

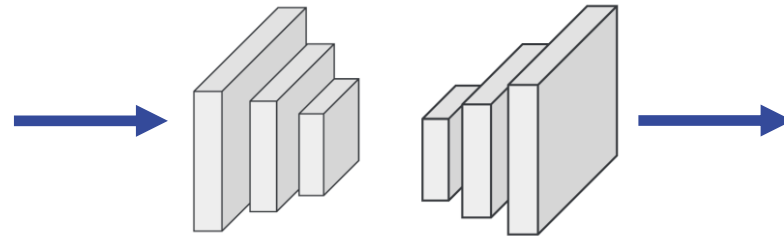
# SkyEye: Self-Supervised Bird's-Eye-View Semantic Mapping Using Monocular Frontal View Images

Nikhil Gosala\*, Kürsat Petek\*, Paulo L. J. Drews-Jr, Wolfram Burgard, Abhinav Valada

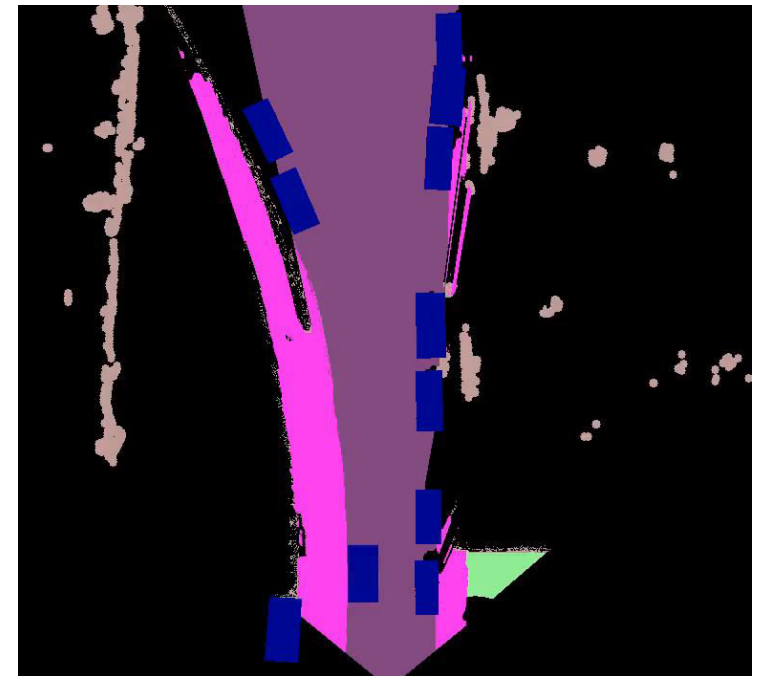
WED-PM-243  
Submission 4355

# Semantic Bird's-Eye-View (BEV) Mapping

Frontal View (FV) Image

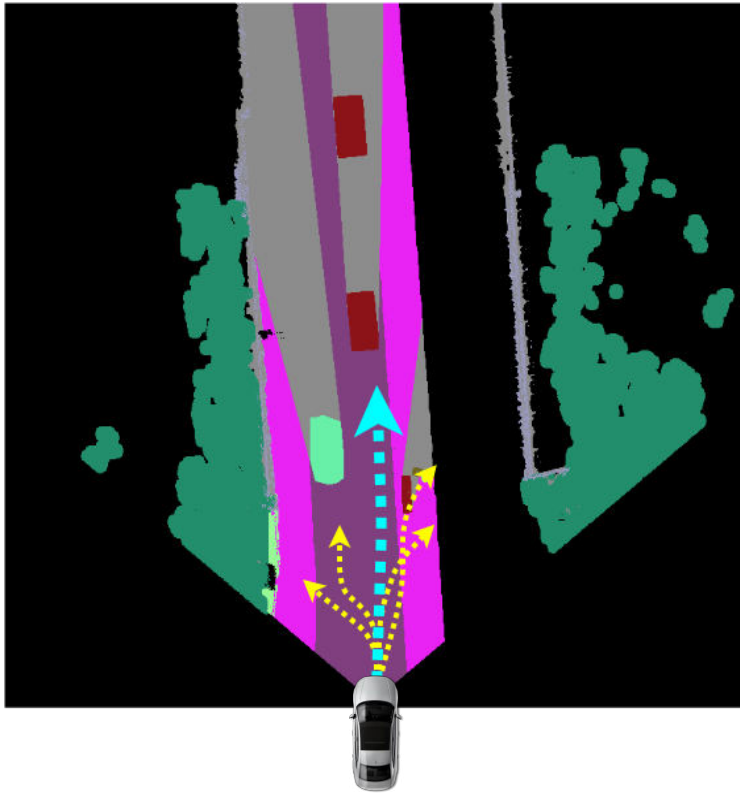


Semantic BEV Map



# Why Bird's-Eye-View Maps?

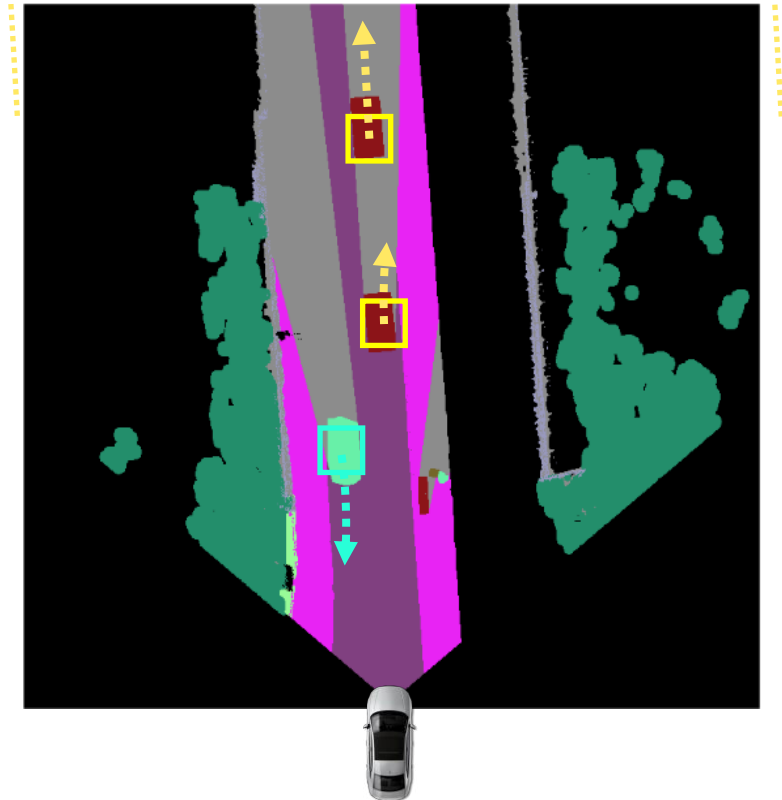
Path Planning



Object Tracking




Trajectory Prediction



# Semantic BEV Mapping: Data Requirements

## Fully-Supervised Approaches

Annotated BEV Maps



Expensive


Simulated BEV Maps



Sim-to-real domain gap

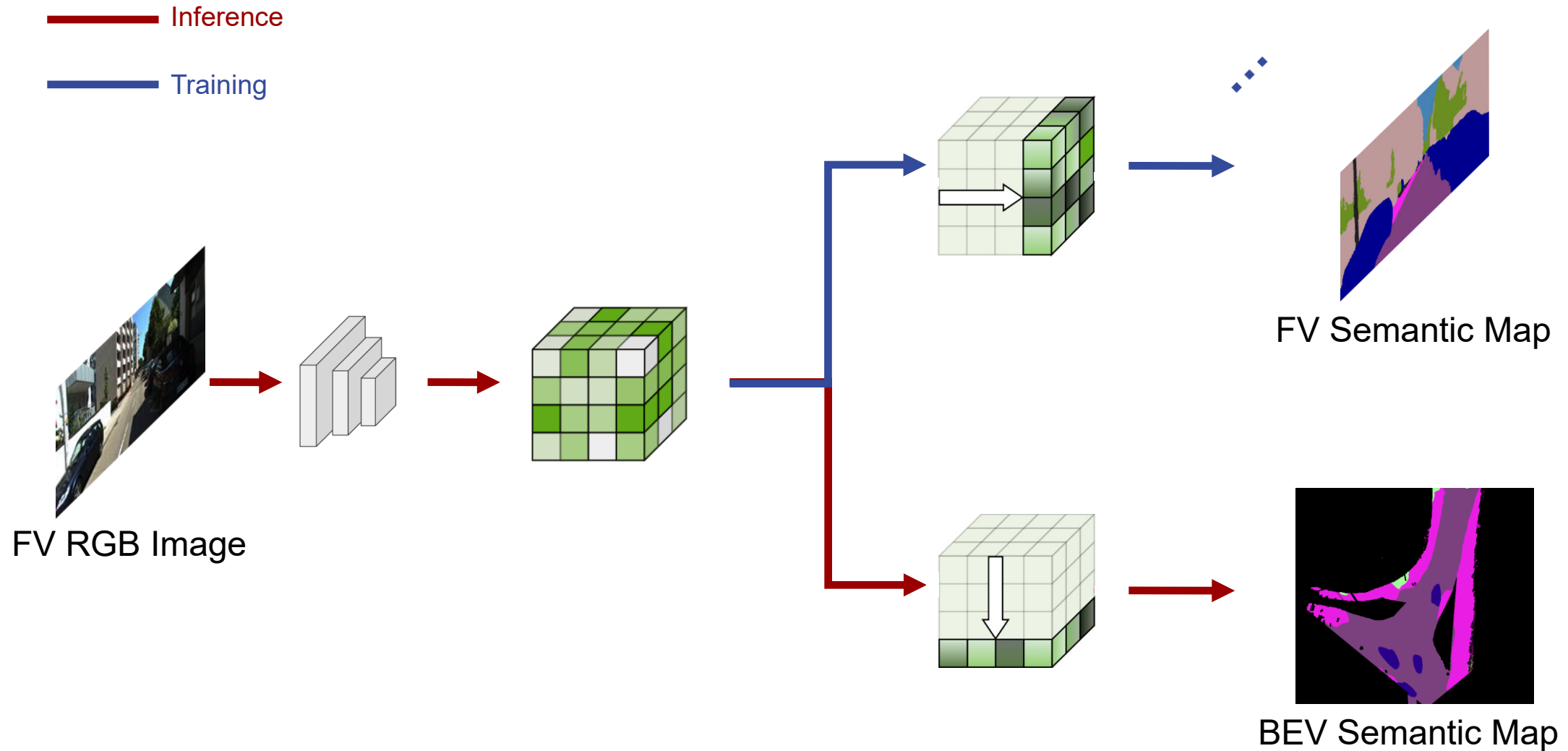
SkyEye

FV RGB Video + FV Semantic Video



All data in FV

# SkyEye: Self-Supervised Semantic BEV Mapping

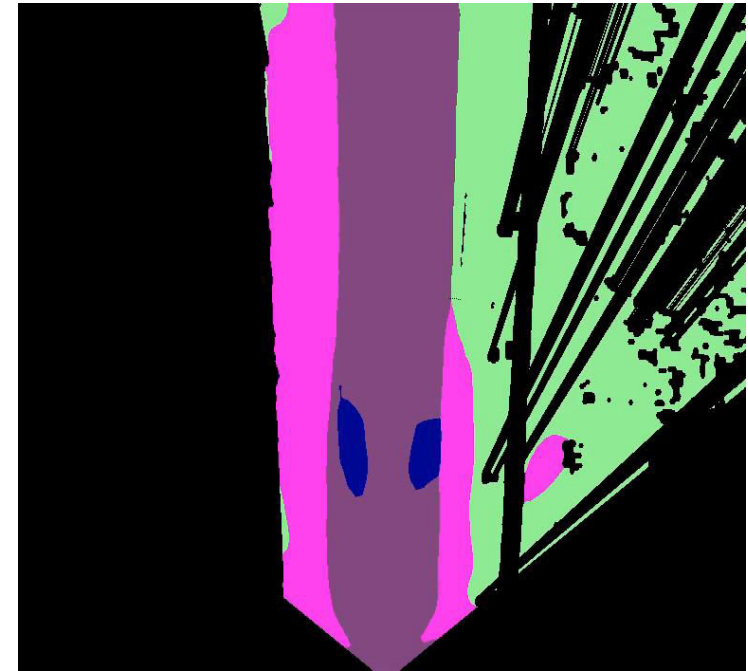


# Prediction of *SkyEye* on KITTI-360

Input Image in FV

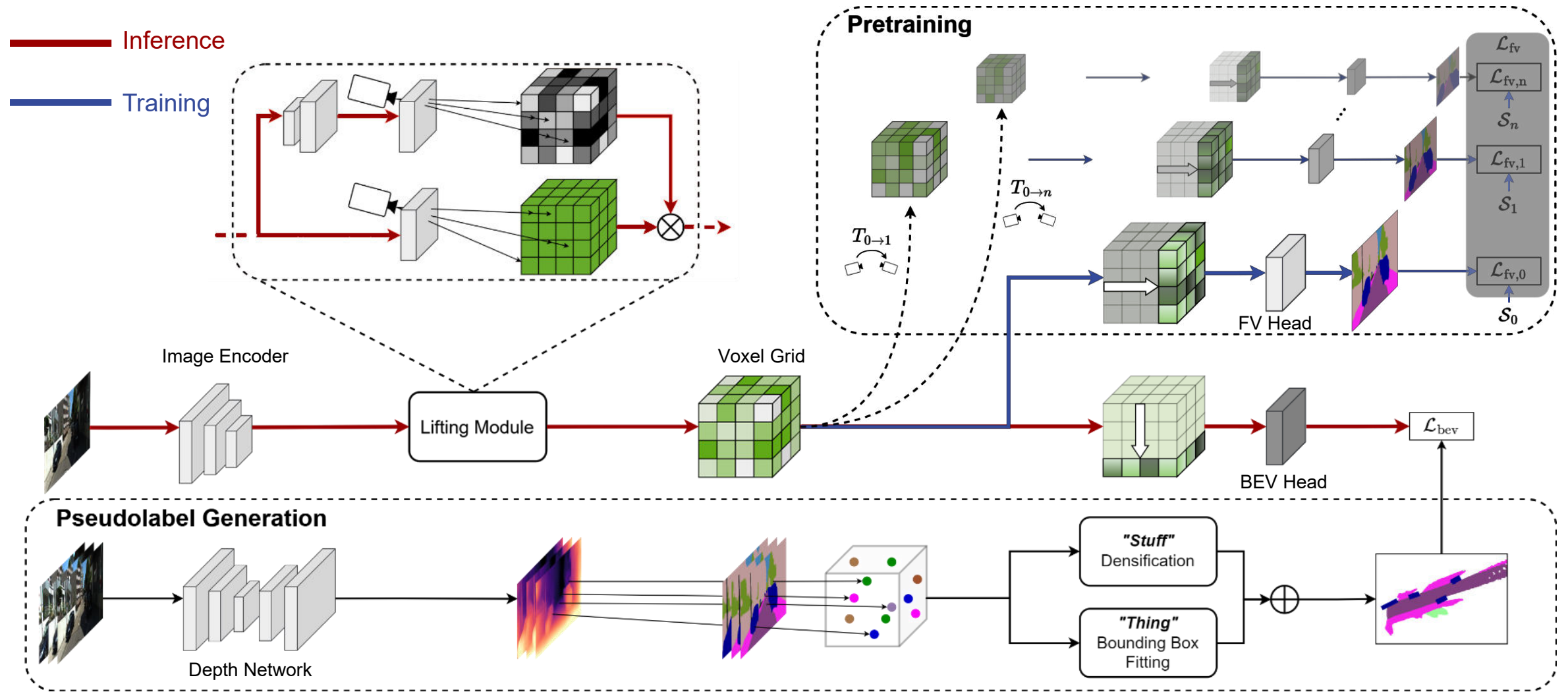


BEV Segmentation Prediction



*SkyEye* is on par with the current state-of-the-art without seeing any ground truth data in BEV

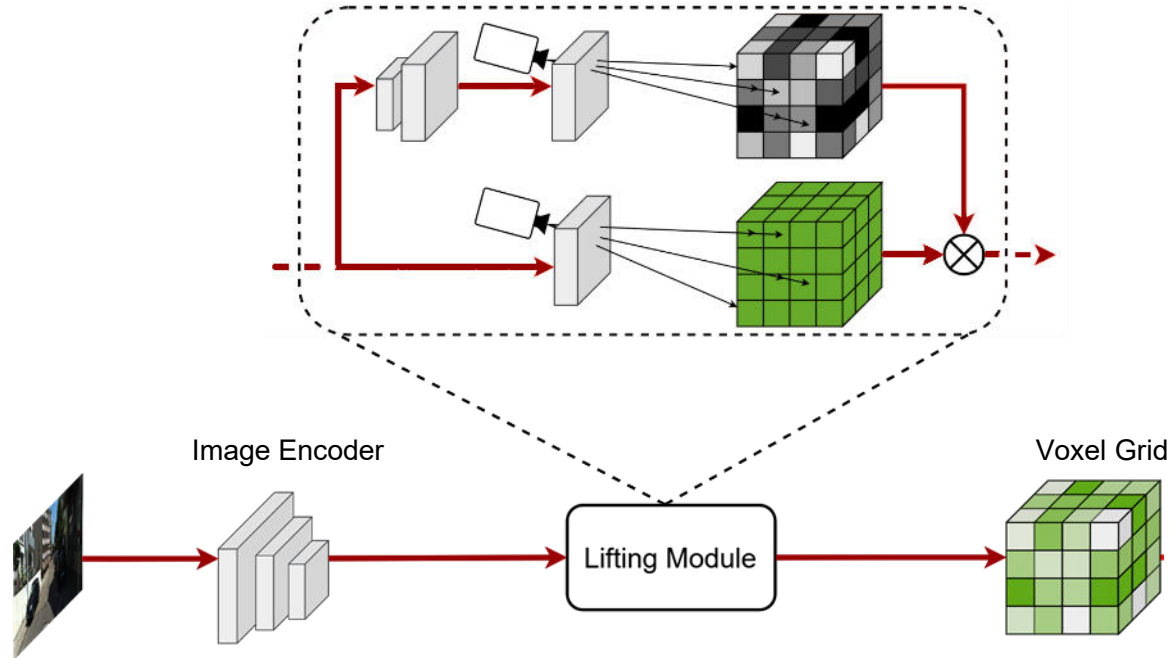
# Approach – SkyEye





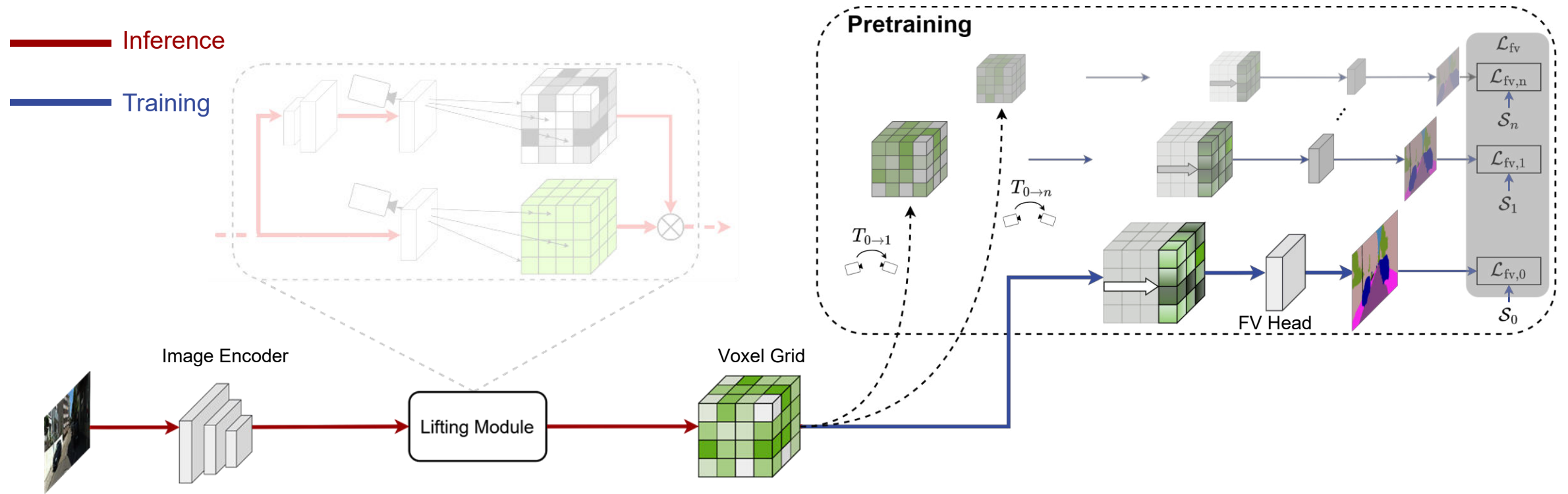
# Approach – Voxel Grid as Joint Representation

— Inference





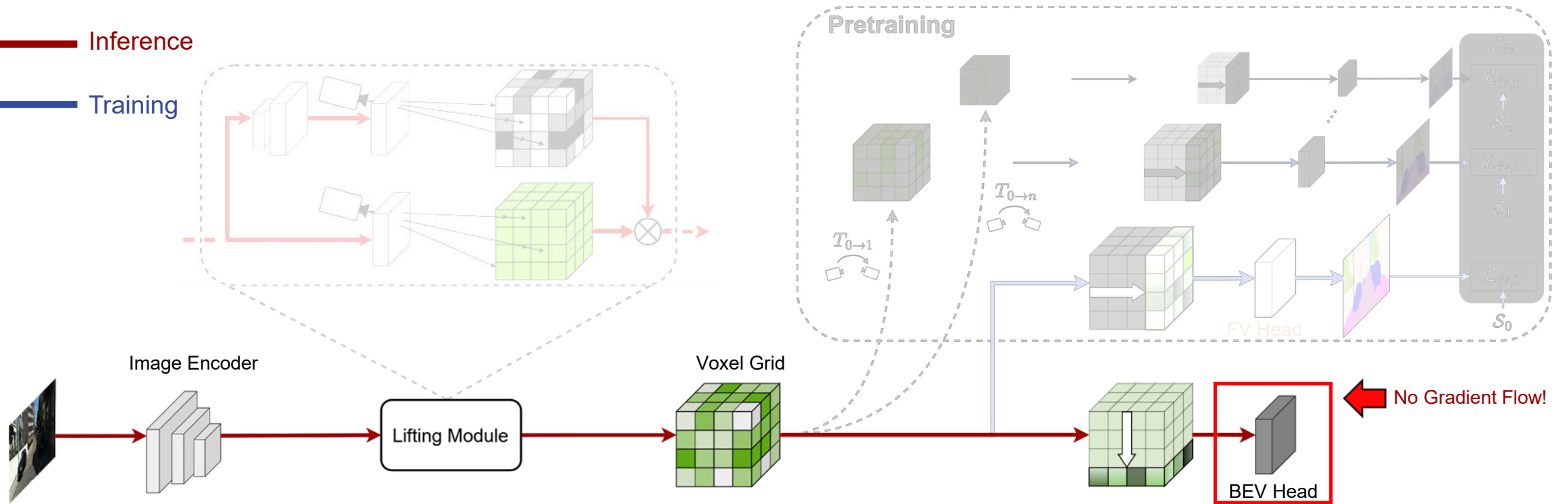
# Approach – Pretraining Using FV Semantic Annotations



# Approach – Finetuning Using Pseudolabels

**Inference**

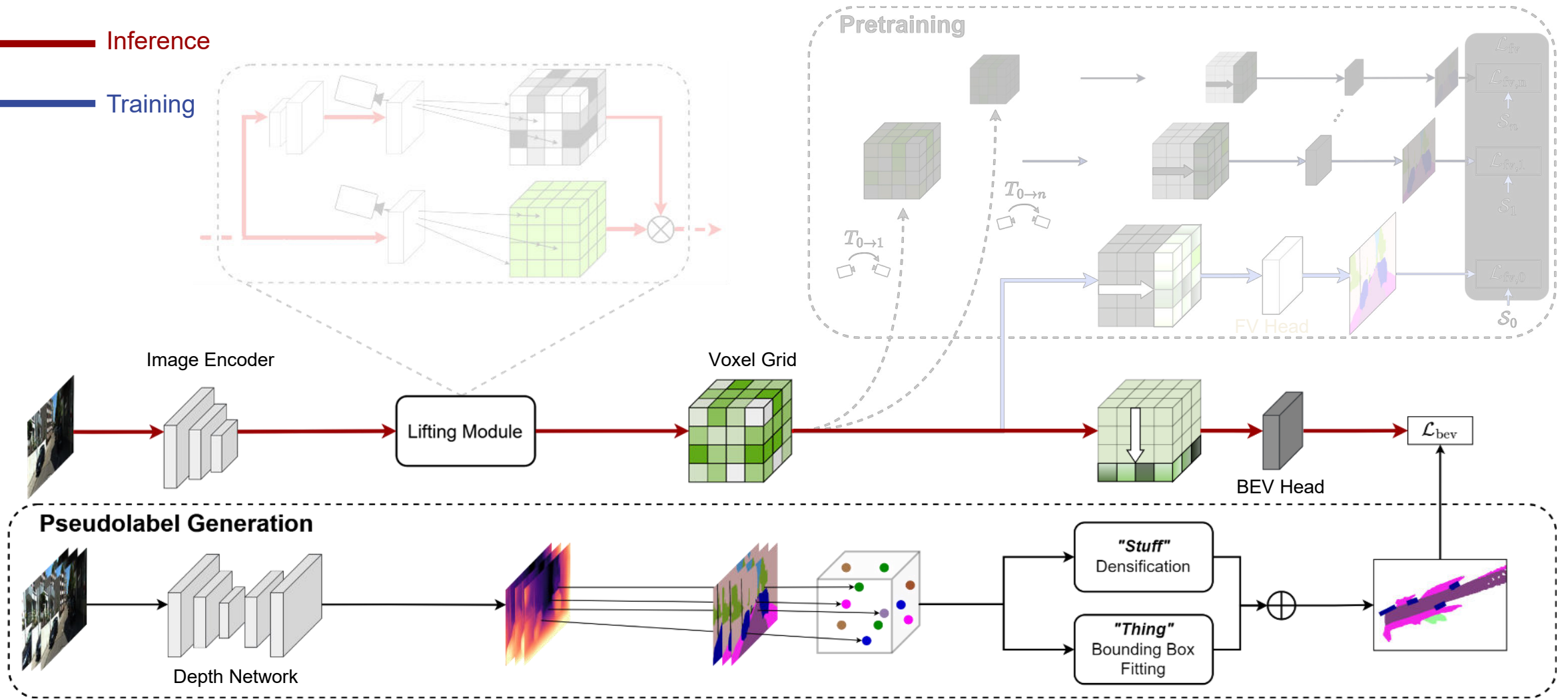
**Training**



# Approach – Finetuning Using Pseudolabels

**Inference**

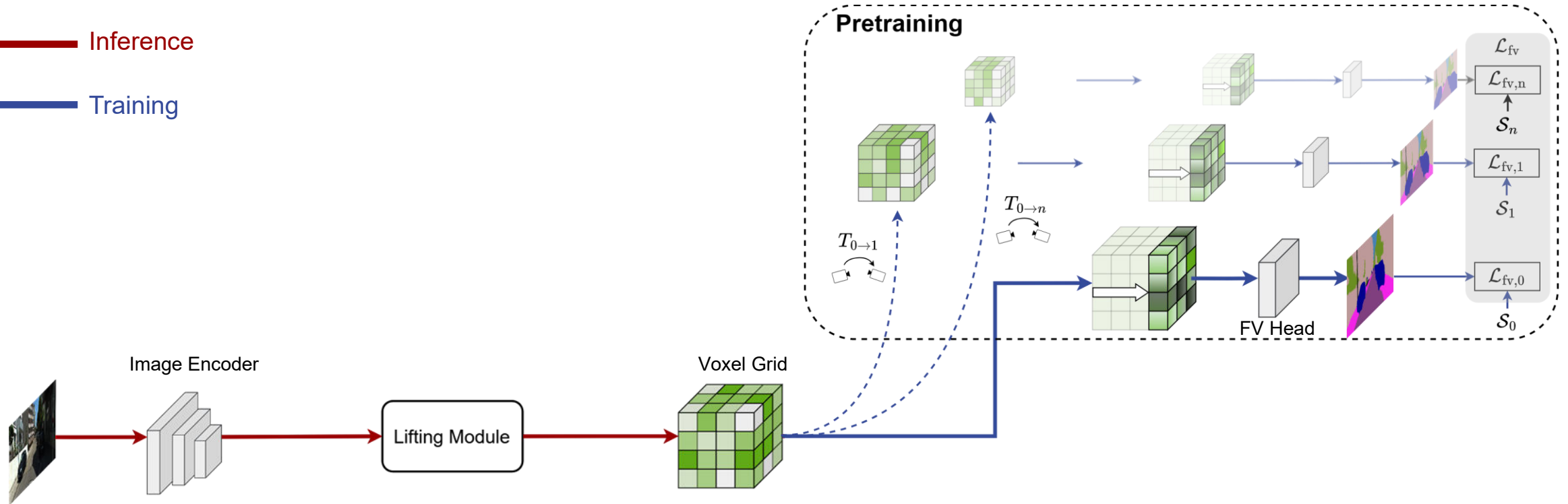
**Training**



# Pretraining: Implicit Supervision

— Inference

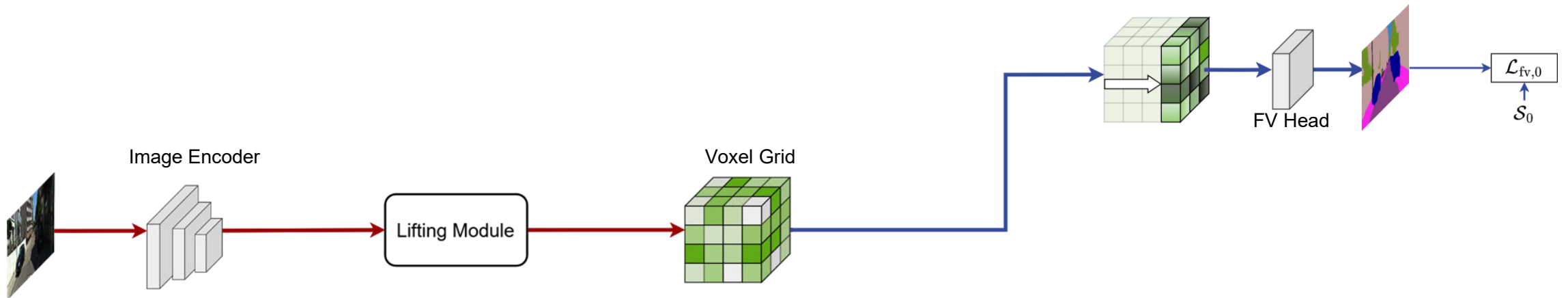
— Training



# Pretraining: Implicit Supervision

— Inference

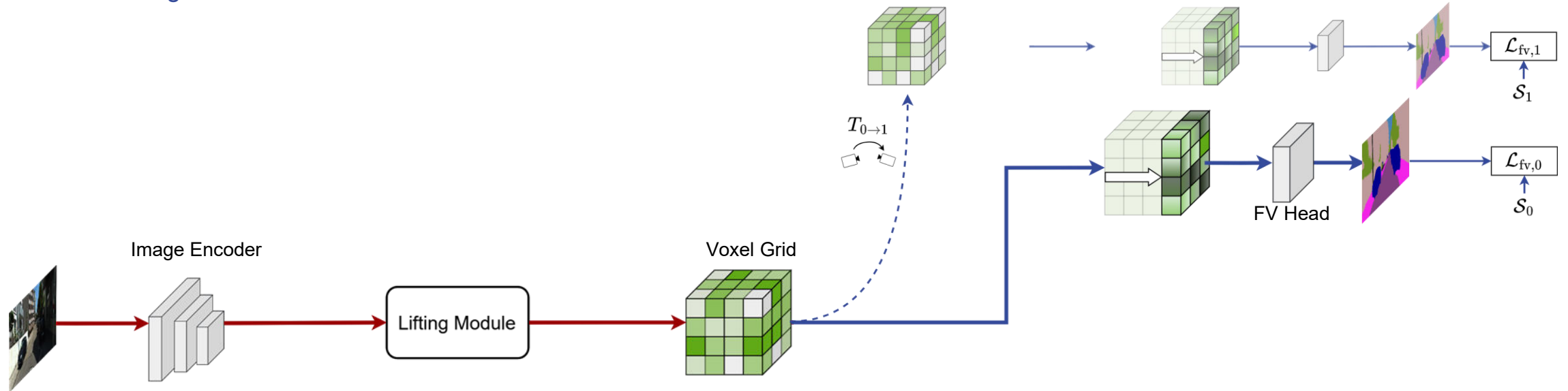
— Training



# Pretraining: Implicit Supervision

— Inference

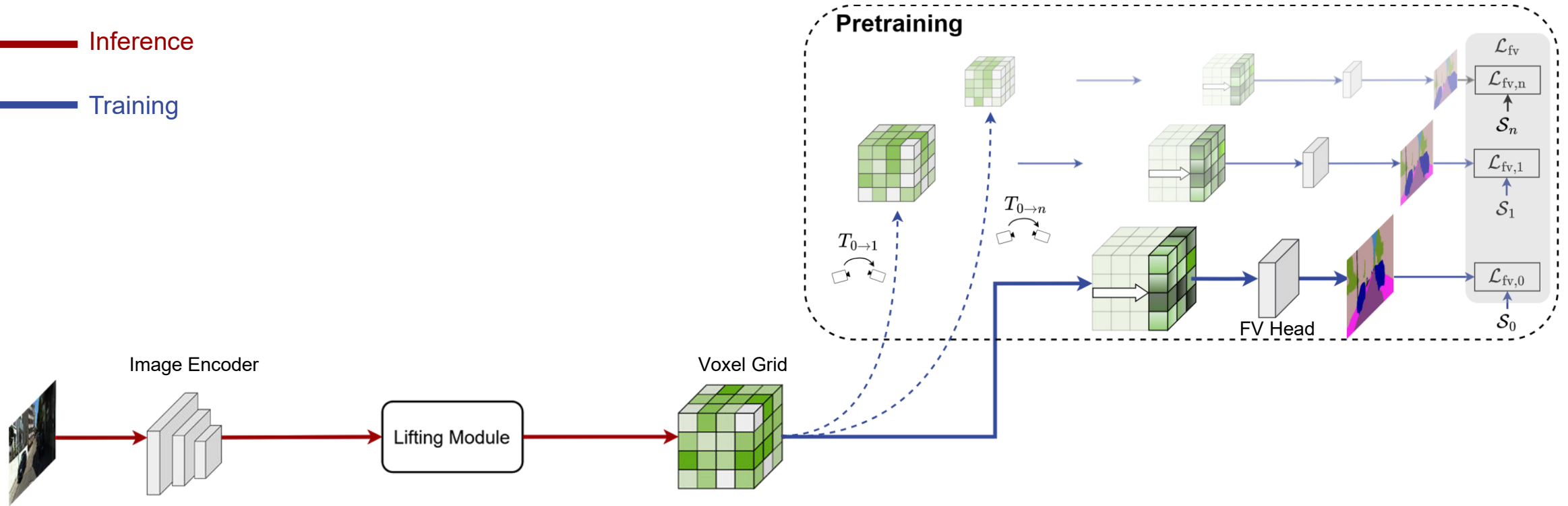
— Training



# Pretraining: Implicit Supervision

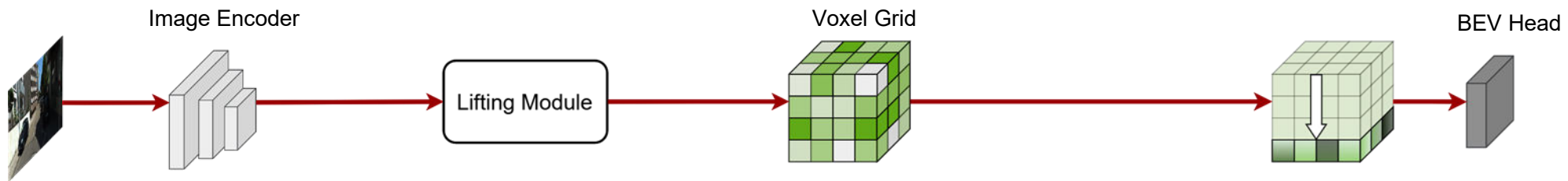
— Inference

— Training

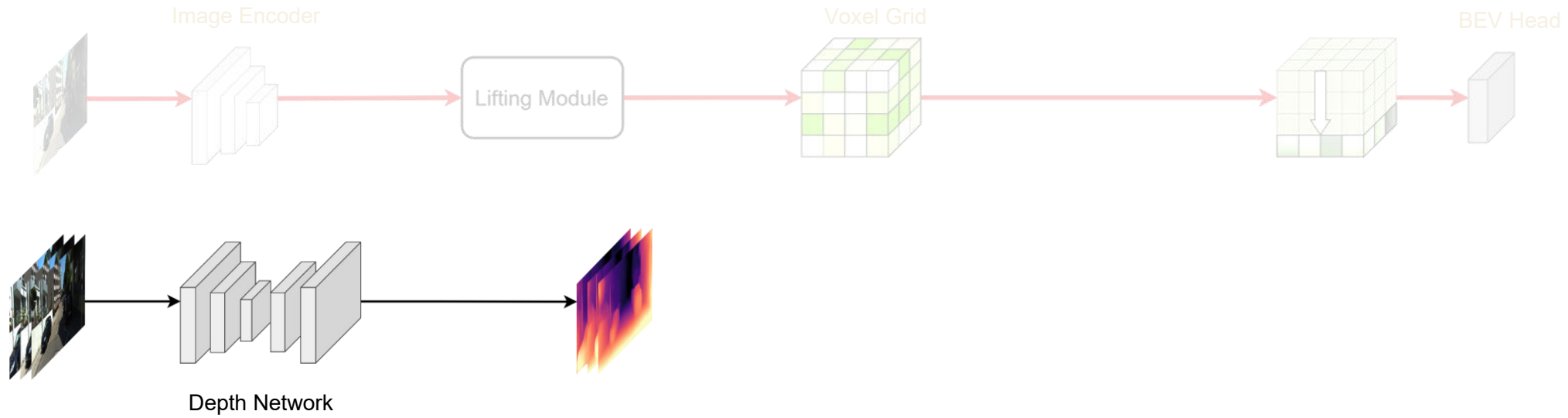




# Finetuning: Explicit Supervision

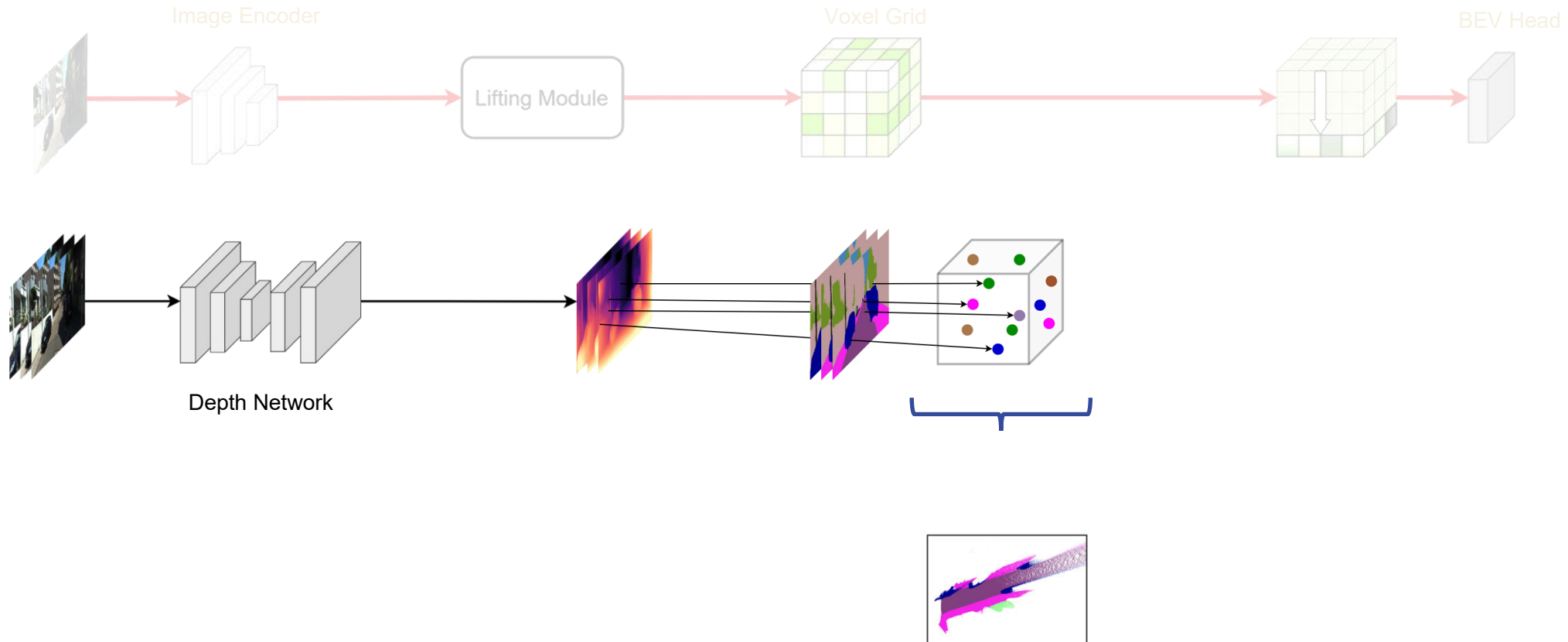


# Finetuning: Explicit Supervision

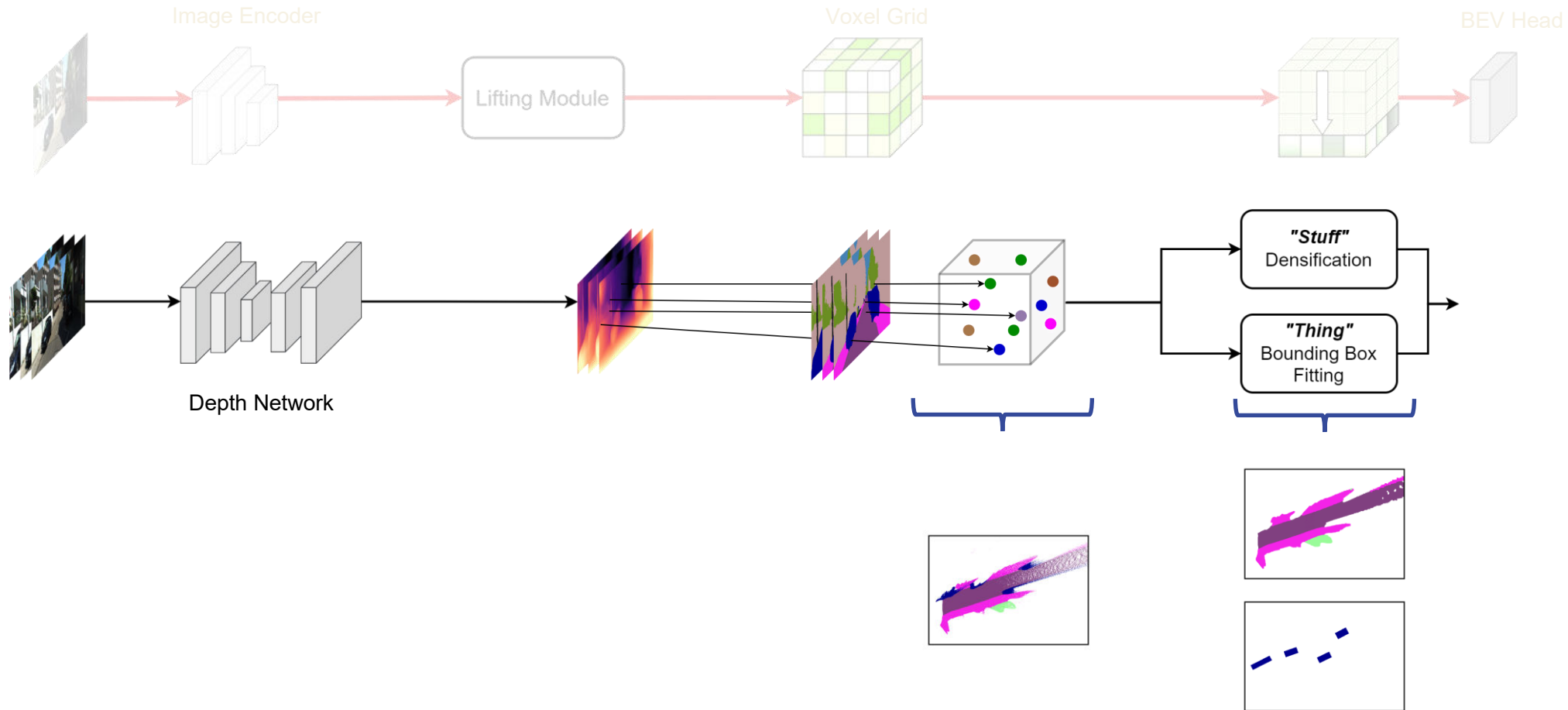




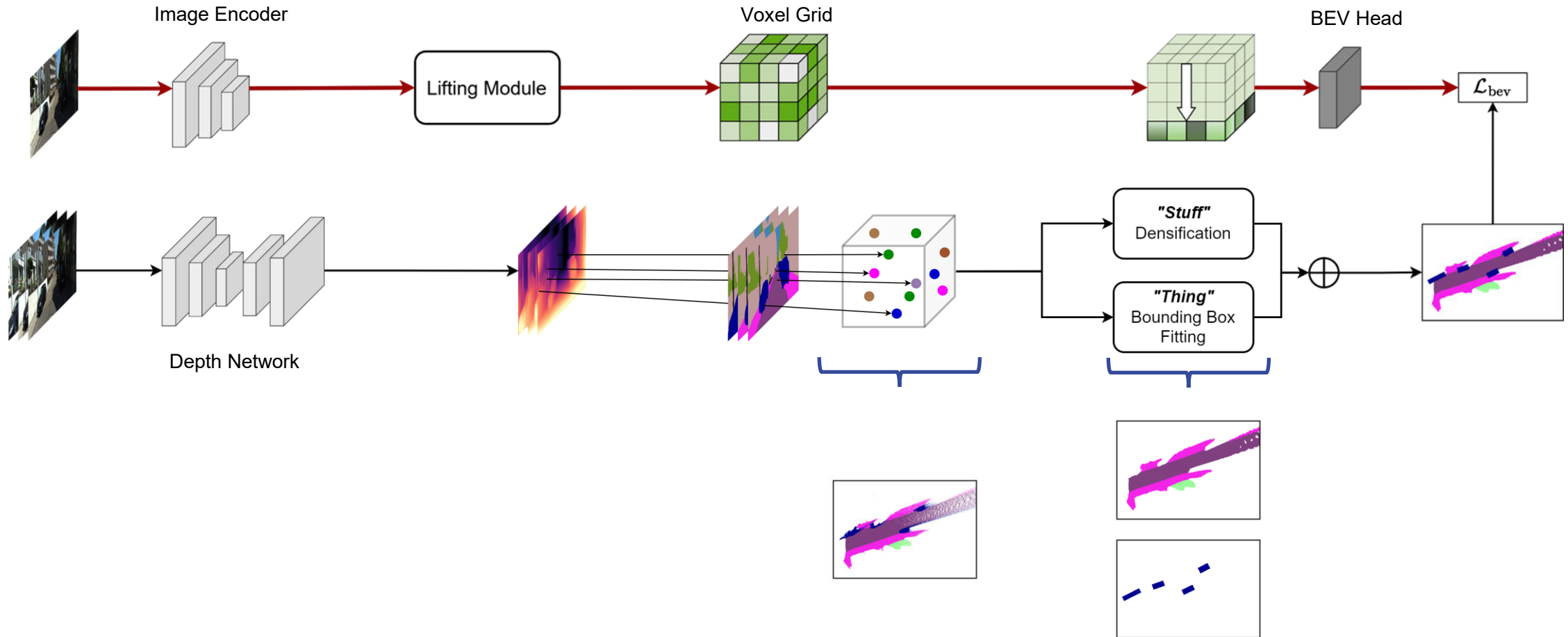
# Finetuning: Explicit Supervision



# Finetuning: Explicit Supervision



# Finetuning: Explicit Supervision



# Evaluation

## Datasets

- KITTI-360-BEV



- Waymo-BEV

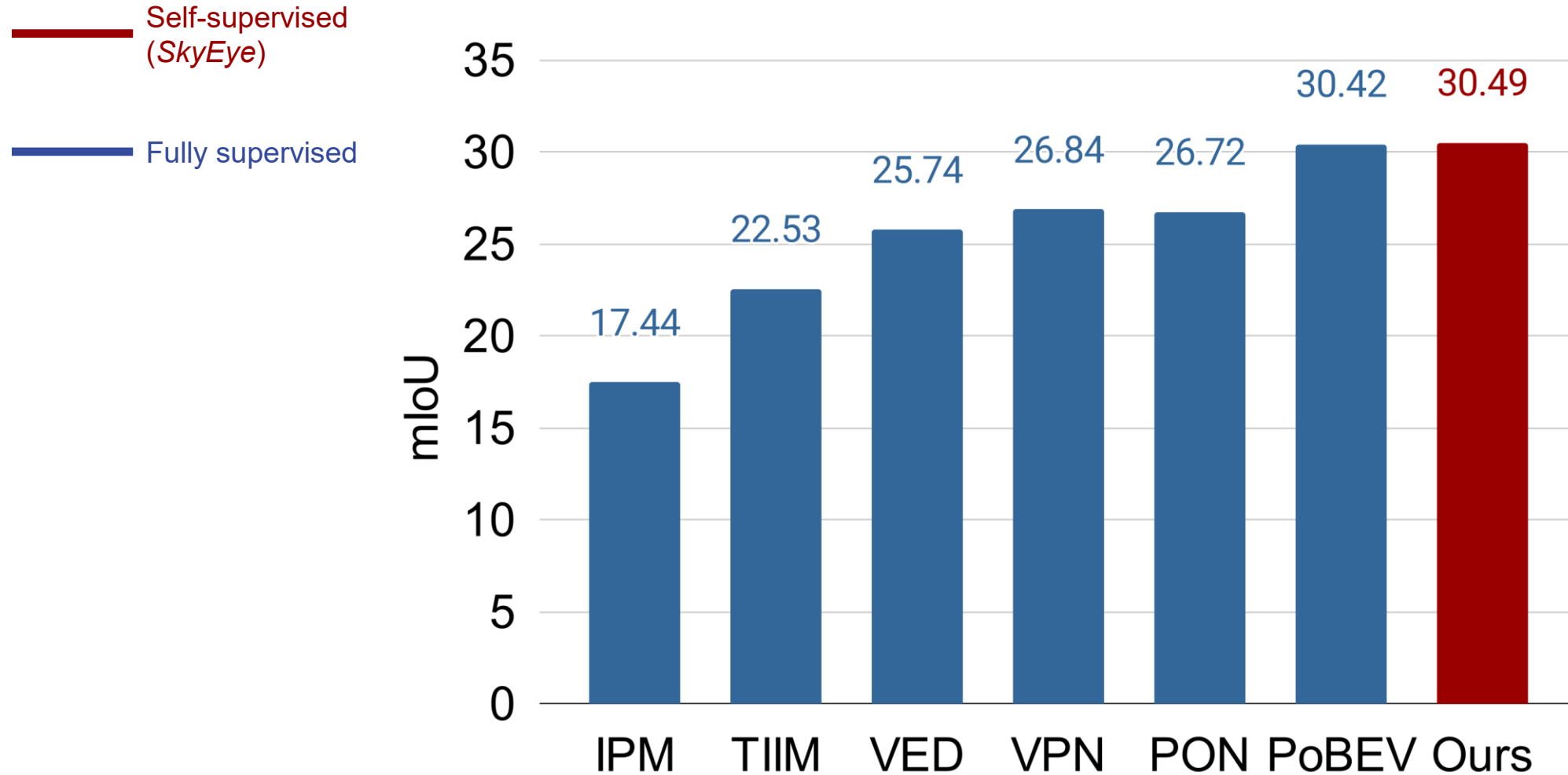


## Experiments

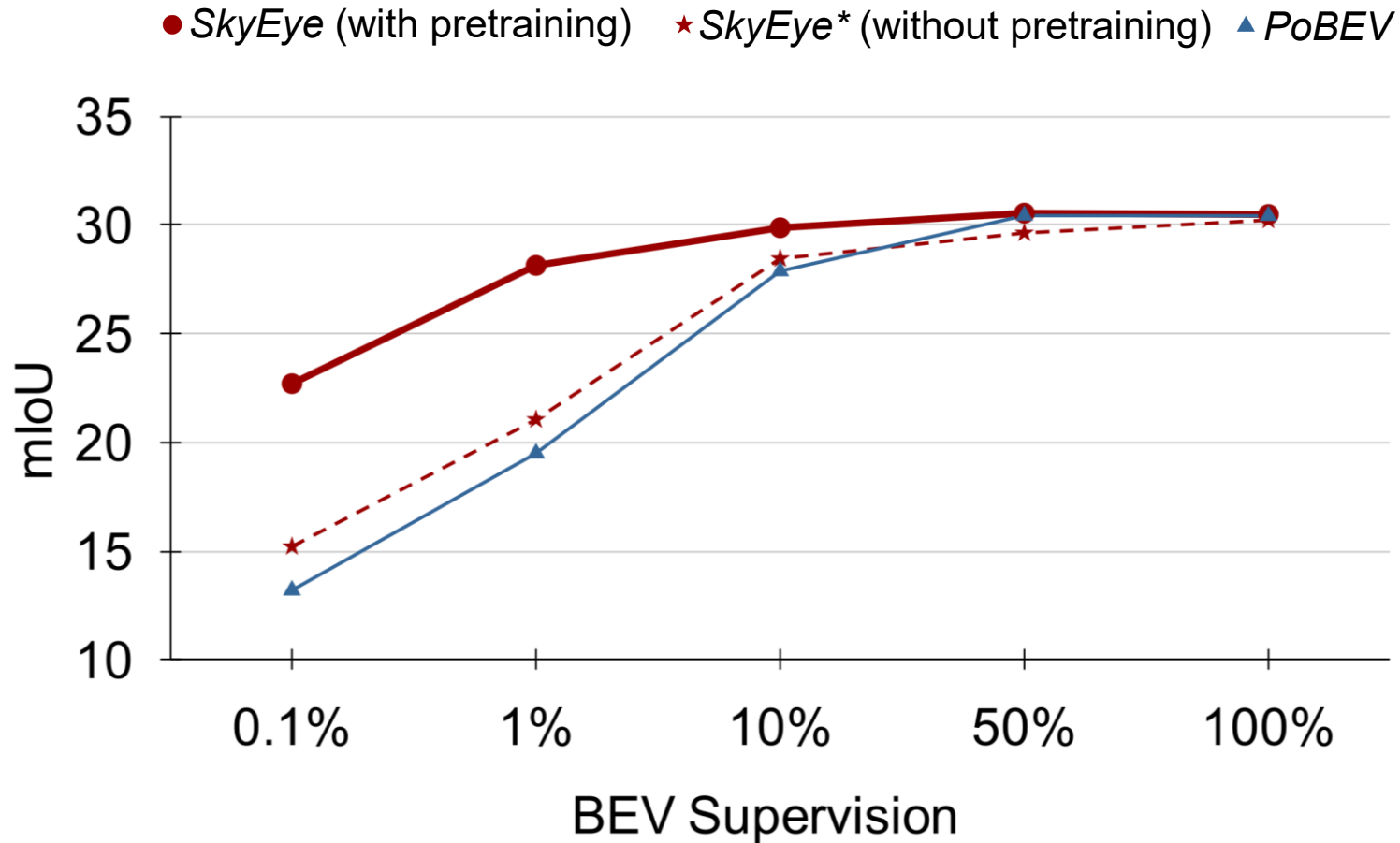
1. Overall Performance of SkyEye (KITTI-360-BEV)
2. Impact of Model Pretraining (KITTI-360-BEV)
3. Generalizability of SkyEye (KITTI-360-BEV → Waymo-BEV)



# Overall Performance of SkyEye: KITTI-360-BEV

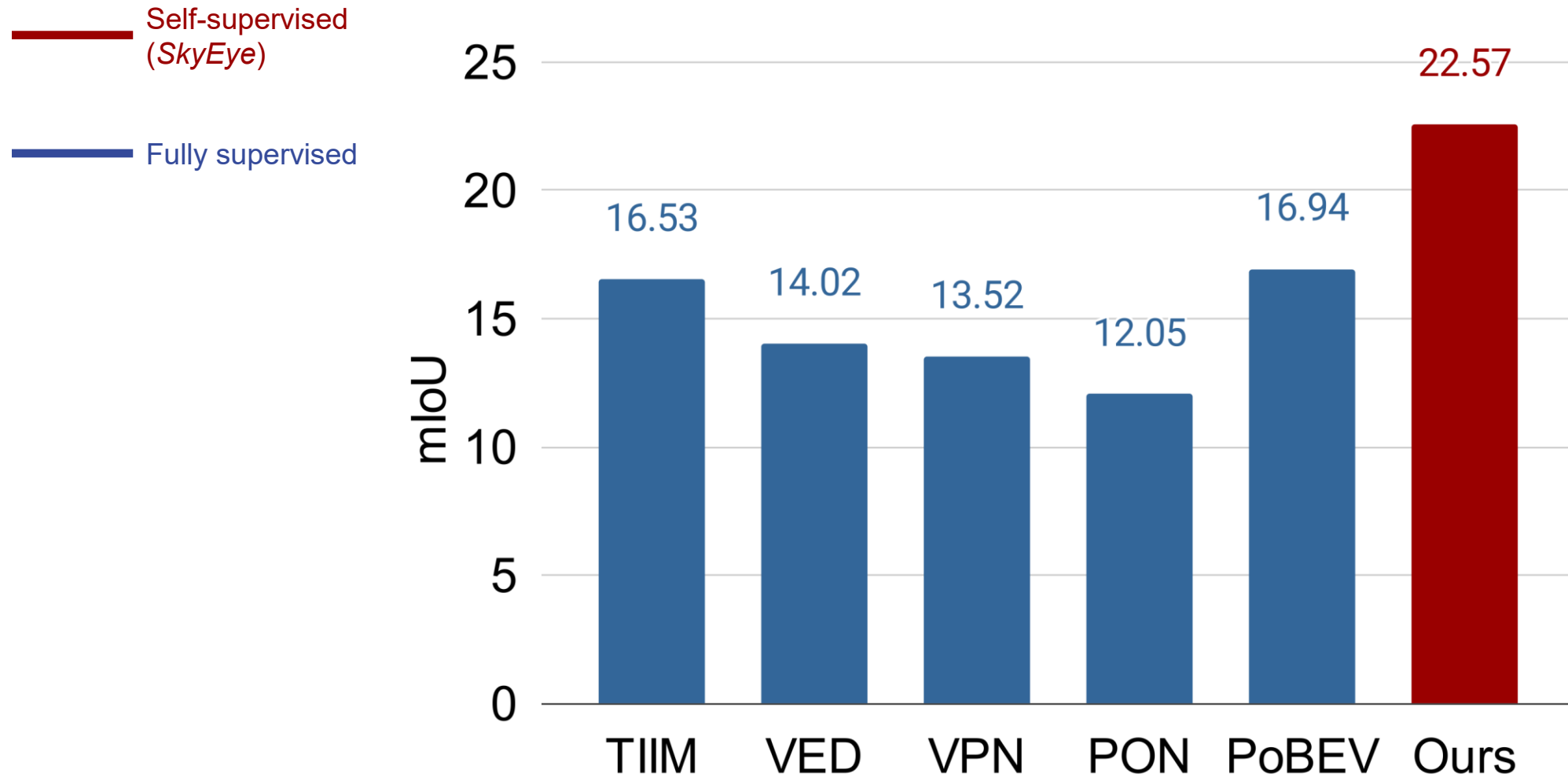


# Impact of Model Pretraining: KITTI-360-BEV



Method	BEV GT	Pretraining
● SkyEye (with pretraining)	×	✓
★ SkyEye* (without pretraining)	×	×
▲ PoBEV	✓	--

# Generalizability of SkyEye: KITTI-360-BEV → Waymo-BEV

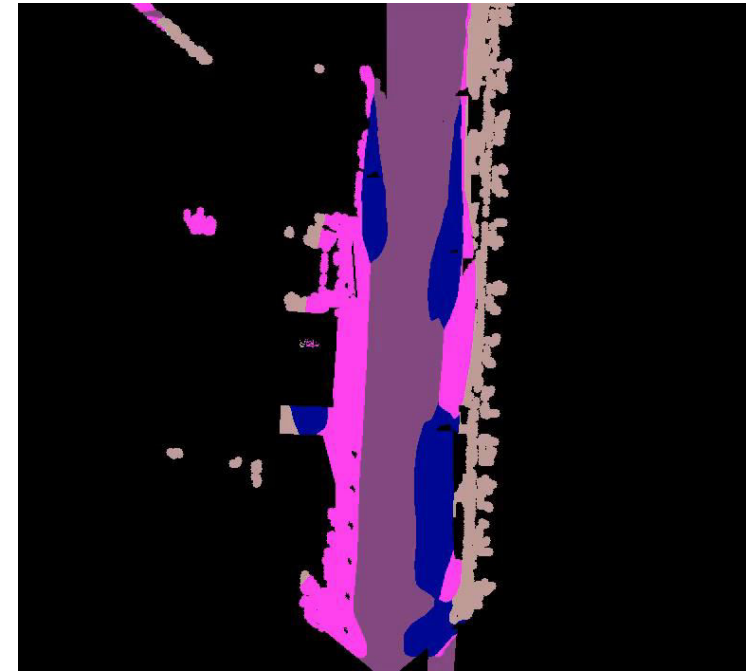


# Qualitative Evaluation: KITTI-360

Input Image in FV



BEV Segmentation Prediction



**SkyEye is on par with the current state-of-the-art without using any ground truth data in BEV**

# Collaborators



Nikhil Gosala



Kürsat Petek



Paulo L. J. Drews-Jr



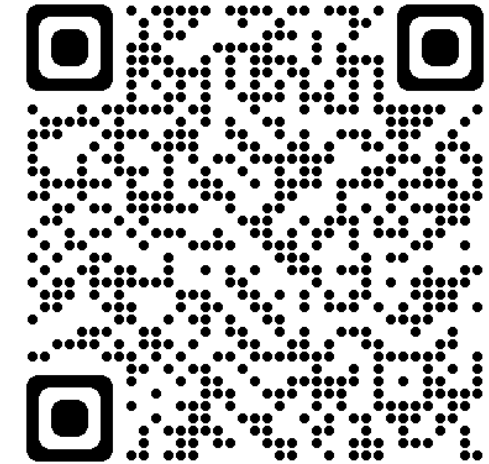
Wolfram Burgard



Abhinav Valada

universität freiburg

UTN  FURG



<http://skyeye.cs.uni-freiburg.de/>

Poster ID: WED-PM-243