Qing Liu Email: qingliu.research@gmail.com Homepage: https://qliu24.github.io

My research interests are in Computer Vision, especially the area of using generative models for image editing tasks, such as inpainting, outpainting, super-resolution, etc. The goal is to provide smart image editing tools that allow better controllability and encourage more creativity for our users. I also work on object recognition and parsing with weak supervision, including but not limited to: weakly supervised learning, domain adaptation (learning from synthetic), few-shot/zero-shot learning. My research improves the data efficiency and the robustness of deep learning models, making them more useful in real world applications.

WORK EXPERIENCE

Research Engineer 2, Adobe 11/29/2021 - Present

345 Park Ave, San Jose, CA 95110

I work for the Media Intelligence Lab at Adobe Research and my responsibility is to explore artificial intelligence based image editing technology and transfer them into high impact features in Adobe products, such as Photoshop and Lightroom. Using my computer vision background, I design and implement new frontier machine learning and deep learning models and test their performance compared with existing techniques. We publish our novel findings at top-tier computer vision journals and conferences, and build interactive web demos to show real use cases. To transfer them into Adobe products, I also work on model deployment for cloud service and mobile devices.

Research Intern, Facebook 05/18/2020-09/11/2020

1 Hacker Way, Menlo Park, CA 94025

I worked for the Facebook AI video team and my responsibility was to conduct computer vision research using video data and deep learning models. I performed a detailed literature search and proposed a novel algorithm for weakly supervised instance segmentation for video objects. I implemented the model using python packages and ran the experiments on GPU clusters. Experimental comparisons with other existing methods were also carried out by me. The model achieved state of the art performance and the work was accepted by the 2021 Conference on Computer Vision and Pattern Recognition.

Applied Scientist Intern, Amazon 06/03/2019-09/06/2019

410 Terry Ave N, Seattle, WA 98109

I worked for the Amazon AWS ReKognition team and my responsibility was to develop a computer vision model that can quickly adapt to unseen object classes. The model should also achieve better performance when more data was processed sequentially. I formulated the problem as a new task for computer vision that combines incremental learning and meta-learning together. I designed a new algorithm that involved prototype based few-shot classification and knowledge distillation, and I implemented it using python and deep learning packages. The proposed method outperformed existing solutions by a large margin and the work was accepted by the 2020 European Conference on Computer Vision.

Applied Scientist Intern, Amazon 05/28/2018-08/24/2018

410 Terry Ave N, Seattle, WA 98109

I worked for the Amazon Transaction Risk Management Systems team and my intern project was “Ladder Neural Network for Continuous Fraud Detection and Queue Rate Reduction”. I used a semi-supervised deep generative model to fit the user transaction data which achieved better performance than the traditional methods on fraud detection using much less labeled data. I implemented the model using python and deep learning packages and experimented with multiple variations. My project won the 2018 Amazon TRMS Science Fair gold award and was deployed into the production line.

EDUCATION

Johns Hopkins University, Baltimore, MD — GPA: 4.00
PhD in Computer Science — 2016/09 – 2021/11
MSE in Computer Science — 2015/01 – 2016/05

The University of Texas at Dallas, Richardson, TX — GPA: 3.85
MS in Molecular and Cell Biology — 2011/08 – 2013/05

Peking University, Beijing, China
BS in Chemistry and Psychology — 2005/09 – 2010/07

RECENT RESEARCH PROJECTS

Object-level Scene Deocclusion

We present a new self-supervised PArallel visible-to-COmplete diffusion framework, named PACO, a foundation model for object-level scene deocclusion. Leveraging the rich prior of pre-trained models, we first design the parallel variational autoencoder, which produces a full-view feature map that simultaneously encodes multiple complete objects, and the visible-to-complete latent generator, which learns to implicitly predict the full-view feature map from partial-view feature map and text prompts extracted from the incomplete objects in the input image. To train PACO, we create a large-scale dataset with 500k samples to enable self-supervised learning, avoiding tedious annotations of the amodal masks and occluded regions. At inference, we devise a layer-wise deocclusion strategy to improve efficiency while maintaining the deocclusion quality. Extensive experiments on COCOA and various real-world scenes demonstrate the superior capability of PACO for scene deocclusion, surpassing the state of the arts by a large margin. Our method can also be extended to cross-domain scenes and novel categories that are not covered by the training set. Further, we demonstrate the deocclusion applicability of PACO in single-view 3D scene reconstruction and object recomposition.

UniHuman: A Unified Model for Editing Human Images in the Wild

We propose UniHuman, a unified model that addresses multiple facets of human image editing in real-world settings. To enhance the model's generation quality and generalization capacity, we leverage guidance from human visual encoders and introduce a lightweight pose-warping module that can exploit different pose representations, accommodating unseen textures and patterns. Furthermore, to bridge the disparity between existing human editing benchmarks with real-world data, we curated 400K high-quality human image-text pairs for training and collected 2K human images for out-of-domain testing, both encompassing diverse clothing styles, backgrounds, and age groups. Experiments on both in-domain and out-of-domain test sets demonstrate that UniHuman outperforms task-specific models by a significant margin. In user studies, UniHuman is preferred by the users in an average of 77% of cases.

Towards Open-World Segmentation of Parts

Segmenting object parts such as cup handles and animal bodies is important in many real-world applications but requires more annotation effort. To address this, we propose to explore a seemingly simplified but empirically useful and scalable task, class-agnostic part segmentation. In this problem, we disregard the part class labels in training and instead treat all of them as a single part class. We argue and demonstrate that models trained without part classes can better localize parts and segment them on objects unseen in training. We then present two further improvements. First, we propose to make the model object-aware, leveraging the

fact that parts are "compositions" whose extents are bounded by objects, whose appearances are by nature not independent but bundled. Second, we introduce a novel approach to improve part segmentation on unseen objects by proposing a self-supervised procedure that iterates between pixel clustering and a supervised contrastive learning that pulls pixels closer or pushes them away. Via extensive experiments on PartImageNet and Pascal-Part, we show notable and consistent gains by our approach, essentially a critical step towards open-world part segmentation.

PUBLICATIONS

- Zhengzhe Liu, **Qing Liu**, Chirui Chang, Jianming Zhang, Daniil Pakhomov, Haitian Zheng, Zhe Lin, Daniel Cohen-Or, Chi-Wing Fu. *Object-level Scene Deocclusion*. SIGGRAPH 2024.
- Nannan Li, **Qing Liu**, Krishna Kumar Singh, Yilin Wang, Jianming Zhang, Bryan A. Plummer, Zhe Lin. *UniHuman: A Unified Model for Editing Human Images in the Wild*. CVPR 2024.
- Jaskirat Singh, Jianming Zhang, **Qing Liu**, Cameron Younger Smith, Zhe Lin, Liang Zheng. *SmartMask: Context Aware High-Fidelity Mask Generation for Fine-grained Object Insertion and Layout Control*. CVPR 2024.
- Bowen Zhang, **Qing Liu**, Jianming Zhang, Yilin Wang, Liyang Liu, Zhe Lin, Yifan Liu. *Amodal Scene Analysis via Holistic Occlusion Relation Inference and Generative Mask Completion*. AAAI 2024 (Oral).
- Jing Gu, Yilin Wang, Nanxuan Zhao, Tsu-Jui Fu, Wei Xiong, **Qing Liu**, Zhifei Zhang, He Zhang, Jianming Zhang, HyunJoon Jung, Xin Eric Wang. *PHOTOSWAP: Personalized Subject Swapping in Images*. NeurIPS 2023.
- Yu Zeng, Zhe Lin, Jianming Zhang, **Qing Liu**, John Collomosse, Jason Kuen, Vishal M. Patel. *SceneComposer: Any-Level Semantic Image Synthesis.* Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.
- Tai-Yu Pan, **Qing Liu**, Wei-Lun Chao, Brian L. Price. *Towards Open-World Segmentation of Parts*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.
- **Qing Liu**, Adam Kortylewski, Zhishuai Zhang, Zizhang Li, Mengqi Guo, Qihao Liu, Xiaoding Yuan, Jiteng Mu, Weichao Qiu, Alan Yuille. *Learning Part Segmentation through Unsupervised Domain Adaptation from Synthetic Vehicles*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.
- **Qing Liu**, Vignesh Ramanathan, Dhruv Mahajan, Alan Yuille, Zhenheng Yang. *Weakly Supervised Instance Segmentation for Videos with Temporal Mask Consistency*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
- Nicholas Ichien*, **Qing Liu***, Shuhao Fu, Keith J Holyoak, Alan Yuille, Hongjing Lu. *Visual analogy: Deep learning versus compositional models*. Proceedings of the 43rd Annual Meeting of the Cognitive Science Society. 2021.
- Adam Kortylewski, **Qing Liu**, Angtian Wang, Yihong Sun, Alan Yuille. *Compositional convolutional neural networks: A robust and interpretable model for object recognition under occlusion*. International Journal of Computer Vision (2020): 1-25.
- **Qing Liu**, Orchid Majumder, Alessandro Achille, Avinash Ravichandran, Rahul Bhotika, Stefano Soatto. *Incremental Meta-Learning via Indirect Discriminant Alignment*. Proceedings of the European Conference on Computer Vision. 2020.
- Adam Kortylewski, Ju He, **Qing Liu**, Alan Yuille. *Compositional Convolutional Neural Networks: A Deep Architecture with Innate Robustness to Partial Occlusion*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.
- Adam Kortylewski, **Qing Liu**, Huiyu Wang, Zhishuai Zhang, Alan Yuille. *Combining Compositional Models and Deep Networks For Robust Object Classification under Occlusion*. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2020.
- **Qing Liu**, Lingxi Xie, Huiyu Wang, Alan Yuille. *Semantic-Aware Knowledge Preservation for Zero-Shot Sketch-Based Image Retrieval*. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.

- **Qing Liu***, Yutong Bai*, Lingxi Xie, Yan Zheng, Weichao Qiu, Alan Yuille. *Semantic Part Detection via Matching: Learning to Generalize to Novel Viewpoints from Limited Training Data*. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.
- Adam Kortylewski, **Qing Liu**, Huiyu Wang, Zhishuai Zhang, Alan Yuille. *Localizing Occluders with Compositional Convolutional Networks*. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshop. 2019.
- Hongjing Lu, **Qing Liu**, Nicholas Ichien, Alan L. Yuille, Keith J. Holyoak. *Seeing the Meaning: Vision meets Semantics in Solving Pictorial Analogy Problems.* Proceedings of the 41st Annual Meeting of the Cognitive Science Society. 2019.
- Boyang Deng, **Qing Liu**, Siyuan Qiao, Alan Yuille. *Few-shot Learning by Exploiting Visual Concepts within CNNs.* arXiv preprint arXiv: 1711.08277.