Animal Image Retrieval Using Deep Learning and FAISS

BY

BENSON BABY THOMAS

crbenson7@gmail.com

07553358087

Table Of Contents

1. Introduction.

Image retrieval systems are essential for applications ranging from digital asset management to content-based image search. Traditional image retrieval methods often rely on handcrafted features or low-level image descriptors, which may not capture the complexity of visual content. This project enhances the retrieval process using a deep learning model to extract high-level features and FAISS (Facebook AI Similarity Search) for efficient and scalable similarity search.

2. Dataset.

The dataset used in this project comprises a collection of 4,738 images featuring various animals. The images are captured in different settings and conditions, providing a rich and diverse set of visual content. The dataset includes multiple species and animal types, ensuring a broad range of features and visual characteristics for effective retrieval.

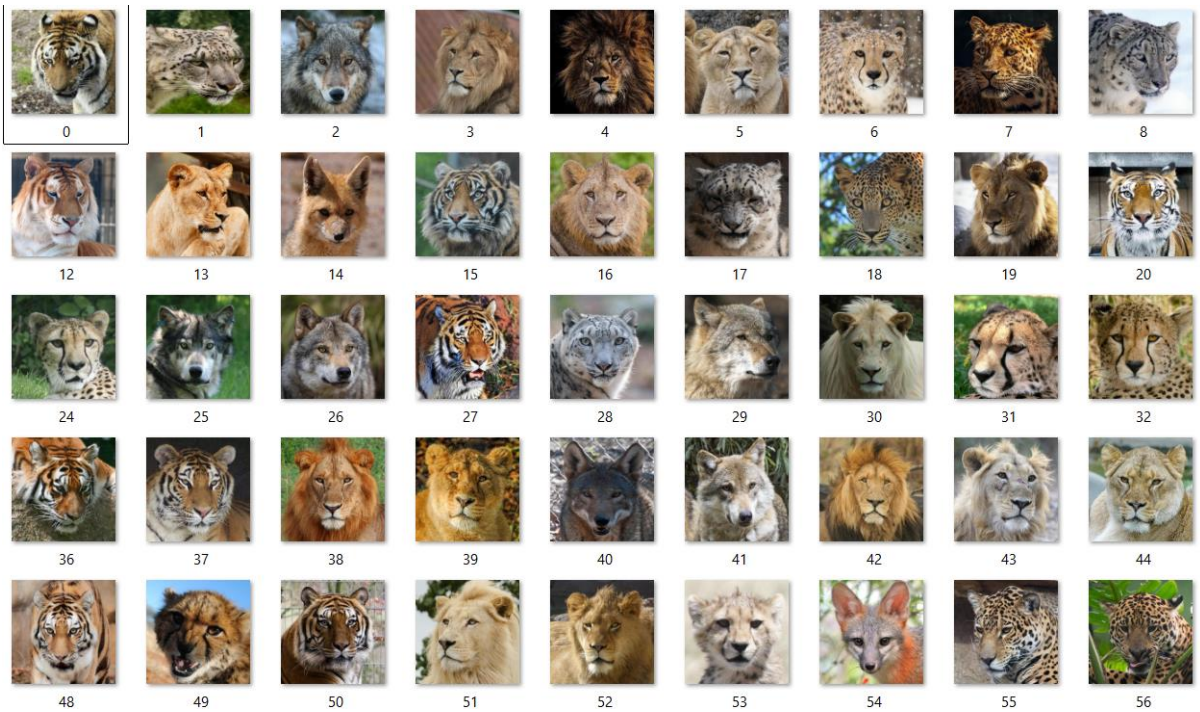

Figure 1: Dataset Sample

3. Methodology.

3.1. Feature Extraction.

To capture rich semantic features from images, we use a pre-trained ResNet-50 model. ResNet-50, a deep convolutional neural network, is known for its ability to extract robust feature representations due to its residual learning framework. The model is pre-trained on the ImageNet dataset, which provides a strong starting point for various image recognition tasks.

Model Setup: The ResNet-50 model is loaded with pre-trained weights and set to evaluation mode to ensure that no additional training occurs during feature extraction. The model is moved to GPU if available to leverage faster computations.

Transformations: Each image undergoes a series of transformations including resizing, center cropping, conversion to tensor, and normalization to match the input requirements of the ResNet-50 model.

The extract_features function processes each image to generate a feature vector representing the image's content. These feature vectors are then flattened and stored for later use.

3.2. Indexing with FAISS.

FAISS is employed to index and efficiently search through the high-dimensional feature vectors. FAISS provides an optimized library for similarity search, enabling fast retrieval even with large datasets.

Index Creation: An IndexFlatL2 instance is used, which performs exact nearest neighbour search using L2 distance. The dimensionality of the index is determined by the length of the feature vectors.

Feature Storage: Features extracted from all images are added to the FAISS index, enabling rapid similarity queries.

3.3. Querying and Retrieval.

For querying, the system extracts features from the query image and uses the FAISS index to find the most similar images.

Query Processing: The query_faiss function extracts features from the query image, reshapes them to match the index input format, and performs a search to retrieve the top-k most similar images.

Results Display: The display_images function visualizes the query image alongside the retrieved similar images. This function also measures and displays the time taken for the retrieval process to assess the efficiency of the system.

4. Results and Discussion.

The implementation of the image retrieval system demonstrates the effectiveness of using deep learning features and FAISS for similarity search. The use of ResNet-50 ensures high-quality feature extraction, while FAISS provides efficient and scalable search capabilities. The results are validated through visual inspection of retrieved images, showcasing the system's ability to return visually similar images.

Figure 2: Result of Image Retrieval

The system successfully demonstrated its capability by retrieving a query image of a tiger. The query image, sourced from Google, was processed and used to find top 7 similar images in the dataset of 4,738 animal images.

5.Conclusion.

The proposed image retrieval system successfully integrates deep learning with FAISS for a robust and efficient similarity search. Future improvements may include experimenting with different deep learning models, incorporating additional indexing techniques, and optimizing the system for even larger datasets.

This approach provides a powerful tool for various image retrieval applications, offering both high accuracy and fast retrieval times.