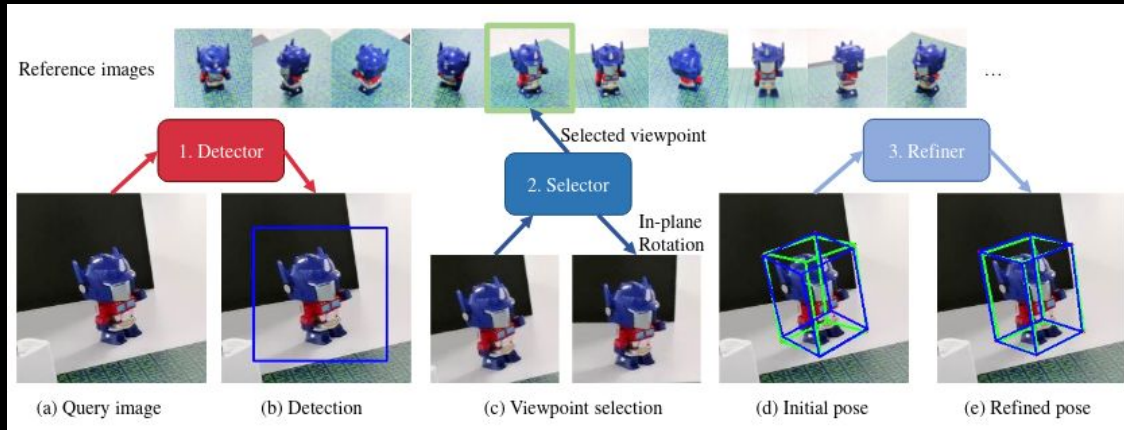


MCVM 2022/23

Seed Paper Introduction

Gen6D [Pengyuan]

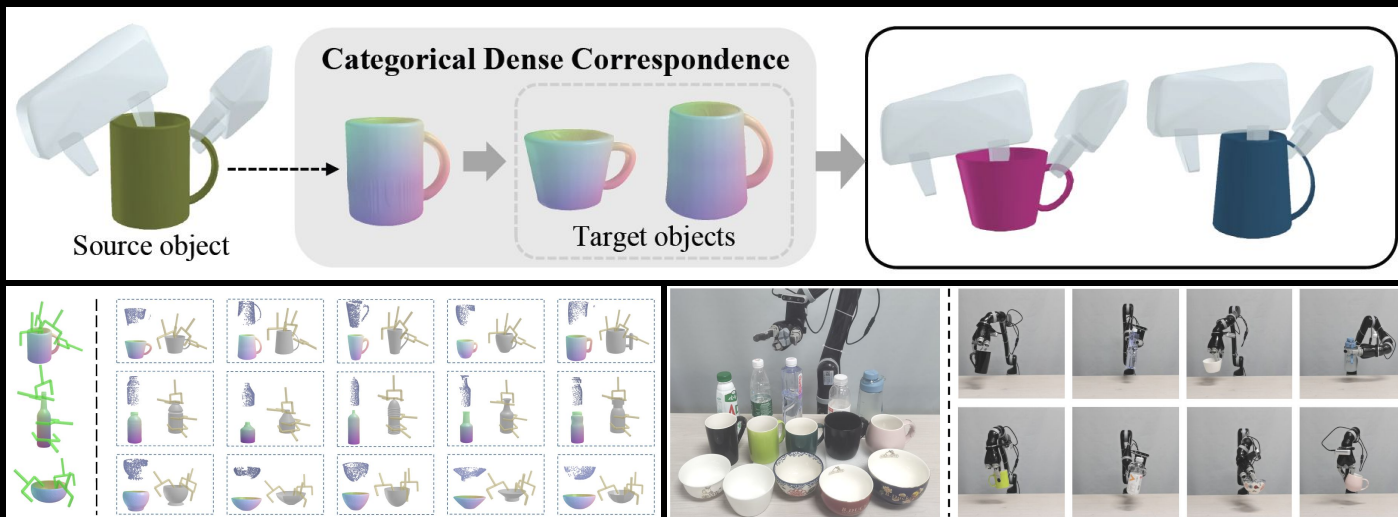
- Few shot 6D pose estimation based on RGB images
- Detector and selector to pick the most similar image in the reference images
- 3D Volume CNN for pose refinement



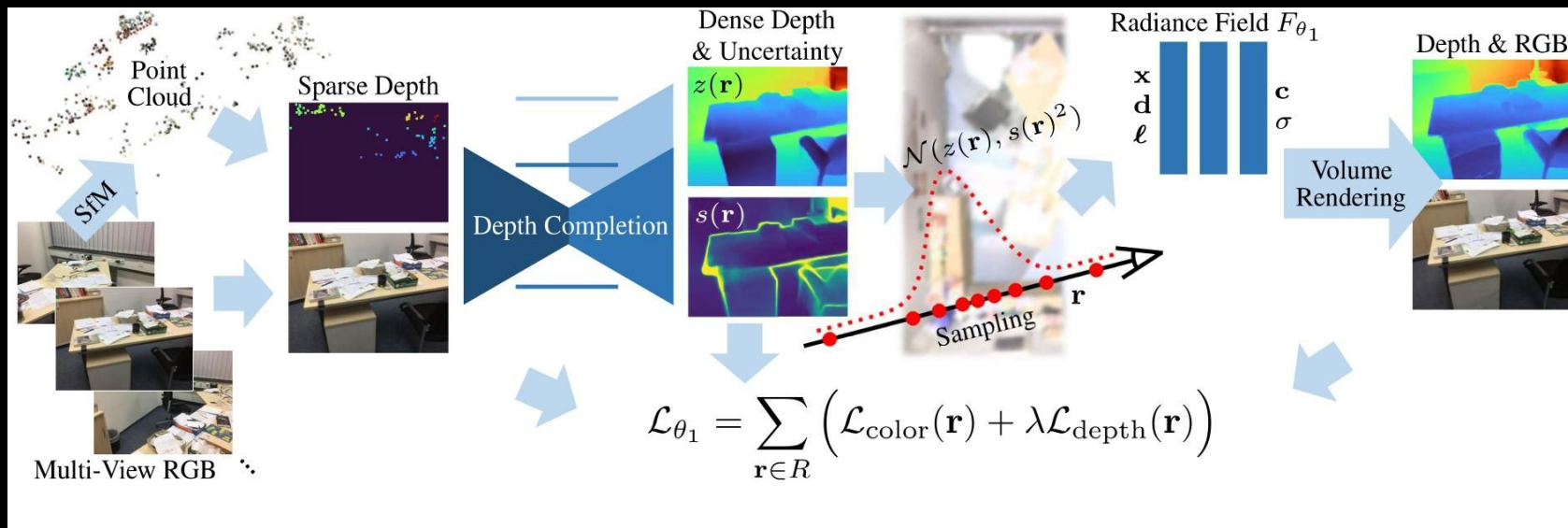
TransGrasp [Guangyao]

<https://github.com/yanjh97/TransGrasp>

- A model predicts grasp poses of a category of objects **by only labeling one instance.**
- Transfer and correct grasp pose across a category of objects.
- Both simulation and real robot system achieve high quality grasps.



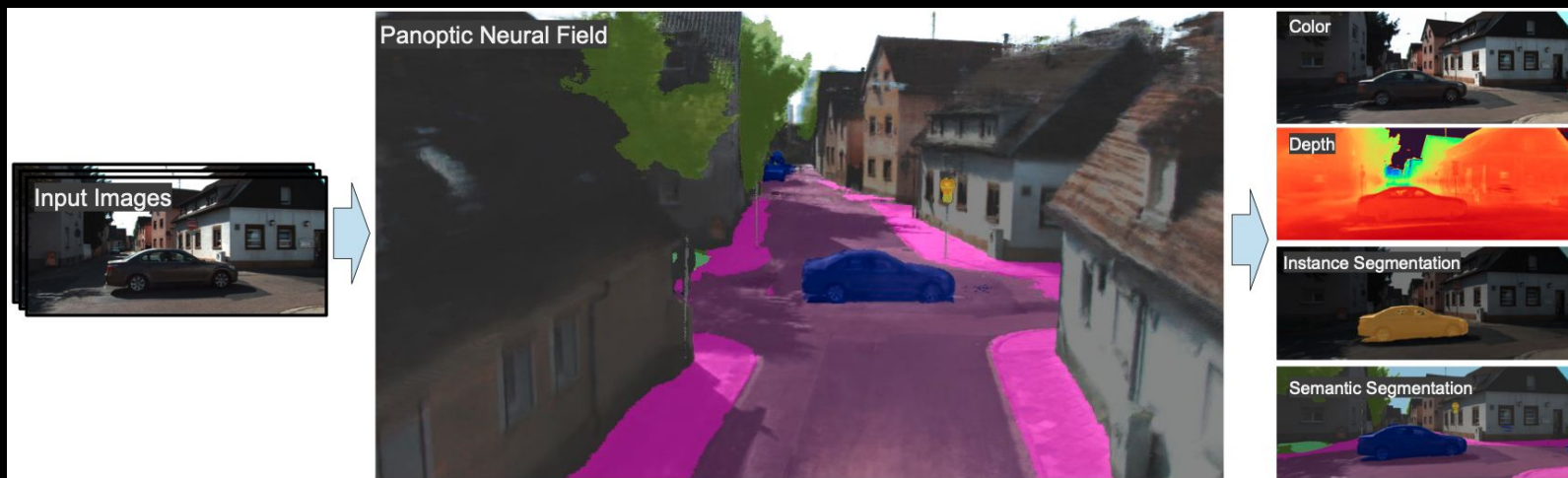
Dense Depth Priors [Patrick]



Panoptic Neural Fields (NeRF) [Stefano]

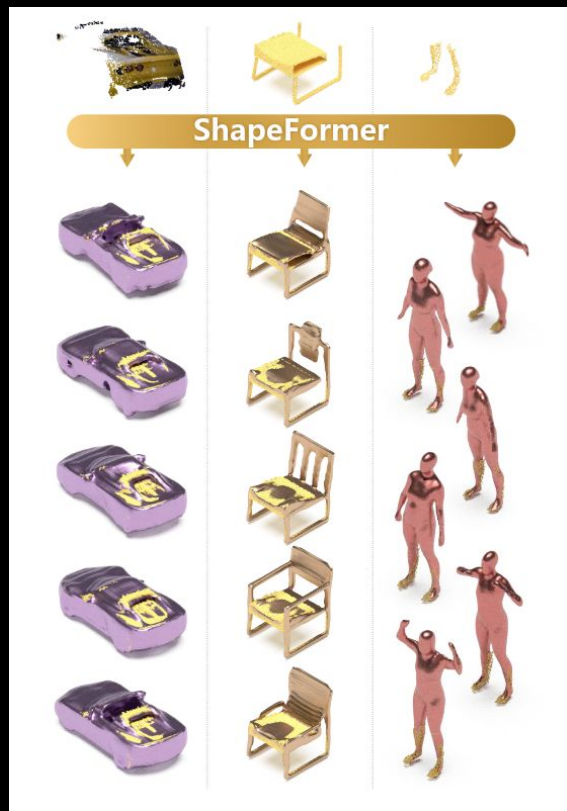
- object-aware neural scene representation
- scene into objects (things) and amorphous stuff background
- individual MLPs for objects and stuff

multiple scene understanding tasks

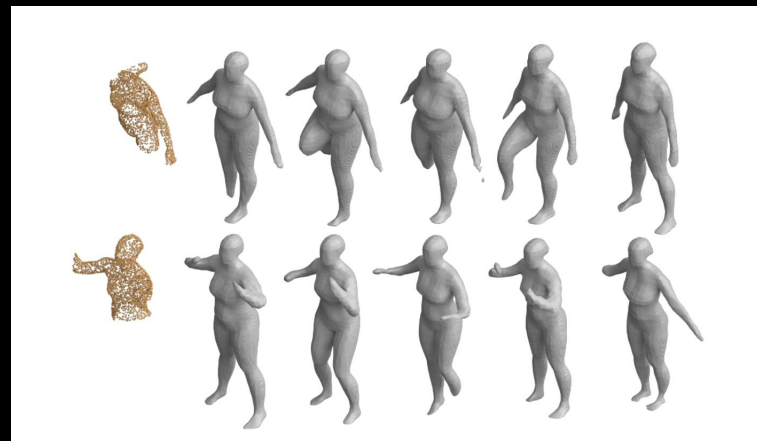
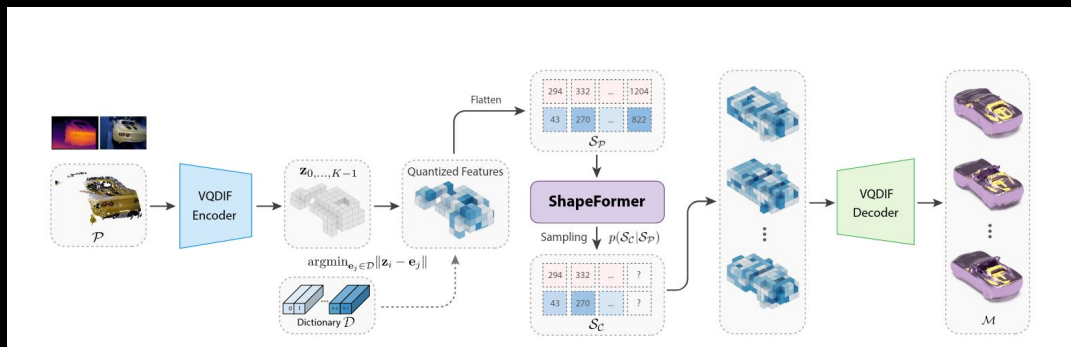


ShapeFormer [Lennart]

- Inferring missing information from noisy or partial scans has numerous real world applications
- Explore a novel approach using transformers and deep implicit functions
- Evaluate its performance and applications on several datasets including PartNet and Faust



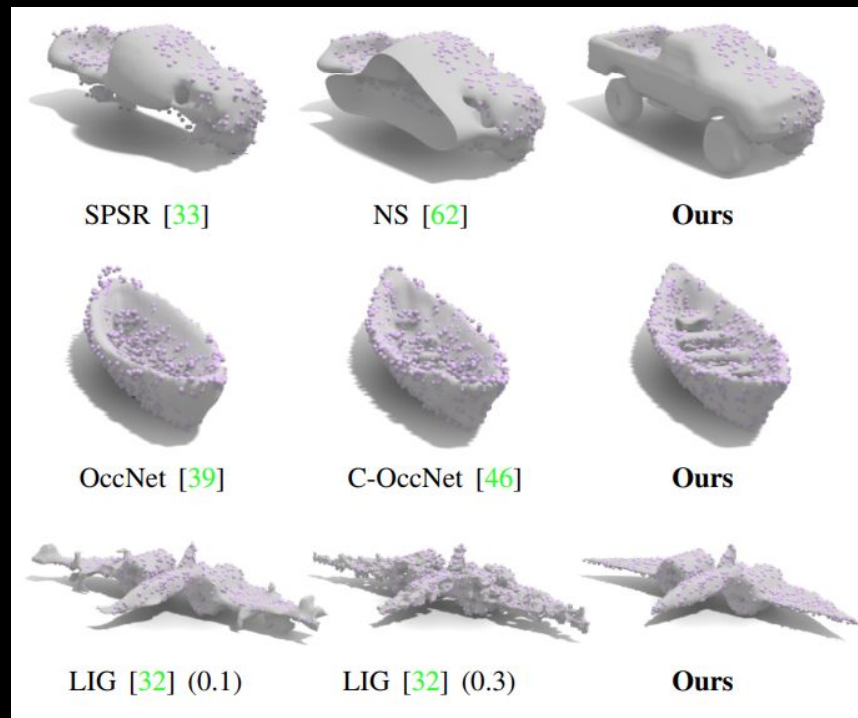
ShapeFormer [Lennart]



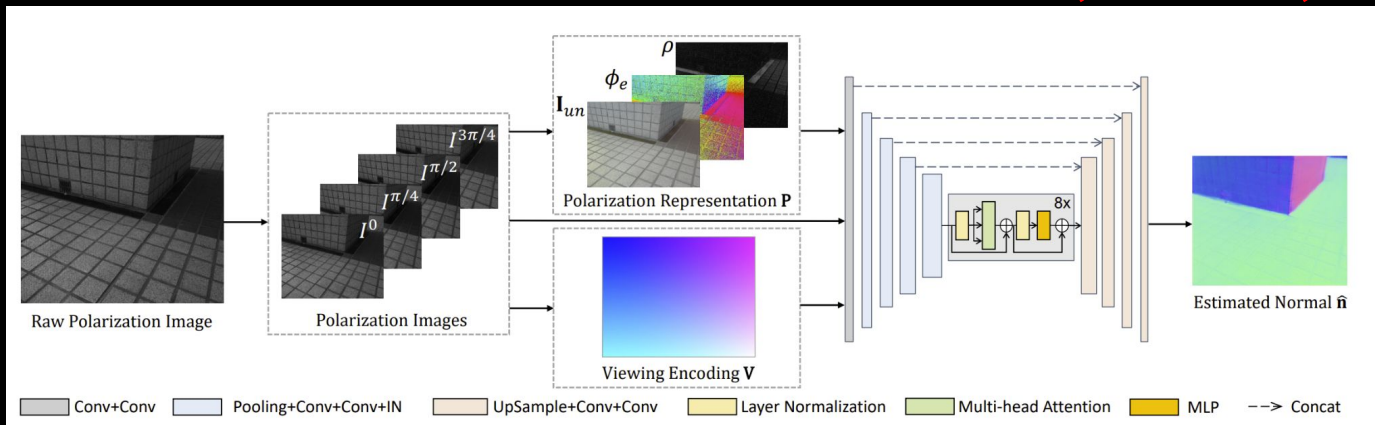
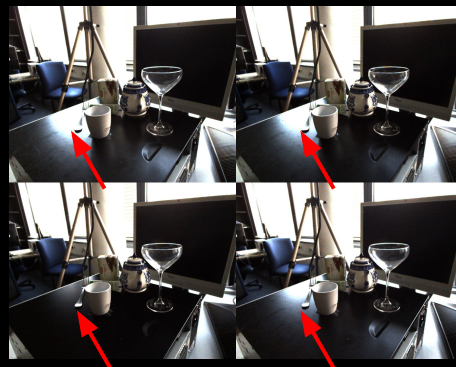
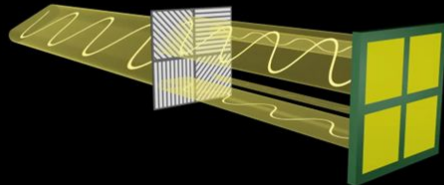
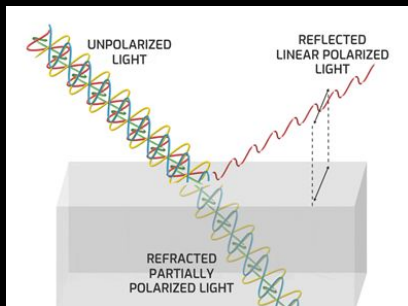
Neural Fields & 3D Reconstr. [Shun-Cheng]

Encoder-decoder network tends to overfit training shapes and cannot generalize to unknown shapes.

This paper enables inference-time optimization using partially known inputs to estimate accurate shapes.

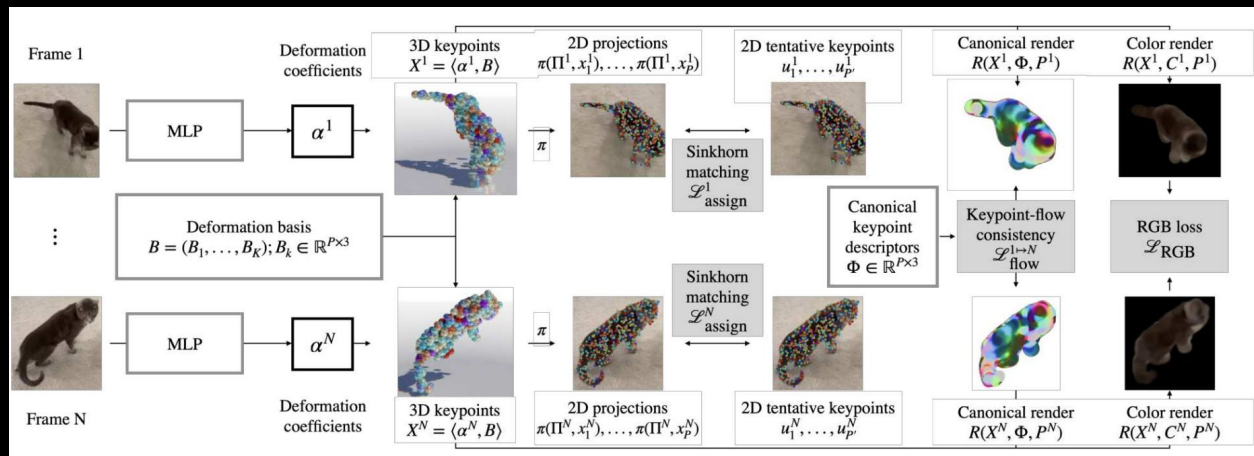


Shape From Polarization in the Wild [Hyun]



KeyTr: Keypoint Transporter [Mahdi]

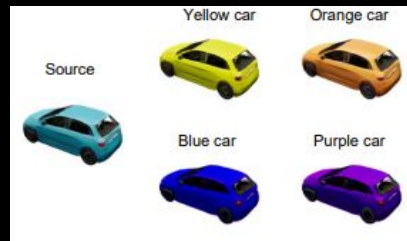
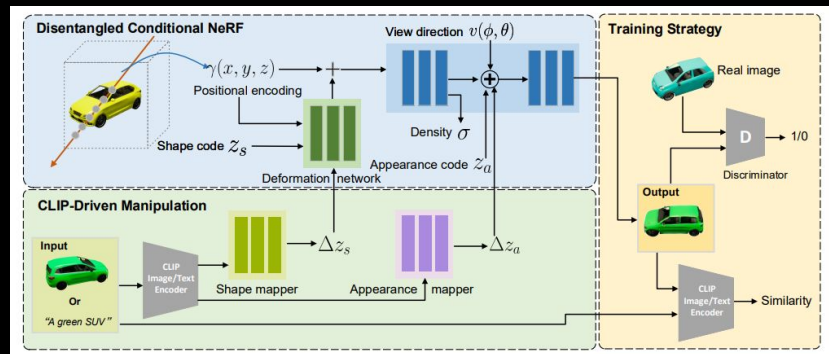
- Reconstructs 3d deformable objects from casual RGB frames
- Predicts and track keypoints
- Uses optical flow and optimal transport for consistency



CLIP-NeRF [Hannah]

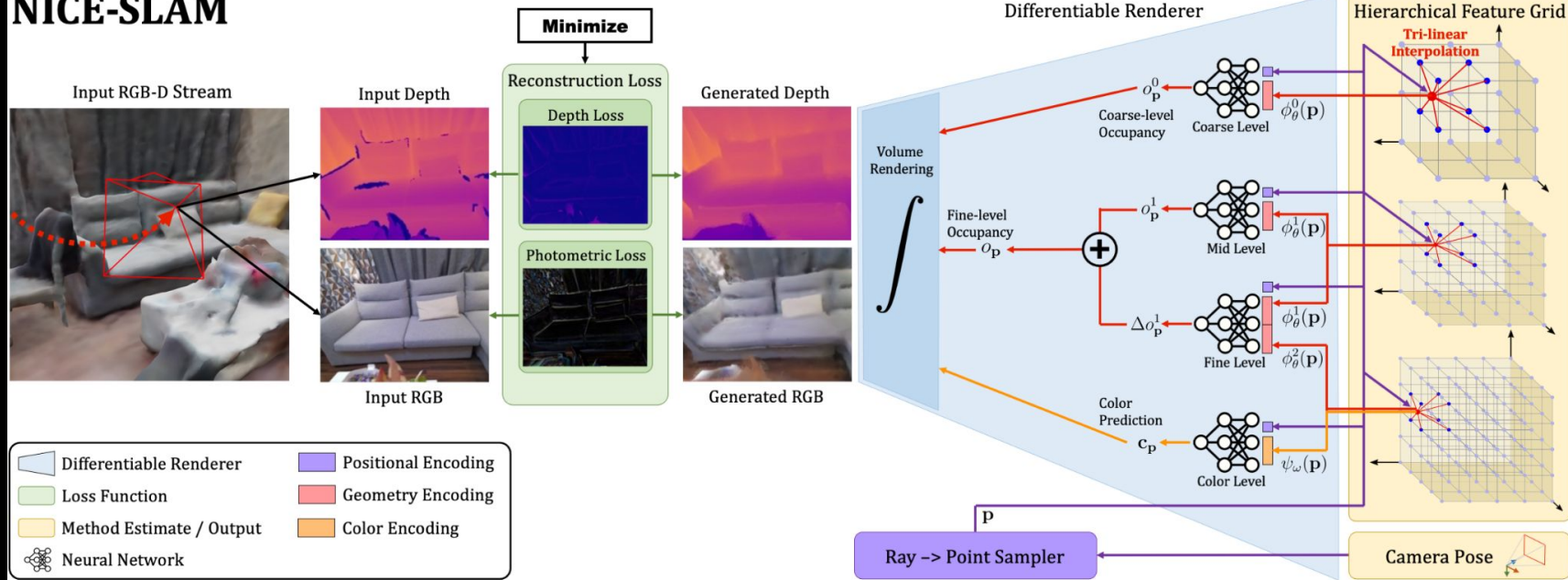
<https://github.com/cassiePython/CLIPNeRF>

- utilizes CLIP
- manipulating NeRFs leveraging joint language-image embedding (text-image driven manipulation)
- disentangled conditional NeRF which allows disentangling shape and appearance
- allow editing of existing data

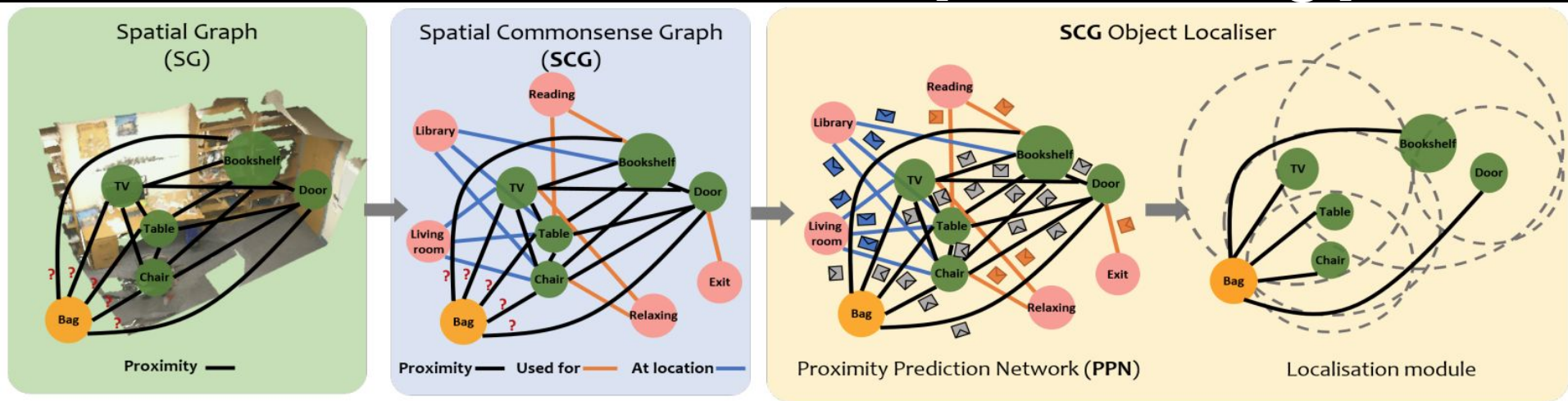


NICE-SLAM [Patrick]

NICE-SLAM



Commonsense Localization [Shun-Cheng]



“A method for estimating the unknown position of an object given a partial 3D scan of a scene.”

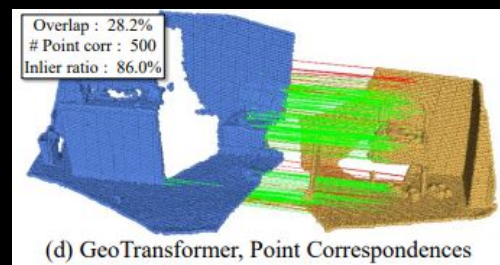
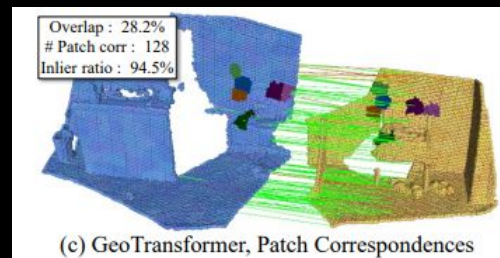
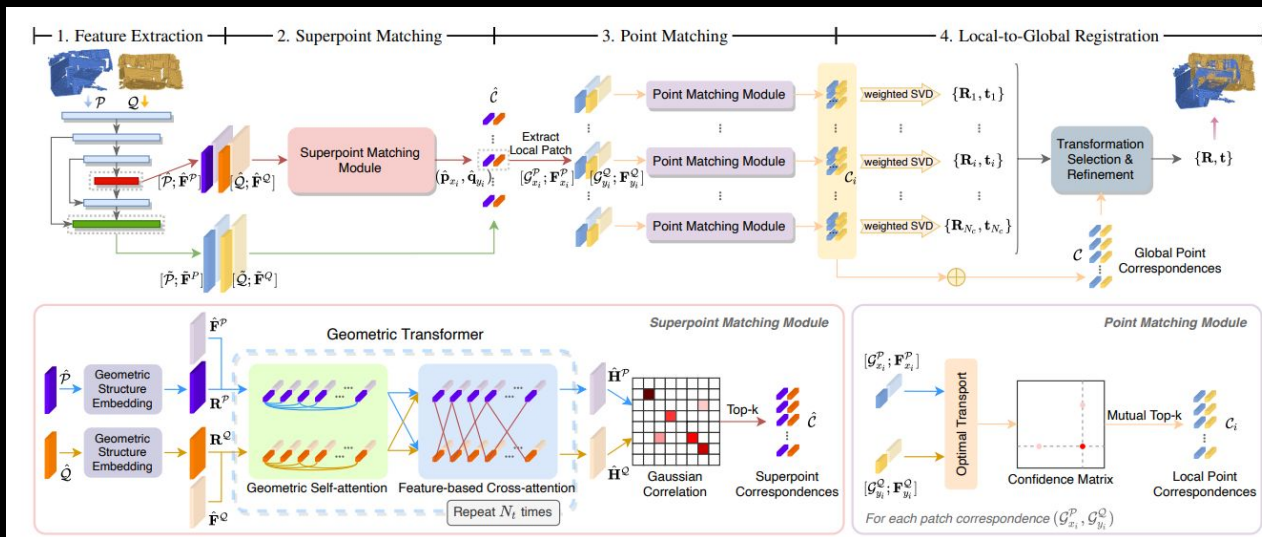
Spatial Commonsense Graph (SCG).

“... by injecting **commonsense knowledge** with in a scene graph representation, so that a machine can also reasonably localize an object in the unseen part of the scene, without the use of any visual/depth information.”



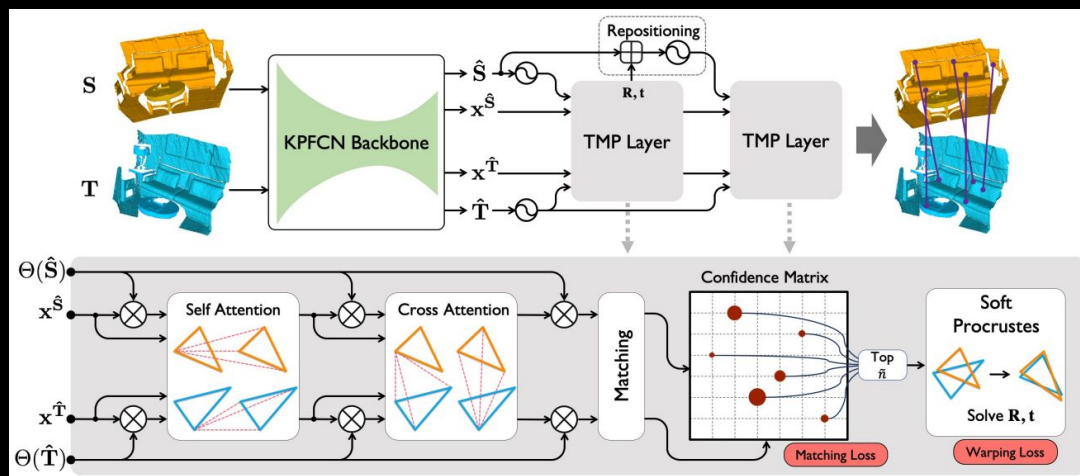
Geometric Transformer [Hao]

- Coarse-to-fine point matching in both rigid and deformable cases.
- Pairwise geometry-based self-attention module.
- Ransac-free point cloud registration.



Lepard [Mert]

- Rigid and deformable point cloud registration.
- An architecture that disentangles the feature and position space.
- Positional encoding on relative distance.
- A repositioning layer for improved matching.



Diffeomorphic Flow Registration [Lennart]

- Deep implicit templates for organ reconstruction
- Traditional shape correspondence works select a single template, which introduces bias
- Here a template is represented as a deep implicit function, and we learn a flow field to each shape while considering the topology of the shape

