Electronic Design Automation
TUM Department of Electrical and Computer Engineering
Technical University of Munich

Bachelor's Thesis, Master's Thesis, Forschungspraxis

Neural Architecture Search for Efficient Vision Transformer

Vision Transformers (ViTs) have shown superior predictive performance than traditional Convolutional Neural Networks (CNNs) in vision applications [1,2]. However, typical ViTs demands large memory and compute due to their huge number of parameters. This restricts their applicability on edge devices with limited memory. Neural Architecture Search (NAS) is a method to automate the design of neural network. Numerous NAS frameworks have effectively found high-performing, efficient neural networks that can be deployed on edge devices [3].

Once-For-All (OFA) [4] proposes a NAS framework that employs progressive shrinking to effectively train a SuperNet as a one-time cost. Specialized sub-networks can then be derived from the SuperNet without additional training. However, OFA has only been applied in CNN-based search spaces.

The project aims to develop a NAS framework to find specialized ViT-based networks for TinyML applications [5]. We are interetested in automated design of efficient ViTs with high accuracy performance and low computational and memory overhead.

[1] Dosovitskiy, Alexey. "An image is worth 16x16 words: Transformers for image recognition at scale." arXiv preprint arXiv:2010.11929 (2020).

[2] Touvron, Hugo, et al. "Training data-efficient image transformers & distillation through attention." International conference on machine learning. PMLR, 2021.

[3] Benmeziane, Hadjer, et al. "A comprehensive survey on hardware-aware neural architecture search." arXiv preprint arXiv:2101.09336 (2021).

[4] Cai, Han, et al. "Once-for-all: Train one network and specialize it for efficient deployment." arXiv preprint arXiv:1908.09791 (2019).

[5] Banbury, Colby, et al. "MLPerf tiny benchmark." arXiv preprint arXiv:2106.07597 (2021).

Prerequisites

- Proficiency in Python.Familiarity with deep learning libraries, such as PyTorch, is a plus.
- Good knowledge of deep neural networks architectures.
- Experience with Vision Transformers would be of advantage.

Contact

If inteteresed, please apply and attach your CV and current transcipt of records to:

mikhael.djajapermana@tum.de

Advisors

Mikhael Djajapermana