Assessing the Generalizability of Deep Learning-Based Compression Techniques for Multibodypart X-ray Medical Images: A Comparative Study

Amina Fettah[1,*], Rafik Menassel[2] and Abdeljalil Gattal[2]

[1]Department of Mathemathics and Informatics, Laboratory of Mathematics, Informatics and System (LAMIS), Echahid Cheikh Larbi Tebessi University Tebessa, Algeria

[2]Department of Mathemathics and Informatics, Laboratory of Vision and Artificial Intelligence (LAVIA), Echahid Cheikh Larbi Tebessi University, Tebessa, Algeria

Abstract

For effective storage and transmission of healthcare data, efficient compression of medical images became necessary. Deep learning outperformed traditional image compression in terms of size reduction and image quality maintaining. However,this study is an extension of previously applied deep learning based compression techniques including an Autoencoder, a Deep Convolutional Autoencoder, originally on an x-ray medical imaging dataset (MXID), to OPEN-I and JSRT datasets with different characteristics. The main goal of this research is an evaluation of the generalizability of the different models across these datasets, and a comprehensive evaluation of the models' performance, highlighting the impact of the tree models on other dataset training in which remarkable findings achieved on larger datasets, while acceptable ones on smaller datasets. This promising contributions was intended for an efficient medical image compression by preserving the images quality, reducing its size, and minimize data loss while preserving image resolution which is crucial for healthcare domain, highlighting the influence of dataset size on compression performance.

Keywords

Deep Convolutional Autoencoder, Medical Image Compression, Deep Learning, MXID x-ray Dataset

1. Introduction

In contemporary healthcare, medical imaging plays a pivotal role by delivering important information about a patient's health. Transmission, analysis, and storage of MRIs, CT scans, and X-rays modalities presents significant challenges in file size and the need of reliable speed network, also in risk of losing crucial details which effects diagnostic accuracy. Another challenge is the volume data of these high-resolution images requires scalable and robust solutions. by focusing on the reduction of file size, lossy compression introduces some degradation in image quality nevertheless it significantly reduces image size [1]. It is essential to deploy effective compression strategies that reduce the size of data while preserving the diagnostically regions of medical imaging to address the presented challenges. The concept of Autoencoders (AEs) were originally introduced by LeCun in [2]. An AE comprises two neural network components: an encoder and a decoder as indicated in Figure 1. Furthermore, these two components are used in the convolutional autoencoder (CAE) which is an extension of the AE that employs convolutional neural networks (CNNs). However, CNNs consists of convolutional layers and pooling layers, convolutional layers comprise multiple nodes that process 2D feature maps, the learnable parameters within these layers are the elements of the filter matrices [3].

Maintaining a balance between image quality and compression efficiency is highly required in medical image field. Deep learning based methods compresses the images into an efficient latent representation that can be reconstructed providing a comparable result and a considerable achievement in the field

TACC'2024The 4th Tunisian-Algerian Conference on applied Computing, December 17-18, 2024, Constantine, Algeria

*Corresponding author.
amina.fettah@univ-tebessa.dz (A. Fettah); r.menassel@univ-tebessa.dz (R. Menassel); abdeljalil.gattal@univ-tebessa.dz (A. Gattal)
0009-0001-6187-9894 (A. Fettah); 0000-0003-3412-2551 (R. Menassel); 0000-0002-7986-2492 (A. Gattal)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

CEUR-WS.org/Vol-3973/paper13.pdf

of medical image compression while providing a balance between compression efficiency and high compressed image quality as discussed in [4] [5] [6].

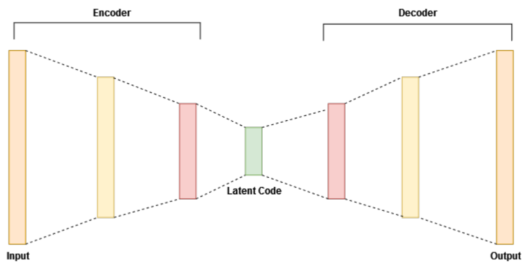


Figure 1: Autoencoders' Structure

The deep learning based image compression techniques applied to MXID X-ray dataset which was assembled to encompass 18 distinct body parts of anatomical structures from Abdomen to lung [7], was extended and evaluated in this study to two additional datasets OPEN-I,and JSRT for a broader evaluation of the technique's effectiveness. In compressing these images through our experiments, an exceptional performance demonstrated by the DCAE model on MXID, and OPEN-I datasets while preserving high-quality reconstructions, as evaluated by several metrics such as the Mean Squared Error (MSE), and Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM). In this paper, the application of the several deep learning based compression techniques to the JSRT dataset and the Open-I dataset remains under-explored. This study aims to address this gap by applying our previously developed models on these datasets that include a broad selection of X-ray images with a more different characteristics. Furthermore, to assess the models' robustness and generalizability on smaller and larger datasets, the techniques applied to these two additional datasets to compress a larger variety of medical imaging datasets while maintaining high image quality for accurate diagnosis.

This paper is organized as follows: Section 2 presents related work in image compression based autoencoder models. Section 3 details the methodology and results. Section 4 reviews the results and discussion. Finally, Section 5 conclusion and for future research.

2. Related Works

Recent works on autoencoder-based approaches dedicated to medical images compression, especially convolutional autoencoders (CAE) which preserves crucial diagnostic features by learning compact representations achieving an efficient image compression results.

To begin with, Mishra et al. [8] proposed a two-stage autoencoder framework for lossy compression of malaria-infected red blood cell images. For extracting unique features for image reconstruction, the method uses a residual-based dual autoencoder network which incorporates color structural similarity to maintain the quality of chrominance information, which is crucial for medical image analysis. This method demonstrated high performance in the different metrics such as Color SSIM, PSNR, And MS-SSIM. Moreover, dual autoencoders and residual learning highlights the potential of deep learning techniques for efficient medical image compression. On the other hand, Saravanan and Juliet [9] propose a deep autoencoder architecture for medical image reconstruction using a Deep Boltzmann Machine. Their aim was the improvement of reconstructed images quality by the power of ity reduction, their results attains high performance, lower reconstruction error and faster convergence. Efficient pre-trained Deep Boltzmann Machine lead to effective learning process in this work. Moving to Cheng et al. [10] , their research compares the performance of three deep learning architectures Convolutional Autoencoders (CAEs), Generative Adversarial Networks (GANS), and super-resolution (SRCNN) for image compression. However, SR-based compression leverages machine learn-ing based super resolution filters and BPG algorithms achieves the best coding performance. On the other hand, CAEs demonstrates their ability in the extraction of compact image features which is a promise tool for image compression domain. Concluding with M. Venugopal and K. Palanisamy, "Wavelet based Convolutional Autoencoder for Medical Image Compression" [11]. In this study, authors aim to increase the compression ratio whereas preserving diagnostic information for a high-quality reconstruction. integrate Convolutional Autoencoders (CAEs) with wavelet transforms as a hybrid technique for medical image compression. This approach uses the wavelet transform's multi-resolution analysis capability to decompose medical images into different frequency components, which are then encoded using the CAE. Wavelet-based Convolutional Autoencoder (W-CAE) outperforms standard CAEs and traditional techniques by an efficient compression quality.

On another research, and using convolutional neural networks (CNNs) and wavelet transformation on MRI, x-ray, and CT scans, Shukla et al [12] proposed an approach that compresses medical images using three hidden layer network, which outperformed lossy and lossless image compression techniques.

The following sections provide an overview of the used datasets, used evaluation metrics, the achieved results, and the discussion part. Subsequently, we explore the potential of the DCAE on the different medical imaging datasets and its performance evaluation and discussion.

3. Methodology

This section is organized as follows: datasets used, performance evaluation metrics chosen, and results and discussion.

3.1. Dataset

3.1.1. MXID dataset

MXID dataset [7] is a collection of 6,869 high-resolution X-ray images with 1024x1024 pixels size from AOUINET Hospital in Tébessa, Algeria. Classified manually into 18 distinct body parts ranging from the lung and abdomen to the wrist and pelvic basin.

the DCAE model was applied on the following additional datasets:

3.1.2. OPEN-I dataset

OPEN-I [13] is a de-identified Indiana chest X-ray collection from the Indiana Network for Patient Care consists of 8121 associated images and 3996 radiology reports. Authors performed de-identification on reports and images, automatic de-identification of images required manual verification and was not perfect, the reports' manual coding resulted in improved retrieval precision.

3.1.3. JSRT dataset

JSRT is a digital database [14] consisting chest radiographs collected from 14 medical centers, with a total of 247 images with and without lung nodules at high resolution (2048 x 2048 matrix size) with a 12-bit grayscale. These chest x-ray radiographs grouped based on the lung nodules' subtlety with varying characteristics.

The same pre-processing steps applied on the mentioned datasets, including resizing to a smaller resolution of 256x256 pixels and normalization, to ensure uniformity in input data format, with a batch size of 4 due to memory limitations. The datasets were divided into three distinct sets: training (60%), testing (20%), and validation (20%).

3.2. Evaluation Metrics

We used a set of quantitative metrics to evaluate the Deep learning based model's performance:

3.2.1. Mean Squared Error (MSE)

MSE used to measure the average of the squared differences between the reconstructed and original image; It is widely used in image processing, signal processing, and machine learning.

$$\text{MSE} = \frac{1}{mn}\sum_{x=1}^{m}\sum_{y=1}^{n}(I(x,y) - K(x,y))^2 \tag{1}$$

3.2.2. Multi-Scale Structural Similarity Index (MS-SSIM)

Multi-Scale Structural Similarity Index (MS-SSIM) used to measure the similarity between two images based on luminance, contrast, and structure. Derived from SSIM to evaluate quality considering structural information.

$$\text{MS-SSIM}(x,y) = [l_m(x,y)]^{\alpha M}\prod_{j=1}^{M}[C_j(x,y)]^{\beta_j}\,[S_j(x,y)]^{\gamma_j} \tag{2}$$

3.2.3. Peak Signal-to-Noise Ratio (PSNR)

PSNR is a measure of the difference of the quality for the reconstructed image compared to the original image derived from MSE and expressed in decibels (dB). It is widely used to evaluate the performance of compression methods. It represents the ratio between power of corrupting noise and the maximum possible power of a signal.

$$\text{PSNR} = 20 \cdot \log_{10}(\text{MAX}) - 10 \cdot \log_{10}(\text{MSE}) \tag{3}$$

3.3. Results and Discussion

3.3.1. Performance on MXID Dataset

- deep convolutional autoencoder (DCAE): For deep convolutional autoencoder (DCAE) architecture, the results has show an outstanding level of performance in X-ray medical image reconstruction. After 51 epochs as depicted in Fig. 2 with a loss of 0.002 and using the PRELU activation function, for ensuring the trade-off between diagnostic information overall quality and compression efficiency, as demonstrated in Figure 11. Also, a PSNR value of 46.78 dB with a minimal loss of information, with an MS-SSIM value of 0.99 suggesting an elevated similarity between original and reconstructed image.
- Autoencoder (AE): For the application of the autoencoder on MXID dataset , results have shown a PSNR of 37.61 dB, MSE of 0.014, and an MS-SSIM of 0.61 in [7], as depicted in Fig. 5, representing a certain loss of details in the image quality , which is important for medical imaging applications. While the overall structural integrity of the images has been preserved proposing an enhancement in the auto-encoder model's architecture for more accurate metrics values and a better image quality, as illustrated in Figure 11.
- Convolutional Neural Network (CNN): CNN's results indicates a certain loss after 21 epochs on MXID dataset, with an MSE of 0.006, PSNR of 41.43 dB, and an MS-SSIM of 0.77 [7] showing a better results than autoencoder's especially in terms of PSNR. deeper architecture may reflects better preservation of fine details which is crucial for medical imaging diagnosis, especially in the regions of interest.

3.3.2. Performance on OPEN-I Dataset

- deep convolutional autoencoder (DCAE): when applying the DCAE on OPEN-I dataset, results outperforms all the other techniques with a loss of 0.0001 which is relatively lower than the application of the same DCAE architechture on MXID and JSRT datasets. Furthermore, a PSNR of 47.14 dB indicates a higher visual quality and a superior preservation of the different anatomical regions with an MS-SSIM value of 0.99 presenting a very high similarity between the input image and reconstructed

one, as demonstrated in Figure 12.
- Autoencoder (AE): For comparing the autoencoders' results on OPENI dataset with other datasets, after 21 epochs autoencoder model achieved a PSNR of 39.07 dB, an MSE of 0.011, and an MS-SSIM of 0.70 which is the highest performance compared to the results on MXID dataset and JSRT dataset; indicating a better compression quality and effective generalizability on larger dataset.

- Convolutional Neural Network (CNN): Similarly, the application of the CNN on the OPENI dataset has achieved an MSE of 0.004, PSNR of 42.40 dB, and an MS-SSIM of 0.80 which indicates the better performance compared to MXID and JSRT results as described in Table 1.

3.3.3. Performance on JSRT Dataset

- Deep Convolutional AutoEncoder (DCAE): the DCAE's performance on the JSRT dataset, characterized by a PSNR of 45.37 dB, an SSIM of 0.97, and an MSE of 0.001 after 21 epochs as illustrated in Figure 7, indicates a slight reduction in image quality compared to the MXID and Open-I datasets due to the JSRT imaging conditions and its potentially smaller dataset size, Figure 13 presents the visual quality results for more clarity. Nevertheless, the subtle differences in SSIM and compression metrics values imply that dataset factors, including size, variability, and image quality, impact the model's performance.
- Autoencoder (AE): Moving to autoencoder's performance on JSRT dataset, and after 18 epochs the model acheived a slightly superior PSNR of 37.84 dB compared to MXID, an MSE of 0.014 which is similar to autoencoder on MXID results, and MS-SSIM of 0.77 which represents the highest value compared to the two other datasets, reflecting that the original and reconstructed images are more similar.

- Convolutional Neural Network (CNN): Model achieved an MSE of 0.016, PSNR of 36.96 dB, and an MS-SSIM of 0.71 after 35 epochs as mentioned in table 1, which represents the lowest performance of the CNN application in comparison to MXID, and OPEN-I datasets leading to under-performance issues. Therefore, smaller datasets may not be suitable for CNN models which requires to learn complex patterns.

Ultimately, the smaller JSRT dataset's size and variability in image quality, such as noisier images or lower quality affected the model's ability to generalize, also struggling to maintain crucial fine details for an accurate diagnosis in the reconstructed images. All in all, smaller datasets may increase the overfitting risk, especially for models with high parameters, such as deep convolutional architecture. Additionally, AEs capture less complex characteristics due to limited data diversity, CNNs outperform AEs in terms of feature extraction, while it need a larger dataset to generalize better. Finally, the DCAE model extracts features combining the advantages of AE, and CNN's models; however, with the JSRT dataset, the DCAE model captures the features better than AEs but may overfit.

3.3.4. Discussion

The performance of the different techniques across the three different X-ray datasets is particularly higher on larger and more diverse datasets including MXID, and OPEN-I highlighting the models potential for different clinical application. However, subtle variations in reconstructed image quality between datasets suggest the models sensitivity to specific dataset characteristics. This includes factors along with size, image diversity, and inherent variations within the data. These findings enables the models to learn more features on larger datasets.

Furthermore, for MXID dataset PSNR and MS-SSIM values indicates a high quality in reconstructed images across the three models (AE, CNN, DCAE), preserving fine structural details all over the 18 body parts by maintaining and capturing structural information-based bone nuances, nerve structures, joint spaces, skeletal structures, vertebral structures, intervertebral discs, organs, tissues, and cardiovascular details, which reflects a balance between image quality preservation and compression for multi-body parts imaging datasets, Figure 11, 12, 13 present a side-by-side visual comparison of the results.

Moreover, OPEN-I chest x-ray dataset presents a slightly higher PSNR values across the three deep learning models, due to the dataset's larger size that offered more comprehensive training data enabling the models to capture more features related to chest anatomy. In addition, larger datasets reduce the overfitting's risk, and improve the model's reconstruction quality.

For the third dataset, a smaller JSRT chest X-ray dataset, indicates a slightly lower PSNR values than OPEN-I and MXID datasets across the three deep learning models (AE, CNN, DCAE) due to the smaller dataset's size that limits the models ability to capture more critical diagnostic characteristics. Additionally, the image quality variability in this dataset may lead to a decreased PSNR and MS-SSIM values. Also, the blurriness produced in the reconstructed images could mask early signs of diseases such as pneumonia or fibrosis in medical, with an existing loss of contrast leading to a subtle differences in tissue density that are important for diagnosis.

Table 1
DCAE, AE, CNN Performance Comparison on: MXID vs. OPEN-I vs. JSRT

Methods	Dataset	Epochs	MSE	PSNR	MS-SSIM
DCAE	MXID	51	0.0002	46.78	0.99
	OPENI	**3**0	**0**.0001	**4**7.14	**0**.99
	JSRT	21	0.001	45.37	0.97
AE	MXID	35	0.014	37.61	0.61
	OPENI	**2**1	**0**.011	**3**9.07	**0**.70
	JSRT	18	0.014	37.84	0.77
CNN	MXID	21	0.006	41.43	0.77
	OPENI	**3**5	**0**.004	**4**2.40	**0**.80
	JSRT	35	0.016	36.96	0.71

Traditional methods, such as JPEG and JPEG2000 are widely used for medical imaging compression due to their computational efficiency, therefore, the JPEG method may achieve higher compression ratio values due to the noticeable loss of the crucial details, which is crucial for the diagnosis process by specialists. On the other hand, JPEG2000 may preserve better quality at high compression levels, it also outperforms JPEG in terms of image quality preservation [15], while balancing image fidelity and storage needs, yet it requires more processing power, and still struggling in complex features and details caption compared to deep learning techniques that offer potential preservation while achieving high compression ratios.

Furthermore, Applying deep learning techniques to the MXID and other datasets presents several challenges. The MXID dataset exhibits several body parts with varied characteristics, luminosity, and contrast levels; However, while the autoencoder reduces the dimensions and noise, it fails to capture complex features for reconstructed medical images, which limits the capture of details in the different anatomical regions.

On the other hand, although CNNs give better results in terms of complex feature extraction, they need larger datasets to avoid overfitting and give better reconstructed images, Figure 13 demonstrate the impact of the CNN model on the JSRT dataset images appears highly blurred and lacks clear anatomical details, indicating that the model struggles to capture fine structures. The DCAE, model still achieves high-quality compression with impressive results exceptionally on the OPEN-I dataset in which the reconstruction error is minimized and the fine details are well preserved for a better accurate diagnosis; Nevertheless, regarding its powerful encoding efficiency in maintaining structural information details through convolutional layers, and its adaptability to noise which is essential in medical imaging, DCAEs are sensitive to training data quality, once it contains artifacts or noise the model maintain these characteristics too, which affects the reconstructed image quality, also, the training process for large datasets with high-resolution can be time-consuming.

These findings highlight the potential of the different techniques across different complex and larger medical imaging datasets while presenting limitations with smaller dataset characteristics such as noise levels, especially in the JSRT dataset.

4. Conclusion

This research examines the DCAE, AE, and CNN models' generalizability and robustness in compressing diverse X-ray datasets. The models demonstrated high compression PSNR values and maintained image quality, as measured by SSIM and MSE, on the MXID and OPEN-I datasets, presenting outstanding results on a larger dataset with different characteristics with a wider range of images, and successfully compressing images while preserving their high quality.

Moreover, its performance was not as evident on a lower-quality, varying noise quality, and smaller datasets as mentioned in the JSRT dataset which has fewer features extraction's opportunities, this may indicate certain limitations in processing images with decreased clarity or detail, and decreasing its generalizability. Future work will entail the integration of the different optimization strategies to enhance generalizability , such as transfer learning techniques pretrained on larger datasets, then fine-tuned on smaller datasets. Also, data augmentation strategies, including scaling, rotation, could increase dataset variability to generalize better different image conditions , and improve models performance.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] A. Fettah, R. Menassel, A. Gattal, Machine learning for medical image analysis: A survey, in: International Conference on Advanced Intelligent Systems for Sustainable Development, Springer, 2022, pp. 148–164.

[2] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (2015) 436–444.

[3] J. Zhai, S. Zhang, J. Chen, Q. He, Autoencoder and its various variants, in: 2018 IEEE international conference on systems, man, and cybernetics (SMC), IEEE, 2018, pp. 415–419.

[4] Y. Raut, T. Tiwari, P. Pande, P. Thakar, Image compression using convolutional autoencoder, in: ICDSMLA 2019, Springer, 2020, pp. 221–230.

[5] G. Guerrisi, F. Del Frate, G. Schiavon, Convolutional autoencoder algorithm for on-board image compression, in: IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2022, pp. 151–154.

[6] Z. Cheng, H. Sun, M. Takeuchi, J. Katto, Deep convolutional autoencoder-based lossy image compression, in: 2018 Picture Coding Symposium (PCS), IEEE, 2018, pp. 253–257.

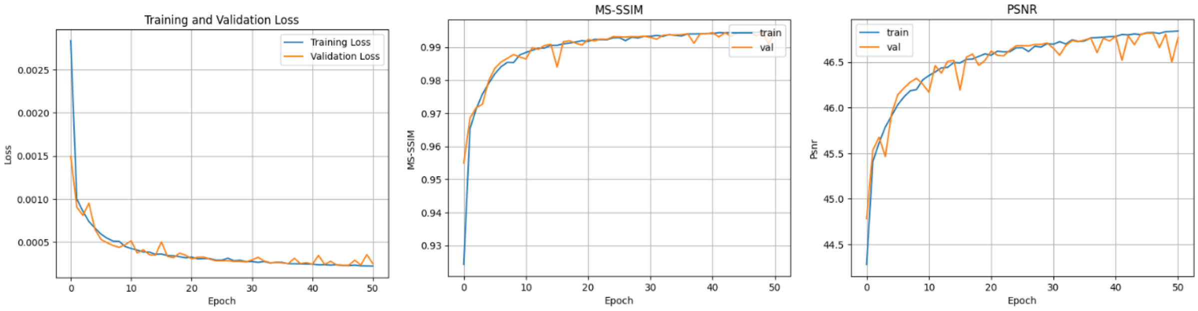
[7] A. Fettah, R. Menassel, A. Gattal, A. Gattal, Convolutional autoencoder-based medical image compression using a novel annotated medical x-ray imaging dataset, Biomedical Signal Processing and Control 94 (2024) 106238.

[8] D. Mishra, S. K. Singh, R. K. Singh, Lossy medical image compression using residual learning-based dual autoencoder model, in: 2020 IEEE 7th Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON), IEEE, 2020, pp. 1–5.

[9] S. Juliet, et al., Deep medical image reconstruction with autoencoders using deep boltzmann machine training, EAI Endorsed Transactions on Pervasive Health and Technology 6 (2020).

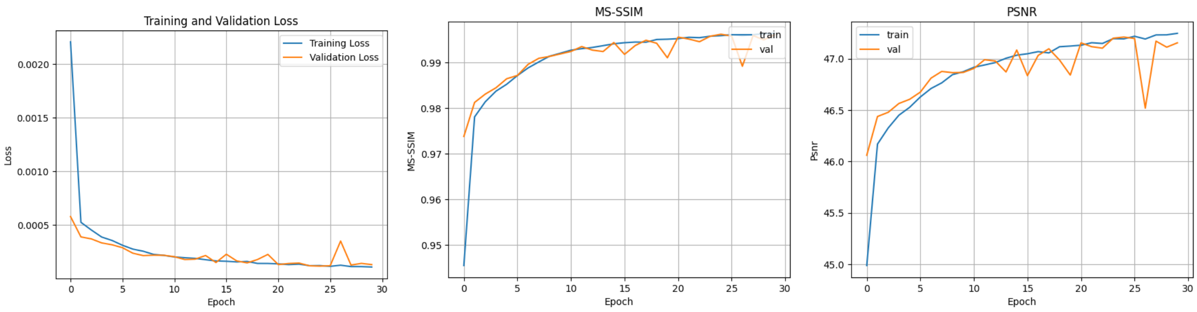
[10] Z. Cheng, H. Sun, M. Takeuchi, J. Katto, Performance comparison of convolutional autoencoders, generative adversarial networks and super-resolution for image compression, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 2613–2616.

[11] M. Venugopal, K. Palanisamy, Wavelet based convolutional autoencoder for medical image compression, in: 2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), IEEE, 2023, pp. 83–88.



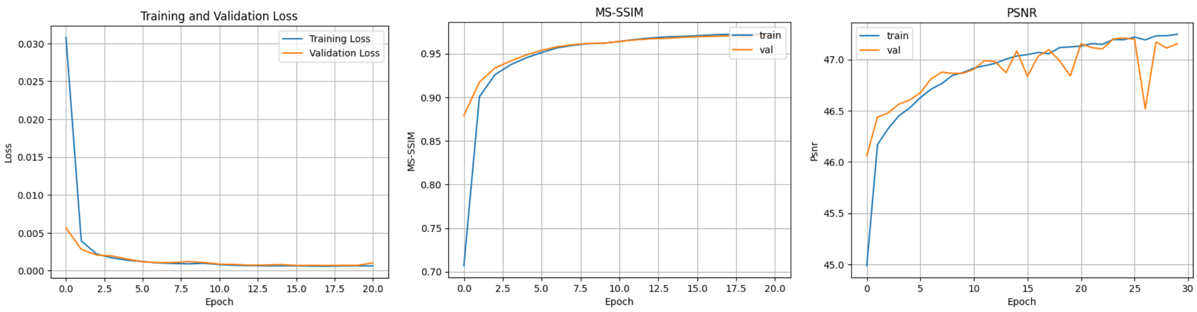
(a) (b) (c)

Figure 2: MXID DCAE Results: MSE, MS-SSIM, PSNR [7]



(a) (b) (c)

Figure 3: OPEN-I DCAE Results: MSE, MS-SSIM, PSNR



(a) (b) (c)

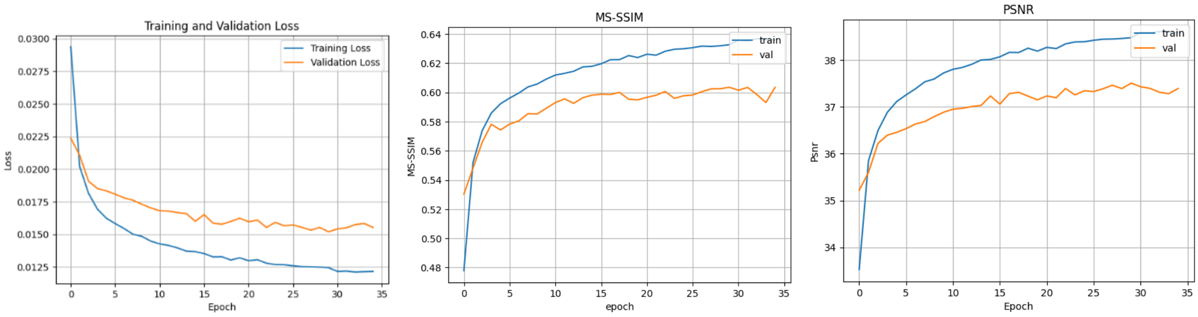
Figure 4: JSRT DCAE Results: MSE, MS-SSIM, PSNR

[12] S. Shukla, A. Srivastava, Medical images compression using convolutional neural network with lwt, International Journal of Modern Communication Technologies and Research 6 (2018) 265086.

[13] D. Demner-Fushman, M. D. Kohli, M. B. Rosenman, S. E. Shooshan, L. Rodriguez, S. Antani, G. R. Thoma, C. J. McDonald, Preparing a collection of radiology examinations for distribution and retrieval, Journal of the American Medical Informatics Association 23 (2016) 304–310.

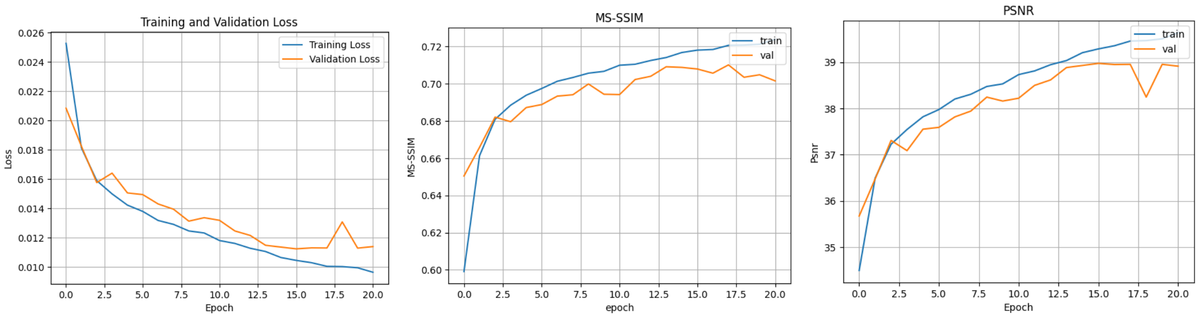
[14] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K.-i. Komatsu, M. Matsui, H. Fujita, Y. Kodera, K. Doi, Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules, American journal of roentgenology 174 (2000) 71–74.

[15] Y.-H. Shiao, T.-J. Chen, K.-S. Chuang, C.-H. Lin, C.-C. Chuang, Quality of compressed medical images, Journal of Digital Imaging 20 (2007) 149–159.



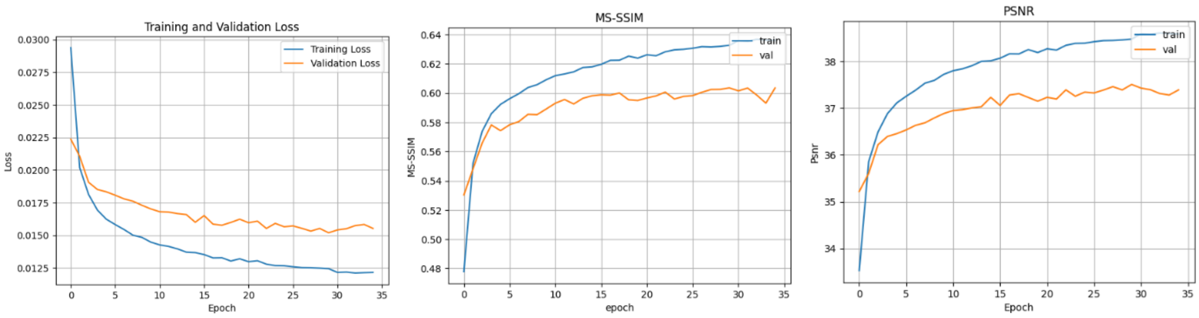
(a) (b) (c)

Figure 5: MXID AE Results: MSE, MS-SSIM, PSNR [7]



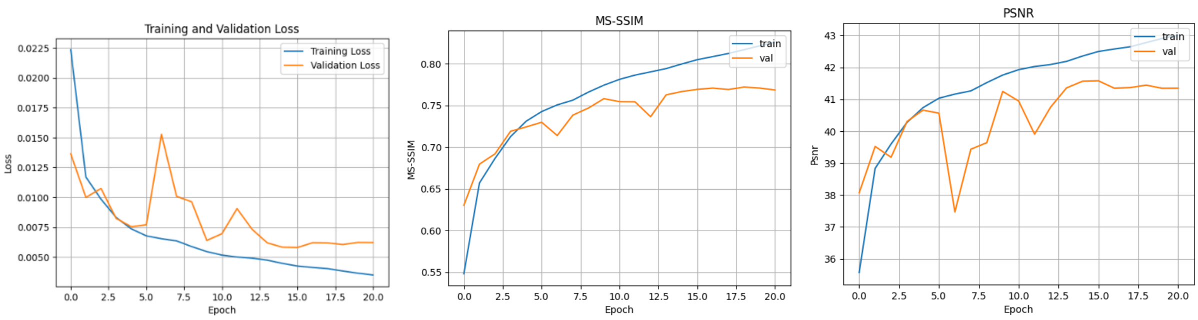
(a) (b) (c)

Figure 6: OPEN-I AE Results: MSE, MS-SSIM, PSNR



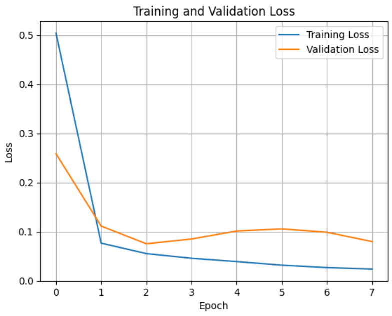
(a) (b) (c)

Figure 7: JSRT AE Results: MSE, MS-SSIM, PSNR

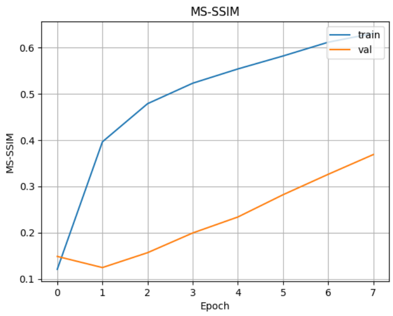


(a) (b) (c)

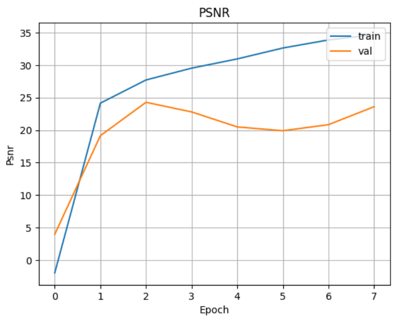
Figure 8: MXID CNN Results: MSE, MS-SSIM, PSNR [7]



(a)

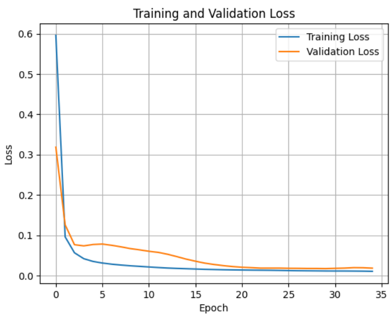


(b)

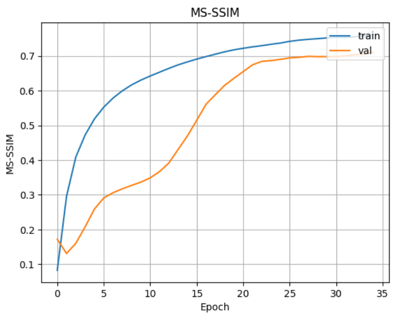


(c)

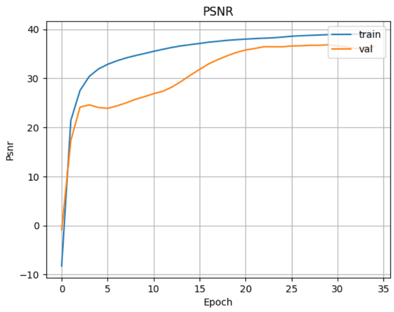
Figure 9: OPEN-I CNN Results: MSE, MS-SSIM, PSNR



(a)



(b)



(c)

Figure 10: JSRT CNN Results: MSE, MS-SSIM, PSNR

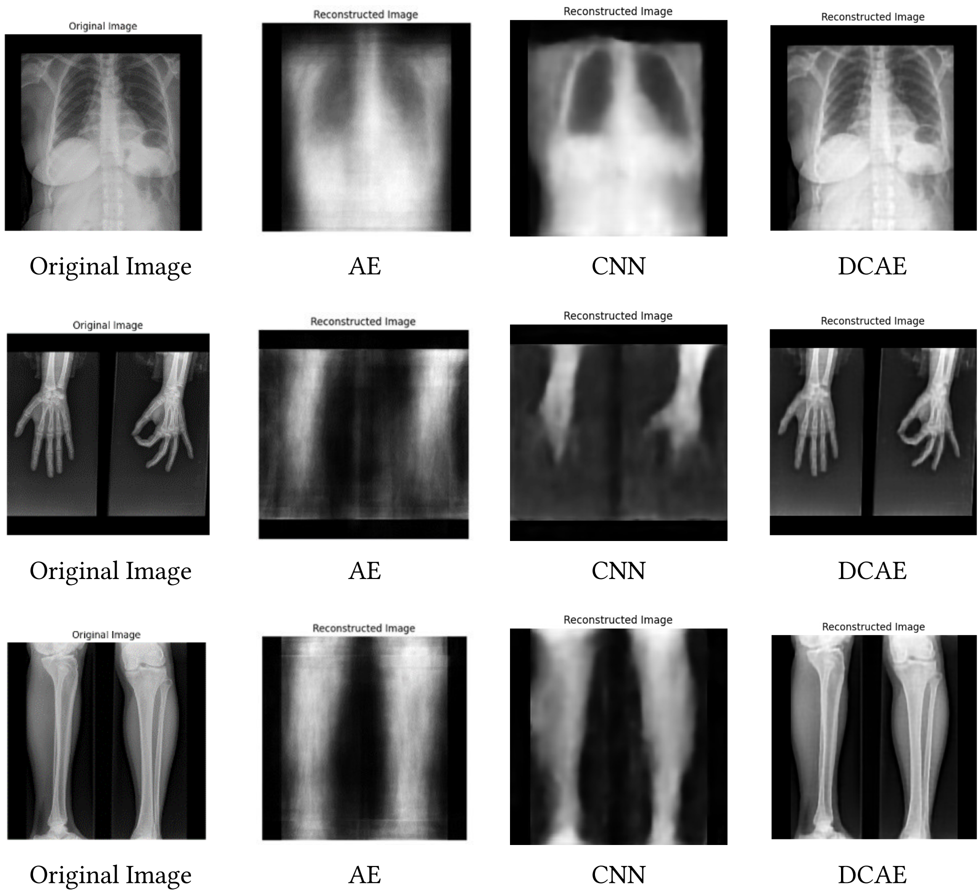


Figure 11: CNN vs. AE vs DCAE. results on MXID

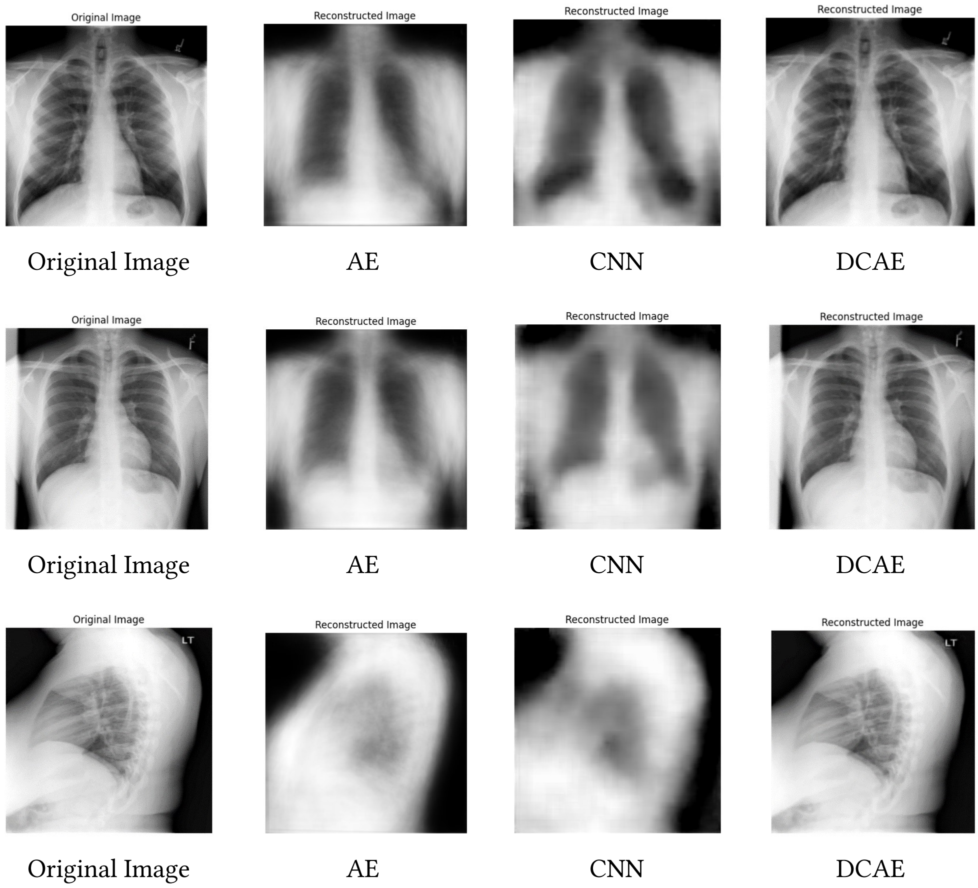


Figure 12: CNN vs. AE vs DCAE. results on OPEN-I

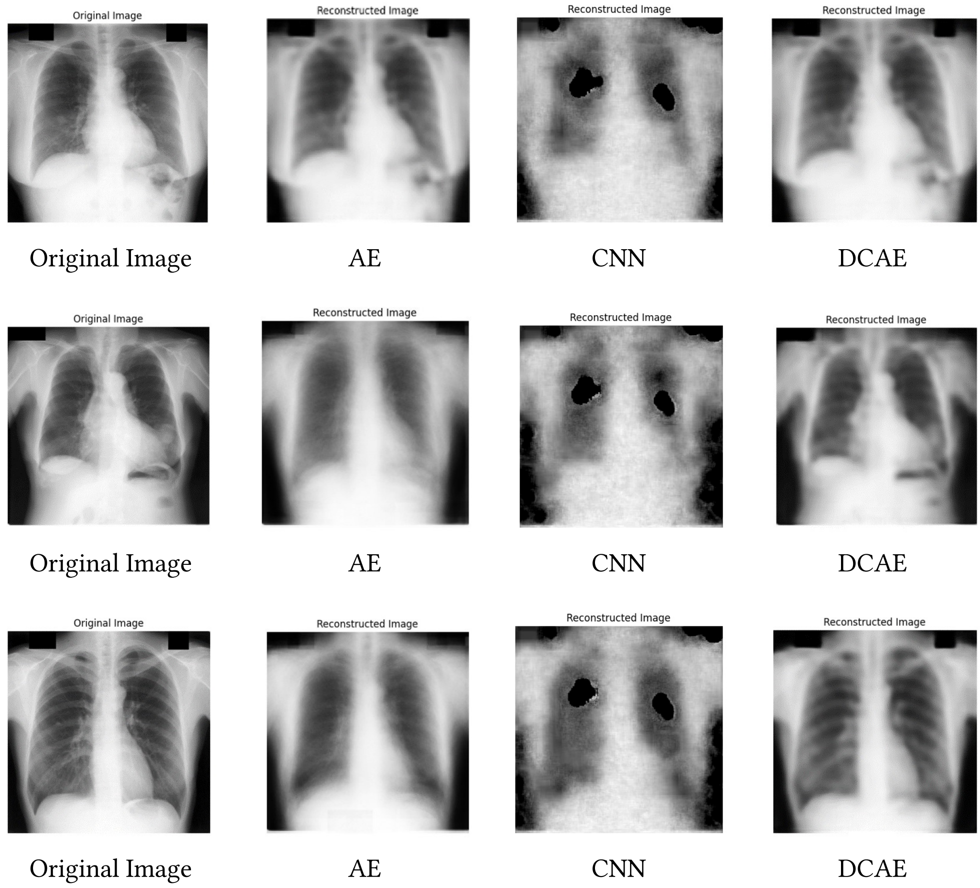


Figure 13: CNN vs. AE vs DCAE. results on JSRT