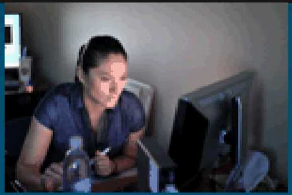
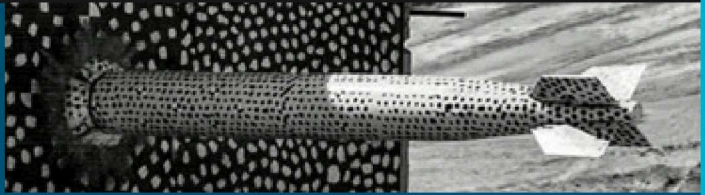
SAND2018-3961C

Machine and Deep Learning Exploration for Spectral X-ray Computed Tomography Classification Applications





PRESENTED BY

Dr. Edward S. Jimenez

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Acknowledgements to our team

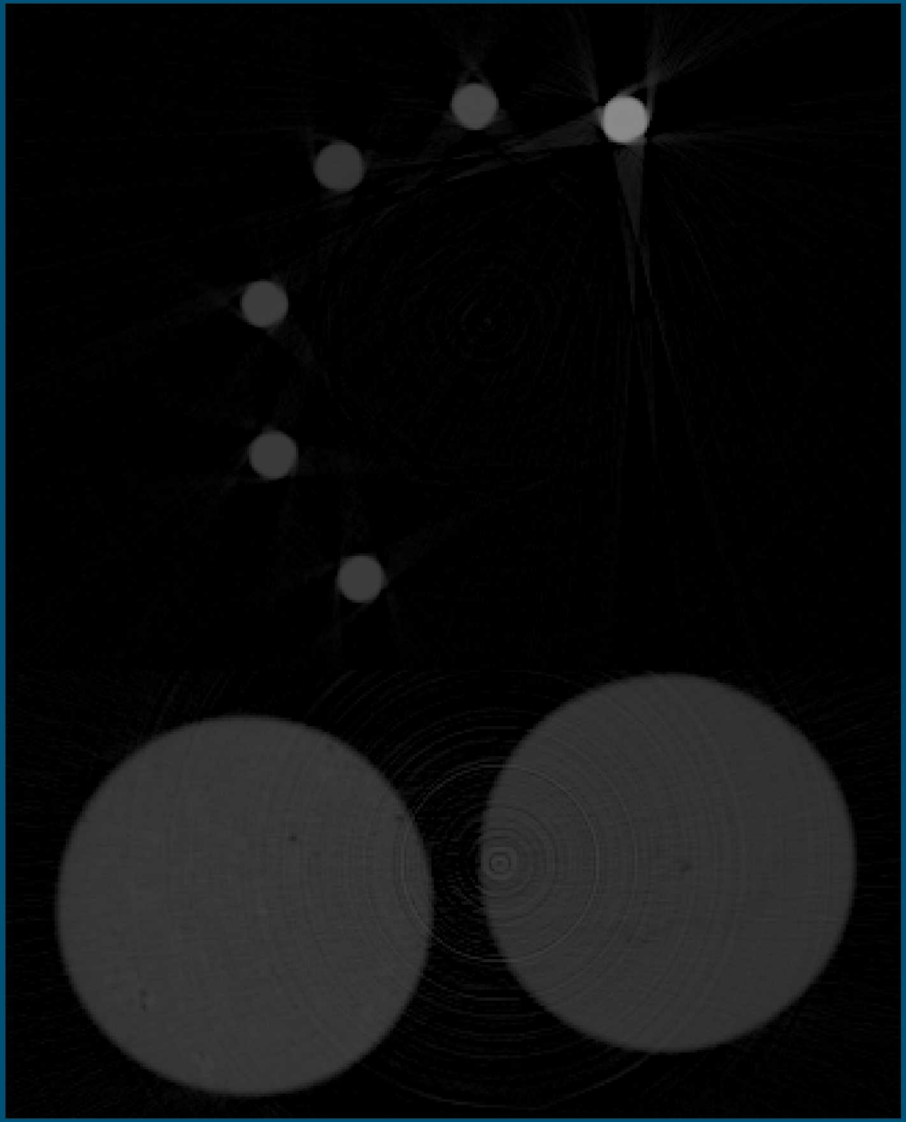
- Srivathsan Koundinyan
- April Suknot
- Ryan Goodner
- Kyle Thompson
- Gabriella Dalton

Ushering in the Age of Color Computed-Tomography

- This work seeks to develop a new capability to non-destructively discriminate between similar materials leveraging nascent technology and data science.

- Impossible with current state-of-the-art.

- Sandia National Laboratories has developed an energy-resolved CT system for R&D for industrial applications.

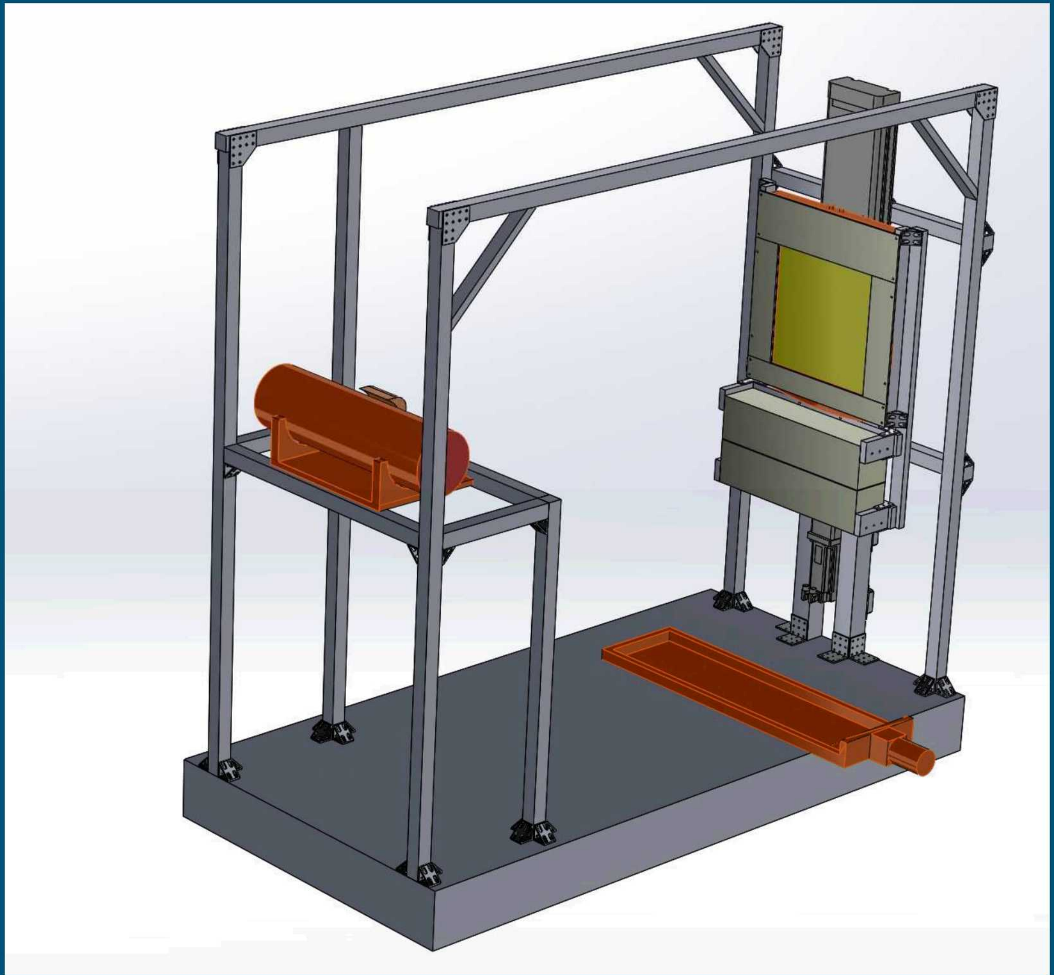


Unique Capability

- Practical high-energy spectral x-ray
  - Double the energy.
  - Data acquisition in approximately 30 minutes.

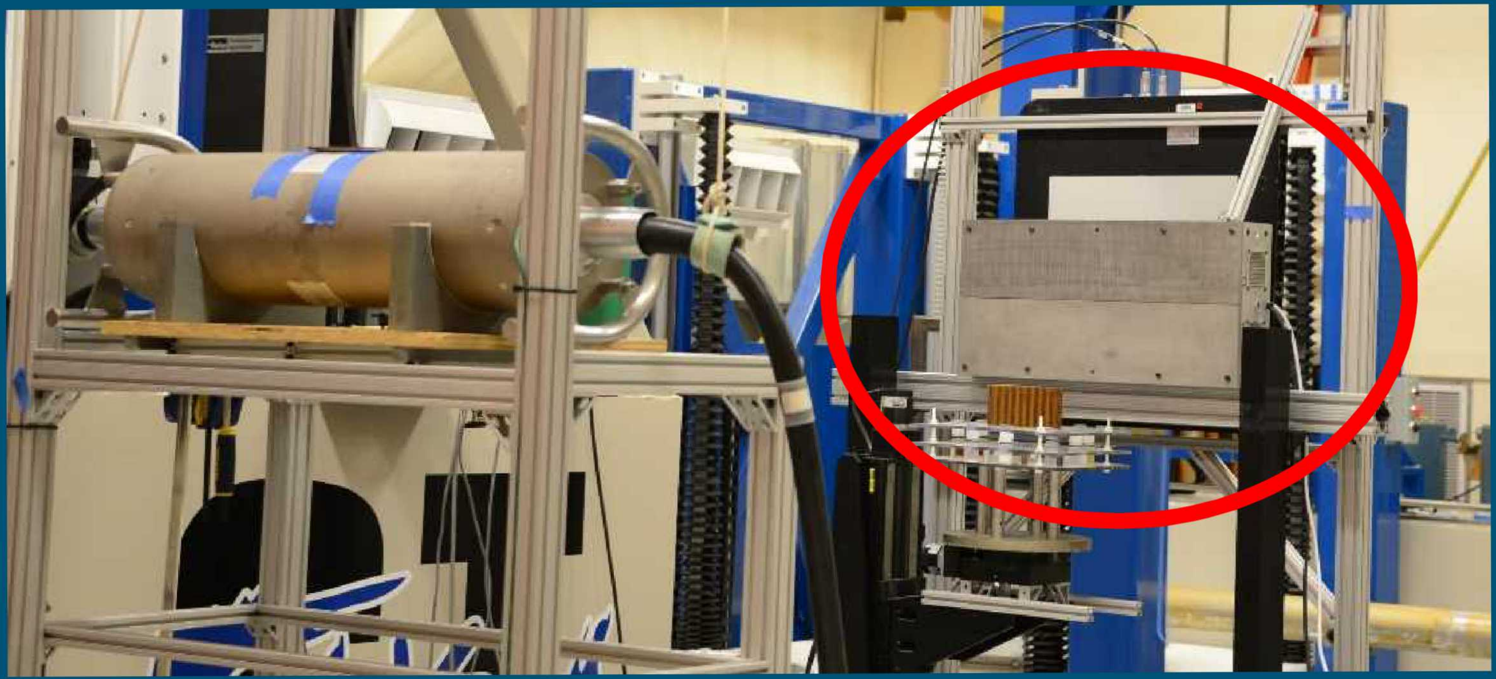
- Revolutionary Machine Learning Applications to identify similar materials with little training.

- Irregular Algorithm development to specifically address this big-data application.
  - 100x more data than traditional approaches
  - High performance reconstruction algorithms



Spectral CT at Sandia National Laboratories

- Spectral Detector.
  - Calibrated for 300keV
  - 640 pixels wide
  - 0.8 mm pixel pitch
  - 128 Channels



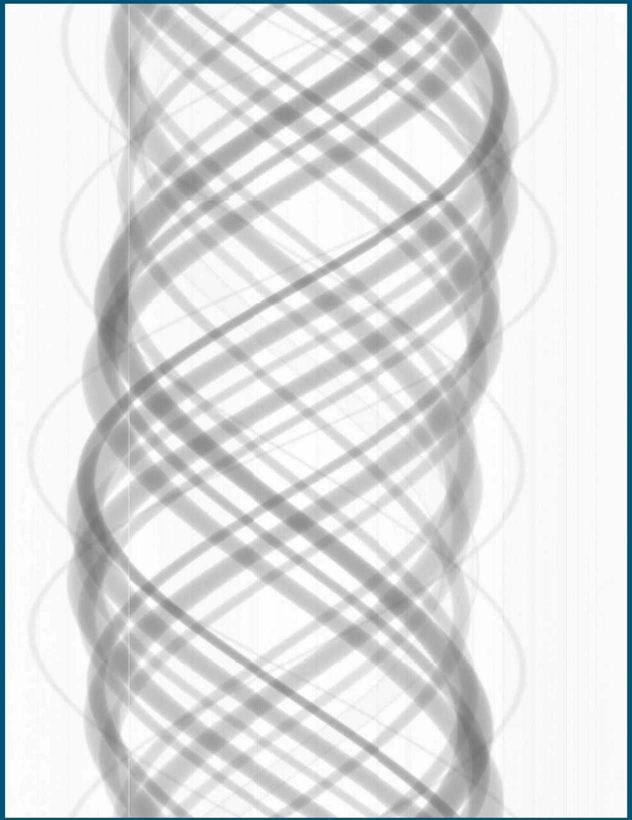
- 4 axis motion control
  - 3 axis object manipulator
  - 1 axis detector stage

- FOV up to half meter wide
  - 9 feet tall!

- In use for data acquisition as of May 2017

Data– Traditional vs. Spectral X-ray Input Data

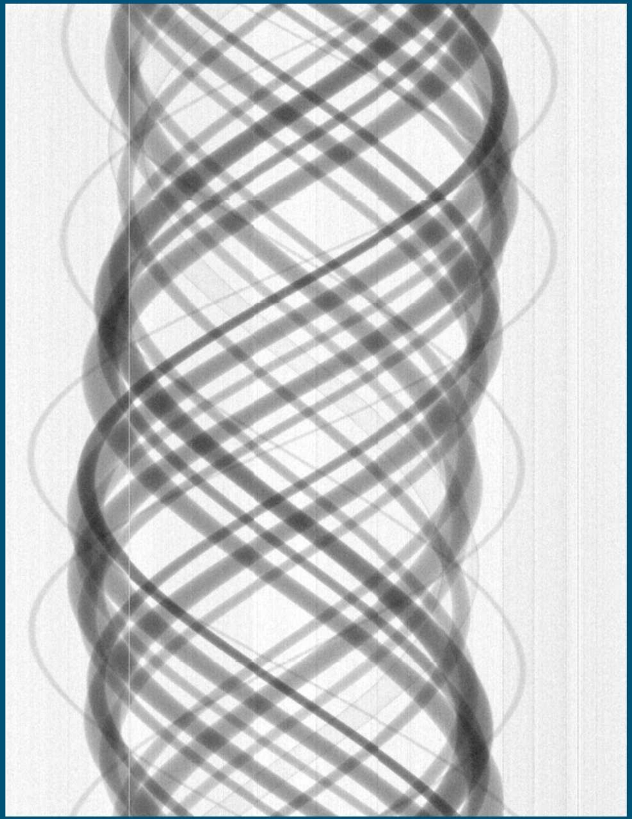
Traditional X-ray image



Traditional X-ray Image –Single Gray-scale Image per Scan

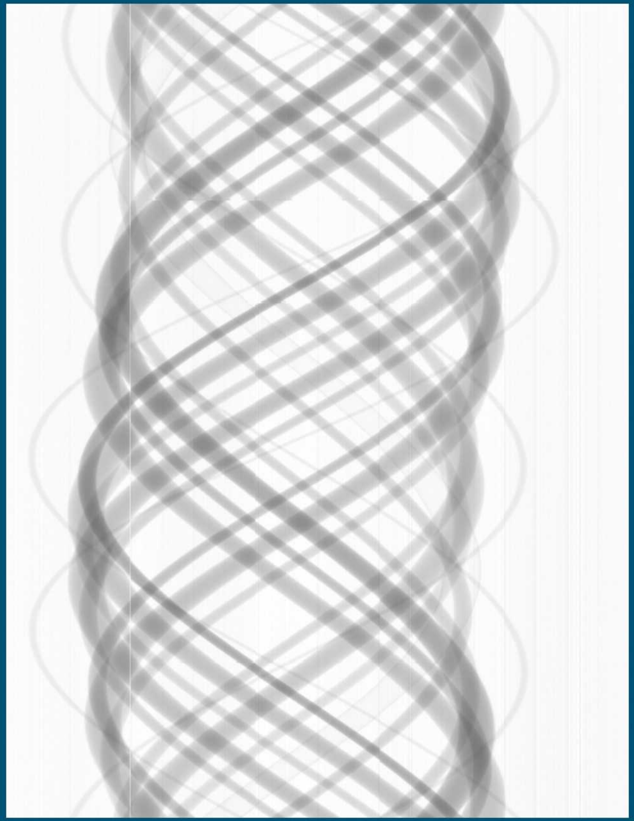
Data– Traditional vs. Spectral X-ray Input Data

Bin 0 – 2 keV



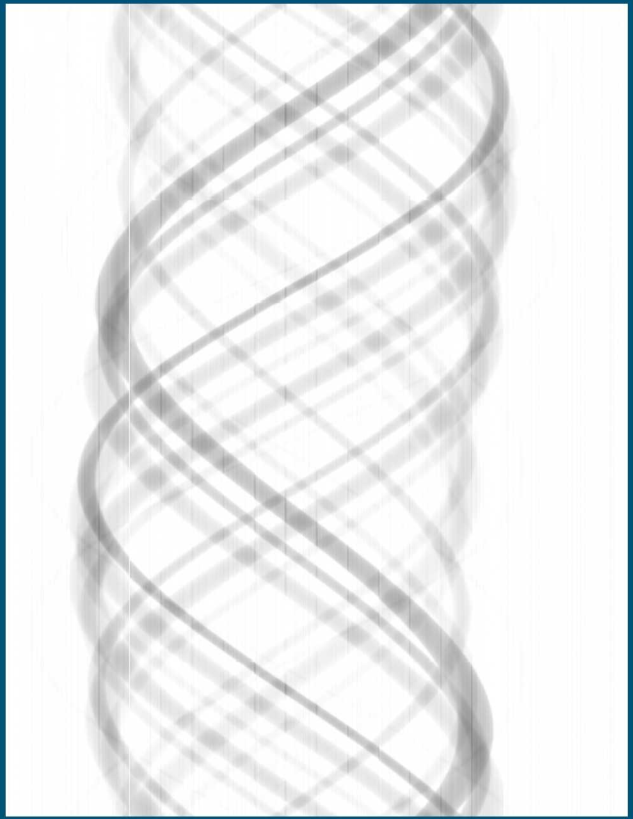
…

Bin 63 – 150 keV



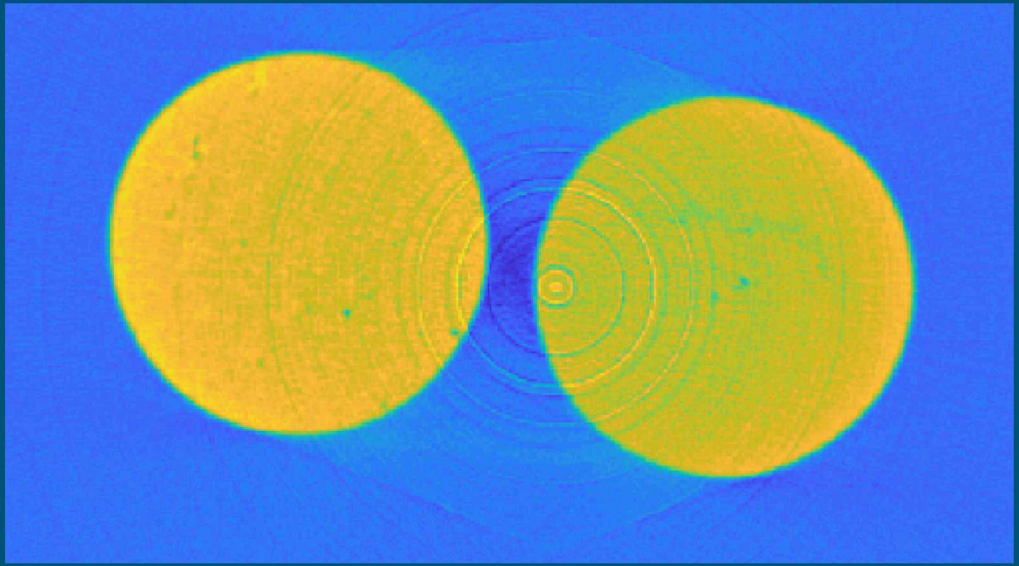
…

Bin 127 – 300 keV



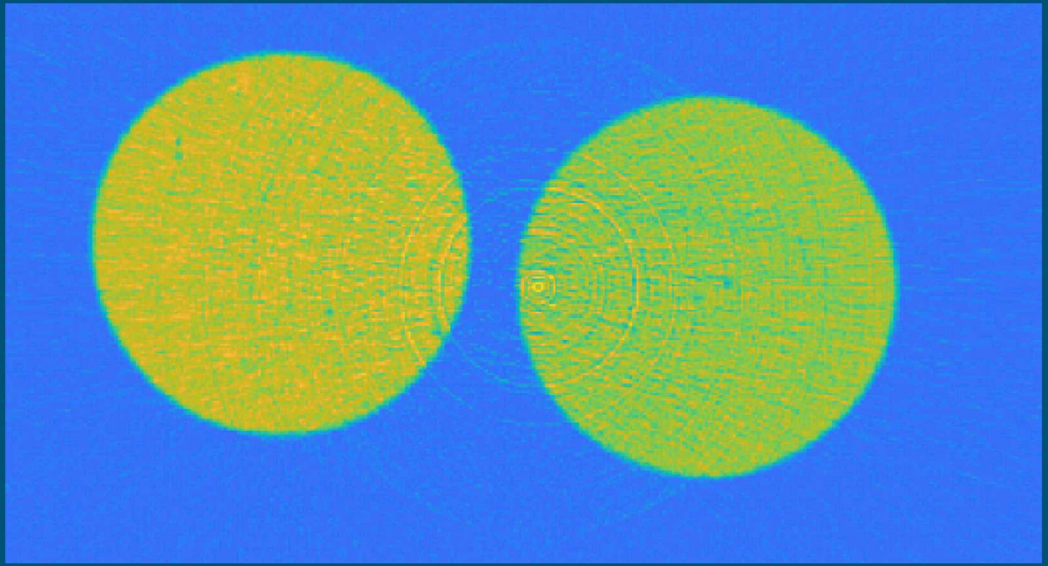
Spectral X-ray Images – 128 Images per Scan

Reconstruction - Even Simple Objects have Artifacts!



Traditional Reconstruction

- Summed bin data
- Streaking emanating from objects
- Beam hardening

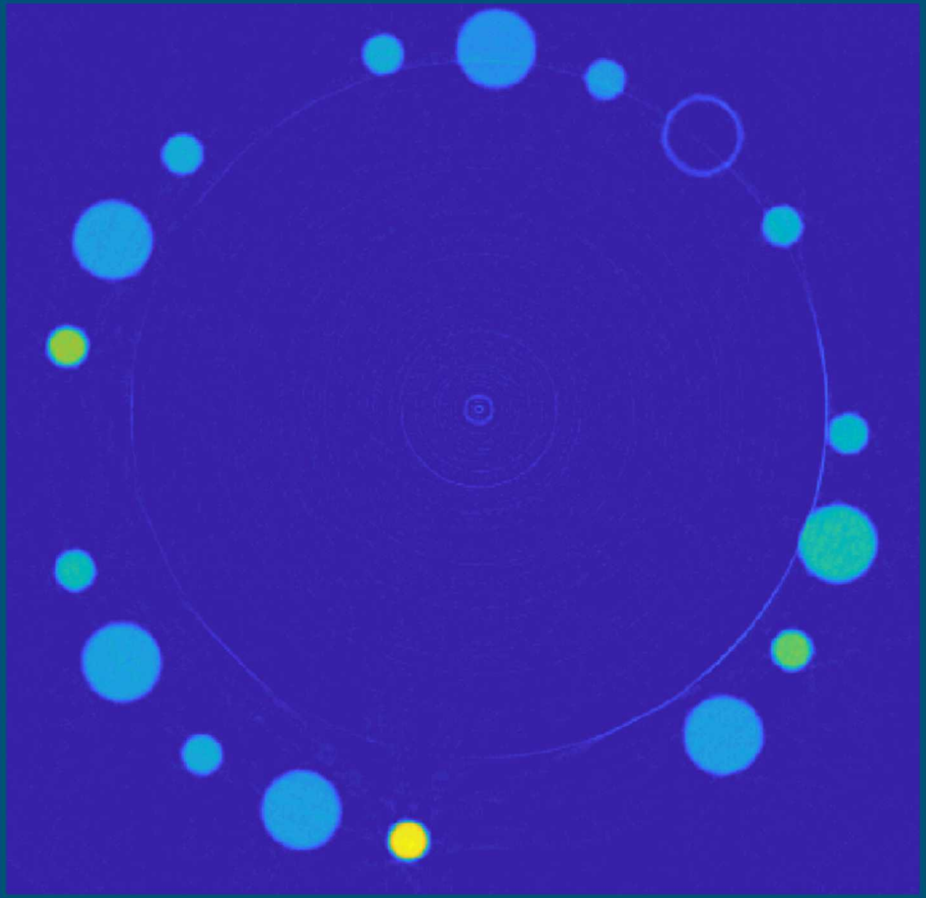


Single Bin Reconstruction

- Identical Reconstruction Algorithm
- Reduced Artifacts
- Uniform Sample Value

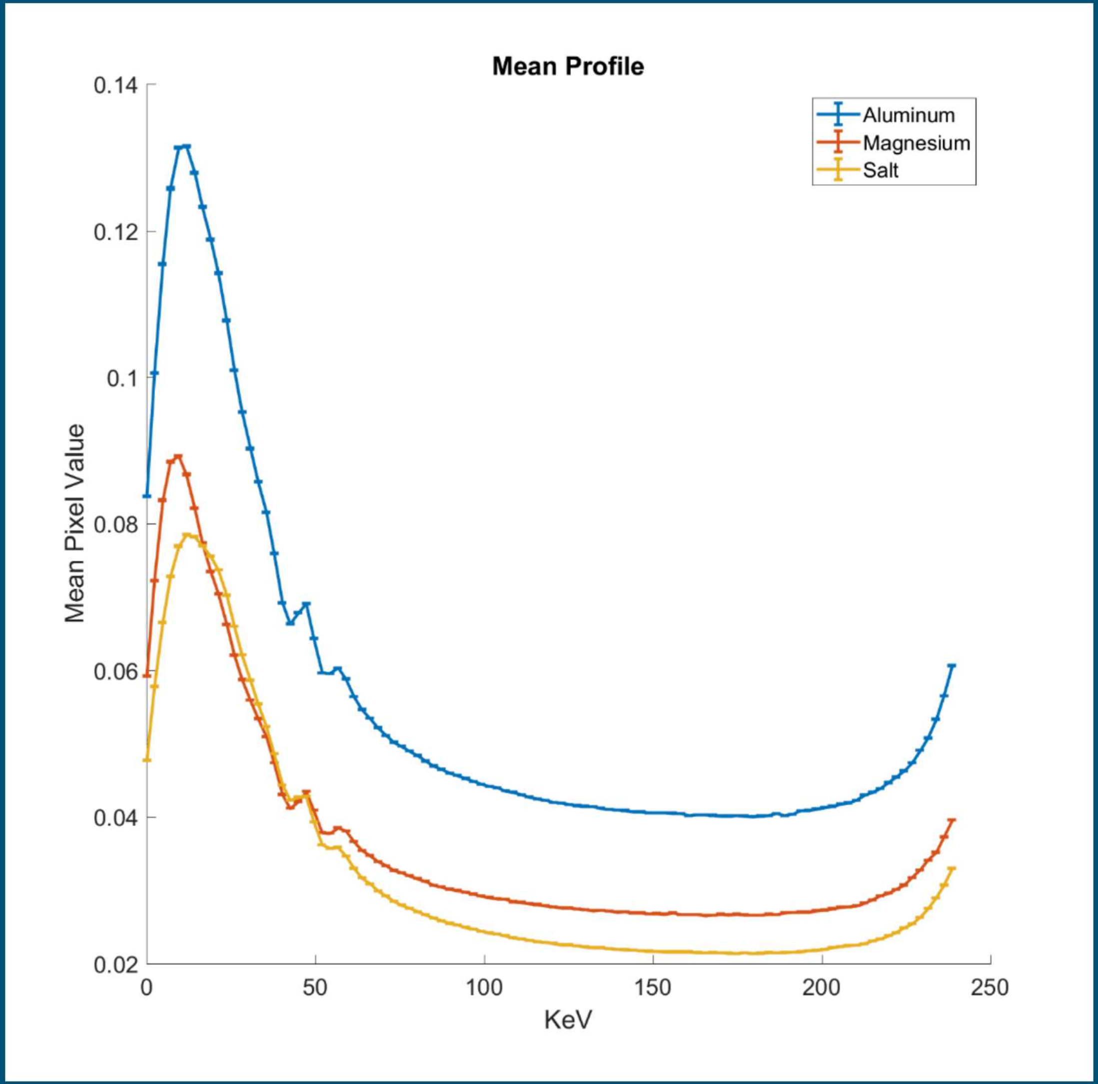
Broad Spectrum of Material: Various Samples

17 Materials



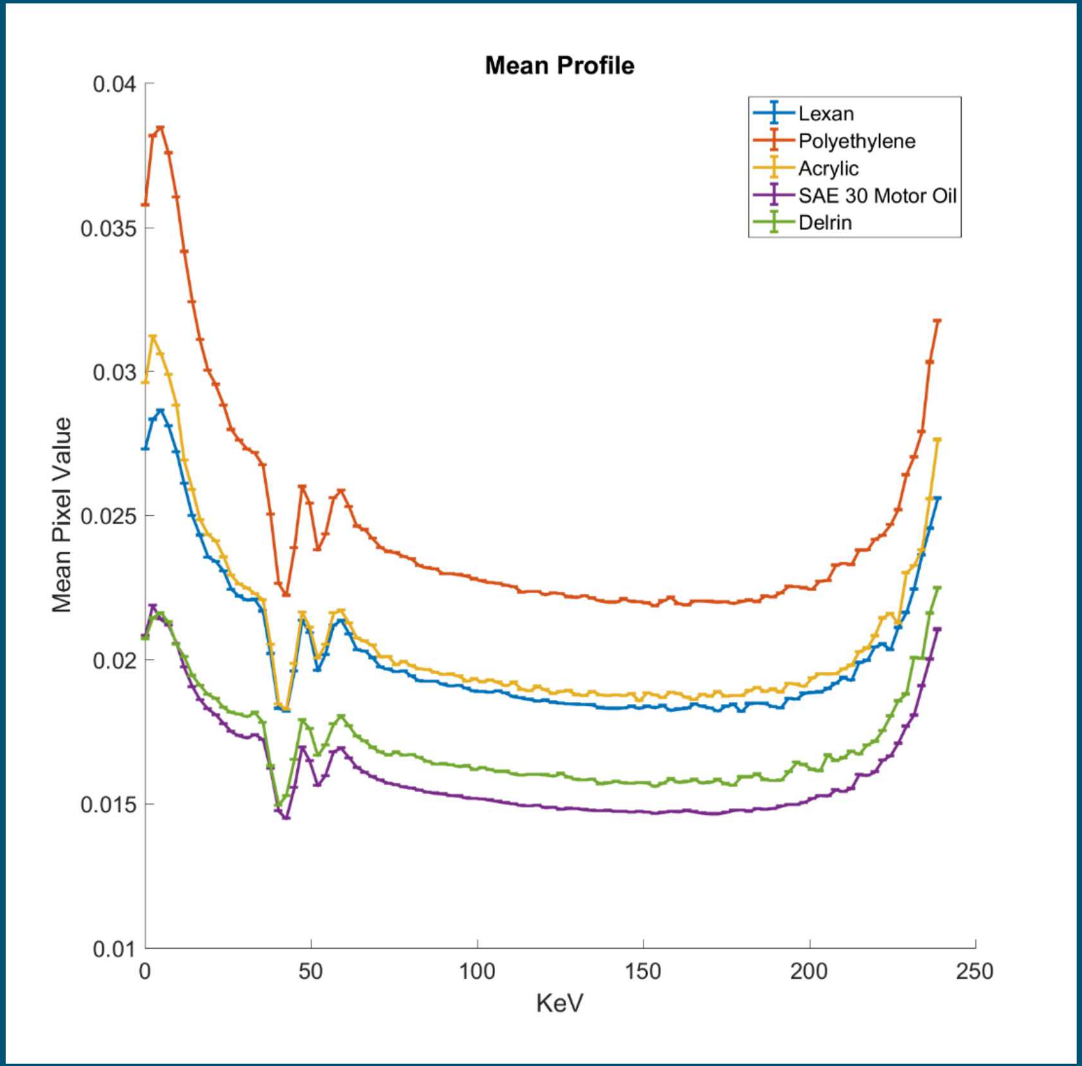
Sample Reconstruction

Metals visualized in Spectral CT



- Within the samples
  - Aluminum
  - Magnesium
  - Salt
- Profiles distinct shape and scale
  - Traditional CT - nearly identical value
  - Mg and NaCl overlap on low end.
- Significance
  - Objects with variety of metals within.
  - Segmentation without interface

Difficult Materials to Distinguish - Plastics



- Plastics have wide variation
  - Not typically addressed in medicine
  - Difficult to distinguish with CT.
  - “inorganic material”
- The profiles imply the ability to distinguish
  - Very close profiles
  - Possible fidelity limitation of the system?

The Information Buried in Spectral Data

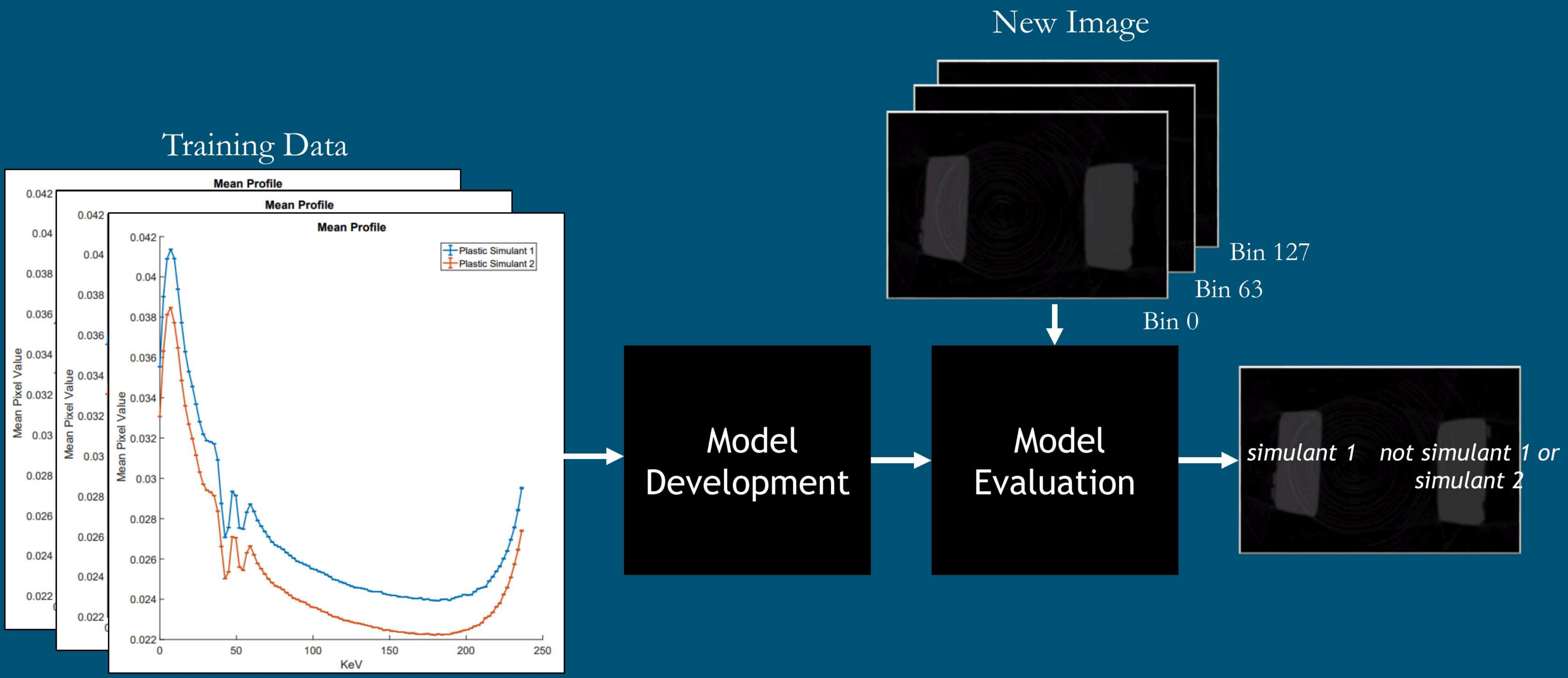
In all our samples:
- Waveforms appear distinct, bur reconstructing attenuation profile directly is difficult
- Information in the waveform correlates to attenuation

Machine Learning-based Approaches
- Go beyond visual discrimination
- Ability to detect subtle correlations
- Detect/Classify material

Supervised and Unsupervised Exploration
- Support Vector Machines with Kernelization (SVM-K)
- Multilayer Perceptrons (MLP)
- K-Means Clustering (KM)

Machine Learning – Beyond Visual Inspection

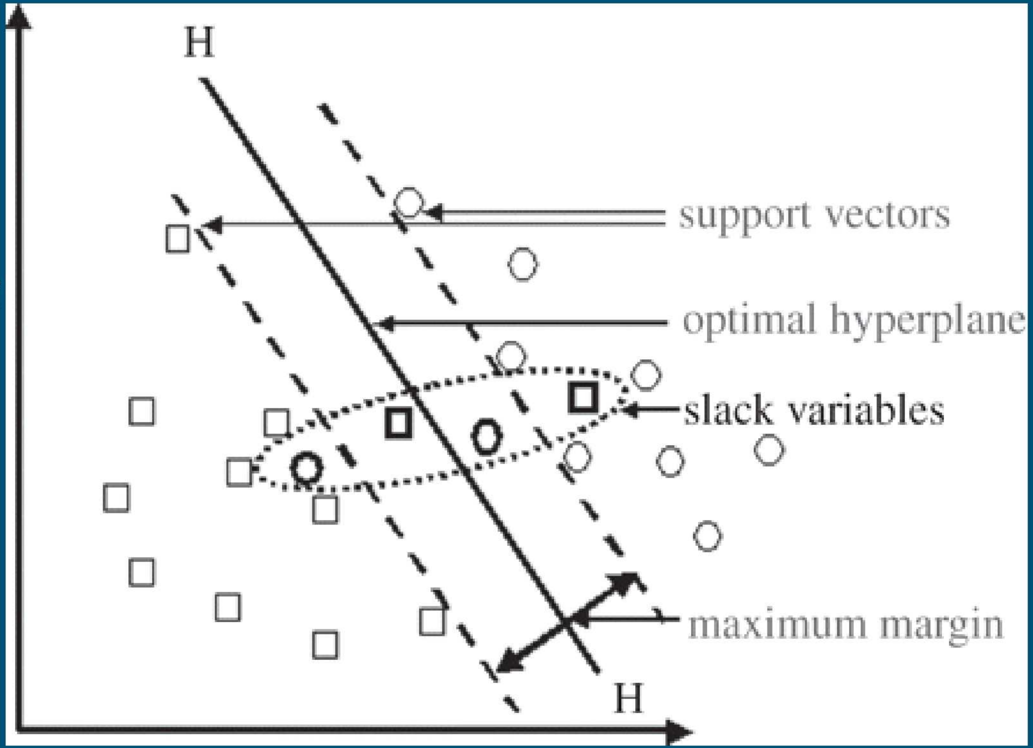
- Investigate the utility of machine learning techniques
  - Can algorithms be developed for automated material classification?
  - Which algorithms perform optimally and why?



Models & Classification Techniques – SVM-K

Kernelized soft-margin support vector machine

- Binary regression model
- Consideration of non-linearly separable data.
- **Performs well with small training sets.**
- Maps input features to higher-dimensional space prior to model calibration
- Augmented with Error Correcting Output Code.

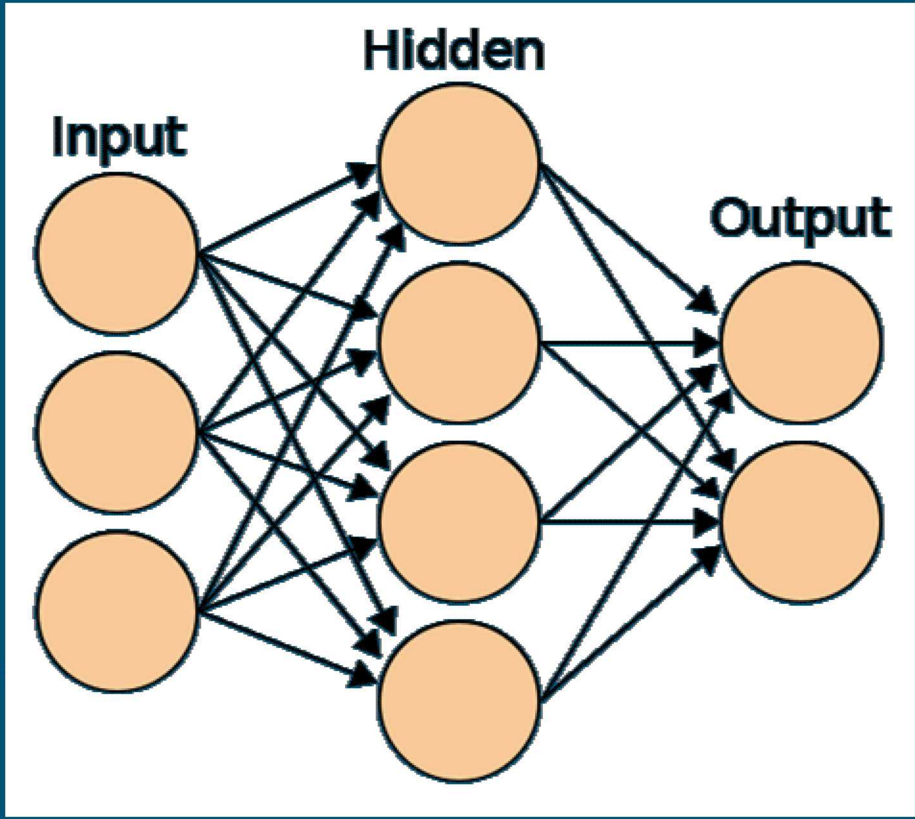


Motivation

- Attempts at simple methods were not effective.
- All x-ray imaging systems are non-linear systems.

Multilayer Perceptrons - First Steps into Deep Learning

Multilayer perceptron
- Species of Artificial Neural Network
- Single hidden layer with 300 neurons
- Model complex non-linear relationships.
- Softmax Activation Function
- Cross-Entropy Cost Function



Motivation
- Every photon counting detector is known to have complex noise.
- Simple Neural Network as precursor to more Complex Neural Networks that require larger training sets.

Classification Technique – K-Means

K-Means Clustering

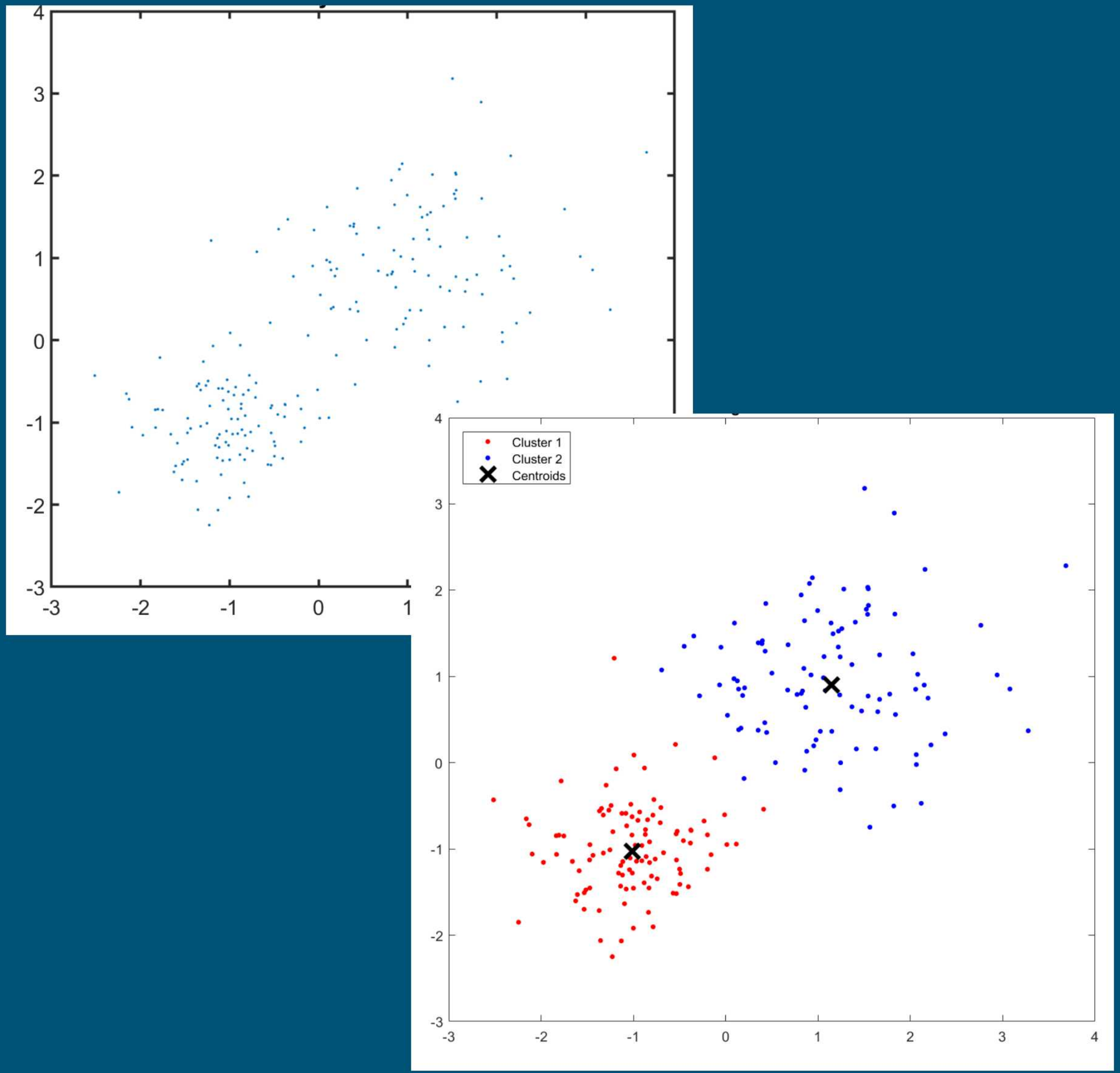
- Identify differing materials without knowledge of material signatures beforehand.
- Distinguish relative differences between waveforms

Advantages

- No prior knowledge of materials necessary
- Simple and Robust in many applications
- Low computational cost
- Simple to interpret

Disadvantages

- Number of clusters needs to be known
- Sensitive to outliers
- Linearly/Spherically separable data needed.



Implementation

Support Vector Machine with Kernalization
- TensorFlow and Matlab
- Size of training and test sets ~1700 samples per scan

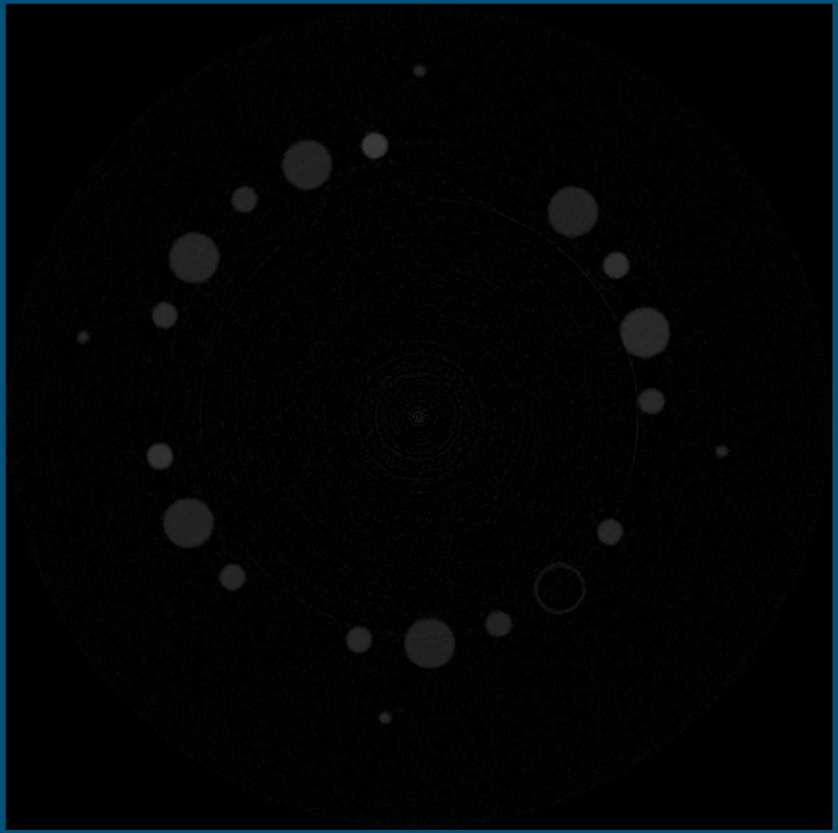
Multilayer Perceptron
- Tensorflow
- Epocs – 30
- 300 Neurons
- Single hidden layer

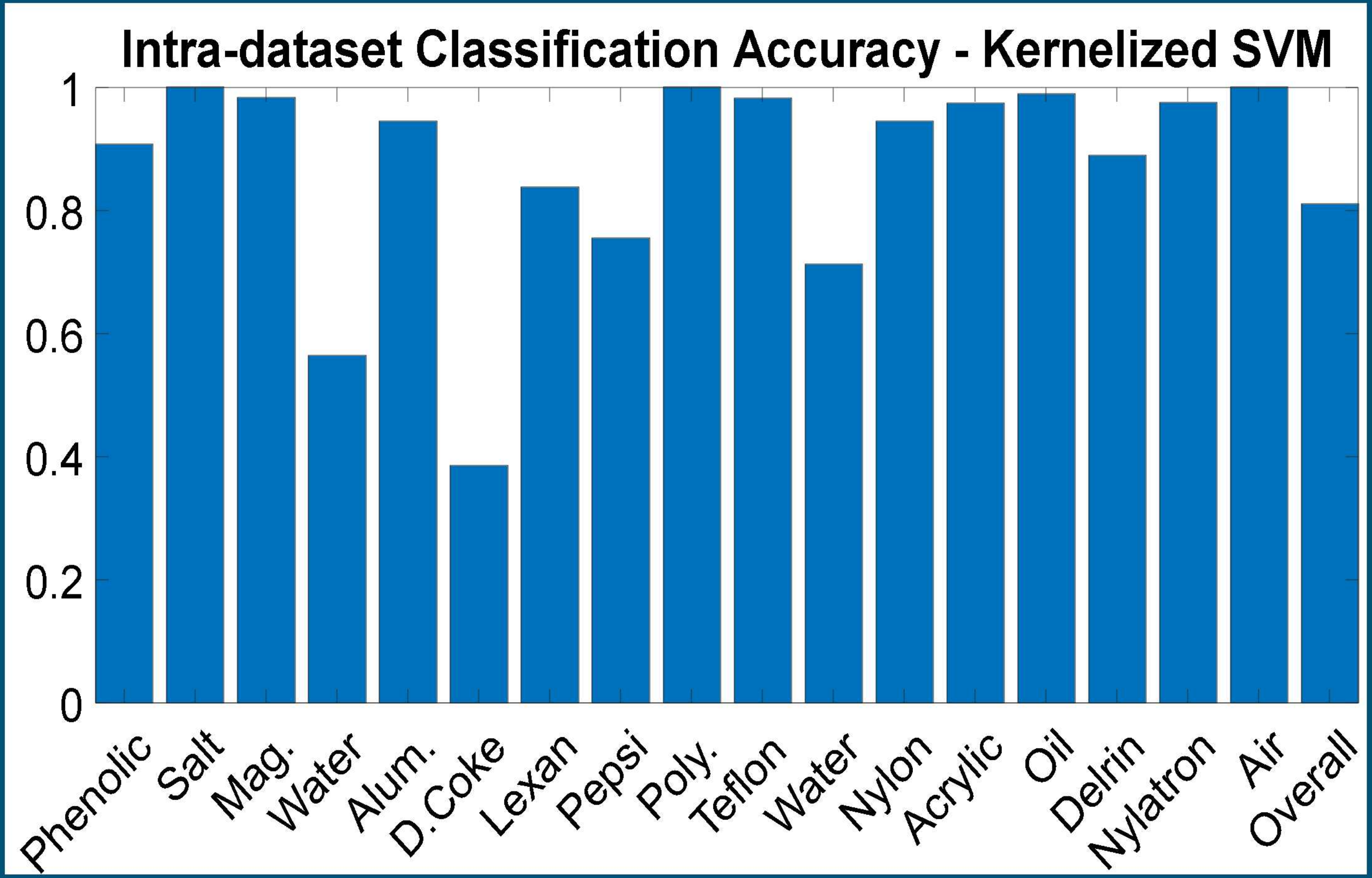
K-Means
- Matlab 2017b – Machine Learning and Statistics toolbox
- PCA to visualize number of clusters

Scans
- 250keV/0.5mA/720 projections/100 Line reads-10ms integration time
- 17 material set
- 6 ceramics

SVM-K – Excellent Intra-dataset Material Identification

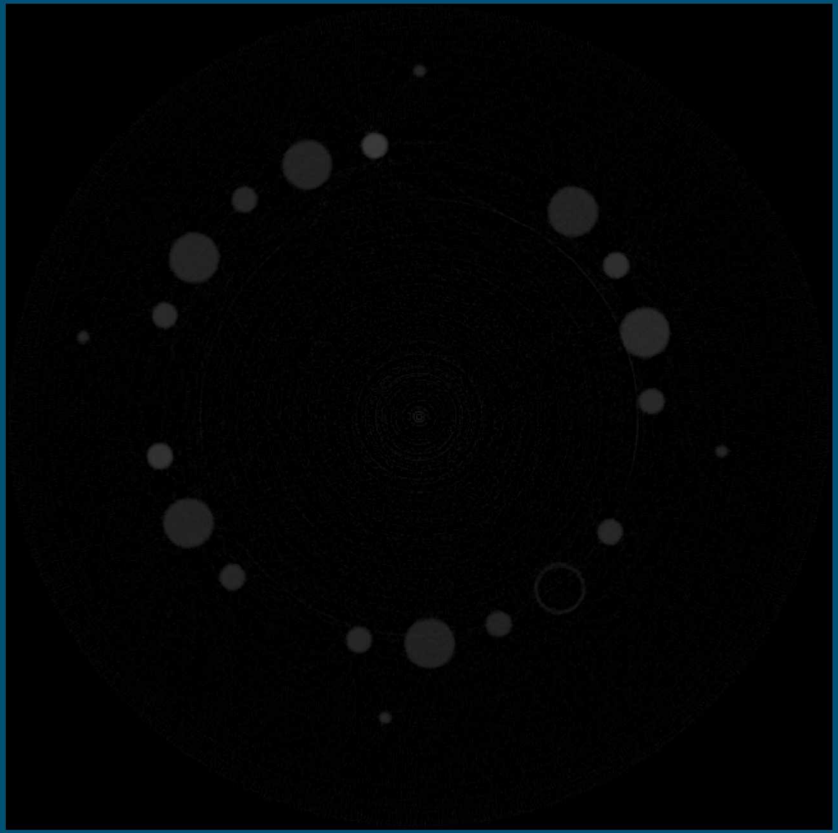
- **Overall performance of 83 percent!**
- Misdetections tended towards similar materials (i.e. Diet Coke identified as water)

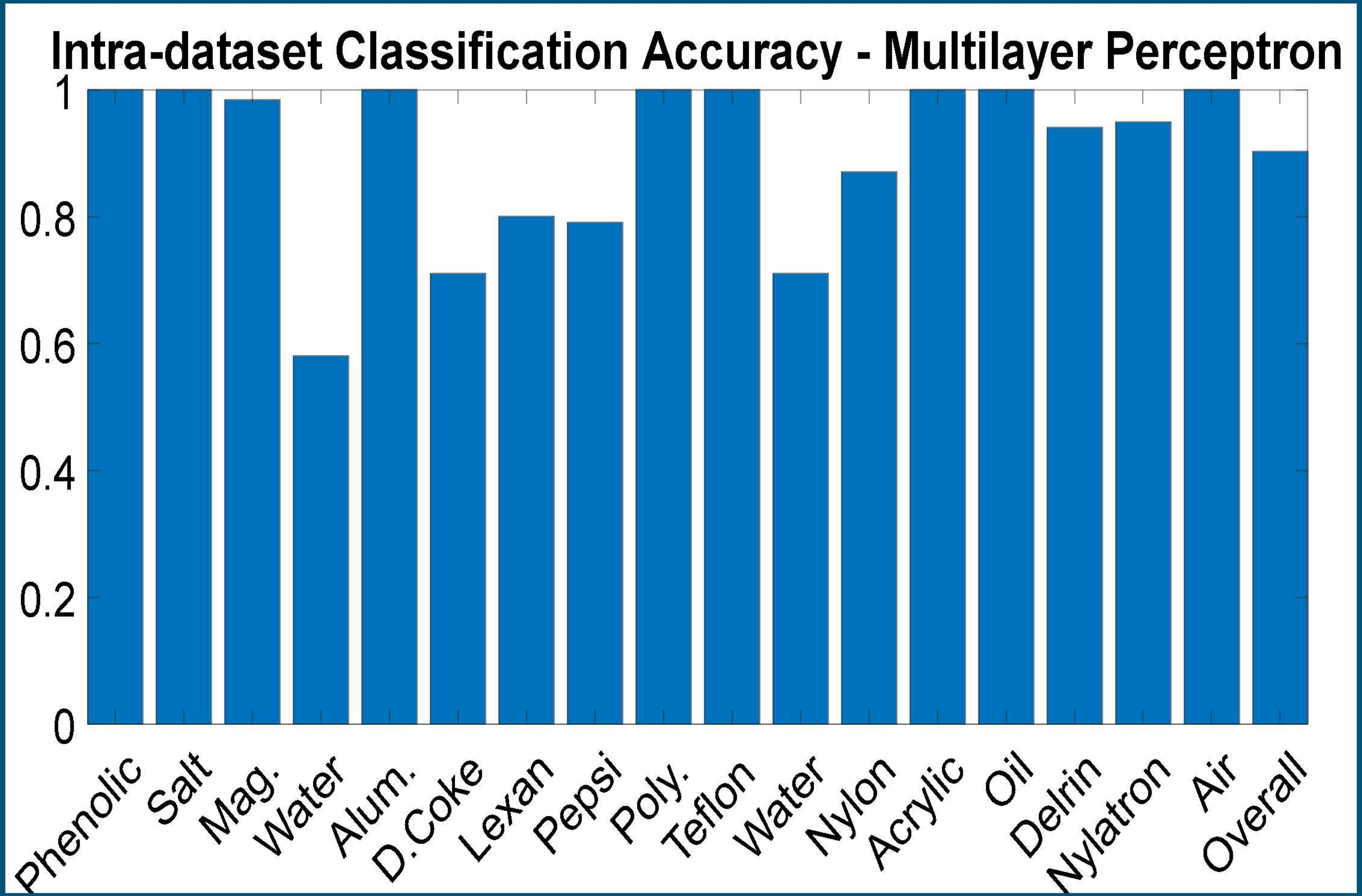




Multilayer Perceptron Improves Performance

- **Performance up to 93 percent accuracy in identifying materials.**
  - Possible over-fitting due to small training set.

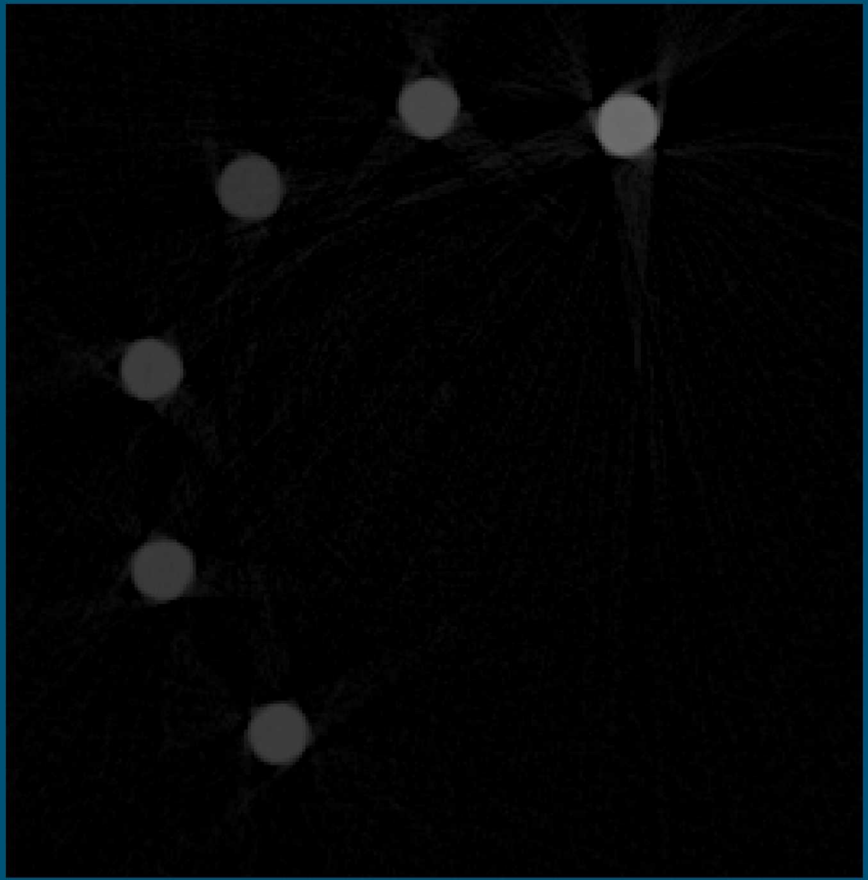


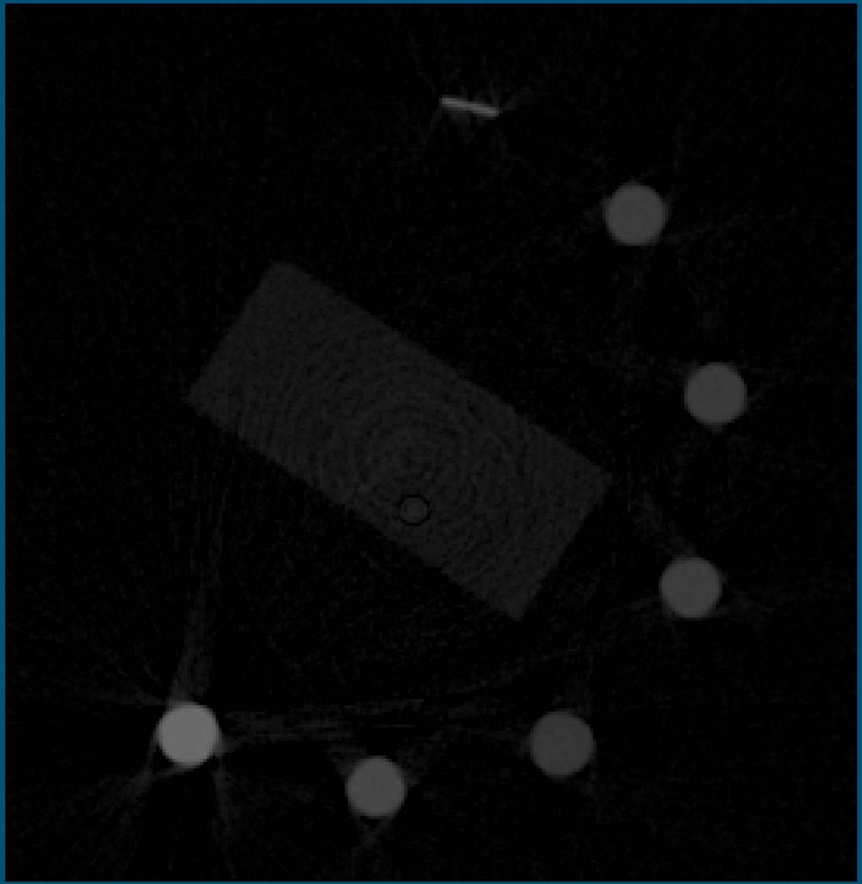


Inter-dataset Performance

- **Six ceramic variants**

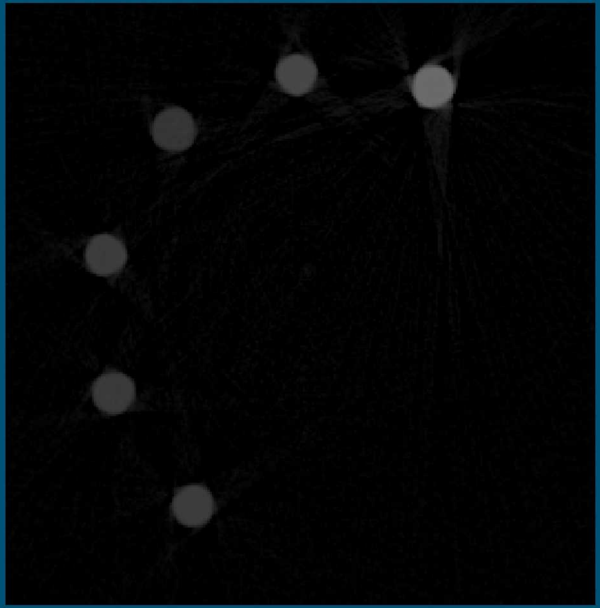
- Second scan: same ceramics but with the addition of steel penny and wood sample as interference.

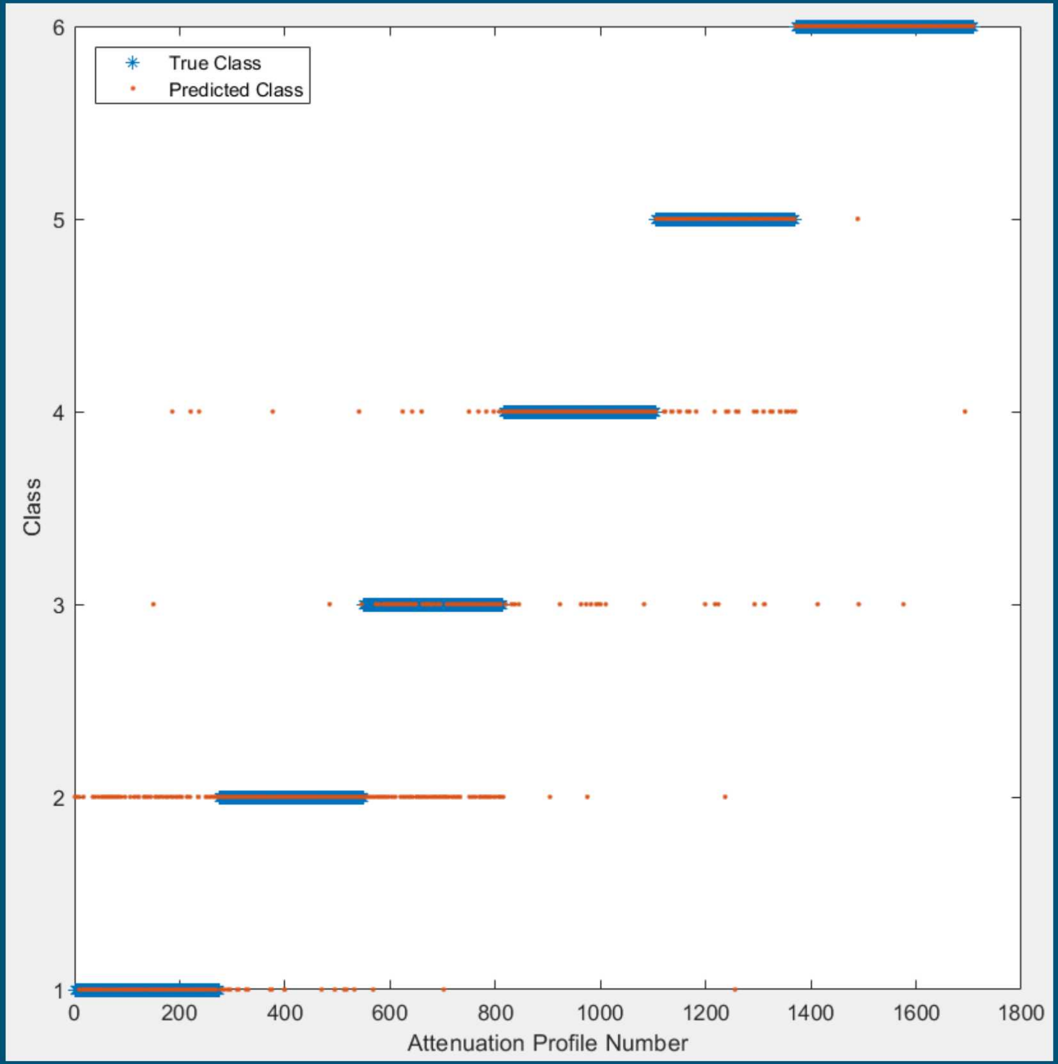


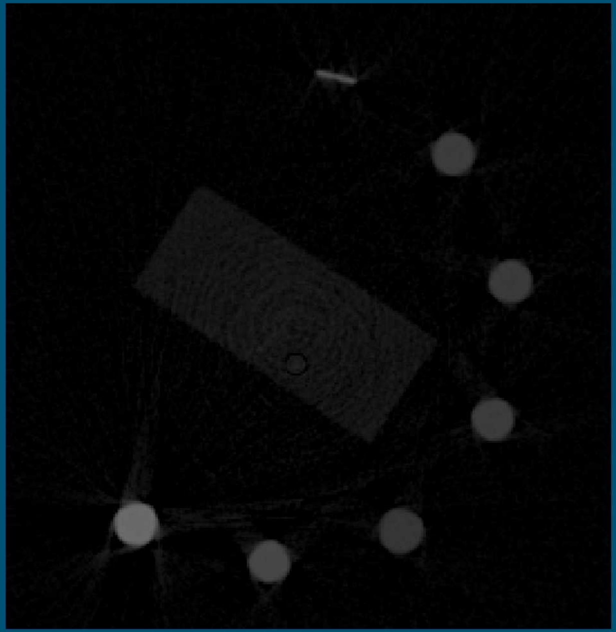


Inter-dataset Performance – SVM-K

- **Six ceramic variants**
- **Voxel by Voxel: 85%**
- **Majority voting: 100%**

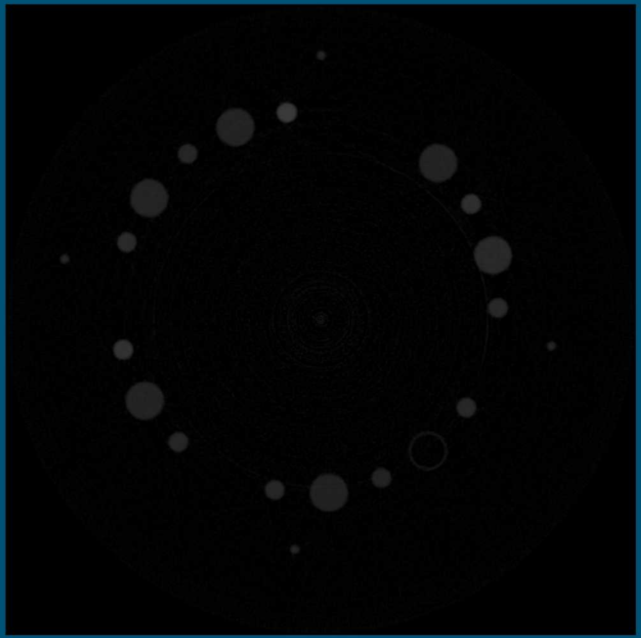


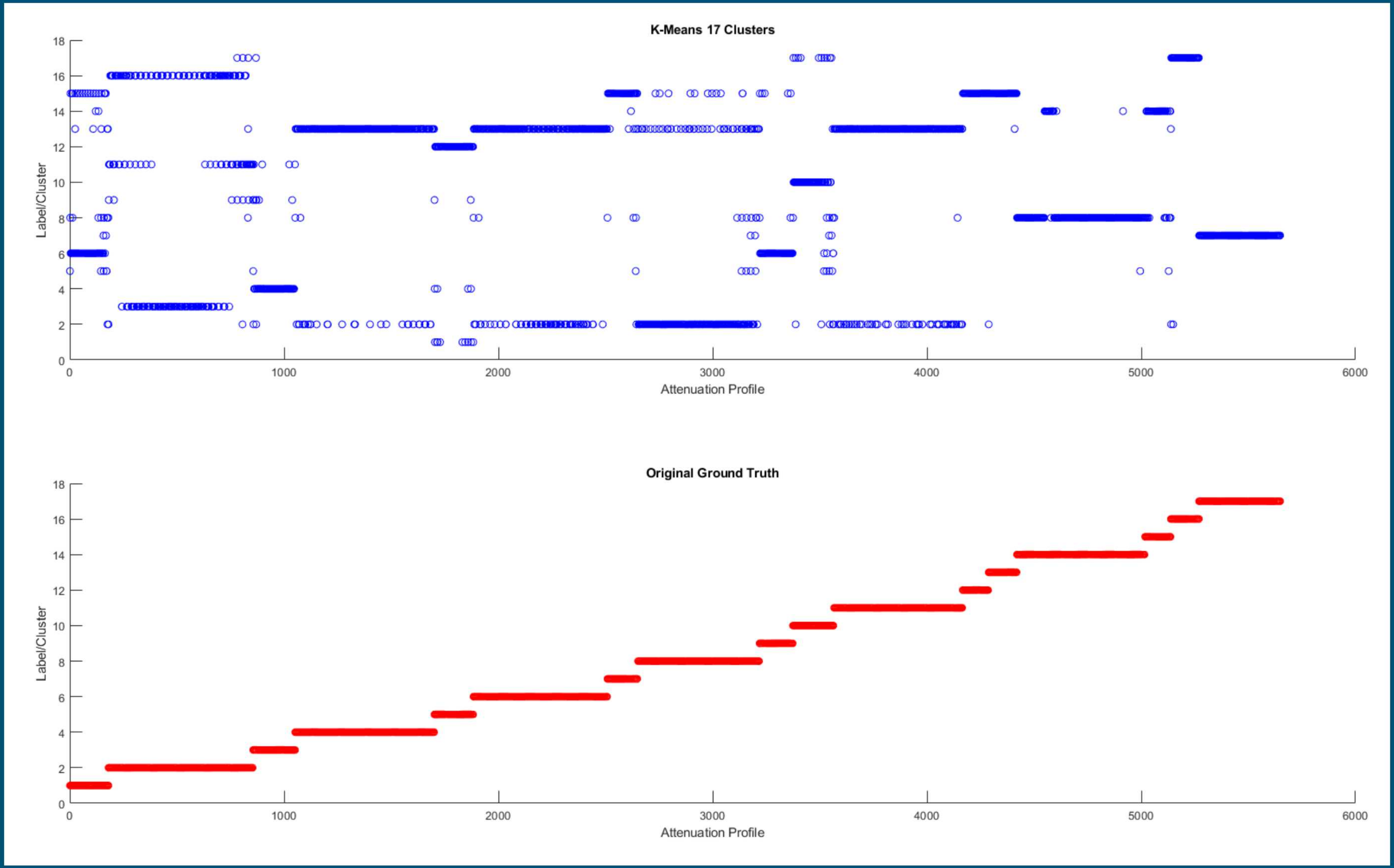




K-Means Performance – 17 Material set

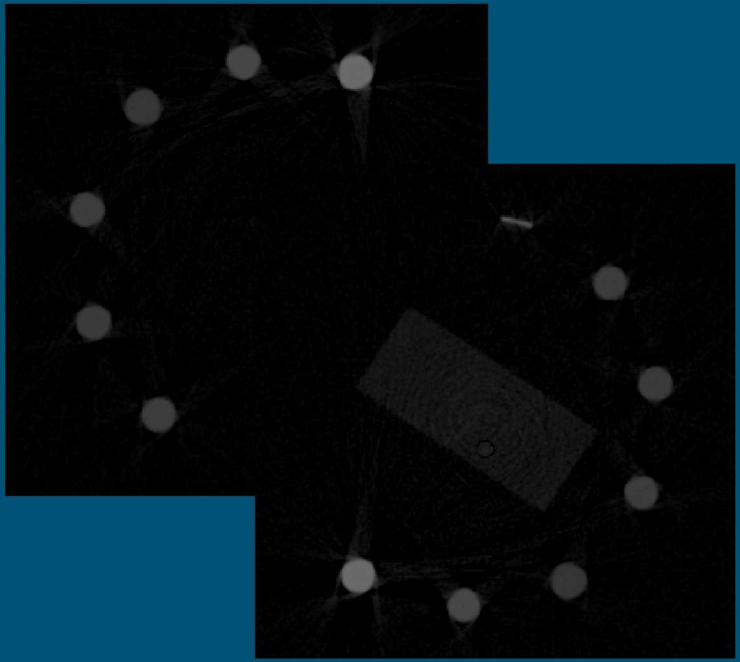
- **Number of clusters: 17**
- **Clusters well.**
- **Metrics: Euler Squared**

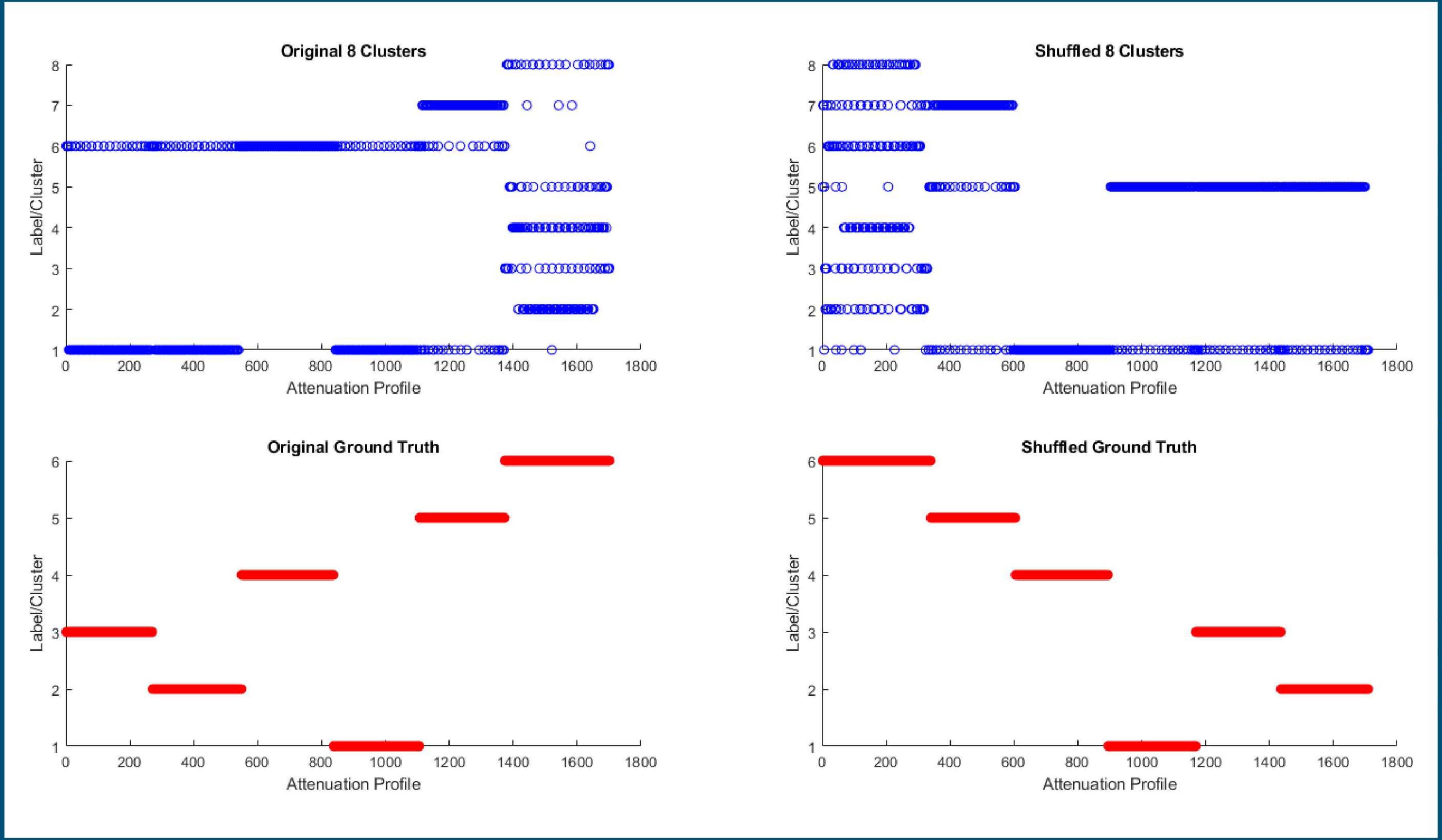




K-Means Performance – Ceramic set

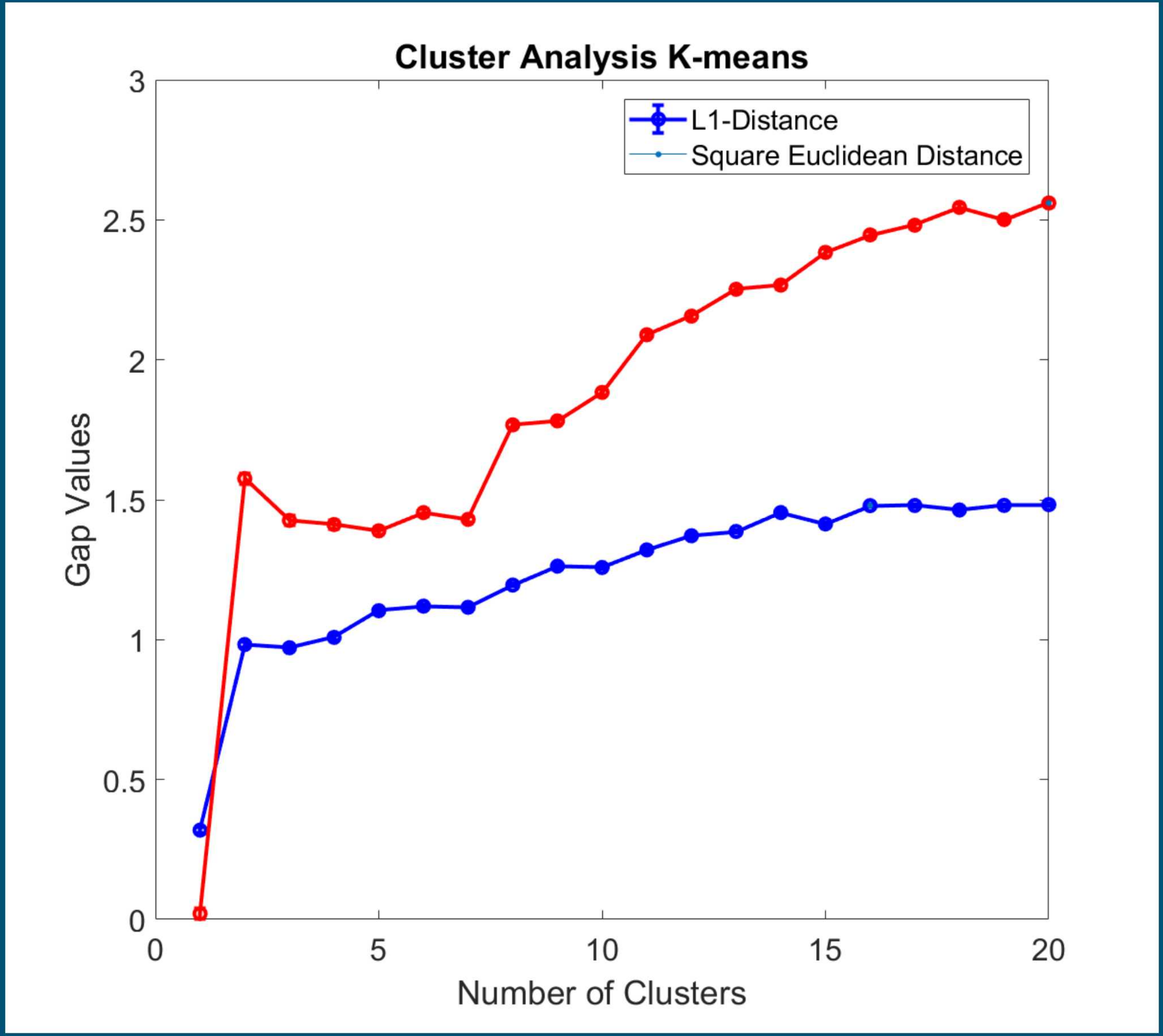
- **Number of clusters:**
- **Clusters not well defined.**
- **Metrics: Euler Squared**





K-Means Performance – L1 vs Euler Squared

- According to gap analysis:
  - Ideal 16 to 18 clusters
  - Regardless of metric
- Linearly inseparable?



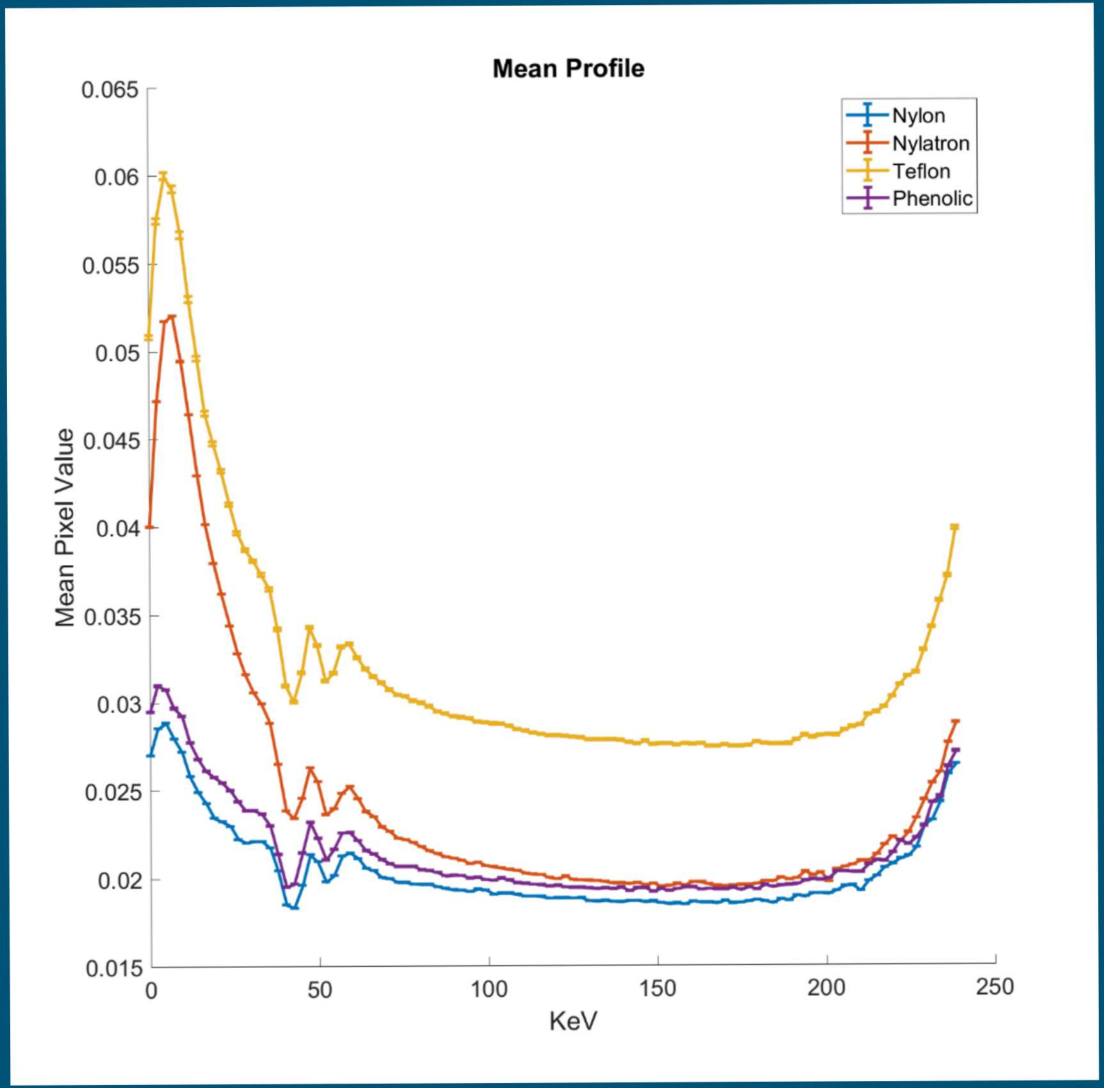
Conclusion- Looking to the future

- This work has developed a revolutionary capability in non-destructive material identification.
  - Machine learning allows for Information Extraction from Spectral Reconstructions.
  - Merging data science and engineering.

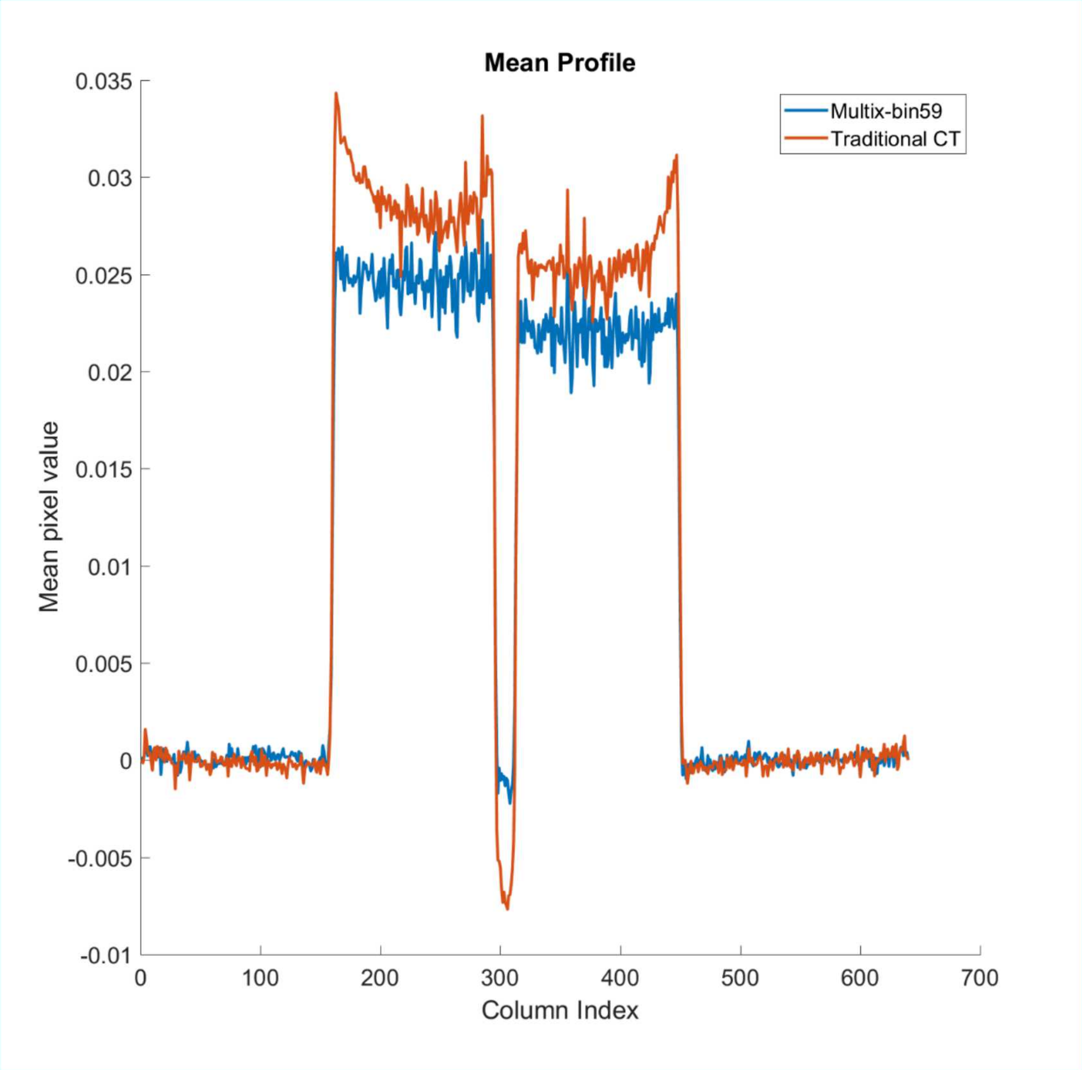
- We look to dig deeper on the potential of Machine Learning
  - Kernelized Unsupervised Learning
  - Must test capability on objects with more complex geometry
  - Larger training sets.
  - Convolutional Neural networks – Waveform Shape Detection

- Create larger training sets with particle transport code
  - PHITS
  - MCNP

Mean Profile: Nylon



Reconstruction: No Additional Cost Artifact Reduction



- Traditional CT exhibits beam hardening artifacts.
  - Unavoidable with a single set of data.
  - Further obfuscates the material identification task
- Spectral CT dramatically reduces these artifacts
  - Single Channel Reconstruction
  - Essentially bandpass filtered data.
  - No additional processing necessary!
  - Single dataset acquisition.