Ankit Shah

aps1@andrew.cmu.edu | 412-626-8671 | www.cs.cmu.edu/ aps1

Education

Carnegie Mellon University — Pittsburgh, PA
Masters in Language Technologies (GPA: 3.65/4.0) — August 2019

- Graduate Research Assistant of Informedia Group, supervised by Prof. Alexander Hauptmann.
- Relevant Courses: Introduction to Machine Learning, Machine Learning for Signal Processing, Introduction to Deep Learning, Computer Vision, Large Scale Multimedia Analysis.

National Institute of Technology Karnataka Surathkal — Mangalore, India
Bachelor of Technology in Electronics and Communication Engineering (GPA: 8.58/10) — November 2015

- Relevant Courses: Data Structures and Algorithms, Speech and Audio Processing, Digital Signal Processing.

Research Experience

Graduate Research Assistant, Language Technologies Institute — September 2017 - present
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 under low resource budget to perform zero shot based learning for rare events such as texting and talking on phone in surveillance videos.
- Analyzed the system integration tasks and developed the R-C3D pipeline to perform end to end activity recognition.

Drawing inferences from Acoustic Analysis of Gunshot Recordings

- Created the first guntype detection system capable of accurately deciphering guntype from cell phone recordings through wavelet analysis of gunshot recordings.
- Developed 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.

Published several papers on research work. Kindly review my webpage and publications to learn in detail

Research Scholar, CMU-NITK Winterschool — December 2014 - present
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, 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 performed audio content analysis with an intuitive web interface to obtain user feedback for semi-supervised learning accessible at nels.cs.cmu.edu.
- Enhanced sound detectors through database of crawled sounds, multiple feature extraction techniques and classifier fusion.

Professional Experience

Verification Engineer at ARM — Bangalore, India
System and Software Group — July 2015 - August 2017

- 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.
- 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.

Component Design Intern at ARM — Bangalore, India
Systems and Software Group — May - July 2014

- Designed 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.

Ankit Shah

aps1@andrew.cmu.edu | 412-626-8671 |www.cs.cmu.edu/ aps1

Publications/Submitted Manuscripts

- **Ankit Shah**, Alexander Hauptmann, Joint Deciphering of GunType Hierarchy and Gunshot using Acoustic Analysis, under review IEEE Journal on Selected Topics in Signal Processing, (**JSTSP**), 2019.
- **Ankit Shah**, Anurag Kumar, Alexander Hauptmann, Bhiksha Raj, Webly Supervised Learning of Sound Events, under review IEEE Transactions on Audio, Speech and Language Processing, (**TASLP**) 2019
- Anurag Kumar, **Ankit Shah**, Alexander Hauptmann, Bhiksha Raj, Learning Sound Events from Webly labeled data, to appear in IEEE International Conference on Acoustics, Speech and Signal Processing, (**ICASSP**) April 2019 . WeblyNet
- **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
- **Ankit Shah**, Anurag Kumar, Alexander G. Hauptmann, Bhiksha Raj, "A Closer Look at Weak Label Learning for Audio Events", under review IEEE Transactions on Multimedia, 2018 Link
- 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
- 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
- **Ankit Shah**, Tyler Vuong, Natural Language Person Search Using Deep Reinforcement Learning, arXiV 2018. Link
- **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
- **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
- 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
- **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
- 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
- **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
- 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
- **Ankit Shah**, Ajith Bhat, Saurabh Saxena, Rashmin Mantri, "Repeatability and Scalability of Code at Top Level Verification," Regional Engineering Conference 2016, ARM.
- 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

Talks

- **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

Ankit Shah

aps1@andrew.cmu.edu | 412-626-8671 |www.cs.cmu.edu/ aps1

Awards

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 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
2014 IEEE Student Enterprise Award, Region 10 for "Mapping of greenhouse gases using Wireless Sensor Networks" - $1500
2013 IEEE Student Enterprise Award, Region 10 for "Quadcopter" project - $1500
2007 Regional Mathematical Olympiad in Mumbai District, Third Prize

Selected Academic Projects

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

- Detected activity on a large scale through Moments in Time dataset using various activity detection frameworks.
- 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 Aug 2017 - Dec 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 percent over the counterparts.

Natural Language Person Search Using Deep Reinforcement Learning Aug 2017 - Dec 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 an accurate localization of person

Triple Attention Network Architectures for MovieQA Feb 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.

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

- Developed FSM design of floating point HCORDIC using verilog with performance comparable to CORDIC IP, with improved computation throughput on hardware of trigonometric functions and other mathematical operations used in signal processing applications.
- Achieved multi-fold 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** percent increase over other methods.

Ankit Shah

aps1@andrew.cmu.edu | 412-626-8671 | www.cs.cmu.edu/ aps1

Skills

- **Deep Learning Frameworks**: Pytorch, Keras, Caffe2, Tensorflow, Theano, CUDA
- **Programming Languages and Tools**: Python, Verilog, Shell Scripting (Bash, awk), C, MATLAB, R, C++ and Java

Academic/Voluntary Activities

Reviewer

- Conference - IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2019
- Workshop - IEEE Detection and Classification of Acoustic Scenes and Events (DCASE) 2018 and 2017
- Research Symposium - LTI Student Research Symposium 2018

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

- Teaching Assistant for Topics in Deep Learning Course - 10707 CMU Machine Learning Department.

Organizer of Task 4 DCASE 2019 Challenge January 2019 - present
Carnegie Mellon University *Mentor*: Prof. Alexander Hauptmann

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

Contributor to 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".

Mentor at The Junior Academy, Global STEM Alliance January 2016 - May 2017

- Mentored a team of young students on wearables challenge implementing an innovative water filtration system.
- Demonstrated a working prototype of filter resulting in safe drinking water, thus **won** the wearables challenge.

Peer Mentoring Programme, NITK Surathkal July 2013 - April 2015

- Teaching assistant for Elements of Electronics and Communication [EC 110] during 2013-14
- Teaching Assistant for Data Structures and Algorithms [EC 232] during 2014-15.

Leadership

- Membership Chair, Indian Graduate Students Association, Carnegie Mellon University - Nov 2017 - present
- President and Vice President of ARM Talkies - Toastmasters Club at ARM - June 2016 - July 2017.
- Joint Convener and Treasurer of Robotics Club 2013 - 2015.
- Vidyut Chairman - IEEE NITK Chapter and Computer Society of India, 2012 - 2015.
- Executive Lead of EFOREA (Engineer's Forum for Entrepreneurship Awareness), and Joint Convener of Electronics Committee NITK 2013 - 2015.