BHAVAN JASANI

bjasani@alumni.cmu.edu | https://bhavanj.github.io/ | (412) 618 – 9200
Google scholar | www.linkedin.com/in/bhavan-jasani | https://github.com/bhavanj

EDUCATION

Carnegie Mellon University, School of Computer Science — Pittsburgh, PA
M.S. in Robotics (Research-based) | *CGPA: 3.95/4.33* — August 2017 – August 2019

***Courses**:* Computer Vision, Machine Learning, Visual Learning & Recognition, Deep Reinforcement Learning, Math Fundamentals For Robotics

Birla Institute of Technology & Science (BITS), Pilani – K.K. Birla Goa Campus — Goa, India
Dual degree: M.Sc. Physics + B.E. Electrical & Electronics Engineering | CGPA: 9.32/10 — August 2011 – August 2016

SKILLS

Python, C/C++, PyTorch, TensorFlow, Scikit, OpenCV, MATLAB, AWS, ELAN, Blender

RESEARCH EXPERIENCE

Amazon Web Services (AWS) AI — Palo Alto, CA
Applied Scientist *(AWS Rekognition and Video team)* — September 2019 – present

- Product focused research at the intersection of Computer Vision and Natural Language Processing
- Developing Visual Question Answering based models for extracting structured data from scanned documents, that goes beyond optical character recognition (OCR) to understand the semantic relationship between different entities in a document

Carnegie Mellon University, Robotics Institute, School of Computer Science — Pittsburgh, PA
Research Assistant (*under Prof. Jeffrey Cohn and Dr. Laszlo Jeni*) — October 2017 – August 2019

- Built multi-modal (video + audio) human emotion recognition system on noisy real-time annotated data which has a variable temporal lag between the video segments and the corresponding annotated emotion labels
- Extracted 3D facial landmarks, head pose, body pose, and facial action units. Used classical models which take time series of these features and deep learning models based on RNN's and CNN's
- Discovered and quantified the influence of head movements, facial expressions and body pose on the behavior of people in interpersonal conversations [details]

Nanyang Technological University, School of Computer Science & Engineering — Singapore
Research Staff — August 2016 – May 2017

- Implemented parallel and hardware efficient (requires 40% fewer hardware resources) object detection algorithm for real-time pedestrian detection on an embedded system (Altera FPGA and Terasic camera)

Internship (*part of bachelor's thesis)* — January 2016 – July 2016

- Developed a novel approach for hardware acceleration by finding optimal bit-width, results in 45% reduction of bit-width of Harris Corner Detector with just 0.57% decrease in accuracy, and runs at high fps (335) on HD videos [details]

SELECTED PUBLICATIONS

- ***B Jasani**, R Girdhar, D Ramanan,* "Are we asking the right questions in MovieQA?", IEEE International Conference on Computer Vision (ICCV) Workshops, 2019 **[spotlight oral presentation] [project page]**
- ***B Jasani**, A Mazagonwalla*, "Skeleton based Zero-Shot Action Recognition in Joint Pose-Language Semantic Space", arXiv:1911.11344, 2019
- *Y Patel, K Chitta,* ***B Jasani***, "Learning sampling policies for domain adaptation of image classifiers", arXiv:1805.07641
- ***B Jasani**, SK Lam, PK Meher, M Wu*, "Threshold-guided design and optimization for Harris corner detector architecture", IEEE Transactions on Circuits and Systems for Video Technology, 2017
- *SK Lam, TC Lim, M Wu, B Cao,* ***B Jasani***, "Data-path unrolling with logic folding for area-time-efficient FPGA-based FAST corner detector", Journal of Real-Time Image Processing, 2019

ACADEMIC PROJECTS

Deleting 3D Objects in Augmented Reality using ORB-SLAM March – May 2019

- Combined 2D image segmentation from Mask RCNN with 3D point cloud generated by RGBD-SLAM and image inpainting for removing objects from a scene [details]

Discovering biases in Visual Question Answering *(under* ***Prof. Deva Ramanan****)* May – Dec 2018

- Discovered and quantified language biases in video-based visual question answering datasets and how deep neural networks learn to cheat
- Exploited these biases to develop a simple model which only looks at the question, and yet achieves state of the art accuracy on leader board of MovieQA dataset
- Proposed adversarial ways to mitigate these language biases to make visual information useful [**project page - ICCV 2019**]

Zero-shot Learning for Action Recognition April – June 2018

- Built a zero-shot body pose based action recognition system (in PyTorch), which learns joint semantics between word embeddings of class labels and the video features extracted from a spatio-temporal graph convolutional network (STGCN)
- STGCN takes time series of body pose of the action performer as the input & learns pose representation [details]

Domain Adaptation for Image Classification March – May 2018

- Took the predictions of a source domain trained network on target domain data as noisy labels
- Implemented an RL agent which learns a policy to sample from this data for training a new classifier for target-dataset, to maximize the classification accuracy of a small annotated partition (that acts as a. reward) of the target-dataset [details]

Adversarial Image Generation using GAN's November 2017

- Implemented a generative adversarial network (GAN) in TensorFlow for generating adversarial images which can fool a neural network (black box attack) for CIFAR 10 dataset [details]

Distracted Driver Detection May – July 2016

- Implemented a fast RCNN based neural network for detecting and classifying the state of the driver from images taken from car's dashboard camera facing the driver, for State Farm Distracted Driver Detection competition hosted on Kaggle

AWARDS & ACHIEVEMENTS

- **Research Assistantship,** Spring 2018 to Spring 2019, Carnegie Mellon University
- **DAAD WISE, 2014 scholarship**, awarded by German Academic Exchange Service for a summer research internship at a German research institution
- **Innovation in Science Pursuit for Inspired Research (INSPIRE), 2011 – 2016 fellowship,** from Department Of Science And Technology, Government of India, awarded to bright students majoring in sciences