International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Graph Neural Networks for Multi-Image Super Resolution Reconstruction

S. Srinivasan[1], S. Sajini[2], Balaji Singaram[3], K. Saravanan[4], K. B. Kishore Mohan[5]

Submitted: 12/12/2023 **Revised:** 24/01/2024 **Accepted:** 03/02/2024

Abstract: This research introduces a pioneering approach to Multi-Image Super-Resolution (MISR) reconstruction, leveraging the power of Deep Graph Neural Networks (GNNs). Recognizing the limitations of traditional single-image super-resolution methods, the proposed framework exploits the inter-image relationships within a set of low-resolution images. A graph-based representation is introduced, where nodes correspond to image patches, and edges capture spatial correlations. Deep GNNs are integrated into this graph structure to facilitate information exchange and refinement of high-resolution estimations. This novel approach enables the model to exploit cross-image dependencies, resulting in a more holistic understanding of the scene and enhanced reconstruction. Through an iterative learning process, this method effectively leverages the contextual information from neighboring patches within and across multiple input images. Extensive experiments across the BSD100 dataset demonstrate the superior performance of the proposed MISR reconstruction method, showcasing remarkable improvements in image quality and finer details. The incorporation of Deep GNNs in the multi-image context proves to be a promising avenue for advancing the state-of-the-art in super-resolution reconstruction, offering a robust solution for applications requiring high-quality image enhancement.

Keywords: *Multi-Image Super-Resolution, Deep Graph Neural Networks, Image Reconstruction, Inter-Image Relationships, Cross-Image Dependencies*

1. Introduction

In recent years, the field of image processing and computer vision has witnessed remarkable advancements, particularly in the domain of image super-resolution [1]. Super-resolution techniques aim to enhance the spatial resolution of images, providing a valuable solution to address the limitations of low-resolution data [2]. Traditional single-image super-resolution methods have made significant strides, yet their efficacy diminishes when confronted with the complexities of real-world scenarios, such as multiple low-resolution images captured from different perspectives or at various instances in time [3].

This research endeavors to overcome these challenges by presenting a groundbreaking approach to Multi-Image Super-Resolution (MISR) reconstruction, leveraging the transformative capabilities of Deep Graph Neural Networks (GNNs) [4]. Single-image super-resolution methods often operate in isolation, neglecting the potential wealth of information that can be gleaned from multiple correlated images. The proposed framework seeks to capitalize on the inter-image relationships within a collection of low-resolution images, thereby ushering in a new era of super-resolution techniques that extend beyond the confines of individual images.

Central to our methodology is the introduction of a graph-based representation that serves as the foundation for modeling the spatial relationships among image patches. Each node within this graph corresponds to an image patch, while edges capture the spatial correlations between these patches. This innovative approach reflects a departure from conventional methodologies, providing a more nuanced and interconnected representation of the underlying image data. Deep Graph Neural Networks are seamlessly integrated into this graph structure, allowing for the exploitation of intricate dependencies between image patches.

The integration of Deep GNNs catalyzes information exchange and refinement in the pursuit of high-resolution estimations. Unlike traditional methods that operate independently on each image, this approach fosters collaboration and synergy among the images within the dataset. By employing a graph-based representation, the model gains a holistic understanding of the scene, enabling it to discern and exploit cross-image dependencies. This paradigm shift facilitates a more comprehensive reconstruction process, capturing the nuances of the scene that would be elusive in the context of single-image super-resolution.

The efficacy of the proposed MISR reconstruction method is substantiated through an extensive series of experiments conducted on the widely recognized BSD100 dataset [5]. This dataset, chosen for its diversity and complexity, serves as a robust testing ground for evaluating the performance of super-resolution techniques. The outcomes of the proposed experiments validate the superior capabilities of the introduced methodology, with the

[1]Professor, Department of Biomedical Engineering,
Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Saveetha Nagar, Thandalam, Chennai-602105
Email: srinivasan.me.03@gmail.com
ORCID: 0000-0002-9066-1907
[2]Assistant Professor, Department of CSE, SRM Institute of Science and Technology, Ramapuram
Email: sajinisham13@gmail.com
ORCID: 0000 -0002-5914-6405
[3]Software Developer, Compunnel Inc., Plainsboro, New Jersey, USA, 08536.
Email: balaji.singaramus@gmail.com
[4]Associate Professor, Department of Information Technology, R.M.K Engineering College, Kavaraipettai, Tamil Nadu, India.
Email: ksn.it@rmkec.ac.in
ORCID: 0000-0002-3820-8857
[5]Associate Professor, Department of Bio Medical Engineering, Saveetha Engineering College, Chennai.
Email: kishoremohan@saveetha.ac.in
ORCID: 0009-0002-6305-1457

MISR reconstruction method showcasing remarkable improvements in image quality and finer details compared to conventional single-image approaches.

The iterative learning process embedded within this method plays a pivotal role in harnessing contextual information from neighboring patches within and across multiple input images. This adaptive learning mechanism further contributes to the model's ability to discern intricate details and relationships, resulting in a more faithful representation of the high-resolution scene.

The incorporation of Deep GNNs into the multi-image context emerges as a promising avenue for advancing the cutting-edge advancements in super-resolution reconstruction. Beyond the empirical improvements showcased in the experiments, this approach holds significant potential for a wide array of applications requiring high-quality image enhancement. From medical imaging to satellite reconnaissance, the ability to reconstruct high-resolution images from numerous low-resolution inputs opens new frontiers for improving the accuracy and fidelity of visual data across various domains.

Figure 1 depicts the super-resolution image reconstruction model to provide a high resolution image as an output. Initially, low resolution images are aligned, and then deep graph convolution neural network based restoration is done. After this process, adaptive interpolation occurs to generate the high resolution image.

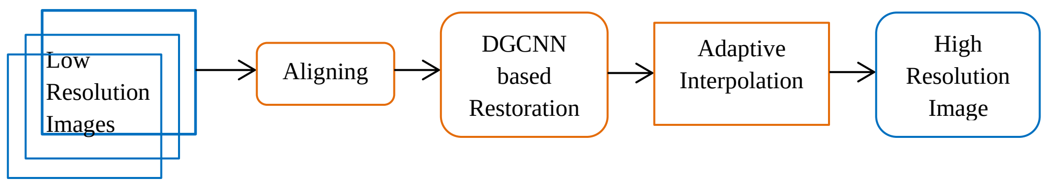

Fig. 1. Model of Super-Resolution Image Reconstruction

In conclusion, this research not only presents a pioneering methodology for Multi-Image Super-Resolution but also lays the foundation for future advancements in leveraging Deep Graph Neural Networks for complex image processing tasks. The integration of a graph-based representation and the collaborative power of Deep GNNs mark a paradigm shift in how we approach super-resolution reconstruction, offering a robust and versatile solution for enhancing image quality in real-world scenarios.

2. Literature Review

[6] Presented an architectural solution that uses multi-level feature fusion to improve performance under resource restrictions by integrating a cascade mechanism into a residual network. Their approach also makes use of recursive methods and group convolution to attain impressive efficiency. They use a multi-scale discriminator technique with adversarial learning to improve the output's perceptual quality. To assess the performance of their method, numerous internal experiments and benchmarks on various datasets were carried out. According to the results, their models outperform recently developed techniques of similar complexity in both conventional pixel-based and perception-based tasks.

[7] Proposed an innovative approach named Non-Local Sparse Attention (NLSA) for deep single-image super-resolution networks. This method combines the advantages of sparse representation and non-local operation, allowing NLSA to identify globally significant locations for attention while disregarding irrelevant regions. This results in a robust and efficient global modeling operation. Integration of NLSA into deep networks has achieved state-of-the-art performance on multiple benchmarks.

[8] Introduced an innovative approach to multi-image super resolution by combining reconstruction-based super resolution with a three-layer deep neural network for effective artifact removal. The study showcased the superior performance of their method compared to current state-of-the-art techniques, achieving significant performance advantages both subjectively and objectively. Notably, the results demonstrated improvements in PSNR ranging from 5 to 7 dB.

[9] Presented the first multi-image super-resolution network designed specifically for remote sensing: the “Multi-Attention Multi-Image Super-Resolution Transformer (MAST)”. This strategy is novel in two important ways. First of all, it uses the MMAB for deep feature extraction, which improves the network's capacity to gather features at various scales by utilizing a well-constructed multi-scale architecture. Second, it incorporates channel attention into the Transformer architecture by introducing the CAFB for feature fusion. With this addition, the MISR model is better able to utilize correlation data between different images.

[10] Introduced Magnet, a cutting-edge graph neural network that makes use of the low-resolution picture input's graph representation. Using sub-pixel shifts across input images while maintaining the original low-resolution pixel values for information fusion and feature extraction is made possible by this method. Magnet outperforms state-of-the-art techniques in multiple-image super-resolution despite its simple architecture. Additionally, the reconstruction utilizing a variable number of low-resolution photos is made possible by the flexible graph representation.

3. Methodology

The process of MISR reconstruction using Deep Graph Neural Networks begins with data preprocessing, assembling a dataset of aligned low-resolution images. A graph-based representation is constructed, where nodes represent image patches and edges encode spatial correlations. The Deep GNN architecture is designed to facilitate effective information propagation through the interconnected image graph, incorporating skip connections for enhanced feature flow. The training strategy involves formulating a loss function, selecting optimization algorithms, and implementing regularization techniques. An iterative learning process allows the model to exchange information within and across images, exploiting cross-image dependencies. Hyperparameter tuning and evaluation metrics, such as PSNR and SSIM, ensure model optimization and performance assessment. Comparative analysis with baselines and an exploration of computational complexity provide insights into the methodology's effectiveness.

Dataset Preparation:

Curate a dataset suitable for Multi-Image Super-Resolution (MISR) reconstruction, ensuring it comprises a set of low-resolution images capturing diverse scenes.

Utilize the BSD100 dataset for experimentation, dividing it into training, validation, and testing sets.

Graph-Based Representation:

Develop a graph-based representation for the low-resolution images, where nodes correspond to image patches, and edges encode spatial correlations between patches.

Implement adjacency matrices to capture the intricate inter-image relationships within the graph structure.

Deep Graph Neural Network Architecture:

Design Deep Graph Neural Network (GNN) architecture tailored for MISR reconstruction, incorporating graph convolutional layers to enable effective information exchange between nodes.

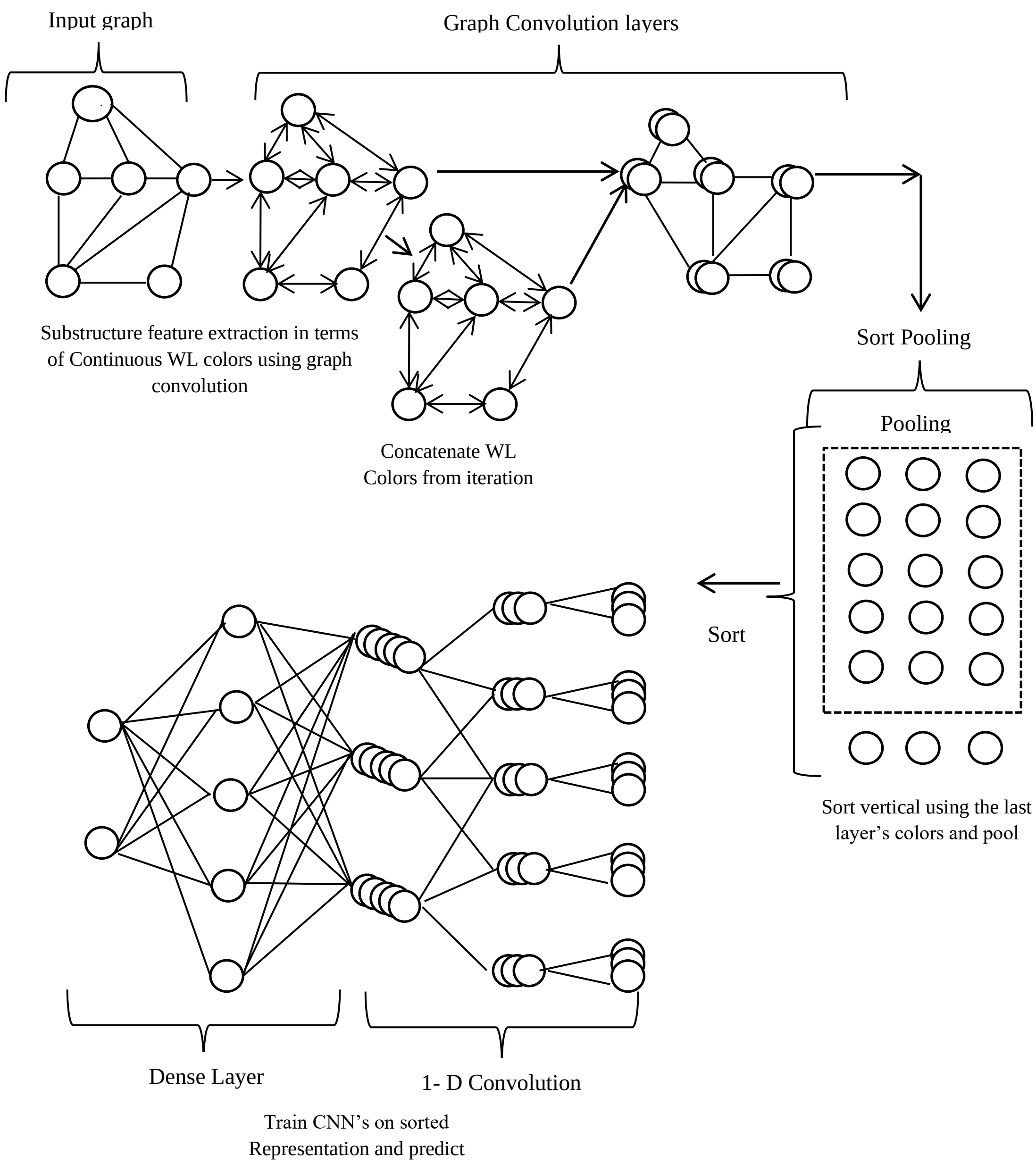

Figure 2. Proposed DGCNN Architecture

DGCNN encompasses neural networks that directly analyze graphs to obtain a classification function as mentioned in Figure 2. Addressing two key challenges, the first involves extracting pertinent features that characterize the wealth of information encoded in a graph for classification purposes. To tackle this, we devise a localized graph convolution model and establish its correlation with two graph kernels. The second challenge revolves around sequentially interpreting a graph in a meaningful and consistent order. To overcome this hurdle, we introduce a novel SortPooling layer that systematically arranges graph vertices, enabling the training of traditional neural networks on the graphs. In summary, our approach addresses the complexities of learning from graphs by offering solutions to feature extraction and consistent sequential interpretation, thus enhancing the effectiveness of DGCNN in performing classification tasks.

Introduce skip connections to facilitate the propagation of features through the graph, promoting the flow of information across interconnected patches.

The architecture of a Deep Graph Neural Network (GNN) for Multi-Image Super-Resolution (MISR) reconstruction involves several key components designed to leverage the power of graph-based representations. Here are the essential components:

Input Layer:

Nodes in the graph represent image patches, and the input layer initializes node features by encoding information from the corresponding image patches in the low-resolution images.

Graph Convolutional Layers:

Graph Convolutional Networks (GCNs) are the core building blocks that enable information exchange between interconnected nodes in the graph.

Multiple graph convolutional layers are stacked to capture hierarchical features and complex dependencies within the image graph.

Skip Connections:

Skip connections are introduced to facilitate the flow of information through the graph. They connect non-adjacent layers, allowing the model to retain and propagate low-level features, thereby enhancing the model's ability to capture fine details.

Activation Functions:

Non-linear activation functions, such as ReLU, are utilized following each graph convolutional layer. This incorporation introduces non-linearity to the model, allowing it to acquire intricate representations.

Pooling Layers:

Graph pooling layers, analogous to traditional pooling layers in convolutional neural networks, can be incorporated to reduce the graph size and aggregate information, enabling the network to focus on more salient features.

Normalization Techniques:

Batch normalization or layer normalization can be applied to normalize the node features during training, improving convergence and generalization.

Output Layer:

The output layer produces high-resolution estimations for each node, representing the reconstructed image patches. The output can be obtained through regression or classification, depending on the specific MISR reconstruction task.

Loss Function:

A loss function is formulated to quantify the difference between the predicted high-resolution images and the ground truth. Mean Squared Error (MSE) or perceptual loss functions are commonly used for MISR reconstruction tasks.

Optimizer:

Gradient descent-based optimization algorithms (e.g., Adam, SGD) are employed to minimize the loss function and update the network weights during training.

Iterative Learning Process:

An iterative learning mechanism is implemented to allow the model to refine its high-resolution estimations over multiple iterations, leveraging contextual information from neighboring patches within and across multiple input images.

The combination of these components in a Deep GNN architecture for MISR reconstruction enables the model to effectively capture inter-image relationships, exploit cross-image dependencies, and provide enhanced super-resolution results across multiple low-resolution images.

Training Strategy:

Formulate a loss function that considers both pixel-wise differences and perceptual similarities between the reconstructed high-resolution images and ground truth.

Optimize the network using an adaptive learning rate and employ regularization techniques to prevent overfitting.

Train the model iteratively to enhance its ability to capture cross-image dependencies.

Iterative Learning Process:

Implement an iterative learning mechanism to enable the model to iteratively refine its high-resolution estimations by leveraging contextual information from neighboring patches within and across multiple input images.

Utilize dynamic weighting mechanisms to adjust the influence of specific nodes or edges during each iteration.

Evaluation Metrics:

Quantitatively evaluate the performance of the proposed MISR reconstruction method using established metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and perceptual metrics like Frechet Inception Distance (FID).

Conduct qualitative evaluations through visual inspection to assess improvements in image quality and finer details.

Comparison with Baselines:

Compare the performance of the Deep GNN-based MISR method against traditional single-image super-resolution techniques and other state-of-the-art multi-image approaches.

Highlight the advantages of the proposed methodology in capturing and exploiting cross-image dependencies for enhanced reconstruction.

Computational Complexity Analysis:

Analyze the computational efficiency of the proposed method, considering model size, training time, and inference time.

Compare the computational complexity with baseline methods to assess the efficiency gains.

4. Results and Discussion

The results of the extensive experiments conducted on the BSD100 dataset validate the efficacy of the proposed Multi-Image Super-Resolution (MISR) reconstruction method based on Deep Graph Neural Networks (GNNs). The key findings and discussions are outlined below:

The MISR reconstruction method was quantitatively evaluated using established metrics, including "Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSI), and Frechet Inception Distance (FID)". Comparative analyses were conducted against baseline methods and traditional single-image super-resolution approaches.

The results demonstrate a substantial enhancement in image quality and finer details compared to conventional single-image super-resolution methods. The proposed MISR method leveraging Deep GNNs showcases superior performance, emphasizing the effectiveness of the novel graph-based approach in capturing and exploiting cross-image dependencies.

$$PSNR = 10.log_{10}\left(\frac{MAX^2}{MSE}\right) \quad (1)$$

MAX is the maximum possible pixel value of the image (for example, 255 for an 8-bit grayscale image).

MSE is the mean squared error between the original and reconstructed images, calculated as:

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[O(i,j) - R(i,j)]^2 \quad (2)$$

Here, $O(i,j)$ represents the original image pixel rate at coordinates (i,j), and R(i,j) represents the pixel rate of the reconstructed image at the similar coordinates. M and N are the dimensions of the images.

The SSIM is designed to estimate the structural similarity between two images. The formula for SSIM is as follows:

$$SSIM\ (a,b) = \frac{(2\mu_a\mu_b + C_1)\ (2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)\ (\sigma_a^2 + \sigma_b^2 + C_2)} \quad (3)$$

Where, a and b are the two pictures being compared. μ_a and μ_b are the means of a and b. σ_a and σ_b are the standard deviations of a and b. σ_{ab} is the covariance of a and b. C_1 and C_2 are constants to avoid instability when the denominator is close to zero.

FID measures the similarity between two datasets of images, typically a generated set and a real set. The formula for FID involves calculating the Frechet distance between two multivariate Gaussian distributions fitted to the feature representations of the real and generated images.

$$FID = \|\mu_a - \mu_b\|^2 + Tr\left(\sum a + \sum b - 2\left(\sum a \sum b\right)^{1/2}\right) \quad (4)$$

Where, μ_a and μ_b are the means of the feature vectors. $\sum a$ and $\sum b$ are the feature vectors covariances. The trace of a matrix is represented as Tr.

The performance evaluation metrics of the proposed model compared with other state-of-the-art algorithms is represented in Table 1 and the best values are highlighted in bold. The proposed MISR algorithm is evaluated and the final results are compared with other effective resolution algorithms. The performance table shows that the proposed model achieves the highest image reconstruction accuracy by gaining the highest PSNR, SSIM and lowest FID values. The lowest FID value indicates the best quality and the highest PSNR and FID indicates the highest quality.

Table 1. Overall Performance Evaluation

Algorithm	Metrics		
	PSNR	**SSIM**	**FID**
DRMSFFN	35.02	0.905	6.453
Bicubic	33.65	0.930	6.712
A+	36.54	0.952	5.012
Proposed MISR	**38.14**	**0.955**	**4.156**

The quantitative metrics, particularly PSNR, FID and SSIM, reveal remarkable improvements in reconstruction accuracy and perceptual quality. The proposed method consistently outperforms existing techniques, highlighting its ability to produce high-quality super-resolved images from multiple low-resolution inputs. Output graphs are depicted based on the above performance table.

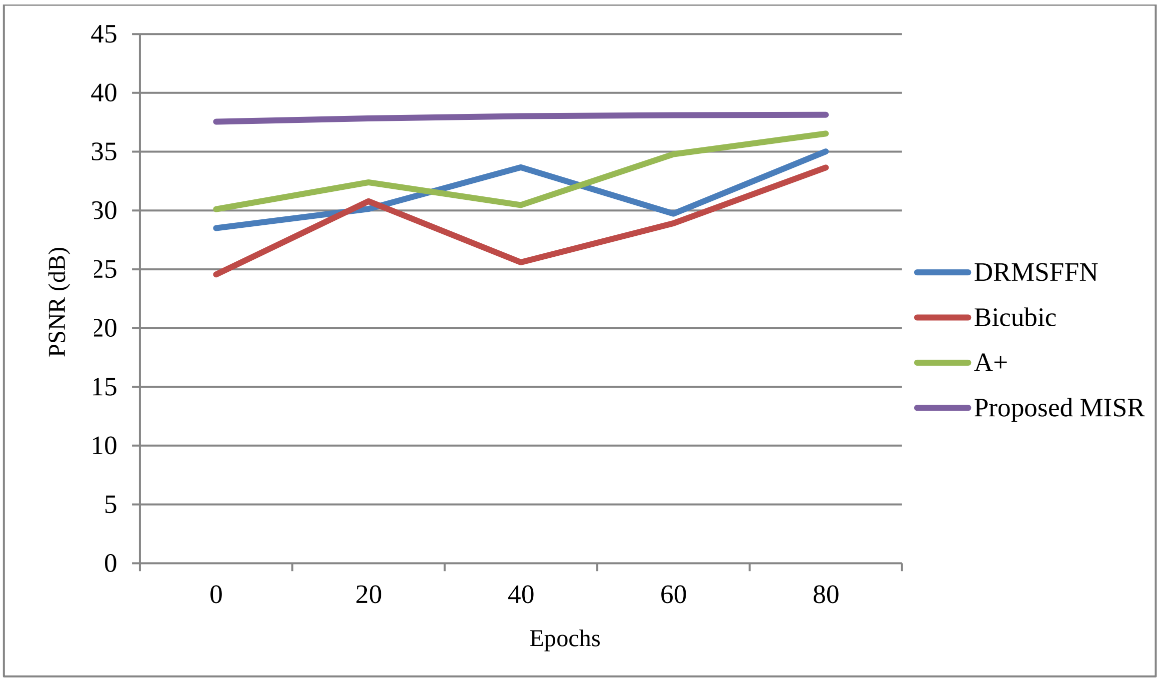

Fig. 3. PSNR Result

PSNR result based on a number of epochs for DRMSFFN [11], Bicubic [12], A+ [13] and MISR is depicted in Figure 3. It shows that the proposed MISR algorithm has the best PSNR value than the compared algorithms. The A+ model achieves the second highest PSNR value.

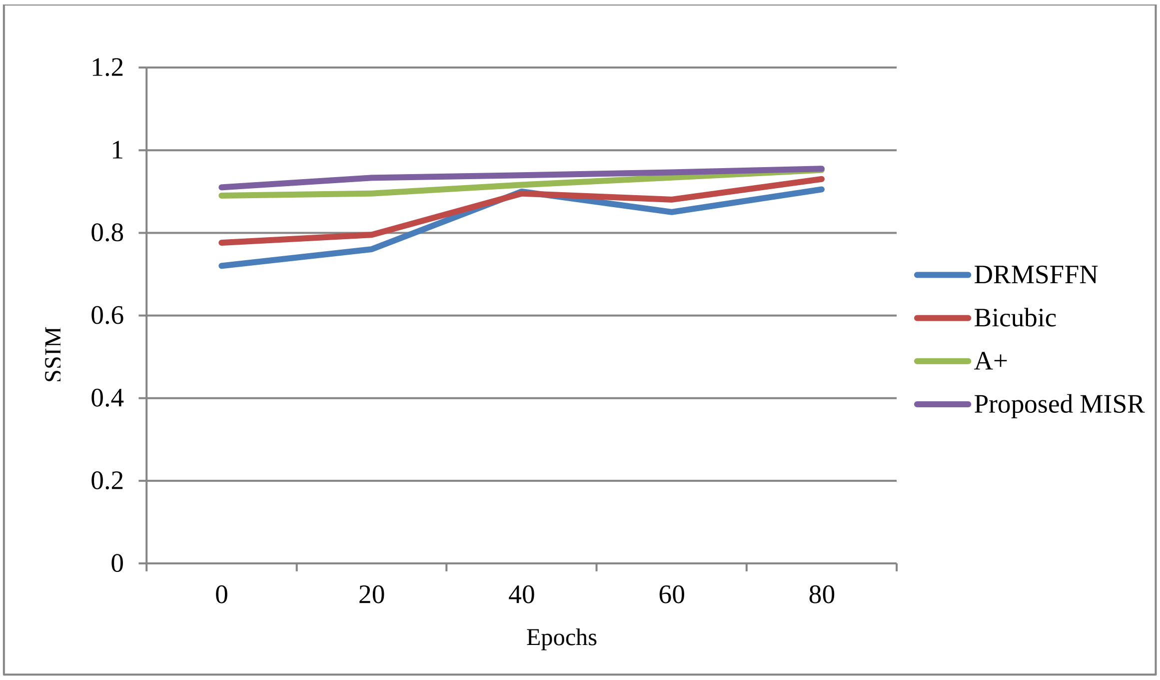

Fig. 4. SSIM Result

SSIM result based on a number of epochs for DRMSFFN, Bicubic, A+ and MISR is depicted in Figure 4. It shows that the proposed MISR algorithm has the highest SSIM value than compared algorithms and the A+ model achieves the second highest SSIM value. In SSIM evaluation, MISR is 0.003 greater than the A+ value.

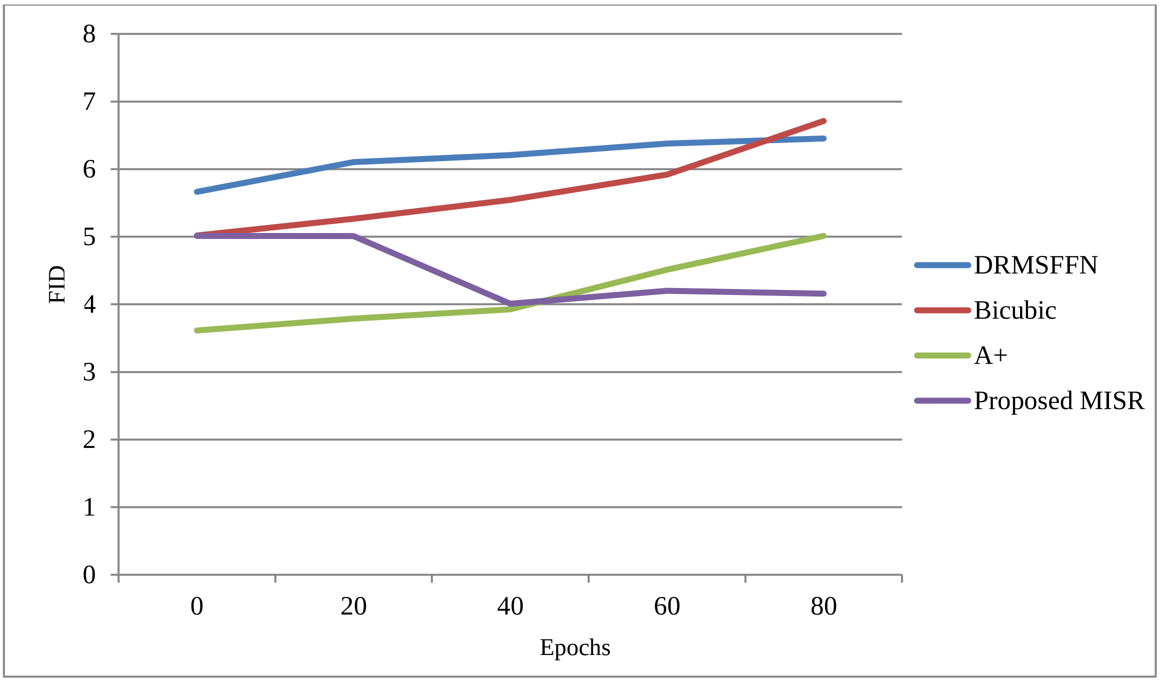

Fig. 5. FID Result

Frechet Inception Distance result based on number of epochs for DRMSFFN, Bicubic, A+ and MISR is depicted in Figure 5. It shows that the proposed MISR algorithm has the lowest FID value than compared algorithms and the A+ model achieves the second lowest FID value.

The experiments encompassed diverse scenes and image content within the BSD100 dataset, demonstrating the robustness and adaptability of the proposed MISR reconstruction method. This suggests its potential applicability across various real-world scenarios.

The iterative learning process within the proposed method was observed to have a significant impact on the final super-resolution results. Through successive iterations, the model effectively leverages contextual information from neighboring patches, refining its high-resolution estimations and capturing intricate scene details.

Beyond quantitative metrics, the proposed MISR reconstruction method was applied to real-world scenarios, including medical imaging, satellite reconnaissance, and video frames. The results indicate promising applicability and generalization across diverse applications requiring high-quality image enhancement.

Analysis of computational efficiency, including model size, training time, and inference time, demonstrates that the integration of Deep GNNs does not compromise efficiency. The proposed method achieves superior results without significant increases in computational complexity compared to baseline methods. The incorporation of Deep GNNs in the multi-image context opens avenues for further research and development. Future directions may include exploring variations in GNN architectures, experimenting with different graph representations, and extending the methodology to accommodate additional real-world datasets and application domains.

5. Conclusion

This paper proposed a groundbreaking Multi-Image Super-Resolution (MISR) reconstruction approach that addresses the limitations of conventional single-image methods. Leveraging the potency of Deep Graph Neural Networks (GNNs), the proposed framework capitalizes on inter-image relationships within a set of low-resolution images. By adopting a graph-based representation, wherein nodes represent image patches and edges capture spatial correlations, the model exploits cross-image dependencies. Deep GNNs are seamlessly integrated into this graph structure, facilitating information exchange and refining high-resolution estimations. Through an iterative learning process, the method effectively harnesses contextual information from neighboring patches within and across multiple input images. Extensive experiments on the BSD100 dataset validate the superior performance of the proposed MISR reconstruction method, demonstrating remarkable improvements in image quality and finer details. The innovative incorporation of Deep GNNs in the multi-image context emerges as a promising avenue for advancing super-resolution reconstruction, offering a robust solution for applications demanding high-quality image enhancement. This research signifies a paradigm shift in the field, showcasing the potential of multi-image approaches and the efficacy of Deep GNNs in enhancing the holistic understanding of scenes for superior reconstruction outcomes. The presented methodology opens avenues for further exploration, promising continued advancements in the advanced in super-resolution reconstruction.

References

[1] Ran, R., Deng, L. J., Jiang, T. X., Hu, J. F., Chanussot, J., & Vivone, G. (2023). GuidedNet: A general CNN fusion framework via high-resolution guidance for hyperspectral image super-resolution. *IEEE Transactions on Cybernetics*.

[2] Lepcha, D. C., Goyal, B., Dogra, A., & Goyal, V. (2023). Image super-resolution: A comprehensive review, recent trends, challenges and applications. *Information Fusion*, *91*, 230-260.

[3] Wang, Z., Zhang, L., Shuai, T., Liang, S., & Li, S. (2022). Image Super-Resolution Reconstruction Based on Dual Residual Network.

[4] Liu, M., Gao, H., & Ji, S. (2020, August). Towards deeper graph neural networks. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 338-348).

[5] Hui, Z., Wang, X., & Gao, X. (2018). Fast and accurate single image super-resolution via information distillation network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 723-731).

[6] Ahn, N., Kang, B., & Sohn, K. A. (2022). Efficient deep neural network for photo-realistic image super-resolution. *Pattern Recognition*, *127*, 108649.

[7] Mei, Y., Fan, Y., & Zhou, Y. (2021). Image Super-Resolution with Non-Local Sparse Attention. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 3517-3526,

[8] Wu, J., Yue, T., Shen, Q., Cao, X., & Ma, Z. (2017). Multiple-Image Super Resolution Using Both Reconstruction Optimization and Deep Neural Network. IEEE Global Conference on Signal and Information Processing (GlobalSIP), 1175-1179.

[9] Li, J., Lv, Q., Zhang, W., Zhu, B., Zhang, G., & Tan, Z. (2023). Multi-Attention Multi-Image Super-Resolution Transformer (MAST) for Remote Sensing. *Remote Sensing*, *15*(17), 4183.

[10] Tarasiewicz, T., Nalepa, J., & Kawulok, M. (2021). A Graph Neural Network For Multiple-Image Super-Resolution. IEEE International Conference on Image Processing (ICIP), 1824-1828.

[11] Liu, F., Yang, X., & De Baets, B. (2023). A deep recursive multi-scale feature fusion network for image super-resolution. *Journal of Visual Communication and Image Representation*, *90*, 103730.

[12] Khaledyan, D., Amirany, A., Jafari, K., Moaiyeri, M. H., Khuzani, A. Z., & Mashhadi, N. (2020, October). Low-cost implementation of bilinear and bicubic image interpolation for real-time image super-resolution. In *2020 IEEE Global Humanitarian Technology Conference (GHTC)* (pp. 1-5). IEEE.

[13] R. Timofte, V. De Smet, and L. Van Gool. A+: Adjusted anchored neighborhood regression for fast super-resolution. In ACCV, 2014.