

REMOTE SENSING IN THE FIGHT AGAINST ENVIRONMENTAL CRIMES: THE CASE STUDY OF THE CATTLE-BREEDING FACILITIES IN SOUTHERN ITALY

FRANCESCO GARGIULO¹, CESARIO VINCENZO ANGELINO¹, LUCA CICALA¹,
GIUSEPPE PERSECHINO¹ & MASSIMILIANO LEGA²

¹CIRA, the Italian Aerospace Research Center, Capua, Italy.

²Department of Engineering, University “Parthenope” Naples, Italy.

ABSTRACT

Enforcement of environmental regulation is a persistent challenge and timely detection of the violations is key to holding the violators accountable. The use of remote sensing data is becoming an effective practice in the fight against environmental crimes. In this work, a novel and effective approach for the detection of potentially hazardous cattle-breeding facilities, exploiting both synthetic aperture radar and optical multispectral data together with geospatial analyses in the geographic information system (GIS) environment, is proposed. Experiments on data available for the area of Caserta (Southern Italy), show that the proposed technique provides very high detection capability, up to 90%, with a acceptable false alarm rate, becoming a useful tool in the hand of agencies engaged in the protection of territory.

Keywords: environmental hazard detection, GIS, image analysis, multisensor data fusion, multispectral images, remote sensing, SAR.

1 INTRODUCTION

Enforcement of environmental regulation is a persistent challenge, and timely detection of violations is key to holding violators accountable. Continuous and targeted monitoring allows for efficient police action [1]. The operational environmental surveillance for the successful identification and prosecution of pollution culprits requires an integrated system based on data from several sources including space, air, waterways, and land monitoring [2].

Environmental forensic studies can produce complex data that are difficult to represent, and problems of great importance and impact on society, like those related to the waste cycle, have received no much attention. The main reason is that they call for a more innovative use of remote sensing tools, which integrate various data sources and different competencies. For these purposes, the use of Geographic Information Systems (GIS) is steadily increasing thanks to the availability of a large bulk of remote-sensing imagery, both in optical and radar modalities.

A first step in this direction is found in [3] where information provided by a ground positioning system (GPS) are integrated in a GIS to analyze source, extent and transport of contaminants. More recent papers rely heavily, and in various ways, on remote sensing imagery, as in [4], where the authors propose a method for the identification of uncontrolled landfills by means of multiresolution Ikonos data. In [5] a GIS-assisted system for hazardous waste site monitoring, based on the integration of multispectral and lidar data with numerous types of thematic information is proposed. [6] reviews the literature of remote sensing and overhead imaging in the context of hazardous waste and discusses future trends with special

attention to multispectral and hyperspectral remote sensing data. [7] uses hyperspectral images to characterize vegetation at hazardous waste sites, with different analysis methods (vegetation indices, red-edge positioning, and machine learning).

There are several other remote sensing data products that can be used in the environmental monitoring, including fusion of optical data with synthetic aperture radar data to detect cattle ranching [8] and use of thermal imagery to monitor landfills [9] and detect illegal dumping [10]. In addition, remote sensing data can be strategically combined with other data layers in GIS to monitor the vulnerability of cultural sites [11] and anticipate environmental violations [1, 2].

In this paper, we propose a new methodology, using optical and SAR remote sensing data, for the detection of small cattle breeding areas, potentially responsible of hazardous littering. This information is then used in a

GIS-based processing workflow to identify new facilities of the same type that are unknown to the official census. Experiments on a test area, with available specific ground truth, prove that the proposed detection and recognition process is characterized by very large detection capability and a negligible false alarm rate. A lot of unregistered buffalo breeding farms can be successfully identified, by inspecting only a small portion of the area of interest.

2 ENVIRONMENTAL HAZARDS RELATED TO BUFFALO BREEDING

Pollutants from manure, litter, and process wastewater can heavily affect human health and the environment. More than 40 diseases found in manure can be transferred to humans, including the causative agents for Salmonellosis, Tuberculosis, and Leptospirosis. Exposure to waterborne pathogen contaminants can result from both recreational use of affected surface water (accidental ingestion of contaminated water and dermal contact during swimming) and from ingestion of drinking water derived from either contaminated surface water or groundwater.

This specific environmental problem is very relevant to the Caserta area, in southern Italy. In fact, the Caserta area is one of the main production sites of the “Mozzarella di Bufala Campana” the worldwide-famous fresh cheese holding the status of a protected designation of origin under the European Union. In 2006 Campania produced 34.000 t of mozzarella, about 80% of national production.

Concerning buffalo breeding facilities (BBFs), many small factories exist which are not even known to the productive activity register, and are not easily monitored and surveyed. Attention on this difficult census has been raised by many recent cases of pollution due to illicit spills involving BBFs reported in the course of inspections carried out by the Forestry members in collaboration with the local agencies in Campania. In particular, as described in many papers and news, these investigations made clear that some holders did not properly accumulate and download all heaps of manure, but several cubic meters had been downloaded in a few square meters.

This is in open violation of the established specific rules, indeed wastewaters cannot spread on the soil. Manure cannot be accumulated in a small area as this represents a serious source of pollution, a bad habit that becomes a serious danger when the BBFs are located in proximity of rivers or urban centers.

The main goal of this project is the detection of BBFs in a specific area of Caserta province, not included in the official census. The area of interest is the catchment area of the river Agnena and surrounding areas as shown in Fig. 1. The Agnena is a river of about 30 km which crosses 8 municipalities in the province of Caserta. In particular, the area of study is that within a buffer of 10 km around Agnena, that is an area of about 300 square kilometers.

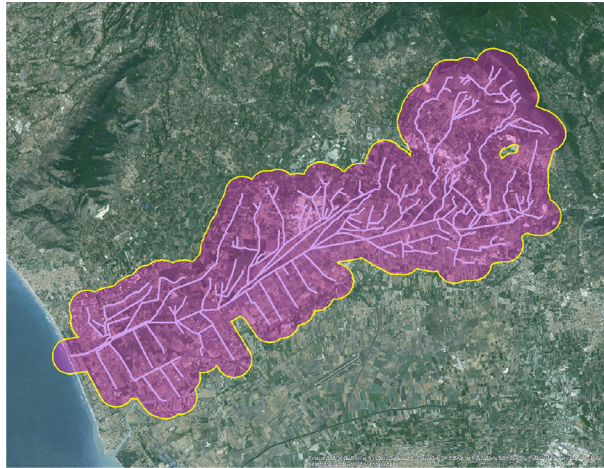


Figure 1: The area of interest around the Agnena river.

Because of the considerable extension of the area of interest, there is a need to establish an in situ verification strategy based on priority.

3 PROPOSED APPROACH

The adopted approach is inspired by the Johnson's criteria [12] and it includes the following phases:

- **Detection:** The goal of this phase is to determine if an object is present. In particular we search, in the remote sensed images, for areas to which objects belong that are similar to parts of a typical buffalo farm.
- **Recognition:** In this phase the type object can be discerned. In particular an expert human photointerpreter can discriminate the buffalo farms among the detected structures.
- **Identification:** A specific object can be discerned in the identification phase. In particular the recognized buffalo farms are compared to the official register, in order to identify the unregistered breeding activities.

Finally, a strategy is formulated in order to physically inspect the unregistered BBFs based on a vulnerability analysis. Indeed, the aim of the presented approach is reducing the need of a large area in-situ inspection, while guaranteeing a level of reliability suitable for real world applications such as environmental monitoring and/or environmental forensics. This is a requirement for typical end users (prefecture, environmental police, prosecutor's office, etc.), who have been interviewed for the purpose of this work. In particular, they ask for an analysis with the following requirements:

- known degree of reliability;
- low cost of the needed imagery;
- high repeatability of the process, with low costs in the preprocessing and design phases.

The first consideration leads to the definition of semi-automated process. The use of automated workflow is limited to the detection and identification, while the phase of recognition

is entrusted to a human photointerpreter. In addition, the tuning of the detection system is still made by a human operator.

The second necessity leads to acquire multispectral data on a single date, while the approach proposed in [8] relies on a temporal analysis on two dates. This choice reduces by half the cost, though with a significant degradation in the precision of the detection (very high false alarm rate). Instead, the use of SAR data has not be limited compared to [8], because our end users, as an Italian institution, have free access to historical series Cosmo Sky-Med Stripmap MapItaly program, which maintains a comprehensive coverage of the Italian territory every 16 days.

The last requirement of the end user, convinced us to opt for binary decision trees, one for each of the two materials to locate, instead of maximum likelihood classifiers based on Gaussian model and operating on many classes, as proposed in [8]. This choice reduces the design efforts of the classifiers used during detection and it does not compromise the performance.

With respect to previous works of the same authors on the same theme, the whole process is discussed. Moreover, the proposed solutions are a practical trade-off between performance optimization and operative needs, such as reducing the costs of the needed remote sensed data and simplifying the design of the detection algorithms.

4 PROCESSING WORKFLOWS

4.1 Detection workflow

In a BBF the following two distinguishing features are usually present: (i) outdoor areas dedicated to the breeding of buffaloes (fences and containments of manure) that are typically covered by animal waste, (ii) the metal roofs of the barns as shown in Fig. 2.



Figure 2: Typical appearance of a buffalo breeding farm. Yellow boxes represent the fenced the fenced areas, typically covered with manure, while Green boxes represent metal roofs.

Based on this a priori, the detection workflow searches for areas having spectral characteristics similar to such structures. Unfortunately, the spectral responses of these types of objects may be confused with other types of materials.

We used a decision tree classifier which performs multistage classifications by using a series of binary decisions to place pixels into classes. The parameters are tuned in order to achieve the desired probability of detection, without care of the false alarm rate. Using decision trees, it is possible to locate areas with spectral responses very similar to those ones with metal roofs or with manure. A further filtering of such zones, based on their geographical area (typically lesser than 8,000 m² for roofs and lesser than 800 m² for manure), reduces false alarms without compromising the chances of correct detection.

It has been made possible a substantial reduction of false alarms through an analysis of proximity between roofs and manure detectors. In particular, the areas distant from each other more than 50 m have been deleted (see Fig. 3c). In order to further reduce false alarms, highly urbanised areas have been removed. An updated map that identifies highly urbanised zones was obtained with SAR data by means the methodology described in [8], based on SAR interferometry.

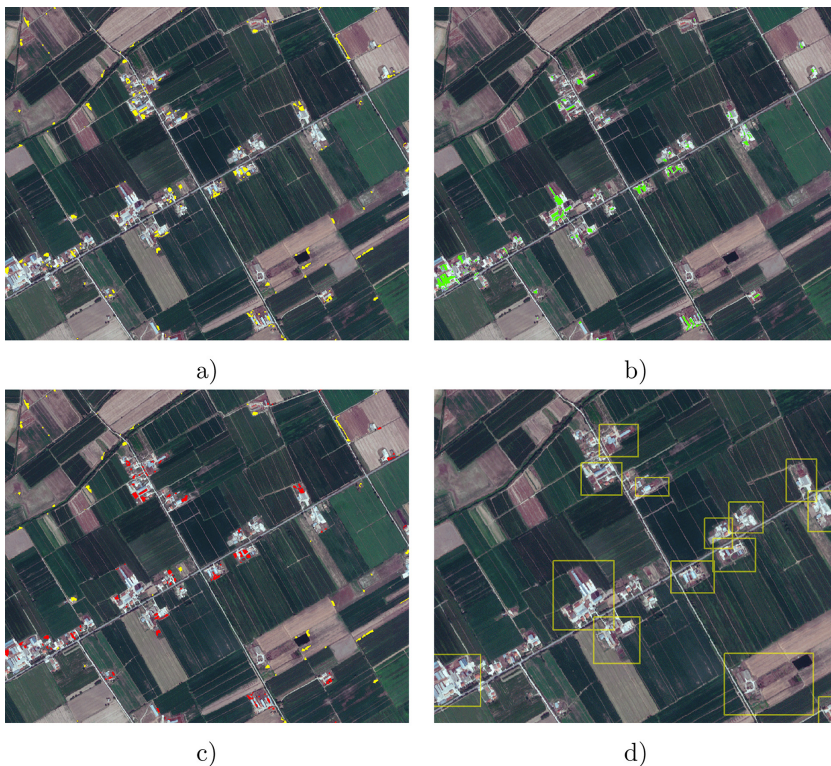


Figure 3: Different phases of the proposed workflow. (a) Detection of manure areas. (b) Detection of metal roofs. (c) Filtering of the detected manure areas performing a proximity analysis to metal roofs (only red manure detectors are kept). (d) Recognition: results of photo-interpretation (yellow boxes).

4.2 Recognition and identification process

The process of photointerpretation was accomplished only in a limited area around the detectors (about 150 meters). In this way, it was possible to visually inspect a much smaller portion of territory. The inspection took place on GeoEye-1 image after pansharpening, which increased spatial resolution to 0.5 m per pixel as showed in Fig. 3d.

The identification of the farms, previously recognized through the process of photointerpretation, is carried out automatically by means of a geospatial analysis in a GIS environment. In particular, each farm recorded in Province of Caserta is associated with the closest farm among those recognized in the satellite image. If a recognized farm does not match with any entry in the database, it is marked as unregistered.

4.3 Strategy for in situ verification

The last step is the creation of a predicted vulnerability map of the area of interest. Firstly, an index V , that assigns to each recognized farm $r \in R$ a value $V(r)$ representing the vulnerability of r , is defined, and then a geostatistical model based on V is elaborated. The result is a vulnerability prediction classification map, where each class is associated with a range of vulnerability values. In particular, the vulnerability $V(r)$ is defined as follows:

$$V(r) = \frac{|I(r)|}{dist(r)} \quad (1)$$

where $I(r)$ is a neighborhood of r with radius δ , $|I(r)|$ represents the number of detections contained in $I(r)$ and $dist(r)$ is the distance of r from the catchment area of the Agnena river. The value of $V(r)$ decreases if r moves away from a catchment area while it increases as the concentration of recognized buffalo breeding farms grows. In order to avoid null or indeterminate values the following assumption were made:

1. $r \in I(r)$ then $|I(r)| \neq 0$;
2. if $dist(r) = 0$ then $dist(r) = \min_r dist(r)$.

5 EXPERIMENTAL RESULTS

In this case study, we used a GeoEye-1 multispectral image taken in 2012 together with a time series of SAR images spanning the period 2009–2012. These latter, coming from the Italian Cosmo Sky-Med coverage program MapItaly. The GeoEye-1 multispectral image has a set of four (blue-green-red-NIR) spectral bands with 2 m/pixel of ground resolution and a panchromatic band with 0.5 m/pixel of resolution, while the Cosmo Sky-Med images are in Stripmap mode with a ground resolution of 3 m/pixel.

5.1 Performance of the detection workflow

In order to achieve a good trade-off between performance and complexity, we tested three different workflows, of incremental complexity, defined as follows:

1. A workflow without the SAR (to filter false alarms in high density urban areas) and neither the proximity analysis;
2. A workflow without the SAR with the proximity analysis;
3. The complete workflow as described in the previous section.

The performance of the automatic detection system was assessed in terms of precision P and recall R. These measures are defined, with reference to a generic target class T, as

$$P = \Pr(c = T | \hat{c} = T) \tag{2}$$

$$R = \Pr(\hat{c} = T | c = T) \tag{3}$$

where Pr stands for probability, c and \hat{c} indicate the true and the selected class/hypothesis. A high precision indicates that when the target class is detected, the decision is very likely correct. A high recall indicates that when the target class is present, it will be very likely detected. Therefore, both measures are desired to be large, and, by tuning the classification parameters, one may increase one of them but decrease the other.

The performances of the three workflows are reported in Table 1. It is worth noting that the proximity analysis between roofs and manures significantly reduces false alarms (improvement of precision of +9.93%), even if it slightly reduces the performance in terms of recall (-6.92%).

As showed in [8], masking highly dense urban areas, in which BBFs are obviously not present, improves the precision (+3.39%), without degradation of the recall (see Table 1). Moreover, the use of such a mask can significantly reduce the working area to simplify the recognition task. However, this marginally applies to the tested area, because of the low urban density around the Agnena river (see Table 2).

The overall results, in terms of F-measure, defined as $\frac{2PR}{P+R}$, are significantly worse than those reported in [8], in which a persistence analysis on two dates instead of a proximity analysis on a single date was proposed, although a direct comparison between the two methods in the same area is not reported in this paper. However, the processing workflow proposed in this paper strongly reduce the quantity of multispectral images to be used and does not need of a very accurate coregistration process between two dates of optical images. This matches the typical user requirements of low cost and low design effort workflows. The use of Cosmo Sky-Med SAR data is free for the Italian public institutions, so that does not add costs to the process.

Table 1: Detection performance with different variants of the proposed procedure.

Workflow	Proximity	Masking	Precision (%)	Recall (%)	F-measure
1	NO	NO	12.62	96.23	0.20
2	YES	NO	22.55	89.31	0.23
3	YES	YES	25.95	89.31	0.26

Table 2: Detection performance with different variants of the proposed procedure. The total area is about 295 km².

Workflow	Proximity	Urban mask	Inspection area	Saving (%)	Recall (%)
1	NO	NO	168 km ²	43.01	99.69
2	YES	NO	73 km ²	75.28	95.60
3	YES	YES	69 km ²	76.69	95.60

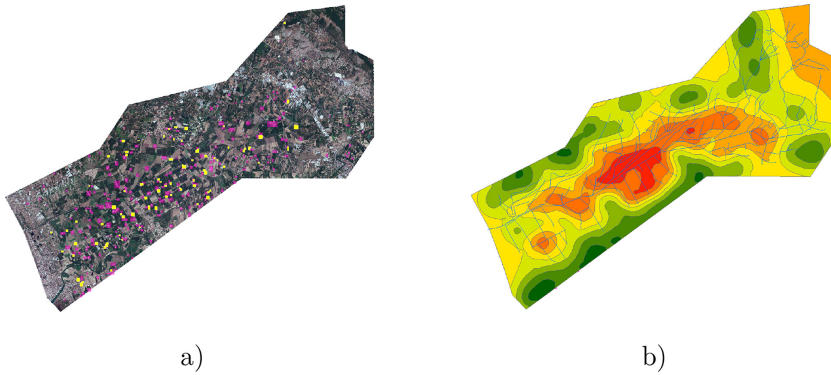


Figure 4: (a) Unregistered buffalo farms (yellow boxes) in the whole area of interest. (b) Predicted Vulnerability Map. Reddish colors indicate areas with increasing vulnerability.

5.2 Complexity of the recognition process

The aim of this approach is to reduce the work of a photointerpreter in the recognition phase. Starting from the detected areas, the suggested approach is to classify the detectors (confirming or rejecting them) and then to visually inspect a restricted area in a range of 150 meters all around the detectors. This latter step is effective because cattle-breeding farms form spatial clusters.

The amount of reduction of the work of the photo-interpreter can be seen comparing the inspected area w.r.t. the whole area of interest. The percentage of reduction is reported in the column “saving” in Table 2. The use of the third workflow (proximity and masking filters) leads to a reduction of 76.69% of the inspected area with a recall of 95.60%. A better recall (99.69%) can be achieved using the first workflow (no false alarms filtering), but in this case the inspected area is reduced only of 43.01%.

5.3 Results of the identification process

To test the identification process, the recognized BBFs have been compared to an official database (updated to 2006). As a result, the 27% of the recognized BBFs were not registered in the database (see Fig. 4a).

Finally, a vulnerability map was produced, as explained in Section 4.3. The value chosen for the radius is $S = 500$ m, while the interpolation is the ordinary Kriging method. The resulting predicted vulnerability map, with 30 classes, is depicted in Fig. 4b, where reddish colours indicate areas with increasing vulnerability.

6 CONCLUSIONS

This paper proposed a novel and effective methodology for the detection of BBFs, using a single multispectral image, SAR data and previously collected vector data layers. The proposed process, unlike [8] requires a lower budget for data acquisition. Moreover, another time demanding preprocessing and algorithm design tasks, as like as fine images coregistration and many classes maximum likelihood training are skipped. This comes at the price of a lower performance of the detection algorithms in terms of false alarm rate. However, it significantly reduces the workload of the photo-interpreter in the recognition phase.

The performance of the overall proposed process is quite satisfactory, showing that it can be a valuable tool in contrasting environmental crimes with the continuous interaction with the institutions.

REFERENCES

- [1] Lega, M. & Napoli, R., Aerial infrared thermography in the surface waters contamination monitoring. *Desalination and Water Treatment*, **23**(1–3), pp. 141–151, 2010.
<http://dx.doi.org/10.5004/dwt.2010.1988>
- [2] Lega, M., Kosmatka, J., Ferrara, C., Russo, F., Napoli, R. & Persechino, G., Using advanced aerial platforms and infrared thermography to track environmental contamination. *Environmental Forensics*, **13**(4), pp. 332–338, 2012.
<http://dx.doi.org/10.1080/15275922.2012.729002>
- [3] Brilis, G., van Waasbergen, R., Stokely, P. & Gerlach, C., Remote sensing tools assist in environmental forensics: part ii - digital tools. *Environmental Forensics*, **2**(3), pp. 223–229, 2001.
<http://dx.doi.org/10.1006/enfo.2000.0033>
- [4] Silvestri, S. & Omri, M., A method for the remote sensing identification of uncontrolled landfills: formulation and validation. *International Journal of Remote Sensing*, **29**, pp. 975–989, 2008.
<http://dx.doi.org/10.1080/01431160701311317>
- [5] Jensen, J., Hodgson, M., Garcia-Quijano, M., Im, J. & Tullis, J., A remote sensing and gis-assisted spatial decision support system for hazardous waste site monitoring. *Photogramm Eng Remote Sensing*, **75**, pp. 169–178, 2009.
<http://dx.doi.org/10.14358/PERS.75.2.169>
- [6] Slonecker, T., Fisher, G., Aiello, D. & Haack, B., Visible and infrared remote imaging of hazardous waste: A review. *Remote Sensing*, **2**, pp. 2474–2508, 2010.
<http://dx.doi.org/10.3390/rs2112474>
- [7] Im, J., Jensen, J., Jensen, R., Gladden, J., Waugh, J. & Serrato, M., Vegetation cover analysis of hazardous waste sites in utah and arizona using hyperspectral remote sensing. *Remote Sensing*, **4**, pp. 327–353, 2007.
<http://dx.doi.org/10.3390/rs4020327>
- [8] Errico, A., Angelino, C.V., Cicala, L., Persechino, G., Ferrara, C., Lega, M., Vallario, A., Parente, C., Masi, G., Gaetano, R., Scarpa, G., Amitrano, D., Ruello, G., Verdoliva, L. & Poggi, G., Detection of environmental hazards through the feature-based fusion of optical and sar data: a case study in southern italy. *International Journal of Remote Sensing*, **36**(13), pp. 3345–3367, 2015.
<http://dx.doi.org/10.1080/01431161.2015.1054960>
- [9] Lega, M. & Napoli, R., A new approach to solid waste landfills aerial monitoring. *WIT Transactions on Ecology and the Environment*, 109, pp. 193–199, 2008.
<http://dx.doi.org/10.2495/WM080211>
- [10] Persechino, G., Lega, M., Romano, G., Gargiulo, F. & Cicala, L., Ides project: An advanced tool to investigate illegal dumping. *WIT Transactions on Ecology and the Environment*, 173, pp. 603–614, 2013.
<http://dx.doi.org/10.2495/SDP130501>
- [11] Lega, M. & Persechino, G., Gis and infrared aerial view: Advanced tools for the early detection of environmental violations. *WIT Transactions on Ecology and the Environment*, 180, pp. 225–235, 2014.
<http://dx.doi.org/10.2495/WM140191>
- [12] Harney, R., *Combat systems vol. 1: Sensors*, 2005.