In-Browser Raster Data Processing

Vishnu Raj Cherukudi Mattuvayil

April 21, 2024

Type of work:	Research Project
Student:	M.Sc. Vishnu Raj Cherukudi Mattuvayil
Supervisor:	Dr.-Ing. Marcus Paradies
Chair:	Prof. Dr.-Ing. habil. Kai-Uwe Sattler
Department:	Department of Databases and Information Systems
Semester:	Summer semester 2024

Keywords: WebTensor, WASM, Raster Data

Motivation

In recent years there has been a sudden growth in the amount of raster data that is being generated. This is from satellites that are analyzing the geographical features of Earth. Two types of spatial data are generally modelled: Raster data and Vector data. Satellite imagery is an example of raster data and is normally represented in the form of multi-dimensional arrays. Vector data is represented as a set of points, lines, and polygons, and is used to represent geographical features such as GPS locations, regional boundaries, and roads [1].

In this project, the focus is on raster data and its processing in the browser with WebAssembly(WASM). Traditional usage of any data processing is done in a server where the data is located or the data is located in another data server and loaded into another server for processing. After that, the data is presented to the client-side using Javascript and HTML. To make the processing faster, rather than using server-side processing the idea is to use WebAssembly for client side processing of the raster data and link the WASM files with javascript to achieve almost native execution speed [2].

Task Description

- Implement a raster data processing library of selected operations using WebAssembly.
  - Research on how to implement the WASM program.
  - After implementation link the program with JavaScript and HTML for testing.

- Implement a data loader from a storage service (e.g. S3) for the above operation into the browser.
  - Research on how to implement the data loader for raster data.
  - Test with localhost then move on to any external server for data consumption.
- Evaluate and compare the performance of the implemented functions.
  - Try to implement the same program with JavaScript alone or with some other traditional programming methodology.

Scientific Approach

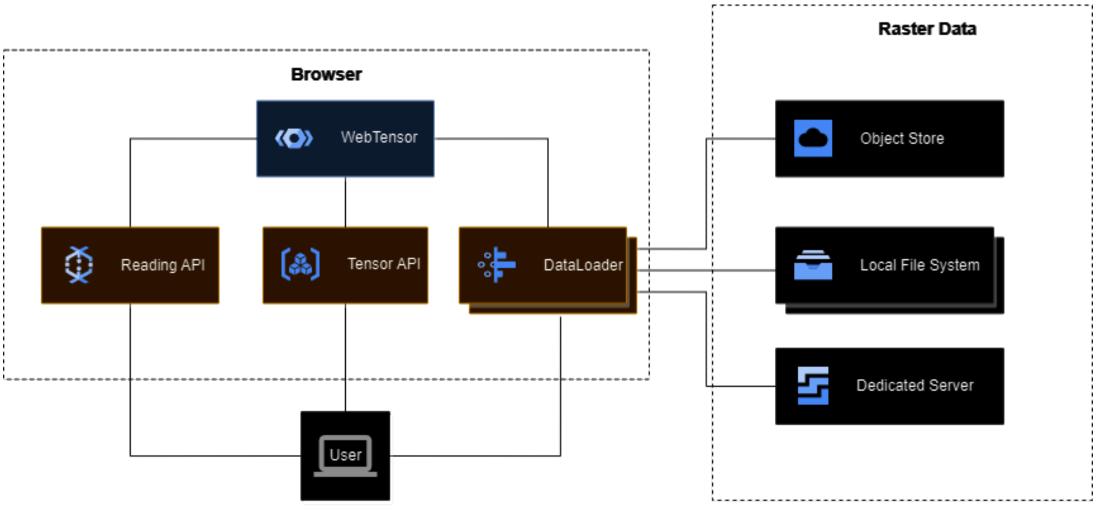

Figure 1: Architecture of WebTensor (orange: JavaScript components) [2]

The objective of this project is to implement a raster data processing library containing the components given in Figure 1. The raster data may be stored in an external server or in the local system itself and it is loaded through the dataloader for further processing with the WASM program. After the data processing, it should be visualized in the browser by linking the results with JavaScript and HTML. Further explanations of components and tasks are given below. The sample raster data for testing may be used with this : Natural Earth Raster Data

0.0.1 WebTensor

WebTensor is the chunked tensor implementation for raster data processing using WASM. This design is used to serve as a backend for JavaScript programs [2]. The programming language to implement the raster data processing library can be selected from the currently available languages. From the written program, it can be compiled to WASM using a compiler compatible with the selected programming language. The operations used inside the WebTensor can be data accesses, aggregations and more.

0.0.2 DataLoader

The dataloader is used to enable JavaScript programs to load data to the WebTensor available from various backends. Since WASM programs cannot access JavaScript memory directly, data should be loaded from the backend onto the WASM heap [2]. The data can be loaded from an object store, local file system or a dedicated server.

0.0.3 Benchmarks

To compare the performance of WASM programs, the same WASM program can be implemented with JavaScript. Then the time required for the operations can be compared from both the programs.

Time Schedule

Time period	**Tasks**
April 21	Finishing and uploading this exposé
April 22 - May 1	Research dataloader implementation methods
May 1 - May 24	Develop dataloader for raster data
May 27 - June 3	Research methods for implementing raster data processing library
June 4 - June 28	Develop the WASM program
July 1 - July 12	Integration and testing
July 15 - July 30	GUI development
September 1	Finishing and uploading project report

References

[1] S. Singla, A. Eldawy, T. Diao, A. Mukhopadhyay, and E. Scudiero, “Experimental study of big raster and vector database systems,” in *2021 IEEE 37th International Conference on Data Engineering (ICDE)*, pp. 2243–2248, 2021.

[2] L. F. Naumann, “WebTensor: Towards high-performance raster data analysis in the browser,” in *Datenbanksysteme für Business, Technologie und Web (BTW 2023), 20. Fachtagung des GI-Fachbereichs „Datenbanken und Informationssysteme" (DBIS), 06.-10, März 2023, Dresden, Germany, Proceedings*, vol. P-331 of *LNI*, pp. 1083–1089, Gesellschaft für Informatik e.V., 2023.