[Samsung PRISM] End Review Report

Video Enhancement And Restoration Using Neural Radiance Field

1.College Professor(s): Uma Mudenagudi uma@kletech.ac.in
Ramesh Ashok Tabib ramesh_t@kletech.ac.in

2.Students:
Ajeet Shyanavad 01fe21bcs277@kletech.ac.in
Vineet Gajanan Pai 01fe21bcs273@kletech.ac.in
P Sai Sudheer 01fe21bcs096@kletech.ac.in
Srivatsa Malipatil 01fe21bcs038@kletech.ac.in

•3.Department: Computer Science And Engineering

Date: 14 Aug 2024

Work-let Area - AI, ML | Video enhancement & Restoration using Neural Radiance Fields(Nerf)

PRISM

[Introduction]

- The problem involves video restoration, specifically the task of enhancing low-quality frames to obtain high-quality ones, such as in video super-resolution, Video Denoising and Video Deblur. Unlike single image restoration, video restoration necessitates leveraging temporal information from multiple adjacent frames, which are often misaligned. Current deep learning approaches address this challenge through the use of either a sliding window strategy or a recurrent architecture. However, these methods are either limited to frame-by-frame restoration or lack the capability to effectively model long-range dependencies in the video data. To overcome these limitations, a solution employing Neural Radiance Field can be explored

[Problem Statement]

- Design and develop an Video Enhancement model improve Video Quality
- Build a deep learning based model that are trained on huge database covering different location, background and

DB Collection	DL Engine Building	Testing	Primary tasks

[References]

Inverting the Imaging Process by Learning an Implicit Camera Model
Xin Huang, Qi Zhang, Ying Feng, Hongdong Li, Qing Wang

[Output Expectations and KPI]

- The objective of the project is to recommend a better pose based on background, location and foreground.
  - Develop a robust learning based solution to predict a pose
  - KPI:
    - Accuracy: Better than Current SOTA
    - Speed: < TBD
    - Deliverables
      - International conference paper co-authored
      - Mobile portable solution
      - Detailed technical documentation

3 Members | 5 Months

- Mentor
  - Mayank Anchlia (m.anchlia, +7387925811)
  - Tanisha (tanisha.2, +91-74-8875-0264)

[Trainings]

- Good knowledge on fundamentals of computer vision & Deep Learning Concepts
- Good hands on in Deep learning development using frameworks like Tensorflow/Keras
  - Training & Testing using CPU/GPU
- Basic knowledge on Camera Model

[Timeline]

Kick Off < 1&2nd Month >
- Problem Understanding
- Data Collection & Preparation
- Complete Trainings
- Literature Survey

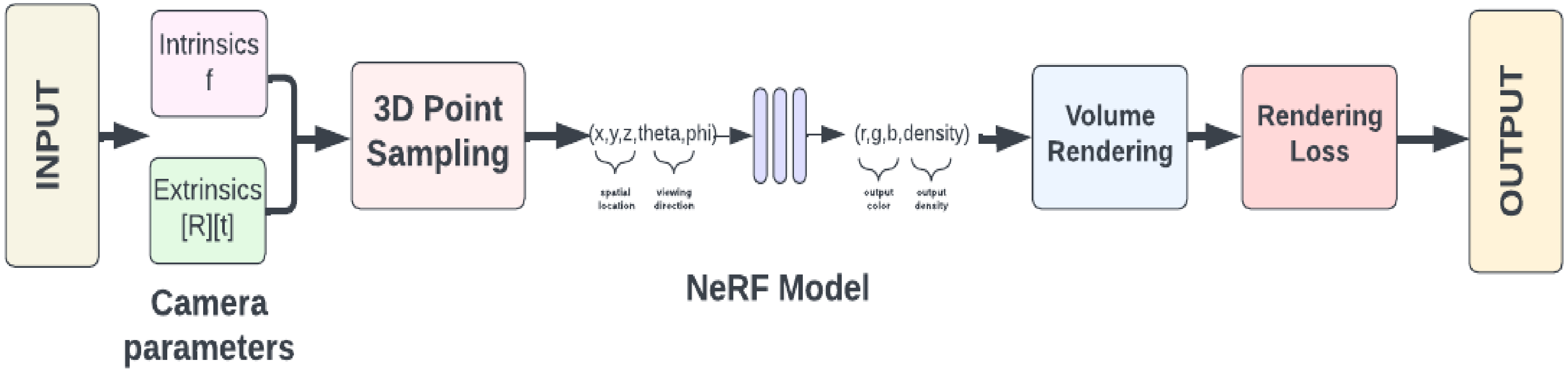
Milestone 1 < 3rd Month >
- State of art approach implementation on PC
- Preliminary results on PC
- 1st Paper draft

Milestone 2< 4th Month >
- Enhancements
- Model Optimization for Mobile

Closure < 5th Month >
- Final paper
- Final Demo on Mobile Device
- Technical Paper
- Final Review Presentation

Approach / Solution

- <u>Concept Diagram</u> :
  (Clear detailed schematic / block diagram / flow chart depicting the proposed concept / solution)



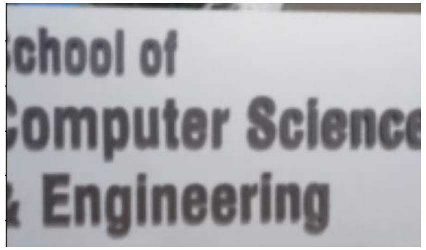
Dataset Description

- <u>Dataset Capture / Preparation / Generation</u> :

- **Standard Datasets**: Diverse NeRF datasets, including NeRF-Synthetic, Blender Scenes, and LLFF, provide a range of synthetic and real-world scenes with complex geometries and lighting conditions.
- **Custom Dataset**: Frames from three static videos were extracted and tailored to the academic environment for specific model training and evaluation.
- The combination of standard and custom datasets ensures the NeRF model is trained to handle a wide variety of real-world scenes.

Experimental Results

- **Results :**

GROUND TRUTH	INPUT (Degraded)		OUTPUT
		PSNR: 32.66 SSIM: 0.89	
		PSNR: 32.66 SSIM: 0.91	
		PSNR: 34.32 SSIM: 0.93	

Experimental Results

- Zoomed In Results :

Ground truth

Input (Degraded)

Output

Experimental Results

GROUND TRUTH

INPUT (Degraded)

PSNR: 25.25

SSIM: 0.78

OUTPUT

Comparison of PSNR values for different methods

Model	PSNR
NeRF	27.54(10000 epochs)
NeRF-Bi	26.42(10000 epochs)
NeRF-Liif	27.07(10000 epochs)
NeRF-Swin	26.34(10000 epochs)
Our model	25.25(5000 epochs)

Experimental Results

- **Major Observations / Conclusions & Challenges :**

- Our model was trained on three custom datasets, which included frames from static videos with varying complexities, as well as standard datasets like NeRF-Synthetic, Blender Scenes, and LLFF. This comprehensive approach exposed the model to a broad spectrum of real-world conditions.

- Our model showcased superior performance, achieving PSNR values ranging from 33 to 35 on custom datasets for 2000 epochs and 25.25 for 5000 epochs on standard datasets.

- The combination of custom and standard data validated the training technique, enhancing the model's precision in handling complex geometries and lighting, and proving its robustness in producing higher-quality outputs.

Deliverable

- **Paper :**

Abstract—Video enhancement and restoration have emerged as critical tasks in computer vision, especially for applications like video streaming, film remastering, and surveillance. Traditional methods often struggle with artifacts, loss of details, and inadequate handling of temporal consistency. Neural Radiance Fields (NeRF), originally designed for novel view synthesis in 3D scenes, have shown promise in addressing these challenges. This paper explores the application of NeRF in video enhancement and restoration, leveraging its ability to model complex spatial and temporal dependencies. We propose a method that uses NeRF to generate high-quality, temporally consistent frames from low-resolution or degraded video sequences. Experimental results demonstrate significant improvements in video quality, surpassing state-of-the-art methods in both subjective and objective evaluations.

- **KPIs Delivered :**

1. Accuracy: Better than Current SOTA
2. International conference paper.

Work-let Closure Details

- **Code Upload details:**

Items	Details
KLOC (Number OF Lines of codes in 000's)	1.5
Model and Algorithm details	Tiny NeRF
Is Mid review, end review report uploaded on Git ?	yes
Link for Git	https://github.ecodesamsung.com/SRIB-PRISM/KLE_23VIS50_Video_enhancement_and_Restoration_using_Neural_Radiance_Fields_Nerf

- **Data details (if applicable):**

Items	Data folder 1	Data folder 2	Data Folder 3
Name & Type of Data (Audio/Image/Video)	Video1	Video2	Video3
Number of data points	30	60	61
Source of Data (self collected, Scrapped, available on open source)	self collected	self collected	self collected
Google drive link/ git link to access data	https://drive.google.com/file/d/1bzUk1brCQJerFwCZqoRmZFP2LDb0M5uf/view?usp=sharing	https://drive.google.com/file/d/1bzUk1brCQJerFwCZqoRmZFP2LDb0M5uf/view?usp=sharing	https://drive.google.com/file/d/1bzUk1brCQJerFwCZqoRmZFP2LDb0M5uf/view?usp=sharing

Work-let Name: Video Enhancement And Restoration Using Neural Radiance Field

Worklet Details

1. Worklet ID: 23VIS50
2. College Name: KLE Technological University

Professor Comments

- Generated an enhanced video as the output of the model.
- Explored more scenes with different lighting conditions.
- Completed Project within given time

Date: 14/08/24

Thank you