

Brian Lu

<https://brianlu.me> | contact@brianlu.me | +1 (630) 717-9618 | [LinkedIn: greencappuccino](#) | [Github: GreenCappuccino](#)

Education

Purdue University | 3.21 GPA | August 2022 – May 2025

B.S. in Computer Engineering

- ECE 30100 — Signals and Systems — Fall 2023
- ECE 20875 — Python for Data Science — Fall 2023
- ECE 26400 — Advanced C Programming — Spring 2023
- ECE 36800 — Data Structures — Fall 2023
- ECE 27000 — Introduction to Digital Design — Fall 2022
- ECE 33700 — ASIC Design Lab — Fall 2023
- ECE 36900 — Discrete Math for Computer Engineering — Spring 2023
- MA 26600 — Differential Equations — Fall 2022

Skills: C, Python, Verilog, SystemVerilog, EDA, Synopsys, Test Benches, Validation/Verification, Circuit Analysis, Signal Analysis

Experience

Merck Sharp & Dohme | Current | August 2023 – Present

Undergraduate Researcher

- Designing a **Visual Document Understanding** (VDU) solution for use in converting scanned lab documents to JSON
- Performing **transfer learning** for the Naver Clova AI Donut Swin **Transformer/BART** model on synthetic documents
- Generating synthetic documents for training using Python, ReportLab, and **Dask** that model real-world characteristics
- Restored tools for the 1994 UniPen on-line handwriting dataset from SunOS 5 to Linux Flatpak
- Writing tools to generate **commercially-viable** synthetic handwriting for use in **synthetic document generation**
- Designing RDF **graph data storage system** based on oxigraph in **Rust** for storage of ingested documents

Skills: Visual Document Understanding, Computer Vision, Natural Language Processing, PyTorch, Python, ReportLab, Dask, LSTM, Transformer, Rust, RDF, Graph Databases, UNIX, Synthetic Data, Transfer Learning/Fine-tuning

The Data Mine @ Purdue University | May 2023 – August 2023

Software Engineering Contractor - Eli Lilly & Company

- Developed automated documentation generation reconciling with **Confluence**, replacing legacy system
- Optimized **Apache Airflow** image building operations on both Python Poetry and Docker build steps
- Replicated production deployment on k3d to simulate and resolve **Helm** deployment errors

Skills: Python, Airflow, Kafka, Neo4j, Jenkins, JFrog Artifactory, AWS, Elastic Kubernetes Service

Merck Sharp & Dohme | August 2022 – May 2023

Student Researcher

- Architected Safety Data Sheet (SDS) parsing and insights software with team
- Software cuts down on **hours** of work per day reading SDS documents, currently being **integrated**
- Wrote **low-level** PDF parsing engine to extract document hierarchies, fields, and images
- Created **OpenCV**-based **classifier** to accurately detect GHS pictograms

Skills: Python, pdfminer, OpenCV, Tabula, LaTeX, Jinja2, FastAPI

High Oak Robotics | May 2021 – January 2023

Robotics Coach

- Lead design of, wrote, and maintained Sequoia **asynchronous real-time** task scheduling library
- Taught robotics **control, navigation**, and programming concepts to team members
- Maintenance of sensor fusion pipeline combining dead wheels and VSLAM for positioning

Skills: Java, Kotlin, async, Motion Profiling, PID, VSLAM, Dead Reckoning, Sensor Fusion

Projects

Shamrock Cluster – <https://shamrock.systems> | September 2020 – Present

- 3-node **High Availability** homelab cluster targeting **cloud native** technologies and **Hyper-Converged Infrastructure**
- 3x hyper-converged Dell PowerEdge r710 servers each with 24 logical CPUs, 128 GB RAM, 12 TB Storage
- Currently undergoing a live-migration from Proxmox VE to **RHEL** with **OpenStack Kolla**, with virtualized **OKD/OpenShift**
- 2x virtualized Windows Server 2022 Domain Controllers with failover provide identity services with Azure AD Connect

Skills: Red Hat Enterprise Linux (RHEL), OpenStack, Kolla, Docker / Podman, Galera, AWX / Ansible, Terraform / OpenTF, Ceph, Windows Server 2022, Windows Domain Controllers, Active Directory, Azure Active Directory, OKD / OpenShift, Kubernetes, Istio, Argo CD, MetalLB, mTLS, Computer Networking, VLAN, VXLAN, Open vSwitch, Dovecot + Postfix