Study Portfolio

COMPUTER VISION

Varit Kobutra
Professor Patricia McManus
Houston Community College
3 August 2024
Portfolio GitHub Link: https://github.com/henrykobutra/itai-1378-portfolio

Computer vision is not about teaching computers to see, it's about teaching computers to understand the world around us.

- Jensen Huang, CEO of NVIDIA

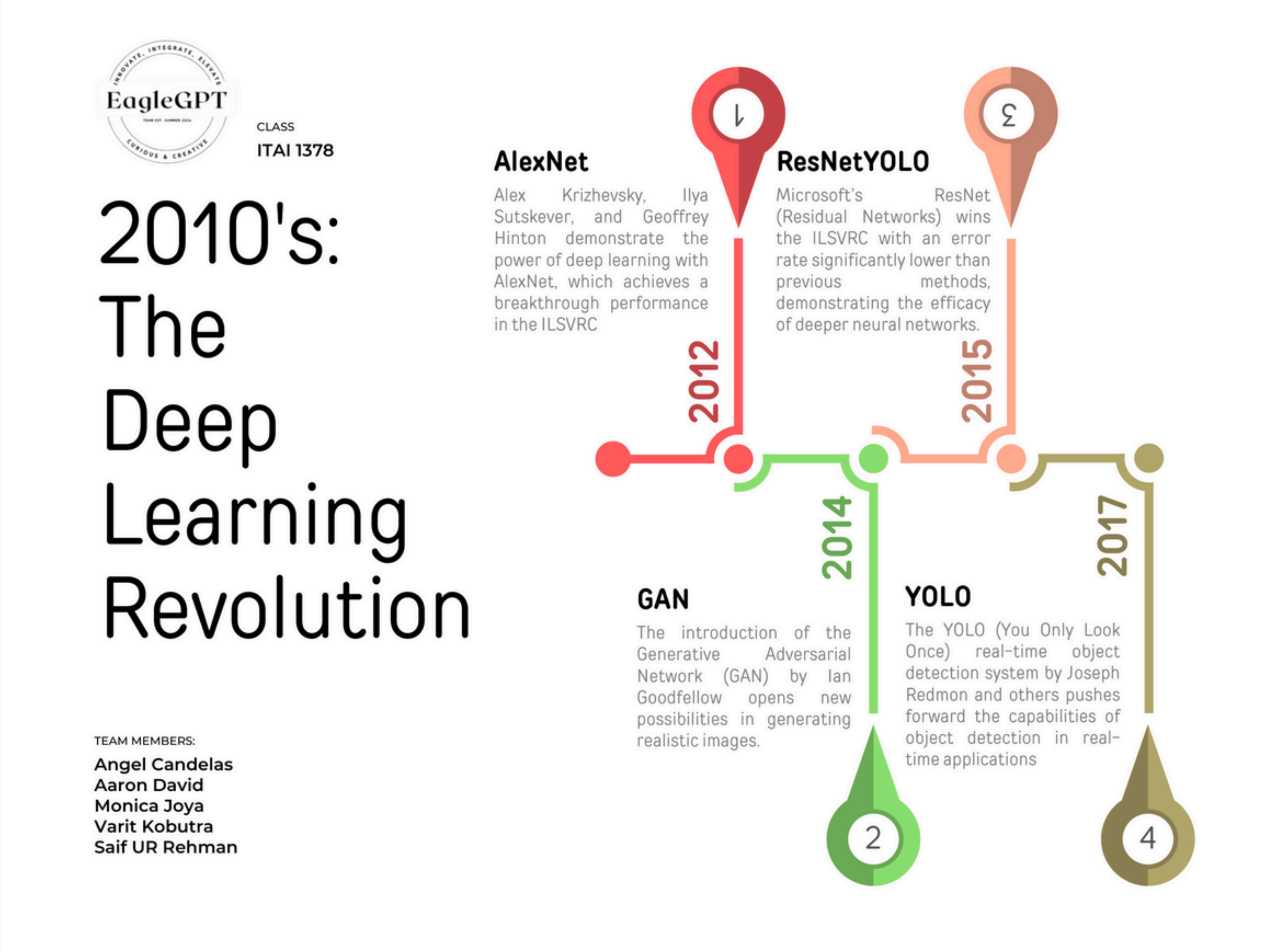
WHAT WE LEARNED

- Historical Timeline of Computer Vision
- Jupyter Notebooks and GitHub
- Image Processing Fundamentals
- Machine Learning in Computer Vision
- Neural Network Building Blocks
- Convolutional Neural Networks
- Advanced CNN Architectures
- Object Detection and Recognition
- The Future of Generative AI and AI Agents

HISTORICAL TIMELINE OF COMPUTER VISION

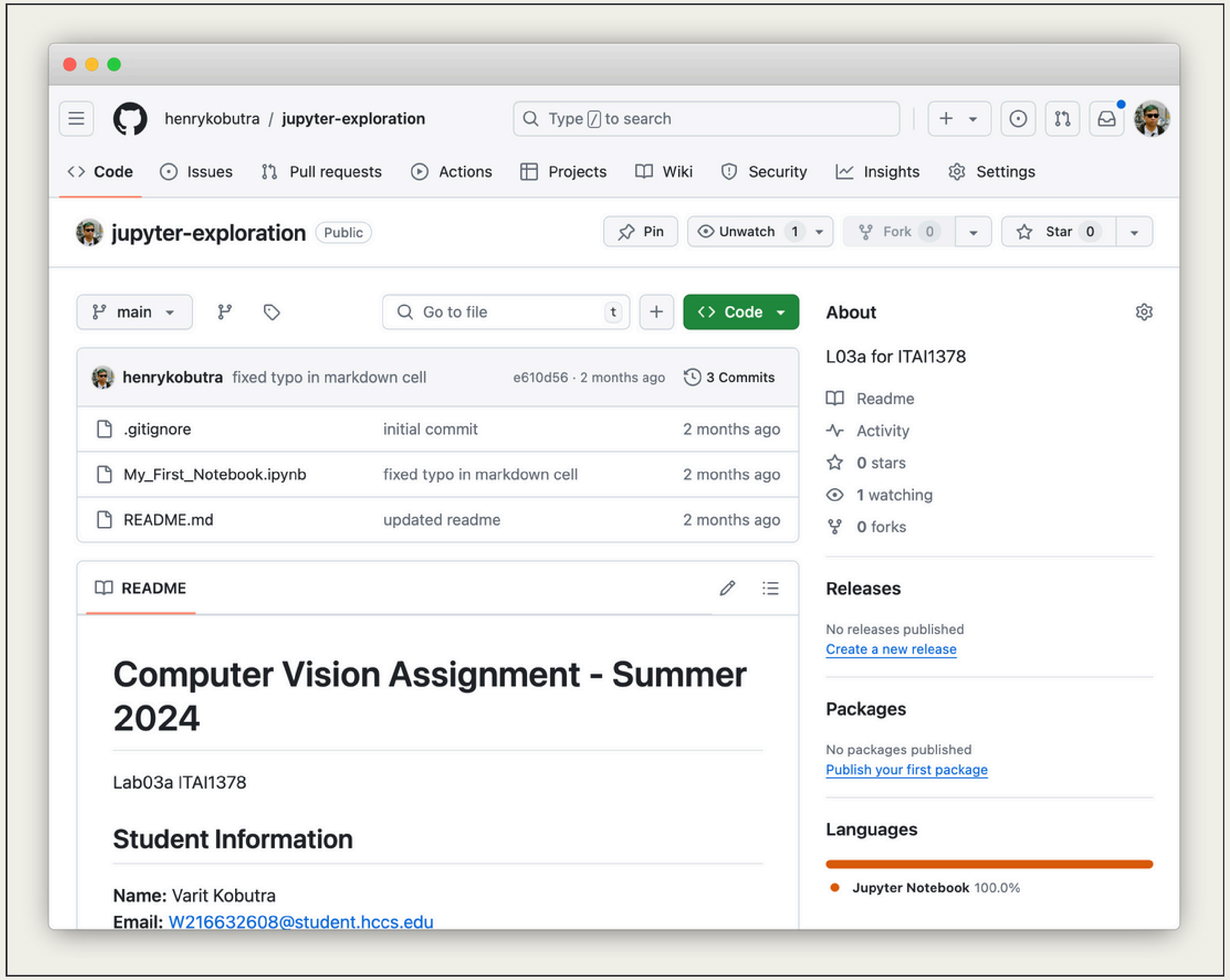
The success of these methods on benchmark competitions has inspired a growing number of researchers and companies to adopt and further develop deep learning methods for computer vision. This has arguably led to a breakthrough in the effectiveness of computer vision in several areas" (Voulodimos et al., 2018, p. 2)

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). *Deep learning for computer vision: A brief review*. Computational Intelligence and Neuroscience, 2018, 1-13. https://doi.org/10.1155/2018/7068349



JUPYTER NOTEBOOKS AND GITHUB

Jupyter Notebooks and GitHub form a powerful tandem in computer vision workflows, with Notebooks providing an interactive environment for code execution, data visualization, and documentation, while GitHub offers version control, collaboration features, and a platform for sharing and discovering cutting-edge computer vision projects and datasets.



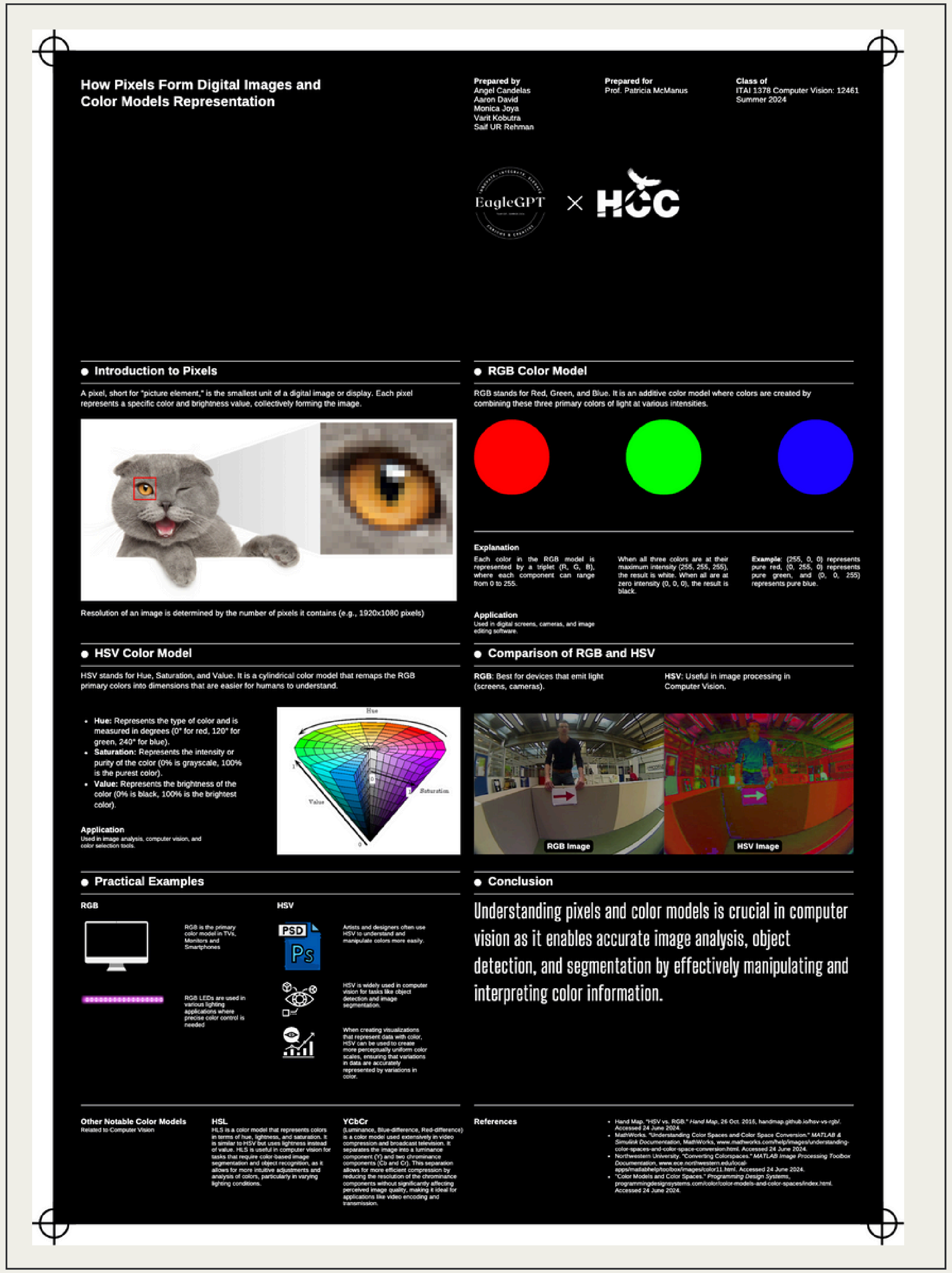
Learn more:

- https://jupyter.org/
- https://github.com

IMAGE PROCESSING FUNDAMENTALS

As Gonzalez and Woods (2018) note, "Image processing is a core technology in computer vision, forming the foundation upon which higher-level vision tasks are built" (p. 1). This underscores the critical role of image processing techniques in enabling sophisticated computer vision applications.

Gonzalez, R. C., & Woods, R. E. (2018). *Digital image processing (4th ed.)*. Pearson.

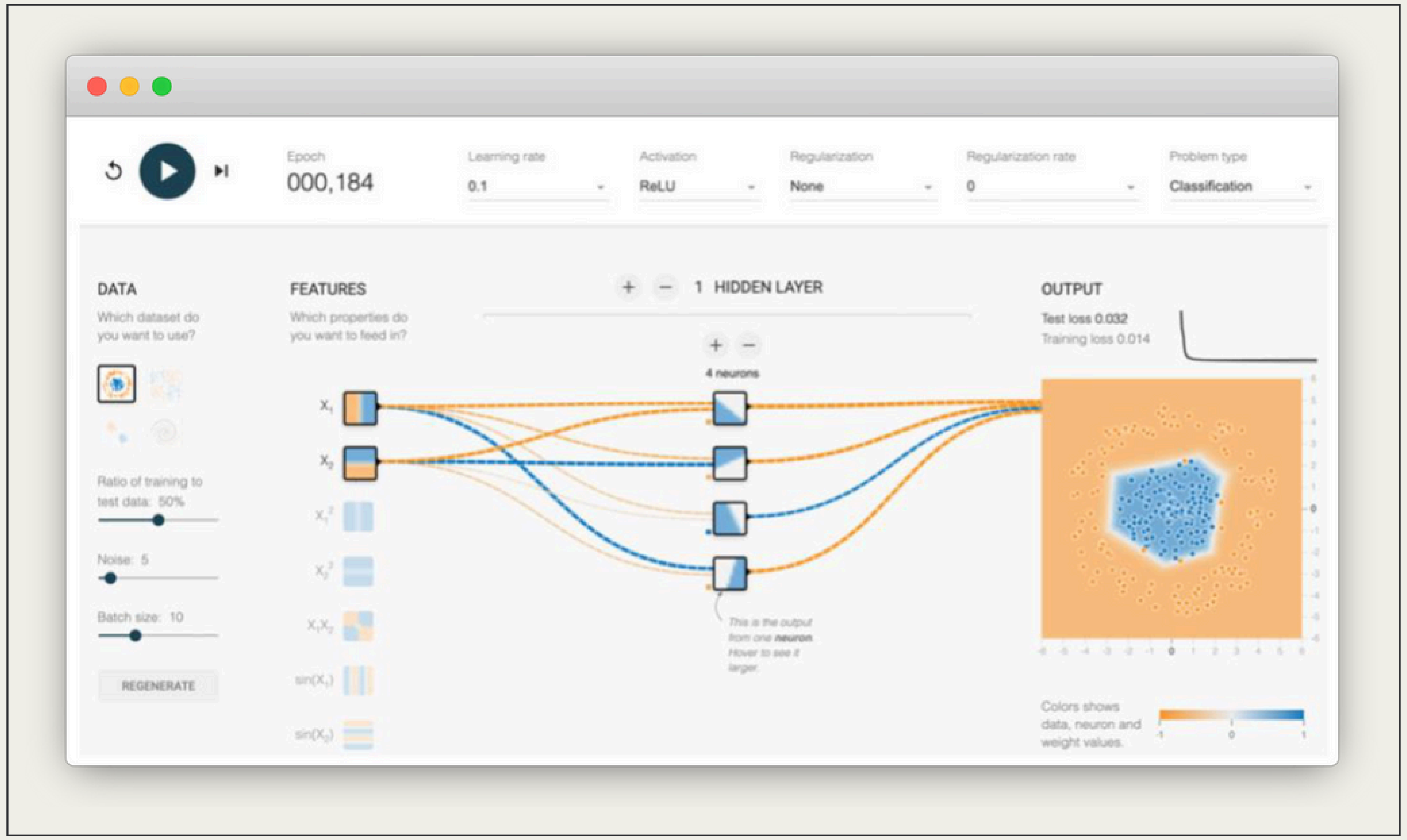


MACHINE LEARNING IN COMPUTER VISION

Machine learning forms the backbone of modern computer vision systems, enabling tasks like image classification, object detection, and scene understanding through algorithms that learn from data. In our course, we explored this synergy hands-on by implementing an SVM classifier for the CIFAR-10 dataset, achieving 31.2% accuracy across 10 diverse object categories and gaining practical insights into challenges like small image sizes and class diversity.

NEURAL NETWORK BUILDING BLOCKS

TensorFlow Playground provided an interactive platform to visualize and experiment with neural network architectures, allowing us to gain intuitive understanding of key concepts like layers, neurons, and activation functions — a approach that aligns with Smilkov et al.'s (2017) finding that "interactive visualizations can help build an intuitive understanding of machine learning concepts" (p. 1).

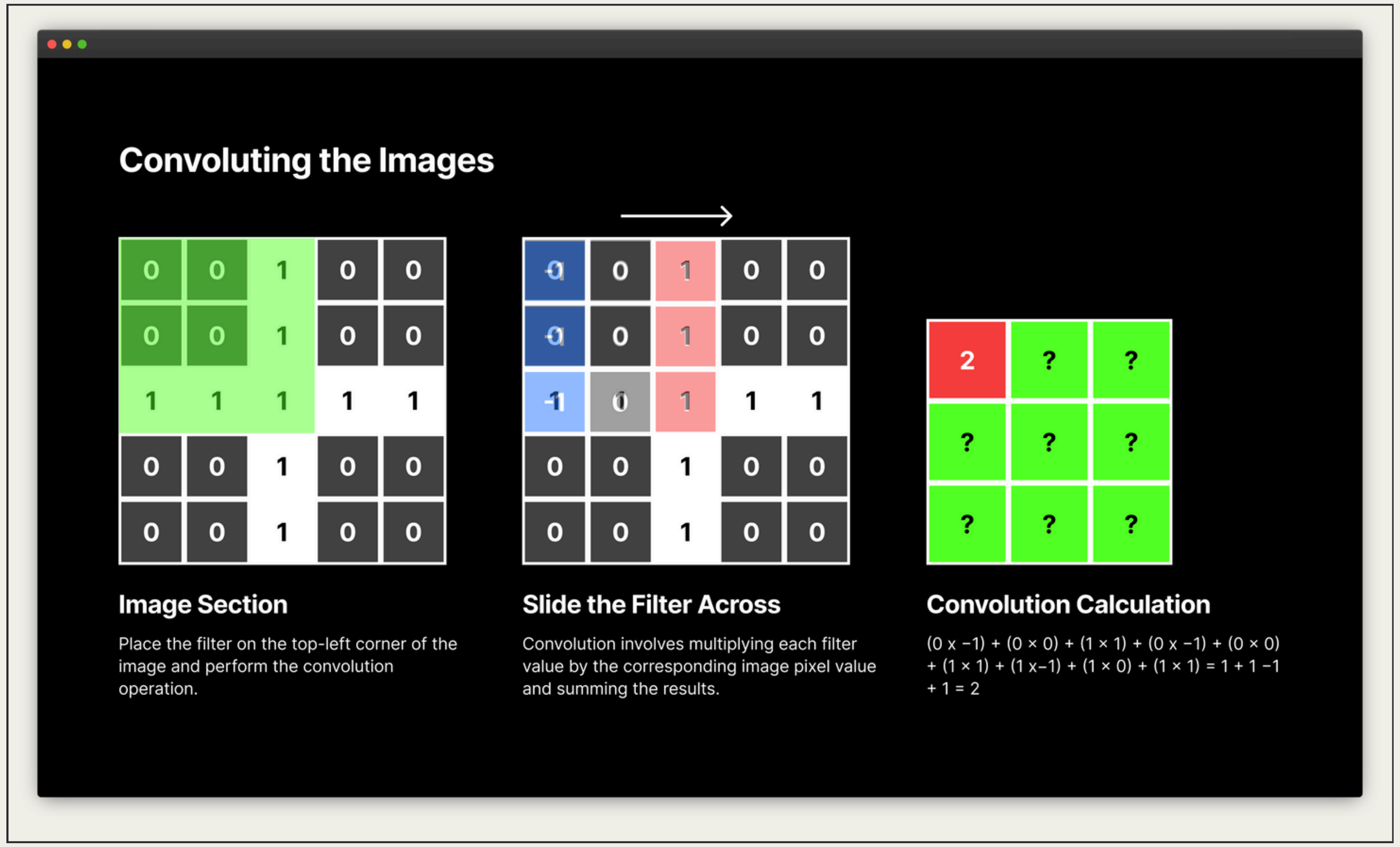


Smilkov, D., Carter, S., Sculley, D., Viégas, F. B., & Wattenberg, M. (2017). *Direct-manipulation visualization of deep networks*. arXiv preprint arXiv:1708.03788.

HCC

CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks revolutionized computer vision by introducing specialized layers that mimic the human visual cortex, enabling efficient processing of grid-like data such as images. As LeCun et al. (2015) note, "CNNs have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent networks have shone light on sequential data such as text and speech" (p. 436).



LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. Nature, 521(7553), 436–444.

ADVANCED CNN ARCHITECTURES

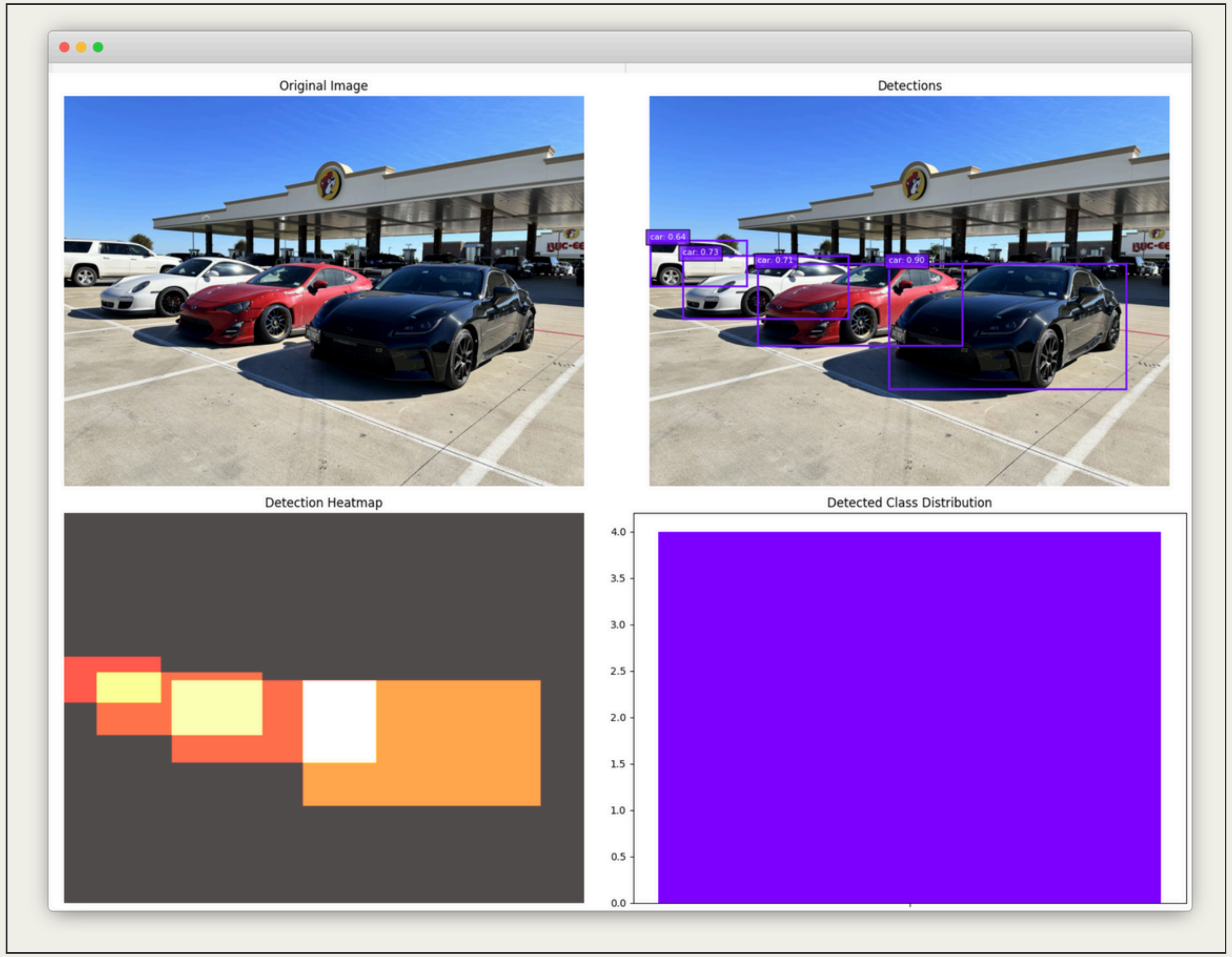
The rapid evolution of object detection architectures, from R-CNN to YOLO, demonstrates the field's progress in balancing accuracy and speed. As Zhao et al. (2019) observe, "Object detection has achieved significant breakthroughs in recent years with the rapid development of deep convolutional neural networks" (p. 1).

Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). *Object detection with deep learning: A review*. IEEE transactions on neural networks and learning systems, 30(11), 3212-3232.

OBJECT DETECTION AND RECOGNITION

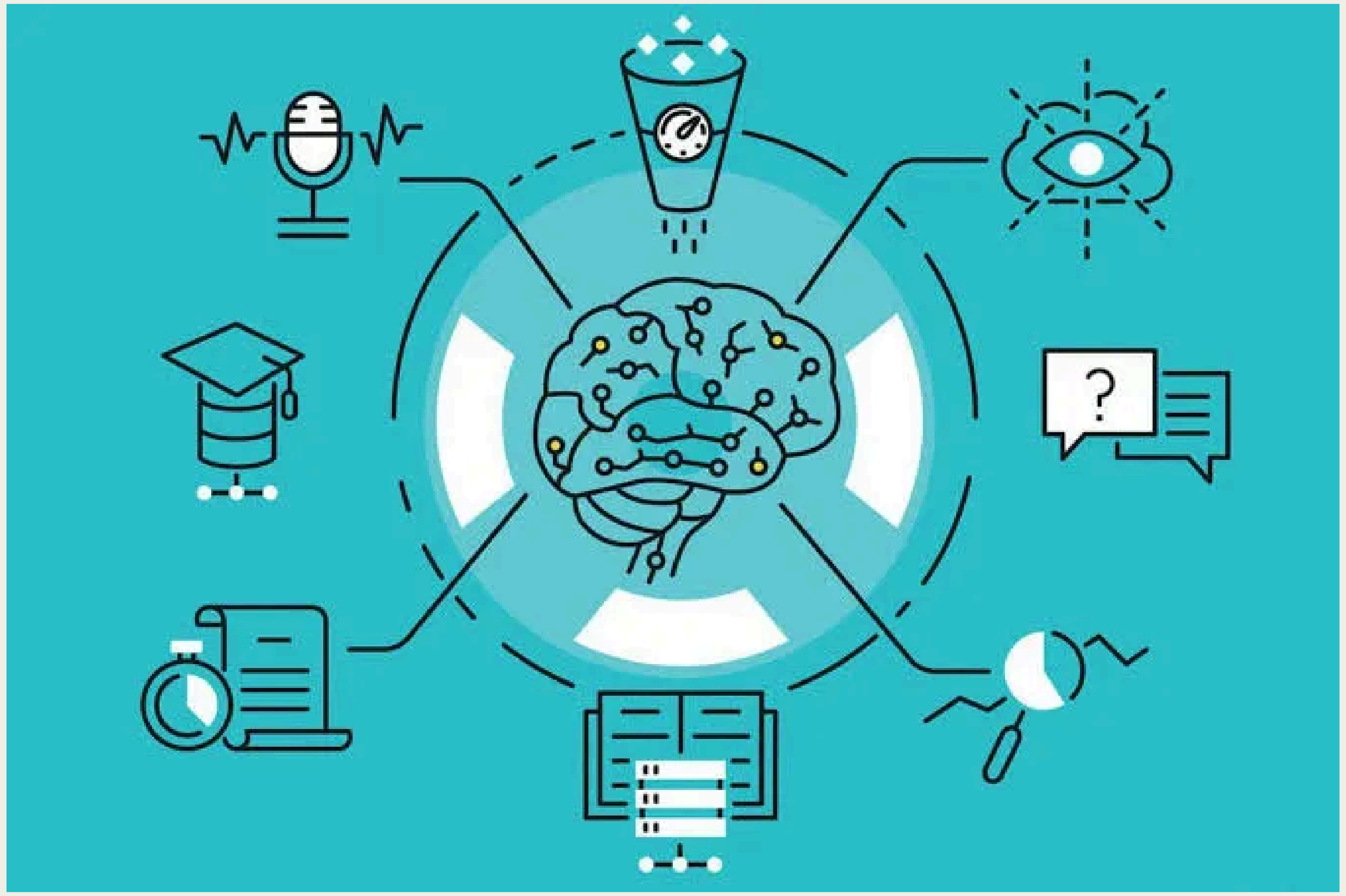
Object detection and recognition extend image classification by not only identifying what objects are in an image, but also where they are located. As Zou et al. (2019) explain, "Object detection can be viewed as a combination of classification and localization tasks, significantly increasing the complexity and utility of computer vision systems" (p. 2).

Zou, Z., Shi, Z., Guo, Y., & Ye, J. (2019). *Object detection in 20 years: A survey*. arXiv preprint arXiv:1905.05055.



THE FUTURE OF GENERATIVE AI AND AI AGENTS

The convergence of Generative AI and AI Agents with computer vision promises to revolutionize how machines perceive and interact with the world. As Ramesh et al. (2022) note, "Large language models have recently been shown to be capable of performing a wide range of tasks, from generating coherent text to manipulating images, by following natural language instructions" (p. 1), suggesting a future where AI systems can not only understand visual data but also create and manipulate it with unprecedented sophistication.



Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). *Hierarchical text-conditional image generation with clip latents*. arXiv preprint arXiv:2204.06125.

Thank you!

VISION ACHIEVED

Varit Kobutra
Professor Patricia McManus
Houston Community College
3 August 2024
Portfolio GitHub Link: https://github.com/henrykobutra/itai-1378-portfolio