

# Ankit Shah

aps1@andrew.cmu.edu | <https://ankitshah009.github.io/>

---

## Education

<b>Carnegie Mellon University, School of Computer Science</b>	Pittsburgh, USA
<b>Ph.D. Candidate in Language Technologies</b> (GPA: 4.04) (Advisor - Prof. Bhiksha Raj)	June 2024
<b>Master of Science in Language Technologies</b> (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.	
<b>Udacity Nanodegree</b>	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
<b>National Institute of Technology Karnataka Surathkal</b>	Mangalore, India
<b>Bachelor of Technology in Electronics and Communication Engineering</b> (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

**Coding Languages: Expert:-** Python, Verilog, Shell Scripting (Bash, awk), C, MATLAB, R **Intermediate:-** C++ and Java  
**Deep Learning Frameworks:** TensorFlow, PyTorch, JAX, Keras, Caffe2, Chainer, Theano, CUDA  
**Cloud Computing Platforms:** AWS, Microsoft Azure, Google Cloud Platform

---

## Professional Experience

<b>Accenture</b>	Mountain View, CA USA
<b>LLM Architecture Associate Director</b> (Mentor: Wei Wei)	Nov 2023 - present
<ul style="list-style-type: none"><li>• End-to-end full stack development of Large Language Models to build custom solutions for Accenture clients.</li><li>• Led the effort to build Foundation models using various data sources collected and build an end-to-end stack with production-quality code pipelines on Azure Platform</li><li>• Performed finetuning of the model drastically enhancing the base model capabilities of Mistral, Llama, Mixtral models for domain-specific tasks.</li></ul>	
<b>Robert Bosch</b>	Pittsburgh, PA USA
<b>Machine Learning Research Intern</b> (Mentor: Bingqing Chen, Ho-Hsiang Wu)	May 2023 - August 2023
<ul style="list-style-type: none"><li>• Enhancing various downstream tasks using grounding Audio Visual and Textual pre-training.</li></ul>	
<b>Google</b>	New York, NY USA
<b>Research Intern</b> (Mentor: Avner May and Dmitriy Serdyuk)	May 2022 - August 2022
<ul style="list-style-type: none"><li>• 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</li></ul>	
<b>MERL</b>	Cambridge, MA USA
<b>Research Intern</b> (Mentor: Chiori Hori)	May 2021 - August 2021
<ul style="list-style-type: none"><li>• Developed novel speech alignment algorithm using multimodal cues to align spoken language reference in video to relevant events.</li><li>• Organized challenge on <a href="#">Reasoning for Audio Visual Scene Aware Dialog</a> 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.</li></ul>	

---

# Ankit Shah

aps1@andrew.cmu.edu | <https://ankitshah009.github.io/>

---

## ReviveMed Deep Learning Scientist

Cambridge, MA USA  
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 Verification Engineer in Systems and Software Group

Bangalore India  
July 2015 - August 2017

- First engineer at ARM to verify and enable chip-to-chip communication via interconnect. Fixed 5 critical RTL bugs across design delivery cycles impacting multiple teams across the globe.
- Devised an in-house verification tool to simultaneously verify 52 subsystems delivered to partners 15% ahead of schedule.
- Identified and designed directed test cases to verify unique power transition scenarios for multiple IPs.

## Component Design Intern, Systems and Software Group

May 2014 - July 2014

- Built a Low Power Interface for AMBA 4 and ACE Protocol, now a fundamental component in ARM designs.
  - Developed a module to handle synchronization of events between CPU's in a multi-cluster CPU configuration.
- 

## Research Experience

### Graduate Research Assistant, Language Technologies Institute

Carnegie Mellon University, School of Computer Science

September 2020 - Nov 2023  
Research Advisor: Prof. Bhiksha Raj

#### *Semi Weak Label Learning*

- A novel Semi-weak label learning paradigm as a middle ground to mitigate problem introduced wherein we utilize semi-weak label data defined as data where presence/absence of a given class and exact count of the given class is present.
- Experimental analysis show the feasibility of our approach wherein the framework introduced outperforms baseline model trained on weakly labeled datasets.
- 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

Carnegie Mellon University, School of Computer Science

September 2017 - August 2019  
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.
-

# Ankit Shah

aps1@andrew.cmu.edu | <https://ankitshah009.github.io/>

---

## Graduate Research Assistant, Language Technologies Institute

Carnegie Mellon University, School of Computer Science

September 2017 - August 2019

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

*Never Ending Learning of Sound* [NELS]

December 2014 - December 2018

Mentors: Prof. Bhiksha Raj and Prof. Rita Singh

- Developed a web-based artificial intelligence system to crawl the web 24x7, detect 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](http://nels.cs.cmu.edu). Enhanced sound detectors via feature extraction and classifier fusion.
- 

## Selected Talks and Tutorial

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

**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

---

## Publications

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

[2] Muhammad Ahmed Shah, Roshan Sharma, Hira Dharmyal, 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](#)

[3] 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](#)

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

[5] **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](#)

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

[7] Joseph Konan, Ojas Bhargave, Shikhar Agnihotri, Shuo Han, Yunyang Zeng, **Ankit Shah**, Bhiksha Raj Psychoacoustic Challenges Of Speech Enhancement On VoIP Platforms, arXiv 2023 [Link](#)

---

# Ankit Shah

aps1@andrew.cmu.edu | <https://ankitshah009.github.io/>

---

## Publications

- [8] 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, Transactions on Machine Learning Research, TMLR 2024 [Link](#)
- [9] **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](#)
- [10] **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](#)
- [11] 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](#)
- [12] **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.
- [13] 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](#)
- [14] 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](#)
- [15] **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.
- [16] **Ankit Shah**, Anxiang Zhang, Bhiksha Raj, Training Image Classifiers using Semi-Weakly Labeled Data, rejected at International Conference on Learning Representations (ICLR) 2023. [Link](#)
- [17] **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.
- [18] **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.
- [19] 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](#)
- [20] Larry Tang, Po Hao Chou, Yi Yu Zheng, Ziqian Ge, **Ankit Shah**, Bhiksha Raj, Ontological Learning from Weak Labels, available at arXiv 2023, [Link](#)
- [21] Yanwen Wang, **Ankit Shah**, Xunwen Qiu, Yanran Cao, Yizi Xu, Bhiksha Raj, MCCLA: An improved contrastive learning structure for audio representations, under submission to ICASSP 2023.
- [22] **Ankit Shah**, Srishti Singh, Shih-Yen Tao, Feature extraction and evaluation for BioMedical Question Answering, available at arXiv 2021 [Link](#)
- [23] **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](#)
- [24] **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.
- [25] **Ankit Shah**, Shijie Geng, Peng Gao, Anoop Cherian, Tim K. Marks, Chiori Hori, Reasoning for Audio Visual Scene-Aware Dialog (AVSD) Track in DSTC10
- [26] 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](#)
-

# Ankit Shah

aps1@andrew.cmu.edu | <https://ankitshah009.github.io/>

---

## Publications

- [27] **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](#)
- [28] 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](#)
- [29] Anurag Kumar, **Ankit Shah**, Alexander Hauptmann, Bhiksha Raj, Learning Sound Events from Webly labeled data, International Joint Conferences on Artificial Intelligence (IJCAI) 2019. [WeblyNet](#)
- [30] **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](#)
- [31] **Ankit Shah**, Anurag Kumar, Alexander G. Hauptmann, Bhiksha Raj, "A Closer Look at Weak Label Learning for Audio Events", arXiv, 2018 [Link](#)
- [32] 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](#)
- [33] 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](#)
- [34] **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](#)
- [35] 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](#)
- [36] 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](#)
- [37] **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](#)
- [38] **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.
- [39] **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](#)
- [40] **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](#)
- [41] 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](#)
- [42] **Ankit Shah**, Tyler Vuong, Natural Language Person Search Using Deep Reinforcement Learning, arXiv 2018. [Link](#)
- [43] 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](#)
- [44] **Ankit Shah**, Ajith Bhat, Saurabh Saxena, Rashmin Mantri, "Repeatability and Scalability of Code at Top Level Verification," Regional Engineering Conference 2016, ARM.
- [45] 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](#)
- [46] **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.
- [47] **Ankit Shah**, Chao Li, Bhiksha Raj, Learning strategies for unsupervised adaptation of test set, under preparation.
-

# Ankit Shah

aps1@andrew.cmu.edu | <https://ankitshah009.github.io/>

---

## Awards

- 2023 New York - Open AI Sponsored Generative AI Hackathon Winnner 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
  - 2017 Carnegie Mellon University Language Technologies Institute Fellowship
  - 2018 Udacity Lyft Scholarship for Self Driving Cars Nanodegree Program
  - 2018 CMU LTI Conference Travel Grant for ICASSP 2018
  - 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 Suicidal 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*

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

# Ankit Shah

aps1@andrew.cmu.edu | <https://ankitshah009.github.io/>

---

## 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<sup>th</sup> 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 Vijayasanen*

- 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.
- 

## Professional Service

### Reviewer - Conferences

Journal	EURASIP Journal on Audio Speech and Music Processing, 2023
Conference	Neural Information Processing Systems, (NeurIPS) 2022, 2023
Conference	Conference on Computer Vision and Pattern Recognition, (CVPR) 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".

---

# Ankit Shah

aps1@andrew.cmu.edu | <https://ankitshah009.github.io/>

---

## Professional Service

### Organizer of Task 4 DCASE 2018 Challenge

Carnegie Mellon University

Organizer of DCASE Task 4 "Large-scale weakly labeled semi-supervised sound event detection in domestic environments".

March 2018 - November 2019

*Mentor:* Prof. Alexander Hauptmann

### Organizer of Task 4 DCASE 2017 Challenge

Carnegie Mellon University

Organizer of IEEE-DCASE 2017 challenge - Task 4 "Large Scale weakly supervised sound event detection for smart cars".

April 2017 - Nov 2017

*Mentor:* Prof. Bhiksha Raj

---

## Teaching and Mentorship

### Teaching Assistant for 11788- Computational Forensics

Carnegie Mellon University

Jan 2022 - May 2022

*Instructor:* Prof. Rita Singh

### Teaching Assistant for 11767-On Device Machine Learning

Carnegie Mellon University

Sept 2021 - Dec 2021

*Mentor:* Prof. Yonatan Bisk and Prof. Emma Strubell

### Teaching Assistant for 10707 - Topics in Deep Learning Course

Carnegie Mellon University

January 2019 - May 2019

*Mentor:* Prof. Ruslan Salakhutdinov

### Mentor at The Junior Academy, Global STEM Alliance

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.

July 2016 - May 2017

### Teaching Assistant for Data Structures and Algorithms, Elements of ECE

National Institute of Technology Karnataka

January 2014 - June 2015

*Mentor:* Prof. Arulalan Rajan

---

## Leadership

Membership Chair, Indian Graduate Students Association, Carnegie Mellon University

Joint Convener and Treasurer of Robotics Club, NITK

Vidyut Chairman - Computer Society of India and IEEE NITK

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

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

---

## Selected Press Coverage

**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