



# High-Performance and Energy-Saving Autonomous Navigation System for Aerial-Ground Robots

MPhil Candidate : Junming Wang

Supervisor : Prof.Heming Cui



## Traditional Mobile Robots



Husky UGV



Agilex Robotics



DJI M350 RTK



Meituan Robotics



## Aerial-Ground Robots (AGR)s is Coming !

(Transcending Limitations of Traditional Robots)

### Academia



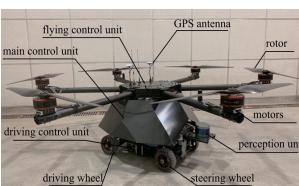
Nature Communications'23



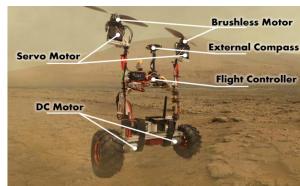
RA-L'22



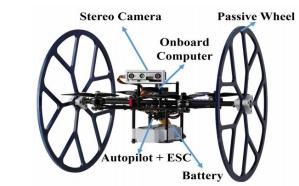
RA-L'22



T-Mech'21



IROS'23



IROS'23



IROS'22



AIM'23

### Industry



AFWERX



Pegasus III



Xiaopeng Huitian X1



Xiaopeng Huitian X2



Jetson One



Pal-V



Aska

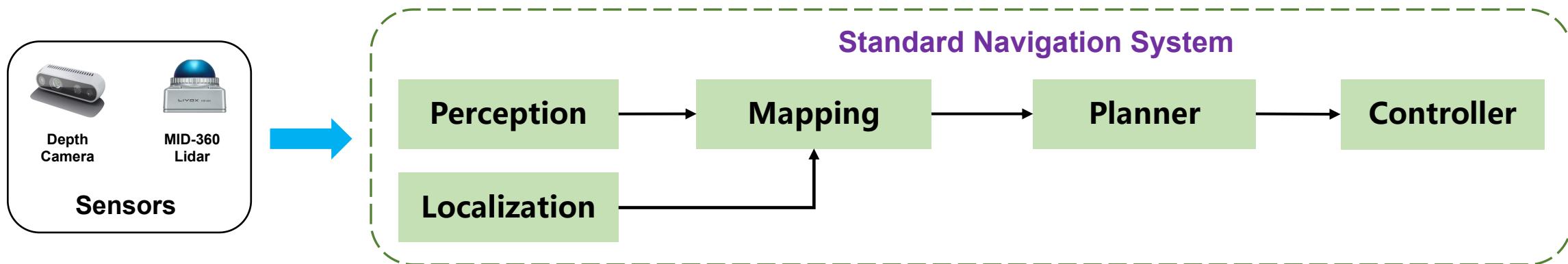


SkyDrive

### Limitations of Traditional UAVs and UGVs Necessitating Aerial-Ground Robotic Systems:

- Limited mobility:** UAVs can only fly, while UGVs are restricted to ground movement.
- Short endurance:** UAVs have limited flight time, and UGVs consume high energy in complex environments.
- Constrained workspace:** UAVs operate primarily in the air, while UGVs are limited to the ground.
- Poor environmental adaptability:** UAVs struggle in adverse weather or complex environments, and UGVs face challenges in extreme terrains.

## ■ Autonomous Navigation in Complex Environments is Very Important For AGRs



## ■ Challenges in Achieving High-performance and Energy-efficient Autonomous Navigation

### Challenge 1 - Limitations of the **Perception Module**:

- The sensor's limited field of view obstructs the perception of obstacle distribution behind occlusions, resulting in an **incomplete local map with unknown and occluded areas**.

### Challenge 2 - Limitations of the **Path Planner**:

- Existing planners fail to address AGR-specific requirements, particularly **energy efficiency** and **dynamic constraints**. AGRs' inherent non-holonomic constraints make directly migrating and using UAV planners impractical.

### Challenge 3 - Prediction Result **Update**:

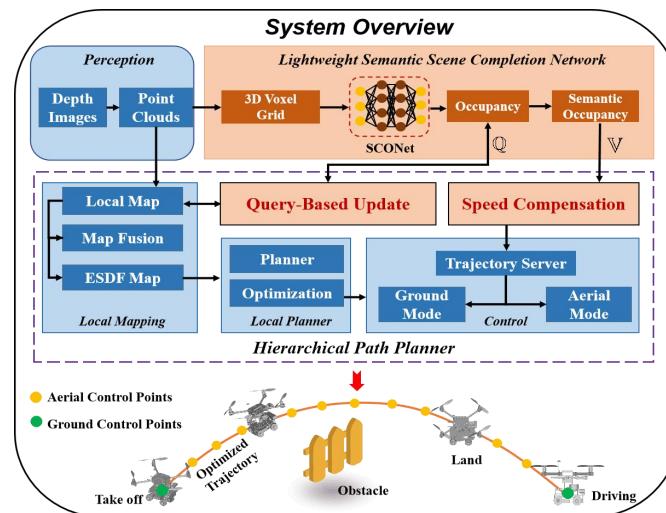
- While a new 3D semantic occupancy network can predict obstacle distribution in occluded areas, exploring **low-latency** methods to update these predictions into the local map remains challenging.

## Outline: High-Performance and Energy-Saving Autonomous Navigation System for Aerial-Ground Robots

- ◆ **Basic idea:** By proposing a novel perception network and a path planner tailored specifically for Aerial-Ground Robots (AGR), we can address all three challenges posed by existing works and achieve an optimal trade-off between **performance and efficiency** in two distinct navigation systems:

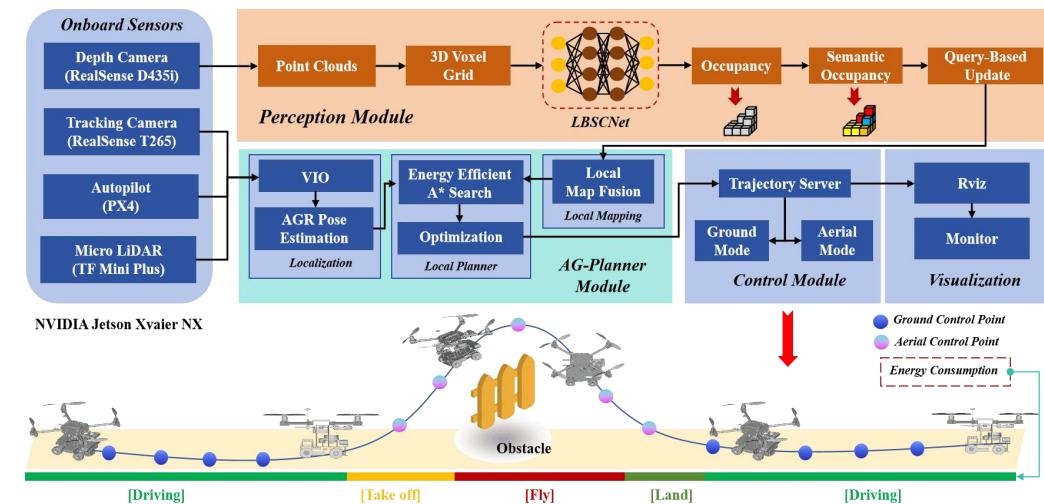
### 1. AGRNav: Efficient and Energy-Saving Autonomous Navigation for Air-Ground Robots in Occlusion-Prone Environments

### 2. HE-Nav: A High-Performance and Efficient Navigation System for Aerial-Ground Robots in Cluttered Environments



**AGRNav**

*The First AGR-Tailored Occlusion-Aware Navigation System*





# AGRNav: Efficient and Energy-Saving Autonomous Navigation for Air-Ground Robots in Occlusion-Prone Environments

# Background

- AGR combines the advantages of unmanned ground vehicle (UGV) and unmanned aerial vehicles (UAV).



UGV

**Advantages:**

- ✓ long endurance
- ✓ low energy consumption

**Disadvantages:**

- ◆ Limited field of view
- ◆ Low velocity



UAV

**Advantages:**

- ✓ High mobility
- ✓ Fast velocity
- ✓ Wide field-of-view

**Disadvantages:**

- ◆ High energy consumption
- ◆ Low load capacity



- **Air-ground robots (AGRs)**, which are known for their outstanding mobility and long endurance, have been gaining significant interest lately and show great potential for applications in search and rescue tasks.



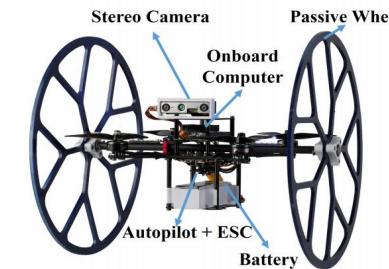
Nature Communications'2023 [1]



RAL'2022 [2]



T-Mech'2021 [3]



RAL'2022 [4]

[1] Sihite E, Kalantari A, Nemovi R, et al. Multi-Modal Mobility Morphobot (M4) with appendage repurposing for locomotion plasticity enhancement[J]. Nature communications, 2023, 14(1): 3323.

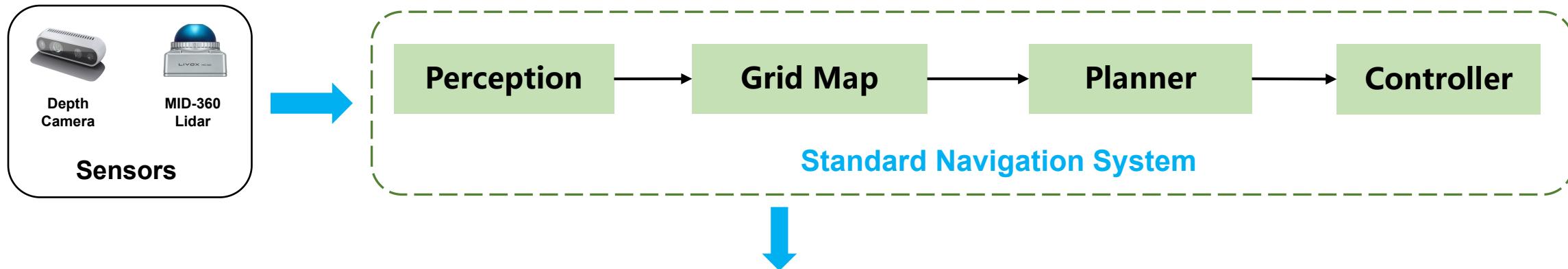
[2] Yang J, Zhu Y, Zhang L, et al. SytaB: A class of smooth-transition hybrid terrestrial/aerial bicopters[J]. IEEE Robotics and Automation Letters, 2022, 7(4): 9199-9206.

[3] Tan Q, Zhang X, Liu H, et al. Multimodal dynamics analysis and control for amphibious fly-drive vehicle[J]. IEEE/ASME Transactions on Mechatronics, 2021, 26(2): 621-632.

[4] Zhang R, Wu Y, Zhang L, et al. Autonomous and adaptive navigation for terrestrial-aerial bimodal vehicles[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 3008-3015.

# Requirements

- To achieve **collision-free** autonomous navigation, the ***perception***, ***planning***, and ***control*** modules within a standard robot navigation system are deployed on the onboard computer (i.e., NVIDIA Jetson Xavier NX) and run asynchronously.



- An ideal AGR navigation system needs to meet the following requirements:

Efficient

◆ **Planning Success Rate:** No collision occurs during movement from the starting point to the endpoint.

◆ **Moving Time:** The movement time from the starting point to the endpoint is short.

Energy-saving

◆ **Energy Consumption:** Minimal power consumption during navigation.

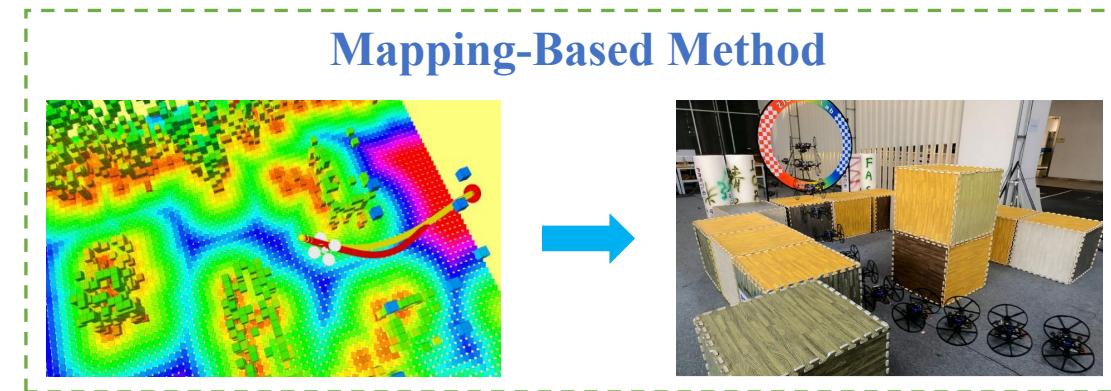
◆ **Mode Switching:** Switch to flight mode only when encountering insurmountable obstacles.



## Existing AGR Navigation Systems **Cannot** Meet all Requirements

- Existing robot navigation systems include **Mapping-Based** methods and **Learning-Based** methods.
- **Mapping-Based Methods** (e.g., TABV [RAL' 2022]; HDF [IROS' 2019])
- **How it works**

- ① Perceive surrounding obstacles and build an **occupany grid map**.
- ② Construct Euclidean Signed Distance Field (**ESDF**) map.
- ③ A path planner based on the **A\* algorithm** searches for trajectories.
- ④ Tracking the trajectory using the PX4 controller.



- **Advantages:** High navigation **success rate** in simple and unobstructed scenarios. 

- **Limitations:** High collision risk and suboptimal energy consumption in occluded and unknown scenarios. 

Inefficient 

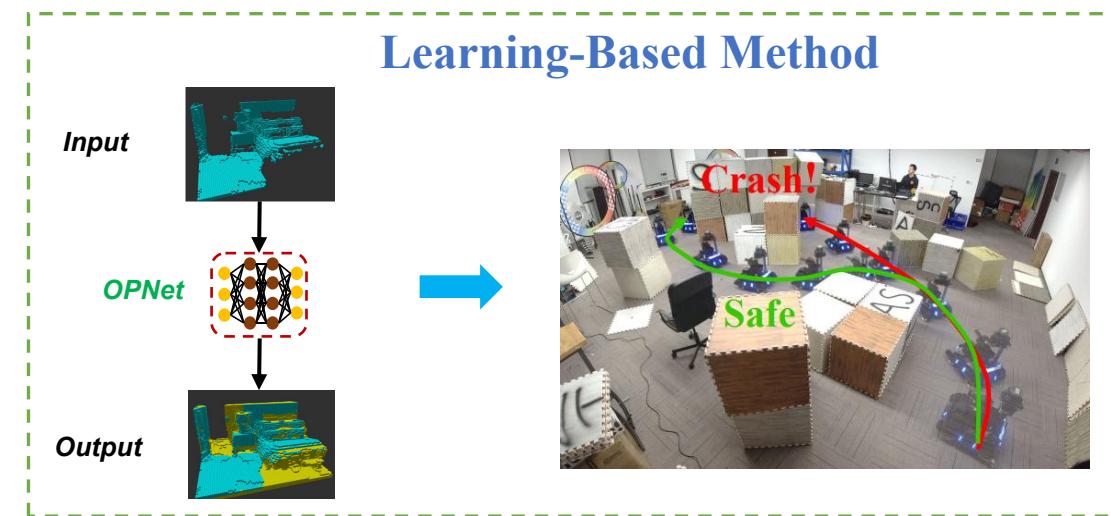
 *Redundant ground trajectories and flight trajectories result in additional energy consumption*

# Challenge

- Learning-Based Methods, **only** designed for ground robots (e.g., OPNet [IROS' 2021])

- **How it works**

- ① Perceive surrounding obstacles and build an **original grid map**.
- ② Use **OPNet** to predict obstacle distribution in occluded areas.
- ③ **Fusion** of predicted local map and original local map.
- ④ Path planner planning trajectory.



- **Advantages:** Predict the distribution of obstacles in blocked areas in advance to **improve** planning success rate. 😊

- **Limitations:**

- The prediction **accuracy** is **low**. 😞
- The map update **delay** is **high**. 😞



# Our Question and Observation

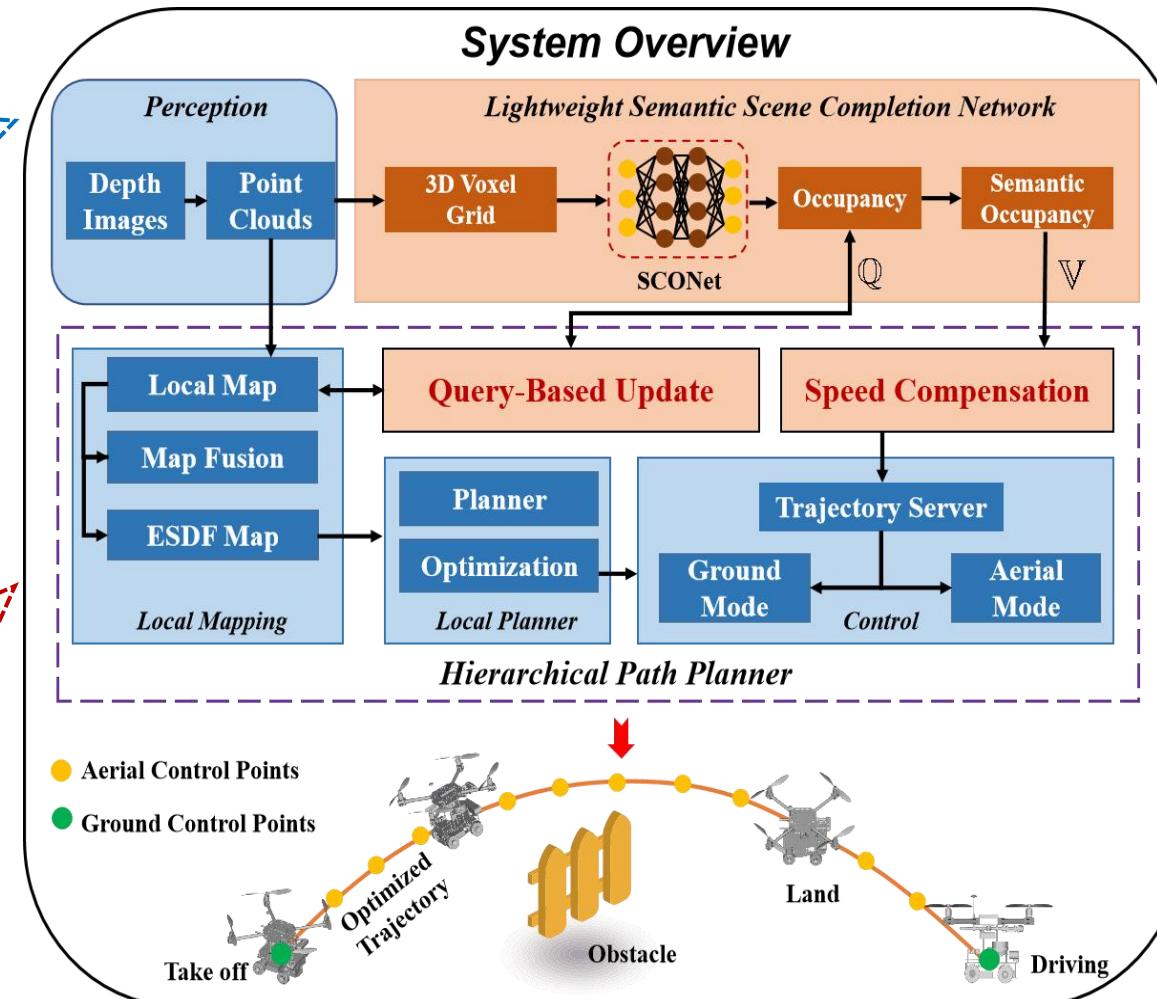
Can we address the limitations in mapping-based and learning-based methods to create an AGR navigation system that is both **highly efficient** and **energy-saving**?

- A lightweight semantic completion network is proposed to predict the distribution of obstacles in occlusion areas in real time.
  - ✓ Depthwise separable convolution;
  - ✓ self-attention mechanism;
- A novel update strategy is proposed to avoid high-latency occupancy updates (e.g., directly merging two maps).
  - ✓ Query-based method;
- A hierarchical path planner is proposed to search energy-efficient, safe, smooth and dynamically feasible air-ground hybrid paths.

Methods	Category	Efficiency Metrics		Energy Metric	Additional Network Inference Metrics		Suitable for AGRs
		Planning Success Rate	Moving Time		Energy Consumption	Prediction Accuracy	
HDF [IROS'19]	Mapping-Based	:(	:(	:(		None	None
TABV [RAL'22]	Mapping-Based	:(	:(	:(	:(	None	None
OPNet [IROS'21]	Learning-Based	None	None	None	:(	:(	X
AGRNav (Ours)	Prediction-Based	:(	:(	:(	:(	:(	:(

# ARGNav: Efficient and Energy-Saving

## ◆ The Overview of Our Proposed System: AGRNav.



### ESDF-Based Planner

- (i) a Kinodynamic A\* path searching front-end,
- (ii) an ESDF-based trajectory optimization back-end.
- (iii) a post-refinement procedure.

### Mode Switching

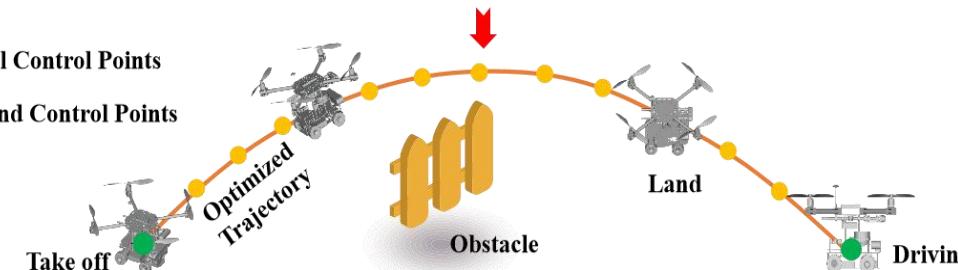
- When the z-axis coordinate of the next control point is greater than the ground threshold, that is, when mode switching is required, an additional trigger signal will be sent to the controller (i.e., PX4 Autopilot).

### SCONet

- To enable SCONet to capture rich and dense contextual information as well as features of occlusion areas, it integrates two self-attention mechanisms.

### Query-based Update Strategy

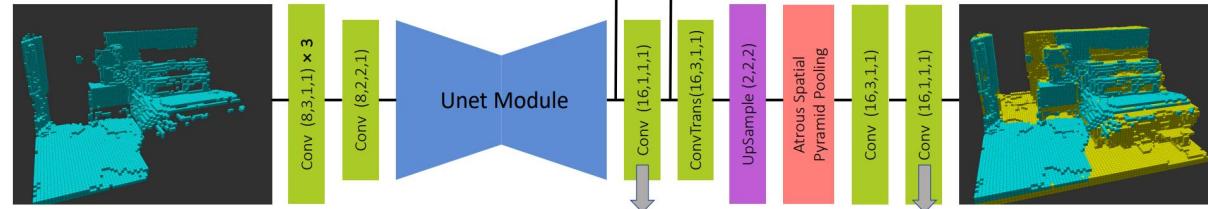
- After mapping, only the occupancy status of free voxels is queried and updated instead of fusing the predicted map and the original map to ensure low latency.



# Contribution ①



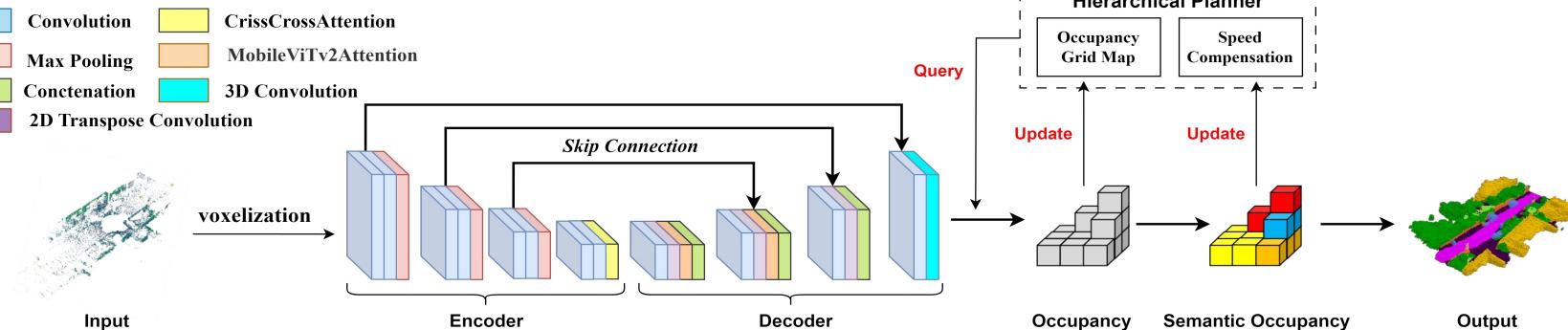
## A Lightweight Semantic Scene Completion Network (SCONet)



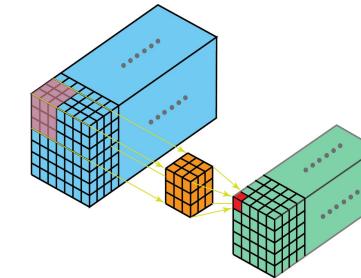
**Network Limitation:**

- Computationally intensive 3D convolution
- High latency update strategy
- Unable to capture fine-grained features

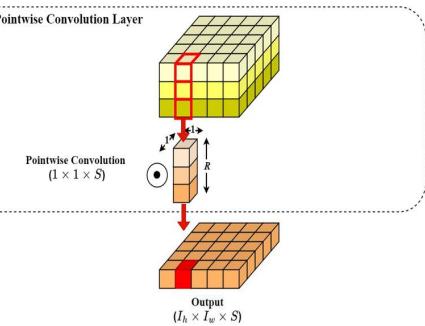
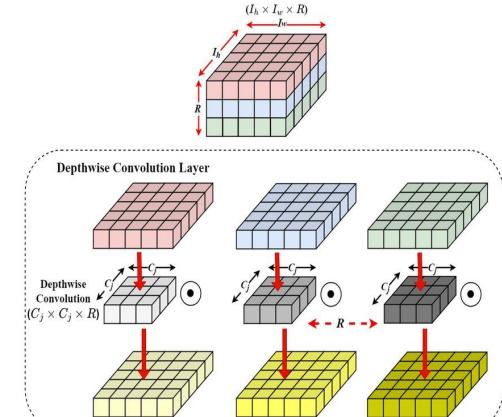
Convolution	CrissCrossAttention
Max Pooling	MobileViTv2Attention
Concatenation	3D Convolution
2D Transpose Convolution	



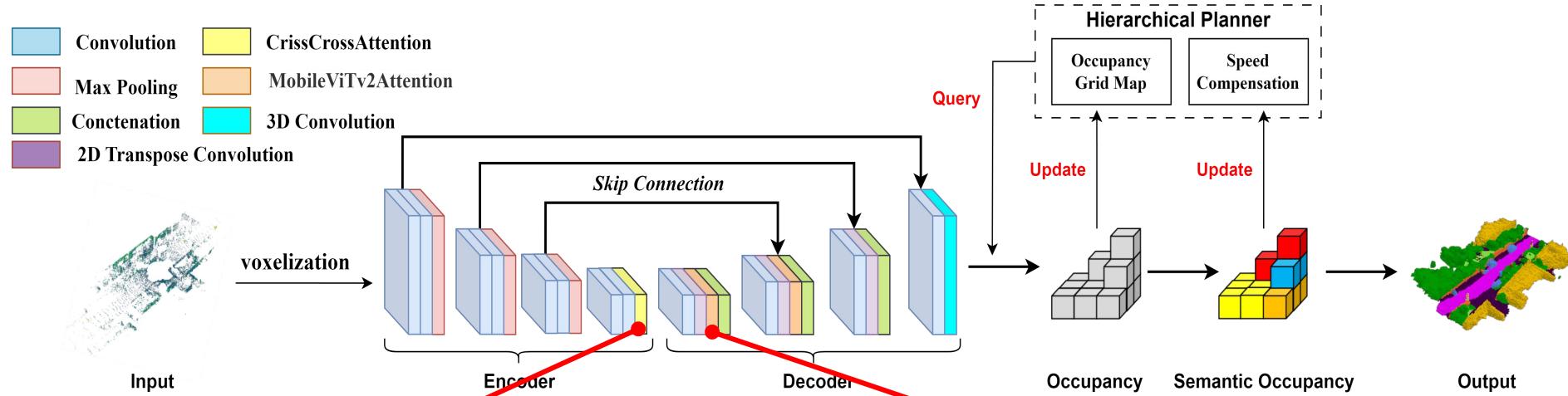
Standard Convolution



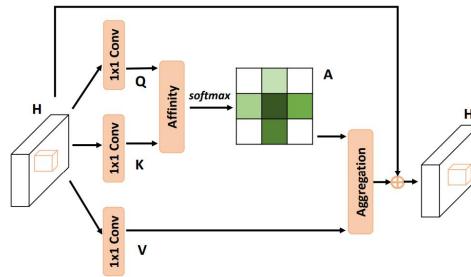
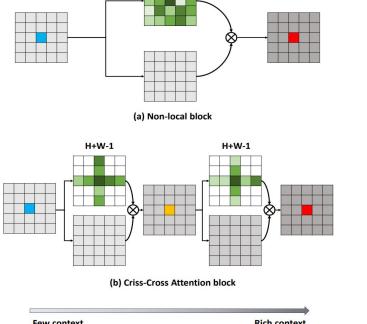
Depthwise Separable Convolution



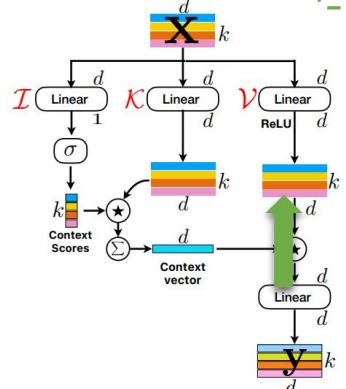
# Contribution ①



## Criss-Cross Attention (CCA)



## MobileViT-v2 Attention

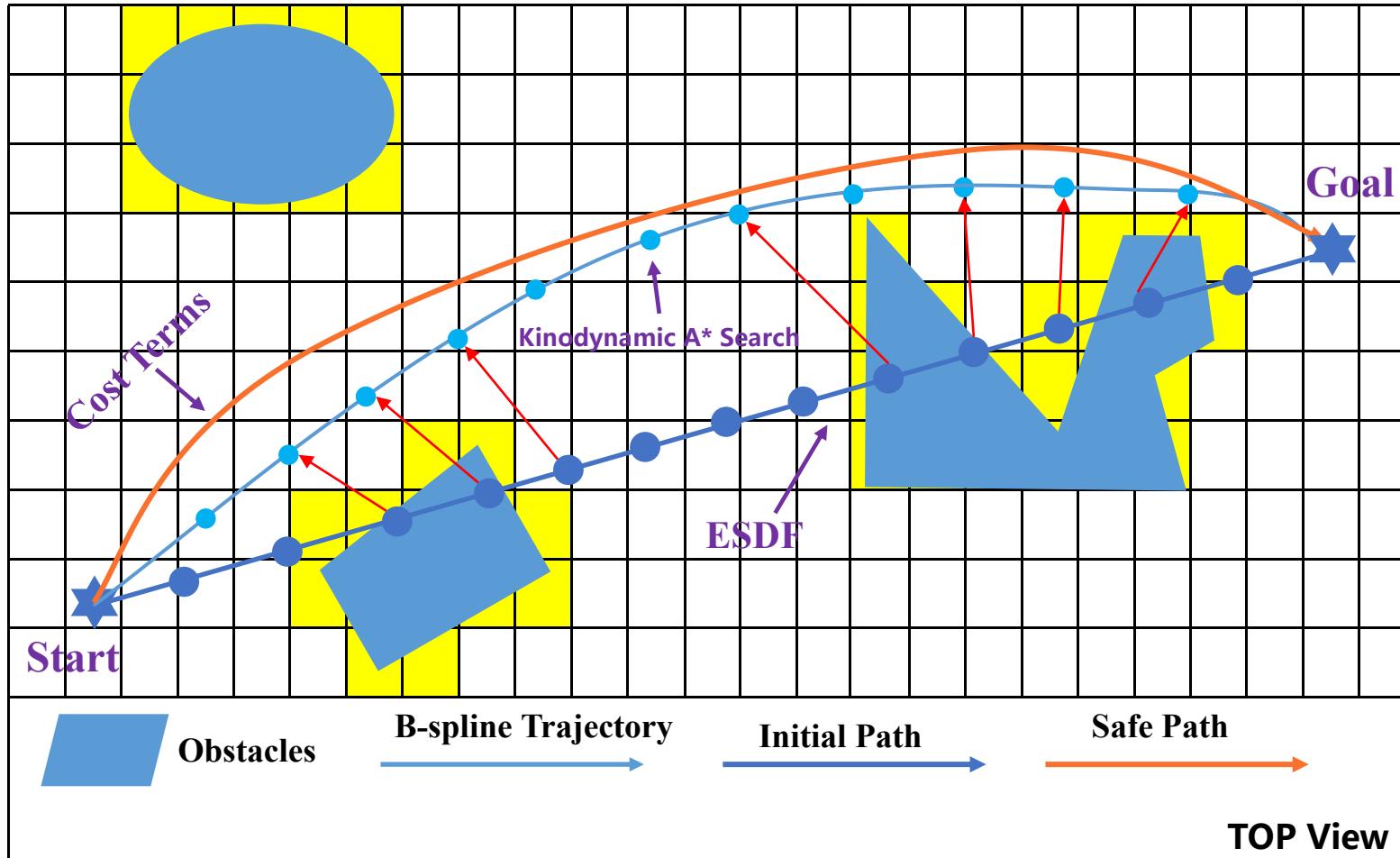


$$y = \left( \sum_{c_v \in \mathbb{R}^d} \left( \underbrace{\sigma(xW_I) * xW_K}_{\text{Context Scores}} \right) * \text{ReLU}(xW_V) \right) W_O$$

★ Broadcasted element-wise multiplication  
 σ Softmax  
 Σ Element-wise sum  
 ● Dot-product  
 || Concatenation

## Contribution ②

# A Hierarchical ESDF-Based Path Planner



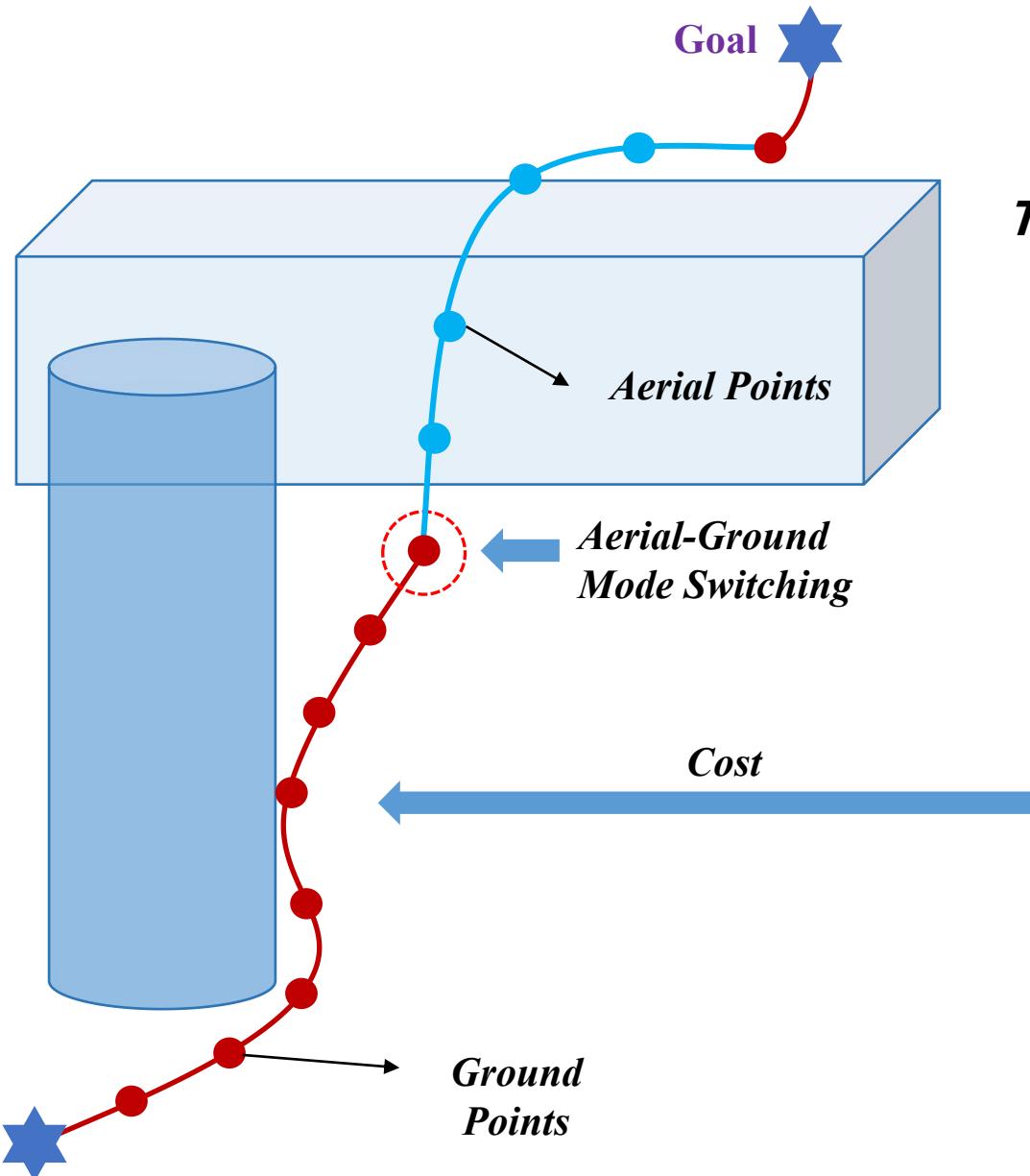
### Algorithm 1: Kinodynamic Path Searching

```

1 Initialize();
2 while  $\neg \mathcal{P}.\text{empty}()$  do
3    $n_c \leftarrow \mathcal{P}.\text{pop}()$ ,  $\mathcal{C}.\text{insert}(n_c)$ ;
4   if ReachGoal( $n_c$ )  $\vee$  AnalyticExpand( $n_c$ ) then
5     return RetrievePath();
6   primitives  $\leftarrow \text{Expand}(n_c)$ ;
7   nodes  $\leftarrow \text{Prune}(\text{primitives})$ ;
8   for  $n_i$  in nodes do
9     if  $\neg \mathcal{C}.\text{contain}(n_i) \wedge \text{CheckFeasible}(n_i)$  then
10       $g_{\text{temp}} \leftarrow n_c.g_c + \text{EdgeCost}(n_i)$ ;
11      if  $\neg \mathcal{P}.\text{contain}(n_i)$  then
12         $\mathcal{P}.\text{add}(n_i)$ ;
13      else if  $g_{\text{temp}} \geq n_i.g_c$  then
14        continue;
15       $n_i.parent \leftarrow n_c$ ,  $n_i.g_c \leftarrow g_{\text{temp}}$ ;
16       $n_i.f_c \leftarrow n_i.g_c + \text{Heuristic}(n_i)$ ;

```

## Contribution ②



*The overall objective function is formulated as follows:*

$$f_{total} = \lambda_s f_s + \lambda_c f_c + \lambda_f (f_v + f_a) + \lambda_n f_n.$$

$$\left\{ \begin{array}{l} f_s = \sum_{i=p_b-1}^{N-p_b+1} \left\| \underbrace{(\mathbf{Q}_{i+1} - \mathbf{Q}_i)}_{\mathbf{F}_{i+1,i}} + \underbrace{(\mathbf{Q}_{i-1} - \mathbf{Q}_i)}_{\mathbf{F}_{i-1,i}} \right\|^2 \quad \text{Smoothness Cost} \\ f_n = \sum_{i=1}^{M-1} F_n(\mathbf{Q}_{ti}) \quad \text{Curvature Cost} \\ f_c = \sum_{i=p_b}^{N-p_b} F_c(d(\mathbf{Q}_i)) \quad \text{Collision Cost} \\ f_v = \sum_{\mu \in \{x,y,z\}} \sum_{i=p_b-1}^{N-p_b} F_v(V_{i\mu}), \quad f_a = \sum_{\mu \in \{x,y,z\}} \sum_{i=p_b-2}^{N-p_b} F_a(A_{i\mu}) \quad \text{Dynamical Feasibility Cost} \end{array} \right.$$

# Implementation and Evaluation

## ◆ Implementation Details:

- **SCONet** was trained and tested on a server with 4 NVIDIA RTX 3090 GPUs using the SemanticKITTI dataset, achieving optimal map completion accuracy after 80 epochs of training with data augmentation and deployed offline.
- **AGRNav** is developed using C++ and packaged as a ROS package for deployment on the onboard computer (i.e., Jetson Xvaier NX).

## ◆ Baseline AGR Navigation Systems:

- HDF [IROS'2019]
- TABV [ICRA'2022]

## ◆ Evaluation settings:

- **Simulation Experiments** were executed on a laptop equipped with Ubuntu 20.04, an i9-13900HX CPU, and an NVIDIA RTX 4060 GPU to simulate aerial-ground robotic navigation within complex environments.
- We employed AGRNav on a custom AGR platform for **indoor and outdoor experiments**, using Prometheus software, with a RealSense D435i depth camera, a T265 tracking camera, and a Jetson Xavier NX computer.



Mode	Average Power	Time (mins)
Fly	987.61 W	14
Hover	532.07 W	26
Ground	197.52 W	55



# Evaluation Questions

01

◆ How is AGRNav' s end-to-end **navigation efficient** compared to baselines ?

02

◆ How does the **energy consumption** of AGRNav compare to the baseline ?

03

◆ How does the **prediction accuracy** and **real-time** performance of key components SCONet compare to the baseline?

# End-to-End Performance

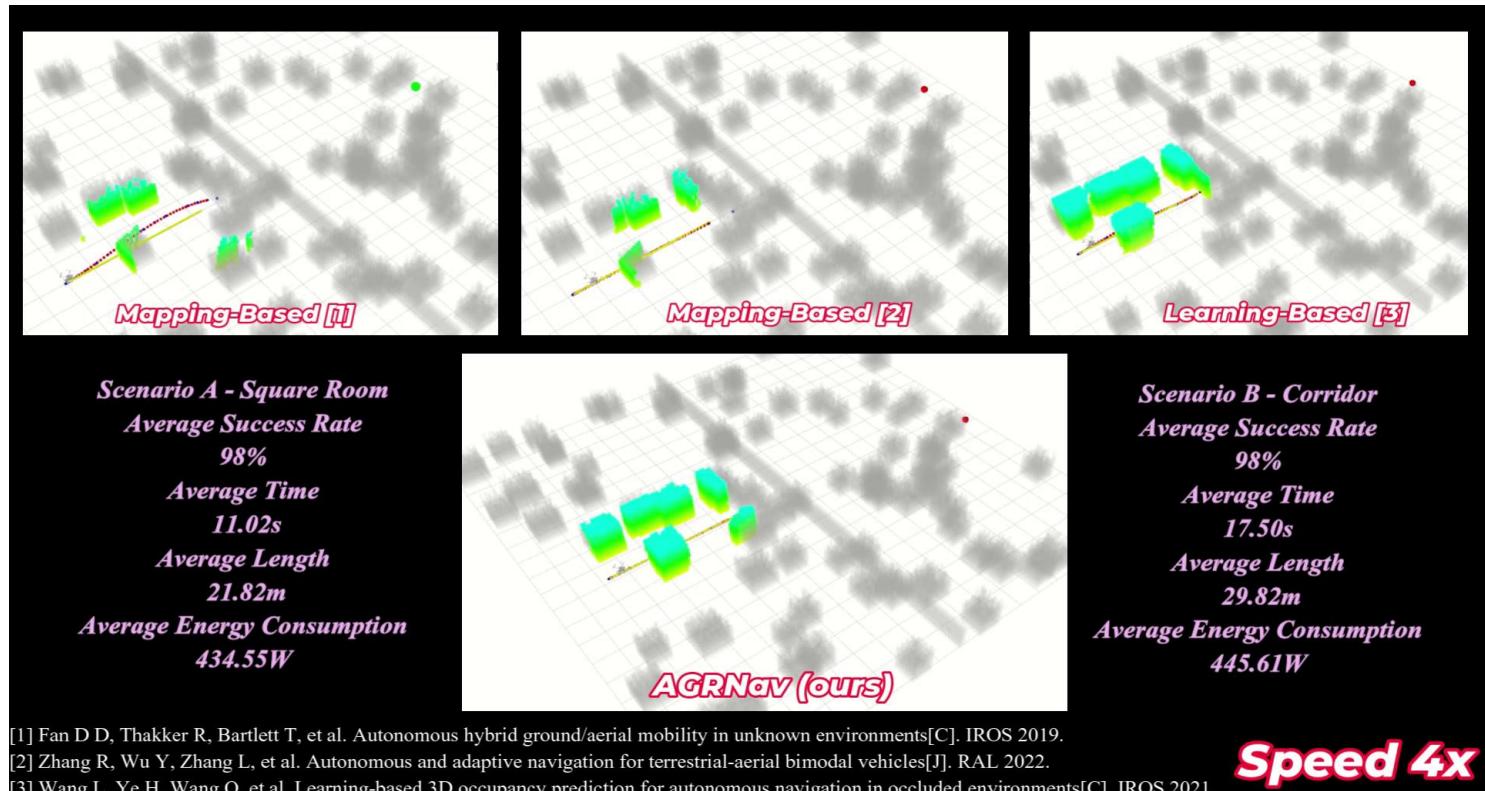
## ◆ Simulation Experiments:

Table 1

Mode	Average Power	Time (mins)
Fly	987.61 W	14
Hover	532.07 W	26
Ground	197.52 W	55

Table 2

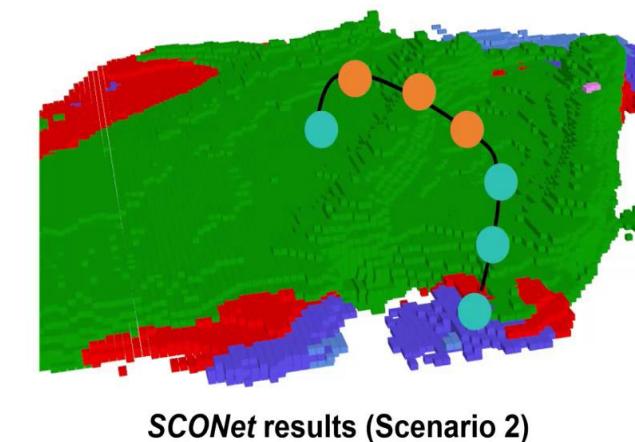
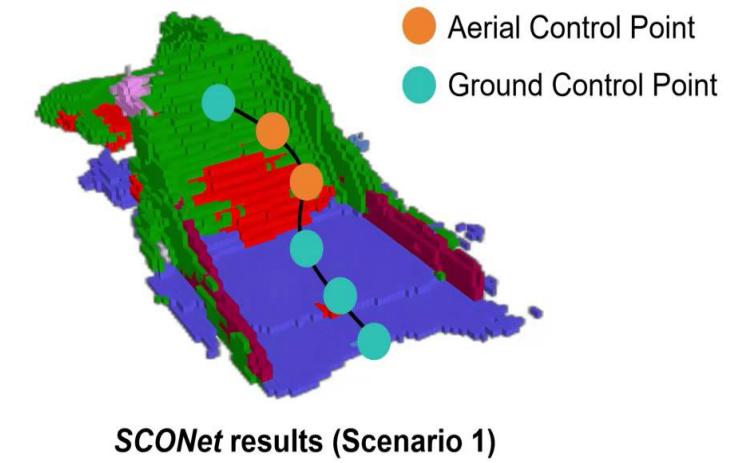
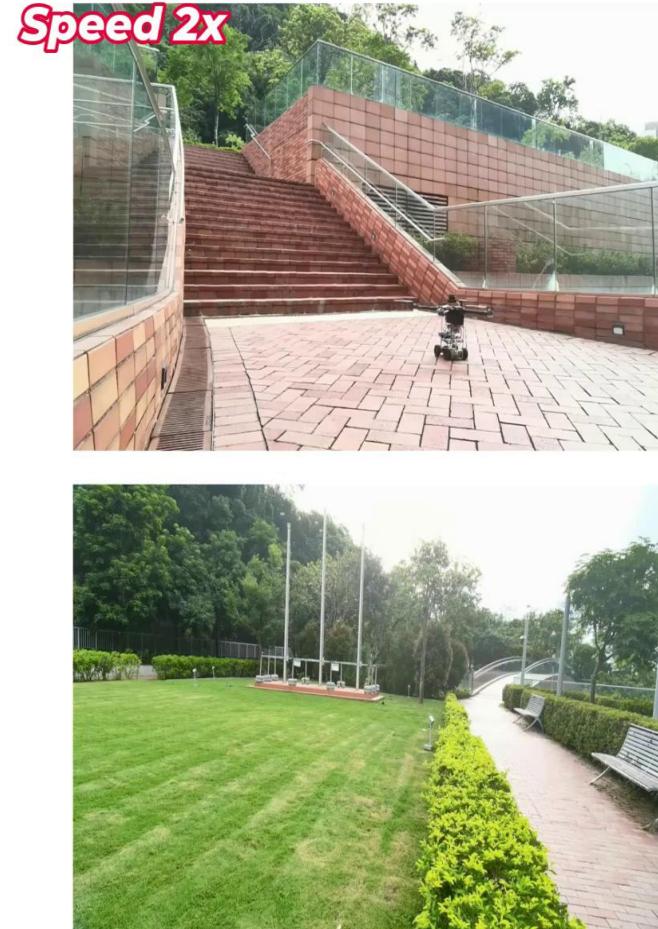
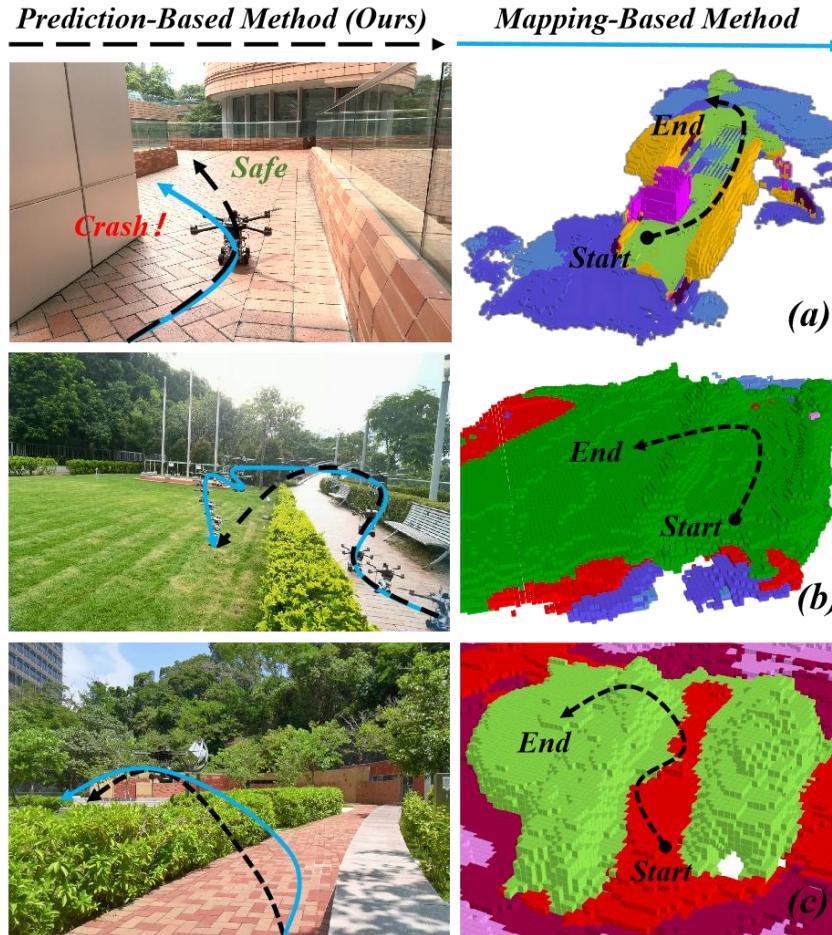
Env.	Method	Succ. (%)	Time (s)	Leng. (m)	Power (W)
Square	Fan's [5]	85.0	13.13	33.79	919.07
	Zhang's [4]	95.0	12.05	23.09	793.30
Room	OPNet [12]	91.0	12.90	32.12	888.04
	AGRNav(Ours)	<b>98.0</b>	<b>11.02</b>	<b>21.82</b>	<b>434.55</b>
Corridor	Fan's [5]	88.0	21.24	33.10	565.24
	Zhang's [4]	97.0	<b>16.97</b>	30.69	519.20
	OPNet [12]	90.0	18.45	32.85	534.11
	AGRNav(Ours)	<b>98.0</b>	17.50	<b>29.82</b>	<b>445.61</b>



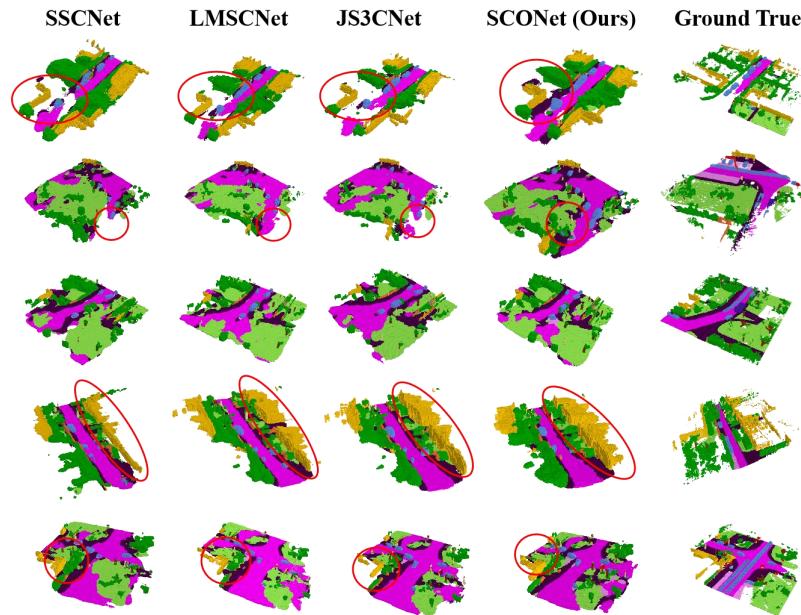
- **Table 2** : our AGRNav outperforms the other three approaches, achieving the **highest success rate (98%)**.
- **Table 2** : our AGRNav substantially reduces redundant paths and cuts energy consumption by half (i.e., **average consumption per second is 434.55 W**) in a square room.
- **Table 2** : in the corridor scene, while the average travel time of is **shorter (16.97 s)**, its average energy consumption is higher due to the inability to predict occlusion areas and a greater reliance on aerial paths.

# End-to-End Performance

## ◆ Real-World Experiments:



# SCONet Performance



**Table 3**

Method	<i>IoU</i>	<i>Prec.</i>	<i>Recall</i>	<i>FPS</i>	<i>mIoU</i>
SSCNet [10]	53.20	59.13	<b>84.15</b>	12.00	14.55
SG-NN [21]	31.26	31.60	54.50	12.00	9.90
J3S3Net [22]	51.10	40.23	61.09	1.73	23.80
LMSCNet [13]	54.89	82.21	62.29	18.50	14.13
S3CNet [23]	45.60	48.79	77.13	1.82	<b>29.50</b>
TDS [24]	50.60	72.43	78.61	1.70	17.70
<b>SCONet (our)</b>	<b>56.12</b>	<b>85.02</b>	63.47	<b>20.00</b>	17.61

- **Table 3:** our SCONet outperforms its rivals, registering the highest IoU completion metric score of **56.12**.
- **Table 3:** despite a slightly lower mIoU compared to S3CNet and J3S3Net, SCONet's inference speed is significantly enhanced, being about **20 times faster**. This is primarily due to the adoption of depthwise separable convolutions instead of the resource-heavy 3D convolutions in its encoder, enabling real-time efficiency with **20 FPS** on an RTX 3090 GPU.



Method	IoU ↑	mIoU ↑
SCONet (ours)	55.50	16.10
w/o Depth-Separable Convolution	54.15	15.76
w/o Criss-Cross Attention	53.20	15.11
w/o MobileViT-v2 Attention	53.86	15.37

Ablation studies on the SemanticKITTI validation set highlight the significance of two key components in our network: **self-attention mechanisms** and **depth-separable convolutions**.

- The CCA mechanism substantially impacts completion and semantic prediction by effectively aggregating context across rows and columns.
  - ◆ Without CCA causes a **4.14%** and **6.15% drop** in completion and semantic completion, respectively.
- MobileViT-v2 Attention captures local scene features, such as occluded areas, with low computational overhead.
  - ◆ Without MobileViT-v2 Attention leads to a **2.95% decline** in IoU.
- Furthermore, depth-separable convolutions significantly reduce the number of parameters.



# AGRNav Conclusion and Future Work

- AGRNav, the first efficient and energy-saving autonomous navigation system for air-ground robots.
  - ◆ SCONet, which outperforms state-of-the-art models in prediction accuracy and inference time;
  - ◆ a hierarchical path planner, improved by a query-based low-latency update method, considers obstacles in occluded areas to generate paths;
  - ◆ not only minimizes collision risk but also reduces energy consumption by 50% compared to the baseline;
  - ◆ The system's robustness has been extensively validated through experiments in both simulated and real-world environments.

□ The construction of ESDF Map **takes up 70%** of the path planning time. In the future, our work will eliminate the process of building ESDF map and achieve faster and real-time path planning.

**AGRNav open source code link: <https://github.com/jmwang0117/AGRNav>**



# HE-Nav: A High-Performance and Efficient Navigation System for Aerial-Ground Robots in Cluttered Environments

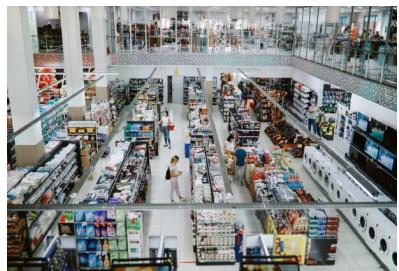


## ➤ Limitations of Existing ESDF-based AGR Navigation System

- The existing Euclidean Signed Distance Field (ESDF)-based navigation system for Aerial-Ground Robots (AGRs) has achieved success in ***lightly occluded environments***, but its performance in ***heavily cluttered or complex environments*** remains a challenge.



Low Bushes



Indoor

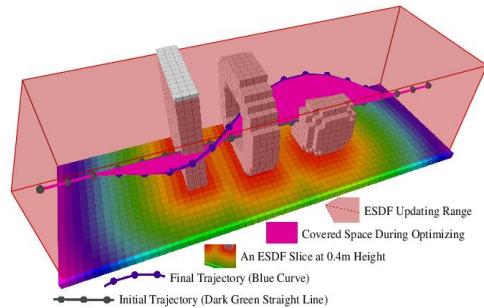


Dense Forest



Large Campus

# Challenge



- While EGO-Planner devised an ESDF-free path planner for **quadcopters**, it fails to address **AGR-specific requirements**, particularly energy efficiency and dynamic constraints.

## EGO-Planner

### Perception Challenges:

- 1. Incomplete Perception: Lightweight 3D semantic occupancy networks like SCONet generate incomplete local maps in cluttered environments due to low prediction accuracy, resulting in high-risk collision paths.
- 2. Computational Accuracy Trade-off: Employing transformers or 3D CNNs can improve accuracy but are impractical for resource-constrained AGRs, while lightweight designs offer computational efficiency but compromise on precision.

### Planning Challenges:

- 1. Inefficient ESDF-based Planning: Existing ESDF-based AGR path planners are inefficient due to redundant calculation times in building the ESDF map, failing to meet real-time requirements.
- 2. Energy Consumption: Flight-centric trajectory generation methods, like those proposed by Zhou et al., result in elevated energy consumption for AGRs.
- 3. Non-holonomic Constraints: The inherent non-holonomic constraints of AGRs make it impossible to naively migrate and use flight-centric planners.
- 4. Mode Switching: Inefficiencies in existing AGR path planners stem from inherent flaws and incomplete local maps, leading to overly conservative or aggressive mode switching (e.g., choice of landing point location), undermining energy efficiency.



# Our Question and Observation

Can we design a more accurate semantic occupancy network and an AGR-tailored, ESDF-free path planner to achieve high-performance navigation for Aerial-Ground Robots in **Cluttered Environments**?

## ■ Perception Observations:

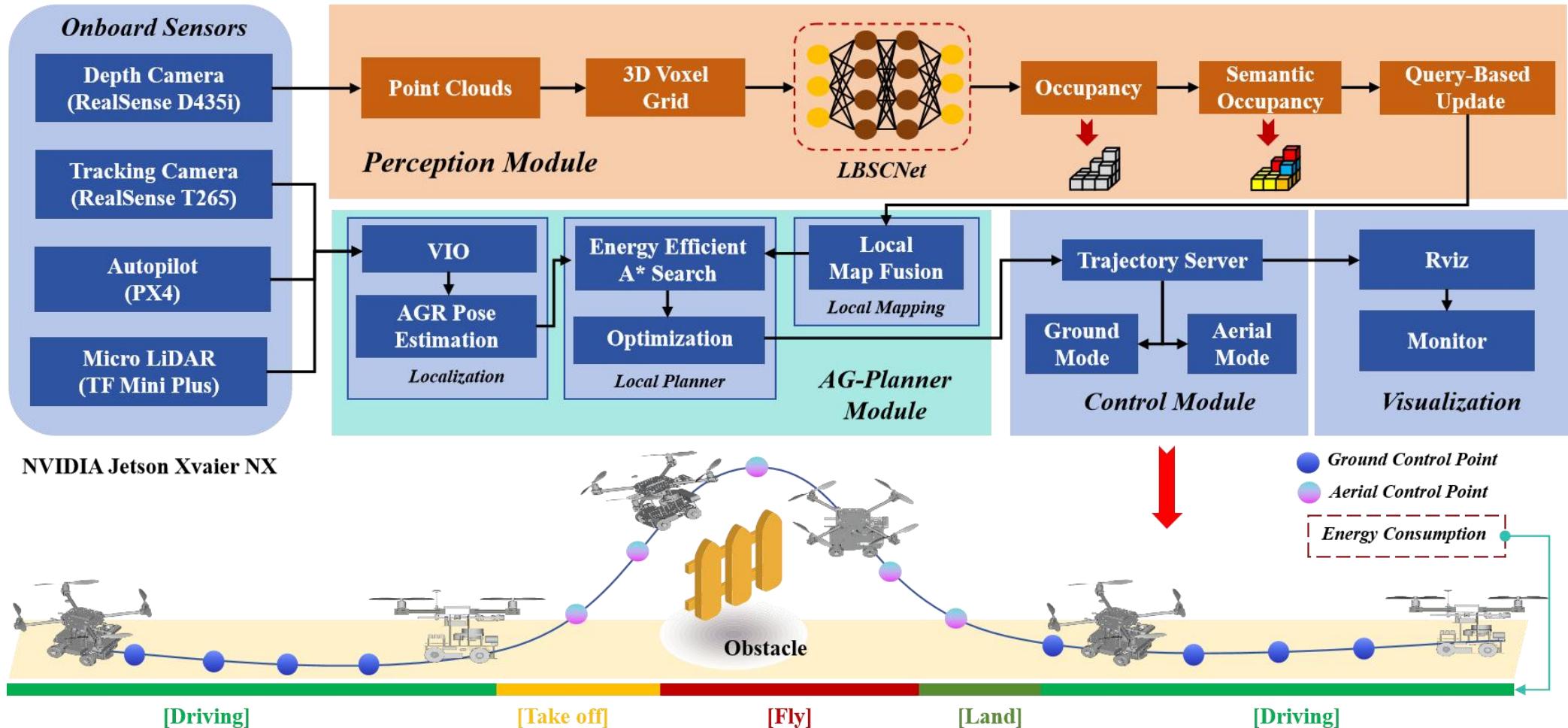
1. Introduce a novel lightweight SSC network that separates the conventional SSC network, which typically learns semantics and geometry together, into distinct network branches, thereby decoupling the processes that were previously jointly learned in networks. This separation enables each branch to focus on acquiring domain-specific features, thereby enhancing the overall performance.
2. Drawing inspiration from BevFusion, transition the feature fusion process to the Bird's Eye View (BEV) space, which has the potential to diminish computational complexity and ensure high-speed inference, culminating in the generation of a complete local map.

## ■ Planning Observations:

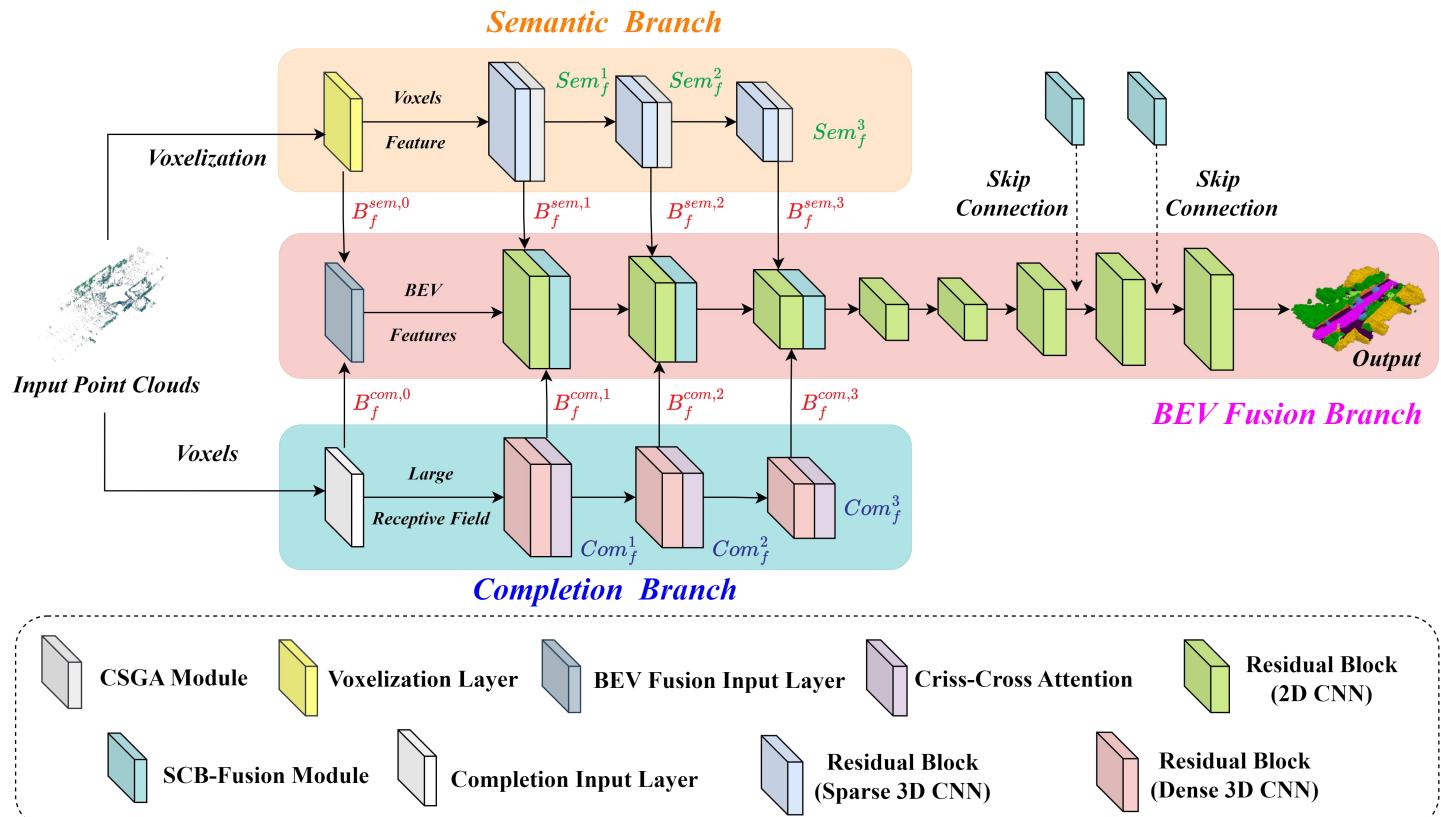
1. Building upon the complete map produced by our SSC network, design a path planner that accommodates the non-holonomic constraints inherent to AGRs to ensure energy-efficient and real-time planning.
2. Remove redundant ESDF calculations and incorporate energy costs associated with different modes (e.g., flying and driving) to facilitate judicious mode switching and promote energy conservation.

# HE-Nav: High-Performance and Efficient

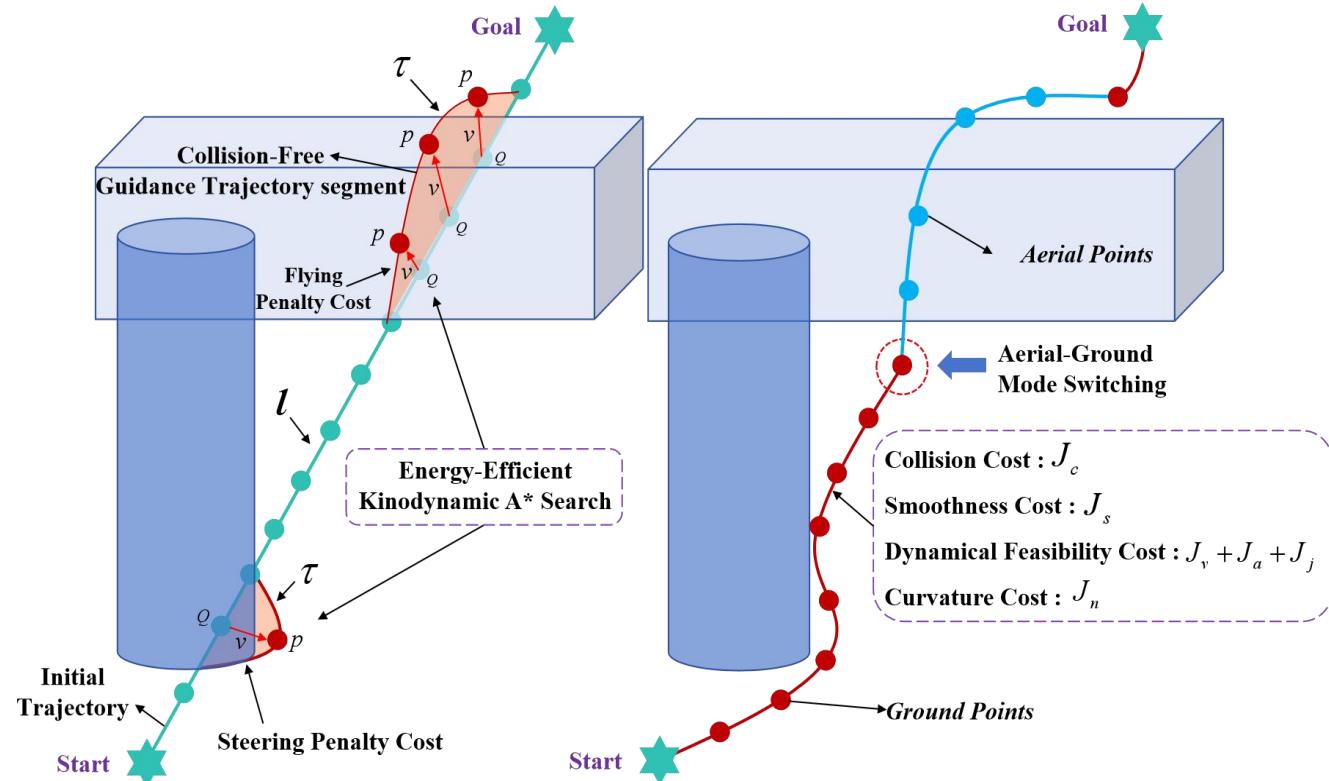
## ◆ The Overview of Our Proposed System: HE-Nav.



# Contribution ①



# Contribution ②



# Implementation and Evaluation

## ◆ Implementation Details:

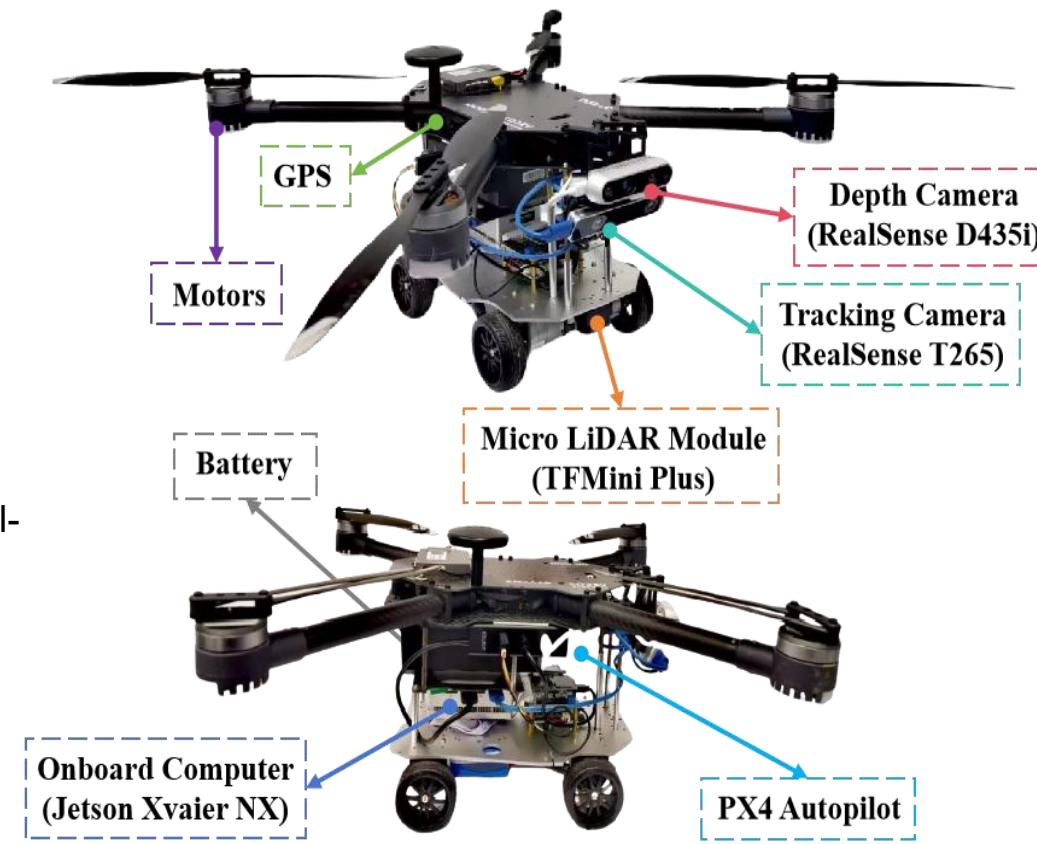
- **LBSCNet** was trained and tested on a server with 4 NVIDIA RTX 3090 GPUs using the SemanticKITTI dataset, achieving optimal map completion accuracy after 80 epochs of training with data augmentation and deployed offline.
- **HE-Nav** is developed using C++ and packaged as a ROS package for deployment on the onboard computer (i.e., Jetson Xavier NX).

## ◆ Baseline AGR Navigation Systems:

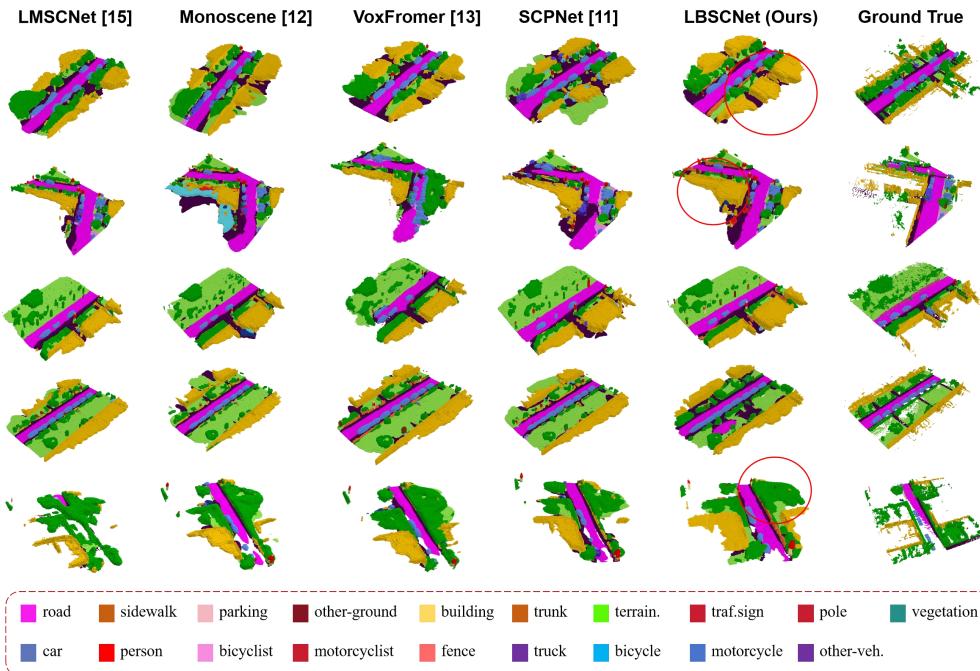
- HDF [IROS'2019]
- TABV [ICRA'2022]
- AGRNav [ICRA'2024]

## ◆ Evaluation settings:

- **Simulation Experiments** were executed on a laptop equipped with Ubuntu 20.04, an i9-13900HX CPU, and an NVIDIA RTX 4060 GPU to simulate aerial-ground robotic navigation within complex environments.
- We employed **HE-Nav** on a custom AGR platform for **indoor and outdoor experiments**, using Prometheus software, with a RealSense D435i depth camera, a T265 tracking camera, and a Jetson Xavier NX computer.



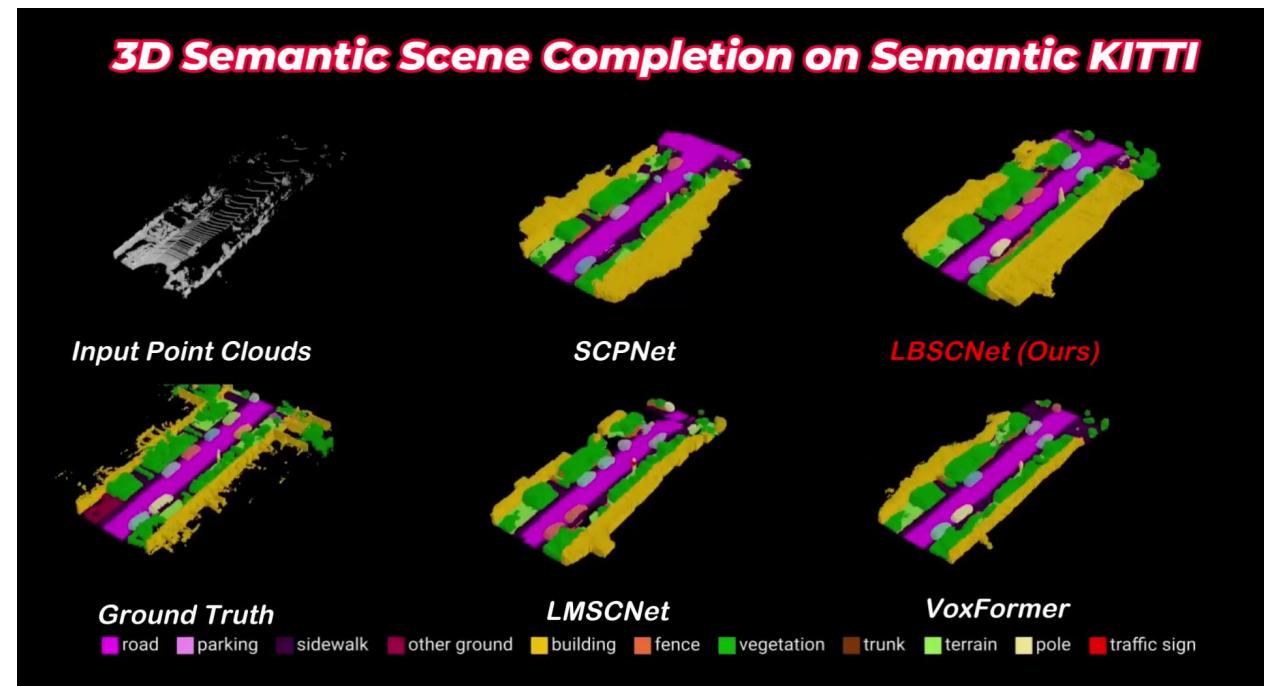
# LBSCNet Performance



Method	<i>IoU</i>	<i>mIoU</i>	<i>Prec.</i>	<i>Recall</i>	<i>FPS</i>
SSCNet [35]	53.20	14.55	59.13	<b>84.15</b>	12.00
LMSCNet [12]	55.32	17.01	77.11	66.19	13.50
LMSCNet-SS [12]	56.72	17.62	81.55	65.07	13.50
SCONet [3]	56.12	17.61	<b>85.02</b>	63.47	18.50
S3CNet [36]	45.60	29.50	48.79	77.13	1.20
Monoscene [9]	38.55	12.22	51.96	59.91	< 1
VoxFormer-T [10]	57.69	18.42	69.95	76.70	< 1
VoxFormer-S [10]	57.54	16.48	70.85	75.39	< 1
SCPNet [8]	56.10	<b>36.70</b>	72.43	78.61	< 1
<b>LBSCNet (Ours)</b>	<b>59.71</b>	23.58	77.60	71.29	<b>20.08</b>

ABLATION STUDY ON THE SEMANTICKITTI VALIDATION SET.

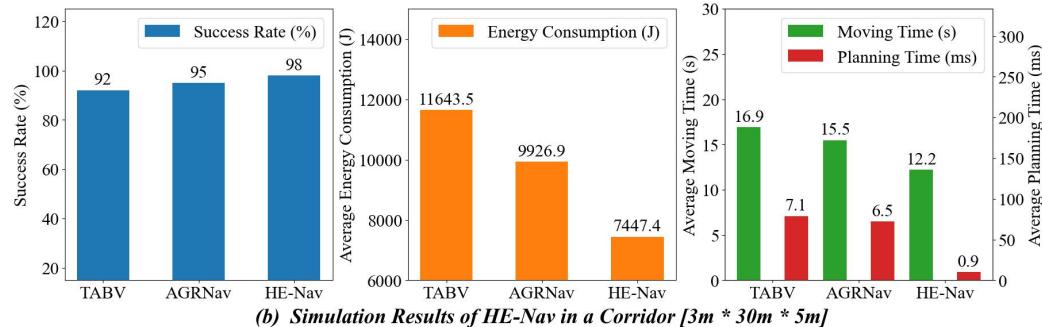
Method	<i>IoU</i> ↑	<i>mIoU</i> ↑
LBSCNet (ours)	58.34	22.74
w/o SCB-Fusion Module	57.05	21.26
w/o Criss-Cross Attention	57.20	22.17



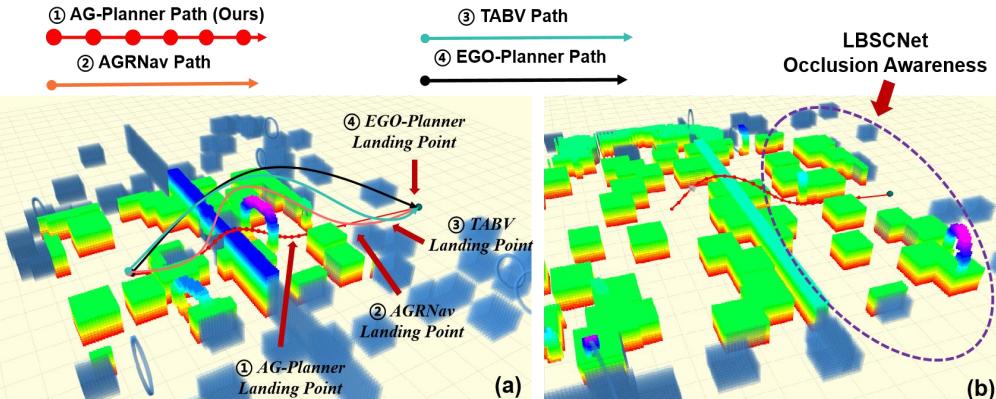
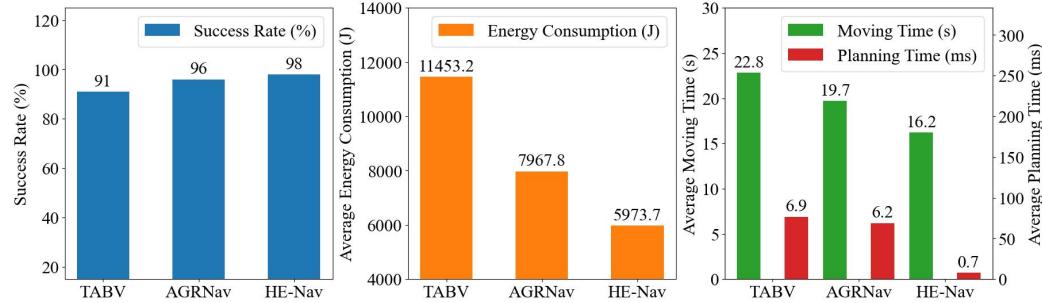
# End-to-End Performance

## ◆ Simulation Experiments:

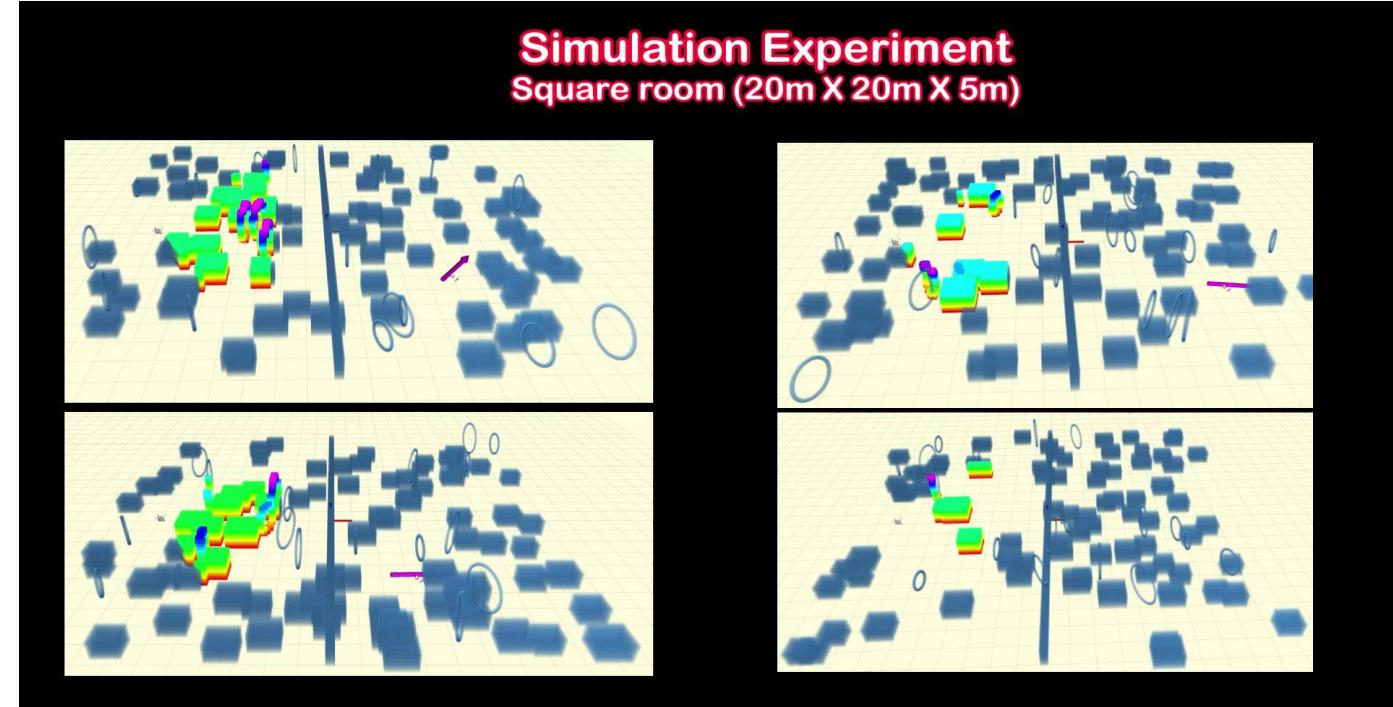
(a) Simulation Results of HE-Nav in a Square Room [20m \* 20m \* 5m]



(b) Simulation Results of HE-Nav in a Corridor [3m \* 30m \* 5m]



Simulation Experiment  
Square room (20m X 20m X 5m)

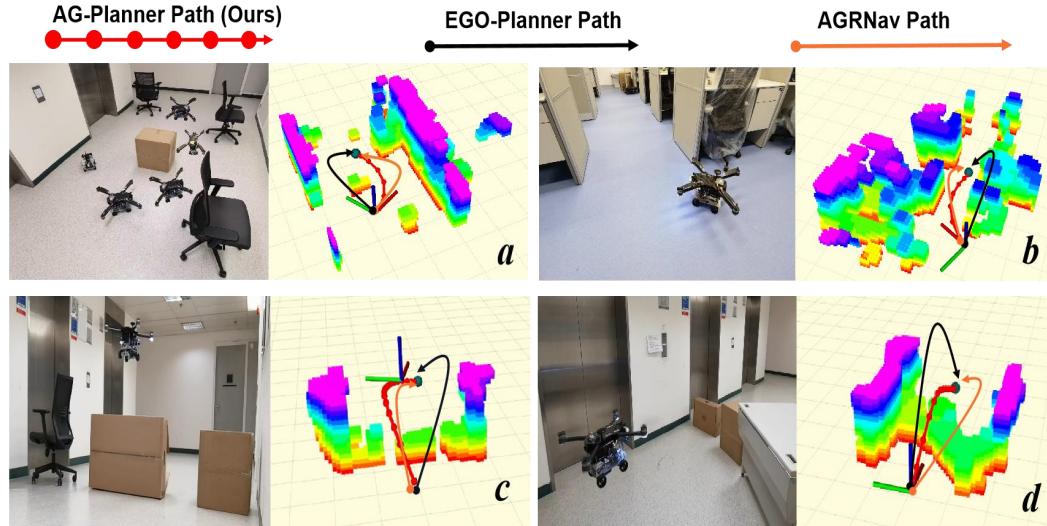


BATTERY AND ENERGY CONSUMPTION PARAMETERS

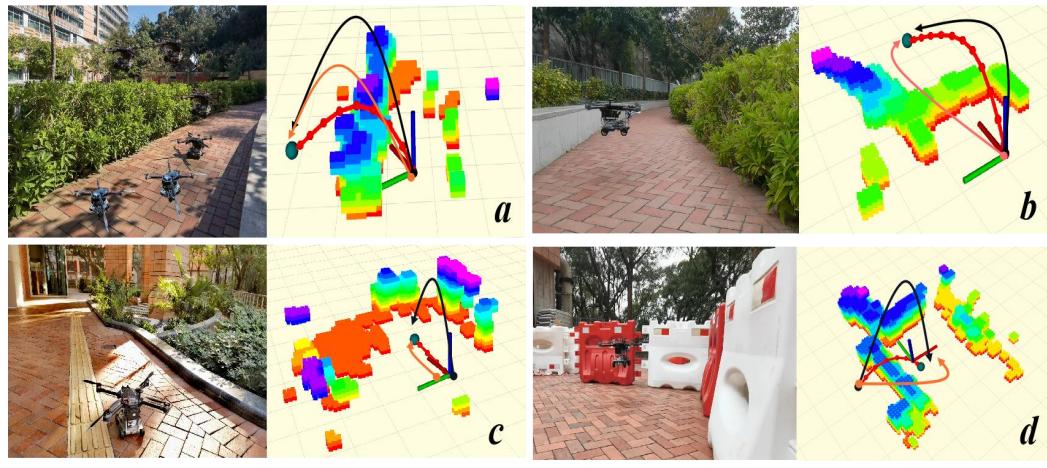
Parameter	Value
Battery Capacity	10000 mAh
Battery Weight	1008 g
Rated Power	231 Wh
Operating Voltage	23.05 V
Driving Energy Consumption	$\approx 251.45 \text{ J/s}$
Flying Energy Consumption	$\approx 988.33 \text{ J/s}$

# End-to-End Performance

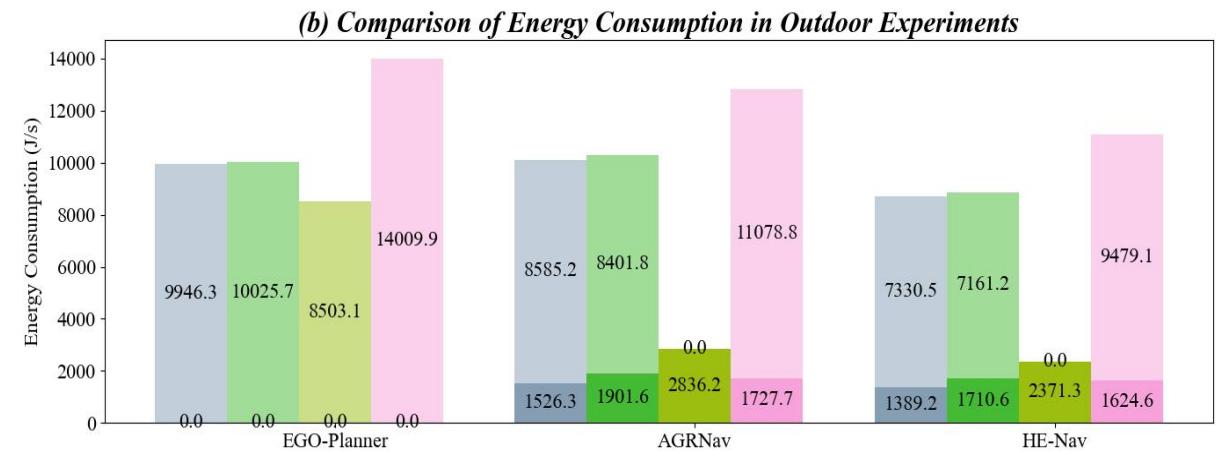
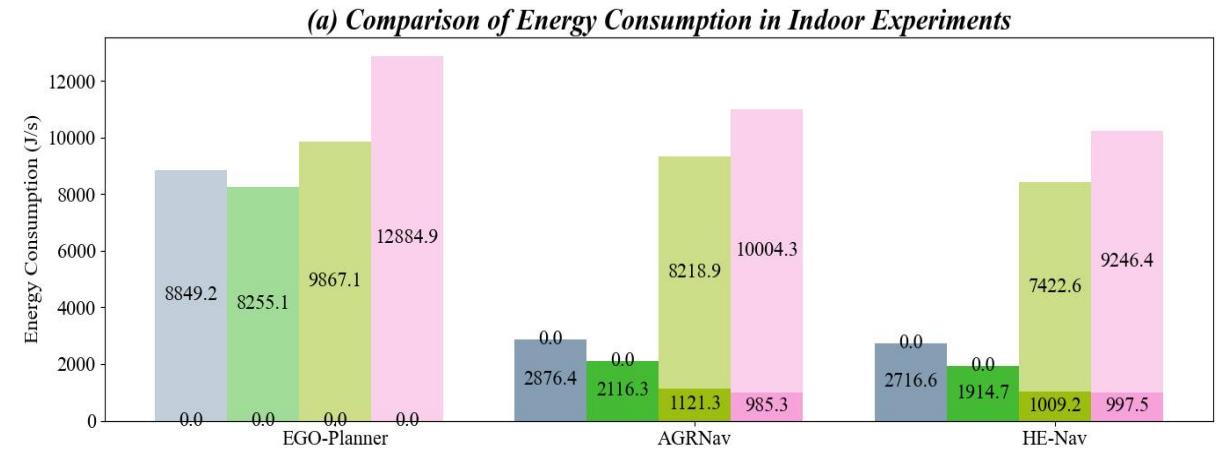
## ◆ Real-World Experiments:



(a) Indoor Real-World Experiments.

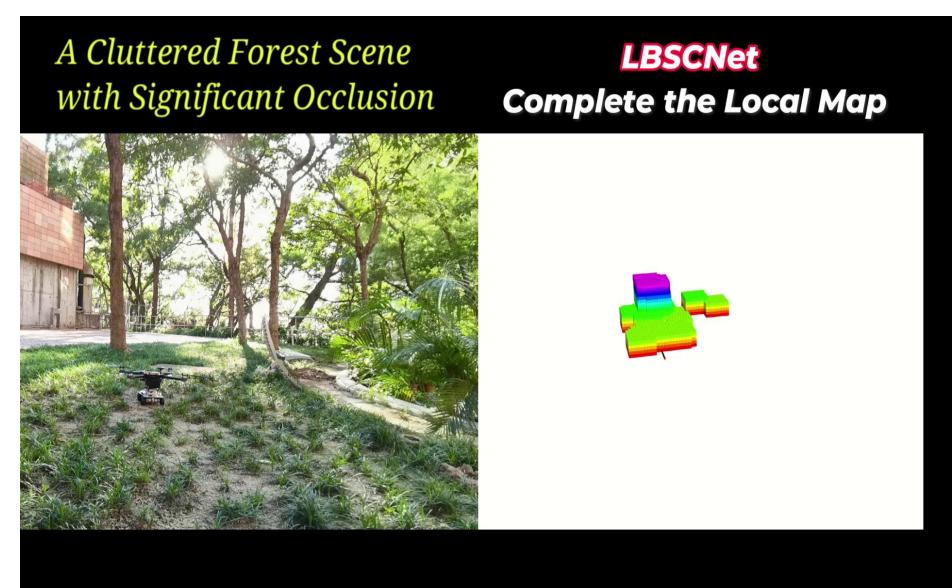
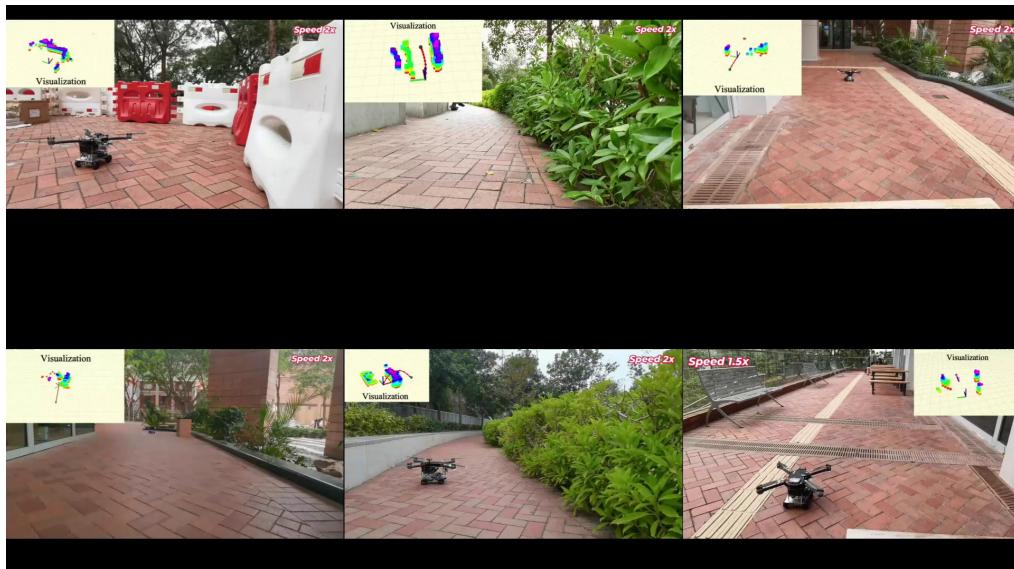


(b) Outdoor Real-World Experiments.



# End-to-End Performance

## ◆ Real-World Experiments:





# HE-Nav Conclusion

1. HE-Nav: The first high-performance, efficient, and ESDF-free navigation system specifically designed for aerial-ground robots (AGR)s.
2. Key Components:
  - LBSCNet: Lightweight BEV-guided semantic scene completion network.
  - AG-planner: Aerial-ground motion planner.
3. Capabilities:
  - Predicting obstacle distributions in occluded areas.
  - Generating low-collision risk, energy-efficient trajectories in real-time ( $\approx 1$  ms).
4. Extensive Evaluation:
  - Simulations and real experiments.
  - Significantly outperforms recent AGR navigation systems and multicopter path planners in performance and efficiency.

HE-Nav open source code link: <https://jmwang0117.github.io/HE-Nav/>



## ➤ First Author:

- [ICRA 2024] AGRNav: Efficient and Energy-Saving Autonomous Navigation for Air-Ground Robots in Occlusion-Prone Environments
  - ✓ ***The First AGR-Tailored Occlusion-Aware Navigation System***
- [RA-L Under Review] HE-Nav: A High-Performance and Efficient Navigation System for Aerial-Ground Robots in Cluttered Environments
  - ✓ ***The First AGR-Tailored ESDF-Free Navigation System***
- [RA-L Under Review] OMEGA: Efficient Occlusion-Aware Navigation for Air-Ground Robot in Dynamic Environments via State Space Model
  - ✓ ***The First AGR-Tailored Dynamic Navigation System***

## ➤ Other work I help with:

- [ICCC 2024] Prediction-based Hierarchical Reinforcement Learning for Robot Soccer
- [ApPLIED Workshop @PODC 2024] Hybrid-Parallel: Achieving High Performance and Energy Efficient Distributed Inference on Robots
- [AIML Workshop @COMPSAC 2023] New Problems in Active Sampling for Mobile Robotic Online Learning
- [Under Review] RRT0: A High-Performance Transparent Offloading System for Model Inference on Robotic IoT
- [Under Review] CacheInf: Collaborative Edge-Cloud Cache System for Efficient Robotic Visual Model Inference



# Conclusion and Future Work

- In this thesis, I present:

- AGRNav: The First AGR-Tailored, Occlusion-Aware Navigation System Enabling Efficient and Energy-Efficient Autonomous Navigation in Occlusion-Prone Environments.
- HE-Nav: The First AGR-Tailored, ESDF-Free Navigation System Achieving High-Performance and Efficient Navigation in Cluttered Environments.

- I have four ideas for future work:

Imitation learning for AGRs

Diffusion model for AGRs

LLMs for AGRs

DRL for AGRs



香港大學  
THE UNIVERSITY OF HONG KONG

# Thanks for listening and questions are welcome!

Name of Speaker: Junming Wang (MPhil Student)