Performance Study of Upscaling Techniques for Massive Remote Sensing Data in Low Bandwidth Environments

Nicolas Escobedo[1], Jason Phan[1], Kalin Zaluzec[1], John Korah[1]

[1]California State Polytechnic University, Pomona, United States

Abstract

Satellite images of land cover are crucial for precision agriculture, but the size and volume of the data sets present challenges for efficient transfer between cloud and edge devices. We aim to identify a lightweight and accurate super-resolution model to upscale low-resolution images rather than download the full-size image. Laplacian Pyramid Super-Resolution Network (LapSRN) based super-resolution models have been shown to work well with various image datasets. We analyzed the performance of two LapSRN baseline models, a LapSRN-Fusion model, and two traditional interpolation techniques to upscale land cover images. The performance study of these models was conducted using land cover data of the Central Valley of California and analyzed using Mean Squared Error, Structural Similarity Index Measure and Runtime. We discovered that the LapSRN-Fusion performed the best in all accuracy metrics tested. Concerning runtime, it outperformed the other machine learning models tested but not conventional interpolation strategies. The experiment demonstrated the efficiency of the LapSRN-Fusion model and its accuracy when applied to satellite images of land cover. These findings have identified an effective methodology for upscaling imagery that is both accurate and does not have strict hardware and network requirements, making it more accessible to devices running on a low bandwidth network.

1. Introduction

When bandwidth is limited or the internet connection is unstable, downloading immense volumes of land cover imagery for precision agriculture can be cumbersome. In this context, we aim to reduce image data volume to minimize transmission time under bandwidth constraints. Our work examines the efficiency of multiple upscaling implementations for receiving lower-resolution images and upscaling the data into higher-resolution images. A practical method of upscaling will help achieve high accuracy at a low computation cost while lowering the effects of network latency and allowing imagery to be sent from a database at a lower resolution.

This would also allow users of precision agriculture to utilize cloud computing resources while leveraging local resources on their edge devices, giving them the ability to perform critical analyses including crop suitability modeling. As a step towards achieving this objective, our research goal is to explore and identify efficient methods for upscaling large remote sensing images. We will also consider the tradeoff between performance and data volume. We evaluate multiple image upscaling strategies including interpolation and deep learning-based LapSRN variants (see Section 4 for model descriptions). The goal is to identify methods that can quickly and accurately reconstruct high-resolution images with minimal computational overhead, enabling timely and reliable analysis in precision agriculture applications. A comparative performance study was conducted using a dataset of satellite images. Each image was upscaled by the various upscaling techniques and compared to the ground truth to calculate Mean Squared Error and Structural Similarity Index Measure. Additionally, we measured the runtime of each technique. We hypothesize that the LapSRN-Fusion model will perform the best in the accuracy metrics and will be less computationally intensive than the original LapSRN model. In the remainder of the paper, we review our problems and objectives, provide a technical background of the upscaling techniques, our performance comparison methodology and a description of the experimental results and analysis.

2. Literature Review

Remote Sensing data has become pivotal in the agricultural industry and the internet and bandwidth can potentially be technical barriers for rural stakeholders. Many times, these stakeholders do not have access to consistent high bandwidth internet. This hinders appropriate use of precision agriculture-based analysis [13]. This point is made evident by cloud data analytics being used at a low 21% [7] across farms in the mid-west.

We will be testing refined super-resolution techniques to overcome the hardware and internet constraints for precision agriculture applications. Traditionally, lightweight upscaling techniques such as Bicubic [4] and Bilinear interpolation algorithms are commonly used to upscale images. Despite this, these algorithms do not meet the accuracy levels of the more recent deep learning-based methods. Within this domain, certain deep learning approaches have proven to be very successful with upsampling specific image sets [3]. For remote sensing however, there are no consistent land cover datasets and metrics to validate such models. Our work helps to bridge this technical gap.

Many deep learning models have been developed for other super-resolution applications and their performances have been well documented and analyzed. For example, *Wang et al.* [1] provides a comparative performance analysis for a deep neural network architecture based on Generative Adversarial Networks (GANs) and GANs performed accurate upscaling. However, two issues arise when using GANs for our case study. Firstly, GANs are computationally expensive and require a significant amount of computing power to train. Secondly, datasets with low variation such as ours can cause model collapse within GANs [9,10].

LapSRN [8] is a well-known framework that works efficiently and provides accurate results for images including text, people, and geometric shapes. For text in particular, LapSRN outperformed all other upscaling techniques with regard to SSIM. A modified LapSRN based technique named LapSRN-Fusion has also been created as a super-resolution method that does spatio-temporal fusion [12]**.** This will use the feature map of a past, high-resolution image and combine it with the present low-resolution image. LapSRN-Fusion's will output a high-resolution image of the current time. At a lower computational cost, LapSRN has proven its ability to output accurate predictions comparable to other costly deep learning models. We hypothesize that the LapSRN-Fusion based techniques are a better fit for upscaling satellite images and provide experimental validation for this, because of the availability of temporally sparse high-resolution data.

3. Technical Background

This section introduces image super-resolution techniques and frameworks that we use in our study. For our work, we used an improved LapSRN based framework called LapSRN-Fusion. We compared this upscaling technique with the baseline LapSRN and other traditional upscaling techniques such as Bilinear Interpolation and Bicubic Interpolation.

3.1 Baseline LapSRN Model

The LapSRN framework [8], which is widely used in super-resolution and serves as a baseline model for our comparative performance analysis. The model receives a low-resolution image as input and progressively predicts the residuals (features) of the high-resolution image using feature maps that form a Laplacian Pyramid. Feature extraction and image reconstruction are the key branches in the model (see Figure 1) with the feature extraction branch providing inputs to the image reconstruction branch.

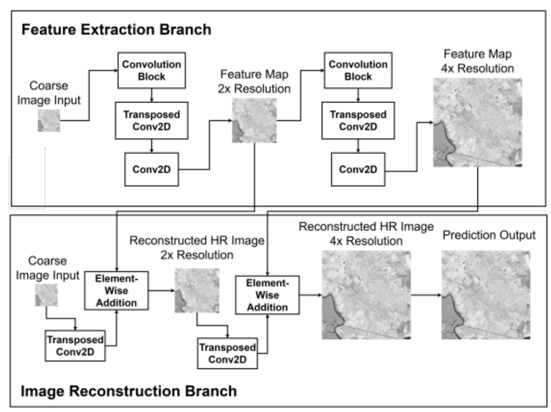

Figure 1. Architecture of LapSRN Model [8].

Feature Extraction Branch: This branch extracts features by using several convolutional layers followed by a transposed convolution layer to control the output size. Each iteration of image reconstruction doubles the image size. Each level refines the feature map at different resolutions and prepares it for the reconstruction branch. In Figure 1, a convolution block refers to a cascade of convolution layers. As seen in Figure 1, the feature extraction branch generates feature maps at different resolutions to be used at different stages in the image reconstruction branch.

Image Construction Branch: This branch uses a transposed 2D convolution network to increase the resolution of the image, doubling in size every pass through. The residuals predicted by the feature extraction branch are reapplied to the upscaled image using element-wise addition to facilitate reconstruction. After two passes, the image resolution becomes four times the original resolution. LapSRN utilizes the robust Charbonier loss function which is similar to MSE that allows the model to be less negatively impacted by outliers and high entropy.

3.2 LapSRN-Fusion Model

The LapSRN-Fusion model [2] builds upon LapSRN by incorporating high resolution images from the past. The LapSRN-Fusion model requires both a low-resolution image and a previous high-resolution image of the same region.

In this model, the feature extraction branch creates the residuals from the past high-resolution image, which is then used by the reconstruction branch. Rather than increasing the size, the transposed convolution layer in LapSRN-Fusion either keeps the resolution or reduces it to match the target resolution. The size of the output tensor is controlled by the stride parameter of the 2D convolution layer.

3.3 Bilinear and Bicubic Interpolation

Bilinear Interpolation estimates the pixel value of an image by averaging the four nearest pixels in a 2D grid. It performs linear interpolation first in one direction (e.g., x-axis) and then in the other (y-axis). Bicubic Interpolation considers the 16 nearest pixels (4×4) and applies cubic interpolation in both the x and y directions for more accurate upscaling. These interpolation techniques are used as a basis for comparison for the LapSRN frameworks. We will compare the speed and accuracy of the LapSRN models against the interpolation techniques as a baseline.

4 Methodology

We conducted a comparative performance study to examine the effectiveness of the LapSRN based techniques across months of land cover data. We will use a variety of metrics to evaluate the computational time and accuracy of each model.

4.1 Comparative Model Performance

The LapSRN-Fusion model is a variant of the fusion model that was formulated in previous work by co-authors Zaluzec *et al.* [2] Compared to the original implementation, the tested model had the convolutional neural network layers reduced from 10 to 4 layers. This modification was made after preliminary testing which showed that the reduction in layers provided a lower computational cost model without significantly diminishing accuracy values. A model with lower computational is more relevant to our focal use cases-edge computing for agricultural applications.

The LapSRN-Fusion model will be compared against the two variations of Lai's LapSRN models, namely LapSRN-Baseline1 and LapSRN-Baseline2 models. LapSRN-Baseline1 is a pretrained model, provided by Lai et al[1]. This model performed well on the Urban100 and the BSD100 [8] datasets which consist of images of buildings and landscape. However, its performance on land cover imagery has not been studied. LapSRN on BSD100 and Urban100 datasets, we applied the framework to Sentinel-2 satellite imagery with agricultural applications, this model is our LapSRN-Baseline2.

The training process for the LapSRN-Fusion and LapSRN-Baseline2 model are similar and use the same land cover imagery dataset. An important distinction is that our LapSRN-Fusion model's training dataset also includes past images. Also included in the performance study are two existing implementations of the traditional Bicubic and Bilinear interpolation upscaling techniques. These lightweight algorithms were chosen due to their wide use across the domain as well as their lower computational requirements. These algorithms deserve consideration because they do not require any training overhead.

4.2 Accuracy Metrics

The following metrics, specifically focusing on runtime and accuracy, will be used to compare the performance. The first set of analyses will compare the runtime for processing single images in the dataset. The next set of analyses will focus on accuracy. All accuracy analyses will compare the model's output to the high-resolution image that was retrieved from google earth.

We will use Mean Squared Error (MSE) to compare

[1]https://github.com/phoenix104104/LapSRN

accuracy across the model outputs. Mean Square Error is calculated using the formula $\frac{1}{n}\sum_{i=1}^{n}(y_i - p_i)^2$, where y is the observed value and p is the predicted value. MSE calculates the average accuracy across multiple datapoints and quantifies the total difference of our predicted image compared to the ground truth. Finally, Structural Similarity Index Measure (SSIM) [11] is another accuracy measure that will be used in the analysis. Given two non-negative values x and y chosen based on the range of pixel values (0 –255) and their means, μ_x and μ_y as well as their variance, σ_x and σ_y and finally two smoothening constants $C_1$and C_2, the SSIM is by the following: $SSIM(x,y) = \frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}$ SSIM complements MSE as MSE quantifies global errors across every point while SSIM prioritizes structural errors and is a metric for human perception accuracy. SSIM values range between 0-1, with higher values being more favorable [11].

4.3 Research Hypotheses

We hypothesize that LapSRN-Fusion would perform better than LapSRN-Baseline1 and LapSRN-Baseline2 in both the runtime and accuracy metrics by a significant margin. This is because the convolutional layers were minimized. We also believed that since we are using high-resolution data to create our image residuals, our model will produce higher quality images with respect to MSE and SSIM. We further hypothesize that both Bilinear interpolation and Bicubic interpolation will perform similarly in all our performance metrics, and significantly better in runtime than the machine learning models. However, we predict that all the three LapSRN based models will outperform the interpolation algorithms in the accuracy metrics.

For experimental validation, we will compare the performance of the five techniques on a large land cover image dataset using MSE, SSIM and runtime measures.

5 Experimental Validation

This section describes the experiment conducted and data observed.

5.1 Experimental Set Up

The data used to test the models consisted of satellite imagery of land cover across 665 regions of central California, captured in each April and July of 2022. These images were captured by the Sentinel-2 satellite and are retrieved via google earth [2]monthly. We retrieved low-resolution images and the high-resolution images of each location, directly from google earth, at sizes 128x128 and 512x512, respectively. We had a 90%-10% ratio between our test set and validation set. The models were run on a NVIDIA T4 GPU via Google Colab.

5.2 Experimental Results and Analysis

The graphs show the average of each datapoint where T_0 is March or June and $T_1$is April or July. The averages are then graphed to compare the models.

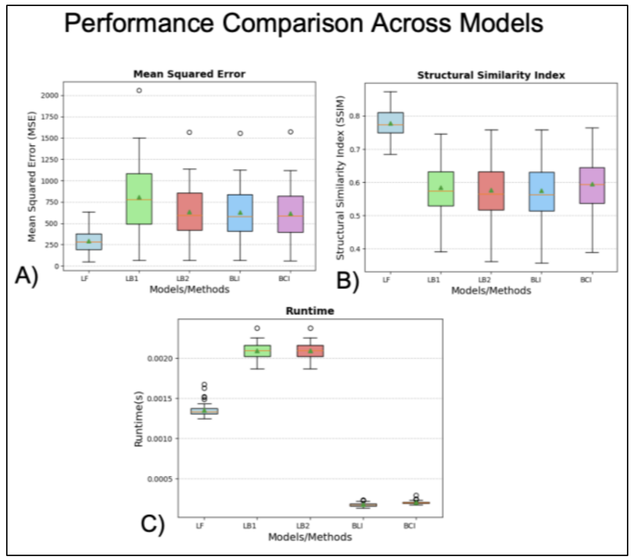

Figure 2: Performance comparison of models using the metrics: A) MSE B) Structural Similarity Index Measure (SSIM) and C) Runtime. Each model name represents: LapSRN-Fusion (LF), LapSRN-Baseline1 (LB1), LapSRN-Baseline2 (LB2), Bilinear Interpolation (BLI), Bicubic Interpolation (BCI)

As seen in Figure 2, the experiment shows that LapSRN-Fusion achieved the lowest MSE, followed by LapSRN-Baseline2, the interpolation models and finally LapSRN-Baseline1. This is due to LapSRN-Fusion incorporating the past high-resolution image into the upscaling. This high-resolution data is used by the other models, which is reflected in the substantial difference in MSE values. Similarly, LapSRN-Fusion outperforms all other models in SSIM (Figure 2(B)). This is due to the origin image of our extracted edges being a higher resolution image.

[2] https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_HARMONIZED

As expected, the interpolation algorithms had the fastest runtime due to their low computational cost. They outperformed all the machine learning models, while LapSRN-Fusion also outperformed the two LapSRN-Baseline models in this regard. It is important to note that the figures shown in the paper are an average across two months. While still outperforming the other tested models with respect to the accuracy metrics, the LapSRN-Fusion model performed better in July than in April T_0 and T_1 images. This is potentially because of large changes in land cover over the observed month.

Based on our findings, the LapSRN-Fusion is a practical choice for upscaling land cover satellite images due to its high accuracy to computational cost ratio when compared to the other machine learning models.

6. Conclusion and Future Directions

We would like to continue applying LapSRN-Fusion to agricultural imagery, testing a larger dataset and testing different architectures. This includes examining the difference in accuracy between a model with SSIM included in the loss function versus without. We also plan to test our framework in other, more volatile regions than Central Valley, California.

The Lai *et al.* [8] implementation produces comparable SSIM and MSE values to the Interpolation techniques but at a slower rate. The Fusion model has a faster runtime than the Baselines, while remaining slower than the Bicubic and Bilinear interpolation algorithms. The higher computational cost comes with an increased accuracy across all data points. The preferred algorithm will depend on the circumstances of use. If accuracy is of the highest priority, then a user might want to directly download the ground truth image despite potentially high bandwidth costs. If both bandwidth and accuracy are of equal concern, then we would recommend LapSRN-Fusion. Finally, if accuracy is only a secondary concern, then a user might opt to download small images and use conventional interpolation algorithms to upscale them. As it is often the case that we value both image quality and bandwidth, we believe that LapSRN fusion is an alternative to existing algorithms that offers a reasonable tradeoff between accuracy and bandwidth conservation.

7 References

[1] X. Wang, L. Sun, A. Chehri, and Y. Song, "A Review of GAN-Based Super-Resolution Reconstruction for Optical Remote Sensing Images," *Remote Sensing*, vol. 15, no. 20, Art. no. 20, Jan. 2023.

[2] K. Zaluzec and J. Korah, "'An Edge Computing Framework for Fusing Geospatial Data Using Laplacian Super Resolution Networks'," California State Polytechnic University, Pomona, KRG-05-2025-1.

[3] N. Yokoya, "Chapter 2 - Deep learning for super-resolution in remote sensing," in *Advances in Machine Learning and Image Analysis for GeoAI*, S. Prasad, J. Chanussot, and J. Li, Eds., Elsevier, 2024, pp. 5–26.

[4] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981.

[5] K. Li, W. Xie, Q. Du, and Y. Li, "DDLPS: Detail-Based Deep Laplacian Pansharpening for Hyperspectral Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 8011–8025, Oct. 2019.

[6] R. Wongso, F. A. Luwinda, and Williem, "Evaluation of Deep Super Resolution Methods for Textual Images," *Procedia Computer Science*, vol. 135, pp. 331–337, Jan. 2018.

[7] G. S. Hundal, C. M. Laux, D. Buckmaster, M. J. Sutton, and M. Langemeier, "Exploring Barriers to the Adoption of Internet of Things-Based Precision Agriculture Practices," *Agriculture*, vol. 13, no. 1, Art. no. 1, Jan. 2023.

[8] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Fast and Accurate Image Super-Resolution with Deep Laplacian Pyramid Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, Oct. 2017.

[9] D. Saxena and J. Cao, "Generative Adversarial Networks (GANs): Challenges, Solutions, and Future Directions," *ACM Comput. Surv.*, vol. 54, no. 3, p. 63:1-63:42, May 2021.

[10] Y. Kossale, M. Airaj, and A. Darouichi, "Mode Collapse in Generative Adversarial Networks: An Overview," in *2022 8th International Conference on Optimization and Applications (ICOA)*, Oct. 2022, pp. 1–6.

[11] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, Nov. 2003, pp. 1398-1402 Vol.2.

[12] J. Cai, B. Huang, and T. Fung, "Progressive spatiotemporal image fusion with deep neural networks," *International Journal of Applied Earth Observation and Geoinformation*, vol. 108, p. 102745, Apr. 2022.

[13] S. B. Damsgaard, N. J. Hernández Marcano, M. Nørremark, R. H. Jacobsen, I. Rodriguez, and P. Mogensen, "Wireless Communications for Internet of Farming: An Early 5G Measurement Study," *IEEE Access*, vol. 10, pp. 105263–105277, 2022.

Address for correspondence:
Nicolas Escobedo
3801 W Temple Ave, Pomona, CA
Nescobedo@cpp.edu