Age Detection Using Convolutional Neural Networks

A. Introduction

Advancements in computer vision have enabled applications like facial recognition and image classification, with age detection emerging as a vital tool for verifying age from facial images. However, age detection is complex due to variations in facial features, expressions, lighting, and ageing patterns. Traditional methods, relying on hand-crafted features and machine learning models like Support Vector Machines (SVMs), often lacked accuracy in capturing the subtle changes associated with ageing.

Convolutional Neural Networks (CNNs) have transformed age detection by automatically learning and extracting meaningful features from raw images, handling variations more effectively than traditional methods. CNN-based age detection has applications in verifying age for websites with mature content, offering senior citizen discounts in e-commerce, and ensuring compliance with age-restricted products like alcohol in vending machines. This report explores various CNN architectures, including VGGNet [1], DenseNet [2], and EfficientNet [3], evaluating their performance on age detection datasets. It compares factors like dataset size, computational resources, and real-time processing requirements to recommend fine-tuning techniques and the most suitable architecture for practical deployment.

B. Datasets

The study uses three datasets found online as a part of previous studies on image classification. These datasets consist of both processed as well as “in the wild” images which need to be processed. The first dataset, namely *UTKFace* [4] is a popular dataset which consists of 20,000 images. The age groups in this dataset range from 0 to 116-year-olds. The second dataset namely *Facial Age* [5] consists of 9,778 images with ages ranging from 0 to 90-year-olds. The third dataset i.e. *Adience* [6] consists of 26,580 images with eight age groups between 0 and 90 years. Two out of the three datasets don’t have pre-defined age group classifications so before training the models, the images of these datasets are mapped to age groups defined by us by extracting age from the image labels and assigning the class to them. For the UTKFace [4] dataset we aim to divide images into 20 classes and for the FacialAge [5] we aim to divide the images into 4 classes. The project moves on to align and crop all images of the dataset to bring consistency to all three datasets. The cumulation of all three datasets gives the model a diversity which will help in enhancing the results of the entire study. The datasets consist of a humble mix of images from different age groups, sexes, racial backgrounds and ethnicities thus providing a diverse training set for the model.

C. Methodology

In this section we go through the data preprocessing steps, the models used to tackle our problem, and the evaluation metrics for the said models.

C.1. Data Preprocessing

The data preprocessing stage consists of multiple stages in order to streamline the images as input. The initial task is to clean the dataset and get rid of any images which have bleeding edges or have extreme lighting conditions. We then apply techniques to standardize all images like random cropping (to crop the faces) and reorientation, flipping or rotating wherever needed, and data manipulation (brightness and contrast normalization). Finally, we normalize the pixel values to a common range for consistent input and output.

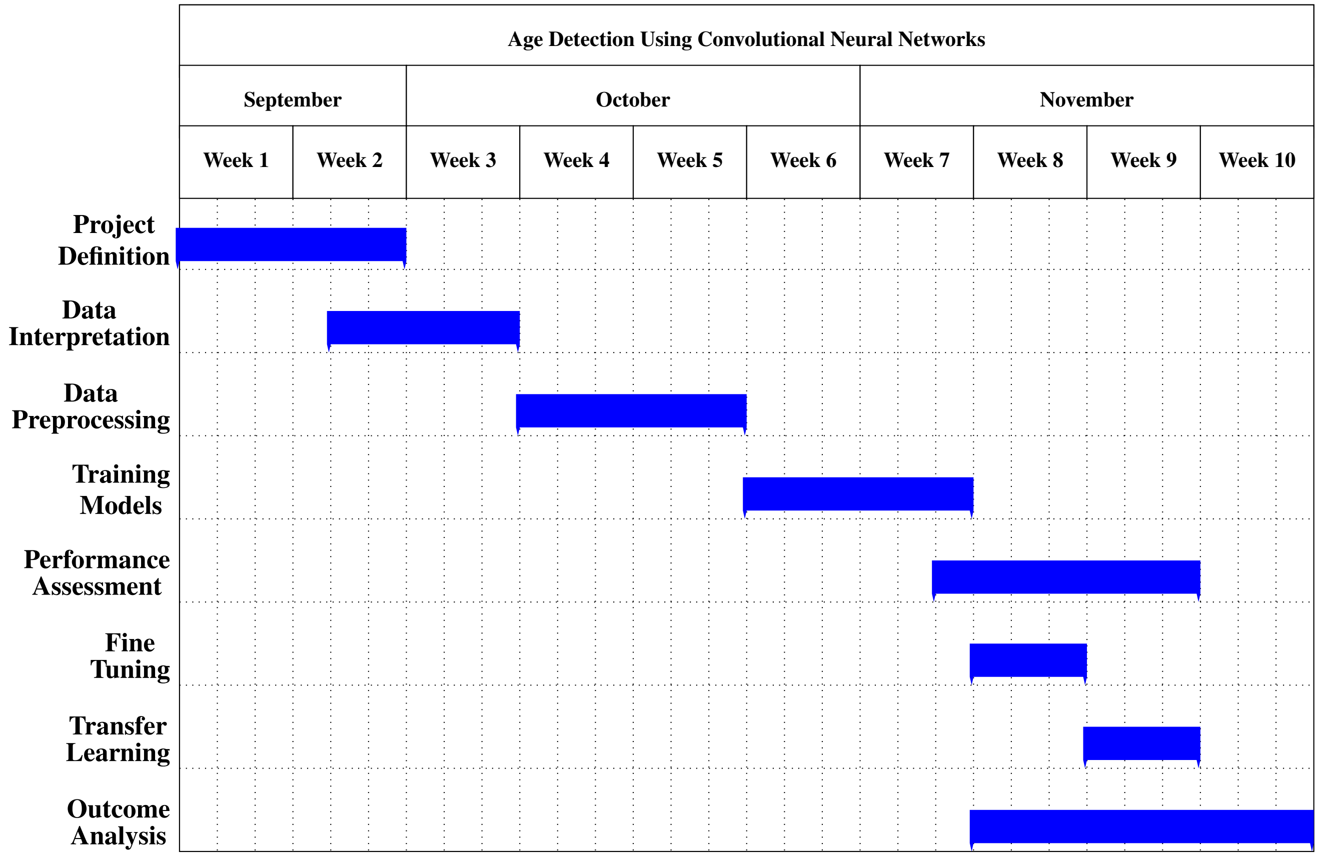
C.2. Model Development

In this study, we employ three deep-learning architectures for age classification. We choose VGGNet [1] as the first test architecture, utilizing convolutional layers to extract features and pooling layers to downsample, reducing computational costs. Secondly, we use DenseNet [2] as it enhances feature propagation through dense connections, allowing earlier layers to capture basic features like edges, while later layers focus on details such as skin texture and wrinkles. The third architecture, EfficientNet [3], is included for its strong ability to recognize visual patterns. Additionally, we apply transfer learning by initializing two models with pre-trained weights from the earlier architectures, improving performance and reducing training time. ResNet [7] is incorporated to address the vanishing gradient problem with residual layers, which helps interpret blurred images more effectively. Lastly, ShuffleNet [8], known for its lightweight design and efficient feature extraction, is utilized to support transfer learning in computationally constrained environments. Testing the problem with multiple architectures provides a comprehensive approach, balancing accuracy and efficiency for age classification tasks and allows us to conclude by drawing comprehensive comparisons across architectures.

C.3. Evaluation Metrics

We train each model with all three datasets after the preprocessing to bring forward varying results. The results are then compared with each other using metrics such as accuracy, precision, recall and F1 score. We also try to draw up a confusion matrix to assess what features from the image are contributing to the model prediction. We also aim to use TSNE and Grad-CAM to visualize the performance of our models.

Gantt Chart



References

[1] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015. 1

[2] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely connected convolutional networks, 2018. 1

[3] Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural networks, 2020. 1

[4] Zhifei Zhang, Yang Song, and Hairong Qi. Age progression/regression by conditional adversarial autoencoder, 2017. 1

[5] Fazle Rabbi. Facial age dataset. https://www.kaggle.com/datasets/frabbisw/facial-age. 1

[6] Eran Eidinger, Roee Enbar, and Tal Hassner. Age and gender estimation of unfiltered faces. *IEEE Transactions on Information Forensics and Security*, 9:2170–2179, 2014. 1

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. 1

[8] Xiangyu Zhang, Xinyu Zhou, Mengxiao Lin, and Jian Sun. Shufflenet: An extremely efficient convolutional neural network for mobile devices, 2017. 1