

# Novelty Assessment Report

**Paper:** CheXGenBench: A Unified Benchmark for Fidelity, Privacy and Utility of Synthetic Chest Radiographs

**PDF URL:** <https://openreview.net/pdf?id=u1OWn3ayY1>

**Venue:** ICLR 2026 Conference Submission

**Year:** 2026

**Report Generated:** 2026-01-01

## Abstract

Structured benchmarks have advanced text-conditional image generation for real-world imagery, however, no such benchmark exists for synthetic radiograph generation. Despite being a highly active area of research, existing studies continue adopting inconsistent evaluation protocols and lack a unified assessment of the three most critical criteria: generative fidelity, privacy risk, and downstream utility. To address these limitations, we introduce CheXGenBench, the first unified evaluation framework for synthetic chest radiograph generation that simultaneously assesses fidelity, privacy risks, and clinical utility across frontier text-to-image (T2I) generative models. Our evaluation protocol, comprising over 20 quantitative metrics, covers 11 leading T2I architectures with plug-and-play integration for newer models. Through a rigorous and fair evaluation protocol, we establish a new SoTA in synthetic chest X-ray generation. Furthermore, our results uncover several critical limitations in the applicability of current generative models, which include (1) even SoTA models struggle with long-tailed medical distributions, (2) models pose high privacy risks regardless of fidelity quality, and (3) synthetic data offers limited utility for downstream multimodal tasks. Drawing from these results, we propose concrete research directions to advance the field. Finally, we curate and release SynthCheX-75K, a high-quality synthetic dataset comprising 75K radiographs generated by our top-performing model (Sana 0.6B). The fine-tuned models and the SynthCheX-75K dataset would be released after acceptance, while the anonymised code is available at <https://anonymous.4open.science/r/CheXGenBench-52F0/README.md>

### Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: [mingzhang23@m.fudan.edu.cn](mailto:mingzhang23@m.fudan.edu.cn)

## Core Task Landscape

This paper addresses: **Synthetic Chest Radiograph Generation Evaluation**

A total of **50 papers** were analyzed and organized into a taxonomy with **17 categories**.

### Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Generative Model Architectures and Training Approaches**
- **Controlled and Conditional Synthesis**
- **Evaluation Methodologies and Benchmarking Frameworks**
- **Application-Driven Synthesis and Data Augmentation**
- **Cross-Domain and Multi-Modal Synthesis**
- **Methodological Reviews and Comparative Studies**

### Complete Taxonomy Tree

- Synthetic Chest Radiograph Generation Evaluation Survey Taxonomy
- Generative Model Architectures and Training Approaches
  - GAN-Based Synthesis Methods
  - Standard GAN Architectures (10 papers)
    - [1] Synthetic chest X-ray data generation for tuberculosis infection detection using generative adversarial networks (O. Nascimento, 2025) [View paper](#)
    - [5] Medical Image Synthesis Using DCGAN for Chest X-Ray Images (M C Sai Akhil, 2024) [View paper](#)
    - [6] Method for generating synthetic data combining chest radiography images with tabular clinical information using dual generative models (Kikuchi Tomohiro, 2023) [View paper](#)
    - [18] Synthesizing chest X-ray pathology for training deep convolutional neural networks (H. Salehinejad, 2018) [View paper](#)
    - [21] Generative X-Ray Synthesis for Enhanced Disease Prediction in Medical Imaging (Smitha Rai, 2024) [View paper](#)
    - [23] Chest x-ray image synthesis using deep convolutional gans (Nada Ahmed Alhamdi, 2024) [View paper](#)
    - [24] CoVaD-GAN: An efficient Data Augmentation technique for COVID CXR Image Classification (Archana Dash, 2023) [View paper](#)
    - [26] Evaluation of deep convolutional generative adversarial networks for data augmentation of chest x-ray images (S Kora Venu, 2020) [View paper](#)
    - [34] Generative adversarial networks for the synthesis of chest x-ray images (Mai Feng Ng, 2023) [View paper](#)
    - [46] Cross-Modality Synthetic Data Augmentation using GANs: Enhancing Brain MRI and Chest X-ray Classification (Kunaal Dhawan, 2024) [View paper](#)
  - Progressive Growing and Style-Based GANs (5 papers)
    - [4] Progressive GAN Framework for Realistic Chest X-Ray Synthesis and Data Augmentation (Kiran Kumar Maguluri, 2025) [View paper](#)
    - [10] Disease-Aware Chest X-Ray Style GAN Image Generation and CatBoost Gradient Boosted Trees (Andi Besse Firdausiah Mansur, 2024) [View paper](#)
    - [16] Image Turing test and its applications on synthetic chest radiographs by using the progressive growing generative adversarial network (Miso Jang, 2023) [View paper](#)

- [19] Evaluating the clinical realism of synthetic chest x-rays generated using progressively growing gans (Rubin, 2021) [View paper](#)
    - [39] Advancing Medical Imaging: A Comparative Exploration of Generative Adversarial Networks for Chest X-ray Synthesis (R Keerthi, 2024) [View paper](#)
  - Domain Adaptation and Cycle-Consistent GANs (2 papers)
    - [41] CycleGAN generated pneumonia chest x-ray images: Evaluation with vision transformer (Gerardo Lugo-Torres, 2024) [View paper](#)
    - [48] Generative domain adaptation for chest X-ray image analysis (Baocai Yin, 2021) [View paper](#)
  - Diffusion-Based Synthesis Methods (6 papers)
  - [11] Denoising diffusion probabilistic models for addressing data limitations in chest X-ray classification (Evi M.C. Huijben, 2024) [View paper](#)
  - [25] Beware of diffusion models for synthesizing medical images—a comparison with GANs in terms of memorizing brain MRI and chest x-ray images (Muhammad Usman Akbar, 2023) [View paper](#)
  - [29] Adapted generative latent diffusion models for accurate pathological analysis in chest X-ray images. (D. I. Morás, 2024) [View paper](#)
  - [36] Evaluating and Improving the Effectiveness of Synthetic Chest X-Rays for Medical Image Analysis (Prakash, 2024) [View paper](#)
  - [45] Exploring Foundation Models for Synthetic Medical Imaging: A Study on Chest X-Rays and Fine-Tuning Techniques (Davide Clode da Silva, 2024) [View paper](#)
  - [50] Evaluating the feasibility of using Generative Models to generate Chest X-Ray Data (Malik, 2023) [View paper](#)
  - VAE-Based and Hybrid Generative Approaches (1 papers)
  - [43] Synthetic Chest X-Ray Augmentation via Generative Variational Autoencoding for Pneumonia Detection (K. Hartomo, 2025) [View paper](#)
- Controlled and Conditional Synthesis
  - Text-Conditioned Generation (4 papers)
  - [8] Roentgen: vision-language foundation model for chest x-ray generation (Chambon, 2022) [View paper](#)
  - [15] Synthetic chest X-ray images from text prompts (Daniel Truhn, 2025) [View paper](#)
  - [20] A Generative Foundation Model for Chest Radiography (JI YUANFENG, 2025) [View paper](#)
  - [40] Recognizing Pneumonia in Real-World Chest X-rays with a Classifier Trained with Images Synthetically Generated by Nano Banana (Jiachuan Peng, 2025) [View paper](#)
  - Mask-Based and Attribute-Conditioned Synthesis (2 papers)
  - [9] Improving Performance, Robustness, and Fairness of Radiographic AI Models with Finely-Controllable Synthetic Data (Bluethgen, 2025) [View paper](#)
  - [32] Feature Controlled Synthetic Medical Image Generation using Conditional Generative Adversarial Network (A. Robert Singh, 2025) [View paper](#)
  - Disease-Aware and Disentangled Synthesis (3 papers)
  - [2] Disentangled Contrastive Learning From Synthetic Matching Pairs for Targeted Chest X-Ray Generation (Euyoung Kim, 2025) [View paper](#)
  - [12] Multiscale attention generative adversarial networks for lesion synthesis in chest X-ray images. (Md Shariful Alam, 2025) [View paper](#)
  - [28] Prototype-Driven Class-Conditional Synthesis for High-Quality Chest X-ray Image Generation. (Bowen Guo, 2025) [View paper](#)
- Evaluation Methodologies and Benchmarking Frameworks
  - Unified Benchmarking and Multi-Metric Evaluation ★ (3 papers)
  - [0] CheXGenBench: A Unified Benchmark for Fidelity, Privacy and Utility of Synthetic Chest Radiographs (Anon et al., 2026) [View paper](#)
  - [17] Introducing SPINE: A Holistic Approach to Synthetic Pulmonary Imaging Evaluation Through End-to-End Data and Model Management (Nikolaos Ntampakis, 2024) [View paper](#)
  - [30] You Don't Have to Be Perfect to Be Amazing: Unveil the Utility of Synthetic Images (Xiaodan Xing, 2023) [View paper](#)
  - Fidelity and Clinical Realism Assessment (2 papers)
  - [7] Evaluating diagnostic content of AI-generated chest radiography: A multi-center visual Turing test (Youho Myong, 2023) [View paper](#)
  - [37] High-Resolution Chest Synthesis Data Generation for Medical Image Analysis: An Advanced GAN Approach with Conditional Image Synthesis (iijí½ííí, 2025) [View paper](#)
  - Privacy Risk and Memorization Analysis (2 papers)
  - [31] The Devil is in the Prompts: De-Identification Traces Enhance Memorization Risks in Synthetic Chest X-Ray Generation (Dutt, 2025) [View paper](#)
  - [42] Sensitivity, Specificity, and Consistency: A Tripartite Evaluation of Privacy Filters for Synthetic Data Generation (Ziller, 2025) [View paper](#)
  - Downstream Task Utility Evaluation (1 papers)
  - [3] A Critical Assessment of Generative Models for Synthetic Data Augmentation on Limited Pneumonia X-ray Data (Daniel Schaudt, 2023) [View paper](#)
- Application-Driven Synthesis and Data Augmentation
  - COVID-19 Detection and Augmentation (3 papers)
  - [13] A preliminary analysis of AI based smartphone application for diagnosis of COVID-19 using chest X-ray images (Aravind Krishnaswamy Rangarajan, 2021) [View paper](#)
  - [14] Generation of synthetic chest X-ray images and detection of COVID-19: A deep learning based approach (Yash Karbhari, 2021) [View paper](#)
  - [27] Improving Covid-19 cxr detection with synthetic data augmentation (Daniel Schaudt, 2021) [View paper](#)
  - Pneumonia Detection and Classification (2 papers)
  - [35] Chest radiography few-shot image synthesis for automated pathology screening applications (Martim Quintas E Sousa, 2021) [View paper](#)
  - [38] Robust Approaches for Deep Learning for Chest X-ray Assessment Using Synthetically Generated Illness (E. Hughes, 2025) [View paper](#)
  - Multi-Disease Classification and Rare Condition Augmentation (1 papers)
  - [49] Nodule Detection and Generation on Chest X-Rays: NODE21 Challenge. (Sogancioglu, 2024) [View paper](#)

- Cross-Domain and Multi-Modal Synthesis (3 papers)
  - [33] Machine-Learning Based Detection of Coronary Artery Calcification Using Synthetic Chest X-Rays (Dylan Saeed, 2025) [View paper](#)
  - [44] Development of a deep neural network for generating synthetic dual-energy chest x-ray images with single x-ray exposure. (Dongâ€Hoon Lee, 2020) [View paper](#)
  - [47] Image projective transformation rectification with synthetic data for smartphone-captured chest X-ray photos classification. (Chak Fong Chong, 2023) [View paper](#)
- Methodological Reviews and Comparative Studies (1 papers)
  - [22] Synthetic data in radiological imaging: current state and future outlook (Sizikova, 2024) [View paper](#)

## Narrative

Core task: synthetic chest radiograph generation evaluation. The field has matured into a structured landscape organized around six main branches. Generative Model Architectures and Training Approaches explore foundational techniques—ranging from early DCGAN Medical Synthesis[5] and Progressive GAN Framework[4] to more recent diffusion-based methods like Denoising Diffusion Probabilistic[11] and Adapted Latent Diffusion[29]—that establish how synthetic chest X-rays are produced. Controlled and Conditional Synthesis focuses on fine-grained manipulation, enabling disease-specific or attribute-driven generation through works such as Disease Aware StyleGAN[10] and Feature Controlled Synthesis[32]. Evaluation Methodologies and Benchmarking Frameworks address the critical question of how to assess realism and clinical utility, with studies like Visual Turing Test[7] and Clinical Realism Evaluation[19] proposing human-expert protocols alongside automated metrics. Application-Driven Synthesis and Data Augmentation examines practical deployment for tasks like COVID Detection Synthesis[14] and Enhanced Disease Prediction[21], while Cross-Domain and Multi-Modal Synthesis investigates translation between imaging modalities or text-to-image generation as in Text to CXR[15]. Finally, Methodological Reviews and Comparative Studies, including Current State Outlook[22] and Comparative Exploration GANs[39], synthesize progress and identify open challenges across these threads.

Several active lines reveal ongoing tensions between model sophistication, controllability, and rigorous validation. On one hand, generative architectures have grown increasingly powerful, yet concerns about clinical fidelity and potential biases persist, as highlighted by Beware Diffusion Models[25] and Critical Assessment Pneumonia[3]. On the other hand, application-driven efforts demonstrate that synthetic data can improve downstream classifiers when carefully validated, though the gap between perceptual quality and diagnostic utility remains debated. CheXGenBench[0] situates itself squarely within the Unified Benchmarking and Multi-Metric Evaluation cluster, proposing a comprehensive framework that bridges multiple evaluation dimensions—visual realism, diagnostic consistency, and downstream task performance. This positions it alongside SPINE[17] and Utility Synthetic Images[30], which similarly advocate for holistic assessment protocols rather than isolated metrics, addressing the field's need for standardized benchmarks that can guide both model development and clinical adoption.

## Related Works in Same Category

The following **2 sibling papers** share the same taxonomy leaf node with the original paper:

### 1. Introducing SPINE: A Holistic Approach to Synthetic Pulmonary Imaging Evaluation Through End-to-End Data and Model Management

**Authors:** Nikolaos Ntampakis, Vasileios Argyriou, Konstantinos Diamantaras, Konstantinos Goulianas, Konstantinos I. Diamantaras, et al. (8 authors total) | **Year/Venue:** 2024 • IEEE Open Journal of Engineering in Medicine and Biology | **URL:** [View paper](#)

#### Abstract

In the evolving field of medical imaging and machine learning (ML), this paper introduces a novel framework for evaluating synthetic pulmonary imaging aiming to assess synthetic data quality and applicability. Our study concentrates on synthetic X-ray chest images, crucial for diagnosing respiratory diseases. We employ SPINE (Synthetic Pulmonary Imaging Evaluation) framework, a threefold synthetic images evaluation method including expert domain assessment, statistical data analysis and adversar...

#### Relationship Analysis

Both papers belong to the unified benchmarking and multi-metric evaluation category, establishing comprehensive frameworks for assessing synthetic chest radiograph generation across multiple dimensions. They overlap in evaluating fidelity, utility, and clinical applicability of synthetic images using diverse metrics. However, the original paper (CheXGenBench) focuses on benchmarking 11 frontier T2I models with 20+ standardized metrics including privacy/re-identification risks and downstream task utility, while the candidate paper (SPINE) introduces a threefold evaluation approach (expert domain assessment, statistical analysis, adversarial evaluation) with emphasis on post-market clinical validation and GAN-based generation rather than comparing multiple modern architectures.

### 2. You Don't Have to Be Perfect to Be Amazing: Unveil the Utility of Synthetic Images

**Authors:** Xiaodan Xing, Federico Felder, Yang Nan, Giorgos Papanastasiou, Simon Walsh, et al. (6 authors total) | **Year/Venue:** 2023 • International Conference on Medical Image Computing and Computer-Assisted Intervention | **URL:** [View paper](#)

#### Abstract

Synthetic images generated from deep generative models have the potential to address data scarcity and data privacy issues. The selection of synthesis models is mostly based on image quality measurements, and most researchers favor synthetic images that produce realistic images, i.e., images with good fidelity scores, such as low Fr'echet Inception Distance (FID) and high Peak Signal-To-Noise Ratio (PSNR). However, the quality of synthetic images is not limited to fidelity, and a wide spectrum ...

#### Relationship Analysis

Both papers belong to the unified benchmarking and multi-metric evaluation category, establishing comprehensive frameworks for assessing synthetic chest radiographs across fidelity, privacy, and utility dimensions. They overlap in evaluating multiple generative models using diverse metrics (FID, privacy measures, downstream task performance) to provide holistic assessments of synthetic image quality. However, the original paper (CheXGenBench) introduces a standardized benchmark with 20+ metrics across 11 T2I models including recent architectures, enhanced captions, and pathology-conditional analysis, while the candidate paper focuses on analyzing trade-offs between fidelity, variety, and privacy across 100k+ images, demonstrating that high utility does not require perfect fidelity or variety.

## Contributions Analysis

**Overall novelty summary.** The paper introduces CheXGenBench, a unified evaluation framework for synthetic chest radiograph generation that simultaneously assesses fidelity, privacy risks, and clinical utility across 11 text-to-image architectures. It resides in the 'Unified Benchmarking and Multi-Metric Evaluation' leaf, which contains only three papers total, indicating a relatively sparse research direction. The sibling papers (SPINE and Utility Synthetic Images) similarly advocate for holistic assessment protocols, suggesting this leaf represents an emerging consensus around comprehensive benchmarking rather than isolated metrics.

The taxonomy reveals that while generative architectures (GANs, diffusion models) and application-driven synthesis are well-populated branches, the evaluation methodologies branch remains comparatively underdeveloped. Neighboring leaves focus on single-aspect assessments: 'Fidelity and Clinical Realism Assessment' examines visual quality via radiologist studies, 'Privacy Risk and Memorization Analysis' addresses data leakage concerns, and 'Downstream Task Utility Evaluation' measures classifier performance. CheXGenBench's multi-faceted approach bridges these fragmented evaluation threads, positioning it at the intersection of previously siloed assessment dimensions.

Among 30 candidates examined, none clearly refute the three core contributions. The unified framework contribution examined 10 candidates with zero refutable matches, suggesting limited prior work proposing simultaneous fidelity-privacy-utility benchmarks at this scale. The state-of-the-art model and evaluation protocol contribution similarly found no refutations across 10 candidates, though the search scope cannot confirm exhaustive novelty. The SynthCheX-75K dataset contribution also showed zero refutations among 10 examined papers, indicating potential novelty in dataset scale or composition within the limited search window.

Based on the top-30 semantic matches and taxonomy structure, the work appears to occupy a genuine gap in comprehensive benchmarking for medical image synthesis. However, the limited search scope means adjacent evaluation frameworks in broader computer vision or alternative medical imaging domains may not have been captured. The analysis covers synthetic chest X-ray generation specifically but does not extend to evaluation methodologies in other radiology subfields or general-purpose image synthesis benchmarks.

---

This paper presents **3 main contributions**, each analyzed against relevant prior work:

### **Contribution 1: CheXGenBench unified evaluation framework**

**Description:** A comprehensive benchmark framework that evaluates synthetic chest X-ray generation models across three dimensions: generative fidelity and mode coverage, privacy and patient re-identification risks, and downstream clinical utility. The framework includes over 20 quantitative metrics and supports plug-and-play integration of new models.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

#### **1. Scorecard for synthetic medical data evaluation**

URL: [View paper](#)

##### **Brief Assessment**

Scorecard Evaluation[69] proposes a general documentation framework for synthetic medical data across multiple modalities, while CheXGenBench is a specialized benchmark specifically for chest X-ray generation with automated metric computation and plug-and-play model integration. The candidate focuses on standardized reporting templates rather than automated evaluation pipelines.

---

#### **2. Eyes Tell the Truth: GazeVal Highlights Shortcomings of Generative AI in Medical Imaging**

URL: [View paper](#)

##### **Brief Assessment**

GazeVal[73] focuses on eye-tracking-based evaluation of synthetic medical images using radiologist gaze patterns, not a comprehensive benchmark framework assessing fidelity, privacy, and clinical utility with 20+ quantitative metrics.

---

#### **3. Generating high-fidelity synthetic patient data for assessing machine learning healthcare software**

URL: [View paper](#)

##### **Brief Assessment**

High Fidelity Patient Data[68] focuses on generating synthetic patient data for primary care records using probabilistic graphical models, not on evaluating synthetic chest X-ray generation models across fidelity, privacy, and clinical utility dimensions.

---

#### **4. Clinical evaluation of medical image synthesis: a case study in wireless capsule endoscopy**

URL: [View paper](#)

##### **Brief Assessment**

Wireless Capsule Endoscopy[65] focuses on evaluating synthetic wireless capsule endoscopy images for inflammatory bowel disease diagnosis, not chest X-ray generation. The domains, imaging modalities, and clinical applications are fundamentally different.

---

#### **5. Generative Artificial Intelligence in Medical Imaging: Foundations, Progress, and Clinical Translation**

URL: [View paper](#)

##### **Brief Assessment**

Foundations Clinical Translation[67] is a broad review paper on generative AI in medical imaging that proposes a general three-tiered evaluation framework (pixel-level fidelity, feature-level realism, task-level clinical relevance) for medical imaging applications. This differs from CheXGenBench, which is a specific benchmark implementation for chest X-ray generation with concrete metrics, datasets, and model comparisons rather than a conceptual framework proposal.

---

#### **6. Interpretable Similarity of Synthetic Image Utility**

URL: [View paper](#)

##### **Brief Assessment**

Interpretable Similarity Utility[72] focuses on assessing similarity between synthetic and real medical images for clinical decision support systems across multiple modalities (endoscopic, dermoscopic, fundus, x-ray, ultrasound), not on unified benchmarking of chest X-ray generation models across fidelity, privacy, and clinical utility dimensions.

---

#### **7. Generative Adversarial Networks for Synthetic Biomedical Data: Ensuring Data Fidelity and Privacy Preservation**

URL: [View paper](#)

##### **Brief Assessment**

Fidelity Privacy Preservation[70] addresses synthetic biomedical data broadly across multiple modalities (genomic sequences, medical imaging, clinical data), while the original paper specifically focuses on chest X-ray generation with over 20 metrics across 11 T2I models. The candidate's scope and methodology differ substantially from the original's specialized radiograph benchmark.

---

#### **8. SynthVal: A Framework for Validating Synthetic Medical Images**

URL: [View paper](#)

##### **Brief Assessment**

SynthVal[71] focuses on general validation methods for synthetic medical images across modalities, while the original paper presents a specialized benchmark specifically for chest X-ray generation with 20+ metrics across fidelity, privacy, and clinical utility dimensions. The candidate's limited context does not demonstrate prior work on this specific comprehensive chest radiograph evaluation framework.

---

## 9. You Don't Have to Be Perfect to Be Amazing: Unveil the Utility of Synthetic Images

URL: [View paper](#)

### Brief Assessment

Utility Synthetic Images[30] focuses on establishing trade-offs between fidelity, variety, privacy, and utility metrics for synthetic chest X-rays, but does not present a unified benchmark framework with standardized protocols, plug-and-play model integration, or the comprehensive 20+ metric suite that CheXGenBench provides.

---

## 10. Design and development of a systematic validation protocol for synthetic melanoma images for responsible use in medical artificial intelligence

URL: [View paper](#)

### Brief Assessment

Melanoma Validation Protocol[66] focuses on dermoscopic melanoma images using GANs with qualitative expert assessments, while the original paper addresses chest X-ray generation with diffusion models using quantitative metrics across fidelity, privacy, and clinical utility dimensions.

---

## Contribution 2: New state-of-the-art model and evaluation protocol

**Description:** The authors establish new state-of-the-art performance in synthetic chest radiograph generation by evaluating 11 leading text-to-image architectures using standardized training protocols and identifying Sana 0.6B as the top-performing model through their comprehensive benchmark.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

## 1. Cxr-irgen: An integrated vision and language model for the generation of clinically accurate chest x-ray image-report pairs

URL: [View paper](#)

### Brief Assessment

CXR-IRGen[55] focuses on generating paired chest X-ray images and radiology reports using an integrated vision-language model, not on establishing a comprehensive benchmark for evaluating text-to-image models across fidelity, privacy, and utility dimensions as the original paper does.

---

## 2. Chest-diffusion: a light-weight text-to-image model for report-to-cxr generation

URL: [View paper](#)

### Brief Assessment

Chest Diffusion[51] focuses on a lightweight text-to-image model for chest X-ray generation with computational efficiency goals, not on establishing comprehensive benchmarking protocols or evaluating multiple architectures systematically.

---

## 3. Evaluating and Improving the Effectiveness of Synthetic Chest X-Rays for Medical Image Analysis

URL: [View paper](#)

### Brief Assessment

Evaluating Effectiveness[36] focuses on best practices for synthetic chest X-ray generation for specific downstream tasks (classification/segmentation) using latent diffusion models, rather than establishing a comprehensive benchmark framework with standardized evaluation protocols across multiple architectures as claimed in the original paper.

---

## 4. Covid-19 pneumonia chest x-ray pattern synthesis by stable diffusion

URL: [View paper](#)

### Brief Assessment

Stable Diffusion Pneumonia[57] focuses specifically on COVID-19 pneumonia pattern synthesis using prior preservation strategy, not on establishing comprehensive benchmarks or evaluating multiple text-to-image architectures with standardized protocols across diverse chest radiograph pathologies.

---

## 5. Spot the fake lungs: Generating Synthetic Medical Images using Neural Diffusion Models

URL: [View paper](#)

### Brief Assessment

Spot Fake Lungs[58] is an early exploratory work using pre-trained DALLE2 and basic Stable Diffusion training (3165 images), evaluated only through radiologist labeling. It does not establish standardized evaluation protocols or benchmark multiple architectures systematically.

---

## 6. Cxr-clip: Toward large scale chest x-ray language-image pre-training

URL: [View paper](#)

### Brief Assessment

CXR CLIP[53] focuses on vision-language pre-training for chest X-rays using contrastive learning between images and text, not on text-to-image generation or synthetic radiograph creation. The candidate addresses a fundamentally different task (discriminative vision-language modeling) than the original paper (generative text-to-image synthesis).

---

## 7. Synthetic lung x-ray generation through cross-attention and affinity transformation

URL: [View paper](#)

### Brief Assessment

Cross Attention Affinity[56] focuses on semantic mask generation from synthetic lung X-rays using cross-attention mapping, not on establishing state-of-the-art performance through comprehensive benchmarking of text-to-image architectures with standardized evaluation protocols.

---

## 8. Any-to-Any Vision-Language Model for Multimodal X-ray Imaging and Radiological Report Generation

URL: [View paper](#)

## Brief Assessment

Any to Any[54] focuses on multimodal generation (multi-view chest X-rays and clinical reports) rather than establishing a comprehensive benchmark for evaluating text-to-image models. The candidate does not challenge the original paper's claim of establishing new SOTA through systematic evaluation of 11 T2I architectures with standardized protocols.

---

## 9. Generative AI Techniques in Medical Imaging Analysis: A Systematic Review

URL: [View paper](#)

### Brief Assessment

Systematic Review Techniques[52] appears to be a systematic review paper covering generative AI techniques in medical imaging broadly, including text-to-image synthesis for chest X-rays. However, the provided context is too fragmentary to determine if it establishes prior state-of-the-art benchmarks or evaluation protocols that would refute the original paper's claims about being the first unified benchmark for synthetic chest radiograph generation.

---

## 10. A Generative Foundation Model for Chest Radiography

URL: [View paper](#)

### Brief Assessment

Generative Foundation Model[20] focuses on a unified framework for multi-modal synthesis (text, mask, bounding box) and downstream utility demonstrations, rather than establishing a comprehensive benchmark with standardized evaluation protocols across multiple architectures as claimed in the original paper.

---

## Contribution 3: SynthCheX-75K synthetic dataset

**Description:** A curated dataset of 75,000 high-quality synthetic chest radiographs generated using the benchmark's best-performing model. This dataset can serve as a standalone training resource, augment existing datasets for rare conditions, or function as an out-of-distribution test set.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

### 1. Chest x-ray image synthesis using deep convolutional gans

URL: [View paper](#)

#### Brief Assessment

Deep Convolutional GANs[23] generates 31,142 unlabeled chest x-ray images using DCGAN, whereas the original paper creates a curated dataset of 75,000 high-quality synthetic radiographs with multi-label pathology annotations and radiology reports. The candidate lacks the scale, curation process, quality filtering, and rich annotations that define SynthCheX-75K's novelty as a comprehensive training resource.

---

### 2. Roentgen: vision-language foundation model for chest x-ray generation

URL: [View paper](#)

#### Brief Assessment

Roentgen[8] focuses on developing a vision-language foundation model for chest x-ray generation conditioned on text prompts, with evaluation using image quality metrics. The candidate does not present a curated synthetic dataset comparable to SynthCheX-75K.

---

### 3. Synthetically enhanced: unveiling synthetic data's potential in medical imaging research

URL: [View paper](#)

#### Brief Assessment

Unveiling Synthetic Potential[63] discusses synthetic data augmentation in medical imaging but does not present a specific large-scale curated synthetic chest radiograph dataset comparable to SynthCheX-75K. The candidate focuses on analyzing effects of synthetic data rather than releasing a standalone dataset resource.

---

### 4. Progressive GAN Framework for Realistic Chest X-Ray Synthesis and Data Augmentation

URL: [View paper](#)

#### Brief Assessment

Progressive GAN Framework[4] focuses on a GAN-based synthesis method for chest X-rays but does not describe creating or releasing a large-scale curated synthetic dataset like SynthCheX-75K. The candidate discusses model architecture and evaluation metrics, not dataset curation or release.

---

### 5. Mitigating risk in medical AI: balancing X-ray datasets for reliable detection

URL: [View paper](#)

#### Brief Assessment

Mitigating Risk Balancing[61] focuses on balancing datasets and mitigating risk in medical AI detection tasks, not on creating large-scale synthetic chest radiograph datasets for training augmentation and out-of-distribution testing as described in the original paper's SynthCheX-75K contribution.

---

### 6. Dual-domain explainability-driven data augmentation for enhanced COVID-19 detection in chest X-rays

URL: [View paper](#)

#### Brief Assessment

Dual Domain Explainability[64] focuses on explainability-driven data augmentation for COVID-19 detection using CGANs, not on creating large-scale general-purpose synthetic chest radiograph datasets for training augmentation and out-of-distribution testing.

---

### 7. Gan-based data augmentation for chest x-ray classification

URL: [View paper](#)

#### Brief Assessment

GAN Based Augmentation[62] focuses on using GANs to augment the CheXpert dataset for improving classification of underrepresented classes, not on creating a standalone curated synthetic dataset like SynthCheX-75K. The candidate does not present a large-scale synthetic dataset release.

---

### 8. Generative models improve fairness of medical classifiers under distribution shifts

URL: [View paper](#)

## Brief Assessment

Fairness Distribution Shifts[59] focuses on using synthetic data to improve fairness and robustness of medical classifiers under distribution shifts, not on creating large-scale synthetic datasets for training augmentation or out-of-distribution testing as standalone resources.

---

## 9. A vision-language foundation model for the generation of realistic chest x-ray images

URL: [View paper](#)

### Brief Assessment

Vision Language Foundation[60] focuses on vision-language foundation models for chest X-ray generation, not on creating large-scale curated synthetic datasets for training augmentation and out-of-distribution testing as described in the original paper's SynthCheX-75K contribution.

---

## 10. Disentangled Contrastive Learning From Synthetic Matching Pairs for Targeted Chest X-Ray Generation

URL: [View paper](#)

### Brief Assessment

Disentangled Contrastive Learning[2] focuses on disentangled attribute control for generating matching pairs of chest X-rays with controllable disease attributes, not on creating large-scale standalone synthetic datasets for training augmentation or out-of-distribution testing.

---

## Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

---

## References

- [0] CheXGenBench: A Unified Benchmark for Fidelity, Privacy and Utility of Synthetic Chest Radiographs [View paper](#)
- [1] Synthetic chest X-ray data generation for tuberculosis infection detection using generative adversarial networks [View paper](#)
- [2] Disentangled Contrastive Learning From Synthetic Matching Pairs for Targeted Chest X-Ray Generation [View paper](#)
- [3] A Critical Assessment of Generative Models for Synthetic Data Augmentation on Limited Pneumonia X-ray Data [View paper](#)
- [4] Progressive GAN Framework for Realistic Chest X-Ray Synthesis and Data Augmentation [View paper](#)
- [5] Medical Image Synthesis Using DCGAN for Chest X-Ray Images [View paper](#)
- [6] Method for generating synthetic data combining chest radiography images with tabular clinical information using dual generative models [View paper](#)
- [7] Evaluating diagnostic content of AI-generated chest radiography: A multi-center visual Turing test [View paper](#)
- [8] Roentgen: vision-language foundation model for chest x-ray generation [View paper](#)
- [9] Improving Performance, Robustness, and Fairness of Radiographic AI Models with Finely-Controllable Synthetic Data [View paper](#)
- [10] Disease-Aware Chest X-Ray Style GAN Image Generation and CatBoost Gradient Boosted Trees [View paper](#)
- [11] Denoising diffusion probabilistic models for addressing data limitations in chest X-ray classification [View paper](#)
- [12] Multiscale attention generative adversarial networks for lesion synthesis in chest X-ray images. [View paper](#)
- [13] A preliminary analysis of AI based smartphone application for diagnosis of COVID-19 using chest X-ray images [View paper](#)
- [14] Generation of synthetic chest X-ray images and detection of COVID-19: A deep learning based approach [View paper](#)
- [15] Synthetic chest X-ray images from text prompts [View paper](#)
- [16] Image Turing test and its applications on synthetic chest radiographs by using the progressive growing generative adversarial network [View paper](#)
- [17] Introducing SPINE: A Holistic Approach to Synthetic Pulmonary Imaging Evaluation Through End-to-End Data and Model Management [View paper](#)
- [18] Synthesizing chest X-ray pathology for training deep convolutional neural networks [View paper](#)
- [19] Evaluating the clinical realism of synthetic chest x-rays generated using progressively growing gans [View paper](#)
- [20] A Generative Foundation Model for Chest Radiography [View paper](#)
- [21] Generative X-Ray Synthesis for Enhanced Disease Prediction in Medical Imaging [View paper](#)
- [22] Synthetic data in radiological imaging: current state and future outlook [View paper](#)
- [23] Chest x-ray image synthesis using deep convolutional gans [View paper](#)
- [24] CoVaD-GAN: An efficient Data Augmentation technique for COVID CXR Image Classification [View paper](#)
- [25] Beware of diffusion models for synthesizing medical images—a comparison with GANs in terms of memorizing brain MRI and chest x-ray images [View paper](#)
- [26] Evaluation of deep convolutional generative adversarial networks for data augmentation of chest x-ray images [View paper](#)
- [27] Improving Covid-19 cxr detection with synthetic data augmentation [View paper](#)
- [28] Prototype-Driven Class-Conditional Synthesis for High-Quality Chest X-ray Image Generation. [View paper](#)
- [29] Adapted generative latent diffusion models for accurate pathological analysis in chest X-ray images. [View paper](#)
- [30] You Don't Have to Be Perfect to Be Amazing: Unveil the Utility of Synthetic Images [View paper](#)
- [31] The Devil is in the Prompts: De-Identification Traces Enhance Memorization Risks in Synthetic Chest X-Ray Generation [View paper](#)
- [32] Feature Controlled Synthetic Medical Image Generation using Conditional Generative Adversarial Network [View paper](#)
- [33] Machine-Learning Based Detection of Coronary Artery Calcification Using Synthetic Chest X-Rays [View paper](#)
- [34] Generative adversarial networks for the synthesis of chest x-ray images [View paper](#)
- [35] Chest radiography few-shot image synthesis for automated pathology screening applications [View paper](#)
- [36] Evaluating and Improving the Effectiveness of Synthetic Chest X-Rays for Medical Image Analysis [View paper](#)
- [37] High-Resolution Chest Synthesis Data Generation for Medical Image Analysis: An Advanced GAN Approach with Conditional Image Synthesis [View paper](#)
- [38] Robust Approaches for Deep Learning for Chest X-ray Assessment Using Synthetically Generated Illness [View paper](#)
- [39] Advancing Medical Imaging: A Comparative Exploration of Generative Adversarial Networks for Chest X-ray Synthesis [View paper](#)
- [40] Recognizing Pneumonia in Real-World Chest X-rays with a Classifier Trained with Images Synthetically Generated by Nano Banana [View paper](#)
- [41] CycleGAN generated pneumonia chest x-ray images: Evaluation with vision transformer [View paper](#)
- [42] Sensitivity, Specificity, and Consistency: A Tripartite Evaluation of Privacy Filters for Synthetic Data Generation [View paper](#)
- [43] Synthetic Chest X-Ray Augmentation via Generative Variational Autoencoding for Pneumonia Detection [View paper](#)

- [44] Development of a deep neural network for generating synthetic dual-energy chest x-ray images with single x-ray exposure. [View paper](#)
- [45] Exploring Foundation Models for Synthetic Medical Imaging: A Study on Chest X-Rays and Fine-Tuning Techniques [View paper](#)
- [46] Cross-Modality Synthetic Data Augmentation using GANs: Enhancing Brain MRI and Chest X-ray Classification [View paper](#)
- [47] Image projective transformation rectification with synthetic data for smartphone-captured chest X-ray photos classification. [View paper](#)
- [48] Generative domain adaptation for chest X-ray image analysis [View paper](#)
- [49] Nodule Detection and Generation on Chest X-Rays: NODE21 Challenge. [View paper](#)
- [50] Evaluating the feasibility of using Generative Models to generate Chest X-Ray Data [View paper](#)
- [51] Chest-diffusion: a light-weight text-to-image model for report-to-cxr generation [View paper](#)
- [52] Generative AI Techniques in Medical Imaging Analysis: A Systematic Review [View paper](#)
- [53] Cxr-clip: Toward large scale chest x-ray language-image pre-training [View paper](#)
- [54] Any-to-Any Vision-Language Model for Multimodal X-ray Imaging and Radiological Report Generation [View paper](#)
- [55] Cxr-irgen: An integrated vision and language model for the generation of clinically accurate chest x-ray image-report pairs [View paper](#)
- [56] Synthetic lung x-ray generation through cross-attention and affinity transformation [View paper](#)
- [57] Covid-19 pneumonia chest x-ray pattern synthesis by stable diffusion [View paper](#)
- [58] Spot the fake lungs: Generating Synthetic Medical Images using Neural Diffusion Models [View paper](#)
- [59] Generative models improve fairness of medical classifiers under distribution shifts [View paper](#)
- [60] A vision-llm language foundation model for the generation of realistic chest x-ray images [View paper](#)
- [61] Mitigating risk in medical AI: balancing X-ray datasets for reliable detection [View paper](#)
- [62] Gan-based data augmentation for chest x-ray classification [View paper](#)
- [63] Synthetically enhanced: unveiling synthetic data's potential in medical imaging research [View paper](#)
- [64] Dual-domain explainability-driven data augmentation for enhanced COVID-19 detection in chest X-rays [View paper](#)
- [65] Clinical evaluation of medical image synthesis: a case study in wireless capsule endoscopy [View paper](#)
- [66] Design and development of a systematic validation protocol for synthetic melanoma images for responsible use in medical artificial intelligence [View paper](#)
- [67] Generative Artificial Intelligence in Medical Imaging: Foundations, Progress, and Clinical Translation [View paper](#)
- [68] Generating high-fidelity synthetic patient data for assessing machine learning healthcare software [View paper](#)
- [69] Scorecard for synthetic medical data evaluation [View paper](#)
- [70] Generative Adversarial Networks for Synthetic Biomedical Data: Ensuring Data Fidelity and Privacy Preservation [View paper](#)
- [71] SynthVal: A Framework for Validating Synthetic Medical Images [View paper](#)
- [72] Interpretable Similarity of Synthetic Image Utility [View paper](#)
- [73] Eyes Tell the Truth: GazeVal Highlights Shortcomings of Generative AI in Medical Imaging [View paper](#)