Homework #4 - Machine Learning for Robotics (RBE 577) Self-Supervised Depth Estimation

Eric Viscione, Paul Crann
RBE577: Machine Learning For Robotics

Introduction:

Depth estimation from single images have many applications in robotics, autonomous navigation, augmented reality, and 3D scene understanding, but need lots of labeled data to train in a supervised manner. Getting labeled, real world, depth maps is very challenging and costly, making the supervised approach unappealing. MonoDepth2 implements a different approach to the problem with an architecture capable of self-supervised training on unlabeled, real world, image sequences. Here we implement the MonoDepth2 library on the syndrome dataset and compare the ground truth depth maps to the MonoDepth2 depth prediction network.

Model:

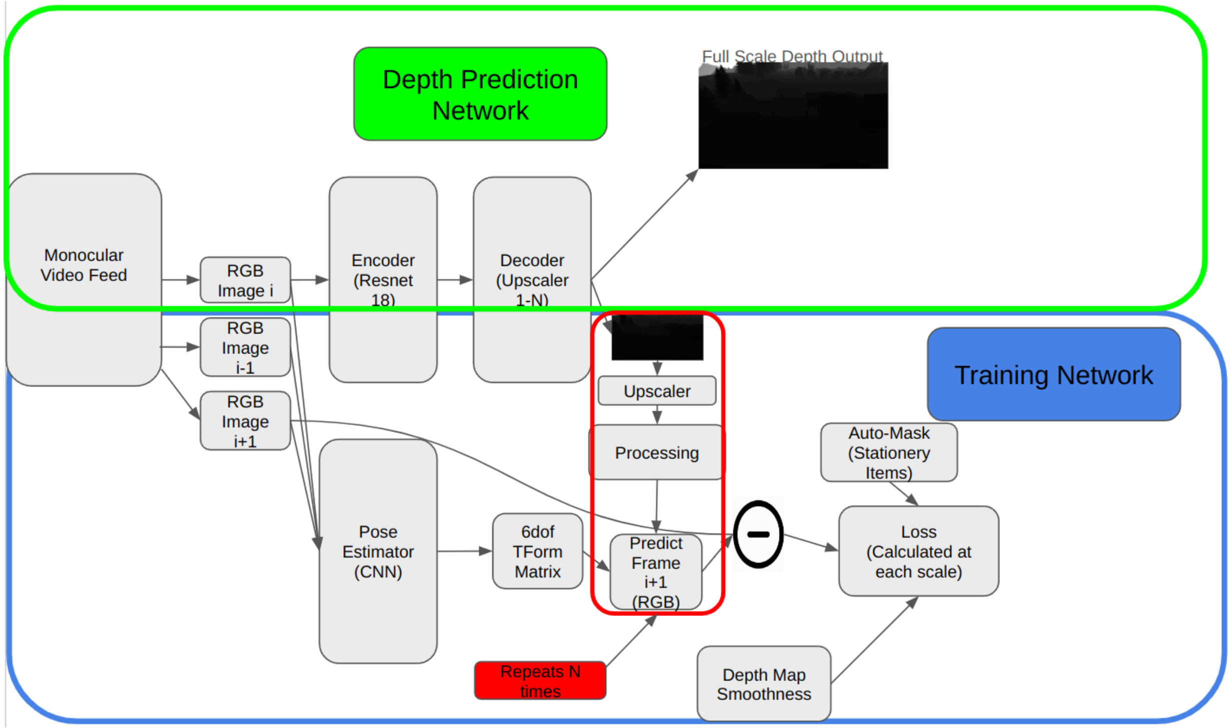
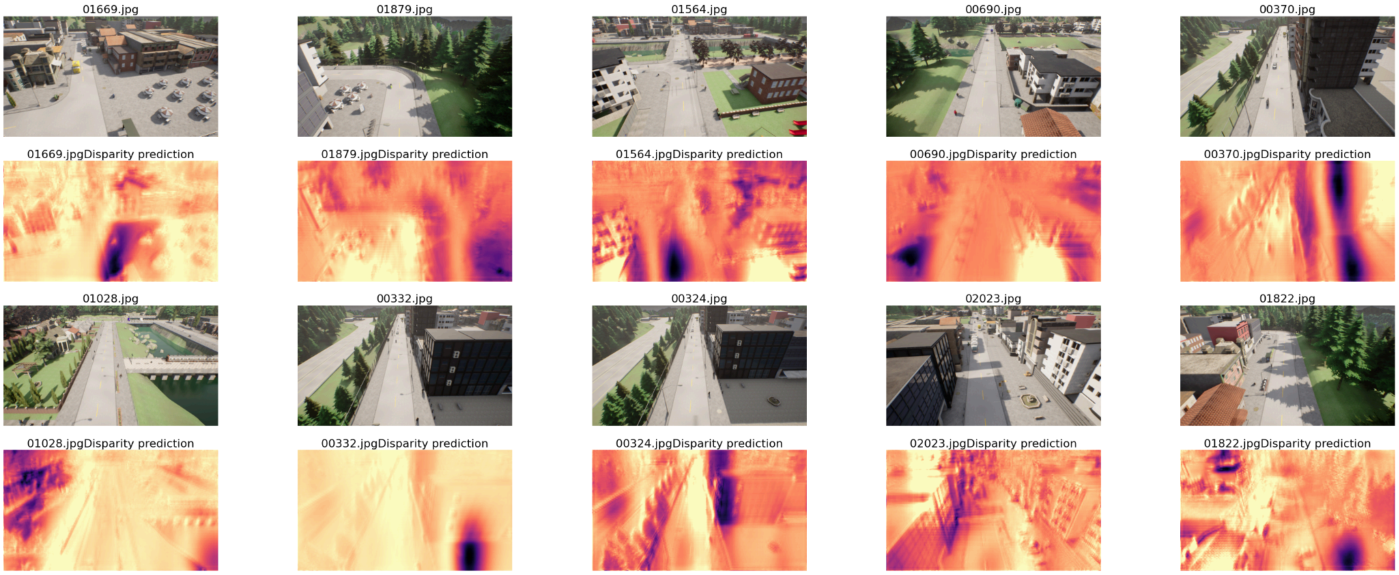


Figure 1: Model Architecture Diagram

Model Architecture:

The model uses two main networks to train, a depth prediction network, and a training pose estimator network. The depth prediction network features a U-Net style encoder-decoder. The encoder uses a pretrained ResNet18, and the decoder utilizes upsampling layers along with skip connections from the encoder to produce a full scale depth map. The pose estimation network uses a modified ResNet18 model to intake sets of images [t-1, t, t+1] and predict 6-DoF relative poses between frames. This network is only utilized during training. Using the pose estimator network during training allows the depth prediction network to be trained on unlabeled datasets by computing the loss using the predicted depth images.

Results:


01669.jpg
01879.jpg
01564.jpg
00690.jpg
00370.jpg
01669.jpgDisparity prediction
01879.jpgDisparity prediction
01564.jpgDisparity prediction
00690.jpgDisparity prediction
00370.jpgDisparity prediction
01028.jpg
00332.jpg
00324.jpg
02023.jpg
01822.jpg
01028.jpgDisparity prediction
00332.jpgDisparity prediction
00324.jpgDisparity prediction
02023.jpgDisparity prediction
01822.jpgDisparity prediction

Loss Vs Epoch

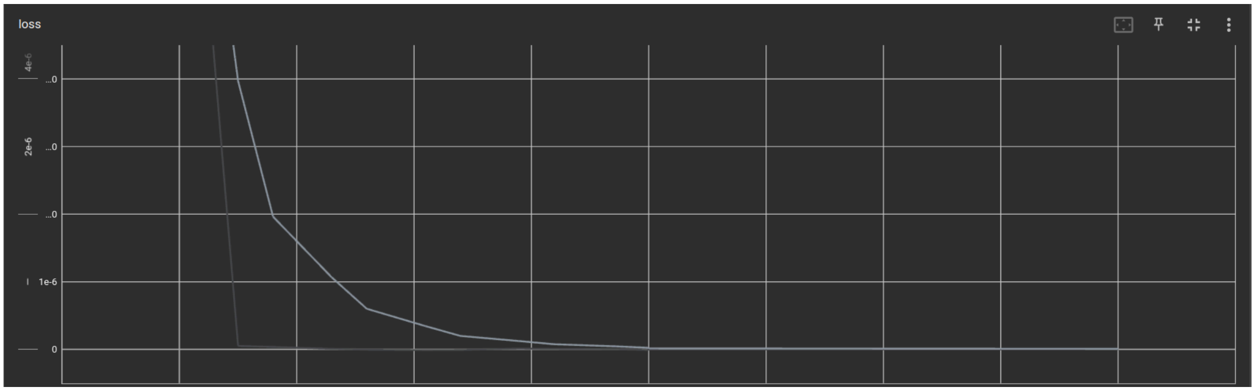


Figure 2: Loss Vs Epoch