

# Haichao Zhang

Email: [zhang.haich@northeastern.edu](mailto:zhang.haich@northeastern.edu) | Mobile: +1 (857)-313-5537 | Boston, MA

Homepage: <https://zhanghaichao.xyz/> | [LinkedIn](#)

## EDUCATION

---

<b>Northeastern University (NEU)</b>	Boston, MA
• <b>PhD candidate in Computer Engineering</b>	Sep 2022 - Present
• Advisor: IEEE Fellow, Prof. <a href="#">Yun Raymond Fu</a>	
• GPA: 3.83/4.0 • Research Assistant at SMILE LAB • Focusing on computer vision and machine learning	
<b>Zhejiang University (ZJU)</b>	Hangzhou, China
• <b>Master of Science in Computer Science and Technology</b>	Sep 2018 - Mar 2021
• Research Assistant at Media Confrontation Lab • Academic Excellence Scholarship	
<b>Tiangong University (TGU)</b>	Tianjin, China
• <b>Bachelor of Engineering in Electronic and Computer Engineering</b>	Sep 2014 - Jul 2018
• Overall Ranking: 1/63 • Student leader in undergraduate electronic design lab • Recipient of the Dean's Scholarship four times.	
<b>The University of California, San Diego (UCSD)</b>	June 2021 - Sep 2021
• Summer Research Intern • Advisor: Prof. Xiaolong Wang	

## RESEARCH INTEREST

---

Computer Vision and Artificial Intelligence. Aims to explore the potential of GenAI for AIGC, Autonomous Driving, and VLM. Video Understanding, Image and Video Synthesis/Editing, Trajectory Prediction, Multimodal Learning, Diffusion, AIGC, VLM, LLM.

## SELECTED WORK EXPERIENCE

---

<b>Amazon AWS AI Lab</b>   Applied Scientist Intern   Bellevue, WA	Jun.2024- Aug.2024
Mentor: Dr. <a href="#">Shuai Bin</a> , Dr. <a href="#">Taewan Kim</a> , Mr. Xunyu Lin	
Topic: Video Understanding, Vision-Language Model, LLM	
• Conducted research on sketch-to-video generation and human face video synthesis.	
<b>Tencent (腾讯)</b>   GY-Lab   Research Intern   (Top-2 IT company in China)	Jul.2020- May.2021
Mentor: Dr. <a href="#">Gang Yu</a>	
Topic: Image and Video Synthesis, Generative AI	
• Conducted research on sketch-to-video generation and human face video synthesis.	
• Published 2 publications as a result of this research.	
• Contributions were incorporated into the WeChat application, enhancing user engagement and experiences.	
<b>Toyota Motor North America</b>   InfoTech Lab   Lab Project Collaboration	Dec.2022- May.2024
Mentor: Dr. <a href="#">Hongsheng Lu</a>	
Topic: Multimodal Learning, Generative Model for Autonomous Driving	
• Proposed a novel algorithm for predicting severely obstructed trajectories, combining wireless and visual modalities.	
• Innovated a groundbreaking framework for predicting trajectories of objects outside of the line of sight.	
• Published two significant papers on the topics in CVPR and ACM MM, contributing to the field of multimodal learning and generative models for autonomous driving.	

## FIRST-AUTHOR PUBLICATIONS

---

### OOTraj: Out-of-Sight Trajectory Prediction With Vision-Positioning Denoising

Haichao Zhang, Yi Xu, Hongsheng Lu, Takayuki Shimizu, Yun Raymond Fu

Published in IEEE/CVF Computer Vision and Pattern Recognition Conference 2024 (CVPR'24) [[project page](#)][[code](#)][[arxiv](#)][[paper](#)]

### Layout Sequence Prediction from Noisy Mobile Modality

Haichao Zhang, Yi Xu, Hongsheng Lu, Takayuki Shimizu, Yun Raymond Fu

Published in proceeding of 31<sup>st</sup> ACM International Conference on Multimedia (ACM MM'23) [[project page](#)][[arxiv](#)][[paper](#)]

## **Camouflaged Image Synthesis Is All You Need to Boost Camouflaged Detection**

**Haichao Zhang**, Can Qin, Yu Yin, Yun Raymond Fu

Under review in IJCAI [[arxiv](#)]

## **Sketch Me A Video: Video Creation with Hand-Drawn Sketches**

**Haichao Zhang**, Gang Yu, Tao Chen, Guozhong Luo. Preprint available. [[arxiv](#)]

## **Fine-grained Identity-Preserving Landmark Synthesis for Face Reenactment.**

**Haichao Zhang**, Tao Chen, Gang Yu, Weixi Zhang, Youcheng Ben, Bin Fu. Preprint available. [[arxiv](#)]

## **Restore DeepFakes Video Frames via Identifying Individual Motion Styles**

**Haichao Zhang**, Zhe-Ming Lu, Hao Luo, Ya-Pei Feng

Published in Electronics Letters. [[link](#)]

## **HONORS & AWARDS**

---

- ACM MM Travel Grant Award 2023, ACM SIGMM
- PhD Network Travel Grant, Northeastern University
- “TI” Cup National Biomedical Engineering Innovative Design Competition **National First Prize**
- Challenge Cup Competition of Science Achievement in China **Provincial Grand Prize**
- Microcontroller Application Design Competition, Tianjin (**6/453, < 1.3%**) **Provincial First Prize**
- Mobile Application Innovation Contest of Northern China **Northern China First Prize**
- Tianjin IOT Innovation and Engineering Application Design Competition **Provincial First Prize**
- Tianjin Undergraduate Robotics Competition **Provincial First Prize**
- Northern China Robotics Competition **Northern China Second Prize**

## **ACADEMIC SERVICE**

---

### **Conference Reviewer:**

- NeurIPS 2023, NeurIPS 2024
- ACM MM 2024
- ICCV 2023 Workshop on Analysis and Modeling of Faces and Gestures (ICCVW)
- CVPR 2024 AI for Content Creation workshop (CVPRW)

### **Journal Reviewer:**

- Multimedia Tools and Applications (MTA)
- ACM Transactions on Knowledge Discovery from Data (TKDD)

### **Teaching Assistant**

- DS5020 at NEU CS, Summer 2023/Spring 2024
- DS5110 at NEU CS, Fall 2023

## **SKILLS**

---

- Deep Learning Tools: PyTorch, TensorFlow
- Languages& Libraries: Python, C/C++, MATLAB, Verilog, OpenCV, Numpy, Labview
- AI Models: Generative Models, including Video Large Language Models, Diffusion Models, Generative Adversarial Networks (GANs), and Variational Autoencoders (VAE); Transformer models and Neural Radiance Fields (NeRF) for advanced AI applications
- AI Tasks: Image and Video Generation, Trajectory Prediction, Autonomous Driving, Video Understanding, Human-centered Generation

## **RESEARCH EXPERIENCE**

---

### ***Toyota InfoTech Research Project***

*Aug 2023-Nov 2023*

#### ***Out-of-sight Pedestrian Trajectory Prediction***

- Initiated a novel research task focused on predicting the trajectories of out-of-sight objects.
- Developed a vision-positioning denoising model that projects out-of-sight agents into camera frames, enhancing prediction accuracy.

### ***Toyota InfoTech Research Project***

*Jan 2023 - Jun 2023*

#### ***Vision-Wireless Modality Fusion Trajectory Prediction (Layout Sequence Prediction from Noisy Mobile Modality)***

- Broadened the scope of Trajectory Prediction by extending from 2D coordinates to 3D world Layout Sequence representation.
  - Formulated a Denoising Diffusion Model (LTrajDiff) capable of predicting trajectories from extremely short or randomly obstructed observations by fusing visual and noisy sensor mobile modalities.
- Northeastern Univ Research Project** *Sep 2022-Dec 2023*  
*Camouflage Image Synthesis*
- Conducted pioneering research highlighting the challenges in camouflage image synthesis.
  - Proposed and developed a camouflage generator model to create synthetic camouflage images, subsequently improving the performance of camouflage object detection systems.
- UCSD Research Project** *Jul.2021- Sep 2021*  
*Compositional Video Synthesis with Neural Radiance Graphs*
- Integrated Neural Radiance Fields (NeRF) into video prediction, pioneering a compositional approach.
  - Innovated a method for decomposing the initial frame into NeRF components.
  - Developed a Graph Neural Network (GNN) model for predicting the pose and actions of components within a scene.
- Tencent Research Project** *Dec.2020-April.2021*  
*Sketch Me A Video: Interactive Video Creation Through Two Rough Hand-Drawn Sketches.*
- Introduced an interactive system allowing users to create videos from start and end sketches.
  - Proposed a pipeline to abstract input sketches for out-of-domain video creation using feature retrieval and projection.
  - Implemented a Variational Autoencoder (VAE) to ensure smooth, natural motion transitions in videos.
- Tencent Research Project** *June.2020-Sept.2020*  
*Fine-grained Identity-Preserving Landmark Synthesis for Face Reenactment*
- Participated in Tencent face reenactment research
  - Identifying artifacts in synthesized guiding landmarks out of distribution, leading to artifacts through identity leaking.
  - Proposed a landmark refinement module to mitigate identity leakage and improve video generation fidelity.
- Zhejiang University (ZJU) Research Project** *Feb.2019-Dec.2019*  
*Restore DeepFakes Video Frames by Identifying Individual Motion Styles*
- Devised a method to identify DeepFakes by learning individual motion styles, a novel approach at the time.
  - Created a technique to restore the original appearance of speakers in DeepFakes videos.
- The Cyberspace Administration of China (CAC) Research Project** *Dec.2018-Mar.2020*  
*Audio and video deep forensic detection analysis prototype system*
- Led and developed the "DeepFakes Detection System" by reproducing each class of fake face detection algorithms and generation algorithms in that time. Investigate and reproduce various deepfake detection and generation algorithms.
- Shenzhen Research Institute of Big Data, Chinese University of Hong Kong, Research Project** *Mar.2020-June.2020*  
*3D contour-based annotation algorithm for medical images*
- Developed an algorithm for segmenting medical images by adapting a 2D deep snake model to 3D image data.
- Alibaba-ZJU Joint Research Institute of Frontier Technologies Research Project** *Mar.2018-Nov.2018*  
*Research on Multimedia Information Hiding Technology of Unstructured Data*
- Developed "Shared Memory Based Code Hiding Platform" in C++, a tool for research in data security and information hiding.
  - Developed "Video Watermarking" Algorithm.
- Tiangong University, Tianjin, China** *Oct.2015-Sept.2017*  
 Student Leader of Electronic Design Innovation Lab for Undergraduate Students. (selected projects).  
 Mainly focused on embedding platforms and robots.
- Wheelchair Control System via analysis eye-blinking EMG and EEG for paralyzed patients*
- Innovated a method to discern intense eye blink EMG signals within EEG data to navigate wheelchair direction.
  - Analyzed EEG patterns to determine the patient's level of tension or relaxation, controlling wheelchair speed accordingly.
- Sign language recognition system of wearable bending sensor gloves*
- Programmed an embedded microprocessor to process analog signals from bending sensors on gloves.
  - Applied an algorithm to interpret sign language, with results displayed on an application interface.