Ankit Shah

aps1@alumni.cmu.edu | https://ankitshah009.github.io/

Education

Carnegie Mellon University, School of Computer Science — Pittsburgh, USA

Ph.D. in Language Technologies(GPA: 4.04) (Advisor - Prof. Bhiksha Raj) — October 2024

Master of Science in Language Technologies (GPA: 3.88) — August 2019

Relevant Courses: Visual Learning and Recognition, Computer Vision, Machine Learning, Machine Learning for Signal Processing, Introduction to Deep Learning, Large Scale Multimedia Analysis, Algorithms for NLP.

Udacity Nanodegree — Online Education

AWS Machine Learning Nanodegree (Scholarship) — June 2021

Natural Language Processing Nanodegree (Scholarship) — June 2020

AI Programming with Python Nanodegree (Scholarship) — May 2020

Introduction to Self Driving Cars Nanodegree Program (Scholarship) — May 2019

Secure and Private AI (Facebook Scholarship) — September 2019

National Institute of Technology Karnataka Surathkal — Mangalore, India

Bachelor of Technology in Electronics and Communication Engineering (GPA: 8.58) — November 2015

Relevant Courses: Data Structures and Algorithms, Speech and Audio Processing, Digital Signal Processing, Introduction to Computer Programming, Mathematics for Electronics and Communication.

Skills

Programming Languages: **Expert**: Python, Shell Scripting (Bash), SQL **Intermediate**: C++, Java, Verilog, MATLAB, R

LLM & GenAI: LangChain, LlamaIndex, OpenAI API, Anthropic API, Azure OpenAI, Google Gemini API, RAG (Retrieval-Augmented Generation), Vector Databases (Pinecone, ChromaDB, FAISS, Weaviate), Prompt Engineering, LLM Fine-tuning, RLHF, LoRA/QLoRA, vLLM, Ollama, Semantic Kernel, AutoGen, CrewAI, Model Context Protocol (MCP)

Deep Learning Frameworks: PyTorch, TensorFlow, JAX, HuggingFace Transformers, Keras, CUDA, DeepSpeed, ONNX, TensorRT, Caffe2, Chainer, Theano

Cloud & MLOps: AWS (SageMaker, Bedrock), Microsoft Azure (Azure ML, Azure AI), Google Cloud (Vertex AI), Docker, Kubernetes, MLflow, Weights & Biases, Ray, Triton Inference Server

Professional Experience

Accenture — Mountain View, CA USA

LLM Architecture Associate Director & Core Tech Lead (Mentor: Wei Wei) — November 2023 - Present

- Core architect of **AI Refinery**, Accenture's enterprise agentic AI platform built on NVIDIA AI Foundry, enabling organizations to deploy custom AI agents, knowledge bases, and domain-specific foundation models at scale.
- Led development of **Fortune Analytics**, an LLM-powered business intelligence platform that won the **iF Design Award 2025** and resulted in a European patent (EP4660825A1) for data visualization using large language models.
- **Physical AI Lead** driving enterprise embodied AI initiatives, integrating humanoid robotics and digital twins using NVIDIA Omniverse for industrial automation and simulation.
- Architected production LLM pipelines on Azure, performing fine-tuning of Llama, Mistral, and Mixtral models with LoRA/QLoRA and RLHF techniques for domain-specific enterprise applications.

Robert Bosch — Pittsburgh, PA USA

Machine Learning Research Intern (Mentor: Bingqing Chen, Ho-Hsiang Wu) — May 2023 - August 2023

- Developed grounded audio-visual pre-training methods inspired by GLIP/GLIGEN, aligning audio-visual representations with language to improve downstream tasks such as sound event detection and audio captioning.

Google — New York, NY USA

Research Intern (Mentor: Avner May and Dmitriy Serdyuk) — May 2022 - August 2022

- Designed self-supervised approaches for audio-visual automatic speech recognition to use video of a speaker along with corresponding audio to improve speech recognition. The approach focuses on methods to use unsupervised data to pre-train audio-visual representations with less transcribed data.

MERL — Cambridge, MA USA
Research Intern (Mentor: Chiori Hori) — May 2021 - August 2021

- Developed novel speech alignment algorithm using multimodal cues to align spoken language reference in video to relevant events.
- Organized challenge on Reasoning for Audio Visual Scene Aware Dialog where participants are expected to learn and produce answers for a dialog without captions. Developed baseline system on Audio Visual Transformer to perform the reasoning task to generate answers for the relevant dialog.

ReviveMed — Cambridge, MA USA
Deep Learning Scientist — September 2019 - August 2020

- Designed pipeline to convert LC-MS (liquid chromatography-mass spectrometry) data to multimedia images which is fed to a deep Neural Network architecture to discover drugs that have effects on disease and healthy patients.
- Discovered bottleneck in peak picking algorithms to design better Artificial Intelligence (AI) capable of leveraging tens of thousands of metabolomic data points to discover novel biology and the most impactful therapeutics. Designing effective unsupervised methods and tackling mislabeled data is a critical bottleneck.

ARM — Bangalore, India
Verification Engineer in Systems and Software Group — July 2015 - August 2017

- First engineer at ARM to verify and enable chip-to-chip communication via interconnect. Resolved **5** critical RTL bugs during design delivery cycles, driving improvements across global teams.
- Devised an in-house verification tool to simultaneously verify 52 subsystems delivered to partners 15% ahead of schedule.
- Identified critical power transition scenarios and engineered targeted test cases to ensure validation across multiple IPs.

Component Design Intern, Systems and Software Group — May 2014 - July 2014

- Spearheaded the creation of a Low Power Interface for AMBA 4 and ACE Protocol, now a critical element in ARM designs.
- Devised a scalable module to handle synchronization of events between CPU's in a multi-cluster CPU configuration.

Research Experience

Graduate Research Assistant, Language Technologies Institute — September 2020 - November 2023
Carnegie Mellon University, School of Computer Science — *Research Advisor:* Prof. Bhiksha Raj

Semi Weak Label Learning

- Formulated a novel Semi-weak label learning paradigm to bridge the gap between weak and strong supervision, incorporating information such as event count and sequence order alongside traditional weak labels indicating event presence or absence.
- Demonstrated the framework's feasibility through experiments, outperforming baseline models trained on weakly labeled datasets with over 10% improvement.
- Thesis topic - Computational Audition from Imprecise Labels.

COVID-19 detection from Voice

- Demonstrated the feasibility of using voice as a diagnostic signal for COVID-19 and worked on the creation of web collection tool - Record your voice to help AI beat Covid!
- Conducted experimental analysis to detect COVID-19 under various circumstances to achieve an AUC of 0.9 and work in progress to finalize the right feature selection process to deduce COVID-19 from cough and speech signals.

DSTA - Voice profiling

- Research Project on deducing human bio-physical parameters through voice profiling. For example: built the models to predict the human's height, weight, age from their voice.
- Continual learning to improve performance on Audio classifiers for custom DSTA audio files which are particularly rare sound classes and hard to disambiguate.

Graduate Research Assistant, Language Technologies Institute September 2017 - August 2019
Carnegie Mellon University, School of Computer Science *Research Advisor:* Prof. Alexander Hauptmann

IARPA DIVA: Deep Intermodal Video Analytics

- Improved small object detection for activity recognition and improved identification of activities such as texting and talking on phone, person-vehicle interaction activities and vehicle only activities in surveillance videos.
- Analyzed the system integration tasks and developed the R-C3D pipeline to perform end to end activity recognition.
- Developing efficient pose estimation models in comparison with OpenPose to detect human activities.

Graduate Research Assistant, Language Technologies Institute September 2017 - August 2019
Carnegie Mellon University, School of Computer Science *Research Advisor:* Prof. Alexander Hauptmann

Drawing inferences from Acoustic Analysis of Gunshot Recordings

- Pioneered the first guntype detection system capable of accurately deciphering guntype from cell phone recordings through wavelet analysis of gunshot recordings.
- Engineered a user interface to display gunshots and performed testing on the real world incidents such as Las Vegas Mass Shooting, Florida School Shooting recordings to validate the model predictions.

Webly and Weakly supervised learning of sound events

- Quantified the effect of label corruption and label density for weakly supervised training of Audio Events with state-of-the-art performance on AudioSet data through WALNet.
- Designed WeblyNet architecture with two neural networks co-teaching each other to learn from audio data using the web with no manual labelling, 17% improvement over strong baseline performance with WALNet.

Research Scholar, CMU-NITK Winterschool December 2014 - December 2018
Never Ending Learning of Sound [NELS] *Mentors*: Prof. Bhiksha Raj and Prof. Rita Singh

- Developed a web-based artificial intelligence system to crawl the web 24x7, distinguish and index sounds and automatically learn their meanings, associations, and semantics to events, objects and places with minimal human intervention.
- Created sound vocabularies and analyzed audio files with an intuitive web interface to obtain user feedback for semi-supervised learning accessible at nels.cs.cmu.edu. Enhanced sound detectors via feature extraction and classifier fusion.

Selected Talks and Tutorial

Ankit Shah, **Invited Speaker** at Microsoft Ignite 2025, "Training and Deploying Reasoning Models with Microsoft Foundry and Azure ML", Microsoft Ignite 2025 Link

Ankit Shah, **Panelist** at "AI Product Showdown ft. experts from NVIDIA, Accenture, and Autodesk", Product @ CMU, San Francisco, 2024 Link

Ankit Shah, Bhiksha Raj, Anurag Kumar, **Tutorial** on "Learning from Weak labels", Interspeech 2022 Link

Ankit Shah, Anurag Kumar, Alexander Hauptmann, Bhiksha Raj "Learning Sound Events from Webly labeled data", International Joint Conferences on Artificial Intelligence, 2019

Ankit Shah, Alexander Hauptmann, "Deciphering Guntype hierarchy using Acoustic analysis of Gunshot Recordings", Language Technologies Institute, Student Research Symposium, 2018

Ankit Shah, Anurag Kumar, Alexander Hauptmann, Bhiksha Raj, "A Closer Look at Weak Label Learning for Audio Events", Language Technologies Institute, Student Research Symposium, 2018

Thesis

Ph.D. Thesis: Computational Audition with Imprecise Labels, Ankit Parag Shah, CMU-LTI-24-018 submitted in partial fulfillment of requirements for Doctor of Philosophy in Language and Information Technology, Carnegie Mellon University, 2024 Link

Undergraduate Thesis: Hardware Architecture for High-Radix Adaptive CORDIC Algorithm, Ankit Parag Shah, Saharsh Samir Oza, Tarun Thokala, Pratik Gujjar, submitted in partial fulfillment of requirements for Bachelor of Technology in Electronics and Communication Engineering, National Institute of Technology Karnataka, 2015 Link

Patents

[1] Wei Wei, Yujia Bao, **Ankit Parag Shah**, Su Min Park, et al., "Application Server and Method for Generating Data Visualizations Using Large Language Models", European Patent EP4660825A1, Accenture Global Solutions Ltd, Filed: June 2025, Published: December 2025 Link

Publications

[1] **Ankit Parag Shah**, Mohammad-Parsa Hosseini, Su Min Park, Connie Miao, Wei Wei, Small Language Models: Architecture, Evolution, and the Future of Artificial Intelligence, Preprints 2026 Link

[2] Zhenting Wang, Qi Chang, Hemani Patel, Shashank Biju, Cheng-En Wu, Quan Liu, Aolin Ding, Alireza Rezazadeh, **Ankit Shah**, Yujia Bao, Eugene Siow, Mcp-bench: Benchmarking tool-using llm agents with complex real-world tasks via mcp servers submitted to The Thirteenth International Conference on Learning Representations (ICLR) 2026, Link

[3] **Ankit Shah**, Rita Singh, Bhiksha Raj, Alexander Hauptmann, Deciphering GunType Hierarchy through Acoustic Analysis of Gunshot Recordings submitted at International Conference on Acoustics, Speech, and Signal Processing, (ICASSP) 2026, Link

[4] Deepak Pandita, Tharindu Cyril Weerasooriya, **Ankit Parag Shah**, Isabelle Diana May-Xin Ng, Christopher M Homan, Wei Wei, ProRefine: Inference-time Prompt Refinement with Textual Feedback, The Thirty-Ninth Annual Conference on Neural Information Processing Systems, NeurIPS 2025 Link

[5] Shafiuddin Rehan Ahmed, **Ankit Parag Shah**, Quan Hung Tran, Vivek Khetan, Sukryool Kang, Ankit Mehta, Yujia Bao, Wei Wei, Enhancing Retrieval for ESGLLM via ESG-CID–A Disclosure Content Index Finetuning Dataset for Mapping GRI and ESRS, arXiv 2025 Link

[6] Yaxuan Wang, Jiaheng Wei, Chris Yuhao Liu, Jinlong Pang, Quan Liu, **Ankit Parag Shah**, Yujia Bao, Yang Liu, Wei Wei, LLM Unlearning via Loss Adjustment with Only Forget Data, Thirteenth International Conference on Learning Representations (ICLR) 2025, Link

[7] Jinlong Pang, Jiaheng Wei, **Ankit Parag Shah**, Zhaowei Zhu, Yaxuan Wang, Chen Qian, Yang Liu, Yujia Bao, Wei Wei, Improving Data Efficiency via Curating LLM-Driven Rating Systems, accepted at The Thirteenth International Conference on Learning Representations (ICLR) 2025 Link

[8] Ksheeraja Raghavan*, Samiran Gode*, **Ankit Shah***, Surabhi Raghavan, Wolfram Burgard, Bhiksha Raj, Rita Singh, Did You Hear That? Introducing AADG: A Framework for Generating Benchmark Data in Audio Anomaly Detection, arXiv 2024 Link

[9] Hao Chen*, **Ankit Shah***, Jindong Wang, Ran Tao, Yidong, Wang, Xing Xie, Masashi Sugiyama, Rita Singh, Bhiksha Raj, Imprecise Label Learning: A Unified Framework for Learning with Various Imprecise Label Configurations, 2024 Conference on Neural Information Processing Systems, NeurIPS 2024 Link

[10] Minghao Liu, Zonglin Di, Jiaheng Wei, Zhongruo Wang, Hengxiang Zhang, Ruixuan Xiao, Haoyu Wang, Jinlong Pang, Hao Chen, **Ankit Shah**, Hongxin Wei, Xinlei He, Zhaowei Zhu, Haobo Wang, Lei Feng, Jindong Wang, James Davis, Yang Liu, Automatic Dataset Construction (ADC): Sample Collection, Data Curation, and Beyond, arXiv 2024 Link

[11] Yujia Bao*, **Ankit Parag Shah***, Neeru Narang*, Jonathan Rivers, Rajeev Maksey, Lan Guan, Louise N Barrere, Shelley Evenson, Rahul Basole, Connie Miao, Ankit Mehta, Fabien Boulay, Su Min Park, Natalie E Pearson, Eldhose Joy, Tiger He, Sumiran Thakur, Koustav Ghosal, Josh On, Phoebe Morrison, Tim Major, Eva Siqi Wang, Gina Escobar, Jiaheng Wei, Tharindu Cyril Weerasooriya, Queena Song, Daria Lashkevich, Clare Chen, Gyuhak Kim, Dengpan Yin, Don Hejna, Mo Nomeli, Wei Wei, Harnessing Business and Media Insights with Large Language Models, arXiv 2024 Link

[12] Joseph Konan, Ojas Bhargave, Shikhar Agnihotri, Shuo Han, Yunyang Zeng, **Ankit Shah**, Bhiksha Raj Psychoacoustic Challenges Of Speech Enhancement On VoIP Platforms, Syn4DataGenAI Workshop, co-organized with Interspeech 2024 Link

[13] Hao Chen, Jindong Wang, **Ankit Shah**, Ran Tao, Hongxin Wei, Xing Xie, Masashi Sugiyama, Bhiksha Raj, Understanding and Mitigating the Label Noise in Pre-training on Downstream Tasks, The International Conference on Learning Representations (ICLR) 2024. Link

[14] **Ankit Shah**, Fuyu Tang, Zelin Ye, Rita Singh, Bhiksha Raj, Importance of negative sampling in weak label learning, International Conference on Acoustics, Speech, and Signal Processing, (ICASSP) 2024. Link

[15] Oscar Chang, Dmitriy Serdyuk, Hank Liao, **Ankit Shah**, Olivier Siohan, Conformer is all you need for Visual Speech recognition, IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2024. Link

[16] Ashwin Pillay, Sage Betko, Ari Liloia, Hao Chen, **Ankit Shah**, Exploring Domain-Specific Enhancements for a Neural Foley Synthesizer, arXiv 2023 Link

[17] Ashwin Pillay, Sage Betko, Ari Liloia, Hao Chen, **Ankit Shah**, DCASE Task 7: Foley Sound Synthesis, DCASE 2023 Challenge Technical Report Link

[18] Avner May, Dmitriy Serdyuk, **Ankit Shah**, Otavio Braga, Olivier Siohan, Audio-visual fine-tuning of audio-only ASR models, arXiv, 2023 Link

[19] Muhammad Ahmed Shah, Roshan Sharma, Hira Dhamyal, Raphael Olivier, **Ankit Shah**, Dareen Alharthi, Hazim T Bukhari, Massa Baali, Soham Deshmukh, Michael Kuhlmann, Bhiksha Raj, Rita Singh, Loft: Local proxy fine-tuning for improving transferability of adversarial attacks against large language model, arXiv 2023 Link

[20] Mark Lindsey, **Ankit Shah**, Francis Kubala, Richard M Stern, Online Active Learning For Sound Event Detection, available at arXiv 2023 Link

[21] **Ankit Shah**, Larry Tang, Po Hao Chou, Yi Yu Zheng, Ziqian Ge, Bhiksha Raj, An Approach to Ontological Learning from Weak Labels, ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2023. Link

[22] **Ankit Shah**, Shuyi Chen, Kejun Zhou, Yue Chen, Bhiksha Raj, Approach to Learning Generalized Audio Representation Through Batch Embedding Covariance Regularization and Constant-Q Transforms, arXiv 2023. Link

[23] Joseph Konan, Ojas Bhargave, Shikhar Agnihotri, Hojeong Lee, **Ankit Shah**, Shuo Han, Yunyang Zeng, Amanda Shu, Haohui Liu, Xuankai Chang, Hamza Khalid, Minseon Gwak, Kawon Lee, Minjeong Kim, Bhiksha Raj, Improving Perceptual Quality, Intelligibility, and Acoustics on VoIP Platforms, arXiv, 2023 Link

[24] Koichiro Yoshino, Yun-Nung Chen, Paul Crook, Satwik Kottur, Jinchao Li, Behnam Hedayatnia, Seungwhan Moon, Zhengcong Fei, Zekang Li, Jinchao Zhang, Yang Feng, Jie Zhou, Seokhwan Kim, Yang Liu, Di Jin, Alexandros Papangelis, Karthik Gopalakrishnan, Dilek Hakkani-Tur, Babak Damavandi, Alborz Geramifard, Chiori Hori, **Ankit Shah**, Chen Zhang, Haizhou Li, João Sedoc, Luis F D'haro, Rafael Banchs, Alexander Rudnicky, Overview of the Tenth Dialog System Technology Challenge: DSTC10, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2023 Link

[25] Clive Gomes, Hyejin Park, Patrick Kollman, Yi Song, Iffanice Houndayi, **Ankit Shah**, Automated Audio Captioning and Language-Based Audio Retrieval, arXiv 2022 Link

[26] **Ankit Shah**, Anxiang Zhang, Bhiksha Raj, Training Image Classifiers using Semi-Weakly Labeled Data, rejected at International Conference on Learning Representations (**ICLR**) 2023. Link

[27] **Ankit Shah**, Takaaki Hori, Jonathan Le Roux, Chiori Hori, DSTC10-AVSD Submission System with Reasoning using Audio-Visual Transformers with Joint Student-Teacher Learning, The Tenth Dialog System Technology Challenge Workshop at Association for the Advancement of Artificial Intelligence (**AAAI**) 2022. Link

[28] **Ankit Shah**, Shijie Geng, Peng Gao, Anoop Cherian, Tim Marks, Chiori Hori, Overview of Audio Visual Scene-Aware Dialog with Reasoning Track for Natural Language Generation in DSTC10, The Tenth Dialog System Technology Challenge Workshop at Association for the Advancement of Artificial Intelligence (**AAAI**) 2022. Link

[29] Rita Singh, **Ankit Shah**, Hira Yasin, An Overview of Techniques for Biomarker Discovery in Voice Signal, Speech and Audio in the Northeast 2022. (SANE 2022), Link

[30] Larry Tang, Po Hao Chou, Yi Yu Zheng, Ziqian Ge, **Ankit Shah**, Bhiksha Raj, Ontological Learning from Weak Labels, arXiv 2023, Link

[31] Yanwen Wang, **Ankit Shah**, Xunwen Qiu, Yanran Cao, Yizi Xu, Bhiksha Raj, MCCLA: An improved contrastive learning structure for audio representations, arXiv 2023.

[32] **Ankit Shah**, Srishti Singh, Shih-Yen Tao, Feature extraction and evaluation for BioMedical Question Answering, arXiv 2021 Link

[33] **Ankit Shah**, Tzu-Hsiang Lin, Shijie Wu, Triple Attention Network Architecture for MovieQA, arXiv 2021 Link

[34] **Ankit Shah**, Bhiksha Raj, Rita Singh, On the pragmatism of using binary classifiers over data-intensive neural network classifiers for detection of COVID-19 from voice, Speech and Audio in the Northeast 2022. Link

[35] **Ankit Shah**, Shijie Geng, Peng Gao, Anoop Cherian, Takaaki Hori and Tim Marks, and Jonathan Le Roux, Chiori Hori, Audio-Visual Scene-Aware Dialog and Reasoning using Audio-Visual Transformers with Joint Student-Teacher Learning, IEEE International Conference on Acoustics, Speech and Signal Processing, (**ICASSP**) 2022.

[36] **Ankit Shah**, Shijie Geng, Peng Gao, Anoop Cherian, Tim K. Marks, Chiori Hori, Reasoning for Audio Visual Scene-Aware Dialog(AVSD) Track in DSTC10

[37] Romain Serizel, Nicolas Turpault, **Ankit Shah**, Justin Salamon, Sound event detection in synthetic domestic environments, IEEE International Conference on Acoustics, Speech and Signal Processing, (**ICASSP**) 2020. Link

[38] Bhiksha Raj, Rita Singh, **Ankit Shah**, Benjamin Striner, Shahan Ali Memon, Vedant Sanil, Soham Deshmukh, Ruan Kangrui, Mahmoud Al Ismail, Hira Dhamyal, Majd Sakr, Nicholas Wolfe, Shmuel Ur, CMU Sounds for COVID Project, CMU Language Technologies Institute Technical Report 2020 Link

[39] **Ankit Shah***, Vaibhav Vaibhav*, Vasu Sharma*, Mahmoud Alismail*, Louis-Philippe Morency , Multimodal Behavior Markers Exploring Suicidal Intent in Social Media Videos, 21st ACM International Conference on Multimodal Interaction, (**ICMI**), 2019. Link

[40] Nicolas Tarpault, Romain Serizel, **Ankit Shah***, Justin Salamon, Sound event detection in domestic environments with weakly labeled data and soundscape synthesis, Detection and Classification of Acoustic Scenes and Events (DCASE) 2019. Link

[41] Anurag Kumar, **Ankit Shah**, Alexander Hauptmann, Bhiksha Raj, Learning Sound Events from Webly labeled data, International Joint Conferences on Artificial Intelligence (**IJCAI**) 2019. Link.

[42] **Ankit Shah**, Jean Baptiste Lamare, Tuan Nguyen, Alexander Hauptmann, CADP: A Novel dataset for CCTV Traffic Camera based Accident Analysis, International Workshop on Traffic and Street Surveillance for Safety and Security, 2018. Link

[43] **Ankit Shah**, Anurag Kumar, Alexander G. Hauptmann, Bhiksha Raj, "A Closer Look at Weak Label Learning for Audio Events", arXiv, 2018 Link

[44] George Larionov, Zachary Kaden, Hima Varsha Dureddy, Gabriel Bayomi T Kalejaiye, Mihir Kale, Srividya Pranavi Potharaju, **Ankit Shah**, Alexander I Rudnicky, Tartan: A retrieval-based socialbot powered by embeddings and a robust intent model, 2nd Proceedings of Alexa Social Prize, 2018. Link

[45] Romain Serizel, Nicolas Tarpault, Hamid Eghbal-Zadeh, **Ankit Shah**, "Large-Scale Weakly Labeled Semi-Supervised Sound Event Detection in Domestic Environments", Detection and Classification of Acoustic Scenes and Events (**DCASE**) 2018. (**Spotlight Paper Presentation**) Link

[46] **Ankit Shah**, Rohan Badlani, Benjamin Elizalde, Anurag Kumar, Bhiksha Raj, "Framework for evaluation of sound event detection in web videos", IEEE International Conference on Acoustics, Speech and Signal Processing, (**ICASSP**) 2018. Link

[47] Pranay Manocha, Rohan Badlani, Anurag Kumar, **Ankit Shah**, Benjamin Elizalde, Bhiksha Raj, "Content-based Representations of audio using Siamese neural networks", IEEE International Conference on Acoustics, Speech and Signal Processing, (**ICASSP**) 2018. Link

[48] Annamaria. Mesaros, Toni Heittola, Aleksandr Diment, Benjamin. Elizalde, **Ankit Shah**, Rohan Badlani, Emmanuel Vincent, Bhiksha Raj, and Tuomas Virtanen."DCASE 2017 challenge setup: tasks, datasets and baseline system" in proceedings of the Detection and Classification of Acoustic Scenes and Events 2017 Workshop (**DCASE**), 2017. (**Spotlight Paper**) Link

[49] **Ankit Shah**, Benjamin Elizalde, Rohan Badlani, Anurag Kumar, Bhiksha Raj, "NELS - Never-Ending Learner of Sounds", ML4Audio at 31st Conference on Neural Information Processing Systems (**NIPS**) 2017, Long Beach, CA, USA. Link

[50] **Ankit Shah**, Rohan Badlani, Benjamin Elizalde, Anurag Kumar, Bhiksha Raj, "A Framework towards Large Scale Learning of Sound Events", 2017 Language Technologies Institute Student Research Symposium, August 2017.

[51] **Ankit Shah**, Harini Kesavamoorthy, Poorva Rane, Pramati Kalwad, Alexander Hauptmann, Florian Metze, Activity Recognition on a Large Scale in Short Videos - Moments in Time Dataset, PrePrint, arXiv 2018. Link

[52] **Ankit Shah**, Rohan Badlani, Anurag Kumar, Benjamin Elizalde, Bhiksha Raj, "An Approach for Self-Training Audio Event Detectors Using Web Data," 25th European Signal Processing Conference (**EUSIPCO**) 2017. Link

[53] **Ankit Shah**, Bhiksha Raj, Does representation size matter more than network depth for transfer learning, under preparation for Advances in Transfer Learning: Theory, Algorithms, and Applications.

[54] Benjamin Elizalde, Anurag Kumar, **Ankit Shah**, Rohan Badlani, Emmanuel Vincent, Bhiksha Raj, Ian Lane, "Experiments on the DCASE Challenge 2016: Acoustic Scene Classification and Sound Event Detection in Real Life Recording.", Detection and Classification of Acoustic Scenes and Events (**DCASE**), 2016. (**Oral Paper Presentation**) Link1

[55] **Ankit Shah**, Tyler Vuong, Natural Language Person Search Using Deep Reinforcement Learning, arXiv 2018. Link

[56] Anurag Kumar, Benjamin Elizalde, **Ankit Shah** , Rohan Badlani, Emmanuel Vincent, Bhiksha Raj, Ian Lane, "DCASE challenge task 1", Detection and Classification of Acoustic Scenes and Events (DCASE), 2016. Link

[57] **Ankit Shah**, Ajith Bhat, Saurabh Saxena, Rashmin Mantri, "Repeatability and Scalability of Code at Top Level Verification," Regional Engineering Conference 2016, ARM.

[58] Saharsh Oza, **Ankit Shah**, Tarun Thokala, Sumam David, "Pipelined implementation of high radix adaptive CORDIC as a coprocessor," 2015 International Conference on Computing and Network Communications (CoCoNet), 2015. Link

[59] **Ankit Shah**, Alexander Hauptmann, Joint Deciphering of GunType Hierarchy and Gunshot using Acoustic Analysis, - under preparation for review IEEE Journal on Selected Topics in Signal Processing, (**JSTSP**) - incomplete.

[60] **Ankit Shah**, Chao Li, Bhiksha Raj, Learning strategies for unsupervised adaptation of test set, under preparation.

[61] **Ankit Shah**, Anurag Kumar, Bhiksha Raj, Rita Singh, Review of past, present and future directions for weak and semi-weak label learning: Past, Present and Future directions, under preparation for IEEE Transactions on Audio, Speech and Language Processing.

Awards

2025 iF Design Award 2025 Winner - Fortune Analytics, LLM-powered platform (Accenture & Fortune) Link
2023 New York - OpenAI Sponsored Generative AI Hackathon Winner for FactGPT project amongst Top 100 participants for the Hackathon Link
2022 Meta Ph.D. Fellowship - Finalist
2020 Carnegie Mellon University Language Technologies Institute Fellowship
2020 Udacity Scholarship for Nutanix Hybrid Cloud Scholarship
2020 Coursera AI for Medicine Specialization Top Student
2020 Mentor for Coursera AI for Medicine Diagnosis, AI for Medicine Prognosis and AI for Medicine Treatment
2020 Udacity Top Student Scholarship Honor for Deep Learning Nanodegree
2019 Udacity and Amazon Scholarship for pursuing AWS Machine Learning Coursework
2019 Intel - Udacity Scholarship for pursuing Intel Edge AI Nanodegree program
2018 Udacity Lyft Scholarship for Self Driving Cars Nanodegree Program
2018 CMU LTI Conference Travel Grant for ICASSP 2018
2017 Carnegie Mellon University Language Technologies Institute Fellowship
2017 R D Sethna Education Scholarship and Lotus Education Trust Scholarship
2017 Gandhian Young Technological Innovator Award 2017 for work on 'Never Ending Learning of Sound' project, amongst **39** projects selected in a pool of over 2915 submissions
2017 Recognition as Competent Communicator and Competent Leader for exceptional achievements in Toastmasters International Communication and Leader Program
2017 ARM Wearable Technology Challenge Winner - AquaeVitae portable water filtration system solving Water Crisis
2017 Bravo Award at ARM for exceptional delivery of tasks
2016 Winner of ARM Wearable Design Challenge - Junior Academy
2016 Bravo Award at ARM for fixing critical RTL bugs in multiple key projects
2016 Mentor at Junior Academy - Wearable challenge under the project titled 'Aquaevitae' at ARM
2014 IEEE Student Enterprise Award, Region 10, "Mapping of greenhouse gases using Wireless Sensor Networks"

Selected Academic Projects

Projects at CMU

Multimodal Behavior Markers of Suidical Intent in Social Media Videos December 2018 - May 2019
Multimodal Affective Computing Course project *Mentor*: Prof. Louis-Philippe Morency

- Explored behavior markers for suicidal intent in social media videos and found use of death related words, profane language as verbal markers, pitch variations and long pauses as acoustic markers whereas slouched shoulders and frequent hand movement as visual markers for flagging suicidal intent.

Action Recognition on a Large Scale in Short Videos - Moments in Time January 2018 - May 2018
Large Scale Multimedia Analysis Course Project *Mentors*: Prof. Alexander Hauptmann and Prof. Florian Metze

- Performed ablation study across various feature representations such as 3D Resnext, I3D features, ResNet features, Temporal Relation Networks as well as detailed analysis of audio features.
- Designed model is 89.23 % accurate in Top-5 accuracy as metric, a significant improvement over the baseline TRN models.

AudioNet Feature Visualization for Large Scale Audio Event Recognition August 2017 - December 2017
Machine Learning for Signal Processing Course Project *Mentor*: Prof. Bhiksha Raj

- Visualized audio features learned at different stages of AudioNet architecture and associated meaning to the features extracted and created analogies to various phases of the audio learnt at each level.

- Concluded that log mel spectrogram is a superior feature representation for Convolutional Network architecture over Mfcc and spectrogram with an absolute improvement of 5% over the counterparts.

Natural Language Person Search Using Deep Reinforcement Learning August 2017 - December 2017
Introduction to Deep Learning Course Project *Mentor*: Prof. Bhiksha Raj

- Detected the person in an image based on their natural language description using deep reinforcement learning which optimizes to the box estimates and rewards network based on accurate localization of person.

Triple Attention Network Architectures for MovieQA February 2018 - May 2018
Introduction to Machine Learning Course Project *Mentor*: Prof. Manuela Veloso and Prof. Pradeep Ravikumar

- Devised a triple attention architecture with focused efforts on quantifying effect of visual and audio modality. Results indicate audio modality provides complementary information to textual based question answering.

Undergraduate Project

Sound Event Detection in Real Life Audio: Example-Based Retrieval June 2016 - September 2016
DCASE 2016 challenge - Results *Mentor*: Prof. Bhiksha Raj

- Developed an optimized machine learning pipeline to retrieve test audio label using a trained multi-class classifier.
- Obtained a significant reduction in Segment-based Error Rate of 0.48 as compared to baseline performance of 0.91, thus ranked 9^{th} amongst 87 participants in DCASE challenge.

Hardware Architecture for High-Radix Adaptive CORDIC June 2014 - April 2015
Undergraduate Thesis *Mentor*: Prof. Sumam David

- Formulated FSM design of floating point HCORDIC using verilog with performance comparable to CORDIC IP and achieving over 1.5x computation throughput on hardware for trigonometric and signal processing related math operations.
- Achieved 2-3x speed up with HCORDIC, enhancing traditional CORDIC through use of hardware multipliers.

Lossless Compression of Medical Images February 2015 - March 2015
Digital Signal Compression Project *Mentor*: Prof. Deepu Vijayasenan

- Analyzed and compared compression ratios for various lossless compression techniques and found Discrete Haar Wavelet transform provided an average compression of 2.3 over original image, **30**% increase over other methods.

Teaching and Mentorship

Teaching Assistant for 11788- Computational Forensics Jan 2022 - May 2022
Carnegie Mellon University *Instructor*:Prof. Rita Singh

Teaching Assistant for 11767-On Device Machine Learning Sept 2021 -Dec 2021
Carnegie Mellon University *Mentor*: Prof. Yonatan Bisk and Prof. Emma Strubell

Teaching Assistant for 10707 - Topics in Deep Learning Course January 2019 - May 2019
Carnegie Mellon University *Mentor*: Prof. Ruslan Salakhutdinov

Mentor at The Junior Academy, Global STEM Alliance July 2016 - May 2017
Mentored a team of young students on wearables challenge implementing an innovative water filtration system. My team demonstrated a working prototype of filter resulting in safe drinking water, thus **won** the wearables challenge.

Teaching Assistant for Data Structures and Algorithms, Elements of ECE January 2014 - June 2015
National Institute of Technology Karnataka *Mentor*: Prof. Arulalan Rajan

Professional Service

Reviewer - Conferences

Journal	EURASIP Journal on Audio Speech and Music Processing, 2023
Conference	AAAI Conference on Artificial Intelligence (AAAI) 2026 (Meta-Reviewer)
Conference	Neural Information Processing Systems, (NeurIPS) 2025, 2024, 2023, 2022
Conference	Conference on Computer Vision and Pattern Recognition, (CVPR) 2026, 2024
Conference	International Conference on Learning Representations (ICLR) 2025, 2024
Conference	IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2024, 2023, 2022, 2021, 2019, 2018
Conference	International Conference on Multimodal Interfaces (ICMI) 2023, 2022, 2020, 2019
Journal	IET Signal Processing 2020
Workshop	IEEE Detection and Classification of Acoustic Scenes and Events (DCASE) 2023, 2021, 2019, 2018, 2017
Research Symposium	LTI Student Research Symposium 2021, 2022, 2018, 2017

Academic Editor at PLOS ONE Journal Oct 2023 - present

Organizer of Reasoning for Audio Visual Scene-Aware Dialog Track 4 - DSTC-10 May 2021 - August 2021
Mitsubishi Electric Research Laboratories *Mentor*: Dr. Chiori Hori
Organizer of Reasoning for Audio Visual Scene Aware Dialog Track 4 at DSTC-10. Responsible for the Code Development of the baseline system and addressing concerns of participants for the challenge ".

Organizer of Task 4 DCASE 2020 Challenge January 2020 - August 2020
Carnegie Mellon University *Mentor*: Prof. Bhiksha Raj
Organizer of IEEE DCASE Task 4 "Sound event detection and separation in domestic environments".

Organizer of Task 4 DCASE 2019 Challenge January 2019 - August 2019
Carnegie Mellon University *Mentor*: Prof. Alexander Hauptmann
Organizer of IEEE DCASE Task 4 "Sound event detection in domestic environments".

Organizer of Task 4 DCASE 2018 Challenge March 2018 - November 2019
Carnegie Mellon University *Mentor*: Prof. Alexander Hauptmann
Organizer of DCASE Task 4 "Large-scale weakly labeled semi-supervised sound event detection in domestic environments".

Organizer of Task 4 DCASE 2017 Challenge April 2017 - Nov 2017
Carnegie Mellon University *Mentor*: Prof. Bhiksha Raj
Organizer of IEEE-DCASE 2017 challenge - Task 4 "Large Scale weakly supervised sound event detection for smart cars".

Leadership

Membership Chair, Indian Graduate Students Association, Carnegie Mellon University

President and Vice President of ARM Talkies - Toastmasters Club at ARM.

Joint Convener and Treasurer of Robotics Club, NITK

Vidyut Chairman - Computer Society of India and IEEE NITK

Executive Lead of Engineer's Forum for Entrepreneurship Awareness, and Joint Convener of Electronics Committee, NITK

Selected Press Coverage

iF Design Award, 2025	Fortune Analytics - iF Design Award 2025 Winner (User Interface category)
Prodigy Finance, 2025	The International Student Podcast: From India to Carnegie Mellon & Accenture - Ankit Shah's international student journey
Fortune, 2024	Using Generative-AI to uncover meaningful business insights
LTI CMU, 2023	LTI Student Wins NYC AI Hackathon, PhD Student Ankit Shah's team win for their project "FactGPT"
Washington Post, 2022	Evin on fire: What really happened inside Iran's most notorious prison: Provided analysis on gunshot analysis and types of gun used
Business Insider 2020,	Do I sound sick to you? Researchers are building AI that would diagnose COVID-19 by listening to people talk.
Pittsburgh News, 2020	Coronavirus Detected By Voice? Carnegie Mellon Researchers Develop App To 'Listen' For Signs Of COVID-19
NYAS Junior Academy, 2017	I am NYAS: Ankit Shah
GYTI Techpedia, 2017	Gandhian Young Technological Innovation Award
Elsevier, 2016	On winning the ARM wearable technology challenge
NYAS, 2016	The New York Academy of Sciences - Meet the 2016 Wearables Challenge Winning Team
Deccan Herald, 2014	Never Ending Learning of Sound, 44 students prepare 10 projects at Winter School