

# Siddhant Bansal

138, New Boys Hostel, IIIT Hyderabad, 500032  
[Email](#) | [Website](#) | [Google Scholar](#) | +91-9904768487

## EDUCATION

**International Institute of Information Technology, Hyderabad, India** **2020 - Present**  
MS by Research in Computer Science (CGPA: 9.2/10)

**Vishwakarma Government Engineering College, Ahmedabad, India** **2015 - 2019**  
Bachelor of Engineering in Electronics and Communication (CPI: 8.09/10)

## WORK EXPERIENCE

**CVIT, IIIT-Hyderabad | Research Fellow | Advisors: Dr. C.V. Jawahar and Dr. Chetan Arora** **August 2019 - Present**

- My View is the Best View: Procedure Learning from Egocentric Videos | [ECCV 2022](#) | [Project Page](#)
  - Proposed the EgoProceL dataset for Procedure Learning using Egocentric Videos.
  - Created the Correspond and Cut (CnC) framework for Procedure Learning that outperformed the state-of-the-art on the benchmark ProceL and CrossTask datasets by 5.2% and 6.3%, respectively.
- Ego4D: Around the World in 3000 Hours of Egocentric Videos | [CVPR 2022 \(ORAL\)](#) | [Project Page](#)
  - Created the I3D ResNet baseline for the Hand and Object benchmark in Ego4D.
  - Worked on IRB application, consent forms, and de-identification for egocentric data collection.
- Improving Word Recognition using Multiple Hypotheses and Deep Embeddings | [ICPR 2020](#) | [Project Page](#)
  - Proposed EmbedNet and Confidence based Accuracy Booster (CAB) to achieve an improvement of 10% in the word recognition accuracy.
- Word Retrieval and Recognition using OCR text and Deep Embeddings | [DAS 2020 \(ORAL\)](#) | [Project Page](#)
  - Improved word accuracy by 1.6% for Indian languages by merging text hypotheses and deep embeddings.
  - Improved word retrieval by 11.12% for Hindi by proposing techniques like Naive Merge and Query Expansion.

**IIT Gandhinagar | Research Intern | Advisor: Dr. Shanmuganathan Raman | [Internship Website](#)** **March 2019 - August 2019**

- **3D Modeling and Geometry Processing:** Implemented a custom autoencoder for solving the problem statement of point cloud completion (Details and results on internship webpage). Also worked on ICP and Triangulation algorithms.

**Meditab Software (India) Pvt. Ltd. | Artificial Intelligence Intern | [Internship Website](#)** **Sept 2018 - March 2019**

- **ELOPE (Evolutionary Layout Optimization and Evaluator):** Generated optimal facility layouts, by implementing ELOPE and using it with the Genetic Algorithm. Decreased traveling time by 75% for DosePacker robots

**Bennett University, Greater Noida | Artificial Intelligence Research Intern | [Internship Website](#)** **June 2018 - July 2018**

- **Footprint Classification:** Successful in classifying humans with up to 5 different foot sizes, by developing a custom Convolutional Neural Network trained on a dataset created with the help of 180 volunteers, using a paper scanner.

**Bioscan Research, Ahmedabad | Data Analyst Intern** **April 2018 - June 2018**

- **Patient Tracker:** Created software capable of tracking patients using Python and SQLite, leading to better workflow for the people working on collecting brain scans.

## PROJECTS

**Google Summer of Code 2022: Enhancing DBpedia with image-based querying | [Project Page](#) | [GitHub](#)** **June 2022 - Present**

- Working on a) framework to create an image-based knowledge graph out of existing DBpedia entries; b) using the graph created to perform tasks like image querying, text + image search, and using relevant input images to add more images to existing articles.

**Scene Text Detection and Recognition** **July 2020 - August 2020**

- Worked on detecting and recognising text in the images from Mobility Assistant for Visually Impaired (MAVI).

**Automatic Garbage Detection and Collection** **May 2018 - April 2019**

- Created a device capable of autonomously detecting and picking up the garbage by detecting waste bottles using CNN (MobileNets), developed algorithms for estimating the depth of the garbage, and estimating the path using PID.

## PUBLICATIONS

- Siddhant Bansal, Chetan Arora, and C.V. Jawahar, "My View is the Best View: Procedure Learning from Egocentric Videos", in [ECCV 2022](#). ArXiv: [2207.10883](#)
- K. Grauman, ..., Siddhant Bansal, ..., C.V. Jawahar, ..., Jitendra Malik, "Ego4D: Around the World in 3,000 Hours of Egocentric Videos", in [CVPR 2022 \(ORAL\)](#). ArXiv: [2110.07058](#)
- Siddhant Bansal, Praveen Krishnan, and C.V. Jawahar, "Improving Word Recognition using Multiple Hypotheses and Deep Embeddings", in [ICPR 2022](#). ArXiv: [2010.14411](#)
- Siddhant Bansal, Praveen Krishnan, and C.V. Jawahar, "Fused Text Recogniser and Deep Embeddings Improve Word Recognition and Retrieval", in [DAS 2020 \(ORAL\)](#). ArXiv: [2007.00166](#)
- Siddhant Bansal, Seema Patel, Ishita Shah, Prof. Alpesh Patel, Prof. Jagruti Makwana, and Dr. Rajesh Thakker. "AGDC: Automatic Garbage Detection and Collection." ArXiv:[1908.05849](#)