AUTOMATED EXTRACTION OF PLANTATIONS FROM IKONOS SATELLITE IMAGERY USING A LEVEL SET BASED SEGMENTATION METHOD

K. Vogt*, B. Scheuermann, C. Becker, T. Büschenfeld, B. Rosenhahn, J. Ostermann

Institut für Informationsverarbeitung, Leibniz Universität Hannover, Appelstraße 9a, 30167 Hannover, Germany –
{vogt, scheuerm, becker, bfeld, rosenhahn, ostermann}@tnt.uni-hannover.de
http://www.tnt.uni-hannover.de/project/wipka/

KEY WORDS: Land Use, Vegetation, Segmentation, GIS, Automation

ABSTRACT:

In this article we present a method that extracts plantations from satellite imagery by finding and exploiting appropriate feature space projections. Segmentation is done using an automatic two-region segmentation based on the level set method. The behaviour of this algorithm is defined by a statistical region model that describes the similarity of regions using distances in arbitrary feature spaces. Subsequently different feature spaces will be evaluated regarding their plantation classification quality in an automatic fashion. The segmentation quality of our method is assessed by testing several orthophotos depicting a wide range of landscape types and comparing them with a manual segmentation. We show that a combination of simple texture based features like the structure tensor and the Hessian matrix are sufficient to facilitate an effective plantation segmentation scheme.

1. INTRODUCTION

When looking at land cover analysis, digital imagery is not only used for interpretation, but also for the verification and updating of tagged regions stored in GIS databases. Thus, the efficiency of today's workflow can be greatly improved by using automatic classification algorithms. In our knowledge-based interpretation and verification system GeoAIDA (Becker et al., 2008), several operators are used to hierarchically analyse image data. Area processing is done using a classification system which incorporates a Markov/Gibbs random field model, as proposed by Gimel'farb (Gimel'farb, 1996). Using this model, plantations (as seen in Figure 1) are particularly difficult to separate from general vegetation areas, since they may not notably differ in their statistical properties. To tackle this challenging task we propose a segmentation on a region model that works with a general feature space representation. We analyse several potential feature spaces, each constructed from the local region of a pixel. By analysing distances in these feature spaces, we can calculate the likelihood that the pixel belongs to a certain region of the current segmentation. For region detection we decided to use a state-of-the-art level set based segmentation framework that allows us to test our region model in a straightforward manner. The insights from the semi-automatic feature space comparison are then incorporated into the segmentation framework and finally tested on real orthophotos.

2. LEVEL SET SEGMENTATION

The variational approach for image segmentation used in our framework is based on the works of (Chan and Vese, 2001, Malladi et al., 1995, Paragios and Deriche, 1999, Rosenhahn et al., 2007). Using the level set formulation for the general problem of image segmentation has several advantages. To allow a convenient and sound interaction between constraints that are imposed on the contour itself and constraints that act on the two regions separated by the contour, the 1-D curve is embedded into a 2-D, image-like structure. Another important advantage of the level set representation is the natural given possibility to handle topological changes of the 1-D curve. This makes it particularly useful for plantation segmentation, where no reliable constraints for the region shapes can be formulated.

* Corresponding author.

In the case of a two-region segmentation, the level set function $\varphi : \Omega \rightarrow \mathbb{R}$ splits the image domain Ω into the two regions $\Omega_1, \Omega_2 \subseteq \Omega$ with

$$\varphi(x) = \begin{cases} \geq 0, \text{ if } x \in \Omega_1 \\ < 0, \text{ if } x \in \Omega_2 \end{cases} . \quad (1)$$

The boundary between the object that is to be extracted and the background is represented by the zero-level line of the function φ. Like most of the works on level set segmentation, we will focus on this special segmentation case with two regions. In our work these will represent the object classes plantations and non-plantations. As shown in (Chan and Vese, 2001), we will require an optimal contour to minimise the so called *Chan-Vese energy functional* that is:

$$E(\varphi) = -\int_\Omega \Big(H(\varphi) \log p_1 + (1 - H(\varphi)) \log p_2 \Big)\, d\Omega + \nu \int_\Omega |\nabla H(\varphi)|\, d\Omega \quad (2)$$

where $\nu \geq 0$ is a weighting parameter between the two given constraints, p_i are probabilities and $H(s)$ is a regularised Heaviside function. The regularised Heaviside function is needed to build the Euler-Lagrange equation and to make it possible to indicate which region a pixel belongs to at each iteration step. Minimising the first term maximises the total a-posteriori probability given the the two probabilities p_1 and p_2 of Ω_1 and Ω_2, i.e., pixels are assigned to the most probable region according to the Bayes rule. The second term minimises the length of the contour and acts as a smoothing term.

Minimisation of the Chan-Vese energy functional (2) can be performed by solving the corresponding Euler-Lagrange equation to φ

$$\frac{\partial \varphi}{\partial t} = \delta(\varphi) \left(\log \frac{p_1}{p_2} + \nu \operatorname{div} \left(\frac{\nabla \varphi}{|\nabla \varphi|} \right) \right) , \quad (3)$$

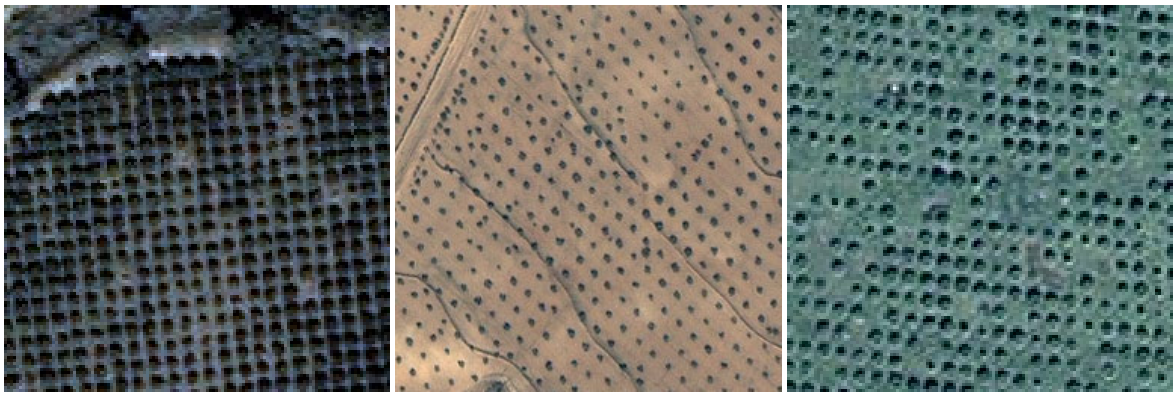

Figure 1: Plantation samples from different locations

where $\delta(s)$ is the derivative of $H(s)$, with respect to its argument. Starting with some initial contour φ^0 and given the probabilities p_1 and p_2 one has to solve the following initial value problem:

$$\begin{cases} \varphi(x,0) = \varphi^0 \text{ for } x \in \Omega \\ \dfrac{\partial \varphi}{\partial t} = \delta(\varphi)\left(\log \dfrac{p_1}{p_2} + \nu \operatorname{div}\left(\dfrac{\nabla \varphi}{|\nabla \varphi|}\right)\right) \end{cases} . \quad (4)$$

Figure 2 shows the workflow of our segmentation algorithm. The way the two probabilities p_1 and p_2 are modelled is a very important factor for the quality of the segmentation process. The next section will, therefore, introduce a statistical region model capable of estimating probability densities in arbitrary feature spaces.

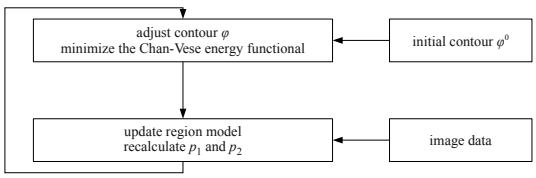

Figure 2: Workflow of the segmentation algorithm

3. STATISTICAL REGION MODEL

The purpose of the region model is to provide a model of the statistical properties underlying the region that is to be extracted by our segmentation algorithm. The interface between the region model and segmentation algorithm is defined by a function $p(x, y)$, mapping each pixel position to the probability that this pixel belongs to the modelled region of the current segmentation. Since $p(x, y)$ depends on the current contour of the region, which will change during the progression of the segmentation, a good presegmentation is vital in order to achieve high quality results. To estimate $p(x, y)$, we use the function $f(x, y)$, which defines a projection from image to feature space. We can now easily calculate the mean distance between each position and the modelled region in an arbitrary feature space. The probability function $p(x, y)$ is now defined by scaling the resulting distances to the range $[0; 1]$, while considering that larger distances will map to a lower probability and vice versa.

In case that several functions $f_c(x, y)$ are available, each will provide one independent feature channel c. The function $p(x, y)$ can then be calculated by multiplying the individual probability functions $p_c(x, y)$.

3.1 Distance Function

When incorporating several feature channels, a consistent scaling of the feature distances might become quite complex. We will, therefore, present a distance function based on the normalised cross correlation (NCC). The feature vectors $i, j \in \mathbb{R}^n$ will be interpreted as discrete signals of length n. The NCC is then defined as

$$\varrho_{ij} = \frac{1}{n-1} \sum_{t=0}^{n-1} \frac{(i(t) - \bar{i})\,(j(t) - \bar{j})}{\sigma_i \sigma_j}, \quad (5)$$

whereby $\bar{i}$ and $\bar{j}$ represent the mean component values of i and j. Since the range of the NCC is known, scaling the distances is quite trivial. Handling of anti-correlations depends on the application so that ϱ_{ij} may either be mirrored around or clamped at 0, keeping all values positive.

$$d_{NCC_1}(i,j) = 1 - |\varrho_{ij}| \quad (6)$$

$$d_{NCC_2}(i,j) = 1 - [\varrho_{ij}]^+, \quad (7)$$

$$with\ [x]^+ = \begin{cases} 0 & , x < 0 \\ x & , else \end{cases}$$

3.2 Presegmentation

In most cases, a good presegmentation can be derived from existing geodata. If the data is not available or too outdated, there is an easy method for generating one by analysing the frequency spectrum of image blocks from the source image. Typical images of plantations show a peak in a certain frequency range. After an appropriate range was learned from existing image data (see section 3.3.1), each image block is tested for the occurrence of such a peak and assigned to either the inner or outer region of the presegmentation. Figure 3 shows the result of the presegmentation step for one test case.

3.3 Feature Selection

Finding appropriate feature spaces is no easy task and is often done manually. Instead of designing one that is optimised for plantation segmentation, our strategy is to generate a wide array of generic feature spaces and compare them in a systematic manner. This selection process is guided by a list of training sets and can also be used to find feature spaces for other object classes. The general workflow shown in figure 4 is further explained in the following sections.

Figure 3: The presegmentation mask for a sample scene

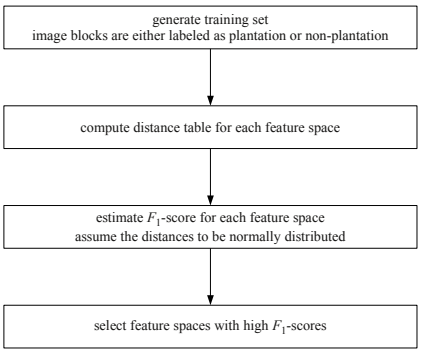

Figure 4: Workflow of the feature selection

3.3.1 Training Sets. In total, eight training sets were produced, each of them chosen to represent a different type of geographical location. Together, the training sets cover an area of approximately $75km^2$ subdivided into 4400 image blocks. These image blocks are then manually classified as being either of the plantation or non-plantation type.

For each training set, a distance table can be generated. A distance table contains the distances between all image blocks for a given feature space. It is convenient to sort the entries so that all plantation and non-plantation blocks are grouped together. The table can now be subdivided into four quadrants:

- Intra-distances between non-plantations
- $2\times$ Inter-distances
- Intra-distances between plantations

A distance table from a small training set is shown in Figure 5.

3.3.2 Classification Model. Our goal is to find feature-spaces that produce small intra-distances and large inter-distances. Of course, *small* and *large* are entirely subjective terms. It is sufficient that intra- and inter-distances are separable. To that end, we will treat that task as a classical binary classification problem.

Using the distance tables we can find normal distributions describing the intra- and inter-distances. As shown in Figure 6, by analysing the overlap of the normal distributions, we can compute the true positive (TP), false positive (FP), true negative (TN) and false negative (FN) rates. It is obvious that features with minimal overlaps will generate better classification results. To get a better understanding of the classification capacity of a given feature space, we will also calculate precision, recall and the corresponding F_1-score (Van Rijsbergen, 1979):

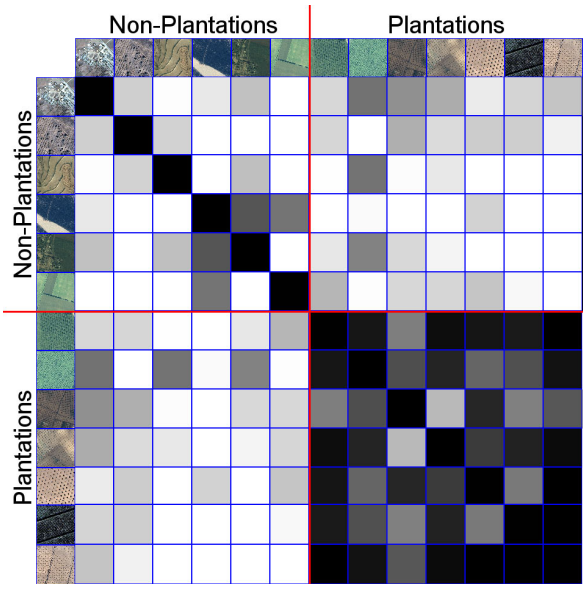

Figure 5: The distance table of a feature contains the distances between all image blocks in the training set. Dark entries signify a small distance in a given feature space and therefore a high similarity.

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (10)$$

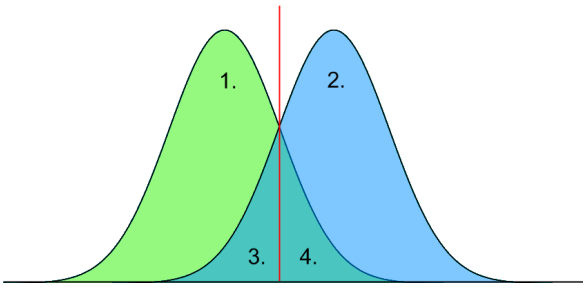

Figure 6: Intra-distances (green) and inter-distances (blue) are approximated by normal distributions. The numbered areas represent: 1. True Positive 2. True Negative 3. False Negative 4. False Positive

3.3.3 Tested Features. Features are generated in a two-step process. First, the image blocks are transformed during a preprocessing step. The resulting image blocks are then projected into a feature space of arbitrary dimension. The following lists present the image and feature space transformations that were tested:

Image transformations: Fourier transformation, structure tensor, Hessian matrix, local entropy, local fractal dimension, normalised difference vegetation index (*NDVI*)

Feature transformations: Histogram, co-occurrence matrix, autocorrelation, frequency spectrum

Some of our features were from the literature (Pratt, 2001), (Brox et al., 2006), (Soille and Rivest, 1996) and (Rouse et al., 1974).

Table 7 contains the results of the different transformation combinations in terms of their F_1-scores. Higher scores indicate better

	Histogram	Co-occurrence matrix	Autocorrelation	Frequency spectrum
Identity	0.78	0.76	0.61	0.75
Fourier transformation	0.83	0.79	0.78	0.73
Local entropy	0.68	0.67	0.52	0.71
Local fractal dimension	0.54	0.53	0.51	0.57
NDVI	0.76	0.53	0.59	0.75
Structure tensor	0.83	0.72	0.61	0.78
Hessian matrix	0.89	0.89	0.58	0.82

Table 7: Median of the F_1-scores across the training sets. Best results are emphasised.

classification results, whereby a 1.0 would imply a perfect classification. Overall, there is a significant variability between each training set, but texture based features generally generate better results. Surprisingly, the NDVI feature only works well on half the training sets, while autocorrelation results tend to be situated at the bottom of the rankings, despite the strikingly regular structure usually displayed by plantations.

4. RESULTS

Input scenes for the training sets and the plantation extraction tests were selected from IKONOS satellite images with a 1m resolution in panchromatic mode. Great care was exercised to select scenes that feature a wide variety of landscape types, and especially most plantation subtypes.

Table 8 shows the segmentation results for five test scenes, using different feature space combinations. The feature spaces used for the segmentations in Figure 9 - 11 were selected according to the results of the feature selection step in table 7. As such, a combination of three feature spaces was used:

- *Histogram from Hessian matrix*
- *Histogram from structure tensor*
- *Histogram from Fourier transformation*

The left image shows a manually generated reference mask, while the right image shows the result from our segmentation algorithm. Overall, our algorithm produces excellent recall, but only mediocre sensitivity results, which were due to plantation regions heavily expanding into background regions. This effect is emphasised if the background region contains object classes that do not appear homogeneous in the chosen feature space.

5. CONCLUSIONS AND FUTURE WORK

Our results have shown that feature spaces well suited for the extraction of specific object classes can be selected in an automatic manner. For some scenes, a segmentation quality of up to 0.94 (F_1-score) was reached. Still, the choice of a two-region segmentation scheme proved to be too restricting, since non-plantation object classes do not appear homogeneous in any single feature space that was tested. Our plan is, therefore, to update our framework to support a multi-region segmentation algorithm, allowing us to define and extract multiple object classes from a scene at the same time. Simultaneously, the catalogue of evaluated feature spaces will be expanded to get a better understanding of which feature space encapsulates the statistical properties of each object class best.

ACKNOWLEDGEMENTS

The work is funded by BKG, Frankfurt (Main) Germany. We gratefully acknowledge this support.

REFERENCES

Becker, C., Ziems, M., Büschenfeld, T., Heipke, C., Müller, S., Ostermann, J. and Pahl, M., 2008. Multi-hierarchical quality assessment of geospatial data. ISPRS.

Brox, T., Weickert, J., Burgeth, B. and Mrazek, P., 2006. Nonlinear structure tensors. Image and Vision Computing 24(1), pp. 41–55.

Chan, T. and Vese, L., 2001. Active contours without edges. IEEE Transactions on Image Processing 10(2), pp. 266–277.

Gimel'farb, G. L., 1996. Texture modeling by multiple pairwise pixel interactions. IEEE Trans. Pattern Anal. Mach. Intell. 18(11), pp. 1110–1114.

Malladi, R., Sethian, J. and Vemuri, B., 1995. Shape modelling with front propagation: A level set approach. IEEE Transaction on Pattern Analysis and Machine Intelligence 17(2), pp. 158–174.

Paragios, N. and Deriche, R., 1999. Unifying boundary and region based information for geodesic active tracking. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, IEEE Computer Society Press, Forth Collins, Colorado, pp. 300–305.

Pratt, W. K., 2001. Digital Image Processing: PIKS Inside. John Wiley & Sons, Inc., New York, NY, USA.

Rosenhahn, B., Brox, T. and Weickert, J., 2007. Three-dimensional shape knowledge for joint image segmentation and pose tracking. International Journal of Computer Vision 73(3), pp. 243–262.

Rouse, W., J., H., R., Haas, A., J., Shell, D.W., Deering, J.C. and Harlan, 1974. Monitoring the vernal advancement of retrogradation of natural vegetation. Final report, Greenbelt, MD: NASA/GSFC.

Soille, P. and Rivest, J.-F., 1996. On the validity of fractal dimension measurements in image analysis. Journal of Visual Communication and Image Representation 7(3), pp. 217–229.

Van Rijsbergen, C. J., 1979. Information Retrieval, 2nd edition. Dept. of Computer Science, University of Glasgow.

	Scene 1			Scene 2			Scene 3			Scene 4			Scene 5		
	Recall	Precision	F_1-score	Recall	Precision	F_1-score	Recall	Precision	F_1-score	Recall	Precision	F_1-score	Recall	Precision	F_1-score
1. Only presegmentation	0.85	0.65	0.74	0.20	0.97	0.33	0.46	0.65	0.54	0.88	0.59	0.70	0.92	0.59	0.72
2. Histogram from structure tensor	0.98	0.68	0.80	0.28	0.79	0.41	0.62	0.64	0.63	0.91	0.75	0.82	0.82	0.64	0.72
3. Histogram from Hessian matrix	0.98	0.69	0.81	0.90	0.82	0.86	0.84	0.61	0.71	0.96	0.66	0.78	0.92	0.60	0.73
4. Frequency spectrum from image	0.85	0.65	0.74	0.20	0.97	0.33	0.46	0.65	0.54	0.88	0.59	0.70	0.92	0.59	0.72
5. Histogram from NDVI	0.77	0.65	0.70	0.92	0.87	0.89	0.96	0.62	0.75	0.84	0.50	0.63	0.94	0.52	0.67
6. Feature combination (2) + (3)	0.98	0.68	0.80	0.91	0.80	0.85	0.73	0.64	0.69	0.93	0.72	0.81	0.89	0.63	0.74
7. Feature combination (2) + (4)	0.98	0.68	0.80	0.28	0.77	0.41	0.62	0.64	0.63	0.91	0.75	0.82	0.82	0.64	0.72
8. Feature combination (2) + (5)	0.97	0.75	0.85	0.93	0.87	0.90	0.96	0.62	0.75	0.85	0.50	0.63	0.94	0.52	0.67
9. Feature combination (3) + (4)	0.98	0.70	0.82	0.90	0.82	0.85	0.84	0.61	0.71	0.96	0.66	0.78	0.92	0.60	0.73
10. Feature combination (3) + (5)	0.94	0.76	0.84	0.92	0.87	0.89	0.96	0.62	0.75	0.85	0.50	0.63	0.94	0.52	0.67
11. Feature combination (4) + (5)	0.77	0.65	0.70	0.93	0.87	0.90	0.96	0.62	0.75	0.84	0.50	0.63	0.94	0.52	0.67
12. Feature combination (2) + (3) + (4)	0.98	0.68	0.80	0.91	0.80	0.85	0.73	0.64	0.69	0.93	0.72	0.81	0.89	0.63	0.74
13. Feature combination (2) + (3) + (5)	0.98	0.74	0.85	0.92	0.87	0.90	0.96	0.62	0.76	0.85	0.50	0.63	0.94	0.52	0.67
14. Feature combination (2) + (4) + (5)	0.97	0.75	0.85	0.93	0.87	0.90	0.96	0.62	0.75	0.85	0.50	0.63	0.94	0.52	0.67
15. Feature combination (3) + (4) + (5)	0.94	0.76	0.84	0.92	0.87	0.89	0.96	0.62	0.76	0.85	0.50	0.63	0.94	0.52	0.67
16. Feature combination (2) + (3) + (4) + (5)	0.98	0.75	0.85	0.92	0.87	0.90	0.96	0.62	0.76	0.85	0.50	0.63	0.94	0.52	0.67

Table 8: Segmentation results for five test scenes using different feature space combinations. Best results for each scene are emphasised.

Figure 9: Manual reference segmentation (left) and our fully automatic segmentation (right)

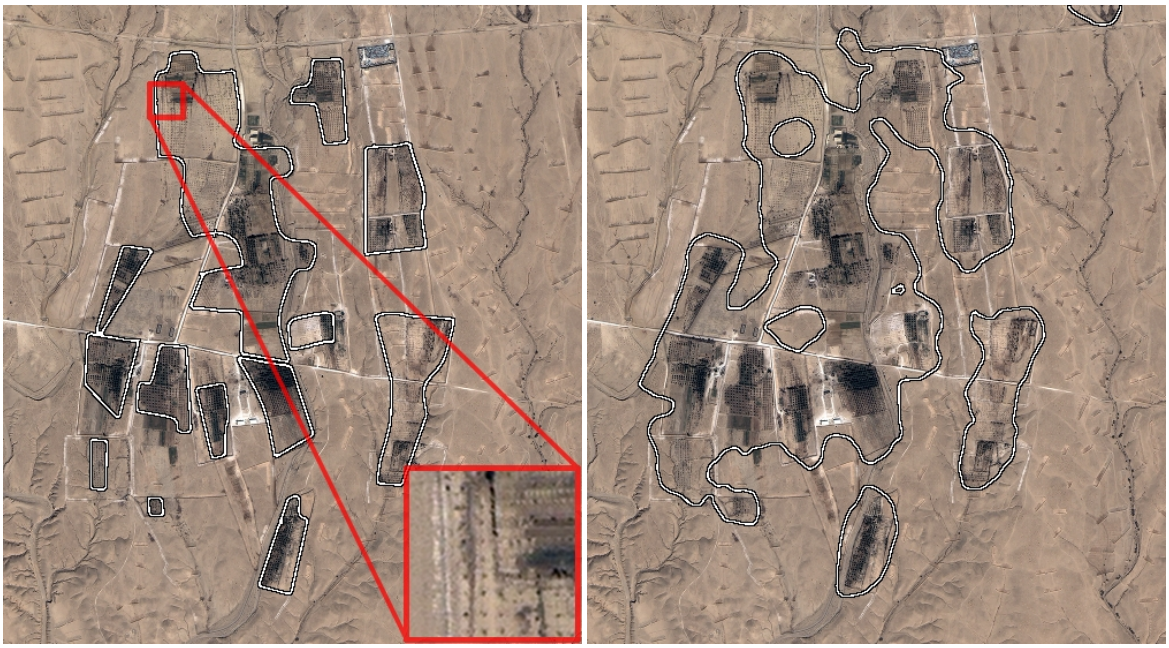

Figure 10: Manual reference segmentation (left) and our fully automatic segmentation (right)

Figure 11: Manual reference segmentation (left) and our fully automatic segmentation (right)