One Dimensional Convolutional Neural Networks for Spectral Analysis

Michael S. Primrose*, Jay Giblin, Christian Smith, Martin R. Anguita, Gabriel H. Weedon
Physical Sciences Inc. 20 New England Business Center Drive, Andover MA, 01255.

ABSTRACT

A platform for building sensor specific machine learning detection algorithms has been developed to classify spectroscopic data. The algorithms are focused on long wave infrared reflectance (LWIR) and Raman spectroscopies. The classification algorithm is based on a one dimensional (1D) convolutional neural network (CNN) architecture. Training data is generated using an appropriate signal model that is combined with sensor specific characteristics such as spectral range, spectral resolution, and noise.

Within this paper, the performance of trained CNNs for both LWIR and Raman sensor systems has been evaluated. The evaluation uses both real and synthetic data to benchmark the performance in terms of the discriminant signal. The evaluation data consists of various chemical representations and varied noise levels. The performance of the 1D CNN approach has demonstrated high classification accuracies on data with low discriminant signals. Specifically, the CNNs have demonstrated a classification accuracy >90% for infrared reflectance data down to a wavelength averaged discriminant SNR>1. For Raman systems, we have demonstrated classification accuracies >90% for data with a peak discriminant SNR of approximately 6.

Keywords: Raman spectroscopy, LWIR spectroscopy, CNNs, Deep learning, Spectral modeling, Spectral classification

1. INTRODUCTION

There is a continued need for developing automated algorithms for classification of trace and bulk chemicals with techniques such as infrared (IR) reflectance spectroscopy and Raman spectroscopy. To date, approaches for IR reflectance have included a straightforward spectral angle mapper for bulk materials, and likelihood ratio tests for trace contaminants that implement linear [1] and nonlinear [2] models of surface transflection. Machine learning and deep learning approaches for spectral classification have also been developed based on support vector machines [3], clustering [4], and neural networks [5]. Artificial neural networks (ANNs) [6] and convolutional neural networks (CNNs) [5] [7] have been used as spectral classifiers for a variety of IR and Raman sensors.

Historically, CNNs have been popular for their usage in image classification. Early and well known CNNs such as AlexNet [8] require a large amount of training data, which can make the process of acquiring training datasets a costly endeavor (e.g. the CIFAR-10 dataset has 6000 images per class). In order to reduce the volume of training data that needs to be acquired, and to help prevent overfitting, data augmentation can be applied to smaller datasets [7]. Furthermore, sensor specific models can be used to generate training data in order to support the development of a "sensor customized" CNN, highlighting the versatility of the approach to classification.

In this work, we describe the performance of a one dimensional (1D) CNN architecture that is trained using synthetic data sets. This framework is being used to train sensor customized classification algorithms while requiring minimal user inputs. Specifically, the approach only requires the spectral range, sampling, and resolution of the sensor hardware to be used. We have developed and demonstrated this approach for IR reflectance and Raman spectroscopies, helping show that this approach is a step forward towards building a platform that supports sensor customized CNNs.

*mprirmose@psicorp.com; phone 1 978-738-8294; psicorp.com

2. CNN ARCHITECTURE AND TRAINING DATA

2.1 CNN Architecture

One of the well-known uses for CNNs has been object classification in two dimensional (2D) imagery. Krizhevsky's work on AlexNet demonstrated how CNNs act as a trainable feature extractor, where the CNN "learns" the important features required to classify objects [8]. Pre-trained 2D CNNs are widely available and can serve as a basis for training a spectral classifier; however, this requires turning the inherent 1D spectral data product into a format matching the input data format. Further, using the pre-trained 2D CNNs still requires training to tune the network to learn the spectral representations [9]. Using a 1D CNN offers two salient advantages:

- The 1D CNN utilizes the inherent 1D data product without needing to process the data into a 2D format
- The 1D CNNs have fewer trainable parameters than the 2D counterparts which can help prevent overfitting, decrease the training time, and increase the inference time.

The CNNs we have developed, and presented within this paper, are a modified 1D analog of VGGNet [10]. The pooling layers of the original VGGNet were removed in favor of a strided convolution, and only a single fully-connected layer was used at the end of the network. Lastly, we started our convolutional layers with 128 feature maps, and progressed through the network following the VGG13 architecture. Figure 1 shows a schematic diagram of a 1D CNN with focus on the first set of convolutional kernels for the first hidden layer. The input spectra is convolved with a set of convolutional kernels (i.e. learnable filters) where each kernel produces a feature map. The convolution of the spectra and learnable filters is passed through an activation function to produce the first hidden layer. Subsequent hidden layers are produced using feature maps and new convolutional kernels. These subsequent hidden layers encode more complex spectral correlations. The dense layer maps the final set of feature maps to the output layer. In the output layer, the softmax classifier is used to return a set of scores for each class that is used to make a target prediction.

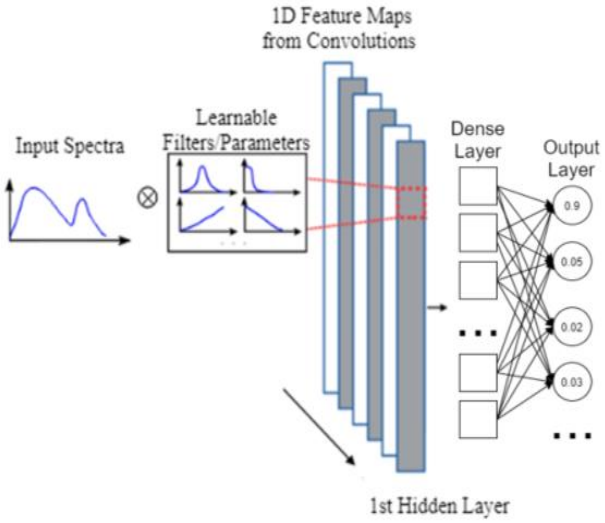

Figure 1: Schematic diagram of a 1D CNN architecture including hidden layers and a dense layer at the end. The architecture used in this paper has eight hidden layers before the dense layer

The training data was produced using mathematical and physics based models including sensor characteristics such as spectral range and resolution. For IR reflectance, the data is generated using the reflection, absorption, and particle scattering parameters derived from measured quantities (n and k scattering parameters). The Raman model requires only a single high fidelity measured reference spectrum for each target or background; however, additional spectra could be incorporated. The spectra are convolved with a Gaussian instrument line shape consistent with the sensor's spectral resolution. Finally, the spectra are interpolated to the spectral sampling of the senor.

The CNNs are trained by minimizing a categorical cross-entropy loss function with the Adam optimization algorithm. All of the CNNs discussed below were trained using the same set of hyperparameters and number of samples seen during the

training. In order to work towards the goal of having a generalized platform for training sensor specific CNNs, hyperparameter optimization has not been studied.

2.2 IR Reflectance Model

The IR reflectance model uses scattering from chemical films, particulates, and a surface material given by equations 1, 2, and 3 respectively:

$$R_{film} = \phi_t R_c + R_o(1 - R_C)^2 Exp[-2\alpha z] \quad (1)$$

$$R_{part} = \sum_i \left\{ \phi_i Q_{b,i} + \phi_i R_o \left(1 - Q_{ext,i} + Q_{f,i}^2\right)^2 \right\} \quad (2)$$

$$R_{Surface} = \left(1 - (\phi_t + \sum_i \phi_i)\right) R_o \quad (3)$$

In this model, the film fill fraction $\boldsymbol{\phi_t}$ and particle fill fractions $\boldsymbol{\phi_i}$ are based on the fraction of a target material distributed on the surface with a reflectance R_o. The film representation models the target reflectance R_c and the reflectance from the surface material attenuated by the film with a thickness z and absorption coefficient α. The reflectance of the i^{th} particle is the result of backscattering $Q_{b,i}$, forward scattering $Q_{f,i}$ and extinction $Q_{ext,i}$ calculated using Mie Theory [10]. A more detailed description and experimental validation of this model is presented in [2].

The training data for the IR CNN are generated from Equation 4:

$$R_{train} = R_{film} + R_{part} + R_{Surface} + \mathcal{N} \quad (4)$$

where $\mathcal{N}$ is a Gaussian noise term. The training data was produced by sampling through a range of target fill fractions, film thicknesses, and various numbers of particles. Training spectra were filtered by noise levels to ensure that at least one point in each spectrum had a discriminant SNR >1. The discriminant SNR is given by Equation 5:

$$Disc.SNR = \frac{|R_{train} - R_o|}{\mathcal{N}} \quad (5)$$

2.3 Raman Linear Mixing Model

The training data was generated using a linear mixing model given by Equation 6:

$$R_{train} = \phi R_{target} + (1 - \phi) R_{bg} + \mathcal{N} \quad (6)$$

where ϕ is the fraction of the target spectrum R_{target}, and R_{bg} is the background (or container) spectrum. The generated spectra includes a Gaussian noise term $\mathcal{N}$ that is varied for every spectra in order to produce a wide variety of noise signatures. The peak discriminant SNR given by Equation 7 was used to filter out any training if the peak discriminant SNR< 1.

$$Disc.SNR = \frac{|R_{train} - R_o|}{\mathcal{N}} \quad (7)$$

3. EVALUATION OF CNN PERFORMANCE AGAINST IR REFLECTANCE DATA

The IR CNNs were typically trained using the target list in Table 1 except one CNN which only used targets available in the validation data. The CNNs were also trained with a background class corresponding to a single type of background substrate. The multiple CNNs were trained for roughened aluminum, high density polyethylene, and cardboard substrates. The CNNs were trained using a set of default training parameters. The CNNs were trained using two particle size distributions. The smaller particle size distribution with a mean particle size of 13 microns and a standard deviation of 3 microns. The larger particle size distribution was based on [2]. These CNNs focused on the long wave IR region specifically from 830-1430 cm^{-1} with a spectral resolution of 8 cm^{-1} and sampled every 4 cm^{-1}. The CNNs trained with aluminum and high density polyethylene (HDPE) substrates were evaluated using synthetic data provided by the Naval Research Labs (NRL).

Table 1. Target list for IR CNN

Chemicals		
2-Chloroethyl Ethyl Sulfide	DMMP	Potassium Chlorate
Acetaminophen	DNT	RDX
Acetonitrile	HMX	Sucrose
Ammonium Chloride	Hydroxypropyl Cellulose	TBP
Caffeine	Isosorbide Dinitrate	TEP
Calcium Stearate	Naproxen Sodium	TMP
Cyclohexanol	Nicotine	TNT
DEMP	n-Octane	Toluene
Dephenylamine	PETN	Triethyl Phosphite
D-Mannitol	Poly(vinylidene fluoride)	

3.1 Synthetic Data Evaluation

The CNNs trained with roughened aluminum and HDPE substrates were evaluated using a set of synthetic data provided by NRL. The synthetic data consisted of target signatures corresponding to surface loadings up to 100 μg/cm^2 with different levels of simulated sensor noise. The only preprocessing done on these data was to convolve and resample the data to match the spectral range and resolution that was used for the CNN. The target signatures were purely particulate with a mean particle size of 13 microns and a standard deviation of 3 microns.

The evaluation data was filtered to only include spectra that had a wavelength averaged discriminant SNR >1. Table 2 has the breakdown of the percentage of correctly classified spectra for the 13 targets in the validation data that were also in the training data. The percentage of correctly classified spectra is further broken down into the substrate type and the particle size distribution used for the training data.

Table 2. Evaluation of CNNs trained for IR spectra. Models were trained with two different particle size distributions. Entries show percentage of correctly identified spectra with an average discriminant SNR>1.

	Aluminum Substrate		**HDPE Substrate**		
	Small Particles	**Large Particles**	**Small Particles**	**Large Particles**	**Small Particles (13 targets)**
Acetaminophen	100%	100%	100%	100%	100%
TNT	100%	96%	100%	100%	100%
DNT	100%	100%	100%	82%	100%
RDX	100%	100%	100%	100%	100%
Naproxen	100%	100%	62%	58%	71%
HMX	100%	100%	100%	100%	100%
Sucrose	97%	14%	69%	23%	77%
Caffeine	97%	97%	88%	24%	88%
DEMP	100%	99%	100%	91%	100%
DMMP	100%	48%	0%	0%	0%
PETN	100%	100%	95%	90%	90%
TEP	96%	95%	12%	0%	25%
Warfarin	100%	100%	100%	10%	95%

3.2 Cardboard Sieved Material Performance

A small set of target material (acetaminophen, caffeine and potassium chlorate) was sieved onto cardboard substrates using different sieve sizes and surface loadings. Spectra were collected using a Nicolet 4700 FTIR with an integrating sphere. The range of surface loadings and sieve sizes can be seen in Figure 2**.** Figure 2 shows the distribution of correctly/incorrectly classified spectra at the various surface loading/sieved particle sizes using color coded points. Further, Figure 2 shows all measured spectra. The average discriminant SNR (Equation 5) was estimated from the measured spectrum using the measured average cardboard background. The measured target spectra were filtered to only include spectra with an average discriminant SNR>1. The CNN correctly classified 60, 80 and 96% of the remaining potassium chlorate, acetaminophen and caffeine spectra. The CNN correctly classified 80.6% of all spectra with an average discriminant SNR >1.

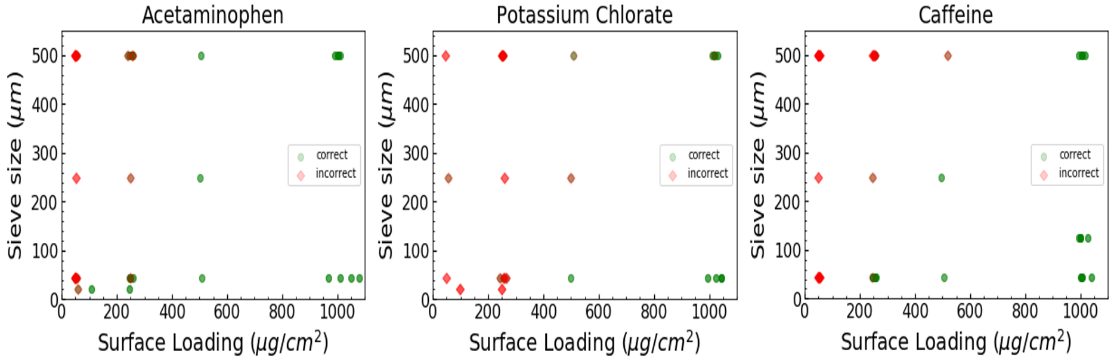

Figure 2. Surface loadings and sieve size used to prepare samples measured on the Nicolet 4700 FTIR. Green and red points correspond to correctly/incorrectly classified spectra respectively. All collected spectra are shown.

4. CNN PERFORMANCE AGAINST RAMAN DATA

The Raman CNN was trained with the targets in Table 3. Background spectra for glass, HDPE, cardboard, and leather were measured using a B&W Tek iRaman system. Spectra for explosive targets (a list of the specifically targets can be seen in Table 3) were obtained from the University of Rhode Island (URI) explosives database. Spectra for sucrose, caffeine, acetaminophen, guanidine hydrochloride and hydroxypropyl cellulose were measured with the iRaman system and added to the training library. Two CNNs were trained based on different sensor characteristics. The first CNN discussed was trained based on the First Defender characteristics, and the second CNN was trained based on data from a Raman microscope used at NRL.

Table 3. Raman target list.

Chemicals

1,3 Dinitrobenzene	Erythritol Tetranitrate (ETN)
1,3,3-Trinitroazetidine (TNAZ)	Ethylene Glycol Dinitrate (EGDN)
1,3,5-Trinitrobenzene	Guanadine Hydrochloride
1,3,5,7-Tetranitro-1,3,5,7-Tetrazocine (HMX)	Hexamethylenetriperoxidediamine (HMTD)
1,4 Dinitrobenzene	Hexanitrohexaazaisowurtzitane (CL-20)
2,3-Dimethyl-2,3-Dinitrobutane (DMNB)	Hexhydro-1,3,5-Trinitro-1,3,5-Triazine (RDX)
2,4,6-Trinitrophenol (Picric Acid)	Hydroxypropyl Cellulose
2,4,6-Trinitrotoluene (TNT)	Lead Azide
2,4-Dinitrotoluene	Mercury 5-Nitrotetrazole (DXN-1)
2,6-Bis(picrylamino)-3,5 Dinitropyridine (PYX)	Meso-Erythritol
2,6-Dinitrotoluene	Methyl-2,4,6-Trinitrophenylnitramine (Tetryl)
2-Amino-4,6-Dinitrotoluene	Methylnitroguanidine
2-Nitrotoluene	Nitrourea
3-Nitro-1,2,4-Triazol-5-One (NTO)	Pentaerythritol Tetranitrate (PETN)
3-Nitrotoluene	Potassium Chlorate
4-Amino-2,6-Dinitrotoluene	Potassium Nitrate
Acetaminophen	Potassium Perchlorate
Ammonium Dinitramide	Sodium Chlorate
Ammonium Nitrate	Sodium Nitrate
Ammonium Perchlorate	Sucrose
Caffeine	Tetraamine-Cis-Bis(5-Nitro-2H-Tetrazolato-N2)Cobalt(III) (BNCP)
Dextrinated Lead Azide	Triacetonetriperoxide (TATP)
Diacetonediperoxide (DADP)	Urea Nitrate
Dicyandiamide	

4.1 CNN for First Defender Data

A CNN was trained with the spectral sampling of the first defender and a spectral range of 450-2630 cm^{-1}. We evaluated the performance of the CNN on sets of data collected using the First Defender as summarized in Table 4. The collected data was taken by measuring bulk samples through various containers including glass, polypropylene (PP), polyethylene sandwich bags (PE) and polyethylene terephthalate (PET) bottles. Data was collected while varying both the integration times and laser power rather than relying on the automatic collection mode.

The collected data was classified by the CNN. Table 4 reports the CNN's top 1 and 3 classification accuracy as well as the accuracy of the native algorithm. A spectrum is considered correctly identified by the native algorithm if the correct target appears anywhere within the potential list of targets. The CNN had a typically higher top 1 classification accuracy than the native algorithm. The CNN always had a top 3 classification accuracy higher than the native algorithm. The CNN demonstrated a classification accuracy of >90% for all of the targets in glass containers except erythritol. Additionally, the CNN had a >90% classification accuracy for TNT through polypropylene and HMX through polyethylene.

Table 4. Evaluation of the CNN trained with the First Defender characteristics. Top 1 and Top 3 classification accuracy were found for the eight measured target compounds. Results were compared to the native algorithm.

	Container	Number of Measured Spectra	Top 1 Classification Accuracy	Top 3 Classification Accuracy	Native Algorithm Accuracy (any)
Potassium Chlorate	Glass	93	93.50%	100%	94.60%
	PP	93	80.60%	97.80%	84.90%
	PET	93	86.00%	98.90%	82.80%
Erythritol	Glass	93	86.00%	89.20%	39.80%
	PP	93	41.90%	51.60%	12.90%
	PET	93	29.00%	40.90%	7.50%
1,3 DNB	Glass	93	100%	100%	77.40%
	PP	93	82.80%	93.50%	38.70%
2,6 DNT	Glass	92	95.70%	100%	87.00%
	PP	93	88.20%	92.50%	65.60%
HMX	Glass	16	100%	100%	81.30%
	PP	38	92.10%	94.70%	71.10%
	PE	38	89.50%	92.10%	71.10%
	Glass and Paper	32	18.80%	15.60%	28.10%
RDX	Glass	22	91.00%	95.50%	81.80%
	PP	44	59.00%	72.20%	81.80%
	PE	42	59.50%	88.10%	73.80%
	Glass and Paper	40	15.00%	60.00%	30.00%
TATP	Glass	22	91.00%	100%	77.30%
	PP	44	86.40%	90.10%	72.70%
	PE	44	77.30%	77.30%	63.60%
	Glass and Paper	40	35.00%	10.00%	45.00%
TNT	Glass	27	96.30%	100%	70.40%
	PP	45	88.90%	93.30%	75.60%
	PE	34	91.20%	100%	67.60%
	Glass and Paper	24	54.20%	62.50%	29.20%

The incident energy on the containers was determined by measuring the laser power at each power setting and using the known integration time. The data in Table 4 was broken down by the known incident energies to produce the laser incident energy vs classification curves in Figure 3. The reported error bars are the standard deviations that were returned from a bootstrapping analysis [11] that was done by a repeated random sampling with replacement of the measured spectra. The CNN outperforms the native algorithm by demonstrating higher classification accuracies at lower incident energies. At higher incident energies, the performance of the CNN and the native algorithm converge. The overall classification accuracy shown in Figure 4 further demonstrates these trends.

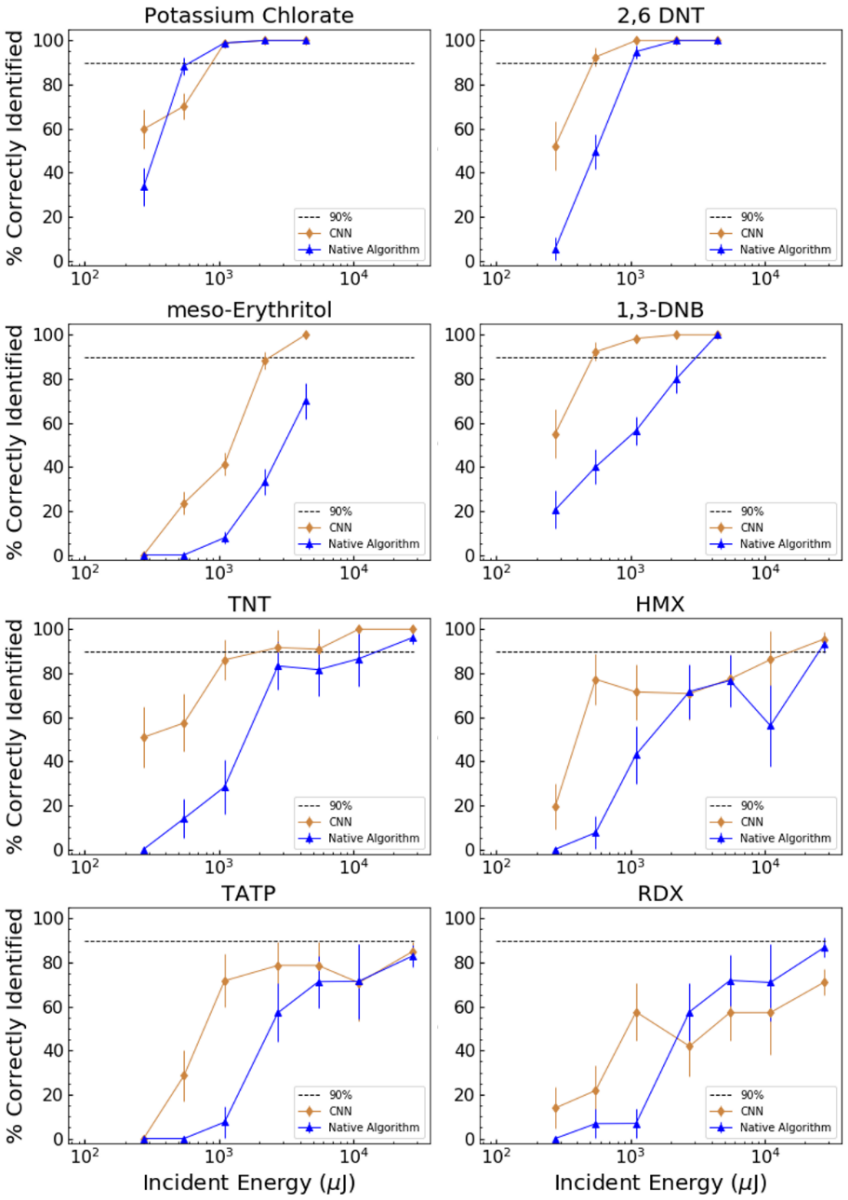

Figure 3. Percentage of correctly classified spectra at each energy level. Incident energy levels were fixed by the First Defender settings for laser power and integration times. Orange curves show the classifications made by the CNN. Blue curves show classifications made by the native algorithm.

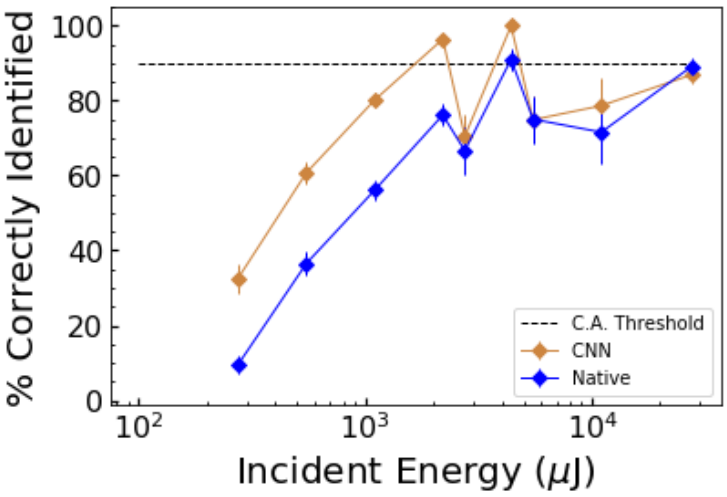

Figure 4. Overall classification accuracy of spectra taken at different incident energy levels.

4.2 CNN for Raman Microscope Data

The second CNN was trained to evaluate Raman microscope data that NRL provided. The CNN was trained using the spectral sampling of the Raman microscope (150-2368 cm^{-1}). The provided data consisted of PETN and RDX sieved onto the surface of a substrate. Substrates of aluminum, nylon, cardboard and glass were prepared for each target. The samples were prepared by passing the target material through a 20 micron sieve. An image of PETN on aluminum is shown with a sample grid in Figure 5. For each of the target/substrate combinations 300 spectra were taken.

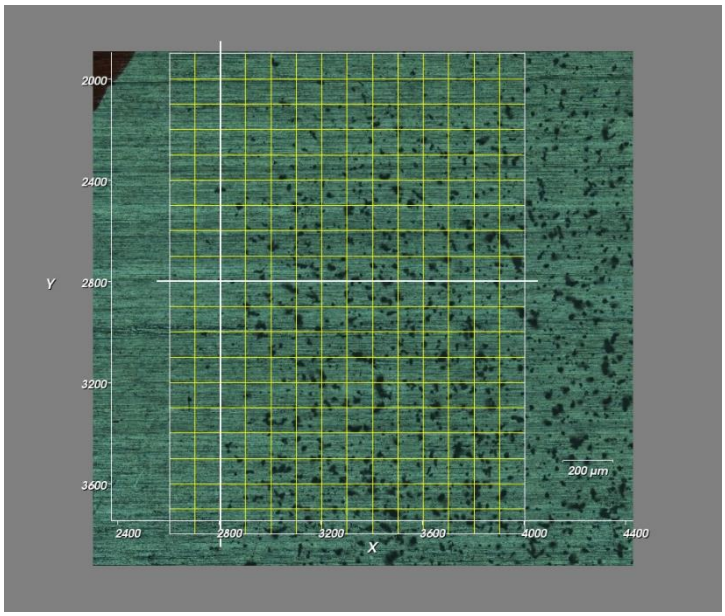

Figure 5. Reference image of ETN sieved onto an aluminum substrate. Points were sampled at the intersection of the gridlines shown.

Spectra for the different substrates was added to the training library by selecting a representative background spectrum from the measured points. . The CNN was trained using the target chemicals in Table 3, and the same hyperparameters used for all of the described CNNs. A classification maps for PETN on aluminum is shown in Figure 6.

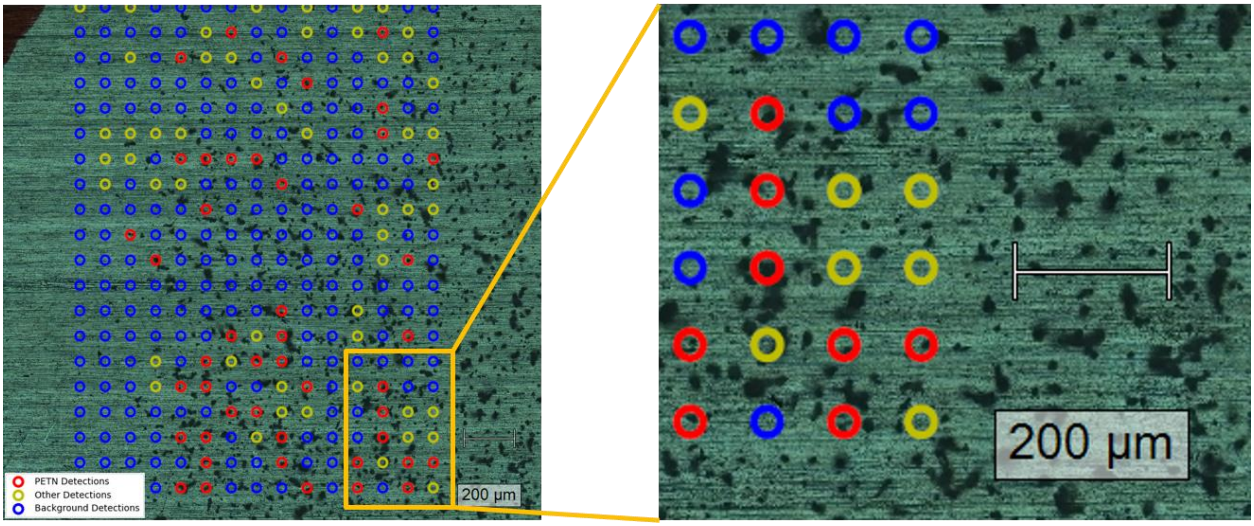

Figure 6. Classification map of sampled points of PETN on aluminum. Red indicated a classification of PETN, yellow indicates a detection of some other compound, while blue indicated a background detection of aluminum.

The points identified as the correct target clearly show particles (dark spots) inside of the circled area where the data was measured. Table 5 reports the raw number of spectra that were classified as the target. No ground truth was provided in order determine a percentage of correctly classified spectra. Still, this qualitative analysis demonstrates the capability of using a 1D CNN to produce a sensor customized classification algorithm that can incorporate user-supplied data.

Table 5. Number of spectra that were classified as "target" by the CNN

	Aluminum	Glass	Nylon	Cardboard
PETN	46	22	41	0
RDX	28	9	44	1

5. CONCLUSION

A general CNN architecture and set of training parameters were established for training a sensor customized classification algorithm. This approach has been demonstrated for both IR and Raman systems through the evaluation of real and synthetic data that was not part of the training data. The CNN demonstrates classification accuracies of >90% for a range of different target representations and collection scenarios.

6. ACKNOWLEDGEMENTS

This material is based upon work supported by U.S. Department of Homeland Security Science and Technology Directorate under Contract Number 70RSAT21C00000015. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of U.S. Department of Homeland Security Science and Technology Directorate and no official endorsement should be inferred.

REFERENCES

[1] A. E. Stevens, P. T. Rourke and E. A. Rietman, "Explosive detection in the presence of clutter by processing Raman spectra with a kernel adatron," Proc. SPIE 7345, Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2009,, 2009.

[2] M. G. Madden and A. G. Ryder, "Machine learning methods for quantitative analysis of Raman spectroscopy data," Proc. SPIE 4876, Opto-Ireland 2002: Optics and Photonics Technologies and Applications,, 2003.

[3] C. J. Breshike, C. A. Kendziora, N. Budack, Y. Yoon, R. Furstenberg, V. Nguyen and R. A. McGill, "Active LWIR hyperspectral imaging and algorithms for rapid standoff trace chemical identification," in *Proc. SPIE 10986, Algorithms, Technologies, and Applications for Multisepctral and Hyperspectral Imagery XXV*, Baltimore, 2019.

[4] S. Hwan An, J. A. Martin and J. Meola, "Longwave Infrared Hyperspectral Subpixel Target Detection with Machine Learning (Preprint)," 2020.

[5] J. Liu, M. Osadchy, L. Ashton, M. Foster, C. J. Solomon and S. J. Gibson, "Deep Convolutional neural networks for Raman spectrum recognition: a unified solution," *The Analyst*, no. 21, 2017.

[6] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems*, 2012.

[7] C. A. Kendziora, R. Furstenberg, C. J. Brekshike, D. Finton, D. C. Kendziora, T. Huffman and R. A. McGill, "Algorithms for identification of trace explosives by active infrared backscatter hyperspectral imaging," Proc. SPIE 11727, Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXVII, 2021.

[8] C. F. Bohren and D. R. Huffman, Absorption and Scattering of Light by Small Particles, Weinheim: Wiley-VCH, 2004.

[9] J. P. Giblin, J. R. Dupuis, J. P. Dixon, J. M. Hensley, D. J. Mansur and W. J. Marinelli, "Active Standoff Chemical Identification Detector," in *Proc. SPIE 10629, Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XIX*.

[10] B. &. T. R. Efron, An introduction to the bootstrap, London: Chapman & Hall/CRC, 1993.

[11] D. Manolakis, E. Truslow, M. Pieper, T. Cooley and M. Brueggeman, "Detection Algorithms in Hyperspectral Imaging Systems: An Overview of Practical Algorithms," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 24-33, 2014.

[12] M. D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks," in *ECCV*, 2014.

[13] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *CoRR*, vol. abs/1409.1556, 2014.