Evaluating the accuracy of manual classification in satellite images using supervised algorithms

ARTICLE HISTORY
Received 10 October 2025
Accepted 20 November 2025
Published 6 January 2026

Zulema Yamileth Cervantez Hernández
Universidad Veracruzana
Facultad de Estadística e Informática
Xalapa, Veracruz
zS21023196@estudiantes.uv.mx
ORCID: 0009-0000-9453-2582

Emmanuel Morales García
Universidad Veracruzana
Facultad de Estadística e Informática
Xalapa, Veracruz
emmorales@uv.mx
ORCID: 0000-0002-6837-9227

Cecilia Cruz López
Universidad Veracruzana
Facultad de Estadística e Informática
Xalapa, Veracruz
ceccruz@uv.mx
ORCID: 0000-0002-9156-5669

Itzel Alessandra Reyes Flores
Uniersidad Veracruzana
Facultad de Estadística e Informática
Xalapa, Veracruz
itreyes@uv.mx
ORCID: 0000-0003-0733-8453

Z. Cervantes, E. Morales, C. López, and I. Reyes, "Evaluating the accuracy of manual classification in satellite images using supervised algorithms", Latin-American Journal of Computing (LAJC), vol. 13, no. 1, 2026.

Evaluating the accuracy of manual classification in satellite images using supervised algorithms

Zulema Yamileth Cervantez Hernández
Universidad Veracruzana
Facultad de Estadística e Informática
Xalapa, Veracruz
zS21023196@estudiantes.uv.mx

Emmanuel Morales García
Universidad Veracruzana
Facultad de Estadística e Informática
Xalapa, Veracruz
emmorales@uv.mx

Cecilia Cruz López
Universidad Veracruzana
Facultad de Estadística e Informática
Xalapa, Veracruz
ceccruz@uv.mx

Itzel Alessandra Reyes Flores
Uniersidad Veracruzana
Facultad de Estadística e Informática
Xalapa, Veracruz
itreyes@uv.mx

Abstract— This research focused on evaluating the manual classification of land cover using Sentinel-2 imagery. Supervised algorithms were applied to validate and improve this process. Three algorithms were selected based on their computational efficiency: K-Nearest Neighbors (KNN), Random Forest (RF), and Support Vector Machine (SVM). The results show that KNN achieved optimal performance, demonstrating a solid balance between accuracy, F1 score, and execution time compared to RF. RF, for its part, obtained greater accuracy, indicating its superior ability to correctly identify classes; however, it requires more computational resources. SVM exhibited lower performance in the evaluated metrics but achieved a shorter execution time. It was identified as the algorithm with the greatest limitations for separating classes within this dataset derived from the different study areas. Overall, the comparison confirmed that the manual classifications developed in QGIS are supported and validated by the application of these supervised methods. The use of such algorithms contributes to improving the accuracy, consistency, and efficiency of geospatial classification tasks.

Keywords— *Remote sensing, Machine learning, Infrared imaging, Data science*

I. Introduction

Analyzing satellite images offers benefits such as the ability to visualize spatiotemporal patterns of the Earth, the environment, and climate change. This type of study enables the monitoring and understanding of these processes, leading to greater precision [1][2],[3]. Another advantage of studying images, such as multispectral images, is that they provide more information about the Earth's surface and vegetation. Currently, the Sentinel-2 remote sensor of the European Space Agency (ESA) provides the most detailed images of the Earth from space [4].

As noted in [5], the diversity of missions carried out by Sentinel-2 within the Copernicus program is highly relevant to Earth observation. Its advanced design and capabilities have driven significant progress in land cover monitoring, precision agriculture, natural disaster management, and ecosystem studies. Sentinel-2 has two satellites (2A and 2B) equipped with the Multispectral Instrument (MSI), which captures images in thirteen spectral bands ranging from visible to shortwave infrared. Its spatial resolution ranges from ten to sixty meters, and its swath width is 290 km [6],[7].

The characteristics of the remote sensor allow for continuous and highly detailed monitoring of study areas. However, using traditional classification methods presents complications in terms of reproducibility and efficiency, especially when applied to heterogeneous areas [8].

Furthermore, the spectral and spatial benefits provided by Sentinel-2 generate a large volume of data that require processing with advanced methods such as machine learning or image processing techniques. In this context, supervised algorithms have broad and optimal potential for classifying information from satellite images. Some models used are Support Vector Machines (SVM), Random Forests (RF), and K-Nearest Neighbors (KNN). These algorithms can process spectral data and identify patterns in images [9],[10],[11]. These algorithms perform well; however, their accuracy varies depending on environmental conditions and the land cover being studied.

Remote sensing and machine learning combined make efforts to transform this type of data into information for decision-making on environmental and territorial issues [12],[13]. Currently, the literature explores the processes, development, and application of supervised algorithms in satellite image classification. A widely cited study is that of [14], who conducted a thorough analysis of the application of RF in satellite image processing, demonstrating its optimal capacity for processing large volumes of data and its robustness against overfitting.

This background highlights the relevance of employing algorithms such as RF, using satellite images from Sentinel-2. To complement this information, [15] conducted a study on the application process of remote sensing techniques, analyzing six commonly used algorithms, such as SVM and RF. They conclude that these types of algorithms offer the capacity to solve high-dimensional problems. This background reinforces the idea that comparing methods is beneficial for verifying operational performance in real geospatial scenarios, as proposed in this research conducted in Veracruz, Mexico. Based on the above, the objective of this research is to identify which supervised machine learning algorithms achieve the best accuracy in classifying areas in satellite imagery. This involves evaluating the classification of Sentinel-2 satellite images in various geospatial scenarios in Veracruz, Mexico.

The analysis compares the results of manual classification with those obtained using SVM, RF, and KNN in three distinct regions: the city of Xalapa, Pico de Orizaba, and Cofre de Perote. These areas were selected for their unique environmental characteristics, including urban centers, bodies of water, diverse vegetation types, and high-altitude snow-capped mountains, allowing for a direct evaluation of the algorithms' performance under different topographic and ecological conditions.

II. Problem Statement

One of the main challenges in satellite image analysis is the accurate classification of land cover, particularly in heterogeneous areas characterized by varying vegetation density and expanding urban zones. Examples of such environments include Xalapa, Cofre de Perote, and Pico de Orizaba in the state of Veracruz, Mexico.

Traditional (or manual) classification methods often present obstacles when attempting to differentiate spectrally similar classes, such as dense and sparse vegetation, urban areas, or arid zones, especially when using medium-resolution imagery. These limitations can compromise the reliability of studies focused on land-use monitoring or urban planning. Therefore, it is essential to conduct post-hoc evaluations using supervised algorithms to assess the accuracy of classifications obtained through manual methods (e.g., manual class selection in QGIS). Supervised algorithms help determine whether classes were assigned correctly, thus strengthening the results through greater accuracy and better generalization across complex landscapes.

III. Related Works

In this study, the authors [16] identified and classified greenhouses in the Anamur region of Mersin, Turkey, using Sentinel-2 MSI medium-resolution and SPOT-7 high-resolution images. This research focused on object-based image analysis (OBIA) using KNN, RF, and SVM algorithms to evaluate which of the different methods and sensors is most effective for greenhouse classification. Multispectral images taken on August 2, 2018, were used, along with field data and visual validation. The methodology consisted of several stages, including atmospheric correction for images with cloud cover, segmentation of the Sentinel-2 MSI images using the ESP-2 tool, extraction of spectral and textural features, as well as the NDVI and NDWI indices, and finally, the application and comparison of the algorithms. The results indicate that the most accurate methods were KNN and RF with SPOT-7 images, achieving an overall accuracy of 91.43% and a Kappa coefficient of 0.88. On the other hand, KNN was the best classified greenhouses in Sentinel-2 MSI images, as it had the highest accuracy (88.38%) and a Kappa coefficient of 0.83. In conclusion, both sensors demonstrated good effectiveness for greenhouse classification, despite having different resolutions. Furthermore, KNN and RF proved to be the most accurate methods.

In another study [17], the performance of Random Forest (RF), Support Vector Machine (SVM), and a combination of both algorithms, known as Stack, was evaluated for classifying satellite images in rural and urban areas of Bangladesh, specifically in the Bhola and Dhaka regions. Images from Landsat-8, Sentinel-2, and Planet satellites were used to determine which sensor and algorithm combination offers the greatest accuracy for detecting land use and land cover changes (LULC) in areas considered fragmented. Geometric and atmospheric corrections were applied, a training dataset was created for each region, and the objects were classified using RStudio software. This classification yielded six classes for Bhola: water bodies, tree vegetation, rainfed agriculture, wetland agriculture, fallow land, and urbanized or swampy areas. In Dhaka, only five landform classes were identified: water bodies, tree cover, grassland or agricultural land, urbanized areas, and landfills. The analysis showed that the Sentinel-2 sensor and SVM were the most accurate and highest-performing in both study areas, with an accuracy of 0.969 in Bhola and 0.983 in Dhaka, and Kappa coefficients of 0.948 and 0.968, respectively. However, SVM performed better in classifying water and vegetation-related categories, while RF and Stack were more effective at distinguishing urbanized areas and landfills. This study concluded that Sentinel-2 is well-suited for classifying different areas at a small scale and that SVM offers better results when the dataset is limited.

This study [18] compared Support Vector Machine (SVM), Random Forest (RF), and Classification and Regression Tree (CART) algorithms with the Google Earth Engine (GEE) platform to map and analyze land cover changes over Lake Urmian in Iran from 2000 to 2020. A total of 55 satellite images from Landsat 5, 7, and 8 were used over time. Additionally, 20,000 training points were obtained from previous field studies and Google Earth maps, while the control set consisted of 6,000 ground points for validation. Before processing, the images were filtered to detect cloud cover, as this could lead to misclassification. Subsequently, the different algorithms were applied and evaluated using the confusion matrix, the Kappa coefficient, and the global accuracy metric. These algorithms were also subjected to a spatial uncertainty analysis using Dempster-Shafer theory (DST) with the Idrisi program. The results of this research showed that the classified maps detected a reduction in the lake's surface area of approximately 1000 to 1400 hectares, and an increase in agricultural and urban areas. Furthermore, SVM proved to be the most efficient algorithm for multi-temporal classification with an accuracy of 92–95%, followed by RF (82–87%) and CART (63–70%).

This research evaluated and compared the Random Forest (RF), K-Nearest Neighbors (KNN), and Gaussian Mixture Model (GMM) algorithms for generating urban land cover maps of Quezon City in the Philippines [19]. The aim was to determine which algorithms are the most accurate in classifying urban and non-urbanized areas, as well as to monitor urban sprawl in the city. For this purpose, pre-corrected Sentinel-2A satellite images downloaded on July 1,

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2026
DOI: 10.5281/zenodo.17781394

2021, were used. A training set was created with 70% of the images manually labeled and 30% for validation. Spectral bands with a resolution of 10 and 20 meters were also used. Image processing was performed using QGIS software with the Semi-Automatic Classification (SCP) and Dzetsaka plugins. Three machine learning algorithms were applied and evaluated using confusion matrices, producer accuracy, user accuracy, and overall accuracy. The results indicated that the algorithms achieved high classification accuracy, with RF showing the highest at 99.32%, followed by KNN at 98.05%, and GMM at 97.17%. However, execution times varied: RF took 7 minutes, KNN approximately 2.5 minutes, and GMM less than 5 seconds. In conclusion, the applied methods proved effective for classifying urban areas, with RF exhibiting the highest accuracy, while GMM stood out for its faster data processing.

The potential of satellite imagery from Sentinel-1 with its SAR sensor, Sentinel-2 in multispectral mode, and the combination of both sensors was also studied using Linear Regression (LR), Classification and Regression Trees (CART), and Random Forest (RF) algorithms, along with airborne LiDAR data. This was done to identify which datasets best model canopy height in the Atlantic Forest of Paraná, Brazil [20]. Additionally, raw, ind, and all features were extracted from the satellite combinations. These methods were evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 metrics with training and test data samples, using R software. The findings showed that Sentinel-2 and the sensor combination are the most suitable for modeling canopy height. While Random Forests performed better than the other algorithms, achieving an RMSE of 4.92 m and an R^2 of 0.58 using only Sentinel-1 data, and an RMSE of 4.86 m and an R^2 of 0.60 using Sentinel-2 data, Sentinel-2 data demonstrates good accuracy in estimating canopy height.

In this study, the performance of Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors (KNN), and Naive Bayes was compared using high-resolution satellite imagery from Cartosat-2E, Cartosat-3, and LISS-4 over the Jaipur area of India [21]. The objective of this study was to compare two types of classification techniques, object-based and pixel-based—to determine which method offers better results in land use and land cover (LULC) classification. The data used in the study were multispectral and orthorectified images downloaded from the National Remote Sensing Agency (NRSC), which were classified using QGIS and the Orfeo Toolbox. The images also underwent a segmentation process, separability indices were calculated between the different classes, and the algorithms were compared using accuracy, recovery, F1 score, and Kappa coefficient. According to the results, for the object-based technique, the algorithm with the highest accuracy was Decision Trees with a Kappa coefficient of 0.90. In contrast, with the pixel-based method, both KNN and SVM performed best, especially with images related to the Cartosat-2E and Cartosat-3 sensors.

Finally, for this research, spectral indices (NDWI, MNDWI, AWEI_SH, AWEI_NSH, AWEI_BOTH) and machine learning algorithms (SVM, RT, MLC, KNN) were evaluated to detect water surfaces in Sentinel-2 images [22]. The aim was to determine which method offers the best classification accuracy using a pixel-based approach, as well as to identify image characteristics that could cause erroneous classification results. For this purpose, images from the Red River, Sylvia Grinnell, Rivière-Rouge, and Fraser Rivers—areas with rivers located in Canada—were used. High-resolution images (PLÉIADES, WorldView-2, TripleSat, and KOMPSAT-3) were also used as a validation dataset. The images were corrected for cloud cover issues, and spectral indices were obtained. The threshold values for the algorithms were manually adjusted and evaluated using the Critical Success Index (CSI) and the Root Mean Square Error (RMSE) to estimate river width. The results showed that the AWEI_NSH index and the SVM algorithm performed best for classification across all study areas, demonstrating greater consistency and accuracy. Furthermore, the study revealed that features increasing the likelihood of misclassification are most closely related to environmental factors such as vegetation, urban infrastructure, water turbidity, and shadows cast by objects.

IV. Methodology

A methodological framework was designed to guide the development of this research and to achieve the stated objec- tive (Figure 1).

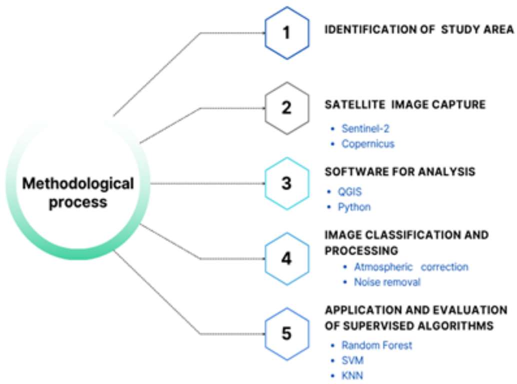

Fig. 1: Methodological process for this research

A. *Identification of the study area*

For the selection of the study areas, regions with diverse environmental scenarios—such as vegetation, urban zones, water bodies, and barren land—were considered in order to include a greater variety of classes. Three regions within the state of Veracruz were selected: Xalapa (urban area), Pico de Orizaba (protected natural area), and Cofre de Per- ote (national park). This environmental heterogeneity makes these areas ideal for assessing the accuracy of land-cover classification methods.

B. *Satellite image*

The images used in this study were obtained from the Copernicus platform via the Sentinel-2 satellite. The selected period, from February to April, was chosen to ensure favorable weather conditions for satellite observation characterized by lower precipitation and minimal cloud coverthus guaranteeing higher image quality. It is impor- tant to note that the input images were obtained in RGB format.

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 1, January 2026

C. Analysis tools

For the initial process—manual classification of the image classes—QGIS version 3.40.3 was used. This stage involved area delimitation, user-assigned classifications, and the creation of a database containing pixels and their corresponding classes. For the supervised classification algorithms, Python was employed, as it allows efficient handling of large data volumes.

D. Image processing

In this phase, the images were preprocessed to correct atmospheric noise, cloud cover, and shadows. These corrections were performed using QGIS. The objective of this step was to optimize the accuracy of spectral identification. It is worth noting that the true-color (RGB) images were converted into false-color composites using the B08, B04, and B03 bands, as the B08 band enhances vegetation, facilitating classification in each image. Following this, spectral features were obtained from the B2, B3, B4, B8, B11, and B12 bands, chosen for their ability to distinguish land cover.

E. Manual classification of images

Manual classification was performed, in which the user assigns classes based on predefined criteria image content, identifying various types of land cover in each study area. A total of six classes were defined: urban areas, water bodies, dense vegetation, sparse vegetation, no vegetation, and snow. Five classes were identified in Xalapa and Pico de Orizaba, while six categories were present in Cofre de Perote. After classification, a comma-delimited (CSV) file containing the pixels of each image and their corresponding class was exported.

F. Supervised algorithms

RF is a supervised algorithm that combines multiple decision trees to generate a more robust model. Developed by American statistician Leo Breiman in 2001, it is based on the principle of bagging (bootstrap aggregation), where multiple decision trees are built and trained with random subsets of data and features. This approach reduces correlation between trees and improves model extension. At each node, a random subset of variables is selected to determine the optimal split, increasing the diversity of the forest. Final predictions are derived from the results of all trees. While each individual tree represents a weak classifier, their combination results in a robust, stable, and low-variance ensemble model [23], [24].

KNN is based on the principle that a new data point can be classified or predicted by analyzing its k nearest neighbors within the feature space. The algorithm's performance depends heavily on the distance metric used and the appropriate selection of the hyperparameter k. In classification, the class of a new data point is determined by the average of its nearest neighbors, while in regression, the predicted value corresponds to the average of the neighbors' outputs. It is important to note that selecting an appropriate value for k is crucial, as choosing a value that is too high can increase the prediction error and negatively impact the model's performance [25].

SVM is used in classification and prediction. Its main function is to find the optimal hyperplane that separates data points of different classes with the maximum margin; that is, the greatest possible distance between the hyperplane and the nearest data points of each class, known as support vectors. SVM combines the maximum margin principle, which improves the model's generalizability, with the kernel method, which allows the algorithm to handle nonlinearly separable data by projecting it onto a higher-dimensional feature space where linear separation becomes possible [26], [27].

G. Characteristics of the models and their process

This section describes how the classes used to generate the classification were heterogeneously distributed, which directly influenced the model performance. The category with the most data was vegetation, followed by urban areas, and then the remaining classes.

Hyperparameters:

- For RF, 200 trees were used.
- For KNN, k=5 and Euclidean distance were used.
- For the SVM, an RBF kernel with C=1.0 and gamma=0.1 was used.

For all three methods, cross-validation was used to strengthen the results; a k-fold = 5 was performed.

H. Metrics for the evaluation of algorithms

Confusion Matrix

Table I shows the characteristics of a confusion matrix.

TABLE I. Example of the confusion matrix

		Prediction	
		Positive	**Negative**
Observation	**Positive**	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

Where:

- **TP**: the model correctly predicts a positive case.
- **TN**: the model correctly predicts a negative case.
- **FP**: the model predicts positive when it is negative (Type I error).
- **FN**: the model predicts negative when it is positive (Type II error).

Other important metrics:

- F1 score: An F1 score close to 1 indicates a good model, while a value of 0 indicates poor predictive performance [29].
- Accuracy: Measures the proportion of positive cases correctly predicted [28].
- Recall: Identifies the proportion of actual positive cases [28].
- ROC curve: Helps with the overall model evaluation; it is expressed by the AUC (area under the curve). A

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 1, January 2026

value close to 1 represents a good model [30].

V. Results

A. . *Images analyzed*

The main results of this research are presented below, covering the process from image acquisition to the evaluation of manual classifications. To perform the classification, Sentinel-2 satellite images were cropped to delimit the study area. The images used correspond to true color composites (RGB with bands B4, B3, and B2). Within the delimited areas, objects in each image were manually classified using QGIS.

Subsequently, the original image was transformed into a false-color composite by incorporating the near-infrared band in place of one of the visible bands. This approach highlights vegetation and other elements with higher reflectance in the infrared spectrum.

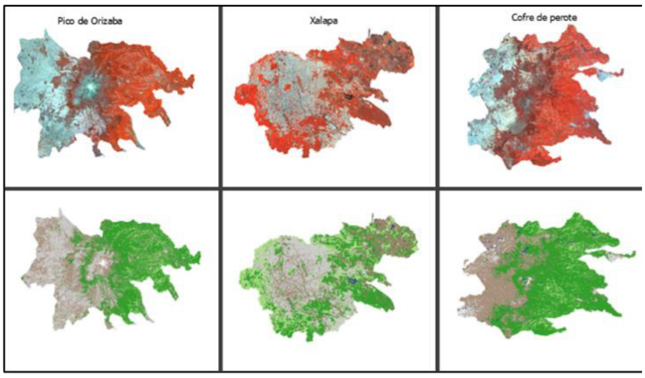

Fig. 2. Study areas cut out and classified

Figure 2 presents the set of satellite images corresponding to the three study areas: Pico de Orizaba, Xalapa, and Cofre de Perote. The top row displays the color composites of each area, highlighting differences in land cover: vegetation (reddish tones), urban areas (bluish tones), and sparse vegetation (gray tones). It is important to note that these images were generated using infrared bands to enhance the accuracy of manual classification.

The bottom row shows the final classification results. This representation enables a clearer comparison of the spatial distribution of vegetation cover versus urban land area in each of the study areas. Table II below summarizes the classifications produced in QGIS.

TABLE II. Land Cover Classes and Coding

Color	Land cover classes	Coding
	Urban area	UA (0)
	Water bodies	WB (1)
	Dense vegetation	DV (2)
	Sparse Vegetation	SV (3)
	No vegetation	NV (4)
	Snow	S (5)

Table II presents the classifications obtained manually in QGIS. Six general groups were defined: Urban Zone (UZ), representing built-up areas; Water Bodies (CA), corresponding to rivers, lakes, or dams; Dense Vegetation (VD), associated with areas of high vegetation cover; Sparse Vegetation (PV), referring to areas with intermediate or degraded cover; No Vegetation (NV), representing arid surfaces; and Snow (N), comprising areas covered by ice and snow. This last category appears only in the images of Pico de Orizaba and Cofre de Perote.

B. *Evaluation of manual classification*

The database used to run the algorithms corresponds to the classified pixels from the three satellite images (Xalapa: 352,338; Pico de Orizaba: 572,412; Cofre de Perote: 359,827 classified pixels in each image). These datasets were exported from QGIS in CSV format. Each record represents a pixel with its spectral values per band, along with its coordinates and the assigned class label (e.g., Urban Area, Water Bodies, Snow, etc.).

For algorithm training and evaluation, each pixel set was divided into two subsets: 80\% of the data were used for training, and the remaining 20\% for testing. Subsequently, 5-fold cross-validation was applied to reduce bias and increase the robustness of the algorithm results. The outcomes are summarized in Table III.

TABLE III. Classification Metrics by Zone and Algorithm

Zones	Metrics	RF	SVM	KNN
Xalapa	Accuracy	0.9808	0.9293	0.9811
	F1 Score	0.9808	0.9288	0.9811
Pico de Orizaba	Accuracy	0.9806	0.7254	0.9776
	F1 Score	0.9805	0.7151	0.9776
Cofre de Perote	Accuracy	0.9874	0.8855	0.9828
	F1 Score	0.9873	0.8746	0.9827

The evaluation of the metrics shows that, for the classifications of the Xalapa image, both Random Forest (RF) and KNN achieved nearly identical accuracy, with an F1 score of 0.98, indicating optimal and balanced performance. In contrast, SVM produced more variable results (0.70–0.93), reflecting a higher number of classification errors compared to the other algorithms. For Pico de Orizaba and Cofre de Perote, RF and KNN also achieved higher accuracy, while SVM still provided acceptable performance.

These results were obtained through 5-fold cross-validation, which reduces overfitting and provides more reliable performance estimates. The execution times of each algorithm for the study areas are shown below.

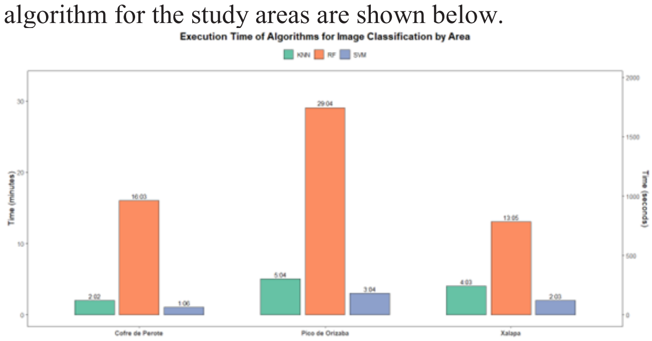

Fig. 3: execution time in algorithms

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 1, January 2026

Another key factor in evaluating the algorithms was execution time. Random Forest (RF) demonstrated consistent and robust performance across the previously discussed metrics.

However, it was also the algorithm with the highest computational cost for image classification, with execution times ranging from 13 to 30 minutes. In contrast, KNN and SVM required substantially less time (between 1 and 5 minutes), making them more practical when the goal is to accelerate classification and reduce processing time (Figure 3).

Overall, these results indicate that KNN provides a more favorable balance between efficiency and accuracy, supporting its use as a reliable complement to manual image evaluation. Conversely, RF remains a strong alternative when the priority is to maximize classification accuracy, regardless of execution time.

C. *Visualization of the confusion matrix and ROC curve*

Below are some of the results obtained from the different classification algorithms. Figures 4, 5, and 6 illustrate the performance of Random Forest in Xalapa, SVM in Pico de Orizaba, and KNN in Cofre de Perote. Although all three algorithms were applied to each study area, only representative cases—ranging from lower to higher classification performance—are presented.

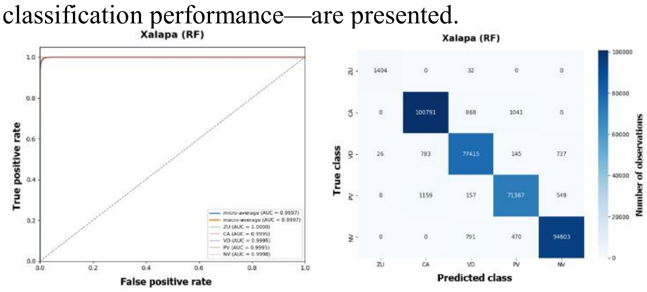

Fig.4. ROC curve and confusion matrix for the classification of Xalapa (RF)

In the case of Xalapa, the Random Forest (RF) model achieved an AUC value close to 1, indicating an outstanding ability to distinguish between classes. The confusion matrix further confirms this accuracy, as most observations are concentrated along the main diagonal with minimal classification errors. These results demonstrate the robustness of the algorithm in a heterogeneous urban–vegetation environment (Figure 4).

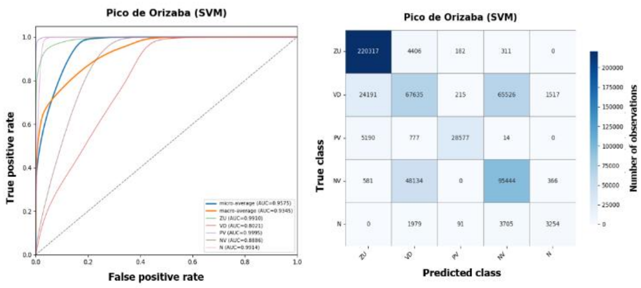

Fig.5. ROC curve and confusion matrix for the classification of Pico de Orizaba (SVM)

Another example is Pico de Orizaba, where the SVM algorithm exhibited more variable performance. Some classes, such as Urban Areas and Sparse Vegetation, achieved AUC values greater than 0.99, whereas other classes, including Dense Vegetation and No Vegetation, showed values between 0.80 and 0.88, indicating difficulties in discriminating among land-cover types. The confusion matrix, in turn, reveals notable misclassifications between Dense Vegetation and other land-cover types, as well as between No Vegetation and Sparse Vegetation (Figure 5).

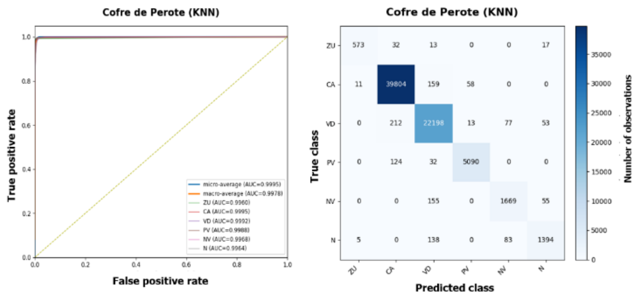

Fig.5. ROC curve and confusion matrix for the classification of Cofre de Perote (KNN)

Finally, for the Cofre de Perote image, the KNN algorithm achieved high AUC values, and the confusion matrix shows that, despite the overall accuracy, some misclassifications occurred between areas such as Dense Vegetation and No Vegetation (Figure 6).

VI. Discussion

According to the results obtained in this research, supervised algorithms are essential for evaluating manual land cover classifications derived from satellite imagery. Consistent with the findings of [16], [19], and [20], the RF algorithm demonstrated the highest accuracy in distinguishing the classes created during the experimental phase. Its robustness, ability to handle spectral variability, and efficiency in managing complex pixel data confirm its excellent performance in validating image classifications in heterogeneous regions such as Xalapa, Cofre de Perote, and Pico de Orizaba, areas characterized by urban zones, dense vegetation, and bodies of water.

The K-Nearest Neighbors (KNN) algorithm also performed well, yielding results comparable to those of RF. However, the main difference lies in the execution time, with KNN being considerably faster. As noted in [19], KNN stands out for its computational efficiency, which was also confirmed in the experimental phase of this study. Therefore, this algorithm can be considered an optimal alternative when speed and accuracy are required.

Conversely, the Support Vector Machine (SVM) algorithm showed lower accuracy in this study compared to the other two classifiers, unlike the results reported by [17] and [18], where the SVM achieved superior performance. This discrepancy can be attributed to the distinctive characteristics of the study areas in this research (high variability, cloud cover, and dense vegetation), which can reduce the spectral separation of the categories, thus improving the performance of the SVM algorithm, as well as the high spatial variability. Finally, compared to the other two algorithms used, SVM is governed by a linear function and is therefore sensitive to a lack of homogeneity, unlike RF and KNN.

VII. Conclusion

The utility of supervised algorithms allows for the validation and reinforcement of classifications performed manually in QGIS, corroborating their objectivity and computational efficiency to support decision-making in land cover analysis. While user-assigned classifications depend on their judgment and experience (supported by the QGIS software), supervised algorithms provide a quantitative and reproducible framework that minimizes subjectivity in the process. Therefore, the use of RF, KNN, and SVM not only produces accurate results but also offers a scientific basis for confirming visually delineated boundaries, thus improving the reliability of the resulting classifications and their applicability in remote sensing studies.

References

[1] Ouchra, H., Belangour, A., & Erraissi, A. (2023). Machine learning algorithms for satellite image classification using Google Earth Engine and Landsat satellite data: Morocco case study. IEEE Access, 11, 71127–71142. https://doi.org/10.1109/ACCESS.2023.3293828.

[2] Wang, Y., Sun, Y., Cao, X., Wang, Y., Zhang, W., and Cheng, X. (2023). A review of regional and Global scale Land Use/Land Cover (LULC) mapping products generated from satellite remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 206, 311-334.

[3] Joshi, P., Saxena, P., & Sharma, A. (2024). Trends and Advancements in Satellite Image Classification: From Traditional Methods to Machine Learning Approaches. ResearchGate. 10.13140/RG.2.2.13239.23201.

[4] European Space Agency. (n.d.). Introducing Sentinel-2. ESA. https://www.esa.int/Applications/Observing the Earth/Copernicus/Sentinel- 2/Introducing Sentinel-2.

[5] Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., and Ranagalage, M. (2020). Sentinel-2 data for land cover/use mapping: A review. Remote sensing, 12(14), 2291.

[6] Segarra, J., Buchaillot, M. L., Araus, J. L., & Kefauver, S. C. (2020). Remote sensing for precision agriculture: Sentinel-2 improved features and applications. Agronomy, 10(5), 641.

[7] Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... & Bargellini, P. (2012). Sentinel-2: ESA's optical high- resolution mission for GMES operational services. Remote sensing of Environment, 120, 25-36.

[8] Fragoso-Campo´n, L., QUIRO´ S, R. E., & GUTIE´RREZ, G. J. (2020). Clasificacio´n supervisada de ima´genes PNOA-NIR y fusio´n con datos LiDAR-PNOA como apoyo en el inventario forestal. Caso de estu- dio: Dehesas. CUADERNOS DE LA SOCIEDAD ESPAN˜OLA DE CIENCIAS FORESTALES: Sociedad Espanola de Ciencias Forestales, 45(3), 77-96.

[9] Zhang, C., Liu, Y., & Tie, N. (2023). Forest land resource information acquisition with sentinel-2 image utilizing support vector machine, k- nearest neighbor, random forest, decision trees and multi-layer percep- tron. Forests, 14(2), 254.

[10] Yuh, Y. G., Tracz, W., Matthews, H. D., & Turner, S. E. (2023). Application of machine learning approaches for land cover monitoring in northern Cameroon. Ecological informatics, 74, 101955.

[11] Thanh Noi, P., & Kappas, M. (2017). Comparison of random forest, k- nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. Sensors, 18(1), 18.

[12] Knudby, A. (2021). Accuracy assessment. Pressbooks. https://ecampusontario.pressbooks.pub/remotesensing/chapter/ch apter- 7-accuracy-assessment/

[13] Tehsin, S., Kausar, S., Jameel, A., Humayun, M., & Almofarreh, D. K. (2023). Satellite image categorization using scalable deep learning. Applied Sciences, 13(8), 5108. https://doi.org/10.3390/app13085108.

[14] Belgiu, M., & Dragut¸ L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24–31.

[15] Maxwell, A. E., Warner, T. A. & Fang, F. (2018). "Implementation of machine-learning classification in remote sensing: An applied review." International Journal of Remote Sensing, 39(9), 2784-2817.

[16] Balcik, F. B., Senel, G., & Goksel, C. (2020). Object-Based Classifi- cation of Greenhouses Using Sentinel-2 MSI and SPOT-7 Images: A Case Study from Anamur (Mersin), Turkey. IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, 13, 2769- 2777. https://doi.org/10.1109/jstars.2020.2996315

[17] Rahman, A., Abdullah, H. M., Tanzir, M. T., Hossain, M. J., Khan, B. M., Miah, M. G., & Islam, I. (2020). Performance of different machine learning algorithms on satellite image classification in rural and urban setup. Remote Sensing Applications Society And Environment, 20, 100410. https://doi.org/10.1016/j.rsase.2020.100410

[18] Feizizadeh, B., Omarzadeh, D., Garajeh, M. K., Lakes, T., & Blaschke, T. (2021). Machine learning data-driven approaches for land use/cover mapping and trend analysis using Google Earth Engine. Journal Of Environmental Planning And Management, 66(3), 665-697. https://doi.org/10.1080/09640568.2021.2001317

[19] Santiago, R. M., Gustilo, R., Arada, G., Magsino, E., & Sybingco, E. (2021). Performance Analysis of Machine Learning Algorithms in Generating Urban Land Cover Map of Quezon City, Philippines Using Sentinel-2 Satellite Imagery. 2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communi- cation and Control, Environment, and Management (HNICEM), 1–6. https://doi.org/10.1109/hnicem54116.2021.9731856

[20] Torres, C., Jéssica Gerente, Rodrigo, Caruso, F., Providelo, L. A., Marchiori, G., & Chen, X. (2022). Canopy Height Map- ping by Sentinel 1 and 2 Satellite Images, Airborne LiDAR data and machine learning. Remote Sensing, 14(16), 4112-4112. https://doi.org/10.3390/rs14164112.

[21] Kumar, A., Kumar, G., Patil, D. S., &Gupta, R. (2024). Evaluating Machine Learning Classifiers for IRS High Resolution Satellite Images Using Object-Based and Pixel-Based Classification Techniques. Journal Of The Indian Society Of Remote Sensing. https://doi.org/10.1007/s12524-024-02084-w.

[22] Purwanto, A. D., Wikantika, K., Deliar, A., & Darmawan, S. (2022). Decision Tree and Random Forest Classification Algorithms for Mangrove Forest Mapping in Sembilang National Park, Indonesia. Remote Sensing, 15(1), 16. https://doi.org/10.3390/rs15010016.

[23] Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/a:1010933404324.

[24] Farhadi, Z., Bevrani, H., Feizi-Derakhshi, M., Kim, W., & Ijaz, M. F. (2022). An Ensemble Framework to Improve the Accuracy of Prediction Using Clustered Random-Forest and Shrinkage Methods. Applied Sciences, 12(20), 10608. https://doi.org/10.3390/app122010608.

[25] Tamamadin, M., Lee, C., Kee, S., \& Yee, J. (2022). Regional Typhoon Track Prediction Using Ensemble k-Nearest Neighbor Machine Learning in the GIS Environment. Remote Sensing, 14(21), 5292. https://doi.org/10.3390/rs14215292.

[26] Adugna, T., Xu, W., & Fan, J. (2022b). Comparison of Random Forest and Support Vector Machine Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images. Remote Sensing, 14(3), 574. https://doi.org/10.3390/rs14030574.

[27] Du, K., Jiang, B., Lu, J., Hua, J., & Swamy, M. N. S. (2024b). Exploring Kernel Machines and Support Vector Machines: Principles, Techniques, and Future Directions. Mathematics, 12(24), 3935. https://doi.org/10.3390/math12243935.

[28] S. Visa, B. Ramsay, A. Ralescu, and E. Van Der Knaap, "Confusion matrix-based feature selection," CEUR Workshop Proc., vol. 710, pp. 120-127, 2011.

[29] Pikabea, I. (2022). Review of learning models in the field of question answering [Undergraduate thesis, University of the Basque Country, Faculty of Informatics]. University of the Basque Country.

[30] Nahm, F. (2022). Receiver Operating Characteristic Curve: Overview and Practical Use for Clinicians [Epub 2022 Jan 18]. Korean Journal of Anesthesiology, 75 (1), 25-36. https://doi.org/10.4097/kja.21209

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 1, January 2026

Z. Cervantes, E. Morales, C. López, and I. Reyes,
"Evaluating the accuracy of manual classification in satellite images using supervised algorithms",
Latin-American Journal of Computing (LAJC), vol. 13, no. 1, 2026.

AUTHORS

Zulema Yamileth Cervantez

Estudiante de octavo semestre de la Licenciatura en Estadística en la Facultad de Estadística e Informática de la Universidad Veracruzana. Su formación académica se ha orientado al análisis de datos, modelación estadística y computación aplicada, con un interés particular en el uso de tecnologías emergentes para la solución de problemas complejos. Actualmente desarrolla su tesis de investigación en el área de percepción remota, enfocándose en el procesamiento de imágenes satelitales, Machine Learning y técnicas de detección de objetos. Su trabajo integra métodos estadísticos con algoritmos de visión por computadora para la identificación automatizada de patrones espaciales, con aplicaciones potenciales en monitoreo ambiental y análisis territorial. Ha participado en actividades académicas relacionadas con programación en Python y R, así como en cursos sobre minería de datos, aprendizaje supervisado y análisis espacial. Su formación combina fundamentos teóricos sólidos con competencias prácticas que fortalecen su perfil como futura profesionista en ciencia de datos y estadística aplicada.

Emmanuel Morales García

Licenciado en Ciencias y Técnicas Estadísticas y Especialista en Métodos Estadísticos por la Universidad Veracruzana, con Maestría en Ciencias de la Información Geoespacial por el Centro Geo, CDMX (Centro CONAHCYT), actualmente cursando el Doctorado en Ciencias de la Computación en la Universidad Veracruzana. Profesor en la Licenciatura en Estadística, en la Especialidad en Métodos Estadísticos y la Maestría en Economía y Sociedad de China y América Latina en la misma universidad, donde he dirigido 15 tesis de licenciatura y 4 de especialidad. También tengo experiencia como Analista Estadístico en la Oficina del Programa de Gobierno del Estado de Veracruz. Mis líneas de investigación incluyen metodologías de cómputo, programación estadística, estadística multivariada, análisis espacial, ciencia de datos y modelos estadísticos, con aplicaciones en biología, medicina, ciencias administrativas y sociales. He participado en diversos congresos nacionales e internacionales.

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 1, January 2026

Z. Cervantes, E. Morales, C. López, and I. Reyes,
"Evaluating the accuracy of manual classification in satellite images using supervised algorithms",
Latin-American Journal of Computing (LAJC), vol. 13, no. 1, 2026.

AUTHORS

Cecilia Cruz López

Profesora de tiempo completo en la Facultad de Estadística e Informática de la Universidad Veracruzana (UV). Es Doctora en Investigación Educativa por la UV, Maestra en Ciencias con especialidad en Estadística Aplicada por el ITESM Campus Monterrey, así como Especialista en Métodos Estadísticos y Licenciada en Estadística por la UV. Actualmente coordina la Especialización en Métodos Estadísticos y pertenece al Sistema Nacional de Investigadores (SNII) como Candidata.

Sus líneas de investigación abarcan la Educación Estadística, la Metodología Estadística Aplicada y la integración de la estadística con tecnologías emergentes como Machine Learning, Ciencia de Datos y Análisis Espacial. Ha colaborado en proyectos sobre sustentabilidad, alfabetización digital y actitudes hacia la estadística en estudiantes latinoamericanos.

Entre sus publicaciones recientes destacan trabajos sobre aprendizaje supervisado, formación en consultoría estadística y aplicaciones multivariantes, además de capítulos sobre alfabetización digital y redes sociales. Ha dirigido más de 70 tesis, combinando docencia, investigación e impulso al uso estratégico de la estadística para el desarrollo sostenible.

Itzel Alessandra Reyes Flores

Es Licenciada en Informática, Maestra en Sistemas Interactivos Centrados en el Usuario y Doctora en Ciencias de la Computación, con formación completa en la Universidad Veracruzana. Se desempeña como Técnica Académica de Tiempo Completo en la Facultad de Estadística e Informática de la UV, donde colabora en actividades de docencia, investigación, acompañamiento académico, gestión académica y apoyo en el análisis de procesos administrativos de la universidad. Posee experiencia sólida en desarrollo de software, diseño de interfaces de usuario y experiencia de usuario (UX/UI), diseño de cursos e-learning y desarrollo de aplicaciones móviles, además de contar con certificaciones en programación de software que fortalecen su práctica profesional y amplían su campo de acción. Es autora y coautora de artículos de investigación en Interacción Humano-Computadora, Inteligencia Artificial, Trabajo Colaborativo Asistido por Computadora y Tecnología Educativa, contribuyendo de manera significativa al desarrollo académico y al avance del conocimiento en estas áreas.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2026