Minh Tran

+1-479-332-8166 | minhtranquangcr@gmail.com

trqminh.github.io | Linkedin | trqminh | Google Scholar

Research Interest

My research interests are algorithms for visual generative tasks (amodal completion, image inpainting, virtual try-on, etc.); visual perception tasks (object detection, segmentation, and tracking); and vision and language interaction (VLMs, text-guided generation).

Education

- **University of Arkansas** — *Aug 2021 - Expected September 2026*
  Ph.D. in Computer Science, (Advisor: Ngan Le) — AR, USA
- **Viet Nam National University, University of Science** — *Sep 2016 - Oct 2020*
  B.Sc. Honors in Computer Science — Ho Chi Minh City, Vietnam

Skills

Research, Computer Vision, Amodal Perception, Diffusion Models, Image Segmentation, Multiple object Tracking, Detectron2, Deep Learning, Large-scale Training, Slurm, PyTorch, PyTorch Lightning, Python, Linux, Software Engineering.

Publications

Conferences

ICCV 2025 **CT-ScanGaze: A Dataset and Baselines for 3D Volumetric Scanpath Modeling** [paper, code]
Trong Thang Pham, Akash Awasthi, Saba Khan, Esteban Duran Marti, Tien-Phat Nguyen, Khoa Vo, Minh Tran, Son Nguyen, Cuong Tran, Yuki Ikebe, Anh Totti Nguyen, Anh Nguyen, Zhigang Deng, Carol C Wu, Hien Nguyen, Ngan Le
Proceedings of the IEEE/CVF International Conference on Computer Vision, ***Highlight paper****, 2025*

ICCV 2025 **DualFit: A Two-Stage Virtual Try-On via Warping and Synthesis** [paper]
Minh Tran, Johnmark Clements, Annie Prasanna Manoharan, Tri Nguyen, Ngan Le
Proceedings of the IEEE/CVF International Conference on Computer Vision, Retail Vision, 2025

SGSMA 2024 **SolarFormer: Multi-scale transformer for solar PV profiling** [paper]
Adrian De Luis, Minh Tran, Taisei Hanyu, Anh Tran, Liao Haitao, Roy McCann, Alan Mantooth, Ying Huang, Ngan Le
2024 International Conference on Smart Grid Synchronized Measurements and Analytics (SGSMA), 2024

ICRA 2024 **Open-fusion: Real-time open-vocabulary 3d mapping and queryable scene representation** [paper, code]
Kashu Yamazaki, Taisei Hanyu, Khoa Vo, Thang Pham, Minh Tran, Gianfranco Doretto, Anh Nguyen, Ngan Le
2024 IEEE International Conference on Robotics and Automation (ICRA), 2024

IJCNN 2024 **Shapeformer: Shape prior visible-to-amodal transformer-based amodal instance segmentation** [paper, code]
Minh Tran, Winston Bounsavy, Khoa Vo, Anh Nguyen, Tri Nguyen, Ngan Le
2024 International Joint Conference on Neural Networks (IJCNN), 2024

ACCV 2024 **Amodal Instance Segmentation with Diffusion Shape Prior Estimation** [paper]
Minh Tran, Khoa Vo, Tri Nguyen, Ngan Le
Proceedings of the Asian Conference on Computer Vision, 2024

NeurIPS 2024 **Henasy: Learning to assemble scene-entities for interpretable egocentric video-language model** [paper]
Khoa Vo, Thinh Phan, Kashu Yamazaki, Minh Tran, Ngan Le
Advances in Neural Information Processing Systems, 2024

CVPR 2023 **DNA: Deformable Neural Articulations Network for Template-free Dynamic 3D Human Reconstruction from Monocular RGB-D Video** [paper]
Khoa Vo, Trong-Thang Pham, Kashu Yamazaki, Minh Tran, Ngan Le
Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023

ISBI 2022 **SS-3DCapsNet: Self-supervised 3D Capsule Networks for Medical Segmentation on Less Labeled Data** [paper]
Minh Tran, Loi Ly, Binh-Son Hua, Ngan Le
2022 IEEE International Symposium on Biomedical Imaging (ISBI), ***Oral Presentation****, 2022*

ICPR 2022 **3DConvCaps: 3DUnet with Convolutional Capsule Encoder for Medical Image Segmentation** [paper, code]
Minh Tran, Viet-Khoa Vo-Ho, Ngan TH Le
2022 26th International Conference on Pattern Recognition (ICPR), 2022

BMCV 2022 **AISFormer: Amodal Instance Segmentation with Transformer** [paper, code, page]
Minh Tran, Khoa Vo, Kashu Yamazaki, Arthur Fernandes, Michael Kidd, Ngan Le
The 33rd British Machine Vision Conference, 2022, 2022

MICCAI 2021 **Multiple meta-model quantifying for medical visual question answering** [paper]
Tuong Do, Binh X. Nguyen, Erman Tjiputra, Minh Tran, Quang D. Tran, Anh Nguyen
Medical Image Computing and Computer Assisted Intervention–MICCAI 2021, 2021

IV 2021 **Deep Federated Learning for Autonomous Driving** [paper]
Anh Nguyen, Tuong Do, Minh Tran, Binh X Nguyen, Chien Duong, Tu Phan, Erman Tjiputra, Quang D Tran
2022 IEEE Intelligent Vehicles Symposium (IV), 2021

NICS 2020 **Mobile Robot Planner with Low-cost Cameras Using Deep Reinforcement Learning** [paper, code]
Minh Tran, Ngoc Q Ly
2020 7th NAFOSTED Conference on Information and Computer Science (NICS), 2020

Journals

TSG 2025 **S3Former: A Deep Learning Approach to High Resolution Solar PV Profiling** [paper]
Minh Tran, Adrian De Luis, Haitao Liao, Ying Huang, Roy McCann, Alan Mantooth, Jack Cothren, Ngan Le
IEEE Transactions on Smart Grid, 2025

TIA 2025 **SolarFormer++: Multi-scale Transformer for Solar PV Profiling and Obstruction Localization for Degradation Mitigation** [paper, code]
Esteban Duran, Minh Tran, Malachi Massey, Adrian Gracia, Taisei Hanyu, Anh Tran, Roy McCann, Haitao Liao, Jackson Cothren, Meredith Adkins, Chase Rainwater, Ying Huang, Alan Mantooth, Ngan Le
IEEE Transactions on Industry Applications, 2025

IMAVIS 2025 **A2VIS: Amodal-Aware Approach to Video Instance Segmentation** [paper, code, page]
Minh Tran, Thang Pham, Winston Bounsavy, Tri Nguyen, Ngan Le
Image and Vision Computing, 2025

RS 2025 **Land8Fire: A Complete Study on Wildfire Segmentation Through Comprehensive Review, Human-Annotated Multispectral Dataset, and Extensive Benchmarking** [paper, code]
Anh Tran, Minh Tran, Esteban Marti, Jackson Cothren, Chase Rainwater, Sandra Eksioglu, Ngan Le
Remote Sensing, 2025

RS 2024 **Aerialformer: Multi-resolution transformer for aerial image segmentation** [paper, code]
Taisei Hanyu, Kashu Yamazaki, Minh Tran, Roy A McCann, Haitao Liao, Chase Rainwater, Meredith Adkins, Jackson Cothren, Ngan Le
Remote Sensing, 2024

PS 2024 **CarcassFormer: an end-to-end transformer-based framework for simultaneous localization, segmentation and classification of poultry carcass defect** [paper]
Minh Tran, Sang Truong, Arthur FA Fernandes, Michael T Kidd, Ngan Le
Poultry Science, 2024

TMI 2022 **Light-weight deformable registration using adversarial learning with distilling knowledge** [paper, code]
Minh Tran, Tuong Do, Huy Tran, Erman Tjiputra, Quang D Tran, Anh Nguyen
IEEE transactions on medical imaging, 2022

Others

DLMIA 2024 **CapsNet for medical image segmentation** [paper]
Minh Tran, Viet-Khoa Vo-Ho, Kyle Quinn, Hien Nguyen, Khoa Luu, Ngan Le
Deep Learning for Medical Image Analysis, 2024

2023 **aistron: Amodal Instance Segmentation Toolbox and Benchmark** [code]
Minh Tran
2023

Experience

- **Catalyze&Shine** *Aug 2024 - Present*
  Research Collaborator Remote
  - Develop virtual try-on system that helps fashion retailers cut photo production costs and boost sales using AI-generated catalogs from flay lay or pack shot photos.
- **AICV Lab, University of Arkansas** *Aug 2021 - Present*
  Research Graduate Fayetteville, AR
  - *Visual Generative AI*:
    - Developed a video amodal completion model and dataset, (LANCOF), designed to reconstruct objects as complete entities from videos, even when partially occluded. The model employs optical flow as motion priors for consistent mask completion and text guidance as semantic priors for accurate content reconstruction.
    - Developed a two-stage virtual try-on framework via warping and synthesis (DualFit) to faithfully maintaining high-frequency garment details, striking an effective balance between reconstruction accuracy and perceptual realism.
  - *Image Segmentation*:
    - Developed amodal instance segmentation modelscapable of predicting the full shape of objects, including occluded region. Various models with different technique has been developed: Transformer (AISFormer), Shape Priors Modeling (ShapeFormer), Diffusion Models: (AISDiff). These models achieved state-of-the-art performance on amodal benchmarks such as KINS and COCOA.
    - Developed a high-resolution semantic image segmentation framework that beats SOTAs on custom datasets for photovoltaic solar profiling (SolarFormer) and poultry imagery (CarcassFormer).
  - *Multiple object tracking and segmentation*:
    - Developed A2VIS, a multiple-object tracking and segmentation model that integrates amodal segmentation capabilities to enhance tracking robustness. The model outperforms state-of-the-art methods in both object tracking and video amodal segmentation.
    - Developed a gait monitoring system (MiGa) for a small-sized pen housing multiple chickens. The system detects birds and their pose, tracks them over time, and estimates their gait scores. The system is delivered to Cobb Vantress, Inc., with the potential to automate chicken legs disease detection.
  - *Medical Imaging:*
    - Developed a medical image segmentation approach using a hybrid convolution-capsule network (3DConvCaps, ICPR'22) and a self-supervised learning technique for medical image segmentation on low-label datasets (SS-3DCapsNet, ISBI'22).
  - *Self-supervised learning:*
    - Developed self-supervised learning frameworks for medical image segmentation and photovoltaic solar profiling on low-label datasets (SS-3DCapsNet, S3Former).
- **AIOZ AI** *Aug 2020 - Aug 2021*
  Researcher Ho Chi Minh City
  - Indoor Delivery Robot: Developed algorithms for localization module of an indoor self-delivery robot (BeetleBot). The robot won runner-up at Qualcomm Innovation Challenge 2021.
  - Medical Imaging: Developed a light-weight model (LDR-ALDK) for medical image registration. One paper got accepted at Transactions on Medical Imaging

Professional Services

Conference Reviewer at ICCV, CVPR, ACCV, MICCAI, IJCNN
Journal Reviewer at IEEE Transactions on Medical Imaging, Image and Vision Computing

Teaching

CSCE 4133: Algorithms, University of Arkansas, *Teaching Assistant* *Fall 2023*
CSCE 4613: Artificial Intelligence, University of Arkansas, Teaching Assistant Fall 2024
NACME Google AMLI Summer Program, *Teaching Assistant* *Summer 2022*

Honors and Awards

University of Arkansas Graduate Assistantship *2021-2026*
Department of Electrical Engineering and Computer Science Fellowship *Jan 2025*
Asian Conference on Computer Vision International Travel Grant *Dec 2024*
Conference on Computer Vision and Pattern Recognition DEI Award *June 2024*
Rodger S. Kline Chair Graduate Scholarship *Jan 2023*
2nd Place - Qualcomm Innovation Challenge *2021*

References

Dr. Ngan Le, *Associate Professor, EECS, University of Arkansas.*
thile@uark.edu