# AMIT KUMAR RANA

+49–17676432044 ◇ Aachen (NRW), Germany

amitranavsr94@gmail.com ◇ akrana94 ◇ amitrana001 ◇ amitrana001.github.io

## EDUCATION

**Master of Science, Media Informatics** 2020 - 2023
**CGPA : 1.4/4.0** *Best: 1.0, Sufficient: 4.0*
RWTH Aachen University *Aachen, Germany*

**Bachelor of Technology, Computer Science and Engineering** 2013 - 2017
**CGPA : 7.3/10.0** *Best: 10.0, Sufficient: 4.0*
Indian Institute of Technology *Kanpur, India*

## SKILLS

**Languages:** Python [PyTorch, TensorFlow, Keras] **Basic:** C, C++, Java, SQL, PHP
**Tools:** VS Code, Pycharm, Google Colab, Jupyter Notebook, Git
**Others:** Distributed Training of Deep Networks, Slurm Batch Manager

## RESEARCH

**Computer Vision Group, RWTH Aachen University** Aachen, Germany
Master's Thesis Student *June 2022 - May 2023*

- Developed a **novel and efficient** architecture for segmenting multiple object instances using user interactions as queries to a **Transformer Decoder** without needing to re-compute image features during refinement.
- **ImageNet** pre-trained **Swin Transformer** is used as a backbone to extract features, with a **multi-scale deformable-attention Transformer** on top to extract multi-scale features.
- The queries are initialized using multi-scale features based on user interactions and then are updated via multi-layered **cross-attention and self-attention** modules in the Transformer decoder.
- During the refinement stage, new interactions are generated in the error region between the predicted and ground truth segmentation masks.
- Achieved **state-of-the-art** results on multiple existing interactive segmentation benchmarks and the proposed **new multi-instance benchmark**.

## EXPERIENCE

**Machine Learning Engineer** Jun 2017 - Nov 2020
Samsung R&D Institute *Noida, India*

- Designed a mechanism to detect **pose** on the live camera preview. Also led the development of a **TensorFlow-lite** and **OpenPose**-based library, capable of automatically matching and saving the best pose.
- Designed a **music recommendation system** for the Samsung music application. Implemented a library that could collect user data along with extracting **audio/music features** (Pitch, genre, entropy, etc.) and train its recommendation model (using **libSVM**) on an android device.
- Developed an automatic volume control system based on conversation detection between the device user and another person in real-time. It included **voice activity detection (VAD)** and voice recognition using **Gaussian Mixture Models (GMMs)**.
- Mentored two teams of **four interns** each in the development of **Generative Adversarial Networks (GANs)** based **Facial Attribute** editing application and CNN based **image-grouping** application.

**Data Analyst Intern** May 2016 - July 2016
Barclays Technology Center *Pune, India*

- Analyzed and presented the working of **py4j** and **PySpark** to the data team. Assisted the data team in the task of transferring large data sets from JVM to python using PySpark.

## PROJECTS

**Video Future Frames Generation** Oct 2021 - Mar 2022
University of Bonn, Lab MA-INF 4308 *Bonn, Germany*

- Implemented **Video Ladder Network** based architecture; a deep encoder-decoder based hierarchical model, augmented by recurrent and feed forward connections at all layers, for future frames generation.
- Trained the models autoregressively on **Moving MNIST** and **KTH Action** datasets.
- Benchmarked this against well-known recurrent models including **LSTM, Conv LSTMs, and GRUs** along with different feed-forward connections including **DCGAN, ResNet, and VGGNet** based models.
- Code and report: https://github.com/Dhagash4/video-prediction

**Analysis of Loss Landscape Topology for Biased Training Data** Apr 2021 - Sep 2021
University of Bonn, Lab MA-INF 4306 *Bonn, Germany*

- Analyzed topology of loss landscape to study the correlation between flatness around minima and generalization behavior of networks trained on multiple biases **(mislabeling, Gaussian noise, and skewness)**.
- Experimented with **Resnet18** and **ShaResNet** architecture on multiple versions of the **CIFAR-10** dataset with induced biases.
- Investigated and implemented **filter-normalized random directions** techniques to visualize loss landscape.
- Code and report: https://github.com/ali-mohammadi-scrc/ML_Lab

**Real Time Vehicle Detection and Classification** Jan 2016 - Apr 2016
IIT Kanpur, Machine Learning Techniques *Kanpur, India*

- Built an end-to-end system for vehicle **detection and classification** for Surveillance Video in real-time.
- Experimented with different feature representations **(HAAR, SIFT, HOG)** along with different learning algorithms such as **SVM, Random Forests, Neural Networks** etc.

## PUBLICATIONS

**A. K. Rana**, S. Mahadevan, A. Hermans, B. Leibe (2023). DynaMITe: Dynamic Query Bootstrapping for Multi-object Interactive Segmentation Transformer. *arXiv*:2304.06668

## TEACHING EXPERIENCE

**MA-INF 4228 - Foundations of Data Science** Apr 2022 - Sep 2022
Dr. Michael Nuesken, University of Bonn *Bonn, Germany*

- Responsible for organizing tutorials and preparing assignment problems.
- Assisted in formulating quizzes for all the video lectures to help review and grasp the lecture content.

## AWARDS AND ACCOMPLISHMENTS

- Recipient of **Dean's List** in the academic year 2020-21 at RWTH Aachen University.
- Recipient of the **best innovation award**, only awarded to **1% of Samsung R&D** engineers.
- Ranked in the top **0.01%** (out of 1.2M candidates) in IIT - Joint Entrance Exam, 2013.

## RELEVANT COURSEWORK

- Machine Learning Techniques
- Advanced Machine Learning
- Computer Vision
- Advanced Computer Vision
- Introduction to Computer Graphics
- Foundations of Data Science
- Principles of Database System
- Bayesian Data Analysis
- Introduction to Game Theory
- Data Structure and Algorithms
- Discrete Math
- Linear Algebra & ODE
- Abstract Algebra
- Probability and Statistics
- Fundamentals of Calculus