Duke UNIVERSITY

Context-Dependent Feature Selection for Classification of Simulated Ground-Penetrating Radar Data

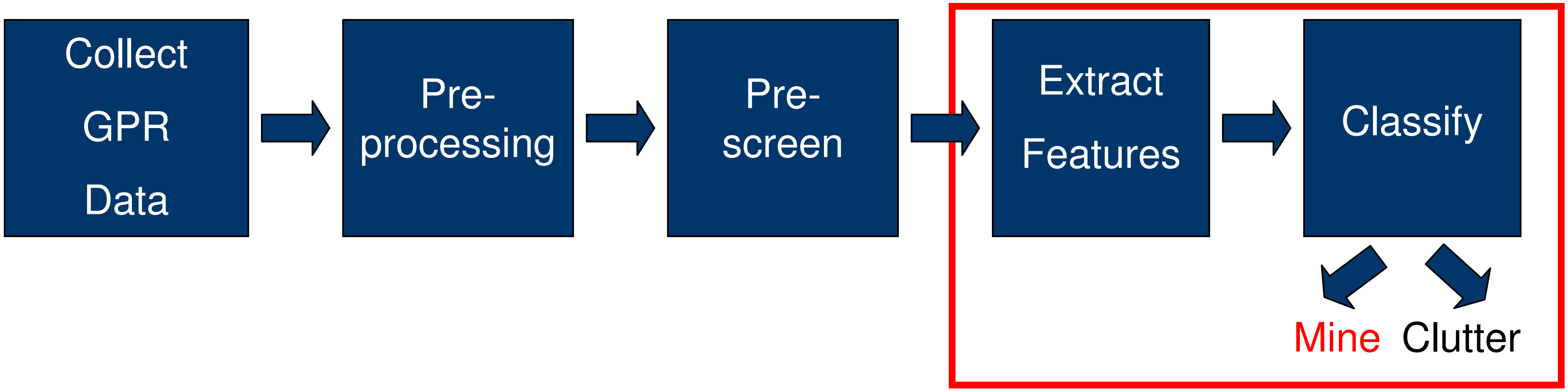
Christopher Ratto, Peter Torrione, Leslie Collins

NISS QMDNS Meeting

05.21.2008

SSPACISS

Processing GPR Data



- Preprocessing and prescreening algorithms remove noise and isolate anomalies ("alarms") in GPR data
- Feature-based classification schemes are then used to determine if alarms are caused by landmines or non-mine "clutter" objects

Environmental Caveat for GPR

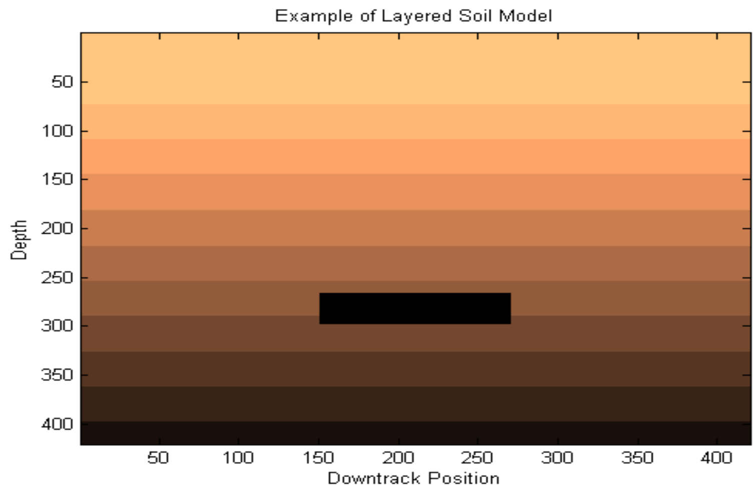
- GPR measures reflections of an electromagnetic pulse caused by changes in subsurface electrical properties (permittivity and permeability)
- The performance of GPR classification algorithms is highly dependent on the environment from which data was collected
  - Moisture changes the dielectric contrast between target and ground (Lensen, et al., 2001) (Miller, et al., 2002)
  - Surface roughness causes random scattering and adds noise to GPR data (Rappaport, 2004)
- A possible solution is context-dependent feature selection
  - Find the best features for classifying GPR signatures collected in a particular environment
  - Find robust features that allow for good classifier performance regardless of soil type, moisture, or roughness

Outline of Experiment

1) Consider one of two environmental scenarios
   a) Soil moisture
   b) Surface Roughness
2) Simulate GPR signatures of mine and clutter objects occurring within that scenario
3) Extract features from simulated data
4) Select the best features for classification
5) Evaluate classifier performance on selected features

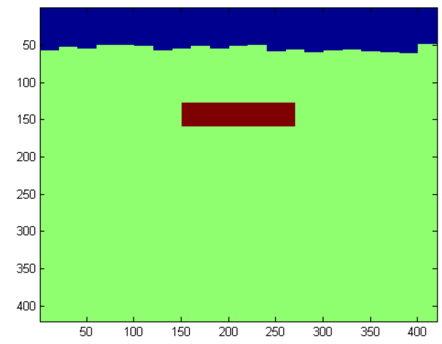
Scenario 1: Soil Moisture

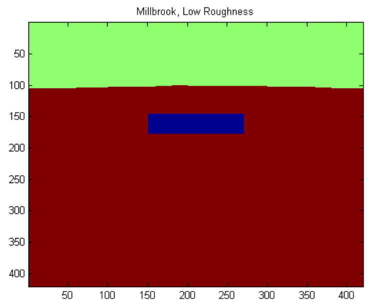
- Hendrickx, et al (1999) previously investigated hydrology of soils near landmines in Bosnia and Kuwait
  - Simulated measurements of soil water content
- Can estimate the relative permittivity from volumetric water content via a calibration curve (Topp, et al., 1980)
- 10-layer soil model used to investigate moisture effects in 4 simulated data sets
  - Kuwait Loam
  - Kuwait Loamy Sand
  - Bosnia Loam
  - Bosnia Loamy Sand

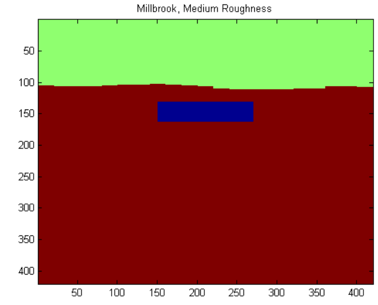


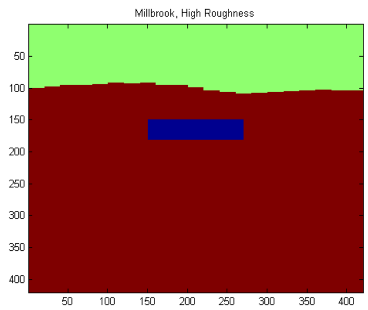
Scenario 2: Surface Roughness

- Rough surface usually realized by white Gaussian process (Tsihrintzis, et al., 1998), (Rappaport and El-Shenawee, 2000)
  - Sharp corners not found in nature
- Instead, we model surface roughness with an autoregressive (AR) model
  - Train 4-th order AR model on real GPR data from 3 different test sites
    - Yuma Proving Ground (Yuma, AZ)
    - Millbrook Proving Ground (UK)
    - Ft. Leonard Wood (St. Louis, MO)
  - "Degree" of roughness is determined by increasing the gain of the AR model's power spectrum
    - Low, medium, and high



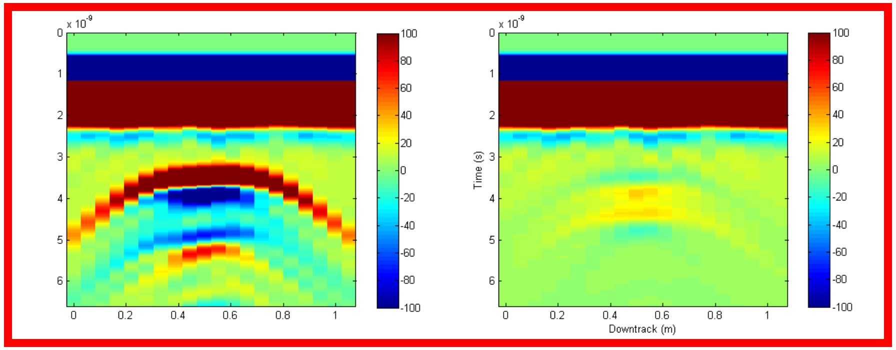


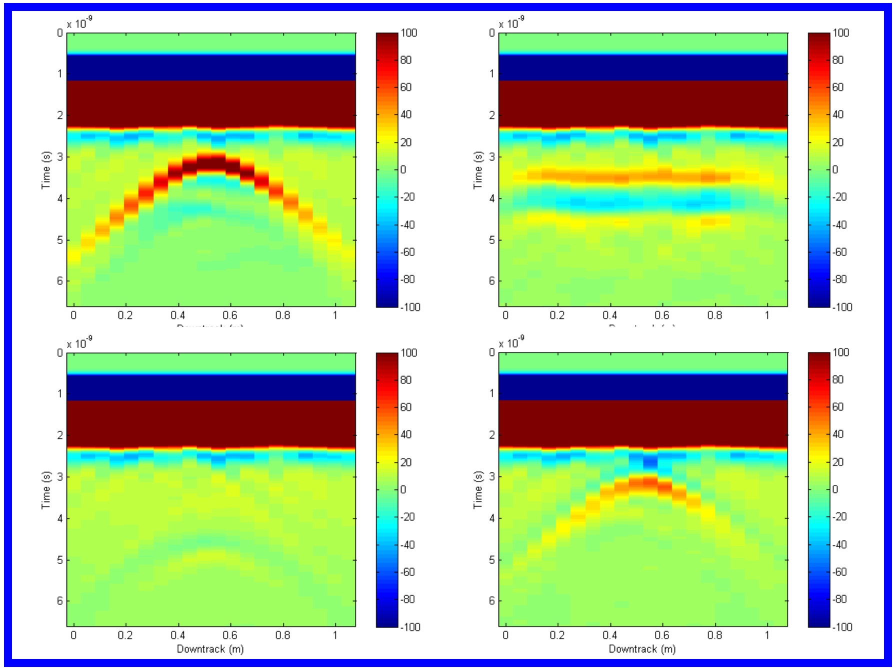




Simulated GPR Signatures

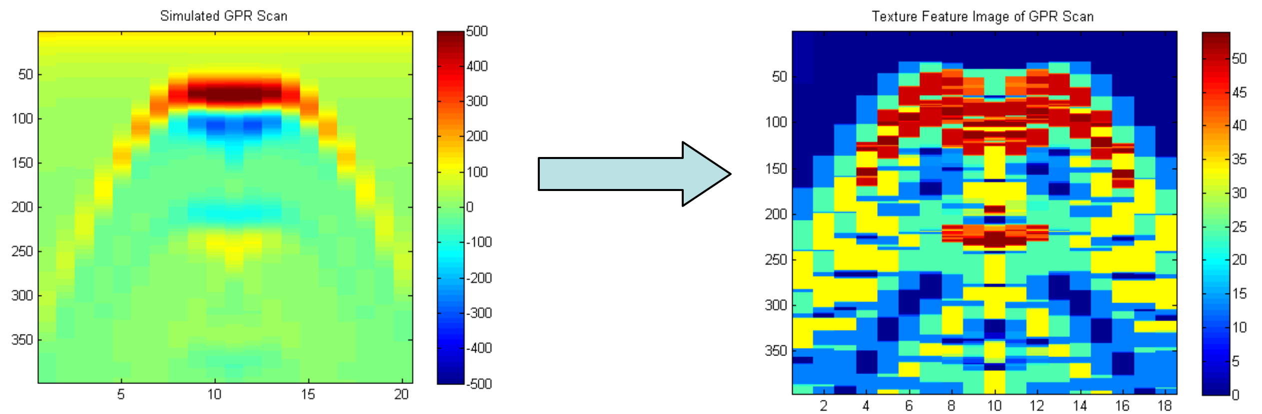
- FDTD method used for calculating electromagnetic fields
- Targets
  - Metal and plastic antitank landmines (30 x 10 cm)
- Clutter
  - Rebar, concrete slab, triangular rock, circular void
- Each object is buried at 10 different depths
  - 60 observations per data set
- Preprocessing
  - Removal of antenna coupling effects
  - Down-sample images by 2 to expedite feature extraction





Feature Extraction

- Texture Feature Coding Method (Horng, 2003)
  - Previously applied to GPR classification (Torrione and Collins, 2007)
  - Transforms a grayscale image into a “feature image”



  - Co-occurrence matrix used to estimate the probability distribution of texture feature numbers within an image
  - 13 statistical measurements of the co-occurrence matrix used as features for classification
  - Features are normalized to zero mean, unit variance

Feature Selection

- Wrapper Method
  - Hybrid forward/backward search
  - Search for 6 features forward, 3 features backward
  - Provides 3 features in relative good context
- Filter Method
  - Mutual information between features and class labels

$$I(i) = \sum_{x_i} \sum_{y} P(X = x_i, Y = y) \log \frac{P(X = x_i, Y = y)}{P(X = x_i) P(Y = y)}$$

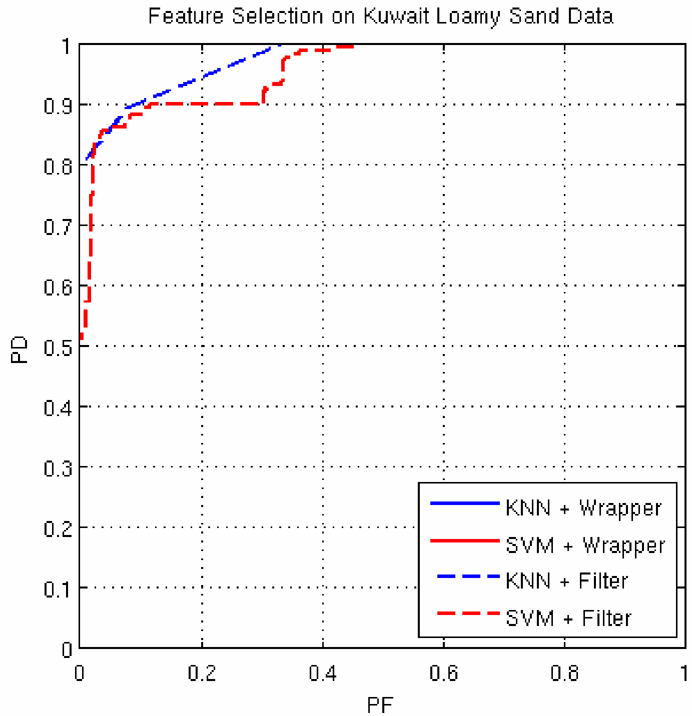
  - Probability distributions estimated with histograms
    - Becomes correlation coefficient if X and Y are Gaussian r.v.'s
  - Choose 3 features with highest *I*

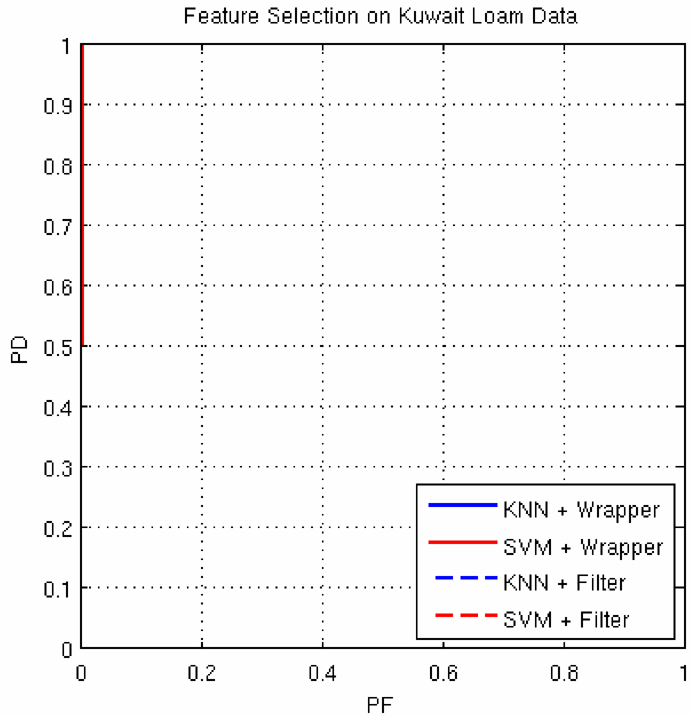
Classifiers

- We desire nonlinear decision boundaries
  - Presence of strong and weak GPR scatterers in both classes
  - Overlap in feature space caused by noise in rough surface data
- Classifiers used
  - K-nearest neighbor (K = 7)
  - Support Vector Machine (Gaussian kernel)
- The wrapper method selects features that maximizes AUC of each classifier
- The filter method chooses the same features for each classifier

Results of Feature Selection for Soil Moisture Scenario

Results of Feature Selection for Soil Moisture Kuwait Loamy Sand and Loam

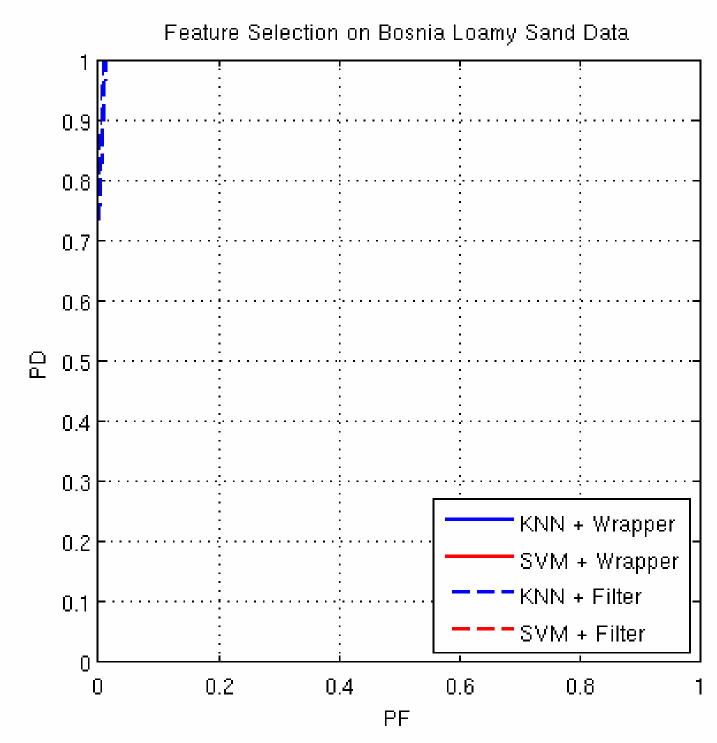


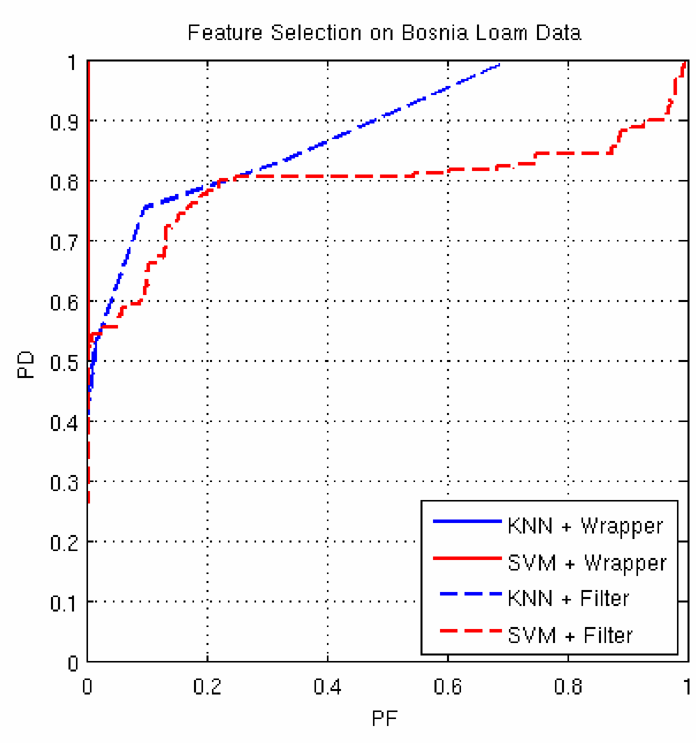


SVM	3	11	10
KNN	3	2	10
FILTER	9	2	10

SVM	10	9	13
KNN	10	13	6
FILTER	9	10	2

Results of Feature Selection for Soil Moisture Bosnia Loamy Sand and Loam



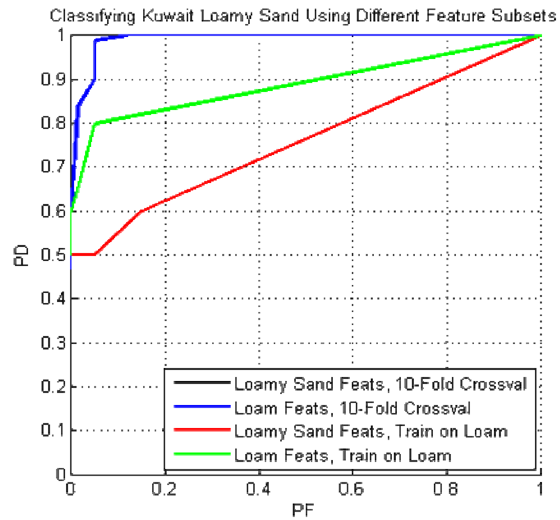


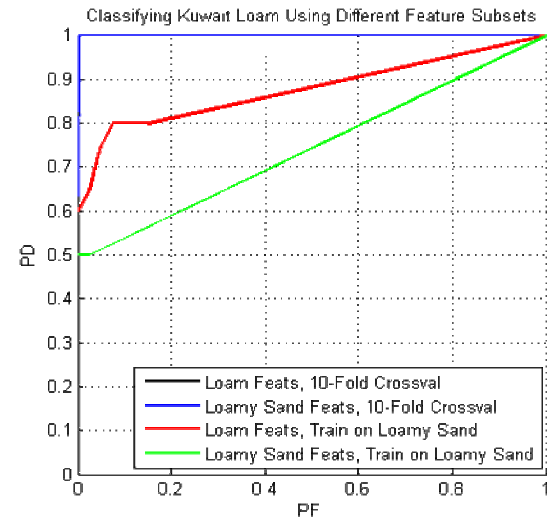
SVM	1	9	6
KNN	2	1	10
FILTER	2	9	11

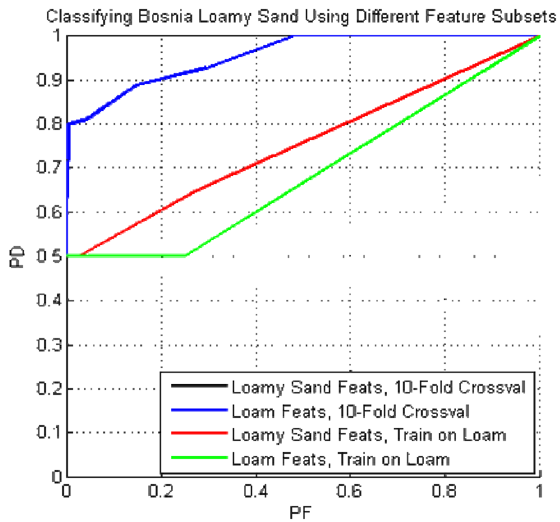
SVM	10	13	12
KNN	10	9	6
FILTER	9	4	13

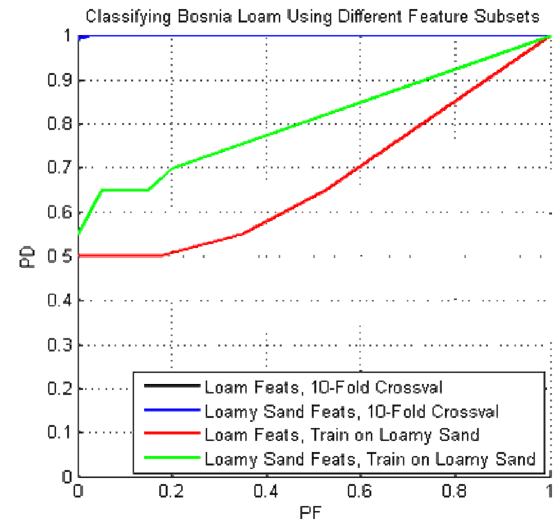
Results of Feature Selection for Soil Moisture Train/Test on Different Soils and Feature Subsets

- Compare KNN performance trained on different soil types using different feature subsets
  - Best features, 10-fold cross-val
  - Other soil type's features, 10-fold cross-val
  - Best features, train on other soil type
  - Other soil type's features, train on other soil type





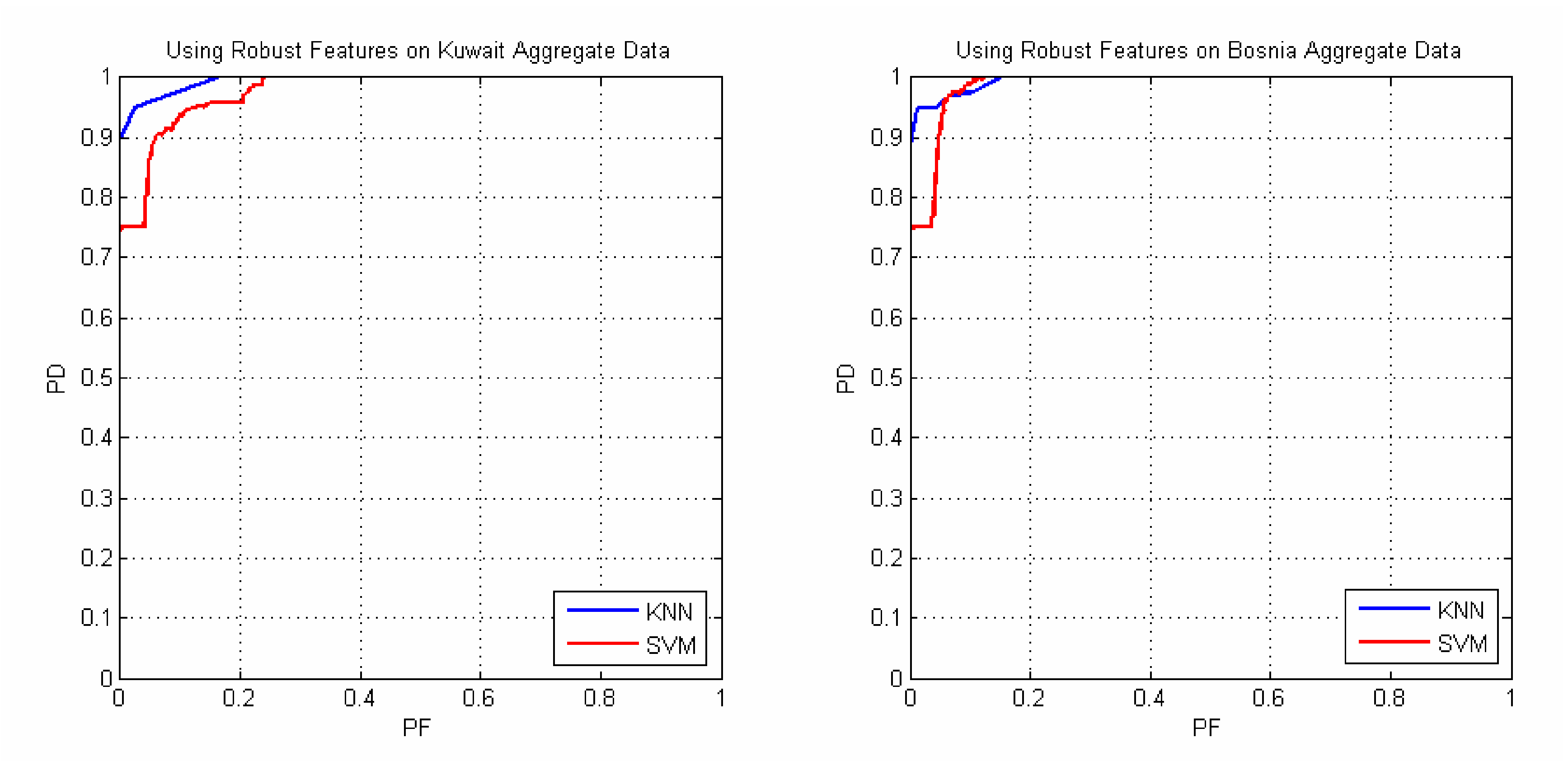




Results of Feature Selection for Soil Moisture Discussion

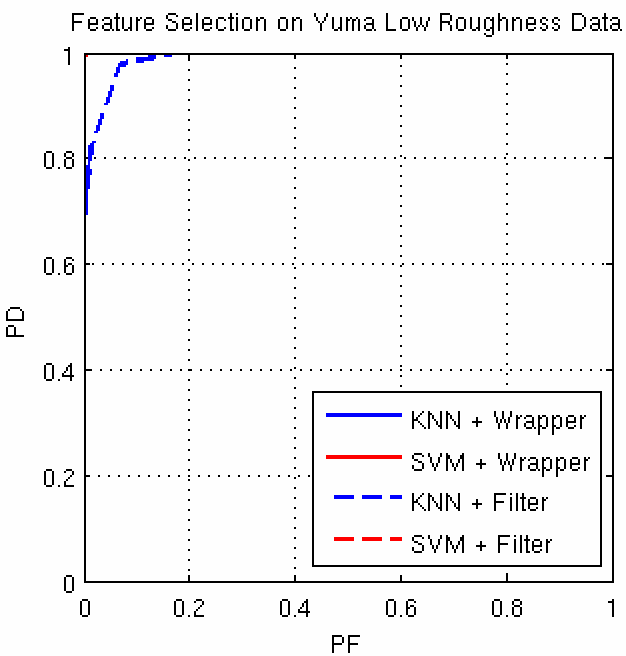
- The wrapper method chooses features for better classification than the filter method
- Best classification achieved by training/testing on same soil type, using best features
  - Using best features for other soil types weakens classifier performance
- Some features are selected regardless of soil type
  - 10 "Energy Distribution 2"
  - 9 "Energy Distribution 1"
- Some features are selected more often for a particular soil type
  - 2 "Code Variance" (loamy sand)
  - 13 "Code Similarity" (loam)

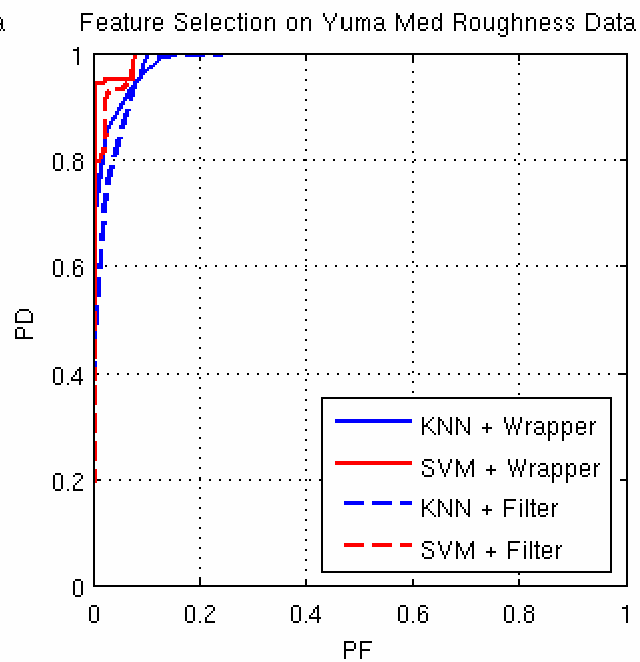
Results of Feature Selection for Soil Moisture
Classifying Aggregate Data Using Features 10-9-2

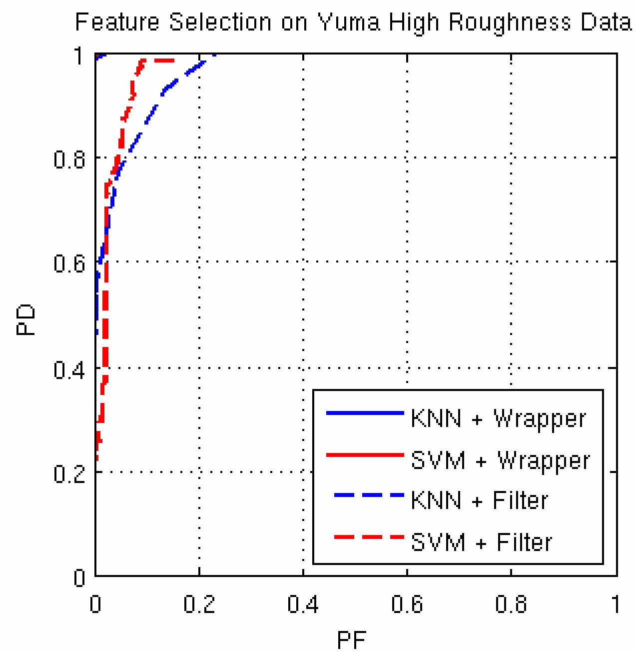


Results of Feature Selection for Surface Roughness Scenario

Results of Feat. Selection for Surface Roughness Yuma Proving Ground – Separated by Roughness





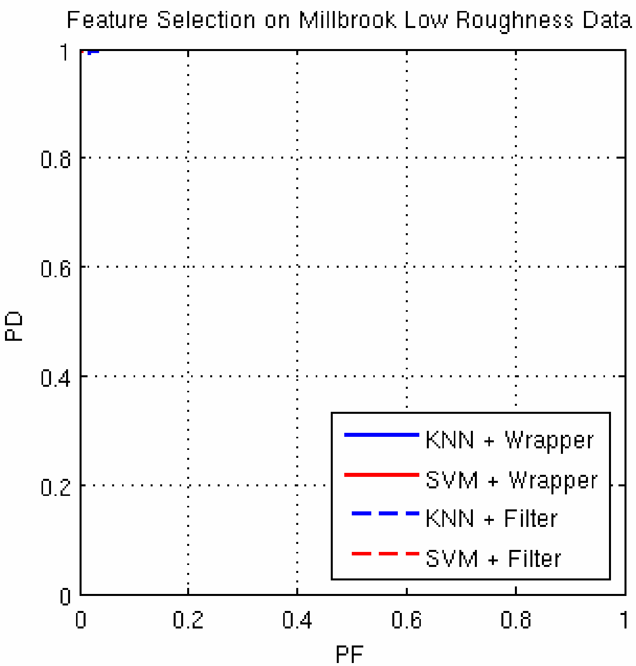


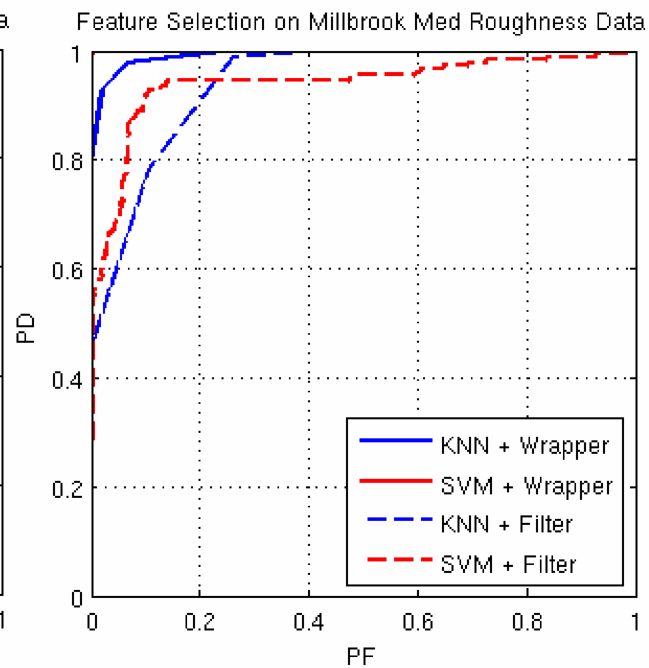
KNN	11	9	7
SVM	12	11	13
FILTER	12	7	8

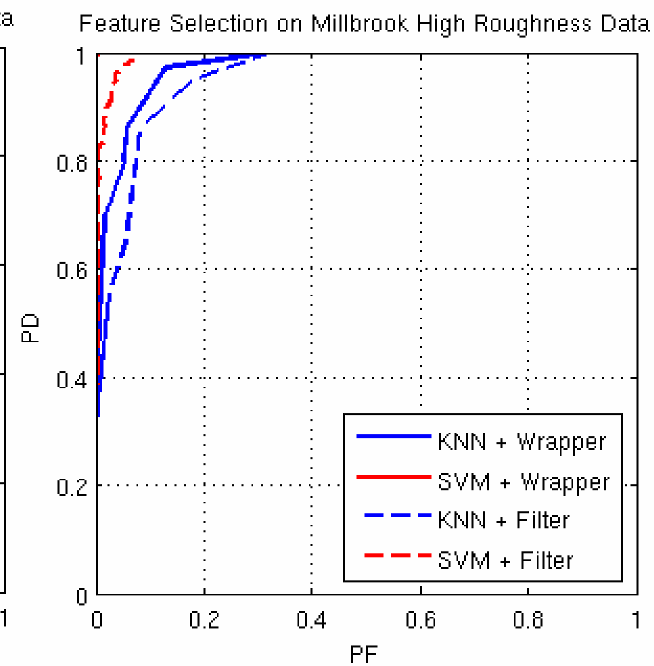
KNN	12	9	1
SVM	12	6	1
FILTER	12	2	7

KNN	3	6	9
SVM	3	6	10
FILTER	12	11	13

Results of Feat. Selection for Surface Roughness Millbrook – Separated by Roughness





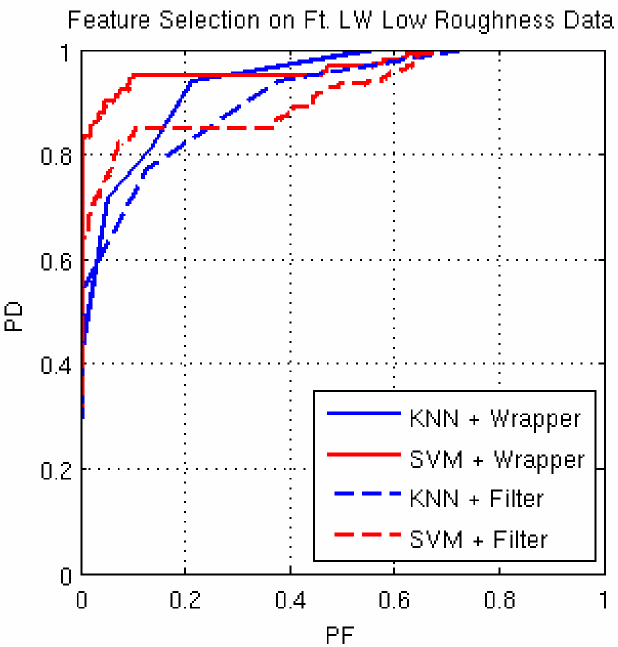


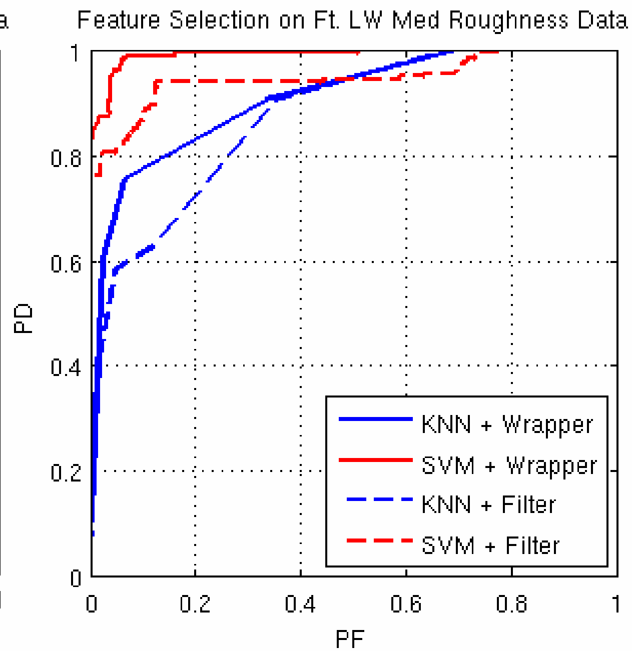
KNN	2	13	12
SVM	2	13	12
FILTER	5	11	13

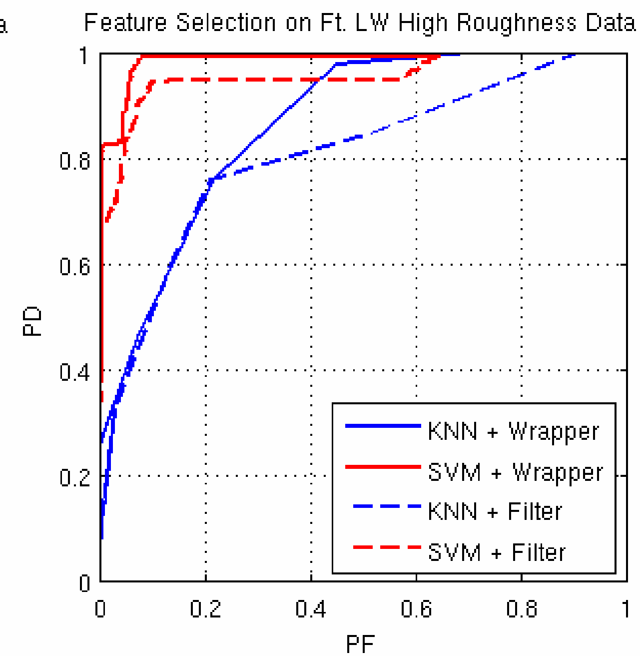
KNN	11	4	2
SVM	11	13	12
FILTER	4	13	12

KNN	11	3	1
SVM	11	13	6
FILTER	3	11	12

Results of Feat. Selection for Surface Roughness Ft. Leonard Wood – Separated by Roughness







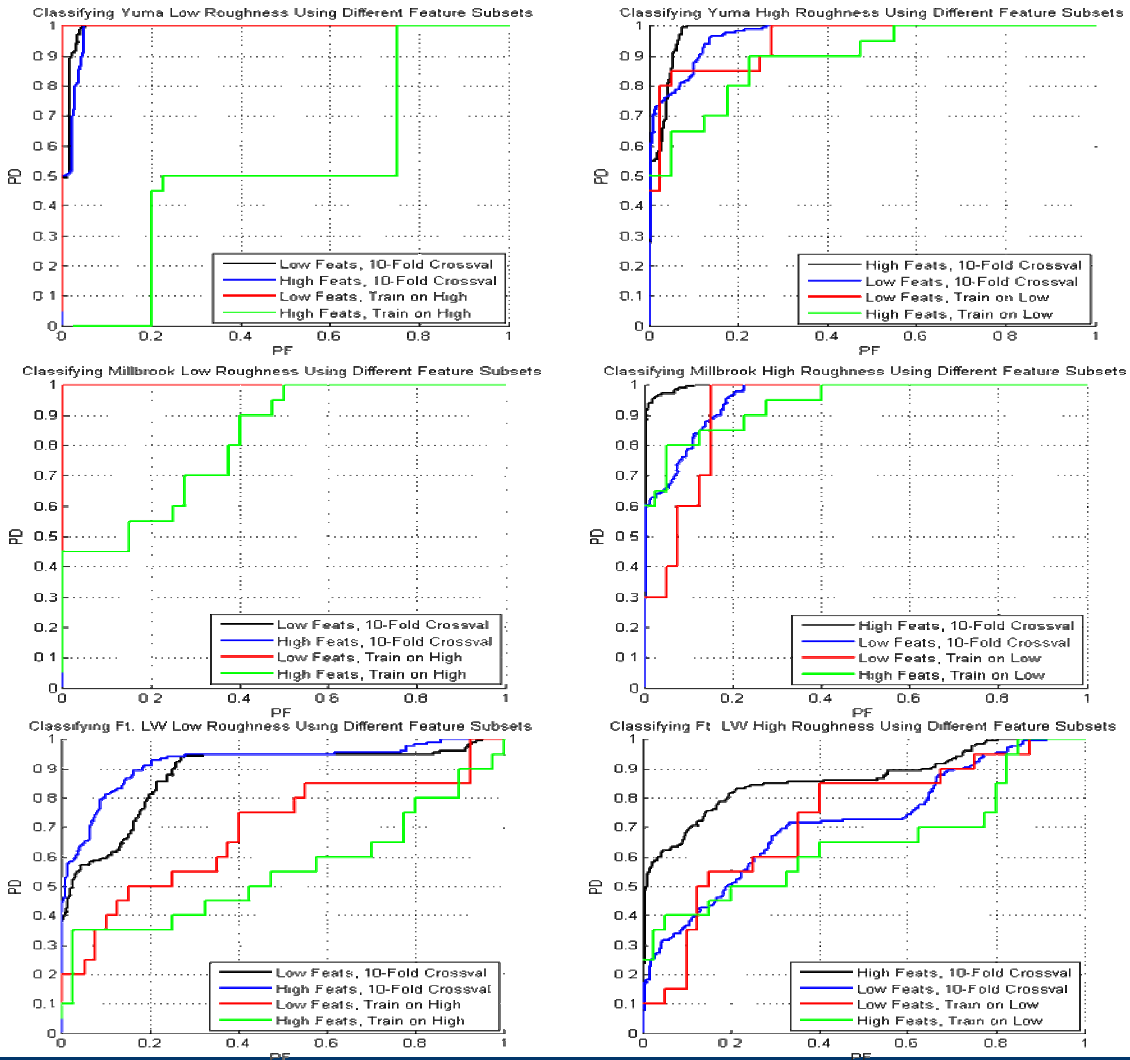
KNN	5	2	11
SVM	1	4	6
FILTER	9	13	4

KNN	6	11	9
SVM	5	2	10
FILTER	1	11	3

KNN	13	2	4
SVM	11	10	13
FILTER	11	1	9

Results of Feat. Selection for Surface Roughness Train/Test on Different Data and Feature Subsets

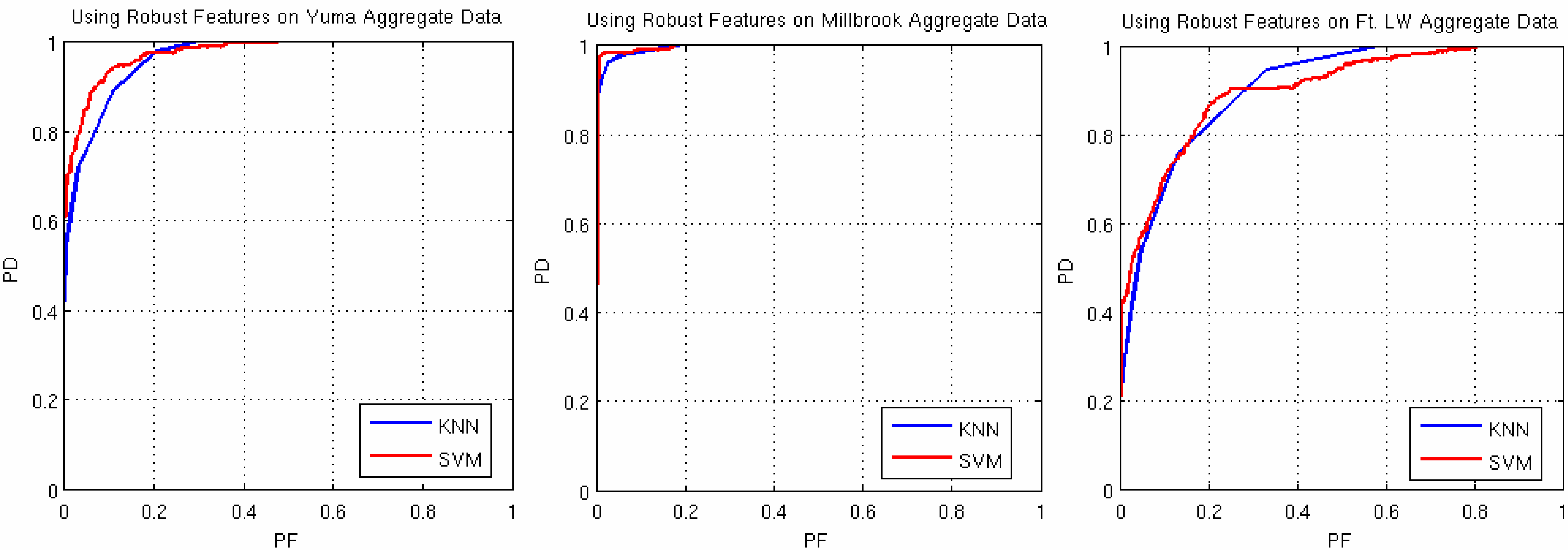
- Compare KNN performance trained on different roughness using different feature subsets
  - Best features, 10-fold cross-val
  - Other roughness features, 10-fold cross-val
  - Best features, train on other roughness
  - Other roughness features, train on other roughness



Results of Feat. Selection for Surface Roughness Discussion

- SVM maintains high AUC despite increasing roughness
  - Selects different support vectors to maximize margin between classes
- Better to train on data with higher roughness than the test data
  - "Worst-case scenario" approach to drawing decision boundaries
  - Over-compensate for overlapping classes in feature space
- Some features are selected regardless of location
  - 11 – "Energy Distribution 3"
  - 13 – "Code Similarity"
- Some features are features dependent on location, but perform well across roughness
  - 12 – "Homogeneity" (Yuma/Millbrook)

Results of Feat. Selection for Surface Roughness Classify Aggregate Data using Features 11-12-13



- Worse performance on Ft. LW data since feature 12 was never selected for that set
- Should include universally robust features as well as location-specific features to achieve good classifier performance

Conclusions

- Context-dependent feature selection can help maintain a high AUC for classification
  - Feature subsets exist that separate data in certain environments better than in other environments
  - Features exist that are robust to environmental changes
- However, context-dependent feature selection requires knowledge of roughness, soil type, and moisture *a priori*
  - Difficult in fielded scenarios
- Target classification in fielded systems should incorporate information regarding environmental conditions before implementing a decision boundary
  - Motivates context-dependent learning