JUST NeRF IT !!

Omri Green, Dheeraj Bhogisetty, Siyuan Huang,
Shreya Bang, Joel John

CONTENT

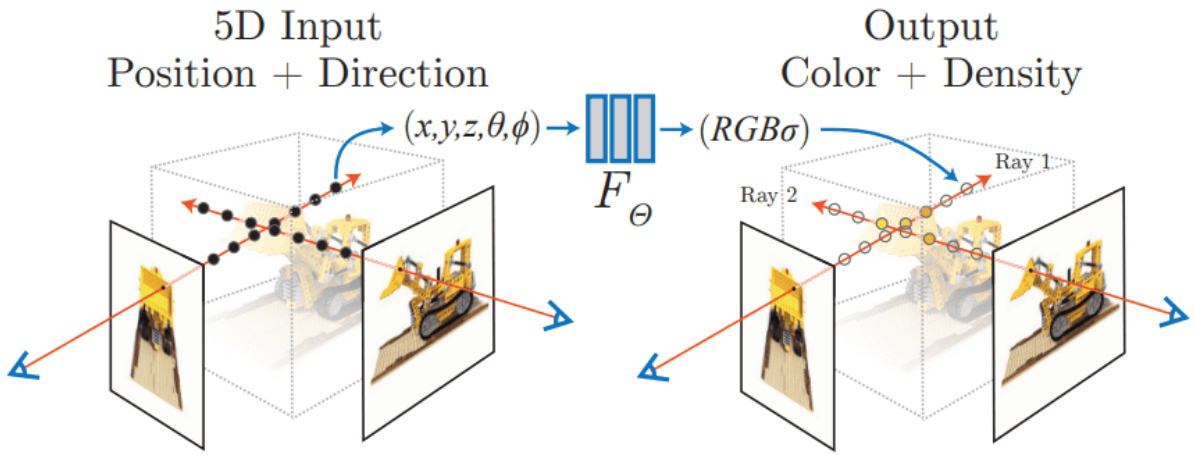
- Objective
- Introduction
- Pipeline
- Dataset Generation & COLMAP
- NeRF Architecture
- Results & Observations
- Challenges
- Future Scope

OBJECTIVE

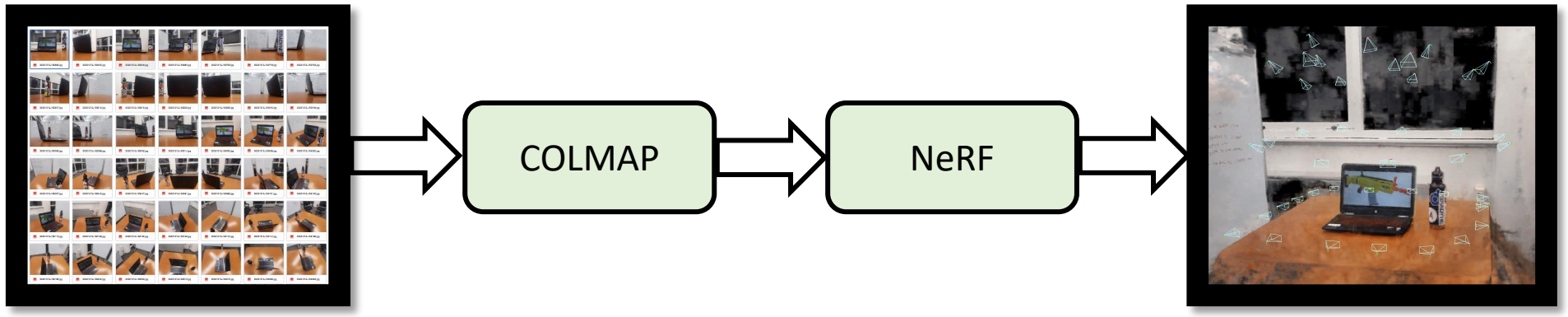
- To create a continuous volumetric (3D) scene given a sparse (discrete) set of images of the scene from different viewing angles and generates novel views utilizing a small number of input images. This is done using NeRF.
- To add labels to the generated scene which will enable users to replace the existing objects with other volumetric structures or simply add new objects to the scene.

INTRODUCTION

- **NeRF(Neural Radiance Field)** is a strategy to solve view synthesis problem without using a CNN.
- The Image is Synthesized by
  - Sampling 5D coordinates along camera rays
  - MLP outputs a color and volume density
  - Render to get the image

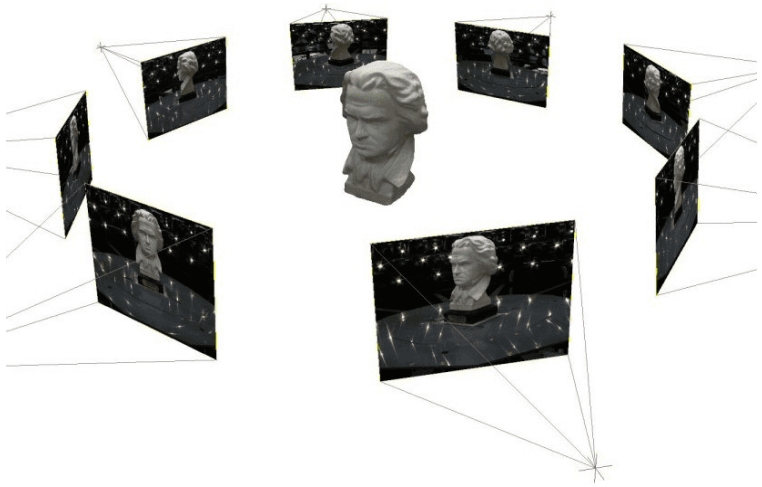


PIPELINE

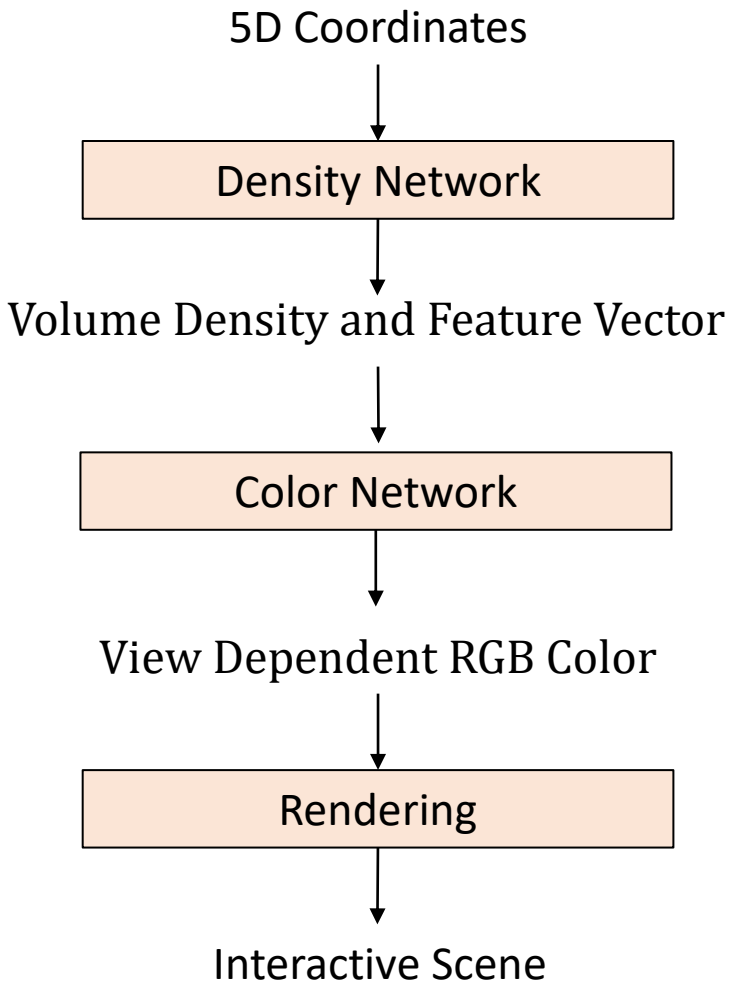


DATASET GENERATION & COLMAP

- 5D input vector - contains the spatial location and the viewing direction i.e., the pose of the camera.
- Camera pose - estimated using computer vision methods.
- COLMAP – open-source SfM package
- COLMAP calculates the intrinsic and extrinsic properties of the camera and the images. The output is a JSON file that maps each image with its pose.



NeRF ARCHITECTURE



- Inputs:

  - 5D Coordinates – Spatial Location and Viewing Direction

- Outputs:

  - Volume Density and View-Dependent Emitted Radiance

- 9 Fully Connected Layers of the Network

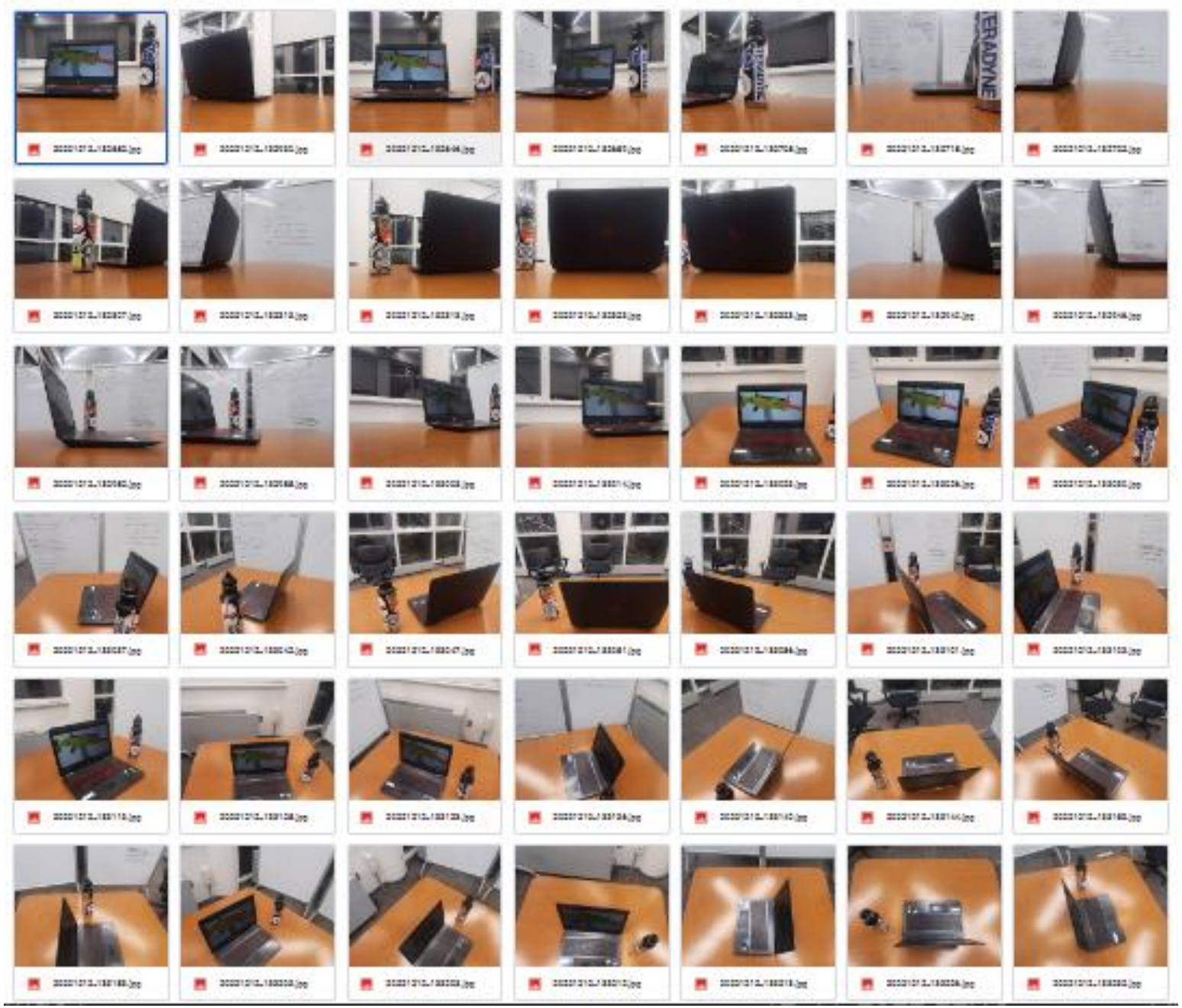
  - First 8 Layers produce volume density and a 256-dimension feature vector for feature convergence

  - Last Layer outputs a view dependent RGB color for rendering

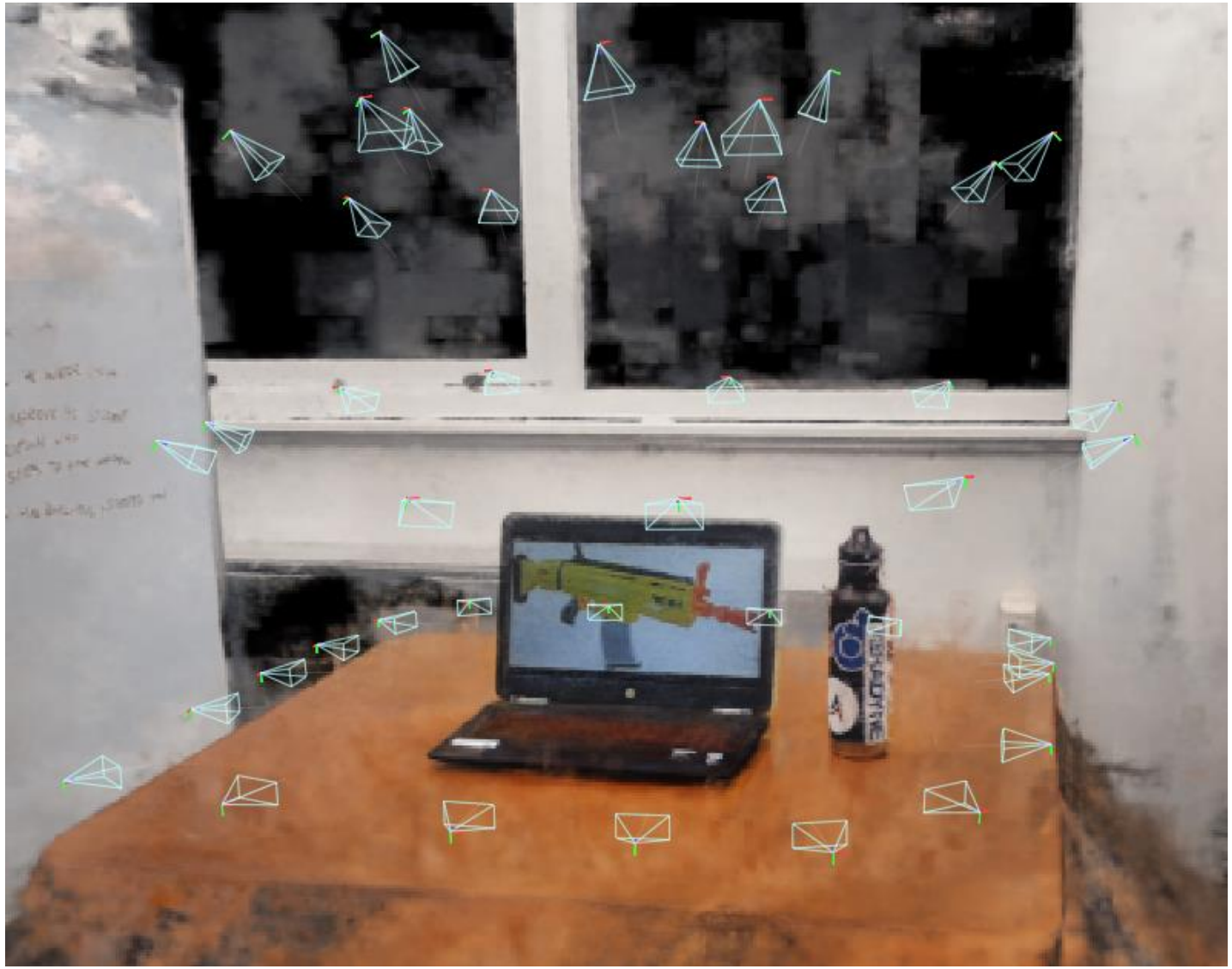
RESULTS

DATASET

- 42 images
- Taken at regular intervals covering the complete upper hemisphere



INPUT IMAGES WITH FRAMES

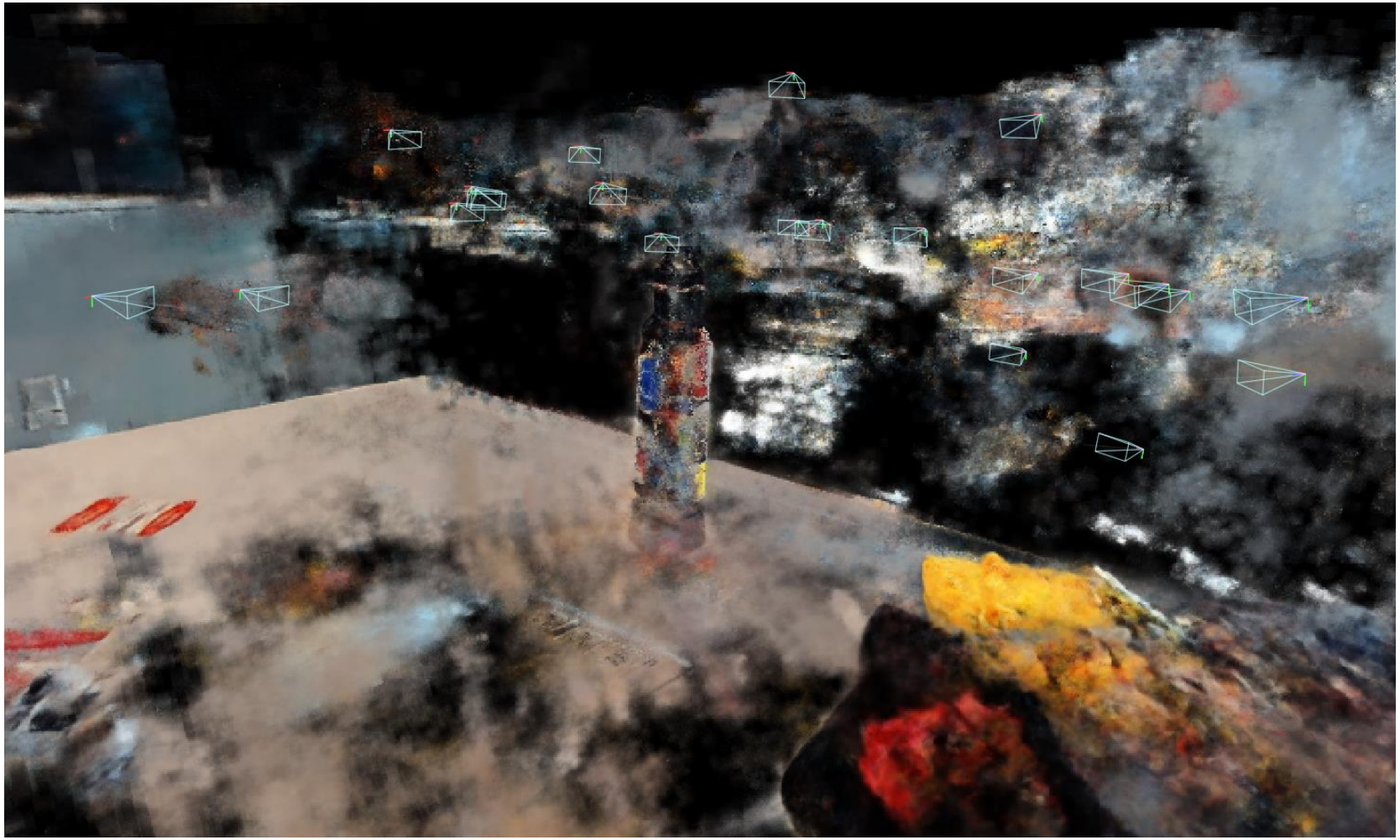



FORTNITE

OBSERVATIONS: PART- 1

- Reflections can be captured by NeRF
- Fully converged scene requires inputs images to capture the scene from all directions
- Convergence from real world scenes is not as good as synthetic scenes

RESULTS FROM A DIFFERENT DATASET



RESULTS FROM A DIFFERENT DATASET

OBSERVATIONS: PART- 2

- The input dataset should cover the scene from all angles to get good convergence.
- Does not work for dynamic scenes – objects in the scene must be static

CHALLENGES/LIMITATIONS

- Achieving faster processing
- Convergence for real world scenes
- Limitation on the Computational power
- Adding the extra heuristic to instant nerf

FUTURE SCOPE

- To enable user to add, remove or replace a object in a 3D volumetric scene -

1. Tiny Object Loader: To add 3D objects to the scene

2. Objects can be removed by modifying the NeRF network to include object label

QUESTIONS?

THANK YOU!