https://doi.org/10.7840/kics.2024.49.5.762

Enhancing Chest X-ray Image Super-Resolution through Residual Group and Channel Attention

Anudari Khishigdelger◆, Ahmed Salem*, Hyun Soo Kang°

ABSTRACT

Medical imaging plays a crucial role in medical diagnosis, with chest X-ray imaging being a widely employed method for screening and diagnosing pulmonary diseases due to its cost-effectiveness. However, low-resolution images generated by expensive equipment and suboptimal imaging techniques often lead to a loss of critical features and acceptable texture. The acquisition of high-quality medical images is paramount for accurate disease diagnosis. This study introduces an innovative approach for reconstructing super-resolution medical images using deep learning techniques, explicitly targeting chest X-ray images. The proposed method, the Pathologically Invariant Remaining Enhanced Channel Attention Block (RECAB), incorporates the Exponential Linear Unit (ELU) activation function. The primary objective is to accurately recover high-resolution (HR) chest X-ray images from their low-resolution (LR) counterparts, leveraging a channel attention mechanism and convolution layer with the ELU activation function—the evaluation of the proposed method involved two datasets, X-Ray 2017 and X-Ray 2014. A comprehensive comparison was conducted with several state-of-the-art techniques, including GAN-based super-resolution, deep learning-based super-resolution, and interpolation-based super-resolution. The quality of the preprocessed images was assessed using the Structural Similarity Index (SSIM) and the Multi-Scale Structural Similarity Index (MSIM). The results prove the higher performance of the proposed method, which outperformed the average of 8.6% and 11.6% in SSIM and MSIM values by 11% and 12.14% on the two datasets, respectively. This research signifies a significant advancement in enhancing the resolution and quality of chest X-ray images, holding substantial potential for improving diagnostic accuracy and aiding in medical decision-making.

Key Words : Chest X-ray, Super-Resolution, Remaining Enhanced channel attention block, Exponential Liner Unit, Structure Similarity Index, Multi Structure Similarity Index.

Ⅰ. Introduction

Chest X-ray imaging is crucial in diagnosing various pulmonary diseases and is a significant part of global healthcare efforts[1]. Accurate and efficient analysis of chest X-ray images is essential for prompt and precise medical diagnosis. Over the years, researchers have explored various techniques to improve the accuracy and effectiveness of chest X-ray image analysis[1]. These techniques include

※ This research was supported by the Korea Institute of Marine Science & Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries (20200611), and supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2024-2020-0-01462,50%) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

◆ First Author : Department of Information and Communication Engineering, School of Electrical and Computer Engineering, Chungbuk National University, Khanudari1999@gmail.com, 학생회원

° Corresponding Author: Department of Information and Communication Engineering, School of Electrical and Computer Engineering, Chungbuk National University hskang@cbnu.ac.kr, 종신회원

* Electrical Engineering Department, Faculty of Engineering, Assiut University, Assiut 71515, Egypt, ahmeddiefy@chungbuk.ac.kr

논문번호 : 202310-093-0-SE, Received October 6, 2023; Revised December 19, 2023; Accepted February 16, 2024

deep learning methodologies, skip connections, and traditional image enhancement techniques. Deep learning, particularly convolutional neural networks (CNNs), has emerged as the leading approach in medical imaging analysis, drastically transforming chest X-ray diagnostics[2]. CNNs are particularly well-suited for this task, as they can autonomously learn features from raw pixel data and improve diagnostic capabilities[3-5]. Furthermore, prevalent architectures such as skip connections are instrumental in preserving high-resolution features throughout the encoding and decoding processes, leading to significant improvements in segmentation and reconstruction tasks[6].

In addition to CNN architectures, traditional image enhancement methods have been investigated to improve the interpretability and quality of chest X-ray images. Other state-of-the- art super-resolution approaches, such as SRGAN[7], SNSRGAN[8], RDN[9], and FSRCNN[10], VDSR[21], have also contributed significantly to this domain.

Our proposed method promised the following steps. Firstly, integrating the Residual Group (RG) and Remaining Channel Attention Blocks (RECAB) within the model is pivotal in enhancing image resolution. This enhancement aids in better diagnosis by providing finer details for accurate medical assessments[11]. Secondly, incorporating the Exponential Linear Unit (ELU) activation function within RECAB is instrumental in preserving subtle information, particularly in the darker regions of grayscale images. This prevents the loss of vital information during activation, ensuring that critical features are retained, which is particularly beneficial for

X-ray image analysis[12]. Thirdly, our model demonstrates efficient resource utilization, striking an optimal balance between model complexity and performance. This optimization allows for practical deployment in resource-constrained environments, broadening its potential applications. Lastly, leveraging Remaining Channel Attention Networks (RECAN) advancements, our model achieves state-of-the-art super-resolution results, positioning it as a highly competitive contender in the image enhancement domain. These advantages collectively underscore the potential and effectiveness of our proposed model in advancing medical image analysis and diagnosis.

In the subsequent sections, we will comprehensively describe our proposed model, outline the architecture in detail, and explain our experimental methodology and the evaluation results, including comparative analyses with other relevant models in the domain.

Ⅱ. Related work

Recently, deep learning based approaches to computer vision have dramatically outperformed traditional approaches. This section overviews various super-resolution methods, categorized into three main approaches: Interpolation-based super-resolution, GAN-based super-resolution, and deep learning-based super-resolution.

2.1 Interpolation Based Super Resolution

Interpolation-based super-resolution techniques wield predefined mathematical models to bolster image resolution in image enhancement. Particularly, nearest neighbor and bicubic interpolation are foundational interpolation approaches critical for upscaling low-resolution images.

The nearest neighbor method operates by straightforwardly duplicating the value of the nearest pixel, making it computationally efficient but often at the cost of introducing blocky artifacts and struggling to encapsulate intricate image features[19].

Conversely, bicubic interpolation takes a nuanced route, employing weighted averaging of neighboring pixels, resulting in smoother outcomes than its nearest neighbor counterpart. It risks over-smoothing fine details despite mitigating blackness, potentially overlooking vital image features[20].

2.2 GAN-Based Super Resolution

The advent of Generative Adversarial Networks (GANs) has revolutionized super-resolution, levera-

ging adversarial training to generate high-resolution images from their lower-resolution counterparts. Often, spectral normalization is incorporated to stabilize the training process and foster improved convergence[7].

SRGAN (Super-Resolution GAN)[7] is a pioneer in this realm, a groundbreaking work that significantly propelled GAN-based super-resolution. SRGAN adopts a GAN architecture enriched with a perceptual loss function, advocating for preserving high-level features. However, they present a challenge in training and stabilization, frequently manifesting issues like mode collapse and noticeable artifacts, ultimately influencing the overall quality of the resultant image[7].

The SNSRGAN architecture utilizes a perceptual loss function to generate high-resolution images from low-resolution inputs. The incorporation of spectral normalization helps stabilize the training process, enhancing convergence[8]. This network has been trained with a scaling factor of x4 and includes six residual blocks. These blocks span from the first residual block to the output, contributing to the discriminator's capabilities. The discriminator block is crucial for discriminating high resolution in super-resolution tasks, leading to improved accuracy. Each residual block comprises a convolutional layer, batch normalization (BN)[8], parametric rectified linear unit (PReLu) activation, and pixel shuffling for Up Sampling. In parallel, the discriminator blocks include spectrally normalized convolutional layers, batch normalizations, leaky ReLu activations, and fully connected layers. The total number of layers in this network depends on the number of Up Sampling layers, which total 128, encompassing all architecture components. Despite the advantages of the SNSRGAN architecture in generating high-resolution images, there are certain limitations to consider. One notable disadvantage lies in the complexity of the model, particularly with the inclusion of 128 layers. This extensive layer count may increase computational demands during training and inference, potentially making the network less practical for deployment on resource-constrained devices. Additionally, the intricate structure of the network could result in longer training times, limiting its efficiency for real-time applications. As such, careful consideration of computational resources and time constraints is necessary when employing the SNSRGAN architecture.

2.3 Deep Learning-based Super Resolution

The Fast Super-Resolution Convolutional Neural Network (FSRCNN)[10] is focused on computational efficiency, utilizing a lightweight architecture with a blend of convolutional and deconvolution layers. While FSRCNN achieves speed improvements, it may exhibit limitations in handling extremely large scale factors. The Efficient Sub-Pixel Convolutional Neural Network (ESPCN)[22] is designed to efficiently upscale low-resolution images using sub-pixel convolutional layers. ESPCN excels in real-time applications due to its computational efficiency, but it might face difficulties with extremely challenging super-resolution tasks. Each model presents distinct advantages and disadvantages, and the choice among them depends on the specific requirements of the task. The Very Deep Super-Resolution Network (VDSR)[21] is characterized by its profound architecture with skip connection for image super-resolution tasks. Despite its depth contributing to enhanced feature learning, very deep networks like VDSR may encounter challenges such as vanishing gradients and exploding gradients during training. The Residual Channel Attention Network (RCAN)[23], specialized for image super-resolution, utilizes residual blocks with channel attention.

While demonstrating remarkable image enhancement, its reliance on ReLu activations can increase computational complexity, potentially affecting real-time efficiency and resource-constrained environments.

The Residual Dense Network (RDN) uses residual network. Also, this network is tailored for image super-resolution and employs a combination of residual learning and dense connections between

layers, which include global feature learning and local feature learning. This design facilitates efficient feature reuse across the network, enhancing its overall performance. However, the dense connections introduce increased model complexity, potentially leading to higher computational costs. Lastly, leveraging Remaining Channel Attention Networks (RCAN) advancements, our model achieves state-of-the-art super-resolution results, positioning it as a highly competitive contender in the image enhancement domain. These advantages collectively underscore the potential and effectiveness of our proposed model in advancing medical image analysis and diagnosis spacing.

Ⅲ. Method

In this section, we present the general methodology for our proposed network. This section encompasses The Remaining Enhanced Channel Attention Network (RECAN), explained as section 3.1, in Residual Groups (RG), explained as section 3.2, and Remaining Channel Attention Blocks (RECAB), explained as section 3.3.

3.1 Remaining Enhanced Channel Attention Network

In this section, we elucidate the RECAN architecture, which is structured around three key steps: the convolution group, residual group, and Up Sampling group. The initial phase involves the input image undergoing two sequential passes through a 3x3 convolution layer, enhancing its feature representation. Following this, the processed data advances to the first Residual Group (RG), in our model, there are four such groups. Each RG comprises four remaining channel attention blocks (RECAB) with convolution layers connected for each block layer. It is interconnected with a short skip connection originating from the second convolution layer. This design ensures the preservation and integration of crucial information.

The information then progresses to the sub-sequent 3x3 convolution layer, undergoing further feature extraction. This layer incorporates a long skip connection, facilitating data flow from the initial convolution layer. This strategic connection enhances the model's ability to capture local and global features.

Finally, the output is subjected to an Up Sampling layer, followed by a 1x1 convolution. This step contributes to refining and reconstructing the spatial details of the image. Combining skip

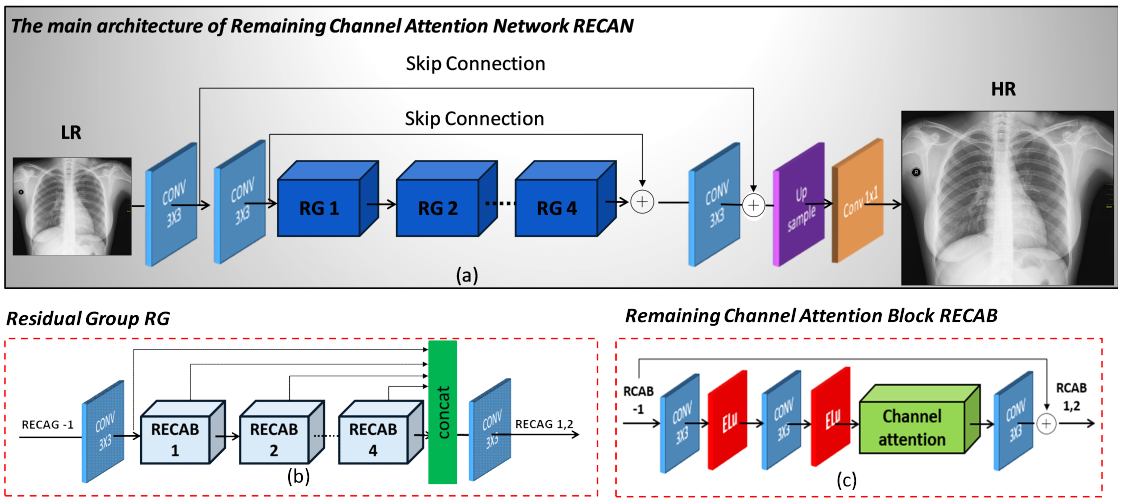


Fig. 1. (a) The main architecture of proposed Remaining enhanced channel attention network (RECAN), (b) figure illustrate residual group RG architecture, (c) figure illustrate remaining channel attention network architecture (RECAB)

connections, Remaining Channel Attention Blocks, and the Up Sampling layer collectively augments the RECAN architecture's effectiveness in feature extraction, information retention, and image reconstruction. The main architecture of the network is illustrated in Figure 1.

3.2 Residual Group

In this section, we elucidate the architectural design of the Residual Group. The entire network comprises four residual groups, each consisting of a convolutional layer and four additional residual groups. Notably, we have implemented dense feature fusion (DFF) within each Residual Group to optimize the utilization of features. Specifically, elements extracted from Remaining Channel Attention Blocks (RECABs) are concatenated with the input and fused through a 3x3 convolutional layer.

3.3 Remaining Channel Attention Block

This section details the critical component, the Remaining Channel Attention Block (RECAB). As well as we discuss the Residual Group (RG), an essential element in our research, including the integration of Remaining Channel Attention Blocks with the ELU activation function. The Remaining Channel Attention Block holds a fundamental position within the residual group of the Remaining Enhanced Attention Network (RECAN), initiating the journey of input information.

The RECAB input undergoes a convolutional layer, followed by an ELU activation function. This crucial step crops a non-zero output for negative inputs, proving especially advantageous for grayscale images. Subsequently, the output experiences another convolution using a 3x3 filter and ELU activation again. This property of ELU aids in preserving subtle details, particularly in the darker regions of the image, preventing the loss of valuable information during activation. The crux of this block lies in incorporating a channel attention mechanism, significantly enhancing the network's capability to focus on pivotal features. The channel attention mechanism enhances feature representation by selectively emphasizing informative channels, thereby improving convolutional neural networks' discriminative power and performance. Following the channel attention, another convolutional layer (3x3 filter) follows, accompanied by a skip connection originating from the first convolutional layer, a vital conduit for upholding and propagating essential information. The output from this block then progresses to the subsequent block for further processing and feature extraction.

In this way, four Remaining Channel Attention Block RCAB form 1 Residual Group. The intricate design of the Remaining Channel Attention Block is visually represented in Algorithm 1 and Figure 1 (c), encapsulating its core architecture.

Ⅳ. Experiments

4.1 Dataset

This study uses two distinct chest X-ray datasets: Chest X-ray 2017[14] and X-ray 2017 in this paper. The Chest X-ray 2014[15] X-Ray 2014 in this paper, X-Ray 2017 dataset comprises 5856 normal chest images from the American National Library of Medicine NIH. This dataset is divided into separate training and testing sets. The training set comprises 1194 normal images, while the test set comprises 120. The dataset explains the table shown in Table 1.

On the other hand, Chest X-ray 14[15], called X-Ray 2014 in this paper, comprises 112,120 frontal-view chest X-ray images acquired from 30,805 unique patients. Each image has dimensions of 1024 × 1024 pixels with 8-bit grayscale values. Moreover, this dataset includes annotations of 880 bounding boxes for eight pathologies performed by board-certified radiologists. We designated the 118 annotated images as the testing set for our analysis, while we randomly selected 1194 images

Table 1. The explanation of the dataset

HR	scale	LR	X-Ray 2014 [14]		X-Ray 2017 [15]	
			Train	Test	Train	Test
1024x1024	x4	256x256	1300	120	1194	118

for the training set.

4.2 Training Settings

In this research, we harnessed a potent hardware setup, leveraging an Intel(R) Core (TM) i7-11700K processor with an impressive 16 CPUs and a substantial 32GB RAM. Furthermore, we capitalized on the robust capabilities of an NVidia Titan XP boasting 16GB RAM for our training endeavors. To transform high-resolution (HR) images into low-resolution (LR) counterparts, we embraced a bicubic kernel-based down-sampling technique, employing a down-sampling factor of r = 2^k (where k ∈). This approach aligns with the established methodology delineated in SNSRGAN[8]. The training process of our model was meticulously optimized using the ADAM optimizer[16], fine-tuned with parameters β1 = 0.9, β2 = 0.999, and ε = 10^-8. Initially, we set the learning rate to 2 × 10^-4, implementing an exponential reduction of 0.1 every 120 epochs to enhance convergence. Our CNN training involved the utilization of different loss functions, encompassing L2 (sum of squared differences) and L1 (sum of absolute differences)[4]. In their image quality assessments employing metrics like MSIM[17] and SSIM[18], the L1 loss consistently showcased superior performance compared to the L2 loss. Building on this insight, we oriented our study to minimize the L1 distance between the original chest X-Ray images' input and their corresponding ground-truth images. The entire training process was orchestrated on the Tensor Flow platform.

To assess the efficacy of our proposed methodology, we conducted a comprehensive comparative analysis with a spectrum of state-of-the-art approaches, including Nearest-neighbor (NN)[19], Bicubic Interpolation[20], VDSR[21], ESPCN[22], FSRCNN[10], RDN[9], SRGAN[7], and SNSRGAN[8]. These methodologies were strategically categorized into two groups: interpolation-based methods (NN, Bicubic Interpolation), GAN-based methods (SRGAN, SNSRGAN), and deep-learning-based methods (ESPCN, FSRCNN, VDSR, RDN, and our proposed RECAN).

4.3 Equations

We assessed the enhanced images' quality using The Structural Similarity Index (SSIM) and the Multi-Scale Structural Similarity Index (MSIM).

The SSIM index is a well-established measure of image quality assessing the structural similarity between two images[18]. The SSIM index produces a value between -1 and 1, where 1 indicates perfect similarity. As well as SSIM index is calculated based on three components:

- Luminance (l): It measures the perceived brightness of the images.
- Contrast (c): It reflects the difference in intensity between an image and its surroundings.
- Structure (s): It considers the patterns or textures in the images.

The overall SSIM index is computed as a combination of these three components, and the equation (1).

$$SSIM(x,y) = (I(x,y)^{\alpha} * c(x,y)^{\beta} * s(x,y)^{\gamma}) \quad (1)$$

where, α, β, and γ are constants used to adjust the importance of each component.

MSIM[17] is an extension of SSIM that considers different scales or levels of details in images. Using a Gaussian pyramid, it breaks down the images into multiple scales and computes SSIM for each scale. The final MSIM score is a combination of SSIM scores across all scales. MSIM equation is explained as equation (2).

$$MSIM(x,y) = \frac{1}{N}\sum_{i=1}^{N} SSIM(x_i, y_i) \quad (2)$$

where N is depicting of total number of scales.

Ⅴ. Results and Ablation Study

5.1 Results

In this section, we present the experimental results of our research on grayscale chest X-ray images. We utilized two datasets: The Chest X-ray 2014 and the Chest X-ray 2017. We applied three super-resolution methods: Interpolation-based super-resolution, GAN-based super-resolution, and deep learning-based super-resolution.

For interpolation-based super-resolution using nearest neighbor (NN), we achieved SSIM values of 0.701 and 0.637 on X-Ray 2014 and X-Ray 2017, respectively. Moreover, the MSIM values were 0.713 and 0.668 for X-Ray 2014 and X-Ray 2017 datasets. Bicubic interpolation shaped SSIM scores of 0.687 and 0.615 on each X-Ray 2014 and X-Ray 2017 image. When trained on the ESPCN, SSIM values were 0.795 and 0.756 for each dataset, and the MSIM scores were 0.815 and 0.804 for X-Ray 2014 and X-Ray 2017, respectively. For FSRCNN, we obtained SSIM values of 0.921 and 0.913, in MSIM 0.962 and 0.968, on X-Ray 2014 and X-Ray 2017, respectively.

Using the SRGAN method, SSIM values of 0.856 and 0.809 were achieved for X-Ray 2014 and X-Ray 2017, respectively. The MSIM scores were 0.897 and 0.923 for each dataset. In the SNSRGAN method, SSIM values of 0.927 and 0.911 were for X-Ray 2014 and X-Ray 2017, respectively. The corresponding MSIM values were 0.988 and 0.983.

For VDSR, SSIM scores were 0.823 and 0.821, while MSIM scores were 0.92 and 0.923 for each X-Ray 2014 and X-Ray 2017 image. In the RDN method, SSIM scores were 0.927 and 0.931 for X-Ray 2014 and X-Ray 2017, respectively. The MSIM values were 0.976 and 0.982 for each X-Ray 2014 and X-Ray 2017 image.

In our proposed method, RECAN, we achieved impressive SSIM values of 0.938 and 0.939 for the X-Ray 2014 and X-Ray 2017 datasets, respectively. Moreover, the MSIM values were 0.99 and 0.989 for each dataset.

In summary, our proposed method, RECAN, demonstrated superior performance to other deep learning-based and GAN-based super-resolution methods, showcasing its potential for enhancing grayscale chest X-ray images.

5.2 Ablation study

Table 3 shows the conducted ablation study is to specific architectural modifications on the performance of the proposed RECAN model. To maintain consistency, all variants within the considered network adhere to the same RECAN configuration, with RG=4 and RECAB=4, and are subject to identical implementation details and experimental setups.

Table 2. Comparison of SSIM and MSIM results for different Super-Resolution Methods. Bold text is the best result, and underlined text is near the best result

METHOD	SCALE	X-Ray 2014		X-Ray 2017	
		SSIM	MSIM	SSIM	MSIM
NN	X4	0.701	0.713	0.637	0.668
Interpolation	X4	0.687	0.681	0.615	0.644
ESPCN	X4	0.795	0.815	0.756	0.804
FSRCNN	X4	0.917	0.953	0.897	0.953
SRGAN	X4	0.844	0.897	0.821	0.923
SNSRGAN	X4	0.927	<u>0.988</u>	0.911	<u>0.983</u>
VDSR	X4	0.823	0.920	0.821	0.923
RDN	X4	0.927	0.976	<u>0.931</u>	0.982
RECAN(ours)	X4	**0.938**	**0.990**	**0.939**	**0.989**

Table 3. Comparison of SSIM and MSIM results for different configurations of RECAN. Bold text indicates the best result, and underlined text represents results that are near the best.

RECAN		RG		RECAB			X-Ray 2014		X-Ray 2017	
		Concatenation	Skip connection	Elu	ReLu	CA	ssim	msim	ssim	msim
4RG 4RECAB	1	✓	✗	✓	✗	✓	**0.938**	**0.99**	**0.939**	**0.989**
	2	✗	✓	✓	✗	✓	0.83	0.872	0.826	0.881
	3	✗	✓	✗	✓	✓	0.861	0.932	0.901	0.952
	4	✓	✗	✗	✓	✓	<u>0.927</u>	<u>0.979</u>	<u>0.931</u>	<u>0.978</u>

5.2.1 Impact of Activation Function in RECAB with Skip Connection

The initial focus of our investigation lies in scrutinizing the role of skip connections within the proposed model. Surprisingly, our observations indicate that including skip connections when utilizing RECAB in the ELU function fails to yield optimal results. The second line of Table 3 depicts a substantial decline in performance (on X-ray 2014: SSIM 0.830, MSIM 0.872; on X-ray 2017: SSIM 0.826, MSIM 0.881) compared to the RG using concatenation.

However, when RECAB is configured with a ReLU function, it performs superior to ELU. Specifically, skip connections paired with RECABs containing ReLU functions exhibit a notable enhancement in performance. The third line of Table 3 substantiates this improvement (on X-ray 2014: SSIM 0.861, MSIM 0.932; on X-ray 2017: SSIM 0.901, MSIM 0.952). Consequently, for RECABs connected with skip connections, ReLU functions outperform ELU functions.

5.2.2 Comparative Analysis with Proposed RECAN Model

To underscore the significance of these architectural choices, we conducted a comparative analysis with variations involving concatenation in RG and ReLU functions in RECAB. Intriguingly, this alternative configuration yielded superior results, emphasizing the critical role of the selected architecture. Table 3, line 4, represents the utilization of concatenation in RG, with each residual group (RG) containing RECAB configured with ReLU functions.

However, this alternative configuration resulted in inferior performance on both X-ray 2014 and X-ray 2017 datasets (on X-ray 2014: SSIM 0.927, MSIM 0.979; on X-ray 2017: SSIM 0.931, MSIM 0.978). Conversely, our proposed RECAN model, showcased in Table 3, line 1, outperforms the alternative configuration when using RECAB with ReLU functions (X-ray 2014: SSIM 0.938, MSIM 0.990; on X-ray 2017: SSIM 0.939, MSIM 0.989).

5.2.3 Ablation Study Conclusion

The ablation study conducted in this investigation provides critical insights into the underlying factors influencing the performance of our proposed RECAN model. By systematically exploring specific architectural modifications and maintaining consistency in the RECAN configuration, we gained a nuanced understanding of the model's behavior under varied conditions. Notably, the ablation study revealed intriguing patterns in the impact of skip connections and the choice of activation functions within the network.

Our findings underscore that skip connections, while commonly employed for feature propagation, do not yield optimal results when used in conjunction with RECABs configured with the ELU activation function. Surprisingly, RECABs incorporating ReLU functions in conjunction with skip connections demonstrated superior performance, highlighting the importance of thoughtful activation function selection in the presence of skip connections.

Moreover, the comparative analysis with alternative configurations further substantiated the critical role of the chosen architecture. The unexpected outcome, where concatenation in RG and ReLU functions in RECAB resulted in inferior

performance, accentuates the unique strengths of our proposed RECAN model. This analysis emphasizes that the superior performance achieved by our model is not merely a product of individual components but rather a result of the intricate interplay between architectural choices.

Ⅵ. Conclusion

This study comprehensively explored and compared diverse super-resolution methods to enhance grayscale chest X-ray images by evaluating three distinct categories of super-resolution techniques—Interpolation-based, GAN-based, and deep learning-based—our analysis focused on assessing their performance using the Chest X-ray 2014 and Chest X-ray 2017 datasets.

As detailed in Table 2, the experimental findings underscore the efficacy of various super-resolution methods, gauged through Single-Scale Structural Similarity Index (SSIM) and Multi-Scale Structural Similarity Index (MSIM) scores. Our proposed RECAN method emerged as the top-performing approach, surpassing benchmarks set by nearest neighbor (NN), bicubic interpolation, SRGAN, SNSRGAN, and other deep learning-based methodologies.

The unexpected outcomes revealed by the ablation study further enhance the significance of our findings, shedding light on the nuanced interdependencies within the proposed RECAN architecture. Overall, this investigation contributes valuable insights into the promising capabilities of deep learning-driven super-resolution techniques in medical imaging. The robust performance of the RECAN model signifies a significant advancement, holding potential for improving diagnostic precision and fostering advancements in medical research and practice.

In conclusion, our investigation illuminates the promising capabilities of deep learning-driven super-resolution techniques in medical imaging. The robust performance of the proposed RECAN model signals a valuable stride forward, offering enhanced image quality that is integral to improving diagnostic precision and advancing the broader landscape of medical research and practice.

References

[1] A. A. Reshi, et al., "An efficient CNN model for COVID-19 disease detection based on X-ray image classification," *Complexity 2021*, pp. 1-12, 2021. (https://doi.org/10.1155/2021/6621607)

[2] Zhao, C. Y., Jia, R. S., Liu, Q. M., Liu, X. Y., Sun, H. M., & Zhang, X. L. (2021). Chest X-ray images super-resolution reconstruction via recursive neural network. Multimedia Tools and Applications, 80, 263-277. (https://doi.org/10.1007/s11042-020-09773-x)

[3] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proc. ECCV*, pp. 286-301, 2018. (https://doi.org/10.48550/arXiv.1807.02758)

[4] A. Salem, H. Ibrahim, and H.-S. Kang, "Light field reconstruction using residual networks on raw images," *Sensors*, vol. 22, no. 5, p. 1956, 2022. (https://doi.org/10.3390/s22051956)

[5] B. Lim, S. Son, H. Kim, S. Nah, and M. K. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE CVPR Wkshps.*, pp. 136-144, 2017. (https://doi.org/10.48550/arXiv.1707.02921)

[6] K. Dashdondov, M.-H. Kim, and M.-H. Song, "Deep autoencoder-based multivariate outlier detection for the classification of hypertension: Case study COVID-19," 2023. (https://doi.org/10.21203/rs.3.rs-2872456/v1)

[7] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. CVPR*, pp. 4681-4690, 2017 (https://doi.org/10.48550/arXiv.1609.04802)

[8] L. Xu, X. Zeng, Z. Huang, W. Li, and H. Zhang, "Low-dose chest X-ray image super-

resolution using generative adversarial nets with spectral normalization," *Biomed. Sign. Process. and Control*, vol. 55, p. 101600, 2020. (https://doi.org/10.1016/j.bspc.2019.101600)

[9] Y. Zhang, et al., "Residual dense network for image super-resolution," in *Proc. IEEE Conf. CVPR*, 2018.(DOI: 10.1109/CVPR.2018.00262)

[10] L. S. Passarella, S. Mahajan, A. Pal, and M. R. Norman, "Reconstructing high resolution ESM data through a novel fast super resolution convolutional neural network (FSRCNN)," *Geophysical Res. Lett.*, vol. 49, no. 4, e2021GL097571, 2022. (https://doi.org/10.1029/2021GL097571)

[11] M.-I. Georgescu, R. T. Ionescu, A.-I. Miron, O. Savencu, N.-C. Ristea, N. Verga, and F. S. Khan, "Multimodal multi-head convolutional attention with various kernel sizes for medical image super-resolution," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, pp. 2195-2205, 2023. (https://doi.org/10.48550/arXiv.2204.04218)

[12] A. Khishigdelger, A. Salem, and H. S. Kang, "Chest X-ray image super-resolution enhanced the ELU function with channel attention," in *Proc. KICS Conf.*, pp. 1386-1387, 2023.

[13] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (COVID-19) based on deep features," 2020.

[14] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, and F. Yan, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *cell*, vol. 172, pp. 1122-1131, e1129, 2018. (https://doi.org/10.1016/j.cell.2018.02.01)

[15] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proc. IEEE Conf. CVPR*, pp. 2097-2106, 2017. (https://doi.org/10.1109/CVPR.2017.369)

[16] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd ICLR'15*, p. 500, San Diego, 2015. (https://doi.org/10.48550/arXiv.1412.6980)

[17] Elizar, E., Zulkifley, M. A., Muharar, R., Zaman, M. H. M., & Mustaza, S. M. (2022). A review on multiscale-deep-learning applications. Sensors, 22(19), 7384. (https://doi.org/10.3390/s22197384)

[18] U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study," *J. Comput. and Commun.,* vol. 7, no. 3, pp. 8-18, 2019. (https://doi.org/10.4236/jcc.2019.73002)

[19] U. Khandelwal, A. Fan, D. Jurafsky, L. Zettlemoyer, and M. Lewis, "Nearest neighbor machine translation," *arXiv preprint arXiv: 2010.00710*, 2020. (https://doi.org/10.48550/arXiv.2010.00710)

[20] S. Gao and V. Gruev, "Bilinear and bicubic interpolation methods for division of focal plane polarimeters," *Optics express,* vol. 19, pp. 26161-26173, 2011. (https://doi.org/10.1364/OE.19.026161)

[21] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. CVPR*, 2016. (https://doi.org/10.1109/CVPR.2016.182)

[22] G. Sparacino, C. Tombolato, and C. Cobelli, "Maximum-likelihood versus maximum a posteriori parameter estimation of physiological system models: The C-peptide impulse response case study," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 6, pp. 801-811, 2000. (https://doi.org/10.1109/10.844232)

Anudari Khishigdelger

Jun. 2021 : B.S. degree, The National University of Mongolia

Feb. 2022~2024 : M.S. degree, Chungbuk National University

Feb. 2024~current : Ph.D degree, Chungbuk National University

<Research Interests> Nanoscience, Deep learning, Image processing, Bioinformatics, Aptamer, and Microbiology.

[ORCID:0009-0009-4201-3559]

Ahmed Salem

Jul. 2010 : B.S. degree, Assiut University, Egypt

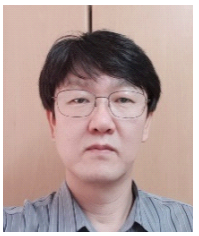
Oct. 2016 : M.S. degree, Egypt-Japan University of Science and Technology, Egypt

Aug. 2022 : Ph.D. degree, Chungbuk National University

<Research Interests> Machine Learning, Computer vision, Multi-view synthesis, and image restoration.

[ORCID:0000-0002-4682-0368]

Hyun Soo Kang



Feb. 1991 : B.S. degree, Kyungpook National University

Feb. 1994 : M.S. degree, Korea Advanced Institute of Science & Technology

Mar. 1991 : Ph.D. degree, Korea Advanced Institute of Science & Technology

Mar. 2005~Current : Professor, Chungbuk National University (research institute for computer and information communication)

<Research Interests> Computer Vision, Deep Learning

[ORCID:0000-0002-4333-2852]