

IMPLEMENTATION OF THE RT-DETR MODEL FOR FOOD NUTRITION ESTIMATION ON A WEBSITE-BASED MONITORING SYSTEM FOR PROGRAM MAKAN BERGIZI GRATIS (MBG)

PENERAPAN MODEL RT-DETR UNTUK ESTIMASI NUTRISI MAKANAN PADA *WEBSITE MONITORING* PROGRAM MAKAN BERGIZI GRATIS (MBG)

Muhammad Alif Fadillah^{1‡}, Eneng Tita Tosida², and Asep Saepulrohman³

^{1,2,3}Department of Computer Science, Universitas Pakuan, Indonesia

[‡]corresponding author: 065122109@student.unpak.ac.id

Copyright © 2026 Muhammad Alif Fadillah, Eneng Tita Tosida, and Asep Saepulrohman. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

The Program Makan Bergizi Gratis (MBG) requires a more efficient nutrition monitoring system to replace the time-consuming and error-prone manual evaluation process. This study integrates a Real-Time Detection Transformer (RT-DETR) model for food object detection and nutritional estimation, alongside a Long Short-Term Memory (LSTM) model for predicting daily nutritional trends, into an MBG monitoring website. The RT-DETR model was trained on annotated food images, while the LSTM model processes time-series data from daily nutritional history to forecast nutritional needs. Evaluation results show that RT-DETR achieved precision and recall values of 93.48% on test data. The LSTM model with a 90:10 data split delivered the best performance, with an RMSE of 38.42 and a MAPE of 28.25%. The web system, built using AstroJS and Supabase, enables real-time nutritional monitoring. However, nutritional estimates currently rely on fixed reference values per 100 grams, so future research is recommended to adopt more advanced segmentation methods, such as SegFormer, to calculate nutrition based on actual portion weight.

Keywords: Computer Vision, Food Nutrition Estimation, LSTM, RT-DETR, Website Monitoring System.

Abstrak

Program Makan Bergizi Gratis (MBG) membutuhkan sistem pemantauan nutrisi yang lebih efisien untuk menggantikan proses evaluasi manual yang memakan waktu dan rawan kesalahan. Penelitian ini mengintegrasikan model *Real-Time Detection Transformer* (RT-DETR) untuk deteksi objek makanan dan estimasi kandungan nutrisi, serta model *Long Short-Term Memory* (LSTM) untuk memprediksi tren nutrisi harian, ke dalam sebuah *website monitoring* MBG. Model RT-DETR dilatih menggunakan gambar makanan yang telah dianotasi, sementara model LSTM memproses data deret waktu dari riwayat gizi harian untuk meramalkan kebutuhan nutrisi. Hasil evaluasi menunjukkan RT-DETR mencapai nilai presisi dan *recall* sebesar 93,48% pada data pengujian. Model LSTM dengan pembagian data 90:10 menghasilkan kinerja terbaik dengan RMSE sebesar 38,42 dan MAPE sebesar 28,25%. Sistem web yang dibangun menggunakan AstroJS dan Supabase ini memungkinkan pemantauan gizi secara real-time. Namun, estimasi nutrisi saat ini masih bergantung pada nilai referensi tetap per 100 gram, sehingga penelitian selanjutnya disarankan mengadopsi metode segmentasi yang lebih canggih, seperti SegFormer, untuk menghitung nutrisi berdasarkan berat porsi yang sebenarnya.

Kata Kunci: *Computer Vision*, Estimasi Nutrisi Makanan, LSTM, RT-DETR, Sistem *Monitoring Website*.

1. Pendahuluan

Program Makan Bergizi Gratis (MBG) merupakan kebijakan intervensi pangan nasional yang dicanangkan oleh Pemerintah Republik Indonesia untuk menurunkan prevalensi malnutrisi dan stunting serta meningkatkan status gizi kelompok rentan, terutama anak sekolah, ibu hamil, dan ibu menyusui (Widyasari *et al.*, 2025). Keberhasilan program sangat bergantung pada kemampuan sistem pengelolaan dan pemantauan nutrisi makanan yang disajikan secara presisi, terukur, dan berkelanjutan. Namun pada praktiknya, evaluasi kandungan gizi masih dilakukan secara manual oleh petugas gizi, sehingga memakan waktu lama, rawan kesalahan pencatatan, dan kurang dapat diskalakan untuk implementasi tingkat nasional.

Perkembangan kecerdasan buatan, khususnya *computer vision* dan *deep learning*, mendorong tumbuhnya sistem analisis makanan berbasis citra yang lebih otomatis dan akurat. Menurut Chai *et al.* (2021), *Computer Vision* adalah disiplin yang memungkinkan komputer memproses dan mengekstrak informasi bermakna dari data visual berdimensi tinggi untuk mendukung pengambilan keputusan. Sarker (2021) menyatakan bahwa *deep learning* efektif untuk pembelajaran *end-to-end* pada data tidak terstruktur seperti citra digital. Dalam konteks estimasi nutrisi, Minaee *et al.*

(2022) menunjukkan bahwa segmentasi citra berbasis *deep learning* mampu memetakan batas objek secara presisi, sehingga relevan untuk analisis komponen makanan.

Berbagai studi mengeksplorasi estimasi nutrisi dari citra makanan. Ahn (2024) mengembangkan model *uncertainty-driven* dengan $R^2 = 0,98$ dan $RMSE = 0,40$. Bianco *et al.* (2025) membandingkan delapan arsitektur dan menemukan ViT-B-16 serta ResNet101 terbaik untuk prediksi komposisi nutrisi 2D. Feng *et al.* (2025) memperkenalkan *framework* multi-modal RGB-D berbasis FLAVA *contrastive learning*, sementara Revesai (2025) menghadirkan model ringan ~97% akurat untuk aplikasi mobile. Ahmad dan Tribuana (2024) menggunakan SSD untuk deteksi kalori makanan tradisional Indonesia; Annisa *et al.* (2025) dan Ash-shafa & Andono (2025) mengeksplorasi CNN dan Mask R-CNN untuk pengenalan dan penghitungan kalori otomatis. Namun, belum ada penelitian yang mengintegrasikan deteksi objek berbasis transformer dengan sistem monitoring berbasis web untuk program gizi pemerintah. RT-DETR (*Real-Time Detection Transformer*), arsitektur deteksi objek berbasis *Vision Transformer* yang bekerja *end-to-end* tanpa NMS, menawarkan inferensi stabil dan efisien (Zhao *et al.*, 2023; Madan dan Reich, 2025), sehingga cocok untuk mendeteksi banyak komponen makanan pada nampian MBG.

Berdasarkan permasalahan dan kajian tersebut, penelitian ini bertujuan: (1) menerapkan model RT-DETR untuk deteksi dan estimasi nutrisi objek makanan dari citra digital, (2) menerapkan model LSTM untuk prediksi tren nutrisi harian, dan (3) mengintegrasikan kedua model dalam sebuah *website monitoring* MBG yang dapat diakses secara *real-time* oleh pengelola dan umum.

2. Metodologi

2.1 Bahan dan Data

Data untuk model RT-DETR dikumpulkan secara primer melalui proses pengambilan citra langsung di salah satu SPPG yang berlokasi di Kecamatan Kemang. Pengambilan gambar dilakukan menggunakan kamera dengan resolusi 1080×1920 piksel, mencakup berbagai variasi sudut pengambilan serta kondisi pencahayaan. Informasi nilai kandungan nutrisi dari setiap jenis makanan ditentukan berdasarkan Tabel Komposisi Pangan Indonesia (TKPI). Dataset yang diperoleh terdiri dari 1.548 citra, kemudian dilakukan proses anotasi menggunakan platform Roboflow dengan metode *bounding box* serta pelabelan kelas sebagai *ground truth*. Selanjutnya, *dataset* tersebut secara otomatis dibagi menjadi data pelatihan (*training*), validasi (*validation*), dan pengujian (*testing*).



Gambar 1: Dataset Menu MBG.

Data yang digunakan pada model LSTM berupa dataset riwayat nutrisi harian MBG yang dikumpulkan secara primer dari salah satu SPPG di Kecamatan Kemang. Dataset ini disusun dalam bentuk deret waktu dan mencakup atribut total energi (kkal), lemak (g), protein (g), karbohidrat (g), serta serat (g), sebagaimana ditunjukkan pada Tabel 1.

Tabel 1: Contoh Data Gizi Harian MBG.

Tanggal	Porsi	E (kkal)	L (g)	P (g)	Karbo (g)	S (g)
20/12/2025	Besar	625	26	18	81	5,5
20/12/2025	Kecil	525	17	6	55	2,0
21/12/2025	Besar	570	19	20	85	3,4
21/12/2025	Kecil	480	18	19	65	2,9
22/12/2025	Besar	645	25	15	95	0,5

Keterangan: E=Energi, L=Lemak, P=Protein, Karbo=Karbohidrat, S=Serat

2.2 Metode Penelitian

2.2.1. Data Preprocessing dan Transformation

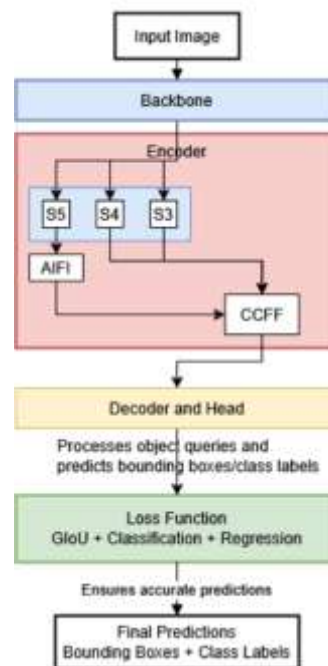
Preprocessing untuk model RT-DETR terdiri dari empat tahap utama: *Auto-Orient* untuk memperbaiki orientasi gambar, *Resize* menjadi 640×640 piksel agar ukuran input seragam, konversi ke *Grayscale* untuk menonjolkan fitur bentuk dan tekstur, serta *Auto-Adjust Contrast* untuk memperjelas objek (Jumadi dan Sartika, 2021). Setelah itu dilakukan augmentasi berupa flip *horizontal*, rotasi 90° dan $\pm 15^\circ$, *crop/zoom* 0–20%, *shear* $\pm 15\%$, penyesuaian saturasi ($\pm 25\%$) dan *brightness* ($\pm 15\%$), serta penambahan *blur* (maks. 2 piksel) dan *noise* (maks. 5% piksel).

Preprocessing untuk model LSTM mencakup konversi tanggal ke format *datetime*, pengurutan data secara kronologis, penanganan *missing value*, dan normalisasi menggunakan *Min-Max Scaling* ke rentang (Suryana dan Silaswara, 2024). Transformasi lanjut meliputi

pembuatan *sliding window* untuk *sequence*, pemformatan input-output (X, Y), *reshaping* ke bentuk 3D (*samples, timesteps, features*), serta pemisahan data *training* dan *testing* berdasarkan urutan waktu untuk mencegah kebocoran data.

2.2.2. Arsitektur Model RT-DETR

Model yang digunakan adalah RT-DETR *Large* (rtdetr-l) yang diimplementasikan menggunakan *framework* Ultralytics berbasis PyTorch. Secara garis besar arsitektur terbagi menjadi tiga komponen utama: *backbone*, *hybrid encoder*, dan *transformer decoder*. *Backbone* mengekstrak fitur visual berhirarki dari citra input; *hybrid encoder* menggabungkan fitur multi-skala lewat mekanisme *attention*; sedangkan *transformer decoder* menghasilkan prediksi *bounding box* dan kelas objek secara *end-to-end* tanpa memerlukan *Non-Maximum Suppression* (NMS) (Zhao *et al.*, 2023). Skema arsitektur RT-DETR ditampilkan pada Gambar 2.



Gambar 2: Arsitektur RT-DETR.

Berdasarkan arsitektur tersebut, spesifikasi teknis RT-DETR *Large* (rtdetr-l) tercantum pada Tabel 2, sedangkan rincian tiap komponen utama arsitektur disajikan pada Tabel 3.

Tabel 2: Spesifikasi Model RT-DETR *Large* (rtdetr-l).

Parameter	Nilai
Total Layer	465
Total Parameter	32.886.221
Kompleksitas Komputasi	108,2 GFLOPs

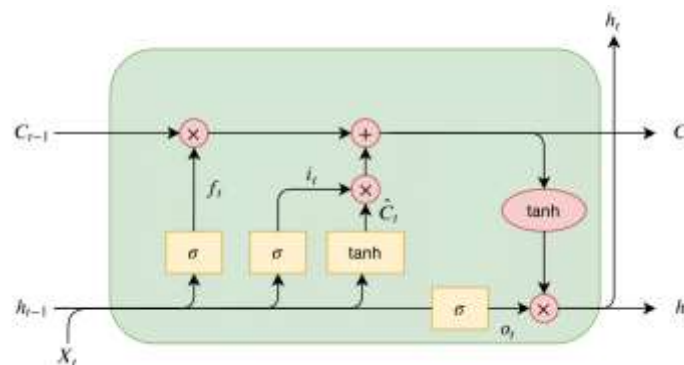
Ukuran Input Citra	640 × 640 piksel
Framework Implementasi	Ultralytics (PyTorch)

Tabel 3: Komponen Arsitektur RT-DETR.

Komponen	Modul Utama	Fungsi
Backbone	HGStem, HGBlock, DWConv	Ekstraksi fitur visual dan representasi multi- skala
Hybrid Encoder	Conv, AIFI, Upsample, Concat, RepC3	Integrasi dan penguatan fitur antar skala via attention
Decoder	RTDETRDecoder	Prediksi bounding box dan klasifikasi objek tanpa NMS

2.2.3. Arsitektur Model LSTM

Model LSTM yang digunakan berupa arsitektur stacked dengan dua lapisan LSTM, dilengkapi dua lapisan *Dropout* untuk mengurangi *overfitting*, serta dua lapisan *Dense* sebagai output. LSTM dipilih karena kemampuannya mengatasi *vanishing gradient* pada RNN konvensional melalui mekanisme gerbang *forget gate*, *input gate*, dan *output gate* yang mengatur aliran informasi di *cell state* (Liu, 2024; Sianturi et al., 2023). Skema umum arsitektur LSTM disajikan pada Gambar 3.



Gambar 3: Arsitektur LSTM.

Forget gate pada LSTM berperan memilih informasi dari *cell state* sebelumnya yang harus dipertahankan atau dibuang, seperti dirumuskan pada Persamaan (1).

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (1)$$

Keterangan: f_t merupakan nilai *forget gate*, h_{t-1} adalah *hidden state* sebelumnya, x_t merupakan input pada waktu ke- t , sedangkan W_f dan b_f masing-masing menyatakan bobot dan bias.

Selanjutnya, *input gate* digunakan untuk menentukan informasi baru yang akan dimasukkan ke dalam memori, sebagaimana ditunjukkan pada Persamaan (2).

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2)$$

Kandidat *cell state* baru dihitung menggunakan fungsi aktivasi hiperbolik tangen seperti pada Persamaan (3).

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (3)$$

Nilai *forget gate* dan *input gate* kemudian digunakan untuk memperbarui *cell state* sebagaimana pada Persamaan (4).

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (4)$$

Tahap terakhir adalah menentukan informasi yang akan dikeluarkan sebagai *hidden state*. Proses ini dilakukan melalui *output gate* yang dirumuskan pada Persamaan (5) dimana menghasilkan nilai *output gate* yang akan menyaring informasi dari *cell state*.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

Nilai akhir *hidden state* dihitung dengan Persamaan (6), yang menyatakan bahwa *hidden state* diperoleh dengan menyaring *cell state* melalui *output gate*. Nilai ini selanjutnya diteruskan ke langkah waktu berikutnya maupun ke lapisan jaringan berikutnya.

$$h_t = o_t \odot \tanh(C_t) \quad (6)$$

Dengan mekanisme tersebut, LSTM mampu menyimpan informasi penting dalam jangka panjang sehingga cocok untuk mempelajari pola temporal pada data nutrisi harian MBG. Rincian arsitektur model yang dipakai dalam penelitian ini disajikan pada Tabel 4.

Tabel 4: Arsitektur Model LSTM.

Layer (Type)	Output Shape	Jumlah Parameter
LSTM (64 units)	(None, 7, 64)	19.200
Dropout	(None, 7, 64)	0
LSTM (32 units)	(None, 32)	12.416
Dropout	(None, 32)	0
Dense (32 units)	(None, 32)	1.056
Dense (10 units)	(None, 10)	330
Total	—	33.002

2.2.4. Implementasi Sistem Website Monitoring

Sistem ini dikembangkan sebagai aplikasi web berbasis AstroJS yang terhubung dengan Supabase (PostgreSQL) dan Drizzle ORM untuk menjaga konsistensi pengelolaan data. Struktur basis data relasionalnya mencakup empat entitas utama, yaitu tabel *detections* untuk menyimpan hasil identifikasi objek dan estimasi nutrisi, tabel *foods* sebagai referensi komposisi gizi, tabel *users* untuk autentikasi, serta tabel *daily_menus* sebagai acuan menu standar harian MBG.

3. Hasil dan Pembahasan

3.1 Evaluasi Deteksi Objek Model RT-DETR

Evaluasi model RT-DETR dilakukan untuk menilai kemampuan deteksi dan klasifikasi objek makanan pada citra *foodtray* Program Makan Bergizi Gratis (MBG) menggunakan *confusion matrix* pada data uji. Evaluasi ini digunakan untuk melihat kesesuaian antara hasil prediksi model dan label sebenarnya berdasarkan jumlah prediksi benar maupun kesalahan klasifikasi pada tiap kelas makanan. Confusion matrix terdiri dari empat komponen utama, yaitu *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN). Bentuk umum *confusion matrix* ditunjukkan pada Tabel 5.

Tabel 5: Asumsi Confusion Matrix.

Kelas Prediksi	Kelas Aktual	
	Positive	Negative
<i>Positive</i>	TP	FP
<i>Negative</i>	FN	TN

Berdasarkan hasil evaluasi pada data testing, diperoleh confusion matrix untuk beberapa kelas representatif yang memiliki jumlah data dominan dalam dataset. Hasil confusion matrix lima kelas representatif ditunjukkan pada Tabel 6.

Tabel 6: *Confusion Matrix* 5 Kelas Representatif.

Aktual \ Prediksi	Nasi	Ayam Goreng	Pisang	Tahu	Tempe Goreng
Nasi	89	1	0	0	0
Ayam Goreng	1	21	0	0	0
Pisang	0	0	46	0	0
Tahu	0	0	0	43	2
Tempe Goreng	0	0	0	2	40

Berdasarkan *confusion matrix*, performa model dievaluasi menggunakan metrik akurasi, presisi, dan *recall*. Akurasi digunakan untuk mengukur tingkat ketepatan prediksi model secara keseluruhan, presisi menilai ketepatan hasil prediksi positif, sedangkan *recall* mengukur kemampuan model dalam mendeteksi seluruh objek yang seharusnya dikenali. Perhitungan akurasi dilakukan untuk mengetahui tingkat kesesuaian prediksi model terhadap seluruh data pengujian sebagaimana dirumuskan pada Persamaan (7).

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (7)$$

Evaluasi presisi mengukur seberapa akurat model Ketika menghasilkan prediksi positif dibandingkan dengan label sebenarnya seperti dirumuskan pada Persamaan (8).

$$\text{Presisi} = \frac{TP}{TP + FP} \times 100\% \quad (8)$$

Evaluasi *recall* mengukur kemampuan model untuk menemukan semua objek actual yang seharusnya terdeteksi seperti dirumuskan pada Persamaan (9).

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (9)$$

Model RT-DETR dilatih untuk mengenali 37 kelas komponen makanan MBG. Pada data pengujian, diperoleh total *True Positive* (TP) sebesar 1.175, *False Positive* (FP) sebesar 82, dan *False Negative* (FN) sebesar 82. Dalam konteks deteksi objek, metrik evaluasi yang digunakan meliputi presisi dan recall, yang masing-masing mencapai nilai sebesar 93,48%. Kesamaan nilai kedua metrik tersebut disebabkan oleh jumlah *False Positive* (FP) dan *False Negative* (FN) yang identik pada data pengujian, yaitu masing-masing sebesar 82. Kondisi ini tidak terjadi secara kebetulan, melainkan menunjukkan adanya keseimbangan antara ketepatan prediksi (*precision*) dan kelengkapan deteksi (*recall*) yang dihasilkan oleh model.

Hasil deteksi memperlihatkan bahwa RT-DETR efektif dalam mengenali komponen makanan MBG dengan akurasi tinggi. Sebagian besar kesalahan klasifikasi muncul pada kelas sayuran yang memiliki kemiripan tekstur dan warna, sedangkan kelas yang visualnya lebih khas seperti Pisang mencapai presisi dan recall sempurna yaitu 100%.

3.2 Evaluasi Prediksi Nutrisi Model LSTM

Evaluasi model LSTM dilakukan menggunakan metrik RMSE (*Root Mean Square Error*) dan MAPE (*Mean Absolute Percentage Error*) pada tiga scenario pembagian data untuk menentukan konfigurasi terbaik (Aji dan Idifitriani, 2024; Suryana dan Silaswara, 2024). RMSE mengukur rata-rata selisih antara nilai prediksi dan nilai nyata dalam bentuk kuadrat, sehingga lebih peka terhadap kesalahan besar, nilai RMSE yang lebih kecil menandakan kemampuan prediksi model yang lebih baik. Sedangkan MAPE menilai rata-rata persentase kesalahan prediksi relatif terhadap nilai actual, memberikan ukuran akurasi dalam bentuk persentase semakin kecil nilai MAPE, semakin dekat prediksi dengan nilai sebenarnya.

Evaluasi RMSE menghitung akar dari rata-rata kuadrat selisih antara nilai aktuan dan nilai prediksi seperti dirumuskan pada Persamaan (10)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10)$$

Keterangan: y_i merupakan nilai aktual, $\hat{y}_i$ merupakan nilai prediksi, dan n

Adalah jumlah data pengujian. Selanjutnya, evaluasi menggunakan MAPE bertujuan mengukur rata-rata persentase kesalahan prediksi dibandingkan nilai actual, sebagaimana dirumuskan pada Persamaan (11).

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (11)$$

Keterangan: y_i merupakan nilai aktual, $\hat{y}_i$ merupakan nilai prediksi, dan n adalah jumlah data pengujian. Hasil evaluasi model pada tiga scenario pembagian data ditunjukkan pada Tabel 7.

Tabel 7: Perbandingan Performa LSTM pada Tiga Skenario Pembagian Data.

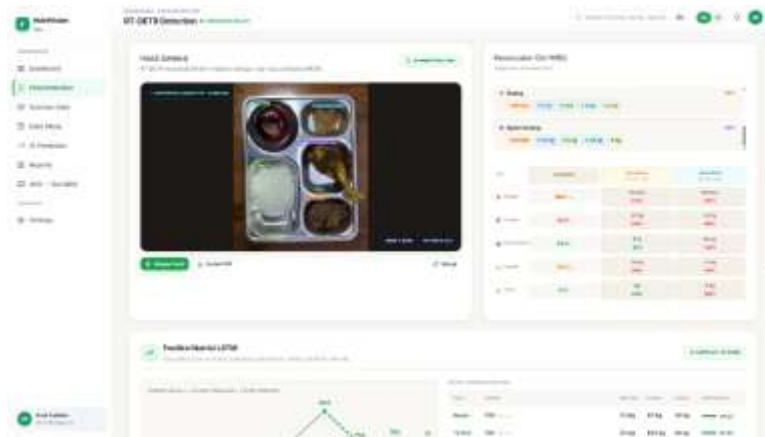
Skenario Split	RMSE	MAPE
90:10	38,42	28,25%
80:20	47,93	38,31%
70:30	53,90	81,78%

Hasil evaluasi mengkonfirmasi bahwa rasio 90:10 menghasilkan performa terbaik dengan RMSE = 38,42 dan MAPE = 28,25%, menunjukkan bahwa proporsi data pelatihan yang lebih besar memungkinkan model mempelajari pola temporal nutrisi secara lebih komprehensif. Penurunan performa yang signifikan pada rasio 70:30 (MAPE = 81,78%) mengindikasikan bahwa model LSTM membutuhkan volume data latih yang memadai untuk menangkap pola deret waktu nutrisi yang bervariasi.

3.3 Implementasi dan Tampilan Sistem Website

Sistem *website monitoring* MBG berhasil dikembangkan dengan halaman utama Halaman Data Nutrisi, dan Halaman Deteksi Menu Makanan. Sistem bekerja dengan menerima unggahan citra *foodtray* dari pengguna melalui *website*, kemudian model RT-DETR melakukan deteksi komponen makanan secara *real-time* dan menghasilkan *bounding box*, label kelas, serta *confidence score*.

Proses deteksi diintegrasikan dengan modul perhitungan nutrisi, di mana setiap objek makanan yang berhasil diidentifikasi dipetakan ke entri pada tabel *foods* menggunakan nama kelas sebagai kunci pencarian. Tabel tersebut berisi informasi kandungan energi (kkal), protein (g), lemak (g), karbohidrat (g), dan serat (g) per 100 gram untuk masing-masing jenis makanan. Mengingat sistem belum mampu memperkirakan berat porsi secara aktual, nilai nutrisi yang digunakan mengacu pada nilai standar per 100 gram untuk setiap objek yang terdeteksi. Total kandungan nutrisi dalam satu nampan diperoleh dengan menjumlahkan seluruh nilai nutrisi dari objek yang terdeteksi dalam satu citra. Implementasi tampilan pada *website monitoring* MBG disajikan pada Gambar 4.



Gambar 4: Halaman Deteksi Menu Makanan pada *Website Monitoring* MBG.

Selain menampilkan hasil estimasi nutrisi secara *real-time*, sistem juga memanfaatkan model LSTM untuk memprediksi kebutuhan nutrisi hari berikutnya berdasarkan pola konsumsi MBG sebelumnya. Hasil prediksi ini digunakan sebagai acuan rekomendasi menu harian agar tetap sesuai standar Angka Kecukupan Gizi (AKG), sehingga *website monitoring* berfungsi sebagai alat pemantauan sekaligus pendukung perencanaan menu berbasis data.

4. Simpulan dan Saran

Penerapan model RT-DETR digunakan untuk tugas deteksi dan estimasi nutrisi makanan dalam Program Makan Bergizi Gratis, yang selanjutnya diintegrasikan ke dalam sistem *monitoring* berbasis *web* menggunakan AstroJS dan Supabase. Hasil evaluasi menunjukkan bahwa model RT-DETR memperoleh nilai presisi dan *recall* keseluruhan masing-masing sebesar 93,48% pada data pengujian. Sementara itu, model LSTM menunjukkan kinerja optimal pada skenario pembagian data 90:10, dengan nilai RMSE sebesar 38,42 dan MAPE sebesar 28,25%.

Sistem yang dikembangkan mampu mendukung proses pemantauan kandungan gizi makanan secara *real-time*. Namun demikian, sistem ini masih memiliki dua keterbatasan utama. Pertama, estimasi nutrisi masih didasarkan pada nilai referensi tetap per 100 gram, sehingga belum mampu merepresentasikan kandungan gizi berdasarkan berat atau ukuran porsi aktual yang diekstraksi dari citra makanan. Kedua, model LSTM masih menunjukkan tingkat kesalahan prediksi yang relatif tinggi (MAPE = 28,25%), sehingga hasil prediksi tren nutrisi harian lebih tepat dimanfaatkan sebagai indikasi umum dibandingkan sebagai acuan kuantitatif yang presisi.

Pengembangan selanjutnya disarankan fokus pada peningkatan estimasi porsi dan berat makanan berbasis citra misalnya dengan metode segmentasi seperti SegFormer agar perhitungan nutrisi lebih akurat sesuai

porsi sebenarnya.

Daftar Pustaka

- Widyasari, S. Y., Larasati, A., & Alam, W. Y. (2025). *Evaluasi Kebijakan Makan Bergizi Gratis di Sekolah Dasar : Implikasi Terhadap Kesehatan Anak dan Pemberdayaan Ekonomi Lokal*. 5, 1727–1736.
- Chai, J., Zeng, H., Li, A., & Ngai, E. W. T. (2021). Machine Learning with Applications Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*, 6(March), 100134. <https://doi.org/10.1016/j.mlwa.2021.100134>
- Sarker, I. H. (2021). Deep Learning: A Comprehensive Overview on Techniques , Taxonomy , Applications and Research Directions. *SN Computer Science*, 2(6), 1–20. <https://doi.org/10.1007/s42979-021-00815-1>
- Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (n.d.). *Image Segmentation Using Deep Learning : A Survey*. 1–22.
- Feng, Y., Wang, Y., Wang, X., Zhao, B., Bi, J., Han, S., Xiao, Z., & Luo, Y. (2026). RGB-D food nutrient estimation supported by FLAVA contrastive learning. *Journal of Food Composition and Analysis*, 150(August 2025), 108821. <https://doi.org/10.1016/j.jfca.2025.108821>
- Revesai, Z. (2025). *Lightweight Interpretable Deep Learning Model for Nutrient Analysis in Mobile Health Applications*. 1, 1–31.
- Ahmad, A., & Tribuana, D. (2024). *Calorie Detection of Traditional Indonesian Food Using the Single Shot Multibox Detector (SSD) Method Deteksi Kalori Makanan Tradisional Indonesia Menggunakan Metode Single Shot Multibox Detector (SSD)*. 4(July), 819–829.
- Ash-shafa, R., & Andono, P. N. (2025). *Pendeteksi visual makanan dan jumlah kalornya menggunakan algoritma mask r-cnn berbasis bot telegram*. 10(1), 641–651.
- Annisa, R., Wintoro, P. B., Pradipta, R. A., Komala, R., Informatika, T., Lampung, U., Dokter, P., Lampung, U., Buatan, K., Makanan, P. C., & Kalori, P. (2025). *Pengenalan Citra Makanan dan Perhitungan Kalori Otomatis Berbasis Convolutional Neural Network*. 7, 2025–2029.
- Zhao, Y., Lv, W., Xu, S., Wei, J., Dang, Q., Liu, Y., Chen, J., & Wang, G. (2023). *DETRs Beat YOLOs on Real-time Object Detection*. 1. <https://doi.org/https://arxiv.org/pdf/2304.08069v1>
- Madan, M., & Reich, C. (2025). Strengthening Small Object Detection in Adapted RT-DETR Through Robust Enhancements. *Electronics*.