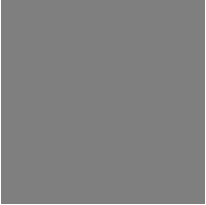
**Methods/Network architecture**

The network we used is an AutoEncoder-like network. Which means it has an encoder, a narrow bottleneck and a decoder part. The difference is in the input and output. In an AE network we usually use the same data as input and output and goal is to compress the data. In our case we give the grayscale image as input and the a and b channels of the colorized pictures as the output. So not just the input and output are different, but moreover their shapes are also dissimilar.

Firstly let’s take a look at the input, which is a 256x256x1 sized matrix. It means that it has 256x256 pixels and only 1 channel.

1

256

256

It goes through the encoder first which was built from several convolutional layers:

1. input: 1@256x256

filter: 3x3 @ 64 (3x3 sized filter with 64 depth)

stride: 2x2

padding: