In this work, we devise two architectures that learn the features that encode a certain text style and are able to transfer them to other text instances while preserving their content. We thus recast the ideas of style transfer, previously applied mostly to paintings, to the text domain.

our method is able to automatically change the style of text regions in natural scene images, generating realistic images with the same textual content but with different text styles.

In machine printed text images, we are able to train models that stimulate a change of the text font. In handwritten text, we are able to transfer the writing style of a particular writer to another.

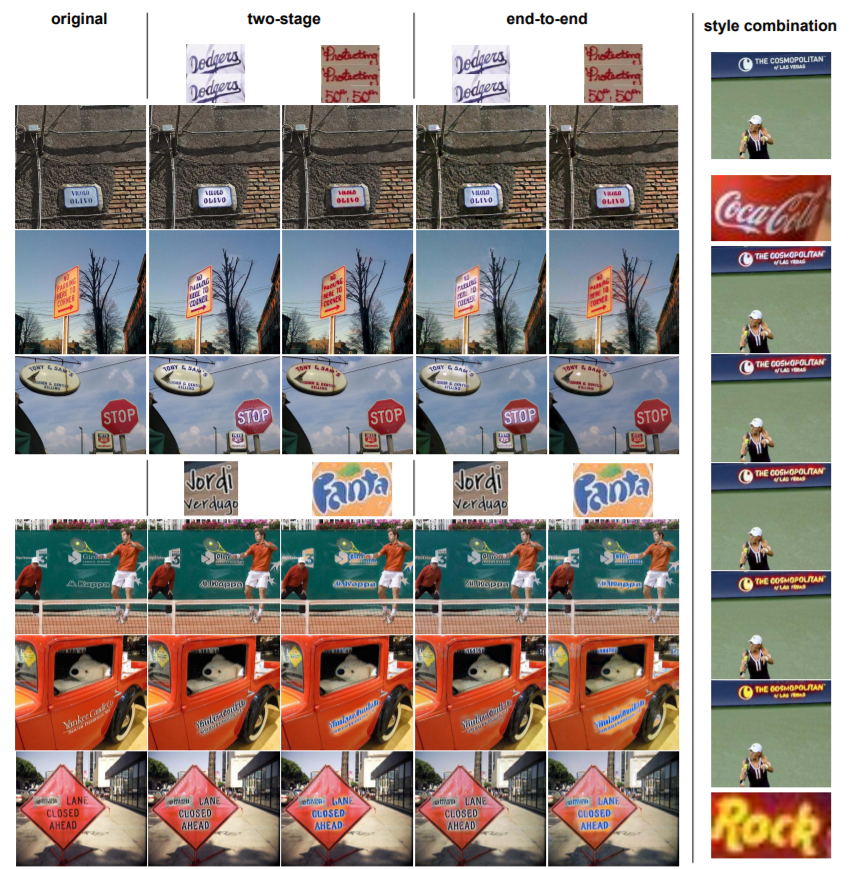
we will demonstrate that such an approach is useful as a data augmentation technique (see Figure 1), in order to deal with the problem of annotated data scarcity.

We use a single CNN model as our baseline. This model can learn to transfer different source styles (up to 32 in their experiments), allowing to generate images with combined styles.

EXPERIMENTS

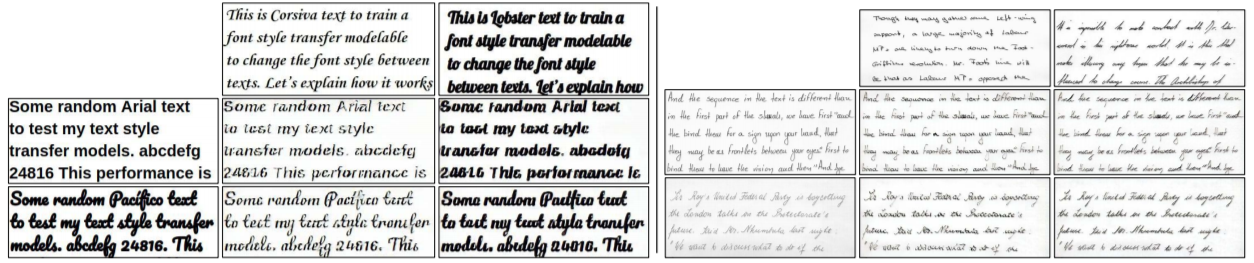
We train 3 models to explore the capabilities of text style transfer in various text domains: (1) a model to transfer scene text styles, (2) a model to transfer machine printed text fonts styles, and (3) a model to transfer handwritten styles.

(1) Scene Text: The performance is appealing, transferring the source styles with high fidelity in both character shapes and colors to a wide diversity of scene texts. The text content is preserved quite well in most images, and only in some cases where the task is very complex due to the original text size or tangled style the result is illegible. Figure 5 shows some of the results.

Figure 5: Applying different styles to various scene text images using the two-stage and end-to-end architectures with the cropped word styles (left). Extrapolating from “Coco-Cola” to “Rock” style on COCO-Text images (Right). (Best viewed in color.)

(2) Machine Printed Text: It successfully transfers the main features of the source font style, such as line width, text orientation, and main font character style. However, it fails transferring the specific styles of some characters, and the final output is influenced by the initial image. Figure 6 shows some of the results.

(3) Handwritten text: The model transfers correctly the main features of the text, as the tight characters and the thick stroke of the style in the first column, and the elongated and italic style of the writer in the second column. However, it fails on transferring more fine-grained characteristics of the source writer style, and some words of the resulting text are blurry. Figure 6 shows some of the results.

Figure 6: Results of the machine printed text model (left) and handwritten model (right). In each text domain, the styles of the images on top are transferred to the images on the left.

Cross Domain??

We have shown that a style transfer model is able to learn text styles as the characters shapes, line style, and colors, and to transfer it to an input text preserving the original characters. We have explored the performance of text style transfer in 3 text modalities: scene text, machine printed text and handwritten text and in cross-modal scenarios, proving the usefulness of text style transfer as a data augmentation technique to train scene text detectors. Cross-modal experiments show the potential of this pipeline in virtual reality scenarios to style an arbitrary text with a given scene style, and the realism of the generated images suggest that the pipeline could be useful in artistic applications, such as graphic design.