zeros and Identities have low shares
  ▷ Amplified distribution for informed operation sampling
- Dataset problems
  ↳ Low correlation between val accuracy and other data splits
  ↳ Train accuracy lower than validation accuracy
- Focal loss helps against class imbalance
- Parallel operations are used
- BOHB chooses more skips with increasing depth



5 Future Works

- HPO before (and after) DARTS
- Decouple cell from template search space
- Additional block types
- Train for more than 20 epochs
- Pipeline for more than 2 hours
- More representative train, val and test split