

# BOSC: A toolbox for aerial imagery mapping

Ricard Durall<sup>1</sup>, Laura Montilla , Esteban Durall

<sup>1</sup>ricard.durall.lopez@gmail.com

**Abstract**— Accurate and efficient label of aerial images is essential for informed decision-making and resource allocation, whether in identifying crop types or delineating land-use patterns. The development of a comprehensive toolbox for manipulating and annotating aerial imagery represents a significant leap forward in remote sensing and spatial analysis. In this report, we introduce BOSC, a toolbox that enables researchers and practitioners to extract actionable insights with unprecedented accuracy and efficiency, addressing a critical need in today's abundance of drone and satellite resources.

For more information or to explore BOSC, please visit our [repository](#).

**Keywords**—Tree classification, Agriculture monitoring, Urban planning, Satellite mapping, UAV

## 1. Introduction

In recent years, advancements in technology have made aerial imagery indispensable across diverse domains, spanning agriculture [3], urban planning [7], and environmental monitoring [4]. The accessibility and cost-effectiveness of acquiring high-resolution aerial images have surged, thanks to the proliferation of resources like drones and satellites. However, harnessing the full potential of this visual data necessitates sophisticated tools capable of efficiently processing and labeling images to derive meaningful insights.

To address the increasing demand for streamlined aerial image analysis, we introduce BOSC, a comprehensive toolbox tailored to the evolving needs of researchers and practitioners. BOSC offers a versatile range of functionalities designed specifically for manipulating and labeling aerial images with exceptional precision and ease. Notably, it enables the drawing of masks and the assignment of class labels on aerial images, whether through manual or automatic means, leveraging cutting-edge artificial intelligence models.

## 2. BOSC Toolbox

### 2.1. Structure

This toolbox consists of two primary components: the backend and the frontend (see Figure 1).

**Backend:** At the heart of the toolbox lies the backend, where all heavy computation operations are executed in a server. This includes tasks such as segmentation, clustering, mapping, and other complex analytical processes. By centralizing these operations in the backend, we ensure swift and accurate processing of large datasets without compromising on performance. Leveraging powerful computational resources, the backend handles the computational heavy lifting, allowing users to interact with the toolbox seamlessly.

**Frontend:** Complementing the backend, the frontend serves as the user interface, providing an intuitive platform for users to access and utilize the toolbox's functionalities. Here, users can easily upload and manipulate aerial images, perform annotations, and visualize analysis results in real-time. The frontend is designed with simplicity and ease of use in mind, shielding users from the technical complexities of the

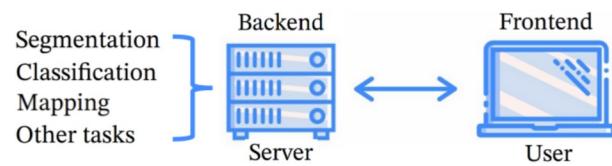


Figure 1. Client-Server architecture of BOSC.

underlying computational processes. By offering a streamlined and user-friendly interface, the frontend enhances the overall user experience, enabling users to focus on their tasks without being encumbered by unnecessary complexities.

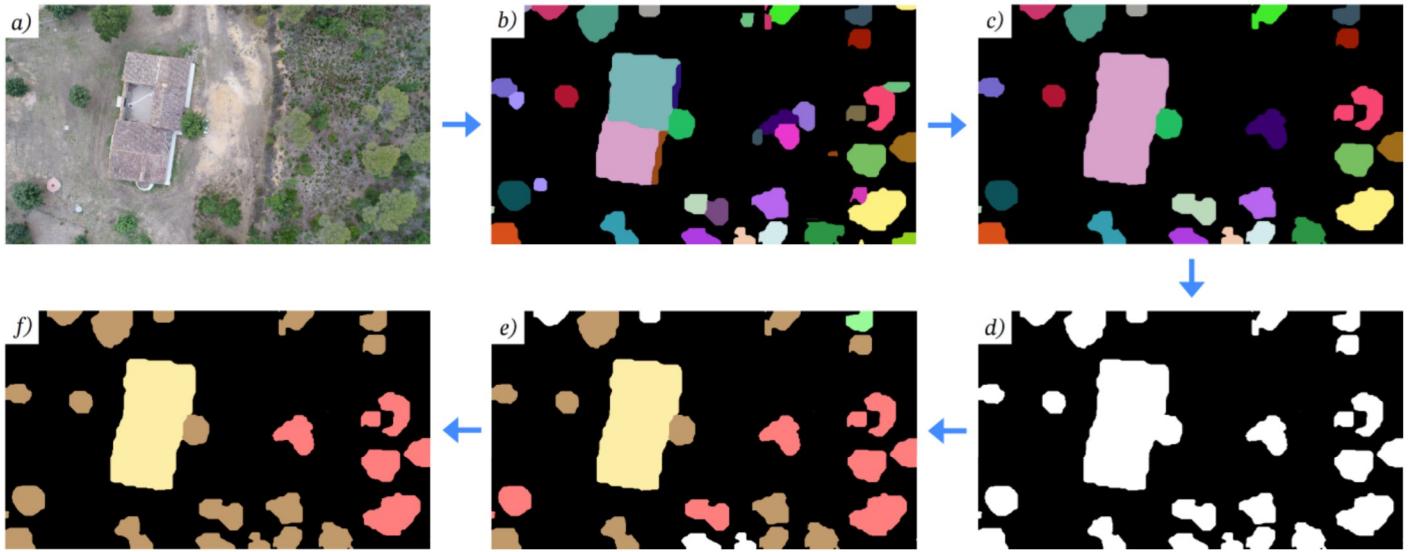
### 2.2. Method

This toolkit comprises three primary methods.

**Segmentation Engine:** Our implementation presents a robust and generalizable image segmentation model utilizing advanced deep learning techniques. Drawing inspiration from the Segment Anything model [6] and Fast Segment Anything [8] model, our approach effectively captures intricate spatial dependencies and hierarchical features within images. This ensures high accuracy in segmenting a diverse array of objects. Trained on extensive labeled datasets, the model delivers pixel-wise predictions during inference, enabling rigorous analysis and supporting various downstream applications like object tracking and scene comprehension. Engineered for efficiency and scalability specifically for aerial imagery, our model consistently integrates into existing workflows, offering users potent tools for thorough image analysis in aerial domains.

**Classification Engine:** Our classification engine utilizes unsupervised classification techniques, specifically a hierarchical clustering algorithm [2], for object classification. Operating without the need for labeled data, it discerns patterns by extracting features from pre-trained ConvNeXt model [5] and grouping similar features into clusters. Additionally, users have the option to provide labeled objects to guide and refine classification accuracy. The user-friendly interface enables interactive input definition and visualization of clustering results, regardless of the complexity of the data. This clustering approach presents a robust solution for object identification and classification in image analysis tasks.

**Mapping Engine:** Powered by [1], our mapping function employs geometric transformations to align the user input image with a reference map. To achieve that users select identifiable features (3 points) common to both the image and the map, serving as control points for transformation. Affine transformations are applied to warp the image, adjusting for translation, rotation, scaling, and shearing. Additional refinement steps ensure accurate alignment, enhancing spatial analysis reliability. The implementation prioritizes efficiency and accuracy, utilizing optimized algorithms for rapid and precise transformations. The user interface allows intuitive point selection and real-time visualization of mapped results.



**Figure 2.** Example of crown-tree classification: a) Raw input image. b) Automatic segmentation output. c) Manually fine-tuned segmentation output. d) Default classification. e) Automatic classification output (no labeled object provided). f) Manually fine-tuned classification output.



**Figure 3.** Example of mapping between an input image (left) and its corresponding location on a third-party map (center). Both images display three selected points used for the mapping. The final outcome (right) shows the new layer containing the labeled segmentation image overlaid on the interactive map.

### 3. Pipeline

To demonstrate the functionality of BOSC, we present a practical example of crown-tree classification, outlining the process from input image to mapping the labeled segmentation solution. In Figure 2, starting from the upper left corner and moving clockwise, we have:

- a) **Selection:** The raw input image containing the objects of interest, such as an aerial image capturing a landscape, agricultural field, or suburban area.
- b) **Segmentation:** The chosen input is processed by a segmentation engine, which identifies individual objects within the image. This results in a segmentation image, where each color corresponds to a separate object.
- c) **Refined Segmentation:** Users have the flexibility to correct any segmentation errors by utilizing the drawing palette as needed.
- d) **Initial Classification:** By default, all segmented objects are initially assigned to the same class (label), visually represented by white color.
- e) **Classification:** After the segmentation step, the classification algorithm generates an initial solution, assigning a color (representing a class) to each segmented object. For example, yellow for buildings and brown for oaks.
- f) **Refined Classification:** Once again, users have the flexibility to adjust any incorrect classes to produce the final labeled segmentation image.

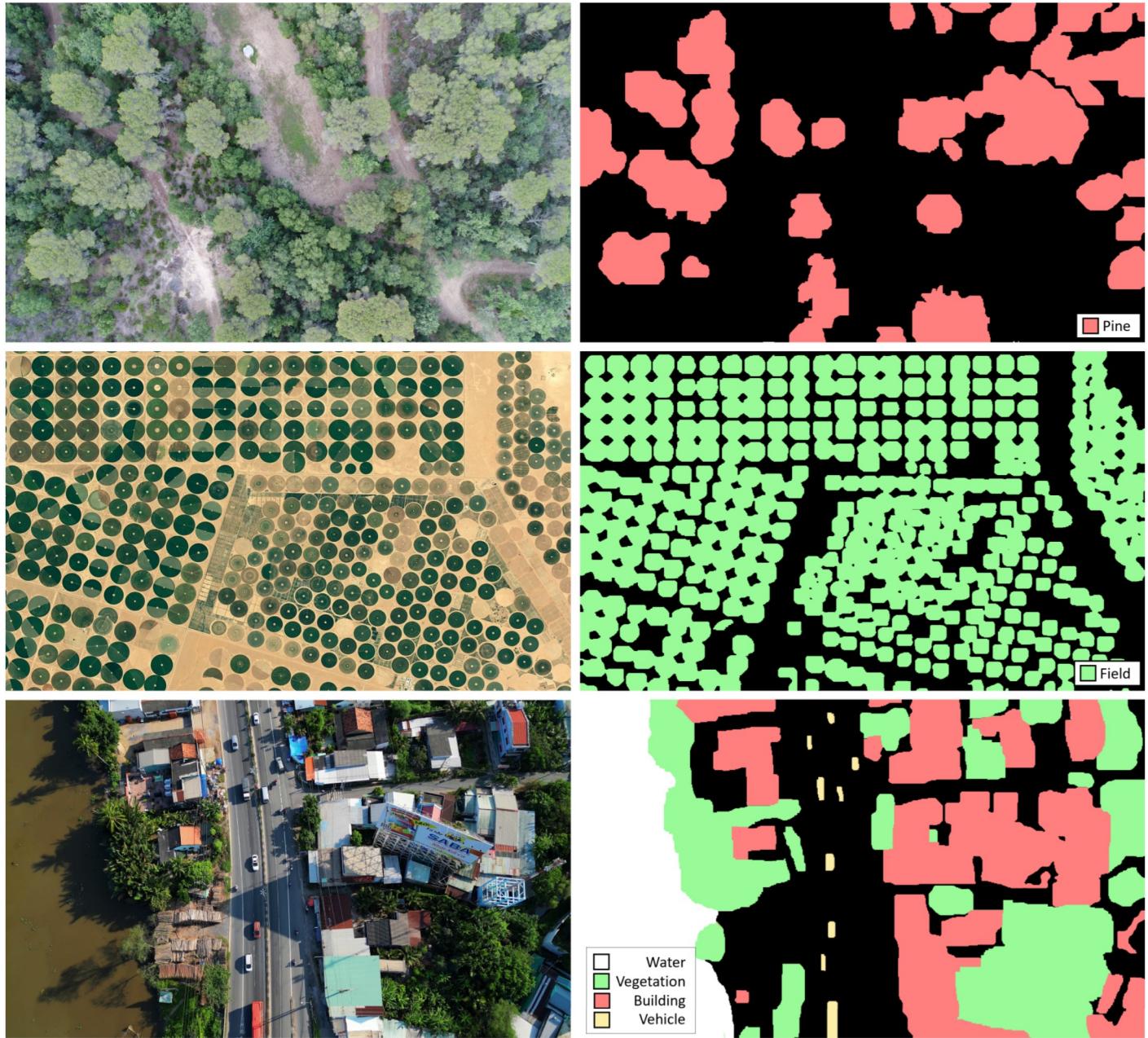
After the completion of the classification process, the next step employs the mapping engine, as depicted in Figure 3.

**Mapping:** Our toolbox enables the seamless overlay of our labeled segmentation image onto a third-party map. This is achieved by selecting three common points in both the raw input image and the target map.

**Final Result:** Finally, we can visualize our results on the map, providing a comprehensive view of the data. This includes relevant statistical information, such as the number of instances of a particular class and the total coverage area of specific classes, and it can be easily extended to other pertinent metrics. This detailed visualization allows for a deeper understanding and analysis of the mapped data.

### 4. Results

In this section, we present various outcomes generated using BOSC. As illustrated in Figure ??, this toolbox is versatile and adaptable to multiple scenarios, as its internal engines are designed to generalize independently of the target domain. This adaptability makes BOSC highly flexible, enabling its efficient use in agricultural settings for tasks such as forest population management and optimization of agricultural plantations. Additionally, we explore further use cases, such as the classification of urban or semi-urban images, highlighting the tool's extensive applicability and versatility.



**Figure 4.** From top to bottom, the first pair of images illustrates an example where BOSC is employed to map only pine trees. This scenario is useful for forest inventory, particularly in estimating the carbon fixation coefficient (net oxygen production), which is tree-dependent. The second example depicts irrigated agriculture in the desert (1 km diameter) in Saudi Arabia. By using the segmentation mask, the status of the crops can be monitored and farming practices optimized. Finally, the third example demonstrates a scenario where multiple classes are mapped, showcasing BOSC's capability to handle diverse scenarios such as semi-urban or urban environments.

## 5. Conclusions

In summary, the BOSC toolbox offers a comprehensive solution for aerial image analysis, incorporating advanced segmentation, classification, and mapping functionalities. By streamlining the image processing workflow and providing intuitive user interfaces, BOSC enables efficient and accurate classification of objects within images. Additionally, its ability to incorporate user inputs and refine segmentation results enhances flexibility and adaptability. Overall, BOSC represents a significant advancement in remote sensing and spatial analysis, empowering researchers and practitioners with powerful tools for extracting actionable insights from aerial imagery.

## References

- [1] G. Bradski, “The OpenCV Library”, *Dr. Dobb’s Journal of Software Tools*, 2000.
- [2] M. L. Zepeda-Mendoza and O. Resendis-Antonio, “Hierarchical agglomerative clustering”, *Encyclopedia of systems biology*, vol. 43, no. 1, pp. 886–887, 2013.
- [3] M. Onishi and T. Ise, “Explainable identification and mapping of trees using uav rgb image and deep learning”, *Scientific reports*, vol. 11, no. 1, p. 903, 2021.
- [4] S. W. Son, J. J. Yu, D. W. Kim, H. S. Park, and J. H. Yoon, “Applications of drones for environmental monitoring of pollutant-emitting facilities”, *Proceedings of the National Institute of Ecology of the Republic of Korea*, vol. 2, no. 4, pp. 298–304, 2021.
- [5] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, “A convnet for the 2020s”, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 11 976–11 986.
- [6] A. Kirillov, E. Mintun, N. Ravi, et al., “Segment anything”, in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015–4026.
- [7] A. M. Muhamad Kamarulzaman, W. S. Wan Mohd Jaafar, M. N. Mohd Said, S. N. M. Saad, and M. Mohan, “Uav implementations in urban planning and related sectors of rapidly developing nations: A review and future perspectives for malaysia”, *Remote Sensing*, vol. 15, no. 11, p. 2845, 2023.
- [8] X. Zhao, W. Ding, Y. An, et al., “Fast segment anything”, *arXiv preprint arXiv:2306.12156*, 2023.