Pulkit Gera

Machine Learning Engineer

+49 1742767785 pgera@mpi-inf.mpg.de Saarbrücken, Germany
pulkit-gera darthgera123 Portfolio

Profile

Passionate Machine Learning and Computer Vision enthusiast with extensive hands-on experience in developing cutting-edge projects, including relightable photorealistic rendering of humans and HDR radiance field representations for indoor scenes.Previously worked on a wide variety of full-stack development and other projects, contributing to significant advancements in AR/VR applications and achieving publications in renowned conferences.

Education

MS by Research, Computer Science **IIIT Hyderabad, India**
CGPA: 10.0/10.0 2021-2022
Thesis:Casual Scene Capture and Editing for AR/VR Applications.

BTech, Computer Science **IIIT Hyderabad, India**
CGPA: 8.77/10.0 2017-2021
Relevant Courses: Software Engineering, Data Structures, Operating Systems, Computer Vision, Computer Graphics, Unix Tools and Scripting, Mobile Robotics & Statistical Methods in AI

Professional Experience

Research Intern **Saarbrücken, Germany**
VCAI, Max Planck Institute for Informatics Advisor: Dr Christian Theobalt 10/2022 - present
- Formulated a novel framework based on 3D Gaussian Splatting that enables real-time photorealistic rendering of real humans from novel views under arbitrary input environment map illuminations.
- Setup and calibrated a specialized USC Lightstage consisting of 331 programmable lights and 40 RED Komodo cameras. Curated preprocessing pipelines for processing terabytes of captured data.
- Curated a novel dataset consisting of 70 human avatars under varying light patterns such as OLAT and spherical gradients.
- Curated synthetic datasets using RenderPeople assets under varying illumination using Blender and scripted using Python.

Research Intern **Quebec City, Canada**
Université de Laval Advisor: Dr Jean Francois Lalonde 05/2021 - 10/2022
- Designed a novel NeRF-based representation for a causally captured full HDR radiance field of a large indoor scene without any elaborate setups or complex capture protocols.
- Predicts full HDR panoramas from any location of the scene and can synthesize correct lighting effects which enable augmentation of indoor scenes with synthetic objects.
- Funded by MITACS-Globalink. Presented at BMVC 2022 and OmniCV2022,CVPR-W.

Research Assistant **Hyderabad, India**
CVIT, IIIT Hyderabad Advisor: Dr P.J. Narayanan 06/2019 - 05/2022
- Developed a novel neural rendering framework for simultaneous view synthesis and appearance editing of a scene with known environmental illumination captured using a mobile camera.
- Disentangled the appearance and learned an independent lighting representation which enables editing of the appearance of real scenes in interesting and non-trivial ways. Presented at ICVGIP 2021.

Deep Learning Intern **Remote(Bengaluru,India)**
Segmind Solutions 08/2020 - 10/2020
- Ported various object detection and semantic segmentation algorithms into a single unified tensorflow library
- Abstracted framework reduced time for prototyping and increased ease of setup by reducing to 5 lines of code.

Technical Staff **Remote(Geneva, Switzerland)**
AICrowd 08/2020 - 10/2020
- Designed over 200 machine learning challenges of different modalities such as tabular data, images, etc that were used in various contests.
- Made benchmarks and tested different research-based challenges of different topics.

Computer Vision Intern **Hyderabad,India**
DreamVu 06/2020 - 08/2020
- Designed image-processing tools to benchmark data captured by the PAL omnidirectional vision system.
- Assisted in capturing data ind improving image quality captured by the camera system.

Publications

- **Casual Indoor HDR Radiance Capture from Omnidirectional Images.** **Pulkit Gera**, Mohammad Reza Karimi Dastjerdi, Charles Renaud, P. J. Narayanan, Jean-François Lalonde. British Machine Vision Conference 2022
- **Neural view synthesis and appearance editing from unstructured images.** **Pulkit Gera**, Aakash KT, Dhawal Srikonda, PJ Narayanan. ACM Indian Conference on Computer Vision, Graphics and Image Processing 2021

Skills

- **Languages:** Python, C++, Bash, Javascript, GO, Matlab, SQL, Ruby on Rails
- **Frameworks:** Pytorch, Tensorflow, Metashape, COLMAP, Mitsuba3, OpenGL, Langchain
- **Technologies:** Git, Meshlab, Blender, Docker, SLURM

Projects

RAG Agents *GenAI, Information Retrieval*

Developed a Retrieval-Augmented Generation (RAG) system leveraging LLMs with LangChain for backend processing and integrated with a Streamlit-based frontend to answer questions across various databases and modalities, including coffee orders, research papers, and image datasets.

UV-Relighting *Computer Vision*

Predict relit UV map based on input environment maps, involving mesh and UV map recovery for captured 3D subjects and employing techniques like latent vector conditioning and diffusion-based relighting to achieve accurate illumination under various conditions.

LANet *Computer Vision*

End-to-end trainable luminance attentive network with two streams for HDR reconstruction trained on Laval Indoor HDR dataset.

Relight My NeRF *Computer Vision*

Pytorch port of ReNe. NeRF based novel view synthesis of an object under novel OLAT conditions.

Ray Tracing *Computer Graphics*

Ray Tracing engine built in C++ than can perform volume rendering and handle basic geometries

Airplane Simulator *Computer Graphics*

3D aircraft simulator using OpenGL and C++, featuring a mini-world that includes an airplane, checkpoint islands, a volcano island, and enemy parachutes.

POSIX Shell *Operating Systems*

An implementation of the POSIX Shell from scratch in C using only system calls which supports functionality like piping and redirection along with a few new custom commands.

QuizApp *Fullstack Development*

Interactive quiz portal with a Go (Gin) backend and a React frontend, allowing administrators to create and manage various quiz types while providing users with a dynamic and engaging platform to participate in quizzes.

Multimodal Summarization of News Articles *NLP*

Generate high-quality summarizations of news articles by taking text and image modalities and leveraging visio-linguistic transformers like OSCAR and UNITER.

Guess the name *Information Retrieval*

Created a crowdsourcing platform with a Flask frontend to fill gaps in Hindi Wikidata by connecting loosely related entities.

Additional Experience and Achievements

- Ranked 1841 in JEE Main 2017 out of approximately 1 million applicants.
- Winner of Megathon 2019, Big Data Track (India's Largest student organized hackathon).
- Dean's merit list awardee given to top 3% of the class.
- Finished 3rd in AICrowd Blitz,out of 350+ teams.
- Student Placement Cell, IIIT Hyderabad
- Head Finance, E-Cell IIIT Hyderabad