Digbalay Bose

Curriculum Vitae

3710 McClintock Ave
Los Angeles, CA-90007
✆ +1 323 356 7550
✉ dbose@usc.edu
 digbose92

Research Interests

Multimodal Machine Learning, Computer Vision, Affective computing, Machine learning for Healthcare

Education

2018–Present **University of Southern California(USC)**.
- Ph.D in Electrical and Computer Engineering; **GPA: 3.88/4.00**
- **Advisor:** Prof. Shrikanth Narayanan

2014–2016 **Indian Institute of Technology Bombay**.
- M.Tech in Electrical Engineering; **CPI: 9.51/10**
- **Specialization**: Control and Computing (Rank: 2/16)
- **Advisor:** Prof. Subhasis Chaudhuri

2010-2014 **Jadavpur University**.
- B.E. in Electronics and Telecommunication Engineering; **CPI: 9.34/10**
- **Specialization rank**: 2/46

Selected Publications

Please see **Google Scholar** for the complete list of publications

- **Does video summarization require videos? Quantifying the effects of language in video summarization**
  Yoonsoo Nam, Adam Lehavi, Daniel Yang, *Digbalay Bose*, Swabha Swayamdipta, Shrikanth Narayanan | ***IEEE International Conference on Acoustics, Speech and Signal Processing 2024***

- **Domain Adaptation for Sentiment Analysis Using Robust Internal Representations**
  Mohammad Rostami, *Digbalay Bose*, Shrikanth Narayanan, Aram Galstyan | ***Findings of EMNLP 2023*** [pdf]

- **MM-AU:Towards Multimodal understanding of advertisement videos**
  Digbalay Bose, Rajat Hebbar, Tiantian Feng, Krishna Somandepalli, Anfeng Xu, Shrikanth Narayanan | ***ACM Multimedia 2023*** [pdf]

- **SEAR: Semantically-grounded Audio Representations**
  Rajat Hebbar, Digbalay Bose, Shrikanth Narayanan | ***ACM Multimedia 2023 ACM Multimedia 2023*** [pdf]

- **Contextually-rich human affect perception using multimodal scene information**
  Digbalay Bose, Rajat Hebbar, Krishna Somandepalli, Shrikanth Narayanan | ***IEEE International Conference on Acoustics, Speech and Signal Processing 2023*** [pdf]

- **FedMultimodal: A Benchmark for Multimodal Federated Learning**
  Tiantian Feng, Digbalay Bose, Tuo Zhang, Rajat Hebbar, Anil Ramakrishna, Rahul Gupta, Mi Zhang, Salman Avestimehr, Shrikanth Narayanan | ***ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2023***

- **A dataset for Audio-Visual Sound Event Detection in Movies**
  Rajat Hebbar, Digbalay Bose, Krishna Somandepalli, Veena Vijai, Shrikanth Narayanan | ***IEEE International Conference on Acoustics, Speech and Signal Processing 2023*** [pdf]

- **MovieCLIP: Visual Scene Understanding in Movies**
  Digbalay Bose, Rajat Hebbar, Krishna Somandepalli, Haoyang Zhang, Yin Cui, Kree-Cole Mclaughlin, Huisheng Wang, Shrikanth Narayanan | ***Winter Conference on Applications of Computer Vision (WACV) 2023***[pdf] [project page]

- **Understanding of Emotion Perception from Art**
  Digbalay Bose, Krishna Somandepalli, Souvik Kundu, Rimita Lahiri, Jonathan Gratch, Shrikanth Narayanan | ***4th ICCV Workshop on Closing the Loop Between Language and Vision(CLVL)***, 2021 [pdf]

Patents

- **Visually Guided Query Processing**
  Ashok Pon Kumar Sree Prakash, Ayushi Dalmia, Amith Singhee, *Digbalay Bose*, Sumanta Mukherjee, Raghavendra Singh, Vikas C. Raykar| ***US Patent (US10878291B2)***, 2020 [pdf]

Work Experience

May 2023-August 2023   **NVIDIA Corporation**, *Computer Vision and Graphics Intern*.
- Developed end-to-end deep learning models for controllable portrait video animation as part of NVIDIA Maxine.

May 2022-August 2022   **NVIDIA Corporation**, *Software Engineering Intern*.
- Developed end-to-end visual and audio-visual deep learning models for high-fidelity facial animation of avatars as part of Maxine ARSDK.

July 2016-June 2018   **IBM Research Lab, India**, *Research Software Engineer*.
- Developed an end-to-end soil moisture extraction system from satellite images by incorporating image interpolation techniques as a part of IBM Geospatial Analytics suite.
- Developed explainable deep learning models in the domains of image classification and visual search as a part of retail and operations effort

May 2013 - July 2013   **Indian Statistical Institute, Kolkata**, *Research Intern*.
Advisor: Prof. Subhamoy Maitra, Applied Statistics Unit
- Developed a key recovery scheme based on the properties of Slid Pairs for stream cipher Salsa20.

Summer schools

July 2020 - August 2020   **Oxford Machine Learning Summer School**.
- Acceptance rate: 15% [Certificate]

Research Experience

- *Research Assistant,* **Signal Analysis and Interpretation Laboratory, University of Southern California (2018 - Present)**
  - **Advisor:** Prof. Shrikanth Narayanan
  - **Automated analysis of advertisement videos**[**ACM MM 2023**]:
    · Introduced large-scale advertisement benchmark dataset and multimodal models for semantic video understanding tasks.
  - **Context driven human affect perception:** [**ICASSP 2023**]:
    · Developed multimodal context fusion module for apparent emotion recognition in EMOTIC and CAER-S datasets.
  - **Multimodal federated learning**[**KDD 2023**]:
    · Co-developed multimodal benchmark tasks and baseline models for federated learning applications.
  - **Visual scene understanding** [**WACV 2023**]:
    · Proposed a large-scale weakly labeled dataset (**MovieCLIP**) of movie shots with automatic method for visual scene labeling.
    · Developed deep learning models for scene and genre classification from short video clips in HVU and MovieScopes datasets.
    · Work done in collaboration with Google Research.
  - **Automated analysis of facial paralysis patients** [**Facial Plastic Surgery and Aesthetic Medicine**]:
    · Developed a configurable web application for recording and segmenting clinical sessions involving facial paralysis patients.
    · Developed a facial landmark based video pipeline involving novel asymmetry measures for predicting standardized scores in a linear mixed effects modeling setup.
    · Work done in collaboration with Dr. Amit Kochhar and Dr. Courtney Voelker.
  - **Understanding emotion perception in art work** [**ICCV CLVL Workshop 2021**]:
    · Developed multimodal transformer (MMBT) based architectures with configurable image features for 9 class evoked emotion recognition using art images and captions in Artemis dataset.
  - **Cross-domain emotion recognition from text**:
    · Co-developed a method for few shot emotion recognition by transferring knowledge from GoEmotions dataset of Reddit comments to SemEval tweet corpus using various label representation methods.
  - **Computational analysis of gender portrayal in media**[**US TV Show Study 2022**] [**India TV Show Study 2023**]:
    · Collaborated with Geena Davis Institute on the Seejane Project to computationally analyze TV shows and advertisements from 2020-2022.
- *Research Scholar,* **Vision and Image Processing Laboratory, Indian Institute of Technology, Bombay (2015 - 2016)**
  - **Advisor:** Prof. Subhasis Chaudhuri
  - **Applications of sparsity and metric learning based methods in classification problems (Master's thesis)**
    · Developed a hierarchical scheme of fine-grained image classification based on a self tuning variant of spectral clustering followed by application of large margin nearest neighbor algorithm.

Skills

- **Languages:** Python, C, C++, R, Javascript, HTML, Bash
- **Machine Learning Frameworks:** Pytorch, Tensorflow, Keras, Caffe, Scikit-learn
- **Computer Vision Frameworks:** OpenCV, Scikit-Image, PIL

- **Softwares:** Maya, Blender, VTK
- **NLP Frameworks:** Spacy, StanfordCoreNLP

Selected Academic Projects

- **StyleIT: Style Guided Image Captioning** (CSCI 699, USC) [Report]
- **Future sales prediction using ensemble models (CSCI 567, USC)** [Report] [Code]
  - Obtained a world rank of 80 among 8292 teams in the Kaggle future sales competition
- **Visual Question Answering : Attention and Fusion based approaches (CSCI 599, USC)** [Report] [Code]
  - Awarded the best project by the poster session sponsors, Neudesic.
- **Multimodal Emotion Recognition from speech utterances (EE 599, USC)** [Slides]

Professional Service

- **Conference Review:** CVPR 2024, CoNLL 2023, Neurips Phys 2023, ICASSP 2024, ACII 2023, ACL 2023, WACV 2023, EMNLP 2023, COLING 2022, ACM MM 2022, ICME 2021, ICASSP 2020, ICME 2020, HiPC 2017

Invited Talk

- Invited talk on **Visual Scene Recognition in movies** at Netflix Research, 2023
- Invited talk on **Understanding context in movies: Taxonomies, Benchmarks and Challenges** at the 3rd Media Understanding Workshop on Context and Environment by Google Research and Center for Computational Media Intelligence, USC.

Selected Awards and Honors:

- **2021:** Phase 1 Finalist in **OpenCV AI Competition** as part of USC SAIL Team (top 200 teams globally).
- **2018:** Awarded **Annenberg Fellowship** by **University of Southern California** .
- **2016: Managers choice award** by **IBM Research, India** for research contributions in data-driven soil moisture modeling.
- **2016: Academic excellence award** by **IIT Bombay** for ranking among top 1% students of Masters in Electrical Engineering.
- **2014:** Secured All India Rank **251** out of 216367 candidates in **Graduate Aptitude Test in Engineering (GATE) 2014**

Selected Coursework

- **USC:** Advanced Computer Vision, Grounding Natural Language, Machine Learning, Deep Learning and its Applications, Affective Computing, Random Processes.
- **IIT Bombay:** Computer Vision, High-Performance Scientific Computing, Matrix Computations.

Teaching Experience

- **Fall 2020:** Teaching Assistant, EE 599: Deep Learning Systems (USC)

Mentoring

- **exploreCSR:** Mentored 3 senior and 1 freshman student for exploreCSR workshop.
- **USC Viterbi:**
  - Eshna Gupta - Freshman (Computer Science)
  - Yoonsoo Nam - 1st year Masters (Computer Science)
  - Haoyang Zhang - Sophomore (Computer Science)
  - Selina Martinez - Sophomore (Electrical and Computer Science)
  - Kishan Narashima Murty - 2nd year Masters (Computer Science)