7th Intercontinental Geoinformation Days

igd.mersin.edu.tr

Assessing the contribution of RGB VIs in improving building extraction from RGB-UAV images

Richmond Akwasi Nsiah[*1], **Saviour Mantey** [2], **Yao Yevenyo Ziggah** [3]

[1]University of Mines and Technology, Faculty of Geosciences and Environmental Studies, Geomatic Department, Tarkwa, Ghana

Keywords
Building Extraction
UAV
GeoBIA
RGB Vegetative Indices
Random Forest

Abstract

Buildings are a fundamental component of the built environment, and accurate information regarding their size, location, and distribution is vital for various purposes. The ever-increasing capabilities of unmanned aerial vehicles (UAVs) have sparked an interest in exploring various techniques to delineate buildings from the very high-resolution images obtained from UAVs. However, UAV images have limited spectral information, and VIs have been adopted to increase the spectral strength of UAVs for building classification. This study aims to assess the contribution of four VIs, the green leaf index (GLI), red-green-blue vegetation index (RGBVI), visual atmospherically resistant index (VARI), and triangular greenness index (TGI), in improving building classification using geographic object-based image analysis (GeoBIA) approach and random forest classifier. For this purpose, five datasets were created and comprised of the RGB-UAV image and the RGB VIs. The experimental result indicated that the RGB + VARI dataset had the best improvement in the building classification based on four evaluation metrics: overall accuracy (0. 9799), precision (0. 9806), recall (0. 9806), and F1-score (0. 9806). The combination of all the VIs with the RGB image, on the other hand, attained results lower than the standalone RGB image: accuracy (0. 9507), precision (0. 9570), recall (0. 9368), and F1-score (0. 9468).

1. Introduction

Among the myriad of urban features, buildings represent a fundamental component (Schlosser et al., 2020), and obtaining accurate and detailed information on buildings is crucial for urban planning, infrastructure development, disaster management, and other applications (Hu et al., 2021). Recent advancements in unmanned aerial vehicle (UAV) technologies and the sophistication of imaging sensor systems have sparked an interest in exploring various methods to delineate building objects from very high-resolution (VHR) UAV imagery.

One such method is geographic object-based image analysis (GeoBIA), which has emerged as a powerful approach for automating the extraction of objects from remote sensing data (Comert & Kaplan, 2018). GeoBIA integrates machine learning algorithms, spatial information, and spectral characteristics to segment and classify image objects, making it particularly suited for building extraction (Aminipouri et al., 2009; Guo & Du, 2017).

While GeoBIA has shown considerable promise in building classification and segmentation, the spectral limitations of UAV-RGB imagery pose a challenge, especially when distinguishing between building materials and other urban (Li et al., 2022). Researchers have since used various ancillary datasets, such as vegetative indices (VI), to address this drawback in the classification process. VIs can capture subtle spectral variations and, when combined with GeoBIA's spatial context analysis, offer a promising avenue for improving building classification and segmentation urban (Öztürk & Colkesen, 2021; Schlosser et al., 2020).

Although some research works have focused on improving building classification using RGB VIs, a comprehensive comparison of the effect of each VI on classification accuracy has not been conducted. Consequently, the primary objective of this study is to investigate the impacts of integrating RGB-based VIs into the GeoBIA classification pipeline for building extraction. To achieve this objective, four well-established VIs: the green leaf index (GLI), red-green-blue vegetation index (RGBVI), visual atmospherically resistant index (VARI),

* **Corresponding Author**

*(ransiah94@gmail.com) ORCID: 0009-0003-4100-9570
(smantey@umat.edu.gh) ORCID: 0000-0002-8210-3577
(yyziggah@umat.edu.gh) ORCID: 0000-0002-9940-1845

Cite this study

Nsiah, R. A., Mantey, S., & Ziggah, Y. Y. (2023). Assessing the contribution of RGB VIs in improving building extraction from RGB-UAV images. Intercontinental Geoinformation Days (IGD), 7, 157-160, Peshawar, Pakistan

and triangular greenness index (TGI) were employed and combined with UAV-RGB imagery. The efficacy of each amalgamation was assessed using key performance metrics such as overall accuracy (OA), F-1 score, and Kappa coefficient.

2. Method

The methodological framework adopted for this research is presented in the subsequent subsections.

2.1. Study Area

The New Mankessim community is within the administrative jurisdiction of the Tarkwa Nsuaem Municipal Assembly, located approximately 19.30 kilometres southwest of the municipal capital, Tarkwa, in the Western Region of Ghana. Geographically, the community is positioned at latitude 5°5’ 29.45” N and longitude 2°6’ 4.70” W, nestled at an average altitude of 55 meters above mean sea level. A resettlement program initiated by one of the prominent mining companies operating in the region, primarily to accommodate the evolving dynamics of mining activities in the area, led to the relocation of community members from previous dwellings to the current location. As such, a notable feature of the New Mankessim community is the uniformity in architectural designs across the settlement. The community’s well-planned layout is marked by a consistent architectural style, reflecting a cohesive and deliberate approach to urban development. Fig. 1 depicts the UAV Image of the study area.

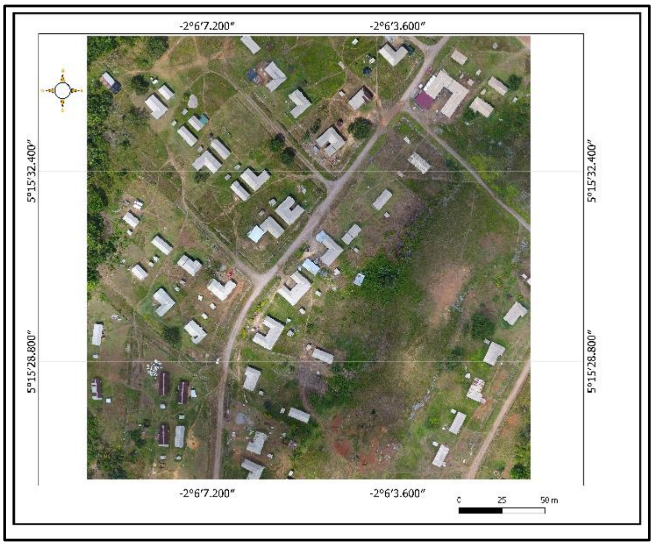

Figure 1. UAV Image of New Mankessim

2.2. Geographic Object-Based Image Analysis (GeoBIA)

GEOBIA is an image analysis approach commonly applied to VHR remote sensing data. It serves various purposes, including land-cover mapping and identifying specific geographic objects like buildings, cars, and trees (Kucharczyk et al., 2020). The workflow of the GeoBIA approach involved image segmentation, feature selection, image classification, and accuracy assessment and was carried out using the Google Earth Engine platform.

2.3 Image Segmentation

This step involved segmenting images into image objects, groups of neighbouring pixels representing objects within the drone image based on spectral and spatial attributes. There are various methods for performing image segmentation. However, the simple linear iterative clustering (SLIC) algorithm (Achanta et al., 2012) was utilised. SLIC is a seed-based clustering technique that effectively utilises a modified k-means clustering strategy to create superpixels with high efficiency. In contrast to prior methodologies, SLIC excels in preserving boundaries while offering improved speed and memory efficiency (Liao et al., 2022).

It also enhances segmentation performance and can be extended for super voxel generation. This method carefully incorporates considerations of both color homogeneity and shape uniformity, achieving a well-balanced trade-off between these aspects (Zhang & Zhu, 2019). Utilising SLIC requires defining several parameters, such as compactness, seed size, and grid type, to obtain optimum and homogenous image objects. For this work, the parameters were determined using a trial-and-error method.

2.4 Feature Selection

The spectral attributes, including the mean values of the red, green, and blue (RGB) bands and VIs within each image object, were selected as the primary features for building extraction. These features capture colour information for distinguishing building objects from other urban features. In total, 916 samples were selected, 456 representing buildings and the remaining non-building objects. These were divided into training (80%) and validation (20%) sets.

2.5 Classification

The step involves using a machine learning classifier to classify the segments into respective classes. For this research, the random forest (RF) classifier proposed by Breiman (2001) was employed to classify the selected features as either buildings or non-buildings. RF is an ensemble machine learning algorithm that combines multiple decision trees to make predictions. Each tree in the forest is trained on a different subset of the data with bootstrapping and random feature selection. The final prediction is determined by a majority vote or averaging of individual tree predictions, making it robust, accurate, and less prone to overfitting, making it robust and effective in handling complex classification tasks (Kumar & Sinha, 2020; Xiao et al., 2020). For this research, the RF classifier was trained using the selected samples, with the mean values serving as input features. Like the image segmentation step, RF also has several parameters that need to be fine-tuned for optimum classification, which were defined using a trial-and-error approach.

2.6 Evaluation Metrics

A comprehensive validation approach is adopted to assess the accuracy of the building classification. The trained RF classifier is applied to the validation data to classify buildings and non-buildings. The results are compared with ground truth data to evaluate classification performance using overall accuracy, precision, recall, and F1-score, which were computed using a confusion matrix—equations (1) to (4) give the mathematical formulations for the evaluation metrics.

$$\text{Recall, R} = \frac{TP}{TP+FN} \quad (1)$$

$$\text{Precision, P} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Overall Accuracy, OA} = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

$$\text{F1-score, F1} = \frac{2*Precision*Recall}{Precision+Recall} \quad (4)$$

Where TP represents correctly identified building pixels, TN indicates correctly identified non-building pixels, FP represents pixels erroneously classified as buildings but are not, and FN denotes pixels overlooked as non-building despite being so.

2.7 RGB-Vegetative Indices

Vegetation indices (VIs) are derived through mathematical equations applied to two or more spectral bands to highlight specific vegetation attributes (Öztürk & Colkesen, 2021). Several VIs that utilise the RGB bands have been created and developed. The RGB VIs utilised in this research are depicted in Fig. 2, and the formulae are in Table 1.

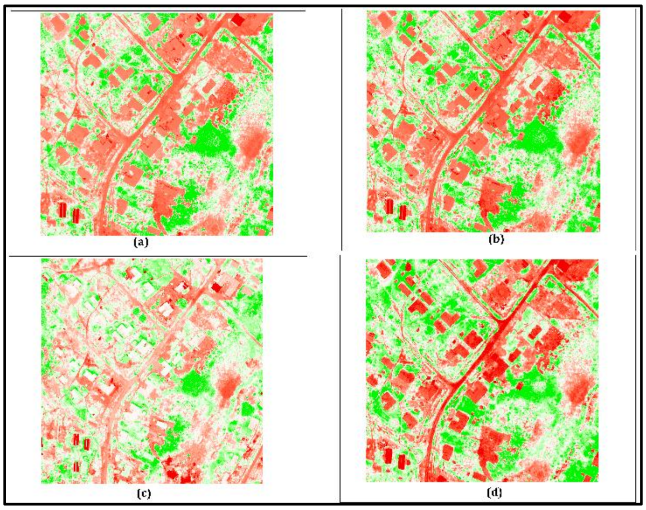

Figure 2. RGB VIs (a)GLI, (b)RGBVI, (c)VARI, and (d)TGI

Table 1. RGB-VIs Utilised

VI	Formula	Reference
Green Leaf Index	GLI = $\frac{(2\times Green)-Red-Blue}{(2\times Green)+Red+Blue}$	Hunt et al. (2012)
Red-Green-Blue Vegetation Index	RGBVI = $\frac{Green^2-Blue\times Red}{Green^2+Blue\times Red}$	Bendig et al. (2015)
Visual Atmospherically Resistant Index	VARI = $\frac{Green-Red-Blue}{Green+Red+Blue}$	Gitelson et al. (2002)
Triangular Greenness Index	TGI = $Green - (0.39 \times Red) + (0.61 \times Blue)$	Louhaichi et al. (2001)

3. Results

For this study, five datasets were created by combining the RGB VIs with the UAV-RGB image. These were RGB and GLI, RGB and RGBVI, RGB and VARI, RGB and TGI, and RGB and all indices. Spectral and spatial information were subsequently selected from each combination and used to train and validate the RF classifier. The evaluation results obtained for each combination, based on the evaluation metrics, are provided in Table 2.

Table 2. Performance of Various Dataset Combinations

Dataset	Metric			
	OA	P	R	F1
UAV-RGB only	0.9565	0.9643	0.9529	0.9586
RGB + GLI	0.9632	0.9897	0.9411	0.9648
RGB +RGBVI	0.9660	0.9671	0.9671	0.9671
RGB + VARI	0.9799	0.9806	0.9806	0.9806
RGB + TGI	0.9714	0.9880	0.9535	0.9704
RGB + All Indices	0.9507	0.9570	0.9368	0.9468

Fig.3. illustrates the extraction results produced by the random forest classifier for each dataset combination.

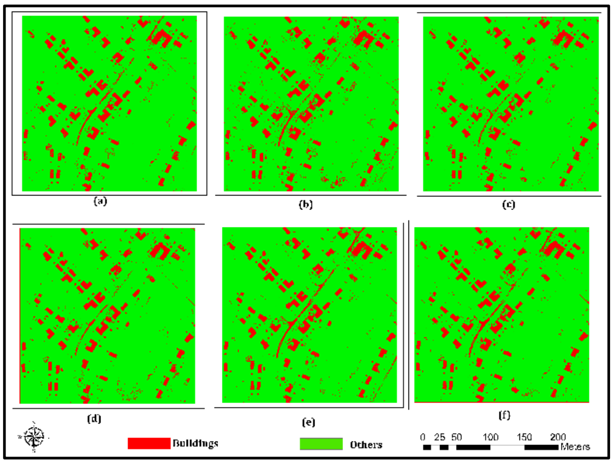

Figure 3. Building Extraction Results RF Classifier (a) UAV Image, (b) RGB + GLI, (c) RGB + RGBVI, (d) RGB + VARI, (e) RGB + TGI, and (f) RGB + All Indices

4. Discussion

The results presented in Table 2 show that the RGB VIs had a significant impact on the building extraction task. All the evaluation metrics were generally improved when the VIs were added to the RGB-UAV image. Notably, the RGB + VARI dataset achieved the highest OA at 0.9799, highest recall at 0.9806, and highest F1 at 0.9806, indicating that the VARI index contributed significantly to accurate building extraction, was influential in capturing most building features and exhibited a strong balance between precision and recall. For precision, however, the RGB + GLI dataset attained the highest precision at 0.9897, indicating minimal false positives.

Surprisingly, combining all available indices, i.e., the RGB + All Indices dataset, resulted in a lower OA and F1 score than some individual index combinations. The thematic maps show that the similar visual outputs were produced by the datasets. Regardless, it is observed that the RGB + VARI dataset had few false positives compared to the others.

5. Conclusion

This study aimed to assess the contribution of four RGB VIs, GLI, RGBVI, VARI, and TGI, in improving building classification tasks from UAV imagery. To that aim, four datasets containing a combination of these VIs and RGB-UAV were created, and a GeoBIA approach was adopted to classify building features from these datasets. In addition, a fifth dataset was created by combining all the RGB VIs and the UAV image. The experimental results highlight the advantages of integrating vegetative indices into building extraction from UAV-RGB imagery. The RGB + VARI dataset emerged as the top-performing combination, achieving the highest overall accuracy, precision, recall, and F1-score levels. However, it is worth noting that the RGB + GLI dataset stood out for its exceptional precision, rendering it particularly suitable for applications where minimizing false positives is paramount. However, combining all the RGB VIs with the RGB image produced lower metric scores than the standalone RGB image.

The consistent performance of RGB + VARI across various metrics accentuates its effectiveness as a standalone index.

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Susstrunk, S. (2012). SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(11), 2274–2281.

Aminipouri, M., Sliuzas, R., & Kuffer, M. (2009). Object-Oriented Analysis of Very High-Resolution Orthophotos for Estimating the Population of Slum Areas, A Case of Dar-Es-Salaam, Tanzania. ISPRS Hannover Workshop 2009 High-Resolution Earth Imaging for Geospatial Information, Hannover Germany, W5.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M. L., & Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible and near-infrared vegetation indices for biomass monitoring in barley. International Journal of Applied Earth Observation and Geoinformation, 39, 79–87.

Breiman, L. (2001). Random Forests. Machine Learning - Springer, 45(1), 5–32.

Comert, R., & Kaplan, O. (2018). Object Based Buıldıng Extractıon And Buıldıng Perıod Estımatıon From Unmanned Aerıal Vehicle Data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4(3), 71–76.

Gitelson, A. A., Kaufman, Y. J., Stark, R., & Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. Remote Sensing of Environment, 80(1), 76–87.

Guo, Z., & Du, S. (2017). Mining parameter information for building extraction and change detection with very high-resolution imagery and GIS data. GIScience and Remote Sensing, 54(1), 38–63.

Hu, Q., Zhen, L., Mao, Y., Zhou, X., & Zhou, G. (2021). Automated building extraction using satellite remote sensing imagery. Automation in Construction, 123, 103509.

Hunt, E. R., Doraiswamy, P. C., McMurtrey, J. E., Daughtry, C. S. T., Perry, E. M., & Akhmedov, B. (2012). A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. International Journal of Applied Earth Observation and Geoinformation, 21(1), 103–112.

Kucharczyk, M., Hay, G. J., Ghaffarian, S., & Hugenholtz, C. H. (2020). Geographic object-based image analysis: A primer and future directions. Remote Sensing, 12(12), 1–33.

Kumar, A., & Sinha, N. (2020). Classification of forest cover type using random forests algorithm. In Lecture Notes in Networks and Systems,94, 395–402.

Li, J., Huang, X., Tu, L., Zhang, T., & Wang, L. (2022). A review of building detection from very high-resolution optical remote sensing images. In GIScience and Remote Sensing 59(1), 1199–1225).

Liao, N., Liu, H., Li, C., Ren, X., & Guo, B. (2022). Simple Linear Iterative Clustering with Efficiency. In Smart Innovation, Systems and Technologies, 277, 109–117.

Louhaichi, M., Borman, M. M., & Johnson, D. E. (2001). Spatially located platform and aerial photography for documentation of grazing impacts on wheat. Geocarto International, 16(1), 65–70.

Öztürk, M. Z., & Colkesen, I. (2021). The impacts of vegetation indices from UAV-based RGB imagery on land cover classification using ensemble learning. Mersin Photogrammetry Journal, 3(2), 41–47.

Schlosser, A. D., Szabó, G., Bertalan, L., Varga, Z., Enyedi, P., & Szabó, S. (2020). Building extraction using orthophotos and dense point cloud derived from visual band aerial imagery based on machine learning and segmentation. Remote Sensing, 12(15), 1–28.

Xiao, Y., Huang, W., & Wang, J. (2020). A Random Forest Classification Algorithm Based on Dichotomy Rule Fusion. ICEIEC 2020 - Proceedings of 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication, 182–185.

Zhang, H., & Zhu, Y. (2019). KSLIC: K-mediods clustering based simple linear iterative clustering. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, 1185, 519–529.