Implementing Face Recognition on a Social Scrabble-playing Robot

Vicky Zeng, Reid Simmons

Motivation

- Personalization improves engagement and helps establish intimate, long-term relationships between Victor and players
- Face recognition is needed for Victor to identify the players
- Little research is done on local optimizations of open-source face recognition

Problem

Chosen face recognition model:
Dlib over OpenFace (Dlib had significantly higher Asian accuracy)

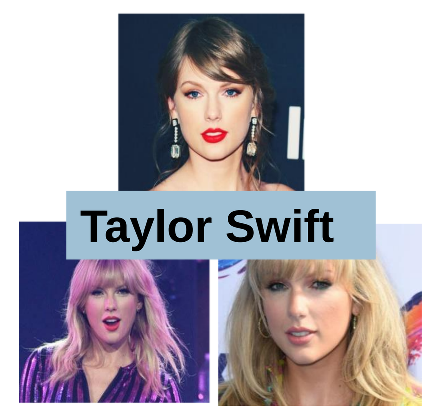
Factors to optimize:
⇧ **Accuracy** (correctly labeling known faces)
⇧ **Unknown** Detection (correctly labeling unknown faces as unknown)
⇩ **False Positives** (labeling a known face as another face)

Baseline numbers:
82% accuracy, **15%** false positive, **43%** unknown detection

Methods

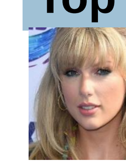
1) Applying a classifier

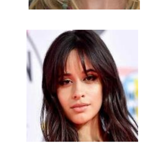
	Original Model	With Classifier
Input	**One** image per label	**Multiple** images per label
Output	**Top** match	**Weighted vote of top 5** matches



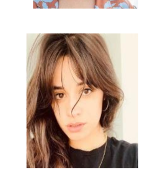
Unlabeled Image

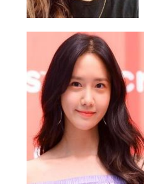
Top 5 matches:


Label: Taylor Swift
Distance: 0.428


Label: Camila Cabello
Distance: 0.431


Label: Camila Cabello
Distance: 0.439


Label: Camila Cabello
Distance: 0.445


Label: Im Yoona
Distance: 0.523

Improvement:
95% accuracy, **5%** false positive, **64%** unknown detection

2) Finding patterns that indicate label uncertainty

Patterns based on distance values and labels

A) - Separate true and false positives, **test simultaneously**
- **Derive related patterns** when noticeable difference occurs

B) - Conduct data collection **only on false positives**
- When a promising pattern shows, **test it against true positives**

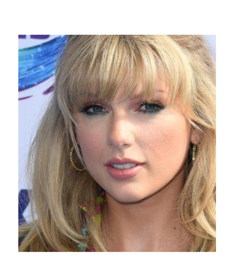
Patterns: Unique labels, Thresholds, Differences between matches
Types: Universal, In-between and Local (with respect to dataset size)

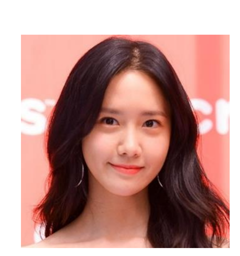
Universal: Top three matches are unique labels and the difference between the first two matches is less than 0.08

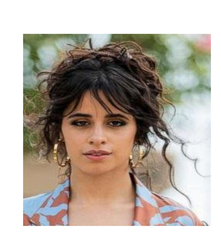
Local (20): There are more than four unique labels
Local (200): … the first match's distance is greater than 0.45
Local (400): … difference between the two labels' averaged distances is less than 0.03.

In-between: … difference between the two distances is: 0.46 (20), 0.45 (200), 0.43 (600)

Example conditions →


Taylor Swift
0.428


Im Yoona
0.462


Camila Cabello
0.481


Camila Cabello
0.535


Barack Obama
0.613

Results

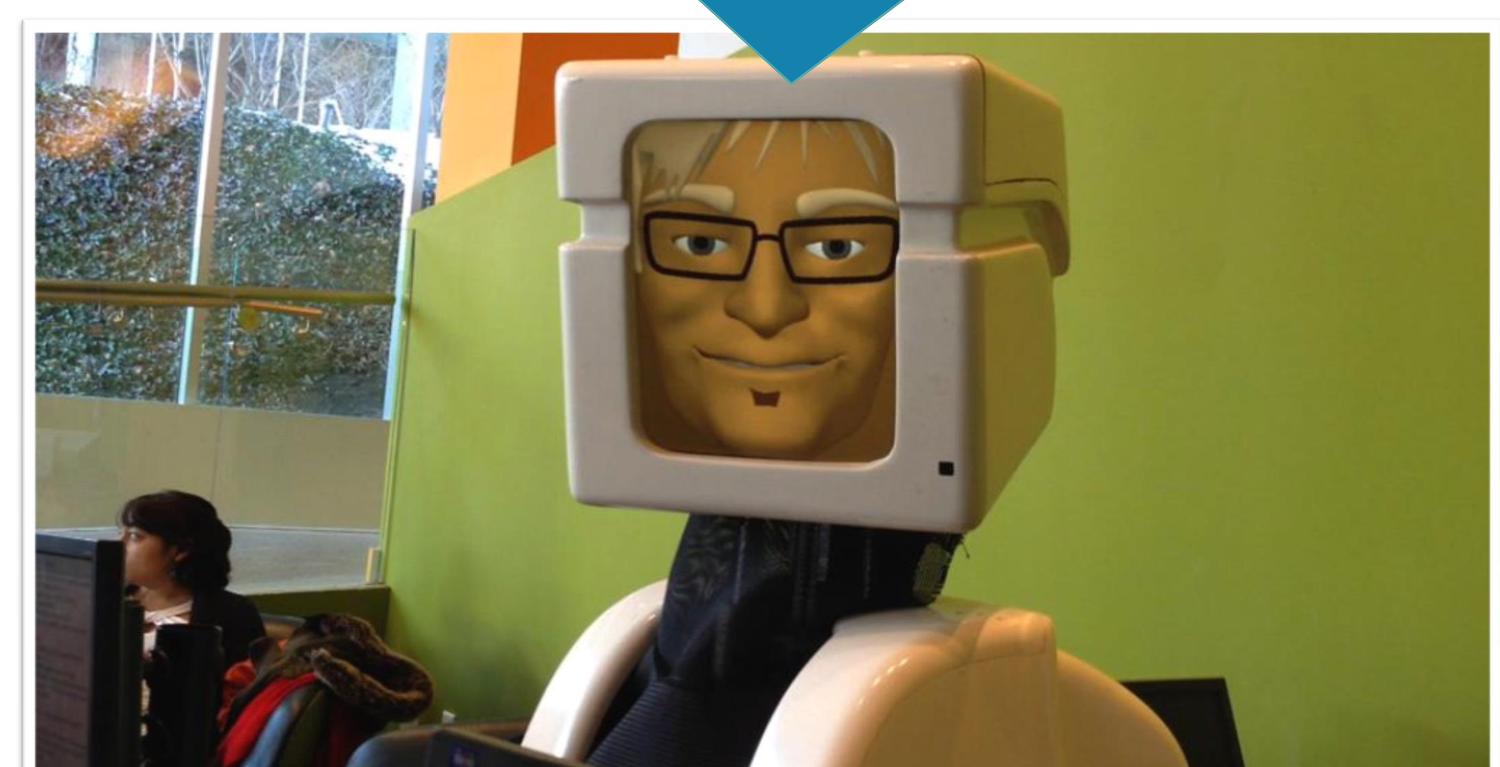
Table 1. Results before and after optimization. The numbers are **averaged from five runs** of each dataset, and the numbers in parentheses represent their **standard deviations.**

Dataset size	Accuracy (%)	False Positive (%)	Unknown (%)	Unknown Detection (%)
200 (BO*)	**83.5 (3.43)**	**11.4 (2.25)**	**4.6 (1.21)**	**44.4 (2.52)**
200 (WC)**	**94.8 (0.75)**	**5.2 (0.62)**	**0.2 (0.39)**	**64.4 (1.62)**
20	95.8 (1.90)	1.2 (1.15)	2.4 (1.46)	93.7 (1.8)
200	**89.1 (0.63)**	**0.8 (0.26)**	**10.1 (0.39)**	**89.4 (1.5)**
400	88.3 (0.85)	1.0 (0.19)	10.6 (0.81)	88.2 (1.1)
600	87.9 (0.71)	1.2 (0.15)	10.9 (0.8)	86.7 (1.2)

*BO: Before optimization **WC: With classifier optimization only

Our work produced higher accuracy, lower false positives, and higher unknown detection.

Today was fun! Guess I'll see you next Wednesday at our usual time?



Future Direction

- Integrate face recognition into Victor's current system
  - Consider physical responses such as greeting and looking in the direction of the person
- Keep interaction logs between Victor and the players. Use that information to personalize subsequent interactions
  - **Game-focused:** Skill level, average response time, types of turns
  - **Time-focused**: Game duration, visit frequency, days and times of visits
  - **Interaction**: Levels of snarkiness, simple Q&A on personal events
  - **External**: Reaching out to players through online platforms

Acknowledgements

The author would like to thank Dr. Reid Simmons for his continuous guidance and support. Also, thank you Rachel Burcin, Dr. John Dolan and Mikayla for organizing RISS and making this all possible.