Vinh Quang TRAN
New York City, New York
Email: tqvinhcs@gmail.com | LinkedIn | GitHub | Google Scholar | : (+1) 631-923-5832 | Visa: H-1B

RESEARCH & TECHNICAL INTERESTS

Computer vision, Generative AI (GenAI), Machine learning, Human action understanding in videos (time series prediction), Scene-text understanding, Object detection, Representation learning, Multimodal learning, and Optimization algorithms.

EDUCATION

Stony Brook University, Stony Brook, NY — Aug 2015 – May 2023
Ph.D. in Computer Science
Advisor: Prof. Minh Hoai Nguyen

University of Science, Ho Chi Minh City, Vietnam — Sep 2009 – Dec 2011
M.S. in Information Systems - (Top 1%)

University of Science, Ho Chi Minh City, Vietnam — Sep 2005 – May 2009
B.S. in Information Technology (Computer Science) - (Top 5%)

INDUSTRY & RESEARCH EXPERIENCE

Senior Applied Scientist, SpreeAI Corporation (Remote, NY) — Jun 2023 – Present

- Led the development of Generative AI–based digital human systems based on stable diffusion for exemplar-guided virtual try-on, multiple pose try-on, targeting realistic garment deformation and identity preservation.
- Trained the end-to-end virtual try-on pipeline, garment alignment, and geometric warping for high-fidelity synthesis.
- Designed and trained pixel-level human and garment parsing models using deep semantic segmentation.
- Collaborated cross-functionally with product and engineering teams to translate research prototypes into production.

Research Assistant, Stony Brook University — Aug 2016 – May 2023

- Developed a language-enhanced video action anticipation model that leverages linguistic knowledge to improve the prediction of future and unseen human actions.
- Cross-Modal Text–Video Retrieval with language and vision models: Exploiting internal knowledge alignment between caption-to-caption linguistic similarity and video-to-video visual similarity for bidirectional video–caption ranking.
- Proposed knowledge distillation frameworks to improve early action anticipation, transferring temporal knowledge from full-sequence action recognition models.
- Developed multi-stage distillation pipelines leveraging bidirectional RNNs to enhance early action recognition accuracy under limited observation.

Visiting Research Scientist, VinAI Research (now Qualcomm AI Research) — Jul 2019 – Aug 2019

- Improved scene text recognition by incorporating lexicon-based constraints during both training and inference.

Teaching Assistant, Stony Brook University — Aug 2021 – Dec 2021

- CSE512: Machine Learning
- CSE527: Introduction to Computer Vision

Research Intern, National Institute of Informatics, Tokyo, Japan — Jan 2013 – Mar 2013

- Developed a video retrieval system leveraging semantic concepts, color layout, and faces for large-scale video exploration.

PUBLICATIONS & SERVICE

- **Publications**: Google Scholar.
- **Reviewer**: BMVC 2020, ICCV 2021, BMVC 2021, CVPR 2022, ECCV 2022, ACCV 2022, CVPR 2023, ICCV 2023, BMVC 2023, CVPR 2025, CVPR 2026

HONORS & AWARDS

- **Special CS Department Chair Fellowship**: ($4,000), Stony Brook University, 2015
- **Video Browser Showdown Winner**: The International Conference on Multimedia Modeling (MMM), 2013

TECHNICAL SKILLS

- **Languages:** Python, C++, Java, C#, MATLAB
- **Deep Learning:** PyTorch(8+ years), TensorFlow(8+ years), Hugging Face(3+ years), Scikit-learn(8+ years)
- **Databases:** MySQL(8+ years), MS SQL Server(8+ years)
- **Others:** Fast-paced environment, strong communication skills and the ability to synthesize complex information.