

Deep Unsupervised Pixelization

Supplementary Materials

Chu Han, Qiang Wen, Shengfeng He, Qianshu Zhu, Yinjie Tan,
Guoqiang Han, Tien-Tsin Wong

1 Network Structure

In order to simplify the explanation, we define a 7×7 Convolution-InstanceNorm-ReLU layer with k filters and stride 1 as C7s1- k . 9R- k indicates 9 residual blocks with k filters. D3s $\frac{1}{2}$ - k means a 3×3 fractional-strided-Convolution-InstanceNorm-ReLU layer with k filters and stride $\frac{1}{2}$.

GridNet consists of:

C7s1-64, C3s2-128, C3s2-256, 9R-256(C7s1-3), 8C3s2-256(C7s1-3), 8C3s2-256(C7s1-3).

Note that for the last three layers of this network, we add C7s1-3 repectively to produce our multiple outputs.

PixelNet and DepixelNet have the same architecture, they both consist of:

C7s1-64, C3s2-128, C3s2-256, 9R-256, D3s $\frac{1}{2}$ -128, D3s $\frac{1}{2}$ -64, C7s1-3.

Discriminators consist of:

C4s2-64, C4s2-128, C4s2-256, C4s2-512.

After the last layer of each discriminator, we add one convolution to produce one dimensional result.

2 More Results



Figure 1: More pixel art results (1). (©Nintendo.)



Figure 2: More pixel art results (2). (©Nintendo.)



Figure 3: More pixel art results (3). (©Nintendo.)

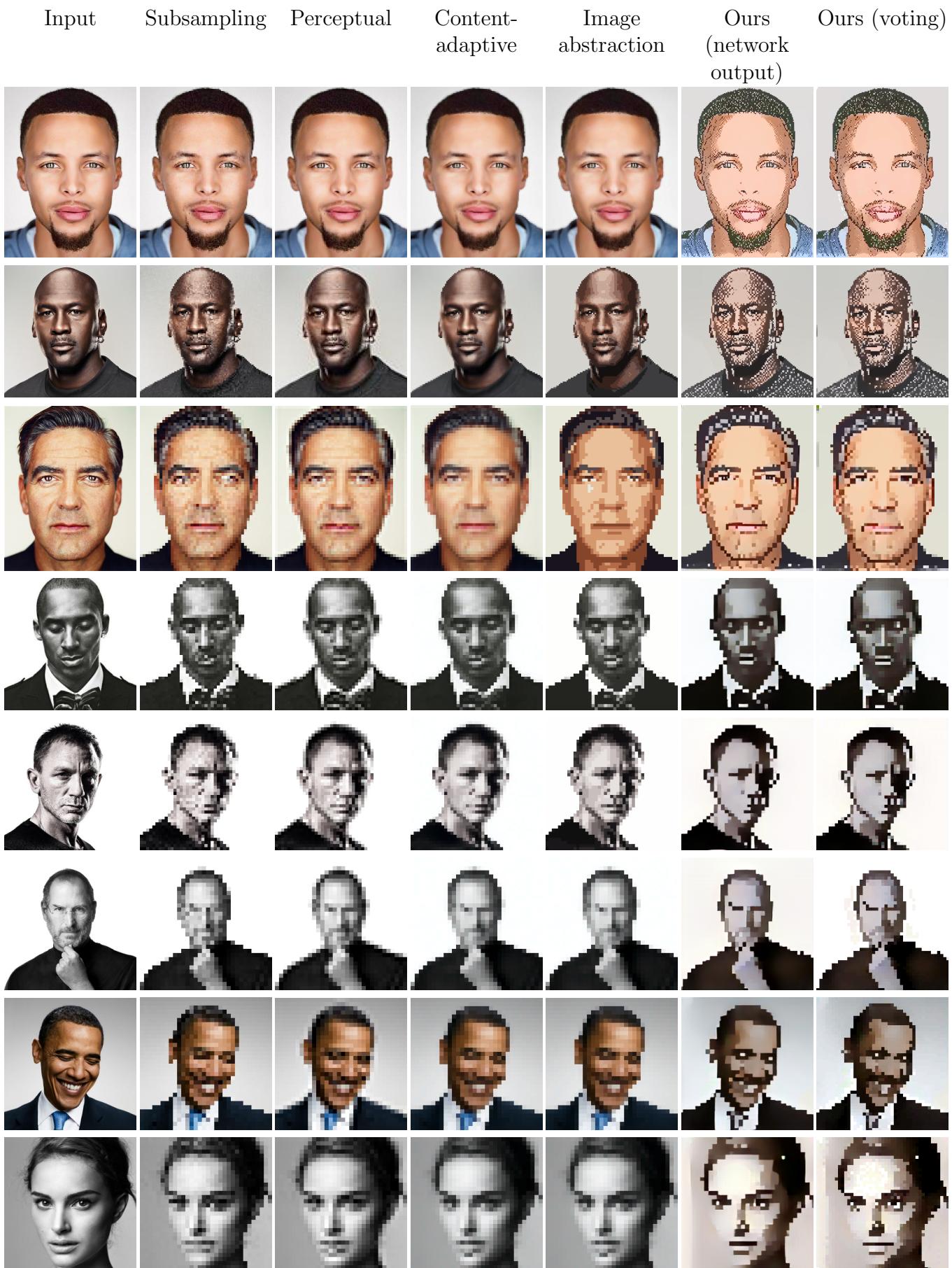


Figure 4: More portrait results (1).

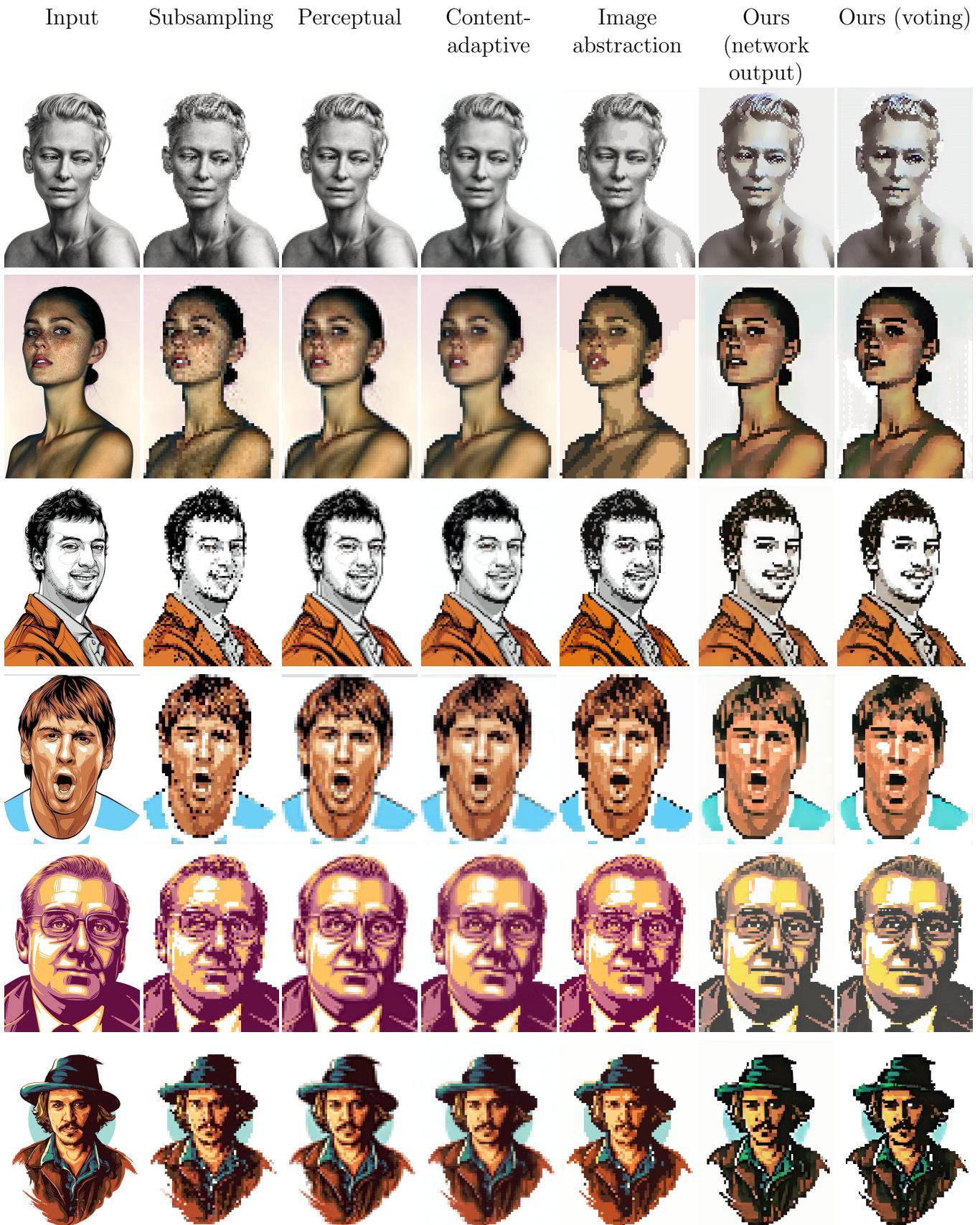


Figure 5: More portrait results (2).

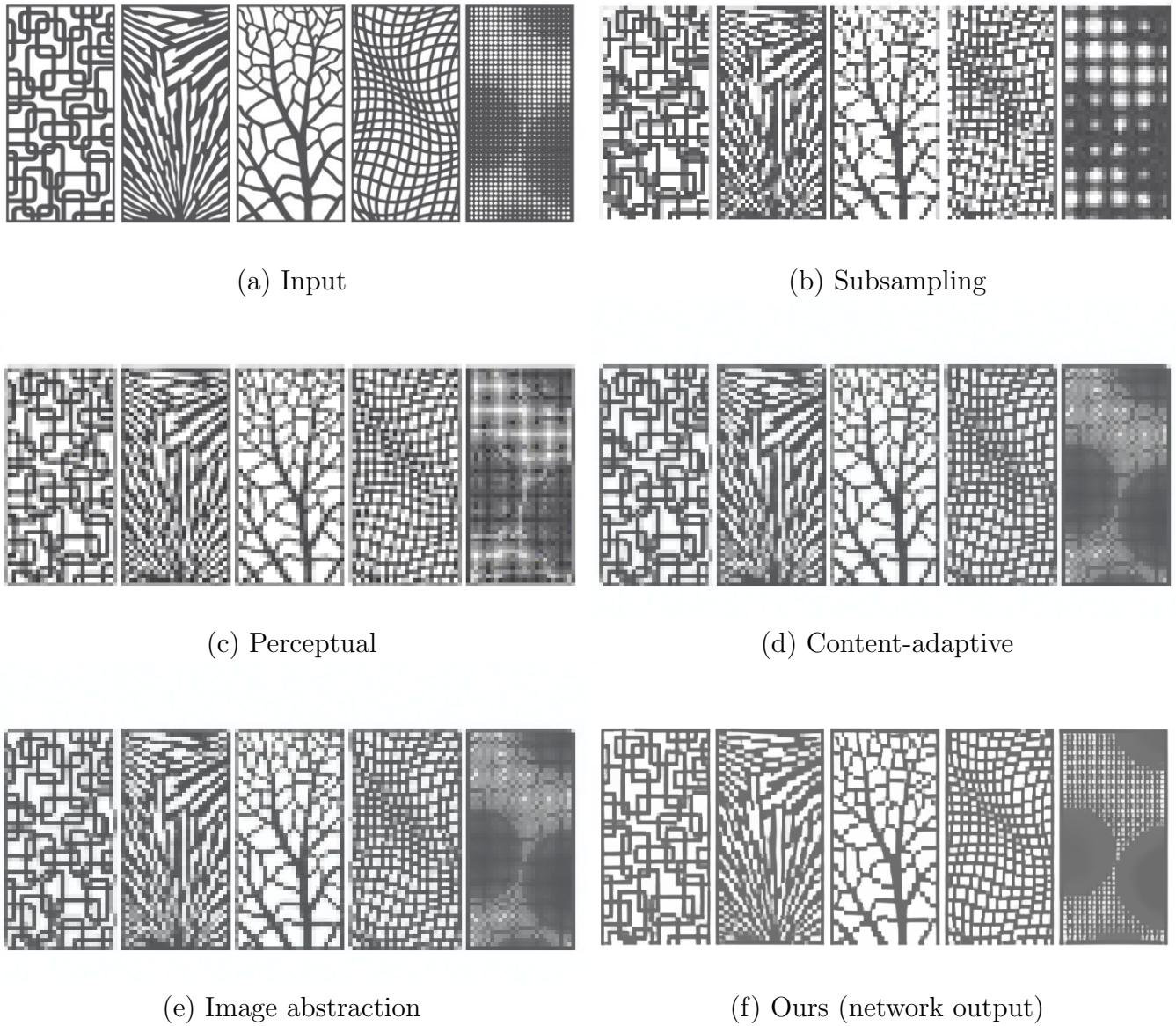


Figure 6: More results.

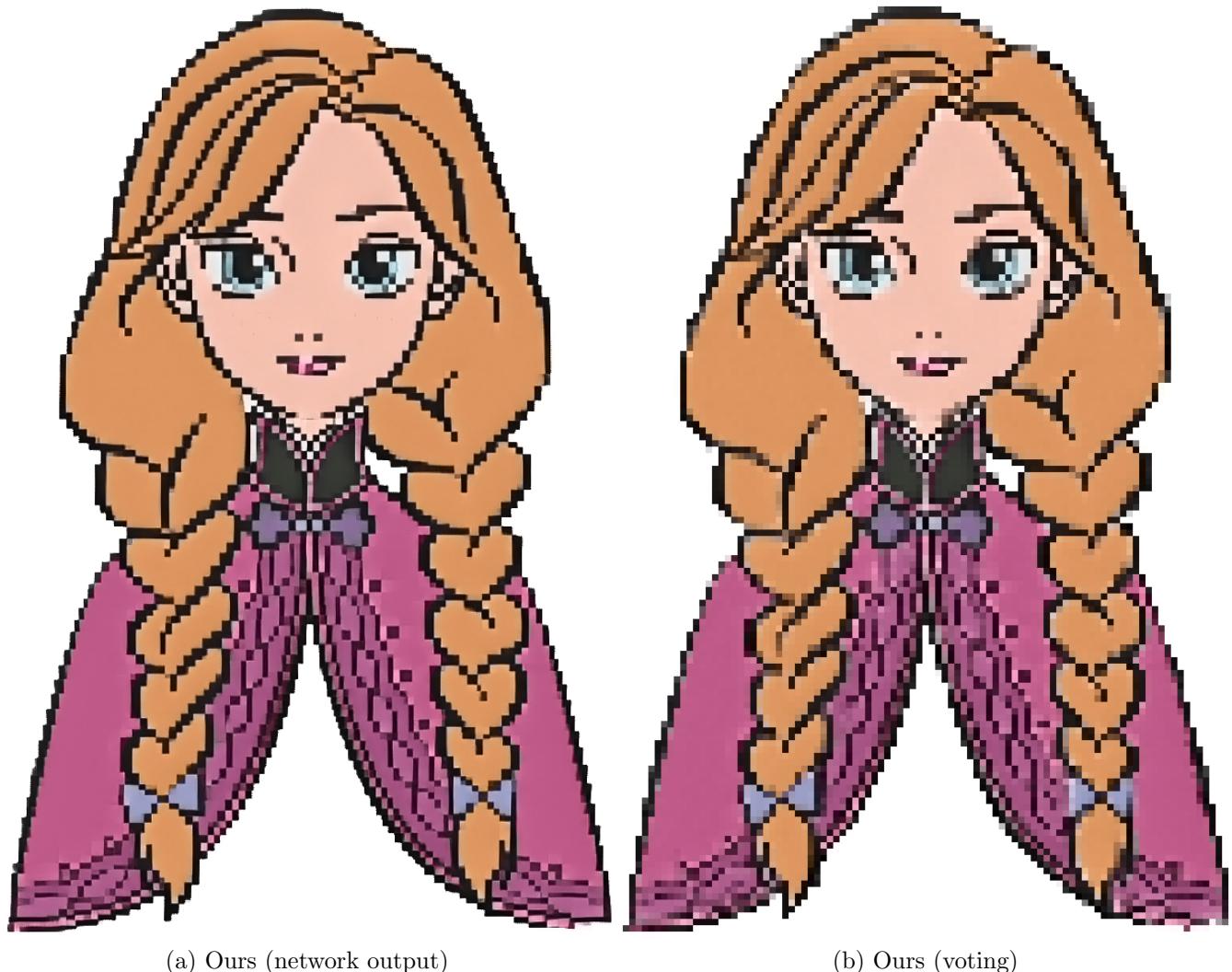


Figure 7: Voting result of teaser.