

# Xijun Wang

Tel: (+1) 240-888-3789      Email: xijun.wang.cs@gmail.com  
4404 Mohegan Place, College Park, MD 20740

## EDUCATION

---

- GAMMA research group, University of Maryland, College Park** 2021-Now  
**Supervisor:** Prof. Ming C. Lin (ACM/IEEE Fellow) & Prof. Dinesh Manocha (ACM/IEEE Fellow)  
*Doctor of Philosophy in Computer Science*
- Visual Information Processing and Learning (VIPL) Lab, Institute of Computing Technology, Chinese Academy of Sciences** 2018-2021  
**Supervisor:** Prof. Meina Kan & Prof. Shiguang Shan (IEEE Fellow) & Prof. Xilin Chen (ACM/IEEE Fellow)  
*Master of Computer Technology*
- School of Information Engineering, China University of Geosciences (Wuhan)** 2014-2018  
(Graduated from the Li Siguang Program Class that selects 30/4000 among freshmen per year)  
*Bachelor of Software Engineering*  
GPA: 91.6/100 (ranking No. 1 among 66 students)

## PROJECT RESEARCH

---

- Unmanned Aerial Vehicle (UAV) Video Activity Recognition (GAMMA)** 2021-Now
- Proposed a novel Fourier Object Disentanglement method (FO) to bestow the network with the ability to intrinsically recognize the moving human actor from the background. FO operates in the frequency domain dictated by the spectrum of the Fourier transform corresponding to the temporal dimensions of the video. It characterizes the motion of the human actor based on the magnitude and rate of temporal change of feature maps that encode information about the spatial pixels of the video.
  - Presented Fourier Attention (FA) to encapsulate context and long range space-time dependencies within a video. Fourier attention works in the frequency domain corresponding to the space-time dimensions of the video.
  - ECCV2022.
- METEOR: A Massive Dense & Heterogeneous Behavior Dataset for Autonomous Driving (GAMMA)** 2021-Now
- Presented a new dataset, METEOR, corresponding to the dense, heterogeneous, and unstructured traffic in India.
  - METEOR has more than 2 million labeled frames and 13 million annotated bounding boxes for 16 unique traffic agents, and GPS trajectories for the ego-agent.
  - Every video in METEOR is tagged using a diverse range of factors including weather, time of the day, road conditions, and traffic density
  - ICRA2023.
- Lesson from a Multi-source Teacher: Auxiliary Modality Distillation for Autonomous Steering (GAMMA)** 2021-Now
- Proposed an expandable learning framework that is compatible with unconstrained form of auxiliary modalities in any arbitrary number of shots, achieving better information extraction, higher efficiency, and fewer data and modalities needed.
  - Proposed a novel training algorithm for modality distillation, and outperform other state-of-the-art methods on the task of modality distillation in autonomous steering.
  - Submitted to ECCV2022.
- Knowledge Distillation (VIPL Lab)** 2020-2021
- Proposed a novel knowledge distillation schema that grew from a tiny network to a target network, which provides a different framework from the existing ones.
  - Developed a novel Triplet Distillation module to collaboratively incorporate the knowledge learned from a teacher network, a large peer model, and a student model, which is inspired by the classroom learning of our human.
  - Provided theoretical analysis in a statistical perspective.
- Model Compression and Acceleration (VIPL Lab)** 2018-2019
- Proposed a fully learnable group convolution module (FLGC) that can be embedded in any deep neural networks for acceleration and automatically optimize the group structure in a fully end-to-end manner by gradient-descent, leading to better structure than the existing predefined, two-steps, or iterative strategies.
  - Accepted by CVPR2019.

## INTERN EXPERIENCE

---

- Research Intern, The Basemodel Group, Megvii Co., Ltd.** 2020&2021  
**Mentor:** Dr. Xiangyu Zhang & Prof. Jian Sun (*ResNet's Authors*)

**Main Duties:** Provided base model support for the company's high-level visual tasks

### Dynamic Region-Aware Convolution

- Proposed a convolution called Dynamic Region-Aware Convolution (DRConv) that can automatically assign multiple filters to corresponding spatial regions where features have similar representation.
- Accepted by CVPR2021.

### Spatial Cross-scale Convolution Module: An Approach to Strengthen both CNNs and Transformers

- Presented a high capacity and effective convolution module named SCSC, which can dynamically combine a large range of receptive fields in one layer to enhance the presentation ability.
- Architectures applied with our method can obtain better performance with fewer parameters and computational cost. Furthermore, SCSC is a general module and can be applied to strengthen both CNNs and Transformers.
- SCSC module can power the classical neural networks (e.g., ResNet50) to achieve comparable performance with strong Transformers (e.g., Swin).
- ICCV2023W.

### Applied Scientist Intern, Amazon/A9

2022 summer

Mentor: *Dr. Shan Yang*

2023 summer

**Main Duties:** Furniture layout and recommendation

#### Paper:

- ICAR: Image-based Complementary Auto Reasoning, IAAA2024
- VLAP: Efficient Video-Language Alignment via Frame Prompting and Distilling for Video Question Answering, submitted to CVPR 2024

### **PUBLICATIONS** (#: first author)

---

- [1] **Xijun Wang**#, Meina Kan, Shiguang Shan, and Xilin Chen. Fully learnable group convolution for acceleration of deep neural networks. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR2019).
- [2] Jin Chen#, **Xijun Wang**#, Zichao Guo, Xiangyu Zhang, Jian Sun. Dynamic Region-Aware Convolution. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR2021).
- [3] Divya Kothandaraman, **Xijun Wang**, Tianrui Guan, Shuowen Hu, Dinesh Manocha. FAR: Fourier Disentangled Space Time Attention for UAV Activity Recognition. (ECCV2022)
- [4] **Xijun Wang**#, Meina Kan, Shiguang Shan, and Xilin Chen. Triplet-Distillation with Model Growing. (under review)
- [5] **Xijun Wang**#, Xiaojie Chu, Xiangyu Zhang, Jian Sun. Spatial Cross-scale Convolution Module: An Approach to Strengthen both CNNs and Transformers. (ICCV2023W)
- [6] Rohan Chandra#, **Xijun Wang**#, Mridul Mahajan, Rahul Kala, Rishitha Palugulla, Chandrababu Naidu, Alok Jain, Dinesh Manocha. METEOR: A Massive Dense & Heterogeneous Behavior Dataset for Autonomous Driving. (ICRA2023)
- [7] **Xijun Wang**#, Ruiqi Xian#, Tianrui Guan, Celso M de Melo, Stephen M Nogar, Aniket Bera, Dinesh Manocha. AZTR: Aerial Video Action Recognition with Auto Zoom and Temporal Reasoning. (ICRA2023)
- [8] Yu Shen, Luyu Yang, **Xijun Wang**, Ming Lin. Small-shot Multi-modal Distillation for Vision-based Autonomous Steering. (ICRA2023)
- [9] Yu Shen, **Xijun Wang**, Peng Gao, Ming Lin. Auxiliary Modality Learning with Generalized Curriculum Distillation. (ICML 2023)
- [10] Tianrui Guan, Aswath Muthuselvam, Montana Hoover, **Xijun Wang**, Jing Liang, Adarsh Jagan Sathyamoorthy, Damon Conover, Dinesh Manocha. CrossLoc3D: Aerial-Ground Cross-Source 3D Place Recognition. (ICCV2023)
- [11] Divya Kothandaraman, **Xijun Wang**, Tianrui Guan, Sean Hu, Ming Lin, Dinesh Manocha. Frequency-based aerial video recognition. (Open Architecture/Open Business Model Net-Centric Systems and Defense Transformation 2023)
- [12] Ruiqi Xian#, **Xijun Wang**#, Dinesh Manocha. MITFAS: Mutual Information based Temporal Feature Alignment and Sampling for Aerial Video Action Recognition. (WACV2024)
- [13] Ruiqi Xian, Xijun Wang, Divya Kothandaraman, Dinesh Manocha. PMI Sampler: Patch similarity guided frame selection for Aerial Action Recognition. (WACV2024)
- [14] **Xijun Wang**#, Anqi Liang, Junbang Liang, Ming Lin, Yu Lou, Shan Yang. ICAR: Image-based Complementary Auto Reasoning (IAAA2024)
- [15] **Xijun Wang**#, Junbang Liang, Chun-Kai Wang, Kenan Deng, Yu Lou, Ming Lin, Shan Yang. VLAP: Efficient Video-Language Alignment via Frame Prompting and Distilling for Video Question Answering (Submitted to CVPR2024)
- [16] **Xijun Wang**#, Ruiqi Xian#, Tianrui Guan, Dinesh Manocha. Prompt Learning for Action Recognition. (Submitted to ICRA2024)
- [17] Tianrui Guan, Fuxiao Liu, Xiyang Wu, Ruiqi Xian, Zongxia Li, Xiaoyu Liu, Xijun Wang, Lichang Chen, Furong

## ***HONORS & AWARDS***

---

- National Scholarship (Graduate) 2019.12
- 4<sup>th</sup> place in Crop Disease Detection, Global AI Challenge 2018.12
- Top 10 Pacesetter, China University of Geosciences (Wuhan) (10/18,000) 2017.05
- National Second Prize, Esri Secondary Development Competition 2016.11
- National Inspirational Scholarship 2016.11
- Advanced Vision Group Champion and Simulation Group Runner-Up, 2016 China Robot Competition 2016.10
- National Scholarship (Undergraduate) 2015.11

## ***SKILLS***

---

- **Computer:** Pytorch, Python, C++.
- **Interests:** Travelling, Hiking, Watching movie, Porcelain, Cook.

## ***SELF-ASSESSMENT***

---

- Competent learner who consolidates the solid theoretic foundation of computer technology and software engineering after years of systematic education.
- Active achiever who obtains several high-level awards and submits my paper to core academic conferences.
- Excellent team-worker who completes various research projects with the capacities of communication, cooperation, and coordination.
- Experienced cross-cultural negotiator who is proficient in English and can adapt to a new environment.