Project Phase 1

Group Members:

- Aditya Rao - 1229544344 (arao87@asu.edu)
- Dev Patel - 1229417087 (dpate243@asu.edu)
- Gaurang - 1230219135 (gnolas35@asu.edu)
- Abhinav Prasad Yasaswi - 1229563753 (abhinavp@asu.edu)
- Ansh Ray - 1229548426 (aray49@asu.edu)
- Dhruv - 1229609591 (dpraja11@asu.edu)

Abstract:
This project delves into computer vision and image analysis, focusing on image feature extraction, vector modeling, and similarity metrics. Group members individually undertake the tasks while collaboratively sharing ideas. Essential tools include Python, PyTorch, and torchvision. The project utilizes the ResNet50 architecture and the Caltech 101 dataset. Tasks encompass implementing programs for visualizing and extracting feature descriptors, including Color Moments and Histograms of Oriented Gradients (HOG), and employing ResNet models. Feature descriptors are stored, and the project concludes with a program to find and visualize the most similar images based on selected models and distance measures.

Keywords: Feature Extraction, Vector Models, Similarity Measures Deep Learning, Computer Vision

Introduction:
In the world of computer vision, we're about to dive into some important concepts. This project is all about understanding key terms, defining our goals, and acknowledging some basic assumptions.
Terminology: To start, let's get familiar with some words and phrases that are essential for our project. We'll be talking about things like extracting features from images, using vectors, and measuring how similar or different images are.
Project Goal Description: Our main aim here is to give everyone a hands-on experience in working with images and understanding how to compare them. We have divided this project into tasks: Task 0 will get us comfortable with the tools we'll use, Task 1 will teach us how to extract features from images, Task 2 is about handling data, and Task 3 will show us how to find similar images.
Assumptions: There are a few things we're assuming as we start this project. We assume that everyone will use certain programming tools and libraries (Assumption 1). We also assume that we have access to specific pre-made models and a dataset (Assumption 2). Lastly, we encourage group members to work together and share ideas while each person works on their own tasks (Assumption 3).

System Description:
The proposed solution requires to store the extracted features vectors of each of the 5 features models into csv files. Hence, one csv file contains feature vectors of all the images in the Caltech-101 dataset. Similarly, 5 csv files for each of the 5 models. Now while executing the

code, we need to enter the user id, select the feature model and input the value of k (k-similar images). In the execution flow, we first store the data of the specified feature model from the csv file to a dataframe. Then read the query image from the id given by the user. Then, we resize the image to (300,100) and convert to a numpy array. we then extract feature of the query image and then compare with the database to get k similar images.

Interface specifications:
Currently, the interface for the user is the terminal. The user needs to run the code and enter image id, feature model and k as their input. The k-similar images will be automatically stored in the mentioned folder.

System requirements/installation and executuion instructions:

The system needs to have the latest version of python installed. The following python libraries will be required:

- torch
- torchvision
- PIL
- matplotlib
- numpy
- scipy
- pandas
- cv2
- math
- sklearn

Related Work:
[1] Efficient Handwritten Digit Recognition based on Histogram of Oriented Gradients and SVM

- In this paper, the authors have proposed an appearance feature-based approach that processes data using Histogram of Oriented Gradients (HOG).
- Given that it is a gradient-based descriptor, HOG is an efficient feature descriptor for handwritten digits that is stable under changes in illumination.
- Moreover, linear SVM has been employed as a classifier and an overall accuracy of 97.25% has been achieved.

[2] Histograms of Oriented Gradients for Human Detection

- In this paper, the authors have shown experimentally that grids of Histograms of Oriented Gradient (HOG) descriptors significantly outperform existing feature sets for human detection
- They have presented the report on the influence of each stage of the computation on performance, concluding that fine-scale gradients, acceptable orientation binning, relatively coarse spatial binning, and high-quality local contrast normalization in overlapping descriptor blocks are all important for good results

[3] Feature Selection for Facial Emotion Recognition Using Cosine Similarity-Based Harmony Search Algorithm

- In this paper, the authors propose a feature selection (FS) technique called the supervised filter harmony search algorithm (SFHSA) based on cosine similarity and minimal-redundancy maximal-relevance (mRMR) to remove the redundant/irrelevant features that do not have any significant impact on the classification process in feature extraction.
- The obtained findings show that their method significantly increased the overall classification accuracy while successfully optimizing the feature vectors.

[4] Analysis of Euclidean distance and Manhattan distance measure in Face recognition

- Here, For determining how similar two feature vectors are to one another, two distance metrics—the L1 metric (Manhattan Distance) and the L2 metric (Euclidean Distance)—have been proposed in the literature.
- Manhattan distance and Euclidean distance are often employed in content-based image retrieval systems to compare two images. Here, classification is done using face photos of three participants with various expressions and perspectives.
- Experimental results are compared and the results show that the Manhattan distance performs better than the Euclidean Distance

[5] A Review on Image Feature Extraction and Representation Techniques

- In this paper, the writers give a thorough overview of image feature representation approaches and concentrate our analysis on the most recent advances in image feature extraction.
- They calculated mean, standard deviation, and skewness to find color features.

[6] Low-key features for Image retrieval based on Extraction of directional binary patterns and its oriented gradients histogram.

- The author of this study presents a unique method for retrieving images based on the extraction of low level features utilizing the Directional Binary Code (DBC), Haar Wavelet transform, and Histogram of Oriented Gradients (HOG) approaches.
- While the limited Binary Patterns (LBP) descriptor takes into account the relationship between a specific pixel and its immediate neighbors, the DBC texture descriptor captures the spatial relationship between any pair of neighboring pixels in a limited area in a particular direction.
- Last but not least, they have used Histogram of Oriented Gradients (HOG) to represent the shape and local properties of wavelet converted pictures for content-based image retrieval.

References:

[1] Ebrahimzadeh, R., & Jampour, M. (2014). Efficient handwritten digit recognition based on histogram of oriented gradients and SVM. International Journal of Computer Applications, 104(9).

[2] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) (Vol. 1, pp. 886-893). Ieee.

[3] Saha, S., Ghosh, M., Ghosh, S., Sen, S., Singh, P. K., Geem, Z. W., & Sarkar, R. (2020). Feature selection for facial emotion recognition using cosine similarity-based harmony search algorithm. Applied Sciences, 10(8), 2816.
[4] Malkauthekar, M. D. (2013, October). Analysis of Euclidean distance and Manhattan distance measure in Face recognition. In Third International Conference on Computational Intelligence and Information Technology (CIIT 2013) (pp. 503-507). IET.
[5] Ping Tian, D. (2013). A review on image feature extraction and representation techniques. International Journal of Multimedia and Ubiquitous Engineering, 8(4), 385-396.
[6] Nagaraja, S., & Prabhakar, C. J. (2015). Low-level features for image retrieval based on extraction of directional binary patterns and its oriented gradients histogram. Computer Applications: An International Journal (CAIJ), 2(1).