

Kinshuk Sarabhai

Senior ML Research Scientist - Computer Vision

kinshuksarabhai@gmail.com | +91 99809 01948 | Bangalore, India

Career Summary

Seasoned backend engineer with 11+ years of experience designing, developing, and optimizing scalable solutions in computer vision, machine learning, and deep learning. Excels at transforming innovative prototypes into robust, production-ready systems.

Keywords:

Agentic AI systems, Gaussian Splatting, Vision-Language Model finetuning, Hybrid (text+visual) recommendation systems, Fine-grained image classifiers, 3D reconstruction

Technical Skills

- Languages: Proficient in Python; Prior experience in C++, C, Java, JavaScript, SQL, HTML/CSS, Matlab etc.
- CV/ML Libraries: Pytorch, HF Transformers, OpenCV, Scikit-learn, catboost etc.
- Engineering Tools: Experience with FastAPI, Celery, MySQL, MongoDB, AWS, GCP, Docker, Kubernetes etc.

Work Experience

Principal ML Scientist @ Phronetic AI, Bangalore, India : March 2024 - Present (~1.25 years)

Working on a low-code multimodal streaming AI platform for developers.

Key projects:

- AI Agent Builder: Design and implementation of a system for development of agentic AI workflows to process text, audio, and video streams.
 - Platform Link: <https://www.phronetic.ai/developer-platform>
- Real-time Talking Face Generation: Adapted Gaussian Splatting based model for real-time rendering to integrate it into agent builder platform.
- ABM (AI Business Manager) Pipelines: Designed and implemented AI-based activity monitoring pipelines for few domain-specific use-cases.
- Vision Planner: Designed video stream processing system
- Owlet Model: Adapted a lightweight vision-language model for video understanding.
 - Paper: <https://arxiv.org/abs/2407.15046>
 - Blog: <https://www.phronetic.ai/blogs/owlet-a-family-of-lightweight-models-for-video-understanding>

Technology Consultant (CV / ML): Nov 2023 - March 2024 (~ 4.5 months)

Focused on multimodal search in fashion e-commerce, video intelligence platform design, and custom ML research (latent diffusion etc.)

Team Lead - Vision Team @ Streamoid Technologies, Bangalore, India : Dec 2013 - Oct 2023 (~10 years)

Led the development of Fashion AI platform.

Notable projects include:

- **New Recommendation System:** Designed and implemented a hybrid (text+visual) search engine leveraging recent advancements in few-shot learning methods like CLIP
 - Tech used: Pytorch, Qdrant, OpenCV, Pillow, FastAPI
- **Allbirds Data Science Projects:** Developed predictive models for cart abandonment and active shopper behavior (likelihood of ordering) using catboost
 - Client projects done in individual capacity as a data science consultant
 - Tech used: Snowflake, Dagster, Catboost, Shap, MLflow, Docker
- **Catalogix:** Designed and implemented APIs for workflows using modern image editing tools like BG removal, smart auto-resizing, automatic shadow generation etc.
 - Involved in fine-tuning models for above use-cases.
 - Tech used: PyTorch, OpenCV, Pillow
- **Cataloging:** Designed and implemented backend for cataloging solution which helps in enrichment of ecommerce products feed using Autoscribe solution (see below).
 - Developed an online learning system with human-in-the-loop to improve models without retraining.
 - Tech used: Werkzeug, Celery, Redis, MongoDB
- **Autoscribe:** Created an auto-scalable model pipeline for extracting fashion attributes, improving inference times significantly.
 - Tech used: Werkzeug, Python AsyncIO, Google Datastore, Google Cloud Run
- **AI Studio:** Designed and implemented backend APIs to train, evaluate and deploy image classifiers on-demand for fine-grained attributes related to fashion
 - Developed a system to show insights to help domain experts improve the data used for attribute classification
 - Tech used: PyTorch, Scikit-learn, Pillow, Werkzeug, Celery, Redis, MongoDB
- **Visual Search & Similar Products:** Improved real-time feature extraction and developed systems for visual product recommendations.
 - Did feature engineering for extracting and matching multiple visual features
 - Tech used: OpenCV, Python, C++, Boost-python, LMDB
- **Animator / piQto:** Reduced image payload size by experimenting with image resolution and JPEG quality leading to faster image upload times by 80%
 - Tech used: OpenCV, Gearman, Django

Software Engineer @ Samsung Research Institute (SRI) - Delhi, Noida, India : Aug - Dec 2013 (~4 months)

Developed a prototype for removing backgrounds in real-time video calls using depth sensors.

Education

M.S (Research) - IIT Delhi, 2010 - 2013

Thesis title: Multi-view Reconstruction using Relaxation Labeling

Description: Proposed a novel approach for multi-view 3D reconstruction and compared it with state of the art algorithms.

B.Tech (CS) - College of Engineering Roorkee, 2006 - 2010

Bachelor project: Offline Handwritten Devnagari Character Recognition

Tech used: MATLAB, LibSVM

(Last updated: 29 June 2025)