Tanya Marwah

✉ tmarwah@cs.cmu.edu
⌂ https://tm157.github.io

Education

2017–   M.S. in Robotics,
Carnegie Mellon University, GPA: 4.08*
Awards: Siebel Scholar 2019.

2013–2017   B.Tech with Honors in Electrical Engineering,
Indian Institute of Technology, Hyderabad, GPA: 9.11, Honors GPA: 10.00.

Research Interests

Computer Vision and Machine Learning: Scene Understanding, Generative Models, Learning with Limited Supervision

Artificial Intelligence and Robotics: Reinforcement Learning, Inverse Reinforcement Learning, Interactive Perception

Publications

○ Learning Textured Articulated Human from Single Image - **Tanya Marwah**, Ye Yuan, Kris M. Kitani, *IEEE Conference on Computer Vision and Pattern Recognition* **(CVPR)**, *2019, Long Beach, CA* (Under Review)

○ Attentive Semantic Video Generation using Captions - **Tanya Marwah**, Gaurav Mittal, Vineeth N. Balasubramanian, *IEEE International Conference on Computer Vision* **(ICCV)**, *2017, Venice, Italy*

○ Sync-DRAW: Automatic Video Generation using Deep Recurrent Attentive Architectures - **Tanya Marwah**, Gaurav Mittal, Vineeth N. Balasubramanian, *ACM, Multimedia* **(ACM-MM)**, *2017, Mountain View California*

○ Sentiment dynamics in social media news channels - Rakshita Nagalla, Nagendra Kumar, **Tanya Marwah**, Manish Singh, *Online Social Networks and Media, Journal, 2018*, (Journal Paper)

○ On the Relevance of Very Deep Networks for Diabetic Retinopathy Diagnostics - **Tanya Marwah**, Akilesh B, Vineeth N. Balasubramanian, Kumar Rajamani *Applications of Cognitive Computing Systems and IBM Watson, 2016, Bengaluru, India*

Experience

2017–   **Research Assistant**, *Carnegie Mellon University*, Pittsburgh, PA.
Advisor: Dr. Kris M. Kitani
Generating realistic human animations from single images; End-to-End 3D texture and pose estimation using SMPL models; Learning human motion policies using Imitation Learning and Reinforcement Learning.

2016–2016   **Research Intern**, *Robert Bosch Engineering and Business Solutions Limited*, Bengaluru, India.
Business Unit Development (BUD)
Advisor: Kumar Rajamani
Worked on early detection of Diabetic Retinopathy in retinal images; Ensemble of Residual Neural Networks trained on dataset acquired by Bosch in India and Germany; Integrated trained network with Bosch servers for field testing and clinical trials.

2015–2015   **Research Intern**, *Indian Institute of Technology*, New Delhi, India.
Advisor: Dr. K.K. Biswas
Action Recognition in temporally untrimmed videos; Classified the videos into 101 action classes; THUMOS challenge 2015.

2014–2014   **Intern**, *Broadcom*, Hyderabad, India.
Advisor: Dr. Ramesh Yerraballi (UT-Austin)
Teamed up with students from UT Austin for "Embedded Systems and Industry Experience" and worked on a Network Packet Analyzer using Raspberry Pi; Library built in C/C++.

Projects

Generating Realistic Human Animations from Single Images

Dr. Kris M. Kitani [Report]

Generated 3D human models that capture the visual apprearance as well as the structure of a person from a single monocular image. Predicted the texture, pose and shape of SMPL model which are combined using a differentiable renderer to generate 3D human models. Animate the 3D models using pose conditioned motion policies that conform to the physical dynamics of human motion.

Using Imitation Learning for Learning Culinary Skills

Dr. Oliver Kroemer [Report]

Graph-based representations to learn control policies for basic cooking skills such as cutting, drizzling, pouring. Used a 7-DOF Franka Arm for imitation learning, demonstrations for various skills were collected using kinesthetic teaching. Used parameterized Dynamic Motor Primitives (DMPs) to represent the low level policy.

Incremental Image Generation Using Scene Graphs

Dr. Louis-Philippe Morency [Report]

Composite generation using scene-graphs. Use Graph Convolution Networks (GCN) to model inter-object relationship in a given image. Use a cascade refinement network (CRN) type decoder to generate an image. Use of attention mechanism along with residual networks to improve the quality of the generated image.

Task-Based Language Grounding using Disentangled Representations

Dr. Ruslan Salakhutdinov, Dr. Katernia Fragkiadaki [Report]

Incorporate β-Variational Autoencoders with recurrent attention mechanism to encode images into structured and factorized latent space representations. Use Asymmetric Advantaged Actor Critic algorithm to learna policy for language instruction-based tasks on VizDoom environment.

Realistic Image Compositing Using GANs

Dr. Abhinav Gupta [Report]

Use Spatial Tranformer Networks (STN) based GANs on CLEVR dataset to predict warp parameters for geometrically correct image compositing. We devise a novel approach by using reinforcement learning to determine how many objects should be added to a scene.

Generating Video using Variational Autoencoders (VAEs)

Dr. Vineeth N. Balasubramanian [Report]

Use Variational Autoencoders to generated videos with spatio-temporal coherence. Recurrent "attention" mechanism is used to attend to each frame separately. Condition the videos on text descriptions to generate videos from a caption.

Analysis of Geographic Atrophy in Fundus Autofluorescence (FAF) Images

Dr. Soumya Jana and Dr. Jay Chabblani (LV Prasad Eye Institute, Hyderabad, India) [Report]

Software tool for opthalmologists that places an ETDRS grid over the atrophic region. Use SIFT and HOG features from the sectors of the ETDRS grid as features. Train SVM and use the distance of the datapoint from the decision boundary as a metric to ascertain the severity of the disease.

Apparent Age Estimation

Dr. Vineeth N. Balasubramanian

Use a Convolutional Neural Network (CNN) to get facial part responses from a person's facial image to segment out their eyes, nose, lips, forehead and neck region. Train 5 small CNNs on these part responses to ascertain a person's age separately. Take an expectation of the result of the 5 network to ascertain a person's apparent age.

Image to Latex Equation Converter and Solver

Dr. Sumohana Channappayya, Best Project Award

Using Hough Transform, SIFT and wavelet filters to segment out the equation characters from the image. Neural Network to recognize the characters present in the image. Combine the characters to form an equation and solve them using a customized solver.

Using Sequences of Photos for Travel Recommendation.

Dr. K Sri Rama Murthy

Used Hidden Markov Models to identify underlying patterns in tourist travels. Optimization of the Recommendation System based on these image sequences.

Teaching Assistant

Carnegie Mellon University

○ Computer Vision, 16-720   *Fall-2018*

○ Computer Vision, 16-720   *Spring-2018*

Indian Institute of Technology, Hyderabad

○ Pattern Recognition and Machine Learning   *Spring-2017*

○ Digital Signal Processing   *Spring-2016*

○ Introduction to Probability   *Spring-2015*

Relevant Graduate Coursework

Learning for Manipulation, Machine Learning, Computer Vision, Deep Reinforcement Learning, Math Fundamentals for Robotics, Monte Carlo Inference on Bayesian Networks, Adaptive Signal Processing, Speech Signal Processing

Skills

Languages   C, C++, Python, Lua, Matlab

Framework   PyTorch, Tensorflow, Theano, ROS, VREP, Mujoco, CUDA, OpenCV, OpenGL