Anurag **Bagchi**

(+412) 287 3504 | anuragba@andrew.cmu.edu | https://miccooper9.github.io/

Experience

Carnegie Mellon University & Toyota Research — *Pittsburgh*

Research Associate III under Dr. Martial Hebert, Robotics Institute — *Aug 2023 - Present*

- [**Submitted** at **CVPR 26** | **Results**] *Led*, *formulated* & *implemented* **Action** conditioned *Video* **World Models** that **beats NWM** (*CVPR 25 Best Paper Hon'ble*) at **3-DoF** *egocentric* navigation by **47%** while being **6**× **lower** in *latency* and **2**× **higher** *resolution*.
- Seamlessly extends to the **1x EVE humanoid** showing **25-DoF** *egocentric* Joint Angle Control, for *Navigation* and *Manipulation*.
- The World Model also demonstrates ***Zero-shot*** Navigation in **Paintings** and *real-world* **Unseen** environments.
- [**Published** at **ICCV 25** | **Twitter Thread**] *Led*, *formulated* & *implemented* **ReferEverything** which **beats LLM + SAM2** (*CVPR 25*) in ***zero-shot* OOD** *non-object* refer segmentation by **44.5%** while using **2000** × fewer training masks.

Bytedance/TikTok AI Lab — *Sinapore*

Computer Vision Engineer, TikTok Brand Safety — *March 2023 - Aug 2023*

- I worked in **Prof . Song Bai's** team, where I *launched* **11** Multimodal models (**Vision**, **Audio**, & **Text**) for policy-detection tasks and *automated* the iteration process for **21** ASR models, enhancing TikTok's policy-violation detection at ×**10M scale**.

TikTok R&D Singapore — *Sinapore*

Machine Learning Engineer, Video & Push Recommendation — *March 2021 - Feb 2023*

- End-to-End ML at ×**100M scale**, from problem formulation & feature engineering to model design , A/B testing and deployment.
- *Led* & *designed* the ML pipeline to leverage *user-feedback* from notifications *as a training signal* for TikTok's Global Recommendation.
- Improved **DAU** (**+0.07%**), user **staytime**(**+3%**), **click-rate** (**+15%**), system **latency** (**-3%**) and **memory** usage(**-2.5%**).

IIIT Hyderabad — *Hyderabad, India*

Research Assistant under Dr. Ravi Kiran and Dr. Makarand Tapaswi, CVIT — *Oct 2020 - Feb 2021*

- *Proposed and implemented* the **first-ever** Audio-Visual **framework** for *Temporal Action Localisation* in Videos.
- Achieved SoTA on **ActivityNet-v1.3** and **Thumos14** datasets at the time of publication.

Samsung Research — *Bangalore, India*

Intern , Imaging R&D Team — *May 2019 - July 2019*

- Received a **Full Time Offer** for modifying different functions in the mobile *camera service suite* using **ToF depth** data.
- I was awarded the *Samsung* **Advanced Programmer** certificate for solving a *3 hour* long *open-ended coding challenge* on my first try.

Indian Statistical Institute Kolkata — *Kolkata, India*

Research Intern under Dr. Swagatam Das, ECSU — *Jan 2020 - Dec 2020*

- Researched and developed *Differential Evolution* based methods for *Adversarial attacks* on SoTA image classifiers.
- Experimented with *Learning* based *Clustering* algorithms

Artificial Intelligence Lab, Jadavpur University — *Kolkata, India*

Research Assistant under Dr. Amit Konar — *2018 - 2020*

- *Designed & implemented* a Vision actuated system for **detecting** and **grasping** objects with **brain commands**.
- This work was later published in the **Journal** of *Biomedical Signal Processing* and *Control*.
- Published two other papers applying *Convolutional Neural Networks* to process **EEG** signals from *Brain computer Interfaces*.

Education

Carnegie Mellon University — *Pittsburgh, USA*

Master of Science in Robotics, School of Computer Science — *2023 - 2025*

- **Advisor : Dr. Martial Hebert, 100% scholarship**, **GPA** : **3.95/4.0**

Jadavpur University — *Kolkata, India*

Bachelor Of Electronics and Telecommunication Engineering — *2016 - 2020*

- Academia **CGPA : 9.35/10** | Standardised Test Scores : **GRE (331/340)** **TOEFL iBT (113/120)**

Skills & Interests

Programming Python, C, C++, MATLAB

Deep Learning Keras, PyTorch, TensorFlow, OpenCV, High-Performance clusters

Interests Computer Vision, Multimodal Learning, Reinforcement Learning, Deep Learning, Machine learning, A.I.

Academic Service

Reviewer **NeurIPS-24,25; ICLR-25,26; CVPR-25; ICML-25**, **TPAMI**, Elsevier (IF : **8.139**) *Aug 2021 - Present*

Publications

Egocentric Action-Conditioned Video World Models ***CVPR 26***

Anurag Bagchi, Zhipeng Bao, Homanga Bharadhwaj, Yu-Xiong Wang, Pavel Tokmakov, Martial Hebert ***Under Submission***

- Video Diffusion models, Humanoids, Action Controlled World Models.[**View results**]

ReferEverything: Towards Segmenting Everything We Can Speak of in Videos ***ICCV 2025***

Anurag Bagchi, Zhipeng Bao, Yu-Xiong Wang, Pavel Tokmakov, Martial Hebert

- Text-to-Video Diffusion model, Refer Segmentation, Vido-Language grounding [**View paper**][**Project Page**]

Hear Me Out: Fusional Approaches for Audio Augmented Temporal Action Localization ***VISIGRAPP 2022***

Anurag Bagchi, Jazib Mahmood, Dolton Fernandes, Ravi Kiran Sarvadevabhatla ***Oral***

- Video Understanding, Temporal Action Localisation, Graph Neural Networks, Audio-Visual fusion[**View paper**][**View code**]

UniGen-AR: Unifying Visual Generation with Auto-Regressive Modeling ***CVPR 26***

Zhipeng Bao, Zhen Zhu, Nupur Kumari, **Anurag Bagchi**, Yu-Xiong Wang, Pavel Tokmakov, Martial Hebert ***Under Submission***

- Autoregressive Image Generation, Unified perception.

Autonomous grasping of 3-D objects by a vision-actuated robot arm using Brain–Computer Interface ***Biomedical Signal Processing and Control***

Arnab Rakshit; Shraman Pramanick *; **Anurag Bagchi***; Saugat Bhattacharyya *IF : 5*

- RGB-D Grasp prediction, Object Detection, Inverse Kinematics, BCI, EEG classification [**View paper**]

Unpublished Projects

Deep Power K-means for high dimensional clustering ***Autoencoders - Deep Clustering - Representation Learning***

Indian Statistical Institute, ECSU *Aug 2020 – Dec 2020*

- Developed a deep clustering framework based on Power K-means that jointly optimises the power means objective together with the auto-encoder loss from Deep K-means while simultaneously learning low dimensional cluster representations in each iteration. The differentiable surrogate in deep K-means is effectively replaced with the kolmogorov mean, while the annealing step is modified to achieve a series of smoother alternatives to the K-means objective.

Novel sparse whitebox adversarial attack on image classifier networks using differential evolution. ***Deep Differential Evolution - Adversarial Attacks***

Indian Statistical Institute, ECSU *Jan 2020 – Aug 2020*

- Implemented and investigated a tunable Differential Evolution based scheme for whitebox adversarial attacks on multiple state-of the-art classifiers like Resnet 50, 101 and VGG-16.. The parameters for evolution are learned using a gradient descent scheme to reach the ideal balance between exploitation and exploration and increase the fooling rate.

Achievements

2023	**Full(100%) Scholarship for MS-Research**, Carnegie Mellon University	*Pittsburgh, USA*
2020	**University Rank 1 in Final semester (grade : 10/10)**, Jadavpur University	*Kolkata, India*
2019	**Samsung Advanced Programming Certificate**, Samsung Research	*Bangalore, India*
2016	**Top 0.1%(99.9 percentile)**, State Engineering Entrance Exam	*West Bengal*
2016	**Top 0.4%(99.6 percentile)**, National Engineering Entrance Exam	*India*

Relevant Coursework

Undergraduate	Linear Algebra, Pattern Analysis and Machine Intelligence,Data Structures and Algorithms, Digital Image Processing, Probability and Statistics, Operating Systems, Digital Signal Processing
Graduate	3D Vision, Robot Learning, Advanced Computer Vision, Math for Robotics, Mechanics of Manipulation