A Comprehensive Analysis of Generative Adversarial Networks for Single Image Super Resolution

Jessica Pinto, John Korah (Faculty Mentor)

Department of Computer Science
California State Polytechnic University, Pomona

Introduction

- Single image super resolution adds finer details to low resolution images
- Critical during emergencies and reconstruction efforts for mapping
- Expensive to load and store a multitude of high resolution images

Background

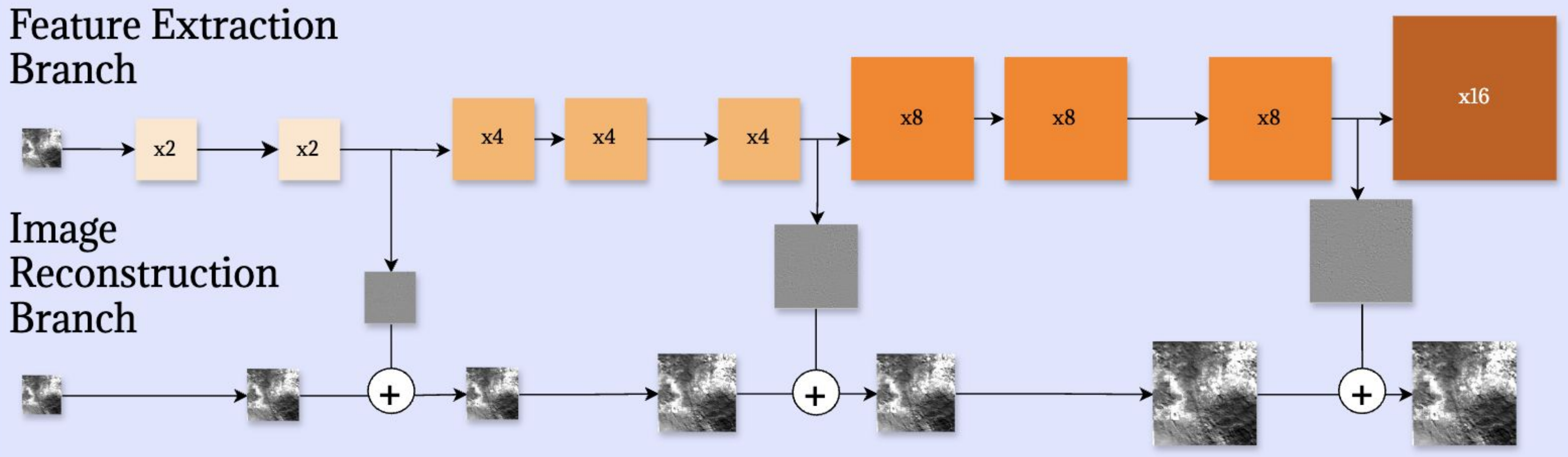
- ***SISR -*** Single Image Super Resolution
  - Model high resolution (HR) images from low resolution (LR) input
- ***LapSRN -*** Laplacian Pyramid Super Resolution Network
- ***Real-ESRGAN -*** Real Enhanced Super Resolution Generative Adversarial Network
- *State of the art:*
  - Densely Residual Laplacian Network (DRLN+)
  - Residual Channel Attention Network (RCAN)
  - Residual Dense Network (RDN)

Research Objectives

- Make high resolution images more accessible
- Use trained models to increase the resolution of low resolution images by adding more detail
- Minimize artifacts and hallucinations

Methodology

LapSRN



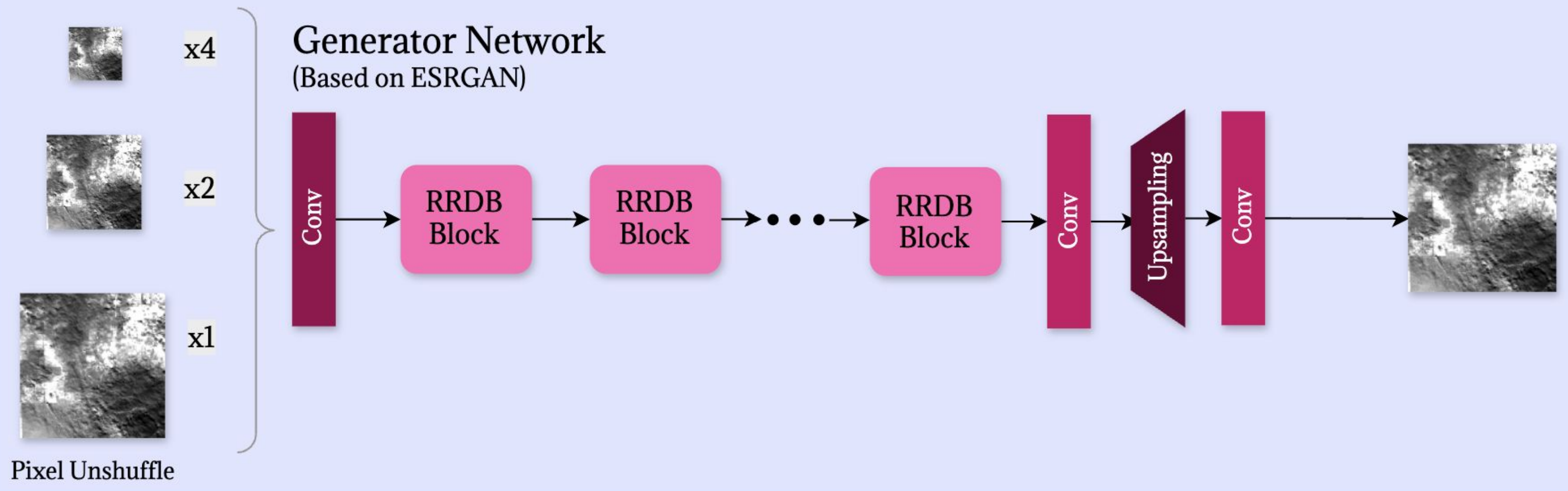
- Learn to predict residuals in training
- Charbonnier loss

$$L(x, y) = \sqrt{(x-y)^2 + \epsilon^2}$$

x = model output
y = ground truth
∈ = constant value

- Upsampled Gaussian pyramid[3]

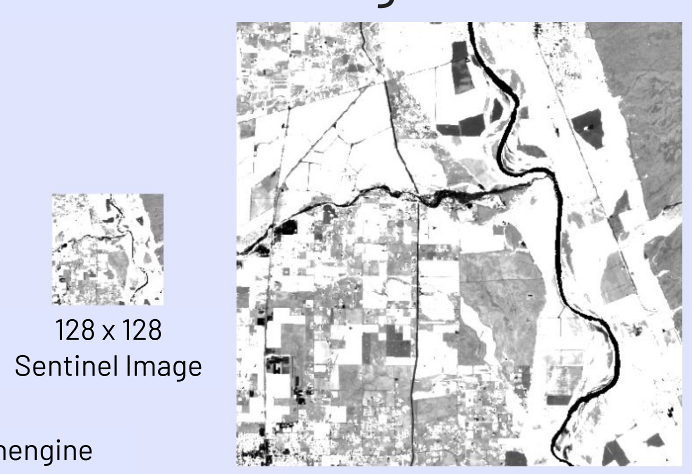
Real-ESRGAN



- Minimize generator loss, maximize discriminator loss[4]
- Residual-in-Residual Dense Blocks (RRDB)
- DScore = 1 ⟶ Real

Experimental Setup

- Images taken from **Google Earth Engine** *
  - LR = 128 x 128, HR = 512 x 512
- Trained as pairs of labeled HR and LR images
  - Taken from May, June, July, August
- Total of 50 epochs, 40 mins - 1.5 hours
- Real-ESRGAN takes HR and LR folders
- LapSRN takes single HDF5 file



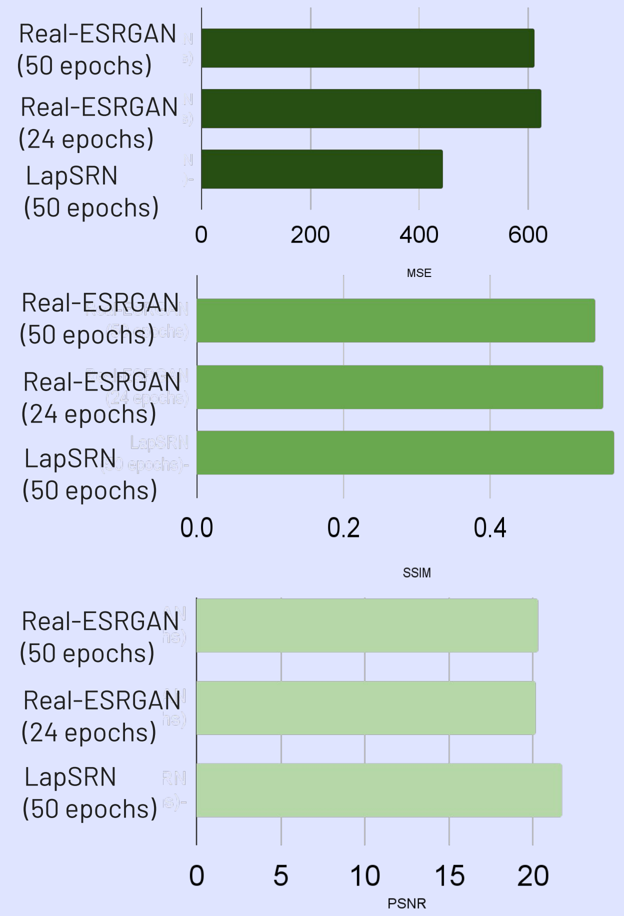
128 x 128 Sentinel Image

512 x 512 Sentinel Image

*https://earthengine.google.com/

Results

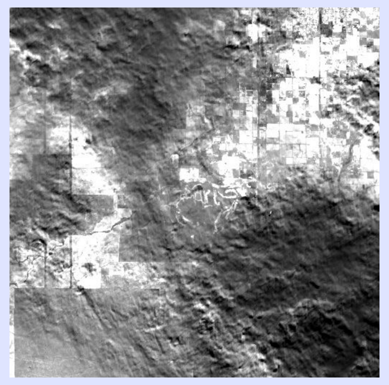
- Quantitative results ⟶ LapSRN's output surpasses Real-ESRGAN[2]
- Perceptual results ⟶ Real-ESRGAN's output surpasses LapSRN[1]



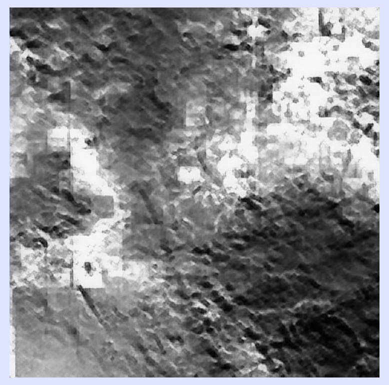
- **MSE** - Mean squared error (lower)
- **SSIM** - Structural similarity index measure (higher)
- **PSNR** - Peak signal-to-noise ratio (higher)

Conclusion & Future Work

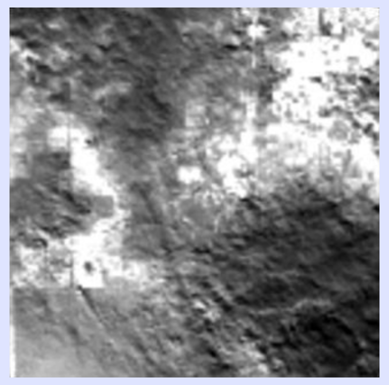
- Perceptual performance of Real-ESRGAN surpasses that of LapSRN
- Quantitative metrics represent the opposite conclusion
- Need for more accurate quantitative evaluation measures
- Incorporate features for geographic identification



Ground Truth HR Image



Real-ESRGAN Output



LapSRN Output

References

[1] J. XU, "TWTYGQYY/Pytorch-LapSRN: Pytorch implementation for lapsrn (CVPR2017)," GitHub, https://github.com/twtygqyy/pytorch-LapSRN (accessed Nov. 11, 2025).
[2] X. Wang, "Xinntao/real-ESRGAN: Real-ESRGAN aims at developing practical algorithms for General Image/video restoration.," GitHub, https://github.com/xinntao/Real-ESRGAN (accessed Nov. 11, 2025).
[3] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep laplacian pyramid networks for fast and accurate super-resolution," arXiv.org, https://doi.org/10.48550/arXiv.1704.03915 (accessed Nov. 11, 2025).
[4] X. Wang, L. Xie, C. Dong, and Y. Shan, "Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data," arXiv.org, https://doi.org/10.48550/arXiv.2107.10833 (accessed Nov. 11, 2025).

Acknowledgements

Thank you to Nico Escobedo for guidance on LapSRN models and implementations.