# Multi-sensor feature fusion for very high spatial resolution built-up area extraction in temporary settlements

Patrick Aravena Pelizari*, Kristin Spröhnle, Christian Geiß, Elisabeth Schoepfer, Simon Plank, Hannes Taubenböck

*German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, 82234 Wessling, Germany*

## ARTICLE INFO

## ABSTRACT

Detailed and up-to-date knowledge on the situation in temporary settlements of forced migrants plays an important role for effective humanitarian assistance. These settlements emerge as planned or spontaneous camps or camp-like structures, characterized by a small-scale physical morphology and high dynamics. Information on the built-up area (BUA; i.e. areas occupied by buildings) in these settlements provides important evidence on the local situation. The objective of this work is to present a generic procedure for the detailed extraction of BUA in complex temporary settlements from very high spatial resolution satellite data collected by different sensor types. The proposed approach is embedded in the methodological framework of object-based image analysis and is compound of i) the computation of an exhaustive set of spectral-spatial features aggregated on multiple hierarchic segmentation scales, ii) filter based feature subset selection and iii) supervised classification using a Random Forest classifier. Experimental results are obtained based on Pléiades multispectral optical and TerraSAR-X Staring Spotlight Synthetic Aperture Radar satellite imagery for six distinct but representative test areas within the refugee camp Al Zaatari in Jordan. The experiments include a detailed assessment of classification accuracy for varying configurations of considered feature types and training data set sizes as well as an analysis of the feature selection (FS) outcomes. We observe that the classification accuracy can be improved by the use of multiple segmentation levels as well as the integration of multi-sensor information and different feature types. In addition, the results show the potential of the applied FS approach for the identification of most relevant features. Accuracy values beyond 80% in terms of κ statistic and True Skill Statistic based on significantly reduced feature sets compared to the input underline the potential of the proposed method.

## 1. Introduction

### 1.1. Temporary settlement analyses – the benefit of Earth observation

As a consequence of the numerous ongoing crises, large scale displacement of people has reached an unprecedented level in recent history. According to the United Nations High Commissioner for Refugees (UNHCR, 2016), there were about 65.3 million forcibly displaced people worldwide in 2015, including 21.3 million refugees, 40.8 million internally displaced people (IDPs) and 3.2 million asylum seekers. Being forced to flee their homes due to conflict situations, man-made or natural disasters, these people belong to the most vulnerable in the world. Most of them seek protection and shelter in urban environments (Taubenböck et al., 2018), but there are still a large number of refugees living in self-settled or planned camps (UNHCR, 2016). Having arrived in a camp, refugees are generally exposed to poor living

conditions with limited access to water, nutrition, medical care and sanitary facilities. Although they are supposed to be temporary, most of these camps are maintained for years or even decades. Thereby complex settlement structures emerge and originally extemporary buildings are solidified and extended (Herz, 2006; Dalal, 2014). At the end of 2015, about 4 million of the world's refugees gather in camps (UNHCR, 2016). Most of these camps are managed or supported by national or international relief organizations, which supply the camps with essential facilities for survival. For effective camp management and decision-making, humanitarian organizations require reliable and up-to date information about the situation on the ground (Bjørgo, 2000; UNHCR, 2000, 2005). In this context, population distribution and numbers are crucial information for relief operations, e.g. to enhance the logistical support of aid agencies (UNHCR, 2007; Ehrlich et al., 2009).

Remote sensing data provide independent, area-wide and up-to-date information on the camp situation and thus can complement

---

information usually collected by field observations (e.g. Schöpfer et al., 2015). In some situations, Earth observation (EO) is the only reliable or independent source of information. This accounts particularly in situations where the on-ground situation is unclear due to uncontrolled growth and arrival of new migrants as well as for cases in which field assessments are either unsafe due to conflict situations, or provide false information due to stakeholder bias and/or politics (Bjørgo, 2000; UNHCR, 2000).

Especially with the continuous advent of satellite sensors providing data of increasing spatial and temporal resolution, the role of EO-based applications for support to humanitarian relief becomes more and more important (Kranz et al., 2010; Kuffer et al., 2016a). In case of population estimation, the use of very high spatial resolution (VHR) satellite data allows for the detection of single dwelling units or the built-up area (BUA) in a camp. An approximate figure of the number of people living in a camp can then be derived based on estimated occupancy rates (i.e. average number of people per shelter or per sqm). Furthermore, if total counts of population are available (e.g. registration figures), such information can form the spatial base entity for population disaggregation. In this way information on the spatial pattern of the population (e.g. population density) and respective changes over time can be given (e.g. Lang et al., 2010). The present study focuses on the BUA rather than on single dwelling units. BUA is defined here as the detailed spatial delineation of areas occupied by buildings. Thereby, the term building refers to any type of man-made temporary settlement structure with a roof.

### 1.2. Image features for Earth observation-based built-up area detection

The development of methods for the derivation of thematic land use/land cover (LULC) information such as BUA from remote sensing imagery has been a major research subject of the remote sensing community in the past decades. EO data from latest VHR sensors such as the multispectral (MS) sensors WorldView (WV) 1–4 or Pléiades allow for LULC mapping at an unprecedented level of spatial detail. Simultaneously, the benefits of increasing spatial resolution are associated with an increasing mismatch between single pixels and the real world objects they depict. Those real world objects are typically represented by numerous pixels. Particularly in heterogeneous environments such as settlement areas this induces high intra-class and low inter-class variability of the different classes in the spectral domain. This poses challenges for an accurate classification (Blaschke, 2010). In addition to MS data, VHR Synthetic Aperture Radar (SAR) data such as imagery from the sensor systems TerraSAR-X (TS-X) or COSMO-Skymed additionally exhibit new potential for accurate settlement mapping (Taubenböck et al., 2012; Chini et al., 2009). The SAR backscatter signal represents a complex combination of various sources, which can provide additional information about objects on the ground due to their distinctive backscatter signature as induced by characteristic geometric as well as dielectric properties and surface roughness. However, being subject to geometric perturbations due to the side looking geometry of SAR sensors (i.e. double bounce, layover, foreshortening and shadowing effects) as well as the speckle effect, the resulting imagery can be difficult to interpret (Brunner et al., 2008; Gamba, 2013).

Recent studies on remote sensing based thematic mapping propose several image processing concepts in order to cope with these challenges. Most of them are based on spectral-spatial homogenization through the extraction of meaningful image features incorporating information beyond the margins of single pixels.

Aiming to capture distinct spatial grey tone dependency patterns, a variety of studies integrate texture features into the classification procedure. A popular and approved approach for computing such measures is the grey-level co-occurrence matrix (GLCM; Haralick et al., 1973). GLCM texture features have demonstrated to beneficially complement MS information in optical data (e.g. Carleer and Wolff, 2006; Pacifici et al., 2009; Geiß et al., 2015; Kuffer et al., 2016b) as well as in single-

(e.g. Gamba et al., 2011; Ban et al., 2015; Uhlmann and Kiranyaz, 2014) and multi-polarized (e.g. Du et al., 2015; Masjedi et al., 2016; Wurm et al., 2017) SAR backscatter information.

An additional group of image features being able to complementary encode spectral and spatial information constitutes morphological profiles (MPs). These kinds of features are built upon mathematical morphology (MM) operations (Soille, 2004). MPs rely on the idea of comprehensively describing image structures by their morphological intrinsic characteristics exploited through sequential morphological transformations of the image data applying a structuring element (SE) of increasing size (Pesaresi and Benediktsson, 2001). MPs and its derivatives have proven particularly effective for classification of urban land cover in VHR MS and hyperspectral imagery (Benediktsson et al., 2003; Fauvel et al., 2008; Tuia et al., 2009; Dalla Mura et al., 2010; Ghamisi et al., 2015; Geiß et al., 2016b). None the less, the utility of MPs with regard to the classification of SAR data is sparsely documented in literature. Among the few available studies, Chini et al. (2009) assess the potential of anisotropic MPs for the classification of urban land cover based on single polarized VHR TS-X Stripmap data (6 m spatial resolution) achieving promising accuracy levels. Du et al. (2015) integrate MPs into an array of polarimetric image descriptors for LULC classification based on Radarsat-2 fine quad-pol data (8 m spatial resolution) and report a significant boost in overall accuracy. Wurm et al. (2017) deploy MPs supplementary to texture information for the mapping of inner urban structures, specifically informal settlements using dual-pol TS-X Stripmap mode imagery. They conclude that MPs computed from 6 m spatial resolution imagery do not allow for a meaningful representation of individual objects such as slum dwellings. This indicates that a detailed extraction of temporary settlement BUA under the consideration of MPs might require higher resolved SAR data.

A prominent methodological concept to face the challenges of VHR remote sensing image classification is object-based image analysis (OBIA). OBIA relies on the aggregation of pixel values to meaningful image objects using a segmentation procedure (Benz et al., 2004; Blaschke, 2010). An object-based representation of the imagery allows for a straight forward regularization of the data based on common measures of central tendency or spread (e.g. *mean*, *median* or *standard deviation*). Additionally, it facilitates the spectral-spatial integration of the pixel information (e.g. spectral values, indices and texture) with geometric characteristics (i.e. object shape and extent) as well as object-based contextual measures (e.g. topological relationships) from single or multi source data into the classification procedure (Stumpf and Kerle, 2011; Geiß et al., 2015). The integration of object-based features calculated from a sequence of multiple hierarchical segmentation-levels for classification has been shown to be superior to single-level approaches (Bruzzone and Carlin, 2006; Taubenböck et al., 2010). Such a multi-level strategy on the one hand allows for more adequate object-based feature representations of all different classification targets of interest (e.g. building rooftops of different types and sizes) and on the other hand for the consideration of their spatial context, as represented by affiliated super-object information.

### 1.3. Studies on building extraction in temporary settlements

Emerging planned or spontaneous temporary settlements such as refugee camps adapt to different natural, social and political conditions. Their physical morphology differs from settlements intended to be permanent. They are composed of small-scale ground-level dwellings of different types (e.g. tents, containers, huts) and materials (e.g. plastic or metal sheet, loam or wood). High temporal dynamics (e.g. due to variations in population pressure), spatial limitations and haphazardly building (e.g. due to uncontrolled population influx) typically induce heterogeneous patterns of BUA often exhibiting high densities. These characteristics impose specific challenges for satellite-based building extraction with respect to temporary settlements.

A variety of methods focusing on the extraction of buildings in

refugee camps based on VHR optical satellite imagery has already been presented in literature, ranging from time consuming and labor intensive visual image interpretation (e.g. Bjørgo, 2000; Giada et al., 2003; Checchi et al., 2013; Spröhnle et al., 2014) to more robust and time efficient (semi-) automated methods (e.g. Giada et al., 2003; Lang et al., 2010; Kemper et al., 2011; Wang et al., 2015; Spröhnle et al., 2014, 2017). Several case studies were applied to different camp types showing promising results for computer-assisted building extraction, including pixel-based classification (e.g. Giada et al., 2003), object-based classification rule-sets (e.g. Lang et al., 2010; Spröhnle et al., 2014, 2017) and approaches based on MM operations combined with thresholding (e.g. Giada et al., 2003; Kemper et al., 2011; Heinzel and Kemper, 2014; Wang et al., 2015).

The investigation of SAR data for building extraction in settlement environments featuring the morphological characteristics of temporary settlements is still in the early stages. Spröhnle et al. (2017) explore the benefits of single polarized SAR data (i.e. a TS-X High Resolution Spotlight image with 1.5 m spatial resolution) as well as its integration with VHR optical data for single dwelling extraction based on object-based rule set classification schemes. Their results reveal that SAR data proves useful especially for the detection of metal sheet dwellings. Furthermore, they show that particularly in complex camp areas the combined use of SAR and optical data may outperform the use of either one of those data sets.

In contrast to previous studies on building extraction in temporary settlements from VHR imagery our approach integrates the concepts quoted in Section 1.2. Thereby, the combined use of VHR MS and TS-X Staring Spotlight SAR data represents a particular novelty. The proposed workflow aims at the accurate extraction of BUA by fusing relevant features in the feature space. In this course, the *object-based profile of local variation* (OPV) is introduced for information extraction from VHR SAR imagery and revealed to be specifically relevant. The *Multiple Correlation-Based Feature Selection* (MCFS), proposed for an efficient multi-view feature selection (FS) constitutes a further contribution of this study. Our results give detailed insights on the classification accuracy of the approach as well as the relevance of the deployed features. The remainder of this paper is organized as follows. Section 2 provides a detailed description of the study site and the deployed data. The related methods are pointed out in Section 3. Experimental results are presented and discussed in Section 4. Finally Section 5 concludes this paper and provides some research perspectives.

## 2. Study site and data

### 2.1. Study site

The Al Zaatari refugee camp is located in the Mafraq governorate in northern Jordan, approximately 15 km from the Syrian border (Fig. 1). It was officially opened on July 29, 2012 to host Syrian refugees fleeing the war in Syria (UNHCR, 2014). As a result of the high population dynamics, the physical camp structure changed over time (Dalal, 2014). On April 07, 2014 – at satellite data acquisition time – there were according to the UNHCR (2014) about 106,442 people registered in the camp. At that stage, the whole camp area was characterized by buildings which are heterogeneous in size, orientation and materials giving the camp a chaotic layout. Building types were mainly flat roofed, corrugated sheet metal containers and tents, made of plastic sheets. Building density varied all over the camp. In many areas, camp structures were strongly condensed with no spaces between the roofs.

### 2.2. Satellite data

With regard to the optical data, a VHR Pléiades satellite image (Fig. 1a) acquired on April 06, 2014 was used for the analysis. The original data set consists of a panchromatic image with 0.5 m spatial resolution as well as a 4 band multispectral (MS; i.e. blue, green, red,

near-infrared) stack with 2 m spatial resolution. Preprocessing steps applied to the Pléiades image included ATCOR-2 (Richter, 1996) atmospheric correction, orthorectification using a Shuttle Radar Topography Mission digital elevation model as well as pansharpening to 0.5 m spatial resolution based on the Gram-Schmidt algorithm (Laben and Brower, 2000).

The SAR data set (Fig. 1b) comprises a single (HH)-polarized TS-X Staring Spotlight X-band image acquired on April 07, 2014. The scene has been ordered in enhanced ellipsoid corrected (EEC) format, which implies a multi-look detected scene, delivered geocoded and orthorectified. This data product and the applied processing steps are described in detail in Fritz et al. (2008). The scene has been resampled from its original pixel spacing of 0.239 m to 0.25 m spatial resolution to allow for a 4-to-1 pixel-to-pixel match with the optical scene. Both images were co-registered considering the optical scene as reference.

### 2.3. Training and test data

A rigorous and unbiased statistical assessment of classification accuracy (Section 3.5.3) requires accurate and representative reference data that is spatially independent from training data (Congalton and Green, 2008; Geiß et al., 2017). Therefore, six pairs of 200 × 200 m test areas and training areas were defined (Fig. 1a, b) in a way, that i) each pair (A–F) represents a distinct pattern of BUA (building types, materials and density) and non-BUA (soil characteristics, path way and road infrastructure, walls and fences) and ii) the ensemble of all pairs is representative for the whole camp.

Pair A and B are located in the unplanned older part of the camp that is mainly characterized by close standing flat roofed sheet metal containers of rectangular shape, in many cases partly covered with blankets. As a result of self-settlement of refugees in the early stages of the camp, this district shows an unplanned layout and high dwelling density (Dalal, 2014). The pairs C–F are chosen from the newer parts of the camp, which originally was planned with a well-structured pattern of blocks of tents and containers with regular spaces between the dwellings. However, as the influx of refugees increased, they were spontaneously filling the spaces between residential blocks and infrastructure (Dalal, 2014). While the pairs D and F are characterized by a mixture of containers and groups of tents (mainly hexagonal UNHCR tents made of plastic sheets), pair E is primarily composed of containers. Pair C is dominated by containers in the northern part and UNHCR tents in the southern part (Fig. 1).

Since accurate and up-to-date ground truth was lacking, test data was extracted within the defined test areas by means of visual image interpretation of the pan-sharpened Pléiades image. Each test area comprises 160,000 pixels. All together the test data consists of 227,323 BUA and 732,677 non-BUA pixels (Fig. 1c). The training areas served for the generation of a pool of labelled segments used for the sampling of training data for classification model learning (Section 3.4).

## 3. Methods

The approach we propose (Fig. 2), addresses the specific challenges of detailed BUA detection in complex temporary settlements from VHR MS and SAR data and involves a multi-level image segmentation procedure (Section 3.1), multi-level feature calculation (Section 3.2), MCFS (Section 3.3) and supervised classification deploying a Random Forest (RF) classifier (Section 3.4). The setup pursued within the performed classification experiments is described in Section 3.5 including i) feature grouping in order to explore the benefits of the considered feature types as well as to evaluate the proposed MCFS approach, ii) a proper sampling scheme for the collection of independent training data and the quantification of the influences of training data set sizes on the classification results and iii) a rigorous accuracy assessment taking class imbalance into account.
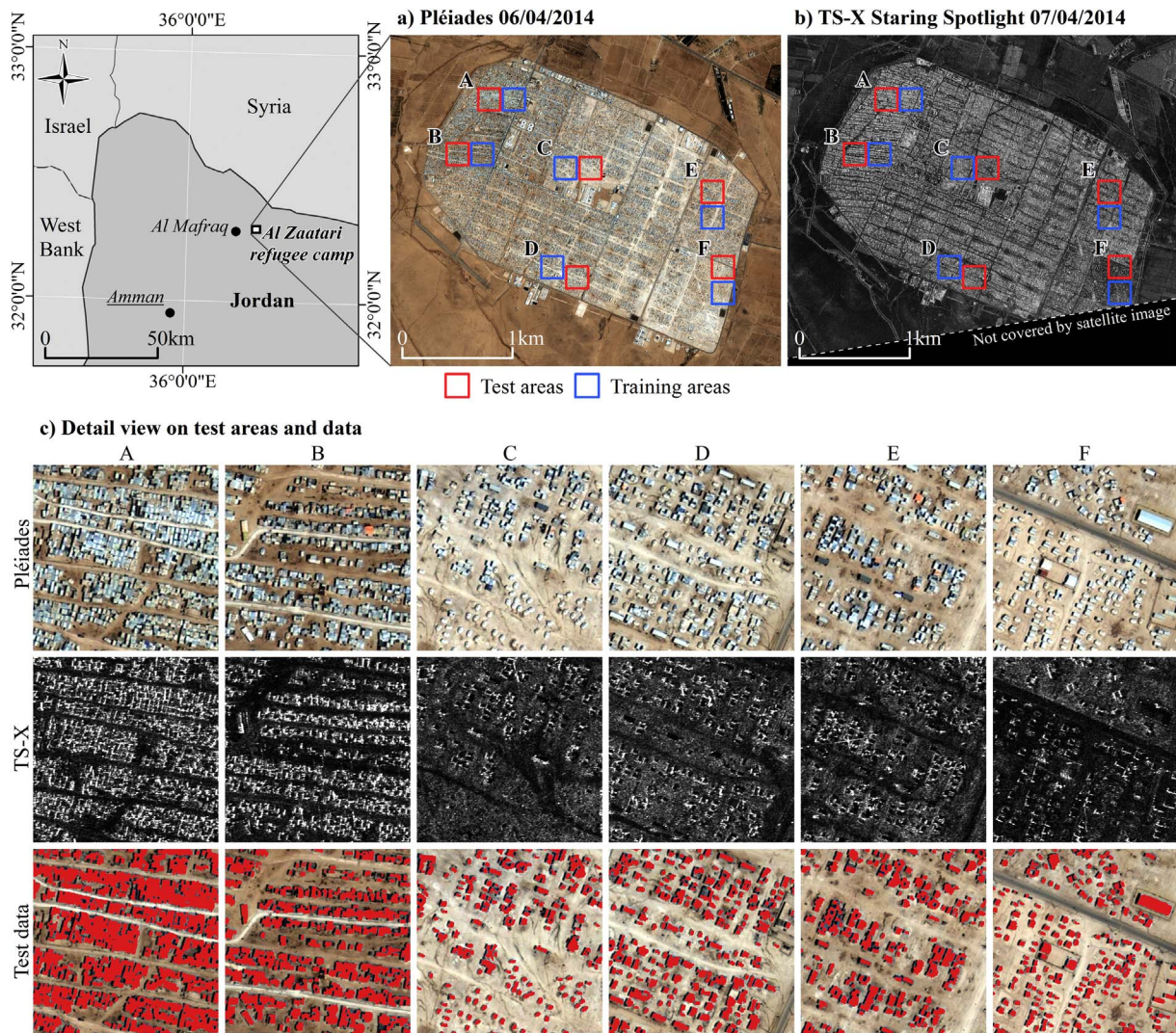
**Fig. 1.** Study site, data sets and deployed 200 × 200 m test and training areas a) Pléiades satellite data with 0.5 m spatial resolution acquired on 06/04/2014, b) TerraSAR-X Staring Spotlight satellite data with 0.25 m spatial resolution acquired on 07/04/2014, c) detail view on the test areas (A-F) and the extracted test data (BUA in red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 3.1. Image segmentation and multi-level concept

In order to capture the spectral-spatial heterogeneity of dynamic settlement environments, a sequence of multiple hierarchical segmentation levels is created based on the MS imagery. Thereby, segmentation at a generic segmentation level $l$ divides the image into $N^l$ objects $O_n^l (n = 1, 2, ..., N^l)$. The multiple segmentation levels comply with an unambiguous hierarchy by incorporating the following relation:

$$\bigcup_{O_n^l \subseteq O_m^{l+1}} O_n^l = O_m^{l+1},$$

(1)

which ensures that a considered object $l$ is included in only one object at level $l + 1$. This concept is consistent with other approaches that integrate super-segment information (e.g. Bruzzone and Carlin, 2006; Geiß et al., 2016a).

We deploy three hierarchic segmentation levels in order to exhaustively account for the BUA and its non-BUA environment (Fig. 3).
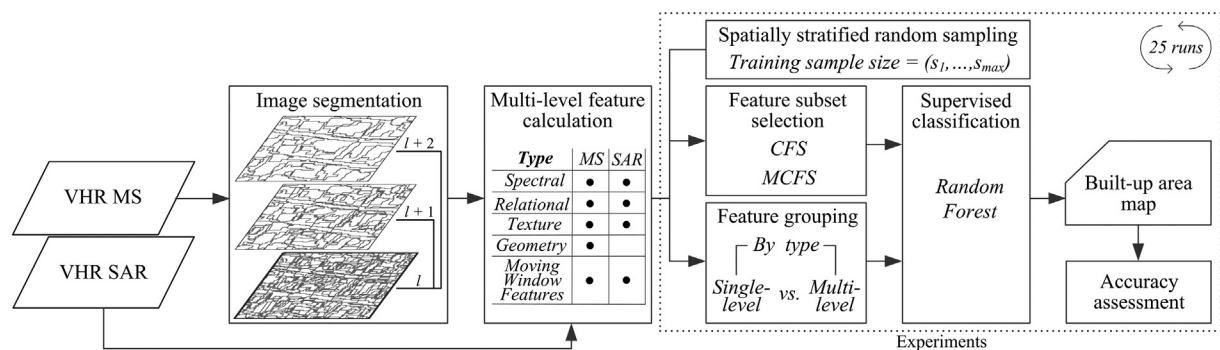


**Fig. 2.** Flowchart of the BUA extraction methodology and experiments.
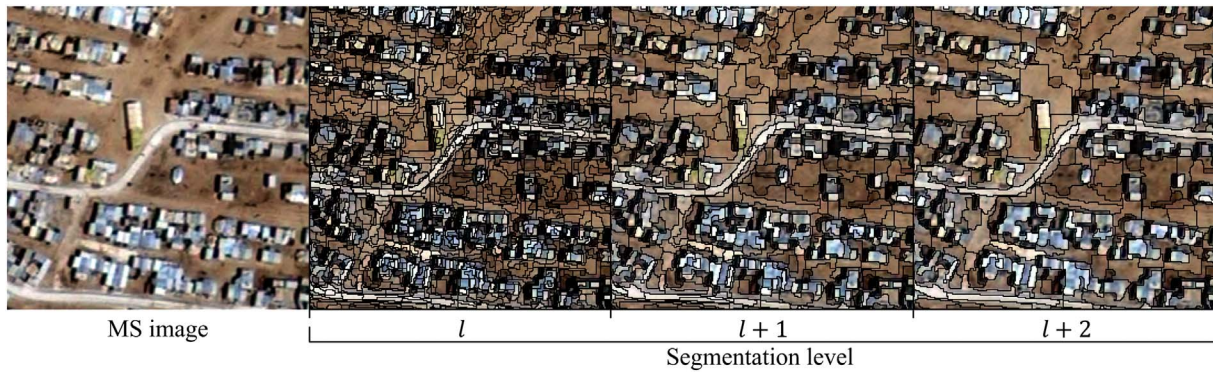
**Fig. 3.** Hierarchical creation of multiple segmentation levels.

The finest level subsequently denoted as *l* represents smallest real world objects (e.g. small buildings or homogeneous roof parts) and tends to be over-segmented with regard to medium and large BUA patches. Segmentation level *l* + 1 represents medium size real world objects (e.g. medium size buildings and BUA patches). The coarsest segmentation level *l* + 2 represents large real world objects (e.g. large buildings and BUA patches) and tends to be slightly under-segmented with regard to small and medium size BUA patches. Respecting the associated classification procedure (Section 3.4), we deploy the objects of the finest level *l* as base-entities. In this way, computation times are kept considerably lower compared to a pixel-based classification, while the details of small-scale settlement structures imaged in VHR data, are still captured.

For the creation of such a hierarchical network of segmentation levels we apply a bottom-up region-merging algorithm i.e. the Fractal Net Evolution Algorithm (FNEA; Baatz and Schäpe, 2000). Starting from individual pixels, the FNEA consecutively merges adjacent regions as long as the user-defined criteria for the maximum allowed heterogeneity of the resulting image segments are not exceeded. These criteria are expressed by the free heterogeneity weight parameters scale ($S$), color ($h_c$) and shape ($h_s$) with the shape parameter integrating smoothness ($h_{ss}$) and compactness ($h_{sc}$). Thereby, larger values of $S$ induce larger within segment heterogeneities, generally leading to larger segments (Benz et al., 2004).

After preliminary trials, we carried out a multi-segmentation procedure using the MS bands with the following scales: $S_l = 25$, $S_{l+1} = 50$, $S_{l+2} = 75$. In accordance with other object-based land cover studies (e.g. Martha et al., 2011; Geiß and Taubenböck, 2015; Geiß et al., 2016a) we set up our FNEA procedure taking into account that, in contrast to natural features, man-made features such as buildings or roads are characterized by distinct shape and size properties. Correspondingly, we place emphasis on the shape heterogeneity rather than on spectral heterogeneity and use the following parametrization: $h_s = 0.7$; $h_c = 0.3$; $h_{ss} = 0.5$; $h_{sc} = 0.5$. Several strategies for an automated and objective identification of optimized segmentations are proposed in literature (Espindola et al., 2006; Esch et al., 2008; Drăgut et al., 2010; Martha et al., 2011; Geiß et al., 2016a). Nonetheless, numerous studies have shown that accuracies of object-based supervised classifications are less dependent on segmentation (Belgiu and Drăgut, 2014), and indicate that over-segmentation is generally preferable (e.g. Stumpf and Kerle, 2011; Johnson and Xie, 2013; Rougier et al., 2016). Thus, in order to keep computational costs low, we applied segmentation without an extensive optimization. The choice of most adequate feature representations within the range of the applied segmentation levels (*l*, *l* + 1 and *l* + 2) is subsequently addressed within the FS procedure (Section 3.3).

### 3.2. Feature calculation

The objects defined by segmentation level *l* are characterized by means of an exhaustive set of descriptive features being calculated based on the multiple segmentation levels associated to the MS and SAR imagery (Table 1). Accordingly, for a generic object $O^l$ of base entity level *l* its resulting affiliated stacked multi-level feature vector $F(O^l)$ is composed of three hierarchic single-level related sub-vectors $f$ of $N$ features and can be written as:

$$F(O^l) = (f(O^l)_{n=1}^N, f(O^{l+1})_{n=1}^N, f(O^{l+2})_{n=1}^N). \tag{2}$$

Overall, $N = 179$ different features are calculated on the three segmentation levels resulting in a total vector of $N_{F(O^l)} = 537$ features. Thereby, the chosen image descriptors aim to cover the feature types typically used in state-of-the art LULC studies based on VHR data (Section 1.2). Among these, metrics of central tendency, spread, band ratios, metrics capturing the relationship of a considered object to its adjacent neighbor objects, texture features, geometric features as well as a sequence of object-based representations of pixel-based moving window (MW) operators was computed (henceforth denoted as Moving Window Features (MWF)).

#### 3.2.1. VHR multispectral features

With regard to the optical data, spectral metrics assessing the object-based central tendency and spread (i.e. *mean*, *median*, *standard deviation* and *interquartile range* (IQR)) of the MS bands as well as their first 3 *principal components* (PCs; Bishop, 2006) are extracted. In addition a pre-defined *brightness* layer (Bn) is calculated as the sum of the MS band values $C_{i(v)}$ divided by the number of bands $n_v$:

$$Bn = \frac{1}{n_v} \sum_{i=1}^{n_v} C_{i(v)}. \tag{3}$$

Bn was further used in order to compute the maximum difference between the available MS channels relative to Bn (MaxDiff; Stumpf and Kerle, 2011) as well as object-based rotation-invariant GLCM texture measures (Haralick et al., 1973; Trimble, 2014). The object-based mean share of the single MS bands compared to the sum of all MS bands is calculated as follows:

$$Sh_{C_{i(v)}} = \frac{C_{i(v)}}{\sum_{i=1}^{n_v} C_{i(v)}}. \tag{4}$$

For a more in depth conception of the spatial context of image-objects, two relational features, reflecting the *mean boarder weighted difference to darker* and *brighter adjacent neighbors* ($\overline{\Delta}_L^d$ and $\overline{\Delta}_L^b$ respectively) with regard to the MS-bands and the first 3 *PCs* are determined.

Considering an image layer L, $\overline{\Delta}_L^d$ and $\overline{\Delta}_L^b$ of an arbitrary object $O^{sl}$ related to a segmentation level *sl* with the darker neighbors $N_{O^{sl}}^d$ and the brighter neighbors $N_{O^{sl}}^b$ are computed as follows:

$$\overline{\Delta}_L^d(O^{sl}) = \frac{1}{w} \sum_{n \in N_{O^{sl}}^d} w_n(|\overline{c}_L(O^{sl}) - \overline{c}_L(n)|) \tag{5}$$

**Table 1**
Object-based features. The subscripts indicate the information layers the respective features were calculated on. The numbers of considered features are denoted in brackets.

| Type (number) | | | Features | Comment/reference |
|---|---|---|---|---|
| VHR Multispectral (115) | Spectral (51) | Central tendency, spread & ratios (37) | $\text{Mean}_{R, G, B, NIR, 1st PC, 2nd PC, 3rd PC, Brightness, MaxDiff}$<br>$\text{Median}_{R, G, B, NIR, 1st PC, 2nd PC, 3rd PC}$<br>$\text{Standard deviation}_{R, G, B, NIR, 1st PC, 2nd PC, 3rd PC}$<br>$\text{IQR}_{R, G, B, NIR, 1st PC, 2nd PC, 3rd PC}$<br>$\text{Mean share of }_{R, G, B, NIR}$<br>Mean normalized green blue index<br>Mean normalized difference vegetation index<br>Mean normalized difference water index | Eq. (4)<br>$(G - B) / (G + B)$<br>Rouse et al., 1974<br>McFeeters, 1996 |
| | | Relational (14) | $\text{Mean w. diff. to darker neighb.}_{R, G, B, NIR, 1st PC, 2nd PC, 3rd PC}$<br>$\text{Mean w. diff. to brighter neighb.}_{R, G, B, NIR, 1st PC, 2nd PC, 3rd PC}$ | Eq. (5)<br>Eq. (6) |
| | Texture (8) | GLCM (8) | $\text{Angular 2nd moment}_{Brightness}$<br>$\text{Contrast}_{Brightness}$<br>$\text{Dissimilarity}_{Brightness}$<br>$\text{Standard deviation}_{Brightness}$<br>$\text{Mean}_{Brightness}$<br>$\text{Homogeneity}_{Brightness}$<br>$\text{Entropy}_{Brightness}$<br>$\text{Correlation}_{Brightness}$ | Haralick et al., 1973 |
| | Geometry (14) | Extent (6) | Area<br>Perimeter<br>Width<br>Length<br>Length/width<br>Area/perimeter | |
| | | Shape (8) | Asymmetry<br>Shape index<br>Elliptic fit<br>Rectangular fit<br>Compactness<br>Density<br>Border index<br>Roundness | Trimble, 2014 |
| | Moving window (42) | OMPs (40) | $\text{Opening}_B$<br>$\text{Closing}_B$<br>$\text{Opening by top-hat}_B$<br>$\text{Closing by top hat}_B$ | Eq. (7); Geiß et al., 2016b |
| | | Edges (2) | $\text{Mean bright Lee Sigma edges}_B$<br>$\text{Mean dark Lee Sigma edges}_B$ | Lee, 1983 |
| VHR SAR (64) | Intensity (6) | Central tendency & spread (4) | $\text{Mean}_{SAR\ Intensity}$<br>$\text{Median}_{SAR\ Intensity}$<br>$\text{Standard deviation}_{SAR\ Intensity}$<br>$\text{IQR}_{SAR\ Intensity}$ | |
| | | Relational (2) | $\text{Mean w. diff. to darker neighb.}_{SAR\ Intensity}$<br>$\text{Mean w. diff. to brighter neighb.}_{SAR\ Intensity}$ | Eq. (5)<br>Eq. (6) |
| | Texture (8) | GLCM (8) | $\text{Angular 2nd moment}_{SAR\ Intensity}$<br>$\text{Contrast}_{SAR\ Intensity}$<br>$\text{Dissimilarity}_{SAR\ Intensity}$<br>$\text{Standard deviation}_{SAR\ Intensity}$<br>$\text{Mean}_{SAR\ Intensity}$<br>$\text{Homogeneity}_{SAR\ Intensity}$<br>$\text{Entropy}_{SAR\ Intensity}$<br>$\text{Correlation}_{SAR\ Intensity}$ | Haralick et al., 1973 |
| | Moving window (50) | OMPs (40) | $\text{Opening}_{SAR\ Intensity}$<br>$\text{Closing}_{SAR\ Intensity}$<br>$\text{Opening by top-hat}_{SAR\ Intensity}$<br>$\text{Closing by top-hat}_{SAR\ Intensity}$ | Eq. (7); Geiß et al., 2016b |
| | | OPV (10) | $\text{Local coefficient of variation}_{SAR\ Intensity}$ | Eqs. (8), (9) |
| (Total: 179) | | | R = red; G = green; B = blue; NIR = near-infrared; PC = principal component | |

$$\overline{\Delta}_L^b(O^{sl}) = \frac{1}{w} \sum_{n \in N_{O^{sl}}^b} w_n (|\overline{c}_L(O^{sl}) - \overline{c}_L(n)|) \tag{6}$$

$\overline{c}_L$ denoting the grey-level value and $w_n$ being the common border between the object $O^{sl}$ and an adjacent object $n$.

A compilation of geometric features includes several metrics capturing an objects' spatial extent and shape (Table 1).

Due to the fact that dwelling materials such as plastic as well as sheet metal exhibit a relatively high contrast to the background in the blue region of the electromagnetic spectrum (Heinzel and Kemper, 2014), optical MWF were calculated based on the blue band of the MS image. This characteristic was also utilized in previous shelter extraction studies by Jenerowicz et al. (2011) and Spröhnle et al. (2017). Optical MWF contain set-theory based MM as well as edge-extraction information. With regard to the MM-features, we deploy object-based MPs (OMPs) following the concept recently introduced by Geiß et al. (2016b). We employ the MM operators *opening* (OP), *closing* (CL), *opening by top-hat* (OTH) and *closing by top-hat* (CTH; Soille, 2004). A sequence of squared shaped MWs of increasing size (i.e. $3 \times 3$, $5 \times 5$, $7 \times 7$, $9 \times 9$, $11 \times 11$, $13 \times 13$, $15 \times 15$, $20 \times 20$, $25 \times 25$, and $30 \times 30$ pixels) is utilized as SE. We chose these sizes in consideration of the spatial resolution of the imagery as well as the morphology of temporary settlement environments. Ranging from 1.5 m to 15 m, they are able to generically model the size of the heterogeneous occurring objects. In a next step resulting grey-level values are aggregated on corresponding segmentation levels by applying the *mean* function. Accordingly, the full OMP assigned to a generic object $O^l$ of base entity level $l$ can be written as:

$$OMP(O^l) =$$
$$(\overline{x}_{OP}(O^i)_{k=1}^K, \overline{x}_{CL}(O^i)_{k=1}^K, \overline{x}_{OTH}(O^i)_{k=1}^K, \overline{x}_{CTH}(O^i)_{k=1}^K);$$
$$i \in [l, l+1, l+2] \tag{7}$$

thereby, $k$ reflects the index of the applied MW sizes from 1 to $K$, $i$ the considered segmentation levels and $\overline{x}$ the object-based mean values of the corresponding pixel values resulting from OP, CL, OTH and CTH.

Regarding the edge information, two edge layers are generated by applying the Sigma filter, an edge preserving smoothing filter introduced by Lee (1983). The first layer includes extracted edges of objects brighter than their background, whereas the second layer includes edges of objects darker than their background. Based on the sigma probability of the Gaussian distribution, the Sigma filter smooths variations in the image by replacing the center pixel of a MW with the average of those pixels within a predefined standard deviation range ($\Delta\sigma$). Neighborhood pixels outside $\Delta\sigma$ are excluded from the calculation of the sample mean. As a result, depending on $\Delta\sigma$, edges and linear features are preserved while noise is smoothed out (Lee, 1983). Within the present study $\Delta\sigma$ was set to a ± 5 standard deviations, enforcing strong edge extraction.

### 3.2.2. VHR SAR features

Consistently with the optical features, the set of SAR features comprises object-based metrics of central spread and tendency, relational context features, texture and MWF (Table 1). Thereby, the calculation is based on the SAR backscatter intensity. With regard to the MWF, the difference in spatial resolution between both images (Section 2.2) is considered by doubling the MW sizes (i.e. $6 \times 6$, $10 \times 10$, $14 \times 14$, $18 \times 18$, $22 \times 22$, $26 \times 26$, $30 \times 30$, $40 \times 40$, $50 \times 50$, and $60 \times 60$ pixels), so that resulting MWs cover the same spatial extent. In contrast to the set of optical MWF, we compute an *object-based profile of local variation* (OPV) from the SAR intensity image based on the *local coefficient of variation* (CoV) instead of Lee-Sigma edges. For an arbitrary center pixel c CoV is defined by the standard deviation and the mean of the grey-level values within a deployed MW (Eq. (8)).

$$CoV_c = \frac{\sigma_c}{\mu_c} \tag{8}$$

An entire OPV related to a generic object $O^l$ of base entity level $l$ is defined as:

$$OPV(O^l) = (\overline{x}_{CoV}(O^i)_{k=1}^K); \; i \in [l, l+1, l+2] \tag{9}$$

$k$ reflects the index of the applied MW sizes from 1 to $K$, $i$ the considered segmentation levels and $\overline{x}$ the object-based means of the CoV.

Considering a single SAR scene the CoV is directly proportional to the *speckle divergence* (Esch et al., 2010, 2012), which has already been shown highly beneficial for the delineation of urban BUA from TS-X imagery recorded in coarser spatial resolution imaging modes.

### 3.3. Features subset selection

High dimensional feature sets typically exhibit a large amount of redundancy, often show high inter-correlations and may be affected by the Hughes Phenomenon (Hughes, 1968). Furthermore, such data is likely to contain irrelevant noise inducing features. At the same time the susceptibility to learn overfitting classification models increases with dimensionality. A FS, filtering out the least promising features may attenuate the aforementioned problems. Thereby, more compact feature sets facilitate data interpretation, reduce data storage requirements and provide faster and more cost-effective models (Guyon, 2003). Supervised FS methods can be categorized into *filters, wrappers* and *embedded* approaches. Filters assess the relevance of a feature only based on the intrinsic characteristics of the data (Duch, 2006). Wrappers instead, iteratively evaluate feature sets by using accuracy estimates provided by the actual classification algorithm (Kohavi and John, 1997). Embedded methods are specific to given classification algorithms, e.g. *Recursive Feature Elimination* for Support Vector Machines (SVM; Guyon et al., 2002) and perform FS during the training process. Comparing the three FS families, filters are classifier independent and in general computationally simpler and faster than wrappers or embedded methods, since they do not incorporate classifier learning (Kohavi and John, 1997; Lal et al., 2006).

Therefore, we deploy a multivariate filter method for dimensionality reduction, namely the *Correlation-Based Feature Selection* (CFS) method introduced by Hall (1999, 2000). CFS has already shown promising results within remote sensing data based classification applications, where CFS lead to good classification results with relatively small feature sets compared to other FS approaches, e.g. in Pal and Foody (2010), Geiß et al. (2015) and Ma et al. (2015). Rather than filter methods evaluating individual features, CFS evaluates individual subsets of features based on an evaluation criterion that favors subsets with a high average feature-class correlation ($\overline{C_{fc}}$) and low average feature-feature inter-correlation ($\overline{C_{ff}}$):

$$m_S = \frac{k\overline{C_{fc}}}{\sqrt{k + k(k-1)\overline{C_{ff}}}} \tag{10}$$

$m_S$ being the merit of subset $S$ and $k$ its number of features. $\overline{C_{fc}}$ and $\overline{C_{ff}}$ are calculated by using the entropy based *symmetrical uncertainty* (Hall, 2000). The subset space is searched using a best-first algorithm (Kohavi and John, 1997) for speeding up computation performance.

The consideration of features of different types (e.g. spectral, textural and geometric features) as well as their calculation from different data sources (e.g. imagery acquired by different sensor types) increases the complexity of the entire feature space by imposing complementary subspaces (also referred to as views) with a particular physical meaning and specific statistic characteristics. In general such subspaces show a high heterogeneity and rather low redundancy among each other, whereas features within a certain sub-feature space are likely to be homogeneous and to show high redundancy (Chen et al., 2014, 2017). In order to evenly take into account different views imposed due to

**Table 2**
Composition of feature sets applied for classification. Subscripts denote included feature types. For details on the different features and feature types see Section 3.2.

| Feature sets: | Level | Features |
|---|---|---|
| $SL_{BL}$ | $l$ | $MS_{Spectral}$, $MS_{Geometry}$, $MS_{Texture}$ |
| $ML_{BL}$ | All | $MS_{Spectral}$, $MS_{Geometry}$, $MS_{Texture}$ |
| $SL_{BL,\ MWF}$ | $l$ | $MS_{Spectral}$, $MS_{Geometry}$, $MS_{Texture}$, $MS_{MWF}$ |
| $ML_{BL,\ MWF}$ | All | $MS_{Spectral}$, $MS_{Geometry}$, $MS_{Texture}$, $MS_{MWF}$ |
| $SL_{BL,\ MWF,\ SAR}$ | $l$ | $MS_{Spectral}$, $MS_{Geometry}$, $MS_{Texture}$, $MS_{MWF}$, $SAR_{All}$ |
| $ML_{BL,\ MWF,\ SAR}$ | All | $MS_{Spectral}$, $MS_{Geometry}$, $MS_{Texture}$, $MS_{MWF}$, $SAR_{All}$ |
| CFS | All | CFS(All) |
| MCFS | All | $CFS(MS_{Spectral/Central\ tendency,\ spread\ and\ ratios})$, $CFS(MS_{Spectral/Relational})$, $CFS(MS_{Geometry})$, $CFS(MS_{Texture})$, $CFS(MS_{MWF})$, $CFS(SAR_{Intensity/Central\ tendency,\ spread\ and\ ratios})$, $CFS(SAR_{Intensity/Relational})$, $CFS(SAR_{Texture})$, $CFS(SAR_{MWF})$ |

SL = single-level feature set (based on segmentation level $l$);
ML = multi-level feature set (based on all segmentation levels);
BL = baseline feature set

the different feature and sensor types in our study in a straightforward and efficient way, we propose a modification of CFS, we call *multiple CFS* (MCFS). MCFS decomposes the entire stacked feature vector into multiple predefined typological sub-feature spaces ($f_T$) and applies CFS on these $f_T$. Thereby, we assume that a considered $f_T$ implies homogeneous orders of scales with regard to feature manifestations and feature-class correlations. The stacked vector of the resulting typological feature subsets then constitutes the feature set for classifier learning. For $N$ arbitrary $f_{Ti}$ $i \in [1,...,N]$ forming the entire feature space $F$, the resulting subset for classification $MCFS(F)$ can be defined as follows:

$$MCFS(F) = (CFS_{f_{T1}},\ CFS_{f_{T2}},\ ...,CFS_{f_{TN}}) \qquad (11)$$

with $CFS_{f_{Ti}}$ denoting the CFS subsets of $f_{Ti}$. An additional effect of the MCFS strategy are less computational costs due to a reduction of the feature subset search space, arising from less possible feature combinations. The decomposition scheme applied within this study is itemized in Table 2.

### 3.4. Classification

*Random forest* (RF), a non-parametric decision-tree based ensemble learning algorithm, was chosen for classification (Breiman, 2001). RF is able to cope with the notable amount of redundancy as well as the high dimensionality inherent in the feature sets (Geiß et al., 2015; Wurm et al., 2017). These are particularly induced by the consecutive window sizes of the MWF and the multi-level feature calculation approach of our methodology (Dalla Mura et al., 2010; Geiß et al., 2016b). In order to minimize the correlation between the trees of the ensemble, RF grows each tree on a bootstrapped sample of the input data at each node affiliated with a new random subset of features drawn from the input feature space. The final predictions are determined by a majority vote of the single trees. In this manner RF attains accurate generalization abilities, robustness to noise as well as computational efficiency (Breiman, 2001; Fernandez-Delgado et al., 2014). The two tuning parameters to be defined setting up a RF are the number of trees forming the ensemble ($n_{tree}$) and the number of features randomly selected at each node ($m_{try}$). We set $n_{tree}$ to 500 in order to provide a reliable error estimate, while maintaining the computation times in a reasonable range. This is in a good agreement with the RF parameter study performed by Genuer et al. (2008). $m_{try}$ is constantly set to $\sqrt{p}$ ($p$

denoting the number of input features), which generally yields near optimum classification results (Genuer et al., 2008).

### 3.5. Experimental setup

#### 3.5.1. Feature grouping

To assess the added value in classification accuracy attained through the multi-level approach as well as through the integration of additional feature types, the classification procedure was carried out and evaluated for six different feature groups composed ensuing from a baseline feature set (BL). Thereby, BL includes the features typically employed in the object-based analysis of VHR MS satellite imagery, i.e. spectral, geometry and texture features (e.g. Stumpf and Kerle, 2011; Ma et al., 2015; Geiß et al., 2015; Leichtle et al., 2017a). Then, MWF and SAR features are added in two steps. Results are compared to the outcomes of the CFS and MCFS. The feature sets applied for classification are listed in Table 2, SL indicating the involvement of single-level and ML indicating the involvement of multi-level feature representations.

#### 3.5.2. Sampling of training data

An independent training data pool of 5700 labelled level $l$ segments (i.e. 2750 BUA and 2750 non-BUA samples) which are equally distributed across the six training areas (A-F; Section 2.3; Fig. 1) served for a random sampling of training data. The implemented sampling procedure draws the training data stratified with regard to the training areas (*spatial stratification*) as well as to the a priori class probabilities (*land cover related stratification*). The impact of the reduction of prior knowledge on classification accuracy was assessed by incrementally decreasing the amount of training samples. This was done with the intention to prospectively reduce time consuming and economically expensive reference data acquisition. Hence, ensuing from a maximum size ($s_{max}$) of 5004 samples training set sizes $s$ amount to $10^{1.90}$, $10^{1.80}$, ..., $10^{1.00}$, $10^{0.75}$, $10^{0.50}$, $10^{0.25}$ and $10^{0.05}$% of $s_{max}$ resulting in training data sets of 15 different sizes. Within an iteration of classifications based on the consecutive sizes of training data it was ensured that each set of training samples $s$ is also contained in the corresponding subsequent training sets with a larger number of samples. In this manner an unbiased quantification of the effect of training set size on classification performance is provided.

#### 3.5.3. Accuracy assessment

The accuracy of remote sensing data-based thematic maps is generally evaluated by comparing the results of the automated classification with reference data meaning to represent reality (Congalton and Green, 2008; Foody, 2002). For binary classifications *True Positive* (TP), *False Positive* (FP), *True Negative* (TN) and *False Negative* (FN) values are determined allowing for the calculation of several well-established accuracy measures such as *Precision* (PR, Eq. (12)), *Recall* (RC, Eq. (13)), *Specificity* (SP, Eq. (14)), *Overall Accuracy* (OA, Eq. (15)), $F_1$-score ($F_1$, Eq. (16)) as well as Cohen's *Kappa Coefficient* ($\kappa$, Eq. (17); Cohen, 1960), which are examined in this study. Additionally, we calculate the *True Skill Statistic* (TSS, Eq. (18); Allouche et al., 2006). Unlike $\kappa$, TSS is mostly insensitive to class-imbalance, which is usually exhibited for settlement areas (Klotz et al., 2016; Leichtle et al., 2017b) and also existent within the present data (Section 2.3). Accuracy is assessed based on the test data introduced in Section 2.3 as reference.

$$PR = \frac{TP}{TP + FP} \qquad (12)$$

$$RC = \frac{TP}{TP + FN} \qquad (13)$$
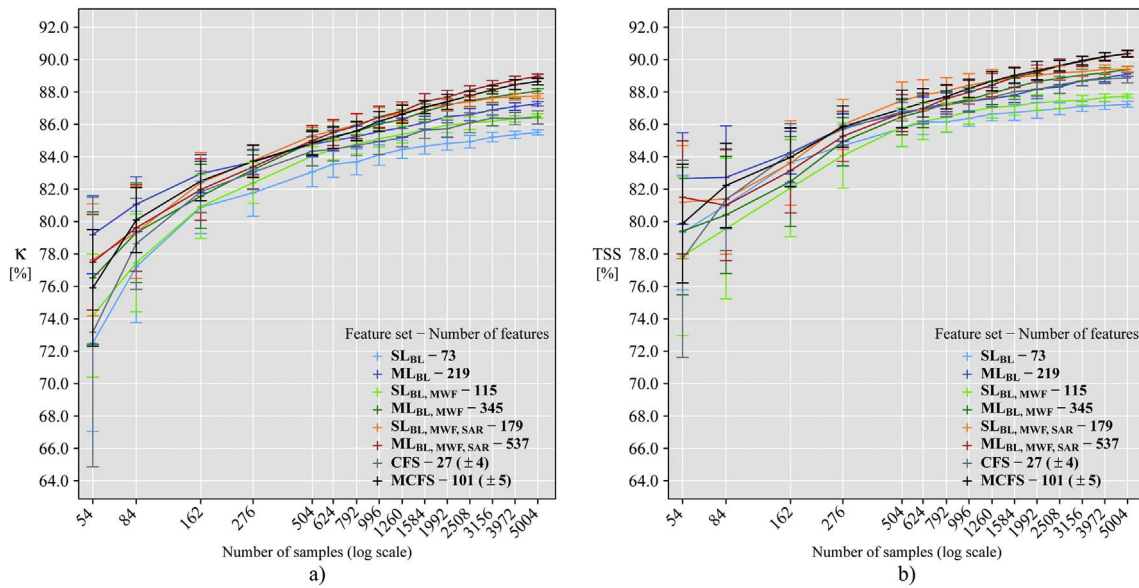
$$SP = \frac{TN}{FP + TN} \qquad (14)$$

**Fig. 4.** κ (a) and TSS (b) classification accuracies of the RF classifier trained on six different feature groups (Table 2) as well as the results of the CFS and MCFS approach as a function of the number of samples. The plots highlight the mean accuracy values over 25 independent runs of the considered classification settings and respective standard deviations. Abscissae are scaled logarithmically.

$$OA = \frac{TP + TN}{TP + FP + TN + FN} \tag{15}$$

$$F_1 = 2 \times \frac{PR \times RC}{PR + RC} \tag{16}$$

$$\kappa = \frac{OA - \frac{(TP+FP)(TP+FN)+(FN+TN)(TN+FP)}{(TP+FP+TN+FN)^2}}{1 - \frac{(TP+FP)(TP+FN)+(FN+TN)(TN+FP)}{(TP+FP+TN+FN)^2}} \tag{17}$$

$$TSS = RC + SP - 1 \tag{18}$$

## 4. Results and discussion

### 4.1. Classification results

#### 4.1.1. Overall evaluation

Fig. 4 depicts the mean classification performances (Fig. 4a: κ; b: TSS) over 25 independent runs and corresponding standard deviations based on the different applied feature sets (Table 2) as functions of the number of input training samples. The legend indicates the feature set composition as well as the number of underlying features. Considering the FS based classification results (i.e. CFS and MCFS) mean feature subset sizes of the 2508 training sample trials and respective standard deviations are itemized. In general, it can be observed that mean accuracy values increase with an increasing number of training samples. Thereby, the slope of the curves decreases, reflecting a proceeding saturation with regard to gains in classification accuracy. Simultaneously, standard deviations decrease, which indicates that the models become more robust. Finally all models culminate in excellent accuracy values ranging from 85.5% ($SL_{BL}$) to 89.0% ($ML_{BL, MWF, SAR}$) in terms of mean κ and 87.2% ($SL_{BL}$) to 90.4% (MCFS and $ML_{BL, MWF, SAR}$) in terms of mean TSS. The functions, of the different feature sets highlight that adding of MS MWF as well as the integration of SAR features induces a boost with regard to achievable classification accuracies. In addition, when comparing SL feature sets with their ML counterparts it can be seen that the latter achieve better results. This demonstrates the benefits of multi-level feature representations for classification and is in accordance with the findings of previous studies (e.g. Geiß et al., 2016a; Johnson and Xie, 2013; Bruzzone and Carlin, 2006). With regard to the feature set $ML_{BL, MWF, SAR}$, however, a relatively large sample (> 1584 samples) is

required until this effect occurs. $ML_{BL}$ performs best with respect to very small training set sizes (< 276 samples) and achieves e.g. a mean κ value of 81.1% and a mean TSS value of 82.7% already with 84 samples.

Fig. 4 further reveals that surpassing an amount of 1584 training samples, $ML_{BL, MWF, SAR}$ and the MCFS subsets outperform the other feature sets. From that point on the MCFS feature subsets attain nearly the same accuracy values as $ML_{BL, MWF, SAR}$ with regard to κ and equal or slightly better accuracy values considering TSS. Thereby MCFS deploys < 20% of the amount of features than $ML_{BL, MWF, SAR}$. E.g., training sets containing 2508 samples affiliated with the $ML_{BL, MWF, SAR}$ feature set result in a mean κ value of 88.1% and a TSS value of 89.6% with 537 features, while MCFS performs with a mean κ value of 87.8% and a mean TSS value of 89.6% with a mean size of 101 ( ± 5) features. This corroborates the ability of the MCFS approach for selecting subsets containing the features most relevant for classification.

Comparing the results of the two FS approaches, MCFS feature subsets outperform the CFS feature subsets with regard to mean κ and mean TSS. Even so, CFS throughout provides fairly good accuracy with a remarkably small number of features (see also Section 4.2). In order to get further information on the potential of the SAR features we also evaluated the results of classifications based on training sets affiliated with the ML SAR features only (Fig. 5). Trials with 2508 training samples resulted in mean κ of 63.3% ( ± 0.5) and mean TSS of 70.1% ( ± 0.5).

#### 4.1.2. Test area-specific evaluation

The results of a test area specific evaluation of the classification accuracies achieved by means of 2508 training samples affiliated with the ML optical features, the ML SAR features, all features together as well as the MCFS subsets are shown in Fig. 5. The plot reveals that the magnitudes of the multi sensor accuracies are mainly driven by the MS features. Additionally, it demonstrates that not all test site specific accuracies equally benefit from the integration of SAR features. The same accounts for the drawing of MCFS feature subsets.

While both, the additional integration of SAR features and the use of MCFS subsets positively influence the accuracy values for test areas A and B, it is apparent that the high accuracy values for test areas C–F can be simply reached by the optical image information. Overall, however, $ML_{BL, MWF, SAR}$ and MCFS perform best.
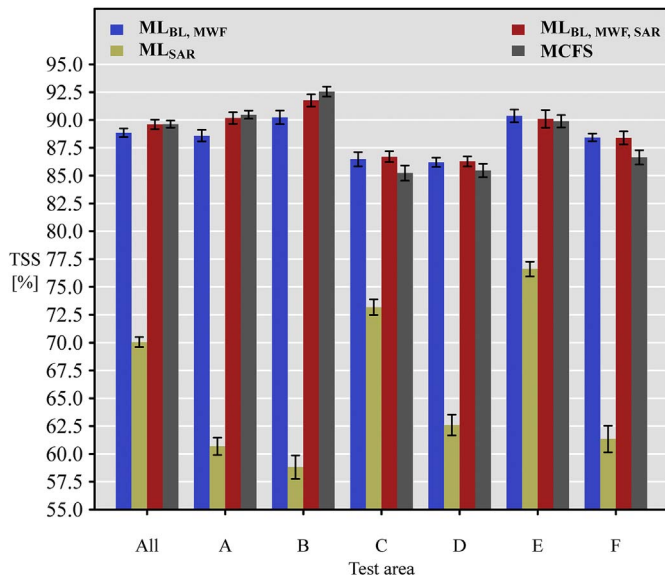
**Fig. 5.** Comparison of TSS accuracy values achieved with 2508 training samples based on all optical features ($ML_{BL, MWF}$), all SAR features ($ML_{SAR}$), all optical and SAR features together ($ML_{BL, MWF, SAR}$) as well as the MCFS subsets. The bars depict mean TSS values over 25 runs, the error bars indicate corresponding standard deviations.

In order to provide additional findings on the relationship between distinct BUA patterns within the Al Zaatari camp, classification accuracy and the amount of training data, Fig. 6 shows the mean accuracies (Fig. 6a: κ; b: TSS) as well as respective standard deviations for 25 independent MCFS trials as a function of training set sizes assessed for the different test areas (A–F) separately. Correspondingly, Table 3 itemizes the mean accuracy values and standard deviations for 25 MCFS classification runs with a training set of 2508 samples. Herein accuracy values for the conjoint evaluation and for each test area are listed separately. The classification result holding the highest κ value of these runs is visualized in Fig. 7 by means of spatially distributed TPs, FPs, TNs and FNs. The figures reveal that the test areas A and B were classified best attaining κ and TSS values beyond 90.0%. These test sites

represent the old part of the camp with high dwelling densities (mainly close standing containers, dusty tents and tarpaulins).

Having a closer look at the MS image (Fig. 1c), it is apparent that there are three main factors influencing the classification result: Dwelling density, dwelling material and soil properties. The test areas with the highest accuracy values (A, B and E) are mainly characterized by sheet metal containers, whereas areas with lower accuracy rates are either partly dominated by tents (as in case of C, southern part) or showing a mixture of containers and tents (D and F). Furthermore, test areas C and D show a bright soil type leading to less contrast between dwellings and the background and thus causing misclassifications between these classes (Fig. 7).

### 4.2. Insights on selected feature subsets

Fig. 8 shows the relationship between the sizes of the feature subsets resulting from the application of CFS and MCFS and the number of input training samples. CFS feature subsets constantly contain less than a mean of 32 ( ± 5) features (i.e. 5.9% of the input feature set). MCFS feature subsets increase continuously until a plateau is reached at 504 features. Henceforth, mean feature subset sizes stabilize remaining in a range between 93 ( ± 6) and 103 ( ± 4) features (i.e. between 17.2% and 19.3% of the input feature set). For small training sets with ≤537 samples (i.e. the total amount of input features; typically referred to as 'small n large p scenario'), which are challenging for some classifiers (Trunk, 1979), both, MCFS as well as CFS almost maintain the resulting number of features smaller than the number of training samples.

Besides the sizes of the subset, we also examined their composition in order to determine the features of most relevance for the temporary settlement classification. A look inside the MCFS subset of an input training set reveals the features of the considered types most correlating with the target classes while exhibiting relatively low feature-feature inter-correlations (Section 3.3). Fig. 9 illustrates the features chosen by MCFS in > 50% of 25 independent FS trials based on 2508 randomly drawn samples (Section 3.5.2) and corresponding to the classification accuracies itemized in Table 3 (resulting in a mean overall κ and TSS accuracy value of 87.8% and 89.6% respectively). We assume that the numbers of choices of the ascertained features reflect their impact on the classification accuracy and therefore can be seen as a measure of
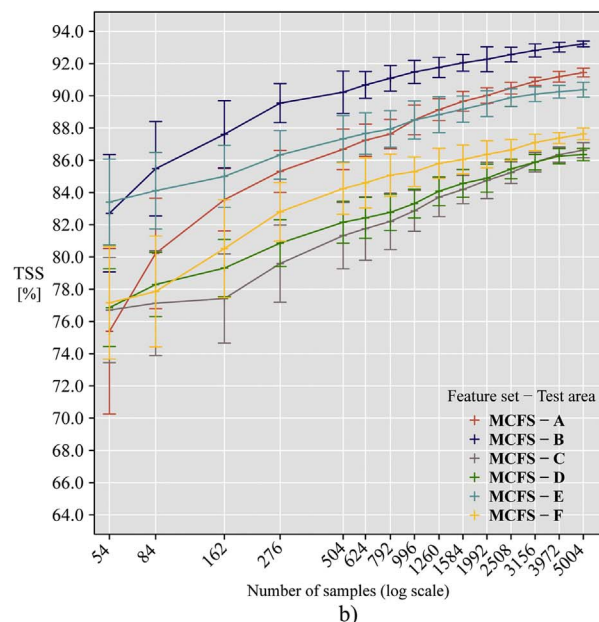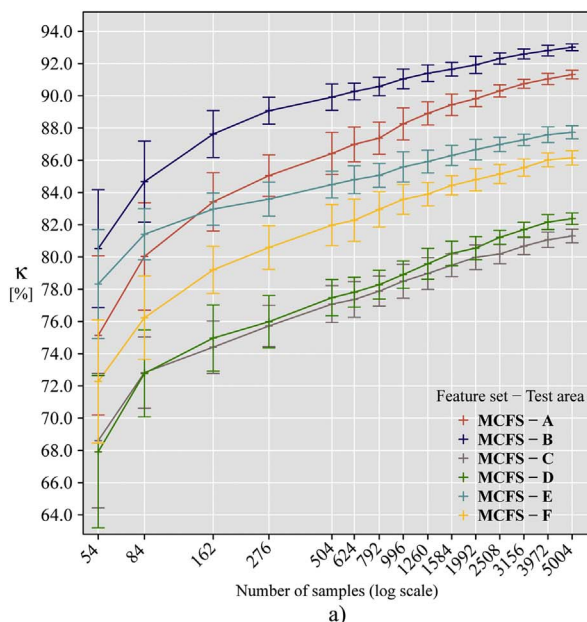


**Fig. 6.** κ (a) and TSS (b) classification accuracies of the RF classifier trained on the MCFS feature subsets as a function of the number of samples for the different test areas (A–F). The colored crosses highlight the mean accuracy values over 25 independent runs of the considered classification settings. The error bars depict respective standard deviations. Abscissae are scaled logarithmically.

**Table 3**

Mean accuracies [%] over 25 classification runs obtained for all test areas together as well as for the single test areas separately with an amount of 2508 training samples affiliated with MCFS feature subsets. Corresponding standard deviations are shown in brackets.

| Test area | Accuracy measures [%] | | | | | | |
|---|---|---|---|---|---|---|---|
| | PR | RC | SP | OA | $F_1$ | κ | TSS |
| All | 88.1 ( ± 0.6) | 93.5 ( ± 0.5) | 96.1 ( ± 0.2) | 95.5 ( ± 0.1) | 90.8 ( ± 0.2) | 87.8 ( ± 0.3) | 89.6 ( ± 0.3) |
| A | 93.5 ( ± 0.6) | 96.4 ( ± 0.5) | 94.0 ( ± 0.6) | 95.2 ( ± 0.2) | 94.9 ( ± 0.2) | 90.3 ( ± 0.4) | 90.5 ( ± 0.4) |
| B | 94.1 ( ± 0.7) | 95.0 ( ± 0.7) | 97.6 ( ± 0.3) | 96.8 ( ± 0.1) | 94.5 ( ± 0.2) | 92.3 ( ± 0.3) | 92.6 ( ± 0.4) |
| C | 77.7 ( ± 1.2) | 89.3 ( ± 0.9) | 95.9 ( ± 0.3) | 95.0 ( ± 0.2) | 83.1 ( ± 0.5) | 80.2 ( ± 0.6) | 85.2 ( ± 0.7) |
| D | 80.0 ( ± 0.8) | 91.0 ( ± 0.8) | 94.5 ( ± 0.3) | 93.8 ( ± 0.2) | 85.1 ( ± 0.3) | 81.2 ( ± 0.4) | 85.5 ( ± 0.6) |
| E | 85.9 ( ± 0.8) | 93.2 ( ± 0.7) | 96.7 ( ± 0.2) | 96.1 ( ± 0.1) | 89.4 ( ± 0.4) | 87.0 ( ± 0.4) | 89.9 ( ± 0.6) |
| F | 85.7 ( ± 1.2) | 89.4 ( ± 0.8) | 97.3 ( ± 0.3) | 96.1 ( ± 0.2) | 87.5 ( ± 0.5) | 85.1 ( ± 0.6) | 86.6 ( ± 0.5) |

PR = *Precision*; RC = *Recall*; SP = *Specificity*; OA = *Overall Accuracy*; $F_1$ = *$F_1$-Score*; κ = *Kappa Coefficient*; TSS = *True Skill Statistic*.
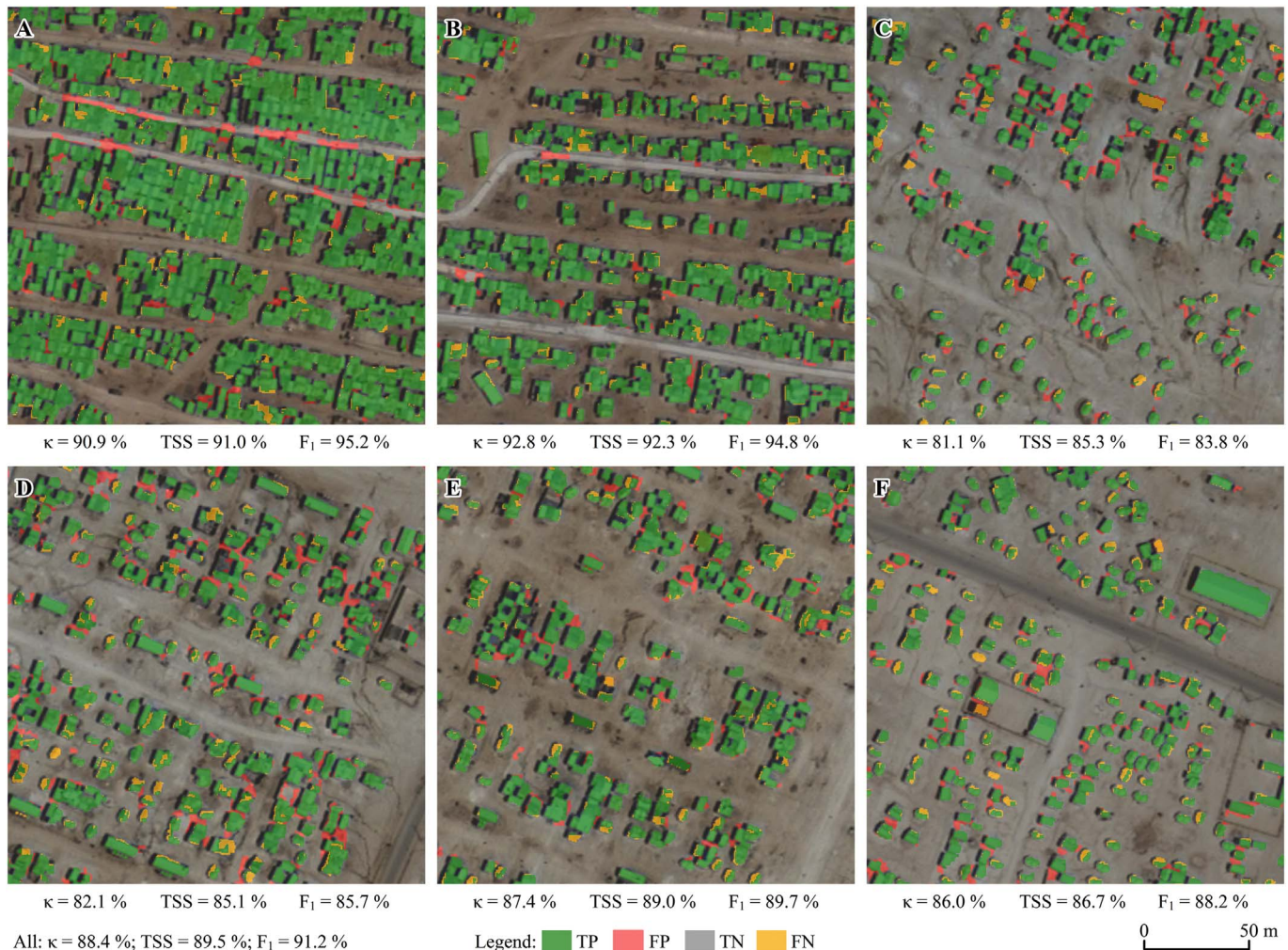


**Fig. 7.** Classification result in terms of TPs, FPs, TNs and FNs (see legend) for the six test areas with the highest overall κ value (κ = 88.4%) obtained out of 25 classification runs with 2508 training samples affiliated with MCFS feature subsets and corresponding accuracy measures (κ, TSS and $F_1$). Conjoint accuracy measures for all test areas (A–F) are given at the bottom left.

importance. First, it can be stated that features calculated based on all considered segmentation levels are selected. Additionally it can be noticed that various features evolve their relevance specifically on certain segmentation levels.

VHR MS first order spectral measures of the *2nd PC* of the MS bands as well as the *NDVI* were chosen most often from the group of MS central tendency spread and ratios and turn out to be relevant at all segmentation levels. Large rates of choice for *2nd PC* features out of the group of relational image descriptors approve its high correlation with the BUA class. Furthermore the blue channel spectral information

shows up to be notably relevant (*median, standard deviation, mean weighted difference to darker neighbor*). This is in line with the findings of previous shelter extraction studies where especially the short wave region of the optical electromagnetic spectrum proved beneficial for refugee camp shelter detection (e.g. Heinzel and Kemper, 2014; Jenerowicz et al., 2011; Spröhnle et al., 2017). Regarding the MS texture measures, *GLCM mean* is identified to be particularly distinctive. Within the group of geometry features the extent measures *area, length* as well as the *area to perimeter* ratio have high numbers of choices across all segmentation levels. The shape measures *shape index* and *density* are
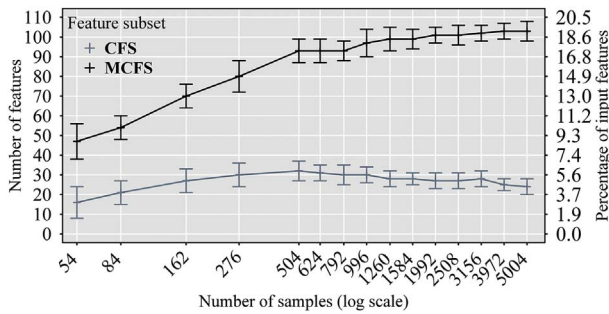
**Fig. 8.** Sizes of CFS and MCFS feature subsets as a function of the number of training samples given in absolute numbers (left ordinate) and in percentage (right ordinate). The cross-marks depict mean values over 25 independent runs; the bars indicate corresponding standard deviation values. The abscissa is scaled logarithmically.

revealed to contribute especially on the coarser segmentation levels $l + 1$ and $l + 2$. This is reasonable, because these levels feature less over-segmentation (Section 3.1) and thus exhibit a more precise delineation of the actual buildings or building patches. From the group of optical MWF the *mean bright Lee Sigma edges* and the *OTH OMP* show a substantial importance.

Regarding the SAR intensity measures of central tendency and spread, *median, standard deviation* and *IQR* show a particular relevance. The *mean difference to darker neighbors* based on the levels $l$ an $l + 1$ is chosen in each of the 25 runs. *GLCM angular 2nd moment* as well as *GLCM entropy* stands out from the SAR texture features. The *OPV* as well as the *CTH OMP* are chosen most often from MWF.

Independently from the satellite sensor, measures of spread are particularly relevant on the two coarser segmentation scales $l + 1$ and $l + 2$. That is plausible, since coarser segment scales obviously exhibit wider ranges of inner segment heterogeneities making such measures more distinctive.

## 5. Conclusion and outlook

This study presents a generic object-based workflow for feature calculation, data fusion and classification within a machine learning framework. The approach was applied for detailed BUA mapping in a complex temporary settlement using VHR multi-sensor (MS and SAR) satellite imagery. Spectral-spatial descriptive features calculated on multiple segmentation levels are considered for learning a RF classifier. In this course, the OPV was introduced for information extraction from VHR SAR data and MCFS was proposed for the selection of the most relevant subset out of the entire feature vector. Attaining accuracy values beyond 80% in terms of κ and TSS, even in complex settlement configurations, the experimental results prove the potential of the method. Thereby, the capability of RF to cope with high dimensional data has been verified. Running classifications on different feature groups (i.e. BL, MWF and SAR), it has been shown that each of the applied feature groups is gainful with regard to resulting accuracies. The same accounts for the ML strategy for the calculation of descriptive features which outperforms the use of a single segmentation level. The experimental results further revealed that the MCFS feature subsets achieve similar high classification accuracy values as the full feature vector. The examination of the relationship between the number of
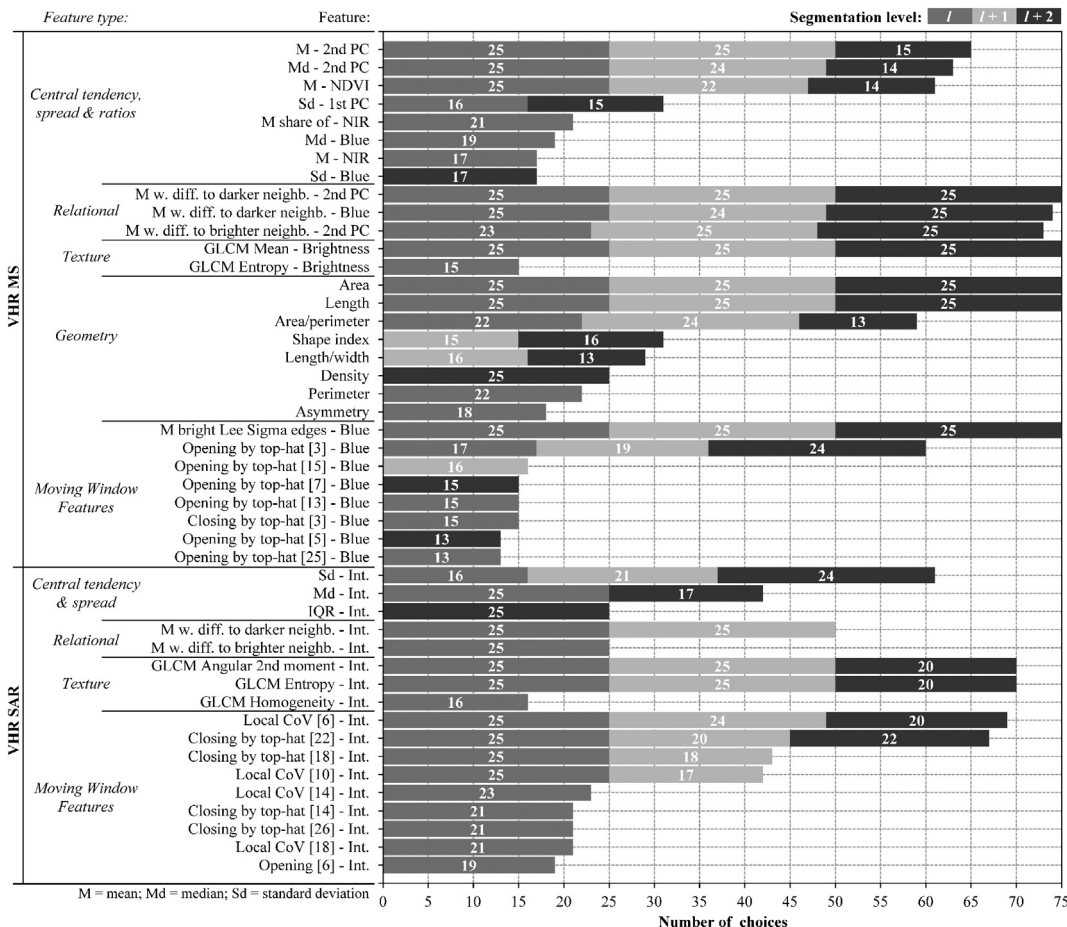


**Fig. 9.** Feature importance analysis: Features selected by MCFS in > 50% of 25 runs (i.e. ≥13 times) with 2508 training samples. The overall bar length indicates the overall number of choices of a feature without discriminating between the underlying segmentation levels. The colors of the stacked bars depict underlying segmentation levels (i.e. $l, l + 1, l + 2$; see legend on top). The actual number of choices of the single features is given in white color in the center of the single bars. [.] indicates the kernel size corresponding to a MWF.

training samples and input features showed that, after an initial rise, the amount of input features selected for classification stabilizes. Resulting MCFS subset sizes remain smaller or equal than a fifth of the input features. Preventing from "small n large p scenarios" the potential of the MCFS approach to substantially reduce dimensionality, without a significant loss in classification accuracy was demonstrated. A detailed analysis of MCFS feature sets allowed for identifying the most relevant features from the considered typological feature groups. Thereby, insights on the relationship between the relevance of a certain feature and its underlying segmentation level were highlighted. OPV features were among the most selected SAR MWF and thus are indicated to be particularly important for BUA extraction using VHR SAR data.

The presented workflow is designed in a modular manner and easily allows for modifications. Since the MCFS is classifier independent future experiments could asses the classification performance resulting from its combination with other state-of-the-art statistical learning algorithms such as a SVM (Cortes and Vapnik, 1995) or a Rotation Forest (Rodriguez et al., 2006). Additionally, the integration of active learning (e.g. Tuia et al., 2011) schemes should be investigated. Such methods guide the user within the sampling process to efficiently collect the most informative labelled samples and thereby reduce the number of training samples required to achieve high magnitude classification accuracies. This might be particularly useful within time-critical rapid mapping activities in the course of ongoing crisis events (Voigt et al., 2007). The application of semi-supervised learning could be tested to also take account of the unlabelled data within the process of classifier learning which has shown to be effective for addressing classification problems characterized by sparse or biased sets of initial training data (Persello and Bruzzone, 2014). With regard to available satellite sensors or SAR imaging modes, future research should investigate i) enhanced spatial and spectral resolution VHR optical data, e.g. from the WV-3 (0.31 m/panchromatic; 1.24 m/8 MS-bands) mission as well as ii) multi-polarized SAR data, e.g. TS-X High Resolution Spotlight dual pol with an azimuth and a ground range resolution up to 2.2 m and 1.2 m respectively, which may offer further potential for BUA detection in temporary settlements.

The proposed approach allows for a detailed automated extraction of BUA in a generic and data-driven way. Based on this information, in conjunction with estimated area-based occupancy rates, an approximate figure of the number of people living in a camp and their spatial distribution pattern can be calculated. The provision of such information is essential to obtain a better understanding of the situation in a camp and draw conclusions for further developments (UNHCR, 2007; Ehrlich et al., 2009).

## Acknowledgements

## References

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43 (6), 1223–1232. http://dx.doi.org/10.1111/j.1365-2664.2006.01214.x.

Baatz, M., Schäpe, A., 2000. Multiresolution segmentation - an optimization approach for high quality multi-scale image segmentation. In: Strobl, J., Blaschke, T., Griesebner, G. (Eds.), Angewandte Geographische Informations-Verarbeitung XII. Wichmann Verlag, Karlsruhe, pp. 12–23.

Ban, Y., Jacob, A., Gamba, P., 2015. Spaceborne SAR data for global urbanmapping at 30m resolution using a robust urban extractor. ISPRS J. Photogramm. Remote Sens. 103, 28–37. http://dx.doi.org/10.1016/j.isprsjprs.2014.08.004.

Belgiu, M., Drăguţ, L., 2014. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. ISPRS J. Photogramm. Remote Sens. 96, 67–75. http://dx.doi.org/10.1016/j.isprsjprs.2014.07.002.

Benediktsson, J.A., Pesaresi, M., Amason, K., 2003. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. IEEE Trans. Geosci. Remote Sens. 41 (9), 1940–1949. http://dx.doi.org/10.1109/TGRS.2003.814625.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multiresolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS J. Photogramm. Remote Sens. 58 (3–4), 239–258. http://dx.doi.org/10.1016/j.isprsjprs.2003.10.002.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, New York (738 pp).

Bjørgo, E., 2000. Using very high spatial resolution satellite sensor imagery to monitor refugee camps. Int. J. Remote Sens. 21 (3), 611–616. http://dx.doi.org/10.1080/014311600210786.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS J. Photogramm. Remote Sens. 65 (1), 2–16. http://dx.doi.org/10.1016/j.isprsjprs.2009.06.004.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. http://dx.doi.org/10.1023/A:1010933404324.

Brunner, D., Bruzzone, L., Ferro, A., Fortuny, J., Lemoine, G., 2008. Analysis of the double bounce scattering mechanism of buildings in VHR SAR data. In: Proc. SPIE7109, Image and Signal Processing for Remote Sensing XIV, 71090Q, http://dx.doi.org/10.1117/12.801670.

Bruzzone, L., Carlin, L., 2006. A multilevel context-based system for classification of very high spatial resolution images. IEEE Trans. Geosci. Remote Sens. 44 (9), 2587–2600. http://dx.doi.org/10.1109/TGRS.2006.875360.

Carleer, A.P., Wolff, E., 2006. Urban land cover multi-level region-based classification of VHR data by selecting relevant features. Int. J. Remote Sens. 27 (6), 1035–1051.

Checchi, F., Stewart, B.T., Palmer, J.J., Grundy, C., 2013. Validity and feasibility of a satellite imagery-based method for rapid estimation of displaced populations. Int. J. Health Geogr. 12 (1), 1–12. http://dx.doi.org/10.1186/1476-072X-12-4.

Chen, X., Li, H., Gu, Y., 2014. Multiview feature selection for very high resolution remote sensing images. In: IEEE 2014 Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control, pp. 539–543. http://dx.doi.org/10.1109/IMCCC.2014.116.

Chen, X., Zhou, G., Chen, Y., Shao, G., Gu, Y., 2017. Supervised multiview feature selection exploring homogeneity and heterogeneity with ℓ1,2-norm and automatic view generation. IEEE Trans. Geosci. Remote Sens. 55 (4), 2074–2088. http://dx.doi.org/10.1109/TGRS.2016.2636329.

Chini, M., Pacifici, F., Emery, W.J., 2009. Morphological operators applied to X-band SAR for urban land use classification. In: 2009 IEEE International Geoscience and Remote Sensing Symposium, Cape Town, pp. IV-506–IV-509. http://dx.doi.org/10.1109/IGARSS.2009.5417424.

Cohen, J., 1960. A coefficient of agreement of nominal scales. Educ. Psychol. Meas. 20 (1), 37–46. Available online: http://journals.sagepub.com/doi/abs/10.1177/001316446002000104, Accessed date: 7 June 2017.

Congalton, R.G., Green, K., 2008. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. CRC Press, Boca Raton, FL, USA (200 pp).

Cortes, C., Vapnik, V., 1995. Support vector networks. Mach. Learn. 20 (3), 273–297. http://dx.doi.org/10.1023/A:1022627411411.

Dalal, A., 2014. Camp cities between planning and practice. Mapping the urbanisation of Zaatari Camp. Stuttgart University & Ain Shams University. Available online: http://iusd.asu.edu.eg/wp-content/uploads/2015/11/2ndInt_Dalal1.pdf, Accessed date: 7 June 2017.

Dalla Mura, M., Benediktsson, J.A., Waske, B., Bruzzone, L., 2010. Morphological attribute profiles for the analysis of very high resolution images. IEEE Trans. Geosci. Remote Sens. 48 (10), 3747–3762. http://dx.doi.org/10.1109/TGRS.2010.2048116.

Drăguţ, L., Tiede, D., Levick, S.R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. Int. J. Geogr. Inf. Sci. 24 (6), 859–871. http://dx.doi.org/10.1080/13658810903174803.

Du, P., Samat, A., Waske, B., Liu, S., Li, Z., 2015. Random Forest and Rotation Forest for fully polarized SAR image classification using polarimetric and spatial features. ISPRS J. Photogramm. Remote Sens. 105, 38–53. http://dx.doi.org/10.1016/j.isprsjprs.2015.03.002.

Duch, W., 2006. Filter methods. In: Guyon, I., Nikravesh, M., Gunn, S., Zadeh, L.A. (Eds.), Feature Extraction: Foundations and Applications. Springer Berlin Heidelberg, 978-3-540-35488-8, pp. 89–117. http://dx.doi.org/10.1007/978-3-540-35488-8_4.

Ehrlich, D., Lang, S., Laneve, G., Mubareka, S., Schneiderbauer, S., Tiede, D., 2009. Can earth observation help to improve information on population? In: Jasani, B., Pesaresi, M., Schneiderbauer, S., Zeug, G. (Eds.), Remote Sensing from Space. Springer, Dordrecht, pp. 211–237. http://dx.doi.org/10.1007/978-1-4020-8484-3_14.

Esch, T., Thiel, M., Bock, M., Roth, A., Dech, S., 2008. Improvement of image segmentation accuracy based on multiscale optimization procedure. IEEE Geosci. Remote Sens. Lett. 5 (3), 463–467. http://dx.doi.org/10.1109/LGRS.2008.919622.

Esch, T., Schenk, A., Thiel, M., Ullmann, T., Schmidt, M., Dech, S., 2010. Land cover classification based on single-polarized VHR SAR images using texture information derived via speckle analysis. In: 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI. 2010. pp. 1875–1878. http://dx.doi.org/10.1109/IGARSS.2010.5650031.

Esch, T., Taubenböck, H., Roth, A., Heldens, W., Felbier, A., Schmidt, M., Mueller, A.A., Thiel, M., Dech, S.W., 2012. TanDEM-X mission-new perspectives for the inventory

and monitoring of global settlement patterns. J. Appl. Remote. Sens. 6 (1). http://dx.doi.org/10.1117/1.JRS.6.061702.

Espindola, G.M., Camara, G., Reis, I.A., Bins, L.S., Monteiro, A.M., 2006. Parameter selection for region-growing image segmentation algorithms using spatial auto-correlation. Int. J. Remote Sens. 27 (14), 3035–3040. http://dx.doi.org/10.1080/01431160600617194.

Fauvel, M., Benediktsson, J.A., Chanussot, J., Sveinsson, J.R., 2008. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. IEEE Trans. Geosci. Remote Sens. 46 (11), 3804–3814. http://dx.doi.org/10.1109/TGRS.2008.922034.

Fernandez-Delgado, M., Cernadas, E., Barro, S., Amorim, D., 2014. Do we need hundreds of classifiers to solve real world classification problems? J. Mach. Learn. Res. 15, 3133–3181. Available online: http://jmlr.org/papers/volume15/delgado14a/delgado14a.pdf, Accessed date: 6 June 2017.

Foody, G., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80 (1), 185–201. http://dx.doi.org/10.1016/S0034-4257(01)00295-4.

Fritz, T., Eineder, M., Mittermayer, J., Schättler, B., Balzer, W., Bruckreuß, S., Werninghaus, R., 2008. TerraSAR-X Ground Segment Basic Product Specification Document. German Aerospace Center, Cluster Applied Remote Sensing, Oberpfaffenhofen, Germany Available online: http://www.dlr.de/PortalData/1/Resources/raumfahrt/weltraum/TX-GS-DD-3302_Basic-Product-Specification-Document_1.5.pdf, Accessed date: 6 June 2017.

Gamba, P., 2013. Human settlements: a global challenge for EO data processing and interpretation. Proc. IEEE 101 (3), 570–581. http://dx.doi.org/10.1109/JPROC.2012.2189089.

Gamba, P., Aldrighi, M., Stasolla, M., 2011. Robust extraction of urban area extents in HR and VHR SAR images. IEEE J. Select. Top. Appl. Earth Obs. Remote Sens. 4, 27–34. http://dx.doi.org/10.1109/JSTARS.2010.2052023.

Geiß, C., Taubenböck, H., 2015. Object-based postclassification relearning. IEEE Geosci. Remote Sens. Lett. 12 (11), 2336–2340. http://dx.doi.org/10.1109/LGRS.2015.2477436.

Geiß, C., Aravena Pelizari, P., Marconcini, M., Sengara, W., Edwards, M., Lakes, T., Taubenböck, H., 2015. Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques. ISPRS J. Photogramm. Remote Sens. 104, 175–188. http://dx.doi.org/10.1016/j.isprsjprs.2014.07.016.

Geiß, C., Jilge, M., Lakes, T., Taubenböck, H., 2016a. Estimation of seismic vulnerability levels of urban structures with multisensor remote sensing. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (5), 1913–1936. http://dx.doi.org/10.1109/JSTARS.2015.2442584.

Geiß, C., Klotz, M., Schmitt, A., Taubenböck, H., 2016b. Object-based morphological profiles for classification of remote sensing imagery. IEEE Trans. Geosci. Remote Sens. 54 (10), 5952–5963. http://dx.doi.org/10.1109/TGRS.2016.2576978.

Geiß, C., Aravena Pelizari, P., Schrade, H., Brenning, A., Taubenböck, H., 2017. On the effect of spatially non-disjoint training and test samples on estimated model generalization capabilities in supervised classification with spatial features. IEEE Geosci. Remote Sens. Lett. 14 (11), 2008–2012. http://dx.doi.org/10.1109/LGRS.2017.2747222.

Genuer, R., Poggi, J.-M., Tuleau, C., 2008. Random forests: some methodological insights. Rapport de recherche RR-6729. Available online: INRIAhttps://arxiv.org/abs/0811.3619, Accessed date: 6 June 2017.

Ghamisi, P., Souza, R., Benediktsson, J.A., 2015. A survey on spectral–spatial classification techniques based on attribute profiles. IEEE Trans. Geosci. Remote Sens. 53 (5), 2335–2353. http://dx.doi.org/10.1109/TGRS.2014.2358934.

Giada, S., DeGroeve, T., Ehrlich, D., Soille, P., 2003. Information extraction from very high resolution satellite imagery over Lukole refugee camp, Tanzania. Int. J. Remote Sens. 24, 4251–4266. http://dx.doi.org/10.1080/0143116021000035021.

Guyon, I., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182.

Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn. 46 (1–3), 389–422. http://dx.doi.org/10.1023/A:1012487302797.

Hall, M.A., 1999. Correlation-based Feature Selection for Machine Learning. Department of Computer Science, The University of Waikato, Hamilton, New Zealand Available online: http://www.cs.waikato.ac.nz/~mhall/thesis.pdf, Accessed date: 6 June 2017 (Ph.D. dissertation).

Hall, M.A., 2000. Correlation-based feature selection for discrete and numeric class machine learning. In: Langley, Pat (Ed.), Proceedings of the Seventeenth International Conference on Machine Learning (ICML '00). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 359–366.

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern. SMC-3, 610–621. http://dx.doi.org/10.1109/TSMC.1973.4309314.

Heinzel, J., Kemper, T., 2014. Use of new coastal spectral band for precise dwelling extraction in the hagaderea refugee camp. In: Proceedings of ESA-EUSC-JRC 9th Conference on Image Information Mining, http://dx.doi.org/10.2788/25852.

Herz, M., 2006. From Camp to City. Refugee Camps of the Western Sahara. Lars Müller Publishers (512 pp).

Hughes, G., 1968. On the mean accuracy of statistical pattern recognizers. IEEE Trans. Inf. Theory 14 (1), 55–63. http://dx.doi.org/10.1109/TIT.1968.1054102.

Jenerowicz, M., Kemper, T., Soille, P., 2011. An automated procedure for detection of IDP's dwellings using VHR satellite imagery. In: Proc. SPIE 8180, Image and Signal Processing for Remote Sensing XVII, 818004, http://dx.doi.org/10.1117/12.898187.

Johnson, B., Xie, Z., 2013. Classifying a high resolution image of an urban area using super-object information. ISPRS J. Photogramm. Remote Sens. 83, 40–49. http://dx.doi.org/10.1016/j.isprsjprs.2013.05.008.

Kemper, T., Jenerowicz, M., Soille, P., Pesaresi, M., 2011. Enumeration of dwellings in Darfur Camps from GeoEye-1 satellite images using mathematical morphology. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 4, 8–15. http://dx.doi.org/10.1109/JSTARS.2010.2053700.

Klotz, M., Kemper, T., Geiß, C., Esch, T., Taubenböck, H., 2016. How good is the map? A multi-scale cross-comparison framework for global settlement layers: evidence from Central Europe. Remote Sens. Environ. 178, 191–212. http://dx.doi.org/10.1016/j.rse.2016.03.001.

Kohavi, R., John, G.H., 1997. Wrappers for feature subset selection. Artif. Intell. 97 (1), 273–324. http://dx.doi.org/10.1016/S0004-3702(97)00043-X.

Kranz, O., Zeug, G., Tiede, D., Clandillon, S., Bruckert, D., Kemper, T., Lang, S., Caspard, M., 2010. Monitoring refugee/IDP camps to support international relief action. In: Altan, O., Backhaus, R., Boccardo, P., Zlatanova, S. (Eds.), Geoinformation for Disaster and Risk Management — Examples and Best Practices. Joint Board of Geospatial Information Societies (JB GIS), United Nations Office for Outer Space Affairs (UNOOSA), pp. 51–56 (ISBN 978-87-90907-88-4).

Kuffer, M., Pfeffer, K., Sliuzas, R., 2016a. Slums from space—15 years of slum mapping using remote sensing. Remote Sens. 8 (6), 455. http://dx.doi.org/10.3390/rs8060455. (2016).

Kuffer, M., Pfeffer, K., Sliuzas, R., Baud, I., 2016b. Extraction of slum areas from VHR imagery using GLCM variance. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (5), 1830–1840. http://dx.doi.org/10.1109/JSTARS.2016.2538563.

Laben, C.A., Brower, B.V., 2000. Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-sharpening. US6011875 A.

Lal, T.N., Chapelle, O., Weston, J., Elisseeff, A., 2006. Embedded methods. In: Guyon, I., Nikravesh, M., Gunn, S., Zadeh, L.A. (Eds.), Feature Extraction: Foundations and Applications, pp. 137–165. http://dx.doi.org/10.1007/978-3-540-35488-8_6.

Lang, S., Tiede, D., Hölbling, D., Füreder, P., Zeil, P., 2010. Earth observation (EO)-based ex-post assessment of IDP camp evolution and population dynamics in Zam Zam, Darfur. Int. J. Remote Sens. 31, 5709–5731. http://dx.doi.org/10.1080/01431161.2010.496803.

Lee, J.-S., 1983. Digital image smoothing and the sigma filter. Comput. Vis. Graph. Image Process. 24 (2), 255–269. http://dx.doi.org/10.1016/0734-189X(83)90047-6.

Leichtle, T., Geiß, C., Wurm, M., Lakes, T., Taubenböck, H., 2017a. Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment. Int. J. Appl. Earth Obs. Geoinf. 54, 15–27. http://dx.doi.org/10.1016/j.jag.2016.08.010.

Leichtle, T., Geiß, C., Lakes, T., Taubenböck, H., 2017b. Class imbalance in unsupervised change detection – a diagnostic analysis from urban remote sensing. Int. J. Appl. Earth Obs. Geoinf. 60, 83–98. http://dx.doi.org/10.1016/j.jag.2017.04.002.

Ma, L., Cheng, L., Li, M., Liu, Y., Ma, X., 2015. Training set size, scale, and features in Geographic Object-Based Image Analysis of very high resolution unmanned aerial vehicle imagery. ISPRS J. Photogramm. Remote Sens. 102, 14–27. http://dx.doi.org/10.1016/j.isprsjprs.2014.12.026.

Martha, T.R., Kerle, N., van Westen, C.J., Jetten, V., Kumar, K.V., 2011. Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis. IEEE Trans. Geosci. Remote Sens. 49 (12), 4928–4943. http://dx.doi.org/10.1109/TGRS.2011.2151866.

Masjedi, A., Valadan Zoej, M.J., Maghsoudi, Y., 2016. Classification of polarimetric SAR images based on modeling contextual information and using texture features. IEEE Trans. Geosci. Remote Sens. 54 (2), 932–943. http://dx.doi.org/10.1109/TGRS.2015.2469691.

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432. http://dx.doi.org/10.1080/01431169608948514.

Pacifici, F., Chini, M., Emery, W.J., 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. Remote Sens. Environ. 113 (6), 1276–1292. http://dx.doi.org/10.1016/j.rse.2009.02.014.

Pal, M., Foody, G.M., 2010. Feature selection for classification of hyperspectral data by SVM. IEEE Trans. Geosci. Remote Sens. 48 (5), 2297–2307. http://dx.doi.org/10.1109/TGRS.2009.2039484.

Persello, C., Bruzzone, L., 2014. Active and semisupervised learning for the classification of remote sensing images. IEEE Trans. Geosci. Remote Sens. 52 (11), 6937–6956. http://dx.doi.org/10.1109/TGRS.2014.2305805.

Pesaresi, M., Benediktsson, J.A., 2001. A new approach for the morphological segmentation of high-resolution satellite imagery. IEEE Trans. Geosci. Remote Sens. 39 (2), 309–320. http://dx.doi.org/10.1109/36.905239.

Richter, R., 1996. A spatially adaptive fast atmospheric correction algorithm. Int. J. Remote Sens. 17 (6), 1201–1214. http://dx.doi.org/10.1080/01431169608949077.

Rodriguez, J.J., Kuncheva, L.I., Alonso, C.J., 2006. Rotation forest: a new classifier ensemble method. IEEE Trans. Pattern Anal. Mach. Intell. 28 (10), 1619–1630. Available online: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1677518, Accessed date: 6 June 2017.

Rougier, S., Puissant, A., Stumpf, A., Lachiche, N., 2016. Comparison of sampling strategies for object-based classification of urban vegetation from Very High Resolution satellite images. Int. J. Appl. Earth Obs. Geoinf. 51, 60–73. http://dx.doi.org/10.1016/j.jag.2016.04.005. (September 2016, ISSN 0303-2434).

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. In: Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, Greenbelt: NASA SP-351, 3010-317, Available online: https://ntrs.nasa.gov/search.jsp?R=19740022614, Accessed date: 6 June 2017.

Schöpfer, E., Spröhnle, K., Aravena Pelizari, P., 2015. Temporäre Siedlungen: Wenn aus Flüchtlingslagern Städte werden. In: Taubenböck, H., Wurm, M., Esch, T., Dech, S. (Eds.), Globale Urbanisierung – Perspektive aus dem All. Springer, pp. 171–178.

Soille, P., 2004. Morphological Image Analysis: Principles and Applications, 2nd ed.

Springer, Berlin, Germany.

Spröhnle, K., Tiede, D., Schoepfer, E., Füreder, P., Svanberg, A., Rost, T., 2014. Earth observation-based dwelling detection approaches in a highly complex refugee camp environment - a comparative study. Remote Sens. 6 (10), 9277–9297. http://dx.doi.org/10.3390/rs6109277.

Spröhnle, K., Fuchs, E.-M., Aravena Pelizari, P., 2017. Object-based analysis and fusion of optical and SAR satellite data for dwelling detection in refugee camps. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (5), 1780–1791. http://dx.doi.org/10.1109/JSTARS.2017.2664982.

Stumpf, A., Kerle, N., 2011. Object-oriented mapping of landslides using random forests. Remote Sens. Environ. 115 (10), 2564–2577. http://dx.doi.org/10.1016/j.rse.2011.05.013.

Taubenböck, H., Esch, T., Wurm, M., Roth, A., Dech, S., 2010. Object-based feature extraction using high spatial resolution satellite data of urban areas. J. Spat. Sci. 55 (1), 117–133. http://dx.doi.org/10.1080/14498596.2010.487854.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., Dech, S., 2012. Monitoring urbanization in mega cities from space. Remote Sens. Environ. 117, 162–176. http://dx.doi.org/10.1016/j.rse.2011.09.015.

Taubenböck, H., Kraff, N.J., Wurm, M., 2018. The morphology of the Arrival City - A global categorization based on literature surveys and remotely sensed data. Appl. Geogr. 92, 150–167. http://dx.doi.org/10.1016/j.apgeog.2018.02.002. (ISSN 0143-6228).

Trimble, 2014. eCognition Developer 9. Reference Book. Trimble Germany GmbH, Munich, Germany.

Trunk, G.V., 1979. A problem of dimensionality: a simple example. IEEE Trans. Pattern Anal. Mach. Intell. 3, 306–307. http://dx.doi.org/10.1109/TPAMI.1979.4766926. (PAMI-1).

Tuia, D., Pacifici, F., Kanevski, M., Emery, W.J., 2009. Classification of very high spatial resolution imagery using mathematical morphology and support vector machines. IEEE Trans. Geosci. Remote Sens. 47 (11), 3866–3879. http://dx.doi.org/10.1109/TGRS.2009.2027895.

Tuia, D., Volpi, M., Copa, L., Kanevski, M., Munoz-Mari, J., 2011. A survey of active

learning algorithms for supervised remote sensing image classification. IEEE J. Sel. Top. Signal Process. 5, 606–617. http://dx.doi.org/10.1109/JSTSP.2011.2139193.

Uhlmann, S., Kiranyaz, S., 2014. Classification of dual- and single polarized SAR images by incorporating visual features. ISPRS J. Photogramm. Remote Sens. 90, 10–22. http://dx.doi.org/10.1016/j.isprsjprs.2014.01.005.

UNHCR, 2000. Key Principles for Decision-making. Refugee Operations and Environmental Management; Engineering and Environmental Service Section, UNHCR, Geneva, Switzerland Available online: http://www.unhcr.org/3b03b24d4.html, Accessed date: 7 June 2017 (June 2000).

UNHCR, 2005. Environmental Guidelines. UNHCR, Geneva, Switzerland Available online: http://www.unhcr.org/3b03b2a04.html, Accessed date: 18 October 2017 (August 2005).

UNHCR, 2007. Handbook for Emergencies, 3rd edition. United Nations High Commissioner for Refugees, Geneva Available online. https://www.humanitarianresponse.info/system/files/documents/files/UNHCR%20Handbook%20for%20Emergencies.pdf, Accessed date: 20 October 2017.

UNHCR, 2014. Syria regional refugee response. Inter-agency information sharing portal. Available online: Zaatari Refugee Camp http://data.unhcr.org/syrianrefugees/settlement.php?id=176&country=107&region=77, Accessed date: 7 June 2017.

UNHCR, 2016. Global Trends – Forced Displacement in 2015. UNHCR, Geneva, Switzerland Available online: https://s3.amazonaws.com/unhcrsharedmedia/2016/2016-06-20-global-trends/2016-06-14-Global-Trends-2015.pdf, Accessed date: 6 June 2017 (June 2016).

Voigt, S., Kemper, T., Riedlinger, T., Kiefl, R., Scholte, K., Mehl, H., 2007. Satellite image analysis for disaster and crisis-management support. IEEE Trans. Geosci. Remote Sens. 45 (6), 1520–1528. http://dx.doi.org/10.1109/TGRS.2007.895830.

Wang, S., So, E., Smith, P., 2015. Detecting tents to estimate the displaced populations for post-disaster relief using high resolution satellite imagery. Int. J. Appl. Earth Obs. Geoinf. 36, 87–93. http://dx.doi.org/10.1016/j.jag.2014.11.013.

Wurm, M., Taubenböck, H., Weigand, M., Schmitt, A., 2017. Slum mapping in polarimetric SAR data using spatial features. Remote Sens. Environ. 194, 190–204. http://dx.doi.org/10.1016/j.rse.2017.03.030.