



AFRL-RY-WP-TP-2020-0136

**LONGWAVE INFRARED HYPERSPECTRAL SUBPIXEL
TARGET DETECTION WITH MACHINE LEARNING
(Preprint)**

**Seung Hwan An, Jacob A. Martin, and Joseph Meola
EO Target Detection & Surveillance Branch (AFRL/RYMT)
Multispectral Sensing & Detection Division**

**MAY 2020
Final Report**

Approved for public release; distribution is unlimited.

See additional restrictions described on inside pages

STINFO COPY

**AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE**

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YY) May 2020	2. REPORT TYPE Conference Paper Preprint	3. DATES COVERED (From - To) 3 April 2020 –3 April 2020
--	--	---

4. TITLE AND SUBTITLE LONGWAVE INFRARED HYPERSPECTRAL SUBPIXEL TARGET DETECTION WITH MACHINE LEARNING (Preprint)	5a. CONTRACT NUMBER N/A
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER N/A

6. AUTHOR(S) Seung Hwan An, Jacob A. Martin, and Joseph Meola	5d. PROJECT NUMBER N/A
	5e. TASK NUMBER N/A
	5f. WORK UNIT NUMBER N/A

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) EO Target Detection & Surveillance Branch (AFRL/RMYT) Multispectral Sensing & Detection Division Air Force Research Laboratory, Sensors Directorate Wright-Patterson Air Force Base, OH 45433-7320 Air Force Materiel Command, United States Air Force	8. PERFORMING ORGANIZATION REPORT NUMBER
---	---

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) <table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border-right: 1px solid black; padding: 5px;">Air Force Research Laboratory Sensors Directorate Wright-Patterson Air Force Base, OH 45433-7320 Air Force Materiel Command United States Air Force</td> <td style="width: 50%; padding: 5px;">Air Force Office of Scientific Research Public Affairs (AFOSR) 875 North Randolph Street, Suite 325, Room 3112 Arlington, VA 22203</td> </tr> </table>	Air Force Research Laboratory Sensors Directorate Wright-Patterson Air Force Base, OH 45433-7320 Air Force Materiel Command United States Air Force	Air Force Office of Scientific Research Public Affairs (AFOSR) 875 North Randolph Street, Suite 325, Room 3112 Arlington, VA 22203	10. SPONSORING/MONITORING AGENCY ACRONYM(S) AFRL/RMYT
	Air Force Research Laboratory Sensors Directorate Wright-Patterson Air Force Base, OH 45433-7320 Air Force Materiel Command United States Air Force	Air Force Office of Scientific Research Public Affairs (AFOSR) 875 North Randolph Street, Suite 325, Room 3112 Arlington, VA 22203	
11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S) AFRL-RY-WP-TP-2020-0136			

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
PAO case number 88ABW-2020-1235, Clearance Date 3 April 2020. To be presented at SPIE Defense +Commercial Sensing, Anaheim, CA, 4/26/20. The U.S. Government is joint author of this work and has the right to use, modify, reproduce, release, perform, display, or disclose the work. Report contains color.

14. ABSTRACT
Hyperspectral imaging has been used to perform automated material detection and identification. However, traditional detection methods based on statistical data processing produce a higher than desired false alarm rate for subpixel targets due to violated assumptions. This paper compares performance of machine learning methods using neural networks in detecting subpixel targets with traditional statistical methods. The assessment will utilize airborne data collected by the SEBASS sensor under a variety of atmospheric conditions from a number of different altitudes. Various methods for atmospheric compensation and temperature-emissivity separation will be used as well to assess robustness of the detection approaches.

15. SUBJECT TERMS
hyperspectral, machine learning

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT: SAR	18. NUMBER OF PAGES 8	19a. NAME OF RESPONSIBLE PERSON (Monitor) Seung Hwan An
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			

Longwave Infrared Hyperspectral Subpixel Target Detection with Machine Learning

Seung Hwan An, Jacob A. Martin, Joseph Meola

Air Force Research Laboratory RYMT, Wright-Patterson AFB, OH 45433

ABSTRACT

Hyperspectral imaging has been used to perform automated material detection and identification. However, traditional detection methods based on statistical data processing produce a higher than desired false alarm rate for subpixel targets due to violated assumptions. This paper compares performance of machine learning methods using neural networks in detecting subpixel targets with traditional statistical methods. The assessment will utilize airborne data collected by the SEBASS sensor under a variety of atmospheric conditions from a number of different altitudes. Various methods for atmospheric compensation and temperature-emissivity separation will be used as well to assess robustness of the detection approaches.

1. INTRODUCTION

Hyperspectral imaging (HSI) has proven to be useful in detecting materials within the longwave infrared (LWIR) region (7-14 μ m). It is capable of detecting during both daytime and nighttime as it measures emitted thermal radiance instead of reflected sunlight. Most LWIR HSI detection algorithms rely on statistics-based template matching. Data-based classification methods have outperformed template matching algorithms for years in other fields and recent results have shown they can out-perform statistics-based methods on LWIR HSI as well.

One advantage of using HSI is the ability to detect targets that are not resolved, generally referred to as sub-pixel targets. Because they are smaller than the sensor's GSD, however, subpixel targets present unique challenges as their spectral features are mixed with the features of the background. A common statistical algorithm known as adaptive coherence estimator (ACE) is generally used to detect subpixel targets. Since a hypercone is extended from the mean of the background clutter in the direction of the target spectra, ACE can capture pixels that are mixtures of target and background. However, if there are a large number of pixels for two or more background types in a scene, the scene mean will be between the two background clutter distributions and physically meaningless for anything but pixels at a boundary between background types. Since the hypercone extends from this point towards the target, it may not capture target pixels that are mixed in low abundance with the true local background. Local mean ACE can circumvent this problem, but without a priori knowledge of the target size, the target will bias the whitening statistics making detection more difficult or even impossible.

This research examines and compares detection performance between statistical methods and neural networks. The statistical algorithm chain consists of multiple atmospheric compensation (AC) and temperature emissivity separation (TES) techniques used in conjunction with ACE. Detection was performed in both radiance and emissivity space. A novel neural network architecture is presented for detecting both pure- and sub-pixel targets. The number of mixed pixels detected at a given false alarm rate are used as a comparison metric. Additionally, general trends with respect to cloud cover and time-of-day are discussed.

2. METHODOLOGY

2.1 Data

The data was collected in Medicine Hat, Alberta in the summer of 2014. The Spatially Enhanced Broadband Array Spectrograph (SEBASS) sensor was used for this study which was mounted on a DHC-6-300 Twin Otter operating at altitudes between 1500' to 9000' above ground level (AGL).¹ The target site had various targets as shown in 1. Targets E1-E8 were 8x8 panels lying flat on the ground. E9-E16 were the same materials but tilted up 45 degrees relative to the ground. E17-E20 were 20x20 panels of varying emissivity lying flat on the ground.

A total of ten radiance cubes from this collection are used in the study varying collection altitude, different levels of cloud cover, and time-of-day/night. Five targets are chosen for this analysis: foamboard, a low emissivity paint (Low-E), glass, sandpaper and flame sprayed aluminum. The spectra of these materials are shown in ??.

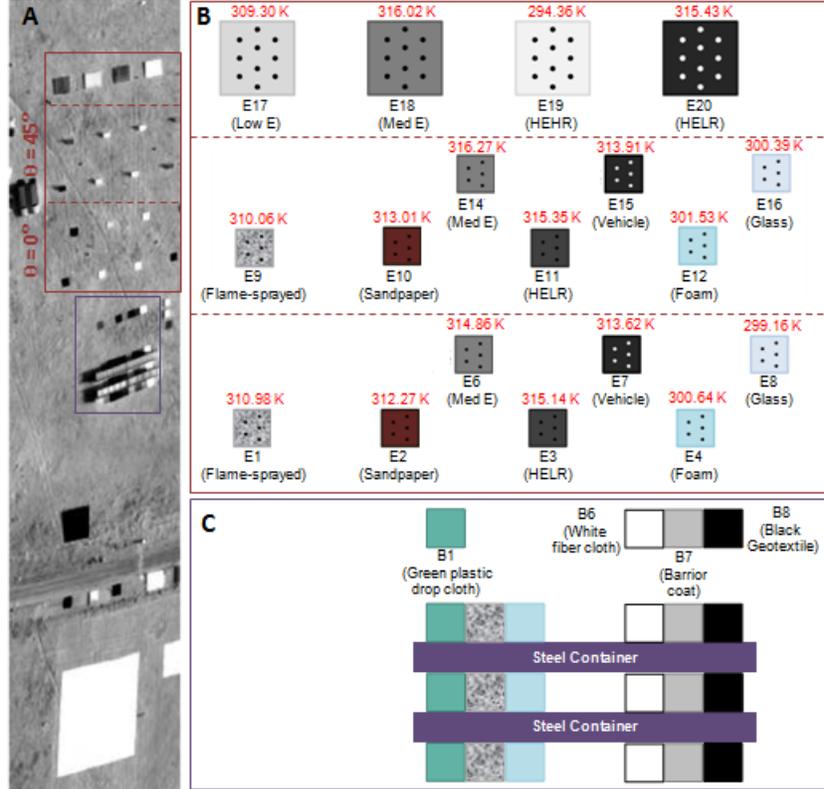


Figure 1: LWIR data and targets used for study

2.2 Statistical Method

A simple model for sensor-reaching radiance in the LWIR is given by,

$$L(\lambda) = L_u(\lambda) + \left[L_d(\lambda)(1 - \epsilon(\lambda)) + \epsilon(\lambda)B(\lambda, T) \right] \tau(\lambda), \quad (1)$$

where $L_u(\lambda)$ is the upwelling radiance, $L_d(\lambda)$ is the downwelling radiance, $\tau(\lambda)$ is atmospheric path transmission, $\epsilon(\lambda)$ is emissivity, and $B(\lambda, T)$ is the radiance of a blackbody at temperature T .² The process of atmospheric compensation seeks to estimate the transmission, upwelling, and downwelling (TUD) components of this model. The four different AC methods are used in this analysis: FLAASH-IR and three other similar LWIR AC algorithms which we refer to here as AC1, AC2, and AC3. These algorithms use known sensor altitude, band centers, band widths and optimize over water vapor, temperature, and ozone search parameters to fit the measured data.³

The max smoothness algorithm is the primary temperature emissivity separation method used in this work.⁴ The temperature range used for max smoothness was the mean brightness temperature of the scene plus/minus five Kelvin. Radiance signatures of the targets were forward modeled from their emissivity signature using TUD estimates and the max apparent spectral temperature of the mean background after conversion to ground leaving radiance.⁵ For all combinations of AC and TES techniques, signed ACE (Equation 2) was used since it is explicitly designed for sub-pixel targets and has proven to be extremely robust with variety of targets.⁵

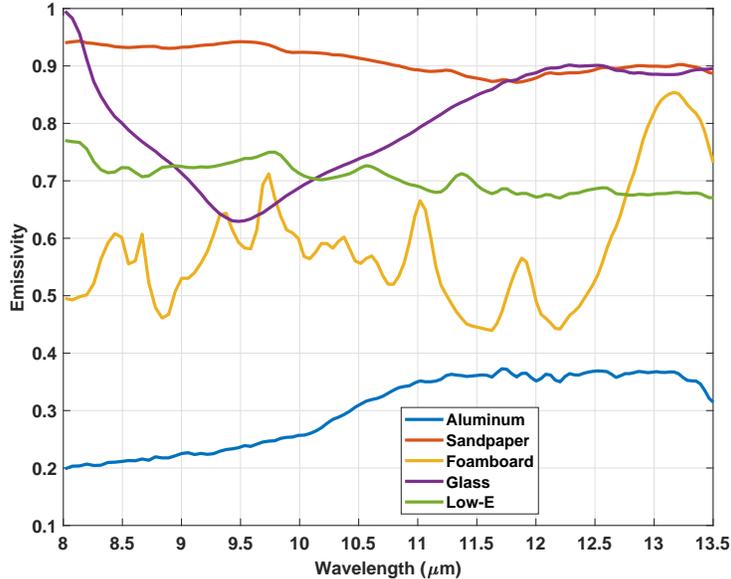


Figure 2: Measured target spectral emissivity

$$r_{ACE}(\mathbf{x}) = \frac{\mathbf{s}^T \hat{\Sigma}^{-1} \mathbf{x}}{\sqrt{\mathbf{s}^T \hat{\Sigma}^{-1} \mathbf{s} \sqrt{\mathbf{x}^T \hat{\Sigma}^{-1} \mathbf{x}}} \quad (2)$$

2.3 Neural Network

Two 50-node single layer neural networks were used for this work, one trained in at-aperture radiance-space and the other trained in emissivity-space requiring AC and TES preprocessing. Training for both of the ANNs was performed solely on synthetic data. This represents a worst-case scenario where no relevant real-world data has been measured. Results would almost certainly improve if the synthetic data was supplemented with at least some measured data. For the foreseeable future, however, at least some synthetic or semi-synthetic data will be required for training machine learning classifiers. No sensor has collected enough data under a wide enough variety of conditions to effectively train a classifier especially when attempting to account for mixed pixels.

To generate training data, ground truth emissivity measurements for the various targets were used in conjunction with a pre-existing background library from another source. Target samples were modeled with a randomly sampled target pixel fill between 5 and 100%. Additionally, the 3x3 pixel local mean was concatenated onto the feature vector. In modeling this 3x3 mean the target abundance was allowed to vary from 100% down to the single pixel target fill divided by nine, assuming all the target energy lands on the center pixel. In the future this could be improved by incorporating knowledge of the sensors point spread function. It is also important to note that even if the entire 3x3 area is covered by the target, the network will still be able to correctly classify the target which cannot be said for local mean ACE.

Target and background temperatures were allowed to vary independently with each being a random draw from a normal distribution with a mean of 320 K and a standard deviation of 20 K. Again, future work could improve results by using different temperature distributions for different background and targets. Additionally, incorporating information about the scene brightness temperature into the feature vector could constrain the range of temperatures to model. To account for atmospheric effects, a number of MODTRAN⁶ runs were performed using the TIGR database⁷ of global radiosonde data. Each of the 2311 TIGR atmospheres was run at 20 different sensor altitudes ranging from 500 to 10000 feet for a total of 46220 different transmission, upwelling, and downwelling (TUD) vectors. These TUD vectors were downsampled to SEBASSs spectral resolution assuming a

triangular spectral response function (SRF). Future work will examine using a Gaussian SRF as well as training multiple networks over a more limited range of altitudes.

Additionally, adjacency effects were added allowing ground shine to account for up to half of the hemispherical reflectance. This will also partially account for cloud cover as some of training samples will include background spectra that are very close to blackbodies. If it is known a priori that target tilt can be ignored, this training could be easily adapted to account solely for partial cloud cover. In total, 150000 samples for each target class were generated according to equations 3,4, and 5.

$$\begin{aligned}
L_1(\lambda, i) = & \tau(\lambda, i)[a_1(i)[1 - a_3(i)][1 - \epsilon_1(\lambda, i) + a_1(i)a_3(i)[1 - \epsilon_1(\lambda, i)B[T_2(i)]] \\
& + [1 - a_1(i)][1 - \epsilon_2(\lambda, i)]L_d(\lambda, i) + a_1(i)\epsilon_1(\lambda, i)B[T_1(i)] \\
& + [1 - a_1(i)]\epsilon_2(\lambda, i)B[T_2(i)]] + L_a(\lambda, i) + n(\lambda, i)
\end{aligned} \tag{3}$$

$$\begin{aligned}
L_2(\lambda, i) = & \tau(\lambda, i)[a_2(i)[1 - a_3(i)][1 - \epsilon_1(\lambda, i) + a_2(i)a_3(i)[1 - \epsilon_1(\lambda, i)B[T_2(i)]] \\
& + [1 - a_2(i)][1 - \epsilon_2(\lambda, i)]L_d(\lambda, i) + a_2(i)\epsilon_1(\lambda, i)B[T_1(i)] \\
& + [1 - a_2(i)]\epsilon_2(\lambda, i)B[T_2(i)]] + L_a(\lambda, i) + n(\lambda, i)
\end{aligned} \tag{4}$$

$$L(\lambda, i) = [L_1(\lambda, i) \quad L_2(\lambda, i)] \tag{5}$$

Where $a_1 \in [0.05, 1]$ represents the target pixel fill in the test pixel, $a_2 \in [a_1/9, 1]$ is the target fill in the 3x3 average, and $a_3 \in [0, 0.5]$ represents the sky fill fraction in the downwelling term. τ , L_a , and L_d are the atmospheric TUD components, $\epsilon_{1,2}$ is the target/background emissivity, $T_{1,2}$ the temperature, and n represents random Gaussian noise. Background samples are generated similarly but without the adjacency term and half of the background training samples are forced to be pure pixels of a single background material.

To generate emissivity training samples, these radiance samples are corrected to emissivity using the atmospheric TUD vector used to generate each sample and an assumed temperature that may differ from that used to generate the radiance sample. How best to account for the errors in the atmospheric compensation process is still very much an open question and the subject of ongoing research. For the results presented in the next section, retrieved emissivities using AC3 were used for this emissivity-space network.

3. RESULTS

The objective of this work is to compare target detection performance of subpixel targets. Truth masks were generated using ACE with in-scene signatures for each cube. The truth mask was then either dilated or eroded with a 3x3 filter. Only pixels in the dilated truth mask and not in the eroded mask were considered as potential sub-pixel targets. Note that we are not using probability of detection, but rather total pixels detected, as a metric because it is unlikely that all of the pixels in the dilated mask have appreciable target energy. It is also very unlikely that a false alarm would occur within this dilated mask for the false alarm rates being used in evaluation. Table 3, 4, and 5 displays the total number of mixed pixels of each target detected by each algorithm chain averaged over all data cubes. The dark green highlights the algorithm with the greatest number of detections and the dark red denotes the worst performing algorithm. The light green blocks represent any algorithm that was within one standard deviation of the best, yellow represents any algorithm within two standard deviation and light red represents any algorithm past two standard deviations. Standard deviation is calculated as the sum in quadrature of the uncertainty in the the number of detected pixels for the two algorithms being compared. The uncertainty in the detected pixel counts is given by the square root of the count.

Algorithm	Aluminum	Sandpaper	Foamboard	Glass	Low-E
FLAASH-Emis	251 (1.4)	78 (9.6)	340 (5.9)	173 (3.9)	365 (7.6)
AFRL01-Emis	249 (1.5)	213 (1.9)	340 (5.9)	253	445 (4.8)
AFRL02-Emis	232 (2.3)	216 (1.7)	353 (5.4)	240 (0.6)	444 (4.9)
AFRL03-Emis	262 (0.9)	236 (0.8)	351 (5.5)	247 (0.3)	462 (4.3)
FLAASH-Rad	245 (1.7)	66 (10.5)	340 (5.9)	168 (4.1)	352 (8.1)
AFRL01-Rad	202 (3.7)	152 (5.0)	331 (6.2)	234 (0.9)	399 (6.4)
AFRL02-Rad	185 (4.5)	223 (1.4)	350 (5.5)	234 (0.9)	436 (5.1)
1D ANN-Rad	283	242 (0.5)	471 (1.3)	232 (1.0)	579 (0.6)
1D ANN Emis	238 (2.0)	253	511	223 (1.4)	601

Figure 3: Number of mixed pixels detected by each algorithm chain for every target with false alarm rate of 10^{-3} . The number in parentheses represents the number of standard deviations away from the maximum for a particular target.

Algorithm	Aluminum	Sandpaper	Foamboard	Glass	Low-E
FLAASH-Emis	202 (1.4)	50 (10.4)	290 (5.0)	141 (3.3)	324 (7.8)
AFRL01-Emis	199 (1.5)	164 (2.9)	292 (4.9)	191 (0.6)	408 (4.8)
AFRL02-Emis	188 (2.1)	181 (2.0)	301 (4.6)	195 (0.4)	419 (4.4)
AFRL03-Emis	212 (0.9)	202 (0.9)	303 (4.5)	195 (0.4)	413 (4.6)
FLAASH-Rad	203 (1.3)	47 (10.6)	291 (5.0)	144 (3.2)	311 (8.3)
AFRL01-Rad	167 (3.2)	128 (4.9)	287 (5.1)	182 (1.1)	342 (7.1)
AFRL02-Rad	164 (3.4)	196 (1.2)	293 (4.9)	203	407 (4.8)
1D ANN-Rad	231	206 (0.7)	376 (1.7)	179 (1.2)	533 (0.7)
1D ANN-Emis	173 (2.9)	220	424	184 (1.0)	556

Figure 4: Number of mixed pixels detected by each algorithm chain for every target with false alarm rate of 10^{-4} . The number in parentheses represents the number of standard deviations away from the maximum for a particular target.

Algorithm	Aluminum	Sandpaper	Foamboard	Glass	Low-E
FLAASH-Emis	157 (1.0)	38 (9.5)	242 (4.8)	111 (1.8)	296 (7.4)
AFRL01-Emis	174 (0.1)	134 (2.4)	257 (4.2)	125 (0.9)	352 (5.2)
AFRL02-Emis	175	152 (1.4)	259 (4.1)	129 (0.6)	371 (4.5)
AFRL03-Emis	168 (0.4)	160 (0.9)	258 (4.1)	125 (0.9)	378 (4.2)
FLAASH-Rad	161 (0.8)	42 (9.1)	245 (4.7)	110 (1.84)	268 (8.5)
AFRL01-Rad	133 (2.4)	104 (4.3)	253 (4.3)	127 (0.7)	293 (7.5)
AFRL02-Rad	140 (2.0)	177	240 (4.9)	128 (0.7)	364 (4.8)
1D ANN-Rad	175	158 (1.0)	306 (2.0)	139	504
1D ANN-Emis	144 (1.7)	161 (0.9)	360	130 (0.6)	486 (0.6)

Figure 5: Number of mixed pixels detected by each algorithm chain for every target with false alarm rate of 10^{-5} . The number in parentheses represents the number of standard deviations away from the maximum for a particular target.

The neural network-based approaches significantly outperform ACE regardless of the combination of AC and TES used for preprocessing. At all three false alarm rates, no single statistical method outperforms the ANN trained in radiance-space for more than one target. Additionally, the radiance-space ANN outperforms ACE regardless of preprocessing method for all targets except glass and 10^{-3} and 10^{-4} FAR and sandpaper at 10^{-5} . Even in these cases, the ANN is still within two standard deviations of the best performing algorithm combination. One possible explanation for the ANN performing comparatively poorly on the glass target is because the ANN was trained under the assumption that all targets are lambertian. Better modeling of BRDF effects would likely improve these results significantly. Even with these limitations it is still the best performing algorithm for glass at 10^{-5} FAR.

The performance of the ANNs trained in either radiance- or emissivity-space is very similar. The emissivity-space detector works slightly better across all targets at 10^{-4} FAR and slightly worse at 10^{-3} and 10^{-5} but none of these differences are statistically significant. The emissivity-space ANN significantly underperforms on the flame-sprayed aluminum target relative to the radiance-space detector. One reason for this may be that the emissivity-space detector is trained assuming the AC algorithm will perfectly estimate the atmosphere but the TES may mis-estimate the temperature. Any errors in the atmospheric correction would likely affect the flame-sprayed aluminum most since it is the most reflective target.

In general, the difference between the ANN techniques and ACE is amplified in both the cloudy and night-time cubes. For example, comparing cubes at 1500 altitude, AC3 detects a total of 41 fewer pixels (221 vs 262) in the night-time cube compared to day at 10^{-4} FAR. The ANN detectors are much more consistent. The radiance-space ANN detectors 297 (day) vs. 300 (night) total pixels. The emissivity-space ANN detects 296 pixels of the day-time cube and 292 for the night-time cube.

AC3 appears to perform the best out of the various pre-processing combinations, but the difference for most targets is not statistically significant. It should also be noted that AC3 contains its own method for TES so max smoothness is not used in that case. The statistics-based detectors operating in radiance-space tend to perform worse than otherwise identical methods that include max smoothness TES. Because the radiance signature of the targets were forward modeled by the mean brightness temperature of the scene whereas max smoothness estimates temperature in every pixel, this result was expected.

4. CONCLUSIONS

This work compares a neural network approach for sub-pixel target detection with ACE, a statistics-based detector designed specifically for detecting sub-pixel targets. General trends noted superior performance using the 1D ANN in emissivity space. Future work will consist of varying hyperparameters within the neural network architecture and testing with different target sites.

REFERENCES

1. B. M. Rankin, J. Meola, D. L. Perry, and J. R. Kaufman, "Methods and challenges for target detection and material identification for longwave infrared hyperspectral imagery," **9840**, 2016.
2. M. T. Eismann, *Hyperspectral Remote Sensing*, SPIE, Bellingham, Washington, 2012.
3. S. M. Adler-Golden, P. Conforti, M. Gagnon, P. Tremblay, and M. Chamberland, "Long-wave infrared surface reflectance spectra retrieved from telops hyper-cam imagery," *Proc. SPIE* **9088**, pp. 90880U–90880U–8, 2014.
4. C. Borel, "Error analysis for a temperature and emissivity retrieval algorithm for hyperspectral imaging data," *Proc. SPIE* **6565**, pp. 65651Q–65651Q–12, 2007.
5. N. P. Wurst, S. H. An, and J. Meola, "Comparison of longwave infrared hyperspectral target detection methods," *Proc. SPIE* **10986**, 2019.
6. A. Berk, P. Conforti, R. Kennett, T. Perkins, F. Hawes, and J. van den Bosch, "MODTRAN6: a major upgrade of the MODTRAN radiative transfer code," *Proc. SPIE* **9088**, pp. 90880H–90880H–7, 2014.
7. Atmospheric Radiation Analysis, "Thermodynamic initial guess retrieval (tigr)." <http://ara.lmd.polytechnique.fr/>, May 2010.