

Yiming Ma

Postdoctoral Research Associate at King's College London

yiming.m@proton.me | GitHub | Google Scholar | LinkedIn

PROFILE

Computer vision and multimodal learning researcher with six publications, five as first author, and experience spanning medical imaging AI, crowd counting, multi-view multimodal driver monitoring, and industry collaboration.

EXPERIENCE

Postdoctoral Research Associate

Jan 2026 - Present

King's College London | London, UK

- Develop AI-driven precision active-surveillance tools for prostate cancer, spanning patient selection, risk stratification, lesion detection, progression assessment, and personalised surveillance.
- Build patient-specific Multi-Modal Fingerprints that integrate longitudinal prostate MRI, clinicopathological variables, demographic data, and radiology reports.
- Investigate foundation models, self-supervised and semi-supervised learning, and longitudinal architectures, with emphasis on missing data, domain shift, and external validation.
- Contribute to clinician-facing standardised reports that translate model outputs into patient selection and follow-up decisions.

Research Assistant

2022 - 2023

University of Warwick | Coventry, UK

- Refined and extended DAD annotations with nine non-driving activities and prepared data for robust benchmarking.
- Designed a multi-view multimodal driver-monitoring system using masked multi-head self-attention, improving DAD AUC-ROC from 88% to 97% and increasing robustness to missing views or modalities.

Teaching Assistant

2023

University of Warwick | Coventry, UK

- Led labs and tutorials for an undergraduate Python and introductory machine-learning module, covering regression, classification, and neural networks.
- Provided one-to-one and small-group support for programming, machine-learning concepts, code troubleshooting, and experiment design.

EDUCATION

PhD in Mathematics of Systems

Sep 2021 - Sep 2025

University of Warwick | Coventry, UK

Supervisors: Prof. Victor Sanchez and Dr Tanaya Guha.

MSc in Mathematics of Systems (Distinction)

2020 - 2021

University of Warwick | Coventry, UK

Research on guide-dog oestrus-interval prediction and digital twins of urban crowds.

BSc in Mathematics and Applied Mathematics

2016 - 2020

Southern University of Science and Technology | Shenzhen, China

GPA: 3.89/4.00.

PUBLICATIONS AND PREPRINTS

Preprint

ZIP: Scalable Crowd Counting via Zero-Inflated Poisson Modeling

Yiming Ma, Victor Sanchez, Tanaya Guha

Use a zero-inflated Poisson likelihood to separate structurally empty regions from discrete counts, without requiring segmentation masks or Gaussian-smoothed targets.

ICME 2025

CLIP-EBC: CLIP Can Count Accurately through Enhanced Blockwise Classification

Yiming Ma, Victor Sanchez, Tanaya Guha

Turn each local count into an unambiguous integer class, match local CLIP image features to natural-language count prompts, and recover a density map from the class probabilities.

ICME 2025

Interact with me: Joint Egocentric Forecasting of Intent to Interact, Attitude and Social Actions

Tongfei Bian, Yiming Ma, Mathieu Chollet, Victor Sanchez, Tanaya Guha

Represent face, body, and hand keypoints as graphs, model their temporal evolution, and forecast intent, attitude, and action through a chained task hierarchy.

CVPRW 2023

Robust Multiview Multimodal Driver Monitoring System Using Masked Multi-Head Self-Attention

Yiming Ma, Victor Sanchez, Soodeh Nikan, Devesh Upadhyay, Bhushan Atote, Tanaya Guha

Fuse top/front and depth/infrared features with source-aware self-attention, and randomly mask feature patches during training to simulate missing views or modalities.

ICIP 2022

FusionCount: Efficient Crowd Counting via Multiscale Feature Fusion

Yiming Ma, Victor Sanchez, Tanaya Guha

Reuse 15 intermediate VGG-16 feature maps, learn spatially varying fusion weights from their contrasts, and decode them with a lightweight saliency-preserving reduction block.

Preprint

Real-Time Driver Monitoring Systems through Modality and View Analysis

Yiming Ma, Victor Sanchez, Soodeh Nikan, Devesh Upadhyay, Bhushan Atote, Tanaya Guha

Use only the last frame of each 16-frame clip, then measure how top/front views and depth/infrared modalities trade accuracy for latency.

SERVICE

Reviewer for IEEE TNNLS, IEEE TMM, IEEE SPL, ECCV, ACM MM, CVPRW, ICME, WACV, BMVC, and ICIP.

SKILLS

Deep learning	Attention, vision transformers, prompt tuning, CLIP, contrastive learning, multimodal alignment, multimodal fusion
Frameworks	PyTorch, TIMM, OpenCLIP, Transformers, TensorBoard, Optuna
Python	NumPy, SciPy, scikit-learn, pandas, OpenCV, Matplotlib, Seaborn
Mathematics and statistics	Probability, statistical inference, optimisation, stochastic processes, time-series modelling, survival analysis, computational statistics, real, complex, functional and Fourier analysis, measure theory
Development tools	SSH, Git, Linux, LaTeX, Markdown, Microsoft Word
Languages	Mandarin Chinese (native); English (IELTS 8.0/9.0)