

Deteksi Stunting Menggunakan Dual-Vae, Gan, Dan Transformer Pada Data Antropometri

Wowon Priatna ^{1,*}, Joni Warta ¹, Agus Hidayat ¹, Adam Arif Budiman ²

¹ Fakultas Ilmu Komputer; Universitas Bhayangkara Jakarta Raya; e-mail: wowon.priatna@dsn.ubharajaya.ac.id, joni.warta@dsn.ubharajaya.ac.id, agus.hidayat@dsn.ubharajaya.ac.id

² Fakultas Teknik; Universitas Darma Persada; e-mail: adam_arif_budiman@ft.unsada.ac.id

* Korespondensi: e-mail: wowon.priatna@dsn.ubharajaya.ac.id

Submitted: 23/09/2026; Revised: 30/09/2026; Accepted: 21/10/2026; Published: 22/10/2026

Abstract

Stunting detection based on anthropometric data requires a classification model capable of capturing complex relationships among child growth characteristics. This study developed a stunting detection model integrating Dual Variational Autoencoder (Dual-VAE), Generative Adversarial Network (GAN), and Transformer using age, sex, weight, and height as predictors. The dataset consisted of 40,071 records of children aged 0–59 months from Jeneponto Regency, Indonesia. Two class-oriented VAE branches were used to learn complementary latent representations of stunted and non-stunted children. GAN-based enrichment was applied to the stunted latent representation, while the Transformer modeled interactions among the resulting latent features. Model performance was evaluated using stratified five-fold cross-validation. The proposed model achieved an accuracy of 0.9960 ± 0.0007 , F1-score of 0.9950 ± 0.0008 , G-Mean of 0.9957 ± 0.0007 , MCC of 0.9917 ± 0.0014 , AUROC of 0.9999 ± 0.0001 , and AUPRC of 0.9998 ± 0.0001 . The results showed that combining class-oriented latent representations, controlled generative enrichment, and feature-interaction modeling produced high and consistent classification performance for anthropometric-based stunting detection.

Keywords: Anthropometric data, Dual-vae, Gan, Stunting detection, Transformer

Abstrak

Deteksi stunting berbasis data antropometri memerlukan model klasifikasi yang mampu menangkap hubungan kompleks antar karakteristik pertumbuhan anak. Penelitian ini mengembangkan model deteksi stunting yang mengintegrasikan Dual Variational Autoencoder (Dual-VAE), Generative Adversarial Network (GAN), dan Transformer menggunakan usia, jenis kelamin, berat badan, dan tinggi badan sebagai prediktor. Dataset yang digunakan terdiri atas 40.071 data anak usia 0–59 bulan di Kabupaten Jeneponto, Indonesia. Dua cabang VAE berorientasi kelas digunakan untuk mempelajari representasi laten yang saling melengkapi antara anak stunting dan non-stunting. Pengayaan berbasis GAN diterapkan pada representasi laten kelas stunting, sedangkan Transformer digunakan untuk memodelkan interaksi antarfitur laten yang dihasilkan. Kinerja model dievaluasi menggunakan stratified five-fold cross-validation. Model memperoleh akurasi sebesar $0,9960 \pm 0,0007$, F1-score $0,9950 \pm 0,0008$, G-Mean $0,9957 \pm 0,0007$, MCC $0,9917 \pm 0,0014$, AUROC $0,9999 \pm 0,0001$, dan AUPRC $0,9998 \pm 0,0001$. Hasil penelitian menunjukkan bahwa kombinasi representasi laten berorientasi kelas, pengayaan generatif terkontrol, dan pemodelan interaksi fitur menghasilkan kinerja klasifikasi yang tinggi dan konsisten untuk deteksi stunting berbasis data antropometri.

Kata kunci: Data antropometri, Deteksi stunting, Dual-vae, Gan, Transformer

1. Pendahuluan

Stunting merupakan masalah pertumbuhan anak yang berkaitan dengan gangguan pertumbuhan linear pada periode awal kehidupan dan dapat berdampak terhadap perkembangan fisik maupun kognitif. Data antropometri seperti usia, jenis kelamin, berat badan, dan tinggi badan menyediakan informasi terukur yang dapat dimanfaatkan untuk mengenali pola pertumbuhan dan mendukung deteksi stunting. Pemanfaatan data tersebut semakin berkembang seiring penerapan machine learning dalam analisis kesehatan anak. Shen et al., (2023) menunjukkan bahwa sejumlah algoritma machine learning dapat digunakan untuk memprediksi stunting pada anak usia di bawah lima tahun, sedangkan Azriani et al., (2025) mengembangkan model prediksi stunting pada populasi anak Indonesia dengan memanfaatkan berbagai karakteristik anak dan keluarga. Pada penelitian sebelumnya, Khoirunnisaa et al., (2024) juga menggunakan algoritma FP-Growth untuk mengidentifikasi pola faktor yang berkaitan dengan stunting pada data Dinas Kesehatan Kota Bekasi.

Berbagai pendekatan machine learning telah dikembangkan untuk meningkatkan kemampuan deteksi stunting. Novalina et al., (2025) membandingkan beberapa algoritma klasifikasi dengan teknik undersampling dan SMOTE pada data Indonesia dan menunjukkan bahwa penanganan ketidakseimbangan data berpengaruh terhadap kinerja prediksi. Sugihartono et al., (2025) juga membandingkan XGBoost, Random Forest, SVM, dan k-NN dengan penerapan SMOTE pada data antropometri dan demografi. Sementara itu, Setiawan et al., (2025) mengembangkan pendekatan berbasis modalitas dengan penyeimbangan data dan reduksi dimensi untuk deteksi dini stunting. Meskipun memberikan kinerja klasifikasi yang baik, sebagian besar pendekatan tersebut masih melakukan pembelajaran langsung dari fitur masukan atau fitur hasil transformasi, sehingga representasi karakteristik stunting dan non-stunting belum dipelajari secara khusus dalam ruang laten yang berbeda.

Perkembangan deep learning membuka peluang untuk mempelajari representasi data secara lebih adaptif. Begashaw et al., (2025) menggunakan jaringan LSTM-FC untuk mengklasifikasikan dan memprediksi status gizi anak, sedangkan Hadikurniawati et al., (2025) menggabungkan ekstraksi fitur berbasis deep learning, indikator antropometri, dan mekanisme attention untuk prediksi stunting. Pendekatan yang lebih dekat dengan pembelajaran representasi probabilistik dikembangkan oleh Hadikurniawati (2026) melalui integrasi Variational Autoencoder (VAE), rekayasa fitur, dan XGBoost untuk klasifikasi data nutrisi. Hasil-hasil tersebut menunjukkan pergeseran dari klasifikasi langsung terhadap variabel antropometri menuju pembelajaran representasi dan interaksi antarfitur, tetapi representasi yang digunakan umumnya masih dibentuk dalam ruang laten bersama dan belum secara eksplisit memisahkan karakteristik kelas stunting dan non-stunting.

Pembelajaran representasi generatif memberikan alternatif untuk memperkaya informasi pada ruang laten sebelum dilakukan klasifikasi. Anshelevich & Katz, (2025) menunjukkan bahwa integrasi VAE dan GAN dapat digunakan untuk menghasilkan data tabular sintetis dengan mempertahankan karakteristik distribusi data, sedangkan Hadikurniawati,

(2026) menunjukkan manfaat VAE dalam memperoleh representasi laten pada data nutrisi. Di sisi lain, Luo et al., (2026) mengembangkan kerangka deep learning untuk identifikasi stunting melalui ekstraksi fitur, fusi fitur, representasi laten, dan mekanisme hierarchical attention. Namun, pendekatan-pendekatan tersebut belum mengintegrasikan pembentukan representasi probabilistik yang berorientasi kelas, pengayaan terkontrol pada ruang laten kelas stunting, dan pemodelan interaksi antarrepresentasi dalam satu kerangka klasifikasi.

Berdasarkan perkembangan tersebut, kesenjangan penelitian terletak pada belum terintegrasinya mekanisme yang membentuk representasi probabilistik secara terpisah untuk karakteristik stunting dan non-stunting, memperkaya representasi kelas stunting secara terkontrol, serta mempelajari interaksi antarfitur laten sebelum proses klasifikasi. Luo et al., (2026) masih melakukan klasifikasi dari variabel antropometri menggunakan beberapa algoritma machine learning, Hadikurniawati (2026) menggunakan satu representasi laten VAE yang dipadukan dengan XGBoost, sedangkan Luo et al., (2026) menitikberatkan pada ekstraksi, fusi fitur, dan mekanisme attention. Oleh karena itu, penelitian ini menerapkan Dual-VAE, GAN, dan Transformer untuk deteksi stunting pada data antropometri.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan komputasional kuantitatif untuk melakukan deteksi stunting berdasarkan data antropometri. Tahapan penelitian meliputi prapemrosesan data, pembentukan representasi laten menggunakan Dual-Variational Autoencoder (Dual-VAE), pengayaan representasi kelas stunting menggunakan Generative Adversarial Network (GAN), serta pemodelan interaksi antarfitur laten menggunakan Transformer. Seluruh proses prapemrosesan, pembelajaran representasi, pengayaan data, dan pemilihan parameter dilakukan hanya menggunakan data pelatihan untuk menghindari data leakage. Kinerja model dievaluasi menggunakan stratified five-fold cross-validation dengan beberapa metrik klasifikasi.

2.1. Data dan Prapemrosesan

Dataset yang digunakan adalah Stunting and Nutritional Status of Toddler from Jeneponto Regency, South Sulawesi, Indonesia yang terdiri atas 40.071 data anak berusia 0–59 bulan dan dikumpulkan pada periode 2021–2024 (RS et al., 2026). Dataset memuat data antropometri dan demografi anak, sedangkan penelitian ini menggunakan usia, jenis kelamin, berat badan, dan tinggi badan sebagai prediktor. Status stunting digunakan sebagai target klasifikasi biner yang terdiri atas kelas stunting dan non-stunting. Variabel identitas serta variabel turunan yang secara langsung merepresentasikan indeks antropometri, meliputi kategori dan Z-score tinggi badan menurut umur, berat badan menurut umur, serta berat badan menurut tinggi badan, dikeluarkan untuk mencegah kebocoran target. Dataset versi 4 tersebut dipublikasikan melalui Mendeley Data dan tetap berisi 40.071 observasi.

Sebelum proses pemodelan, data diperiksa untuk mengidentifikasi nilai hilang, duplikasi, nilai tidak valid, dan inkonsistensi pencatatan. Jenis kelamin dikodekan secara numerik, sedangkan variabel antropometri kontinu dinormalisasi menggunakan RobustScaler.

Parameter normalisasi dihitung hanya berdasarkan data pelatihan pada setiap fold, kemudian parameter yang sama diterapkan pada data validasi dan pengujian. Prosedur ini mengikuti rancangan eksperimen pada penelitian, yaitu memastikan bahwa informasi dari data validasi maupun pengujian tidak digunakan dalam pembentukan parameter prapemrosesan.

2.2. Model Duan-Vae-Gan-Transformer

Model mengintegrasikan Dual-VAE, GAN, dan Transformer dalam satu alur klasifikasi. VAE digunakan karena mampu mempelajari representasi probabilistik dengan memetakan data masukan ke distribusi laten yang direpresentasikan melalui parameter rata-rata dan varians (Hadikurniawati, 2026; Berahmand et al., 2024). Berbeda dengan penggunaan satu VAE, penelitian ini menggunakan dua VAE berorientasi kelas. VAE pertama dilatih menggunakan data non-stunting, sedangkan VAE kedua dilatih menggunakan data stunting sehingga masing-masing encoder mempelajari struktur probabilistik dari kelas yang berbeda. Prinsip VAE dan regularisasi distribusi laten melalui Kullback–Leibler divergence juga merupakan dasar model pada naskah penelitian awal.

Setelah kedua VAE selesai dilatih, setiap observasi dilewatkan melalui kedua encoder tanpa menggunakan label kelas sebenarnya pada tahap inferensi. Jika keluaran encoder non stunting adalah $z_i^{(0)}$ dan keluaran encoder stunting adalah $z_i^{(1)}$, maka representasi komplementer setiap observasi dibentuk melalui.

$$z_i = z_i^{(0)} || z_i^{(1)}$$

Dengan $||$ menunjukkan operasi konkatenasi. Representasi tersebut menggambarkan satu observasi antropometri terhadap dua ruang laten yang dipelajari dari karakteristik non-stunting dan stunting. Penggunaan representasi laten probabilistik untuk klasifikasi data nutrisi sebelumnya telah ditunjukkan oleh (Hadikurniawati, 2026), sedangkan kajian autoencoder menunjukkan kemampuannya dalam menghasilkan representasi data berdimensi lebih rendah (Berahmand et al., 2024).

Representasi komplementer kelas stunting pada data pelatihan kemudian digunakan untuk melatih GAN. Generator mempelajari distribusi laten dan menghasilkan representasi stunting tambahan, sedangkan discriminator membedakan representasi nyata dengan representasi hasil generasi. Pendekatan VAE-GAN telah digunakan untuk pembangkitan data tabular sintetis dan menunjukkan kemampuan menggabungkan pembelajaran probabilistik dengan generatif (Anshelevich & Katz, 2025). Model generatif pada data tabular juga memiliki potensi untuk memperkaya variasi representasi apabila proses pembangkitan dikendalikan sesuai struktur distribusi data (Wang et al., 2024). GAN tidak digunakan untuk melakukan oversampling langsung pada data antropometri dan hanya diterapkan pada representasi laten kelas stunting. Tingkat pengayaan 10%, 20%, dan 30% dievaluasi menggunakan data validasi, sedangkan data validasi dan pengujian tidak mengalami proses generasi.

Representasi laten hasil Dual-VAE dan pengayaan GAN selanjutnya digunakan sebagai masukan Transformer. Setiap dimensi representasi diperlakukan sebagai feature token sehingga mekanisme self-attention dapat mempelajari hubungan antardimensi laten. Pendekatan Transformer pada data tabular memungkinkan interaksi antarfitur dipelajari tanpa mengasumsikan adanya urutan temporal antarobservasi (Wang et al., 2024; Lin et al., 2025). Oleh karena itu, Transformer dalam penelitian ini tidak memodelkan urutan antar anak, melainkan interaksi nonlinear antarfitur laten yang berasal dari ruang representasi non-stunting dan stunting. Keluaran Transformer selanjutnya diteruskan ke lapisan klasifikasi biner untuk menghasilkan probabilitas stunting.

2.3. Skenario Pengujian dan Evaluasi

Eksperimen diimplementasikan menggunakan Python dan PyTorch dengan random seed 42. Evaluasi dilakukan menggunakan stratified five-fold cross-validation sehingga distribusi kelas dipertahankan pada setiap fold. Pada setiap iterasi, sebagian data pelatihan digunakan sebagai validasi internal untuk memilih hiperparameter, tingkat pengayaan GAN, menerapkan early stopping, dan menentukan ambang klasifikasi. Parameter prapemrosesan dipelajari dari bagian pelatihan dan selanjutnya diterapkan pada data validasi dan pengujian.

Optimasi hiperparameter dilakukan menggunakan Genetic Algorithm (GA) dengan populasi enam individu selama tiga generasi, probabilitas crossover 0,80, probabilitas mutasi awal 0,20, elitism dua individu, dan tournament size tiga. Parameter yang dioptimasi meliputi dimensi laten VAE, konfigurasi Transformer, learning rate, batch size, dan tingkat pengayaan GAN. Model dilatih menggunakan AdamW maksimum 20 epoch, gradient clipping sebesar 1,0, learning-rate scheduling, serta early stopping dengan patience delapan epoch. Ambang klasifikasi dipilih berdasarkan prediksi data validasi dengan memaksimalkan G-Mean dan selanjutnya digunakan tanpa perubahan pada data pengujian. Kinerja model diukur menggunakan accuracy, precision, recall, specificity, F1-score, balanced accuracy, G-Mean, Matthews Correlation Coefficient (MCC), AUROC, dan AUPRC. Penggunaan beberapa metrik penting karena akurasi saja dapat memberikan gambaran yang kurang memadai ketika distribusi kelas tidak seimbang (van Zyl & Engelbrecht, 2025; Chicco & Jurman, 2023).

Berdasarkan confusion matrix, TP menunjukkan data stunting yang diprediksi dengan benar, TN merupakan data non-stunting yang diprediksi dengan benar, FP adalah data non-stunting yang salah diprediksi sebagai stunting, dan FN adalah data stunting yang salah diprediksi sebagai non-stunting. Metrik utama dihitung sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP+FN}$$

F1-score menggabungkan precision dan recall melalui:

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

G-Mean mengukur keseimbangan antara sensitivitas dan spesifisitas:

$$G - MEAN = \sqrt{Recall \cdot Specificity}$$

MCC dihitung menggunakan seluruh komponen confusion matrix:

$$MCC = \frac{TP.TN - FP.FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Nilai MCC berada pada rentang -1 sampai 1, dengan nilai yang semakin mendekati 1 menunjukkan kesesuaian yang semakin tinggi antara kelas aktual dan prediksi. MCC relevan untuk klasifikasi biner karena memperhitungkan keempat komponen confusion matrix, terutama ketika distribusi kelas tidak seimbang (van Zyl & Engelbrecht, 2025; Chicco & Jurman, 2023).

Selain metrik berbasis ambang tersebut, AUROC digunakan untuk mengukur kemampuan model membedakan kedua kelas pada berbagai nilai ambang melalui hubungan antara true positive rate dan false positive rate. AUPRC digunakan untuk mengevaluasi hubungan precision–recall pada berbagai ambang dan memberikan informasi tambahan mengenai kinerja kelas positif. Oleh karena itu, AUROC dan AUPRC dilaporkan bersama F1-score, G-Mean, dan MCC agar evaluasi model tidak bergantung pada satu karakteristik kinerja saja (van Zyl & Engelbrecht, 2025; Chicco & Jurman, 2023).

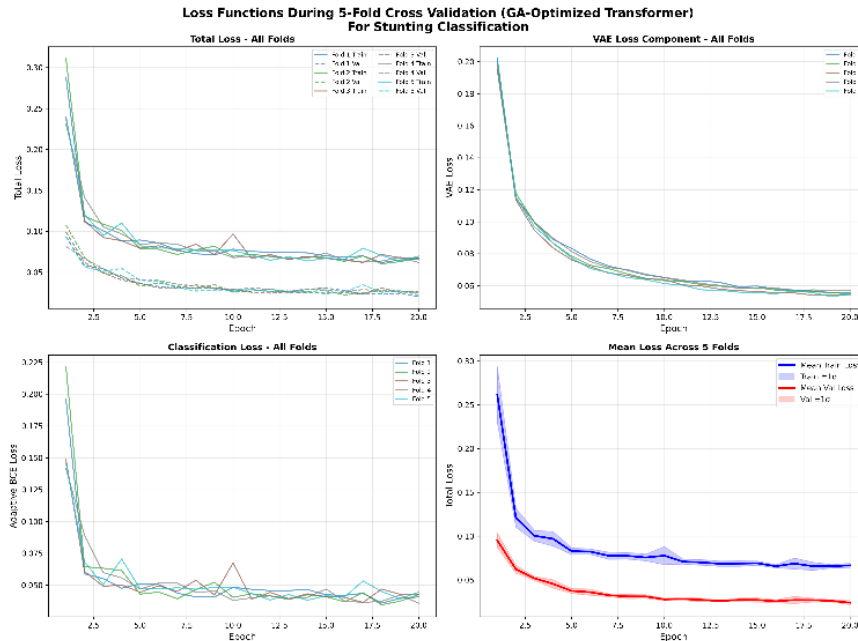
3. Hasil dan Pembahasan

Hasil penelitian dianalisis berdasarkan proses optimasi model, kinerja klasifikasi pada lima fold, perbandingan dengan penelitian terdahulu, serta pengujian ablasi terhadap komponen model. Evaluasi tersebut digunakan untuk mengetahui konsistensi kinerja Dual-VAE–GAN–Transformer sekaligus menilai kontribusi pembelajaran representasi berorientasi kelas, pengayaan laten menggunakan GAN, dan pemodelan interaksi fitur menggunakan Transformer.

3.1. Hasil Pengujian Model

Model Dual-VAE–GAN–Transformer dievaluasi menggunakan stratified five-fold cross-validation setelah optimasi hiperparameter dengan Genetic Algorithm menghasilkan fitness terbaik sebesar 0,9910. Model memperoleh rata-rata accuracy $0,9960 \pm 0,0007$, recall $0,9941 \pm 0,0019$, precision $0,9960 \pm 0,0019$, F1-score $0,9950 \pm 0,0008$, AUROC $0,9999 \pm 0,0001$, AUPRC $0,9998 \pm 0,0001$, G-Mean $0,9957 \pm 0,0007$, dan MCC $0,9917 \pm 0,0014$. Nilai tersebut menunjukkan kinerja yang konsisten pada lima fold pengujian.

Perilaku pelatihan pada Gambar 1 menunjukkan bahwa total loss, VAE loss, dan classification loss menurun terutama pada beberapa epoch awal dan kemudian cenderung stabil. Pola yang relatif serupa pada seluruh fold menunjukkan bahwa proses pembelajaran berlangsung konsisten hingga akhir pelatihan.



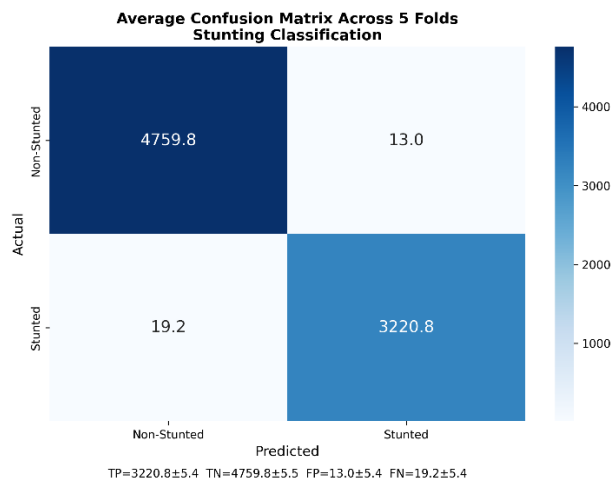
Sumber: Hasil Pengolahan Data (2026)

Gambar 1. Perkembangan total loss, VAE loss, dan classification loss selama proses five-fold cross-validation

Hasil klasifikasi rata-rata pada lima fold ditunjukkan oleh confusion matrix pada Gambar 2. Secara rata-rata, model menghasilkan $4759,8 \pm 5,5$ true negative, $3220,8 \pm 5,4$ true positive, $13,0 \pm 5,4$ false positive, dan $19,2 \pm 5,4$ false negative. Jumlah kesalahan klasifikasi yang relatif kecil pada kedua jenis error menunjukkan bahwa model mampu membedakan kelas stunting dan non-stunting dengan baik. Nilai false positive yang rendah menunjukkan bahwa model jarang mengklasifikasikan anak non-stunting sebagai stunting, sedangkan nilai false negative yang juga rendah menunjukkan bahwa kasus stunting yang terlewat relatif sedikit. Kondisi ini sejalan dengan tingginya nilai precision, recall, F1-score, dan MCC yang diperoleh model.

Kemampuan model dalam membedakan kelas positif juga diperlihatkan melalui analisis precision-recall pada Gambar 3. Kurva precision-recall pada lima fold terlihat menempel di area atas grafik, yang menunjukkan bahwa model mampu mempertahankan nilai precision dan recall yang sangat tinggi secara bersamaan. Boxplot distribusi AUPRC juga menunjukkan penyebaran yang sangat kecil antar fold, dengan nilai rata-rata $0,9998 \pm 0,0001$. Temuan ini memperlihatkan bahwa performa model tidak hanya tinggi secara rata-rata, tetapi juga stabil pada seluruh pembagian data. Dengan demikian, hasil pengujian mendukung bahwa pendekatan berbasis Dual-VAE, GAN, dan Transformer mampu menghasilkan deteksi stunting berbasis data antropometri dengan performa yang sangat konsisten.

Secara keseluruhan, hasil pengujian menunjukkan bahwa model yang diusulkan memiliki performa klasifikasi yang kuat pada dataset antropometri Jeneponto. Tingginya nilai F1-score, G-Mean, dan MCC menunjukkan bahwa model tidak hanya baik dalam memaksimalkan prediksi benar, tetapi juga mampu menjaga keseimbangan klasifikasi antara kelas stunting dan non-stunting. Selain itu, AUROC dan AUPRC yang mendekati 1 menegaskan bahwa model mempunyai kemampuan diskriminasi yang sangat tinggi pada berbagai ambang keputusan. Oleh karena itu, model ini memiliki potensi yang baik untuk digunakan sebagai alat bantu deteksi stunting berbasis data antropometri.



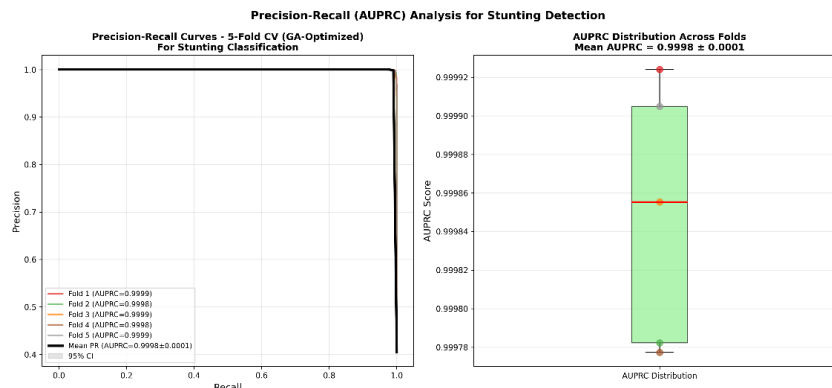
Sumber: Hasil Pengolahan Data (2026)

Gambar 2. Average confusion matrix hasil klasifikasi stunting pada lima fold pengujian

3.2. Perbandingan Dengan Penelitian Sebelumnya

Berdasarkan Tabel 1, model Dual-VAE–GAN–Transformer menghasilkan kinerja yang lebih tinggi pada metrik yang tersedia dibandingkan beberapa penelitian sebelumnya. Model usulan memperoleh accuracy $0,9960 \pm 0,0007$, recall $0,9941 \pm 0,0019$, precision $0,9960 \pm 0,0019$, dan F1-score $0,9950 \pm 0,0008$. Sebagai pembanding, XGBoost oleh Wicaksono et al., (2025) menghasilkan accuracy 0,9757 dan F1-score 0,9697, Random Forest oleh Resha & Toding (2026) memperoleh accuracy $0,9581 \pm 0,0028$ dan F1-score $0,9484 \pm 0,0034$, sedangkan tuned Gradient Boosting oleh (Rs et al., 2026) mencapai accuracy dan F1-score sebesar 0,9754.

Selain itu, model usulan mencapai AUROC $0,9999 \pm 0,0001$, AUPRC $0,9998 \pm 0,0001$, G-Mean $0,9957 \pm 0,0007$, dan MCC $0,9917 \pm 0,0014$, yang menunjukkan kinerja yang konsisten pada berbagai ukuran evaluasi. Hasil ini mendukung bahwa kombinasi representasi laten berorientasi kelas melalui Dual-VAE, pengayaan laten menggunakan GAN, dan pemodelan interaksi fitur menggunakan Transformer memberikan manfaat pada deteksi stunting berbasis data antropometri. Namun, perbandingan tersebut tetap perlu dipahami sesuai dengan perbedaan rancangan eksperimen dan metrik yang dilaporkan pada masing-masing penelitian.



Sumber: Hasil Pengolahan Data (2026)

Gambar 3. Kurva precision–recall lima fold dan distribusi nilai AUPRC antar fold

Tabel 1. Komparasi Kinerja Model Pada Dataset Stunting Jeneponto

Model	Accuracy	Recall	Precision	F1	AUROC	AUPRC	G-Mean	MCC
XGBoost –								
(Wicaksono et al., 2025)	0,9757	0,962	0,9774	0,9697	-	-	0,9734	0,9494
Random Forest –								
(Resha & Toding, 2026)	0,9581 ± 0,0028	0,9517 ± 0,0034	0,9451 ± 0,0038	0,9484 ± 0,0034	0,9907 ± 0,0012	-	-	0,9130 ± 0,0060
XGBoost –								
(Dewandra et al., 2026)	0,9681	0,95	-	-	-	-	-	-
Tuned Gradient Boosting (Rs et al., 2026)								
	0,9754	-	-	0,9754	-	-	-	-
Dual-VAE–GAN–Transformer (Usulan)								
	0,9960 ± 0,0007	0,9941 ± 0,0019	0,9960 ± 0,0019	0,9950 ± 0,0008	0,9999 ± 0,0001	0,9998 ± 0,0001	0,9957 ± 0,0007	0,9917 ± 0,0014

Sumber: Hasil Pengolahan Data (2026)

3.3. Analisis Ablasi Model

Berdasarkan Tabel 2, model lengkap menunjukkan perbedaan yang signifikan terhadap seluruh konfigurasi ablasi pada F1-score, AUPRC, G-Mean, dan MCC $p < 0,05$. Pada AUROC, perbedaannya tidak signifikan terhadap Feature-Token Transformer $p = 0,0510$ dan Single-VAE + Transformer $p = 0,1240$, sedangkan perbedaan signifikan terlihat pada Dual-VAE + Transformer $p = 0,0113$ dan Dual-VAE + GAN + MLP $p = 0,0102$. Pola ini menunjukkan bahwa pengaruh perubahan arsitektur lebih jelas terlihat pada metrik yang mempertimbangkan keseimbangan prediksi kelas dibandingkan AUROC saja.

Tabel 2. Hasil Uji Signifikansi terhadap Konfigurasi Ablasi

Konfigurasi pembanding	F1	AUROC	AUPRC	G-Mean	MCC
Feature-Token Transformer	0,00735	0,051	0,0237	0,0127	0,00846
Single-VAE + Transformer	0,00193	0,124	0,0043	0,00133	0,00218
Dual-VAE + Transformer	0,000581	0,0113	0,00219	0,000277	0,000513
Dual-VAE + GAN + MLP	0,000052	0,0102	0,00594	0,00000143	0,0000695

Sumber: Hasil Pengolahan Data (2026)

Hasil memperlihatkan bahwa ketika salah satu bagian model dikurangi atau diganti, perubahan kinerja muncul secara konsisten pada sebagian besar metrik evaluasi. Temuan ini memberi dasar bahwa kombinasi Dual-VAE, pengayaan laten berbasis GAN, dan Transformer berperan dalam kinerja model secara keseluruhan. Namun, hasil ablasi ini tidak dimaksudkan untuk menyatakan bahwa satu komponen tertentu menjadi satu-satunya penyebab peningkatan kinerja, karena pengujian dilakukan pada konfigurasi model yang saling berinteraksi.

3.4. Pembahasan Dan Keterbatasan

Hasil pengujian menunjukkan bahwa Dual-VAE–GAN–Transformer mampu mempertahankan kinerja yang tinggi pada lima fold pengujian. Hasil perbandingan pada Tabel 1 dan analisis ablasi pada Tabel 2 menunjukkan bahwa pembentukan representasi laten berorientasi kelas, pengayaan laten menggunakan GAN, dan pemodelan interaksi fitur dengan Transformer berkaitan dengan peningkatan kinerja model. Namun, hasil tersebut tetap perlu dipahami dalam konteks eksperimen yang dilakukan dan belum dapat digeneralisasikan secara langsung ke seluruh populasi. Penelitian ini menggunakan 40.071 data anak usia 0–59 bulan dari Kabupaten Jeneponto periode 2021–2024 dengan empat variabel utama, yaitu usia, jenis kelamin, berat badan, dan tinggi badan. Variabel kategori antropometri dan Z-score yang berkaitan langsung dengan penentuan status stunting telah dikeluarkan untuk mengurangi risiko target leakage. Meskipun demikian, karena status stunting memang berkaitan erat dengan pengukuran antropometri, tingginya nilai evaluasi perlu ditafsirkan secara proporsional.

4. Kesimpulan

Penelitian ini mengembangkan model Dual-VAE–GAN–Transformer untuk deteksi stunting menggunakan data antropometri. Hasil pengujian lima fold menunjukkan kinerja yang konsisten dengan accuracy sebesar $0,9960 \pm 0,0007$, recall $0,9941 \pm 0,0019$, precision $0,9960 \pm 0,0019$, F1-score $0,9950 \pm 0,0008$, AUROC $0,9999 \pm 0,0001$, AUPRC $0,9998 \pm 0,0001$, G-Mean $0,9957 \pm 0,0007$, dan MCC $0,9917 \pm 0,0014$. Hasil perbandingan dan pengujian ablasi menunjukkan bahwa penggunaan representasi laten berorientasi kelas melalui Dual-VAE, pengayaan laten dengan GAN, serta pemodelan interaksi fitur menggunakan Transformer berkontribusi terhadap kinerja model secara keseluruhan. Meskipun demikian, hasil tersebut masih terbatas pada data anak dari Kabupaten Jeneponto dan belum dapat langsung digeneralisasikan ke populasi lain. Penelitian selanjutnya perlu menguji model menggunakan

data independen dari wilayah berbeda serta mempertimbangkan faktor non-antropometri agar kemampuan generalisasi model dapat dievaluasi secara lebih luas.

Ucapan Terima Kasih

Penulis menyampaikan terima kasih kepada Universitas Bhayangkara Jakarta Raya, khususnya Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM), atas dukungan pendanaan melalui Program Hibah Penelitian Internal. Artikel ini merupakan salah satu luaran dari kegiatan penelitian yang didanai melalui program tersebut.

Daftar Pustaka

- Anshelevich, D., & Katz, G. (2025). Knowledge-Based Systems Synthetic tabular data generation using a VAE-GAN architecture. *Knowledge-Based Systems*, 326(June), 113997. <https://doi.org/10.1016/j.knosys.2025.113997>
- Azriani, D., Agustian, D., Zuhairini, Y., Yulita, I. N., & Dhamayanti, M. (2025). Prediction models for stunting at 2-years-old from Indonesian newborn population. *BMC Pediatrics*, 25(718). <https://doi.org/https://doi.org/10.1186/s12887-025-06096-4>
- Begashaw, G. B., Zewotir, T., & Fenta, H. M. (2025). A deep learning approach for classifying and predicting children ' s nutritional status in Ethiopia using LSTM - FC neural networks. *BioData Mining*. <https://doi.org/10.1186/s13040-025-00425-0>
- Berahmand, K., Daneshfar, F., Sadat, E., & Yuefeng, S. (2024). Autoencoders and their applications in machine learning: a survey. In *Artificial Intelligence Review*. Springer Netherlands. <https://doi.org/10.1007/s10462-023-10662-6>
- Chicco, D., & Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining*, 16(1), 1–23. <https://doi.org/10.1186/s13040-023-00322-4>
- Dewandra, P., Gede, I., Mas Diyasa, S., & Rahajoe, A. D. (2026). Seminar Nasional Informatika Bela Negara (SANTIKA) Perbandingan Algoritma Random Forest dan XGBoost dalam Klasifikasi Status Stunting pada Balita Berdasarkan Data Antropometri. *Seminar Nasional Informatika Bela Negara (SANTIKA)*, 6, 108–112.
- Hadikurniawati, W. (2026). Lightweight Hybrid Feature Engineering and Fusion with Variational Autoencoder and XGBoost for Nutritional Data Classification. *Engineering, Technology & Applied Science Research*, 16(1), 30859–30868. <https://doi.org/10.48084/etasr.15310>
- Hadikurniawati, W., Hartomo, K. D., Sembiring, I., & Arthur, C. (2025). A Dual-Fusion Hybrid Model with Attention for Stunting Prediction among Children under Five Years. *Journal of Applied Data Sciences*, 6(3), 1985–1998. <https://doi.org/https://doi.org/10.47738/jads.v6i3.831> This
- Khoirunnisaa, N., Priatna, W., Rasim, & Warta, J. (2024). Stunting Pada Dinas Kesehatan Kota Bekasi Menggunakan Algoritma FP-Growth. *Jurnal Teknologi Dan Ilmu Komputer Prima (JUTIKOMP)*, 7(1), 73–87. <https://doi.org/https://doi.org/10.34012/jutikomp.v7i1.4761>

- Lin, Q., Sheng, J., Zhou, C., Xiao, T., Meng, Y., Lu, M., Zhang, J., Yan, X., Peng, L., Cao, H., & Li, L. (2025). LPItabformer: Enhancing generalization in predicting lncRNA-protein interactions via a tabular Transformer. *Computational and Structural Biotechnology Journal*, 27(January), 2323–2335. <https://doi.org/10.1016/j.csbj.2025.05.050>
- Luo, S., Wang, Y., Wu, R., Yang, Z., Gao, H., Yuan, J., & Wang, Z. (2026). Using deep learning to monitor children ' s nutritional status for child stunting identification. *Frontiers in Pediatrics*, 14(August). <https://doi.org/10.3389/fped.2026.1725474>
- Novalina, N., Amyas, I., Tarigan, A., Kameela, F. K., & Rizkinia, M. (2025). Benchmarking machine learning algorithm for stunting risk prediction in Indonesia. *Bulletin of Electrical Engineering and Informatics*, 14(3), 2252–2263. <https://doi.org/10.11591/eei.v14i3.8997>
- Resha, M., & Toding, A. (2026). Z-Score-Free Stunting Prediction from Basic Anthropometric Measurements Using Machine Learning. *Junral Zetroem*, 08(02), 148–154. <https://doi.org/https://doi.org/10.36526/ztr.v8i2.8736>
- RS, A. H., Purnamawati, Jaya, H., & Arfandi, A. (2026). Machine Learning-Based Stunting Classification Using Puskesmas Nutritional Data : A Comparative Study of Six Algorithms with Hyperparameter Tuning. *International Conference on Artificial Intelligence and the Digital Commons Machine*, 162–175.
- RS, A. H., Purnamawati, P., Jaya, H., Arfandi, A., Suhardi, I., Widodo, S., & Lonang, S. (2026). *Dataset Stunting and Nutritional Status of Toddler from Jeneponto Regency, South Sulawesi, Indonesia*. <https://Data.Mendeley.Com/>.
- Setiawan, Y., Hamim, M., Al, Z., Nizar, M., Ma, P., & Made, I. (2025). Modality-based Modeling with Data Balancing and Dimensionality Reduction for Early Stunting Detection. *JOIN (Jurnal Online Informatika)*, 10(1), 53–65. <https://doi.org/DOI: 10.15575/join.v10i1.1495>
- Shen, H., Zhao, H., & Jiang, Y. (2023). Machine Learning Algorithms for Predicting Stunting among Under-Five Children in Papua New Guinea. *Children*, 10(10), 1638. <https://doi.org/10.3390/children10101638>
- Sugihartono, T., Wijaya, B., Alkayes, A. F., & Anugrah, H. A. (2025). Optimizing Stunting Detection through SMOTE and Machine Learning : a Comparative Study of XGBoost , Random Forest , SVM , and k-NN. *Journal of Applied Data Sciences*, 6(1), 667–682. <https://doi.org/https://doi.org/10.47738/jads.v6i1.494>
- van Zyl, J. P., & Engelbrecht, A. P. (2025). Analysis of classification metric behaviour under class imbalance. *Egyptian Informatics Journal*, 31(August 2024), 100711. <https://doi.org/10.1016/j.eij.2025.100711>
- Wang, A. X., Chukova, S. S., Simpson, C. R., & Nguyen, B. P. (2024). Challenges and opportunities of generative models on tabular data. *Applied Soft Computing*, 166(July 2023), 112223. <https://doi.org/10.1016/j.asoc.2024.112223>
- Wicaksono, A., Prasetyo, D., Mar'atullatifah, Yulaikha, Iswavigra, Utari, D., Himmatunnisa, & Mahmudah5, A. H. (2025). Data Analysis and Explainable Machine Learning for Stunting Prediction. *Journal of Artificial Intelligence and Legal Technology*, 1(1), 35–44.