

# CS698O PROJECT: DEPTH PREDICTION USING SINGLE IMAGE

Divyanshu Shende, Rahul Tudu

Indian Institute of Technology Kanpur

In this project we implement a NIPS 2014 paper by David Eigen, Christian Puhrsch, Rob Fergus. The paper is titled Depth Map Prediction from a Single Image using a Multi-Scale Deep Network.

## The Dataset

The dataset used by the us is the labelled NYU Depth v2 dataset. This contains the 3 dimensional image matrix and also a depth matrix which gives the depth (in metres) of each pixel from the camera.

## The Network

The network used has two component stacks. These are called the coarse and fine layers. The input to both these layers is an 304x228 image with 3 RGB channels. The coarse layer outputs a 1/4 resolution (of the original) depth. This corresponds to a coarse depth map. This is then combined with one of the outputs of the fine layers and then processed to obtain a finer depth map having resolution 1/4-th that of the original image.

## The model architecture (from the paper)

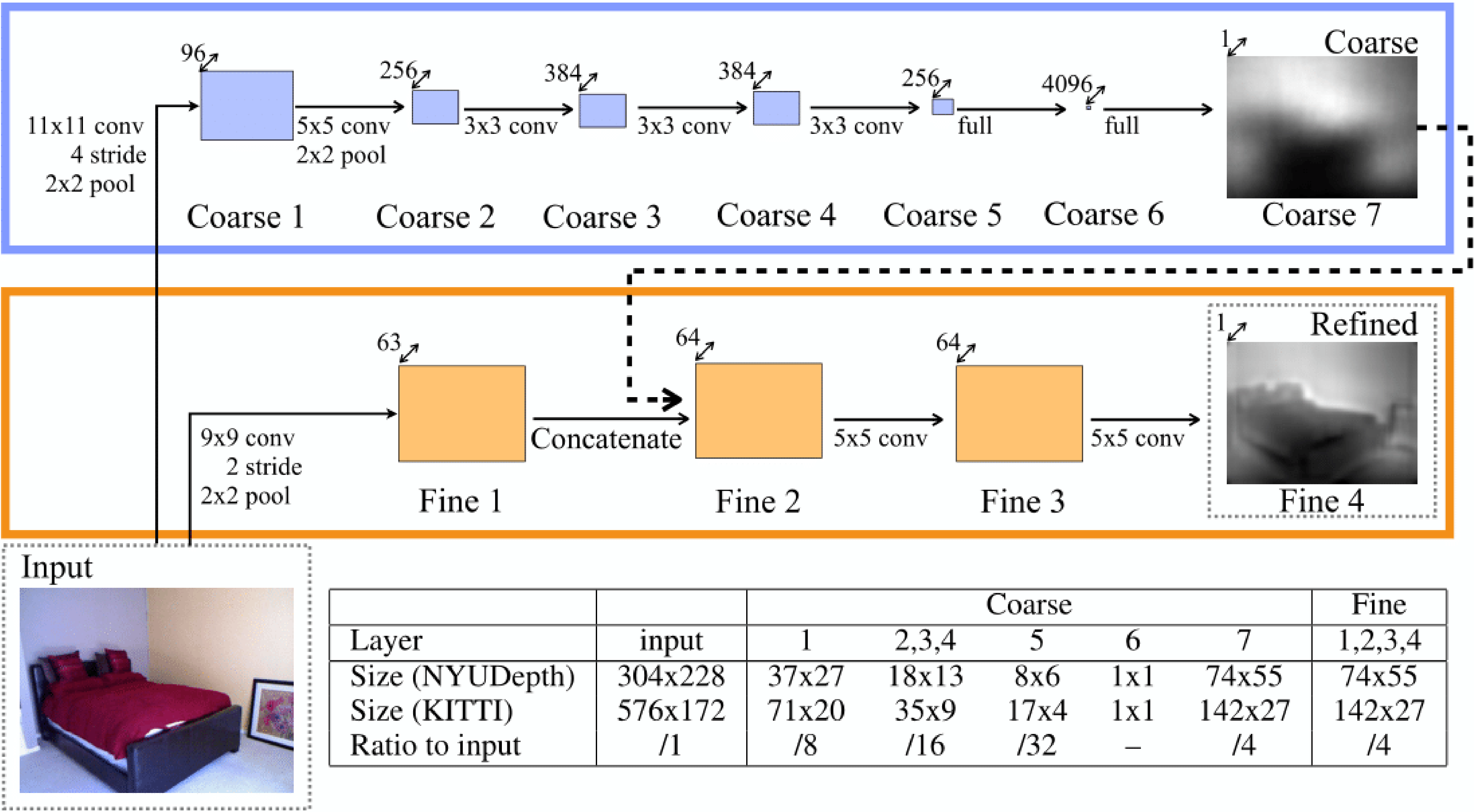


Figure 1: Model architecture.

## Loss function

The scale of the image affects the depth predictions. For this reason, the scale invariant loss function is used by the authors. This scale invariant loss function is given by  $D(y, y*) = \frac{1}{2n} \sum_1^n abc$

## Implementation details

The implementation of the authors was in Theano. However,