U-net architecture with 12 blocks of encoder and 12 of decoder trained with 50 epochs / 4 batch size on 1024x1024 images.

Training

A collage of images of people walking on a bridge

Description automatically generated

Prediction

A collage of images of a bathroom

Description automatically generated