Xutong Ren

(412)535-2277 | xutongr@cs.cmu.edu

EDUCATION

Carnegie Mellon University, School of Computer Science Pittsburgh, PA
Master of Science in Machine Learning | GPA: 4.00 Dec. 2020

Peking University, School of Electronics Engineering and Computer Science Beijing, China
Bachelor of Science in Computer Science | GPA: 3.81 July 2019

SKILLS

- Programming Languages: Python, C/C++, MATLAB, Lua;

- Frameworks: Linux, PyTorch, Torch, Jupyter, Git.

WORK EXPERIENCES

Microsoft Redmond, WA
Data & Applied Science Intern, Security & Compliance Group May 2020 – Aug. 2020

- Formulated phishing email detection as a visual task based on the existence of phishing templates among phishing campaigns.

- Designed and implemented a software (in $C^{\#}$ and Python) that can render synthetic phishing emails with different templates, providing large scale phishing datasets without privacy issues for product testing.

- Proposed an attention framework via class activation maps. Achieved FPR $< 0.1\%$ and TPR $> 98.5\%$ on real/synthetic data.

- Developed an advanced classifier (in PyTorch) that automatically localizes critical visual components (logos, banners, signatures) and detects phishing emails unsing global-local infomation, improve the TPR from $\sim$40% to $\sim$95% for unseen templates.

- Applied image rotation prediction to extract visual embeddings for OPTICS to distinguish email templates and do phishing campaign detection. The template detection performance achieved NMI $> 86.4\%$.

Huawei Beijing, China
Research Intern, Noah's Ark Lab May 2019 – Aug. 2019

- Proposed a generalized coarse-to-fine model (in Lua) with effective information passing to accommodate different image recognition tasks. Brought gains of 2% - 10% in few shot classification, object localization, biomedical semantic segmentation, etc.

- Designed a progressive training strategy to improve learning stability and relieve over-fitting, which re-defines the concept of sample difficulty. Overcame the defects of curriculum learning and made the theory feasible in practical.

- Explored the problem of scene classification with partial feedbacks and equipped current margin-based algorithms with active learning techniques, making online scene identification possible.

- Helped building a self-supervised visual representation learning approach. Implemented a recurrent solution (in TensorFlow) to jigsaw puzzles of arbitrary permutations.

PROJECT EXPERIENCES

Carnegie Mellon University, Biomedical Software Development August. 2020 – Present

- Maintained and solved issues for AITom, an open-source AI driven cellular electron cryo-tomography analysis software.

- Helped develop weakly-supervised approaches to perform 3D scene understanding for the development of downstream biomedical tasks. Implemented few-shot semantic segmentation for cellular component detection of cryo-ET.

Google, AI Machine Learning Competition Jan. 2019

- Focused on image to image translation and realized human face attributes editing in an unsupervised and interactive way.

- Proposed and Implemented a VAE-GAN that is capable of both decoupling human images into identity preserving details and structure reserving skeletons and compositing them into realistic human faces.

- Led a team of three to do data cleaning, ensemble modules into a final pipeline and prepare the demo and presentation. This project won the *Most Technical Award*.

DiDi/Kuaishou, Images & Videos Enhancement May. 2017 – July 2019

- Proposed a joint low-light enhancement and denoising algorithm (in MATLAB) based on a sequential Retinex decomposition, simultaneously solving two related image degradation problems and improving visual quality.

- Developed a low-rank solution to enhance low-light images with heavy noise. Extended image enhancement methodology to videos via 3D implementation and guided filter.

- Applied a mesh-flow denoising technique (in C++) to efficiently remove random noises in videos posted on Kuaishou platform.

- Set up in-vehicle multimedia degradation simulations and built paired in-vehicle multimedia corpus, containing both degraded and ground truth data.