

# M. Alfi Hasan, PhD

Current Position: *Staff Data Scientist at Walmart Global Tech*

**Email** malfihasan19@outlook.com  
**Phone** +1-347-265-5476  
**Address** Dallas, Texas [available without sponsorship]

**Website** [www.malfihasan.com](http://www.malfihasan.com)  
**GitHub** [github.com/malfihasan](https://github.com/malfihasan)  
**LinkedIn** [malfihasangeo](https://www.linkedin.com/in/malfihasangeo)

Data scientist and experienced programmer specialized in computer vision, deep learning models on edge devices, multi-spectral remote sensing, and large vision models across diverse domains, with over 12 years of experience (5+ years in academia, 7+ years in industry).

- As a **Tech Lead** at PhysicalAI team, led the development of device-efficient, gen-ai powered ML models on IoT devices for the barcode intelligence of Walmart and currently scaling the solution to 600+ stores.
- Have domain expertise and hands-on **experience with advanced ML/DL architectures**, such as Generative Adversarial Networks (GANs), the Segment Anything series, the YOLO series, CLIP, Grounding DINO, and Transformers, including the development of custom deep learning architectures across various disciplines.
- Have **advanced proficiency in Python** (including frameworks like PyTorch and TensorFlow). Additionally, have software development working experience with languages such as **C++, Visual Basic, Go and R**.
- Demonstrate a unique strength in adopting **multidisciplinary** domain expertise in cross-functional corporate teams as well as academic environments. Designed and developed **end-to-end intelligence system** from device-level solution to server-side dashboard display while maintaining CI/CD pipelines.
- As a technology advocate in the data science field, led the integration of **in-house generative AI and agentic systems** for several retail solutions.
- The **doctoral research** focused on predicting disease epidemic cycles using real-time satellite imagery through computer vision, machine learning, and statistical techniques.

## EXPERIENCE

**Walmart Global Tech**

**Bentonville, AR, USA**

*Staff Data Scientist*

*May, 2024 - Present*

- **Leading system architecture and end-to-end technical execution** of a barcode intelligence solution, integrating CV-based inference, distributed data pipelines, CI/CD, and MLOps to achieve **measurable OpEx savings across 600+ live stores**. The solution is now integrated with Walmart's associate app and used daily by 10,000+ associates.
- Led the data science technical execution of a **cost-optimization management system**, designing a vision-based conversational intelligence layer over retail camera feeds using production-grade Large Vision Models (LVMs).
- Coordinated and architected an in-house LLM-based framework to **reduce operational costs**, leveraging open-source multimodal models (e.g., qwen3-vl, llama3.2-vision) and advanced prompting paradigms including Chain-of-Thought (CoT), ReAct, and Visual Chain-of-Thought (VCoT).
- Advanced a production-grade multi-barcode reader solution to **improve shelf-stocking efficiency** for Walmart associates by leveraging state-of-the-art deep learning object detection models such as Grounding DINO and YOLOX.
- Built an auto-annotation system using vision and multimodal foundation models (e.g., Grounded SAM) to **reduce labeling costs and time**.
- Contributed to data science hiring by serving as a technical interviewer, evaluating candidates on machine learning, statistical reasoning, and system design.
- Supported and contributed to large-scale open-source object detection projects, including the YOLOX framework, through feature development and performance improvements.

*Senior Data Scientist*

*May 2022 - April 2024*

- Led the development of a cross-platform (Android/iOS) annotation tool, **reducing manual labeling cost and turnaround time** by integrating on-device and backend inference using YOLOv8 and EfficientNet, in collaboration with cross-functional teams.
- Proposed, architected, and led an automated end-to-end model training and inference pipeline, scaling dynamic product detection across multiple Walmart stores through **100+ custom camera deployments** to identify safety compliance issues. Leveraged Looper, Docker, Concord, and GCP for implementation.
- Developed a novel dimension estimation model using MiDaS-inspired monocular depth architectures, **enabling a robust shelf-availability service for retail stocking**. The model is a proprietary in-house deep learning solution and is currently in the patent filing process.

- Deployed a high-throughput image processing framework that serves multiple projects, handling 1,000+ requests per second and integrating object detection and classification pipelines with embedding-based vector search using Milvus and Pinecone.
- **Improved online shopping efficiency** by designing a recommendation system for dynamic product recognition in inventory management, leveraging Siamese neural network architectures and LightGBM model.
- Architected a modular, production-grade ML inference deployment framework that **does not require external managed ML inference engines** (e.g., Vertex AI, SageMaker), implemented in Go, supporting cross-platform execution and integrating PostgreSQL, SQL-based systems, and BigQuery.
- Enabled continuous data science research and delivery by **leading recurring technical knowledge-sharing sessions** and maintaining CI/CD infrastructure for end-to-end ML workflows.

**Bayer Crop Science**

**St. Louis, MO, USA**

**Senior Data Scientist**

*Oct 2021 - May 2022*

- Drove data-informed roadmaps that reduced canola **operational costs by over 50%** by proposing and delivering an automated UAV-based flowering detection system, leveraging semantic segmentation (DeepLabv3), image classification (ResNet backbones), classical CV algorithms (OpenCV), and scalable cloud infrastructure on AWS EC2.
- Developed a **patented** deep learning framework for satellite-to-UAV image super-resolution using GAN (Pix2Pix, StyleGAN) and Transformer-based architectures on multi-spectral geospatial imagery, implemented in PyTorch and supported by SQL-driven data workflows.
- **Delivered model-driven insights to multiple operational teams** by building geospatial visualization pipelines using GDAL and Shapely, surfaced through interactive Tableau dashboards.
- Evaluated emerging satellite data acquisition opportunities and delivered market and vendor analysis to inform procurement decisions, contributing to the R&D organization's 5+ year technology roadmap.
- Supported technical hiring efforts by participating in **candidate evaluation and interview processes** across multiple engineering and data science teams.

**Spatial Data Scientist**

*Mar 2020 - Oct 2021*

- Increased breeding pipeline efficiency for global cotton programs, **delivering nearly \$100K in cost savings** by designing and deploying a UAV-based multi-flight cotton maturity detection framework leveraging semantic segmentation models (U-Net, R-CNN) in combination with decision tree-based modeling.
- Developed core automation modules for the company's first scalable UAV processing pipeline, eliminating manual field note collection across multiple crops and **generating thousands-of-dollar operational savings**; recognized with the 2021 Breeding Excellence Award.
- Orchestrated key components of a global UAV-based soybean maturity automation pipeline, enabling thousands of acres of new maturity operations by developing software tools and models for color calibration, GCP detection, and maturity estimation, resulting in a **patent**.

**Research Data Scientist (Geo-Spatial)**

*Sep 2018 - Mar 2020*

- For **reducing soybean maturity breeding variability** in year-long operations, applied a broad range of machine learning techniques, including ensemble and deep tabular models (Random Forest, XGBoost, TabNet), alongside causal inference and Bayesian modeling approaches.
- **Reduced field monitoring operations by over 50%** by proposing and delivering a satellite-based crop health quality recommendation system, utilizing time-series forecasting and anomaly detection with GAN-based models and temporal modeling with LSTM architectures on high-resolution multi-band imagery.
- **Enabled operational placement optimization** for a global testing network by developing a satellite-based clustered embedding solution leveraging word2vec-style representation learning.

**Remote Sensing Data Scientist (Summer Intern)**

*Jun 2018 - Aug 2018*

- **Automated a cost-free, open-source image stitching solution** for hundreds of UAV flights by designing and deploying a scalable photogrammetry pipeline based on OpenDroneMap (ODM).

**International Maize and Wheat Improvement Center (CIMMYT)**

**Kingston, RI, USA**

**Research Consultant**

*Jun 2016 - May 2018*

- Designed and implemented a mobile irrigation optimization system, utilizing R-based analytical models and Android deployment within a summer research project in collaboration with CIMMYT.

**Graduate Research and Teaching Assistant**

Jan 2015 - May 2018

- Developed a predictive epidemiological model to forecast rotavirus outbreaks up to one month in advance in developing regions, as part of the "Control of Endemic Cholera in Bangladesh" project funded by the Bill & Melinda Gates Foundation, leveraging causal inference, real-time satellite data, and multi-source disease surveillance data (e.g., WHO, CDC) using LSTM, TabNet, and LASSO models.

**Institute of Water and Flood Management (IWMF)**

BUET, Dhaka, Bangladesh

**Research Associate**

Jul 2011 - Dec 2014

- Participated in large-scale international climate research projects focused on high-resolution regional climate modeling and flood forecasting, enabling climate impact assessments for multiple developing countries.

**EDUCATION****Ph.D. in Civil and Environmental Engineering**

Aug 2018

*University of Rhode Island, Kingston, RI, USA***Master of Science in Water Resources Development**

Dec 2014

*Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh***Bachelor of Science in Water Resources Engineering**

Feb 2011

*Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh***LEADERSHIP, AWARDS AND EXTRACURRICULAR ACTIVITIES**

- Led a cross-functional hackathon team to a **1st-place finish in the Generative AI category** and secured a **top-20 overall ranking at Techathon 2023**, hosted by Walmart Global Tech.
- Recognized as an **Exceptional Performer** within a **50+ member organization** at Walmart Global Tech in 2023 for sustained technical and business impact.
- Awarded the **Digital Innovation Award 2021** by Bayer Crop Science for architecting an **end-to-end automated UAV analytics platform**, spanning drone image acquisition, backend processing, deep learning inference, and visualization dashboards.
- Enabled innovation in remote sensing and geospatial analytics by organizing focused **data science meetups**, facilitating collaboration between engineers and scientists that **contributed to multiple patent developments**.
- Served as a corporate mentor for graduate students in remote sensing and agriculture, collaborating with **Purdue University, the University of Illinois at Urbana-Champaign, and the University of Missouri** to provide technical guidance and industry insight.
- Led and managed a student-run blood donation organization as **President of the Badhon-Suhrawardi Hall Unit**, coordinating operations for **1000+ donors** over a two-year term (2008–2009).
- Founded and operated a college-based apparel business ("**Bayna**") for **three years**, managing end-to-end operations including sourcing, production, sales, and finance.

**PUBLICATIONS**

Have published 9 journal papers and 20 conference papers and reviewed many more. Have more than 1000+ citations as of now. Presented talks and posters in various conferences. **More details can be found in google scholar.**