Laxmi Pandey

Email: **laxmip@meta.com**
Website | **GitHub** | **Google Scholar** | | **Linkedin**

SUMMARY

Experienced researcher in machine learning and NLP, specializing in speech recognition. Proven track record of high-impact publications and interdisciplinary collaborations.

RESEARCH INTEREST

Audio-Visual-Textual Speech Recognition, Human-Computer Interaction, Machine Learning

WORK EXPERIENCE (10+ years)

1. Research Scientist at Meta AI - California, USA, (Sept 2022 - Present)
   - Speech and natural language processing.
2. Research Intern at Facebook AI - California, USA, (May 2021 - Sept 2021)
   - Inverse text normalization and domain adaptation for speech processing.
3. Graduate Teaching Assistant at UC Merced - California, USA, (Aug 2018 - May 2022)
   - CSE-005: Introduction to Computer Applications.
   - CSE 155: Introduction to Human-Computer Interaction.
4. Worked as a Research Engineer at Cogknit Semantics - IISC Bangalore, India, (Oct 2017 - Mar 2018)
   - Non-speech sequence Labeling, Speaker Verification, Out-of-vocabulary Words Detection Silence detection and Speech Denoising.
   - Developed an automatic speech recognition models for Indian languages using Kaldi toolkit.
5. Worked as a Research Associate at MIPS Lab, IIT Kanpur, India (July 2012 - July 2017)
   - Developed an automatic speech recognition systems using sphinx3 toolkit .
   - Developed a voice enabled radio news access system using spoken query with database handling using PHP and MySQL.
   - Developed a speech based search engine to search a spoken keyword with in an audio archive for the retrieval of keyword related information.
   - Developed a tool for real time goal detection tool for automatic generation of cheer sound effects at every goal hit in the soccer match.
   - Designed audio and video analyzing tools, scoring the similarity and difference between their contents.

ACADEMIC DETAILS

Examination	Institute	Year	GPA
Ph.D. (Computer Science)	University of California, Merced	2018 - 2022	3.7
M.S. (Electrical Engineering)	Indian Institute of Technology, Kanpur	2015 - 2017	3.5
B.Tech (Electrical Engineering)	Uttar Pradesh Technical University, Lucknow	2008 - 2012	3.5

PUBLICATIONS

1. Minh Tran, Yutong Pang, Debjyoti Paul, Laxmi Pandey, Kevin Jiang, Jinxi Guo, Ke Li, Shun Zhang, Xuedong Zhang, Xin Lei **"R2S: Representation Learning for Speech Recognition with Guidance from Synthetic Data"** [In review].
2. Minh Tran, Yutong Pang, Debjyoti Paul, Laxmi Pandey, Kevin Jiang, Jinxi Guo, Ke Li, Shun Zhang, Xuedong Zhang, Xin Lei **"A Domain Adaptation Framework for Speech Recognition Systems with Only Synthetic data"**, IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2025.
3. Laxmi Pandey, Ke Li, Jinxi Guo, Debjyoti Paul, Arthur Guo, Jay Mahadeokar, Xuedong Zhang **"Towards Scalable Efficient On-device ASR with Transfer Learning"**, arXiv preprint arXiv:2407.16664, 2024.
4. Laxmi Pandey, Ahmed Sabbir Arif **"MELDER: The Design and Evaluation of a Real-time Silent Speech Recognizer for Mobile Devices"**, CHI Conference on Human Factors in Computing Systems, CHI 2024.

5. Laxmi Pandey, Debjyoti Paul, Pooja Chitkara, Yutong Pang, Xuedong Zhang, Kjell Schubert, Mark Chou, Shu Liu, Yatharth Saraf **"Improving Data Driven Inverse Text Normalization using Data Augmentation"**, arXiv preprint arXiv:2207.09674, 2022.

6. Laxmi Pandey, Ahmed Sabbir Arif **"Design and Evaluation of a Silent Speech-Based Selection Method for Eye-Gaze Pointing"**, ACM Interactive Surfaces and Spaces Conference, ISS 2022.

7. Laxmi Pandey, Ahmed Sabbir Arif **"Effects of Speaking Rate on Speech and Silent Speech Recognition"**, CHI Conference on Human Factors in Computing Systems Extended Abstracts, CHI 2022. *[Poster]*

8. Laxmi Pandey, Ahmed Sabbir Arif **"Silent Speech and Emotion Recognition from Vocal Tract Shape Dynamics in Real-Time MRI"**, Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters, SIGGRAPH 2021. *[Poster]*

9. Laxmi Pandey, Ahmed Sabbir Arif **"LipType: A Silent Speech Recognizer Augmented with an Independent Repair Model"**, CHI Conference on Human Factors in Computing Systems, CHI 2021.

10. Laxmi Pandey, Khalad Hasan, Ahmed Sabbir Arif **"Acceptability of Speech and Silent Speech Input Methods in Private and Public"**, CHI Conference on Human Factors in Computing Systems, CHI 2021.

11. Laxmi Pandey, Ahmed Sabbir Arif **"Enabling Text Translation Using the Suggestion Bar of a Virtual Keyboard"**, IEEE International Conference on Systems, Man, and Cybernetics, SMC 2020.

12. Laxmi Pandey, Azar Alizadeh, Ahmed Sabbir Arif **"Enabling Predictive Number Entry and Editing on Touchscreen-Based Mobile Devices"**, ACM Conference on Human Information Interaction and Retrieval, CHIIR 2020.

13. Steven J. Castellucci, I. Scott MacKenzie, Mudit Misra, Laxmi Pandey, Ahmed Sabbir Arif **"TiltWriter: Design and Evaluation of a No-touch Tilt-based Text Entry Method for Handheld Devices"**, ACM International Conference on Mobile and Ubiquitous Multimedia, MUM 2019.

14. Laxmi Pandey, Ahmed Sabbir Arif **"Context-Sensitive App Prediction on the Suggestion Bar of a Mobile Keyboard"**, ACM International Conference on Mobile and Ubiquitous Multimedia, MUM 2019. *[Poster]*

15. Aditya Raikar , Saurya Basu, Laxmi Pandey, Rajesh M. Hegde **"Multi-Channel Joint speech dereverberation and denoising using Deep Priors"**, IEEE India Council International Conference, INDICON 2018.

16. Laxmi Pandey, Karan Nathwani **"LSTM Based Attentive Fusion of Spectral and Prosodic Information for Keyword Spotting"**, Interspeech 2018.

17. Laxmi Pandey, Anurendra Kumar, Vinay Namboodiri **"Monaural Audio Source Separation Using Variational Autoencoders"** , Interspeech 2018.

18. Laxmi Pandey, Nitish Divakar, Krishna D N, Anuroop Iyengar **"Deep Clean: GPU Powered Speech Denoising using Adversarial Learning"** , NVIDIA's GPU Technology Conference, GTC 2018. *[Poster]*

19. Laxmi Pandey, Rajesh M. Hegde **"Keyword Spotting in Continuous Speech using Spectral and Prosodic Information Fusion"**, Circuits Systems and Signal Processing, CSSP Springer, 2018. *[Journal]*

20. Laxmi Pandey, Kuldeep Chaudhary, Rajesh M. Hegde **"Fusion of Spectral and Prosodic Information using Combined Error Optimization for Keyword Spotting"**, IEEE National Conference on Communications, NCC 2017.

21. L. Pandey, K. Nathwani, S. Kaur, I. Hussain, R. Pathak, G. Singh, S. Tiwari, Rajesh M. Hegde **"Domain Specific Audio Indexing Using Linguistic Information"**, IEEE Symposium on Signal Processing and Information Technology, ISSPIT 2014.

THESIS

1. PhD Thesis, UC Merced, [Advisor: Prof. Ahmed Sabbir Arif] (Aug 2018 - Aug 2022)
   Lip Reading as an Active Mode of Interaction with Computer Systems
   Developed image-based silent speech recognition techniques, addressing user perceptions, error tolerance, and privacy concerns, while optimizing models for improved speed, accuracy, and hands-free usability in computer systems.

2. MS Thesis, IIT Kanpur, [Advisor: Prof. Rajesh M Hegde] (Aug 2015 - Aug 2017)
   Fusion of Spectral and Prosodic Information for Keyword Spotting in Continuous Speech
   Developed novel spectral-prosodic fusion methods for improved keyword spotting and syllable prediction in continuous speech using deep learning techniques.

SCHOLARSHIPS

1. Fred and Mitzie Ruiz Fellowship, UC Merced, 2021, $1000.
2. Hatano Cognitive Development Research Fellowship, UC Merced, 2020, $1000.
3. SIGIR Student Travel Grant, 2020, $1000.
4. ACM-W Scholarship Award for Attendance at Research Conferences, 2020, $600.
5. Summer EECS Bobcat Fellowship at UC Merced, 2019, $4200.
6. International Speech Communication Association (ISCA) Travel Grant, 2018, €650.

PROFESSIONAL MEMBERSHIPS AND SERVICES

1. Member, Institute of Electrical and Electronics Engineers (IEEE).
2. Member, Association for Computing Machinery (ACM).
3. Reviewed papers for CHI 2021, GI 2021, ISWC 2020, IUI 2019-2021.
4. Served as a program committee member at Graphics Interface Conference (GI) 2021.
5. Session co-chair at IEEE Int. Conference on Systems, Man, and Cybernetics (SMC) 2020.

PRESENTATIONS AND WORKSHOPS

1. Presented a work *"MELDER: The Design and Evaluation of a Real-time Silent Speech Recognizer for Mobile Devices"*, at CHI Conference on Human Factors in Computing Systems (CHI) 2024, Hawaii.
2. Presented a work *"Silent Speech and Emotion Recognition from Vocal Tract Shape Dynamics in Real-Time MRI"*, at Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH) 2021, Virtual.
3. Presented a work *"LipType: A Silent Speech Recognizer Augmented with an Independent Repair Model"*, at CHI Conference on Human Factors in Computing Systems (CHI) 2021, Yokohama, Japan.
4. Presented a work *"Acceptability of Speech and Silent Speech Input Methods in Private and Public"*, at CHI Conference on Human Factors in Computing Systems (CHI) 2021, Yokohama, Japan.
5. Presented a work *"Enabling Text Translation Using the Suggestion Bar of a Virtual Key-board"*, at Conference on Systems, Man, and Cybernetics (SMC) 2020, Toronto, Canada.
6. Presented a work *"Enabling Predictive Number Entry and Editing on Touchscreen-Based Mobile Devices"*, at Conference on Human Information Interaction and Retrieval (CHIIR) 2020, Vancouver, Canada.
7. Presented a work *"Novel Eyes-Free Text Entry Techniques for Mobile Devices"*, at the 2nd Annual Fall Symposium: Branches of Cognitive Science, Cognitive Science Student Association, University of California, Merced, USA,
8. Participated in SIGCHI Summer School on *Research Methods and Approaches to Text Entry and Other Interaction Techniques*, May 21 - 25, 2018, IIT Bombay, India.

RESEARCH AND ACADEMIC PROJECTS

1. **Design and development of lip reading assistive technology for human-computer interaction**, UC Merced, [Advisor: Prof. Ahmed Arif]: Developing a technique that enables users to enter text on mobile devices using silent speech (or lip reading). It captures the user's face using the front-camera, then segments silently speaking lip sequences to classify them into text using a vision-based approach.

2. **Fusion of Spectral and Prosodic Information for Keyword Spotting in Continuous Speech** *(July 2016 - July 2017), Master's Thesis, IIT Kanpur [Advisor: Prof. Rajesh Hegde]*: Developed a methodology for hierarchical fusion of spectral and prosodic information at feature-level and model-level to improve Automatic Speech Recognition (ASR) performance using Hidden Markov Modeling (HMM).

3. **Build and demonstrate a real time continuous speech recognition system for Hindi language** *(Jan 2016 - Apr 2016), Course project for EE627, Speech Signal processing, IIT Kanpur, [Advisor: Prof. Rajesh Hegde]*: Developed and demonstrated a system for spotting keywords in continuous speech utterances using word boundary identification and language modeling.

4. **Development of Prosodically Guided Phonetic Engine for Hindi** *(July 2012 - Dec 2014), IIT Kanpur, [Advisor: Prof. Rajesh Hegde]*: Developed a phonetic engine capable of performing automatic phonetic transcription, syllabification, pitch marking and break marking in speech utterances.

5. **License Plate Detection and Recognition in Complex Scenes Using Mathematical Morphology and Support Vector Machines** *(Aug 2015 - Nov 2015), Term Paper for EE604, Image Signal processing* : Developed a highly reliable license plate detection and recognition approach using mathematical morphology and support vector machines (SVM).

INDUSTRIAL COLLABORATION

1. **Real Time Goal Detection in Soccer Match using Hidden Markov Model** *(Feb 2016 - Oct 2016), LG Soft India*: Developed a real time goal detection tool capable of generating an automatic cheering effects at every goal hits in the soccer match video using markov modeling.

2. **Development of SMS Compression Techniques for Indian Languages** *(Sept 2015 - Jan 2016), CDAC India*: Developed a standard loss less compression algorithm that can allow more number of characters in SMS. The encoding schemes allow the transmission of around 160 characters for pure Hindi, and multilingual text.

3. **Video Analyzer using Command Line Video Quality Metric** *(March 2015 - Feb 2016), LG Soft India*: Designed and developed a video analyzer tool for scoring the similarity or difference between two videos on the basis of dropped frames and broken macro blocks.

4. **Audio Analyzer using Dynamic Time Warping Algorithm** *(Jan 2014 - Feb 2015), LG Soft India*: Designed and developed an audio analyzer tool for scoring the similarity in audio effects between reference and test device using using Dynamic Time Warping (DTW).

5. **Digital Mandi for the Indian Kissan** *(July 2012 - Mar 2013), BSNL India*: Digital Mandi Application for Indian Kisan is a unique web and cell phone based multimodal agriculture commodity price retrieval system. It has been developed by BITCOE (BSNL IIT-Kanpur Centre of Excellence) at IIT Kanpur. Currently, the application has been deployed on the BSNL national network in Orissa and Haryana. The application service was formally inaugurated by Honorable Minister of Communications and IT on 29th August 2011 and is operational since then.

RELEVANT COURSES

Statistical Signal Processing, Speech Signal Processing, Image Signal Processing, Introduction to Signal Analysis, Digital Signal Processing, Human-Computer Interaction, Computer Animation, Photonics, Networks and Switching.

REFERENCES

1. Prof. Ahmed Arif **[Ph.D. Advisor]** - *asarif@ucmerced.edu*
   Professor, Electrical and Computer Science Engineering, UC Merced, USA

2. Prof. Rajesh M. Hegde **[M.S. Advisor]** - *rhegde@iitk.ac.in*
   Professor, Electrical Engineering, IIT Kanpur, India

3. Prof. Marcelo Kallmann **[Course Instructor]** - *mkallmann@ucmerced.edu*
   Professor, Electrical and Computer Science Engineering, UC Merced, USA

4. Prof. Vinay Namboodri **[Course Instructor]** - *vinaypn@iitk.ac.in*
   Professor, Computer Science and Engineering, IIT Kanpur, India