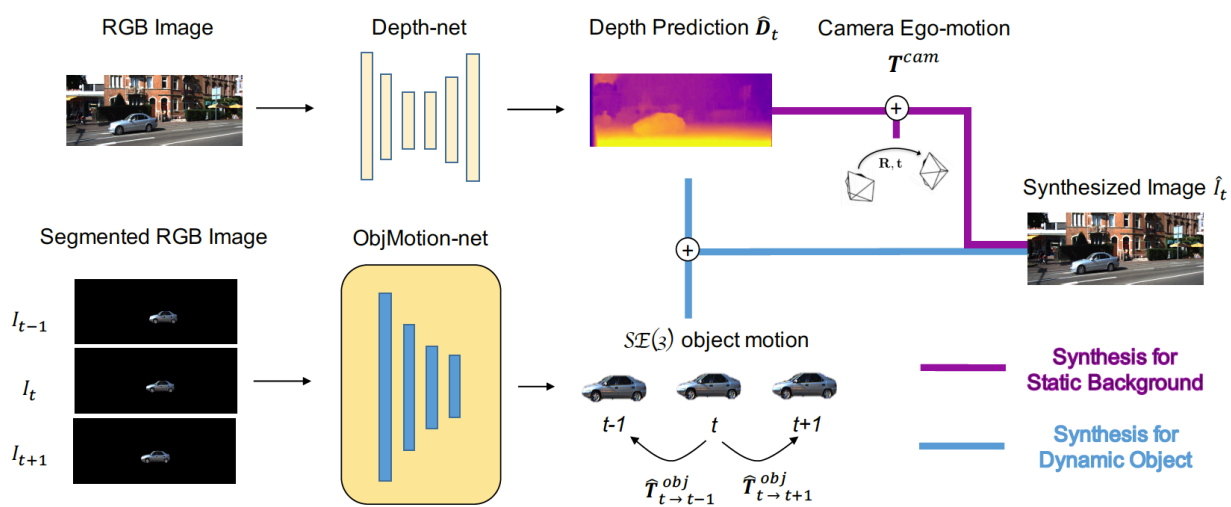


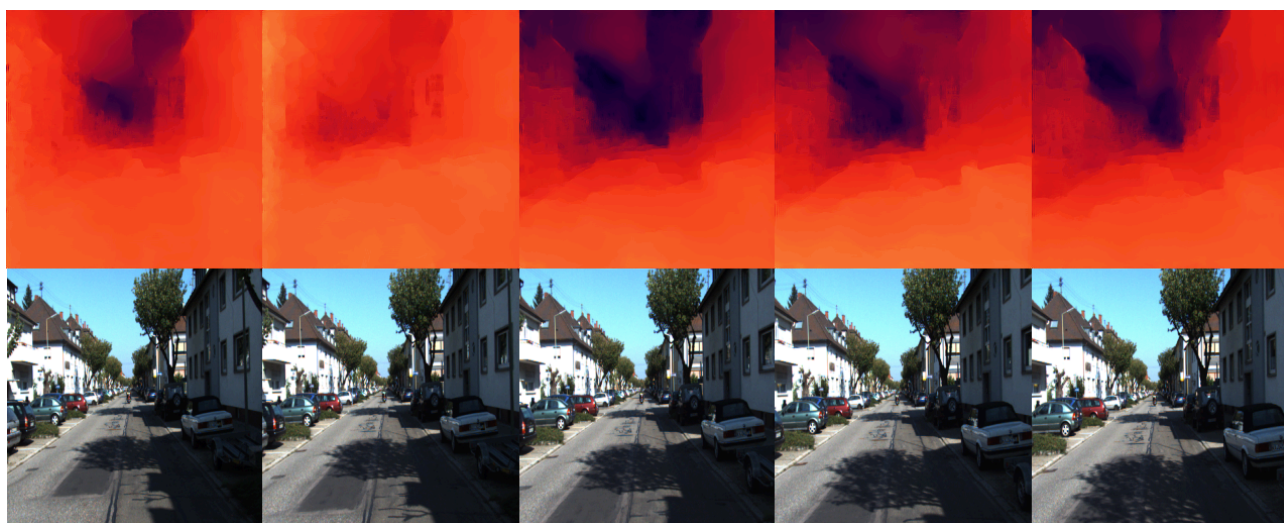
Self-supervised Object Motion and Depth Estimation from Video

Libing Zeng

In this project, I was planning to implement this new paper, Self-supervised Object Motion and Depth Estimation from Video, in PyTorch. But the goal is changing a little bit. Because, after I build the first part of the model, that is depth estimation, I found that the depth prediction results is not temporally consistent. For example, when we generate depth predictions of a sequence of images and then go through the predictions like playing a video, we can see serious flickering artifacts between neighboring depth predictions. And then I read the paper carefully again, and I don't see they take this temporally consistent problem into consideration. Also, they don't have any demo videos.



Currently, I have finished the depth estimation part. As mentioned above, temporal inconsistency would be a natural problem in this kind of single image depth prediction algorithm. Here, I would to show some depth predictions of a sequence of images.



From this figure, we can see that there are severe inconsistency artifacts in the depth predictions, especially for from the first prediction to the second one and then to the third one.

I will try to improve this inconsistency problem. Actually, currently, this is an actively research topic. For example, a paper in siggraph2020, Consistent Video Depth Estimation, which is a test fine-tuning algorithm.