Application of machine learning to skin cancer detection and classification

Andrew C. Terentis*[1], Xingquan Zhu[2], John M. Strasswimmer[3]
[1]Dept. Chem. & Biochemistry, [2]Dept. Computer & Elec. Engineering, FAU, Boca Raton
[3]Dermatology Associates of the Palm Beaches, Delray Beach, FL

We are seeking to utilize the techniques of machine learning for rapid, automated detection of residual skin cancer using Raman spectroscopy following partial laser ablation of the tumor. In a preliminary study we obtained twenty-five tissue samples from eleven patients undergoing Mohs surgery to remove squamous cell carcinomas (SCC). A total of 147 Raman spectra were collected from untreated and partially ablated normal and SCC tissue samples. Principal component analysis (PCA) of the spectra followed by a binary logistic regression (BLR) using the first five PCs as predictors gave 92% sensitivity and 60% specificity for cancer detection for non-ablated samples, and 95% sensitivity and 100% specificity for ablated samples. We also tested a preliminary long short-term memory (LSTM) based neural network and a generic neural network on the same Raman data set for comparison. The results showed that by using LSTM recurrent neural network the classification is significantly improved over the generic neural network, from 71.1% to 91.1% accuracy. These preliminary results demonstrate the feasibility of using Raman spectroscopy to detect residual cancer within partially ablated tumors using machine learning techniques. We are currently working to generate larger training data sets and exploring other machine and deep learning approaches to Raman spectroscopic data analysis.