

Shreshth Saini

📞 +1 737-781-5912 • ✉️ saini.2@utexas.edu • 🌐 shreshthsaini.github.io
🐙 GitHub/shreshthsaini • in LinkedIn/shreshthsaini
🔍 GoogleScholar/ssaini • 🐦 twitter/ssaini

Education

MS+Phd in Electrical and Computer Engineering (DICE Track)

The University of Texas at Austin, USA

2022 – 2026

Supervisor: Prof. Alan C Bovik

Bachelor of Technology in Electrical Engineering

Indian Institute of Technology (IIT) Jodhpur, India

2016 – 2020

Supervisor: Dr. Anil K Tiwari

Research Interests

Generative AI, Diffusion Models, Video Diffusion Models, Image/Video Processing, Inverse Problems, Computer Vision (CV), Deep Learning (DL), Machine Learning (ML), High Dynamic Range (HDR) Videos, Image/Video Quality Assessment

Employment and Research Appointments

Applied Scientist Intern

Amazon - Preception Team

Seattle, Washington

June 2024 – August 2024

- Worked with the Preception team on large-scale synthetic data generation.
- Developed novel edit-bench and T2I-based diffusion model for consistent image/video editing and generation.
- Aiming to conduct Image+Video editing challenge and workshop.

Research Intern

Alibaba Group

Sunnyvale, California

Jan. 2024 – May 2024

- Developed generalizable and robust Vision Model-based Video Quality Assessment (VQA) methods.
- Using Diffusion Model priors as perceptual consistency for IQA (*Paper: under review*).

Co-Founder

Short-X

Austin, Texas

Jan. 2023 – Jan. 2024

- Short-X aims to automate the arduous task of making short-form contents from traditional long-form content.
- Building core AI models and pipelines for Short-X. Working on transcription, extracting semantically meaningful and unique highlights, removing pauses, identifying speaker and smart vertical cropping.

Graduate Research Assistant

YouTube/Laboratory for Image and Video Engineering, UT Austin

Austin, Texas

Aug. 2022 – Present

- Developing scalable vision models for HDR videos for tasks like ITM/TM, gamut expansion & quality assessment.
- Created the largest HDR-SDR dataset for short-form videos (publicly available); Developing video quality assessment methods for HDR videos, which uses Non-Linear expansion of extremes of sub-level luminance.

Machine Learning Engineer

BioMind (Products)

Singapore, Singapore

Feb. 2022 – June 2022

- Developed SOTA multimodal DL models for segmentation and classification of 25+ tumor/non-tumor classes.
- Exploited TFRecords for memory-intense 4D datasets and proposed multi-task model for tumor predictions.

Research Engineer – AI

Arkray, Inc.

Kyoto, Japan (Remote)

Aug. 2020 – Dec. 2021

- Proposed semi-supervised DL models to learn from a large chunk of the private unlabelled and noisy 2D datasets.
- Deployed models for products: UrineSediment Analyzer, and automated BodyFluid Analyzer. (Aution EYE)

Research Assistant

National University of Singapore, Singapore

Singapore, Singapore

May 2019 – July 2019

Supervisor: Dr. Mengling 'Mornin' Feng

- Developed novel deep learning architecture for large-scale public health datasets.
- Published SOTA results with low cost for skin lesion analysis.

Undergraduate Researcher

Image Processing and Computer Vision Lab, IIT Jodhpur

Jodhpur, India

Aug. 2018 – Aug. 2020

Supervisor: Dr. Anil Kumar Tiwari

- Worked on developing ML methods aimed for AI-based diagnosis and treatment support. 1/3
- Developed DL models for retinal vessel & skin lesion segmentation, and diagnosis of left-atrium in 3D GE-MRIs.

Research Intern

The Multimedia Analytics, Networks and Systems Lab, IIT Mandi

Mandi, India

May 2018 – July 2018

Supervisor: Dr. Aditya Nigam

- Developed novel CNN model for iris segmentation which uses cascaded hourglass modules at the bottleneck of encoder-decoder design.

Publications

Conferences:

1. **Prime-EditBench: A Real World Benchmark for Image and Video Editing Task using Diffusion Models (Amazon)**
S Saini, P Korus, S Jin, AC Bovik.
2. **Perceptual Consistency in Diffusion Model: Exploiting Diffusion priors for Image Quality Assessment (Alibaba)**
S Saini, Y Ye, AC Bovik. *Under Review - ICLR*
3. **ITM-DM: Using Diffusion Models for UGC-Video Inverse Tone-Mapping (YouTube)**
S Saini, N Birkbeck, B Adsumilli, AC Bovik. *Under Review - TIP*
4. **Subjective and Objective Analysis of Large-Scale UGC-HDR Videos (YouTube)**
S Saini, N Birkbeck, B Adsumilli, AC Bovik. *Under Review - TIP*
5. **Contrastive HDR-VQA: Deep Contrastive Representation Learning for High Dynamic Range Video Quality Assessment**
S Saini, A Saha, AC Bovik. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024, Waikoloa, Hawaii*
6. **M2SLAe-Net: Multi-Scale Multi-Level Attention Embedded Network for Retinal Vessel Segmentation**
S Saini, G Agrawal. *The IEEE International Symposium on Biomedical Imaging (IEEE ISBI), 2021 Acropolis-France*
7. **(M)SLAe-Net: Multi-Scale Multi-Level Attention Embedded Network for Retinal Vessel Segmentation [Paper]**
S. Saini, G. Agrawal. *9th IEEE International Conference On Healthcare Informatics (IEEE ICHI), 2021 (full Oral Presentation) Victoria, British Columbia, Canada*
8. **B-SegNet Branched SegMentor Network for Skin Lesion Segmentation [Paper]**
S Saini, YS Jeon, M Feng. *Association for Computing Machinery Conference on Health, Inference, and Learning (ACM CHIL), 2021 (full Oral Presentation)*
9. **Detector-SegMentor Network for Skin Lesion Localization and Segmentation [Paper]**
S Saini, D Gupta, AK Tiwari. *National Conference on Computer Vision, Pattern Recognition, Image Processing, & Graphics (NCVPRIPG), 2019 (full Oral Presentation), twin of ICVGIP*

Journals:

1. **ITM-DM: Video Diffusion Model for High Dynamic Range Video Inverse Tone Mapping**
S Saini, N Birkbeck, B Adsumilli, AC Bovik. *Under Review*
2. **PixISegNet: pixel-level iris segmentation network using convolutional encoder-decoder with stacked hour-glass bottleneck [Paper]**
RR Jha¹, G Jaswal¹, S Saini², D Gupta², A Nigam.
The Institution of Engineering and Technology (IET Biometrics), 2019

Book Chapters:

1. **Iris Segmentation in the Wild using Encoder-Decoder based Deep Learning Techniques [Paper]**
S Saini, D Gupta, RR Jha, G Jaswal, A Nigam.
AI and Deep Learning in Biometric Security: Trends, Potential and Challenge
CRC Press (Taylor & Francis Group), 2020

Selected Achievements

- Awarded Cockrell Engineering (UT Austin) Graduate Fellowship for exceptional academic record, **2022-2027**
- Received **Merit-Cum-Means Scholarship** from IIT Jodhpur to cover undergraduate expenses, **2017-2019**
- Won medical imaging track at NUS-MIT datathon, led a team of 10 data scientists and clinicians, **2019**
- Established undergraduate research group (LAMBDA), group publishes in international conferences, **2018**
- Letter of Appreciation from District Collector Sirohi (Rajasthan, India) for Academic Excellence, **2013**

Selected Coursework

- **CS/EE:** Computational and Variational Methods for Inverse Problems, Applied Machine Learning, Machine Learning, Artificial Intelligence, Advanced Computer Vision, Digital Image Processing, Computational Imaging, Digital Video, Vision Systems, Information Theory and Coding, Algorithms.
- **Mathematics:** Probability, Statistics, and Random Processes, Linear Algebra and Calculus, Complex Analysis and Differential Equations, Adv. Probability and Stochastic Processes, Statistical Methods I/II.
- **Others:** Principles of Management, Professional Ethics, IP Management and Exploitation, Technology Manage-

ment, Basic of Leadership, Longhorn Startup.

Technical Skills

- *Languages:* Python, MATLAB, C++.
- *Frameworks:* Tensorflow, Pytorch, Scikit-Learn, OpenCV, FFMPEG, Docker.
- *Misc.:* Git, Bash, LATEX, Linux, SQL.

Leadership

- Student Leader, LAMBDA-IIT Jodhpur, India. *Aug. 2018 – Aug. 2020*
- Volunteer, NUS-MIT Datathon, Singapore. *May 2019 – July 2019*
- Overall Student Head, Entrepreneurship Cell-IIT Jodhpur, India. *May 2018 – May 2019*
- Assistant Head, Counselling Services-IIT Jodhpur, India. *May 2018 – May 2019*
- Volunteer, International Workshop on Deep Learning(IWDL), India. *May 2018 – July 2018*
- Vice Captain, Astronomy Club-IIT Jodhpur, India. *May 2017 – May 2018*

References

- Up to 4 references available on request