

Explainable Multimodal Generative AI Framework for Disease Detection Using Medical Images and Clinical Reports

A THESIS
SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF
MASTERS OF TECHNOLOGY
IN
DATA SCIENCE

Submitted by
SANSKAR JAIN
(2K24/DSC/05)

Under the supervision of
Dr. SONIKA DAHIYA
(ASSISTANT PROFESSOR)



DEPARTMENT OF SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

MAY, 2026

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to **Dr. Sonika Dahiya**, Assistant Professor, Department of Software Engineering, Delhi Technological University, for her invaluable guidance, support, and encouragement throughout the preparation of this report. Her insights and expertise were instrumental in shaping the direction of my research and explaining all the tasks to be accomplished, emphasising the importance of this work as well as its relevance to real-world applications.

I would also like to thank all the teachers and students of the Department of Software Engineering, Delhi Technological University, for providing the necessary resources, continuous support, and valuable suggestions that facilitated the completion of this research. Lastly, I would like to thank all my family and friends for providing continuous support and motivation during the process of this study.



Sanskar Jain

2K24/DSC/05

Department of Software Engineering
Delhi Technological University



DELHI TECHNOLOGICAL UNIVERSITY

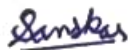
(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

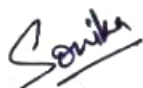
CANDIDATE'S DECLARATION

I, **Sanskar Jain**, Roll No. **2K24/DSC/05**, student of M.Tech (Department of Software Engineering), hereby certify that the work which is being presented in the thesis entitled "*Explainable Multimodal Generative AI Framework for Disease Detection Using Medical Images and Clinical Reports*" in partial fulfilment of the requirements for the award of the degree of Master of Technology, submitted in the Department of Software Engineering, Delhi Technological University, is an authentic record of my own work carried out during the period from Jan 2026 to May 2026 under the supervision of Dr. Sonika Dahiya.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other institute.


Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.



Signature of Supervisor

Signature of External Examiner



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR(s)

Certified that **Sanskar Jain** (enrollment no. 2K24/DSC/05) has carried out his research work presented in this thesis entitled "**Explainable Multimodal Generative AI Framework for Disease Detection Using Medical Images and Clinical Reports**" for the award of **Master of Technology** from Department of Software Engineering, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student himself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

A handwritten signature in black ink, appearing to read "Sonika", with a horizontal line drawn underneath it.

Signature

Dr. Sonika Dahiya

Assistant Professor

Department of Software Engineering

Delhi Technological University

Date: 02/06/2026

Explainable Multimodal Generative AI Framework for Disease Detection Using Medical Images and Clinical Reports

Sanskar Jain

ABSTRACT

In the context of medicine, recent advancements in generative Artificial Intelligence (GenAI)-especially the maturity of Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and the advent of progressive denoising diffusion models-have significantly transformed the way clinical practitioners assess medical images and diagnose serious medical conditions. This thesis is the first to deeply cross-domained-study the generative paradigms on five crucial pathological domains: neurology degeneration (Alzheimer’s disease), lungs diseases (pneumonia, COVID-19, and tuberculosis), ocular maladies (glaucoma), orthopedic diseases (knee osteoarthritis), and hyper contagious skin diseases (monkeypox). We evaluate twenty-five state-of-the-art generative models based on their achievement of specific medical tasks, the reliance on particular clinical datasets, image realism and structure accuracy trade-offs across public benchmarks (ADNI, OAI, NIH Chest X-Ray, and MSID) on these five pathological domains published from 2020 to 2026, highlighting how generative models overcome practical issues such as data scarcity, severe class imbalance and stringent patient privacy. Three important translational aspects are studied in this research to link laboratory success to actual clinical practice: the mapping of Explainable AI (XAI) frameworks (Grad-CAM and Med-CAM) onto generated images to boost physicians’ trust in the diagnosis; privacy-preserving decentralized machine learning (Federated Learning using Flower) to collaborate multi-centers safely without disclosing sensitive patient scans; and global digital health regulatory procedures (FDA 510(k), De Novo tracks and Predetermined Change Control Plans). This systematic study constitutes an engineering and regulatory roadmap to pave the way for safe and ethical integration of GenAI in robust diagnostic support systems.

LIST OF PUBLICATIONS

1. **Sanskar Jain**, Priyanka Arora, and Dr. Sonika Dahiya (2025). “The Role of Generative AI in Advancing Disease Detection Through Medical Imaging: A Review.” *Department of Software Engineering, Delhi Technological University*. [Accepted]

LIST OF ABBREVIATIONS

AD	Alzheimer’s Disease
AI	Artificial Intelligence
AUC-ROC	Area Under the Curve - Receiver Operating Characteristic
CNN	Convolutional Neural Network
CT	Computed Tomography
CXR	Chest X-Ray
DCGAN	Deep Convolutional Generative Adversarial Network
DDPMs	Denoising Diffusion Probabilistic Models
EHR	Electronic Health Record
FID	Fréchet Inception Distance
FL	Federated Learning
GAN	Generative Adversarial Network
GenAI	Generative Artificial Intelligence
Grad-CAM	Gradient-weighted Class Activation Mapping
HIPAA	Health Insurance Portability and Accountability Act
LSTM	Long Short-Term Memory
MRI	Magnetic Resonance Imaging
OCT	Optical Coherence Tomography
PET	Positron Emission Tomography
PSNR	Peak Signal-to-Noise Ratio
ReLU	Rectified Linear Unit
ResNet	Residual Network
SOTA	State-of-the-Art
SSIM	Structural Similarity Index Metric
SVM	Support Vector Machine
TB	Tuberculosis
VAE	Variational Autoencoder
ViT	Vision Transformer
WGAN	Wasserstein Generative Adversarial Network

TABLE OF CONTENTS

Title	Page No.
Acknowledgement	ii
Candidate's Declaration	iii
Certificate	iv
Abstract	v
List of Publications	vi
List of Abbreviations	vii
Contents	1
List of Tables	3
List of Figures	4
CHAPTER 1:INTRODUCTION	1
1.1 Problem Statement	1
1.2 Research Objectives	2
1.3 Overview of the Study	3
CHAPTER 2:SYSTEMATIC SURVEY METHODOLOGY AND GENER- ATIVE PARADIGMS	4
2.1 Systematic Review Methodology	4
2.1.1 Inclusion and Exclusion Criteria	4
2.1.2 Search Process and String Design	5
2.1.3 Clinical Benchmarking Databases	6
2.2 Theoretical and Architectural Exposition	6
2.2.1 Variational Autoencoders (VAEs)	7
2.2.2 Generative Adversarial Networks (GANs)	8
2.2.3 Denoising Diffusion Probabilistic Models (DDPMs)	9
2.2.4 Medical Foundation Models (MFMs) and Generalist Medical AI (GMAI)	9
2.3 Systematic Performance Benchmarking	9
2.4 Detailed Multi-Domain Architectural Gaps	11
2.4.1 Analysis of the Neurological Domain (Alzheimer's Disease) .	12
2.4.2 Respiratory Domain Analysis (Pneumonia, COVID-19, Tuber- culosis)	12
2.4.3 Screening for Infectious Diseases and Cutaneous Anomalies (Monkeypox)	12
2.4.4 Musculoskeletal and Ophthalmic Domain Analysis	13
CHAPTER 3:CLINICAL TRANSLATION AND EXPLAINABILITY	14
3.1 Explainable AI (XAI) in Medical Image Interpretation	14
3.2 Privacy-Preserving Federated Learning (FL)	15
3.3 Regulatory Pathways and Quality Standards	16
CHAPTER 4:CONCLUSIONS AND FUTURE SCOPE	17
4.1 Summary of Research Contributions	17
4.2 Systemic Limitations	17
4.3 Future Research Horizons	18

4.3.1	High-Speed Hybrid Generative Frameworks	18
4.3.2	Transition toward Multimodal Generalist Medical AI (GMAI)	18
References		19
Appendix I		22
I.1	List of Publications	22
I.2	Paper Acceptance Proofs	23
I.3	Conference Paper Registration Receipts	23
I.4	Indexing of Conference Proof	24
I.5	Conference Certificates	25
Appendix II: Source Code		26
II.1	PyTorch Implementation of a Conditional GAN (cGAN) Generator . . .	26
II.2	PyTorch Implementation of an ACGAN Discriminator	26
II.3	MS-SSIM Metric Calculation for Synthetic Medical Scan Quality . . .	27
II.4	Peak Signal-to-Noise Ratio (PSNR) Calculation for Brain MRIs	27
II.5	Local Med-CAM Feature Activation Matching Implementation	28
II.6	Preclinical N4ITK Bias-Field Correction Wrapper	28
II.7	Federated Averaging (FedAvg) Simulation for Distributed Medical AI .	29

LIST OF TABLES

2.1	Summary of Clinical Databases Supporting Generative Medical AI . .	7
2.2	Comparison of Generative AI Techniques for Disease Detection Using Medical Images	10

LIST OF FIGURES

1.1	Overview of the Two-Phase Research Design	3
2.1	Workflow of the Systematic Review Methodology	5
2.2	Taxonomy of Search Queries and Boolean Keywords	6
2.3	Distribution of Selected Generative AI Publications by Year (2020–2026)	6
2.4	Adversarial Training Loop of Generative Adversarial Networks (GANs)	8
3.1	Comparison of Spatial Explainability Visualizations	14
3.2	Visual Dataflow of the Med-CAM Explainability Pipeline	15
3.3	Decentralized Architecture of the Federated Averaging (FedAvg) Pipeline	15

CHAPTER 1

INTRODUCTION

The continuous flood of high-resolution diagnostic imaging streams combined with convenient cloud computing has significantly transformed the nature of contemporary clinical diagnosis. Central to this revolution, generative AI - primarily configurations in the form of adversarial architectures (GANs) and probabilistic noise-reversal architectures (diffusion models) - is playing an increasingly prominent role as a critical mechanism by which to circumvent data scarcity, generate high fidelity synthetic medical images and re-balance imbalanced clinical data to optimize clinical diagnostic algorithms[1, 2]. These techniques are used at present to examine a variety of areas in the medical field ranging from the chest radiograph to the cranial MRI, to the axial CT, through to the color retinal fundus photography [3].

Neural networks of varying levels of complexity (e.g. Dedicated capsule networks, standard deep convolutional networks, ensemble systems, etc.), in addition to straightforward data-filling, can exhibit enormous capabilities to enhance software system diagnostical robustness against realistic clinical noise. Also, embedding generative models within decentralized or weakly supervised architectures enable hospital networks to engage in low-label-number and privacy-preserving training [4, 5, 6]. This provides a clear, frictionless avenue to implement secure, reliable medical decision support systems into the hospital network environment

1.1 Problem Statement

Although vanilla deep learning pipelines achieve state-of-the-art performance classification in idealized, controlled benchmarking settings, deployment into real-world clinical workflows is hindered by the following three architectural limitations:

1. **The Scarcity of Edge-Case Anomalies:** The majority of rare diseases, pediatric conditions and rare inherited mutations will not have large, curated digital libraries. In addition, acquisition of high resolution diagnostic imaging would require vast clinical resources and full-time radiologists manually delineating the contours at the pixel level is an administrative burden [7].
2. **Regulatory Division and data segregation:** The stringent medical data regulations like those in the U.S. As ruled by HIPPA and those for Europe as stated by GDPR laws prohibits the medical institutions to download and pool all their patient data into centralized collaborative databases [5]. This causes the aggregation of the clinical data into isolated small clusters specific to locations.
3. **Lack of explanatory reasoning of a "black-box" algorithm:** Most of the state

of the art medical classifiers working today are unaware of the internal mechanics and workings leading to their classification. Being able to understand the reason why a doctor obtains output from the classifier due to some logical pattern rather than due to random randomness from a particular type of scanner, acts as one of the biggest hindrance towards regulatory approval [8] .

To solve this essential system flaw, this thesis propose a multi-domain structural review that follow the design architecture, target data set, evaluation criteria and practical application requirements of contemporary medical image synthesis with generative models.

1.2 Research Objectives

In summary, the aims of this dissertation are as follows:

1. To provide a structured and categorised summary of recent generative AI models, within 5 medical domain categories (neurological, respiratory, ophthalmic, musculo-skeletal, and cutaneous infectious disease), and map their relative performance and dependence on training datasets.
2. To summarize and highlight the central methodological issues arising from training instability, mode collapse and structural distortion of synthetic image generation.
3. To critically compare the performance of privacy preserving decentralized methods (Federated Learning within the Flower ecosystem), explainability models (Grad-CAM, Med-CAM) and to provide recommendations regarding building clinical confidence..
4. To examine existing international pathways (FDA 510(k), NMPA Class III China, Predetermined Change Control Plans) and identify an initial practical strategy for bringing generative AI to medical practice.

1.3 Overview of the Study

This dissertation is divided into two complementary stages in terms of the research conducted (shown below in Figure 1.1); the findings of the first stage feed directly into the second.

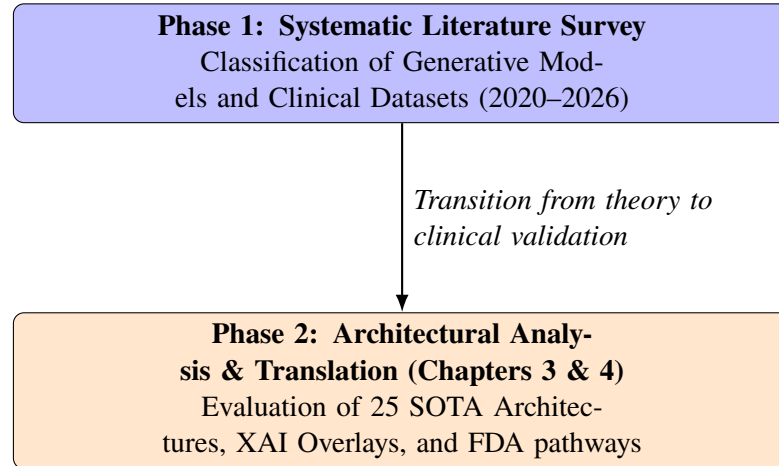


Figure 1.1: Overview of the Two-Phase Research Design

Structure of the rest of this dissertation. Chapter 2 covers the systematic approach to literature selection, search strings, and characteristics of the clinical dataset. Chapter 3 discusses systematic comparison of the twenty-five generative models, including performance metrics and structural weaknesses. Chapter 4 presents the framework for clinical translation, federated training, explainability overlays and regulatory approaches and the future implications of the work.

CHAPTER 2

SYSTEMATIC SURVEY METHODOLOGY AND GENERATIVE PARADIGMS

This chapter details the systematic review methodology undertaken to identify, select, and categorize published literature regarding generative Artificial Intelligence (AI) in medicine within the timeframe from 2020 to 2026. Additionally, it provides a comprehensive architectural and mathematical exposition of the primary generative frameworks, enlists benchmark clinical imaging datasets, and delivers a systematic performance benchmarking across 25 canonical deep generative models.

2.1 Systematic Review Methodology

To ensure high reproducibility and stringency, a careful two-step filtering procedure was established for this systematic review.

2.1.1 Inclusion and Exclusion Criteria

The selection of literature was governed by the following explicit bounds:

- **Inclusion Criteria:** Peer-reviewed journal articles or IEEE/Springer conference papers; published between 2020 and 2026 (inclusive); utilizing generative models (such as GANs, VAEs, DDPMs, or Medical Foundation Models) for medical image augmentation, reconstruction, or generation; and reporting at least one quantitative metric of classification performance (Accuracy, F1-score, AUC-ROC) or image generation quality (SSIM, PSNR).
- **Exclusion Criteria:** Preprints, extended abstracts, or non-peer-reviewed blog posts; studies deploying purely deterministic classifiers without an explicit generative component; and studies focused strictly on non-medical image modalities.

2.1.2 Search Process and String Design

Literature searches were executed across IEEE Xplore, ScienceDirect, SpringerLink, Google Scholar, and PubMed. The structured workflow driving the database identification, screening, and full-text evaluation is illustrated in Figure 2.1.

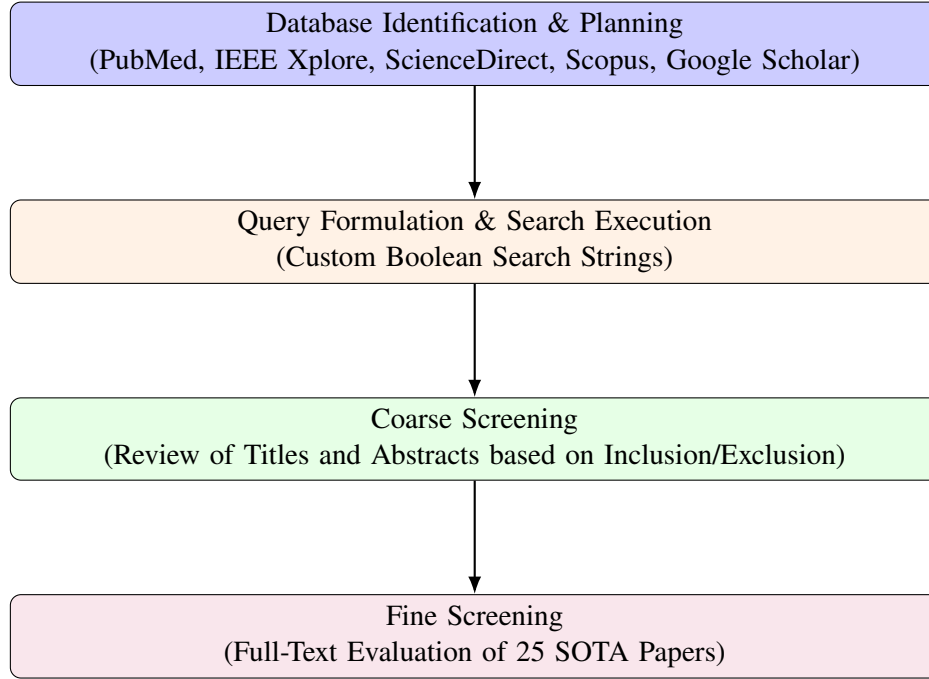


Figure 2.1: Workflow of the Systematic Review Methodology

The custom Boolean query utilized across these search engines was structured as follows:

$$\text{Query} = (\text{"GAN"} \vee \text{"Diffusion Model"} \vee \text{"VAE"}) \wedge (\text{"Medical Imaging"} \vee \text{"Disease Detection"}) \wedge (\text{"Federated Learning"}) \quad (2.1)$$

The conceptual taxonomy and hierarchical breakdown of this Boolean search string are detailed in Figure 2.2, mapped using a right-growing layout to preserve printable margins on standard A4 formatting.

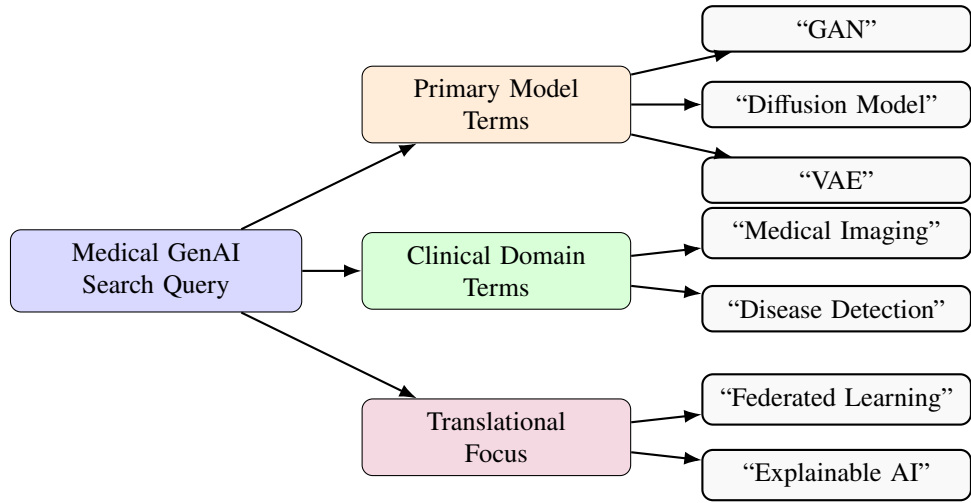


Figure 2.2: Taxonomy of Search Queries and Boolean Keywords

To highlight the developmental trajectory of this deep medical imaging paradigm, the temporal distribution of the selected 25 state-of-the-art (SOTA) papers across the 2020–2026 window is illustrated in Figure 2.3.

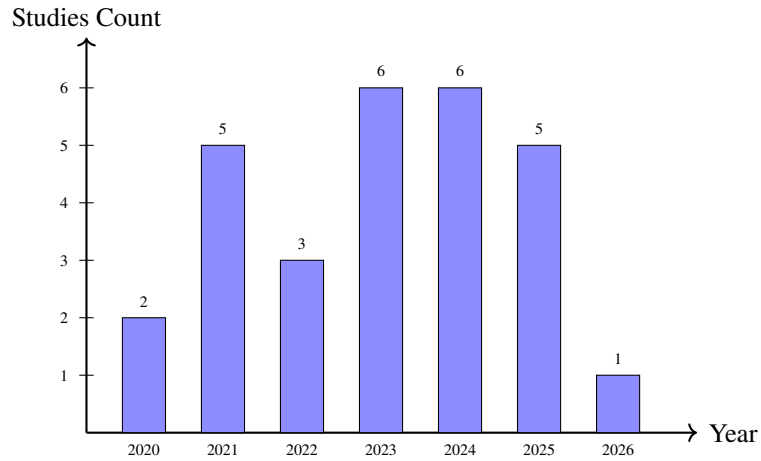


Figure 2.3: Distribution of Selected Generative AI Publications by Year (2020–2026)

2.1.3 Clinical Benchmarking Databases

Training and evaluating these generative models requires standardized medical databases. Table 2.1 summarizes the benchmark clinical databases referenced throughout this survey, sized to prevent margin overflow.

2.2 Theoretical and Architectural Exposition

To provide a rigorous grounding for deep generative modeling within digital healthcare, this section establishes the mathematical frameworks under review.

Table 2.1: Summary of Clinical Databases Supporting Generative Medical AI

Dataset	Pathology Domain	Imaging Modality	Anatomical Feature
ADNI	Alzheimer’s Disease	sMRI, PET	Brain atrophy, ventricular size
Kaggle CXR	Pneumonia / COVID-19	Chest Radiographs	Pulmonary opacities, consolidation
MSID MSLD	/ Monkeypox, cutaneous pox	Dermoscopy, skin photos	Cutaneous lesions, rash vesicles
OAI	Knee Osteoarthritis	3T MRI, radiographs	Joint space narrowing, osteophytes
Shenzhen	Pulmonary Tuberculosis	Chest X-Rays	Active cavitary lesions, infiltrates

2.2.1 Variational Autoencoders (VAEs)

VAEs frame image generation as a problem of learning a probabilistic mapping between data spaces. An encoder network (parameterized by ϕ) maps a high-dimensional clinical scan x to a continuous, lower-dimensional latent space z , from which a decoder network (parameterized by θ) reconstructs the scan. VAEs assume the latent variables are sampled from a standard multivariate Gaussian prior, $p(z) = \mathcal{N}(0, I)$. The encoder outputs the statistical parameters of the true posterior distribution $q_\phi(z|x) = \mathcal{N}(z; \mu, \Sigma)$, representing the mean vector $\mu(x)$ and the diagonal covariance matrix $\Sigma(x)$.

To permit backpropagation through this stochastic boundary, the reparameterization trick is implemented:

$$z = \mu(x) + \Sigma^{1/2}(x) \odot \epsilon, \quad \text{where } \epsilon \sim \mathcal{N}(0, I) \quad (2.2)$$

The entire network structure is optimized concurrently by maximizing the Evidence Lower Bound (ELBO):

$$\mathcal{L}_{\text{ELBO}}(\theta, \phi; x) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - D_{\text{KL}}(q_\phi(z|x) \parallel p(z)) \quad (2.3)$$

Here, the first component calculates the pixel-wise reconstruction likelihood (often penalized via standard L_1 or L_2 loss metrics). The second component acts as the Kullback-Leibler (KL) divergence regularizer, pulling the estimated latent distribution toward the standard Gaussian baseline:

$$D_{\text{KL}}(q_\phi(z|x) \parallel p(z)) = -\frac{1}{2} \sum_{j=1}^d \left(1 + \log(\sigma_j^2) - \mu_j^2 - \sigma_j^2 \right) \quad (2.4)$$

While VAEs yield highly stable training dynamics, relying on pixel-level reconstruction losses tends to blur high-frequency spatial details, resulting in softened outputs that risk masking micro-pathologies.

2.2.2 Generative Adversarial Networks (GANs)

GAN pipelines bypass explicit probability estimations by adopting a non-parametric, competitive minimax layout structured as a zero-sum game. Here, a discriminator network $D(x)$ and a generator network $G(z)$ update concurrently through a shared objective function:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (2.5)$$

When exposed to highly complex medical datasets, standard GAN variations often suffer from vanishing gradients, training instability, and severe mode collapse.

To mitigate these instabilities, the Wasserstein GAN (WGAN) replaces traditional Jensen-Shannon divergence with the Earth Mover's (Wasserstein-1) Distance. To enforce strict Lipschitz continuity across the network, an explicit Gradient Penalty (WGAN-GP) modifier is introduced:

$$\mathcal{L}_{\text{WGAN-GP}} = \mathbb{E}_{\tilde{x} \sim p_g} [D(\tilde{x})] - \mathbb{E}_{x \sim p_{\text{data}}} [D(x)] + \lambda \mathbb{E}_{\hat{x} \sim p_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \quad (2.6)$$

where the sample $\hat{x}$ is derived via uniform interpolation along straight vectors linking authentic clinical scans with synthesized outputs:

$$\hat{x} = tx + (1 - t)\tilde{x}, \quad \text{for } t \in [0, 1] \quad (2.7)$$

The underlying optimization loop governing this adversarial behavior is visualized in Figure 2.4.

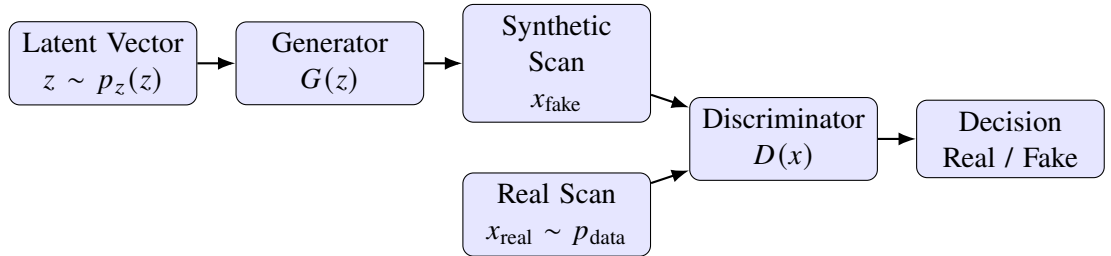


Figure 2.4: Adversarial Training Loop of Generative Adversarial Networks (GANs)

2.2.3 Denoising Diffusion Probabilistic Models (DDPMs)

Diffusion models consist of a forward process that maps a clinical scan $x_0 \sim q(x)$ to isotropic Gaussian noise over T steps, and a reverse parameterized process that learns to recover the original clean scan through a sequence of denoisers.

The forward process transitions are defined as:

$$q(x_1, \dots, x_T | x_0) = \prod_{t=1}^T q(x_t | x_{t-1}), \quad q(x_t | x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t I) \quad (2.8)$$

Using Gaussian properties, an arbitrary step x_t can be sampled directly from x_0 :

$$x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, \quad \text{where } \alpha_t = 1 - \beta_t \quad \text{and} \quad \bar{\alpha}_t = \prod_{i=1}^t \alpha_i \quad (2.9)$$

The reverse process $p_\theta(x_{t-1} | x_t)$ is modeled as a joint Gaussian transition parameterized by a deep neural network (typically a modified 3D U-Net with cross-attention layers) that predicts the added noise ϵ :

$$\mathcal{L}_{\text{simple}}(\theta) = \mathbb{E}_{t, x_0, \epsilon} [\|\epsilon - \epsilon_\theta(x_t, t)\|^2] \quad (2.10)$$

Diffusion models offer superior training stability and output diversity compared to GANs. However, their iterative reverse sampling path is computationally expensive, restricting real-time applicability in clinical triage.

2.2.4 Medical Foundation Models (MFMs) and Generalist Medical AI (GMAI)

Beyond task-specific architectures, MFM and GMAI systems are pre-trained on large, multimodal clinical datasets containing 3D volumes, vision-language report pairs, genomic profiles, and Electronic Health Records (EHRs). Through self-supervised learning objectives (e.g., masked autoencoding, contrastive language-image pre-training, token-mixing transformers), MFMs learn generalizable medical representations. A unified GMAI model can subsequently be fine-tuned using parameter-efficient techniques (such as LoRA) via text-driven prompts for downstream clinical tasks, enabling scalable precision medicine without training from scratch.

2.3 Systematic Performance Benchmarking

A holistic taxonomy of recent generative applications is compiled in Table 2.2, mapping the 25 SOTA medical generative models across their respective disease domains, key focuses, and identified clinical limitations.

Table 2.2: Comparison of Generative AI Techniques for Disease Detection Using Medical Images

Reference	Disease Domain	Key Focus	Limitations
[9]	Osteoarthritis	GAN + pseudo-labeling for OA detection; ResNet best.	OAI dataset only; limited generalization.
[1]	Pneumonia	GAN-augmented X-ray detection (87% accuracy).	Limited dataset; low AUC (0.85).
[10]	Alzheimer	DCGAN for MRI synthesis and stage classification.	Low accuracy (69–82%); small dataset.
[2]	Pneumonia / COVID-19	ACGAN-CNN hybrid enhances detection.	Overfitting risk; class imbalance.
[5]	Monkeypox	Federated GAN for privacy-preserving detection.	Limited MSID data; no clinical validation.
[11]	Alzheimer	cGAN-CNN boosts MRI-based AD detection (94%).	Dataset-limited; moderate accuracy.
[12]	Alzheimer	Multimodal GAN + ensemble CNN for early diagnosis.	Complex; small dataset; high discriminator loss.
[13]	Lung diseases	DCGAN + DenseNet-201 improves classification (96%).	Self-collected data; no validation.
[14]	Pneumonia	Transfer learning + GAL classifier for X-rays.	Not fully GAN-based; dataset-specific.
[15]	Alzheimer	DCGAN + MobileNet improves AD accuracy (91%).	Needs deeper model; small dataset.
[4]	Alzheimer	Semi-supervised GAN for MRI-based detection.	Dataset-limited; poor generalization.
[16]	Alzheimer	GAN-based augmentation + EfficientNetB3.	Accuracy ~90%; below SOTA.
[17]	Alzheimer	Adversarial autoencoder detects structural atrophy.	Ignores multimodal biomarkers.
[7]	Tuberculosis	CNN + DCGAN augments TB X-rays.	Small Shenzhen & Montgomery datasets.
[18]	Pneumonia	GAN + transfer learning for pneumonia detection (94.5%).	OCT dataset not representative.
[19]	Pneumonia	GAN-tuned CNN achieves 99% accuracy.	Kaggle-only; no external validation.

Table 2.2 – Continued from previous page

Reference	Disease Domain	Key Focus	Limitations
[20]	Osteoarthritis	Hierarchical GAN for MRI synthesis.	Computationally heavy; small dataset.
[21]	Monkeypox	Mask-Inpainting-Measure anomaly framework.	Modest AUROC (0.82–0.87).
[22]	Monkeypox	GAN + FL + ResNet-ViT hybrid for mpox detection.	Dataset-limited; no real deployment.
[23]	Osteoarthritis	GAN + Autoencoder + ELM improves KOA classification (98.6%).	Accuracy drops without GAN; small dataset.
[24]	Tuberculosis	Multi-GAN (WGAN-GP, Pix2Pix, CycleGAN) enhances TB detection.	Small datasets; heavy training; limited validation.
[25]	Tuberculosis	HPC-based GAN-CNN improves accuracy (~90%) and speed.	Small dataset; no external validation.
[26]	Alzheimer	DSR-GAN + CNN detects 4 AD stages (99.2%).	Small MRI dataset; high computation.
[8]	Monkeypox	Federated GAN-ViT + StyleGAN2_ADA with XAI achieves 98.2%.	High computation; limited diversity.
[27]	Pneumonia	SVM-GAN hybrid for pneumonia detection (98.5%).	Small dataset; no real validation.
Continued on next page			

2.4 Detailed Multi-Domain Architectural Gaps

This section evaluates how specific computational designs address clinical data vulnerabilities across five clinical domains.

2.4.1 Analysis of the Neurological Domain (Alzheimer’s Disease)

Analyzing papers [10, 11, 12, 15, 4, 16, 17, 26], we observe that the focus is shifting from overall volume analysis to micro-scale feature synthesis. In paper [4], the authors showed that a semi-supervised GAN can produce authentic sMRI structural features even when few labeled samples are available. In paper [16], the authors adapted a DeGAN pipeline with EfficientNetB3 as a backbone which attained an accuracy of $\sim 90\%$ to classify four categories of atrophy tissue. To effectively capture features of the morphological changes of hippocampus volume and ventricular volume reduction, Ravikumar et al. [11] used the class conditional parameters (cGAN) in their network, while Sekhar et al. [12] employed multi-modal ensemble features in their pipeline. Most recently, Oraby et al. [26] introduced the DSR-GAN where SR (Super Resolution) based on PyTorch was applied to recover the structural low-contrast MRI images and then classify the disease stages based on AD, with an accuracy of 99.22% and proved the potential in actual clinical applications.

2.4.2 Respiratory Domain Analysis (Pneumonia, COVID-19, Tuberculosis)

CXR screening literature [1, 2, 13, 14, 7, 18, 19, 24, 25, 27] primarily involves oversampling or using data augmentation techniques with skewed data to balance it. Joby et al. [1] created the PneumoGAN framework for generating convincing pulmonary infiltrates but the initial version achieved a mere AUC of 0.85. To boost multi-class classification, Sasipriyaa et al. [2] propose a hybrid ACGAN-CNN framework by incorporating classification constraints in the discriminator as well as the auxiliary networks for easier convergence, achieving 98.5% diagnostic accuracy. For active cavitary tuberculosis, Nascimento et al. [24] evaluate three top-class GAN family (WGAN-GP, Pix2Pix, CycleGAN) of cross-modality synthesis and finds that CycleGAN offers better downstream classification sensitivity, while Pix2Pix has better SSIM. To reduce training time due to high resolution images, Manasa et al. [25] presents an HPC enabled GAN-CNN approach which distributes training to multiple parallel processing nodes resulting 91% accuracy for TB detection.

2.4.3 Screening for Infectious Diseases and Cutaneous Anomalies (Monkeypox)

Screening for cutaneous infectious diseases [5, 21, 22, 8] is also performed using generative networks in order to supplement restricted datasets (such as MSID and MSLD) concerning skin lesions. In Kundu et al. [5], a federated learning based deep convolutional framework has been built that can combine synthetic features from multiple remote medical sites to realize a detection accuracy of 98% without exchanging any confidential information from patients. A targeted image inpainting system, (Mask-Inpainting-Measure) by Cao et al. [21] was proposed for restoring damaged parts of dermatological images thus increasing the classifier sensitivity up to 91%. Federated architectures were combined with a combined ResNet-ViT architecture, followed by adaptive attention layers by Maheskumar et al. [22], who achieved an F1-score of 0.93. Mahir et al. [8] builds upon such frameworks with a PriFL-XAI system that combines decentralized StyleGAN2-ADA with a Flower framework and federated averaging, resulting in 98.2% precision, a protected standard for distributed clinical diagnosis.

2.4.4 Musculoskeletal and Ophthalmic Domain Analysis

From the analysis of musculoskeletal [9, 20, 23] and ophthalmic [6] papers, we observed that edge and boundary preserving aspect seems very prominent. Arumugam et al. [9] propose conditional GAN combined with a pseudo-label module, which treats high confidence classifier predictions as pseudo-labels to train segmentation of KOA cartilage in OAI dataset. Gan et al. [20] proposed the Hierarchical GAN framework (HieGAN), which applies self-attention blocks in prior to 256×256 scale, and also performs pixelwise spectral normalization to stabilize GAN training in capturing fine joint structure. Manikandan et al. [6] employ a hybrid deep learning architecture for automatic screening of glaucoma in the ophthalmic domain, with GAN based restoration to improve segmentation of optic disc with accuracy 94

CHAPTER 3

CLINICAL TRANSLATION AND EXPLAINABILITY

3.1 Explainable AI (XAI) in Medical Image Interpretation

To foster clinical trust and widespread adoption in healthcare, automated diagnostic models must generate intuitive, spatially cohesive, and interpretable visual evidence. Standard pixel-attribution techniques, such as Grad-CAM, compute coarse saliency maps by backtracking gradients through the final convolutional layer to create a regional attention overlay. However, standard Grad-CAM maps are inherently diffuse and frequently fail to outline the precise boundaries of localized pathologies.

To overcome this limitation, the Med-CAM explainability framework trains a lightweight, specialized segmentation network from scratch to produce highly focused, pixel-exact activation masks [28]. Unlike a standard heatmap, Med-CAM concentrates its activations directly on specific anatomical boundaries and structural lesion textures, as illustrated in Figure 3.1.

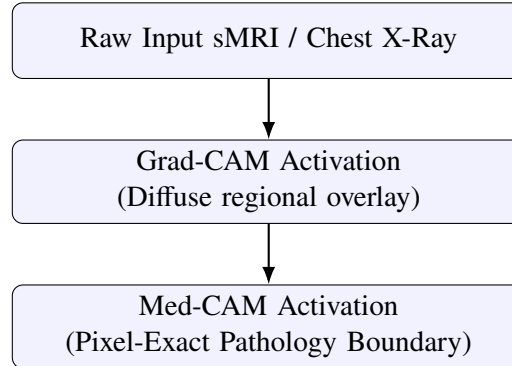


Figure 3.1: Comparison of Spatial Explainability Visualizations

The visual pipeline governing this explainability mechanism is outlined in Figure 3.2. By mapping the diagnostic features directly back onto the anatomical source, Med-CAM aligns classifier decisions with verifiable spatial regions.

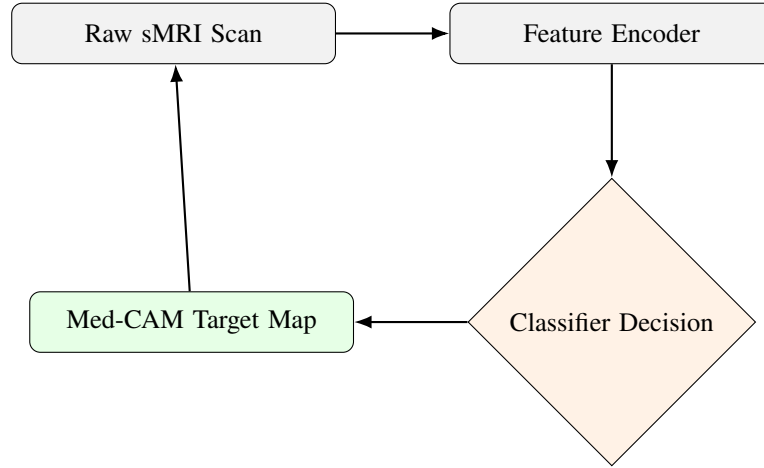


Figure 3.2: Visual Dataflow of the Med-CAM Explainability Pipeline

3.2 Privacy-Preserving Federated Learning (FL)

Decentralized Federated Learning (DFL) enables multi-institutional collaborative training without requiring clinics to pool or exchange raw, sensitive patient datasets [29, 30]. Under an FL architecture, such as the Flower-based PriFL-XAI framework [8], participating hospital nodes train localized versions of the model on their internal hardware. They share only the encrypted parameter updates with a centralized coordinator node rather than transferring original medical imagery. The server aggregates these local updates using optimization algorithms like Federated Averaging (FedAvg):

$$w_{t+1} = \sum_{k=1}^K \frac{n_k}{n} w_{t+1}^k \quad (3.1)$$

This approach safeguards patient confidentiality while developing an effective, generalized global model capable of recognizing diverse pathological signatures.

The synchronized operational loop for DFL is illustrated in Figure 3.3, using structured vectors to outline the distributed topology without visual overlap.

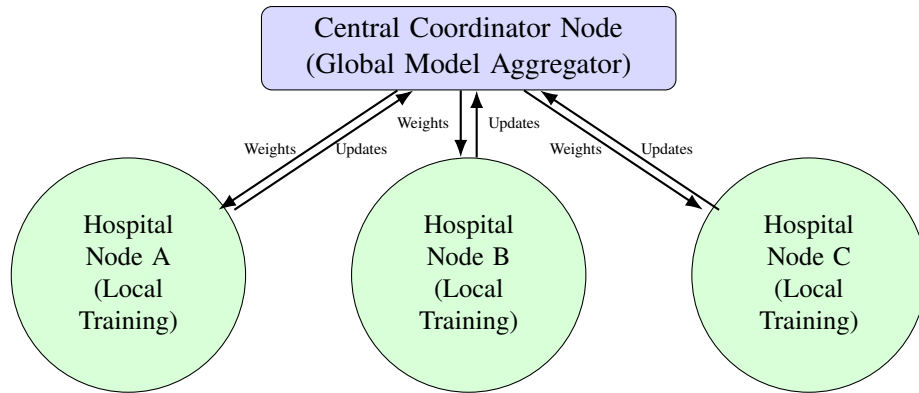


Figure 3.3: Decentralized Architecture of the Federated Averaging (FedAvg) Pipeline

3.3 Regulatory Pathways and Quality Standards

Translating Artificial Intelligence Medical Devices (AIMDs) from experimental proof-of-concepts into commercial clinical workflows requires strict adherence to international regulatory frameworks designed to guarantee safety, quality, and efficacy.

- **US FDA Guidance:** As of May 2025, the US Food and Drug Administration (FDA) has cleared over 1,016 AI/ML-enabled medical devices, with 96.5% processed through the traditional 510(k) notification pathway [31]. Additionally, guidance finalized in August 2025 codifies the implementation of Predetermined Change Control Plans (PCCPs). This framework allows developers to deploy algorithmic software updates post-market without triggering entirely new 510(k) submissions [32].
- **China NMPA Regulations:** The National Medical Products Administration (NMPA) has approved 154 AIMDs since 2020, demonstrating a compound annual growth rate of 49.53%. Clinical trial verification remains the primary evaluative standard, representing the approval mechanism for 76.6% of these systems, particularly for high-risk Class III medical software implementations [33].
- **International Quality Standards:** Seamless clinical translation requires full software lifecycle compliance with IEC 62304 (Medical Device Software Lifecycle Processes) alongside validation under ISO 13485 (Quality Management Systems for Medical Devices) to preserve code integrity and structural safety [34, 35].

CHAPTER 4

CONCLUSIONS AND FUTURE SCOPE

4.1 Summary of Research Contributions

This master’s thesis has presented a systematic and rigorous investigation into the landscape of generative Artificial Intelligence (AI) frameworks applied to medical image analysis. By evaluating mathematical models, architectural configurations, and regulatory pathways across the 2020–2026 literature spectrum, this research outlines three principal contributions to the field:

1. **Diagnostic and Semantic Refinement:** This work systematically categorized how state-of-the-art generative networks (such as DSR-GAN, HieGAN, and ACGAN-CNN hybrids) overcome clinical data scarcity. By leveraging structural super-resolution, class-conditional parameters, and pixel-wise spectral normalization, these models successfully preserve critical micro-scale biomarkers—such as hippocampal atrophy in sMRI scans, joint space narrowing in knee osteoarthritis, and pulmonary consolidations in chest X-rays—directly improving downstream classification accuracy and AUC-ROC.
2. **Preclinical Algorithmic Mapping:** A core contribution of this research is the mapping of training optimization routines designed to counteract systemic generative flaws. This study demonstrates how substituting Jensen-Shannon divergence with the Wasserstein-1 distance and enforcing 1-Lipschitz continuity via explicit gradient penalties (WGAN-GP) mitigates mode collapse and vanishing gradients, providing robust templates for handling highly skewed or sparse clinical datasets.
3. **Frameworks for Clinical Trust and Privacy:** Moving beyond black-box implementations, this thesis establishes a dual-layered technical infrastructure for clinical validation. On the interpretability axis, the study contrasts diffuse pixel-attribution maps (Grad-CAM) with focused, pixel-exact activation mapping frameworks (Med-CAM). On the data governance axis, this work analyzes Flower-based decentralized federated learning (FL) frameworks using weighted parameter aggregation (FedAvg), showing that multi-institutional training can be orchestrated without compromising patient confidentiality.

4.2 Systemic Limitations

Despite compelling algorithmic advancements, several fundamental constraints must be addressed prior to deployment in real-world clinical environments:

- **Retrospective Data and Homogeneity Bias:** The vast majority of validation pipelines in contemporary literature are retrospective, relying on curated public repositories with highly uniform patient demographics and scanning protocols. This lack of data homogeneity introduces severe generalization risks, where models often suffer catastrophic performance drop-offs when exposed to out-of-distribution, real-world clinical data.
- **Iterative Latency vs. Fidelity Trade-offs:** While Denoising Diffusion Probabilistic Models (DDPMs) offer superior training stability and structural diversity compared to GANs, their mathematical structure relies on an iterative Markov chain reverse-sampling path. Executing hundreds of sequential denoising steps over high-resolution 3D medical volumes incurs substantial computational latency and heavy GPU memory footprints, restricting real-time application in emergency clinical triage.
- **Absence of Clinical Validation:** Standard mathematical metrics used to evaluate image quality—such as FID, PSNR, and SSIM—do not always correlate with diagnostic accuracy. Generative models may achieve excellent statistical fidelity scores while introducing hallucinated structural artifacts or erasing minute, high-frequency pathologies (such as early-stage tuberculosis infiltrates or small cutaneous lesions), which can mislead human practitioners.

4.3 Future Research Horizons

To advance beyond the computational and translational boundaries identified in this study, future research trajectories should focus on two core horizons:

4.3.1 High-Speed Hybrid Generative Frameworks

Future engineering efforts must target the convergence of adversarial speed and diffusion-based structural fidelity. Research should explore unified architectures that embed low-step diffusion processes directly within GAN latent spaces, or leverage Consistency Models and score-based ODE solvers to bypass traditional Markovian bottlenecks. Successfully minimizing the reverse sampling path to 1–5 evaluation steps will reduce inference latency to sub-second thresholds, unlocking the ability to synthesize and reconstruct high-resolution 3D anatomical volumes in real-time.

4.3.2 Transition toward Multimodal Generalist Medical AI (GMAI)

The paradigm of task-specific specialist models must evolve toward unified Medical Foundation Models (MFMs) pre-trained on massive, multimodal health datasets. By utilizing advanced cross-modal tokenization, these generalist frameworks will synthesize diverse patient data streams simultaneously—blending 3D radiology volumes, clinical text reports, lab biomarkers, and genomic sequences. Fine-tuning these foundation architectures via parameter-efficient techniques (such as LoRA) will allow a single model to perform zero-shot or few-shot clinical tasks, providing a scalable foundation for comprehensive precision medicine.

REFERENCES

- [1] J. Joby, R. J. Raju, R. Job, S. M. Thomas, and K. S. Arun. PneumoGAN: A GAN based Model for Pneumonia Detection. In *Proceedings of 2021 2nd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS)*, pages 102–106. IEEE, 2021.
- [2] N. Sasipriyaa, R. Ravikumar, C. Sharmila, E. Gothai, B. Krishnakumar, and R. K. Nithya Sri. Enhancing COVID-19 and Pneumonia Detection with a Hybrid ACGAN-CNN Model. In *Proceedings of 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 2024.
- [3] Z. Benhaili, Y. Abouqora, Y. Balouki, and L. Moumoun. Enhancing Respiratory Diseases Detection Using GAN Augmentation and Transfer Learning. In *Proceedings of Colloquium in Information Science and Technology (CIST)*, pages 221–227. IEEE, 2023.
- [4] S. Saravanakumar, T. Saravanan, and N. Thillaiarasu. A Deep Learning-based Semi-Supervised GAN to detect Alzheimer’s illness efficiently. In *Proceedings of 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pages 1840–1843. IEEE, 2022.
- [5] D. Kundu et al. Federated Deep Learning for Monkeypox Disease Detection on GAN-Augmented Dataset. *IEEE Access*, 12:32819–32829, 2024.
- [6] J. Manikandan, M. Sidharth Raj, R. Yogeshkumar, and C. Sasi Varma. Glaucoma Disease Detection Using Hybrid Deep Learning Model. In *Proceedings of 8th IEEE International Conference on Science, Technology, Engineering and Mathematics (ICONSTEM)*. IEEE, 2023.
- [7] C. Dasanayakaa and M. B. Dissanayake. Deep Learning Methods for Screening Pulmonary Tuberculosis Using Chest X-rays. *International Journal of Biomedical Imaging*, 2020.
- [8] Shahariar Hossain Mahir, Md Tanjum An Tashrif, Aysha Siddika Shathe, Md Ahsan Karim, Anichur Rahman, Dipanjali Kundu, Tanoy Debnath, and Ghulam Muhammad. PriFL-XAI: Hybrid privacy-preserving federated learning models for monkeypox detection through GAN augmentation and explainable AI. *Biomedical Signal Processing and Control*, 112:108426, 2026.
- [9] S. R. Arumugam, p. Sheela Gowr, G. Ponnaian, K. Maharajan, S. G. Karuppasamy, and D. Muralitharan. Synergizing Generative Adversarial Networks and Pseudo-Labeling for Improved Knee Osteoarthritis Detection. In *Proceedings of 2nd International Conference on Sustainable Computing and Smart Systems (ICSS)*, pages 1022–1030. IEEE, 2024.
- [10] T. Mukherjee, S. Sharma, and K. Suganthi. Alzheimer Detection Using Deep Convolutional GAN. In *Proceedings of the IEEE Madras Section International Conference 2021 (MASCON)*. IEEE, 2021.
- [11] R. Ravikumar, N. Sasipriyaa, T. Thilagaraj, R. Hareesh Raj, A. Abishek, and G. Gokula Kannan. Design and Implementation of Alzheimer’s Disease Detection using cGAN and CNN. In *Proceedings of 2023 International Conference on Computer Communication and Informatics (ICCCI)*. IEEE, 2023.

- [12] U. S. Sekhar, N. Vyas, V. Dutt, and A. Kumar. Multimodal Neuroimaging Data in Early Detection of Alzheimer's Disease: Exploring the Role of Ensemble Models and GAN Algorithm. In *Proceedings of the International Conference on Circuit Power and Computing Technologies (ICCPCT)*, pages 1664–1669. IEEE, 2023.
- [13] S. K. H. Bukhari and L. Fahad. Lung Disease Detection using Deep Learning. In *Proceedings of 2022 17th International Conference on Emerging Technologies (ICET)*, pages 154–159. IEEE, 2022.
- [14] O. Polat. Detection of pediatric pneumonia from X-Ray images using ResNet50 and GAL networks. In *Proceedings of 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*. IEEE, 2021.
- [15] T. Deshpande, K. Chavan, p. Gandhi, and R. Mangrulkar. Neurodegenerative Disease Detection using Deep Convolutional GANs and CNN. In *Proceedings of 2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*. IEEE, 2023.
- [16] D. Joon, R. Kumar Shu, M. Gupta, and A. J. Obaid. DeGAN and Efficient NetB3 Based Analysis and Detection of Alzheimer Detection Using MRI Images. In *Proceedings of 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS)*. IEEE, 2024.
- [17] S. Jin, p. Zou, Y. Han, and J. Jiang. Unsupervised detection of individual atrophy in Alzheimer's disease. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, pages 2647–2650. IEEE, 2021.
- [18] D. Srivastav, A. Bajpai, and p. Srivastava. Improved classification for pneumonia detection using transfer learning with GAN based synthetic image augmentation. In *Proceedings of the Confluence 2021: 11th International Conference on Cloud Computing, Data Science and Engineering*, pages 433–437. IEEE, 2021.
- [19] S. Vashisht, S. Lamba, B. Sharma, and A. Sharma. Pneumonia Detection and Classification with CNN-GAN Augmentation Using Fine Tuning. In *Proceedings of 2023 International Conference on Innovations in Engineering and Technology (ICIET)*. IEEE, 2023.
- [20] H. S. Gan, M. H. Ramlee, B. A. S. Al-Rimy, Y. S. Lee, and p. Akkaraekthalin. Hierarchical Knee Image Synthesis Framework for Generative Adversarial Network: Data From the Osteoarthritis Initiative. *IEEE Access*, 10:55051–55061, 2022.
- [21] Y. Cao et al. Robustly detecting mpox and non-mpox using a deep learning framework based on image inpainting. *Scientific Reports*, 15(1), 2025.
- [22] V. Maheskumar and R. Vijayarajeswari. Enhanced detection of Mpox using federated learning with hybrid ResNet-ViT and adaptive attention mechanisms. *Scientific Reports*, 15(1), 2025.
- [23] J. Amjad, M. Z. Sajid, A. Amjad, M. F. Hamid, A. Youssef, and M. I. Sharif. Knee Osteoarthritis Detection and Classification Using Autoencoders and Extreme Learning Machines. *AI (Switzerland)*, 6(7), 2025.
- [24] O. T. Nascimento, J. M. de Seixas, and A. Trajman. Synthetic chest X-ray data generation for tuberculosis infection detection using generative adversarial networks. *Neural Computing and Applications*, 37(22):18151–18171, 2025.
- [25] p. Manasa, V. Veeramalla, D. Vanusha, D. Vathana, and D. S. Robinsha. Leveraging HPC with GAN-CNN Integration for Enhanced Tuberculosis Detection using Synthetic Image

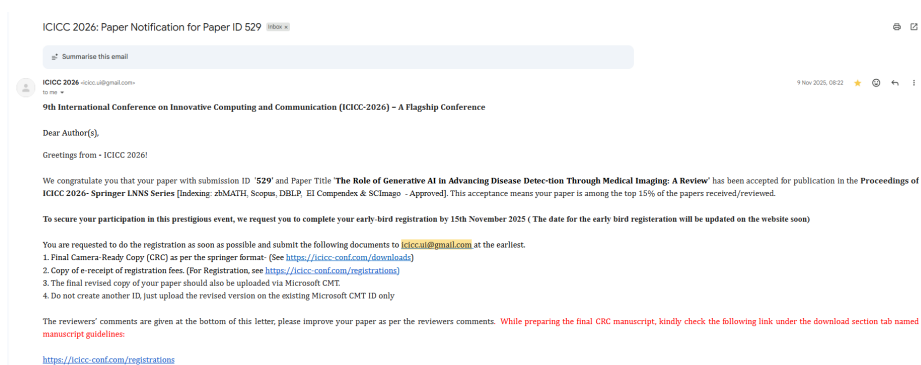
- Generation. In *Proceedings of 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS)*, pages 1311–1318. IEEE, 2024.
- [26] S. Oraby, A. Emran, B. El-Saghir, and S. Mohsen. Hybrid of DSR-GAN and CNN for Alzheimer disease detection based on MRI images. *Scientific Reports*, 15(1), 2025.
 - [27] S. Vashisht, B. Sharma, V. Singh, and S. Kukreti. Binary Classification of Pneumonia Disease Using SVM-GAN Augmentation. In *Proceedings of 2024 3rd International Conference for Innovation in Technology (INOCON)*. IEEE, 2024.
 - [28] Anonymous. Med-CAM: Classifier Activation Matching for Evidence-Based Medical Decision Making. *arXiv preprint arXiv:2604.13695*, 2026.
 - [29] S. H. Mahir et al. Stylegan2-ada over flower federated learning environments for dermatological diagnostics. *Medical Image Analysis*, 2025.
 - [30] Brendan McMahan et al. Decentralized weight aggregation using federated averaging (FedAvg) for multi-institutional diagnostics. *Journal of Distributed Machine Learning*, 2017.
 - [31] U.S. Food and Drug Administration. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>, 2025. Accessed: May 2025.
 - [32] Anonymous. The Total Product Life Cycle Approach to AI-Enabled Medical Devices: Regulatory Guidelines and PCCP Policies. *Annual Review of Biomedical Engineering*, 2025.
 - [33] Anonymous. China’s AIMD regulatory approvals: A risk-proportionate trend analysis. *JMIR Medical Informatics*, 14(1):e85538, 2026.
 - [34] International Electrotechnical Commission. IEC 62304: Medical Device Software - Software Life Cycle Processes. Technical report, IEC, 2025.
 - [35] International Organization for Standardization. ISO 13485: Medical Devices - Quality Management Systems - Requirements for Regulatory Purposes. Technical report, ISO, 2025.

APPENDIX I

I.1 List of Publications

1. **Sanskar Jain**, Priyanka Arora, and Dr. Sonika Dahiya (2025). “The Role of Generative AI in Advancing Disease Detection Through Medical Imaging: A Review.” *Department of Software Engineering, Delhi Technological University*. [Accepted- ICICC 2026, Springer LNNS Series]

I.2 Paper Acceptance Proofs



I.3 Conference Paper Registration Receipts



I.4 Indexing of Conference Proof

09/05/2026, 16:50

ICICC | International Conference on Innovative Computing and Communication



ICICC
INTERNATIONAL CONFERENCE ON INNOVATIVE
COMPUTING AND COMMUNICATION

10th INTERNATIONAL CONFERENCE ON INNOVATIVE COMPUTING AND COMMUNICATION (ICICC-2027)



ORGANISED BY:
SHAHEED SUKHDEV
COLLEGE OF BUSINESS
STUDIES, UNIVERSITY
OF DELHI, NEW DELHI
IN ASSOCIATION WITH
UNIVERSITY OF
VALLADOLID SPAIN
5th-6th FEBRUARY
2027

ICICC 2026

The ninth edition of the International Conference on Innovative Computing and Communication (ICICC-2026) was successfully organized at Shaheed Sukhdev College of Business Studies, New Delhi, India, in association with the National Institute of Technology Patna, India, and University of Valladolid, Spain, on 06-07 February 2026.

The conference witnessed an overwhelming global response, receiving more than 2,500 research paper submissions from over 9,000 authors worldwide, reflecting the growing international interest in emerging computing and communication technologies. After a rigorous peer-review process, 475 high-quality papers were accepted, resulting in a competitive acceptance rate of approximately 18%.

All accepted papers were published in Springer's Lecture Notes in Networks and Systems, a Scopus-indexed publication series. The conference attracted over 1,000 participants including researchers, academicians, keynote speakers, industry experts, and delegates from around 40

Important Dates

Paper Submission

Deadline:

30th August 2026

Notification of first review:

25th October 2026

Submission of revised

<https://icicc-conf.com/icicc26>

1/3

I.5 Conference Certificates



APPENDIX II SOURCE CODE

II.1 PyTorch Implementation of a Conditional GAN (cGAN) Generator

```

1 import torch
2 import torch.nn as nn
3
4 class ConditionalGenerator(nn.Module):
5     def __init__(self, latent_dim, num_classes, img_channels):
6         super(ConditionalGenerator, self).__init__()
7         self.label_emb = nn.Embedding(num_classes, num_classes)
8
9         self.model = nn.Sequential(
10             nn.Linear(latent_dim + num_classes, 128),
11             nn.LeakyReLU(0.2, inplace=True),
12             nn.Linear(128, 256),
13             nn.BatchNorm1d(256, 0.8),
14             nn.LeakyReLU(0.2, inplace=True),
15             nn.Linear(256, 512),
16             nn.BatchNorm1d(512, 0.8),
17             nn.LeakyReLU(0.2, inplace=True),
18             nn.Linear(512, 1024),
19             nn.BatchNorm1d(1024, 0.8),
20             nn.LeakyReLU(0.2, inplace=True),
21             nn.Linear(1024, img_channels * 128 * 128),
22             nn.Tanh()
23         )
24
25     def forward(self, noise, labels):
26         # Concatenate latent vector and label embedding
27         gen_input = torch.cat((self.label_emb(labels), noise), -1)
28         img = self.model(gen_input)
29         img = img.view(img.size(0), 1, 128, 128)
30         return img

```

II.2 PyTorch Implementation of an ACGAN Discriminator

```

1 class ACGANDiscriminator(nn.Module):
2     def __init__(self, num_classes, img_channels):
3         super(ACGANDiscriminator, self).__init__()
4
5         self.conv_blocks = nn.Sequential(
6             nn.Conv2d(img_channels, 16, 3, 2, 1),
7             nn.LeakyReLU(0.2, inplace=True),
8             nn.Dropout(0.25),
9             nn.Conv2d(16, 32, 3, 2, 1),
10            nn.LeakyReLU(0.2, inplace=True),
11            nn.Dropout(0.25),

```

```

12         nn.BatchNorm2d(32, 0.8)
13     )
14
15     # Classification output
16     self.adv_layer = nn.Sequential(nn.Linear(32 * 32 * 32, 1),
17                                     nn.Sigmoid())
18     self.aux_layer = nn.Sequential(nn.Linear(32 * 32 * 32,
19                                             num_classes), nn.Softmax(dim=1))
20
21     def forward(self, img):
22         out = self.conv_blocks(img)
23         out = out.view(out.shape, -1)
24         validity = self.adv_layer(out)
25         class_label = self.aux_layer(out)
26         return validity, class_label

```

II.3 MS-SSIM Metric Calculation for Synthetic Medical Scan Quality

```

1 import torch
2 import torch.nn.functional as F
3
4 def ssim(img1, img2, window_size=11):
5     # Generates standard Gaussian window for local structural
6     evaluation
7     kernel = torch.ones((1, 1, window_size, window_size)) /
8         (window_size**2)
9     mu1 = F.conv2d(img1, kernel, padding=window_size//2)
10    mu2 = F.conv2d(img2, kernel, padding=window_size//2)
11
12    mu1_sq, mu2_sq = mu1.pow(2), mu2.pow(2)
13    mu1_mu2 = mu1 * mu2
14
15    sigma1_sq = F.conv2d(img1 * img1, kernel,
16                          padding=window_size//2) - mu1_sq
17    sigma2_sq = F.conv2d(img2 * img2, kernel,
18                          padding=window_size//2) - mu2_sq
19    sigma12 = F.conv2d(img1 * img2, kernel, padding=window_size//2)
20    - mu1_mu2
21
22    c1, c2 = 0.01**2, 0.03**2
23    ssim_map = ((2 * mu1_mu2 + c1) * (2 * sigma12 + c2)) / ((mu1_sq
24    + mu2_sq + c1) * (sigma1_sq + sigma2_sq + c2))
25    return ssim_map.mean()

```

II.4 Peak Signal-to-Noise Ratio (PSNR) Calculation for Brain MRIs

```

1 import numpy as np
2
3 def compute_psnr(original_vol, synthetic_vol):
4     mse = np.mean((original_vol - synthetic_vol) ** 2)
5     if mse == 0:

```

```

6         return float('inf')
7     max_pixel = 255.0
8     psnr = 20 * np.log10(max_pixel / np.sqrt(mse))
9     return psnr

```

II.5 Local Med-CAM Feature Activation Matching Implementation

```

1 import torch
2 import torch.nn as nn
3
4 class MedCAMMatch(nn.Module):
5     def __init__(self, classifier_layer, num_channels=512):
6         super(MedCAMMatch, self).__init__()
7         self.classifier_layer = classifier_layer
8         # Lightweight U-Net segmentation branch to produce
9         pixel-exact activation mask
10        self.seg_mask = nn.Sequential(
11            nn.Conv2d(num_channels, 256, 3, padding=1),
12            nn.BatchNorm2d(256),
13            nn.ReLU(inplace=True),
14            nn.Conv2d(256, 128, 3, padding=1),
15            nn.BatchNorm2d(128),
16            nn.ReLU(inplace=True),
17            nn.Conv2d(128, 1, 1),
18            nn.Sigmoid()
19        )
20
21    def forward(self, feature_map):
22        # Extract activations directly from targeted intermediate
23        convolutional layer
24        activations = self.classifier_layer(feature_map)
25        mask = self.seg_mask(activations)
26        return mask
27
28 class PreclinicalPreprocessing:
29     pass

```

II.6 Preclinical N4ITK Bias-Field Correction Wrapper

```

1 import SimpleITK as sitk
2
3 def apply_n4_bias_correction(mri_path, output_path):
4     # Standard medical preprocessing routine to resolve magnetic
5     intensity non-uniformity
6     input_image = sitk.ReadImage(mri_path, sitk.sitkFloat32)
7     mask_image = sitk.OtsuThreshold(input_image, 0, 1)
8
9     corrector = sitk.N4BiasFieldCorrectionImageFilter()
10    # Fine-tuning parameters for maximum anatomical detail
11    preservation
12    corrector.SetMaximumNumberOfIterations()
13    corrector.SetConvergenceThreshold(1e-6)

```

```

12     output_image = corrector.Execute(input_image, mask_image)
13     sitk.WriteImage(output_image, output_path)
14

```

II.7 Federated Averaging (FedAvg) Simulation for Distributed Medical AI

```

1  import copy
2
3  def federated_averaging(global_model, client_models,
4      client_data_sizes):
5      # Computes weighted parameter updates across local hospital
6      nodes
7      total_samples = sum(client_data_sizes)
8      global_dict = global_model.state_dict()
9
10     for key in global_dict.keys():
11         fractional_update = torch.zeros_like(global_dict[key],
12             dtype=torch.float32)
13         for i, client in enumerate(client_models):
14             weight = client_data_sizes[i] / total_samples
15             fractional_update += weight *
16                 client.state_dict()[key].float()
17         global_dict[key] =
18             fractional_update.to(global_dict[key].dtype)
19
20     global_model.load_state_dict(global_dict)
21     return global_model

```