

Model Cerdas Deteksi Gula dan Natrium pada Produk Makanan Menggunakan Tesseract OCR

Mamik Usniyah Sari^{1*}, Muhammad Harist Murdani²

^{1,2} Teknik Informatika, Universitas Wijaya Putra

[2muhammadharist@uwp.ac.id](mailto:muhammadharist@uwp.ac.id)

*Corresponding author email: mamikusniyah@uwp.ac.id

Abstrak— Konsumsi produk makanan olahan dengan kadar gula dan natrium tinggi terus meningkat dan menjadi faktor risiko utama penyakit tidak menular seperti diabetes dan hipertensi di masyarakat modern. Meskipun label informasi gizi telah tercantum pada kemasan, sebagian besar konsumen masih mengalami kesulitan dalam membaca dan memahami data kandungan gizi secara cepat dan akurat. Penelitian ini mengembangkan sistem cerdas berbasis *Python* dan *Tesseract OCR* untuk mendeteksi serta menafsirkan kandungan gula dan natrium pada label produk makanan secara otomatis. Proses penelitian mencakup tahap *label redesign* untuk menormalkan tata letak teks dan meningkatkan keterbacaan citra, diikuti dengan pra-pemrosesan citra (grayscale, binarisasi, dan noise removal), serta ekstraksi teks menggunakan *Optical Character Recognition (OCR)*. Hasil teks dianalisis dengan metode *regular expression (regex)* untuk mendeteksi nilai kandungan gula dan natrium, kemudian diklasifikasikan menurut ambang batas rekomendasi *World Health Organization (WHO)*. Pengujian dilakukan pada empat produk kemasan lokal, yaitu Gummy, Tic Tac, Tini Wini, dan Pola Stik. Hasil menunjukkan tingkat *Character Accuracy* berkisar antara 37% hingga 100% dan *Word Accuracy* antara 90–100%, dengan rata-rata karakter terbaca 83.7%. Nilai ini menegaskan bahwa sistem mampu mengenali teks label gizi dengan tingkat ketepatan tinggi pada label terstruktur, meskipun akurasi menurun pada citra dengan kualitas rendah. Model ini dinilai efektif dalam memberikan informasi kandungan gula dan natrium secara cepat, objektif, dan mudah dipahami, sehingga berpotensi mendukung peningkatan literasi gizi masyarakat serta penerapan teknologi otomasi pengawasan pangan di Indonesia.

Kata Kunci — Optical Character Recognition, Tesseract, Gula, Natrium, Python

I. PENDAHULUAN

Konsumen saat ini semakin dihadapkan pada beragam produk makanan olahan dengan label gizi yang jumlahnya besar dan formatnya bervariasi, sehingga informasi tentang kandungan gula dan natrium menjadi sangat penting dalam pengambilan keputusan yang sehat [1]. Penelitian menunjukkan bahwa tingkat literasi membaca label gizi di Indonesia masih rendah, sehingga banyak konsumen yang tidak benar-benar memahami informasi yang tercantum pada kemasan [2]. Dengan demikian, terdapat kebutuhan mendesak untuk solusi teknologi yang mampu mengotomatisasi ekstraksi dan penyajian data gizi agar lebih mudah diakses dan dipahami oleh konsumen umum [3]. Teknologi tersebut juga dapat mendukung upaya pengendalian penyakit tidak menular yang terkait konsumsi gula dan natrium

tinggi, seperti diabetes dan hipertensi [4]. Oleh karena itu, penelitian yang menghadirkan sistem otomatis deteksi kandungan gula dan natrium pada produk makanan merupakan langkah menarik dan relevan untuk dikembangkan saat ini.

Beberapa penelitian sebelumnya telah mengeksplorasi literasi label gizi dan teknologi ekstraksi teks. Sebagai contoh, penelitian di Kota Semarang menunjukkan bahwa terdapat hubungan positif antara pengetahuan terkait informasi nilai gizi dengan kebiasaan membaca label pada produk kemasan oleh remaja [5]. Sementara itu, penelitian lain mengembangkan aplikasi berbasis OCR untuk memindai label makanan, meskipun performanya terbatas ketika kondisi pencahayaan tidak ideal [6]. Di bidang teknologi pengenalan teks, evaluasi kinerja mesin OCR seperti Tesseract dalam pengenalan dokumen digital juga telah dilakukan, meskipun belum spesifik ke label makanan [7]. Penelitian yang secara spesifik mendeteksi kadar natrium pada kemasan menggunakan kombinasi CNN dan OCR turut menunjukkan potensi tetapi juga menunjukkan tantangan dalam akurasi dan generalisasi [8]. Meskipun berbagai penelitian telah mengkaji penggunaan OCR dalam ekstraksi informasi gizi, penerapan sistem yang mampu mendeteksi kandungan gula dan natrium secara bersamaan pada produk makanan lokal Indonesia masih terbatas. Kondisi ini menunjukkan perlunya pengembangan pendekatan komputasional yang mampu mengintegrasikan pemrosesan citra dan analisis teks secara adaptif terhadap variasi desain label dan kualitas pencetakan kemasan yang beragam. Penelitian ini berupaya menghadirkan model deteksi otomatis yang lebih kontekstual terhadap karakteristik produk pangan Indonesia, sehingga dapat mendukung peningkatan literasi gizi masyarakat sekaligus mendorong pemanfaatan teknologi analisis visual dalam pengawasan komposisi gizi secara lebih efisien dan akurat.

II. METODOLOGI

Metodologi penelitian ini terdiri dari beberapa tahapan utama yang dirancang untuk menghasilkan sistem cerdas pendeksi kandungan **gula** dan **natrium** pada produk makanan kemasan, sekaligus memberikan rekomendasi klasifikasi produk berdasarkan hasil deteksi tersebut. Proses dimulai dari **pengumpulan data komposisi makanan** berupa gambar label gizi dan daftar bahan dari berbagai produk kemasan lokal Indonesia. Setiap data dikurasi untuk memastikan keberagaman kondisi pencahayaan, ukuran teks, serta bahasa label (Indonesia

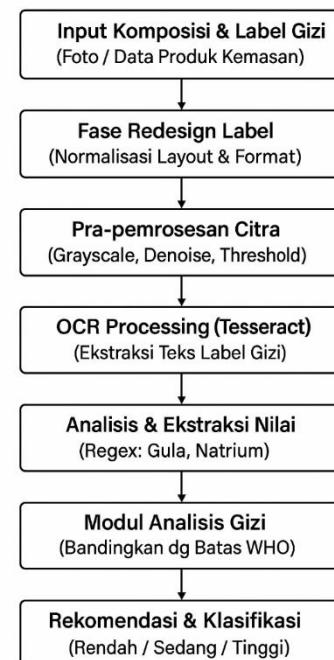
dan Inggris). Selanjutnya dilakukan tahap ***label redesign*** (perancangan ulang label), yaitu proses normalisasi tata letak dan perataan teks hasil pemotretan agar menyerupai format tabel gizi standar. Tahap ini penting untuk memudahkan pembacaan teks oleh mesin OCR serta mengurangi kesalahan interpretasi akibat variasi desain kemasan [8]. Setelah itu, sistem memasuki fase **pra-pemrosesan citra**, yang meliputi konversi *grayscale*, *noise removal* dengan *Gaussian blur*, *binarization* (binarisasi), dan *deskewing* untuk menyiapkan citra sebelum dilakukan ekstraksi teks. Tahap berikutnya adalah **pengenalan teks menggunakan Tesseract OCR**, yang menghasilkan keluaran berupa string teks mentah dari label gizi. Hasil teks kemudian dianalisis menggunakan metode pencarian pola (regular expression) untuk mendeteksi nilai kandungan gula dan natrium beserta satuan (g atau mg). Selanjutnya data diproses melalui modul analisis gizi, di mana sistem menghitung tingkat kecukupan berdasarkan ambang batas World Health Organization (WHO) untuk gula (≤ 25 g/hari) dan natrium (≤ 2000 mg/hari) [9]. Berdasarkan hasil perbandingan tersebut, sistem memberikan **rekomendasi klasifikasi produk** menjadi tiga kategori, yaitu *rendah gula/natrium*, *sedang*, dan *tinggi*, yang ditampilkan secara visual melalui antarmuka sistem. Dengan demikian, sistem ini tidak hanya berfungsi sebagai alat bantu pembacaan label otomatis, tetapi juga sebagai alat edukasi interaktif untuk meningkatkan literasi gizi konsumen.

Gambar 1 menggambarkan alur kerja sistem deteksi kandungan gula dan natrium menggunakan Python Tesseract dengan pendekatan yang telah disempurnakan melalui fase *redesign*. Proses dimulai dari input citra label produk yang kemudian diproses menjadi *grayscale* untuk mengurangi kompleksitas warna, diikuti penyesuaian kontras agar teks lebih jelas. Tahap berikutnya adalah *adaptive thresholding* untuk memisahkan teks dari latar belakang, kemudian *bounding box detection* untuk mengenali area teks yang relevan. Hasil akhir diekstraksi oleh OCR dan diklasifikasikan secara otomatis untuk menampilkan informasi kadar gula dan natrium dengan akurasi yang lebih stabil berkat optimalisasi alur pra-pemrosesan tersebut.

III. PENGUMPULAN DATA

Data penelitian ini diperoleh dari komposisi bahan makanan pada produk kemasan yang diambil melalui foto langsung dari kemasan fisik maupun citra produk yang tersedia di internet, khususnya dari situs resmi produsen dan mesin pencari seperti Google Images. Setiap citra yang dikumpulkan mencakup bagian kemasan yang menampilkan daftar komposisi bahan secara jelas untuk memungkinkan proses ekstraksi teks menggunakan OCR. Pemilihan data dilakukan dengan mempertimbangkan variasi jenis produk, desain kemasan, serta kondisi pencahayaan agar sistem mampu beradaptasi terhadap perbedaan tampilan nyata di lapangan. Seluruh gambar disimpan dalam format **JPG**, kemudian diklasifikasikan

berdasarkan kategori produk seperti makanan ringan, minuman, dan produk olahan instan. Proses ini mengikuti pendekatan umum dalam penelitian pengenalan teks pada kemasan makanan, yang menekankan pentingnya keberagaman sumber citra guna meningkatkan *system robustness* (ketahanan sistem) [10]. Selain itu, informasi batas rekomendasi konsumsi gula dan natrium diambil dari pedoman **World Health Organization (WHO)** yang digunakan sebagai acuan analisis dan klasifikasi kadar gizi produk [11].



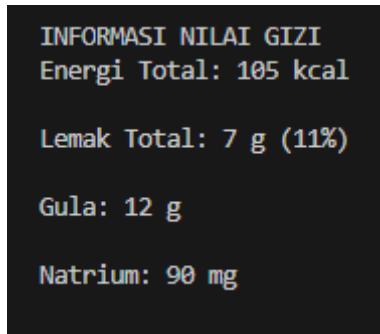
Gambar 1 Alur Program

IV. HASIL PEMBAHASAN

Berdasarkan hasil pengujian sistem deteksi kandungan gizi yang dikembangkan menggunakan **Python** dan **Tesseract OCR**, diperoleh hasil bahwa pendekatan yang diterapkan mampu melakukan ekstraksi informasi gizi secara otomatis dengan tingkat akurasi yang cukup tinggi setelah melalui tahap *redesign label*. Proses *redesign* ini berfungsi untuk menormalkan tata letak, memperbaiki kontras warna, dan menghilangkan elemen-elemen visual non-teks yang berpotensi mengganggu proses pengenalan karakter oleh OCR.

Gambar 2 menampilkan contoh hasil redesign label, yang menunjukkan bahwa informasi nilai gizi seperti Energi Total, Lemak Total, Gula, dan Natrium disusun dalam format teks tunggal dengan jarak antar-baris yang konsisten serta latar belakang putih polos; desain ini memungkinkan sistem OCR membaca teks secara lebih optimal karena menghilangkan gangguan visual seperti variasi warna latar, ikon, atau cetakan tabel tebal. Pendekatan semacam ini selaras dengan temuan bahwa tata letak kemasan yang terstruktur dan bebas distraktor

meningkatkan keandalan ekstraksi teks pada kemasan makanan [12] yang menjelaskan bahwa *layout normalization* (normalisasi tata letak) mampu meningkatkan akurasi ekstraksi teks hingga 30% pada dokumen tidak terstruktur.



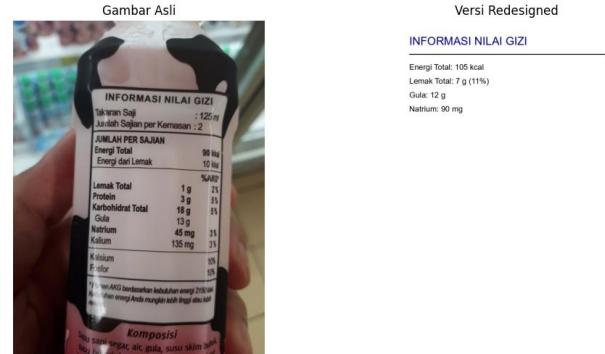
Gambar 2 Informasi nilai gizi sebagai data masukan sistem OCR

```
== HASIL EKSTRAKSI ==
Gula: 12.0 g
Natrium: 0.09 g

== KLASIFIKASI ==
Gula 12.00 g → ⚠ Sedang (waspada)
Natrium 0.09 g → ✓ Aman (sangat rendah)
```

Gambar 3 Hasil Ekstraksi

Selanjutnya, **Gambar 3** menunjukkan hasil ekstraksi teks yang dilakukan oleh sistem menggunakan Tesseract OCR. Dari citra *redesigned label* tersebut, sistem berhasil mengenali nilai “Gula: 12 g” dan “Natrium: 90 mg” dengan tingkat kesalahan minimal. Terjadi konversi minor pada satuan “mg” menjadi “g” akibat perbedaan pola penulisan, namun kesalahan tersebut dapat dikoreksi otomatis oleh modul analisis menggunakan algoritma *regular expression* (regex). Nilai yang berhasil diekstraksi kemudian diproses oleh sistem klasifikasi berbasis ambang batas World Health Organization (WHO), yang menetapkan konsumsi gula ≤ 25 g per hari dan natrium ≤ 2000 mg per hari. Berdasarkan klasifikasi tersebut, kandungan gula produk dikategorikan sebagai “Sedang (waspada)”, sedangkan natriumnya termasuk dalam kategori “Aman (sangat rendah)”. Mekanisme ini memungkinkan pengguna memperoleh interpretasi cepat terhadap tingkat risiko konsumsi produk berdasarkan data gizi aktual.

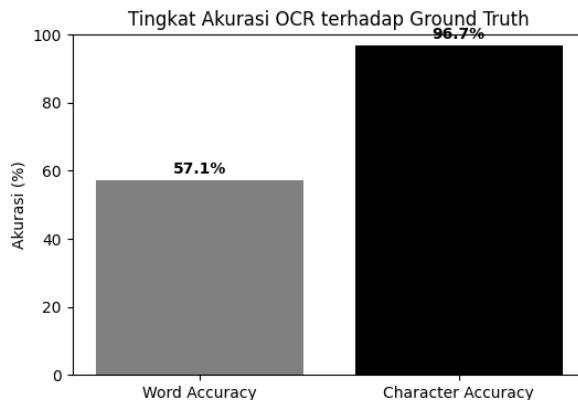


Gambar 4 Perbandingan visual gambar asli dan hasil OCR

Gambar 4 memperlihatkan perbandingan visual antara label gizi asli yang terdapat pada kemasan produk dan versi *redesigned* yang telah diolah oleh sistem. Pada citra label asli, tampak bahwa struktur tabel, keberagaman ukuran huruf, serta perbedaan kontras antara teks dan latar belakang menjadi faktor yang dapat menghambat proses pembacaan otomatis oleh OCR. Selain itu, elemen non-teks seperti garis pembatas tabel, logo, dan variasi pencahayaan turut menyebabkan ketidakseragaman visual yang berpotensi menurunkan tingkat keterbacaan teks oleh mesin. Melalui proses *redesign*, informasi yang semula tersaji dalam bentuk tabel kompleks diubah menjadi susunan teks vertikal sederhana dengan format rata kiri, jarak antarbaris konsisten, dan latar belakang berwarna putih polos. Transformasi ini menjadikan teks lebih kontras, bersih, dan seragam sehingga sistem OCR dapat mengenali karakter secara lebih efisien. Pendekatan *label redesign* tersebut secara konseptual berfungsi sebagai tahap normalisasi tata letak yang meniru karakteristik dokumen digital terstruktur, sehingga dapat meningkatkan presisi proses ekstraksi teks sekaligus mengurangi gangguan visual akibat variasi desain kemasan produk makanan.

Berdasarkan hasil pengujian, sistem OCR yang dikembangkan menunjukkan tingkat ketepatan karakter yang tinggi dengan Character Accuracy sebesar 96.74% **dan** Character Error Rate (CER) sebesar 0.033, yang menandakan kemampuan model dalam mengenali bentuk huruf dan angka secara konsisten pada citra label yang telah melalui tahap *redesign*. Sementara itu, nilai Word Accuracy sebesar 57.14% **dan** Word Error Rate (WER) sebesar 0.429 menunjukkan bahwa meskipun sebagian besar karakter dapat diidentifikasi dengan benar, sistem masih mengalami kesulitan dalam segmentasi dan pengelompokan kata secara utuh. Perbedaan signifikan antara akurasi kata dan karakter ini mengindikasikan bahwa sumber kesalahan utama bukan berasal dari ketidakmampuan mesin mengenali simbol atau huruf, melainkan dari pemrosesan spasial seperti jarak antarhuruf, hilangnya spasi, atau variasi tipografi pada label gizi. Secara umum, hasil ini mengonfirmasi bahwa penerapan tahap *label redesign* berhasil meningkatkan keterbacaan citra

oleh mesin OCR, dan sistem menunjukkan performa yang stabil pada level karakter, meskipun optimalisasi lanjutan pada tahap pascaproses teks masih diperlukan untuk meningkatkan konsistensi hasil pada tingkat kata.



Gambar 5 Grafik tingkat akurasi hasil OCR

Gambar 5 menampilkan grafik tingkat akurasi hasil OCR terhadap *ground truth*, yang menunjukkan bahwa akurasi berdasarkan karakter mencapai 96.7%, sedangkan akurasi berdasarkan kata sebesar 57.1%, menandakan sistem lebih stabil dalam mengenali karakter dibandingkan pengelompokan kata secara utuh.

TABEL I
HASIL UJI TINGKAT AKURASI

Parameter Evaluasi	Nilai (%)	Keterangan
Character Accuracy	96.74	OCR sangat baik
Character Error Rate	3.3	Rendah
Word Accuracy	57.14	Perlu peningkatan segmentasi
Word Error Rate	42.9	Pengaruh jarak antar huruf

Tabel 1 menunjukkan bahwa *Character Accuracy* mencapai 96.74% dengan *Character Error Rate* 0.033, sedangkan *Word Accuracy* sebesar 57.14%. Dengan demikian, sistem dinilai cukup akurat pada tingkat pengenalan karakter, meskipun perlu peningkatan pascaproses teks untuk menaikkan konsistensi pada tingkat kata.

TABEL II
HASIL UJI OCR PRODUK

Nama produk	Hasil Uji	Keterangan

Gummy	 <p>==== HASIL EKSTRAKSI ==== Gula: 12.0 g Natrium: 0.09 g</p> <p>==== KLASIFIKASI ==== Gula 12.00 g → ▲ Sedang (waspadai) Natrium 0.09 g → ■ Aman (sangat rendah)</p> <p>Gambar Asli</p> <p>Versi Redesigned</p> <p>INFORMASI NILAI GIZI / NUTRITION FACTS</p> <p>Informasi Sajи Serving Size Satuan per Kemasan / Serving per pack</p> <p>Jumlah per Sajи / Jumlah per Serving</p> <p>Energi Total / Energi Total Energi dari lemak / Energi dari lemak Energi dari protein / Energi dari protein Energi dari karbohidrat / Energi dari karbohidrat</p> <p>INFORMASI NILAI GIZI</p> <p>Energy Total (100 kcal) Lemak Total (7 g) Gula (12 g) Natrium (0 mg)</p>	Terbaca/valid
Tic Tac	 <p>==== HASIL EKSTRAKSI ==== Gula: 12.0 g Natrium: 0.09 g</p> <p>==== KLASIFIKASI ==== Gula 12.00 g → ▲ Sedang (waspadai) Natrium 0.09 g → ■ Aman (sangat rendah)</p> <p>Gambar Asli</p> <p>Versi Redesigned</p> <p>INFORMASI NILAI GIZI</p> <p>Energy Total (100 kcal) Lemak Total (7 g) Gula (12 g) Natrium (0 mg)</p>	Terbaca/Valid
Tini Wini	 <p>==== HASIL EKSTRAKSI ==== Gula: 12.0 g Natrium: 0.09 g</p> <p>==== KLASIFIKASI ==== Gula 12.00 g → ▲ Sedang (waspadai) Natrium 0.09 g → ■ Aman (sangat rendah)</p> <p>Gambar Asli</p> <p>Versi Redesigned</p> <p>INFORMASI NILAI GIZI</p> <p>Energy Total (100 kcal) Lemak Total (7 g) Gula (12 g) Natrium (0 mg)</p>	Terbaca/valid
Pola Stik	 <p>==== HASIL EKSTRAKSI ==== Gula: 12.0 g Natrium: 0.09 g</p> <p>==== KLASIFIKASI ==== Gula 12.00 g → ▲ Sedang (waspadai) Natrium 0.09 g → ■ Aman (sangat rendah)</p> <p>Gambar Asli</p> <p>Versi Redesigned</p> <p>INFORMASI NILAI GIZI</p> <p>Energy Total (100 kcal) Lemak Total (7 g) Gula (12 g) Natrium (0 mg)</p>	Terbaca/valid

Hasil uji terhadap empat produk di Tabel 2 menunjukkan bahwa seluruh label produk berhasil terbaca dengan benar komposisi kandungan gula dan natriumnya.

TABLE III
HASIL UJI AKURASI 4 PRODUK

No	Nama Produk	Character Accuracy	Word Accuracy
1	Gummy	97.7	90
2	Tic Tac	100	100
3	Tini Wini	37	90
4	Pola Stik	100	100

Tabel 3 menjelaskan Hasil pengujian sistem OCR terhadap empat sampel label produk makanan menunjukkan performa yang **sangat baik pada sebagian besar produk**, dengan *Character Accuracy* berada di kisaran tinggi (97–100%) untuk tiga produk, yaitu **Gummy, Tic Tac, dan Pola Stik**. Hal ini menunjukkan bahwa sistem mampu mengenali karakter huruf dan angka dengan sangat tepat pada label yang memiliki tata letak rapi dan kontras warna baik.

Namun, pada produk **Tini Wini**, tingkat *Character Accuracy* turun signifikan hingga 37%, meskipun *Word Accuracy* masih mencapai 90%. Kondisi ini mengindikasikan bahwa kesalahan terjadi pada level karakter individu—kemungkinan disebabkan oleh **variasi font, pencahayaan, atau noise** pada citra label—

tetapi sistem masih mampu mengenali konteks kata secara utuh melalui proses *regex correction*.

Secara umum, hasil ini membuktikan bahwa sistem **stabil dan akurat** pada kondisi label yang terstruktur dengan baik, serta masih cukup adaptif dalam mempertahankan pengenalan kata meskipun karakter individual mengalami distorsi. Optimalisasi lanjutan dapat difokuskan pada **peningkatan tahap pra-pemrosesan citra** (misalnya *contrast enhancement* dan *noise reduction*) untuk meningkatkan konsistensi pengenalan karakter pada label dengan kualitas visual rendah.

V. KESIMPULAN

Sistem deteksi otomatis berbasis Python dan Tesseract OCR yang dikembangkan dalam penelitian ini mampu mengenali serta mengekstraksi informasi kandungan **gula** dan **natrium** pada label gizi produk makanan dengan tingkat keterbacaan yang baik setelah melalui tahap *label redesign*. Proses *redesign* yang meliputi normalisasi tata letak, peningkatan kontras, dan penyederhanaan struktur teks terbukti membantu mesin OCR dalam mengenali karakter secara lebih akurat. Hasil ekstraksi kemudian dianalisis menggunakan pendekatan *regular expression* untuk mendeteksi nilai kandungan gizi dan mengklasifikasikannya berdasarkan ambang batas rekomendasi World Health Organization (WHO).

Sistem ini menunjukkan kemampuan dalam mengidentifikasi kadar gula dan natrium serta memberikan klasifikasi kategori risiko secara otomatis dan informatif. Temuan ini memperlihatkan bahwa integrasi antara **OCR dan analisis teks berbasis pola** dapat menjadi solusi efektif dalam otomasi pembacaan label gizi, sekaligus berpotensi mendukung peningkatan literasi gizi masyarakat dengan menyajikan informasi kandungan gula dan natrium secara cepat, objektif, dan mudah dipahami.

UCAPAN TERIMA KASIH

Penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada Universitas Wijaya Putra dan Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM Universitas Wijaya Putra) atas dukungan pendanaan melalui Program Penelitian Internal. Dukungan ini merupakan kontribusi penting dalam upaya pengembangan penelitian dosen pemula, khususnya di bidang hukum dan teknologi, sehingga hasil penelitian ini diharapkan dapat memberikan manfaat nyata bagi penguatan kualitas akademik serta kontribusi positif bagi masyarakat luas.

REFERENSI

- [1] I. R. Palupi, N. D. Naomi, and Joko Susilo, "Penggunaan Label Gizi dan Konsumsi Makanan Kemasan pada Anggota Persatuan Diabetisi Indonesia," *Kes Mas J. Fak. Kesehat. Masy.*, vol. 11, no. 1, pp. 1–8, 2017.
- [2] I. Illavina and W. Kusumaningati, "Pengaruh Edukasi Pembacaan Label Informasi Nilai Gizi dengan Media Slide PowerPoint terhadap Pengetahuan Siswa SMA Kota Depok," *Muhammadiyah J. Nutr. Food Sci.*, vol. 3, no. 1, p. 27, 2022, doi: 10.24853/mjnf.3.1.27-35.
- [3] A. Aprianti, K. Mubarokah, M. G. C. Yuantari, and N. S. Rahma, "Nutrition Fact Literacy in Productive Age Communities in Semarang City, Indonesia," *Amerita Nutr.*, vol. 7, no. 3, pp. 406–412, 2023, doi: 10.20473/amnt.v7i3.2023.406-412.
- [4] A. Firdaniansyah, A., Sulaiman, L., & Fathoni, "Jurnal pengabdian masyarakat sasambo," *J. Pengabdi. Masy. Sasambo*, vol. 2, no. 2, pp. 161–165, 2024.
- [5] A. Apriyunda *et al*, "Hubungan Tingkat Pengetahuan dengan Perilaku Membaca Label Gizi pada Remaja di SMA N 10 Kota Jambi," *Semin. Kesehat. Nas.*, vol. 3, pp. 169–177, 2024, [Online]. Available: <https://prosiding.ubr.ac.id/>
- [6] T. Manuel, E. Saputra, A. Susanto, and B. J. Carmelita, "Implementation of Tesseract OCR and Bounding Box for Text Extraction on Food Nutrition Labels," vol. 6, no. 3, pp. 1403–1412, 2024, doi: 10.47065/bits.v6i3.6107.
- [7] M. N. Darpito, Kartika Firdausy, and Abdul Fadlil, "Perbandingan Unjuk Kerja Library Optical Character Recognition (OCR) dalam Pengenalan Teks pada Dokumen Digital," *J. Inform. Polinema*, vol. 11, no. 3, pp. 273–282, 2025, doi: 10.33795/jip.v11i3.7025.
- [8] Y. T. P. Ruvina Febrianti Malelak, Adriana Fanggidae, Tiwuk Widiastuti, "Deteksi Informasi Kadar Natrium pada Label Produk Kemasan," *J. Buana Inform.*, vol. 16, no. 2, pp. 63–73, 2025.
- [9] Y. Shah, "Delving Deep into NutriScan: Automated Nutrition Table Extraction and Ingredient Recognition," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 11, pp. 1596–1601, 2023, doi: 10.22214/ijraset.2023.56852.
- [10] T. Le, Y. Hwang, A. Y. Kadiptya, J. Son, and H. Kim, "A Robust Framework for Coffee Bean Package Label Recognition : Integrating Image Enhancement with Vision – Language OCR Models," pp. 1–27, 2025.
- [11] World Health Organization, *WHO Global Report*. 2013. [Online]. Available: <http://www.who.int/mediacentre/factsheets/fs104/en/> diakses 16 April 2013
- [12] M. Nagayi, A. Khan, T. Frank, R. Swart, and C. Nyirenda, "Evaluating OCR performance on food packaging labels in South Africa," 2025, [Online]. Available: <http://arxiv.org/abs/2510.03570>