SEGMENTATION OF COASTLINES IN RGB IMAGES USING SEGMENTED HYPERSPECTRAL IMAGES AS LABELS

Simen L. Eine, Corrado Chiatante, Samuel Boyle, Dennis D. Langer, Milica Orlandić

Department of Electronic Systems, Norwegian University of Science and Technology (NTNU)

ABSTRACT

Satellites can be equipped with multiple camera payloads, such as a Hyperspectral Imaging (HSI) and a Red, Green and Blue (RGB) camera, enabling on-board data fusion. Combining information from these sensors can add value across various applications, such as semantic image segmentation. The HSI typically has a larger ground sample distance than the RGB camera, but it offers higher spectral resolution and therefore more reliable discrimination between semantic classes. Such segmentation can support applications such as indirect georeferencing, where ground features like coastlines are compared with known reference ground-truth data.

This paper proposes an efficient method for segmenting RGB imagery by combining the superior spectral information of the HSI with the higher spatial resolution of the RGB sensor. The method registers hyperspectral class labels to the RGB image and uses the aligned labels as training data for a classifier that segments the coastline region. The approach is evaluated on data acquired by the HYPSO-2 satellite and achieves an average runtime of 11.79 s across the tested scenes, while improving segmentation accuracy by 12.22 % relative to the registered hyperspectral labels. The resulting high-resolution coastline labels may further improve the indirect georeferencing accuracy of RGB imagery.

***Index Terms*—** Semantic segmentation, Machine learning, Image registration, Multimodal analysis

1. INTRODUCTION

Satellite imagery is typically geolocated using direct georeferencing techniques. The limited precision of navigation and attitude sensors on many CubeSats and cost-constrained satellite missions can result in significant geolocation errors. Satellite missions such as HyperScout-1 and HYPSO-2 address this limitation by applying indirect georeferencing methods to refine the geolocation of the acquired data. These satellites are equipped with Hyperspectral Imaging (HSI) payloads, which generate large volumes of data. Due to limited downlink capacity, both missions aim to perform indirect georeferencing on-board, thereby enabling the transmission of higher-level data products with reduced data volume [1, 2].

In addition to its HSI payload, HYPSO-2 is equipped with a Red, Green and Blue (RGB) camera that provides a higher spatial resolution than the HSI sensor. On-board processing algorithms that combine hyperspectral and RGB data can be used to create data products of higher value, such as enhanced hyperspectral data produced by super-resolution methods or anomaly and object detection algorithms that reduce the size of the data products. Accurate georeferencing is a useful prerequisite for such processing. By applying indirect georeferencing to the higher spatial resolution data from the RGB camera, the geolocation accuracy of the satellite can be improved further.

Semantic segmentation is the process of assigning a class label to each pixel in an image. Hyperspectral images have a higher spectral resolution than RGB images, which enables more reliable discrimination between different semantic classes [3]. Convolutional Neural Network (CNN)s have proved effective for semantic segmentation of hyperspectral imagery, and such models can be executed on-board the HYPSO-2 satellite [4]. Segmented sea-land-cloud labels derived from HYPSO imagery have also been used for indirect georeferencing by matching extracted shorelines to a vector-based ground-truth dataset of ocean and lake coastlines [5].

Image registration is a method for geometrically aligning one image with another. It typically involves feature detection and matching, and then creating and applying a transform model that maps the sensed image onto a reference image. A common application of image registration is multimodal analysis, which combines complementary information from different sensors to extract more information from the same scene [6].

This paper proposes a method to improve the georeferencing accuracy of Earth Observation (EO) RGB images through the following steps:

1. Registering a hyperspectral image to the higher-spatial resolution RGB image, enabling the transfer and alignment of semantic segmentation labels derived from the hyperspectral data. These labels then serve as training data for segmentation of the RGB image.

2. Training a local Random Forest (RF) classifier to segment the RGB image into sea-land-cloud classes based

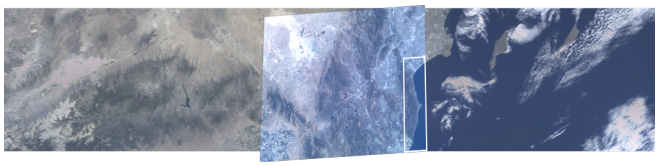


Fig. 1: RGB image registered to the HSI image. The white lines indicate the region of interest used as an example in this paper.

on the registered hyperspectral labels.

3. Segmenting the higher spatial resolution RGB image with the RF classifier to obtain higher spatial resolution sea-land-cloud labels.

The resulting high-resolution semantic labels can be used to perform indirect georeferencing with greater accuracy because the refined coastline boundaries allow more precise feature extraction and matching. By combining the higher spectral resolution of the hyperspectral image with the higher spatial resolution of the RGB image, the proposed approach exploits the complementary strengths of both sensors. This multimodal fusion enables the generation of more accurate semantic labels and, consequently, higher-quality data products for EO satellite applications.

2. METHOD

The methodology presented in this paper uses georeferenced coastline vector data and an RGB image registered to a hyperspectral image to refine the semantic class labels along the coastline. Although the proposed algorithm is intended for future on-board deployment on the HYPSO-2 satellite, this section describes the ground-based procedure used for experimental evaluation.

2.1. Registration of hyperspectral and RGB images

To ensure proper spatial alignment between the datasets, the hyperspectral and RGB images are first registered. Feature detection is performed using the Scale-Invariant Feature Transform algorithm, while feature matching is carried out using K-Nearest Neighbours with Nearest Neighbour Distance Ratio to filter matches. Finally, the Random sample consensus method is used to identify and remove outliers. This workflow, described in [7], is computationally efficient for the alignment of sensor and reference images in remote sensing applications.

The RGB image is registered to the hyperspectral image using a perspective transform derived from a homography that is estimated from the matched features. The result is shown in Figure 1.

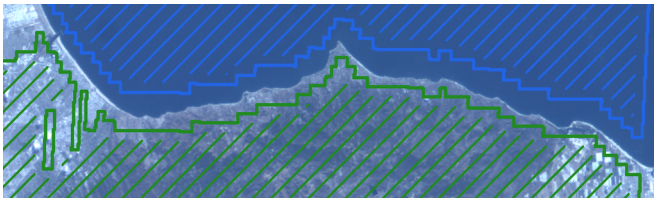


Fig. 2: Training labels on the RGB region of interest. Blue: eroded sea labels. Green: eroded land labels.

2.2. Preparing training labels

A semantically segmented hyperspectral image is used as the source of training labels in the proposed method. These labels are created using the wavelet-based CNN applied to hyperspectral images acquired by HYPSO-2 [8]. The resulting labels are then registered to the RGB image using the homography obtained during the registration step. Because the RGB images have a higher spatial resolution than the hyperspectral imagery, the label boundaries do not always align perfectly with the visible features in the RGB images.

For the HYPSO-2 satellite operating at an altitude of approximately $590\ \mathrm{km}$, the hyperspectral sensor has an across-track Ground sample distance (GSD) of $69.1\ \mathrm{m}$ pointing nadir, while the RGB camera achieves a nadir GSD of $28\ \mathrm{m}$ [9]. The along-track GSD of the HSI depends on both the frame rate and the orbital speed of the satellite [10]. For the scenes considered in this work, the average along-track GSD is $873.98\ \mathrm{m}$, ranging from $871.39\ \mathrm{m}$ to $876.43\ \mathrm{m}$. In this case, the GSD of the HSI is approximately 2.5 times larger than that of the RGB camera across-track and more than 30 times larger along-track.

Due to the substantial difference in GSD, the registered class labels from the hyperspectral image can be significantly misaligned when projected onto the RGB images. To reduce the risk of assigning incorrect labels to pixels, each class label is shrunk using a binary erosion operation. The eroded class labels are then transformed with the homography to overlap the RGB area. This process is illustrated in Figure 2, which shows the coastline in the RGB image from Figure 1, with land labels shown in green and sea labels shown in blue.

2.3. Removing segmentation noise

To ensure that segmentation of the RGB image is applied only to the relevant region, a coastline mask is created. This mask is defined by the intersection of the coastline detected in the hyperspectral class labels and the coastline available in the ground-truth dataset.

To create the coastline mask from the labels, both the land and sea class labels are dilated, and the intersection of the two masks is used as the detected coastline. The coastline ground-truth mask is generated by calculating the pixel coordinates of the ground-truth dataset and rasteriz-

ing the coastline into a mask. A linear interpolation is used to estimate pixel positions from geographical coordinates, and the coastline coordinates are then transformed into image coordinates. The Global Self-consistent, Hierarchical, High-resolution Shoreline (GSHHS) database is used as the coastline ground-truth dataset [11], and the coastline mask is created by rasterizing the calculated pixel positions.

Finally, both the ground-truth coastline mask and the hyperspectral coastline mask are dilated by the number of pixels to be included in the segmentation. This margin should be larger than the distance between the training labels and affects how many pixels are included as training samples for the RF classifier.

In principle, the coastline ground-truth mask is not strictly necessary and could be omitted by using only the boundary between the sea and land classes in the hyperspectral segmentation as the reference coastline. However, such an approach could introduce additional noise if the class labels are inaccurate. It may also lead to segmentation in regions where no corresponding ground control points exist, increasing computational effort without providing useful information for georeferencing corrections.

2.4. Segmenting coastline pixels in the RGB image

RF classifiers perform well for classifying high dimensional remote sensing data such as hyperspectral images [12]. They have also been used for cloud segmentation on RGB images with lower runtimes but lower accuracy and higher peak memory usage compared to CNN methods [13]. The proposed method requires the classifier to be trained on-board the satellite once the hyperspectral labels and RGB data are available, which adds computational cost compared with a pre-trained classifier. Given the limited onboard computational resources of HYPSO-2, the RF classifier is chosen for the experiments conducted in this study.

Prior to training and inference, intensity features are calculated for each pixel by applying Gaussian smoothing at multiple spatial scales. This is achieved by convolving the image with a series of Gaussian kernels whose sizes increase from a predefined minimum to a maximum value. This feature-extraction approach is equivalent to the Gaussian blur features used in Trainable Weka Segmentation [14], where Gaussian filters of varying kernel sizes are used to characterize image intensity patterns across multiple spatial resolutions.

The RF classifier is trained on the extracted features using the aligned and eroded class labels from the hyperspectral image as training labels. The mask of valid coastline pixels is used to select only the necessary samples. Pixels with overlapping training labels are used to train a random forest predictor, while pixels without training labels are used to predict class labels with the trained predictor. The random forest classifier used is the RTrees class from the OpenCV machine learning module.

Step	Average Time [s]
Preprocessing	0.32
Image registration	2.75
Ground-truth mask generation	3.23
RGB intensity feature extraction	2.69
Segmentation	2.81
Total runtime	11.79

Table 1: Average time for each processing step across all scenes.

3. RESULTS

The method proposed in Section 2 was evaluated on ground using seven different scenes captured by the HYPSO-2 satellite. The runtime of each experiment was measured, and the segmentation accuracy was evaluated against a manually georeferenced ground truth of the segmented area.

3.1. On-board implementation considerations

The methodology described in Section 2 is intended to run on-board the HYPSO-2 satellite. The satellite has limited computational resources, and any implementation running on it should be efficient and implemented in a compiled language. OpenCV is a highly optimized computer vision library that can be used with C++ [15], making it a strong candidate for an on-board implementation.

A Python program was created to perform the experiments described in this paper. All steps of the methodology, except the ground-truth mask creation, depend on OpenCV functions and NumPy. The ground-truth mask generation relies on additional Python libraries such as Shapely, GeoPandas, and scikit-learn, and would require more effort to implement in a compiled language for on-board use. This step also depends on a database of reference coastline data being available on board the satellite, which may be a challenge depending on the size of the database. The coastline ground-truth dataset used in this study is the GSHHS database [11], and the full resolution shapefile of ocean boundaries is approximately 154 MB. This size could be reduced by lowering the resolution of the dataset.

3.2. Runtimes and accuracy

All experiments were run on a Dell Pro 14 Plus PB14250. The runtime for each scene is shown in the top plot of Figure 3, and the number of processed pixels is shown in the bottom plot. The average processing times for the seven scenes are summarized in Table 1.

The segmentation accuracy is measured by calculating the proportion of correctly segmented pixels divided by the

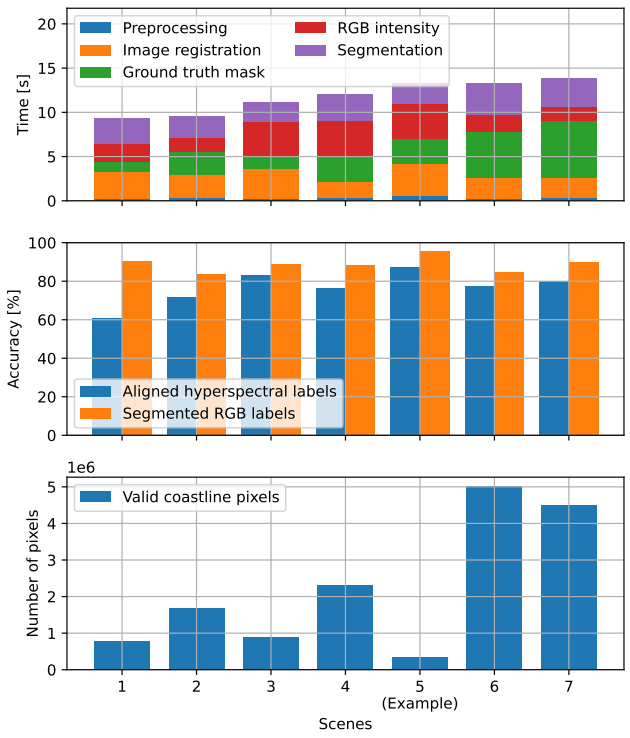
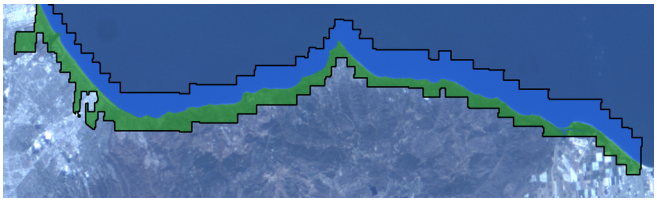
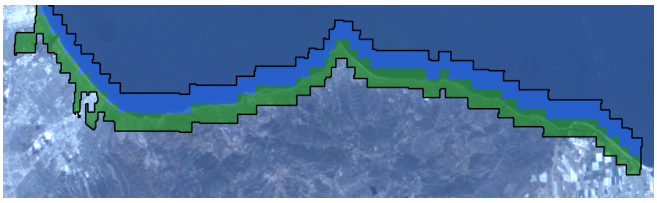


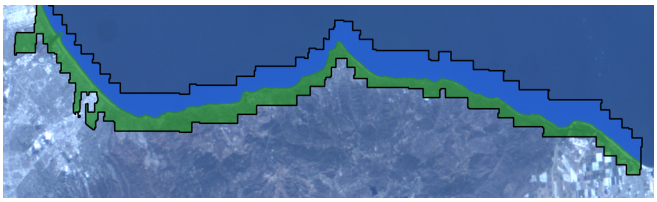
Fig. 3: Runtime for each processing step (top), segmentation accuracy (middle), and amount of processed pixels (bottom) for each scene.



(a) Manually georeferenced image with OpenStreetMap land and water data used as ground truth.



(b) Hyperspectral labels registered to the image.



(c) RF image segmentation produced by the proposed method.

Fig. 4: Visualization of the ground truth and labels over the selected coastline region in the RGB image.

total number of segmented pixels in a region with known ground truth. Areas covered by clouds are excluded, and the ground truth is created from OpenStreetMap land and water data aligned to the scene by manual georeferencing. The middle plot in Figure 3 shows the accuracy of the class labels around the coastline for each scene. The segmented class labels are compared with the hyperspectral class labels aligned to the RGB image. This comparison is visualized in Figure 4, where the class labels overlap scene 5, the example scene used throughout the paper. On average across all tested scenes, the accuracy improved by 12.22 %, with the average accuracy of the segmented RGB labels reaching 88.89 % and the aligned hyperspectral labels reaching 76.67 %.

4. DISCUSSION

The methodology presented in this paper is able to detect coastlines at a higher spatial resolution than the hyperspectral image, which could be used to georeference the RGB image with higher accuracy than the coastline solution based only on indirect georeferencing of the hyperspectral image. The class boundaries from the hyperspectral data are refined, resulting in an average improvement of 12.16 % in the segmented coastline region across the tested scenes. For missions such as HYPSO-2, where indirect georeferencing is performed by matching coastlines [5], more accurate coastline labels could therefore improve the final georeferencing accuracy.

The runtime is relatively consistent across the tested scenes, even though the number of processed pixels varies significantly. The experiments were implemented in Python using optimized libraries such as OpenCV and scikit-learn, but the runtime could be reduced with a more optimized implementation in a compiled language such as C++. One possible optimization would be to skip the step of creating the coastline mask from the ground-truth vector. Although this would reduce processing time, it could also introduce additional noise or increase the number of pixels to be segmented if the hyperspectral class labels are inaccurate or if coastline regions are present in the class labels but are not covered by the ground-truth data. Reducing the resolution of the RGB image would also decrease the runtime, but at the cost of lower spatial accuracy.

5. CONCLUSION

The methodology presented in this paper segments coastlines in RGB images by refining class labels generated from a corresponding hyperspectral image. This multimodal approach uses the high spectral resolution of hyperspectral imagery to generate reliable class labels and the high spatial resolution of RGB imagery to define more accurate coastline boundaries,

while avoiding the higher computational cost of more complex machine learning methods. With an increase of $12.16\ \%$ in segmentation accuracy over the segmented region, the refined labels are promising for improving indirect georeferencing accuracy in RGB images. This approach can help improve the quality or reduce the size of data products from EO satellites.

6. REFERENCES

[1] I. Benhadj, S. Livens, et al., "HyperScout-1 inflight calibration and product validation," *International Journal of Remote Sensing*, vol. 45, no. 7, pp. 2486–2517, Apr. 2024.

[2] C. Chiatante, D. D. Langer, et al., "Onboard Hyperspectral Classification Enables Georeferencing," in *2023 13th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Athens, Greece, Oct. 2023, pp. 1–5, IEEE.

[3] Y. Huang, S. You, Y. Fu, and Q. Shen, "Weakly-supervised Semantic Segmentation in Cityscape via Hyperspectral Image," July 2021, arXiv:2012.10122 [cs.CV].

[4] S. Boyle, P. Mikulowski, and M. Orlandić, "Onboard Accelerator of Hyperspectral Classification For HYPSO," in *2025 14th Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro, June 2025, pp. 1–4, IEEE.

[5] C. Chiatante, H. S. Løvås, et al., "Memory-efficient georeferencing of segmented satellite imagery," Submitted to IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2026.

[6] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[7] M. Ihmeida and H. Wei, "Image Registration Techniques and Applications: Comparative Study on Remote Sensing Imagery," in *2021 14th International Conference on Developments in eSystems Engineering (DeSE)*, Sharjah, United Arab Emirates, Dec. 2021, pp. 142–148, IEEE.

[8] P. Berg, S. Boyle, A. Alinejad, and M. Orlandić, "Wavelet Convolutional Neural Network for Hyperspectral Image Segmentation," in *2025 15th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Barcelona, Spain, Nov. 2025, pp. 1–5, IEEE.

[9] D. D. Langer, J. Garrett, et al., "The HYPSO RGB cameras," *75th International Astronautical Congress*, 2024.

[10] D. D. Langer, T. A. Johansen, and A. J. Sørensen, "Consistent Along Track Sharpness in a Push-Broom Imaging System," in *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, Pasadena, CA, USA, July 2023, pp. 4486–4489, IEEE.

[11] P. Wessel and W. H. F. Smith, "A global, self-consistent, hierarchical, high-resolution shoreline database," *Journal of Geophysical Research: Solid Earth*, vol. 101, no. B4, pp. 8741–8743, Apr. 1996.

[12] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24–31, Apr. 2016.

[13] S. Kacker, A. Meredith, G. Labrèche, and K. Cahoy, "Machine Learning Image Processing Algorithms Onboard OPS-SAT," .

[14] I. Arganda-Carreras, V. Kaynig, et al., "Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification," *Bioinformatics*, vol. 33, no. 15, pp. 2424–2426, Aug. 2017.

[15] K. Pulli, A. Baksheev, K. Kornyakov, and V. Eruhimov, "Real-time computer vision with OpenCV," *Communications of the ACM*, vol. 55, no. 6, pp. 61–69, June 2012.