Flying Guide Dog: Walkable Path Discovery for the Visually Impaired Utilizing Drones and Transformer-based Semantic Segmentation

Haobin Tan, Chang Chen, Xinyu Luo, Jiaming Zhang, Constantin Seibold, Kailun Yang, Rainer Stiefelhagen
Karlsruhe Institute of Technology

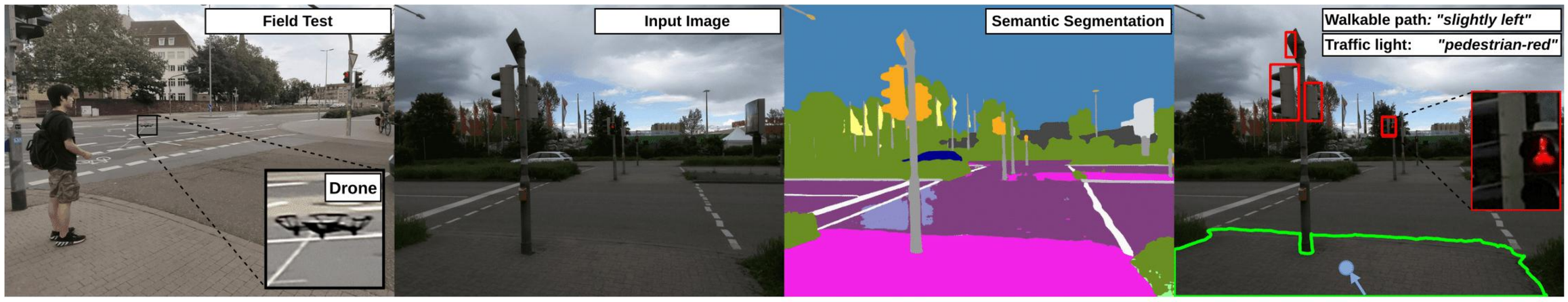

Fig. 1: Flying guide dog prototype

Motivation

- Assisting tools for blind and visually impaired people (BVIP)
- Training a guide dog is time- and money-costly

Methodology

- Novel "flying guide dog" prototype utilizing drone and Visual Transformer-based street view semantic segmentation
- Drone control algorithm for automatically discovering walkable areas and avoiding obstacles
- Pedestrian traffic light recognition
- Interactive voice feedback for supplementary assistance

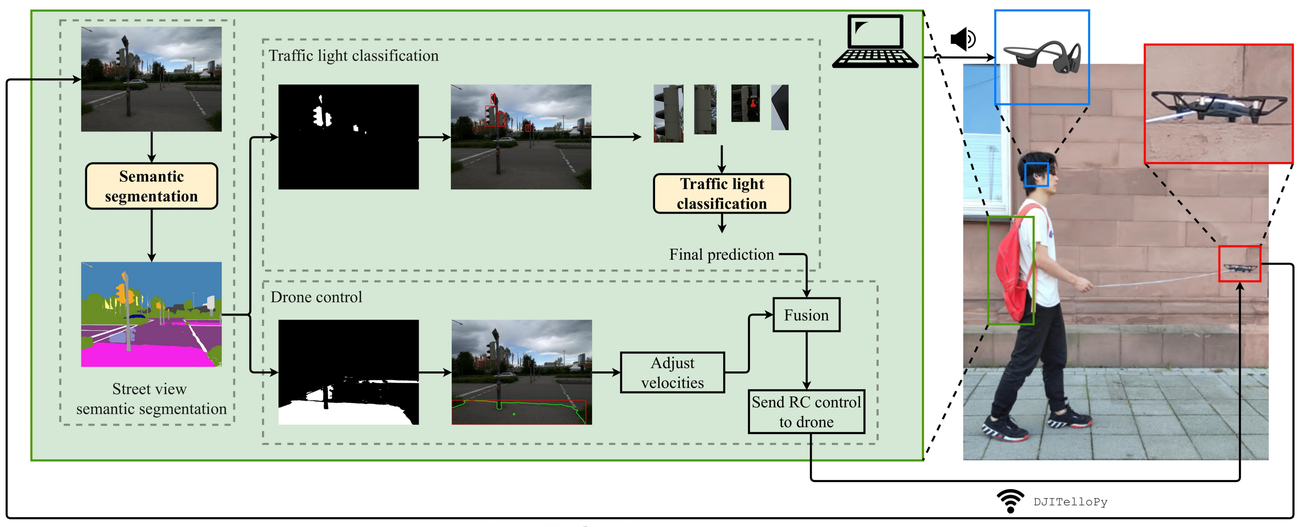

Fig. 2: System overview

Pedestrian and Vehicle Traffic Lights (PVTL) Dataset

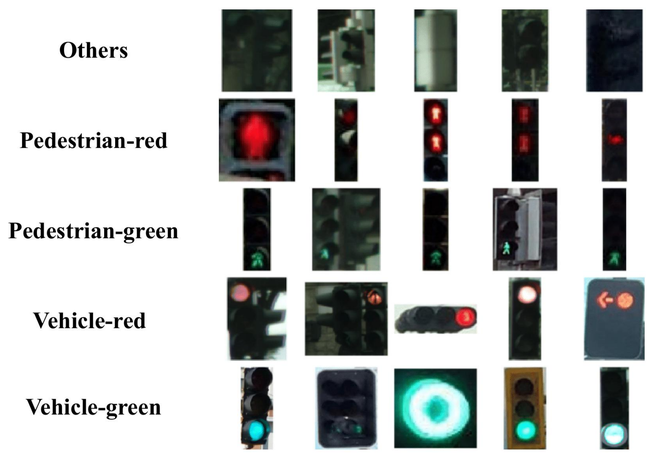

Fig. 3: PVTL dataset

Real-world Scenes User Study

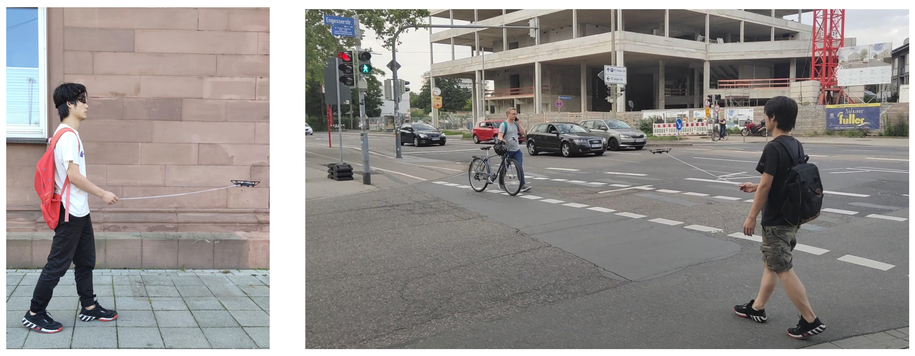

Fig. 5: User study scenarios

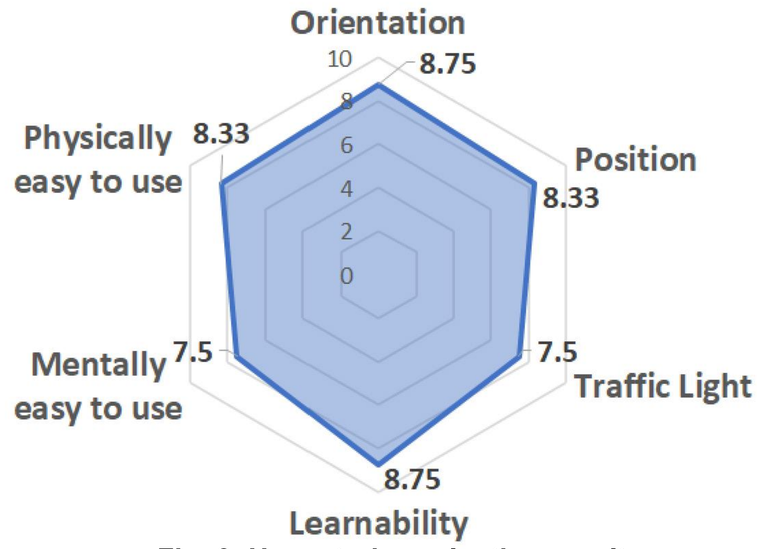

Fig. 6: User study evaluation results

Results

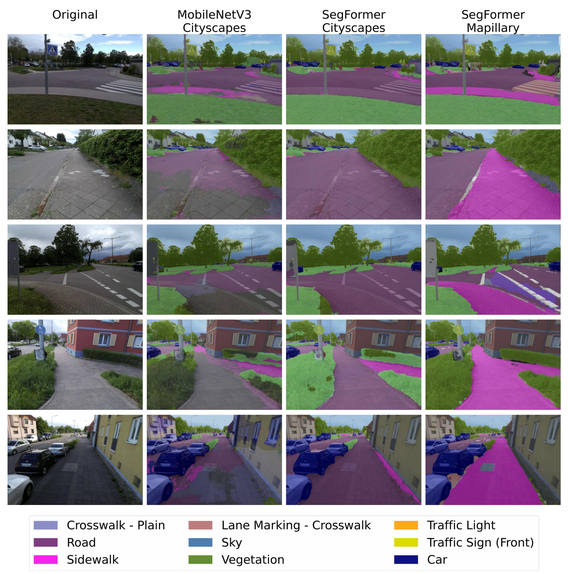

Fig. 4: Qualitative results

Category	IoU
Bike Lane	31.87
Crosswalk - Plain	21.09
Curb	53.20
Pedestrian Area	31.24
Road	85.20
Sidewalk	62.32
Lane Marking - Crosswalk	63.80
Traffic Light	61.06
Traffic Sign (Front)	68.67
Car	86.84

Tab. 1: Per-class IOU on Mapillary Vistas

	Method	FPS ↑	mIoU (%) ↑	Params (M) ↓
Non Real-Time	BiSeNet (R101) [23]	9.3	20.4	50.1
	TASCNet (R50) [24]	**11.9**	46.4	**32.8**
	TASCNet (R101) [24]	8.8	48.8	51.8
	ShelfNet (R101) [25]	9.1	49.2	57.7
	RGPNet (R101) [22]	10.8	50.2	52.2
	RGPNetB(WRN38) [22]	3.4	**53.1**	215.0
Real-Time	RGPNet (HarDNet39D) [22]	34.7	**42.5**	9.4
	RGPNet (R18) [22]	**35.7**	41.7	17.8
	SegFormer-B0 (ours)	15.2	41.8	**3.8**

Tab. 2: Comparison of SOTA methods on Mapillary Vistas

Future Work

- Combine more portable AI computer (*e.g.* Jetson AGX Xavier)
- Use drone with larger battery capacity