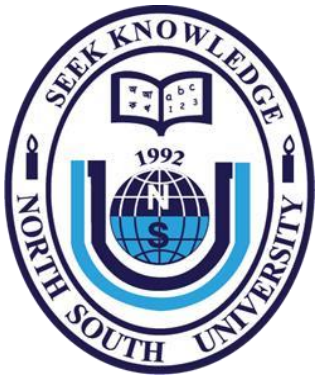
Department of Electrical and Computer Engineering North South University

---



Nearest Neighbour of Different Image Representation

Technical article based on our project

Submitted by:
Ankur Chowdhury ID:1911844042
Sajid Wasif ID: 1912313642

Faculty Advisor:
Dr. Mohammad Ashrafuzzaman Khan
Assistant Professor
North south university

Introduction

The main objective of our project is to leverage the extracted features to build a system that enables similarity search. In this report, we delve into the exploration of several state-of-the-art CNN models, including VGG16, ResNet101, ZFNet, and MobileNet, for feature extraction from a given dataset. By providing an input image, our system identifies and retrieves the ten nearest images based on their feature similarity, thereby facilitating efficient content-based image retrieval. With the dataset in hand, we employ various CNN architectures to extract features from the images. To extract features, we feed the images through the pretrained CNN models, utilizing the learned weights and biases obtained from training on large-scale image classification tasks. By removing the fully connected layers at the end of the models, we access the intermediate feature maps, which serve as rich representations capturing image characteristics. These feature maps are then flattened into 1D arrays to form a compact representation of the images' visual content.

Next, we leverage these extracted features to build a similarity search system. Given a query image, we compare its feature representation with those of the entire dataset. To efficiently compute similarity, we employ various distance metrics such as Euclidean distance or cosine similarity. The images with the closest feature representations are retrieved and presented to the user as the ten most similar images.

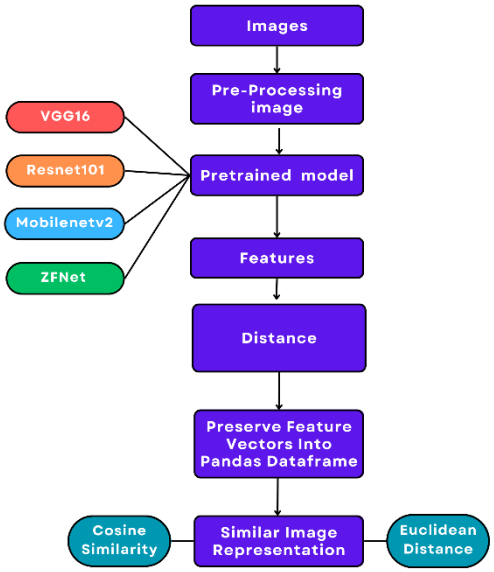
Nearest neighbor of an image

The process of converting an image into a numerical vector representation is known as "image-to-vector." To put it another way, it entails extracting the visual information from a picture and transforming it into a structured, compact numerical representation that can be processed quickly by computers. These tasks include feature extraction, similarity search, and classification.

Image to vector

CNN models retrieved characteristics from photos, converted them into feature vectors, and used the cosine similarity metric and Euclidian distance to compare them. The 10 closest photos were displayed as the system's output, making it possible for users to look through and find related photographs.

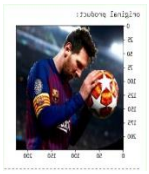
Methodology



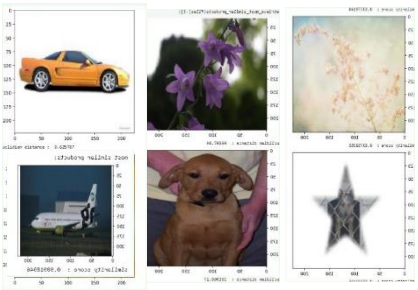
In the beginning, we chose four pretrained models (Vgg16, Resnet101, Zfnet, and mobileNet-v2) to extract features from each picture in our dataset, convert each feature into a single feature vector, and then calculate the distance between these vectors. Finally, we created an input system in which the user can enter a random image, and our system will return the 10 nearest neighbors of that image using cosine similarity and Euclidian distance so that we can see what they represent.

Results

input



output



input

output

In our results, this system uses pixel values to detect objects, extracting features from the input image as feature vectors. Outputs show similarity between different objects, which humans can understand and the system can make them understand. And Object detection can be varied with interesting similarities from different fields, providing an opportunity to explore and observe similarities in depth.

Conclusion and Future Work

The system successfully demonstrated the use of pretrained models and feature extraction techniques to detect similarities between images of various types and from different environments. The system correctly identified related images by computing feature vectors and measuring distances. The results demonstrate the value of cross-domain image similarity in applications such as content-based image retrieval and creative exploration. Deep learning advancements and integration could improve the system's performance and broaden its applications. Overall, this project may help to advance the field of cross-domain image analysis by providing valuable insights for future research and practical applications.

In terms of future work, we are interested in exploring how the features vector influences object detection and using more diverse CNN models to improve similarity score and Euclidian distance.

References

[1] http://poseidon2.feld.cvut.cz/conf/poster/poster2018/proceedings/Poster_2018/Section_IC/IC_057_Listik.pdf

[2] Local Aggregation for Unsupervised Learning of Visual... - Google Scholar

[3] NEAREST NEIGHBOUR STRATEGIES FOR IMAGE UNDERSTANDING (psu.edu)N. Barla, "v7labs," v7, 15 March 2022. [Online]. Available: https://www.v7labs.com/blog/semantic-segmentationguide#:~:text=Semantic%20Segmentation%20follows%20three%20steps, by%20creating% 20a%20segmentation%20mask. [Accessed 24 April 2022].

[4] KHLOE, "Image Recognition Applications," 16 January 2022. [Online]. Available: https://www.datasciencesociety.net/image-recognition-applications-7-essential-future-uses/. [Accessed 25 April 2022].