

Deteksi Menu dan Estimasi Kandungan Nutrisi Menggunakan YOLOv11 pada Program Makan Bergizi Gratis

M. Yogi Dwiyanto^{1*}, Teguh Iman Hermanto², M. Rafi Muttaqin³

^{1,2,3} Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Wastukencana, Purwakarta, Indonesia

Email: m.yogi73@wastukencana.ac.id^{1*}, teguhiman@wastukencana.ac.id², rafi@wastukencana.ac.id

INFORMASI ARTIKEL

Histori artikel:

Naskah masuk: 13 Juni 2026

Direvisi: 16 Juli 2026

Diterima : 31 Juli 2026

Kata Kunci:

computer vision,
YOLOv11,
deteksi makanan,
estimasi nutrisi,
makanan bergizi gratis

ABSTRAK

Abstract- The Free Nutritious Meal Program requires a rapid and standardized mechanism to verify menu components and provide nutritional information. This study developed a computer-vision system using YOLOv11 to detect food items and estimate nutritional content from meal images. Unlike previous studies that generally detect common food items or rely on international nutritional databases, this study specifically targets 30 MBG menu component classes and integrates detection results with the Indonesian Food Composition Table 2020 as a national nutritional reference. The research followed the CRISP-DM stages: business understanding, data understanding, data preparation, modelling, evaluation, and deployment. A dataset of 300 images containing 30 food classes and 1,541 annotated objects was prepared in Roboflow and divided into training, validation, and testing sets at an 80:10:10 ratio. The YOLOv11 model, consisting of 25.37 million parameters, was trained with dynamic augmentation and stopped at epoch 42 through early stopping. Testing produced 75.19% precision, 73.98% recall, a 74.58% F1-score, 82.64% mAP50, and 77.21% mAP50-95. Detection labels were mapped to the Indonesian Food Composition Table 2020 and complementary USDA data to estimate energy, protein, fat, carbohydrate, and fibre per standardized 100-gram portion. The model was deployed through a Streamlit web application that displays bounding boxes, food labels, nutritional estimates, and their contribution to nutritional adequacy values. The results indicate that YOLOv11 can support practical menu identification and standardized nutritional monitoring, although larger and more balanced datasets and automatic portion estimation are still required.

Abstrak- Program Makan Bergizi Gratis memerlukan mekanisme yang cepat dan terstandarisasi untuk memverifikasi komponen menu serta menyajikan informasi nutrisi. Penelitian ini mengembangkan sistem computer vision menggunakan YOLOv11 untuk mendeteksi objek makanan dan mengestimasi kandungan nutrisi melalui citra menu. Berbeda dari penelitian sebelumnya yang umumnya mendeteksi makanan secara umum atau menggunakan basis data nutrisi internasional, penelitian ini secara khusus menargetkan 30 kelas komponen menu MBG dan mengintegrasikan hasil deteksi dengan Tabel Komposisi Pangan Indonesia 2020 sebagai basis data nutrisi nasional. Penelitian mengikuti tahapan CRISP-DM, yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan. Dataset berjumlah 300 citra yang mencakup 30 kelas makanan dan 1.541 objek terannotasi dipersiapkan melalui Roboflow, kemudian dibagi menjadi data latih, validasi, dan pengujian dengan rasio 80:10:10. Model YOLOv11 dengan 25,37 juta parameter dilatih menggunakan augmentasi dinamis dan berhenti pada epoch ke-42 melalui early stopping. Pengujian menghasilkan precision 75,19%, recall 73,98%, F1-score 74,58%, mAP50 82,64%, dan mAP50-95 77,21%. Label hasil deteksi dipetakan ke Tabel Komposisi Pangan Indonesia 2020 dan data pelengkap USDA untuk mengestimasi energi, protein, lemak, karbohidrat, dan serat berdasarkan porsi standar 100 gram. Model diterapkan pada aplikasi web Streamlit yang menampilkan kotak deteksi, label makanan,

estimasi nutrisi, dan kontribusinya terhadap angka kecukupan gizi. Hasil menunjukkan bahwa YOLOv11 dapat mendukung identifikasi menu dan pemantauan nutrisi secara praktis, meskipun masih diperlukan dataset yang lebih besar dan seimbang serta estimasi porsi otomatis.

Copyright © 2026 LPPM - STMIK IKMI Cirebon
This is an open access article under the CC-BY license

Penulis Korespondensi:

M. Yogi Dwiyanto

Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Wastukencana, Purwakarta
Indonesia
Email: m.yogi73@wastukencana.ac.id

1. Pendahuluan

Pemenuhan gizi seimbang pada anak usia sekolah merupakan bagian penting dalam pembangunan sumber daya manusia karena kecukupan energi dan zat gizi berkaitan dengan pertumbuhan fisik, fungsi metabolik, kapasitas kognitif, serta kesiapan belajar [1], [2]. Indonesia masih menghadapi ketidakcukupan konsumsi pangan yang berdampak terhadap kualitas kesehatan dan pendidikan. Kondisi tersebut mendorong pelaksanaan Program Makan Bergizi Gratis yang menargetkan peserta didik dari pendidikan anak usia dini sampai pendidikan menengah. Besarnya volume distribusi harian menuntut mekanisme pengawasan menu yang cepat, konsisten, dan dapat diterapkan pada banyak lokasi.

Pengawasan menu di sekolah pada umumnya masih mengandalkan pemeriksaan visual dan uji organoleptik oleh petugas. Cara tersebut penting untuk memastikan makanan diterima dalam kondisi sesuai, tetapi identifikasi komponen menu dan pemeriksaan informasi nutrisi secara manual berpotensi berbeda antarpetugas, memerlukan waktu, dan sulit diskalakan. Permasalahan ini membuka peluang penerapan computer vision untuk mengenali objek makanan secara otomatis dari citra, kemudian menghubungkan hasil pengenalan dengan basis data komposisi pangan.

Penelitian terdahulu menunjukkan bahwa deteksi makanan berbasis YOLO dapat digunakan untuk mengidentifikasi objek sekaligus memperkirakan kalori melalui pemetaan setiap label ke data nutrisi [3]. Model Single Shot Multibox Detector juga telah mencapai mAP 77,20% pada deteksi beberapa jenis makanan, tetapi performa antar kelas masih beragam karena kemiripan visual dan ketidakseimbangan data [4]. Pendekatan Mask R-CNN menghasilkan akurasi 78%, namun mengalami salah klasifikasi pada objek dengan warna dan tekstur serupa [5].

Sementara itu, integrasi pengenalan bahan dengan basis data nutrisi dilaporkan mampu menghasilkan estimasi atribut gizi sekitar 85% [6]. Kajian sistematis menegaskan bahwa transparansi sumber dan proses perhitungan nutrisi diperlukan agar hasil aplikasi dapat dipahami dan dipercaya pengguna [7].

Dari berbagai pendekatan deteksi objek yang telah dikaji, arsitektur YOLO dipilih karena kemampuannya memproses seluruh citra dalam satu lintasan jaringan (*single-pass*) sehingga menghasilkan kecepatan inferensi yang jauh lebih tinggi dibandingkan pendekatan dua tahap seperti Mask R-CNN, tanpa mengorbankan akurasi secara signifikan [5]. Dibandingkan SSD yang mengandalkan fitur dari satu skala, YOLO menggunakan representasi multiskala yang lebih baik dalam mendeteksi objek berukuran kecil dan berdekatan, kondisi yang umum dijumpai pada citra menu makanan dengan banyak komponen dalam satu wadah [4]. Sementara itu, pendekatan berbasis Transformer seperti DETR menawarkan kemampuan deteksi yang kuat namun memerlukan sumber daya komputasi yang lebih besar dan waktu pelatihan yang lebih lama, sehingga kurang praktis untuk dataset berskala terbatas [14].

Di antara varian YOLO, YOLOv11 dipilih karena menghadirkan peningkatan arsitektural yang relevan untuk tugas deteksi makanan multiobjek. Blok C3k2 meningkatkan efisiensi ekstraksi fitur dengan parameter yang lebih sedikit dibanding pendahulunya, modul SPPF mempertahankan konteks spasial pada berbagai skala, sedangkan blok C2PSA menambahkan mekanisme atensi spasial yang memungkinkan model memusatkan perhatian pada area objek yang relevan meskipun terdapat latar belakang yang kompleks [9]. Dibandingkan YOLOv8 yang menjadi basis pengembangannya, YOLOv11 dilaporkan mencapai akurasi yang lebih tinggi dengan jumlah parameter yang lebih efisien [8], [9]. Studi terkini pada deteksi pangan untuk pencegahan stunting menggunakan

YOLOv11 memperoleh mAP@0.5 sebesar 97%, meskipun menggunakan dataset yang lebih besar dan berorientasi pada edukasi gizi [10]. Karakteristik ini menjadikan YOLOv11 sebagai pilihan yang tepat untuk mendeteksi 30 kelas komponen menu MBG secara bersamaan dalam satu citra dengan kebutuhan komputasi yang dapat dijalankan pada lingkungan pengembangan standar. Pada konteks menu Program Makan Bergizi Gratis, masih diperlukan penelitian yang secara khusus menggabungkan deteksi multiobjek, pemetaan ke Tabel Komposisi Pangan Indonesia, perbandingan dengan Angka Kecukupan Gizi, dan penerapan dalam aplikasi yang mudah digunakan.

Kebaruan penelitian ini terletak pada pengembangan alur terintegrasi yang mendeteksi 30 komponen menu melalui YOLOv11, memetakan setiap label ke data nutrisi TKPI 2020 dan data pelengkap USDA, menghitung kontribusi zat gizi berdasarkan porsi standar, serta menyajikan hasil melalui aplikasi Streamlit. Penelitian bertujuan mengimplementasikan YOLOv11 untuk identifikasi menu MBG dan mengevaluasi kinerjanya menggunakan precision, recall, F1-score, mAP50, mAP50-95, serta confusion matrix.

2. Metode Penelitian

Penelitian menggunakan kerangka Cross-Industry Standard Process for Data Mining (CRISP-DM), yang terdiri atas business understanding, data understanding, data preparation, modelling, evaluation, dan deployment [11]. Pada tahap business understanding, kebutuhan sistem dirumuskan sebagai alat bantu bagi petugas sekolah untuk mengenali komponen menu secara otomatis dan memperoleh informasi nutrisi tanpa menggantikan verifikasi ahli gizi.

2.1 Dataset dan Persiapan Data

Dataset terdiri atas 300 citra menu makanan yang dikumpulkan dari sumber primer dan sekunder. Citra mencakup 30 kelas makanan, antara lain nasi, ayam goreng, telur, tahu, tempe, sayuran campuran, buah, dan susu. Proses anotasi dilakukan pada Roboflow menggunakan bantuan Segment Anything Model 3.0, kemudian hasilnya diperiksa dan disesuaikan secara manual. Seluruh anotasi dikonversi ke format YOLO dan menghasilkan 1.541 objek. Dataset dibagi menjadi 240 citra pelatihan, 30 citra validasi, dan 30 citra pengujian atau rasio 80:10:10.

Eksplorasi data menunjukkan distribusi kelas tidak seimbang. Kelas nasi memiliki 255 objek, ayam goreng 171 objek, dan sayuran campuran 152 objek, sedangkan beberapa kelas lain mempunyai sampel yang lebih sedikit. Kondisi ini dipertimbangkan dalam interpretasi confusion matrix karena kelas dengan sampel terbatas

cenderung menghasilkan variasi performa yang lebih tinggi.

2.2 Pemodelan YOLOv11

Model yang digunakan adalah YOLOv11 dengan bobot pralatih. Arsitektur tersebut memiliki 357 layer dan sekitar 25,37 juta parameter. Pelatihan dilakukan pada lingkungan Google Colaboratory menggunakan GPU NVIDIA 1030GT dengan VRAM 2 GB dan RAM sistem 8 GB. Framework yang digunakan adalah Ultralytics YOLOv11 berbasis Python dengan PyTorch sebagai backend. Pelatihan dikonfigurasi pada citra berukuran 640×640 piksel, batch size 16, learning rate awal 0,01 dengan optimizer otomatis, dan target maksimum 50 epoch. Augmentasi dinamis mencakup rotasi $\pm 15^\circ$, shear $\pm 10^\circ$, perubahan skala hingga 50%, translasi hingga 10%, pembalikan horizontal, mosaic, mixup, serta variasi hue, saturation, dan value. Proses pelatihan berhenti pada epoch ke-42 melalui mekanisme early stopping dengan nilai patience 10, dan diselesaikan dalam waktu sekitar 0,275 jam atau sekitar 16,5 menit.

2.3 Evaluasi dan Estimasi Nutrisi

Performa model dievaluasi pada data pengujian menggunakan precision, recall, F1-score, mAP pada ambang Intersection over Union 0,50, mAP pada rentang 0,50–0,95, dan confusion matrix. F1-score dihitung sebagai rata-rata harmonik precision dan recall. Nilai-nilai tersebut digunakan untuk menilai ketepatan prediksi, kemampuan menemukan objek relevan, keseimbangan keduanya, dan kualitas lokalisasi objek.

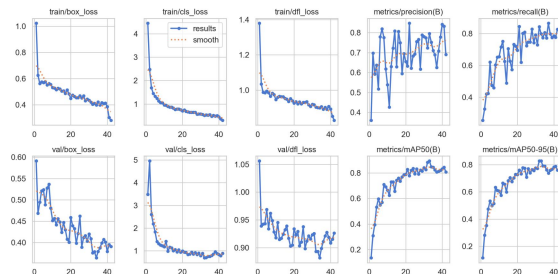
Setiap label makanan dipetakan ke nilai energi, protein, lemak, karbohidrat, dan serat per 100 gram berdasarkan TKPI 2020 [12]. Untuk edamame, kelengkeng, kurma, pakcoy, dan stroberi yang tidak tersedia secara memadai pada sumber utama, digunakan data pelengkap USDA FoodData Central [13]. Total nutrisi diperoleh dengan menjumlahkan nilai setiap objek yang terdeteksi berdasarkan asumsi porsi standar 100 gram. Persentase kecukupan dihitung dengan membagi estimasi nutrisi terhadap nilai referensi AKG, kemudian dikalikan 100%. Sistem diterapkan menggunakan Python dan Streamlit.

3. Hasil dan Pembahasan

3.1 Hasil Pelatihan dan Evaluasi

Pelatihan berhenti pada epoch ke-42 melalui mekanisme *early stopping* dengan nilai *patience* 10, dari target awal 50 epoch. Selama pelatihan, *box loss*, *classification loss*, dan *distribution focal loss* menunjukkan tren penurunan yang konsisten pada data pelatihan maupun validasi. Nilai *precision*, *recall*, *mAP50*, dan *mAP50-95* meningkat secara

bertahap dan mencapai kondisi relatif stabil menjelang akhir pelatihan. Pola konvergensi ini menunjukkan bahwa model berhasil mempelajari representasi fitur visual objek makanan tanpa indikasi overfitting yang signifikan, sebagaimana ditunjukkan pada Gambar 1.



Gambar 1. Grafik loss dan metrik evaluasi selama pelatihan model YOLOv11

Tabel 1. Hasil evaluasi model YOLOv11

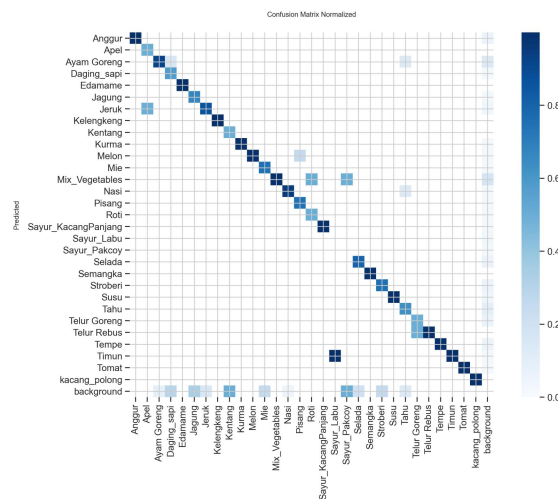
Metrik	Nilai (%)
Precision	75,19
Recall	73,98
F1-score	74,58
mAP50	82,64
mAP50-95	77,21

Precision 75,19% menunjukkan bahwa sekitar tiga perempat prediksi objek sesuai dengan kelas sebenarnya, sedangkan recall 73,98% menunjukkan bahwa sebagian besar objek aktual berhasil ditemukan. Selisih kecil antara precision dan recall menghasilkan F1-score 74,58% yang mencerminkan keseimbangan yang cukup baik antara ketepatan dan kelengkapan deteksi. Nilai mAP50 82,64% menandakan kemampuan lokalisasi yang baik pada ambang IoU 0,50, sedangkan mAP50-95 77,21% menunjukkan bahwa performa tetap relatif kuat ketika kriteria tumpang tindih kotak diperketat hingga IoU 0,95. Selisih sekitar 5,4 poin antara mAP50 dan mAP50-95 mengindikasikan bahwa model masih memiliki ruang peningkatan pada akurasi lokalisasi *bounding box* secara presisi, yang umum terjadi pada dataset dengan variasi sudut pengambilan gambar dan ukuran objek yang beragam.

Dibandingkan penelitian SSD yang memperoleh mAP 77,20% [4], model dalam penelitian ini menghasilkan mAP50 yang lebih tinggi sebesar 5,44 poin. Namun, hasilnya masih berada di bawah studi YOLOv11 untuk pangan tinggi protein yang mencapai mAP@0.5 97% [10]. Perbedaan tersebut dapat dijelaskan oleh ukuran dataset, jumlah kelas, distribusi sampel, variasi visual, dan kompleksitas

citra multiobjek. Dataset penelitian ini hanya terdiri atas 300 citra dengan 30 kelas, sehingga rata-rata contoh per kelas relatif terbatas dibandingkan studi pembandingan.

Analisis *confusion matrix* yang dinormalisasi disajikan pada Gambar 2 untuk memberikan gambaran distribusi kesalahan klasifikasi secara menyeluruh.



Gambar 2. Confusion matrix ternormalisasi model YOLOv11 pada data pengujian

Berdasarkan *confusion matrix*, kelas dengan performa deteksi terbaik adalah nasi, ayam goreng, sayuran campuran, tempe, tahu, jeruk, dan susu, yang ditunjukkan oleh dominasi nilai tinggi pada diagonal utama. Keberhasilan deteksi kelas-kelas tersebut didukung oleh jumlah sampel yang relatif banyak serta karakteristik visual yang cukup khas dan berbeda dari kelas lain.

Sebaliknya, Kesalahan deteksi terjadi pada beberapa kelompok kelas dengan pola yang dapat diidentifikasi. Pertama, kesalahan akibat kemiripan visual warna dan tekstur, di mana beberapa jenis sayuran seperti sayur labu, sayur kacang panjang, dan sayur pakcoy memiliki warna hijau dan tekstur permukaan yang serupa. Kedua, kesalahan akibat ketidakseimbangan jumlah data, di mana kelas dengan sampel terbatas seperti kurma (8 objek), sayur labu (10 objek), dan edamame (11 objek) cenderung rentan terhadap salah klasifikasi. Ketiga, kesalahan akibat kesamaan konteks penyajian, di mana lauk-pauk yang digoreng seperti tahu, tempe, dan ayam goreng memiliki warna kecokelatan yang serupa sehingga berpotensi menimbulkan kebingungan klasifikasi.

3.2 Integrasi Estimasi Nutrisi

Model terlatih berhasil diintegrasikan dengan basis data nutrisi. Ketika citra diunggah, sistem menghasilkan bounding box, label, dan confidence untuk setiap objek. Label tersebut digunakan

sebagai kunci untuk mengambil nilai energi, protein, lemak, karbohidrat, dan serat. Metode pemetaan basis data dipilih karena estimasi berat langsung dari satu citra dua dimensi memiliki ambiguitas skala dan kedalaman. Pendekatan ini sejalan dengan penelitian yang menyatakan bahwa integrasi pengenalan makanan dengan basis data gizi merupakan strategi praktis untuk meningkatkan keterlacakan perhitungan [6], [7].

Tabel 2. Contoh hasil estimasi pada satu citra menu

Komponen	Hasil
Objek terdeteksi	Nasi, ayam goreng, tahu, jeruk, timun
Energi	588 kkal
Protein	52,4 g
Lemak	17,6 g
Karbohidrat	56,6 g
Serat	5,9 g

Pada contoh pengujian, sistem mendeteksi lima objek berupa nasi, ayam goreng, tahu, jeruk, dan timun. Dengan asumsi setiap objek mewakili 100 gram, total estimasi mencapai 588 kkal energi, 52,4 gram protein, 17,6 gram lemak, 56,6 gram karbohidrat, dan 5,9 gram serat. Hasil kemudian dibandingkan dengan referensi AKG sehingga pengguna memperoleh gambaran kontribusi menu terhadap kebutuhan harian.

Nilai estimasi harus dipahami sebagai pendekatan berbasis porsi standar, bukan pengukuran massa aktual. Kelebihannya adalah proses perhitungan sederhana, cepat, transparan, dan dapat ditelusuri ke basis data. Keterbatasannya, dua porsi dengan jenis makanan sama akan memperoleh estimasi identik meskipun ukuran fisiknya berbeda. Dengan demikian, sistem lebih tepat digunakan sebagai alat bantu skrining dan informasi awal daripada sebagai pengganti penilaian ahli gizi.

3.3 Implementasi Aplikasi

Aplikasi Streamlit menyediakan halaman unggah citra, referensi AKG, data nutrisi makanan, serta tampilan hasil deteksi dan estimasi. Pengguna dapat mengunggah format JPG, JPEG, PNG, atau WEBP. Sistem kemudian memproses citra dan menampilkan bounding box, nama kelas, confidence, rincian nutrisi per objek, total nutrisi, dan persentase pemenuhan AKG. Desain ini mendukung keterjelasan proses karena sumber data dan asumsi porsi dapat diketahui pengguna.

Secara praktis, aplikasi dapat membantu petugas melakukan pemeriksaan awal terhadap kesesuaian komponen menu. Secara ilmiah, penelitian menunjukkan bahwa pipeline deteksi objek dapat diperluas menjadi sistem pendukung keputusan nutrisi melalui pemetaan data yang terstruktur. Walaupun demikian, validasi lapangan, pengujian waktu inferensi pada perangkat sekolah, dan evaluasi usability belum dilakukan secara luas sehingga efektivitas operasionalnya masih perlu diuji.

4. Kesimpulan

Model YOLOv11l berhasil diterapkan untuk mendeteksi 30 kelas makanan pada menu Program Makan Bergizi Gratis menggunakan 300 citra dan 1.541 objek teranotasi. Hasil pengujian memperoleh precision 75,19%, recall 73,98%, F1-score 74,58%, mAP50 82,64%, dan mAP50-95 77,21%. Model mampu mengenali mayoritas objek dengan baik, tetapi masih mengalami kesalahan pada kelas yang memiliki kemiripan visual dan jumlah sampel terbatas. Integrasi hasil deteksi dengan TKPI 2020 dan USDA memungkinkan sistem menghitung estimasi energi, protein, lemak, karbohidrat, dan serat, kemudian membandingkannya dengan referensi AKG. Implementasi melalui Streamlit menghasilkan aplikasi yang dapat memberikan deteksi dan informasi nutrisi secara cepat serta transparan.

Secara praktis, sistem yang dikembangkan berpotensi mendukung pelaksanaan Program Makan Bergizi Gratis pada skala operasional dengan menyediakan alat bantu verifikasi menu yang lebih cepat, konsisten, dan dapat didokumentasikan dibandingkan pemeriksaan visual manual. Petugas penerima di sekolah dapat memanfaatkan aplikasi ini untuk melakukan pemeriksaan awal terhadap kelengkapan komponen menu dan memperoleh gambaran kandungan nutrisi secara instan, sehingga mendukung standarisasi pengawasan gizi di berbagai lokasi distribusi secara bersamaan. Dalam konteks program nasional yang menargetkan jutaan penerima manfaat, pendekatan berbasis computer vision seperti ini dapat menjadi komplemen teknologi yang membantu menjaga konsistensi kualitas gizi tanpa menambah beban kerja petugas secara signifikan.

Kelebihan penelitian terletak pada integrasi deteksi multiobjek, basis data nutrisi nasional, evaluasi kecukupan, dan aplikasi web dalam satu alur. Keterbatasannya adalah ukuran dan ketidakseimbangan dataset, asumsi porsi tetap 100 gram, belum adanya estimasi volume atau berat aktual, serta belum dilakukannya pengujian pengguna pada lingkungan sekolah. Penelitian selanjutnya perlu memperbesar dataset, melakukan perbandingan dengan model lain, menambahkan

estimasi porsi berbasis objek referensi atau sensor kedalaman, dan mengembangkan aplikasi mobile yang tervalidasi bersama ahli gizi.

Daftar Pustaka

- [1] D. Rahmawati, R. Hasibuan, L. Nazar, H. Ratnaningsih, and R. Aliyah, "Peran kepedulian orang tua terhadap dampak nutrisi dalam perkembangan kognitif dan motorik anak pada pendidikan anak usia dini," *Jurnal Madinasika Manajemen Pendidikan dan Keguruan*, vol. 6, pp. 158–165, 2025, doi: 10.31949/madinasika.v6i2.13996.
- [2] J. N. Aulia, "Masalah gizi pada anak usia sekolah," *Jurnal Ilmiah Kesehatan Media Husada*, vol. 11, no. 1, pp. 21–25, 2022.
- [3] F. Romadhon et al., "Food image detection system and calorie content estimation using YOLO to control calorie intake in the body," *E3S Web of Conferences*, vol. 465, 2023, doi: 10.1051/e3sconf/202346502057.
- [4] F. A. Puspitasari and H. Pangaribuan, "Deteksi kalori pada citra makanan dengan algoritma Single Shot Multibox Detector," *Jurnal Comasie*, vol. 12, no. 2, pp. 40–51, 2025.
- [5] R. A. Shafa and P. N. Andono, "Pendeteksi visual makanan dan jumlah kalorinya menggunakan algoritma Mask R-CNN berbasis bot Telegram," *JUPI*, vol. 10, no. 1, pp. 641–651, 2025, doi: 10.29100/jupi.v10i1.6972.
- [6] M. Al-Saffar and W. R. Baiee, "Nutrition information estimation from food photos using machine learning based on multiple datasets," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 5, pp. 2922–2929, 2022, doi: 10.11591/eei.v11i5.4007.
- [7] L. M. Amugongo, A. Kriebitz, A. Boch, and C. Lütge, "Mobile computer vision-based applications for food recognition and volume and calorific estimation: A systematic review," *Healthcare*, vol. 11, no. 1, 2023, doi: 10.3390/healthcare11010059.
- [8] M. Kotthapalli, D. Ravipati, and R. Bhatia, "YOLOv1 to YOLOv11: A comprehensive survey of real-time object detection innovations and challenges," *arXiv preprint arXiv:2508.02067*, Aug. 2025. [Online]. Available: <https://arxiv.org/abs/2508.02067>
- [9] R. Khanam and M. Hussain, "YOLOv11: An overview of the key architectural enhancements," *arXiv preprint arXiv:2410.17725*, Oct. 2024. [Online]. Available: <https://arxiv.org/abs/2410.17725>
- [10] A. D. Restu, L. Erba, C. K. R. Meiranadi, S. G. Wilujeng, and M. Wildan, "Implementation of YOLOv11 for food detection to support nutritional information in stunting prevention," *Sinkron*, vol. 10, no. 1, pp. 1–15, 2026, doi: 10.33395/sinkron.v10i1.15553.
- [11] L. E. Dendy and R. Tanamal, "Association rule mining for truck body damage pattern analysis using Apriori and CRISP-DM," *Journal of Applied Informatics and Computing*, 2026.
- [12] Kementerian Kesehatan Republik Indonesia, *Tabel Komposisi Pangan Indonesia 2020*. Jakarta: Kementerian Kesehatan Republik Indonesia, 2020.
- [13] U.S. Department of Agriculture, "FoodData Central," 2026.
- [14] J. P. F. Kristanto, D. A. Prabowo, and Y. S. R. Nur, "Food detection to estimate calories using Detection Transformer," *Jurnal Teknokes*, vol. 18, no. 4, pp. 337–346, 2025, doi: 10.35882/jteknokes.v18i4.132.
- [15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.
- [16] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with Transformers," in *Proc. European Conf. Computer Vision (ECCV)*, Glasgow, UK, Aug. 2020, pp. 213–229, doi: 10.1007/978-3-030-58452-8_13.