Bo Shang

bshang@ccny.cuny.edu
Field Data Engineer | Data Ops & Infrastructure +1-929-520-0488
cnpcshangbo.github.io | linkedin.com/in/bo-shang New York, NY

Field-first data engineer with years of hands-on experience deploying **real-world sensor data collection systems**. Build robust, configuration-portable data pipelines from field sensors through cloud infrastructure. Translate field constraints into working solutions: multi-modal data ingestion (LiDAR, RGB, thermal), quality-assurance workflows, and deployment infrastructure for autonomous systems and infrastructure monitoring.

Experience

- **City University of New York — AI & Mobility Research Lab** New York, NY
  Field Data Engineer — Autonomous Vehicles *Jul 2025 – Present*
  - Operationalized **fixed roadside LiDAR data collection**; built a **configuration-portable inference pipeline** that trains on one site's sensor setup and infers on another with different beam counts and frame rates
  - Designed data ingestion from roadside LiDAR sensors; implemented multi-frame point-cloud reconstruction to densify per-vehicle representations for downstream classification
  - Established data labeling and annotation workflows for vulnerable road user (VRU) classification; coordinated with field teams to validate sensor coverage and detection performance
  - Built reproducible dataset versioning and format-conversion infrastructure across heterogeneous sensor configurations
- **City University of New York — Robotics Lab** New York, NY
  Field Data Engineer — Infrastructure Inspection *Dec 2022 – Dec 2024*
  - Owned the end-to-end data pipeline for **robotic bridge inspection**: field data collection (drone/robot imagery + sensor metadata), ingestion, storage, processing, and delivery to end users
  - Built a custom **WebODM-based data platform** integrating automated segmentation, 3D reconstruction, interactive visualization, and crack measurement for end-user delivery
  - Trained CNNs for crack, spalling, and stain detection on robot-collected imagery; pushed models to **AWS** for scalable cloud inference
  - Contributed to a contrastive-learning approach for robust defect mapping on concrete slabs using impact-echo data; equal-first-author IROS 2025 / IEEE T-ASE 2025 paper
- **Missouri University of Science and Technology** Rolla, MO
  Embedded Data Systems Engineer *Jan 2020 – Nov 2022*
  - Deployed real-time vision processing on **NVIDIA Jetson edge hardware** (Nano, TX2, Xavier) for field robotics; optimized inference pipelines for power and latency constraints
  - Built data quality evaluation framework for visual odometry systems; integrated motion-capture ground truth (VICON) for validation in GPS-denied environments
  - Designed and tested data collection protocols for autonomous flight in harsh conditions; characterized sensor reliability and failure modes
- **Vaughn College of Aeronautics and Technology** New York, NY
  Technical Instructor *2023 – 2024*
  - Taught applied courses on AI systems, research methodology, and robotics control; mentored students on practical deployment of autonomous systems
- **University of California, Merced** Merced, CA
  Field Robotics Systems Engineer *Sep 2015 – Sep 2017*
  - Co-led development of **autonomous robotic platform for GPS-denied exploration**; integrated inertial + visual odometry, sensor fusion, and robust localization for cave deployment
  - Built custom embedded data-collection software on Linux UAVs; validated sensor data in real-world environments with high noise and degraded signals

Key Data Operations Projects

- **Heterogeneous Sensor Data Ingestion (2024–present)**: Built configuration-portable pipelines that handle LiDAR variants differing in beam count and frame rate, training on one sensor configuration and inferring on another. Validated end-to-end data consistency across configurations.
- **Bridge Inspection Data Platform (2022–present)**: Designed and operated a WebODM-based system for drone/robot imagery ingestion, preprocessing, 3D reconstruction, and interactive delivery to end users (city engineers, inspectors).
- **Defect Detection & Cloud Deployment (2022–present)**: Trained CNNs for crack/spalling/stain detection from drone imagery and deployed them on AWS for scalable inference, integrated into the inspection platform.
- **Edge Vision on NVIDIA Jetson (2020–2022)**: Designed image-processing algorithms for girder detection (OpenCV, C++) on NVIDIA Jetson under power and latency constraints, as part of a bridge-inspection drone.

Core Competencies

- **Data Ops & Pipelines**: End-to-end sensor data ingestion, ROS / ROS 2, MCAP, Protobuf, ETL, data versioning, format conversion, quality assurance, metadata management
- **Cloud & Infrastructure**: AWS (S3, EC2, SageMaker), Docker, CI/CD (Git/GitHub Actions), Linux/Ubuntu, model packaging and serving
- **Edge Computing**: NVIDIA Jetson (Nano/TX2/Xavier), embedded Linux, real-time and power/latency constraints
- **Field Engineering**: Multi-modal sensor integration (LiDAR, RGB, thermal, IMU), sensor validation, GPS-denied localization, robustness under real-world constraints
- **Programming**: Python (primary), C++, MATLAB, Fortran, Swift
- **Data Formats**: JSON, CSV, format conversion across sensor configurations, dataset versioning

Education

- **Northeastern University & University of California, Merced** — Shenyang, China / Merced, CA
  Ph.D. in Pattern Recognition & Intelligent Systems — *Aug 2013 – Dec 2020*
- **Northeastern University** — Shenyang, China
  M.Eng. in Pattern Recognition & Intelligent Systems; B.Eng. in Automation — *Aug 2007 – Dec 2013*

Patents & Recognition

- **US 12,296,994** (granted May 2025) — Unmanned vehicle system design and deployment
- **Best System Control** & **Best Mission Planning Awards** — International Aerial Robotics Competition (AUVSI), 2014