Learned Image Compression for Earth Observation: Implications for Downstream Segmentation Tasks

Authors

Christian Mollière[1], Iker Cumplido[2], Marco Zeulner[2], Lukas Liesenhoff[1], Matthias Schubert[2], Julia Gottfriedsen[1]

Affiliations

[1]OroraTech GmbH, St.-Martin-Str. 112-114, 81669 Munich, Germany

[2]LMU Munich, Oettingenstr. 67, 80538 München,Germany

The rapid growth of data from satellite-based Earth observation (EO) systems poses significant challenges in data transmission and storage. We evaluate the potential of task-specific learned compression algorithms in this context to reduce data volumes while retaining crucial information. In detail, we compare traditional compression (JPEG 2000) versus a learned compression approach (Discretized Mixed Gaussian Likelihood) on three EO segmentation tasks: Fire, cloud, and building detection. Learned compression notably outperforms JPEG 2000 for large-scale, multi-channel optical imagery in both reconstruction quality (PSNR) and segmentation accuracy. However, traditional codecs remain competitive on smaller, single-channel thermal infrared datasets due to limited data and architectural constraints.

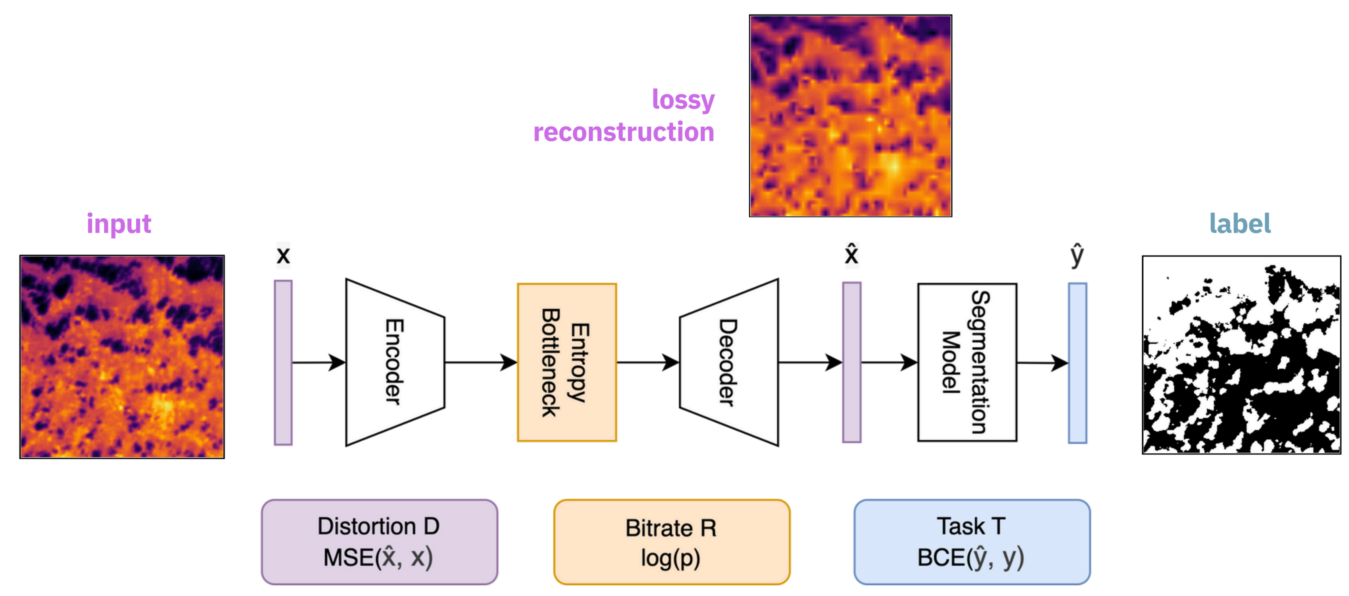

Figure 1 Overview of experimental setup. The compression model (auto-encoder + bottleneck) on the left side, and the segmentation model on the right.

Relevance for Cubesat Platforms

Commercial cubesat platforms operate under significant constraints. Although tasks can be executed on the edge, there is a necessity for raster data product representations to facilitate debugging and verification. This ensures the maintenance of data quality and safety, particularly for tasks where safety is a critical concern.

Experimental Setup

Figure 1 shows an overview of the model architecture used for this work. It concatenates a compression module with a segmentation model. We opt for segmentation on the reconstructed input instead of its embedding to conserve inter-operability with traditional compression methodologies like JPEG.

Benchmark Datasets and Task Baselines

For the purpose of this study we focus on relevant tasks for disaster and emergency response, with low count of spectral features (1-3 channels). **Figure 2** shows exemplary samples of each analysed task. The tasks can be split in two groups, single-channel and multi-channel tasks. All three datasets were evaluated using a baseline segmentation model (UNet + pretrained ResNet34 backbone).

- 🔥 **Fire Detection Dataset** (1 ch)
  - *(Rötzer et al. [2025])*
  - using MWIR @ 3.8 µm at 200 m GSD
- ☁️ **Cloud Segmentation Dataset** (1 ch)
  - *(Wölki et al. [2024])*
  - using LWIR @ 11.5 µm at 200 m GSD
- 🏢 **Building Segmentation** (3 ch)
  - *(Prexl et al. [2023])*
  - using Optical RBG at 10 m GSD

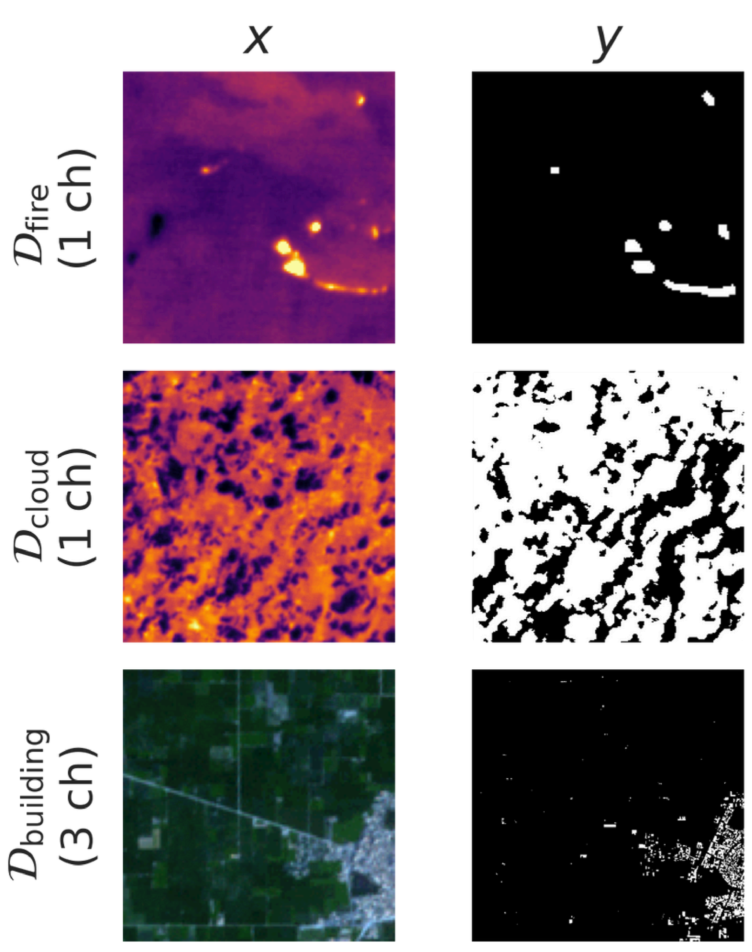

Figure 2 Examples of the used dataset.

Segmentation Performance under Bitrate Reduction

Figure 3 shows the impact of bit rate reduction on segmentation performance (F1), when training on the decompressed image data. This is done to directly quantify the effect of compression on our selection of downstream tasks. For all task, we observe that a certain level of image quality is needed before reaching a plateau in segmentation performance. This is in agreement with existing work, as many segmentation tasks in EO do not seem to require the full fidelity of the used observation data *(Garcia-Sobrino et al. [2020])*.

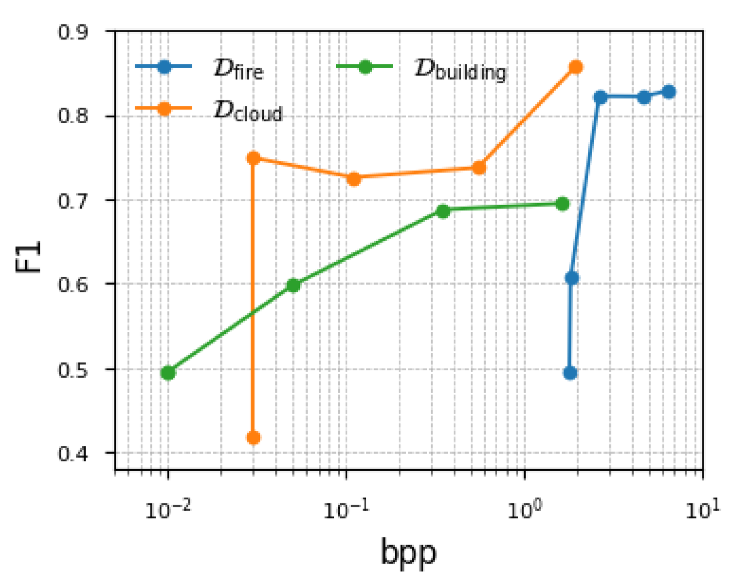

Figure 3 Segmentation performance of benchmarks under bitrate reduction using the JPEG 2000 baseline.

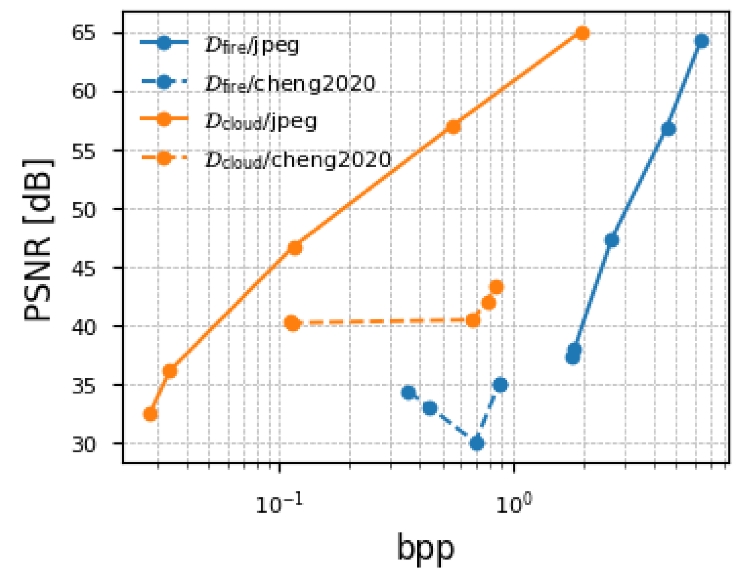

Figure 4 Performance of standalone compression algorithms on single-channel tasks.

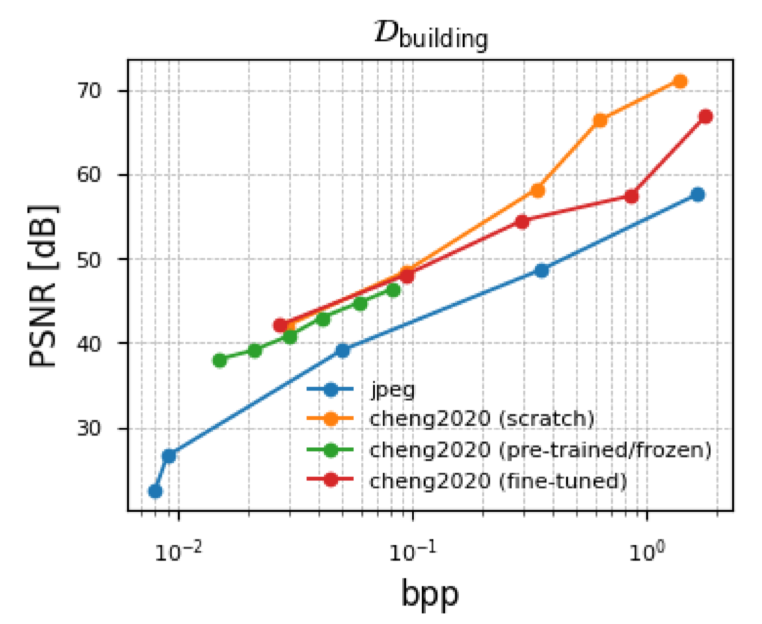

Figure 5 Performance of standalone compression algorithms on multi-channel task.

Standalone Performance of Learned Compression

Figure 4 shows the results on the two single-channel tasks fire detection and cloud segmentation. JPEG2000 remains superior in the monochromatic tasks, despite testing multiple learning architectures including FullyFactorizedPrior, ScaleHyperPrior *(Ballé et al. [2018])* and the Cheng2020Anchor *(Cheng et al. [2020])*. We accredit this result due to the lack of redundant spectral information that is usually present in multi-spectral datasets.

Figure 5 summarizes the performance on the multi-channel task of building segmentation. Here, the learned compression algorithm clearly outperforms the JPEG baseline both, when trained from scratch or fine-tuned. The pretrained models were created by training on RGB still frames of Vimeo90k dataset, which represent a different modality compared to the remote sensing data *(Xue et al. [2017])*.

Figure 6 shows the distribution of symbols within the three benchmark datasets, as well as for the Kodak dataset (Kodak [1993]) which is used as a common benchmark for evaluating compression algorithms. The building and fire dataset show a strong bias towards lower values with a lower overall entropy and less dynamic range when compared to Kodak.

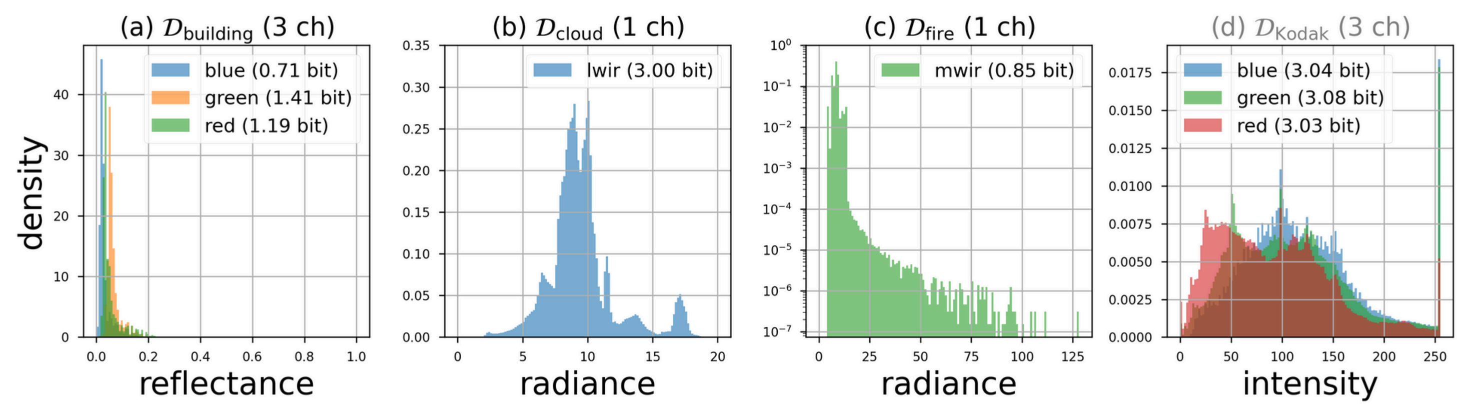

Figure 6 Distribution of unique symbols in each dataset.

Conclusion

- The analysed benchmarks can still be solved operationally using a mild bitrate reduction.
- The performance of learned compression algorithms translates well to multi-channel EO tasks.
- Traditional codecs outperform learned compression methodologies on tested single-channel tasks.

References

- (Rötzer et al. [2025]), Self-Supervised Learning for Fire Segmentation in FOREST-2 Images, ESA–NASA Workshop on AI Foundation Models for EO
- (Wölki et al. [2024]), Exploring Machine Learning for Cloud Segmentation in Thermal Satellite Images of the FOREST-2 Mission, SPACE 2024
- (Prexl et al. [2023]), The Potential of Sentinel-2 Data for Global Building Footprint Mapping with High Temporal Resolution, JURSE 2023
- (Garcia-Sobrino et al. [2020]), Competitive Segmentation Performance on Near-Lossless and Lossy Compressed Remote Sensing Images, IEEE GRSL
- (Ballé et al. [2018]), Variational Image Compression with a Scale Hyperprior, arXiv
- (Cheng et al. [2020]), Learned Image Compression with Discretized Gaussian Mixture Likelihoods and Attention Modules, arXiv
- (Xue et al. [2017]), Video Enhancement with Task-Oriented Flow, arXiv (Vimeo-90k)
- (Kodak [1993]), Kodak Lossless True Color Image Suite (Photo CD Sampler, Final Version 2.0)

OroraTech

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Made with ❤️ by OT Data Science, November 2025