Generative AI Menggunakan DCGAN untuk Analisis Citra Ultrasonografi Jantung Janin

1st M. Razhel Miroza
Jurusan Sistem Komputer,
Fakultas Ilmu Komputer
Universitas Sriwijaya

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

2nd Annisa Darmawahyuni
Intelligent System Research Group,
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia
riset.annisadarmawahyuni@gmail.com

3rd Ade Iriani Sapitri
Intelligent System Research Group,
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia
adeirianisapitri13@gmail.com

Abstrak—Penelitian ini bertujuan untuk menghasilkan citra ultrasonografi jantung janin sintetis menggunakan *Deep Convolutional Generative Adversarial Network* (DCGAN). Model dilatih secara terpisah untuk setiap pasien dengan konfigurasi pelatihan yang sama guna menilai kemampuan DCGAN dalam mereplikasi pola anatomi pada berbagai kategori kelainan jantung bawaan, yaitu *Atrial Septal Defect* (ASD), *Atrioventricular Septal Defect* (AVSD), *Ventricular Septal Defect* (VSD), dan kondisi normal. Evaluasi dilakukan menggunakan metrik *Fréchet Inception Distance* (FID), *Kernel Inception Distance* (KID), dan *Learned Perceptual Image Patch Similarity* (LPIPS) disertai analisis visual. Hasil menunjukkan variasi kualitas antar pasien; sebagian citra memperlihatkan struktur anatomi yang jelas, sedangkan lainnya tampak seragam atau kabur karena keterbatasan data. Nilai FID yang relatif tinggi menandakan distribusi citra sintetis belum sepenuhnya menyerupai citra asli, sementara nilai KID dan LPIPS menunjukkan kemiripan perseptual yang baik. Temuan ini mengindikasikan potensi DCGAN sebagai pendekatan generatif untuk memperluas dataset ultrasonografi medis.

Kata Kunci—*Congenital Heart Disease*, DCGAN, *Fetal Cardiac Ultrasound*, *Synthetic Medical Image Generation*

I. PENDAHULUAN

Perkembangan teknologi *Artificial Intelligence* (AI) telah mendorong transformasi besar dalam bidang kesehatan, khususnya dalam analisis citra medis. Salah satu cabang AI yang memiliki peran penting adalah *Deep Learning*, yang telah terbukti efektif dalam mengekstraksi fitur kompleks dan menghasilkan representasi citra dengan kualitas tinggi [1].

Dalam praktik klinis, ultrasonografi merupakan modalitas pencitraan jantung yang banyak digunakan karena bersifat non-invasif, aman, serta mampu memberikan informasi anatomi dan fungsi jantung secara *real-time*. Namun, citra ultrasonografi sering kali memiliki keterbatasan berupa *speckle noise*, kualitas akuisisi yang bervariasi, serta resolusi rendah, sehingga menyulitkan proses analisis otomatis berbasis komputer. Keterbatasan ini menimbulkan tantangan bagi pengembangan model *deep learning* yang bergantung pada kualitas dan kuantitas data. Beberapa studi menunjukkan bahwa kualitas akuisisi citra ultrasonografi janin sangat berpengaruh terhadap akurasi pengenalan struktur jantung [2]

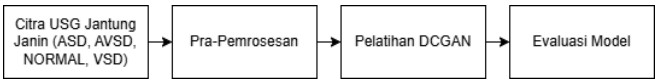
Salah satu pendekatan yang berkembang untuk mengatasi keterbatasan data adalah *Generative Adversarial Network* (GAN). Model ini mampu menghasilkan citra sintetis yang menyerupai data asli, sehingga dapat memperkaya dataset yang terbatas dan berpotensi meningkatkan performa model pembelajaran mesin. Tinjauan terkini menyebutkan bahwa GAN banyak digunakan dalam augmentasi citra medis, termasuk untuk meningkatkan keseimbangan dataset dan kualitas citra [3]. Studi lain juga menunjukkan bahwa citra ultrasonografi sintetis yang dihasilkan GAN dapat dievaluasi secara klinis dan memiliki kualitas mendekati citra asli [4].

Di antara berbagai varian GAN, *Deep Convolutional* GAN (DCGAN) menjadi salah satu arsitektur populer karena mampu menghasilkan citra dengan struktur yang lebih stabil dan realistis melalui pemanfaatan lapisan konvolusi. Sebuah studi empiris membandingkan berbagai varian GAN dan menyebutkan bahwa meskipun DCGAN termasuk arsitektur dasar, model ini tetap kompetitif dalam menghasilkan citra medis dengan kualitas yang dapat diterima [5].

Penelitian ini berfokus pada penerapan DCGAN untuk menghasilkan citra sintetis ultrasonografi jantung janin. Kualitas citra sintetis yang dihasilkan akan dievaluasi menggunakan metrik *Fréchet Inception Distance* (FID) sebagai indikator kesesuaian distribusi citra sintetis terhadap citra asli. Selain FID, metrik *Kernel Inception Distance* (KID) digunakan untuk mengukur kesamaan distribusi tanpa asumsi *Gaussian*, sedangkan *Learned Perceptual Image Patch Similarity* (LPIPS) digunakan untuk menilai kesamaan perseptual antar citra berdasarkan fitur visual tingkat tinggi. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan metode generatif untuk citra medis, khususnya sebagai solusi terhadap keterbatasan dataset.

II. METODOLOGI

Bagian ini menjelaskan rancangan penelitian yang digunakan untuk menghasilkan citra ultrasonografi jantung janin sintetis menggunakan model DCGAN. Proses penelitian terdiri dari beberapa tahapan utama, meliputi pengumpulan dan pra-pemrosesan data, perancangan arsitektur model, pelatihan, serta evaluasi kualitas citra hasil generasi. Alur menyeluruh proses pelatihan dan evaluasi DCGAN ditunjukkan pada Gambar 1, yang memperlihatkan hubungan antar tahap mulai dari pengolahan data hingga pengukuran performa model menggunakan metrik kuantitatif dan analisis visual.



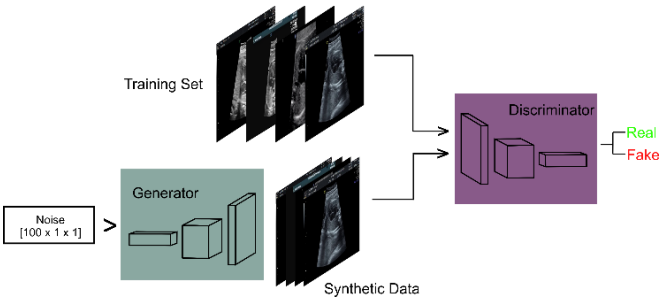
Gambar 1. *Diagram alur proses pelatihan dan evaluasi DCGAN.*

A. *Arsitektur GAN*

GAN merupakan pendekatan *deep learning* yang digunakan untuk menghasilkan data sintetis yang menyerupai data asli. GAN telah banyak digunakan di berbagai modalitas citra medis karena kemampuannya menghasilkan data sintetis tanpa harus secara eksplisit memodelkan distribusi data [6]. Dalam domain ultrasound janin, menunjukkan bahwa GAN dapat membantu klasifikasi citra plane otak janin dengan meningkatkan kinerja model ketika dataset asli terbatas [7].

Arsitektur GAN terdiri dari dua komponen utama, yaitu *Generator* (G) dan *Discriminator* (D), yang dilatih secara bersaing dalam suatu proses yang disebut *adversarial training*. *Generator* berperan menghasilkan citra baru dari input acak (*noise vector*), sedangkan *Discriminator* bertugas membedakan antara citra asli dan citra sintesis. Keduanya dilatih secara bersamaan, *Generator* berusaha menipu *Discriminator* dengan menghasilkan citra yang realistis, sementara *Discriminator* belajar mengenali perbedaan antara citra asli dan citra palsu [8].

Ilustrasi umum arsitektur GAN ditunjukkan pada Gambar 2, di mana *Generator* menghasilkan citra sintetis dari random noise yang kemudian dievaluasi oleh *Discriminator* untuk menentukan apakah citra tersebut tergolong nyata atau palsu.



Gambar 2. *Arsitektur Umum GAN.*

Melalui proses pelatihan berulang, *Generator* secara bertahap mempelajari pola distribusi data sehingga mampu menghasilkan citra yang menyerupai karakteristik data asli. Dalam bidang analisis citra medis, model ini dapat dimanfaatkan untuk memperbanyak data pelatihan, meningkatkan kualitas citra, atau mengurangi ketidakseimbangan kelas pada dataset [3].

B. *Arsitektur DCGAN*

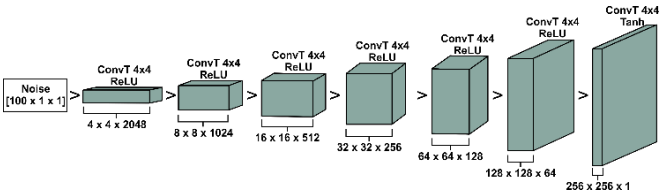
DCGAN adalah pengembangan dari arsitektur GAN konvensional yang mengganti lapisan *fully connected* dengan lapisan konvolusi dan *transposed convolution*, *plus batch normalization* dan nonlinearitas *ReLU/LeakyReLU* agar pelatihan lebih stabil [9]. Contoh aplikasinya adalah pada sintesis citra ultrasound payudara, DCGAN dilatih selama berbagai *epoch* (50, 100, 200, 500, dan 1000) dan dievaluasi oleh dua radiolog untuk keseluruhan kualitas citra, definisi struktur anatomi, dan visualisasi massa, serta kemampuan membedakan citra sintetis vs asli [10]. Hasil menunjukkan bahwa gambar yang dihasilkan pada *epoch* 500 dan 1000 memiliki kualitas tinggi, dan sejumlah citra sintetis sulit dibedakan dari citra asli menurut evaluasi pakar [10].

Arsitektur DCGAN terdiri atas dua komponen utama, yaitu *Generator* dan *Discriminator*, yang bekerja secara bersaing dalam proses pelatihan *adversarial*. *Generator* berusaha menghasilkan citra sintetis yang menyerupai citra asli, sedangkan *Discriminator* dirancang untuk membedakan citra asli dari hasil generasi.

Generator menerima input berupa *random noise vector z* yang diambil dari distribusi *Gaussian* atau *uniform*, kemudian mengubahnya menjadi citra sintetis melalui beberapa lapisan *transposed convolution*. Setiap lapisan diikuti oleh *Batch Normalization* dan fungsi aktivasi *ReLU*, kecuali lapisan keluaran yang menggunakan fungsi *Tanh* agar nilai piksel berada dalam rentang $[-1,1]$ [11]. Secara matematis, proses generasi dapat dinyatakan di persamaan (1).

$$G(z) = Tanh(W_n * ReLU(\ldots ReLU(W_1 * z + b_1)\ldots) + b_n) \quad (1)$$

di mana $W_i$dan $b_i$masing-masing merepresentasikan bobot dan bias pada lapisan ke-i, sedangkan tanda * menunjukkan operasi konvolusi. Struktur umum *Generator* ditunjukkan pada Gambar 3, yang memperlihatkan proses transformasi dari vektor acak berdimensi rendah menuju citra sintetis berdimensi tinggi.

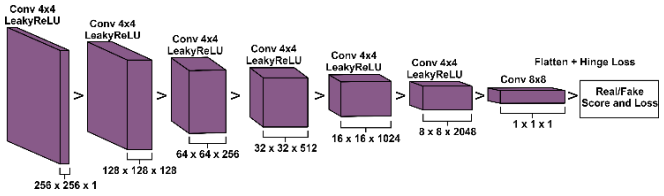


Gambar 3. *Arsitektur Generator pada DCGAN.*

Discriminator merupakan jaringan konvolusi yang berfungsi menilai apakah suatu citra tergolong asli atau hasil generasi. Lapisan-lapisan konvolusi mengekstraksi fitur visual secara hierarkis, diikuti fungsi aktivasi *LeakyReLU* untuk menjaga aliran gradien selama pelatihan. Lapisan akhir menggunakan fungsi aktivasi *Sigmoid* untuk menghasilkan probabilitas antara 0 dan 1 [12]. Secara matematis, fungsi *Discriminator* dirumuskan di (2).

$$D(x) = \sigma(W_m * f(\ldots f(W_1 * x + b_1)\ldots) + b_m) \quad (2)$$

dengan $f$adalah fungsi *LeakyReLU* dan σmenunjukkan fungsi *Sigmoid*. Nilai keluaran $D(x)$ mendekati 1 jika citra dikenali sebagai data asli, dan mendekati 0 jika citra hasil *Generator* [13]. Arsitektur *Discriminator* ditunjukkan pada Gambar 4, yang menggambarkan proses ekstraksi fitur bertingkat dari citra masukan hingga keputusan klasifikasi akhir.



Gambar 4. *Arsitektur Discriminator pada DCGAN.*

III. EKSPERIMEN DAN HASIL

Bagian ini memaparkan hasil eksperimen berdasarkan metodologi yang dijelaskan sebelumnya. Model DCGAN dilatih secara terpisah untuk setiap pasien guna menghasilkan citra ultrasonografi jantung janin sintetis dengan mempertimbangkan variasi anatomi dan jumlah data. Pelatihan dilakukan menggunakan konfigurasi arsitektur dan parameter yang seragam agar perbandingan hasil antar pasien tetap objektif. Citra sintetis kemudian dievaluasi menggunakan tiga metrik utama FID, KID, dan LPIPS serta melalui analisis visual pola anatomi. Evaluasi ini bertujuan menilai kemampuan DCGAN mereplikasi karakteristik citra ultrasonografi janin secara realistis dan mengamati variasi hasil antar pasien dengan kondisi jantung bawaan berbeda.

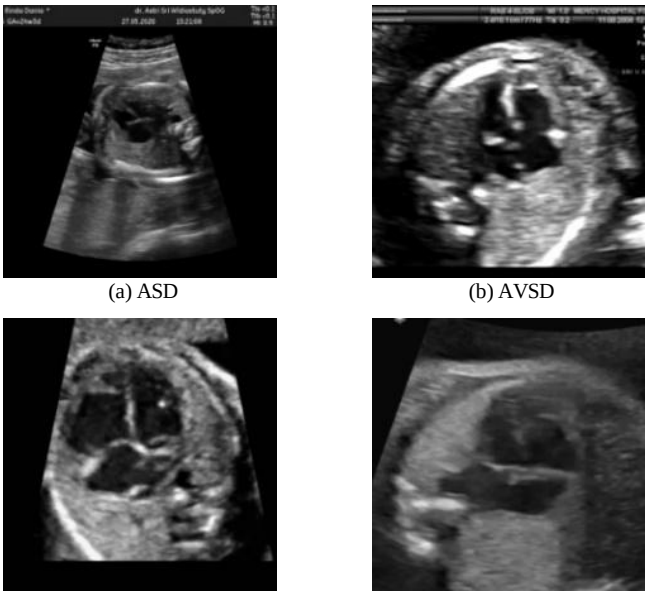
A. Dataset dan Pra-pemrosesan

Dataset yang digunakan pada penelitian ini terdiri dari citra ultrasonografi jantung janin dengan empat kategori kondisi jantung, yaitu *Atrial Septal Defect* (ASD), *Atrioventricular Septal Defect* (AVSD), *Ventricular Septal Defect* (VSD), dan NORMAL. Total citra yang digunakan ditampilkan pada Tabel 1, dengan jumlah bervariasi untuk setiap pasien. Ukuran awal setiap citra adalah 400 × 300 piksel dengan mode warna RGB. Posisi dan orientasi jantung pada citra berbeda-beda karena variasi anatomi dan sudut pengambilan gambar pada setiap pasien, sehingga menambah kompleksitas proses pembelajaran model generatif [14].

Tabel 1. *Jumlah citra ultrasonografi jantung janin berdasarkan kategori kondisi jantung*

Kategori Kondisi Jantung	Pasien	Total Gambar
ASD	AB6_ASD	290
	AB16_ASD	77
AVSD	AB10_AVSD	839
	AB11_AVSD	1157
NORMAL	NR2	351
	NR10	298
VSD	AB19_VSD	2472
	AB28_VSD	498

Tahapan pra-pemrosesan dilakukan untuk menyeragamkan karakteristik citra sebelum digunakan dalam pelatihan model DCGAN. Langkah-langkah yang diterapkan meliputi konversi ke *grayscale* untuk mengurangi kompleksitas warna, *resizing* ke ukuran seragam 256 × 256 piksel, dan *cropping* pada area fokus jantung untuk menghilangkan bagian luar yang tidak relevan. Selanjutnya dilakukan normalisasi nilai piksel ke rentang $[-1,1]$ agar sesuai dengan fungsi aktivasi *tanh* pada *Generator*. Selain itu, dilakukan augmentasi berupa manipulasi *brightness* dan *contrast* guna meningkatkan variasi visual dataset dan mengurangi risiko *overfitting* selama pelatihan [15]. Contoh citra hasil pra-pemrosesan dari masing-masing kategori kondisi jantung ditunjukkan pada Gambar 5.



(a) ASD (b) AVSD (c) NORMAL (d) VSD

Gambar 5. *Contoh citra hasil pra-pemrosesan untuk masing-masing kategori kondisi jantung janin.*

B. Pelatihan DCGAN

Pelatihan DCGAN dilakukan secara terpisah untuk setiap pasien menggunakan arsitektur dan parameter pelatihan yang sama. Citra hasil pra-pemrosesan dari masing-masing kategori kondisi jantung digunakan sebagai data masukan, dengan ukuran tetap 256 × 256 piksel dan format *grayscale*. Proses pelatihan dilakukan menggunakan *batch size* 16 selama maksimum 200 *epoch*, dengan optimisasi *Adam* ($\beta_1 = 0.0$, $\beta_2 = 0.9$) serta laju pembelajaran sebesar 2×10^{-4} untuk *Generator* dan *Discriminator*. Inisialisasi bobot menggunakan distribusi normal dengan deviasi standar 0.02, dan fungsi aktivasi utama yang digunakan adalah *ReLU* pada *Generator* serta *LeakyReLU* pada *Discriminator*.

Sejumlah penelitian telah memanfaatkan kombinasi fungsi *loss* untuk meningkatkan stabilitas pelatihan model generatif. Sebagai contoh, MelGAN menerapkan hinge loss bersama feature matching loss [16]. Sementara itu, U-Patch GAN menggunakan kombinasi feature matching loss dan *VGG-16-based perceptual loss* [17]. Dalam penelitian ini digunakan kombinasi ketiga komponen *hinge loss*, *perceptual loss* berbasis VGG-16, dan *feature matching loss* dengan tujuan meningkatkan stabilitas dan kualitas hasil generasi. Selain itu, *instance noise* ditambahkan pada citra nyata dan hasil generasi untuk mengurangi osilasi gradien selama proses *adversarial* [18]. Model *Generator* menghasilkan citra sintetis dari vektor laten berdimensi 128, sedangkan *Discriminator* dilengkapi dengan *Spectral Normalization* guna mencegah *gradient explosion* [19]. Mekanisme *early stopping* diterapkan untuk

menghentikan pelatihan ketika tidak terjadi peningkatan signifikan terhadap nilai loss gabungan (*combined loss*) selama sepuluh *epoch* berturut-turut [20]. Karena jumlah data tiap pasien bervariasi, proses pelatihan berhenti pada *epoch* yang berbeda-beda, umumnya mencapai konvergensi antara *epoch* ke-60 dan ke-110, tergantung kompleksitas data pasien. Seluruh proses pelatihan per pasien rata-rata memerlukan waktu beberapa jam, bergantung pada jumlah citra dan kompleksitas data masing-masing pasien.

C. Evaluasi Kualitas Citra Sintesis

Evaluasi terhadap citra sintesis dilakukan untuk menilai tingkat kemiripan dan realisme citra sintetis yang dihasilkan oleh model DCGAN dibandingkan dengan citra ultrasonografi jantung janin asli. Pada penelitian ini digunakan tiga metrik utama, yaitu FID, KID, dan LPIPS. Ketiga metrik ini dipilih karena mampu memberikan gambaran yang komprehensif terhadap kualitas citra dari sisi statistik, stabilitas, dan persepsi visual.

1) *Fréchet Inception Distance*

FID mengukur jarak antara distribusi fitur citra asli dan citra hasil generasi berdasarkan representasi dari jaringan Inception-V3. Semakin kecil nilai FID, semakin mirip distribusi kedua kelompok citra tersebut. Secara matematis, FID dirumuskan sebagai di persamaan (3).

$$FID = \| \mu_r - \mu_g \|^2 + Tr(\Sigma_r + \Sigma_g - 2\,(\Sigma_r \Sigma_g)^{\frac{1}{2}}) \quad (3)$$

dengan μ_r, Σ_r adalah *mean* dan kovarians dari fitur citra asli, sedangkan μ_g, $\Sigma_g$merupakan *mean* dan kovarians dari fitur output model. Nilai FID yang rendah menunjukkan bahwa citra buatan memiliki distribusi yang mendekati citra nyata secara global [21].

2) *Kernel Inception Distance*

KID merupakan varian dari FID yang menggunakan pendekatan *polynomial kernel* untuk mengukur kesamaan distribusi tanpa mengasumsikan bentuk *Gaussian*. Metrik ini lebih stabil ketika ukuran dataset kecil atau distribusi datanya tidak seimbang, seperti pada data ultrasonografi yang bervariasi antar pasien. Nilai KID yang lebih kecil menandakan kesamaan distribusi fitur yang lebih baik antara citra sintetis dan citra asli [22].

3) *Learned Perceptual Image Patch Similarity*

LPIPS menilai kesamaan perseptual antara dua citra dengan menghitung jarak antar fitur yang diekstraksi oleh jaringan konvolusional terlatih (misalnya VGG). Metrik ini lebih mencerminkan persepsi manusia dibanding kesamaan piksel mentah. Nilai LPIPS yang kecil menunjukkan citra buatan tampak lebih mirip secara visual dengan citra asli, terutama dalam hal tekstur dan kontras lokal yang penting pada citra medis [23].

Hasil perhitungan ketiga metrik evaluasi terhadap citra hasil DCGAN ditunjukkan pada Tabel 2.

Tabel 2. *Evaluasi DCGAN*

Pasien	FID	KID	LPIPS
AB6_ASD	136.3123	0.3547 ± 0.0073	0.3627 ± 0.0130
AB16_ASD	197.3654	0.4093 ± 0.0038	0.4687 ± 0.0063
AB10_AVSD	215.3582	0.3647 ± 0.0157	0.5481 ± 0.0302
AB11_AVSD	117.2991	0.3408 ± 0.0095	0.4211 ± 0.0103
NR2	98.0550	0.1759 ± 0.0041	0.3870 ± 0.0053
NR10	264.1900	0.6135 ± 0.0083	0.5502 ± 0.0057
AB19_VSD	136.5699	0.2991 ± 0.0027	0.3950 ± 0.0056
AB28_VSD	153.6776	0.3554 ± 0.0169	0.5587 ± 0.0120

Berdasarkan hasil evaluasi, nilai FID yang diperoleh bervariasi antara 98 hingga 264, menunjukkan bahwa kualitas distribusi citra sintetis berbeda-beda antar pasien. Sementara itu, nilai KID dan LPIPS relatif konsisten pada kisaran 0.17–0.61 dan 0.36–0.55, yang mengindikasikan bahwa meskipun model belum sepenuhnya meniru distribusi citra asli, citra hasil generasi masih mempertahankan kemiripan visual secara perseptual.

D. Analisis Visual

Analisis visual dilakukan untuk menilai kualitas hasil generasi DCGAN dengan membandingkannya terhadap citra ultrasonografi jantung janin asli pada setiap kategori kondisi jantung, seperti ditunjukkan pada Gambar 6. Secara umum, model DCGAN mampu menghasilkan citra sintetis dengan pola anatomi yang menyerupai bentuk jantung janin pada data asli. Struktur utama seperti ruang atrium dan ventrikel terlihat pada sebagian besar hasil generasi, meskipun ketajaman detailnya belum sepenuhnya setara dengan citra asli.



(a) Real ASD



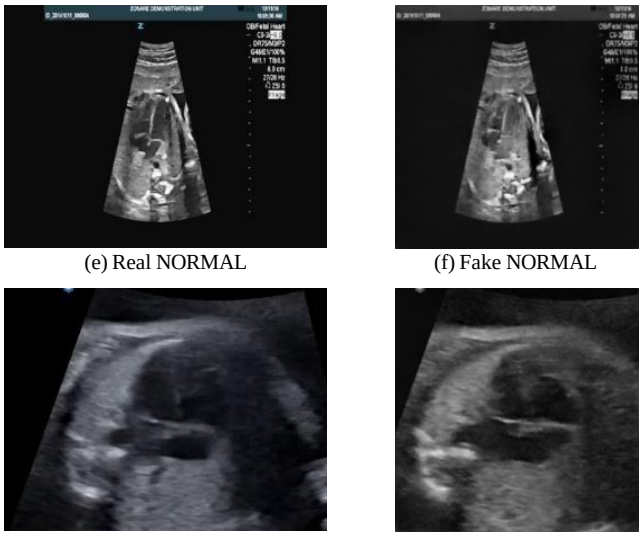
(b) Fake ASD



(c) Real AVSD



(d) Fake AVSD



(e) Real NORMAL (f) Fake NORMAL

(g) Real VSD (h) Fake VSD

Gambar 6. *Perbandingan citra ultrasonografi jantung janin asli dan hasil generasi DCGAN pada empat kategori kondisi jantung: (a–b) ASD, (c–d) AVSD, (e–f) NORMAL, dan (g–h) VSD.*

Output model pada kategori NORMAL dan VSD menunjukkan tekstur speckle khas ultrasonografi yang cukup realistis, dengan tingkat kontras dan pencahayaan yang relatif stabil. Sementara itu, pada kategori ASD dan AVSD masih ditemukan beberapa artefak visual dan area buram yang mengindikasikan ketidakstabilan selama proses pelatihan adversarial [24]. Perbedaan kualitas antar kategori ini dipengaruhi oleh variasi jumlah data dan kompleksitas bentuk anatomi jantung pada tiap pasien.

Meskipun demikian, hasil visual memperlihatkan bahwa DCGAN telah berhasil mempelajari distribusi pola intensitas dan karakteristik tekstur citra ultrasonografi. Citra sintetis yang dihasilkan dapat dianggap cukup representatif untuk menggambarkan bentuk jantung janin secara umum, serta menunjukkan potensi penggunaan model generatif ini dalam memperluas dataset pelatihan untuk penelitian medis berbasis citra.

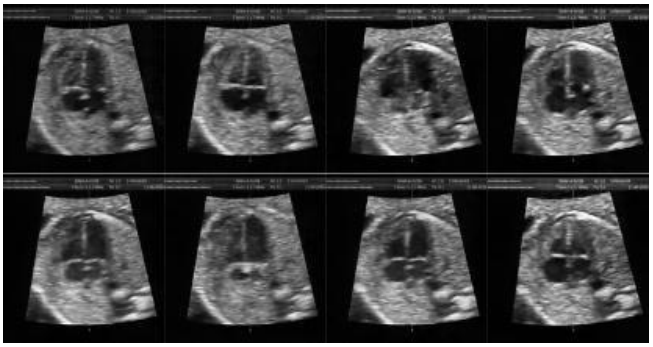
E. Pembahasan

Hasil generasi citra pada tiap pasien menunjukkan variasi kualitas yang dipengaruhi oleh jumlah data, kompleksitas anatomi, dan stabilitas pelatihan. Secara umum, model DCGAN mampu menghasilkan citra sintetis yang menyerupai pola anatomi jantung janin, meskipun belum sepenuhnya stabil pada seluruh pasien. Pada pasien AB6_ASD, anatomi empat ruang jantung masih dapat dikenali dengan jelas, menandakan bahwa *Generator* telah mempelajari bentuk global jantung dengan cukup baik. Namun, meskipun citra terlihat realistis, nilai FID pasien ini masih relatif tinggi, sementara KID dan LPIPS menunjukkan kesamaan perseptual yang baik. Perbedaan ini menandakan adanya ketidaksesuaian antara kualitas visual dan hasil kuantitatif.

Kondisi tersebut dipengaruhi oleh keterbatasan model *Inception-V3* yang digunakan dalam perhitungan FID. Model ini dilatih pada dataset citra alami (*ImageNet*), sehingga fitur yang diekstraksi kurang mewakili karakteristik ultrasonografi [5]. Akibatnya, meskipun citra sintetis terlihat mirip secara visual, distribusi fiturnya berbeda secara statistik dan menghasilkan nilai FID tinggi. Untuk memperoleh penilaian yang lebih seimbang, digunakan juga metrik KID dan LPIPS yang lebih stabil pada dataset kecil dan lebih sesuai dengan persepsi manusia terhadap tekstur serta pola lokal citra medis. Kombinasi ketiga metrik ini memberikan evaluasi yang lebih komprehensif terhadap kualitas citra hasil generatif.

Pola yang berbeda juga terlihat pada pasien lain dengan variasi jumlah data dan kompleksitas anatomi yang beragam. Pada pasien AB16_ASD, hasil generasi kurang stabil dengan beberapa area jantung tampak gelap atau buram serta variasi anatomi yang terbatas. Kondisi ini sejalan dengan jumlah data yang sangat sedikit, sehingga model belum mampu mempelajari pola intensitas dan tekstur secara menyeluruh. Sementara itu, pada pasien AB10_AVSD, citra sintetis tampak lebih beragam dan struktur jantung empat ruang masih dapat terlihat, meskipun terdapat sedikit ketidakteraturan pada tingkat pencahayaan. Nilai FID dan LPIPS yang sedang menunjukkan keseimbangan antara kualitas visual dan variasi tekstur yang masih perlu ditingkatkan.

Pada pasien AB11_AVSD, citra sintetis menunjukkan variasi anatomi yang cukup baik dengan hanya sedikit artefak kabur. Nilai FID dan KID tergolong moderat, menandakan bahwa model relatif stabil pada kategori ini. Gambar 7 menampilkan contoh hasil generasi pasien tersebut yang memperlihatkan keragaman visual antar sampel. Sebaliknya, pasien NR2 dan NR10 menunjukkan hasil yang paling tidak beragam. Kedua pasien ini menghasilkan citra dengan pola anatomi hampir identik di setiap keluaran, hanya berbeda sedikit pada tingkat kecerahan. Meskipun tampak jelas secara visual, rendahnya variasi menunjukkan terjadinya *mode collapse*, yaitu kondisi ketika *Generator* hanya mampu menghasilkan pola serupa berulang kali [25]. Fenomena ini sejalan dengan nilai FID dan KID yang rendah namun LPIPS sedang, menandakan bahwa kemiripan perseptual tinggi tidak selalu mencerminkan keragaman visual yang baik.



Gambar 7. *Contoh citra hasil generasi DCGAN untuk pasien AB11_AVSD yang menunjukkan variasi anatomi jantung dan tekstur speckle yang berbeda antar citra sintetis.*

Pada pasien AB19_VSD, pola serupa juga muncul di seluruh keluaran model. Walaupun memiliki jumlah data terbanyak, citra yang dihasilkan cenderung menampilkan bentuk anatomi yang sama dengan perbedaan intensitas dan kecerahan yang minimal. Kondisi ini mengindikasikan *mode*

collapse, di mana Generator terlalu fokus pada satu distribusi dominan dalam dataset [25]. Gambar 8 memperlihatkan contoh citra hasil generasi pasien tersebut yang menunjukkan pola berulang dengan variasi sangat kecil. Berbeda dengan itu, pasien AB28_VSD memperlihatkan variasi citra yang lebih beragam, meskipun sebagian hasil tampak kabur. Nilai LPIPS yang tinggi mengindikasikan adanya perbedaan perseptual yang besar antar citra, namun hal ini juga menunjukkan bahwa konsistensi anatomi belum sepenuhnya terjaga.



Gambar 8. *Contoh citra hasil generasi DCGAN untuk pasien AB19_VSD yang menunjukkan kemunculan pola anatomi seragam dan variasi visual yang rendah (mode collapse).*

Hasil yang diperoleh menunjukkan bahwa performa DCGAN sangat dipengaruhi oleh ukuran dan keragaman data pada setiap pasien. Nilai FID yang relatif tinggi menunjukkan bahwa distribusi citra sintetis belum sepenuhnya menyerupai citra asli, meskipun secara visual model telah mampu mengenali pola anatomi dasar jantung janin. Kekaburan citra serta rendahnya variasi anatomi umumnya muncul pada pasien dengan jumlah data terbatas, mengindikasikan bahwa kompleksitas dan keseimbangan data berperan penting dalam menjaga stabilitas pelatihan.

IV. KESIMPULAN

Penelitian ini menunjukkan bahwa kemampuan DCGAN dalam menghasilkan citra ultrasonografi jantung janin sintetis dipengaruhi oleh karakteristik data tiap pasien. Variasi hasil, di mana sebagian citra tampak realistis sementara lainnya kurang optimal, disebabkan oleh perbedaan jumlah data, kompleksitas anatomi, dan stabilitas pelatihan. Nilai FID yang tinggi meski kualitas visual baik menunjukkan perbedaan distribusi fitur antara citra sintetis dan asli akibat penggunaan model *Inception-V3* dari domain citra alami. Sementara itu, nilai KID dan LPIPS yang lebih konsisten menunjukkan kesamaan perseptual yang baik antara citra sintetis dan citra asli.

Fenomena *mode collapse* pada beberapa pasien, seperti AB19_VSD dan NR2, menunjukkan bahwa *Generator* belum mampu merepresentasikan seluruh variasi data dan cenderung menghasilkan pola berulang. Untuk mengatasinya, penelitian selanjutnya dapat menerapkan strategi stabilisasi pelatihan seperti *label smoothing*, *minibatch discrimination*, atau *feature matching loss* untuk menjaga keseimbangan antara *Generator* dan *Discriminator* [21].

Selain itu, arsitektur alternatif seperti WGAN-GP dan StyleGAN berpotensi digunakan untuk mengurangi *mode collapse* dan meningkatkan keragaman citra sintetis. WGAN-GP memanfaatkan fungsi kerugian berbasis jarak *Wasserstein* dengan penalti gradien untuk menjaga stabilitas pelatihan [26], sementara StyleGAN memperkenalkan mekanisme *style mapping* dan *adaptive instance normalization* yang memisahkan kontrol struktur global dan detail lokal, menghasilkan citra yang lebih beragam serta konsisten secara anatomis [27].

Ke depan, diperlukan penyesuaian khusus pada tahap pra-pemrosesan maupun arsitektur model untuk setiap kelompok data, agar model generatif dapat menyesuaikan diri dengan karakteristik visual yang unik pada tiap pasien dan menghasilkan citra sintetis yang lebih stabil serta konsisten secara anatomis. Penelitian lanjutan juga dapat mengeksplorasi varian arsitektur GAN lain, serta mengintegrasikan hasil generasi dengan model segmentasi jantung guna menilai konsistensi anatomi dan meningkatkan pemanfaatan citra sintetis dalam analisis klinis.

REFERENCES

[1] D. N. Sindhura, R. M. Pai, S. N. Bhat, and M. M. M. Pai, "A review of deep learning and Generative Adversarial Networks applications in medical image analysis," *Multimed Syst*, vol. 30, no. 3, p. 161, Jun. 2024, doi: 10.1007/s00530-024-01349-1.

[2] H. Wu *et al.*, "Application of Artificial Intelligence in Anatomical Structure Recognition of Standard Section of Fetal Heart," *Comput Math Methods Med*, vol. 2023, no. 1, Jan. 2023, doi: 10.1155/2023/5650378.

[3] A. Makhlouf, M. Maayah, N. Abughanam, and C. Catal, "The use of generative adversarial networks in medical image augmentation," *Neural Comput Appl*, vol. 35, no. 34, pp. 24055–24068, Dec. 2023, doi: 10.1007/s00521-023-09100-z.

[4] S. Zama *et al.*, "Clinical Utility of Breast Ultrasound Images Synthesized by a Generative Adversarial Network," *Medicina (B Aires)*, vol. 60, no. 1, p. 14, Dec. 2023, doi: 10.3390/medicina60010014.

[5] Y. Skandarani, P.-M. Jodoin, and A. Lalande, "GANs for Medical Image Synthesis: An Empirical Study," *J Imaging*, vol. 9, no. 3, p. 69, Mar. 2023, doi: 10.3390/jimaging9030069.

[6] M. H. Akpinar *et al.*, "Synthetic Data Generation via Generative Adversarial Networks in Healthcare: A Systematic Review of Image- and Signal-Based Studies," *IEEE Open J Eng Med Biol*, vol. 6, pp. 183–192, 2025, doi: 10.1109/OJEMB.2024.3508472.

[7] A. Montero, E. Bonet-Carne, and X. P. Burgos-Artizzu, "Generative Adversarial Networks to Improve Fetal Brain Fine-Grained Plane Classification," *Sensors*, vol. 21, no. 23, p. 7975, Nov. 2021, doi: 10.3390/s21237975.

[8] J. Fan, "Analysis of Generative Adversarial Networks (GANs) and Their Variants Based on Encoders and Decoders," INSTICC, Sep. 2024, pp. 343–348. doi: 10.5220/0012937900004508.

[9] M. F. Bin Hossen, S. Ummie, M. E. Rahman, M. N. A. Siddiky, and M. R. Rahman, "An Overview of Generative

Adversarial Networks and Their Variants," Apr. 01, 2025. doi: 10.20944/preprints202504.0071.v1.
[10] T. Fujioka *et al.*, "Breast Ultrasound Image Synthesis using Deep Convolutional Generative Adversarial Networks," *Diagnostics*, vol. 9, no. 4, p. 176, Nov. 2019, doi: 10.3390/diagnostics9040176.
[11] P. Sun, X. Liu, L. Weng, and Z. Liu, "Generative Adversarial Network Based on Self-Attention Mechanism for Automatic Page Layout Generation," *Applied Sciences*, vol. 15, no. 5, p. 2852, Mar. 2025, doi: 10.3390/app15052852.
[12] F. Mercaldo, F. Martinelli, and A. Santone, "Deep Convolutional Generative Adversarial Networks in Image-Based Android Malware Detection," *Computers*, vol. 13, no. 6, p. 154, Jun. 2024, doi: 10.3390/computers13060154.
[13] P. Purwono, A. N. E. Wulandari, A. Ma'arif, and W. A. Salah, "Understanding Generative Adversarial Networks (GANs): A Review," *Control Systems and Optimization Letters*, vol. 3, no. 1, pp. 36–45, Feb. 2025, doi: 10.59247/csol.v3i1.170.
[14] Y. Sang, A. Banerjee, M. Beetz, and V. Grau, "Deep conditional generative model for personalization of 12-lead electrocardiograms and cardiovascular risk prediction," *Front Digit Health*, vol. 7, Apr. 2025, doi: 10.3389/fdgth.2025.1558589.
[15] S. R. Brandigampala, A. F. Al-Battal, and T. Q. Nguyen, "Data Augmentation Methods For Object Detection and Segmentation In Ultrasound Scans: An Empirical Comparative Study," in *2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS)*, IEEE, Jul. 2022, pp. 288–291. doi: 10.1109/CBMS55023.2022.00057.
[16] K. Kumar *et al.*, "MelGAN: Generative Adversarial Networks for Conditional Waveform Synthesis."
[17] C. Fan, H. Lin, and Y. Qiu, "U-Patch GAN: A Medical Image Fusion Method Based on GAN.," *J Digit Imaging*, vol. 36, no. 1, pp. 339–355, Feb. 2023, doi: 10.1007/s10278-022-00696-7.
[18] S. Jenni and P. Favaro, "On Stabilizing Generative Adversarial Training with Noise," Sep. 2019, [Online]. Available: http://arxiv.org/abs/1906.04612
[19] Z. Lin, V. Sekar, and G. Fanti, "Why Spectral Normalization Stabilizes GANs: Analysis and Improvements," Apr. 2021.
[20] M. M. Saad, M. H. Rehmani, and R. O'Reilly, "Early Stopping Criteria for Training Generative Adversarial Networks in Biomedical Imaging," May 2024.
[21] M. M. Saad, R. O'Reilly, and M. H. Rehmani, "A survey on training challenges in generative adversarial networks for biomedical image analysis," *Artif Intell Rev*, vol. 57, no. 2, p. 19, Jan. 2024, doi: 10.1007/s10462-023-10624-y.
[22] E. Betzalel, C. Penso, and E. Fetaya, "Evaluation Metrics for Generative Models: An Empirical Study," *Mach Learn Knowl Extr*, vol. 6, no. 3, pp. 1531–1544, Jul. 2024, doi: 10.3390/make6030073.
[23] S. Abbasi, H. Lan, J. Choupan, N. Sheikh-Bahaei, G. Pandey, and B. Varghese, "Deep learning for the harmonization of structural MRI scans: a survey," *Biomed Eng Online*, vol. 23, no. 1, p. 90, Aug. 2024, doi: 10.1186/s12938-024-01280-6.
[24] J. Liang, H. Zeng, and L. Zhang, "Details or Artifacts: A Locally Discriminative Learning Approach to Realistic Image Super-Resolution," Mar. 2022.
[25] Y. Zhao, "GAN-based image generation," *Applied and Computational Engineering*, vol. 92, no. 1, pp. 128–135, Sep. 2024, doi: 10.54254/2755-2721/92/20241743.
[26] H. Anaya-Sánchez, L. Altamirano-Robles, R. Díaz-Hernández, and S. Zapotecas-Martínez, "WGAN-GP for Synthetic Retinal Image Generation: Enhancing Sensor-Based Medical Imaging for Classification Models," *Sensors*, vol. 25, no. 1, p. 167, Dec. 2024, doi: 10.3390/s25010167.
[27] A. H. Bermano *et al.*, "State-of-the-Art in the Architecture, Methods and Applications of StyleGAN," *Computer Graphics Forum*, vol. 41, no. 2, pp. 591–611, May 2022, doi: 10.1111/cgf.14503.