MACHINE LEARNING APPLIED TO RAMAN SPECTRA OF PANCREATIC CANCER CELLS TO IDENTIFY MOLECULAR SIGNATURES, PATTERNS AND PROTEIN EXPRESSIONS.

CHRISTOPHER MANDRELL, *Physics, Southern Illinois University Carbondale, Carbondale, IL, USA*; SAKINEH E.M. ABADI, FARHAN CHOWDHURY, *Mechanical Engineering and Energy Processes, Southern Illinois University-Carbondale, Carbondale, IL, USA*; P SIVAKUMAR, *Physics, Southern Illinois University Carbondale, Carbondale, IL, USA*.

Pancreatic cancer is the third leading cause of cancer-related deaths in the United States, with the life expectancy for patients diagnosed in the late stage ranging from 6-12 months. This study is part of a group effort to examine the hypothesis that, difficult to detect, tumor initiating cells (TIC) exist in small numbers in the solid tumors and are responsible for the cancer's progression and relapse. The pancreatic cell line MIA PaCa-2 is seeded and grown in 90 Pa 3D fibrin gels for 10 days then samples are collected and Raman spectra obtained from a 784nm laser in the 150-1800nm and 2500-3500nm ranges. These spectra are analyzed via combinations of data preprocessing, wavelength or dimension reduction and machine learning classification algorithms. We extend to other cell lines such as CFPAC-1 and PANC-1. Support vector machine (SVM) and k-nearest neighbors (kNN) supervised machine learning classifiers are applied to the raw and pre-processed data sets and with various statistical and machine learning dimension reduction protocols. These combinations are compared to determine which performed best at classifying cancer and normal cell samples, and which led to selection of the same or similar dimensions. Identification of the best performing dimensions/wavelengths is then attempted from the Raman spectra by comparing them to existing biological molecule Raman databases to identify the patterns in the spectra and any unique molecular signatures or protein expressions that could prove useful to better understanding and therefore treating pancreatic cancer.