

Depth Estimation - KIWIBOT

David Choqueluque Roman
dchoqueluque@gmail.com

The problem

Build a model for depth estimation using all 4 cameras. The model should process the 4 cameras at the time and output the corresponding 4 depth maps

My solution

To solve the problem I have used a simple U-Net. The main idea is to train the model as a simple depth estimator using the four cameras in the dataset as different samples (Figure 1).

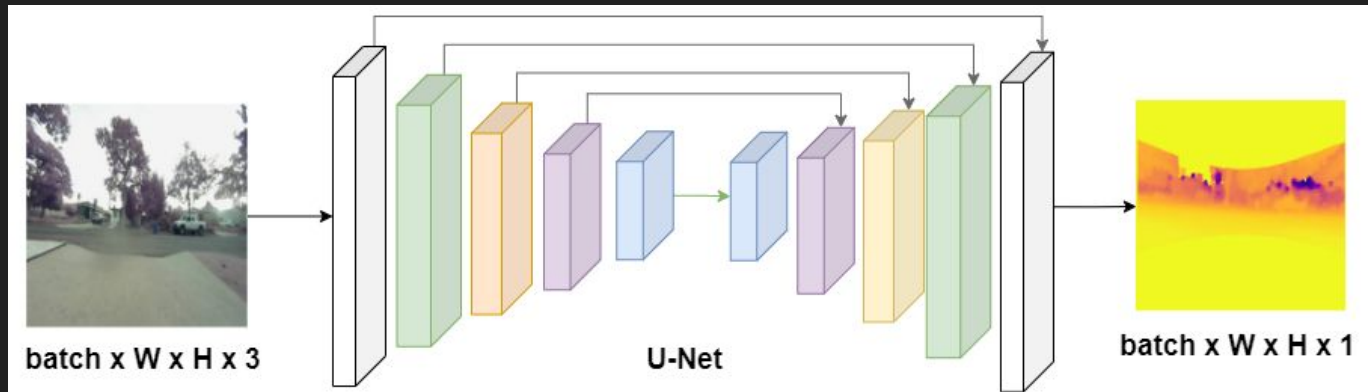


Figure 1. Training step

My solution (cont.)

During inference time, to process four inputs in one forward pass, the four images (left, right, raw, and back) were stacked in the batch dimension (See Figure 2). This approach allowed me to train a simple U-Net with images of the four cameras but it could compute depth maps in parallel. The model works in parallel performing like a four stream CNN with shared weights. For example, if we want to process 3 robots the input data dimensions are $(3 \times 4 \times W \times H \times 3)$, but the model reshapes the input in the inference step to $(12 \times W \times H \times 3)$. This approach works because the model was trained using images from all cameras.

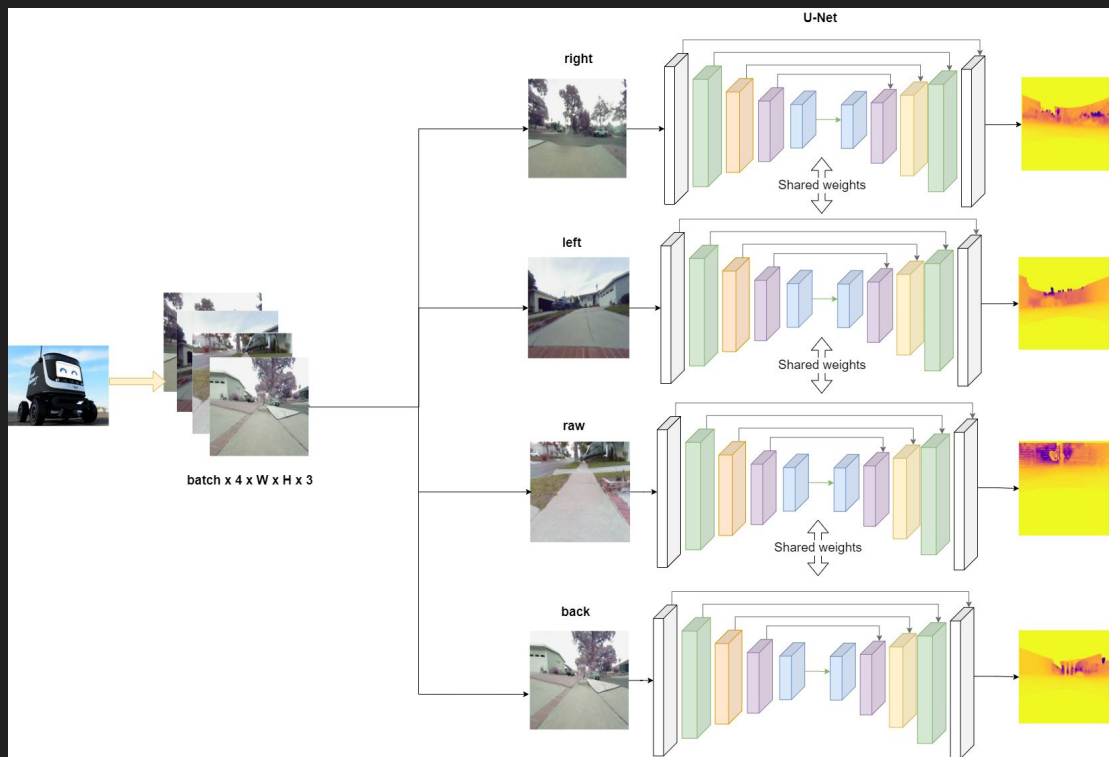
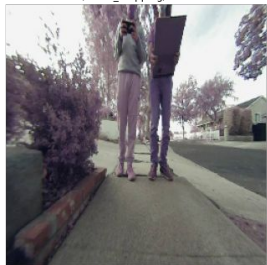


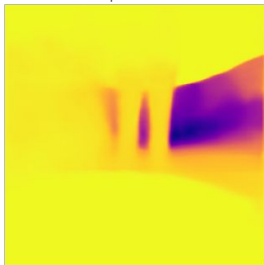
Figure 2. Inference step

Results

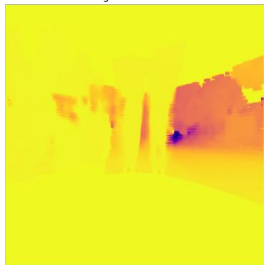
/video_mapping/back



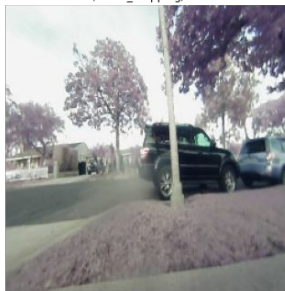
prediction



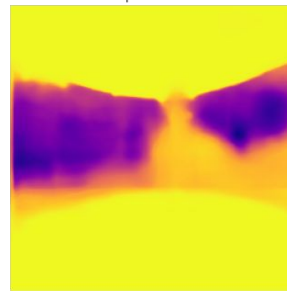
ground truth



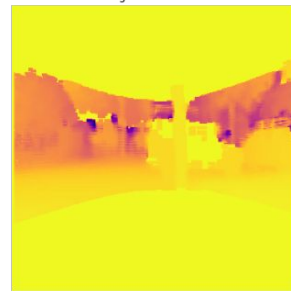
/video_mapping/left



prediction



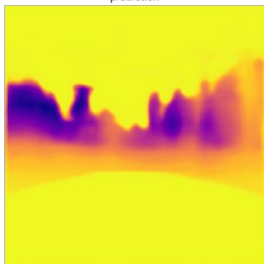
ground truth



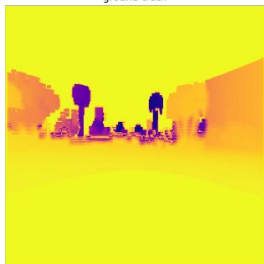
/video_mapping/left



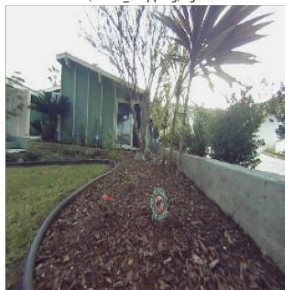
prediction



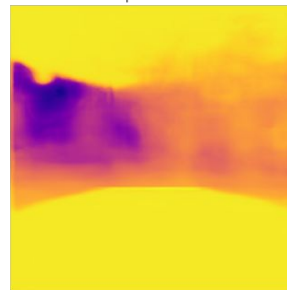
ground truth



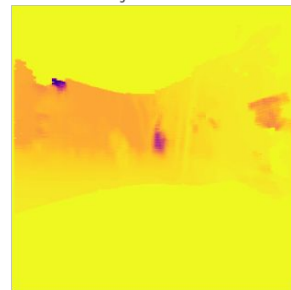
/video_mapping/right



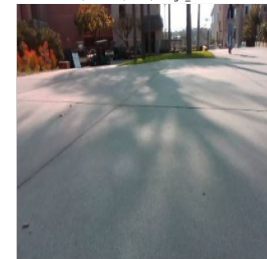
prediction



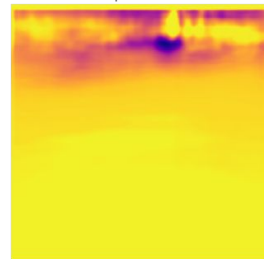
ground truth



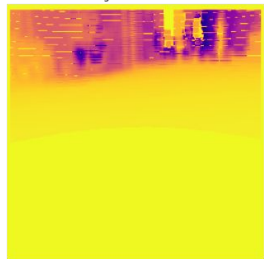
/camera/color/image_raw



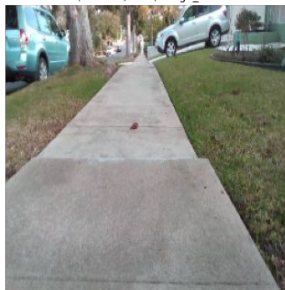
prediction



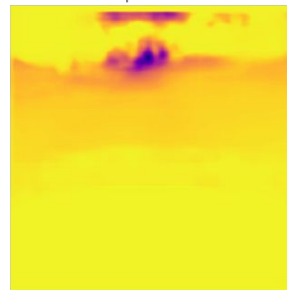
ground truth



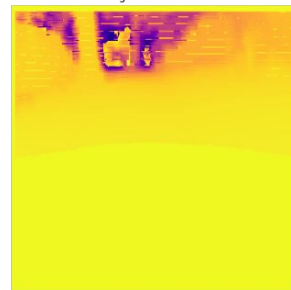
/camera/color/image_raw



prediction



ground truth



References

- https://keras.io/examples/vision/depth_estimation/
- <https://kargarisaac.medium.com/self-supervised-monocular-depth-estimation-in-autonomous-driving-636b0c01eb61>
- <https://medium.com/mlearning-ai/monocular-depth-estimation-using-u-net-6f149fc34077>
- [High Quality Monocular Depth Estimation via Transfer Learning](#)