7.2 Solution:

Receptive field size = (output_size - 1) * stride + Ksize.

Input -> C64 -> 16 x 16 -> C128 -> 7 x 7 -> C256 -> 4x4 -> C512 -> 1x1

7.3 Solution:

Because L_1 could force the low-frequency correctness. If we only minimize L_1, the low-frequency correctness will be weakened, which will result into a blurry image.