Dr. Deepak Kumar

Ph.D.
Machine Vision Lab
Department of Computer Science and Engineering
IIT Roorkee, Roorkee, Uttarakhand, India, 247667

Email: d_kumar@cs.iitr.ac.in
Mobile: +91-941-1611-295
Github: https://deepakkumar-iitr.github.io/

CURRENT POSITION

I recently completed my Ph.D. at the Machine Vision Lab, *Indian Institute of Technology Roorkee, India,* under the supervision of **Prof. Balasubramanian Raman**. My research focused on Affective Computing, Multimodal human behavior analysis leveraging audio, text, visual, and physiological signal modalities. Additionally, I gained industry experience at Samsung Research Institute Bangalore (SRI-B), where I worked on Vision-Language Models (VLMs).

EDUCATION

Ph.D., Computer Science and Engineering, CGPA: 8.59/10   Dec 2021 - July 2025
Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, India
Thesis Title: From Signals to Visuals: Multimodal Affect and Behavior Analysis
Supervisor: Prof. Balasubramanian Raman

M.Tech, Software Engineering, CGPA: 8.25/10   Jul 2014 - June 2016
Motilal Nehru National Institute of Technology Allahabad, Prayagraj, Uttar Pradesh, India
Thesis Topic: Enhanced Hierarchial Particle Swarm Optimization using Time-Varying Acceleration Coefficient
Supervisor: Prof. Rajesh Tripathi

B.Tech, Computer Science and Engineering, 67.80 %   Jul 2008 - June 2012
Uttar Pradesh Technical University Lucknow, Uttar Pradesh, India

WORK EXPERIENCE

Research Intern   (Feb 17, 2025 - Aug 14, 2025)
***Samsung Research Institute (SRI-B)**, Bangalore, India.*
Working on Vision-Language Models (VLMs) to make them compatible for on-device deployment, including preprocessing of text documents, OCR outputs, scene text, handwritten text images, VQA data, and image captioning datasets. The goal was to build a lightweight model capable of handling diverse text extraction tasks across use cases while running efficiently on edge devices.

Work-studentship   (Aug 2023 - Jan 2024)
***Deloitte**, US-India Offices.*
Involved in cleaning, visualizing and analyzing URL data. I worked on building an automated phishing URL detection classifier using machine and deep learning. Performed error analysis, model selection and made inferences.

Teaching Assistant   (July 2022 - Dec 2024)
Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, India.
Assisted Prof. Balasubramanian Raman and Prof. Shahbaz Khan in lab and tutorial sessions for Computer Graphics (CSN-372), Machine Learning (CSN-382), Dynamic Graph Algorithm (CSN-531), Data Structures (CSC-102), and Advanced Algorithm (CSC-501) courses.

Assistant Professor under TEQIP-III (NPIU-MHRD)   (Oct 2018 - Dec 2021)
College of Technology, GB Pant University of Agriculture and Technology Pantnagar, Uttarakhand, India.

Involved in the research activities at department and college level, and Taught Python Programming, C-Programming, Data Structures, and Digital Electronics courses to the B.Tech, and B.Sc Agriculture students.

Teaching Assistant (July 2014 - June 2016)
Motilal Nehru National Institute of Technology Allahabad, Prayagraj, Uttar Pradesh, India.
Assisted Prof. Rajesh Tripathi in Lab and Quiz sessions for Computer Graphics and Digital Computer Organization (CA-3103) courses.

RESEARCH INTERESTS

- Deep Learning, Computer Vision

- Large Language models

- Vision Language models

- Multi-Modal Affective Computing

- Human-Computer Interaction (HCI)

- Multi-Modal Information Analysis

PUBLICATIONS

Journals

[J-2] **D. Kumar**, P. Singh and B. Raman, 'Affective State Detection via Hierarchical Cross-Modal Attention Networks on Multi-Modal Physiological Signals.', *Computers in Biology and Medicine* (Elsevier), **(SCI, Q1, IF: 7.0). [Under Review]**

[J-1] **D. Kumar**, P. Singh and B. Raman, 'All signals point to personality: A dual-pipeline LSTM-attention and symbolic dynamics framework for predicting personality traits from Bio-Electrical signals.', *Biomedical Signal Processing and Control* (Elsevier), **vol. 96**, pp. 106609, 2024. **(SCI, Q1, IF: 4.9)**

Conferences

[C-8] **D. Kumar**, A. P. Singh, P. Kumar, Xi. Li, and B. Raman, 'GAViD: A Large-Scale Multimodal Dataset for Context-Aware Group Affect Recognition from Videos.', Proceedings of *the 33rd ACM International Conference on Multimedia (MM'25)*, October 27 - 31, 2025, Dublin, Ireland. **[CORE A* | ERA A]. [Under Review]**

[C-7] **D. Kumar**, S. Madan, P. Singh, A. Dhall and B. Raman, 'Towards Engagement Prediction: A Cross-Modality Dual-Pipeline Approach using Visual and Audio Features.', Proceedings of *the 32nd ACM International Conference on Multimedia (MM'24)*, pp. 11383-11389, October 28 - November 1, 2024, Melbourne, Australia. **[CORE A* | ERA A].**

[C-6] **D. Kumar**, P. Dhamdhere and B. Raman, 'Fusing Multimodal Streams for Improved Group Emotion Recognition in Videos.', Proceedings of *the 27th International Conference on Pattern Recognition (ICPR 2024)*, pp. 403-418, December 01-05, 2024, Kolkata, INDIA. **[CORE B | Qualis A1].**

[C-5] **D. Kumar**, P. Singh, Richa, KB Nampalle and B. Raman, 'Integrating Physiological Signals with Dynamical Attention Networks for Personality Trait Analysis.', Proceedings of *the 33rd International Joint Conference on Neural Networks (IJCNN 2024)*, pp. 1-8, June 30- July 5, 2024, Pacifico Yokohama, Yokohama, Japan. **[CORE B | Qualis A2].**

[C-4] **D. Kumar**, P. Singh, A. Kumar, S. Ghosh and B. Raman, 'Neuro-Emotional Mapping of Human Emotions via EEG Signals.', Proceedings of *the 18th International Conference on Automatic Face and Gesture Recognition (FG 2024)*, pp. 1-5, May 27-31, 2024, SDKM, ITU Campus, Istanbul, Turkey. **[CORE B | Qualis A1].**

[C-3] **D. Kumar** and B. Raman, 'Speech-based Automatic prediction of Interview Traits.', Proceedings of *the 7th IAPR International Conference on Computer Vision and Image Processing (CVIP 2022)*, pp. 586-596, November 4-6, 2022, VNIT Nagpur, INDIA. [**IAPR Endorsed**].

[C-2] **D. Kumar**, S. Madan, and A. Singh, 'Modified HPSO using TVAC and Analysis using CEC Benchmark Functions.', Proceedings of *8th International Conference System Modeling and Advancement in Research Trends (SMART 2019)*, pp. 7-9, November 20-22, 2019, CCSIT, TMU, Moradabad, INDIA.

[C-1] S. Madan, **D. Kumar**, and A. Agnihotri, 'Privacy-preserving data aggregation in wireless sensor.', Proceedings of *7th International Conference on System Modeling and Advancement in Research Trends (SMART 2018)*, pp. 165-167, November 21-23, 2018, CCSIT, TMU, Moradabad, INDIA.

PROJECTS

Ph.D. Course Work Project

Emotion recognition using Speech with Deep Learning Techniques

Invloved in cleaning and analyzing the speech data. In this project, we predicted the emotions(Happy, sad, fear, anger, disgust) using speech with deep learning techniques (1D CNN with batch normalization, RNN, LSTM). In this project, we have used four available data sets (SAVEE, TESS, CREMA-D. Ravdess).

B.Tech Project

Developed College Website Using PHP and MySQL

I developed a college website using HTML, CSS, JavaScript, PHP, and MySQL for this project. The main objective was to develop a website that can be used in any college or school.

AWARDS & SCHOLARSHIPS

1. Awarded **Travel Grant** to attend the 5th Indian Symposium on Machine Learning (IndoML) hosted by the BITS Pilani Goa Campus during December 21-23, 2024.

2. Achieved place in **top-3** papers in the Multimediate Challenge at the ACM MM (Core A*) Conference 2024.

3. Spearheaded the "Syntax" team in the "Brain Responses to Emotional Avatars Challenge" in 18th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2024), securing the **second position**.

4. Played a role as a **volunteer** at the 14th The Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP) 2023 from 15 - 17 December at **IIT Ropar**, India.

5. **First rank** in the Data Science and Artificial Intelligence Quiz conducted by **CommBank India** during the 12th International Conference on Soft Computing for Problem Solving (SocProS 2023) held at Indian Institute of Technology Roorkee, Roorkee, India. during 11-13 August 2023.

6. Registration and travel support from **Microsoft Research Grant** to present a research paper at the 7th International Conference on Computer Vision and Image Processing (CVIP 2022).

7. Received **MHRD Scholarships** during M.E. (2014-16) and Ph.D. (2022-26).

8. Qualified **GATE** Computer Science (Rank 2571, percentile 98.85) in 2013, (Rank 1523, percentile 99.02) in 2014, and (Rank 1459, percentile 98.74) in 2015.

9. Qualified **UGC NET** Computer Science in Dec 2015, Nov 2017, and July 2018.

Attended workshop/Training/Conferences

1. Attended 18th International Conference on 'Automatic Face and Gesture Recognition (FG 2024)', Istanbul, Turkey.   May 2024

2. Attended the 7th International Conference on 'Computer Vision and Image Processing (CVIP 2022)', VNIT Nagpur, India.   Nov 2022

3. Attended 6th Summer School on AI, IIIT Hyderabad.   July 2022

4. One week of training on Machine Learning and Data Analytics Using Python Conducted by **Indian Institute of Technology Roorkee**. (Online Mode)   April 2020

5. Attended 10 days short Data Science and Analytics course at **Indian Institute of Technology, Gandhinagar** funded by **World Bank under TEQIP-III MHRD Govt. of India**. Feb 2020

6. Attended the one-week short-term course on Advanced Pedagogies: Active Learning and Digital Tools at **IIT Delhi** funded by **World Bank under TEQIP-III MHRD Govt. of India**. June 2019

7. Attended a one-week training program on "Machine Learning and Its Applications" organized by **Electronics and ICT Academy, IIT Roorkee**.   Oct 2018

Talks/ Workshops delivered

1. Served as Resource Person in the workshop on "Advanced Programming Concepts Using C and Python" on September 21st and 23rd, 2019, organized by FOECS, TMU Moradabad, Uttar Pradesh, India.

2. Conducted a 45-day short-term course on "Android: Apps Development", a skill-oriented value-added program from 26th June to 10th August 2017 at TMU Moradabad, Uttar Pradesh, India.

3. Conducted a 45-day short-term course on "Android: Mobile Apps Development" skill-oriented program from 18th June to 4th August 2018 at TMU Moradabad, Uttar Pradesh, India.

Technical Skills

Programming Languages :	C/C++ (proficient), Python (proficient), Core Java (proficient)
Development Tools :	TensorFlow, PyTorch, Keras, Anaconda, Spyder, Pandas, Visual Studio Code, Code::Blocks, Eclipse.
Database :	Microsoft SQL DB Server
Web Development :	HTML, CSS, PHP, JavaScript
Others :	LaTeX, XAMPP, Android Development (prior experience).

Professional Activities

Professional Memberships: Student member of IEEE, Professional member of ACM (Association for Computing Machinery) and Associate member of IEI (The Institution of Engineers (India)).
Research Review: Reviewed research papers in various Journals and conferences, including IEEE Signal Processing Letters journal, IJCNN 2024, CVIP 2024, CINE 2024, ICPR 2024, ACM Multimedia 2025, and ACM Multimedia 2025 1MDeepfake Challenge.
Program Committee member in the A2I: Affective Artificial Intelligence (ICPR 2024) workshop as part of the 27th International Conference on Pattern Recognition (ICPR 2024), December 01-05, 2024, Kolkata, India.

Personal Details

Father's Name :	Sh. Ajab Singh	Mother's Name :	Smt. Nirmesh Devi
Date of Birth :	Aug 09, 1991	Marital Status :	Married
Nationality :	Indian	Passport No. :	P0332791

Permanent Address: Vill. Bhawalpur Wasli, P.O. Sarakara Kamal, Dist. Sambhal, Uttar Pradesh, India

REFERENCES

Prof. Balasubramanian Raman

Professor, CSE Department, Indian Institute of Technology Roorkee, Roorkee, India.

E-mail: bala@cs.iitr.ac.in

Prof. Abhinav Dhall

Associate Professor, Monash University, Australia.

E-mail: abhinav.dhall@monash.edu

Prof. Rajesh Tripathi

Associate Professor, CSE Department, MNNIT Allahabad, Prayagraj, India.

E-mail: rajeshtcsed@mnnit.ac.in