

Krishna Srikar Durbha

Graduate Research Assistant at The University of Texas at Austin

[in](#) krishna-srikar-durbha [📧](#) krishnasrikard [📞](#) +1 512-593-3339 [@](#) krishna.durbha@utexas.edu [</>](#) krishnasrikard.github.io

Education

Ph.D in Electrical and Computer Engineering at The University of Texas at Austin

Aug 2022 - Present

- **CGPA:** 3.8 / 4
- **Courses:** Generative Models in Machine Learning, Advanced Topics in Computer Vision, Digital Video, etc.

B.Tech in Electrical Engineering at Indian Institute of Technology Hyderabad

Jul 2018 - May 2022

- **CGPA:** 9.11 / 10.
- **Courses:** Image and Video Processing, Representation Learning, Probabilistic Graphical Models, Statistical Inference, etc.

Work Experience

The University of Texas at Austin | Graduate Research Assistant

Aug 2022 - Present

- Working under the guidance of **Professor Alan Bovik**.
- Collaborating with Meta Video Infrastructure Team on Video Engineering.

Samsung Research America | MPI Research Intern

Summer 2025

- Worked on high-resolution text-to-image generation using diffusion models.
- Curating the work for CVPR 2026 submission.

Skydio | Wireless Systems Intern

Summer 2024

- Worked on improving the quality of video streamed from the drone to the controller by employing adaptive video streaming and video quality assessment techniques.
- Curated a comprehensive dataset of videos by conducting experiments to demonstrate the necessity of frame-rate changes under different visual environments, user requirements, and severe network conditions.
- Experimented with GoP size and resolutions for live video streaming under bitrate settings, frame-rate settings, different visual environments, and network conditions.
- Proposed both simple and complex frame-rate switching mechanisms for different perceptual conditions, flying environments, and network conditions, etc using various visual features, feedback from drones, etc.

Oracle | Application Development Intern

Summer 2021

- Developed an AI-powered conversational interface with Oracle Digital Assistant, leveraging Natural Language Understanding to automate user interactions and tasks within the Journeys application.
- Curated a comprehensive training dataset to optimize the chatbot's intent classification and entity recognition capabilities, underpinning its ability to accurately interpret user queries and perform actions within the Journeys application.
- Architected the backend infrastructure, integrating the chatbot with the Journeys ecosystem using RestAPI, OBotML, Client Web SDK, and Oracle HCM Cloud, facilitating seamless data flow and functionality. Ensured code quality and reliability through comprehensive testing practices, utilizing Mocha Chai and Jasmine frameworks.

Publications

K. S. Durbha, A. K. Venkataramanan, R. Sureddi, and A. C. Bovik. "Perceptual Classifiers: Detecting Generative Images using Perceptual Features." arXiv preprint arXiv:2507.17240 (2025).

K. S. Durbha, H. Tmar, C. Stejerean, I. Katsavounidis and A. C. Bovik, "Bitrate Ladder Construction Using Visual Information Fidelity," 2024 Picture Coding Symposium (PCS), Taichung, Taiwan, 2024, pp. 1-4, doi: 10.1109/PCS60826.2024.10566405.

K. S. Durbha and S. Amuru, "AutoML Models for Wireless Signals Classification and their effectiveness against Adversarial Attacks," 2022 14th International Conference on COMMunication Systems & NETWORKS (COMSNETS), 2022, pp. 265-269, doi: 10.1109/COMSNETS53615.2022.9668448.

Preprints

K. S. Durbha and A. C. Bovik, "Constructing Per-Shot Bitrate Ladders using Visual Information Fidelity," arXiv preprint arXiv:2408.01932, 2024. (Under Review in TIP)

K. S. Durbha, H. Tmar, P. H. Wu, I. Katsavounidis and A. C. Bovik, "Leveraging Compression to Construct Transferable Bitrate Ladders" (Finalizing for Submission in TIP)

Projects

Impact of Backbone Models and Dataset on Cartoonization Performance

[Code](#)

- Explored the impact of backbone architecture namely VGG19 and ViT-B/16 in the cartoonization process particularly in structural and content loss functions.
- Creating and Understanding a cartoonization process to a specific style to cartoon rather than a collection of styles.
- Observed a high color pallet and coarse surface with a transformer backbone (ViT) rather than a CNN backbone (VGG19). We also observed an improvement in the cartoonization process as the size of the dataset increases particularly the ViT backbone.

An Efficient Approach to Super-Resolution with Fine-Tuning Diffusion Models

[Code](#)

- Explored the potential of pre-trained diffusion model SR3, specifically fine-tuning and zero-shot approaches for the task of image super-resolution.
- Demonstrated the generalization ability of fine-tuning process of SR3. The fine-tuning process is evaluated with limited time steps, iterations and data samples.
- Evaluated the zero-shot approach of using range-null space decomposition for super-resolution using unconditional DDPM with using a conditional DDPM SR3 trained from scratch.

Optical Flow Less Video Frame Interpolation

[Code](#)

- Designed a lightweight video restoration transformer to capture long-range interactions, for fast inference and smaller training requirements for video frame interpolation. The model employs self-attention for feature extraction and mutual-attention as a surrogate to motion estimation to capture temporal information and feature alignment.
- Created a training procedure to predict intermediate frames of the video which are continuous with subsequent frames by only looking at the previous frames essentially following causality.
- Achieved comparable results with other SoTA video interpolation models.

Similarities between local-patch quality maps of NR IQA algorithms and saliency maps of computer vision classification models

[Code](#)

- Achieved an understanding of similarities between perception of images by humans and classification models. NR-IQA models trained on human judgments/quality ratings are used to replicate the perception of humans. Local-patch quality maps provide the key areas focused on while rating an image.
- Using PaQ-2-PiQ to create local-patch quality maps for images. ResNet18 is trained on images rated as good-quality images by PaQ-2-PiQ and saliency maps are generated using Grad-CAM.
- Compared the variation in local-patch quality maps and saliency maps due to natural scene distortions like brightness, contrast, jpeg-compression, motion-blur, zoom-blur, etc.

Skills

Programming Languages	C++, Python, MATLAB, Java, Javascript
Libraries	PyTorch, TensorFlow, AutoKeras, Scikit-Learn
Web Development	HTML, CSS, Node.js, SQL
Other Tools	LaTeX, Microsoft Office
Operating Systems	Linux, Windows