Sanjay Kumar

(352)-573-8960 | svk2121@columbia.edu | linkedin.com/in/sanjayvkumar31 | github.com/sanjito31

Education

Columbia University — **New York, NY**
M.S. in Computer Science - Machine Learning Track — *Expected Dec 2026*

University of California, Los Angeles (UCLA) — **Los Angeles, CA**
B.S. in Microbiology, Immunology, and Molecular Genetics — *Sep 2016 - Jun 2020*

Experience

Cedars-Sinai Regenerative Medicine Institute, Svendsen Lab — **Los Angeles, CA**
Research Associate II — *Feb 2022 - May 2024*

- Applied machine learning to data by developing three custom Python scripts and macros to automate high-throughput imaging analysis, processing over 10,000 microscopy images using computer vision techniques with CellProfiler, Random Forest, and R.
- Conducted large-scale biomedical research studying gut microbiome-CNS relationships in Parkinson's Disease patients using 50+ micro-fluidic chip models, generating complex datasets requiring advanced statistical analysis and pattern recognition.
- Analyzed clinical research data using ANOVA and statistical methods to summarize quantitative PCR, ELISA, and Western Blot results across 50+ experiments, demonstrating expertise in healthcare data interpretation and validation.
- Investigated therapeutic interventions by testing anti-sense oligonucleotides for SARS-CoV2 prevention in primary lung tissue and mouse models across 5 major experiments, contributing to understanding of precision medicine approaches.
- Managed complex laboratory workflows by mastering 20+ standard operating procedures for stem cell differentiation into neurons, gut cells, and immune cells while mentoring undergraduate researchers in biomedical techniques.

Arthritis & Pain Relief Medical Center (Dr. Soha Dolatabadi, MD) — **Los Angeles, CA**
Medical Scribe — *Sep 2020 - Jan 2022*

- Streamlined clinical workflows by documenting and transcribing medical histories, medications, physical examinations, and treatment plans in real-time, gaining deep understanding of patient care processes and documentation requirements that insurance companies require.
- Optimized healthcare data management by maintaining detailed electronic health records (EHR) for over 200 patients, ensuring seamless information flow between healthcare providers and insurance entities through organized documentation of lab results, imaging reports, and clinical notes.
- Developed automation solutions by creating custom scripts and macros to eliminate repetitive charting tasks, resulting in 67% increased patient throughput (15 to 25 patients daily) and demonstrating early application of technical skills to solve healthcare inefficiencies.
- Facilitated patient-provider communication by interfacing directly with patients in both Spanish and English to explain care instructions, diagnostic procedures, and address questions, providing insight into patient experience pain points in the healthcare journey.

Projects

Sanjay's Pics - Photography Portfolio and EXIF Parser — **www.sanjayspics.com**
Full Stack Web Application — *May 2025 - Aug 2025*

- Architected full-stack photography portfolio with Next.js 15 and TypeScript, implementing Server Components, App Router, and PostgreSQL database with Prisma ORM for type-safe development and optimized performance.
- Reverse-engineered proprietary Fujifilm EXIF parser to extract 50+ undocumented maker note metadata fields including film simulations, color profiles, and custom camera recipes, surpassing capabilities of standard EXIF libraries.
- Built comprehensive admin dashboard with role-based authentication using BetterAuth, featuring image management, batch operations, EXIF data visualization, and recipe organization with secure protected routes.
- Designed cloud-optimized image pipeline integrating Cloudinary CDN for automatic optimization, Vercel Blob for large file handling, and responsive gallery with lazy loading for enhanced user experience.

RGB Information Panel
Full Stack IoT Display System — *May 2025 - Aug 2025*

- Developed full-stack IoT information display system using C++ on ESP32 microcontroller with 64x32 RGB LED matrix, implementing HUB75 DMA display protocol and custom object-oriented architecture for real-time data visualization.
- Built RESTful backend API with FastAPI and Docker to aggregate data from multiple external APIs including OpenWeatherMap, NYC MTA, Spotify, and Formula 1 services, deployed on containerized home server infrastructure.
- Engineered modular display framework with custom C++ objects for different information types (weather, transit, music, sports), utilizing ArduinoJson for efficient data parsing and Adafruit GFX for graphics rendering.
- Implemented Wi-Fi enabled embedded system with real-time data fetching capabilities, designing scalable architecture for future WEBP animation support and over-the-air update functionality.

Technical Skills

- **Programming Languages**: C, C++, Java, Python, R, HTML, CSS, JavaScript/TypeScript.
- **Data Analysis**: Cell Profiler, ImageXpress, ilastik, FIJI, Anaconda, ggplot2, NumPy, Tidyverse, RandomForest. XGBoost.
- **Other**: Docker, React, Flask, FastAPI, Node.JS, Next.JS, SFTP, SMB, SSH, Git, PostgreSQL, PlatformIO, Arduino.