**Automated Interpretation of Digital Pathology Images**

*Joseph Picone*

Professor, Department of Electrical and Computer Engineering

Co-Director, The Neural Engineering Data Consortium

College of Engineering, Temple University

Philadelphia, Pennsylvania, USA

picone@temple.edu

**Abstract:**

Whole-slide imaging in digital pathology is rapidly advancing the need for automated digital image analysis. Ever-evolving guidelines for manual review of pathology images underscore the importance of AI-based algorithms that increase efficiency, improve productivity, and reduce fatigue. In this talk, we will review a collaboration between Fox Chase Cancer Center, Temple University Hospital, and the College of Engineering to detect and characterize cancerous cells in digitized images of pathology slides. There are two main components to the project: (1) the development of an open source corpus of 100,000 digital pathology images to support training and development of state of the art machine learning systems, and (2) the development of deep learning-based technology to automatically segment and classify images.

We will discuss the pipeline we have developed for digitizing and annotating pathology images and introduce a machine learning system that segments breast tissue images and performs patch-level classifications at accuracies above 80%. Digital pathology is one of the grand challenges of the current generation of deep learning systems that have been enabled by breakthroughs in machine learning algorithms for large, complex networks. There are two reasons this is an extremely challenging problem for deep learning. First, the images are extremely high resolution (50K x 50K pixels). The events of interest in these images are extremely small and highly confusable – often referred to as a needle in the haystack problem. Second, the amount of data, which can approach petabytes, makes it a great computational challenge. Data must be carefully staged, and algorithms must be designed to support incremental training and large scale multiprocessing to make the research feasible.

The data and open source tools developed in this project will accelerate pathologist productivity by prioritizing slides that need manual review. Such a translational capability will be transformative and evolve into additional areas of medical science involving images (e.g., prostate cancer) and multi-channel signals (e.g., electroencephalograms).

**Biography:**

Dr. Joseph Picone received his Ph.D. in Electrical Engineering in 1983 from the Illinois Institute of Technology. He is currently a professor in the Department of Electrical and Computer Engineering at Temple University. He has spent significant portions of his career in academia (MS State), research (Texas Instruments, AT&T) and the government (NSA). His primary research interests currently are applications of machine learning in the health sciences. Dr. Picone’s research funding sources over the years have included NSF, NIH, DoD, DARPA as well as the private sector. He has also been involved in several startup companies in healthcare. For over 40 years, his research groups have been known for producing many innovative open source materials for education and research, including the first state of the art public domain speech recognition system and the TUH EEG Corpus, which has over 7,000 subscribers (see [www.isip.piconepress.com](http://www.isip.piconepress.com)). He is currently developing a large corpus of 100,000 annotated digital pathology images. Dr. Picone is a Senior Member of the IEEE, holds several patents in human language technology.