Impact of Learned Domain Specific Compression on Satellite Image Object Classification

Alexander Bayerl[1], Manuel Keglevic[1], Matthias Wödlinger[1] and Robert Sablatnig[1]

[1]Computer Vision Lab, TU Wien, Favoritenstraße 9/193-1, Vienna, Austria

Abstract

This paper proposes a methodology for learned compression for satellite imagery. The proposed method utilizes an image patching and stitching approach to address the high resolution of satellite images. We present rate-distortion metrics showing that this methodology outperforms JPEG2000, currently used on satellites. In addition, we demonstrate that using satellite images to train the compression model leads to superior performance compared to using non-domain-specific data. Furthermore, a detailed evaluation of the compression algorithm in a downstream classification task is conducted. The results demonstrate that 77.83% classification accuracy is still achievable for highly compressed images with a bitrate of 0.02 BPPs when the classification model is trained on images from the same compression model. The downstream classification task evaluation highlights that the performance of the classification model is highly dependent on the type of compression applied to the training data. When trained with learned compression images, the model can only classify images with an acceptable level of accuracy (>77%) if they had also undergone learned compression. Likewise, a model trained with JPEG images can only classify JPEG images with acceptable accuracy (>89%).

Keywords

Learned Image Compression, Satellite Imagery, Remote Sensing, Image Classification, Machine Learning

1. Introduction

As remote sensing technology develops, satellites take photos with increasing spatial, temporal, and spectral resolution. This leads to an increasing amount of produced data per day, which is a challenge for data storage [1]. In addition to data storage, transferring satellite images from satellites to terrestrial nodes is a bottleneck in this process as well. Compression algorithms specialized for the satellite image domain have been developed to alleviate this problem [2, 3, 4, 5].

Since image compression is a ubiquitous and fundamental operation, it is a well-studied topic. Improvements in image compression enable faster image data transfer and reduced storage costs. The invention of the discrete cosine transformation in 1972 by Nasir Ahmed et al. [6] led to the definition of the JPEG-Format in 1992, which is still dominant. Ballé et al. [7] showed in 2016 that using compression models trained by artificial neural networks can outperform traditional image compression algorithms like JPEG-Discrete Cosine Transformation in terms of image quality and bitrate.

For a specific image domain, further enhancements in learned compression can be achieved by limiting the training data to images from this domain. For example, Tsai et al. [8] show that using domain-specific training data can significantly enhance the compression performance of video game images. Similarly, Wödlinger et al. [9] demonstrate superior performance in stereo image compression compared to other approaches by designing a custom-built architecture and training it using domain-specific data.

For satellite images following difficulty must be taken into account: currently, 27 satellites with a spatial resolution of less than 10 m per pixel are active, 19 of which have been launched in the last 20 years [10]. This results in increasing file sizes per satellite image [11] which has to be considered when processing such images on neural network hardware accelerators. Even though a simple method for handling this is dividing the image into processable patches and compressing each patch independently, this leads to stitching artifacts on the border between two patches in the decompressed image.

This work examines learned image compression in the context of satellite photography:

- We propose a methodology to alleviate border artifacts when stitching patches of compressed images.
- The proposed method is evaluated on a classification downstream task (see Figure 1) using the "Functional Map of the world"-satellite image data set published in 2017 by John Hopkins University Applied Physics Laboratory [12] .
- Furthermore, we investigate the influence of domain-specific training data on the rate- dis-

26th Computer Vision Winter Workshop, Robert Sablatnig and Florian Kleber (eds.), Krems, Lower Austria, Austria, Feb. 15-17, 2023

alexanderbayerl95@gmail.com (A. Bayerl);
keglevic@cvl.tuwien.ac.at (M. Keglevic);
mwoedlinger@cvl.tuwien.ac.at (M. Wödlinger);
sab@cvl.tuwien.ac.at (R. Sablatnig)
0000-0002-4644-2723 (M. Keglevic); 0000-0002-3872-7470 (M. Wödlinger); 0000-0003-4195-1593 (R. Sablatnig)

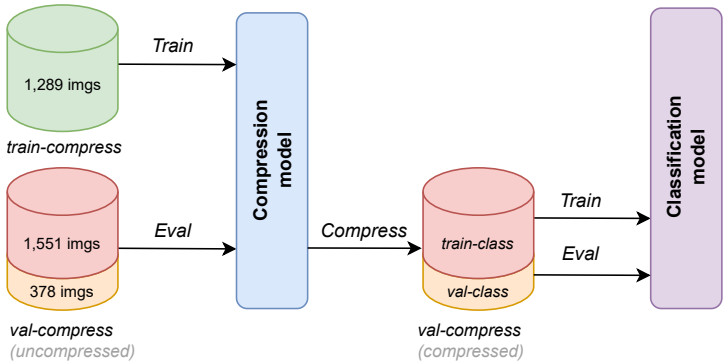


Figure 1: Separation of the dataset used in this work: Subset of fMoW have been used to train the compression. Subset of fMoW have been compressed. The compressed images themselves are separated into a train and into validation set for classification training

tortion metric and the classification downstream task.
- We show that even with compression ratios as low as 0.02 BPP, a classification accuracy of 77.83% can be achieved as long as domain-specific data is utilized for training.

2. State of the art

Lossy image compression is the process of reducing the size of digital image data without sacrificing its overall quality. This differs from lossless image compression, which does not permit any information loss during the compression process.

2.1. Traditional Image Compression

A.J. Hussain et al. [13] conducted an exhaustive survey on the subject of lossy image compression. The authors separate the compression approaches into predictive coding, transform coding, vector quantization, and neural network approaches.

JPEG, the most popular lossy image codec, is based on transform coding, which uses the Discrete Cosine Transformation to convert an image from pixel-space to frequency-space [6]. The method utilizes the fact that the human visual system is less susceptible to variations in high-frequency components. By applying wavelet transformations on the image, JPEG2000 improves on that to achieve better rate-distortion metrics [14].

More recently, Fabrice Bellard developed the BPG format (Better Portable Graphics) that outperforms JPEG and JPEG2000 in terms of rate and distortion [15]. This format relies on the intraframe encoding of HEVC [15].

2.2. Learned Image Compression

Recently, image compression models based on artificial neural networks have outperformed traditional compression methods in terms of rate and distortion. Jamil et al. [16] provide a survey on that subject. According to the findings of this survey, autoencoders are the most common learning-driven lossy image compression architectures. These models utilize an encoder to transform image data into a low-dimensional latent space. A decoder is then employed to reconstruct the original image from this encoding. The seminal work of this approach is from Balle et al. [7]. They learn a probability distribution of the latent space jointly with the encoder and decoder networks trained to reconstruct the original image. Subsequent works employ hyperpriors and auto-regressive context models to decorrelate the spatial information in the latent space [17].

Similarly, Toderici et al. [18] show that Recurrent Neural Network (RNN) architectures can be used for learned image compression. Their model leverages feedback loops to iteratively compress an image to the desired bit rate.

Furthermore, Generative Adversarial Networks (GANs) have also been used in image compression. According to Jamil et al. [16], GAN compression outperforms traditional image compression algorithms in terms of visual quality, albeit with the disadvantage of higher deployment costs.

2.3. Satellite image compression

Indradjad et al. [19] compare four different approaches for satellite image compression with transform codings: a wavelet approach by Delaunay et al. [20], bandelets [21], JPEG 2000 [14], and a discrete wavelet transformation method by the CCSDS (Consultative Committee for Space

Data Systems) [2]. Of these approaches, JPEG 2000 yields the highest peak signal-to-noise ratio (PSNR) as well as the second shortest compression and decompression times.

More recently, de Oliviera et al. [4] investigated neural networks for the compression of satellite images. An autoencoder with learned hyper-prior is utilized to learn compression models for satellite imagery. The proposed method outperforms the CCSDS wavelet compression [2] currently used on French satellites in terms of rate and distortion.

Bacchus et al. [3] investigate the use of learned methods for onboard satellite image compression, to address high memory and complexity constraints in this domain. The authors also employ a hyperprior-based architecture and incorporate data augmentations as a preprocessing step. Their method performs better than JPEG2000, and the authors concluded that its relatively low inference time makes it well-suited for use on satellites.

3. Methodology

This section provides an overview of the methodology proposed in this work. It begins with a brief introduction to learned image compression, followed by an explanation of how the technique is adapted to suit high-resolution satellite images.

3.1. Learned Image Compression

This work is based on the compression model by Balle et al. [17]. Figure 2 shows an overview of the architecture. The model has an autoencoder structure, and the distribution of the quantized latent $p_{\hat{\boldsymbol{y}}}$ is modeled using a learned hyperprior g_h and a context model g_{cm} that predicts the parameters of a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma})$. The autoregressive component utilizes already decoded pixels for decoding further pixels. This yields superior rate-distortion results, with the disadvantage that decoding has to be done iteratively and not in parallel.

We directly train the model with the trade-off between the distortion D of the original image and the compression rate R:

$$L = D + \lambda \cdot R \tag{1}$$

Here λ controls the trade-off between rate and distortion. For the distortion D the Mean Squared Error (MSE) is used, which computes the averaged pixel-wise quadratic difference between original image and distorted image:

$$D = \mathbb{E}_{\boldsymbol{x} \sim p_{\boldsymbol{x}}} ||\boldsymbol{x} - \hat{\boldsymbol{x}}||_2^2 \tag{2}$$

The compression rate R is estimated by the cross-entropy between the entropy model distribution $p_{\hat{\boldsymbol{y}}}$ and the actual marginal distribution $m(\boldsymbol{y})$, where $\boldsymbol{y}$ denotes the latent encoding. Similarly, the rate of the hyperprior $\boldsymbol{z}$ is calculated which leads to the following definition for the rate-loss R:

$$R = \underbrace{\mathbb{E}_{\boldsymbol{x} \sim p_{\boldsymbol{x}}} \left[-\log_2 p_{\hat{\boldsymbol{y}}}(\hat{\boldsymbol{y}}) \right]}_{\text{rate (latents)}} + \underbrace{\mathbb{E}_{\boldsymbol{x} \sim p_{\boldsymbol{x}}} \left[-\log_2 p_{\hat{\boldsymbol{z}}}(\hat{\boldsymbol{z}}) \right]}_{\text{rate (hyper-latents)}} \tag{3}$$

3.2. Stitching

As discussed in the introduction, a limitation of satellite imagery is that image samples have resolutions of up to 14798×14802 pixels, which causes issues for the training and inference on neural network hardware accelerators such as GPUs. Since dividing the input into patches and processing the patches independently of each other leads to visible artifacts on the borders between the patches in the stitched images, our approach resolves this issue by compressing overlapping patches. For the stitched image, the average value of both patches (or four patches in corners) is used for the overlapping regions. Figure 3 illustrates the overlapping regions of a 1496×1496 image with a patch size of 256×256 pixels. A step size of 248 pixels in either the X or Y dimension is employed, resulting in an overlapping region of 8 pixels. A disadvantage of this method is that it leads to the pixels in the overlapping regions being compressed multiple times, i.e., 5.14% of the total pixels in the previous example.

In Figure 4 the influence of this blending process can be seen. As a result, the boundaries of each patch are less visible in the blended image on the right.

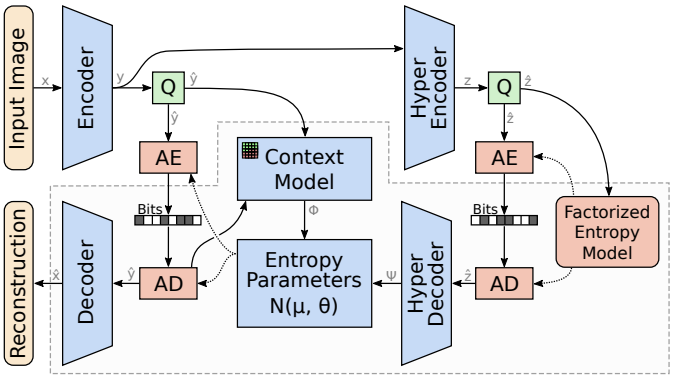
4. Evaluation

This section provides an overview of the evaluation process and presents the results of this work.

Firstly, the utilized data set is described in detail, and how it was employed in this work. Subsequently, the results of the proposed compression algorithm on the data set are highlighted and discussed. Finally, the results of the downstream classification task on the compressed images are presented.

4.1. Dataset

The dataset used in this work is the Functional Map of the world (fMoW). It was created at the John Hopkins University Applied Physics Laboratory in Laurel, Maryland (United States) and is publicly available at https://github.com/fMoW/dataset [12]. This dataset was compiled to facilitate research in computer vision for



Component	Symbol
Input Image	$\boldsymbol{x}$
Encoder	$f(\boldsymbol{x};\boldsymbol{\theta}_e)$
Latents	$\boldsymbol{y}$
Latents (quantized)	$\hat{\boldsymbol{y}}$
Decoder	$g(\hat{\boldsymbol{y}};\boldsymbol{\theta}_d)$
Hyper Encoder	$f_h(\boldsymbol{y};\boldsymbol{\theta}_{he})$
Hyper-latents	$\boldsymbol{z}$
Hyper-latents (quant.)	$\hat{\boldsymbol{z}}$
Hyper Decoder	$g_h(\hat{\boldsymbol{z}};\boldsymbol{\theta}_{hd})$
Context Model	$g_{cm}(\boldsymbol{y}_{<i};\boldsymbol{\theta}_{cm})$
Entropy Parameters	$g_{ep}(\cdot;\boldsymbol{\theta}_{ep})$
Reconstruction	$\hat{\boldsymbol{x}}$

Figure 2: Compression architecture used in this work [17].

Figure 3: 1496×1496 image divided into 36 patches (256×256 pixels) with an overlapping region of 8×256 pixels between two patches.

remote sensing applications. It includes over 1 million images of objects taken from satellites, categorized into 63 categories, such as airports, tunnel openings, zoos, and towers. Christie et al. [12] highlight the importance of obtaining a geographically distributed data set to minimize geographical bias.

Overall the dataset contains about 628,000 training images and about 100,730 images for validation. The photographs are provided as compressed JPEG- and lossless TIFF-color images. Each object has been photographed in a variety of environmental settings (weather, time, season). Since this work explicitly focuses on high-resolution satellite images, only images with a resolution of at least 1024×1024 pixels are considered.

The partitioning of the data set used in this work is shown in Figure 1. For compression training, 1,289 images from the fMoW train set, uniformly distributed over all 63 categories, are used (*train-compress*). These 1,289 images are from 1,038 objects. As such, for some objects, there are multiple images taken under different environmental conditions.

Another set, denoted as the (*val-compress*), consists of 1,929 images from 1,038 objects from the fMoW validation set. The *val-compress* serves two purposes: one is evaluating the compression, and another is evaluating the downstream classification task. For the latter, the *val-compress* is split again into 1,551 images for classification training (*train-class*) and 378 images for classification validation (*val-class*).

4.2. Compression Evaluation

With the parameter λ in Equation 1 the trade-off between rate and distortion can be controlled, i.e. increasing the λ leads to a smaller MSE but therefore more BPP. To evaluate our model for different bitrates, we train the model with different values for the parameter lambda. In Figure 5 the compression results with bitrates ranging from 0.003 BPP to 0.68 BPP are shown for an example image. The BPP of the compressed image is calculated directly by dividing the file size of the encoded image by the amount of pixels in the respective image.

The Peak-signal-to-noise-ratio (PSNR) metric is used to evaluate the distortion. The distortion is calculated using the MSE between the compressed and the corresponding uncompressed images. The PSNR is defined as:

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{255}{\text{MSE}}\right) \tag{4}$$

As depicted on the Rate-Distortion-Curve in Figure 6, the results indicate that the proposed learned compression methodology outperforms JPEG and is also superior

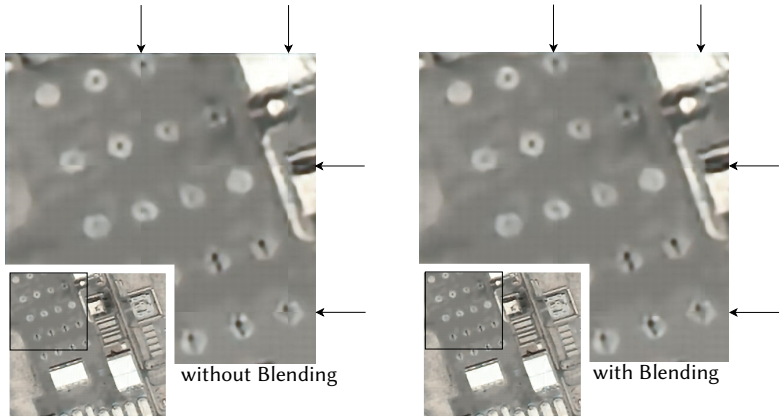


Figure 4: In the left image, simple patching without blending is shown. In this case the connection line between two patches can be seen. In the right image the patches are blended. The connection lines are denoted with arrows in the images.

Table 1
Results of domain specific training and mon domain specific training

	PSNR	*BPP*
ImageNet trained	32.8	0.84
FMoW trained	32.92	0.67

to the JPEG2000 compression format, which is frequently used in satellite applications.

To verify that training the compression model with domain-specific satellite images improves the downstream classification task, another compression model was trained on 1,749 non-domain-specific samples from the ImageNet data-set.

The results in Table 1 show that the domain-specific compression model trained with satellite images outperforms the model trained with ImageNet samples. For a PSNR of 33 it yields a lower bit rate of 0.67 BPP compared to 0.84 BPP achieved by the domain-agnostic model.

5. Classification Evaluation

In addition to the evaluation in terms of image quality, in this section, compressed quality is represented by the accuracy of a classification downstream task, i.e., identifying objects in satellite images. As mentioned in the Section 4.1, compressed and uncompressed versions of the *val-compress* set are used, with 1,554 images used for training (*train-class*) the classification model, and 375 images used to validate the model (*val-class*).

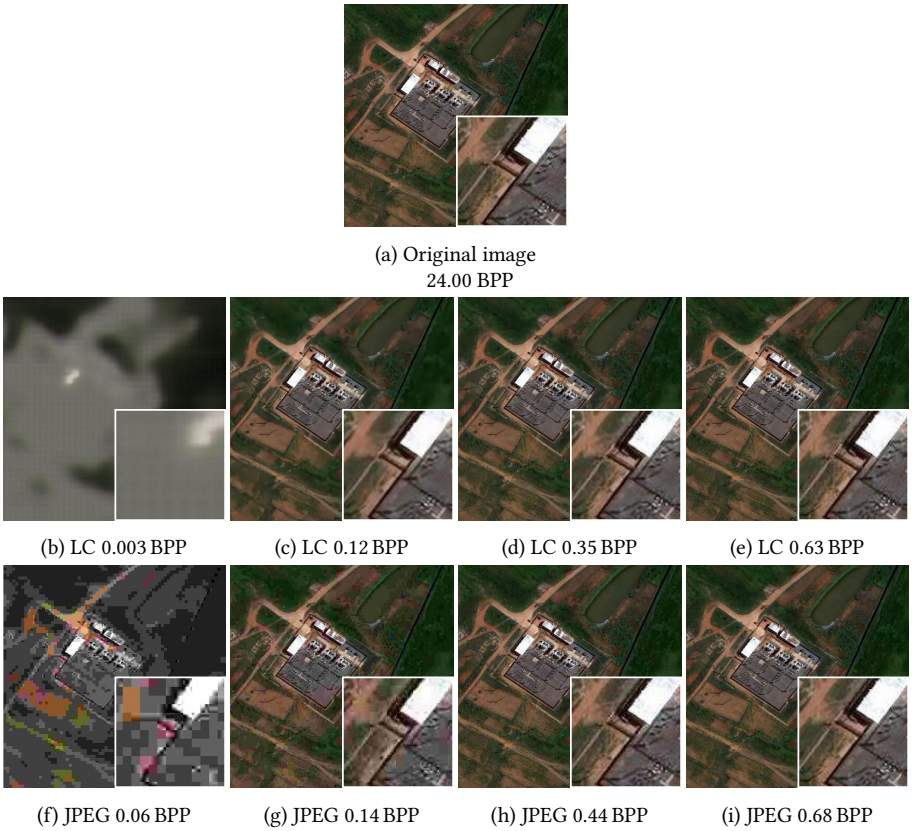
A dual path network[1] [22] is utilized for this evaluation. For the evaluation of the classification downstream task, classification models have been trained with the following 4 training sets:

- *train-class* compressed by the learned compression model with 0.02 BPP
- *train-class* compressed by the learned compression model with 0.67 BPP
- *train-class* compressed by the learned model that was trained with 1,749 non-domain specific images (ImageNet [23]) with 0.84 BPP
- *train-class* compressed in JPEG format with 0.77 BPP

Each of these classification models has been used to validate data sets in different compression scenarios: JPEG data sets (0.31 BPP, 0.76 BPP, 1.55 BPP), Learned-compression-compressed (LC) datasets (0.02 BPP, 0.67 BPP, 1.07 BPP), dataset retrieved from ImageNet-trained learned Compression (0.84 BPP) and one without compression. The result for this classification validations are shown in Table 2. The columns denote the data set the classifier was trained on, the rows denote the data set that was classified during validation.

The results show that a classification model works best when classifying images that were compressed with the same algorithm (JPEG or learned compression) as the images on which it was trained, i.e., the JPEG-trained classifiers classified JPEG images with accuracies over 89%. In contrast, the JPEG-trained classifier only achieves an accuracy of up to 35.21% on images compressed by learned compression. Similarly, the accuracy of the LC-trained classifiers was at least 77% when classifying LC images (except for very low bitrate of 0.02 BPP), and no more than 39.38% when classifying JPEG-compressed images.

[1] https://github.com/fMoW/first_place_solution



(a) Original image
24.00 BPP

(b) LC 0.003 BPP (c) LC 0.12 BPP (d) LC 0.35 BPP (e) LC 0.63 BPP

(f) JPEG 0.06 BPP (g) JPEG 0.14 BPP (h) JPEG 0.44 BPP (i) JPEG 0.68 BPP

Figure 5: Differences between an original image, JPEG compressed versions of this images, and results of our learned compression models (LC), with given BPP.

Table 2
Classification accuracy for datasets with varying compression distortion; the columns denote classifiers trained with various image compression formats; the columns denote the data set, that was compressed; LC= Learned Compression

Validated with:	Classifier trained on images compressed with:			
	LC 0.02 BPP	*LC 0.67 BPP*	*LC ImageNet 0.84 BPP*	*JPEG 0.77 BPP*
LC 0.02 BPP	**77.83%**	27.12%	10.58%	12.91%
LC 0.67 BPP	25.57%	**80.67%**	15.67%	32.89%
LC 1.07 BPP	26.9%	78.81%	16.63%	35.21%
LC ImageNet 0.84 BPP	14.94%	14.67%	**78.92%**	11.22%
JPEG 0.31 BPP	14.41%	35.81%	11.22%	89.81%
JPEG 0.77 BPP	16.59%	38.07%	11.85%	**91.55%**
JPEG 1.55 BPP	15.34%	39.38%	14.32%	91.55%
Uncompressed	15%	39.52%	12.97%	91.29%

The ImageNet-trained compression model demonstrates that the training data is also crucial for the downstream classification task. It classified 78.92% of the images created by the same compression model, while other datasets, even with high bitrate, could not attain an accuracy greater than 16.63% for any other evaluation. It fails to classify all other data sets, including the LC-compressed ones trained with satellite images. This suggests that traditional compression methods lead to a more versatile encoding that is not as dependent on the specific domain.

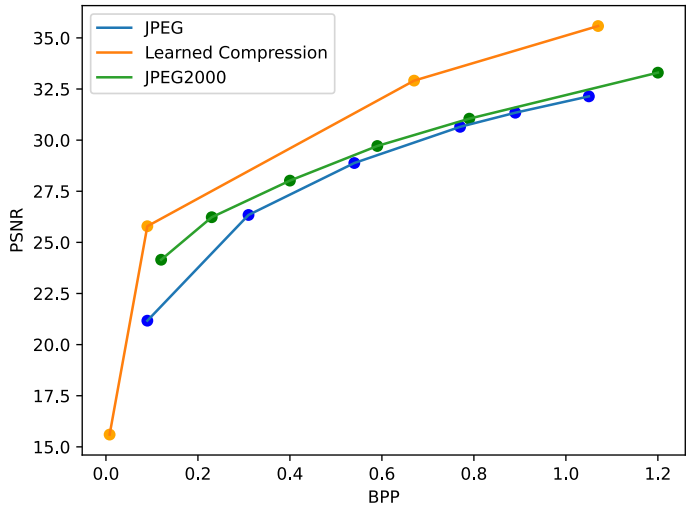


Figure 6: Rate distortion graph for JPEG, JPEG2000 and proposed learned compression methodology evaluated on satellite images (*val-compress* dataset).

6. Conclusion

In this work, we propose a satellite compression methodology that outperforms traditional methods (JPEG, JPEG2000) in terms of rate and PSNR. We show that images that exceed the memory of typical neural network hardware accelerators can be compressed by feeding in patch-wise parts of the image. To remove artifacts at the connection line between two patches, the connection region is smoothed by compressing overlapping patches and combining the pixels in these regions. The proposed methodology offers superior performance compared to JPEG, and JPEG2000, commonly used for satellite imaging. We assess the effects of compression on the performance of an object classification downstream task. We demonstrate that a classification model can learn to classify images with an accuracy of 77.83% even for images compressed with a bitrate as low as 0.02 BPP. Furthermore, we show that using differently encoded images for training and inference can deteriorate classification accuracy significantly. As such, classification models trained with JPEG images only achieve acceptable results when tested on JPEG images. Similarly, classification models trained with images compressed with learned compression models fail when tested with JPEG images.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 965502.

References

[1] H. Guo, Z. Liu, H. Jiang, C. Wang, J. Liu, D. Liang, Big earth data: A new challenge and opportunity for digital earth's development, International Journal of Digital Earth 10 (2017) 1–12.

[2] P.-S. Yeh, P. Armbruster, A. Kiely, B. Masschelein, G. Moury, C. Schaefer, C. Thiebaut, The new ccsds image compression recommendation, 2005, pp. 4138 – 4145. doi:10.1109/AERO.2005.1559719.

[3] P. Bacchus, R. Fraisse, A. Roumy, C. Guillemot, Quasi lossless satellite image compression, IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium (2022) 1532–1535.

[4] V. A. de Oliveira, M. Chabert, T. Oberlin, C. Poulliat, M. Bruno, C. Latry, M. Carlavan, S. Henrot, F. Falzon, R. Camarero, Satellite image compression and denoising with neural networks, IEEE Geoscience and Remote Sensing Letters 19 (2022) 1–5.

[5] F. E. Hassan, G. I. Salama, M. S. Ibrahim, R. M. Bahy, Investigation of on-board compression techniques for remote sensing satellite imagery, in: International Conference on Aerospace Sciences and Aviation Technology, volume 11, The Military Technical College, 2011, pp. 937–946.

[6] N. Ahmed, T. Natarajan, K. Rao, Discrete cosine

transform, IEEE Transactions on Computers C-23 (1974) 90–93. doi:10.1109/T-C.1974.223784.

[7] J. Ballé, V. Laparra, E. P. Simoncelli, End-to-end optimized image compression, in: 5th International Conference on Learning Representations, ICLR 2017, 2017.

[8] Y.-H. Tsai, M.-Y. Liu, D. Sun, M.-H. Yang, J. Kautz, Learning binary residual representations for domain-specific video streaming, Proceedings of the AAAI Conference on Artificial Intelligence 32 (2018). URL: https://ojs.aaai.org/index.php/AAAI/article/view/12259. doi:10.1609/aaai.v32i1.12259.

[9] M. Wödlinger, J. Kotera, J. Xu, R. Sablatnig, Sasic: Stereo image compression with latent shifts and stereo attention, in: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 651–660. doi:10.1109/CVPR52688.2022.00074.

[10] Y. Lai, J. Zhang, Y. Song, Surface water information extraction based on high-resolution image, IOP Conference Series: Earth and Environmental Science 330 (2019) 032013. URL: https://dx.doi.org/10.1088/1755-1315/330/3/032013. doi:10.1088/1755-1315/330/3/032013.

[11] Q. Zhao, L. Yu, Z. Du, D. Peng, P. Hao, Y. Zhang, P. Gong, An overview of the applications of earth observation satellite data: Impacts and future trends, Remote Sensing 14 (2022) 1863. doi:10.3390/rs14081863.

[12] G. Christie, N. Fendley, J. Wilson, R. Mukherjee, Functional map of the world, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6172–6180.

[13] A. Hussain, A. Al-Fayadh, N. Radi, Image compression techniques: A survey in lossless and lossy algorithms, Neurocomputing 300 (2018) 44–69. doi:https://doi.org/10.1016/j.neucom.2018.02.094.

[14] A. Skodras, C. Christopoulos, T. Ebrahimi, The jpeg 2000 still image compression standard, IEEE Signal Processing Magazine 18 (2001) 36–58. doi:10.1109/79.952804.

[15] G. J. Sullivan, J.-R. Ohm, W.-J. Han, T. Wiegand, Overview of the high efficiency video coding (hevc) standard, IEEE Transactions on Circuits and Systems for Video Technology 22 (2012) 1649–1668. doi:10.1109/TCSVT.2012.2221191.

[16] S. Jamil, M. J. Piran, MuhibUrRahman, Learning-driven lossy image compression; a comprehensive survey, 2022. URL: https://arxiv.org/abs/2201.09240. doi:10.48550/ARXIV.2201.09240.

[17] D. Minnen, J. Ballé, G. D. Toderici, Joint autoregressive and hierarchical priors for learned image compression, Advances in neural information processing systems 31 (2018).

[18] G. Toderici, D. Vincent, N. Johnston, S. Jin Hwang, D. Minnen, J. Shor, M. Covell, Full resolution image compression with recurrent neural networks, in: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2017, pp. 5306–5314.

[19] A. Indradjad, A. S. Nasution, H. Gunawan, A. Widipaminto, A comparison of satellite image compression methods in the wavelet domain, IOP Conference Series: Earth and Environmental Science 280 (2019) 012031. doi:10.1088/1755-1315/280/1/012031.

[20] X. Delaunay, M. Chabert, V. Charvillat, G. Morin, Satellite image compression by post-transforms in the wavelet domain, Signal Processing 90 (2010) 599–610. doi:10.1016/j.sigpro.2009.07.024.

[21] S. Mallat, G. Peyré, A Review of Bandlet Methods for Geometrical Image Representation, Numerical Algorithms 44 (2007) 205–234. URL: https://hal.archives-ouvertes.fr/hal-00359744.

[22] J. L. Yuppen Chen, International Journal of Computer Applications (2017).

[23] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: A large-scale hierarchical image database, in: 2009 IEEE conference on computer vision and pattern recognition, Ieee, 2009, pp. 248–255.