

Sahil Mondal

Los Angeles, CA | sahilmondal@gmail.com | (213) 663-4475 | linkedin.com/in/sahil-mondal | [Portfolio](#)

EDUCATION

University of Southern California, Los Angeles, USA

August 2022–May 2024

Master's of Science in Computer Science (GPA: 3.5/4.0)

Relevant Courses: Applied Natural Language Processing, Web Technologies, ML for Data Science, Analysis of Algorithms

Dr. Vishwanath Karad MIT World Peace University, Pune, India

July 2018–June 2022

Bachelor of Technology in Computer Science and Engineering (GPA: 9.74/10.0)

TECHNICAL SKILLS

Languages: Python, R, C, C++, C#, Java, Flutter, SQL, HTML/CSS, Javascript, PERL

Frameworks & Technologies: Tensorflow, Pytorch, OpenCV, React Js, Express Js, Node Js, Bootstrap, Flask, MySQL, MongoDB, Unity, Hugging Face, Keras, Android, Scikit-learn

Cloud Platforms: Google Cloud Platform (GCP), Amazon Web Services(AWS)

Developer Tools: Git, MATLAB, Postman, Tableau, Power BI, Docker, WordPress, Linux, Kubernetes

EXPERIENCE

USC Institute for Creative Technologies, LA, USA | Computer Vision Intern

May 2023–December 2023

- Integrated **Segment Anything** and **Lama** models, driven by textual prompts, to build an **image inpainting system**.
- Optimized 3D terrain reconstruction through application of diverse **Neural Radiance Field (NeRF)** manipulation techniques adopting **Learned Perceptual Image Patch Similarity (LPIPS)** loss and extending it to **NerfStudio framework** code, resulting in an **80% improvement** over conventional NeRF models.
- Crafted a **3D mesh** capturing terrain details to create a simulated environment tailored for **US Army training** needs.

Quidich Innovation Labs, Mumbai, India | Software Developer Intern

February 2022–July 2022

- Integrated **Yolov5**, **Fast R-CNN** with **Non-maximum suppression and Confluence** to model an accurate **real-time object tracking** system for cricket achieving **58 frames per second**.
- Attained a **98% accuracy** rate in real-time player **face recognition** by leveraging **Yolov5-face** model for face detection and **python face_recognition** library to extract facial encodings to match with encodings from test data.
- Collaborated with 2 interns to deploy the **dockerized** application for **20% more scalability** during prestigious cricket tournaments, including the World Cup 2021 and the Indian Premier League 2022. The live broadcast reached a viewership of **400 million** television viewers.

Canspirit Ai, Pune, India | Machine Learning Intern

March 2021–June 2021

- Spearheaded a team to create a **Flask web app** to extract crucial user info from PDFs using **text extraction** and **Named Entity Recognition models** obtaining an **94% overall accuracy** rate.
- Employed **Optical Character Recognition** tools like **Pytesseract**, **EasyOCR** with a **99% text extraction accuracy**, **UbiAi text annotation tool** for auto-labeling of custom entities and transformer language models (**BERT**, **ROBERTA**) to train on custom NER tags and respective tokens.
- Boosted a **25% increase** in company revenue following product launch, along with a case study.

PROJECTS

Counter COVID-19 Solutions | MobilenetV2, YoloV3, Python, OpenCV, Local Binary Patterns Histograms (LBPH)

- Designed a system incorporating three integral features for corporations to aid COVID-19 prevention norms.
- Implemented a Mobilenetv2 architecture for face mask detection, YoloV3 model with a metric of Euclidean Distance to monitor social distancing in real-time and LBPH face recognition algorithms enabling swift identification of violators who are then promptly issued warnings.

Yelp Business Fetch Website | React.js, Node.js, Java, XML, Bootstrap, HTML/CSS, Javascript, JSON, GCP

- Developed a fully responsive website utilizing React.js for the frontend and Node.js for the backend, fetching business details in the United States via the Yelp API and deployed it on Google Cloud Platform (GCP). Created an Android app with similar features using Java and XML.
- Pioneered 5+ user-friendly features such as input autocomplete, tabular display of businesses, business card with details, Google Map integration, past business reviews, and storing reservations with http localStorage.

AI-Driven Visual Narratives | Python, Text2Video/Text2Music models, Ollama, LLMs

- Engineered a dynamic visual-story enriched with integrated music through the fusion of DAMO Vilab's Text2Video diffusion model and Meta AI's Text2Music MusicGen model in response to user prompts.
- Fragmented user-prompt into sentence clusters based on semantic similarity using K-means and sentence embeddings.
- Transformed user prompts into music-context-aware cues leveraging Ollama and Mistral 7B LLM.

HONORS AND AWARDS

- Secured top 1 percentile in National Engineering Olympiad (April 2020) amongst 75000 students.
- Achieved top 2 percentile score out of 4500 applicants in course of 'Python for Data Science' offered by NPTEL.