

Bor-Shiun Wang

PHD CANDIDATE · COMPUTER SCIENCE AND ENGINEERING

☎ (+886) 978-561-844 | ✉ eddie1998221@gmail.com | 🏠 eddie221.github.io | 🎓 Bor-Shiun Wang

Summary

PhD candidate in Computer Science at National Yang Ming Chiao Tung University, specializing in interpretable AI for vision and multimodal systems. My research focuses on understanding and improving how deep models encode knowledge, with the goal of making AI systems more transparent and reliable. I am interested in translating interpretability research into practical solutions for real-world AI systems.

Publications

- **Bor-Shiun Wang**, Chien-Yi Wang*, Wei-Chen Chiu*, “Uncovering the Why: Interpretable CLIP Similarity via Dual Modalities Decomposition”, Under Peer Review, 2026.
- **Bor-Shiun Wang**, Chien-Yi Wang*, Wei-Chen Chiu*, “MCPNet++: An Interpretable Classifier via Multi-Level Concept Prototypes”, In IEEE Transactions on Pattern Analysis & Machine Intelligence (TPAMI), 2026.
- **Bor-Shiun Wang**, Chien-Yi Wang*, Wei-Chen Chiu*, “MCPNet: An Interpretable Classifier via Multi-Level Concept Prototypes”, In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2024.
- **Bor-Shiun Wang***, Ping-Yang Chen*, Yi-Kuan Hsieh, Jun-Wei Hsieh, Ming-Ching Chang, JiaXin He, Shin-You Teng, HaoYuan Yue, Yu-Chee Tseng, “PRB-FPN+: Video Analytics for Enforcing Motorcycle Helmet Laws”, In IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW) on the AI City Challenge, 2023.
- **Bor-Shiun Wang**, Jun-Wei Hsieh, Yi-Kuan Hsieh, Ping-Yang Chen, “COFENet: Co-Feature Neural Network Model for Fine-Grained Image Classification”, In IEEE International Conference on Image Processing (ICIP), 2022.
- **Bor-Shiun Wang**, Jun-Wei Hsieh, Ping-Yang Chen, Ming-Ching Chang, Lipeng Ke, Siwei Lyu, “LDW-Pooling: Learnable Discrete Wavelet Pooling for Convolutional Networks”, The British Machine Vision Conference (BMVC), 2021.

Research Experience

Interpretable MLLMs/LLMs

SKILL: PYTORCH, MLLMs, LLMs, HUGGING FACE

On-going

Sep. 2025 - Present

- Investigating **knowledge localization** in LLMs/MLLMs to understand how information is encoded across layers.
- Designing **probing and editing experiments** to analyze the relationship between internal representations and model outputs.
- Providing **preliminary insights** into the role of internal representations in model behavior.

Interpretable cross-modality similarity (CLIP-DMD)

SKILL: PYTORCH, DISTRIBUTED TRAINING, OPENAI API, MLLMs, HUGGING FACE

Under Peer-reviewed

Sep. 2024 - Aug. 2025

- Proposed a **cross-modal decomposition framework** that explains CLIP similarity by disentangling visual-textual interactions in a shared concept space.
- Achieved **20%–30% improvement** over prior methods (e.g., SpLiCE) on cross-modal retrieval and alignment metrics.
- Developed a **generalizable framework** applicable to multiple downstream tasks requiring interpretable multimodal reasoning.

Enhanced Interpretable Classifier (MCPNet++)

SKILL: PYTORCH, OPENAI API, DISTRIBUTED LEARNING, MLLMs, IMAGE EDITING, HUGGING

TPAMI 2026

FACE

Aug. 2023 - Aug. 2024

- Extended MCPNet to **transformer-based architectures**, enabling interpretable classification beyond CNNs.
- Achieved over **80% alignment** between discovered concepts and human-interpretable semantics, improving explanation faithfulness.
- Improved classification performance by **1.12%** on average across multiple benchmark datasets while maintaining interpretability.

Multi-Level Concept Prototype Network (MCPNet)

SKILL: PYTORCH, OPENAI API, DISTRIBUTED LEARNING, MLLMS

CVPR 2024

Aug. 2022 - Aug. 2023

- Proposed a **hierarchical prototype-based architecture** that bridges low-level visual features and high-level semantic concepts.
- Enabled **concept-level explanations** for model predictions without sacrificing classification performance.
- Learned **generalized concept representations** that can be leveraged for downstream tasks.

College/University Student Research Project

SKILL: PYTORCH, DISTRIBUTED LEARNING, LABELME

ICIP 2022

Jul. 2019 - Feb. 2020

- Proposed COFENet, a deep learning architecture for **fine-grained image classification** under high intra-class similarity.
- Designed a **spatial-structural relation module** to capture pairwise and geometric relationships between feature channels.
- Improved classification robustness for **small, blurry, and texture-rich objects** through structured feature modeling.

Applied Projects

Cassava Leaf Disease Classification

SKILL: PYTORCH

Nov. 2020 - Feb. 2021

- Developed a deep learning model for fine-grained plant disease classification under high intra-class similarity.
- Applied **soft-label training** and **mix-up augmentation** to capture inter-class relationships.
- Achieved a public leaderboard score of **0.8957**, improving the baseline by **~4.5%**.

Education

National Yang Ming Chiao Tung University (NYCU)

PH.D. IN INSTITUTE OF COMPUTER SCIENCE AND ENGINEERING

Hsinchu, Taiwan

Feb. 2022 - Present

- Focusing on explainable AI, interpretable AI.

National Chiao Tung University (NCTU)

M.S. IN INSTITUTE OF INTELLIGENT SYSTEMS (GPA: 4.24 / 4.30)

Hsinchu, Taiwan

Feb. 2020 - Jan. 2022

- Focusing on computer vision and machine learning.

National Taiwan Ocean University (NTOU)

B.A. IN DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (GPA: 3.89 / 4.00)

Keelung, Taiwan

Sep. 2016 - Jan. 2020

- Conduction a College/University Student Research Application
- Received the Academic Excellence Award, Fall 2018 and Spring 2019

Scholarship

Subsidy for Domestic Graduate Students to Attend International Academic Conferences

Taiwan

NATIONAL SCIENCE AND TECHNOLOGY COUNCIL (NSTC)

2024

- Awarded for presenting the paper "MCPNet" at the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2024) in Seattle, USA.

Technical Skills

Programming C/C++, Python, HTML/CSS, SQL

AI/ML Deep Learning, Computer Vision, Vision-Language Models, Explainable AI

Frameworks PyTorch, Hugging Face, Scikit-learn, NumPy, OpenAI API

Cloud/DevOps Git/Github, Docker

Reviewer Services

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
- International Conference on Learning Representations (ICLR)
- Neural Information Processing Systems (NeurIPS)
- IEEE Journal on Emerging and Selected Topics in Circuits and Systems (JETCAS)