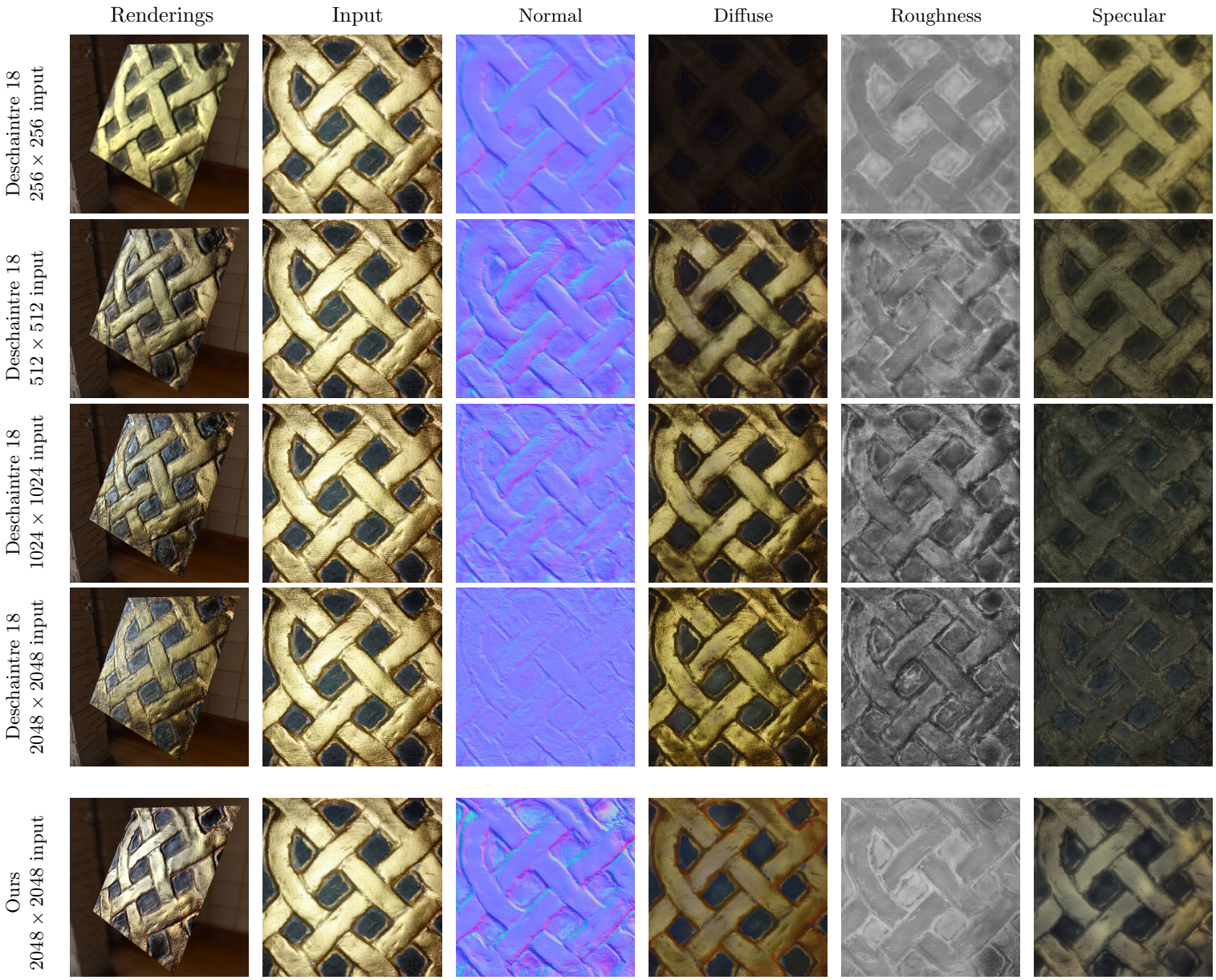

Figure 1: In this example, we have applied the deep network of Deschaintre et al. [2018] on images of increasing resolution, in a fully-convolutional manner. The network was trained at resolution 256×256 pixels. As resolution increases, the relative network footprint decreases, preventing the network to perceive all visual cues at once. In particular, at large resolution the flash highlight covers more than an area of 256×256 pixels, which never occurs in the training data. As a consequence, the network progressively interprets all image variations as due to the diffuse term, flattening the normal and specular maps proportionally. The last row provides results of our method applied at high resolution, using crops of the 256×256 maps predicted by Deschaintre et al. as exemplars. Since the input image is captured under a strong collocated flash lighting, our method tends to under-estimate the specular albedo and over-estimate the diffuse albedo. Nevertheless, our results are much closer to the ones obtained by Deschaintre et al. on the 256×256 input than the results obtained using their method in a fully-convolutional manner.