

Tsai-Shien Chen

☎ (+209) 261-6926 | ✉ tsaishienchen@gmail.com | 🏠 tsaishien-chen.github.io | 📷 tsaishien-chen | 🎓 Tsai-Shien Chen

Education

University of California, Merced

Merced, CA

PH.D. IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

Aug. 2022 - Now

- Work with [Ming-Hsuan Yang](#).
- GPA: 4.00 / 4.00

National Taiwan University

Taipei, Taiwan

M.S. IN ELECTRONICS ENGINEERING

Sep. 2015 - March 2022

- Work with [Shao-Yi Chien](#).
- GPA: 4.30 / 4.30, Rank: 1st / 128.

B.S. IN ELECTRICAL ENGINEERING

- GPA: 4.23 / 4.30, Rank: 5th / 190.

Research Interests

Deep Learning for Computer Vision

- Synthesis, Generation, and Creation for Image, Video, and 3D data
- Representation Learning and Discriminative Model

Selected Publications

1. **Panda-70M: Captioning 70M Videos with Multiple Cross-Modality Teachers** [\[Paper\]](#) [\[Website\]](#)
Tsai-Shien Chen, Aliaksandr Siarohin, Willi Menapace, Ekaterina Deyneka, Hsiang-wei Chao, Byung Eun Jeon, Yuwei Fang, Hsin-Ying Lee, Jian Ren, Ming-Hsuan Yang, Sergey Tulyakov
Conference on Computer Vision and Pattern Recognition (CVPR), 2024
2. **Snap Video: Scaled Spatiotemporal Transformers for Text-to-Video Synthesis** [\[Paper\]](#) [\[Website\]](#)
Willi Menapace, Aliaksandr Siarohin, Ivan Skorokhodov, Ekaterina Deyneka, **Tsai-Shien Chen**, Anil Kag, Yuwei Fang, Aleksei Stoliar, Elisa Ricci, Jian Ren, Sergey Tulyakov
Conference on Computer Vision and Pattern Recognition (CVPR), 2024 **[Highlight]** (Acceptance Rate: 2.8%)
3. **Motion-Conditioned Diffusion Model for Controllable Video Synthesis** [\[Paper\]](#) [\[Website\]](#)
Tsai-Shien Chen, Chieh Hubert Lin, Hung-Yu Tseng, Tsung-Yi Lin, Ming-Hsuan Yang
arXiv preprint, 2023
4. **Incremental False Negative Detection for Contrastive Learning** [\[Paper\]](#)
Tsai-Shien Chen, Wei-Chih Hung, Hung-Yu Tseng, Shao-Yi Chien, Ming-Hsuan Yang
International Conference on Learning Representations (ICLR), 2022
5. **Orientation-aware Vehicle Re-identification with Semantics-guided Part Attention Network** [\[Paper\]](#) [\[Website\]](#)
Tsai-Shien Chen, Chih-Ting Liu, Chih-Wei Wu, Shao-Yi Chien
European Conference on Computer Vision (ECCV), 2020 **[Oral]** (Acceptance Rate: 2.1%)
6. **Viewpoint-Aware Channel-Wise Attentive Network for Vehicle Re-Identification** [\[Paper\]](#)
Tsai-Shien Chen, Man-Yu Lee, Chih-Ting Liu, Shao-Yi Chien
Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2020

Work Experiences

Research Intern at Snap Inc.

Santa Monica, CA

PANDA-70M: CAPTIONING 70M VIDEOS WITH MULTIPLE CROSS-MODALITY TEACHERS [CVPR'24]

May. 2023 - Now

- Collected a dataset with 70M high-quality video-caption pairs through an automatic pipeline including video splitting and captioning process.
- The proposed dataset can significantly improve the training of video captioning, retrieval, and synthesis models.

SNAP VIDEO: SCALED SPATIOTEMPORAL TRANSFORMERS FOR TEXT-TO-VIDEO SYNTHESIS [CVPR'24 (HIGHLIGHT)]

- Built a Text-to-Video model by extending the EDM framework and replacing the U-Net with transformer-based architecture for better efficiency.
- The efficiency allows us to scale up a T2V model with billions of parameters and achieve state-of-the-art visual quality and inference speed.

- Highlighted the unfavorable effect from false negatives for self-supervised contrastive learning.
- Proposed a strategy to incrementally detect more reliable false negatives when the embedding space becomes more semantically structural.

Research Experiences

Ph.D. student at UC Merced

Merced, CA

MOTION-CONDITIONED DIFFUSION MODEL FOR CONTROLLABLE VIDEO SYNTHESIS

Aug. 2022 - Now

- Explored an effective conditional mechanism for video synthesis, where a user can input a starting image frame and a set of strokes.
- Proposed the first stroke-based video diffusion model and achieved state-of-the-art visual quality and motion control fidelity.

BIOMETRIC RECOGNITION AND IDENTIFICATION AT ALTITUDE AND RANGE (BRIAR)

- BRIAR is a large-scale collaborative project among industry, academia, and government with an annual research funding of 12 million dollars.
- Proposed a blind face restoration algorithm which is delivered to several federal organizations and institutions.

Master student at National Taiwan University

Taipei, Taiwan

DENSE CONTRASTIVE PRE-TRAINING ON LARGE-SCALE UNLABELED DATASET FOR SCENE TEXT RECOGNITION

Sep. 2019 - March 2022

- Built a large-scale unlabeled scene text dataset which contains around 8M word boxes captured from 300 metropolises around the world.
- Introduced a novel dense contrastive learning framework to pre-train a strong scene text recognition model on the proposed dataset.

ORIENTATION-AWARE VEHICLE RE-IDENTIFICATION WITH SEMANTICS-GUIDED PART ATTENTION NETWORK [ECCV'20 (ORAL)]

- Proposed a network that can predict the localization of different vehicle views given only image-level labels during training.
- Proposed a distance metric that places greater emphasis on co-occurrence vehicle views when evaluating the feature distance of two images.

VIEWPOINT-AWARE CHANNEL-WISE ATTENTIVE NETWORK FOR VEHICLE RE-IDENTIFICATION [CVPRW'20]

- Proposed an attention mechanism to make the framework channel-wisely reweigh feature maps based on the viewpoint of the vehicle image.
- Explored the interpretability of how our channel-wise attention mechanism improves the learning framework.

Honors & Awards

2024	Graduate Student Opportunity Program Fellowship (top 2 in university) , University of California, Merced
2024	CVPR Travel Grant , Computer Vision and Pattern Recognition
2023	Best Master Thesis of the Year Award , Graduate Institute of Electronics Engineering, National Taiwan University
2022	Honorary Member (top 3% in college) , Phi Tau Phi Scholastic Honor Society
2020	ECCV Oral Paper (2% acceptance rate) , European Conference on Computer Vision
2019	Valedictorian , Department of Electrical Engineering, National Taiwan University
2015-2019	4-time Presidential Award (top 5% in department) , National Taiwan University

Professional Activities

Conference Reviewer	<ul style="list-style-type: none"> - Conference on Computer Vision and Pattern Recognition (CVPR): 2022, 2023, 2024 - International Conference on Computer Vision (ICCV): 2021, 2023 - European Conference on Computer Vision (ECCV): 2022, 2024 - Asian Conference on Computer Vision (ACCV): 2024 - Conference on Neural Information Processing Systems (NeurIPS): 2023, 2024 - International Conference on Learning Representations (ICLR): 2024 - International Conference on Machine Learning (ICML): 2024
Journal Reviewer	<ul style="list-style-type: none"> - IEEE Transactions on Intelligent Transportation Systems (T-ITS) - Neurocomputing
Teaching Assistant	<ul style="list-style-type: none"> - University of California, Merced: CSE-185 Introduction to Computer Vision (Spring 2023) - University of California, Merced: CSE-024 Advanced Programming (Fall 2022) - National Taiwan University: EEE-5053 Computer Vision (Spring 2021)