https://doi.org/10.53360/2788-7995-2025-4(20)-11

IRTSTI: 28.23.15

M.K. Ryspayeva[1,2*], S.O. Salykova[1]
[1]Akhmet Baitursynuly Kostanay Regional University,
Baytursynov Street 47, 110000, Kostanay, Kazakhstan
[2]Astana IT University,
Mangilik El, block C.1, 010000, Astana, Kazakhstan,
[*]e-mail: marya.rys1@mail.ru

EVALUATION OF STYLEGAN2 AND STYLEGAN3 FOR SYNTHETIC MEDICAL IMAGE GENERATION ON BUSI AND CBIS-DDSM DATASETS

Abstract: *Deep learning from medical images is typically hindered by limited access to images and severe imbalance of classes that reduces the effectiveness of typical machine learning algorithms. Generative adversarial networks can be employed to address such issues by creating natural-appearing synthetic images to complement training sets. In this study, we compare two advanced GAN architectures, StyleGAN2 and StyleGAN3, using two publicly available breast imaging datasets: BUSI (ultrasound, 210 malignant cases) and CBIS-DDSM (mammography, 509 malignant cases). Evaluation was based on Fréchet Inception Distance and Kernel Inception Distance. On BUSI, StyleGAN3 achieved FID = 140.7 and KID = 0.06 at 1000 epochs, whereas StyleGAN2 achieved FID = 259.7 and KID = 0.25. On CBIS-DDSM, StyleGAN3 achieved FID = 90.6 and KID = 0.06, and StyleGAN2 achieved FID = 124.8 and KID = 0.10. These results demonstrate that StyleGAN3 has a tendency to synthesize images that are more natural and diversified under limited dataset conditions, at the cost of increased training times, whereas StyleGAN2 provides similar quality at less expensive computational costs. The results indicate the potential of generating medical images and the trade-off between image quality and efficiency in data augmentation for breast cancer image improvement.*

Key words: *Generative Adversarial Networks; StyleGAN2; StyleGAN3; FID; KID; Breast Ultrasound; Mammography.*

Introduction

Generative adversarial networks (GANs) currently represent the prominent research focus of ongoing information technology research, bringing forth techniques for generating data, enriching it, and running simulations of it in application domains wherein the data scarcity is a fundamental challenge. Their ability to generate realistic, high-dimensional samples makes them particularly attractive for small-sample environments, where traditional machine learning methods suffer from overfitting and poor generalization [1].

One of the application domains wherein overfitting is especially ruthless is medical imaging, where datasets are typically small, imbalanced, and costly to annotate. From the perspective of IT research, medical data constitute the desirable test case to investigate generative models under limited resources and limited data. Traditional GAN realizations, such as Wasserstein GANs (WGAN, WGAN-GP) [2, 3], have demonstrated the potential of image generation by synthesis but typically produce samples of restricted realism and unstable trainability. Latest models, such as those of the StyleGAN family, feature advanced architecture innovations that significantly upgrade realism and stability [4]. StyleGAN2 replaced the AdaIN operator of StyleGAN by weight demodulation, appended path-length regularization, and optimized generator/discriminator blocks, and all of them together established novel quality baselines for synthetic images. Nevertheless, StyleGAN2 remained prone to aliasing artifacts that may affect geometric consistency for longer-term training [5]. StyleGAN3 resolved the issue by adopting alias-free continuous phase transform-based architecture to boost stability and ensure better preservation of structural details [6].

In the above prior work of the first author (Ryspayeva & Salykova, 2025), dataset balancing was studied through a proprietary version of GAN (DGAN-WP-TL), and StyleGAN2 was tested over three medical sets. That work, though, was limited to 500 epochs, and StyleGAN3 was omitted. The present paper expands upon that comparison by systematically comparing StyleGAN2 and StyleGAN3, training the two models to 1000 epochs, and comparing their output on two medical image test sets: BUSI (ultrasound) and CBIS-DDSM (mammography) [7]. In our experiments, we focus on synthesizing malignant cases, which represent the minority class in both BUSI and CBIS-

DDSM datasets. The motivation is to mitigate class imbalance by generating realistic samples for the underrepresented malignant category. The research aim is therefore systematic and computational:

– To evaluate StyleGAN2 and StyleGAN3 under conditions of extreme data scarcity, with medical images as a standard case scenario.

– To investigate their quality (FID, KID, SWD) and efficiency (training time), and therefore highlight architecture trade-offs that hold for limited-sample IT applications.

In this in-depth analysis, we aim to achieve novel insights into the merits and limitations of StyleGAN-based architectures in medical images and shed light upon their promise of making up for dataset scarcity and breast cancer data skewness.

Methodology

This study was executed from two publicly accessible breast imaging datasets: the Breast Ultrasound Images Dataset (BUSI) and the Curated Breast Imaging Subset of DDSM (CBIS-DDSM). The breast ultrasound image dataset comprises 210 grayscale images of malignant cases. The images were all reduced to 512×512 pixels and normalized to the [0,1] interval [8]. The CBIS-DDSM dataset contains 509 mammograms that are labeled as malignant. The images were reduced to 256×256 pixels and normalized [9]. For this study, only the malignant subsets of BUSI (210 images) and CBIS-DDSM (509 images) were used. The generative task was specifically targeted at producing synthetic malignant samples, reflecting the minority class in both datasets.

Two of the most recent GAN architectures were taken into account in this study: StyleGAN2 and StyleGAN3. StyleGAN2 incorporates adaptive instance normalization, residual connections, and path length regularization that, together, enable synthesizing images that are plausible to human eyes. However, the architecture is known to be vulnerable to the aliasing artifacts that may be seen under extended training [5]. StyleGAN3 addresses the above under the alias-free design that is founded upon continuous phase-based transformations and filtered activations [6]. The above innovation introduces geometric stability and reduces inconsistencies, making it particularly suitable for medical images for which fine details matter. It is not restricted to medical imaging and can be generalized to additional domains for synthetic data generation to support machine learning under data-scarce conditions.

The training conditions were for the continuation of our previous research (Ryspayeva & Salykova, 2025). In the case of the BUSI and CBIS-DDSM tests, the batch size was 2. That decision was not only dictated by computational needs, however, but by the experience that such a setup provides for stronger convergence for those specific sets of information. The models were all trained for 1000 epochs, doubling the training period of the corresponding previous research, to interrogate the eventual StyleGAN architecture convergence behavior. The training was conducted using the Adam optimizer with $\beta 1 = 0.0$ and $\beta 2 = 0.99$ as hyperparameters, and the learning rate was maintained at a constant value of 0.002. Optimizing was done for the non-saturating GAN loss under R1 regularization. Adaptive augmentation was employed to achieve higher generalization in cases where data availability was limited. All of the experiments were run from a single GPU, with memory usage limited to approximately 10 GB [7].

Three standard metrics quantified performance.

– Fréchet Inception Distance (FID): measures similarity of distributions between real-world and artificial images in the Inception feature space [10].

– Kernel Inception Distance (KID): calculates an unbiased similarity estimate by utilizing a polynomial kernel, incredibly robust despite limited sample sizes [11].

– Sliced Wasserstein Distance (SWD): measures multi-scale distributional correspondence, structure-sensitive [12].

For visualization purposes, we employed Principal Component Analysis (PCA) to project higher-dimensional image representations into two-dimensional space in such a way that real and synthetic data distributions can be comparatively analyzed. We perform PCA over (224×224) size flattened grayscale images to visualize real and synthetic sample distribution alignment. The principal components of the PCA are calibrated only over the real dataset and subsequently applied to synthetic samples [13].

By integrating quality and efficiency analysis, it is possible to enable the holistic analysis of StyleGAN architectures for small sample sizes. It is not limited to medical imaging since it can be extended to synthetic data generation for machine learning support under data-scarce cases for different domains.

Experiments and Results

Two breast image datasets, CBIS-DDSM and BUSI, were tested with StyleGAN2 and StyleGAN3. The quantitative quality of images was examined by applying the FID and KID. Both distances were calculated for various epochs of training to track the change in the performance of the models as they converged.

Table 1 demonstrates FID and KID scores calculated for the BUSI dataset. StyleGAN3 was better than StyleGAN2 at every epoch. At 100 epochs, StyleGAN3 had already achieved significantly lower FID (281.17 vs. 565.37) and KID (0.27 vs. 0.78), indicating stabler convergence. After longer training, the gap continued to grow wider: at 1000 epochs, StyleGAN3 had achieved KID = 0.06 and FID = 140.67, while StyleGAN2 achieved KID = 0.25 and FID = 259.65

Table 1 – FID, KID and SWD values for BUSI dataset

Epoch	StyleGAN2			StyleGAN3		
	KID ↓	FID↓	SWD↓	` KID ↓	FID↓	SWD ↓
100	0.78±0.02	565.37	-	0.27±0.02	281.17	-
500	0.42±0.01	383.48	-	0.08±0.01	149.27	-
1000	0.25±0.01	259.65	5.62	0.06±0.01	140.67	3.21

The CBIS-DDSM dataset showed comparable trends, as shown in Table 2. StyleGAN2 showed continuous improvements over training, achieving 0.10 KID and 124.79 FID at 1000 epochs. StyleGAN3 remarkably achieved better outcomes at every checkpoint: at 1000 epochs, it achieved 0.06 KID and 90.58 FID, indicating ~27% FID improvement compared to StyleGAN2.

Table 2 – FID and KID values for CBIS-DDSM dataset

Epoch	StyleGAN2			StyleGAN3		
	KID ↓	FID↓	SWD↓	` KID ↓	FID↓	SWD↓
100	0.56±0.01	401.20	-	0.11	129.84	-
500	0.13±0.01	143.75	-	0.07	97.27	-
1000	0.10±0.01	124.79	25.00	0.06	90.58	24.32

Taken together, these results highlight the superiority of StyleGAN3 for medical image synthesis in both BUSI and CBIS-DDSM. On BUSI, StyleGAN3 reduced FID by nearly 46% and KID by more than 4× compared to StyleGAN2 at 1000 epochs. On CBIS-DDSM, StyleGAN3 achieved a 27% reduction in FID and a 40% reduction in KID at the same epoch. These findings emphasize that the alias-free design of StyleGAN3 can better capture subtle texture patterns and structural consistency, which are critical in medical imaging.

In addition to FID and KID, SWD, another proxy for quantifying synthetic-to-real distributional similarity, was also computed. StyleGAN3 had much lower SWD values, verifying its superior structure information retention capability: 3.21 vs. 5.62 on BUSI, and 24.32 vs. 25.00 on CBIS-DDSM.

In addition to image quality, the computational cost of model training was also of concern to us (Table 3). On BUSI, StyleGAN2 was trained for ~13.2 h for 1000 epochs, while StyleGAN3 nearly doubled the computational hours (~26.1 h). On CBIS-DDSM, StyleGAN2 was trained for ~20.5 h, with StyleGAN3 taking it to ~38.0 h. While costly, the massive quality gains observed with StyleGAN3 may be justified when the realism of structure is of utmost importance.

Table 3 – Training time statistics for StyleGAN2 and StyleGAN3

Dataset	Model	Time / epoch, s	sec/kimg, s	100 epochs	500 epochs, hours	1000 epochs, hours
CBIS-DDSM	StyleGAN2	74.0	145.4	~2.05 h	~10.3	~20.5
	StyleGAN3	136.7	268.6	~3.8 h	~19.0	~38.0
BUSI	StyleGAN2	47.5	226.2	~1.32 h	~6.6	~13.2
	StyleGAN3	94.0	447.6	~2.61 h	~13.1	~26.1

In addition to the quantitative results, we also qualitatively examined the generative potential of StyleGAN2 and StyleGAN3 for both of the datasets, BUSI and CBIS-DDSM. As indicated by Figure 1, the evolutionary trend of the FID and KID values demonstrates StyleGAN3's faster convergence that achieves smaller values and variability compared to StyleGAN2, in particular for BUSI. The numerical pattern is also evidenced by examination of distributions in Figure 2, for which the visualizations under PCA indicate a more comparable correspondence of the real and synthetic samples for StyleGAN3. The SWD corresponding values also prove the finding: StyleGAN3 achieved SWD of 3.21 for BUSI compared to 5.62 for StyleGAN2, and 24.32 for CBIS-DDSM compared to 25.00 for StyleGAN2. It suggests that StyleGAN3 is capable of achieving representations that are closest to the manifold of real medical images.

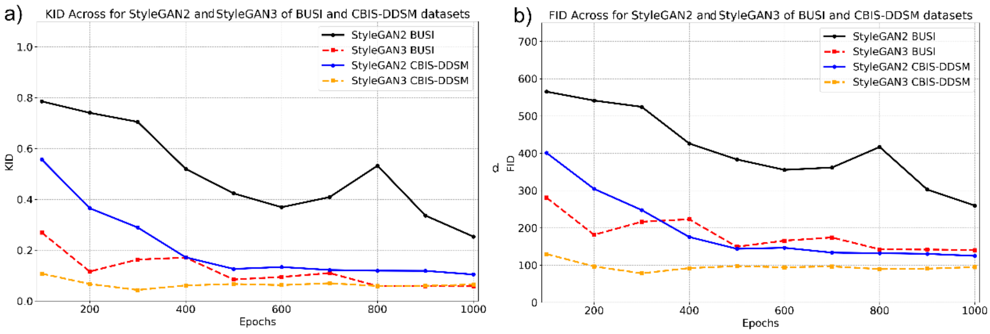

Figure 1 – Evolution of generative performance across training epochs for BUSI and CBIS-DDSM datasets:
(a) KID curves for StyleGAN2 and StyleGAN3, (b) FID curves for StyleGAN2 and StyleGAN3

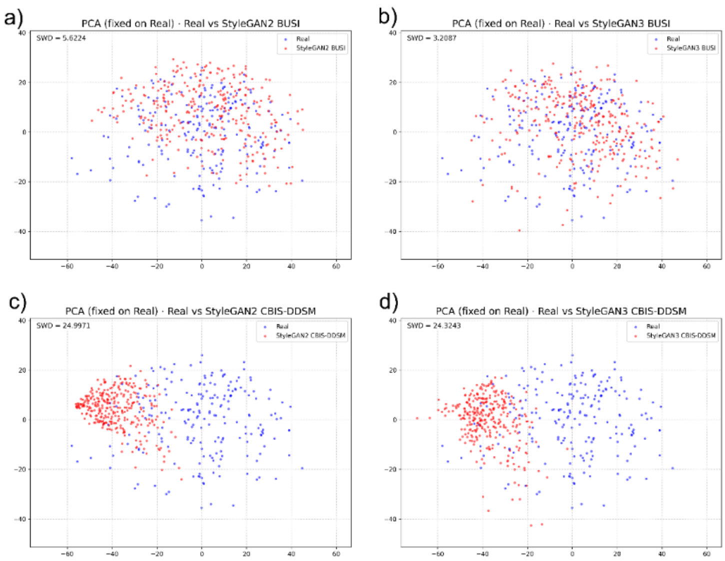

Figure 2 – PCA visualization of real and generated images from BUSI and CBIS-DDSM datasets:
(a) Real vs. StyleGAN2 on BUSI, (b) Real vs. StyleGAN3 on BUSI,
(c) Real vs. StyleGAN2 on CBIS-DDSM, (d) Real vs. StyleGAN3 on CBIS-DDSM

Qualitative analysis of the synthetic images also verifies these findings. In ultrasound synthesis of BUSI (Figure 3, (a)), StyleGAN3 possesses sharper texture and more natural-looking lesions with less artifact of smoothing when compared to StyleGAN2. Likewise, for CBIS-DDSM mammography images (Figure 3, (b)), the two models synthesize the general breast silhouette; however, the StyleGAN2 output images are subject to over-smoothing. In comparison, StyleGAN3 generates more detailed morphologies of parenchymal tissue with sharper anatomical structures, although with noticeable subtle blurring evident.

In short, these results show that StyleGAN3 not only improves numerical outcomes but also enhances the perceptual realism of simulated medical images. Such improvements are important particularly for medical imaging cases, for which visual plausibility has a direct impact upon diagnostic model learning and clinical interpretation of synthetic images.

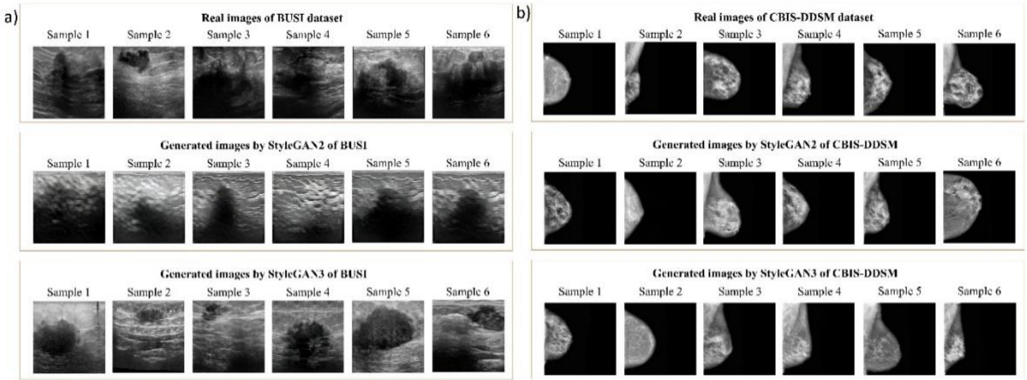

Figure 3 – (a) Qualitative comparison of real and synthetic images. (a) BUSI ultrasound dataset: first row – real samples, second row – StyleGAN2-generated, third row – StyleGAN3-generated. (b) CBIS-DDSM mammography dataset: first row – real samples, second row – StyleGAN2-generated, third row – StyleGAN3-generated

Discussion

Our experiment confirms that StyleGAN3 does make remarkable quality improvements of synthetic medical images over StyleGAN2, particularly for the BUSI dataset. On 100, 500, and 1000 epochs, StyleGAN3 kept achieving reduced FID and KID values, showing that its aliasing-free structure is better for learning fine-grained texture structures from ultrasound images. On CBIS-DDSM, StyleGAN2 and StyleGAN3 displayed confident convergences, and StyleGAN3 achieved extra reductions of FID (90.58 vs. 124.79) and KID (0.06 vs. 0.10) after epochs of 1000. These results verify that the advantages of StyleGAN3 over StyleGAN2 in BUSI actually extend to mammography, even when limited data are used. Previous works in Table 4, such as IGAN [14], 2S-BUSGAN [15], and GSDA [16], achieved FID scores of 40-100, but often for reduced image sizes (256 × 256 × 256) and larger datasets (≈780 images). Our own prior work with DGAN-WP-TL [7] achieved an FID of 179.4 and a KID of 0.1448 on BUSI (509 images), as it was experiencing issues learning from smaller sets. In contrast, StyleGAN3 trained on just 210 BUSI images reached an FID of 140.7 and a KID of 0.06 after 1000 epochs, representing a meaningful step forward. These results underscore the potential of advanced architectures to mitigate data scarcity and produce high-quality images despite limited training samples.

The results indicate that while StyleGAN3 does not yet surpass the very low FID values reported in some larger-scale studies (e.g., IGAN), it significantly improves over our previous DGAN-WP-TL baseline and StyleGAN2 under the same data constraints (Table 4).

Table 4 – Comparison of GAN-based methods for BUSI dataset synthesis

Author & Year	Method	FID	KID	Dataset
Alruily et al. (2023) [14]	IGAN	41.86	–	BUSI
Luo et al. (2023) [15]	2S-BUSGAN	101.00	0.6238	BUSI
Liu et al. (2023) [16]	GSDA	68.78	–	BUSI
Ryspayeva & Salykova (2025) [7]	DGAN-WP-TL	179.42	0.1448	BUSI
This work	StyleGAN2	259.65	0.25	BUSI
This work	StyleGAN3	140.67	0.06	BUSI

We also benchmarked StyleGAN2 and StyleGAN3 on CBIS-DDSM and compared them with prior studies on mammogram synthesis (Table 5). Methods such as CycleGAN [17], BreastGAN [18] and StyleGAN-XL [19] reported very low FID scores, but they typically relied on larger and more diverse datasets. In contrast, our experiments used only 509 malignant samples from CBIS-DDSM, which explains the relatively higher FID values. Importantly, StyleGAN3 achieved a substantial improvement, reducing FID from 124.8 (StyleGAN2) to 90.6 and KID from 0.10 to 0.06.

This comparison underlines the difficulty of working with CBIS-DDSM due to its small class size and limited diversity. Nevertheless, our StyleGAN3 results represent a clear improvement over both StyleGAN2 and DGAN-WP-TL under identical dataset constraints, and provide evidence that alias-free architectures improve mammogram synthesis as well.

Table 5 – Comparison of GAN-based methods for mammogram synthesis

Author & Year	Method	FID	KID	Dataset
Garrucho et al. (2023) [17]	CycleGAN	73.16	-	OPTIMAM
Fan et al. (2021) [18]	BreastGAN	21.25	-	DDSM
Prodan et al. (2023) [19]	StyleGAN-XL	9.8	-	ADMANI
Ryspayeva & Salykova (2025) [7]	DGAN-WP-TL	182.35	0.1795	CBIS-DDSM
This work	StyleGAN2	124.79	0.10	CBIS-DDSM
This work	StyleGAN3	90.58	0.06	CBIS-DDSM

While StyleGAN3 necessarily improves generative quality, it does so at non-negligible computational cost. Combined training for 1000 epochs increased from 13.2 hours for StyleGAN2 to 26.1 hours for StyleGAN3 for BUSI. For CBIS-DDSM, the analogous phenomenon was observed: 20.5 hours for StyleGAN2 and 38.0 hours for StyleGAN3. These findings highlight the trade-off between quality and efficiency: StyleGAN3 is preferable when fidelity and structural consistency are critical, while StyleGAN2 may remain a practical alternative in resource-limited environments.

Conclusion

Two state-of-the-art generative adversarial network models, StyleGAN2 and StyleGAN3, are experimented upon under breast imaging datasets (BUSI and CBIS-DDSM) in the current research. The research novelty is in the information technology field: alias-free generative model assessment under stringent data scarcity and imbalance conditions. It was shown after extending the training to 1000 epochs that StyleGAN3 significantly outperformed StyleGAN2 under all experimental conditions with appreciably diminished FID and KID values for two datasets. More specifically, StyleGAN3 achieved FID = 140.67 and KID = 0.06 for the dataset BUSI, and FID = 90.58 and KID = 0.0 By focusing on malignant case synthesis, this work demonstrates that advanced GAN architectures can enhance the representation of minority classes, offering a practical approach to dataset balancing in breast cancer imaging6 for the dataset CBIS-DDSM, hence establishing stronger medical image synthesis task baselines.

The implications are both methodological and computational. Firstly, they demonstrate that state-of-the-art alias-free generative design (StyleGAN3) translates to practical improvements in stability, faithful reproduction, and output sample diversity even for sparse datasets. Secondly, they reveal the computational-efficiency/image-quality trade-off: StyleGAN3 almost takes twice as much time to train as StyleGAN2, yet delivers much more realistic outcomes. These outcomes benefit the IT field by providing a plausible experimental framework, performance measures, and design trade-off insights for GAN-based synthesis under data-constrained conditions.

Future work will aim to scale these techniques to massive data sets, to combine pipelines of augmentation and transfer learning, and to construct hybrid architectures that combine alias-free generators with task-specialized priors. These areas will extend the use of generative models beyond healthcare even further, making data augmentation, simulation, and machine learning for small-sample domains even more of a core technology.

References

1. Generative adversarial nets / I. Goodfellow et al // Advances in Neural Information Processing Systems. – 2014. – Vol. 27. – P. 2672-2680.
2. Arjovsky M. Wasserstein Generative Adversarial Networks / M. Arjovsky, S. Chintala, L. Bottou // Proceedings of the 34th International Conference on Machine Learning (ICML). – 2017. – Vol. 70. – P. 214-223.
3. Improved training of Wasserstein GANs / I. Gulrajani et al // Advances in Neural Information Processing Systems (NeurIPS). – 2017. – P. 5767-5777.
4. Karras T. A Style-Based Generator Architecture for Generative Adversarial Networks / T. Karras, S. Laine, T. Aila // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2019. – P. 4401-4410.
5. Analyzing and Improving the Image Quality of StyleGAN / T. Karras et al // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2020. – P. 8110-8119.
6. Alias-Free Generative Adversarial Networks / T. Karras et al // Advances in Neural Information Processing Systems (NeurIPS). – 35th Int. Conf. on Neural Information Processing Systems. – 2021. – P. 1190-1201.

7. Ryspayeva M. Multi-domain synthetic medical image generation and dataset balancing with DGAN-WP-TL / M. Ryspayeva, O. Salykova // Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization. – 2025. – Vol. 13, № 1. https://orcid.org/10.1080/21681163.2025.2556687.
8. Dataset of Breast Ultrasound Images / W. Al-Dhabyani et al // Data in Brief. – 2019. – Vol. 28. – P. 104863. https://orcid.org/10.1016/j.dib.2019.104863.
9. Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) / R. Sawyer-Lee et al // The Cancer Imaging Archive. – 2016. https://orcid.org/10.7937/K9/TCIA.2016.7O02S9CY. – URL: https://www.cancerimagingarchive.net/collection/cbis-ddsm/.
10. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium / M. Heusel et al // Advances in Neural Information Processing Systems (NeurIPS). – 2017. – Vol. 30. – P. 6626-6637.
11. Demystifying MMD GANs / M. Bińkowski et al // International Conference on Learning Representations (ICLR). – 2018.
12. Progressive Growing of GANs for Improved Quality, Stability, and Variation / T. Karras et al // International Conference on Learning Representations (ICLR). – 2018. – Poster 2018.
13. Jolliffe I.T. Principal Component Analysis / I.T. Jolliffe. – 2nd ed. – New York: Springer, 2002. – 487 p. – (Springer Series in Statistics).
14. IGAN: Improved Generative Adversarial Network for Breast Ultrasound Image Synthesis / M. Alruily et al // Diagnostics. – 2023. – Vol. 13, № 15. – P. 2562. https://orcid.org/10.3390/diagnostics13152562.
15. 2S-BUSGAN: Two-Stage GAN for Breast Ultrasound Image Synthesis / W. Luo et al // Journal of Imaging. – 2023. – Vol. 9, № 3. – P. 54. https://orcid.org/10.3390/jimaging9030054.
16. Liu C. GSDA: Generative Semantic Data Augmentation for Breast Ultrasound Image Synthesis / C. Liu, S. Ding, W. Luo // Frontiers in Oncology. – 2023. – Vol. 13. – P. 1156789. https://orcid.org/10.3389/fonc.2023.1156789.
17. Cross-domain synthesis of mammograms using CycleGAN for data augmentation in breast cancer screening / L.Garrucho et al // Computers in Biology and Medicine. – 2023. – Vol. 165. – P. 107375. https://orcid.org/10.1016/j.compbiomed.2023.107375.
18. Fan J. BreastGAN: Adversarial Data Augmentation for Mammogram Classification / J. Fan, R. Huang, Z. Liu // Computers in Biology and Medicine. – 2021. – Vol. 135. – P. 104601. https://orcid.org/10.1016/j.compbiomed.2021.104601.
19. Prodan R. StyleGAN-XL for high-resolution mammogram synthesis with limited annotations / R. Prodan, P.F. Jäger, A. Kopp-Schneider // Medical Image Analysis. – 2023. – Vol. 87. – P. 102842. https://orcid.org/10.1016/j.media.2023.102842.

М.К. Рыспаева[1,2*], О.С. Салыкова[1]
[1]Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университеті,
110000, Қостанай, Қазақстан Байтурсынов көшесі 47,
[2]Astana IT University,
010000, Астана, Қазақстан, Мәңгілік Ел, С 1 блогы,
*e-mail: marya.rys1@mail.ru

STYLEGAN2 ЖӘНЕ STYLEGAN3 МОДЕЛЬДЕРІНІҢ BUSI ЖӘНЕ CBIS-DDSM ДЕРЕКҚОРЛАРЫНДА СИНТЕТИКАЛЫҚ МЕДИЦИНАЛЫҚ СУРЕТТЕРДІ ГЕНЕРАЦИЯЛАУҒА БАҒАЛАНУЫ

Медициналық кескіндер негізінде терең оқытуды қолдану көбінесе суреттердің шектеулі қолжетімділігіне және сыныптар теңгерімсіздігінің айқын байқалуына байланысты күрделене түседі. Мұндай жағдайда дәстүрлі машиналық оқыту алгоритмдерінің тиімділігі төмендейді. Генеративтік қарсылас желілер (GAN) бұл мәселені шешу үшін табиғи көрінетін синтетикалық суреттерді жасап, оқыту жиынтықтарын толықтыруға мүмкіндік береді. Бұл зерттеуде біз екі озық GAN архитектурасын, StyleGAN2 және StyleGAN3, екі ашық қолжетімді сүт безі кескіндерінің деректер жиынтығы негізінде салыстырдық: BUSI (ультрадыбыс, 210 қатерлі іс жағдайы) және CBIS-DDSM (маммография, 509 қатерлі іс жағдайы). FID және KID бойынша жүргізілді. BUSI деректерінде StyleGAN3 1000 эпохада FID = 140.7 және KID = 0.06 нәтижесін көрсетті, ал StyleGAN2 үшін FID = 259.7 және KID = 0.25 болды. CBIS-DDSM деректерінде StyleGAN3 FID = 90.6 және KID = 0.06 нәтижесін берді, ал StyleGAN2 сәйкесінше FID = 124.8 және KID = 0.10 көрсетті. Бұл нәтижелер

StyleGAN3 моделі шектеулі деректер жағдайында табиғиға ұқсас әрі әртүрлі суреттерді синтездей алатынын, бірақ сонымен бірге оқыту уақытының ұлғаюына әкелетінін көрсетеді. Ал StyleGAN2 салыстырмалы түрде төмен есептеу шығындарымен ұқсас сапаға қол жеткізе алады. Зерттеу нәтижелері медициналық суреттерді генерациялау әлеуетін және сапа мен тиімділіктің деректерді көбейту (data augmentation) үшін тепе-теңдігін көрсетеді.

Түйін сөздер: *Генеративтік қарсылас желілер (GAN); StyleGAN2; StyleGAN3; FID; KID; Сүт безі ультрадыбысы; Маммография.*

М.К. Рыспаева[1,2*], О.С. Салыкова[1]
[1]Костанайский региональный университет имени Ахмета Байтурсынулы,
110000, Костанай, Казахстан, ул. Байтурсынова 47
[2]Astana IT University,
010000, Астана, Казахстан Мангилик Ел, блок С.1
*e-mail: marya.rys1@mail.ru

ОЦЕНКА STYLEGAN2 И STYLEGAN3 ДЛЯ СИНТЕТИЧЕСКОЙ ГЕНЕРАЦИИ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ НА ДАТАСЕТАХ BUSI И CBIS-DDSM

Глубокое обучение на медицинских изображениях обычно затруднено ограниченным доступом к данным и сильным дисбалансом классов, что снижает эффективность традиционных алгоритмов машинного обучения. Генеративные состязательные сети (GAN) могут использоваться для решения таких проблем путем создания реалистичных синтетических изображений, дополняющих обучающие выборки. В данном исследовании мы сравниваем две передовые архитектуры GAN, StyleGAN2 и StyleGAN3, на основе двух общедоступных наборов данных изображений молочной железы: BUSI (ультразвук, 210 случаев злокачественных опухолей) и CBIS-DDSM (маммография, 509 случаев злокачественных опухолей). Оценка проводилась с использованием метрик FID и KID. На BUSI StyleGAN3 при 1000 эпохах достиг FID = 140.7 и KID = 0.06, тогда как StyleGAN2 показал FID = 259.7 и KID = 0.25. На CBIS-DDSM StyleGAN3 достиг FID = 90.6 и KID = 0.06, а StyleGAN2 – FID = 124.8 и KID = 0.10. Эти результаты демонстрируют, что StyleGAN3 имеет тенденцию синтезировать более реалистичные и разнообразные изображения в условиях ограниченных данных, но при этом требует большего времени обучения, тогда как StyleGAN2 обеспечивает сопоставимое качество при меньших вычислительных затратах. Результаты указывают на потенциал генерации медицинских изображений и компромисс между качеством и эффективностью для задач аугментации данных при улучшении изображений молочной железы.

Ключевые слова: *Генеративные состязательные сети, StyleGAN2, StyleGAN3, FID, KID, Ультразвук молочной железы, Маммография.*

Сведения об авторах

Марья Куанышевна Рыспаева* – докторант, Костанайский региональный университет имени Ахмет Байтұрсынұлы, Костанай, преподаватель, Astana IT University, Астана, Казахстан; e-mail: marya.rys1@mail.ru. ORCID: https://orcid.org/0000-0001-5055-4149.

Ольга Салыкова – кандидат технических наук, и.о. ассоциированного профессора Костанайский региональный университет имени Ахмет Байтұрсынұлы, Костанай Казахстан; e-mail: solga0603@mail.ru. ORCID: https://orcid.org/0000-0002-8681-4552.

Авторлар туралы мәліметтер

Марья Рыспаева* – докторант, Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университеті, Қостанай, Astana IT University оқытушы, Астана, Қазақстан; e-mail: marya.rys1@mail.ru. ORCID: https://orcid.org/0000-0001-5055-4149.

Ольга Салыкова – т.ғ.к., қауымдастырылған профессор, Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университеті, Қостанай Қазақстан; e-mail: solga0603@mail.ru. ORCID: https://orcid.org/0000-0002-8681-4552.

Information about the authors

Marya Ryspayeva* – doctoral student, Akhmet Baitursynuly Kostanay Regional University, Kostanay, Astana IT University lecturer, Astana, Kazakhstan; e-mail: marya.rys1@mail.ru. ORCID: https://orcid.org/0000-0001-5055-4149.

Olga Salykova – candidate of technical sciences, associate professor, Akhmet Baitursynuly Kostanay Regional University, Kostanay, Kazakhstan; e-mail: solga0603@mail.ru. ORCID: https://orcid.org/0000-0002-8681-4552.

Received 19.09.2025
Revised 28.10.2025
Accepted 03.11.2025