LIACS
Universiteit Leiden

Towards Neural Architecture Search for robust neural networks

Annelot Bosman, Jan van Rijn, Holger H. Hoos and Marcel Baumann

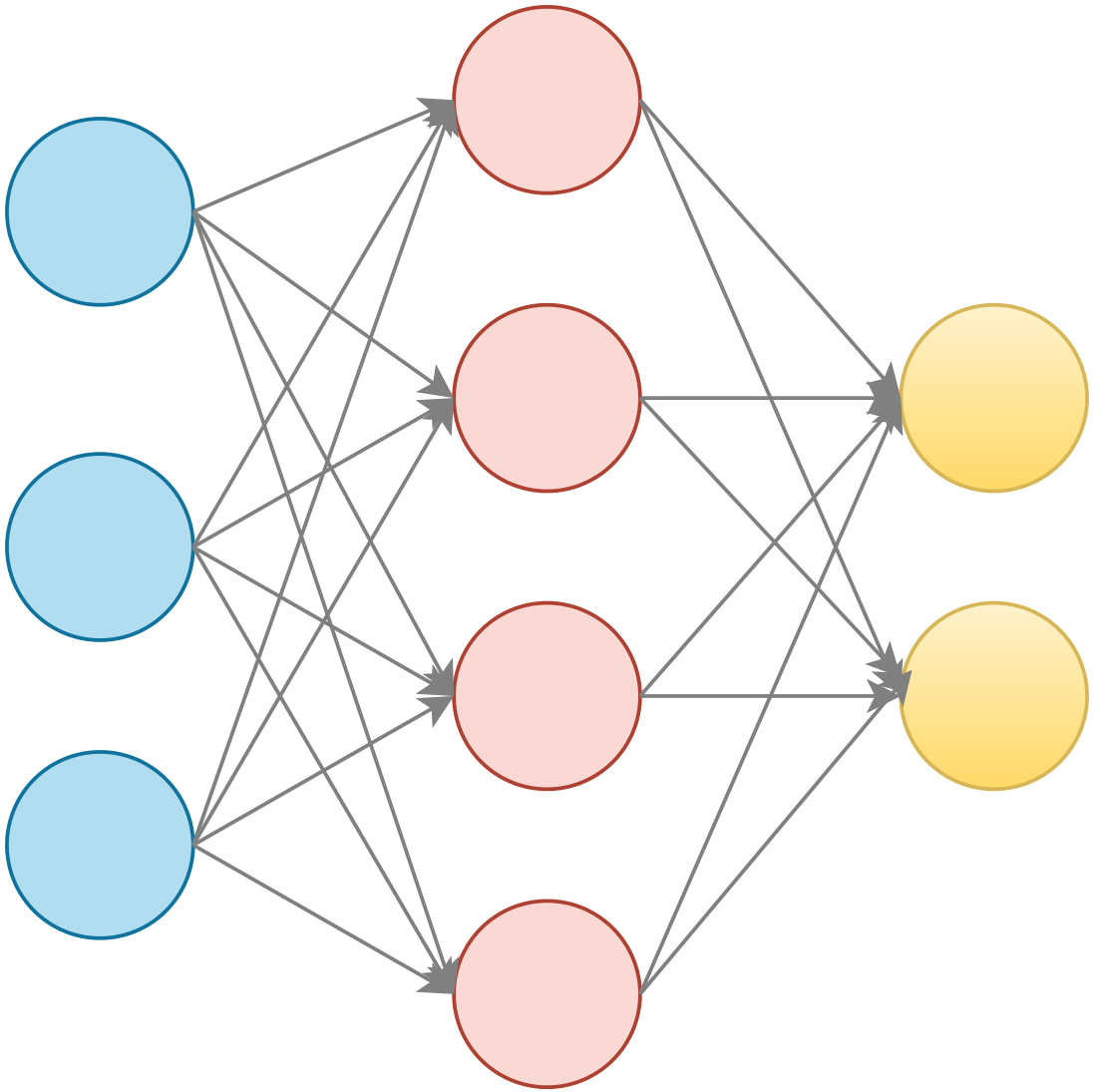
Neural network verification

- Expensive
- Not user friendly
- Considered a posteriori

Neural architecture search

- Expensive
- User friendly and safe options available
- No network defined and trained

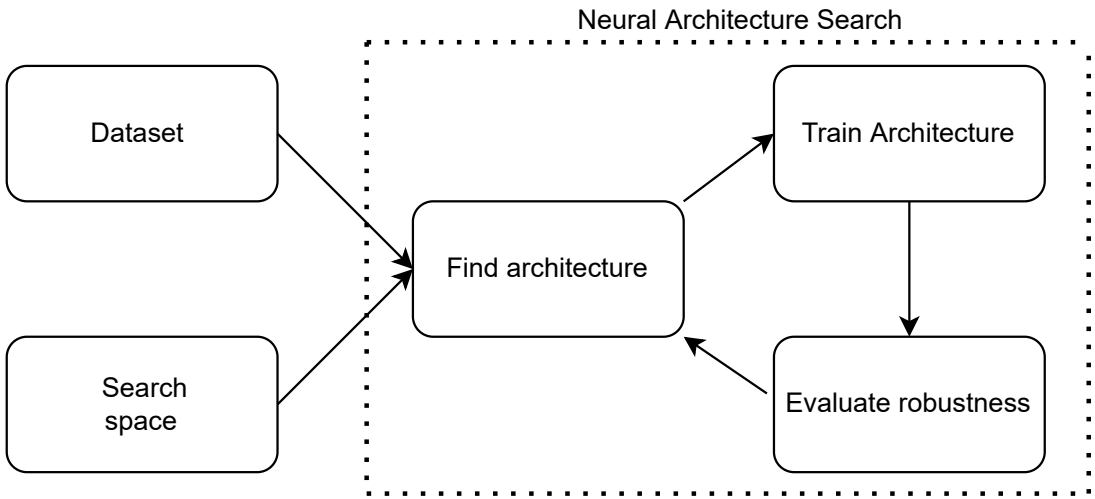
Neural Architecture Search



Neural Architecture Search

- Search space
- Search strategy
- Performance estimation

Components


Neural Architecture Search
Dataset
Search space
Find architecture
Train Architecture
Evaluate robustness

Questions

- Which verification method should we use?
- Which Neural Architecture Search method should we use?
- Should we use exact methods or inexpensive methods during searching?

Questions

- Which verification method should we use?
  DNNV [Shriver et al., 2021]
- Which Neural Architecture Search method should we use?
  Auto-keras [Jin et al., 2019]
- Should we use exact methods or inexpensive methods during searching?

Topics under investigation

- How expensive is including robustness?
- Do we find more robust networks?

Future research

- What features do robust networks have?
- How can robustness be including in training process?

Bibliography

Haifeng Jin, Qingquan Song, and Xia Hu. Auto-keras: An efficient neural architecture search system. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 1946–1956, 2019.

David Shriver, Sebastian Elbaum, and Matthew B. Dwyer. DNNV: A Framework for Deep Neural Network Verification. In *Proceedings of the 33rd International Conference on Computer Aided Verification (CAV 2021)*, pages 137–150, 2021.