Bo (Bryan) Cao

Ph.D. Candidate | Multimodal AI | Efficient Machine Learning

@ boccao@cs.stonybrook.edu   Stony Brook, NY, United States   scholar.google.com/citations?user=czcKhLAAAAAJ

 github.com/bryanbocao   in linkedin.com/in/bryanbocao   bryanbocao.github.io

Ph.D. researcher and industry research scientist with extensive experience at Nokia Bell Labs. Specializing in multimodal learning, vision–wireless association, efficient machine learning, few-class neural networks, and model merging. Proven record of translating theoretical research into real-world systems for edge AI and robotics.

🎓 EDUCATION

2026 (Expected)	**Ph.D. Computer Science** (GPA : **3.9/4.0**) Stony Brook University, Stony Brook, NY, US
	Research : Vision-Mobile Multi-Modal Association and Learning, Efficient ML, Advisor : Prof. Shubham Jain
2018	**M.S. Computer Science** (GPA : **3.9/4.0**) University of Colorado Boulder, Boulder, CO, US
	Research : Deep Learning on Human Robot Interaction, Advisor : Prof. Dan Szafir
2013	**MSc Computer Science** (**Merit**) The University of Sheffield, Sheffield, SY, UK
2011	**B.Eng. Computer Science and Technology** Guang Dong University of Technology, Guangzhou, GD, China

📖 SELECTED PUBLICATIONS AND PREPRINTS

> **Few-Class Arena : A Benchmark for Efficient Selection of Vision Models and Dataset Difficulty Measurement**
> Bryan Bo Cao, Lawrence O'Gorman, Michael Coss, Shubham Jain ICLR 2025

> **A Landmark-Aware Visual Navigation Dataset for Map Representation Learning**
> Faith Johnson*, Bryan Bo Cao*[1], Kristin Dana, Shubham Jain, Ashwin Ashok HRI2025

> **A Lightweight Measure of Classification Difficulty from Application Dataset Characteristics**
> Bryan Bo Cao, Lawrence O'Gorman, Michael Coss, Shubham Jain ICPR 2024

> **OVIDA : Orchestrator for Video Analytics on Disaggregated Architecture**
> Manavjeet Singh, Sri Pramodh Rachuri, Bryan Bo Cao, Abhinav Sharma, Venkata Bhumireddy, Francesco Bronzino, Samir R. Das, Anshul Gandhi, Shubham Jain SEC 2024

> **Representation Similarity : A Better Guidance of DNN Layer Sharing for Edge Computing without Training**
> Bryan Bo Cao, Lawrence O'Gorman, Michael Coss, Shubham Jain MobiCom 2024 S3 Workshop

> **Feudal Networks for Visual Navigation**
> Faith Johnson, Bryan Bo Cao, Kristin Dana, Shubham Jain, Ashwin Ashok CVPR2024 Embodied AI Workshop

> **ViFiT : Reconstructing Vision Trajectories from IMU and Wi-Fi Fine Time Measurements**
> Bryan Bo Cao, Abrar Alali, Hansi Liu, Nicholas Meegan, Marco Gruteser, Kristin Dana, Ashwin Ashok, Shubham Jain
> arxiv.org/pdf/2310.03140 MobiCom 2023 ISACom Workshop

> **ViTag : Online WiFi Fine Time Measurements Aided Vision-Motion Identity Association in Multi-person Environments**
> Bryan Bo Cao, Abrar Alali, Hansi Liu, Nicholas Meegan, Marco Gruteser, Kristin Dana, Ashwin Ashok, Shubham Jain
> github.com/bryanbocao/vitag **Best Demonstration Award**🏆 SECON 2022

> **Vi-Fi : Associating Moving Subjects across Vision and Wireless Sensors**
> Hansi Liu, Abrar Alali, Bryan Bo Cao, Nicholas Meegan, Hongyu Li, Marco Gruteser, Shubham Jain, Kristin Dana, Ashwin Ashok
> github.com/vifi2021/Vi-Fi IPSN 2022

> **StatsMerging : Statistics-Guided Model Merging via Task-Specific Teacher Distillation**
> Ranjith Merugu*, Bryan Bo Cao*, Shubham Jain arxiv.org/pdf/2506.04567 *(Under Submission)*

> **YOPO-Nav : Visual Navigation using 3DGS Graphs from One-Pass Videos**
> Ryan Meegan, Adam D'Souza, Bryan Bo Cao, Shubham Jain, Kristin Dana arxiv.org/pdf/2512.09903 *(Under Submission)*

🏆 HONORS AND AWARDS

DEC 2024	Few-Class project was selected for Bell Labs **Breakthrough Project** (Publisherd in ICPR 2024, ICLR 2025).
AUG 2024	Awarded **Outstanding Innovation Award (top 10%)** at Nokia Bell Labs.
AUG 2023	Awarded **Outstanding Student Research Award (top 10%)** at Nokia Bell Labs.
SEP 2022	Awarded **Best Demonstration Award** for the ViTag paper in SECON 2022.
SEP 2017	Awarded **Beverly Sears Graduate Student Grant Award** at University of Colorado Boulder.
JUN 2010	Awarded **First Class Scholarship for Academic Excellence (top 3%)** at Guangdong University of Technology.
JUN 2010	Awarded **Second Class Scholarship for Students' All-Round Development (top 8%)** at Guangdong University of Technology.

---

1. * : Equal contribution

🖼️ Industry Research Experience

Fall 2025 — Jun 2021 | Research Intern / Co-op, Nokia Bell Labs, Murray Hill, NJ, US
- Mentor : Dr. Lawrence O'Gorman, Dr. Xin Yuan. **Manager :** Michael Coss. Featured LinkedIn Post (click).
- Designed **Structural Pruning** for **CNN (ResNet)**, and **Transformer (ViT)**. Experiments on pruning 20% parameters while retaining the performances at 89.80% and 92.03% of the original model based on neuron dependency graph, respectively.
- Proposed **Class-Aware Structural Pruning (CASP)** using Grad-CAM gradients for pruning extraneous parameters (large to few classes). CASP outperformed the magnitude-based method by pruning **2.5x** more filters, **11%** reduction of computation and **36%** fewer number of parameters.
- Developed **Class-Similarity Guided Neural Network Model Selection**, a novel strategy guided by Class-similarity to select models and class group for an application specifications. It assisted in selecting a sub-YOLO model **42%** smaller than the baseline YOLOv5-nano for a 3-class group, **85%** smaller for a 2-class group in a real-world robot application.
- Integrated tracking evaluation into BMVC codec system implementation.
- Improved CenterTrack on lower quality videos.

`Few-Class` `Pruning` `Class-Similarity` `CNN` `Transformers` `ViT` `YOLO` `Python` `PyTorch`

Aug 2019 — May 2019 | Intern, Sunrise Technology at CEWIT, Stony Brook, NY, US
- Mentor : Dr. Dantong Yu
- Improved data collection, model training and deployment pipeline for autonomous driving.
- Trained and deployed an End-to-End CNN model to drive the Jetson TX1 race car autonomously.

`End-to-end Training` `Python` `Keras`

Aug 2017 — May 2017 | Research Intern, Ericsson Silicon Valley, Santa Clara, CA, US
- Mentor : Alvin Jude
- Developed an augmented reality collaboration system using HoloLens.

`AR` `JavaScript` `WebGL` `Node.js`

Nov 2014 — May 2014 | Test Engineer, IBM International System Technology Co. Ltd (ISTC), Shenzhen, Guangdong, China
- **Manager :** Alex Zeng, **Team Lead :** Kayla Zhang
- Tested enterprise System X servers in Linux environments.
- Implemented Front-end work of Redfish Project for report auto-generation.

`Python` `JavaScript` `Linux` `web.py`

🏛️ Academic Research and Teaching Experience

May 2026 — Jan 2021 | Research Assistant, Stony Brook University, Stony Brook, NY, US
- **Reality-Aware Networks** Project Link : ashwinashok.github.io/realityawarenetworks
  Advisor : Shubham Jain. Co-PIs : Ashwin Ashok, Kristin Dana & Marco Gruteser (now @Google)
  Leverage deep neural networks to the challenges of
  (a) Multimodel Association on visual and phone tracklets from camera and smartphone domains using IMU and FTM (Bidirectional LSTM, Encoder-decoder, Contrastive Learning, 2021-2022);
  (b) Reconstructing visual trajectories from phone data (Transformer, 2023);
  (c) Robot Visual Navigation (Hierarchical RL, Human-in-the-loop, Topological Map, Metric Learning, 3D Gaussian Splatting 2024-2025).
- Led research on **vision–wireless association** using **camera**, **IMU**, and **Wi-Fi FTM** data.
- Developed sequence models (BiLSTM, Transformers, contrastive learning) for cross-modal trajectory reconstruction.
- Investigated Robot navigation using hierarchical RL, metric learning and 3D Gaussian Splatting.

`LSTM` `Transformer` `IMU` `Wi-Fi` `FTM` `Contrastive Learning` `RL` `Multimodal Learning` `Multimodal Association` `Vision Wireless Fusion` `Visual Navigtion` `Feudal Network` `3D Gaussian Splatting` `PyTorch` `Transformers`

Dec 2020 — Aug 2018 | Teaching Assistant, Stony Brook University, Stony Brook, NY, US
- Fall 2020   CSE114   Introduction to Object Oriented Programming, Prof. Praveen Tripathi
- Fall 2019   CSE527   Introduction to Computer Vision, Prof. Dimitris Samaras
- Spring 2019 CSE216-01 Programming Abstractions, Dr. Ritwik Banerjee
- Fall 2018   CSE219-02 Computer Science III, Dr. Richard McKenna

🧪 TECHNICAL SKILLS

Programming	Python, Java
Machine Learning	PyTorch, TensorFlow, Keras, Scikit-learn
Computer Vision	CNNs, Vision Transformers, YOLO, OpenCV
Systems	Docker, Linux
Simulation	Habitat-Sim, Habitat-Lab

📇 DATASETS

YOPO-Campus (Under Sub. 2026)	A visual navigation dataset collected on Busch Campus, Rutgers University (New Jersey, United States) for visual navigation and visual place recognition. Link : ajd324.github.io
LAVN (HRI 2025)	A Landmark-Aware Visual Navigation (LAVN) dataset to allow for supervised learning of human-centric exploration policies and map building. Link : huggingface.co/datasets/visnavdataset/lavn
Vi-Fi (IPSN 2022)	A large-scale multi-modal dataset to for vision-wireless systems. Link : sites.google.com/winlab.rutgers.edu/vi-fidataset/home

💼 COMMUNITY SERVICE

Reviewer	NeurIPS, ICLR, CVPR, ECCV, WACV, IMWUT, MM, TIP, TNNLS, PR

🎓 MENTORSHIP EXPERIENCE

Present – Jan 2025

Ranjith Merugu, 2025 SUMMER RESEARCH ASSISTANT @PICASSOLAB,
Implemented Weight Statistics-Guided Model Merging by Task-Specific Teacher Distillation and benchmarked on Computer Vision and NLP datasets. In submission to **ICLR 2026**.
`Model Merging` `Python` `PyTorch`

Dec 2024 – Aug 2024

Venkata Nithish Reddy Bhumireddy, 2024 SUMMER INTERN, Full-Time SWE @Snowflak
Measured CKA-based representation similarity on YOLOv11 for object detection and YOLOv11-seg for segmentation.
`CKA Similarity` `YOLO` `Segmentation` `Python` `PyTorch`

May 2024 – Jan 2023

Abhinav Sharma, 2023 SUMMER INTERN, 2024 SPRING CO-OP @NOKIA BELL LABS, Full-Time AI Models Engineer, Efficient Generative AI @AMD
> Measured CKA-based representation similarity on EfficientNet and EfficientDet, Object detection, license plate recognition, OCR and pose estimation evaluation on Jetson Nano.
> Published a paper at **MobiCom 2024 S3 Workshop**.

`CKA Similarity` `Efficient AI` `Object Detection` `OCR` `Pose Estimation` `Jetson Nano` `Python` `PyTorch`

Dec 2022 – Jan 2022

Purna Rao Mallepaddi, 2022 SUMMER INTERN @META, ExMLE @Whoop, MLE @Tiktok Search
> Expanded Computer Vision application pipelines, including OCR on vehicle plate recognition, pedestrian detection and counting in PyTorch, Kubernetes on Jetson Nano.

`Efficient AI` `Object Detection` `OCR` `Pose Estimation` `Jetson Nano` `Python` `PyTorch`

Dec 2021 – Aug 2021

Omkar Manjrekar, 2022 SUMMER INTERN @META, Ex-ThoughtSpot, MLE @Otter.ai
> Refactored CV codebase. Explored fast super-resolution in PyTorch, Kubernetes on Jetson Nano.

`PyTorch` `Kubernetes` `Jetson Nano`