SOUMYADEEP BANIK

Mumbai, Maharashtra, India

✉ soumya.d.banik@gmail.com 🌐 soumyadbanik ⧉ soumyadbanik in soumya-d-banik 📱 +91 8478860607

About

Computer Vision researcher with 4+ years of experience across academic research and industry, focusing on human pose reconstruction, multi-view geometry, person re-identification, and spatio-temporal video understanding. My work spans healthcare monitoring and sports broadcasting, with current research interests in inclusive 3D/4D human modelling, motion understanding, and geometry-aware learning.

Research Interests

- 3D & 4D Human Mesh Recovery (monocular & multi-view)
- Spatio-temporal human motion understanding
- Geometry-aware and graph-based representation learning
- Model Optimization, Real-time and edge-efficient vision systems

Research Experience

Indian Statistical Institute (ISI) **Kolkata**
Topic - Distributed Cognitive Systems for Healthcare

Project Linked Junior Research Fellow under **Prof. Ashish Ghosh** *Oct 2021 - Dec 2022*

- Conducted research on abnormal human activity recognition using RGB video and skeletal data
- Studied spatio-temporal modelling of human motion for healthcare monitoring
- Focused on **Graph Representation Learning** for human skeleton-based action recognition
- Investigated movement patterns related to medical conditions such as falls, fatigue, and prolonged inactivity

Teaching Assistant *Aug 2022 – Sep 2022*
- Machine Learning-I

Teaching Assistant *Feb 2022 – Jul 2022*
- Advanced Image Processing - M.Tech. Computer Science

Industry Experience

Computer Vision Engineer - *Quidich Innovation Labs* **Mumbai**
Building state-of-the-art AI solutions for sports broadcasting *March 2024–Present*

- **Decision Review Systems (DRS)**
  - Developed and deployed real-time automated no-ball and run-out detection systems from high-speed cameras (100–120 fps)
- **StiQy**
  - Designed ground perspective tracking algorithms for real-time (50 fps) virtual graphics overlays on cricket fields, significantly enhancing viewer engagement and experience.
  - **AI Keyer**- Developed an end to end player segmentation pipeline for accurate keying of players in virtual ads on the ground. Optimized the pipeline with CUDA acceleration to achieve a high fps (90+ fps) inference on NVIDIA RTX 3080 Ti while maintaining high mask accuracy
  - Worked on optimization of SOTA VLM and segmentation models (e.g. SAM-based architectures) for domain-specific use cases.
- **HyperView**
  - Developing advanced real-time pose estimation algorithms.
  - Reconstructed 3D player poses using 3D triangulation (90% accuracy) with a calibrated multi-camera setup on the cricket pitch, integrating BVH (BioVision Hierarchy) motion data into 3D graphics and motion builder software (Unreal Engine, Blender, Maya) to enhance realism.

Junior Data Scientist - *HappyMonk.ai* **Bangalore**
Building edge-centric vision models for video surveillance *December 2022-January 2024*

- **Human Activity Recognition**
  - Led the entire project lifecycle, overseeing data collection, preparation, preprocessing, model development to deployment on edge devices
  - Leveraged **SLOWFAST** network to develop a **spatio-temporal** deep learning model to accurately detect **human activities** within a scene
  - Automated the entire pipeline to streamline the production processes
- **Space Utilization**
  - Employed image processing techniques to implement an algorithm to show the empty spaces within an extensive inventory area
- **Object detection**
  - Collaborated in multiple projects on developing the dataset, object detection models(e.g. YOLOv5, YOLOv8) across different domains, addressing business problems such as **face detection**, **object counting**, **emotion recognition**, and **fire-smoke detection** etc.
  - Collaborated with the team to efficiently deploy these models on edge devices

Skillset

Topics of Interest	Pose Estimation(2D, 3D), Depth Estimation, 3D point cloud, Object detection, 3D Computer Vision, Human Activity Recognition, Object tracking, Semantic segmentation, OCR(Optical Character Recognition),
Tools & Techniques	Machine Learning(Regression, Classification, Decision Trees, Random Forests, Time series Forecasting and Clustering), Deep learning(DNN, RNN, LSTM, Encoder-Decoder Models, ViTs, Diffusion-based models),Graph Representation learning, GNN(Graph Neural Networks), Data Structure, OOPS
Frameworks	PyTorch, Ultralytics, Flask, FastAPI, AWS Lambda, NVIDIA DeepStream, TAO, CUDA, cuDNN
Data Analysis	Pandas, NumPy, Matplotlib, Seaborn
Model Optimization	Onnx, TensorRT
Programming Languages	Python, C, HTML/CSS, SQL
Technical	Git, Linux, LATEX, Docker, ngrok, MLOps, CVAT (annotation tool), OpenCV, Jupyter lab, Google Colab

Selected Projects

SpecsKart - Building a virtual specs try-on platform and recommendation system [*In collaboration with AI-LENS*] *January 2024-Present*
- Projecting the chosen specs onto the user's face using facial keypoint features, accommodating various facial orientations.
- Implementing a recommendation system based on past specs choices.
- **Technologies:** Python3, Matplotlib, MediaPipe, OpenCV, NumPy, Streamlit app, Scipy, FastApi, AWS Lambda

Surveying Social Distancing in Aerial Videos using Computer Vision 📄 � *May 2021-July 2021*
- Built a web app for monitoring the number of social distancing violations in a video by Multiple object tracking.
- Worked with a large scale video dataset i.e. Stanford Drone Dataset (link)
- Used Deep SORT for performing the Multiple object tracking and YOLOv4 for the detection task
- **Technologies:** Python3, Matplotlib, Seaborn, Pandas, Darknet, TensorFlow, HTML, Flask API

Quadtree based Image Saliency mapping ⌂ *October 2020-December 2020*
- Proposed a Quadtree Based approach for getting Salient features from an image which can significantly reduce the computational complexity of a CNN model from $O(M^2N^2C)$ to $O(M^2NC)$.
- **Technologies:** PyTorch, Python3

Video Steganography using LSB Technique by a Unique Frame Selection Method ⌂ *March 2019-May 2019*
- Proposed a unique frame selection method that takes the first number from the pseudo-random number sequences generated by our modular function (a Linear congruential generator) by taking a seed value as the secret key defined by the sender and increased the security.
- **Technologies:** OpenCV, NumPy,Pandas,Python3, Anaconda

Education

Ramakrishna Mission Vivekananda Educational and Research Institute **Belur, Howrah**
M.Sc in Computer Science *2019-2021*
Thesis Project - Surveying Social Distancing in Aerial Drone Videos 📄 (Grade- 9.43/10)
CGPA- 7.16/10

Ramakrishna Mission Residential College(Autonomous) **Narendrapur, Kolkata**
B.Sc(Hons) in Computer Science *2016-2019*
Thesis Project - Video Steganography using LSB Technique by a Unique Frame Selection Method 📄 (Grade- 9/10)
CGPA- 7.49/10 (First Class)

Honours, Scholarships & Activities

- Junior Research Fellowship (MeitY, Government of India), supporting research at the Indian Statistical Institute
- Swami Vivekananda Merit-cum-Means (SVMCM) Scholarship, Government of West Bengal, for postgraduate studies
- Departmental Service: Co-organised ENVISION'19, a technical event hosted by the Department of Computer Science, RKMRC Narendrapur
- Languages: Bengali (native), Hindi, English
- Interests: Photography, illustration, and visual design