

# Qitong Wang (He/Him/His)

✉ wqtwjt@udel.edu

🐦 @QitongWang2242

🌐 wqtwjt2242

🌐 <https://wqtwjt1996.github.io/>

(Updated on Oct. 14th, 2025.)

## Education

- 2021 – Now     📖 **Ph.D. University of Delaware**; Computer & Information Sciences.  
Advisor: *Christopher Rasmussen*.
- 2018 – 2020     📖 **M.S. Boston University**; Computer Science.  
Advisor: *Margrit Betke*.
- 2014 – 2018     📖 **B.Eng. Wuhan University of Technology**; Software Engineering.  
GPA: 91.04/100; Rank: 2/228.

## Research Publications

### Journal & Conference

- 1     **Q. Wang**, E. Chinkaka, R. Richaud, M. Haghdadi, C. Wolk, K. V. Oromeng, K. F. Davis, F. Bianco, X. Peng, and J. M. Klinger.  
“MO-SAM: Testing the reliability and limits of mine feature delineation using Segment Anything Model to democratize mine observation and research.”  
*PLOS Sustainability and Transformation*, 2025.
- 2     **Q. Wang**, T. Li, K. X. Nguyen, and X. Peng  
“Beyond Accuracy: On the Effects of Fine-tuning Towards Vision-Language Model’s Prediction Rationality.”  
*Association for the Advancement of Artificial Intelligence (AAAI)*, 2025.  
▷ Acceptance rate 23.4%; Top conference in Artificial Intelligence.
- 3     **Q. Wang**, L. Zhao, L. Yuan, T. Liu, and X. Peng  
“Learning from Semantic Alignment between Unpaired Multiviews for Egocentric Video Recognition.”  
*International Conference on Computer Vision (ICCV)*, 2023.  
▷ Acceptance rate 26.1%; Top conference in Computer Vision & Pattern Recognition.
- 4     **Q. Wang**, B. Fu, M. Li, J. He, X. Peng, and Y. Qiao  
“Region-aware Arbitrary-shaped Text Detection with Progressive Fusion.”  
*IEEE Transactions on Multimedia (TMM)*, 2022.  
▷ Impact factor 7.3; Top journal in Multimedia.  
▷ The first two authors contributed equally to this work.
- 5     Y. Zou, J. Choi, **Q. Wang**, and J.-B. Huang  
“Learning representational invariances for data-efficient action recognition.”  
*Computer Vision and Image Understanding (CVIU)*, 2022.  
▷ Impact factor 4.5; Prestigious journal in Computer Vision & Pattern Recognition.
- 6     Y. Zheng, **Q. Wang**, and M. Betke  
“Semantic-Based Sentence Recognition in Images Using Bimodal Deep Learning.”  
*IEEE International Conference on Image Processing (ICIP)*, 2021.  
▷ Prestigious conference in Computer Vision & Pattern Recognition.
- 7     **Q. Wang**, Y. Zheng, and M. Betke  
“A Method for Detecting Text of Arbitrary Shapes in Natural Scenes That Improves Text Spotting.”  
*Workshops of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPRW)*, 2020.

## Internship & Collaboration

Sep 2021 – Now	📌 <b>Graduate Research Assistant</b> , University of Delaware.
Jun 2025 – Sep 2025	📌 <b>PhD Research Intern</b> , Dolby Laboratories.
Sep 2021 – Nov 2022	📌 <b>Research Collaboration</b> , Google Research.
Jun 2021 – Aug 2021	📌 <b>Applied Science Intern</b> , Amazon Web Services (AWS).
May 2020 – Aug 2020	📌 <b>Visiting Student</b> , Shenzhen Institute of Advanced Technology.

## Selected Projects

Jun. 2025 - Now	📌 <b>Dolby Laboratories</b> ; Introduced a new approach that improves both the reliability and efficiency of Vision-Language Models (VLMs) for video learning and understanding; <u>work in submission</u> .
May. 2024 - May. 2025	📌 <b>Deep-REAL Lab, University of Delaware</b> ; Introduced a new dataset for video understanding and benchmarked various models, including Multimodal Large Language Models such as Video-LLaVA and VideoLLaMA2.
Sep. 2023 - Aug. 2024	📌 <b>Deep-REAL Lab, University of Delaware</b> ; Designed experiments to investigate the impact of fine-tuning on the prediction rationality of Vision-Language Models (e.g., CLIP); <u>work accepted by AAAI-2025</u> .
Mar. 2023 - Apr. 2024	📌 <b>Dept. of Geog., University of Delaware</b> ; Proposed a method utilizing the “Segment Anything Model (SAM)” to detect mine features in satellite imagery; <u>work accepted by PLOS Sustainability and Transformation, 2025</u> .
Sep. 2021 - Nov. 2022	📌 <b>Google Research (Collaboration)</b> ; Developed a method to align unpaired multiview videos with varying cross-view semantic information, utilizing the capabilities of Large Language Models; <u>work accepted by ICCV-2023</u> .
May. 2021 - Aug. 2021	📌 <b>Amazon Web Services</b> ; Developed a method for image-text retrieval.
May. 2020 - Nov. 2020	📌 <b>Shenzhen Inst. of Adv. Tech., Chinese Academy of Sciences</b> ; Developed a new method for scene text detection; <u>work accepted by TMM-2022</u> .

## Awards


2024	📌 <b>Outstanding Conference Travel Award</b> , CIS department of University of Delaware.
	📌 <b>CIS Distinguished Graduate Student Award</b> , University of Delaware.
2023	📌 <b>Outstanding Conference Travel Award</b> , CIS department of University of Delaware.

## Services

Conference Reviewer	📌 ECCV 2024, CVPR 2025-2026, ICCV 2025, ACMMM 2025, ACMMM 2025 Datasets, BMVC 2024-2025.
Journal Reviewer	📌 IEEE TIP, IEEE TMM, IEEE Access, PLOS ONE.
Program Committee	📌 AAAI 2026
Vol. Conf. Reviewer	📌 CVPR 2023, NeurIPS 2023, AAAI 2024, ICLR 2025.
Vol. Jour. Reviewer	📌 IEEE TAI, IEEE TPAMI, ACM TIST.


## Invited Talk


---

Sep. 2023     **Extraction, Effluent, and Enumeration in Extraglobal Geopolitics;** Off-Earth Geopolitics Workshop. University of Oxford (virtually).

## Skills

---

Coding     Python, Pytorch,  $\LaTeX$ , ...

Large Models     CLIP, SAM, BERT, Video MLLMs (Video-LLaVA, VideoLLaMA2)...