Project Proposal
Bias Analysis in Facial Expression Recognition (FER) Datasets

Gergo Gyori **Katalin Literáti-Dobos**

1 Introduction

Facial Expression Recognition (FER) systems are increasingly adopted across industries like healthcare, human-computer interaction, and security. However, these systems often rely on datasets that may introduce biases, such as imbalances in demographic representation, gender, and cultural diversity, which can impact the fairness and accuracy of model predictions.

This project aims to identify and analyze potential biases in commonly used FER datasets, examining their implications for real-world applications. To ensure relevance, the analysis will focus on recent high-impact papers (2020-2024) with over 100 citations.

2 Objectives

- To analyze the usage patterns of widely cited FER datasets in research papers from 2020 to 2024.

- To identify the most frequently used FER datasets and analyze their reported performance (accuracy, F1 score, etc.).

- To investigate the possible biases within these datasets, including demographic, gender, and cultural biases, and assess their impact on FER model outcomes.

3 Methodology

3.1 Identify Widely Used FER Datasets

- Data Source: Use academic paper repositories (e.g., Google Scholar, OpenAlex, or arXiv) to collect research papers on FER from 2020-2024.

- Inclusion Criteria: Papers that have been cited over 100 times, and that explicitly use FER datasets.

- FER Datasets: Identify the most commonly used FER datasets (e.g., CK+, FER2013, AffectNet, etc.) based on the frequency of use in the papers.

3.2 Analyze Paper Performance Metrics

- Extract performance metrics (e.g., accuracy, F1 score) reported in the papers for different FER models trained on these datasets.

- Summarize the reported performance across models and datasets to establish benchmarks.

3.3 Assess Potential Biases

This part of the project will be finalized once the relevant datasets are identified. If dataset descriptions lack sufficient detail, we will download the datasets and generate statistics to assess potential biases effectively.

- Demographic Bias: Analyze whether datasets include a diverse range of ethnicities, age groups, and genders.

- Cultural Bias: Assess whether datasets are over-represented with facial expressions from specific cultural backgrounds, potentially limiting the generalizability of models.

- Gender Bias: Check the representation of male and female subjects in the datasets. Bias Impact Analysis: Investigate how the identified biases might affect model performance and fairness, especially when applied to diverse populations.

3.4 Evaluation

- Perform a literature review on bias detection methods used in FER.

- Investigate any existing bias mitigation techniques applied in papers using the identified datasets.