SCIENCE · PASSION · TECHNOLOGY

Open Thesis / Project
Folded Object Detection Models

Embedded Learning and Sensing Systems Group

Motivation

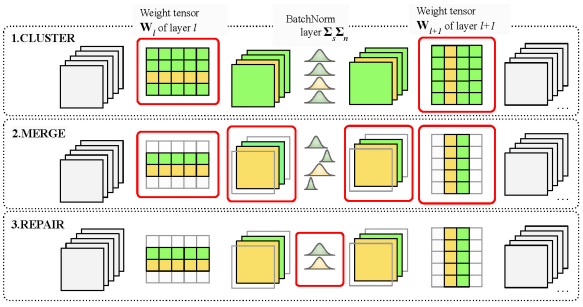
As machine learning models become larger and more complex, they increasingly demand substantial computational resources and memory for training and deployment. Resource-efficient machine learning algorithms address these challenges by optimizing performance for systems with limited resources. As the use of deep learning models on edge devices grows, these models often struggle to fast inference time and smaller model size. Contrast to pruning, quantization, and distillation, Model Folding [1] was proposed to reduce parameter size by clustering similar neurons and preserving data statistics. In object detection, it is unclear how model folding will improve SOTA object detection models like YOLO11n, and can you combine model folding and quantization on YOLO models. The goal of this project/thesis is to implement an efficient, on-mobile object detection algorithm by folding YOLO models without hurting its performance. **Interested? Please contact us for more details!**

Target Group

Students in ICE, Computer Science or Software Engineering.

Thesis Type

Master Project / Master Thesis.



Goals and Tasks

- Thorough literature research on model efficiency and model compression for object detection models;
- Implement model folding for YOLOv11n and/or explore quantize a folded YOLO model.
- Evaluate the performance of folded YOLO models;
- Present the results of your work and summarize the outcomes in a written report.

Requirements:

- Knowledge of neural networks;
- Programming skills, Python;
- Knowledge of deep learning frameworks (eg. PyTorch or TensorFlow).

Used Tools & Equipment

- A computation cluster of TU Graz and an Android phone.

References

[1] Dong Wang et al. "Forget the Data and Fine-tuning!
Just Fold the Network to Compress". In: *Proceedings of the International Conference on Learning Representations (ICLR)*. 2025. URL: https://openreview.net/forum?id=W2Wkp9MQsF.

Contact Persons

- Dong Wang (dong.wang@tugraz.at)
- Assoc. Prof. Dr. Olga Saukh (saukh@tugraz.at)