People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University Echahid Cheikh Larbi Tebessi- Tebessa

Faculty of Exact Sciences and Natural and Life Sciences
Department of computer science/mathematics
Domiciliation laboratory: laboratory of mathematics, informatics and systems (LAMIS)

Thesis

Presented and Publicly Defended for Obtaining the Diploma of Doctorate in Third Cycle

By:
Fettah Amina

Domain: mathematics and computer science
Branch: computer science
Speciality: Artificial Intelligence

Title

Machine Learning Techniques for Medical Image Compression

Defended on : 22 April 2025 In front of the jury composed of:

Full name	Rank	University	
Mr. Djeddi Chawki	**Pr**	Univ. of Tebessa	President
Mr. Menassel Rafik	Dr	Univ. of Tebessa	Supervisor
Mr. Gattal Abdeljalil	Pr	Univ. of Tebessa	Co- supervisor
Mr. Moukhati Farid	Pr	Univ. of Badji Mokhtar	Examiner
Mr. Sari Toufik	Pr	Univ. of OEB	Examiner
Mr. Laimeche Lakhdar	Pr	Univ. of Tebessa	Examiner

Academic year: 2024/2025

This work is dedicated to my cherished parents, my supportive brothers, and my beloved husband…

"What we have once enjoyed we can never lose. All that we love deeply becomes a part of us."

— Helen Keller

ACKNOWLEGMENT

Most of all, I would like to express my deepest gratitude to my supervisor, **Dr. Rafik Menassel**, whose untiring supervision, ongoing encouragement, and helpful feedback have touched every aspect of this research. Your expertise has been a steady source of motivation, and your support has been priceless in helping me navigate the complexities and challenges of this work. I am truly fortunate to have had you as a mentor throughout this process.

I am also greatly beholden to my co-supervisor, **Pr. Abdeljalil Gattal**, for his valuable suggestions and thoughtful support in refining the methodologies used in this research. His advice and proposals have been instrumental in enhancing the quality of the research and aligning it with the requirements of the discipline.

Special gratitude to El-Aouinet Hospital in Tebessa for granting access to the X-ray image datasets and for their cooperation in this project. I would also like to extend my heartfelt thanks to **Dr. Gattal Abdelhak,** the radiologist, whose expertise in medical imaging played a significant role in the dataset collection process. His endeavors in verifying and making sure that the images were correctly annotated were of great value in upholding the quality and integrity of the dataset.

I would also like to thank from the bottom of my heart the board of examiners **Pr. Moukhati Farid, Pr. Sari Toufik**, and **Pr. Laimeche Lakhdar** for agreeing to read and evaluate my thesis. I would also like to thank **Pr. Chawki Djeddi** for agreeing to preside over my defense jury. Your remarks and recommendations have contributed a lot to the final shape of this work.

To my beloved father, whose rightful words of encouragement and motivation during the codification process were the real pillars upon which this research was founded. Your passion and leadership have been the radiant light that has ever shone on me to reach greater heights and perform better.

To my mother, whose unending love, support, and belief in me have been the cornerstone on which all my actions. Your presence was a never-ending fountain of courage, inspiration, and the well upon which my dreams were fed.

To my brothers, whose camaraderie has been a force of laughter and support, illumination on the blackest nights. Your love has been the continual reminder of life's greatest riches—the ties that cannot be shattered, regardless of what we lived through.

To my husband, my best friend, whose love, patience, and deep understanding have been the constant anchor that kept me steadfast through it all. Not just my support, but the love of my life too, always with me, supporting my dreams and desires. You being in my life is what gives me warmth and makes my life blissful, and your belief in me is what makes me hang in there, makes me strong enough to carry on. I am eternally grateful for our love.

To my in-laws, whose warmth, generosity of spirit, and love have been comforting at all times. Your life has been a cause of inspiration, support, and happiness, for which I thank you forever.

Finally, a heartfelt thanks to **Dr. Takoua LAIB** for her help in the verification of this thesis. Your support has been an inspiration, best wishes in your academic journey.

Abstract

Medical imaging plays a vital part in modern-day healthcare practice, enabling diagnosis, treatment planning, and follow-up in the patient. Among various imaging modalities, X-ray imaging remains one of the most common due to its effectiveness, widespread availability, and diagnostic relevance. However, the continuous increase in the volume and resolution of medical images, especially X-rays, brings data storage, transmission, and processing challenges of utmost importance. These challenges normally lead to costly and complex management issues, especially in large healthcare systems.

To address these challenges, both traditional compression techniques and more recent machine learning-based techniques have been explored. Traditional techniques, such as JPEG and PNG, offer basic compression but are not tailored to the specific needs of medical imaging since they are unable to preserve diagnostic quality or exploit the unique structural patterns in medical images. Yet, deep learning techniques, particularly autoencoders, convolutional neural networks (CNNs), and Variational autoencoders (VAEs), have demonstrated excellent potential in improving both image quality and compression ratios, surmounting the limitations of conventional techniques.

This research investigates the potential of deep learning-based image compression models—like Autoencoders (AE), Deep Convolutional Autoencoders (DCAE), CNNs, and VAEs—for achieving better medical X-ray image compression. It also incorporates machine learning algorithms such as Principal Component Analysis (PCA) and K-means clustering for the optimization of image storage and transmission. Additionally, this study introduces the Medical X-ray Imaging Dataset (MXID), a high-quality dataset of X-ray images from AOUINET Hospital in Tebessa, Algeria, to facilitate various medical imaging tasks, including compression, classification, and machine learning applications. The study's findings demonstrate that deep learning-based models, specifically DCAEs, outperform alternative compression techniques in terms of image compression efficiency and quality retention yielding a PSNR of 46,78 dB.

The ultimate objective of this research in the long term is to contribute to the development of efficient, high-quality medical image compression techniques that will be able to improve the handling of healthcare data, thereby facilitating more effective healthcare delivery through better storage and transmission.

Keywords: X-ray Medical Image, Deep Learning, Machine Learning, Image Compression, Medical X-ray Imaging Dataset (MXID).

Résumé

L'imagerie médicale joue un rôle essentiel dans la pratique moderne des soins de santé, permettant le diagnostic, la planification du traitement et le suivi du patient. Parmi les différentes modalités d'imagerie, la radiographie reste l'une des plus courantes en raison de son efficacité, de sa grande disponibilité et de sa pertinence diagnostique. Cependant, l'augmentation continue du volume et de la résolution des images médicales, en particulier des rayons X, pose des problèmes de stockage, de transmission et de traitement des données de la plus haute importance. Ces défis conduisent généralement à des problèmes de gestion coûteux et complexes, en particulier dans les grands systèmes de soins de santé.

Pour relever ces défis, les techniques de compression traditionnelles et les techniques plus récentes basées sur l'apprentissage automatique ont été explorées. Les techniques traditionnelles, telles que JPEG et PNG, offrent une compression de base mais ne sont pas adaptées aux besoins spécifiques de l'imagerie médicale, car elles sont incapables de préserver la qualité du diagnostic ou d'exploiter les modèles structurels uniques des images médicales. Cependant, les techniques d'apprentissage profond, en particulier les autoencodeurs, les réseaux neuronaux convolutifs (CNN) et les autoencodeurs variationnels (VAE), ont démontré un excellent potentiel pour améliorer à la fois la qualité de l'image et les taux de compression, surmontant ainsi les limites des techniques conventionnelles.

Cette recherche étudie le potentiel des modèles de compression d'images basés sur l'apprentissage profond, tels que les autoencodeurs (AE), les autoencodeurs convolutionnels profonds (DCAE), les CNN et les VAE, afin d'obtenir une meilleure compression des images radiologiques médicales. Elle intègre également des algorithmes d'apprentissage automatique tels que l'analyse en composantes principales (ACP) et le regroupement K-means pour l'optimisation du stockage et de la transmission des images. En outre, cette étude présente le Medical X-ray Imaging Dataset (MXID), un ensemble de données de bonne qualité d'images radiographiques provenant de l'hôpital AOUINET de Tebessa, en Algérie, afin de faciliter diverses tâches d'imagerie médicale, notamment la compression, la classification et les applications d'apprentissage automatique. Les résultats de l'étude démontrent que les modèles basés sur l'apprentissage profond, en particulier les DCE, surpassent les autres techniques de compression en termes d'efficacité de compression des images et de rétention de la qualité, ce qui donne un PSNR de 46,78 dB.

L'objectif ultime de cette recherche à long terme est de contribuer au développement de techniques de compression d'images médicales efficaces et de haute qualité qui seront en mesure d'améliorer le traitement des données de soins de santé, facilitant ainsi une prestation de soins de santé plus efficace grâce à un meilleur stockage et à une meilleure transmission.

Les mots clé: Image médicale à rayons X, apprentissage profond, apprentissage automatique, compression d'image, ensemble de données d'imagerie médicale à rayons X (MXID).

الملخص

يلعب التصوير الطبي دورًا حيويًا في ممارسة الرعاية الصحية في العصر الحديث، حيث يتيح التشخيص والتخطيط للعلاج ومتابعة المريض. من بين طرائق التصوير المختلفة، يظل التصوير بالأشعة السينية أحد أكثرها شيوعًا نظرًا لفعاليته وتوافره على نطاق واسع وأهميته التشخيصية. ومع ذلك، فإن الزيادة المستمرة في حجم ودقة الصور الطبية، وخاصةً الأشعة السينية، تجلب تحديات تخزين البيانات ونقلها ومعالجتها ذات أهمية قصوى. وعادةً ما تؤدي هذه التحديات إلى مشاكل إدارية مكلفة ومعقدة، خاصةً في أنظمة الرعاية الصحية الكبيرة .

ولمواجهة هذه التحديات، تم استكشاف كل من تقنيات الضغط التقليدية والتقنيات الحديثة القائمة على التعلم الآلي. وتوفر التقنيات التقليدية، مثل JPEG وPNG ، ضغطًا أساسيًا ولكنها ليست مصممة خصيصًا لتلبية الاحتياجات المحددة للتصوير الطبي لأنها غير قادرة على الحفاظ على جودة التشخيص أو استغلال الأنماط الهيكلية الفريدة في الصور الطبية. ومع ذلك، أظهرت تقنيات التعلّم العميق، ولا سيما تقنيات التشفير التلقائي والشبكات العصبية التلافيفية (CNNs) والتشفير التلقائي المتغير(VAEs) ، إمكانات ممتازة في تحسين جودة الصورة ونسب الضغط، متجاوزةً بذلك قيود التقنيات التقليدية.

يبحث هذا البحث في إمكانات نماذج ضغط الصور القائمة على التعلّم العميق - مثل برامج الترميز التلقائي (AE) والترميز التلقائي التلافيفي العميق (DCAE) ونماذج الشبكات المترابطة الشبكية (CNNs) ونماذج VAEs - لتحقيق ضغط أفضل لصور الأشعة السينية الطبية. كما أنها تتضمن أيضًا خوارزميات التعلم الآلي مثل تحليل المكونات الرئيسية (PCA) وتجميع K-means لتحسين تخزين الصور ونقلها. بالإضافة إلى ذلك، تقدم هذه الدراسة مجموعة بيانات التصوير الطبي بالأشعة السينية (MXID)، وهي مجموعة بيانات عالية الجودة لصور الأشعة السينية من مستشفى العوينات في تبسة بالجزائر، لتسهيل مهام التصوير الطبي المختلفة، بما في ذلك الضغط والتصنيف وتطبيقات التعلم الآلي. توضح نتائج الدراسة أن النماذج القائمة على التعلم العميق، وتحديدًا DCAEs، تتفوق على تقنيات الضغط البديلة من حيث كفاءة ضغط الصور والاحتفاظ بالجودة مما يؤدي إلى PSNR قدره 46،78 ديسيبل.

الهدف النهائي من هذا البحث على المدى الطويل هو المساهمة في تطوير تقنيات ضغط الصور الطبية الفعالة وعالية الجودة التي ستكون قادرة على تحسين التعامل مع بيانات الرعاية الصحية، وبالتالي تسهيل تقديم الرعاية الصحية بشكل أكثر فعالية من خلال تحسين التخزين والنقل.

الكلمات المفتاحية: صورة طبية بالأشعة السينية، التعلم العميق، التعلم العميق، التعلم الآلي، ضغط الصور، مجموعة بيانات التصوير الطبي بالأشعة السينية (MXID).

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

MRI	Magnetic Resonance Imaging
CT	Computed Tomography
JPEG	Joint Photographic Experts Group
JPEG2000	Advanced Image Compression Standard by The Joint Photographic Experts Group
CNN	Convolutional Neural Network
RLE	Run-Length Encoding
HEIF	High Efficiency Image File Format
GIF	Graphics Interchange Format
PNG	Portable Network Graphics
BMP	Bitmap
Tiff	Tagged Image File Format
DICOM	Digital Imaging and Communications In Medicine
RAW	Raw Image Format
SVG	Scalable Vector Graphics
RGB	Red, Green, Blue
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
MPEG	Moving Picture Experts Group
ISO	International Organization for Standardization
VQ	Vector Quantization
SPIHT	Set Partitioning in Hierarchical Trees
LZW	Lempel-Ziv-Welch
ASCII	American Standard Code For Information Interchange
DPCM	Differential Pulse Code Modulation
ADPCM	Adaptive Differential Pulse Code Modulation
PET	Positron Emission Tomography
AI	Artificial Intelligence
RNN	Recurrent Neural Network
KNN	K-Nearest Neighbors
SVM	Support Vector Machine
PCA	Principal Component Analysis

ANN	Artificial Neural Network
RELU	Rectified Linear Unit
NLP	Natural Language Processing
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
ECG	Electrocardiogram
Tanh	Hyperbolic Tangent
AE	Autoencoder
VAE	Variational Autoencoder
DCAE	Deep Convolutional Autoencoder
CAE	Convolutional Autoencoder
GSA-PSO	Gravitational Search Algorithm-Particle Swarm Optimization
ROI	Region Of Interest
NON-ROI	Non-Region Of Interest
RNN-Conv	Recurrent Neural Network-Convolutional
CR	Compression Ratio
DT	Decision Tree
BPNN	Backpropagation Neural Network
GBA	Gradient Boosting Algorithm
RBFNN	Radial Basis Function Neural Network
OCT	Optical Coherence Tomography
MS-SSIM	Multi-Scale Structural Similarity Index
DIR	Deformation Image Registration
NON-DIR	Non-Deformation Image Registration
GBS	Graph Based Segmentation
FF-NNs	Feedforward Neural Networks
PSNR	Peak Signal-To-Noise Ratio
MSE	Mean Squared Error
SSIM	Structural Similarity Index
AAPCA	Adaptive Asymmetric Principal Component Analysis
BTRPCA	Block-To-Row Principal Component Analysis Algorithm
CoC	Coefficient Of Compression
WDR	Wide Dynamic Range
PCADTT	Principal Component Analysis with Discrete Tchebichef Transform
DTT	Discrete Tchebichef Transform

DA-RBM	Deep Autoencoder With Restricted Boltzmann Machine
DA-MBP	Deep Autoencoder With Multiple Back Propagation
Op-CNN	Optimized Convolutional Neural Network
CAE-NN	Convolutional Autoencoder Neural Networks
SRCNN	Super Resolution Convolutional Neural Network
GANs	Generative Adversarial Networks
BPG	Better Portable Graphics
UKSSL	Underlying Knowledge-Based Semi-Supervised Framework
SVD	Singular Value Decomposition
MSVD	Modified Singular Value Decomposition
MXID	Medical X-Ray Imaging Dataset
JPEG-XT	Joint Photographic Experts Group- Extended
OPT-JPEG-XT	Optimized JPEG-XT
SSR	Space Saving Rate
LAVIA-MXID	Laboratory of Vision and Artificial Intelligence-Medical X-Ray Imaging Dataset
NIH	National Institutes of Health
JSRT	Japanese Society of Radiological Technology
RSNA	Radiological Society of North America
ISIC	International Skin Imaging Collaboration
LIDC-IDRI	The Lung Image Database Consortium (LIDC) And Image Database Resource Initiative (IDRI)
CAD	Computer Aided Detection
GPU	Graphics Processing Unit
ReLU	Rectified Linear Unit
PReLU	Parametric Rectified Linear Unit Activation Layer

GENERAL INTRODUCTION

With technological advancements, the amount of visual data created in nations throughout the world has skyrocketed across a variety of industries, including communication, entertainment, scientific study, and even healthcare. However, this has resulted in further challenges with image processing, transmission, and storage. Image compression is a key technique for addressing these challenges, which makes it useful in a variety of industries. Compression methods are used to minimize image file size while preserving key features.

High-resolution X-rays, MRI's, CT scans images aid the medical field with their precise diagnosis and treatment formulation. Unfortunately, the sheer size of the datasets poses a significant logistical and computational problem regarding image management. In medical analysis, real-time image processing is crucial. Therefore, image avoidance preservation techniques, image compression, should be employed to address this problem. Image avoidance preservation allows for the safeguard of critical diagnostic details in aid of seamless network data transfer, unrivaled storage, and boost processing speed.

1.1. PROBLEMATIC

In today's digital era of healthcare, medical imaging has become an integral part of diagnosis work. Among the many imaging modalities, X-ray Imaging is one of the most accepted due to its cost, accessibility, and diagnostic utility. However, the exponential growth in the number and resolution of X-ray images, and efficient image transmission, especially in telemedicine systems has introduced significant challenges in terms of efficient image transmission, bandwidth consumption, and data storage. To solve these problems, several traditional compression methods such as JPEG, JPEG2000, PNG, and TIFF have been used according to their simplicity of implementation; however, these methods have limitations due to their use with medical images, and non-context-aware techniques using the same compression rules for all images regardless of structural or semantic content. By contrast, deep learning-based approaches—i.e., those employing autoencoders and convolutional neural networks (CNNs)—can learn task-specific optimal representations of medical images. Such models can learn to compress images adaptively by preserving diagnostic key features while reducing storage size drastically. Their use in the medical imaging community is an ongoing area of research, with performance, generalizability, and clinical acceptability issues. This thesis addresses this gap

by designing, training, and evaluating deep convolutional autoencoder models specifically for medical applications.

1.2. MOTIVATION

The purpose of this work is to create image compression technologies that are specifically designed for usage in the medical industry. Conventional algorithms such as JPEG and JPEG2000 are usually unsuitable for more difficult medical imaging jobs because medical images contain complex patterns and tiny features that help in diagnosis. The goal of this effort is to develop novel compression algorithms that use machine learning to improve compression efficiency and diagnostic feature preservation.

1.3.OBJECTIVES

The fundamental goal of this research is to investigate and create imaging approaches, particularly in medicine, that use machine learning for picture compression. More precisely, the research aims to achieve the following objectives:

- Determine if autoencoder or CNN-based algorithms are suitable for compressing medical photos.
- Evaluate the effectiveness of the methods proposed by looking at the compression ratio, the quality of the images, and the speed of the calculations.
- Try to find the practical uses and advantages of using machine learning techniques for compression in medical imaging.

1.4. SCOPE

This study focuses on exploring and assessing machine learning strategies for image compression in the field of medical imaging. It pertains to the creation of new compression methods based on autoencoders and CNNs, as well as experiments with publicly available medical images. The study includes grayscale medical images where the utmost image fidelity in integration with the existing medical imaging methods is sought.

1.5. OVERVIEW OF THE THESIS

The thesis is primarily on image compression, with a detailed survey of the existing techniques. Some of the key highlights of the research work included are:

1. Detailed review and classification of traditional and modern image compression methods.

2. Implementation and comparison of machine learning and deep learning-based compression models.

3. Construction and contribution of a new high-quality X-ray dataset (MXID) for medical imaging research.

4. Comparative study of various models with various performance measures across various datasets.

1.6. Organization of the thesis

Chapter One is the introductory chapter presents a detailed review of existing image compression techniques, both traditional and recent, and it provides background to the significance of image compression in medical imaging.

Chapter Two includes the machine learning and deep learning methods, and discussion regarding the strengths and weaknesses of each technique.

The Third Chapter is focused on machine learning techniques for compressing medical images. It presents a state-of-the-art overview of unsupervised learning approaches—such as Principal Component Analysis (PCA), K-means clustering, and various types of autoencoders (AE, DCAE, VAE)—and analyzes their applicability and performance when processing medical image data.

The Fourth Chapter discusses the datasets used in the study, including how the Medical X-ray Imaging Dataset (MXID) was created, its characteristics, and its importance. It also includes a comparison to existing public medical datasets.

Chapter Five is where methodology and experimental setup enter the scene, discussing the usage of various compression models (PCA, K-means, AE, DCAE, CNN, and VAE) and how the performance is measured, and discusses the experimental results, with a detailed analysis of the findings. The performance of all the models is compared, and the contributions of the study are evaluated in relation to prior work.

The General Conclusion and Perspectives summarizes the key findings of the thesis, highlights contributions to medical image compression research, and offers guidance on how future research can be enhanced to further aid compression techniques at the cost of diagnostic image quality.

2 IMAGE COMPRESSION

2.1 Introduction

The fast-forward development of technologies has enabled the multi-production, storing, and transmission of Visual Data in the multimedia, communication, healthcare, and scientific fields. The sources of High-Resolution photographs such as digital cameras, medical imaging and satellites have further increased the need for data management. Image compression caters to these problems by serving as a principal tool to maximize efficiency in storage, transmission, and processing.

Simply put, image compression is the action of shrinking pictures to a small file size without sacrificing an acceptable level of visual integrity to their intended use. Compression operations are capable of achieving big reductions in storage expense, bandwidth utilization, and processing time by locating and removing unnecessary or redundant data. Therefore, compression operations improve the handling of data, enable quick processing of pictures, and improve real-time processing capacity.

The primary objectives of image compression can be summarized as follows:

- Reducing Storage Requirements: Storage units have the capacity to hold more information without a hike in cost using compressed image data.
- Reducing Bandwidth Use: Compressed images take less bandwidth, thus can be transferred quicker and with less network congestion.
- Improving Transmission Speed: Compression makes it possible for visual information to be sent fast via various communication media, such as satellite communication, wireless connectivity, and web platforms.
- Maintenance of Visual Integrity: Maintaining the integrity of the image is a core function of compression algorithms, ensuring proper visual information for correct interpretation is preserved.
- Improving Real-Time Processing: Real-time image analysis programs like video conferencing, surveillance, and medical imaging are enabled by lightweight compression.

As image compression has evolved, it has benefited from advances in information theory, signal processing, and computer techniques. Through significant innovation, the industry has consistently increased picture compression efficiency, beginning with foundational approaches such as Huffman coding, run-length encoding (RLE) and proceeding to more contemporary standards such as JPEG2000, JPEG and High-Efficiency Video Coding.

2.2 Standard Image Formats

Image formats specify how pixels are saved, compressed, and displayed. Different formats are appropriate for a variety of uses, such as medical imaging, online browsing, and printing.

The following are some of the most popular standard picture formats presented in **Table 2-1**:

Table 2-1 Image File Formats and Its Extension

	Image Format	**Image Extension**	**Application**	**Benefits**	**Limitations**
Lossy Image Formats	**JPEG** (Joint Photographic Experts Group)	.jpg, .jpeg	Web images, digital artwork, screenshots	High quality, supports transparency	Larger file size than JPEG
	WEBP (Web Picture Format by Google)	.webp	Photography, web images	High compression, small file size	Loss of quality, artifacts at high compression
	HEIF (High-Efficiency Image Format)	.heif, heif	Printing, professional photography, medical imaging	High quality, multiple compression options	Large file size
Lossless Image Formats	**GIF** (Graphics Interchange Format)	.gif	Simple animations, web graphics	Supports animations, small file size	Limited to 256 colors
	PNG (Portable Network Graphics)	.png	Windows applications, icon design	High-quality images, simple format	Very large file size
	Tiff (Tagged Image File Format)	.tif, .tiff	Smartphone photos,	High efficiency,	Limited software compatibility

			modern image storage	better quality than JPEG	
Specialized Formats	**DICOM** (Digital Imaging and Communications in Medicine)	.dcm	Medical imaging (X-rays, CT scans, MRI)	Stores patient data with images	Requires specialized software
	RAW (Unprocessed Camera Image Format)	**Canon** → .cr2, .cr3 **Nikon** → .nef **Sony** → .arw **Adobe RAW** → .dng	Professional photography, post-processing	Maximum detail for editing	Very large file size, requires conversion
	SVG (Scalable Vector Graphics, XML-based format)	.svg	Logos, icons, scalable graphics	Infinite scalability, small file size	Not suitable for complex images like photos

2.3 Image Compression

In the Image compression task, the input image undergoes different steps to compress and reconstruct it. These steps start from input image to decompressed image for reducing its size while maintaining the quality. **Figure 2-1** represent the process of the image compression.

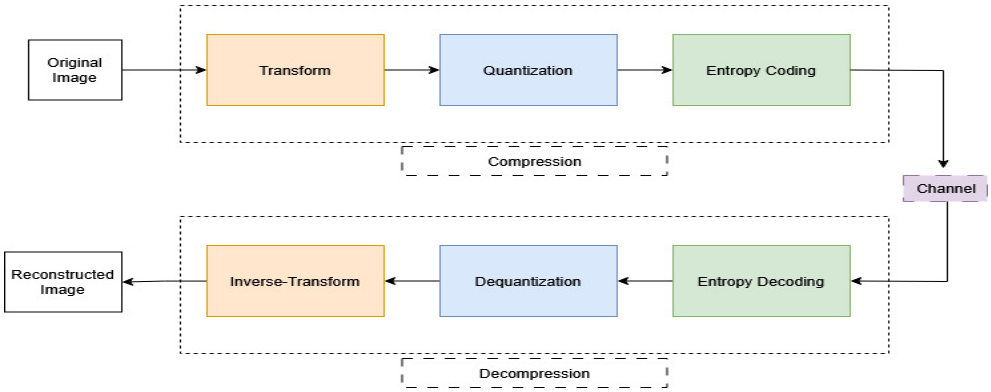


Figure 2-1 A General Compression Scheme

Following the steps in detail:

2.3.1 Input Image (Source File):

Starting from the input image, which refers to the original uncompressed image consisting of pixels that represents the visual information, stored in digital formats in wide color space (RGB or Grayscale).

2.3.2 Encoder (Compression):

The encoder serves to reduce the redundancies and irrelevant information by using several compression-based techniques (traditional or machine learning).

2.3.3 Encoded Image (Compressed File):

After the compression process, the image is compressed and represented in compact format with a smaller file size containing essential information with a slight quality reduction.

2.3.4 Decoder (Decompression):

In the decompression step, the decoder is responsible of compressed image reconstruction by reversing the compression process based on compression technique used either lossy or lossless.

2.3.5 Output Image (Decompressed File):

The reconstructed image is obtained after decoding its compressed format back into its approximate format; its quality depends on the applied compression technique.

2.4 Compression Techniques

The importance of image compression techniques is evident in various applications, including image transmission, storage, and processing. To tackle the challenge of reducing image size while preserving image quality, several image compression techniques have been developed. Some of the commonly utilized techniques are:

2.4.1 Image Compression Based Traditional Techniques

2.4.1.1 Lossy Compression

Lossy compression is a method of data compression that minimizes file sizes by approximating certain information, resulting in a reduction in fidelity but also achieving higher compression ratios. Lossy compression eliminates unnecessary data by reducing the file size and quality compared to the original one and it is reversible compression because after decompression the

original data could be recovered perfectly. Lossy compression entails relinquishing some degree of data integrity for the sake of enhancing efficiency, is typically suggested for situations in which reducing storage or transmission requirements takes precedence over maintaining the precise particulars of the original data. This is particularly pertinent in domains such as multimedia streaming, digital photography, and internet communication.

2.4.1.1.1 Fractal Coding

Hutchinson [1] and Barnsley [2] introduced Fractal coding, which is a lossy compression technique, it is a technique that employs the intricate mathematical principles of fractal geometry, offering an alternative solution for image compression. This technique locates self-similarity in images and reduces the amount of data desired for image reproduction [3]. Fractal image coding involves breaking down an image into affine transformations, each range block is encoded by selecting the domain block with the best match from the domain pool [4].

2.4.1.1.2 Transform Coding (DCT, DWT)

Transform coding is based on the principle that the pixels in an image exhibit a specific level of correlation with their adjacent pixels [5]. It is a crucial element of various image and video compression algorithms, including MPEG and JPEG. This technique involves converting data from the spatial domain, which consists of raw pixel values, to the frequency domain using mathematical operations. The Discrete Wavelet Transform (DWT) and the Discrete Cosine Transform (DCT) are commonly used methods for this type of transformation.

2.4.1.1.2.1 Discrete Cosine Transform (DCT) Coding

Discrete Cosine Transform (DCT) coding compresses an image or video into a smaller size while retaining significant details; it is a two-dimensional mathematical transform that converts real-valued input data from the spatial domain into the frequency domain [6]. It works by dividing the image into small blocks and determining how much of each color shade shows at various "brightness levels". This allows to the image to be saved more effectively, similar to the code that recreates the picture, resulting in reduced file sizes in formats such as JPEG and MPEG.

2.4.1.1.2.2 Discrete Wavelet Transform (DWT) Coding

Discrete Wavelet Transform (DWT) is a widely used coding technique in the digital image processing. The technique can be utilized for transforming and compressing image, as well as for steganography [7]. Wavelet transformations are built using fundamental, unalterable

components that serve a variety of purposes. These components are derived from a predetermined function, known as the mother wavelet, which remains constant [8].

2.4.1.1.3 JPEG Standards

The Joint Photographic Experts Group [9] developed the JPEG standard in 1992, the latter is a lossy image compression method based on the Discrete Cosine Transform to reduce image size while maintaining an acceptable level of quality. The JPEG standard provides three lossy encoding modes: progressive, sequential, and hierarchical, along with a single lossless encoding technique [10].

2.4.1.1.4 Jpeg-2000 Standards

Likewise, the Joint Photographic Experts Group (JPEG) developed JPEG 2000 and published it as an ISO standard in 2000. Unlike JPEG, which divides images into blocks measuring 8x8 pixels and compresses them, JPEG 2000 compresses each horizontal line into a signal that is then transformed into wavelets [11], achieving superior image quality and compression efficiency. Its versatility in handling various image types, including continuous-tone, bi-level, grayscale, and multi-component images, as well as its resilience to errors, which make it suitable for a wide range of applications. Although it has several advantages, its adoption has been limited by factors such as slower processing times and compatibility issues with current systems.

2.4.1.1.5 Vector Quantization (VQ)

The Vector quantization (VQ) is a compression technique that has been used in many different processing techniques such as image compression. It includes three different parts: code-book generation, encoding and decoding, In the process of code-book generation, images of interest are initially segmented into k-dimensional training vectors [12]. Then in the partitioning into regions, each of them is associated with one of the vectors. In the encoding or the quantization, each vector is assigned the closest vectors from the codebook. In the decoding procedure, using the same codebook, a codeword represents the original data vector from its corresponding vector for image reconstruction [13].

2.4.1.1.6 Set Partitioning in Hierarchical Trees (SPIHT)

The Set Partitioning in Hierarchical Trees (SPIHT) algorithm, initially proposed by Amir Said and William A. Pearlman [14] in 1996; based wavelet-based image compression, developed for image compression technology. SPIHT efficiently encodes images by dividing them into different levels of resolution based on their significance, the most significant sub-blocks are

then encoded and transmitted first. This hierarchical approach allows SPIHT to achieve efficient compression ratios and preserving image quality. With its ability to balance compression efficiency and computational complexity, this algorithm takes advantage as a popular choice for different applications, such as satellite imaging, and medical imaging.

2.4.1.2 Lossless Compression

Lossless compression technique reduces image file size with image quality preservation and without loss of details and original information. This method ensures the perfect image reconstruction from compressed file which requires high image fidelity and quality reconstruction, making them ideal for different applications especially medical one. Following section discusses the lossless techniques from Run-Length Encoding (RLE) to Arithmetic coding.

2.4.1.2.1 RUN-LENGTH ENCODING (RLE)

The Run-length encoding (RLE) is lossless compression technique; it compresses data by recognizing and decreasing runs, which are sequences of successive occurrences of the same alphabet or character. As a result, it assists to minimize the overall size of repeated symbol in the runs. In addition, the length of each run and the character in it should be recorded [15]. The fundamental concept of this method is to exchange consecutive repetitions of a symbol with a single instance of the symbol, accompanied by the count of its appearances [16].

2.4.1.2.2 HUFFMAN CODING (Entropy Coding)

Huffman coding is proposed in 1952 by Huffman, D [17], this technique uses variable-length codes to represent the input characters or symbols in order to evaluate the probability of the source characters; the less common character has a larger code than the most common one. To evaluate the occurrence rate of each symbol, Huffman coding commences by examining the input data for the frequency analysis stage. The subsequent step entails constructing the Huffman tree based on the frequency analysis, wherein each leaf node signifies a symbol. In the course of constructing the tree, less frequent symbols are allotted longer codewords, while more frequent symbols are allocated shorter codewords. The codewords are then assigned to their respective positions in the tree, guaranteeing that no codeword serves as a prefix to another codeword, thereby ensuring that the encoded data can be uniquely decoded. The input data are symbolically encoded using the generated Huffman tree and code words. The same Huffman tree is utilized in the decoding phase of the compressed data. The encoded data are traversed bit by bit, adhering to the path in the Huffman tree until a leaf node is reached. The symbol

represented by the leaf node is then output, and the process continues until all encoded data are decoded. Although it is a useful technique, it may not be as effective as more advanced compression methods such as run-length encoding or arithmetic coding, which consider symbol relationships in order to achieve greater compression ratios.

2.4.1.2.3 Lempel–Ziv–Welch (LZW) Compression

The compression system of Lempel–Ziv–Welch (LZW) [18], This compression starts with the dictionary creation that contains all potential byte or character sequences. Then it operates by sequentially reading input data and constructing a dictionary containing byte sequences identified in the input stream. As it processes the data, it searches for the longest sequence in the dictionary that is identical to the current input and produces the index of that sequence in the dictionary. When a novel sequence is detected but not yet present in the dictionary, it is immediately added to it, and its corresponding index is assigned. The dictionary grows in size as additional sequences are encountered. The compressed data is composed of a succession of indices that direct attention to entries in the dictionary. Finally, during decompression, the compressed data is scanned, and the corresponding sequences in the dictionary are retrieved. As each sequence is retrieved, it is incorporated into the dictionary for future reference.

2.4.1.2.4 Shannon–Fano Coding

The Shannon-Fano algorithm, developed by Claude Shannon and Robert Fano in 1949, is a lossless compression technique [19]. It uses variable-length code words, where characters that are more frequent are assigned shorter codes, and less frequent characters are represented by longer codes [20]. This technique starts with building probability table, then Organize the table by frequency, with the highest frequency character appearing at the top, after that is splitting the table into two sets that the combined probability of both groups is as close as possible, maintaining the original meaning and context. Next step is assigning the value of “1” in the right sets and the “0” in the left sets, and finally undertaking the two previous steps until successfully dividing all the symbols in a separated group.

2.4.1.2.5 Arithmetic Coding

Arithmetic coding is a type of data compression technique that utilizes entropic coding and is employed in lossless data compression. This method typically employs a set number of bits per character, similar to the ASCII code. By converting a string to arithmetic encoding, frequently used characters are assigned fewer bits, while fewer common characters are assigned more bits, thus reducing the overall bit count.

2.5 Comparison of Traditional Techniques

Traditional compression techniques comparison emphasizes the differences between various algorithms lossy and lossless. For the lossy compression, it reduces the file size and can affect the quality by losing information in the reconstructed images which is often acceptable. On the other hand, lossless techniques ensure that the reconstructed image quality well preserved where image quality maintain take precedence. **Table 2-2** presents various traditional compression methods with its type, applications, strengths, and weakness for each, while lossy techniques prioritize efficiency and lossless techniques focus on accuracy.

Table 2-2 Traditional Image Compression comparison

Techniques	Compression Type	Applications	Strengths	Weakness
Run-Length Encoding (RLE)	Lossless	Used in binary images and bitmaps based repetitive image patterns	- Simple in terms of implementation. - Robust for identical pixels images with long runs.	- Less effective for high complex image variation. - Limited compression ratio.
JPEG (DCT-based)	Lossy	Utilized in photographic digital and online images	- Natural images with high compression ratio values. - Widely supported. - Customizable balance quality/ compression.	- Blurriness (blocks effect) at high compression ratios with artifact images. - inconvenience with sharpened and edged images.
Huffman Coding	Lossless	Lossless compression technique applied on PNG, and GIF images	- Highly preserves original images details. - high performance with imbalanced probability distributions.	- Less efficiency for image with uniform distributions. - symbol probabilities knowledge required.
Lempel-Ziv-Welch (LZW)	Lossless	Lossless compression for GIFs, and TIFFs formats	- No requirements for image distribution prior knowledge.	- Less efficient with complex images. Inefficient for complex images with a moderate compression ratio.
Fractal Coding	Lossy	Compression of natural images (e.g., landscapes)	- High compression ratios. - Independent resolution (scalable images).	- Exceedingly slow encoding process. - Inappropriate for all image types.
Discrete Cosine Transform (DCT)	Lossless	Widely used in JPEG compression	- Low complexity	- Elevated compression ratio with artifacts. - For text images or sharp edges is not suitable.

Predictive Coding (DPCM/ADPCM)	Lossy	Lossless/lossy compression in image processing and speech	- Simple implementation. - Appropriate for predictable data sequences.	- Inappropriate for high-frequency variations. - Not suitable for images with details.
Arithmetic Coding	Lossless	Lossless compression	- compared to Huffman coding, it gives better compression results. - More flexible.	- Computationally complex. - Time-consuming compression.
Vector Quantization (VQ)	Lossy	Used in video coding image and low-bit-rate	- Applied in low-bit-rate with acceptable compression ratio. - Redundancy reduction.	- Possibility for image distortion. - Complex generation of the codebook.
Wavelet Transform	Lossy	medical imaging, satellite imagery, JPEG2000	- representation of multiresolution images. - Better handling of sharp edges than DCT. - efficient for complex images.	- Complex regarding computation. - Large memory requirements.
SHANON-FANO CODING	Lossless	Early lossless coding algorithm, mostly of historical interest	- easy to understand. - Conceptually simple. - Basis for more advanced algorithms like Huffman coding.	- Less efficient compared to Huffman coding. - Optimal prefix codes are not always optimal.

Next section, present medical imaging modalities which comes with large size requiring file size reduction while preserving its quality for an optimized transmission and storage.

2.6 Medical Image Modalities

Medical imaging modalities used to visualize the different representations and details of the human body using a specific medical image equipment and technologies to help radiologists and clinicians to diagnose several diseases. Each technique differs in term of information representation, details of the specific body part for effective treatment. These modalities are: X-ray, Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) scans, Computed Tomography (CT), and ultrasound, as illustrated in **Figure 2-2.**

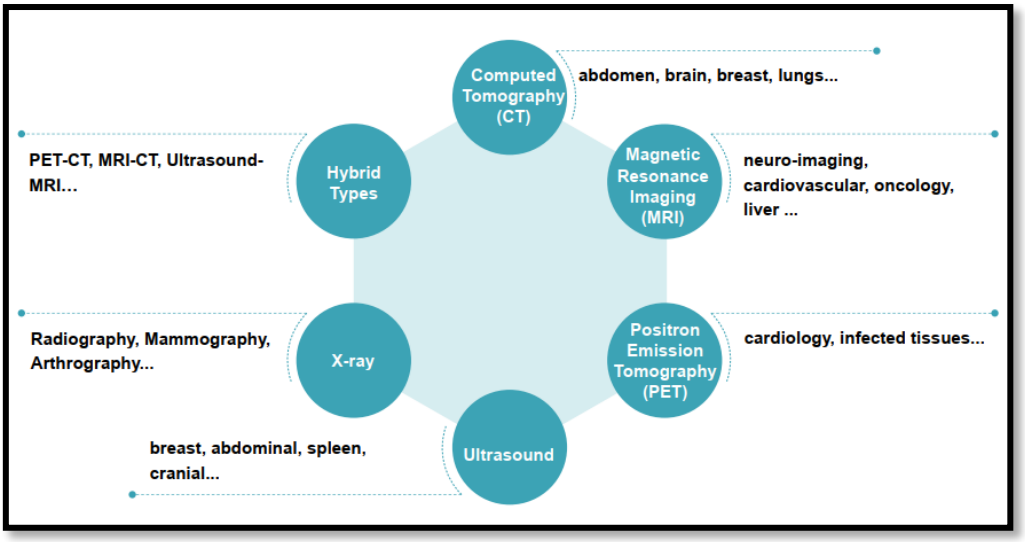


Figure 2-2 Various Medical Imaging Techniques and Their applications *[21]*

One of the most used technique is x-ray imaging, it generates images of dense structures to diagnose a range of diseases, ranging from chest problems to bone fractures and dental issues by employing ionizing radiation to capture these images. Furthermore, Magnetic Resonance Imaging (MRI) produces images of a high-resolution quality for capturing some minor tissues such as muscles, brain, and ligaments using radio waves. On the other hand, Computed Tomography (CT) displays complex structures, soft tissues and organs by combining x-ray images to create cross-sectional view from different angles. For the Positron Emission Tomography (PET) scans uses gamma rays from injected radioactive tracers for displaying bones images and organs scans. Finally, Ultrasound uses sound waves to generate abdominal images, bones, and soft tissues as described in **Table 2-3**. Each technique generates and visualize specific organs, soft tissues, bones, tumors, and others providing accurate diagnosis, and treatment planning.

Table 2-3 Medical Image Comparison *[21]*

Techniques	Anatomy	Functionality	Advantages	Disadvantages
X-ray Radiography	Chest	used to produce images of bones, tumors, and other dense matter.	-The procedure is non-invasive, rapid, and painless. -The insertion of an arterial catheter and guidewire is not required.	-No real-time information. -Cannot be done without contrast (allergy, toxicity).
Magnetic resonance imaging (MRI)	Abdomen Brain Heart	used to produce detailed images of organs, soft tissues, bones, ligaments, and cartilage.	-There are no ionizing rays. -The spatial resolution is excellent.	-Sensitivity is relatively low. -The scan and post-processing times are long. comparatively costly

Nuclear medicine imaging including positron emission tomography (PET)	Brain Heart	-A radioactive 'tracer 'is injected, inhaled, or swallowed. -The scanner uses the gamma rays emitted by this material to display images of bones and organs.	-Can aid in the diagnosis, treatment, or prediction of a variety of conditions. -Can determine how far a cancer has spread and how well treatment is working.	-Some people may experience allergic or injection. -site reactions as a result of radioactive materials.
Computed tomography (CT scans)	Abdomen Brain Heart	Multiple X-rays are used to create cross-sectional layers that show detailed images of bones, organs, tissues, and tumors inside the body.	-Immediate and painless. -aid diagnosis and treatment of a wider range of conditions than a standard X-ray.	-utilizes higher doses of radiation than a X-ray. -Anesthesia is required for some procedures.
Ultrasound	Heart	-High-frequency soundwaves are used to generate moving images of the inside of the body on a screen, including organs, soft tissues, bones, and an unborn baby.	-High-resolution images. -Real-time information	-Time consuming. -Veins can't be seen in their whole.

Medical imaging modalities plays a vital role in diseases diagnosis, on the other hand confronting challenges in transmission and storage due to its large file size. The application of the different compression techniques reduces image size while preserving its quality. Thus, next section introduces the compression techniques in details.

2.7 Conclusion

Image compression has become crucial for digital images' size reduction while maintaining image quality; it is a critical tool for data storage optimization, faster transmission, and reconstruction for a wide range of applications. Throughout this chapter, various traditional compression techniques have been explored including lossy and lossless methods, their mechanism, suitability for different types of images, and a comparison based on compression type, application, and strengths and weaknesses of each.

As the demand of high image resolution continue to grow, especially in medical imaging domain such as CT, X-rays, and MRIs scans in terms of volume and complexity, selection of appropriate compression technique is required for a balance of compressed image size and high-quality reconstructed image. Machine learning and deep learning techniques revolutionized image compression field for more efficient compression, offering adaptability in handling larger datasets, and minimizing quality loss for accurate diagnosis and clinical interpretation.

Part 2

Machine Learning Techniques

3 MACHINE LEARNING TECHNIQUES

3.1 Introduction

Machine learning is an essential subfield of artificial intelligence (AI), it focuses on creating algorithms that allow computers to learn and adapt without human involvement. Machine learning models constantly improve their performance by analyzing earlier data and experiences, resulting in increased efficiency over time.

This chapter looks into machine learning approaches, such as unsupervised learning, supervised learning, and deep learning approaches that will be used in our research. With their ongoing development, both machine learning and deep learning technologies have come to be extremely efficient in solving challenging problems in a wide range of fields. Various methods, such as support vector machines, decision trees, and clustering algorithms, come under the category of machine learning. All of these techniques are designed to identify patterns, categorize information, and predict from input features. All these approaches come under unsupervised, supervised, and reinforcement learning. These are extremely helpful when handling structured data.

Deep learning is a subset of broad machine learning; it uses multi-layered artificial neural networks to handle large volumes of unstructured data such as photos, text, and audio. Recurrent Neural Networks (RNNs) and Convolutional neural networks (CNNs) excel in image classification, natural language processing, and time-series forecasting because they can detect subtle hierarchical patterns in data.

This chapter provides a thorough examination of deep learning and machine learning techniques, comparing their applications, advantages, and limitations.

3.2 Machine learning techniques

Machine learning techniques are categorized into three main types: unsupervised learning, supervised learning, and semi-supervised learning. This research introduces and explains the main algorithms of each technique starting by supervised leaning techniques to deep learning models, as stated in **Figure 3-1**.

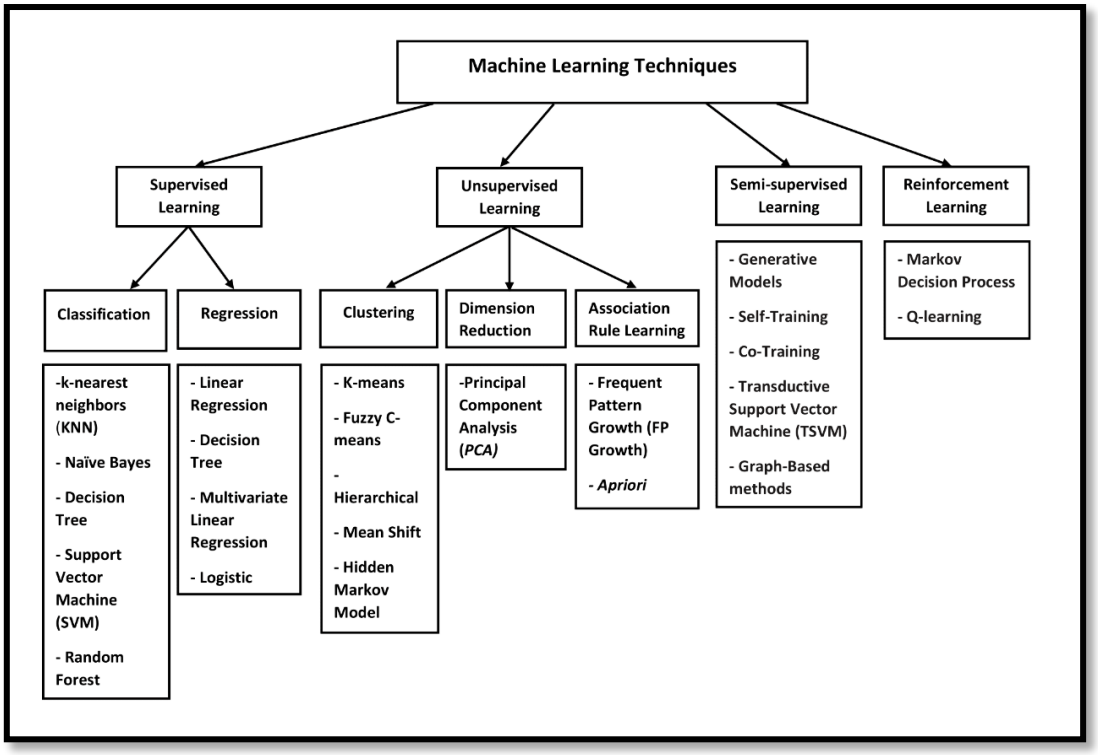


Figure 3-1 Different machine learning techniques *[21]*

3.2.1 Supervised learning

Supervised learning is a subfield machine learning in which models are modeled on labeled data, its aim is that algorithms minimizes the error between its predictions and true outputs. It is grouped into regression and classification algorithms where a known set of outputs is used to guide the learning process.

3.2.1.1 Classification

In supervised learning, classification is a used to train models on labeled data into predefined classes, it aims to learn features to distinguish one class from others. Its types are binary classification involving two classes in which the model is trained to differentiate between two categories, also Multiclass classification is applied to a model that predicts more than two classes, and multilabel classification in which data points belong to multiple classes simultaneously.

3.2.1.1.1 K-Nearest Neighbors (KNN)

K-nearest neighbors (KNN) is a sort of a supervised classification technique, it uses the similarity principle to classify data based on the defined ‘K’ parameter; this technique calculates the distance using Euclidean, Manhattan, or Minkowski distances as a similarity measure between target and training datasets’ points. For each new point x, the process begins

by identifying its K-nearest neighbors from the training data before making a prediction, then the new assigned point will be classified as the most common class among its k-nearest neighbors. There exist several strategies to choose K value; the simple one is running the algorithm repeatedly with different K values to choose the one with higher performance [22]. The KNN presents a simple and effective method while sensitive in terms of 'K' value choice, small 'K' value can lead to a higher sensitivity to noise, while large 'K' value is more robust to noise.

For instance, Euclidean distance metrics is defined as follows:

$$d(x,\ x_i) = \sqrt{\sum_{j=1}^{n}\left(x_j - x_{i,j}\right)^2} \qquad (3\text{-}1)$$

$$knn = \{k - min(d(x, x_i))\} \qquad (3\text{-}2)$$

Where: x_j and $x_{i,j}$ represent the j-the features of points x and x_i, respectively.

3.2.1.1.2 Support Vector Machines (SVM)

Support Vector Machine (SVM) is a supervised nonlinear binary classification learning model, it was introduced by Vanpik in 1992. SVM aims to identify the optimal hyper plane that divides data points in a high dimensional space. Also, serves to optimize the margin between different classes, making it effective for both linearly separable and non-linearly separable data. SVM's application are different in several tasks, such as in medical diagnosis, face recognition, text classification, financial forecasting.

3.2.1.1.3 Naive Bayes

The Naive Bayes classifier is a probabilistic model that applies Bayes' theorem; it performs effectively with limited data, because it assumes conditional independence between features given the class label. This classifier frequently demonstrates remarkable effectiveness and is extensively utilized due to its ability to surpass more complex classification techniques in performance [23]. Variants of Naive Bayes are Gaussian, Bernoulli, and Multinomial Naive Bayes. Calculating posterior probability as follows:

$$P(C \mid x) = P(x \mid C)\ .P(C)\ /\ P(x) \qquad 3\text{-}3)$$

$$P(C|x) = P(C)\ \prod_{i=1}^{n} P(x_i \mid C) \qquad (3\text{-}4)$$

Where: $P(C \mid x)$ is the **conditional probability.**

$P(x \mid C)$ is the **likelihood.**

$P(C)$ is the **prior probability** of class C.

$P(x)$ is the **predictor's prior probability.**

Naive Bayes classification involves calculating $P(C \mid x)$ for every class C and assigning x to the class with the highest conditional probability.

3.2.1.1.4 Decision Tree

Decision tree algorithm is also a supervised machine learning technique employed for both classification (classification tree) and regression (regression tree). It partitions data into subsets based on attribute's value, starting from root node and at each node, the tree recursively divides the data according to a feature until the stopping criteria are met. Each branch represents the result of a decision, while each leaf node indicates a class label or predicted value. Using two well-known criteria or metrics such as Entropy and Gini to split and determine attribute's value for each node defined as follows, recursively:

$$Gini = 1 - \sum_{i=1}^{n} p_i^2 \qquad (3\text{-}5)$$

$$Entropy = -\sum_{i=1}^{n} p_i \, log_2(p_i) \qquad 3\text{-}6)$$

Where p_i is the probability of class i at that node.

3.2.1.1.5 Random Forest

Random Forest represents an aggregation of multiple classifiers introduced by Beirman. [24] To build or constructs a single classifier [25], it serves as an ensemble learning technique for classification, proving highly effective on complex datasets. By building this classifier, its performance can be better than individual classifier does [26], to enhance model's accuracy and robustness this model builds multiple decision trees then combines their outputs each tree is selected randomly with replacement (bagging), which increases generalization and reduces overfitting. In this method increased number of trees no longer significantly to improve the performance according to Beirman.

3.2.1.2 Regression

Regression is a technique within supervised learning. that predict outcomes based on input features or variables and to model the relationship between them. Most common regression technique is introduced in the next section.

3.2.1.2.1 Linear Regression

Linear regression is a supervised machine learning model and a basic statistical technique. It aims to model the relationship between one or more independent variables and a continuous dependent variable. Its equation is as follows:

$$y = \beta_0 + \beta_1 x_1 \qquad (3\text{-}7)$$

Where: y is the dependent variable

β_0 is the predicted value when x value is 0

β_1 is coefficient for each feature

x_1 is the independent feature

3.2.2 Unsupervised learning

In contrast to supervised learning, unsupervised learning is a type of machine learning that focuses on training a model without labeled data, and without predefined classes. It includes clustering and dimensionality reduction.

3.2.2.1 Clustering

Clustering is the process of dividing data into groups, with each group containing similar characteristics based on a similarity measure and by minimizing the distance between points and the cluster center. Most common clustering algorithms includes k-means (used in our study) which is discussed in the next section.

3.2.2.1.1 K-Means

K-Means serves as an unsupervised learning algorithm based on data partitioning into Into a specified number of clusters "K", determined by the similarity of each data point to another using Euclidian distance that minimizes the total squared distance between the cluster centers and the data points in each group. defined as follows:

$$min \sum_{i=1}^{k} \sum_{x \in C_i} \| x - \mu_i \|^2 \qquad (3\text{-}8)$$

Where: K is the number of clusters.

x represents each data point.

C_i is the i-th cluster.

μ_i is the centroid of the *i*-th cluster.

$\| x - \mu_i \|^2$ is the squared Euclidean distance.

K-means steps are as follows:

- Selection of K initial centroids.
- Each data point is grouped with the nearest cluster center point based on distance measure.
- The centers are updated by recalculating the average of all points assigned to each of the K clusters.
- This process continues until convergence or until a predefined number of iterations is reached.

3.2.2.2 Dimensionality Reduction

Dimensionality reduction refers to the process of converting data with many features into a space with fewer dimensions, while retaining the most important information. A widely adopted method for this purpose is Principal Component Analysis (PCA), which is utilized in our study and described in the following section.

3.2.2.2.1 Principal Components Analysis (PCA)

Principal Component Analysis (PCA) is a linear data dimension reduction method with the objective of preserving the most significant information in a data set and converting it into a new set of variables called principal components. PCA makes data visualization more efficient by representing it in the lower-dimensional space and reducing the data by lowering its complexity.

3.2.2.2.2 Autoencoders

An autoencoder, originally proposed by LeCun in his doctoral thesis [27], It consists of two main parts: an encoder that converts input data into a compressed latent space representation, and a decoder that maps back the input from this representation. These two components work together to compress data into a compact form without compromising its quality, and it is therefore applicable for successful image transmission and storage. The use of autoencoders in image compression provides a foundation for the attainment of successful input data representations [28]. **Figure 3-2** illustrates the architecture of an autoencoder.

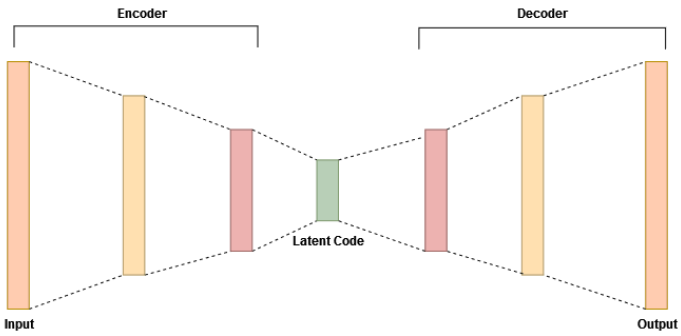


Figure 3-2 Architecture of an autoencoder

3.2.3 Semi- Supervised Learning

Semi-supervised learning is a machine learning approach that combines the strengths of supervised and unsupervised learning by training on a large set of unlabeled data alongside a smaller set of labeled data. Aiming to address the challenge of scarce labeled data, which is often costly and time-consuming to acquire, this approach helps reduce the effort and expense involved in building high-performing models. It is particularly beneficial for labeling large datasets required in supervised learning.

3.2.4 Reinforcement learning

Another type of machine learning is reinforcement learning (RL), the latter is about interacting with the environment using an agent that takes different actions and receive feedbacks as penalties or rewards. These outcomes improve the decision-making task and reinforce agent's knowledge with optimal strategies to attain lasting goals through this process. This learning technique is a powerful tool for different intelligent systems' development such as robotics, and autonomous vehicles.

3.2.5 Deep learning

Deep learning is a subset of machine learning that concentrates on algorithms built around artificial neural networks with different multiple layers, mimicking the human's brain structure and function, it extracts high-level representations from raw input data, deep learning techniques have evolved over decades; driven by progress in computational power, the availability of large datasets, and advancements in training algorithms.

3.2.5.1 Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a type of deep learning model and a specialized subclass of Artificial Neural Networks (ANNs); CNNs were developed for many purposes,

such as, image segmentation, visual data processing and evaluation, image classification, and object detection. It consists of different multiple layers, starting from convolutional layers which applies convolution operations to detect features (textures, edges, and patterns), and each convolutional layer learn to extract complex features using RELU (Rectified Linear Unit) function. Then, pooling layers that reduces spatial dimensions (down sampling) while maintaining crucial features using max pooling and average pooling methods. Finally, fully connected layers (dense layers), which interpret learned features and map them to output classes, like in classification tasks. CNN's capacity performs in features extraction from raw pixel data; as presented in **Figure 3-3**.

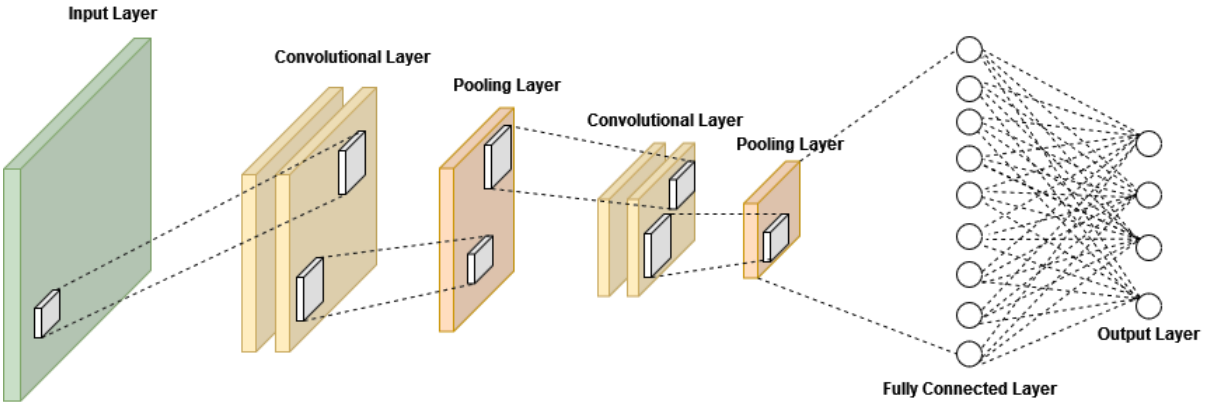


Figure 3-3 The Architecture of CNN's Model *[29]*

3.2.5.2 Recurrent neural networks (RNN)

Recurrent neural networks (RNNs) are a class of neural networks designed to detect patterns in a sequence of data, where the order of data points is important [30]. Unlike traditional feed forward neural networks, RNNs utilize recurrent connections to preserve information from previous inputs via hidden states, which makes them well suited for tasks like medical data analysis, time-series data, and natural language processing (NLP). However, standard RNNs struggles from vanishing gradient problems, making process of learning long-term dependencies difficult. To tackle this issue, advanced variants such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) are employed; enabling learning that is more efficient over long sequences. RNNs applications are varied such as in the text generation tasks, speech-to-text-systems and medical sequence analysis (e.g., ECG signal classification).

3.2.5.3 Deep Generative Models

The core concept of generative models is to learn the underlying probabilistic distribution of a data class, enabling the generation of new, similar data instances. [31]. Generative models can generally be classified into two primary categories:

- Cost function-based models: Generative adversarial networks, adversarial autoencoders.
- Energy-based models: Boltzmann Machines, Restricted Boltzmann Machines, Deep Boltzmann Machines, Deep Belief Networks, Generative Stochastic Networks, and Binary Boltzmann Machines.

3.2.5.4 The Activation functions in deep learning models

The Activation functions play an important role in deep learning models, controlling how neurons compute and transmit information. They introduce nonlinearity, allowing neural networks to learn sophisticated patterns rather than linear transformations.

1. **Linear Activation Function**

Example:

$$f(x) = ax \qquad (3\text{-}9)$$

- The network has no non-linearity, resulting in straightforward linear regression.
- Used in regression models, not deep networks.

2. **Non-Linear Activation Functions**

   Non-linearity is required for deep networks to learn complicated patterns. These can be further divided as:

a) **Sigmoid-Based Functions**

- **Sigmoid**:

$$f(x) = \frac{1}{1+e^{-x}} \qquad (3\text{-}10)$$

→ Used for binary classification.

- **Tanh (Hyperbolic Tangent)**:

$$f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad (3\text{-}11)$$

→ Zero-centered output.

b) **ReLU-Based Functions** (Piecewise Linear)

- **ReLU (Rectified Linear Unit)**:

$$f(x) = \max(0, x) \qquad (3\text{-}12)$$

→ Fast and commonly used in deep networks.

- **Leaky ReLU**: Allows small negative values to prevent dead neurons.
- **Parametric ReLU (PReLU)**: Learns the negative slope.
- **Exponential Linear Unit (ELU)**: Smoothens ReLU to improve learning.

c) **Probability-Based Functions**

- **Softmax**: Converts logit values to probabilities for multi-class categorization.

Specialized Classification of Activation Functions

In addition to Linear vs. Non-Linear, several scholars classify activation functions according to:

- **Monotonic vs. non-monotonic:** Sigmoid and Tanh are monotonic, whereas ReLU is not.
- **Unbounded vs. Bounded:** ReLU is unbounded, but Sigmoid and Tanh are bound.
- **Continuous vs. Discontinuous:** Most are continuous, but Hard Threshold functions are discontinuous.

3.3 Conclusion

Machine learning is today a common technology on a range of applications. This chapter examined an overview of some of the most significant machine learning and deep learning methods; which have revolutionized several domains by learning patterns from data and making intelligent decisions. We have discussed machine learning techniques encompassing supervised learning, unsupervised learning in details, each introduced with its algorithms, which has provided foundational methods for tasks like clustering, classification, and regression. Then, we discussed deep learning and its different models' architectures such as CNNs, RNNs, and deep generative models. However, deep learning and machine learning models will continue to dive innovation with the ongoing research, shaping artificial intelligence's future across industries.

4 MACHINE LEARNING FOR MEDICAL IMAGE COMPRESSION (STATE OF THE ART)

4.1 Introduction

Compressed Medical Image Compressing medical images is a critical field of medicine, the reflection of growing demand for successful storing, transmission, and processing good-quality medical images. Some traditional compression techniques such as JPEG2000 and SPIHT are now widely employed methods; nevertheless, they will certainly struggle to achieve high compression ratios without losing good diagnostic data.

Medical image compression using machine learning has been classified into three categories: supervised learning, unsupervised learning, and deep learning. In the field of medical image compression, a number of approaches are addressed, with a focus on the strengths and limitations of each supervised method that utilizes labeled data to enhance the effectiveness of compression algorithms.

This chapter offers a detailed survey of recent advancements in medical image compression using machine learning, focusing on supervised and unsupervised approaches. PCA and clustering methods have been noted to emphasize their usage in data structuring and dimension reduction. Autoencoders (AE), Convolutional Neural Networks (CNN), Deep Convolutional Autoencoders (DCAE), and Variational Autoencoders (VAE) have been vital in medical image compression in deep learning. Based on the review of the key studies and new releases, we discuss the merits, demerits, and areas of study in the existing methods. In totality, the existing chapter provides insightful remarks regarding how machine learning is poised to reshape the future of medical image compression based on its different techniques and practical applications.

4.2 Medical Image Compression based Machine Learning Techniques

Medical image compression has been well-researched using supervised as well as unsupervised machine learning techniques. Supervised approaches tend to utilize labeled data with models at the cost of diagnostic quality to reduce compression. PCA and K-means clustering are examples of such unsupervised approaches that employ natural data patterns to compress without needing labeled input. In order to present the efficiency, application, and limitations of different methods for compressing medical images, various studies have been presented in this section.

4.2.1 Supervised Learning Techniques for Medical Image Compression

As part of this investigation, we carried out this research with a thorough survey [21], which covered several techniques, such as Support Vector Machines (SVMs), Decision Trees, and different deep learning architectures. The use of supervised learning methods for medical image compression depends on training the models on labeled datasets that optimize the reconstruction process. Nevertheless, the survey also noted a number of drawbacks, including high processing demands and restricted applicability to other medical imaging modalities. As explained in [21], these realizations influenced the later development of our CAE-

based compression strategies.

Medical image compression, which includes multiple methods of compression, including the widely used lossless compression technique for different medical picture modalities, is one of the most important processing tasks in many medical imaging applications. Due to its size, the medical image needs to be compressed while maintaining the regions of interest using various machine learning and deep learning techniques. In [32], a hybrid technique that combines lossless and lossy compression, by using recurrent neural networks and the discrete walvet transform on brain images that are identical to CT and MRI images. Therefore, RNNs are enhanced by integrating the Gravitational Search Algorithm with Particle Swarm Optimization (GSA-PSO) to reduce losses during lossy compression; the non-ROI and ROI portions of the brain images were compressed separately using optimized RNN and, DWT, which together produced a higher PSNR than the previous techniques.

In various deep learning applications, the widely used U-net segmentation neural network is employed to segment medical images and define the region of interest (ROI) within the image, and then the information is compressed using an auto-encoder neural network, that uses a convolutional neural network [33]. The findings suggest that the proposed solution effectively preserves most of the image's visual information, particularly in the ROI, while reducing the required storage size.Additionally, [34] offers a method for convolutional recurrent neural networks (RNN-Conv) to compress x-ray pictures. With high compression ratios (CR) of 40 or more, the suggested method works better than any other method for low-resolution photos; nevertheless, noise from an alternative image capture technology causes a little degradation at low compression. A medical X-ray image compression system was also tested using machine learning algorithms, including decision trees (DT), support vector machines (SVM), back-propagation neural networks (BPNN), gradient boosting algorithm (GBA), logistic regression (LR), and radial basis function neural network (RBFNN), the effectiveness of the proposed discrete cosine transform (DCT)-based image compression with nine compression ratios was evaluated. RBFNN produced the best overall result.

Moreover, CNNs [35] are used to compress medical images in three steps: compression, segmenting the compressed sections, and data preprocessing to minimize OCT image noise. This facilitates the effective training of CNNs for both reconstruction and compression. The preprocessing module's semantic segmentation images were used to train these two networks concurrently. The test results demonstrate that the proposed framework outperforms existing methods in terms of MS-SSIM and visualization, with the difference being especially noticeable at higher compression ratios. Lossless and near-lossless techniques, as proposed in [36] apply a prediction-based compression method using the CLEF Med 2009 dataset for medical images. These techniques retain image quality in the diagnostically important region (DIR) even after compression by separating the image into foreground (DIR) and background (NON-DIR) regions

through a graph-based segmentation (GBS) technique. For the DIR parts, near-lossless and lossless compression methods are used, while a lossy compression technique is applied to the non-DIR parts. During the compression and decompression stages, feed-forward neural networks (FF-NNs) (Predictors) are trained using a Gravitational Search Algorithm and a Practical Swarm Algorithm to predict current pixel values based on neighboring pixels. Experimental results show that the gravitational search algorithm outperforms the particle swarm algorithm in near-lossless prediction. Moreover, the lossless compression approach increases the PSNR and reduces the CR values compared to the near-lossless approach. In contrast, the gravitational search algorithm demonstrates better performance in near-lossless prediction than the particle swarm algorithm. A detailed comparison of the different machine learning techniques is provided in **Table 4-1**.

Table 4-1 Comparative Study of the Different Machine Learning Techniques for Medical Image Compression *[21]*

PURPOSE OF THE	Authors(s) (year)	Domain of Application	Information from the Database	Learning Approach	ALGORITHM	Outcomes	Ref
Compression	S.Sabbaravapu, S.Gottapu et al., [2021]	- Compression of medical images including MRIs and CT scans	-Collected from MedPix Database 2020	Supervised learning	- DWT and RNNs	-Medical image compression performance evaluation: **PSNR:** 35.2362 dB **CR:** 23.3428 **MSE:** 20.9673 **SS (Space saving):** 95.7095	[37]
	K. Dmililer., [2021]	-X-ray image compression	-80 X-ray photographs collected from the Famagusta General Hospital's Radiology Department in Cyprus, comprising X-ray images for various regions of the body.	Supervised learning	-DT, SVM, GBA, RBFNN, LR, BPNN	**Accuracy:** -**Scenario 1:** - DT and SVM: 100% -GBA and SVM: 89% -LR, RBFNN, and SVM: 78% when accuracy rate of 67% is considered. **-Scenario 2:** -RBFNN: 100% and 89% -GBA, RBFNN, and SVM: 78% - All of the algorithms produced the findings when 67% accuracy was considered.	[38]
	A.Khaleel, N. Zahri, M.Ahmad., [2021]	Medical image compression	-CLEF MED 2009 dataset	Supervised learning	Hybrid compression method based on CNN	-**performance measures:** **-for α=0.07:** -Average **CR**: 9.68 -**Average PSNR** : ROI : 49.07 dB NROI : 46.74 dB Entire Image: 47.79 dB **-for α=0.08:** -Average **CR** : 6.73 -**Average PSNR** :	[33]

					ROI : 54.27 dB. NROI : 45.12 dB Entire Image: 49.06 dB	
P.Guo, D.Li, X.Li., [2020]	-Optical coherence tomography image compression	-Open-source dataset from Duke University's Vision and Image Processing (VIP) Laboratory.	Supervised learning	-CNNs	-Similarity index **MS-SSIM**: CR=80: 0.985 CR=40: 0.989 CR=20: 0.992 CR=10: 0.994 computation time:0.015 s/image	[35]
A.Sushmit, S.Zaman, A.Humayun et al., [2019]	-X-Ray Image Compression	-NIH Clinical Center's Chest X-Ray8 Dataset	Supervised learning	-RNN-Conv	- **Quality evaluation metrics: Compression ratio (CR)= 8:** For image size **128*128:** -SSIM: 0.9579 -PSNR (dB): 35.9325 For image size **256*256:** -SSIM: 0.9509 -PSNR (dB): 34.8701	[34]
M. Ayoobkhan, E. Chikkannan, et al., [2018]	Medical Image Compression	-CLEF med 2009	Supervised learning	-LLP-GS and LLP-PS	- In comparison to near-lossless compression approaches, the proposed method obtains a PSNR of 41.46 and a CR of 8. 85	[36]

4.2.2 Unsupervised Learning Techniques for medical Image Compression based K-MEANS – PCA

According to [39], traditional machine learning techniques are typically thought to be easier to understand and converge more quickly than deep learning techniques. For various medical imaging applications, efforts have been made to improve medical image compression and accuracy by exploring different approaches, such as PCA and k-means; **Table 4-2** provides a thorough comparison of the data and highlights the differences between the various methods. The important contributions in this field are briefly summarized in the next section. First, k-means clustering compression was incorporated into logistic chaotic map encryption in [40] by shrinking the input image and enlarging the encryption key space. Several metrics, including the Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), Structural Similarity Index (SSIM), and correlation coefficients, are used to assess the efficacy of the suggested approach. Additionally, this study examines both modern and old lossy image compression techniques using the Kodak dataset [41], with a particular emphasis on autoencoders, Kmeans, DWT, and PCA or principal component analysis. The study reveals a relationship between SSIM and K values, where higher K values lead to longer processing times, and concludes that the K-means algorithm operates best when K

is set to 16.

However, in [42] ,the authors propose an automated arbitrary principal component analysis (AAPCA) as a regional compression technique, which uses a block-to-row principal component analysis algorithm (BTRPCA) based on factorization to compress the ROI after separating it from the background using brain symmetry. Principal Component Analysis (PCA) was used as a dimensionality reduction algorithm for brain MRI scans. In comparison to current compression techniques, the results show better segmentation, higher compression rates overall, and improved performance metrics (PSNR and Correlation coefficient (CoC)). It also achieved higher compression rates for non-ROI (NROI) while optimally compressing the ROI at a lower rate.

Besides, [43] introduces a novel medical image compression technique based on block-to-row bidirectional PCA with a focus on ROI. The method's superiority in terms of providing increased compression ratios and peak signal-to-noise ratios was demonstrated in **Table 4-2** by comparing its performance with the performance of other existing methods, such as ROI-based block-by-block PCA and ROI-based block-to-row PCA. Added to that, the authors of [44] assess the effectiveness of PCA and Wavelet Difference Reduction (WDR) on rib cage X-ray and lower abdomen CT scans. For both CT and X-ray pictures, the increase in the number of main components results in a drop in MSE and an improvement in PSNR. Compared to CT scans with comparable components, X-ray scans show inferior compression efficiency despite having a higher PSNR. Compared to CT scans of the same constituents, X-ray scans exhibit a low compression efficiency for a better PSNR. Therefore, a new hybrid color image compression technique was put forth [45] that combines the strengths of PCA (PCADTT) and Discrete Tchebichef Transform (DTT). It utilizes PCA to provide the reduction in dimensions of images and DTT to improve image quality for best image compression. The experimental results show that this approach is better than current methods in compression ratios, processing complexity, and image quality for a range of categories of image contents.

Table 4-2 Comparative Analysis of State-of-the-Art Lossy Image Compression Techniques, Including PCA, Kmeans, CNN, AE, DCAE, and VAE *[29]*

Learning Technique	**Ref**	**Year**	**Model**	**Dataset**	**Evaluation Metrics**			
					MSE	**PSNR**	**MS-SSIM**	**CR**
PCA	[45]	2021	PCA-DTT	Different Color Image	-	33.43	0.97	-
	[44]	2019	*PCA And Wavelet Difference Reduction (WDR)*	Lower Abdomen CT Scan	0.21	53.00	-	-
				Rib Cage X-Ray Scan	0.62	47.16	-	-

	[43]	2017	Block-By-Block PCA	MRI, CT, X-RAY	-	39.01	-	-
			Block-To-Row PCA		-	38.79	-	-
			Block-To-Row Bi-Directional PCA		-	39.02	-	-
	[42]	2022	AAPCA	MRI Brain Images	-	47.96	-	-
Kmeans	[41]	2022	Kmeans-Based Lossy Compression	Kodak Dataset	43.97	31.69	0.90	4.58
	[40]	2021	Kmeans-Based Lossy Compression	Two Colored Images	4.73×10^{-4}	33.24	0.97	-

4.2.3 Deep Learning Based Approaches for Medical Image Compression

Medical image compression methods based on deep learning utilize the strength of neural networks to learn compact and efficient representations of medical images. In this part, Autoencoders, Convolutional Autoencoders (CAE), and Variational Autoencoders (VAE) has been presented for a detailed comparative description of experts' research in this field.

4.2.3.1 Autoencoder-based Methods for Medical Image Compression

Turning to autoencoders, Senapati et al., [46] used a three-layer autoencoder architecture and a Kaggle dataset to study the compression and denoising of grayscale medical images. The proposed model was trained to predict the original, clean information from the noisy input using denoising autoencoders with non-linear activation characteristics on images corrupted with Gaussian noise. The findings from the comparative analysis of findings generated from methods like total variation and wavelet denoising showed that the model proposed performed better than others, specifically the SSIM. Furthermore, layer-by-layer pre-training was achieved by implementing the Deep Autoencoder for high dimensionality reduction using the Deep Boltzmann Machine as an approximation inference technique [47]. When compared to the performance of other autoencoders, such as Deep Autoencoder with Restricted Boltzmann Machine (DA-RBM), Deep Autoencoder with multiple back propagation (DA-MBP), and Deep Convolutional Autoencoder with RBM (DCA-RBM), the results showed that the suggested approach was effective and produced the best results in terms of PSNR, SSIM, and MSE. A two-stage autoencoder system for compressing malaria red blood cell (RBC) image patches is suggested in [48]. A residual-based dual autoencoder network, the suggested decompressor module reconstructs the original image while identifying distinguishing features. The quality of the chrominance component in cell pictures is assessed using the Color-SSIM metric after decompression. In comparison to previous neural network-based compression techniques for medical images, the suggested method shows gains in PSNR, Color SSIM,

and MS-SSIM of roughly 35%, 10%, and 5%, respectively.

Liu et al. [49], introduce a new medical image compression technique using variational autoencoders (VAEs) that combine VAE and residual network models. The objective is to obtain high image reconstruction and low information loss. Enhanced PSNR and MS-SSIM values are shown in experimental results compared to traditional approaches. Conversely, Bale et al. present a fresh image compression with nonlinear transforms [50] by using an application of a form of stochastic gradient descent; the model resolves zero-gradient artifacts of quantization and learns rate-distortion performance. The rate-distortion performance and image quality of the proposed model are comparable to baseline JPEG and JPEG 2000 compression. Vikraman and Jabeena [51] proposed a hybrid machine learning approach with two segmentation and optimization stages as a bridge to CNN. The approach divided the images into ROIs and NROIs using the Grey Wolf Optimization with Fuzzy C-Means algorithm. The ROI was later compacted by an optimized convolutional neural network (Op-CNN), while the NROI was compacted using recurrent neural networks (RNNs). With a PSNR of 45.502, the new approach outshone the existing approaches in quality. Conversely, Surbhit Shukla and Anugrah Srivastava [52] introduced a technique that makes use of wavelet transformation and convolutional neural networks (CNNs). A three-hidden-layer network was trained using a variety of medical imaging, such as CT, MRI, and X-ray scans. With fewer hidden neurons than image pixels, the network was able to compress target images by storing activation values and weights after training. For all examined medical photos, the algorithm performed better than alternative techniques in both lossy and lossless compression.

By using Convolutional Autoencoder Neural Networks (CAE-NN) for both the compression and decompression operations, the authors used DCAE in [53], [54], and utilized ADAM Optimizer to reduce loss and noise in the reconstructed image. While the outcomes were not very noteworthy, CAE Networks outperformed more traditional encoders like quantization and DCT in terms of efficiency.

Further, Guerrisi et al proposed a lossy image compression technique based on convolutional autoencoders (CAEs) [54] . Sentinel-2 images were utilized by the authors to evaluate the CAE architecture, measuring the quality of images in terms of SSIM and PSNR. Further, the authors explored the compressed images with respect to their usefulness in object detection applications. In another paper [55], three lossy image compression architectures were introduced by the authors: Super Resolution (with SRCNN structure), Generative Adversarial Networks (GANs), and Convolutional Autoencoders (CAEs). CAEs, in the authors' view, are valid alternatives to traditional transforms like the wavelet transform and the Discrete Cosine Transform (DCT) since they learn efficient compact features for improved coding efficiency over JPEG. Two application domains where GANs exhibit promising benefits are higher compression ratios and good subjective quality reconstruction. Super-resolution achieved the best rate distortion (RD) performance of the architectures and it is equal to that of BPG. It is notable to mention the

recent progress in weakly supervised learning techniques, particularly in medical image analysis, though our research is mostly focused on unsupervised learning from compression techniques. We are motivated to investigate similar approaches in compression issues, by the success of semi-supervised techniques in learning from unlabeled data, as illustrated by the UKSSL model proposed by Ren et al.[56], The potential advantages of having both unlabeled and labeled data in our compression methods are suggested by the use of contrastive learning models, e.g., MedCLR, for extracting knowledge and later fine-tuning with sparse labeled data. Also, the use of poorly supervised approaches, viz., instance-based, bag-based, and hybrid approaches, suggests how strong the models are under noisy or sparse labels [57]. Although its primary domain is medical image analysis, this weakly supported learning observation has a broader impact on increasing the flexibility of compression algorithms when presented with faulty or incomplete data. Finally, the research of [58] is grounded in the idea of a Quantum-amplified Artificial Neural Network framework to foster medical image compression. The key idea is to use quantum computing to maximize feature extraction and later integrate it with traditional neural network architectures. By projecting classical data into quantum states and further utilizing these states and then subjecting them to quantum circuits, the model outputs amplified feature vectors that are taken as inputs by a traditional neural network for compressing images.

The QANN model leverages quantum entanglement, parallelism, and superposition to achieve image compression via feature extraction and compression of their representation with the assistance of quantum processing. Compression is then succeeded by inverse operations or classical image processing techniques that recover the image. The model has demonstrated enhanced peak signal-to-noise ratio and structural similarity index, and reduces image size for MRI (73.3%), X-ray (74.1%), and CT-SCAN (71.8%). The combination of quantum feature extraction with traditional neural networks proves that quantum computing can significantly increase the efficiency of medical image processing solutions. **Table 4-3** illustrates a comparative analysis of state of the art.

Table 4-3 Comparative Analysis of State-of-the-Art: Lossy Image Compression Techniques, Including AE, DCAE, and VAE *[29]*

Learning Technique	Ref	Year	Model	Dataset	Evaluation Metrics			
					MSE	PSNR	MS-SSIM	CR
AE	[48]	2020	Lossy Compression Based-Autoencoder	Malaria Cell Images Dataset	-	35.91	0.97	-
	[47]	2020	Deep Autoencoder with Deep Boltzmann Machines	NIH Clinical Center's Chest X-Ray Image Dataset	0.01	37.02	0.99	-

	[46]	2022	Compression And Denoising AE	Kaggle Dataset-Medical MNIST Dataset	-	23.91	0.92	-
VAE	[50]	2018	End-To-End Compression	Kodak Test Set	-	32.43	0.97	-
	[49]	2022	Lossy Compression Based-VAE	TCGA-LUAD Lung Cancer CT Image Dataset	-	Higher PSNR And MS-SSIM Values Than Ballé 2018		-
CNN	[51]	2023	Optimized Convolutional Neural Network (Op-CNN)	Brain MRI Segmentation	2.33	45.50	-	27.45
	[52]	2018	Compression Based-Neural Networks (CNN + LWT)	MRI, X-Ray, Computer Tomography Images	/	39.59	-	
DCAE	[54]	2023	Lossy Compression Based-Convolutional Neural Network	Images In the RGB Spectrum Obtained from Sentinel-2 Data.	-	25.65	0.87	-
	[53]	2020	Convolutional Auto Encoder Neural Network (CAE-NN)	MNIST Dataset	108.2	-	-	-
	[55]	2018	CAE	IMAGENET Database	-	26.48	0.82	-

4.2.3.2 Convolutional Autoencoders (CAEs) Vs. Run-Length Encoding (RLE) Based Compression Techniques

This section provides a concise summary of current approaches for compressing medical images using both conventional and deep learning techniques.

First, using the BRATS dataset, the authors of [59] suggest an SPIHT picture compression method based on regions of interest. In comparison to the traditional SPIHT, our technique prioritizes clinically essential regions and operates well at low bit rates while demonstrating higher visual quality. Additionally, Prasantha et al. [60] used MSVD, which divides the image into 64 × 64 blocks and applies SVD to each block, rather than using the SVD approach on the full image. Particularly in the ROI, MSVD exhibits better visual quality at the same bit rate than SPIHT while simultaneously lowering computing cost when compared to conventional SVD.

On the other hand, [61] authors offer a no-reference method of compressed medical image evaluation that mimics the formations of medical images by classifying the shapes of the images using a convolutional neural network. CNN classification accuracy is calculated by calculating the loss of information because

of irreversible operations. Using various interpolation methods, studies discovered that the model could distinguish between photo quality encoded by JPEG and JPEG2000 with great accuracy.

Nagoor et al. [62] use deep neural networks to apply lossless compression techniques for volumetric medical images by taking advantage of structural similarity and shared feature representation across multiple scans with the same resolution and acquisition parameters, as well as local sampling (3D cube and 3D pyramid) within a single volume. Flattened 1D vectors of 3D nearby voxel sequences were used to train the model. When compared to the most advanced lossless compression techniques, the results show a greater compression ratio. In a different use of deep learning, [63] suggested a lossless method for compressing medical images by segmenting them into sections according to anatomical traits using anatomical information and a deep neural network. After that, the suggested neural network is trained to produce the best predictors for every area. With a 38%, gain in compression performance over JPEG2000, this "split and conquer" strategy produces better compression.

Additionally, Shukla et al. [52] included a left wavelet transformation and used the backpropagation technique to train a three-layer convolutional neural network. The goal of this research is to outperform feed-forward neural networks in terms of compression. Furthermore, a high PSNR for MRI and CT scan images with good image quality retention was obtained by integrating the wavelet-SPIHT approach after compression. Furthermore, [29] suggested CNN, AE, and DCAE compression algorithms, which were then applied to the medical x-ray imaging dataset (MXID). With a PSNR value of while maintaining a high-quality reconstructed image and important features, DCAE performed better in experiments than CNN and AE.

For 16-bit depth medical image compression, a better JPEG-XT algorithm (OPT_JPEG-XT) was introduced in [64] it multiplies the discrete cosine transform coefficients by N, which allows for lossy and lossless compression. The truncated decimal parts of the DCT coefficients in standard JPEG-XT are kept by OPT_JPEG-XT, according to the experiments of the authors. With the maintenance of lossless compression within the upper sub-images, results show a very high compressed image quality with MSE less than 0.08 and higher PSNR values at a saving space rate (SSR) of 60%.

When investigating the application of the Discrete Wavelet Transform to the compression of numerous images. Agarwal et al. [65] compressed Magnetic Resonance Imaging (MRI) using wavelet families and resolutions. Large medical images can be sent using this DWT-based technique, which uses Haar, Daubechies type 4, and Bi-orthogonal wavelets to produce a high compression ratio with a satisfactory visual image quality.

Additionally, two identical feed-forward neural networks (FF-NNs) (predictors) using gravitational search and particle swarm algorithms were applied in lossless and near-lossless approaches using predictive image coding techniques in [36] on the CLEF med 2009 dataset. Additionally, the method

separates images into diagnostically important regions (DIRs) and non-DIRs using a graph-based segmentation (GBS) technique. With a Markov model and arithmetic coding, the prediction error is significantly reduced. This method produced better results than near-lossless methods, with a PSNR of 41.46 dB and a CR of 8.85. Further, utilizing a deep convolutional neural network to learn compact representations of medical images, followed by a Huffman encoder, Navaneethakrishnan & Shanmugam [66] created a novel method for compressing medical images. Using the databases Kaggle and Medpix. The results demonstrate the efficacy of the suggested method with high SSIM and SSS values, which show a high degree of similarity between the compressed and original images, good compression ratios, which range from 25.76 to 29.97 and PSNR values, which range from 43.51 dB to 46.29 dB.

4.3 Conclusion

Machine learning has greatly facilitated medical image compression research using adaptive and data-driven techniques that surpass traditional methods both in efficiency and information retention. In this chapter, different machine learning techniques (supervised, and unsupervised) has been presented alongside with the deep learning-based researches on (Autoencoders, CNNs, and VAEs).

This chapter outlined recent developments in medical image compression using machine learning, outlining various methodologies, their applications, and relative strengths. As promising as deep learning has rendered the area of compression, a detailed state of the art has been presented, where we have tried to recover the most similar studies by researchers in medical image compression field. The next chapter will explain thoroughly the details on the newly created MXID dataset in this research for different medical imaging tasks, such as classification, segmentation, compression, and detection.

5 CONSTRUCTED MXID DATASET

Contents

5.1 Introduction

The storage and processing of medical imaging, and more specifically X-ray scans, are at the center of the modern practice of healthcare diagnostics. The usefulness of the technique is that it is cost-effective, efficient, accessible, non-invasive, and has great potential for diagnosis, for example, the identification of fractures, identification and assessment of infection, and the evaluation of a range of pathological conditions. However, many problems are created by the increasing amount of data, including storage and transmission. To address these challenges, a comprehensive dataset named MXID (Medical X-ray Imaging Dataset) has been developed in this research. This dataset facilitates image compression, enhancement, classification, and various other processing tasks at an efficient scale. Designed to be both high-quality and widely accessible, it serves as a valuable resource for evaluating and training machine learning and deep learning models.

Comprising 6,869 high-quality X-ray images, the MXID dataset has been collected from AOUINET Hospital in Tebessa, it includes X-ray images from 18 different anatomical regions classified and categorized based on patient gender and body part identification (type). The dataset is constructed to support different image processing tasks, including image classification, segmentation, detection, augmentation, compression, machine learning and deep learning-based techniques.

This chapter includes a full explanation of the MXID dataset, including dataset collection, preprocessing, labeling, classification, characteristics, and a comparison to other datasets. In addition, an expanded set of 3,120 X-ray pictures has been included for the 18 distinct classes in MXID, dubbed LAVIA MXID—an improved version with higher diversity and a richer collection of medical images. This extended dataset is an invaluable resource for developing and training machine learning and deep learning approaches for a wide range of medical image processing jobs.

5.2 MXID Dataset Construction

MXID dataset creation is a neatly organized process for a quality medical X-ray image dataset suitable for analysis activities across a range of different tasks. The initial step is tackling the real acquisition of images from medical sources to ensure accuracy and authenticity. Preprocessing methods can include scaling, normalization, and removal of noise to enhance and preserve uniformity of quality images. For a well-balanced anatomical shape of different body areas, the images are carefully annotated in different zones of anatomy. The final dataset is divided into a test set, a validation set, and a training set for preparations on its intended uses, namely, image compression as well as the training of models for diagnostic use.

5.2.1 Overview of our MXID dataset

Specialist radiologists and doctors have painstakingly built and labeled the various elements of the medical body. Through labelling each image with the right body part, an accurate assessment of compression performance is attempted. The suggested MXID dataset was built sequentially, as seen in **Figure 5-1**. The dataset is divided into train, validation, and test sets with a trained split. This approach ensures that body parts are distributed equitably among subgroups, which reduces bias, and enables the development and test of models without doubt.

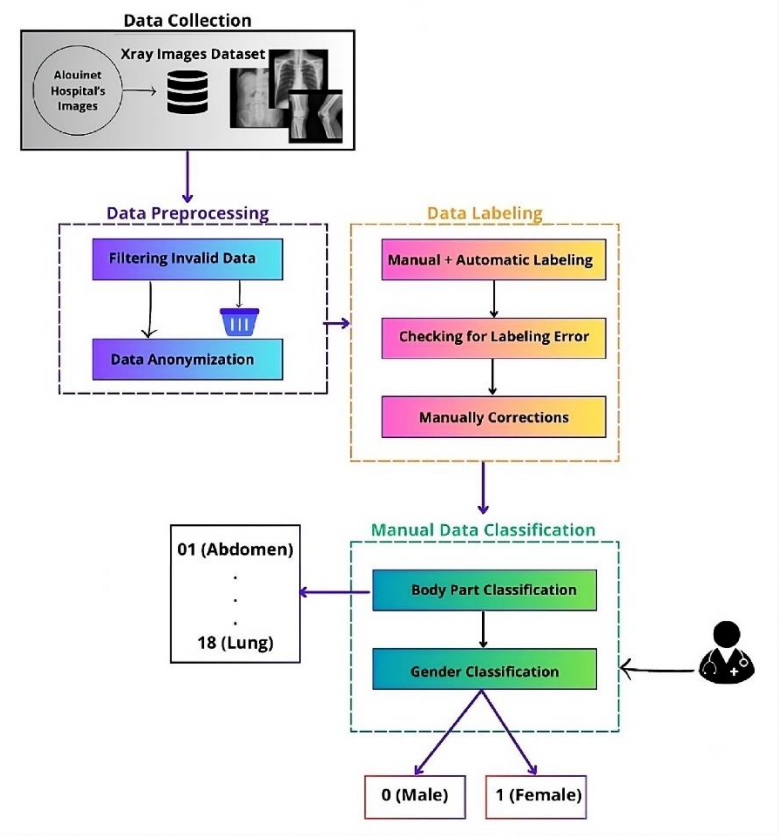


Figure 5-1Representation of the MXID Dataset Development Process *[29]*

5.2.2 Dataset Collection and preparation

This study is based on a collection of 6,869 medical images acquired from AOUINET Hospital, Tebessa City, Algeria. The images were collected over two periods of time: August-October 2022 and January-April 2023. All of them have a resolution of 1024 × 1024 pixels and were taken from patients undergoing diagnostic imaging scans. The acquisition technique utilized computed radiography (CR) plates to effectively position patients inside the treatment area for lung block form design.

For digitization of images, CR 15-X tabletop system was used. It is a good cost-benefit ratio with excellent image quality by Agfa HealthCare's established technology. It is designed for flexibility

with adjustable speed, easy workflow, and compatibility with a broad range of digital radiography applications. To boost imaging capability further, it can be combined with an NX workstation, a DRYSTAR 5302 digital thermal printer, or an SE software package.

Optimal imaging is achieved in the CR 15-X with the MUSICA image processing algorithm that can automatically analyze all the features of all images and automatically adjust the processing parameters. This leads to optimal imaging results irrespective of body area variation or X-ray dose levels, with less retakes and post-processing. Moreover, with a DR system, MUSICA provides an image consistency of computed radiography (CR), but with increased diagnostic accuracy. Being executable on its own in the form of an algorithm, there is no calibration and preprocessing needed with human intervention, hence being readily implementable with or without training.

For ensuring patient data confidentiality, the database was anonymized in its entirety by removing names and all other identifiable personal information. 152 incomplete or faulty images were picked up during preliminary scanning in the quality control process. They were then erased to ensure the accuracy and reliability of the dataset. This systematic identification and rejection of erroneous images also contributes to improving the overall quality of the dataset, reducing potential bias and making downstream experiments and analysis more reliable.

Lastly, to preserve the quality and integrity of the dataset, an equally careful screening was also performed to eliminate duplicate images. Duplicate images that recurred repeatedly were scanned and eliminated one by one, leaving behind a single instance of each unique image. In both data collection processes, over a hundred duplicates were found and eliminated. **Figure 5-2** illustrates some example images of the proposed MXID dataset.

Figure 5-2Visual Samples of the MXID Dataset *[29]*

5.2.3 Dataset labeling

Following the preparation phase, a thorough manual verification was conducted to confirm the accuracy of the labels of the dataset. The images were verified one by one to confirm that the label accurately represented the visual information. Any errors identified were corrected on the spot, and comments were made. This rigorous verification went on until all the images of the dataset had been verified. Following the manual review, a script was run to verify that no labeling errors were committed on the entire dataset. The script searched for any discrepancy between the filenames and image labels. The images that were flagged were then intensively searched for any errors in an effort to identify and correct any errors to guarantee the correctness of the dataset.

The medical record dataset was built on a highly structured naming convention for each image. The norm has fields in the captioned format of PatientID, TypeID, Image_Count, Gender, and Serial Number. In this approach, each patient is assigned a unique identification that is tagged as PatientID to make it easy to manage medical data and identify various X-ray images of a subject. TypeID classifies the category of image, and the part of the body or condition it shows. Such a systematic nomenclature system guarantees consistency and comprehension in processing the dataset.

Furthermore, the Image Copies attribute provides the number of various images that were taken for a patient, each one being of a different body part in a particular medical case. A valuable demographic attribute is added to the database by the Gender attribute, with 0 indicating males and 1 indicating females. In addition, as shown in **Figure 5-3**, a Serial Number is given to every image that is a unique identifier for the vast majority of images of the same modality and patient.

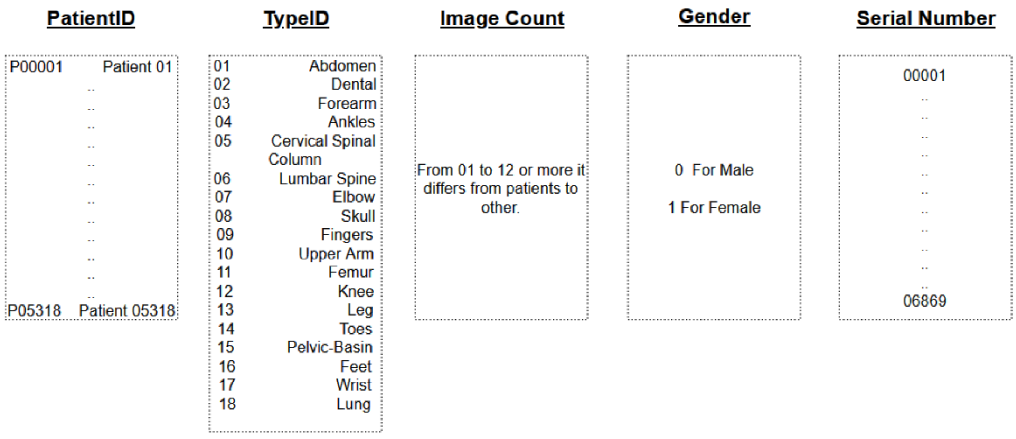


Figure 5-3 Comprehensive Image Labeling Displaying Specific Identifiers for PatientID, TypeID, Image_Count, Gender, and Serial_Number *[29]*

This annotated data provides a solid basis for detailed examination and analysis. It was developed with extreme attention to detail and extensively validated by an experienced panel of experts,

including radiologists, to be accurate and reliable.

5.2.4 Dataset for body part and gender classification

X-ray imaging is the most widely used technique in medical imaging, which employs a variety of techniques to produce precise images of the body's internal organs for diagnostic purposes. The lung, belly, tooth, skull, leg, upper arm, knee, finger, cervical spinal column, lumbar spine, ankle, foot, wrist, forearm, elbow, and pelvic basin are among the 18 carefully categorized body component categories found in the MXID collection. Every one of these X-ray types provides vital medical data that enables physicians to identify and diagnose a variety of illnesses. Each image is classified as either male or female in order to provide a demographic dimension. The MXID dataset's photo distribution by gender and body part is shown in **Figure 5-4**.

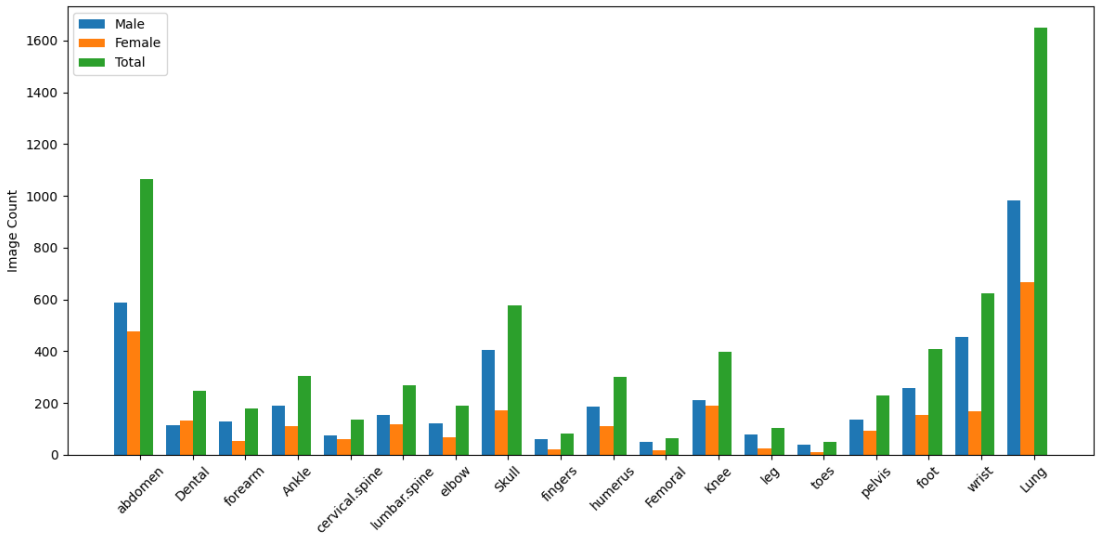


Figure 5-4 Image Distribution in the MXID Dataset by Body Part and Gender *[29]*

5.2.5 Dataset Split

The database was strictly divided into different subsets based on an automated multiclass methodology for providing sufficient training, testing, and validation of the model. The process divided the photographs equally among the three sets. Ten percent were reserved for validation, twenty percent for testing, and seventy percent for training based on a predefined ratio. The division maintained that all the subgroups had a representative class label in balanced form. Distribution of the X-ray images on the basis of gender and body part is described below. **Table 5-1** displays the training, testing, and validation samples utilized, while the gender distribution for body part is presented in **Table 5-2**.

Table 5-1 Strategy for Dividing the MXID Dataset into Training, Testing, and Validation Sets *[29]*

TypeID	Study	Train	Test	Validation	Total
01	ABDOMEN	742	221	101	**1064**
02	DENTAL	177	47	24	**248**
03	FOREARM	129	35	15	**179**
04	ANKLE	209	57	37	**303**
05	CERVICAL SPINAL COLUMN	96	27	11	**134**
06	LUMBAR SPINE	177	61	30	**268**
07	ELBOW	133	34	21	**188**
08	SKULL	404	111	63	**578**
09	FINGERS	56	16	8	**80**
10	UPPER ARM	212	51	36	**299**
11	FEMUR	44	13	8	**65**
12	KNEE	276	82	42	**400**
13	LEG	76	19	8	**103**
14	TOES	35	10	5	**50**
15	PELVIC-BASIN	159	44	25	**228**
16	FEET	282	82	45	**409**
17	WRIST	434	133	56	**623**
18	LUNG	1159	343	148	**1650**
	Total	**4800**	**1386**	**683**	**6869**

Table 5-2 Gender Allocation Across Various Body Parts in the MXID Dataset *[29]*

TypeID	Study	Male	Female	Total
01	ABDOMEN	588	476	**1064**
02	DENTAL	115	133	**248**
03	FOREARM	127	52	**179**
04	ANKLE	191	112	**303**
05	CERVICAL SPINAL COLUMN	75	59	**134**
06	LUMBAR SPINE	152	116	**268**
07	ELBOW	120	68	**188**
08	SKULL	406	172	**578**
09	FINGERS	61	19	**80**
10	UPPER ARM	187	112	**299**

11	FEMUR	49	16	**65**
12	KNEE	211	189	**400**
13	LEG	79	24	**103**
14	TOES	39	11	**50**
15	PELVIC-BASIN	135	93	**228**
16	FEET	256	153	**409**
17	WRIST	455	168	**623**
18	LUNG	984	666	**1650**
	Total	**4230**	**2639**	**6869**

5.2.6 Dataset for patient identification

By identifying likely relationships and patterns between a large number of body areas, this multi-dimensional approach improves diagnostic accuracy. The 5,318 multibody component imaging dataset includes a strong foundation for using machine learning to improve medical diagnosis and inform the building of personalized treatment regimens. The images are organized and archived systematically after being assigned new names, after renaming with the PatientID_TypeID-ImageCount_Gender_SerialNumber scheme. **Figure 5-5** clearly illustrates this naming convention.

- Patient identifier: P00051.
- Type identifier: 06, 15, 17, 03, 08, 18, 01, 12. (representing different body parts)
- Image count: from 01 to 10.
- Gender: 0 ('0' represents male)
- Serial number: unique identifier for each image from 00075 to 00085.

The sample image depicts a patient P00051 receiving X-rays of different parts of the body, e.g., Lumbar Spine (06), Pelvic-Bassin (15), Wrist (17), Forearm (03), Skull (08), Lung (18), Abdomen (01), and Knee (12). The patient received 10 different sessions of X-ray imaging for different parts of the body, as illustrated in **Figure 5-5.**

Figure 5-5 Patient Identification – Illustrating the Association of a Single Patient with Multiple X-ray Images from Different Body Regions *[29]*

5.2.7 MXID dataset for medical image compression

6,869 medical X-ray images from the MXID dataset have been utilized in this research study. This was resized to 256 × 256 pixels memory-efficiently when implementing it. Due to this, division of the provided dataset into various groups was done automatically using an automatic multiclass partition method so that there would be adequate training, testing, and validation of the model. Partitioning was performed automatically to achieve balanced sample distribution and proportionate class label representation. The dataset was divided into three subsets with a fixed proportion of 60% for training, 20% for testing, and 20% for validation. Automatic multiclass partitioning gives class labels balanced in subsets, leading to overall better balance in the dataset. Besides, preprocessing methods like scaling, normalization, and feature extraction were also applied to improve the images prior to analysis. The increasing demand for high-resolution medical images has posed serious challenges in picture storage, transmission, and retrieval. Thus, there is an urgent need for new compression and data management techniques to utilize these precious diagnostic tools efficiently, which will be discussed in the subsequent section.

5.3 MXID Dataset's Characteristics

MXID dataset is a diversified structured medical X-ray image dataset for addressing a wide variety of medical imaging analytical tasks. The key features are:

- Collected 9,989 high-quality X-ray images.
- **Image resolution:** 1024 by 1024 pixels.

- **Body Part Categories:** Images are labeled based on anatomical regions in order to maintain balanced distribution.
- Data was collected at AOUINET Hospital, Tébessa city, over two years.
- The images are kept in PNG to maintain quality.
- The dataset is unlabeled for pathology and can be used for any of the image analysis applications.
- **Dataset partitioned into:** training, validation, and testing subsets for machine learning applications.
- **Diversity:** Covers many body areas, making it suitable for deep learning models in medical imaging.

5.4 Comparison with Existing Datasets

The advancement and assessment of machine learning algorithms in the field of medical imaging are significantly dependent on the quality and balance of the datasets utilized. To pinpoint areas requiring enhancement and further investigation, it is essential to analyze the existing datasets and understand their constraints. **Table 5-3** offers a summary of the characteristics of these datasets along with a concise description.

The MURA dataset [67] consists of a large set of musculoskeletal images aimed at aiding the training of deep learning models for the diagnosis of skeletal conditions. While the dataset serves as a valuable resource, it may fall short in precisely depicting certain anatomical regions, as some variations might not be fully reflected in the X-ray images. Additionally, the NIH Chest X-ray dataset [68] comprises a vast set of frontal chest X-ray images that capture various conditions, including pneumonia, lung nodules, and other abnormalities. Although the dataset is a valuable resource, it may fall short in its ability to reflect some regions of anatomy, as some of the variations may not be fully reflected in the X-ray images. The NIH Chest X-ray dataset [68] also includes a vast repository of frontal chest X-ray images that reflect various conditions like pneumonia, lung nodules, and so on. Although its focus is primarily on chest disorders, limiting its explanation to other parts of the body, it remains a useful resource in creating and validating machine learning models. Alternatively, the Chest X-ray 8 dataset [69] is a collection of chest X-ray images and their corresponding radiological reports. It encompasses more than one abnormality type, such as lung nodules, pneumonia, and cardiomegaly. This database has served as a standard for building and testing algorithms that are programmed to read chest X-rays. The JSRT dataset [70], likewise, primarily concerns pulmonary diseases and has served extensively in the research to build and test methods for lung abnormality detection and classification.

The RSNA Pneumonia dataset [71] is a chest X-ray image database that specializes in pneumonia

detection. It serves as a benchmark dataset for training and validation of algorithms aimed at the detection and diagnosis of pneumonia. Likewise, the INbreast dataset [72] contains mammographic images for breast cancer screening. The dataset is suitable for additional research studies and algorithm development for the detection of breast cancer. Meanwhile, the ISIC Skin Lesion dataset [73] has a wide variety of skin lesions and is a core dataset for training models in skin lesion classification and diagnosis. It discusses a number of skin diseases, including melanoma, with a sharp interest in lesion analysis.

The LIDC-IDRI dataset [74] features chest CT scans that are predominantly centered on lung nodules. It has been frequently employed in the development and evaluation of computer-aided detection (CAD) systems designed for the identification and classification of these nodules. Likewise, the Open-I dataset [75] represents a multimodal medical imaging repository that encompasses images from multiple body regions. This dataset is a significant resource for researchers and medical practitioners engaged in a range of medical image processing endeavors. This dataset encompasses a diverse range of medical imaging modalities, such as X-rays, CT scans, and MRI scans. Furthermore, the CAMELYON16 dataset [76] is specifically focused on breast cancer detection, serving as a benchmark for the automated diagnosis and classification of breast cancer, and comprises digitized histopathology slides. **Figure 5-6** represents a comparison of related datasets based on images number.

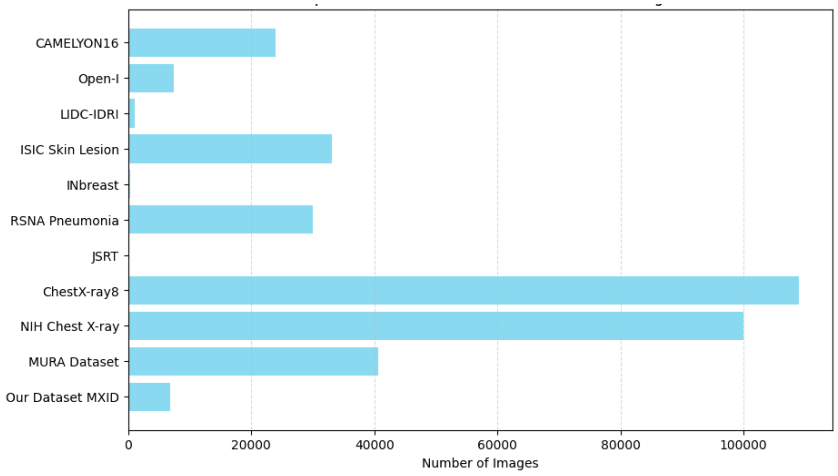


Figure 5-6 Comparison of Related Datasets Based on the Number of Images *[29]*

Advancing machine learning in medical imaging depends on the development of comprehensive and diverse X-ray datasets. Maintaining consistency and accuracy in model training is essential for reliable performance. Addressing these challenges can lead to more accurate diagnostic algorithms, ultimately improving patient care and overall healthcare outcomes. **Table 5-3** illustrates the availability of publicly accessible medical X-ray datasets with corresponding ground truth.

Table 5-3 Summary of Publicly Accessible Medical Radiography Image Datasets *[29]*

Datasets	Refs	Clinical Context	The number of images	Body parts	Data source	Availability
Our Dataset MXID	/	Multi-body parts	6869 images	Skull, legs, lungs, shoulders, knees, fingers, cervical spinal column, lumbar spine …	AOUINET Hospital	Publicly available
MURA Dataset	[67]	Musculoskeletal disorders	40561 images	Elbow, finger, forearm, hand, humerus, shoulder, and wrist	Stanford University	Publicly available
NIH Chest X-ray Dataset	[68]	Chest conditions	100,000 images	Chest	National Institutes of Health	Publicly available
ChestX-ray8	[69]	Chest conditions	108,948 images	Chest	National Institutes of Health	Publicly available
JSRT	[70]	Pulmonary diseases	247 images	Chest	Japanese Society of Radiological Technology	Available (limited access)
RSNA Pneumonia	[71]	Pneumonia	30,000 images	Chest	Radiological Society of North America	Publicly available
INbreast	[72]	Breast cancer screening	410 images	Breast	University of São Paulo	Available (limited access)
ISIC Skin Lesion	[73]	Skin lesions	33,126 images	Skin	International Skin Imaging Collaboration	Publicly available
LIDC-IDRI	[74]	Lung nodules	1,018 images	Lung	National Institutes of Health	Available (limited access)
Open-I	[75]	Various medical images	7470	Multiple body parts	National Library of Medicine	Publicly available

CAMELYON16	[76]	Breast cancer detection	23916	Breast	Radboud University Medical Center	Available (limited access)

Image quality problems can make it difficult to identify details in medical X-ray images, and compression techniques can make the task even more difficult. Maintaining image integrity is the main goal of medical image compression [77]. Pixel-level compression accuracy should be precisely evaluated as part of an efficient evaluation process for X-ray image compression [78]. The two main categories of image compression methods are lossless and lossy. Compared to the original, lossy compression drastically lowers file size but degrades image quality [21]. Conversely, lossless compression reduces file size without sacrificing data, which makes it appropriate for applications that need accurate medical evaluations. Image compression is covered in more detail in the following section.

5.5 Overview of LAVIA-MXID dataset

The LAVIA-MXID dataset is an upgrade of the MXID dataset supplemented with 6869 X-ray images in that this upgraded version is accomplished with greater diversity and a more extensive set of medical images. It now comprises 9989 images selected and handled by expert radiologists and medical experts. Every image in the LAVIA-MXID dataset is meticulously labeled with its corresponding body area as displayed in **Figure 5-7**, thus allowing precise evaluation of various medical analysis techniques. Due to its extremely fine-grained annotations, the dataset can even be utilized for a wide range of machine-learning and deep-learning applications.

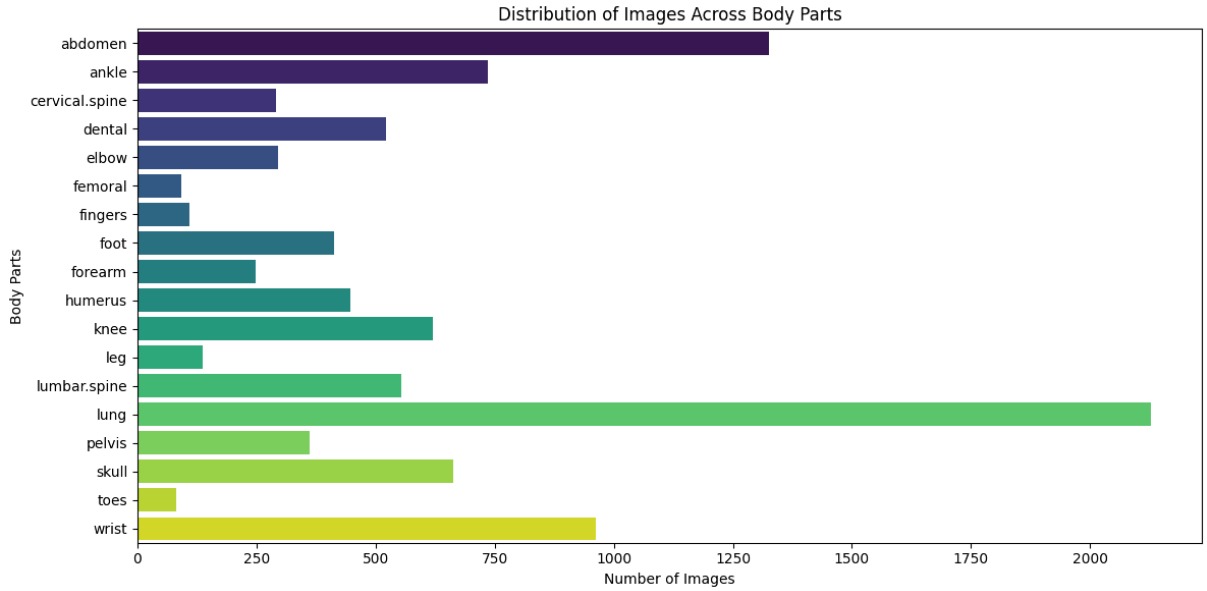


Figure 5-7 LAVIA-MXID Dataset Distribution Across Body Parts

The process of dataset preparation is such that it ensures a balanced composition of different body parts to avoid bias while training and testing models. **Figure 5-8** demonstrates that all images are also logically divided into train, validation, and test sets, which are beneficial for model development and model testing. The dataset includes 9,989 images collected in a time span of approximately two years from the AOUINET Hospital in Tebessa City. All of the medical images are of a size of 1024 × 1024 pixels. LAVIA-MXID dataset is ready for a range of image analysis goals other than compression of images, and images are categorized by body part and tested for consistency. However, it does not possess any annotations regarding pathological conditions because it is to be used along with a range of image analysis goals rather than being restricted to pathology-specific studies.

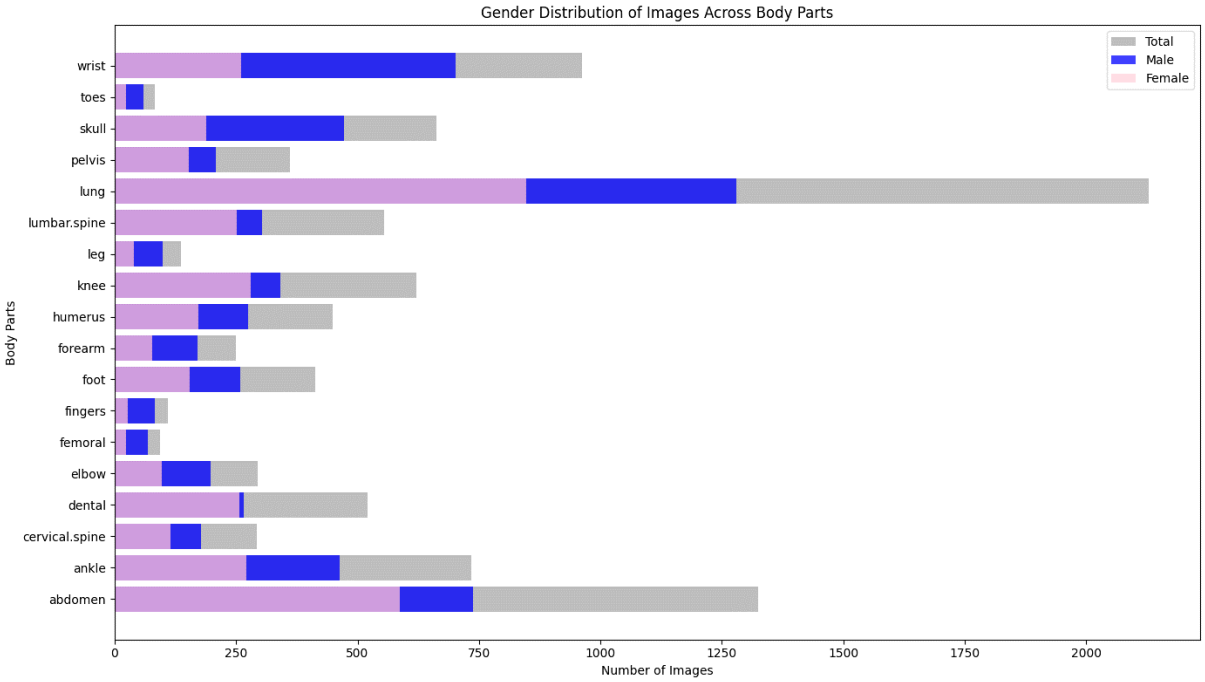


Figure 5-8 LAVIA-MXID Dataset Distribution Based Gender Across Body Parts

Table 5-4Comparison of Sample Counts in the LAVIA-MXID and MXID Datasets, Categorized by Body Part and Gender

N°	body part	Dataset	Male	Female	Total	N°	body part	Dataset	Male	Female	Total
1	**Abdomen**	LAVIA-MXID	738	587	**1325**	2	**Dental**	LAVIA-MXID	265	256	**521**
		MXID	588	476	**1064**			MXID	115	133	**248**
3	**Forearm**	LAVIA-MXID	171	78	**249**	4	**Ankle**	LAVIA-MXID	463	272	**735**
		MXID	127	52	**179**			MXID	191	112	**303**

5	**Cervical Spine**	LAVIA-MXID	177	115	**292**	6	**Lumbar Spine**	LAVIA-MXID	303	252	**555**
		MXID	75	59	**134**			MXID	152	116	**268**
7	**Elbow**	LAVIA-MXID	198	97	**295**	8	**Skull**	LAVIA-MXID	473	189	**662**
		MXID	120	68	**188**			MXID	406	172	**578**
9	**Fingers**	LAVIA-MXID	83	27	**110**	10	**Upper arm**	LAVIA-MXID	275	173	**448**
		MXID	61	19	**80**			MXID	187	112	**299**
11	**Femoral**	LAVIA-MXID	69	24	**93**	12	**Knee**	LAVIA-MXID	341	280	**621**
		MXID	49	16	**65**			MXID	211	189	**400**
13	**Leg**	LAVIA-MXID	98	39	**137**	14	**Toes**	LAVIA-MXID	59	23	**82**
		MXID	79	24	**103**			MXID	39	11	**50**
15	**Pelvic-Basin**	LAVIA-MXID	209	152	**361**	16	**Feet**	LAVIA-MXID	259	154	**413**
		MXID	135	93	**228**			MXID	256	153	**409**
17	**Wrist**	LAVIA-MXID	702	260	**962**	18	**Lung**	LAVIA-MXID	1280	848	**2128**
		MXID	455	168	**623**			MXID	984	666	**1650**

Table 5-5 LAVIA-MXID Dataset Partitioning Across Training, Testing, and Validation Sets

Body part	MXID Dataset				**LAVIA-**MXID Dataset			
	Train	Test	Validation	**Total**	Train	Test	Validation	Total
ABDOMEN	659	307	98	**1064**	913	272	140	**1325**
DENTAL	150	71	27	**248**	352	111	58	**521**
FOREARM	106	56	17	**179**	165	65	19	**249**
ANKLES	181	93	29	**303**	520	148	67	**735**
CERVICAL SPINAL COLUMN	79	43	12	**134**	210	55	27	**292**
LUMBAR SPINE	158	81	29	**268**	395	103	57	**555**
ELBOW	122	55	11	**188**	209	48	38	**295**
SKULL	338	171	69	**578**	448	134	80	**662**

FINGERS	48	23	10	**81**	73	24	13	**110**
UPPER ARM	188	81	30	**299**	316	88	44	**448**
FEMUR	37	26	2	**65**	72	11	10	**93**
KNEE	232	131	36	**399**	437	120	64	**621**
LEG	67	27	9	**103**	102	24	11	**137**
TOES	29	13	7	**49**	56	17	9	**82**
PELVIC-BASIN	138	58	32	**228**	259	67	35	**361**
FEET	254	112	44	**410**	307	54	52	**413**
WRIST	383	176	64	**623**	682	188	92	**962**
LUNG	963	517	170	**1650**	1537	403	188	**2128**
Total	4132	2041	696	**6869**	7053	1932	1004	**9989**

5.6 Comparative Description of MXID and LAVIA-MXID Dataset

The LAVIA-MXID dataset is an upgraded version of the MXID dataset, with increased variety and dataset size. **Table 5-6** is a comparative analysis of important differences between the two:

Table 5-6 Comparison of the MXID and LAVIA-MXID datasets

ASPECT	**MXID DATASET**	**LAVIA-MXID DATASET**
DATASET SIZE	6,869	9,989
DIVERSITY	Moderate, primarily covering specific body parts	Enhanced with a more balanced representation across various anatomical areas.
LABELS	Labeled by body region and gender	Labeled with expert validation for accuracy.
DATA COLLECTION	Sourced from AOUINET Hospital	Expanded dataset with added images to enhance diversity.
PIXEL DENSITY	1024 × 1024 pixels	1024 × 1024 pixels
DATA FORMAT	PNG	PNG
DATA SEGMENTATION	Divided into training, validation, and testing sets.	Preserves the structured partitioning to enhance model assessment.

APPLICATION	Primarily used for image analysis, compression, and classification	Extended to support deep learning-based diagnostic research and wider applications

LAVIA-MXID is a polished and fully-fledged version of MXID with a more extensive and balanced dataset that has further increased its use in medical image analysis, compression methods, and deep learning-based solutions.

5.7 Conclusion

This chapter covered the entire process of dataset preparation, organization, and generation, emphasizing the key significance of high-quality medical X-ray datasets for most image analysis tasks. The chapter introduced LAVIA-MXID as a more sophisticated MXID, a dataset that was developed with the vision of offering a more rich and organized set of images to support the development of accurate machine-learning models. Additionally, a comparative analysis of relevant datasets demonstrated the exclusive strengths and advantages of the proposed dataset. The chapter also discussed significant features such as dataset partitioning strategies and image preprocessing techniques, demonstrating their impact on model performance. By presenting a well-structured and well-organized dataset, the subsequent steps serve as the foundation for applications in image compression, model training for diagnostics, and overall medical imaging research, leading to further development of AI-based healthcare solutions.

6 IMPLEMENTATION AND RESULTS

6.1 Introduction and Motivation

Medical imaging plays a crucial role in healthcare sector, it enables accurate diagnosis and aid in treatment planning, and patient monitoring. However, the increasing volume of the different modalities of high resolution, such as X-ray, CT scans, and MRIs poses challenges in the storage, transmission, and processing; leading to increased costs and management complexities. Thus, effective image compression is required to reduce/optimize transmission and storage while ensuring no critical loss in diagnostic information. Several models have been developed to address these challenges, including Machine Learning-Based Compression of Principal Component Analysis (PCA) and K-means clustering, and Deep learning models such as Autoencoders (AE), Convolutional Autoencoders (CAE), and Variational Autoencoders (VAE) have shown promising results in learning compact representations of medical images.

This chapter presents the implementation details and results of the proposed contributions, the datasets used, the objectives of deep leaning models training, and a detailed analysis of different results is provided, highlighting comparisons and key outcomes. Additionally, the performance of the different deep learning models (AE, DCAE, CNN, and VAE) is assessed across multiple X-ray datasets, and the results analyzed in relation to existing works.

This chapter's aim is to provide a comprehensive understanding of the contributions in this research. Findings are analyzed to highlight the strengths and limitations of each approach.

Medical image compression progress is required to tackle the increasing issues of storage and transmission in healthcare domain. However, the high volume of medical images demands a large storage and transmission constraints, which poses challenges for healthcare systems. Efficient compression techniques are necessary to improve transmission efficiency and reduce storage costs while maintaining crucial diagnostic quality of images. Traditional compression techniques often struggle to provide high image quality and maintain the balance with compression ratio values; to this end, machine learning and deep learning-based techniques offer high performance compression strategies depending on its architecture and datasets.

This chapter introduces the primary contributions of this research, with emphasis on developing and testing machine learning and deep learning-based techniques for compressing medical X-ray images, the contributions in this chapter span multiple areas, including:

- The application of **four deep learning models**—**CNN, AE, DCAE, and VAE**—for X-ray medical image compression.

- The use of machine learning algorithms such as Principal Component Analysis (PCA) and K-means clustering.

This research aims to contribute to the design of efficient, high-quality medical image compression techniques for improving data management in healthcare networks.

This chapter provides research that contributed to the publication of the results of the study in a published article titled "Convolutional Autoencoder-Based medical image compression using a novel annotated medical X-ray imaging dataset", throughout the Biomedical Signal Processing and Control journal, volume 94, August 2024, with the following authors: Amina Fettah, Rafik Menassel, Abdeljalil Gattal, Abdelhak Gattal [29].

Also, provides two researches of two submitted papers which are under review (one is on the LAVIA-MXID, and the other one is on the comparative study of the Convolutional autoencoder vs. RLE). Additionally, the generalizability of the different deep learning models (AE, CNN, and DCAE) on two additional datasets (OPENI and JSRT), titled “Assessing the Generalizability of Deep Learning-Based Compression Techniques for Multibodypart X-ray Medical Images: A Comparative Study “, as result of our participation in an international conference in Constantine, Algeria.

6.2 Datasets Selection and Description

6.2.1 Details of the Primary Dataset (MXID Dataset)

The dataset used in the present study consisted of 6,869 medical X-ray pictures from the MXID collection. During implementation, the images were reduced to a normal 256 × 256-pixel size to account for memory constraints. [29].

6.2.2 Description of the Two Additional Datasets Used for Validation

OPEN-I Dataset. The Indiana Network for Patient Care's de-identified OPEN-I [2 collection of chest X-rays includes 3996 radiology reports and 8121 related images. The authors carried out de-identification on both reports and photographs; manual coding of the reports enhanced retrieval precision, while automatic de-identification of the images needed manual verification and was not flawless.

JSRT Dataset. JSRT is a digital database [3] that includes 247 high quality, 12-bit grayscale chest radiographs with and without lung nodules that were collected from 14

medical facilities. The complexity of the lung nodules, which differed in characteristics, was used to classify these chest x-ray images.

6.2.3 Rationale behind Datasets Selection

The datasets used here in were selected carefully to facilitate the validation of deep learning and machine learning-based compression methods. Maintaining image quality, effectiveness of compression, and generalizability across imaging modalities are the goals of the selection.

1. **Relevance to Compression of Medical Images**
    - High-resolution X-ray images, which are often utilized in medical diagnostics and necessitate effective storage and transmission solutions, are included in the datasets.
    - X-ray images are a perfect test case for compression strategies since they must maintain important anatomical information while minimizing file size.
2. **Diversity and Generalization**
    - To evaluate the generalizability of the model across various medical imaging sources, several datasets are used.
    - A more thorough assessment of compression techniques is made possible by the datasets' variances in picture resolution, contrast levels, and anatomical areas.
3. **Standardization and Comparability**
    - The chosen datasets adhere to accepted medical imaging formats, making it easier to compare them to previous research and guaranteeing that deep learning architectures can use them.
    - They offer a regulated setting for comparing various compression techniques, such as CNN, AE, DCAE, VAE, PCA, and K-means clustering.
4. **Availability and Practical Application**
    - The datasets are made to match actual medical imaging problems or are made publicly available.
    - They are set up to facilitate tests that mimic actual transmission and storage limitations in healthcare systems.
5. **Alignment with Research Objectives**
    - By employing measurements like PSNR, MS-SSIM, and compression ratio, the datasets allow for both a quantitative and qualitative assessment of compression techniques.

- They are perfect for testing hybrid and deep learning-based methods since their selection supports the objective of enhancing medical image compression while preserving diagnostic integrity.

This research guarantees that the suggested compression strategies are both efficient and suitable to actual medical imaging scenarios by using a variety of clinically relevant datasets.

6.3 Training Setup

6.3.1 Dataset Splits

In order to enhance the efficiency of training, testing, and validation of the models, the dataset was divided automatically by applying a multiclass splitting technique. The process ensured an equal representation of class labels and maximized the sample distribution over an enormous number of subsets. 60% of the samples were assigned to the training set, 20% to the testing set, and 20% to the validation set, using a fixed split ratio. This automatic multiclass split, enriching the dataset representation, essentially captures the resulting subsets' class label distribution. The images are preprocessed by methods involving feature extraction, scaling, and normalization.

6.3.2 Loss Function

To train the models based on autoencoders, the Mean Squared Error (MSE) loss function was employed. MSE is a standard reconstruction loss function used in image compression applications because it computes the mean of the squared errors between the input image $y_{true,i}$ and reconstructed image $y_{pred,i}$. It is defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_{true,i} - y_{pred,i})^2 \qquad (6\text{-}1)$$

where n represents the number of pixels in the image. This loss encourages the autoencoder-based models to minimize the pixel-wise differences between the original and the reconstructed images, which ensures fidelity in reconstruction.

Though MSE does not directly measure perceptual quality, it is effective for measuring pixel-level similarity. In addition to this, other measurements introduced in the next section, such as Peak Signal-to-Noise Ratio (PSNR) and Multi-Scale Structural Similarity Index (MS-SSIM) were computed in training, providing feedback regarding perceptual quality and structural detail preservation.

6.3.3 Evaluation Metrics

To assess the performance of the deep learning-based model, we employed a number of quantitative metrics:

6.3.3.1 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.

$$MSE = \frac{1}{(m \times n)} \sum_{x=1}^{m} \sum_{y=1}^{n} (I(x,y) - K(x,y))^2 \qquad (6\text{-}2)$$

6.3.3.2 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.

$$MS - SSIM(x,y) = [l_m(x,y)]^{\alpha_M} . \prod_{j=1}^{M} [C_j(x,y)]^{\beta_j} . [S_j(x,y)]^{y_j} \qquad (6\text{-}3)$$

6.3.3.3 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.

$$PSNR = 20 \times log_{10}(MAX) - 10 \times log_{10}(MSE) \qquad (6\text{-}4)$$

6.3.3.4 Compression Ratio (CR)

For a high-quality picture, the Compression Ratio (CR) should be lower. The following equation is used to measure it.

$$CR = \frac{Size\ of\ Uncompressed\ Data}{Size\ of\ Compressed\ Data} \qquad (6\text{-}5)$$

- For all the dataset in machine learning techniques, we used the Average calculating

as follow:

$$Averaged\ MS-SSIM\ =\ (sum\ of\ MS-SSIM\ of\ each\ image\ )/(Number\ of\ tested\ images) \quad (6\text{-}6)$$

$$Averaged\ PSNR\ =\ (sum\ of\ PSNR\ of\ each\ image\ )/(Number\ of\ tested\ images) \quad (6\text{-}7)$$

$$Averaged\ MSE\ =\ (\ sum\ of\ MSE\ of\ each\ image\ )/(Number\ of\ tested\ images) \quad (6\text{-}8)$$

$$Averaged\ CR\ =\ (\ sum\ of\ CR\ of\ each\ image\ )/(Number\ of\ tested\ images) \quad (6\text{-}9)$$

6.4 Machine Learning and Deep Learning Techniques for Medical Image Compression: Models Overview, Experiments, and Findings

6.4.1 Machine Learning Based Approaches for Medical Image Compression

This section outlines the six primary forms of lossy image compression based on unsupervised learning: Principal Component Analysis (PCA), K-Means clustering, Convolutional Neural Networks (CNNs), Deep Convolutional Autoencoders (DCAEs), Autoencoders (AEs), and Variational Autoencoders (VAEs).

6.4.1.1 Principal Component Analysis (PCA)

PCA is a dimensionality reduction technique that aims to keep as much of the original variance as possible while converting high-dimensional data to a lower-dimensional representation. PCA can aid in picture compression by reducing the amount of data required to represent an image while retaining its essential components.

6.4.1.2 K-Means

One more method for image compression is k-means clustering. A procedure of unsupervised learning, called K-means clustering, divides data into a predefined number of clusters based on the degree of similarity of the data to one another. Clustering with various values of k will give varying outcomes in the output finally [79]. K-means can be used to combine similar colors and represent an image with fewer color clusters for image

reduction.

6.4.1.3 Experiments and Findings Based on MXID Dataset

6.4.1.3.1 First experiment: Unsupervised learning techniques using only one image

In the first experiment, one single medical image out of the newly proposed MXID dataset in 256 x 256 PNG pixels was used for comparison between the PCA-based lossy compression approach and the K-means-based lossy compression approach. For the unusual application of pulmonary imaging, where strict isolation of structures within the lung is critical in diagnosis, an image used within this study **Figure 6-3** was chosen as a sample to demonstrate the ability of our compression method. In trying to compress well and preserve fine features, the complex structures of lungs, including the blood vessels, airways, and lung tissues, must be handled carefully in the process of compression. This decision was made in order to draw attention to the strengths and weaknesses of the unsupervised learning strategies that were being considered. The quality of the compressed images was assessed using the MS-SSIM, compression ratio, PSNR, and MSE metrics. The performance of the PCA-based and K-means-based approaches with different parameters is shown in **Figure 6-1** and **Figure 6-2**, respectively.

In medical image compression, choosing the ideal K value for K-means clustering is a challenging problem that varies from picture to image and is dependent on a number of variables. The results for the chosen image show the efficacy of the K-means-based approach with k=64 and the PCA-based method with the number of components set between 400 and 500. These results are representative of the features of the selected image.

Grayscale medical images, for example, found a compromise between compression efficiency and image quality at moderate K values like 16 to 32, but binary images, such as bone X-rays, profited from a K value of 2 for foreground-background separation in this proposed dataset. Even though this leads to bigger file sizes, higher K values (e.g., K=64 or more) may be chosen in situations where fine detail preservation is necessary (e.g., lung, abdomen).

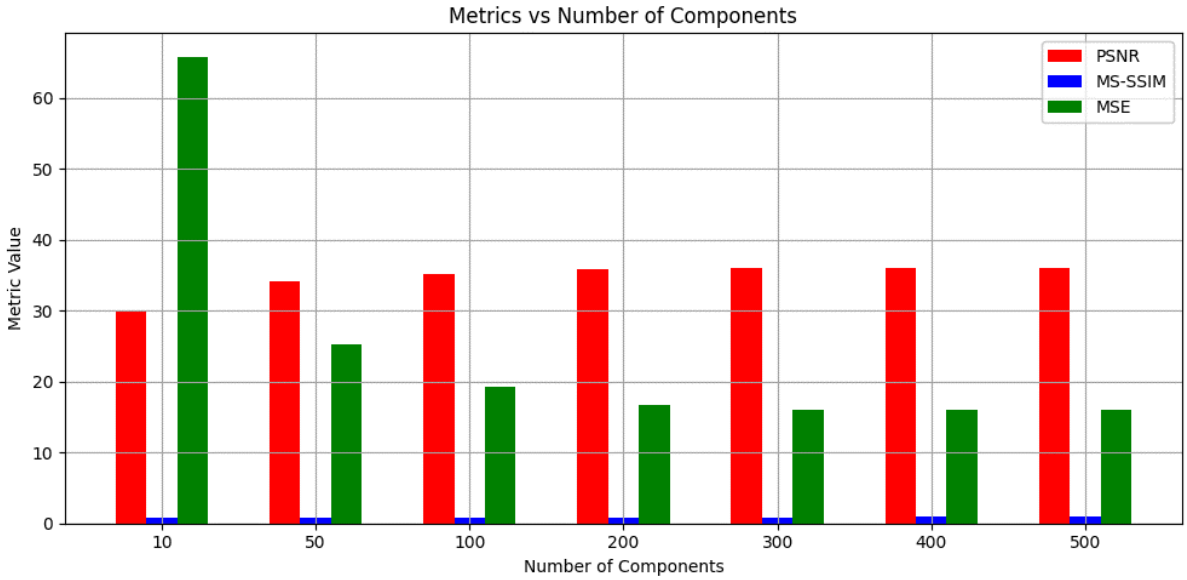


Figure 6-1 Quantitative Assessment of PSNR, MSE, and MS-SSIM for the PCA-Based Lossy Compression Method. The Reported Outcomes Stem from Methodical Experimentation, Offering Insights into Performance Dynamics Across Different Principal Components *[29]*

Table 6-1 Comprehensive Comparison of Single Image Compression Results for PCA-Based and K-means-Based Lossy Compression *[29]*

Methods	Parameters	MSE	PSNR	CR	MS-SSIM
PCA	10	65.88	29.94	-	0.77
	50	25.34	34.09	-	0.83
	100	19.34	35.27	-	0.87
	200	16.71	35.90	-	0.89
	300	16.11	36.06	-	0.89
	400	15.97	36.10	-	0.89
	500	**15.97**	**36.10**	**-**	**0.89**
K-means	2	1062.82	17.87	99.15%	0.39
	4	187.88	25.39	98.31%	0.63
	8	46.94	31.42	96.61%	0.83
	16	11.40	37.56	93.22%	0.95
	32	3.09	43.23	86.44%	0.98
	48	1.47	46.46	79.66%	0.99
	64	**1.00**	**48.11**	**72.88%**	**0.99**

The data presented in **Figure 6-2** illustrate many performance indicators for the selected image. K-means outperformed Ahuja and Doriya's results in terms of image quality, as seen by greater PSNR and MS-SSIM values for k=64 [40]. In addition, **Table 6-1** gives the complete findings of the initial experiments' performances for both k-means and PCA.

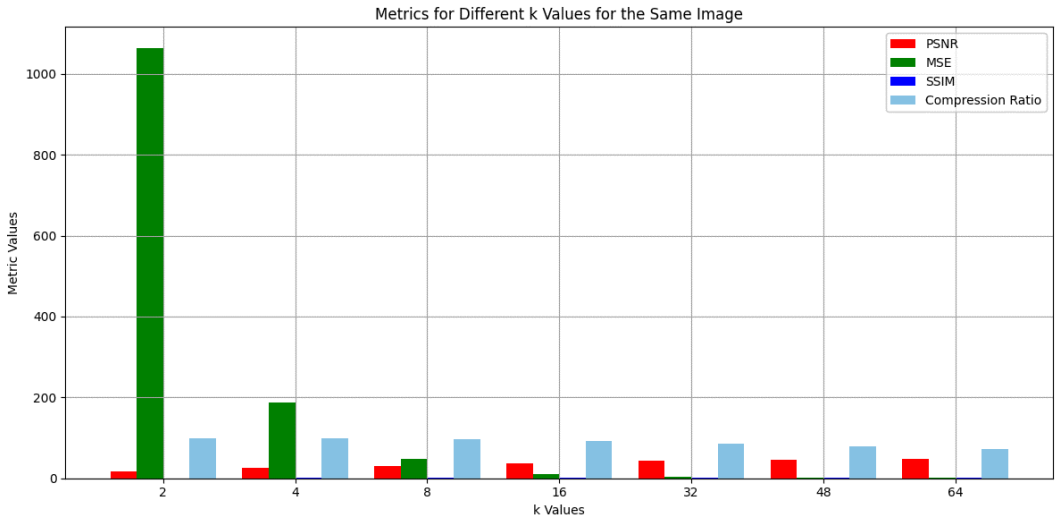
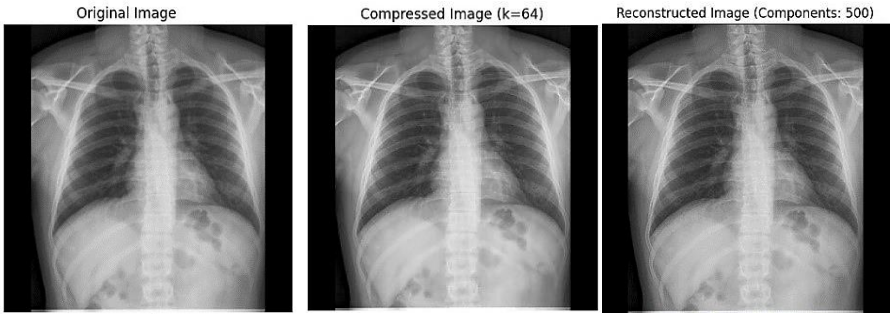


Figure 6-2 Single-Image Evaluation of PSNR, MSE, and MS-SSIM for the Kmeans-Based Lossy Compression Method *[29]*



(a) **(b)** **(c)**

Figure 6-3 Result of Machine learning and Deep Learning algorithms (a) Input medical image considered for evaluation, (b) Compressed image using PCA-based lossy compression where components = 500, and (c) Compressed image using K-MEANS-based lossy compression where k=64 *[29]*

6.4.1.3.2 Second Experiment: Unsupervised learning techniques using all images

A total of 6869 medical images from the proposed MXID dataset in PNG format, sized 256 x 256 pixels, are utilized in the second experiment to compare the performance of the various approaches. Due to available computing resources, i.e., GPU memory, the batch size is four. A normalization stage subsequent to rescales pixel values to an interval of [0, 1] so that the image dataset is well ordered and prepared for input. The image qualities compressed are

approximated in terms of the compression ratio, MS-SSIM, PSNR, and MSE.

The results of the PCA study offer important new information about the effect of dimensionality reduction on the various quality measures. Also, some PCA components' mean evaluation metric values (10, 50, 100, 150, 200, and 250) offer useful trade-offs between image quality and changes in the growth of compression levels.

First, from **Table 6-2** it can be seen that increased numbers of components result in improved mean PSNR, MSE, and MS-SSIM significantly, demonstrating superior reconstruction accuracy**. Figure 6-4** illustrates trends against the number of components. There is consistent performance with better-quality images, indicated by superior PSNR, decreasing MSE, and extremely high structure similarity, displaying pixel-wise error was nil.

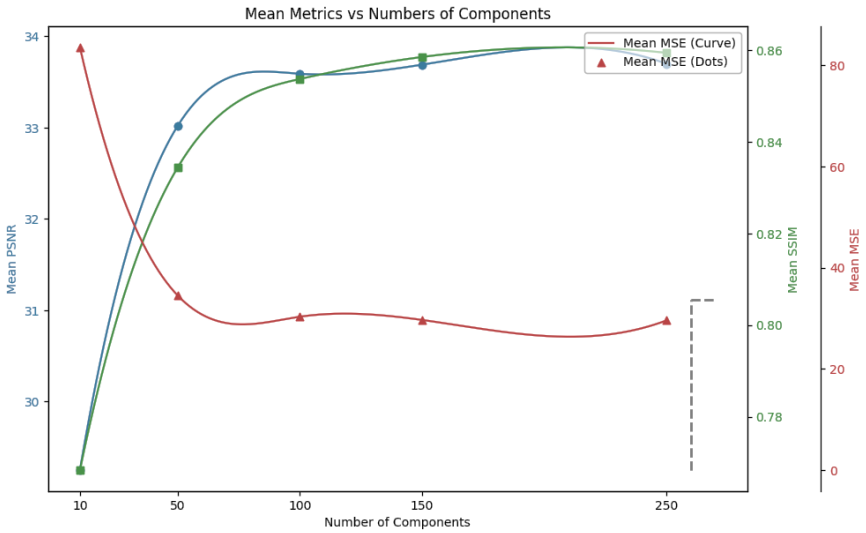


Figure 6-4 Evaluation of PSNR, MSE, and MS-SSIM for the PCA-Based Lossy Compression Method, showcasing performance across varying principal components (10, 50, 100, 150, 250) *[29]*

Table 6-2 Comparative Compression Results for K-means-Based Lossy Compression Methods Across Various Clustering Levels *[29]*

Methods	Parameters	Mean MSE	Mean PSNR	Mean CR	Mean MS-SSIM
PCA	**10**	83.56	29.25	-	0.77
	50	34.48	33.02	-	0.83
	100	30.29	33.59	-	0.85
	150	29.65	33.69	-	0.86
	250	**29.56**	**33.70**	-	**0.86**
Kmeans	**2**	756.90	19.60	99.11%	0.43
	4	156.78	26.34	98.23%	0.67
	8	36.99	32.60	96.46%	0.87
	16	9.42	38.53	92.91%	0.95
	32	2.57	44.16	85.82%	0.98
	48	1.28	47.19	78.73%	0.99
	64	**0.81**	**49.16**	**71.64%**	**0.99**

The results obtained for 250 elements were consistent with previous ones presented by Elamparuthi and were 33.70 dB in terms of PSNR. Consistency of these results with those presented by Elamparuthi et al. [80] suggests there is strong consistency between the two studies, which validates the given results in terms of validity and consistency. Moreover, experimental results indicate that the methods presented in [43], [81] demonstrated superior performance against numerous state-of-the-art compression techniques, particularly in segmentation, compression ratios, and overall performance. Even though X-ray scans come with a higher PSNR, they are less efficient when compressed.

In medical image compression, k-means clustering is crucial; a rise in k number enhances grouping tightness, compression, and image quality. Higher values of PSNR and MS-SSIM as discovered in **Table 6-2** indicate that this can potentially preserve very critical diagnostic information. We also selected the values for k in this data after careful consideration on image quality and compression in a hospital setting.

K=2 means a binary option, which will produce a higher compression ratio but lower image quality as there are fewer intensity values. If we consider k=4.8, then we began seeing an increasingly better image of images, and we ended up compromising on image quality over compression efficiency that can save time and space but compromise diagnostically important information.

Moving to k=16 improves picture quality by dividing the data into numerous clusters, resulting in greater MS-SSIM and PSNR values of 38.53 dB compared to [41], as well as lower MSE and CR values, indicating more dependable findings. Finally, k=64 provides the best amount of clustering among the tested options, if correct diagnosis is critical and there are sufficient storage or transmission resources. However, as the value of 'k' grows so does the processing time. **Figure 6-5** demonstrates an apparent pattern between the various assessment metrics.

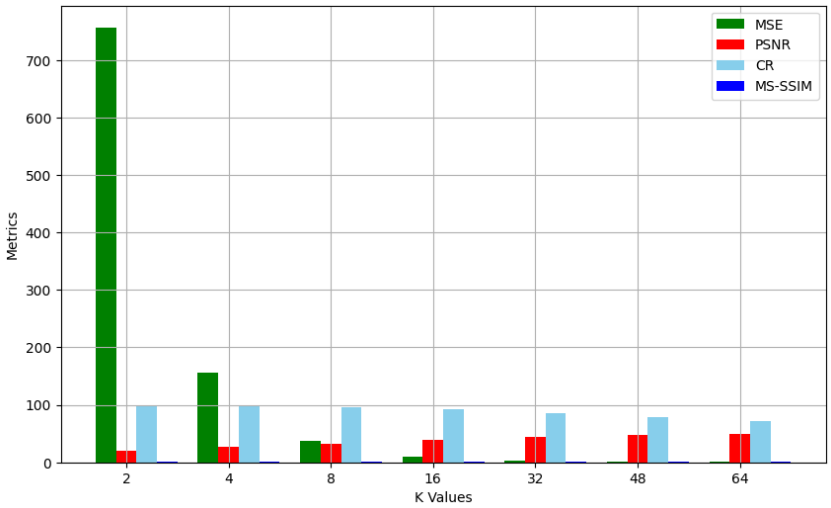


Figure 6-5 PSNR, MSE, and MS-SSIM Evaluation for Kmeans-Based Lossy Compression Method at Different Clustering Levels (k=2, k=4, k=8, k=16, k=32, k=48, k=64) *[29]*

The following section will include an in-depth discussion of the deep learning models employed, and the training setup, methodologies, and parameters used to achieve optimal performance.

6.4.2 Deep Learning Based Models for Medical Image Compression

6.4.2.1 Overview of Deep Learning Models: AE, CNN, DCAE, VAE, and CAE

6.4.2.1.1 Autoencoders (AEs)

The suggested Autoencoder (AE) model, shown in **Figure 6-6**, follows the guidelines of a basic simple feed forward autoencoder and employs a single-channel grayscale input format with 256 × 256-pixel resolution, according to standards used in research such as [82]. The architecture begins by flattening the input images to a 1D vector and then two dense feature extraction layers of 128 and 64 units, respectively, are applied to this vector with the Rectified Linear Unit (ReLU) activation function, as demonstrated in **Table 6-3**. A popular selection for the imposition of non-linearity in neural networks for increased training speed and better generalization over the sigmoid function [83].

Simultaneously, compressed representation, as a 64-bit 1D vector, is calculated using two thick layers of 128 and 256x256x1. ReLU activation for the first layer and Sigmoid activation in the final layer are generally used to reconstruct images, as detailed in **Table 6-4**; it increases an image's clarity and gives a probabilistic explanation, where the output appears as the probability of the activation of the pixel [27]. With past experience incorporated into lessons learned from existing work, our approach effectively learns and representations salient features in input data and remedies some compression issues.

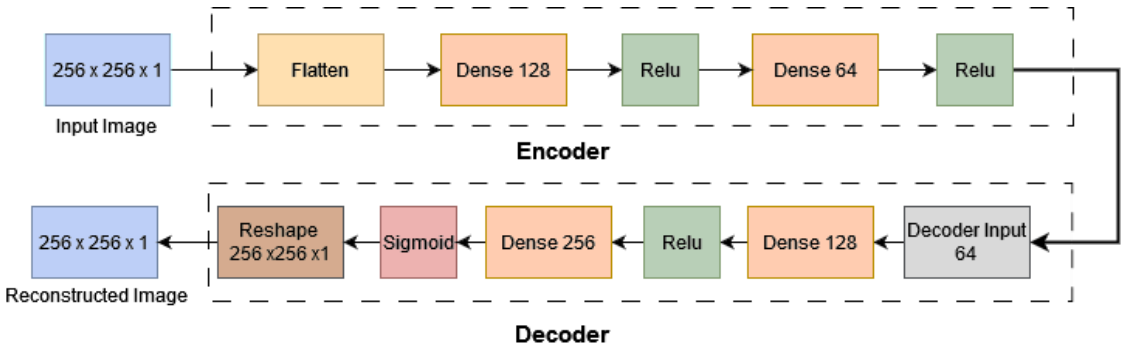


Figure 6-6 Architectural Diagram of the Proposed Autoencoder for Image Compression

6.4.2.1.1.1 Autoencoder's architecture details

Encoder

Table 6-3 AE Encoder's Architecture Details

Layer	Type	Output Shape	Activation	Description
Input	Input Layer	(256, 256, 1)	-	Grayscale medical image input
Flatten	Flatten	(65,536)	-	Converts 2D image into 1D vector
Dense 1	Fully connected	(128)	RELU	Reduces dimensionality while retaining features
Dense 2	Fully connected	(64)	RELU	Compresses features into a **64-dimensional latent space**

Decoder

Table 6-4 AE Decoder's Architecture Details

Layer	Type	Output Shape	Activation	Description
Input	Input Layer	(64)	-	Grayscale medical image input
Dense 3	Fully connected	(128)	-	Converts 2D image into 1D vector
Dense 4	Fully connected	(65,536)	RELU	Reduces dimensionality while retaining features
Reshape	Reshape	(256, 256, 1)	RELU	Compresses features into a **64-dimensional latent space**

6.4.2.1.2 Convolutional Neural Network (CNN)

As far as maintaining image quality without compromising file size, Convolutional Neural Network (CNN) image compression algorithms tend to perform better than conventional techniques such as JPEG. These kinds of CNN-based algorithms are especially ideal for those

applications that require high image integrity, e.g., commercial photo, remote sensing, and medical imaging. The developed CNN model is composed of an encoder and decoder. The encoder applied convolutional layers, batch normalization, and pool layers to reduce the dimension of the medical image presented in **Table 6-5,** before a follow-up dense layer which compressed the relevant features to a latent space with 128 dimensions. The compressed representation was then reconstructed by the decoder into a compressed picture using upsampling and convolutional layers, as shown in **Table 6-6**, with the fundamental architecture optimized to maintain significant spatial properties in medical images.

6.4.2.1.2.1 Convolutional Neural Network (CNN) architecture's details

The architecture of the Convolutional Neural Network (CNN) encoder module is to effectively learn hierarchical features from the input data. The architecture consists of a sequence of convolutional layers, batch normalization layers, and max-pooling layers that work together to learn spatial hierarchies and down sample. **Table 6-5** provides a comprehensive description of the encoder module, including the layer types, activation functions, feature maps, and convolutional kernel sizes applied at each level.

Table 6-5 An Overview of the Encoder Component of the Convolutional Neural Network (CNN)

NODE	LAYER TYPE	ACTIVATION FUNCTION	FEATURE MAPS	CONVOLUTIONAL KERNEL
encoder input	Input	-	-	(256, 256, 1)
conv2d	Convolutional	ReLU	64	(3,3)
batch_norm	Batch Normalization	-	-	-
max_pooling	Max Pooling	-	-	(2,2)
conv2d_1	Convolutional	ReLU	128	(3,3)
batch_norm_1	Batch Normalization	-	-	-
max_pooling_1	Max Pooling	-	-	(2,2)
conv2d_2	Convolutional	ReLU	256	(3,3)
batch_norm_2	Batch Normalization	-	-	-
max_pooling_2	Max Pooling	-	-	(2,2)
flatten	Flatten	-	-	-

Encoder output	Dense	ReLU	128	-

Table 6-6 Overview of the Decoder Component of a Convolutional Neural Network (CNN)

NODE	LAYER TYPE	ACTIVATION FUNCTION	FEATURE MAPS	CONVOLUTIONAL KERNEL
decoder_input	Input	-	-	(128,)
dense_1	Dense	ReLU	262144	-
reshape	Reshape	-	(32, 32, 256)	-
conv2d_transpose_1	Conv2DTranspose	ReLU	256	(3,3)
up_sampling_1	UpSampling2D	-	-	(2,2)
batch_norm_3	Batch Normalization	-	-	-
conv2d_transpose_2	Conv2DTranspose	ReLU	128	(3,3)
up_sampling_2	UpSampling2D	-	-	(2,2)
batch_norm_4	Batch Normalization	-	-	-
conv2d_transpose_3	Conv2DTranspose	ReLU	64	(3,3)
up_sampling_3	UpSampling2D	-	-	(2,2)
batch_norm_5	Batch Normalization	-	-	-
conv2d_transpose_4	Conv2DTranspose	Sigmoid	1	(3,3)

Table 6-6 is a representation of the decoder part of the Convolutional Neural Network (CNN) model. It shows the most significant layers, activation functions, and feature map transformations while decoding. The decoder builds the image from compressed shape systematically in layers such as dense, convolutional transpose, and up sampling to increase the resolution of output gradually up to the final image.

6.4.2.1.3 Variational Autoencoder (VAE)

Variational Autoencoder (VAE) is a generative model that combines autoencoder and probabilistic modeling components and is designed for feature representation and unsupervised learning purposes. It comprises of an encoder and a decoder, as shown in **Figure 6-7**. This concept gives the encoding process a probabilistic twist. Instead of immediately converting the input data into a fixed-size latent representation, VAEs

convert it to a probability distribution in the latent space. As a result, VAEs may generate varied and useful samples from data distributions while also performing exact reconstructions. Unlike traditional compression approaches that depend on heuristics, VAEs learn a probabilistic model of data distribution, allowing them to preserve important visual features while reducing pictures into compact latent representations.

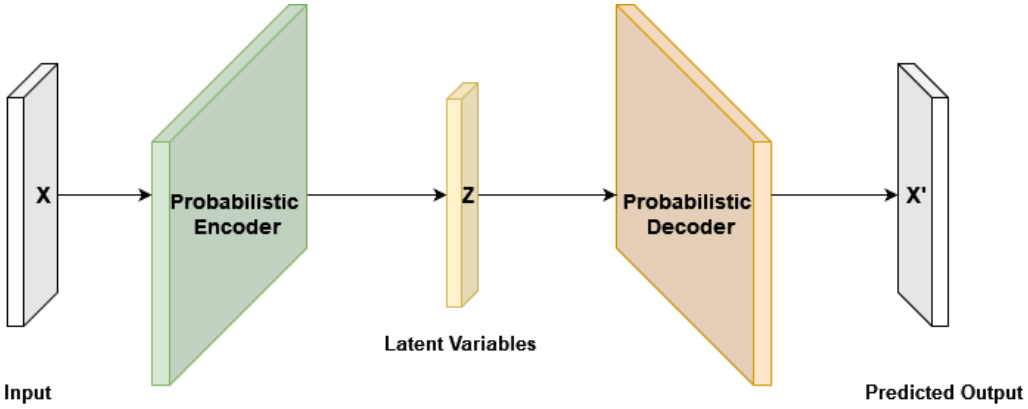


Figure 6-7 Variational Autoencoder (VAE) Architectural Overview *[29]*

6.4.2.1.3.1.1 Variational Autoencoder (VAE) architecture's details

Encoder

Table 6-7 VAE Encoder's Details

LAYER	TYPE	OUTPUT SHAPE	ACTIVATION	DESCRIPTION
INPUT	Input Layer	(256 × 256)	-	Flattened grayscale image input
DENSE 1	Fully Connected	(1000)	**RELU**	Reduces dimensionality while extracting features
DENSE 2	Fully Connected (μ)	(32)	-	Mean (μ) of the latent space distribution
DENSE 3	Fully Connected ($\log\sigma^2$)	(32)	-	Log variance ($\log\sigma^2$) for reparameterization

Latent Space

Table 6-8 VAE Latent space's Description

OPERATION	DESCRIPTION

REPARAMETERIZATION TRICK	Samples latent vector **z** using $z = \mu + \epsilon \times \sigma, where\ \epsilon \sim N(0,1)$

Decoder

Table 6-9 VAE Decoder's Details

LAYER	TYPE	OUTPUT SHAPE	ACTIVATION	DESCRIPTION
DENSE 4	Fully Connected	(1000)	**RELU**	Expands the latent vector
DENSE 5	Fully Connected	(256 × 256)	**Sigmoid**	Outputs reconstructed image with pixel values between **0 and 1**

Loss Functions

Table 6-10VAE's Loss Functions

LOSS COMPONENT	DESCRIPTION
RECONSTRUCTION LOSS (MSE)	Measures the difference between the original and reconstructed images
KL DIVERGENCE LOSS	Regularizes the latent space to follow a normal distribution

6.4.2.1.4 Proposed Approach Deep Convolutional Autoencoder (DCAE) Architecture

The deep convolutional autoencoder architecture was particularly intended to recreate X-ray medical pictures while retaining essential diagnostic information. The architecture is made up of two connected components: the decoder and the encoder. **Figure 6-8** shows how the encoder extracts detailed picture characteristics from the MXID dataset using a convolutional layer with a kernel size of 3 × 3, 128 filters, as demonstrated in **Table 6-12**, with a ReLU activation function.

The decoder in **Table 6-13**, on the other hand, converts compressed information into a recognizable picture and has the same structure as the encoder [84]. Starts with a 3x3 kernel convolutional layer with 32 filters to reinforce the compressed representation and prepare data for subsequent upsampling, which is consistent with comparable techniques [85]. These 'UpSampling2D' layers, which are 2x2 in size, efficiently recover spatial dimensions and

delicate characteristics by reversing the impact of maxpooling applied during encoding. The reconstruction is then reinforced by two further convolutional layers, each with 64 and 128 filters, then ReLU activation is used to boost accuracy. Lastly, the final convolutional layer uses the'sigmoid' activation function with a 3x3 kernel and one filter to make sure pixel values in the reconstructed image are in the correct range. Parametric Rectified Linear Unit (PReLU) and ReLU can be effective activation functions for image compression when used with CAEs. However, the ReLU activation function has been shown to possess the capability of speeding up deep learning training for quite some time now, and proven to outperform traditional sigmoid functions in terms of gradient dispersion [86]. Adoption of the PReLU function in this work has been motivated by the work of Cheng et al., [55]. The activation function has proven to be good at extracting additional information and producing improved results.

This architecture effectively meets the criteria of X-ray medical image compression, successfully reducing dimensionality in the encoder while keeping essential diagnostic properties and speedily reconstructing pictures in the decoder. The whole architecture shown in **Figure 6-8** demonstrates a strong dedication to minimize data loss during the compression phase, making it an attractive alternative for healthcare applications requiring high picture resolution and resource economy.

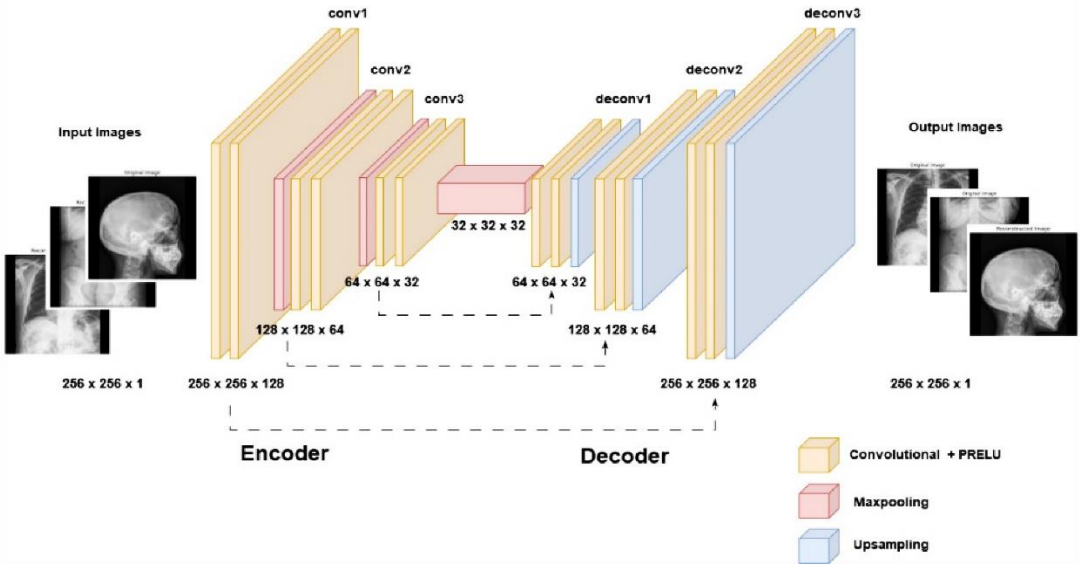


Figure 6-8 Proposed Deep Convolutional Autoencoder (DCAE) Architecture for Image Compression Using TensorFlow and Keras

Table 6-11Pseudo code of the model's training algorithm

Algorithm

- Initialize the weights and biases of the DCAE model using the initializer ***I***.
- Set the optimizer ***O*** with initial learning rate ***r***.
- **For** each epoch **e = 1** to ***E*** **do**
  a. Shuffle the dataset ***X***.
  b. **For** each batch $x \subset X$ of size ***B*** **do**
  i. Forward propagate the batch ***x*** through the encoder to obtain the latent representation ***z***.
  ii. Decode ***z*** using the decoder to generate the reconstructed output $\hat{x}$.
  iii. Compute the reconstruction loss $L(x, \hat{x})$.
  iv. Back-propagate the loss ***L*** through the network.
  v. Update the weights and biases using the optimizer ***O***.
  c. Optionally adjust learning rate ***r*** based on schedule or validation performance.
  d. Save model weights if current loss is the best so far.
- **End For (epochs)**
- **Return** the final reconstructed images $\hat{X}$.

6.4.2.1.4.1 Deep Convolutional Autoencoder (DCAE) Architecture's Details

A. With RELU Activation Function

Encoder

Table 6-12 DCAE's Encoder Details

Layer	*Type*	*Output Shape*	*Activation*	*Description*
Input	Input Layer	(256, 256, 1)	-	Grayscale image input
Conv 1	Convolutional (128 filters, 3×3)	(256, 256, 128)	**RELU**	Extracts low-level features
Maxpool 1	Max Pooling (2×2)	(128, 128, 128)	-	Reduces spatial dimensions
Conv 2	Convolutional (64 filters, 3×3)	(128, 128, 64)	**RELU**	Extracts deeper features
Maxpool 2	Max Pooling (2×2)	(64, 64, 64)	-	Further reduces spatial dimensions
Conv 3	Convolutional (32 filters, 3×3)	(64, 64, 32)	**RELU**	Encodes high-level representations
Maxpool 3	Max Pooling (2×2)	(32, 32, 32)	-	Latent representation

Decoder

Table 6-13 DCAE's Decoder Details

Layer	*Type*	*Output Shape*	*Activation*	*Description*
Conv 4	Convolutional (32 filters, 3×3)	(32, 32, 32)	**RELU**	Begins decoding process
Upsmaple 1	UpSampling (2×2)	(64, 64, 32)	-	Upscales spatial dimensions
Conv 5	Convolutional (64 filters, 3×3)	(64, 64, 64)	**RELU**	Restores mid-level features
Upsmaple 2	UpSampling (2×2)	(128, 128, 64)	-	Further upscales
Conv 6	Convolutional (128 filters, 3×3)	(128, 128, 128)	**RELU**	Restores low-level details
Upsmaple 3	UpSampling (2×2)	(256, 256, 128)	-	Final upscaling
Conv 7	Convolutional (1 filter, 3×3)	(256, 256, 1)	**Sigmoid**	Outputs reconstructed grayscale image

B. With PRELU Activation Function

Similar to the DCAE model the only changeset is in the activation function (PRELU).

6.4.2.1.5 Convolutional Autoencoder (CAE) Architecture

a) **Encoder :**

1. **Convolutional Layers:** The encoder starts with a 256 x 256 grayscale input image (single channel) and goes on to use three convolutional layers, **Table 6-14** presents layer type, output shape, activation function, and description of the encoder. The first layer has 32 filters with a 3 × 3 kernel and ReLU activation function. The second layer has 64 filters with the same kernel and activation function, and the third layer has 128 filters with the same kernel and ReLU activation function. In addition, the padding capability is employed to preserve the original dimensions without compromising the necessary picture elements.
2. **MaxPooling Layers:** Following each convolutional layer, a 2 × 2 Maxpooling layer is used to minimize spatial dimensions while preserving important x-ray picture properties.

3. **Global Average Pooling:** This is a critical layer in this autoencoder. After the last convolutional layer, this layer converts the feature maps into a single vector to minimize overfitting, reduces the number of parameters, and improves the model's generalizability.
4. **Dense Layer:** The output of the global average pooling layer is then passed through a dense layer of 128 units, resulting in the encoded output.

Table 6-14CAE Encoder's Details

Layer	Type	Output Shape	Activation	Description
Input Layer	Input (256 × 256 × 1)	(256, 256, 1)	-	Grayscale X-ray image
Conv 1	Conv2D (32 filters, 3×3)	(256, 256, 32)	RELU	Extracts low-level features
Maxpooling 1	MaxPooling2D (2×2)	(128, 128, 32)	-	Reduces spatial dimensions
Conv 2	Conv2D (64 filters, 3×3)	(128, 128, 64)	RELU	Extracts deeper features
Maxpooling 2	MaxPooling2D (2×2)	(64, 64, 64)	-	Further reduces spatial dimensions
Conv 3	Conv2D (128 filters, 3×3)	(64, 64, 128)	RELU	Extracts high-level representations
Maxpooling 3	MaxPooling2D (2×2)	(32, 32, 128)	-	Final spatial reduction
Global AVG Pooling	GlobalAveragePooling2D	(128)	-	Reduces overfitting, improves generalization
Dense Layer	Dense (128 units)	(128)	-	Encoded latent space

b) **Decoder:**

1. **Reshaping (4 × 4 × 8):** Following the dense layer, a reshaping layer is required to prepare the encoded output for upsampling, as demonstrated in **Table 6-15**.
2. **Convolutional and UpSampling Layers:** It is composed of up of three convolutional layers (128, 64, and 32) with a ReLU activation function and upsampling, which mimic the encoder's structure.

3. **Output Layer:** This layer consists of a single convolutional filter with a 3 × 3 kernel and utilizes a sigmoid activation function.

Table 6-15 CAE Decoder's Details

Layer	Type	Output Shape	Activation	Description
ReshapeLayer	Reshape (4 × 4 × 8)	(4, 4, 8)	-	Prepares for upsampling
Conv 4	Conv2D (128 filters, 3×3)	(4, 4, 128)	RELU	Starts decoding
Upsampling 1	UpSampling2D (2×2)	(8, 8, 128)	-	Upscales spatial dimensions
Conv 5	Conv2D (64 filters, 3×3)	(8, 8, 64)	RELU	Intermediate reconstruction
Upsampling 2	UpSampling2D (2×2)	(16, 16, 64)	-	Further upscaling
Conv 6	Conv2D (32 filters, 3×3)	(16, 16, 32)	RELU	Restores finer details
Upsampling 3	UpSampling2D (2×2)	(256, 256, 32)	-	Final upscaling
Output Layer	Conv2D (1 filter, 3×3)	(256, 256, 1)	Sigmoid	Generates final reconstructed image

This convolutional autoencoder was designed to take an original picture, compress it into a lower-dimensional latent space in order to retain all of the critical data, and then rebuild the encoded image as closely as possible to the original while maintaining quality and minimizing size. **Figure 6-9** illustrates this architecture.

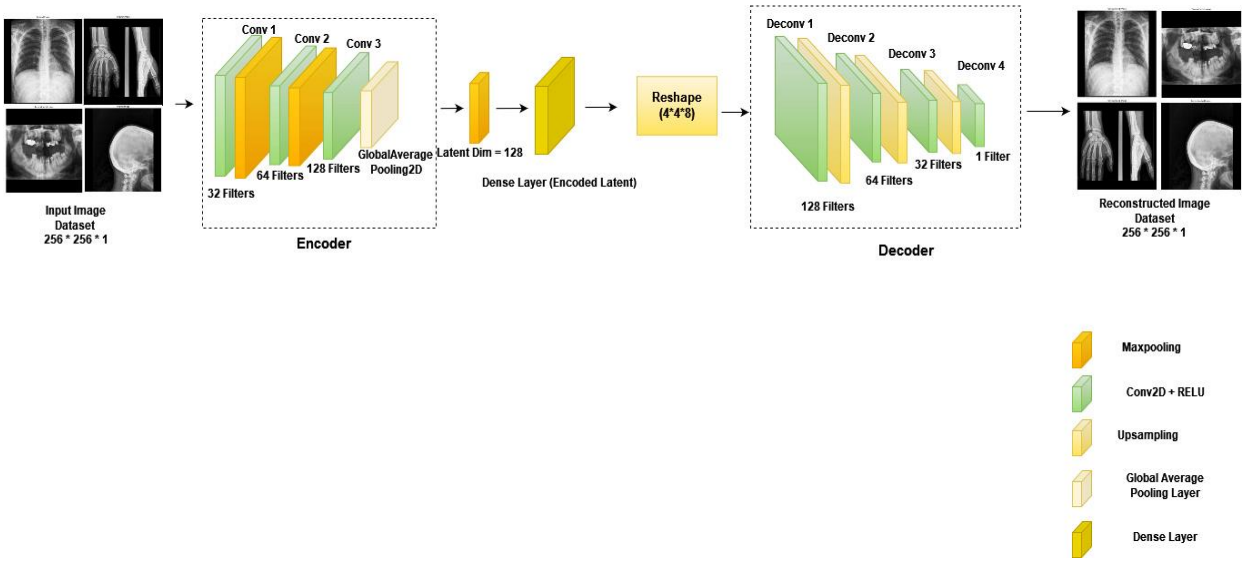


Figure 6-9 Convolutional Autoencoder Architecture for X-ray Medical Image Compression

6.4.2.1.6 RLE

The RLE architecture functions in the following way:

Starting from a 2D grayscale image, as illustrated in **Figure 6-10**, the encoding transforms it into a 1D array for easier processing. The 1D array is then compressed using RLE, resulting in a list of value/count pairs. The compressed data is decompressed back into a 1D array and reshaped into its original 2D grayscale picture format.

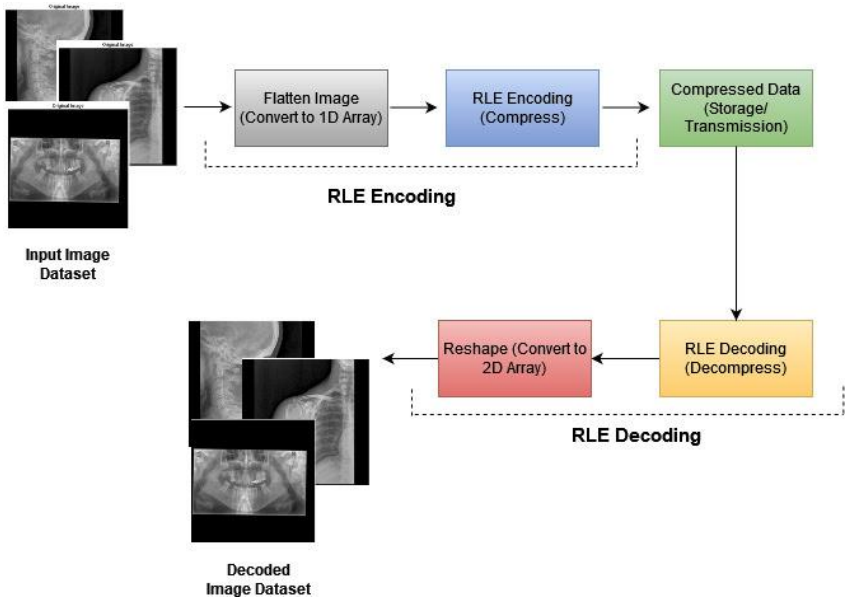


Figure 6-10 RLE Architecture

6.4.2.1.7 Hybridization of both CAE and RLE Techniques (CAE-RLE)

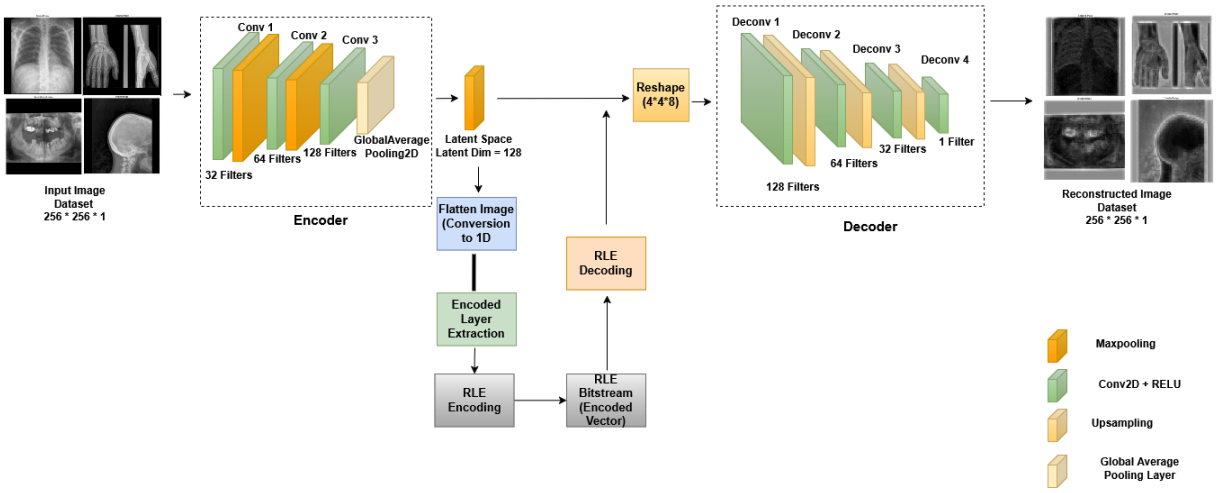


Figure 6-11 Hybridization of the CAE-RLE for medical image reconstruction

6.4.2.1.7.1 CAE-RLE Architecture Details

The purpose of this architecture to reduce the dimensionality of the image while maintaining the critical features. This hybrid approach presented in **Figure 6-11** was selected with the aim of achieving higher compression efficiency in comparison with CAE and RLE individually while preserving important regions and key features of the medical images. If we assume that latent space has several repeated values, RLE can reduce the size of the compressed latent vector by encoding only the value and its count rather than storing each pixel value individually. Applying RLE after the CAE encoding process ensures that no important information is lost in the compression process as the loss of small details impact diagnosis accuracy, as long as the latent space has sufficient redundancy. Additionally, our aim is to reduce file size while obtaining an optimized compression, which means to maximize data size reduction without sacrificing image quality.

Although the hybridization of CAE and RLE is theoretically sound, poor quality reconstructed images may arise due to lack of sufficient redundancy from the latent vectors that were generated by the CAE; Also, the latent space could be varied and dense, in which the RLE might not perform as well as expected, or even worse; it might remove critical information leading to blurry reconstructions. Additionally, the decoding process in this architecture may not be able to recover the original image quality from the compressed latent space if the RLE has compressed the latent space in an unsuitable way, which impact the reconstructed image quality, **Table 6-16** Present a comparative discussion of the methods: CAE, RLE, CAE-RLE, respectively.

In conclusion, this hybrid CAE-RLE architecture presents a limitation in this research, while

the CAE provided a good balance between image quality and compression. In addition, RLE works effectively for repetitive data and given a good result in comparison to the CAE-RLE outcomes in terms of image quality. Overall, this hybrid technique faces challenges due to mismatches of handling the latent space in each method. To improve this hybrid CAE-RLE method further optimization of the latent space representation or switching to other hybrid compression technique could be more suited for the preservation of the crucial information in the medical image compression domain.

Table 6-16 Comparison of the AE, RLE, and AE-RLE Techniques

Features	***CAE***	***RLE***	***CAE-RLE***
Compression Type	Lossy	Lossless	Hybrid (Lossy+Lossles)
Compression Effectiveness	Moderate to High	Good for redundant Patterns Only	
Visual Quality	Minor Quality Loss	Preserved Exactly	Noticeable Quality Loss
Suitability for Medical Imaging	Risk if overly lossy	Integrity Preserved	Blanced : AE preserves crucial features, while RLE can cause distortion
Speed	Requires Time for The Training Process	Very Fast	Moderate (Training + fast encoding)
Storage Saving	Depends on Latent Size	Depends on Pixel redundancy	Generally best savings due to double-stage reduction

6.4.2.2 Hyperparameters Configuration

Table 6-17 outlines the hyperparameter settings used across various deep learning techniques used in this study. The parameters include the initial learning rate, batch size, epochs, number of trainable parameters, optimizer, error function, and activation functions. The configuration choice is critical in optimizing the performance of the specific models, and the table specifies the specific setting used for Autoencoders (AE), Convolutional Neural Networks (CNN), Variational Autoencoders (VAE), Deep Convolutional Autoencoders

(DCAE), and Convolutional Autoencoders (CAE).

-Adam Optimizer: The training process employed the Adam optimizer, which is a stochastic gradient descent optimization algorithm that learns the learning rate for each parameter independently. Adam is an extension of the advantages of two very successful stochastic gradient descent extensions. Adam estimates the first moment (mean) and second moment (uncentered variance) of the gradients for each parameter in order to create adaptive learning rates.

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \qquad (6\text{-}10)$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g^2{}_t \qquad (6\text{-}11)$$

Where:

- β_1 and β_2 are exponential decay rates for the moment estimates.

In this work, the Adam optimizer was initialized with a learning rate of 0.001, which was an optimal balance between training speed and stability.

Table 6-17 Hyperparameters Settings across AE, CNN, VAE, DCAE, and CAE models

Technique Used	Initial Learning Rate	Batch Size	Epochs	Trainable Parameters	Optimizer	Error Function	Activation Function
Autoencoder (AE)	0.001	8	Fixed in 50	16859456	Adam Optimizer	Mean Square Error (MSE)	RELU, Sigmoid
Convolutional Neural Network (CNN)	0.001	4	Fixed in 50	1008449	Adam Optimizer	Mean Square Error (MSE)	ReLU
Variational Autoencoder (VAE)	0.001	8	Fixed in 100	-	Adam Optimizer	Mean Square Error (MSE)	
Deep Convolutional Autoencoder (DCAE)	0.001	4	Fixed in 100	12156609	Adam Optimizer	Mean Square Error (MSE)	ReLU and PReLU
Convolutional Autoencoder (CAE)	0.001	4	Fixed in 200	332801	Adam Optimizer	Mean Square Error (MSE)	RELU, Sigmoid

The following tests were carried out using Python 3.7.16, TensorFlow 2.7.0, and Keras API on GPU NVIDIA and 16 GB RAM. For each training iteration, we kept the best outputs, as well as the improved DCAE model weights and biases, using Keras callbacks such as

ModelCheckpoint. To stress the capability of the autoencoder to learn effective latent representations from medical images, we conducted an experiment where the model was trained under various settings: fifty epochs, one hundred epochs... These experiments were performed with consistent training parameters to investigate the improvement in performance with time. In the following section, the results of reconstruction achieved by our model, initially trained on our well-curated X-ray dataset, are presented.

6.4.2.3 Experiments and Findings Based on MXID Dataset

6.4.2.3.1 Third Experiment: Deep Learning Models: CNN, AE, DCAE, and VAE

- Autoencoder (AE): The implementation of the autoencoder on the MXID dataset yielded a PSNR of 37.61 dB, an MSE of 0.014, and an MS-SSIM of 0.61 in [29], as shown in **Figure 6-12**. These values indicate a significant loss of detail, which is crucial for medical imaging applications. While the primary structural integrity of the images remains intact, as depicted in **Figure 6-20**.

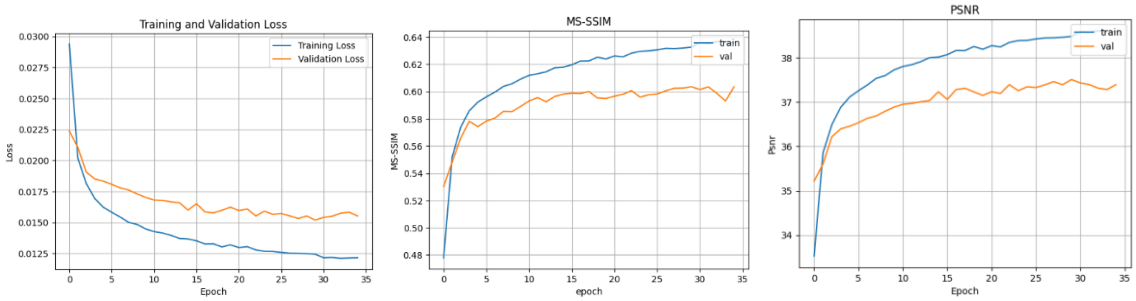


Figure 6-12 MXID AE's Results: MSE, MS-SSIM, PSNR *[29]*

- Convolutional Neural Network (CNN): The Convolutional Neural Network (CNN) model was initially trained on the proposed MXID dataset for 21 epochs, which yielded high performance values. The model achieved a mean squared error (MSE) of 0.006, a maximum signal-to-noise ratio (PSNR) of 41.43 dB, and a mean structural similarity index measure (MS-SSIM) of 0.77, as shown in **Figure 6-13**. These values reflect improved performance compared to both the Variational Autoencoder (VAE) and the proposed Autoencoder (AE) models.

Figure 6-20 show the reconstructed pictures employing the CNN, which have improved visual quality compared to previous methodologies [52], with the combined CNN and wavelet transformation achieving a PSNR of 39.59 dB. On the other hand, [51] offered a region-based segmentation and optimization approach that combined fuzzy C-means with hybrid Grey Wolf

Optimization. This model had a higher PSNR value, indicating that the segmentation method maintained diagnostically essential areas critical to healthcare experts.

Deep learning image compression studies have explored a number of architectures for enhanced image quality and efficient compression ratios. In [87], a dual CNN architecture was proposed, where ComCNN learns a compressed image representation for image compression, and RecCNN generates high-fidelity reconstruction in decompression. It demonstrated considerable improvement over traditional compression methods by successfully mitigating artifacts at the same bit rates, outperforming standard image codecs. Similarly, [34] proposed a convolutional RNN model which achieved a 4.85 dB improvement in Peak Signal-to-Noise Ratio (PSNR) and a 4.86% improvement in Structural Similarity Index (SSIM) compared to the LSTM Conv model of Toderici et al. [88] for an 8x compression ratio on 256 × 256 images. Additionally, their method achieved a 9.85 dB gain in PSNR over the JPEG-2000 standard, which has particular significance in medical imaging. By way of comparison, the proposed CNN model performed even better, achieving an MSE of 0.006, a PSNR of 41.43 dB, and an MS-SSIM of 0.77, outperforming the earlier reported results, particularly in PSNR.

Additionally, a comparative visual image analysis is provided by K-Means, PCA, and CNN, and some features are referred to, as illustrated in **Figure 6-20**. However, due to the high contrast, K-Means loses some of the color information and can no longer preserve small image details. Moreover, images generated based on CNN reconstruction were discovered to be blurry, presenting a potential drawback in preserving critical areas. Although the higher k value produces a more precise replication, fuzziness in CNN-painted images is a sign that there is a need for adjustment to increase the network's ability to preserve details in the course of compression.

Deeper Convolutional Neural Networks (CNNs) are effective in learning hierarchical features, enhancing the retention of fine details when compressing images. However, with the thoughtfully balanced architecture of the proposed moderate-depth CNN, experiments proved that there were constraints in the faithful preservation of complex structures critical in medical imaging. The outcome depicts the trade-off between network depth and detail retention. Apart from this, other primary architectural features such as Rectified Linear Unit (ReLU) activations, batch normalization, and convolutional kernels also contributed significantly to model performance regarding compression. Nevertheless, the reconstructed images could not achieve the level of desired quality. To counter these problems, future work will study

enhancements in the network architecture, loss function optimization, and the inclusion of more advanced post-processing techniques. These improvements aim at blurriness elimination and overall improvement of compressed medical X-ray image quality.

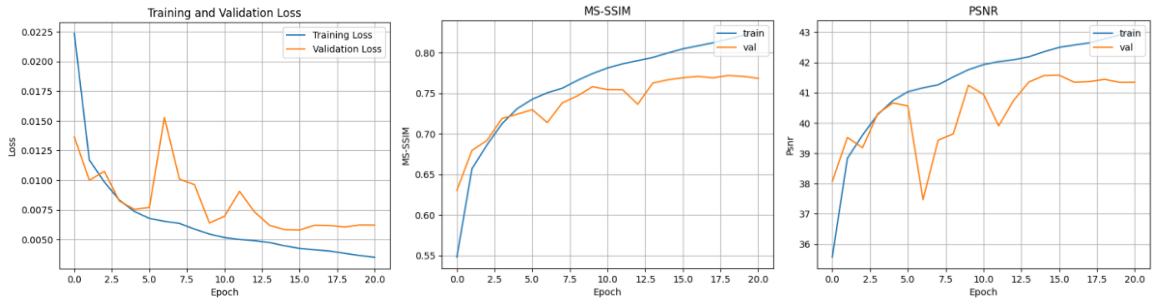


Figure 6-13 Evaluation of Mean Squared Error (MSE), Multi-Scale Structural Similarity Index (MS-SSIM), and Peak Signal-to-Noise Ratio (PSNR) depicting the performance dynamics of the CNN model over 21 training epochs *[29]*

Table 6-18 Comparative Compression Results for Various Deep Learning Models *[29]*

Model	Epochs	MSE	PSNR	MS-SSIM
AE	35	0.014	37.61	0.61
VAE	25	-	21.64	0.75
CNN	21	0.006	41.43	0.77
Proposed DCAE(RELU)	40	0.0004	46.37	0.98
Proposed DCAE(PRELU)	**51**	**0.0002**	**46.78**	**0.99**

Table 6-18 shows that the training process results in a decreased loss of 0.014, an increased MS-SSIM of 0.61, and a higher PSNR value of 37.61 dB when compared to the Mishra et al. approach [27][48] of 35.91 dB, the Juliet et al. approach [47] of 37.02 dB, and the Senapati et al. approach [46] of 23.91 dB. **Figure 6-14** provides more information on PSNR, MS-SSIM, and MSE details. This means that the autoencoder is effectively learning to compress and reconstruct pictures while retaining their structural features. However, as seen in **Figure 6-15** the compressed pictures were fuzzy due to a lack of finer details.

Figure 6-14 Evolution of Mean Squared Error (MSE), Multi-Scale Structural Similarity Index (MS-SSIM), and Peak Signal-to-Noise Ratio (PSNR) across 35 epochs for the Autoencoder Model Performance *[29]*

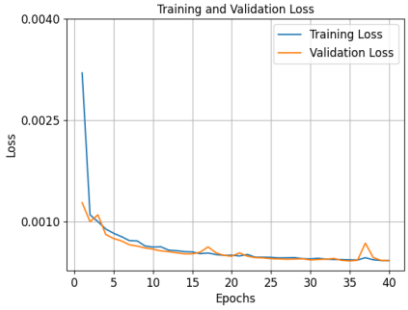
- Variational Autoencoder (VAE): Variational Autoencoder (VAE) architecture was also employed for the proposed Medical X-ray Imaging Dataset (MXID), which is made up of fully linked encoding and decoding layers that were trained for 25 epochs with an early stopping strategy to reduce overfitting. The data suggest picture preservation issues, which result in early termination due to minor reductions in validation loss, as seen in Figure **6-15**. Furthermore, we ran the trained VAE model on a variety of pictures from the dataset, producing and comparing PSNR and MS-SSIM metrics for each. The findings indicate that **Figure 6-20** produced the best possible outcomes. Furthermore, Ballé et al. [89] and Liu et al. [49] produced exceptional findings that outperformed our investigation, as shown in **Table 6-18.**

Figure 6-15Evaluation of Mean Squared Error (MSE) Loss for VAE Compression Technique *[29]*

- Deep Convolutional Autoencoder (DCAE): The deep convolutional autoencoder (DCAE) uses a symmetric encoding-decoding structure that is optimized to learn a compact representation of input images in a manner that ensures high-quality reconstruction. The

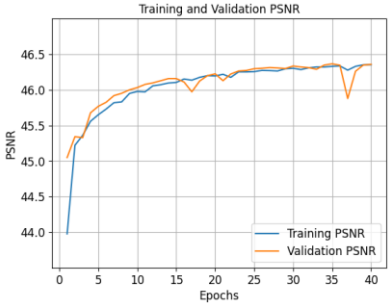
experimental results demonstrate that the method performs higher than existing deep learning-based methods for X-ray medical image reconstruction. The model achieved a loss of 0.002, structural similarity index (MS-SSIM) of 0.99, and peak signal-to-noise ratio (PSNR) of 46.78 dB after 51 training epochs with a batch size of four. The outcome indicates that there is negligible loss of diagnostic information, with the reconstructed images being highly similar to the originals. The use of the PReLU activation function played a major role in the performance of the model, which performed better than the ReLU activation function, as illustrated in **Figure 6-16**. However, it enables one to effectively build up variances of input data with better reconstruction quality across different classes of anatomy. For this reason, the DCAE model effectively preserved essential information that is vital to accurate medical diagnosis in different body areas, each presenting a unique combination of challenges concerning anatomical complexity, diagnostically relevant regions, and image characteristics, ranging from dental radiographs to the lumbar spine and cervical spinal column. An example representative is shown in **Figure 6-17**. The model was successful. These images show the well-preserved vertebral elements, intervertebral discs, and nerve elements. In addition, it is higher than CNN in preserving significant regions by maintaining structural information such as organs, tissues, and bony structures from lungs and abdomen.

Furthermore, the ability of the model to project finer details within bone architecture, joint compartments, and soft tissue was crucial for diagnostic accuracy in orthopedic examinations and disorders related to the joints. Where skeletal morphology was dense, say, in the skull, leg, and arm, the DCAE ensured accurate preservation of details necessary for effective clinical use. Preservation of bone density, facial structure, and musculoskeletal elements demonstrates the adaptability of the DCAE to disparate imaging circumstances.

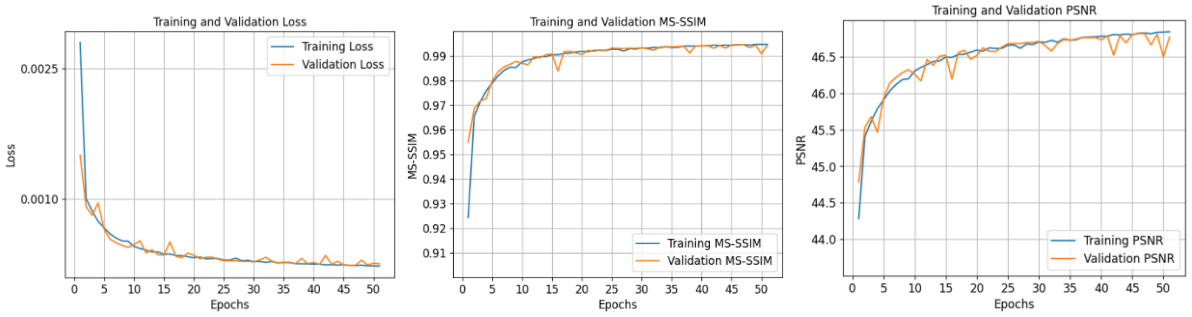


MSE (RELU)

MS-SSIM (RELU)

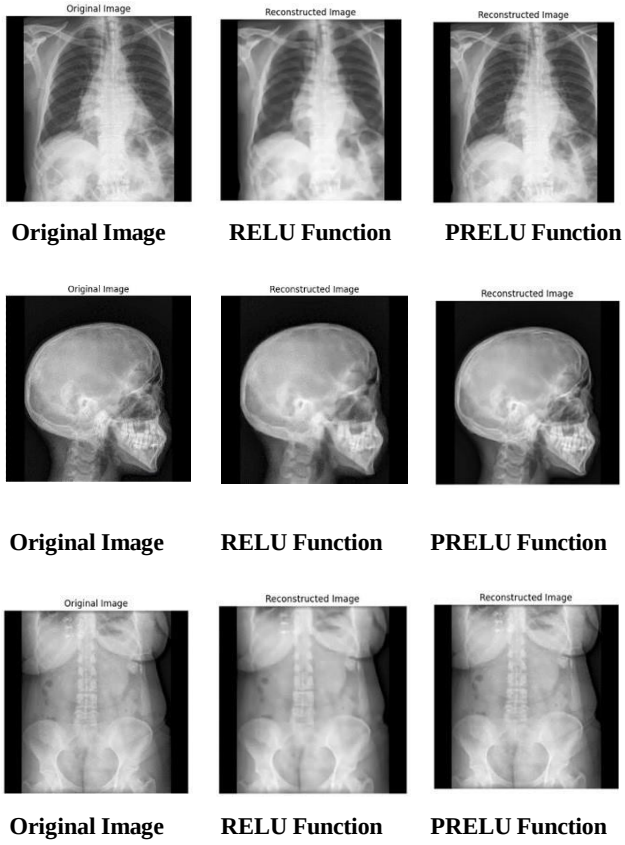


PSNR (RELU)



MSE (PRELU) **MS-SSIM (PRELU)** **PSNR (PRELU)**

Figure 6-16 Evaluation of DCAE Performance: Comparative Analysis of MSE, PSNR, and MS-SSIM with Both Activation Functions *[29]*



Original Image **RELU Function** **PRELU Function**

Original Image **RELU Function** **PRELU Function**

Original Image **RELU Function** **PRELU Function**

Figure 6-17 Deep Convolutional Autoencoder Results from the MXID Dataset comparing the impact of ReLU and PReLU activation functions on the compressed images: Representative Samples *[29]*

6.4.2.3.2 Fourth Experiment with Proposed Convolutional Autoencoder (CAE)

a) MXID DATASET

The proposed CAE structure is really designed for compression of medical images, maintaining necessary feature retention important in diagnosis in the clinical setting but

minimizing loss of information.

In MXID dataset, performance is exceptional where the model achieved a PSNR of 47.318 dB, MSE 0.00008, and MS-SSIM of 0.997 at 94 epochs. This is higher than the deep convolutional autoencoder of [29], which had a PSNR of 46.78 dB, an MSE of 0.0002, and an MS-SSIM of 0.99 at 51 epochs. The outcomes are shown in Error! Reference source not found. and **Figure 6-30** which closely approximate those of the OPEN-I dataset. This very high reconstruction accuracy ensures that the structural integrity of the original images is adequately maintained, which is essential for precise medical diagnoses. Additionally, the model's effectiveness and robustness are further cemented by its consistent performance across 18 distinct anatomical classes, each with different structural details and complexities.

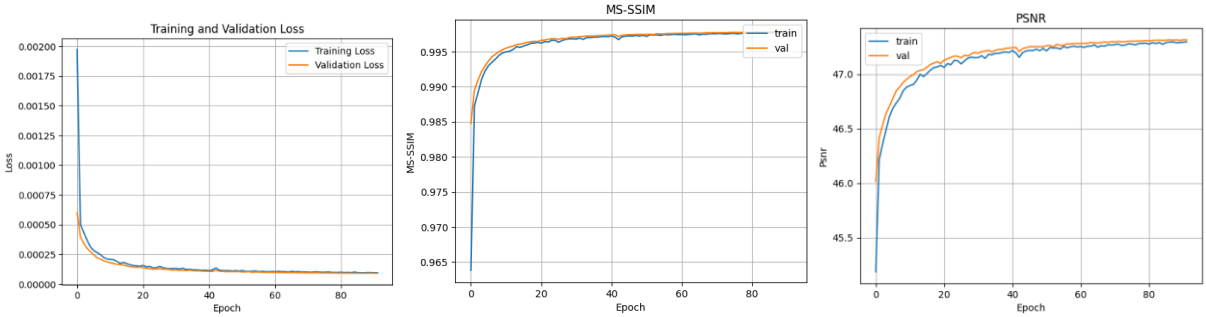


Figure 6-18The CAE model's performance trends on MXID dataset were evaluated using Mean Squared Error (MSE), Multi-Scale Structural Similarity Index (MS-SSIM), and Peak Signal-to-Noise Ratio (PSNR)

Figure 6-19 shows how the model perfectly reconstructs various parts of the body, including abdominal regions, small vessels, tissue differentiation, bone structures, and fine details in more complex regions such as organ boundaries, intervertebral discs, dental images with visible fillings, and other critical information. These results demonstrate that the model collected critical diagnostic information while maintaining a balance between picture fidelity and compression efficiency for reliable clinical interpretation over a wide variety of anatomical locations.

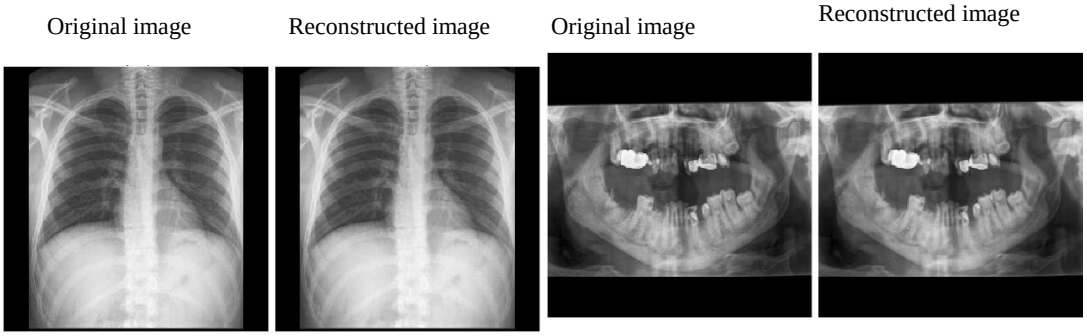


Figure 6-19 Representative Samples from The MXID Dataset Using CAE Compression Process

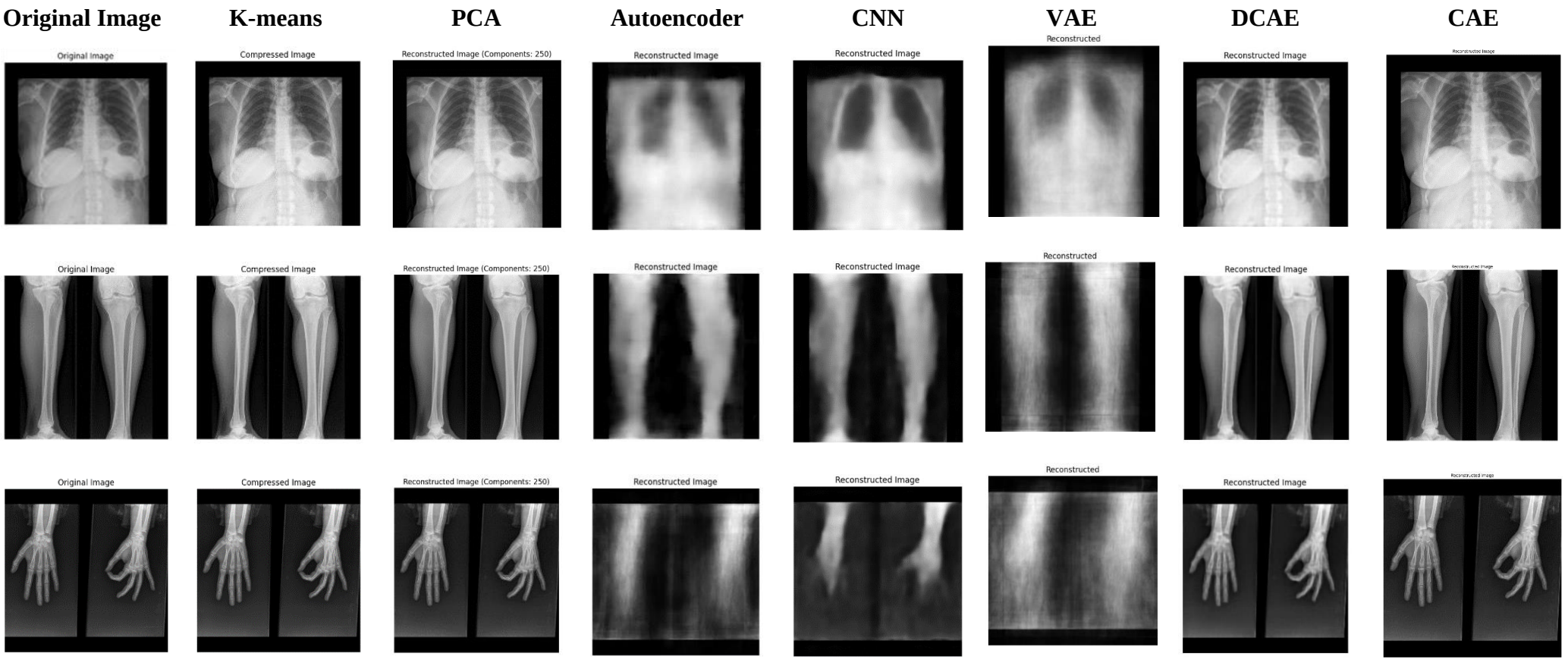


Figure 6-20 A comprehensive comparison of sample image reconstructions is presented, featuring results obtained through K-Means, PCA, Autoencoder, CNN, VAE, DCAE, and CAE Methods. The original image is juxtaposed with the corresponding reconstructed images, respectively *[29]*

6.4.2.4 Generalization Across the Additional Datasets

6.4.2.4.1 Fifth Experiment: Performance of DCAE, AE, and CNN on OPEN-I Dataset and JSRT Dataset

6.4.2.4.1.1 Performance on OPENI Dataset

- deep convolutional autoencoder (DCAE): The DCAE algorithm performs better than other algorithms on the OPEN-I dataset, with a loss of just 0.0001, and in the MXID and JSRT datasets, the loss is higher. Moreover, a PSNR of 47.14 dB indicates superior visual quality and superior anatomical area preservation, and an MS-SSIM of 0.99 indicates very high similarity between the input and reconstructed images, as shown in **Figure 6-27**.

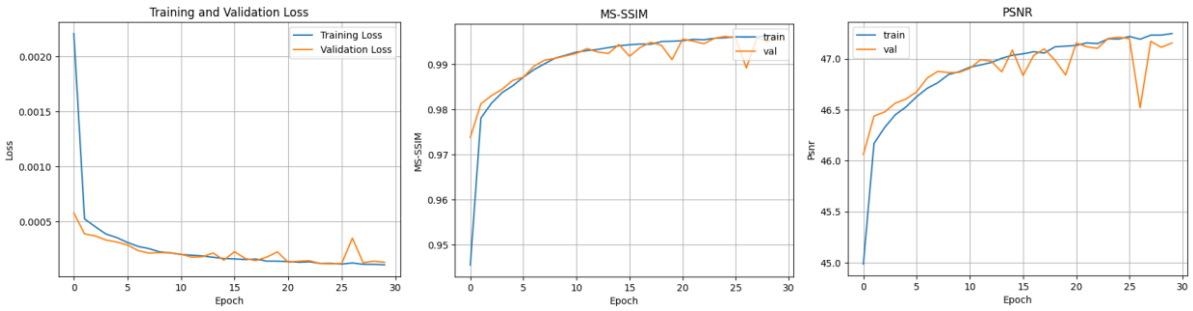


Figure 6-21 **OPEN-I DCAE's Results: MSE, MS-SSIM, PSNR**

- Autoencoder (AE): On 21 epochs, the autoencoder achieved 39.07 dB PSNR, 0.011 MSE, and 0.70 MS-SSIM on the OPEN-I database, higher than on MXID and JSRT databases. This is equivalent to higher compression quality and higher generalizability, particularly on larger databases. This improved performance illustrates the ability of the model to maintain important image details without sacrificing efficient compression, and this is a very promising approach to medical imaging where both diagnostic relevance and storage economy are required.

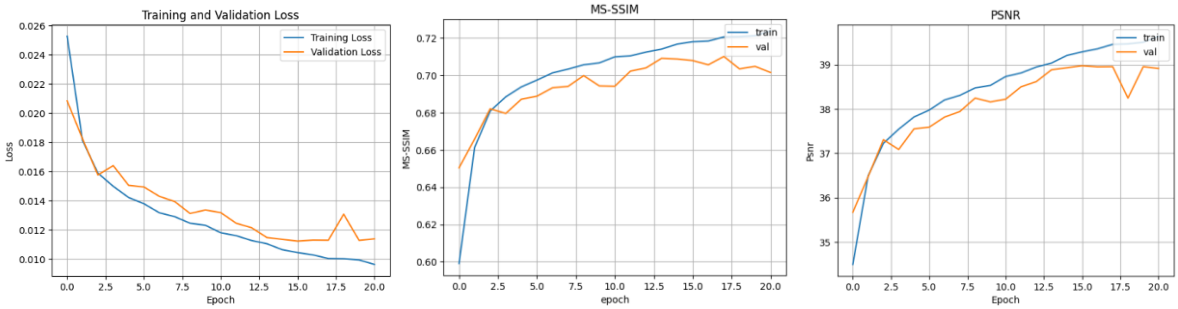


Figure 6-22 **OPEN-I AE's Results: MSE, MS-SSIM, PSNR**

- Convolutional Neural Network (CNN): Similarly, Similarly, applying the CNN to the OPENI dataset yielded an MSE of 0.004, a PSNR of 42.40 dB, and an MS-SSIM of

0.80, indicating superior performance compared to the MXID and JSRT findings shown in **Table 6-19**.

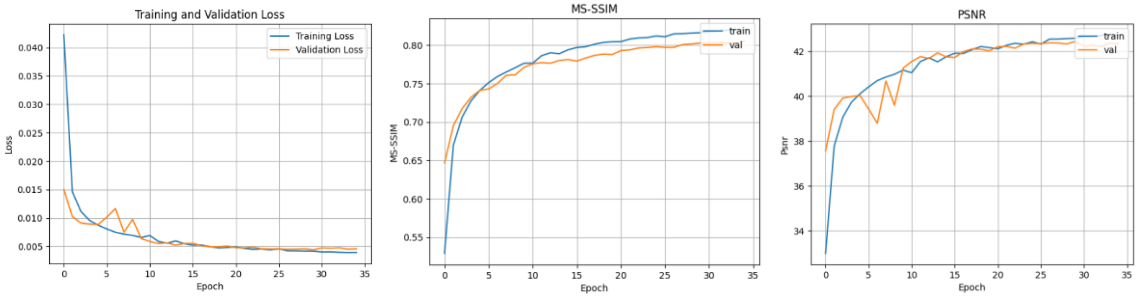


Figure 6-23 **OPEN-I CNN's Results: MSE, MS-SSIM, PSNR**

6.4.2.4.1.2 Performance on JSRT Dataset

- Deep Convolutional AutoEncoder (DCAE): The DCAE performed well on the JSRT dataset, with a PSNR of 45.37 dB, an SSIM of 0.97, and an MSE of 0.001 after 21 epochs as depicted in **Figure 6-25**. However, image quality was slightly lower compared to the MXID and Open-I datasets due to the JSRT imaging conditions and smaller dataset size, as shown in **Figure 6-28**. Nonetheless, the slight variances in SSIM and compression metric values suggest that dataset characteristics such as size, variability, and picture quality influence the model's performance.

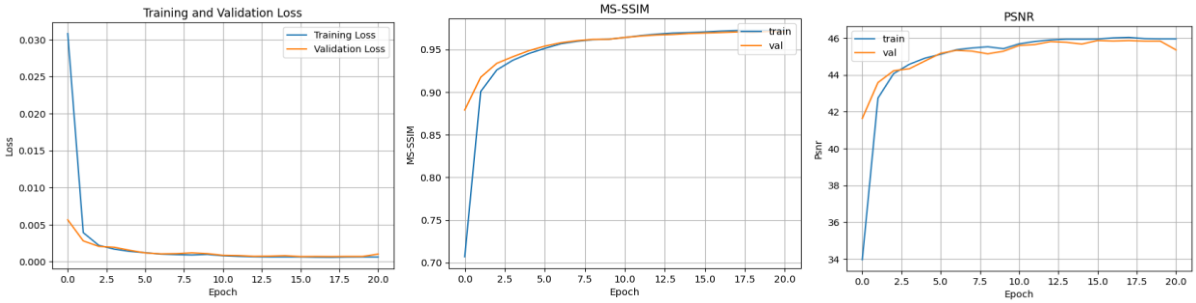


Figure 6-24 **JSRT DCAE's Results: MSE, MS-SSIM, PSNR**

- Autoencoder (AE): Moving on to the autoencoder's performance on the JSRT dataset, after 18 epochs, the model achieved a slightly higher PSNR of 37.84 dB than MXID, an MSE of 0.014, which is similar to the autoencoder's MXID results, and an MS-SSIM of 0.77, which is the highest value compared to the two other datasets, indicating that the original and reconstructed images are more similar.

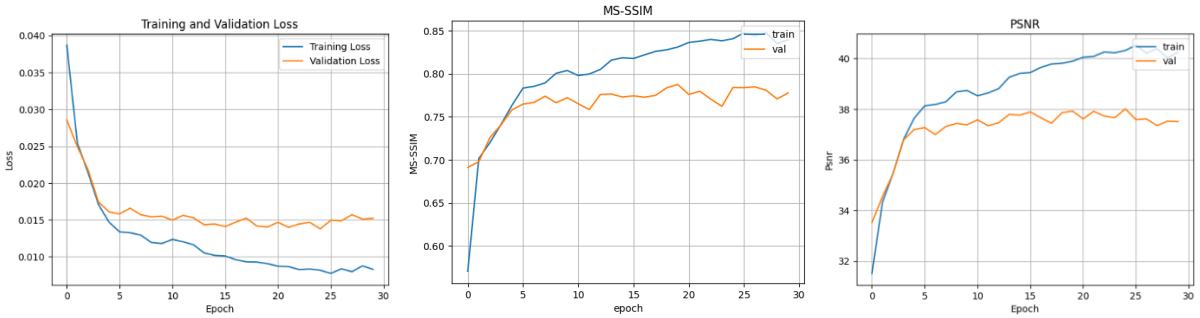


Figure 6-25 **JSRT AE's Results: MSE, MS-SSIM, PSNR**

- Convolutional Neural Network (CNN): The model attained an MSE of 0.016, PSNR of 36.96 dB, and MS-SSIM of 0.71 at 35 epochs, as indicated in **Table 6-19**. The performance is the poorest across the tested datasets and indicates poor performance of the CNN on the JSRT dataset compared to MXID and OPEN-I. The lower and more volatile the dataset, i.e., more noisy or lower image resolution, for example, it influenced the model in terms of how generalizable or how much detail it can retain critical to accurate diagnosis. Smaller datasets will easily overfit, particularly with deep convolutional models with high parameters. While autoencoders (AEs) learn in terms of less complex features using less diversified data, feature learning in CNNs is good but utilizes larger datasets for better generalization. DCAE model incorporates the strengths of both CNNs and AEs but on the JSRT database performed better feature extraction than AEs but was also prone to overfitting.

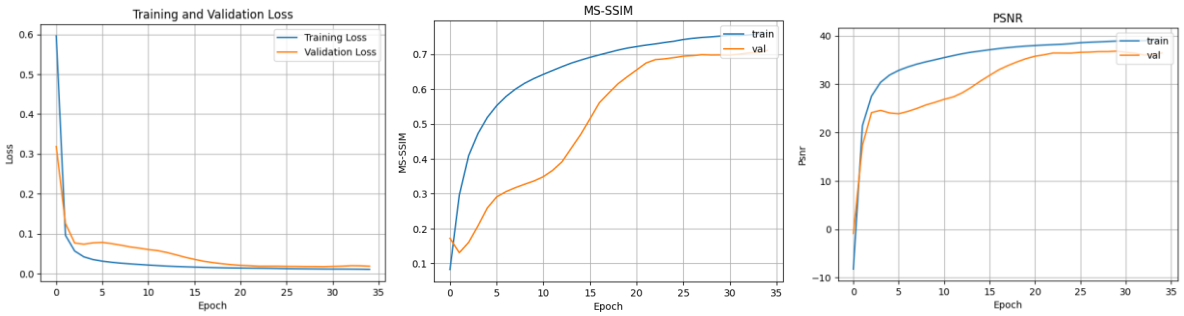


Figure 6-26 **JSRT CNN's Results: MSE, MS-SSIM, PSNR**

6.4.2.4.1.3 Discussion

The performance of the various approaches on the three X-ray datasets is much better on larger and diverse datasets such as MXID and OPEN-I, indicating their clinical use potential. However, differences in the quality of reconstructed images across datasets indicate that the models are sensitive to dataset-specific properties, i.e., size, diversity, and inherent variations. These allow the models to pick up more features when trained on larger datasets. In addition,

the PSNR and MS-SSIM values of the MXID dataset reveal high-quality reconstructed images in the three models (AE, CNN, and DCAE). The models are able to preserve fine structural details in 18 anatomical regions, such as key features including bone contours, nerve structures, joint spaces, skeletal and vertebral details, intervertebral discs, organs, tissues, and cardiovascular features. This suggests a fair trade-off between image quality maintenance and compression efficiency for multi-body part imaging datasets. The visual side-by-side comparison of such results is given in **Figure 6-20, Figure 6-27**, **Figure 6-28**.

The OPEN-I chest X-ray dataset indicates marginally greater PSNR values for all three deep learning models because it is larger in scope and has more extensive training data. This allows the models to learn more precise features compared to anatomy of the chest. Big datasets also lower the risk of overfitting and enhance the quality of reconstructed images.

In contrast, the JSRT chest X-ray small dataset produces lower PSNR values than the OPEN-I and MXID datasets for the three models (AE, CNN, and DCAE). The small dataset size restricts the model from extracting very little diagnostic information. The variation of image quality in the dataset causes low PSNR and MS-SSIM values. The blurring induced within reconstructed images by this effect can hide initial symptoms of disease like pneumonia or fibrosis. The simultaneous decrease in contrast between such tissues can decrease detection of minute differences in tissue density, which is critical to generating accurate medical diagnoses.

Table 6-19 DCAE, AE, CNN Performance Comparison on: MXID vs. OPEN-I vs. JSRT

Applied Technique	Dataset	Epochs	MSE	PSNR	MS-SSIM
DCAE	**MXID**	51	0.0002	46.78	0.99
	OPEN-I	**30**	**0.0001**	**47.14**	**0.99**
	JSRT	21	0.001	45.37	0.97
AE	**MXID**	35	0.014	37.61	0.61
	OPEN-I	**21**	**0.011**	**39.07**	**0.70**
	JSRT	18	0.014	37.84	0.77
CNN	**MXID**	21	0.006	41.43	0.77
	OPEN-I	**35**	**0.004**	**42.40**	**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; thus, the JPEG method may achieve higher compression ratio values despite the noticeable loss of crucial details, which is critical 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 [70], while balancing image fidelity and storage requirements; however, it requires more processing power and is still struggling with complex features and details captioning when compared to deep learning techniques that offer potential preservation while achieving high compression ratios.

Besides, the application of deep learning on MXID and other datasets is confronted with great challenges. MXID dataset consists of several parts of the human body with various characteristics, luminosity, and contrasts. Nevertheless, while the autoencoder reduces dimensions and noise, it cannot obtain intricate features for reconstructed medical images, which limits the process of capturing the details in the various anatomical areas.

Although CNNs are superior in extracting intricate features, they need more extensive datasets to avoid overfitting and produce better-quality reconstructed images. As indicated in **Figure 6-28**, the CNN model performs poorly with the JSRT dataset, generating extremely blurred images with no discernible anatomical details, reflecting its inability to capture fine structures.

Conversely, the DCAE model generates high-quality compression, particularly excelling on the OPEN-I dataset, with reconstruction error minimized and fine details well-preserved, enhancing diagnostic accuracy. While it is effective in encoding structural information with convolutional layers and is robust to noise—both desirable in medical imaging—DCAEs are extremely sensitive to the quality of training data. When the data set contains artifacts or noise, these defects are maintained in the reconstructed images, which can affect diagnostic reliability. Large high-resolution data sets' training for DCAEs is also computationally costly and time-consuming.

These findings demonstrate the promise of many methodologies across various complicated and bigger medical imaging datasets, while also highlighting limits with smaller dataset features such as noise levels, particularly in the JSRT dataset.

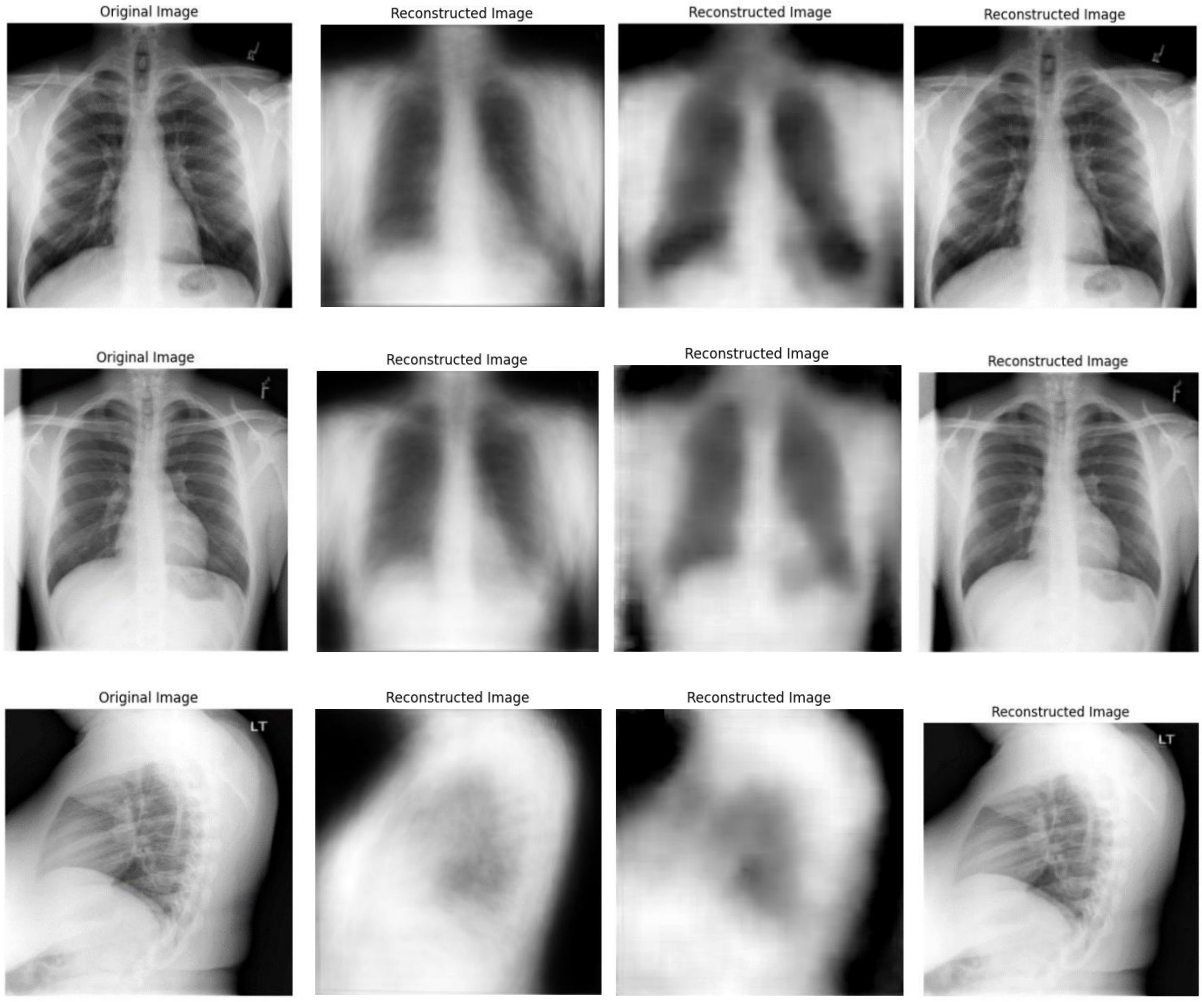
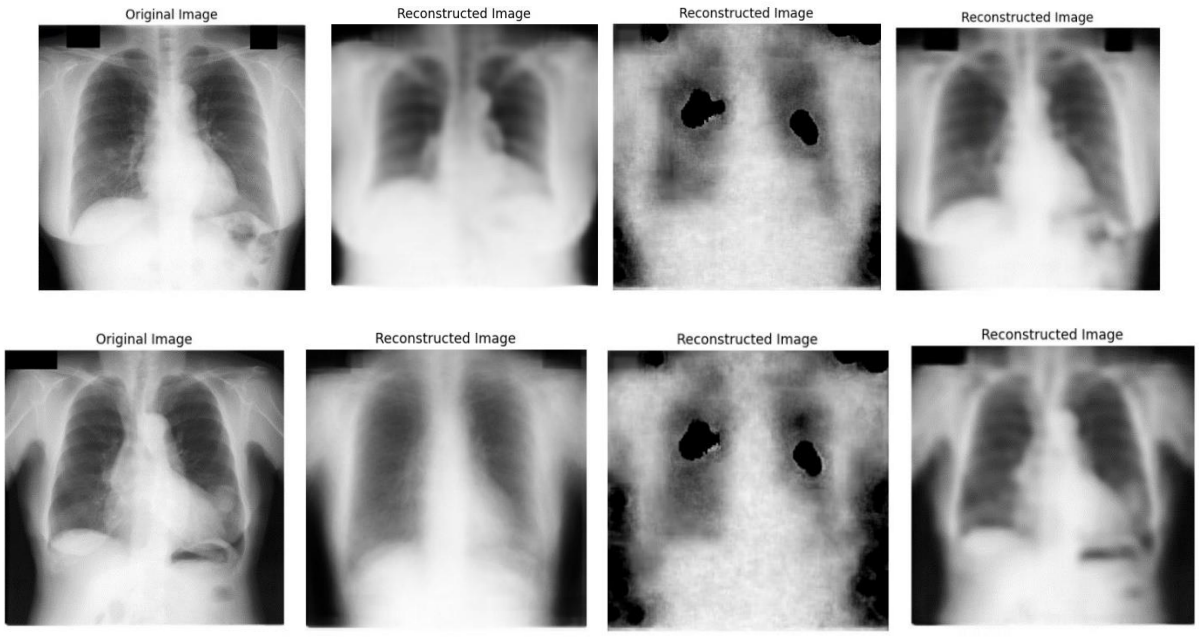


Figure 6-27 CNN vs. AE vs DCAE Results on OPEN-I Dataset



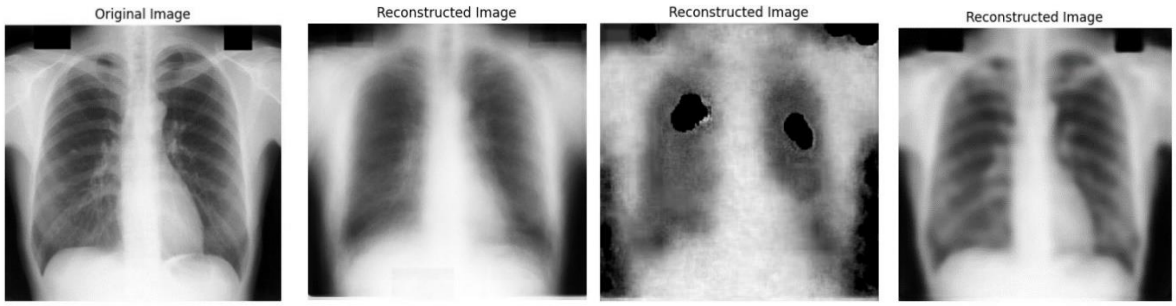


Figure 6-28 **CNN vs. AE vs DCAE Results on JSRT Dataset**

6.4.2.4.2 Sixth Experiment: Comparative study of CAE vs. RLE on Open-I and JSRT datasets

a) OPENI DATASET

Table 6-20 shows that the compression performance of the OPEN-I dataset was better than that of the MXID and JSRT datasets. With a significantly reduced MSE value of 0.00004, PSNR of 47.528 dB, and MS-SSIM of 0.998 at 58 epochs, as shown in **Figure 6-29**, indicating that the reconstructed image was of extremely high quality, had excellent detail preservation, and no distortion.

Figure 6-30 illustrates how the quality, diversity, and resolution of OPEN-I images result in improved reconstruction. Furthermore, the application of CAE architecture is most appropriate for this dataset, with minimal fine feature loss, allowing data reduction and image significant information for precise diagnosis, and an optimal image size for economical storage and transfer.

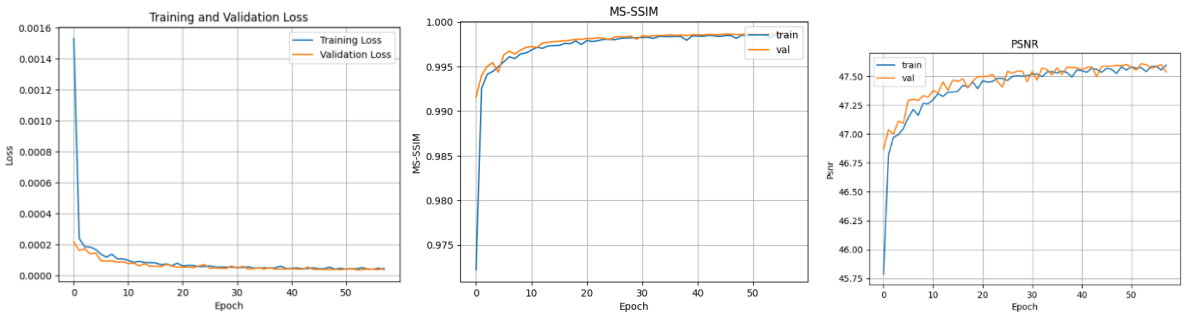


Figure 6-29 Assessment of Mean Squared Error (MSE), Multi-Scale Structural Similarity Index (MS-SSIM), and Peak Signal-to-Noise Ratio (PSNR) illustrating the performance trends of the CAE model on OPEN-I dataset

The OPEN-I dataset exhibits excellent preservation in the reconstructed pictures, with minimum loss of information in the ribs, lung arteries, and heart veins apparent, as well as in the broncho-vascular continuity, which is critical for diagnosis accuracy and precise

evaluation. Overall, this CAE approach shows a significant reduction in picture size while maintaining image quality and accurate, trustworthy diagnosis.

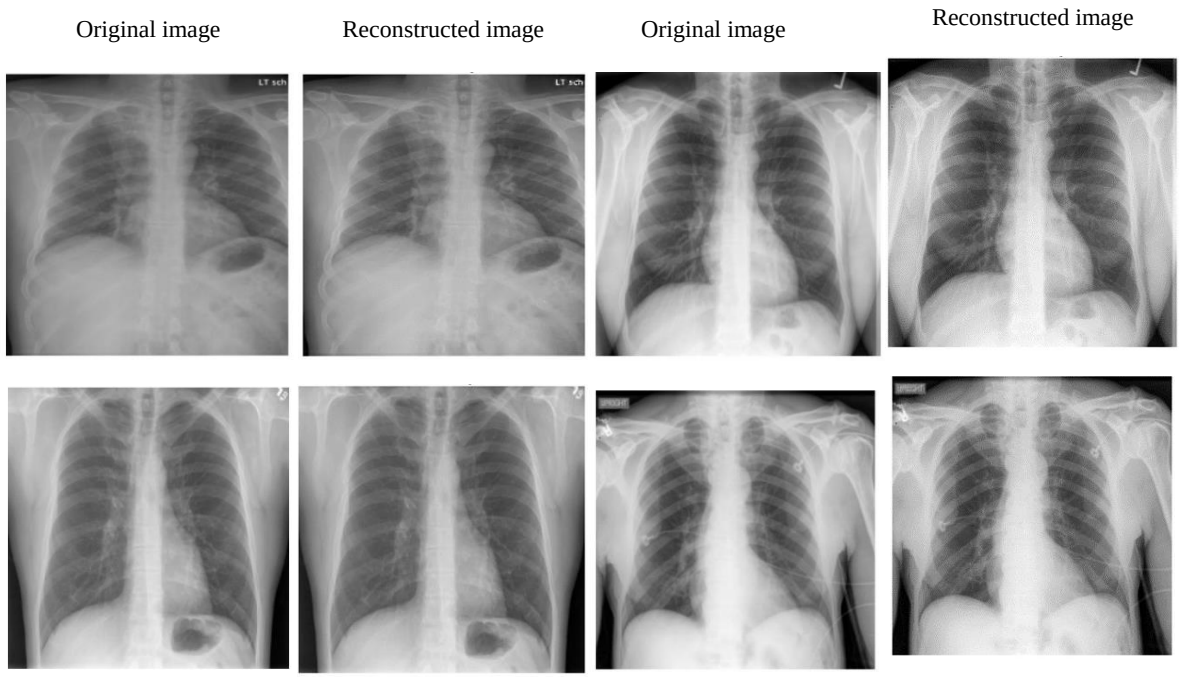


Figure 6-30 Outstanding demonstration of CAE's implementation on the OPEN-I dataset

b) JSRT DATASET

The model performance in the JSRT dataset is poorer than in the MXID and OPEN-I datasets, indicating loss of partial image quality, as shown in **Table 6-20**. However, the results indicate minor loss with an MSE of 0.00013, PSNR of 47.109 dB, and an MS-SSIM score of 0.997 after 45 iterations, as seen in **Figure 6-31**, such that the compression process achieved a decent visual quality that was quite close to the original. Furthermore, these results show equality of original and compressed image quality with diagnostic regions preserved for effective medical diagnosis.

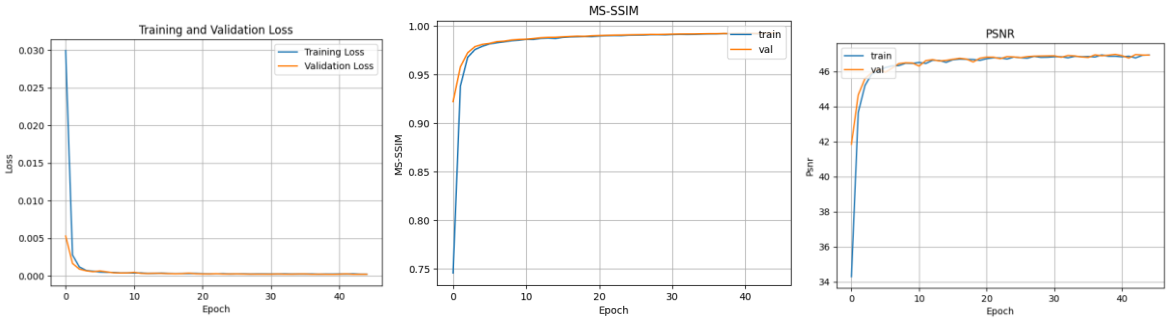


Figure 6-31 The performance trends of the CAE model were demonstrated through the evaluation of Mean Squared Error (MSE), Multi-Scale Structural Similarity Index (MS-SSIM), and Peak Signal-to-Noise Ratio (PSNR) on JSRT dataset

The visual outcome of the radiologists' and clinicians' diagnosis depends on the complexity of the case of the patient. Overall, anatomical structures are generally maintained in compressed images, as one can observe from **Figure 6-32**. Diagnosis would need more of the finer details to be maintained. Crucial structures such as sharp edges of rib cages, vascular veins in lungs and heart, pulmonary nodules, and tissue density are deformed in compressed images. This blurring of acuities can impair an expert's diagnostic accuracy by leading to overlooking findings, particularly in high-precision applications such as pathology at early stages or small lesions based on subtle grayscale differentiation.

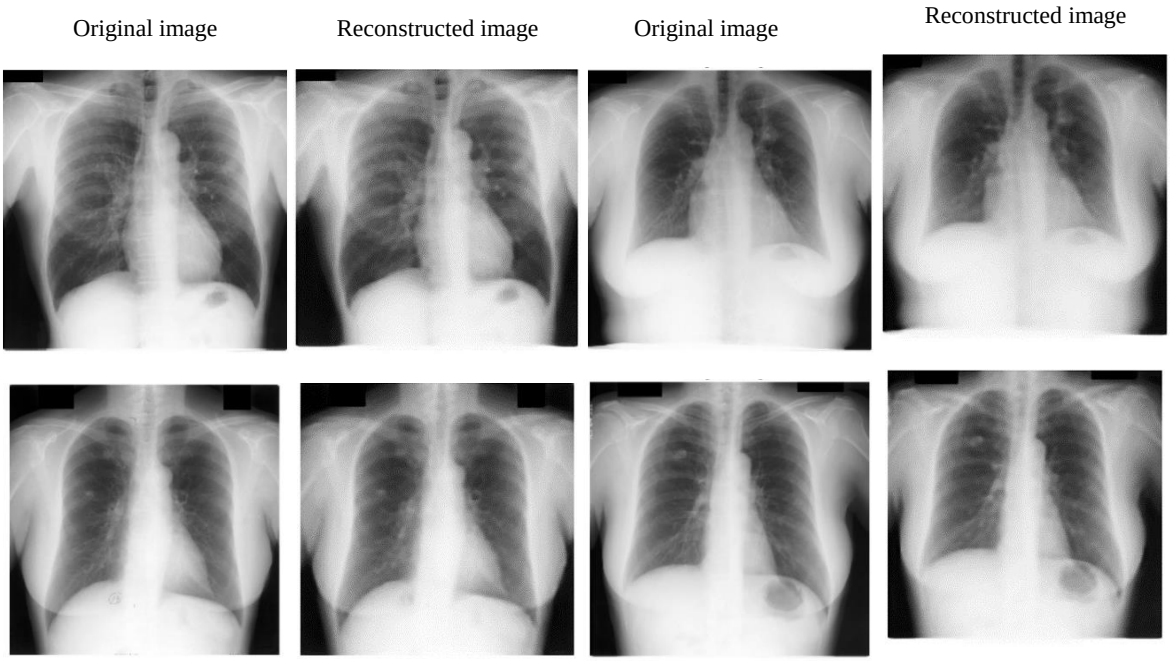


Figure 6-32 The Impact of the CAE on the JSRT Dataset: Representative Samples

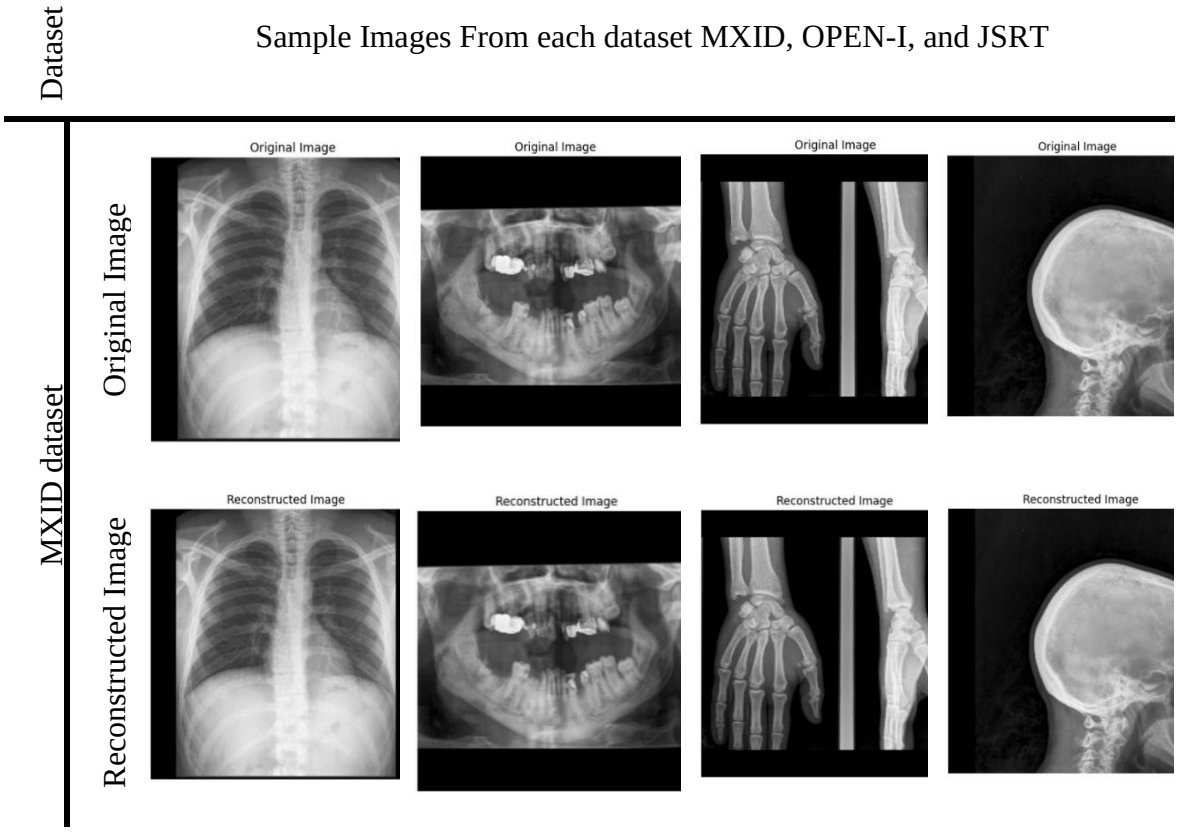
Finally, the solution proposed had decent performance on the three sets of data in balancing compressed image quality and data compression, preserving vital information, and at the same time, presenting an efficient avenue through which medical images could be transmitted and stored. The performance of compression for the OPEN-I dataset outperformed the performance of the MXID and JSRT datasets.

Table 6-20 Assessment of the CAE on the MXID, JSRT, and OPEN-I datasets

Datasets	Epochs	MSE	PSNR (dB)	MS-SSIM
MXID dataset	94	0.00008	47.318	0.997
JSRT dataset	45	0.00013	47.109	0.997
OPEN-I dataset	58	0.00004	47.528	0.998

6.4.2.4.3 Experiment with run-length encoding (RLE)

Run-Length Encoding (RLE) compression of Open-I, JSRT, and MXID datasets compressed the image size without sacrificing on the visual quality of the image as observed from **Figure 6-33.** This makes RLE an appropriate method for storage and transmission of medical images with good balance between the image compression ratio and reconstruction. For the MXID dataset, compression was specified as loss of quality that was imperceptible, with best trade-off between image fidelity and size reduction. For Open-I and JSRT datasets, RLE showed good performance in preserving salient features of the datasets, i.e., pixel intensity change. Reconstructed images were similar to the original images without loss of quality. These values achieved—PSNR: "inf" dB, MS-SSIM: 1.0, and MSE: 0.0—all verify that RLE is maintaining ideal structure features. These are all suggestive of how much it can be trusted as a lossless compression method, particularly for medical X-ray images with pixel repetition patterns.



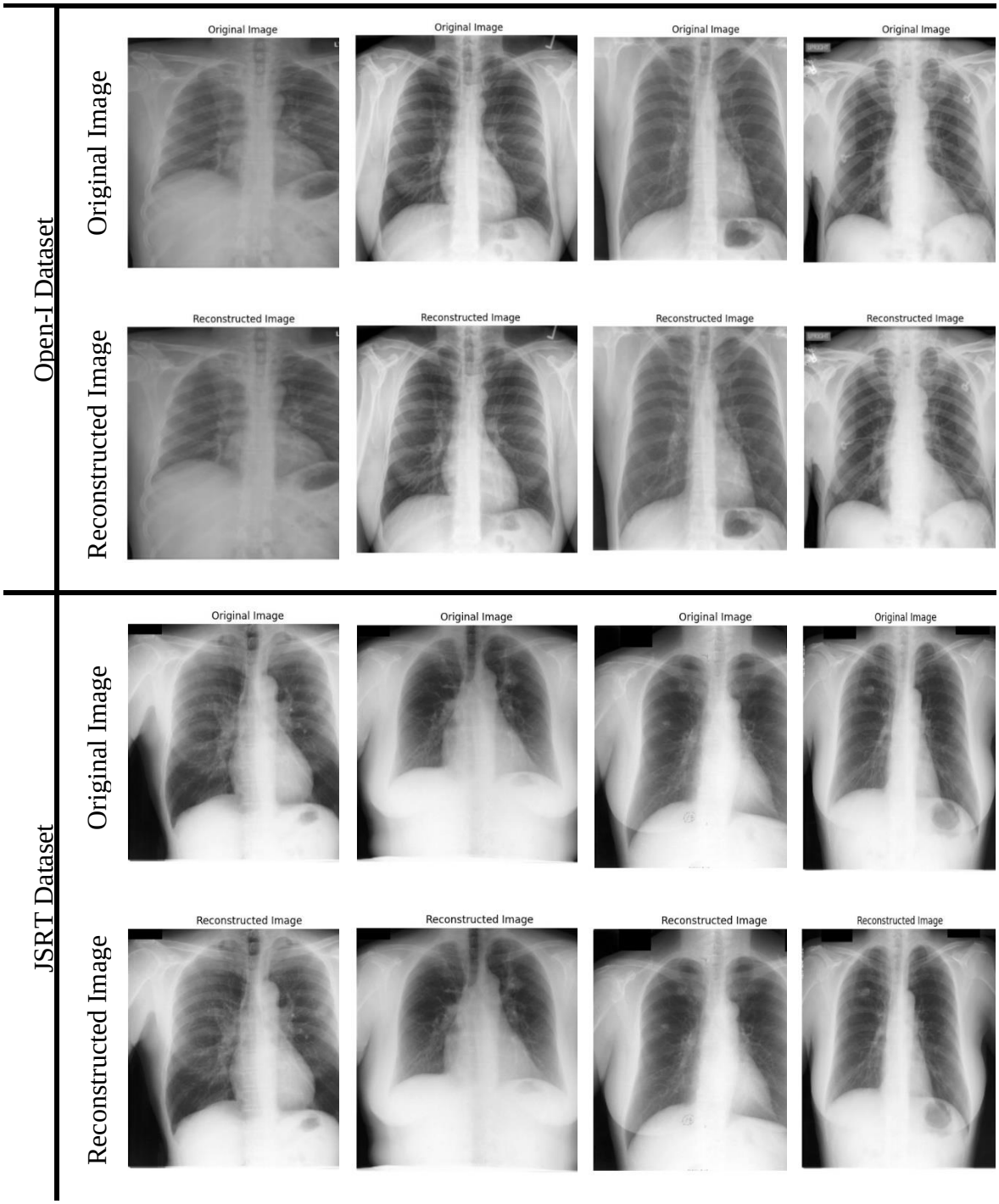


Figure 6-33 Illustrative example of the impact of RLE on MXID, OPEN-I, and JSRT datasets

6.4.2.4.4 Comparative Study Evaluation

The proposed CAE model exhibits improved performance on the OPEN-I dataset with maximum PSNR and MS-SSIM and minimum MSE than that of the JSRT dataset, which exhibited smaller values and weaker compression. The model also preserves notable fine details in 18 anatomical areas well, ensuring robust performance depending on dataset

characteristics. **Table 6-21** contrasts sample images from the MXID, OPEN-I, and JSRT datasets with four being selected from each, as shown in **Figure 6-19**, **Figure 6-30**, and **Figure 6-32**, respectively. The OPEN-I dataset was consistently better with higher PSNR values and lower MSE. Comparing RLE and CAE compression techniques from the three datasets showed that CAE compressed image size better. This is characteristic of the autoencoder's capacity to learn compact latent representations, effectively reducing image dimensionality without sacrificing the important structural information.

Table 6-21 Performance Assessment of the Proposed Autoencoder and RLE on grayscale samples with normalized size of 256 x 256 from the three datasets, using PSNR, MSE, and MS-SSIM metrics for a precise evaluation

Dataset	Images samples	CAE			RLE		
		PSNR (dB)	MSE	MS-SSIM	PSNR (dB)	MSE	MS-SSIM
MXID Dataset	**Image 1** ‘P00001_18-01_0_00001.png’	47.113	0.00013	0.995	Inf dB	0.0	1.0
	Image 2 ‘P00006_02-01_0_00007.png’	47.061	0.00015	0.997	Inf dB	0.0	1.0
	Image 3 ‘P00007_17-02_0_00009.png’	46.931	0.00019	0.997	Inf dB	0.0	1.0
	Image 4 ‘P00009_08-03_1_00012.png’	47.129	0.00013	0.994	Inf dB	0.0	1.0
OPENI Dataset	**Image 1** ‘CXR1000_IM-0003-1001.png’	47.50	0.000051	0.998	Inf dB	0.0	1.0
	Image 2 ‘CXR1005_IM-0006-1001.png’	47.30	0.000091	0.997	Inf dB	0.0	1.0
	Image 3 ‘CXR1007_IM-0008-3001.png’	47.52	0.000049	0.998	Inf dB	0.0	1.0
	Image 4 ‘CXR1008_IM-0009-4004.png’	47.15	0.00012	0.997	Inf dB	0.0	1.0
JSRT Dataset	**Image 1** ‘JPCLN001.png’	46.73	0.00025	0.990	Inf dB	0.0	1.0
	Image 2 ‘JPCLN003.png’	46.99	0.00017	0.992	Inf dB	0.0	1.0
	Image 3 ‘JPCLN004.png’	46.68	0.00027	0.990	Inf dB	0.0	1.0
	Image 4 ‘JPCLN006.png’	46.84	0.00021	0.992	Inf dB	0.0	1.0

On the other hand, RLE is not superior to CAE, with only moderate compression, as seen from **Table 6-22**. For example, image 'P00001_18-01_0_00001.png' in MXID dataset was originally 262,144 bytes, and encoded size is 131,072 bytes. After the application of RLE compression, file size was 209,540 bytes, showing the ability of RLE in encoding redundancy but without reducing the file size much. Nevertheless, RLE also maintains some anatomy information better, as indicated by **Figure 6-33**. It is because the complexity and diversity coping incapability of RLE in X-ray medical images, and CAE with enhanced feature extraction and dimension reduction.

CAE outperformed RLE greatly in terms of compression ratio since it did an optimal minimization of image size. **Table 6-22** is evidence that the 'CXR1000_IM-0003-1001.png' image of the OPEN-I dataset was subjected to an RLE compression ratio of 1.10, while its CAE-coded counterpart had a CR of 2.0. That indicates that CAE can do file size minimization well without sacrificing useful image features needed for proper medical diagnosis.

Table 6-22 Comparison of the original image size, CAE encoded image size, RLE Encoded image size, and compression ratio's values four samples from the three different datasets

Original Image Size (4 * 256 *256 * 1) Bytes = 262144 Bytes		
Encoded Images Size = 131072 Bytes **CR = 2.0**		
IMAGES' DATASET SAMPLES	**RLE Encoded Image Size (Bytes)**	**CR - RLE Encoded Image**
MXID Dataset		
Image 1 **'**P00001_18-01_0_00001.png**'**	209540	1.25
Image 2 'P00006_02-01_0_00007.png'	170388	1.53
Image 3 'P00007_17-02_0_00009.png'	151224	1.73
Image 4 'P00009_08-03_1_00012.png'	204188	1.28
OPENI Dataset		
Image 1	238220	1.10

'CXR1000_IM-0003-1001.png'		
Image 2 'CXR1005_IM-0006-1001.png'	254512	1.03
Image 3 'CXR1007_IM-0008-3001.png'	246752	1.06
Image 4 'CXR1008_IM-0009-4004.png'	253452	1.03
JSRT Dataset		
Image 1 'JPCLN001.png'	251056	1.04
Image 2 'JPCLN003.png'	249240	1.05
Image 3 'JPCLN004.png'	249948	1.05
Image 4 'JPCLN006.png'	245144	1.07

6.4.2.5 Experiment Based on LAVIA-MXID Dataset

9,989 images from the suggested LAVIA-MXID dataset in PNG (256x256 pixels) were employed to test the compression methods. Due to limited GPU memory, a batch size of four was used and images were normalized to a pixel value range of [0, 1]. Several deep learning techniques were attempted, including deep learning-based autoencoders, and the performance of these techniques was evaluated using the aforementioned metrics, and the findings are summarized in **Table 6-23**:

Table 6-23Result of handcrafted machine learning and deep learning method for image compession using our LAVIA MXID dataset compared to MXID dataset

Methods	**Parameters**	**MXID dataset**			**LAVIA MXID dataset**			
		MSE	**PSNR**	**MS-SSIM**	**Parameters**	**MSE**	**PSNR**	**MS-SSIM**
AE	**Epochs** =35	0.014	37.61	0.61	**Epochs** = 48	0.012	38.38	0.63
VAE	**Epochs** =25	-	21.64	0.75	**Epochs** = 29	-	21.00	0.73
CNN	**Epochs** =21	0.006	41.43	0.77	**Epochs** = 35	0.006	41.00	0.74
DCAE	**Epochs** =30	0.0002	46.78	0.99	**Epochs** = 74	0.0001	47.00	0.99

The comparison of outputs for various epochs for models AE, VAE, CNN, and DCAE on MXID version 1 and version 2 identifies improvement with the new LAVIA-MXID dataset to some degree by providing better-quality images and minimal reconstruction artifacts. The comparisons bellow draws the discussion of the models on the MSE, PSNR, and MS-SSIM parameters as well as between the versions of datasets and the performance of models.

6.4.2.5.1 Comparisons by Models

For Autoencoder (AE), the accuracy increased from the MXID to LAVIA-MXID dataset, but AE was not good with complex feature extraction against the DCAE model. The DCAE model, due to its deep architecture benefit, achieved better PSNR and MS-SSIM performance with the LAVIA-MXID dataset. For CNN model, it was comparatively good with MSE and PSNR but needed to be better with feature extraction and classification tasks. The VAE, with good performance in probabilistic modeling, demonstrated good performance in PSNR and MS-SSIM with some artifacts due to the stochastic behavior of the model.

6.4.2.5.2 Comparisons by Dataset's Version

The larger LAVIA-MXID dataset provided more diverse samples, resulting in better generalization on all the metrics. Still, the small improvements in CNN and VAE indicate that dataset version 1 might have been representative enough for these models. On the other hand, AE and DCAE had marginal but consistent improvement on all metrics with the LAVIA-MXID dataset. The slow progress in CNN and VAE implies possible structural limitations or trivial performance gains notwithstanding the bigger dataset.

6.5 Comparison with Related Works in The Literature

Table 6-24 presents a comparative study of some of the most advanced image compression techniques in terms of Peak Signal-to-Noise Ratio (PSNR) and Multi-Scale Structural Similarity Index Measure (MS-SSIM). These indices are important in quantifying the quality and fidelity of visually reconstructed images in lossy compression systems.

As can be seen from the results, the Proposed Convolutional Autoencoder (CAE) and its variants surpass most of the current methods in PSNR and MS-SSIM. Particularly, the Proposed Convolutional Autoencoder had a PSNR of 47.31 dB and MS-SSIM of 0.99, which is the highest performance among all compared methods. Among the earlier approaches, the Op-CNN [51] was the earlier best PSNR of 45.50 dB and other techniques like Convolutional Neural Networks with Wavelet Transformation [52] also exhibited strong performance (PSNR = 39.59 dB), but without MS-SSIM reporting. Lossy

Compression based Autoencoder [48] and End-to-End VAE [50] demonstrated relatively high MS-SSIM values (0.97) but also their PSNR values were significantly lower than those of the proposed methods.
Surprisingly, while some models such as the Compression and Denoising Autoencoder [46] or Lossy Compression Based-CNN [54] focused on visual resemblance (MS-SSIM of 0.92 and 0.87, respectively), they sacrificed reconstruction precision, as revealed by their low PSNR values (23.91 dB and 25.65 dB). In contrast, AE-DBM [47] had relatively equal performance with 37.02 dB PSNR and 0.99 MS-SSIM, but nonetheless lagged behind the introduced DCAE models. The three models proposed — Deep Convolutional Autoencoder Based ReLU, PReLU, and the final Proposed Convolutional Autoencoder — all performed better than previous models on both objective scores. The ReLU-based model achieved 46.37 dB / 0.98, the PReLU-based model slightly better at 46.78 dB / 0.99, and the final model peaked at 47.31 dB / 0.99. This indicates the effectiveness of the proposed architectural improvements and activation function modifications.

In [57], authors applied our proposed DCAE model and other models on the different medical imaging modalities, then comparing the results of each with their results of the proposed Quantum-enhanced Artificial Neural Network (QAAN) model achieving a high average PSNR values (ranging from 43.51 dB to 46.29 dB). Results were very close for each model, the difference in Δ value is minor when DCAE is compared with the other current methods and QANN method. While our proposed CAE outperformed all the models in the literature with a PSNR value of **47.31 dB**.

Table 6-24 Comparative Analysis of PSNR and MS-SSIM Values across Various State-of-the-Art Image Compression Techniques *[29]*

Technique	**PSNR Values (dB)**	**MS-SSIM**
Lossy Compression based Autoencoder [48]	35.91	0.97
Deep Autoencoder with Deep Bolzmann Machines (AE-DBM) [47]	37.02	0.99
Compression and Denoising Autoencoder [46]	23.91	0.92
Proposed Autoencoder [29]	**37.61**	**0.61**
Convolutional Neural Networks and Wavelet Transformation [52]	39.59	-

Optimized Convolutional Neural Network (Op-CNN) [51]	45.50	-
Proposed CNN	**41.43**	**0.77**
End-to-End compression-based VAE [50]	32.43	0.97
Integration of Residual Networks and Variational Autoencoder [49]	Higher PSNRand MS-SSIM than Ballé	
Variational Autoencoder [29]	**21.64**	**0.75**
Lossy Compression Based-Convolutional Neural Network [54]	25.65	0.87
Convolutional Autoencoder Neural networks (CAE-NN) [53]	-	-
Convolutional Autoencoder (CAE) [55]	26.48	0.82
a Quantum-enhanced Artificial Neural Network model for Efficient Medical Image Compression [57]	43.51 to 46.29	0.99
Proposed Deep Convolutional Autoencoder Based ReLU [29]	46.37	0.98
Proposed Deep Convolutional Autoencoder Based PReLU [29]	**46.78**	**0.99**
Proposed Convolutional Autoencoder	**47.31**	**0.99**

6.6 Conclusion

In this chapter, a series of experiments and findings of machine learning and deep learning-based compression techniques on medical X-ray images, the findings demonstrated the effectiveness of the models in striking a balance between storage and diagnostic quality preservation. Also, the LAVIA-MXID dataset, intended for different applications, such as image compression, and proves the success of machine learning and deep learning models such as the Deep Convolutional Autoencoder (DCAE) in achieving higher compression quality. The findings were analyzed to compare different architectures based on their strength and weaknesses, and identifying optimum configurations for medical image compression. Besides, the comparison with existing approaches provided a sense of the advantages and drawbacks of deep learning-based solutions. Some trade-offs between compression ratio, image quality, and computational complexity are nevertheless still open to be explored. The potential future directions for the development of medical imaging compression techniques, we are attempting to apply a Generative Adversarial Network (GAN) which would require some more time, we will complete it shortly.

GENERAL CONCLUSION AND PERSPECTIVES

The growing volume of high-quality medical images is presenting enormous challenges in terms of processing, transmission, and storage. Traditional techniques of compression, although adequate for the majority of instances of general image compression, are not up to the mark for medical images, as they are not designed to consider the unique structural features on which diagnostic dependability rests. It therefore becomes increasingly evident that new, advanced methods of compression must be explored in new health systems.

With the use of deep learning methods such as Autoencoders (AE), Deep Convolutional Autoencoders (DCAE), Convolutional Neural Networks (CNNs), and Variational Autoencoders (VAEs), this study has been conducted on machine learning and deep learning-based methodology in addressing problems with compressing medical X-ray images. It has also demonstrated the ability of such new methods to not just boost compression ratios but keep intact the fundamental structural information essential for accurate diagnosis. Research into traditional machine learning algorithms like Principal Component Analysis (PCA) and K-means clustering.

The research introduced the Medical X-ray Imaging Dataset (MXID), a large and high-quality dataset of 6,869 X-ray images collected at AOUINET Hospital, Tebessa, Algeria. This 18-region anatomical data set with varying body part and gender splits is a valuable resource for a wide range of image processing tasks such as image classification, image segmentation, image detection, image compression, and so on. MXID dataset is the yardstick by which future comparative medical image compression, classification, and machine learning research will be compared, illustrating the potential offered by high quality representative data sets in facilitating medical imaging innovation.

The findings of the study confirm that deep learning-based models, particularly DCAEs and CAEs, are better than other compression algorithms in terms of image compression efficiency and image quality retention. Deep learning models performed better on PSNR and SSIM, and findings confirm that high-quality feature preservation of medical images is preserved. However, even as positive progress is observed, there remain issues in preserving high fidelity of intricate anatomical structures of utmost relevance to clinical diagnosis, especially in highly detailed feature images such as bones, tissues, and organs.

Second, this study also demonstrates the importance of high quality and diverse dataset during training and testing of deep learning models. The MXID dataset provides not only good-quality images but also correct annotations, which enhances the performance and accuracy of compression models.

Despite the advancements triggered, there are some points left to explore in the future. Progressing the structure of models, optimizing loss functions, and including post-processing and pre-processing techniques, such as in the hybrid AE-RLE model incorporating the RLE, can further enhance compressed medical image quality.

In general, this work has a significant implication for image compression in medicine using deep learning with better compression ratios and preservation of image quality compared to traditional schemes. The study shows the potential of deep learning to transform the management of medical data such that storage, transmission, and access to critical healthcare information is more efficient. With increased popularity of medical imaging and healthcare system development, future development can be based on generative methods (GANs) and the MXID dataset, which can lead to improved healthcare in the context of quicker and more accurate diagnosis, reduced storage cost, and improved overall management of medical images within clinical settings.

SCIENTIFIC CONTRIBUTIONS

- Fettah, R. Menassel, and A. Gattal, “Machine Learning for Medical Image Analysis: A Survey”, *Lecture Notes in Networks and Systems*, vol. 713 LNNS, no. July, pp. 148–164, 2023, doi: 10.1007/978-3-031-35248-5_14.
- Fettah, R. Menassel, A. Gattal, and A. Gattal, “Convolutional Autoencoder-Based medical image compression using a novel annotated medical X-ray imaging dataset,” *Biomed Signal Process Control*, vol. 94, p. 106238, Aug. 2024, doi: 10.1016/J.BSPC.2024.106238.
- Fettah, R. Menassel, and A. Gattal, “Assessing the Generalizability of Deep Learning-Based Compression Techniques for X-ray Medical Images: A Comparative Study” *in international conference:* Trends and Advances in Collaborative applied Computing (TACC 2024), November 25–26, 2024, Constantine, Algeria.

BIBLIOGRAPHY

[1] J. E. Hutchinson, “FRACTALS AND SELF SIMILARITY.”

[2] M. F. Barnsley and S. Demko, “ITERATED FUNCTION SYSTEMS AND THE GLOBAL CONSTRUCTION OF FRACTALS.,” *Proc R Soc Lond A Math Phys Sci*, vol. 399, no. 1817, pp. 243–275, Jun. 1985, doi: 10.1098/rspa.1985.0057.

[3] R. Menassel, I. Gaba, and K. Titi, “Introducing BAT inspired algorithm to improve fractal image compression,” *International Journal of Computers and Applications*, vol. 42, no. 7, pp. 697–704, Oct. 2020, doi: 10.1080/1206212X.2019.1638631.

[4] N. A. Backiam, “Fast Fractal Image Compression based on Fisher’s classification scheme.”

[5] A. Deshlahra, G. S. Shirnewar, and A. K. Sahoo, “A Comparative Study of DCT, DWT & Hybrid (DCT-DWT) Transform.”

[6] F. Alfiah *et al.*, “Discrete Cosine Transform DCT Methods on Compression RGB and Grayscale image**********-********** Discrete Cosine Transform DCT Methods on Compression RGB and Grayscale image,” *International Journal of Computer Techniques*, vol. 4, [Online]. Available: http://www.ijctjournal.org

[7] A. Kurniawan, T. W. Purboyo, and A. L. Prasasti, “Implementation of Image Compression Using Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT),” 2017. [Online]. Available: http://www.ripublication.com

[8] Tp. Joshua, “International Journal of Computer Science and Mobile Computing Comparison of DCT and DWT Image Compression,” 2016. [Online]. Available: www.ijcsmc.com

[9] “JPEG-00125072”.

[10] A. Saffor, A. R. Ramli, and K.-H. Ng, “A COMPARATIVE STUDY OF IMAGE COMPRESSION BETWEEN JPEG AND WAVELET,” 2001.

[11] R. Mohammed, “A Multi-agent System for Image Compression,” *Int J Comput Appl*, vol. 53, no. 9, pp. 38–45, Sep. 2012, doi: 10.5120/8451-2248.

[12] S. M. Hosseini and A. R. Naghsh-Nilchi, “Medical ultrasound image compression using contextual vector quantization,” *Comput Biol Med*, vol. 42, no. 7, pp. 743–750, Jul. 2012, doi: 10.1016/j.compbiomed.2012.04.006.

[13] T.-C. Lu and C.-Y. Chang, “Journal of Information Hiding and Multimedia Signal Processing c ◯2010 ISSN,” 2010.

[14] A. Said, W. A. Pearlman, and S. Member, “A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees,” 1996.

[15] “Fundamental Data Compression.”

[16] M. J. Haque, M. N. Huda, and M. Coordinator, “Study on Data Compression Technique,” 2017.

[17] D. A. Huffmant, “A Method for the Construction of Minimum-Redundancy Codes*.”

[18] M. J. Knieser, F. G. Wolff, C. A. Papachristou, D. J. Weyer, and D. R. Mcintyre, “A Technique for High Ratio LZW Compression,” 2003.

[19] S. S and R. L, “TEXT COMPRESSION ALGORITHMS - A COMPARATIVE STUDY,” *ICTACT Journal on Communication Technology*, vol. 02, no. 04, pp. 444–451, Dec. 2011, doi: 10.21917/ijct.2011.0062.

[20] S. G. Fathahillah and R. Zain, “Homogeneous Image Compression Techniques with the Shannon-Fano Algorithm,” *Int. J. Environ. Eng. Educ*, vol. 1, no. 2, pp. 59–66, 2019, doi: 10.5281/zenodo.3490205.

[21] A. Fettah, R. Menassel, and A. Gattal, “Machine Learning for Medical Image Analysis: A Survey,” *Lecture Notes in Networks and Systems*, vol. 713 LNNS, no. July, pp. 148–164, 2023, doi: 10.1007/978-3-031-35248-5_14.

[22] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, “KNN Model-Based Approach in Classification.”

[23] K. Vembandasamyp, R. R. Sasipriyap, and E. Deepap, “Heart Diseases Detection Using Naive Bayes Algorithm,” 2015. [Online]. Available: www.ijiset.com

[24] L. Breiman, “Random Forests,” 2001.

[25] T. G. Dietterich, “An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization,” 2000.

[26] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, “An assessment of the effectiveness of a random forest classifier for land-cover classification,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 67, no. 1, pp. 93–104, Jan. 2012, doi: 10.1016/j.isprsjprs.2011.11.002.

[27] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” May 27, 2015, *Nature Publishing Group*. doi: 10.1038/nature14539.

[28] D. Mishra, S. K. Singh, and R. K. Singh, “Deep Architectures for Image Compression: A Critical Review,” *Signal Processing*, vol. 191, p. 108346, 2022, doi: 10.1016/j.sigpro.2021.108346.

[29] A. Fettah, R. Menassel, A. Gattal, and A. Gattal, “Convolutional Autoencoder-Based medical image compression using a novel annotated medical X-ray imaging dataset,” *Biomed Signal Process Control*, vol. 94, p. 106238, Aug. 2024, doi: 10.1016/J.BSPC.2024.106238.

[30] R. M. Schmidt, “Recurrent Neural Networks (RNNs): A gentle Introduction and Overview,” Nov. 2019, [Online]. Available: http://arxiv.org/abs/1912.05911

[31] *2018 International Conference on Intelligent Systems and Computer Vision : 2-4 April 2018, Fez, Morocco*. Institute of Electrical and Electronics Engineers, 2018.

[32] S. Rani, S. Sasibhushans, R. Gottapu, and P. Rao, “A discrete wavelet transform and recurrent neural network based medical image compression for MRI and CT images,” *J Ambient Intell Humaniz Comput*, no. 0123456789, 2020, doi: 10.1007/s12652-020-02212-7.

[33] A. I. Khaleel, N. Adilah, H. Zahri, and M. I. Ahmad, “A Hybrid Compression Method for Medical Images Based on Region of Interest Using Artificial Neural Networks,” vol. 2021, 2021.

[34] A. S. Sushmit, S. U. Zaman, A. I. Humayun, T. Hasan, and M. I. H. Bhuiyan, “X-Ray image compression using convolutional recurrent neural networks,” *2019 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2019 - Proceedings*, pp. 19–22, 2019, doi: 10.1109/BHI.2019.8834656.

[35] P. Guo, D. Li, and X. Li, “Deep OCT image compression with convolutional neural networks,” *Biomed Opt Express*, vol. 11, no. 7, p. 3543, Jul. 2020, doi: 10.1364/boe.392882.

[36] M. U. A. Ayoobkhan, E. Chikkannan, and K. Ramakrishnan, “Feed-Forward Neural Network-Based Predictive Image Coding for Medical Image Compression,” *Arab J Sci Eng*, vol. 43, no. 8, pp. 4239–4247, Aug. 2018, doi: 10.1007/s13369-017-2837-z.

[37] A. S. Methil, “Brain Tumor Detection using Deep Learning and Image Processing,” *Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021*, no. March, pp. 100–108, 2021, doi: 10.1109/ICAIS50930.2021.9395823.

[38] N. Tajbakhsh *et al.*, “Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?,” *IEEE Trans Med Imaging*, vol. 35, no. 5, pp. 1299–1312, 2016, doi: 10.1109/TMI.2016.2535302.

[39] B. Ahuja and R. Doriya, “An Unsupervised Learning Approach for Visual Data Compression with Chaotic Encryption,” *2021 4th International Conference on Electrical, Computer and Communication Technologies, ICECCT 2021*, 2021, doi: 10.1109/ICECCT52121.2021.9616827.

[40] A. Thakker, N. Namboodiri, R. Mody, R. Tasgaonkar, and M. Kambli, “Lossy Image Compression-A Comparison Between Wavelet Transform, Principal Component Analysis, K-Means and Autoencoders,” *5th IEEE International Conference on Advances in Science and Technology, ICAST 2022*, pp. 569–576, 2022, doi: 10.1109/ICAST55766.2022.10039613.

[41] S. Ting Lim and N. Bin Abd Manap, “A Region-based Compression Technique for Medical Image Compression using Principal Component Analysis (PCA).” [Online]. Available: www.ijacsa.thesai.org

[42] Z. Fan, X. Rong, and X. Yu, “REGION OF INTEREST BASED MEDICAL IMAGE COMPRESSION USING BLOCK-TO-ROW BI-DIRECTIONAL PRINCIPAL COMPONENT ANALYSIS,” *ICIC Express Letters ICIC International ©2017*, vol. 11, pp. 1765–1772, 2017.

[43] G. S. Rao, S. S. Rani, and B. P. Rao, “Performance Analysis of CT Scan and X-Ray Medical Images Compression Using PCA and WDR techniques,” *SSRN Electronic Journal*, Mar. 2019, doi: 10.2139/SSRN.3352401.

[44] S. Elamparuthi and N. Puviarasan, “Implementation Of A Hybrid Color Image Compression Technique Using Principal Component Analysis And Discrete Tchebichef Transform,” 2021.

[45] R. K. Senapati, R. Badri, A. Kota, N. Merugu, and S. Sadhul, “Compression and Denoising of Medical Images Using Autoencoders,” *Proceedings - 2022 International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems, ICMACC 2022*, pp. 466–470, 2022, doi: 10.1109/ICMACC54824.2022.10093634.

[46] S. Saravanan and S. Juliet, “Deep medical image reconstruction with autoencoders using deep boltzmann machine training,” *EAI Endorsed Trans Pervasive Health Technol*, vol. 6, no. 24, pp. 1–9, 2020, doi: 10.4108/eai.24-9-2020.166360.

[47] D. Mishra, S. K. Singh, and R. K. Singh, “Lossy Medical Image Compression using Residual Learning-based Dual Autoencoder Model,” *7th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering, UPCON 2020*, Nov. 2020, doi: 10.1109/UPCON50219.2020.9376417.

[48] X. Liu *et al.*, “Medical Image Compression Based on Variational Autoencoder,” *Math Probl Eng*, vol. 2022, 2022, doi: 10.1155/2022/7088137.

[49] J. Ballé, V. Laparra, and E. P. Simoncelli, “End-to-end Optimized Image Compression,” *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, Nov. 2016, Accessed: Dec. 31, 2023. [Online]. Available: https://arxiv.org/abs/1611.01704v3

[50] B. P. Vikraman and A. Jabeena, “Segmentation based medical image compression of brain magnetic resonance images using optimized convolutional neural network,” *Multimed Tools Appl*, pp. 1–19, Sep. 2023, doi: 10.1007/S11042-023-16559-4/METRICS.

[51] S. Shukla and A. Srivastava, “Medical images Compression using convolutional neural network with LWT,” *International Journal of Modern Communication Technologies & Research (IJMCTR)*, vol. 6, no. 6, pp. 34–37, 2018, doi: 10.31873/ijmctr.6.6.45.

[52] Y. Raut, T. Tiwari, P. Pande, and P. Thakar, “Image Compression Using Convolutional Autoencoder,” *Lecture Notes in Electrical Engineering*, vol. 601, pp. 221–230, 2020, doi: 10.1007/978-981-15-1420-3_23/COVER.

[53] G. Guerrisi, F. Del Frate, and G. Schiavon, “Convolutional Autoencoder Algorithm for On-Board Image Compression,” *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2022-July, pp. 151–154, 2022, doi: 10.1109/IGARSS46834.2022.9883256.

[54] Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, “Deep Convolutional AutoEncoder-based Lossy Image Compression,” *2018 Picture Coding Symposium, PCS 2018 - Proceedings*, pp. 253–257, Sep. 2018, doi: 10.1109/PCS.2018.8456308.

[55] Z. Ren, X. Kong, Y. Zhang, and S. Wang, “UKSSL: Underlying Knowledge based Semi-Supervised Learning for Medical Image Classification,” *IEEE Open J Eng Med Biol*, 2023, doi: 10.1109/OJEMB.2023.3305190.

[56] Z. Ren, S. Wang, and Y. Zhang, “Weakly supervised machine learning,” Sep. 01, 2023, *John Wiley and Sons Inc*. doi: 10.1049/cit2.12216.

[57] R. P. N, R. Rajavel, M. Malarvel, and A. Shankar, “Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000. A Quantum-enhanced Artificial Neural Network Model for Efficient Medical Image Compression”, doi: 10.1109/ACCESS.2024.0429000.

[58] P. K. Sran, S. Gupta, and S. Singh, “Segmentation based image compression of brain magnetic resonance images using visual saliency,” *Biomed Signal Process Control*, vol. 62, Sep. 2020, doi: 10.1016/j.bspc.2020.102089.

[59] H. S. Prasantha, “NOVEL APPROACH FOR IMAGE COMPRESSION USING MODIFIED SVD,” 2020. [Online]. Available: www.ijcrt.org

[60] I. Urbaniak and M. Wolter, “Quality assessment of compressed and resized medical images based on pattern recognition using a convolutional neural network,” *Commun Nonlinear Sci Numer Simul*, vol. 95, Apr. 2021, doi: 10.1016/j.cnsns.2020.105582.

[61] O. H. Nagoor, J. Whittle, J. Deng, B. Mora, and M. W. Jones, “LOSSLESS COMPRESSION FOR VOLUMETRIC MEDICAL IMAGES USING DEEP NEURAL NETWORK WITH LOCAL SAMPLING,” 2020.

[62] Q. Min, X. Wang, B. Huang, and Z. Zhou, “Lossless medical image compression based on anatomical information and deep neural networks,” *Biomed Signal Process Control*, vol. 74, Apr. 2022, doi: 10.1016/j.bspc.2022.103499.

[63] Z. Li *et al.*, “An optimized JPEG-XT-based algorithm for the lossy and lossless compression of 16-bit depth medical image,” *Biomed Signal Process Control*, vol. 64, Feb. 2021, doi: 10.1016/j.bspc.2020.102306.

[64] R. Agarwal, C. S. Salimath, and K. Alam, “Multiple Image Compression in Medical Imaging Techniques using Wavelets for Speedy Transmission and Optimal Storage,” *Biomedical and Pharmacology Journal*, vol. 12, no. 1, pp. 183–198, Mar. 2019, doi: 10.13005/bpj/1627.

[65] S. Navaneethakrishnan and G. Shanmugam, “A novel compression methodology for medical images using deep learning for high-speed transmission,” *International Journal of Reconfigurable and Embedded Systems*, vol. 13, no. 2, pp. 262–270, Jul. 2024, doi: 10.11591/ijres.v13.i2.pp262-270.

[66] M. Radiographs *et al.*, “MURA : Large Dataset for Abnormality Detection in,” no. Midl 2018, pp. 1–10.

[67] R. W. Filice *et al.*, “Crowdsourcing pneumothorax annotations using machine learning annotations on the NIH chest X-ray dataset,” 2019.

[68] 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,” 2017, doi: 10.1109/CVPR.2017.369.

[69] J. Shiraishi *et al.*, “Development of a Digital Image Database for Chest Radiographs With and Without a Lung Nodule : Receiver Operating Characteristic Pulmonary Nodules,” no. January, pp. 71–74, 2000.

[70] I. Pan, A. Cadrin-Chênevert, and P. M. Cheng, “Tackling the radiological society of North America pneumonia detection challenge,” *American Journal of Roentgenology*, vol. 213, no. 3, pp. 568–574, May 2019, doi: 10.2214/AJR.19.21512/ASSET/IMAGES/LARGE/09_19_21512_07.JPEG.

[71] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, “INbreast: Toward a Full-field Digital Mammographic Database,” *Acad Radiol*, vol. 19, no. 2, pp. 236–248, Feb. 2012, doi: 10.1016/J.ACRA.2011.09.014.

[72] R. Ali, R. C. Hardie, M. S. De Silva, and T. M. Kebede, “Skin Lesion Segmentation and Classification for ISIC 2018 by Combining Deep CNN and Handcrafted Features,” pp. 1–4, 2018.

[73] S. G. Armato *et al.*, “The Lung Image Database Consortium „ LIDC … and Image Database Resource Initiative „ IDRI … : A Completed Reference Database of Lung Nodules on CT Scans,” no. February, pp. 915–931, 2011.

[74] D. Demner-Fushman *et al.*, “Preparing a collection of radiology examinations for distribution and retrieval,” *Journal of the American Medical Informatics Association*, vol. 23, no. 2, pp. 304–310, Mar. 2016, doi: 10.1093/jamia/ocv080.

[75] M. Jaber, “A deep learning-based iterative digital pathology annotation tool,” pp. 1–22, 2021.

[76] M. R. Lone, “A high speed and memory efficient algorithm for perceptually-lossless volumetric medical image compression,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 2964–2974, Jun. 2022, doi: 10.1016/J.JKSUCI.2020.04.014.

[77] C. Tensmeyer and T. Martinez, “Historical Document Image Binarization: A Review,” *SN Comput Sci*, vol. 1, no. 3, p. 173, May 2020, doi: 10.1007/s42979-020-00176-1.

[78] Z. Ahmed, K. Mohamed, S. Zeeshan, and X. Q. Dong, “Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine,” 2020, *Oxford University Press*. doi: 10.1093/database/baaa010.

[79] S. Elamparuthi and N. Puviarasan, “Implementation Of A Hybrid Color Image Compression Technique Using Principal Component Analysis And Discrete Tchebichef Transform,” *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 10, pp. 5374–5387, Apr. 2021, doi: 10.17762/TURCOMAT.V12I10.5339.

[80] S. Ting Lim and N. Bin Abd Manap, “A Region-based Compression Technique for Medical Image Compression using Principal Component Analysis (PCA),” *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 13, no. 2, p. 2022, Accessed: Dec. 31, 2023. [Online]. Available: www.ijacsa.thesai.org

[81] D. R. Nayak, N. Padhy, P. K. Mallick, and A. Singh, “A deep autoencoder approach for detection of brain tumor images,” *Computers and Electrical Engineering*, vol. 102, Sep. 2022, doi: 10.1016/j.compeleceng.2022.108238.

[82] F. Hussain and J. Jeong, “Efficient deep neural network for digital image compression employing rectified linear neurons,” *J Sens*, vol. 2016, 2016, doi: 10.1155/2016/3184840.

[83] S. Zebang and K. Sei-Ichiro, “Densely connected autoencoders for image compression,” in *ACM International Conference Proceeding Series*, Association for Computing Machinery, 2019, pp. 78–83. doi: 10.1145/3313950.3313965.

[84] Y. Raut, T. Tiwari, P. Pande, and P. Thakar, *Image Compression Using Convolutional Autoencoder*, vol. 601. Springer Singapore, 2020. doi: 10.1007/978-981-15-1420-3_23.

[85] T. Jiang and J. Cheng, “Target Recognition Based on CNN with LeakyReLU and PReLU Activation Functions,” in *Proceedings - 2019 International Conference on Sensing, Diagnostics, Prognostics, and Control, SDPC 2019*, Institute of Electrical and Electronics Engineers Inc., Aug. 2019, pp. 718–722. doi: 10.1109/SDPC.2019.00136.

[86] F. Jiang, W. Tao, S. Liu, J. Ren, X. Guo, and D. Zhao, “An End-to-End Compression Framework Based on Convolutional Neural Networks,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, pp. 3007–3018, Oct. 2018, doi: 10.1109/TCSVT.2017.2734838.

[87] G. Toderici *et al.*, “Full resolution image compression with recurrent neural networks,” *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 5435–5443, 2017, doi: 10.1109/CVPR.2017.577.

[88] J. Ballé, V. Laparra, and E. P. Simoncelli, “End-to-end optimized image compression,” *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, 2017.