



3-1-2024

Can Neural Networks Reach Human Vision Levels on Object Recognition Tasks?

Luke D. Baumel

Loyola University Chicago, lbaumel@luc.edu

Mikayla Cutler

Loyola University Chicago, mcutler2@luc.edu

Matt Hyatt

Loyola University Chicago


Joseph Tocco

Loyola University Chicago, jtocco1@luc.edu

William Friebe

Loyola University Chicago, wfriebe@luc.edu

Follow this and additional works at: https://ecommons.luc.edu/psychology_facpubs

 *See next page for additional authors*

 Part of the [Computer Engineering Commons](#), and the [Psychology Commons](#)

Recommended Citation

Baumel, Luke D.; Cutler, Mikayla; Hyatt, Matt; Tocco, Joseph; Friebe, William; Baker, Nicholas Dr.; and Thiruvathukal, George K. Dr.. Can Neural Networks Reach Human Vision Levels on Object Recognition Tasks?. Chicago Society for Neuroscience Annual Meeting, , , 2024. Retrieved from Loyola eCommons, Psychology: Faculty Publications and Other Works,

This Presentation is brought to you for free and open access by the Faculty Publications and Other Works by Department at Loyola eCommons. It has been accepted for inclusion in Psychology: Faculty Publications and Other Works by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



This work is licensed under a [Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 License](#).

© The Authors, 2024.

Authors

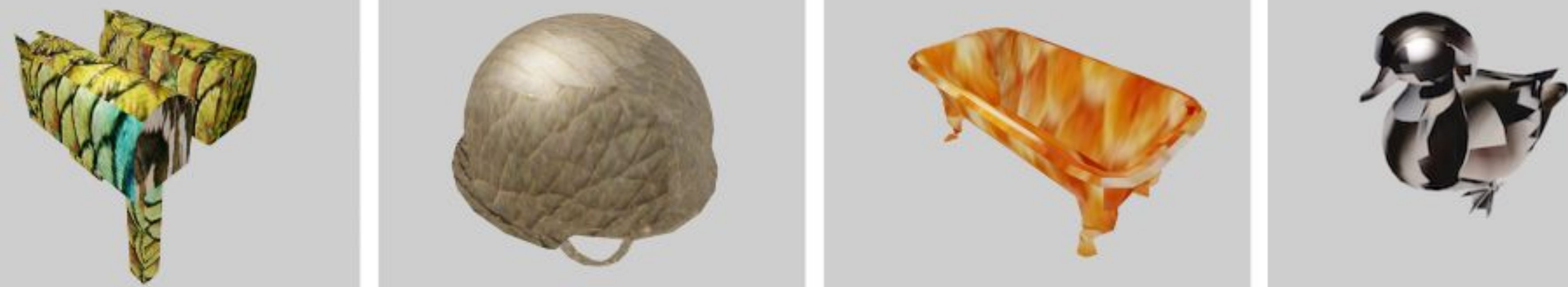
Luke D. Baumel, Mikayla Cutler, Matt Hyatt, Joseph Tocco, William Friebe, Nicholas Baker Dr., and George K. Thiruvathukal Dr.

Can Neural Networks Reach Human Vision Levels on Object Recognition Tasks?

Luke D. Baumel¹, Mikayla Cutler¹, Matthew Hyatt², Joseph Tocco¹, Will Friebe¹,
Dr. Nicholas Baker¹, Dr. George K. Thiruvathukal²

¹Department of Psychology, Loyola University Chicago, Chicago, IL ²Department of Computer Science, Loyola University Chicago, Chicago, IL

Preparing people to lead extraordinary lives



Where AI Has Failed Before

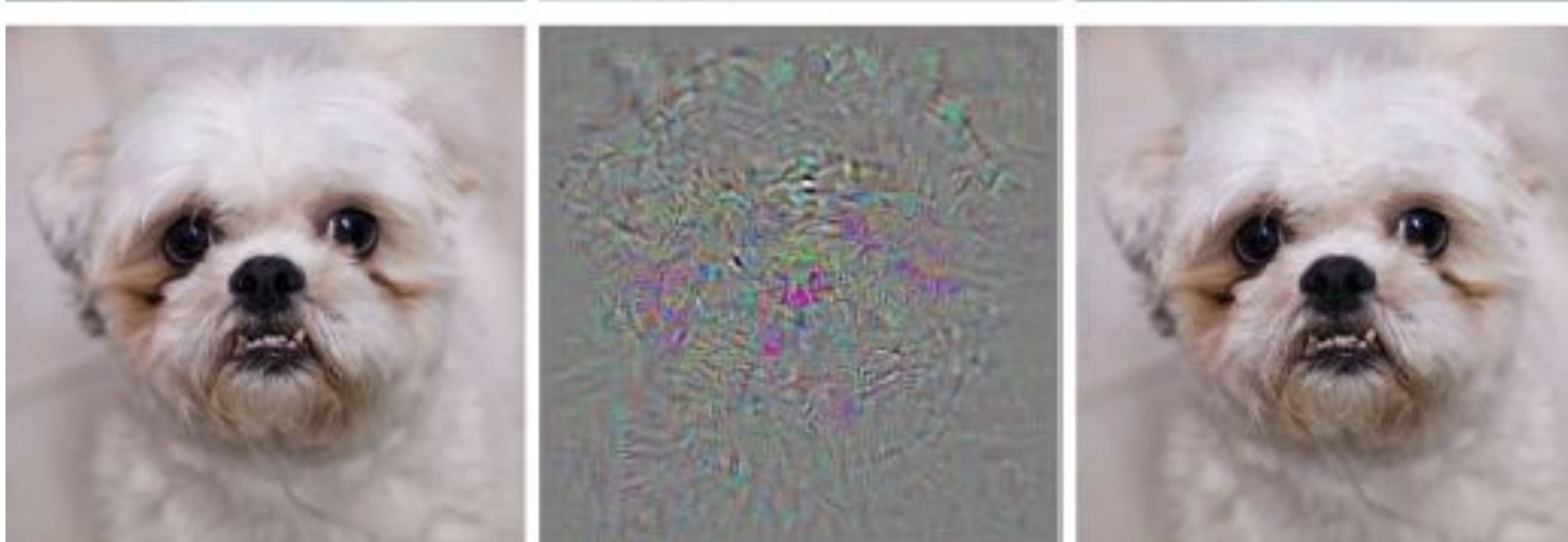


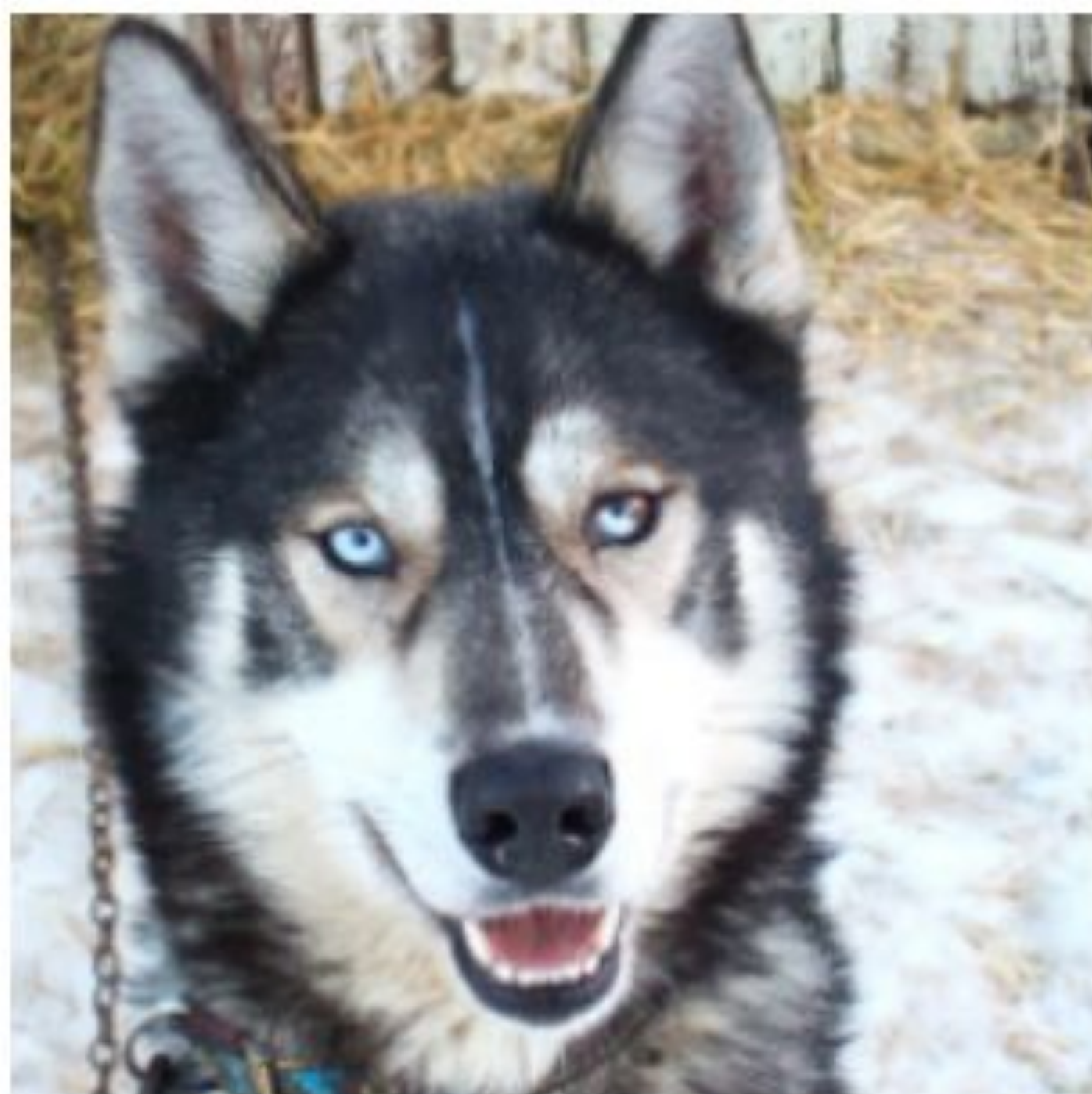
Image incorrectly classified when only a few pixels were changed (Szegedy 2014)



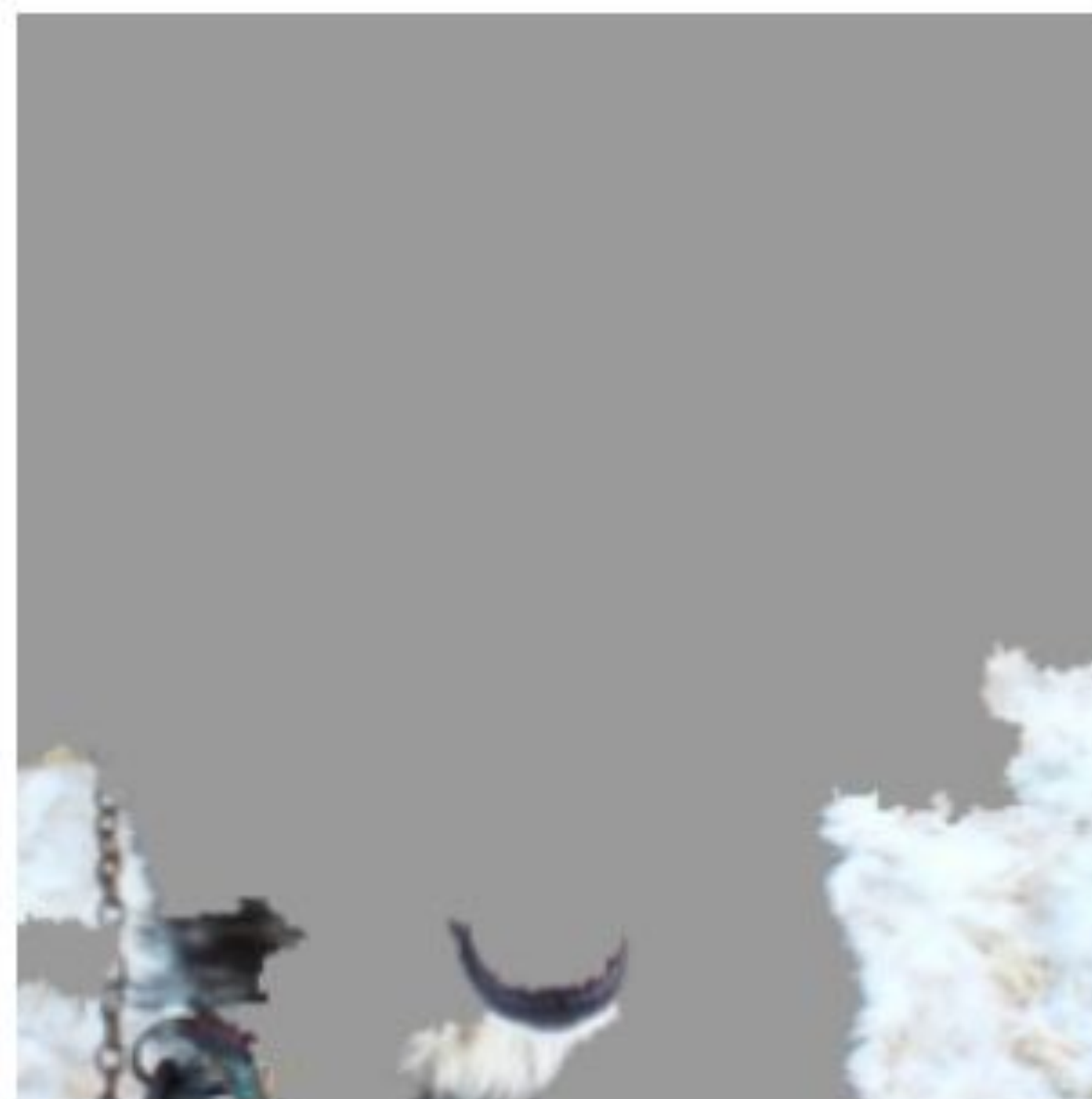
Goose
(0.006%)

Web site
(10.13%)

Glass goose incorrectly categorized as a website (Baker 2018)



(a) Husky classified as wolf



(b) Explanation

Husky misclassified as a wolf due to background (Tulio Ribeiro et al. 2016)

Our Methodologies

- Stimuli creation via scripting in Blender. Capture 12 viewpoints of 3-dimensional stimulus objects with mapped characteristic and uncharacteristic texture
- Test Convolutional Neural Networks (CCNs) with novel stimuli and sieve through data to determine what categories the CNN assigned the highest probability to
- Run human experiment using same novel stimuli to understand if shape or texture plays a more significant role in perception



Human Experiment Results

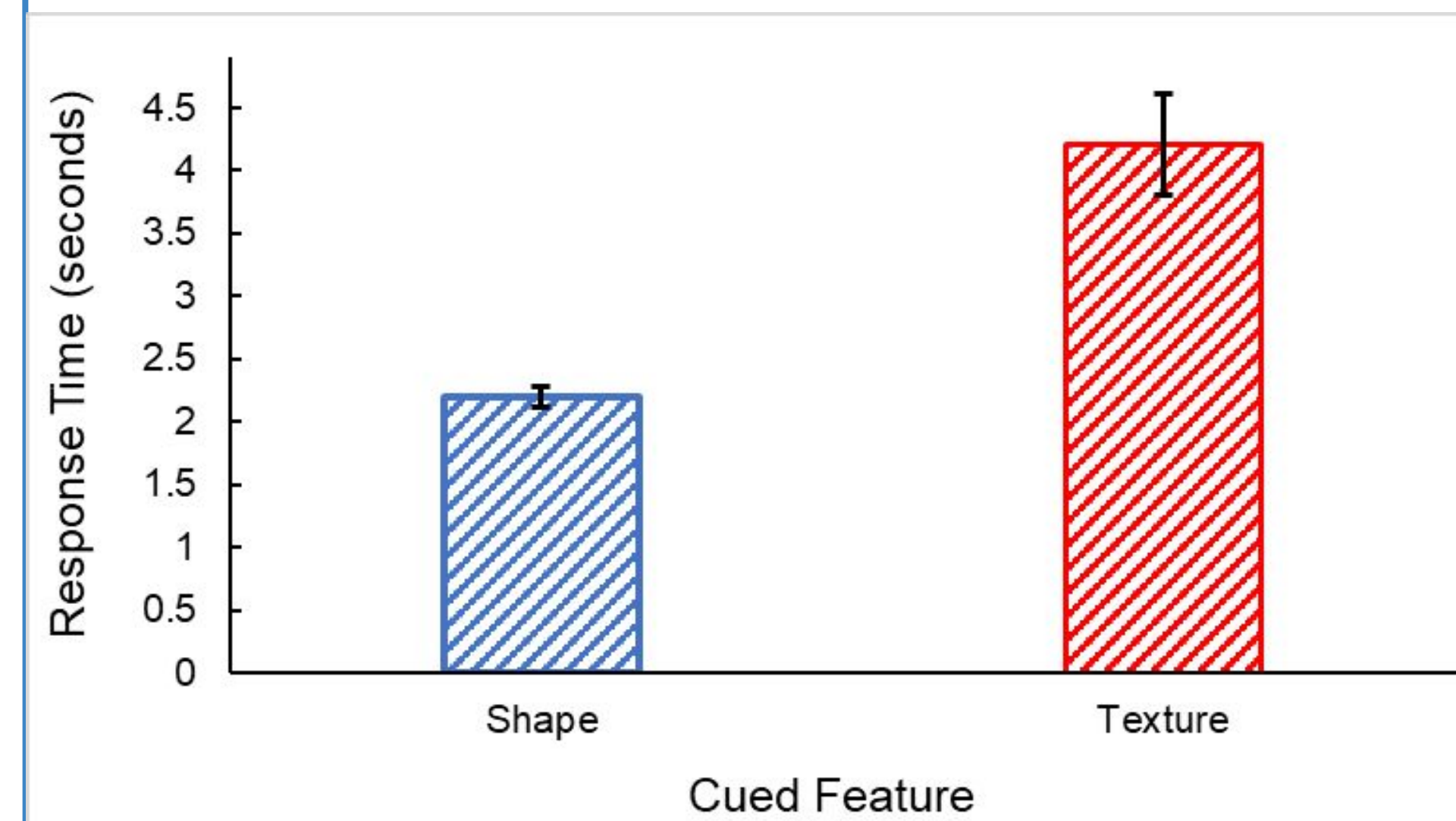


Figure 1. Response Time in Shape vs. Texture Trials
Response time is nearly double for texture trials (2.2s for shape and 4.2s for texture)

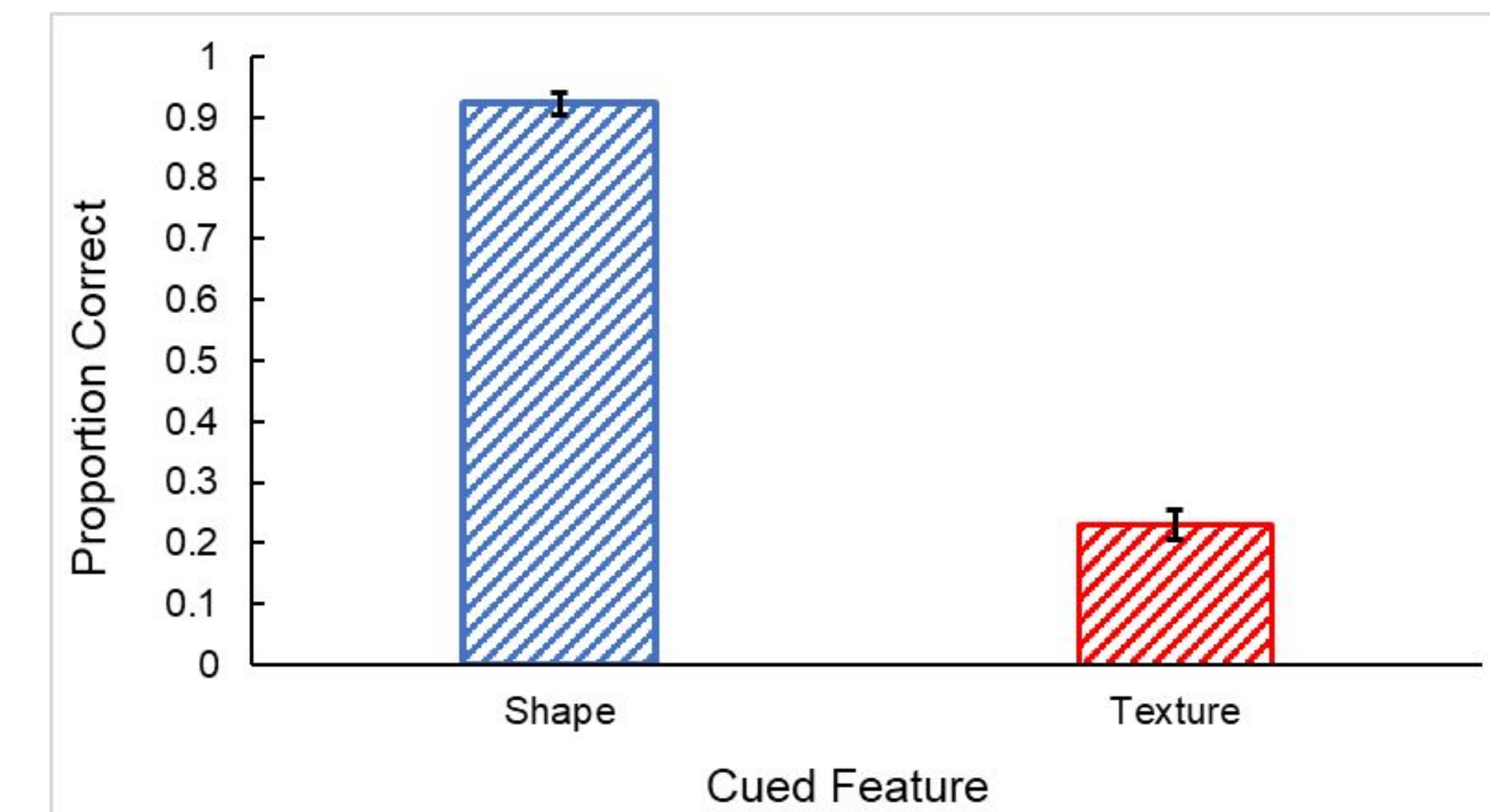


Figure 2. Prop. of Correct Selection in Shape vs. Texture Trials
Shape distinguished correctly in 92-93% of trials whereas texture distinguished correctly in ~23% of trials

Preliminary Network Results

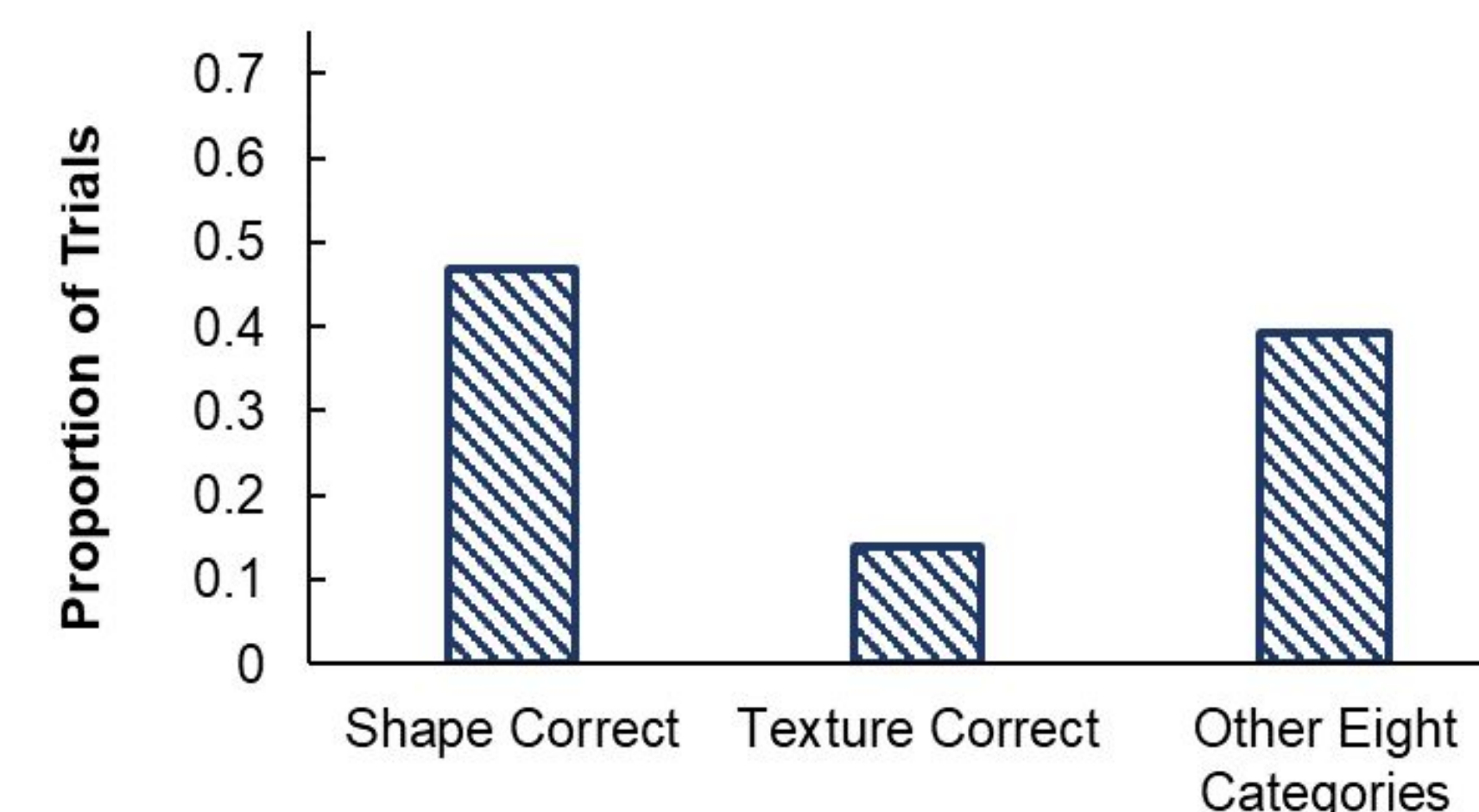


Figure 3. ResNet Classification of Image
Network correctly classified by shape in 47% of trials, texture in 13% of trials, and within the other 8 categories the other 39% of the trials

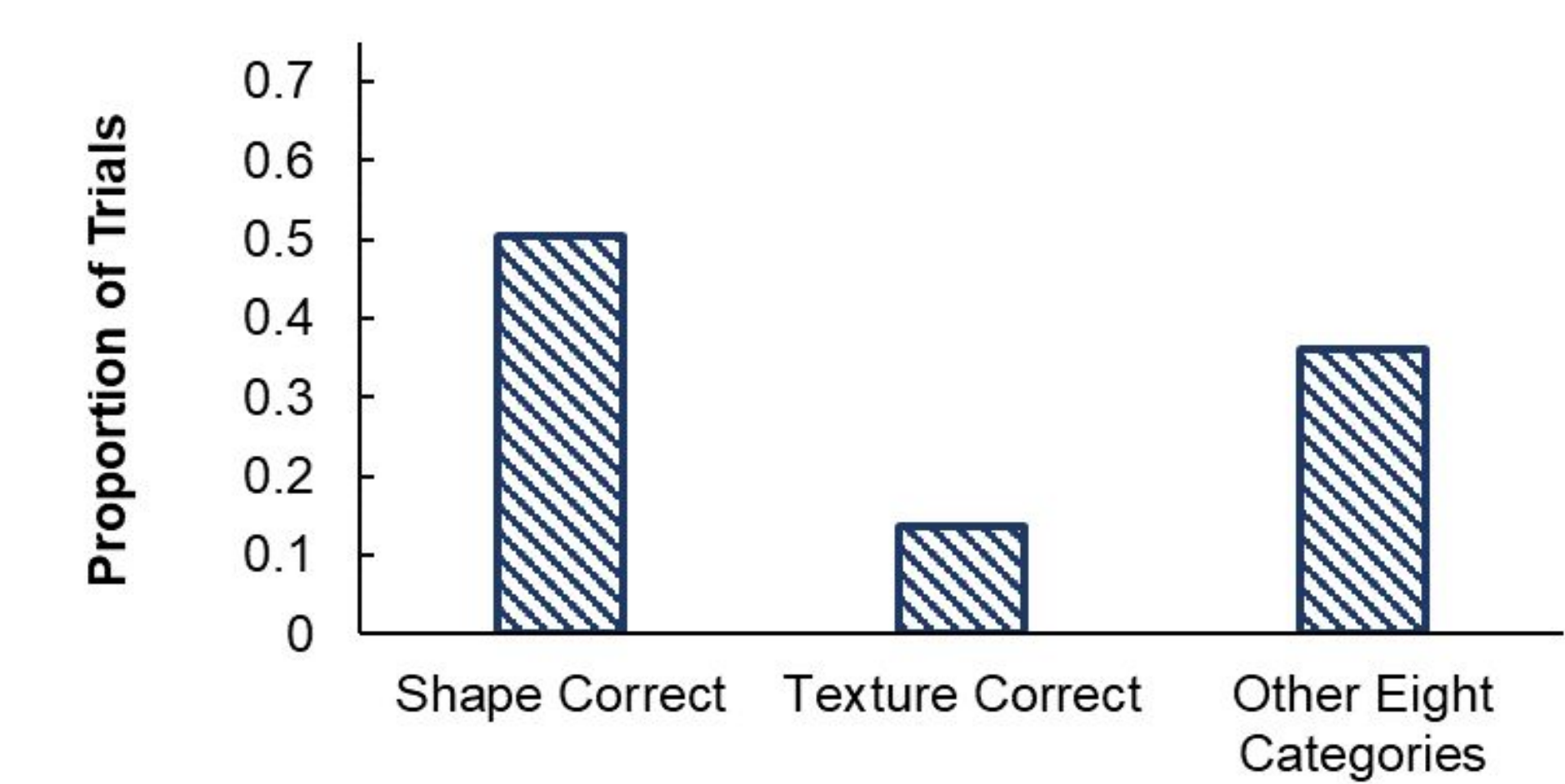


Figure 4. ViT Classification of Image
Network correctly classified by shape in 50% of trials, texture in 15% of trials, and within the other 8 categories the other 40% of trials

Next Steps

- Continue analysis to verify CNN's lack use of texture in image classification
- Investigate why 3-D renderings produce this effect whereas 2-D images do not
- Create models to preserve texture essence to determine if this carries more useful information for classification
- Test networks using other phenomena that humans use to see and interpret their visual environment

Impacts

- Enhance comprehension of CNN perception capabilities
- Improve technology for image classification-related fields
- Optimize technology physicians use to aid in analyzing medical scans
- Advance the algorithms and systems that autonomous vehicles rely on for better decision-making

References and Acknowledgements

Baker, N., Lu, H., Erlikhman, G., & Kellman, P. J. (2018). Deep convolutional networks do not classify based on global object shape. *PLOS Computational Biology*, 14(12). <https://doi.org/10.1371/journal.pcbi.1006613>

Szegedy C, Zaremba W, Sutskever I, Bruna J, Erhan D, Goodfellow I, Fergus R. Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199. 2013 Dec 21.

Tulio Ribeiro, M., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*. <https://doi.org/10.18653/v1/n16-3020>

Thank you to our mentors Dr. Nicholas Baker and Dr. George K. Thiruvathukal for their support of our research and to the Carbon Family for funding our research through the Carbon Fellowship.