Hiran Sarkar

67, Badridas Temple Street
Kolkata, 700004
(+91) 7980543067
hiransarkar2001@gmail.com

Website
aroundstar
Google Scholar

Area(s) of interest: Rigid and deformable 3D/4D reconstruction & generation, implicit deep learning, human pose estimation, generative modeling, video understanding, few-shot/incremental/open-set learning & object detection

EDUCATION

Netaji Subhash Engineering College, Kolkata - *Bachelor's of Technology* JULY 2019 - JULY 2023

Major in Computer Science and Engineering

CGPA: 9.14 / 10

EXPERIENCE

Technical University of Munich, CAMP - *Research Intern* APRIL 2023 - Present

Working on **dynamic scene representation from a monocular video** using NeRF and Neural ODE advised by *Dr. Benjamin Busam*.

- We learn the underlying physics of scenes with deterministic dynamics, using Neural Radiance Field that learns the spatial information and neural ODE that learns the temporal dynamics.

  This setup enables:

  - *Reconstruction of the scenes:* Accurately recreating the dynamics of the scenes
  - *Extrapolation of the scenes:* Predicting the evolution of the dynamics beyond the training data.
  - *Generalization to different physical parameters:* For eg., varying damping and friction. Trained on sparse values of these parameters, the model can generate novel trajectories for unseen values.
  - *Generalization to different initial conditions:* For eg., varying the initial position and velocity of dynamic objects. Trained on sparse values of initial conditions, the model can generate novel trajectories for unseen configurations.
- *Results:* **[Link]**. *The work is in progress and will be submitted to ICCV 2025.*

Sony Research India - *Research Intern* APRIL 2023 - Present

- Worked on **open-set object detection** where we addressed the high unknown misclassification problem. We introduced a clip based clustering based module that forms clusters in the semantic space forming a clear boundary separation especially between classes of close semantic proximity. The paper is accepted at **WACV 2024** as an oral presentation. **[Link]**
- Worked on a survey paper on **few-shot object detection (FSOD)** that covers different variants of the field including standard FSOD, generalized FSOD, incremental FSOD, open-set FSOD, and domain adaptive FSOD. The work is currently under submission in **ACM Computing Surveys**. **[Link]**
- Worked on **generative modeling**: text conditioned image to video generation, LCM-LoRA for faster diffusion model inference and real to anime video generation.

Independent Research JUNE 2022 - OCT 2022

Non-rigid 3D reconstruction, human pose estimation *advised by Dr. Rishabh Dabral*

- Modified the Part Affinity Fields code with triplet based bipartite matching.
- Intensive literature review and study of the problem of non-rigid SFM and dynamic scene reconstruction.

Indian Institute of Science, Bangalore, Vision & AI Lab - *Research Intern* JUNE 2021 - MAY 2022

I have worked on **domain adaptation** and **domain generalization** using self-supervision across multiple settings under *Dr. Venkatesh Babu*.

- Worked on unsupervised source-free domain adaptation technique using a subsidiary pretext task which acts as an aid to help attain domain invariance thereby minimizing domain discrepancies. The paper is accepted at **ECCV 2022**. **[Link]**
- Developed an unsupervised non-source-free domain adaptation technique using a Bag-of-VisualWords (BoW) like representation. We looked into the tradeoff between negative transfer risk and domain invariance exhibited at different layers of the network. The paper is accepted at **NeurIPS 2022**. **[Link]**

Teamcognito, Kolkata - *Research Intern* OCT 2020 - JAN 2021

- Worked on a speech recognition model using Recurrent Neural Network Transducer. Trained on the Common-Voice dataset.
- Worked on "Malware Detection on IoT devices based on traffic Meta-Data"

RESEARCH PAPERS

In Progress

Learning Physics from Visual Observations: NeRF meets Neural ODE for Dynamic Reconstruction [Results]
Target Conference: ICCV 2025
Hiran Sarkar, Benjamin Busam

Published

Beyond Few-shot Object Detection: A Detailed Survey [Link]
In Submission (ACM Computing Surveys)
Hiran Sarkar*, Vishal Chudasama*, Pankaj Wasnik, Vineeth Balasubramanian, Jayateja Kalla

Open-Set Object Detection By Aligning Known Class Representations [Link]
Winter Conference on Applications of Computer Vision (WACV) 2024. ***(Oral Presentation)***
Hiran Sarkar, Vishal Chudasama, Naoyuki Onoe, Pankaj Wasnik, Vineeth Balasubramanian

Subsidiary Prototype Alignment for Universal Domain Adaptation [Link]
Neural Information Processing Systems (NeurIPS) 2022
Jogendra N. Kundu*, Suvaansh Bhambri*, Akshay R. Kulkarni*, **Hiran Sarkar**, Varun Jampani, Venkatesh Babu R.

Concurrent Subsidiary Supervision for Unsupervised Source-Free Domain Adaptation [Link]
European Conference on Computer Vision (ECCV) 2022
Jogendra N. Kundu*, Suvaansh Bhambri*, Akshay R. Kulkarni*, **Hiran Sarkar**, Varun Jampani, Venkatesh Babu R.

PROJECTS

Image Style Transfer - [Link]

Implemented the paper "*Rethinking and Improving the Robustness of Image Style Transfer*" in TensorFlow. This paper examines why pre-trained ResNet networks hinder Image Style Transfer compared to VGG, despite ResNet's superiority in tasks like classification. It employs the L-BFGS optimizer to accelerate training.

Cartoon GAN - [Link]

This project aims to transform real-world photos into cartoon-style images using a generative adversarial network (GAN) comprising a generator and a discriminator. It employs two types of losses: adversarial loss (discriminator loss), which guides the generator to achieve the desired manifold transformation, and content loss (perceptual loss), which ensures the preservation of image content during stylization.

Emojify - [Link]

Returns the emoji that matches the input text, using LSTM network.

Neural Machine Translation - [Link]

Performed neural machine translation from English to Spanish using Bahdanau Attention mechanism and deployed it using Flask.

REVIEWER

- **ICLR 2023,** Neural Fields across Fields Workshop

OPEN SOURCE

Winter of Code at DeepFusionAI, Remote - *Open Source Developer* DEC 2020 - FEB 2021

- Worked on Social Distance Detector using MobileDets.

Contributor at TensorFlow

- Fixed various bugs and added a feature in the main TensorFlow GitHub repository.
    - Added the axis argument in tensorflow.keras.losses.categorical_crossentropy() and in tensorflow.keras.losses.binary_crossentropy(). [Link]
    - Fix for Conv1DTranspose, Conv2DTranspose, Conv3DTranspose layers when filter=0 by adding a ValueError. [Link]

SKILLS

Language(s): Python, JAVA, C, C++, HTML, CSS

Frameworks: PyTorch, TensorFlow, Keras

Tools: Blender, VSCode, PyCharm, Colmap

Awards

Sony Research India Annual Awards - FY 23: Best paper-Gold Category for “Open-Set Object Detection By Aligning Known Class Representations”. [Link]

5G Hackathon: Phase 1 winner, among Top 100 ideas on our project “Fertilizer Management System” in 5G Hackathon 2021. Our team won a prize of Rs 1,00,000. [Link]

Prayas 2020: A national level inter-college project competition. Our team stood 3rd for building real-time object detection on mobile devices.

Hult Prize: Our team stood second in Hult Prize OnCampus at Netaji Subhash Engineering College 2020.