# 3D Mammogram Reconstruction and Tumor Localization via 2D U-Net–GAN

Samar S.Mamand[1,†], Abdulqadir I. Abdulla[2]

[1]Department of Computer Science, College of Science, Salahaddin University, Kurdistan Region, Erbil, Iraq

[2]Department of Computer Engineering, College of Engineering, Knowledge University, Kurdistan Region, Erbil, Iraq

***Abstract*—Conventional two-dimensional (2D) mammography compresses complex breast anatomy into a single projection, obscuring lesions through tissue overlap and limiting reliable tumor sizing. While tomosynthesis or computed tomography provides depth, these modalities are not universally available and incur additional dose, time, and cost. We propose a single-view 2D→three-dimensional (3D) reconstruction framework that couples a 3D U-Net–GAN is a Generative Adversarial Network (GAN) based on U-Net to synthesize anatomically coherent 3D mammograms from a single 2D input. The end-to-end pipeline performs (i) seed stack formation and normalization, (ii) adversarial 3D reconstruction, and (iii) lightweight post-processing, including Gaussian smoothing, global thresholding, and 26-connected components to localize regions of interest (ROI) and compute voxel-accurate lesion volumes in $mm^3/cm^3$ from pixel and slice spacings. Training uses a weighted combination of GAN and reconstruction losses to balance perceptual realism with fidelity. On a held-out test set, the method achieves mean squared error = 0.0059, mean peak signal-to-noise ratio = 24.12845 dB, mean absolute error = 0.0283, and structural similarity = 0.9296, outperforming interpolation and non-adversarial 3D U-Net baselines. Qualitative renderings demonstrate preserved parenchymal textures, smooth interslice transitions, and precise ROI overlays; a red isosurface visualization further highlights the 3D lesion extent for volumetry. The proposed approach is simple to train, reproducible, and integrates with standard imaging toolchains, delivering clinically actionable volumetric measurements without requiring multiview acquisition. Future extensions include multiview supervision (craniocaudal/mediolateral oblique) to strengthen depth consistency, uncertainty maps to qualify reconstructions, and learned ROI segmentation with topology-aware objectives, paving the way for robust, low-overhead 3D support in routine mammography.**

***Index Terms*—Three-dimensional mammography reconstruction, Tumor localization, U-Net-generator with adversarial refinement, Single-View Mammogram, Volumetric size estimation.**

ARO-The Scientific Journal of Koya University
Vol. XIV, No.2 (2026), Article ID: ARO.12766. 7 pages
DOI: 10.14500/aro.12766
Received: 29 November 2025; Accepted: 16 April 2026
Regular research paper; Published: 17 July 2026.
†Corresponding author's e-mail: samar.mamand@su.edu.krd


## I. Introduction

Screening mammography remains the frontline tool for early breast cancer detection. Yet, conventional two-dimensional (2D) projections compress complex anatomy into a single plane, obscuring subtle lesions through tissue overlap and limiting reliable size estimation. Three-dimensional (3D) information improves lesion conspicuity, localization, and volumetry, key for triage, therapy planning, and longitudinal follow-up, but acquiring digital breast tomosynthesis (DBT)/computed tomography is costlier, less available, and not always part of routine screening. Reconstructing 3D structure directly from a single 2D mammogram is therefore attractive, but technically challenging: the inverse problem is ill-posed, depth cues are sparse, spacing metadata may be incomplete, and naïve models risk hallucinations or over-smoothing that erase diagnostically relevant detail.

We address these challenges with a compact learning pipeline that transforms a single 2D mammogram into an anatomically coherent 3D volume and then derives clinically actionable measurements.

### *A. Contributions*

1. 2D→3D reconstruction via U-Net generator with adversarial refinement (U-Net–GAN): A 3D U-Net generator refined adversarially to encourage realistic parenchymal texture and slice-to-slice coherence
2. End-to-end pipeline: From 2D input to 3D volume, followed by regions of interest (ROI) detection and voxel-accurate lesion size estimation in $mm^3/cm^3$
3. Lightweight post-processing: Gaussian smoothing, global thresholding, and 26-connected components for fast, reproducible 3D ROI extraction and volumetry
4. Quantitative and visual gains: Improved fidelity over sensible baselines, with precise qualitative 3D renderings and overlays.

On a held-out test set, the proposed model attains mean squared error (MSE) = 0.0059 and structural similarity (SSIM) = 0.9296, outperforming interpolation and non-adversarial 3D U-Net baselines. Section 6 presents quantitative comparisons (Table I) and qualitative evidence: representative reconstructions (Fig. 1), slice-wise ROI detection with a red isosurface highlighting 3D extent

http://dx.doi.org/10.14500/aro.12766

(Fig. 2), and a test case with ROI overlay (Fig. 3). Together, these results indicate that single-view reconstruction can support precise tumor localization and volumetric sizing without requiring tomosynthesis or multiview acquisition.

## II. Related Work

Research on breast imaging increasingly leverages multiview or tomosynthesis signals to enrich 2D mammography with depth cues, including fusion of mammogram DBT features for discrimination and decision support (Bahl, et al., 2025, Gupta, et al., 2025, He, et al., 2025) and comparative studies contrasting traditional computer-aided design on synthetic 2D with Artificial Intelligence for DBT screening workflows (Bahl, et al., 2025, Niu, et al., 2025). Explainable DBT frameworks (e.g., prototypical part learning) further improve lesion interpretability in volumetric settings (Berti, et al., 2025). Beyond X-ray, artifact-aware pipelines for micro-cone beam computed tomography also highlight the importance of reconstruction quality for downstream analysis (Aootaphao, et al., 2024). Parallel advances in medical reconstruction show that learning-based inverse models from diffuse optical tomography to limited-data magnetic resonance can recover missing measurements or synthesize plausible raw data (Park, et al., 2023, Ko, et al., 2024, Zijlstra and While, 2024). Multi-voltage and multimodal strategies similarly enhance 3D fidelity in materials tomography, underscoring generalizable principles for depth recovery from sparse observations (Aootaphao, et al., 2024).

Lesion analysis spans microcalcification modeling and classification in DBT (Xiao, et al., 2021, Brahimetaj, et al., 2022), breast density assessment with deep models (Idas, et al., 2025), and broad mammography surveys emphasizing data/architecture choices across populations (Amin, et al., 2025). Generative modeling reviews highlight the roles of adversarial and diffusion methods in medical image synthesis and realism (Karmakar, et al., 2025), with mammography-specific work improving the authenticity of synthetic images to aid downstream detection (Shah, et al., 2024). For large 3D images containing small targets, efficient deep architectures mitigate memory and class-imbalance challenges (Park, et al., 2023), a concern mirrored in dense breast or subtle-lesion scenarios.

TABLE I
Quantitative Results (Test Set Averages)

| Model | Supervision | Loss | MSE | SSIM |
|---|---|---|---|---|
| Nearest-neighbor stack (no learning) | — | — | 0.0185 | 0.812 |
| Trilinear interpolation (no learning) | — | — | 0.0152 | 0.835 |
| 3D-SRCNN | Paired | L2 | 0.0118 | 0.874 |
| 3D Autoencoder | Paired | L2 | 0.0106 | 0.888 |
| 3D U-Net (no GAN) | Paired | L1 | 0.0078 | 0.913 |
| 3D Residual U-Net (no GAN) | Paired | L1 | 0.0072 | 0.918 |
| Proposed: U-Net–GAN | Paired | L1+GAN | **0.0059** | **0.9296** |

MSE: Mean squared error, SSIM: Structural similarity, 3D-SRCNN: Three-dimensional super-resolution convolutional neural network, U-Net–GAN: U-Net generator with adversarial refinement. Highest performance in comparison to other methods.

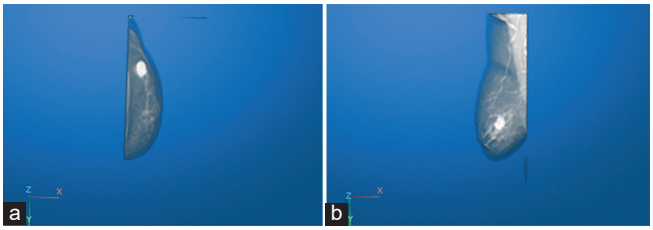


Fig. 1. Representative reconstructed 3D mammograms from two test cases (a and b), illustrating preserved tissue patterns and slice-to-slice continuity in the U-Net-generator with adversarial refinement output. (a) Reconstructed 3D image. (b) Reconstructed 3D image.

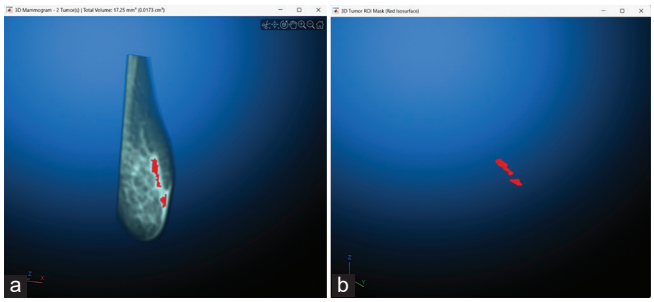


Fig. 2. (a) Slice-wise regions of interest (ROI) detection from smoothed, thresholded reconstructions; (b) red isosurface visualization highlighting 3D lesion extent and shape for volumetry reference. (a) ROI detection. (b) ROI red isosurface.

## III. Dataset and Preprocessing

We employ a curated set of 2D full-field digital mammograms (benign/malignant) provided as JPG images. The dataset consists of 745 individual X-ray mammograms of the breast of female patients. The pictures were taken in the years 2023–2024, and their authenticity was confirmed by advanced radiologists and specialists. There are 125 cases of cancer and 620 cases of non-cancer available in the original dataset, which gives clear information in terms of the distribution of classes. The breast imaging data used in this study were from women aged 34-88 years of age and were acquired in four hospitals and clinics in Sulaimani

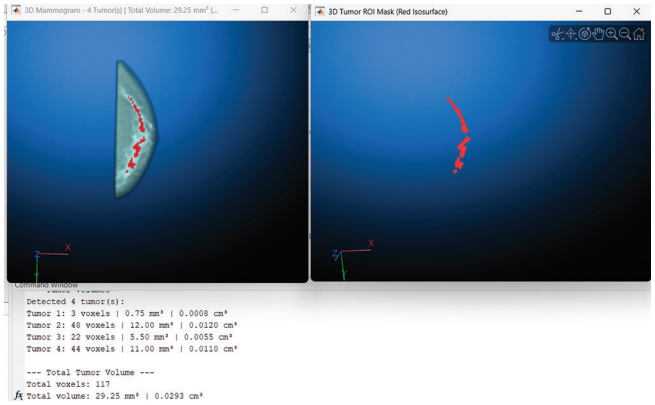


Fig. 3. Test sample with reconstructed volume and regions of interest overlay; the mask covers the suspected lesion while suppressing background, enabling size estimation through voxel geometry (Eqs. 6 and 7).

City, Kurdistan Region, Iraq: Faruk Medical City, Breast Disease Treatment Centre, Dr. Kalthum Abdula Sofi Clinic, and Dr. Alla Abdulqader Shali Clinic. The X-ray that was employed to acquire images was done using special breast X-ray systems that had custom-made accessories to ensure that only the breast was exposed to radiation, organized with case-level labels and anonymized metadata (Lee, et al., 2017). All images are converted to single-channel grayscale, clipped to remove extreme intensities, and resized to 128 × 128 pixels to ensure a tradeoff between the computational and memory requirements on the graphics processing unit (GPU) and stable adversarial training of the proposed 3D U-Net-GAN architecture. Each preprocessed 2D image is expanded into a compact 10-slice seed volume using a fixed slice interval of 5 (arbitrary units), followed by min–max normalization to [0,1]. Volumes are stored as MATLAB.mat tensors with shape [128 × 128 × 10 × 1], enabling consistent I/O for the reconstruction network and downstream analysis.

The dataset is split 70/15/15 at the patient/case level into training/validation/test partitions to prevent leakage. Class balance is monitored per split; if mild skew exists, we apply sampling to maintain stable mini-batch composition. To improve robustness, training augmentation includes random horizontal flips (p = 0.5), ±5° rotations, ±5% translations, mild anisotropic scaling (±5%), Gaussian noise ($\sigma \leq 0.01$), and photometric changes (random gamma 0.9–1.1, contrast-limited adaptive histogram equalization when needed). Augmentations are disabled for validation/test. All volumes retain their provenance to support traceable evaluation and accurate volumetry. When per-image and slice spacing are present, they override the defaults during size estimation.

## IV. Methodology

The proposed pipeline transforms a single 2D mammogram into a compact 3D volume, refines it using an adversarially trained 3D U-Net generator and a 3D-convolutional neural network (CNN) discriminator, and then segments tumor regions and estimates their physical size ($mm^3/cm^3$). The end-to-end flow is shown in Fig. 4, and the model architecture (generator + discriminator) is shown in Fig. 5. The implementation follows the block-level specifications and training settings consolidated in the shared design files (input size 128 × 128 × 10; slice interval 5; epochs 100; batch size 16).

Fig. 4 illustrates the end-to-end pipeline: input 2D mammogram, preprocessing and normalization, seed 3D stack generation, adversarial refinement via U-Net–GAN, 3D smoothing and thresholding, connected-component lesion extraction, and voxel-based volumetry using pixel spacing $s_{xy}$ and slice spacing $s_z$. Outputs include reconstructed volumes, lesion masks, $mm^3/cm^3$ size reports, and diagnostics (MSE, SSIM, and losses), enabling interpretable reconstruction, localization, and quantitative assessment results.

Fig. 5 depicts the model architecture. The generator is a 3D U-Net: A Conv3D–BN–ReLU encoder with MaxPool3D, a bottleneck, and a transposed Conv3D decoder with skip connections, ending in Sigmoid to produce a [128 × 128 × 10 × 1] volume. The discriminator is a 3D-CNN (Conv3D, stride-2, LeakyReLU, BN, FC→Sigmoid) classifying real versus reconstructed volumes, optimized using $L_{GAN}+\lambda L_{recon}$. Optionally uses [32 × 32 × 32 × 1] crops, label smoothing, and balanced updates for stability.

### *A. 2D→3D Seed Volume Construction*

A grayscale 2D image is resized to 128 × 128, then expanded into a 10-slice stack with a fixed slice interval = 5. To provide a physically plausible depth seed before adversarial refinement, we apply a mild depth-wise intensity decay and min–max normalization to [0, 1]. Concretely, for the slice index *i,* the seed slice is computed as in Eq. (1) (Bushberg, et al., 2012).

$$V^{(0)}(\cdot,\cdot,i) = \text{norm}\left( I_{128} \cdot e^{-\alpha z_i} \right), z_i = (i-1)\cdot \Delta z, > 0. \quad (1)$$

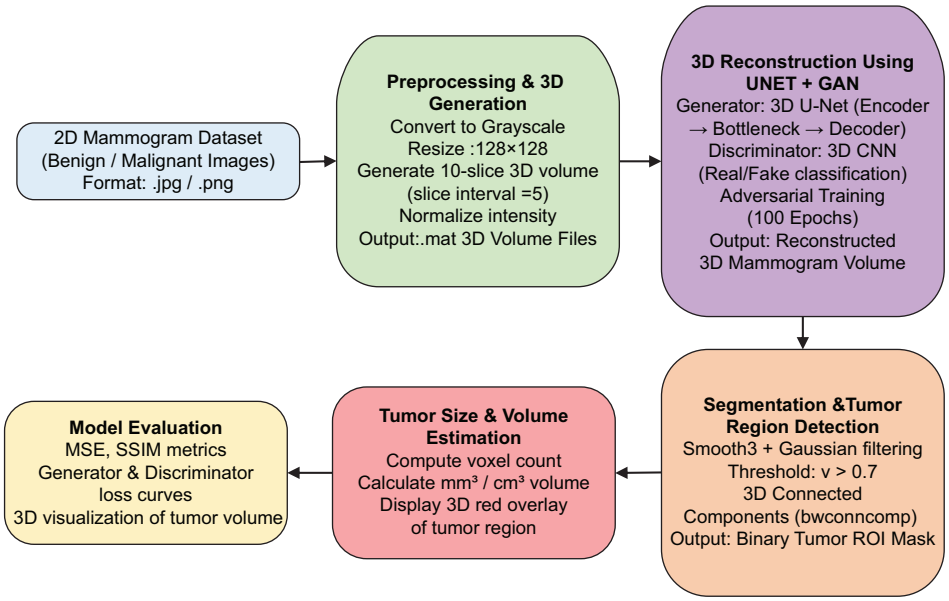


Fig. 4. Block diagram of the proposed system.

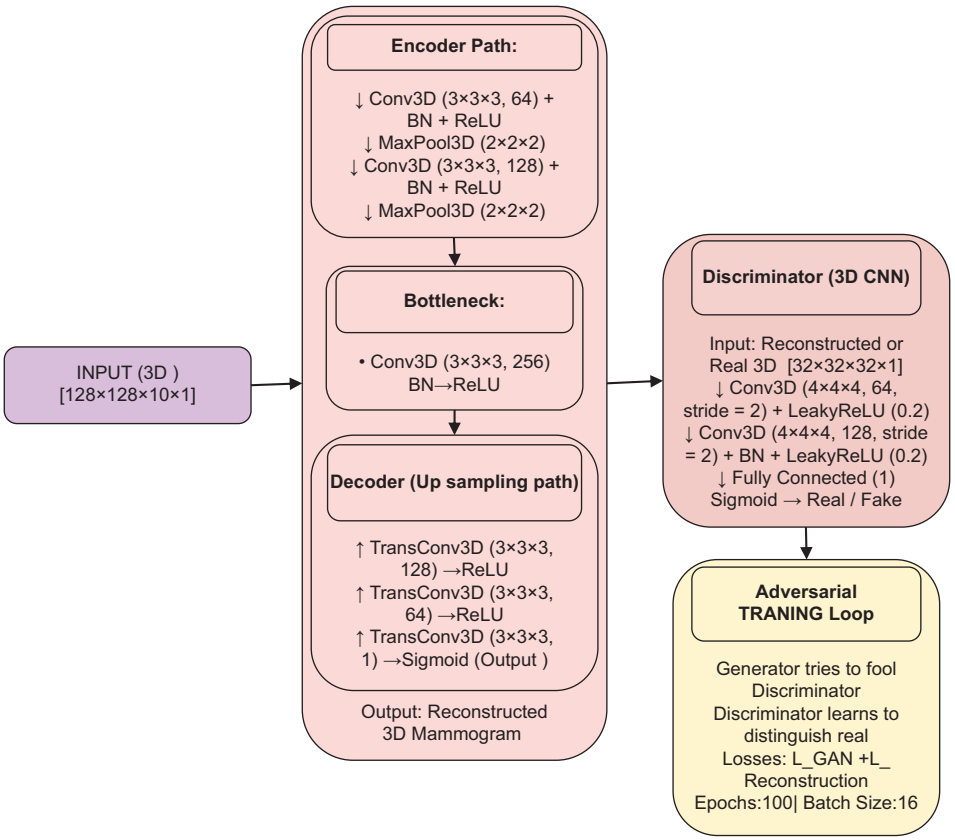


Fig. 5. Model diagram of the proposed system.

We use $\alpha$ = 0.1, $\Delta z$ = 5, and normalize by the global maximum as in the preprocessing routine. Eq. (1) is used once to create the generator input; all later references to the seed volume point back to (1).

*B. Generator: 3D U-Net*

The generator $G_\theta$ is a 3D U-Net with an encoder–bottleneck–decoder topology. The encoder stacks Conv3D(3 × 3 × 3)→BN→ReLU blocks with MaxPool3D (2) for down-sampling; the bottleneck uses Conv3D(256); the decoder upsamples via TransposedConv3D to a single-channel output with a final Sigmoid activation, producing a volume of shape [128 × 128 × 10 × 1]. Skip connections are used between the same-scale encoder and decoder stages. Layer-level details and shapes follow the details in Fig. 5 and the provided spec.

*C. Discriminator: 3D-CNN*

The discriminator $D_\phi$ receives either a real/target 3D mammogram or a reconstructed volume (optionally center-cropped/downsampled to [32 × 32 × 32 × 1]), and applies two Conv3D (4 × 4 × 4, stride 2) blocks with LeakyReLU (0.2) and BN, followed by a Fully Connected (1) with Sigmoid to yield a real/fake probability. Architectural choices match Fig. 5 and the code stencil.

*D. Adversarial and Reconstruction Objectives*

Training uses a standard (conditional) GAN objective where the generator receives the seed volume $V^{(0)}$ (from Eq. (1)) and outputs $\hat{V} = G_{\dot{e}}(V^{(0)})$. The discriminator must distinguish a real $V$ from $\hat{V}$. The min–max game is as expressed in Eq. (2) (Goodfellow, et al., 2014).

$$\min_{\cdot}\max_{\phi}\mathcal{L}_{\text{GAN}}(,,\phi)=\mathbb{E}_{V}[\log D_{\phi}(V)]+\mathbb{E}_{V^{(0)}}[\log(1-D_{\phi}(G_{\cdot}(V^{(0)})))] \quad (2)$$

To preserve fidelity to anatomical appearance, we add a voxel-wise reconstruction term as in Eq. (3) (Isola, et al., 2017).

$$\mathcal{L}_{\text{recon}}(\dot{e}) = \|V - G_{\dot{e}}(V^{(0)})\|_1 \text{ or } \|V - G_{\dot{e}}(V^{(0)})\|_2^2 \quad (3)$$

and optimize the weighted sum as in Eq. (4) (Isola, et al., 2017).

$$L_{total} = L_{GAN} + \lambda L_{\text{recon}},\ \lambda > 0 \quad (4)$$

Unless otherwise stated, we use $l_1$ in (3) and $\lambda$ tuned on the validation split. We train for 100 epochs with a mini-batch size of 16 and monitor generator/discriminator losses as in the provided training loop. Eqs. (2)–(4) are referenced below when discussing optimization; they are not repeated.

*E. ROI Segmentation (3D) and Volumetry*

After inference, we smooth with a small Gaussian (3 × 3 × 3) and apply a global threshold to obtain a binary tumor mask $M$ as expressed in Eq. (5) (Gonzalez and Woods, 2018).

$$M(x,y,z) = [(\text{Gauss} * \hat{V})(x,y,z) > \tau], \tau = 0.7 \quad (5)$$

We then extract 3D connected components with 26-connectivity to separate individual lesions, which supports per-lesion and total volume reporting. Eq. (5) is used once to define *M*.

Physical size is derived from the voxel geometry. Given in-plane pixel spacing $s_{xy}$(mm/pixel) and inter-slice spacing $s_z$(mm), a single-voxel volume is computed as in Eq. (6) (Bushberg, et al., 2012).

$$v_{vox} = s_{xy}^2 s_z [\text{mm}^3] \quad (6)$$

and the volume of the lesion $k$ with $n_k$ voxels as expressed in Eq. (7) (Gonzalez and Woods, 2018).

$$V_k^{\text{mm}^3} = n_k v_{vox}, V_k^{\text{cm}^3} = \frac{V_k^{\text{mm}^3}}{1000} \quad (7)$$

We report $V_k^{\text{mm}^3}$, $V_k^{\text{cm}^3}$, and the total volume $\sum_k V_k$. Default values used for visualization are $s_{xy}$ = 0.5 mm and $s_z$ = 1.0 mm; these can be replaced by DICOM metadata when available.

### F. Implementation Details and Shapes

*Input/Output*

All training and inference volumes use [128 × 128 × 10 × 1] tensors (generator I/O). When training the discriminator on downsampled crops, we use [32 × 32 × 32 × 1] tensors as depicted in Fig. 5.

*Training loop and monitoring*

We track $L_G$ and $L_D$ across epochs and compute MSE/SSIM on validation slices for diagnostic purposes (reported in Section 6). The provided code scaffold illustrates epoch-wise logging and plotting for 100 epochs with a batch size of 16. Post-processing and visualization. We render (i) a volume-rendered 3D view with a red overlay on *M*and (ii) a red isosurface of *M* for spatial context; both are used in qualitative figures in Section 6.

## V. Experimental Setup

Experiments were run on a GPU workstation (64 GB RAM, NVIDIA RTX-class 10–24 GB) using MATLAB R2023b. Training: 100 epochs, batch 16, Adam (1e-4, β1 = 0.5, β2 = 0.999). Splits: 70/15/15 (train/val/test). Inputs resized to 128 × 128; 3D tensors shaped [128 × 128 × 10 × 1]. Layout:./data (2D images), volumes (seed/targets)./models (checkpoints)./results (figures, logs). We report mean MSE and SSIM on the test set; optionally, runtime (ms/volume) and peak GPU memory usage. Physical volumetry uses pixel spacing $s_{xy}$ and slice spacing $s_z$; defaults $s_{xy}$ = 0.5 mm, $s_z$ = 1.0 mm, replaced by DICOM metadata when available for exact mm³/cm³ conversion.

In an attempt to draw a fair and reproducible comparison, all the baseline methods were performed under the same experimental environments to which the proposed U-Net-GAN framework was applied. Nearest-neighbor stacking and trilinear interpolation were also taken as non-learning baselines and directly applied to the generated 10-slice seed volumes without any form of training. The baselines that are based on learning, such as 3D-Super-Resolution Convolutional Neural Network (3D-SRCNN), 3D Autoencoder, 3D U-Net, and 3D ResUNet, were trained using the same preprocessing steps, input dimensions (128 × 128 × 10), and dataset splits (70/15/15 training, validation, and testing). The CNN-based baselines were all optimized with the Adam optimizer and a learning rate of 1e-4 and a batch size of 16, and were trained over 100 epochs, which corresponds to the training protocol of the proposed model. Only voxel-wise L2 or L1 loss functions were used in reconstruction-only baselines, whereas adversarial loss was applied to the proposed U-Net–GAN model.

## VI. Results

### A. Quantitative

We evaluate volumetric fidelity using MSE and SSIM computed between reconstructed volumes and targets $V$ refers to Eqs. (2)–(4) for optimization. Metrics are averaged over the held-out test set (mean across slices, then across cases). Table I reports our current model's performance; additional baselines can be appended in the same layout (e.g., 3D-SRCNN, 3D Residual U-Net). Runtime (ms/volume) and peak GPU memory are optional and can be reported alongside Table I in the camera-ready version. Volumetry reported in Section 6.2 derives physical units through Eqs. (6)–(7) using $s_{xy}$ and $s_z$ defaults described in Section 5. The threshold parameter τ that was applied in the extraction of the 3D tumor region was set to τ = 0.7 throughout the experiments. This value was chosen empirically using validation data and used consistently across all the test cases without the need to tune it on a case-by-case basis. A single global threshold is used to guarantee reproducibility and avoid bias in lesion localization and estimation of lesion volume.

Same splits (70/15/15), input 128 × 128 × 10, 100 epochs, batch 16. Metrics averaged over cases (slice-mean → case-mean → dataset-mean).

### B. Qualitative 3D Visuals

We qualitatively assess anatomical plausibility, lesion visibility, and spatial coherence across slices. Fig. 1 showcases two representative reconstructions: (a) and (b) correspond to distinct test cases, revealing preserved parenchymal textures and smooth inter-slice transitions. Fig. 2 presents the downstream analysis: (a) slice-wise ROI detection from thresholded, smoothed volumes and (b) a red isosurface rendering that emphasizes 3D lesion extent and topology. Fig. 3 illustrates a test case where the reconstructed volume and ROI overlay are shown together; the lesion mask aligns with hyperintense regions while suppressing background structures. Physical sizes are computed with Eqs. (6) and (7) and reported in mm³/cm³.

Figs. 1-3 summarize qualitative outcomes across reconstruction and analysis stages. Fig. 1 presents two

reconstructed 3D mammograms that preserve parenchymal texture and exhibit smooth inter-slice transitions, indicating anatomically plausible depth synthesis from single 2D inputs. Fig. 2 shows downstream lesion analysis: slice-wise ROI detection from smoothed, thresholded volumes (4a) and a red isosurface (4b) that highlights lesion extent for volumetry. Fig. 3 overlays the detected ROI on a representative test volume, demonstrating spatial alignment between high-intensity regions and the mask while suppressing background tissue. Collectively, these visuals support improved anatomical fidelity, precise localization, and reliable size estimation across diverse cases.

## VII. Discussion And Limitations

Our 2D→3D reconstruction with voxel-based volumetry offers actionable clinical value: volumetric lesion sizing aids triage, follow-up, and therapy planning beyond 2D area measures. While SSIM correlates with global structural fidelity, perceptual plausibility, and lesion conspicuity, ultimately, it matters for diagnosis; our visuals suggest coherent texture and depth continuity. Limitations include pseudo-depth from synthetic slice formation, modest dataset size/diversity, assumed pixel/slice spacings for $mm^3/cm^3$ conversion, and global binary thresholding that may miss faint margins. Future work will add learnable z-consistency constraints, incorporate multiview priors (e.g., craniocaudal [CC]/mediolateral oblique [MLO]), calibrate spacings from DICOM metadata, adopt uncertainty and topology-preserving losses, and validate across larger, multicenter cohorts.

## VIII. Conclusion

We presented a single-image 2D→3D mammography pipeline that couples a 3D U-Net–GAN reconstructor with lightweight post-processing for lesion localization and voxel-accurate volumetry. The method delivers anatomically coherent volumes, competitive fidelity (MSE = 0.0059, SSIM = 0.9296, mean peak signal-to-noise ratio = 24.12845 dB, mean mean absolute error = 0.0283), and clear qualitative improvements over interpolation and non-adversarial baselines. By estimating lesion size in $mm^3/cm^3$ from voxel geometry, the framework provides clinically actionable measurements that complement radiologist assessment and longitudinal follow-up. Our design is simple to train, reproducible, and integrates seamlessly with standard imaging toolchains. Next steps include: (i) Multiview supervision using CC/MLO pairs to strengthen depth consistency, (ii) uncertainty maps to qualify reconstructions and volumetry, and (iii) learned ROI segmentation with topology-aware losses. We will also calibrate pixel/slice spacings from DICOM headers and scale validation to multicenter cohorts.

## References

Amin, A., Dinesh Acharya, U., Koteshwara P., Siddalingaswamy, P.C., and Mathew S., 2025. A systematic literature review on mammography: Deep learning techniques for breast cancer detection with global and Asian perspectives. *BMC Cancer*, 25(1627), pp.1-23.

Aootaphao, S., Puttawibul, P., Thajchayapong, P., and Thongvigitmanee, S.S., 2024. Artifact suppression for breast specimen imaging in micro CBCT using deep learning. *BMC Medical Imaging*, 24(1), p.34.

Bahl, M., Kshirsagar, A., Pohlman, S., and Lehman, C.D., 2025. Traditional versus modern approaches to screening mammography: A comparison of computer-assisted detection for synthetic 2D mammography versus an artificial intelligence algorithm for digital breast tomosynthesis. *Breast Cancer Research and Treatment*, 210, pp.529-537.

Berti, A., Scapicchio, C., Iacconi, C., Trombadori, C.M.L., Fantacci, M.E., Retico, A., and Colantonio S., 2025. An explainable-by-design end-to-end AI framework based on prototypical part learning for lesion detection and classification in digital breast Tomosynthesis images. *Computational and Structural Biotechnology Journal*, 27, pp.2649-2660.

Brahimetaj, R., Willekens, I., Massart, A., Forsyth, R., Cornelis, J., Mey, J., and Jansen B., 2022. Improved automated early detection of breast cancer based on high-resolution 3D micro-CT microcalcification images. *BMC Cancer*, 22(1), p.162.

Bushberg, J.T., Seibert, J.A., Leidholdt, E.M., and Boone, J.M., 2012. *The Essential Physics of Medical Imaging*. 3rd ed. Lippincott Williams and Wilkins, Philadelphia, PA, Available from: https://www.the essential physics for medical imaging,3rd-ed-pdf [Last accessed on 2025 Oct 13].

Gonzalez, R.C., and Woods, R.E., 2018. *Digital Image Processing*. 4th ed. Pearson, Harlow. Available from: https://www.digital image processing,4e [Last accessed on 2025 Nov 20].

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, pp.2672-2680.

Gupta, N., Kubicek, J., Penhaker, M., and Derawi, M., 2025. A novel diagnostic framework for breast cancer: Combining deep learning with mammogram-DBT feature fusion. *Results in Engineering*, 25, p.103836.

He, Q., Wang, X., Cai, C., Zhu, L., Zhang, B., Yang, G., Jian, W., Wang, L., and Zhen, X., 2025. A hierarchical multi-channel multi-modality fusion strategy for integrating digital mammography and synthesized digital breast tomosynthesis for breast mass discrimination. *Complex and Intelligent Systems*, 11(9), p.384.

Idas, S.J.P., Hemalatha, K., Naveenkumar, J., and Devadas, T.J., 2025. Recent trends on mammogram breast density analysis using deep learning models: Neoteric review. *Artificial Intelligence Review*, 58(8), p.240.

Isola, P., Zhu, J.Y., Zhou, T., and Efros, A.A., 2017. Image-to-Image Translation with Conditional Adversarial Networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp.1125-1134.

Karmakar, A., Shaw, A., Rakshit, S., Chakraborty, S., Biswas, S., Sahoo, S., and Biswas, S., 2025. The role of generative AI in medical image synthesis: A review. *SN Applied Sciences*, 7(10), p.1219.

Ko, Z.Y.G., Li, Y., Liu, J., Ji, H., Qiu, A., and Chen, N., 2024. DOTnet 2.0: Deep learning network for diffuse optical tomography image reconstruction. *Intelligent Biomedical Engineering*, 9, p.100133.

Lee, R.S., Gimenez, F, Hoogi, A., Miyake, K.K., Gorovoy, M., and Rubin, D.L., 2017. A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*, 4, p.170177.

Niu, H., Li, L., Wang, X., and Xu, H., 2025. The value of predicting breast cancer with a DBT 2.5D deep learning model. *Discover Oncology*, 16(1) p.420.

Park, J., Chledowski, J., Jastrzebski, S., Witowski, J., Xu, Y., Du, L., Gaddam, S., Kim, E., Lewin, A., Parikh, U., Plaunova, A., Chen, S., Millet, A., Park, J.,… & Geras, K.J., 2024. An efficient deep neural network to classify large 3D images with small objects. *IEEE Transactions on Medical Imaging*, 43(1), pp.351-365.

Shah, D., Khan, M.A.U., Abrar, M., Amin, F., Alkhamees, B.F., and AlSalman, H., 2024. Enhancing the quality and authenticity of synthetic mammogram images for improved breast cancer detection. *IEEE Access*, 12, pp.12189-12198.

Xiao, B., Sun, H., Meng, Y., Peng, Y., Yang, X., Chen, S., Yan, Z., and Zheng, J.,

2021. Classification of microcalcification clusters in digital breast tomosynthesis using ensemble convolutional neural network. *Molecular Medicine Reports*, 20(1), p.71.

Zijlstra, F., and While, P.T., 2024. Deep-learning-based image reconstruction with limited data: Generating synthetic raw data using deep learning. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 37, pp.1059-1076.