Scientific African 26 (2024) e02431

Contents lists available at ScienceDirect

Scientific African

journal homepage: www.elsevier.com/locate/sciaf

Precise LULC classification of rural area combining elevational and reflectance characteristics using UAV

Ke Zhang [a,*], Lameck Fiwa [b], Madoka Kurata [c], Hiromu Okazawa [d], Kenford A.B. Luweya [e], Mohammad Shamim Hasan Mandal [f], Toru Sakai [g]

[a] Rural Development Division, Japan International Research Center for Agricultural Sciences, 1-1-1 Owashi, Tsukuba, Japan
[b] Faculty of Agriculture, Lilongwe University of Agriculture and Natural Resources, Lilongwe P.O. Box 219, Malawi
[c] Japan International Cooperation Agency, 2 chome, 5-25 Chiyoda, Tokyo, Japan
[d] Faculty of Regional Environment Science, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya, Tokyo, Japan
[e] Graduate School of Agro-Environmental Science, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya Tokyo, Japan
[f] Forestry Division, Japan International Research Center for Agricultural Sciences, 1-1-1 Owashi, Tsukuba, Japan
[g] Social Sciences Division, Japan International Research Center for Agricultural Sciences, 1-1-1 Owashi, Tsukuba, Japan

ARTICLE INFO

Editor: DR B Gyampoh

Keywords:
Remote sensing
LULC
UAV
Image classification
Machine learning
Thatched house
Rural development

ABSTRACT

With the development of unmanned aerial vehicle (UAV) in the recent decade, very high-resolution aerial imagery has been used for precise land use/land cover classification (LULC). However, special structures in rural areas of developing countries such as traditional thatched houses have posed challenges for precise LULC classification due to their undistinctive appearance and confusable characteristics in both reflectance and structure. LULC mapping is essential particularly in rural areas which have high data scarcity and vulnerability to natural disasters. With high-resolution observation has been achieved by UAVs, it is important to propose high-precision LULC classification methods which can fully use the advantages of UAVs. To emphasize the differences among the common LULC types in rural areas, this study proposed an original index, the rural residence classification index (RCI). RCI was calculated as the product of the above ground height and the square of the difference between the NDVI value and one. Then, a comprehensive classification method was established by combining the RCI, the traditional threshold method and a machine learning method. As a result of the comparison with the traditional threshold method, object-based image analysis, and random forest methods, the method by this study achieved the highest overall accuracy (overall accuracy = 0.903, kappa = 0.875) and classification accuracy for detecting thatched houses (user's accuracy = 0.802, producer's accuracy = 0.920). These findings showed the possibility on identifying the confusable structures in rural areas using remote sensing data, which was found difficult by the previous studies so far. The method by this study can promote the further utility of UAVs in LULC classification in rural areas in developing countries, thereby providing precise and reliable material for hydrological, hydraulic or ecosystem modelling, which eventually contributes to more accurate natural hazard risk assessment, rural development, and natural resource management.

* Corresponding author.
E-mail address: ergouqidao@affrc.go.jp (K. Zhang).

https://doi.org/10.1016/j.sciaf.2024.e02431
Received 4 March 2024; Received in revised form 14 October 2024; Accepted 15 October 2024
Available online 18 October 2024

Introduction

Land use/land cover (LULC) classification is essential not only for natural resource management and urban planning, but also for natural hazard risk assessment, disaster response after disasters (such as floods and earthquakes), and post-disaster reconstruction [1]. As a landlocked country in sub-Saharan Africa, Malawi lies in the eastern rift of the Great Rift Valley and is bordered by Lake Malawi to the east of the country. Geolocation has caused natural hazards in Malawi, such as cyclones, floods, and earthquakes [2]. In recent years, the disaster vulnerability of Malawi has increased because of climate change, population growth, and agricultural expansion, particularly for rural households with semi-permanent buildings [3], such as grass thatched roof houses, the details of which will be introduced later within the section. For instance, the flood caused by the tropical cyclone Freddy devastated southern Malawi in March 2023 has killed 679 people, injuring 2178, displacing 143,487, and affecting 2514,913 [4]. Natural hazard impacts can be reduced by applying suitable risk assessments and hazard mapping [5]. It has been suggested that for flood risk hazard mapping with hydrological models, where LULC data are important basic parameters, LULC data with higher resolution and accuracy could provide better modelling performance [6]. Highly accurate and precise LULC mapping is not only useful for hazard mapping, but also important to improve the understanding of both societal and environmental exposure to natural hazards, such as cyclones, floods, and earthquakes, by providing precise information of household, residence, infrastructure, and natural resource.

As the major material for LULC mapping, artificial satellites launched by the National Aeronautics and Space Administration (NASA), the European Space Agency (ESA), Maxar and other national or private aerospace development organisations have become an essential platform for modern remote sensing based on advances in sensor technologies, aerospace platforms, and computer sciences since the 1960s. Nowadays, satellite imagery has been used for LULC mapping at national or regional level all over the world, with well-developed classification technology and algorithms. On the other hand, satellite observations still have limitations in terms of both spatial and temporal resolutions, which cause difficulty when precise LULC mapping at local level is necessary. For example, the satellites that are utilised worldwide, such as Landsat, Terra, and SPOT, have a ground resolution of several tens of meters, and satellite data cannot be obtained under certain conditions, such as heavy cloud cover [7]. Particularly in rural areas of Malawi, where the area of single housing is often smaller than 100 m^2 (10 $\times$ 10 m), it is difficult to identify a single household using open-access satellite imagery. Therefore, the development of new remote sensing platforms and classification methods that are suitable for high-precise LULC identification has become necessary. Since the 2010s, unmanned aerial vehicles (UAVs) have made rapid and significant progress in performance and have been used worldwide for disaster relief, civil engineering surveys, pesticide spraying, crop monitoring, and infrastructure inspection [8]. Compared with satellite data, UAVs have advantages in terms of spatial and temporal resolution, mobility, and manoeuvrability. At present, with the development of the Global Positioning System (GPS), Geographic Information System (GIS), and optimal sensor technologies, and improvement in aviation law systems, UAVs are already being seen as a stable and reliable platform for LULC mapping.

Several classification techniques have been developed to map LULC patterns and dynamics from the remotely sensed imagery, and in recent years, the application of machine learning algorithms has become one of the most popular approaches [9]. Among the supervised machine learning classifiers that are commonly used for LULC mapping, such as maximum likelihood, random forest (RF), support vector machine (SVM), spectral angle mapper, fuzzy adaptive resonance theory-supervised predictive mapping, Mahala Nobis distance, radial basis function, decision tree, multilayer perception, and naïve Bayes, it has been found that RF and SVM have the best performance for LULC classification [10]. It has been pointed out that RF modelling is one of the most precise classification methods because it can model the complexity of input variables, handle outliers, treat noise efficiently, and avoid overfitting [11]. RF is a non-parametric ensemble algorithm based on decision trees that produces multiple estimators during the modelling process and combines them to generate the final prediction. In a supervised machine learning classifier, the dataset is separated into training and testing data, whereas in RF, the training data are further separated into bootstrap samples, which are used for each decision tree, and out-of-bag samples, which are used to evaluate the performance of the constructed model. Two parameters, the number of the trees and the number of features of each split, must be defined when setting up an RF model. According to the developer of RF, although a large number of trees tends to provide better accuracy, more trees than required do not improve the classification result [12]. Despite the high potential of machine learning and also deep learning on LULC classification or ground object identification, the accuracy of machine learning and deep learning has been proven to be restricted in the studies so far due to a variety of factors, including the selection of an efficient wavelength, spatial resolution, and the selection and tuning of hyperparameters [13], and studies worldwide have been dedicated to the improvement of classification accuracy or efficiency of machine learning or deep learning approaches by different kinds of approaches. For example, Hap et al. (2021) developed a new automated weed detecting system using Convolution Neural Network (CNN) classification and snapshots to choose the optimal parameters for the model, and eventually achieved the highest overall accuracy (0.99) among the studies so far at the similar field [13].

Traditionally, machine learning classifiers are used at the pixel level, which means assigning the given classes to all pixels of the study area based entirely on the spectral feature information of each single pixel. This approach can yield fine performance when the aerial data have medium- or low-ground resolution because one pixel typically represents an area of > 100 m^2 [14]. However, with the rapid development of UAVs, aerial data with a high or even ultrahigh resolution of < 10 cm can be easily obtained. In high-resolution imagery, the variance in multiple pixels representing one ground object makes it difficult to separate a vast number of pixels in one image [15]. Differing from traditional pixel-based classification methods, the object-based image analysis (OBIA) method first separates the image into segments which are small polygons constructed of several neighbouring and similar-valued pixels, then, with appropriate training samples, the classification is performed by dividing the segments into different classes according to their shape, size, and spectral content [16]. The classification based on these segmented objects instead of single pixels has been proven to have better performance at defining the precise shape of the ground object and separating the pixels located at the border between two

ground objects [17].

Although pixel- or object-based classification using supervised machine learning is the most efficient, convenient, and commonly used method in remote sensing to create LULC maps, the technique has one limitation factor which is that the classification accuracy is determined by user's-trained samples that define the statistics of the classes of interest [18]. In other words, human error has a direct and significant effect on classification accuracy. When the ground object category is simple inside the study site, instead of the automatic identification methodology, some studies use the threshold of certain waveband information, calculated index values, or other pixel values, such as elevation or the true height value, to extract ground objects with special attributes. One representative example of this approach is the normalized difference vegetation index (NDVI) threshold method used to extract vegetation areas or even identify vegetation species [19].

Despite the various classification methods introduced above, some certain ground object types remain difficult to identify using remote sensing data. For example, Wouters et al. found it difficult to identify thatched houses when constructing a flood damage assessment tool in southern Malawi. Although the study tried to improve the classification accuracy by combining OBIA and SVM, and further utilizing the digital surface model (DSM) data, the classification accuracy remained moderate (F1 score of the OBIA-SVM method: 0.53; OBIA-SVM-DSM method: 0.72) [20]. Thatch is a material made from tall grass mainly used for roofing structures and has similar reflection characteristics to soil and dead wood. This characteristic results in a very close RGB appearance of thatched houses with the soil and a very close multispectral characteristics of thatched houses with the relatively low trees inside the settlement with lower stem density. The potential biases in identifying thatched houses are that when using RGB value as the main parameter to automatically detect thatched houses, it is difficult to separate them from the bare ground with the same colour. When using height or vegetation indices such as NDVI as the main parameters, thatched houses cannot be separated from the lower trees. Furthermore, this

Fig. 1. Details of the study site: (a) Location of the study site; (b) Ground object categories inside the study site.

problem cannot be solved by using all the RGB value, height information and NDVI as parameters respectively since the similarity among the three kinds of ground type would cause confusion during the procedure of machine learning and eventually a result with mixed classes. This is the reason why thatched houses are difficult to separate using remote sensing imagery, and this difficulty has been proven by Chan et al., (2022) when they tried to identify the thatched houses in a refugee camp in east Africa using very high-resolution (VHR) UAV imagery with the U-Net architecture. As the result of the study, the thatched houses inside the refugee camp were completely missing from the residence class and all classified as bare ground [21]. Similarly, when Chen et al., (2021) used VHR imagery with OBIA approach to classify the LULC in rural Mozambique, the user's accuracy of the thatched houses (0.58) was extremely lower than the class types (metal roof: 0.94, Corral: 0.92, Other: 0.98) [22]. Thatched houses are traditional housing found in rural areas throughout Africa, Asia, and Europe. There are many opinions suggesting that the thatched houses are "semi-permanent" and can increase the risk of disease such as Malaria, and thus should be replaced with iron-cedar materials [23]. Conversely, there are also studies on architecture showing that, with the right housing technique on roof angle and packing method, and with proper maintenance, structures in Africa that have thatched roofs last for a century or longer [24]. Further opinion suggests that the people in villages in the rural areas of Malawi still feel distrustful of the western way structures which lack an understanding of the basic African traditions and customs [25]. Regardless of the different opinions on thatched houses, they are home to numerous Malawian families at present and should not be ignored or treated as irrelevant during precise LULC mapping in Malawi and all other countries and areas worldwide with similar cultures. However, no study has focused on LULC considering the characteristics of thatched houses.

Based on this background and facing the same difficulties on rural area LULC classification like the previous studies, the objectives of this study were to (1) revise the specific characteristics of thatched houses in the remote sensed imagery and find the key difference between thatched houses and other ground objects in rural area; (2) develop an original index for identifying thatched houses by combining both optical and structural features based on the key differences; and (3) establish a comprehensive methodology for high-accuracy LULC classification using UAV imagery, compare the accuracy of the results produced by the three most commonly used classification methods (RF, OBIA, and a ensembled threshold method), to finally suggest the optimal method for precise LULC mapping in rural area. The methodology of this study was specifically identifying the complex differences between thatched houses and other ground objects, and proposing an original detection index by combining the spectral and height information instead of using them as individual parameters which causes mis classification, enhancing the differences between ground objects. Based on the introduction of LULC classification studies so far in the previous paragraphs in this section, since RF has been proven to be one of the most proper algorithms for LULC classification [11], RF has been chosen as the pixel-based machine learning method as the comparison with the method proposed by this study. On the other hand, since the OBIA has been considered as a more suitable classification approach especially for UAV imagery classification, and is becoming a popular way at LULC classification, it has been chosen as the object-based method for the comparison with the proposed method. Furthermore, the conventional ensemble threshold method has the advantages of efficiency and versatility, it has been chosen as the conventional method for the comparison with the proposed method. Finally, for the validation procedure was conducted on pixel-level for all methods, which was exactitude to less than three centimetres. This study has used >2000 accuracy assessment points which was a very high density for the size of the study area, to evaluate the user's accuracy, producer's accuracy, overall accuracy and kappa coefficient. This study also analysed the confusion matrix created by the accuracy assessment points to unravel the detailed advantages and also disadvantages of the proposed method, in order to better contribute to the practical use of it in rural areas of developing countries.

Materials and methods

Study site

The location of the study site is shown in Fig. 1. Chembe Village is a small settlement located on the southern shore of Lake Malawi. The village is part of Lake Malawi National Park, Mangochi District, Malawi, where the climate is classified as tropical savanna. The yearly average minimum and maximum temperatures are 29.9 °C and 33.6 °C respectively, coupled with a mean annual rainfall of 1690 mm.

Adjacent to Lake Malawi, the village maintains a traditional way of life based on fishing and cultivating crops such as corn, rice, and vegetables. Houses with iron or grass thatched roofs are densely distributed in the village, with trees and grass interspersed among them. The roads inside the village are not paved, and the soil surface is classified as Cambic Arenos by the FAO Soil Group. The study site is located in the centre of the village between 34.836E to 34.843E and 14.023S to 14.030S, with an area of 240,000 m^2 (300 × 800 m), which is approximately one-third of the total area of the village.

UAV and ground surveys

The ground survey and UAV photography were conducted on 11 March 2023 under sunny and no-wind conditions. During the aerial surveys with UAVs, the use of ground control points (GCPs) is one of the most commonly used methods to calibrate the geolocation information of UAV imagery. Proven by many previous studies, the calibration accuracy is affected by both the number and the spatial distribution of the GCPs [26]. Awasthi et al. (2020) have compared the geolocation accuracy of UAV products calibrated by different GCP distribution patterns and have found that GCPs evenly distributed inside the target area with a minimum number of five, which forms the shape of a die could provide proper calibration result [27]. Based on these instructions, before conducting the surveys of this study, GCPs were evenly distributed at the study site with four at the corners and one at the centre (Fig. 1**a**). Zhang et al. (2022) have tested 88 kinds of GCP distribution patterns in an experimental field and established a method to previously evaluate the GCP

calibration accuracy using the number and surround area percentage by GCPs [28]. Based on the accuracy prediction method suggested by the previous study, the five GCPs surrounding an area of 112,116 m^2 (target area size: 212,786 m^2) would at least provide calibration with an error of no >15 cm. The calibration accuracy could be improved with a greater number of GCPs, however, considering that the time of the survey of this study was limited, and the geolocation accuracy of an error <15 cm was acceptable due to the objective of this study, the authors have decided this distribution strategy could accomplish accurate coordinate calibration for this study. Anti-aircraft signals (size: 40 × 40 cm) were settled on the GCPs to make them recognisable in UAV imagery. All the GCPs had a clear view of the sky and maintained a distance of at least 5 m from trees or houses to ensure the accuracy of the coordinate measurement. The coordinates of the GCPs were determined using two sets of real-time kinematic global navigation satellite system (RTK-GNSS) receivers, Drogger DG-PRO1RWS (BizStation Corp., Matsumoto, Japan), in RTK mode. During the ground survey, definition of the categories of the ground objects has been conducted (Fig. 2**b**).

UAV photography was conducted after the GCPs were fixed and measured. The UAV used in this study was a Phantom 4 Multispectral (DJI Inc., Shenzhen, China). The camera incorporates one RGB lens and five multispectral lenses (red, green, blue, red-edge, and near-infrared bands). The resolution for all lenses was 1600 × 1300 pixels. During photography, the camera angle was set vertically downward using the auto flight application DJI GS Pro (DJI Inc.). The flying altitude was 50 m above the ground, and the overlap was 75 % for the top and 75 % for the side. The total flight distance and time were 16 km and 160 min, respectively, and 10 batteries were used. The flights covered a total area of 24 ha, and 16,680 images (2780 for each lens) were collected. The flight height was decided considering the balance between flight time and the ground resolution of the UAV imagery. 50 m flight height with DJI P4 provides a ground resolution of 2.6 cm, which is considered precise enough for the detection of all kinds of ground objects inside the study site, and the flight time by this flight height was acceptable. The top and side overlap rates were decided based on the recommendation from the commonly used UAV image processing software, which is at least 75 % frontal and at least 60 % side overlap. The top overlap of this study was the same as the recommendation, and the side overlap rate was relatively higher than the recommendation considering the possible mis capturing, image blur and overexposure.

Image processing

The Structure from Motion (SfM) process was conducted using Metashape Professional ver. 2.0.1 (Agisoft LLC., Saint Petersburg, Russia) to reconstruct the 3D structure of the study site. After inputting the UAV images into the SfM software, a tie-point cloud was generated by identifying the significant and stable points in the overlapping areas among the aerial images. During this procedure, Metashape uses a key matching algorithm to relate the imagery and a greedy algorithm to find approximate camera locations and refines them using a bundle-adjustment algorithm. The coordinates of the GCPs obtained by the RTK-GNSS were then adapted to the tie-point cloud to correct for accurate absolute geolocation by a georeferencing algorithm. Based on the geographically corrected tie-point cloud, a dense point cloud using a patch-based multi-view stereo algorithm for depth-map computation, 3D polygonal mesh

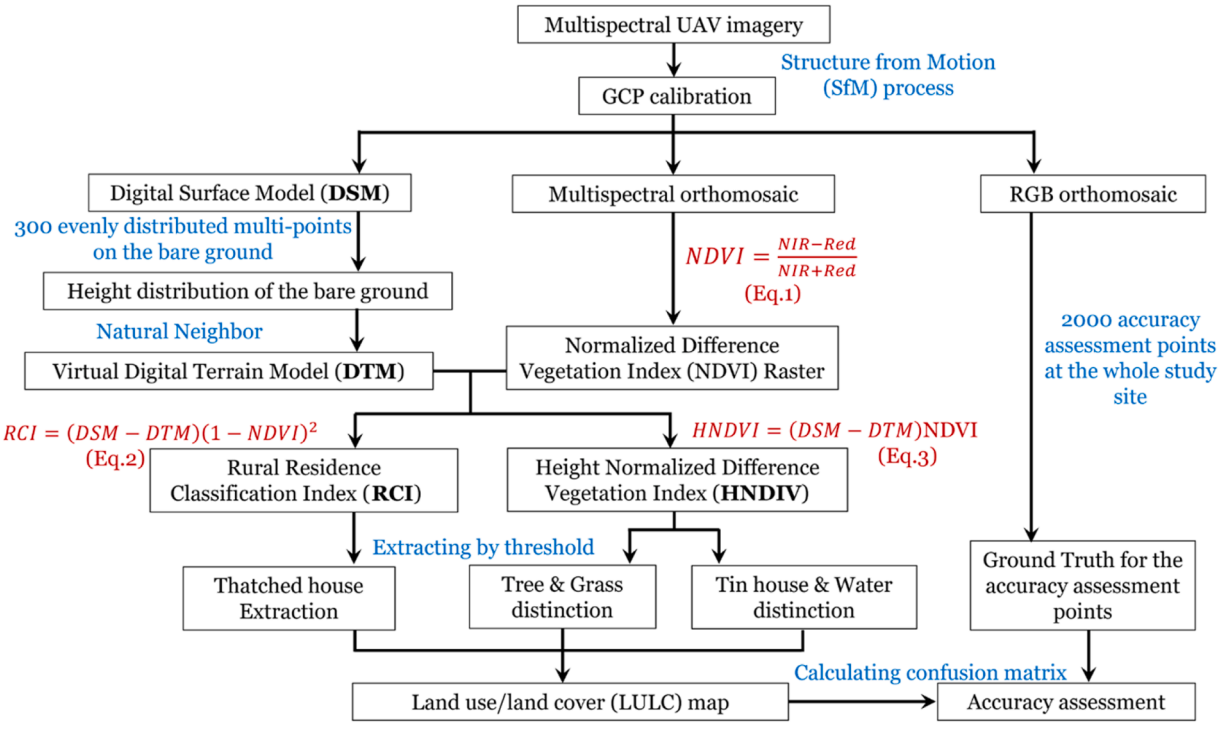


Fig. 2. Workflow of the LULC classification modified using the RCI.

using the surface reconstruction algorithm based on Delauney triangulation or Position surface reconstruction, and texture model using the surface complexity and atlas filling coefficient ('Texture mapping') algorithm were generated sequentially. Finally, the RGB orthomosaic, multispectral orthomosaic per band, and DSM were constructed. The ground resolutions of the final products were 2.6 cm for the orthomosaic and 5.2 cm for the DSM. A raster of the NDVI was then calculated using the raster calculator tool within ArcGIS Pro ver. 3.1.0 (Esri Inc., Redlands, CA, U.S.) using Eq. (1). Finally, one raster combined with the RGB, NDVI, and DSM values of the study site was generated for the use in the classifiers.

$$NDVI = (NIR - Red) \ / \ (NIR + Red) \tag{1}$$

LULC classification methods

Ensembled threshold method

The threshold method is possible for LULC classification when a study site has simple ground object categories with significant reflectance characteristics. The study site of this study is a small village with six main LULC types, which are considered to be able to be classified using the proper parameters. An ensembled threshold method using two parameters, NDVI and DSM, were used to extract all six types of ground objects in this study.

Firstly, as shown in Eq. (1), the NDVI which represents the normalized value of the difference between the reflectance of the NIR and red bands of light [29] was calculated within ArcGIS Pro. During the photosynthesis of healthy green vegetation canopies, chlorophyll strongly absorbs red light while reflecting almost all of the NIR light, causing an extreme difference between the reflectance of NIR and red light, and eventually a high NDVI value. Therefore, the NDVI is often considered the optimal threshold for extracting vegetation coverage. NDVI can not only be the standard to identify vegetation, but also some objects that have the opposite reflectance to vegetation. For example, water reflect approximately 15 %~5 % of the visible light, while absorbs almost all of the infrared light. This characteristic caused a very low NDVI of water. Moreover, tin materials reflect >80 % of the visible light but reflect only about 50 % of the near-infrared light. Because of this characteristic, houses with tin roofs also have negative NDVI value. In this study, the NDVI threshold was used to separate the classes with extremely high NDVI value (NDVI $\geqq$ 0.3) which were the tree and grass class, and also the classes with extremely low NDVI (NDVI < -0.2), which were the water and tin house class. The classes left after removing the high NDVI and low NDVI areas were soil whose NDVI value was near 0, and thatched houses, whose NDVI value varies between -0.2 to 0.4 depending on the status of the thatch and the surrounding objects. After the NDVI thresholding step, the six ground objects were separated into three groups which would be further thresholding using the second parameter.

In addition to the optical characteristics, the height value is another commonly used threshold to separate ground objects, particularly when they have similar optical features [30]. For example, UAV photography took place during the rainy season in Malawi, when grass was flourishing and had similar visual colour and NDVI values as the tree canopies. In this case, the height threshold makes it easy to separate trees and grass. Moreover, water and tin houses that both have negative NDVI could be separated using height information. However, the DSM created from UAV images by SfM process cannot be used directly as a threshold parameter. The DSM value represents the elevation of a ground object, which is the distance between the top of the object and sea level. Because of the effect of undulating terrain, certain ground objects with a low absolute height (the distance between the top of the object and the ground surface) located at a highland may have a higher elevation value than the objects with high absolute height at a lowland. To remove the effect of upward and downward terrain, this study created a virtual Digital Terrain Model (DTM) which represents only the elevation of the ground surface. To do this, three hundred multiple points equally distributed in the bare ground area (such as roads and yards) inside the study site were created, of which the elevation value was extracted from the DSM raster. Subsequently, a virtual surface raster was constructed based on multiple points using the natural neighbour geo-process in ArcGIS Pro ver. 3.1.0. The DTM fabricated in this step was a smooth surface that only represented the relief of the bare ground. Finally, the difference in the height between the DSM and the DTM, which is the absolute height of the ground objects, was calculated and used as the height threshold. Among the three groups made by the NDVI thresholding, each group had one class with high height value (tin houses, thatched houses and trees) and one class with low height value (water, soil and grasses). Thus, the three groups were further separated by a height threshold of 2.0 m. To this point, all the six LULC classes inside the study site were identified using two threshold parameter and combined into one LULC raster for the final accuracy assessment.

Comprehensive classification modified by the rural residence classification index (RCI)

Altthough the ensembled threshold method could manage LULC in small settlements with simple ground objects, the method could face difficulty when separating thatched houses from other objects. Therefore, in African villages, where there are a vast number of thatched houses, this method may not provide satisfactory classification accuracy. Thatch, known as deadwood in the ecosystem service field, presents challenges in identification because of the lack of significant spectral characteristics. Because of this, traditional remote sensing approaches have shown limitations in classifying deadwood using spectral and spatial information, such as the NDVI and height values [31]. Even with machine learning approaches, deadwood and bare ground misclassifications have been addressed using RF and SVM models [32,33]. Because of the instability of the reflectance characteristic of the thatch made by tall grasses and tree branches, NDVI value of thatched housed can cover a wide range of -0.2 to 0.5, which includes many other ground objects such as low-density trees and dry soil. This means that when separating vegetation using NDVI threshold, if the threshold is set as 0.3, some of the thatched houses will be extracted with the vegetation. Conversely, if the NDVI threshold is set lower, a portion of the trees and grasses will be classified as thatched houses, because there are multiple small, unhealthy or low-density parts of trees with an NDVI value of 0.2 ~ 0.3. This difficulty also remains even when using he second parameter which is the height because the specific parts of

trees often have close height value to thatched house. To overcome this limitation for identifying thatches or deadwoods, this study observed the NDVI and height of the objects inside the village closely and listed the common feature and difference of them, which were:

(1) Tin houses have low NDVI (< 0.2) and high altitude (> 3.0 m).
(2) Thatched houses have moderate NDVI ($-0.2 \sim 0.4$) and high altitude (> 3.0 m), which are similar with specific parts of trees. However, when NDVI value are similar, thatched houses have higher altitude than trees. This is because the specific parts of trees with lower NDVI are at the edge of the canopies, and have lower density, sometimes include exposed ground surface among them. According to previous studies, during the construction of DSM, which is the original material of the altitude data, a smooth filter algorithm was adapted into the dense point cloud extracted from the UAV images, in order to remove the noise points. Because of the smooth filter, the height of the ground objects with low density, small size and sharp shapes would be underestimated.
(3) The main parts of trees have very high NDVI ($0.6 \sim 1.0$) and very high altitude (> 4.0 m). On the other hand, the edge of the canopy has similar NDVI ($0.3 \sim 0.5$) with thatched houses, but lower altitude (< 3.0 m) than thatched houses.
(4) Grasses have high NDVI (> 0.4) and very low altitude ($0.0 \sim 1.0$ m).
(5) Water has very low NDVI (< -0.2) and very low altitude ($-0.5 \sim 0.0$ m).
(6) Soil has moderate NDVI ($-0.2 \sim 0.3$) depending on the moisture, and low altitude ($\fallingdotseq 0.0$ m).

The list above showed that although some objects could be easily classified using the ensembled threshold method, the common features of thatched houses and trees on NDVI and height cannot be separated simply based on the two parameters. However, by combining the NDVI and height by certain calculation, the difference between them could become outstanding. This named a novel index called which was the Rural Residence Classification Index (RCI), calculated using Eq. (2).

$$RCI = (DSM - DTM) \times (1 - NDVI)^2 \tag{2}$$

The index comprises two parts: the spatial part (DSM – DTM), which is the height value, and the spectral part $(1 - \text{NDVI})^2$. The tin houses had a negative NDVI value and high altitude, resulting the highest RCI among all the ground objects. The main parts of trees with deep green colour and dense canopies had extremely high NDVI value close to 1.0, making the $(1 - \text{NDVI})^2$ part close to 0.0, and resulted the smallest RCI among the ground objects despite their high altitude. For the low trees with a low density of leaves and low altitudes, the NDVI value was relatively high ($0.3 \sim 0.5$), resulting a low value of $(1 - \text{NDVI})^2$; while the height value was relatively low ($1.5 \sim 3.0$), which eventually resulted in a second low RCI value among the four types of ground objects. Finally, for the thatched houses, the height was higher than that of the small trees, whereas the NDVI was lower than that of the small trees, making their RCI larger than that of the small trees but lower than that of the tin houses. As above using this index, the two types of information, height and NDVI were combined into one figure, instead of used sequentially and independently. This calculation strengthened the characteristics of each ground object and solved the problem of the conflict when deciding the optimal thresholds sequentially. Therefore, the RCI was used to separate thatched house and trees in this study.

Additionally, another index, the HNDVI suggested by Wang, N., et al. (2022) [34] combining NDVI and height by a simpler calculation was used to divide trees, grass, tin houses and water. To this point, a comprehensive classification method was developed by combining the RCI and HNDVI thresholding. The workflow of the final method is illustrated in Fig. 2.

$$HNDVI = (DSM - DTM) \times NDVI \tag{3}$$

Computational efficiency is a considerable factor for the comprehensive evaluation of a novel method [35]. The processing speed of this method was measurer on a hosting devise with an 12th Gen Intel Core i7–12700F processor and an NVIDIA GeeForce RTX 3070 graphics processing unit. After constructing the DSM and multispectral orthomosaic within the SfM software and input the aerial products into ArcGIS Pro, it took approximately 15 min to create the 300 multipoints for bare ground elevation extraction, 8 s to extract the ground elevation using the 300 points, and 3 s to create the virtual digital terrain model using the ground elevation. Using the multispectral orthomosaic, it took 92 s to calculate the NDVI value of the study site. Using the DSM, DTM, and NDVI indices, it took 55 and 48 s to calculate the RCI and HNDVI index of the study site. Finally, it took averagely 18 s to extract thatched houses, trees, grass, tin house, water area and soil using the threshold value of each index. Another 77 s were needed to reconstruct the LULC map using the extracted classes. A total processing time of approximately 22 min was needed for this method including the manually adjustment in order to construct the DTM which was necessary for all methods mentioned in this study. The time consumed decides the DTM construction was 7 min, which was more efficient than supervised machine learning approaches considering the time used for manually creating the training samples.

RF classification

As one of the comparative method, the RF classification of this study was conducted using Python 3. The python library 'nivapy3' and 'numpy' were used to access and appearance the multispectral orthomosaic constructed in 2.3. and rasterising the vector training data created within ArcGIS Pro. The training samples were created within ArcGIS Pro as polygons, which were then transformed into coordinate data by python to extract the pixel value from the combined raster. During the RF modelling, among averagely 60 polygon samples containing approximately 1500,000 pixels per class, 80 % of the samples were used as the training data, while 20 % were used for testing. When creating the polygon samples, since every pixel inside the polygons would be used as one sample during the machine

learning procedure later, every polygon contained completely only the pixels from one particular class and no pixel of other classes was contained in the polygon. The pixels on the edge between two kinds of ground objects were not contained in the polygon samples since it would lead to individual bias caused by subjective judgement. The 'sklearn.model (train_test_split)' library was used to separate the dataset. There were two reasons for using polygon samples instead of point samples for RF modelling, which were (1) since the very high ground resolution caused a large size of the input raster (31,180 × 25,680 pixel), sampling enough training data by points became very time and labour consuming; (2) the classification results by RF would be compared with OBIA afterwards, and training samples for OBIA method would be based on polygons. After setting the dataframe, this study used the python library 'utils' and 'RandomForest Classifier (sklearn.ensemble)' to train a basic model based on the training samples and implement classification prediction. According to previous studies, in LULC classification, a typical forest should contain 200–500 classification trees [36,37]. Therefore, for the RF classification in this study, the maximum number of trees was set to 500, whereas the maximum tree depth was 30. Finally, based on the trained classifier, a classification raster containing six classification classes (which were then named as Tin house, Thatched house, Tree, Grass, Water area and Soil) were exported as a geotiff data by 'nivapy (array_to_gtiff)' for further accuracy assessment. The same classes were used for all four classifiers verified in this study.

Object based image analysis

The OBIA classification in this study was conducted using eCognition Developer ver. 10.3 (Trimble Inc., Westminster, CO, U.S.), which is also based on Python 3 language. After creating a new project and inputting the multispectral orthomosaic constructed in 2.3., the first step of the OBIA was segmentation which separate the combined raster into small polygons containing several adjacent pixels depending on their spectral, geometrical, and spatial properties. The algorithm used for the segmentation was the 'multiresolution segmentation' which creates the objects using an iterative algorithm where the pixels are gathered within a certain threshold which define the upper object variance. This threshold is named as the scale factor in eCognition user interface. The scale factor which determines the average size of the segments was determined based on trial and error in this study, and finally set to be 100 considering the average number of pixels for each ground object in the UAV imagery. Consequently, a total of 730,933,869 pixels of the combined raster at the study site were separated into 43,010 objects. After the segmentation, a class hierarchy containing the six classes (as described in the RF method) was created, and then 60 segments were selected as the training samples for each class. The sample number was fairly equal to RF classification descripted at 2.4.1. After the sampling, the 43,010 objects were classified using the hierarchical classification algorithm, which evaluates the membership value of the basic image among the active classes based on the mean pixel values, mean brightness, standard deviation, and position and shape. The classification result was then exported as a raster for further accuracy assessment.

Accuracy assessment method

After conducting five types of classification, the accuracy of each method was evaluated by calculating the confusion matrix. More than two thousand assessment points equally distributed across the study site were created within ArcGIS Pro using the 'Creating Accuracy Assessment Points' geoprocessing tool based on the 'Stratified Random' strategy, which means instead of random creating evenly distributed points like a fishnet inside the interested area, the study area is first split into strata and random samples are generated within each stratum, making sure that there are same number of assessment points for all the classes despite the area percentage of them. Originally, 2500 assessment points were generated, and then every assessment point was checked manually to confirm if they are located in the ambiguous or confusing edge between different classes. Those points of which the class cannot be objectively judged were deleted to avoid biased accuracy assessment. After this procedure, eventually 2151 assessment points were used for further analysis. The attribute table of the assessment points included the ground truth and classification result fields. The ground truth was determined by visual judgement since the very high ground resolution (2.6 cm) allowed a clear judgement of the class of the assessment points located in the main part of the ground objects. The classification result was extracted from the classification raster by each classification method. An accuracy assessment was then conducted using the 'Compute Confusion Matrix' geoprocessing tool by calculating the user's accuracy, producer's accuracy, overall accuracy and kappa coefficient for each method. The equations of the four indicators are shown in Eqs. (4–7).

$$User's\ Accuracy = x_i / x_r \tag{4}$$

$$Producer's\ Accuracy = x_i / x_c \tag{5}$$

$$Overall\ Accuracy = \sum_{i=1}^{m} x_i \Big/ N \tag{6}$$

$$\text{Kappa} = \left[N \sum_{i=1}^{m} x_i - \sum_{i=1}^{m} (x_c \times x_r) \right] \Big/ \left[N^2 - \sum_{i=1}^{m} (x_c \times x_r) \right] \tag{7}$$

where x_r is the total point number due to the reference ground truth, x_c is the total point number due to the classification result, x_i is the total number of the accurate classification (calculated by the diagnosis of the confusion matrix), N is the total number of the accuracy assessment points, and m is the number of classes [38].

Results

Optical and structural characteristics of the study site

Fig. 3 shows the RGB orthomosaic, height map, and NDVI map of the study site, where an enlarged view of each map is shown on the right side of the figure. As shown in Fig. 3**(a)**, the difference between the reflectance of visible light of the thatched house and that of the soil was not significant. Although the thatched houses could be identified by visual judgement based on the shadows around them, the RGB values on pixel level in the two classes were extremely close. Furthermore, in the NIR band of light, which is shown as the NDVI value in Fig. 3**(b)**, there was no difference between the thatched houses and the soil, making the thatched houses completely invisible on the NDVI map. The average R, G, B, and NDVI values of all the pixels were 120 ± 36, 106 ± 33, 92 ± 30, and 0.15 ± 0.1, respectively, for the thatched house class; and 114 ± 41, 96 ± 35, 77 ± 31, and 0.15 ± 0.12, respectively, for the soil class. The results of the *t*-test ($p = 0.78$) show no significant difference between the optimal pixel values of the two classes, indicating that the classification based on optimal characteristics could not produce satisfactory performance when identifying thatched houses or soil, which is the reason for the necessity of utilizing hight information as another parameter.

Fig. 3**(c)** shows the heights of the ground objects from the ground. As showed by the colour of ground surface, the effect of the terrain relief was excluded by calculating the differential values between the DSM and DTM. Because of this process of creating the virtual DTM, the thatched houses, tin houses, and trees were displayed clearly on the height map. However, the enlarged views on the right side of the graph show that although the main part of the tree class had extremely high values for the NDVI and height, overlapping areas existed between the thatched houses and the trees because some parts (most likely the outer edges) of the trees had

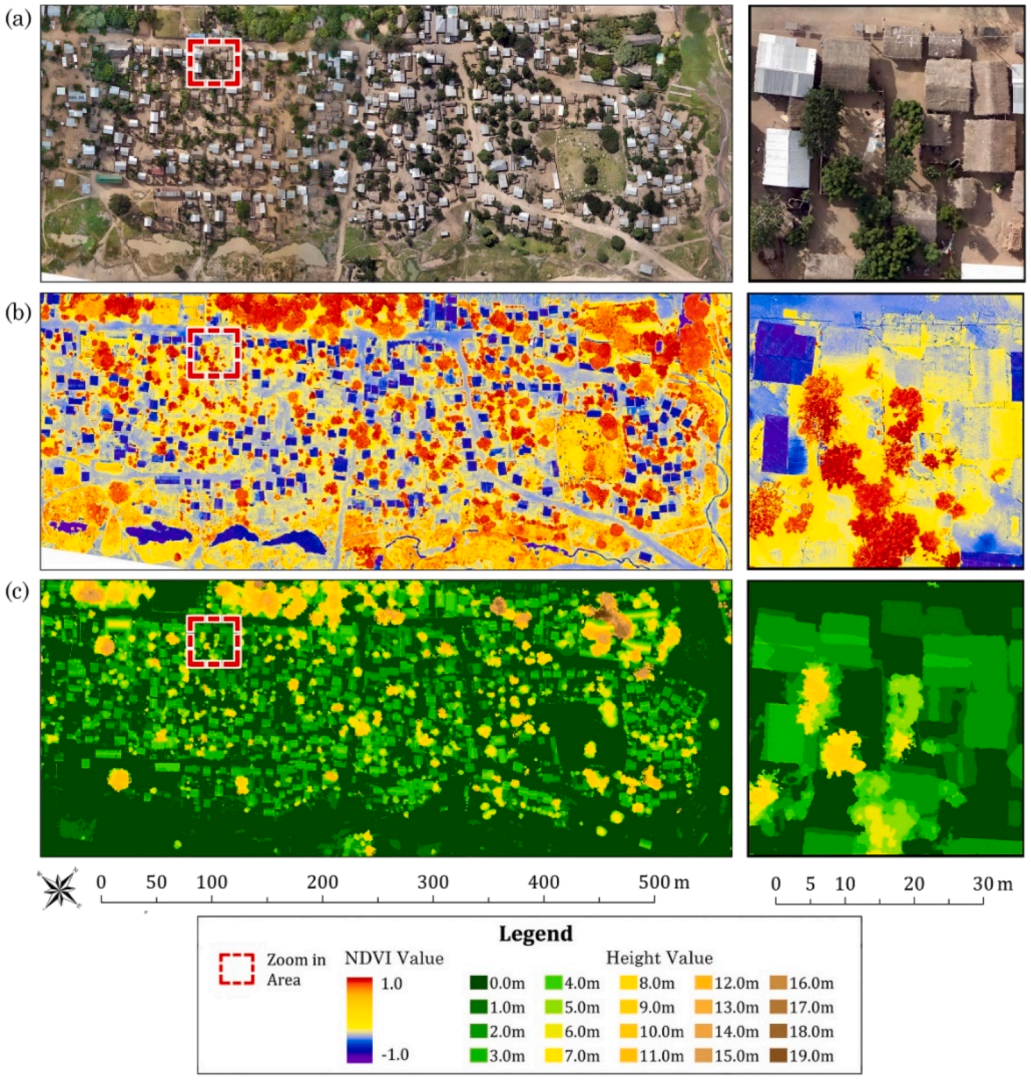


Fig. 3. Optical and height map of the study site: (a) Orthomosaic; (b) NDVI map; (c) Height map.

similar height as the thatched house. The outer edge of a tree often has a lower height and lower leaf density, which causes a lower NDVI value and a more similar RGB colour to brown than green, making them more similar to deadwood than the green canopy of the tree. These parts of individual trees are often excluded from the tree class when the classification is used to identify trees in fields such

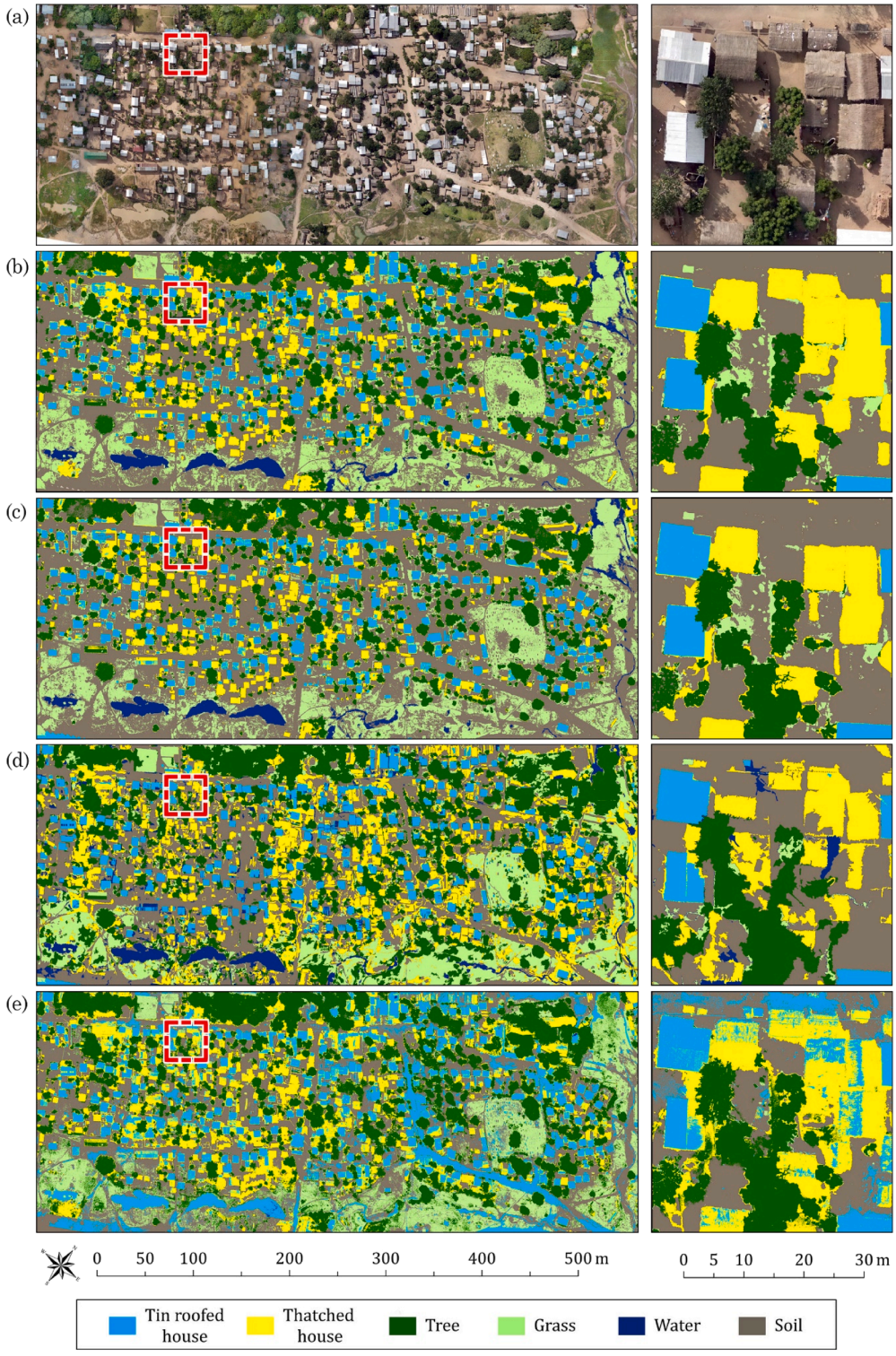


Fig. 4. Comparison between real landscape and the classification results: (a) Real landscape; (b) RCI method; (c) Ensembled method; (d) OBIA method; (e) RF method.

as forestry, agroforestry, and ecosystem services. However, when conducting precise LULC classification in rural areas, excluding these areas could lead to poor classification accuracy by mis excluding objects with similar outlooks, such as the thatched house in this study. In other words, this overlapping area between the thatched house and lower trees is considered a potential challenge for separating the thatched house from the lower parts of the trees, even when combining optical and structural information. In order to better explain the concept of RCI, the facts above were summarized and listed up succinctly earlier at 2.4.4.

Classification results comparison

Fig. 4 shows the results of the four classification methods used in this study.

Based on the overall visual appearance of the classification result raster, the RCI, threshold, and OBIA methods were able to separate the thatched house class, whereas the RF method mixed a part of thatched houses and tin houses, and also mis classified a part of soil into tin houses.

Specifically, the detailed views showed that compared with the results of the RCI method, fewer ground objects were classified as thatched houses by the ensembled threshold method, while somewhere that was not thatched houses was classified as hatched houses by the OBIA method. The RF method yielded similar result of the location of thatched houses, but contained noises, and also mis-classified a small part of soil as thatched houses. All the three methods except the RCI method classified the edge of trees as thatched

Table 1
Confusion matrices of the classification results by the five methods.

(a) Confusion Matrix of RCI Method

Class	Tin	Thatch	Tree	Grass	Water	Soil	Total	U_Accuracy	OA
Tin	332	11	0	0	1	1	345	0.962	
Thatch	21	327	4	1	0	41	394	0.830	
Tree	1	1	369	1	0	3	375	0.984	
Grass	3	0	25	298	1	21	348	0.856	
Water	0	0	0	1	246	35	282	0.872	
Soil	0	8	3	15	10	371	407	0.912	
Total	357	347	401	316	258	472	2151	0.000	
P_Accuracy	0.930	0.942	0.920	0.943	0.953	0.786	0.000	0.903	0.903
Kappa								0.884	

(b) Confusion Matrix of the ensembled Method

Class	Tin	Thatch	Tree	Grass	Water	Soil	Total	U_Accuracy	OA
Tin	332	11	0	0	1	1	345	0.962	
Thatch	18	241	4	0	0	39	302	0.798	
Tree	1	1	367	1	0	4	374	0.981	
Grass	3	1	25	293	1	23	346	0.847	
Water	0	0	0	1	244	34	279	0.875	
Soil	3	93	5	21	12	371	505	0.735	
Total	357	347	401	316	258	472	2151	0.000	
P_Accuracy	0.930	0.695	0.915	0.927	0.946	0.786	0.000	0.859	0.859
Kappa								0.830	

(c) Confusion Matrix of OBIA method

Class	Tin	Thatch	Tree	Grass	Water	Soil	Total	U_Accuracy	OA
Tin	311	2	1	0	3	2	319	0.975	
Thatch	10	228	5	5	15	99	362	0.630	
Tree	4	21	354	105	1	25	510	0.694	
Grass	0	3	32	190	2	42	269	0.706	
Water	18	7	0	2	227	18	272	0.835	
Soil	14	86	9	14	10	286	419	0.683	
Total	357	347	401	316	258	472	2151	0.000	
P_Accuracy	0.871	0.657	0.883	0.601	0.880	0.606	0.000	0.742	0.742
Kappa								0.688	

(d) Confusion Matrix of RF method

Class	Tin	Thatch	Tree	Grass	Water	Soil	Total	U_Accuracy	OA
Tin	300	58	1	0	230	93	682	0.440	
Thatch	54	279	8	3	0	79	423	0.660	
Tree	2	8	388	85	1	31	515	0.753	
Grass	1	0	3	220	0	19	243	0.905	
Water	0	0	0	0	0	0	0	0.000	
Soil	0	2	1	8	27	250	288	0.868	
Total	357	347	401	316	258	472	2151	0.000	
P_Accuracy	0.840	0.804	0.968	0.696	0.000	0.530	0.000	0.668	0.668
Kappa								0.598	

houses. For vegetation classes, the RCI, threshold, and OBIA methods successfully separated tree and grass classes. However, all the three methods except the RF method mis classified a part of trees as grass, showing a strength of pixel-based machine learning methods on classifying vegetation. All the three methods except the RF method were able to extract the water area. The RF method mis classified all water area as tin houses, while the OBIA method mis classified some parts of soil as water. The visual LULC mapping results showed that the four methods had both advantages and disadvantages at certain points or certain object identification. Therefore, a quantitative evaluation was showed at the next section.

Accuracy assessment results

A confusion matrix is a common and comprehensive accuracy assessment tool for LULC classification. The accuracy assessment results for the four classification methods are presented as confusion matrices based on the accuracy assessment points in Table 1.

In the confusion matrix, a row and a column are both named as "Total". The numbers in "Total" row shows the ground truth obtained by visual judgement. As shown in the "Total" row in each confusion matrix, among the randomly created 500 accuracy assessment points, there were 357, 347, 401, 316, 258, and 462 points for the tin house, thatched house, tree, grass, water area, and soil class, respectively. In contrast to the "Total" row, the "Total" column shows the classification results. If the value of the "Total" column is higher than that of the "Total" row for a ground object class, it means that there were more points that were classified as that class than the actual number; and vice versa when the "Total" column is lower than the "Total" row. This means, for example, as shown in the "Tin" row in Table 1**(d)**, there were 682 points that were classified as the tin house by the RF method, while only 300 of them really belonged to the tin house class, and 58 of them actually belonged to thatched house class, 1 point actually belonged to the tree class, 230 actually belonged to water class, and 93 actually belong to the soil class. The data in the "Tin" row of the RF methods showed that RF used in this study had a serious problem mis classifying other ground objects as tin houses. Likewise, as shown in the "Tin" column, there were 357 points that actually belonged to the tin house class, while only 300 of them were classified accurately by the RF method, and 54 of them were misclassified as the thatched house, 2 as the tree, and 2 as the grass. The "Tin" column showed that the RF method had a slight problem missing some points that should had been classified as the tin house, and mis classified them as something else.

Based on the information in the confusion matrix, four indicators were calculated and are listed in Table 1, which are the user's accuracy (U_Accuracy), producer's accuracy (P_Accuracy), overall accuracy (OA), and the kappa index. The user's accuracy presents a false positive by showing the percentage of true answers among all classification results. On the other hand, the producer's accuracy presents a false negative by showing the percentage of correct classification results among the entire ground truth. For example, the user's accuracy of the tin house class by the RF method was 0.440, meaning that among all the points that were classified as tin houses by this method, only 44 % were classified correctly, whereas the producer's accuracy was 0.840, meaning that among all the points that should have been classified as tin houses, 84 % were classified correctly using this method. The overall accuracy is one of the most basic accuracy measures which is calculated by dividing the correctly classified points (sum of he values in the main diagonal) by the

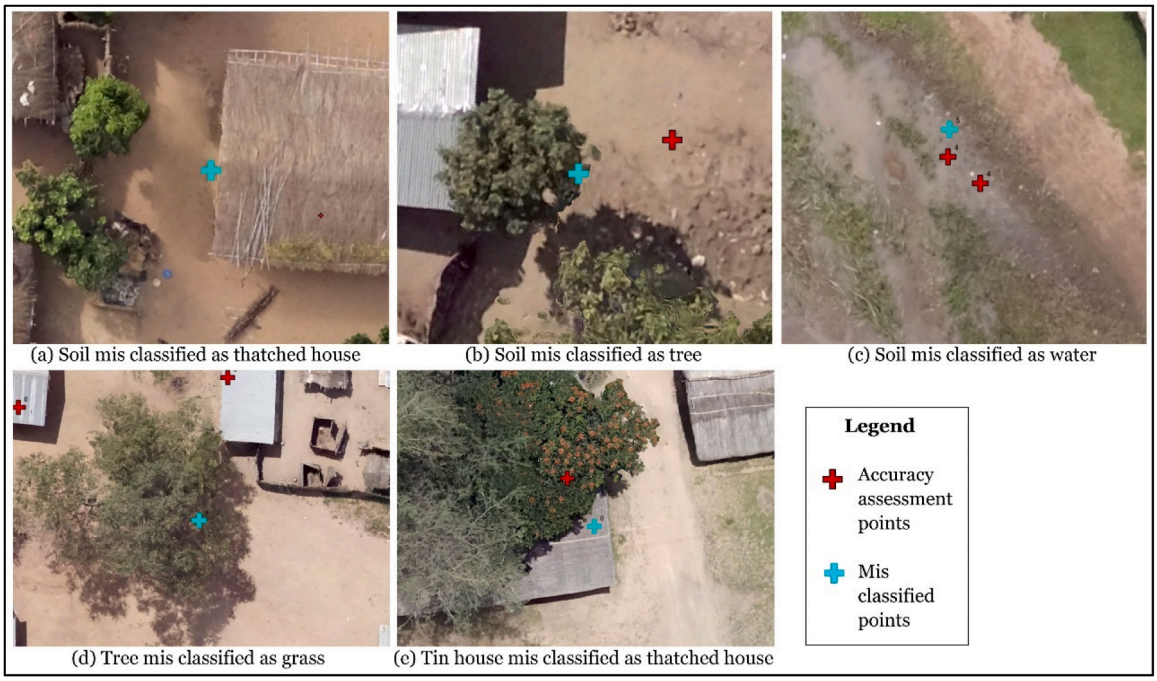


Fig. 5. Representative examples of mis classification by the RCI method.

total number of the accuracy assessment points. The kappa index is a statistical indicator used for overall classification performance testing and is considered to have higher objectivity than the overall accuracy by considering the non-elements of the confusion matrix in the calculation [39].

The RCI method had the highest classification accuracy (OA = 0.903, kappa = 0.884), followed by the ensembled threshold method, which had a slightly lower accuracy (OA = 0.859, kappa = 0.830). The OBIA method yielded moderate accuracy (OA = 0.742, kappa = 0.688), whereas the RF method had relatively low accuracies (OA = 0.668, kappa = 0.598).

Specifically, although the RCI method yielded the highest overall accuracy, it had both advantages and disadvantages. For all the classes except soil, both user's accuracy and producer's accuracy were above 0.800, which were considered as excellent. For the soil class, although user's accuracy (0.912) was high, producer's accuracy (0.786) was significantly low compared with the other classes. This fact showed the most important problem of the RCI method, which is mis classifying soil as other ground objects. According to the "Soil" column, among all the 472 accuracy assessment points that belonged to the soil class, the RCI method only succeeded at identifying 371 (79 %) of them. Among the 101 points that missed by the RCI method, 41 were classified as thatched houses, 35 were classified as water, and 21 was classified as grass. To unravel the reason of this problem of the RCI method, this study observed all the points that were mis classified by the RCI method and summarized the representative examples in Fig. 5**(a–c)**. The examples shoed that the soil area near to the thatched houses, tall trees and very shallow water area were possible to be mis classified. In contrast to the performance of soil class, for the classes of thatched house, grass and water area, the user's accuracy (< 0.900) was lower than the producer's accuracy (> 0.900). According to each row of these classes in the confusion matrix, there were 21 tin house points mis classified as thatched house by the RCI method, 25 grass points mis classified as tree, and 35 water point mis classified as soil. The examples of these mis classification were also showed in Fig. 5**(**d, e**)**. This fact showed the second problem of the RCI method, which was slightly overclassifying thatched houses, grass and water area, when the optical or structural features or the object edges were confusing or blurred. Finally, for the classes of tin house and tree, both the user's accuracy and the producer's accuracy were above 0.900, indicating the RCI had excellent performance identifying tin houses and trees.

The ensembled threshold method yielded the second highest overall accuracy (OA = 0.859, kappa = 0.830) next to the RCI method. Since the RCI method was based on the ensembled threshold method with modification using original index for thatched houses identification, the accuracy for the classes other than thatched house by these two methods did not show significant differences. Without the modification by RCI, the threshold had both lower user's accuracy (0.798) and lower producer's accuracy (0.695) for the thatched house classification. Same accuracy pattern was observed for the soil class too. This fact showed that instead of over classifying or under classifying a certain class like the RCI method, the ensembled threshold method had a problem of mixing up two classes.

As the comparison to the RCI method, the OBIA method and the RF method, both basing on the machine learning algorithms showed some notable differences. Firstly, for the tin house class and water area class, the OBIA method yielded excellent user's accuracy (0.975 and 0.835) and producer's accuracy (0.871 and 0.880). On the other hand, the RF method showed poor user's accuracy (0.440) for tin house class because it failed to identify water area. This was because the size of water area inside the study area was small, causing a relatively small training sample size. Although the lack of training samples also existed in the OBIA method, the segmentation process before excluding machine learning algorithm improve the accuracy of classes with regular and smooth surfaces such as tin houses and water area. This difference between the pixel-based and object-based machine learning also effected the accuracy for thatched house classification. Although the RF end to over classify the thatched houses, for example classify tin houses and trees as thatched houses, it succeeded in identifying most of the thatched houses and showed a better performance than the OBIA method. On the other hand, because of the segmentation process, a part of thatched houses which had similar reflectance with the surrounding area such as soil was mixed up with the soil even before the machine learning algorithm. This fact showed that the pixel-based RF method was more suitable for thatched house classification compared with the OBIA method. Finally, both the RF and OBIA method showed an acceptable accuracy for tree and grass.

Discussion

Comparison with previous studies

The thatch used for roofing in Malawi is mainly made from tall grass or wood branches growing natively in the bush or forest near communities and is considered a forest resource. The thickness of the roofs varies between 8 and 16 in, with the thatches packed tightly in a near-vertical position on top of the roof to ensure drainage. The height of the roofs varies between 2 and 3 m, and the area ranges between 20 and 600 m^2. These characteristics have made thatched houses a unique LULC type directly made from natural resources in Malawi, as well as in other countries in Africa, Asia, and Europe. Previous studies have suggested that the performance of LULC varies for different ground objects, regardless of whether the classification is conducted using machine learning or threshold approaches, OBIA, or pixel-based image analysis [40]. The accuracy of a classification result depends not only on the techniques, but also on the location of the site of interest, the contents inside the site, and the characteristics of the objects, suggesting the necessity of moderating or originally developing a comprehensive classification method which is optimal for each type of ground object inside the site of interest. In this study, variations in different ground objects using the same classifier were observed for all four classification methods, and among all six ground object types, the classification results for the thatched house were the most notable because the RCI method showed a much higher accuracy on this class than the machine learning algorithms.

During satellite remote sensing, thatched houses were not a major consideration for LULC classification because of the limitations of the spatial resolution of most satellite imagery. However, the importance of monitoring the number and dynamics of thatched

houses as a traditional and relatively major house style in the rural areas of Malawi has been gathering attention lately. In recent years, the rapid development of UAVs has made it possible to obtain high-resolution aerial imagery (ground resolution: 1–10 cm) of a certain area at a low cost, and to implement clear observations of thatched houses and other small ground objects. There have been several studies that compared the performance of satellite and UAV imagery on LULC classification and declared a higher classification accuracy by UAVs. However, UAV remote sensing will never replace satellite remote sensing but should be a complementary approach for a more practical development of remote sensing because it focuses on different points: large-area or precision observations. Higher precision does not necessarily result in higher accuracy. With the breakthrough implementation of ground resolution, traditional and unitary classification methods cannot achieve higher classification accuracy and may cause more noticeable misclassification. In other words, the more detailed the information that is observed, the newer and more comprehensive the classification needed to achieve proper classification. Although UAV remote sensing has achieved clear vision on thatched houses, the method to automatic identify them still need to be modified. This was the motivation of this study to developing an original index specially for separating thatched houses from other confusing objects in rural areas.

Because of the similar reflectance characteristics of soil and low-density tree branches, identifying thatched houses or deadwood resources has been challenging for many studies in the forestry, agroforestry, and ecosystem service fields. Chen et al. (2021) used high-resolution RGB satellite imagery (QuickBird, ground resolution: 2.4 m) to monitor the changes in the percentage of thatched houses in a rural setting in Mozambique between 2004 and 2012 using the OBIA method. However, the accuracy of the thatched house classification varied greatly in different years, particularly the classification by the satellite imagery taken in 2004, which showed high accuracy (user's accuracy: 0.58; producer's accuracy: 0.88) [22]. Chan et al. (2022) used RGB UAV imagery (ground resolution: 15 cm) for LULC mapping in refugee camps in Malawi and Kenya using deep learning (U-Net architecture) in 2022. The classification results showed high accuracy in detecting tin roofs, whereas the thatched houses, which were the main structures at the refugee camps, were completely undetectable by the deep learning algorithm [21]. These results indicate that classification based simply on spectral characteristics cannot provide satisfactory results for detecting thatched houses. Moreover, when simply using optical data, the OBIA method performs better than machine learning and deep learning approaches, which tend to completely ignore thatched houses and classify them as bare ground. Therefore, Wouters et al. (2021) conducted LULC classification using RGB UAV imagery (ground resolution: 11 cm) in southern Malawi in 2018 using the OBIA method and obtained similar results. The spectral difference caused by the materials led to a higher F1-score for tin houses (0.89), but a low F1-score for thatched houses (0.53). The study further moderated the classification model by considering height information from the DSM. After adding height data as a parameter for the OBIA classification, the F1-score of the thatched houses increased, but was still considered moderate (0.72) [20]. This finding showed that considering height information could improve the classification of thatched houses, but the improvement was not significant when height data were added as another parameter of the training samples. The results from previous studies showed the difficulty in accurately detecting thatched houses has become a crucial problem for community planning and disaster risk assessment in rural areas in developing countries.

Against this background, the novel method suggested by this study specialised in detecting thatched houses and different ground objects to accomplish a comprehensive classification methodology for LULC mapping in rural areas. As explained in the Results, thatched houses, bare ground, and low-height and low-density tree canopies have similar characteristics (spectral and structural) to each other. When using the threshold method to separate these objects, as explained in Methods, using two thresholds would cause over-exclusion of the thatched house, which can be observed in Fig. 4**(c)**. Fig. 4**(a and b)** showed that a similar problem occurs when using machine learning approaches. The training samples with either close spectral features or close height information that have been defined as three different classes cause more noise information and confusion when constructing the classification model and eventually result in classifying all three classes as one, which happened in the RF method of this study and in the deep learning method of Chan et al. (2022) [21]. This is also the reason why the OBIA method, which also applies the machine learning algorithm only at the object base, did not achieve significant accuracy improvement even after considering the height information [22]. By the comparison with the related studies so far, the authors found that all the previous studies and traditional classification methods used the optical information and the structure information as two different parameters, which is the primary problem causing the misclassification. Therefore, we developed an index containing the NDVI and (DSM–DTM) values. As introduced in the Methods and Eq. (2), the calculation using the NDVI and height data enhanced the differences between tin houses, thatched houses, trees, and bare ground, and eventually achieved the highest classification accuracy in the study site.

Just as no matter how precise the spatial resolution is, UAV cannot replace satellite in the field of remote sensing, the threshold method, no matter how many parameters are used for modification, cannot replace supervised machine learning in the field of LULC mapping. However, there is possibility that they can be mutual complementation to each other, to better achieve the diverse purpose by different users. Without doubt, the classification accuracy by machine learning will be improved with a drastic increase of the training samples and model tuning. The only problem is it may not be the most efficient approach in certain cases, for example, the LULC classification at a small settlement with simple ground object category as the study site of this study. In these cases, there is space to use experience and logical calculation to optically suit the certain and sometimes unique local requirement efficiently. Under this philosophy, the RCI method suggested an alternative way that uses logical calculation as a shortcut aiming straightforward to the LULC classification in rural areas with the intermingling of thatched houses, tin roofed houses, tree, grass, and water area.

Limitation and future tasks

Although the RCI achieved the highest average accuracy for all classes, the average accuracy ($\leqq 0.900$) for thatched houses, grass and soil still needs to be improved. Although the user's accuracy, producer's accuracy, overall accuracy and kappa index are practical

when comparing different classification methods, more detailed characteristics for each particular classification method can be unravelled from the confusion matrices by further analysing the mis classified points. Instead of using the number of the misclassification as one error indicator, calculating the percentage of each kind of mis classification can provide the information of that which classes were most easily confused by a certain classification method. For example, in Table 1 **(a)**, out of all the 57 mis classified points, 54 points were mis classified as thatched house. This means, by the RF method, tin houses were most easily mis classified as thatched house with the percentage of approximately 15 % (54/357). This kind of information can evaluate every single classification method more comprehensively and specifically. Based on this consideration, five most easily confused classes for the RCI method are listed as below in the order of seriousness. If the degree of mis classification is not acceptable for the demand of accuracy when conducting LULC classification in rural areas, it is recommended to conduct further adjustment of the threshold for both RCI and HNDVI through trial and error.

(1) A notable trend (approximately 9 %) of mis classifying soil as thatched house.
(2) A relatively notable trend (approximately 7 %) of mis classifying soil as water.
(3) A relatively notable trend (approximately 6 %) if mis classifying tree as grass.
(4) A slight trend (approximately 5 %) of mis classifying thatched house as tin roofed house.
(5) A slight trend (approximately 5 %) of mis classifying tree as grass.
(6) A very slight trend (approximately 4 %) of mis classifying soil as water.

Moreover, this study used limited training data to maintain fairness on time and labour cost between the compared methods. A higher classification accuracy can be expected by increasing the number of training samples or tuning models, although this may also cause regional deviations which should be considered. Furthermore, the practicality of the method suggested in this study should be verified and further modified at various locations to improve its versatility.

Application values of the developed LULC methodology

This study is one of the few studies which intent to achieve high precision LULC classification specifically in rural areas in developing areas using remote sensing data. A new point of view of combing the elevational and spectral reflectance characteristics into one single index, which can maximize the difference among the typical LULC types in rural areas. The attempt of creating one index by standardizing two totally different types of information to separate the ground objects with similar appearance has not been reported by other studies in the field of LULC classification so far to the authors' knowledge, which makes it the most important novelty of this study. The novel comprehensive classification method proposed by this study not only achieved higher overall accuracy than the traditional threshold method or commonly used machine learning methods used as comparison inside the same study site, but also achieved higher classification accuracy on specific land use types such thatched houses than the previous studies conducted in similar field, which were described detailly at **4.1**. A high precision of LULC map data provided by this study will be of immediate interest for users from the wider research community, such as the fields of hydrological model, hydraulic model, agricultural engineering, land resources, forests/agroforestry and ecosystem, where LULC data is essential for the monitoring, simulating and predicting accuracy. For example, a higher accuracy of flood risk hazard map with precision LULC information can provide solutions for both disaster prevention countermeasures and post disaster recovery in rural areas, especially in Africa where thatched houses are homes for many but have been continually ignored in the field of remote sensing. This contributes to both climate resilient communities, and a high standard of living, quality of life and well-being for all citizens, which meet the first and seventh goal of the African's Union's Agenda 2063. Furthermore, a higher accuracy of forest resource modelling can be achieved by precision mapping of thatched houses since they are one of the main consumers of the wood resource. A better management of both forest and national parks using the precise LULC data contributes to achieve sustainable communities and also climate actions, which meet the third and thirteenth goal of the SDGs.

Conclusion

Ultra high-resolution and multispectral aerial imagery taken by UAVs and SfM technology have made it possible to conduct highly precise LULC classification in rural areas which contributes to natural hazard assessment, community planning, national surveillance, and natural resource management. However, with the breakthrough improvement in the ground resolution of aerial imagery, a review and reappraisal of the classification methodology considering the ability to separate small-scale ground objects with less distinctive features by considering their complex characteristics has become necessary.

In the case of rural areas in developing countries, identifying thatched houses using aerial imagery has been challenging in many previous studies and has caused difficulties in LULC mapping. By comparing the slight and complex optical and structural differences between thatched houses and other ground objects, this study suggests an original index (RCI) by combining the spectral and height information as a single figure, which can enhance the differences between the ground objects. As a case study in rural areas of Africa, the UAV and ground surveys of this study were employed at a small human settlement inside Lake Malawi National Park, after which multispectral and height maps were constructed using the SfM process. Four classification methods, namely the RCI method, the ensembled threshold method, OBIA method, and RF method, were used to classify the main ground objects inside the village, and the classification accuracy was then compared using confusion matrices. As the result, the RCI method achieved the highest overall accuracy and kappa index and performed particularly well in detecting thatched houses. The RCI method, providing an efficient and low cost way particular for the LULC classification in rural areas can potentially contribute to the damage control of natural disasters and

rural development, and eventually to a higher quality of life for villagers in developing countries.

CRediT authorship contribution statement

Ke Zhang: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Lameck Fiwa:** Investigation, Methodology, Resources, Supervision, Writing – review & editing. **Madoka Kurata:** Conceptualization, Investigation, Project administration, Resources, Writing – review & editing. **Hiromu Okazawa:** Project administration, Resources, Supervision, Funding acquisition, Writing – review & editing. **Kenford A.B. Luweya:** Investigation, Writing – review & editing. **Mohammad Shamim Hasan Mandal:** Methodology, Software, Writing – review & editing. **Toru Sakai:** Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

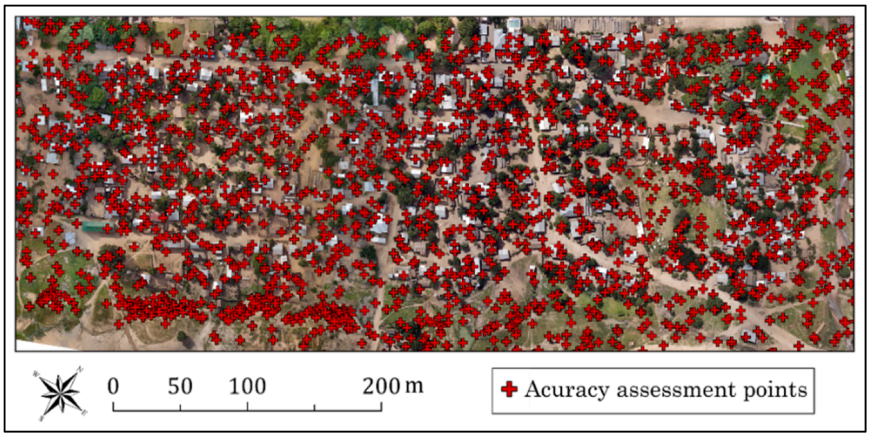
Acknowledgements

The authors would like to thank the IntNRMS (Establishment of a Sustainable Community Development Model based on Integrated Natural Resource Management System in Lake Malawi National Park) project supported by Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA) for providing the equipment and other resources for the surveys conducted in this study. The authors also acknowledge Japan International Research Center for Agricultural Sciences (JIRCAS) for supporting this study. Finally, we thank Elsevier Language Services for the language editing.

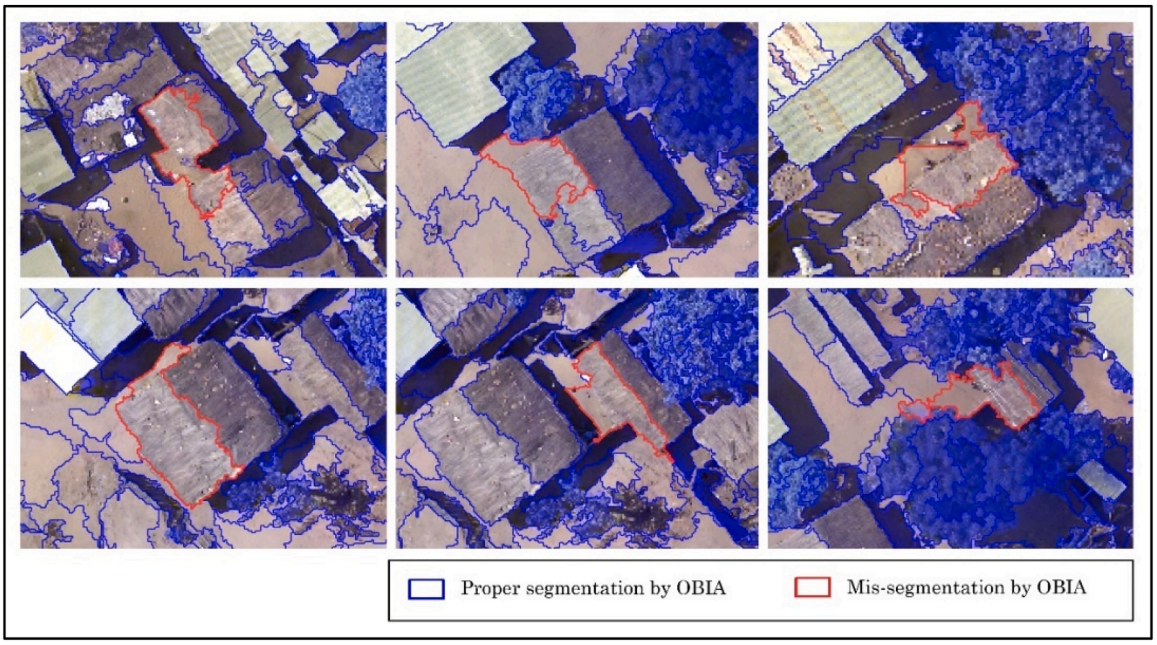
Abbreviations

DSM	Digital Surface Model
DTM	Digital Terrian Model
GCP	Ground Control Point
LULC	Land Use/Land Cover
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
OBIA	Object-based Image Analysis
RCI	Rural Residence Classification Index
RF	Random Forest
RTK-GNSS	Realtime Kinematic-Global Navigation Satellite System
SfM	Structure from Motion
SVM	Support Vector Machine
UAV	Unmanned Aerial Vehicle

Appendix A. 2000 accuracy assessment points distribution



Appendix B. Examples of mis-segmentation by OBIA (eCognition Developer ver. 10.3)



Appendix C. Average accuracy for each class by different methods

	RF	OBIA	Threshold	RCI
Tin House	0.640	0.923	0.946	0.946
Thatched House	0.732	0.657	0.746	0.886
Tree	0.860	0.788	0.948	0.952
Grass	0.801	0.654	0.887	0.900
Water	0.000	0.857	0.910	0.913
Soil	0.699	0.644	0.760	0.849
Total	0.668	0.742	0.859	0.903

References

[1] A. Elamin, A. El-Rabbany, UAV-based multi-sensor data fusion for urban land cover mapping using a deep convolutional neural network, Remote Sens (Basel) 14 (17) (2022) 4298.
[2] I. Ngoma, M. Sassu, Sustainable African housing through traditional techniques and materials: a proposal for a light seismic roof, in: Proceedings of the 13th World Conference on Earthquake Engineering, Vancouver, BC, Canada, 2004, pp. 1–6.
[3] Government of Malawi, 2015 Floods PDNA Report, Lilongwe, Malawi, 2015.
[4] DIEM-Impact assessment, FAO, (2023). Tropical cyclone Freddy, Malawi, 2023-The impact of tropical cyclone Freddy on agriculture and livelihoods.
[5] M. Abdulrazzak, A. Al-Shabani, K. Noor, A. Elfeki, A. Kamis, Integrating hydrological and hydraulic modelling for flood risk management in a high-resolution urbanized area: case study Taibah University campus, KSA, in: Recent Advances in Environmental Science from the Euro-Mediterranean and Surrounding Regions: Proceedings of Euro-Mediterranean Conference for Environmental Integration (EMCEI-1), Tunisia 2017, Springer International Publishing, 2018, pp. 827–829.
[6] T.N. Wickramaarachchi, Water availability assessment in cultivation and non-cultivation seasons to identify water security in a tropical catchment: gin catchment, Sri Lanka, Paddy and Water Environment (2023) 1–13.
[7] A. Khaliq, L. Comba, A. Biglia, D. Ricauda Aimonino, M. Chiaberge, P. Gay, Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment, Remote Sens (Basel) 11 (4) (2019) 436.
[8] L. Kapustina, N. Izakova, E. Makovkina, M. Khmelkov, The global drone market: main development trends, in: SHS Web of Conferences 129, EDP Sciences, 2021, p. 11004.
[9] M.A. Haq, G. Rahaman, P. Baral, A. Ghosh, Deep learning based supervised image classification using UAV images for forest areas classification, J. Indian Soc. Remote Sens. 49 (3) (2021) 601–606.
[10] L. Ma, M. Li, X. Ma, L. Cheng, P. Du, Y. Liu, A review of supervised object-based land-cover image classification, ISPRS J. Photogramm. Remote Sens. 130 (2017) 277–293.

[11] M.A. Haq, Planetscope nanosatellites image classification using machine learning, Computer Systems Science & Engineering 42 (3) (2022).
[12] L. Breiman, Random forests, Mach Learn 45 (2001) 5–32.
[13] M.A. Haq, CNN based automated weed detection system using UAV imagery, Computer Systems Science & Engineering 42 (2) (2022).
[14] D. Ventura, A. Bonifazi, M.F. Gravina, A. Belluscio, G. Ardizzone, Mapping and classification of ecologically sensitive marine habitats using unmanned aerial vehicle (UAV) imagery and object-based image analysis (OBIA), in: Remote Sens (Basel), 10, 2018, p. 1331. Ma, L., Li, M., Ma, X., Cheng.
[15] A.M. Lechner, A. Fletcher, K. Johansen, P. Erskine, Characterizing upland swamps using object-based classification methods and hyper-spatial resolution imagery derived from an unmanned aerial vehicle, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 1 (2012) 101–106.
[16] K. Zhang, S. Maskey, H. Okazawa, K. Hayashi, T. Hayashi, A. Sekiyama, L. Fiwa, Assessment of three automated identification methods for ground object based on UAV imagery, Sustainability 14 (21) (2022) 14603.
[17] B. Yang, T.L. Hawthorne, H. Torres, M. Feinman, Using object-oriented classification for coastal management in the east central coast of Florida: a quantitative comparison between UAV, satellite, and aerial data, Drones 3 (3) (2019) 60.
[18] Y. Lanthier, A. Bannari, D. Haboudane, J.R. Miller, N. Tremblay, Hyperspectral data segmentation and classification in precision agriculture: a multi-scale analysis, in: IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium 2, IEEE, 2008, pp. II–585.
[19] E. Neinavaz, A.K. Skidmore, R. Darvishzadeh, Effects of prediction accuracy of the proportion of vegetation cover on land surface emissivity and temperature using the NDVI threshold method, Int. J. Appl. Earth Obs. Geoinf. 85 (2020) 101984.
[20] L. Wouters, A. Couasnon, M.C. De Ruiter, M.J. Van Den Homberg, A. Teklesadik, H. De Moel, Improving flood damage assessments in data-scarce areas by retrieval of building characteristics through UAV image segmentation and machine learning–a case study of the 2019 floods in southern Malawi, Natural Hazards and Earth System Sciences 21 (10) (2021) 3199–3218.
[21] C.Y.C. Chan, M. Weigand, E.A. Alnajar, H. Taubenböck, Investigating the capability of UAV imagery for AI-assisted mapping of Refugee Camps in East Africa, in: Proceedings of the Academic Track at State of the Map 2022, 2022, pp. 45–48.
[22] D. Chen, T.V. Loboda, J.A. Silva, M.R. Tonellato, Characterizing Small-Town Development Using Very High Resolution Imagery within Remote Rural Settings of Mozambique, Remote Sens (Basel) 13 (17) (2021) 3385, https://doi.org/10.3390/rs13173385.
[23] L.S. Tusting, M.M. Ippolito, B.A. Willey, I. Kleinschmidt, G. Dorsey, R.D. Gosling, S.W. Lindsay, The evidence for improving housing to reduce malaria: a systematic review and meta-analysis, Malar. J. 14 (1) (2015) 1–12.
[24] G. Kritika, M. Thussu, Redefining cultural identity through architecture, Journal of Emerging Technologies and Innovative Research 7 (7) (2020) 1445–1457.
[25] S.G. Tambala, Redefining Lost Culture Through Architecture, DalTech-Dalhousie University, 1999.
[26] J.K.S. Villanueva, A.C. Blanco, Optimization of ground control point (GCP) configuration for unmanned aerial vehicle (UAV) survey using structure from motion (SFM), Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 42 (2019) 167–174.
[27] B. Awasthi, S. Karki, P. Regmi, D.S. Dhami, S. Thapa, U.S. Panday, Analyzing the Effect of Distribution Pattern and Number of GCPs on Overall Accuracy of UAV Photogrammetric Results, in: K. Jain, K. Khoshelham, X. Zhu, A. Tiwari (Eds.), Lecture Notes in Civil Engineering, Proceedings of the International Conference on Unmanned Aerial System in Geomatics, Roorkee, India, 6–7 Apil, Springer, Cham, Switzerland, 2019, pp. 339–354.
[28] K. Zhang, H. Okazawa, K. Hayashi, T. Hayashi, L. Fiwa, S. Maskey, Optimization of ground control point distribution for unmanned aerial vehicle photogrammetry for inaccessible fields, Sustainability 14 (15) (2022) 9505.
[29] R. Pamuji, A.I. Mahardika, N. Wiranda, N.A.B. Saputra, M.H. Adini, D. Pramatasari, Utilizing Electromagnetic Radiation in Remote Sensing for Vegetation Health Analysis Using NDVI Approach with Sentinel-2 Imagery. Kasuari, Physics Education Journal (KPEJ) 6 (2) (2023) 56–64.
[30] O. Tsuji, M. Yoneyama, M. Kimura, T. Muneoka, Estimation of Weed Community Area in Hayfield Using Small UAV, Journal of Japanese Society of Irrigation, Drainage and Reclamation Engineering 85 (10) (2017) 939–942.
[31] M.C. Jutras-Perreault, T. Gobakken, E. Næsset, H.O. Ørka, Comparison of Different Remotely Sensed Data Sources for Detection of Presence of Standing Dead Trees Using a Tree-Based Approach, Remote Sens (Basel) 15 (9) (2023) 2223.
[32] K. Zielewska-Büttner, P. Adler, S. Kolbe, R. Beck, L.M. Ganter, B. Koch, V. Braunisch, Detection of Standing Deadwood from Aerial Imagery Products: two Methods for Addressing the Bare Ground Misclassification Issue, Forests (11) (2020) 801.
[33] A. Haara, S. Nevalainen, Detection of dead or defoliated spruces using digital aerial data, For. Ecol. Manage. 160 (1–3) (2002) 97–107.
[34] N. Wang, Y. Du, F. Liang, H. Wang, J. Yi, The spatiotemporal response of China's vegetation greenness to human socio-economic activities, J. Environ. Manage. 305 (2022) 114304.
[35] M.A. Haq, CDLSTM: a novel model for climate change forecasting, Computers, Materials and Continua 71 (2) (2022) 2363–2381.
[36] D.R. Cutler, T.C. Edwards Jr, K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, J.J Lawler, Random forests for classification in ecology, Ecology 88 (11) (2007) 2783–2792.
[37] V. Lebourgeois, S. Dupuy, É. Vintrou, M. Ameline, S. Butler, A. Bégué, A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM). Remote Sensing 9 (3) (2017) 259.
[38] Banko, G. (1998). A review of assessing the accuracy of classifications of remotely sensed data and of methods including remote sensing data in forest inventory.
[39] R.G. Congalton, A review of assessing the accuracy of classifications of remotely sensed data, Remote Sens. Environ. 37 (1) (1991) 35–46.
[40] J. Rogan, J. Franklin, D. Stow, J. Miller, C. Woodcock, D. Roberts, Mapping land-cover modifications over large areas: a comparison of machine learning algorithms, Remote Sens. Environ. 112 (5) (2008) 2272–2283.