

# Joseph P. Robinson, Ph.D.

LEAD AI AND SOFTWARE ENGINEER · MACHINE LEARNING RESEARCHER · MULTI-MODAL SYSTEM SPECIALIST

+1 (978) 764 0003 | [jobs.vision@gmail.com](mailto:jobs.vision@gmail.com) | [jobsvision.com](http://jobsvision.com) | [github.com/visionjo](https://github.com/visionjo) | [linkedin.com/in/jrobby](https://linkedin.com/in/jrobby) | [M jvision](https://www.jvision.com) | [scholar](https://scholar.google.com)

"Be like water making its way through cracks. Do not be assertive, but adjust to the object, and you shall find a way around or through it. If nothing within you stays rigid, outward things will disclose themselves." –Bruce Lee

## Experience

**Amazon Robotics** | *Software Developer Engineer* | Westborough, MA

Feb. 2025 - CUR

- Implemented depth perception for robots deployed in the [next-generation fulfillment centers](#).
- More to come!

**BitHuman.io** | *Head of Rapid Product Integration and Development (RAPID)* | Remote

Aug. 2023 - Jan. 2025

- Architected and built the next-generation virtual assistant, replacing Unreal Engineer with a fully generative AI solution: AI modules included, but not limited to, talking face synthesizer, speech synthesizer, speech-to-text, and large-language models (LLMs).
- Reduced latency by >70% via strategic message passing, hybrid computing (on-prem and cloud), and robust caching (Redis).
- Invented dual-layered LLM for more natural conversations [P2] and improved accuracy [P3]: mean rate of hallucinations, 20% to 7%.
- Presented products at conferences and client meetings, nationally and globally, as a company leader and technology liaison.
- Designed goal-orientated paradigm [P1] required for release, securing two significant contracts worth >\$3M.
- Worked closely with stakeholders with product iteration, gathering feedback and other data points for critical decision-making.
- Built and managed company roadmap, driving RAPID team via agile and company OKRs, aligning short-term to long-term vision.
- Researched and incorporated the latest works, ensuring our critical AI components performed on par with SOA.
- Implemented computer vision algorithms to detect the user via webcam and use lip sync to cancel background noise [P4].
- Led a team of seven engineers (i.e., AI and software) via Agile (Scrumban) workflow: designed software, facilitated code reviews, and set high-quality standards for coding, documentation, and model deployment.

**Northeastern University** | *Machine Learning Software Engineer* | Remote

Jan. 2023 - Dec. 2023

- Optimized GPU utilization and SLURM configurations, resulting in a 35% performance improvement for large-scale AI/ML models.
- Built microservices and Q&A systems using dual LLMs for short-term and long-term memory, reducing operational overhead by automating [HPC cluster](#) processes and improving [user-facing documentation](#), reducing service ticket counts by >40%.
- Provided MLOps to the cluster to improve model training for data labeling, hyperparameter selection, and model deployment.
- Delivered various workshops for effective ML on the cluster, including PyTorch best practices, model/data version control, and others.

**Tufts University** | *Part-Time Lecturer- Machine Learning* | Medford, MA

Jan. 2022 - Dec. 2022

- Taught a project-based machine learning course to 120 students, using real-world datasets and hands-on programming exercises.
- Automated grading using [Otter-Grader](#), providing students instant feedback and reducing grading time by 60%.

**Vicarious Surgical** | *AI Engineer- Team Perception* | Waltham, MA

March 2021 - July 2022

- Architected and deployed scalable 3D depth estimation for robotic systems, achieving sub-millimeter accuracy in real-world testing.
- Designed fault-tolerant, high-performance deep learning models optimized for edge devices in C++ and CUDA for 40% reduced latency.
- Developed robotic algorithms for sensor-based motion control systems, enabling precise surgical instrument manipulation through end-to-end software integration implemented using C++ and ROS: testing prototyped concepts and features via simulation and robot.
- Collaborated with HW & SW teams to test and validate perception systems in a lab setting, using feedback for continuous refinement.
- Prototype and test concepts or features, both through simulation and emulators and with live robotic equipment
- Led data collection, labeling real and (Gazebo) simulated data to fine-tune deep learning models via multi-task & transfer learning.

**ISMConnect** | *Grad Student Intern* | Cambridge, MA

May 2019 - Aug. 2019

- Researched bias in facial recognition systems, proposing a novel method grounded in signal detection theory to evaluate fairness.
- Developed the [Bias Faces in the Wild \(BFW\)](#) dataset, improving bias measurement across under-represented demographic groups, resulting in a 15% increase in model fairness [4]. Learn more on [GitHub](#).
- Improved SOA CNN models for facial recognition by addressing demographic imbalances, balancing accuracy across subgroups [1].
- Invented feature learning scheme to de-bias face features and balance subgroup performances, increasing overall accuracy by >6%.

**Snap Inc. (Snapchat)** | *Grad Student Intern* | Santa Monica, CA

May 2018 - Aug. 2018

- Developed a landmark localization model using a novel Laplace KL-divergence objective, improving prediction certainty by >20% [5].
- Designed adversarial framework to train on unlabeled data, achieving real-time processing speeds (>20 FPS) with models 1/8th size.
- Conducted ablation studies showing model robustness, outperforming state-of-the-art in facial landmark detection by >7%.

**Systems & Technology Research (STR)** | *Grad Student Intern* | Woburn, MA

May 2016 - Sep. 2017

- Built C++ API for clustering algorithms (K-means, GMM, Agglomerative), achieving top rankings in NIST evaluations for [IARPA's JANUS](#).
- Accelerated nearest neighbor search via product quantization with inverted file indexing, gaining 200x speedup comparable accuracy.
- Designed Python APIs for [IARPA's Odin Program](#)– adversarial attacks by benchmarking black-box attacks on deep CNN models.

**MIT Lincoln Labs (MIT-LL)** | *Grad Student Intern* | Lincoln, MA

May 2014 - Aug. 2014

- Led a team in the [TRECVID's Multimedia Event Detection \(MED\)](#) challenge, securing 3rd place by developing a system that combined pre-trained CNNs and SVMs to detect complex events in unstructured videos [8].
- Integrated object and scene recognition models into a unified event detection pipeline, significantly reducing false positives while improving detection accuracy in a diverse video dataset [[presentation](#)].

**BBN Technology** | *DSP Engineering Co-Op* | Cambridge, MA

Jan. 2013 - Sep. 2013

- Improved the [Helicopter Alert and Threat Termination Acoustic System](#) via wavelets, increasing detection accuracy to 92% (+6%).
- Developed a Java-based tool to analyze and visualize audio data, leading to increased accuracy.

**Analogic Corporation** | *EE Image Reconstruction Co-Op* | Peabody, MA

Jan. 2012 - Sep. 2012

- Optimized [CT bag-scanning](#) algorithms, reducing hardware requirements by 33% using GPGPU and Intel Vector libraries.
- Developed interactive GUI in Java to train employees on visual threat detection, improving compliance with safety protocols.

**Joe Robinson Construction** | *Business Proprietor* | Salisbury, MA

Aug. 2005 - Dec. 2008

- Worked with customers on design, hired and managed employees, handled weekly payroll, maintained tools, and advertised services.
- Ran day-to-day operations on residential job sites as the licensed contractor, overseeing new construction, roofs, siding, and decks.

## Education

---

**Ph.D. in Computer Engineering**, ML-CV (Thesis), Northeastern University (NEU) | Boston, USA

Jan 2015 - Dec 2020

**BS in Electrical & Computer Engineering**, Northeastern University | Boston, USA

Jan 2011 - May 2014

**Awards:** Best Teacher (Data Science, 2019) | Best TA (Digital Image Processing & Circuits Lab, 2017) | **Huntington 100** | 1<sup>st</sup>-place, **Senior Capstone** | 1<sup>st</sup>-place, ECE Remote Control Contest | Best Reviewer, IEEE FG | Best Student Reviewer, IEEE FG (3x) | Best Senator (2x).  
**15+ International Conferences:** Attended CVPR, ACM-MM, ICCV, ECCV, AMFG, ICME, AAAI | Hosted workshops (e.g., **RFIW2021** & **AMFG2023**) and tutorials at numerous top-tier conferences (e.g., [6]) | Organized and Hosted **New England Computer Vision** Conference at NEU.  
**Extra Curriculum:** President, **IEEE@NEU** | Lead Research Ambassador, Student Research Engagement Committee | Student Senator, SGA | **Research Experiences for Undergraduates (REU)** participant (2010 & 2011) | ECE Department Ambassador | SAC board member.  
**Courses:** Assistive Robotics | Deep Learning | NLP | Optimization | ML | Computer Vision | Digital Image Proc. | Parallel Processing.

## Selected Publications

---

30+ peer review papers. See complete list on **Scholar**. Also, check out my Medium blog: <https://medium.com/@jvision>.

- [1] **JP Robinson** et al. "Balancing Biases and Preserving Privacy on Balanced Faces in the Wild". In: *TIP*. 2023.
- [2] **JP Robinson** et al. "Families In Wild Multimedia: A Multi-Modal Database for Recognizing Kinship". In: *TMM* (2021).
- [3] **JP Robinson** et al. "Survey on the Analysis and Modeling of Visual Kinship: A Decade in the Making". In: *TPAMI* (2021).
- [4] **JP Robinson** et al. "Face Recognition: Too Bias, or Not Too Bias?". In: *CVPR* (2020).
- [5] **JP Robinson** et al. "Laplace landmark localization". In: *ICCV* (2019).
- [6] **JP Robinson** et al. "To Recognize Families In the Wild: A Machine Vision Tutorial". In: *ACM MM*. 2018.
- [7] **JP Robinson** et al. "Visual kinship recognition of families in the wild". In: *TPAMI* (2018).
- [8] **JP Robinson** et al. "Deep models for detecting complex events in unconstrained videos". In: *SATBCA*. 2016.
- [9] **JP Robinson** et al. "Families in the wild (fiw): Large-scale kinship image database and benchmarks". In: *ACM MM*. 2016.

## Selected Projects

---

### Data Collection and Depth Estimation for Robotic Surgery

- Led data collection for depth estimation, capturing real and simulated data, i.e., human cadavers, 3D-printed shapes, Gazebo/ROS (i.e., from abundant source to target domain). Integrated ROS-based simulation for automated robotic movement and fine-tuning.
- Trained depth perception models with sub-millimeter precision using transfer learning (from driving datasets to surgical applications), reducing the required real-world ground truth data and ensuring domain adaptation (i.e., boot-strapping model training).

### Visual Kinship Recognition – Families In the Wild (FIW) Dataset [Project Homepage]

- Created the largest and most comprehensive visual kinship recognition dataset, featuring 1,000 families with 10+ images each [9].
- Led a global team for data collection, developed a semi-supervised labeling model, and built a Java-based GUI to accelerate dataset annotations [7], then fed a novel multi-modality algorithm with feedback to automate the collection of multimedia data [2].
- Designed novel models and algorithms that significantly advance kinship recognition research.
- Proposed a generative model to predict child appearances from parent images, exploring age and gender variations in latent space.
- Grew as an expert in visual kinship recognition, contributing more than 12 peer-reviewed papers, multiple challenges (e.g., **Kaggle competition** with 550+ teams), tutorials at top-tier conferences (e.g., CVPR, AMFG [6]), and a PAMI survey [3]. Released data on **FiftyOne**.

### Tunnel Inspecting Robot

- Collaborated with MASS-DOT to design and build an autonomous tunnel inspection robot to automate labor-intensive processes, reduce safety risks for inspectors and drivers, and accelerate inspection times from  $\approx 5$  days to  $< 1$  day per site.
- Engineered a cost-effective solution ( $< \$300$ ) using a Raspberry Pi for real-time video streaming and Arduino for precise motor control.
- Implemented computer vision algorithms to detect, track, and measure structural defects, enabling long-term analysis of deterioration rates and digital documentation organized spatially and temporally.
- Developed a robust JAVA-based GUI for remote operation, integrating features for user login, progress tracking, database access, and dual joystick control of the robot's base and adjustable camera arm via an Xbox controller.
- Optimized real-time video processing via TCP to relay data from device to host, where vision algorithms processed feed in parallel.
- Led daily stand-ups, weekly deep dives, and biweekly sync-ups w/ stakeholders, ensuring alignment w/ MASS-DOT requirements.
- Awarded 1<sup>st</sup> place in the Senior Capstone Competition for technical innovation, live system demonstration, & effective simulation.

## Selected Patents

---

[P1] **Joseph Robinson**, Yun Fu, Method of Providing Personalized Customer Interactions with Adaptive AI. *US Patent 18,732,610*.

March 2024

[P2] **Joseph Robinson** Yun Fu, Dual-Layered Artificial Intelligence System. *U.S. Patent 18,604,504*.

March 2024

[P3] **Joseph Robinson**, Y. Fu, Dual-Layered AI System w Large Language Models & Different Visual Agents. *US Patent 18,634,939*.

April 2024

[P4] **Joseph Robinson**, Yun Fu, Background Noise Filtering System. *U.S. Patent 18,634,991*.

April 2024

## Selected Certificates

---

2023	<b>Yoga Teaching Training (200 Hours)</b> , The Hot Yoga Studio	Seabrook, NH
2012	<b>Gordon Engineering Leadership Program</b> , CENSSIS – Northeastern University	Boston, MA
2010	<b>Computer Aided Drafting Certificate</b> , Northern Essex Community College	Haverhill, MA
2006	<b>Massachusetts Construction Supervisor's License (CSL)</b> , Merrimac College	Merrimac, MA

## Skills

---

<b>Programming</b>	C++/C, Python, CUDA, JAVA, Matlab, Bash, LaTeX, Catkin, CMake, Emacs, SQL, R.
<b>Frameworks</b>	TensorFlow, PyTorch, OpenCV, Keras, Spark, Scikit, ROS, Dlib, Open3D.
<b>Software &amp; Tools</b>	Linux, Docker, ROS, Gazebo, AWS, Kubernetes, Slurm, MPI, Git, Jenkins, JIRA, NVIDIA Jetson.
<b>Software Development</b>	Architecture design, fault-tolerant systems, CI/CD pipelines, source control (Git).
<b>Computer Vision</b>	3D depth estimation, object detection/classification, scene understanding, real-time processing, Segmentation.
<b>Soft Skills</b>	Strategic planning, team mentorship, cross-functional collaboration.

Last modified: Friday 14<sup>th</sup> February, 2025