# Generative Breast Cancer Imaging: Lesion Synthesis and Risk Prediction with Latent Diffusion Models

*Nicola Todaro, Thomas Chen*

**Abstract**. In this work, we present a novel pipeline that leverages latent diffusion models (LDMs) and variational autoencoders (VAEs) to generate synthetic mammograms. By conditioning on lesion-specific descriptors and patient features (e.g., BI-RADS scores, mass density, calcifications), our method constructs a low-dimensional latent space that preserves diagnostic attributes essential for breast cancer imaging. Our approach not only augments rare cases in existing datasets (such as VinDr-Mammo and INbreast) but also opens pathways for improved in-silico patient profiling and risk-based imaging studies. Comparative evaluations against VAE baselines and diagnostic reliability tests (via reconstruction loops and segmentation/classification checks) demonstrate the promise of guided synthesis as a tool for medical imaging research.

**Keywords.** Latent Diffusion Models, Variational Autoencoders, Synthetic Mammograms, Breast Cancer Imaging, Data Augmentation, In-Silico Patient Profiling

## Introduction

Breast cancer remains one of the leading causes of mortality among women worldwide. Accurate imaging is key for timely diagnosis; however, conventional data are often limited—especially for rare cancer types. Recent advances in generative models, particularly latent diffusion models and VAEs, have shown potential in overcoming data scarcity by synthesizing realistic medical images. In this paper, we outline a framework that learns a latent representation embedding key diagnostic features such as mass density, calcifications, and tissue asymmetry, thereby enabling the generation of synthetic mammograms conditioned on lesion descriptors or patient metadata.

**Related Work**

Our study builds upon multiple breakthroughs in high-fidelity image synthesis. Notable related works include:

- **High-Resolution Image Synthesis with Latent Diffusion Models,** which demonstrates that latent spaces can be structured to maintain detailed image features.
- **Anatomically-Controllable Medical Image Generation with Segmentation-Guided Diffusion Models,** highlighting the importance of segmentation and anatomical priors in guiding image synthesis.
- **Towards Generalizable Tumor Synthesis,** emphasizing the capability of generative models to recreate diverse tumor morphologies.

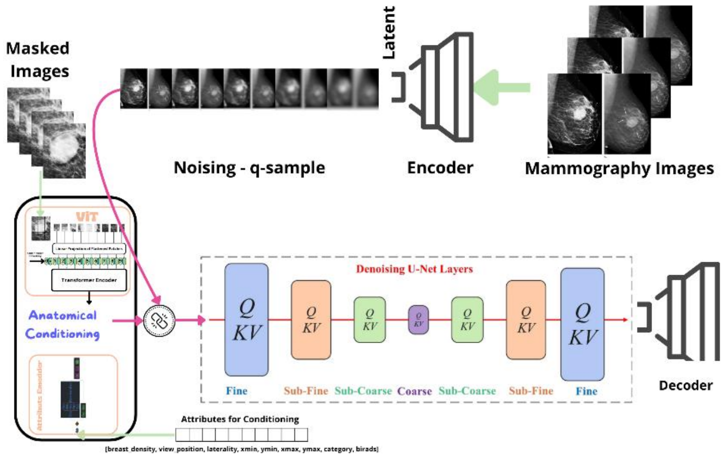
These contributions form the cornerstone of our approach as we extend their insights into breast cancer imaging applications.

# 1 Methodology

## 1.1 Pipeline Architecture

Our system consists of two primary components: a latent diffusion model and a VAE-based reconstruction pipeline. The process is illustrated in the attached diagram[1]:

- **Anatomical Conditioning:** Mammograms (or masked input images) are first processed by a Vision Transformer (ViT), embedded with key attributes (breast density, view position, laterality, bounding box coordinates, category, and BI-RADS scores).
- **Latent Encoding & Noising:** The conditioned input undergoes a noising process (q-sampling) before being encoded into a compact latent space. This stage leverages VAEs to guarantee that detailed cancer-relevant information is preserved.
- **Denoising U-Net:** A multi-scale denoising U-Net refines latent representations at various scales (from fine to coarse layers), ensuring that the generated images represent realistic lesion patterns.
- **Decoding:** Finally, a decoder reconstructs high-fidelity synthetic mammograms from the refined latent distribution.




[1]

## 1.2 Conditioning Mechanisms and Extensions

For improved control over the synthesis, we incorporate a conditioning mechanism using cross-attention and FiLM layers. This allows the pipeline to consider textual embeddings (derived from radiology reports or simple descriptors such as “mass in upper left quadrant”) potentially generated via CLIP or BioGPT encoders.

An additional reconstruction loop is introduced in groups of three:

- Synthetic images are analyzed by a pre-trained classifier/segmentation model.
- A cycle-consistency loss ensures the preservation of lesion descriptors through the reconstruction process.

Furthermore, our work explores domain translation: generating lesion-present images from healthy scans, simulating disease progression through unpaired GANs or diffusion-to-diffusion translation modules.

# 2. Experiments and Evaluation

## 2.1 Datasets and Training

We utilize publicly available datasets (e.g., VinDr-Mammo, INbreast) to train and validate our models. Data augmentation techniques are applied, and the training protocols are adapted to ensure stability during diffusion model training.

## 2.2 Evaluation Metrics

Evaluation is twofold:

- **Image Quality:** Measured via perceptual similarity metrics and fidelity assessments compared to real mammograms.
- **Diagnostic Reliability:** Evaluated by a reconstruction loop method where generated images are passed through diagnostic classifiers. Cycle-consistency loss provides additional checks on lesion integrity.

**Results and Discussion.** Our experiments demonstrate that the proposed latent diffusion model captures essential cancer-related features, improving the diversity and quality of synthetic mammograms. The diagnostic loops reveal that the generated lesions preserve key attributes necessary for reliable risk assessment. Comparative studies with VAE baselines indicate significant improvements in realism and morphological accuracy.

**Future Work.** Potential enhancements include:

- **Longitudinal Lesion Evolution Modeling:** Generating sequences to simulate temporal changes in lesions.
- **Counterfactual Synthesis:** Creating what-if scenarios by generating healthy variants of cancerous images.
- **Multimodal Generation:** Extending to simultaneous synthesis of complementary imaging modalities (e.g., ultrasound with corresponding biopsy images).

**Conclusion.** We have developed a robust framework for synthetic mammogram generation that augments traditional imaging datasets with realistic, medically relevant images. Our latent diffusion and VAE-based pipeline not only offer a promising tool for medical research but also pave the way for advanced diagnostic aids and explainable AI applications in oncology.

**References**


1. High-Resolution Image Synthesis with Latent Diffusion Models. Link
2. Anatomically-Controllable Medical Image Generation with Segmentation-Guided Diffusion Models. Link
3. Towards Generalizable Tumor Synthesis. Link


*Additional details, parameters, and repository code are accessible via the ISPAMM GitHub repositories and our public repository.*