

EDUCATION

- **New York University** New York, USA
Ph.D. in Electrical and Computer Engineering; GPA: 3.95/4.00 *Sep 2017 – Present*
- **Tianjin University** Tianjin, China
B.Eng. in Electronic Science and Technology; GPA: 3.86/4.00 *Sep 2013 – July 2017*
- **Nanyang Technological University** Singapore
Exchange student in Electrical and Computer Engineering; GPA: 5.00/5.00 *Aug 2016 – Dec 2016*
- **University of Toronto** Toronto, Canada
Exchange student in Electrical and Computer Engineering; GPA: 4.00/4.00 *Sep 2015 – Dec 2015*

EXPERIENCE

- **NYU Video Lab** Brooklyn, NY, USA
Research Assistant with Prof. Yao Wang *Oct 2017 – Present*
 - **Detailed 3D Face Reconstruction (WIP)**: Proposed to learn a detailed UV images and texture images as supplementary for the over-smoothed 3DMM face models; Implemented ResNet50 and U-Net for 3DMM parameter learning and residual UV/texture map; Realized detailed face reconstruction from single 2D image with more vivid looking and face details.
 - **TV Show Understanding via Video (WIP)**: Proposed a graph based method for TV show understanding and scene segmentation; Introduced object and background based graph representation for shots; Employed detectron for dense scene parsing and object detection.
 - **TV Show Understanding via Text**: Performed multi-class multi-label classification of high level semantic concepts for closed captions of TV shows at the scene level; Introduced focal loss for extremely imbalanced label distribution; Leveraged Transformer architecture (GPT & BERT) for topic classification in taxonomy. Joint work with *Viacom*.
 - **Salient 360**: Proposed a 2-stage strategy for saliency detection on 360 images. Introduced equi-rectangular mapping for sphere projection and fine-tuned the SalGAN for saliency detection on projected cubes; Designed a de-noising auto-encoder with dilated convolution to stitch predicted cubic saliency maps and remove stitching artifacts; Win a third place in the 2018 ICME Salient360! Grand Challenge.
- **Samsung Research American** Mountain View, CA, USA
Research Intern with Dr. Madhukar Budagavi *Jun 2019 – Aug 2019*
 - **Voice Over 3D Character Animation (VOCA)**: Implemented a deep neural net to regress the vertex offset of 3D FLAME head model given the audio. Employed deepspeech to convert the raw audio input into audio features. Achieved realistic human head speaking animation based on audio input only.
 - **Video Over 3D Face Animation**: Implemented an end-to-end coarse-to-fine system for detailed and textured 3D face shape and pose estimation from monocular videos. Implemented a pipeline for detailed geometry, albedo, lighting, pose and projection parameters reconstruction in real-time. Proposed to employ Neural Differentiable Renderer to learn face parameters in unsupervised manner. Achieved mean vertex distance less than 5mm for face reconstruction.
 - **3DMM face model fitting over head scan**: Fast fitting of 3DMM Face models to high-resolution 3D human head scans with PyTorch. Designed a 2-stage pipeline for 3D face model registration on GPU with face landmark loss and chamfer distance. Our pipeline provides robust and precised face fitting solution in high speed.
 - **3D Talking Face with Personalized Pose Dynamics**: Learn audio based personalized pose dynamics for Obama with conditional Generative Adversarial Networks; Introduced a pose guided VOCA model for talking face generation from audio; Joint work with Prof. Xiaohu Guo's group, results submitted to CVPR 2020
- **AT&T Labs Research, AT&T** Middletown, NJ, USA
Research Intern with Dr. Zhu Liu *Jun 2018 – Aug 2018*
 - **Story Segmentation for TV News Closed Captions**: Proposed a hierarchical model that utilized GRU and ConvNet to perform story segmentation; A bi-directional GRU with self attention was leveraged to summarize each sentence, and followed by a ConvNet to aggregate information at the sentence level; Outperformed the SOTA by 3% in F1 score.

Surgical Planning Lab, BWH, Harvard Medical School

Boston, MA, USA

Research Intern with Prof. Jayender Jagadeesan

Dec 2016 – May 2017

- **Breast Deformation Estimation with VAE:** Performed MRI images dimension and spacing preprocessing, prone and supine MRI breast images alignment with 3D slicer; Proposed to apply Variational Auto-encoder (VAE) to estimate the deformation of breast tissue in MRI images over different poses through the mapping in the latent space; Applied adversarial training as fine tuning; Obtained structural similarity index of 0.65.
- **Lung Nodule Malignancy Classification:** Performed 3D Non-Uniform Sampling of the Lung Nodule from MRI images. Trained a 3D CNN to classify the nodule using the image only, classification accuracy was twice higher than the human radiology experts, surpassing the previous method using SVM with hand crafted features by 5%. Implemented and Evaluated the Multi-cropping network Architecture.

PROJECTS

- **Joint learning of disentangle representations across modalities:** Designed a bi-auto-encoder architecture for joint learning of disentangled representations (content v.s. pose) across video and audio modality using adversarial training. Evaluated on the synthesized bouncing MNIST dataset augmented with sound. Successfully verified the clear and complete disentanglement of learned representations within each modality and shareable pose representations across modalities
- **Image Style Transformation:** Conducted image style transformations between real scene driving images and virtual (game) driving images (GTA V) via cycleGAN, realized high-resolution scene conversion between natural scene and game scene. Converted driving videos recorded from virtual scene (game) to real scene via the trained cycleGAN model.
- **Neural Machine Translation:** Implemented seq2seq, seq2seq with dot-product attention, convolutional attention model and Transformer for machine translation between English and Chinese, English and Vietnamese.

PROGRAMMING SKILLS

- **Languages:** Python, C/C++, Matlab **Deep Learning Library:** PyTorch, Keras **Others:** OpenGL