CLIPFont: Text Guided Vector WordArt Generation

songyiren@sjtu.edu.cn
zyx153@sjtu.edu.cn

Font design is a repetitive job that requires specialized skills. Unlike the existing few-shot font generation methods, this paper proposes a zero-shot font generation method called CLIPFont for any language based on the CLIP model. The style of the font is controlled by the text description, and the skeleton of the font remains the same as the input reference font. CLIPFont optimizes the parameters of vector fonts by gradient descent and achieves artistic font generation by minimizing the directional distance between text description and font in the CLIP embedding space. CLIP recognition loss is proposed to keep the category of each character unchanged. The gradients computed on the rasterized images are returned to the vector parameter space utilizing a differentiable vector renderer. Experimental results and Turing tests demonstrate our method's state-of-the-art performance. Project page: https://github.com/songyiren98/CLIPFont
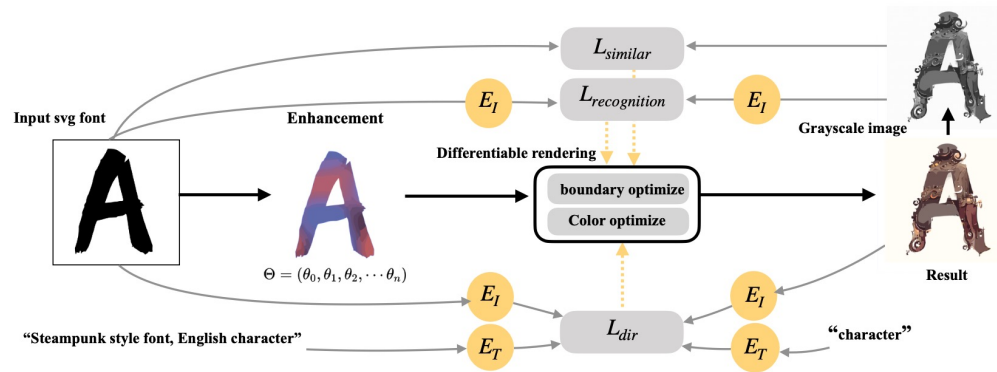
Fig.1 Overall schematics of CLIPFont. CLIPFont iteratively synthesizes stylized character through gradient descent. Starting from the input vector font, the color and boundary of polygons are optimized to best match the given textual prompt.

Fig.2 CLIPFont's WordArt generation results. Input a font, CLIPFont optimizes the shape boundary of each character to match the text description and can handle any language characters.