Deep Chakraborty

dchakraborty@umass.edu
deepc94.github.io

Summary

Machine learning researcher with 7+ years of experience in computer vision and self-supervised learning. Published at top venues (ECCV, CVPR, AISTATS) with industry research experience at Apple. Expertise in foundation model pre-training, deep learning for vision and speech, and hardware-optimized architectures.

Education

University of Massachusetts Amherst — Amherst, MA
Ph.D. and M.S. in Computer Science — Sep 2017 – May 2026

- **Dissertation:** Information-Theoretic Methods for Understanding and Improving Representations in Neural Networks, advised by Erik Learned-Miller
- **Courses:** Computer Vision, Deep Learning, Advanced ML, Advanced NLP, Distributed Systems

Manipal University — Manipal, India
B.Tech. in Electronics and Communication Engineering — Jul 2012 – Jul 2016

Professional Experience

Apple — Sunnyvale, CA
Machine Learning Research Intern, Video Computer Vision — Jun – Sep 2021

- Developed neural architecture search pipelines for hardware-optimized image classification models at ImageNet scale using unsupervised learning objectives

Apple — Sunnyvale, CA
Machine Learning Research Intern, Video Computer Vision — Jun – Sep 2020

- Built end-to-end scene graph generation system achieving joint improvements in object detection and relationship estimation on Visual Genome scale data
- Work shipped in commercial AR/VR product and resulted in granted patent (US 12,315,237)

Signify (*formerly* Philips Lighting) — Cambridge, MA
Research and Development Intern, Speech Processing and Deep Learning — May 2018 – Aug 2018

- Developed a generative method to convert the original emotion in spoken language to a target emotion without using paired training data

Academic Experience

University of Massachusetts Amherst — Amherst, MA
Graduate Research Assistant — Jan 2018 – Present

- Developed self-supervised representation learning methods achieving 3-5% improvement on image classification, object detection, and tracking benchmarks
- Built automatic Mars terrain categorization system with 83.6% retrieval precision on a new 2.4M-image dataset collected from the Curiosity rover with a 123-class taxonomy
- Designed deep saliency methods for pedestrian detection from thermal images, improving log-average miss rate by 13-20% and mean average precision by 6-7%
- Teaching assistant for graduate and undergraduate ML/AI courses, designed course material, led discussion sections, and delivered guest lectures

Selected Publications

[1] **Improving Pre-Trained Self-Supervised Embeddings Through Effective Entropy Maximization**
Deep Chakraborty, Yann LeCun, Tim G.J. Rudner, Erik Learned-Miller
International Conference on Artificial Intelligence and Statistics (AISTATS), 2025
→ Proposed entropy maximization framework for continued pre-training of foundation models

[2] **Self-Supervised Learning to Guide Scientifically Relevant Categorization of Martian Terrain Images**
T. Panambur*, **D. Chakraborty***, M. Meyer, R. Milliken, E. Learned-Miller, M. Parente
Computer Vision and Pattern Recognition Workshops (CVPRW), 2022 (**Oral**)

[3] **Pedestrian Detection in Thermal Images using Saliency Maps**
D. Ghose*, S.M. Desai*, S. Bhattacharya*, **D. Chakraborty***, M. Fiterau, T. Rahman
Computer Vision and Pattern Recognition Workshops (CVPRW), 2019 (**Spotlight**)
→ Deep saliency methods to massively improve pedestrian detection in thermal images

[4] **Unsupervised Hard Example Mining from Videos for Object Detection**
S. Jin*, A. Roy Chowdhury*, H. Jiang, A. Singh, A. Prasad, **D. Chakraborty**, E. Learned-Miller
European Conference on Computer Vision (ECCV), 2018
→ Self-supervised method for mining hard negatives from unlabeled videos to improve detection

[5] **Nonparallel Emotional Speech Conversion**
J. Gao, **D. Chakraborty**, H. Tembine and O. Olaleye
Annual Conference of the International Speech Communication Association (INTERSPEECH), 2019 (**Oral**)
→ GAN-based approach for emotion conversion without parallel training data

[* = equal contribution] See full list and citations on Google Scholar.

Projects

See full list of projects on github.com/deepc94

- **Probing Comparative Reasoning Abilities of Vision-Language Models**   Spring 2023

  Created a curated dataset of 307 images to test VLM's understanding of comparative adjectives. Evaluated CLIP and BLIP models on image retrieval tasks, identifying a 40% performance gap compared to human accuracy.

- **Sentiment Style Transfer in Text**   Fall 2018

  Implemented a sequence-to-sequence RNN to convert negative reviews to positive reviews using unpaired training data. Enhanced sentiment classification accuracy from 63% to 72% by leveraging weak supervision from noisy data pairs.

Skills

- **ML Frameworks:** PyTorch (expert), NumPy (expert), Hugging Face Transformers, Lightning AI, TensorFlow
- **Languages:** Python (expert), C++, Bash, LaTeX
- **Computer Vision:** Self-Supervised Learning (expert), Object Detection (expert), Tracking, Neural Architecture Search, Scene Graph Generation
- **Distributed Systems:** SLURM clusters (expert), Multi-GPU Training, AWS

Awards

- **Outstanding Reviewer**, ICCV 2025 (top 2.6% of 12,171 reviewers)   Oct 2025
- **Dissertation Writing Fellowship**, UMass Amherst CICS   Dec 2023
- **Riseman and Hanson scholarship**, UMass Amherst   Jun 2020