iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.51850

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Super-Resolution of Images using Contrastive Self-Learning

Gourav Meher[1], Ashutosh Mandhani[2], Dr. P. Akilandeshwari[3]
Department of Computing Technologies, SRMIST Kattankulathur

Abstract: ***Super-resolution (SR) techniques aim to enhance the resolution and quality of low-resolution images. In recent years, deep learning-based approaches have achieved remarkable success in this field. However, training deep SR models often requires large-scale datasets and computationally expensive operations, limiting their practicality. Different approaches have been proposed to tackle computation and time constraints. Our novel approach involves a combination of knowledge distillation and contrastive loss function to train a compact and efficient contrastive self-distillation (CSD) model. In the framework, a teacher network is initially trained on a large dataset using a traditional supervised learning approach. The teacher network learns to generate high-resolution images from low-resolution inputs. Subsequently, a student network is initialized with the same architecture as the teacher network but with fewer parameters. The student network is then trained to mimic the behavior of the teacher network by utilizing a contrastive loss. The contrastive loss is formulated by constructing positive and negative pairs of low-resolution and high-resolution image patches. The proposed CSD method achieves competitive performance compared to state-of-the-art SR methods while requiring significantly fewer parameters and computational resources. Furthermore, it demonstrates improved generalization capability by effectively reconstructing details in real-world images.***
Keywords: Images, Super-resolution, Knowledge Distillation, Contrastive Loss

I. INTRODUCTION

The method of super resolution involves generating high-resolution images from low-resolution ones. Computer vision and image processing remain integral areas of research for the past few decades. Early super resolution techniques relied on pixel duplication through interpolation to enlarge images and form their foundation. These methods hampered their proficiency in generating high-quality images. Advanced machine learning techniques have enabled researchers to generate high-quality images through more sophisticated super resolution algorithms developed over the years. Achieving super resolution often involves utilizing neural networks, which are popularly employed. The name given to machine learning algorithms that are inspired by the structure and function of the human brain is neural networks. The capability to learn intricate patterns and relationships in data is possessed by artificial intelligence models. Additionally, they can be utilized for a vast array of duties, including image manipulation and visual perception. The usage of neural networks within super resolution algorithms has brought about substantial improvement in their performance over time. There are various methods for super resolution, including interpolation-based methods, reconstruction-based methods, and learning-based methods. Interpolation-based methods involve simple interpolation of low-resolution images using techniques such as bicubic interpolation. Reconstruction-based methods involve the use of inverse problem techniques, such as total variation regularization and non-local means, to reconstruct high-resolution images. Learning-based methods use machine learning algorithms, such as deep neural networks, to learn the mapping between low-resolution and high-resolution images and generate high-quality output. Super resolution has become an active research area, and numerous methods and algorithms have been proposed in recent years. The development of deep learning techniques, such as Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs), has led to significant improvements in super resolution performance and has made super resolution an important and widely-used technology in various fields.

II. LITERATURE SURVEY

Super resolution of images is a challenging problem in computer vision and image processing and has been an active area of research for several decades. The goal of super resolution is to generate high-resolution images from low-resolution input images. In recent years, researchers have made significant progress in developing super resolution algorithms, particularly those based on deep learning techniques.

One of the earliest works in super resolution was proposed by Irani and Peleg in 1991. They proposed a technique called super-

©IJRASET: All Rights are Reserved

resolution from a single image, which involved the use of a linear spatially variant filter to generate a high-resolution image from a low-resolution image. The technique was based on the assumption that the low-resolution image is a blurred version of the high-resolution image. The filter was designed to estimate the high-resolution image from the low-resolution image, by exploiting the spatially varying nature of the blur.

Another early work in super resolution was proposed by Borman and Stevenson in 1998. They proposed a technique called super resolution using sparse representations, which involved the use of sparse coding to generate a high-resolution image from a set of low-resolution images. The technique assumed that the low-resolution images have a common high-resolution representation. The technique involved learning a dictionary of basic functions from the low-resolution images and using sparse coding to generate the high-resolution image. In 2004, Freeman et al. proposed a neural network-based approach to super resolution. They used a multi-layer perceptron (MLP) to generate high-resolution images from low-resolution images. The MLP was trained to learn the mapping between low-resolution and high-resolution images, by minimizing the difference between the predicted and actual high-resolution images. Later in 2006, Glasner et al. proposed a super resolution technique based on the use of a Markov random field (MRF) model. The technique involved the use of an MRF to estimate the high-resolution image from the low-resolution image. The MRF was designed to capture the spatial dependencies between neighboring pixels in the image. The early works in super resolution involved the use of various techniques such as interpolation, sparse coding, wavelet transforms, neural networks, and Markov random fields. These techniques laid the foundation for the development of more sophisticated super resolution algorithms based on deep learning techniques.

In 2014, Dong et al. proposed a deep convolutional neural network (CNN) based approach to super resolution called SRCNN. The SRCNN consisted of three layers, with each layer consisting of multiple convolutional filters. The network was trained to learn the mapping between low-resolution and high-resolution images using a mean squared error (MSE) loss function. The SRCNN achieved state-of-the-art performance on several benchmark datasets. Ledig et al. Later in 2017, proposed a generative adversarial network (GAN) based approach to super resolution called SRGAN. The SRGAN consisted of a generator network and a discriminator network. The generator network was trained to generate high-resolution images from low-resolution images, while the discriminator network was trained to distinguish between the generated high-resolution images and the real high-resolution images. The SRGAN achieved state-of-the-art performance on several benchmark datasets. In the years of 2018-2020 Zhang et al. proposed a network called EDSR, which stands for Enhanced Deep Super-Resolution. The EDSR consisted of 32 layers, with each layer consisting of multiple convolutional filters. The network was trained using a MSE loss function and achieved state-of-the-art performance on several benchmark datasets. Then, Haris et al. proposed a network called ESRGAN, which stands for Enhanced Super-Resolution Generative Adversarial Networks. The ESRGAN improved upon the SRGAN by introducing a residual-in-residual (RIR) block and a perceptual loss function. The network achieved state-of-the-art performance on several benchmark datasets.

III. IMPLEMENTATION

For the model training, we have utilised DIV2K dataset. The DIV2K dataset is a popular dataset for image super-resolution tasks, specifically aimed at the development and evaluation of deep learning models for this task. The DIV2K dataset is split into two parts: a training set and a validation set. The training set contains 800 images, while the validation set contains 100 images. The remaining 100 images are reserved for testing and are not publicly available, but researchers can submit their results to an online leaderboard for evaluation.

The major components involved in the model are as follows:

A. Channel-Splitting Super-Resolution Network

The basic idea behind CSSR-Net is to divide the input image into multiple channels, and then process each channel separately using a separate neural network. This allows the network to learn more accurate and complex representations of the image, which can then be combined to produce a high-resolution output. The student and teacher separately produce a high-resolution image constrained by the reconstruction loss. The knowledge is constructed by contrastive loss and explicitly transferred from teacher to student, where the student-channel's output is pulled to closer to the teacher-channel's output and pushed far away from the negative samples in the embedding space. The network architecture is shown in fig.1.

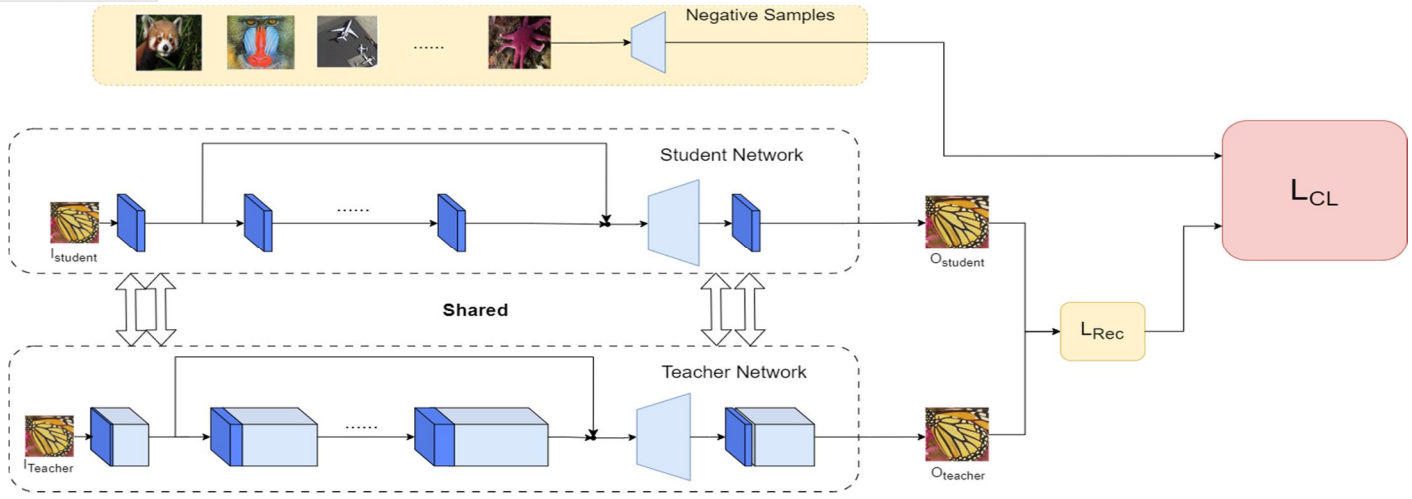


Fig.1: CSSR-Net

B. Contrastive Loss Function

The contrastive loss function is a popular loss function used in super-resolution algorithms for learning similarity between high-resolution and low-resolution image patches. The contrastive loss function aims to minimize the distance between the feature representations of two similar patches and maximize the distance between the feature representations of two dissimilar patches. In CSSRNet, the loss function incorporates contrasting influences that simultaneously attract the output of the CSSR network towards the output of its teacher network while pushing output of student network towards negative samples. This creates opposing forces that guide the training process.

C. Overall Loss

The complete loss function in our Contrastive Self-Distillation (CSD) approach is formed by combining the reconstruction loss and the contrastive loss. This can be expressed as : L = λ1 * Lrec + λ2 * Lcon, where Lrec represents reconstruction loss and Lcon represents contrastive loss. Here λ is a hyperparameter necessary to balance the two loss functions.

IV. RESULTS

We chose EDSR+ model for the backbone of the CSSR-Network and and compressed it, thereby creating a smaller student model with only 2.7M parameters using a compression ratio of 16.0x in comparison to the original model. The proposed CSD model achieved high PSNR and SSIM values. The EDSR+ 0.25× model achieved a compression ratio of 16x with only a 0.2dB loss in PSNR at the SR scale of 4 on DIV2K. The model showed an average PSNR value of 34.46 and SSIM value of 0.9453 on DIV2K dataset and on other public datasets like Urban100 and BSD100 it showed PSNR value of 32.55 and 32.23 and SSIM value of 0.9322 and 0.9011 respectively.

An examination of the effect of number of negative samples used in the training process showed that it had a linear relationship with performance. However it was also seen that it led to an increase in time for training as well as memory cost. For the tradeoff between performance and efficiency, we chose 10 negative samples for all our tests.

V. CONCLUSION & LIMITATIONS

We comprehensively evaluated the performance of the CSD scheme on various SR models using standard benchmark datasets and demonstrate superior performance compared to existing methods. We found that this approach could be extended to other low-level vision tasks like dehazing, denoising, and deblurring. The proposed method has several advantages, including significantly reducing the size and complexity of the SR models, resulting in faster inference times and requiring less computational resources. Additionally, the CSD scheme can achieve state-of-the-art performance while maintaining high-quality output images, making it an effective solution for real-world applications. The approach used in this project is essentially for individual small-scale firms to integrate our methodologies with their setup for improving their results and further research on this could lead to its integration in various other applications such as real-time video streaming process, reconstruction of historical records, etc.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

VI. FUTURE SCOPE

Future work has possibility to explore the generalizability of the CSD scheme to other low-level vision tasks beyond SR, as well as investigating the impact of different hyperparameters on the model's performance. Overall, the CSD scheme represents a promising approach to compressing and accelerating SR models while maintaining high-quality output images.

REFERENCES

[1] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee, “Enhanced Deep Residual Networks for Single Image Super-Resolution,” ,arXiv:1707.02921, Cornell University, 2017

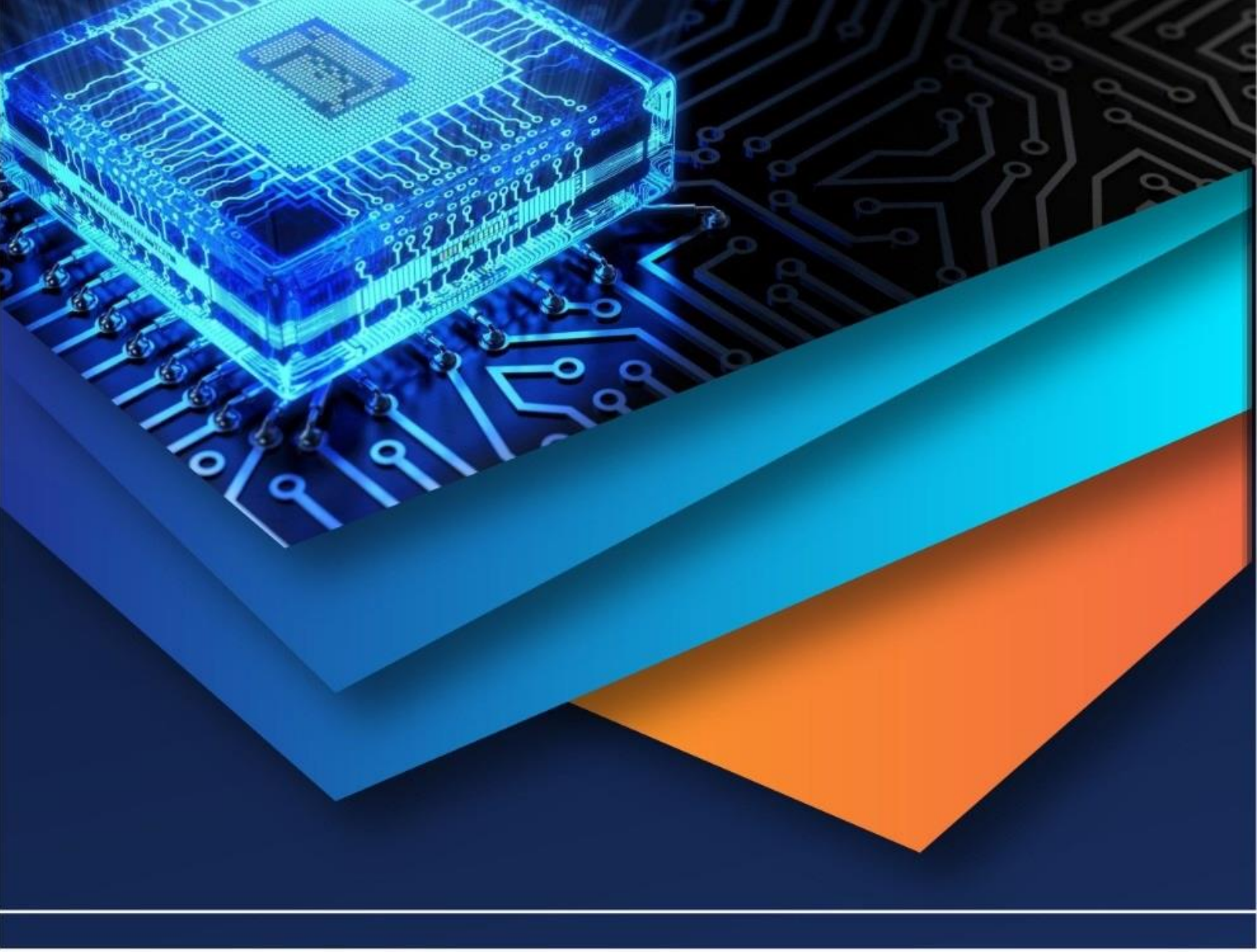
[2] Namhyuk Ahn, Byungkon Kang, Kyung-Ah Sohn, “Fast, Accurate, and Lightweight Super-Resolution with Cascading Residual Network”, European Conference on Computer Vision (ECCV),

[3] Geoffrey Hinton, Oriol Vinyals, Jeff Dean, “Distilling the Knowledge in a Neural Network” arXiv:1503.02531, 2015.

[4] Feng Wang, Huaping Liu, ”Understanding the Behaviour of Contrastive Loss”; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 2495-2504

[5] Y. Jo, S. Yang and S. Kim, "Investigating Loss Functions for Extreme Super-Resolution," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA,

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)