SAND2021-10684C

Sandia National Laboratories

# Hyperspectral Image Target Detection Using Deep Ensembles for Robust Uncertainty Quantification

Rajeev Sahay, Daniel Ries, Joshua D. Zollweg, Christopher G. Brinton

55th Asilomar Conference on Signals, Systems, and Computers,

U.S. DEPARTMENT OF ENERGY NNSA

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.

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.

# Presentation Outline

- Hyperspectral Image (HSI) Target Detection
- Deep Learning for HSI Target Detection
- Uncertainty Quantification Using Deep Ensembles
- Experiments
  - Dataset
  - Implementation
  - Results and Comparison with Bayesian Neural Networks (BNNs)
- Conclusion

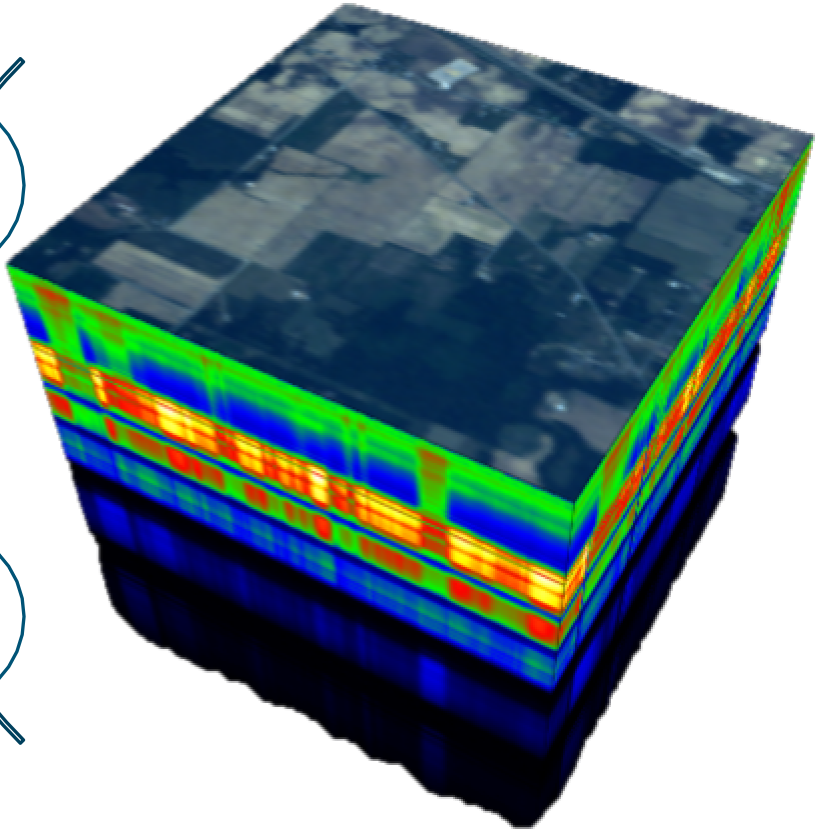
# Hyperspectral Image (HSI) Target Detection

## Hyperspectral Imagery

Strong material discriminant captured in spectral bands

High dimensionality

High co-linearity between spectral bands



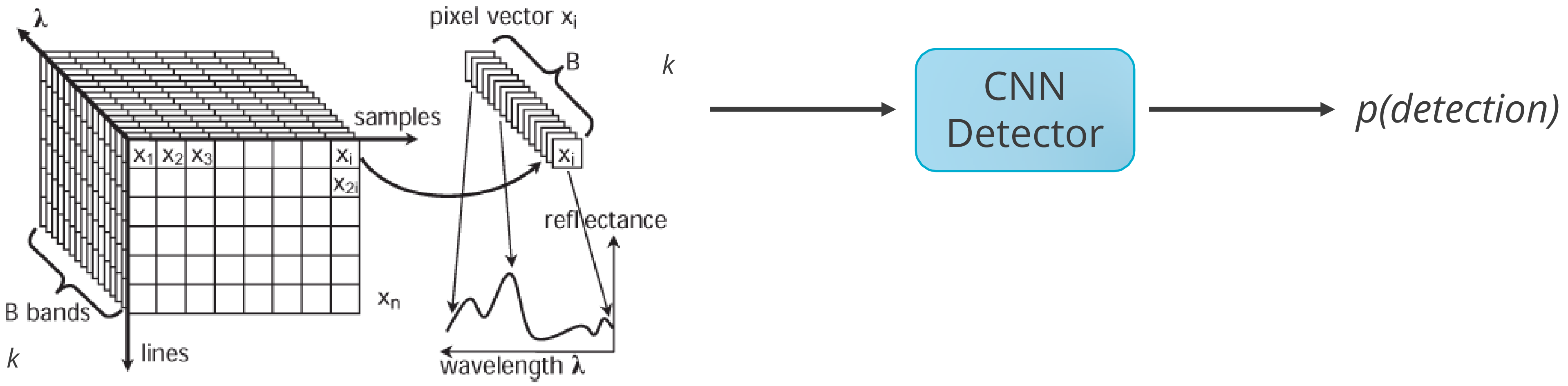
## Target Detection

Identification of one (or more) target material(s)

Class Imbalance - few target pixels per image

Low target abundance levels

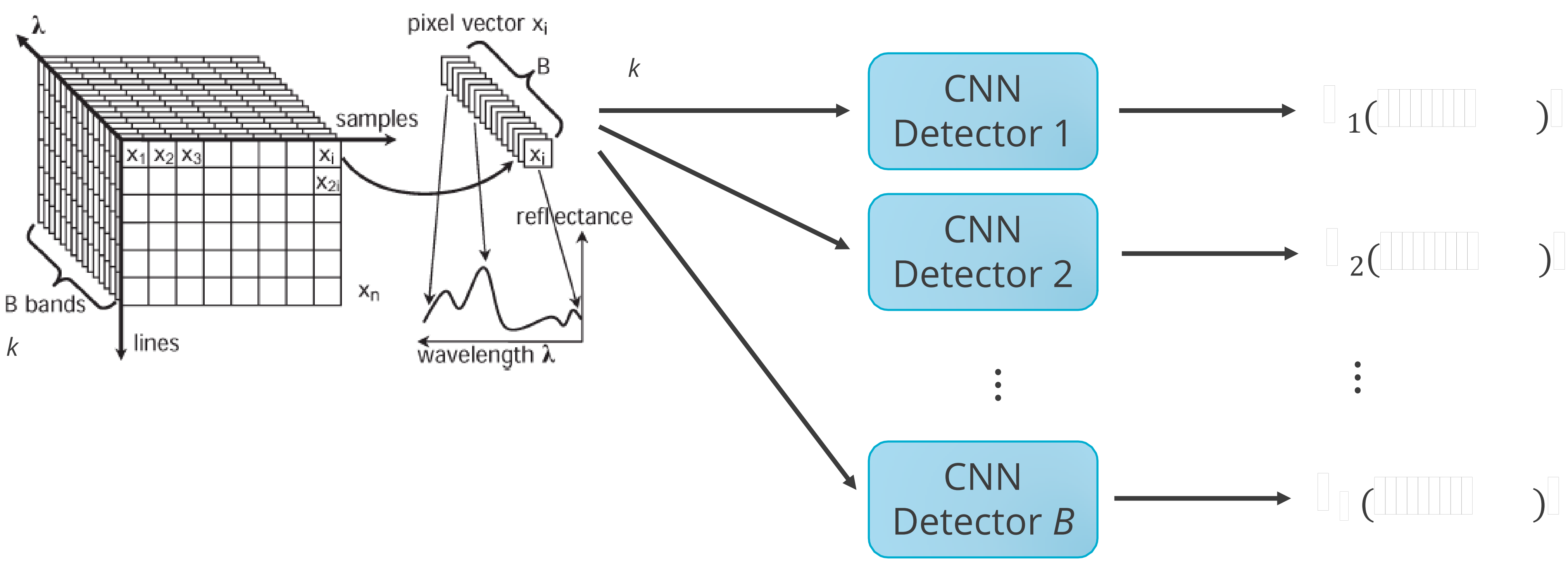
# Deep Learning for Hyperspectral Image (HSI) Target Detection




- CNN-based detectors have been found to be robust predictors for HSI target detection
- However, standard DL models produce point estimates at test time with no associated measure of uncertainty
- *Uncertainty quantification* (UQ) is vital in national security problems – especially on sub-filled pixels in which target abundance level is very little but we still need to be confident in its detection
- Note that 'probabilistic interpretation' of output models is not a good measure of confidence due to adversarial samples, wrong prediction with high p(detection), inability to detect trace materials, etc.
- We build upon the success of CNNs for HSI target detection by incorporating them into a UQ framework

Image modified from https://www.researchgate.net/figure/Hyperspectral-image-cube-and-a-typical-SRC-of-a-pixel-1_fig1_261500739

# Uncertainty Quantification Using Deep Ensembles




- Use an ensemble of identical CNNs to generate a distribution of predictions – then, use this distribution to characterize the (un)certainty associated with the model's prediction

- Train each detector using the same training data but with different random initializations – this creates diverse models
  - Also incorporate adversarial training to improve detection performance

# Uncertainty Quantification Using Deep Ensembles

Metrics of Interest

- The probability of detection: $\mathrm{p}(detection) = \frac{1}{\mathrm{B}} \sum_{i=1}^{B} p_i(detection)$

- Detection performance: area under the ROC curve (AUC)

- Confidence interval (CI): $p(detection) \pm z_\alpha \frac{s_i}{\sqrt{B}}$
  - $s_i$ - standard deviation of prediction over ensemble
  - $z_\alpha$ - $\left(1 - \frac{\alpha}{2}\right)$ quantile of Gaussian distribution

- High confidence (HC) set
  - Setting $\alpha = 0.2$, each sample for which (i) at least an 80% chance of containing the target, or (ii) at most a 20% chance of containing the target, is added to the HC set

- Probability of detection (PD) at a constant false alarm rate (CFAR)

# Experiments – Dataset

Employed Dataset: Megascene

- HSI scenes from upstate New York
- Targets: Green paint (manually inserted) with abundance varying between 0 – 100%

Three considered atmospheric scenes

1. Mid-latitude summer (MLS)
2. Sub-arctic summer (SAS)
3. Tropical (TROP)

Three considered times of day

1. 1200
2. 1430
3. 1545

Training scene: MLS 1200

# Experiments – Implementation

Consideration of 2 architectures

- Each architecture operates on the 25 functional principal components of the data as input

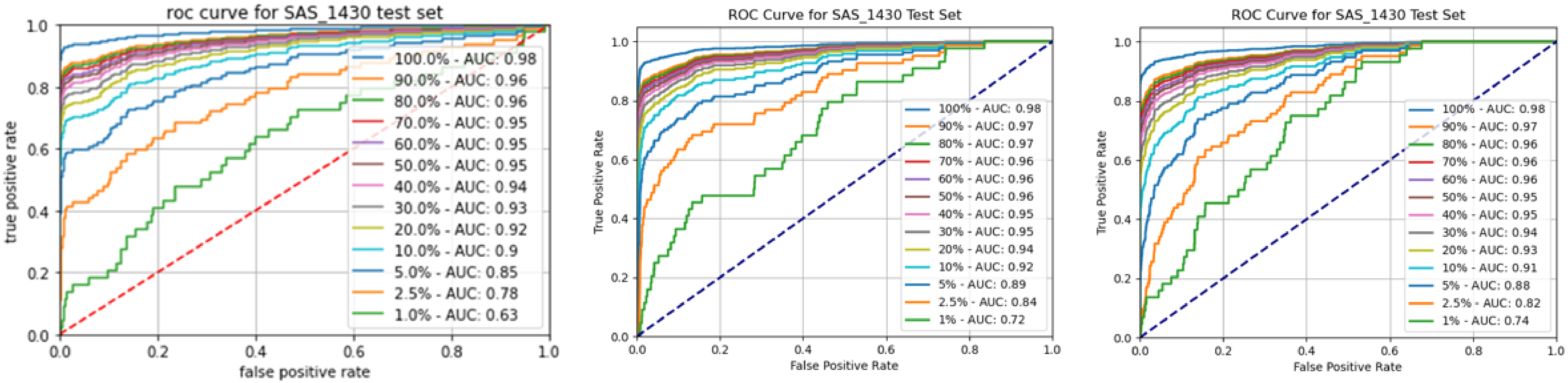
1. Convolutional Neural Network Detector
    - Hidden layer 1 (convolutional ReLU): $64 \times 4 \times 1$
    - Hidden layer 2 (convolutional ReLU): $32 \times 3 \times 1$
    - Hidden layer 3 (Dense ReLU): 128

2. Fully Connected Neural Network Detector
    - Hidden layer 1 (dense sigmoid): 10
    - Hidden layer 2 (dense sigmoid): 10
    - Hidden layer 3 (dense sigmoid): 10

Comparison UQ model: Bayesian Neural Network (BNN) trained using MCMC

- Same architecture as fully connected NN described above with $\mathcal{N}(0, 10)$ priors on all weights

# Experiments – Results



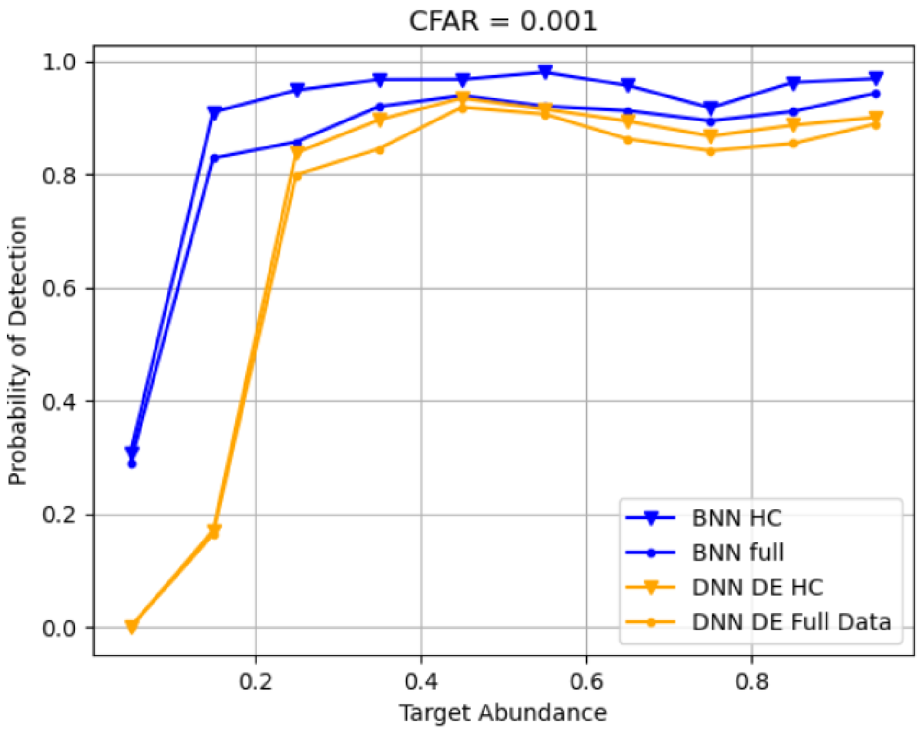

ROC curves for different target abundance levels on BNN (left), DNN DE (middle), and CNN DE (right) on SAS 1430.

# Experiments – Results

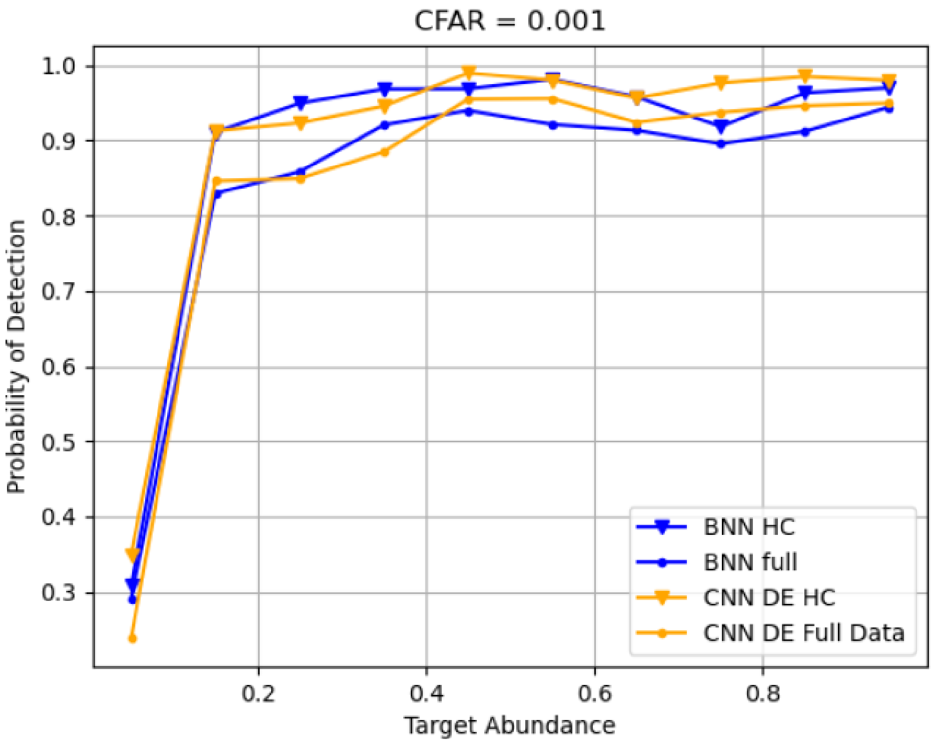
| Scene | BNN | DNN DE | CNN DE |
|---|---|---|---|
| MLS 1200 | 0.81 | 0.99 | 0.91 |
| MLS 1430 | 0.79 | 0.99 | 0.90 |
| MLS 1500 | 0.74 | 0.98 | 0.90 |
| SAS 1200 | 0.66 | 0.99 | 0.90 |
| SAS 1430 | 0.67 | 0.99 | 0.90 |
| SAS 1545 | 0.88 | 0.99 | 0.89 |
| TROP 1200 | 0.51 | 0.97 | 0.91 |
| TROP 1430 | 0.50 | 0.96 | 0.91 |
| TROP 1545 | 0.53 | 0.96 | 0.90 |

Proportion of samples captured in the HC set for each considered model

# Experiments – Results




Comparison between BNN and DNN DE of PD at CFAR = 0.001 (averaged across all nine scenes)




Comparison between BNN and CNN DE of PD at CFAR = 0.001 (averaged across all nine scenes)

# Conclusion

- Deep learning-based HSI target detectors demonstrate robust performance, but standard models lack measurements of uncertainty

- Our deep ensemble framework successfully incorporates the robustness of DL detectors while also providing UQ characteristics

- We showed that our framework serves as a robust detector and provides high probabilities of detection on low CFARs – often desirable for national security problems

Uncertainty quantification is critical in security sensitive deep learning applications