K. N. Toosi University of Technology

Overview of Image Super Resolution Methods

Professor: Dr. Farnaz Sheikhi

Presenter: Ali Ghanbari

November 28, 2022

Overview

1. Introduction
2. Applications
3. Evaluation Index
4. Datasets
5. Methods
6. Interpolation Based Methods
7. CNN Based Methods
8. GAN Based Methods
9. Conclusion

Introduction

Image super-resolution reconstruction refers to a technique of recovering a high-resolution (HR) image from a low-resolution (LR) degraded image [8]

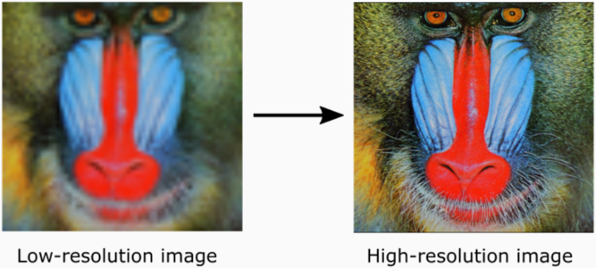


Figure 1: Transform your low-resolution images to high-resolution images [5]

Applications

Image super resolution is popularly used in the following applications [1]:

- Surveillance
- Medical
- Media

Datasets

A lot of different datasets available for training and testing super resolution models.

Popular training datasets:

- 91 Image
- BSD300
- Flickr2K

Popular testing datasets:

- Set5
- Set14
- Urban100

Evaluation Index

Current mainstream objective evaluation methods:

- Peak signal to noise ratio (PSNR)
- Structural similarity (SSIM)
- Perceptual index (PI)
- Root MSE (RMSE)

Methods

Some of the methods that we are going to look at:

- Based on Interpolation
- Based on CNN
- Based on GAN

Interpolation Based Methods

Interpolation

Interpolation works by using known data to estimate values at unknown points.

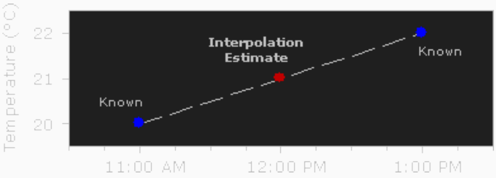


Figure 2: Linear interpolation for temperature at noon [6]

Image Interpolation

Image interpolation works in two directions, and tries to achieve a best approximation of a pixel's color and intensity based on the values at surrounding pixels.

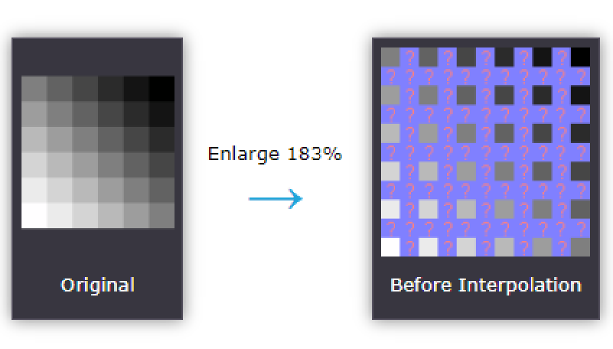


Figure 3: How resizing / enlargement works [6]

Interpolation Methods

List of interpolation algorithms:

- Nearest Neighbor Interpolation
- Bilinear Interpolation
- Bicubic Interpolation
- Method based on Directional Bicubic Interpolation
- Method based on DWT and BI

Bicubic Interpolation

Bicubic considers the closest 4x4 neighborhood of known pixels — for a total of 16 pixels. Since these are at various distances from the unknown pixel, closer pixels are given a higher weighting in the calculation. Compared to previous methods it:

- Produces Sharper Images
- Offers a Balance Between Time and Output Quality
- is the Standard Method Used in Many Image Editing Programs

Bicubic Interpolation

Bicubic considers the closest 4x4 neighborhood of known pixels — for a total of 16 pixels. Since these are at various distances from the unknown pixel, closer pixels are given a higher weighting in the calculation. Compared to previous methods it:

- Produces Sharper Images
- Offers a Balance Between Time and Output Quality
- is the Standard Method Used in Many Image Editing Programs

However it only interpolates the image edges horizontally and vertically so that the edges are vulnerable to artifacts.

Interpolation Artifacts

All non-adaptive interpolators attempt to find an optimal balance between three undesirable artifacts: edge halos, blurring and aliasing.

Figure 4: Image interpolation Artifacts [6]

Method based on Directional Bicubic Interpolation

This methods interpolates lost pixels based on local intensity and direction to better preserve sharp edges and details.

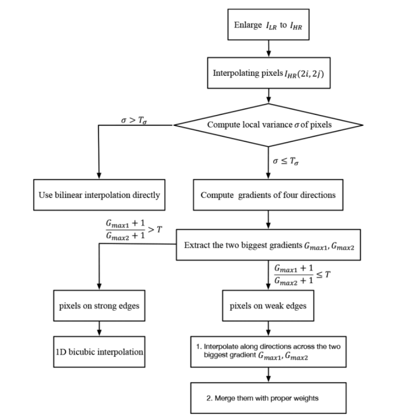


Figure 5: Flowchart of directional BI method

Method based on DWT and BI

a super-resolution technique based on the interpolation of high-frequency sub-band images obtained by the discrete wavelet transform (DWT) and input images.

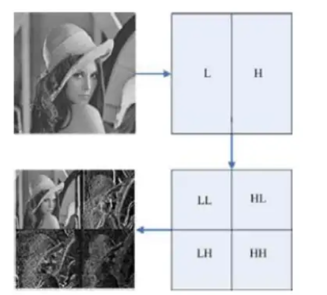


Figure 6: Discrete wavelet transform(DWT) example [10]

Method based on DWT and BI

Signal filters used on the image:

- HPF (High Pass Filter)- to extract edges
- LPF (Low Pass Filter)-for approximation

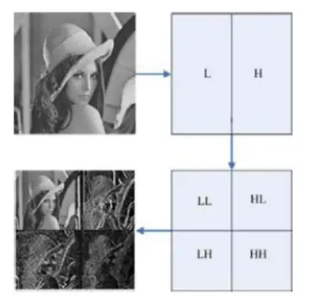


Figure 6: Discrete wavelet transform(DWT) example [10]

Method based on DWT and BI

The image is decomposes into 4 sub-band:

- LL — gives an approximation
- LH — Horizontal features (HPF along rows)
- HL — Vertical features(HPF along col.)
- HH — Diagonal features

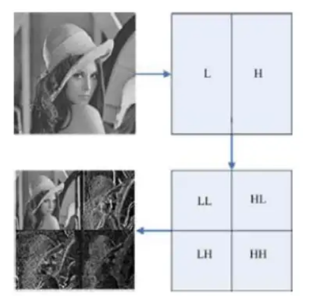


Figure 6: Discrete wavelet transform(DWT) example [10]

Method based on DWT and BI

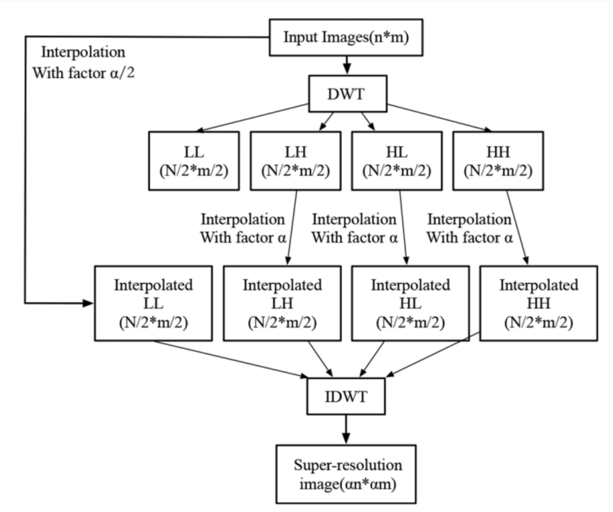


Figure 6: Flowchart of DWT and BI method

CNN Based Methods

CNN

A convolutional neural network (CNN) is a type of artificial neural network used primarily for image recognition and processing, due to its ability to recognize patterns in images [9]

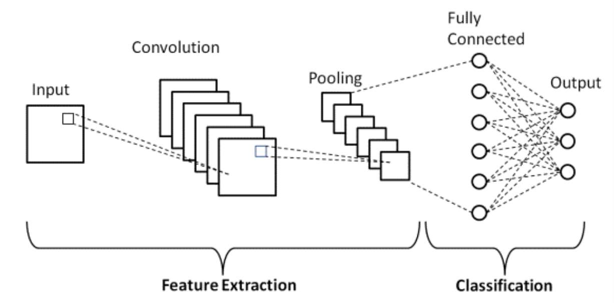


Figure 7: Basic CNN Architecture [2]

Convolution Layer

This type of layer is used to **extract features** from the input images. The output is termed as the Feature map which gives us information about the image such as the corners and edges.

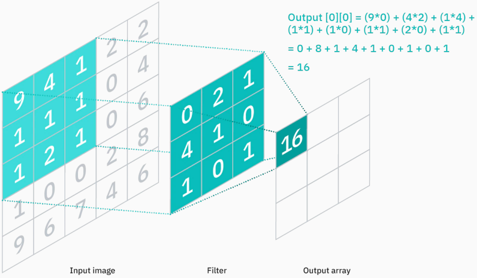


Figure 8: Basic CNN Architecture [2]

CNN Models

Most important CNN based models for Image Super Resolution:

- SRCNN
- EDSR
- WDSR
- RDN
- DSRN
- LRFNet
- RCAN
- CARN
- IDN

SRCNN

SRCNN is the pioneering work of deep learning applications in the field of super-resolution reconstruction. SRCNN's structure is very simple, using only three convolution layers.

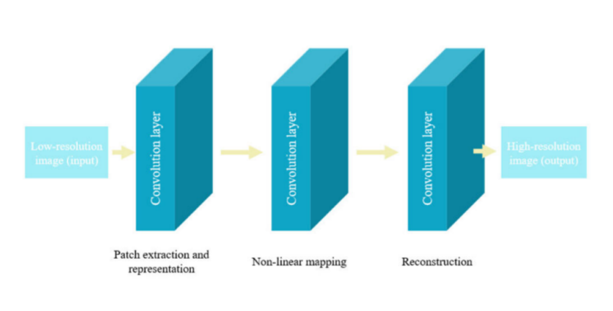
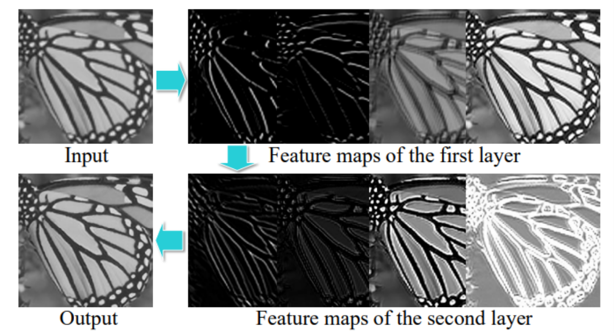


Figure 9: SRCNN structure [8]

SRCNN

For a given LR image, it uses the bicubic algorithm to zoom to the target size first and then run that through the three convolution layers:

1. Feature extraction and representation
2. Non-linear mapping
3. Reconstruction



SRCNN

SRCNN uses MSE as a loss function, which is beneficial to obtain a higher PSNR.

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^{n} \left\| F(Y_i - X_i) \right\|^2$$

The network structure of EDSR is based on the improvement of SRResNet. The batch normalisation (BN) layer in the residual block is removed, and the ReLU active layer is not set outside the residual block.

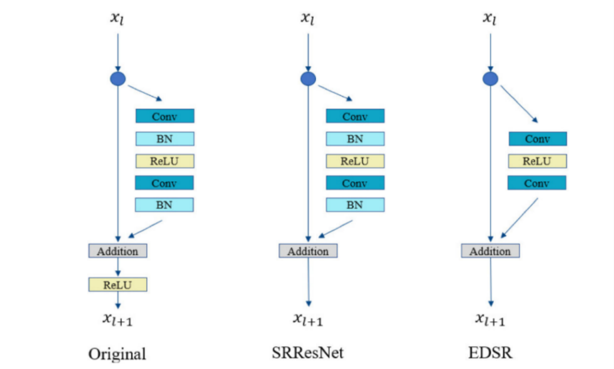


Figure 9: Comparison of the original, SRResNet and EDSR residual blocks

EDSR

Since the BN layer consumes the same amount of memory as the convolutional layer. removing it means that EDSR can stack more network layers or extract more features for each layer for better performance with the same computing resources.

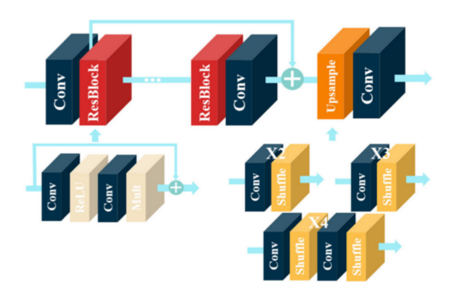


Figure 10: Network structure of EDSR

EDSR

EDSR uses the L_1 loss function to optimise the network model:

$$L_1(P) = \frac{1}{N} \sum_{p \in P} |x(p) - y(p)|$$

CNN Models

Most important CNN based models for Image Super Resolution:

- SRCNN
- EDSR
- WDSR
- RDN
- DSRN
- LRFNet
- RCAN
- CARN
- IDN

GAN Based Models

GAN

A **generative adversarial network** consists of Two neural networks that contest with each other in the form of a zero-sum game, where one agent's gain is another agent's loss.

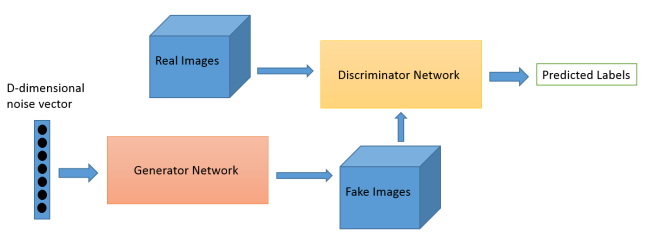


Figure 10: discriminator-generator feedback loop [3]

GAN Models

- SRGAN
- ESRGAN
- ProSR
- CinCycle
- SRFeat

SRGAN

This network is the first framework to recover $4\times$ down-sampled images. Here is the architecture of generator and discriminator network [8]:

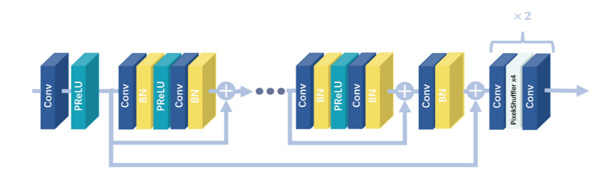


Figure 11: Generation network structure of SRGAN [8]

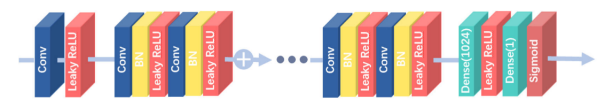


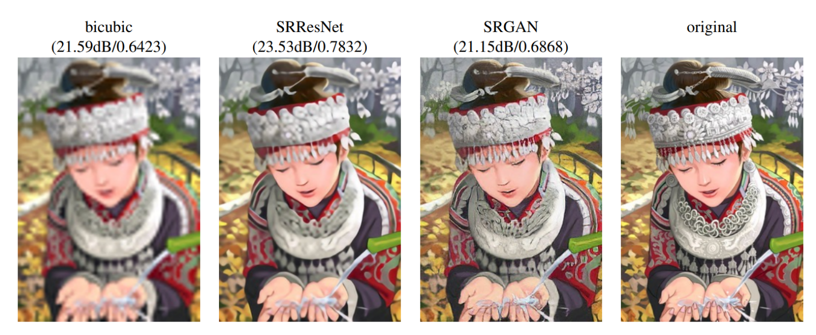
Figure 12: Discriminant network structure of SRGAN [8]

SRGAN

The structure also modified the loss function from the mean squared loss function to a new **perceptual loss** function consisting of resistance loss and content loss.

SRGAN

An example photo-realistic image that was super resolved with a $4\times$ upscaling factor is shown below with corresponding PSNR and SSIM are shown in brackets [7]:



ESRGAN

ESRGAN is a model proposed based on the improvement of SRGAN. it aims to improve the three key parts of SRGAN:

- Network Structure
- Adversarial Loss
- Perceived Loss

ESRGAN - Network Structure

First, the the BN layers in residual blocks from SRGAN are removed:

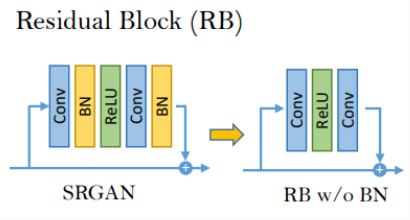


Figure 11: New Residual Block Structure [11]

ESRGAN - Network Structure

Second, RRDB block is used in our deeper model and β is the residual scaling parameter:

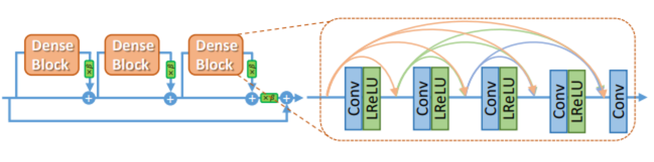


Figure 12: Residual in Residual Dense Block [11]

ESRGAN - Adversarial Loss

which estimates the probability that one input image x is real and natural, a relativistic discriminator tries to predict the probability that a real image x_r is relatively more realistic than a fake one x_f

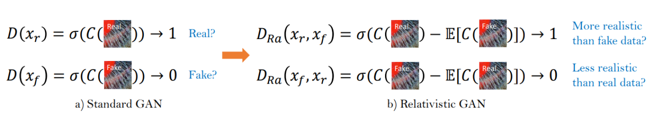


Figure 13: Difference between standard discriminator and relativistic discriminator [11]

ESRGAN - Perceived Loss

ESRGAN proposes a more efficient perceptual domain loss, using pre-activation features. This will overcome two shortcomings:

1. Sparse activated features
2. Remove inconsistent reconstructed brightness compared with the ground-truth image

ESRGAN

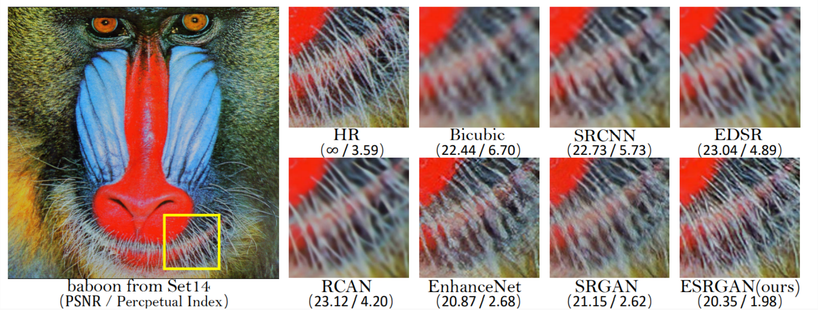


Figure 14: ESRGAN produces more natural textures [11]

GAN Models

- SRGAN
- ESRGAN
- ProSR
- CinCycle
- SRFeat

Conclusion

- Evaluation indexes: PSNR, SSIM, PI and RMSE
- Interpolation-based SR: BI-based, DWT-based
- CNN-based SR: SRCNN, EDSR, ...
- GAN-based SR: SRGAN, ESRGAN, ...

References i

- A Review of Image Super-Resolution. https://blog.paperspace.com/image-super-resolution/.
- Basic CNN Architecture: Explaining 5 Layers of Convolutional Neural Network — upGrad blog — upgrad.com. https://www.upgrad.com/blog/basic-cnn-architecture/.
- A beginner's guide to generative adversarial networks (gans) — pathmind. https://wiki.pathmind.com/generative-adversarial-network-gan.
- Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. **Learning a deep convolutional network for image super-resolution.** In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 184–199, Cham, 2014. Springer International Publishing.

References ii

Priya Dwivedi.
Super-resolution imaging using deep learning algorithms — becominghuman.ai.
https://becominghuman.ai/super-resolution-imaging-using-deep-learning-algorithms-a206ad04ff

Cambridge in Colour.
Understanding digital image interpolation.

Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew P. Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi.
Photo-realistic single image super-resolution using a generative adversarial network.
CoRR, abs/1609.04802, 2016.

References iii

Kai Li, Shenghao Yang, Runting Dong, Xiaoying Wang, and Jianqiang Huang.
Survey of single image super-resolution reconstruction.
IET Image Processing, 14(11):2273–2290, 2020.

Arm Ltd.
What is a convolutional neural network (CNN) — arm.com.
https://www.arm.com/glossary/convolutional-neural-network.

Reshu Singh.
2D-DWT : A brief intro — medium.com.
https://medium.com/image-vision/2d-dwt-a-brief-intro-89e9ef1698e3.

Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, and Xiaoou Tang.
ESRGAN: enhanced super-resolution generative adversarial networks.
CoRR, abs/1809.00219, 2018.

Thank you for your attention!