LA-UR- 09-01648

Approved for public release; distribution is unlimited.

Title: Detection of Facilities in Satellite Imagery using Semi-Supervised Image Classification and Auxiliary Contextual Observables

Author(s): Neal R. Harvey
Christy Ruggiero
Steven Brumby
Norma Pawley
Brian MacDonald
Lee Balick
Alden Oyer

Intended for: Proceedings of 2009 SPIE Defense, Security & Sensing Conference

Detection of Facilities in Satellite Imagery using Semi-supervised Image Classification and Auxiliary Contextual Observables

Neal R. Harvey*, C. Ruggiero, S. Brumby,
N.H. Pawley, A. Oyer, R. Hixson,
L. Balick and B. MacDonald

Los Alamos National Laboratory,
Los Alamos, NM 87545, USA

ABSTRACT

Detecting complex targets, such as facilities, in commercially available satellite imagery is a difficult problem that human analysts try to solve by applying world knowledge. Often there are known observables that can be extracted by pixel-level feature detectors that can assist in the facility detection process. Individually, each of these observables is not sufficient for an accurate and reliable detection, but in combination, these auxiliary observables may provide sufficient context for detection by a machine learning algorithm.

We describe an approach for automatic detection of facilities that uses an automated feature extraction algorithm to extract auxiliary observables, and a semi-supervized assisted target recognition algorithm to then identify facilities of interest. We illustrate the approach using an example of finding schools in Quickbird image data of Albuquerque, New Mexico. We use Los Alamos National Laboratory's Genie Pro automated feature extraction algorithm to find a set of auxiliary features that should be useful in the search for schools, such as parking lots, large buildings, sports fields and residential areas and then combine these features using Genie Pro's assisted target recognition algorithm to learn a classifier that finds schools in the image data.

Keywords: facility detection, feature extraction, target recognition, machine learning, remote sensing

1. INTRODUCTION

Detecting complex targets in satellite imagery data, such as industrial facilities, is a difficult task which has a data and time intensity that quickly overwhelms human analysts. Accordingly, automated target detection is a huge area of research. Target detection in satellite imagery tends to be focused on the detection of more compact objects, such as vehicles or particular buildings. These targets generally have a well defined structure. Target detection algorithms generally involve finding some defining characteristics or features of the target of interest and looking for the best match to these characteristics or features in the data.[1,2]

Facility detection is a related area to target detection, if one simply considers that the specific facility type is the target (a facility being defined as a building or complex of buildings, installations, or place that provides a particular service or is used for a particular industry). However, there are differences between facility detection and what is commonly thought of as target detection. Facilities, as targets, tend to be larger, more complex and more varied from one example to another. Facilities of a particular type will obviously have similar characteristics, but in general the variability of the interrelationships between useful features is much greater.

There has been some research into developing algorithms to find "target" that are larger and more complex than the traditional targets. Zhenwei et al[3] describe an approach for regions of interest detection in remote sensing imagery. Their approach involves breaking the image up into tiles, calculating some linear-feature measure for each tile, based on length of edges and then grouping the tiles using the mean-shift segmentation method. This approach allows the detection of regions of interest such as airports, based on the detection of the large linear features of the runways.

Further author information: Neal R. Harvey: E-mail: harve@lanl.gov, Telephone: +1 505 667 9077

Sengupta et al[4] describe an approach for detection of regions of human settlement. The approach involves clustering of the images based on spectral information, then breaking the image into tiles and calculating features from the class statistics for the tiles, dimensionality reduction of the feature space followed by clustering of the reduced-dimensionality data and labeling as human settlement tiles based on the class-label statistics calculated in the first spectral clustering. A second human-settlement labeling is done using co-registered panchromatic imagery, based on tiling of the image, calculation of the number of corner and edge points in each tile and a tile labeled as human settlement by thresholding based on the corner and edge-point numbers (or clustering based on these numbers and some other gray-level texture measures). The intersection of the first spectral-labeling technique and the second gray-level labeling technique produces the final human settlement labeling of the tiles.

In this paper we present a new approach to detecting complex targets, such as facilities, by combining an automated feature extraction approach, using the Genie Pro automated feature extraction (AFE) software developed at LANL with a semi-supervised assisted target recognition (ATR)/target detection approach, using appropriate modifications to the Genie Pro software. The goal in feature extraction or classification is to assign a label to every pixel in an image, designating to which particular class it belongs. The assignment of class labels is usually performed based on a set of features calculated for each pixel. The goal in target detection is to define the locations of targets of interest. This can be achieved by identifying only a single pixel for each target. The distinction between target detection and feature extraction is subtle. This distinction is further discussed in sections 2.1.1 and 2.1.2. While this distinction may be subtle, the application of this combination approach provides dramatic results for detecting specific facility types in satellite imagery.

2. ALGORITHMS

2.1 Genie Pro

Turning raw satellite imagery into semantically meaningful maps and overlays is an important area of activity in remote sensing. Image analysts have the task of transforming imagery into maps of terrain classification such as crop types, road networks, buildings, etc. To be useful, this information needs to be up to date and it is therefore necessary that the analysis be performed quickly and as accurately as possible. Manual extraction of these features is time consuming and expensive and thus there is an urgent need for automated tools to assist analysts in deriving these maps. Genie Pro is one such tool.

Genie Pro is a general purpose, interactive, adaptive tool for developing pixel classification algorithms for image analysis, using training data provided by a human expert.[5,6] Genie Pro is a rewrite of the original GENIE software,[7–9] incorporating enhancements in usability, stability, reliability, performance and responsiveness.

2.1.1 Automated Feature Extraction

Most mapping tasks can be seen as variations on the 'pixel labeling' problem: assigning a label to every pixel in an image that represents a category or feature of the scene under that pixel. Accurately assigning a label to a pixel requires making use of the information in the image in a highly data- and task-specific way. Genie Pro knows very little about specific labeling tasks. Instead it has the ability to learn from examples provided by an image analyst. Genie Pro has a flexible and general purpose approach to pixel classification that can be applied to a great variety of different targets and different imagery types. The essential idea is that rather than deciding in advance what kinds of attributes or characteristics will be useful in classifying pixels, Genie Pro has a "toolbox" of image processing operators with which it can develop attributes that it determines to be most useful. It searches the space of possible "attribute extractors" that derive numerical values from local pixel neighborhoods. These attributes can then be combined using a more conventional machine learning framework (statistical classifier) to produce a final predicted class label for the pixel. Attribute extractors are themselves composed of simple image processing operations joined together into small image processing pipelines. The resulting flexible structure allows the attribute extractors to derive numerical measures that describe a variety of image characteristics: from spectral characteristics, through texture properties on to morphological properties and spatial context.

For further details about the Genie Pro tool and how it works, the interested reader is referred to the previously published work in the area.[5,6,10]

2.1.2 Assisted Target Recognition (ATR)/Target Detection

While pixel-by-pixel classification tools such as Genie Pro, as described above, are useful in broad area mapping applications, they are less well-suited to the detection of discrete objects. Pixel classifiers, however, do have many advantages: they are simple to design, they can readily employ formal machine learning tools and they are widely available on a variety of platforms. By making some minor modifications, these pixel classifiers can be used in object/target-detection settings.

For target detection problems, what is ultimately desired is not necessarily an image in which each pixel has been individually labeled as target or non-target. Often, the desired output is not even an image, but a list of target locations. To be useful to an analyst, a target detector need identify only a single pixel for each object. Once the attention of the analyst has been focused, there is little gain in having multiple pixels identified for a single target.

Thus, for the target-detection scenario, instead of having a metric that attempts to precisely delineate every pixel comprising the objects/targets of interest, we define a new metric that aims to focus the analyst's attention to these targets, without the distraction of many false alarms. The metric used is related to the multiple-instance problem[11–13] in machine learning. In this metric, a target is a set of pixels (typically compact and contiguous), and the classification associated with the target is given by the "or" of the classifications assigned to each of the pixels. That is, if any of the pixels in the target are "hits" then the target is detected. Any hits outside the target are false alarms. To apply this metric, the targets in the training data must be fully delineated and separately identified. Since the output of many pixel classifiers is a continuous real-valued quantity (to which a threshold is applied to make the ultimate decision as to the pixel label), this target-detection scoring is implemented by replacing all the pixels in each target object with the maximum-valued pixel in that object. This usually permits a much higher threshold than would be needed to hit all or most of the pixels in a target; and an increased threshold will, in general, produce fewer false alarms outside the object.

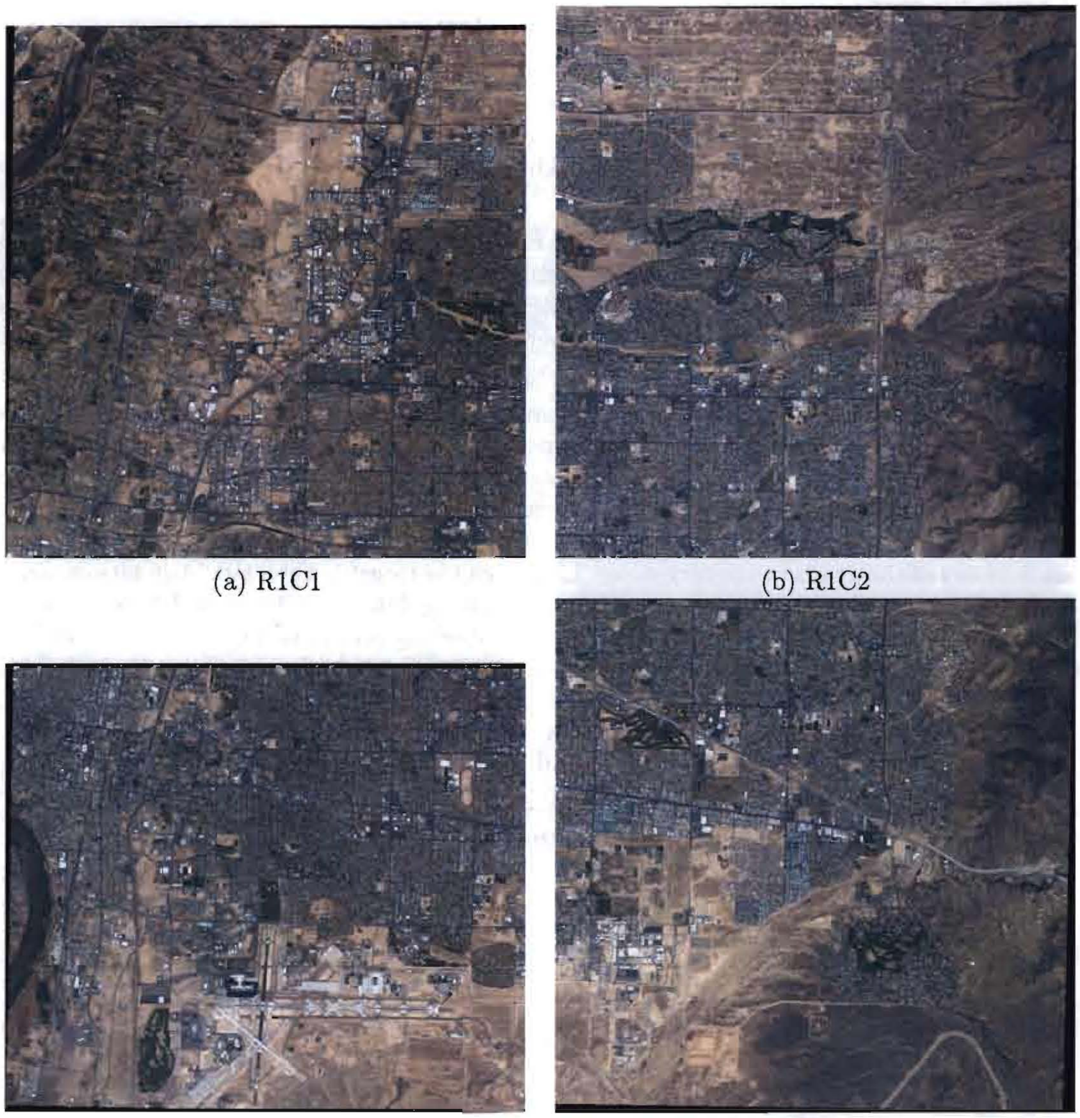
For further details about the target-detection (focus-of-attention) metric and modifications, the interested reader is referred to the previously published work in the area.[14, 15]

3. DATA SETS AND EXPERIMENTS

3.1 Data

The data we used for the experiments described here was Quickbird[16] multi-spectral imagery (4 band: Blue, Green, Red and Near Infra-Red). The particular set of Quickbird data used was that taken over Albuquerque on April 23rd 2002. The "whole", scene covering the city of Albuquerque, consisted of 4 images, each being approximately 4,000 pixels on a side:

1. 02APR23174958-M2AS_R1C1-005560582010_01_P001: labeled as R1C1 in Fig. 1 (a)
   North-western part of the city of Albuquerque.
2. 02APR23174958-M2AS_R1C2-005560582010_01_P001: labeled as R1C2 in Fig. 1 (b)
   North-eastern part of the city of Albuquerque.
3. 02APR23174958-M2AS_R2C1-005560582010_01_P001: labeled as R2C1 in Fig. 1 (c)
   South-western part of the city of Albuquerque.
4. 02APR23174958-M2AS_R2C2-005560582010_01_P001: labeled as R2C2 in Fig. 1 (d)
   South-eastern part of the city of Albuquerque.



(a) R1C1 (b) R1C2

(c) R2C1 (d) R2C2

Figure 1. Quickbird imagery used for experiments

3.2 Experiments

The basic experiment undertaken was to develop a method for finding facilities of interest within multi-spectral imagery using the Genie Pro feature extraction tool. The particular type of facility of interest for these experiments was schools - specifically, High Schools. The reason for choosing High Schools is that they are fairly large facilities, there are a reasonable number of examples with which to train and test detection algorithms and there is a lot of information available regarding their locations which can be used to develop training and test data sets and to determine performance measures for the algorithms, such as Google, Google Maps, Google Earth, Yahoo Maps, and the Community Link web page and map for Albuquerque.[17]

Limitations of time and space preclude a detailed description of our initial experiments in trying to develop facility-detection algorithms using just the raw multi-spectral image data - where we left it entirely up to the Genie Pro automated feature extraction software to find appropriate characteristic features of the facilities of interest, from only the raw image data, and the best combination of these features. Suffice it to say that the facility-detection algorithms developed using just the multi-spectral image data had an unacceptably high false-alarm rate.

We restrict ourselves in this paper to describing our multi-stage approach that uses an expert's knowledge of the problem to assist our software in its search. This approach uses automated feature extraction (AFE) algorithms to extract a set of pre-defined (by the human expert) auxiliary observables followed by an assisted target recognition (ATR) algorithm to then identify facilities of interest. The multi-stage approach proceeds as follows. In the first stage, a subject-matter expert generates a list of features (what we refer to as auxiliary observables) that may be useful in detecting a particular type of facility. These auxiliary observables, on their own, may not be sufficient to enable accurate and reliable detection of the facilities of interest. However, in combination, they could provide information that the software can use to develop good detection algorithms. In the second stage, therefore, Genie Pro is used (in AFE mode) to derive algorithms capable of finding each of the defined auxiliary observables in the available data. In the third stage, the outputs of applying the auxiliary observable-detection algorithms to the available image data are stacked as additional 'pseudo-bands', together with the original image data, and provided as input to the Genie Pro software (now operating in ATR mode), which is used to develop an algorithm to detect the facilities of interest.

3.2.1 Auxiliary Features

For the High School detection problem, the list of auxiliary observables that were determined by the human analyst to be potentially useful was as follows:

1. Athletic fields
2. Parking lots
3. Large buildings
4. Residential areas

None of these auxiliary observables, individually, is sufficient to define the presence of a High School. However, they are all features that one can see at most High Schools. How one should use this information, or combine the features, together with any additional useful information that can be gleaned from the original image data, is what we are asking our Genie Pro software to determine.

For each of the auxiliary observable-detection problems, the optimization was a two-class problem: the feature of interest against a background of everything else. Figure 2 shows some examples of the training data provided for developing the AFE algorithms for finding the auxiliary observables. Figure 2 (a) shows an example for athletic fields, (b) for parking lots, (c) for large buildings and (d) for residential areas. In these images, for each of the auxiliary observable-detection problems the training data provided for the particular feature of interest is shown overlaid onto the original data in green, while the training data provided for the background is shown overlaid in red.

3.2.2 High Schools

Ground truth used by the human analyst to mark-up the training and test data for the High Schools was obtained from Google, Google Maps, Google Earth, Yahoo Maps and the Community Link web page and map for Albuquerque.[17]

For training data, an analyst marked up (delineated) the campus area for four High Schools, two in image R1C1: Valley High School and Sandia Preparatory School and two in image R1C2: Albuquerque Academy and Sandia High School. In addition to the training data for the class of interest (High Schools), background training data was provided that included examples of industrial, commercial and retail areas, purely residential areas and golf courses.

Figure 3 shows the training data provided for High Schools. Figure 3 (a) shows all of the training data provided for image R1C1, (b) shows all of the training data provided for image R1C2, (c) shows a close-up of the training data provided for one of the High Schools used in training (Sandia Preparatory School) and (d) shows a close up of some background regions, containing large buildings, parking lots, roads and residential areas. In

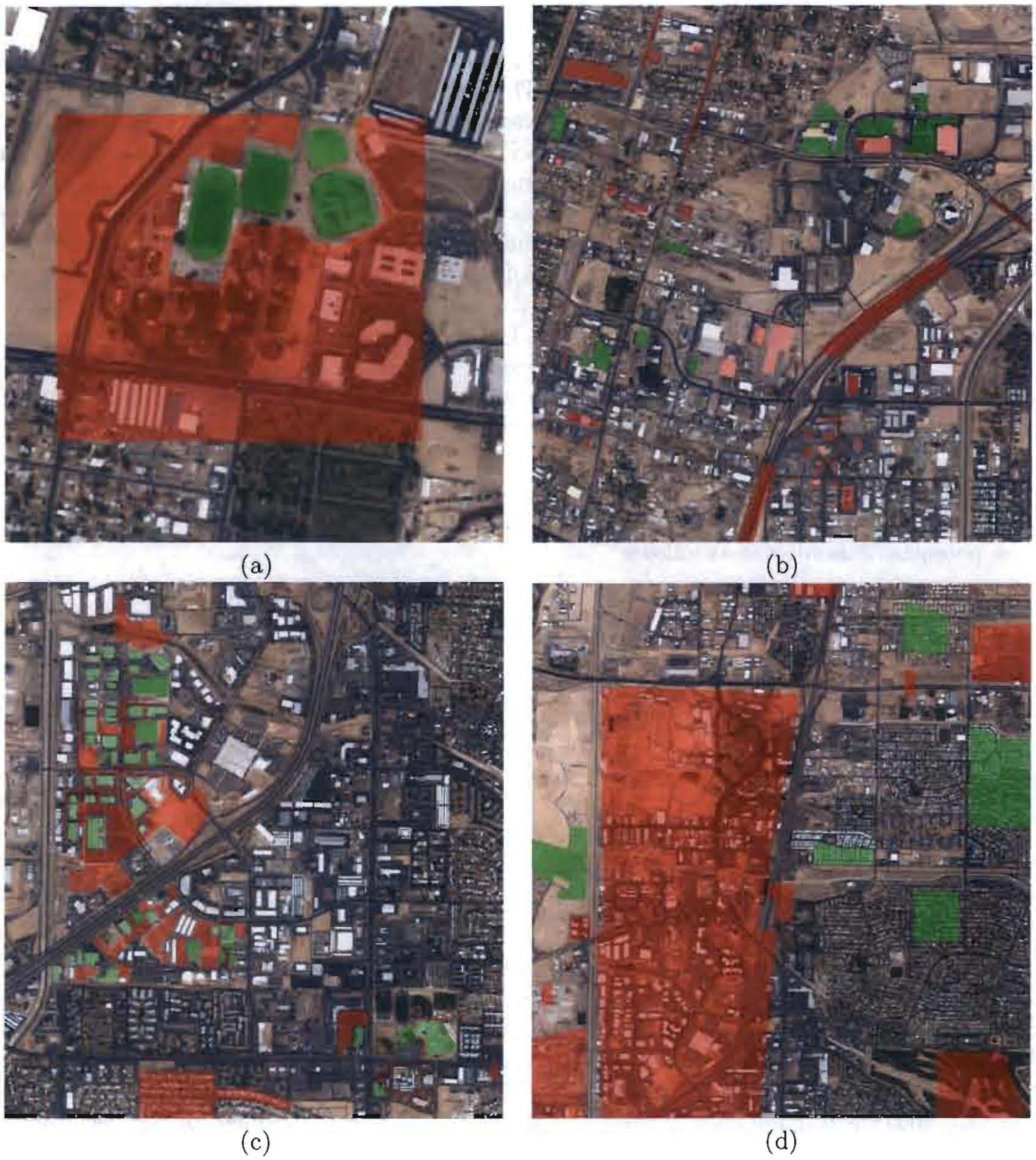
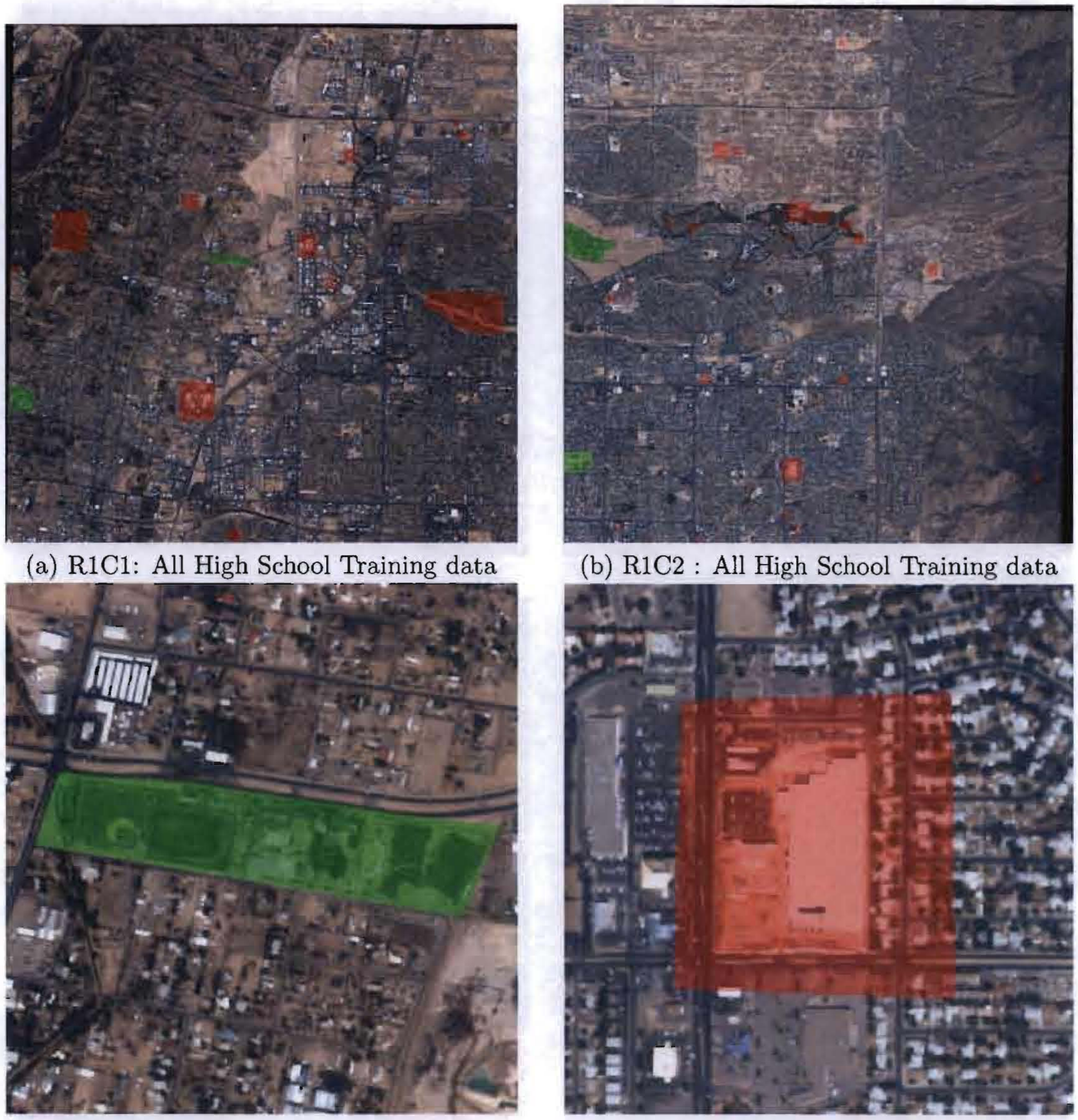


Figure 2. Example training data Auxiliary Features: (a) Athletic Fields; (b) Parking Lots; (c) Large Buildings; (d) Residential Areas. Note images shown are not necessarily at the same spatial scale.

these images, just as for the auxiliary observables, the training data provided for the facility of interest is shown overlaid onto the original data in green, while the training data provided for the background is shown overlaid in red.

Note that large sections of the images that contain the training data (R1C1 and R1C2) are unmarked and were therefore not seen/used during training and thus this data can be considered testing data in addition to the two images containing no training data at all (R2C1 and R2C2).



(a) R1C1: All High School Training data (b) R1C2 : All High School Training data

(c) R1C1: Training Data School Example (d) R1C2: Training Data Background Example

Figure 3. Training data used for High School detection experiments. Green overlay areas show high schools marked up for training. Red overlay areas show background marked up for training. (a) Shows all training data for R1C1 image; (b) Shows all training data for R1C2 image; (c) Shows an example close up of the training data marked-up for a school-Sandia Preparatory School; (d) Shows an example close up of a background region - containing large buildings, parking lots, roads and residential areas

4. RESULTS

4.1 Auxiliary Features

Figure 4 shows the results of applying the auxiliary observable-finding algorithms to some of the testing data - i.e. data not seen during training. Fig. 4 (a) shows the results for athletic fields, (b) shows the results for parking lots, (c) shows the results for large buildings and (d) shows the results for residential areas. In these images, just as for the training data shown above, the pixels determined by the auxiliary observable extraction algorithm to be the class of interest are overlaid in green and those determined to be background are overlaid in red.

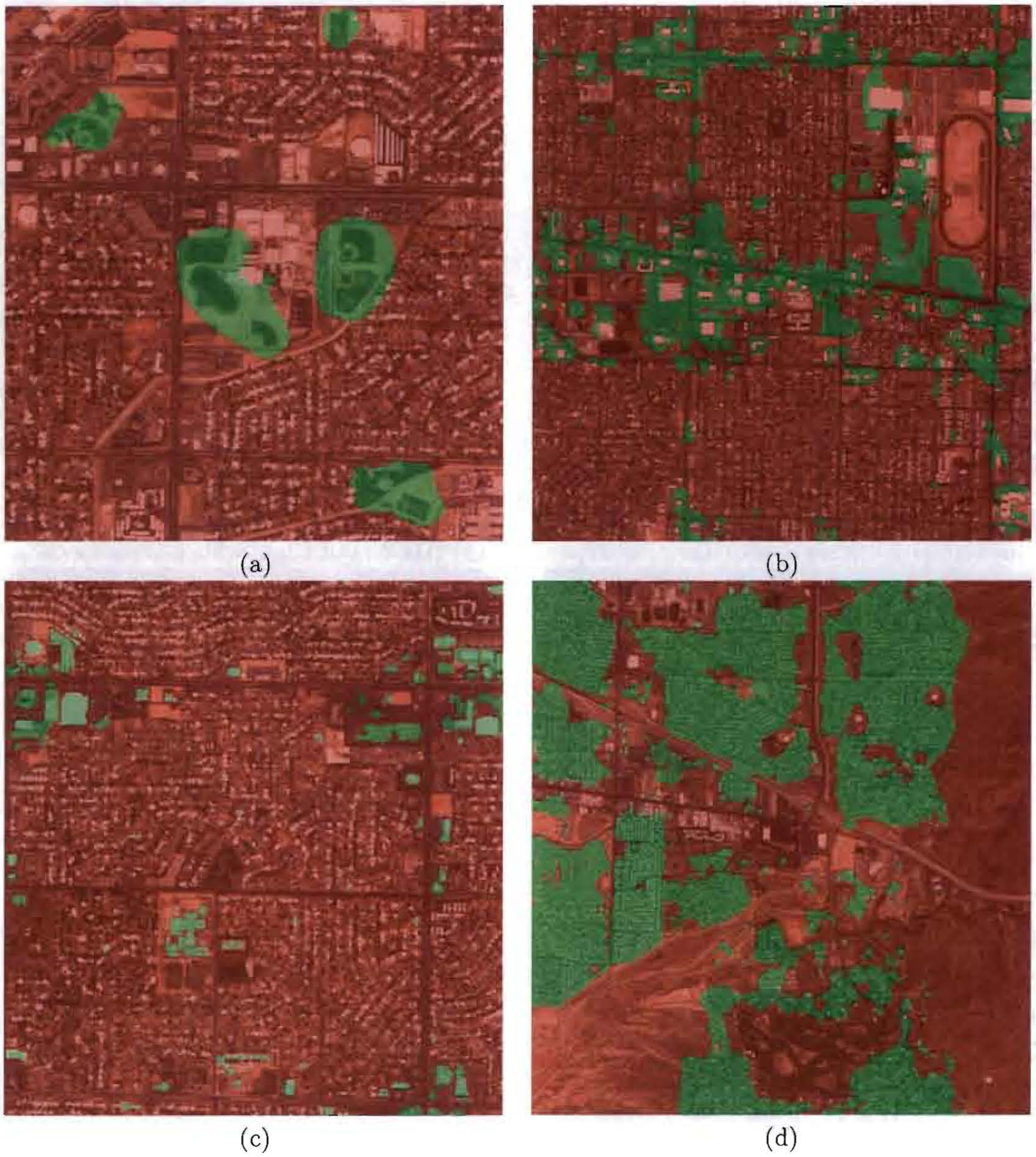


Figure 4. Example results applied to testing data (data not seen during training) for Auxiliary Features: (a) Athletic Fields; (b) Parking Lots; (c) Large Buildings; (d) Residential Areas. Note images shown are not necessarily at the same spatial scale.

4.2 Schools

There are a total of twelve High Schools known to be present in the regions covered by the Albuquerque Quickbird images. Four of these schools were used in training, leaving eight High Schools for testing.

On the training data the High School detection algorithms performed extremely well. The algorithms had a detection rate (DR) of 100% and had zero false alarms, i.e. a false alarm rate (FAR) of 0%.

On the test data, the High School detection algorithms detected 7 of the 8 High Schools, a detection rate (DR) of 87.5%. The one High School that was not detected was the one that does not have athletic fields on campus. With respect to false alarms these fell into several categories, as described below.

Table 1. Number of false alarms (FAs) for the categories enumerated

Category Number	Number of FAs
1	15
2	7
3	2
4	2
5	12
6	19
7	13
8	9
9	2

1. Bonus detections. These are detections of schools, other than High Schools: Elementary Schools and Middle Schools. We do not consider these true false alarms, as they are actually schools. However, they are not High Schools, which was the specific type of school provided in the training data.
2. UNM Campus facilities. It is hard to decide whether these should be considered bonus detections or not. These are hits within educational facilities. But, they are not really "school", as such.
3. US Army RSRV Training Center. Similar to the UNM campus facilities, it is hard to decide whether or not these should be considered bonus detections. These are hits within an "educational facility" of sorts.
4. Unused industrial/construction areas, consisting of large patches of vegetation, near empty ground (which look a lot like a parking lot to the naked eye) and some large buildings.
5. Large houses next to golf courses.
6. Large buildings with parking lots next to golf courses.
7. Municipal Parks, consisting of large buildings with parking lots next to large green areas.
8. Large roads/runways next to golf courses.
9. Electrical sub-station (consisting of some large buildings and a parking area, etc.) adjacent to a golf course.

Figure 5 shows the results of applying the High School-finding algorithms to some of the test data - i.e. data not seen during training. Fig. 5 (a) and (b) show the detections for two High Schools: (a) is Del Norte High School and (b) is Albuquerque High School, (c) and (d) shows the detections for some true false alarms: (c) is Ross Enchanted Park and (d) are large buildings adjacent to UNM Golf Course, (e) and (f) show bonus detections: (e) is Onate Elementary School and (f) is Wilson Middle School. In these images the centroids of the regions detected are overlaid as light blue squares, over a general red background overlay.

5. DISCUSSION AND CONCLUSIONS

The High School detection algorithm developed using the approach described here had very good performance. On the training data the algorithm had perfect results: 100% DR and 0% FAR. But that is on data seen during training and one would expect that an algorithm optimized on a set of training data would work well on that data. On the test data, the developed High School detection algorithm was able to successfully detect all but one of the High Schools in the Albuquerque area. The High School it wasn't able to detect (Freedom High School), was the only one that didn't have Athletic Fields on campus. Given that all the High Schools used in training had associated athletic fields and the human analyst had determined that athletic fields were a sufficiently useful feature for High School detection (they had been chosen as one of the auxiliary observables) the missed detection is perhaps not so surprising.

With regard to false alarms on the test data, there were a total of 79 hits on non High School areas. Of these 79 false alarms, 15 were detections of schools other than actual High Schools: Elementary and Middle schools. Given that they are actual schools detections, but are not the High Schools that were the type used in training, we don't classify these as "true" false alarms, but instead use the term "bonus" detections. Similar to the Elementary and Middle School bonus detections, 7 of the 79 false alarms were hits on UNM campus facilities. Given that these are hits on "educational" facilities we could perhaps put these hits into the bonus detection category as well. An additional 2 of the 79 false alarms were hits on a US Army RSRV Training Center. Now, it is possible to consider this an "educational" facility, but perhaps the link gets a little tenuous. However, looking at the campus of the Training Center it is very easy to see similarities to a High School, and it is understandable why our algorithm should label this as a High School. Of the remaining false alarms, there were certain similarities between them. They consisted of regions having large buildings, near large green areas, such as parks and golf courses and areas that looked like parking lots (such as runways and large, empty stretches of interstate), that were not very far from residential areas. In fact, for some of the false alarms that hit on Municipal Parks, a human analyst wasn't able to differentiate these areas from Middle or Elementary schools just from looking at the data. It was necessary to gather additional information from maps, etc. in order to make the determination. If we ignore the Middle and Elementary bonus detections, we have a total of 66 false alarms (that could be reduced to 57 if we took out the UNM Campus and Army Training Center hits). Our quickbird mage data covered an area of approximately 366 Km^2. For our 66 false alarms, then, our school detection algorithm therefore had a false alarm rate of 0.18 / Km^2 or 1 false alarm / 5.5 Km^2.

Our experiments have shown that our multi-stage approach of defining a set of auxiliary contextual obervables for which we develop automated feature extraction algorithms and use the outputs from these as inputs to the facility detection stage works very well, for the specific example facility type chosen. We used High Schools as our example "facility of interest" and chose a set of auxiliary observables that were relevant to that particular problem. For other facility detection problems it will be necessary to determine an appropriate set of auxiliary obervables for which suitable feature extraction algorithms can be developed. What these auxiliary observables will be will depend on both the particular facility of interest and the type of data to be used.

ACKNOWLEDGMENTS

This work was funded by the US Department of Energy.

REFERENCES

[1] Bhanu, B., "Automatic target recognition: State of the art survey," *IEEE Trans. Aerospace and Electronic Systems* **22**(4), 364–379 (1986).

[2] Dudgeon, D. and Lacoss, R., "An overview of automatic target recognition," *The Lincoln Laboratory Journal* **6**(1), 3–9 (1993).

[3] Shenwei, G., Wenxian, Y., Yongmei, J., and Yi, S., "Regions of interest detection in large remote sensing images," in [*Proc. SPIE*], **6044** (November 2005).

[4] Sengupta, S., Lopez, A., and Poland, D., "Class-label statistics: A basis for fusing information from multi-temporal imagery with applications to unsupervised detection of human settlement," in [*Proc. IEEE Int. Geoscience & Remote Sensing Symposium (IGARSS)*], **2**, 1084–1087 (2004).

[5] Perkins, S., Edlund, K., Esch-Mosher, D., Eads, D., Harvey, N., and Brumby, S., "Genie pro: Robust image classification using shape, texture, and spectral information," in [*Proc. SPIE*], **5806**, 139–148 (2005).

[6] "Genie pro: Genetic imagery exploitation." http://www.geniepro.lanl.gov/.

[7] Theiler, J., Harvey, N., Brumby, S., Szymanski, J., Alferink, S., S.J.Perkins, Porter, R., and Bloch, J., "Evolving retrieval algorithms with a genetic programming scheme," in [*Proc. SPIE*], **3753**, 416–425 (1999).

[8] Harvey, N. R., Perkins, S., Brumby, S. P., Theiler, J., Porter, R. B., Young, A. C., Varghese, A. K., Szymanski, J. J., and Bloch, J. J., "Finding golf courses: The ultra high tech approach," in [*Lecture notes in computer science: Real-World Applications of Evolutionary Computing*],

[9] Harvey, N., Theiler, J., Brumby, S., Perkins, S., Bloch, J. S. J., Porter, R., Galassi, M., and Young, A., "Comparison of genie and conventional supervised classifiers for multispectral image feature extraction," *IEEE Trans. Geoscience and Remote Sensing (TGARS)* **40**, 393–404 (February 2002).

[10] Perkins, S., Harvey, N., brumby, S., and Lacker, K., "Support vector machines for broad area feature extraction in remotely sensed images," in [*Proc. SPIE*], **4381**, 286–295 (2001).

[11] Dietterich, T. G., Lathrop, R. H., and Lozano-perez, T., "Solving the multiple-instance problem with axis-parallel rectangles," *Artificial Intelligence* **89**, 31–71 (1997).

[12] Maron, O. and Ratan, A., "Multiple-instance learning for natural scene classification," in [*Proc. 15th International Conference on Machine Learning*], 341–349, Morgan Kaufmann (1998).

[13] Zhang, Q. and Goldman, S., "Em-dd: An improved multiple-instance learning technique," in [*Advances in Neural Information Processing Systems*], 1073–1080, MIT Press (2001).

[14] Harvey, N. and Theiler, J., "Focus-of-attention strategies for finding discrete objects in multispectral imagery," in [*Proc. SPIE*], **5546**, 179–189 (2004).

[15] Theiler, J., Harvey, N., David, N., and Irvine, J., "Approach to target detection based on relevant metric for scoring performance," in [*Proc. 33rd Applied Image Pattern Recognition Workshop (AIPR '04)*], 184–189 (2004).

[16] "Digitalglobe — an imagery and information company."
http://www.digitalglobe.com/.

[17] "Community link: Albuquerque, new mexico map."
http://www.communitylink.com/us/nm/albuquerque/map/city.

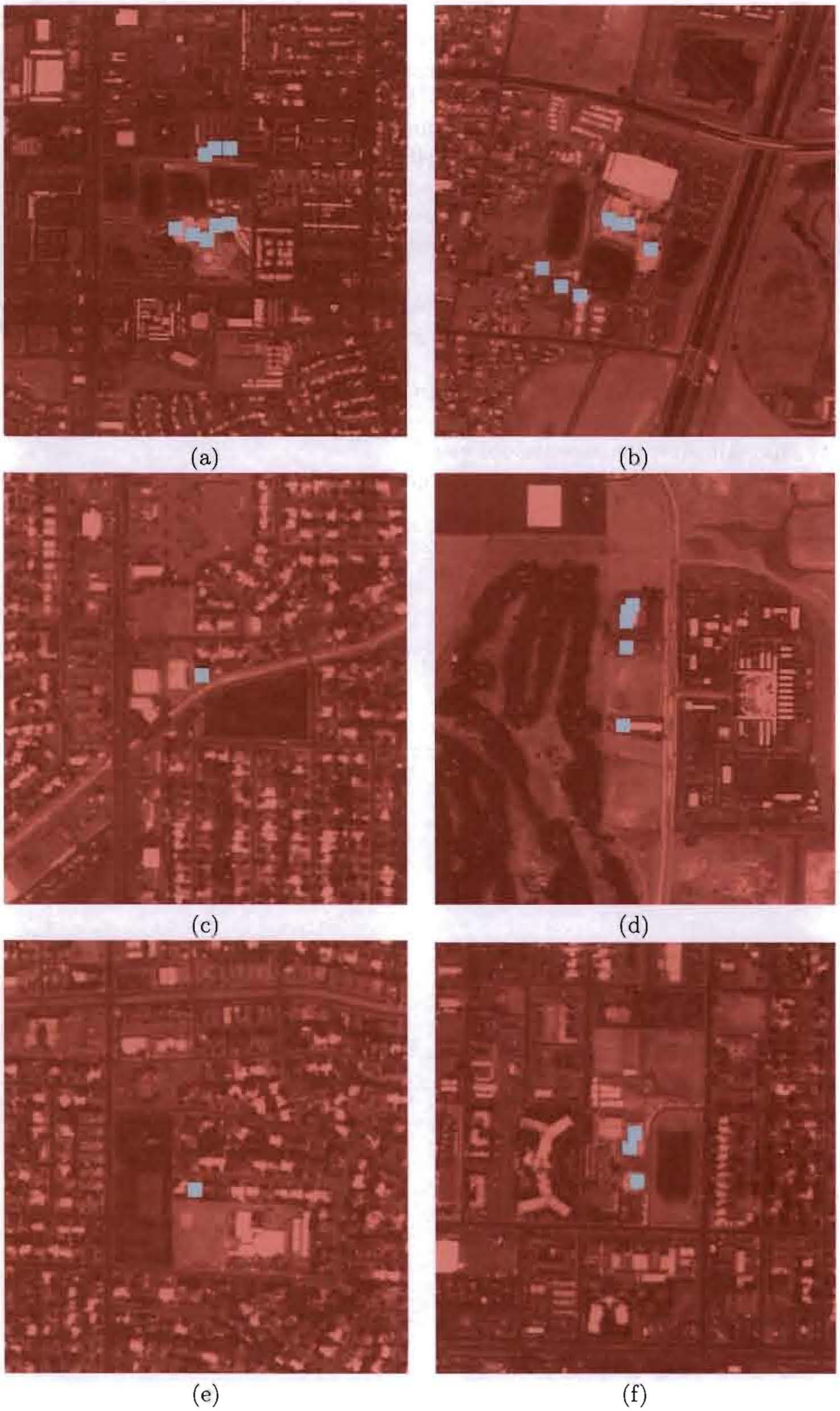


Figure 5. Example results for High Schools' Detection: (a) and (b) True Positives; (c) and (d) False Positives; (e) and (f) Bonus Detections (schools detected, but not High Schools). Overlaid blue squares highlight the centroids of the regions detected. Note images shown are not necessarily at the same spatial scale.