Vinay Kumar Verma

AI Researcher | ML Engineer | Computer Vision
PhD-MT24002 | vinayv@iiitd.ac.in
South Delhi, 110020, India
linkedin.com/in/vermavinay982 github.com/vermavinay982

Education

Indraprastha Institute of Information Technology, Delhi — CGPA: 8.54 (Till 2nd semester)
PhD Scholar, Human Machine Interaction Lab (CSE)
2024 - Present

Indraprastha Institute of Information Technology, Delhi — CGPA: 8.54 (Till 2nd semester)
M.Tech Research (CSE)
Thesis: Localized Perception for Constrained Vision Systems (ICASSP)
2024 - 2026

Indraprastha Institute of Information Technology, Delhi — CGPA: 9.17
PG Diploma (Data Science and Artificial Intelligence with IBM)
2022 - 2023

ADGITM, GGSIPU, Delhi — CGPA: 7.79
B.Tech (ECE)
2016 - 2020

Skills

Programming:	Python, C++, Java, Go, Node.js, Flask, Git, Kubernetes, Docker
ML-DL Frameworks:	PyTorch, TensorFlow, OpenCV, ONNX, Triton Inference Server
Cloud/Edge:	Nvidia Jetson (Nano, Orin), AWS, GCP, Azure, Raspberry Pi
Technical Electives	Digital Image Proc.(8/10), Computer Vision(8/10), Deep Learning(8/10), Artificial Intel.(7/10), Object Oriented Programming(7/10), Research Method.(10/10)
Expertise Area	Computer Vision, Deep Learning, Machine Learning, Cloud & Edge Deployments
Programming	Python, C++, C, Java, Go, Node.js
Tools and Tech	PyTorch, OpenCV, TensorFlow, TensorFlow.js, ONNX, AWS, GCP, Azure, Git, Jetson, Raspberry Pi, Triton Server, Kubernetes, Docker

Work Experience

Research Scholar (Advisor Dr. Jainendra Shukla) (Jul, 25 - Present)
IIIT Delhi, India

- Designing **Temporal Interaction Grammars**, a falsifiable framework for formalizing and validating Human-Robot Interaction (HRI).
- Defining interaction primitives $(\sigma, \rho, \tau, \alpha)$ and operators (sequence, repair) with explicit **temporal budgets** to prevent breakdowns[cite: 1029, 1038].
- Applying this grammar to multimodal (gaze, posture) behavioral analysis for early autism diagnosis, in collaboration with **AIIMS New Delhi**[cite: 261, 941, 942].
- Collaborating with AIIMS New Delhi for clinically usable solutions.

Researcher (Advisor Prof. A.V. Subramanyam) (Oct, 23 - Jul, 25)
IIIT Delhi, India
Focus: Efficient Large Image Segmentation for 3D Medical, Satellite Imaging (Jetson)

- Developed segmentation algorithms for large satellite and volumetric medical datasets.
- Presented at ICASSP 2025, Hyderabad (Paper: Resource-Efficient Perception).
- Designed efficient 3D segmentation pipelines for memory constraints on edge platforms.

Computer Vision Engineer 2 (Nov, 22 - Nov, 23)
Stats Perform, London, UK [Remote]
Focus: Real-time football analytics & player tracking for live broadcasts

- Improved player tracking accuracy from 85% → 98% using temporal smoothing
- Designed robust Jersey OCR (TensorFlow), achieving 90%+ accuracy in live scenarios.

Machine Learning SDE - 1 (Apr, 22 - Nov, 22)
Fynd - Shopsense Retail (Reliance), Mumbai, India
Focus: High-traffic computer vision APIs for e-commerce

- Built CV APIs (e.g., OCR, smart crop, shadow removal) handling 2.5M+ req/day.
- Deployed models via Kubernetes, with Pub/Sub queue for bulk async processing.

Senior Computer Vision Engineer (Aug, 19 - Apr, 22)
Wobot Intelligence, Gurugram, India
Focus: Surveillance AI and video analytics across multi-camera systems (300+ CCTV Cams)

- Designed multi-camera vehicle tracking for retail & drive-through analytics.
- Built lighting-invariant color detection algorithm for CCTV, achieving 95%+ real-world accuracy.

Projects

CVPR 2025: Foundation Models for 3D Biomedical Image Segmentation (Apr, 25)

- Developed architecture to create universal 3D segmentation foundational model.
- Boosted ultrasound dice score from 0.30 → 0.70. Ranked 5th out of 210 global entries.

Vehicle Speed Estimation - Homography (Feb, 25)

- Designed lightweight system for estimating real-time vehicle speed on campus CCTV.
- Useful for low-resource pipeline suitable for smart fines and alert systems.

VigenPy - OpenSource Python Package (Jul, 23)

- Inspired by Lego blocks to stack videos along any axis (e.g., A—B—A, A+B).
- Useful for comparing algorithms or generating simulation videos.

Tex-Tractor - Live Text Searching in Videos (Nov, 24)

- Real-time text detection & OCR system from live camera feeds with 95% accuracy.
- Keywords: Motion detection, OCR, OpenCV.

Face Features Based Shopping Site (Jan, 24)

- Created recommendation system to suggest gifts based on recipient facial features.
- Keywords: TensorFlow, Keras, OpenCV, Flask, MTCNN, VGG-Face.

Publications

- **Resource-Efficient Perception for Vision Systems** Link
  - A. V. Subramanyam, N. Singal and V. K. Verma, "Efficient Localized Perception for Resource-Constrained Vision Systems," ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hyderabad, India, 2025, pp. 1-5, doi: 10.1109/ICASSP49660.2025.10888483
  - Optimized computer vision for embedded/edge systems under resource constraints.
- **Medical Face Identity** Link
  - V. K. Verma, V. Kansal and P. Bhatnagar, "Patient Identification using Facial Recognition," 2020 International Conference on Futuristic Technologies in Control Systems & Renewable Energy (ICFCR), Malappuram, India, 2020, pp. 1-7, doi: 10.1109/ICFCR50903.2020.9250002.
  - Used facial features to assign unique patient IDs, improving hospital workflows

Awards and Achievements

- Judge & Mentor, Smart India Hackathon 2024, ISRO Nodal Centre, Gujarat
- Judge, Toycathon India 2022, Delhi
- JK Pal Best Student Award – IEEE Delhi Section (2021)
- Winner – Smart India Hackathon 2018, CSIR Pune (Dengue Prediction App)
- Vice Chair – IEEE NIEC (Organized 10+ workshops and SIGs)

Interests and Hobbies

- Astronomy, Astrophotography, AI in Space Imaging, Long Exposure Photography

Declaration: The above information is correct to the best of my knowledge.
Vinay Kumar Verma
Date: November 2, 2025