

Image Clustering Project Documentation

Introduction

In this project, we undertook the task of clustering a set of images based on their content. The original intent was to use ImageBind LLM embeddings for this purpose. However, due to accessibility constraints, we opted to utilize the embeddings from a pre-trained ResNet50 model provided by TensorFlow's Keras API, which has been extensively trained on the ImageNet dataset.

Methodology

Feature Extraction:

- Model Used: ResNet50, excluding the top classification layer to obtain feature representations.
- Image Preprocessing: Images were resized to 224x224 pixels, the expected input size for ResNet50, and preprocessed to match the model's training data format.
- Embedding Extraction: Each image was passed through the model to obtain a 2048-dimensional embedding vector from the penultimate layer.

Clustering:

- Algorithm: K-means clustering.
- Determining 'k': The Elbow Method was employed to ascertain the optimal number of clusters. The plot indicated a potential elbow at $k=4$, which was chosen for the clustering.
- Clustering Execution: We applied the K-means algorithm with $k=4$ to the extracted features to segment the images into four distinct clusters.

Results

Clustering Visualization:

- t-SNE: A dimensionality reduction technique was used to visualize the embeddings in two dimensions. The t-SNE plot displayed discernible clusters, corroborating the effectiveness of the feature extraction and clustering process.
- Cluster Analysis: Upon visual inspection of the images in each cluster, clear thematic content was identified, validating the clustering's intuitive correctness.

Conclusion

The project successfully demonstrates the utility of transfer learning for image clustering tasks. By leveraging the powerful feature extraction capabilities of a pre-trained CNN, we could segment the images into meaningful clusters without direct access to the original ImageBind LLM embeddings.

Future Work

Further work could include a more granular analysis of the cluster contents, experimenting with different clustering algorithms, and exploring other pre-trained models for feature extraction.