International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307 OPEN ACCESS

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT2410492

Convolution Neural Network Approach for Single Image Super Resolution

Lajja Dave[1], Neha Patel[2], Nayana Suresh[1]

[1]Assistant Professor, Sardar Patel College of Engineering, Bakrol, Gujarat, India

[2]Teaching Assistant, Sardar Patel College of Engineering, Bakrol, Gujarat, India

ARTICLEINFO

Article History:

Accepted : 30 Nov 2024

Published: 31 Dec 2024

Publication Issue

Volume 10, Issue 6

November-December-2024

Page Number

2403-2408

ABSTRACT

The goal of Super-Resolution (SR) is to generate a higher-resolution image from lower-resolution input images. High-resolution images offer more pixel density, thus capturing finer details of the original scene. Single Image Super-Resolution (SISR) seeks to restore a high-resolution image from a single low-resolution input, which is a significant challenge in computer vision. This process involves using the low-resolution image as the input and the high-resolution image as the reference, with the SR model producing the predicted high-resolution output. This paper proposes a neural network-based approach utilizing convolutional layers to improve Peak Signal-to-Noise Ratio (PSNR) and reduce processing time compared to traditional methods. The architecture consists of a convolutional layer, a max-pooling layer, and a reconstruction layer.

Keywords: Super-Resolution, Image Restoration, Convolutional Neural Network, PSNR, Image Processing.

Introduction

super-resolution has become a crucial aspect of image processing aiming to generate high-resolution hr images from one or multiple low-resolution images of the same scene during image capture the resulting images are often low in resolution due to limitations in equipment performance or environmental factors [1] hr images are essential for extracting more detailed information and are required for further image processing and analysis not only do hr images provide better visual quality but they also enable more effective analysis and feature extraction in various applications such as medical imaging eg mri and pet satellite imaging and target recognition for instance in satellite imaging applications like remote sensing and landsat multiple images of the same area are typically available allowing super-resolution techniques to enhance the targets resolution another application is the conversion of ntsc video signals to hdtv where sr techniques are used to display sdtv signals on hdtv screens without visual artifacts [3] the spatial resolution of captured images is largely determined by the sensors size and the density of the detectors that form the sensor increasing spatial resolution through hardware involves either reducing the detector size or equivalently increasing detector density or enlarging the sensor size however smaller detectors have drawbacks including lower dynamic range reduced low-light sensitivity higher dark signal increased diffraction sensitivity and higher similarity all of which negatively impact image quality additionally hardware costs [4] increase with higher detector density or larger sensor size limiting the maximum achievable resolution using hardware alone besides sensor constraints other factors also limit the capture of hr images such as lens blur atmospheric disturbances finite shutter speed aperture limitations object movement sensor noise and media turbulence a common hardware solution to increase spatial resolution is to reduce pixel size ie increase pixel density but this reduces the amount of light captured leading to shot noise and degraded image quality high costs for precision optics and image sensors are also significant concerns in hr imaging as a result there is a growing need for new approaches to increase spatial resolution addressing the limitations of sensor and optics manufacturing technology.

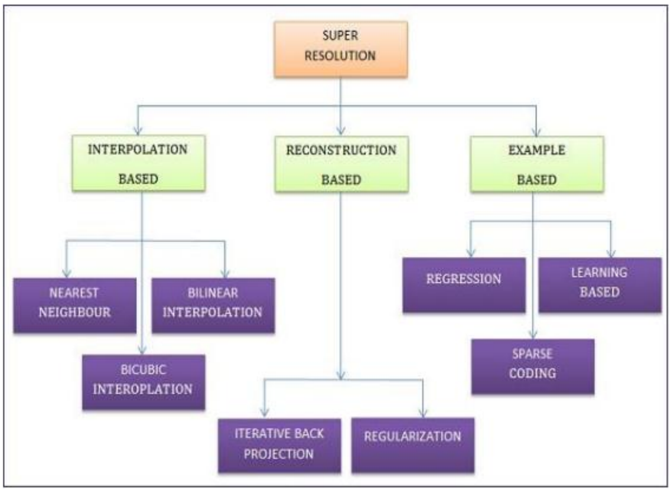
Literature Review:

The super resolution image enhancement is an ill-posed inverse problem it means that solution to the problem is not unique. The problem of super resolution can be expressed:

Y= DHX+N

Where X the unknown is image to be estimated, Y is the observed image, H represents the blurring operator, D represent the Downsampling matrix and N is additive noise. [1] The linear system expresses the relation between the LR images and HR images. Super resolution methods are able to solve the inverse problem and estimate the HR images. [3] Size of the image, H and D are unknown in real application and it is necessary to estimate from the available LR images. The linear system is ill-posed. So, proper prior regularization is required

Methodology:



1. Interpolation-Based Methods:

These methods are the simplest and oldest techniques for image super-resolution. They work by estimating pixel values in a high-resolution image based on neighboring pixels.

Common Techniques:

- **Nearest Neighbor Interpolation**
- **Bilinear Interpolation**
- **Bicubic Interpolation** [17]

Advantages:

- **Low computational cost**: These methods are computationally efficient and fast to implement, making them suitable for real-time applications.
- **Simple to use**: They are straightforward and do not require complex models or training data.
- **No training data needed**: Unlike learning-based methods, interpolation methods do not rely on large datasets or model training.

Disadvantages:

- **Blurring effects**: These methods often produce blurred images, especially when the upscaling factor is high. They cannot recover fine details.
- **Limited quality improvement**: The resolution enhancement is often modest, and it's difficult to reconstruct high-frequency information like textures or sharp edges.
- **Inaccurate for large scaling factors**: As the scaling factor increases, the quality degrades significantly.

2. Example-Based Methods:

These methods utilize machine learning or deep learning techniques to learn the relationship between low-resolution and high-resolution image patches by training on a dataset of example images.

Common Techniques:

- **Sparse coding-based super-resolution**
- **Dictionary learning-based methods**
- **Deep learning-based models (CNNs, GANs) [12]**

Advantages:

- **Superior image quality**: They can recover high-frequency details, such as textures and edges, leading to sharper and more realistic images.
- **Learning from data**: The models are trained on large datasets, allowing them to generalize well and produce high-quality images even for complex textures and structures.
- **Flexible and adaptable**: These methods can adapt to different types of images and noise levels by training on specific datasets.

Disadvantages:

- **High computational cost**: Training the models is computationally expensive, and running them on large images may also require significant resources.
- **Requires large datasets**: The quality of the output heavily depends on the size and diversity of the training data. Poor or limited datasets can lead to suboptimal results.
- **Risk of overfitting**: If the model is trained on a small or biased dataset, it may not generalize well to new images.
- **Time-consuming training process**: Building the model and training it can be time-intensive, especially for deep learning models.

3. Reconstruction-Based Methods:

These methods use prior knowledge (such as smoothness, edge information, or sparsity) to iteratively reconstruct the high-resolution image by enforcing constraints that the reconstructed image should satisfy. [7]

Common Techniques:

- **Total variation regularization**
- **Maximum a posteriori (MAP) estimation**
- **Regularized deconvolution techniques**
- **Iterative back-projection (IBP)**

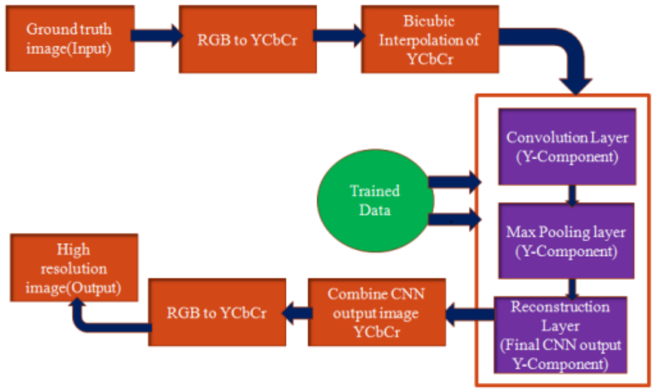
Advantages:

- **Improved accuracy**: These methods are good at maintaining image fidelity and preserving fine details like sharp edges and textures.
- **Incorporates prior knowledge**: By using priors, these methods can enforce constraints that produce more natural-looking images and reduce artifacts.
- **Versatile**: These methods can be tailored for various types of images and noise models by adjusting the priors or constraints.

Disadvantages:

- **Slow convergence**: Reconstruction-based methods are often iterative and can take a long time to converge, especially for large images or high scaling factors.
- **Sensitive to priors**: The quality of the output depends on the selection of the prior knowledge. Inappropriate priors can lead to artifacts or unnatural images.
- **Computationally expensive**: Although more accurate, these methods can be computationally intensive, especially when combined with complex priors or constraints.
- **Difficulty in handling large scaling factors**: Like interpolation methods, they can struggle to handle large scaling factors without introducing artifacts or inaccuracies.

In the propose work approach to the convolution neural network CNN will use for the faster computation output as well as the comparative better resolution of the image.



Neural networks can be applied to computer vision tasks and to achieve good generalization performance it is advantageous to embed prior knowledge into the network architecture the convolutional neural network CNN is a specialized form of a multilayer perceptron designed specifically for processing and identifying two-dimensional image information a convolutional neural network CNN consists of one or more convolutional layers often accompanied by a subsampling layer followed by one or more fully connected layers similar to a standard neural network in CNNs the convolutional layers serve as feature extractors but unlike traditional models these layers are not hand-designed instead the weights of the convolutional filter kernels are learned during the training process convolutional layers are effective at extracting local features by limiting the receptive fields of the hidden layers to be local CNNs are applied across various fields including image and pattern recognition speech recognition natural language processing and video analysis CNNs have gained prominence for several reasons in traditional pattern recognition models feature extractors were manually designed however in CNNs both the weights of the convolutional layers used for feature extraction and the fully connected layers used for classification are learned during training the improved architecture of CNNs reduces memory and computational complexity while delivering superior performance in tasks where the input has local correlations CNNs leverage the spatial relationships between pixels in an image making them fast to train which facilitates the development of deep multi-layer networks today deep convolutional networks or their variants are used in most image recognition and computer vision tasks CNNs are highly efficient algorithms widely used in pattern recognition and image processing a typical CNN architecture includes multiple layers with the convolutional and max-pooling layers being the most common in the proposed work the architecture consists of three layers convolutional layer max-pooling layer and reconstruction layer.

ALGORITHM:

Input: Ground truth image and trained data of CNN

Output: High resolution RGB image

Steps of Algorithm:

1. Read the RGB input image.
2. Convert the image to YCbCr color space.
3. Apply bicubic interpolation to downscale and upscale the Y-component.
4. Process the Y-component through the CNN.
5. Merge the enhanced Y-component with the original Cb and Cr components.
6. Convert the final image back to the RGB color space.

Results:

PEAK SIGNAL TO NOISE RATIO (PSNR)

Peak Signal-to-Noise Ratio (PSNR) is a metric commonly used to evaluate the quality of a reconstructed or processed image. It quantifies the difference between the original image and the processed image in terms of signal and noise. Higher PSNR values indicate better image quality.

PSNR is calculated using the Mean Squared Error (MSE) between the original and processed images. The formula for PSNR is:

$$\mathrm{PSNR} = 10 \times \log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right)$$

Where,

MAX is the maximum possible pixel value of the image (often 255 for 8-bit images).

MSE is the Mean Squared Error between the original and processed images.

IMAGE	PSNR	RMSE	TIME	SSIM
BABOON	23.5219	16.9998	21.3594	0.7083
BUTTERFLY	32.7525	5.8738	19.3281	0.9641
PEPPER	36.8743	3.6545	56.375	0.9192
HEAD	35.7228	4.1725	23.9844	0.8861
FLOWERS	27.8984	10.2712	16.2186	0.87953
ZEBRA	33.494	5.3931	49.6094	0.9423
BIRD	40.9149	2.2951	23.5313	0.9858
MONARCH	27.7737	10.4197	18.8125	0.9376

Here in the table shown PSNR, RMSE, SSIM and execution of proposed method. In the
table we can see that the PSNR, RMSE, SSIM for each image is increase compared to the
Existing TV method as well as the execution time also reduced.

Discussion:

Image	PSNR		RMSE		TIME		SSIM	
	Existing Work	Proposed Work	Existing Work	Proposed Work	Existing Work	Proposed Work	Existing Work	Proposed Work
BABOON	20.792	23.5219	6.0731	16.9998	89.2188	21.3594	0.7531	0.7083
BUTTERFLY	26.0212	32.7525	6.0419	5.8738	82.6563	19.3281	0.9536	0.9641
PEPPER	30.3123	36.8743	6.0223	3.6545	920.156	56.375	0.9813	0.9192
HEAD	31.2259	35.7228	6.0182	4.1725	121.875	23.9844	0.9177	0.8861
FLOWERS	22.9773	27.8984	6.0558	10.2712	49.6875	16.2186	0.8584	0.8795
ZEBRA	21.9719	33.494	6.0665	5.3931	67.5	49.6094	0.8940	0.9423
BIRD	34.1861	40.9149	6.0146	2.2951	110.625	23.5313	0.9878	0.9858
MONARCH	23.8016	27.7737	6.0541	10.4197	80.1563	18.8125	0.9499	0.9376

The performance was evaluated using PSNR, RMSE, and SSIM metrics. The proposed CNN method showed higher PSNR and better quality compared to existing TV-PSSR techniques while reducing processing time

Conclusion:

The CNN-based approach successfully improves image resolution while maintaining low computational costs. The architecture effectively extracts image features and provides a superior alternative to traditional methods.

Acknowledgement:

I would like to express my gratitude to the organization and supporters for their invaluable support in completing this survey paper. First of all, I would like to express my sincere gratitude to Mr. Jigar Dalvadi for his support and guidance throughout the explaining about survey paper. I am also grateful to Niteen Vaghela and Nayana Suresh for their assistance with data collection, analysis, and interpretation, which made the study possible and contributed to enhancing the quality of this survey paper. Finally, I thank to Sardar Patel College of Engineering in collaboration with BRDC cell for providing the necessary facilities and resources to conduct my experiments and survey paper. I extend my gratitude to all for your invaluable contributions that significantly enhanced the quality of the paper.

References

[1]. Zhou, Ying, Zheng, Z., & Sun, Q. (2024). Bi-directional projection framework for fast Single Image Super Resolution. Recent Patents on Engineering, 18(9). doi:10.2174/0118722121248802231004053522

[2]. Zhou, Yichao, Tang, Z., & Hu, X. (2014). Fast Single Image Super Resolution Reconstruction via Image Separation. Journal of Networks, 9(7). doi:10.4304/jnw.9.7.1811-1818

[3]. Nath, A. G., Nair, M. S., & Rajan, J. (2015). Single image super resolution from compressive samples using two level sparsity based reconstruction. Procedia Computer Science, 46, 1643–1652. doi:10.1016/j.procs.2015.02.100

[4]. Purkait, P., & Chanda, B. (2012). Super resolution image reconstruction through Bregman iteration using morphologic regularization. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 21(9), 4029–4039. doi:10.1109/TIP.2012.2201492

[5]. Freeman, W. T., Thouis, R., & Jones, E. C. (2002). Example-based superresolution. Computer Graphics and Applications, 22, 56–65.
[6]. Giaconi, J. C., Allen, C. R., & Steinbach, L. S. (2009). Anterior cruciate ligament graft reconstruction: clinical, technical, and imaging overview. Topics in Magnetic Resonance Imaging: TMRI, 20(3), 129–150. doi:10.1097/RMR.0b013e3181d657a7
[7]. Trimeche, M. (n.d.). Super-Resolution Image Reconstruction Using Non-Linear Filtering.
[8]. Nielsen, M. A. (2015). Neural Networks and Deep Learning". Determination Press.
[9]. Farrell, J., Xiao, F., & Kavusi, S. (n.d.). Resolution and light sensitivity trade off.
[10]. Goto, T. (2012). Learning-based super-resolution image reconstruction on multicore processor. IEEE Transactions On, 58(3), 941–946.
[11]. Xiaoqiang, Y., & Yuan, P. (n.d.). Alternatively constrained dictionary.
[12]. Liu, D., Wang, Z., & Huang, T. (n.d.). Efficient Image and Video SuperResolution Using Deep Convolutional Neural Networks.
[13]. Liu, C., Shum, H.-Y., & Freeman, W. T. (2007). Face hallucination: Theory and practice. International Journal of Computer Vision, 75(1), 115–134. doi:10.1007/s11263-006-0029-5
[14]. Naval Criminal Investigative Services (NCIS)" Ikena Forensic video enhancement</i>. (n.d.).
[15]. No title. (n.d.). Retrieved 19 September 2024, from https://www.motiondsp.com/ncis-surveillance-video/
[16]. (N.d.). Retrieved 19 September 2024, from http://america.pink/bilinear-interpolation_661297.html
[17]. G_, S. (2011, August 7). Zoom an image with different interpolation types. Retrieved 19 September 2024, from CodeProject website: http://www.codeproject.com/Articles/236394/Bi-Cubic-and-Bi-Linear-Interpolation
[18]. Bevilacqua, M., Roumy, A., Guillemot, C., & Morel, M. L. A. (n.d.). Low complexity singleimage super-resolution based on nonnegative neighbor embedding. British Machine.
[19]. Jin, M., Jin, J., & Hao, G. (2012). Optimization design for eye-in-hand robotics Based on A New Binocular Stereo Vision Servoing Model. Proceedings of the 2nd International Conference on Electronic and Mechanical Engineering and Information Technology (2012). Presented at the 2nd International Conference on Electronic and Mechanical Engineering and Information Technology, China. doi:10.2991/emeit.2012.118
[20]. Low-Complexity Single-Image Super-Resolution. (n.d.). Retrieved 19 September 2024, from http://people.rennes.inria.fr/Aline.Roumy/results/SR_BMVC12.html
[21]. Research. (n.d.). Retrieved 19 September 2024, from https://sites.google.com/site/romanzeyde/research-interests