

**PEMBUATAN DATA SINTESIS DATASET ANAK AUTISME
MENGUNAKAN METODE *GENERATIVE ADVERSARIAL
NETWORKS* (GANS)**

TUGAS AKHIR

Diajukan Oleh:

MUHAMMAD ARIF

NIM. 200705088

Mahasiswa Fakultas Sains dan Teknologi

Program Studi Teknologi Informasi



**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI AR-RANIRY
BANDA ACEH
2025 M / 1447 H**

LEMBAR PERSETUJUAN

PEMBUATAN DATA SINTESIS DATASET ANAK AUTISME MENGUNAKAN METODE *GENERATIVE ADVERSARIAL NETWORKS* (GANS)

TUGAS AKHIR

Diajukan kepada Fakultas Sains dan Teknologi
Universitas Islam Negeri (UIN) Ar-Raniry Banda Aceh
Sebagai Salah Satu Beban Studi Memperoleh Gelar Sarjana (S1)
Dalam Program Studi Teknologi Informasi

Oleh:

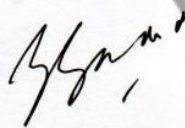
MUHAMMAD ARIF

200705088

**Mahasiswa Fakultas Sains dan Teknologi
Program Studi Teknologi Informasi**

Disetujui Untuk Dimunaqasyahkan Oleh:

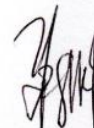
Pembimbing I,



Khairan AR. M.Kom

NIP.198607042014031001

Pembimbing II,

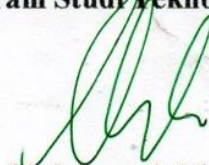


Raihan Islamadina, M.T

NIP.198901312020122011

Mengetahui,

Ketua Program Studi Teknologi Informasi



Malahayati, M.T

NIP.198301272015032003

LEMBAR PENGESAHAN

PEMBUATAN DATA SINTESIS DATASET ANAK AUTISME MENGUNAKAN METODE *GENERATIVE ADVERSARIAL NETWORKS* (GANS)

TUGAS AKHIR

Telah Diuji Panitia Ujian Munaqasah Tugas Akhir
Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh dan Dinyatakan Lulus
Serta Diterima Sebagai Salah Satu Beban Studi Program Sarjana (S-1)
Dalam Program Studi Teknologi Informasi

Pada Hari/Tanggal : Rabu, 20 Agustus 2025

26 Safar 1447 H

Di Darussalam, Banda Aceh

Panitia Ujian Munaqasah Tugas Akhir:

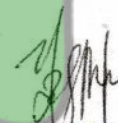
Ketua,



Khairan A.R. M. Kom

NIP.198607042014031001

Sekretaris,



Raihan Islamadina, M.T

NIP.198901312020122011

Penguji I,



Dr. Hendri Ahmadian, M.I.M

NIP.198301042014031002

Penguji II,



Sarini Vita Dewi, M.Eng

NIP.198301272015032003

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Ar-Raniry Banda Aceh




Prof. Dr. Ir. Muhammad Dirhamsyah, M.T., IPU

NIP.196210021988111001

LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini:

Nama : Muhammad Arif
NIM : 200705088
Program Studi : Teknologi Informasi
Fakultas : Sains dan Teknologi
Judul : Pembuatan Data Sintesis Dataset Anak Autisme
Menggunakan Metode Generative Adversarial Networks
(Gans)

Dengan ini menyatakan bahwa dalam penulisan tugas akhir ini, saya:

1. Tidak menggunakan ide orang lain tanpa mampu mengembangkan dan mempertanggungjawabkan;
2. Tidak melakukan plagiasi terhadap naskah karya orang lain;
3. Tidak menggunakan karya orang lain tanpa menyebutkan sumber asli atau tanpa izin pemilik karya;
4. Tidak memanipulasi dan memalsukan data;
5. Mengerjakan sendiri karya ini dan mampu bertanggungjawab atas karya ini.

Bila dikemudian hari ada tuntutan dari pihak lain atas karya saya, dan telah melalui pembuktian yang dapat dipertanggungjawabkan dan ternyata memang ditemukan bukti bahwa saya telah melanggar pernyataan ini, maka saya siap dikenai sanksi berdasarkan aturan yang berlaku di Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh.

Demikian pernyataan ini saya buat dengan sesungguhnya dan tanpa paksaan dari pihak manapun.

Banda Aceh, 20 Agustus 2025

Yang Menyatakan,



Muhammad Arif

ABSTRAK

Autism Spectrum Disorder (ASD) merupakan gangguan perkembangan saraf yang memengaruhi kemampuan sosial, komunikasi, serta perilaku individu sejak usia dini. Di Indonesia, jumlah anak dengan autisme terus meningkat, namun keterbatasan data yang valid dan representatif menjadi hambatan serius dalam pengembangan sistem diagnosis berbasis kecerdasan buatan. Salah satu jenis data penting dalam penelitian ini adalah gambar tulisan tangan anak autisme yang mencerminkan pola motorik halus dan kemampuan kognitif. Namun, proses pengumpulan data tersebut memiliki tantangan besar dari segi etika, biaya, dan ketersediaan. Penelitian ini bertujuan untuk mengatasi keterbatasan data dengan menghasilkan data sintesis menggunakan metode *Generative Adversarial Networks (GANs)*, yaitu arsitektur jaringan saraf yang terdiri dari dua komponen: *generator* dan *discriminator*. Melalui pelatihan yang bersifat kompetitif, GANs mampu menghasilkan data sintesis visual yang menyerupai data asli. Dengan pendekatan ini, diharapkan jumlah dan variasi dataset dapat ditingkatkan secara signifikan untuk mendukung pelatihan model klasifikasi dan analisis perilaku anak autisme. Hasil dari penelitian ini diharapkan dapat berkontribusi dalam pengembangan sistem pendukung diagnosis yang lebih efisien, hemat biaya, dan akurat. Selain itu, penggunaan data sintesis juga membuka peluang untuk riset lanjutan di bidang kesehatan anak dengan tetap memperhatikan validitas dan etika penggunaannya.

Kata Kunci: *Autisme, Data Sintesis, Tulisan Tangan, Generatif Adversarial Networks (GANs), Kecerdasan Buatan.*

ABSTRACT

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that affects an individual's ability to interact socially, communicate, and exhibit appropriate behaviors, typically emerging in early childhood. In Indonesia, the number of children with autism continues to rise, yet the lack of valid and representative data remains a significant barrier to developing artificial intelligence-based diagnostic systems. One crucial type of data in this context is handwriting samples from children with autism, which can reflect fine motor patterns and cognitive abilities. However, collecting such data presents major challenges in terms of ethics, cost, and availability. This study aims to address the data scarcity issue by generating synthetic data using the *Generative Adversarial Networks (GANs)* method, a neural network architecture consisting of two components: a generator and a discriminator. Through a competitive training process, GANs are capable of producing synthetic visual data that closely resemble real samples. This approach is expected to significantly expand the amount and diversity of available datasets, supporting the training of classification models and behavioral analysis of children with autism. The results of this study are expected to contribute to the development of more efficient, cost-effective, and accurate diagnostic support systems. Additionally, the use of synthetic data opens up new opportunities for further research in child health, while ensuring the validity and ethical considerations of its application.

Keywords: Autism, Synthetic Data, Handwriting, Generative Adversarial Networks (GANs), Artificial Intelligence.

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, taufik, dan hidayah-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul “*Pembuatan Data Sintesis Dataset Anak Autisme Menggunakan Metode Generative Adversarial Networks (GANs)*” sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi, UIN Ar-Raniry Banda Aceh.

Penyusunan skripsi ini merupakan hasil dari proses pembelajaran, penelitian, serta bimbingan dari berbagai pihak yang telah memberikan dukungan baik secara akademik maupun moral. Oleh karena itu, dengan penuh rasa hormat dan terima kasih, penulis menyampaikan apresiasi kepada:

1. Pembimbing I dan Pembimbing II yang telah meluangkan waktu, memberikan bimbingan, arahan, serta masukan yang sangat berarti selama proses penyusunan skripsi ini.
2. Bapak/Ibu dosen di lingkungan Program Studi Teknologi Informasi yang telah memberikan ilmu, wawasan, dan motivasi selama masa studi.
3. Kedua orang tua tercinta yang senantiasa memberikan doa, dukungan, serta semangat yang tak ternilai harganya.
4. Rekan-rekan mahasiswa dan semua pihak yang turut memberikan kontribusi, baik secara langsung maupun tidak langsung, dalam penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan dan keterbatasan. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan untuk penyempurnaan karya ini di masa yang akan datang. Semoga skripsi ini dapat memberikan manfaat serta menjadi kontribusi positif dalam pengembangan ilmu pengetahuan, khususnya dalam bidang kecerdasan buatan dan teknologi kesehatan anak.

Banda Aceh, 07 Agustus 2025
Yang Menyatakan,

DAFTAR ISI

COVER	1
LEMBAR PERSETUJUAN	i
LEMBAR PENGESAHAN	3
LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR	4
LEMBAR ABSTRAK	5
KATA PENGANTAR	6
DAFTAR ISI	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
DAFTAR LAMPIRAN	x
BAB I PENDAHULUAN	1
I.1 Latar Belakang	1
I.2 Rumusan Masalah	2
I.3 Tujuan Penelitian	3
I.4 Batasan Masalah	3
I.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	4
II.1 Penelitian Terdahulu	4
II.2 Kajian Teoritis	7
II.2.1 Pengertian Autisme	7
II.2.2 Pengertian Data Sintesis	7
II.2.3 Etika dan Privasi dalam Penggunaan Data Sintesis	8
II.2.4 Deteksi Gambar Tulisan Dan Teknik Pengolahan Citra	8
II.2.5 Deep Learning	9
II.2.6 CNN (Convolutional Neural Network)	10
II.2.7 Generative Adversarial Networks (GANs)	11
A. Varian GANs	11
B. GANs Untuk Data Sintesis	12
C. Proses Sintesis Data dengan GANs	13
D. Keuntungan Penggunaan GAN untuk Data Sintesis	13
E. Tantangan dalam Menggunakan GAN untuk Sintesis Data	14
II.2.8 Perbandingan GAN Dengan Teknik Lain	14
II.2.9 Evaluasi Kualitas Data Sintesis dengan Data Asli	15
BAB III METODELOGI PENELITIAN	17
III.1 Jenis Penelitian	17

III.2 Metode Penelitian.....	17
III.3 Tahapan Penelitian	17
III.3.1 Studi Literatur.....	18
III.3.2 Pengumpulan Dataset Asli Tulisan Anak Autisme.....	18
III.3.3 <i>Preprocessing</i> Data.....	20
III.3.4 Implementasi Model GAN (<i>Generative Adversarial Networks</i>)	20
A. Arsitektur <i>Generator</i>	21
B. Arsitektur <i>Discriminator</i>	21
III.3.5 Pelatihan Model GAN	22
III.3.6 Evaluasi Hasil Data Sintesis	25
III.3.7 Analisis Dan Visualisasi Hasil.....	27
III.4 <i>Flowchart</i> Alur Sistem	28
III.5 Kebutuhan Perangkat Keras Dan Lunak.....	31
BAB IV HASIL DAN PEMBAHASAN.....	34
VI.1 Hasil Implementasi.....	34
A. <i>Setup</i> dan Ekstraksi Dataset	34
B. Import Library dan Reprocessing Dataset.....	37
C. Pembangunan Arsitektur GAN	42
D. Inisialisasi Model & <i>Optimizer</i>	47
E. Pelatihan Model.....	48
F. Evaluasi Gans	55
IV.2 Pembahasan.....	76
A. Kualitas Visual Gambar Sintetik	76
B. Evaluasi Stabilitas Pelatihan dari <i>Kurva Loss</i>	77
C. Evaluasi Klasifikasi Menggunakan <i>Confusion Matrix</i>	77
D. Analisis Distribusi Fitur Menggunakan <i>t-SNE</i>	77
E. Evaluasi Kualitas Warna dengan <i>Histogram RGB</i>	78
F. Evaluasi Kemiripan Struktural Menggunakan <i>SSIM</i>	78
BAB V KESIMPULAN DAN SARAN	80
V.1 Kesimpulan.....	80
V.2 Saran	80
DAFTAR PUSTAKA.....	83
DAFTAR LAMPIRAN	81

DAFTAR GAMBAR

Gambar II. 1	Arsitektur CNN	11
Gambar II. 2	Arsitektur GAN.....	13
Gambar III. 1	Dataset Anak Autisme	19
Gambar III. 2	Dataset Anak Autisme	19
Gambar III. 3	Dataset Anak Autisme	20
Gambar III. 4	Proses Train Loop Saat Generate Per Epoch	23
Gambar III. 5	Flowchart Alur Sistem.....	29
Gambar IV. 1	Hasil Ekstraksi Perkelas	37
Gambar IV. 2	Preveie Hasil Ekstraksi Perkelas	37
Gambar IV. 3	Hasil Ekstraksi Dihitung dan diberi Label	41
Gambar IV. 4	Hasil Visualisasi Loss Generator dan Diskriminator	54
Gambar IV. 5	Kurva Loss Generator dan Diskriminator	54
Gambar IV. 6	Hasil Generate Data Sintesis Saat di Visualiasasikan	56
Gambar IV. 7	Hasil Generate Data Sintesis Saat di Visualiasasikan	56
Gambar IV. 8	Hasil Generate Data Sintesis Saat di Visualiasasikan	56
Gambar IV. 9	Hasil Generate Data Sintesis Saat di Visualiasasikan	56
Gambar IV. 10	Hasil Generate Data Sintesis Saat di Visualiasasikan	56
Gambar IV. 11	Hasil Generate Data Sintesis Saat di Visualiasasikan	57
Gambar IV. 12	Hasil Generate Data Sintesis Saat di Visualiasasikan	57
Gambar IV. 13	Hasil Generate Data Sintesis Saat di Visualiasasikan	57
Gambar IV. 14	Hasil Generate Data Sintesis Saat di Visualiasasikan	57
Gambar IV. 15	Hasil Generate Data Sintesis Saat di Visualiasasikan	57
Gambar IV. 16	Hasil Generate Data Sintesis Saat di Visualiasasikan	58
Gambar IV. 17	Hasil Generate Data Sintesis Saat di Visualiasasikan	58
Gambar IV. 18	Hasil Generate Data Sintesis Saat di Visualiasasikan	58
Gambar IV. 19	Hasil Generate Data Sintesis Saat di Visualiasasikan	58
Gambar IV. 20	Hasil Generate Data Sintesis Saat di Visualiasasikan	58
Gambar IV. 21	Hasil Generate Data Sintesis Saat di Visualiasasikan	59
Gambar IV. 22	Hasil Generate Data Sintesis Saat di Visualiasasikan	59
Gambar IV. 23	Hasil Generate Data Sintesis Saat di Visualiasasikan	59
Gambar IV. 24	Hasil Generate Data Sintesis Saat di Visualiasasikan	59
Gambar IV. 25	Hasil Generate Data Sintesis Saat di Visualiasasikan	60
Gambar IV. 26	Hasil Sintesis Per 10 Epoch.....	67
Gambar IV. 27	Hasil Evaluasi Confusion Matrix	68
Gambar IV. 28	Akurasi Perkelas Confusion Matrix	70
Gambar IV. 29	Evaluasi Visulisasi t-SNE	72
Gambar IV. 30	Hasil Nilai SSIM	74
Gambar IV. 31	Evaluasi Hasil Histogram RGB Real dan Fake	75

DAFTAR TABEL

Tabel II. 1	Penelitian Terdahulu.....	6
Tabel II. 2	Perbandingan GAN Dengan Teknik Lain.....	14
Tabel III. 1	Perangkat Keras (Hardware)	31
Tabel III. 2	Perangkat Lunak (Software).....	32
Tabel III. 3	Library Tambahan	32
Tabel IV. 1	Penjelasan Hasil Evaluasi Confusion Matrix	68
Tabel IV. 2	Penjelasan Akurasi Perkelas Confusion Matrix	71



DAFTAR LAMPIRAN

Codingan Setup & Ekstraksi Dataset	88
Codingan Import Library & Persiapan Dataset	89
Codingan DcGAN Arsitektur (U-Net Style Generator) untuk Generator	90
Codingan DcGAN Arsitektur untuk Discriminator (ResNet-style).....	91
Codingan Inisialisasi Model & Optimizer	92
Codingan Training Loop GAN.....	92
Codingan Evaluasi Plotly Kurva Loss G dan D	94
Codingan Dowload Gambar Hasil Sintesis Tiap Epoch.....	94
Codingan Evaluasi Visualisasi Gambar Sintetik Keseluruhan Terbaik	95
Codingan Evaluasi Tampilkan Visualisasi Untuk Perkelas, Confusion Matrix + Akurasi Perkelas, SSIM, Visualisasi t-SNE + PCA.....	95
Codingan Evaluasi Histogram Warna Perkelas Real/Sintesis Berdampingan	99



BAB I

PENDAHULUAN

I.1 Latar Belakang

Autism Spectrum Disorder (ASD) atau Gangguan Spektrum Autisme merupakan salah satu gangguan perkembangan saraf yang memengaruhi cara individu berinteraksi secara sosial, berkomunikasi, serta mengekspresikan perilaku dan emosinya. Gangguan ini bersifat jangka panjang dan biasanya terdeteksi sejak usia dini, bahkan sebelum anak menginjak usia tiga tahun. ASD memiliki spektrum yang luas, artinya gejala dan tingkat keparahannya dapat sangat bervariasi antar individu.

Di Indonesia, kasus autisme menunjukkan peningkatan yang signifikan dari tahun ke tahun. Berdasarkan pernyataan Wakil Menteri Kesehatan RI pada tahun 2024, jumlah anak dengan autisme diperkirakan mencapai lebih dari 2,4 juta jiwa. Namun, data nasional yang valid dan menyeluruh masih sulit ditemukan. Hal ini disebabkan oleh kurangnya sistem pendataan yang konsisten, serta masih minimnya kesadaran masyarakat dan tenaga medis terhadap pentingnya deteksi dan diagnosis dini autisme. Keterbatasan data tersebut menjadi kendala serius dalam pengembangan berbagai riset dan teknologi, khususnya sistem diagnosis berbasis kecerdasan buatan (*Artificial Intelligence*).

Salah satu bentuk data yang sangat bermanfaat dalam penelitian autisme adalah gambar tulisan tangan anak autisme, yang dapat merepresentasikan pola motorik halus, gaya kognitif, serta kemampuan koordinasi visual-motorik anak. Data ini sering digunakan untuk menganalisis karakteristik perkembangan dan perilaku anak. Sayangnya, proses pengumpulan data semacam ini sangat tidak mudah. Selain membutuhkan waktu dan biaya yang besar, pengumpulan data juga harus melalui prosedur yang etis dan tidak mengganggu kondisi psikologis subjek. Oleh karena itu, jumlah data yang tersedia sering kali sangat terbatas, tidak seimbang, dan tidak representatif, sehingga dapat memengaruhi akurasi model yang dikembangkan dan menimbulkan risiko *overfitting*.

Sebagai solusi atas keterbatasan tersebut, pendekatan melalui pembuatan *data sintesis* menjadi salah satu metode yang mulai banyak dikembangkan dalam dunia penelitian, terutama yang melibatkan penerapan *machine learning*. Salah satu metode paling populer dan efektif dalam menghasilkan data sintesis visual adalah *Generative Adversarial Networks (GANs)*. GANs merupakan arsitektur jaringan saraf yang terdiri

dari dua komponen utama, yaitu *generator* yang bertugas menciptakan data baru dan *discriminator* yang bertugas mengevaluasi apakah data tersebut asli atau palsu. Proses pelatihan GANs bersifat kompetitif, di mana kedua jaringan tersebut saling memperbaiki performa satu sama lain hingga menghasilkan data yang sangat mirip dengan data asli.

Penelitian ini secara khusus memanfaatkan metode GANs untuk menghasilkan data sintesis berupa gambar tulisan tangan anak autisme. Tujuan dari pembuatan data sintesis ini adalah untuk memperluas jumlah dan variasi dataset yang tersedia, sehingga dapat mendukung pelatihan model klasifikasi atau analisis perilaku anak autis secara lebih optimal untuk citra medis. Selain itu, data sintesis juga dapat membantu peneliti mengurangi biaya dan waktu dalam pengumpulan data, serta meminimalkan potensi bias yang mungkin terjadi jika hanya mengandalkan data asli dari jumlah sampel terbatas.

Penggunaan GANs dalam konteks ini sangat relevan karena memiliki beberapa keunggulan, seperti kemampuan untuk belajar tanpa label, menghasilkan data yang bervariasi dan realistis, serta fleksibel dalam berbagai jenis input visual. Dengan memanfaatkan keunggulan ini, diharapkan proses penelitian dan pengembangan sistem pendukung diagnosis anak dengan autisme dapat dilakukan secara lebih efisien, cepat, dan akurat. Meski demikian, tetap perlu diperhatikan aspek validitas dan etika dalam penggunaan data sintesis, terutama dalam ranah kesehatan anak.

Berdasarkan latar belakang tersebut, peneliti mengangkat judul “*Pembuatan Data Sintesis Dataset Anak Autisme Menggunakan Metode Generative Adversarial Networks (GANs)*” yang diharapkan mampu memberikan kontribusi nyata dalam pengembangan sistem berbasis AI untuk mendukung diagnosis, terapi, maupun penelitian lanjutan mengenai anak-anak dengan autisme di Indonesia.

I.2 Rumusan Masalah

Rumusan masalah pada penelitian ini dapat diuraikan sebagai berikut

1. Bagaimana cara menghasilkan data sintesis yang *realistis* dan *representatif* berdasarkan tulisan tangan anak autis yang terbatas dengan menggunakan metode GANs?
2. Bagaimana memastikan validitas dan kualitas data sintesis yang dihasilkan dengan mengevaluasi melalui Visualisasi gambar sintesis, t-SNE+PCA Confusion Matrix+Akurasi perkelas, Histogram RGB dan SSIM

Perumusan masalah ini diharapkan dapat menjadi dasar dalam proses pembuatan data sintesis berdasarkan dataset tulisan tangan anak autisme dengan menggunakan metode *Generative Adversarial Networks (GANs)* secara efektif dan terarah.

I.3 Tujuan Penelitian

1. Menggunakan metode GANs dapat menghasilkan data sintesis yang realistis dan representatif berdasarkan tulisan tangan anak autisme yang terbatas.
2. Memastikan validitas dan kualitas data sintesis yang dihasilkan dengan mengevaluasi melalui Visualisasi gambar sintesis, t-SNE+PCA Confusion Matrix+Akurasi perkelas, Histogram RGB dan SSIM.

I.4 Batasan Masalah

Pada proses pembuatan data sintesis anak autisme memiliki batasan masalah sebagai berikut:

- a) Membuat data sintesis berdasarkan dataset anak autisme menggunakan metode *Generative Adversarial Networks (GANs)*
- b) Menggunakan *Google Collab* dan bahasa pemrograman *Python*.
- c) Gambar tulisan tangan anak autisme sebagai acuan untuk pembuatan data sintesis.

I.5 Manfaat Penelitian

- a) Bagi Mahasiswa
Untuk mengaplikasikan serta meningkatkan ilmu yang diperoleh selama perkuliahan dan untuk menyelesaikan mata kuliah Tugas Akhir.
- b) Bagi Instansi/Profesi pengembang AI
Mempermudahkannya dalam mendapatkan dataset yang sulit untuk ditemukan sehingga dengan adanya data sintesis ini dapat membantu proses dalam pencarian dataset.