International Information and Engineering Technology Association

Advancing the World of Information and Engineering

An Efficient Multi-stage Object-Based Classification to Extract Urban Building Footprints from HR Satellite Images

Gopala Krishna VSSN Pendyala[1*], Hemantha Kumar Kalluri[2], Venkateswara C. Rao[1]

[1] National Remote Sensing Centre, Hyderabad 500037, Telangana, India
[2] Department of Computer Science & Engineering, Vignan's Foundation for Science, Technology and Research, Vadlamudi 522213, Andhra Pradesh, India

Corresponding Author Email: gopalakrishna_pvssn@nrsc.gov.in

ABSTRACT

Urban building information can be effectively extracted by applying object-based image segmentation and multi-stage thresholding on High Resolution (HR) remote sensing satellite imageries. This study provides the results obtained using this method on the images of Indian remote sensing satellite, CARTOSAT-2S launched by the Indian Space Research Organization (ISRO). In this study, a method is developed to extract urban building footprints from the HR remote sensing satellite images. The first step of the process consists of generating highly dense per pixel Digital Surface Model (DSM) by using semi global matching algorithm on HR satellite stereo images and applying robust ground filtering to generate Digital Terrain Model (DTM). In the second step, multi-stage object-based approach is adopted to extract building bases using the PAN sharpened image, normalized Digital Surface Model (nDSM) derived from DSM and DTM, and Normalised Difference Vegetation Index (NDVI). The results are compared with the manual method of drawing building footprints by cartographers. An average precision of 0.930, recall of 0.917, and f-score of 0.922 are obtained. The results are found to be in a match with the method using the high resolution Airborne LiDAR DSM by providing a solution for large areas, low cost and low time.

1. INTRODUCTION

Growth assessment, planning, development, monitoring, risk management, and climate monitoring in urban areas is achieved in a planned way using the urban information system. This is essentially a Geographic Information System (GIS) based data information and management system used for planning and managing urban habitat facilities [1]. The basic building block of this information system is the individual building data with the location and size of the buildings from which, the extent of human settlements, their exact locations together with other attributes such as the cluster density, and proximity to utilities are estimated.

The manual method of collecting individual building information with attributes is very expensive and time consuming. Automatic extraction of building information using High Resolution (HR) remote sensing images is one of the widely used methods globally since it allows large area information retrieval at economical and faster rates.

This paper provides an efficient multi-stage object-based methodology to extract individual building information from High Resolution remote sensing images. The rest of the paper is organized as follows: Section two provides a literature study, Section three gives data sets and study area, Section four discusses the methodology adopted, Section five presents the results and analysis and conclusions are drawn in the last chapter.

2. LIRERATURE STUDY

Various HR satellite images are globally available with the resolution/Ground Sample Distance (GSD) varying from 0.5 m to 1.0 m, such as IKONOS (image GSD 1.0 m), QuickBird (0.6 m), WorldView-2 (0.5 m), GeoEye (0.46 m) and Cartosat -2S (0.6 m) [2]. Using the Panchromatic (PAN) and MS (Multi Spectral) images acquired from these satellites, various methods are adopted by researchers to extract individual building information. The primary data which is in the form of images, are directly used or the derived secondary data such as PAN sharpened image, Digital Surface Model (DSM), Digital Terrain Model (DTM), normalised DSM (nDSM), and Normalized Difference Vegetation Index (NDVI) are used for this purpose. These images are also used in conjunction with very dense and accurate Airborne LiDAR-derived DSM for this purpose.

The models used on these datasets are broadly grouped into three categories as shown in Figure 1. The first model is based on using the image data directly; the second model uses the NDVI, nDSM derived together with the images; and the third model uses the hybrid approach of two independent sensor data, where the image is used in conjunction with the Airborne LiDAR-DSM. The processing algorithms used to mine the building information are supervised classification, morphological operations, object-based approach with segmentation, machine learning based approach, Fully Convolution Network (FCN), etc.

Table 1. Results obtained HR satellite image data (Model 1)

Reference	Author	Data set	Method	Precision P TP/(TP+FP)	Recall R TP/(TP+FN)	f-score 2*((P*R)/(P+R))
[3]	Jin and Davis	IKONOS PAN	Differential Morphological Profile	0.67	0.62	0.644
[4]	Tiwari et al.	IKONOS Fused	Object-based approach with segmentation	0.66	0.73	0.693
[5]	Lefèvre et al.	QuickBirdPAN	Advanced morphological operations	0.79	0.635	0.704
[6]	Chandra and Ghosh	SAR, multispectral images	Using shadow information	0.794	0.643	0.710
[7]	Liasis and Stavrou	Google Earth Satellite images 1 m to 10 m	Segmentation using active color and color features	0.705	0.740	0.722
[8]	Sefercik et al.	IKONOS Pan sharpened	Object-based segmentation	0.701	0.826	0.758
[9]	Benarchid et al.	GeoEye Pan sharpened	Object-based approach with shadow information	0.90	0.66	0.761
[8]	Sefercik et al.	QuickBird PAN &MS	Object-based segmentation	0.669	0.909	0.770
[10]	Chen et al.	RGB high-resolution images from Google Earth	Object-based and machine learning-based approach	0.902	0.724	0.803
[11]	Li et al.	QuickBird, Pléiades 0.5 m	Morphological building indices (MBIs) and mask with hierarchical probable model	0.7562	0.863	0.806
[12]	Li et al.	HR data with 0.31 m PAN and 1.24 m MS	Semantic segmentation	0.896	0.763	0.824
[13]	Sefercik et al.	WorldView 2 Pan sharpened	Novel Fusion approach	0.874	0.79	0.829
[14]	Attarzadeh and Momeni	QuickBird	Object-based approach from buildings characteristics	0.809	0.893	0.849
[15]	Gavankar and Ghosh	IKONOS PAN	Morphological based (automatic)	0.895	0.915	0.904

Table 2. Results obtained by HR satellite image data HR satellite image + derived products (Model 2)

Reference	Author	Data set	Method	Precision P TP/(TP+FP)	Recall R TP/(TP+FN)	f-score 2*((P*R)/(P+R))
[16]	Davydova et al.	WorldView, nDSM	Neural network	0.9	0.735	0.809
[17]	Bittnera et al.	WorldView, nDSM	Fully Convolution Network (FCN)	0.86	0.78	0.818
[8]	Sefercik et al.	WorldView, nDSM	Novel Fusion approach	0.896	0.815	0.853
[8]	Sefercik et al.	IKONOS, nDSM	Object-based segmentation	0.904	0.829	0.864
[18]	Shaker et al.	IKONOS, nDSM	Supervised classification	0.937	0.84	0.885

Table 3. Results obtained with HR satellite/aerial image + LiDAR-nDSM (Model 3)

Reference	Author	Data set	Method	Precision P TP/(TP+FP)	Recall R TP/(TP+FN)	f-score 2*((P*R)/(P+R))
[19]	Demir et al.	Aerial RGB 0.125 m, LiDAR-nDSM, NDVI	Supervised classification	0.76	0.9	0.824
[20]	San and Turker	IKONOS, nDSM, NDVI	SVM, edge detection, vectoring and grouping	0.78	0.95	0.856
[19]	Demir et al.	Aerial RGB 0.125 m, LiDAR-nDSM, NDVI	Classification over images by applying LiDAR height information	0.86	0.87	0.865
[19]	Demir et al.	Aerial RGB 0.125m, LiDAR-nDSM, NDVI	NDVI classification + voids from LiDAR-DTM	0.87	0.87	0.87
[19]	Demir et al.	Aerial RGB 0.125 m, LiDAR-nDSM, NDVI	LiDAR classification	0.92	0.83	0.872
[21]	Yan et al.	Aerial Image + LiDAR-DSM 0.09m	Stacked Sparse Auto encoder (ANN)	0.909	0.907	0.908
[22]	Ni et al.	ALS Point cloud	Supervised segmentation, Random Forests based Classification	0.954	0.867	0.908

The results obtained for each model varied with the type of dataset and the processing model adopted. Quantification of the results is reported in terms of precision, recall, and f-score using the number of True Positives (TP), True Negatives (TN), and False Positives (FP) detected by that particular method. The detailed results obtained by various researchers are shown in Tables 1, 2, and 3.

It is found that, the f-score achieved by various researchers for HR satellite image data (model 1) is 0.644 – 0.904, for HR satellite image + derived products (model 2) is 0.809 – 0.885 and for HR satellite/aerial image + LiDAR-nDSM (model 3) is 0.824 – 0.908. It is also observed that the HR satellite/aerial image data in conjunction with LiDAR data used in model 3 gives the best results. However, LiDAR data

is based on the aerial/ drone platform, hence, using this model is time consuming and expensive. On the other hand, the HR satellite data with the derived product of nDSM, NDVI used in model 2 offers the advantage of faster, large area coverage, and less cost. Hence, the objective is to derive a methodology by improving model 2 to achieve the same results obtained by model 3.

Model 1: Using HR satellite images

Model 2: Using HR satellite image + derived products nDSM, NDVI

Model 3: Using HR satellite/aerial image + LiDAR-nDSM

Figure 1. Models used to extract building information

3. DATA SET AND STUDY AREA

3.1 Data set

The CARTOSAT-2S satellite of the Indian Space Research Organization (ISRO) provides images in a single PAN and four MS bands at a spatial resolution of 0.60 m and 1.6 m simultaneously at 11-bit radiometric resolution. The PAN image can take panchromatic (black and white) in a selected portion of the visible and near-infrared spectrum (0.50 - 0.85 μm). The 4-band MS records data of spectral resolutions in Blue (0.43 - 0.52 μm), Green (0.52 - 0.61 μm), Red (0.61 - 0.69 μm) and NIR (0.76 - 0.90 μm) in 4 separate bands.

3.2 Study area

CARTOSAT-2S satellite images of the part of West Hyderabad acquired on January 1, 2018, and January 28, 2018, simultaneously in PAN and MS bands with across track stereo mode are used for investigation purposes. An

area covering 3 km² was a subset for study purposes. The scene includes urban settlements consisting of very dense to sparse buildings. The heights of the buildings vary from a single floor to high-rise apartments with concrete/sheet rooftops. The location of the study area and the images of the PAN and MS are shown in Figure 2.
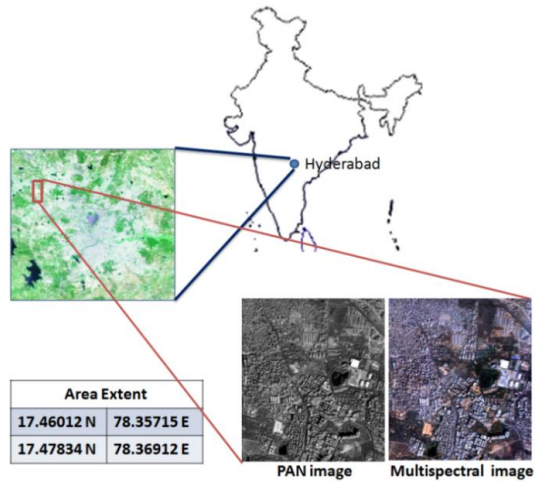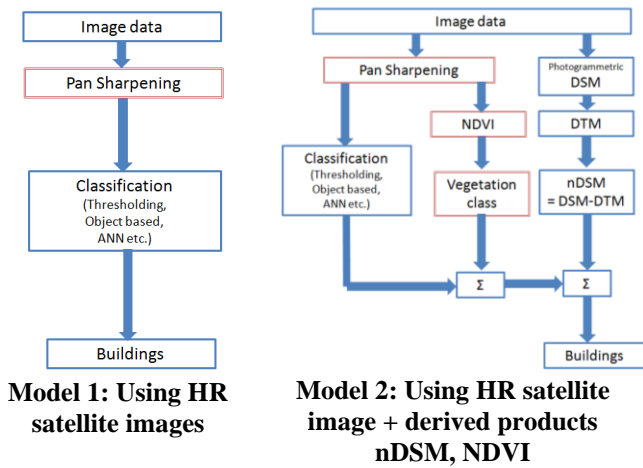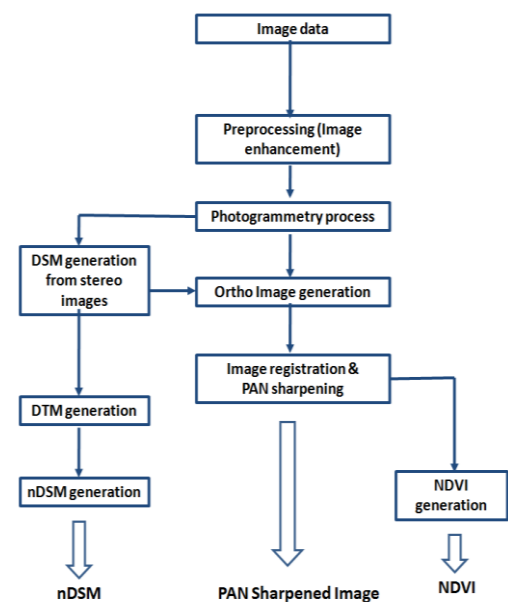

Figure 2. Study area

4. METHODOLOGY

As mentioned in the previous section, the best results are achieved by using LiDAR-nDSM compared to photogrammetrically derive satellite DSM, due to the high point density. Hence, the approach of this study is based on improving the derived nDSM from satellite stereo images to match the density of LiDAR-nDSM supplemented with customised preprocessing.

The process is carried out as (i) Preprocessing and secondary data generation as per the flowchart shown in Figure 3 (ii) Multi-stage object-based classification as per the flowchart in Figure 4 and (iii) Regularization of building shapes.

Figure 3. Flow chart of preprocessing and secondary data generation

PAN sharpened Image
↓
Water body and shadows removal
↓
NDVI → Masking for water body and shadows
↓
Vegetation filtering using NDVI threshold → Water body & Shadow layer
→ Vegetation layer
↓
nDSM → Height layer extracted using nDSM threshold
↓
Spectral filter applied on Height layer using Multiband characteristics
↓
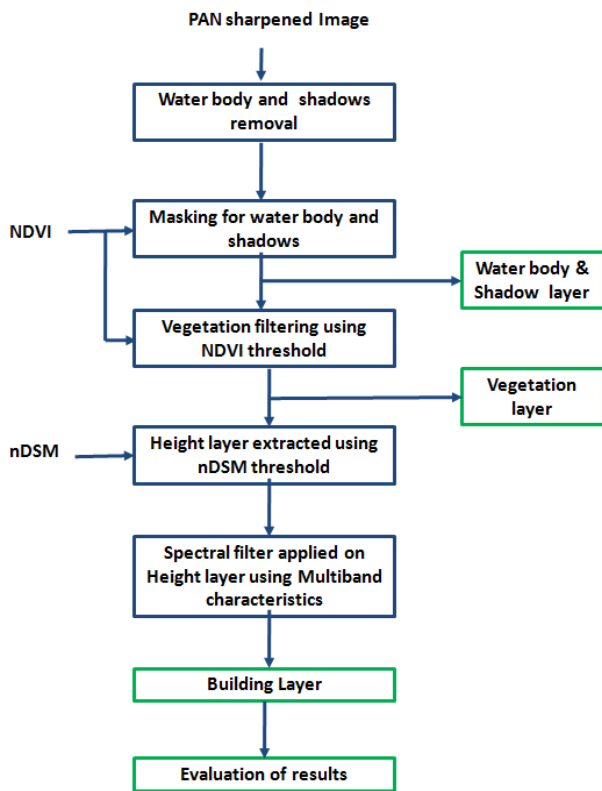Building Layer
↓
Evaluation of results

Figure 4. Flow chart of multi stage object-based classification

4.1 Preprocessing and secondary data generation

These is achieved by (a) Image enhancement and PAN sharpening (b) generating nDSM from highly dense per-pixel DSM using a semi-global matching algorithm and (c) applying robust ground filtering method based on Discrete Cosine Transform (DCT) to generate DTM.

PAN Sharpening: After preliminary image enhancements, high resolution (0.6 m) PAN image is fused with the low resolution MS image (1.6 m) to generate high resolution (0.6 m) PAN sharpened image. This process injects the spectral details of the MS image into the HR PAN image. Partial Replacement Adaptive Component Substitution (PRACS) method is used after experimenting with various PAN sharpening algorithms with the data set and evaluating the resultant image quantitatively.

nDSM generation: Per-pixel DSM is generated using the semi-global matching method [23]. This method could derive DSM point cloud from the HR satellite images with a point density of 2.79 points / sq. m. with a height accuracy of 1.79 m. The filtering of ground points (DTM) from DSM is achieved by DCT-based ground surface interpolation as it offers better noise filtering [24]. The accuracy of DSM generation using this method is reported in K-values of 0.91, 083 and 0.90 respectively for the areas containing vegetation, small close buildings and high rise buildings respectively [24]. The terrain variations are in all respects are reflected in this DTM as the point density is very high (which is almost like LiDAR DTM). nDSM is generated by subtracting DTM from DSM.

NDVI Generation: Normalised Difference Vegetation Index gives a quantitative estimation of vegetation growth and biomass. This is generated using the NIR band and the Red band of the fused image.

4.2 Multi-stage object-based classification

This process consists of (a) image segmentation and (b) multi-stage classification. Image segmentation is the process of dividing the images into segments / objects based on its spectral properties of the image that uniquely represent a small set of areas. Multi-stage thresholding is applied to the created objects to derive the buildings.

Object Generation: Multiresolution segmentation is used to generate the objects having a pixel count of fewer than 14 pixels.

Shadows and water body removal: This is a threshold level classification using the NDVI and Blue band of the image to remove shadows and water bodies.

Vegetation removal: NDVI Threshold is determined from the bimodal histogram of the image having dominant buildings and vegetation. This step differentiates the vegetation from buildings.

Height object extraction: Threshold-based classification with a height limit of 2.0 m on nDSM is applied above that and small height objects can be removed.

Extraction of buildings from height objects: All height objects do not represent buildings; hence multispectral (Red, Green, Blue bands) thresholding is used to extract buildings from height objects.

4.3 Regularization of building shapes

Regularization is carried out to bring the created objects to the shapes of the buildings. Holes in the buildings are subsumed into large buildings and the smaller noise appearing as buildings are removed by area threshold. Finally, the edges of the buildings are regularized to get the proper corners.

5. RESULTS AND ANALYSIS

The process experiments on 4 subsets of areas extracted from the study area as shown in Figure 5. The subsets are identified as Villa, Dense, High-rise, and Sparse areas as per the distribution of buildings.

Figure 5. Subset of areas selected for study

5.1 Results

The processing of the data is carried out using the e-cognition software as per the methodology provided in the previous section. The multiresolution segmentation and the multi-stage object-based classification using thresholding are applied sequentially and the results are provided in Figure 6, and the building shape regularization is shown in Figure 7.
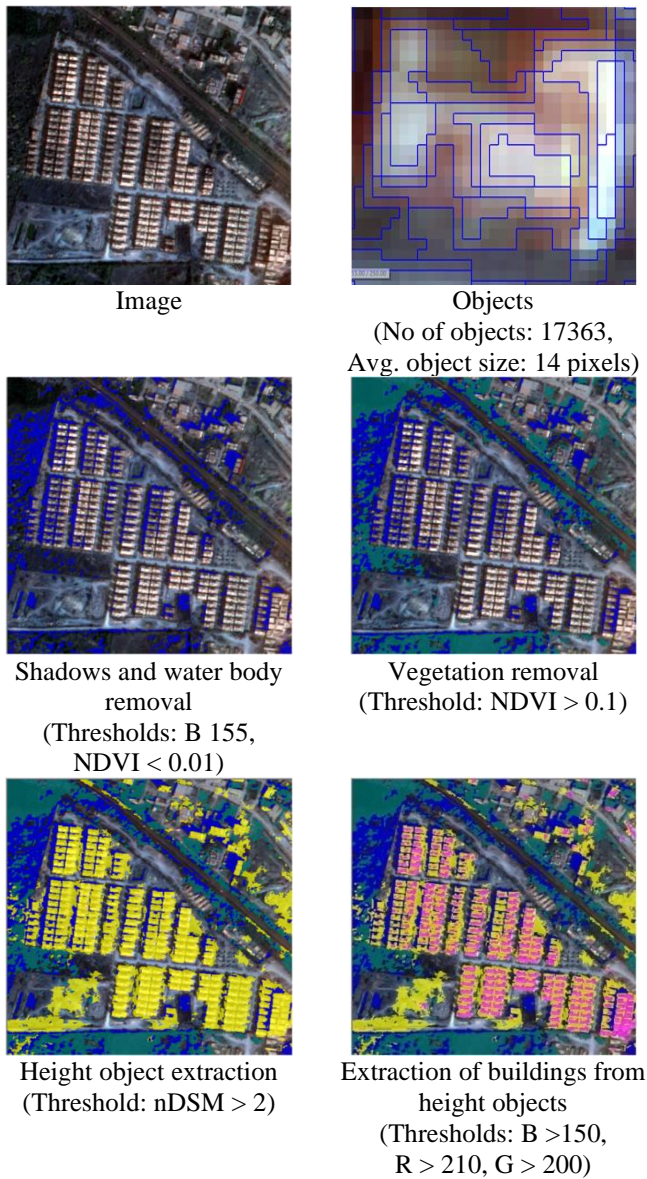

| Image | Objects (No of objects: 17363, Avg. object size: 14 pixels) |

Shadows and water body removal (Thresholds: B 155, NDVI < 0.01)

Vegetation removal (Threshold: NDVI > 0.1)

Height object extraction (Threshold: nDSM > 2)

Extraction of buildings from height objects (Thresholds: B >150, R > 210, G > 200)

Figure 6. Applied thresholds and the resultant objects

Actual Building footprints
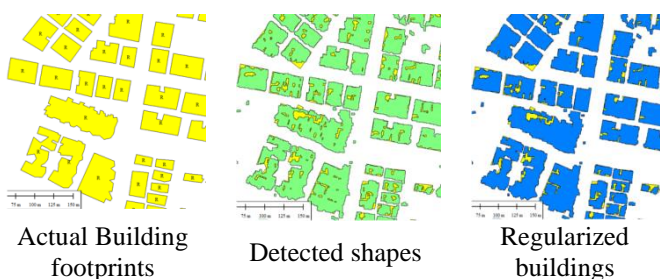
Detected shapes

Regularized buildings

Figure 7. Regularization of building shapes

5.2 Analysis

The reference data is provided by manually drawing all the building boundaries residing in that particular study area. The reference data is compared with the buildings extracted by the developed multi-stage object-based classification methodology. The results are tabulated in Tables 4a and 4b.

Table 4a. Number of buildings detected in the study areas

Sample Type	No of Buildings detected		
	Correct Detection TP	Undetected FN	Wrong Detection FP
Villa	176	26	1
High-Rise	119	7	22
Dense	124	10	6
Sparse	234	18	18

Table 4b. Results obtained from the study areas

Sample Type	Precision P TP/ (TP+FP)	Recall R TP/ (TP+FN)	f-score 2*((P*R)/ (P+R))
Villa	0.994	0.871	0.929
High-Rise	0.844	0.944	0.891
Dense	0.954	0.925	0.939
Sparse	0.929	0.929	0.929
Average	**0.930**	**0.917**	**0.922**

It is observed that the obtained results are comparable with the f-score achieved by using the LiDAR-nDSM. The false negatives or undetected buildings occurred due to the low spectral reflectivity from the roofs covered with dirt or fading. This is apparent as high undetected buildings in the villa area (increased false negative) as reflectance from the houses is not bright and the reflectance is similar to the surrounding features on the ground. The wrong detection or false positives occurred due to the similarity of the spectral characteristics between the rooftops and metal/concrete roads in the areas where it is highly urbanized. This result is reflected in the high- rise area (increased false positive) due to the concrete road bright reflections falsely recorded as a building.

6. CONCLUSIONS

Urban building detection using High Resolution satellite images provides a viable solution in terms of large area coverage with less cost and time. Multi-stage object-based classification combined with a highly accurate and dense DSM can provide better building detection results.

A combination of high resolution PAN sharpened satellite image with carefully extracted nDSM derived from Semi-global matching algorithm and NDVI data sets can be used to extract buildings efficiently from urban areas by applying the multi-stage object-based segmentation and classification methods. An f-score of 0.922 is achieved, which is comparable to the results that use Airborne LiDAR-nDSM. Utilizing the dense DTM provides the same levels of detection efficiency in case of terrain variations from flat to steep terrain. However, the accuracy of the algorithm slightly varies on influence of season and location as the vegetation effects the DSM generation and DTM filtering accuracy.

REFERENCES

[1] Pendyala, V.G.K., Kalluri, H.K., Rao, C.V. (2018). A Multi stage approach for urban building extraction from remote sensing satellite images. International Journal of Engineering & Technology, 7(4.24): 95-99. https://doi.org/10.14419/ijet.v7i4.24.21864

[2] Jacobsen, K. (2013). DEM generation from high resolution satellite imagery. PFG 2013/5 Report, pp. 483-493, Stuttgart. https://doi.org/10.1127/1432-8364/2013/0194

[3] Jin, X.Y., Davis, C.H. (2005). Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information. EURASIP Journal on Applied Signal Processing, 2005(14): 2196-2206. https://doi.org/10.1155/ASP.2005.2196

[4] Tiwari, P.S., Pande, H., Nanda, B.N. (2004). Building footprint extraction from IKONOS imagery based on multi-scale object oriented fuzzy classification for urban disaster management. The Intl. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 34.

[5] Lefèvre, S., Weber, J., Sheeren, D. (2007). Automatic building extraction in VHR images using advanced morphological operators. 2007 Urban Remote Sensing Joint Event, Paris, France, pp. 1-5. https://doi.org/10.1109/URS.2007.371825

[6] Chandra., N., Ghosh, J.K. (2017). A cognitive method for building detection from high-resolution satellite images. Current Science, 112(5): 1038-1044. https://doi.org/10.18520/cs/v112/i05/1038-1044

[7] Liasis, G., Stavrou, S. (2016) Building extraction in satellite images using active contours and colour features. International Journal of Remote Sensing, 37(5): 1127-1153. https://doi.org/10.1080/01431161.2016.1148283

[8] Sefercik, U.G., Karakis, S., Bayik, C., Alkan, M., Yastikli, N. (2014). Contribution of normalized DSM to automatic building extraction from HR mono optical satellite imagery. European Journal of Remote Sensing, 47(1): 575-591. https://doi.org/10.5721/EuJRS20144732

[9] Benarchid, O., Raissouni, N., Samir, E.A., Abbous, A.E., Azyat, A., Achhab, N.B., Lahraoua, M., Asaad, C. (2013). Building extraction using object-based classification and shadow information in very high resolution multispectral images, a case study: Tetuan, Morocco. Canadian Journal on Image Processing and Computer Vision, 4(1).

[10] Chen, R.X., Li, X.H., Li, J. (2018). Object-based features for house detection from RGB high-resolution images. Remote Sensing, 10(3): 451. https://doi.org/10.3390/rs10030451

[11] Li, S.D., Tang, H., Huang, X., Mao, T., Niu, X.N. (2017). Automated detection of buildings from heterogeneous VHR satellite images for rapid response to natural disasters. Remote Sensing, 9(11): 1177. https://doi.org/10.3390/rs9111177

[12] Li, W.J., He, C.H., Fang, J.R., Zheng, J.P., Fu, H.H., Yu, L. (2019). Semantic segmentation-based building footprint extraction using very high-resolution satellite images and multi-source GIS data. Remote Sensing, 11(4): 403. https://doi.org/10.3390/rs11040403

[13] Sefercik., U.G., Karakis., S., Atalay, C., Yigit., I., Gokmen, U. (2017). Novel fusion approach on automatic object extraction from spatial data: Case study Worldview-2 and TOPO5000, Geocarto International. https://doi.org/10.1080/10106049.2017.1353646

[14] Attarzadeh, R., Momeni, M. (2018). Object-based rule sets and its transferability for building extraction from high resolution satellite imagery. Journal of Indian Society of Remote Sensing, 46(2): 169-178. https://doi.org/10.1007/s12524-017-0694-6

[15] Gavankar, N.L., Ghosh, S.K. (2018). Automatic building footprint extraction from high-resolution satellite image using mathematical morphology. European Journal of Remote Sensing, 51(1): 182-193. https://doi.org/10.1080/22797254.2017.1416676

[16] Davydova, K., Cui, S.Y., Reinartz, P. (2016). Building footprint extraction from digital surface models using neural networks. Proc. SPIE 10004, Image and Signal Processing for Remote Sensing, XXII: 100040J. https://doi.org/10.1117/12.2240727

[17] Bittnera, K., Cuia, S., Reinartza, P. (2017). Building extraction from remote sensing data using fully convolutional network. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XLII-1/W1, pp. 481-486. https://doi.org/10.5194/isprs-archives-XLII-1-W1-481-2017

[18] Shaker, I.F., Abd-Elrahman, A., Abdel-Gawad, A.K., Sherief, M.A. (2011). Building extraction from high resolution space images in high density residential areas in the Great Cairo Region. Remote Sensing, 3(4): 781-791. https://doi.org/10.3390/rs3040781

[19] Demir, N., Poli, D., Baltsavias, E. (2008). Extraction of buildings and trees using images and Lidar data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVII(Part B4): 313-318. https://doi.org/10.3929/ethz-b-000011960

[20] San, D.K., Turker, M. (2010). Building extraction from high resolution satellite images using hough transform. International Archives of the Photogrammetry, Remote Sensing And Spatial Information Science, Volume XXXVIII, Part 8, Kyoto Japan.

[21] Yan, Y.M., Tan, Z.C., Su, N., Zhao, C.H. (2017). Building extraction based on an optimized stacked sparse autoencoder of structure and training samples using Lidar DSM and optical images. Sensors, 17(9): 1957. https://doi.org/10.3390/s17091957

[22] Ni, H., Lin, X.G., Zhang, J.X. (2017). Classification of ALS point cloud with improved point cloud segmentation and random forests. Remote Sensing, 9(3): 288. https://doi.org/10.3390/rs9030288

[23] Hirschmuller, H. (2008). Stereo processing by semi-global matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2): 328-341. https://doi.org/10.1109/TPAMI.2007.1166

[24] Pendyala, V.G.K., Kalluri, H.K., Rao, C.V. (2019). Dense DSM and DTM point cloud generation using CARTOSAT-2E satellite images for high-resolution applications. Journal of the Indian Society of Remote Sensing, 47(12): 2085-2096. https://doi.org/10.1007/s12524-019-01051-0