

# Haystack with Similarity Search

Final Project Presentation

By: Sunkari Sai Satvik (2021101117)  
Rushil Kaul (2021101063)



# Statement of the Problem & Scope

Statement of the Problem: Implement Haystack File System and add a visual similarity search feature to retrieve images based on visual content.

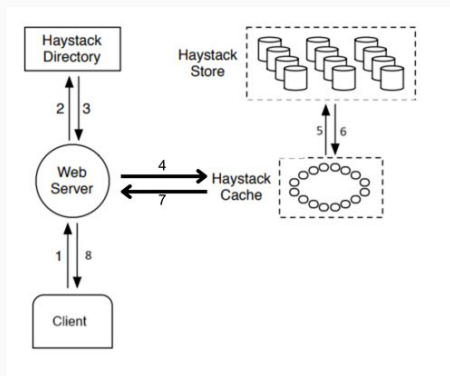
Scope:

1. Implement Haystack-style image storage.
2. Develop a visual similarity search to organise and retrieve images based on visual content using feature vectors.

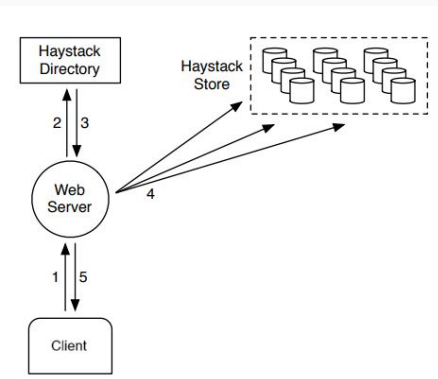
# Haystack Overview & Core Operations

Haystack File System: Metadata and raw data separation for efficient retrieval.

Main operations: Read, Write, Delete.



Read/Delete Operations



Write Operations

# Extension: Visual Similarity Search

Objective: Implement visual similarity search using image features.

How We Implemented:

Feature Extraction: Using MobileNetV2 (64-dimensional feature vectors).

Faiss: For fast similarity search with L2 distance.

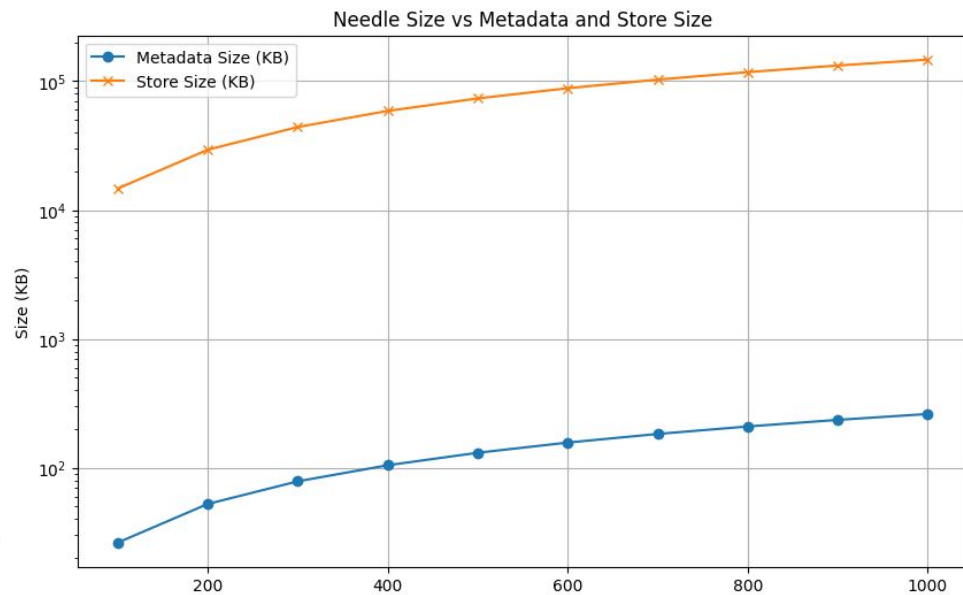
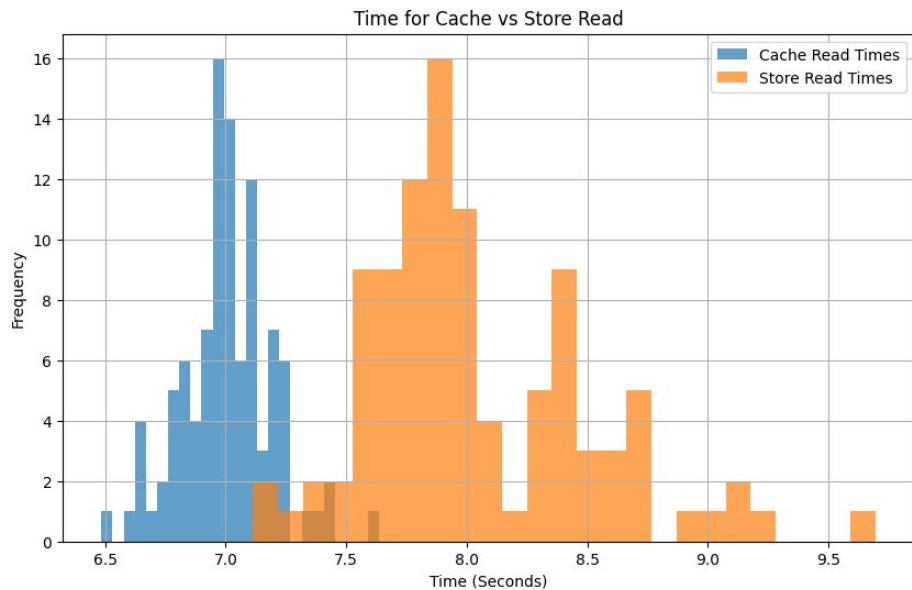
# Feature Extraction & Distributed Faiss

Feature Extraction: Pre-trained MobileNetV2 model generates compact feature vectors.

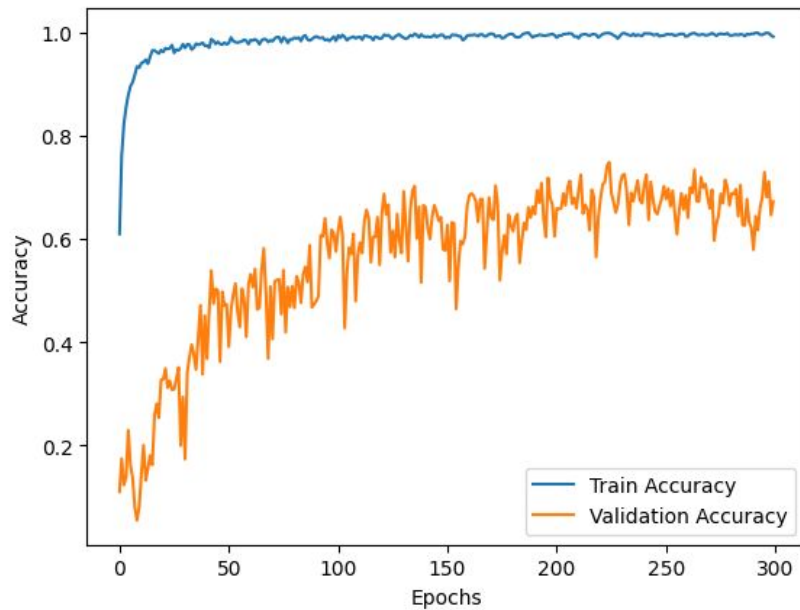
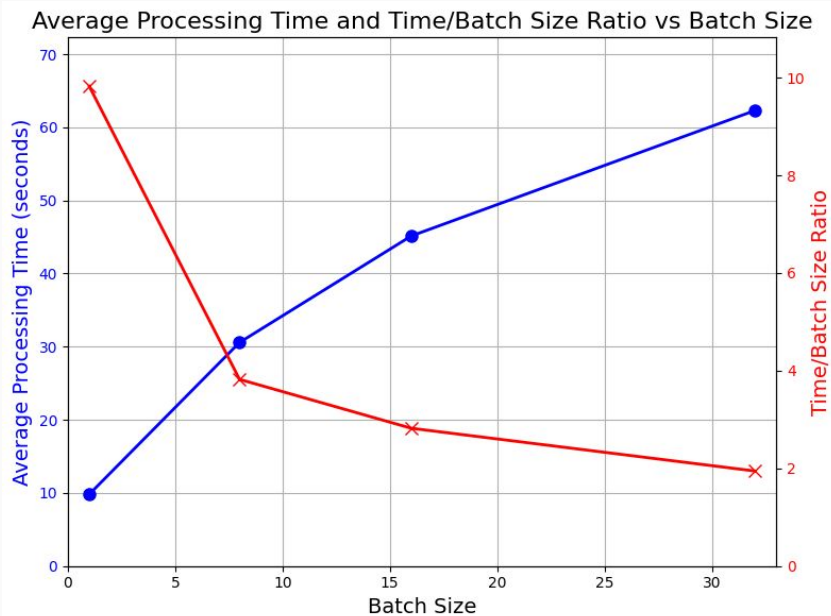
Distributed Faiss: Used for efficient similarity search through indexed feature vectors.

Query Pipeline: Query image  $\rightarrow$  Extract feature  $\rightarrow$  Faiss search  $\rightarrow$  Retrieve image.

# Some Benchmarks and Graphs



# Benchmarks and Graphs



# Conclusion

Visual similarity search improves image retrieval by enabling content-based search.

References:

"Finding a Needle in Haystack: Facebook's Photo Storage"

"Faiss: A Library for Efficient Similarity Search"

GitHub, Documentation, Tutorials links.



Thank you!