Karan Taneja

Hostel 18, IIT Bombay, Mumbai – 400076

karantaneja@iitb.ac.in • krntneja.github.io

Research Interests: Machine Learning, Computer Vision, Language and Speech Processing, and intersections thereof.

Education

- **Indian Institute of Technology Bombay** — **Mumbai, India**
  Dual Degree (B.Tech. + M.Tech.), Electrical Engineering — *2015–2020*
  - **CPI** (Overall): 9.57 /10, 3^{rd} among 78 students
  - **Specialization**: Signal Processing and Communication
  - **Minor**: Computer Science and Engineering

Publications

- **Exploiting Monolingual Speech Corpora for Code-mixed Speech Recognition**
  Karan Taneja, Satarupa Guha, Preethi Jyothi, Basil Abraham
  Proceedings of Interspeech 2019, 15-19 September 2019, Graz, Austria
- **A Deep CNN Framework to Reconstruct k-t-Undersampled Resting-fMRI**
  Karan Taneja, Prachi H. Kulkarni, S. N. Merchant, Suyash P. Awate
  Under Review for IEEE International Conference on Image Processing (ICIP) 2020, Abu Dhabi, UAE
- **A Bayesian Deep CNN Framework for Reconstructing k-t-Undersampled Resting-fMRI**
  Karan Taneja, Prachi H. Kulkarni, S. N. Merchant, Suyash P. Awate
  Under Review for IEEE JSTSP Special Issue on Domain Enriched Learning for Medical Imaging

Achievements and Awards

- **Institute Academic Prize (IAP)** — **IIT Bombay**
  Awarded with IAP for ranking 1^{st} among 78 students in EE department for year 2018-19 — *2018-2019*
- **Exceptional Performance Grade (AP)** — **IIT Bombay**
  AP grade in 3 courses: Image Processing, Control Systems, Computer Programming and Utilization — *2015-2020*
- **International Speech Communication Association (ISCA) Travel Grant** — **ISCA–Interspeech'19**
  Awarded with ISCA Travel Grant for attending Interspeech 2019, Graz, Austria. — *2019*
- **Kishore Vaigyanik Protsahan Yojana (KVPY) Fellow** — **Government of India**
  Awarded with KVPY fellowship, ranked 437 among 100 thousand candidates nationwide — *2014*
- **National Talent Search Scheme (NTSS) Fellow** — **Government of India**
  Awarded with NTSE fellowship, in top 1000 students among 300 thousand candidates nationwide — *2013*

Research Projects

- **Fast R-fMRI k,t-undersampled Acquisition and Reconstruction Using a CNN Framework**
  Masters Thesis, IIT Bombay, Guide: Prof. Suyash Awate, Prof. Shabbir Merchant — *Jan 2019–Present*
  - Implemented a two-stage CNN framework model for R-fMRI reconstruction from under-sampled k,t-space acquisition to optimize for the structural similarity of the functional connectivity maps of the human brain.
  - Generalized the previously used GRAPPA scheme in the first stage to use a deep CNN network for the k-space filling and another network, in the second stage, for time upsampling and image enhancement.
  - Showed that our method outperforms previous methods that use dictionary priors and sparsity constraints.
  - Extended our framework to estimate uncertainty over the mean output to get more insight about the quality of functional connectivity maps and showed improved performance over previous training objectives.

- **Code-mixed Automatic Speech Recognition (CM-ASR) Using Monolingual Corpora**
  Research Intern, Microsoft IDC, Hyderabad, India, Guide: Prof. Preethi Jyothi — *Oct 2018–Feb 2020*
  - Proposed two linguistically motivated algorithms to create synthetic CM speech while preserving span length distributions and phone transition probability distributions at switch points.
  - Investigated the effect of using synthetic CM speech to train acoustic models, used transcripts from the synthetic data to train language models and examined their effect on ASR performance in real Hindi-English CM speech.
  - Showed improvements over naive linguistically-unmotivated methods using monolingual corpora for CM-ASR.
  - Working on voice conversion for introducing speaker variability in utterances from CM text-to-speech systems.

- **Spherical CNNs for Human Detection Task in 360-degree Images**
  Research Intern, Sony Semiconductor Solutions, Kanagawa, Japan *May 2018–Jul 2018*
  - Extended the work on spherical CNNs to propose much lighter convolution operations on sphere S^2 and rotation group $SO(3)$ by factorization into depth-wise and point-wise spherical convolution.
  - Enabled the use of deeper architectures such as ResNet, DenseNet and others which were previously infeasible with spherical CNNs and showed improved performance in human-detection on a Theta 360-degree image dataset.
  - Published an internal technical report titled *Spherical CNNs for Human Detection Task in 360-degree Images*.

- **Attention-based and Segmental Models for Speech Recognition**
  R&D Project, IIT Bombay, Guide: Prof. Preethi Jyothi, Prof. Sunita Sarawagi *Jan 2018–May 2018*
  - Proposed *Segmentation via Attention model* that maximizes joint conditional probability of output sequence of an attention model by introducing a latent segmentation for phoneme recognition on the TIMIT dataset.
  - Devised an algorithm to marginalize over segmentations by interpreting attention probabilities over encoder states as conditional probabilities of segment boundaries and doing a one-step look-ahead for end-of-segment boundary.
  - Enabled tractability by using a top-k beam over segmentation-lattice and implemented the model in TensorFlow.

- **Patient-motion Detection in MRI Scans Using Deep Neural Networks**
  Research Intern, Philips Innovation Campus, Bangalore, India *May 2017–Jul 2017*
  - Implemented a 3D CNN classifier and cascaded 2D CNN-RNN classifier where RNN takes in the features extracted by CNN over slices of MRI scans for detecting the level of patient motion detection.
  - Implemented several feature-based classifiers that used wavelet analysis, spectral analysis, edge information and histogram of gradient information and performed feature selection for patient motion detection.
  - Showed empirically that CNN-RNN model outperforms 3D-CNN model which outperforms feature-based classifiers.

Other Select Projects

- **Semantic Image Inpainting with Deep Generative Models**
  Course: Medical Image Computing, Spring 2018-19
  - Implemented a Generative Adversarial Network (GAN) based technique for image inpainting with backpropagation-to-input and Poisson blending in PyTorch, inspired from a paper with same title by Yeh et al., in CVPR 2017.
  - Extended the ideas to Variational Auto-encoder (VAE) based model and showed that VAEs perform better than GANs for MRI image inpainting owing to sparsity in modes of variation in MRI.
- **Compressed Sensing and Dictionary Learning for Video Capturing**
  Course: Recent Topics in Analytical Signal Processing, Spring 2018-19
  - Implemented a method for video reconstruction using coded sampling, K-SVD algorithm for dictionary learning and Orthogonal Matching Pursuit (OMP) algorithm for sparse reconstruction.
  - Proposed different sampling strategies to improve performance without hardware constraints and experimented with variations in sampling schemes, dictionary parameters, etc. to quantify the robustness of different modules.
- **Experiments with Differentiable Neural Computer (DNC)**
  Course: Advanced Machine Learning, Spring 2018-19
  - Experimented with the DNC on top-k sorting, shortest-path in a graph and connectedness in a graph tasks.
  - Evaluated the performance of DNC with different parameter settings and tasks to gain insights about its capacity.
- **Voice Conversion using Generative Adversarial Networks (GANs)**
  Course: Automatic Speech Recognition, Autumn 2017-18
  - Replicated the results of voice conversion using Variational Autoencoding Wasserstein GAN by Hsu et al. 2017.
  - Extended the VAE-WGAN method by conditioning the GAN on sentence embeddings as content representation.
- **Batch Arm Pulls for Stochastic Multi-Armed Bandits**
  Course: Advanced Concentration Inequalities, Autumn 2019-20
  - Designed algorithms for the setting of jointly pulling a batch of K arms among N arms with stochastic rewards.
  - Proposed a novel algorithm for simple regret minimization with approximately N/K factor reduction in sample complexity over state-of-the-art and empirically confirmed its validity through simulations.
- **The Music Box – Modelling, Rendering, and Animation**
  Course: Computer Graphics, Autumn 2018-19
  - Modelled a music box with a humanoid and a giraffe kept in a realistic room with varying lighting and textures.
  - Rendered animation of dancing characters using interpolation, and moving camera on user-specified Bezier curve.
- **Vector Representation of Words Using Neural Networks**
  Winter Project: Center for Indian Lanuage Technology, Winter 2016-17
  - Implemented continuous bag of words Word2Vec model for an English corpus with 1.7b words in TensorFlow and experimented with network architecture, training algorithms and related hyper-parameters.

- **Music Genre Detection using Machine Learning**
  Foundations of Machine Learning, Spring 2016-17
  - Used Music Information Retrieval (MIR) toolbox for feature extraction from audio files of five genres and compared the performance of several machine learning algorithms and achieved 78% accuracy on the GTZan dataset.
  - Visualized extracted data from music files using t-Distributed Stochastic Neighbour Embedding (t-SNE).
- **Isolated Words Speech Recognition System**
  Course: Speech Processing, Autumn 2018-19
  - Obtained features for the utterances using Mel frequency cepstral coefficients (MFCCs) from scratch.
  - Implemented vector quantization codebook, clustering, and dynamic time warping for speech recognition.
- **Adaptive Echo Cancellation (AEC)**
  Course: Digital Signal Processing, Spring 2017-18
  - Implemented Least Mean Square (LMS), Normalized LMS and Recursive Least Square algorithms for AEC.
  - Comparatively analyzed computation cost, convergence time and stability of the algorithms.

Teaching

- **Teaching Assistant**

Introduction to Machine Learning (CS 419)	Prof. Sunita Sarawagi	Autumn 2018, IIT Bombay
Data Analysis and Interpretation (EE 223)	Prof. Shabbir Merchant	Summer 2019, IIT Bombay
Data Analysis and Interpretation (EE 223)	Prof. Prasanna Chaporkar	Autumn 2019, IIT Bombay

- **Bootcamps**

Coding Bootcamp for Interview Preparation	Undergraduate Academic Council	Autumn 2018
Web Development and Programming Camp	Camp K–12, Mumbai	Summer 2016

Relevant Coursework

Machine Learning	Machine Learning, Advanced Machine Learning, Reinforcement Learning, Speech Processing, Speech Recognition, Image Processing, Advanced Image Processing, Medical Imaging
Signal Processing	Digital Signal Processing, Analytical Signal Processing, Estimation & Identification, Probability & Random Processes, Markov Chains & Queuing Systems, Data Analysis
Computer Science and Mathematics	Computer Graphics, Data Structures and Algorithms, Matrix Computations, Control Systems, Optimization, Game Theory and Applications, Computer Networks, Operating Systems, Microprocessors, Discrete Structures, Error Correcting Codes, Random Graphs

Technical Aptitude

Programming	Python, C/C++, Java, MATLAB/Octave, VHDL, Web development
Packages	SciPy stack, PyTorch, TensorFlow, OpenCV, OpenGL
Other tools	Bash, Vim, Jekyll, LaTeX, Eagle, Ngspice, Quartus, SolidWorks

Events & Activities

- Attended, volunteered and presented a paper at *Interspeech 2019, Graz, Austria* organized by ISCA and TU Graz.
- Attended *Symposium on Recent Advances in Speech Prosody Research*, 2018 at IIT Bombay.
- Attended *National Science (Vijyoshi) Science Camp 2014*, organized by IISc Bangalore and IISER Kolkata.
- Won *Honorable Mention Prize* in an essay competition organized by Vigilance Department (IIT Bombay) during *Vigilance Awareness Week 2015* on moral values and responsibilities.
- Awarded with *Technical Color Prize*, Hostel 2, IIT Bombay for notable contribution in inter-hostel competitions.
- Volunteered for social initiatives under *National Service Scheme* at IIT Bombay, like health camps, cloth collection drive, assisting construction workers and awareness initiatives.
- Mentored three first-year undergraduate students in the *Summer of Science* study-project organized by *Math and Physics Club, IIT Bombay* on introductory topics in Machine Learning and Artificial Intelligence.
- Speaker for *Reflections Session* by Web and Coding Club, IIT Bombay to share my research experience at IITB.

References

Prof. Suyash Awate
Computer Science and Engineering
IIT Bombay

Prof. Preethi Jyothi
Computer Science and Engineering
IIT Bombay