

Paint Transformer: Feed Forward Neural Painting with Stroke Prediction

(Supplementary Material)

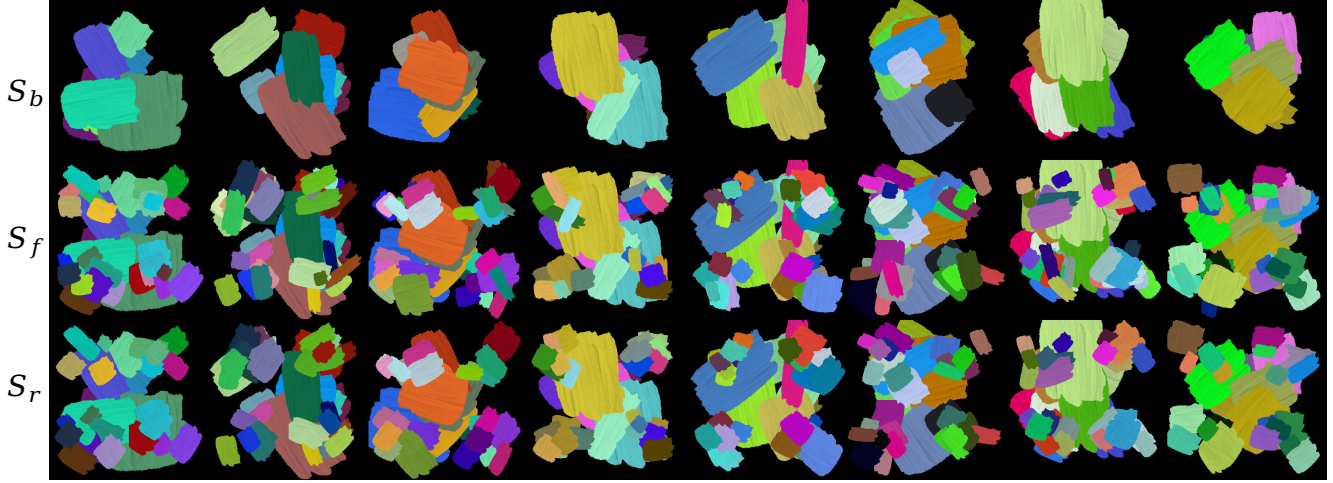


Figure 1. Canvas-target-predict pairs in training.

1. More Training Details

In our training, firstly, we sample 8 strokes on a canvas with 64×64 resolution as background. Then, the background canvas is divided into 4 blocks with 32×32 size. For each block, we further sample another 8 strokes as foreground ones based on the background. The stroke predictor learns to predict these extra strokes. Such operations encourage the stroke predictor to paint from coarse to fine. Thus, it always tends to generate refined strokes to minimize the differences between current canvas and target. During inference, the coarse-to-fine inference process can gradually fill in the canvas and reduce the differences

between canvas and the real image. Therefore, our stroke predictor can be generalized from randomly-synthesized dataset to real-world images successfully. More canvas-target-predict pairs (denoted as S_b , S_f , and S_r respectively) during training period are shown in Fig. 1.

2. More Inference Results

We provide more results including high-resolution (1024×1024) results of our algorithm in Fig. 2. The animated painting process can be found in the attachment or [our code page](#).



Figure 2. More inference result.