Real-time emotion recognition in virtual reality

Jeremy Di Dio

Sruti Bhattacharjee

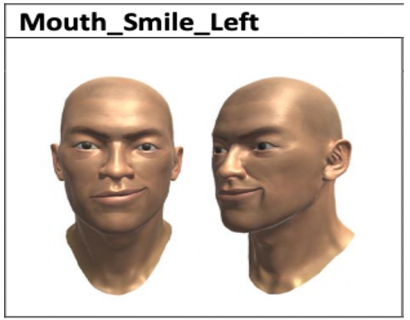
Semester Project Fall 2021

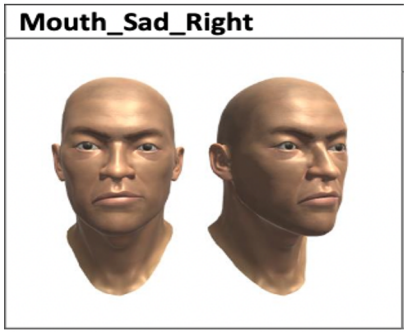
Outline

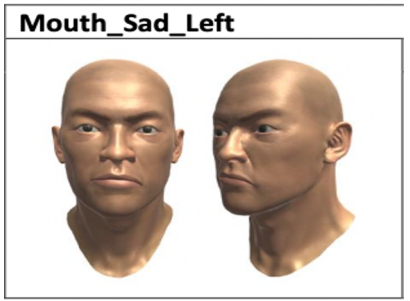
- Project Description
- Project Road Map
- Database Creation
- Models Implementation
- Evaluation
- Future work
- Conclusion

Project Description

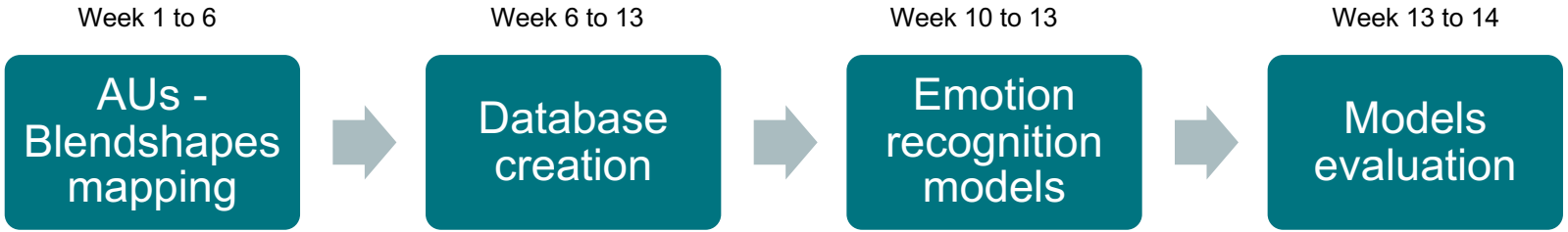
- New HTC Vive facial tracker
- Blendshapes data
- Identification of 6+1 basic emotions
  - Happiness, Fear, Anger, Surprise, Disgust, Sadness
  - Neutral
- Real-time emotion recognition







Project Road Map



Database Creation

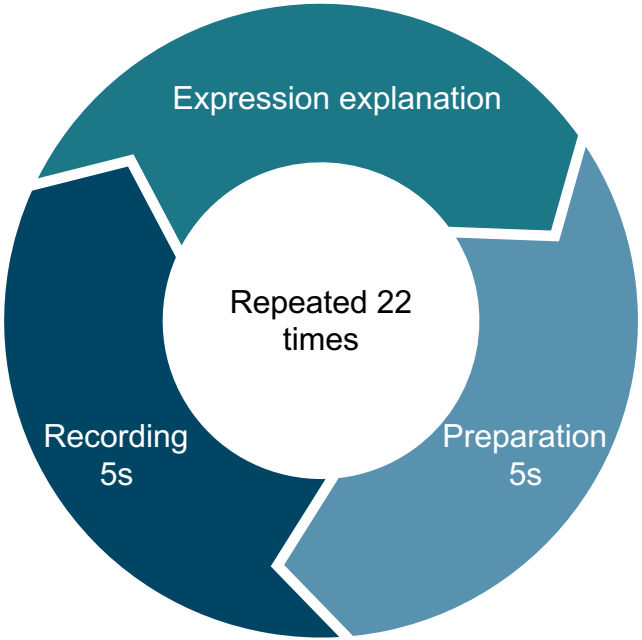
- Extended Cohn-Kanade database (CK+)
- Facial Action Coding System
- Collected expressions
  - 6 Basic emotions
  - 16 relevant AUs

AU 9

AU 12

Farnsworth, B. (2019, August 18). *Facial Action Coding System (FACS) – A Visual Guidebook*. Imotions. Retrieved October 1, 2021, from https://imotions.com/blog/facial-action-coding-system/

Recording Protocol



Recording process

Participant being recorded

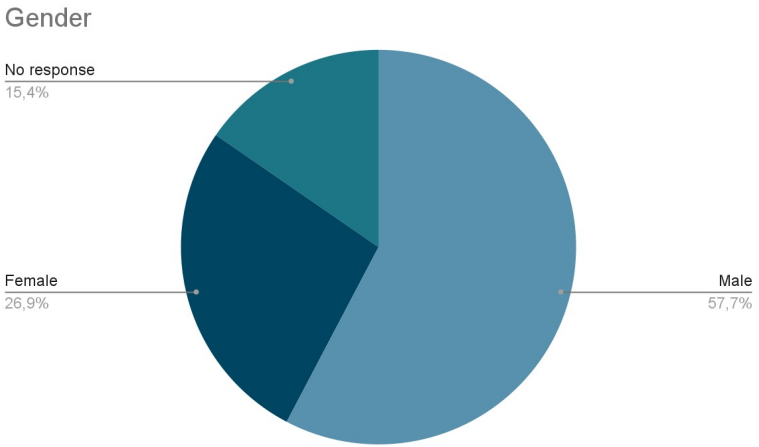
Collected Data

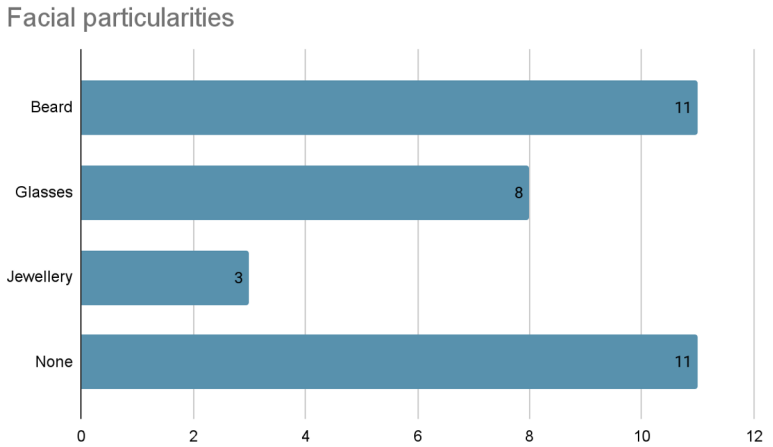
- Blendshapes
- Raw lips video

Expression	Multiple expression	Frame	Apex	Eye_Left_Blink	Eye_Left_Wide	Eye_Left_Right	Eye_Left_Left
AU1	3	0	False	0.0	0.246164	0.000000	0.049763
AU1	3	1	False	0.0	0.232747	0.000000	0.049385
AU1	3	2	False	0.0	0.232747	0.000000	0.049385
AU1	3	3	False	0.0	0.219523	0.000000	0.049394
AU1	3	4	False	0.0	0.219523	0.000000	0.049394
...	...	...	...	...	...	...	...
Surprise	1	247	False	0.0	0.521865	0.037978	0.019441
Surprise	1	248	False	0.0	0.599759	0.037845	0.019582
Surprise	1	249	False	0.0	0.599759	0.037845	0.019582
Surprise	1	250	False	0.0	0.646012	0.037488	0.019277
Surprise	1	251	False	0.0	0.646012	0.037488	0.019277

Participants Details

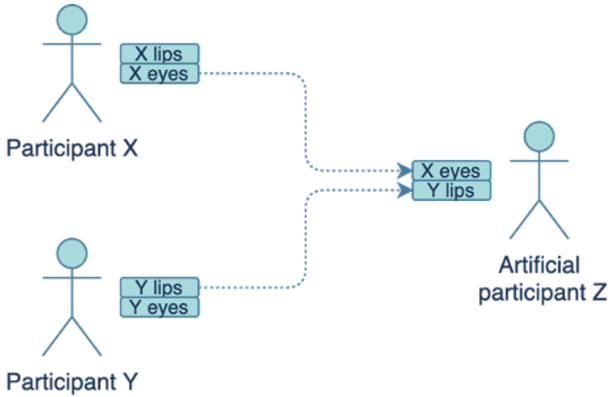
- Total of 27 participants
- Average age of 26
- Estimation of 5’500 samples by participants





Models Implementation

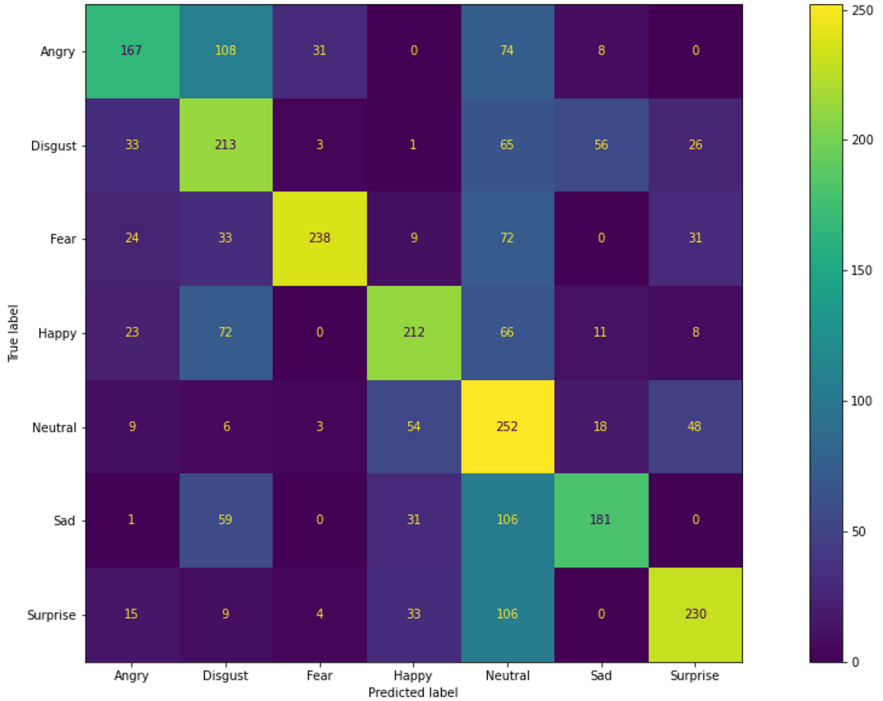
- Dataset preparation
  - Remove inconsistent data (one participant removed)
  - Selection of relevant samples (210 samples per participant)
  - Data augmentation to a total of ~15’000 samples
- Random Forest and LightGBM
  - Most suitable models considering the dataset characteristics
- By participant cross-validation
  - 10 round of randomized 80/20 split of training and testing set



Data augmentation illustration

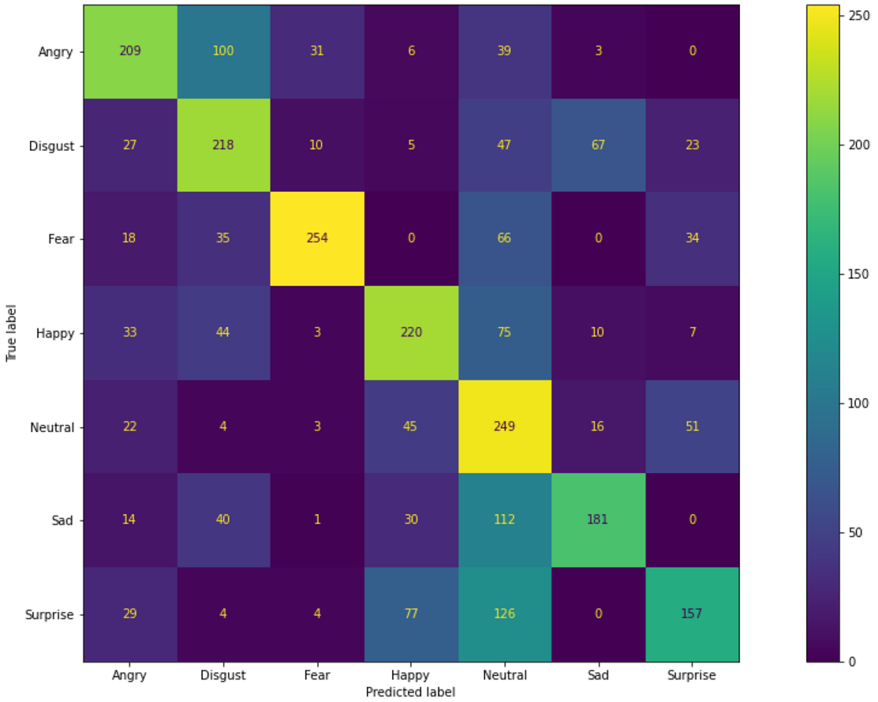
Random Forest

- Scikit-learn library
- Hyperparameters tuned
  - Maximum depth, Number of trees
- Accuracy
  - 48% with D.A
  - 37% without D.A



LightGBM

- LightGBM library
- Hyperparameters tuned
  - Number of trees, Learning rate, Number of leaves
- Accuracy
  - 52% with D.A
  - 38% without D.A



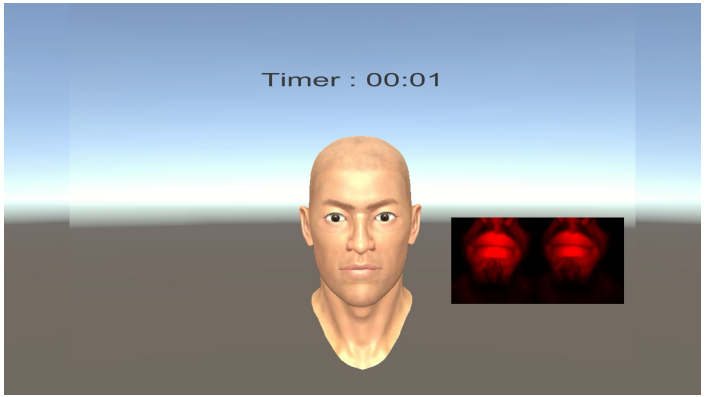
Evaluation

- Offline evaluation
  - posed and unposed emotion recording
  - Unity integration complexity
- Real-time evaluation simulation
  - The goal is to simulate a real-time prediction on pre-recorded data
- Evaluation visualization web application
  - We build a webapp to visualize and simulate real-time evaluation

Evaluation Set

- Same recording app as for training dataset
- Two continuous recording of the 6 basic emotions
  - Predefined order VS undefined order
  - Two 25 seconds recording
- 2 participants in total
  - Both Male, both little beard, one with glasses

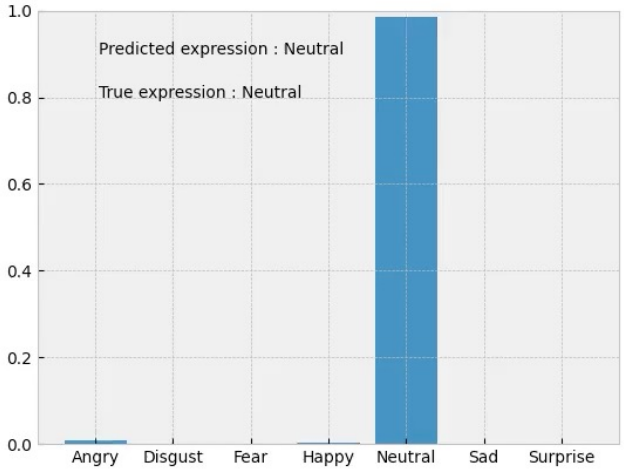
Screen recording of participant's view



Emotions' order: neutral, happy, sad, fear, angry, surprise, disgust

Real-Time Simulation

- Live prediction simulation
  - Based on pre recorded data
- 50Hz rate
- Probability view
  - Each emotion is predicted by a percentage of certainty



Emotion prediction sample

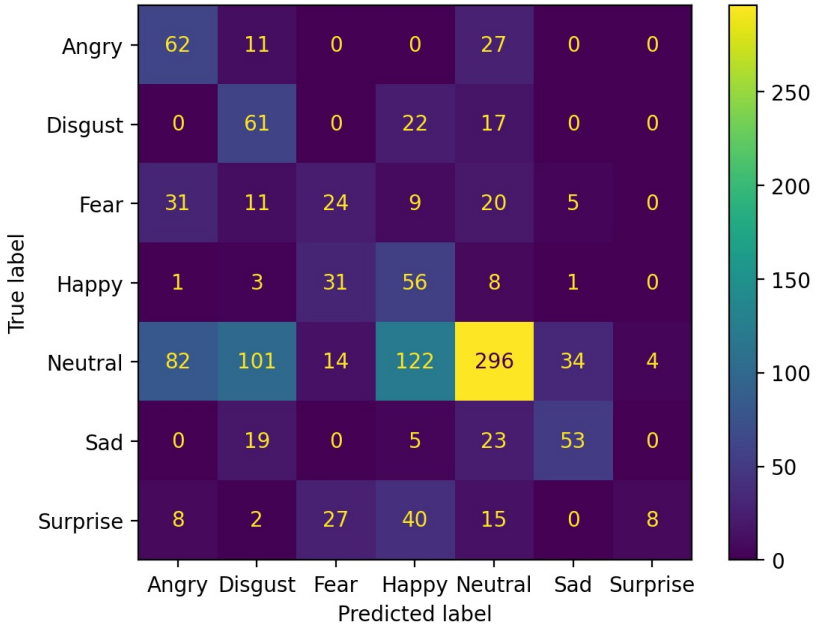
Results

Emotions	Random Forest	LightGBM
Anger	55%	62%
Disgust	62%	61%
Fear	0%	24%
Happiness	61%	56%
Sadness	55%	53%
Surprise	20%	8%
Neutral	40%	45%
Total	**41%**	**45%**

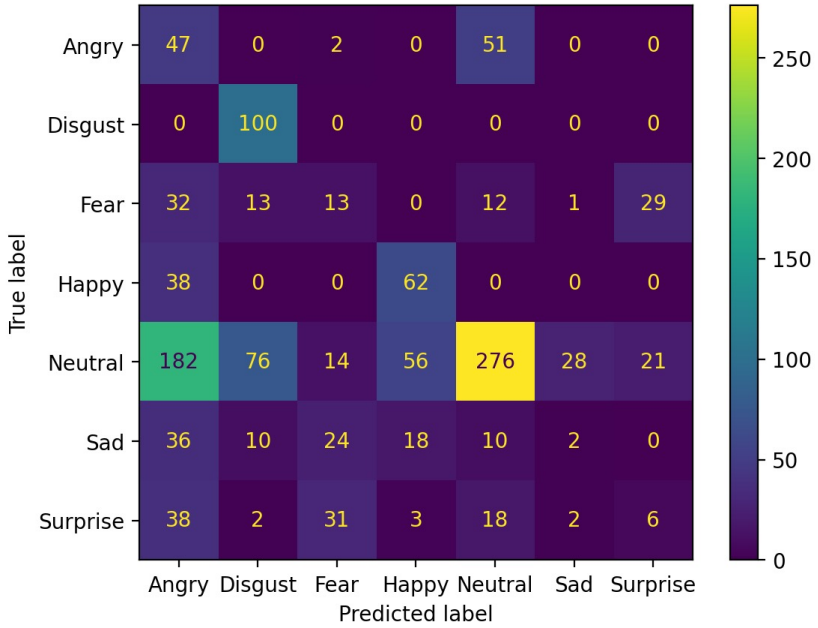
Results of our best recording

- Fear bad performance
  - Negative emotion
  - Frowning and lowering lips corner
  - Close to anger and disgust
- Surprise bad performance
  - Positive emotion
  - Participants smiled
  - Close to happiness

Eyeglasses influence performances



Participant without eyeglasses



Participant with eyeglasses

Future Work

- Increase training database size
  - More diverse participants
  - Also for evaluation
- Online evaluation
  - Build a VR application enabling live emotion recognition
  - Lab streaming layer or OSC
- Action Units recognition
  - Build different emotion recognition system based on AUs
  - Use AUs-Emotions mapping system

Conclusion

- Limitations
  - Number of participants
  - Accuracy of models
- Blendshapes usage for emotion recognition

Frooxius. (2021, March 13). *Setting up avatars for VIVE Facial Tracker in Neos VR & Custom Scripting (tutorial)* [Video]. Youtube. https://www.youtube.com/watch?v=OHyl90_X9SQ

Many thanks!