

Jonathan Skaza¹ Shravan Murlidaran² Apurv Varshney³ Ziqi Wen³ William Wang³ Miguel P. Eckstein¹²³ Michael Beyeler¹²³

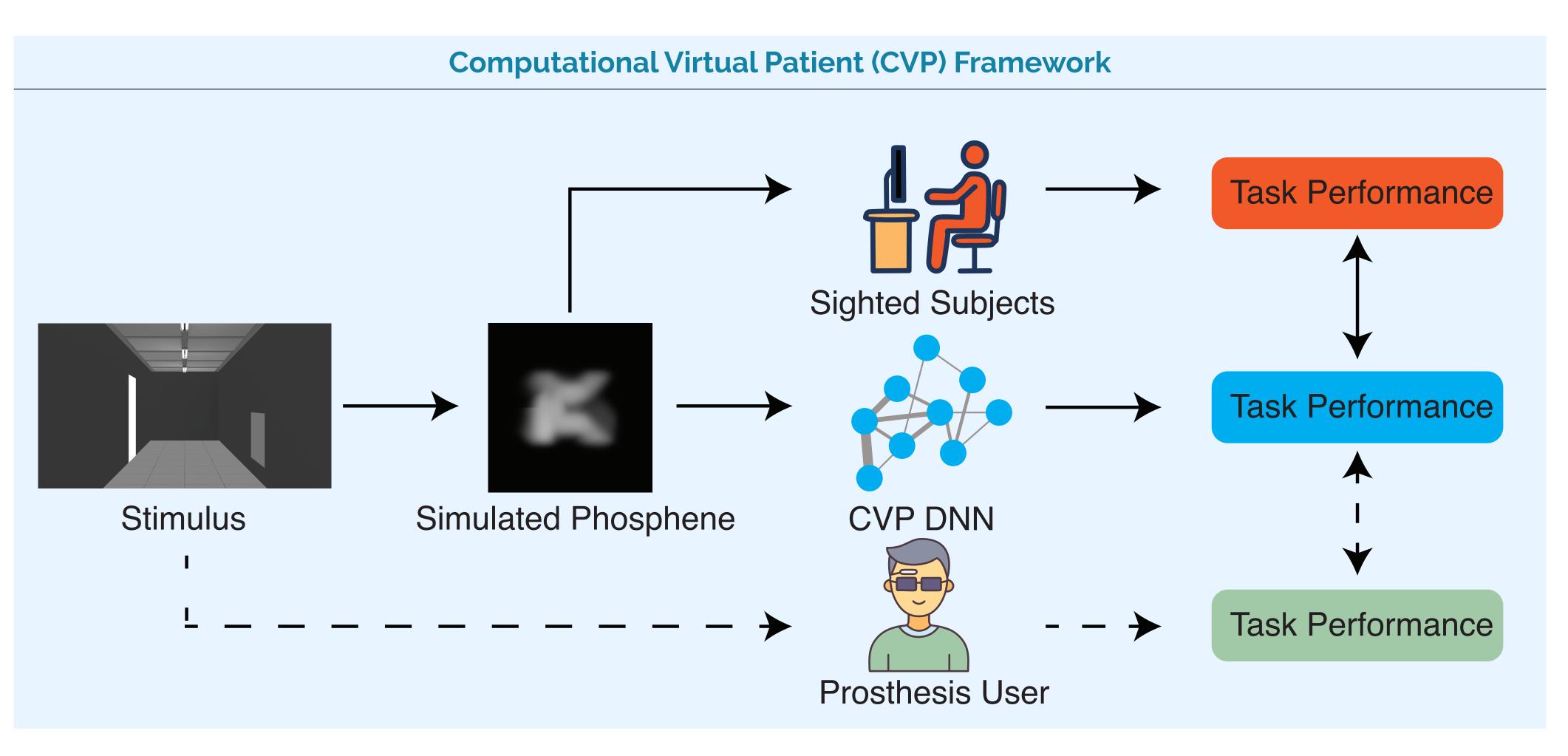


¹Graduate Program in Dynamical Neuroscience ²Department of Psychological and Brain Sciences ³Department of Computer Science

University of California, Santa Barbara

Why Simulate Prosthetic Perception?

- Current visual prostheses elicit rudimentary phosphenes that differ significantly from natural vision
- Users are typically evaluated post-implantation using the Functional Low-Vision Observer Rated Assessment (FLORA) (Geruschat et al., 2015)
- A fundamental challenge is predicting perceptual capabilities before implantation



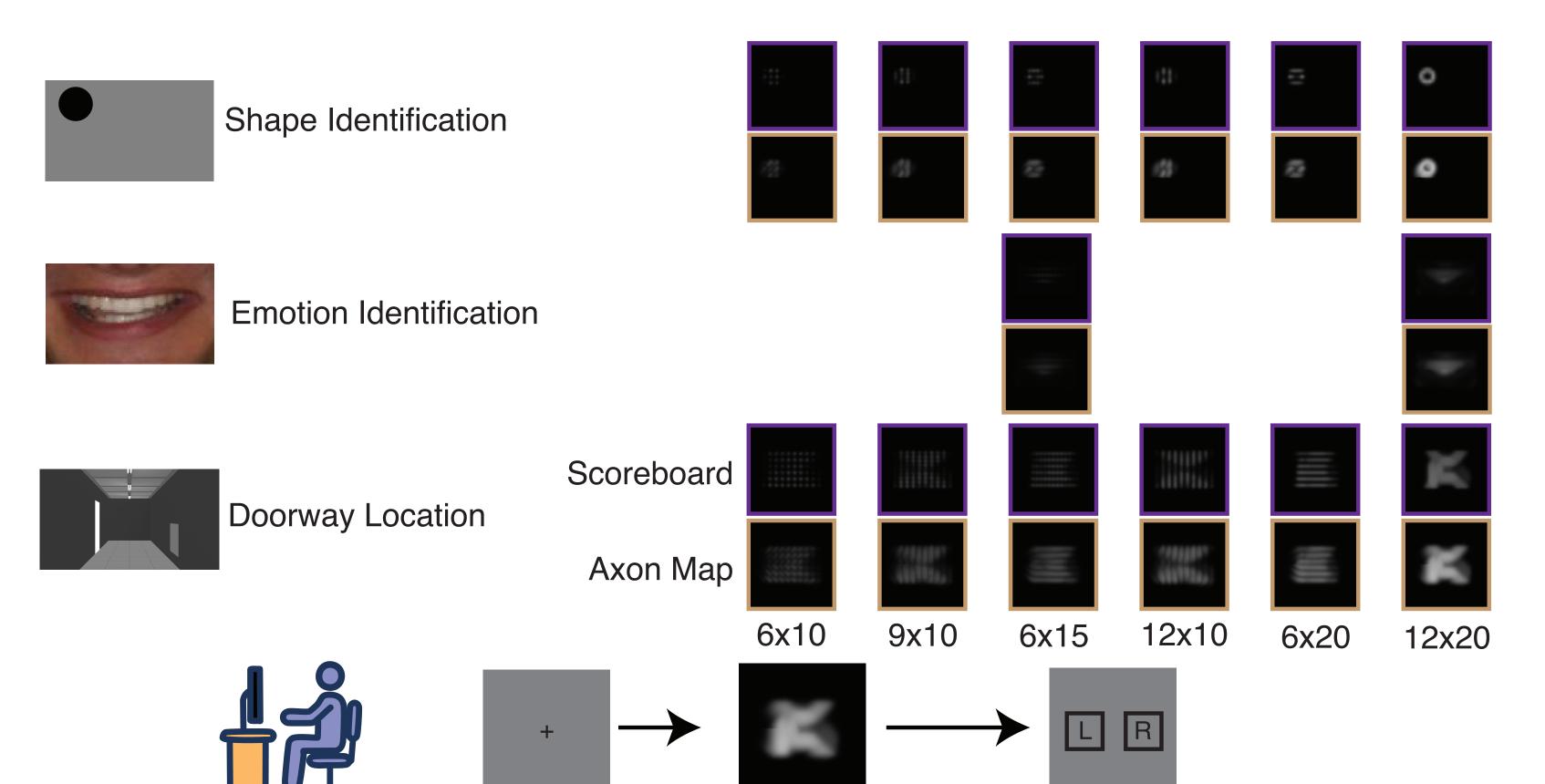
Simulating and Evaluating Perception with Prosthetic Vision

Tasks: Shape identification, emotion identification, person facing direction, doorway location, person location, window location Phosphene Simulation:

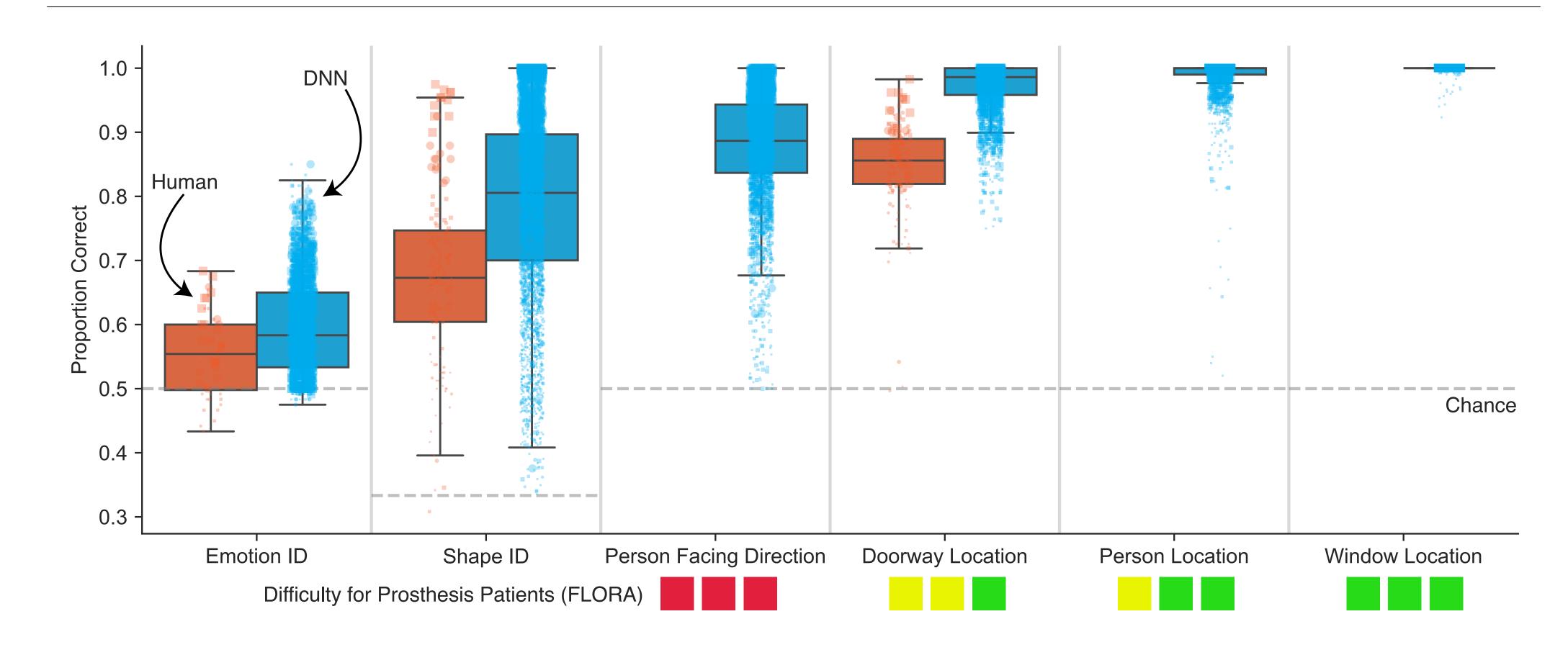
- Two phosphene generation models: the 'scoreboard' model (producing localized points of light) and the 'axon map' model (which incorporates the anatomical layout of retinal ganglion cell axons) (Beyeler et al., 2017, 2019)
- Six electrode configurations: 6×10, 9×10, 6×15, 12×10, 6×20, and 12×20

Evaluation:

- Human observers in a forced-fixation, forced-choice task
- Deep neural networks (DNNs) trained to perform the same tasks: ResNet, VGG, EfficientNet, RegNetY, MobileNetV2, Vision Transformers (ViT), and ConvNeXt
- We compared human and DNN performance across all tasks, electrode configurations, and phosphene models

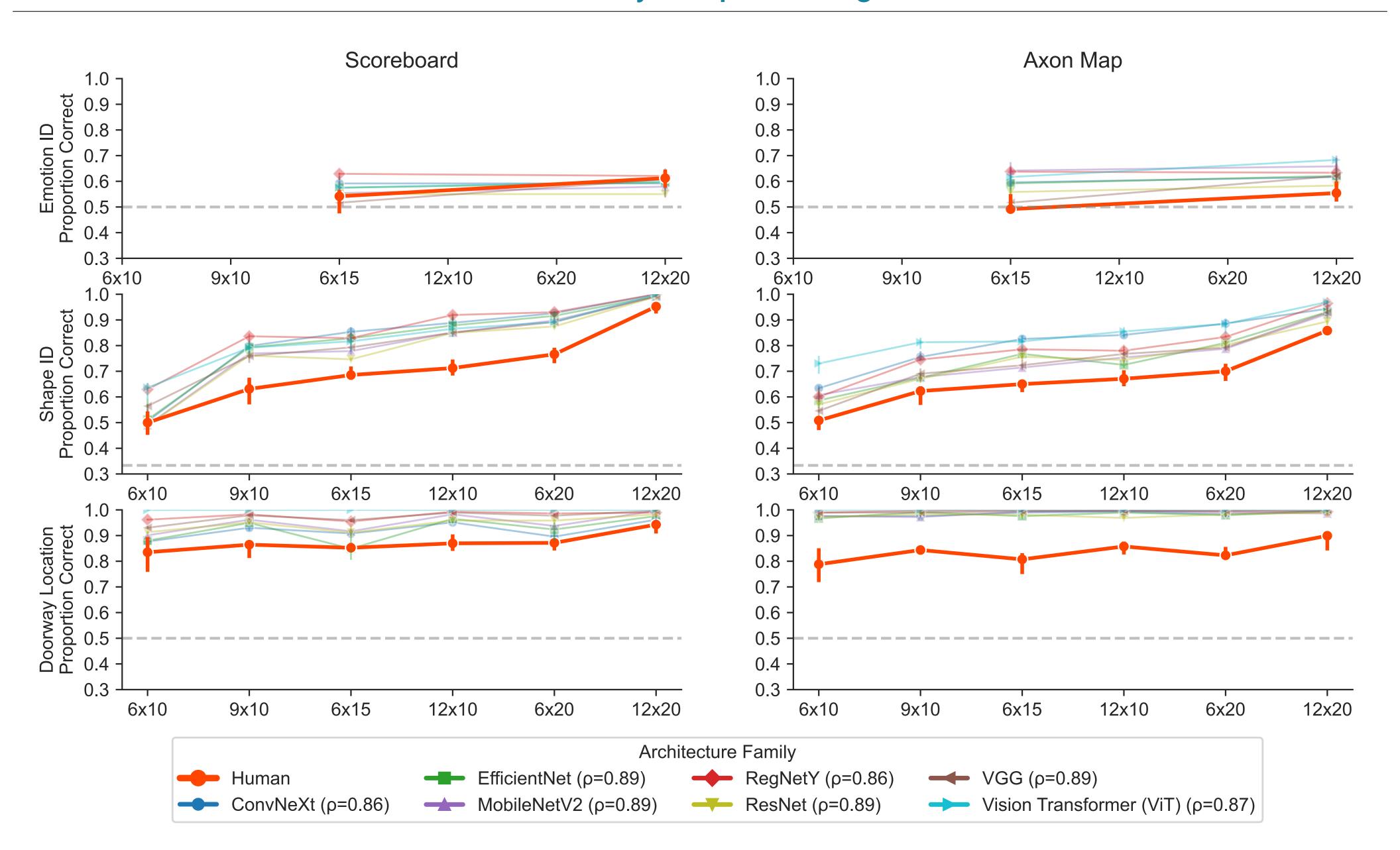


Human-Model Agreement Across Visual Tasks



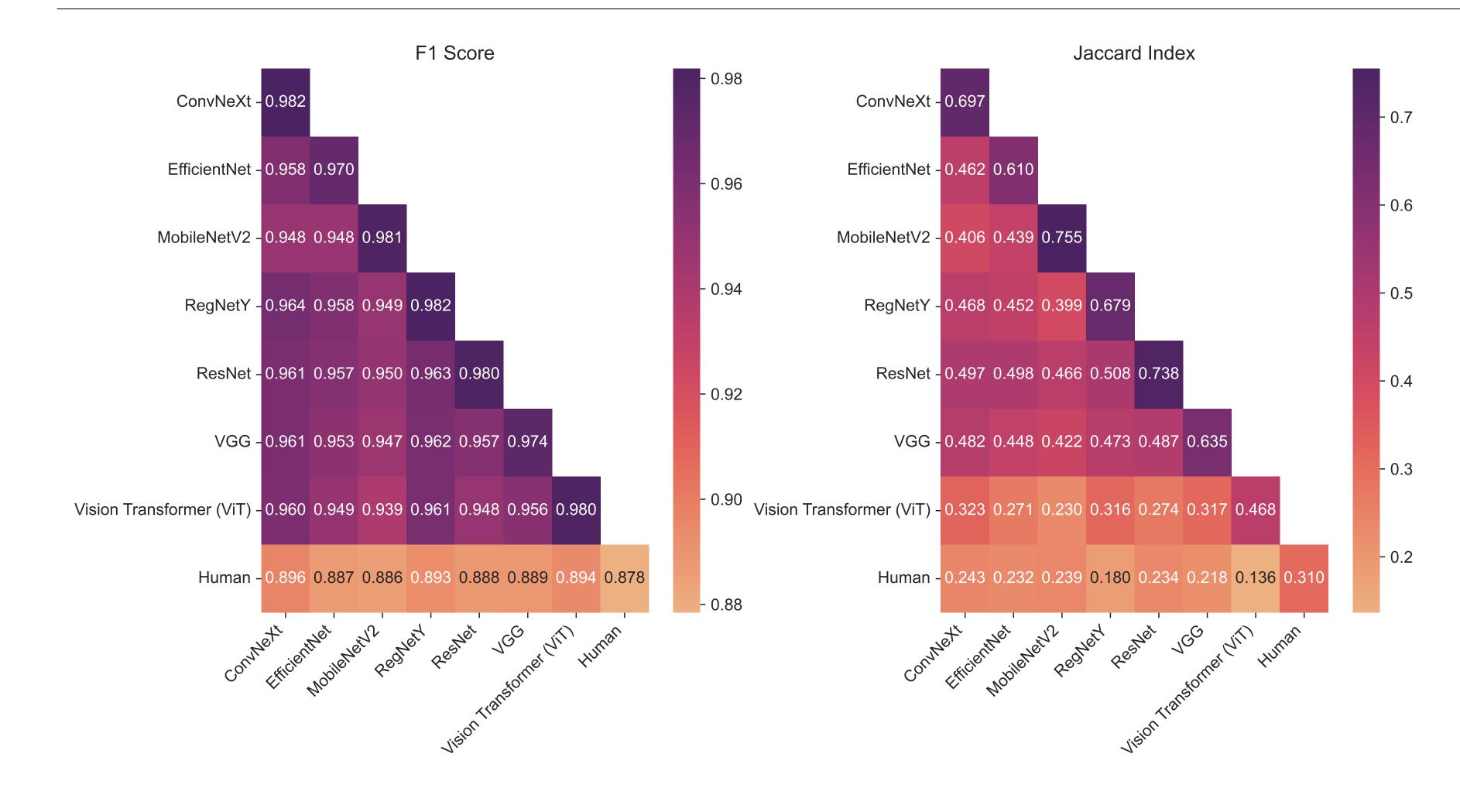
- DNNs tended to outperform human participants
- Human observers and DNNs displayed similar performance variation across tasks
- Task-level performance rankings were consistent between human psychophysics and DNNs
- FLORA task difficulty assessments were preserved in DNN performance patterns

Shared Sensitivity to Implant Configurations



- Humans and DNNs showed strong agreement in performance across implant designs (Spearman's ρ = 0.86 to 0.89)
- Human-DNN alignment persisted across both 'scoreboard' and 'axon map' models, and across DNN architectures, despite perceptual distortions
- Performance generally increased with higher electrode counts
- In both humans and DNNs, shape task performance declined under the 'axon map' model for high-density configurations (e.g., 12×20 , 6×20 , 12×10 ; Δ = -0.042 to -0.094; p < 0.05)

Trial-Level Agreement and Visual Attribution



- F1 score (agreement on correct classifications) was consistently high between humans and DNNs:
- $2 \cdot (both correct)$ $2 \cdot (both correct) + (A wrong, B correct) + (A correct, B wrong)$
- Jaccard index (overlap in misclassified items) revealed distinct misclassification patterns:

$$Jaccard = \frac{(both wrong)}{(A wrong) \cup (B wrong)}$$

 Attribution analysis: Separate CNNs trained to predict (i) ground truth and (ii) human choices revealed divergent saliency patterns in cases of disagreement, suggesting reliance on different perceptual cues

Key Takeaways

Computational Virtual Patients (CVPs) offer a powerful framework for predicting task difficulty and assessing the impact of electrode configuration on performance. Our results highlight their potential as a tool for guiding task-specific evaluation and enhancing visual prosthesis design.

Acknowledgements

Ground Truth DNN: Neutral

This research was supported by the UC Noyce Initiative, which fosters collaborative research in digital innovation across University of California campuses. For more information, visit https://ucnoyce.org/.

This is joint work with the Bionic Vision Lab at UCSB.

Ground Truth DNN: Right

Ground Truth DNN: Square

References

Michael Beyeler, Geoffrey M Boynton, Ione Fine, and Ariel Rokem. pulse2percept: A python-based simulation framework for bionic vision. Proceedings of the 16th Python in Science Conference (SciPy), 2017.

Michael Beyeler, Devyani Nanduri, James D Weiland, Ariel Rokem, Geoffrey M Boynton, and Ione Fine. A model of ganglion axon pathways accounts for percepts elicited by retinal implants. Scientific reports, 9(1):9199, 2019.

Duane R Geruschat, Marshall Flax, Nilima Tanna, Michelle Bianchi, Andy Fisher, Mira Goldschmidt, Lynne Fisher, Gislin Dagnelie, Jim Deremeik, Audrey Smith, et al. Flora™: Phase i development of a functional vision assessment for prosthetic vision users. Clinical and Experimental Optometry, 98(4):342-347, 2015.

