Segmentation of wildfires, smoke plumes, and burn scars using multi-sensor input and unsupervised and supervised machine learning for improved spatiotemporal coverage and facilitation of automated tracking

Nicholas LaHaye

Collaborators: Kyo Lee, Kyongsik Yun, Michael J. Garay, Alex Goodman, Olga Kalashnikova

ECMWF-ESA Workshop on Machine Learning for Earth Observation and Prediction
November, 15, 2022

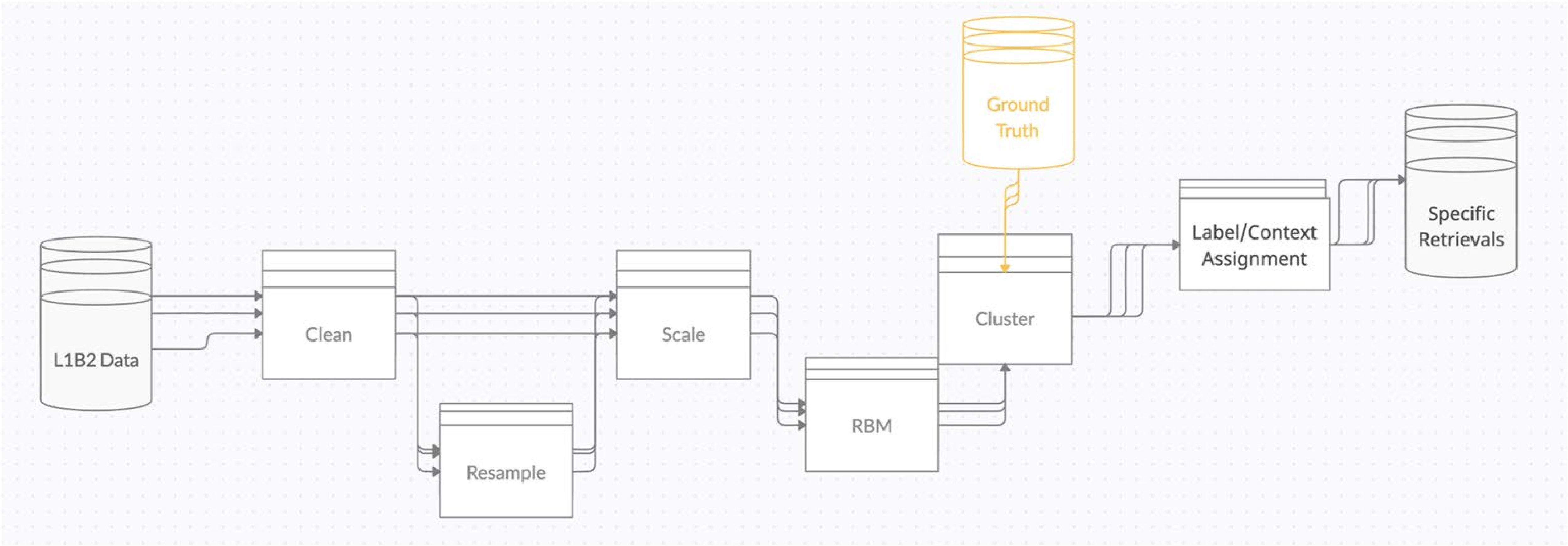
Jet Propulsion Laboratory
California Institute of Technology

Background

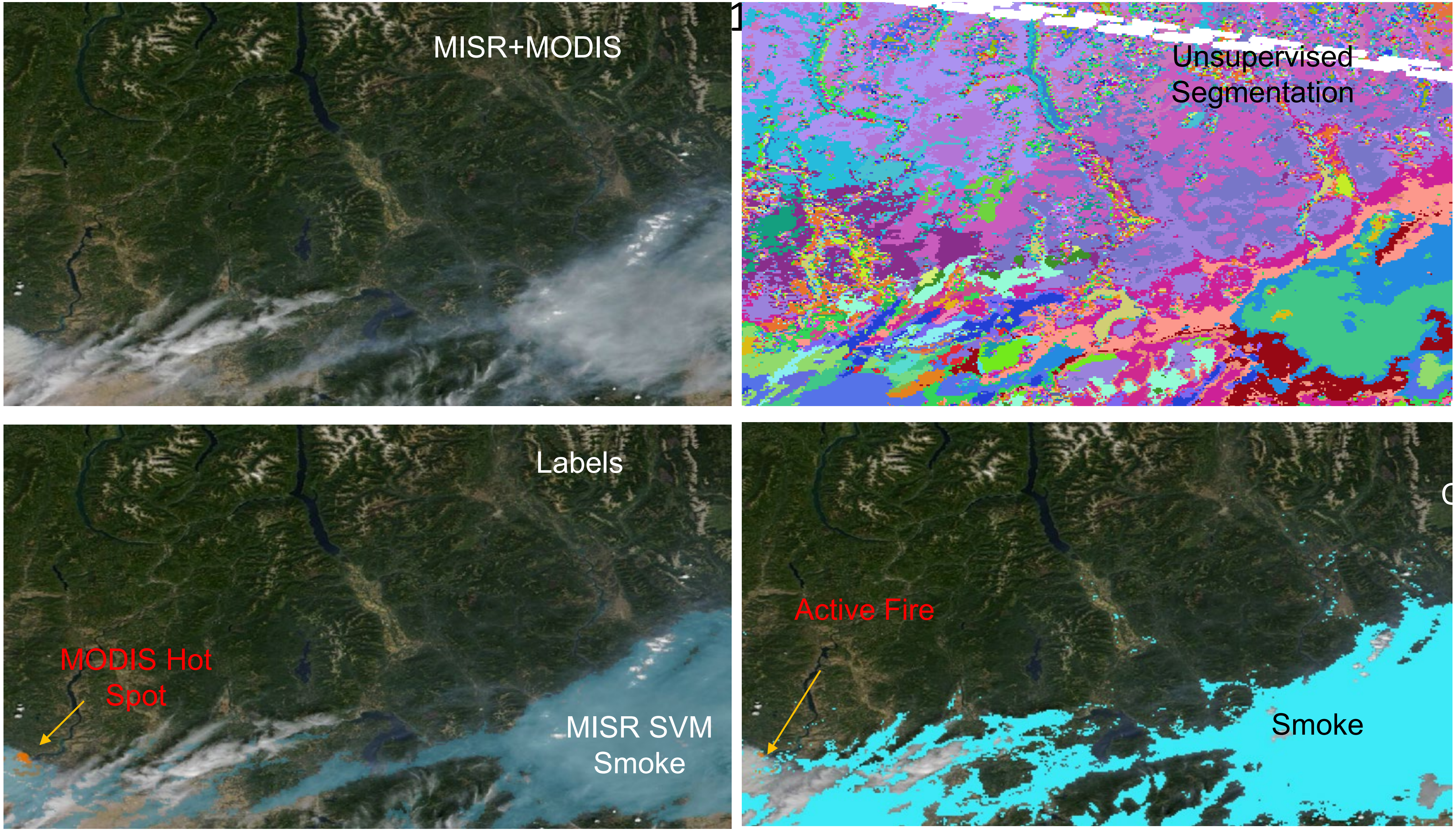
Vision for the Future

- Sensor Web
  - Ways to use the data generically
  - Automatic recognition of latent patterns
  - Simple combination of information
  - Tiered, interconnected view of data
- Issues
  - Different resolutions/grids
  - Instrument-specific modalities
  - Complexity increases with more data

Current Flow Chart - SIT-FUSE



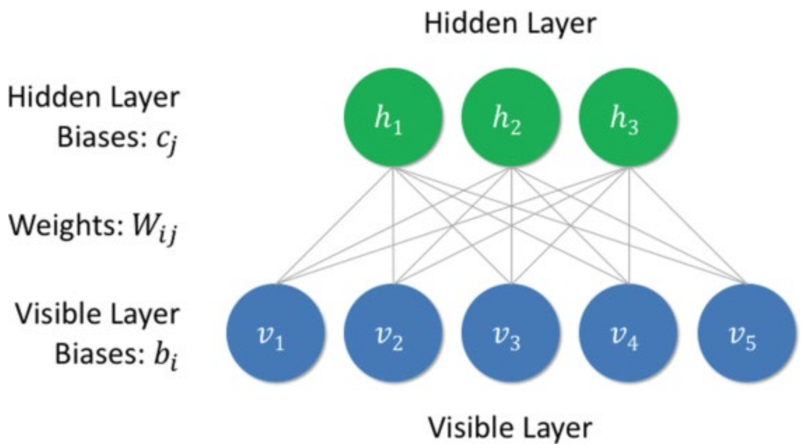
Williams Flats Fire 8 August

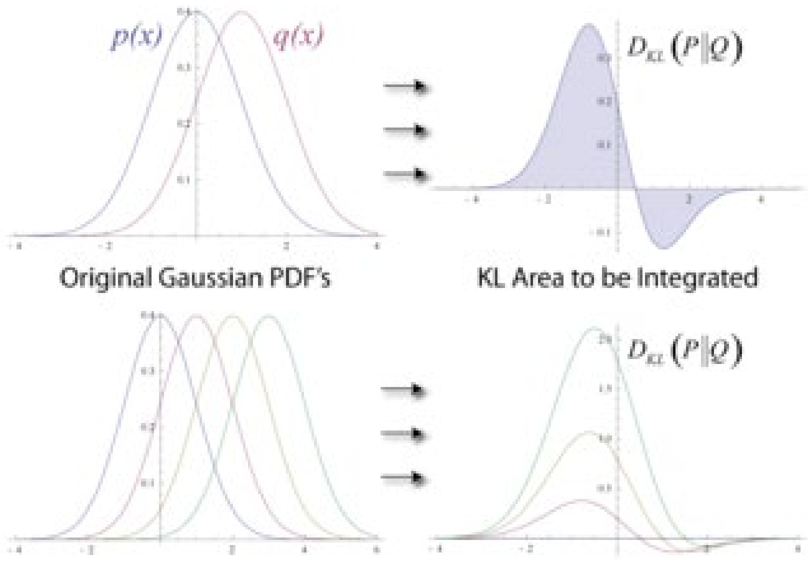


Burn Scar

RBMs

- Simple 2 layer learning architecture
  - Input layer
  - Hidden layer
- Unsupervised
- Generative
- "Restrictions"
  - Fully connected inter-layer
  - No intra-layer connections
- Goal
  - Reconstruct distribution of input as output
    - Generically
- Stackable - Creates Deep Belief Network (DBN)





RBM Learning

- Initialize
- Approximate samples from input distribution, given output (Gibb's Sampling)
- Attempt to reconstruct input, based on approximated samples, W, c, and b
- Adjust W, c, and b based on desire to minimize reconstruction error
- Repeat until convergence
    - Minimization of difference of free energy between p(x) and p_hat(x)

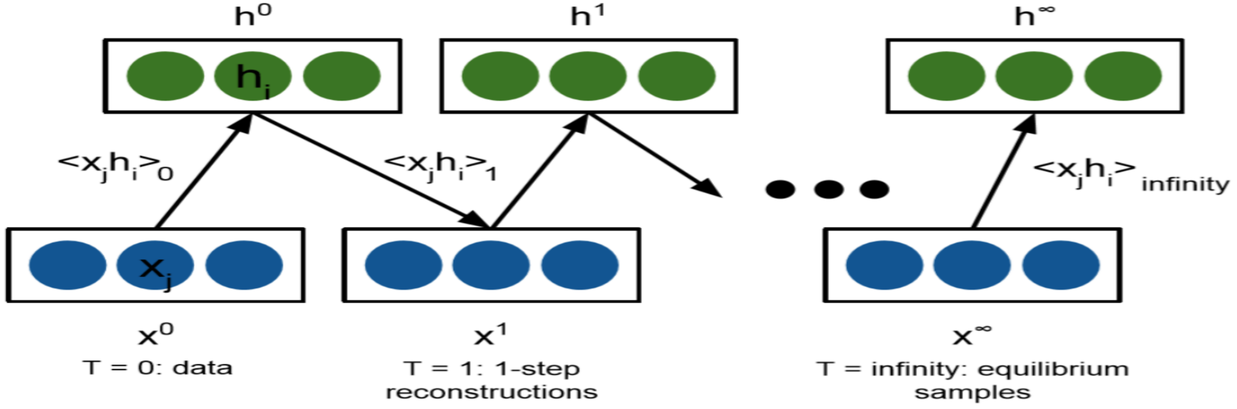
$$E(x,h) = -\left(\sum_{i=1}^{n}\sum_{j=1}^{m} x_j w_{ij} h_i + \sum_{j=1}^{m} b_j x_j + \sum_{i=1}^{n} c_i h_i\right)$$

$$p(x) = \frac{1}{Z}\sum_h \exp\left(-E(x,h)\right)$$

$$\mathcal{L} = \ln\frac{1}{Z}\sum_h \exp\left(-E(x,h)\right) = \ln\sum_h \exp\left(-E(x,h)\right) - \ln\sum_{x,h}\exp(-E(x,h))$$

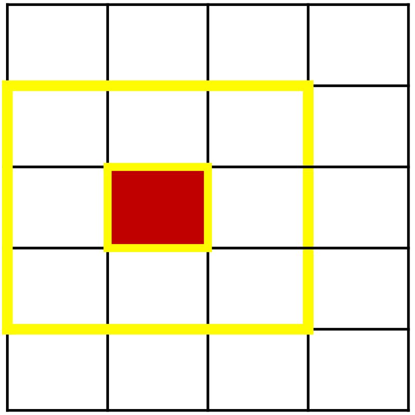
$$\nabla_{w_{ij}}\mathcal{L}_L = \langle x_j h_i\rangle_{data} - \langle x_j h_i\rangle_{model}$$

$$\nabla_{w_{ij}}\mathcal{L} \approx p(h_i = 1|x^0)x_j^0 - p(h_i = 1|x^k)x_j^k$$

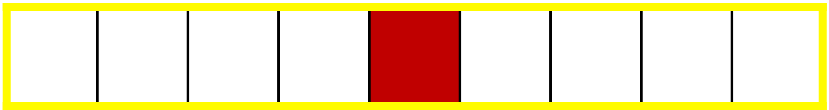


What Are We Reconstructing?

- Image Segmentation
  - Agnostic to scene size
  - Architecture can be used for 1-D and 2-D data
  - For 2-D, still require each input to represent its surroundings
    - Vectorize each pixel + all neighboring pixel
      - 2-D to 1-D Transformation
      - Each sample's feature-set consists of 9 pixels X instruments M channels
- What are we reconstructing?
  - The distribution of sets of pixels across scene(s) from a single instrument or set of instruments

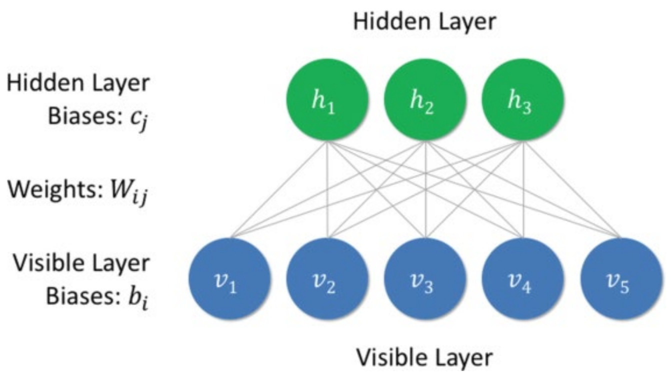


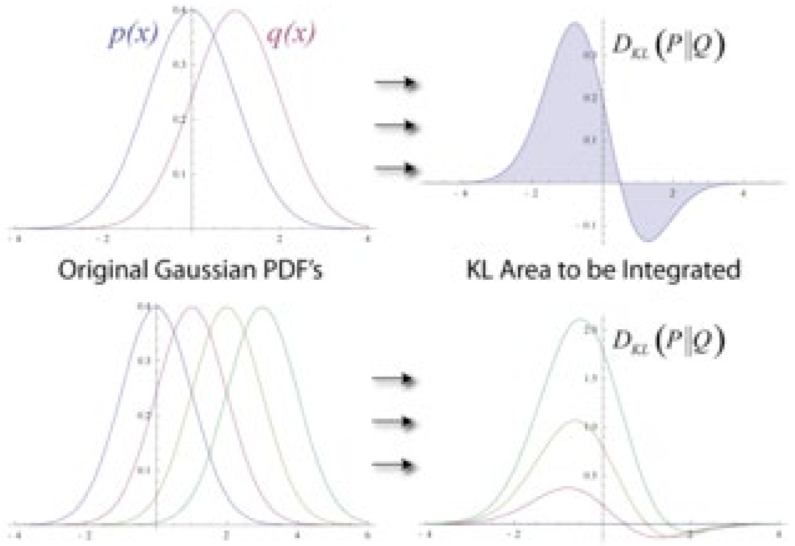
- In the future
- Test out convolutional RBMs in architecture similar to those that do supervised image/instance segmentation
  - Like Fully Convolutional Networks (FCNs)



RBM Architecture

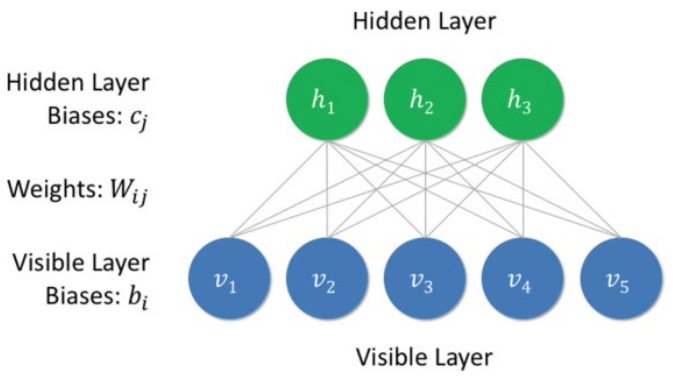
- Output/Hidden Layer size
  - < input size
    - Compressed representation
      - Potential loss of specificity
  - > input size
    - Expanded dimensionality
      - Non-linear relationships in original feature space more easily found/represented

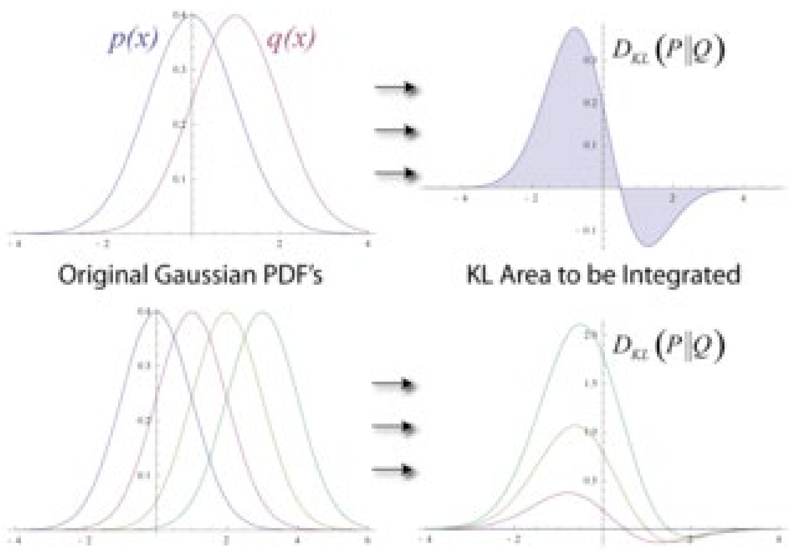




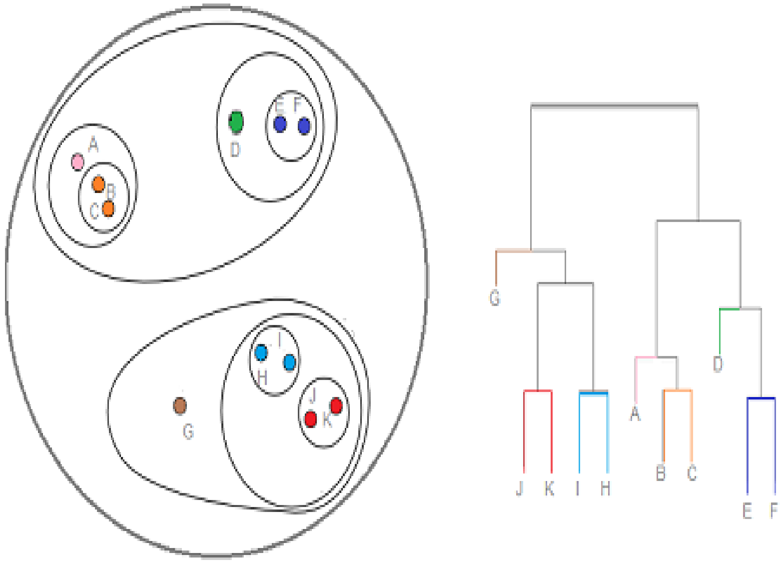
RBM - Output

- Machine interpretable matrix
  - N X H
  - Useful for further unsupervised learning
  - Not so useful for validation or further human-in-the-loop processing
- Need - A way to translate information back to human readable form



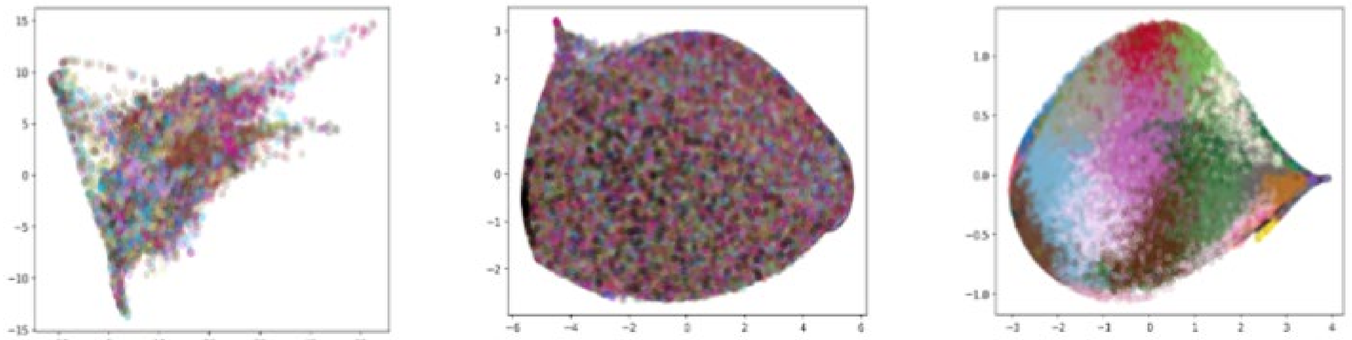


Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH)



- Agglomerative Clustering
  - Each sample is its own cluster
  - Merge based on heuristic until user-specified number of clusters remains
    - Heuristic here is merge that minimizes variance/variance increase
- BIRCH is memory efficient and fast(er)
  - Unlike many other agg. clustering methods
- Contains intermediate step of algorithm-defined number of clusters

Structure Visualization



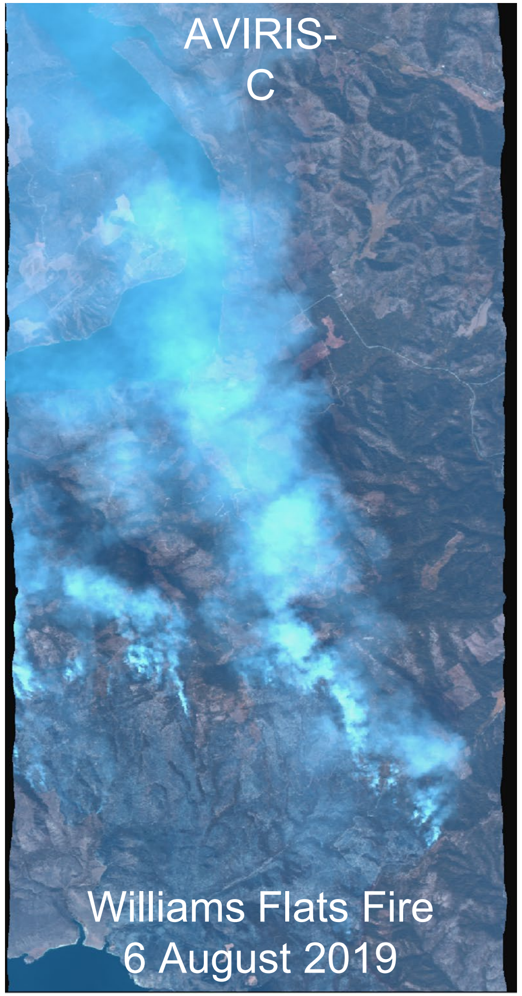
(a) Raw Landsat-8 + Sentinel-2 Fusion PCA

(b) Raw Landsat-8 + Sentinel-2 Fusion t-SNE

(c) RBM Landsat-8 + Sentinel-2 Fusion t-SNE

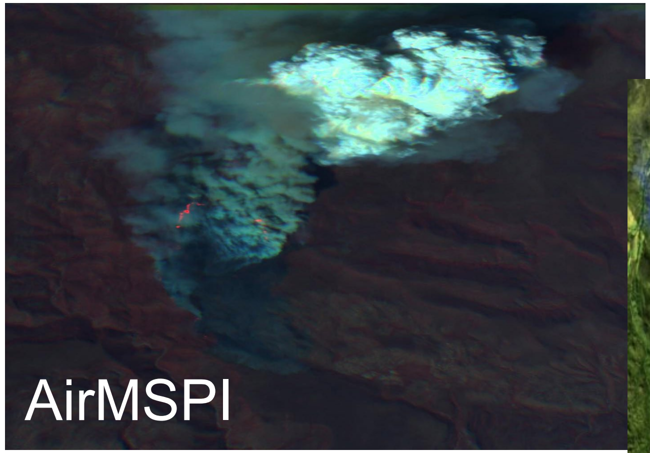
Application: Detection of Smoke Plumes and Active Fires

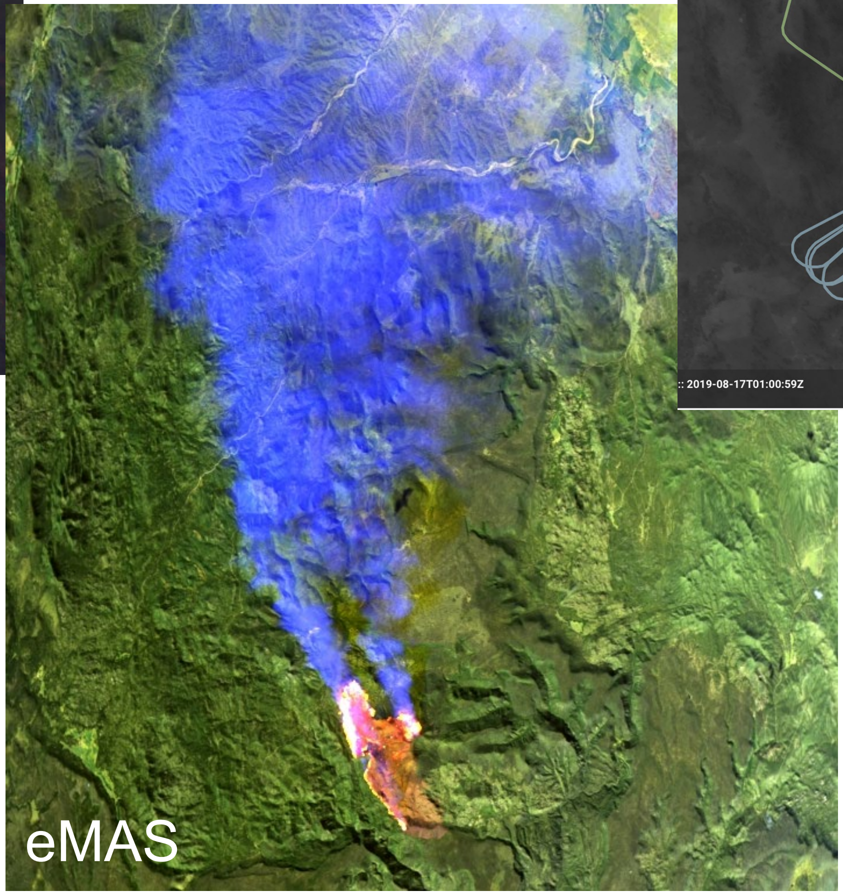
- Needs
  - Fire and smoke products are available from some instruments
    - Not easily accessible
    - Different data formats, content, structure
    - Not interoperable
  - Instrument specific implementation
- Issues
  - Confusion with clouds
  - Limited fire identification
- Methodology to make data more "plug-and-play"
- Per-instrument manual segmentation for supervised machine learning efforts would be very time consuming
  - Estimate ~30 minutes/ scene for millions of scenes

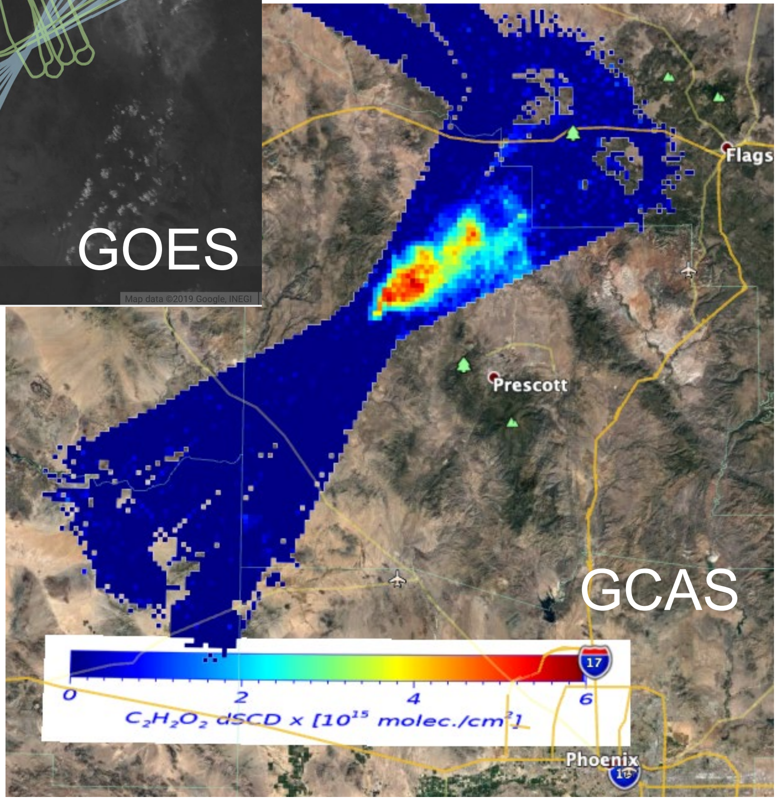


Background

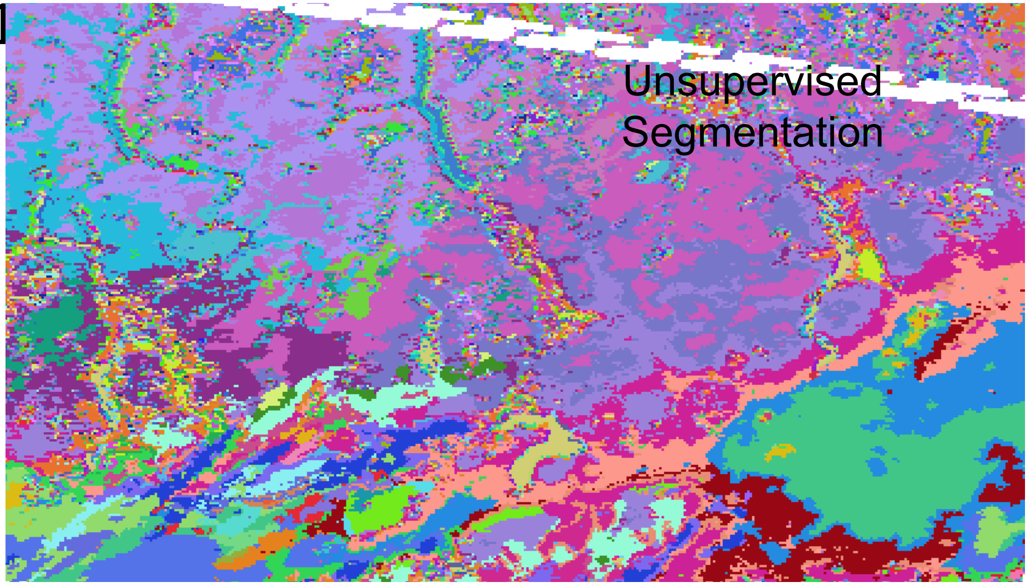
Remote Sensing Views of the Sheridan Fire from FIREX-AQ (16 August 2019)

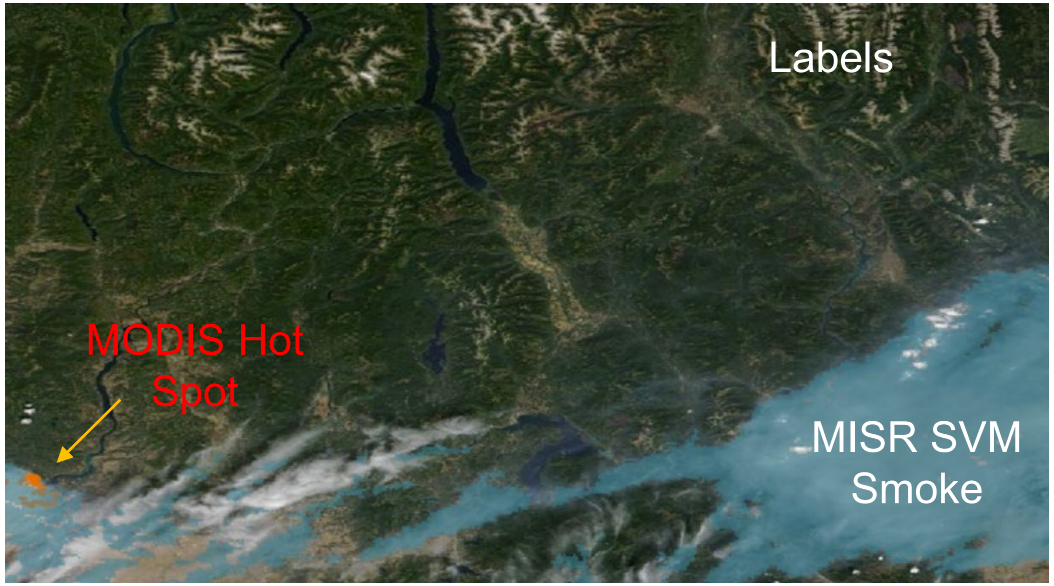






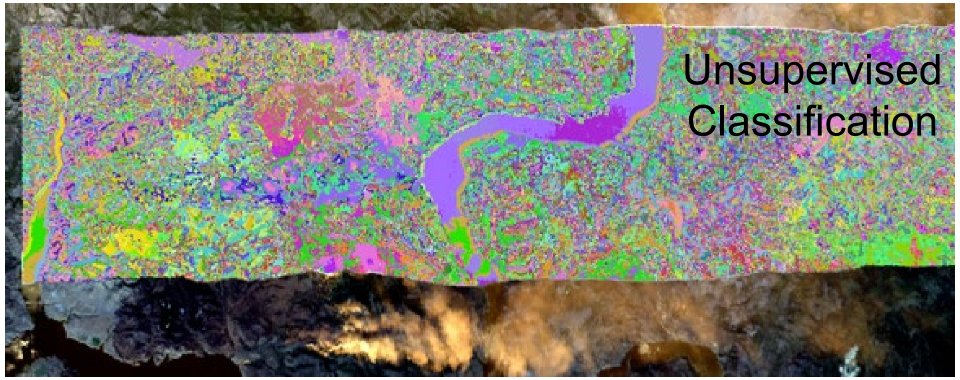
Williams Flats Fire 8 August





Burn Scar

Williams Flats Fire 6 August 2019



Williams Flats Fire 8 August 2019

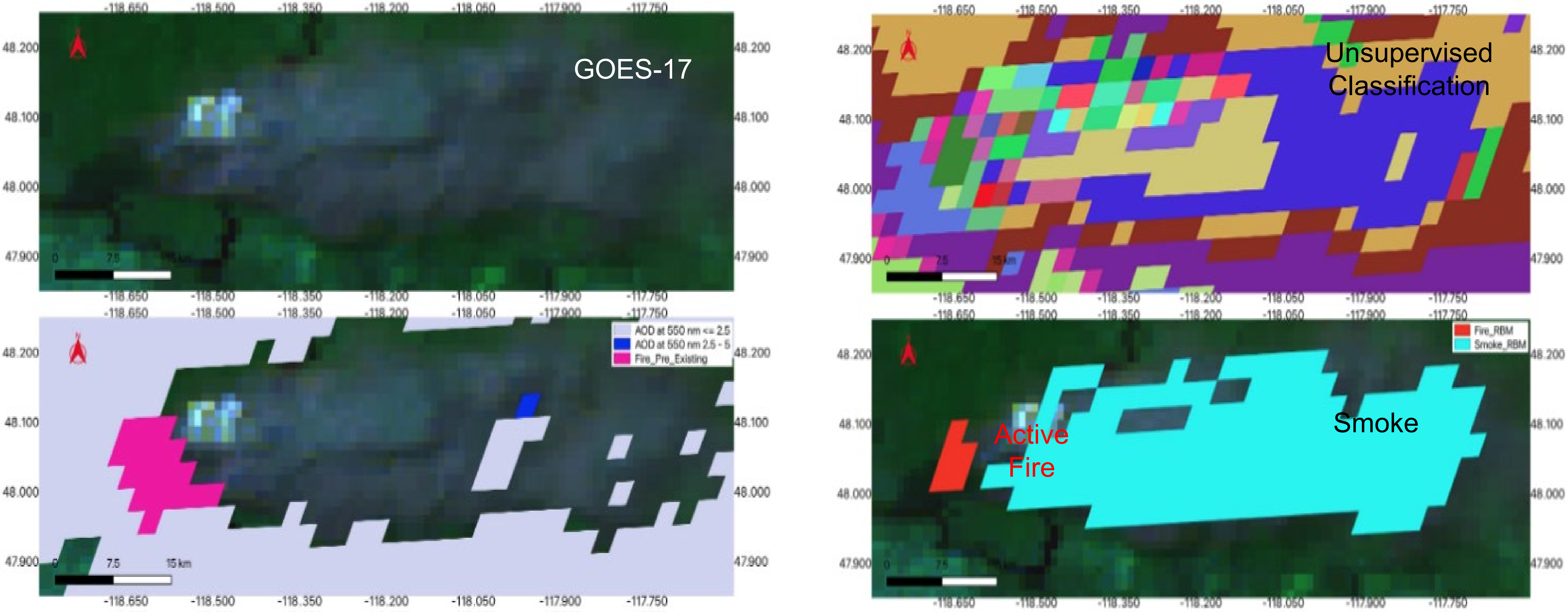


Table 3: Summary of Fire Detection Comparison of subset of produced products to pre-Existing products. Total pixel count is the total number of pixels tested. Pre-existing fire pixel count is the number of pixels labeled as fire in the pre-existing product. RBM-based fire pixel count is the number of pixels labeled as fire in our product. % Agreement is the percentage of fire pixels in the pre-existing product that are also identified as fire in our product. % False positive is the % of fire pixels in our product that are clearly mislabeled. δ % True positive is the % change of fire pixel count within the tested pixels from the pre-existing products to our products. The first two rows evaluate our MASTER + eMAS fusion product against both the pre-existing MASTER fire detection product, and the eMAS fire detection product.

Dataset	Total Pixel Count	Pre-Existing Product Fire Pixel Count	RBM-Based Fire Pixel Count	% Agree-ment	% False Positive	δ % True Positive
MASTER	22087108	147369	131067	72.9	0.0	-11.0
eMAS	43614368	9255	21710	71.7	1.3	234.6
eMAS + MASTER Fusion	2492030	284	1227	86.6	0.0	432.0
VIIRS	40129041	41	374	81.8	0.0	912.1
GOES	18750000	38	31	75.0	0.0	-18.0
MISR + MODIS Fusion	5497240	77	55	63.2	0.0	-28.8

Table 4: Summary of smoke detection comparison of subset of produced datasets to manually segmented products. Total pixel count is the total number of pixels tested. Man. seg. smoke pixel count is the number of pixels labeled as smoke in the manually segmented product or in MISR's SVM product's smoke class. RBM-based smoke pixel count is the number of pixels labeled as smoke in our product. % Agreement is the percentage of smoke pixels in the pre-existing product that are also identified as fire in our product. % False positive is the % of fire pixels in our product that are clearly mislabeled. δ % True positive is the % change of smoke pixel count within the tested pixels from the manually segmented products to our products.

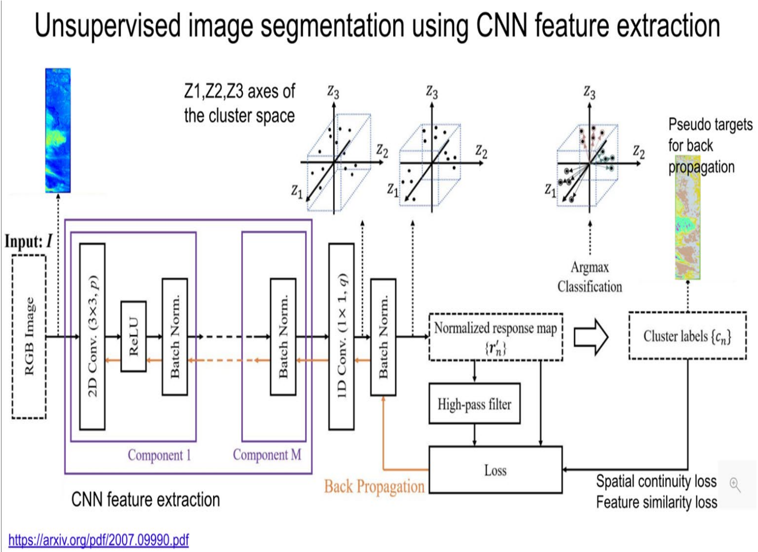
Dataset	Total Pixel Count	Baseline Smoke Pixel Count	RBM-Based Smoke Pixel Count	% Agree-ment	% False Positive	δ % True Positive
MASTER	22087108	43613537	42824132	91.4	7.5	-0.71
eMAS	11205711	2738986	2602037	85.6	2.1	-5.0
eMAS + MASTER Fusion	2492030	248108	230244	90.7	0.0	-7.2
GOES	18750000	50960	40462	73.6	5.8	0.0
MISR + MODIS Fusion	5497240	30884	34716	74.3	0.0	11.4

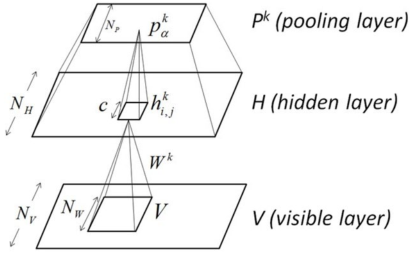
Use Cases

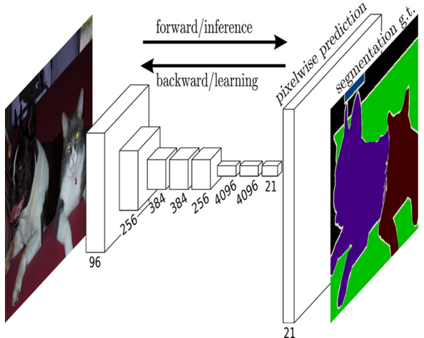
- Image Segmentation
  - Static scenes
    - Single and multi-sensor
    - Instrument and algorithm development
      - Minimal ground truths available
    - Current Use-Cases
      - Wildfire, smoke, and burn scar
      - Harmful algal bloom characterization
      - Water turbidity
      - Water bodies & ice
      - Ionospheric anomalies
    - Potential for subset classifications
      - Multiple clusters per final retrieval
    - Labels can be generated at a large enough scale and fed to a lightweight supervised model for operational use
      - On ground and onboard detection capabilities

Current Research

- Code optimization
- Uncertainty Quantification
- Convolutional Architectures for unsupervised feature extraction
- Expansion of regions + instrument sets for fire + smoke plume identification & tracking
- Include other kinds of instruments/tasks
  - Utilization of SAR, etc. for improvement of burn scar detection capabilities
- Open source release of data and software







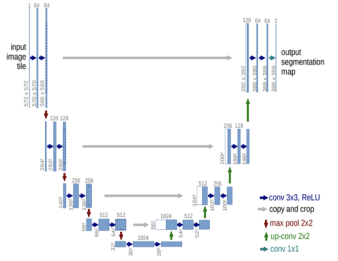
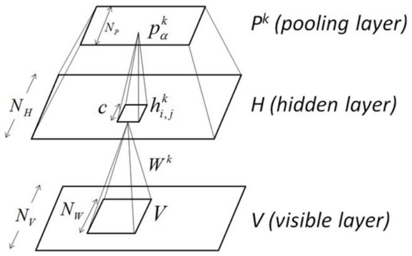
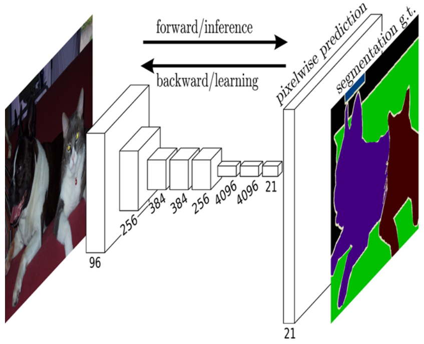


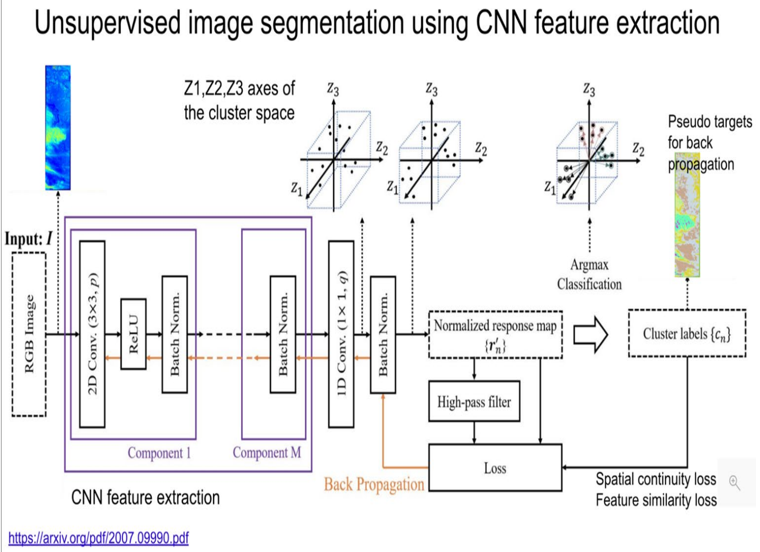
Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

Current Research - Convolutional Architectures

- Utilization of ConvRBM layers
  - Architecture like that of Fully Convolutional Network (FCN) - supervised instance segmentation architecture
  - Apply per-pixel clustering to result
- Future Work
  - Test other architectures that are successful in supervised domains
    - Traditional CNN
    - U-Net architecture with ConvRBM layers
  - Test semi-supervised approach with clustering integrated in training







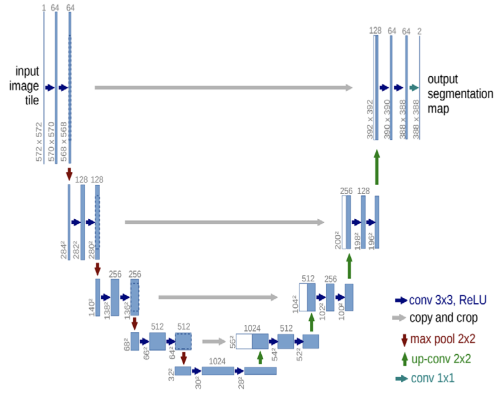
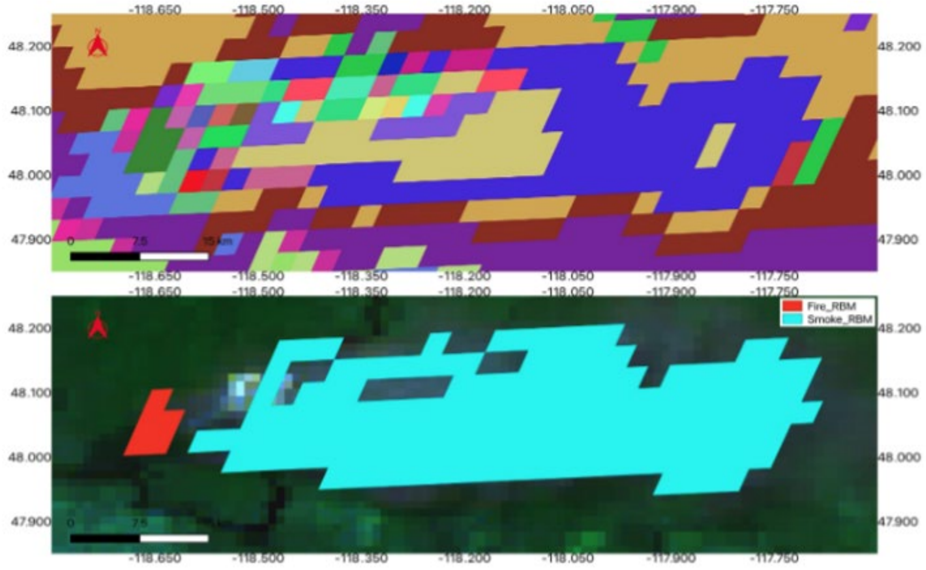
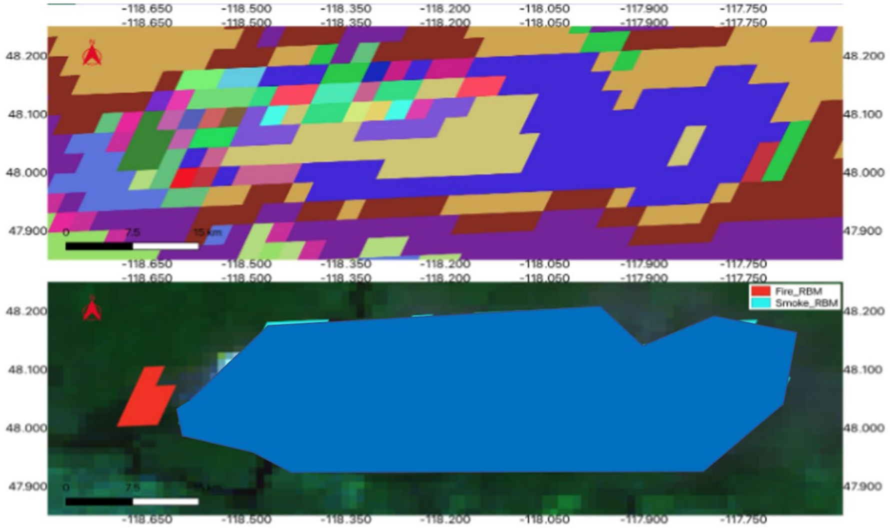


Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

Current Research - Shape Approximation/Tracking Metrics

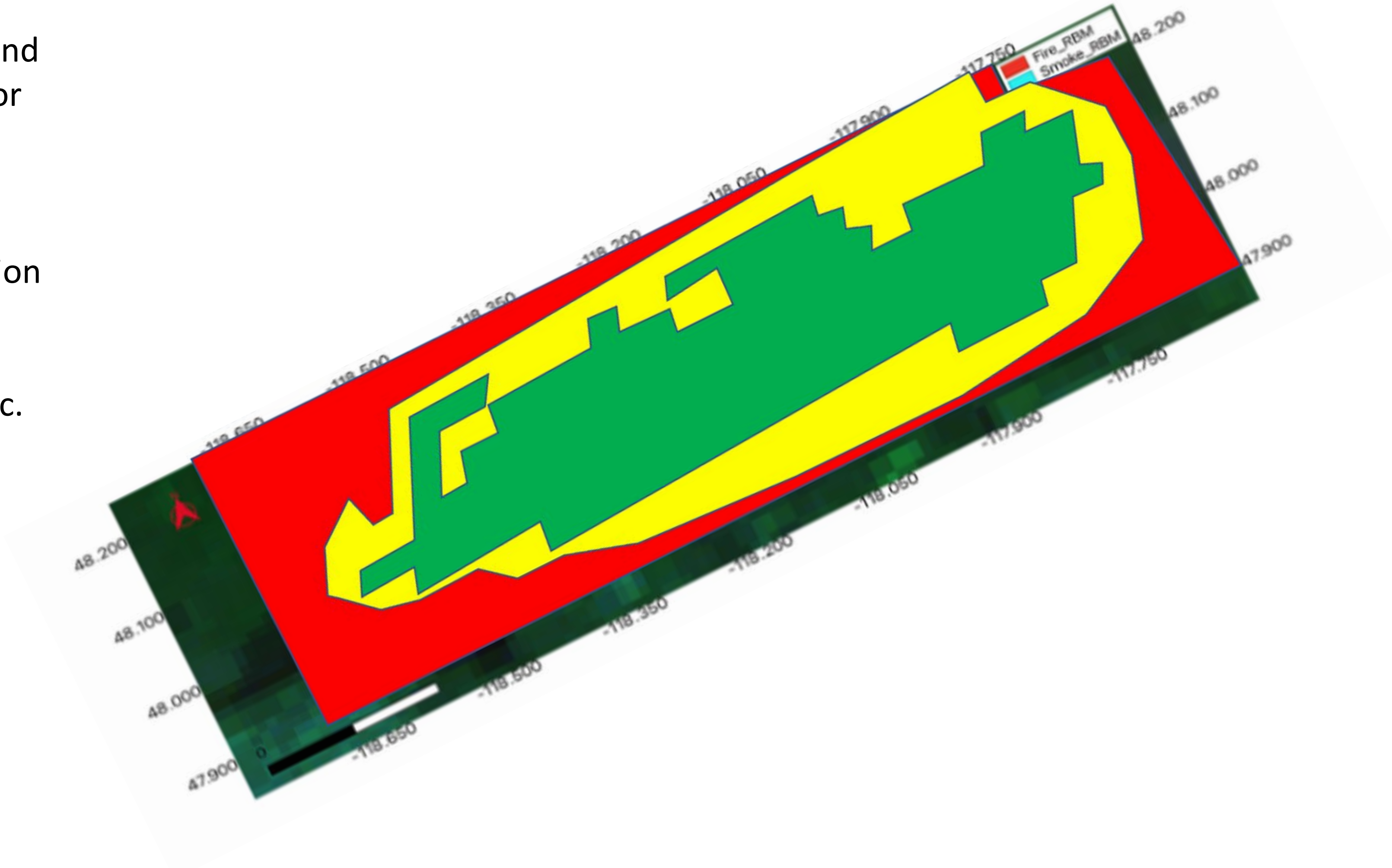
- Utilization of shape approximation and cluster distribution to provide a certainty/similarity of plumes in 2 separate scenes
    - Manually segmented “ground truth” proves insufficient for boundaries of things like aerosol plumes
    - Providing uncertainty associated with segmentation would prove useful when trying to only analyze main areas of aerosol plumes, etc.



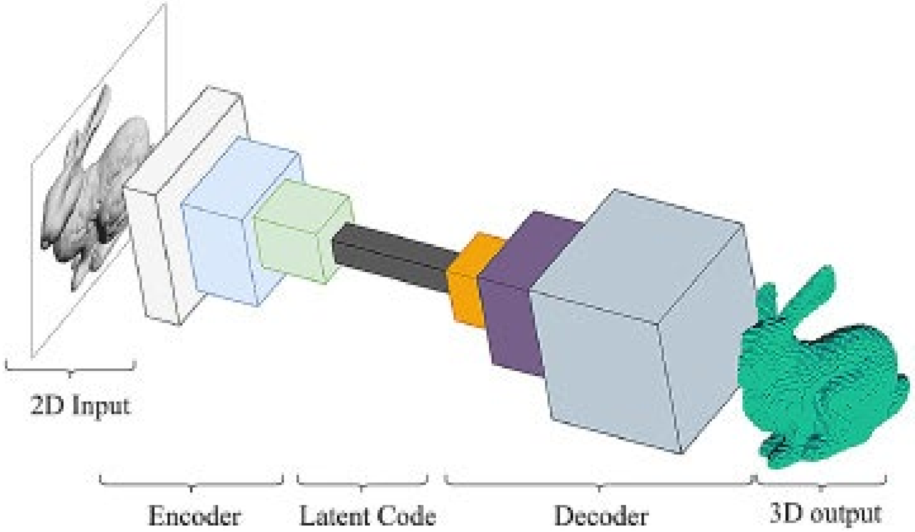


Current Research -
Uncertainty Quantification

- Utilization of pseudo-log-likelihood from RBM and uncertainty associated with clustering
  - Manually segmented “ground truth” proves insufficient for boundaries of things like aerosol plumes
  - Providing uncertainty associated with segmentation would prove useful when trying to only analyze main areas of aerosol plumes, etc.



- **Looking Forward**
  - Expansion of regions + instrument sets for current use cases
  - Updating web interface to allow for SME user corrections to be captured and applied via online learning
  - Provide support for future/new satellite-based missions like MAIA, EMIT, SBG, as well as future airborne missions
  - Utilize transfer learning to allow for hybrid ‘knowledge’ to be used in simpler agile onboard ML architectures
  - Test potential for 2-D and 3-D reconstruction



References and Funding Acknowledgement

1. Nicholas LaHaye, Jordan Ott, Michael J. Garay, Hesham El-Askary, Erik Linstead "Multi-Modal Object Tracking and Image Fusion With Unsupervised Deep Learning", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, issue 8, Aug. 2019, pages 3056–3066.
2. Nicholas LaHaye, Michael J. Garay, Brian Bue, Hesham El-Askary, Erik Linstead, "A Quantitative Validation of Multi-Modal Image Fusion and Segmentation for Object Detection and Tracking. Remote Sensing 2021", Remote Sensing, 2021, 13, 2364. https://doi.org/10.3390/rs13122364
3. Nicholas LaHaye, Kyongsik Yun, Huikyo Lee, Michael J. Garay, Alex Goodman, Hesham El-Askary, Krzysztof Gorski, Olga V. Kalashnikova, Erik Linstead, "Development and Application of Unsupervised Machine Learning for Smoke Plumes and Active Fires Identification from the FIREX-AQ Datasets", Remote Sensing, 2022, under revision.

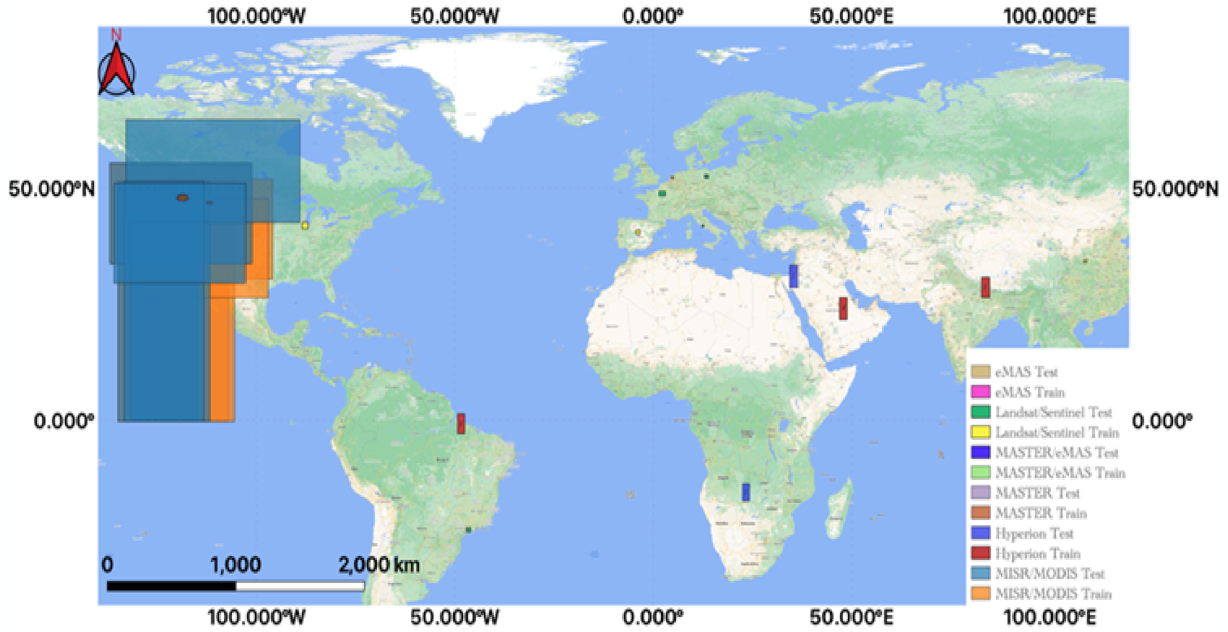
Funding Acknowledgement: Research reported in this presentation was supported by the NASA ROSES AIST project - Innovative Geometric Deep Learning Models for Onboard Detection of Anomalous Events

Thank you!

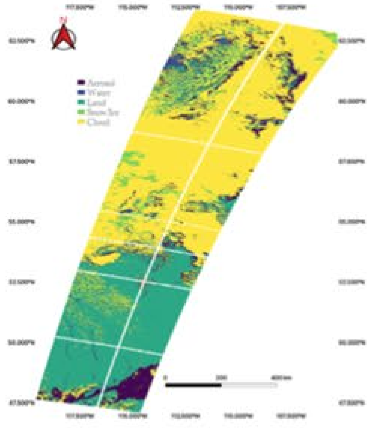
Backup Slides

Large Scale Validation

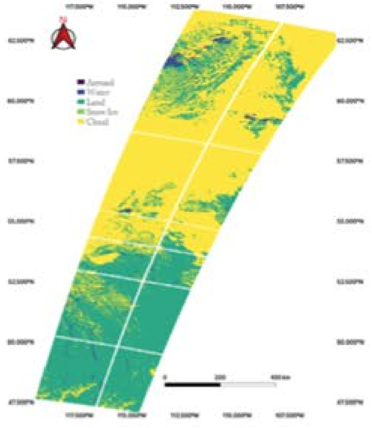
- We know the models can represent structure in data, but how well?
  - Evaluate datasets from different instruments over large areas
    - Compare against ground truth labels
  - Measure inside + outside training extent
- Is multi-instrument fusion possible?
  - Is there value added?
  - Is this a resource-hungry venture
- Whats the performance within large coarse-scale problems vs. finer-scale problems



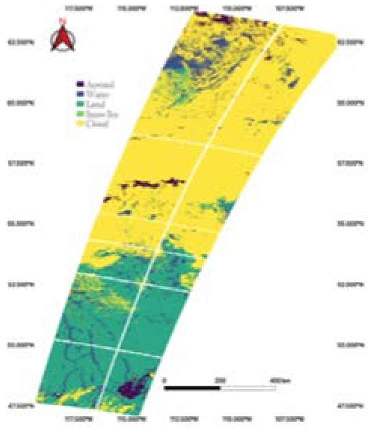
Classifier/Comparison Uncertainty



(a) MISR SVM Truth



(b) Mapped MISR SVM



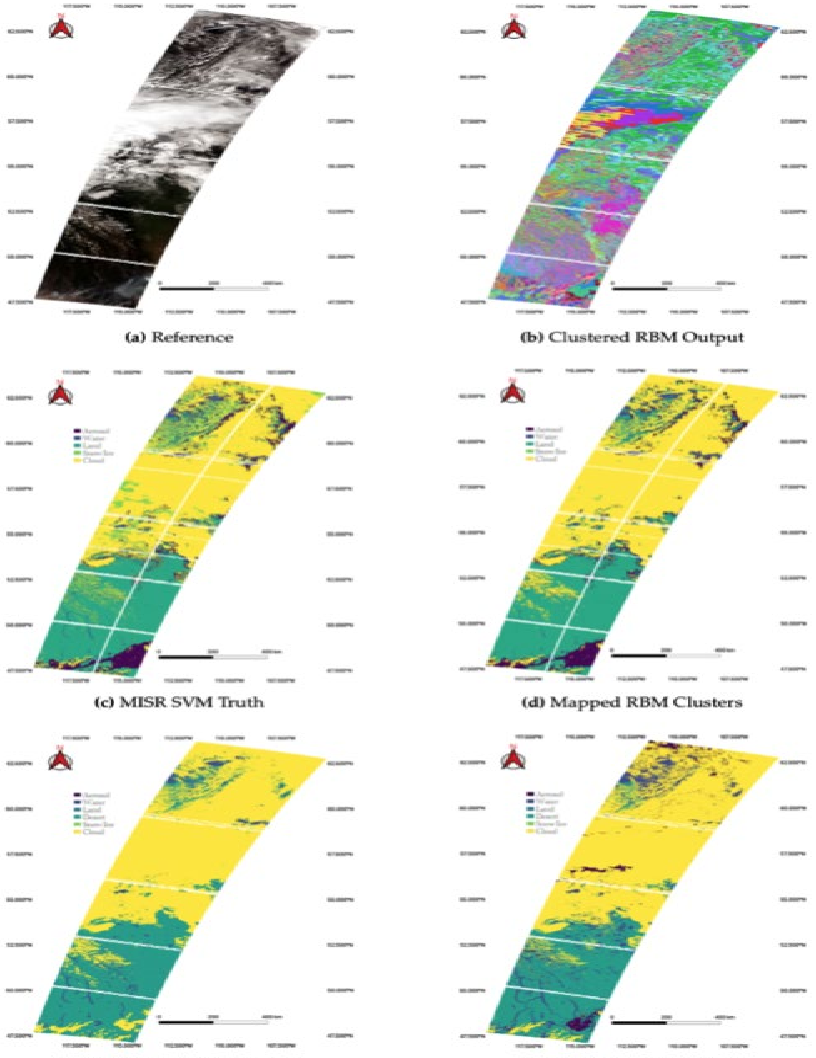
(c) Actual MODIS Cloud Mask

MISR SVM vs. MODIS Cloud Mask

N = 1948222	Aerôsol	Waîer	Lând	Sñow	Clôud
Aerosol	0	487	8358	0	29103
Water	0	157930	98087	0	8591
Land	0	2673	992812	0	16203
Snow	0	0	0	0	0
Cloud	0	8470	76085	0	549423

N = 1948222	Aerôsol	Waîer	Lând	Sñow	Clôud
Aerosol	0.0	1.3	22.0	0.0	76.7
Water	0.0	59.7	37.1	0.0	3.2
Land	0.0	0.3	98.1	0.0	1.6
Snow	0.0	0.0	0.0	0.0	0.0
Cloud	0.0	1.3	12.0	0.0	86.7

Comparisons



(a) Reference (b) Clustered RBM Output

(c) MISR SVM Truth (d) Mapped RBM Clusters

(e) MODIS Cloud Mask Truth (f) Mapped RBM Clusters

MISR + MODIS RBM vs. MODIS Cloud Mask

N = 1845900	Aerosol	Water	Land	Desert	Snow	Cloud
Aerosol	10315	577	195	1422	0	23363
Water	856	180758	5717	63630	0	9641
Land	191	1078	128486	46640	0	3982
Desert	1083	11000	34652	746302	0	15344
Snow	0	0	0	0	0	0
Cloud	2608	16747	291	27752	0	557515

N = 1845900	Aerosol	Water	Land	Desert	Snow	Cloud
Aerosol	28.8	1.6	0.5	4.0	0.0	6.5
Water	0.3	69.4	2.2	24.4	0.0	3.7
Land	0.1	0.6	71.2	25.9	0.0	2.2
Desert	0.1	1.4	4.3	92.3	0.0	1.9
Snow	0.0	0.0	0.0	0.0	0.0	0.0
Cloud	0.4	2.8	0.1	4.6	0.0	92.2

MISR + MODIS RBM vs. MISR SVM

N = 1845900	Aerosol	Water	Land	Snow	Cloud
Aerosol	61248	582	14732	11	6094
Water	475	151643	9477	1	3502
Land	8888	3908	1001212	30	24437
Snow	480	18	2443	1529	26375
Cloud	6268	4729	28211	622	488916

N = 1845900	Aerosol	Water	Land	Snow	Cloud
Aerosol	74.1	0.7	17.8	0.0	7.4
Water	0.3	91.9	5.7	0.0	2.1
Land	0.9	0.4	96.4	0.0	2.4
Snow	1.6	0.0	7.9	5.0	85.5
Cloud	1.2	0.9	5.3	0.1	92.5