ZEKI

Zentrum für erlebbare
Künstliche Intelligenz
und Digitalisierung e.V.

2023 Internship Group Presentations

Berlin, 31.08.2023

Topic Overview / Agenda

BeIntelli

1. TP2: AI Middleware - Edge Cloud
2. TP3: Extrinsic Camera and Lidar Calibration
3. TP3: Localization by Visual Place Recognition
4. TP3: Free Parking Space Detection
5. TP4: Smart Parking Recommendation
6. TP4: AI Platform and Tools

COBRA-5G

1. Pod/Container-Monitoring in K8s Clusters

Go-KI

1. Smart Fridge - Object Detection
2. Ground Assistance System - Voice Interaction

BeIntelli TP3

Localization by Visual Place Recognition

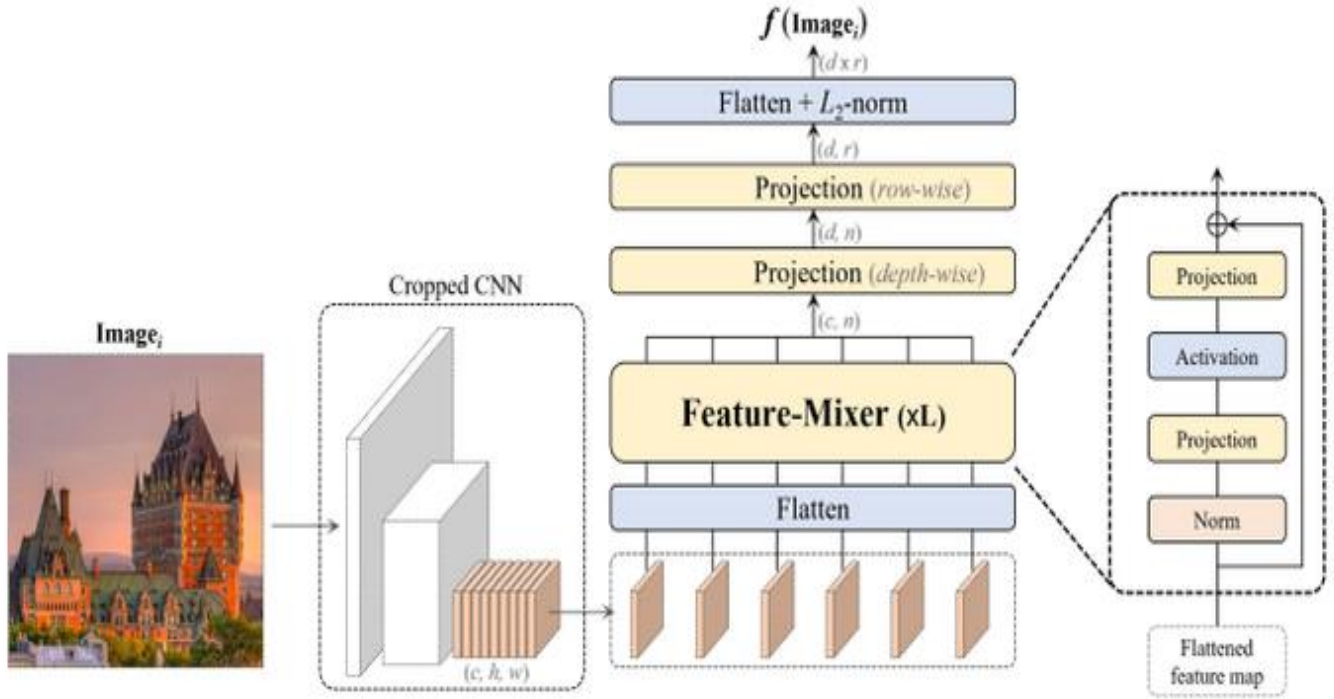
Group member:

- Ahmet Alperen Güngör

Supervisor:

- Ketan Motlag / Srinivas Ravuri

- To improve the initialization of the localization algorithm[1]
- Recognize environment by camera data using a NN-based regression model (MixVPR)[2]
- Locate where our vehicle is and output GNSS coordinates.



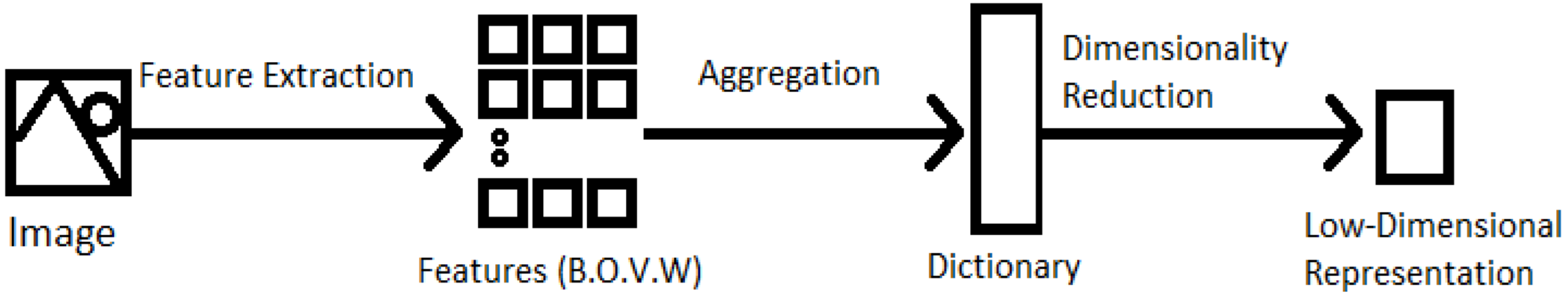
Literature Review - Manual Approaches:

Challenges:[3][4]

1)Viewpoint Change 2)Perceptual Aliasing 3)Appearance Change

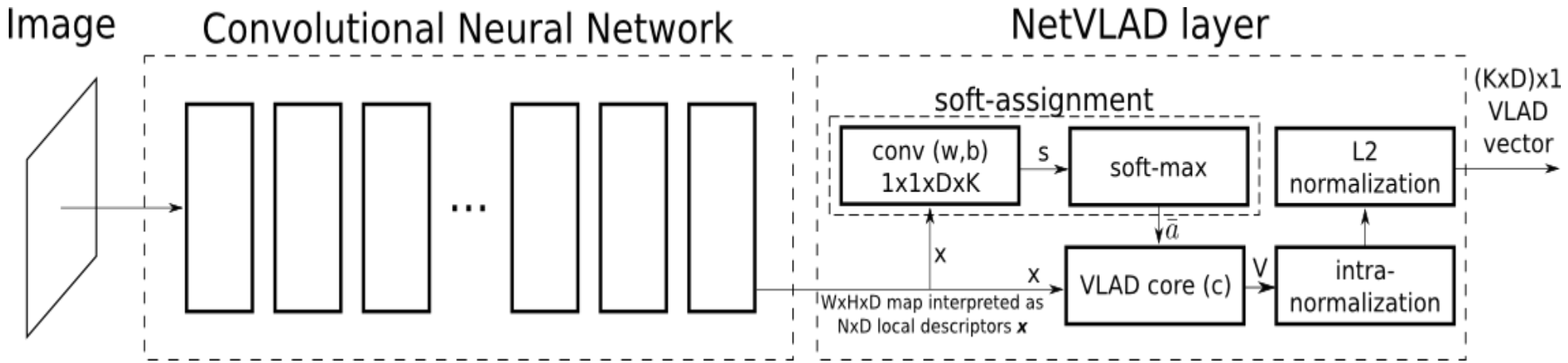
Process:

- Local (SIFT[5] /SURF[6]) or Global (GIST[7]) Feature Extraction
- Aggregation (VLAD[8][9])
- Dimensionality Reduction (PCA)

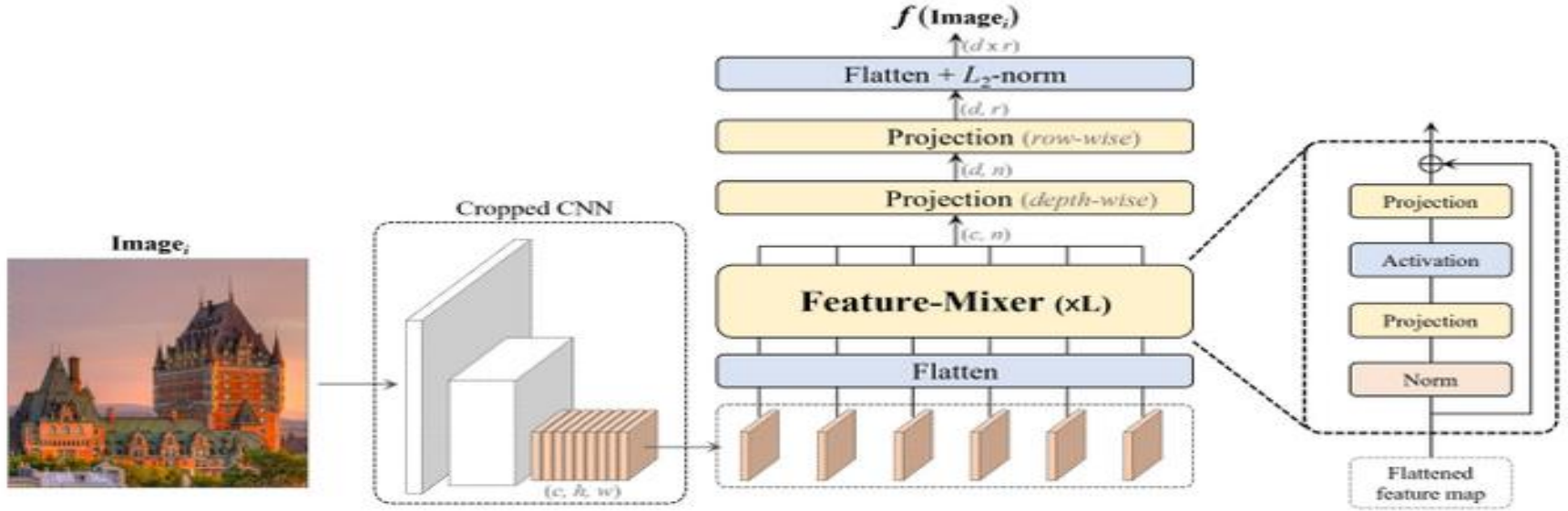


End-to-End Deep Learning Approaches:

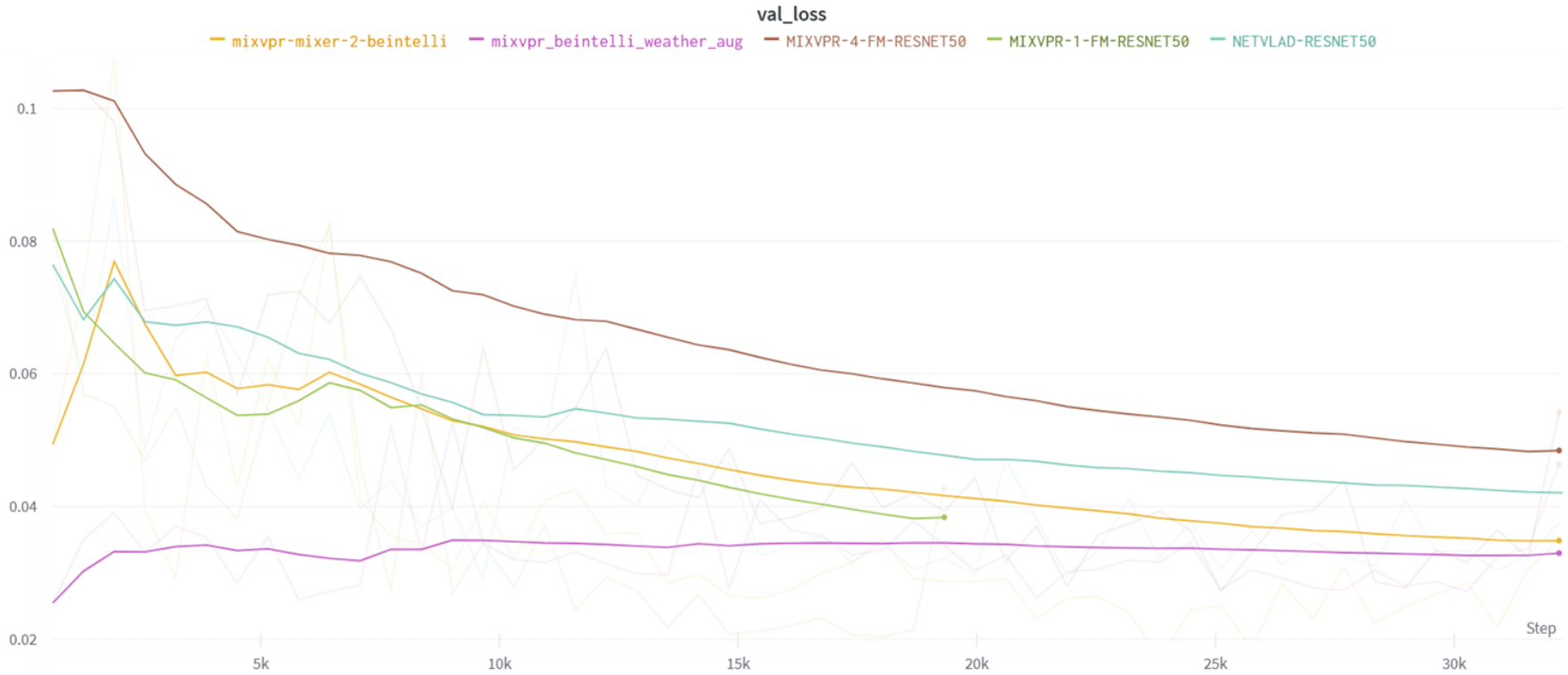
NetVLAD[10][11][12]**:**

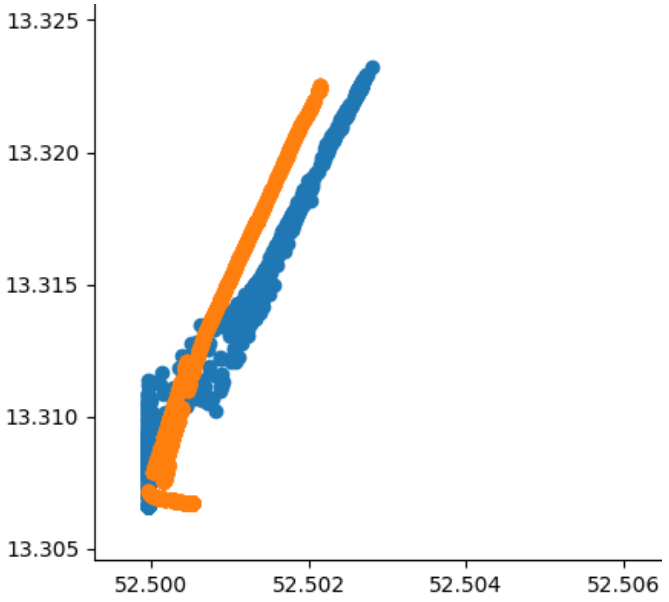


MixVPR:



Results:





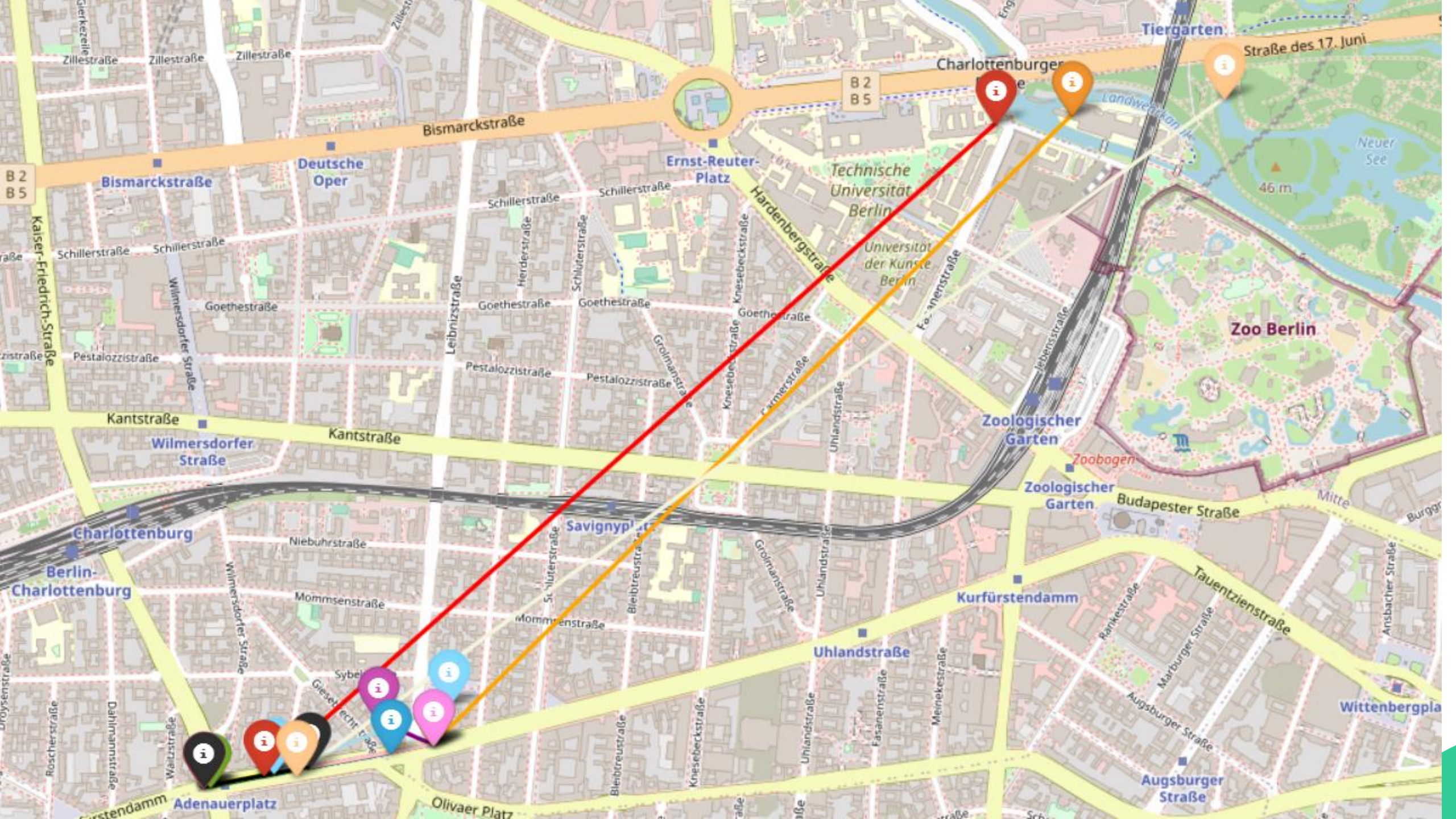
examples

pred: [52.50215322 13.32168739], target: [52.5019887 13.32144691]

pred: [52.50109027 13.31465805], target: [52.50082308 13.31404815]

pred: [52.50203281 13.3214075], target: [52.50212322 13.32229852]

pred: [52.50136279 13.31668175], target: [52.50134043 13.31736934]


Bismarckstraße
B 2
B 5
Deutsche Oper
Ernst-Reuter-Platz
Technische Universität Berlin
Universität der Künste Berlin
Hardenbergstraße
Charlottenburger Straße
Tiergarten
Straße des 17. Juni
Landwehrkanal
Neuer See
46 m
Zoo Berlin
Zoologischer Garten
Zoobogen
Budapester Straße
Mitte
Kaiser-Friedrich-Straße
Schillerstraße
Goethestraße
Pestalozzistraße
Wilmersdorfer Straße
Leibnizstraße
Herderstraße
Schlüterstraße
Kantstraße
Wilmersdorfer Straße
Charlottenburg
Berlin-Charlottenburg
Savignyplatz
Niebuhrstraße
Mommsenstraße
Kurfürstendamm
Uhlandstraße
Tauentzienstraße
Rankestraße
Marburger Straße
Augsburger Straße
Wittenbergplatz
Ansbacher Straße
Adenauerplatz
Olivaer Platz
Roscherstraße
Dahlmannstraße
Waitzstraße
Giesebrechtstraße
Bleibtreustraße
Knesebeckstraße
Grolmanstraße
Meinekestraße
Fasanenstraße
Jebensstraße

Future Works:

- Increasing accuracy (less loss):
  - More hyperparameter tuning and testing with current MixVPR model
  - Instead of a basic CNN feature extraction, a more complex one such as attention-based transformers[13]
  - Improvements in feature-mixer layer - a more complex and accurate layer if possible

- Improved robustness to temporary and repeating objects like cars, humans and trees/buildings (ex. Zoo road and rural areas)[14][15][16]

- Combining RGB image data with lidar data[17]

References:

[1] https://autoware.org/autoware-challenge-2023/
[2] A. Ali-bey, B. Chaib-draa, and P. Giguere, “Mixvpr: Feature mixing ` for visual place recognition,” in WACV, 2023.
[3] S. Lowry et al., "Visual Place Recognition: A Survey," in IEEE Transactions on Robotics, vol. 32, no. 1, pp. 1-19, Feb. 2016, doi: 10.1109/TRO.2015.2496823.
[4] Barros, T., Pereira, R., Garrote, L., Premebida, C., & Nunes, U. J. (2021). Place recognition survey: An update on deep learning approaches. *arXiv preprint arXiv:2106.10458*.
[5] Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60, 91–110 (2004). https://doi.org/10.1023/B:VISI.0000029664.99615.94
[6] Bay, H., Tuytelaars, T., Van Gool, L. (2006). SURF: Speeded Up Robust Features. In: Leonardis, A., Bischof, H., Pinz, A. (eds) Computer Vision – ECCV 2006. ECCV 2006. Lecture Notes in Computer Science, vol 3951. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11744023_32
[7] Oliva A, Torralba A. Building the gist of a scene: the role of global image features in recognition. Prog Brain Res. 2006;155:23-36. doi: 10.1016/S0079-6123(06)55002-2. PMID: 17027377.
[8] Arandjelovic, Relja & Zisserman, Andrew. (2013). All about VLAD. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 1578-1585. 10.1109/CVPR.2013.207.
[9] H. Jégou, M. Douze, C. Schmid and P. Pérez, "Aggregating local descriptors into a compact image representation," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 2010, pp. 3304-3311, doi: 10.1109/CVPR.2010.5540039.
[10] Arandjelovic, R., Gronat, P., Torii, A., Pajdla, T., & Sivic, J. (2016). NetVLAD: CNN architecture for weakly supervised place recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5297-5307).
[11] Uy, M. A., & Lee, G. H. (2018). Pointnetvlad: Deep point cloud based retrieval for large-scale place recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4470-4479).
[12] Hausler, S., Garg, S., Xu, M., Milford, M., & Fischer, T. (2021). Patch-netvlad: Multi-scale fusion of locally-global descriptors for place recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 14141-14152).
[13] Wang, R., Shen, Y., Zuo, W., Zhou, S., & Zheng, N. (2022). Transvpr: Transformer-based place recognition with multi-level attention aggregation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 13648-13657).
[14] Keetha, N., Mishra, A., Karhade, J., Jatavallabhula, K. M., Scherer, S., Krishna, M., & Garg, S. (2023). AnyLoc: Towards Universal Visual Place Recognition. arXiv preprint arXiv:2308.00688.
[15] Michael McHenry, Yang Cheng, and Larry Matthies "Vision-based localization in urban environments", Proc. SPIE 5804, Unmanned Ground Vehicle Technology VII, (27 May 2005); https://doi.org/10.1117/12.603973
[16] Schubert, S., Neubert, P., Garg, S., Milford, M., & Fischer, T. (2023). Visual Place Recognition: A Tutorial. arXiv preprint arXiv:2303.03281.
[17] Komorowski, J., Wysoczańska, M., & Trzcinski, T. (2021, July). MinkLoc++: lidar and monocular image fusion for place recognition. In 2021 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.