

Predicting Stunting Levels In Toddlers Using Machine Learning Based On Health Data And Image Processing

Gathan Raihan Al Saleh¹⁾; Juju Jumadi²⁾; Ahmad Arib Al Farisy³⁾

^{1,2,3)} Universitas Dehasen Bengkulu, Indonesia

Email: ¹⁾ gathanraihan21@gmail.com

How to Cite :

Saleh, G.R.A., Jumadi, J., Farisy A.A.A. (2026). Predicting Stunting Levels In Toddlers Using Machine Learning Based On Health Data And Image Processing. Jurnal Media Computer Science, 5(4)

ARTICLE HISTORY

Received [22 September 2026]

Revised [23 Oktober 2026]

Accepted [26 Oktober 2026]

KEYWORDS

Convolutional Neural Network,
Ensemble Learning, Image
Processing.

This is an open access article under the
[CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



ABSTRAK

Stunting merupakan salah satu permasalahan gizi kronis pada balita yang dapat menghambat pertumbuhan fisik dan perkembangan kognitif anak. Keterbatasan kualitas dan kelengkapan data kesehatan yang dicatat secara manual di Posyandu menjadi tantangan dalam proses identifikasi dini stunting. Penelitian ini bertujuan membangun model prediksi tingkat stunting balita menggunakan pendekatan Machine Learning berbasis data kesehatan dan image processing. Landasan teori yang digunakan meliputi Machine Learning, Support Vector Machine (SVM), Random Forest, Convolutional Neural Network (CNN), Image Processing, dan Ensemble Learning dengan teknik Hard Voting. Metode penelitian yang digunakan adalah penelitian eksperimen kuantitatif dengan kerangka kerja CRISP-DM. Data diperoleh melalui observasi dan pengumpulan data kesehatan balita di Posyandu Desa Lubuk Sahung, Kabupaten Seluma, Bengkulu, yang didukung oleh dataset publik Kaggle serta citra tubuh balita. Tahapan penelitian meliputi preprocessing data, ekstraksi fitur citra menggunakan CNN, feature fusion antara data kesehatan dan fitur citra, pelatihan model SVM dan Random Forest, serta penggabungan model menggunakan Ensemble Voting. Hasil penelitian menunjukkan bahwa model SVM dan Ensemble Voting memperoleh performa terbaik dengan accuracy 91,67%, precision 100%, recall 84,62%, dan F1-score 91,67%, sedangkan Random Forest memperoleh accuracy 83,33%. Berdasarkan analisis generalisasi, SVM memiliki stabilitas terbaik dengan selisih akurasi training dan testing sebesar 4,16%. Penelitian ini menunjukkan bahwa kombinasi data kesehatan dan fitur citra mampu mendukung prediksi tingkat stunting secara efektif sebagai alat bantu pemantauan awal kondisi balita.

ABSTRACT

Stunting is one of the most prevalent chronic nutritional disorders affecting toddlers, with significant impacts on their physical growth and cognitive development. The limited quality and completeness of health records collected manually at community health posts (Posyandu) present challenges for the early identification of stunting. This study aimed to develop a predictive model for stunting levels in toddlers using a machine learning approach based on health data and image processing. The theoretical framework incorporated Machine Learning, Support Vector Machine (SVM), Random Forest, Convolutional Neural Network (CNN), Image Processing, and Ensemble Learning using the Hard Voting technique. This study employed a quantitative experimental design following the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework. The data were obtained through observations and the collection of child health records at the Posyandu in Lubuk Sahung Village, Seluma Regency, Bengkulu Province, complemented by a public dataset from Kaggle and body images of children under five. The research process included data preprocessing, image feature extraction using CNN, feature fusion of health data and image features, model training using SVM and Random Forest, and model integration through Ensemble Voting. The results demonstrated that the SVM and Ensemble Voting models achieved the best predictive performance, with an accuracy of 91.67%, precision of 100.00%, recall of 84.62%, and an F1-score of 91.67%, while the Random Forest model achieved an accuracy of 83.33%. Based on the generalization analysis, the SVM model exhibited the highest stability, with a training-testing accuracy gap of only 4.16%. These findings indicate

PENDAHULUAN

Perkembangan teknologi informasi dan kecerdasan buatan (Artificial Intelligence) telah memberikan kontribusi signifikan dalam berbagai bidang, termasuk sektor kesehatan. Teknologi ini memungkinkan analisis data yang kompleks serta pengambilan keputusan yang lebih akurat, sehingga membuka peluang untuk meningkatkan kualitas layanan kesehatan dan mempermudah akses layanan kesehatan bagi masyarakat. AI terdiri dari berbagai teknik, salah satunya Machine Learning (ML), yang memungkinkan sistem komputer untuk mengolah data kesehatan dan menghasilkan prediksi berdasarkan pola yang diperoleh dari data yang tersedia. (Thaariq et al., 2024).

Penelitian ini dilaksanakan di Desa Lubuk Sahung, Kabupaten Seluma, Provinsi Bengkulu, yang merupakan wilayah pedesaan dengan karakteristik sosial ekonomi menengah ke bawah dan masih menghadapi permasalahan tingkat stunting pada balita. Pemantauan kesehatan balita di desa ini dilakukan melalui kegiatan Posyandu yang berperan dalam pencatatan dan pengumpulan data kesehatan balita secara manual dan berkala. Secara ideal, data kesehatan balita diharapkan tercatat secara lengkap, akurat, dan konsisten sebagai dasar dalam identifikasi tingkat stunting sejak dini.

Data kesehatan merupakan informasi yang berkaitan dengan kondisi kesehatan individu yang diperoleh melalui proses pemeriksaan dan pengukuran kesehatan. Dalam pemantauan kesehatan balita, data kesehatan digunakan untuk menilai status gizi dan pertumbuhan anak berdasarkan indikator antropometri (Kementerian Kesehatan Indonesia, 2022). Dalam penelitian ini, data kesehatan balita yang diperoleh dari kegiatan Posyandu di Desa Lubuk Sahung umumnya meliputi berat badan (BB), tinggi badan (TB), lingkaran lengan atas (LLA), lingkaran kepala (LKA), serta umur. Variabel-variabel tersebut termasuk jenis data kesehatan antropometri yang digunakan untuk menilai pertumbuhan, perkembangan, dan status gizi balita dalam identifikasi stunting. Namun, berdasarkan kondisi aktual di lapangan, proses pencatatan data tersebut masih dilakukan secara manual dan sering mengalami berbagai kendala. Pada saat pengukuran, balita sering kali menangis atau tidak kooperatif, sehingga petugas mengalami kesulitan dalam melakukan penimbangan dan pengukuran secara akurat. Akibatnya, sebagian data tidak dapat dicatat secara lengkap atau memiliki tingkat ketelitian yang rendah. Kondisi ini menyebabkan kualitas data kesehatan balita menjadi kurang optimal dan menyulitkan proses analisis apabila hanya mengandalkan pencatatan manual. Kesenjangan antara kebutuhan akan data yang akurat dan kondisi pencatatan di lapangan inilah yang menjadi salah satu permasalahan utama.

Beberapa penelitian terdahulu menunjukkan bahwa penerapan Machine Learning dalam bidang kesehatan mampu meningkatkan akurasi analisis dan prediksi permasalahan gizi anak, termasuk stunting. Ningrum et al., (2025) menyatakan bahwa metode klasifikasi menggunakan Support Vector Machine (SVM) mampu memberikan tingkat akurasi yang baik dalam penentuan status stunting pada wilayah lokal. Selain itu, Fitriyani et al., (2024) lewat hasil pengujian menunjukkan bahwa Naïve Bayes memiliki performa tertinggi dengan akurasi = 98,57%, recall dan presisi yang sangat tinggi, namun Random Forest juga memberikan hasil yang baik dengan akurasi = 98,29%, dalam analisis dataset kesehatan. Di sisi lain, studi yang dilakukan oleh Kusuma, (2024) menjelaskan bahwa penerapan deep learning dalam pemrosesan citra digital memiliki kemampuan dalam melakukan ekstraksi fitur secara otomatis, sehingga berpotensi dimanfaatkan sebagai data pendukung dalam analisis data kesehatan. Berdasarkan hasil penelitian-penelitian tersebut, penggunaan berbagai algoritma Machine Learning dengan dukungan data pendukung yang relevan diharapkan dapat membantu praktisi kesehatan dalam mengidentifikasi tingkat stunting secara dini. Model prediksi yang optimal selanjutnya dapat digunakan sebagai alat informasi pendukung dalam proses pemantauan tingkat stunting.

Berdasarkan permasalahan kualitas dan kelengkapan data kesehatan balita yang diperoleh dari kegiatan Posyandu, penelitian ini menawarkan pendekatan prediksi tingkat stunting menggunakan metode Machine Learning algoritma Ensemble Learning dengan teknik Voting. Data kesehatan balita digunakan sebagai sumber data utama, sementara dataset publik dan citra digital balita dimanfaatkan sebagai data pendukung mengingat keterbatasan data primer yang tersedia. Citra digital balita diolah menggunakan teknik pemrosesan citra digital untuk mengekstraksi fitur visual yang kemudian digabungkan dengan data kesehatan sebagai input model prediksi akibat kendala di lapangan, seperti balita yang tidak kooperatif saat pengukuran. Dengan penerapan prediksi tingkat stunting menggunakan metode Machine Learning dengan pendekatan Ensemble Learning secara terbatas ini, diharapkan dapat membantu tenaga kesehatan sebagai bahan pendukung pengambilan keputusan awal.

LANDASAN TEORI

Stunting merupakan kondisi gagal tumbuh pada anak akibat kekurangan gizi kronis yang terjadi dalam periode waktu yang lama, terutama pada 1.000 hari pertama kehidupan. Anak yang mengalami stunting memiliki tinggi badan yang lebih rendah dibandingkan dengan standar pertumbuhan anak seusianya. Menurut World Health Organization and Unicef, (2023) Stunting didefinisikan secara antropometri sebagai kondisi ketika nilai tinggi badan menurut umur (TB/U) berada di bawah -2 standar deviasi (< -2 SD) berdasarkan standar pertumbuhan anak yang telah ditetapkan secara internasional. Jika nilai berada di bawah -3 SD, maka dikategorikan sebagai stunting berat (severely stunted). Di Indonesia, ketentuan mengenai klasifikasi status gizi ini diatur secara resmi melalui Peraturan Menteri Kesehatan (PMK) No. 2 Tahun 2020).

Machine Learning merupakan salah satu cabang dari Artificial Intelligence yang memungkinkan sistem komputer untuk belajar dari data tanpa diprogram secara eksplisit. Machine Learning bekerja dengan memanfaatkan algoritma yang mampu mengenali pola dari data dan menghasilkan prediksi atau keputusan berdasarkan pola tersebut. Dalam bidang kesehatan, Machine Learning banyak digunakan untuk analisis data medis, klasifikasi penyakit, serta prediksi kondisi kesehatan pasien. Menurut Wardhana et al. (2023), Machine Learning merupakan metode komputasi yang memungkinkan sistem untuk belajar dari data dan meningkatkan kinerjanya secara otomatis melalui pengalaman. Machine Learning digunakan untuk menemukan pola tersembunyi dalam data serta menghasilkan prediksi berdasarkan informasi yang tersedia, sehingga dapat membantu dalam pengambilan keputusan secara lebih efektif dan mengungkapkan bahwa proses implementasi Machine Learning terhadap prediksi tingkat kasus penyakit dengan algoritma klasifikasi dapat dilakukan dengan melakukan upaya normalisasi angka kasus penyakit serta menerapkan teknik pemodelan dengan algoritma klasifikasi. Pemodelan dengan teknik decision tree dalam penelitian ini menghasilkan akurasi yang sangat baik.

Random Forest merupakan salah satu algoritma dalam Machine Learning yang termasuk dalam metode ensemble learning. Algoritma ini bekerja dengan membangun sejumlah pohon keputusan (decision tree), kemudian menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan keputusan akhir yang lebih akurat. Metode Random Forest pertama kali diperkenalkan oleh Leo Breiman sebagai pengembangan dari metode decision tree yang memiliki kelemahan terhadap masalah overfitting pada proses klasifikasi data. Menurut Ratnawati & Sulistyaningrum, (2020) Random Forest merupakan metode klasifikasi yang dikembangkan dari algoritma Decision Tree dengan pemilihan atribut secara acak pada setiap node untuk menentukan proses klasifikasi. Pada proses klasifikasinya, metode ini menggunakan pendekatan suara terbanyak (majority voting) dari seluruh pohon keputusan yang dihasilkan sehingga mampu menghasilkan model klasifikasi yang lebih stabil

METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah penelitian eksperimen dengan pendekatan kuantitatif. Pendekatan ini dipilih karena penelitian berfokus pada pembangunan dan evaluasi model prediksi tingkat stunting menggunakan algoritma Machine Learning berdasarkan data kesehatan dan citra tubuh balita. Penelitian ini tidak mengembangkan sistem aplikasi secara menyeluruh, melainkan menitikberatkan pada proses pemodelan, pelatihan (training), serta pengujian (testing) model untuk mengetahui tingkat kinerja prediksi yang dihasilkan. Dalam pelaksanaannya, penelitian ini menggunakan model proses CRISP-DM (Cross-Industry Standard Process for Data Mining) sebagai kerangka kerja pengembangan model. CRISP-DM dipilih karena menyediakan tahapan sistematis dalam proses pengolahan data dan pembangunan model Machine Learning

HASIL DAN PEMBAHASAN

Hasil

Hasil penelitian menunjukkan bahwa dataset yang digunakan terdiri atas 120 citra balita dan satu file Excel berisi data kesehatan balita. Berdasarkan proses validasi, seluruh citra berhasil dibaca dengan baik, tidak ditemukan citra rusak, dan dataset dinyatakan layak untuk digunakan dalam proses penelitian. Pada tahap pemeriksaan data numerik, ditemukan missing value pada atribut LILA sebanyak 13 data dan LIKA sebanyak 10 data, dengan total 23 data kosong. Setelah dilakukan proses penanganan missing value, seluruh atribut memiliki nilai lengkap dan tidak ditemukan data duplikat, sehingga dataset dinyatakan bersih dan siap digunakan pada tahap preprocessing serta pemodelan.

Pada tahap preprocessing citra, dilakukan beberapa proses meliputi resize citra menjadi 224 × 224 piksel, normalisasi nilai piksel ke rentang 0–1, cropping area tubuh balita, dan denoising. Tahapan ini bertujuan untuk menyeragamkan ukuran citra, memfokuskan objek utama, serta mengurangi gangguan visual pada gambar. Hasil preprocessing menunjukkan bahwa citra telah memiliki kualitas yang lebih baik dan siap digunakan untuk proses ekstraksi fitur menggunakan Convolutional Neural Network (CNN).

Ekstraksi fitur citra dilakukan menggunakan CNN yang terdiri atas convolution layer, fungsi aktivasi ReLU, pooling layer, dan flatten layer. Proses ini menghasilkan fitur visual numerik yang merepresentasikan karakteristik tubuh balita. Selanjutnya, dilakukan feature fusion dengan menggabungkan 64 fitur visual hasil ekstraksi CNN dan 3 fitur kesehatan balita. Hasil penggabungan tersebut menghasilkan matriks fitur gabungan sebanyak 120 data dengan total 68 kolom fitur yang digunakan sebagai input pada proses klasifikasi.

Pada tahap pelatihan model, dilakukan hyperparameter tuning menggunakan Grid Search Cross-Validation terhadap algoritma Support Vector Machine (SVM) dan Random Forest. Model SVM memperoleh akurasi training sebesar 79,16% dan akurasi testing sebesar 75,00%, dengan selisih akurasi sebesar 4,16%. Selisih yang relatif kecil menunjukkan bahwa model SVM memiliki kemampuan generalisasi yang baik dan tidak mengalami overfitting yang signifikan. Sementara itu, model Random Forest memperoleh akurasi training sebesar 90,62% dan akurasi testing sebesar 62,50%. Selisih akurasi sebesar 28,12% menunjukkan adanya kecenderungan overfitting karena model terlalu baik dalam mempelajari data training, tetapi kurang optimal saat diuji pada data testing.

Model Ensemble Voting Classifier dibangun dengan menggabungkan SVM dan Random Forest menggunakan metode voting berdasarkan probabilitas prediksi. Bobot SVM diberikan lebih besar karena memiliki performa pengujian yang lebih baik. Hasil pelatihan menunjukkan bahwa Ensemble Voting memperoleh akurasi training sebesar 79,17% dan akurasi testing sebesar 75,00%, dengan selisih sebesar 4,17%. Hal ini menunjukkan bahwa model Ensemble Voting memiliki kemampuan generalisasi yang baik dan mampu mempertahankan performa yang stabil pada data testing.

Tabel 1. Perbandingan Kinerja Model

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM	75,00	73,33	84,62	78,57
Random Forest	62,50	62,50	76,92	68,97
Ensemble Voting	75,00	73,33	84,62	78,57

Berdasarkan perbandingan kinerja model, SVM dan Ensemble Voting memperoleh hasil terbaik dengan accuracy sebesar 75,00%, precision sebesar 73,33%, recall sebesar 84,62%, dan F1-Score sebesar 78,57%. Random Forest memperoleh accuracy sebesar 62,50%, precision sebesar 62,50%, recall sebesar 76,92%, dan F1-Score sebesar 68,97%. Nilai recall dan F1-Score yang lebih tinggi pada SVM dan Ensemble Voting menunjukkan bahwa kedua model lebih baik dalam mendeteksi kasus stunting serta memiliki keseimbangan yang lebih baik antara precision dan recall.

Secara keseluruhan, hasil penelitian menunjukkan bahwa penggabungan fitur citra dan data kesehatan dapat digunakan sebagai dasar klasifikasi tingkat stunting balita. Model SVM menjadi model terbaik dalam penelitian ini karena menghasilkan performa klasifikasi yang baik, memiliki akurasi testing tertinggi, serta menunjukkan stabilitas yang lebih baik dibandingkan Random Forest. Meskipun Ensemble Voting mampu menghasilkan performa yang sama dengan SVM, metode tersebut belum memberikan peningkatan akurasi dibandingkan penggunaan SVM secara individual.

Pembahasan

Penelitian ini bertujuan untuk mengklasifikasikan tingkat stunting balita menggunakan kombinasi data kesehatan dan citra balita dengan pendekatan machine learning berbasis feature fusion. Sebelum dilakukan proses klasifikasi, data terlebih dahulu melalui tahap preprocessing untuk meningkatkan kualitas data yang digunakan.

Pada tahap preprocessing citra dilakukan proses resize, normalisasi, cropping area tubuh balita, dan pengurangan noise. Tahapan ini bertujuan untuk menghasilkan citra yang lebih seragam dan mengurangi informasi yang tidak relevan sehingga dapat meningkatkan kualitas fitur yang diekstraksi.

Selanjutnya, ekstraksi fitur citra dilakukan menggunakan Convolutional Neural Network (CNN). Hasil ekstraksi menghasilkan 64 fitur visual yang merepresentasikan karakteristik tubuh balita. Fitur visual tersebut kemudian digabungkan dengan 3 fitur kesehatan melalui proses feature fusion sehingga diperoleh total 68 fitur yang digunakan sebagai input model klasifikasi.

Berdasarkan hasil pengujian, model Support Vector Machine (SVM) memperoleh accuracy sebesar 75,00%, precision sebesar 73,33%, recall sebesar 84,62%, dan F1-Score sebesar 78,57%. Hasil

tersebut menunjukkan bahwa model SVM mampu melakukan klasifikasi tingkat stunting dengan baik. Nilai recall yang tinggi menunjukkan bahwa sebagian besar kasus stunting berhasil dideteksi oleh model, sedangkan nilai precision menunjukkan bahwa sebagian besar data yang diprediksi sebagai stunting sesuai dengan kondisi sebenarnya.

Model Random Forest memperoleh accuracy sebesar 62,50%, precision sebesar 62,50%, recall sebesar 76,92%, dan F1-Score sebesar 68,97%. Hasil tersebut menunjukkan bahwa performa Random Forest masih berada di bawah model SVM. Selain itu, selisih akurasi training dan testing yang cukup besar menunjukkan adanya kecenderungan overfitting, sehingga kemampuan model dalam melakukan generalisasi terhadap data baru masih kurang optimal.

Metode Ensemble Voting Classifier yang menggabungkan SVM dan Random Forest menghasilkan accuracy sebesar 75,00%, precision sebesar 73,33%, recall sebesar 84,62%, dan F1-Score sebesar 78,57%. Hasil tersebut menunjukkan bahwa metode ensemble mampu mempertahankan performa yang diperoleh model SVM. Namun demikian, penggunaan Ensemble Voting belum mampu meningkatkan performa klasifikasi dibandingkan dengan model SVM secara individual.

Berdasarkan analisis generalisasi model, SVM memiliki selisih akurasi training dan testing sebesar 4,16%, diikuti oleh Ensemble Voting sebesar 4,17%, sedangkan Random Forest memiliki selisih sebesar 28,12%. Hasil ini menunjukkan bahwa SVM merupakan model yang paling stabil dan memiliki kemampuan generalisasi terbaik pada penelitian ini.

Secara keseluruhan, penelitian ini menunjukkan bahwa kombinasi fitur kesehatan dan fitur citra hasil ekstraksi CNN mampu digunakan untuk melakukan klasifikasi tingkat stunting balita dengan baik. Model Support Vector Machine (SVM) menjadi model terbaik dalam penelitian ini dengan accuracy sebesar 75,00%, nilai recall sebesar 84,62%, serta kemampuan generalisasi yang paling baik dibandingkan dengan model lainnya.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian dan pembahasan mengenai prediksi tingkat stunting balita menggunakan Machine Learning berbasis data kesehatan dan Image Processing, dapat disimpulkan bahwa penelitian ini berhasil menerapkan metode Machine Learning dengan memanfaatkan data kesehatan balita yang terdiri dari berat badan, tinggi badan, lingkaran lengan atas (LILA), lingkaran kepala (LIKA), umur, serta fitur citra yang diperoleh melalui proses Image Processing dan ekstraksi fitur menggunakan Convolutional Neural Network (CNN). Tahapan preprocessing yang dilakukan, meliputi penanganan missing value, validasi data, normalisasi, resize citra, cropping area tubuh balita, dan ekstraksi fitur CNN, mampu menghasilkan dataset yang siap digunakan dalam proses pelatihan dan pengujian model.

Hasil pengujian menunjukkan bahwa algoritma Support Vector Machine (SVM) dan Ensemble Voting Classifier memperoleh nilai accuracy sebesar 75,00%, precision sebesar 73,33%, recall sebesar 84,62%, dan F1-Score sebesar 78,57%, sedangkan Random Forest memperoleh accuracy sebesar 62,50%, precision sebesar 62,50%, recall sebesar 76,92%, dan F1-Score sebesar 68,97%. Berdasarkan hasil tersebut, SVM dan Ensemble Voting menjadi model dengan performa terbaik pada dataset yang digunakan dalam penelitian ini.

Analisis generalisasi menunjukkan bahwa model SVM memiliki selisih akurasi training dan testing sebesar 4,16%, Ensemble Voting sebesar 4,17%, dan Random Forest sebesar 28,12%. Hasil tersebut menunjukkan bahwa model SVM memiliki kemampuan generalisasi yang paling baik dan performa yang lebih stabil dibandingkan dengan model lainnya.

Meskipun demikian, hasil yang diperoleh belum menunjukkan peningkatan performa yang signifikan dari penggunaan metode Ensemble Learning maupun penambahan fitur citra. Hal ini terlihat dari nilai evaluasi Ensemble Voting yang identik dengan model SVM tunggal. Dengan kata lain, penggabungan model menggunakan metode voting belum mampu memberikan peningkatan akurasi yang berarti dibandingkan dengan penggunaan satu model klasifikasi saja.

Kondisi tersebut dipengaruhi oleh beberapa faktor teknis. Pertama, jumlah dataset yang digunakan masih relatif terbatas sehingga model belum memperoleh variasi data yang cukup untuk mempelajari pola stunting secara lebih mendalam. Kedua, ukuran data uji yang kecil menyebabkan perubahan prediksi pada beberapa sampel saja dapat memengaruhi nilai evaluasi secara signifikan. Ketiga, kualitas dan konsistensi citra balita yang digunakan masih bervariasi, baik dari sisi pencahayaan, posisi tubuh, jarak pengambilan gambar, maupun latar belakang, sehingga fitur visual yang dihasilkan belum mampu memberikan informasi tambahan yang kuat bagi model klasifikasi. Keempat, variabel kesehatan yang

digunakan sudah memiliki kontribusi yang dominan dalam proses prediksi, sehingga penambahan fitur citra belum memberikan peningkatan performa yang berarti.

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa metode Machine Learning yang diterapkan telah mampu melakukan klasifikasi tingkat stunting balita dengan cukup baik pada dataset penelitian. Model Support Vector Machine (SVM) menjadi model terbaik karena menghasilkan nilai accuracy dan kemampuan generalisasi yang lebih baik dibandingkan dengan model lainnya.

Saran

Berdasarkan hasil kesimpulan yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya. Jumlah dataset yang digunakan sebaiknya diperbanyak dan diperoleh dari berbagai posyandu, puskesmas, atau wilayah yang berbeda agar model dapat mempelajari pola data yang lebih beragam serta memiliki kemampuan generalisasi yang lebih baik. Penambahan jumlah data juga diharapkan dapat menghasilkan evaluasi model yang lebih representatif dan mengurangi kemungkinan bias akibat keterbatasan sampel penelitian.

Kualitas citra balita yang digunakan dalam proses ekstraksi fitur juga perlu ditingkatkan dengan menerapkan standar pengambilan gambar yang lebih baik, seperti pengaturan pencahayaan yang seragam, posisi tubuh yang konsisten, serta latar belakang yang lebih sederhana. Dengan demikian, fitur visual yang dihasilkan dapat lebih representatif dan memberikan kontribusi yang lebih besar terhadap proses klasifikasi.

Penelitian selanjutnya juga disarankan untuk menggunakan metode evaluasi yang lebih komprehensif, seperti K-Fold Cross Validation, sehingga hasil pengujian tidak hanya bergantung pada satu kali pembagian data latih dan data uji. Metode tersebut dapat memberikan gambaran performa model yang lebih stabil dan objektif dibandingkan dengan metode hold-out validation yang digunakan pada penelitian ini.

Selain itu, penelitian berikutnya dapat mencoba algoritma lain seperti XGBoost, LightGBM, CatBoost, maupun pendekatan Deep Learning end-to-end untuk mengetahui kemungkinan peningkatan performa klasifikasi. Pengembangan metode Ensemble Learning yang lebih kompleks, seperti Stacking atau Boosting, juga dapat dilakukan mengingat metode Hard Voting pada penelitian ini belum mampu memberikan peningkatan performa yang signifikan dibandingkan dengan model tunggal.

Penambahan variabel lain yang berkaitan dengan stunting, seperti riwayat gizi ibu, berat badan lahir, pola makan, status imunisasi, kondisi sanitasi, dan faktor sosial ekonomi keluarga, juga dapat dipertimbangkan untuk meningkatkan kemampuan model dalam mempelajari faktor-faktor yang memengaruhi stunting secara lebih menyeluruh.

Terakhir, hasil penelitian ini dapat dikembangkan ke dalam bentuk aplikasi berbasis web atau mobile sebagai alat bantu skrining awal bagi tenaga kesehatan dan kader posyandu. Namun, sebelum diimplementasikan secara luas, model perlu diuji menggunakan dataset yang lebih besar dan lebih beragam untuk memastikan bahwa performa yang dihasilkan benar-benar stabil dan dapat diterapkan pada kondisi nyata di lapangan.

DAFTAR PUSTAKA

- ABRO, A. A., TALPUR, M. S. H., JUMANI, A. K., SIDDIQUE, W. A., & YAŞAR, E. (2022). Voting combinations-based ensemble: A hybrid approach. *Celal Bayar Üniversitesi Fen Bilimleri Dergisi*, 18(3), 257–263. <https://doi.org/10.18466/cbayarfbe.1014724>
- Aftab, J., Iqbal, M. J., Mehmood, R., & Hussain, M. (2025). Artificial intelligence based classification and prediction of medical imaging modalities using deep learning. *Scientific Reports*, 15, Article 93718. <https://doi.org/10.1038/s41598-025-93718-7>
- Ashillah, S., Gulo, S. A., Rahmi, Y., & Niska, D. Y. (2025). Implementasi algoritma pengolahan citra digital untuk perbaikan kualitas gambar. *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, 8(3), 274–280. <https://doi.org/10.47970/siskom-kb.v8i3.823>
- Azhar, A., Siswoyo, B., Pratama, D., Anam, K., & Susana, H. (2024). Penerapan algoritma Convolutional Neural Network (CNN) untuk diagnosa tumor otak. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 1797–1801. <https://doi.org/10.36040/jati.v8i2.8242>
- Chen, C., Liu, X., Ding, M., Zheng, J., & Li, J. (2025). A review of convolutional neural network based methods for medical image classification. *Computers in Biology and Medicine*, 185, Article 109566. <https://doi.org/10.1016/j.combiomed.2024.109566>

- Dafa Handika N., & Rahmadewi, R. (2021). Pengelolaan citra digital: Perbandingan studi kasus antara klasifikasi warna menggunakan pengolahan model warna HSV dan penerapan metode konvolusi dalam pengolahan citra digital. *Jurnal Teknologi Informasi*, 32(3), 167–186.
- Febriyanti, N. R., Kusriani, K., & Hartanto, A. D. (2025). Analisis perbandingan algoritma SVM, Random Forest dan Logistic Regression untuk prediksi stunting balita. *Edumatic: Jurnal Pendidikan Informatika*, 9(1), 149–158. <https://doi.org/10.29408/edumatic.v9i1.29407>
- Fitriyani Sahamony, N., & Rianto, H. (2024). Analysis of performance comparison of machine learning models for predicting stunting risk in children's growth. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 413–422.
- Hamid, M. A., Supriyanto, C., & Wibowo, A. (2025). Performance comparison of Random Forest, SVM, and XGBoost algorithms with SMOTE for stunting prediction. *Journal of Applied Informatics and Computing*, 9(2), 1–10.
- Jain, A., Kumar, A., & Susan, S. (2021). Evaluating deep neural network ensembles by majority voting cum meta-learning scheme. *Neural Processing Letters*, 53, 3471–3490. <https://doi.org/10.1007/s11063-021-10515-2>
- Kementerian Kesehatan Republik Indonesia. (2022). *Profil kesehatan Indonesia 2022*. Kementerian Kesehatan Republik Indonesia.
- Kementerian Kesehatan Republik Indonesia. (2023). *Survei kesehatan Indonesia (SKI) 2023 dalam angka*. Badan Kebijakan Pembangunan Kesehatan, Kementerian Kesehatan Republik Indonesia.
- Kusuma, T. (2024). Studi algoritma deep learning dalam pengolahan citra digital. *Jurnal Dunia Data*, 1(5), 1–15.
- Li, Y., Daho, M. E. H., Conze, P. H., Zeghlache, R., Le Boité, H., Tadayoni, R., Cochener, B., Lamard, M., & Quéllec, G. (2024). A review of deep learning-based information fusion techniques for multimodal medical image classification. *Information Fusion*, 107, Article 102305. <https://doi.org/10.1016/j.inffus.2024.102305>
- Lase, Y. Y., Prasetyo, H., & Nugroho, A. (2026). Optimization performance of Extreme Gradient Boosting and Random Forest algorithms for stunting classification based on economic factors. *Jurnal Teknik Informatika (JUTIF)*, 7(1), 1–10.
- Marthasari, G. I., Hayatin, N., & Yuniarti, M. (2022). Content classification based-on Latent Semantic Analysis and Support Vector Machine (LSA-SVM). *Jurnal Transformatika*, 19(2), 144–150. <https://doi.org/10.26623/transformatika.v19i2.2745>
- Mienye, I. D., Sun, Y., & Wang, Z. (2025). Deep convolutional neural networks in medical image analysis: A review. *Information*, 16(3), 195. <https://doi.org/10.3390/info16030195>
- Mustafa, W. A., Yazid, H., & Kader, M. M. A. (2025). Advances in medical image classification using deep learning and feature fusion techniques. *Diagnostics*, 15(8), Article 988. <https://doi.org/10.3390/diagnostics15080988>
- Ningrum, S., Barata, M. A., & Mahmudah, N. (2025). Klasifikasi status stunting pada balita di Kecamatan Singahan dengan algoritma machine learning. *SMARTICS Journal*, 11(1), 28–33. <https://doi.org/10.21067/smartics.v11i1.11862>
- Pratama, I. P. A., Putra, I. M. A., & Dewi, N. K. C. (2024). Evaluating the performance of voting classifier in multiclass classification. *International Journal of Data Science*, 5(2), 85–94.
- Saeed, E. M. H. (2025). An ensemble voting classifier based on machine learning models for phishing detection. *International Journal of Scientific Research in Science, Engineering and Technology*, 12(1), 15–27. <https://doi.org/10.32628/ijsrset251211>
- Sarvamangala, D. R., & Kulkarni, R. V. (2021). Convolutional neural networks in medical image understanding: A survey. *Evolutionary Intelligence*, 15, 1–22. <https://doi.org/10.1007/s12065-020-00540-3>
- Singh, P., Kumar, R., & Sharma, A. (2025). An ensemble-driven machine learning framework for classification and prediction using voting classifier. *Discover Artificial Intelligence*, 5, Article 1467. <https://doi.org/10.1007/s43621-025-01467-4>
- Siregar, A. P., Rahman, F., & Putri, D. A. (2024). Comparative analysis of machine learning models for early detection of stunting risk in toddlers. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(4), 721–730.
- Thaariq, M. A., Baskara, M. D. M., & Chaniago, R. A. (2024). Systematic literature review: Analisis penerapan machine learning dalam bidang kesehatan. *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, 168–173.
- Unjung, J., Kim, H., & Lee, S. (2025). Soft voting ensemble model to improve Parkinson's disease prediction using machine learning. *International Journal of Artificial Intelligence*, 14(1), 45–56.

- UNICEF, WHO, & World Bank Group. (2021). *Levels and trends in child malnutrition: UNICEF/WHO/World Bank Group joint child malnutrition estimates—Key findings of the 2021 edition*. World Health Organization.
- UNICEF, WHO, & World Bank Group. (2023). *Levels and trends in child malnutrition: UNICEF/WHO/World Bank Group joint child malnutrition estimates—Key findings of the 2023 edition*. World Health Organization.
- UNICEF, WHO, & World Bank Group. (2025). *Levels and trends in child malnutrition: UNICEF/WHO/World Bank Group joint child malnutrition estimates—Key findings of the 2025 edition*. World Health Organization.
- Wardhana, R. G., Wang, G., & Sibuea, F. (2023). Penerapan machine learning dalam prediksi tingkat kasus penyakit di Indonesia. *Journal of Information System Management (JOISM)*, 5(1), 40–45.
- Widyawati, S. B. D., Prasetyo, A., & Setiawan, R. (2025). Comparative analysis of Random Forest, Logistic Regression, and Decision Tree for toddler nutritional health prediction. *International Journal of Intelligent Systems and Applications*, 13(4), 115–124.
- World Health Organization. (2021). *WHO child growth standards and the identification of severe acute malnutrition in infants and children*. World Health Organization.
- World Health Organization. (2022). *Global nutrition targets 2025: Stunting policy brief*. World Health Organization.
- World Health Organization. (2023). *Child malnutrition: Key facts and global estimates*. World Health Organization.
- World Health Organization. (2024). *Malnutrition in children: Global health estimates and prevention strategies*. World Health Organization.
- World Health Organization. (2025). *Nutrition and food safety: Child growth and malnutrition monitoring*. World Health Organization.
- Yuliana, R., Sari, D. P., & Nugroho, H. A. (2023). Implementation of Support Vector Machine for classification of nutritional status in toddlers. *Indonesian Journal of Artificial Intelligence and Data Mining*, 6(2), 90–98.
- Zhang, Y., Wang, L., Chen, X., & Liu, H. (2023). Feature fusion-based deep learning model for medical image classification. *Biomedical Signal Processing and Control*, 86, Article 105295. <https://doi.org/10.1016/j.bspc.2023.105295>
- Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., van Ginneken, B., Madabhushi, A., Prince, J. L., Rueckert, D., & Summers, R. M. (2021). A review of deep learning in medical imaging: Imaging traits, technology trends, case studies, and progress highlights. *Proceedings of the IEEE*, 109(5), 820–838. <https://doi.org/10.1109/JPROC.2021.3054390>
- Zulfahmi, M., Saputra, R., & Hidayat, T. (2025). Application of XGBoost-based machine learning methods for stunting prediction using infant health data. *International Journal of Artificial Intelligence Research*, 9(1), 1–12.
- Zulfa, N., Widodo, A., & Kurniawan, D. (2024). Feature extraction using Convolutional Neural Network for image-based classification. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 8(4), 512–520.