

# Deep internal learning for single SWIR satellite image super resolution

Yakov Geltser, Shimrit Maman, Stanley Rotman, Dan G. Blumberg.

Ben-Gurion University of the Negev, Israel.

## ABSTRACT

The compact dimensions of CubeSats limit the optical equipment they can carry, which in turn affects the spatial resolution of the images they capture. BGUSAT, a 3U CubeSat, gathers Short Wave Infra-Red (SWIR) images between 1.55-1.7 micrometers with a spatial resolution of 600 meters per pixel. Leveraging deep learning techniques for enhancing satellite imagery resolution, particularly for CubeSats like BGUSAT, offers significant improvements of data analysis for remote sensing applications. Traditional deep learning super-resolution algorithms require a large amount of training data. However, if there is a shortage of available satellite imagery or if a single existing image requires enhancement through super-resolution techniques, it may not be sufficient to rely on traditional methods.

Single image super-resolution methods, such as bicubic interpolation, do not consider the complexity of features within the image and provide very limited enhancement results. Satellite imagery characteristics vary significantly across sensors, altitudes, and spectral bands. Pre-trained models in supervised learning may yield inaccurate predictions with new sensor data. Thus, a self-supervised method, Zero Shot Super Resolution (ZSSR), which focuses on the unique internal features of each image to extract latent information, was adopted. Our proposed approach using the ZSSR algorithm operates without reference data, ensuring high-quality, image-specific data enhancement using a single image.

A single BGUSAT image was super-resolved using our approach, with scale factors ranging from 2 to 9. several evaluation methods were applied to compare the quality of super-resolved images using ZSSR against traditional bicubic interpolation: visual interpretation, and three non-reference evaluation methods.

**Keywords:** Super-resolution, Remote sensing, Single band, CubeSat, Internal learning, Nano satellite.

## 1. INTRODUCTION

Super-resolution methods are widely used in a variety of applications of image processing and enhancement, including the field of earth observation [1]. Satellite imagery data is available at various scales and formats. In 2017, BGUSAT, a nano-satellite, was launched with a single-band short-wave infrared (SWIR) sensor operating between 1.55-1.7 micrometers of wavelength. Its imagery is used for various purposes such as flood detection and climate change monitoring. However, BGUSAT's main disadvantage is its spatial resolution, 600m/pixel. The imagery data generated by BGUSAT can be enhanced and used for much more detailed research by using novel deep learning methods such as super resolution. Performing super resolution image enhancement by deep learning methods is a challenge when training data is scarce and limited.

The idea of increasing image resolution is based on the addition of pixels and features to an original image to expose details that cannot be seen on the original image like specific features, lines, and shapes. The most common method to conduct super resolution is by adding nearest neighbor pixels or using interpolations such as bicubic interpolation. [2]. These methods do increase the resolution of the image but do not allow for the detection of finer high frequency details. The goal of this research is to provide a sharp image with high-frequency details. Utilizing the power of deep learning, it is possible to obtain high frequency details through a learning process. Convolutional neural networks are widely used in the field of image processing for various cases, image super resolution (SR) is one of them. The prime methods which are used for SR require an extensive training on a large dataset, methods such as SRGAN [3] use training dataset to train a generator and discriminator which provides a SR image based on the generated data. SREDSR [4] is an encoder decoder model based on SRResNet architecture, the model requires an extensive training on a large dataset. MIP [5] Is a generator-discriminator based method and is used to generate random noise and uses its previously trained discriminator, Self-Fuse Net [6] uses a

fusion network along with a customized UNet and divides the features inside the image to low, mid and high frequency features, and uses them to fuse the high resolution (HR) image. In the field of satellite imagery SR, HighRes-net [7] is used with pre-trained data, and RAMS [8] uses a multi-image method which allows for the collection of more samples of the same place with a slight spatial and temporal shift.

In the BGUSAT case, the amount of data are limited, in some cases there is only one image of the targeted area, where pre-trained SR models are trained on a considerably large datasets, so the case of BGUSAT imagery does not fit the model of been used as training data for a pre-trained model, also using pre-trained models might be prone to overfitting, pretrained data in the field of remote sensing depends on the kind of data the model was trained on, whether its resolutions fit the requirements, of spatial and spectral resolutions thus an image specific method is required to fit this kind of data. A method for providing SR in multiple scales, that can work on a single band, on a single image, where sensor type, satellite orientation, atmospheric conditions, ground segments and shadowing caused by different times of image acquisition is required. Relying on the internal scene statistics [9], a great amount of missing information is required for SR and can be found **within** the image when considering recurring features. For example, one can expect that features that are specific to a corn field can reoccur in the near area around a corn field, this reoccurrence provides the possibility to estimate missing pixels and features. Zero Shot Super-Resolution (ZSSR) [10] makes use of the internal features method to learn which features are more likely to reoccur in a specific image and use them to predict missing pixels when applying SR.

## 2. METHODOLOGY

In this study, the ZSSR algorithm was used for super-resolution of images at scales from 2 up to 9, based on learning the internal statistics of an image as mentioned in [11]. The ZSSR algorithm is built as a fully convolutional network with 8 hidden layers, each with 64 channels, and ReLU activation on each layer. The network input is interpolated to the output size. In this system, only the residual between the interpolated low-resolution (LR) and its HR parent is learned. An  $L_1$  loss function is used with the ADAM optimizer. The learning rate acts as a stop function, starting at 0.001 and periodically being divided by 10 based on the linear fit of the reconstruction error. If the standard deviation exceeds a factor of the slope of the linear fit, the learning rate decreases. Training stops when the learning rate reaches  $10^{-6}$ . The weights are initialized randomly.

The process works as follows: an input image is downscaled by the required factor and used as input for the network training phase. The network begins to train by augmenting the image and rotating it to learn internal features in various orientations. It also divides the image into "father" and "son" pairs, where each pair is an augmentation of the original image. The learning rate drops as the mean squared error (MSE) decreases below a predetermined value. When the stopping criterion is met, determined by the learning rate, the original input image is inserted into the network prediction phase where the weights have already been set by the previous step. The image is then upscaled to the predetermined scale factor using the filters to complete the low-resolution features to higher resolution. The architecture of the algorithm can be seen in Figure 1.

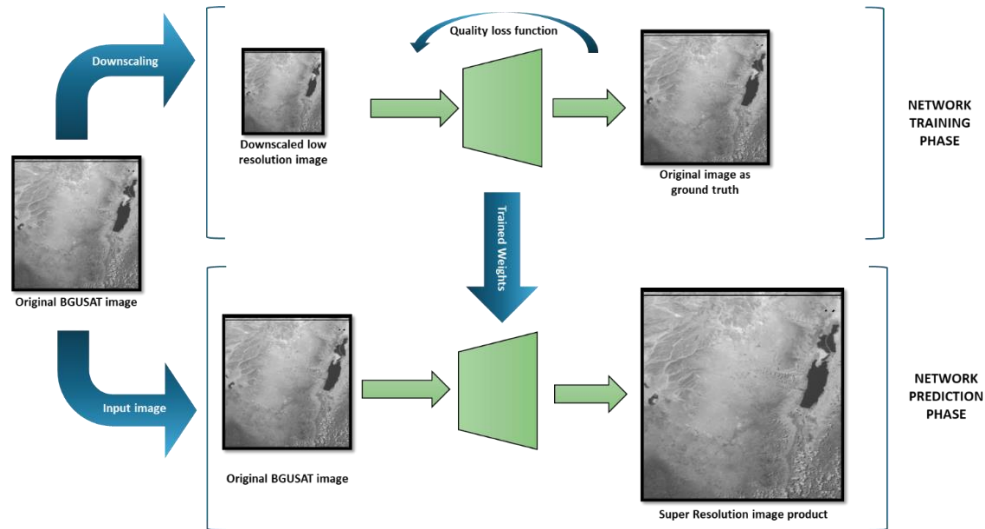


Figure 1. System architecture schematic – modified after (Irani 2017).

### 3. RESULTS

This chapter describes the results generated by applying deep learning methods to test and improve the SR methods applied to a single band, low resolution satellite image. We applied the ZSSR as described by [10] on BGUSAT images. To the best of our knowledge ZSSR has been tested on satellite images on [12], [6] where in the first paper the multi band results were evaluated by using a full-reference methods, The latter paper used single band image but didn't show and measure the appearance of hidden features after the process of super resolution. To provide SR results a sequence of growing SR images was produced at increasing scale factors, In Figure 2 the SR images are presented with their scale factor, where the area of the dead sea is enlarged to show features that are not observable in the original and lower resolution images like the evaporation pool borders. The image subset of the dead sea was further chosen for visually evaluating the super resolution results. This specific area exhibits a variety of features and can represent the extraction of hidden features. The enhancement in the Figure 3 shows an improvement in image resolution by demonstrating less pixelized image with sharper edges. Features that were not present in the original and low-SR images appear in the higher resolution ones, for example the borders of the evaporation pools in the dead sea became detectable on higher resolution enhancement. This result is especially notable when comparing the original image to the upscaled by 2 and upscaled by 4 images. In the upscaled by 8 image the enhancement is very mild.

The following results are summarized in Figure 2, the BGUSAT image is upscaled by the ZSSR algorithm to scale factors of 2,4 and 8 A, B, C, D respectively and it is evaluated by both qualitatively and quantitatively using visual image interpretation as qualitative method (Figure 3). For qualitative methods, non-reference image quality methods were selected, these methods are based on natural scene statistics [9], sharpness estimation and local variance. The quantitative results show a comparison of the upscaled image using ZSSR vs. upscaled image using standard bicubic interpolation as a basic method for SR [2].

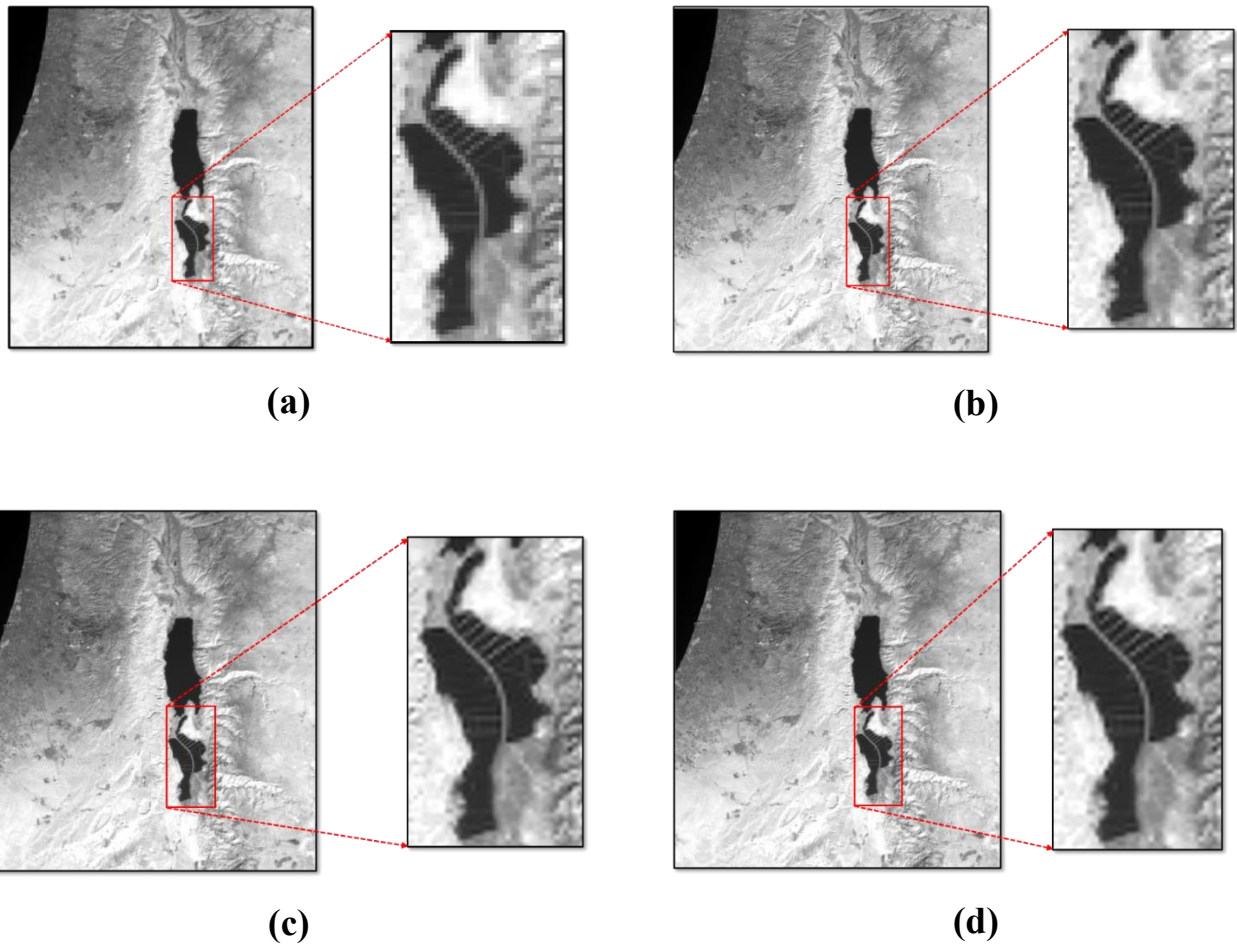


Figure 2. BGUSAT image with different scales of super-resolution using ZSSR, showing an enlarged area of the Dead Sea. (a) Original BGUSAT image, (b) ZSSR scale factor of 2, (c) ZSSR scale factor of 4, (d) ZSSR scale factor of 8.

Zooming in and examining the images in Figure 3 it is possible to see through visual interpretation that in image (b) the lines which are the evaporation pool borders in the Southern part of the Dead sea are not visible due to low resolution, as the resolution increases by using the ZSSR algorithm, it is possible to see the traces of the borders, especially when increasing the resolution by a scale factor of four. The appearance of pool borders proves the ability of the algorithm to predict hidden details by using only the internal features of the image.

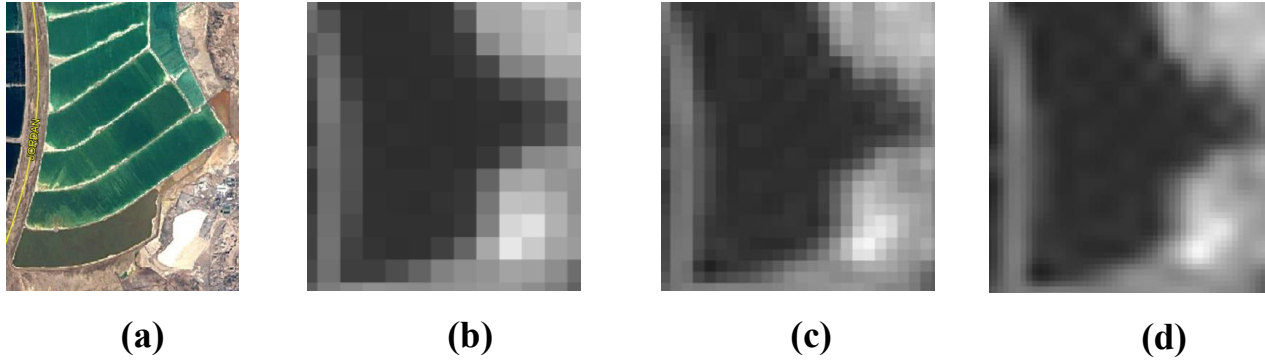


Figure 3. The above figure shows the scene at high and super resolution. (a) High-resolution image taken by Landsat/Copernicus satellites for comparison, (b) Original BGUSAT image, (c) 2x super-resolution image using the ZSSR algorithm, (d) 4x super-resolution image using the ZSSR algorithm.

By using a cross section, it is possible to get quantitative insights about the pool borders appearance. In Figure 4 a cross section was measured across the original image, x2 and x4 images. The cross section crosses the pool borders from north to south. The pixel values of the cross sections were recorded and plotted, the plots show an unrecognizable pattern in the original image, in the x2 image there is a hint of a pattern, but it is still not clear. In the x4 image cross section plot a clear pattern can be seen with the exact amount of pool borders as intensity peaks. The width of each of these pool borders is about 30 meters with wide areas which span up to 150 meters. It is remarkable that a satellite operating on the spatial resolution of  $600 \left[ \frac{m}{pixel} \right]$  can provide the insights that these pool borders exist, and it is possible to measure the distance between them by utilizing the ZSSR method.

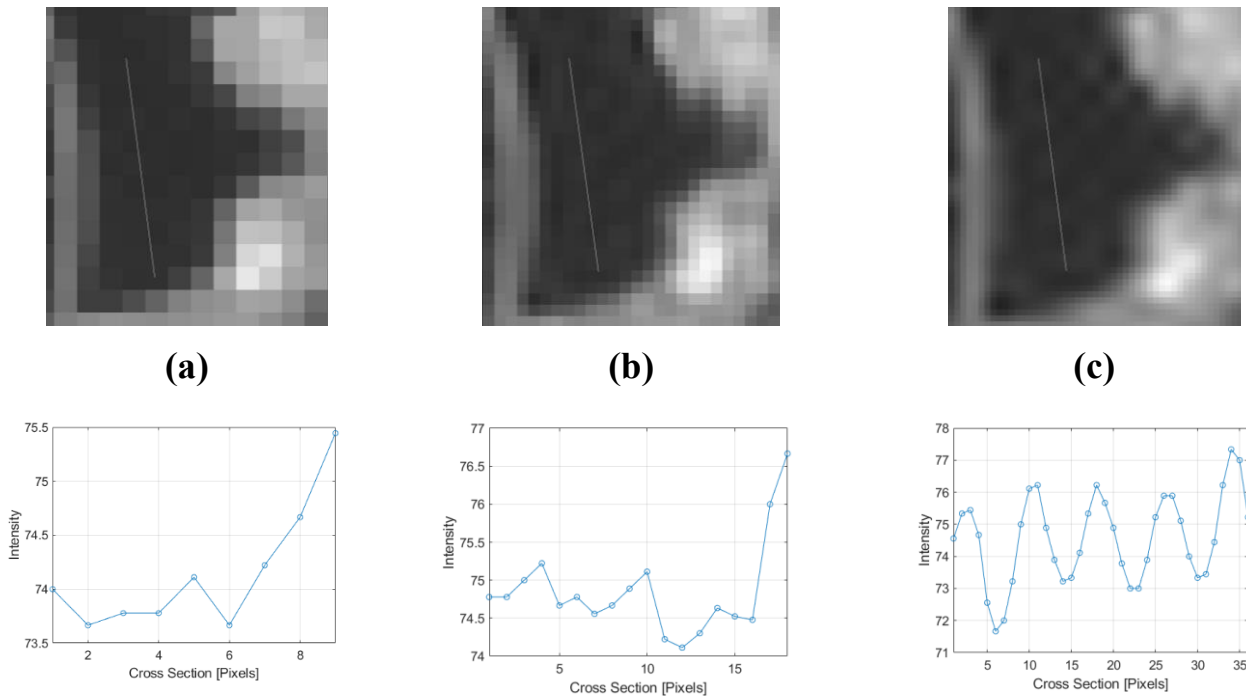


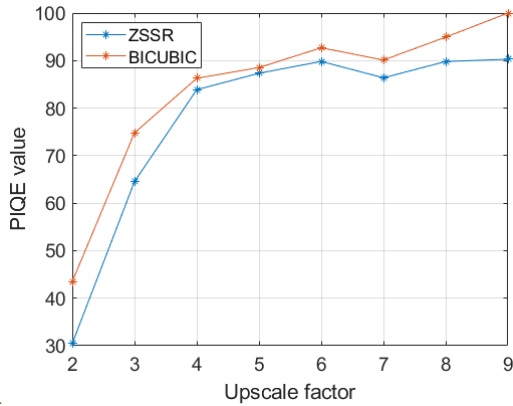
Figure 4. Dead sea evaporation pools cross-section. (a) Original image, (b) Upscaled image using ZSSR with a scale factor of 2, (c) Upscaled image using ZSSR with a scale factor of 4. The plots represent the intensity of pixels per cross-section in pixels.

For evaluating the results in a quantitative way, blind image and non-reference image quality assessment methods were chosen, used along with the visual interpretation, the quantitative non-reference methods are BRISQUE [13], NIQE [14], PIQE [15], and  $\Delta DoM$  [16]. BRISQUE and NIQE methods are based on natural scene statistics (NSS) [9] these two metrics are proven by their author to perform better statistically than the full-reference image quality (IQ) methods such as structural similarity index measure (SSIM) and the peak signal-to-noise ratio metric (PSNR). These NSS are based on different kinds of image and not necessarily on remote sensing image, which means they are not specific to remote sensing imagery. These methods show that as the results are numerically higher, it's further from the NSS, which means – higher numerical results, lower image quality. PIQE is based on block-wise distortion and local variance and is also evaluated as higher numerical results – lower image quality.  $\Delta DoM$  was originally used for detecting the sharpness of documents but is also a considerable method to determine the blurriness of the image, and is evaluated as the lower numerical score, the higher the image quality. The results of the ZSSR algorithm were compared to bicubic interpolation which is the widespread method for increasing image resolution without the need for external data or algorithm fitting. Standard bicubic interpolation focuses solely on the pixels that are near the targeted pixel, while ZSSR considers the near area along with the features from the whole image. The results of the evaluation methods per image in comparison to the standard bicubic interpolation presented in Table 1.

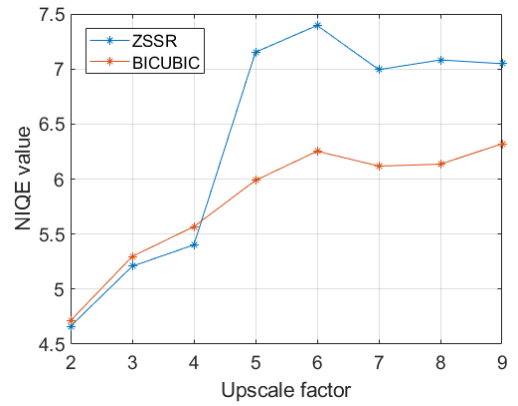
Table 1 . Quantitative results of ZSSR compared to bicubic interpolation for a single image. Lower values of PIQE, NIQE, and BRISQUE indicate better image quality, while higher values of  $\Delta DoM$  indicate better image quality.

	$\Delta DoM$	PIQE	NIQE	BRISQUE
Original	1.01	20.2	3.74	21.63
X2 Cubic Interpolated	0.98	43.45	4.71	34.88
X2 ZSSR	<b>1.05</b>	<b>32.77</b>	<b>4.58</b>	<b>27.74</b>
Difference in % (between enhancements)	7%	32%	3%	20%
X4 Cubic Interpolated	0.66	86.27	5.57	56.32
X4 ZSSR	<b>0.79</b>	<b>84.52</b>	<b>5.34</b>	<b>52.46</b>
Difference in % (between enhancements)	19%	2%	4%	7%

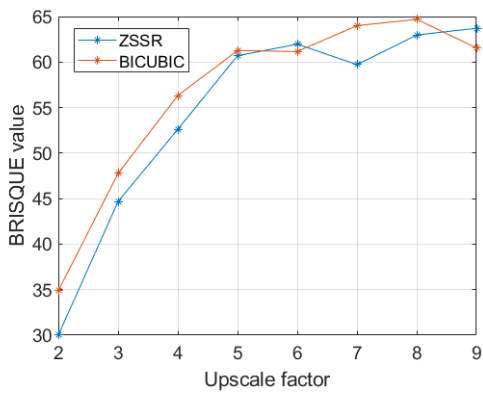
As can be seen in Table 1, by all metrics the ZSSR algorithm provides better results in comparison to the standard bicubic interpolation - lower values in BRISQUE NIQE and PIQE mean higher image quality, and higher values in  $\Delta DoM$  mean better image sharpness, although in some cases the improvement is mild, it still shows. It is important to note that these matrices mostly measure image quality disregarding the fact that super resolution might provide false information, this information is further validated using the visual interpretation. These metrics mostly measure noise and blurriness of the image and demonstrate that by applying super resolution either by interpolation or by deep learning algorithms, noise and blurriness are added to the image. Resulting in biased statistics from the natural scene. while using x4 ZSSR, the  $\Delta DoM$  sharpness matric shows that the image is sharper by 19% than the interpolated image, while all other metrics show up to 7% improvement compared to the interpolation, which means that in the case of sharpness, the ZSSR algorithm maintains better sharpness than standard interpolation on higher scales. In Figure 5 it is shown that across all scale factors, the ZSSR presents better results in both PIQE and  $\Delta DoM$  metrics, while NIQE shows a slight advantage to the ZSSR up to the scale factor of 4, the bicubic interpolation shows better results in higher scale factors, as for BRISQUE the results of ZSSR are better up to a scale factor of 6. Both BRISQUE and NIQE are based on predetermined image statistics values, where on higher scales can determine image artifacts as unnatural.



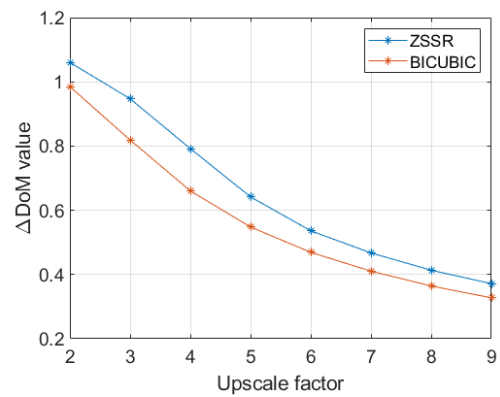
(a)



(b)



(c)



(d)

Figure 5. Plots showing the differences in metric values between ZSSR and bicubic interpolation across 9 scale factors. (a) PIQE values – lower values indicate better image quality, (b) NIQE values – lower values indicate better image quality, (c) BRISQUE values – lower values indicate better image quality, and (d)  $\Delta DoM$  values – higher values indicate better image sharpness.

#### 4. DISCUSSION

In [12], the author measured the image quality by using full reference methods like SSIM and PSNR, and showed the improvement on an RGB image, in our case, we only use non-reference image quality assessment metrics, and we show the appearance of features by measuring the change between pixel intensity on a cross-section line. [6] also doesn't measure appearance of those features in ZSSR or any other algorithm. It is necessary to assess the reveal of hidden features in the case of super resolution as one of the goals is to reveal information that was hidden beforehand. When using evaluation methods without any reference image, it's difficult to determine false information as those methods rely mostly on image quality and sharpness. Although showing that the ZSSR method provides a sharper image quality and shows results that are closer to a natural scene statistic when using internal features in comparison to a standard bicubic interpolation it is important to understand the limitations of this method when considering an anomaly in a reoccurring space. If the anomaly goes below Rayleigh criterion or the sampling is much below the Nyquist frequency a false information or a case of information ignorance may occur, an anomaly may be considered as noise.

In the case of using non-reference image quality assessment methods based on natural scene statistics, it is important to consider on which images the method was trained on. NSS may have a slight difference when looking solely at remote sensing earth observation images. The images tested are single band images of SWIR wavelength, it is difficult to estimate accuracy of the evaluation method when the main used methods were not tested on these specific settings.

While using the deep neural network, network adjustments for optimization might provide better results, in the case of remote sensing images, low level features are more common, thus a network with less layers might yield better results. As for the filters, there are many changes, a wider network can be considered for the extraction of more reoccurring features. By using the ZSSR method, all the limitations are yet to be seen, and should be further studied.

## 5. CONCLUSIONS

Utilizing deep internal learning significantly enhances the super resolution of remote sensing imagery compared to conventional methods such as interpolations. Evaluation of product quality involves qualitative visual assessment and diverse quantitative analyses encompassing NSS and image sharpness determination. Findings indicate substantial enhancements in various resolution scale factors through single image super resolution via internal learning. This approach is adaptable to diverse satellite image types without necessitating prior training and is tailored to individual images. Research demonstrates significant outcomes for a single spectral band, revealing the detection of previously imperceptible features through the Zero-Shot Super Resolution (ZSSR) technique. Implementing the algorithm enables a fourfold increase in image resolution, surpassing the quality of standard bicubic interpolation and, in certain quality metrics, achieving up to a ninefold scale factor enhancement

## ACKNOWLEDGEMENTS

Thanks to Prof. Ofer Hadar, Dr. Itay Dror, Mrs. Divya Mishra, Prof. Daniel Choukroun and Dr. Alexander Shyriayev for their helpful comments. This research was supported by the Ministry of Innovation, Science, and Technology of Israel (Grant No. 96924).

## REFERENCES

- [1] P. Wang, B. Bayram and E. Sertel, "A comprehensive review on deep learning based remote sensing image super-resolution methods," *Earth-Sci. Rev.*, vol. 232, pp. 104110, 2022.
- [2] M. Sdraka *et al*, "Deep learning for downscaling remote sensing images: Fusion and super-resolution," *IEEE Geoscience and Remote Sensing Magazine*, vol. 10, (3), pp. 202-255, 2022.
- [3] C. Ledig *et al*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
- [4] B. Lim *et al*, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017.
- [5] J. Wang *et al*, "Unsupervised remoting sensing super-resolution via migration image prior," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*, 2021.



- [6] D. Mishra and O. Hadar, "Self-FuseNet: data free unsupervised remote sensing image super-resolution," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 1710-1727, 2023.
- [7] M. Deudon *et al*, "Highres-net: Recursive fusion for multi-frame super-resolution of satellite imagery," *arXiv Preprint arXiv:2002.06460*, 2020.
- [8] F. Salvetti *et al*, "Multi-image super resolution of remotely sensed images using residual attention deep neural networks," *Remote Sensing*, vol. 12, (14), pp. 2207, 2020.
- [9] D. Ruderman and W. Bialek, "Statistics of natural images: Scaling in the woods," *Advances in Neural Information Processing Systems*, vol. 6, 1993.
- [10] A. Shocher, N. Cohen and M. Irani, "'zero-shot' super-resolution using deep internal learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [11] M. Zontak and M. Irani, "Internal statistics of a single natural image," in *Cvpr 2011*, 2011.
- [12] Y. Junzhi, "Zero-shot super resolution for satellite remote sensing images," in *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*, 2019.
- [13] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, (12), pp. 4695-4708, 2012.
- [14] A. Mittal, R. Soundararajan and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, (3), pp. 209-212, 2012.
- [15] N. Venkatanath *et al*, "Blind image quality evaluation using perception based features," in *2015 Twenty First National Conference on Communications (NCC)*, 2015.
- [16] J. Kumar, F. Chen and D. Doermann, "Sharpness estimation for document and scene images," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, 2012.