

GANwriting: Content-Conditioned Generation of Styled Handwritten Word Images

Supplementary Material

Lei Kang^{*†}, Pau Riba^{*}, Yaxing Wang^{*}, Marçal Rusiñol^{*}, Alicia Fornés^{*},
Mauricio Villegas[†]

^{*}Computer Vision Center, Universitat Autònoma de Barcelona, Spain
{lkang, pribal, yaxing, marcal, afornes}@cvc.uab.es

[†]omni.us, Berlin, Germany
{lei, mauricio}@omni.us.com

1 Video Interpolation

To better showcase the meaningfulness of the learned stylistic embedding space, we have attached a video where we animate a much finer interpolation than the one pictured in the paper, between different calligraphic styles of several words, composing the first sentence of Ernest Hemingway’s “*The Old Man and The Sea*”. We appreciate how the generator is able to provide a smooth transition between different writing styles for a given static content. We provide some screenshots of such video in Figure 1.

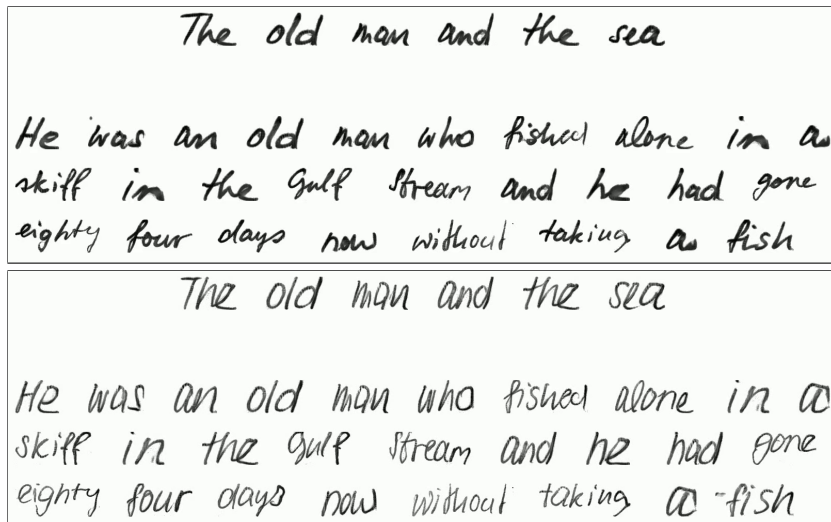


Fig. 1. Sample frames of the interpolation video.

2 Limitation of the proposed method when dealing with calligraphic styles

We evidence in Figure 2 the limitations of the proposed approach on imitating calligraphic styles. Unlike in [2], where characteristic glyphs from a given writer were manually cropped to perfectly compose a fraudulent text excerpt as if it was written by a certain person, our approach is not able to produce such levels of mimicking. When the model, trained with the IAM dataset, is fed with an unconventional calligraphic style, the proposed approach is not able to convey such stylistic aspects to the generated word samples. In Figure 2, we injected word samples written by Mary Shelley, and, the reader will appreciate how the rendered results are not able to imitate the visual aspect of such handwriting. However, the proposed generative method is still able to correctly render the textual contents, regardless of the provided calligraphic style.

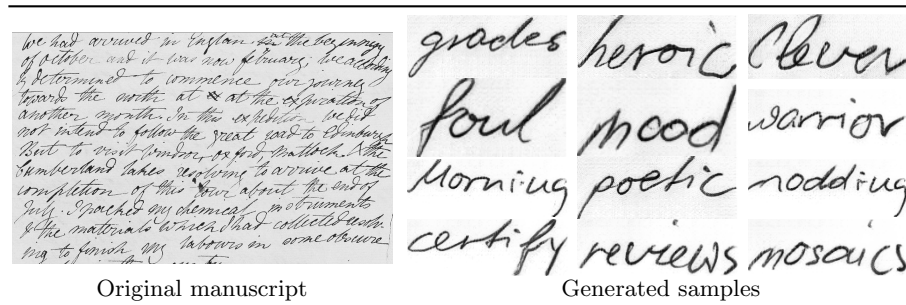


Fig. 2. Limitations of the proposed approach when mimicking Mary Shelley’s handwriting style.

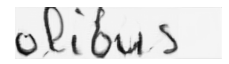
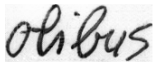
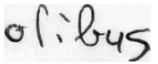
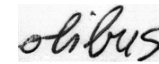




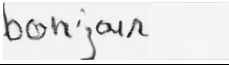

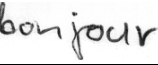

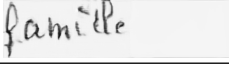

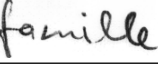
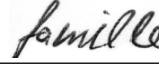
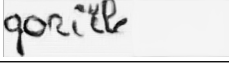
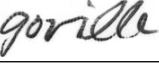
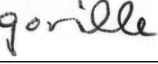
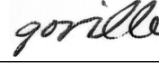
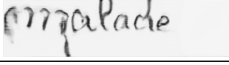
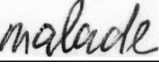
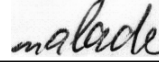

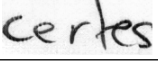
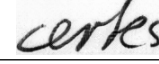

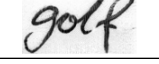
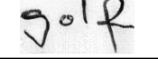
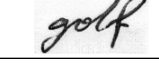


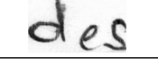

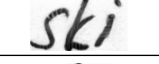
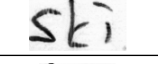
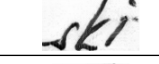

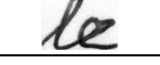
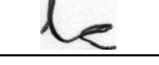
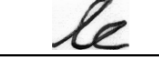
3 Qualitative comparison with Alonso *et al.* [1]

We present in Table 1, a qualitative comparison with the work of Alonso *et al.* [1]. We can appreciate how our proposed method clearly produces much credible generated images while being able to render the same content word with different calligraphic styles. Whereas [1] suffers from the mode collapse problem, always tending towards producing similar glyphs, our proposed method is able to yield different stylistic instances of the same textual content.

4 t-SNE Embedding visualizations

Due to space constraints, we are aware that the t-SNE plot presented in the paper in Figure 5 is shown at a quite small scale. This difficult its inspection. We provide here in Figures 3, 4, 5 and 6, four different t-SNE plots for images generated with the same textual content and for various calligraphic styles.

Table 1. Qualitative comparison with Alonso *et al.*. Images reprinted from [1].

Content	Alonso <i>et al.</i> [1]	Ours		
		Style A	Style B	Style C
"olibus"				
"reparer"				
"bonjour"				
"famille"				
"gorille"				
"malade"				
"certes"				
"golf"				
"des"				
"ski"				
"le"				

References

1. Alonso, E., Moysset, B., Messina, R.: Adversarial generation of handwritten text images conditioned on sequences. In: Proceedings of the International Conference on Document Analysis and Recognition (2019)
2. Haines, T.S., Mac Aodha, O., Brostow, G.J.: My text in your handwriting. ACM Transactions on Graphics **35**(3), 1–18 (2016)

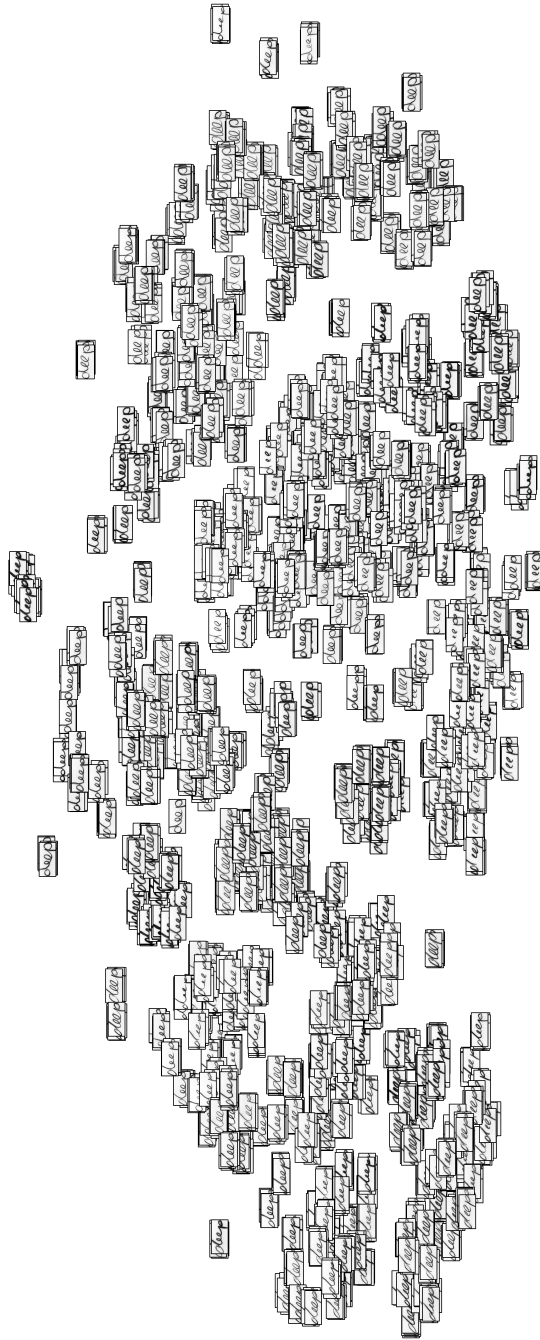


Fig. 3. t-SNE embedding visualization of 2,500 generated instances of the word "deep".

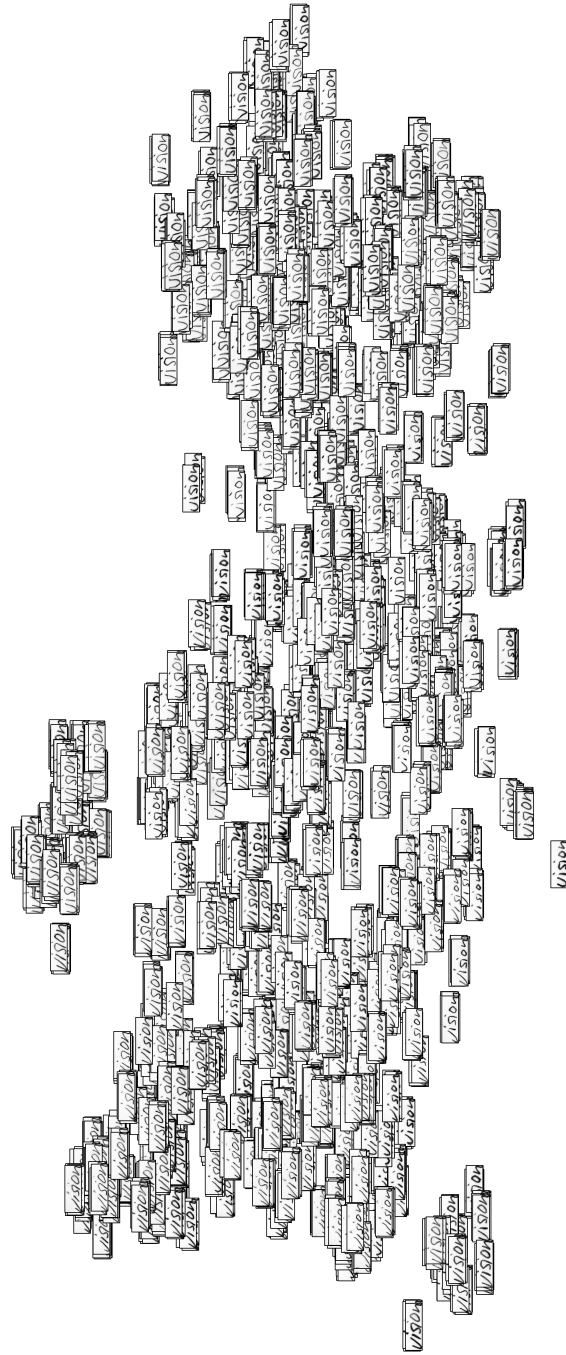


Fig. 4. t-SNE embedding visualization of 2,500 generated instances of the word "vision".

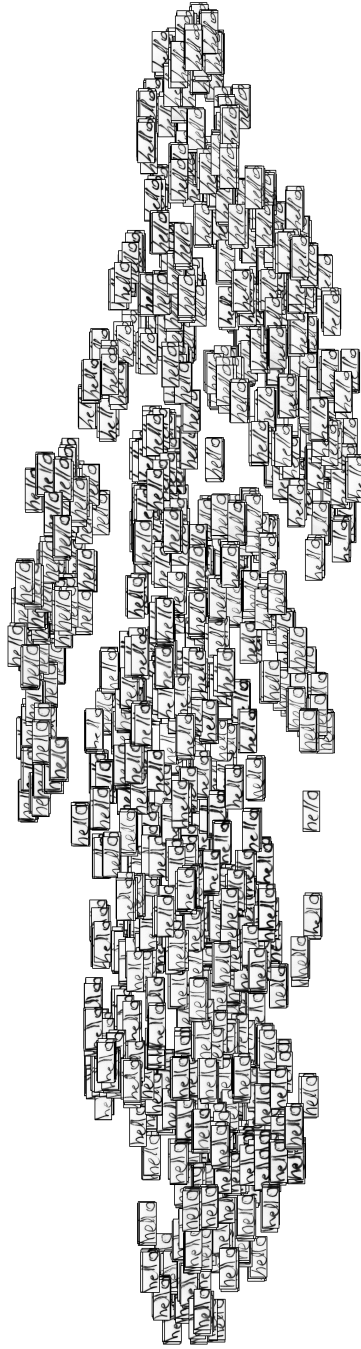


Fig. 5. t-SNE embedding visualization of 2,500 generated instances of the word "hello".

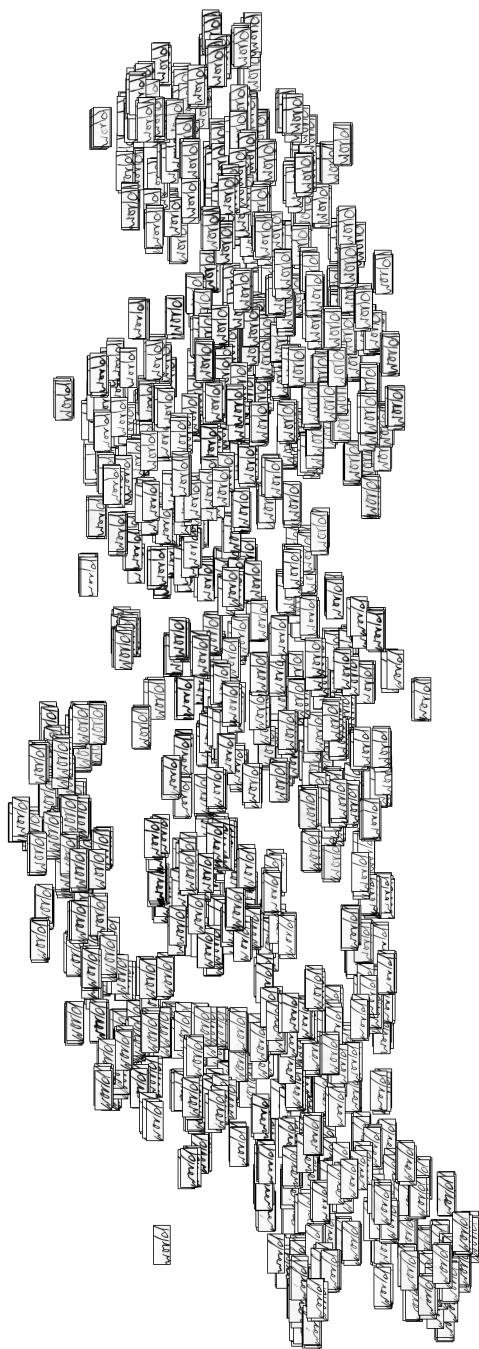


Fig. 6. t-SNE embedding visualization of 2,500 generated instances of the word "world".