

SANDIPAN BANERJEE

CONTACT INFORMATION

EMAIL: sandipan9008@gmail.com | [Personal Webpage](#) | LOCATION: New York, NY.

RESEARCH INTERESTS

Computer Vision - Classification, synthesis, self/semi-supervised learning, domain invariance.

Deep Learning - Video Diffusion, VLMs, Multi-modal CNNs, VAE, (VQ)GAN, Transformers.

Biometrics - Face (identity/authentication/recognition) and iris (segmentation).

WORK EXPERIENCE

SEPT 2025-PRESENT	Applied Scientist II at Amazon <ul style="list-style-type: none">- Design video diffusion architecture for one-shot video editing for ads localization.- Build multi-LLM based synthetic video quality estimation pipeline for ads localization.
NOV 2023-SEPT 2025	Senior Machine Learning Research Engineer at Pipio <ul style="list-style-type: none">- Pipeline for diffusion based (audio, image, text, ...) to avatar video gen & editing.- Prior network architecture for lipsync improvement in talking head renderer.- Generate natural looking motion in video sequences of virtual avatars.
JUN 2022-SEPT 2023	Staff Research Scientist at Samsung Research America (NEON) <ul style="list-style-type: none">- Framework for fast image to image rendering built around lightweight GANs.- Multi-modal synthesis models for realistic behavior generation in NEONs.
JUL 2019-JUN 2022	Sr. Computer Vision Scientist at Affectiva (acquired by Smart Eye) <ul style="list-style-type: none">- Universal encoding space that can classify facial attributes, identity, lighting all at once, invariant to camera type and location inside a car.- GAN models to manipulate facial attributes, using losses driven by human perception.- Weakly supervised gaze prediction models, robust to changes in camera characteristics.- Hallucinating synthetic samples in facial attribute feature space to detect temporal events like drowsiness from driver videos.
MAY 2014-MAY 2019	Graduate Research Assistant at University of Notre Dame <ul style="list-style-type: none">- Generation of natural looking synthetic faces (2D & 3D) to augment training datasets.- Multi-scale cascaded network of GANs for hallucinating context and background pixels directly from a single face mask.- Created the Notre Dame Synthetic Face Dataset, containing 2M images of 12K synthetic identities, that can be used without copyright or privacy concerns.- PAD Project: detecting counterfeit pharmaceutical drugs.
SUMMER 2016	Research Intern at Xerox Palo Alto Research Center (PARC) <ul style="list-style-type: none">- Keypoint detection in the forehead and cheek regions using texture information.- Artificial aging of face images (CNN, GAN, blending) for predicting skin conditions.
AUG 2013-MAY 2014	Graduate Teaching Assistant at University of Notre Dame <ul style="list-style-type: none">- <i>Computer Networks (CSE 30264)</i>, and <i>Ethical and Professional Issues (CSE 40175)</i> courses.
JUL 2012-APR 2013	Associate Engineer at Unisys Global Services India <ul style="list-style-type: none">- Plug-in development for Unisys' proprietary Clearpath mainframe servers.

EDUCATION

AUG 2013-AUG 2019	PhD in COMPUTER SCIENCE, UNIVERSITY OF NOTRE DAME, USA ADVISERS: Patrick Flynn & Kevin Bowyer . DISSERTATION: <i>Exploring the Effects of Frontalization & Data Synthesis on Face Recognition</i> . [PDF]
AUG 2008-MAY 2012	B.Tech in COMPUTER SCIENCE & ENGINEERING, NIT DURGAPUR, India

PUBLICATIONS (* DENOTES EQUAL CONTRIBUTION)

- J. Flynn*, W. Paier*, D. Dinev, S. Nguyen, H. Poghosyan, M. Toribio, **S. Banerjee**, and G. Gafni, *EditYourself: Audio-Driven Generation and Manipulation of Talking Head Videos with Diffusion Transformers*, pre-print. [\[paper\]](#)
- S. Banerjee**, A. Joshi, and J. Turcot, *The Universal Face Encoder: Learning Disentangled Representations Across Different Attributes*, in IEEE Conference on Computer Vision & Pattern Recognition (CVPR) Workshops, 2023. [\[paper\]](#)
- S. Banerjee**, W. Scheirer, K. Bowyer, and P. Flynn, *Analyzing the Impact of Shape & Context on the Face Recognition Performance of Deep Networks*, in IEEE International Conference on Automatic Face and Gesture Recognition (FG), 2023. [\[paper\]](#)
- S. Banerjee**, A. Joshi, J. Turcot, B. Reimer, and T. Mishra, *Driver Glance Classification In-the-wild: Towards Generalization Across Domains and Subjects*, in IEEE International Conference on Automatic Face and Gesture Recognition (FG), 2021. [\[paper\]](#)
- S. Banerjee**, A. Joshi, P. Mahajan, S. Bhattacharya, S. Kyal, and T. Mishra, *LEGAN: Disentangled Manipulation of Directional Lighting and Facial Expressions by Leveraging Human Perceptual Judgements*, in IEEE Conference on Computer Vision & Pattern Recognition (CVPR) Workshops, 2021. [\[paper\]](#) (Best Paper Runner Up)
- S. Banerjee**, A. Joshi, A. Ghoneim, S. Kyal, and T. Mishra, *Synthesize & Learn: Jointly Optimizing Generative and Classifier Networks for Improved Drowsiness Detection*, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021. [\[paper\]](#)
- D. Saavedra, **S. Banerjee**, and D. Mery, *Detection of Threat Objects in Baggage Inspection with X-ray Images using Deep Learning*, in Neural Computing and Applications (NCAA), 2020. [\[paper\]](#)
- A. Joshi, S. Kyal, **S. Banerjee**, and T. Mishra, *In-The-Wild Drowsiness Detection from Facial Expressions*, in IEEE Intelligent Vehicles Symposium (IV) Workshops, 2020. [\[paper\]](#)
- S. Banerjee**, W. Scheirer, K. Bowyer, and P. Flynn, *On Hallucinating Context and Background Pixels from a Face Mask using Multi-scale GANs*, in IEEE Winter Conference on Applications of Computer Vision (WACV), 2020. [\[paper\]](#)
- S. Banerjee**, W. Scheirer, K. Bowyer, and P. Flynn, *Fast Face Image Synthesis with Minimal Training*, in IEEE Winter Conference on Applications of Computer Vision (WACV), 2019. [\[paper\]](#)
- S. Banerjee***, J. Brogan*, J. Krizaj, A. Bharati, B. RichardWebster, V. Struc, P. Flynn, and W. Scheirer, *To Frontalize or Not To Frontalize: Do We Really Need Elaborate Pre-processing To Improve Face Recognition?*, in IEEE Winter Conference on Applications of Computer Vision (WACV), 2018. [\[paper\]](#)
- D. Mery, and **S. Banerjee**, *Recognition of Faces and Facial Attributes using Accumulative Local Sparse Representations*, in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2018. [\[paper\]](#)
- S. Banerjee***, J. Bernhard*, W. Scheirer, K. Bowyer, and P. Flynn, *SREFI: Synthesis of Realistic Example Face Images*, in IAPR/IEEE International Joint Conference on Biometrics (IJCB), 2017. [\[paper\]](#)
- D. Mery, E. Svec, M. Arias, V. Rizzo, J. Saavedra, and **S. Banerjee**, *Modern Computer Vision Techniques for X-ray Testing in Baggage Inspection*, in IEEE Trans. on Systems, Man, and Cybernetics: Systems (SMC), 47 (4), pp. 682 - 692, 2017. [\[paper\]](#)
- W. Scheirer, et al., *Report on the BTAS 2016 Video Person Recognition Evaluation*, in IEEE International Conference on Biometrics Theory, Applications and Systems (BTAS), 2016. [\[paper\]](#)
- S. Banerjee**, J. Sweet, C. Sweet, and M. Lieberman, *Visual Recognition of Paper Analytical Device Images for Detection of Falsified Pharmaceuticals*, in IEEE Winter Conference on Applications of Computer Vision (WACV), 2016. [\[paper\]](#)

S. Banerjee, and D. Mery, *Iris Segmentation using Geodesic Active Contours and GrabCut*, PSIVT Workshop on 2D & 3D Geometric Properties from Incomplete Data, 2015. [\[paper\]](#)

PATENTS

A. Liot, A. Unnikrishnan, S. Sadi, S. Banerjee, V. Gokul, J. Palan, H-J. Kang, and O. Texler, *Cache-based Content Distribution Network*, U.S. Patent 12,477,159 (granted in 2025).

S. Banerjee, R. el Kaliouby, A. Joshi, S. Kyal, and T. Mishra, *Synthetic Data for Neural Network Training Using Vectors*, U.S. Patent 11,769,056 (granted in 2023).

R. Lokesh, S. Banerjee, H-J. Kang, O. Texler, and S. Sadi, *Lightweight Rendering System With On-device Resolution Improvement*, U.S. Provisional Patent Application Serial No. 63/456,337.

A. Joshi, S. Banerjee, and J. Turcot, *Neural Network Multi-Attribute Facial Encoder and Decoder*, U.S. Provisional Patent Application Serial No. 18/212,742.

T. Mishra, S. Banerjee, and A. Joshi, *Neural Network Synthesis Architecture using Encoder-Decoder Models*, U.S. Provisional Patent Application Serial No. 63/071,401.

SKILL SET

PROGRAMMING LANGUAGES: Python; C++

APPLICATIONS: Pytorch; Tensorflow; Keras; OpenCV; Hugging Face; scikit-learn; NumPy

PLATFORMS: EC2; SageMaker; Paperspace; FoundryML; GCP; Git; Docker; Jupyter; WandB

OTHER RELEVANT INFORMATION

AREA CHAIR & PROGRAM COMMITTEE MEMBER:

- IEEE Senior Member (elevated in 2024).
- IEEE Winter Conference on Applications of Computer Vision ([WACV](#)) 2023, 2024, 2025, 2026.
- The Annual AAAI Conference on Artificial Intelligence ([AAAI](#)), 2025.
- Human Behavior Understanding ([HBU](#)) Workshop at IEEE WACV 2021.
- Synthetic Data for Face and Gesture Analysis ([SD-FGA](#)) Workshop at IEEE FG 2024.

JOURNAL REVIEWER:

- IEEE Trans. on Pattern Analysis & Machine Intelligence ([TPAMI](#))
- IEEE Trans. on Image Processing ([TIP](#))
- Pattern Recognition ([PR](#))
- IEEE Trans. on Information Forensics & Security ([TIFS](#))
- IEEE Trans. on Biometrics, Behavior, and Identity Science ([TBIOM](#))
- IEEE Journal of Selected Topics in Signal Processing ([JSTSP](#))
- IEEE Transactions on Human-Machine Systems ([SMC](#))
- IEEE [Sensors](#)
- The Visual Computer Journal ([TVC](#)).

CONFERENCE REVIEWER:

- IEEE Conf. on Computer Vision & Pattern Recognition ([CVPR](#)) 2020-26 ([Outstanding Reviewer Award](#), '21)
- IEEE International Conference on Computer Vision ([ICCV](#)) 2021/23
- European Conference on Computer Vision ([ECCV](#)) 2022/24/26
- IEEE Winter Conference on Applications of Computer Vision ([WACV](#)) 2020-22
- International Conference on Pattern Recognition ([ICPR](#)) 2022
- Asian Conference on Computer Vision ([ACCV](#)) 2020
- IEEE/IAAPR International Joint Conference on Biometrics ([IJCB](#)) 2020-25
- IEEE Conference on Biometrics: Theory, Application & Systems ([BTAS](#)) 2019
- IAPR International Conference on Biometrics ([ICB](#)) 2019.

INVITED TALKS:

- Dept. of Computer Science, Colorado State University (04/2021)
- The MIT AgeLab/Advanced Vehicle Technology ([AVT](#)) Consortium (01/2021)

- Doctoral Consortium at [IEEE WACV 2019](#)
- [Midwest Vision Workshop 2018](#)
- Amazon Graduate Research Symposium 2017
- NSF Data Science Workshop 2016.

MENTORING: Vignesh Gokul (UC San Diego), Kunjal Panchal (UMass Amherst), Prashant Mahajan (Amazon).

REFERENCES

Available upon request.