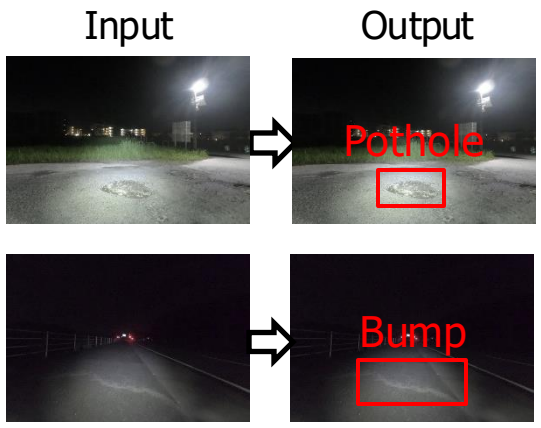
CLIP-Guided Cross-Modal Feature Fusion and Domain-Adaptive Feature Aggregation based Few-Shot Learning for Night Pavement Defect Detection

沓 崇崇 池永研究室 修士課程修了

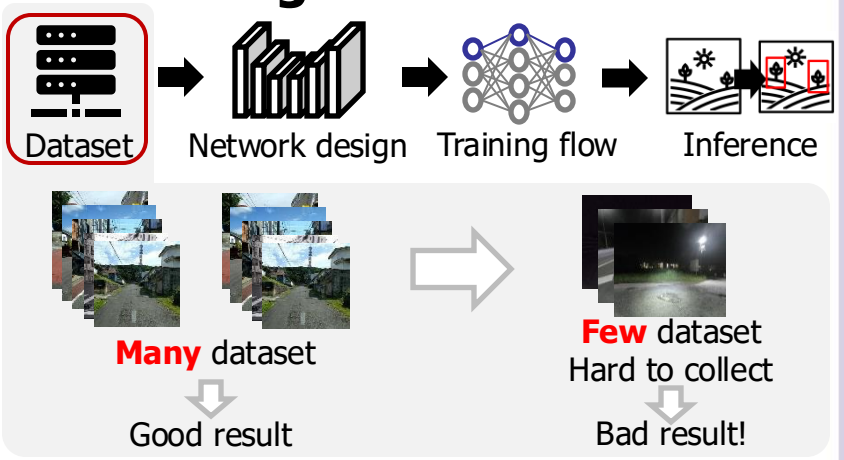
Background

Application

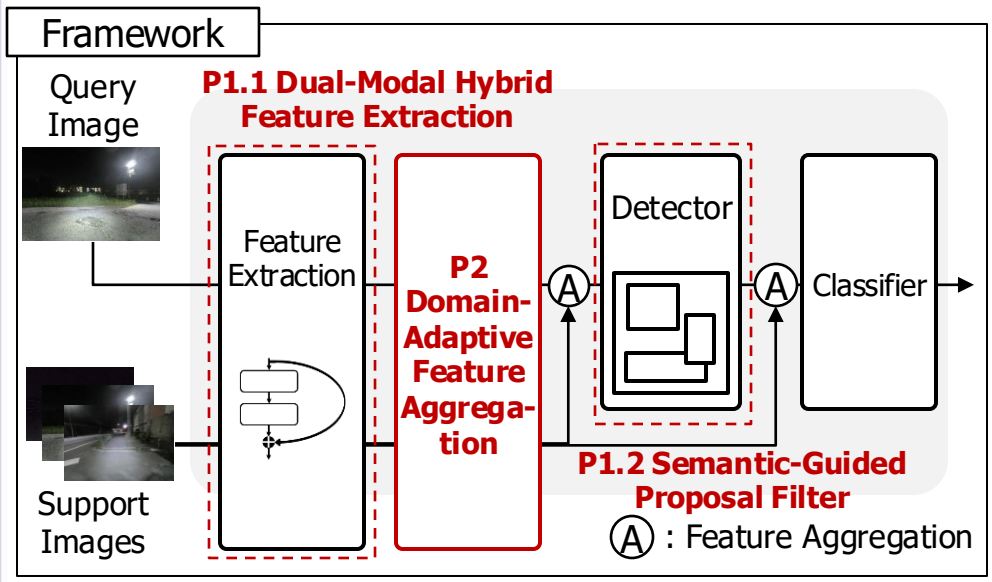
- Traffic safety
- Road maintenance
- Vehicle quality
- .....

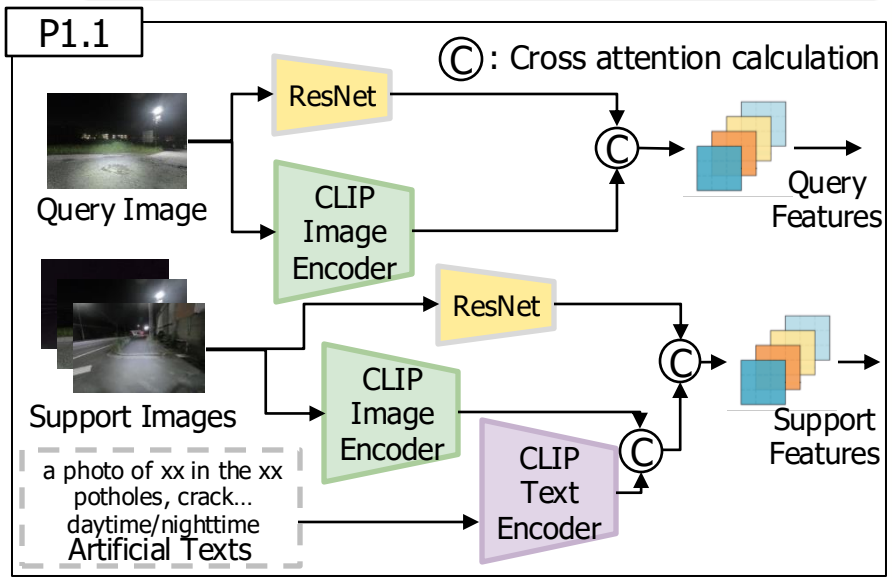


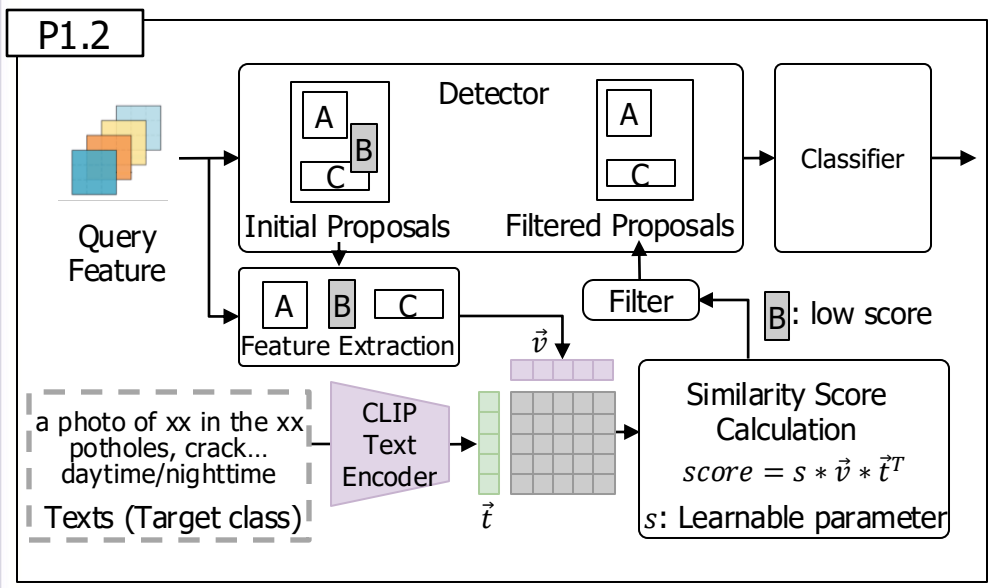
Challenges

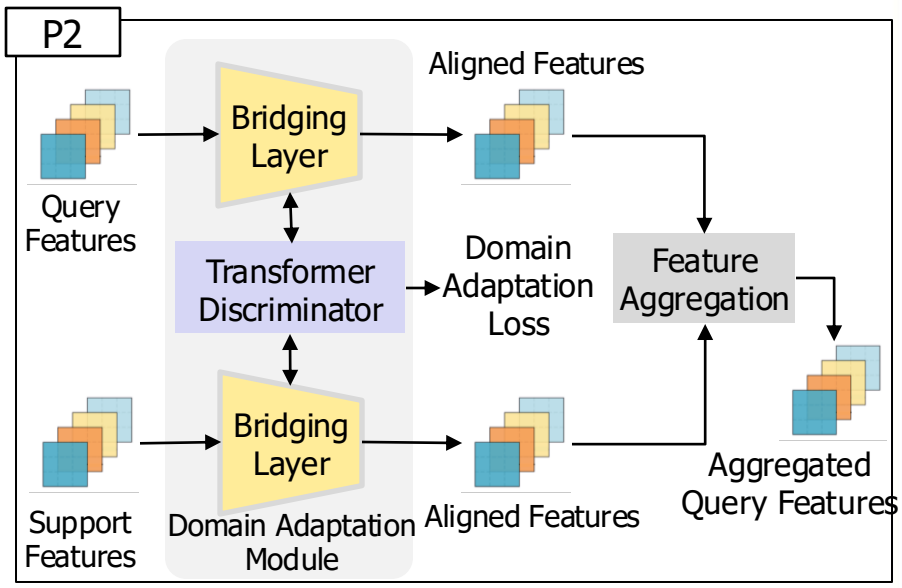


Proposals









Experiment Results

Method	1shot			3shot			5shot			10shot		
	mAP	Prec.	Rec.	mAP	Prec.	Rec.	mAP	Prec.	Rec.	mAP	Prec.	Rec.
Baseline	9.1	10.5	24.5	19.2	18.2	41.5	23.4	22.2	48.1	29.9	34.7	54.6
P1	18.7	12.5	27.4	33.2	22.9	48.9	39.1	27.3	54.9	45.1	38.4	59.8
P1+P2	26.2	14.6	22.8	41.8	27.5	52.7	49.6	35.3	60.2	55.6	41.5	62.1

Conclusion

- Evaluation results show that P1+P2 achieves mAP 55.6% at 10-shot, outperforming the baseline by 25.7%

Graduate School of Information, Production and Systems
Waseda University