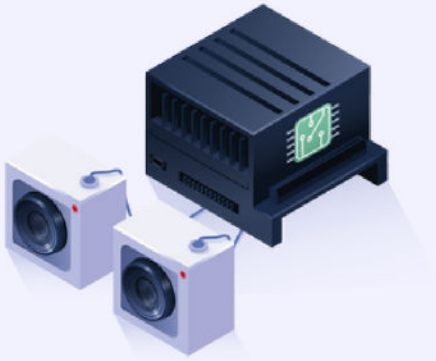
aloception

for autonomous & assisted mobility



Aloception Xavier

Cameras

Any monocular or stereo camera setup. Tested with ZED2, STURDeCAM20-IP67, ZED 2i, Realsense D435, OAK-D-Lite

Jetson AGX Xavier

GPU 384-Core Volta GPU with 48 Tensor cores 5.5 TFLOPS (FP16) 11.1 TOPS (INT8). DLA (2x) NVDLA Engines 4.1 TFLOPS (FP16) 8.2 TOPS (INT8). CPU 6-core Carmel ARM v8.2 64-bit CPU, 8MB L2 + 4MB L3
Memory 8GB 256-bit LPDDR4x 1333 MHz - 85.3 GB/s
Module Power 10-15-30W. Volatge 5V, 9V-20V. 30W Required by Aloception

Latent space & Features

Alopix : Pixel to 3D world latent representation from stereo or monocular camera, FOV < 180. Included features are : monocular depth estimation, stereo estimation with or without external prior, free space & obstacles.

Aloception development kit

Docker deploiement is supported with Jetpack 5.0.2 7 Jetpack 8.4. ROS Wraper can be used at inference.

Aloception Xavier

Aloception's core technology involves building a 3D latent representation of the environment which captures the geometry of diverse scenes (warehouses, ADAS, water, offices ...) This representation provides relevant information to robotics for autonomous scene comprehension and navigation. The concept of latent space for neural networks in robotics involves learning a lower-dimensional representation of the environment that preserves important features while reducing the complexity of the data. In this way, the neural network can make more efficient use of limited computational resources and can better generalize to new environments. Using this 3D latent representation, Aloception enables robots to better understand their environment and navigate it more effectively.

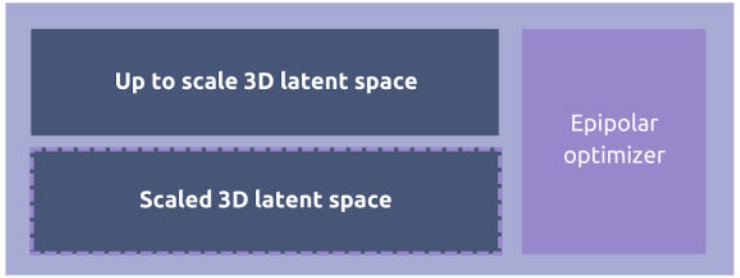
Quality in diverse environments

The aloception Xavier solution performance is measure with data from various environments to ensure that the system is robust and can handle various real-world scenarios.
Testing the system in different environments can help identify weaknesses and areas for improvement. Additionally to metrics we use manual visual inspection to help identify sources of errors or inaccuracies in the system's output that may not be apparent through quantitative metrics alone.
This can help improve the system's accuracy and increase its usefulness in practical applications. Aloception capabilities are tested in various environments including warehouses, ADAS, water, offices, and off-road environments. Please find attached to this document sample of videos from these environments that we used internally to measure the system performance.

Neural architecture

Left frame | Right frame | Single frame | Depth prior

The system can works with multiple combination of the above



Free space | Obstacles | Stereo depth | Monocular depth

System features

Free space

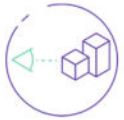
Any potential navigable area. We define as navigable any part of the floor that could potentially be used by a mobile robot to navigate, such as grass, floor, or water. Specific navigable areas can be targeted on-demand.

Obstacles

Any group of points that is relatively close to the camera and not part of the free space. This formulation makes it possible to detect any potential obstacles without prior knowledge of the obstacle.

Sereo depth

Any potential navigable area. We define as navigable any part of the floor that could potentially be used by a mobile robot to navigate, such as grass, floor, or water. Specific navigable areas can be targeted on-demand.



Monocular depth

Monocular up-to-scale depth estimation. This is useful in this setup to distinguish between close and far obstacles.