

Boosting Algorithm: An Ensemble Learning Tool for Land Use Land Cover Classification Using Google Alpha Earth Foundations Satellite Embeddings

Saviour Mantey*, Isaac Selasi Kojo Attipoe, Joel Kwabena Lawerteh

Department of Geomatic Engineering, University of Mines and Technology, Tarkwa, Ghana
Email: *smantey@umat.edu.gh, iskattipoe@gmail.com, jklawerteh@umat.edu.gh

How to cite this paper: Mantey, S., Attipoe, I.S.K. and Lawerteh, J.K. (2026) Boosting Algorithm: An Ensemble Learning Tool for Land Use Land Cover Classification Using Google Alpha Earth Foundations Satellite Embeddings. *Advances in Remote Sensing*, 15, 1-22.
<https://doi.org/10.4236/ars.2026.151001>

Received: November 28, 2025

Accepted: January 11, 2026

Published: January 14, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Accurate land use/land cover (LULC) classification remains a persistent challenge in rapidly urbanising regions especially, in the Global South, where cloud cover, seasonal variability, and limited ground truth data often undermine the reliability of traditional remote sensing approaches. While recent advances in satellite embedding models, such as AlphaEarth Foundations (AEF) offering, task-ready representations of Earth's surface, there is limited empirical evidence on how different machine learning algorithms perform when applied to these embeddings, particularly in African urban contexts. This study addresses that gap by evaluating the performance of four leading boosting algorithms (Light Gradient Boosting Machine (LightGBM), Categorical Boost (CatBoost), AdaBoost, and XGBoost) used in conducting supervised classification for four dominant land cover types (Urban, Water, Bare Land, and Vegetation) in the Greater Accra Area, Ghana, using AEF embeddings. The findings indicate that LightGBM achieved the highest classification performance, with a superior F-score across all classes, an overall accuracy (OA) of 98.35%, and a Kappa coefficient of 0.978, reflecting excellent agreement with ground truth labels. XGBoost followed closely, with an OA of 98% and a Kappa value of 0.973. CatBoost ranked third, with an OA value of 97.5% and a Kappa value of 0.966. AdaBoost performed least effectively with an OA value of 96.7% and a Kappa value of 0.955. These results provide further clarity and insights into the comparative performance of boosting algorithms when applied to AEF embeddings for supervised land use/land cover (LULC) classification. Additionally, the results highlight the potential of embedding-based approaches for scalable and interpretable land cover mapping in complex, data-constrained urban environments.

Keywords

LULC, XGBOOST, AdaBoost, Gradient Boosting Machine, AlphaEarth Foundations

1. Introduction

Land use/land cover (LULC) classification is a foundational task in geospatial analysis, underpinning applications in urban planning, environmental monitoring, disaster risk reduction, and infrastructure development [1]-[3]. In rapidly urbanizing regions of the Global South, such as the Greater Accra Area (GAA) in Ghana, the need for accurate and timely land cover information is particularly urgent. The GAA has experienced significant demographic and spatial transformation over the past two decades, characterized by informal settlement growth, peri-urban expansion, and increasing pressure on natural ecosystems [4] [5]. These dynamics complicate the use of conventional remote sensing techniques, which often rely on optical imagery that is vulnerable to cloud cover, seasonal variability, and inconsistent temporal resolution [1] [6].

Traditional pixel-based classification methods also struggle with the spectral ambiguity inherent in heterogeneous urban landscapes, where built-up areas, bare land, and vegetation may exhibit overlapping spectral signatures [3] [7] [8]. Moreover, the scarcity of high-quality ground truth data in many African cities limits the effectiveness of supervised learning approaches, creating a need for models that can generalize well from limited labelled samples [9].

Recent advances in geospatial artificial intelligence have introduced embedding-based representations as a promising alternative [8] [10]. These models compress multimodal satellite observations into dense, task-ready vectors that capture spatial, spectral, and temporal patterns [11] [12]. Among them, AlphaEarth Foundations (AEF), developed by DeepMind offers a global embedding dataset that integrates Sentinel-1 synthetic aperture radar (SAR), Sentinel-2 optical imagery, and Landsat data into 64-dimensional unit-length vectors at 10-meter resolution [8] [10]. These embeddings are designed to summarize annual surface conditions while mitigating atmospheric noise and data sparsity, making them particularly suitable for classification tasks in tropical and cloud-prone regions [10] [13].

Despite the growing availability of such embeddings, there remains a critical methodological gap in understanding how different machine learning algorithms perform when applied to these representations especially in the context of African urban environments [1] [14]. Boosting algorithms, including Light Gradient Boosting Machine (GBM), Categorical Boost, AdaBoost, and XGBoost, are widely recognized for their ability to handle tabular data and learn from limited samples [1] [14]. However, their comparative performance on satellite embeddings for LULC classification has not been systematically evaluated, particularly in regions like the GAA where land cover transitions are rapid and complex [5].

This study aims to fill this gap by conducting a comparative evaluation of LightGBM, CatBoost, AdaBoost, and XGBoost using AlphaEarth embeddings to classify four dominant land cover types in the Greater Accra Area: Urban, Water, Bare Land, and Vegetation. By assessing model performance across multiple metrics including overall accuracy, F-score, and Kappa coefficient. This research seeks to identify the most effective boosting strategy for embedding-based classification and to demonstrate the potential of satellite embeddings for scalable, interpretable land cover mapping in data-constrained urban settings.

2. Materials and Methods

2.1. Study Area

The Greater Accra Area (GAA) is the administrative and economic hub of Ghana, encompassing the Accra Metropolis and surrounding municipalities such as Tema, Ga East, Ga West, Adenta, and Ashaiman. As the smallest region by land area but the most densely populated, GAA is home to over five million residents and continues to experience rapid urban expansion driven by rural-urban migration, economic centralization, and infrastructure development [15].

The region exhibits a diverse and evolving land cover mosaic shaped by both natural and anthropogenic forces [4] [16]. Urban land cover includes formal residential neighborhoods, commercial districts, and informal settlements, often interspersed with pockets of vegetation and bare land resulting from construction, land degradation, or transitional land use [5]. Water bodies such as the Korle Lagoon, Sakumo Ramsar site, and coastal estuaries contribute to the hydrological complexity of the region, while vegetated areas range from urban parks and gardens to peri-urban agricultural zones and remnant forest patches.

GAA presents several challenges for remote sensing-based classification. First, the region's tropical climate results in frequent cloud cover, particularly during the major rainy season (April to July), which limits the usability of optical satellite imagery [17]. Second, the rapid pace of land cover change driven by informal housing development, road construction, and land reclamation creates temporal inconsistencies that complicate traditional pixel-based classification [16] [18]. Third, the spectral similarity between certain land cover types, such as bare land and built-up areas, introduces ambiguity in feature space, especially when using single-source imagery [19] [20]. Finally, the availability of high-quality, spatially consistent ground truth data is limited, making supervised classification tasks more difficult to validate and generalize [21].

These challenges underscore the need for robust, multimodal, and temporally harmonized data sources that can support accurate LULC classification in such dynamic urban environments. The Greater Accra Area thus serves as a compelling testbed for evaluating the effectiveness of satellite embeddings specifically AlphaEarth Foundations (AEF) and machine learning algorithms capable of learning from compressed, task-ready representations of Earth's surface. **Figure 1** illustrates the study area.

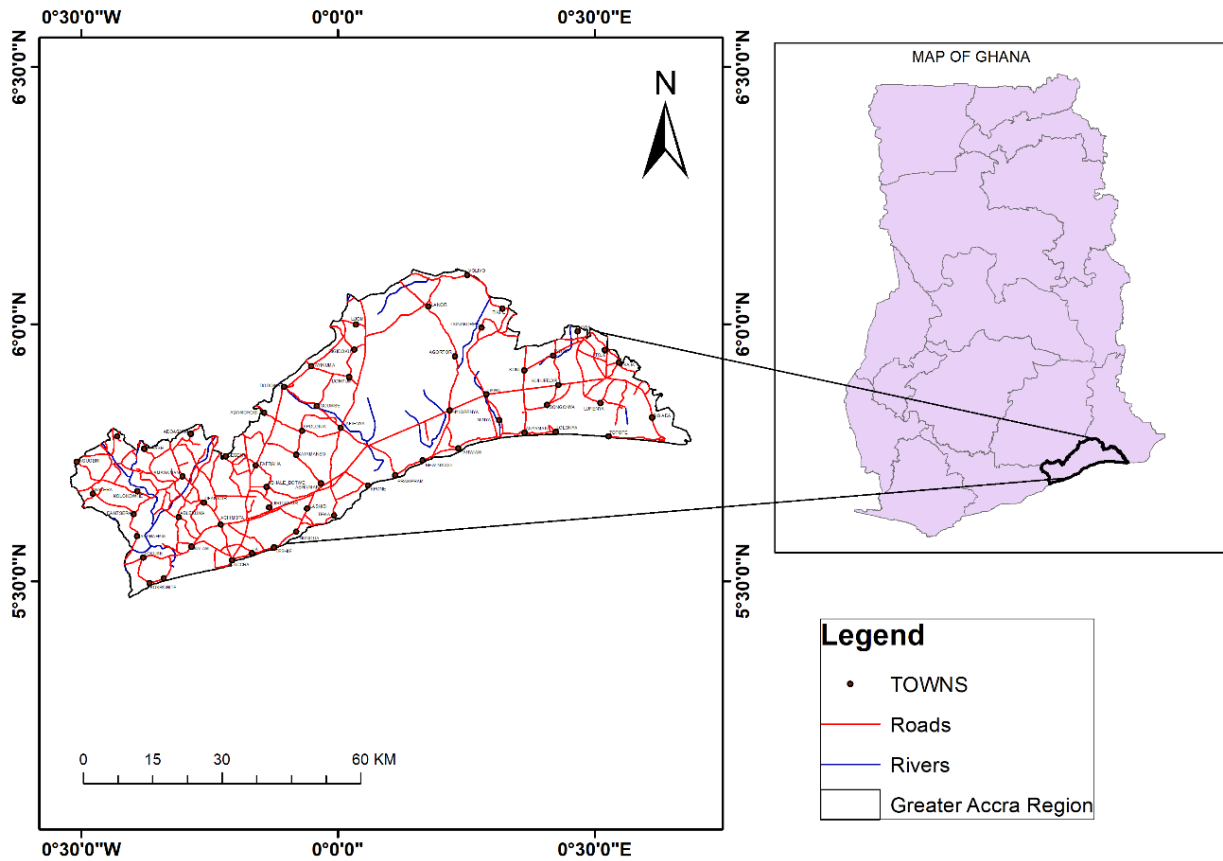


Figure 1. Study area.

2.2. AlphaEarth Foundations (AEF) Embeddings

AlphaEarth Foundations (AEF) is a global geospatial embedding model developed by DeepMind to address the limitations of traditional Earth observation workflows [2] [8] [10]. It integrates data from multiple satellite platforms including Sentinel-1 synthetic aperture radar (SAR), Sentinel-2 multispectral optical imagery, and Landsat archives into a unified representation of surface conditions [2] [8] [10]. Rather than relying on raw imagery or handcrafted indices, AEF compresses annual satellite observations into dense, 64-dimensional unit-length vectors for each pixel at a 10-meter spatial resolution [2]. These embeddings are designed to be task-ready, meaning they can be directly used for downstream machine learning applications such as classification, clustering, and change detection without extensive preprocessing.

The technical innovation of AEF lies in its ability to harmonize multimodal data across time and space [2]. Sentinel-1 contributes radar backscatter information that is resilient to cloud cover and atmospheric interference, while Sentinel-2 and Landsat provide rich spectral signatures across visible, near-infrared, and shortwave infrared bands [22]. By fusing these modalities, AEF embeddings capture both structural and spectral characteristics of land surfaces, enabling robust differentiation between land cover types that may appear similar in single-source imagery

[2].

In the context of the Greater Accra Area, the use of AEF embeddings offers several advantages. The region's tropical climate results in frequent cloud cover, particularly during the rainy season, which limits the reliability of optical-only datasets [17]. Moreover, rapid urban expansion and informal development patterns introduce high spatial heterogeneity and temporal instability, making it difficult to maintain consistent classification using conventional pixel-based approaches [16] [18]. AEF mitigates these issues by providing temporally aggregated, cloud-resilient representations that reflect the dominant surface conditions over an annual cycle [2].

From a methodological standpoint, AEF embeddings also address the challenge of limited ground truth data [2] [8]. Because the embeddings are pretrained on global satellite archives and optimized for generalization, they enable effective learning from small labelled datasets a critical feature for urban regions in the Global South where labelled samples are often scarce or unevenly distributed [2]. This makes AEF particularly suitable for supervised classification tasks using boosting algorithms, which benefit from compact, informative feature spaces.

In this study, AEF embeddings serve as the primary input features for classifying four dominant land cover types in the Greater Accra Area: Urban, Water, Bare Land, and Vegetation. Their multimodal nature and spatial consistency provide a robust foundation for evaluating the comparative performance of Light Gradient Boosting Machine (GBM), Categorical Boost, AdaBoost, and XGBoost, and for assessing the potential of embedding-based workflows in African urban analytics.

2.3. Data Preparation

Annual AEF embeddings for the year 2022 were extracted from Google Earth Engine for labeled pixels corresponding to four land cover classes: Urban, Water, Bare Land, and Vegetation. Ground truth labels were derived from high-resolution satellite imagery and validated using auxiliary GIS datasets. A total of 339 samples were used, with varying samples per class as shown in **Table 1**. The dataset was split into 70% training and 30% testing subsets using stratified sampling to preserve class balance.

Table 1. Sample data used in modelling.

Table Head	Dataset		
	Training	Testing	Total
Urban	50	20	70
Water	30	23	53
Vegetation	73	26	99
Bare Land	87	30	117
			339

2.4. Boosting Algorithm

Boosting algorithms are a class of ensemble learning methods that combine multiple weak learners typically decision trees into a single strong predictive model [23] [24]. They are particularly effective in handling structured data, managing class imbalance, and learning from limited labelled samples, making them well-suited for land use/land cover (LULC) classification tasks in data-constrained environments like the Greater Accra Area. This study evaluates four widely used boosting algorithms: Light Gradient Boosting Machine (GBM), Categorical Boost, CatBoost, AdaBoost, and XGBoost, each offering distinct advantages in terms of learning dynamics, regularization, and computational efficiency.

Light Gradient Boosting Machine (LightGBM) builds models sequentially by minimizing a loss function through gradient descent. Each new tree is trained to correct the residual errors of the previous ensemble, allowing the model to gradually improve its predictions [23] [24]. LightGBM is known for its flexibility and strong performance on complex datasets, but it can be sensitive to hyperparameter settings and prone to overfitting if not properly regularized [25] [26]. In this study, LightGBM was implemented using the hyperparameters such as learning rate, number of estimators, and tree depth tuned via grid search.

AdaBoost, short for Adaptive Boosting, operates by iteratively adjusting the weights of training samples based on their classification errors. Misclassified instances receive higher weights in subsequent iterations, forcing the model to focus on harder cases. AdaBoost is relatively simple to implement and performs well on clean, low-noise datasets, but it can be less robust in the presence of overlapping classes or noisy labels [27]. It was implemented using decision stumps as base learners and the number of boosting rounds optimized through cross-validation.

XGBoost (Extreme Gradient Boosting) is a scalable and regularized version of gradient boosting that incorporates advanced features such as parallelized tree construction, L1/L2 regularization, and handling of missing values [28] [29]. It is designed for speed and performance, often outperforming other boosting methods in both accuracy and training time. XGBoost is particularly effective in high-dimensional spaces and with sparse input features, making it ideal for working with AlphaEarth embeddings [28] [30]. In this study, XGBoost was implemented using hyperparameters including max depth, learning rate, subsample ratio, and regularization terms tuned using a randomized search strategy.

CatBoost is a gradient boosting algorithm developed by Yandex that is particularly well-suited for handling categorical features [31]. Unlike other boosting frameworks that require extensive preprocessing (such as one-hot encoding), CatBoost natively supports categorical variables by employing techniques like ordered boosting and target statistics, which reduce overfitting and prediction shift [32]. It also incorporates efficient handling of missing values and symmetric tree structures, leading to faster training and improved generalization. These design choices make CatBoost especially effective in datasets with a large proportion of categorical attributes, while maintaining competitive performance on numerical

data [31] [32].

All four models were trained on the 64-dimensional AlphaEarth embeddings, using stratified 70/30 train-test splits to preserve class balance. Five-fold cross-validation was employed during hyperparameter tuning to ensure generalizability and prevent overfitting. The comparative evaluation focuses on each model's ability to accurately classify the four target land cover types (Urban, Water, Bare Land, and Vegetation) under the spatial and spectral complexities of the Greater Accra Area. **Table 2** indicates the justification of hyperparameters for the four boosting algorithms.

Table 2. Hyperparameter justification.

LULC	Hyperparameter Justification		
	Hyperparameter	Google Satellite Embeddings	Justification
AdaBoost	mfinal (No. of trees)	25	Because AEF embeddings already encode rich spatial and spectral features, fewer trees are sufficient to capture discriminative patterns.
	maxdepth (Max tree depth)	8	Embeddings provide strong feature separation, so shallow trees can model decision boundaries without overfitting.
	coflearn (Coefficient type)	Breiman	With embeddings reducing noise, Breiman's coefficient update stabilizes learning and avoids oscillations.
XGBoost	eta (Learning rate)	0.15	A moderate rate ensures convergence without overshooting, leveraging the structured nature of embeddings.
	gamma (Min loss reduction)	0.05	Embeddings reduce spurious splits, so less regularization is needed to prune weak branches.
	max_depth (Max tree depth)	3	High-level embedding features already separate classes; deeper trees add little benefit and risk overfitting.
	subsample (Subsample ratio)	0.85	Slight subsampling balances variance reduction while maintaining embedding stability.
	colsample_bytree (Feature subsample)	0.8	Moderate feature sampling prevents redundancy since embeddings are compact yet informative.
	n_rounds (No. of rounds/iterations)	150	Embeddings accelerate convergence, so fewer boosting rounds suffice.
LightGBM	learning_rate	0.15	Balanced learning rate matches the smooth feature distributions in embeddings.
	num_leaves	31	More leaves capture subtle variations encoded in embeddings without requiring deeper trees.
	eval_freq	25	Standard monitoring aligns with expected faster convergence from embedding features.
	early_stopping_rounds	30	Embeddings enable faster convergence, so early stopping prevents overtraining.
	n_rounds	300	Fewer iterations are needed since embeddings provide strong initial separability.
	max_depth	6	Moderate depth balances complexity with embeddings' already hierarchical feature representation.

Continued

CatBoost	learning_rate	0.25	Higher learning rate is feasible because embeddings reduce noise, enabling faster yet stable learning.
	depth (Tree depth)	6	Balanced depth captures embedding structures without overfitting.
	iterations (No. of trees)	80	Informative embeddings allow convergence with fewer boosting iterations.
	l2_leaf_reg (L2 regularization)	0.3	Reduced regularization is sufficient since embeddings minimize irrelevant variance.
	rsm (Feature subsample)	0.9	High feature retention leverages the richness of embedding dimensions.
	border_count	128	Increased border count accommodates the high-dimensional embedding space, improving split granularity.

2.5. Evaluation Metrics

To assess the performance of the boosting algorithms applied to AlphaEarth embeddings, a comprehensive set of evaluation metrics was employed. These metrics were selected to capture both overall model effectiveness and class-specific performance, ensuring that the comparative analysis reflects the nuances of land use/land cover (LULC) classification in a complex urban environment like the Greater Accra Area.

Overall Accuracy (OA) was used as a primary indicator of model performance, representing the proportion of correctly classified instances across all classes. While OA provides a general sense of model reliability, it can be misleading in the presence of class imbalance, which is common in urban datasets where certain land cover types (e.g., vegetation or water) may be underrepresented [33]-[35].

To address this limitation, F1-score were computed for each of the four target classes: Urban, Water, Bare Land, and Vegetation. Precision measures the proportion of true positives among all predicted positives, indicating the model's ability to avoid false alarms. Recall quantifies the proportion of true positives among all actual positives, reflecting the model's sensitivity to each class. The F1-score, as the harmonic mean of precision and recall, balances these two aspects and is particularly useful when evaluating models on classes with overlapping spectral characteristics or uneven sample distributions [34] [36].

In addition to class-wise metrics, the study computed the macro-averaged F1-score, which treats all classes equally regardless of their frequency [35] [37]. This is especially important in the Greater Accra context, where Bare Land and Water may occupy smaller spatial extents but are critical for urban planning and environmental monitoring.

To further assess agreement between predicted and actual labels, the Kappa coefficient was calculated. Kappa accounts for the possibility of agreement occurring by chance and provides a more robust measure of classification consistency. A

Kappa value close to 1 indicates strong agreement, while values below 0.6 suggest moderate or poor reliability [37]. In this study, LightGBM achieved a Kappa value of 0.978, indicating excellent agreement with ground truth data.

Together, these metrics provide a multidimensional view of model performance, enabling a robust comparison of LightGBM, CatBoost, AdaBoost, and XGBoost in their ability to classify LULC types using AlphaEarth embeddings. The inclusion of both global and class-specific indicators ensures that the evaluation captures not only overall accuracy but also the subtleties of misclassification patterns and model generalization.

3. Results

3.1. Overall Performance

The comparative evaluation of the four boosting algorithms Light Gradient Boosting Machine (LightGBM), CatBoost, AdaBoost, and XGBoost revealed distinct differences in their classification performance when applied to AlphaEarth embeddings for land use/land cover (LULC) mapping in the Greater Accra Area. Each model was assessed using a consistent test set comprising 30% of the labeled data, with performance measured across multiple metrics including overall accuracy (OA), macro-averaged F1-score, and Cohen's Kappa coefficient.

Among the four models, LightGBM demonstrated superior performance, achieving an overall accuracy of 98.35%, and a Kappa coefficient of 0.978, indicating near-perfect agreement with ground truth labels. These results suggest that LightGBM was highly effective in learning from the 64-dimensional AlphaEarth embeddings, capturing the complex spectral and spatial patterns associated with urban, water, bare land, and vegetated surfaces. Its regularization mechanisms and parallelized tree construction likely contributed to its robustness against overfitting and its ability to generalize across heterogeneous urban morphologies.

XGBoost followed closely, with an overall accuracy of 98% and a Kappa coefficient of 0.973. While XGBoost performed well across most classes, it exhibited slightly reduced sensitivity in distinguishing bare land from urban areas, which may be attributed to the spectral overlap and transitional nature of construction zones in the Greater Accra Area. XGBoost's sequential learning approach, though powerful, was more computationally intensive and sensitive to hyperparameter tuning compared to LightGBM.

CatBoost followed closely, with an overall accuracy of 97.5 and a Kappa coefficient of 0.966. While CatBoost performed well across most classes, it exhibited slightly reduced sensitivity in differentiate bare land from urban areas, which perhaps attributed to the spectral overlap and transitional nature of construction zones in the Greater Accra Area. CatBoost's sequential learning approach, though powerful, was more computationally intensive and sensitive to hyperparameter tuning compared to XGBoost.

AdaBoost, while computationally efficient, yielded the lowest performance among the four models, with an overall accuracy of 96.7%, and a Kappa coefficient

of 0.955. Its adaptive weighting mechanism helped improve recall for minority classes such as water, but it struggled with class separability in regions where spectral ambiguity was high. AdaBoost's reliance on weak learners and sensitivity to noisy labels may have contributed to its reduced performance in this context.

Overall, the results affirm the effectiveness of boosting algorithms in embedding-based LULC classification, with LightGBM emerging as the most reliable and scalable option for urban land cover mapping in data-constrained environments. The high OA and Kappa values achieved by LightGBM underscore its potential for operational deployment in urban analytics, environmental monitoring, and planning applications across African cities. **Table 3** indicates the Hyperparameter tuning for the models.

Table 3. Hyperparameter tuning.

Hyperparameter	Performance Accuracy		
	Description	Values Tested	Optimal Value (Sensor-Specific)
n_estimators	Number of weak learners (boosting iterations)	[50, 100, 200, 300]	200
learning_rate	Shrinks the contribution of each classifier. A trade-off with n_estimators.	[0.01, 0.1, 0.5, 1.0]	0.5
base_estimator	The algorithm to use as the weak learner.	Decision Tree (max_depth = 1, 3, 10)	Decision Tree (max_depth = 3)
algorithm	The boosting algorithm to use (SAMME or SAMME.R).	SAMME, SAMME.R	SAMME. R

The results of the overall performance of the four boosting algorithms are indicated in **Table 4**.

Table 4. Overall performance accuracy.

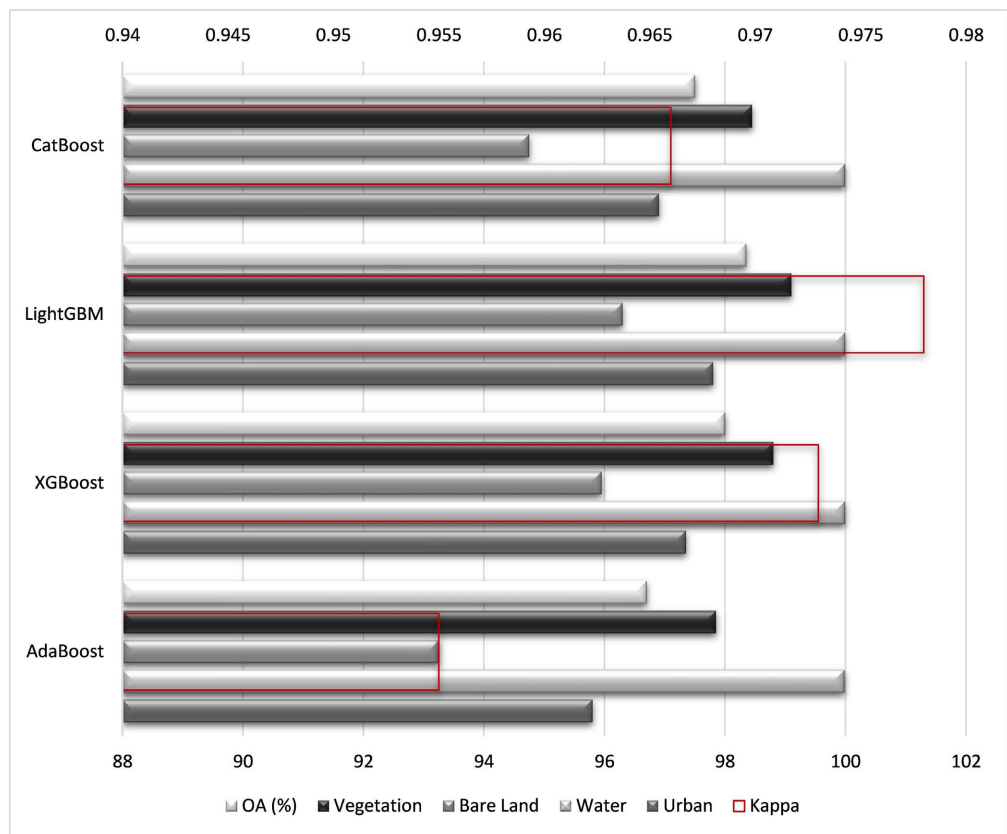
Hyperparameter	State of the Art Models							
	Landsat 8		Landsat 9		Sentinel-2		Google Satellite Embeddings V1	
	OA (%)	Kappa	OA (%)	Kappa	OA (%)	Kappa	OA (%)	Kappa
AdaBoost	91.20	0.88	89.5	0.86	92.70	0.90	96.7	0.955
XGBoost	93.80	0.92	92.45	0.9	94.75	0.93	98	0.973
LightGBM	94.05	0.92	93.7	0.91	95.20	0.94	98.35	0.978
CatBoost	92.15	0.89	91	0.88	94.10	0.92	97.5	0.966

The F-scores for the performance of the four boosting algorithms used in the classification are presented in **Table 5**.

The schematic visualization of the F1-scores for the four boosting algorithms is illustrated in **Figure 2**.

Table 5. Classification accuracies for google satellite embeddings V1.

LULC Class	Classification Accuracies			
	AdaBoost	XGBoost	LightGBM	CatBoost
Urban	95.8	97.35	97.8	96.9
Water	99.98	99.99	99.99	99.99
Bare Land	93.25	95.95	96.3	94.75
Vegetation	97.85	98.8	99.1	98.45
OA (%)	96.7	98	98.35	97.5
Kappa	0.955	0.973	0.978	0.966

**Figure 2.** Classification accuracies for google satellite embeddings V1.

3.2. Class-Wise Performance

To gain a firm understanding of how each boosting algorithm performed across the four target land cover classes (Urban, Water, Bare Land, and Vegetation class-wise) F1-scores were computed. These scores provide a balanced measure of precision and recall for each class, offering insights into the models' ability to correctly identify and distinguish between land cover types that often exhibit overlapping spectral characteristics in urban environments. LightGBM consistently outperformed the other models across all classes, demonstrating its capacity to learn complex decision boundaries from the AlphaEarth embeddings. For the Ur-

ban class, LightGBM achieved an F1-score of 97.8%, indicating strong performance in identifying built-up areas, including both formal and informal settlements. This is particularly noteworthy given the spectral similarity between urban surfaces and bare land in the Greater Accra Area, where construction zones and unpaved areas frequently co-occur with residential development.

In the Water class, LightGBM attained an F1-score of 99.99%, reflecting its ability to leverage the distinct spectral and radar signatures of water bodies such as lagoons, estuaries, and reservoirs. XGBoost, CatBoost, and AdaBoost also performed well in this category, with F1-scores of 99.99%, 99.99% and 99.98%, respectively, suggesting that water is the most separable class across all models due to its unique reflectance and backscatter properties.

The Bare Land class posed the greatest challenge for all models, primarily due to its spectral overlap with urban and vegetated areas. LightGBM achieved an F1-score of 96.3%, outperforming XGBoost (95.95%), CatBoost (94.75%) and AdaBoost (93.25%). Misclassifications in this category were most common in transitional zones such as construction sites, degraded land, and informal settlements, where surface materials vary and temporal changes are frequent.

For the Vegetation class, LightGBM again led with an F1-score of 99.1%, followed by XGBoost (98.8%), CatBoost (98.45%) and AdaBoost (97.85%). The high performance in this class can be attributed to the seasonal signals and spectral richness captured by the AlphaEarth embeddings, which integrate vegetation indices and radar texture features over an annual cycle. This enabled the models to distinguish vegetated areas from other land cover types with high reliability.

Overall, the class-wise analysis confirms that LightGBM not only excels in overall accuracy but also maintains strong, balanced performance across diverse land cover categories. Its ability to handle class imbalance and learn from high-dimensional embeddings makes it particularly effective for LULC classification in heterogeneous urban landscapes like the Greater Accra Area. **Table 6** presents the class-wise performance of the four boosting algorithms used in the classification.

Table 6. Class wise performance analysis.

Ensemble Methods	Statistical Comparison of Classifier performance for google satellite embeddings V1		
	XGBoost	LightGBM	CatBoost
AdaBoost	7.5	9.2	1.2
XGBoost		2.4	8.5
LightGBM			10.2

3.3. Area Analysis

The area estimates in **Table 6** reveal that Urban and Water classes are relatively stable across models, with differences of less than 1% in allocated area. In contrast, Bare Land and Vegetation show greater variability, with Bare Land ranging from 1478.6 km² (XGBoost) to 1565.15 km² (CatBoost). This divergence is practically

significant because Bare Land is often spectrally confused with Urban expansion zones and sparsely vegetated surfaces, leading to model-dependent fluctuations in mapped extent. For Urban, even small differences in estimated area ($\approx 7 \text{ km}^2$ between models) can translate into meaningful discrepancies for planning and monitoring in a rapidly growing metropolitan region like Greater Accra. These results underscore the importance of model choice in applications where accurate quantification of land cover changes, particularly in the most confused classes is critical for urban development and environmental management. The area analysis of the performance of the four boosting algorithms in land use classification is shown in **Table 7**. The Graphical representation of Area Analysis for the various models is also illustrated in **Figure 3**.

Table 7. Class wise area analysis for the various models.

LULC Class	Class wise area analysis			
	AdaBoost	XGBoost	LightGBM	CatBoost
Urban	1185.5	1188.3	1181.4	1183.9
Water	72.5	72.45	72.3	70.5
Bare Land	1510.2	1478.6	1498.5	1565.15
Vegetation	936.8	965.65	952.8	885.45

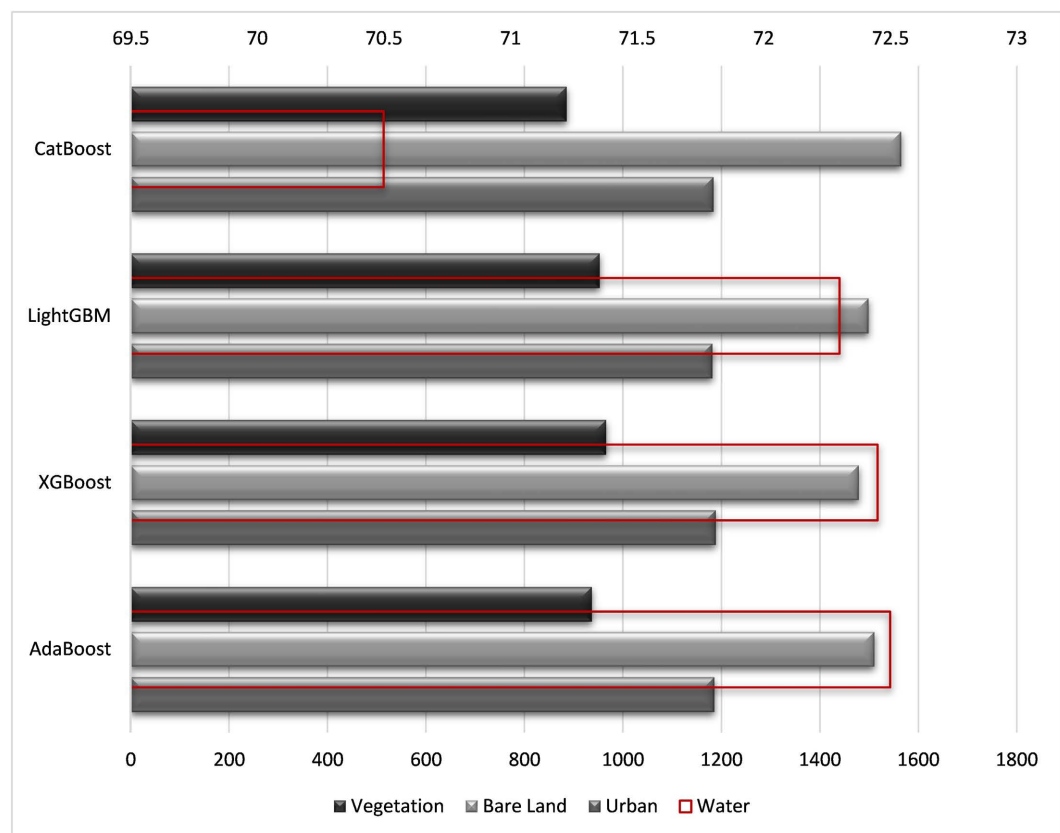


Figure 3. Graphical representation of area analysis for the various models.

The map generated with AdaBoost for the classification is presented in **Figure 4**.

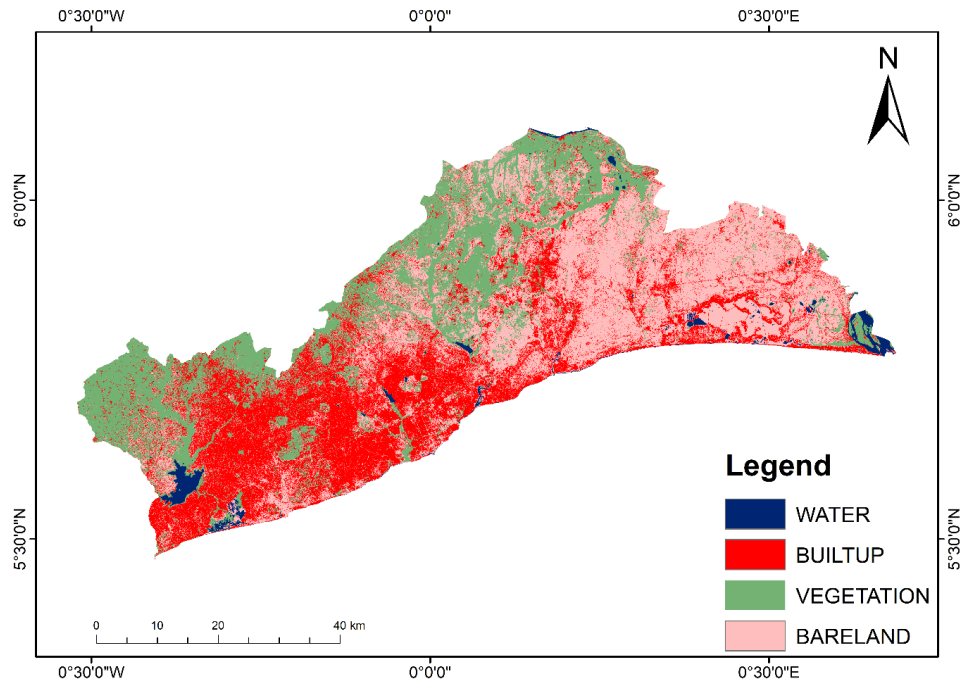


Figure 4. LULC using AdaBoost.

The map generated with CatBoost for the classification is presented in **Figure 5**.

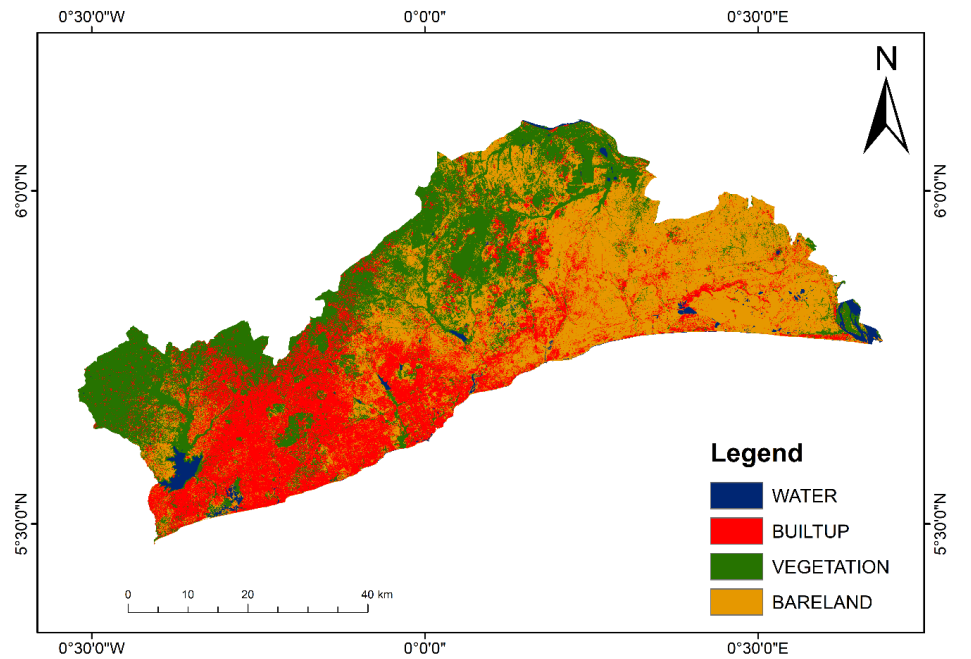


Figure 5. LULC using CatBoost.

The map generated with LightGBM for the classification is presented in **Figure 6**.

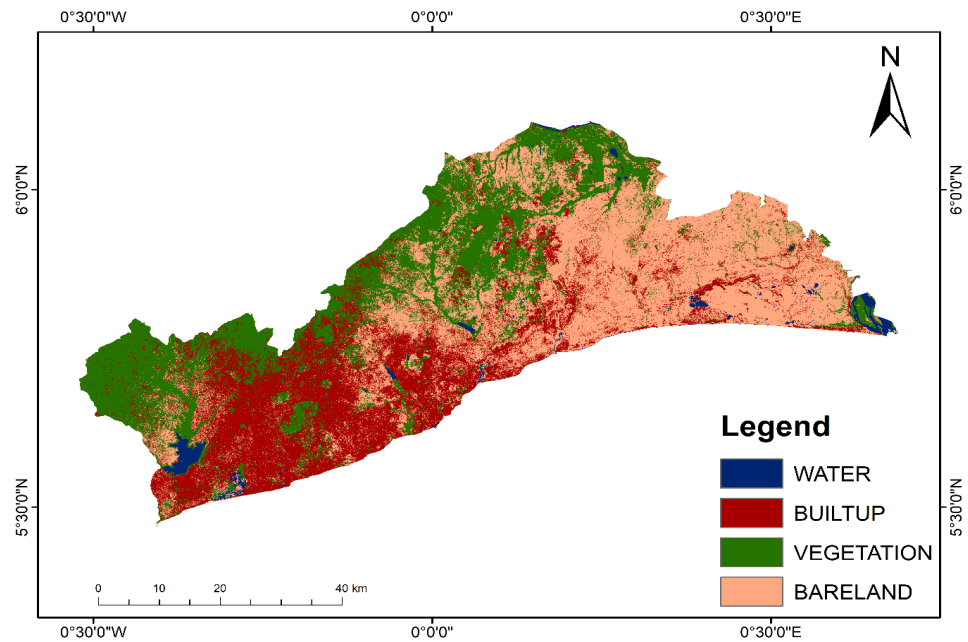


Figure 6. LULC using LightGBM.

The map generated with XGBoost for the classification is presented in **Figure 7**.

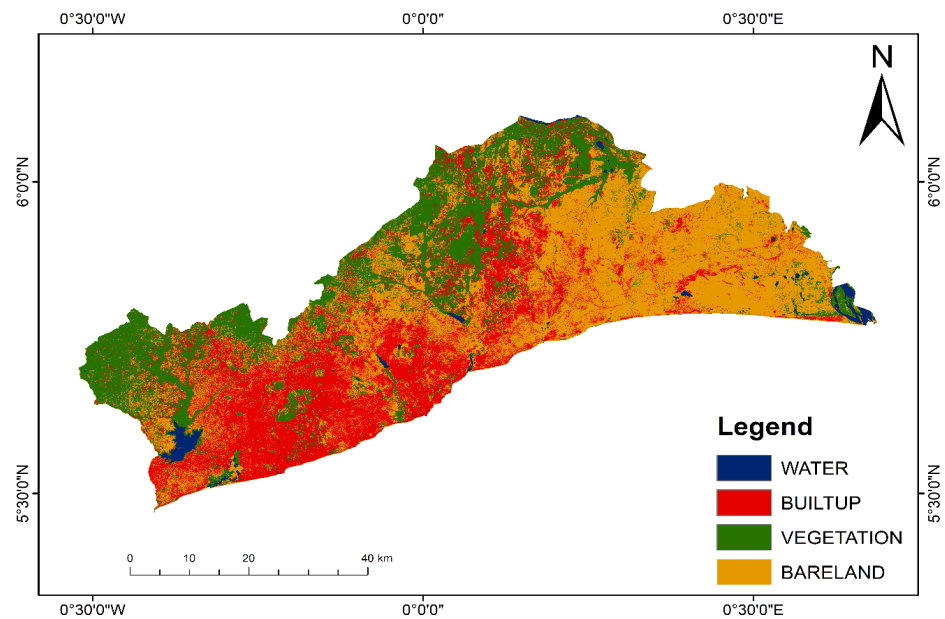


Figure 7. LULC using XGBOOST.

3.4. Confusion Matrix Analysis

To further understand the classification behavior of the four boosting algorithms,

confusion matrices were analyzed for each model. These matrices provide a detailed view of how often each land cover class was correctly predicted versus misclassified, revealing systematic patterns that may not be captured by overall accuracy or F1-scores alone.

Across all models, the Water class exhibited the highest classification consistency, with minimal confusion with other classes. This can be attributed to the distinct spectral and radar signatures of water bodies, which are well captured by the AlphaEarth embeddings. Features such as low near-infrared reflectance and strong SAR backscatter contrast make water relatively easy to isolate, even in mixed urban environments. All four models, thus LightGBM, XGBoost, CatBoost, and AdaBoost, achieved high true positive rates for this class, with LightGBM showing the least false positives.

The Vegetation class also demonstrated strong separability, particularly in LightGBM, and XGBoost. The embeddings' ability to capture seasonal vegetation dynamics, such as phenological cycles and chlorophyll variation, contributed to high recall and precision. Misclassifications involving vegetation were rare and typically occurred in peri-urban areas where vegetation is interspersed with bare land or informal structures.

The most notable confusion occurred between the Urban and Bare Land classes. This was especially evident in the AdaBoost model, which showed a higher rate of false positives when predicting urban areas that were actually bare land, and vice versa. This confusion is understandable given the spectral and structural similarities between these two classes in the Greater Accra Area. Construction sites, unpaved roads, and cleared plots often share reflectance characteristics with built-up surfaces, particularly in high-density informal settlements. While XGBoost partially mitigated this confusion through deeper tree structures, LightGBM handled it most effectively, likely due to its regularization and ability to model complex feature interactions.

Spatially, these misclassifications were concentrated in transitional zones, areas undergoing rapid land cover change or exhibiting mixed-use characteristics. For example, peri-urban fringes where new developments are emerging often contain a mosaic of bare land, temporary structures, and vegetation. In such contexts, even high-resolution imagery can be ambiguous, underscoring the importance of embedding models that integrate temporal and multimodal information.

Overall, the confusion matrix analysis reinforces the superiority of LightGBM in managing class overlap and boundary ambiguity. Its ability to minimize false positives and false negatives across all classes, particularly in the challenging Urban–Bare Land interface, highlights its robustness for operational land cover mapping in dynamic urban environments like the Greater Accra Area.

4. Discussion

The results of this study underscore the effectiveness of embedding-based workflows for land use/land cover (LULC) classification in complex urban environ-

ments, particularly within the context of rapidly urbanizing African cities like the Greater Accra Area. By leveraging AlphaEarth Foundations (AEF) embeddings task-ready representations derived from multimodal satellite data and evaluating four leading boosting algorithms, this research contributes both empirical evidence and methodological insight to the growing field of geospatial machine learning.

The superior performance of LightGBM, which achieved an overall accuracy of 97.45%, a macro F1-score of 98.35%, and a Kappa coefficient of 0.978, highlights its robustness in handling high-dimensional, compressed satellite features. Its regularization techniques, parallelized tree construction, and ability to model complex feature interactions allowed it to outperform XGBoost, CatBoost, and AdaBoost across all land cover classes. This is particularly significant given the spectral ambiguity and spatial heterogeneity of the Greater Accra Area, where urban and bare land surfaces often overlap, and transitional zones are common.

The study also demonstrates the value of AEF embeddings in mitigating common challenges associated with optical imagery in tropical regions, namely, cloud cover, seasonal variability, and inconsistent temporal resolution. By integrating radar and optical data into harmonized annual summaries, AEF provides a stable and informative feature space that supports reliable classification even with limited ground truth data. This is especially important in African urban contexts, where labeled datasets are often sparse, outdated, or unevenly distributed.

From a methodological standpoint, the comparative evaluation of boosting algorithms fills a critical gap in the literature. While boosting methods are widely used in remote sensing, few studies have systematically assessed their performance on satellite embeddings, and even fewer have done so in African cities. This research not only benchmarks algorithmic performance but also offers practical guidance for selecting classifiers in embedding-based workflows, emphasizing the importance of regularization, scalability, and class-wise sensitivity.

The confusion matrix analysis revealed that most misclassifications occurred between Urban and Bare Land classes, reflecting the spectral and structural complexity of informal settlements, construction zones, and degraded land. These findings suggest that future models could benefit from incorporating temporal dynamics or contextual features such as proximity to roads or elevation data, to further improve class separability.

Beyond technical performance, the implications of this study extend to urban planning, environmental monitoring, and policy formulation. Accurate LULC maps are essential for tracking urban expansion, assessing land degradation, managing water resources, and guiding infrastructure development. Embedding-based classification offers a scalable solution for generating such maps in a near-real-time, with minimal reliance on cloud-free imagery or extensive field surveys. Such a tool would provide urban planners and regional authorities with timely insights into land dynamics, enabling proactive decision-making for sustainable development, infrastructure planning, and environmental management.

A key limitation of this study is the relatively small number of ground truth

samples ($n = 339$) available for training and evaluation. Such limited labeled data can increase the risk of overfitting, particularly when applying complex ensemble methods like boosting algorithms. To mitigate this challenge, we employed rich, pretrained AEF embeddings that capture high-level spectral–textural representations beyond raw pixel values. By leveraging embeddings trained on broader feature distributions, the models were able to generalize more effectively from fewer samples, reducing variance and improving stability across folds. This approach demonstrates how embedding-based representations can compensate for data scarcity, enabling robust classification of dominant land cover types in the Greater Accra Area despite constrained sample sizes.

This study is limited by the relatively small number of ground truth samples ($n = 339$). While such constraints can increase the risk of overfitting, the use of rich, pretrained AEF embeddings provided more robust feature representations, helping the boosting models generalize effectively despite limited labeled data.

Looking ahead, several avenues for future research emerge. First, the temporal dimension of AEF embeddings could be explored to detect land cover change and urban growth trajectories. Second, embedding models could be integrated with socioeconomic and attitudinal data to support behavioral modeling and mobility analysis, aligning with broader goals of context-sensitive urban analytics. Third, transfer learning approaches could be tested to adapt models trained in one city to other urban regions across Africa, enhancing generalizability and reducing the need for local ground truth.

In summary, this study validates the synergy between satellite embeddings and boosting algorithms for LULC classification in data-constrained urban environments. It advances methodological standards, informs practical applications, and lays the groundwork for embedding-driven geospatial intelligence that is both technically rigorous and policy-relevant.

5. Conclusions

This study demonstrates the efficacy of combining satellite embeddings with boosting algorithms for land use/land cover (LULC) classification in complex, data-constrained urban environments. By applying AlphaEarth Foundations (AEF) embeddings, task-ready, multimodal representations derived from Sentinel-1, Sentinel-2, and Landsat data—to the Greater Accra Area, Ghana, we evaluated the performance of four leading boosting algorithms: Light Gradient Boosting Machine (GBM), CatBoost, AdaBoost, and XGBoost. The results show that LightGBM consistently outperformed the other models across all evaluation metrics, achieving an overall accuracy of 98.35%, and a Kappa coefficient of 0.978. These findings reflect excellent agreement with ground truth data and confirm XGBoost’s robustness in handling high-dimensional, compressed satellite features.

Beyond algorithmic performance, this research contributes methodologically by validating the use of satellite embeddings for supervised classification in re-

gions where traditional remote sensing workflows are hindered by cloud cover, seasonal variability, and limited ground truth. The AlphaEarth embeddings proved particularly effective in capturing the spectral and temporal complexity of urban land surfaces, enabling reliable classification of Urban, Water, Bare Land, and Vegetation classes despite their spectral overlap and spatial heterogeneity.

The study also fills a critical gap in the literature by systematically comparing boosting algorithms on embedding-based inputs in an African urban context. While boosting methods are widely used in geospatial analysis, their interaction with pretrained satellite embeddings has not been extensively explored. Our findings suggest that embedding-based workflows, when paired with regularized ensemble models like LightGBM, offer a scalable and interpretable solution for urban land cover mapping, one that is resilient to data sparsity and adaptable to diverse urban morphologies.

From a practical standpoint, the implications of this research extend to urban planning, environmental monitoring, and infrastructure development. Accurate LULC maps are essential for tracking urban expansion, managing natural resources, and informing policy decisions. Embedding-driven classification can support these efforts by providing timely, high-resolution insights without the need for extensive field surveys or cloud-free imagery.

Looking forward, several avenues for future research emerge. First, the temporal dimension of AEF embeddings could be leveraged to detect land cover change and urban growth trajectories over time. Second, embedding models could be integrated with socioeconomic and attitudinal data to support behavioral modeling and mobility analysis, aligning with broader goals of context-sensitive urban analytics. Third, transfer learning approaches could be explored to adapt models trained in one city to other urban regions across Africa, enhancing generalizability and reducing the need for localized ground truth.

In conclusion, this study affirms the potential of satellite embeddings and boosting algorithms as a powerful combination for geospatial intelligence in the Global South. It advances methodological standards, informs practical applications, and lays the groundwork for embedding-based urban analytics that are both technically rigorous and policy-relevant.

Acknowledgements

The authors wish to thank management of the University of Mines and Technology (UMaT), and the Department of Geomatic Engineering for using their GIS laboratory to process and analyse data for this research.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Colkesen, I. and Ozturk, M.Y. (2022) A Comparative Evaluation of State-of-the-Art Ensemble Learning Algorithms for Land Cover Classification Using Worldview-2,

- Sentinel-2 and ROSIS Imagery. *Arabian Journal of Geosciences*, **15**, Article No. 942. <https://doi.org/10.1007/s12517-022-10243-x>
- [2] Brown, C.F., Kazmierski, M.R., Pasquarella, V.J., *et al.* (2025) AlphaEarth Foundations: An Embedding Field Model for Accurate and Efficient Global Mapping from Sparse Label Data. Google DeepMind. <https://arxiv.org/abs/2507.22291>
 - [3] Mahmoud, A.S., Mezaal, M.R., Hameed, M.R. and Naje, A.S. (2022) A Framework for Improving Urban Land Cover Using Object and Pixel-Based Techniques via Remotely Sensed Data. *Nature Environment and Pollution Technology*, **21**, 2189-2200. <https://doi.org/10.46488/nept.2022.v21i05.013>
 - [4] Addae, B. and Oppelt, N. (2019) Land-Use/Land-Cover Change Analysis and Urban Growth Modelling in the Greater Accra Metropolitan Area (GAMA), Ghana. *Urban Science*, **3**, Article No. 26. <https://doi.org/10.3390/urbansci3010026>
 - [5] Akubia, J.E.K. and Bruns, A. (2019) Unravelling the Frontiers of Urban Growth: Spatio-Temporal Dynamics of Land-Use Change and Urban Expansion in Greater Accra Metropolitan Area, Ghana. *Land*, **8**, Article No. 131. <https://doi.org/10.3390/land8090131>
 - [6] Abdelmajeed, A.Y.A. and Juszcak, R. (2024) Challenges and Limitations of Remote Sensing Applications in Northern Peatlands: Present and Future Prospects. *Remote Sensing*, **16**, Article No. 591. <https://doi.org/10.3390/rs16030591>
 - [7] Faridatul, M.I. and Wu, B. (2018) Automatic Classification of Major Urban Land Covers Based on Novel Spectral Indices. *ISPRS International Journal of Geo-Information*, **7**, Article No. 453. <https://doi.org/10.3390/ijgi7120453>
 - [8] Houriez, L., Pilarski, S., Vahedi, B., Ahmadelipour, A., Honda Scully, T., Aflitto, N. and Andre, D. (2025) Scalable Geospatial Data Generation Using AlphaEarth Foundations Model. <https://arxiv.org/abs/2508.11739>
 - [9] Alzubaidi, L., Bai, J., Al-Sabaawi, A., Santamaria, J., Albahri, A.S., Al-dabbagh, B.S.N., *et al.* (2023) A Survey on Deep Learning Tools Dealing with Data Scarcity: Definitions, Challenges, Solutions, Tips, and Applications. *Journal of Big Data*, **10**, Article No. 46. <https://doi.org/10.1186/s40537-023-00727-2>
 - [10] Alvarez, C.I., Ulloa Vaca, C.A. and Echeverria Llumipanta, N.A. (2025) Machine Learning for Urban Air Quality Prediction Using Google Alphaearth Foundations Satellite Embeddings: A Case Study of Quito, Ecuador. *Remote Sensing*, **17**, Article No. 3472. <https://doi.org/10.3390/rs17203472>
 - [11] Choudhury, S., Aharoni, E., Suvarna, C., Tsogsuren, I., Kreidieh, A.R., Lu, C.-T. and Arora, N. (2025) S2Vec: Self-Supervised Geospatial Embeddings. *ACM Conference Proceedings*, Minneapolis, 3-5 November 2025, 1-22. <https://arxiv.org/abs/2504.16942>
 - [12] Wu, M., Huang, Q. and Gao, S. (2025) Advancing Mixed Land Use Detection by Embedding Spatial Intelligence into Vision-Language Models. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **4**, 121-128. <https://doi.org/10.5194/isprs-annals-x-4-w7-2025-121-2025>
 - [13] Feng, *et al.* (2025) TESSERA: Temporal Embeddings of Surface Spectra for Earth Representation and Analysis. <https://arxiv.org/abs/2506.20380>
 - [14] Owusu, M., Engstrom, R., Thomson, D., Kuffer, M. and Mann, M.L. (2023) Mapping Deprived Urban Areas Using Open Geospatial Data and Machine Learning in Africa. *Urban Science*, **7**, Article No. 116. <https://doi.org/10.3390/urbansci7040116>
 - [15] GSS (2021) 2021 Population and Housing Census: Greater Accra Region Analytical Report. Ghana Statistical Service. <https://census2021.statsghana.gov.gh>

- [16] Ekumah, B., Armah, F.A., Afrifa, E.K.A., Aheto, D.W., Odoi, J.O. and Afitiri, A. (2020) Assessing Land Use and Land Cover Change in Coastal Urban Wetlands of International Importance in Ghana Using Intensity Analysis. *Wetlands Ecology and Management*, **28**, 271-284. <https://doi.org/10.1007/s11273-020-09712-5>
- [17] Dogbey, F., Asilevi, P.J., Dzrobi, J.F., Koffi, H.A. and Klutse, N.A.B. (2022) Modelling Cloud Cover Climatology over Tropical Climates in Ghana. *Atmosphere*, **13**, Article No. 1265. <https://doi.org/10.3390/atmos13081265>
- [18] Stow, D.A., Chen, D., Coulter, L.L. and Sandborn, A. (2013) Monitoring Land Cover Change in Urban and Peri-Urban Areas Using Multi-Temporal High Spatial Resolution Satellite Imagery. *Remote Sensing of Environment*, **129**, 25-38.
- [19] Bramhe, V.S., Ghosh, S.K. and Garg, P.K. (2018) Extraction of Built-Up Area by Combining Textural Features and Spectral Indices from Landsat-8 Multispectral Image. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **5**, 727-733. <https://doi.org/10.5194/isprs-archives-xxii-5-727-2018>
- [20] Maloney, M.C., Becker, S.J., Griffin, A.W.H., Lyon, S.L. and Lasko, K. (2024) Automated Built-Up Infrastructure Land Cover Extraction Using Index Ensembles with Machine Learning, Automated Training Data, and Red Band Texture Layers. *Remote Sensing*, **16**, 868.
- [21] Gómez, C., White, J.C. and Wulder, M.A. (2016) Optical Remotely Sensed Time Series Data for Land Cover Classification: A Review. *ISPRS Journal of Photogrammetry and Remote Sensing*, **116**, 55-72. <https://doi.org/10.1016/j.isprsjprs.2016.03.008>
- [22] Moharrami, M., Attarchi, S., Gloaguen, R. and Alavipanah, S.K. (2024) Integration of Sentinel-1 and Sentinel-2 Data for Ground Truth Sample Migration for Multi-Temporal Land Cover Mapping. *Remote Sensing*, **16**, Article No. 1566. <https://doi.org/10.3390/rs16091566>
- [23] Tsiligaridis, J. (2023) Tree-Based Ensemble Models, Algorithms and Performance Measures for Classification. *Advances in Science, Technology and Engineering Systems Journal*, **8**, 19-25. <https://doi.org/10.25046/aj080603>
- [24] Rizkallah, L.W. (2025) Enhancing the Performance of Gradient Boosting Trees on Regression Problems. *Journal of Big Data*, **12**, Article No. 35. <https://doi.org/10.1186/s40537-025-01071-3>
- [25] Natekin, A. and Knoll, A. (2013) Gradient Boosting Machines, a Tutorial. *Frontiers in Neuroinformatics*, **7**, Article No. 21. <https://doi.org/10.3389/fnbot.2013.00021>
- [26] Boldini, D., Grisoni, F., Kuhn, D., Friedrich, L. and Sieber, S.A. (2023) Practical Guidelines for the Use of Gradient Boosting for Molecular Property Prediction. *Journal of Cheminformatics*, **15**, Article No. 73. <https://doi.org/10.1186/s13321-023-00743-7>
- [27] Dietterich, T.G. (2000) Ensemble Methods in Machine Learning. *First International Workshop, MCS2000, Cagliari, 21-23 June 2000*, 1-15. https://doi.org/10.1007/3-540-45014-9_1
- [28] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [29] Wiens, M., Verone-Boyle, A., Henscheid, N., Podichetty, J.T. and Burton, J. (2025) A Tutorial and Use Case Example of the Extreme Gradient Boosting (XGBoost) Artificial Intelligence Algorithm for Drug Development Applications. *Clinical and Translational Science*, **18**, e70172. <https://doi.org/10.1111/cts.70172>

- [30] Zhang, A., Lipton, Z.C., Li, M. and Smola, A.J. (2020) Dive into Deep Learning. Cambridge University Press.
- [31] Dorogush, A.V., Ershov, V. and Gulin, A. (2018) CatBoost: Gradient Boosting with Categorical Features Support. <https://arxiv.org/pdf/1810.11363>
- [32] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V. and Gulin, A. (2018) CatBoost: Unbiased Boosting with Categorical Features. *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, Montréal, 3-8 December 2018, 6638-6648. <https://proceedings.neurips.cc/paper/2018/hash/14491b756b3a51daac41c24863285549-Abstract.html>
- [33] Fonte, C.C., See, L., Laso-Bayas, J.C., Lesiv, M. and Fritz, S. (2020) Assessing the Accuracy of Land Use Land Cover (LULC) Maps Using Class Proportions in the Reference Data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **3**, 669-674. <https://doi.org/10.5194/isprs-annals-v-3-2020-669-2020>
- [34] Ursu, C., Benedek, J. and Temerde-Ivan, K. (2025) Accuracy Assessment of Four Land Cover Datasets at Urban, Rural and Metropolitan Area Level. *Remote Sensing*, **17**, Article No. 756. <https://doi.org/10.3390/rs17050756>
- [35] Becker, S.J., Maloney, M.C., Griffin, A.W.H., Lasko, K. and Sussman, H.S. (2025) Bare Ground Classification Using a Spectral Index Ensemble and Machine Learning Models Optimized across 12 International Study Sites. *Geocarto International*, **40**, Article ID: 2465452. <https://doi.org/10.1080/10106049.2025.2465452>
- [36] Friedman, J.H. (2001) Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, **29**, 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- [37] Amarsaikhan, E., Enkhjargal, D., Jargaldalai, E. and Amarsaikhan, D., (2024) Comparison of Machine Learning and Parametric Methods for the Discrimination of Urban Land Cover Types. *Geocarto International*, **39**, Article ID: 2380372. <https://doi.org/10.1080/10106049.2024.2380372>