A Computational Study of
Cross-Cultural Facial Analysis Bias

Background and Motivation

Many studies have found that facial recognition algorithms are largely biased towards Caucasian men, and are more likely to misidentify the faces of non-Caucasian individuals. A study on facial analysis software by Joy Buolamwini from the MIT Media Lab showed that gender was misidentified in 35% of darker-skinned females and 12% of darker-skinned males, as compared to 7% in lighter-skinned females and 1% in lighter-skinned males [5]. These facial recognition softwares have a variety of application from law enforcement to social media analysis, which makes the repercussions of unintentional algorithmic bias drastic for groups that these algorithms are faulty in identifying. In a similar vein, we believe computational facial analysis software, such as OpenFace, may have inevitably inherited the biases reflected by a largely Caucasian and Eastern Asian team of researchers. As the field of facial recognition grows in prominence, particularly with respect to affective computing, it is important to address the biases these softwares potentially have and either fix them or acknowledge their incompleteness.

We chose to analyze South Asians for two reasons: (1) South Asians have often been combined into one category titled "Asian / Asian American" alongside Eastern Asians, an ethnic group that has remarkably distinctive features from South Asians; (2) South Asians are largely underrepresented as a group in facial expression research or collection (CAS-PEAL, HKPolyU, CUHK, TFEID). For context, the term "East Asians" can classify individuals with origins from, but not limited to, China, Japan, Taiwan, and Hong Kong, whereas the term "South Asians" can refer to those with origins from, but again not limited to, Afghanistan, Bhutan, India, Nepal, and Pakistan. Clearly, there is rich diversity even within these two categories, and so we imagine that generalizing across the term "Asian" while addressing the ethnic groups within can lead to unintentional algorithmic bias. For the purpose of this class project, we came to the decision that we would exclude Eastern Asians from this study because there can be a large origin overlap between Eastern and Southern Asians that we may not be able to account for, and we hope to elucidate the merits of South Asians independently. Thus, we will compare facial features between South Asians and Caucasians because many of these algorithms are optimized for Caucasian features [5].

Related Works

There are a range of common differences we expect to observe between ethnic and cultural backgrounds, including variance in skintones, group prosody, shared dialects, and even our belief systems, but one distinction that's not ordinarily considered is our facial expressiveness to a shared emotional state.

There is a widely-held assumption that people from all ethnic backgrounds can express an emotional state with the same facial expressions. In the 1960s, Paul Ekman traveled to a remote jungle in Papua New Guinea and discovered that tribes that were untouched by western civilization exhibited six "universal" facial expressions [1]. He also went on to develop an emotion recognition test only using Caucasian actors [2]. Later studies have looked into differences between cross-cultural facial expressions [3,8], cross-cultural recognition of emotional expression [4,6], and have addressed gender and skintone algorithmic-bias in widely-available, facial analysis software [5]. Understanding that there is varied perception of the term "Asian" and whether South Asians fit into that category [7], we expect datasets titled "Asian" to likely be biased with more Eastern Asian subject data.

Description of Project, Steps To Complete

Our group will be exploring the inter- and intra-cultural differences in affect recognition between two ethnic groups, South Asians and Caucasians. We will study how, according to an emotional model of choice such as PAD, both groups rate photos of members from their own group and photos of members from the other group. We will also use one or more computational tools to rate the same photos of both groups on the same emotional state model. Our choice of computational tool will heavily be influenced by what analyses it and our participants can jointly score (eg. rating on a PAD scale or basic emotions, but AUs would be impractical for our participants to score). Given these ratings from both ethnic groups and the computational tool, we will be able to compare and contrast our results to test against our hypotheses.

Before our study can be completed, we will need to accomplish a few key tasks. First and foremost, we will need to gather a set of photos of members from the two ethnic groups. To limit confounding variables, we will use photos of each group with the same, posed negative emotion [citation 8 addresses recognition of negative facial expressions]. This specific emotion will be chosen by findings in relevant readings on posed facial expressions in lab studies. We will first check to see if these photos can be found from existing databases online, but if not, we will capture them ourselves. Next, we will need to choose an emotional state model and a computational tool for recognizing affective properties on this model. These will be chosen through a thorough search of open-source tools and an understanding of what emotional models they utilize.

Expected Results

We hypothesize that members of different ethnic groups express and recognize affect differently due to cultural and physical differences. We also hypothesize that some, if not all, of the computational tools used to recognize and identify affect in humans underestimate the effects of these cultural differences, and thus derive incorrect or even biased results.

Citations

[1] Ekman, Paul. "Handbook of Cognition and Emotion, Ch 3." *Basic Emotions*, Paul Ekman, 1999, www.paulekman.com/wp-content/uploads/2013/07/Basic-Emotions.pdf.

[2] Ekman, Paul. "Pictures of Facial Affect (POFA)." *Paul Ekman Group*, 1993, www.paulekman.com/product/pictures-of-facial-affect-pofa/.

[3] Ekman, P. (1972). Universals and cultural differences in facial expressions of emotions . In J. Cole (Ed.), Nebraska Symposium on Motivation, 1971 (pp. 207–283). Lincoln: University of Nebraska Press.

[4] Boucher, J. D., & Carlson, G. E. (1980). Recognition of facial expressions in three cultures. Journal of Cross-Cultural Psychology , 11 , 263–280.

[5] Hardesty, Larry, and MIT News Office. "Study Finds Gender and Skin-Type Bias in Commercial Artificial-Intelligence Systems." *MIT News*, 11 Feb. 2018, news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212.

[6] Schimmack, U. (1996). Cultural influences on the recognition of emotion by facial expressions: Individualistic or Caucasian cultures? Journal of Cross-Cultural Psychology , 27 , 37–50.

[7] Wang, Alton, and Karthick Ramakrishnan. "Who We Count as Asian American: #BrownAsians." *Data Bits*, 20 Nov. 2017, aapidata.com/blog/who-we-count-as-asian-american-brownasians/.

[8] Jack, Rachael E., et al. "Cultural Confusions Show That Facial Expressions Are Not Universal." *Current Biology*, vol. 19, no. 18, 2009, pp. 1543–1548., doi:10.1016/j.cub.2009.07.051.