DeepRob

Discussion 4
Overview of Final Project Topics I
University of Michigan and University of Minnesota

RGB-D Architectures

Core List

1. PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes, Xiang et al., 2018
2. A Unified Framework for Multi-View Multi-Class Object Pose Estimation, Li et al., 2018
3. PVN3D: A Deep Point-Wise 3D Keypoints Voting Network for 6DoF Pose Estimation, He et al., 2020
4. Learning RGB-D Feature Embeddings for Unseen Object Instance Segmentation, Li et al., 2021

Object Pose, Geometry, SDF, Implicit surfaces

Core List

1. SUM: Sequential scene understanding and manipulation, Sui et al., 2017
2. DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation, Park et al., 2019
3. Implicit surface representations as layers in neural networks, Michalkiewicz et al., 2019

Pre-tra… Archit…

Core List

1. SORNet: Spatial Object-Centric Representations for Sequential Manipulation, Yuan et al., 2021
2. CLIPort: What and Where Pathways for Robotic Manipulation, Shridhar et al., 2021
3. Masked Visual Pre-training for Motor Control, Xiao et al., 2022
4. R3M: A Universal Visual Representation for Robot Manipulation, Nair et al., 2022
5. Do As I Can, Not As I Say: Grounding Language in Robotic Affordances, Ahn et al., 2022
6. RT-1: Robotics Transformer for Real-World Control at Scale, Brohan et al., 2022

Tactile Perception for Grasping and Manipulation

Core List

1. More Than a Feeling: Learning to Grasp and Regrasp using Vision and Touch, Calandra et al., 2018
2. Tactile Object Pose Estimation from the First Touch with Geometric Contact Rendering, Bauza et al., 2020
3. Visuotactile Affordances for Cloth Manipulation with Local Control, Sunil et al., 2022
4. ShapeMap 3-D: Efficient shape mapping through dense touch and vision, Suresh et al., 2022

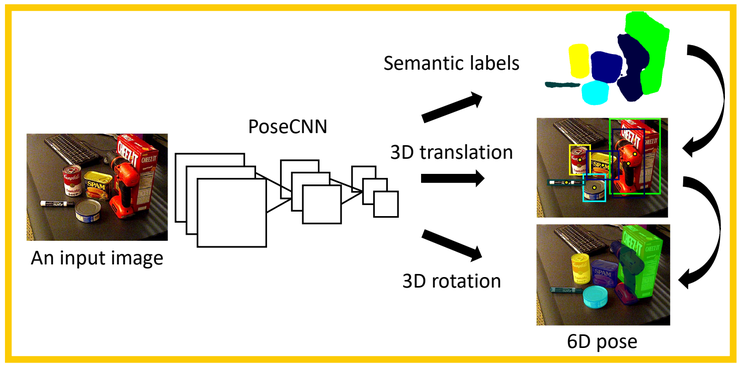
Neural Radiance Fields and Implicit Representations

Core List

1. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall et al., 2020
2. Object-Centric Neural Scene Rendering, Guo et al., 2020
3. Neural Descriptor Fields: SE(3)-Equivariant Object Representations for Manipulation, Simeonov et al., 2021
4. NeRF-Supervision: Learning Dense Object Descriptors from Neural Radiance Fields, Yen-Chen et al., 2022
5. NARF22: Neural Articulated Radiance Fields for Configuration-Aware Rendering, Lewis et al., 2022

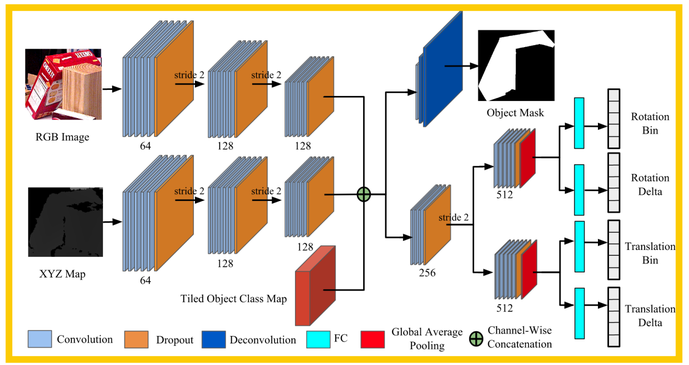
RGB-D Architectures

PoseCNN



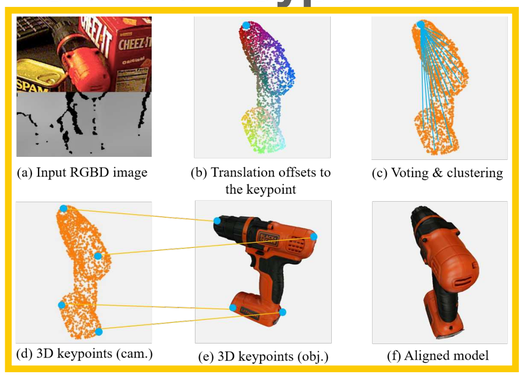
[1] Xiang et al., RSS 2018

Unified Multi-Class Pose Estimation



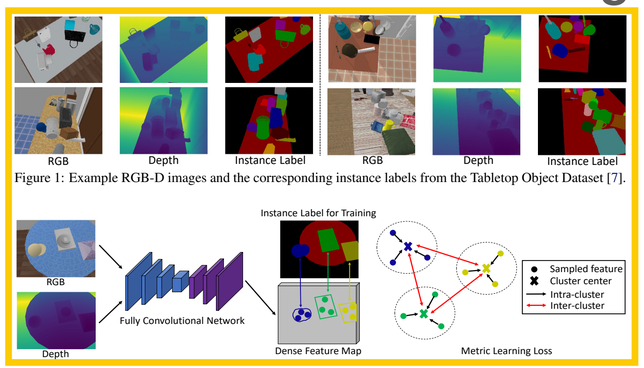
[2] Li et al., ECCV 2018

PVN3D: 3D Keypoint Voting



[3] He et al., CVPR 2020

RGB-D Feature Embeddings

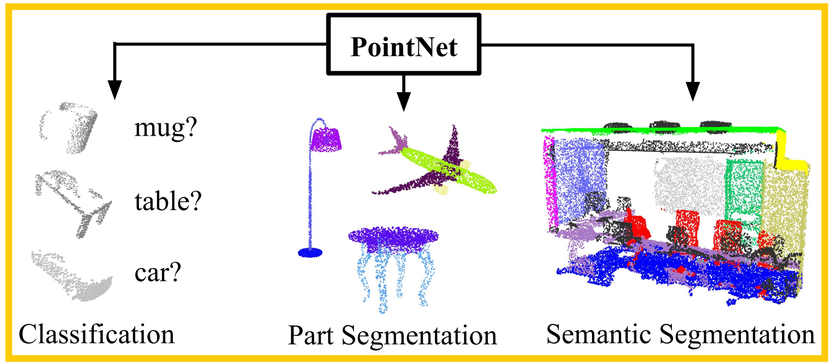


[4] Xiang et al., CoRL 2021

[1] Yu Xiang, Tanner Schmidt, Venkatraman Narayanan, Dieter Fox. “PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes” RSS, 2018.
[2] Chi Li, Jin Bai, Gregory D. Hager. “A Unified Framework for Multi-View Multi-Class Object Pose Estimation” ECCV, 2018.
[3] Yisheng He, Wei Sun, Haibin Huang, Jianran Liu, Haoqiang Fan, Jian Sun. “PVN3D: A Deep Point-wise 3D Keypoints Voting Network for 6DoF Pose Estimation” CVPR, 2020.
[4] Yu Xiang, Christopher Xie, Arsalan Mousavian, Dieter Fox. “Learning RGB-D Feature Embeddings for Unseen Object Instance Segmentation” CoRL, 2021.

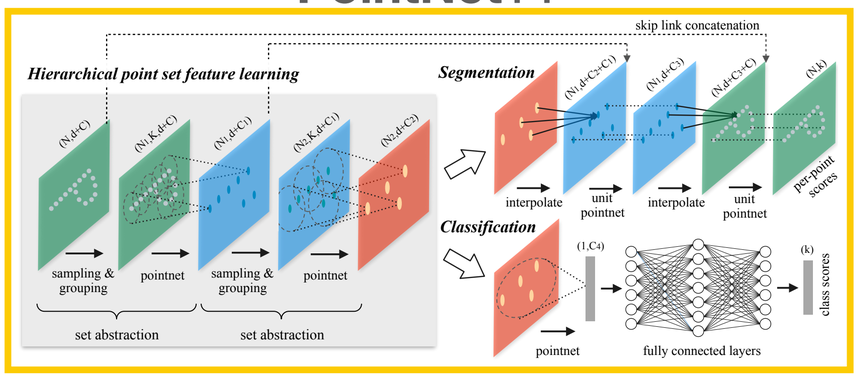
Pointcloud Processing

PointNet



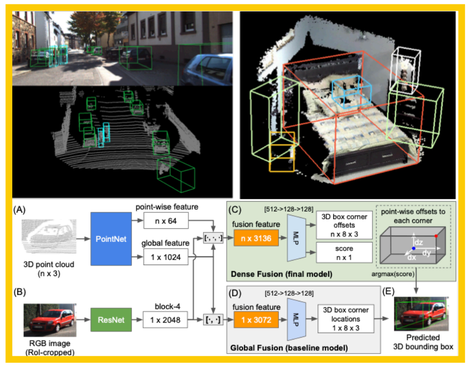
[1] Qi et al., CVPR 2017

PointNet++



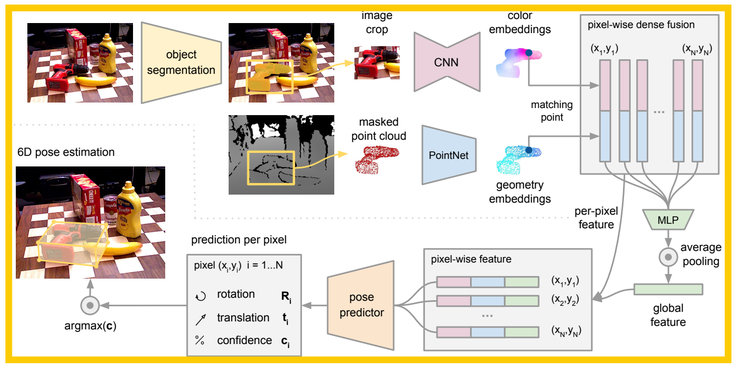
[2] Qi et al., NeurIPS 2017

PointFusion



[3] Xu et al., CVPR 2018

DenseFusion



[4] Wang et al., CVPR 2019

[1] Charles R. Qi, Hao Su, Kaichun Mo, Leonidas J. Guibas. “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation” CVPR, 2017.

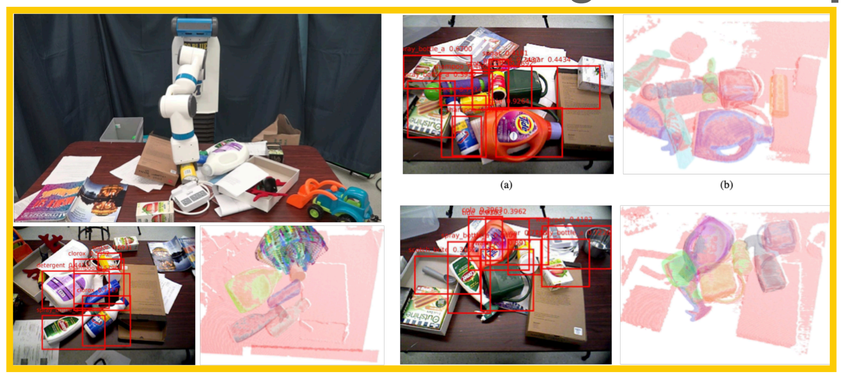
[2] Charles R. Qi, Li Yi, Hao Su, Leonidas J. Guibas. “PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space” NeurIPS, 2017.

[3] Danfei Xu, Dragomir Anguelov, Ashesh Jain. “PointFusion: Deep Sensor Fusion for 3D Bounding Box Estimation” CVPR, 2018.

[4] Chen Wang, Danfei Xu, Yuke Zhu, Roberto Martín-Martín, Cewu Lu, Li Fei-Fei, Silvio Savarese. “DenseFusion: 6D Object Pose Estimation by Iterative Dense Fusion” CVPR, 2019.

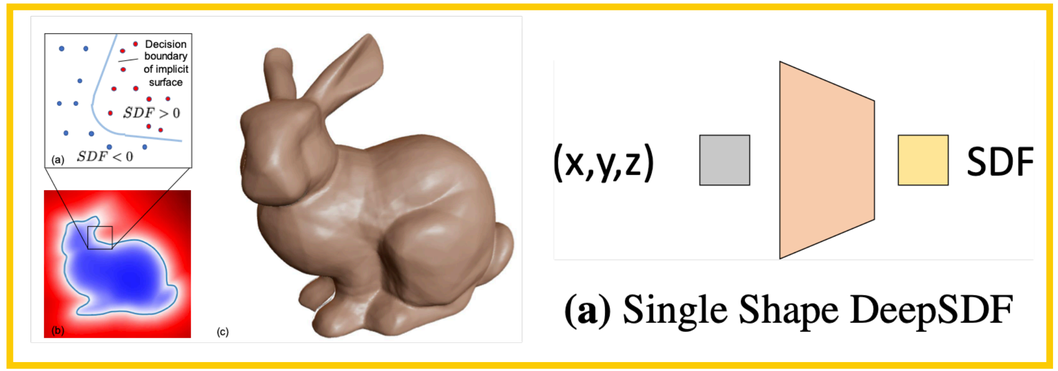
Object Pose, Geometry, SDF, Implicit Surfaces

Sequential Scene Understanding and Manipulation



[1] Sui et al., IROS 2017

DeepSDF



[2] Park et al., CVPR 2019

Implicit Surface Representations

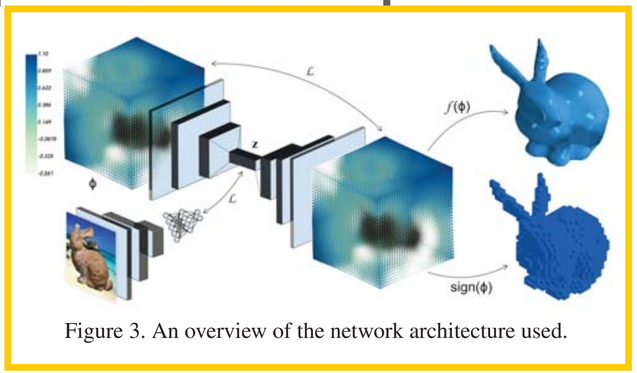


Figure 3. An overview of the network architecture used.

[3] Michalkiewicz et al., ICCV 2019

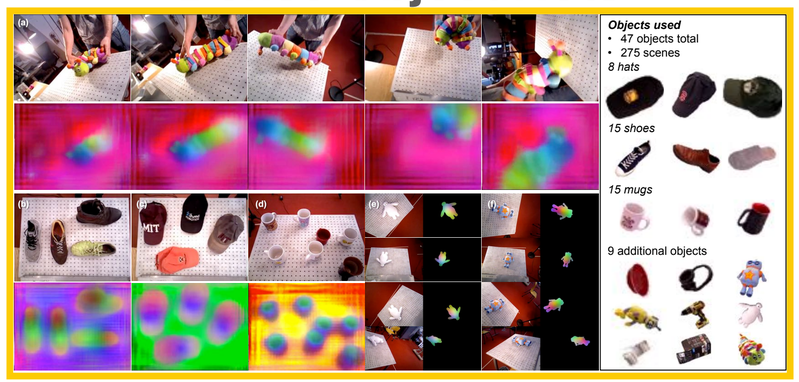
[1] Zhiqiang Sui, Zheming Zhou, Zhen Zeng, Odest Chadwicke Jenkins. “SUM: Sequential scene understanding and manipulation” IROS, 2017.

[2] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, Steven Lovegrove. “DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation” CVPR, 2019.

[3] Mateusz Michalkiewicz, Jhony Kaesemodel Pontes, Dominic Jack, Mahsa Baktashmotlagh, Anders Eriksson. “Implicit Surface Representations As Layers in Neural Networks” ICCV, 2019.

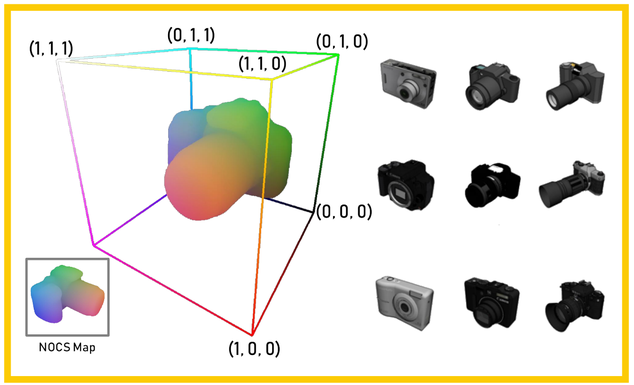
Dense Object Descriptors and Category-Level Representations

Dense Object Nets



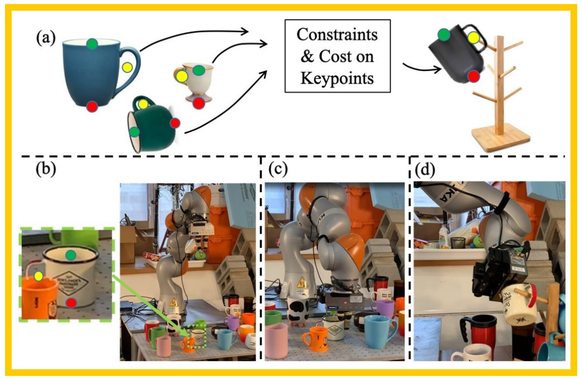
[1] Florence et al., CoRL 2018

NOCS



[2] Wang et al., CVPR 2018

kPAM



[3] Manuelli et al., ISRR 2019

Category-Level Pose from RGB

[4] Lin et al., ICRA 2022

[1] Peter R. Florence, Lucas Manuelli, Russ Tedrake. “Dense Object Nets: Learning Dense Visual Object Descriptors By and For Robotic Manipulation” CoRL, 2018.

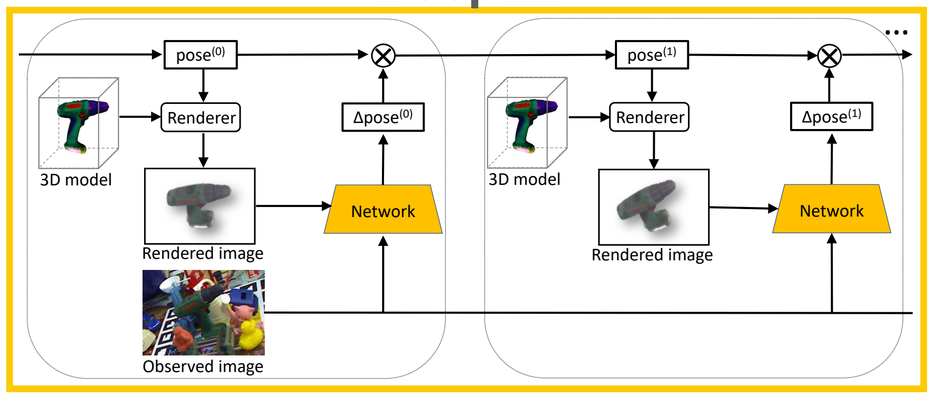
[2] He Wang, Srinath Sridhar, Jingwei Huang, Julien Valentin, Shuran Song, Leonidas J. Guibas. “Normalized Object Coordinate Space for Category-Level 6D Object Pose and Size Estimation” CVPR, 2019.

[3] Lucas Manuelli, Wei Gao, Peter Florence, Russ Tedrake. “kPAM: KeyPoint Affordances for Category-Level Robotic Manipulation” ISRR, 2019.

[4] Yunzhi Lin, Jonathan Tremblay, Stephen Tyree, Patricio A. Vela, Stan Birchfield. “Single-Stage Keypoint-Based Category-Level Object Pose Estimation from an RGB Image” ICRA, 2022.

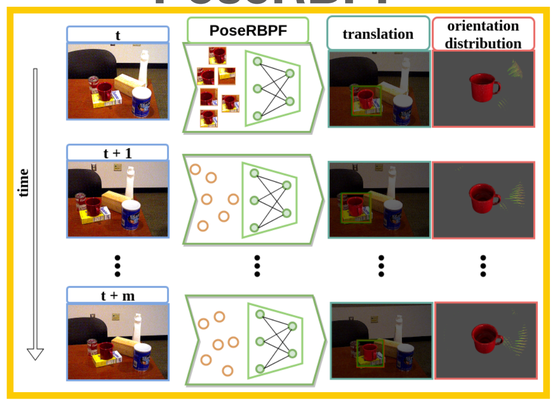
Recurrent Networks and Object Tracking

DeepIM



[1] Li et al., ECCV 2018

PoseRBPF

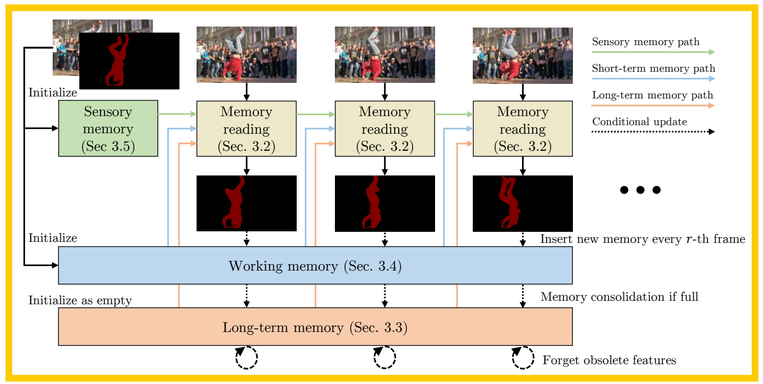


[2] Deng et al., RSS 2019

6-PACK

[3] Wang et al., ICRA 2020

XMem



[4] Cheng and Schwing, ECCV 2022

[1] Yi Li, Gu Wang, Xiangyang Ji, Yu Xiang, Dieter Fox. “DeepIM: Deep Iterative Matching for 6D Pose Estimation” ECCV, 2018.

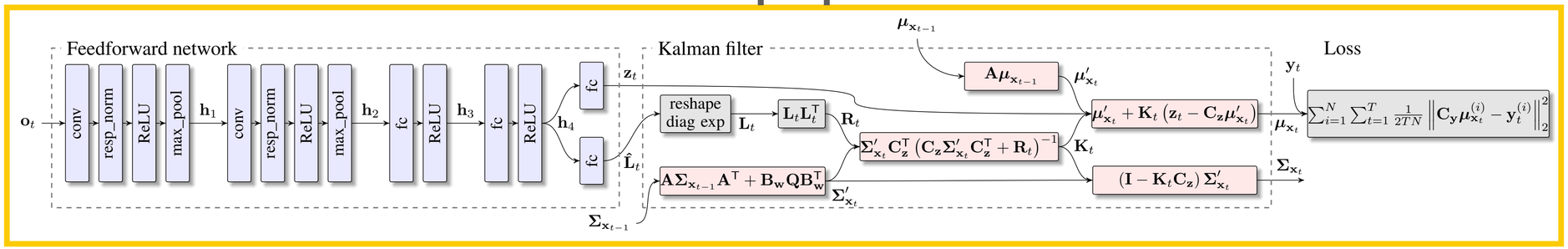
[2] Xinke Deng, Arsalan Mousavian, Yu Xiang, Fei Xia, Timothy Bretl, Dieter Fox. “PoseRBPF: A Rao-Blackwellized Particle Filter for 6D Object Pose Tracking” RSS, 2019.

[3] Chen Wang, Roberto Martín-Martín, Danfei Xu, Jun Lv, Cewu Lu, Li Fei-Fei, Silvio Savarese, Yuke Zhu. “6-PACK: Category-level 6D Pose Tracker with Anchor-Based Keypoints” ICRA, 2020.

[4] Ho Kei Cheng, Alexander G. Schwing. “XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model” ECCV, 2022.

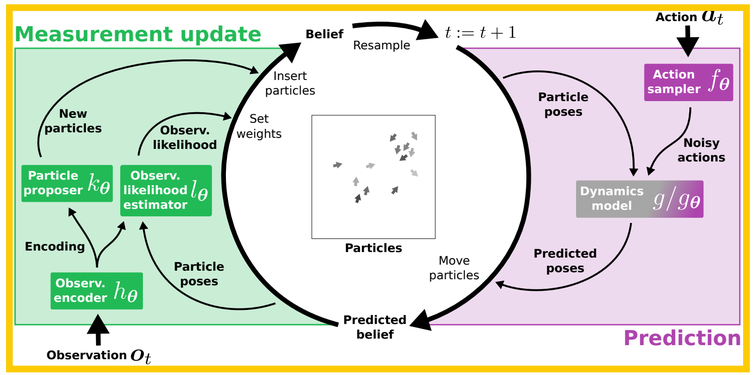
Visual Odometry and Localization

Backprop KF



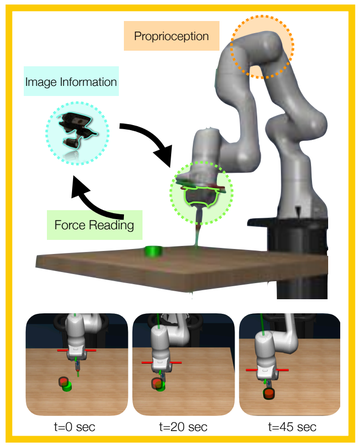
[1] Haarnoja et al., NeurIPS 2016

Differentiable Particle Filters



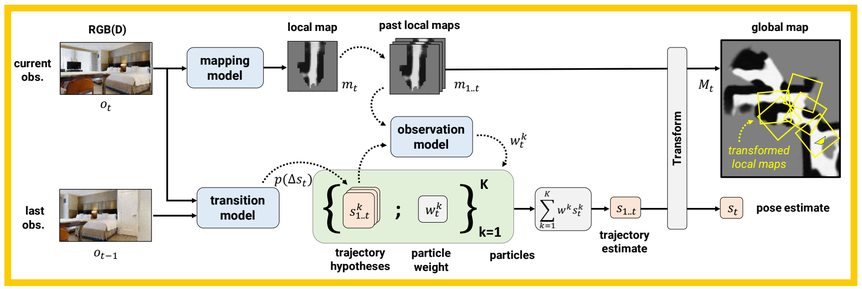
[2] Jonschkowski et al., RSS 2018

Multimodal Fusion



[3] Lee et al., IROS 2020

Differentiable SLAM-net

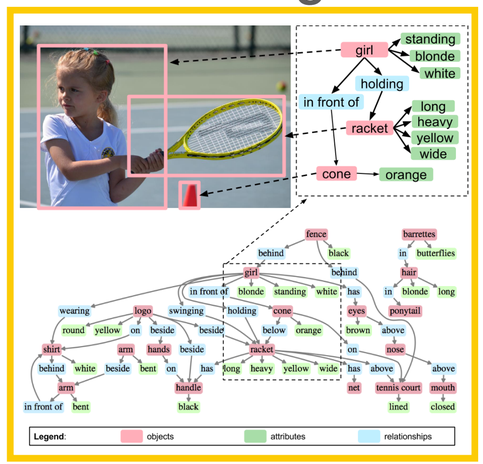


[4] Karkus et al., CVPR 2021

[1] Tuomas Haarnoja, Anurag Ajay, Sergey Levine, Pieter Abbeel. “Backprop KF: Learning Discriminative Deterministic State Estimators” NeurIPS, 2016.
[2] Rico Jonschkowski, Divyam Rastogi, Oliver Brock. “Differentiable Particle Filters: End-to-End Learning with Algorithmic Priors” RSS, 2018.
[3] Michelle A. Lee, Brent Yi, Roberto Martín-Martín, Silvio Savarese, Jeannette Bohg. “Multimodal Sensor Fusion with Differentiable Filters” IROS, 2020.
[4] Peter Karkus, Shaojun Cai, David Hsu. “Differentiable SLAM-net: Learning Particle SLAM for Visual Navigation” CVPR, 2021.

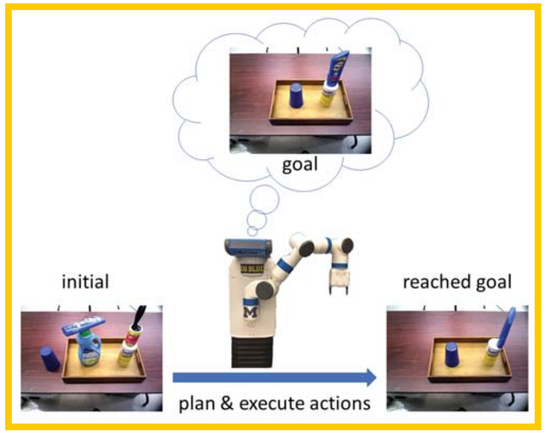
Semantic Scene Graphs and Explicit Representations

Image Retrieval using Scene Graphs



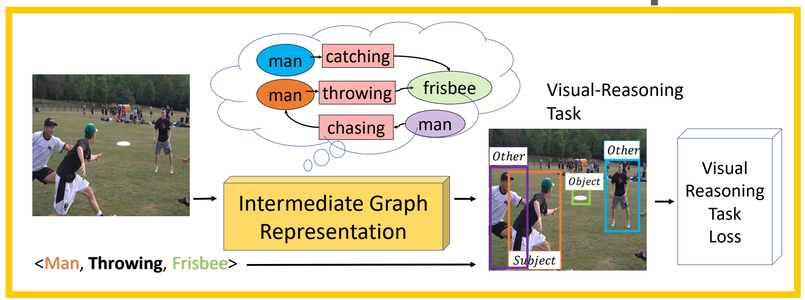
[1] Johnson et al., CVPR 2015

Semantic Robot Programming



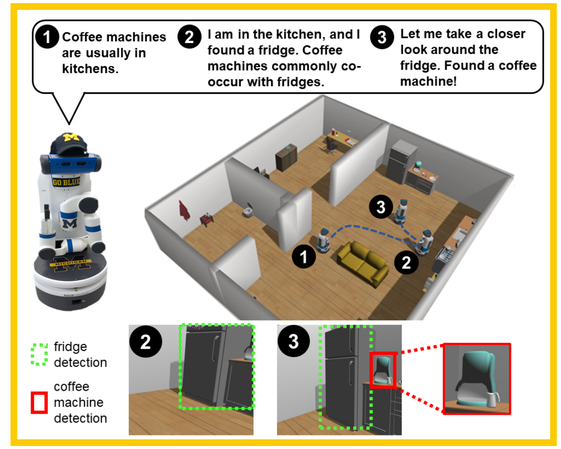
[2] Zeng et al., ICRA 2018

Differentiable Scene Graphs



[3] Raboh et al., WACV 2020

Semantic Linking Maps



[4] Zeng et al., ICRA 2020

[1] Justin Johnson, Ranjay Krishna, Michael Stark, Li-Jia Li, David A. Shamma, Michael S. Bernstein, Li Fei-Fei. “Image Retrieval using Scene Graphs” CVPR, 2015.

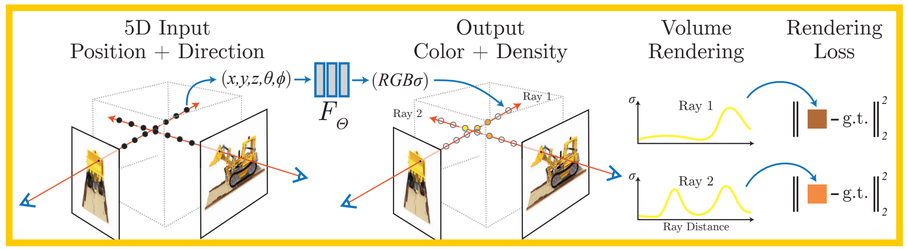
[2] Zhen Zeng, Zheming Zhou, Zhiqiang Sui, Odest Chadwicke Jenkins. “Semantic Robot Programming for Goal-Directed Manipulation in Cluttered Scenes” ICRA, 2018.

[3] Moshiko Raboh, Roei Herzig, Gal Chechik, Jonathan Berant, Amir Globerson. “Differentiable Scene Graphs” WACV, 2020.

[4] Zhen Zeng, Adrian Röfer, Odest Chadwicke Jenkins. “Semantic Linking Maps for Active Visual Object Search” ICRA, 2020.

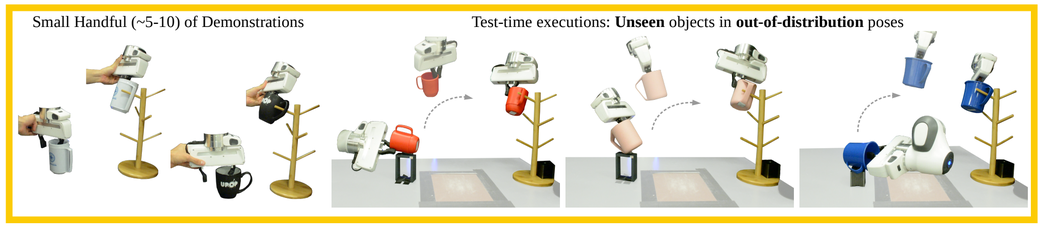
Neural Radiance Fields and Implicit Representations

NeRF



[1] Mildenhall et al., ECCV 2020

Neural Descriptor Fields

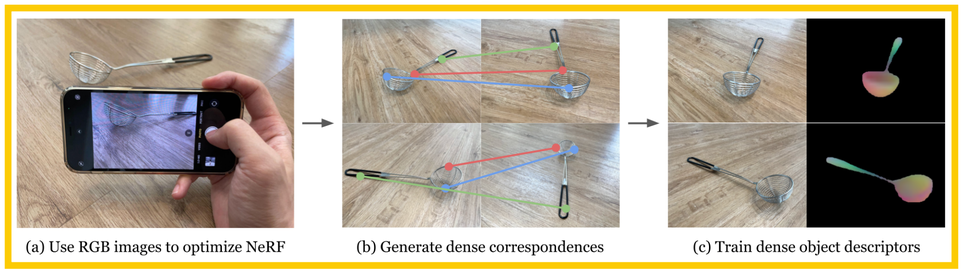


[3] Simeonov et al., ICRA 2022

Object-Centric NeRFs

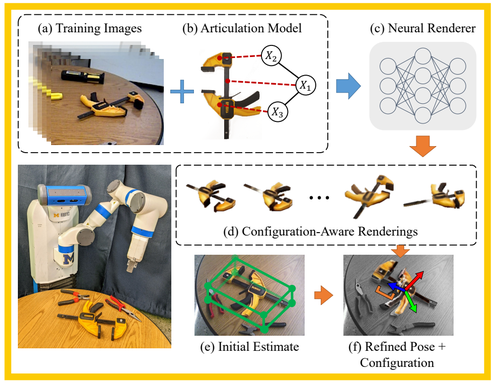
[2] Guo et al., 2020

NeRF-Supervision



[4] Yen-Chen et al., ICRA 2022

NARF22



[5] Lewis et al., IROS 2022

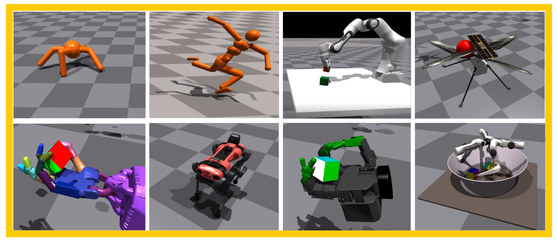
[1] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, Ren Ng. “NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis” ECCV, 2020.
[2] Michelle Guo, Alireza Fathi, Jiajun Wu, Thomas Funkhouser. “Object-Centric Neural Scene Rendering” 2020.
[3] Anthony Simeonov, Yilun Du, Andrea Tagliasacchi, Joshua B. Tenenbaum, Alberto Rodriguez, Pulkit Agrawal, Vincent Sitzmann. “Neural Descriptor Fields: SE(3)-Equivariant Object Representations for Manipulation” ICRA, 2022.
[4] Lin Yen-Chen, Pete Florence, Jonathan T. Barron, Tsung-Yi Lin, Alberto Rodriguez, Phillip Isola. “NeRF-Supervision: Learning Dense Object Descriptors from Neural Radiance Fields” ICRA, 2022.
[5] Stanley Lewis, Jana Pavlasek, Odest Chadwicke Jenkins. “NARF22: Neural Articulated Radiance Fields for Configuration-Aware Rendering” IROS, 2022.

Datasets

Large-Scale Data Collection with Robots

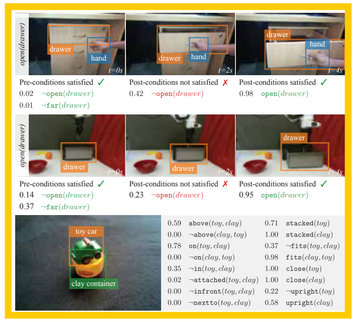
[1] Levine et al., IJRR 2018

Simulating Robot Datasets



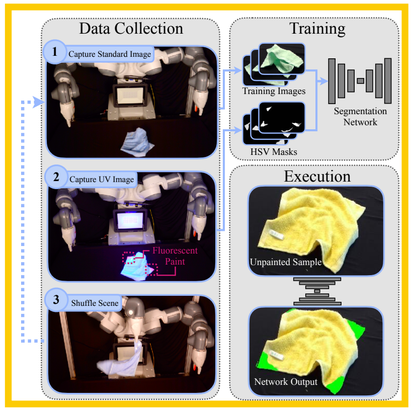
[2] Makoviychuk et al., 2021

Grounding Predicates through Actions



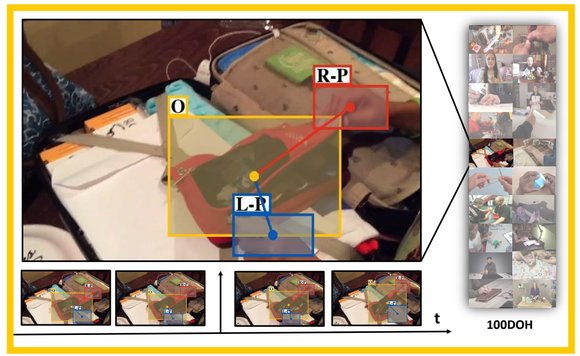
[3] Migimatsu and Bohg, ICRA 2021

All You need is LUV



[4] Thananjeyan et al., IROS 2022

Internet Scale



[5] Shan et al., CVPR 2020

[1] Sergey Levine, Peter Pastor, Alex Krizhevsky, Deirdre Quillen. “Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection” IJRR, 2018.

[2] Viktor Makoviychuk, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, Nikita Rudin, Arthur Allshire, Ankur Handa, Gavriel State. “Isaac Gym: High Performance GPU-Based Physics Simulation For Robot Learning” 2021.

[3] Toki Migimatsu, Jeannette Bohg. “Grounding Predicates through Actions” ICRA, 2021.

[4] Brijen Thananjeyan, Justin Kerr, Huang Huang, Joseph E. Gonzalez, Ken Goldberg. “All You Need is LUV: Unsupervised Collection of Labeled Images using Invisible UV Fluorescent Indicators” IROS, 2022.

[5] Dandan Shan, Jiaqi Geng, Michelle Shu, David F. Fouhey. “Understanding Human Hands in Contact at Internet Scale” CVPR, 2020.

Even More!

- Grasp Pose Detection
- Tactile Perception for Grasping and Manipulation
- Transformer Architectures
- Deformable Object Representations
- Interpretable Models

RGB-D Architectures

Core List

1. PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes, Xiang et al., 2018
2. A Unified Framework for Multi-View Multi-Class Object Pose Estimation, Li et al., 2018
3. PVN3D: A Deep Point-Wise 3D Keypoints Voting Network for 6DoF Pose Estimation, He et al., 2020
4. Learning RGB-D Feature Embeddings for Unseen Object Instance Segmentation, Li et al., 2021

Object Pose, Geometry, SDF, Implicit surfaces

Core List

1. SUM: Sequential scene understanding and manipulation, Sui et al., 2017
2. DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation, Park et al., 2019
3. Implicit surface representations as layers in neural networks, Michalkiewicz et al., 2019

DeepRob

Discussion 4
Overview of Final Project Topics I
University of Michigan and University of Minnesota

Tactile Perception for Grasping and Manipulation

Core List

1. More Than a Feeling: Learning to Grasp and Regrasp using Vision and Touch, Calandra et al., 2018
2. Tactile Object Pose Estimation from the First Touch with Geometric Contact Rendering, Bauza et al., 2020
3. Visuotactile Affordances for Cloth Manipulation with Local Control, Sunil et al., 2022
4. ShapeMap 3-D: Efficient shape mapping through dense touch and vision, Suresh et al., 2022

Pre-tra… Archit…

Core List

1. SORNet: Spatial Object-Centric Representations for Sequential Manipulation, Yuan et al., 2021
2. CLIPort: What and Where Pathways for Robotic Manipulation, Shridhar et al., 2021
3. Masked Visual Pre-training for Motor Control, Xiao et al., 2022
4. R3M: A Universal Visual Representation for Robot Manipulation, Nair et al., 2022
5. Do As I Can, Not As I Say: Grounding Language in Robotic Affordances, Ahn et al., 2022
6. RT-1: Robotics Transformer for Real-World Control at Scale, Brohan et al., 2022

Neural Radiance Fields and Implicit Representations

Core List

1. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall et al., 2020
2. Object-Centric Neural Scene Rendering, Guo et al., 2020
3. Neural Descriptor Fields: SE(3)-Equivariant Object Representations for Manipulation, Simeonov et al., 2021
4. NeRF-Supervision: Learning Dense Object Descriptors from Neural Radiance Fields, Yen-Chen et al., 2022
5. NARF22: Neural Articulated Radiance Fields for Configuration-Aware Rendering, Lewis et al., 2022