3rd Place Solution to NAVER LABS Mapping & Localization Challenge 2020: Outdoor Track

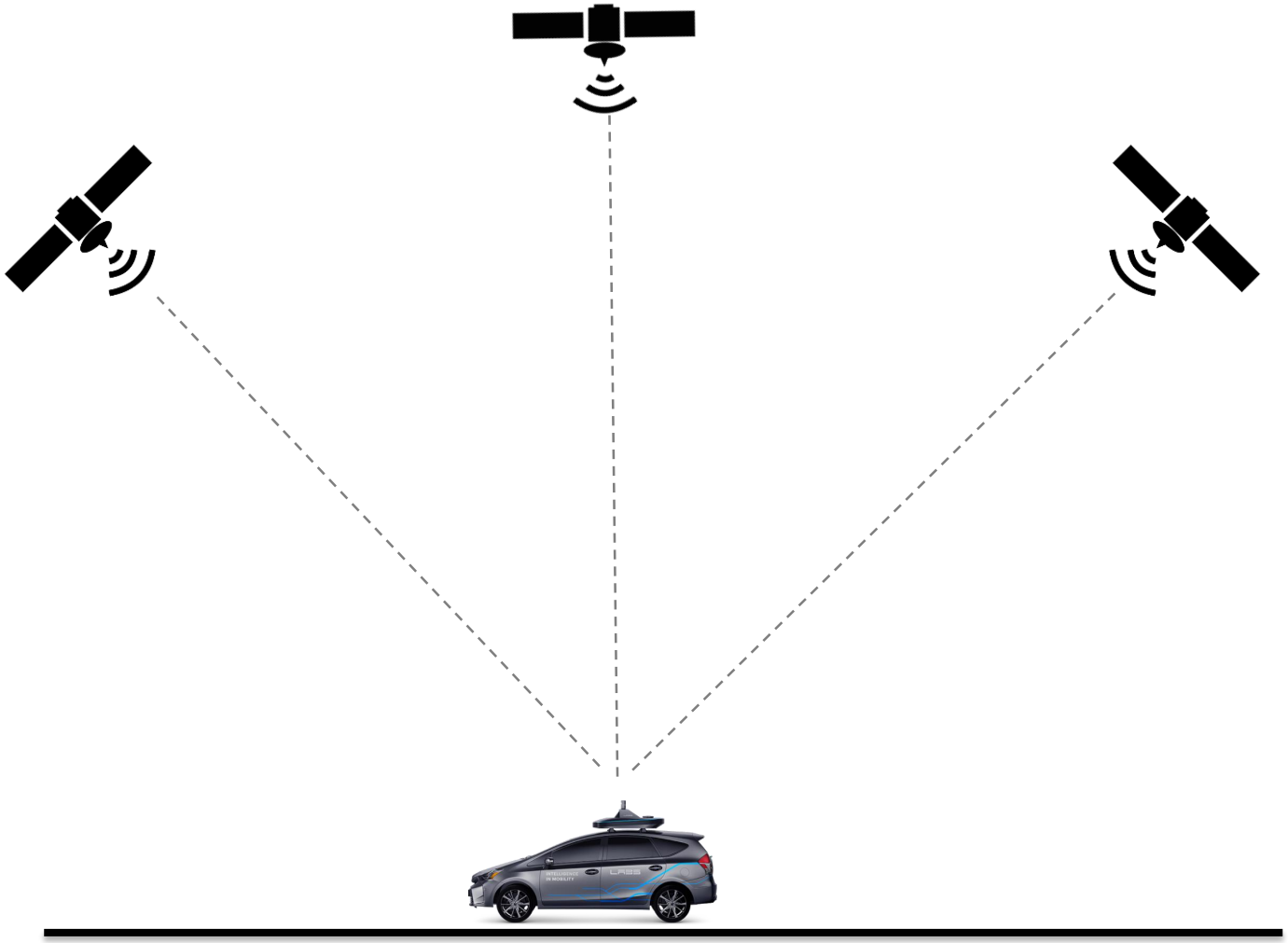
세종대학교 Robotics and Computer Vision(RCV) 연구실

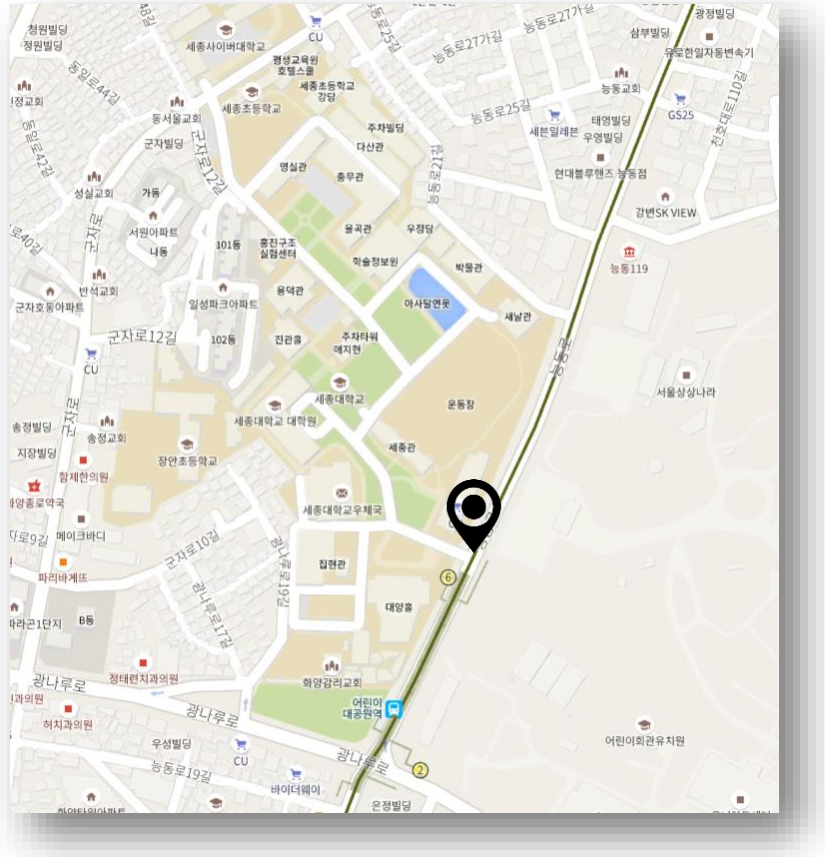
김지원*, 김태주*, 황유진*, 최유경†

세종대학교 SEJONG UNIVERSITY

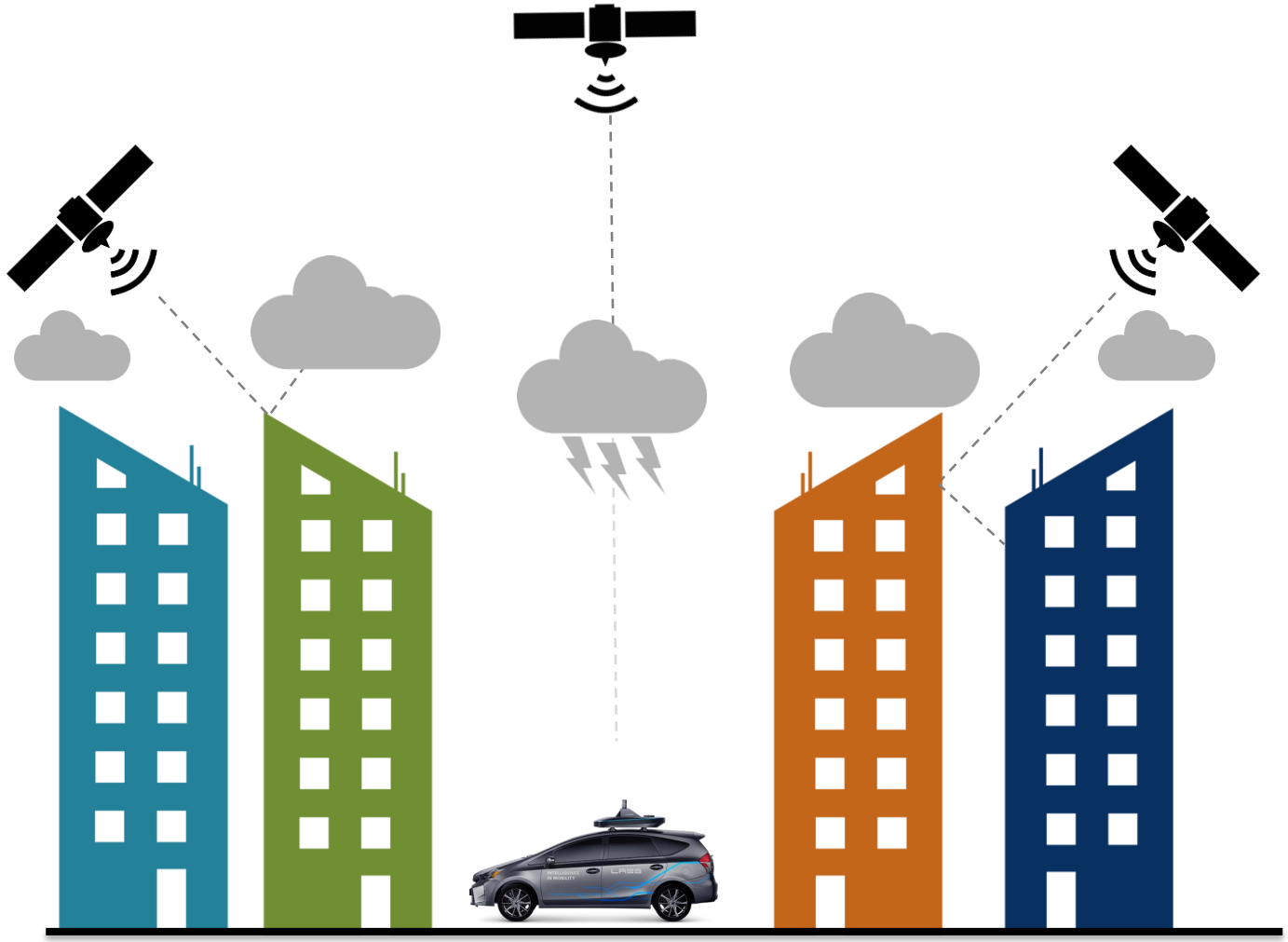
Sejong RCV

Visual Mapping & Localization

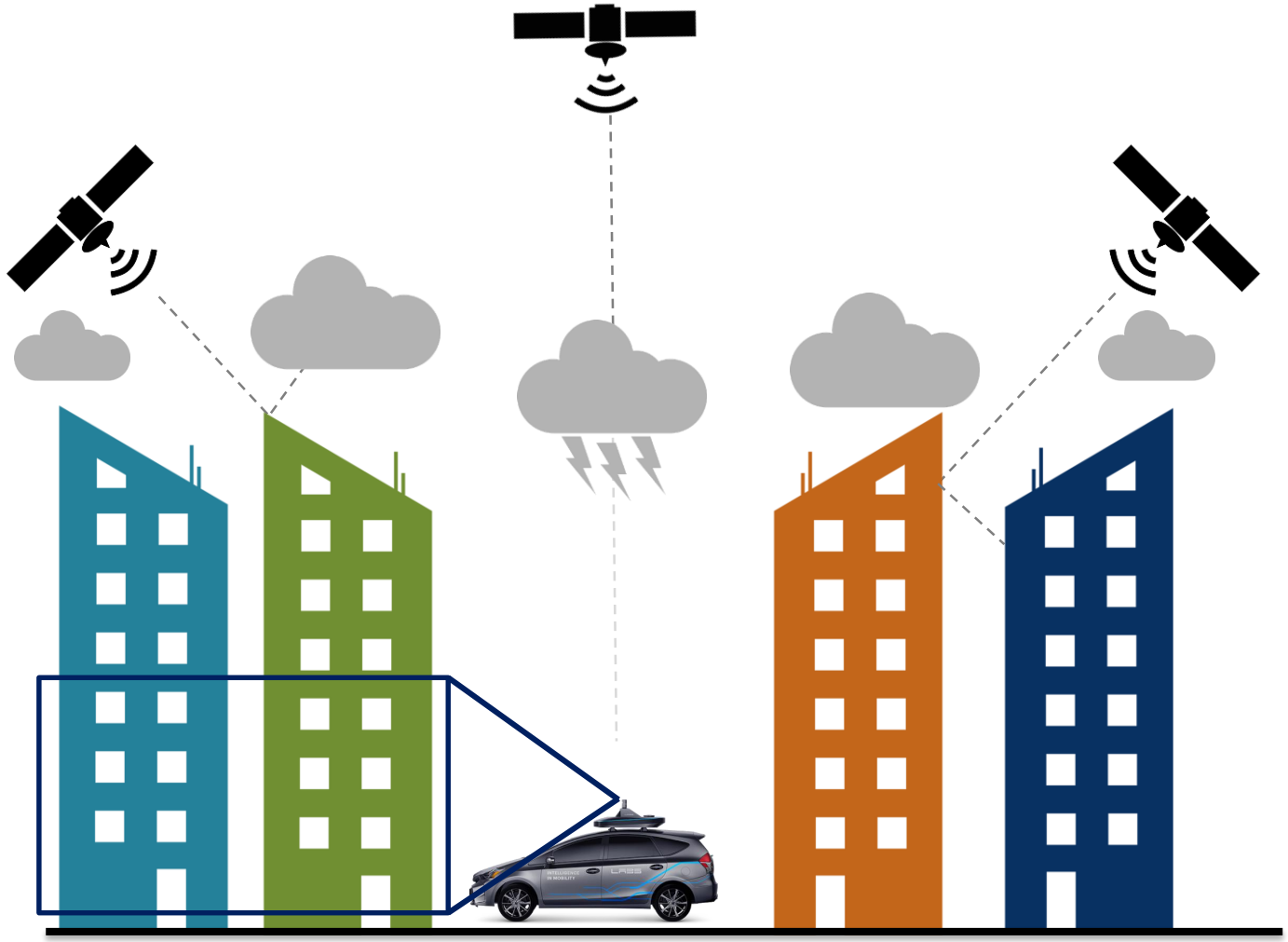


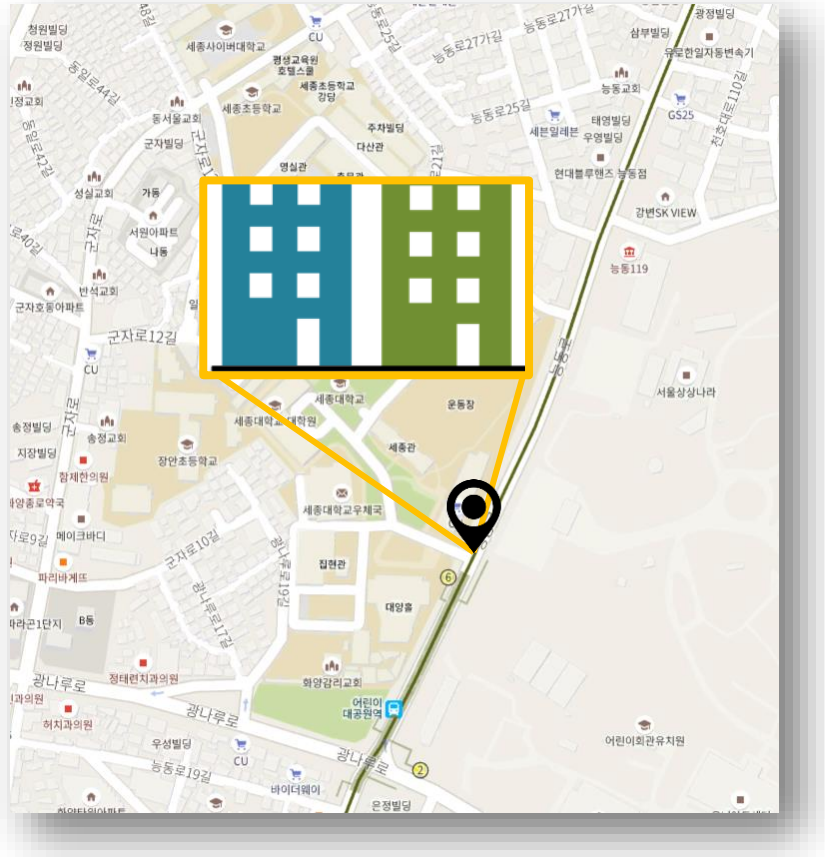


Visual Mapping & Localization



Visual Mapping & Localization



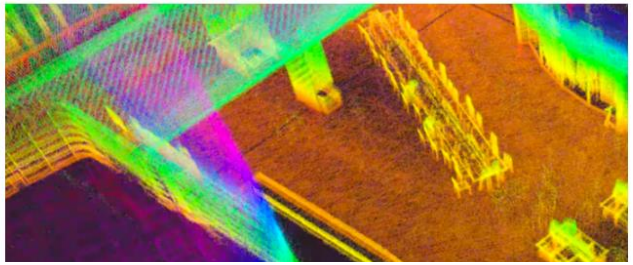


NAVER LABS Mapping & Localization Challenge

NAVER LABS
Mapping & Localization Challenge

2020
4/8 - 8/18

네이버랩스는 도심의 각 공간을 다양한 자율주행 머신들로 연결하고, 공간 데이터를 정보화하여 공간 간 이동이 자동화되는 도시(A-CITY)를 위한 기술을 연구합니다. 이 기술을 실현하기 위한 첫 번째 단계는 매핑 장비로 '고정밀 지도'를 만드는 것입니다. 정보화된 공간에서 자율주행 머신의 위치를 정밀하게 파악할 수 있는 '고정밀 측위 기술' 또한 필요합니다. 네이버랩스는 사진 한 장만으로 위치와 포즈를 정확히 파악할 수 있는 Visual Localization 기술 연구에 집중하고 있습니다. 네이버랩스는 Mapping & Localization 분야 연구가 더욱 활발해질 수 있도록 그동안 만들어 온 실내외 데이터셋을 국내 대학 연구자들에게 공개하고, 서로의 Visual Localization 기술을 겨룰 수 있는 챌린지를 준비하였습니다.



Indoor Top3

1st
2nd
3rd SejongRCV-Indoor

Outdoor Top3

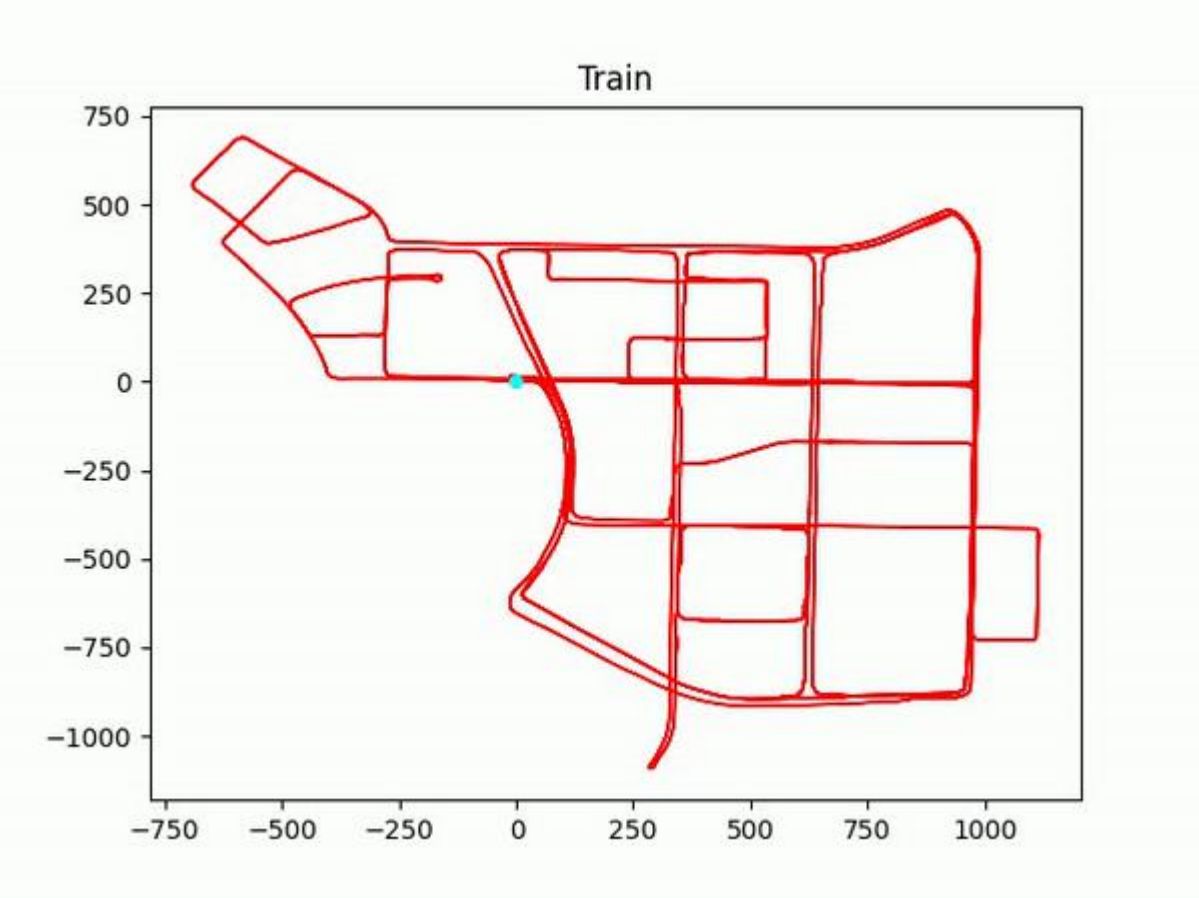
1st
2nd
3rd SejongRCV-Outdoor

참가자격

- 참가 대상은 국내 대학교에서 관련 분야 연구 중인 내국인 대학생 및 대학원생으로 한정합니다.
- 참가팀은 최소 1명에서 최대 3명으로 구성할 수 있습니다.

NAVER LABS Mapping & Localization Challenge

주어진 판교/여의도 아웃도어 데이터셋을 Train(90%)/Validation(10%) 으로 나누어 평가



	판교	여의도
Total	48107	86275
DB	43297	77647
Query	4810	8628

NAVER LABS Mapping & Localization Challenge

Place Recognition

DB

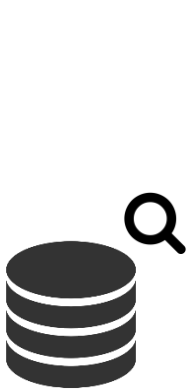
Query

NetVLAD

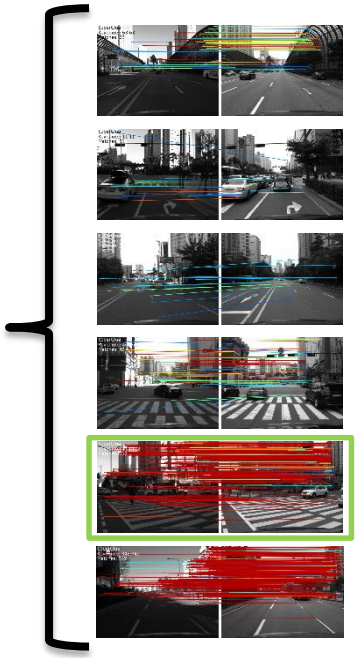
Global Descriptor

NetVLAD

Global Descriptor

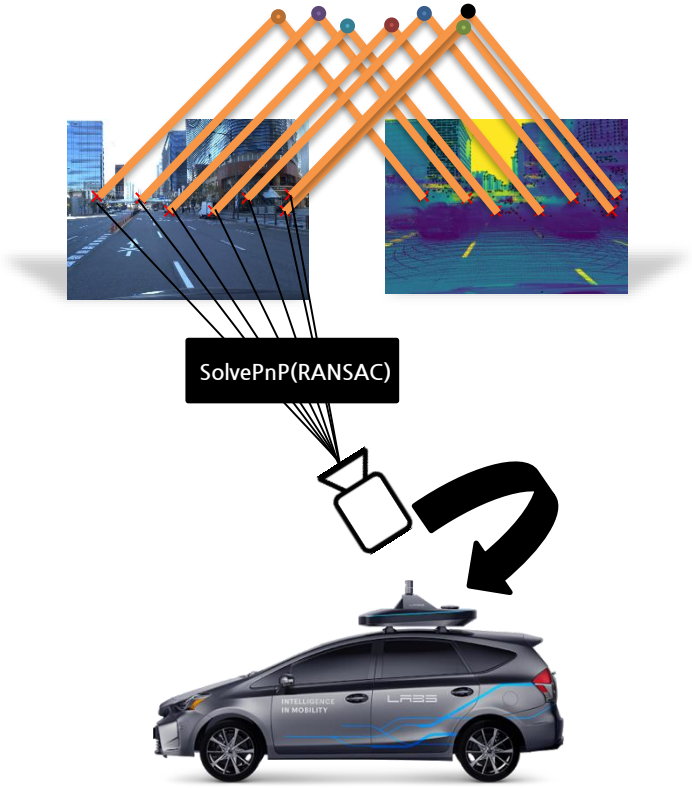


NN Search



Re-ranking

Pose Estimation



NAVER LABS Mapping & Localization Challenge

Place Recognition

DB

NetVLAD

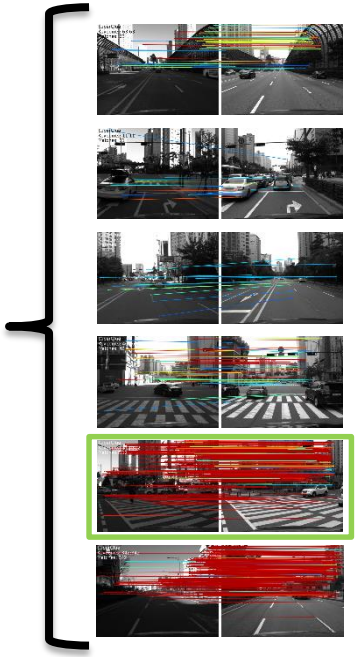
Global Descriptor

Query

NetVLAD

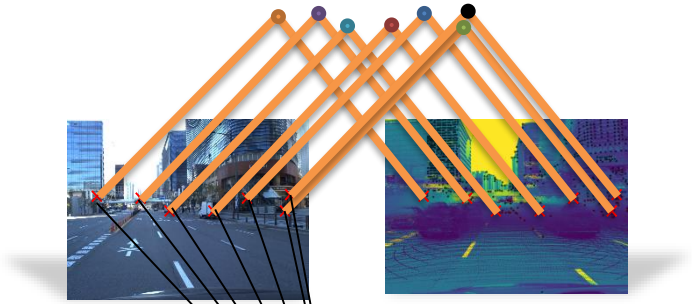
Global Descriptor

NN Search



Re-ranking

Pose Estimation

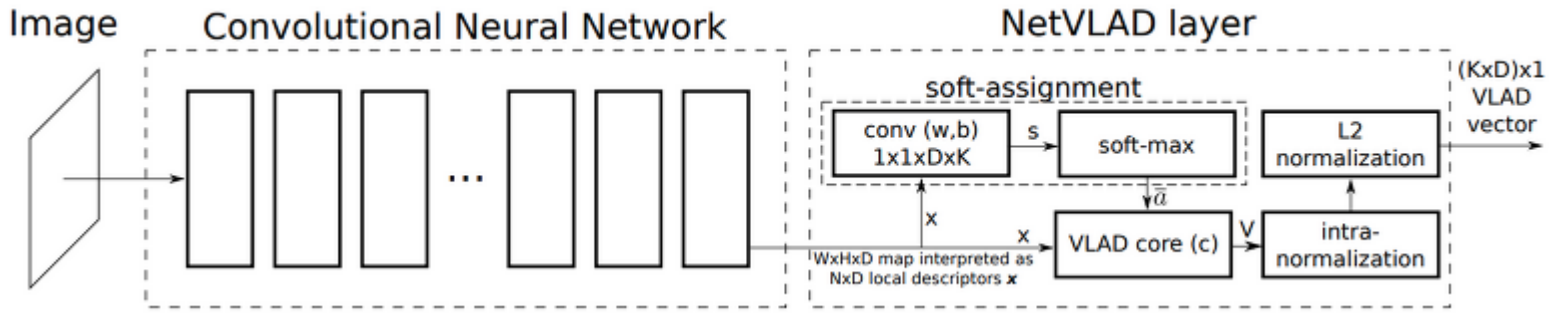


SolvePnP(RANSAC)

Place Recognition

Place Recognition

Global Descriptor **NetVLAD** : CNN architecture for weakly supervised place recognition For Top-10



(a) Mobile phone query

(b) Retrieved image of same place

[1] NetVLAD: CNN architecture for weakly supervised place recognition

Place Recognition

Global Descriptor **NetVLAD** : CNN architecture for weakly supervised place recognition For Top-10

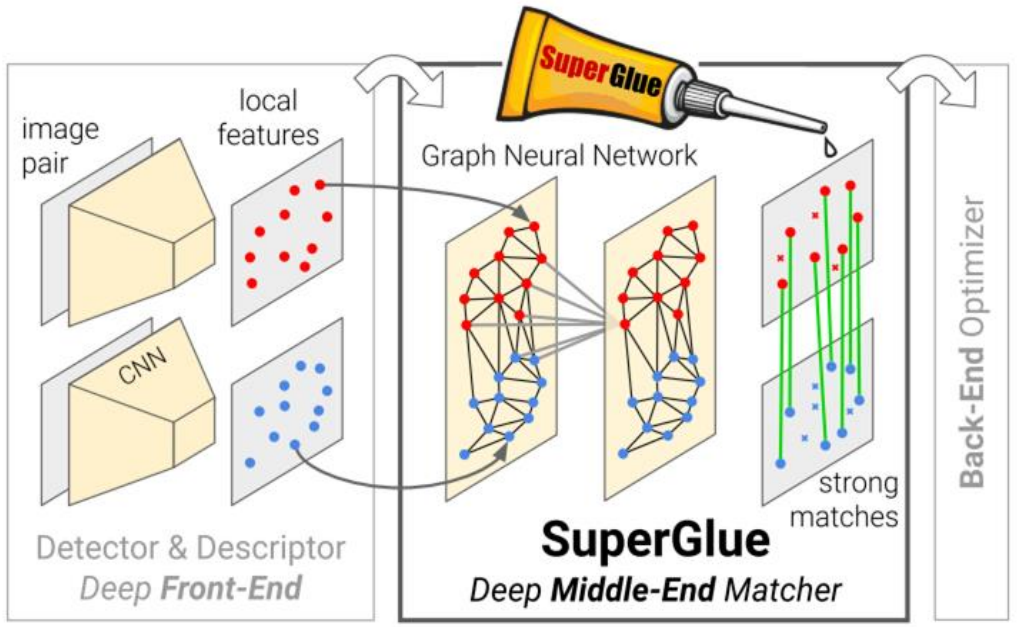
NetVLAD로 추출한 DB(Train)의 GT를 기준으로 계산한 QR(Validation) Place Recognition의 성능

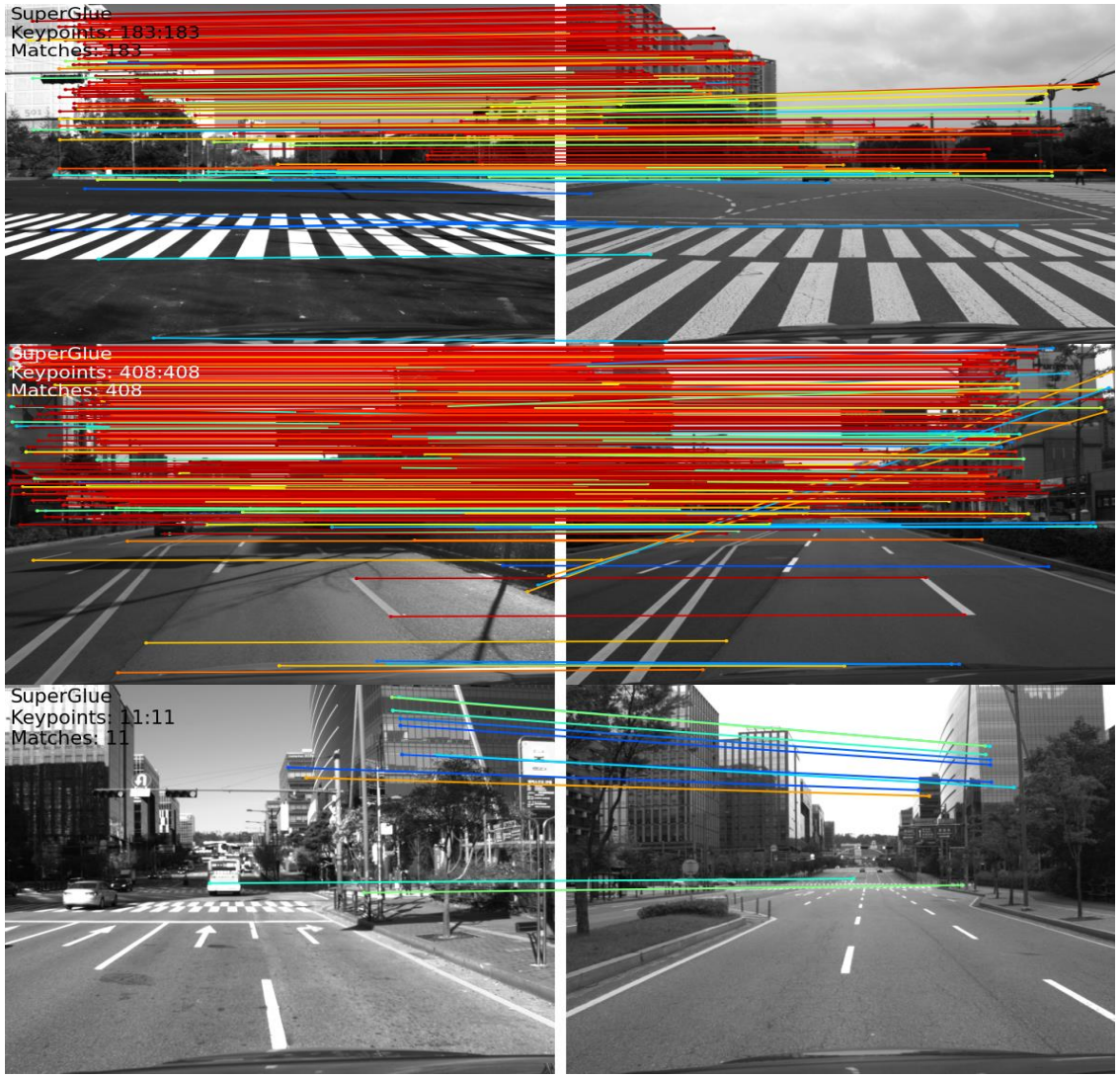
Dataset	Metric	recall@1	recall@5	recall@10
판교	0.5m / 2.0°	28.42	34.38	35.7
	1.0m / 5.0°	96.404	97.83	97.88
	5.0m / 10.0°	100.0	100.0	100.0
여의도	0.5m / 2.0°	32.08	36.64	39.88
	1.0m / 5.0°	86.6	89.11	90.16
	5.0m / 10.0°	100.0	100.0	100.0

[1] NetVLAD: CNN architecture for weakly supervised place recognition

Place Recognition

SuperPoint & SuperGlue





[2] SuperPoint: Self-Supervised Interest Point Detection and Description
[3] SuperGlue: Learning Feature Matching with Graph Neural Networks

Place Recognition : Result

Place Recognition : Result

NAVER LABS Mapping & Localization Challenge

Place Recognition

DB

NetVLAD

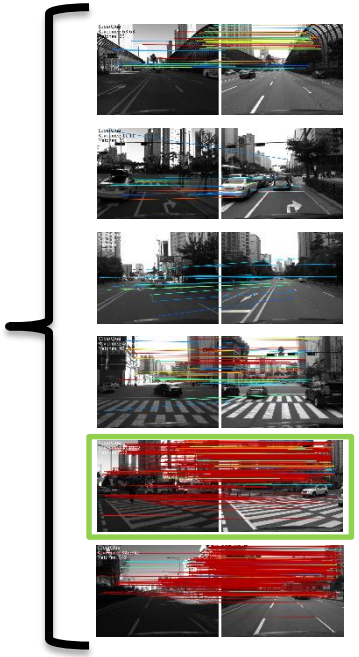
Global Descriptor

Query

NetVLAD

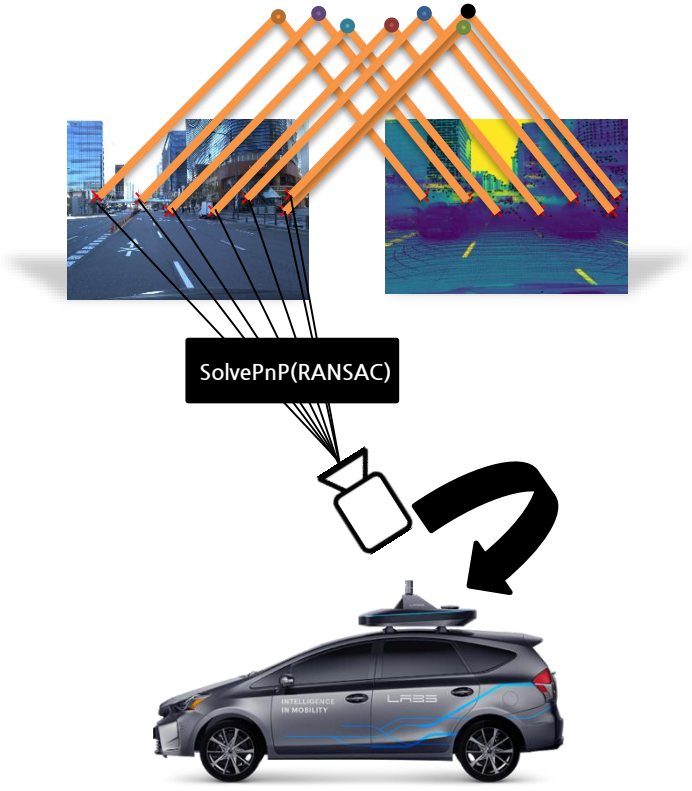
Global Descriptor

NN Search



Re-ranking

Pose Estimation

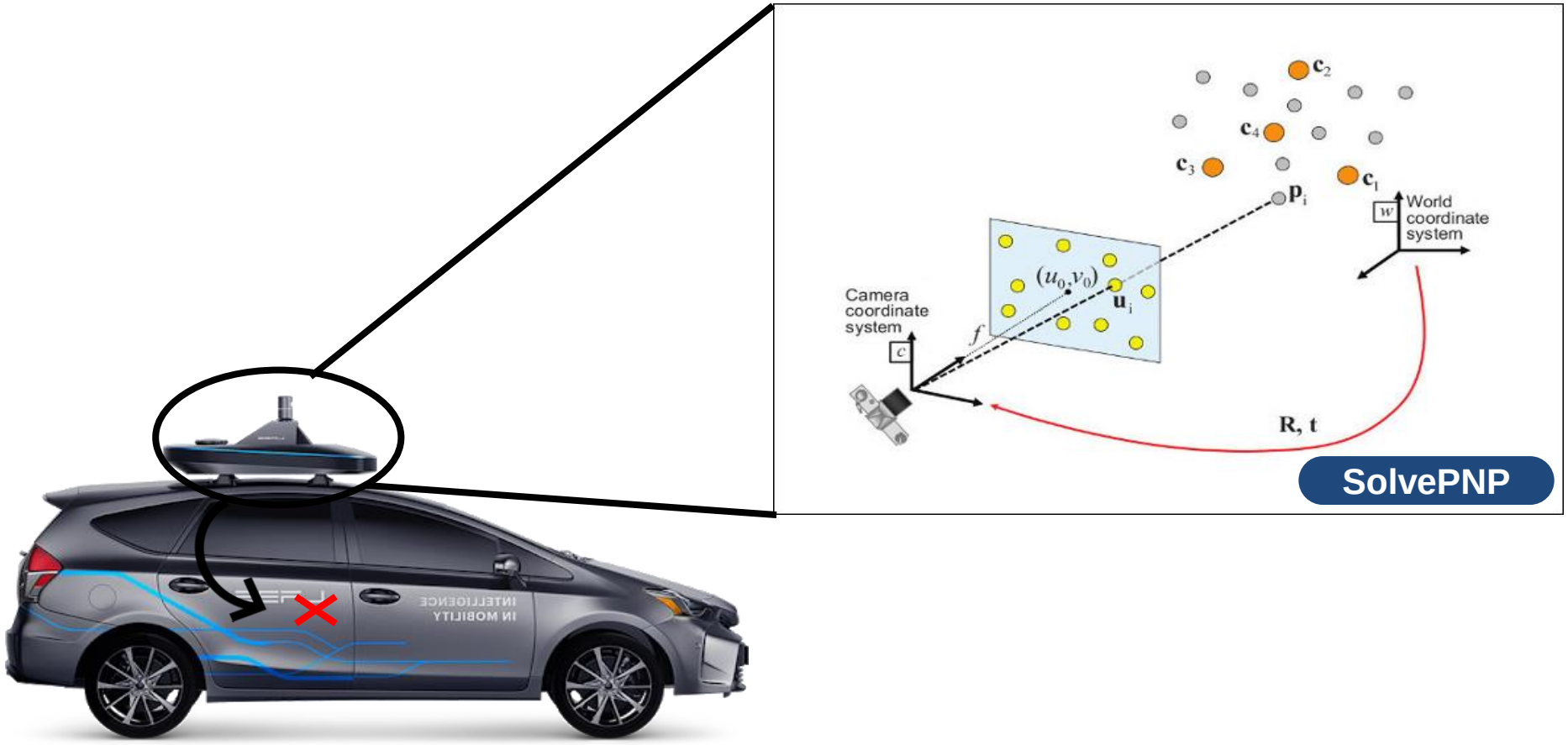


Pose Estimation

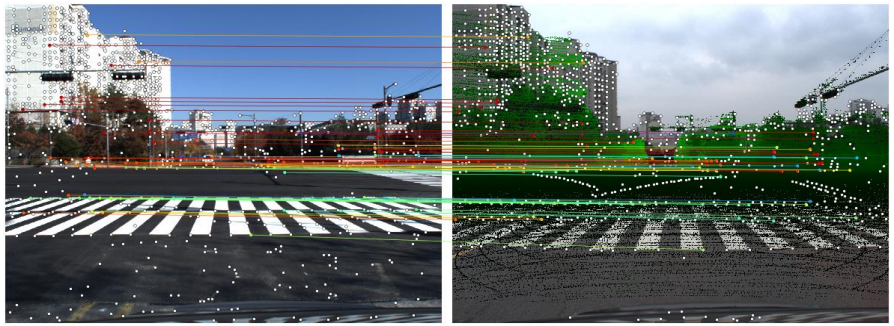
Pose Estimation

PnP algorithm

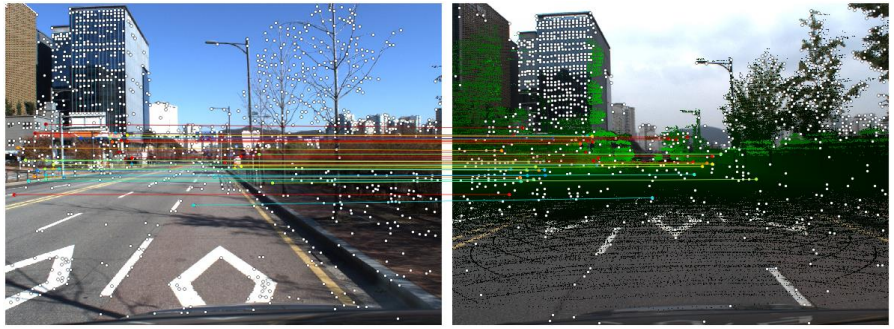
SolvePnP : Estimate the orientation of a 3D object in a 2D image



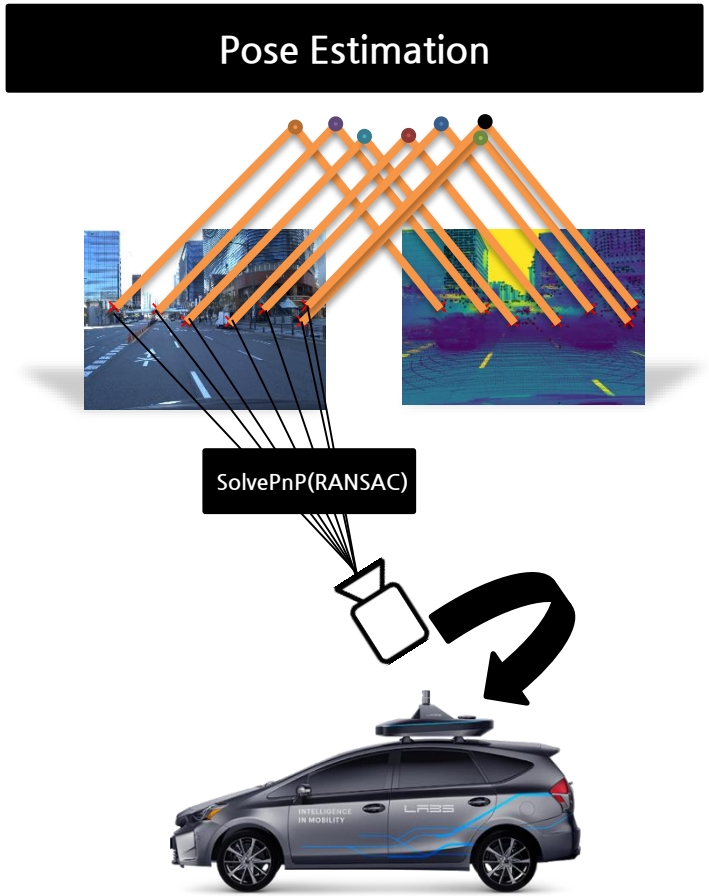
Pose Estimation



QR DB



QR DB



투영(Projection)한 DB 이미지에서 QR-DB 매칭 키포인트(2D)에 대응되는 3D 라이다 포인트 파싱

2D-3D 대응점으로 PnP 알고리즘 수행하여 QR의 Pose를 계산

Pose Estimation

PnP algorithm

SolvePnP : Estimate the orientation of a 3D object in a 2D image

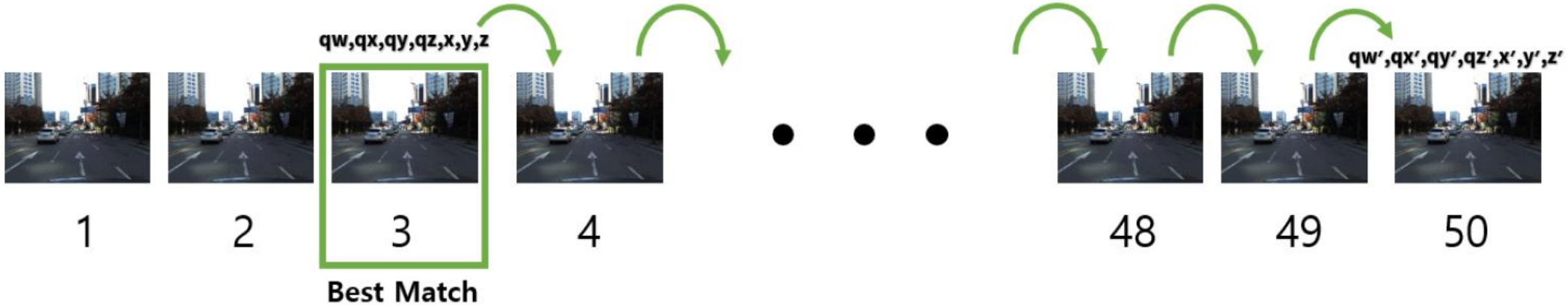
Dataset	Metric	recall@1
판교	0.5m / 2.0°	28.42
	1.0m / 5.0°	96.404
	5.0m / 10.0°	100.0
여의도	0.5m / 2.0°	32.08
	1.0m / 5.0°	86.6
	5.0m / 10.0°	100.0

Dataset	Metric	recall@1
판교	0.5m / 2.0°	99.792
	1.0m / 5.0°	100.0
	5.0 / 10.0°	100.0
여의도	0.5m / 2.0°	98.71
	1.0m / 5.0°	99.51
	5.0 / 10.0°	100.0

PnP 알고리즘을 통해서 추정한 Pose는 0.5m 이내에서 99%의 성능을 나타냄

Visual Odometry

Visual Odometry



키포인트 매칭이 가장 많은 QR에서 Pose를 예측하고
이후 Visual Odometry를 통해 최종 QR의 Pose를 예측

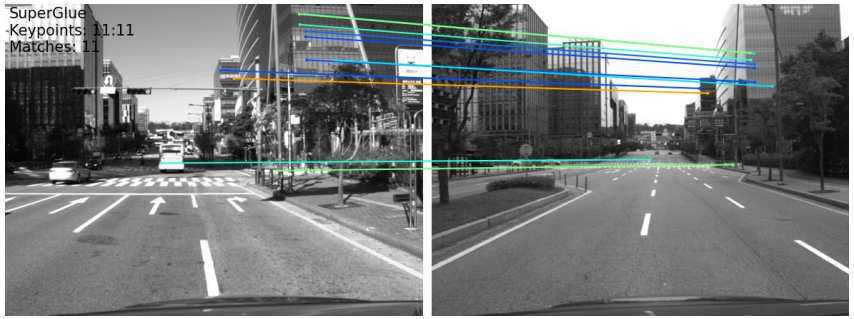
Visual Odometry

1번째 QR

40번째 QR

41번째 QR

50번째 QR



Best Match

Conclusion

Conclusion

Dataset	Team	0.5m / 2.0°	1.0m / 5.0°	5.0m / 10.0°
판교	**Our**	**44.00%**	**58.00%**	**68.00%**
	Baseline	24.00%	42.00%	52.00%
여의도	**Our**	**82.00%**	**86.00%**	**92.00%**
	Baseline	30.00%	54.00%	72.00%

***Baseline : R2D2(tuned) + NetVLAD**

Thank you for attention


Jiwon Kim,


TaeJoo Kim,


Yujin Hwang,


Yukyung Choi

jwkim, yjhwang, tjkim, ykchoi@rcv.sejong.ac.kr

GitHub

Visit to our GitHub! : https://github.com/sejong-rcv/SejongRCV-Outdoor

Reference

[1] NetVLAD: CNN architecture for weakly supervised place recognition https://arxiv.org/abs/1511.07247

[2] SuperPoint: Self-Supervised Interest Point Detection and Description https://arxiv.org/abs/1712.07629

[3] SuperGlue: Learning Feature Matching with Graph Neural Networks https://arxiv.org/abs/1911.11763

[4] Opencv - Real Time pose estimation of a textured object https://docs.opencv.org/master/dc/d2c/tutorial_real_time_pose.html