Capture and Classification of Roofs for Building Characterization from High – Resolution Images

Luisa Fernanda Rodríguez Ocaño, Alexander Páez Lancheros and Johan Andrés Avendaño Arias

Paper 12653

Instituto Geográfico Agustín Codazzi (IGAC) of Colombia

May 2024

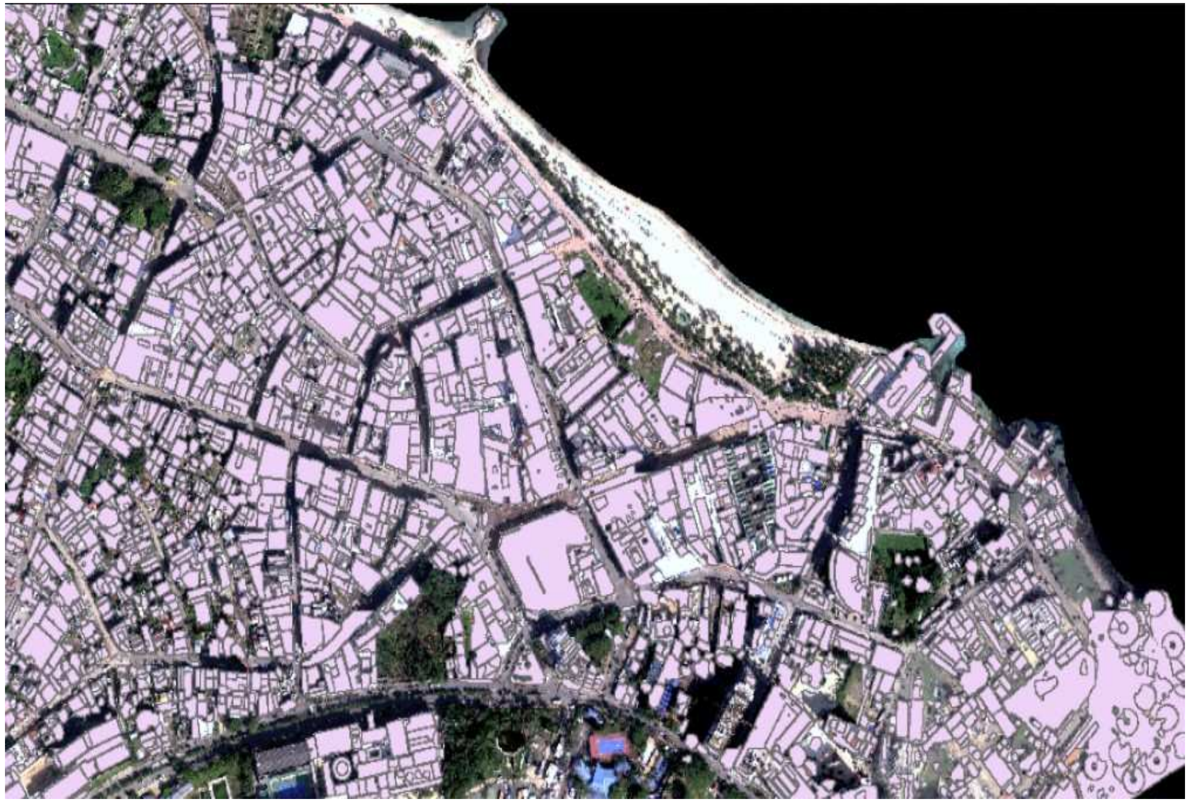
Introduction

Updated cadastral information is crucial for strengthening governance in territory administration. Various methods, including direct, indirect, declarative, and collaborative approaches, can be employed to gather data on the physical, legal, and economic components of properties within a territory. In direct methods, information is collected through on-site visits and identification of each property, involving slow and costly processes.

On the other hand, indirect methods allow for obtaining information about property attributes through secondary sources, particularly remote sensing images that offer increasingly higher spatial and temporal resolution (IGAC, 2023).

The value of properties is a key component of cadastral information for enhancing municipal finances, as it forms the basis for property taxes. The property value is determined through various methods, some of which separately assess the value of the land and buildings (IGAC, 2017).

Key Words: Orthoimage, Classify Roofs, Segmentation, Machine Learning Detect Constructions, Satellite Images, SAM.



Materials

Data: High-resolution orthoimages from urban areas of three municipalities in Colombia were selected, representing diverse conditions in construction characteristics. The features of these orthoimages, obtained from the database on the Colombia en Mapas platform (IGAC, 2023), are described in Table 1.

MUNICIPALITY	YEAR	AREA (Ha)	SPATIAL RESOLUTION	SENSOR
SAN ANDRES	2020	2845,4253	10cm	sensor ADS 120
QUIMBAYA	2018	321,81	15cm	sensor WILD RC-30
CARMEN DE BOLIVAR	2021	600.72	10cm	sensor S.O.D.A.

The data was standardized to three bands (RGB), with the resolution also adjusted to 20 cm. This decision was made considering that the SAM algorithm requires three bands for its operation. Additionally, it was considered that the IGAC establishes the standard for cartographic inputs in RGB format. The objective of this choice is to achieve experiment replicability, allowing the use of available inputs in the institute.

Development Environment: For the development of this research, Python programming language version 3.8, ARCGIS 10.8.1, and a computer with 16 GB RAM, an AMD Ryzen 5 3600G processor at 3.60 GHz, and a 64-bit operating system were utilized.

Method(s)

The first involved selecting the study area and standardizing the orthoimages with resampling to a resolution of 20 cm.

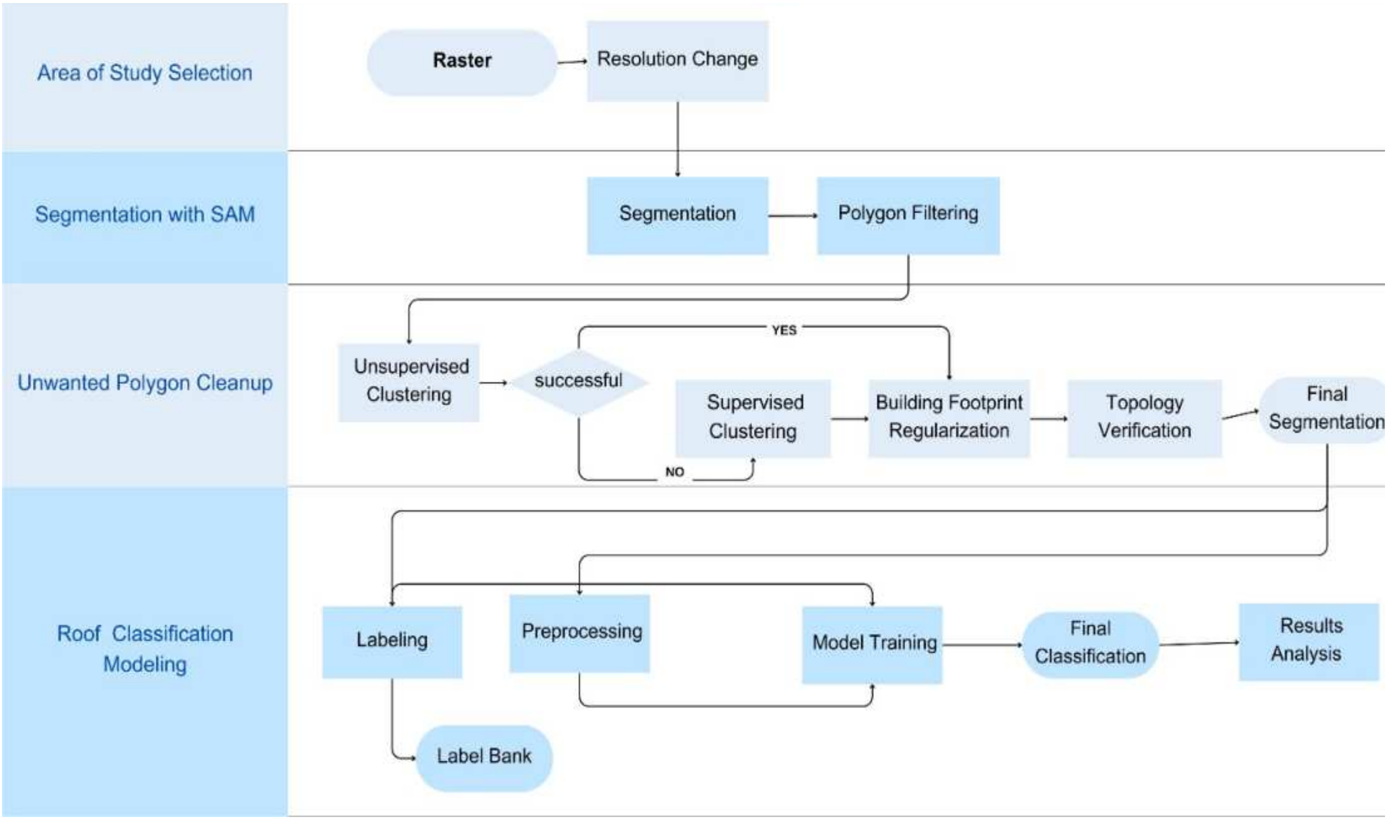
In the second phase, segmentation was carried out using the SAM.

For the next phase, unwanted polygons were removed, resulting in a final segmentation of the roofs.

In the final phase, pixel-level modeling based on *Random Forest* was implemented to classify roofs based on construction material.

Specific models for each municipality were then trained using processed images and generated polygons, with the dataset divided into 70% for training and 30% for testing.

The parameters used were a maximum of 50 trees, a maximum depth of 30, and a maximum of 3000 examples per class.



Results (Quimbaya)

In the municipality of Quimbaya, the following labels were used: zinc, amianto, clay, imitation clay, and oxidized zinc.
The obtained metrics on the test sets are presented below.

Metrics on Quimbaya Test Sets.

Roof	Original			Processing 1			Processing 2		
	acuraccy			acuraccy			acuraccy		
	0.77			0.78			0.80		
	Precision	Recall	f1-score	Precision	Recall	f1-score	Precision	Recall	f1-score
ZINC	0.76	0.70	0.73	0.81	0.67	0.73	0.80	0.70	0.75
AMIANTO	0.88	0.87	0.87	0.89	0.89	0.89	0.89	0.87	0.88
CLAY	0.77	0.76	0.77	0.77	0.80	0.78	0.91	0.78	0.84
IMITATION CLAY	0.26	0.75	0.39	0.40	0.77	0.53	0.36	0.73	0.48
OXIDIZED ZINC	0.35	0.36	0.35	0.26	0.29	0.27	0.40	0.65	0.49

Classification Processing 2.

Confusion matrix on the Quimbaya test sets.

	Original					Processing 1					Processing 2				
Roof	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
ZINC (1)	649	170	5	5	90	637	165	8	0	145	652	181	4	3	97
AMIANTO (2)	127	1884	48	6	109	104	1984	33	0	98	110	1904	15	3	152
CLAY (3)	3	30	617	88	51	8	26	657	65	63	0	16	629	62	104
IMITATION CLAY (4)	0	51	12	50	2	3	0	1	48	10	0	0	6	43	10
OXIDIZED ZINC (5)	74	36	105	19	133	35	66	156	7	109	51	29	40	8	238

Results (San Andrés)

For the municipality of San Andrés, the assigned labels were green shingle, red shingle, blue shingle, amianto, zinc, and oxidized zinc. The obtained metrics on the test sets are presented below.

Metrics on San Andrés Test Sets

Roof	Original			Processing 2		
	acuraccy			acuraccy		
	0.75			0.78		
	Precision	Recall	f1-score	Precision	Recall	f1-score
GREEN SHINGLE	0.29	0.88	0.93	0.31	0.87	0.46
ZINC	0.95	0.88	0.88	0.97	0.91	0.94
AMIANTO	0.91	0.92	0.91	0.92	0.52	0.66
RED SHINGLE	0.17	0.97	0.96	0.09	0.86	0.16
OXIDIZED ZINC	0.35	0.68	0.46	0.36	0.57	0.44
BLUE SHINGLE	0.86	0.94	0.96	0.92	0.95	0.94

Confusion matrix on the San Andrés test sets

	Original						Procesamiento 2					
Cubierta	1	2	3	4	5	6	1	2	3	4	5	6
GREEN SHINGLE (1)	71	0	4	0	4	5	73	0	1	1	7	2
ZINC (2)	11	1757	10	2	79	63	14	1756	5	33	66	48
AMIANTO (3)	143	30	438	9	452	18	122	20	563	54	331	0
RED SHINGLE (4)	0	0	0	15	5	2	0	0	1	19	2	0
OXIDIZED ZINC (5)	14	39	26	62	254	13	15	25	41	90	232	2
BLUE SHINGLE (6)	10	18	4	1	11	629	9	8	1	12	6	637

Classification Processing 2.

Results (Carmen de Bolívar)

The assigned labels for the municipality of Carmen de Bolívar are amianto, zinc, and oxidized zinc, with these being the predominant roofing materials
The obtained metrics on the test sets are presented below.

Metrics on Carmen de Bolívar Test Sets

Roof	Original			Processing 1			Processing 2		
	acuraccy			acuraccy			acuraccy		
	0.83			0.85			0.85		
	Precision	Recall	f1-score	Precision	Recall	f1-score	Precision	Recall	f1-score
ZINC	0.9	0.75	0.82	0.91	0.8	0.85	0.91	0.78	0.84
AMIANTO	0.88	0.92	0.9	0.89	0.88	0.88	0.9	0.92	0.91
OXIDIZED ZINC	0.7	0.87	0.77	0.71	0.89	0.79	0.72	0.89	0.8

Confusion matrix on the Carmen de Bolívar test sets

	Original			Procesamiento 1			Procesamiento 1		
Roof	1	2	3	1	2	3	1	2	3
ZINC (1)	988	96	240	1055	83	186	1034	76	214
AMIANTO (2)	51	892	29	46	853	73	49	893	30
OXIDIZED ZINC (3)	59	31	613	53	21	629	49	26	628

Classification Processing 1.

Conclusions

The roof detection using the Segment Anything Model and the filtering process has shown positive results in solving the problem. Although there is a 16% increase in generated polygons compared to manual digitization, this difference is attributed to the model's ability to detect separately roofs with two or more waters in some cases.

Roofs with different types of materials on the same structure also influence the identification and separation of roofs.

The implementation of the Random Forest algorithm for pixel-level roof classification in 3 municipalities with different construction characteristics demonstrated the model's adaptability to various geographical and climatic conditions reflected in the images.

While the model achieves accuracies above 75% for pixel classification in urban contexts, specific challenges were identified in the detection of particular materials, such as oxidized zinc, which exhibits confusion with clay.

Fieldwork is required to increase the performance and quality of the models.

SUSTAINABLE DEVELOPMENT GOALS

International Federation of Surveyors supports the Sustainable Development Goals

Commission

Commission's name

Serving Society for the Benefit of People and Planet