Faculty of Mathematics and Information Science
WARSAW UNIVERSITY OF TECHNOLOGY

Introduction to Machine Learning (IML) Project

Warsaw,

David Abraham
Tsitsi Nyamutswa
Akhil Ajith Ambalapurath
Daniel Ndashimye

Contents

1. Task Description

The goal is to develop a desktop computer vision application for human age detection. The following list provides the deliverables that we aim to produce by the end of this project:

- A suitable dataset for human age detection
- Efficient data handling from image/camera/video input
- Exploratory Data Analysis to gain insight into the dataset
- Appropriate Image Augmentation to ensure best possible results
- Fully developed face recognition algorithm
- First developed Age Detection Algorithm (milestone 2)
- Partially completed GUI application (milestone 2)

2. Facial Recognition Algorithm

For the face detection ad recognition we used Multi-task Cascaded Convolutional Networks (MTCNN) and Facenet, respectively. MTCNN allowed us to detect faces and add a bounding box for each detected face and the FaceNet model from the facenet-pytorch library is used specifically for feature extraction and recognition of the the faces.

UPDATE: After creating our model, we realized that the new model did not work well with the MTCNN facial detection. We decided to implement this using the Blazeface face detection model from the mediapipe library for detecting faces in each frame of a video or webcam feed.

A set of images and videos were randomly selected and evaluation is based on whether or not a human face is detected in the image, and if so a bounding box is added. The following ratio is then computed:

$$\text{Ratio of non-human faces detected} = \frac{\text{Number of non-human faces detected}}{\text{Total Number of Images}}$$

We evaluated two sets of images, one with animal faces and one with human faces, each with 100 images. The human faces where of ages 10-50 and the animal images included lions, cats, bears, cheetahs, tigers, monkeys etc. We obtained the following results as percentages:

	Set of Human Faces	Set of Animal Faces
Non-human faces	10%	81%

Table 2.1. Accuracy of Facial Recognition Algorithm on Animal Faces and Human Faces

The comparison showed that when our face recognition algorithm is run on images and videos of animals there is a 10% chance of some parts of the animal being detected as a human face. This was especially visible on the noses of some animals as well as the prints on the tigers and cheetahs. However when it came to the images of humans there was only a 5% chance that the face was not detected.

3. Exploratory Data Analysis

Exploratory Data Analysis (EDA) is a critical phase in the data analysis process that involves understanding the main characteristics of a dataset. The primary goal is to gain insight into the characteristics of the dataset by employing various techniques.

3.1. Cleaning the dataset

We used the UTKFace dataset as our starting point and went through the raw dataset examining the pictures in each folder. Our main goal was to ensure that each image was clear and the persons face was not obstructed. At this point it became very clear that our dataset was imbalanced, this is shown clearly in the graph below:

During this preprocessing phase, we also split our dataset into multiple folders to organize images based on key demographic attributes, specifically age, ethnicity and gender in order to allow for any targeted analyses in future training of the model. The distribution of this data across the categories is not uniform with certain subgroups over-represented , while others are under-represented. This non-uniform distribution can significantly impact the performance of our model. To illustrate this distribution we have generated graphs for each ethnicity category to provide insight into this:

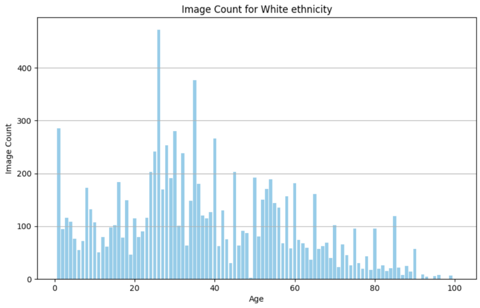

Figure 3.2. Distribution of White Demographic in dataset

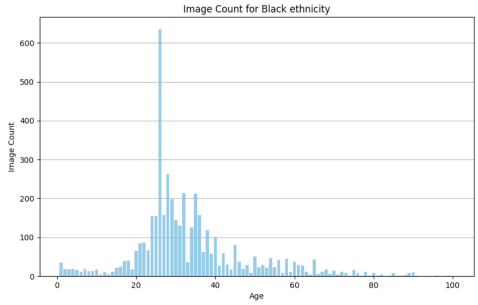

Figure 3.3. Distribution of Black Demographic in dataset

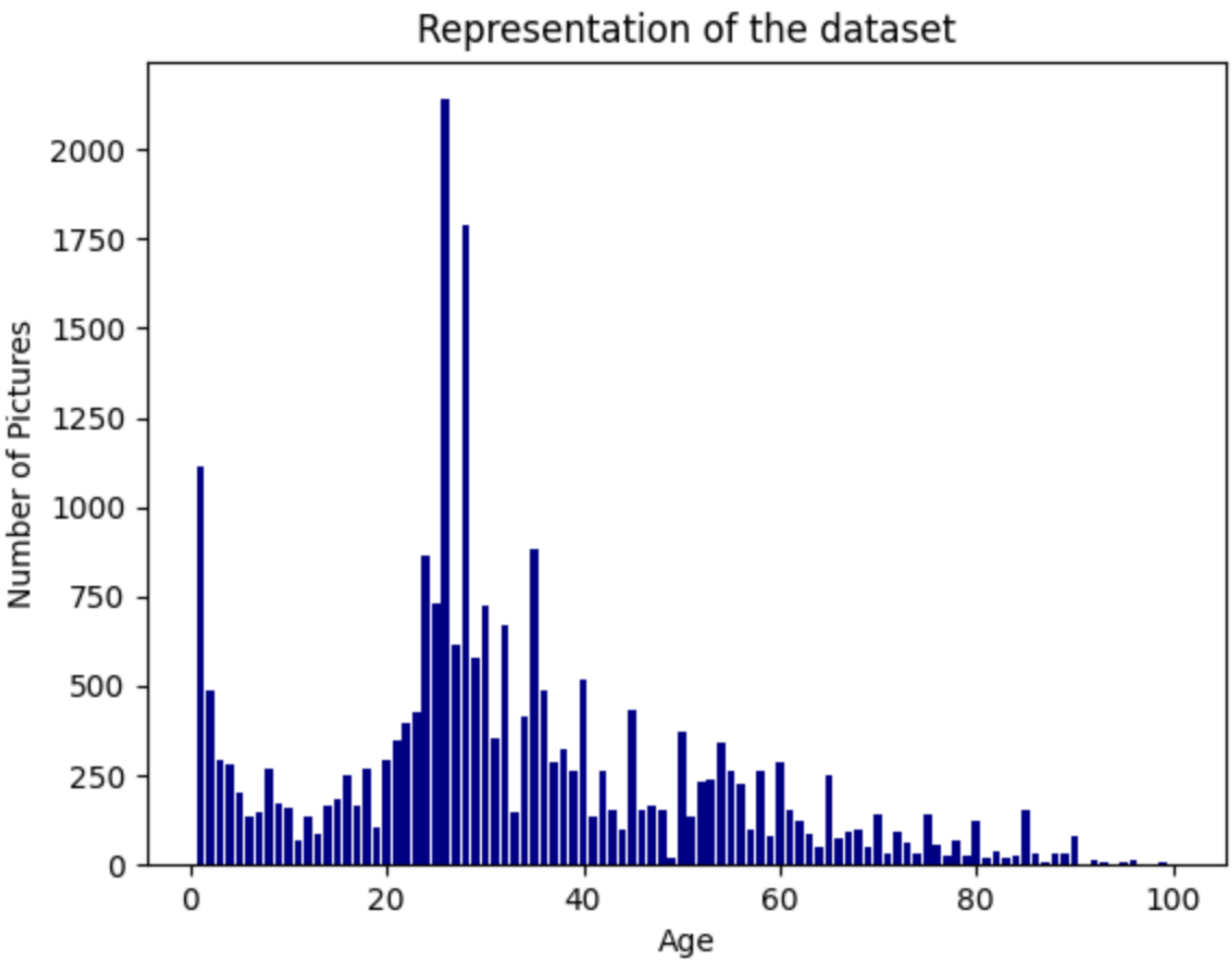

Figure 3.1. Representation of the Dataset after cleaning

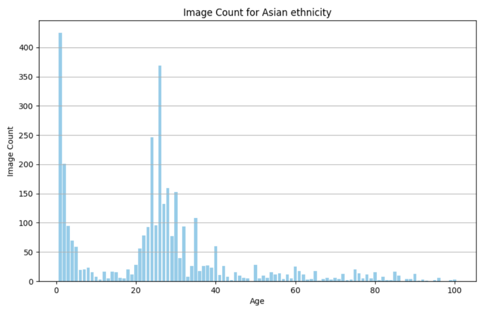

Figure 3.4. Distribution of Asian Demographic in dataset

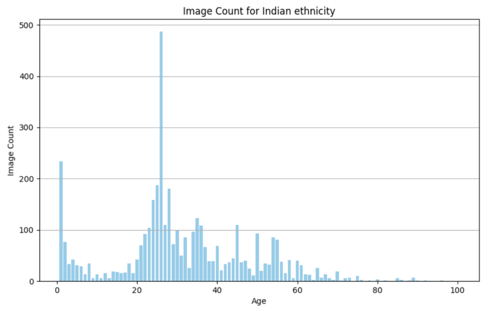

Figure 3.5. Distribution of Indian Demographic in dataset

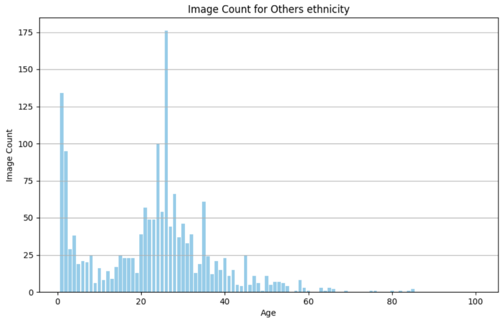

Figure 3.6. Distribution of 'Other' Demographic in dataset

The figures above reveal non-uniform distributions of various ethnicities across different age groups. Among these distributions, the White Representation stands out with a more uniform pattern, distinct from other ethnicities that exhibit peaks within the age range of 23-26. Notably, for ages 60 and above, the demographic is predominantly composed of the white ethnicity. This uneven distribution may adversely affect our model's performance in age detection for non-white ethnicities within that specific age demographic.

The accompanying graph details the number of individuals within each ethnicity, we observe the variations in our dataset. This emphasizes the dataset's diversity and helps with further understanding of the studied population which could potentially affect subsequent analyses.

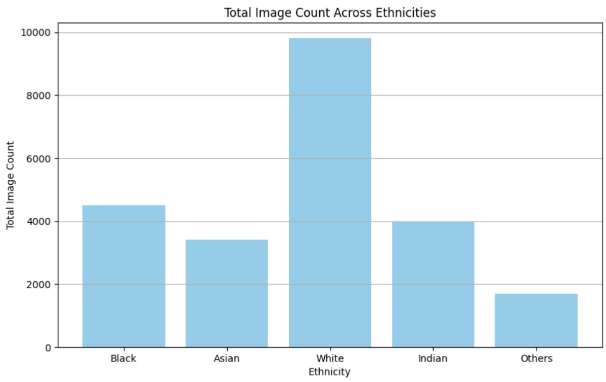

Figure 3.7. Distribution of all Ethnicities

3.2. Selecting subsets for Exploratory Data Analysis

The next stage involved selecting a subset of 100 images from the young people, and another of 100 from the old people. To do this we tried 3 different pairs of subsets and this helped us to visualize and compare the normalised histograms of the red, blue and green channels for these 100 images in both groups. The pairs we used are as follows:

1. Young people between ages 10-25 and Old people between ages 61-88
2. Young people between ages 1-10 and Old people between ages 84-100
3. Young people between ages 3-5 and Old people between ages 69-88

when we plotted these three pairs into normalised histograms we achieved the following results:

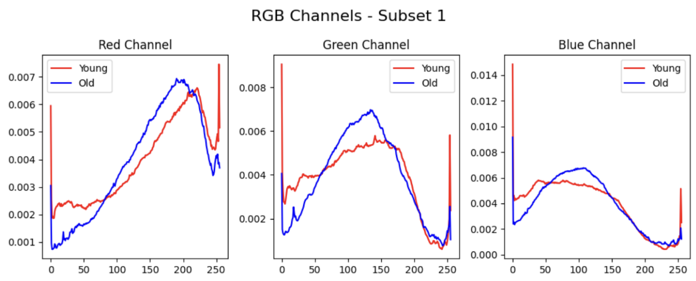

Figure 3.8. Young: 10-25 Old: 61-88

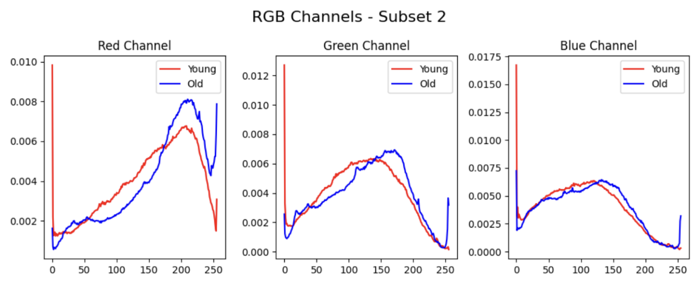

Figure 3.9. Young: 1-10 Old: 84-100

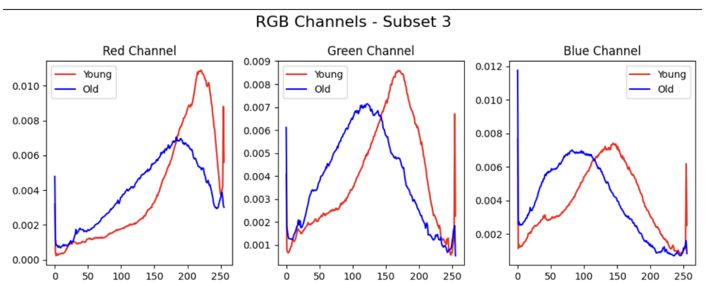

Figure 3.10. Young: 3-5 Old: 69-88

The clearest deviation was from subset 3 and so we applied multiple augmentation methods to this specific subset with the goal of creating a bigger gap between the color channels of the Young and Old people.

3.3. Image Augmentation

In this stage, we implement augmentation techniques to introduce variations in the appearance of images, with the goal of improving the model's ability to make appropriate predictions across different ages. Our augmentation pipeline involves a set of transformations, including rotation, flipping, addition of noise, contrast enhancement, color jittering, zooming, edge enhancement, blurring, and conversion to grayscale. The goal is to create a more comprehensive and representative dataset by enhancing the overall performance and reliability of out machine learning model. The augmentaion methods that showed the best results are recorded below.

1. **Greyscale Augmentation**
   This involves converting images to black and white to reduce the impact of color information. This method accentuates facial features, enhances contrast and potentially helps with age-related feature extraction.
2. **Zoomed Augmentation**
   We introduce variations in the image scale in order to emphasize facial details provideing the model with broader perspective on age related visual cues.
3. **Split and Mirror Augmentation**
   This method involves splitting the images vertically and mirroring one half in order to create a symmetrical effect. Through this we explore how symmetry impacts age detection as symmetry is typically associated with youthfulness.

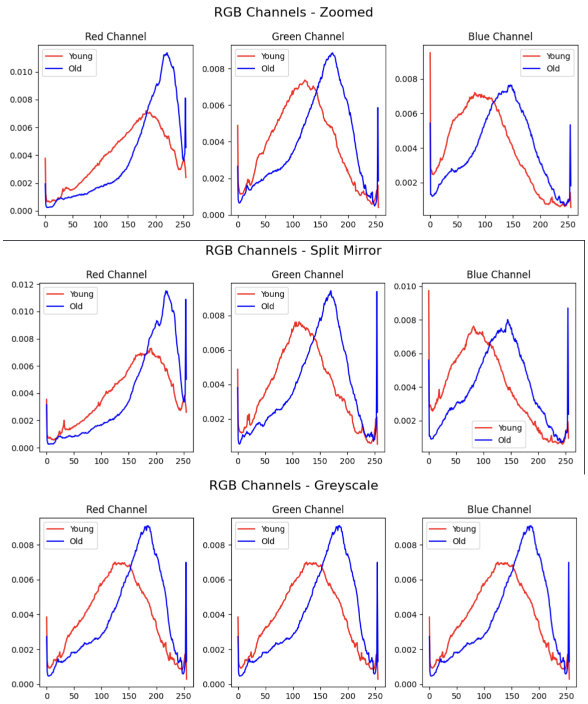

Figure 3.11. Results of Image Augmentation

Combining Greyscale and zoomed augmentation allowed us to remove color distractions while the zoomed augmentation introduced scale variations. This combination seeks to enhance the model's ability to discern age-related patterns in a more focused and nuanced manner. The graph below show the result of this combination:

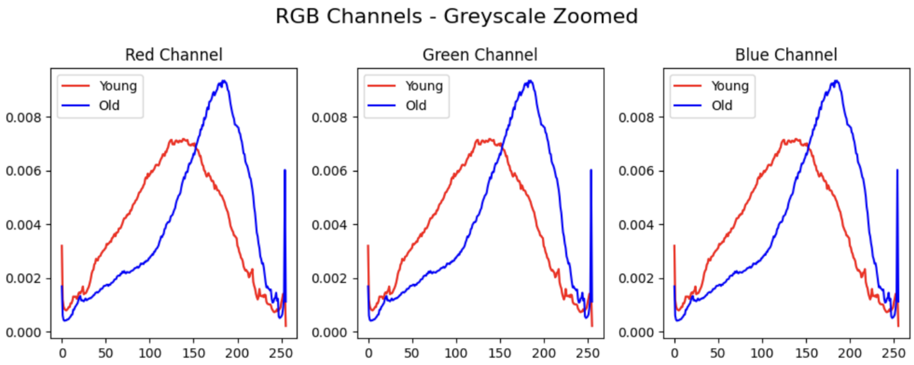

Figure 3.12. Zoomed Greyscale

4. Bibliography - Model

- Lecture Notes
- github/deepankarvarma
- Medium.com