Hauke Kirchner

Benchmarking tree species classification with synthetic data and deep learning

Table of contents

1 Motivation

2 Methods

3 Tools

Why is it important to benchmark the training process of neural networks?

Training speed

Energy efficiency

Training speed

- training of neural networks is computational intensive
- ⇒ workflow needs to be optimized for available hardware
  - How many GPUs are worth to request?
  - What is the best set of software? (pytorch, cuda, ...)
- ⇒ high impact of deep learning applications on energy consumption

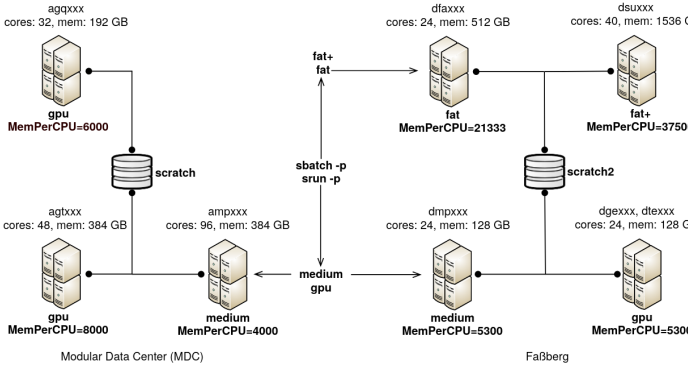
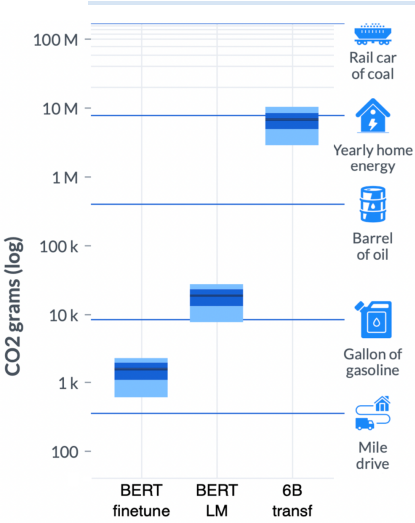


Image source: https://www.gwdg.de/web/guest/hpc-on-campus/scc, Accessed on: 09.11.2022

Energy efficiency ⚡

- asdfasdf

⇒ as deep learning is emerging in several fields the impact on energy consumption and consequently our climate optimized training processes are essential



CO_2 Relative Size Comparison

Image source: Adapted from todo

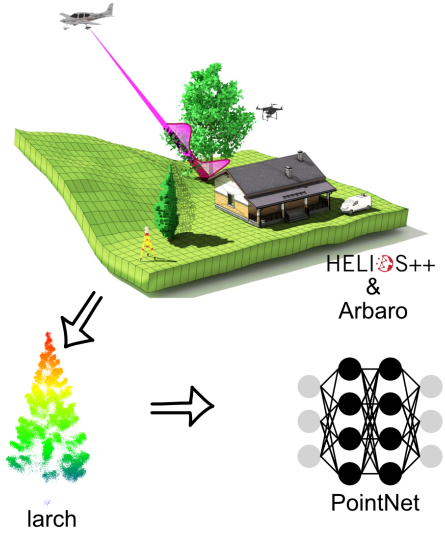
Use case

Training of PointNet with synthetic data

- lack of pre-trained models

Tree species classification

-



Outline

1 Motivation

2 Methods

3 Tools

Methods

- data loading
- trainin time

Tools - Overview

tool	purpose
tensorboard	
Vtune	
likwid	
PyTorch - built-in	

Benchmarking is the first step of optimizing

Image generated with stable diffusion:
"Sherlock Holmes locates the best graphical processing unit inside the data center for his deep learning workflow"

- "Stable Diffusion v1 version of the model requires 150,000 A100 GPU Hours for a single training" session[a]

[a] https://syncedreview.com/2022/11/09/almost-7x-cheaper-colossal-ais-open-source-solution-accelerates-aigc-at-a-low-cost-diffusion-pretraining-and-hardware-fine-tuning-can-be/, Accessed on: 10.11.2022

References