Visual Prostheses

Eye Detection and Face Recognition

Hesham M. Moneer

Supervisors:
Dr. Seif El Dawlatly
Eng. Reham Elnabawy

German University in Cairo

July 1, 2022

Outline

1. Introduction
2. Background
3. Methodologies
4. Results
5. Conclusion
6. References

Motivation

Normal Sight

Motivation

Acquired Blindness (Retinitis Pigmentosa [peripheral vision loss])

Motivation

Prosthetic Vision (low spatial resolution)

Objective

Improve visual prostheses implantees efficiency in group conversations.

In a group conversation setting, it is very valuable for a group member to do the following efficiently:

Objective

1) Recognize other group members

Objective

2) Distinguish alike faces

Objective

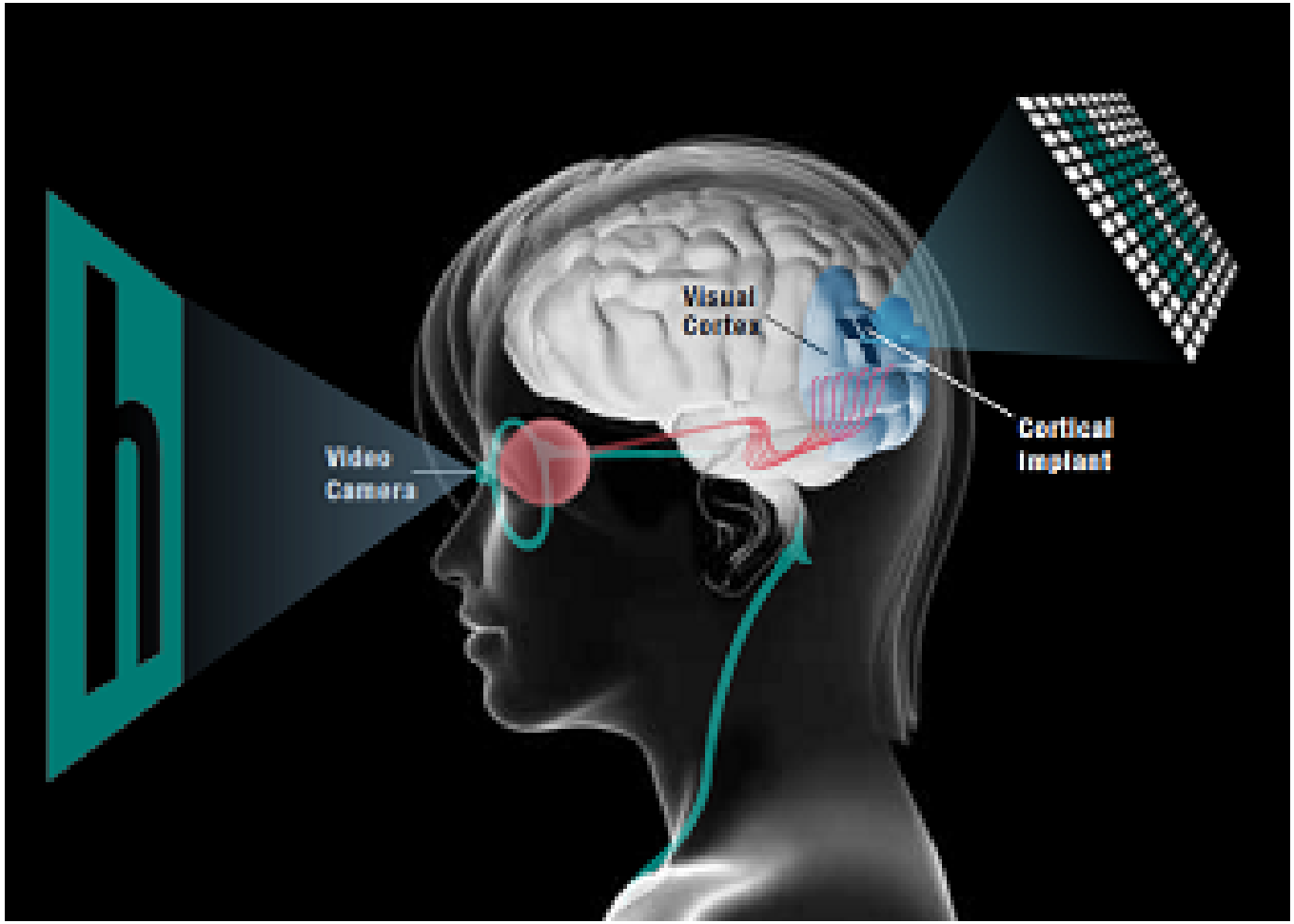
3) Identify the talking person(s)

Objective

4) See facial expressions (convey emotions words cannot portray)

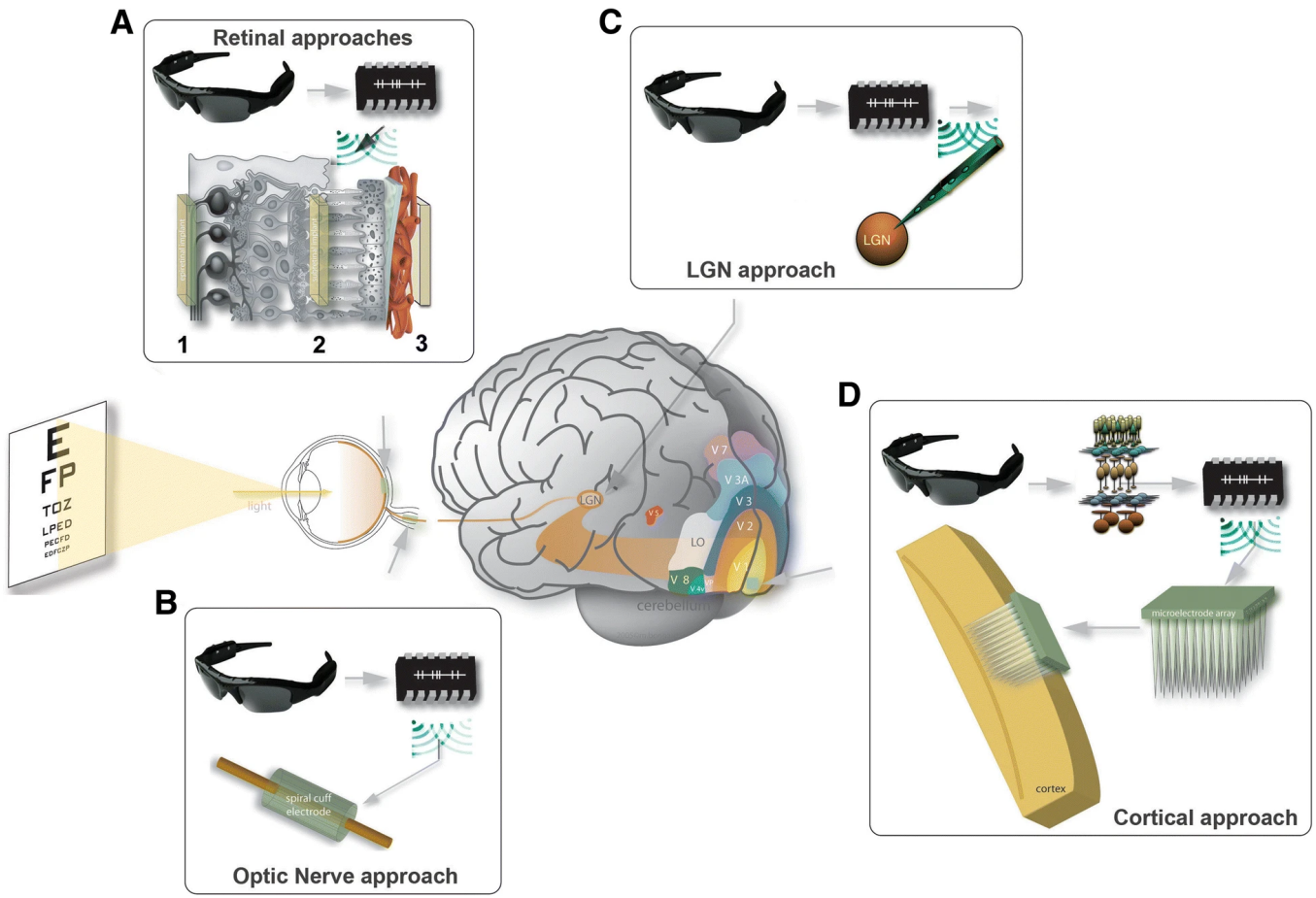
Visual Prostheses

Implantable medical device that cause perception of phosphenes [1]



Visual Prostheses

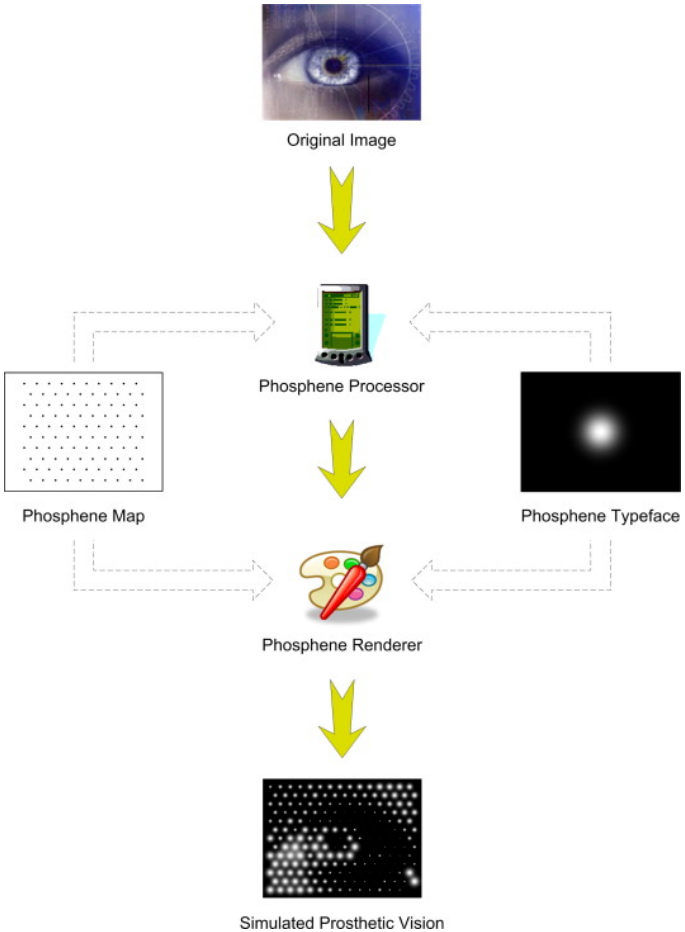
Different approaches along the visual pathway [1]



Simulations of Prosthetic Vision (SPVs)

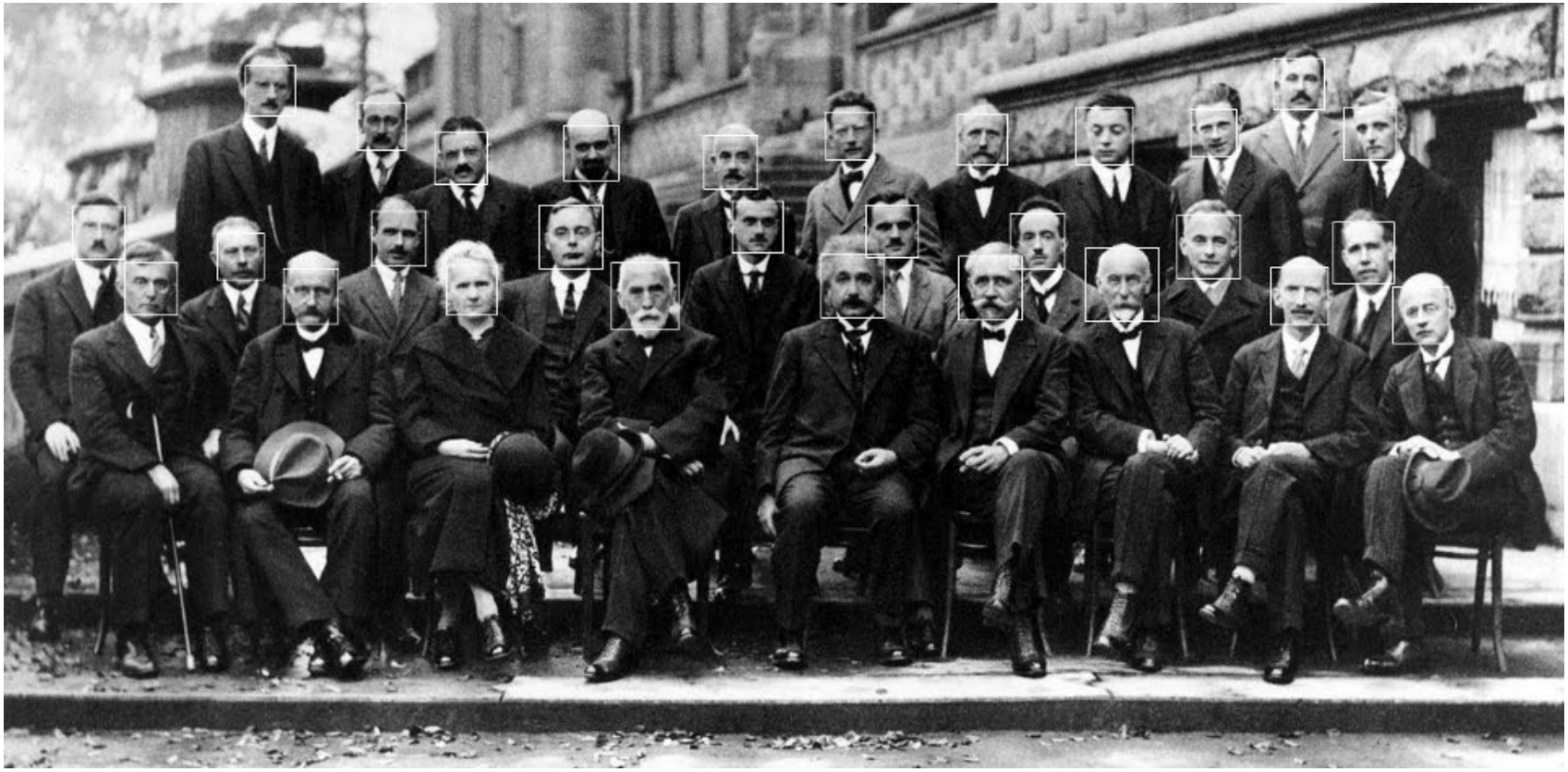
No access to actual implantees

SPVs applied on normally sighted subjects in VR [2]



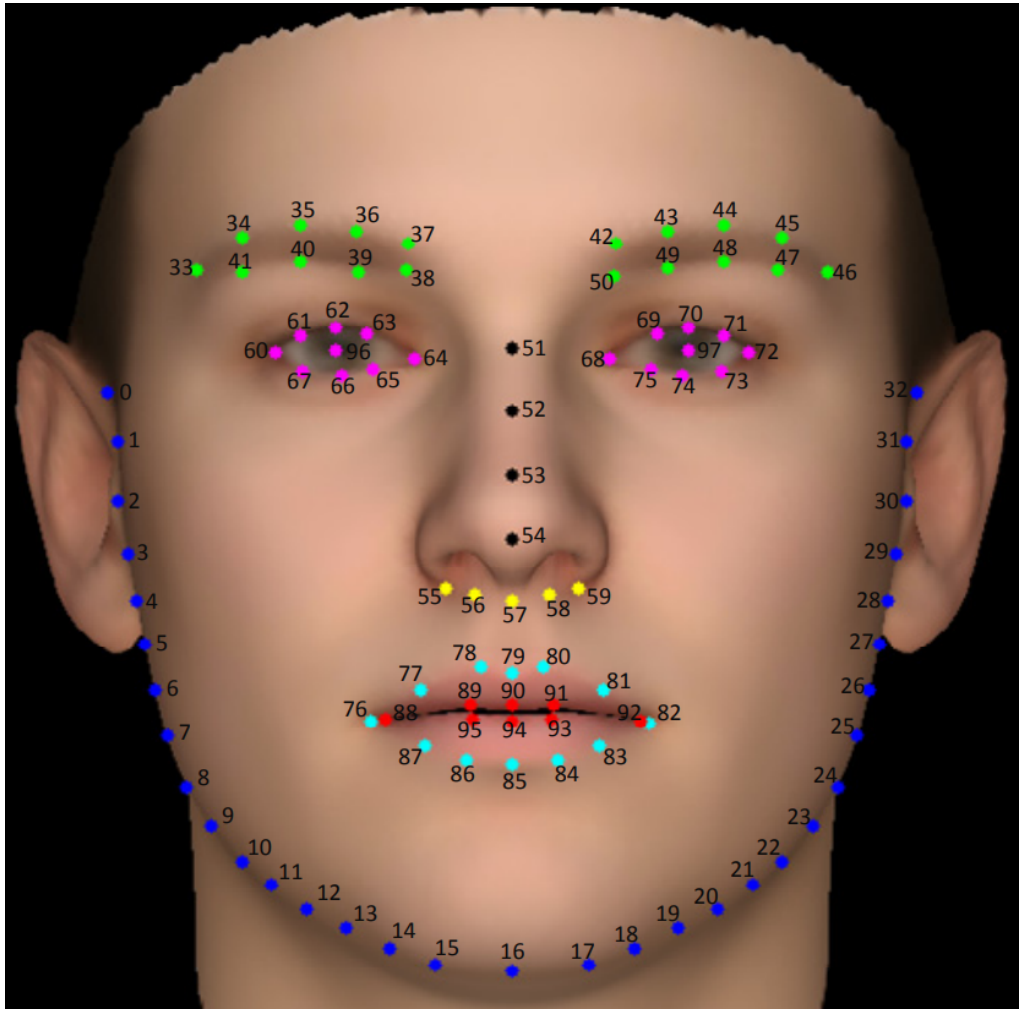
Facial Detection and Landmarks

Viola Jones Haar Cascade Classifier [3]



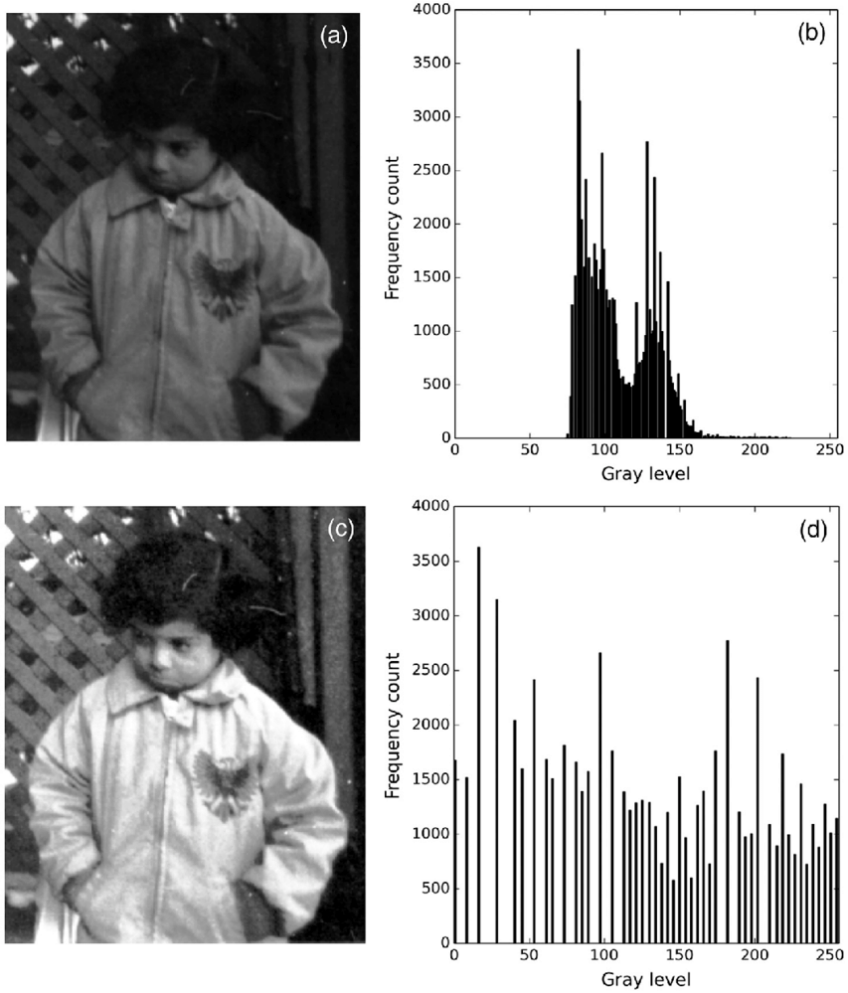
Facial Detection and Landmarks

Face landmarks detection [4]



Enhancement Techniques

Histogram Equalization [5]

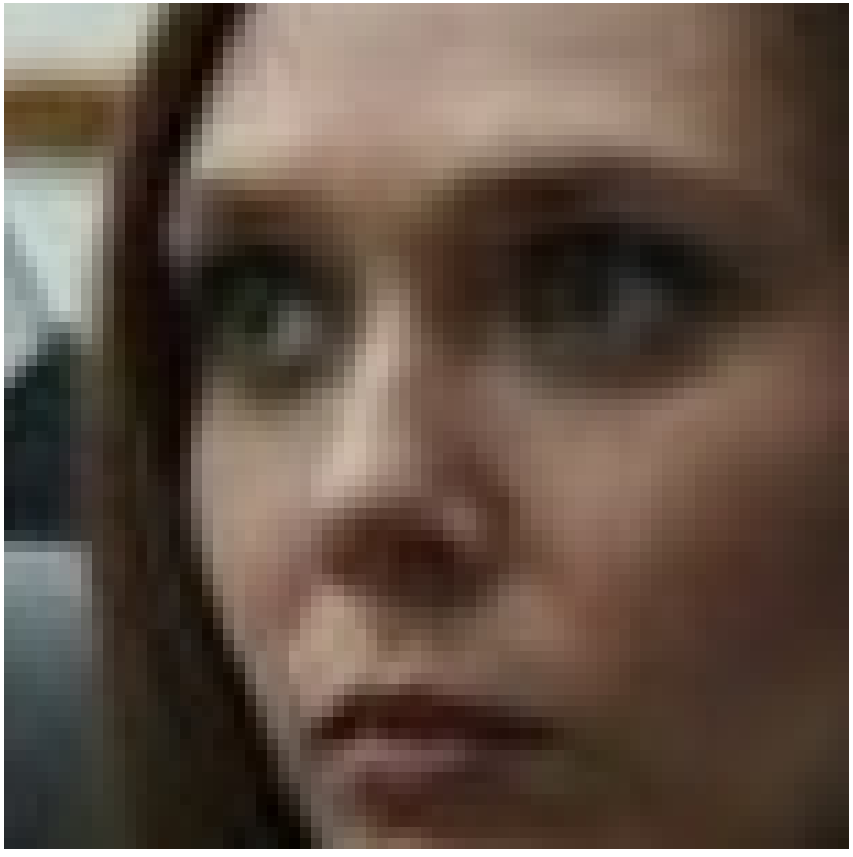


Enhancement Techniques

Region of interest (ROI) magnification (VJFR vs SFR) [6]

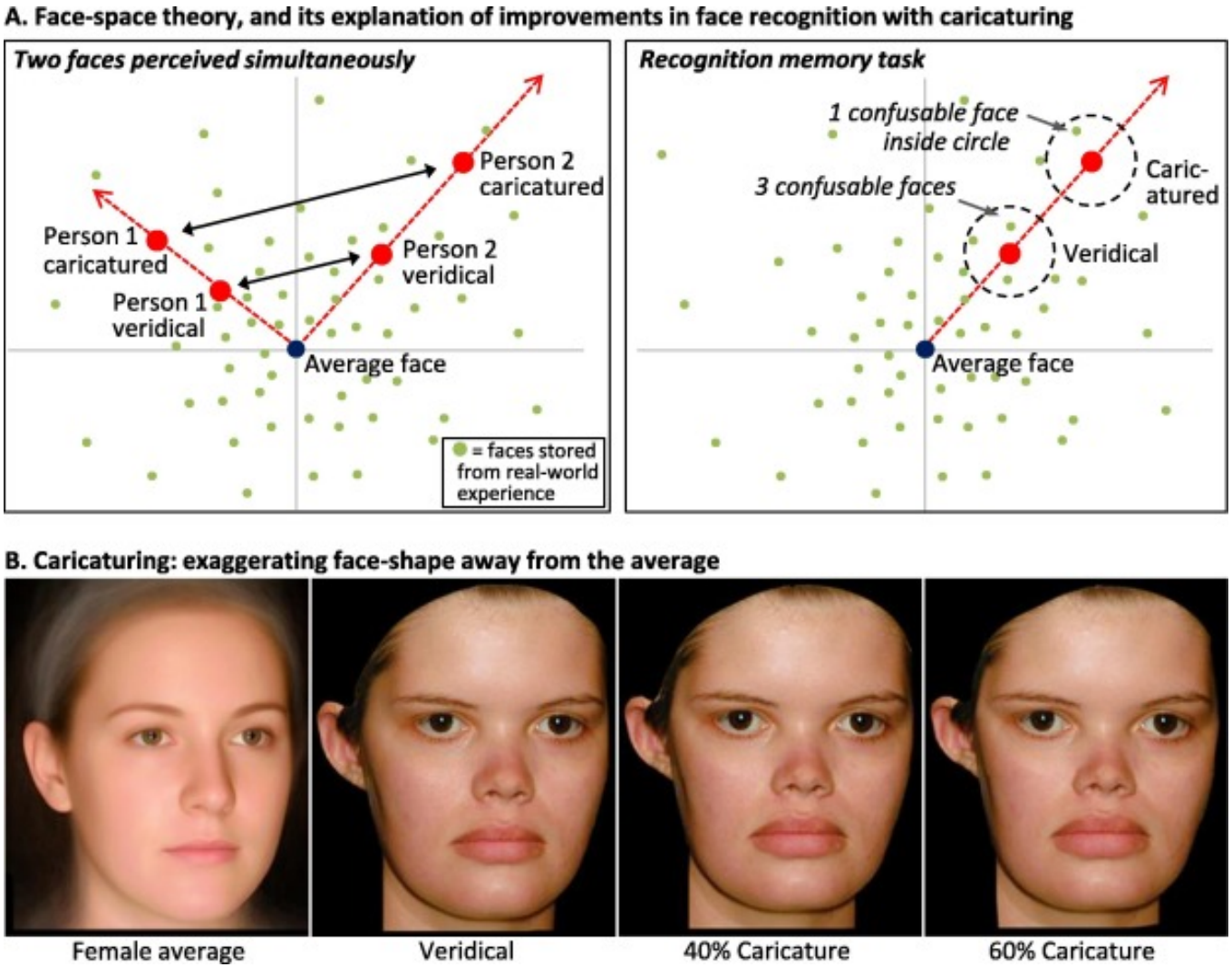
Enhancement Techniques

Region of interest (ROI) magnification (VJFR vs SFR) [6]



Enhancement Techniques

Caricaturing on perceptual face-space [7]

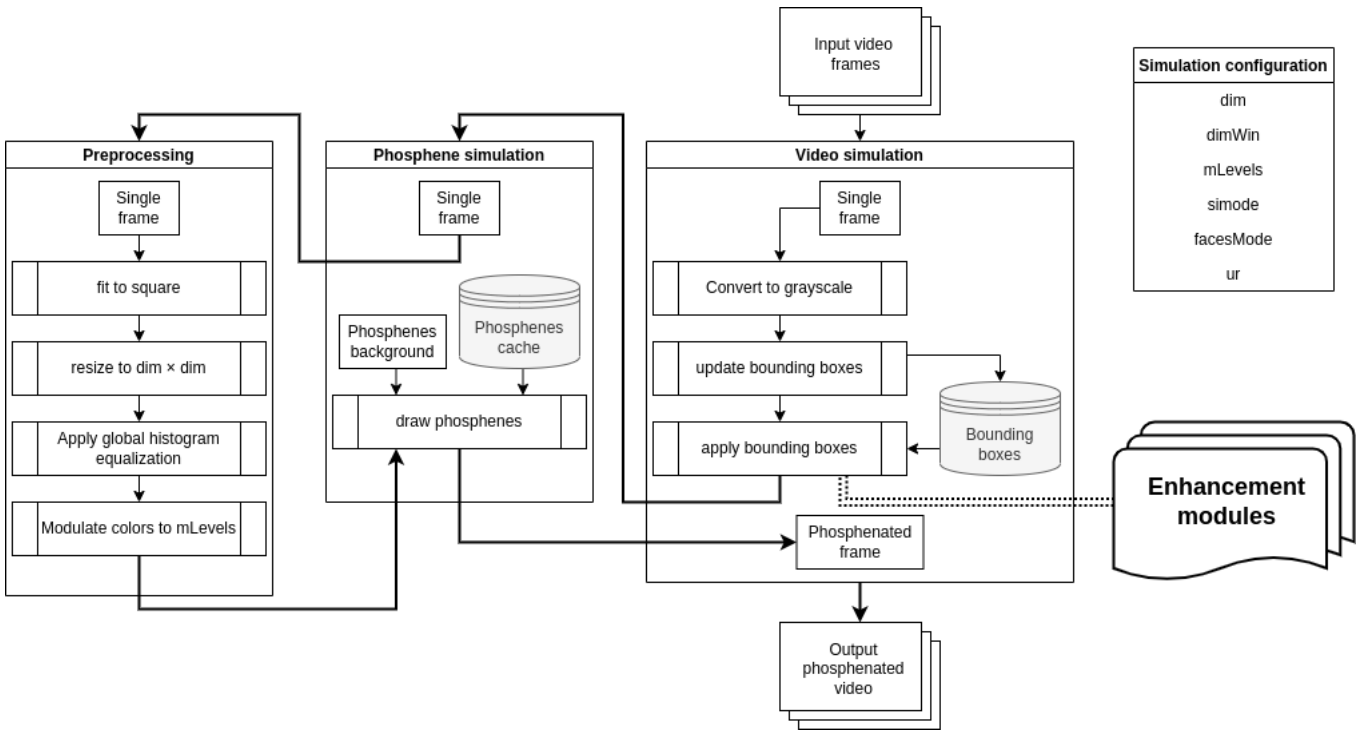


Simulation

1. Python
2. Numpy
3. OpenCV
4. Dlib

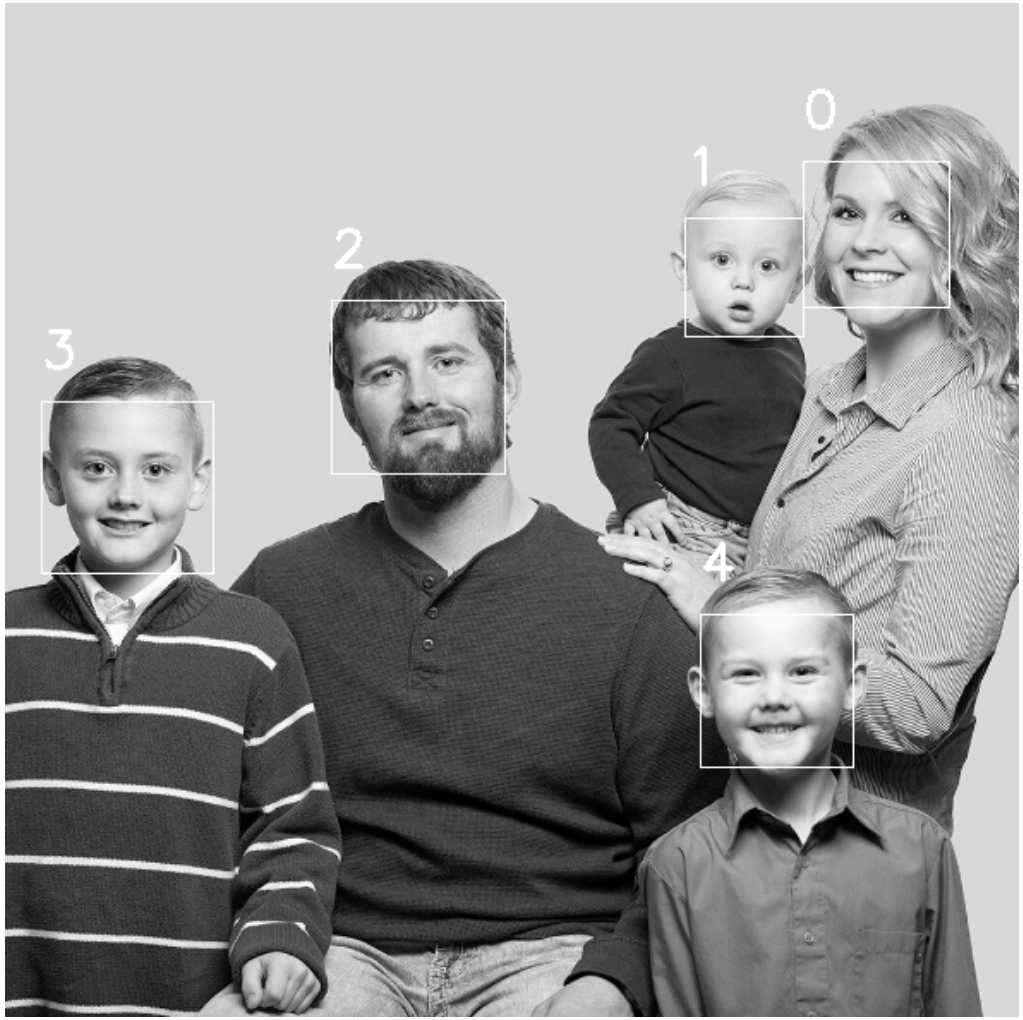
Simulation

https://github.com/HeshamMoneer/Phosphenes-Simulation



Simulation

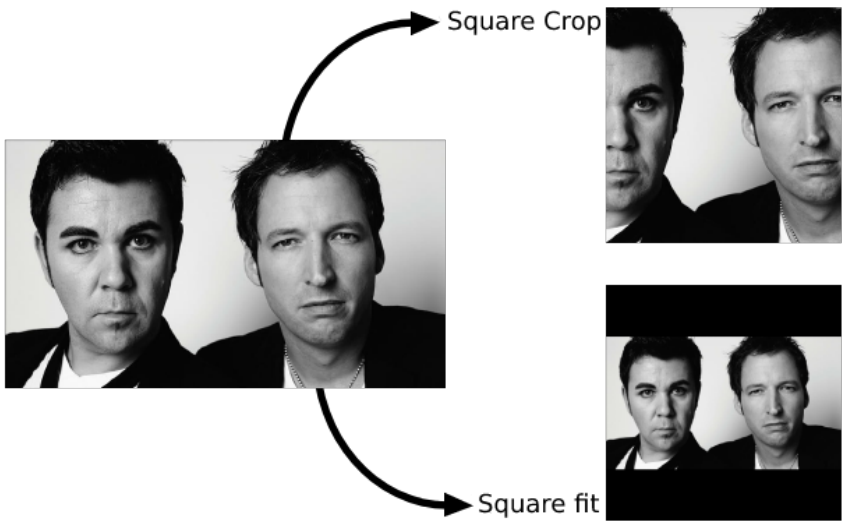
Faces Bounding Boxes



Simulation

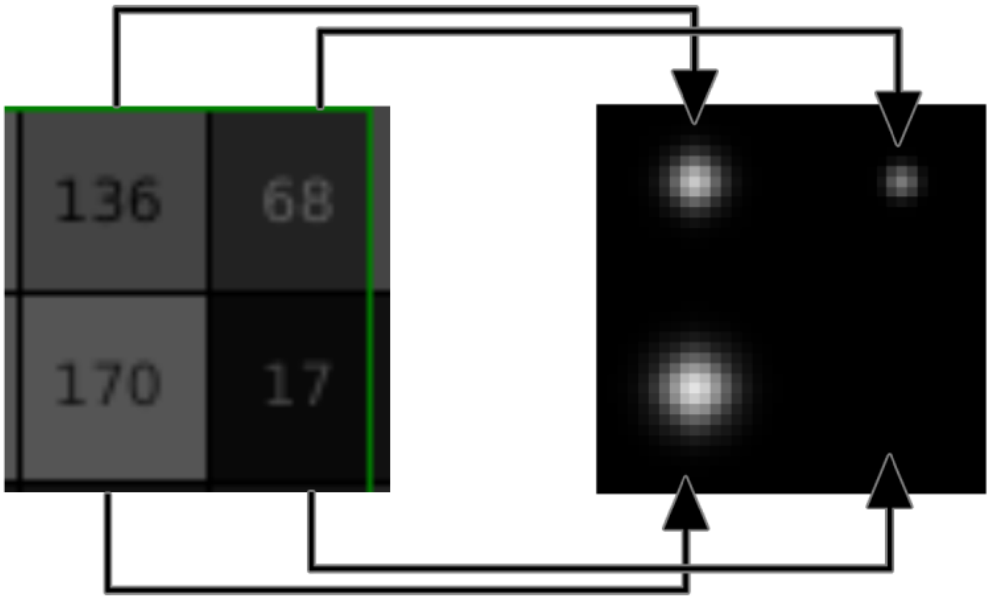
Image preprocessing

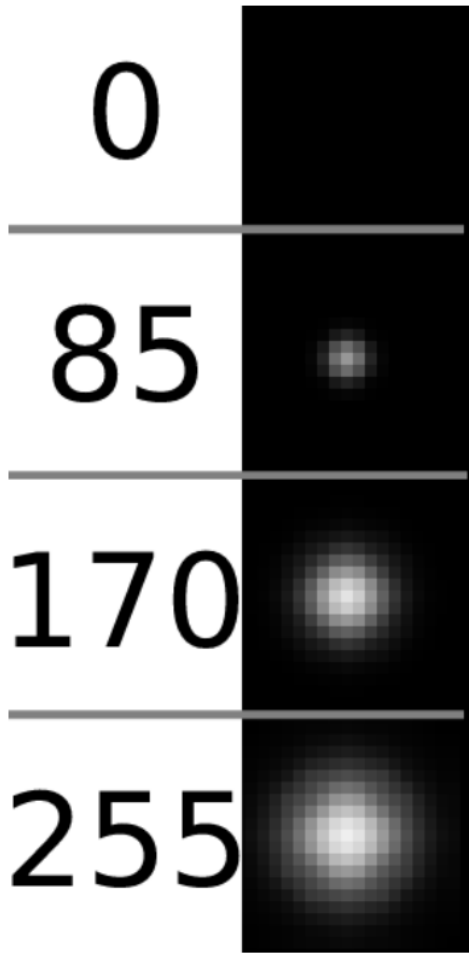
1. Square fit
2. Resize to dim $\times$ dim (bilinear interpolation)
3. Apply Histogram equalization
4. Modulate colors to mLevels



Simulation

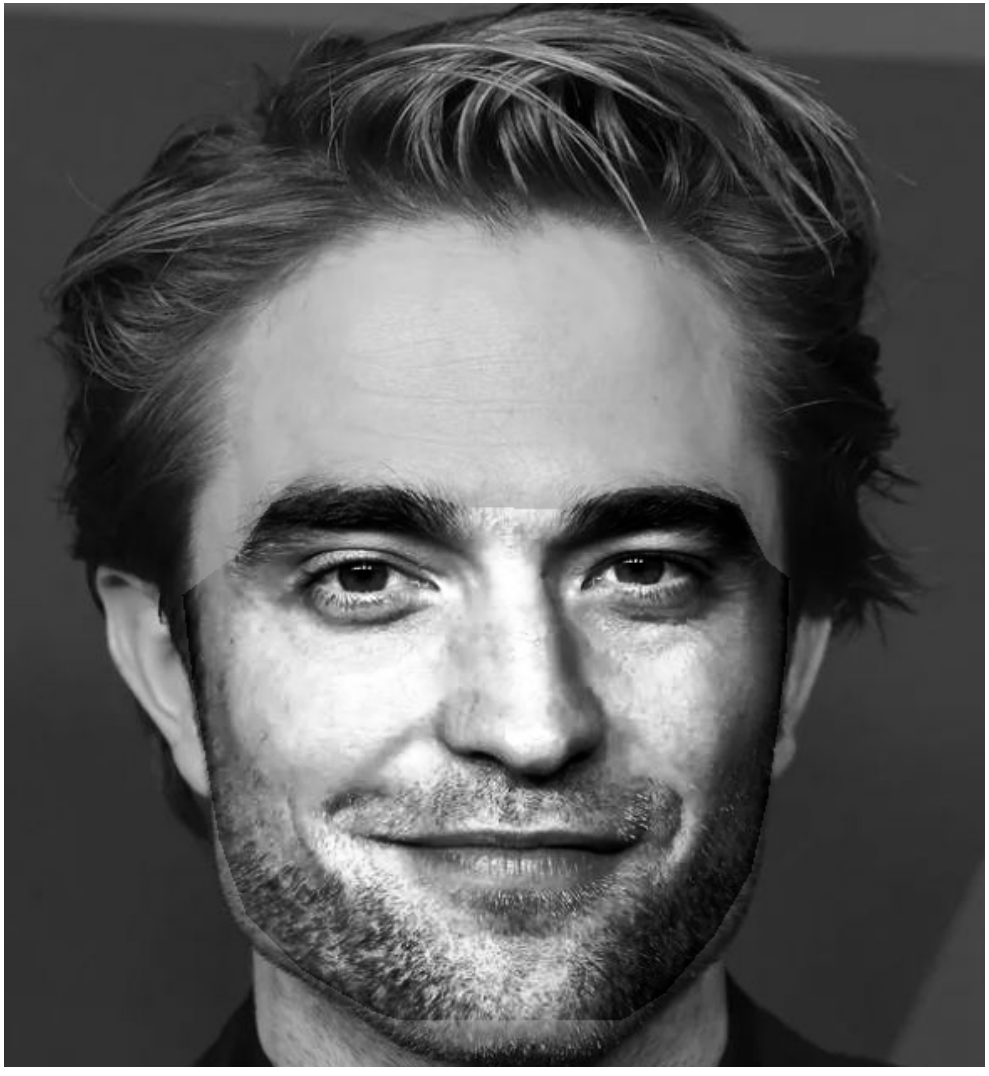
Draw Phosphenes





Enhancement Modules

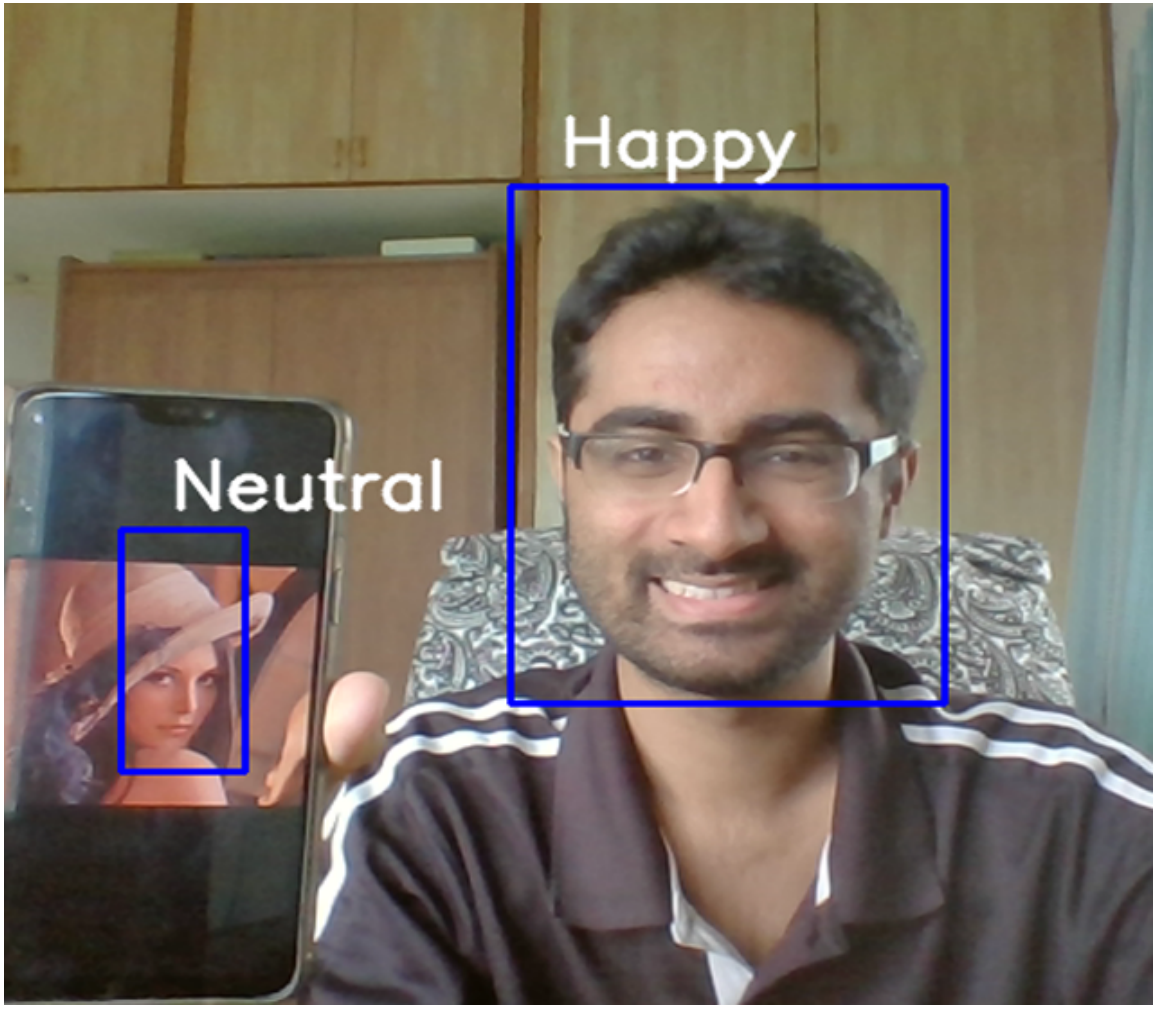
Face-specific histogram equalization (FSHE)



Enhancement Modules

Emotion Recognition (CNN)

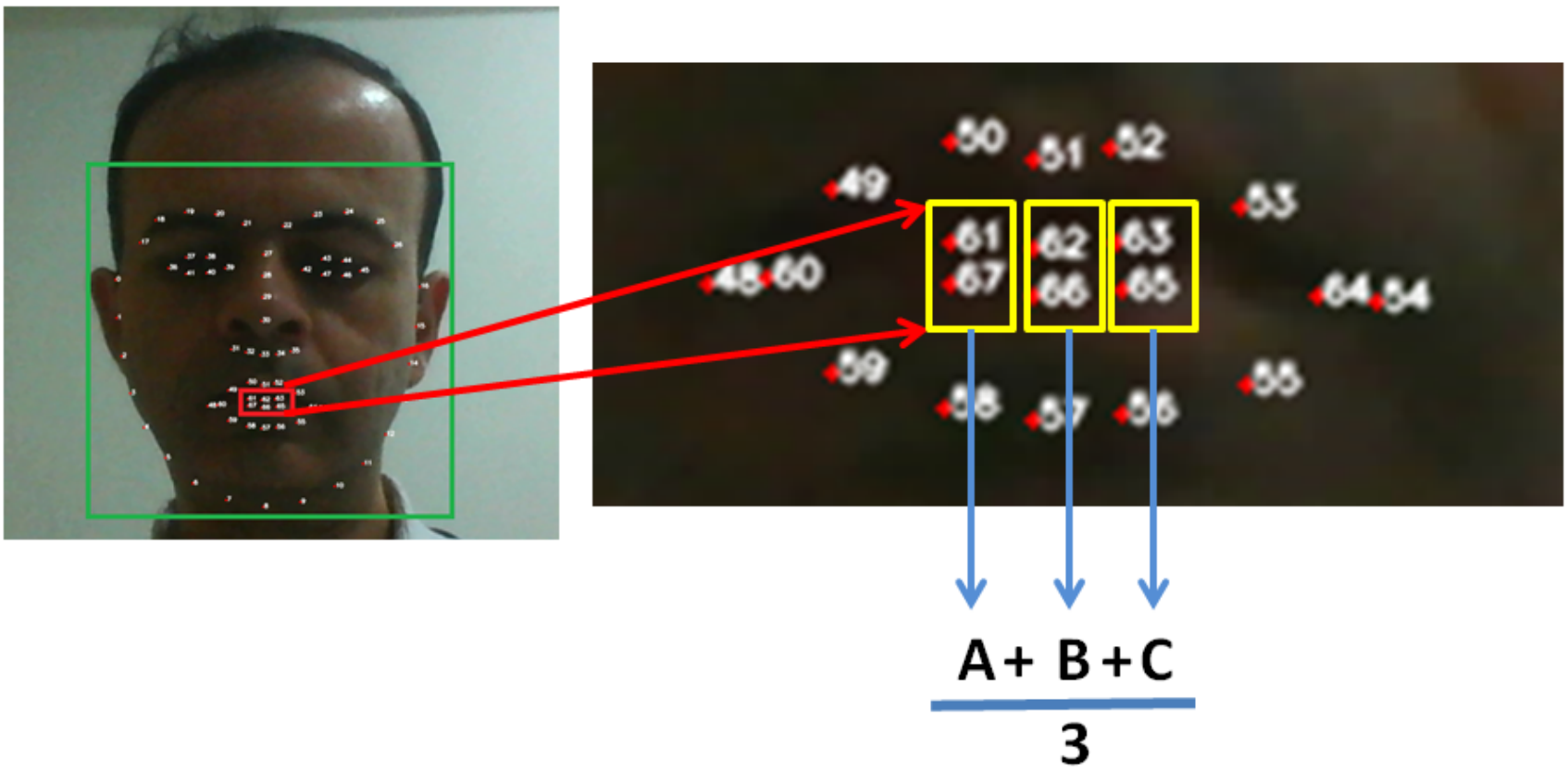
https://github.com/atulapra/Emotion-detection



Enhancement Modules

Talking Detection (RNN)

https://github.com/sachinsdate/lip-movement-net



Enhancement Modules

Caricaturing

https://github.com/Azmarie/Face-Morphing

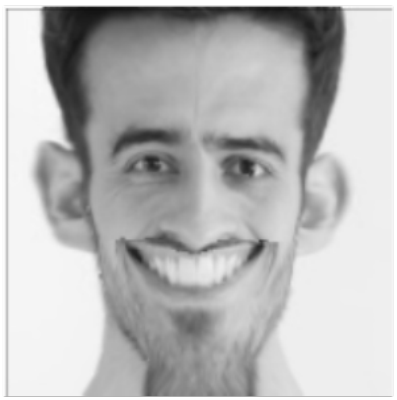
Average Face

Average Landmarks

Veridical Face

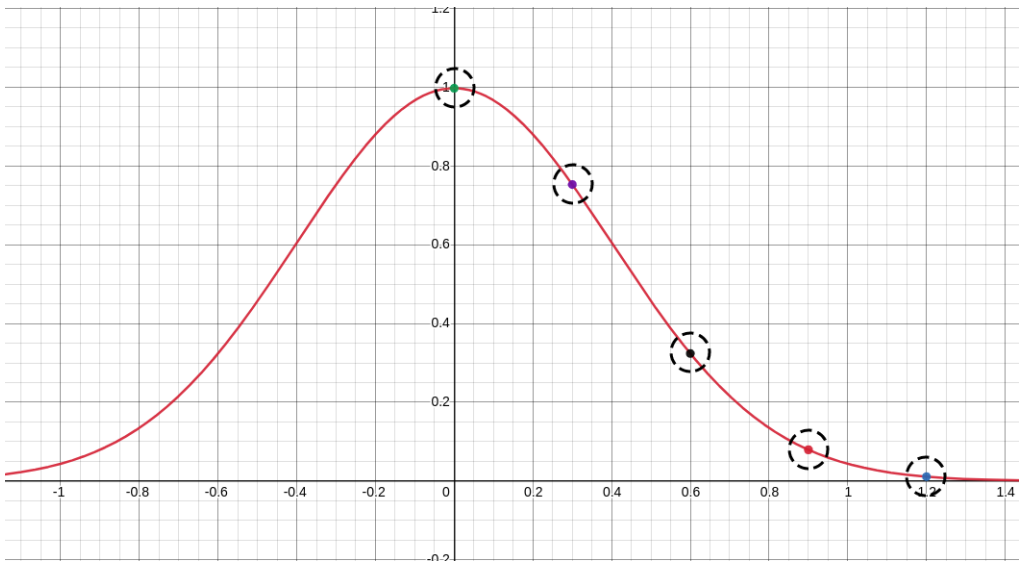
Face Detection
Face Landmarks
Delaunay Triangulation

Caricatured Face



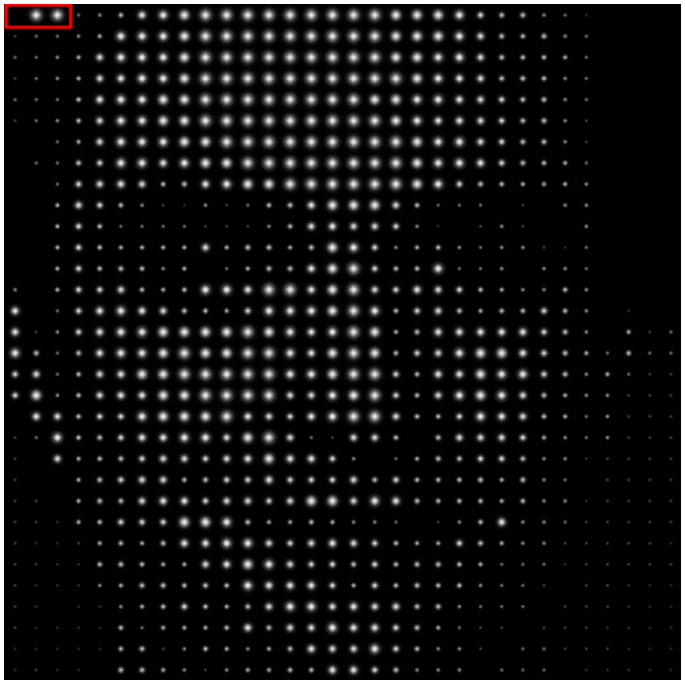
Simulation Mode

1. Blur Color Modulated (BCM)
2. Blur Size Modulate (BSM)
3. Array Color Modulated (ACM)
4. Array Size Modulated (ASM)



Enhancement Mode

1. VJFR or SFR
2. Enhancement Module
   1. Mere magnification
   2. Caricaturing
   3. FSHE
   4. Talking Detection
   5. Emotion Recognition



Computer Screen Experiment

Array Size Modulated (ASM) phosphenes

Experiment Groups

1. Control Group
2. Viola Jones Face region magnification group (VJm)
3. Viola Jones FSHE group (VJhe)
4. Viole Jones Caricaturing group (VJc)
5. Viola Jones talking detection group (VJtd)
6. Viola Jones emotion recognition group (VJer)

Experiment Tests

1. Recognition test
2. Talking test
3. Distinction test

Virtual Reality Experiment

Array Color Modulated (ACM) phosphenes

Experiment Groups

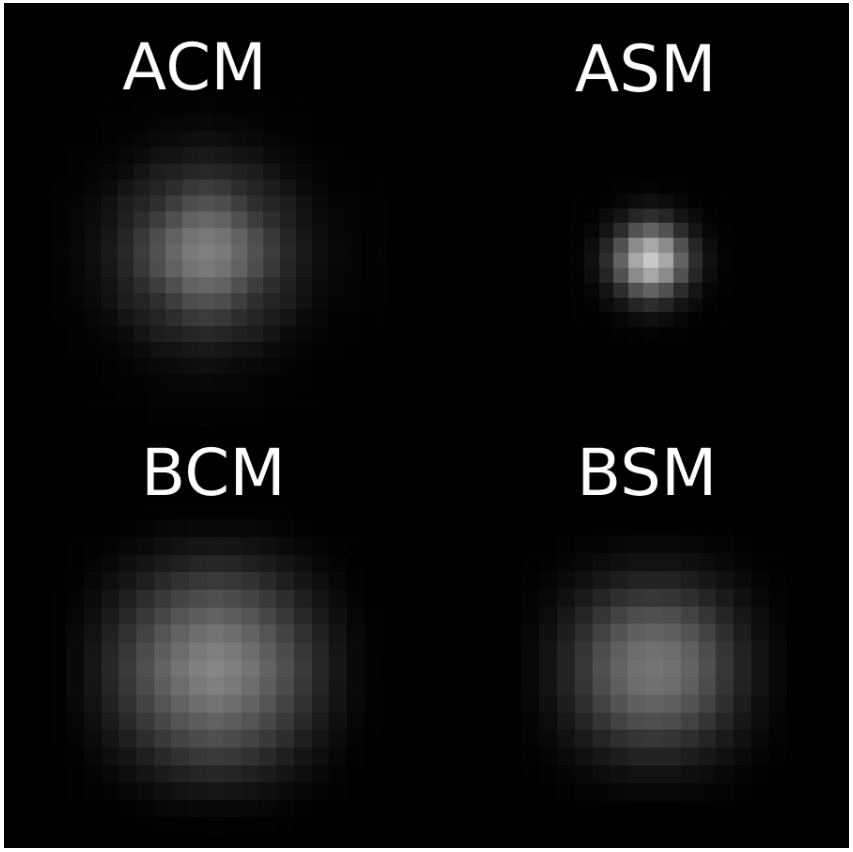
1. Control Group
2. Viola Jones Face region magnification group (VJm)
3. Viole Jones Caricaturing group (VJc)

Experiment Tests

1. Recognition test
2. Expressions test
3. Distinction test

Simulation

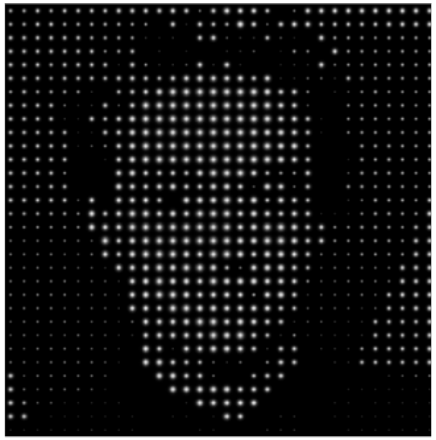
Single phosphene



Simulation

Whole face

ASM



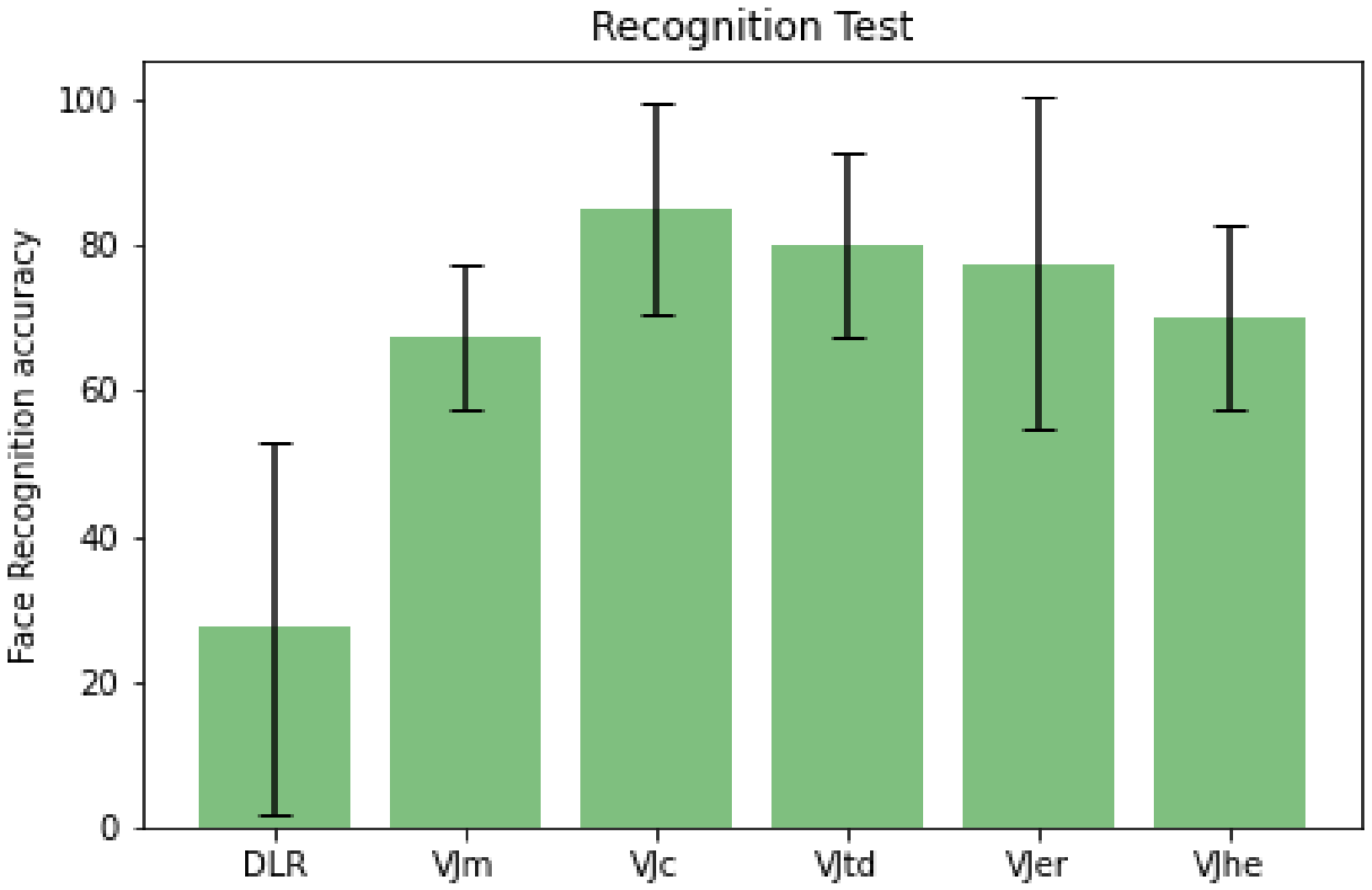
BSM

ACM

BCM

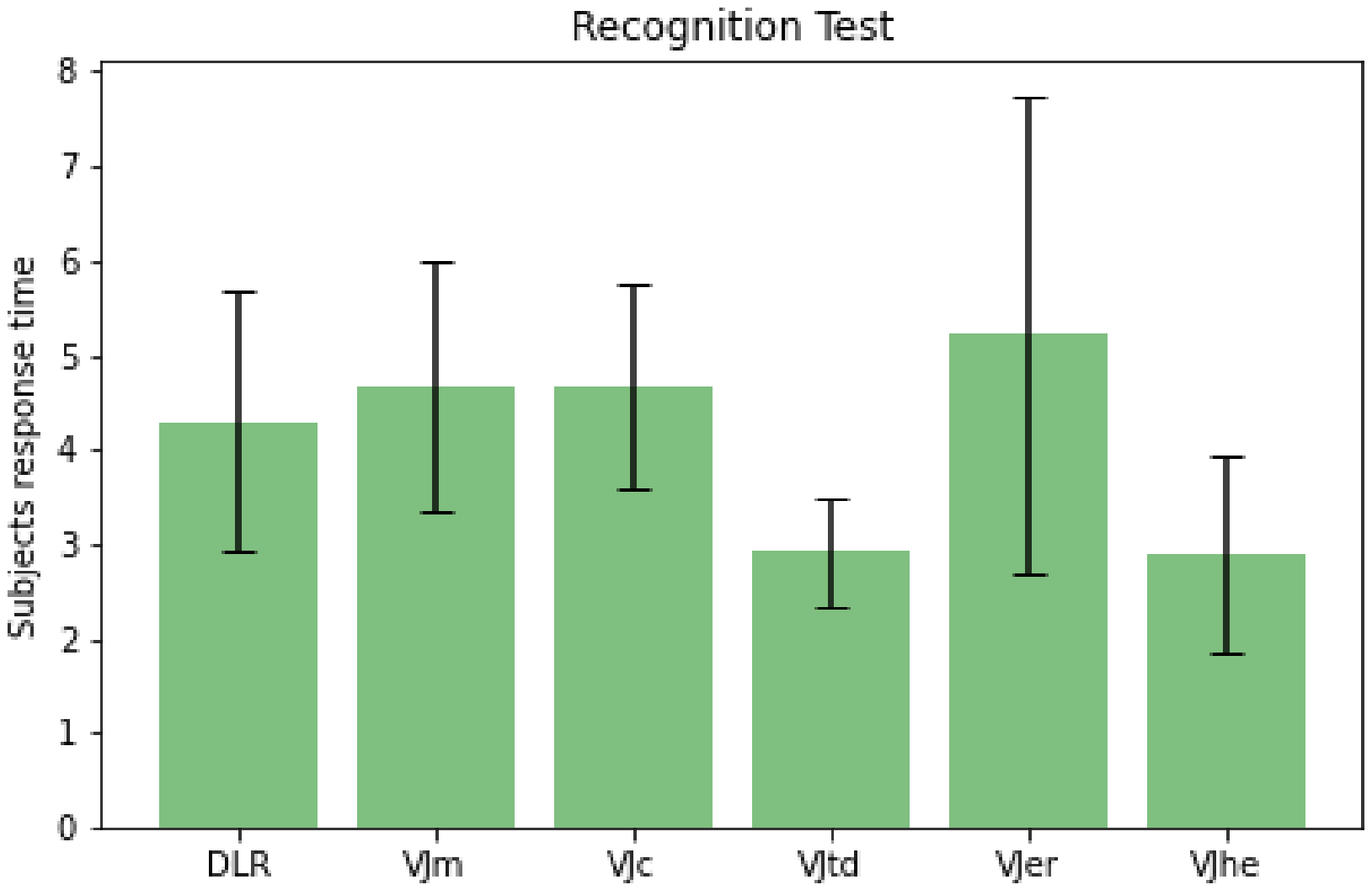
Computer Screen Experiment

Recognition Test accuracy



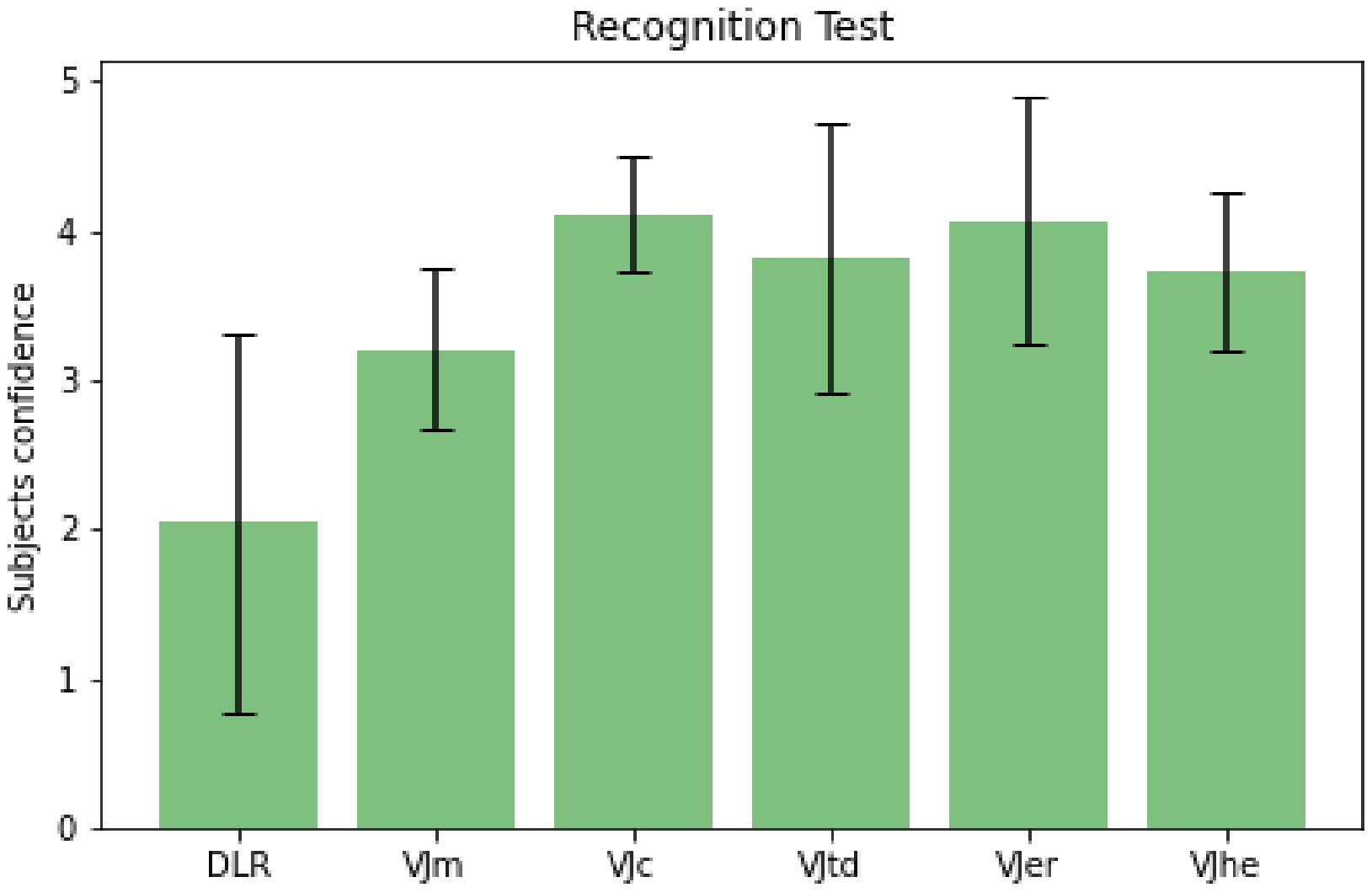
Computer Screen Experiment

Recognition Test response times



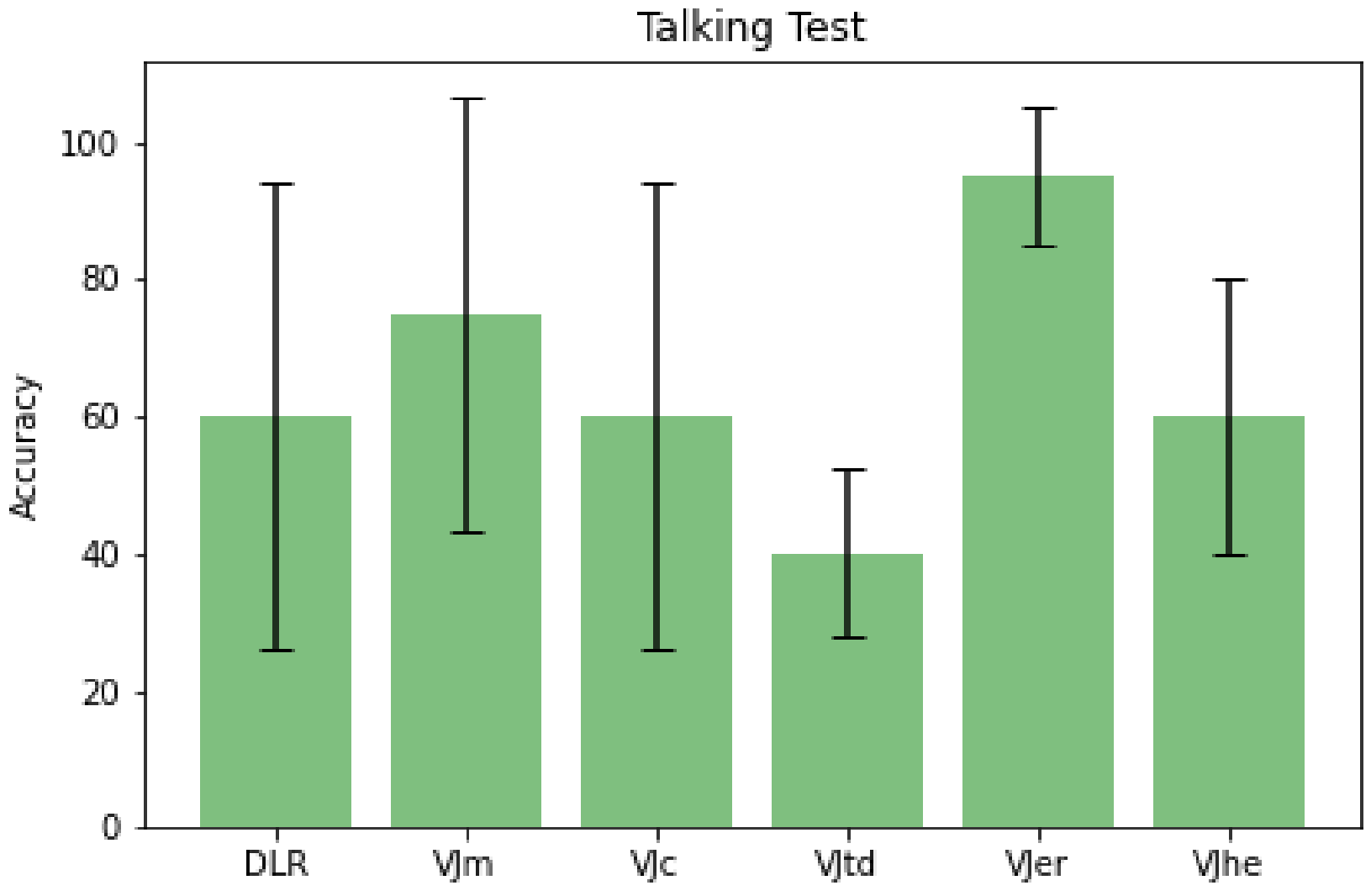
Computer Screen Experiment

Recognition Test confidence levels



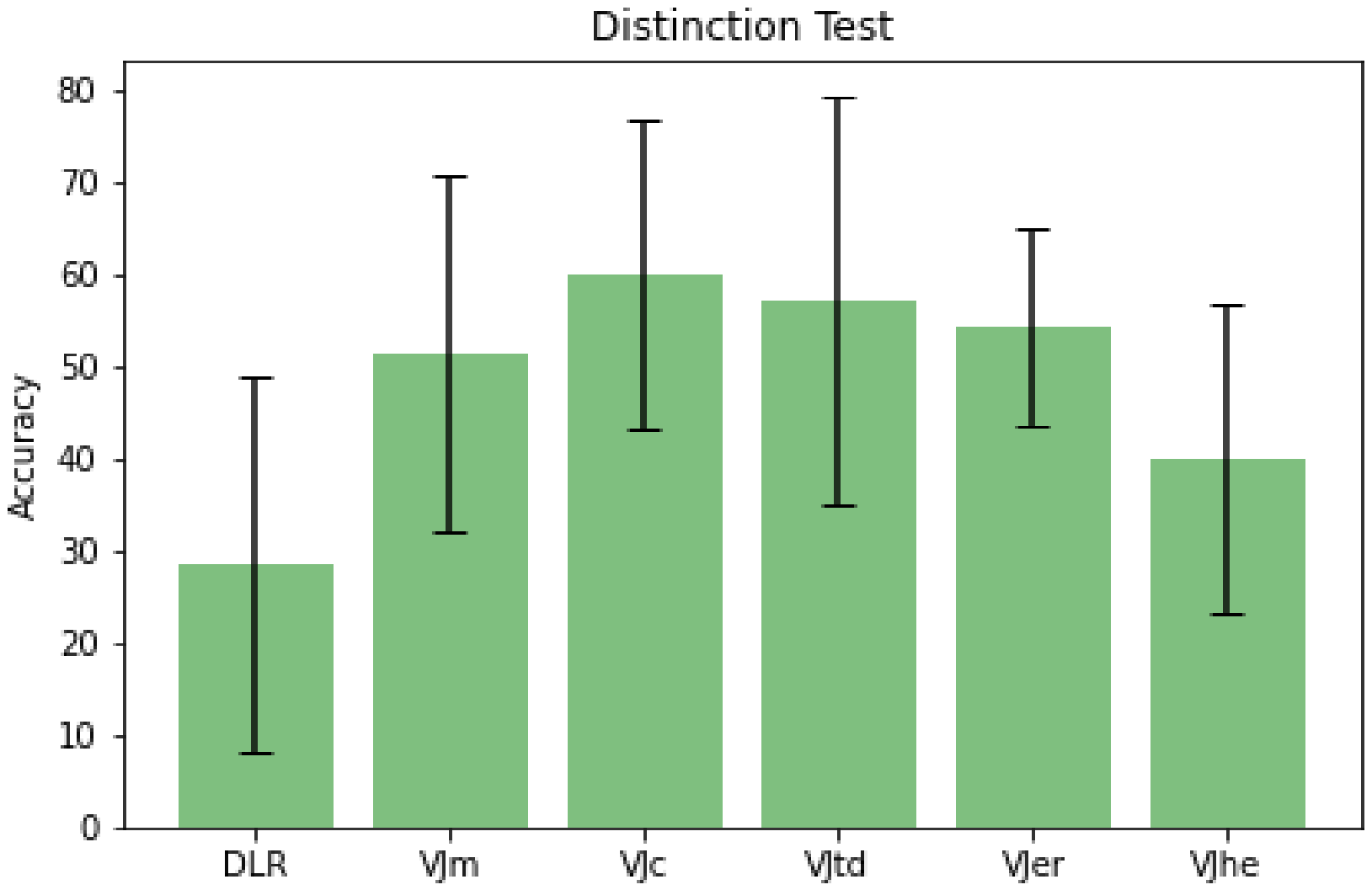
Computer Screen Experiment

Talking Test accuracy



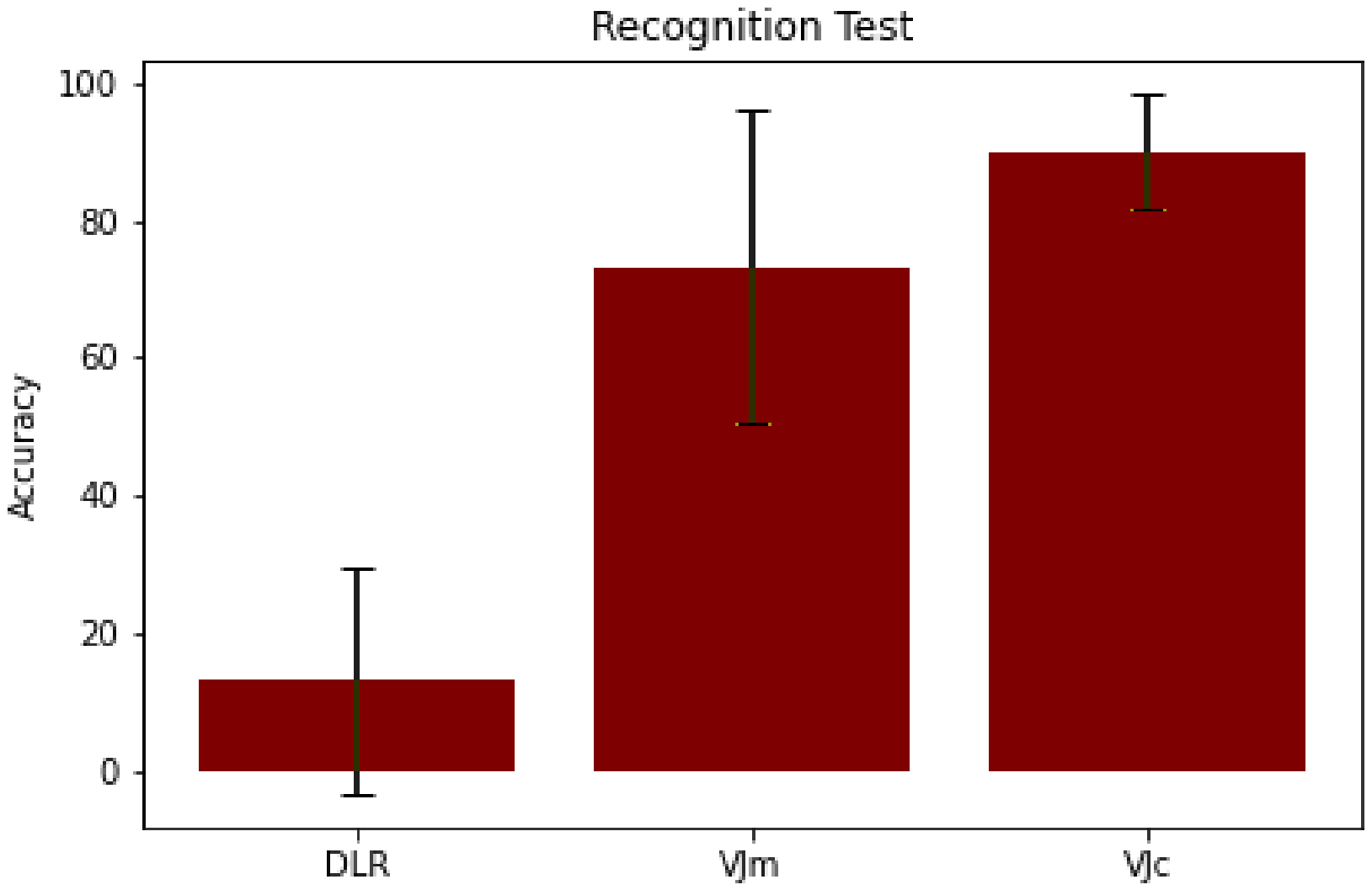
Computer Screen Experiment

Distinction Test accuracy



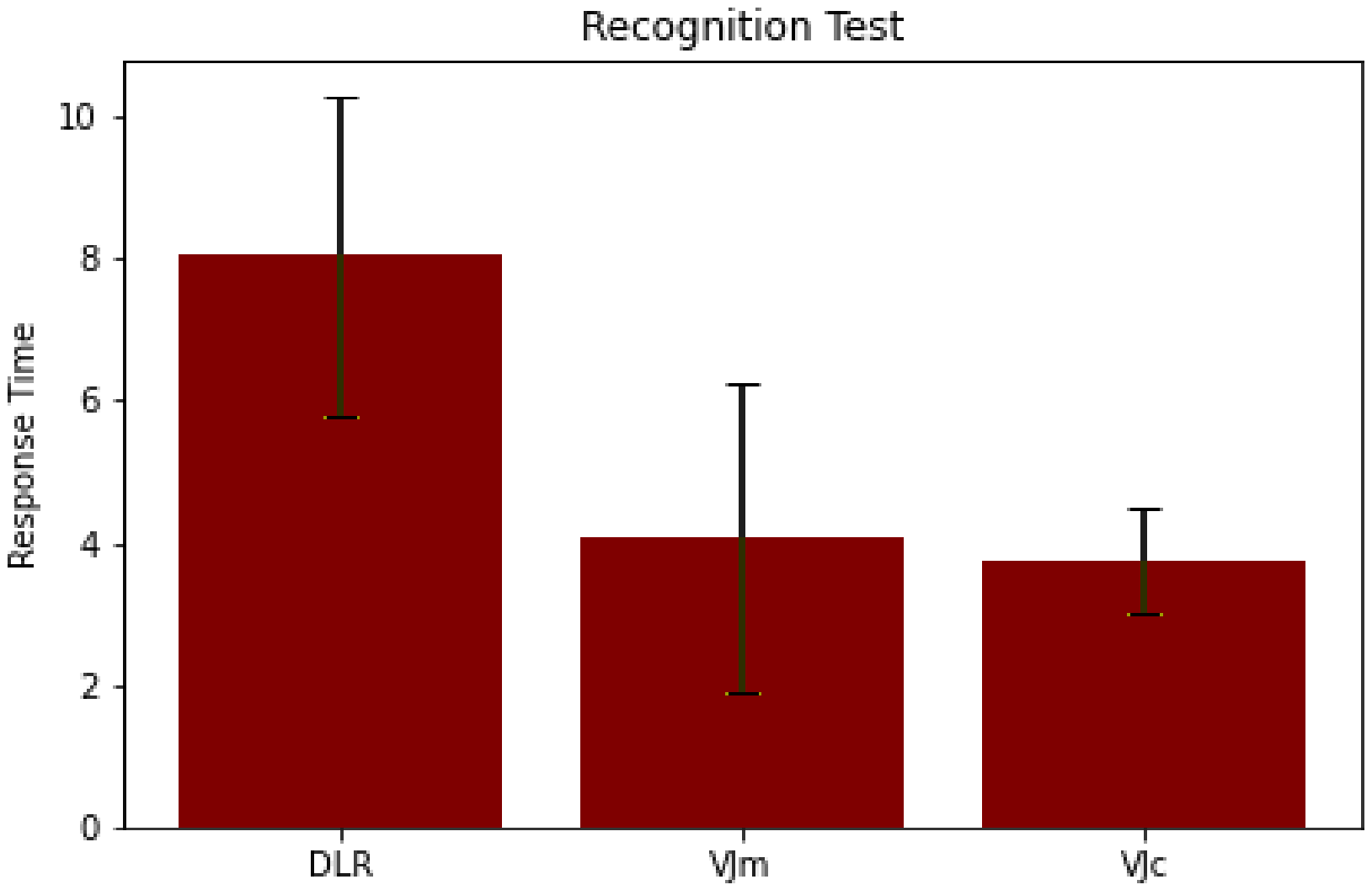
Virtual Reality Experiment

Recognition Test accuracy



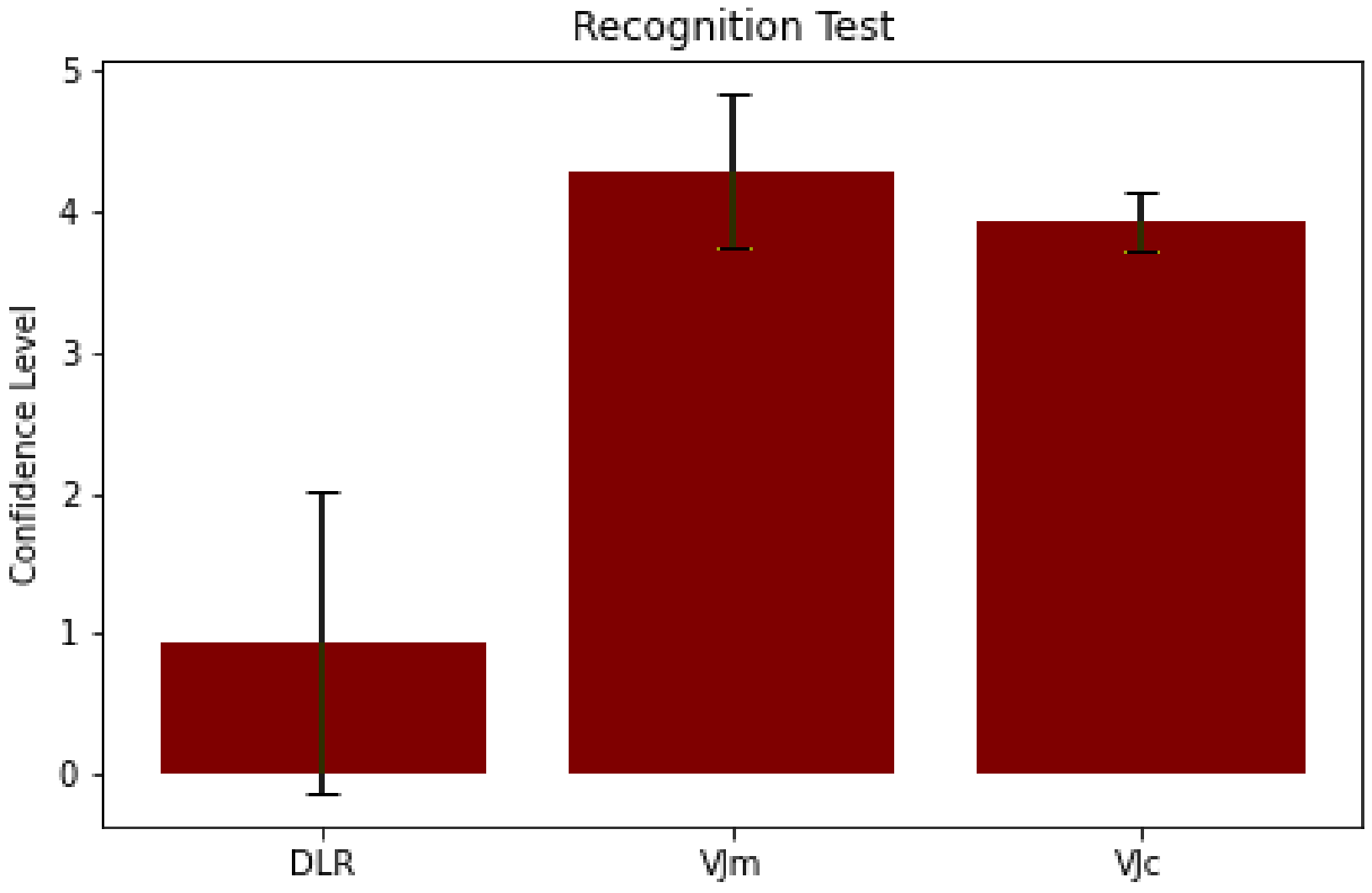
Virtual Reality Experiment

Recognition Test response times



Virtual Reality Experiment

Recognition Test confidence levels

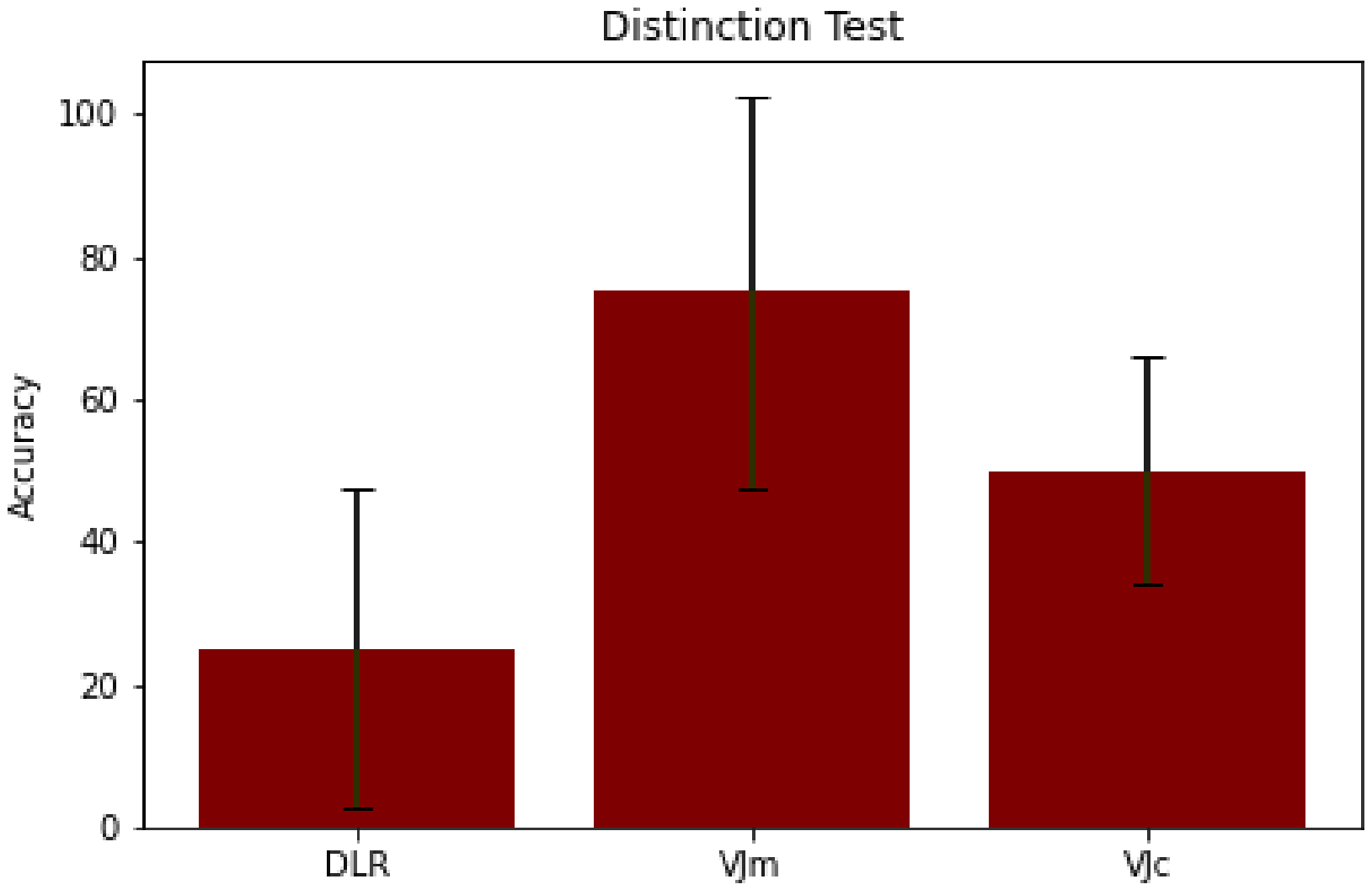


Virtual Reality Experiment

Expressions Test accuracy

Virtual Reality Experiment

Distinction Test accuracy



Applicability in Real-time

- Experiment images and videos were phosphenated in real-time
- No lag was observed by the subjects nor the tests conductor
- The total duration of simulating and playing a video did not exceed the actual duration of the video

Conclusion

- Emotion Recognition, Talking Detection, and FSHE modules did not achieve their hoped results
- Region of interest magnification is a very promising approach
- Caricaturing made facial features more evident

Limitation and Future work

Limitations

- Limited number of participants
- Limited number of questions
- No access to actual implantees
- Caricaturing algorithm was of low quality (to meet real-time requirements)
- Instability of visual field
- Highlighted emotions were distracting
- Talking person detection was depriving the subjects from having total control

Limitation and Future work

Future work

- More participants
- More questions
- Actual implantees experiment
- Compare and contrast between mere magnification and magnification with caricaturing
- Higher quality Caricaturing algorithm (e.g. CariGAN [8])
- More facial landmarks (e.g. Mediapipe 468 landmarks)
- Face tracking for more stability of the visual field
- Separate assessment of each enhancement module

[1] E. Fernandez, "Development of visual neuroprostheses: trends and challenges," *Bioelectronic medicine*, vol. 4, no. 1, pp. 1–8, 2018.

[2] S. C. Chen, G. J. Suaning, J. W. Morley, and N. H. Lovell, "Simulating prosthetic vision: I. visual models of phosphenes," *Vision research*, vol. 49, no. 12, pp. 1493–1506, 2009.

[3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, vol. 1, pp. I–I, Ieee, 2001.

[4] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1867–1874, 2014.

[5] S. S. Agaian, B. Silver, and K. A. Panetta, "Transform coefficient histogram-based image enhancement algorithms

using contrast entropy," *IEEE transactions on image processing*, vol. 16, no. 3, pp. 741–758, 2007.

[6] J. Wang, X. Wu, Y. Lu, H. Wu, H. Kan, and X. Chai, "Face recognition in simulated prosthetic vision: face detection-based image processing strategies," *Journal of neural engineering*, vol. 11, no. 4, p. 046009, 2014.

[7] J. L. Irons, T. Gradden, A. Zhang, X. He, N. Barnes, A. F. Scott, and E. McKone, "Face identity recognition in simulated prosthetic vision is poorer than previously reported and can be improved by caricaturing," *Vision research*, vol. 137, pp. 61–79, 2017.

[8] K. Cao, J. Liao, and L. Yuan, "Carigans: Unpaired photo-to-caricature translation," *arXiv preprint arXiv:1811.00222*, 2018.

Thank You
Any Questions?