International Journal of Remote Sensing

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tres20

Fast building detection using new feature sets derived from a very high-resolution image, digital elevation and surface model

Mehmet Akif Günen

To cite this article: Mehmet Akif Günen (2024) Fast building detection using new feature sets derived from a very high-resolution image, digital elevation and surface model, International Journal of Remote Sensing, 45:5, 1477-1497, DOI: 10.1080/01431161.2024.2313991

To link to this article: https://doi.org/10.1080/01431161.2024.2313991

Published online: 14 Feb 2024.

INTERNATIONAL JOURNAL OF REMOTE SENSING
2024, VOL. 45, NO. 5, 1477–1497
https://doi.org/10.1080/01431161.2024.2313991

Taylor & Francis
Taylor & Francis Group

Fast building detection using new feature sets derived from a very high-resolution image, digital elevation and surface model

Mehmet Akif Günen

Faculty of Engineering and Natural Sciences, Department of Geomatics Engineering, Gümüşhane University, Gümüşhane, Türkiye

ABSTRACT

Detecting building rooftops with very high-resolution (VHR) images is an important issue in many fields, including disaster management, urban planning, and climate change research. Buildings with varying geometrical features are challenging to detect accurately from VHR image due to complicated image scenes containing spectrally similar objects, illumination, occlusions, viewing angles, and shadows. This study aims to detect building rooftops with high accuracy using a new framework that includes VHR image, visible band difference vegetation index, digital surface and elevation models, the terrain ruggedness and the topographic position index. Five distinct feature sets were generated in order of importance by exposing the ten related stacking features to a feature selection procedure using the maximum relevance minimum redundancy method. Then, Auto-Encoder, k-NN, decision tree, RUSBoost, and random forest machine learning algorithms were utilized for binary classification. Random forest yielded the highest accuracy (97.2% F-score, 98.72% accuracy) when all features (F_{10}) were used, while decision tree was the least successful (59.16% F-score, 83.56% accuracy) for RGB feature set (F_{RGB}). It was revealed that classification of F_{10} with random forest increased F-score by about 23% compared to classification with F_{RGB}. Additionally, McNemar's tests showed no statistically significant difference between random forest vs k-NN and decision tree vs RUSBoost.

ARTICLE HISTORY
Received 21 April 2023
Accepted 21 January 2024

KEYWORDS
Building detection; very high-resolution image; machine learning; deep learning

1. Introduction

The advancement of remote sensing (RS) technology has made it possible to obtain very high resolution (VHR) images that provide detailed information about the earth's surface. VHR images include remotely sensed satellite images and photogrammetrically produced orthophotos and aerial images (Günen et al. 2019; Han et al. 2019; Li et al. 2022; Q. Wang et al. 2017). Building detection is defined by the RS community as the detection of specific building components from geo-imagery (Jing et al. 2022; Sritarapipat and Takeuchi 2017). Buildings are the fundamental elements of a city and the basis of an urban map. Images from RS sensors (e.g. aerial image and satellite images) are processed to provide critical

CONTACT Mehmet Akif Günen akif@gumushane.edu.tr Faculty of Engineering and Natural Sciences, Department of Geomatics Engineering, Gümüşhane University, Gümüşhane 29100, Türkiye

information on topics such as building's location, planning, management, and security. They are also essential for rapid building detection after earthquake, destruction, building renovation, and urban regeneration projects. VHR images are becoming more usable and accessible thanks to the quick development of various sensors, and they are a crucial input data for building detection. The geometric shape, colour, inside structure, height and texture pattern of buildings can all be revealed in extensive detail thanks to the VHR images, which can approach a few centimetres or decimetres levels (Huang and Zhang 2011b; Raczko, Krówczyńska, and Wilk 2022; Swan et al. 2022; Voigt et al. 2007). Traditional surveying techniques frequently include extensive field mapping and surveying that is time-consuming and expensive. As a result, a number of techniques have been developed to do building detection using images that have undergone post-processing. Image processing methods such as histogram matching, morphological operations and filters generally cannot produce satisfying results. The use of image processing methods for building detection has the drawback of being susceptible to image noise. Also, they identify and detect certain features based on the shape and structure of the objects in an image. Small specks of noise, for example, can be misinterpreted as a roof, while a true roof can be missed due to the noise. Another problem is that they are not very robust to changes in lighting or viewpoint. They can be sensitive to fluctuations in light intensity, causing the filter to deliver erroneous findings (Huang and Zhang 2011a, 2011b; Singh et al. 2012).

Rule-, image segmentation-, and classification- (e.g. machine learning and deep learning) based methods are used to detect building rooftops. Roof classification from VHR RS images is commonly used for building detection. The resulting roofs are actually representations of buildings. Automatic information extraction is vital in situations where fast data is needed, such as disaster management. Regardless of the methods chosen, the major goal is to perform precise and accurate roof detection, laying the framework for several application areas such as illegal roof use (Shi et al. 2022), 3D city model (Kocaman et al. 2006), real-estate management (T. Wu et al. 2020), assessment of solar photovoltaic power generating capacity (Zhong et al. 2021), and health hazard (Gualtieri et al. 2022).

In the Polish cadastral modernization research, VHR images and aerial Light Detection and Ranging (LiDAR) data were utilized to train the U-Net architecture, which identified buildings with an 89.5% overall accuracy and 80.7% recall (Wierzbicki, Matuk, and Bielecka 2021). Buildings were identified with 93.50% recall and 87.60% precision using 0.25 m spatial resolution digital surface models (DSM) from aerial stereo images and vegetation cover data produced to identify typical urban areas in Shanghai, China using region growing method (B. Wu et al. 2022). The DeepLab v3 model was used to predict the installable areas of rooftop mountable photovoltaic solar panels from Google Earth satellite images with a spatial resolution of 0.25 m. The study found 92% accuracy and 79% precision for Nanjing rooftops (Zhong et al. 2021). Zhang et al. used the Massachusetts buildings dataset with a spatial resolution of 1 m (F-Score is 96.72%), the IAIL dataset with a spatial resolution of 30 cm (accuracy is 96.46%), and the DeepGlobe Building Detection Challenge dataset with a pansharpened 31 cm spatial resolution (accuracy is 96.83%) to test the effectiveness of their proposed local-global dual-stream network in detecting roofs (Zhang et al. 2020). Boonpook et al (Boonpook, Tan, and Xu 2021). present a system for deep learning to recognize buildings from unmanned aerial vehicles (UAV) images. According to the results, increasing the number of features

enhances classification accuracy. The SegNet architecture was used for classification, and the utilization of the other features (e.g. DSM or VDVI), together with RGB, raised the F-Score in the test areas by more than 90%. In order to detect building footprints, the classification of Google Earth images with 0.4 m spatial resolution for urban areas utilizing a number of variables and a Random Forest method have been examined by Thottolil et al. The proposed feature set from the RGB image produced 85.62% accuracy (Thottolil and Kumar 2022). Virtrianaa et al. (Virtriana et al. 2023) used the Random Forest to detect building damage caused by the Anak Krakatau Volcano tsunami. 14 distinct feature sets were generated and classified using the Random Forest for the three data sets xBD, Copernicus, and field survey data in order to confirm and evaluate the results. Using the geometry, texture + vegetation index, and geometry feature sets, the best F-score for the xBD, Copernicus, and field survey data was obtained as 58.6%, 78.3%, and 57.8%, respectively. Aside from the studies mentioned above, the following papers using machine learning approaches have been published: Support Vector Machine (SVM) (Dornaika et al. 2016), AdaBoost, Naïve Bayes, J_{48}, Random Forest and SVM (Cohen et al. 2016), k-nearest neighbour (k-NN) (Chandra and Ghosh 2018), artificial neural networks (Lari and Ebadi 2007).

Pleiades, GeoEye, QuickBird, WorldView, IKONOS, and Gaofen-2 are just a few examples of the satellite and satellite based VHR images that can give a huge amount of spatial detail information. Furthermore, with aerial cameras installed on UAVs and professional aircraft, significantly higher resolution images can be obtained due to the rapid development of photogrammetric software, hardware breakthroughs, and improvements in imaging sensors. Spatial information such as orthophoto, ortho-mosaic, solid model, and digital surface model can be constructed using the images collected with these air vehicles. Due to the heterogeneity of buildings and their environment, robust building detection from VHR images is not a straightforward operation and remains a problem. Numerous researches have explored this topic and presented numerous approaches from various viewpoints to address these issues.

Because of the higher expense and difficulties in analysing LiDAR data due to its distinctive data collection technique, optical images remain the most essential data source. Optical images provide rich information about the colours, textures, and details of buildings. Also, optical images help us better understand the relationships between objects. In the context of building detection, this increases the ability to understand the relationship of buildings to their environments and better predict the intended use of buildings. Optical cameras generally have a wider field of view. This makes it possible to detect buildings in more areas simultaneously, ensuring fast and efficient detection. The primary issue with building detection methods is the mixing up of the building class with other objects including vegetation, the ground, complicated building structures, and shadows. Misclassifying a non-building as a building and mixing up classes like trees-shade are other issues. Due to the intricacy of the classification process, many approaches and strategies, including feature extraction and selection, have been offered to address the problems it has brought up.

In this study, true orthophoto produced photogrammetrically using images obtained with professional aircraft, visible band difference vegetation index (VDVI) produced from this orthophoto, digital surface model (DSM) with 5 m spatial resolution produced by automatic mapping from 30 cm resolution stereo aerial images have been used. Also,

TanDEM-X High Resolution Elevation Exchange (TREx) digital elevation model (DEM) with 12 m spatial resolution has been used. Finally, the topographic position index (TPI) and terrain ruggedness index (TRI) produced from these DSM and DEM have been created. For the binary classification of stacked related features (building and other), the Auto-Encoder deep learning architecture, k-NN, Decision Tree, RUSBoost, and Random Forest classifiers are employed.

Rest of the paper is structured as follows: Section 2 introduces the materials and methods, Section 3 presents the results, Section 4 discusses the experimental results, and Section 4 contains the conclusion.

2. Materials and methods

This research was conducted in the Elbistan district of Kahramanmaraş, Türkiye, the epicentre of the Elbistan (7.6 MW) earthquake on 6 February 2023. The location map of the city is presented in Figure 1. At the time of this study, the earthquake disaster had not yet occurred. Most of the buildings in the study area do not currently exist. It covers about 2.319 km^2 of land. The research region lies between the coordinates of 38°24′45″ – 38°16′30″N and 37°15′14″ – 37°25′04″E. The growing city is Turkey's fourth-largest plain and home to the third-largest lignite

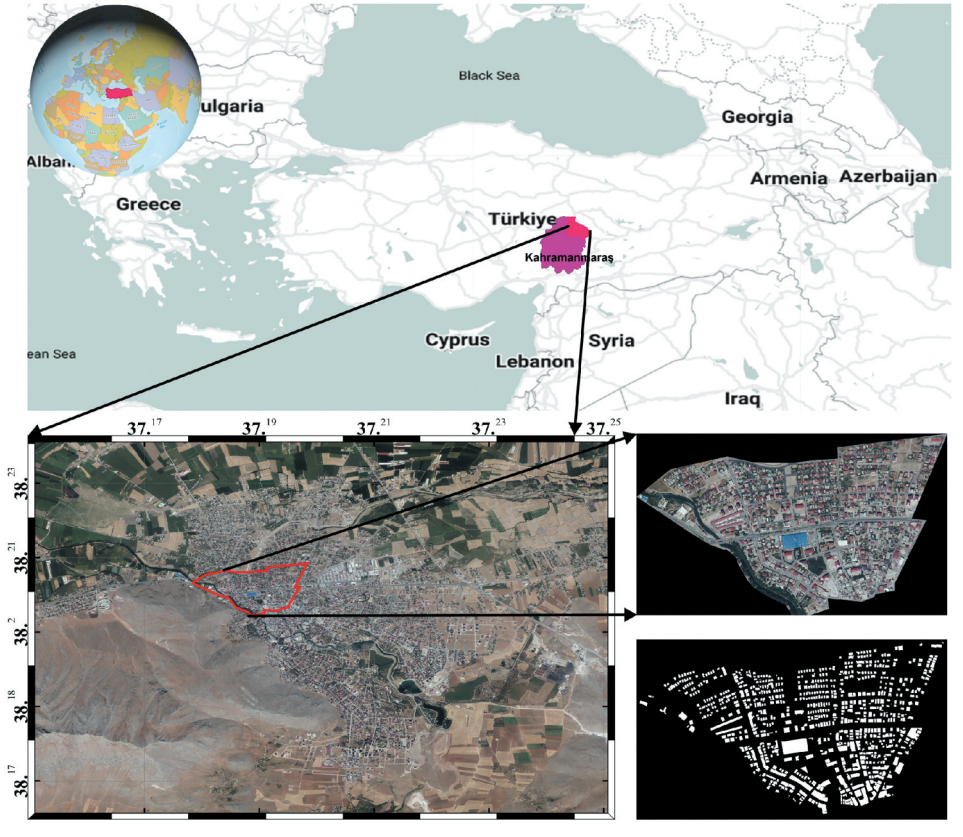

Figure 1. Location map of Elbistan, Kahramanmaraş, Türkiye.

thermal power plant. Due to this, urbanization is accelerating, and the city is home to both new and historic buildings. This city was chosen as the research area because it has the capacity to identify illegal structures, has the potential to produce solar photovoltaic energy because of its geopolitical location (Günen 2021), and may serve as a base map for planning.

2.1. Proposed framework

Various features used with VHR images increase the roof detection performance. A comprehensive literature review revealed that there are studies that use RGB and DSM (Ferrari et al. 2021; Marmanis et al. 2018; Sun and Wang 2018), RGB and VDVI (Y. Wang et al. 2022), DSM and VDVI (Liu et al. 2018; Xiaoqin et al. 2015; Xu et al. 2018), RGB, DSM and VDVI (Boonpook, Tan, and Xu 2021; Chauhan et al. 2022). This study compares several classification algorithms and feature sets obtained from a combination of VHR image, DEM, DSM, including one index (VDVI) calculated from the VHR image and two indices (TRI and TPI) calculated from the DEM and DSM to detect buildings with complex geometry and texture. Furthermore, no study has been found in which roof detection was performed using a feature set that included the TRI and TPI indexes. The work is significant because of the ongoing need for practical and effective algorithms to use for building detection in natural disaster response. TPI and TRI images produced from DSM and DEM data are named as TPI5, TRI5, TPI12 and TRI12, respectively. DSM5 and DEM12 represent the spatial resolution of DSM and DEM, respectively. RGB bands of orthophoto (F_{RGB}) are given in ECW format and are georeferenced using WGS84 geographic coordinate system. It has a 25 cm spatial resolution, and is located on the L38d2 in 1/100,000 map sheet. 5 m grid spacing DSM (DSM5) was created using an automatic matching method from stereo aerial images with a 30 cm resolution. Resampled data from Level-0 data has a vertical accuracy of 3 m at the 90% confidence interval, covering all details like human structure and vegetation in the land topography (URL 2022). The radar interferometry method was used to create the digital elevation model created as part of TanDEM-X High Resolution Elevation Exchange-TREx, which has a spatial resolution of 12 m (DEM12). The raster data for this study were acquired in the third and fourth quarters of 2019. The flowchart shown in Figure 2 was used in this study to implement building detection using classifiers with various architectures on VHR images.

The VDVI, an RGB-based image index, is used to increase vegetation detail while reducing other transmitted data in geo-images. It can be computed using Equation 1, with a value ranging from −1 to 1.

$$\mathrm{VDVI} = \frac{2 \times Green - Red - Blue}{2 \times Green + Red + Blue} \tag{1}$$

The TPI is a measure of the topographic position of a surface. The TPI indicates how the average height around a point affects the altitude measurement at that point. A point's TPI value indicates whether it is higher or lower than the average of its surroundings. If it is positive, the point is higher. Woodlands, rocky regions, and other topographic features are frequently identified using TPI (De Reu et al. 2013; Weiss 2001). TPI can be calculated using Equation 2.

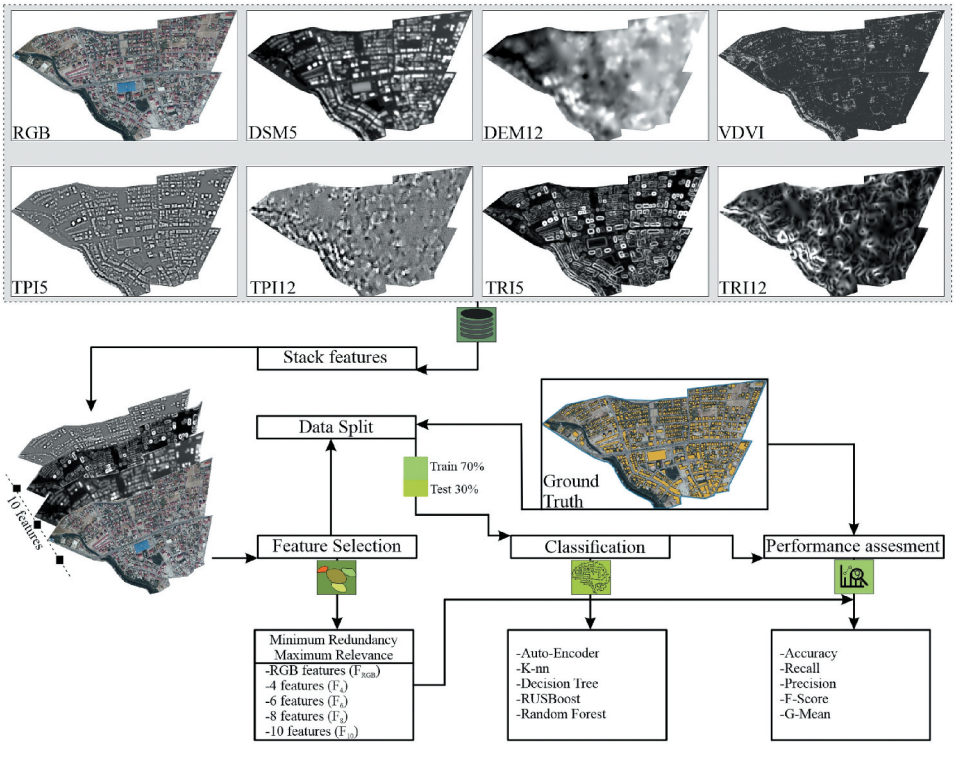

Figure 2. Flowchart of proposed method.

$$TPI = \left(\frac{Z_0 - Z_{mean}}{Z_{std}}\right) * 100 + 0.5 \tag{2}$$

The TRI (Riley, DeGloria, and Elliot 1999; Shepard et al. 2001) is a measure of a terrain's roughness or ruggedness that is calculated by subtracting the maximum and minimum elevation values from a moving window over the terrain surface. TRI is calculated using Equation 3.

$$TRI = \sqrt{\sum (Z_c - Z_0)^2} \tag{3}$$

Where Z_0 is query point and mean, standard deviation and central of the elevation values around the point Z_0 is Z_{mean}, Z_{std}, and Z_c respectively. It is essential to remember that the window size is typically determined by the user and can be altered based on the scope of the research. The larger the window size, the smoother the terrain will appear, while a smaller window size will result in a more detailed representation of the terrain's ruggedness. 3×3 window size was used when creating TRI and TPI images.

2.2. Feature selection

Feature selection is a machine learning strategy. This approach identifies the dataset's most important features. This ensures that worthless or unimportant features in the dataset are not included for the model when it is suspected that they will lower the

model's performance. Feature selection can also be used to eliminate correlations and redundancy among characteristics in a dataset. Benefits such as improving model performance, making the model easier to understand, and boosting the model's working speed can be obtained by selecting features (Günen 2022a; Günen, Atasever, and Beşdok 2020). The Maximum Relevance Minimum Redundancy (MRMR) approach is a feature selection method (Ding and Peng 2005; Günen 2022b). This method analyses the associations between features in the dataset and selects features that have the highest feature-target relationship. The MRMR approach is also intended to eliminate redundancy among dataset features. This strategy is beneficial if the dataset has a large number of features with complex interactions between them. In this investigation, MRMR was utilized since it is a simple and quick method with low calculation costs.

2.3. Classification methods

2.3.1. Auto-encoder

Auto-encoder is one of the deep learning architectures. The autoencoder's primary goal is to use learned encoding to convert the input into a reduced dimensional space. A particular kind of feedforward neural network called an auto-encoder uses the input and the target to be the same. An autoencoder consists of three components: encoder, code, and decoder. The encoder, which also generates a code, compresses the input. The code is then used by the decoder to completely reconstruct the input. Auto-encoders can be used to condense the features and eliminate unnecessary or redundant information. Additionally, auto-encoders can be used to reduce the dimension of the dataset, which facilitates understanding and simpler manipulation of the dataset using other models. Although they can also be utilized for supervised learning issues, auto-encoders are frequently used for unsupervised learning (Günen, Atasever, and Beşdok 2020; Zabalza et al. 2016).

2.3.2. k-NN

This method is used to classify the in the dataset. The k-NN classifier assigns the class of the k data points closest to a query point in the dataset. A distance function is created and used for this and k is determined by the user. The k-NN assesses how similar or close data points are to each other based on this strategy. It is a type of supervised learning since, like any classifier, it needs some training data with predetermined labels (Cover and Hart 1967; Günen 2022c).

2.3.3. Decision tree

A tree structure is used in the decision tree approach to classify in a dataset. The dataset is classified using a tree structure that looks at the correlations and relationships between the in the dataset and uses those observations. The root or root node refers to the initial nodes in a decision tree. According to the root condition, each observation is assigned a 'Yes' or 'No' classification. There is a part called interval nodes or nodes under the stem cells that allows data to be classified. The model's complexity grows as the data size increases. The part named leaf nodes or leaves, located at the bottom of the decision tree, provides the classification result. The main

challenge in taking the decision tree into practice is figuring out which characteristics would be taken into account at each level and as the root node (Günen 2022a; Kavzoglu, Kutlug Sahin, and Colkesen 2015).

2.3.4. RUSBoost

Random Under Sampling with Boosting (RUSBoost) method is used to stabilize unbalanced classes in the dataset. This method randomly selects the data of the lesser number of the classes in the dataset, increasing the data of the fewer classes in the dataset. In this way, it is ensured that the model is trained with the data of the lesser number of the classes in the dataset. The software trains one weak learner during each learning cycle. The RUSBoost method is also used to better understand in the dataset with the Boosting method. The boosting method involves using more than one decision tree and combining the results of these decision trees. This provides a better understanding of in the dataset (Kesikoglu et al. 2016; Seiffert et al. 2008).

2.3.5. Random forest

Random forest, developed by Breiman (Breiman 2001), is one of the most popular ensemble methods. Random forest involves the creation of many decision trees and the combination of their outcomes. As a result, it improves understanding of in the dataset and lowers variation between trees. Using a random selection of features, the random forest method creates decision trees. A random sample of the original training data is used to train each tree. Thus, utilizing diverse decision tree attributes minimizes tree similarity, avoiding overfitting issues. It could incorporate several strategies for combining trees. For example, tree results are averaged or the most frequent outcome is selected. The number of potential splitting attributes and the number of trees are used to fine-tune the random forest (Breiman 2001; Günen 2022a).

2.4. Performance evaluation

The confusion matrix can be used to reveal omission and commission errors for each class. Confusion matrix elements true positive (TP), false positive (FP), true negative (TN), and false negative (FN) are used to statistically analyse classification methods. Confusion matrix elements are presented in Table 1. It is possible to generate a number of classification performance metrics utilizing related factors, including Accuracy, Specificity, User's Accuracy/Precision, Producer's Accuracy/Recall/Sensitivity, G-mean and F-Score given in Equation 4–9.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

Table 1. Confusion matrix elements.

		Model$_i$	
Ground Truth/Model$_j$		Correct	Incorrect
	Correct	TP	FP
	Incorrect	FN	TN

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Specificity = \frac{TN}{TP + FP} \tag{7}$$

$$F - Score = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{8}$$

$$G - mean = \sqrt{Recall \times Specificity} \tag{9}$$

Statistical analysis is performed to determine whether differences between two groups on comparison scales are statistically significant using McNemar's test. This test is designed to evaluate how well two distinct classifiers perform. Confusion matrix is used to calculate the McNemar's test, which is based on the chi-square distribution. The two classifiers perform statistically similar if the calculated statistical value (χ^2) is lower than the value of 3.84 in the 95% confidence interval. Equation 10 used to calculate χ^2 value (Günen 2022c; Mojaddadi Rizeei, Pradhan, and Saharkhiz 2019).

$$\chi^2 = \frac{(|FP - FN| - 1)^2}{FP + FN} \tag{10}$$

3. Results

Developing cities have a uniform urban type as a result of ongoing migration, but they also have unique structures in the old rural type. In terms of planning, it is crucial to continuously observe various types of structures. VHR images allow for the depiction of building details, improving the accuracy of building location, form, and size estimations. Figure 3 depicts the many types of structures and roofs present in Elbistan. It is critical to provide new methodologies for applications where fast building detection process is carried out by detecting the roof from the VHR image. The primary goal of this research is to develop building detection training data with new features and then use this data to identify the model with the best classification performance. These features can be validated by comparing the input and output data.

Republic of Türkiye Ministry of National Defense General Directorate of Maps provides VHR image, DSM, and DEM. The relevant data were georeferenced with the on-site GNSS measurement and the co-registration process was carried out by the General Directorate of Mapping of Turkey. The spatial resolution of the data varies. These data must be expressed at the same spatial resolution in order to be combined and stacked in the geographic information system. Unlike the Wald protocol, low spatial resolution images are up sampled to high resolution images to avoid data loss. Actually, the goal is to make the lower resolution image grids smaller so that the spatial resolution of the data remains same. The bicubic interpolation technique was used to up-sample the DSM and DEM data

Figure 3. Different roof types: a) burgundy painted trapezoidal sheet, b) flat building, c) tile covered roof, d) sinus corrugated sheet, e) rusted sheet roof and f) shingle roof.

to the VHR image. Bicubic interpolation was selected because it allows for more accurate visualization and analysis of topographic data, demonstrates relationships between data points more precisely, and produces more consistent and reliable results by reducing sudden fluctuations in the dataset. The three primary data sources were then cropped to cover the study area. When choosing the study area, care was taken to include various roof types, structures of varying heights, and classes such as asphalt, vegetation, water, and permeable/impermeable surfaces. Later, it was intended to easily determine vegetation classes after the classification process by constructing a VDVI index image from the VHR image. The MRMR technique was used to rank the 10 features (3 features come from RGB) to determine which were the most significant. The MRMR, which takes into consideration feature dependencies, has been chosen because it determines the most important features in the dataset in a robust and quick manner, even when the data size is quite high, by decreasing feature redundancy in the dataset. Feature importance scores are given in Figure 4. Then, performance analysis was carried out for the 4, 6, 8 and 10 feature sets, which have the most importance and are named F_4, F_6, F_8, and F_{10}, respectively. To the best of our knowledge, there is no framework that uses a feature selection strategy to combine these 10 features into a feature set for building detection. For example, F_4 includes DSM5, VDVI, TPI5, blue band features. Four of the ten features were chosen since it is a time-consuming and complex process to assess all features by decreasing them one by one. Because there are many papers in the literature that only employ the RGB feature (F_{RGB}) from VHR images, this feature set has also been examined. It has been determined which feature sets are more important in the detection of buildings. According to the order of importance obtained by MRMI, the most important features are respectively: DSM5, VDVI, and TPI5.

In summary, DSM5 is the most important feature, whereas TRI12 is the least important. Following the feature selection procedure, the data was separated into 70% training and 30% test samples, ensuring that the same samples were utilized in both feature sets. In

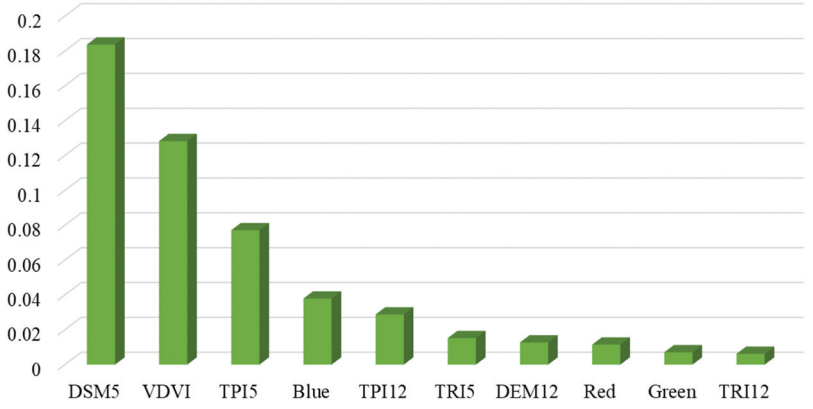

Figure 4. Feature importance scores.

Table 2. The hyperparameter values used in classification methods.

Methods	Hyper-parameters	Value
Auto-Encoder	Epoch	500
	Softmax Epoch	1000
	1st hidden layer	30
	2nd hidden layer	30
	1st transfer functions	logsig
	2nd transfer functions	purelin
	Weight Regularization	0.1
	Sparsity Proportion	0.15
	Sparsity Regularization	4
	Training algorithms	RP
k-NN	Number of Neighbors	9
	Distance Criterion	Euclidean
	Distance Weight	Equal
Decision Tree	Splitting Criterion	Gini
	Maximal number of decision splits	20
RUSBoost	Robust Error Goal	0.15
	Number of ensembles learning cycles	30
Random Forest	Number of ensembles learning cycles	30
	Maximal number of decision splits	20

order to learn the obtained training data set, classifiers with various architecture such as deep learning, boosting, bagging and non-parametric were used. A variety of hyper-parameters are employed with the classifiers. These parameters, as shown in Table 2, have a significant impact on the performance of the classifiers. Because it is not possible to examine the complete data set or the portion used for training in terms of time and expense, a small portion (about 5%) was taken and the optimum values were established via grid search.

Resilient backpropagation algorithm (RP) was chosen for Auto-Encoder training because it produced the best results in previous studies (Günen et al. 2020). Because Ananthi and Sathyabama (Ananthi and Sathyabama 2009) suggested that the nearest neighbour (k) in the k-NN technique be an odd number rather than an even number, a study on odd numbers was done. The Gini index is utilized in the decision tree approach because the calculations are simpler, it is faster, and it produces more stable results for numerous classes in feature selection (Chan and

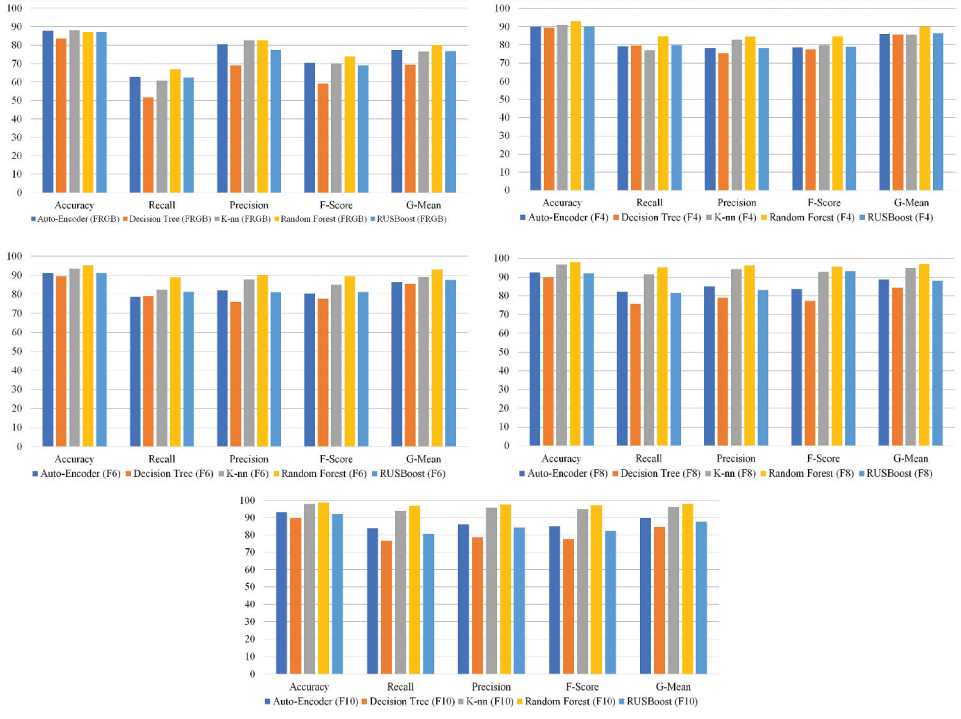

Figure 5. Performance comparison of classification algorithms.

Paelinckx 2008). The number of ensembles learning cycles was searched between 10–100 in multiples of 10 in the RUSBoost and random forest approaches. After the classification process is performed, it is very important to present the results of the test images quantitatively and qualitatively according to the ground truth data. In this context, the classification performance assessment of the test data is given in Figure 5. Here, separate figures are created for each feature set. In general, it is seen that F_{RGB} has the lowest classification performance and F_{10} has the highest classification performance. When comparing the F_4 and F_6 feature sets, it is seen that the TPI and TRI indexes in the proposed framework increase the classification performance. In addition, when F_6 and F_8 are compared, the effect of DEM12 and red band on classification performance is quite effective despite the low feature importance score (Figure 4). It is seen that the classification performance for the same classifier increases as the proposed features are added. The visual representation of the whole dataset as a result of the binary classification is presented in Figure 6, and a qualitative examination of the overlay map obtained with the feature sets giving the best, medium and weakest classification performance is presented in Figure 7.

MSI Delta 15 laptop with AMD Ryzen 9 9500 HX processor, R×6700graphics card, and 32 Gb RAM is used for all applications. The open-source program QGIS was employed for ground truth generation, data clipping, and up-sampling operations, and MATLAB 2021b was used for classification processes.

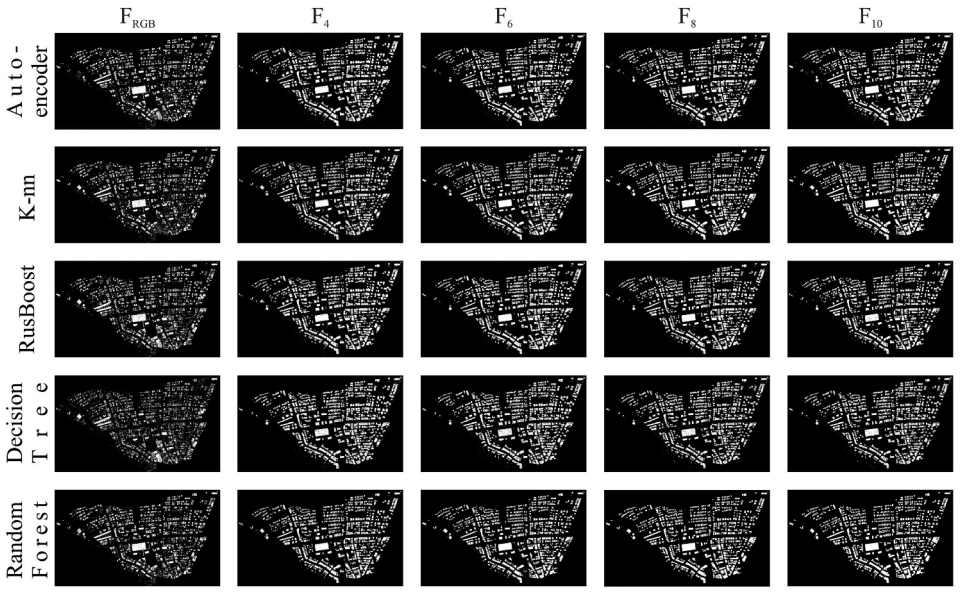

Figure 6. Binary classification results of the dataset.

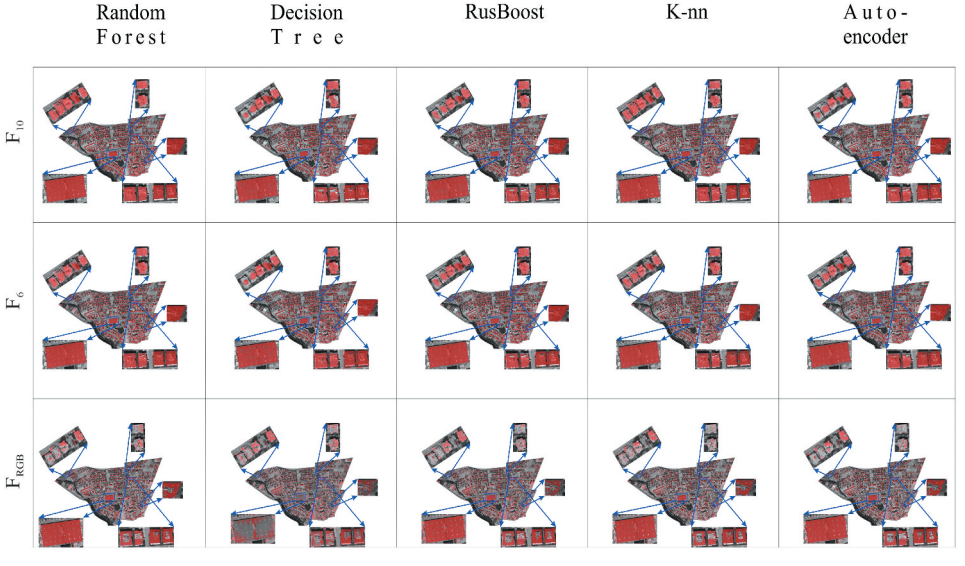

Figure 7. Zoomed result of different roof types. (red areas represent the classification result, gray areas represent the gray scaled VHR).

4. Discussion

In order to enable the learning of main features for fast building detection, it is crucial to provide a sufficient dataset and feature sets for classification algorithms. Since accuracy is insufficient for large unbalanced data sets, recall, precision, F-score, and G-mean metrics were also utilized for performance assessment. Examining the results indicates that the

Random Forest approach yields the best results. k-NN and auto-encoder algorithms follow random forest, respectively. The F_{10} and F_8 feature sets produced the best results in all classifiers. Among all of the classifiers and feature sets, the Decision Tree method produced the least successful result (for F_{RGB}).

The running times are shown in Figure 8. The running times are displayed using logarithms to make the visual presentation clear. When these findings are analysed, it becomes clear that decision tree and k-NN are relatively quick (for instance, decision tree (F_{RGB}) takes around 6 seconds and k-NN (F_{RGB}) takes about 15 seconds). On the other hand, Auto encoder F_{10} and random forest F_{10} took 54,752 and 2,611 seconds, respectively, to complete the training procedure.

In terms of textural, shape, and illumination circumstances, the dataset employed is fairly hard. Geometric and material variations, as shown in Figure 3, make classification difficult. Furthermore, roof components such as antennas, chimneys, terrace, domes, and water tanks make classification difficult. These are the most significant elements affecting F_{RGB} classification success. Furthermore, single-story buildings are frequently dwellings with a garden, and their roofs are usually shaded by tree leaves, trees, and multi-story structures. Powerful classifiers such as support vector machines could be used in addition to the methods used to perform the classification process, but this method was not included in the comparison because it was not feasible in the F_{RGB} feature set, as it performed the classification process in a longer time than the random forest F_{10} feature.

The raster image produced by the classification of the F_{10} feature of random forest, which produced the best results, was turned into a polygon using QGIS software. This converted polygon data serves as the foundation for numerous base maps. Thanks to this extracted polygon information, it is ensured that the vector information of the destroyed buildings is created quickly. This generated polygon data (Figure 9) is expected to be published in an open street map for Elbistan. Building polygon correction, on the other

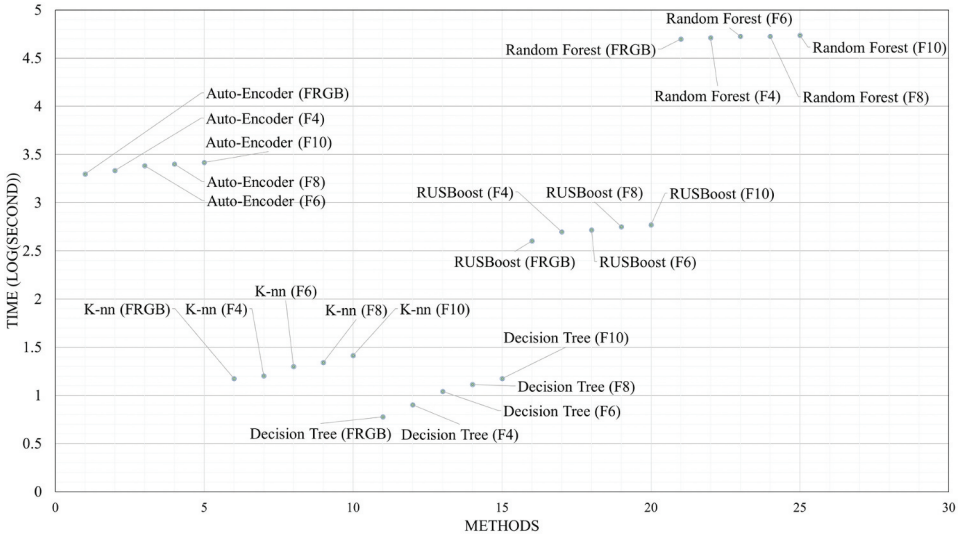

Figure 8. Comparison of the runtimes of the methods.

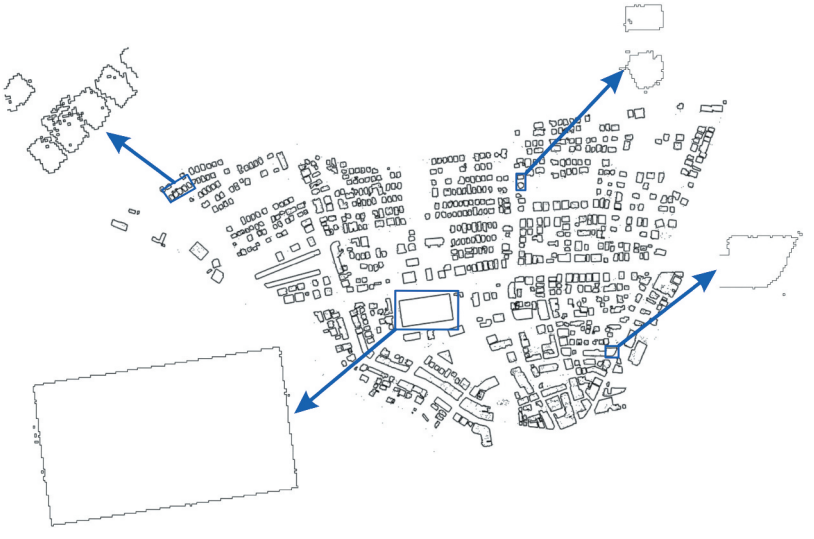

Figure 9. Polygon map of test area.

Table 3. McNemar's test results.

	Auto-Encoder	k-NN	Decision Tree	RUSBoost	Random Forest
Auto-Encoder	-	11.76	14.77	6.50	6.89
k-NN		-	33.35	22.92	**1.79**
Decision Tree			-	**3.03**	25.68
RUSBoost				-	16.25
Random Forest					-

hand, is the subject of a distinct study. This research does not concentrate on polygon regularization.

Table 3 displays the χ^2 values obtained by applying the McNemar's test to assess whether there is a statistically significant difference between the methods. The null hypothesis stated that there are no differences between classifiers. It is rejected if the derived test statistic value is greater than the critical chi-square value ($\chi^2 > 3.84$, $p < 0.05$). McNemar's tests revealed no statistically significant difference between decision tree vs RUSBoost and random forest vs k-NN.

Although the proposed framework's feature sets improve classification performance according to the F_{RGB}, they have drawbacks and limitations, such as the large dimensionality of the feature sets, which makes data classification challenging. When Figure 5 is examined, it is clear that F_{RGB} has the worst classification performances while F_{10} and F_8 have the best classification performances in general. In addition, as can be seen in Figure 8, the increase in dimension also increases the running time of the classifiers and causes cumbersomeness. As a result, the availability of simpler structured classifiers for problem solving by applying classifications in different architectures was explored, and the k-NN approach, which provides a very flexible and simple solution, was chosen to be the best classifier. DSM and DEM data can be easily produced by obtaining VHR images utilizing UAVs, even if the study area's data is not received from an institution, as in this study. Furthermore, depending on the UAVs equipment, the collected data may have

a better spatial resolution. Since the most importance is assigned to DSM5, VDVI and TPI5 after feature selection, orthophoto, DSM, and DEM from VHR images obtained with UAV can be used for detection of collapsed building with k-NN with high classification performance. Therefore, the proposed framework is available for testing by researchers for many geographic areas, making the methodology very applicable. Main problem for classification-based approach is the selection of classifier hyper-parameters. Working with high-dimensional and sampled data is a limitation because determining these parameters with a small portion of the data doesn't really address the entire data. In light of the fact that there is no statistically significant difference between the random forest approach, which yields the best classification result, and the k-NN method, adopting the k-NN method for quick solutions is a more adaptable and practical option. The parameter optimization process moves fast because k-NN operates quickly. It should be emphasized that although while k-NN trains more quickly than Random Forest, the predicted running time can vary significantly depending on a number of variables, including the amount of data and the optimization structure used in k-NN. Random Forest may be preferable in very large datasets since it has a more predictable running time for prediction and could be faster in very large datasets.

5. Conclusion

Accurate detection of building roofs has increasingly become possible with the widespread application of high-resolution remote sensing images from low altitude aerial photogrammetry. The complexity of the various building roof types that are evident in images and the ease with which other elements (e.g. water tank, solar panel, shade) might be mistaken for building roofs, however, pose a significant obstacle to the current approaches. In this study, building roofs are determined with high accuracy by using visible bands and VDVI index produced from them, as well as DSM, DEM and TRI and TPI indexes produced from them, respectively. The effectiveness of the classifiers and the effectiveness of the feature sets were compared by classifying the various feature sets that had been produced. The usage of the k-NN approach is recommended because it operates significantly more quickly. The following are some major conclusions and contributions to the literature that can be derived from the current study:

(1) Given that the proposed F_{10} feature set is not much different from F_{RGB} in terms of time and significantly increases accuracy, it is critical that it be used in various test areas and that it improves the findings of studies conducted solely with F_{RGB} in the literature.
(2) Random Forest outperformed competing algorithms across all performance parameters, despite its poor processing speed. The random forest outperformed the k-NN, Auto-encoder, RUSBoost, and decision tree by approximately 3%, 12%, 15%, and 20%, respectively, when considering the F-Score.
(3) The most important features as a result of MRMR feature selection are DSM5, VDVI, and TPI5. The visible bands are of limited impact on determining the building, but the VDVI created from them and DSM are the most important features.
(4) The ability to convert the resulting raster image into vector data demonstrates that the data may be used for navigation, base maps, and the lod-2 3D building model.

(5) The results of the McNemar's test revealed that, with the exception of k-NN, there was a statistically significant difference between Random Forest and the other algorithms.
(6) There is no statistically significant difference between the decision tree and RUSBoost approaches, despite there being a 2% difference in accuracy and a 5% difference in F-Score.

The 6 February 2023 Elbistan (7.6 MW) earthquake had not yet occurred at the time of this study. According to the early analysis, around 1500 buildings in the test area are mildly to moderately damaged or collapsed. The applicability of the proposed method will be demonstrated in the next study after a possible earthquake by completing a UAV flight, with the aim that collapsed structures will be spotted quickly and will aid people working in disaster management. VHR, DSM and DEM will be produced quickly and practically using UAV. In this way, a geographic information system will be created in order to provide speedy access to citizens who are under the debris in the event of the next possible earthquake. There will also be efforts to improve polygon regulation in order to enhance the effectiveness of this procedure.

Acknowledgements

Thanks to Republic of Türkiye Ministry of National Defense General Directorate of Maps.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Mehmet Akif Günen http://orcid.org/0000-0001-5164-375X

Author contributions

MAG was responsible for the study's design, research, and writing.

Data availability statement

The required references and the accompanying author can provide the datasets.

References

Ananthi, S., and S. Sathyabama. 2009. "Spam Filtering Using K-NN." *Journal of Computer Applications* 2 (3): 20.

Boonpook, W., Y. Tan, and B. Xu. 2021. "Deep Learning-Based Multi-Feature Semantic Segmentation in Building Extraction from Images of UAV Photogrammetry." *International Journal of Remote Sensing* 42 (1): 1–19. https://doi.org/10.1080/01431161.2020.1788742.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Chan, J. C. W., and D. Paelinckx. 2008. "Evaluation of Random Forest and Adaboost Tree-Based Ensemble Classification and Spectral Band Selection for Ecotope Mapping Using Airborne Hyperspectral Imagery." *Remote Sensing of Environment* 112 (6): 2999–3011. https://doi.org/10.1016/j.rse.2008.02.011.

Chandra, N., and J. K. Ghosh. 2018. "A Cognitive Viewpoint on Building Detection from Remotely Sensed Multispectral Images." *IETE Journal of Research* 64 (2): 165–175. https://doi.org/10.1080/03772063.2017.1351320.

Chauhan, N., R. Kumar, S. Mukherjee, A. Hazra, and K. Giri. 2022. "Ultra-resolution unmanned aerial vehicle (UAV) and digital surface model (DSM) data-based automatic extraction of urban features using object-based image analysis approach in Gurugram, Haryana." *Applied Geomatics* 14 (4): 751–764. https://doi.org/10.1007/s12518-022-00466-8.

Cohen, J. P., W. Ding, C. Kuhlman, A. Chen, and L. Di. 2016. "Rapid Building Detection Using Machine Learning." *Applied Intelligence* 45 (2): 443–457. https://doi.org/10.1007/s10489-016-0762-6.

Cover, T., and P. Hart. 1967. "Nearest Neighbor Pattern Classification." *IEEE Transactions on Information Theory* 13 (1): 21–27. https://doi.org/10.1109/TIT.1967.1053964.

De Reu, J., J. Bourgeois, M. Bats, A. Zwertvaegher, V. Gelorini, P. De Smedt, and P. Finke, M. Antrop, P. De Maeyer, P. Finke, M. Van Meirvenne. 2013. "Application of the Topographic Position Index to Heterogeneous Landscapes." *Geomorphology* 186:39–49. https://doi.org/10.1016/j.geomorph.2012.12.015.

Ding, C., and H. Peng. 2005. "Minimum redundancy feature selection from microarray gene expression data." *Journal of Bioinformatics and Computational Biology* 3 (2): 185–205. https://doi.org/10.1142/s0219720005001004.

Dornaika, F., A. Moujahid, Y. El Merabet, and Y. Ruichek. 2016. "Building Detection from Orthophotos Using a Machine Learning Approach: An Empirical Study on Image Segmentation and Descriptors." *Expert Systems with Applications* 58:130–142. https://doi.org/10.1016/j.eswa.2016.03.024.

Ferrari, L., F. Dell'acqua, P. Zhang, and P. Du. 2021. "Integrating EfficientNet into an HAFNet Structure for Building Mapping in High-Resolution Optical Earth Observation Data." *Remote Sensing* 13 (21): 4361. https://doi.org/10.3390/rs13214361.

Gualtieri, A. F., G. M. Lassinantti, V. Scognamiglio, and D. Di Giuseppe. 2022. "Human Health Hazards Associated with Asbestos in Building Materials." *Ecological and Health Effects of Building Materials* 297–325. https://doi.org/10.1007/978-3-030-76073-1_16.

Günen, M. A. 2021. "A Comprehensive Framework Based on GIS-AHP for the Installation of Solar PV Farms in Kahramanmaraş, Turkey." *Renewable Energy* 178:212–225. https://doi.org/10.1016/j.renene.2021.06.078.

Günen, M. A. 2022a. "Adaptive Neighborhood Size and Effective Geometric Features Selection for 3D Scattered Point Cloud Classification." *Applied Soft Computing* 115. https://doi.org/10.1016/j.asoc.2021.108196.

Günen, M. A. 2022b. "Nokta Bulutu Verisi Kullanılarak Elma Bahçesinden Meyve Tespiti." *El-Cezeri* 9 (1): 253–265. https://doi.org/10.31202/ecjse.962269.

Günen, M. A. 2022c. "Performance Comparison of Deep Learning and Machine Learning Methods in Determining Wetland Water Areas Using EuroSAT Dataset." *Environmental Science and Pollution Research* 29 (14): 21092–21106. https://doi.org/10.1007/s11356-021-17177-z.

Günen, M. A., U. H. Atasever, and E. Beşdok. 2020. "Analyzing the Contribution of Training Algorithms on Deep Neural Networks for Hyperspectral Image Classification." *Photogrammetric Engineering & Remote Sensing* 86 (9): 581–588. https://doi.org/10.14358/PERS.86.9.581.

Günen, M. A., Ü. H. Atasever, T. Taşkanat, and E. Beşdok. 2019. "Usage of Unmanned Aerial Vehicles (UAVs) in Determining Drainage Networks." *Nature Sciences* 14 (1): 1–10. https://doi.org/10.12739/NWSA.2019.14.1.4A0062.

Günen, M. A., E. Besdok, P. Civicioglu, and U. H. Atasever. 2020. "Camera Calibration by Using Weighted Differential Evolution Algorithm: A Comparative Study with ABC, PSO, COBIDE, DE,

CS, GWO, TLBO, MVMO, FOA, LSHADE, ZHANG and BOUGUET." *Neural Computing & Applications* 32 (23): 17681–17701. https://doi.org/10.1007/s00521-020-04944-1.

Han, Y., J. Choi, J. Jung, A. Chang, S. Oh, and J. Yeom. 2019. "Automated Coregistration of Multisensor Orthophotos Generated from Unmanned Aerial Vehicle Platforms." *Journal of Sensors* 2019. https://doi.org/10.1155/2019/2962734.

Huang, X., and L. Zhang. 2011a. "Morphological Building/Shadow Index for Building Extraction from High-Resolution Imagery Over Urban Areas." *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing* 5 (1): 161–172. https://doi.org/10.1109/JSTARS.2011.2168195.

Huang, X., and L. Zhang. 2011b. "A Multidirectional and Multiscale Morphological Index for Automatic Building Extraction from Multispectral GeoEye-1 Imagery." *Photogrammetric Engineering & Remote Sensing* 77 (7): 721–732. https://doi.org/10.14358/PERS.77.7.721.

Jing, W., J. Lin, H. Lu, G. Chen, and H. Song. 2022. "Learning Holistic and Discriminative Features via an Efficient External Memory Module for Building Extraction in Remote Sensing Images." *Building & Environment* 222. https://doi.org/10.1016/j.buildenv.2022.109332.

Kavzoglu, T., E. Kutlug Sahin, and I. Colkesen. 2015. "An Assessment of Multivariate and Bivariate Approaches in Landslide Susceptibility Mapping: A Case Study of Duzkoy District." *Natural Hazards* 76 (1): 471–496. https://doi.org/10.1007/s11069-014-1506-8.

Kesikoglu, M. H., U. H. Atasever, C. Ozkan, and E. Besdok. 2016. "The Usage of Rusboost Boosting Method for Classification of Impervious Surfaces." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 40:981–985. https://doi.org/10.5194/isprs-archives-XLI-B7-981-2016.

Kocaman, S., L. Zhang, A. Gruen, and D. Poli. 2006. "3D City Modeling from High-Resolution Satellite Images." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (1/W41). https://doi.org/10.3929/ethz-b-000158058.

Lari, Z., and H. Ebadi. 2007. "*Automatic Extraction of Building Features from High Resolution Satellite Images Using Artificial Neural Networks*." Paper presented at the ISPRS Working Group I/5 and IV/3:"High Resolution Earth Imaging for Geospatial Information.

Li, J., X. Huang, L. Tu, T. Zhang, and L. Wang. 2022. "A Review of Building Detection from Very High Resolution Optical Remote Sensing Images." *GIScience & Remote Sensing* 59 (1): 1199–1225. https://doi.org/10.1080/15481603.2022.2101727.

Liu, Y., B. Fan, L. Wang, J. Bai, S. Xiang, and C. Pan. 2018. "Semantic Labeling in Very High Resolution Images via a Self-Cascaded Convolutional Neural Network." *Isprs Journal of Photogrammetry & Remote Sensing* 145:78–95. https://doi.org/10.1016/j.isprsjprs.2017.12.007.

Marmanis, D., K. Schindler, J. D. Wegner, S. Galliani, M. Datcu, and U. Stilla. 2018. "Classification with an Edge: Improving Semantic Image Segmentation with Boundary Detection." *Isprs Journal of Photogrammetry & Remote Sensing* 135:158–172. https://doi.org/10.1016/j.isprsjprs.2017.11.009.

Mojaddadi Rizeei, H., B. Pradhan, and M. A. Saharkhiz. 2019. "Urban Object Extraction Using Dempster Shafer Feature-Based Image Analysis from Worldview-3 Satellite Imagery." *International Journal of Remote Sensing* 40 (3): 1092–1119. https://doi.org/10.1080/01431161.2018.1524173.

Raczko, E., M. Krówczyńska, and E. Wilk. 2022. "Asbestos Roofing Recognition by Use of Convolutional Neural Networks and High-Resolution Aerial Imagery. Testing Different Scenarios." *Building & Environment* 217. https://doi.org/10.1016/j.buildenv.2022.109092.

Riley, S. J., S. D. DeGloria, and R. Elliot. 1999. "Index That Quantifies Topographic Heterogeneity." *Intermountain Journal of Sciences* 5 (1–4): 23–27.

Seiffert, C., T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano. 2008. "*RUSBoost: Improving Classification Performance When Training Data is Skewed*." In Paper presented at the 19th International Conference on Pattern Recognition, Tampa, FL, USA.

Shepard, M. K., B. A. Campbell, M. H. Bulmer, T. G. Farr, L. R. Gaddis, and J. J. Plaut. 2001. "The Roughness of Natural Terrain: A Planetary and Remote Sensing Perspective." *Journal of Geophysical Research: Planets* 106 (E12): 32777–32795. https://doi.org/10.1029/2000JE001429.

Shi, X., H. Huang, C. Pu, Y. Yang, and J. Xue. 2022. "CSA-UNet: Channel-Spatial Attention-Based Encoder–Decoder Network for Rural Blue-Roofed Building Extraction from UAV Imagery." *IEEE Geoscience & Remote Sensing Letters* 19:1–5. https://doi.org/10.1109/LGRS.2022.3197319.

Singh, D., R. Maurya, A. S. Shukla, M. K. Sharma, and P. R. Gupta. 2012. "*Building Extraction from Very High Resolution Multispectral Images Using NDVI Based Segmentation and Morphological Operators*." In Paper presented at the Students Conference on Engineering and Systems, Allahabad, India.

Sritarapipat, T., and W. Takeuchi. 2017. "Building Classification in Yangon City, Myanmar Using Stereo GeoEye Images, Landsat Image and Night-Time Light Data." *Remote Sensing Applications: Society & Environment* 6:46–51. https://doi.org/10.1016/j.rsase.2017.04.001.

Sun, W., and R. Wang. 2018. "Fully Convolutional Networks for Semantic Segmentation of Very High Resolution Remotely Sensed Images Combined with DSM." *IEEE Geoscience & Remote Sensing Letters* 15 (3): 474–478. https://doi.org/10.1109/LGRS.2018.2795531.

Swan, B., M. Laverdiere, H. L. Yang, and A. Rose. 2022. "Iterative Self-Organizing SCEne-LEvel Sampling (ISOSCELES) for Large-Scale Building Extraction." *GIScience & Remote Sensing* 59 (1): 1–16. https://doi.org/10.1080/15481603.2021.2006433.

Thottolil, R., and U. Kumar. 2022. "*Automatic Building Footprint Extraction Using Random Forest Algorithm from High Resolution Google Earth Images: A Feature-Based Approach*." *Paper presented at the International Conference on Electronics, Computing and Communication Technologies*, Bangalore, India.

URL. 2022. "From Republic of Türkiye Ministry of National Defense General Directorate of Maps." https://www.harita.gov.tr/urun/sayisal-yuzey-modeli-5-m-seviye-0-sym5-l0-/1.

Virtriana, R., A. B. Harto, F. W. Atmaja, I. Meilano, K. N. Fauzan, T. S. Anggraini, and W. Suminar, F. C. Mustika, W. Suminar. 2023. "Machine Learning Remote Sensing Using the Random Forest Classifier to Detect the Building Damage Caused by the Anak Krakatau Volcano Tsunami." *Geomatics, Natural Hazards and Risk* 14 (1): 28–51. https://doi.org/10.1080/19475705.2022.2147455.

Voigt, S., T. Kemper, T. Riedlinger, R. Kiefl, K. Scholte, and H. Mehl. 2007. "Satellite Image Analysis for Disaster and Crisis-Management Support." *IEEE Transactions on Geoscience & Remote Sensing* 45 (6): 1520–1528. https://doi.org/10.1109/TGRS.2007.895830.

Wang, Q., L. Yan, Q. Yuan, and Z. Ma. 2017. "An Automatic Shadow Detection Method for VHR Remote Sensing Orthoimagery." *Remote Sensing* 9 (5): 469. https://doi.org/10.3390/rs9050469.

Wang, Y., S. Li, F. Teng, Y. Lin, M. Wang, and H. Cai. 2022. "Improved Mask R-CNN for Rural Building Roof Type Recognition from Uav High-Resolution Images: A Case Study in Hunan Province, China." *Remote Sensing* 14 (2): 265. https://doi.org/10.3390/rs14020265.

Weiss, A. 2001. "*Topographic Position and Landforms Analysis*." In Paper presented at the ESRI user conference, San Diego, CA.

Wierzbicki, D., O. Matuk, and E. Bielecka. 2021. "Polish Cadastre Modernization with Remotely Extracted Buildings from High-Resolution Aerial Orthoimagery and Airborne LiDAR." *Remote Sensing* 13 (4). https://doi.org/10.3390/rs13040611.

Wu, B., S. Wu, Y. Li, J. Wu, Y. Huang, Z. Chen, and B. Yu. 2022. "Automatic Building Rooftop Extraction Using a Digital Surface Model Derived from Aerial Stereo Images." *Journal of Spatial Science* 67 (1): 21–40. https://doi.org/10.1080/14498596.2020.1720836.

Wu, T., Y. Hu, L. Peng, and R. Chen. 2020. "Improved Anchor-Free Instance Segmentation for Building Extraction from High-Resolution Remote Sensing Images." *Remote Sensing* 12 (18). https://doi.org/10.3390/rs12182910.

Xiaoqin, W., W. Miaomiao, W. Shaoqiang, and W. Yundong. 2015. "Extraction of Vegetation Information from Visible Unmanned Aerial Vehicle Images." *Transactions of the Chinese Society of Agricultural Engineering* 31 (5). https://doi.org/10.3969/j.issn.1002-6819.2015.05.022.

Xu, Y., L. Wu, Z. Xie, and Z. Chen. 2018. "Building Extraction in Very High Resolution Remote Sensing Imagery Using Deep Learning and Guided Filters." *Remote Sensing* 10 (1): 144. https://doi.org/10.3390/rs10010144.

Zabalza, J., J. Ren, J. Zheng, H. Zhao, C. Qing, Z. Yang, and S. Marshall, S. Marshall. 2016. "Novel Segmented Stacked Autoencoder for Effective Dimensionality Reduction and Feature Extraction in Hyperspectral Imaging." *Neurocomputing* 185:1–10. https://doi.org/10.1016/j.neucom.2015.11.044.

Zhang, H., Y. Liao, H. Yang, G. Yang, and L. Zhang. 2020. "A Local-Global Dual-Stream Network for Building Extraction from Very-High-Resolution Remote Sensing Images." In *IEEE Transactions on Neural Networks and Learning Systems*. https://doi.org/10.1109/TNNLS.2020.3041646

Zhong, T., Z. Zhang, M. Chen, K. Zhang, Z. Zhou, R. Zhu, and J. Yan, G. Lü, J. Yan. 2021. "A City-Scale Estimation of Rooftop Solar Photovoltaic Potential Based on Deep Learning." *Applied Energy* 298:117132. https://doi.org/10.1016/j.apenergy.2021.117132.