

Image Processing Pipeline for Satellite and Aerial Image Enhancement

A. Kheskani, C. Scorția, P. Misterka, and V. Petkov

Group 13, Delft University of Technology

Abstract

This report presents an integrated satellite and aerial image processing pipeline addressing common remote sensing challenges. The system combines image stitching for creating seamless mosaics, cloud removal using a dual-branch deep learning architecture, multispectral analysis for vegetation and mineral mapping, single image super-resolution enhancement, semantic segmentation for land cover classification, interactive object removal and insertion, and depth extraction from LiDAR data. All components are unified in an interactive GUI enabling end-to-end processing workflows.

1. Introduction

Aerial surveying and satellite remote sensing generate large volumes of imagery that have downstream applications such as environmental monitoring, geological exploration, urban planning, and disaster response. However, noise, artifacts, varying weather conditions, and the sheer amount of data make it a difficult task.

We present an end-to-end enhancement pipeline that addresses these issues via a set of image processing steps. Section 2 explains and motivates our choice of test datasets, section 3 explains our image processing pipeline design in detail, and section 4 provides a short summary of achieved goals.

2. Datasets

This section describes the aerial and auxiliary datasets used in our project, summarizes their key characteristics, and explains our rationale for their selection.

2.1. AVIRIS hyperspectral surveys

Hyperspectral imaging captures a wider range of electromagnetic spectrum, usually for hundreds of spectral bands in visible, near, and mid-infrared range. This project uses NASA's AVIRIS dataset [GES*98] [NAS24] that provides partial coverage of North America over 284 reflectance bands with high spatial resolution. We use the L2A orthorectified reflectance product, which notably does not correct for atmospheric water-vapor absorption or topographic illumination effects.

2.2. USGS Spectrum Library

Materials absorb different portions of the electromagnetic spectrum differently. To detect different mineral deposits we can identify

Table 1: AVIRIS-NG L2A orthorectified reflectance chunk characteristics of the Fording River mine, B.C., Canada

Property	Value
Spatial dimensions	$\sim 3500 \times \sim 600$ pixels
Spatial sampling distance (SSD)	$5.8 \text{ m} \times 5.8 \text{ m}$
Approx. ground footprint	$\approx 19.4 \text{ km} \times 3.5 \text{ km}$
Spectral bands	425 bands
Spectral range	390–2510 nm
Spectral sampling interval (SSI)	5 nm
Radiance units	$\mu\text{W cm}^{-2} \text{ nm}^{-1} \text{ sr}^{-1}$

their spectral signatures in our hyperspectral images. The USGS Spectral Library provides reference for such spectral signatures of different minerals and compositions thereof. Out of the $\sim 2,500$ spectra samples with 5 purity classes from 3 different spectrometers we limit our reference to the top 3 purity classes and the ASDHR spectrometer.

2.3. Humans in the Loop: Semantic Segmentation of Ariel Imagery

The 'Dubai Satellite Imagery Dataset', found on Kaggle, consists of ariel imagery obtained by MBRSC satellites. The dataset consists of 72 high-resolution RGB satellite images and their corresponding semantic segmentation masks. These masks partitioned the terrain into 6 distinct classes: Water, Land, Road, Building, Vegetation and Unlabeled.

2.4. USGS 3DEP Lidar Data

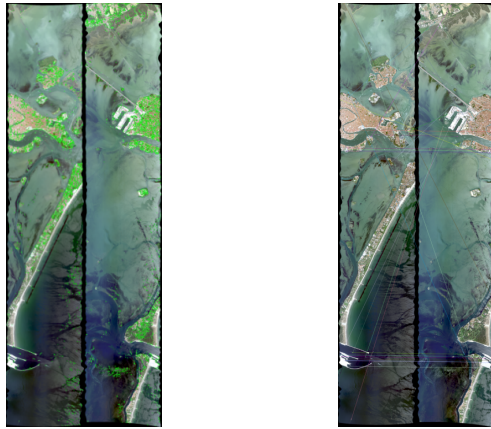
For depth and elevation extraction, we additionally use airborne LiDAR data [U.S24] provided by the USGS. The dataset consists of high-resolution point clouds capturing terrain elevation independent of illumination conditions. LiDAR measurements are processed into reastered depth maps, which serve as a geometric reference for downstream depth and 3D reconstruction tasks. Unlike hyperspectral imagery, LiDAR directly encodes surface height, enabling reliable elevation estimation over vegetation and urban structures.

3. Image processing pipeline

In this section we describe the individual components of our pipeline. Subsections 3.1-3.4 explain preprocessing steps for downstream tasks, and subsections 3.5-3.7 focus on hyperspectral image processing. Lastly, subsections 3.8-3.9 explain semantic segmentation and operations, and subsection 3.10 shows 3D depth estimation.

3.1. Image stitching

Image stitching is performed on hyperspectral scenes from AVIRIS dataset to construct spatially continuous mosaics. Since hyperspectral cubes contain hundreds of spectral bands, all geometric alignment is computed in RGB space and only later applied to the full spectral data when stitching exactly two images. This significantly reduces computational overhead while preserving spectral integrity.

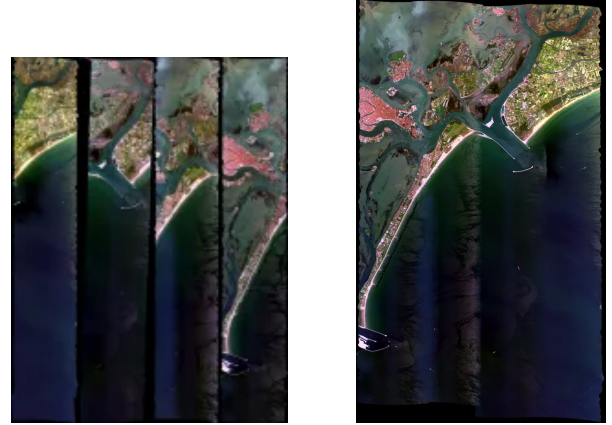


(a) Detected SIFT keypoints (b) kNN matches

Figure 1: Feature detection using SIFT.

Keypoints are detected using the Scale-Invariant Feature Transform (SIFT) [Low04] (Fig. 1a), applied to contrast-enhanced RGB projections of each hyperspectral image. Descriptor matching is performed using k-nearest neighbors with $k = 2$ (Fig. 1b), followed by Lowe’s ratio test [Low04] to reject ambiguous correspondences. A match is accepted only if the ratio between the distances to the closest and second-closest descriptors satisfies

$$\frac{d_1}{d_2} < 0.7.$$



(a) The original AVIRIS images

(b) The stitched result

Figure 2: The stitched result

This threshold, consistent with our implementation, removes weak matches in low-texture regions. The remaining correspondences are then used to estimate a homography via RANSAC [FB81], which robustly rejects outliers and wraps one image into the coordinate frame of the other.

For multiple images (Fig. 2), a connectivity graph is constructed based on geospatial proximity derived from ENVI metadata. Edges represent valid pairwise homographies, and a spanning structure global alignment while minimizing accumulated error. Overlapping regions are blended using distance-based feathering to suppress visible seams.

3.2. Cloud Removal and Inpainting

We developed a dual-branch cloud removal architecture for multispectral satellite imagery, addressing the challenge that thin and thick clouds require fundamentally different reconstruction strategies.

Architecture. As shown in Figure 3, an Opacity Estimation Head first predicts per-pixel cloud opacity $\alpha \in [0, 1]$. This routes inputs to specialised branches: NAFNet [CCZS22] with FFT frequency filtering for thin clouds (where ground features remain partially visible), and SwinIR [LCS*21] with SAR cross-attention for thick clouds (leveraging radar’s cloud-penetrating capability). The final output blends both branches:

$$\hat{y} = \alpha \cdot f_{\text{thick}}(x, x_{\text{SAR}}) + (1 - \alpha) \cdot f_{\text{thin}}(x) \quad (1)$$

Training Strategy. We employ three-phase training on SEN12MS-CR [EMSZ22]: (1) pre-train branches independently with frozen routing, (2) progressively unfreeze layers while training the opacity head with routing supervision to prevent collapse, (3) end-to-end fine-tuning with combined loss:

$$\mathcal{L} = \mathcal{L}_1 + 0.1 \cdot \mathcal{L}_{\text{SSIM}} + 0.01 \cdot \mathcal{L}_{\text{NLL}} \quad (2)$$

where $\mathcal{L}_{\text{NLL}}$ enables uncertainty quantification.

Results. Table 2 shows quantitative improvements over single-branch baselines. Figure 4 demonstrates successful reconstruction



Figure 3: Dual-branch cloud removal architecture with opacity-based soft routing.

across varying cloud densities on AVIRIS hyperspectral data via spectral convolution to simulated Sentinel-2 bands.

3.3. Object Removal

Our object removal pipeline combines precise segmentation with context-aware inpainting. Users draw bounding boxes around targets; Segment Anything Model (SAM) [KMR*23] generates pixel-accurate masks. We then apply DiffusionSat inpainting with crop-based processing—extracting only the masked region plus context padding, processing at native resolution, and seamlessly blending back:

$$I_{\text{out}} = M_f \cdot I_{\text{inpaint}} + (1 - M_f) \cdot I_{\text{orig}} \quad (3)$$

where M_f is a feathered mask for smooth transitions. Figure 5 shows vehicle removal from aerial imagery.

3.4. Object Insertion

For realistic object compositing, we implement Poisson blending [PGB03] to seamlessly integrate source objects into target scenes by solving:

$$\min_f \iint_{\Omega} |\nabla f - \mathbf{v}|^2 \text{ with } f|_{\partial\Omega} = f^* \quad (4)$$

where $\mathbf{v}$ is the guidance field from the source. Figure 5 demonstrates insertion of structures into satellite imagery with automatic colour and lighting harmonisation.

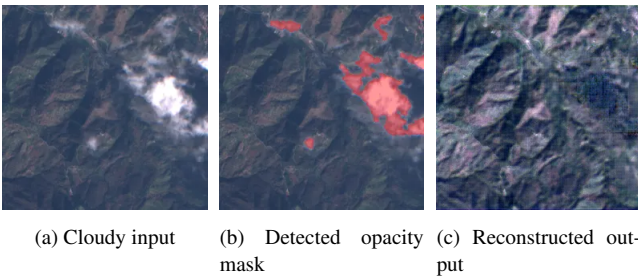


Figure 4: Cloud removal pipeline.

3.5. Hyperspectral image processing

We preprocess the stitched hyperspectral images by removing noisy mid-infrared bands (due to atmospheric moisture) and normalizing band values to remove artifact outliers. We then produce various multispectral indices (see Fig. 6 and Table 3) for downstream tasks.

Table 2: Cloud removal quantitative evaluation on SEN12MS-CR test set.

Method	PSNR $\uparrow$	SSIM $\uparrow$
NAFNet only	26.4	0.82
SwinIR only	27.1	0.84
Ours (dual-branch)	28.9	0.87

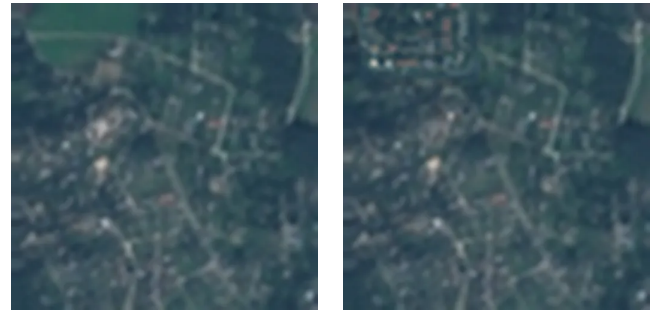
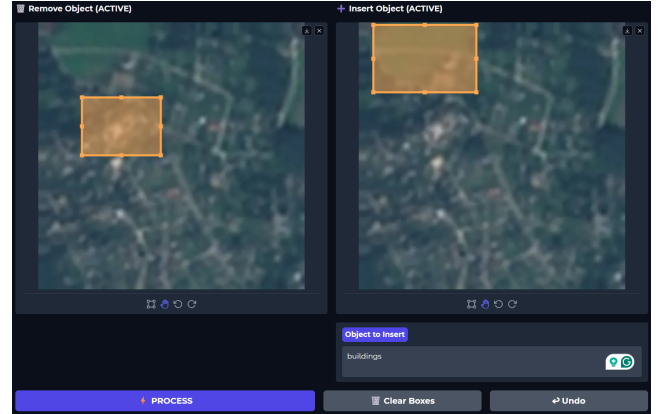


Figure 5: Object removal and insertion pipeline. Top: GUI interface showing user-defined regions for removal (left) and insertion with prompt “buildings” (right). Bottom left: Original input. Bottom right: Processed result.

Vegetation detection We provide the Normalized Difference Vegetation Index (NDVI) Normalized Difference Red Edge Index NDRE indices and the TCARI/OSAVI ratio. By approximating plant chlorophyll content, these indices suggest vegetation strength at different stages of plant lifecycles.

Water body detection We use the Hyperspectral Difference Water Index (HDWI) to detect water bodies. Compared to the well-established Normalized Difference Water Index (NDWI), HDWI works outside noisy SWIR bands around $\sim 1450\text{nm}$, and in our empirical tests provides drastically more precise approximation.

True color We provide a true color approximation from hyperspectral bands with gray balance and tunable gamma/saturation parameters.



Figure 6: Multispectral indices computed from the AVIRIS-NG aerial images. From left: true color approximation, NDVI, NDRE, HDWI, and TCARI/OSAVI. Brighter colors indicate higher index value. Note that TCARI/OSAVI properly identifies region’s dry vegetation, and HDWI discerns the lake near the top of the image.

Index	Formula
NDVI	$\frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R}$
NDRE	$\frac{\rho_{NIR} + \rho_R}{\rho_{NIR} - \rho_{RE}}$
HDWI	$\frac{\rho_{NIR} + \rho_{RE}}{\rho_G - \rho_{NIR}}$
OSAVI	$1.16 \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R + 0.16}$
TCARI	$3 \left[(\rho_{700} - \rho_{670}) - 0.2(\rho_{700} - \rho_{550}) \frac{\rho_{700}}{\rho_{670}} \right]$

Table 3: Multispectral index bands for hyperspectral reflectance for vegetation and water-body detection.

3.6. Hyperspectral single image super-resolution

We implement a hyperspectral version of single-image Zero Shot Super-Resolution (ZSSR) [SCI17] paper with visualizations in Figure 7. ZSSR produces pairs of downsampled low- and high-resolution (LR-HR) pairs of the image to be upscaled. We leverage a multiscale pyramid to progressively train on repeatedly downsampled LR-HR pairs (i.e. x4-x8 midway training) to further improve accuracy. To allow training on hyperspectral images, we make the following modifications to the base ZSSR:

3D convolutions As hyperspectral images are 3-dimensional, we replace the original 2D convolutions with 3D ones.

Trilinear interpolation To account for cross-corellation across different bands we use trilinear interpolation on the entire cube rather than bicubic interpolation per-band.

Spectral loss As we aim to preserve pixel’s spectral response used in section 3.7, we implement an ensemble loss based on the Spectral Angle Mapper (SAM) [YGB92]

$$\text{SAM}(r, s) = \arccos \left(\frac{r^\top s}{\|r\|_2 \|s\|_2} \right)$$

$$\mathcal{L}(\hat{x}, x) = \frac{1}{15} \mathcal{L}_1(\hat{x}, x) + \frac{14}{15} \mathcal{L}_{\text{SAM}}(\hat{x}, x)$$

minimizing spectral angle in highly-dimensional hyperspectral space, and decreasing band error variance.

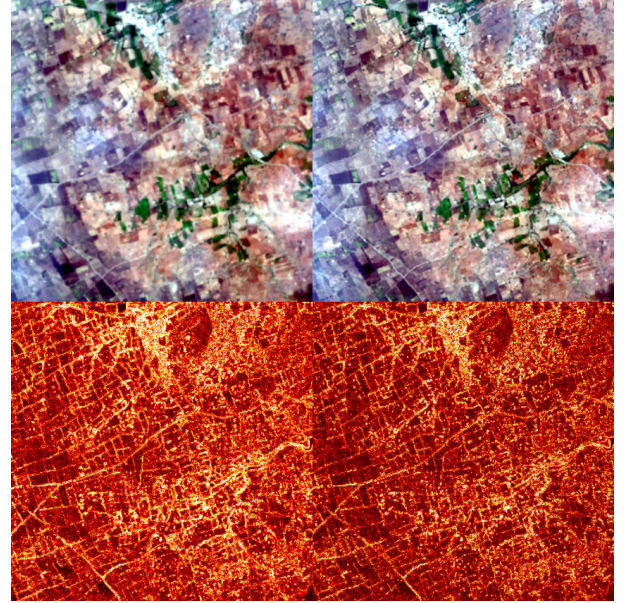


Figure 7: Hyperspectral single image super-resolution comparison and PSNR (peak signal to noise ratio) using trilinear interpolation (left, 24.23dB) and ZSSR (right, 25.57dB) - higher is better. Heat map in the bottom row shows the corresponding error between original and upscaled (brighter is worse). Notice the difference in sharpness near the bottom green area and cities.

3.7. Mineral deposit identification

To identify rock and mineral composition of surveyed areas we compare preprocessed hyperspectral images with laboratory spectra samples from the USGS Spectrum Library.

Band alignment As reference band wavelengths differ from AVIRIS wavelengths, we first convolve material spectra using gaussian bandpass approximation

$$S_i(\lambda) = \exp \left(-\frac{(\lambda - c_i)^2}{2\sigma_i^2} \right), \quad \sigma_i = \frac{\text{FWHM}_i}{2\sqrt{2\ln 2}}.$$

where c_i is the band wavelength center and FWHM_i is the width at half signal strength.

Masking We use indices from section 3.5 to mask vegetation and water bodies from mineral detection.

Spectrum comparison We then measure the difference between the aerial spectrum response and the convolved material spectrum for every pixel in the hyperspectral image using SAM as seen in Figure 8. As SAM runs in $\mathcal{O}(n_{\text{mat}} \cdot n_{\text{bands}} \cdot w \cdot h)$, our implementation provides multithreaded matrix-based batch processing version of SAM that yields $\sim 1000x$ speedup for large images. We note that a limitation of our approach is its inability to distinguish mineral mixtures due to lack of on-site reference spectra.

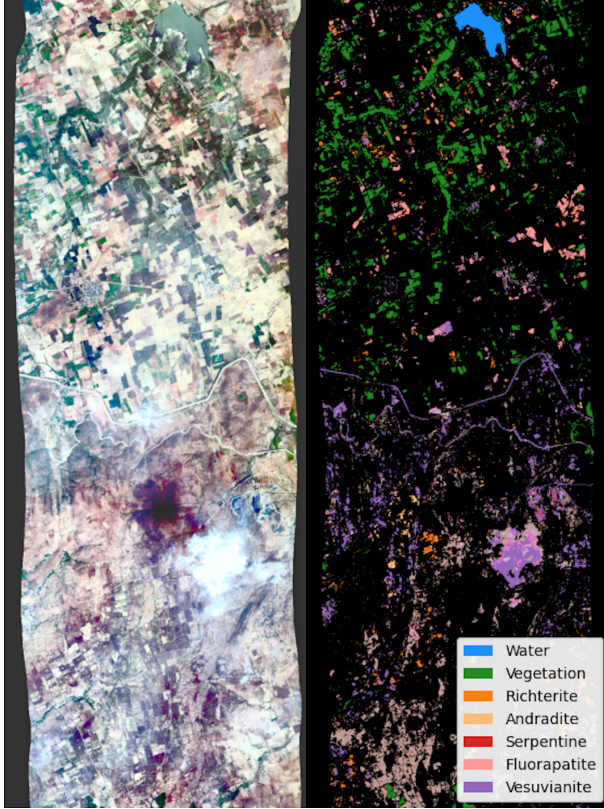


Figure 8: Mineral deposit mask using SAM (right). Top 5 minerals and vegetation and water masks are shown. We intentionally do not remove clouds to demonstrate their misclassifications.

Additionally, we provide a single-pixel function for spectra match comparison using Spectral Feature Fitting (SFF) [CGS90]. We find SFF to be computationally prohibitively expensive to run on the entire image.

$$\text{SFF}(r_c, s_c) = \|r_c - a^* s_c\|_2 \quad a^* = \frac{s_c^\top r_c}{s_c^\top s_c},$$

3.8. Semantic Segmentation

A U-Net architecture was implemented [RFB15] to perform the semantic segmentation, originally utilized and designed for biomedical imaging. We used its skip connections to preserve context which

is essential for remote sensing. This design choice has recently been validated by Goswami et al. (2025) [GDM*25], who demonstrated U-Net’s powerful performance in identifying topographical land-forms from satellite data.

The model was trained using Cross-Entropy Loss to treat the segmentation as a pixel-wise classification task allowing to distinguish between different decision boundaries. The Adam optimizer (learning rate = 0.01) was also needed for ensuring convergence to a low final loss within a constrained training budget of 100 epochs. The resize approach performs a global spatial compression of the input tensor, but for images that are larger we use a sliding window approach.

As shown in Figure 9, the model is able to delineate complex structure, accurately.

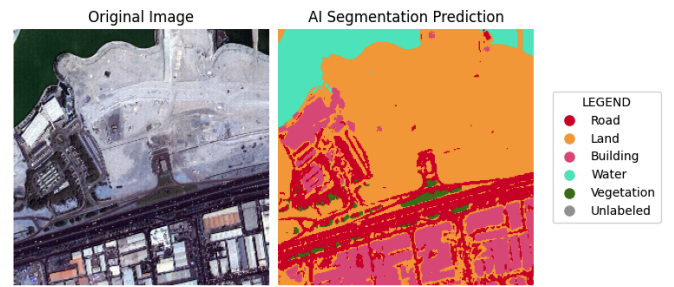


Figure 9: Semantic Segmentation

Testing with new images, the model misclassified certain labels where classes shared similar spectral characteristics. The model struggled to distinguish between water and vegetation, and between roads and buildings. These misclassifications primarily attributed to the limited diversity and number of images. (Only RGB inputs are accepted as hyperspectral images brought many complications)

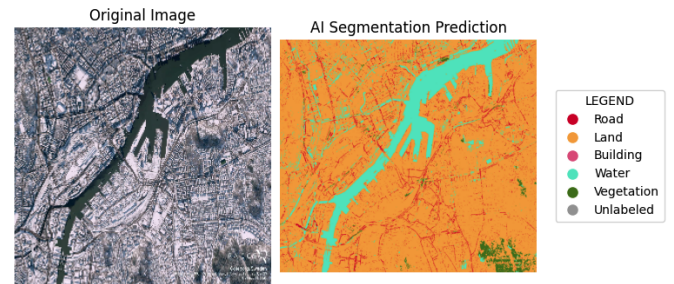


Figure 10: Semantic Segmentation (Test)

3.9. Semantic Operations

We leveraged our understanding and use of the U-Net model by implementing a suite of Semantic Guided Operations.

1. **Semantic Recoloring (Data Augmentation)** We developed a tool that blends user-defined RGB values onto specific segmentation

classes, which consist of: Road, Land, Building, Water, Vegetation and Unlabeled. This operation, which was a concept explored by Toker et al. [TECLT24].

Sakaridis et al. [SDVG18] established a mathematical linear blending formula which we adopted. While their work focused on blending atmospheric light to simulate fog, we generalized their equation to blend arbitrary colors onto target classes:

$$I_{out} = I_{orig} \cdot (1 - \alpha) + C_{target} \cdot \alpha$$

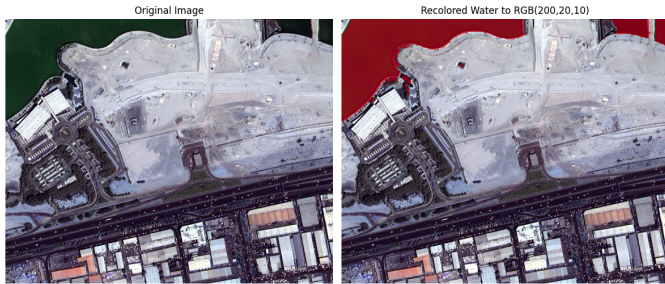


Figure 11: Semantic Recoloring on water with intensity 0.7

2. **Saliency Visualization** The implementation of this operation was chosen to be able to verify the model's focus. This operation converts the image to grayscale while preserving the color of a single chosen (active) class. Isolating spatial elements is essential for understanding deep learning in remote sensing [LSL*23]. By removing context (background color), allows for immediate qualitative assessment of the segmentation boundaries, acting as a visualization attention mechanism.

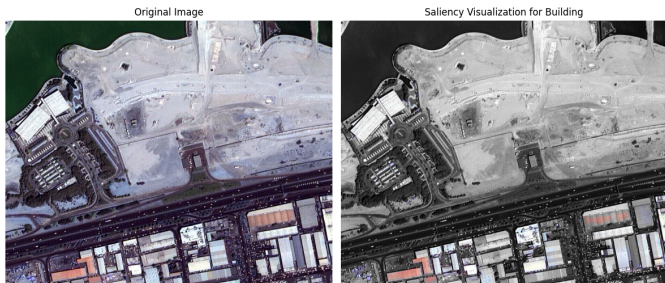


Figure 12: Saliency Visualization for Buildings

3.10. Depth/ 3D Extraction

Depth extraction (Fig. 13) combines airborne LiDAR with hyperspectral imagery to recover both terrestrial elevation and shallow-water bathymetry. Above-water geometry is derived from LiDAR point clouds by generating Digital Surface Models (DSM) and Digital Terrain Models (DTM), whose difference yields a Canopy Height Model (CHM) (Fig. 13b). For underwater regions, depth is estimated using a Spectrally Derived Bathymetry (SBD) approach based on hyperspectral reflectance (Fig. 13d). Water bodies are first identified using the HDWI (Table 3, Fig. 13c) which enables robust

making of aquatic regions. Bathymetry is then estimated using the Strumpf band-ratio method [SHS03]

$$D \propto \frac{\log(\rho_{blue})}{\log(\rho_{green})},$$

with blue and green bands centered at 480 nm and 550 nm, respectively. The resulting relative depths are normalized to a fixed range and fused with the LiDAR-derived CHM, producing a continuous elevation surface across land-water boundaries.

4. Summary

We presented an integrated satellite image processing pipeline combining hyperspectral analysis, deep learning restoration, and semantic operations. Key contributions include: image stitching with spectral integrity preservation for hyperspectral mosaics; multispectral index computation (NDVI, NDWI, BSI) and mineral identification via accelerated SAM matching; single image super-resolution enhancement for improved spatial detail; a dual-branch cloud removal architecture using NAFNet and SwinIR with learned opacity routing; SAM-guided object editing with DiffusionSat inpainting for removal and insertion; U-Net semantic segmentation enabling class-specific recolouring and saliency visualisation; and LiDAR-bathymetry fusion for continuous elevation mapping. All components are unified in an interactive Gradio GUI, enabling end-to-end processing for remote sensing applications.

References

- [CCZS22] CHEN L., CHU X., ZHANG X., SUN J.: Simple baselines for image restoration. In *European Conference on Computer Vision (ECCV)* (2022), Springer, pp. 17–33. [2, 8](#)
- [CGS90] CLARK R. N., GALLAGHER A. J., SWAYZE G. A.: Material absorption band depth mapping of imaging spectrometer data using a complete band shape least-squares fit with library reference spectra. In *AVIRIS Proceedings June 4–5, 1990 (JPL Publication 90-54)* (Pasadena, CA, USA, June 1990), Jet Propulsion Laboratory (JPL), pp. 176–186. Accessed: 2026-01-29. https://aviris.jpl.nasa.gov/proceedings/workshops/90_docs/17.PDF. [5, 8](#)
- [DBMR07] DALTON J. B., BOVE D. J., MLADINICH C. S., ROCKWELL B. W.: Imaging spectroscopy applied to the animas river watershed and silverton caldera. In *Integrated Investigations of Environmental Effects of Historical Mining in the Animas River Watershed, San Juan County, Colorado*, Church S. E., von Guerard P., Finger S. E., (Eds.), no. 1651 in Professional Paper. U.S. Geological Survey, Reston, VA, 2007, ch. E2, pp. 143–159. Version 1.0. Accessed: 2026-01-29. https://pubs.usgs.gov/pp/1651/downloads/Vol1_combinedChapters/vol1_chapE2.pdf. [doi:10.3133/pp1651. 8](#)
- [EMSZ22] EBEL P., MERANER A., SCHMITT M., ZHU X. X.: SEN12MS-CR: A multi-modal multi-temporal dataset for cloud removal in sentinel-2 imagery. *IEEE Transactions on Geoscience and Remote Sensing* (2022). [doi:10.1109/TGRS.2022.3180405. 2, 8](#)
- [FB81] FISCHLER M. A., BOLLES R. C.: Random sample consensus: A paradigm for model fitting. *Communications of the ACM* 24, 6 (1981), 381–395. [2](#)
- [GDM*25] GOSWAMI M., DEY S., MUKHERJEE A., MOHANTY S., PATTAIAK P. K.: Convolutional neural network segmentation for satellite imagery data to identify landforms using u-net architecture, 2025. URL: <https://arxiv.org/abs/2502.05476>, [arXiv:2502.05476. 5](#)

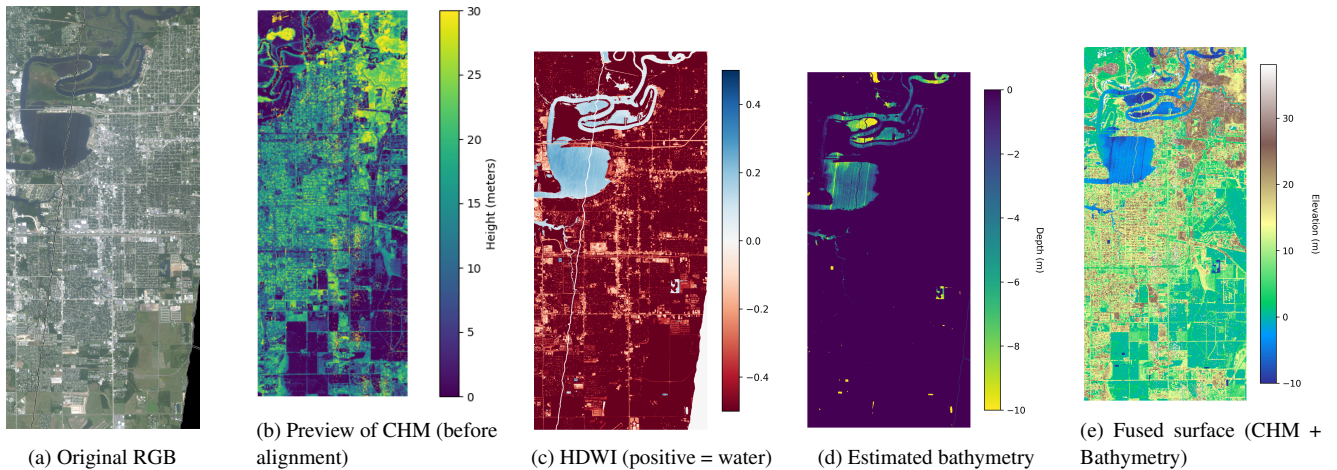


Figure 13: Depth/3D extraction process

- [GES*98] GREEN R. O., EASTWOOD M. L., SARTURE C. M., CHRIEN T. G., ARONSSON M., CHIPPENDALE B. J., FAUST J. A., PAVRI B. E., CHOVIT C. J., SOLIS M., OLAH M. R., WILLIAMS O.: Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris). *Remote Sensing of Environment* 65, 3 (Sept. 1998), 227–248. doi:10.1016/S0034-4257(98)00064-9. 1
- [KMR*23] KIRILLOV A., MINTUN E., RAVI N., MAO H., ROLLAND C., GUSTAFSON L., XIAO T., WHITEHEAD S., BERG A. C., LO W.-Y., DOLLAR P., GIRSHICK R.: Segment anything. *arXiv preprint* (2023). URL: <https://arxiv.org/abs/2304.02643>, arXiv:2304.02643. 3, 8
- [LCS*21] LIANG J., CAO J., SUN G., ZHANG K., VAN GOOL L., TIMOFTE R.: Swinir: Image restoration using swin transformer. In *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)* (2021), pp. 1833–1844. 2, 8
- [Low04] LOWE D. G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, 2 (2004), 91–110. 2
- [LSL*23] LV J., SHEN Q., LV M., LI Y., SHI L., ZHANG P.: Deep learning-based semantic segmentation of remote sensing images: a review. *Frontiers in Ecology and Evolution Volume 11 - 2023* (2023). URL: <https://www.frontiersin.org/journals/ecology-and-evolution/articles/10.3389/fevo.2023.1201125>, doi:10.3389/fevo.2023.1201125. 6
- [NAS24] NASA JET PROPULSION LABORATORY: Aviris – airborne visible / infrared imaging spectrometer. Accessed: 2026-01-12. <https://aviris.jpl.nasa.gov/>. 1
- [PGB03] PÉREZ P., GANGNET M., BLAKE A.: Poisson image editing. *ACM Transactions on Graphics (SIGGRAPH)* 22, 3 (2003), 313–318. doi:10.1145/882262.882269. 3
- [RFB15] RONNEBERGER O., FISCHER P., BROX T.: U-net: Convolutional networks for biomedical image segmentation, 2015. URL: <https://arxiv.org/abs/1505.04597>, arXiv:1505.04597. 5
- [SCI17] SHOCHER A., COHEN N., IRANI M.: "Zero-Shot" Super-Resolution using Deep Internal Learning, Dec. 2017. URL: <https://arxiv.org/pdf/1712.06087>, arXiv:1712.06087, doi:10.48550/arXiv.1712.06087. 4, 8
- [SDVG18] SAKARIDIS C., DAI D., VAN GOOL L.: Semantic foggy scene understanding with synthetic data. *International Journal of Computer Vision* 126, 9 (Mar. 2018), 973–992. URL: <http://dx.doi.org/10.1007/s11263-018-1072-8>, doi:10.1007/s11263-018-1072-8. 6
- [SHS03] STUMPF R. P., HOLDERIED K., SINCLAIR M.: Determination of water depth with high-resolution satellite imagery over variable bottom types. *Limnology and Oceanography* 48, 1 (2003), 547–556. 6
- [TECLT24] TOKER A., EISENBERGER M., CREMERS D., LEAL-ITAIXÉ L.: Satsynth: Augmenting image-mask pairs through diffusion models for aerial semantic segmentation, 2024. URL: <https://arxiv.org/abs/2403.16605>, arXiv:2403.16605. 6
- [TGC20] TRIPATHI M. K., GOVIL H., CHATTORAJ S. L.: Identification of hydrothermal altered/weathered and clay minerals through airborne aviris-ng hyperspectral data in jahajpur, india. *Heliyon* 6, 2 (Feb. 2020), e03487. PMCID: PMC7047179; PMID: 32140599. Accessed: 2026-01-29. <https://pmc.ncbi.nlm.nih.gov/articles/PMC7047179/>. doi:10.1016/j.heliyon.2020.e03487. 8
- [U.S24] U.S. GEOLOGICAL SURVEY: Usgs 3d elevation program (3dep), 2024. Accessed 2026-01. URL: <https://www.usgs.gov/3d-elevation-program>. 2
- [YGB92] YUHAS R. H., GOETZ A. F. H., BOARDMAN J. W.: Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm. In *Summaries of the Third Annual JPL Airborne Geoscience Workshop, Volume 1: AVIRIS Workshop* (June 1992), pp. 147–149. JPL-PUBL-92-14-VOL-1; NASA-CR-194540; NTRS Document ID: 19940012238. 4, 8

Student	Implemented features and its explanation	What I implemented myself	Where and how I used external sources	Where I present in video
Aryan Kheskani Raisinghani	A deep learning-based semantic segmentation pipeline utilizing a U-Net architecture with sliding window inference for high-resolution imagery. Integrated semantic guided operations include semantic recoloring and saliency visualization for class-specific analysis.	Developed the U-Net model architecture and established the training workflow using Cross-Entropy Loss. Developed the custom sliding window inference logic to handle large-scale satellite data and implemented the semantic guided operations module for real-time mask manipulation	Gemini was used to help plan/brain storm algorithms, and also helped in understanding the GUI in gradio. OpenCV was used for processing inputs and creating effects for the guided operations.	02:15
Claudia Scorția	Image stitching of hyperspectral aerial images and depth / 3D extraction using LiDAR and hyperspectral data. Image stitching enables spatially continuous mosaics from multiple AVIRIS scenes, while depth extraction recovers terrain elevation and shallow-water bathymetry by fusing LiDAR-derived surface models with spectrally derived depth estimates.	I implemented SIFT-based feature detection and matching, RANSAC homography estimation, and image blending for stitching. I also implemented LiDAR-based DSM/DTM/CHM generation, water masking using HDWI, spectrally derived bathymetry, and fusion into a single elevation surface.	Academic references informed stitching and bathymetry methods. OpenCV was used for feature detection and geometric transformations. Gemini was used for debugging, and to improve interface.	00:00 (Image stitching) and 03:42 (3D/Depth Extraction)
Paul Misterka	(1) Hyperspectral preprocessing for vegetation and water body detection. (2) Hyperspectral zero-shot single image super-resolution for better ground-level detail. (3) Spectral angle (SAM) mineral mapping with hyperspectral masking for geological surveys (e.g. mining ore content).	(1) Implemented hyperspectral image loading, pre-processing, true color visualization, and 4 multispectral vegetation and water content indices. (2) Implemented a zero-shot 3D convolutional model (ZSSR) with progressive multi-scale pyramid training, variable zoom factor, adapted to hyperspectral imaging with an ensemble L_1 -SAM loss. (3) Implemented multi-threaded batch Spectral Angle Mapping for hyperspectral mineral identification with water and vegetation masking, convolutional band mapping, and bonus comparison against Spectral Feature Fitting.	Single image super-resolution (upscaling) is based on ZSSR paper [SCI17]. Spectral fitting algorithms are based on [CGS90] and [YGB92]. Overall approach to mineral mapping is based on [TGC20] and [DBMR07]. Additionally, used ESA STEP platform for reference band ranges. Used GPT5.2 for brainstorming and help with plot visualization. Used pytorch for model training and trilinear interpolation, scikit for LAB color space transformation and convex hull algorithm, and own project for pipeline design.	04:30
Vladimir Petkov	Cloud Removal Inpainting using a dual-branch architecture combining NAFNet (thin clouds) and SwinIR (thick clouds) with learned opacity-based routing, and Object Editing for removal/insertion using DiffusionSat inpainting with SAM-based segmentation from bounding boxes.	I designed and trained the dual-branch cloud removal model on SEN12MS-CR data, implemented the routing network for cloud opacity estimation, integrated SAM for bounding-box-to-mask conversion, built the DiffusionSat inpainting pipeline for object editing, and created the GUI integration for both modules.	NAFNet [CCZS22] and SwinIR [LCS*21] architectures were adapted for the dual-branch design. Diffusion-Sat was used for satellite-aware inpainting. SAM (Segment Anything) [KMR*23] was used for segmentation. SEN12MS-CR [EMSZ22] dataset was used for training. Claude was used for debugging and code assistance with Gradio.	00:43

Table 4: Group contribution