Milestone 3

Alison Yao, Isidora Diaz, Luis Henrique Simplicio Ribeiro, Isabella Bossa, Elie Attias

April 2023

1 Summary of the Data

The dataset we are using in this project is the Large-scale CelebFaces Attributes (CelebA) Dataset . CelebFaces Attributes Dataset (CelebA) is a large-scale face attributes dataset with more than 200K celebrity images, each with 40 attribute annotations. The images in this dataset cover large pose variations and background clutter. CelebA has large diversities, large quantities, and rich annotations. It includes:

- Number of celebrities: 10,177
- Number of images: 202,599

1.1 Sample Images

A collection of sample images can be found in figure 1.

1.2 Data Files

A histogram displaying the number of images per person can be shown in figure 2. The data files we will consider are the following :

1. `img_align_celeba.zip`: All the face images, cropped and aligned
2. `list_eval_partition.csv`: Recommended partitioning of images into training, validation, testing sets. Images 1-162770 are training, 162771-182637 are validation, 182638-202599 are testing.
3. `list_bbox_celeba.csv`: Bounding box information for each image. "x_1" and "y_1" represent the upper left point coordinate of the bounding box. "width" and "height" represent the width and height of bounding box
4. `list_landmarks_align_celeba.csv`: Image landmarks and their respective coordinates. There are 5 landmarks: left eye, right eye, nose, left mouth, right mouth

2 Deeper Understanding of the Data

We observed class imbalances in the dataset

1. Some celebrities were present in as many as 35 images, while others only had 1 picture that included them as shown in figure 2.
2. Slightly over 58% of the observations belong to female celebrities, while the remaining 42% are from male celebrities. Although there is a slight class imbalance here, the proportions are relatively close to those of the general population and the difference between the two ratios is not too large, so we believe this will not be a problem.
3. Some attributes are more present than others. For example, less than 5% of the images contain bald celebrities. This potential domain shift would only be an issue if we were planning to use the model with a subset of celebrities that had a different distribution of the target attribute.
4. It is difficult to identify outliers in image datasets without manually exploring the data. Similarly, it is difficult to identify the presence of patterns in images without an already trained model that can identify image trends.

Figure 1: Sample Images

3 Clean and Labeled Visualization

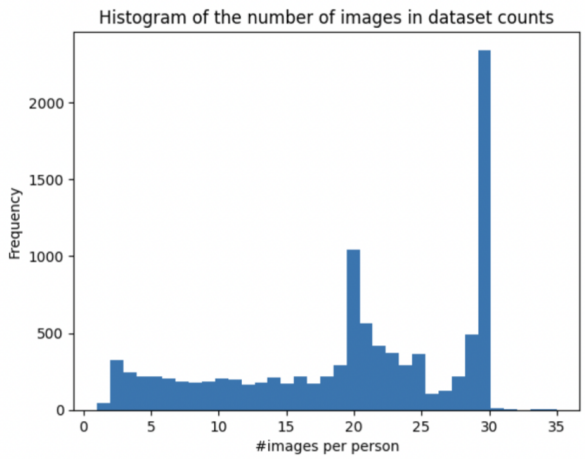
We processed the data in the following way. We first normalised the pixel values to ensure that all pixel values are between 0 and 1. We also augmented our data to enhance our final model's robustness. Such processing is displayed in figure 3.

4 Project Questions

1. Is it possible to utilize transfer learning in order to develop a facial recognition model capable of accurately categorizing celebrity faces sourced from the CelebA dataset?
2. What degree of resilience does the model exhibit in the presence of images featuring high levels of noise?
3. What degree of resilience does the model exhibit in the presence of adversarial inputs?
4. How fast can our model both find faces and classify them in a high-frequency frame video?

5 Meaningful Insights & Noteworthy Findings

1. Most of the celebrities have multiple images in the dataset, so we only need to deal with 44 celebrities who only have one image in the dataset.
2. Images are of different sizes, so when we train the model, we will need to convert them into the same size either by padding, cropping, downsampling, upsampling etc.
3. Both in-the-wild image data and face-aligned data are provided in CelebA, so we can use the in-the-wild as they correspond to the image types we will encounter in our final classification task, where we intend to classify celebrity faces in movies i.e 'in the wild'



	Count
count	10177.000000
mean	19.907537
std	8.925916
min	1.000000
25%	13.000000
50%	21.000000
75%	29.000000
max	35.000000

Figure 2: Histogram of the Number of Images in Dataset Counts

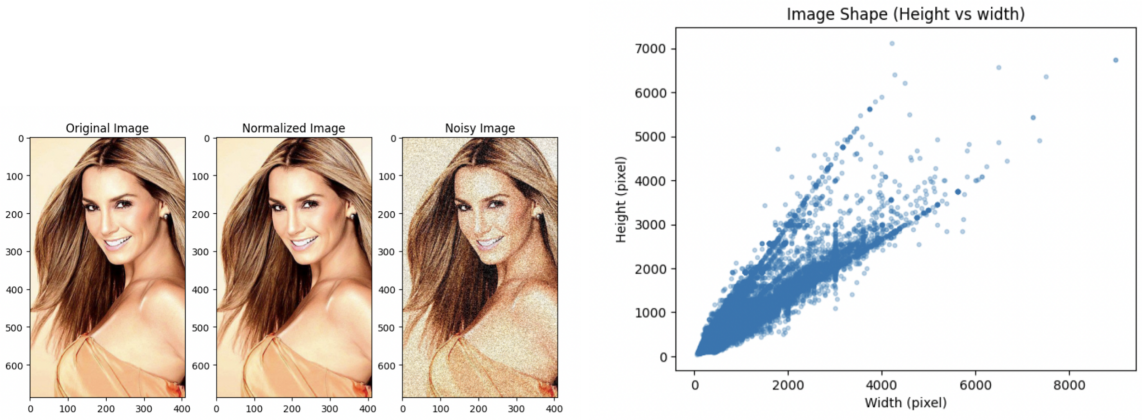


Figure 3: Data Augmentation visualisation and distribution of images shapes

6 Implementation Plan

1. Preprocess the data,
2. Utilize the optimal YOLO model weights for object detection and import them,
3. Implement transfer learning to train our model to detect bounding boxes around celebrity faces in CelebA dataset or use a pre-trained model for this task,
4. Apply the face detection model to the "in the wild" data to extract images of the faces of the celebrities and create a new trained/test dataset focused on recognizing celebrity faces
5. Perform a secondary image processing step to address any inhomogeneities in the YOLO-transformed dataset,
6. Model Development: Develop a contrastive learning-based model to classify facial celebrity images
7. Model Evaluation:
   (a) Test the model's ability to identify and classify celebrity faces in real-time videos
   (b) Evaluate the model's performance on the test set and noisy representations of the train/test set