Augmentasi Citra Pra-Kanker Serviks Menggunakan Arsitektur Pix2pix

Marsella Vindriani
Jurusan Sistem Komputer,
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia

Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
marsellavindriani@gmail.com

Muhammad Naufal Rachmatullah*
Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
naufalrachmatullah@gmail.com

Ade Iriani Sapitri
Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
adeirianisapitri13@gmail.com

Akhiar Wista Arum
Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
wistaarum16@gmail.com

Anggun Islami
Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
anggunislami2@gmail.com

***Abstrak*— Kanker serviks merupakan salah satu penyebab utama kematian akibat kanker pada perempuan di seluruh dunia. Deteksi dini melalui skrining seperti Inspeksi Visual Asam Asetat efektif dan terjangkau, namun masih menghadapi kendala berupa subjektivitas interpretasi hasil serta keterbatasan tenaga medis spesialis. Seiring berkembangnya kecerdasan buatan, model deep learning mulai dimanfaatkan untuk mendukung diagnosis medis, tetapi kinerjanya sangat bergantung pada ketersediaan dataset yang besar dan beragam. Dataset publik *milik International Agency for Research on Cancer*, masih terbatas dalam jumlah, variasi, dan kualitas, sehingga menghambat kemampuan model dalam melakukan generalisasi. Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan pemanfaatan *conditional Generative Adversarial Network*, khususnya arsitektur Pix2pix, sebagai metode augmentasi citra pra-kanker serviks. Model ini menghasilkan citra sintetis fotorealistis berdasarkan masker segmentasi multikelas yang terdiri dari tiga kelas utama, yaitu *Cervical Area*, *Columnar Area*, dan *Lesion*, menggunakan dataset citra *post*-IVA pra-kanker serviks dari RSUP Mohammad Hoesin dan IARC. Hasil penelitian menunjukkan bahwa metode augmentasi menggunakan Pix2pix menghasilkan citra sintetis yang realistis dan konsisten dengan citra asli. Berdasarkan evaluasi *Fréchet Inception Distance*, model dengan resolusi citra 640×640 piksel mencapai nilai FID sebesar 54,37, menunjukkan tingkat kemiripan tinggi dengan citra nyata. Dengan demikian, augmentasi menggunakan arsitektur Pix2pix efektif dalam memperkaya variasi dataset medis dan menjadi solusi potensial untuk mengatasi keterbatasan data dalam pengembangan model segmentasi citra pra-kanker serviks yang lebih akurat dan andal.**

Kata Kunci— kanker serviks, IVA, augmentasi data, Pix2pix, cGAN, GAN, *Deep Learning*

I. PENDAHULUAN

Kanker serviks merupakan salah satu masalah kesehatan global yang menjadi penyebab utama kematian akibat kanker pada perempuan [1] [2]. Menurut data *World Health Organization* (WHO) tahun 2024, sekitar 660.000 wanita di seluruh dunia didiagnosis dengan kanker serviks dan lebih dari 350.000 kematian terjadi setiap tahunnya, dengan sebagian besar kasus berasal dari negara berpenghasilan rendah dan menengah [3], [4]. Di Indonesia, kanker serviks menempati posisi kedua tertinggi setelah kanker payudara pada wanita, dengan lebih dari 36.000 kasus baru dan sekitar 21.000 kematian setiap tahunnya menurut laporan WHO dan Kementerian Kesehatan RI tahun 2024 [5]. Namun, penyakit ini termasuk salah satu jenis kanker yang dapat dicegah dan disembuhkan apabila lesi pra-kanker berhasil dideteksi dan ditangani sejak dini melalui skrining. Program skrining nasional menggunakan metode Inspeksi Visual dengan Asam Asetat (IVA) menjadi pendekatan utama karena biayanya rendah dan mudah dilakukan di fasilitas kesehatan dasar. Prosedur ini dilakukan dengan mengoleskan larutan asam asetat pada leher rahim untuk menonjolkan area lesi pra-kanker sebagai perubahan warna putih (*acetowhite*) [2], [6].

Meskipun metode IVA efektif untuk deteksi dini, hasil interpretasinya sangat bergantung pada pengalaman tenaga kesehatan dan kondisi pencahayaan saat pemeriksaan. Hal ini menyebabkan hasil diagnosis sering bersifat subjektif dan bervariasi antar pemeriksa, terutama di daerah dengan sumber daya terbatas [7]. Tantangan ini menunjukkan perlunya inovasi berbasis teknologi, seperti penerapan kecerdasan buatan (AI) dan computer vision, untuk membantu proses deteksi otomatis lesi pra-kanker serviks secara lebih objektif, konsisten, dan akurat. Seiring dengan kemajuan teknologi kecerdasan buatan (AI), model *deep learning* menawarkan solusi menjanjikan untuk membantu proses deteksi dini kanker serviks secara lebih objektif dan efisien. Arsitektur canggih seperti *You Only Look Once* (YOLO) dan *Segment Anything Model* (SAM) telah menunjukkan kemampuan superior dalam tugas deteksi dan segmentasi citra medis. Namun, performa model-model ini sangat bergantung pada ketersediaan data latih berlabel dalam jumlah besar dan beragam [8] [9] [10]. Saat ini, dataset citra pra-kanker yang tersedia untuk publik, seperti yang disediakan oleh *International Agency for Research on Cancer* (IARC), masih terbatas dari segi kuantitas, variasi, dan kualitas [11]. Kelangkaan data ini menjadi tantangan fundamental yang

menghambat kemampuan model *deep learning* untuk melakukan generalisasi pada data baru dan mengenali variasi lesi atipikal yang jarang muncul [12] [13].

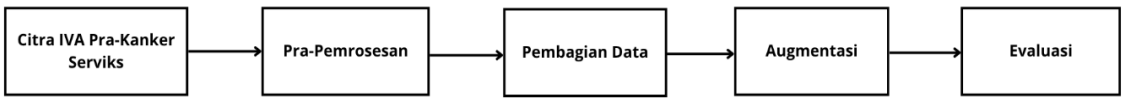
Untuk mengatasi keterbatasan tersebut, augmentasi data menjadi sebuah pendekatan krusial untuk memperkaya dataset pelatihan secara artifisial [14], [15] [16]. Metode augmentasi konvensional seringkali gagal menciptakan variasi data yang benar-benar baru dan realistis secara klinis [17]. Oleh karena itu, diperlukan pendekatan yang lebih canggih. *Generative Adversarial Networks* (GANs) menawarkan paradigma terobosan dengan kemampuannya mensintesis data baru yang sepenuhnya orisinal dan fotorealistis [15], [18]–[20]. Secara khusus, varian *conditional* GAN (cGAN) dengan menggunakan arsitektur Pix2pix memungkinkan proses generasi citra dikontrol oleh input tambahan, seperti masker segmentasi, sehingga dapat menghasilkan citra yang sesuai dengan kondisi yang diinginkan [21][22].

Berdasarkan permasalahan tersebut, penelitian ini berupaya mengisi kekosongan pada penelitian sebelumnya dengan menerapkan arsitektur Pix2pix, untuk melakukan augmentasi citra pra-kanker serviks serta menganalisis kualitas hasil citra sintetis yang dihasilkan. Penelitian ini berfokus pada pembuatan citra sintetis fotorealistis dari masker segmentasi multikelas guna memperkaya dataset medis dan meningkatkan performa model segmentasi otomatis di masa mendatang.

II. METODOLOGI PENELITIAN

A. Alur Kerja Penelitian

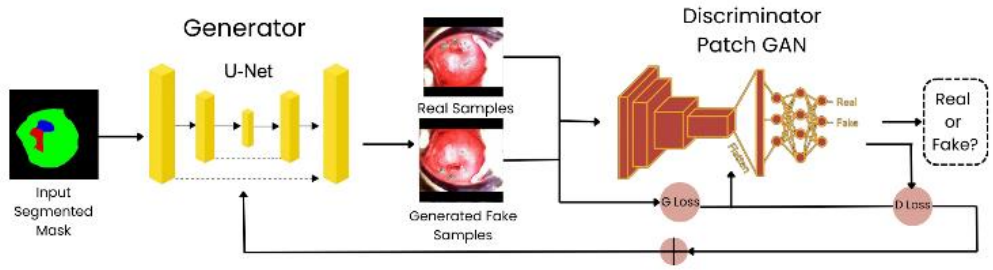
Alur penelitian ditunjukkan pada Gambar 1, dimulai dari pengumpulan citra IVA pra-kanker serviks, kemudian dilakukan pra-pemrosesan berupa anotasi dan pembuatan masker segmentasi. Data yang telah siap selanjutnya dibagi menjadi data latih dan data validasi. Tahap inti adalah proses augmentasi menggunakan arsitektur Pix2pix untuk menghasilkan citra sintetis, yang kemudian dievaluasi baik secara kuantitatif maupun kualitatif guna menilai kualitas dan realisme citra yang dihasilkan.



Gambar 1. Alur Kerja Penelitian

B. Arsitektur Pix2pix

Pix2pix adalah sebuah kerangka kerja cGAN yang dirancang khusus untuk tugas translasi citra-ke-citra (*image-to-image translation*). Arsitektur ini mampu mempelajari pemetaan dari sebuah citra input ke sebuah citra output yang diinginkan [21]. Dalam penelitian ini, Pix2pix digunakan untuk menerjemahkan masker segmentasi menjadi citra serviks yang fotorealistis. Arsitektur ini terdiri dari dua komponen utama yang visualisasinya dapat dilihat pada Gambar 2.



Gambar 2. Arsitektur Pix2pix

Arsitektur ini terdiri dari dua komponen utama, antara lain :

- Generator berbasis U-*Net*: terdiri dari *encoder* dan *decoder* yang dihubungkan skip connections. Hal ini memungkinkan model mempertahankan detail tekstur dan struktur dari masker segmentasi.
- Diskriminator *Patch*GAN: tidak hanya mengevaluasi keseluruhan citra, diskriminator menilai realisme pada *patch* berukuran kecil, sehingga mendorong generator menghasilkan detail frekuensi tinggi yang realistis.

Fungsi objektif dari cGAN yang digunakan dalam Pix2pix dirumuskan pada Persamaan (1) :

$$\mathcal{L}_{cGAN}(G,D) = \mathbb{E}_{x,y}\left[log\, D(x,y)\right] + \mathbb{E}_{z,y}[log(1 - D(G(z,y),y))] \quad (1)$$

Dimana G mencoba meminimalkan fungsi ini, sedangkan D mencoba memaksimalkannya. Variabel y merepresentasikan citra input (masker segmentasi) dan x adalah citra target (citra serviks asli).

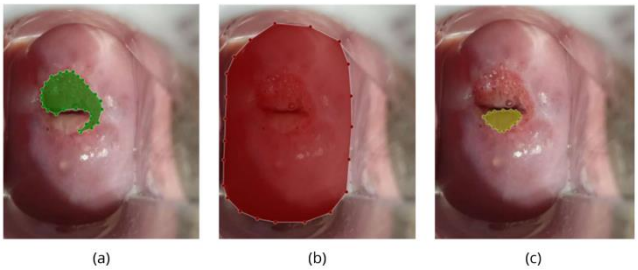
III. DATASET

A. Persiapan Data

Langkah awal dalam penelitian ini adalah persiapan dataset. Data yang digunakan berupa citra *post*-IVA pra-kanker serviks abnormal yang bersumber dari RSUP Mohammad Hoesin sebanyak 118 citra dan IARC sebanyak 159 citra [11].

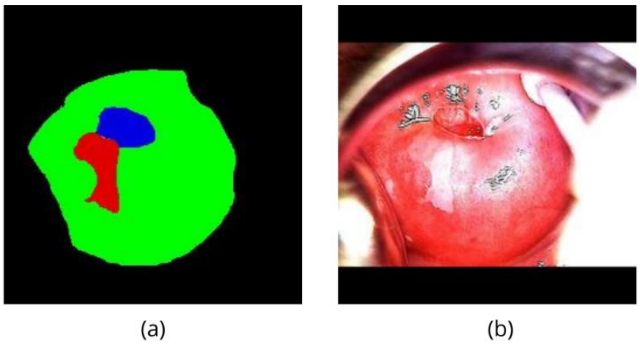
B. Pra-pemrosesan Data

Tahap pra-pemrosesan data mencerminkan metodologi standar yang umum digunakan untuk tugas translasi citra-ke-citra. Setiap citra kemudian melalui proses anotasi objek menggunakan *Software LabelMe* untuk menghasilkan label berupa masker segmentasi. Proses ini dilakukan dengan memberikan label poligon pada tiga kelas objek utama, yaitu *Columnar Area* (CA), *Cervical Area*, dan *Lesion*. Seluruh hasil anotasi divalidasi oleh ahli medis [8], [23]. Visualisasi hasil anotasi ditampilkan pada Gambar 3, yang menunjukkan contoh representasi tiap kelas objek.



Gambar 3. Label Segmentasi: (a) CA, (b) *Cervical Area*, (c) *Lesion*

Hasil dari tahap pra-pemrosesan ini adalah dataset berpasangan yang terdiri dari citra IVA asli dan masker segmentasi multikelas yang sesuai. Dengan demikian, setiap sampel data memiliki pasangan yang jelas antara citra asli dan masker segmentasinya. Adapun visualisasi data ditunjukkan pada Gambar 4.



Gambar 4. Visualisasi Data: (a) Masker Segmentasi (b) Citra Asli

Penelitian ini melakukan dua skenario eksperimen dengan ukuran citra yang berbeda untuk membandingkan hasilnya, antara lain :

1. Ukuran 256x256 piksel: Citra asli di-*resize* ke ukuran 286x286 piksel, kemudian di-*crop* secara acak menjadi 256x256 piksel.

2. Ukuran 640x640 piksel: Citra digunakan dalam ukuran aslinya tanpa melalui proses *resize* dan *crop*.

Pemilihan ukuran 256×256 didasarkan pada arsitektur standar model Pix2Pix, yang secara default dilatih dengan input berukuran 256 piksel agar efisien secara komputasi . Proses *random cropping* dilakukan untuk menambah variasi spasial dan mencegah *overfitting* dengan memaksa model belajar dari posisi objek yang berbeda dalam citra. Sementara ukuran 640×640 dipertahankan untuk menjaga detail anatomi dan tekstur halus pada area lesi, sehingga model dapat menghasilkan citra sintetis dengan struktur yang lebih realistis.

C. Pembagian Data

Setelah melalui tahap pra-pemrosesan, dataset yang digunakan dalam penelitian ini berjumlah 267 pasangan citra, yang merupakan gabungan dari dua sumber utama, yaitu citra *post*-IVA pra-kanker serviks yang diperoleh dari RSUP Mohammad Hoesin dan IARC. Pembagian ini mengikuti alokasi data train dan data valid sebesar 80:20, dapat dilihat pada Tabel 1.

TABLE I. PEMBAGIAN DATA

	Citra Asli (*Ground Truth*)	**Masker Segmentasi (Input)**
Data Latih	216	216
Data Valid	51	51

Dataset tersebut kemudian dibagi menjadi dua bagian terpisah untuk keperluan pelatihan dan validasi model. Sebanyak 216 pasangan citra digunakan sebagai data latih, yang terdiri dari citra asli sebagai *ground truth* dan masker segmentasi sebagai input untuk proses augmentasi. Sementara itu, 51 pasangan citra sisanya digunakan sebagai data validasi untuk mengevaluasi performa model dalam menghasilkan citra sintetis baru.

D. Metrik Evaluasi

Kualitas dan realisme citra sintetis yang dihasilkan oleh model cGAN dievaluasi secara kuantitatif menggunakan *Fréchet Inception Distance* (FID). Metrik ini dipilih karena kemampuannya yang unggul dalam mengukur kesamaan konseptual antara distribusi citra asli dan citra buatan. FID menghitung jarak antara distribusi fitur dari kedua set citra yang diekstraksi menggunakan model *Inception*V3 [24], [25]. Rumus untuk menghitung FID terdapat pada Persamaan (2):

$$\mathrm{FID} = \| \mu_r - \mu_g \|^2 + \mathrm{T_r}\,(\Sigma_r + \Sigma_g - 2\,(\Sigma_r \Sigma_g)^{1/2}) \quad (2)$$

Keterangan:

- μ_r, Σ_r= rata-rata dan kovarians fitur citra asli (*real images*).
- μ_g, Σ_g = rata-rata dan kovarians fitur citra hasil generatif (*generated images*).
- $\mathrm{Tr}(\cdot)$= fungsi *trace* matriks, yaitu jumlah elemen diagonal utama.

Skor FID yang lebih rendah menunjukkan bahwa citra yang dihasilkan memiliki kualitas dan keragaman yang lebih mirip dengan citra asli.

IV. HASIL DAN EVALUASI

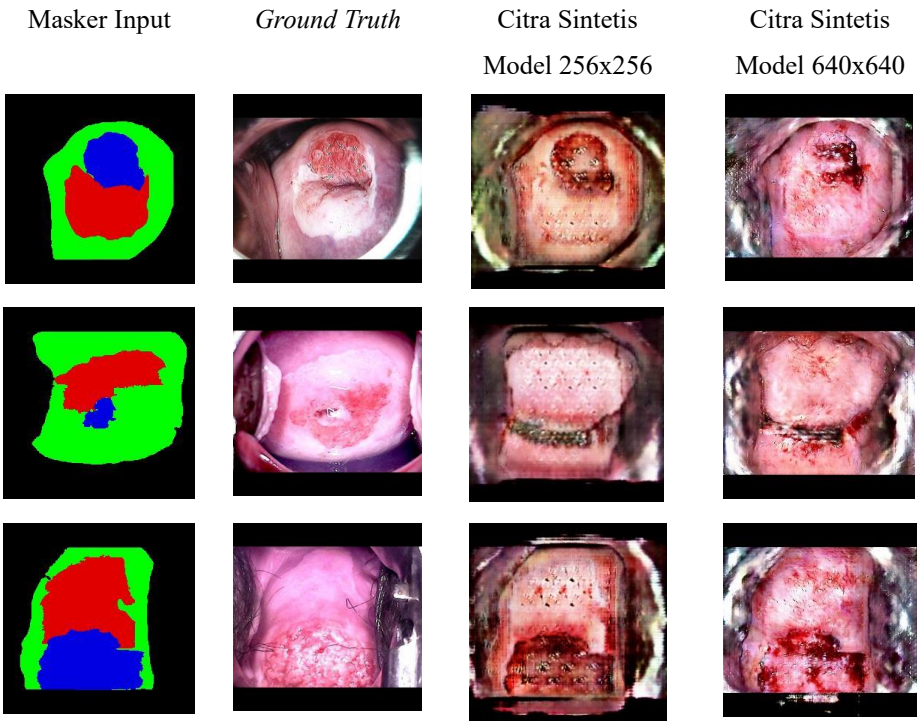
Setelah melalui proses pelatihan dan optimisasi, model augmentasi yang dikembangkan pada kedua skenario berhasil mencapai skor FID pada Tabel 2 berikut.

TABLE II. HASIL

Ukuran citra	*Buffer Size*	*Steps*	Skor FID
256x256	400	40.000	474.12
640x640	**50**	**55.000**	**54.37**

Hasil ini menunjukkan bahwa model yang dilatih pada resolusi lebih tinggi (640x640) mampu menghasilkan citra sintetis dengan tingkat realisme dan variasi yang lebih baik, yang tercermin dari skor FID yang lebih rendah.

Selain evaluasi kuantitatif, dilakukan pula perbandingan visual untuk menilai kualitas citra yang dihasilkan secara kualitatif. Untuk menunjukkan proses translasi citra-ke-citra secara utuh, Gambar 5 menyajikan beberapa contoh visualisasi antara masker input, citra asli (*ground truth*), dan citra sintetis yang digenerasi oleh model 640x640.



Gambar 5. Visualisasi Hasil Augmentasi

Visualisasi pada Gambar 5 menunjukkan bahwa model Pix2Pix mampu menerjemahkan informasi struktural pada masker input menjadi citra serviks yang tampak realistis dan memiliki kesesuaian anatomi dengan ground truth. Detail penting seperti tekstur jaringan, perubahan warna lesi, kontur serviks, serta karakteristik pencahayaan pada citra post-IVA dapat direkonstruksi dengan baik, sehingga hasil augmentasi dinilai konsisten secara klinis dan tidak hanya valid secara kuantitatif berdasarkan skor FID. Selain itu, perbandingan antara model beresolusi 256×256 dan 640×640 menunjukkan perbedaan kualitas yang signifikan. Model 256×256 masih menghasilkan artefak visual seperti blur pada batas lesi, distorsi struktur serviks, serta kurang mampu mempertahankan detail tekstur akibat proses downsampling yang berlebihan. Sementara itu, model 640×640 mampu menghasilkan citra sintetis dengan ketajaman fitur yang lebih tinggi, kontur jaringan yang lebih presisi, serta representasi spasial lesi yang lebih akurat terhadap ground truth. Oleh karena itu, dapat disimpulkan bahwa penggunaan resolusi tinggi pada proses augmentasi berbasis Pix2Pix memberikan kinerja rekonstruksi citra yang lebih baik dan layak diterapkan untuk meningkatkan kualitas dataset pada sistem segmentasi pra-kanker serviks.

DISKUSI

Hasil penelitian ini mengonfirmasi keunggulan metode augmentasi berbasis GAN. Dengan memanfaatkan arsitektur Pix2pix, model mampu menghasilkan citra baru yang bervariasi dan fotorealistis, sehingga memperkaya representasi data untuk pelatihan model segmentasi otomatis.

Nilai FID yang lebih rendah pada resolusi 640×640 piksel dibandingkan 256×256 piksel menunjukkan bahwa citra dengan resolusi tinggi memiliki tingkat realisme dan kesamaan distribusi fitur yang lebih baik terhadap citra asli. Hal ini dapat dijelaskan oleh fakta bahwa resolusi tinggi mempertahankan informasi spasial dan detail anatomi yang lebih lengkap, seperti tekstur jaringan dan kontur serviks. Dalam resolusi rendah, banyak detail mikro seperti variasi

warna jaringan, bayangan, dan batas halus antararea lesi hilang selama proses *downsampling*, yang menyebabkan generator kehilangan konteks lokal penting untuk menghasilkan tekstur realistis. Selain itu, diskriminator *Patch*GAN pada resolusi tinggi memiliki lebih banyak area patch untuk dinilai, sehingga mampu memberikan umpan balik yang lebih kaya dan mendorong generator menghasilkan citra dengan kualitas yang lebih baik.

Meskipun hasilnya menunjukkan potensi besar, penelitian ini memiliki beberapa keterbatasan (*limitations*). Pertama, jumlah dataset yang digunakan masih relatif terbatas hanya 267 pasangan citra, sehingga keberagaman data masih belum cukup untuk merepresentasikan seluruh variasi klinis lesi serviks. Kedua, evaluasi hasil augmentasi baru difokuskan pada metrik FID dan perbandingan visual, belum mencakup pengujian lanjutan terhadap peningkatan kinerja model segmentasi yang dilatih dengan data hasil augmentasi. Ketiga, model Pix2pix masih memiliki risiko menghasilkan *artifacts* atau tekstur buatan yang sulit diidentifikasi secara klinis tanpa validasi dari tenaga medis ahli. Oleh karena itu, penelitian lanjutan perlu mencakup peningkatan jumlah data, penerapan metrik evaluasi tambahan seperti *Structural Similarity Index Measure* (SSIM) dan *Peak Signal-to-Noise Ratio* (PSNR), serta integrasi hasil augmentasi ke dalam *pipeline* segmentasi *end-to-end* untuk menilai dampak langsung terhadap performa model medis.

KESIMPULAN

Penelitian ini mengimplementasikan arsitektur Pix2pix untuk augmentasi citra pra-kanker serviks dan menghasilkan citra sintetis yang fotorealistis serta konsisten secara klinis. Berdasarkan hasil evaluasi, model dengan resolusi 640×640 piksel menunjukkan performa yang jauh lebih baik dengan skor FID sebesar 54,37, dibandingkan dengan resolusi 256×256 piksel yang memperoleh nilai FID sebesar 474,12. Perbedaan ini menunjukkan bahwa resolusi tinggi mampu mempertahankan detail anatomi dan tekstur jaringan yang lebih akurat, sehingga citra hasil generasi memiliki realisme visual yang tinggi dan layak digunakan untuk memperkaya dataset medis. Harapannya, penelitian ini dapat dikembangkan lebih lanjut dengan mengevaluasi pengaruh dataset hasil augmentasi terhadap performa model segmentasi otomatis, serta mengeksplorasi arsitektur generatif lain seperti *Diffusion Model* untuk meningkatkan kualitas dan keragaman citra sintetis yang dihasilkan.

DAFTAR PUSTAKA

[1] J. Ferlay *et al.*, "Cancer statistics for the year 2020: An overview," *Int. J. Cancer*, vol. 149, no. 4, pp. 778–789, 2021, doi: 10.1002/ijc.33588.

[2] R. Viñals, M. Jonnalagedda, P. Petignat, J. P. Thiran, and P. Vassilakos, "Artificial Intelligence-Based Cervical Cancer Screening on Images Taken during Visual Inspection with Acetic Acid: A Systematic Review," *Diagnostics*, vol. 13, no. 5, 2023, doi: 10.3390/diagnostics13050836.

[3] World Health Organization, "Cervical Cancer," World Health Organization. Accessed: Sep. 23, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cervical-cancer

[4] A. Lohiya *et al.*, "Effectiveness of Visual Inspection with Acetic Acid (VIA) Screening on Cervical Cancer Mortality and Incidence, A Systematic Review and Meta,Analysis," *Asian Pacific J. Cancer Prev.*, vol. 23, no. 2, pp. 399–407, 2022, doi: 10.31557/APJCP.2022.23.2.399.

[5] World Health Organization (WHO), "WHO, UNFPA mengapresiasi upaya Indonesia mengeliminasi kanker serviks, mendorong strategi vaksin terpadu, dan memperkuat skrining," World Health Organization Indonesia. Accessed: Sep. 16, 2025. [Online]. Available: https://www.who.int/indonesia/id/news/detail/15-11-2024-who--unfpa-commend-indonesia-s-efforts-to-eliminate-cervical-cancer--urge-streamlined-vaccine-strategy-and-enhanced-screening

[6] S. Wang *et al.*, "Cervical Cancer Screening via Visual Inspection with Acetic Acid and Lugol Iodine for Triage of HPV-Positive Women," *JAMA Netw. Open*, vol. 7, no. 3, p. E244090, 2024, doi: 10.1001/jamanetworkopen.2024.4090.

[7] S. Nurmaini *et al.*, "Real time mobile AI-assisted cervicography interpretation system," *Informatics Med. Unlocked*, vol. 42, no. July, p. 101360, 2023, doi: 10.1016/j.imu.2023.101360.

[8] S. Nurmaini *et al.*, "Robust assessment of cervical precancerous lesions from pre- and post-acetic acid cervicography by combining deep learning and medical guidelines," *Informatics Med. Unlocked*, vol. 52, no. June 2024, p. 101609, 2025, doi: 10.1016/j.imu.2024.101609.

[9] A. S. Allogmani, R. M. Mohamed, N. M. Al-shibly, and M. Ragab, "Enhanced cervical precancerous lesions detection and classification using Archimedes Optimization Algorithm with transfer learning," *Sci. Rep.*, vol. 14, no. 1, pp. 1–14, 2024, doi: 10.1038/s41598-024-62773-x.

[10] Y. Zhang, Z. Shen, and R. Jiao, "Segment anything model for medical image segmentation: Current applications and future directions," *Comput. Biol. Med.*, vol. 171, no. February, p. 108238, 2024, doi: 10.1016/j.compbiomed.2024.108238.

[11] I. A. for R. on Cancer, "IARC Cervical Cancer Image Bank," International Agency for Research on Cancer (IARC). Accessed: Aug. 10, 2024. [Online]. Available: https://screening.iarc.fr/cervicalimagebank.php

[12] S. Yang, K. D. Kim, E. Ariji, N. Takata, and Y. Kise, "Evaluating the performance of generative adversarial network-synthesized periapical images in classifying C-shaped root canals," *Sci. Rep.*, vol. 13, no. 1, pp. 1–10, 2023, doi: 10.1038/s41598-023-45290-1.

[13] H. M. Rai, J. Yoo, S. Agarwal, and N. Agarwal, "LightweightUNet: Multimodal Deep Learning with GAN-Augmented Imaging Data for Efficient Breast Cancer Detection," *Bioengineering*, vol. 12, no. 1, 2025, doi: 10.3390/bioengineering12010073.

[14] S. Chen, D. Gao, L. Wang, and Y. Zhang, "Cervical Cancer Single Cell Image Data Augmentation Using Residual Condition Generative Adversarial Networks," *2020 3rd Int. Conf. Artif. Intell. Big Data, ICAIBD 2020*, no. x, pp. 237–241, 2020, doi: 10.1109/ICAIBD49809.2020.9137494.

[15] H. C. Park, I. P. Hong, S. Poudel, and C. Choi, "Data Augmentation Based on Generative Adversarial Networks for Endoscopic Image Classification," *IEEE Access*, vol. 11, no. April, pp. 49216–49225, 2023, doi: 10.1109/ACCESS.2023.3275173.

[16] Z. Zhang *et al.*, "DiffBoost: Enhancing Medical Image

Segmentation via Text-Guided Diffusion Model," *IEEE Trans. Med. Imaging*, vol. XX, no. Xx, pp. 1–13, 2024, doi: 10.1109/TMI.2024.3519307.

[17] Y. Jiang and V. S. K. Manem, "Data augmented lung cancer prediction framework using the nested case control NLST cohort," *Front. Oncol.*, vol. 15, no. February, pp. 1–12, 2025, doi: 10.3389/fonc.2025.1492758.

[18] G. Matsukawa and A. Yoshitaka, "Data Augmentation with Diffusion Model for Hand Detection," *Proc. - 2024 Int. Symp. Multimedia, ISM 2024*, pp. 170–173, 2024, doi: 10.1109/ISM63611.2024.00037.

[19] Z. Xu *et al.*, "Cross-domain attention-guided generative data augmentation for medical image analysis with limited data," *Comput. Biol. Med.*, vol. 168, no. November 2023, p. 107744, 2024, doi: 10.1016/j.compbiomed.2023.107744.

[20] J. Liu *et al.*, "CycleGAN Clinical Image Augmentation Based on Mask Self-Attention Mechanism," *IEEE Access*, vol. 10, no. September, pp. 105942–105953, 2022, doi: 10.1109/ACCESS.2022.3211670.

[21] M. Tahmid, M. S. Alam, N. Rao, and K. M. A. Ashrafi, "Image-to-Image Translation with Conditional Adversarial Networks," *Proc. 2023 IEEE 9th Int. Women Eng. Conf. Electr. Comput. Eng. WIECON-ECE 2023*, pp. 468–472, 2023, doi: 10.1109/WIECON-ECE60392.2023.10456447.

[22] B. Lakshmipriya, P. K. Singh, and S. Jayalakshmy, "Pix2pix GAN for Image-to-Image Translation," *Gener. AI*, no. August, pp. 73–90, 2025, doi: 10.1002/9781394302932.ch4.

[23] D. G. Saragih, A. Hibi, and P. N. Tyrrell, "Using diffusion models to generate synthetic labeled data for medical image segmentation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 19, no. 8, pp. 1615–1625, 2024, doi: 10.1007/s11548-024-03213-z.

[24] A. Borji, "Pros and cons of GAN evaluation measures: New developments," *Comput. Vis. Image Underst.*, vol. 215, no. December 2021, p. 103329, 2022, doi: 10.1016/j.cviu.2021.103329.

[25] M. J. Chong and D. Forsyth, "Effectively unbiased FID and inception score and where to find them," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 6069–6078, 2020, doi: 10.1109/CVPR42600.2020.00611.