Accelerating Satellite Image Based Large-Scale Settlement Detection with GPU

Dilip R. Patlolla

Anil M. Cheriyadat

Eddie A. Bright

Jeanette E. Weaver

Oak Ridge National Laboratory
Oak Ridge, TN

GPU Technology Conference 2013

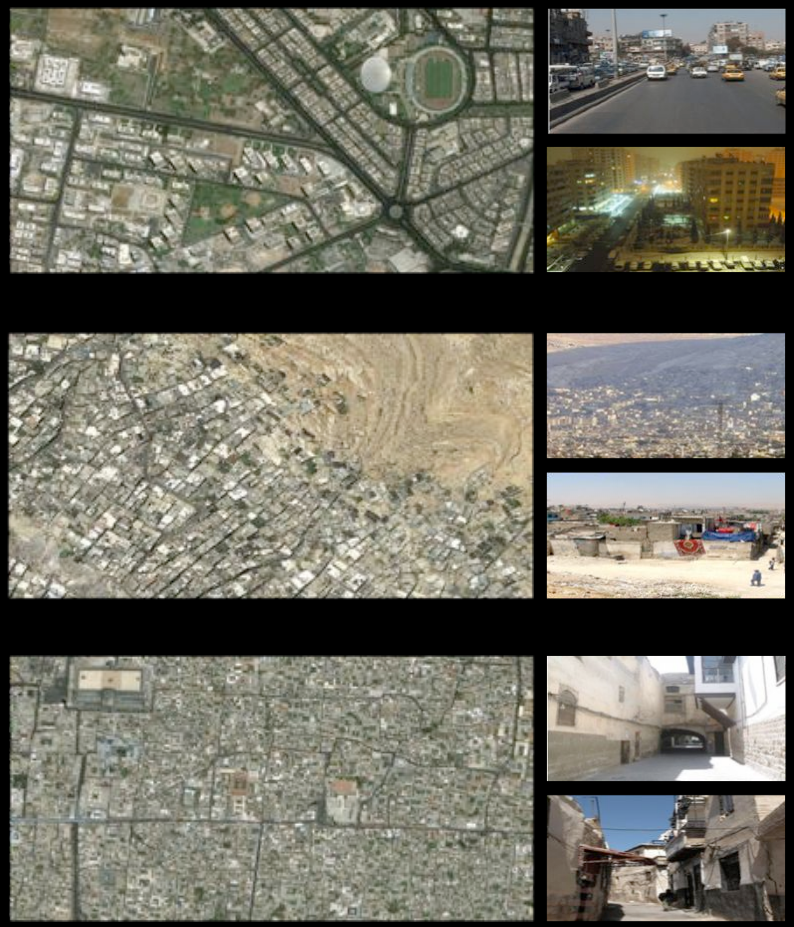
San Jose, CA

Introduction

- Advanced geospatial models built on settlement characteristics such as size, shape, location, growth can reveal social stability and political durability.

- Automated settlement detection process involves extracting, representing, and identifying patterns characterizing settlements in image data.

- Strong demand to understand the settlement characteristics from aerial imagery.

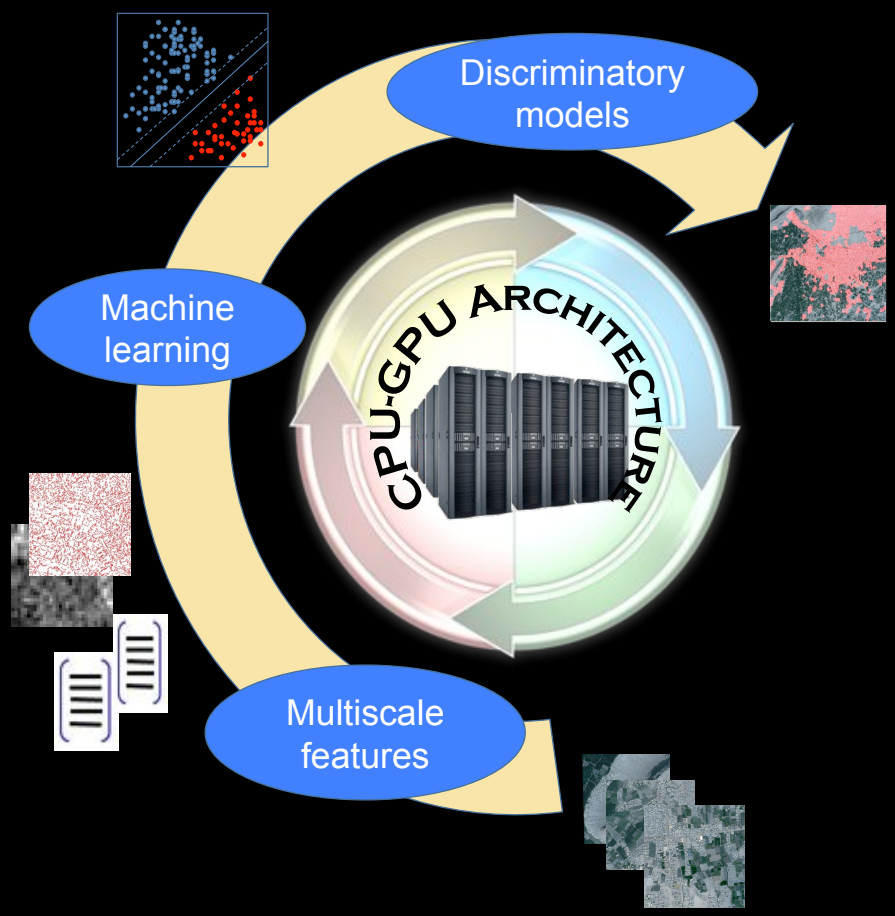


Problem

- Land surface ~150 Million km^2 translates into roughly 600 Trillion pixels to be processed with 0.5 m imagery.
- Pixel to feature descriptor mapping is an expensive, but key step in automated scene analysis.
- Application of settlement detection algorithms on large image databases can easily saturate the processing capabilities of conventional CPUs.

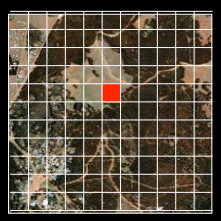
Large Scale Data Analysis

- Parallel feature extraction using Single Instruction Multiple Threads (SIMT) architecture of GPU.
- Efficient GPU implementation of multiscale features.
- Exploitation of multi-GPU architecture for large-scale image analysis.
- Reduce redundant data transfers and computations.

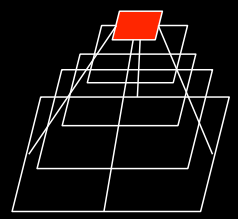


Settlement Mapping Process

Geographic Information Science and Technology



Divide Image into pixel blocks



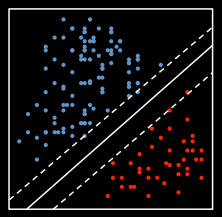
Compute Multi-scale features for each pixel block

HOG Statistics
(5 features x 5 scales)

Pixel Block Intensity
Mean and Variance
(2 features x 5 scales)

GLCM Contrast
(1 feature x 1 scale)

Each pixel block mapped to 36-dimensional vector



Apply linear SVM Model

Output Settlement Image

Features are computed for each scale :

$$d^s = \{\ h^s_{i,j}, c^s_{i,j}, g_{i,j}\ ; s = 0,1,..N\}$$

N is the total number of scales.
i,j represent the index of the pixel blocks

For each pixel-block at every scale, we reduce the high-dimensional feature vectors to a scalar value by projecting the feature vector onto the SVM decision plane,

$$z^s = (W^T)^s\ d^s$$

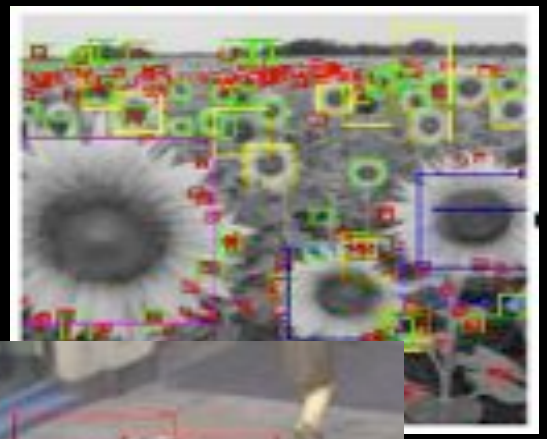
SVM decision function is given by $f(z) = \sum_{s=0}^{N} z^s + b$

At each pixel-block ***(i,j)*** the SVM decision function simply becomes $f(z_{i,j}) > 0$ for settlement.

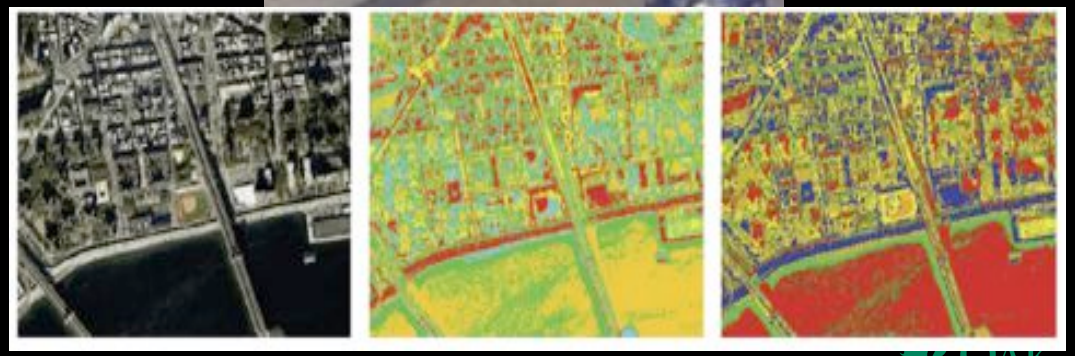
W and *b* are SVM model parameters

GPU Based Vision Systems

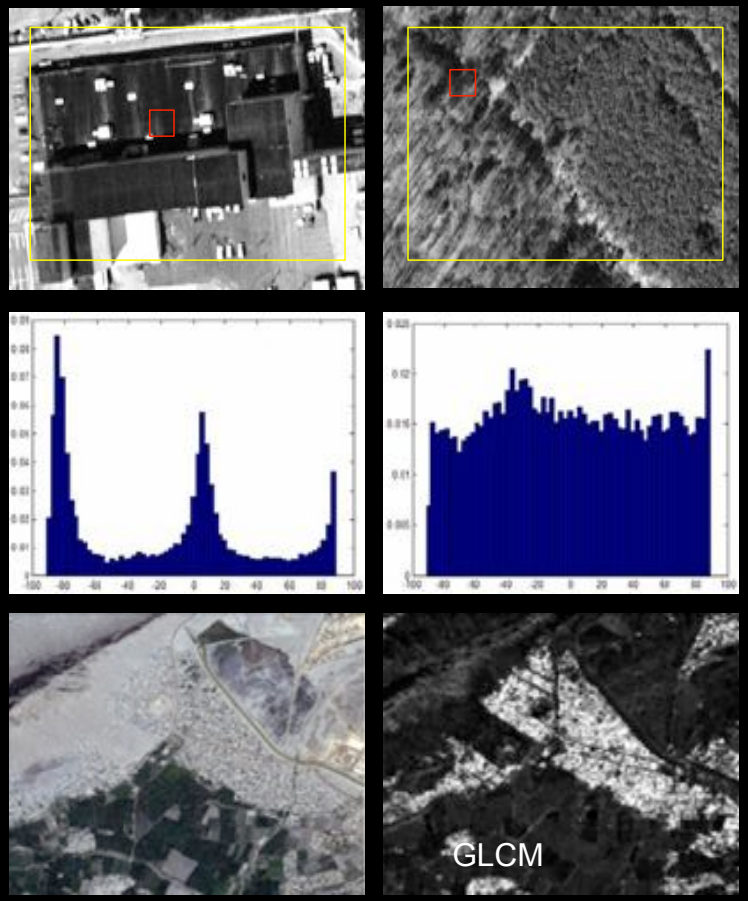
- Cornelius et al. - Scale invariant feature descriptor generation and matching on a GPU
- Bernabe et al. - satellite image classification on GPU, but limited to pre processing filtering, post processing
- GPU based HOG for Pedestrian Detection
- Parallel GLCM by shahbahrami et al. but targeted to a Cell processor



Significant speedups have been achieved in all cases



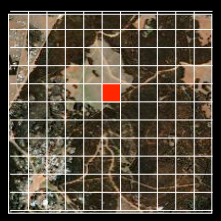
Computer Vision Challenges



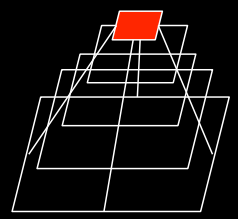
- With high-resolution imagery, pixel features are not discriminatory.
- Multi-scale features provide accurate settlement characterization.
- We aggregate three low level features from 5 spatial scales.
  - Histogram of Gradient Orientations (HOG) : *h*
  - Local pixel intensity statistics (Spectral) : *c*
  - Gray-Level Co-Occurrence Matrix (GLCM) : *g*
- A linear support vector machine (SVM) classifier is used to map the feature vector (*z*) to one of the binary classes (settlement and non-settlement).

Settlement Mapping Process

Geographic Information Science and Technology



Divide Image into pixel blocks



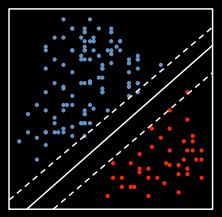
Compute Multi-scale features for each pixel block

HOG Statistics
(5 features x 5 scales)

Pixel Block Intensity
Mean and Variance
(2 features x 5 scales)

GLCM Contrast
(1 feature x 1 scale)

Each pixel block mapped to 36-dimensional vector



Apply linear SVM Model

Output Settlement Image

Features are computed for each scale :

$$d^s = \{\ h^s_{i,j}, c^s_{i,j}, g_{i,j}\ ; s = 0,1,..N\}$$

N is the total number of scales.
i,j represent the index of the pixel blocks

For each pixel-block at every scale, we reduce the high-dimensional feature vectors to a scalar value by projecting the feature vector onto the SVM decision plane,

$$z^s = (W^T)^s\ d^s$$

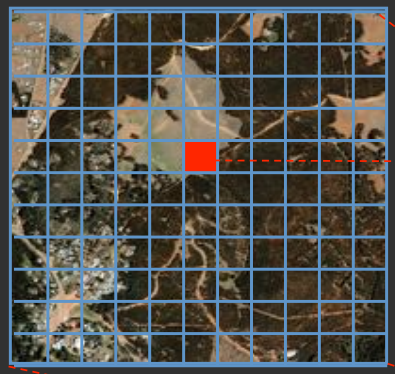
SVM decision function is given by $f(z) = \sum_{s=0}^{N} z^s + b$

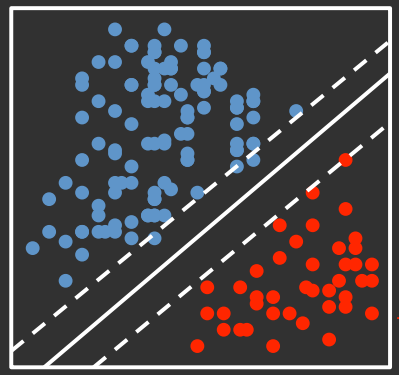
At each pixel-block ***(i,j)*** the SVM decision function simply becomes $f(z_{i,j}) > 0$ for settlement.

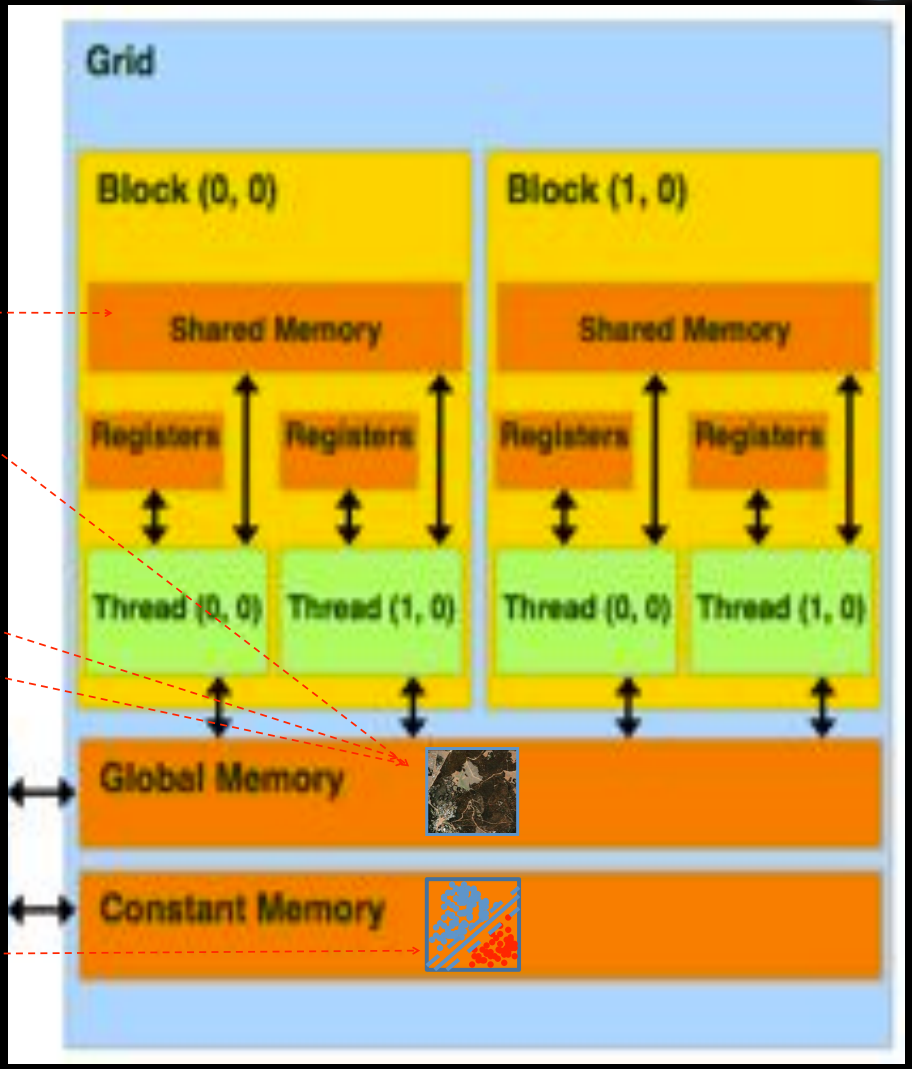
W and *b* are SVM model parameters

DATA LOCALITY

HOST (CPU)



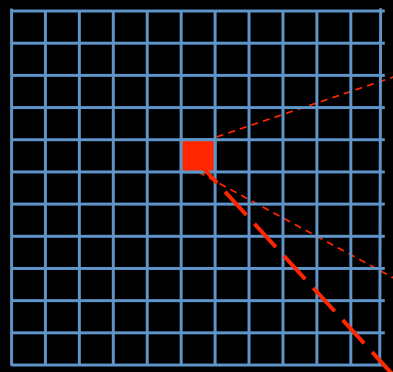


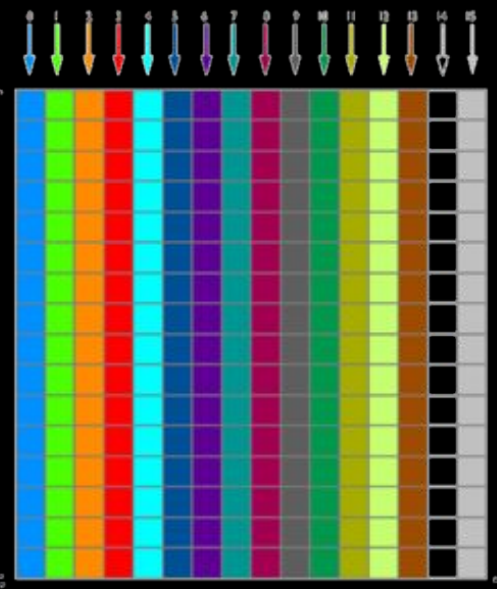


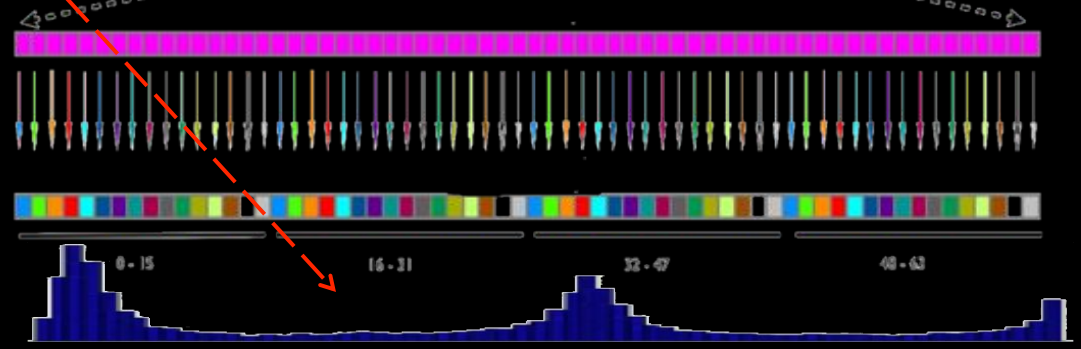
HISTOGRAM OF ORIENTATIONS

Geographic Information Science and Technology

Gradient data



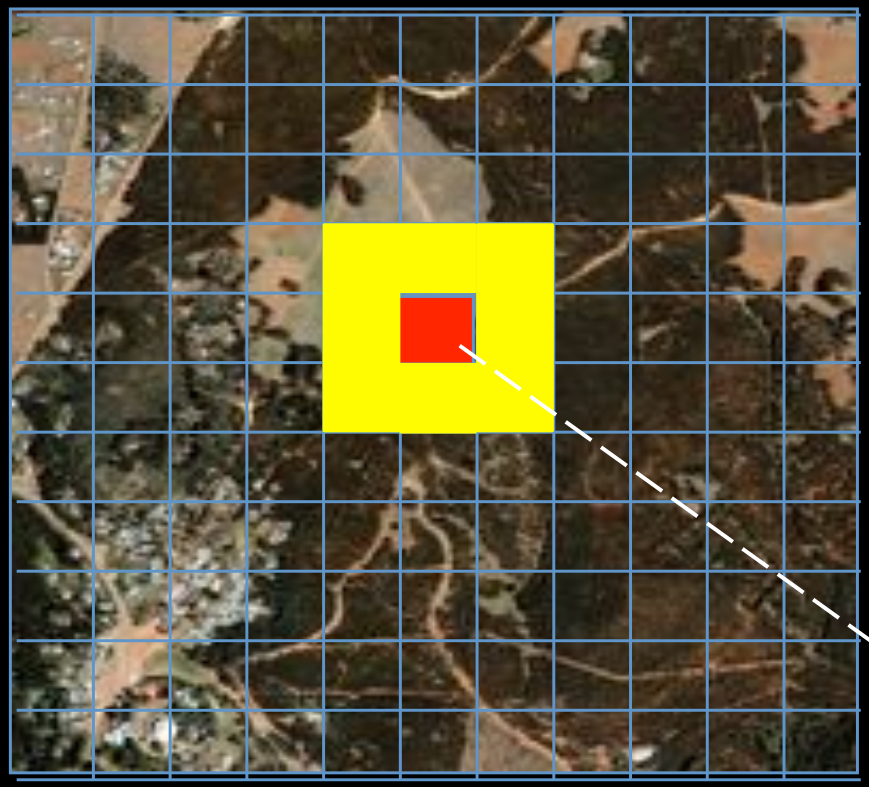


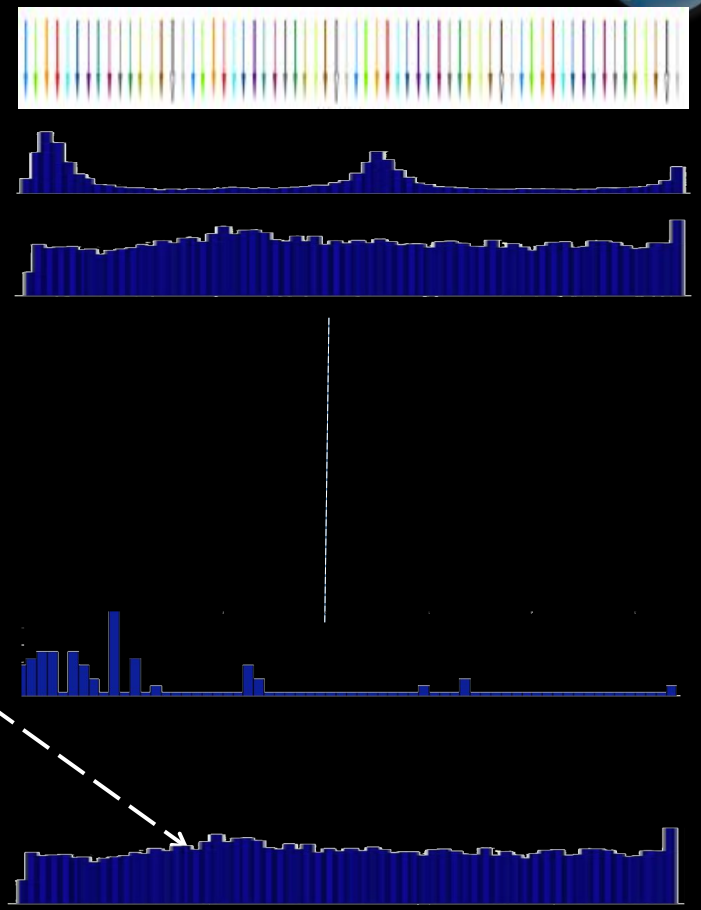


M is the pixel Block size.

```
for j = 1 to M do
    index = (j, baseidx, M)
    OA_bin = OA[index]
    HIST_shared[OA_bin]+=GR[index]
end for
Sync
for k = 1 to Size(HIST)/M do
    stride = M
    Update HIST_global[baseidx; k; stride]
end for
Sync
```

MULTISCALE HOG

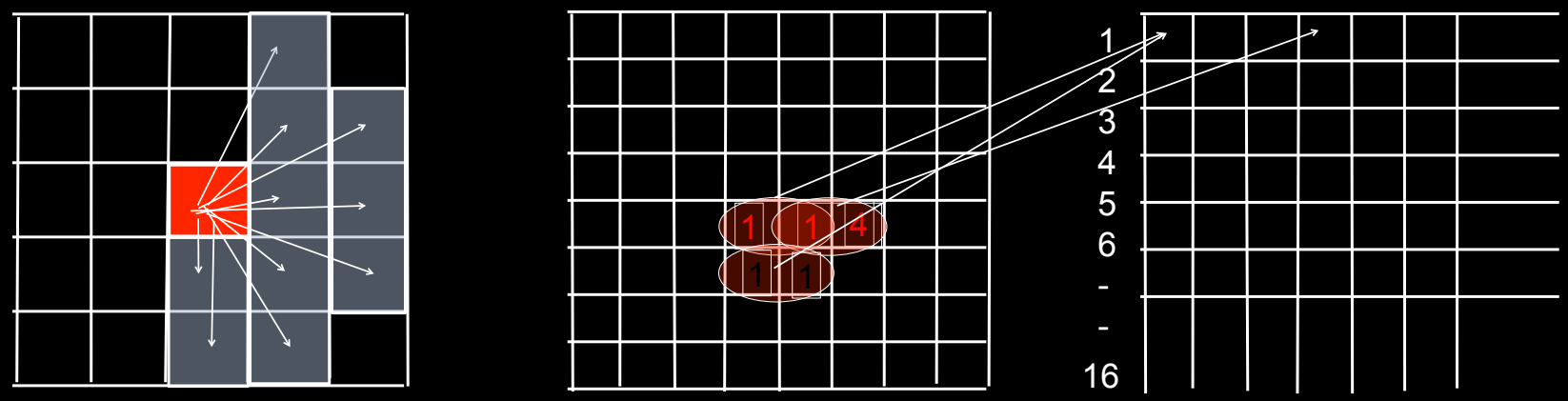




Histogram for scale 2

- 64 threads compute the mean of the base scale histograms
- No Memory Conflicts
- Reduced redundancy

GLCM



N threads working
NxN quantized pixel block
In Shared Memory

$$con(c) = \sum_{m=1}^{Q}\sum_{n=1}^{Q}(m-n)^2\mathbf{P}(c)_{m,n}$$

$$\mathbf{g}^0_{i,j} = con(k)(\mathbf{c}(1)^0_{i,j})^2$$

P is the normalized co-occurrence matrix
Q represents the number of quantized gray levels.
$c(1)^0_{i;j}$ represents the mean pixel intensity for the pixel-block at (i; j) and k = arg $\min_k$ con(c).

Hardware

Xeon + Tesla 4U Server

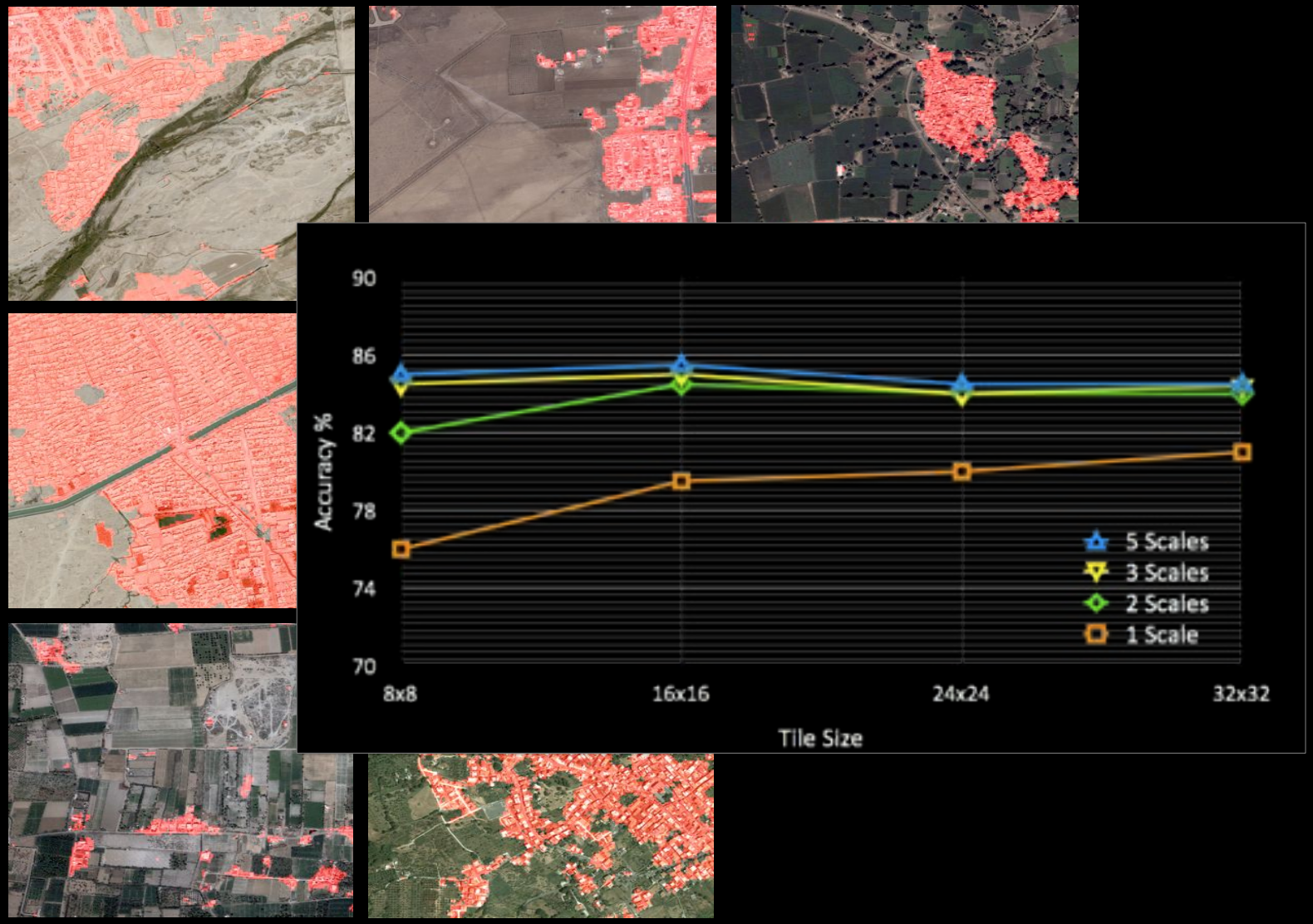
- CPU
  - 2 Intel Xeon E5620 Westmere
    - 2.4 GHz Quad Core
    - 48 GB DDR3 Memory
- GPU
  - 4 Nvidia Tesla C2075
    - 448 CUDA cores
    - 6 GB GDDR5 Memory

Validation

- Settlement Detection Model learned from a dataset consisting of 60 scenes
- Each Scene
- - represents settlement structures from different parts of the globe
- - Size : 2048 x 2048 at 0.6 m spatial resolution
- Tested on a dataset consisting of 55 scenes with 86% accuracy
- Results validated with CPU implementation.

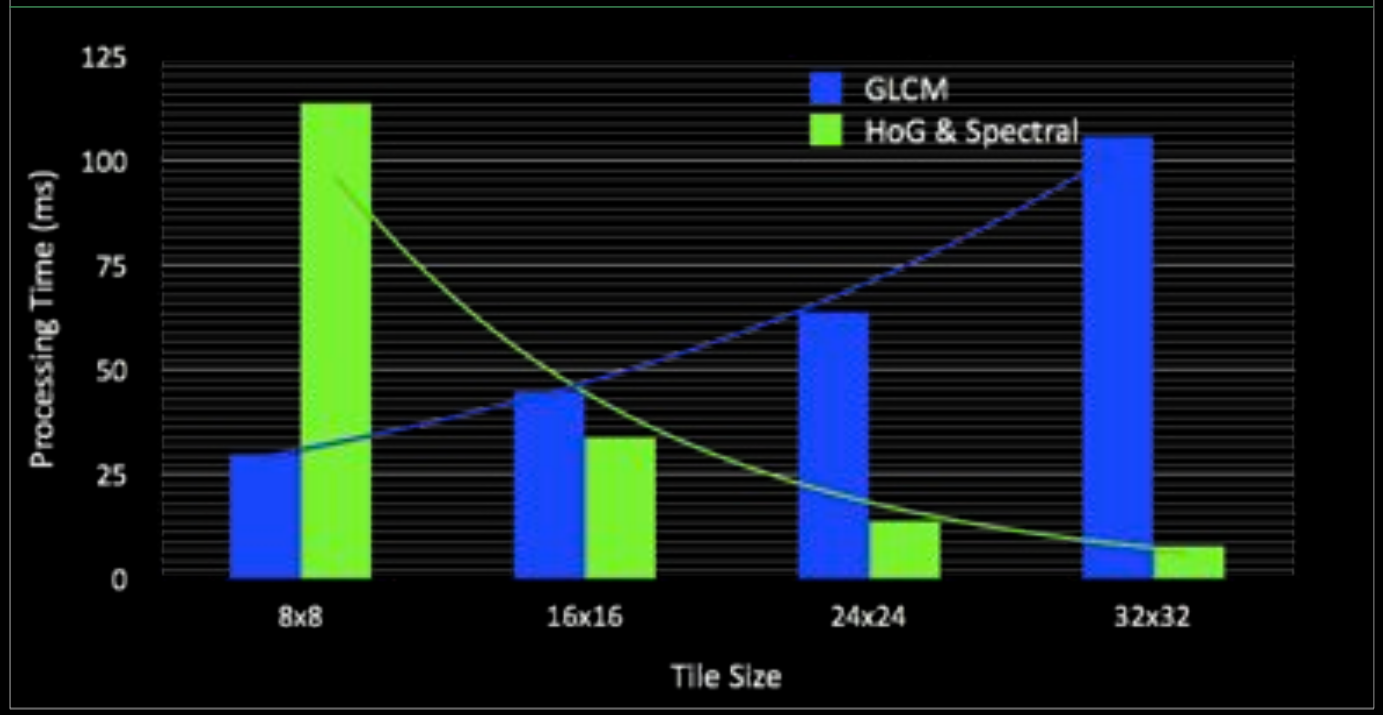
Few scenes

ACCURACY

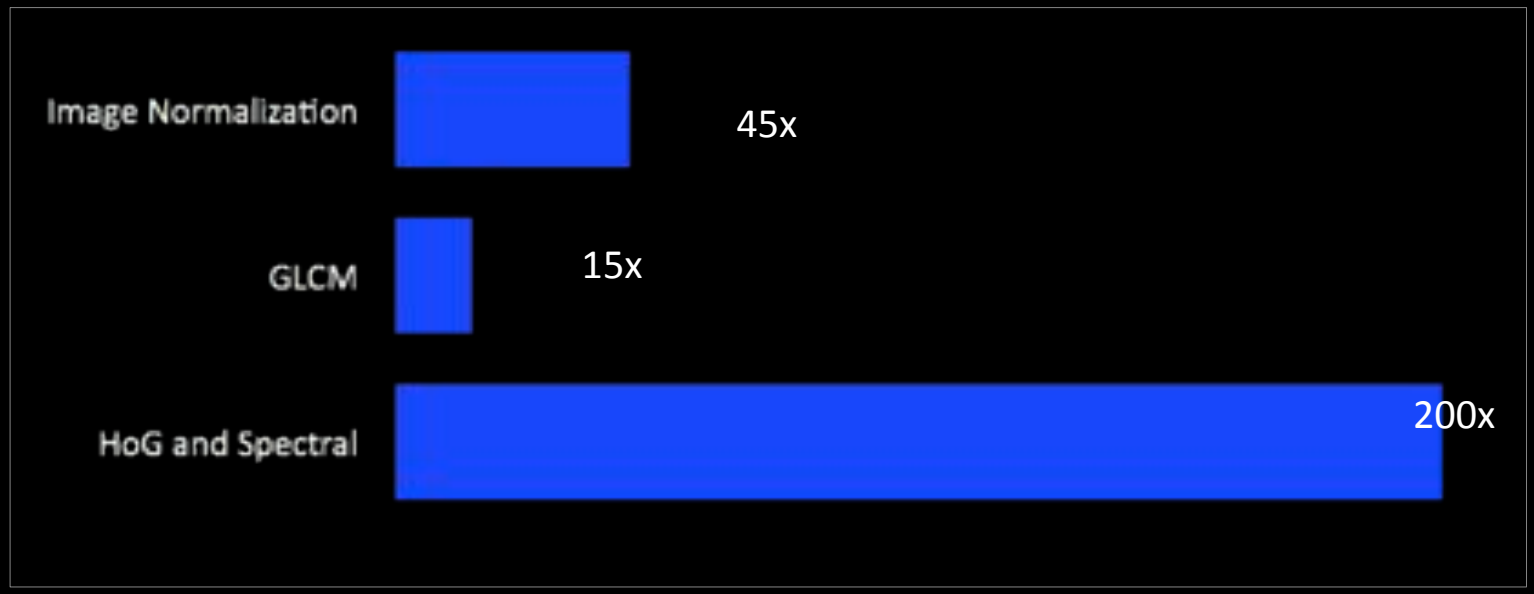

90
86
82
78
74
70
Accuracy %
8x8
16x16
24x24
32x32
Tile Size
5 Scales
3 Scales
2 Scales
1 Scale

PROCESSING TIME

FEATURE RUNTIME



SPEEDUP



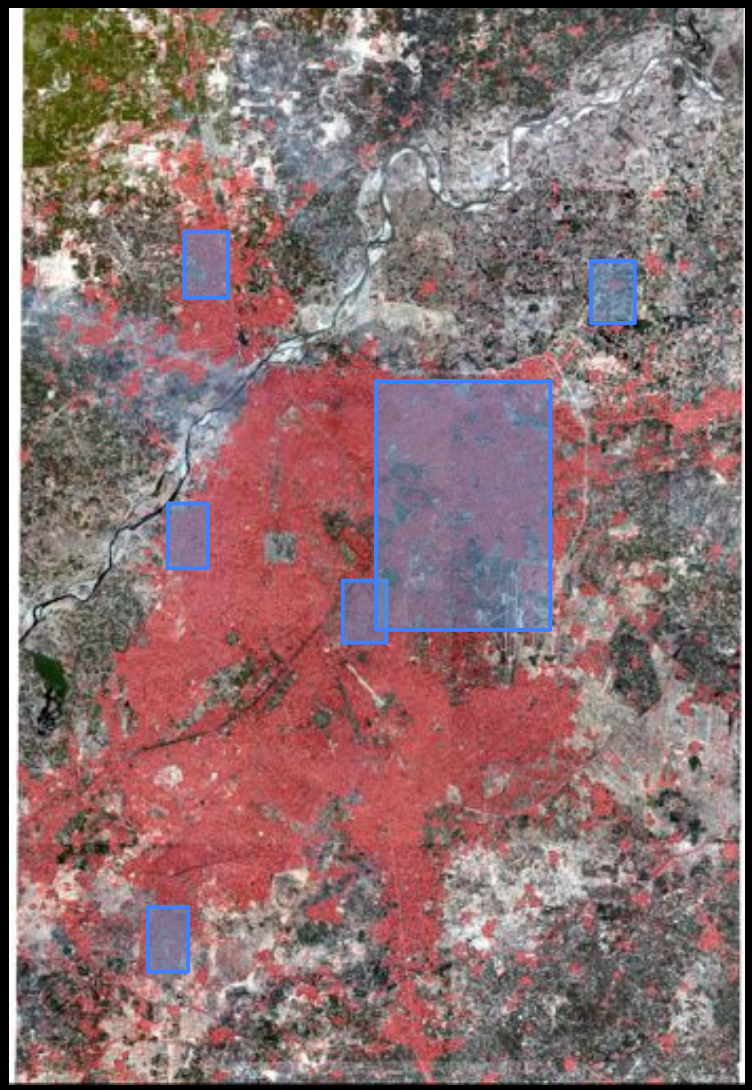
LAHORE SETTLEMENT EXTRACTION

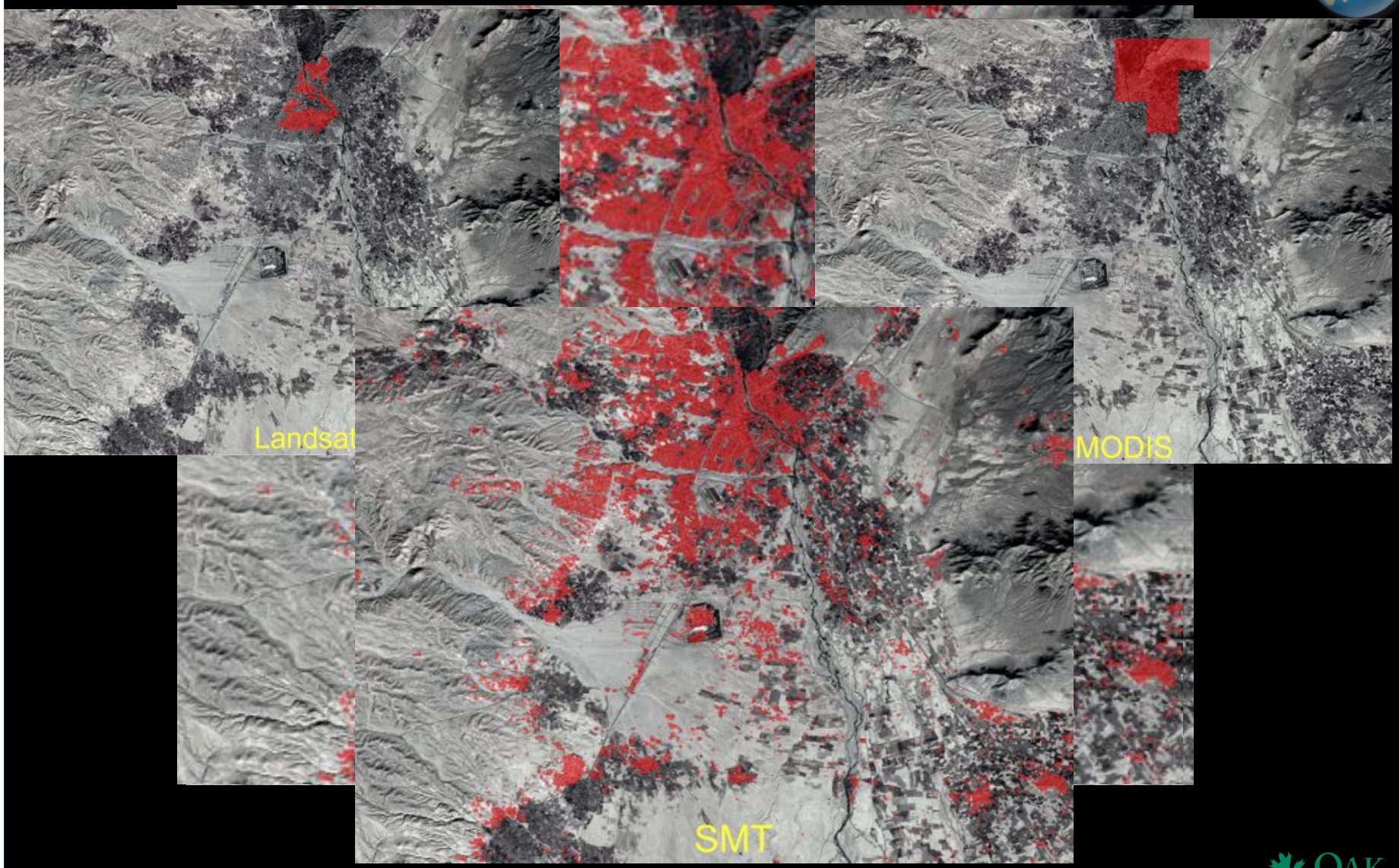
Geographic Information Science and Technology

- Data
  - 0.6 m spatial resolution
  - RGB bands processed
- Processed 1200 sq. km in 56 sec
- Wall to Wall processing

```
Non-settlement accuracy (class 0) is 96.7281%
Settlement accuracy (class 1) is 82.3651%
Overall Accuracy = 0.93
*********************************************
            Predicted
*********************************************
   Actual          0          1
         0      71869       2431
         1       4535      21181
*********************************************
Class      Recall    Precision    Harmonic Mean
  0          0.97        0.94         0.95
  1          0.82        0.90         0.86
```

LAHORE




Landsat
MODIS
SMT

NEXT

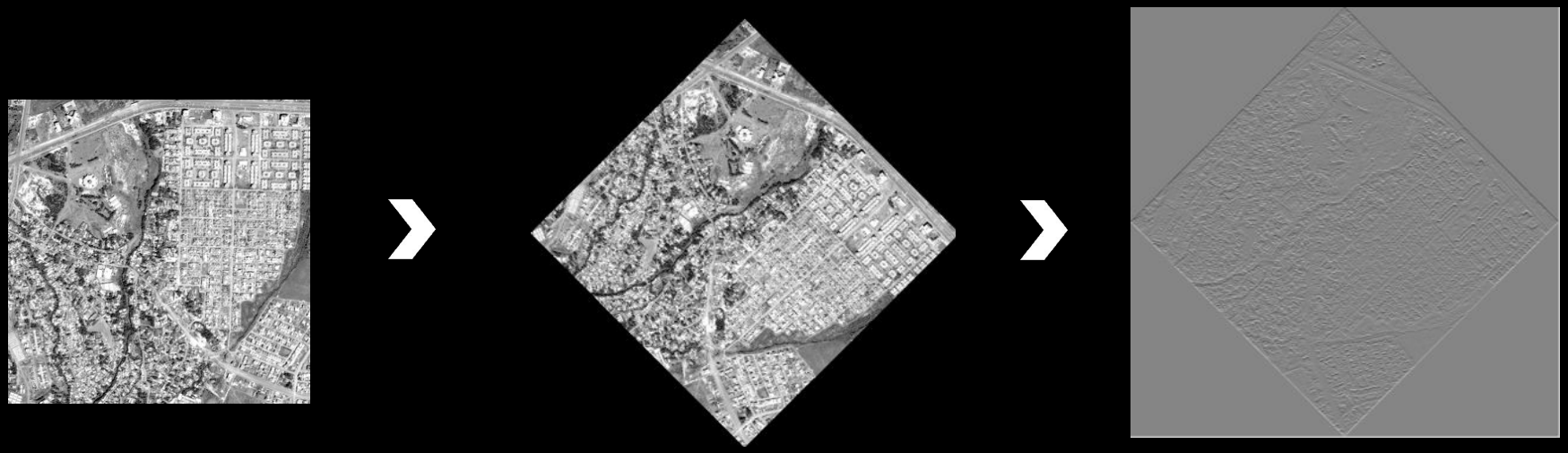
- Increase the accuracy.

Implemented filter responses of the Leung-Malik set. The LM set consists of 48 filters.

Speedup of separable filter response GPU implementation over non separable filter response ~40x

- So what about non separable filters????
- Solution : rotate the image at the cost of extra processing time but still get a speedup of ~30x over non separable

MPI+GPU implementation to tackle country scale imagery on a cluster - already in place. ☺



THANK YOU