Project Report on

Celebrity Recognition

Course - IEC 03 (AI Techniques)

Team Member details

1. **Raman Sharma**
   22114076
   +91 7521829455
   raman_s@cs.iitr.ac.in
   Contribution: Contributed to dataset preparation, searching and testing different available models which works for available dataset.

2. **Vineet Kumar**
   22114107
   +91 9263436403
   vineet_k@cs.iitr.ac.in
   Contribution: Contributed to preparing test dataset and making inference and result analysis for different models.

3. **Anvit Gupta**
   22114009
   +91 9462011044
   anvit_g@cs.iitr.ac.in
   Contribution: Contributed to dataset preparation, model architecture and coding. Comparing performance of different standard models for face detection.

4. **Souvik Karmakar**
   22114096
   +91 8642924659
   souvik_k@cs.iitr.ac.in
   Contribution: Helps in data augmentation and preprocessing steps during implementation.

1. Introduction

With the rapid advancement of computer vision and deep learning technologies, facial recognition has become a powerful tool across various applications, from security and surveillance to entertainment. This project focuses on celebrity face recognition, a specialized application of facial recognition that identifies prominent personalities from a given dataset of images.

The system aims to accurately identify celebrities from large collections of images by leveraging deep learning models and image processing techniques. This application can serve entertainment industries, social media platforms, and content creators who need automated systems to recognize public figures across vast media.

2. Project Objectives

The main objectives of the project are:

- To build a dataset containing images of celebrities.
- To preprocess and augment the dataset for model training.
- To implement and fine-tune a Convolutional Neural Network (CNN) architecture, specifically VGG16.
- To evaluate model performance and predict celebrities from new images.

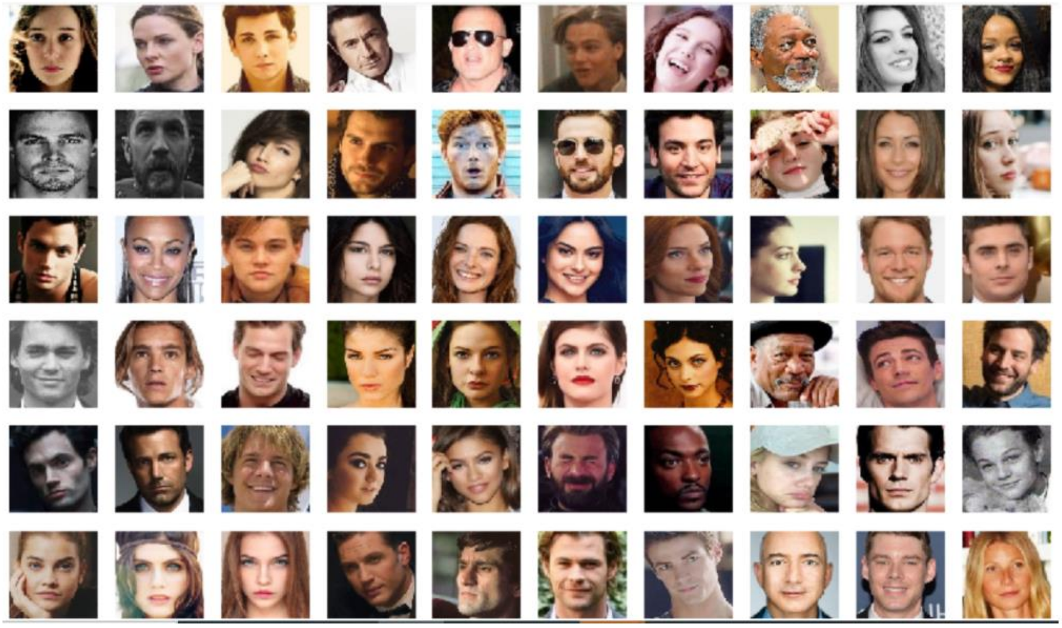
3. Dataset Preparation

The dataset consists of 13 classes, each corresponding to a different celebrity, with each class containing between 250 to 300 images.

The dataset was compiled by merging three publicly available datasets from Kaggle, ensuring a diverse and representative collection of images.

These datasets were selected for their diversity and quality, ensuring a rich variety of facial expressions, poses, and lighting conditions.

The combined dataset provides a robust foundation for training our face recognition model, enhancing its ability to generalize across different scenarios.



4. Methodology

The methodology of the project can be broken down into several key steps:

4.1. Data Preparation

The initial step involved preparing the dataset:

- **Dataset Unzipping**: The dataset was unzipped to access the images.
- **Directory Structure**: Images were organized into separate folders for each celebrity.

4.2. Dataset Splitting

A Python function was implemented to split the dataset into training and validation sets, ensuring a balanced distribution of classes. This involved:

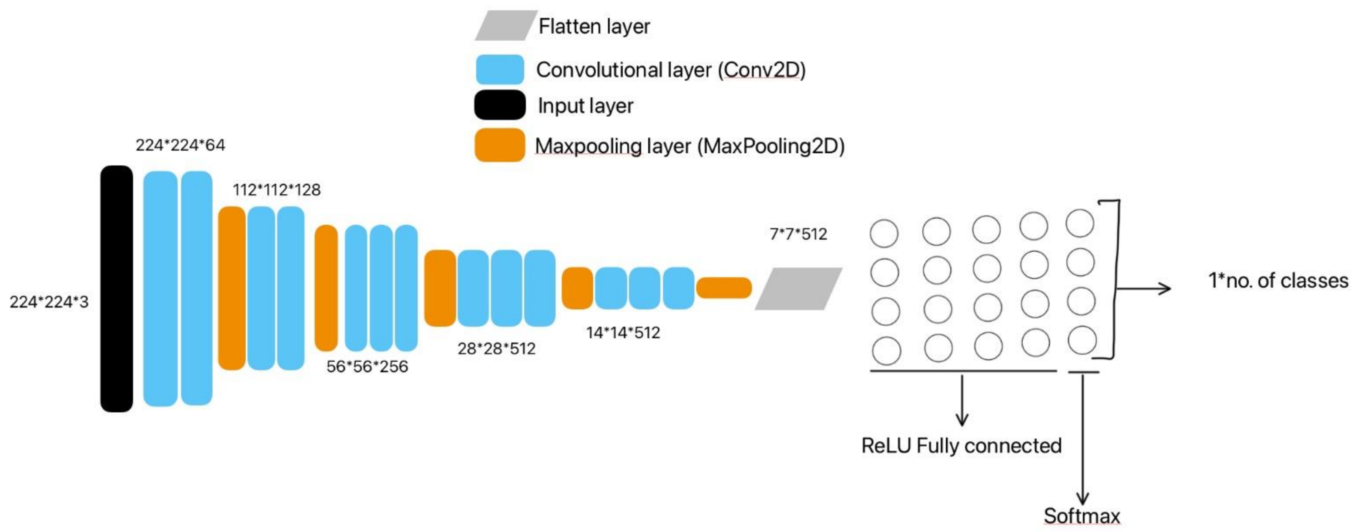
- Iterating through each celebrity folder.
- Randomly dividing images into training (80%) and validation (20%) sets.

4.3. Data Augmentation

To enhance the model's ability to generalize, data augmentation techniques were applied using Keras' **ImageDataGenerator**. This included:

- Rescaling pixel values.
- Randomly rotating, shifting, shearing, and flipping images.

4.4. Model Architecture



The project utilized the VGG16 architecture for feature extraction. Key steps included:

- Loading the VGG16 model with pre-trained weights (ImageNet).
- Freezing the initial layers to retain learned features while training the upper layers.

4.5. Model Training

The model was trained in two phases:

1. **Initial Training**: The model was trained for 10 epochs with a learning rate of 0.0001.
2. **Fine-Tuning**: The last four layers of the base model were unfrozen, and training continued for another 10 epochs with a reduced learning rate of 0.00001.

4.6. Model Evaluation

Training and validation loss and accuracy were plotted to monitor performance. Metrics were analyzed to ensure the model did not overfit.

5. Results

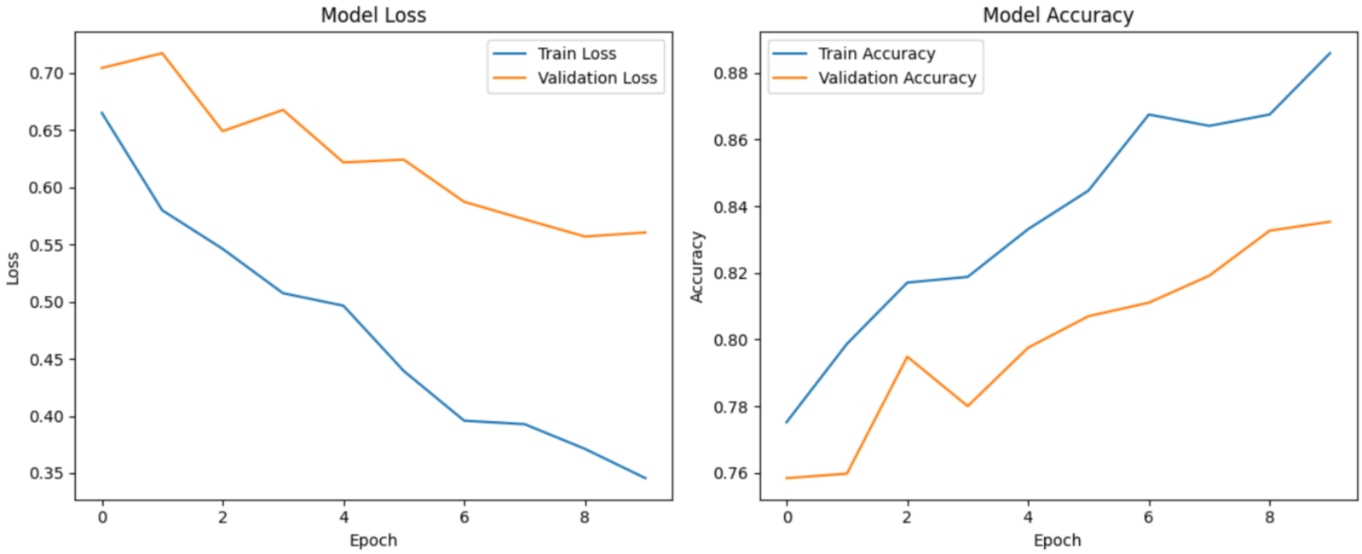
The model achieved significant performance metrics, with training and validation accuracy improving across epochs. The loss curves indicated convergence, validating the training process.

Training Accuracy – 88.59 %

Validation Accuracy – 83.54 %

Training Loss – 0.3456

Validation Loss – 0.8354



The validation accuracy of the model will increase if dataset is large and chances of overfit will reduce.

6. Inference

For testing the model's effectiveness, we defined functions to load and preprocess images for prediction. The model was then used to predict celebrity identities from a test directory containing images of various celebrities.

The prediction process involved:

- Loading images, resizing them, and normalizing pixel values.
- Using the trained model to generate predictions and mapping them back to the corresponding celebrity names.

7. Conclusion

The "Celebrity Face Detection" project successfully demonstrated the effectiveness of deep learning in image recognition tasks. The implementation of VGG16 architecture and data augmentation techniques resulted in a robust model capable of accurately identifying celebrities from new images.

By leveraging CNN architectures and fine-tuning them with a celebrity dataset, the model was able to identify celebrities in a variety of settings with high accuracy. This project highlights the potential of face recognition in entertainment and media management, offering insights into further optimizations and applications.