A Computational Study of Cultural Effects on Facial Expressiveness

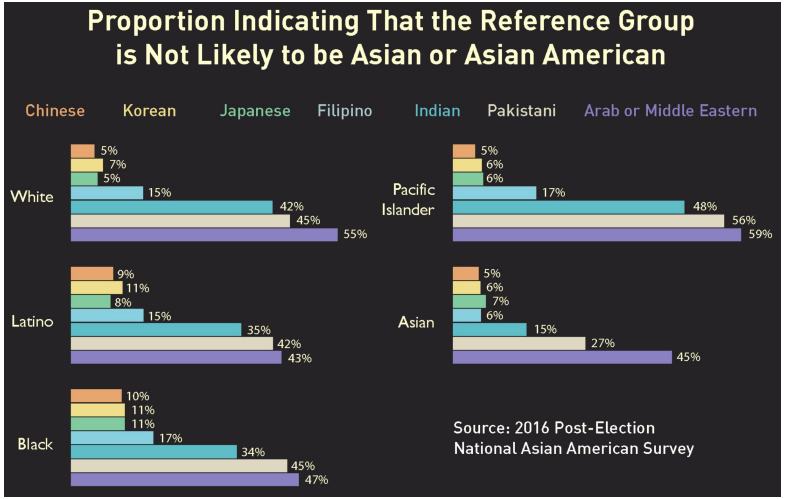
By Rafah Ali, Eshan Tarneja, and Ikhlas Attarwala

Intro + Motivation

- Our favorite discussion questions
  - “How does this model apply to other cultures?”
  - “Is this experiment biased to favor culture x?”
  - “Can this tool work for a different culture?”
- Led us to ask some of our own questions
  - **Do cultures express and perceive emotion the same?**
  - **Can our computational tools account for these potential differences?**
  - **Can we get some love for South Asians!**
- Our process
  - Data collection
  - Survey
  - Computational Analysis
  - Data/Result Analysis

Related Work

- A lot of older research has shown a severe lack in considering ethnic and cultural background.
- Even Paul Ekman created a caucasian-only dataset of facial expressions in 1993 [1]
- In 2009, a paper by Rachael Jack came out challenging the universality of facial expressions [2]
- With available computational tools, challenges such as variance of skin tone and facial structure needed to be addressed [3]
- There is also a major varied perception of the term “Asian”, and whether South Asians fit into this category [4]:



Methods – Photos

- Population (n = 21):
  - Caucasians raised in the US
  - South Asians raised in South Asia

	South Asian	Caucasian
Female	4	6
Male	6	5

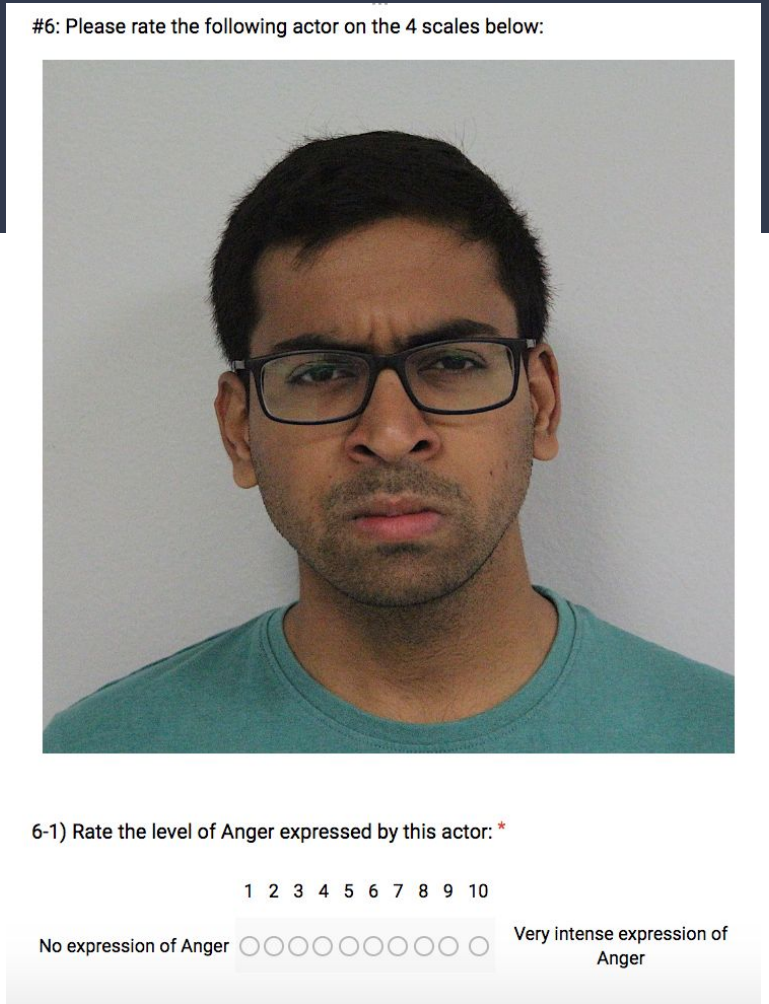
- Age range: 20 to 35 years
- Average age: 23.5 years
- 2 expressions: **Neutral** and **Angry**
- Image properties:
  - Focal length
  - Lighting
  - White balance

Methods - Surveys

- Population (n = 47):
  - Caucasians raised in the US
  - South Asians raised in South Asia

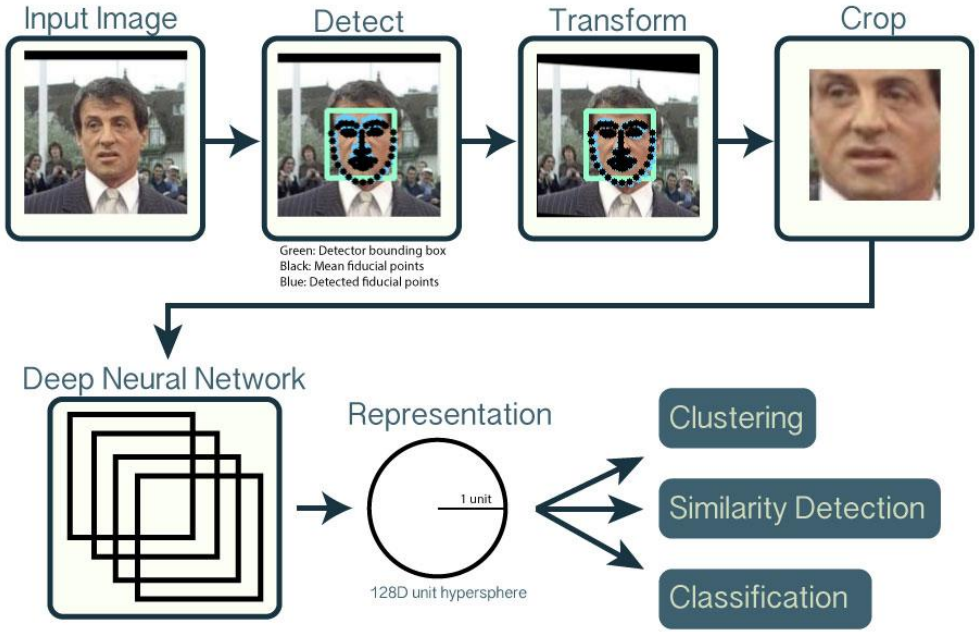
	South Asian	Caucasian
Female	11	11
Male	14	11

- Age range: 20 to 38 years
- Average age: 23.9 years
- Within-subjects design - three groups
- Anger ratings



Methods - OpenFace

- Most interested in FACS
- CASIA-WebFace
  - 10,575 individuals for a total of 494,414 images
- FaceScrub
  - 106,863 face images* of male and female 530 celebrities, with about 200 images per person
- Presence of an AU and its intensity on a 0 to 5 point scale
- Better on video



Findings from our Surveys

- Recap: 2 races rating 2 photos/person from 2 races
- A few ‘neutral faces’ were rated as having higher ‘anger’ scores than the ‘angry face’ itself

Who is angrier?	Both races’ neutral photos	Both races’ angry photos
SA raters	South Asians	South Asians
Caucasian raters	Caucasians	Caucasians

- What did we found within groups?
- What did we find between groups?

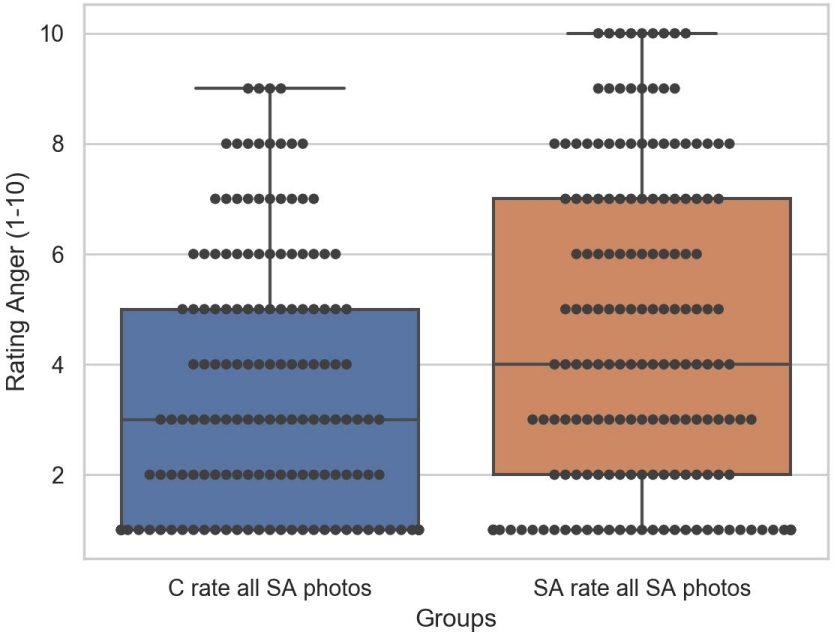
Findings from our Surveys (cont.)

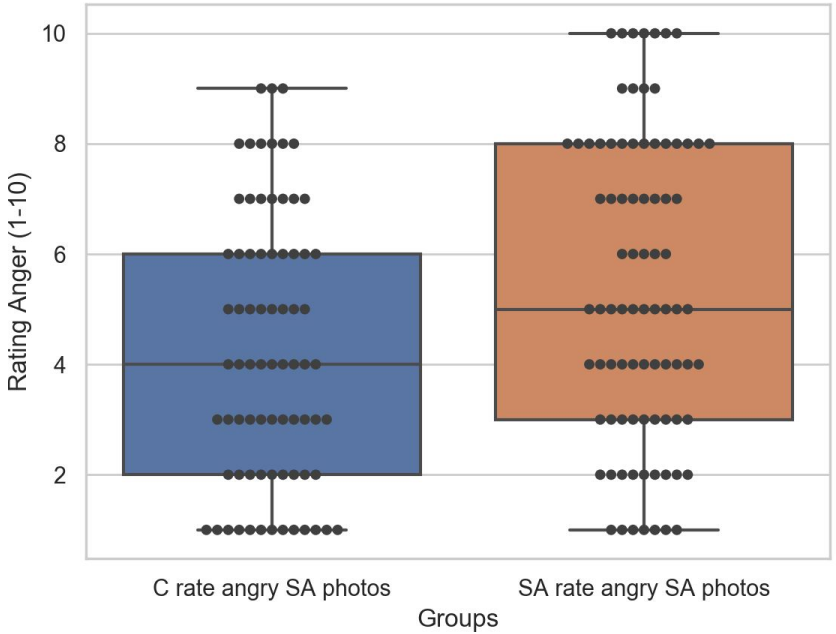
- **What did we found within groups?**
- ie. does each race have a smaller spread of scores for their own race vs. the spread of scores for the other race scoring them?

- **What did we find between groups?**

Rating Caucasians:	P-value = 0.764
Rating ‘neutral’ Caucasians:	P-value = 0.658
Rating ‘angry’ Caucasians:	P-value = 0.993
Rating South Asians:	**P-value = 0.005*** **
Rating ‘neutral’ South Asians:	P-value = 0.146
Rating ‘angry’ South Asians:	**P-value = 0.007*** **

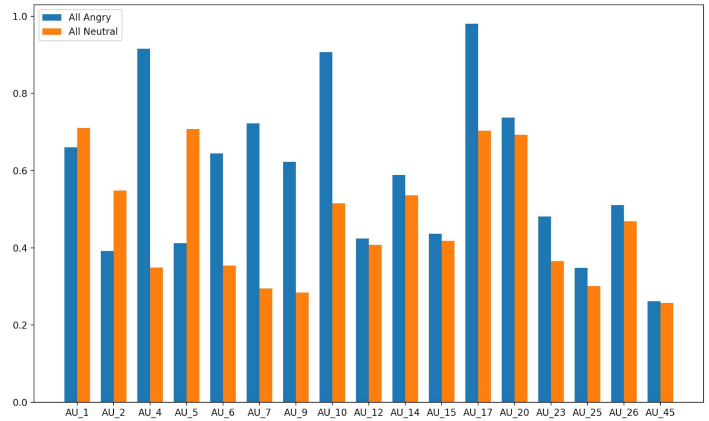
Findings from our Surveys (cont.)





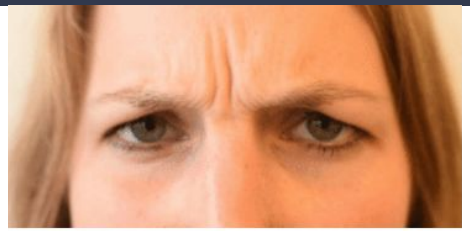
According to OpenFace, do people of different cultures make expressions in different ways?

- 5 categories:
  - Angry vs. neutral faces overall
  - Angry vs. neutral faces in Caucasians,
  - Angry vs. neutral faces in South Asians
  - South Asian angry faces vs. Caucasian angry faces
  - South Asian neutral faces vs. Caucasian neutral faces
- T-test; p<.01
- AUs with interesting results
  - AUs 4, 9, 10 shown in anger across cultures
  - AU 7 high **only** in overall anger
  - AU 23 shown **only** on C anger
  - AU 5 higher **only** in overall when neutral

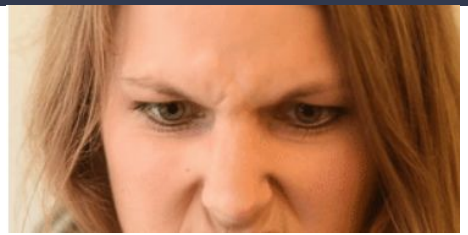


According to OpenFace, do people of different cultures make expressions in different ways?

- AUs 4, 9, 10 shown in anger across cultures
- C show AU 4 most pronounced
- SA show AUs 9,10 during anger **and** neutral
- AU 7 high **only** in overall anger
- AU 23 shown **only** on C anger
- AU 5 higher **only** in overall when neutral
- AU 6 (confounding)



AU 4 - Brow Lowerer

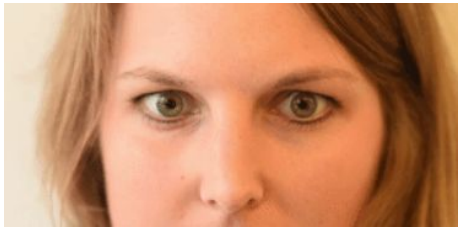


AU 9 - Nose Wrinkler

AU 10 - Upper Lip Raiser

AU 7 - Lid Tightener

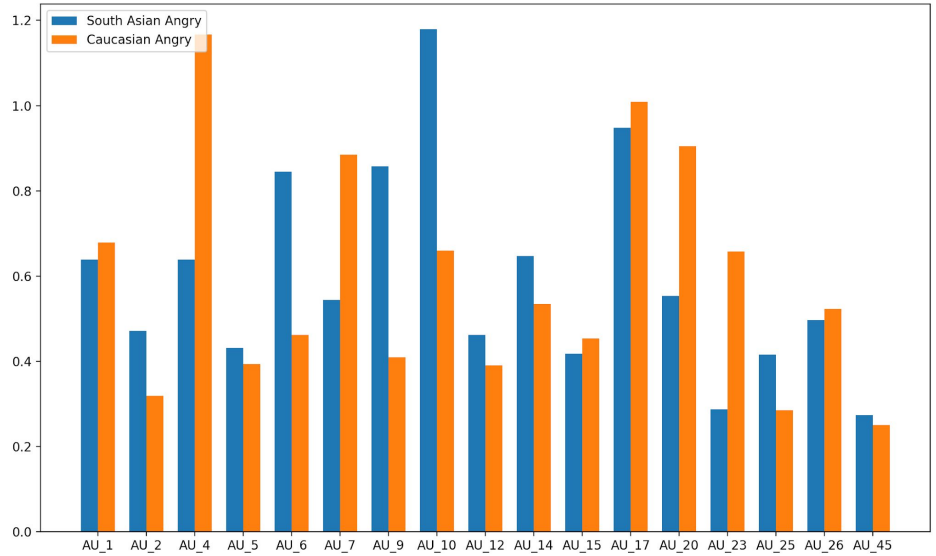
AU 23 - Lip Tightener



AU 5 - Upper Lid Raiser

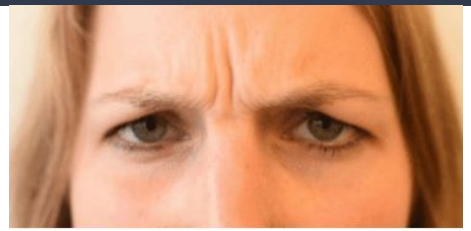
Can OpenFace results indicate differences across cultures' expression and perception of emotion?

- AU 4
  - C on all; SA on SA
- AU 5 = less anger
  - SA on C
- AU 12 = less anger
  - All on all
- AU 14
  - C on C
- AU 17
  - C on C, SA on C, and SA on SA
- AU 26 = less anger
  - All on all
  - Clenched jaw?

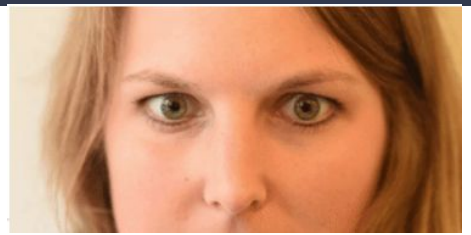


Can OpenFace results indicate differences across cultures' expression and perception of emotion?

- AU 4
  - All categories use it to show anger
  - C on all; SA on SA
- AU 5 = less anger
  - SA on C
- AU 12 = less anger
  - All on all
- AU 14
  - C on C
- AU 17
  - Both categories had in common
  - C on C, SA on C, and SA on SA
- AU 26 = less anger
  - All on all
  - Clenched jaw?

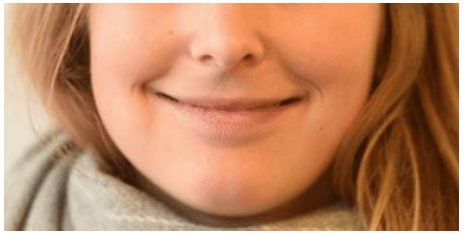


AU 4 - Brow Lowerer



AU 5 - Upper Lid Raiser

AU 14 - Dimpler



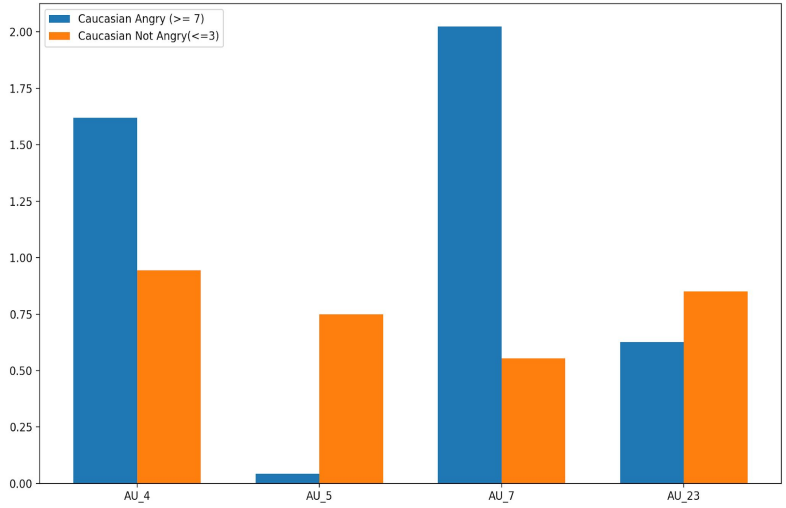
AU 12 - Lip Corner Puller

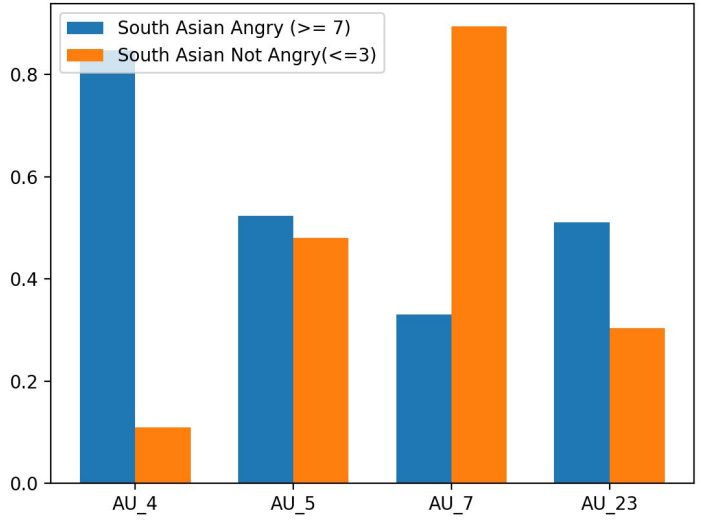
AU 17 - Chin Raiser

AU 26 - Jaw Drop

Does OpenFace read emotional intensity the same way a human does?

Assumption: “South Asians are best at rating South Asians” (Vice-Versa for Caucasasian)
Analysis: Use EMFACS 4 “Angry” AU’s to judge emotion





Discussion – Who's Right?

- No Ground Truth in these types of Analysis
- Assumptions required to test hypothesis
  - "Open face is correctly able to detect action units cross-culturally"
  - "Intra-culture ratings are more accurate than inter-culture ratings
- Circular reasoning
  - Led to some long nights of discussion!
- The solution - tradeoffs
  - Discussions of bias often require these
  - Ex. trading false positive for false negatives

Limitations + Future Work

- Limitations
  - Pool of photographed participants
    - Similar Age
    - Similar Communities
    - Exposed to US culture as students in the US
  - Posing Emotion
    - Confounding Laugh (AU_6)
    - Not trained actors
  - Still photo vs Video Analysis
    - OpenFace performs worse on singular images
    - AU's easier to detect as a change
    - Same for human reading
- Future Work
  - Age, Gender, Emotion Analysis
  - Deeper, more diverse participant pool
  - "Change in emotion" analysis (maybe video)

Thank you! Questions?