Deep neural networks for image segmentation in fine art

Dr. David G. Stork

1. Tell us about yourself

*Dr. Stork is widely considered a pioneer in the application of computer vision, image analysis, machine learning, and artificial intelligence to problems in the history and interpretation of fine art. He is a graduate of MIT and the University of Maryland and studied Art History at Wellesley College. He has held faculty positions in Physics, Mathematics, Computer Science, Electrical Engineering, Statistics, Neuroscience, Psychology, and Art and Art History variously at Wellesley and Swarthmore Colleges and Clark, Boston, and Stanford Universities. His 200+ scholarly publications, eight books (including **Pattern classification**, 2nd ed. by Duda, Hart, and Stork), and 54 US patents have garnered 80,000 citations. He is a Fellow of IEEE, OSA, SPIE, IS&T, IAPR and IARIA and is completing **Pixels & paintings: Foundations of computer-assisted connoisseurship** (Wiley).*

2. Give a brief description of the project

Images of fine art paintings and drawings are some of the most iconic and important images ever created and pose profound and consequential problems to scholars and the general public alike. Recently, computer vision, machine learning, and artificial intelligence have helped solve a number of problems in the history and study of paintings and drawings, such as authentication, interpretation, inferring artists' working methods, and more; such technology promises to revolutionize the study of art.

*This liveProject will extend these early technical successes to the problem of **segmentation** of art images. Segmentation—labeling each pixel in an image such as "road," "person," "sky," "building," etc.—is an essential first step in many automated tasks, including autonomous driving and remote image analysis. Recent efforts in deep learning with massive databases of photographs have led to state-of-the-art segmentation performance on natural photographs. However, artworks differ in numerous ways from such common photographs, satellite images and medical images, for instance in the statistics of color, form, style, and composition, especially amidst variations in style of artwork (Impressionist, Expressionist, Renaissance, abstract, ...).*

This project will build upon current deep network methods for semantic image segmentation and extend them to art images, all in order to provide valuable software tool for art scholars, as well as to shed light upon deep neural networks and how they work.

3. Who the project is for and what they will learn?

This project is ideal for computer engineers with interest in computer vision, deep learning, big data, image analysis, image statistics, artificial intelligence, especially if they are seeking a research career in machine learning, artificial intelligence in image analysis.

4. Project Outline

Milestone 1: Benchmark exisiting segmentation networks on art images

- *Hand label the segmentation of small database of art images (paintings and drawings) to serve as test data*
- *Quantify performance (region accuracy and label accuracy) of existing segmentation networks applied to realistic art images*
- *Quantify performance (region accuracy and label accuracy) of existing segmentation networks applied to Impressionist and Expressionist art images*

Libraries/tools used:

- *Segmentation data*
- *Segmentation code on GitHub*
- *Online segmentation sites*

Deliverable:

- *Segmentation performance; analysis of the role of shape, color, and style upon segmentation accuracy*

Milestone 2: Further train exisiting (trained) segmentation networks with art images; test network performance on art images

- *Automatically scrape art images (of a particular style, e.g., Impressionist) from the web*
- *Train existing segmentation code (available through GitHub) with unsegmented art images*
- *Automatically scrape art images (of a particular style, e.g., Impressionist) from the web*
- *Train existing segmentation code (available through GitHub) with unsegmented art images*

Libraries/tools used:

- *Segmentation data*
- *Segmentation code on GitHub*

• Deep network visualization tools

Deliverable:

• Segmentation performance; analysis of the role of shape, color, and style upon segmentation accuracy
• Comparison of segmentation performance to that in Milestone 1; statistical analysis of the improvement in segmentation accuracy due to additional training data

Milestone 3: Optimize and test segmentation networks with art images; analyze trained networks to understand image features that lead to improved segmentation performance

• Automatically scrape art images (of a particular style, e.g., Impressionist) from the web
• Train existing segmentation code (available through GitHub) with unsegmented art images

Libraries/tools used:

• Tools for scraping web
• Segmentation GitHub
• Online segmentation sites

Deliverable:

• Quantify segmentation performance on wide variety of art images
• Comparison of optimal network with those used in Milestone 1 to identify what network properties most determine improved performance
• Draft report, which could be expanded into publishable paper

5. Recommended Reading

• "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," by Vijay Badrinarayana, Alex Kendall, and Roberto Cipolla, IEEE Transactions on pattern analysis and machine intelligence ***39****(12):2481–2495 (2017)*
• "Scene parsing through ADE20K dataset," Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 633–641 (2017)
• Chapter 2: Image processing from ***Pixels & paintings: Foundations of computer-assisted connoisseurship*** *by David G. Stork (manuscript in preparation)*
• ***Deep learning with PyTorch****, Eli Stevens, Luca Antiga and Thomas Viehmann (Manning, 2020)*

- ***Deep learning for vision systems****, Mohamed Elgendy (Manning, 2020)*
- ***Probabilistic deep learning****, Oliver Dürr, Beate Sick, and Elvis Murina (Manning, 2020)*
- ***Pattern classification*** *(2nd ed.), by Richard O. Duda, Peter E. Hart, and David G. Stork (Wiley, 2001)*