PARIS
LODRON
UNIVERSITY
SALZBURG

COPERNICUS MASTER
IN DIGITAL EARTH

Template Matching

POCH Sarah
Master Copernicus in Digital Earth
2021 - 2022

Abstract

The development of advanced computer vision and artificial intelligence techniques offers many possibilities but requires a lot of computing resources. In these conditions, is it really necessary to use such techniques to perform simple tasks such as counting objects?

Template Matching, a rather old computer vision technique, can be used to perform this task in a simple and resource efficient way.

Another advantage of using such a technique is to couple it with other types of analysis such as OBIA to improve the results. Indeed, one of the limitations of template matching in terms of accuracy can be reduced thanks to this.

Keywords : Template Matching, OBIA, Computer Vision, Pattern Recognition

Summary

1 Introduction

Developed since the 1980s, Template Matching is one of the simplest and oldest techniques used for object recognition and identification. Before explaining the usefulness and implementation of this method, here is a brief look back at the place of Template Matching in Artificial Intelligence (AI) and object detection.

To start at the most general level, Template Matching is a computer vision technique. Computer vision is essential in understanding systems and automating them to achieve what the human eye can do. In other words, the idea is to extract enough information from data such as images or videos to make decisions. Computer vision is useful in a variety of fields. Facial recognition is one example, there are also tumour detection in medicine and counting of objects or organisms, among others.

Within this vast field of computer vision, there is also a sub-category called pattern recognition. Several techniques are applicable in pattern recognition, with a common objective. Training an automated system to recognise one or more patterns in a specific environment and making decisions about categorising them. Template Matching is one of the techniques used in pattern recognition. Indeed, as its name indicates, the aim is to find a "match" between a certain template and an image to analyse.

First of all, the theoretical aspect of Template Matching will give a general understanding of this pattern recognition technique. Then, the application of this technique with the software eCognition (Trimble) will allow to identify and count the number of dwellings in a part of the Lukole refugee camp, in Tanzania.

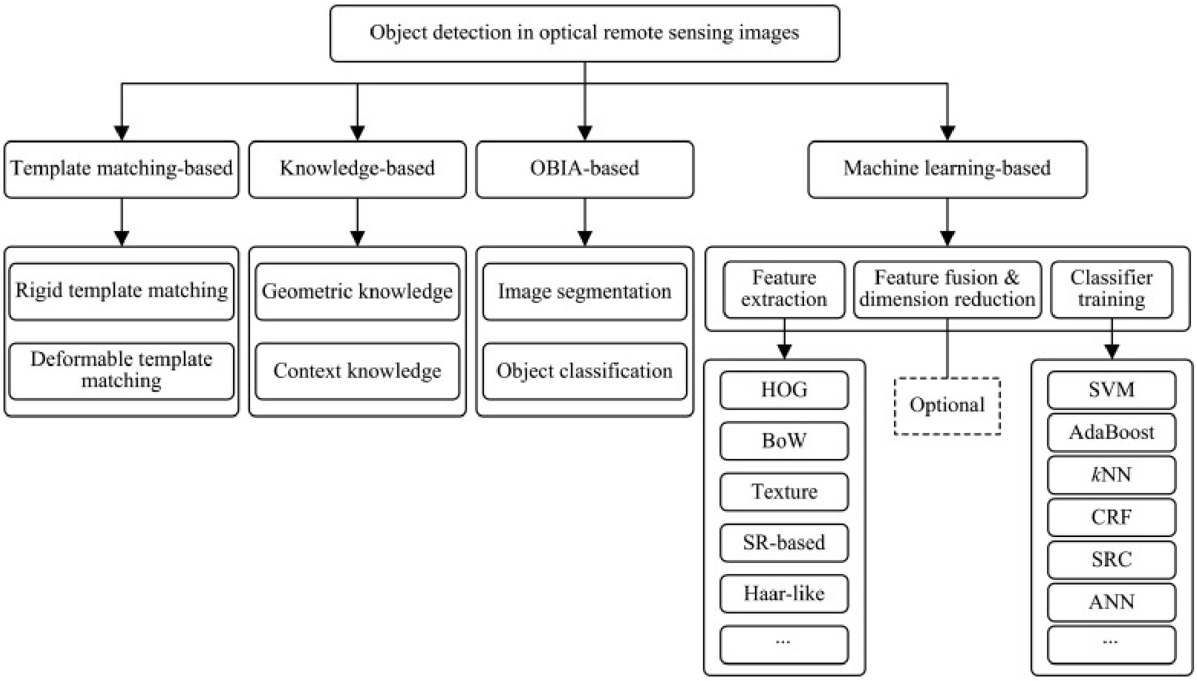

Fig. 1. Taxonomy of methods for object detection in optical RSIs. Rounded rectangles with solid borders illustrate our scope in this paper, from [2]

2 What is Template Matching ?

As mentioned earlier, Template Matching is a pattern recognition technique. The different techniques belonging to this category follow a relatively common procedure. This procedure is implemented as followed. Data acquisition and pre-processing, template generation then, the comparison between the template and the input image. The results of this comparison are then used to make decisions. This procedure, applied to the Template Matching technique, is summarised in figure 2.

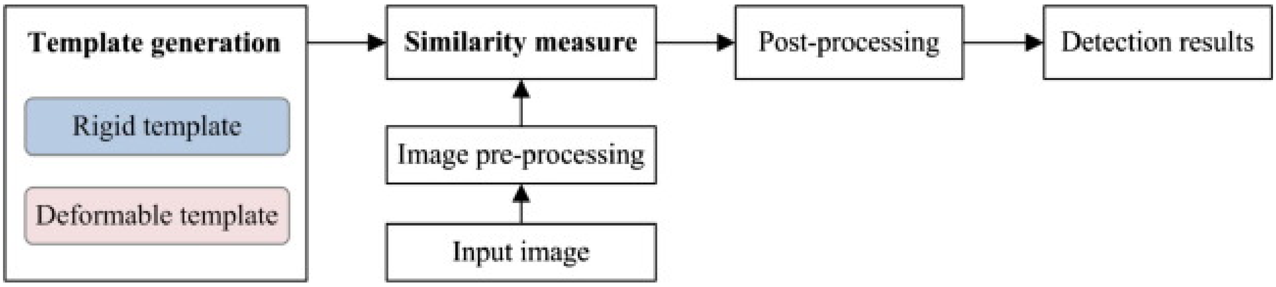

Fig. 2. The flowchart of template matching-based object detection, from [2]

In order to have a general understanding of the technique, only the template generation and the measure of similarities will be explained.

2.1 Template Generation

The essential element for the analysis of an image with the Template Matching technique is the template. This will then be compared with the image. However, what is a a template? A template is obtained from a pattern, in other words, a collection of similar objects. These objects do not have to be completely identical, they just have to have a common denominator.

Indeed, the aim is to reproduce what humans are capable of seeing and interpreting, who are quite capable of recognising and identifying the same object despite variations. These variations can be of various natures such as a change of scale, rotation, point of view, illumination, etc. One of the most common examples is the following. The human eye is able to recognise a person's face, which has fairly constant elements over time, even if the person's expression changes. But other types of variations can be observed, specific to the data acquisition, such as noise or sensor-related variations. The main difficulty remains to implement these variations and to analyse an image taking it into account.

The generation of a template is necessary for each pattern that we wish to identify. And if the latter can be generated "by hand", it is also possible to use a training set. Furthermore, as shown in figure 1, two types of templates can be generated; a rigid template and a deformable template.

2.1.1 Rigid Template

From a chronological point of view, the rigid template is the first type of template to be studied and used. The objects that this type of template can detect must be simple objects that do not experience significant internal variations. Indeed, the template, for a correct analysis must be very precise. This type of template will be used in particular for the detection of buildings or roads.

2.1.2 Deformable Template

Although the use of deformable templates is more recent, they are also more malleable and more efficient. Indeed, unlike its predecessor, the deformable template better accepts different types of variations such as internal variations or variations relating to the shape of the object. This category can also be divided in two types : free-form deformable templates and Parametric deformable templates.

	Rigid template	Deformable template
Strength	Simple and easy to implement	More powerful and flexible than rigid shape matching in dealing with shape deformations and intra-class variations
Limitations	Scale and rotation dependent - Sensitive to shape and viewpoint and change	Need more prior information and parameters of the geometrical shape for template designing - Computationnally expensive
Objects	Buildings, roads, etc.	Airplane, ship, tree crown, etc.

Table 1. Comparison of rigid and deformable template properties adapted from [2].

The choice, made by the user, between using a rigid or a deformable template will be lead by the needs and the computational capabilities of the project.

2.2 Measure of Similarities

Now that a template is available for each class to be detected, it is possible to move on to the calculation of similarities. During this step, the template will be compared to the image that we want to analyse.

The word used here is similarity. Indeed, the goal is not to find an exact match but to determine the best matches. The comparison will therefore be carried out by moving the template over the entire image surface. For a better analysis, it is also possible to take into account 3 degrees of freedom: translation, rotation and change of scale. There are several mathematical principles for calculating the similarities between the image and the template. Four of the most popular techniques will be briefly discussed here: Sum of Absolute Differences (SAD), Sum of Squared Differences (SSD), Normalized Cross-Correlation (NCC) and Euclidean Distance (ED). However, the most commonly used are the SAD and the NCC.

To illustrate these mathematical principles, the following data will be used throughout this section.

Let T = $\begin{pmatrix} 5 & 6 & 1 \\ 4 & 0 & 5 \\ 8 & 2 & 7 \end{pmatrix} \in M_3$ be the template T and I = $\begin{pmatrix} 8 & 7 & 4 & 8 & 1 \\ 1 & 5 & 3 & 2 & 2 \\ 1 & 6 & 1 & 4 & 0 \\ 0 & 0 & 2 & 6 & 7 \\ 5 & 9 & 3 & 8 & 0 \end{pmatrix} \in M_5$ the analysed image I. Moreover, in this example, only a match with the entire template will be take into account and search.

2.2.1 Sum of Absolute Difference

This technique is based on the calculation of a pixel-by-pixel difference. Each pixel in matrix A is subtracted from the value of the pixel with the same coordinates in matrix B. The sum of the absolute values of each result gives the measure of similarity between the two matrices. The lower the value obtained, the more similar the two matrices are. The following equation expresses the calculation of the SAD in a mathematical way.

$$SDA = \sum_{i=1}^{m}\sum_{j=1}^{n} |A_{i,j} - B_{i,j}| \text{ with } A, B \in M_{mn} \tag{1}$$

The result is strictly positive and belong to [0, m.n. | f |] with m and n the number of lines and columns of the matrices A and B and f the maximal absolute value that a pixel can take. To apply this equation to our data, the I matrix is subdivided into 9 sub-matrices belonging to M_3, the values obtain will be such as $SAD \in [0; 81]$. The calculation of the SAD is carried out for each sub-matrix and gives the following result matrix: $R = \begin{pmatrix} 34 & 24 & 25 \\ 35 & \boxed{17} & 29 \\ 25 & 19 & 35 \end{pmatrix}$.

The sub-matrix with the smallest SAD value is the central sub-matrix of $I = \begin{pmatrix} 8 & 7 & 4 & 8 & 1 \\ 1 & 5 & 3 & 2 & 2 \\ 1 & 6 & 1 & 4 & 0 \\ 0 & 0 & 2 & 6 & 7 \\ 5 & 9 & 3 & 8 & 0 \end{pmatrix}$.
This part of the image is therefore the one that most closely matches the template.

2.2.2 Sum of Squared Difference

As with the previous technique, the sum of squared differences technique is also based on a pixel-by-pixel difference. The difference between the two pixel values is squared and then summed. Again, the lower the value, the more similar the template is to the portion of the image being studied. The equation for calculating the SSD is as follows:

$$SSD = \sum_{i=1}^{m}\sum_{j=1}^{n} (A_{i,j} - B_{i,j})^2 \text{ with } A, B \in M_{mn} \tag{2}$$

The result is also positive and $SSD \in [0, m.n.f^2]$. In this case, the values will be such as $SSD \in [0; 729]$. Calculating the SSD between each submatrix of I and T gives the following matrix: $R = \begin{pmatrix} 158 & 90 & 121 \\ 175 & 81 & 123 \\ 115 & \boxed{59} & 175 \end{pmatrix}$.

The submatrix with the smallest SSD value is $I_{\{4,5,6\}\{2,3,4\}}$. The portion most similar to the template in I is therefore the following: $I = \begin{pmatrix} 8 & 7 & 4 & 8 & 1 \\ 1 & 5 & 3 & 2 & 2 \\ 1 & 6 & 1 & 4 & 0 \\ 0 & 0 & 2 & 6 & 7 \\ 5 & 9 & 3 & 8 & 0 \end{pmatrix}$

2.2.3 Euclidean Distance

The calculation of the Euclidean Distance between two points is commonly used and this distance can also be useful in calculating the similarities between the two matrices T and I. To obtain the Euclidean distance between the two matrices, simply calculate the square root of the sum of the squared differences obtained previously.

The value obtained is then in the following range: $[0, \sqrt{m.n.f}]$ with m the number of rows, n the number of columns and f the maximum value that can be assigned to a pixel. The equation to obtain the Euclidean distance is as follows:

$$ED = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(A_{i,j} - B_{i,j})^2} \text{ with } A, B \in M_{mn} \tag{3}$$

In the case of the comparison between T and the submatrices of I, the Euclidean distance obtained will be in the interval $[0, 27]$. The results are presented in the following matrix: $R = \begin{pmatrix} 12.57 & 9.49 & 11 \\ 13.23 & 9 & 11.09 \\ 10.72 & \boxed{7.68} & 13.23 \end{pmatrix}$.

The submatrix with the smallest value of ED is $I_{\{3,4,5\}\{2,3,4\}}$. The portion most similar to the template in I is therefore the following $I = \begin{pmatrix} 8 & 7 & 4 & 8 & 1 \\ 1 & 5 & 3 & 2 & 2 \\ 1 & 6 & 1 & 4 & 0 \\ 0 & 0 & 2 & 6 & 7 \\ 5 & 9 & 3 & 8 & 0 \end{pmatrix}$.

However, this method has a significant disadvantage as it is sensitive to noise. Indeed, the slightest pixel variation can greatly change the result of the similarity calculation and therefore the conclusion of the best match. The last method presented is more robust to this problem.

2.2.4 Normalized Cross-Correlation

Normalized Cross Correlation is a more complex technique for measuring similarities than the previous ones, which makes it more robust and less sensitive to noise. The reduction of noise reduces the number of false positives. The formula for calculating the cross-correlation $CC = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}.B_{i,j}$ is in the numerator. The denominator, on the other hand, allows for the normalisation of the result. This normalisation brings the result to a value between -1 and 1, so that the comparison of results is possible. Here, the pixel value is strictly positive, so the result will be in the range [0; 1]. The equation to obtain the correlation coefficient is as follows:

$$NCC = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}.B_{i,j}}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}^2 \sum_{i=1}^{m}\sum_{j=1}^{n} B_{i,j}^2}} \text{ with } A, B \in M_{mn} \tag{4}$$

In this case, the closer the value is to 1, the more similar the portion of the image studied is to the template. The results are displayed in the following matrix: $R = \begin{pmatrix} 0.635 & 0.807 & 0.682 \\ 0.475 & 0.806 & 0.678 \\ 0.715 & \boxed{0.887} & 0.572 \end{pmatrix}$.

The submatrix with the highest correlation coefficient is $I_{\{3,4,5\}\{2,3,4\}}$. The portion most similar to the template in I is therefore the following: $I = \begin{pmatrix} 8 & 7 & 4 & 8 & 1 \\ 1 & 5 & 3 & 2 & 2 \\ 1 & 6 & 1 & 4 & 0 \\ 0 & 0 & 2 & 6 & 7 \\ 5 & 9 & 3 & 8 & 0 \end{pmatrix}$

3 Application

Now that we have an overview of what Template Matching is and how it works, here is an example of its application. The software used in this case study is eCognition. It allows to generate or import a template and to apply a template matching algorithm. Thanks to this software, it is very easy to combine template matching with other classification techniques.

The objective will be to identify and count the number of dwellings present in a part of the Lukole refugee camp in Tanzania. The result obtained will then be combined with an object-oriented image analysis to improve the accuracy of the results and limit the number of false positives, among other things. This analysis has already been carried out following the workflow described in the figure 3 and developed in the article [12].

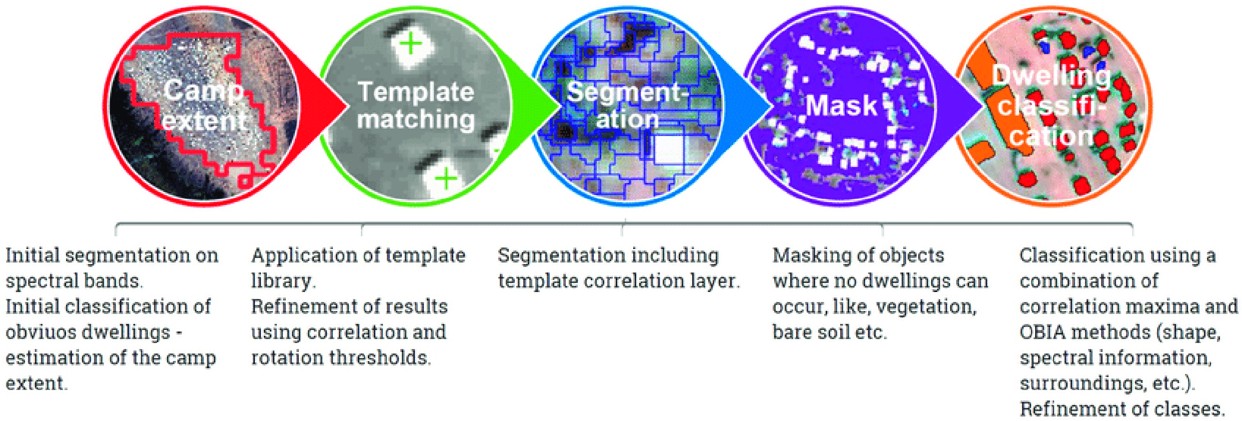

Fig. 3. Integrated workflow combining object-based image analysis (OBIA) and template matching for dwelling extraction from [12]

3.1 Study area and Data

Template Matching is a technique that can be performed with any kind of image (UAV, satellite, etc.) and the object to be identified depends on the resolution of the image to be analysed. The study area is part of the Lukole refugee camp in Tanzania. This camp has been established in 1994 in the context of the Rwanda crisis and closed in 2008.

The image used is a very high resolution image taken by the Ikonos satellite. This image was then pansharpened. It is composed of four spectral bands; blue, green, red and near infrared and its resolution is 1 metre.

Fig. 4. Pansharpened image from Ikonos of the Lukole refugee camp, in Tanzania

3.2 Template Generation

To begin this analysis, it is necessary to generate a tent template using the Template Editor, a specific tool in the software. This essential step is performed in three stages, sample selection, template generation and template testing as shown in figure 5. This process is cyclic and the steps can be repeated until a satisfactory template is obtained. The parameter that determines whether the generated template is satisfactory or not is the sample correlation coefficient and the threshold is set by the user.

Fig. 5. Worflow of the template generation with the software eCognition.

Only one type of tent will be considered, the white-roofed tents with a rectangular shape as pointed in the figure 6. However, if several types of tents were to be identified, a template for each type would be required.

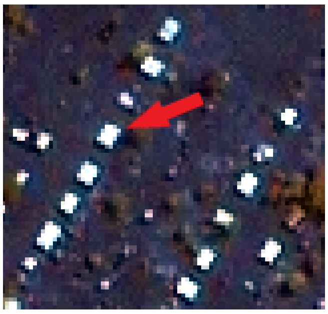

Fig. 6. Dwelling of interest for the template generation

Sample Selection

The samples selected to generate a template are, in this case, 10 pixels by 10 pixels in size and 3 pixels of context are added around the sample. Due to the rectangular shape of the tents, it is necessary to orientate the sample. In addition, as the sample size is standardised, the large tents visible on the left will not be detected, nor will the smaller ones due to the resolution of the image.

The samples will be selected only from a portion of the image. This portion is shown in figure 7.

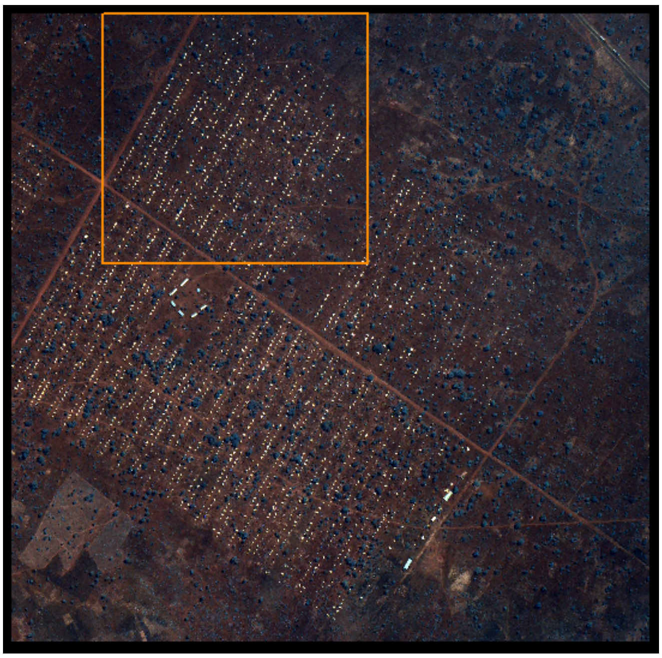

Fig. 7. Area used to select the samples.

Template Generation

With the selected samples, it is possible to generate a template. This template can be generated from each layer of the image. The generation can also be done in several ways, mean, with mask or grouped. In this case, the template is generated by means of the selected samples.

Template Testing

The Template Editor allows the user to test the generated template. The previously defined area is then analysed and compared to the template. The comparison can also be done by rotating the template. This allows the selection of new samples with a high correlation coefficient to the template. The objective of this step is to improve the template and increase its accuracy.

In this case

After performing the steps described above several times, a template with a sample correlation rate of 0.817 is obtained. This template is obtained from 110 samples selected in the search area and visible in figure 8. This template was generated on the red layer, because it is with this layer that the highest correlation rate was obtained. It is therefore on this same layer that the template matching algorithm will be applied. Applying the algorithm to another layer would then be irrelevant and false.

The template obtained in figure 9 therefore represents a rectangular tent with a white roof. It can be seen that this representation does not look very much like what can be seen on the initial image. This is partly due to the resolution of the template which is 10x10 pixels.

Fig. 8. Samples used for the template generation.

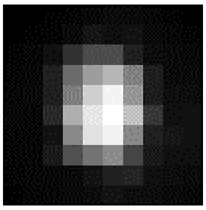

Fig. 9. Mean template generated with the red layer.

3.3 Template Matching Algorithm

Now that a template has been generated, it is possible to apply the template matching. As mentioned above, the analysis is carried out on the same layer as the one used to generate the template. Here, it is the red layer.

The method of similarity calculation is imposed by the software, it is the Normalized Cross-Correlation. The analysis takes into account the following variations: translation and rotation with a step of 10°. However, it is not possible to take into account scale variations. This implies that some large tents or smaller ones appearing in the image will not be taken into account and counted. In these cases, the best solution is to create a template adapted to these different types of tents.

Finally, the essential parameter to choose in this analysis is the minimum correlation coefficient that can exist between the template and a part of the image. The choice of 0.75, rather low compared to the coefficient of the template, will allow to have more results but also more false positives. The choice of this threshold is important because a threshold that is too low implies many false positives and a choice that is too high implies few results and many unaccounted for tents.

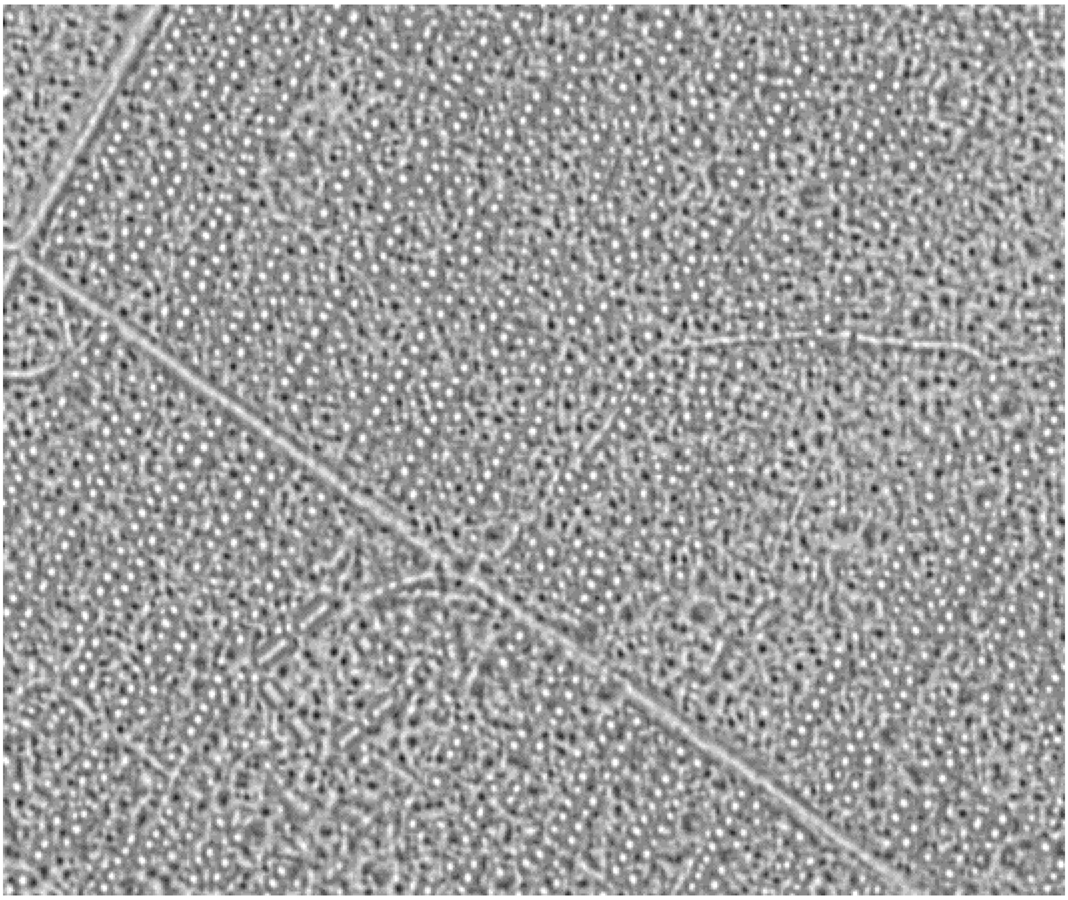

Fig. 10. Output layer obtained after application of the template matching algorithm

Two results are obtained from this analysis, visible in figures 10 and 11. Figure 10 shows the output layer indicating the areas most likely to match the template. The brighter the area, the more likely it is that a tent is in that area. It is also very easy to see the tents lined up, with circular rather than rectangular shapes.

Nevertheless, the limit of this result is quickly visible. For example, the road is very bright. Using this layer to continue the analysis would therefore require the user to quickly eliminate elements that do not have the required shape from his results.

To continue our analysis, the second result is more useful. The result, shown in figure 11, is a thematic layer. All matches found between the template and the image, having a correlation rate greater than or equal to 0.75 are indicated by a cross.

Although the road problem is no longer encountered, false positives can still be observed. For example, in the bottom left corner of the image, trees are identified as tents, among other things.

To identify these false positives, it is possible to improve the accuracy of the result by using our knowledge, which is where Object-Based Image Analysis (OBIA) comes in.

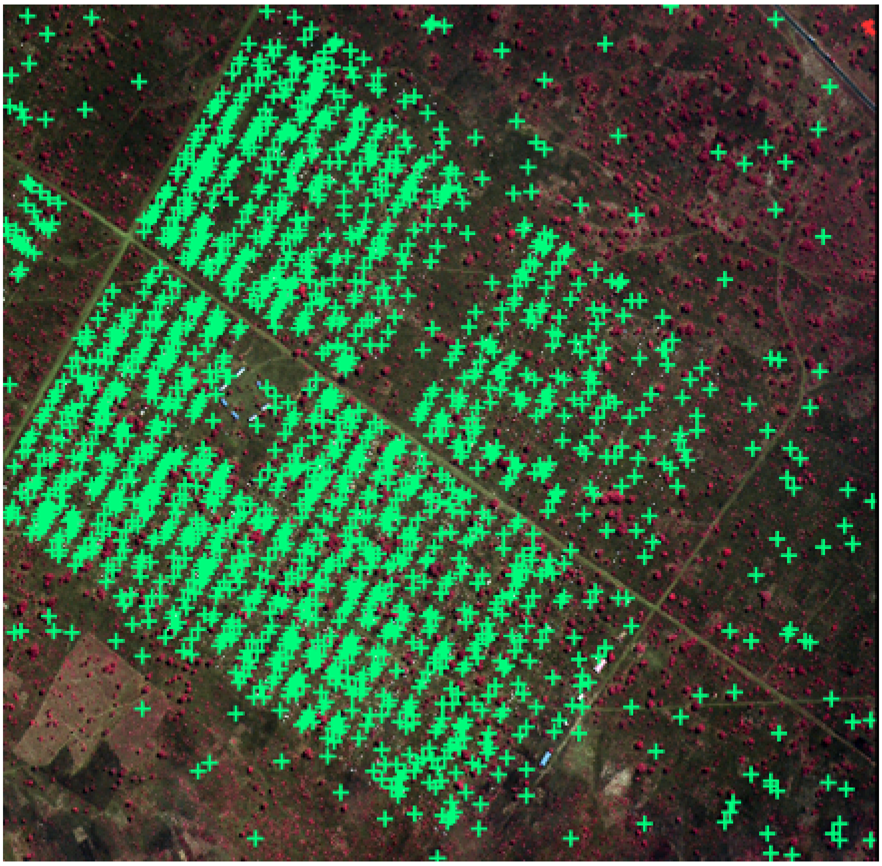

Fig. 11. Thematic layer obtained where the markers shows the template matches.

3.4 Template Matching and OBIA

As mentioned before, one of the limitations of template matching is the high number of false positives. To improve the results it is therefore interesting to combine this method with OBIA. The objective is therefore to establish rules to classify our results as tent or not, based on our knowledge. The following analysis is not very advanced and is based on the use of the Normalized Difference Vegetation Index (NDVI).

The first step is to perform a multiresolution segmentation. This makes it possible to obtain objects and classify them. Thanks to the thematic layer obtained during template matching, it is possible to classify the objects with a marker as potential tents. It is then time to use the NDVI and our knowledge. NDVI is calculated thanks to the following formula :

$$NDVI = \frac{NIR - R}{NIR + R} \tag{5}$$

The software calculates the average NDVI of each object. This then allows a threshold to be set for the classification. Since the roofs of the tents are white and not vegetation, a threshold of 0.3 is set. If the NDVI of a potential tent is below this threshold then it is a tent, otherwise it is a false positive.

The result is shown in Figure 12. Yellow are the tents classified as such and red are the false positives. These false positives can be recognised in some cases by their shape. Indeed, we can see in the upper right-hand corner that a large non-rectangular area has a marker indicating the presence of a potential tent. During the segmentation, we could already tell that this was a false positive, which was later confirmed. The tents with a white roof on the image are tents that were not identified during the template matching.

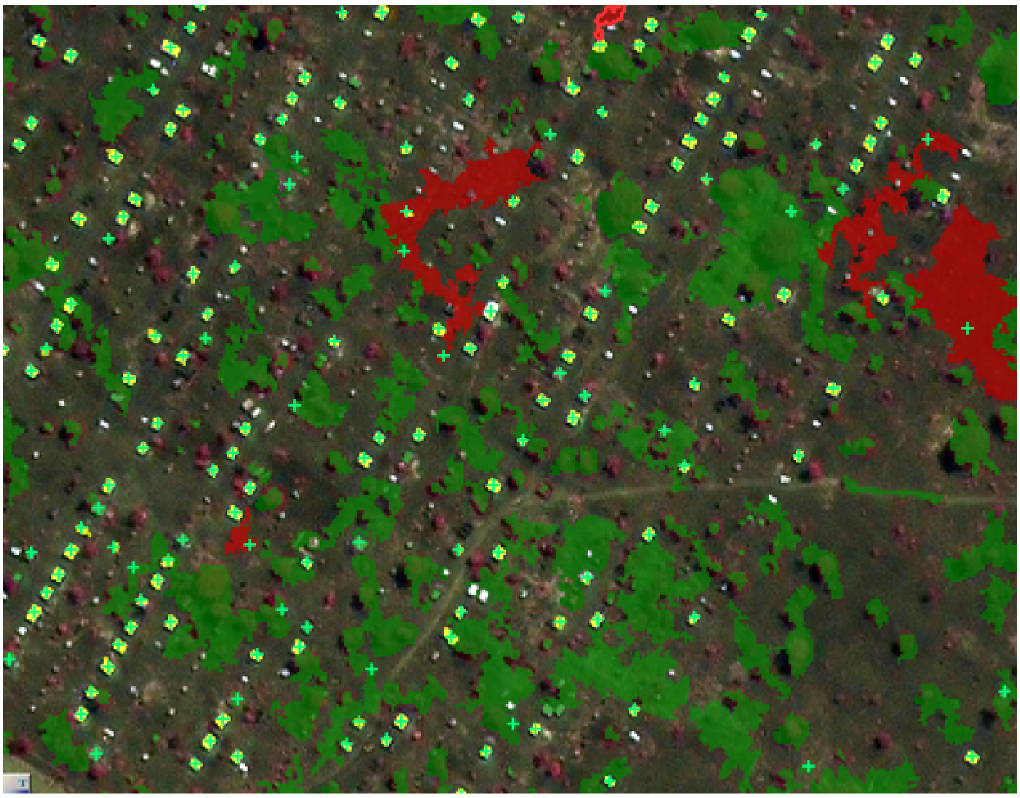

Fig. 12. Results after a short classification (in green vegetation, in red false positive and in yellow dwellings.

The analysis carried out here is not very advanced but already makes it possible to eliminate some of the false positives. Other object attributes can be used to refine the result. For example, since tents are rectangular, it is possible to search for objects with the most rectangular shape possible. The area of the object can also be used. It is also interesting to try to classify the tents that were not detected in the template matching phase. For this the shape, area and color/brightness of the roof can be used.

4 Conclusion

In conclusion, template matching is a pattern recognition technique particularly adapted to object counting. The subject of study here is the counting of tents in a part of the Lukole refugee camp in Tanzania.

Template matching, although a technique dating back to the 1980s, is still useful and continues to be developed, particularly in the field of deformable templates. Moreover, it is a technique that requires few computational resources compared to other computer vision techniques. And it is complementary to other analysis techniques such as OBIA.

References

[1] Roberto Brunelli. *Template matching techniques in computer vision: theory and practice*. John Wiley & Sons, 2009.

[2] Gong Cheng and Junwei Han. A survey on object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 117:11–28, jul 2016.

[3] L. Di Stefano, S. Mattoccia, and M. Mola. An efficient algorithm for exhaustive template matching based on normalized cross correlation. pages 322–327, 2003.

[4] eCognition TV. ecognition webinar, counting features with template matching.

[5] S. Giada, T. De Groeve, D. Ehrlich, and P. Soille. Information extraction from very high resolution satellite imagery over lukole refugee camp, tanzania. *International Journal of Remote Sensing*, 24(22):4251–4266, 2003.

[6] M.B. Hisham, Shahrul Nizam Yaakob, R.A.A Raof, A.B A. Nazren, and N.M. Wafi. Template matching using sum of squared difference and normalized cross correlation. In *2015 IEEE Student Conference on Research and Development (SCOReD)*, pages 100–104, 2015.

[7] Calvin Hung, Mitch Bryson, and Salah Sukkarieh. Multi-class predictive template for tree crown detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68:170–183, 2012.

[8] Anil K Jain, Yu Zhong, and Marie-Pierre Dubuisson-Jolly. Deformable template models: A review. *Signal Processing*, 71(2):109–129, 1998.

[9] Arie Nakhmani and Allen Tannenbaum. A new distance measure based on generalized image normalized cross-correlation for robust video tracking and image recognition. *Pattern Recognition Letters*, 34:315–321, 02 2013.

[10] J.N. Sarvaiya, Suprava Patnaik, and Salman Bombaywala. Image registration by template matching using normalized cross-correlation. pages 819–822, 2009.

[11] Priyanka Sharma and Manavjeet Kaur. Classification in pattern recognition: A review. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(4), 2013.

[12] Dirk Tiede, P. Krafft, Petra Füreder, and Stefan Lang. Stratified template matching to support refugee camp analysis in obia workflows. *Remote Sensing*, 9, 04 2017.

[13] J. Weber and S. Lefèvre. Spatial and spectral morphological template matching. *Image and Vision Computing*, 30(12):934–945, 2012.