

Deteksi Informasi Kadar Natrium pada Label Produk Kemasan Menggunakan FPN, Faster R-CNN, dan OCR

Ruvina Febrianti Malelak^{*1}, Adriana Fanggidae², Tiwuk Widiastuti³, Yulianto Triwahyuadi Polly⁴

^{1,2,3,4} Program Studi Ilmu Komputer, Fakultas Sains Dan Teknik, Universitas Nusa Cendana
Jl. Adisucipto, Penfui, Kupang, 850001, Nusa Tenggara Timur, Indonesia
Email: ¹ruvinamalelak@gmail.com, ²adrianafanggidae@staf.undana.ac.id,
³tiwukwidiastuti@stafundana.ac.id, ⁴yuliantopolly@staf.undana.ac.id

Abstract. *Detection of Sodium Content Information on Packaged Product Labels Using FPN, Faster R-CNN, and OCR.* Nutrition facts, especially sodium content, play a key role in promoting healthier food choices. However, public awareness remains low, and nutrition labels are often overlooked. This study proposes an automated system to detect and extract sodium information from product packaging images. The system applies an object detection approach using a Feature Pyramid Network (FPN) and Faster R-CNN, implemented via Detectron2. Sodium text extraction is then performed with Tesseract OCR. A dataset of 3,028 images was used for training, validation, and testing. The training phase achieved detection accuracy of 98.90% (AP50) and 85.00% (AR). Testing on 400 images produced an average text extraction accuracy of 85.13%. Recognition errors were mainly caused by poor image quality, such as blurriness, tilted labels, reflections, or overly bold text. Despite these limitations, the system demonstrates strong potential as an automated tool for reading nutrition labels, particularly sodium content, with further refinement needed for real-time use.

Keywords: sodium, Feature Pyramid Network (FPN), Convolutional Neural Network (CNN), Optical Character Recognition (OCR), nutrition facts.

Abstrak. Informasi nilai gizi, khususnya kandungan natrium, berperan penting dalam mendorong masyarakat memilih makanan yang lebih sehat. Namun, kesadaran terhadap label gizi masih rendah sehingga sering diabaikan. Sistem otomatis dikembangkan untuk mendeteksi dan mengekstraksi informasi kadar natrium dari citra kemasan produk. Sistem menggunakan pendekatan deteksi objek berbasis *Feature Pyramid Network* (FPN) dan *Faster R-CNN* melalui *framework Detectron2*. Ekstraksi teks kadar natrium dilakukan menggunakan *Tesseract OCR*. Sebanyak 3.028 gambar digunakan dalam proses pelatihan, validasi, dan pengujian FPN. Pada tahap pelatihan diperoleh akurasi deteksi 98,90% (AP50) dan (AR) 85,00%. Sementara itu, pengujian OCR terhadap 400 gambar berbeda dari data pelatihan FPN, menghasilkan rata-rata akurasi ekstraksi teks 85,13%. Kesalahan pengenalan umumnya disebabkan kualitas citra rendah, seperti gambar blur, orientasi label miring, pantulan cahaya, atau teks terlalu tebal. Meskipun terdapat keterbatasan, sistem menunjukkan potensi kuat sebagai alat otomatis membaca informasi gizi, khususnya kandungan natrium, dengan catatan diperlukan penyempurnaan lebih lanjut agar konsisten di kondisi nyata.

Kata Kunci: Natrium, *Feature Pyramid Network* (FPN), *Convolutional Neural Network* (CNN), *Optical Character Recognition* (OCR), informasi gizi.

1. Pendahuluan

Produk pangan dalam kemasan telah menjadi bagian penting dalam kehidupan sehari-hari karena kepraktisannya. Berbagai jenis makanan instan, camilan, minuman kemasan, dan produk olahan lainnya sering dikonsumsi, terutama produk yang mengandung kadar natrium. Natrium merupakan nutrisi esensial yang berperan dalam menjaga keseimbangan cairan tubuh, fungsi saraf, dan aktivitas sel. Data dari *World Health Organization* (WHO) menunjukkan bahwa asupan natrium harian yang direkomendasikan adalah maksimal 2.000 mg atau setara dengan kurang dari lima gram garam [1]. Riset Kesehatan Dasar (Riskesdas) 2023 mencatat 53,5% penduduk Indonesia mengonsumsi natrium melebihi batas aman, sementara WHO menyatakan

bahwa 70-80% asupan natrium berasal dari makanan kemasan seperti mi instan, makanan kaleng, dan camilan asin. Akibat dari konsumsi natrium berlebih sangat serius bagi kesehatan. Studi dari Kementerian Kesehatan Indonesia menunjukkan bahwa hipertensi atau tekanan darah tinggi menjadi penyebab utama penyakit jantung dan stroke, di mana konsumsi natrium yang tinggi merupakan salah satu pemicunya. Di Indonesia, hipertensi telah menjadi penyebab utama kematian, dengan angka kejadian mencapai 34,1% berdasarkan laporan Riskesdas 2018 [2].

Permasalahan utama adalah rendahnya kepedulian dan cenderung mengabaikan label informasi gizi pada kemasan makanan. Hal ini bukan semata karena tidak mengetahui pentingnya, tetapi lebih karena merasa informasi tersebut sulit dipahami, membosankan, atau tidak langsung berdampak. Padahal, kebiasaan mengonsumsi produk tinggi natrium tanpa disadari dapat menimbulkan risiko kesehatan dalam jangka panjang, untuk meningkatkan kepedulian sejak dini, diperlukan solusi yang mempermudah akses dan pemahaman informasi gizi. Teknologi pengolahan citra berbasis *Artificial Intelligent* (AI) dapat menjadi jawaban, karena mampu mengekstraksi informasi kadar natrium secara otomatis dari label kemasan. Sistem ini dapat mendorong masyarakat untuk lebih aktif memilih produk yang mendukung pola makan sehat. Model seperti *Feature Pyramid Network* (FPN) digunakan untuk mendeteksi area label pada kemasan, sedangkan *Optical Character Recognition* (OCR) berfungsi mengekstrak teks dari label tersebut. Pendekatan ini memungkinkan penyajian informasi kadar natrium secara lebih jelas, sehingga konsumen dapat memahami kandungan produk dengan lebih baik.

FPN adalah arsitektur jaringan konvolusi yang meningkatkan deteksi objek dengan memanfaatkan fitur multi-skala dari berbagai lapisan. Arsitektur ini menggunakan jalur bawah ke atas untuk mengekstraksi fitur dari lapisan dengan resolusi rendah hingga tinggi, serta jalur atas ke bawah untuk memperkaya fitur dengan informasi dari lapisan yang lebih tinggi. Kombinasi kedua jalur ini memungkinkan FPN mendeteksi objek dengan berbagai ukuran secara lebih akurat, termasuk objek kecil dengan detail halus tanpa kehilangan konteks global [3]. FPN dikonfigurasi bersama algoritma *Faster R-CNN* untuk mendukung proses deteksi objek. *Faster R-CNN* menggunakan *Region Proposal Network* (RPN) untuk menghasilkan kandidat wilayah objek, yang kemudian dievaluasi dalam satu jaringan *end-to-end*, sehingga meningkatkan kecepatan dan akurasi deteksi [4]. Penelitian [5] menunjukkan bahwa penggunaan *Multi-Level Residual Feature Pyramid Network* (MR-FPN) mampu memperkuat fitur *multi-level* dalam menangkap detail teks dan mencapai akurasi sebesar 80,92% pada *dataset* ICDAR 2015, lebih baik dibandingkan metode konvensional. Penelitian [6] mengusulkan metode deteksi kereta api menggunakan *Faster R-CNN* dengan arsitektur VGG16 untuk mengidentifikasi lokomotif dan gerbong, yang menunjukkan akurasi deteksi sebesar 79,09% untuk lokomotif dan 97,05% untuk gerbong, dengan rata-rata akurasi keseluruhan masing-masing 86,40% dan 97,23%. OCR mengenali karakter huruf atau angka dalam citra dan mengonversinya ke dalam format teks yang dapat diproses oleh komputer, sehingga memungkinkan pengolahan informasi teks dalam gambar secara lebih cepat dan efisien [7]. Penelitian yang terkait dengan OCR antara lain, penelitian [8] menerapkan *Tesseract* OCR untuk mengekstrak informasi nutrisi dari label dengan *Word Error Rate* (WER) maksimum sebesar 10% dan *Character Error Rate* (CER) sebesar 1,6%. Sementara itu, penelitian [9] menggunakan *Tesseract* untuk mendeteksi variasi teks pada media pembelajaran dengan rata-rata akurasi keseluruhan sebesar 91,11%.

Dalam penelitian ini, digunakan kombinasi FPN dan algoritma *Faster R-CNN* untuk mendeteksi area teks pada label produk kemasan secara akurat. Setelah area teks berhasil terdeteksi, proses ekstraksi informasi kadar natrium dilakukan menggunakan *Tesseract* OCR untuk mengenali karakter dalam citra dan mengubahnya menjadi teks digital. Pendekatan ini memungkinkan sistem untuk secara otomatis mengidentifikasi dan membaca informasi kadar natrium pada label kemasan dengan akurasi tinggi.

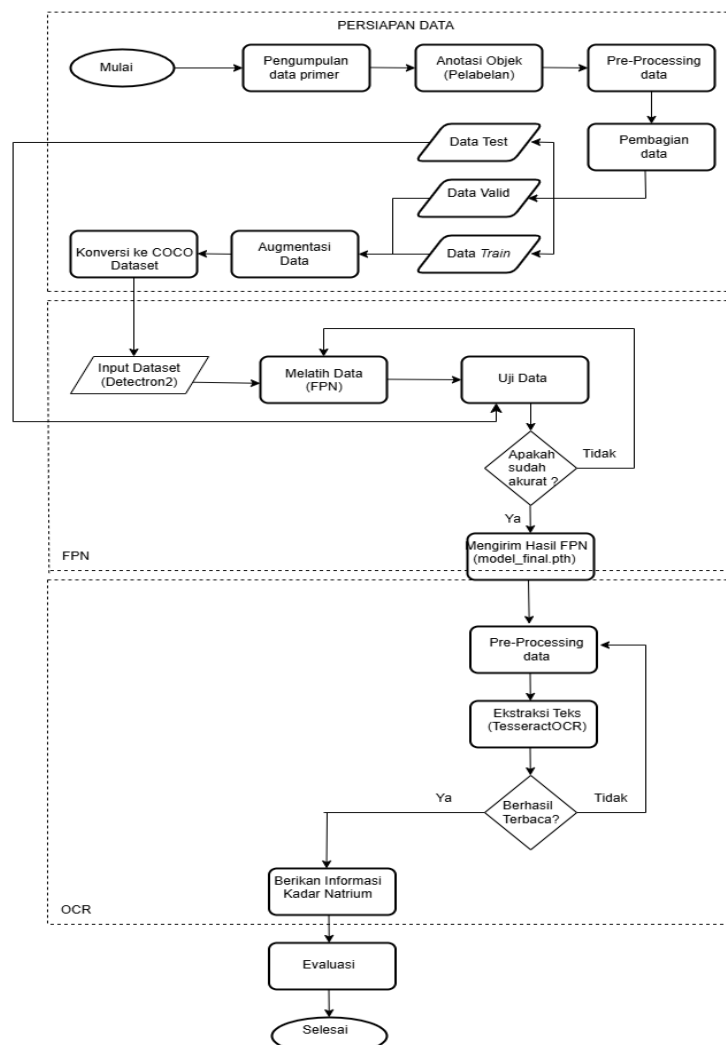
2. Tinjauan Pustaka

Penelitian ini merujuk pada beberapa studi sebelumnya. Penelitian [10] menggunakan *Faster R-CNN* untuk mengenali spesies burung *lovebird* berdasarkan motif atau warna dengan akurasi tinggi. Dengan *dataset* sebanyak 808 gambar dari delapan kelas, model mencapai akurasi

sebesar 96% dan nilai *loss* sebesar 0,2%. Penelitian [11] mengembangkan *Attention Feature Pyramid Network* (AFPN) yang terintegrasi dengan *receptive field* untuk meningkatkan akurasi deteksi, khususnya pada objek kecil. Model dilatih menggunakan *framework* PyTorch dengan *backbone* ResNet-AFPN dan optimasi *Stochastic Gradient Descent* (SGD), menghasilkan mAP sebesar 83,5% pada *dataset* PASCAL VOC (20 kategori) dan 40,6% pada MS COCO (80 kategori). Penelitian [12] melakukan perbandingan kinerja beberapa algoritma deteksi objek, yaitu *Faster R-CNN*, *RetinaNet*, dan *Single Shot Detector* (SSD) dengan berbagai *backbone* ResNet pada pendeteksian kapal laut. Hasil penelitian menunjukkan bahwa *Faster R-CNN* dengan *backbone* ResNet-50 memberikan performa terbaik dengan nilai F1-score sebesar 0,85, melampaui metode lain dalam mendeteksi objek kecil pada citra satelit. Sementara itu, penelitian [13] mengembangkan sistem OCR berbasis *Template Matching Correlation* untuk mengenali karakter dari dokumen cetak yang dikonversi ke format digital. Proses dilakukan melalui tahapan pra-pemrosesan seperti binerisasi, segmentasi, dan normalisasi, dan menghasilkan akurasi pengenalan sebesar 92,90%, menunjukkan efektivitas metode tersebut dalam membangun sistem OCR dengan performa baik.

3. Metodologi Penelitian

Gambar 1 menggambarkan alur kerja sistem yang dirancang dalam penelitian ini, yang mencakup proses deteksi area teks menggunakan FPN dan *Faster R-CNN*, serta ekstraksi informasi kadar natrium menggunakan *Tesseract OCR*.



Gambar 1. Alur Kerja Sistem

3.1 Persiapan Data

Tahap ini diawali dengan proses persiapan data. Karena penelitian ini dengan dua jalur maka, *dataset*nya terpisah. Sebanyak 3.028 data primer berupa citra label produk dikumpulkan dari berbagai supermarket di Kota Kupang untuk digunakan dalam pengujian FPN, sedangkan 400 data terpisah digunakan untuk pengujian OCR, contoh data produk dapat dilihat pada Gambar 2. Proses pengambilan gambar dilakukan menggunakan kamera *smartphone* dengan resolusi 12 MP hingga 50 MP. Jarak pengambilan berkisar antara 10–40 cm dengan posisi kamera tegak lurus terhadap label untuk meminimalisasi distorsi. Label produk difoto dalam kondisi kemasan berdiri tegak di rak maupun saat dipegang, sehingga area informasi gizi dapat terlihat jelas.

Dataset yang terkumpul kemudian dibagi menjadi tiga subset menggunakan *Roboflow*, yaitu 2.622 gambar untuk pelatihan, 219 gambar untuk validasi, dan 187 gambar untuk pengujian. Jumlah data pelatihan dibuat lebih besar agar model memiliki keragaman sampel yang cukup setelah proses augmentasi (rotasi, *shear*, konversi *grayscale*, *blur*, *noise*, dan *auto-orient*), sehingga dapat mencegah *overfitting* dan meningkatkan generalisasi model. [14]. *Grayscale* juga dilakukan untuk menyederhanakan kompleksitas data yang memiliki representasi 8-bit [15]. Proporsi pembagian ini mengikuti praktik umum *machine learning*, yaitu sekitar 80% untuk pelatihan, 10% untuk validasi, dan 10% untuk pengujian. Perbedaan jumlah rasio validasi dan pengujian disebabkan oleh proses pembagian otomatis *Roboflow* yang melakukan pembulatan pada jumlah data.

Proses anotasi dilakukan juga pada *Roboflow* dengan memberikan label “Informasi gizi” pada area yang memuat informasi nutrisi produk. Selanjutnya, dilakukan praproses data berupa *resizing* ke ukuran 640×640 piksel, perataan orientasi gambar (*auto-orient*). Seluruh *dataset* kemudian dikonversi ke format *COCO JSON* agar kompatibel dengan arsitektur FPN yang digunakan dalam proses pelatihan model.



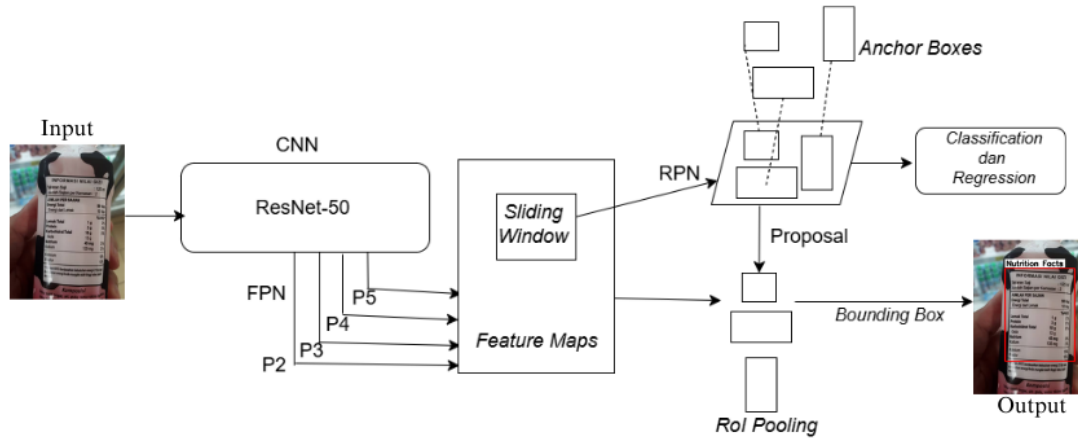
Gambar 2. Contoh produk Kemasan

3.2 Feature Pyramid Network (FPN) dengan Algoritma Faster R-CNN

Arsitektur FPN dengan *Faster R-CNN* dapat dilihat pada Gambar 3. Proses dimulai dengan menginput gambar produk kemasan yang telah melalui *Pre-Processing*. Gambar ini diproses oleh ekstraksi fitur *Backbone CNN*, yaitu ResNet-50, yang berfungsi mengekstraksi fitur penting dari citra. ResNet-50 menggunakan *residual connections* untuk mengatasi masalah *vanishing gradient* pada jaringan dalam melalui penggunaan *residual connections*, dengan ResNet-50 sebagai versi yang paling umum digunakan [16], menghasilkan *feature maps* pada beberapa tahap konvolusi, yaitu Conv2, Conv3, Conv4, dan Conv5. *Feature maps* ini kemudian menjadi input bagi *Feature Pyramid Network* (FPN).

FPN kemudian menggabungkan informasi ini melalui mekanisme *bottom-up* dan *top-down pathway* dengan *lateral connection*, sehingga diperoleh representasi *multi-level* dari tiap level yang lebih kuat [17], berupa *feature maps multi-level* yaitu P2, P3, P4, dan P5. Setiap level

membawa informasi dengan kedalaman berbeda, di mana P2 lebih kaya detail spasial untuk objek kecil, sementara P5 lebih mewakili informasi semantik untuk objek berukuran besar yang memungkinkan deteksi objek dalam berbagai ukuran ke *Faster R-CNN*, yaitu algoritma deteksi objek yang meningkatkan efisiensi dengan menghilangkan ketergantungan pada metode *Selective Search*, serta menghasilkan *region proposals* secara langsung melalui RPN [18].



Gambar 3. Arsitektur FPN dengan Faster R-CNN

Feature maps multi-level kemudian diteruskan ke *Region Proposal Network* (RPN) yang menghasilkan kandidat wilayah (*region proposals*) melalui *anchor boxes* pada tiap level *feature map* [19]. RPN memiliki dua cabang yaitu *Classification Loss*, yang menentukan apakah suatu proposal *anchor box* mengandung objek atau tidak dan *Regression Loss*, yang memperbaiki posisi serta ukuran *anchor box* agar lebih sesuai dengan objek sebenarnya. Dari kedua cabang akan dihitung *Total Loss Function* [20], dapat dilihat pada Persamaan 1, dengan proposal terbaik diseleksi menggunakan *Non-Maximum Suppression* (NMS). Hasil proposal dilanjutkan ke *Region of Interest* (RoI) *Pooling* yang menyelaraskan ukuran sebelum klasifikasi akhir [21], sehingga model dapat memberikan label “Informasi Gizi” sekaligus menghasilkan *bouding box* pada area yang terdeteksi.

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

dimana,

- p_i = probabilitas *anchor box*
- p_i^* = *ground truth* $\begin{cases} 1 \text{ positif anchor} \\ 0 \text{ negatif anchor} \end{cases}$
- N_{cls} = jumlah total *anchor box*
- N_{reg} = total *anchor* positif
- t_i = koordinat prediksi
- t_i^* = koordinat *ground truth*

Pada L_{reg} , menggunakan fungsi *Smooth L1 Loss* didefinisikan sebagai berikut:

$$SmoothL1 = \begin{cases} 0,5(t_i, t_i^*)^2, & \text{jika } |t_i, t_i^*| < 1 \\ |t_i, t_i^*| - 0,5, & \text{jika } |t_i, t_i^*| \geq 1 \end{cases}$$

3.3 Konfigurasi Pelatihan

Pada tahap ini, dilakukan pelatihan model deteksi objek menggunakan arsitektur *Faster R-CNN* dengan *backbone* ResNet-50 dan *Feature Pyramid Network* (FPN) yang diimplementasikan melalui *framework Detectron2*. Pelatihan dimulai dengan memuat bobot awal (*pretrained weights*) dari *dataset COCO*, kemudian dilakukan penyesuaian terhadap *dataset* yang

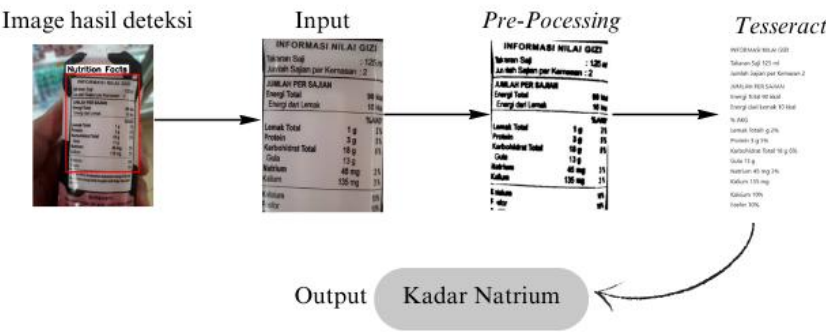
digunakan dalam penelitian. *Hyperparameter* utama yang digunakan dalam proses pelatihan ditunjukkan pada Tabel 1.

Tabel 1. Hyperparameter Pelatihan Model

Hyperparameter	Nilai	Keterangan
<i>Optimizer</i>	SGD	<i>Optimizer</i> bawaan <i>Detectron2</i> untuk pembaruan bobot
<i>Learning Rate (BASE_LR)</i>	0,00025	Pelatihan stabil dan mencegah divergensi
<i>Batch Size (IMS_PER_BATCH)</i>	2	Disesuaikan dengan keterbatasan GPU
<i>Max Iterations (MAX_ITER)</i>	5000	Memastikan konvergensi model
<i>ROI Batch Size per Image</i>	128	Jumlah proposal per gambar
Jumlah Kelas (<i>NUM_CLASSES</i>)	1	Satu kelas objek, yaitu “Informasi Gizi”
<i>Pretrained Weights</i>	COCO	Bobot awal dari <i>COCO dataset</i> untuk <i>transfer learning</i>
<i>Dataloader Workers</i>	2	Jumlah pekerja untuk mempercepat pemrosesan data

3.4 Optical Character Recognition (OCR)

OCR adalah teknologi yang mengubah teks dalam gambar menjadi teks digital yang dapat diedit atau diproses komputer [7]. Proses OCR dapat dilihat pada Gambar 4. Proses ini terdiri dari tiga tahap utama, yaitu *pre-processing*, ekstraksi teks, dan analisis hasil OCR. Pada tahap *pre-processing*, digunakan metode *adaptive thresholding* untuk meningkatkan keterbacaan teks. Ekstraksi teks dilakukan menggunakan *Tesseract OCR* yang didukung oleh teknologi jaringan saraf *Long Short-Term Memory (LSTM)* dan pencocokan pola berbasis *regular expression (regex)* untuk meningkatkan akurasi pengenalan teks [22].



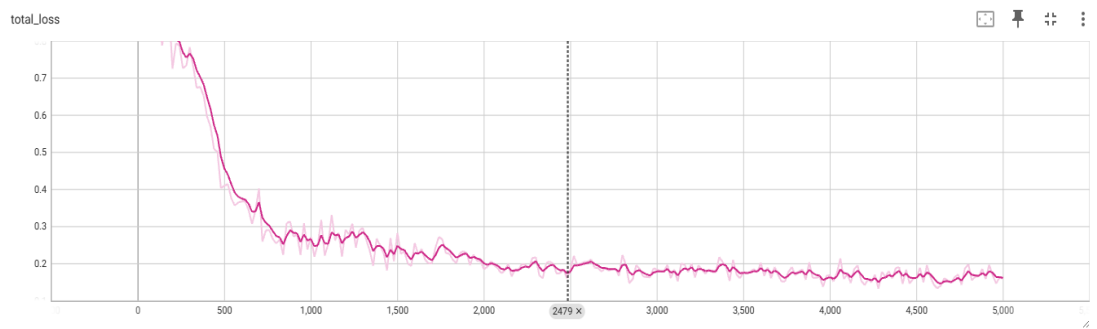
Gambar 4. Proses Optical Character Recognition (OCR)

4. Hasil dan Diskusi

Penelitian ini menggunakan dua jalur, yaitu mendeteksi area informasi gizi pada kemasan produk menggunakan FPN dengan algoritma *Faster R-CNN*, dan mengekstraksi teks dari kadar Natrium menggunakan OCR.

4.1 Feature Pyramid Network (FPN) dengan Algoritma Faster R-CNN

Proses implementasi dilakukan menggunakan *framework* Detectron2 dari *Facebook AI Research (FAIR)*, dengan evaluasi mencakup *loss* pada setiap langkah pelatihan serta analisis metrik *Average Precision (AP)* dan *Average Recall (AR)* pada berbagai ambang batas *Intersection over Union (IoU)*. Berdasarkan log pelatihan hingga iterasi ke-5000, model dengan performa terbaik berdasarkan total *loss* diperoleh pada iterasi ke-4279, sebagaimana ditunjukkan pada Tabel 2 dan grafik Total *loss* dapat dilihat pada Gambar 5.



Gambar 5. Grafik Total Loss

Grafik *loss* pada Gambar 5 menampilkan total *training loss* yang dicatat secara otomatis oleh Detectron2 selama proses pelatihan. Secara default, *framework* ini tidak menampilkan *validation loss* secara langsung dalam grafik, melainkan melakukan evaluasi validasi secara terpisah menggunakan metrik *Average Precision* (AP) dan *Average Recall* (AR) pada data validasi.

Tabel 2. Nilai Loss Iterasi ke-4279

Komponen Loss	Akurasi
Total Loss	0,1362
Classification Loss	0,03198
Box Regression Loss	0,11000
RPN Classification	0,00002488
RPN Localization	0,002076

Pada Tabel 1, nilai total *loss* yang rendah menunjukkan bahwa model belajar secara efisien dan stabil. Selanjutnya, hasil evaluasi menggunakan *COCOEvaluator* menunjukkan bahwa akurasi *Average Precision* (AP) terbaik pada ambang $\text{IoU} = 0,50$ mencapai 98,90%, dan *Average Recall* (AR) sebesar 85,00% dengan maksimum deteksi sebanyak 100 gambar, sebagaimana ditunjukkan pada Tabel 3.

Tabel 3. Hasil Final Evaluasi Model

Evaluasi Metrik	Akurasi
AP [IoU = 0.50]	0,989
AP [IoU = 0.75]	0,909
AP [IoU = 0.50:0.95]	0,789
AR [maxDets = 1]	0,838
AR [maxDets = 10]	0,850
AR [maxDets = 100]	0,850

4.2 Optical Character Recognition (OCR)

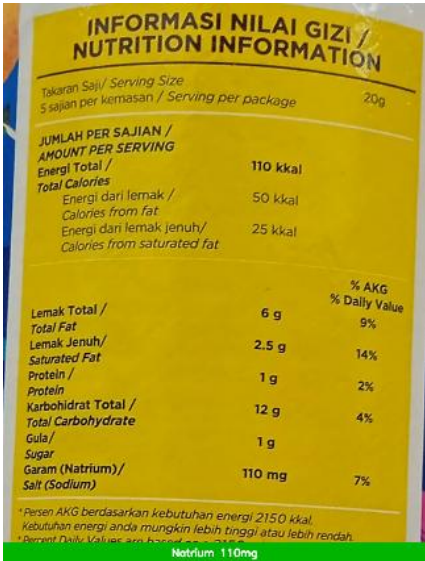
Pada tahap ini, dilakukan proses pengenalan teks kadar natrium dari gambar informasi gizi yang telah terdeteksi oleh model FPN. Berbeda dengan pengujian FPN yang menggunakan 187 gambar subset test dalam pengujian area informasi gizi, evaluasi OCR menggunakan 400 gambar uji terpisah yang dipilih, dibagi ke dalam empat batch. Evaluasi difokuskan pada dua aspek, yaitu keberhasilan membaca kata “natrium” dan keberhasilan mengekstraksi nilai kadarnya (mg), dengan hasil disajikan pada Tabel 4.

Tabel 4. Hasil Evaluasi Tesseract OCR

Batch	Total Gambar	Kata Natrium (terbaca)	Kadar Natrium mg (terbaca)	Akurasi per Batch (%)
1	100	95	86	90,50%
2	100	93	79	86,00%
3	100	87	63	75,00%
4	100	94	84	89,00%
		Rata-rata	85,13%	

Berdasarkan Tabel 3, sistem berhasil mendeteksi kata “natrium” pada seluruh *batch* yang diuji, meskipun terdapat variasi performa di setiap batch. Perbedaan hasil ini menunjukkan adanya tantangan dalam mengenali karakter numerik, yang cenderung lebih rentan terhadap gangguan (*noise*), variasi bentuk teks, serta kualitas gambar yang kurang optimal, misalnya teks buram, miring, atau terkena pantulan cahaya. Hal ini menunjukkan perlunya peningkatan tahap *preprocessing* agar hasil lebih konsisten. Rata-rata akurasi dari keempat *batch* mencapai 85,13%, yang menunjukkan bahwa performa OCR cukup baik dalam mengekstraksi informasi teks kadar natrium. Contoh hasil pengujian citra dapat dilihat pada Tabel 5.

Tabel 5. Contoh Citra Berhasil dan Gagal

Kondisi Citra	Citra Asli	Hasil Deteksi	Hasil
Lurus dan teks jelas	Asli: Natrium 110 mg 	Hasil deteksi: Natrium 110 mg 	Berhasil
Sudut miring dan melengkung	Asli: Natrium 90 mg 	Hasil deteksi: Natrium tidak ditemukan 	Gagal

Kondisi Citra	Citra Asli	Hasil Deteksi	Hasil
Buram, tidak fokus dan pantulan cahaya	Asli: Natrium 115 mg 	Hasil deteksi: Kadar tidak ditemukan 	Gagal
Teks terlalu tebal	Asli : Natrium 460 mg 	Hasil deteksi: Natrium tidak ditemukan 	Gagal

Tabel 5 menampilkan beberapa contoh citra kemasan produk beserta hasil deteksi dan ekstraksi informasi kandungan natrium oleh sistem. Setiap baris pada tabel mencakup kondisi citra, gambar asli, hasil keluaran sistem, serta status akhir: “Berhasil” jika nilai natrium dikenali dengan tepat, dan “Gagal” apabila sistem tidak mendeteksi informasi atau menghasilkan keluaran yang keliru. Sistem mampu mendeteksi kandungan natrium secara akurat pada citra dengan label yang lurus dan teks yang jelas. Namun, kegagalan terjadi pada kondisi visual tertentu, seperti label yang miring atau permukaan kemasan yang melengkung, yang menyebabkan distorsi teks sehingga informasi natrium tidak dapat dikenali. Citra yang buram, tidak fokus, atau terkena pantulan cahaya juga mengakibatkan hasil keluaran tidak akurat, di mana sistem hanya mampu mengenali sebagian teks. Selain itu, pada citra dengan teks yang terlalu tebal, sistem kesulitan mengekstraksi nilai dengan benar; misalnya, nilai asli “460 mg” tidak terbaca dengan baik oleh sistem. Temuan ini menunjukkan bahwa faktor-faktor seperti posisi label, pencahayaan, tingkat fokus, dan ketebalan teks memiliki pengaruh signifikan terhadap akurasi sistem dalam mengekstraksi informasi kandungan natrium dari gambar kemasan produk.

5. Kesimpulan dan Saran

Penelitian ini mengembangkan sistem otomatis untuk mendeteksi dan mengekstraksi informasi nilai gizi dari gambar kemasan produk, dengan fokus pada kandungan natrium sebagai studi kasus. Sistem mengintegrasikan model deteksi objek berbasis *Feature Pyramid Network* (FPN) dan *Faster R-CNN* menggunakan *framework Detectron2*, serta *Tesseract Optical Character Recognition* (OCR) untuk ekstraksi teks. Sebanyak 3.028 gambar digunakan dan dibagi ke dalam set pelatihan, validasi, dan pengujian FPN. Hasil pelatihan menunjukkan

performa model yang baik, dengan total *loss* terendah sebesar 0,1362 pada iterasi ke-4.279. Evaluasi menunjukkan akurasi deteksi sebesar 98,90% (AP50) dan (AR) sebesar 85,00% dengan maksimal 100 deteksi. Setelah proses deteksi, citra diproses lebih lanjut untuk ekstraksi teks. Pengujian OCR pada 400 gambar berbeda menghasilkan rata-rata akurasi ekstraksi teks sebesar 85,13%. Tantangan utama dalam ekstraksi disebabkan oleh kualitas gambar yang kurang baik, seperti buram, miring, terkena pantulan cahaya, atau memiliki teks yang terlalu tebal.

Meskipun penelitian ini difokuskan pada kandungan natrium, metode yang diusulkan bersifat umum dan dapat diperluas untuk membaca keseluruhan teks pada label informasi gizi, seperti protein, karbohidrat, atau gula. Dengan demikian, penelitian ini berpotensi memberikan manfaat yang lebih luas dalam pengembangan sistem pembacaan label gizi otomatis. Untuk pengembangan selanjutnya, disarankan penggunaan *dataset* yang lebih bervariasi dengan kondisi nyata guna meningkatkan kemampuan generalisasi model. Akurasi OCR dapat ditingkatkan melalui proses *fine-tuning* pada *dataset* berbahasa Indonesia. Selain itu, pemanfaatan model kecerdasan buatan terkini juga berpotensi meningkatkan kinerja sistem. Ke depan, sistem ini dapat dikembangkan menjadi aplikasi berbasis *web* atau *mobile* untuk mendukung penggunaan praktis secara *real-time*.

Referensi

- [1] L. Hasibuan, "Batas Aman Konsumsi Gula dan Garam Harian Menurut WHO," CNBC Indonesia. Diakses: 3 Maret 2025. [Daring]. Tersedia pada: <https://www.cnbcindonesia.com/lifestyle/20241203142608-33-593000/batas-aman-konsumsi-gula-dan-garam-harian-menurut-who>
- [2] D. Arlinta, "Lebih dari 50 Persen Masyarakat Indonesia Konsumsi Garam Berlebihan," kompas.id. Diakses: 3 Maret 2025. [Daring]. Tersedia pada: <https://www.kompas.id/baca/humaniora/2023/03/28/lebih-dari-50-persen-masyarakat-indonesia-konsumsi-garam-berlebihan>
- [3] X. Shao *dkk.*, "Multi-Scale Feature Pyramid Network: A Heavily Occluded Pedestrian Detection Network Based on ResNet," *Sensors*, vol. 21, no. 5, hlm. 1820, Jan 2021, doi: 10.3390/s21051820.
- [4] W. Sheng, X. Yu, J. Lin, dan X. Chen, "Faster RCNN Target Detection Algorithm Integrating CBAM and FPN," *Applied Sciences*, vol. 13, no. 12, hlm. 6913, Jan 2023, doi: 10.3390/app13126913.
- [5] J. Kang, M. Ibrayim, dan A. Hamdulla, "MR-FPN: Multi-Level Residual Feature Pyramid Text Detection Network Based on Self-Attention Environment," *Sensors (Basel)*, vol. 22, no. 9, hlm. 3337, Apr 2022, doi: 10.3390/s22093337.
- [6] J. Pardede dan H. Hardiansah, "Deteksi Objek Kereta Api menggunakan Metode Faster R-CNN dengan Arsitektur VGG 16," *MIND (Multimedia Artificial Intelligent Networking Database) Journal*, vol. 7, no. 1, hlm. 21–36, Jun 2022, doi: 10.26760/mindjournal.v7i1.21-36.
- [7] J. Wang, J. Tang, M. Yang, X. Bai, dan J. Luo, "Improving OCR-based Image Captioning by Incorporating Geometrical Relationship," dalam *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA: IEEE, Jun 2021, hlm. 1306–1315. doi: 10.1109/CVPR46437.2021.00136.
- [8] T. M. E. Saputra, A. Susanto, dan B. J. Carmelita, "Implementation of Tesseract OCR and Bounding Box for Text Extraction on Food Nutrition Labels," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, hlm. 1403–1412, Des 2024, doi: 10.47065/bits.v6i3.6107.
- [9] K. A. Nugraha, "Penerapan Optical Character Recognition untuk Pengenalan Variasi Teks pada Media Presentasi Pembelajaran," *Jurnal Buana Informatika*, vol. 15, no. 01, hlm. 69–78, Apr 2024, doi: <https://orcid.org/0000-0003-3401-8033>.
- [10] Fino Charli, Hadi Syaputra, Muhammad Akbar, Siti Sauda, dan Febriyanti Panjaitan, "Implementasi Metode Faster Region Convolutional Neural Network (Faster R-CNN) Untuk

- Pengenalan Jenis Burung Lovebird,” vol. 1, no. 3, Des 2020, doi: 10.51519/journalita.volume1.issue3.year2020.page185-197.
- [11] B. Yu, X. Pan, X. Li, G. Liu, dan Y. Ma, “Object Detection Algorithm Based on Improved Feature Pyramid,” *Scientific Programming*, vol. 2022, no. 1, hlm. 3583399, 2022, doi: 10.1155/2022/3583399.
- [12] R. D. Altarez, “Faster R-CNN, RetinaNet and Single Shot Detector in different ResNet backbones for marine vessel detection using cross polarization C-band SAR imagery,” *Remote Sensing Applications: Society and Environment*, vol. 36, hlm. 101297, Nov 2024, doi: 10.1016/j.rsase.2024.101297.
- [13] S. Hartanto, A. Sugiharto, dan S. N. Endah, “Optical Character Recognition Menggunakan Algoritma Template Matching Correlation,” *Jurnal Masyarakat Informatika*, vol. 5, no. 9, hlm. 1–12, Apr 2015, doi: 10.14710/jmasif.5.9.8435.
- [14] F. A. Rafi, A. Fanggidae, dan Y. T. Polly, “Sistem Pendeteksi Kerusakan Jalan Aspal Menggunakan Canny Edge Detection,” *J-Icon : Jurnal Komputer dan Informatika*, vol. 11, no. 1, hlm. 85–90, Mar 2023, doi: 10.35508/jicon.v11i1.10100.
- [15] A. M. Dwitala, A. Fanggidae, Y. T. Polly, dan B. B. Koten, “Klasifikasi Citra Biji Arbila di Kabupaten Kupang menggunakan k-Nearest Neighbor berdasarkan fitur Warna dan Tekstur,” *Jurnal Ilmiah Telekomunikasi, Kendali dan Elektronika Terapan*, vol. 12, no. 2, hlm. 177–186, Okt 2024, doi: 10.34010/telekontran.v12i2.14207.
- [16] A. Ridhovan dan A. Suharso, “Penerapan Metode Residual Network (Resnet) dalam Klasifikasi Penyakit pada Daun Gandum,” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 1, hlm. 58–65, Feb 2022, doi: 10.29100/jupi.v7i1.2410.
- [17] G. Zhao, W. Ge, dan Y. Yu, “GraphFPN: Graph Feature Pyramid Network for Object Detection,” 8 Januari 2022, arXiv:2108.00580. doi: 10.48550/arXiv.2108.00580.
- [18] G. Yang, J. Lei, Z. Zhu, S. Cheng, Z. Feng, dan R. Liang, “AFPN: Asymptotic Feature Pyramid Network for Object Detection,” 24 September 2023. doi: 10.48550/arXiv.2306.15988.
- [19] X. Dong, Y. Qin, Y. Gao, R. Fu, S. Liu, dan Y. Ye, “Attention-Based Multi-Level Feature Fusion for Object Detection in Remote Sensing Images,” *Remote Sensing*, vol. 14, no. 15, hlm. 3735, Jan 2022, doi: 10.3390/rs14153735.
- [20] S.-H. Tsang, “Review: Faster R-CNN (Object Detection),” TDS Archive. Diakses: 20 Juli 2025. [Daring]. Tersedia pada: <https://medium.com/data-science/review-faster-r-cnn-object-detection-f5685cb30202>
- [21] T. Bai dkk., “An Optimized Faster R-CNN Method Based on DRNet and RoI Align for Building Detection in Remote Sensing Images,” *Remote Sensing*, vol. 12, no. 5, hlm. 762, Jan 2020, doi: 10.3390/rs12050762.
- [22] I. H. A. Amin dan A. Aprilino, “Implementasi Algoritma Yolo dan Tesseract OCR pada Sistem Deteksi Plat Nomor Otomatis,” *Jurnal Teknoinfo*, vol. 16, no. 1, hlm. 54–59, Jan 2022, doi: 10.33365/jti.v16i1.1522.