

Demo: Pictory - Neural Style Transfer and Editing with CoreML

Sebastian Pasewaldt
Digital Masterpieces GmbH
sebastian.pasewaldt@digitalmasterpieces.com

Amir Semmo
Hasso Plattner Institute, Faculty of Digital Engineering,
University of Potsdam, Germany
amir.semmo@hpi.de

Mandy Klingbeil
Hasso Plattner Institute, Faculty of Digital Engineering,
University of Potsdam, Germany
mandy.klingbeil@hpi.de

Jürgen Döllner
Hasso Plattner Institute, Faculty of Digital Engineering,
University of Potsdam, Germany
juergen.doellner@hpi.de

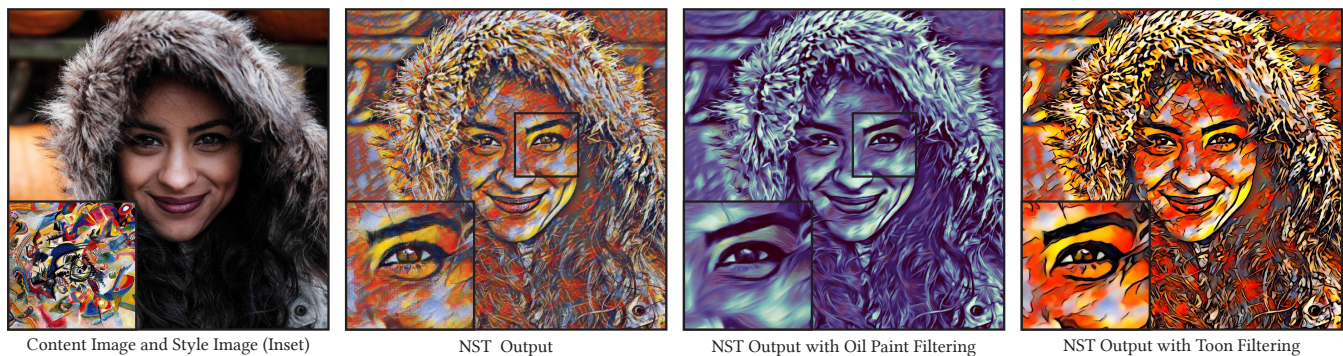


Figure 1: Exemplary results produced with *Pictory* for a 1 MP content image: neural style transfer (NST) of Wassily Kandinsky’s “Composition VII”, post-processing via oil paint filtering including color grading, and post-processing via toon filtering. Content image by Frank Marino and style image by Wassily Kandinsky are in the public domain.

ABSTRACT

This work presents advances in the design and implementation of *Pictory*, an iOS app for artistic neural style transfer and interactive image editing using the *CoreML* and *Metal* APIs. *Pictory* combines the benefits of neural style transfer, e.g., high degree of abstraction on a global scale, with the interactivity of GPU-accelerated state-of-the-art image-based artistic rendering on a local scale. Thereby, the user is empowered to create high-resolution, abstracted renditions in a two-stage approach. First, a photo is transformed using a pre-trained convolutional neural network to obtain an intermediate stylized representation. Second, image-based artistic rendering techniques (e.g., watercolor, oil paint or toon filtering) are used to further stylize the image. Thereby, fine-scale texture noise—introduced by the style transfer—is filtered and interactive means are provided to individually adjust the stylization effects at run-time. Based on qualitative and quantitative user studies, *Pictory* has been redesigned and optimized to support casual users as well as mobile artists by

providing effective, yet easy to understand, tools to facilitate image editing at multiple levels of control.

CCS CONCEPTS

• **Computing methodologies** → **Non-photorealistic rendering; Image processing;**

KEYWORDS

mobile, neural style transfer, image filtering, stylization

ACM Reference format:

Sebastian Pasewaldt, Amir Semmo, Mandy Klingbeil, and Jürgen Döllner. 2017. Demo: Pictory - Neural Style Transfer and Editing with CoreML. In *Proceedings of SA '17 Symposium on Mobile Graphics & Interactive Applications, Bangkok, Thailand, November 27-30, 2017*, 2 pages. <https://doi.org/10.1145/3132787.3132815>

1 MOTIVATION

Smartphones and tablet PCs have become one of the most popular platforms to capture and stylize photos¹. In particular, image stylization apps (e.g., Prisma, BeCasso) enjoy a growing popularity on mobile devices to foster casual creativity [Winnemöller 2013]. With the introduction of neural style transfer (NST) [Gatys et al. 2016]

SA '17 Symposium on Mobile Graphics & Interactive Applications, November 27-30, 2017, Bangkok, Thailand

© 2017 Copyright held by the owner/author(s).

This is the author’s version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *Proceedings of SA '17 Symposium on Mobile Graphics & Interactive Applications, November 27-30, 2017*, <https://doi.org/10.1145/3132787.3132815>.

¹<https://www.flickr.com/cameras> (accessed 6/2017)

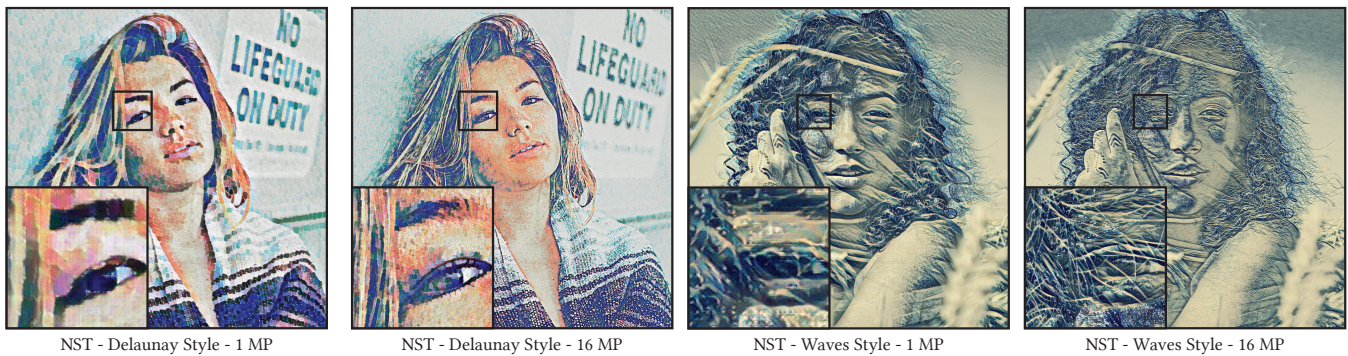


Figure 2: Comparison of input images processed at different resolutions: Although the images have been processed with the same NST configuration (Delaunay or Waves), the 16 MP output is characterized by finer details. Content images by Joe Gardner and Larisa Birta are in the public domain.

as a new approach for image-based artistic rendering, several publications and implementations have been proposed to improve its inherent limitations [Jing et al. 2017], such as high computational costs, high memory consumption and “black-box” processing—vital aspects to use NSTs as interactive tools on mobile devices.

2 TECHNICAL APPROACH

This paper presents ongoing developments in the design and implementation of *Pictory*, a mobile app that combines NSTs with other paradigms of image-based artistic rendering as proposed by Semmo et al. [2017a; 2017b]. This combined approach enables *Pictory* to transform images into high-quality artistic renditions on a global and local scale. Thereby, a NST is performed in a first processing stage to yield an intermediate result. Subsequently, this image is processed using interactive image filtering or processing techniques—e.g., watercolor rendering or oil paint filtering—to reduce fine-scale visual noise and locally inject characteristics of artistic media, which may be interactively refined by a user.

Compared to previous works [Jing et al. 2017] and publicly available apps, the following techniques are implemented and combined in a single mobile app:

- (1) GPU-accelerated implementations of Johnson et al.’s feed-forward NST [Johnson et al. 2016] and IB-AR techniques (e.g., oil paint filtering [Semmo et al. 2016]) are used to perform a style transfer at interactive frame rates and with low memory consumption. At this, the proposed implementation is based on Apple’s CoreML and Metal APIs for GPU-based processing that achieves interactive frame rates on modern devices. For example, the stylization of a 1 MP image including all processing stages takes 0.16 seconds on a first-generation iPad Pro (Apple A9X / PowerVR Series 7XT GPU).
- (2) To cope with the inherent scale-invariance (Figure 2) and memory limitations of NSTs on mobile devices, joint bilateral upsampling [Kopf et al. 2007] is used with the original input image. This way, high-resolution exports (up to 16 MP) can be provided and the visual discrepancies reduced (compared to a low-resolution preview) while maintaining interactive parameterizations of the stylization effects.

- (3) A hierarchical, multi-view GUI is implemented that supports multiple levels of control to support casual users and digital artists in their creative editing process. This ranges from a high level of control—by providing presets—over a low level of control—by global parameter adjustments—, to on-screen painting.

All three enhancements significantly facilitate the implementation of interactive tools that enable to adjust the stylization effects at run-time.

ACKNOWLEDGMENTS

We would like to thank Frank Schlegel, Moritz Hilscher and Hendrik Tjabben for their substantial contributions to the app prototype. This work was funded by the Federal Ministry of Education and Research (BMBF), Germany, for the AVA project 01IS15041.

REFERENCES

- Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2016. Image Style Transfer Using Convolutional Neural Networks. In *Proc. CVPR*. IEEE Computer Society, Los Alamitos, 2414–2423. <https://doi.org/10.1109/CVPR.2016.265>
- Yongcheng Jing, Yezhou Yang, Zunlei Feng, Jingwen Ye, and Mingli Song. 2017. *Neural Style Transfer: A Review*. arXiv.org report 1705.04058. <https://arxiv.org/abs/1705.04058>
- Justin Johnson, Alexandre Alahi, and Li Fei-Fei. 2016. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In *Proc. ECCV*. Springer International, Cham, Switzerland, 694–711. https://doi.org/10.1007/978-3-319-46475-6_43
- Johannes Kopf, Michael F. Cohen, Dani Lischinski, and Matt Uyttendaele. 2007. Joint Bilateral Upsampling. *ACM Trans. Graph.* 26, 3 (2007). <https://doi.org/10.1145/1276377.1276497>
- Amir Semmo, Tobias Isenberg, and Jürgen Döllner. 2017a. Neural Style Transfer: A Paradigm Shift for Image-based Artistic Rendering?. In *Proc. International Symposium on Non-Photorealistic Animation and Rendering (NPAR)*, Holger Winnemöller and Lyn Bartram (Eds.). ACM, New York, 5:1–5:13. <https://doi.org/10.1145/3092919.3092920>
- Amir Semmo, Daniel Limberger, Jan Eric Kyprianidis, and Jürgen Döllner. 2016. Image Stylization by Interactive Oil Paint Filtering. *Computers & Graphics* 55 (2016), 157–171. <https://doi.org/10.1016/j.cag.2015.12.001>
- Amir Semmo, Matthias Trapp, Jürgen Döllner, and Mandy Klingbeil. 2017b. *Pictory: Combining Neural Style Transfer and Image Filtering*. In *Proc. SIGGRAPH Appy Hour*. ACM, New York, 5:1–5:2. <https://doi.org/10.1145/3098900.3098906>
- Holger Winnemöller. 2013. NPR in the Wild. In *Image and Video based Artistic Stylization*, Paul Rosin and John Collomosse (Eds.). Computational Imaging and Vision, Vol. 42. Springer, Chapter 17, 353–374. https://doi.org/10.1007/978-1-4471-4519-6_17