



Article

An Object-Oriented Approach to the Classification of Roofing Materials Using Very High-Resolution Satellite Stereo-Pairs

Francesca Trevisiol , Alessandro Lambertini , Francesca Franci and Emanuele Mandanici *

Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, Viale del Risorgimento 2, 40136 Bologna, Italy; francesca.trevisiol2@unibo.it (F.T.); alessandro.lambertini@unibo.it (A.L.); francescafranci@live.it (F.F.)

* Correspondence: emanuele.mandanici@unibo.it

Abstract: The availability of multispectral images, with both high spatial and spectral resolution, makes it possible to obtain valuable information about complex urban environment, reducing the need for more expensive surveying techniques. Here, a methodology is tested for the semi-automatic extraction of buildings and the mapping of the main roofing materials over a urban area of approximately 100 km², including the entire city of Bologna (Italy). The methodology follows an object-oriented approach and exploits a limited number of training samples. After a validation based on field inspections and close-range photos acquired by a drone, the final map achieved an overall accuracy of 94% (producer accuracy 79%) regarding the building extraction and of 91% for the classification of the roofing materials. The proposed approach proved to be flexible enough to catch the strong variability of the urban texture in different districts and can be easily reproducible in other contexts, as only satellite imagery is required for the mapping.



Citation: Trevisiol, F.; Lambertini, A.; Franci, F.; Mandanici, E. An Object-Oriented Approach to the Classification of Roofing Materials Using Very High-Resolution Satellite Stereo-Pairs. *Remote Sens.* **2022**, *14*, 849. <https://doi.org/10.3390/rs14040849>

Academic Editors: Saeid Homayouni, Ying Zhang and Ali Mohammadzadeh

Received: 22 December 2021

Accepted: 5 February 2022

Published: 11 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: building extraction; roofing classification; WorldView-3; OBIA classification; machine learning; smart city

1. Introduction

The WorldView-3 (WV3) satellite platform offers unprecedented capabilities of very high spectral and spatial resolution. It provides a 0.3 m resolution panchromatic sensor, an eight-band visible and near-infrared imager (operating between 397 and 1039 nanometres) with a ground sample distance (GSD) of 1.2 m, a short-wave infrared (SWIR) sensor with further eight bands (1184–2373 nm) at a resolution of 3.7 m, and finally the CAVIS (Clouds, Aerosol, Vapors, Ice, and Snow) sensor with twelve additional bands at 30 m for the retrieval of atmospheric properties [1]. These capabilities make it particularly suitable for application in urban areas, where the extremely complex morphology requires a higher level of detail [2,3].

The present paper proposes an approach for the detection of buildings and the classification of the major roofing material types entirely based on the satellite WV3 data, without any ancillary information. This method is tested over an area of approximately 100 km², covering an entire city, and aims to be easily repeatable in most of the urban contexts, as the only required data are the satellite images.

Automatic building detection from remote sensing has been widely debated in the scientific literature and a detailed review is beyond the scope of the present paper. Readers are referred to recent review papers [4–6]. Here, only contributions where the classification of roofing materials is considered (alongside the problem of roof detection) are briefly discussed. This problem has been addressed for several purposes, encompassing building inventory for earthquake damage assessment [7], identification of hazardous materials such as asbestos [8,9], and for thermal flux or energy efficiency studies that require proper determination of surface emissivity in a smart city perspective [10,11], among others.

It is to be noted that very often the published tests are performed over very limited areas (typically a few square kilometers) [12–14], with relatively homogeneous urban texture. From the point of view of the main data sources, two approaches were proposed in the literature: the first relies only on images, which can be multispectral [8,9,15] or hyperspectral [16–18], and the second combines images with a digital elevation model of the urban area [14,19]. When included in the processing, the contribution of the elevation model becomes dominant especially for the identification of roof surfaces (by excluding all the other urban elements), while spectral information is used for material identification [19]. Although some spectral libraries of building materials were developed recently [20,21], classification approaches based on training samples or on decision rules are often preferred, especially with multispectral data [13,19].

When working on images only, the classification process is based on the spectral content and requires both a high spatial and spectral resolution. Moreira and Galvão [16] found that near and short-wave infrared bands provide the best discrimination, when using MIVIS data and field spectral signatures as training for the spectral angle mapper classifier. Better results in terms of overall accuracy (up to 97%) were achieved with machine learning algorithms, such as support vector machine (SVM) and random forest [17], or convolutional neural network [22].

A recent work by Kim et al. [23] implemented deep learning algorithms using aerial images with four bands only, but it was limited to a few materials (concrete and metal) and required an extensive training dataset (more than 11,600 images). Usually, when working with multispectral sensors (with a more limited number of bands compared to hyperspectral ones), images are to be coupled with elevation information and an object-oriented approach is to be preferred, in order to achieve comparable accuracy levels [12,15]. Abriha et al. [15] also tested the application of pan-sharpening to WorldView-2 imagery, to investigate possible benefits in terms of accuracy, and found improvements by 2–3% with some drawbacks in shadowed areas.

In all the proposed approaches multispectral images and elevation data come from different sources; usually, digital elevation models come from LiDAR surveys, which are not available everywhere and are expensive [24]. For this reason, the present work proposes an approach entirely based on satellite stereo-pairs, which can provide both spectral and elevation information with lower costs and widespread availability.

The adopted classification strategy is based on OBIA (Object-Based Image Analysis) techniques [25,26]. OBIA refers to classification approaches that investigate geographic entities or phenomena delineating and analyzing image objects rather than pixels [27]. OBIA emerged as particularly useful for extracting and mapping features from high spatial resolution images [28]. OBIA involves two main steps: segmentation and classification. The segmentation phase divides the image into objects, i.e., homogeneous groups of contiguous pixels that have similar spectral and/or spatial characteristics [29]. The segmentation step can be performed in a hierarchical approach, with semantic relationships between objects at different levels [30]. The classification is therefore based on the different spectral, shape, texture, hierarchical, and contextual features of these objects [31].

The validation of OBIA classifications is still debated in the literature and, even though several approaches have been proposed, there is not an obvious solution [32]. The optimal choice depends on the types of OBIA application, which can span from a complete partition of the studied area into polygons to the detection or delineation of selected spatial entities [33]. The confusion matrix, usually expressed in terms of areas or feature count, provides valuable information, even if it is not sufficient to fully characterize the quality of the obtained products [34]. In theory, the geometric quality of the segmentation should be assessed in terms of both location and shape. Many metrics can be defined to evaluate the similarity of mapped polygons against reference polygons, but most of them assume a one-to-one correspondence that is unlikely or even impossible to achieve in many applications, due to under or over-segmentation. This problem can be overcome by

identifying corresponding polygons on the basis of intersection criteria, but their practical implementation varies greatly in the published works.

This paper adopts a separate assessment for the segmentation phase (mainly related to the problem of building identification) and the thematic accuracy of the obtained roofing materials map. The computation of the confusion matrix is adapted to handle the problem of over-segmentation.

2. Materials

The study area is the entire city of Bologna (Italy), which includes very different kinds of urban texture. The city center is mainly composed of a dense pattern of historical buildings and small houses with complex shapes and narrow streets, characterized also by the typical porticoes that became UNESCO World Heritage in 2021 [35]. Conversely, in the peripheral and productive districts a wider variety of building types can be found (e.g., industrial sheds and condominiums), and many open spaces are present. The terrain has a mainly flat topography, with an average elevation of 60 m a.s.l., except for the southern part that includes a hillside, with moderate slopes, higher vegetation coverage and far lower building density.

Three WV3 panchromatic stereo-pairs were acquired (one on 14 September 2017 and two on 20 September 2017); furthermore, two multispectral and SWIR images were also acquired (20 September 2017). Overall, they cover an area of approximately 100 km² (including all the urbanized territory of the municipality, see Figure 1) and can be considered cloud-free and haze-free. All the images were partially preprocessed by the supplier at the Ortho-Ready level.

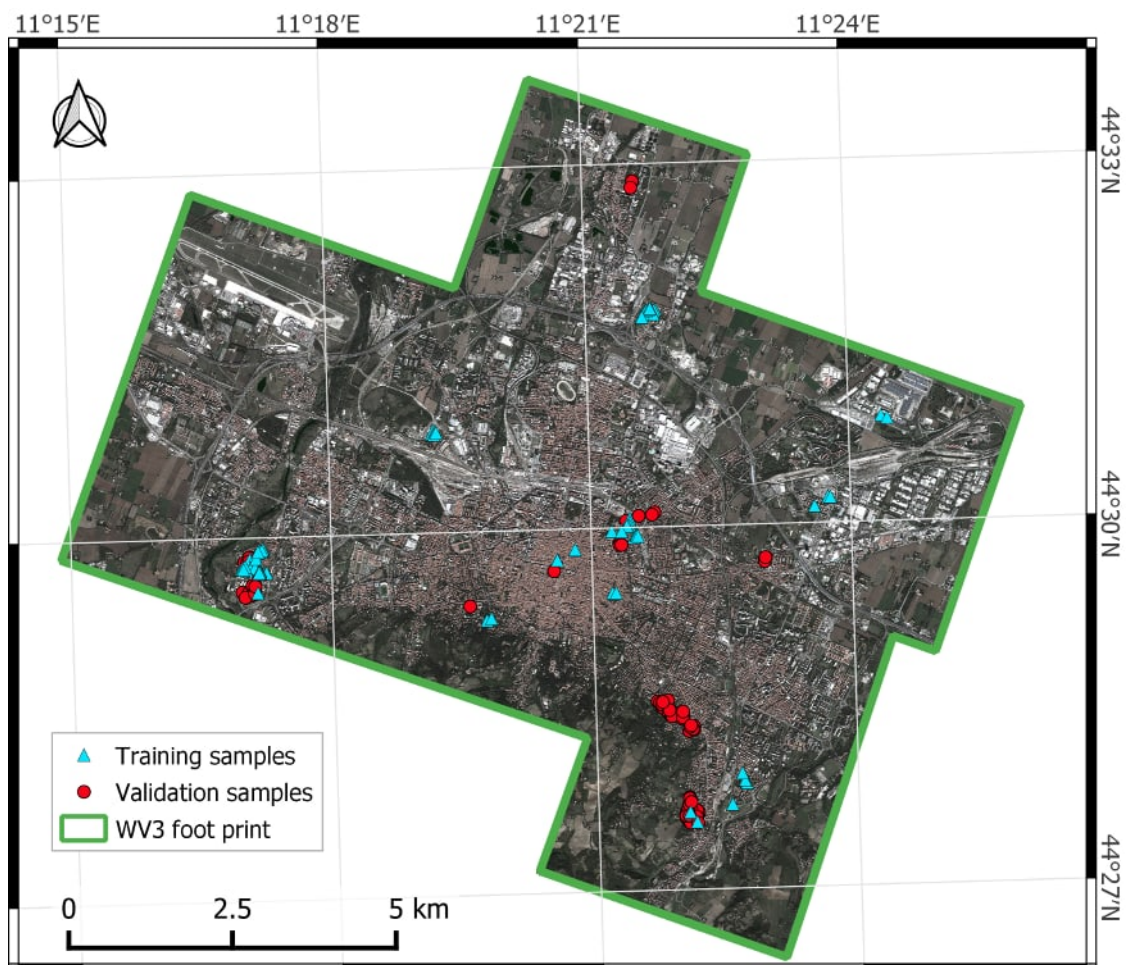


Figure 1. The WV3 image of the study area, with the location of training and validation sites.

For the orthorectification purpose, five ground control points were measured using network real-time kinematic (NRTK) techniques, based on virtual reference stations, with a precision of the measured coordinates of ~2 cm. Furthermore, to perform the atmospheric correction, four spectral signatures of different land cover materials were measured in the field (contemporary to image acquisition) with an SVC HR-768i portable spectroradiometer, and properly georeferenced with NRTK measurements.

For validation purposes, the technical cartography (TC) of the municipality, which comes with a nominal scale of 1:2000 and a tolerance for planimetric coordinates of ~0.5 m, was used to create the ground truth polygons and handle the over-segmentation problems. The cartography is updated to autumn 2016.

All the data were coherently framed in the ETRF2000 (epoch 2008), which is the official geodetic framework for Italy.

To validate the classification, several field data were acquired during dedicated surveys by directly accessing some roofs and using also the information derived by some inspections using UAVs (Unmanned Aerial Vehicles).

2.1. UAV Inspections

Video inspection using a drone was chosen to expand the calibration and validation dataset. Sequences of close-range images were acquired above the surveyed buildings, in order to unequivocally recognize the material of the inspected roofs.

The area analyzed for this project is an urban center and consequently strict restrictions are placed on UAV flights, for the protection of sensitive areas (such as airports and buildings of strategic importance) and to ensure the safety of the population living in densely inhabited area. These restrictions impose a maximum flight altitude above the ground level: it varies from 0 m (No Fly Zone) in the immediate vicinity of the sensitive areas, to 25 and 45 m moving away from them. In most of the city area the limit is 60 m. These limitations were considered in the design of the overflights, as it was necessary to avoid the no-fly zones and also the areas with the 25 m limit, as it is penalizing for roof inspections of tall buildings. Furthermore, the requirement to operate in “visual line of sight” (the pilot never loses sight on the drone) and to prevent interferences on the radio control results in a maximum horizontal distance between the UAV and pilot of 100 m. This implied the need to perform numerous take-offs and landings to inspect sufficiently large portions of the territory.

A small UAV was chosen, with a take-off weight of less than 0.25 kg. It mounts a 12 megapixel optical sensor installed on a gimbal, to allow precise orientation of the camera controlled by remote. Given the equivalent focal length of 25 mm and a maximum flight height of 45 m, the resulting GSD is no wider than 0.02 m.

Because of the premises described above, it was therefore not possible to follow rigorous statistical criteria for the random selection of the sampling points, but these overflight areas necessarily had to fall within those where aerial activity is permitted. Despite these limitations, four areas with heterogeneous roofing were carefully chosen to maximize the amount of information obtained. Overall, 12 flights were performed over these four areas, covering ~0.5 km² in total.

3. Methods

3.1. DSM Extraction

The procedure for the generation of the DSM is already discussed in the technical note in [2], and only the most relevant features are recalled here. The entire photogrammetric process was performed in the OrthoEngine tool of the Geomatica suite (version 2018, PCI Geomatics, Markham, ON, Canada), using a matching strategy based on the semi-global matching technique [36]. The WV3 images were oriented through a rational polynomial model, using the coefficient provided by the supplier. The model was further refined with a first-order polynomial, whose coefficients were estimated using five control points and 73 tie-points, well distributed in the study area and manually collimated. The overall root

mean square error (RMSE) of the residuals was well below one pixel. The DSM extraction was performed with the highest possible resolution, which is the original resolution of the panchromatic band (0.3 m), and no smoothing filters were applied. The DSM was validated with check points and existing models of the area (from LiDAR data and oblique aerial images). Roof surfaces and open spaces are reconstructed with an average error of 1–2 pixels, but severe discrepancies frequently occur in narrow roads and urban canyons (up to several meters in average) [2].

For the present work, the DSM was further processed to remove the elevation of buildings and trees and estimate a DTM. Considering the peculiarity of the urban environment, where built-up surfaces cover most of the area, DTM computation was based on a percentile filter applied on the raster DSM. A large square kernel (300 pixels, equal to 90 m) was considered, in order to include enough pixels at the terrain level. The 10-percentile of the elevation values was computed inside each kernel and assumed as the terrain level, excluding possible outliers due to some residual noise in the model. Considering the overlap of 50% between adjacent kernels, the final DTM comes with a resolution of 45 m. Kernel size and percentile were manually tuned considering some small areas where known points were present. Locally, it is still possible that a particularly dense web of buildings and trees may raise the DTM value of a cell; but this combination of parameters is expected to make it very unlikely to happen. A smaller kernel size (150 pixels) was chosen for the hilly portion of the study area, where terrain elevation varies more rapidly and no large buildings are present. The final DTM was validated against the vertical control points of the TC; the resulting median value of the residuals is 0.5 m (interquartile range 1.3 m).

3.2. Multispectral Image Pre-Processing

WV3 images were radiometrically and geometrically corrected. Using the coefficients provided by the supplier, all the bands were firstly calibrated to top-of-atmosphere reflectance, then to surface reflectance using an empirical line correction. The linear regression was based on the four reflectance spectra measured on the field contemporary to satellite image acquisitions. Multispectral and SWIR bands come with different spatial resolutions and different sets of rational polynomial coefficients for the orthorectification process. To ensure the best possible coherence of the whole dataset, both multispectral and SWIR bands were orthorectified using the DSM derived by the panchromatic stereo-pairs and the same five ground control points used for the generation of the DSM itself.

All the bands were stacked into a unique data cube at the resolution of 1.2 m, by upsampling SWIR bands with a nearest neighbour algorithm. Finally, the two images (acquired on 20 September 2017) were mosaicked to cover the whole study area.

3.3. Object-Based Classification

To automatically detect buildings and differentiate them depending on their roofing material, a new multi-level object-based procedure was developed as a process tree, starting from the approach proposed in [37]. The classification was applied to the WorldView-3 imagery and the derived DSM, using eCognition software (version 10.1, Trimble, Munich, Germany). Figure 2 shows the general workflow of the proposed procedure.

On a DSM, buildings are typically characterized by sudden changes in elevation along the perimeter, because they present very different height values compared to the surrounding environment. Moreover, a feature that allows to distinguish buildings from other landscape elements is the very steep slope at the edges [37,38]. This information was therefore obtained by applying a surface calculation algorithm to the 30 cm pixel size DSM, i.e., the Zevenbergen–Thorne method [39]. Prior to slope layer extraction, the Gauss-Blur low-pass convolution filter using the 3x3 window was applied to the DSM to reduce the effect of data noise (Figure 3b).

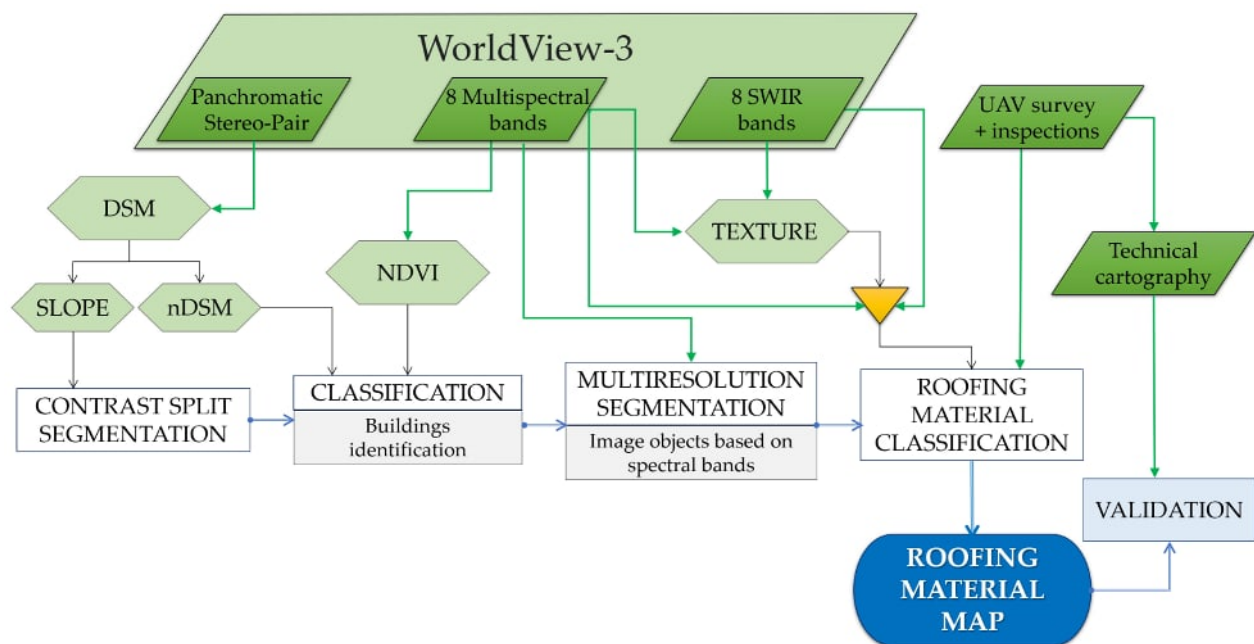


Figure 2. Main steps of the proposed approach.

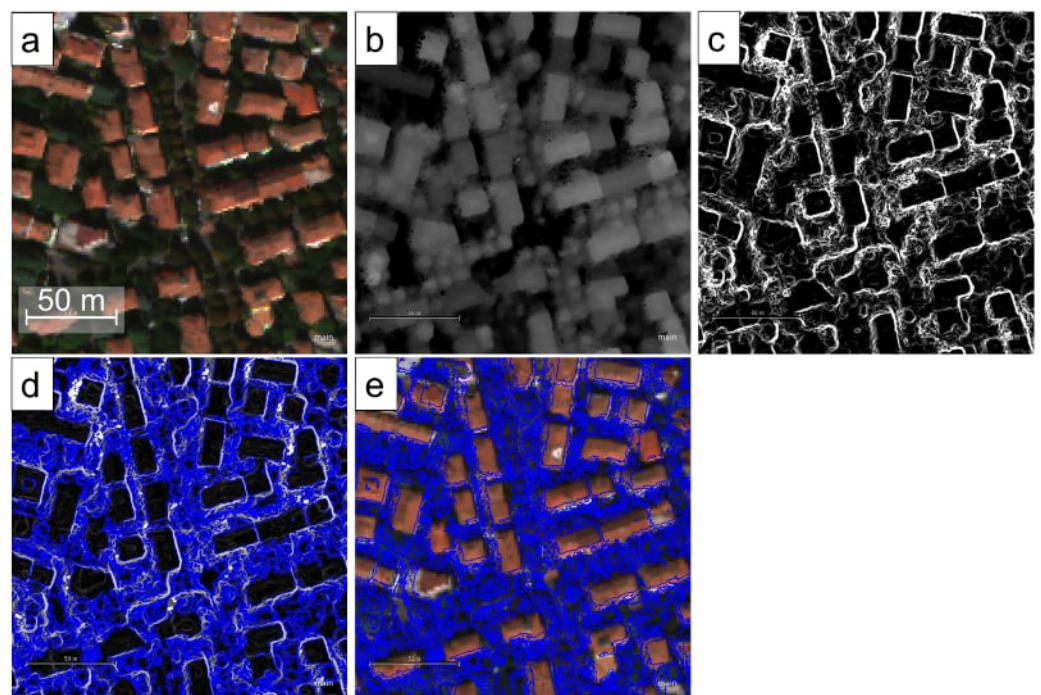


Figure 3. Main steps of the segmentation. (a) the original WV3 image; (b) the photogrammetric DSM; (c) slope layer; (d) objects obtained from the contrast split segmentation (blue lines); (e) objects overlaid to the true color WV3 image.

The “Contrast Split Segmentation” (implemented in the eCognition software) was then performed on the slope layer to generate the image objects. This algorithm was chosen due to its capability to exploit the high contrasts between the steep zones (which appear as bright pixels in Figure 3c) corresponding to building borders, and flat areas (darker pixels). Two segmentation steps were necessary: the first is applied at the pixel level to delineate steep slope objects, while the second one is applied to the first object level to separate areas

with medium slope from flat ones. This segmentation approach allowed to obtain image objects very similar in shape to the buildings (Figure 3e).

The obtained objects were then classified applying sequential rules. A masking approach was performed, in order to detect and exclude from subsequent processing the objects that do not belong to buildings. Firstly, a threshold (equal to 50%) on the slope was set to identify steep objects. The NDVI (Normalized Vegetation Difference Vegetation Index) was used to exclude vegetated areas; considering WV3 imagery, the index was calculated using bands five and seven and a threshold of 0.3 was adopted. To identify the image objects actually corresponding to roofs, a rule based on the object height was applied. This feature was calculated as the difference between the DSM and the DTM derived from WV3 pairs (as described in Section 3.1). Image objects characterized by an actual height equal to or greater than two meters were assigned to the building category. A refinement rule based on the mean elevation above the sea level and the extent of the objects was introduced to avoid the erroneous inclusion of very small objects (area < 30 m²) and some very large objects located in the hilly zone (elevation > 120 m and area > 20,000 m²) that cannot be buildings. In particular, the lower limit of 30 m² can be debatable: it is thought to remove little objects that are detected as building but are actually due to unavoidable residual noise in the DSM and consequently in the slope layer. It has also the advantage of removing temporary canopies or parked truck trailers. Occasionally, it can happen that some little man-made structures are excluded (for example power cabins). However, they are not relevant for the proposed analyses, as also suggested in other studies [40].

After masking out all the objects not belonging to buildings, a multiresolution segmentation was performed, which generated a second level of objects, in order to exploit the spectral information and characterize the buildings according to their roofing material [37,41]. The segmentation algorithm was applied to the eight multispectral bands and not to the SWIR ones, due to their worse spatial resolution. The resulting new objects became the input for the supervised classification. Six informative macro-classes of roofing materials were identified, grouping the most widely used in the city:

1. Clay tiles;
2. Sheath;
3. Metal sheets;
4. Gravel;
5. Gravel tiles;
6. Other materials.

The “Clay tiles” class considers two subclasses: red- and gray-colored tiles. The great majority of buildings of the historic center has the traditional red third or half round ridge clay tiles. The “Sheath” class groups bituminous, slated and reinforced sheaths. Sheets of different alloys and with different coatings are included in the class “Metal sheets”: red, brown, green painted and uncoated metal sheets. Finally, two different classes were considered for the material gravel. “Gravel” refers to flat roof covered by a layer of loose gravel, and “Gravel tiles” refers to concrete tiles with gravel. Examples of roofing material classes are shown in Figure 4.

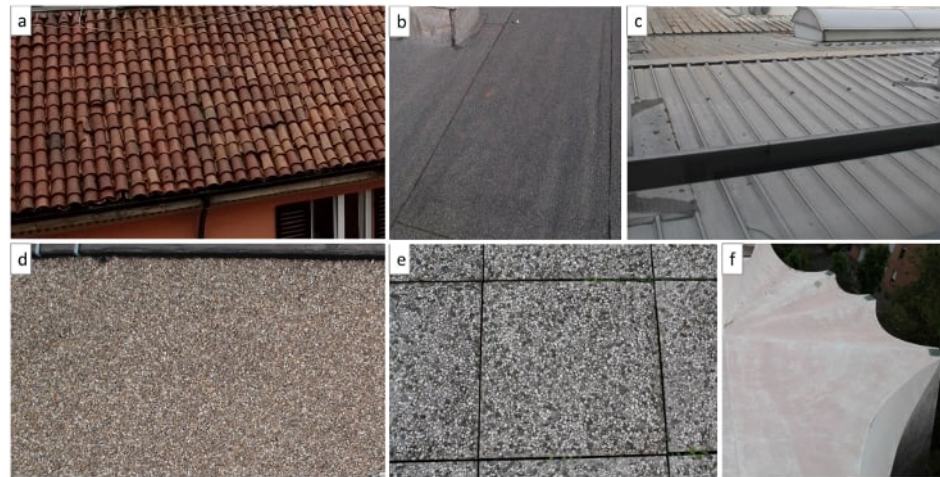


Figure 4. Examples of materials considered in the roofing classification (pictures gathered during the surveys). (a) Red clay tiles; (b) bituminous sheath; (c) metal sheets; (d) layer of loose gravel; (e) gravel tiles; (f) a uniform white-painted concrete roof belonging to “other materials” class.

The training samples for each class (except “other materials”, as explained later in this section) were obtained by the field surveys (direct roof inspections and UAV flights) described in Section 2.1. After the segmentation phase, the segments corresponding to some known roofs, with surfaces as homogeneous as possible, were picked as training. Overall, 49 buildings were selected as samples encompassing the five defined classes, for a total surface of ~24,800 m² (Table 1).

Table 1. Summary of the training samples used for the supervised classification of the roofing materials.

Macro Classes	Area (m ²)	No. of Buildings	No. of Objects	No. of Subclasses
Clay tiles	7764.9	16	741	2
Sheath	5964.9	9	496	2
Metal sheets	6614.4	17	910	4
Gravel	2071.9	3	249	1
Gravel tiles	2428.0	4	365	1
Total	24,844.2	49	2761	10

A supervised sample-based machine learning algorithm was applied. The kernel based classifier SVM [42,43] was chosen, which proved to be successful with both linear and non-linear datasets [44]. Here, the following object-features were considered and computed from training samples: mean and standard deviation for all the 16 bands, texture metrics such as homogeneity and correlation (obtained through the gray level co-occurrence matrix [45]) on Red, Red-edge, and SWIR1 bands. Indeed, wavelengths corresponding to the Red and Red-edge were considered informative for the discrimination of roofing materials, especially for clay [17]. Different kernels and parameters were tested. The optimal set was found in the non-linear radial basis function kernel using parameters C and gamma equal to 100 and 0.1 respectively (see the discussion in Section 5 for further details).

Intrinsically, the SVM classifier assigns all the objects to one of the given classes by partitioning the feature space. However, although the collected samples encompass the most common materials covering the vast majority of the buildings, it is likely that some very rare ones are not included. A membership function was considered in order to exclude from the given classes those objects presenting spectral characteristics too far from the training clusters. These objects were then assigned to the “other materials” class.

Two tests were performed at this point. In the first one all the 16 WV3 bands were included in the feature space for the classification, whilst in the second one only the eight bands belonging to the multispectral sensor were used. The rationale behind this choice is

to evaluate the contribution of the SWIR bands that, on the one hand, provide meaningful spectral features for material discrimination, but, on the other hand, come at a significantly lower resolution (3.4 m instead of 1.2 m).

3.4. Accuracy Assessment

The accuracy assessment is divided in two parts. The first one regards the building identification and it is performed over the whole area, while the second concerns the roofing material classification and it is limited to some check sites (as detailed below).

The first assessment was carried out in GIS environment by comparing the results obtained from the methodology previously described, with the official cartography. To avoid possible unwanted errors caused by the reference vector data, a series of checks on the building geometries and topological correctness of the reference data were performed. Furthermore, the elements not significant for the comparison were removed, i.e., canopies, shacks, cabins, walls, porticoes, kiosks, greenhouses and all the elements less than 2 m high. These elements, indeed, include structures that are temporary, very small or do not fit the definition of building.

All mapping elements were dissolved into a single geometry then separated again obtaining single polygons for each group of adjacent buildings. The output was then filtered by removing also the elements with particularly small area, less than 30 square meters. Finally, a clipping polygon was identified as the intersection between the footprint of WV3 imagery and the Municipality area. Only the area inside this polygon was considered for the accuracy analysis.

Finally, both the segments identified as buildings and the reference vector map were exported into Boolean rasters (0–1) with 0 equal to “no-building” and 1 equal to “building”, and aligned with the same grid spacing of 0.3 m.

The following step is the assessment of the material classification. A confusion matrix was computed by comparing the results of the classification with a validation dataset, made of 100 buildings, covering a total surface of more than 30,000 m². Although random sampling was not possible, the ground truth dataset was gathered from field surveys. In a GIS environment, indeed, after a proper georeferencing of the data, a new field was created in the attribute table of the technical cartography (TC); it contains the description of the roofing material, obtained by the interpretation of the close-range photographs and direct inspections.

The validation was performed considering the fact that even the most homogeneous roofs can have other materials or elements on their surface, such as rooftop HVAC (Heating, Ventilation, and Air Conditioning) units or chimneys. Moreover, trying to overcome oversegmentation problems and to realize a one-to-one object comparison, super-object level was created: all smaller objects included inside a building (as defined in the TC) were merged into a single super-object. To every super-object roof, a single class was assigned with a majority filter, i.e., the class covering the majority of its surface.

To quantitatively assess the quality of the classifications, the classic remote sensing indicators were computed from the confusion matrices, i.e., overall accuracy, producer and user accuracies and K-coefficient, as defined by Campbell [46].

4. Results

Figure 5 shows the roofing materials map obtained (with the procedure described above) using all the 16 multispectral and SWIR bands. In addition, in Figure 6 the effects of majority filter on the classification are presented. Comparing the two products, it is clear that considering every building as a single object and assigning a single material to it reduces remarkably the classification noise due to HVAC elements and smaller components on the rooftop.

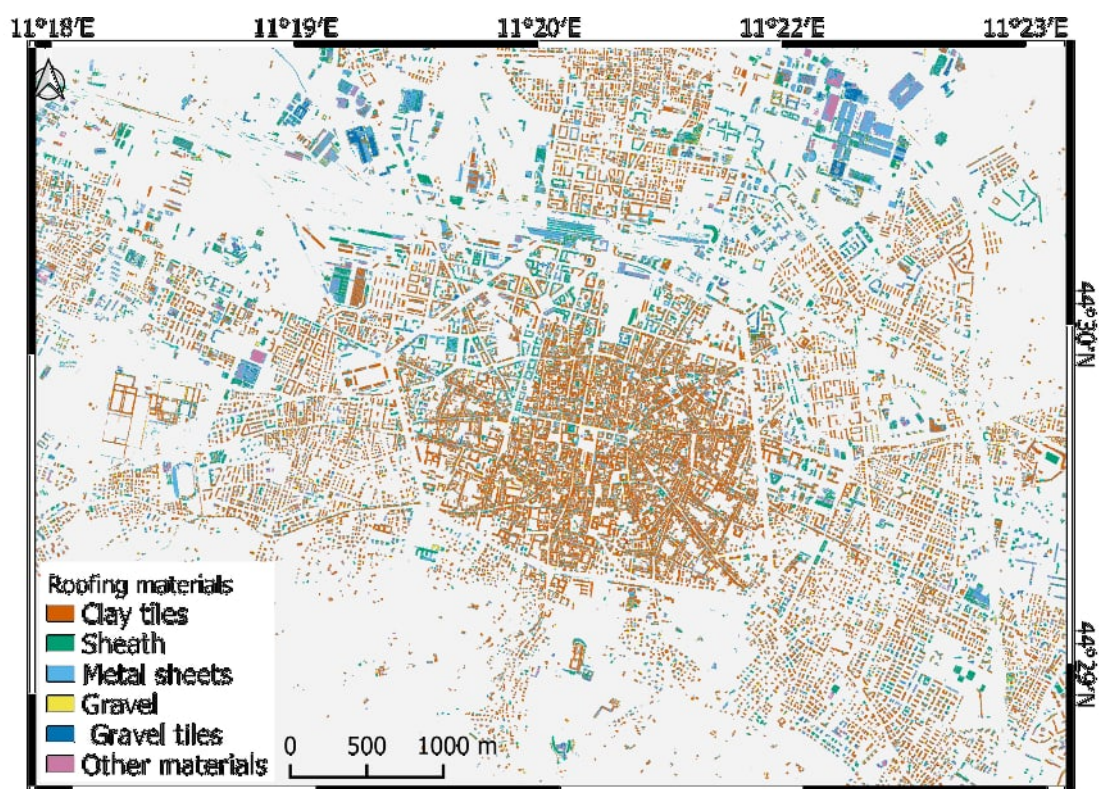


Figure 5. Classification of the roofing materials over the city of Bologna using all the 16 WV3 bands.

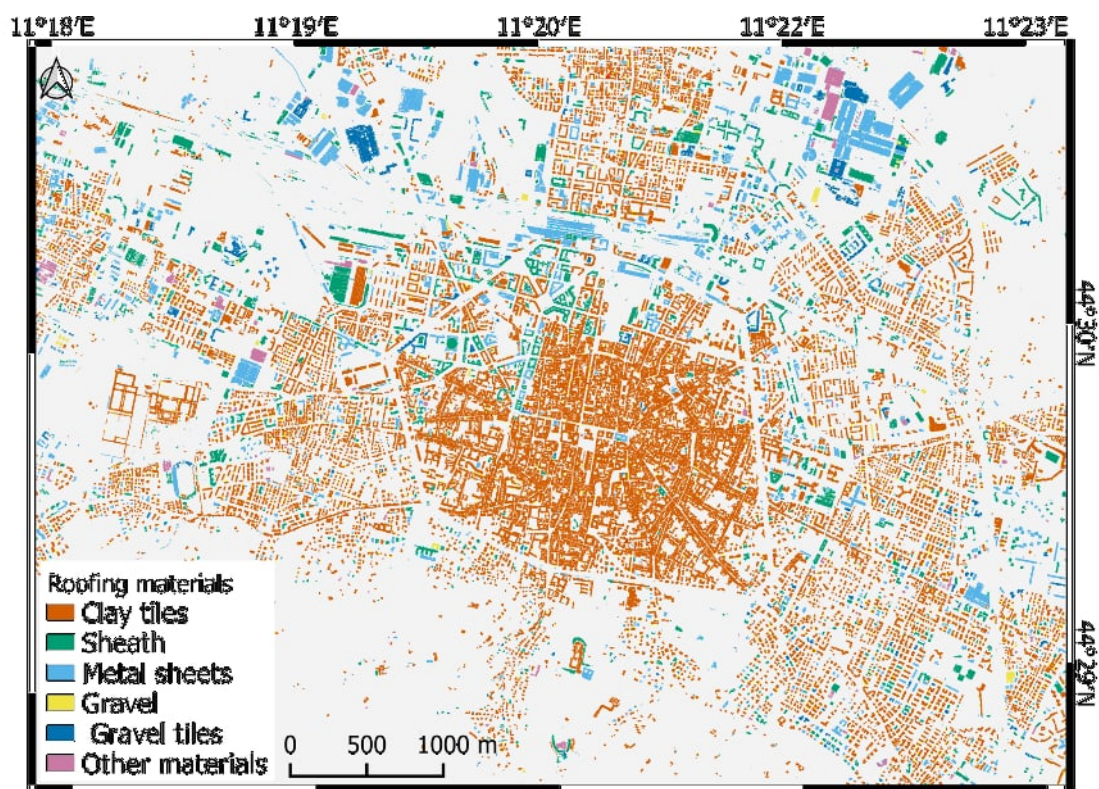


Figure 6. Classification map of the super-object image after the majority filter.

From the final classification map (Figure 6) some statistics about the occurrence of the different materials can be obtained (Figure 7). It is confirmed that the clay tile is the most

used material, covering more than half (54%) of buildings in the study area. The other frequent materials are sheath (20%), followed by metals sheets used in the 15% of cases. The remaining roofs are covered by gravel tiles (6%), gravel (1%); finally, the 4% does not fall in any of the considered classes and is labeled as “other materials”.

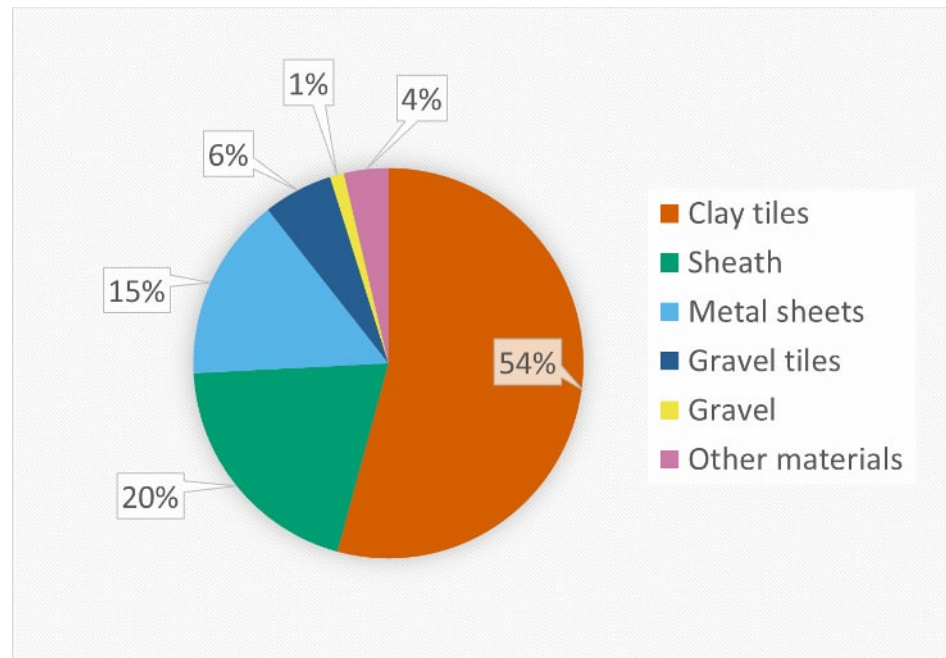


Figure 7. Occurrence of the roofing classes in Bologna (considering the results of the majority filter).

As described in the methodology, the accuracy assessment was performed in two steps. First, the quality of the building detection was assessed over the entire area, then the performance of the material classification was checked against the collected check sites. The accuracy of the building identification procedure is summarized by the confusion matrix in Table 2. The detection reached an overall accuracy of 94.9%. The procedure was able to correctly identify 9.8 km² of buildings and only 2.6 out of 12.4 km² were not recognized, returning an omission error of 21.2%. In addition, a lower commission error (18.5%) was found: only 2.2 km² of the no-building area (total of 82 km²) was misclassified as buildings.

Table 2. Confusion matrix compares areas classified as buildings by the proposed procedure with the reference given by the TC of Bologna. The matrix elements are expressed in km², while the derived accuracy indicators are expressed as percentage.

Classification		Reference	
		Building	No Building
Classification	Building	9.803	2.223
	No building	2.633	79.720
	Producer accuracy	78.8%	97.3%
	User accuracy	81.5%	96.8%
	Overall accuracy	94.9%	
	K coefficient	77.2%	

Second, the accuracy of the roofing materials map was assessed. The validation of SVM classification based on all the 16 WV3 bands was performed by computing the confusion matrix shown in Table 3. An overall accuracy of 91.3% was achieved (K-coefficient 88.9%).

On the other hand, the overall accuracy obtained by training the SVM only on the 8 multispectral bands and on the texture of the red and Red-Edge bands (Table 4) failed to reach the 90% threshold (89.5%, K 86.7%).

Table 3. Confusion matrix (expressed in m²) and accuracy metrics (expressed as percentage) for the classification of roofing materials obtained by considering both the multispectral and the SWIR bands.

		Reference						Sum
		Clay	Sheath	Metal	Gravel Tiles	Gravel	Other	
Classification	Clay tiles	5505.4	127	369.5	0	0	0	6001.7
	Sheath	254.3	6047	215.5	0	0	0	6516.4
	Metal sheets	387.6	647	8384.9	19.4	0	0	9438.8
	Gravel tiles	0	1	51.2	4958.2	0	0	5010.7
	Gravel	0	410	0	4.9	1305.2	0	1720.3
	Other materials	8.9	0	241.6	0	0	2353.4	2603.9
	Sum	6156.3	7231.7	9262.6	4982.5	1305.2	2353.4	
Producer accuracy		89.4%	83.6%	90.5%	99.5%	100.0%	100.0%	
User accuracy		91.7%	92.8%	88.8%	99.0%	75.9%	90.4%	
Overall Accuracy		91.3%						
K coefficient		88.9%						

Table 4. Confusion matrix (expressed in m²) and accuracy metrics (expressed as percentage) for the classification of roofing materials obtained by considering only the eight multispectral bands.

		Reference						Sum
		Clay	Sheath	Metal	Gravel Tiles	Gravel	Other	
Classification	Clay tiles	5312.1	0	1.4	0	257.76	0	5571.3
	Sheath	254.3	6269.0	116.5	24.3	0	0	6664.1
	Metal sheets	0	310.3	8536.0	0	0	480.4	9326.7
	Gravel tiles	0	240.1	367.2	4958.2	0	0	5565.5
	Gravel	581.0	412.2	0	0	1047.4	0	2040.6
	Other materials	8.9	0	241.6	0	0	1873.0	2123.5
	Sum	6156.3	7231.7	9262.6	4982.5	1305.2	2353.4	
Producer accuracy		80.3%	86.7%	92.2%	86.3%	99.5%	79.6%	
User accuracy		51.3%	94.1%	91.5%	95.3%	89.1%	88.2%	
Overall Accuracy		89.5%						
K coefficient		86.7%						

The main difference between the 16-band and 8-band solutions can be found in the clay class, where a sharp decrease in the user accuracy is observed.

5. Discussion

The proposed methodology proved to be effective in detecting and classifying roofing materials over a large area, including very different kinds of urban textures. As for the building extraction, the obtained producer accuracy (79%) is lower compared to the results achieved by some recently proposed deep learning algorithms [47,48]; these algorithms, however, require very large datasets for the training, especially to handle the complexity of different spatial structures and distributions [49]. It is worthwhile to point out that the validation adopted here is based on the TC of the municipality, which, on the one hand, has the advantage to be a fully independent dataset, and on the other hand, it comes with its own errors and small inconsistencies with imagery data. For the analysis reported in Table 2, any manual editing on the data is avoided. Other works, achieving better results, used ground-truth data based on manual delineation of features from the images themselves [50], or data that underwent a huge manual editing [40].

In the proposed detection approach, the variations in the urban texture impact on the complexity of the rule set for the segmentation phase. For example, the historic city center, mainly flat, completely built-up with buildings close or even connected to each other, differs a lot from the residential zone in the southern part of the study area. That zone

indeed is hilly and covered by woods surrounding mainly detached houses. In this context a segmentation based on the slope instead of the original DSM proved to be effective.

The burdensome requirement of a large training set limits the adoption of deep learning algorithms for roofing materials classification, because in many practical cases ground-truth datasets are relatively small and they can be exploited more efficiently with SVM [51]. Here, the achieved accuracies (overall 91%, user and producer reported in Table 3) are comparable (or outperform in some cases) with the reported experiments using similar imagery and OBIA approach [12–14], and also with few works exploiting convolutional neural networks but with a more limited number of material classes [9,23].

A critical step in the described approach is the parameterization of the SVM algorithm. The choice of the set of parameters has a significant impact on the final accuracy. Indeed, different parameters were tested for the optimization of the solution (using 16 bands), as summarized in Figure 8. The adopted set is the one resulting in the highest accuracy.

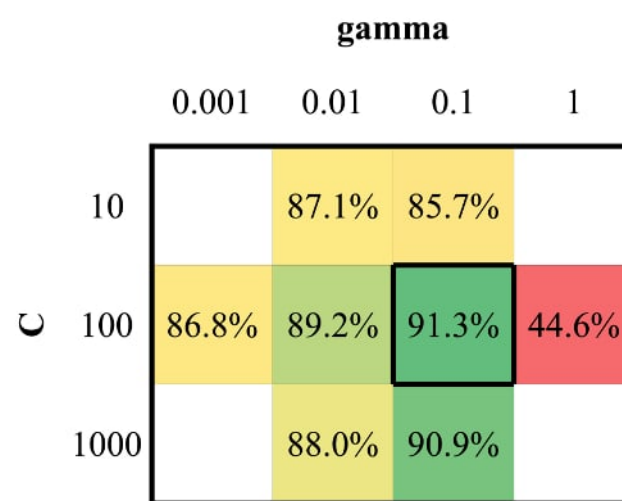


Figure 8. Calibration of the radial basis function kernel parameters of the SVM classifiers. Overall accuracy obtained with different set of C and gamma.

The inclusion of the SWIR bands improved the quality of the final solution, especially for clay. This might be due to the fact that some distinctive spectral features of clay minerals appear in the SWIR region. Overall, the rate of the improvement is a trade-off between the additional spectral information content and the lower spatial resolution, which might smooth spectral signature differences. However, using the texture metrics, the lower resolution of the SWIR band may have the positive effect of reducing the noise due to HVAC units or chimneys on the rooftops.

As for the considered classes, the classification included the most used roofing solutions, but took into account also the possible occurrence of rare materials. One example of such a situation is depicted in Figure 9, where the temporary cover of a swimming pool is correctly assigned to the other material class. The shape of the detected object appears irregular, probably as a consequence of the rounded and convex surface of the dome inflatable structure, which is poorly resolved by the photogrammetric DSM.

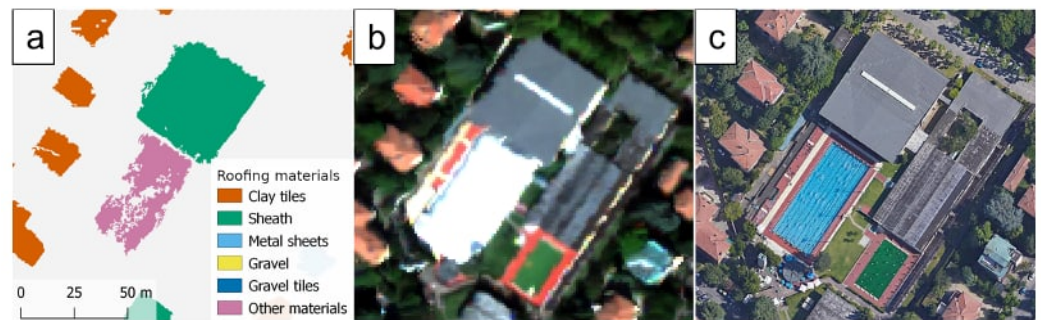


Figure 9. Example of other materials class. The roof classified as other material in (a) belongs to a municipal swimming pool building. The white roof visible in the WV3 image (b) corresponds to the structure used to convert the summer outdoor public pool visible in the Google Earth [52] image (c) into an indoor one during the rest of the year.

Another example of other material is the tensile structure realized for the restoration of the Neptune statue in the city center in 2017, as shown in Figure 10. It can be noted that close to the upper right corner of the “other materials” there is a false detection of a metal roof. Probably, this misclassification is due to an artifact of the DSM occurring in a shadowed area of the original image.

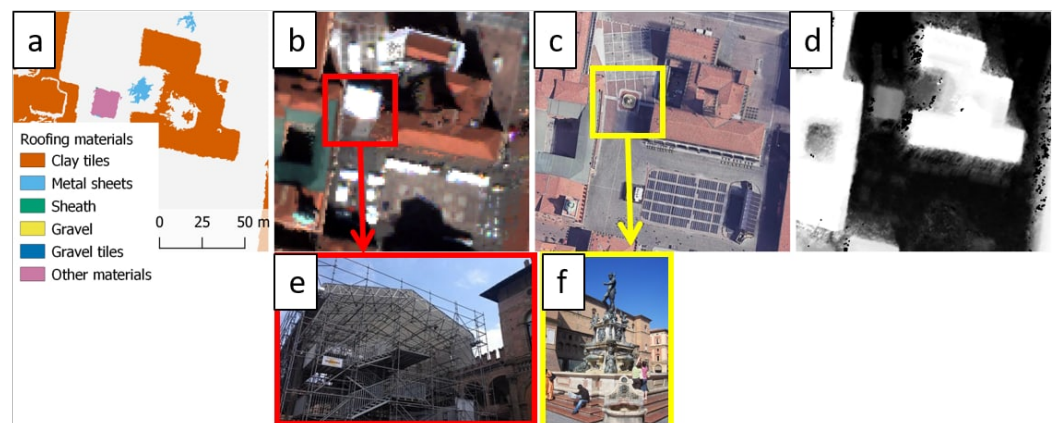


Figure 10. (a) Classification map detail; (b) corresponding WV3 image used for the classification; (c) another high-resolution image of the same area without the temporary structure (from Google Earth [52]); (d) photogrammetric DSM; (e) photograph of the tensile structure; (f) photograph of the statue of Neptune.

Investigating the commission errors outlined by the confusion matrix, it can be pointed out that some errors are related to paintings and coatings of sheath membranes. For example, the two buildings in Figure 11 are classified as metal sheets and bituminous sheath, respectively, when actually they are both covered by sheaths. One of them (indicated by the left arrow in Figure 11a) has an aluminum paint over the whole surface, while the other (right arrow) is only partially painted. When dominant, this aluminum paint led to a misclassification as a metal roof. Regarding the paintings, also aging effects may have consequences on the spectral signature of the material (as highlighted also by [53,54]) and in turn on the assignment of the objects.

The class with the lower user accuracy is the gravel (Table 3). This fact may be related to two possible causes. First, gravel can show large variations in the spectral signature depending on the lithology of the originating rocks. The second aspect is the unbalance of the training dataset (see Table 1, where gravels appears to be the minority class). Sampling unbalance, indeed, typically does not affect the overall accuracy, but has a considerable impact on the accuracy of the minority class [55]. For relatively rare classes, however, it is difficult to obtain a numerous set of samples from surveys, especially given the strict

regulations on drone flights in urban areas. Among the mitigation strategies proposed in the literature [55], the under-sampling, over-sampling or the adjustment of the cost function of the SVM may be suitable for these situations and may deserve further investigations.



Figure 11. (a) Classification map detail, where the two blocks indicated by the arrow are both covered by sheath; (b) corresponding WV3 image used for the classification; (c) another high-resolution image of the same area (from Google Earth [52]).

Another delicate point in the proposed classification approach is the application of the majority filter. This step is necessary for the validation phase in order to establish a correspondence between the ground truth objects derived by the TC and the objects coming from the image segmentation. The oversegmentation problem, indeed, is widespread all over the processed area. Only considering the 100 validation buildings, the median number of segments belonging to one super-object is 35 (with an interquartile range of 50). In half of the super-objects, more than the 70% of the area is covered by the dominant class, while it covers less than 50% of the area only in one quarter of them. This is caused by the occurrence of many small segments of different classes inside almost every building.

As a consequence, the majority filter has a certain impact on the final map. On the one hand, it strongly improves the results by removing some noise due to the presence of chimneys, HVAC units, and skylights, which produce small objects of different classes. It also removes errors occurring at the edges of the roofs, where some segments sometimes include pixels from the facades or the ground. On the other hand, some disadvantages can arise when the roofs of large buildings are covered by different materials over different wings, but they belong to a single super-object. One example of these positive and negative consequences can be seen in Figure 12, where the gravel roof (yellow) appears cleaned from spurious objects after the majority filter, while one correctly classified metal wing (cyan) of the central building is merged with the dominant sheath cover (green).

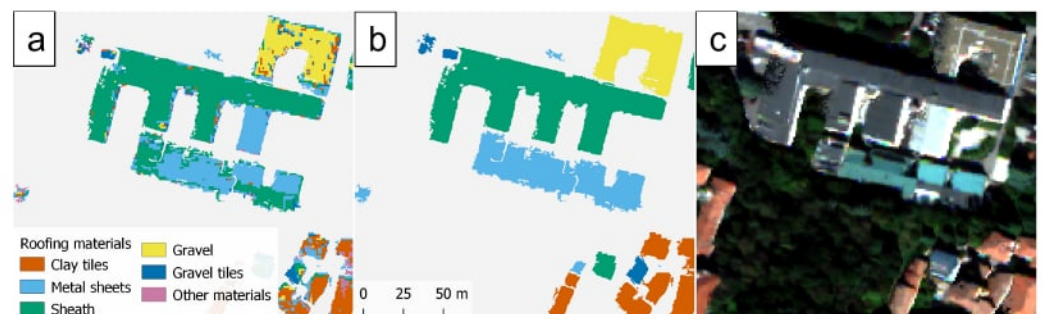


Figure 12. Effects of the majority filter. (a) SVM classification map, (b) application of the majority filter, (c) original WV3 image.

In the perspective of applying the methodology to the management and urban planning of wide metropolitan areas, the adoption of the majority filter produces more homogeneous maps that may fit better the level of detail required by stakeholders.

The obtained classification considers the major categories of roofing materials. Additional experiments and analyses are necessary to verify the capability of multispectral stereo-images to expand the number of classes that can be discriminated. Other aspects to be addressed in future works are a finer analysis on the impact of the segmentation parameters on the SVM results, and how to handle the majority filter without a TC. Regarding this last point, possible alternative sources of vector data may be OpenStreetMap [56] or other crowd mapping initiatives [10].

6. Conclusions

A methodology was presented for the detection of buildings and the classification of the roofing materials over an area of approximately 100 km² and a wide variety of urban textures. It relies only on very high-resolution multispectral images with stereo capabilities acquired by a satellite platform. This makes the procedure highly replicable virtually in any part of the world, since it does not require expensive datasets acquired on the field or by dedicated aerial surveys (such as LiDAR data). The achieved overall accuracy—approximately 91% in roofing classification—is comparable with most of the published works on these topics but applies on a wider area. Most of the observed errors are related to artifacts or defects in the photogrammetric digital surface model, or to the presence of certain coatings that may hide the distinctive spectral features of the roofing materials. Despite these residual errors, the proposed approach can be used to conduct a census of the major roofing materials in a city.

Author Contributions: Conceptualization, E.M.; methodology, F.T. and F.F.; validation, F.T. and A.L.; formal analysis, F.T. and F.F.; investigation, A.L. and E.M.; data curation, A.L.; writing—original draft preparation, F.T., F.F., A.L. and E.M.; writing—review and editing, E.M.; supervision, E.M.; project administration, E.M.; funding acquisition, E.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work is part of the ChoT project, funded by the Italian Ministry of Education, University and Research (MIUR) in the framework of the Scientific Independence of young Researches (SIR) 2014 program (Grant Number RBSI14SYES).

Acknowledgments: The authors would like to thank the personnel of the University of Bologna for their valuable collaboration, especially Andrea Braschi, Roberto Battistini, and Luciano Ricucci (AUTC—Buildings and Works Management) and Paolo Conte, Luca Poluzzi, Luca Tavasci, and Maria A. Tini (Department of Civil, Chemical, Environmental and Materials Engineering) for their help in the field surveys. The authors would like to thank also Andrea Minghetti, Paola Africani, and Elisa Paselli of the “SIT” division of the Municipality of Bologna, for providing the technical cartography.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

DSM	Digital Surface Model
DTM	Digital Terrain Model
GSD	Ground Sampling Distance
HVAC	Heating, Ventilation and Air Conditioning
LiDAR	Light Detection And Ranging
NDVI	Normalized Difference Vegetation Index
OBIA	Object-Based Image Analysis
SVM	Support Vector Machine
SWIR	Short-Wave Infrared
TC	Technical Cartography
UAV	Unmanned Aerial Vehicle
WV3	WorldView-3

References

1. Longbotham, N.; Pacifici, F.; Baugh, B.; Camps Valls, G. Prelaunch assessment of WorldView-3 information content. In Proceedings of the 2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Lausanne, Switzerland, 24–27 June 2014; IEEE: Piscataway, NJ, USA, 2014. [\[CrossRef\]](#)
2. Mandanici, E.; Girelli, V.A.; Poluzzi, L. Metric accuracy of digital elevation models from WorldView-3 stereo-pairs in urban areas. *Remote Sens.* **2019**, *11*, 878. [\[CrossRef\]](#)
3. Georganos, S.; Abascal, A.; Kuffer, M.; Wang, J.; Owusu, M.; Wolff, E.; Vanhuyse, S. Is it all the same? Mapping and characterizing deprived urban areas using WorldView-3 superspectral imagery. A case study in Nairobi, Kenya. *Remote Sens.* **2021**, *13*, 4986. [\[CrossRef\]](#)
4. Ghanea, M.; Moallem, P.; Momeni, M. Building extraction from high-resolution satellite images in urban areas: Recent methods and strategies against significant challenges. *Int. J. Remote Sens.* **2016**, *37*, 5234–5248. [\[CrossRef\]](#)
5. Tomljenovic, I.; Höfle, B.; Tiede, D.; Blaschke, T. Building extraction from airborne laser scanning data: An analysis of the state of the art. *Remote Sens.* **2015**, *7*, 3826–3862. [\[CrossRef\]](#)
6. Luo, L.; Li, P.; Yan, X. Deep learning-based building extraction from remote sensing images: A comprehensive review. *Energies* **2021**, *14*, 7982. [\[CrossRef\]](#)
7. Vu, T.T.; Yamazaki, F.; Matsuoka, M. Multi-scale solution for building extraction from LiDAR and image data. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 281–289. [\[CrossRef\]](#)
8. Bassani, C.; Cavalli, R.M.; Cavalcante, F.; Cuomo, V.; Palombo, A.; Pascucci, S.; Pignatti, S. Deterioration status of asbestos-cement roofing sheets assessed by analyzing hyperspectral data. *Remote Sens. Environ.* **2007**, *109*, 361–378. [\[CrossRef\]](#)
9. Krówczyńska, M.; Raczko, E.; Staniszewska, N.; Wilk, E. Asbestos—cement roofing identification using remote sensing and convolutional neural networks (CNNs). *Remote Sens.* **2020**, *12*, 408. [\[CrossRef\]](#)
10. Abdulkarim, B.; Kamberov, R.; Hay, G. Supporting urban energy efficiency with volunteered roof information and the Google Maps API. *Remote Sens.* **2014**, *6*, 9691–9711. [\[CrossRef\]](#)
11. Bitelli, G.; Conte, P.; Csoknyai, T.; Franci, F.; Girelli, V.A.; Mandanici, E. Aerial thermography for energetic modelling of cities. *Remote Sens.* **2015**, *7*, 2152–2170. [\[CrossRef\]](#)
12. Hamedianfar, A.; Shafri, H.Z.M.; Mansor, S.; Ahmad, N. Improving detailed rule-based feature extraction of urban areas from WorldView-2 image and lidar data. *Int. J. Remote Sens.* **2014**, *35*, 1876–1899. [\[CrossRef\]](#)
13. Gibril, M.B.A.; Shafri, H.Z.M.; Hamedianfar, A. New semi-automated mapping of asbestos cement roofs using rule-based object-based image analysis and Taguchi optimization technique from WorldView-2 images. *Int. J. Remote Sens.* **2016**, *38*, 467–491. [\[CrossRef\]](#)
14. Norman, M.; Mohd Shafri, H.Z.; Mansor, S.; Yusuf, B.; Mohd Radzali, N.A.W. Fusion of multispectral imagery and LiDAR data for roofing materials and roofing surface conditions assessment. *Int. J. Remote Sens.* **2020**, *41*, 7090–7111. [\[CrossRef\]](#)
15. Abriha, D.; Kovács, Z.; Ninsawat, S.; Bertalan, L.; Balázs, B.; Szabó, S. Identification of roofing materials with discriminant function analysis and random forest classifiers on pan-sharpened WorldView-2 imagery—A comparison. *Hung. Geogr. Bull.* **2018**, *67*, 375–392. [\[CrossRef\]](#)
16. Moreira, R.C.; Galvão, L.S. Variation in spectral shape of urban materials. *Remote Sens. Lett.* **2010**, *1*, 149–158. [\[CrossRef\]](#)
17. Samsudin, S.H.; Shafri, H.Z.M.; Hamedianfar, A.; Mansor, S. Spectral feature selection and classification of roofing materials using field spectroscopy data. *J. Appl. Remote Sens.* **2015**, *9*, 095079. [\[CrossRef\]](#)
18. Cilia, C.; Panigada, C.; Rossini, M.; Candiani, G.; Pepe, M.; Colombo, R. Mapping of asbestos cement roofs and their weathering status using hyperspectral aerial images. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 928–941. [\[CrossRef\]](#)
19. Huang, J.; Zhang, X.; Xin, Q.; Sun, Y.; Zhang, P. Automatic building extraction from high-resolution aerial images and LiDAR data using gated residual refinement network. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 91–105. [\[CrossRef\]](#)
20. Kotthaus, S.; Smith, T.E.L.; Wooster, M.J.; Grimmond, C.S.B. Derivation of an urban materials spectral library through emittance and reflectance spectroscopy. *ISPRS J. Photogramm. Remote Sens.* **2014**, *94*, 194–212. [\[CrossRef\]](#)
21. Ilehag, R.; Schenk, A.; Huang, Y.; Hinz, S. KLUM: An urban VNIR and SWIR spectral library consisting of building materials. *Remote Sens.* **2019**, *11*, 2149. [\[CrossRef\]](#)
22. Ye, C.; Li, H.; Li, C.; Liu, X.; Li, Y.; Li, J.; Nunes Gonçalves, W.; Marcato, J.J. A building roof identification CNN based on interior-edge-adjacency features using hyperspectral imagery. *Remote Sens.* **2021**, *13*, 2927. [\[CrossRef\]](#)
23. Kim, J.; Bae, H.; Kang, H.; Lee, S.G. CNN algorithm for roof detection and material classification in satellite images. *Electronics* **2021**, *10*, 1592. [\[CrossRef\]](#)
24. Bitelli, G.; Girelli, V.A.; Lambertini, A. Integrated use of remote sensed data and numerical cartography for the generation of 3D city models. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-2*, 97–102. [\[CrossRef\]](#)
25. Blaschke, T.; Lang, S.; Hay, G.J. (Eds.) *Object-Based Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2008. [\[CrossRef\]](#)
26. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16. [\[CrossRef\]](#)
27. Chen, G.; Weng, Q.; Hay, G.J.; He, Y. Geographic object-based image analysis (GEOBIA): Emerging trends and future opportunities. *GIScience Remote Sens.* **2018**, *55*, 159–182. [\[CrossRef\]](#)
28. Hossain, M.D.; Chen, D. Segmentation for object-based image analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. *ISPRS J. Photogramm. Remote Sens.* **2019**, *150*, 115–134. [\[CrossRef\]](#)

29. Myint, S.W.; Gober, P.; Brazel, A.; Grossman Clarke, S.; Weng, Q. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* **2011**, *115*, 1145–1161. [[CrossRef](#)]
30. Taubenböck, H.; Esch, T.; Wurm, M.; Roth, A.; Dech, S. Object-based feature extraction using high spatial resolution satellite data of urban areas. *J. Spat. Sci.* **2010**, *55*, 117–132. [[CrossRef](#)]
31. Bhaskaran, S.; Paramananda, S.; Ramnarayan, M. Per-pixel and object-oriented classification methods for mapping urban features using Ikonos satellite data. *Appl. Geogr.* **2010**, *30*, 650–665. [[CrossRef](#)]
32. Ye, S.; Pontius, R.G.; Rakshit, R. A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. *ISPRS J. Photogramm. Remote Sens.* **2018**, *141*, 137–147. [[CrossRef](#)]
33. Radoux, J.; Bogaert, P. Good practices for object-based accuracy assessment. *Remote Sens.* **2017**, *9*, 646. [[CrossRef](#)]
34. Whiteside, T.G.; Maier, S.W.; Boggs, G.S. Area-based and location-based validation of classified image objects. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *28*, 117–130. [[CrossRef](#)]
35. UNESCO. World Heritage List. Available online: <https://whc.unesco.org/en/list/1650/> (accessed on 21 December 2021).
36. Hirschmuller, H. Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 328–341. [[CrossRef](#)] [[PubMed](#)]
37. Franci, F.; Lambertini, A.; Bitelli, G. Integration of different geospatial data in urban areas: A case of study. In Proceedings of the Second International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2014), Paphos, Cyprus, 7–10 April 2014; Hadjimitsis, D.G., Themistocleous, K., Michaelides, S., Papadavid, G., Eds.; SPIE: Bellingham, WA, USA, 2014; doi:10.1117/12.2066614. [[CrossRef](#)]
38. Uzar, M. Automatic building extraction with multi-sensor data using rule-based classification. *Eur. J. Remote Sens.* **2014**, *47*, 1–18. [[CrossRef](#)]
39. Zevenbergen, L.W.; Thorne, C.R. Quantitative analysis of land surface topography. *Earth Surf. Process. Landforms* **1987**, *12*, 47–56. [[CrossRef](#)]
40. Ji, S.; Wei, S.; Lu, M. Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 574–586. [[CrossRef](#)]
41. Baatz, M.; Schäpe, A. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informations-Verarbeitung*; Strobl, J., Blaschke, T., Griesbner, G., Eds.; Wichmann Verlag: Karlsruhe, Germany, 2000; Volume XII, pp. 12–23.
42. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
43. Moorthi, S.M.; Misra, I.; Kaur, R.; Darji, N.P.; Ramakrishnan, R. Kernel based learning approach for satellite image classification using support vector machine. In Proceedings of the 2011 IEEE Recent Advances in Intelligent Computational Systems, Kerala, India, 22–24 September 2011; doi:10.1109/raics.2011.6069282. [[CrossRef](#)]
44. Scholkopf, B.; Sung, K.K.; Burges, C.J.C.; Girosi, F.; Niyogi, P.; Poggio, T.; Vapnik, V. Comparing support vector machines with Gaussian kernels to radial basis function classifiers. *IEEE Trans. Signal Process.* **1997**, *45*, 2758–2765. [[CrossRef](#)]
45. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621. [[CrossRef](#)]
46. Campbell, J. *Introduction to Remote Sensing*, 3rd ed.; Guilford Press: New York, NY, USA, 2002.
47. Xu, Y.; Wu, L.; Xie, Z.; Chen, Z. Building extraction in very high resolution remote sensing imagery using deep learning and guided filters. *Remote Sens.* **2018**, *10*, 144. [[CrossRef](#)]
48. Liu, Y.; Gross, L.; Li, Z.; Li, X.; Fan, X.; Qi, W. Automatic building extraction on high-resolution remote sensing imagery using deep convolutional encoder-decoder with spatial pyramid pooling. *IEEE Access* **2019**, *7*, 128774–128786. [[CrossRef](#)]
49. Guo, H.; Shi, Q.; Du, B.; Zhang, L.; Wang, D.; Ding, H. Scene-driven multitask parallel attention network for building extraction in high-resolution remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 4287–4306. [[CrossRef](#)]
50. Qin, Y.; Wu, Y.; Li, B.; Gao, S.; Liu, M.; Zhan, Y. Semantic segmentation of building roof in dense urban environment with deep convolutional neural network: A case study using GF2 VHR imagery in China. *Sensors* **2019**, *19*, 1164. [[CrossRef](#)] [[PubMed](#)]
51. Kuras, A.; Brell, M.; Rizzi, J.; Burud, I. Hyperspectral and lidar data applied to the urban land cover machine learning and neural-network-based classification: A review. *Remote Sens.* **2021**, *13*, 3393. [[CrossRef](#)]
52. Google LLC. Google Earth 7.3.4 (18 June 2019) Bologna, Italy. Available online: <https://www.google.com/earth/index.html> (accessed on 30 January 2022).
53. Nasarudin, N.; Shafri, H. Development and utilization of urban spectral library for remote sensing of urban environment. *J. Urban Environ. Eng.* **2011**, *5*, 44–56. [[CrossRef](#)]
54. Samsudin, S.H.; Shafri, H.Z.M.; Hamedianfar, A. Development of spectral indices for roofing material condition status detection using field spectroscopy and WorldView-3 data. *J. Appl. Remote Sens.* **2016**, *10*, 025021. [[CrossRef](#)]
55. Waske, B.; Benediktsson, J.A.; Sveinsson, J.R. Classifying remote sensing data with support vector machines and imbalanced training data. In *Multiple Classifier Systems*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 375–384. [[CrossRef](#)]
56. OpenStreetMap. OpenStreetMap. Available online: <https://www.openstreetmap.org/> (accessed on 30 January 2022).