

Victor Chen

vic006@ucsd.edu | (415) 802-8728 | San Francisco, CA 94134 | <https://github.com/vichen1>

EDUCATION

University of California, San Diego (UCSD)

B.S. Data Science

Expected June 2027.

Relevant Coursework: Data Structures & Algorithms, Machine Learning Fundamentals, Data Cleaning, Data Visualization, Linear Algebra, Multivariable Calculus, Database Management, Object-Oriented Programming

TECHNICAL SKILLS

- **Languages:** Python, SQL, Java, MATLAB, HTML
- **Libraries:** Pandas, NumPy, scikit-learn, Matplotlib, Plotly, Beautiful Soup
- **Databases & Tools:** PostgreSQL, Git/GitHub, Jupyter Notebooks, VS Code, IntelliJ, JUnit Testing

PROJECTS

Analyzing Healthiness Between Recipes | [Project Link](#)

March 2025

Python, Pandas, NumPy, Jupyter Notebooks, Plotly, scikit-learn, GitHub

- Collected and pre-processed **200K+ Food.com recipes**, cleaning nutritional datasets for feature engineering and prediction.
- Built and optimized Decision Tree and Random Forest models; improved accuracy from **75% → 99%** with GridSearchCV hyperparameter tuning.
- Produced interactive data visualizations with Plotly to communicate insights across nutrition, cuisine, and user ratings.
- Deployed project as a web app, demonstrating end-to-end skills in **data pipeline design, visualization, and documentation**.

Custom RGB Image Processing Utility | *Python, NumPy, OOP, PIL*

March 2024

- Developed an object-oriented image processing framework with a custom **RGBImage** class for pixel-level manipulation.
- Engineered functions for reading, saving, brightness/contrast adjustment, and filtering using efficient NumPy operations.
- Designed robust input validation pipelines, preventing errors from malformed or out-of-range pixel values.
- Collaborated in a two-person team to define requirements and optimize runtime efficiency for large images.

Integrated Data and Image Processing Framework | *Java, Data Structures, OOP*

May 2024

- Built a priority-based data handling system with a dynamic heap structure, enabling **$O(\log n)$** scheduling of large-scale tasks.
 - Applied priority queues to optimize computation and resource allocation, improving performance on high-resolution image datasets.
 - Demonstrated scalable task scheduling by applying heap-based structures to handle large data and image processing workloads.
-