Comparative Analysis of Image Super-Resolution Models - Deep Machine Learning

Gritta Joshy, Qi Chen

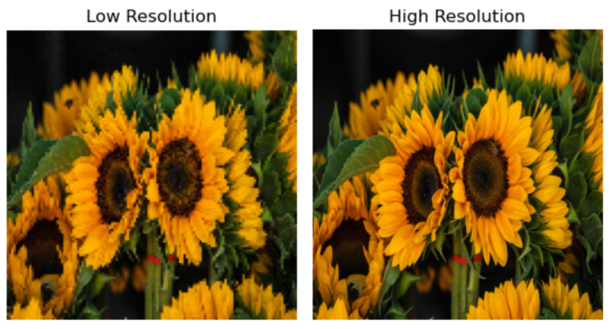
Chalmers University of Technology
Electrical Engineering Department

CHALMERS
UNIVERSITY OF TECHNOLOGY

Introduction

Have you ever taken a photo only to find it's too blurry to capture the moment perfectly?

In our digital age, the demand for high-quality images has never been greater. Image super-resolution is an advanced technique that enhances low-resolution images, bringing out details that were once hidden. This method has far-reaching applications, from improving medical scans for accurate diagnoses to sharpening satellite images for environmental monitoring.



Using the diverse and high-quality DIV2K dataset, we train and compare models to push the boundaries of image clarity and detail.

Objective

The goal of this project is to experiment with three deep learning models: EDSR,VDSR,and DRCN to upscale low resolution images to high resolution images.
The project aims to compare the performance of these models based on several quantitative and qualitative metrics: PSNR,SSIM,PI. Through this comparison, we aim to determine which model provides the best image quality for super-resolution tasks.

Dataset

The DIV2K dataset is one of the most popular datasets used for image super-resolution, which is composed of 800 images for training, 100 images for validation, and each image has a 2K resolution.

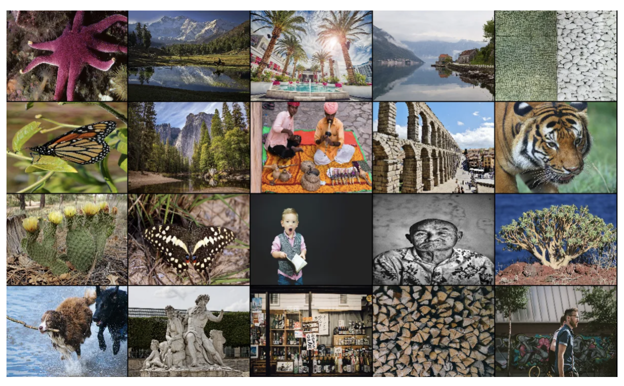
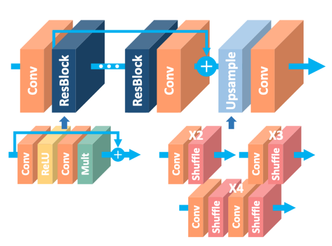


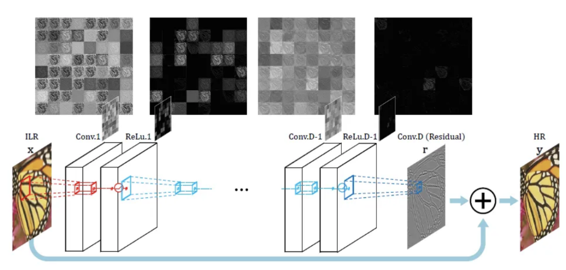
Figure 1: DIV2K Sample

Models

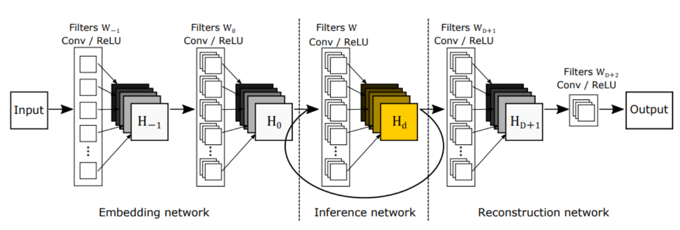
- EDSR - Enhanced Deep Residual Network.[3]



- VDSR - Very Deep Super-Resolution [1]



- DRCN - Deep Recursive Convolutional Network.[2]



Feature	EDSR	VDSR	DRCN
Architecture	- Deep Residual Network - No Batch Normalization - Post-upsampling	- Convolutional Neural Network - Pre-upsampling - Global skip connection	- Embedded , Recursive and Reconstruction network - Post-upsampling
Depth	- 32 convolutional layers - Residual blocks	- 20 convolutional layers - Single global skip connection	- 5 convolutional layers (2 in embedding, 1 recursive, 3 in reconstruction) - 16 recursive applications
Features	- Local and global residual learning - Focuses on high upscaling - Higher computational complexity	- Residual learning for faster convergence - Lightweight architecture	- Recursive structure - Weight sharing in recursive layers - Color-aware Loss - Flexible Scaling

Table 1: Comparison of EDSR, VDSR, and DRCN Architectures.

Methodology

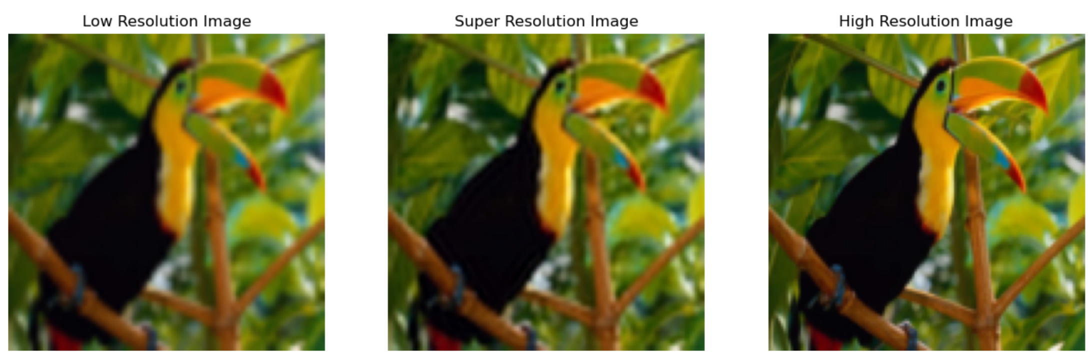
Feature	EDSR	VDSR	DRCN
Training Setup	- High-scale factors (x4) - Epochs: 20 - Optimizer: Adam - Loss: Mean Absolute Error	- Lower-scale factors (x2) - Epochs: 50 - Optimizer: SGD with momentum to 0.9 - Loss: Mean Squared Error	- Scale factor: x4 (default, flexible) - Epochs: 70 - Optimizer: Adam - Loss: Custom DRCN loss (combination of MSE, recursive loss, and color loss)
Hyperparameters	- Learning rate: 10^{-4} - Batch size: 8 - No augmentation	- Learning rate: 10^{-3} - Batch size: 8 - No augmentation	- Learning rate: 10^{-4} - Batch size: 8 - number of recursionS: 16
Evaluation Metrics	- PSNR - SSIM - PI	- PSNR - SSIM - PI	- PSNR - SSIM - PI

Table 2: Comparison of Methodologies for Training EDSR, VDSR, and DRCN

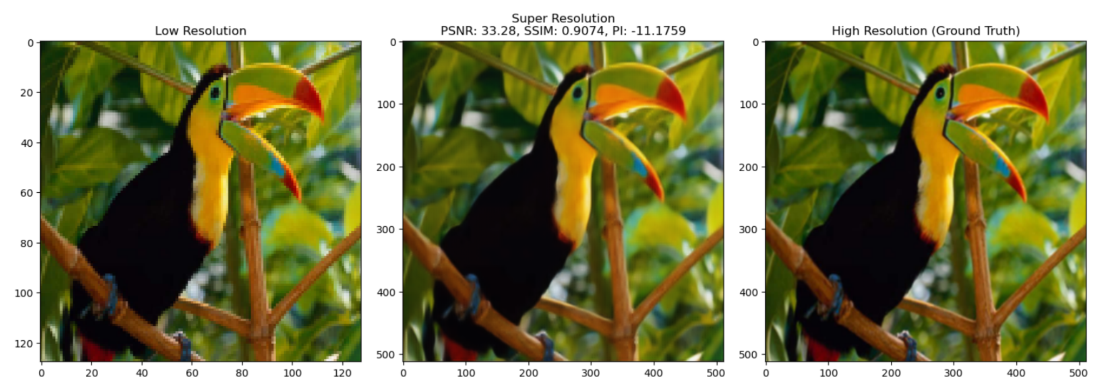
Results

- EDSR

- VDSR



- DRCN



- Average PSNR, SSIM values for each model on Set5 dataset .

Model	PSNR	SSIM	PI
EDSR	30.14	0.8666	-5.5048
VDSR	29.82	0.9574	-5.3873
DRCN	31.86	0.8876	-10.3662

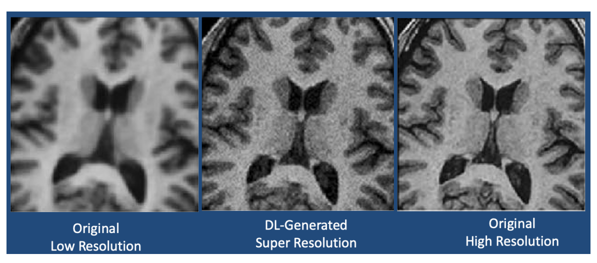
Table 3: PSNR, SSIM, and PI values for three models.

Conclusion

- Each model took at least 3 hours to train over 20 to 70 epochs. While DRCN produced better image quality overall, VDSR was only trained at a 2x scaling factor compared to 4x for the others, impacting direct comparisons.
- DRCN achieved the best image quality, with the highest PSNR and strong SSIM, thanks to its recursive structure, making it ideal for high-quality reconstruction. EDSR also performed well, offering a good balance between quality and computational cost, especially for larger scaling tasks. VDSR excelled in structural similarity but lagged in perceptual quality and PSNR.

Food for thought

In a world where clarity is of great importance, what new possibilities can arise from our advancements in image resolution?



References

[1] J. Kim, J. K. Lee, and K. M. Lee.
Accurate image super-resolution using very deep convolutional networks.
arXiv preprint arXiv:1511.04587, 2016.

[2] J. Kim, J. K. Lee, and K. M. Lee.
Deeply-recursive convolutional network for image super-resolution.
arXiv preprint arXiv:1511.04491, 2016.

[3] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee.
Enhanced deep residual networks for single image super-resolution.
arXiv preprint arXiv:1707.02921, 2017.

gusjoshgr@student.gu.se; gusqichr@student.gu.se