TOWARDS AN ETHNIC-BIAS FREE APPROACH
FOR FACIAL RECOGNITION

Thilan Costa

A dissertation submitted in partial fulfilment of the requirements for
MSc Advanced Software Engineering Degree in Computer Science

Department of Computing
Informatics Institute of Technology, Sri Lanka
in collaboration with
University of Westminster, UK

2021

Abstract

Mostly developed countries use **Facial Recognition** to fight terrorism and to keep their country safe from various incidents like bomb blasts or crimes. So, in those countries facial recognition is widely used by security forces like the military and police. Likewise, most of the private and government institutions and companies also use facial recognition to identify their employees and for security purposes.

One of the problems of facial recognition is identifying false faces as positives. This problem is happening throughout the ethnic minorities according to some studies. This problem justifies that there is ethnic bias in facial recognition systems and algorithms which are being used currently. This dissertation is a result of the project which aims to explore the ethnicity bias issue on the current state of the art facial recognition algorithm Facenet and to provide a novel image selection method to minimize the effect from the ethnicity bias without making changes to the neural network architecture

This approach modifies the existing facenet approach while classifying the ethnicities of images of the dataset to minimize the ethnic bias of the facenet. This approach considers the ethnic distribution of the given data set and uses that to create the small batches of face images which will be used to train the model. The model trained with the proposed method has shown significant improvement when considering the AUC metric which reduces the chance to misclassify a person with underrepresented ethnicity. Thus, reducing the false acceptance rate which increases the reliability of the proposed model as well as minority representation. The clustering evaluation tasks revealed that the model trained with the proposed approach can form near-perfect clusters while the base model struggles to do so. The solution is available as multiple jupyter notebooks which will run on google colab.

Keywords: Facial recognition (FR), Ethnic-bias, Ethnic distribution, Ethnicity classification.

i