Paul Yoon

+1 (713) 320-1032 | pauljy@stanford.edu | linkedin.com/in/pauljinyoon | paulyoon.xyz

EDUCATION

STANFORD UNIVERSITY Palo Alto, CA

Bachelor of Science, Computer Science. Minor in Music. GPA: 3.8/4.0 *Expected June 2027*

Relevant Coursework: Linear Algebra, Differential, and Integral Calculus of Several Variables, Computer Organization and Systems, Algorithmic Analysis, Probability Theory for Computer Scientists, Machine Learning, Computer Vision

PROFESSIONAL EXPERIENCE

STANFORD CENTER FOR ARTIFICIAL INTELLIGENCE IN MEDICINE & IMAGING Palo Alto, CA

Researcher – Principal Investigator: Bao Do, MD *Sept 2025 – Present*

- Implemented a retrieval-augmented generation (RAG) pipeline linking radiology reports with case study and image retrieval across 5+ web APIs, automating ∼90% of manual image search
- Engineered feature-extraction logic deriving 8+ imaging features per report, achieving 100% coverage (on retrieved images) of top-3 differential diagnoses across 50+ cases

SUNDIAL Palo Alto, CA

Data Science Intern *Jul 2024 – Sep 2024*

- *Sundial is a Series A startup building an automated data science and product analytics tool founded by two early Meta executives, one of whom went to Sequoia Capital before founding Sundial*
- Developed classifier identifying fraudulent users using behavior-based thresholds, reducing false positives compared to existing process by 50% and boosting overall detection accuracy by 75%
- Created, trained, and tuned a time series seasonality model, outperforming existing model by 120% as measured by mean absolute percentage error (MAPE)

PROJECTS

Timestamping Video Game Eliminations with Computer Vision

Python, LaTeX *Apr 2025 – Jun 2025*

- Curated a custom detection dataset (165 events, ∼1,650 frames) with color-jitter augmentation strategies
- Fine-tuned YOLOv8-nano (2.5M params) on <200 images, achieving 0.61 F1 and 0.49s mean temporal error, doubling precision compared to a baseline template matcher
- Engineered a lightweight inference pipeline (OpenCV + ffmpeg) that processes a 14-minute VOD in 1.5 minutes on CPU, generating highlight clips with sub-frame accuracy

An Exploratory Analysis of Feature Representation in Music Source Separation

Python, LaTeX *Jan 2025 – Mar 2025*

- Implemented featurization approaches (STFT, Mel-spectrogram) within a Band-Split RNN to isolate vocal tracks
- Integrated HiFi-GAN for Mel-spectrogram inversion, improving audio reconstruction quality by reducing artifacts
- Optimized training via PyTorch AMP, hyperparameter tuning, and data augmentation on the MUSDB18 dataset

Stanford Christian Students App

React Native, TypeScript *Jun 2024 – Aug 2024*

- Overhauled main React dashboard to seamlessly display daily readings, user information, and upcoming events
- Integrated AWS Amplify backend system for user authentication, push notifications, and Bible content retrieval
- Refactored and debugged a 50+ MB codebase, improving code modularity and maintainability

Explicit/Implicit Heap Allocator

Unix, C *May 2024 – Jun 2024*

- Implemented the “malloc”, “realloc”, and “free” functions optimizing for request throughput and memory utilization
- Incorporated an explicit list of nodes to assign optimal locations for new memory requests and lower memory fragmentation
- Achieved 91% memory utilization via testing on heap activity memory requests from Emacs, Cmake, and Firefox

TECHNICAL SKILLS

Languages: Python, LaTeX, TypeScript, JavaScript, Lean, SQL, C++, C, HTML/CSS

Frameworks/Libraries: Pandas, NumPy, Matplotlib, scikit-learn, React, React Native, Next.js

Developer Tools: Git, Unix, Vim, VS Code, Apache Spark, Snowflake, Jupyter Notebook, Qt Creator