



Lightweight ML Artifact Store & Viewer

Project Management and Software Development
for Medical Applications

General Info

Contact Person: Nikolas Brasch

Contact Email: nikolas.brasch@tum.de

Project Abstract

This project aims at building a system to store artifacts (e.g. tables, figures, images, meshes, ...) for 3D machine learning use cases. The goal is to combine the convenience of established file formats in a customizable folder structure on a distributed storage system for easy access with a lightweight, performant and storage efficient meta data store for fast interactive querying, visualization and comparison of artifacts.

Background and Motivation

While there exist many solutions to log, visualize and compare artifacts of an ML development process, most of them use custom representations to store the artifacts, which requires a custom API for accessing the artifacts making the integration of novel artifacts difficult and prevents easy access to specific data for post-processing.

Therefore, this project aims at creating a metadata store and web based visualization system around a file system based storage system of different file types (e.g. tables, images, figures, videos, 3D data). When logging artifacts part of the data is usually duplicated from another source e.g. bounding boxes or segmentations are shown on top of images or point clouds, in order to save storage dynamic linking to existing data on the storage shall be used where possible. To allow for fast queries and comparison a system is needed to create meta data of the logged data while it is

being logged. To allow visualization and comparison of different experiments a web-based system shall be developed that can leverage the data stored in the file system and the metadata to minimize storage overhead.

Student's Tasks Description

The tasks include

1. Development of a logging system that stores artifacts in their native file format
2. A system that generates metadata of the logged data allowing fast queries
3. Development of an API that allows filter queries to retrieve artifacts for specific experiments or value ranges etc.
4. Development of a web-based visualization that leverages both metadata and data stored on the file system and allows the comparison of multiple experiments

Technical Prerequisites

- Programming skills (required)
- Using databases (helpful)
- Web development (helpful)

References

- [1] [Tensorboard](#)
- [2] [TensorBoard3DPlugin](#)
- [2] [Open3D Plugin](#)
- [3] [Weights & Biases](#)