

Ashutosh Chaubey

Education

- 2024 - 2029 **PhD in Computer Science**, *University of Southern California, Los Angeles.*
Advisor - [Prof. Mohammad Soleymani](#)
- 2017 - 2021 **BS in Computer Science**, *Indian Institute of Technology, Roorkee.*
GPA 9.718/10.0 - Third-highest GPA amongst all the graduating students of IIT Roorkee 2021

Publications

- 2024 **ContextIQ : A Multimodal Expert-Based Video Retrieval System for Contextual Advertising.**
Ashutosh Chaubey, Anoubhav Agarwaal, Sartaki Roy, Aayush Agrawal, Susmita Ghose
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) 2025 [\[Paper\]](#)
- 2023 **Meta-Learning Framework for End-to-End Imposter Identification in Unseen Speaker Recognition.**
Ashutosh Chaubey, Sparsh Sinha, Susmita Ghose
IEEE Workshop on Automatic Speech and Understanding (ASRU) 2023 [\[Paper\]](#) [\[Poster\]](#)
- 2022 **Improved Relation Networks for End-to-End Speaker Verification and Identification.**
Ashutosh Chaubey, Sparsh Sinha, Susmita Ghose
Interspeech 2022 [\[Paper\]](#)[\[Poster\]](#)
- 2022 **OPAD : An Optimized Policy-based Active Learning Framework for Document Content Analysis.**
Sumit Shekhar, Bhanu Prakash Reddy Guda, **Ashutosh Chaubey**, Ishan Jindal, Avneet Jain
CVPR 2022 Workshop on Fair, Data Efficient and Trusted Computer Vision [\[Paper\]](#) [\[Patent\]](#)
- 2020 **Universal Adversarial Perturbations : A Survey.**
Ashutosh Chaubey*, Nikhil Agrawal*, Kavya Barnwal, Keerat K. Guliani, Pramod Mehta
arXiv Preprint (**50+** citations) [\[Paper\]](#)
- 2019 **A GAN-based Ensemble Technique for Automatic Evaluation of Machine Synthesized Speech.**
Ashutosh Chaubey*, Jaynil Jaiswal*, Bhimavarapu Sasi Kiran Reddy, Shashank Kashyap, Puneet Kumar, Raman Balasubramanian, Partha Pratim Roy
Asian Conference on Pattern Recognition (ACPR) 2019 [\[Paper\]](#) [\[Poster\]](#)

Research Experience

- Aug 2024 - Present **Graduate Researcher**, Institute for Creative Technologies, University of Southern California.
Advisor - [Prof. Mohammad Soleymani](#)
- Working on diffusion models for pose and expression conditioned video generation.
 - Working on multimodal emotion and facial expression analysis by using VLLMs.
- Apr 2023 - Jul 2024 **Founding Research Engineer**, Anoki Inc..
Advisor - [Dr. Susmita Ghose](#)
- Worked with multimodal models for video retrieval using text, image, and audio. [\[Under review \(WACV 2025\)\]](#)
 - Worked with state-of-the-art text-to-speech and voice cloning technologies to generate AI voice-overs for TV ads.
 - Mentored an intern on emotion recognition in movie scenes using audio and multi-modal approaches.
- Jul 2021 - Mar 2023 **Data Scientist**, LG Ad Solutions (formerly Alphonso Inc.).
Advisor - [Dr. Susmita Ghose](#)
- Worked with state-of-the-art speaker verification models such as RawNet and ECAPA-TDNN to exhibit their generalizability on near-/far-field audios, different languages, and under noisy environments.
 - Proposed a novel relation network-based pipeline for end-to-end speaker recognition, improving the baseline by up to 12% relatively. [\[Paper \(Interspeech 2022\)\]](#) [\[Paper \(ASRU 2023\)\]](#)
 - Worked on neural audio fingerprinting for automatic content recognition on TV devices using contrastive learning.
- May 2020 - Jul 2020 **Research Intern**, Big-data Experience Lab, Adobe Research.
Advisor - [Dr. Sumit Shekhar](#)
- Used deep Q-learning to learn an optimal acquisition function for active learning by modeling the active learning cycle as a Markov Decision Process. [\[Paper \(CVPR 2022 Workshops\)\]](#)
 - Reduced the annotation effort by using a weak learning setting where the annotator just has to verify the current model predictions on acquired samples. [\[US Patent App. 17/170,307\]](#)
- Jan 2019 - Mar 2020 **Research Intern**, Indian Institute of Science (IISc.), Bengaluru | Indian Institute of Technology, Roorkee.
Advisors - [Prof. R Venkatesh Babu](#) | [Prof. R. Balasubramanian](#)
- Worked on multi-person human pose prediction using synthetic dual person dataset to mitigate data limitations.
 - Worked on automatic evaluation of TTS and proposed a GAN-decoder-based scoring mechanism. [\[Paper \(ACPR 2019\)\]](#)
 - Experimented with state-of-the-art universal adversarial perturbation techniques (attacks and defenses) and wrote a survey paper with over 50 citations. [\[Survey Paper\]](#)

Skills

Coding Languages - Python [Advanced], C++ [Intermediate]
Frameworks/Libraries - PyTorch, NumPy, Pandas, Transformers
Tools - VSCode, Git, Anaconda, Docker