# Online Segmentation of LiDAR Sequences: Dataset and Algorithm


Romain Loiseau[1, 2] Mathieu Aubry[1] Loïc Landrieu[2]

romain.loiseau@enpc.fr mathieu.aubry@enpc.fr loic.landrieu@ign.fr

[1]LIGM, Ecole des Ponts, Univ Gustave Eiffel, CNRS, France

[2]LASTIG, Univ. Gustave Eiffel, ENSG, IGN, F-94160 Saint-Mande, France


ECCV TEL AVIV 2022

romainloiseau.fr/helixnet/

## Motivations

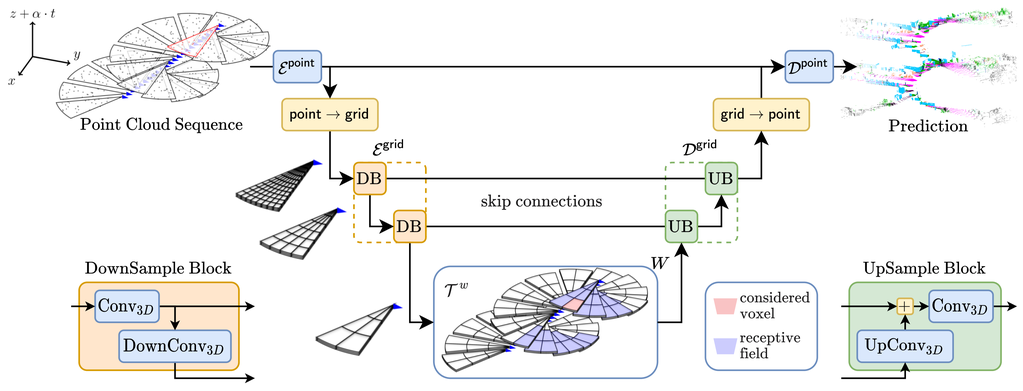
**Autonomous driving:**

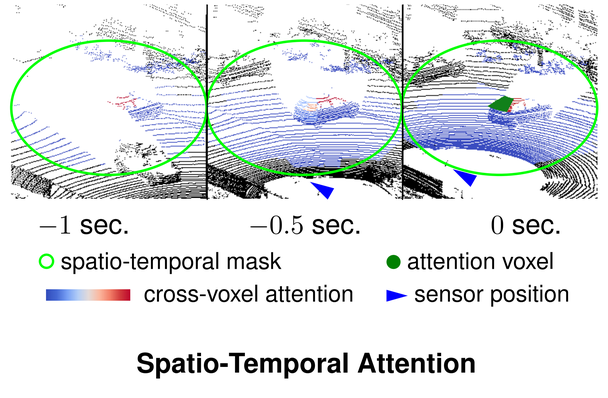
- **Real-time processing of LiDAR sequences** is critical.
- Current datasets do not have all **sensor's data at point level**.

**Contributions:**

- **HelixNet**, a 10-billion points **dataset** with fine-grained timestamps and sensor metadata and rotation information.
- **Helix4D**, a compact and efficient **spatio-temporal transformer** architecture designed for rotating LiDAR sequences.
- **SOTA performances** with a reduction of over $5\times$ in terms of latency and $50\times$ in model size compared to others. **Helix4D is as fast as the fastest and as precise as the most precise models.**

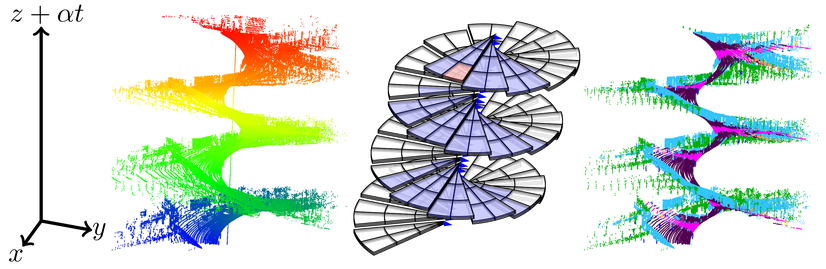
## Helix4D: a spatio-temporal transformer designed for efficient inference




**Spatio-Temporal Attention**

## LiDAR acquisition as a continuous sequence



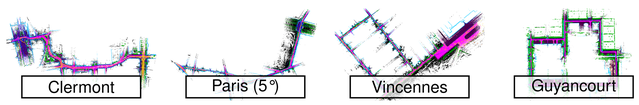

Temporal acquisition Online processing Semantic labels

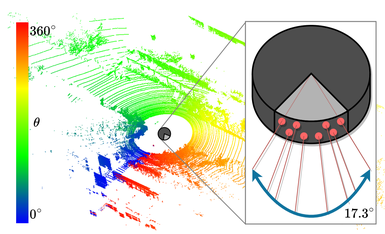
## Semantic Segmentation Results (mIoU)

| Method | Size $\times 10^6$ | Full frame ● 104ms HelixNet | Full frame ● 104ms SemKITTI [1] | Full frame ● 104ms Inf. (ms) | ⅕ frame ▶ 21ms HelixNet | ⅕ frame ▶ 21ms SemKITTI [1] | ⅕ frame ▶ 21ms Inf. (ms) |
|---|---|---|---|---|---|---|---|
| SalsaNeXt [2] | 6.7 | 69.4 | 55.8 | **23** ✓ | 68.2 | 55.6 | **10** ✓ |
| PolarNet [4] | 13.6 | 73.6 | 58.2 | 49 ✓ | 72.2 | 56.9 | 36 ✗ |
| Pan. PolarNet [5] | 13.7 | — | 64.5 | 50 ✓ | — | 60.3 | 44 ✗ |
| SPVNAS [3] | 10.8 | 73.4 | 64.7 | 73 ✓ | 69.9 | 57.8 | 44 ✗ |
| Cylinder3D [6] | 55.9 | 76.6 | **66.9** | 108 ✗ | 75.0 | 65.3 | 54 ✗ |
| **Helix4D (Ours)** | **1.0** | **79.4** | 66.7 | 45 ✓ | **78.7** | **66.8** | 19 ✓ |

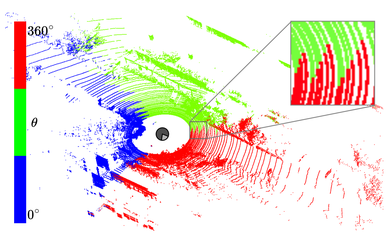
## HelixNet

- **10 billion points** spanning over 6 different French cities.
- **20 sequences** annotated with a **9-classes** nomenclature.
- **Point-level information:** sensor's rotation / position / release time ⟶ method's real-time readiness assessment.



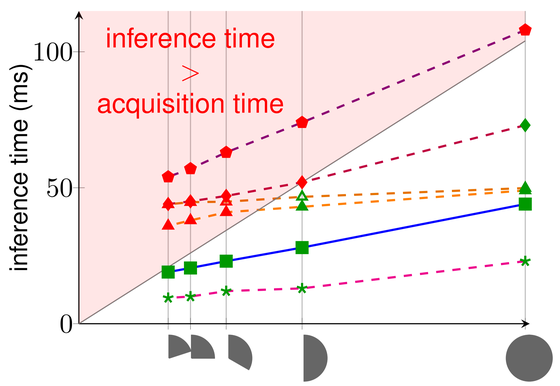



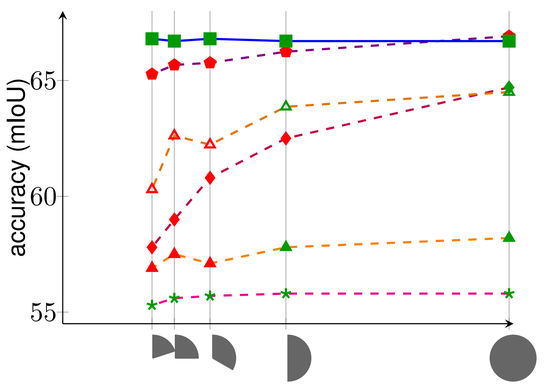

**Rotation of the sensor head**

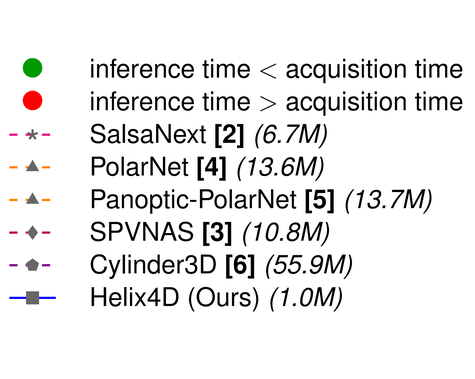



**Slices covering $120°$**

## Influence of Slice Size (SemanticKITTI [1] validation set)




## Bibliography


**[1]** Behley, J. *et al.*, ICCV 2019 ; **[2]** Cortinhal, T. *et al.*, ISVC 2020 ; **[3]** Tang, H. *et al.*, ECCV 2020 ; **[4]** Zhang, Y. *et al.*, CVPR 2020 ; **[5]** Zhou, Z. *et al.*, CVPR 2021 ; **[6]** Zhu, X. *et al.*, CVPR 2021