Computer Vision project proposal

Adheesh Trivedi*, **Aditya Sinha***

[1]Indian Institute of Science Education and Research, Bhopal
adheesh22@iiserb.ac.in adityas22@iiserb.ac.in

Problem Statement

Design and develop a high-performance software solution capable of searching for images within a directory using a provided template image. We try to address the following specific tasks:

1. **Face-based Image Search:** Given an input image containing one or more faces, identify and extract each unique face. For each face extracted from the input image, search through the directory of images to find those that contain the same person.
2. **Texture-based Image Search:** Given an input image with a distinct texture, search the image directory for images that contain the same or highly similar texture patterns. Add support for searching for common textures (e.g., fabrics, surfaces) efficiently.
3. **Class-based Search Using large Datasets:** Allow the user to provide a specific class label (e.g., "dog", "car") as a search query. Use pre-trained models on the CIFAR and ImageNet datasets to classify and find images from the directory that belong to the given class.

Key technical features to note:

- Efficient image feature extraction and matching algorithms (e.g., face recognition, texture descriptors).
- Integration of pre-trained models for CIFAR and ImageNet classification.
- Concurrent search capabilities, leveraging multithreading or multiprocessing to handle multiple search queries simultaneously.

Output: Software

Motivation

The motivation for this project is to build an efficient image search tool capable of handling large datasets by:

- Improving the image retrieval systems
- Efficient face identification
- Learning how applications like google photos create internal photo library using face detection
- Understanding what all techniques google lens may use to identify textures very easily
- Dataset exploration and Classification

Real-time image retrieval demands efficient computation. Optimizing for speed and concurrency allows for simultaneous searches without compromising on performance, making it suitable for large-scale industrial applications and services. Even in our day to day life, we often encounter tasks like searching images of dogs in your gallery. Even we encountered such tasks very often, which greatly motivated us to take up this project.

Brief related work and limitations

Face clustering

In 2015 there was work done by Google on face clustering problem (arXiv:1503.03832 [cs.CV]). They named their solution FaceNet. Their model and dataset remains closed (proprietary). However there are several open source implementations of the paper. We can take inspiration from them.

Another good paper we found on this problem is titled "Linkage Based Face Clustering via Graph Convolution Network" (GCN-based Face Clustering) (arXiv:1903.11306v3 [cs.CV]). We will also look into this approach and compare with FaceNet. FaceNet however will give better accuracy, we will also give speed of detection a good weight while comparing as we want to work on large number of images.

Texture matching

Texture analysis has evolved from traditional methods to modern deep learning techniques. Early approaches, such as Haralick features and Gabor filters, laid the foundation for extracting textural information from images. Local Binary Patterns (LBP) and wavelet transforms have also been significant in texture classification due to their ability to capture features at various scales. Recently, Convolutional Neural Networks (CNNs) have advanced the field by enabling automatic feature extraction and improving classification accuracy. However, these methods face limitations. Traditional techniques are often sensitive to noise and may struggle with multiscale textures, relying heavily on hand-crafted features that might miss critical information. On the other hand, deep

*These authors contributed equally.

learning approaches require large labeled datasets for training and are computationally expensive, raising challenges in real-time applications and interpretability.

Objects in image classification model

We will use datasets like ImageNet and CIFAR for training a model for image classification. A previous work done is very famous, AlexNet, which we will use for reference while building our CNN. We will want to identify multiple classes in a image too.

Datasets

We will use following datasets:

- **ImageNet** and **Cifar** for training our model that will search for pretraining specified classes.
- **LFW (Labeled Faces in WIld)** for measuring performance of our Face clustering algorithm.
- **Texture Datasets** Work on datasets mentioned in this link: https://github.com/J-Mourad/Dataset-crowd?tab=readme-ov-file

Algorithms and Methodology

For face clustering we will be building on algorithms posed by papers mentioned above and integrating them to our system. Also try to coin a algorithm so as to make the image retrieval fast.

For texture matching search we will compare different methods and pick which one performs best. The algorithms we will be comparing can be LBP, SIFT, Gabor Filters, CNN, Template Matching etc. Taking different texture and images in different lighting conditions, scale, rotation and occlusion.

For object classification we will go with building CNN for the given dataset.

Work plan

Team Assignments

- **Adheesh Trivedi** will focus on the problem of **Face Clustering**.
- **Aditya Sinha** will take responsibility for the problem of **Texture and Template Matching**.
- Both team members will collaborate on the **Object Classification** problem.

Project Repository Setup

- We will create a **GitHub repository** to facilitate collaboration, version control, and project tracking.
- The repository will serve as the central hub for code, documentation, and performance metrics.
- Regular commits and branch management will ensure smooth integration of individual components and ease of review.

Exploration of Techniques and Algorithms

- Each member will explore state-of-the-art techniques and algorithms relevant to their assigned tasks:
  - **Face Clustering**: Adheesh will investigate methods like *FaceNet* and *Graph Convolution Networks (GCNs)* for linkage-based clustering.
  - **Texture and Template Matching**: Aditya will explore advanced texture analysis and matching techniques using feature extraction models.
  - **Object Classification**: Both members will evaluate models trained on *CIFAR* and *ImageNet* datasets for object classification.
- The exploration phase will focus on measuring the **performance** and **speed** of these techniques using benchmark datasets.

Model Development and Evaluation

- Each member will develop the best-performing model based on their findings.
- Performance metrics such as accuracy, speed, scalability, and efficiency will be recorded and compared.
- Models will be tested and fine-tuned based on real-world use cases.

Integration into a Unified System

- After the individual models have been developed and optimized, they will be **integrated** into a common software system.
- The integration process will be carried out through the GitHub repository, ensuring seamless collaboration and component interaction.
- A modular design will allow each model (face clustering, texture matching, object classification) to function independently and in harmony within the software.

Review and Refinement

- Each contribution will be carefully reviewed by both members to ensure code quality, efficiency, and compatibility.
- Iterative development cycles will be followed, allowing for regular testing, feedback, and improvement.

Acknowledgement

We would like to express my heartfelt gratitude to Dr. Akshay Agarwal for his invaluable guidance and motivation, which encouraged us to explore innovative ideas and pursue impactful projects in the field of computer vision. His insights and enthusiasm have been instrumental in shaping our approach and inspiring us to push our boundaries.

We would also like to thank our teaching assistants for their constant support and for providing comprehensive tutorials throughout the course. Their assistance has been vital in helping us understand complex concepts and apply them effectively.

Refrences

Florian Schroff, Dmitry Kalenichenko, James Philbin 2015. FaceNet: A Unified Embedding for Face Recognition and Clustering.

Wendong Zhang; Sha Guan; Chunzhi Wang; Yucheng Zhang; Xianjing Zhou 2021.Face Detection and Face Attribute Recognition based on Deep Learning.

Manish H. Bharati, J. Jay Liu, John F. MacGregor 2004. Image texture analysis: methods and comparisons.

Ayushman Ramola, Amit Kumar Shakya, Dai Van Pham 2020. Study of statistical methods for texture analysis and their modern evolutions.

M Venkataramana E Sreenivasa Prof, C H Reddy, Venkata Ramana Mancha 2022. A Review of Recent Texture Classification: Methods.