SAND2020-1870D

Sandia National Laboratories

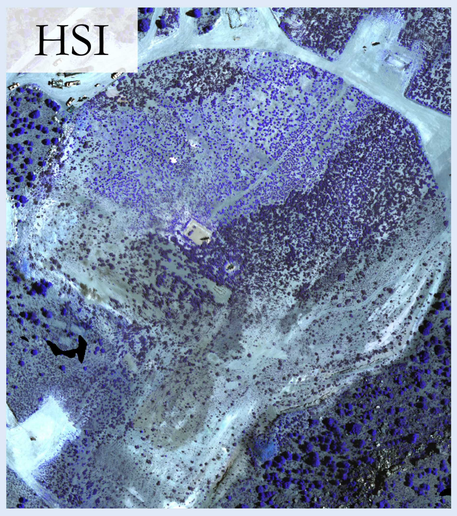
Hyper-spectral and Polarimetric SAR Data Fusion for Terrain Classification Using Probabilistic Feature Fusion

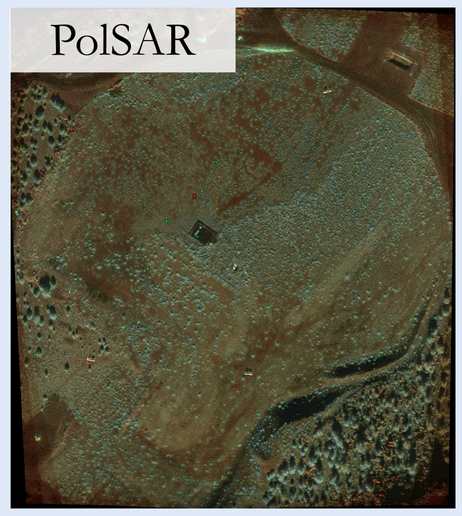
R. Derek West, David A. Yocky, Brian J. Redman, John D. van der Laan, and Dylan Z. Anderson

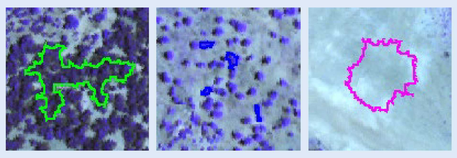
Introduction

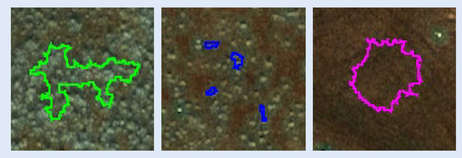
There exist a variety of remote sensing technologies that can be leveraged to extract useful information of a scene, such as determining if terrain-based non-prompt underground nuclear explosion terrain signatures are present. The choice of which technology to use depends on the information desired. It may be the case that multiple technologies are required in order to remove ambiguities in the information. A robust framework for sensor data fusion can leverage the strengths of multiple sensors, simultaneously, and lead to more knowledge. We demonstrate a data fusion framework with polarimetric SAR (PolSAR) and hyper-spectral image (HSI) data and show that fusing the sensor data improves terrain classification.

Images and Terrain Classes





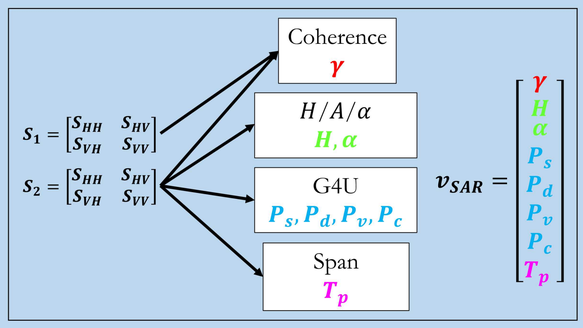


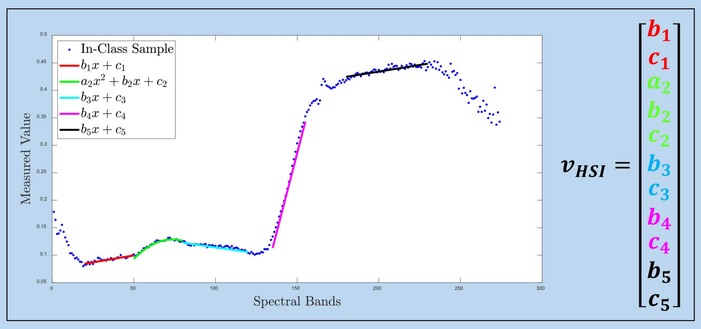


For terrain classification, co-registration of the acquiring data is of paramount importance. Misaligned data can produce erroneous results. We have been working on a method that is tolerant to a small amount of mis-registration of the images sets.

Feature Vector Representation

Feature vectors are an efficient representation of data. The elements should be defined to enable discrimination between terrain classes.



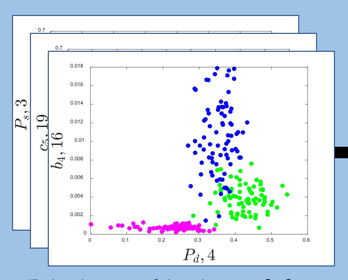


$$v_{HSI} = \begin{bmatrix} b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \\ b_3 \\ c_3 \\ b_4 \\ c_4 \\ b_5 \\ c_5 \end{bmatrix}$$

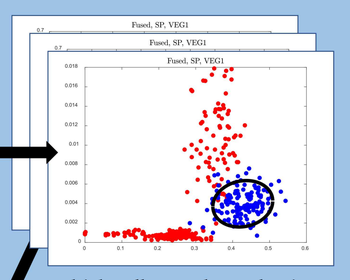
PolSAR features are derived from multiple passes of the scene, decomposition algorithms, and total power.

HSI features are the coefficients derived from fitting polynomials to sub-bands of the sampled spectrum.

Probabilistic Feature Fusion (PFF) Modeling

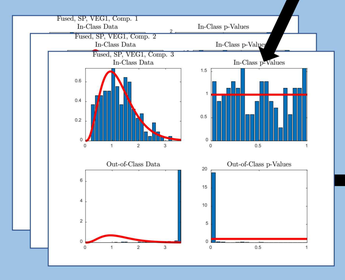


Pairwise combinations of feature vector elements are displayed for each class…

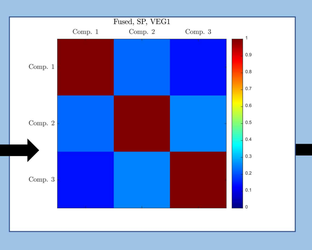


…which allows the selection of features that discriminates between classes.

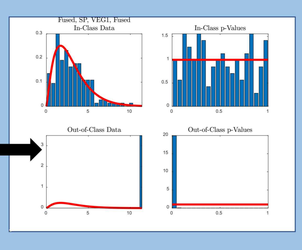
PFF is a frequentist approach for combining multiple observations or sources. It is a one-class classifier and only requires modeling in-class training data.



A metric is computed for the in-class training data for each selected feature pair. Probability distributions are then fit to the metric values.



The correlation is computed between the metric values and is compensated for in the fusion process.

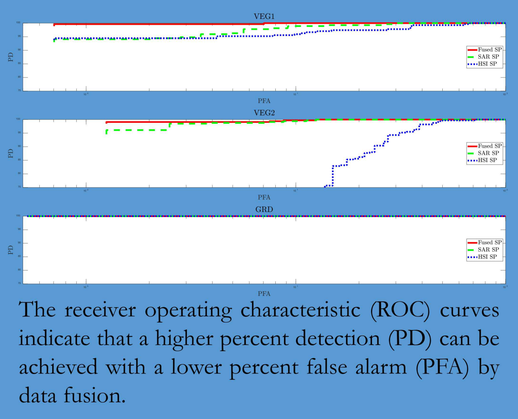


The features are fused together. The fused model produces p-values that give an indication of model consistency.

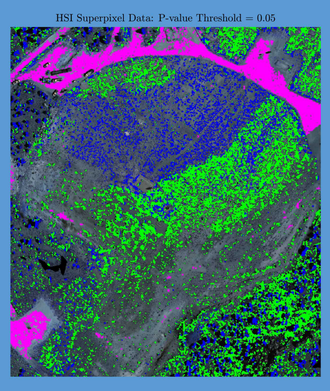
K. M. Simonson, R. D. West, R. L. Hansen, T. E. LaBruyere, and M. H. Van Benthem, "A statistical approach to combining multisource information in one-class classifiers," Statistical Analysis and Data Mining: The ASA Data Science Journal, vol. 10, no. 4, pp. 199–210, 2017. [Online].

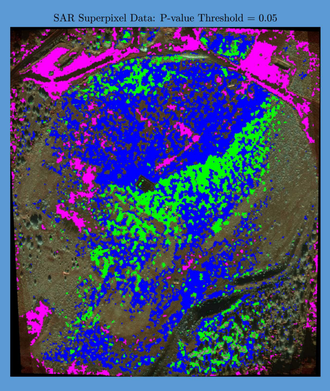
Results

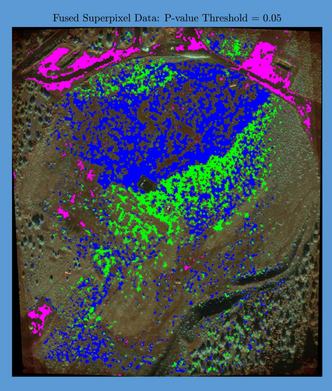
The confusion matrices indicate that fusing the PolSAR and HSI data increases the performance over each modality by its self.



The receiver operating characteristic (ROC) curves indicate that a higher percent detection (PD) can be achieved with a lower percent false alarm (PFA) by data fusion.







Inspection of the separate HSI and SAR PFF models applied to the image data highlights strengths and weakness of each modality. The fused PFF models demonstrate that the combined data is more selective and may be overcoming the shortcomings of the separate modalities and leveraging their strengths.

Conclusion

We demonstrated that with proper feature selection and modeling, data fusion from different remote sensing technologies can improve the terrain content knowledge of a scene. The impact of this is that more useful information can be derived in a more efficient manner. While the data fusion framework exemplar included PolSAR and HSI data for terrain classification, the framework is readily extensible to different sensing modalities and for more general purposes. Future work could include validating the framework on additional data sets.

U.S. DEPARTMENT OF ENERGY
National Nuclear Security Administration

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.
SAND No. _____

NEM 2020
webPMIS# SL19-SAR/HSI Fusion-PD2pa

Sandia National Laboratories