

Xin Cai

✉ xincai00@gmail.com • 🌐 totalvariation.github.io

What lies behind us and what lies before us are
tiny matters compared to what lies within us.

Ralph Waldo Emerson

Education

Sichuan University (Project 985 & 211)

Chengdu, China

M.Sc. (Degree Type: equivalent to MPhil), Computer Science & Technology

Sept. 2017 - Jun. 2020

Overall GPA: 3.58/4.0

Master Thesis: On Deep Learning-Based Dense Pixel-wise Prediction.

Supervisor: Prof. Yi-Fei Pu

The thesis aims at studying a wide range of computer vision tasks, such as semantic segmentation, human pose estimation, depth estimation, and optical flow, in a unified framework, referred to as dense pixel-wise predictions. Besides, a novel method was proposed to efficiently generate high-resolution predictions, considered as one of the greatest challenges for dense pixel-wise predictions.

Shanghai Normal University

Shanghai, China

B.Eng., Telecommunications Engineering

Sept. 2011 - Jun. 2015

Overall GPA: 3.61/4.0

Bachelor Thesis: On HEVC-Based Video Coding Technology.

Thesis Advisor: Assoc. Prof. Qian Zhang

The main contributions of the thesis include: (1) A fast Coding Unit (CU) partition algorithm based on depth levels predicted from spatially adjacent CUs was proposed; (2) An intra-frame candidate mode selection algorithm was proposed to reduce redundant information in the spatial domain, thereby improving the coding efficiency.

Extracurricular Research Experience:

Research Advisor: Prof. Yan-Hua Long

Building an automatic speaker verification system based on Gaussian Mixture Models (GMMs).

Employment

University of Trento & FBK

Trento, Italy

PhD Candidate, Withdrawal

Nov. 2024 - Dec. 2024

Recipient of University Scholarship

Ulster University

Belfast, UK

PhD Candidate, Cancelled, Artificial Intelligence

Sept. 2021 - Dec. 2023

Recipient of Department for the Economy (DfE) International Studentship

PhD Project: Large-Scale Crop Type Mapping from Satellite Image Time Series

In the initial phase of my PhD, I explored the possibilities of constructing deep generative classifiers for analysing Satellite Image Time Series (SITS). In the subsequent stage, I proposed to reformulate the representation of SITS to bridge the gap between SITS analysis and broader research domains, such as time series analysis and computer vision.

Sunrise Asset Management Co., Ltd

Shanghai, China

Broker

Jul. 2015 - Jun. 2017

Stupid, but can't sweep it under the rug

Publications

Cai, X., 2025. One Framework to Rule Them All: Unifying RL-Based and RL-Free Methods in RLHF. arXiv preprint arXiv:2503.19523. [\[Paper Link\]](#)

Cai, X., Bi, Y., Nicholl, P. and Sterritt, R., 2023. Revisiting the Encoding of Satellite Image Time Series. BMVC **Oral** [Work done at Ulster University] [\[Paper Link\]](#), [\[Code Link\]](#)

Cai, X., Bi, Y. and Nicholl, P., 2022. Tampered VAE for Improved Satellite Image Time Series Classification. arXiv preprint arXiv:2203.16149. [Work done at Ulster University] [\[Paper Link\]](#)

Cai, X. and Pu, Y.F., 2019. FlatteNet: a simple and versatile framework for dense pixelwise prediction. IEEE Access, 7, pp.179985-179996. [\[Paper Link\]](#), [\[Code Link\]](#)

Competitions

- AI4Food Security Challenge [\[South Africa Track\]](#) [\[Germany Track\]](#)
The goal of this challenge is to classify crop types based on time series data from Sentinel-1, Sentinel-2 and Planet Fusion Monitoring Data. My submissions (completed all by myself) marked as AIRC_Ulster are ranked 8th and 7th on the South Africa and Germany Track.
- PASCAL VOC 2012 Challenge Competition "comp5" Leaderboard ([\[Link:XC-FLATTENET\]](#))
PASCAL VOC 2012 Challenge is one of the most widely used datasets for benchmarking visual recognition algorithms. There are three main object recognition competitions: classification, detection, and segmentation. My submissions (completed all by myself) are ranked 1st and 3rd out of a total of 32 submissions marked as public on the segmentation leaderboard. (Up to 00:00 24th Apr. 2021)

Skills

Research Interests: Deep Generative Models, Multimodal Learning, Large Language Models, Reinforcement Learning, Time Series Analysis, and Computer Vision

Programming Languages: Python, C/C++, Matlab, Shell Scripting

Deep Learning Libraries: PyTorch, TensorFlow

Systems: LINUX, SLURM

Writing Tools: \LaTeX

Languages: Fluent English, Native Chinese