

Model – final score 0.47

Final model had a tricky architecture. It is tough to explain with words, so I draw the picture in Appendix A. Basic principle is to create system where layers passed along the network to learn nice distinctions within the objects of the scene.

Such a deep model with some of information is being passed all the way through the network allows to catch small details, which is required to detect target from bigger distance. And that was primary issue of the project.

Issues

Provided dataset did not include enough training examples that contain target hero, so model could not converge. Later explorations showed that it is better to neglect provided dataset and collect own images.

Adam optimizer has rapid gradient descend, so learning rate of 0.0005 was used to fight it. Otherwise, model stuck on local minimum.

Data

Total amount of collected images is around 3000. In addition, all of the pictures and masks are flipped horizontally in order to increase dataset. My estimation is – that architecture is space specific; flipping image change where the object is. As the result, it is beneficial to do so.

Extensions

Model can be retrained to follow different target and with its current state it would take significantly less time and effort, as pre-trained model is already knows well the color pallet of the simulator and ways to deconstruct image. However, in order to retrain model to follow new object at least 1000 of new images of new target within the environment would be needed.

However, the network is powerful and can be scaled to include additional classes.

Hyperparameters

- ➔ Epoch – is set to 20, but model was retrained multiple times. In total about 100. Longer training might improve model even further.
- ➔ Batch size – 4 which along with small learning rate decrease rate of gradient descend.
- ➔ Learning rate – 0.005 as stated above
- ➔ Steps per epoch – I used recommended formula “number of training samples” / “batch size”

Appendix A model

