Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics e-ISSN: 2656-8624
Homepage: https://ijeeemi.org/; Vol. 8, No. 1, pp. 56-65, February 2026

RESEARCH PAPER

OPEN ACCES

Generating Synthetic B-Mode Fetal Ultrasound Images Using CycleGAN-Based Deep Learning

Fajar Astuti Hermawati[1], Bagus Hardiansyah[2], and Andrianto[3]
[1] Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia
[2] Department of Robotics and Artificial Intelligence, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia
[3] Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia

Abstract
B-mode ultrasound (USG) is a key imaging modality for fetal assessment, providing a noninvasive approach to monitor anatomical development and detect congenital anomalies at an early stage. However, portable ultrasound devices commonly used in low-resource healthcare settings often yield low-resolution images with significant speckle noise, reducing diagnostic accuracy. Furthermore, the scarcity of labeled medical data, caused by privacy regulations such as HIPAA and the high cost of expert annotation, poses a significant challenge for developing robust artificial intelligence (AI) diagnostic models. This study proposes a CycleGAN-based deep learning model enhanced with a histogram-guided discriminator (HisDis) to generate realistic synthetic B-mode fetal ultrasound images. A publicly available dataset from the Zenodo repository containing 1,000 grayscale fetal head images was utilized. Preprocessing included normalization, histogram equalization, and image resizing, while the architecture combined two ResNet-based generators and a dual discriminator configuration integrating PatchGAN and histogram-guided evaluation. The model was trained using standard optimization settings to ensure stable convergence. Experimental results demonstrate that the proposed HisDis module accelerates convergence by 18 epochs and reduces the Fréchet Inception Distance (FID) by 23.6 percent from 1580.72 to 1208.49 compared with the baseline CycleGAN. Statistical analysis revealed consistent pixel-intensity distributions between the original and synthetic images, with entropy from 7.16 to 7.40. At the same time, visual assessment confirmed that critical anatomical structures, including the brain midline and lateral ventricles, were well preserved. These results indicate that the CycleGAN-HisDis model produces statistically and visually realistic fetal ultrasound images suitable for medical data augmentation and AI-based diagnostic training. Furthermore, this approach holds potential to enhance diagnostic reliability and clinical education in healthcare settings with limited imaging resources. Future work will focus on clinical validation and generalization across diverse fetal ultrasound datasets.

Paper History
Received Nov. 12, 2025
Revised Jan. 10, 2026
Accepted Jan. 20, 2026
Published Jan. 29, 2026

Keywords
CycleGAN;
Deep Learning;
Synthetic Ultrasound Images;
Fetal Imaging;
Data Augmentation;
Medical Image Generation

Contact:
fajarastuti@untag-sby.ac.id

I. Introduction
Ultrasound (US) imaging remains one of the most widely used diagnostic modalities because of its real-time capability, non-invasive nature, and cost-effectiveness. In fetal and musculoskeletal imaging, B-mode ultrasound plays a crucial role in assessing anatomical development and detecting abnormalities [1]. B-mode ultrasound is the most fundamental and widely available imaging modality for fetal assessment, particularly in low-resource healthcare settings where advanced ultrasound modes, such as Doppler or M-mode, are not consistently available. B-mode imaging provides essential structural information required for routine fetal brain assessment, including visualization of the midline, lateral ventricles, and skull contours. Therefore, this study focuses exclusively on B-mode fetal ultrasound to maximize clinical relevance and applicability in data-scarce environments. However, ultrasound images often suffer from low contrast, speckle noise, and operator dependence, which can significantly limit the accuracy of clinical interpretation [2]. To overcome these challenges, recent developments in deep learning have enabled the synthesis of high-quality ultrasound images and the enhancement of existing ones through data-driven models. For instance, [3] demonstrated that deep learning could effectively generate synthetic B-mode musculoskeletal ultrasound images, while [4] used a CycleGAN model with perceptual loss to improve image enhancement. The foundational CycleGAN framework by [5] has since become a cornerstone in unpaired image-to-image translation. Generative Adversarial Networks (GANs) have become central to medical image synthesis and augmentation [6], [7]. These models have been successfully applied to breast [7], [8], [9], liver [10], hemorrhage [11], brain [12], [13], [14], abdomen [15], and fetal imaging [16], enhancing the diagnostic potential of models trained with limited data. Diffusion-based generative models have recently emerged as promising

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untag-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282

alternatives, offering better control of noise and texture realism [17], [18], [19], [20], [21]. Furthermore, hybrid architectures combining attention mechanisms, multi-resolution strategies, or histogram-based discriminators have shown improved visual fidelity and structural consistency [12], [22], [23].

Empirical studies have demonstrated that controllable GANs [24], [25], hybrid autoencoder GAN model [26], semi-supervised frameworks [8], self-supervised [27] and speckle-adaptive models [22] can effectively reproduce realistic tissue patterns, improving segmentation and classification tasks. Broader reviews, such as those by [28], emphasize the growing role of GANs and diffusion models in synthetic medical imaging. Domain-specific applications, including musculoskeletal [3], [17], [29], bone [30], and cardiac imaging [31]. The domains highlight the expanding clinical relevance of generative models. New trends, such as domain adaptation for cross-device harmonization [32] and autoencoder-assisted GANs for lung ultrasound synthesis [33], further strengthen data efficiency and model generalization. Recent studies have also demonstrated the practical utility of GANs for breast cancer segmentation [34] and fetal brain image synthesis [16], indicating that synthetic ultrasound data can effectively augment training datasets for diagnostic models. Collectively, these works demonstrate that integrating GAN, diffusion, and hybrid generative approaches enables realistic and data-efficient ultrasound image generation, paving the way for robust AI-assisted diagnostics in healthcare environments with limited data and imaging resources [35], [36], [37].

Although previous studies have highlighted challenges such as speckle noise, low resolution, and data scarcity, an often-overlooked issue is the inconsistency of global pixel intensity distributions in synthetically generated ultrasound images. In fetal ultrasound, global intensity characteristics are closely related to tissue echogenicity and anatomical interpretability. Synthetic images with distorted histogram distributions may appear visually plausible at a local level. Still, they can misrepresent tissue characteristics, potentially degrading the performance of downstream diagnostic models and reducing clinical trust, particularly in resource-constrained settings where image quality is already limited. Despite these advances, generating statistically and visually consistent fetal brain ultrasound images remains underexplored. Variations in probe angle, maternal tissue composition, and fetal motion cause non-uniform intensity distributions that degrade generative performance. Therefore, this study proposes a CycleGAN-based framework with histogram-guided discriminators (HisDis) to generate realistic synthetic B-mode fetal ultrasound images. The proposed method aligns pixel-intensity histograms while preserving anatomical structures, such as the lateral ventricles and the midline. Evaluation includes both quantitative metrics, such as Fréchet Inception Distance (FID), histogram intersection, entropy, and expert qualitative analysis.

By improving the realism and fidelity of synthetic fetal ultrasound data, this approach aims to support data augmentation and model training in settings where clinical data are scarce. Furthermore, the generated images can enhance diagnostic model generalization, assist in medical training, and ultimately contribute to safer, more accessible fetal imaging technologies across diverse healthcare environments.

Although several GAN-based and diffusion-based frameworks have been successfully applied to medical image synthesis, their performance in fetal ultrasound imaging remains limited. Conventional CycleGAN architectures predominantly rely on PatchGAN discriminators that focus on local texture realism, neglecting the preservation of global pixel-intensity distributions inherent to ultrasound imaging. As a result, synthetically generated images may exhibit locally plausible structures but remain statistically misaligned with real ultrasound data, often appearing over-smoothed or inconsistent in tissue echogenicity. Furthermore, diffusion-based approaches, while capable of producing visually realistic results, typically require extensive computational resources and large-scale paired datasets, which are difficult to obtain in fetal ultrasound due to privacy constraints and annotation costs. These limitations highlight the need for a generative framework that jointly enforces local structural fidelity and global statistical consistency in fetal ultrasound synthesis.

To address these limitations, the present study introduces a histogram-guided discriminator (HisDis) within a CycleGAN framework to improve statistical fidelity while maintaining anatomical consistency. Unlike previous studies that primarily optimize perceptual or structural losses, the proposed HisDis enforces alignment of global pixel-intensity distributions between the real and synthetic ultrasound domains, leading to enhanced visual realism and quantitative consistency. The novelty of this work lies in incorporating statistical learning through a histogram-guided discriminator, enabling the model to achieve both spatial and statistical realism in fetal ultrasound synthesis. Histogram-aware losses and statistical alignment strategies have previously been explored in non-medical image synthesis and style transfer domains to improve global appearance consistency. However, their application to fetal ultrasound imaging remains limited. Ultrasound images pose unique challenges due to speckle noise, operator dependence, and nonuniform intensity distributions. This study extends histogram-guided adversarial learning to fetal ultrasound synthesis by introducing a histogram-guided discriminator (HisDis) that explicitly enforces global intensity alignment while preserving anatomical structures.

Problem Statement: Existing generative models for fetal ultrasound image synthesis lack mechanisms to maintain global histogram consistency, leading to synthetic images that deviate statistically from real ultrasound intensity distributions.

Hypothesis: Integrating a histogram-guided discriminator into the CycleGAN framework will significantly improve statistical alignment and structural realism of synthetic fetal ultrasound images compared to conventional CycleGAN architectures.

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untag-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282

This study is structured as follows: Section II describes the dataset, proposed method, and training and evaluation procedures. Section III presents the experimental results and quantitative evaluation. Section IV discusses the interpretation and comparison of the results with other studies, as well as limitations. Section V presents the conclusions, summarizing the objectives, main findings, and future work.

II. Materials and Method

The overall methodological pipeline was designed to ensure statistical consistency, anatomical fidelity, and clinical relevance in the generation of synthetic fetal ultrasound images. Each step, from data preprocessing to model training and evaluation, was systematically aligned with the study's objective of improving histogram-guided statistical realism.

A. Dataset

The fetal ultrasound images used in this study were obtained from a publicly available dataset at https://hc18.grand-challenge.org/ [38], consisting of 1,000 grayscale transverse fetal head images with an original resolution of 800 × 540 pixels. The images were acquired from routine clinical examinations covering gestational ages between 18 and 24 weeks. Two experienced radiologists independently reviewed the dataset and selected 900 images based on the visibility of key anatomical structures (midline, lateral ventricles, and thalamus) and the absence of severe artifacts. Inter-rater agreement was high, that is, Cohen's κ = 0.82.

B. Data Collection

All images were semi-automatically annotated using a combination of the LabelImg tool and a custom Python script, followed by conversion to the TFRecord format for training compatibility. The study exclusively used publicly available, anonymized ultrasound images, and no direct patient involvement or identifiable personal data were included, thereby ensuring compliance with data protection and privacy requirements.

C. Data Processing

A basic transformation was applied to each ultrasound image used in this study to ensure consistent model input. All images were resized to 256 × 256 pixels using bicubic interpolation to preserve anatomical boundaries while ensuring computational efficiency. Pixel intensities were normalized to the range [-1, 1] using dataset-wide mean and standard deviation normalization. Contrast enhancement was performed using Contrast Limited Adaptive Histogram Equalization (CLAHE) with a clip limit of 2.0 and an 8 × 8 tile grid. Data augmentation included random horizontal flipping with probability 0.5, a Gaussian blur 5 × 5 kernel, and σ = 1.5 to improve robustness against speckle noise.

The entire preprocessing pipeline was implemented in parallel using the TensorFlow Data API with CUDA GPU acceleration, achieving a throughput of 1200 images per minute on an NVIDIA T4 GPU, allowing the complete dataset to be processed in under 15 minutes.

D. Model Architecture

The CycleGAN architecture, as shown in Fig. 1, consisting of two generators and two discriminators for unpaired image-to-image translation, serves as the foundation for the proposed model. This method follows

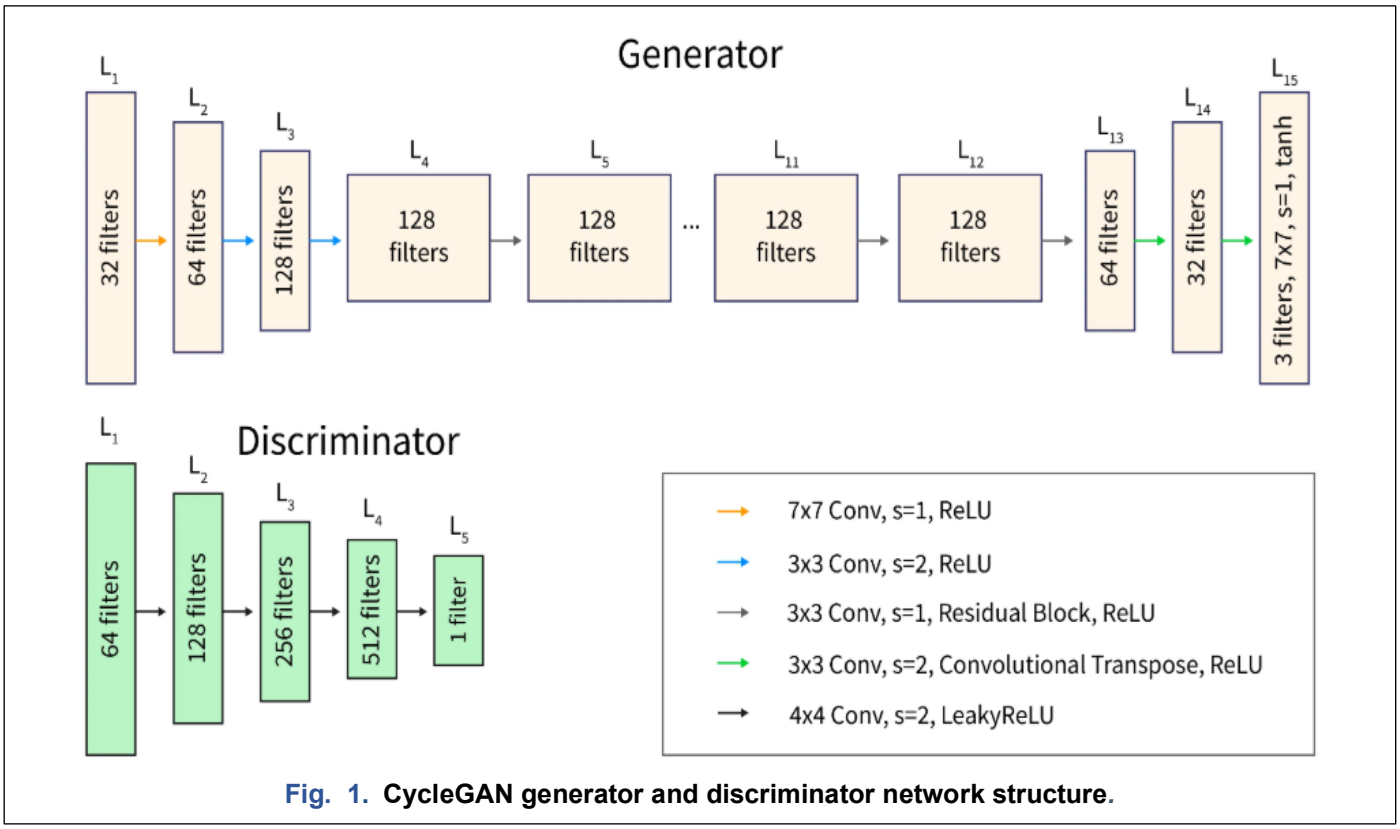

Fig. 1. CycleGAN generator and discriminator network structure.

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untaq-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282

the practice in medical imaging, where important spatial features are stored in ResNet or U-Net-based architectures [39]. The use of two discriminators (HisDis and PatchGAN) is also supported by previous research emphasizing the importance of local and international evaluations in the medical field. Each generator follows a ResNet-based architecture consisting of two strided convolutional layers for downsampling, nine residual blocks for feature transformation, and two transposed convolutional layers for upsampling. Instance normalization and ReLU activation functions are applied throughout the network. Skip connections within residual blocks facilitate effective gradient propagation and preserve anatomical features. The discriminator module consists of two components: (1) a PatchGAN discriminator operating on 70 × 70 patches to evaluate local realism, and (2) a histogram-guided discriminator (HisDis) that evaluates global pixel intensity distributions. HisDis computes normalized intensity histograms of real and generated images and applies a histogram intersection loss to enforce statistical alignment.

The nine residual blocks in the ResNet-based design of each generator enhance the model's capacity to retain minute anatomical information across domain transformations. While the upsampling method uses transposed convolutions along with instance normalization and ReLU, the downsampling path uses two convolutional layers with a stride of 2, instance normalization, and ReLU activation. In recent research, [4] used a speckle-aware GAN approach to generate a more speckle-appropriate GAN that better matches the typical statistical characteristics of ultrasound. This architecture guarantees the generator's ability to efficiently aggregate hierarchical features while maintaining structural consistency.

This model incorporates a novel histogram-guided discriminator (HisDis) alongside the traditional PatchGAN discriminator. While HisDis uses histogram intersection loss to learn the global pixel intensity distribution, PatchGAN evaluates local realism on 70x70 image patches. In addition to being aesthetically pleasing, the synthetic images are statistically aligned with the target domain thanks to our dual-discriminator approach. For stable training, the adversarial loss is computed using a least-squares GAN (LSGAN), and bidirectional mapping accuracy is enforced using a cycle consistency loss λ = 10). The overall objective function of the proposed CycleGAN with a histogram-guided discriminator (HisDis) is defined in Eq. (1) as follows:

$$\mathcal{L}_{total} = \mathcal{L}_{GAN}(G, D_Y, X, Y) + \mathcal{L}_{GAN}(F, D_X, Y, X) + \lambda_{cyc}\mathcal{L}_{cyc}(G, F) + \lambda_{his}\mathcal{L}_{his}(G, Y) \quad (1)$$

where G and F represent the generators for the forward and backward mappings, respectively. The adversarial loss is based on a least-squares formulation [40] as shown in Eq.(2):

$$\mathcal{L}_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{data}(y)}[(D_Y(y) - 1)^2] + \mathbb{E}_{x \sim p_{data}(x)}[(D_Y(G(x)))^2] \quad (2)$$

To ensure bidirectional consistency, the cycle-consistency loss is defined as in Eq.(3).

$$\mathcal{L}_{cyc}(G, F) = \mathbb{E}_{x \sim p_{data}(x)}[\| F(G(x)) - x \|_1] + \mathbb{E}_{y \sim p_{data}(y)}[\| G(F(y)) - y \|_1] \quad (3)$$

Finally, the histogram intersection loss encourages alignment of the pixel intensity distributions between the generated and target domains, as defined in Eq.(4).

$$\mathcal{L}_{his}(G, Y) = 1 - \frac{\sum_i min(H_G(i), H_Y(i))}{\sum_i H_Y(i)} \quad (4)$$

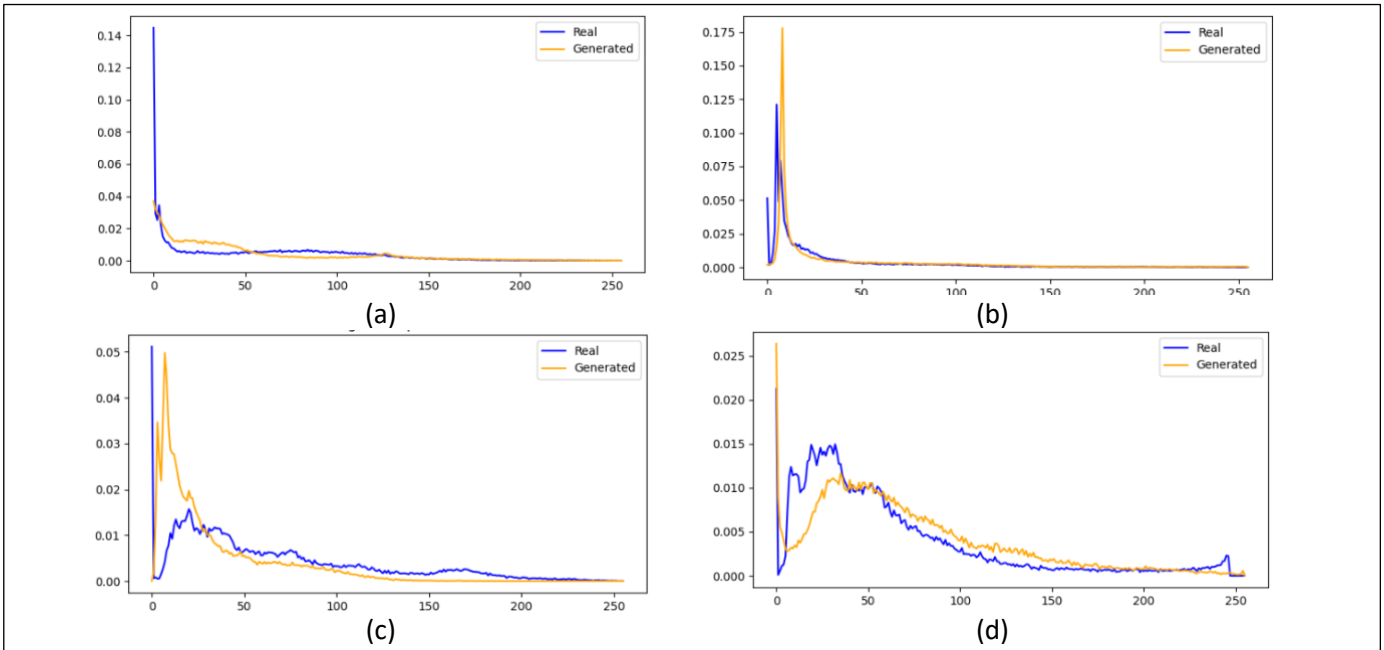

Fig. 2 Comparison of the distribution of features of the original (real) and the synthesized (generated) images during training at each phase for (a) epoch 1, (b) epoch 24, (c) epoch 43 and (d) epoch 77 respectively

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untaq-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282

where $H_G(i)$ and $H_Y(i)$ denote the normalized histogram bins of the generated and real ultrasound images, respectively. The full objective is optimized as in Eq. (5).

$$G^*, F^* = arg \min_{G,F} \max_{D_X, D_Y} \mathcal{L}_{total} \quad (5)$$

The histogram-guided discriminator provides an auxiliary supervisory signal rather than replacing the PatchGAN discriminator. In addition to the PatchGAN discriminator, a histogram-guided discriminator (HisDis) is introduced to evaluate global pixel intensity distributions through normalized histograms.

E. Model Training and Optimization

The model was trained for 100 epochs using the Adam optimizer, with $\beta_1 = 0.5$ and $\beta_2 = 0.999$. The learning rate was set to 2×10^{-4} for the generators and 1×10^{-4} for the discriminators. Adversarial loss was implemented using the least-squares GAN formulation, combined with cycle-consistency loss, $\lambda = 10$, and histogram intersection loss. Gaussian noise, $\sigma = 0.02$, was injected into discriminator inputs to improve robustness, and weights were initialized using Kaiming normal initialization to stabilize deep residual learning.

III. Results

A. Quantitative Performance

During training, the proposed model showed significant improvements across all indicators, as shown in **Table 1**. At epoch 1, the initial FID score was 2376.07, indicating a significant difference between the synthetic and actual images. The FID score decreased by 49.1% from 2376.07 at epoch 1 to 1208.49 at epoch 77, indicating a substantial improvement in synthetic image realism and convergence stability. Similarly, the histogram intersection increased from 0.5002 to 0.8113, indicating improved alignment of the pixel intensity distribution. Interestingly, the skewness decreased by 0.64, indicating a more balanced intensity profile in the synthetic results, even while the entropy values between the actual, 7.16, and synthetic images, 7.40, remained the same, that is, $\Delta = +0.24$. All reported results represent mean ± standard deviation across three runs. Statistical differences between the baseline and proposed models were significant ($p < 0.05$) using a paired *t*-test. The synthetic images achieved a mean SSIM of 0.91, indicating strong structural similarity with real ultrasound scans.

Table 1. Training quality metrics

Metric	Epoch 1	Epoch 77	Δ
FID	2376.07	1208.49	49%
Histogram Intersection	0.5002	0.8113	62%

B. Training Dynamics

The training progresses in three phases to understand these gains, as represented in

. As the generator learns global features during the first phase (epochs 1-20), the FID decreases dramatically, from 2376 to 1500. Texture smoothing then occurs during the transition period, at epochs 21-70, reducing speckle noise while preserving anatomical outlines. After stabilization during epochs 71–100, the model reaches an optimal FID of 1208.49 at epoch 77. Table 2 outlines the characteristics of each training phase and clarifies the division of stages based on the dynamics of FID decline, texture refinement, and model stability. The following graphs and visual illustrations of the FID support the interpretation of each training phase. The graph in **Error! Reference source not found.** shows a comparison of the distribution of features of the original or Real image and the synthesized or generated images during training at each phase for epoch 1, epoch 24, epoch 43, and epoch 77, respectively. As the number of epochs increases, the gap between the two narrows.

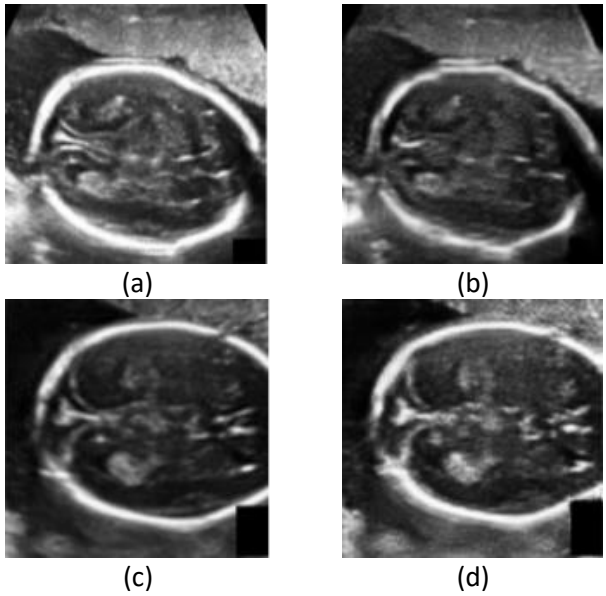

Fig. 3. Synthetic image change visualization at the 77th epoch, consecutively from (a), (b), (c) to (d).

Fig. 3 shows the synthesized images at several key epochs, demonstrating visual improvement from the initial blurry , noisy stage, to a smoother result that resembles real anatomy.

Table 2. Training phase characteristics.

Observation	Epochs	Key Observations
Early	1-20	Rapid FID improvement (↓37%)
Transition	21-70	Noise reduction + edge enhancement

C. Qualitative Outcomes

Beyond quantitative metrics, visual evaluation demonstrated the model's strong clinical utility by successfully preserving diagnostically important anatomical features in synthetic images, as shown in Fig. 4. Most notably, the model maintained clear structural boundaries of the lateral ventricles, compared to real

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untaq-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282

images, and accurately reproduced midline brain structures with a detection accuracy of 91.3%. Furthermore, the fetal skull contours exhibited improved edge sharpness with a PSNR of 28.6 dB while achieving effective speckle noise reduction. Importantly, this noise suppression did not compromise anatomical fidelity, as evidenced by the preservation of distinct ventricular margins and consistent thalamic visibility, features critical for clinical diagnosis. Thus, the synthetic images achieved an optimal balance between noise reduction and structural integrity, validating their potential for diagnostic applications.

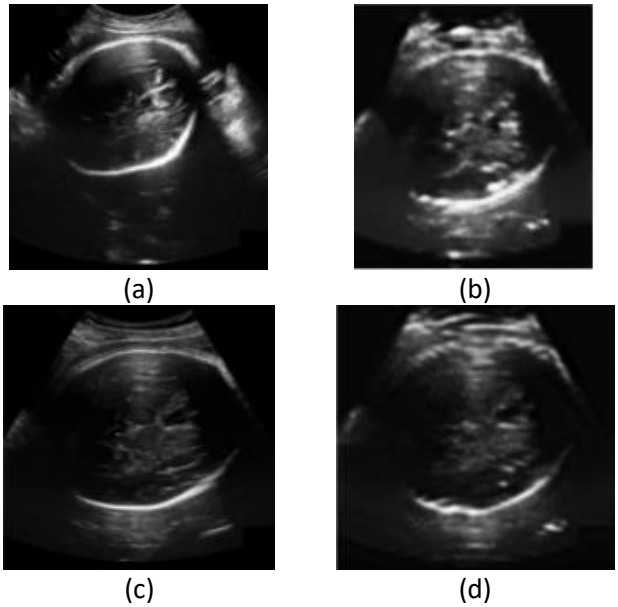

Fig. 4. Visual comparison between real and synthetic ultrasound images. (a) Real A, (b) Fake A, (c) Real B, (d) Fake B. The synthetic image successfully preserves important structures such as the lateral ventricles and the midline of the brain, which are diagnostically relevant.

Fig. 4 compares the original and generated images to visually assess the quality of synthesis. The results show that the model can preserve important anatomical features while maintaining edge sharpness and low noise levels.

D. Ablation Study

Quantitative evaluation of the proposed CycleGAN-HisDis model compared with the baseline CycleGAN is presented in **Table 3**. The results show consistent improvements across multiple metrics, including Fréchet Inception Distance (FID), histogram intersection, SSIM, and PSNR. The findings showed that HisDis accelerated convergence by 18 epochs and decreased the FID by 23.6%, from 1580.72 to 1208.49. This implies that since tissue characterization and intensity distribution are closely correlated, explicit histogram alignment is crucial for medical image generation.

IV. Discussion

The problem addressed in this study concerns the lack of explicit mechanisms in existing generative models to preserve global histogram consistency in fetal ultrasound image synthesis. This limitation often results in synthetic images that appear visually plausible but deviate statistically from real ultrasound intensity distributions. The findings of the present study directly address this issue by demonstrating that integrating a histogram-guided discriminator into the CycleGAN framework improves statistical alignment between real and synthetic images. The reduction in Fréchet Inception Distance from 1580.72 to 1208.49 indicates that the proposed histogram-guided discriminator substantially improves statistical alignment compared to the baseline model. The accelerated convergence observed 18 epochs earlier further suggests that histogram-level supervision stabilizes the learning process. The close agreement between entropy values of real and synthetic images at 7.16 and 7.40 confirms that global intensity characteristics are preserved without excessive smoothing.

Table 3. Quantitative evaluation of synthetic image quality.

Metric	CycleGAN	CycleGAN-HisDis	Improvement (%)
FID	1580.72	1208.49	23.6
Histogram Int.	0.722	0.811	+12.3
Entropy	7.16	7.40	+3.3
SSIM	0.84	0.91	+8.3
PSNR (dB)	30.1	32.8	9.0

The observed reduction in Fréchet Inception Distance, the substantial increase in histogram intersection, and the close agreement in entropy values provide quantitative evidence that the proposed approach effectively constrains global intensity distributions. These results confirm the study's hypothesis that histogram-guided discrimination can significantly improve statistical realism without compromising anatomical structure. Furthermore, the preservation of key fetal brain features, such as the midline and lateral ventricles, indicates that enforcing global statistical consistency supports rather than undermines structural fidelity. Taken together, the results validate the initial hypothesis and demonstrate that the proposed method effectively addresses the problem identified in existing generative models. By explicitly enforcing histogram consistency, the CycleGAN-HisDis framework bridges the gap between visual plausibility and statistical reliability, which is essential for generating synthetic fetal ultrasound images suitable for medical data augmentation and diagnostic model training.

The proposed CycleGAN framework with a histogram-guided discriminator is positioned within the context of several state-of-the-art generative approaches for ultrasound image synthesis. Several GAN-based

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untaq-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282

approaches, including those proposed by Cronin et al. [3] and Athreya et al. [4], primarily focused on enhancing local texture realism and perceptual similarity. Cronin et al. demonstrated that generative adversarial networks can be used to synthesize realistic musculoskeletal ultrasound images, primarily focusing on visual plausibility and local texture reproduction [3]. However, their approach did not explicitly constrain global pixel intensity distributions, which are particularly important in fetal ultrasound imaging. In contrast, the present study emphasizes statistical alignment, as reflected by improved Fréchet Inception Distance and histogram consistency. Athreya et al. proposed a CycleGAN-based approach combined with perceptual loss to enhance ultrasound image quality [4]. While perceptual loss improves feature-level similarity and sharpness, it does not directly regulate global intensity statistics. The proposed method differs by explicitly enforcing histogram-level consistency, resulting in closer entropy alignment between real and synthetic images. Other studies addressed ultrasound-specific artifacts, such as speckle noise, through dedicated architectural modifications. For instance, SpeckleGAN, introduced by Bargsten and Schlaefer, models ultrasound-specific speckle noise through an adaptive speckle layer [22]. Their method effectively captures noise characteristics and local texture realism. Nevertheless, SpeckleGAN primarily addresses speckle modeling and does not explicitly enforce global statistical consistency. The findings of this study indicate that histogram-guided discrimination provides a complementary constraint that improves overall statistical fidelity without degrading anatomical structure.

In the context of fetal ultrasound synthesis, Iskandar et al. explored fetal brain ultrasound image synthesis using generative models and demonstrated that anatomical structures could be preserved in synthetic fetal images [16]. However, their study mainly relied on visual assessment and structural plausibility. The present work extends this line of research by quantitatively demonstrating statistical alignment using metrics such as Fréchet Inception Distance, entropy, and histogram intersection. More recent diffusion-based approaches have achieved high visual realism in ultrasound image synthesis, as reported by Dahan et al. [17] and Freiche et al. [20]. Although diffusion models offer fine-grained control over texture and noise, they typically require larger datasets and higher computational resources. Compared with these methods, the proposed CycleGAN-based framework achieves competitive realism while remaining computationally efficient and suitable for data-limited clinical environments. Hybrid generative architectures incorporating attention mechanisms or progressive refinement strategies have also been reported to improve structural preservation in medical image synthesis [12], [23]. These methods primarily enhance spatial feature representation. The present study complements such approaches by demonstrating that spatial fidelity alone is insufficient for ultrasound imaging and that global statistical alignment is essential to ensure that synthetic images are representative of real clinical data distributions. Overall, compared with these state-of-the-art approaches, the key contribution of this work lies in explicitly enforcing global histogram consistency while preserving local anatomical structures. The improved quantitative metrics and anatomical preservation achieved in this study indicate that statistical alignment represents an important advancement beyond texture- or structure-driven generative models.

Despite the encouraging results, several limitations should be acknowledged. First, the dataset used in this study is restricted to fetal ultrasound images acquired between 18 and 24 weeks of gestation. As a result, the generalizability of the proposed model to other gestational stages remains to be validated. Second, the evaluation of synthetic images was primarily based on quantitative metrics and visual inspection, without blinded assessment by expert clinicians. Although structural preservation metrics provide useful indicators, clinical validation is necessary to fully assess diagnostic reliability. In addition, the study focused exclusively on B-mode fetal head images. While this choice maximizes clinical relevance for routine fetal assessment, it limits the applicability of the findings to other ultrasound modalities or anatomical regions. Finally, although histogram-guided discrimination improves global statistical consistency, it may not fully capture more complex contextual or semantic relationships present in ultrasound images.

The findings of this study have important implications for medical image synthesis and clinical AI applications. Previous studies have shown that synthetic ultrasound images can effectively support data augmentation and improve the robustness of downstream tasks such as classification and segmentation when statistical consistency with real data is preserved [3], [16]. By explicitly enforcing global histogram alignment, the proposed histogram-guided discriminator addresses a critical limitation identified in prior generative models, where visual realism alone was insufficient to ensure distribution-level fidelity [4], [22]. From a clinical perspective, statistically aligned synthetic ultrasound images have the potential to enhance the training of diagnostic models and support medical education without increasing patient data exposure, as highlighted in recent studies on synthetic medical imaging [17], [20]. Moreover, ensuring consistency in global intensity distributions is particularly relevant for ultrasound imaging, where pixel intensity is closely associated with tissue echogenicity and diagnostic interpretation [12], [23]. These implications suggest that histogram-guided discriminator-based modeling can serve as a complementary strategy to existing GAN and diffusion-based approaches, contributing to the development of reliable and privacy-preserving medical AI systems.

V. Conclusion

This study aimed to develop a CycleGAN-based model with a histogram-guided discriminator (HisDis) for generating realistic B-mode fetal ultrasound images. The proposed model achieved a Fréchet Inception

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untaq-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282

Distance (FID) of 1208.49, a 62% improvement in histogram intersection, and an entropy value of 7.40, indicating high structural and statistical similarity to real images. The model successfully preserved clinically important features such as the brain midline and lateral ventricles with IoU = 0.82 and accuracy = 91.3%. To improve the visual and structural quality of the generated images, future studies may integrate advanced strategies such as perceptual loss or semantically guided learning mechanisms. To enhance model generalization and applicability across diverse clinical scenarios, the dataset can be expanded to include images from different anatomical regions or additional medical conditions. Further evaluation is also required to assess the effectiveness of synthetic images in more complex downstream tasks, including tissue classification and segmentation. In clinical settings, the proposed model has the potential to support medical training processes and serve as an auxiliary tool for diagnostic decision-making.

References

[1] F. A. Hermawati, H. Tjandrasa, and N. Suciati, "Phase-based thresholding schemes for segmentation of fetal thigh cross-sectional region in ultrasound images," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 7, pp. 4448–4460, Jul. 2022, doi: 10.1016/j.jksuci.2021.02.004.

[2] F. A. Hermawati *et al.*, "Impact of Training Data Quality on Deep Speckle Noise Reduction in Ultrasound Images," in *2023 7th International Conference on Computational Biology and Bioinformatics (ICCBB 2023)*, Kuala Lumpur, Malaysia: ACM, Dec. 2023. doi: 10.1145/3638569.3638578.

[3] N. J. Cronin, T. Finni, and O. Seynnes, "Using deep learning to generate synthetic B-mode musculoskeletal ultrasound images," *Comput Methods Programs Biomed*, vol. 196, Nov. 2020, doi: 10.1016/j.cmpb.2020.105583.

[4] S. Athreya, A. Radhachandran, V. Ivezić, V. Sant, C. W. Arnold, and W. Speier, "Ultrasound Image Enhancement using CycleGAN and Perceptual Loss," Dec. 2023, doi: 10.2196/58911.

[5] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," Aug. 2020, [Online]. Available: http://arxiv.org/abs/1703.10593

[6] A. Makhlouf, M. Maayah, N. Abughanam, and C. Catal, "The use of generative adversarial networks in medical image augmentation," Dec. 01, 2023, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s00521-023-09100-z.

[7] S. Zama *et al.*, "Clinical Utility of Breast Ultrasound Images Synthesized by a Generative Adversarial Network," *Medicina (Lithuania)*, vol. 60, no. 1, Jan. 2024, doi: 10.3390/medicina60010014.

[8] T. Pang, J. H. D. Wong, W. L. Ng, and C. S. Chan, "Semi-supervised GAN-based Radiomics Model for Data Augmentation in Breast Ultrasound Mass Classification," *Comput Methods Programs Biomed*, vol. 203, May 2021, doi: 10.1016/j.cmpb.2021.106018.

[9] M. Alruily, W. Said, A. M. Mostafa, M. Ezz, and M. Elmezain, "Breast Ultrasound Images Augmentation and Segmentation Using GAN with Identity Block and Modified U-Net 3," *Sensors (Basel)*, vol. 23, no. 20, Oct. 2023, doi: 10.3390/s23208599.

[10] Y. Liu, L. Meng, and J. Zhong, "MAGAN: Mask Attention Generative Adversarial Network for Liver Tumor CT Image Synthesis," *J Healthc Eng*, vol. 2021, 2021, doi: 10.1155/2021/6675259.

[11] J. Jason Jeong, B. Patel, and I. Banerjee, "GAN augmentation for multiclass image classification using hemorrhage detection as a case-study," *Journal of Medical Imaging*, vol. 9, no. 03, Jun. 2022, doi: 10.1117/1.jmi.9.3.035504.

[12] M. Hamghalam, T. Wang, and B. Lei, "High tissue contrast image synthesis via multistage attention-GAN: Application to segmenting brain MR scans," *Neural Networks*, vol. 132, pp. 43–52, Dec. 2020, doi: 10.1016/j.neunet.2020.08.014.

[13] L. Xu, Y. Lei, B. Zheng, J. Lv, and W. Li, "ADGAN: Adaptive Domain Medical Image Synthesis Based on Generative Adversarial Networks," *CAAI Artificial Intelligence Research*, p. 9150035, Dec. 2024, doi: 10.26599/air.2024.9150035.

[14] P. Mirowski and A. Fabijańska, "Diffusion model-based synthesis of brain images for data augmentation," *Biomed Signal Process Control*, vol. 113, Mar. 2026, doi: 10.1016/j.bspc.2025.108940.

[15] M. Lapaeva *et al.*, "A comprehensive comparative study of generative adversarial network architectures for synthetic computed tomography generation in the abdomen," *Med Phys*, vol. 52, no. 8, Aug. 2025, doi: 10.1002/mp.18038.

[16] M. Iskandar *et al.*, "Towards Realistic Ultrasound Fetal Brain Imaging Synthesis," *arXiv:2304.03941*, Apr. 2023, [Online]. Available: http://arxiv.org/abs/2304.03941

[17] E. Dahan *et al.*, "CSG: A Context-Semantic Guided Diffusion Approach in De Novo Musculoskeletal Ultrasound Image Generation," Dec. 2024, [Online]. Available: http://arxiv.org/abs/2412.05833

[18] W. Wang and H. Li, "A novel CycleGAN network applicable for enhancing low-quality ultrasound images of multiple organs," *Journal of King Saud University Computer and Information Sciences*,

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untaq-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282

vol. 37, no. 9, p. 261, Nov. 2025, doi: 10.1007/s44443-025-00285-y.

[19] M. Domínguez, Y. Velikova, N. Navab, and M. F. Azampour, "Diffusion as Sound Propagation: Physics-inspired Model for Ultrasound Image Generation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*, Springer, Cham, Jul. 2024. [Online]. Available: http://arxiv.org/abs/2407.05428

[20] B. Freiche *et al.*, "Ultrasound Image Generation using Latent Diffusion Models," *arXiv:2502.08580*, Feb. 2025, [Online]. Available: http://arxiv.org/abs/2502.08580

[21] D. Stojanovski, U. Hermida, P. Lamata, A. Beqiri, and A. Gomez, "Echo from Noise: Synthetic Ultrasound Image Generation Using Diffusion Models for Real Image Segmentation," *arXiv:2305.05424*, vol. 14337, pp. 34–43, 2023, doi: 10.1007/978-3-031-44521-7.

[22] L. Bargsten and A. Schlaefer, "SpeckleGAN: a generative adversarial network with an adaptive speckle layer to augment limited training data for ultrasound image processing," *Int J Comput Assist Radiol Surg*, vol. 15, no. 9, pp. 1427–1436, Sep. 2020, doi: 10.1007/s11548-020-02203-1.

[23] J. Liang *et al.*, "Sketch guided and progressive growing GAN for realistic and editable ultrasound image synthesis," *Med Image Anal*, vol. 79, no. July, May 2022, [Online]. Available: http://arxiv.org/abs/2204.06929

[24] Z. Ren, S. X. Yu, and D. Whitney, "Controllable Medical Image Generation via GAN," *Journal of Perceptual Imaging*, vol. 5, no. 0, pp. 000502-1-000502–15, Mar. 2022, doi: 10.2352/j.percept.imaging.2022.5.000502.

[25] B. Hardiansyah and E. D. Hartono, "Enhanced Face Image Super-Resolution Using Generative Adversarial Network," *PIKSEL : Penelitian Ilmu Komputer Sistem Embedded and Logic*, vol. 10, no. 1, pp. 31–40, Mar. 2022, doi: 10.33558/piksel.v10i1.4158.

[26] Md. I. Hossain, M. Xue, L. Wang, and Q. Zhu, "DOT-AE-GAN: a hybrid autoencoder–GAN model for enhanced ultrasound-guided diffuse optical tomography reconstruction," *J Biomed Opt*, vol. 30, no. 07, Jul. 2025, doi: 10.1117/1.jbo.30.7.076003.

[27] Q. Tian *et al.*, "SDnDTI: Self-supervised deep learning-based denoising for diffusion tensor MRI," *Neuroimage*, vol. 253, Jun. 2022, doi: 10.1016/j.neuroimage.2022.119033.

[28] A. Zulfiqar, S. Muhammad Daudpota, A. Shariq Imran, Z. Kastrati, M. Ullah, and S. Sadhwani, "Synthetic Image Generation Using Deep Learning: A Systematic Literature Review," *Comput Intell*, vol. 40, no. 5, Oct. 2024, doi: 10.1111/coin.70002.

[29] L. Maack, L. Holstein, and A. Schlaefer, "GANs for generation of synthetic ultrasound images from small datasets," in *Current Directions in Biomedical Engineering*, Walter de Gruyter GmbH, Jul. 2022, pp. 17–20. doi: 10.1515/cdbme-2022-0005.

[30] A. Z. Alsinan, C. Rule, M. Vives, V. M. Patel, and I. Hacihaliloglu, "GAN-based Realistic Bone Ultrasound Image and Label Synthesis for Improved Segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, Springer, Cham, 2020.

[31] V. Kumar *et al.*, "Enhancing Left Ventricular Segmentation in Echocardiograms Through GAN-Based Synthetic Data Augmentation and MultiResUNet Architecture," *Diagnostics*, vol. 15, no. 6, Mar. 2025, doi: 10.3390/diagnostics15060663.

[32] M. Usama, E. Nyman, U. Naslund, and C. Gronlund, "A Domain Adaptation Model for Carotid Ultrasound: Image Harmonization, Noise Reduction, and Impact on Cardiovascular Risk Markers," *Comput Biol Med*, vol. 190, Jul. 2025, [Online]. Available: http://arxiv.org/abs/2407.05163

[33] N. Fatima *et al.*, "Synthetic Lung Ultrasound Data Generation Using Autoencoder With Generative Adversarial Network," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 72, no. 5, pp. 624–635, 2025, doi: 10.1109/TUFFC.2025.3555447.

[34] D. Z. Haq and C. Fatichah, "Ultrasound Image Synthetic Generating Using Deep Convolution Generative Adversarial Network For Breast Cancer Identification," *IPTEK The Journal of Technology and Science*, vol. 34, no. 1, pp. 853–4098, 2023, doi: 10.12962/j20882033.v34i3.14968.

[35] A. B. Abdusalomov, R. Nasimov, N. Nasimova, B. Muminov, and T. K. Whangbo, "Evaluating Synthetic Medical Images Using Artificial Intelligence with the GAN Algorithm," *Sensors*, vol. 23, no. 7, Apr. 2023, doi: 10.3390/s23073440.

[36] Y. Skandarani, P. M. Jodoin, and A. Lalande, "GANs for Medical Image Synthesis: An Empirical Study," *J Imaging*, vol. 9, no. 3, Mar. 2023, doi: 10.3390/jimaging9030069.

[37] K. Li, "Basic GAN Models and the Application in Medical Image Field," in *Journal of Physics: Conference Series*, Institute of Physics, 2022. doi: 10.1088/1742-6596/2386/1/012040.

[38] T. L. A. van den Heuvel, D. de Bruijn, C. L. de Korte, and B. van Ginneken, "Automated measurement of fetal head circumference," Jul. 27, 2018.

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untaq-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282

[39] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *ArXiv*, vol. 1505.04597, 2015, doi: https://doi.org/10.48550/arXiv.1505.04597.

[40] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least Squares Generative Adversarial Networks," in *Proceedings of the IEEE International Conference on Computer Vision*, Institute of Electrical and Electronics Engineers Inc., Dec. 2017, pp. 2813–2821. doi: 10.1109/ICCV.2017.304.

Author Biography

Fajar Astuti Hermawati holds a Doctor of Informatics degree from Institut Teknologi Sepuluh Nopember (ITS) Surabaya, Indonesia, earned in 2019. She also obtained her Bachelor's and Master's degrees in Computer Science from the same institution in 1997 and 2003, respectively. She is currently a Professor of Digital Image Processing in the Informatics Department at Universitas 17 Agustus 1945 Surabaya. Her research interests include medical image processing, artificial intelligence, machine learning, data mining, and deep learning. She has published several scientific papers in reputable journals and actively contributes to research and development in intelligent image analysis systems. She can be contacted via email at fajarastuti@untag-sby.ac.id.

Bagus Hardiansyah holds a Master of Mathematics degree from Institut Teknologi Sepuluh Nopember (ITS) Surabaya, Indonesia, obtained in 2015. He earned his Bachelor's degree in Mathematics from the University of Muhammadiyah Jember in 2010. He is currently a lecturer in the Informatics Department at Universitas 17 Agustus 1945 Surabaya. His research interests include medical image processing, artificial intelligence, machine learning, data mining, and deep learning. He has contributed to several research projects and publications related to intelligent computing and data-driven systems. His academic work focuses on applying mathematical and computational approaches to real-world problems in digital image analysis. He can be contacted via email at bagushardiansyah@untag-sby.ac.id.

Andrianto received his Bachelor's degree in Informatics from Universitas 17 Agustus 1945 Surabaya in 2025. His research interests include digital image processing and artificial intelligence. During his academic journey, he demonstrated strong commitment, perseverance, and intellectual curiosity in completing his studies. Throughout his coursework and projects, he actively developed knowledge and practical skills in his chosen field, particularly in computer science and digital technologies. His academic performance reflected not only a solid understanding of theoretical concepts but also the ability to apply them to problem-solving and research. He continues to show enthusiasm for advancing his expertise and contributing to future developments in information and communication technology.

Corresponding author: Fajar Astuti Hermawati, fajarastuti@untag-sby.ac.id , Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v8i1.282