

Yanwu Xu

👤 1996.03 📍 15224, Pittsburgh, US 📞 +01 7249893729 ✉️ yanwuxu@pitt.edu

Personal Summary

Currently, I am a 4rd-year Ph.D. student in [Intelligent Systems](#) at [The University of Pittsburgh](#). I enjoy the process when we found an interesting research problem and make it a useful application. I want to work with a team and share happiness when we make success. Github:

Education

The University of Pittsburgh

Ph.D., Intelligent System Program.

Advisor: Kayhan Batmanghelich, Co-Advisor: Mingming Gong.

Pittsburgh, PA 15224, USA

Sep. 2019-

Central South University

B.S., Electronic Mechanical Engineering.

Changsha, C.N.

Sep. 2013-2017

Technical Skills

Programing Language: Python , Matlab, C++. Machine Learning Framework: Pytorch(master), Tensorflow.

Work Background

- Computer Vision: text image editing, Nerf 3D reconstruction, 3D image style transfer, video style transfer, video motion tracking, metric learning, few-shot learning, semantic Segmentation, object detection, image generation, image editing.
- Medical Image Analysis: brain tumor detection, lung lesion segmentation.
- Machine Learning: adversarial learning, domain transfer, weakly supervised learning, unsupervised learning.
- Familiar Model: GAN, VAE, F-RCNN, LSTM, etc.

Research Topics

My research topics mainly include the generative models (Diffusion, GAN, VAE) and their application on text-2-image generation, text image editing, weakly supervised learning, domain adaptation, and image-to-image translation. I study the stability of the training generative model on the large-scale dataset and training model with limited data and limited annotations. My current applications are mainly on image style transfer of 2D object, 3D object and videos and Nerf 3D reconstruction . My other research also covers the area of medical image analysis, which is weakly supervised brain tumor segmentation, domain adaptation, and domain generalization on abdominal segmentation tasks.

Work Experience

Sep.2022-Nov.2022: Student researcher, **Google**, Mountain View, CA, USA

During the research work, we target at boost the generation speed of Diffusion model and the text image editing with Diffusion model.

May.2022-Augst.2022: Research intern, **Google**, Mountain View, CA, USA

During the intern, we developed a on-device product for the general style transfer, which can be real-time for photos and videos. In this project, I built a pipeline for training the styleization models and the data preprocessing.

Augst.2022-November.2022: Student researcher, **Google**, Mountain View, CA, USA

In the extension work as student researcher, we are focusing on 3D object stylization and the high fidelity image stylization and editing. We propose two novel task and framework for 3D NeRF stylization and 2D image style transfer based on the recent diffusion generative model.

Research Experience

2017-2018: Research intern, **Universit'e Paris-Est**, Paris, Fr

Advisor: Prof. Chaohui Wang.

Main work: Develop a robust local descriptor model, and ACCV2018 accepts the work.

Jun. 2018: Visiting Scholar, **University of Pittsburgh**, Pitt, US

Advisor: Prof. Kayhan Batmanghelich.

Main work: Propose a multi-scale network based on U-Net for brain tumor segmentation. Our method gets rank 9/66 in the **BraTS challenge**.

2019 -- Current: Ph.D, **University of Pittsburgh**, Pitt, US

Advisor: Prof. Kayhan Batmanghelich, Co-Advisor: Mingming Gong.

Main works:

Conditional GANs: We found that AC-GAN fails on the large scale conditional generation on the image caused by the missing entropy term. To solve this issue, we propose the method TAC-GAN with mutual information maximization player to complement the missing term and achieve the SOTA on multiple datasets, CIFAR10, CIFAR100 and VGGFACE. Accepted by NeurIPS 2019.

Complementary Learning: In this task, we can only access the complimentary labels, which only indicate the negative category. Thus, to improve the previous discriminative method, we propose the generative-discriminative model to generate accurate annotated samples. The experiments show that our method improves this task by a large gap compared to the discriminative method. Accepted by AAAI 2020.

Image-to-Image Translation: The current popular I2I models, such as CycleGAN, GCGAN, and CUT, are too restrictive or weak for specific I2I tasks. For more general applications, we propose a Maximum Spatial Perturbation Consistency (MSPC) model, which designs an adversarial spatial perturbation against the translation network, and we achieve the best results over several popular image translation tasks and our designed face posing transfer dataset. Accepted by CVPR2022.

Domain Generalization: Different from the domain adaptation task, the DG does not have access to the images in the target domain. We can utilize a shallow CNN to generate the synthetic domain to generalize to the unseen domains. We propose an adversarial domain generalization against the main task, which can further improve the generalization ability of the classification and segmentation models. Under review of MICCAI2022.

Research Publication

- Maximum Spatial Perturbation Consistency for Unpaired Image-to-Image Translation, CVPR2022, Yanwu Xu, Shaoan Xie, Wenhao Wu, Kun Zhang, Mingming Gong*, Kayhan Batmanghelich* (* equal advising).
- Adversarial Consistency for Single Domain Generalization in Medical Image Segmentation, MICCAI2022, Yanwu Xu, Shaoan Xie, Maxwell Reynolds, Matthew Ragoza, Mingming Gong, Kayhan Batmanghelich.
- Generative-Discriminative Complementary Learning, AAAI2020, **Yanwu Xu**, Mingming Gong, Junxiang Chen, Tongliang Liu, Kun Zhang, Kayhan Batmanghelich.
- Twin Auxiliary Classifiers GAN, NeurIPS2019 spotlight (0.2%), Mingming Gong*, **Yanwu Xu***, Chunyuan Li, Kun Zhang, Kayhan Batmanghelich (* equal contribution).
- Compressed Self-Attention for Deep Metric Learning, AAAI2020, Ziye Chen, Mingming Gong, **Yanwu Xu**, Chaohui Wang, Kun Zhang, Bo Du.
- Robust Local Descriptor Learning, ACCV2018, **Yanwu Xu**, Mingming Gong, Tongliang Liu, Kayhan Batmanghelich, and Chaohui Wang.
- Multi-scale Masked 3-D U-Net for Brain Tumor Segmentation, BraTS2018 LNCS, **Yanwu Xu**, Mingming Gong, Huan Fu, Dacheng Tao, Kun Zhang and Kayhan Batmanghelich.
- DSANet: Dynamic Segment Aggregation Network for Video-Level Representation Learning, ACM MM 2021, Wenhao Wu, Yuxiang Zhao, **Yanwu Xu**, Xiao Tan, Dongliang He, Zhikang Zou, Jin Ye, Yingying Li, Mingde Yao, Zichao Dong, Yifeng Shi
- Unaligned image-to-image translation by learning to reweight, ICCV 2021, Shaoan Xie, Mingming Gong, **Yanwu Xu**, Kun Zhang
- Multi-scale masked 3-D U-net for brain tumor segmentation, MICCAI Brainlesion Workshop 2019, **Yanwu Xu**, Mingming Gong, Huan Fu, Dacheng Tao, Kun Zhang, Kayhan Batmanghelich.
- 3D-BoxSup: Positive-Unlabeled Learning of Brain Tumor Segmentation Networks From 3D Bounding Boxes, Frontiers in neuroscience 14, 350, **Yanwu Xu**, Mingming Gong, Junxiang Chen, Ziye Chen, Kayhan Batmanghelich

Conference PC

NeurIPS2022, ECCV2022, MICCAI2022, ICPR2022, IJCAI2021, IJCAI2020, CDML2020(NeurIPS workshop), CVPR2021, AAAI2020