Xingrui WANG

(+1) 2136757298 | xwang378@jhu.edu| 3100 Wyman Park Dr, Baltimore, MD 21211

https://xingruiwang.github.io/

EDUCATION BACKGROUND

Whiting School of Engineering, Johns Hopkins University — **Baltimore, MD**

Ph.D. in Computer Science; *08/2023- Present*

Viterbi School of Engineering, University of Southern California — **Los Angeles, CA**

M.S. in Applied Data Science; **GPA**: 3.92 / 4.00 *08/2021- 05/2023*

School of Statistics, Renmin University of China — **Beijing, China**

B.S. in Statistics; Minor in Data Science; **GPA**: 87.04 / 100 *09/2017- 07/2021*

PUBLICATIONS

- **Xingrui Wang**, Wufei Ma, Zhuowan Li, Adam Kortylewski, Alan Yuille. 3D-Aware Visual Question Answering about Parts, Poses and Occlusions. *NeurIPS 2023*.
- Zhuowan Li, **Xingrui Wang**, Elias Stengel-Eskin, Adam Kortylewski, Wufei Ma, Benjamin Van Durme, Alan Yuille. Super-CLEVR: A Virtual Benchmark to Diagnose Domain Robustness in Visual Reasoning. *CVPR 2023* ***(Highlight).***
- Yunhao Ge*, Yao Xiao*, Zhi Xu, **Xingrui Wang**, Laurent Itti. Contributions of Shape, Texture and Color in Visual Recognition. *ECCV 2022*
- Haoyu Liu, Yang Liu, **Xingrui Wang**, Hanfang Yang. Towards Language Hint Attention Reinforcement Learning. IEEE World Congress on Computational Intelligence, *WCCI 2022;*
- **Xingrui Wang**, Xinyu Liu, Ziteng Lu, Hanfang Yang, Large Scale GPS Trajectory Generation Using Map Based on Two Stage GAN, J. data sci. 19(2021), no. 1, 126-141. *DOI 10.6339/21-JDS1004*
- Jinhua Su, Yanbing Bai, **Xingrui Wang**, etc. Technical Solution Discussion for Key Challenges of Operational Convolutional Neural Network-Based Building-Damage Assessment from Satellite Imagery: Perspective from Benchmark xBD Dataset. Remote Sensing. 2020; 12(22):3808. *DOI: 10.3390/rs12223808*

WORKING EXPERIENCE

CCVL, Johns Hopkins University | Research Intern — **Baltimore, MD**

Advisor: Prof. Alan Yuille. *06/2022- 12/2022*

Research Area:

- **(VQA)** Domain shift in visual question answering and neural symbolic model;
- **(VQA)** 3D aware visual question answering with 3D object detection in neural mesh model.

iLab, University of Southern California | Research Assistant — **Los Angeles, CA**

Advisor: Prof. Laurant Itti. *09/2021- 05/2022*

Research Area:

- **(Explainable AI)** Explore the decoupling feature extractor on shape, texture, and color on visual recognition and its application in zero-shot learning and image generation.
- **(Human-centric AI)** The knowledge exchange between human and neural network via knowledge distillation.

Samsung R&D Institute China-Beijing | Research Intern — **Beijing, China**

Advisor: Dr. Yang Liu *12/2020- 06/2021*

Research Area:

- **(Explainable RL)** Human pre-guided attention in reinforcement learning with high efficiency learning.
- **(Embodied AI)** ALFRED Challenge, EAI @ CVPR 2021.

Wangxuan Institute of Computer Technology, Peking University | Research Intern — **Beijing, China**

Advisor: Prof. Yongtao Wang *09/2019- 02/2020*

Research Area:

- **Semantic segmentations** with multi-Scale Feature Learning.

RESEARCH EXPERIENCE

3D aware VQA with neural mesh based 3D objects detection | CCVL, JHU **Baltimore, MD**

Advisor: Prof. Alan Yuille *06/2022- 12/2022*

(NeurIPS 2023)

- Design visual question answering on the parts, 3D pose and occlusions about objects.
- Proposed PO3D-VQA, a VQA model with 3D scene parser based on neural mesh model to detect and reproject objects.
- Our PO3D-VQA ourpreform the mDETR, FiLM, P-NSVQA and its 3D variants on answering parts, poses and occlusions questions

Domain shift in VQA dataset and probabilistic symbolic model | Research Assistant, CCVL, JHU **Baltimore, MD**

Advisor: Prof. Alan Yuille *06/2022- 12/2022*

(CVPR 2023 highlight)

- We propose Super-CLEVR, a new VQA benchmark with controlable domain shift factors
- Design a new symbolic model P-NSVQA, enable probabilistic reasoning module based on NS-VQA;
- Compare the in-domain and out-of-domain performance of FILM, NSVQA, NSCL, mDETR and our P-NSVQA;
- Super-CLEVR reveals the lack of generalization ability of current VQA baseline on out-of-domain while P-NSVQA can outpreform other methods.

Generic Interface for Human-Neural Network Knowledge Exchange | Research Assistant, iLab, USC **Los Angeles, CA**

Advisor: Prof. Laurent Itti. *04/2022- 10/2022*

- Propose a framework (HNI) for human to interact with nerual network using a structural representation of visual concepts
- In image classification, HNI can visualize the logic of inference with class-specific Structural Concept Graphs (SCC)
- Human can directly provide feedback and guide the nerual network by modifying the SCC.
- The logic of SCC can also transfer to the neural network with knowledge distillation
- The experiment on ImageNet shows that the accuracy on target classes can increase by around 4% without much drop on the other classes

Contributions of Shape, Texture, and Color in Visual Recognition | Research Assistant, iLab, USC **Los Angeles, CA**

Advisor: Prof. Laurent Itti. *09/2021- 12/2021*

(ECCV 2022)

- Inspired by the human vision system and proposed a humanoid vision engine (HVE) to separately compute shape, texture and color features in image recongnition
- Built a pipeline to summarize the contributions of these features in image recongnition for a given dataset automaticly
- Compared with the results of questionnaire human participents, verified that the contributions consist to the importances of these features in human's decision process
- Explored the protential application of HVE in open-world zero shot learning and image imaginitation from features

Alfred@EAI workshop, CVPR 2021 | Research Intern, Samsung (SRC-B) **Beijing, China**

Advisor: Dr. Yang Liu. *04/2020- 06/2021*

- Based on the seq2seq baseline, adding a depth estimation module to generate a 2D obstacle map in real-time
- With the segmentation result, ground the object position on the obstacle map
- Use FMM to generate navigation to the grounded object and combine with low level language to guide action
- The proposed method outperforms the baseline by 5% on the unseen success rate.

Coarse-grained Pre-guided Attention in Reinforcement Learning | Research Intern, Samsung (SRC-B) **Beijing, China**

Advisor: Dr. Yang Liu. *12/2020- 06/2021*

(WCCI 2022)

- Proposed an experimental and exploratory method, language hint attention reinforcement learning
- Tracked the template images in the whole frame and use the result as the human coarse-grained pre-guided attention
- Fuse the coarse-grained pre-guided attention with fined-grained attention and sent it to LSTM controller
- The experiments in Atari Game validated that our hint information can improve the sample efficiency significantly

Trajectory Generation based on Two-stage GAN | Research Assistant, Renmin University of China **Beijing, China**
Advisor: Prof. Hanfang Yang. *09/2019- 03/2020*
(Journal of Data Science, 2021)

- Studied many model design methods in image caption, pedestrian trajectory estimation and machine translation
- Built a two-stage GAN model by PyTorch, which could effectively extract image features and generate trajectory data with both large-scale and high-precision
- Calculated the JS-divergence of the generated data and the real data on different indicators such as trajectory point velocity, which indicated that our model fitted the real data better than the previous models
- Adopted the API of OpenStreetMap and Baidu Map to make a trajectory visualization program

Change Detection by Kernel Two Sample Test | Research Assistant, Renmin University of China **Beijing, China**
Advisor: Prof. Hanfang Yang *10/2020- 01/2021*
(Undergraduate thesis)

- Proposed a nonparametric two sample test to identify the changes in the spatial area
- For two sets of images during two specific time periods in various condition of the same area, applied two sample test and decide if the differences are significant
- The experiments under progress showed that the detection result is more robust when converting this task to a statistical problem instead of using the popular deep learning models

Building Damage Assessment from Satellite Imagery | Research Assistant, Renmin University of China **Beijing, China**
Advisor: Prof. Yanbing Bai. *09/2020- 01/2021*
(Remote sensing, 2020)

- Analyzed the current damage detection model in remote sensing image processing area and divide them into object-level model and pixel-level model, and have proposed a metric to evaluate both of them
- Applied GAN model to restore pre-damage images based on post-damage images when only post-damage images are available in the damage assessment task
- The experiments demonstrated our method preformed in detect damage building comparing with previous detection methods based on post-damage image with more detailed experiment in progress

Gated Scale-Transfer Operation for Multi-Scale Feature Learning in Pixel Labeling | Research Assistant **Beijing, China**
Advisor: Prof. Yongtao Wang, *Peking University* *09/2019- 02/2020*

- Added a multi-scale feature fusion module between different scale layers to build a new model, improving the HRNet and enhancing the semantic segmentation effect
- Tested the new model in Pytorch on the Cityscapes, LIP, and Pascal Context dataset and achieved higher MIoU values without additional computation; the MIoU increased by 1.3% and the number of parameters increase by less than 1%
- Used matplotlib to visualize the model attention layer in heat map and compared the result with baseline's, which reflected a better segmentation effect on the object details such as edges and on the objects of different scales

TEACHING EXPERIENCES

University of Southern California **Los Angeles, CA**

- Course Producer: DSCI 552 - Machine Learning for Data Science Spring 2022 - Spring 2023