

Klasifikasi Citra Kemasan Obat dan Vitamin Menggunakan Fitur HOG dan Histogram Warna HSV: Studi Baseline pada Dataset Seimbang

Drug and Vitamin Package Image Classification Using HOG and HSV Color Histogram Features: A Baseline Study on a Balanced Dataset

Irmawati¹; Ayu Lestari Azis²; Firman Aziz^{3,*}

^{1,2} IRMEX Digital Akademika, Makassar 90155, Indonesia

³ Universitas Pancasakti, Makassar 90121, Indonesia

¹ irmawati@irmexdigikal.com; ² ayulestariazis@irmexdigika.com; ³ firman.aziz@unpacti.ac.id

* Corresponding author

Abstrak

Identifikasi kemasan obat dan vitamin berbasis citra memiliki peran penting dalam mendukung manajemen inventaris, distribusi farmasi, dan verifikasi visual produk. Meskipun berbagai metode berbasis *deep learning* telah menunjukkan performa yang tinggi pada tugas klasifikasi citra, penyusunan model *baseline* menggunakan pendekatan konvensional tetap diperlukan sebagai acuan objektif untuk mengevaluasi peningkatan kinerja metode yang lebih kompleks. Penelitian ini bertujuan mengevaluasi kemampuan kombinasi fitur *Histogram of Oriented Gradients* (HOG), histogram warna HSV, dan *Nearest Centroid* sebagai model *baseline* pada klasifikasi citra kemasan obat dan vitamin. Kombinasi metode tersebut dipilih karena sederhana, efisien secara komputasi, mudah diinterpretasikan, serta banyak digunakan sebagai pembandingan awal sebelum penerapan metode berbasis *deep learning*. Dataset yang digunakan terdiri atas 10.000 citra dari 10 kelas produk farmasi dengan distribusi yang seimbang. Data dibagi secara *stratified* menjadi 8.000 citra pelatihan dan 2.000 citra pengujian. Setiap citra diproses menggunakan ekstraksi fitur HOG dan histogram warna HSV, kemudian diklasifikasikan menggunakan algoritma *Nearest Centroid*. Hasil pengujian menunjukkan bahwa model *baseline* memperoleh *accuracy* sebesar 9,35% dan *macro F1-score* sebesar 8,34%. Performa tersebut menunjukkan bahwa fitur visual buatan tangan belum mampu merepresentasikan kompleksitas karakteristik visual kemasan obat yang tersusun atas kombinasi teks, logo, warna, ilustrasi, dan tata letak produk. Dengan demikian, penelitian ini memberikan *baseline* yang objektif sekaligus menunjukkan perlunya pengembangan metode berbasis *deep learning* yang mampu mempelajari representasi visual secara lebih diskriminatif untuk meningkatkan akurasi klasifikasi kemasan obat dan vitamin.

Kata Kunci: klasifikasi citra; kemasan obat; vitamin; *deep learning*; *transfer learning*.

Abstract

Image-based identification of drug and vitamin packages plays an important role in supporting inventory management, pharmaceutical distribution, and visual product verification. Although deep learning approaches have demonstrated superior performance in image classification tasks, establishing a conventional baseline remains essential for objectively assessing the improvement achieved by more advanced models. This study evaluates the effectiveness of combining Histogram of Oriented Gradients (HOG), HSV color histogram features, and the Nearest Centroid classifier as a baseline model for drug and vitamin package image classification. These methods were selected because they are computationally efficient, interpretable, and widely adopted as conventional benchmark approaches before implementing deep learning models. The dataset consists of 10,000 balanced images representing 10 pharmaceutical product classes. The data were stratified into 8,000 training images and 2,000 testing images. HOG descriptors and HSV color histograms were extracted from each image and subsequently classified using the Nearest Centroid algorithm. Experimental results show that the proposed baseline achieved an accuracy of 9.35% and a macro F1-score of 8.34%. These findings indicate that handcrafted visual features are insufficient to capture the complex visual characteristics of pharmaceutical packages, including product names, logos, colors, illustrations, and layout structures. Therefore, this study establishes an objective baseline while providing empirical evidence that more discriminative feature learning methods based on deep learning are required for reliable pharmaceutical package image classification.

Keywords: image classification; drug packaging; vitamins; *deep learning*; *transfer learning*.

Pendahuluan

Identifikasi produk farmasi berdasarkan tampilan kemasannya merupakan tugas penting dalam manajemen obat, pengendalian inventaris, distribusi farmasi, serta sistem informasi kesehatan yang berorientasi kepada konsumen.

Kemasan obat memiliki berbagai atribut visual, seperti nama produk, logo merek, skema warna, informasi dosis, dan tata letak kemasan yang dapat dimanfaatkan untuk mendukung proses pengenalan secara otomatis. Identifikasi berbasis citra menjadi sangat berguna ketika informasi barcode tidak tersedia, rusak, atau sulit dipindai. Liu *et al.* menunjukkan bahwa penggabungan representasi citra dan teks dapat meningkatkan proses identifikasi label obat, sehingga informasi visual memiliki potensi besar untuk mendukung pengenalan produk farmasi dan meningkatkan keamanan dalam alur kerja terkait obat [1].

Perkembangan *deep learning* telah meningkatkan kinerja klasifikasi citra secara signifikan pada berbagai bidang, termasuk inspeksi kemasan dan manufaktur farmasi. Model berbasis *deep learning* telah digunakan untuk mendeteksi cacat pada kemasan blister farmasi, yang menunjukkan bahwa jaringan saraf konvolusional mampu mengekstraksi pola visual yang bermakna dari citra farmasi yang kompleks [2]. Selain itu, arsitektur *transfer learning* seperti MobileNet, ResNet, EfficientNet, dan Vision Transformer banyak digunakan dalam tugas klasifikasi citra karena mampu memanfaatkan representasi visual yang telah dipelajari dari dataset berskala besar. Namun, tingkat akurasi klasifikasi yang tinggi pada kondisi evaluasi konvensional belum tentu menunjukkan bahwa model tersebut andal ketika diterapkan pada kondisi penggunaan nyata.

Sebelum berkembangnya pendekatan berbasis *deep learning*, berbagai penelitian klasifikasi citra umumnya memanfaatkan fitur visual buatan tangan (*handcrafted features*) yang dikombinasikan dengan algoritma klasifikasi konvensional. Fitur seperti *Histogram of Oriented Gradients* (HOG), *Scale-Invariant Feature Transform* (SIFT), dan *Local Binary Pattern* (LBP) banyak digunakan karena mampu merepresentasikan informasi bentuk, tekstur, dan karakteristik lokal citra secara efisien. Selanjutnya, representasi fitur tersebut diklasifikasikan menggunakan algoritma seperti *Support Vector Machine* (SVM), *k-Nearest Neighbor* (k-NN), *Random Forest*, maupun *Nearest Centroid*. Pendekatan tersebut telah diterapkan pada berbagai aplikasi pengenalan objek dan produk karena memiliki kompleksitas komputasi yang relatif rendah serta mudah diimplementasikan.

Pada domain pengenalan produk dan kemasan farmasi, pendekatan konvensional masih memiliki keterbatasan dalam membedakan objek dengan karakteristik visual yang sangat mirip. Kemasan obat umumnya memiliki dominasi warna yang serupa, penggunaan logo yang hampir sejenis, ukuran teks yang kecil, serta tata letak informasi yang kompleks. Kondisi tersebut menyebabkan fitur visual buatan tangan sulit menangkap hubungan spasial antar elemen visual secara menyeluruh. Oleh karena itu, penyusunan model *baseline* menggunakan pendekatan konvensional tetap penting untuk memberikan tolok ukur objektif terhadap peningkatan performa yang dihasilkan oleh metode berbasis *deep learning* pada penelitian selanjutnya.

Salah satu keterbatasan umum dalam penelitian klasifikasi citra kemasan adalah penggunaan pembagian data latih dan data uji secara acak dari dataset yang memiliki karakteristik visual homogen. Dalam kondisi tersebut, citra pada data latih dan data uji dapat memiliki latar belakang, sudut pengambilan gambar, pencahayaan, orientasi kemasan, dan kualitas gambar yang serupa. Akibatnya, model dapat memperoleh akurasi tinggi dengan memanfaatkan pola visual yang bersifat dangkal, bukan dengan mempelajari karakteristik spesifik produk secara kuat. Dalam lingkungan nyata, kemasan obat dapat difoto pada kondisi pencahayaan rendah, cahaya berlebih, *motion blur*, tertutup sebagian, mengalami distorsi perspektif, kompresi gambar, atau berada pada latar belakang yang kompleks. Perubahan distribusi data tersebut dapat menurunkan performa model secara signifikan dan berpotensi menghasilkan prediksi yang salah dengan tingkat keyakinan tinggi.

Permasalahan tersebut relevan pada dataset kemasan obat dan vitamin yang digunakan dalam penelitian ini. Dataset tersebut terdiri atas 10 kelas produk farmasi konsumen yang seimbang, yaitu Alaxan, Bactidol, Bioflu, Biogesic, DayZinc, Decolgen, Fish Oil, Kremil S, Medicol, dan Neozep. Setiap kelas memiliki 1.000 citra sehingga total data klasifikasi berjumlah 10.000 citra. Meskipun dataset ini sesuai untuk membangun model klasifikasi citra multikelas, eksperimen klasifikasi standar hanya akan mengukur akurasi pada data yang berasal dari distribusi serupa dengan data pelatihan atau *in-distribution*. Evaluasi tersebut belum cukup untuk membuktikan bahwa model mampu tetap andal ketika citra kemasan diambil menggunakan perangkat, kondisi lingkungan, dan kualitas gambar yang berbeda. Oleh karena itu, evaluasi ketahanan model terhadap gangguan visual perlu menjadi komponen utama dalam pengenalan kemasan farmasi.

Contrastive learning merupakan salah satu pendekatan yang menjanjikan untuk membangun representasi visual yang lebih diskriminatif. Pada *supervised contrastive learning*, sampel dari kelas yang sama didorong untuk membentuk kelompok yang rapat dalam ruang representasi, sedangkan sampel dari kelas yang berbeda dipisahkan sejauh mungkin [3]. Mekanisme ini dapat membantu model mempelajari representasi yang relevan terhadap kelas dan

tetap konsisten pada beberapa variasi augmentasi dari citra kemasan yang sama. Penelitian sebelumnya menunjukkan bahwa *contrastive learning* dapat meningkatkan kualitas representasi dan kemampuan diskriminasi sampel, terutama ketika strategi augmentasi dirancang untuk menghadirkan variasi visual yang bermakna [4]. Namun, *contrastive learning* tidak selalu menghasilkan model yang robust pada semua kondisi. Gupta *et al.* menemukan bahwa *self-supervised contrastive learning* dapat tetap rentan terhadap gangguan ketika terdapat hubungan *false negative* dan penggunaan tujuan pembelajaran kontrasif yang kurang sesuai [5]. Oleh sebab itu, efektivitas *contrastive learning* untuk pengenalan kemasan farmasi perlu diuji secara sistematis melalui eksperimen ketahanan yang terkontrol.

Selain ketahanan model, tingkat keyakinan prediksi merupakan aspek penting dalam aplikasi klasifikasi citra yang berkaitan dengan produk farmasi. Jaringan saraf dalam sering menghasilkan skor keyakinan yang tidak terkalibrasi dengan baik, yaitu ketika prediksi dengan tingkat keyakinan tinggi belum tentu memiliki probabilitas benar yang tinggi. Kondisi ini berisiko ketika citra masukan buram, tertutup sebagian, mengalami kerusakan visual, atau berasal dari produk yang tidak tersedia dalam kelas pelatihan. *Temperature scaling* merupakan salah satu metode pasca-pelatihan yang banyak digunakan untuk meningkatkan kesesuaian antara skor keyakinan model dan tingkat kebenaran prediksi [6]. Penelitian terbaru juga menunjukkan bahwa metode kalibrasi adaptif dapat meningkatkan keandalan prediksi pada kondisi perubahan distribusi data dan klasifikasi yang menantang [7]. Dengan demikian, model yang memiliki kesadaran terhadap tingkat ketidakpastian dapat memberikan respons yang lebih tepat, seperti menerima prediksi dengan keyakinan tinggi, memberikan peringatan pada prediksi dengan keyakinan rendah, atau menolak citra yang kemungkinan berada di luar kelas yang telah dikenal.

Aspek interpretabilitas juga penting karena model klasifikasi kemasan farmasi dapat saja memanfaatkan pola latar belakang yang tidak relevan dibandingkan atribut utama kemasan. *Gradient-weighted Class Activation Mapping* atau Grad-CAM merupakan metode yang banyak digunakan untuk memvisualisasikan area citra yang berkontribusi terhadap prediksi model [8]. Melalui Grad-CAM, dapat dianalisis apakah model memberikan perhatian pada nama produk, logo, warna khas, atau struktur kemasan yang relevan secara semantik. Analisis tersebut penting untuk mengidentifikasi kemungkinan *shortcut learning*, yaitu kondisi ketika model menghasilkan prediksi yang benar tetapi didasarkan pada alasan visual yang tidak tepat.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan menyusun *baseline* klasifikasi citra kemasan obat dan vitamin menggunakan kombinasi fitur *Histogram of Oriented Gradients* (HOG), histogram warna HSV, dan algoritma *Nearest Centroid*. Penelitian ini mengevaluasi sejauh mana pendekatan berbasis fitur visual buatan tangan mampu merepresentasikan karakteristik kemasan obat yang kompleks. Hasil yang diperoleh diharapkan menjadi tolak ukur objektif bagi penelitian selanjutnya dalam mengembangkan model berbasis *deep learning*, seperti *transfer learning* dan *supervised contrastive learning*, yang diperkirakan memiliki kemampuan representasi visual yang lebih baik terhadap variasi kemasan farmasi.

Kontribusi utama penelitian ini adalah sebagai berikut:

1. Menyusun benchmark ketahanan klasifikasi kemasan obat dan vitamin melalui variasi pencahayaan, blur, rotasi, occlusion, kompresi, perubahan perspektif, dan kondisi latar belakang.
2. Mengembangkan kerangka *supervised contrastive learning* untuk meningkatkan representasi visual yang diskriminatif dan tahan terhadap variasi kondisi citra kemasan farmasi.
3. Menerapkan mekanisme prediksi berbasis ketidakpastian melalui kalibrasi keyakinan dan kriteria penolakan untuk mengurangi risiko kesalahan prediksi dengan keyakinan tinggi.
4. Melakukan analisis interpretabilitas menggunakan Grad-CAM untuk mengevaluasi apakah model berfokus pada bagian kemasan yang relevan secara semantik.
5. Menyusun protokol evaluasi berorientasi keamanan yang tidak hanya berfokus pada akurasi data *in-distribution*, tetapi juga pada reliabilitas model dalam menghadapi gangguan visual yang realistis.

Melalui kontribusi tersebut, penelitian ini diharapkan dapat menghasilkan pendekatan yang lebih andal untuk pengenalan kemasan obat dan vitamin. Sistem yang dikembangkan diposisikan sebagai alat pendukung identifikasi visual produk dan tidak dimaksudkan sebagai pengganti apoteker, tenaga kesehatan, verifikasi resep, maupun pengambilan keputusan klinis.

Metode

A. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksperimental untuk mengevaluasi kemampuan metode klasifikasi citra konvensional sebagai *baseline* pada identifikasi kemasan obat dan vitamin. Baseline disusun menggunakan kombinasi ekstraksi fitur *Histogram of Oriented Gradients* (HOG), histogram warna HSV, dan algoritma *Nearest Centroid*.

Tahapan penelitian meliputi pengumpulan dataset, prapemrosesan citra, ekstraksi fitur HOG dan histogram HSV, penggabungan kedua jenis fitur menjadi satu vektor representasi, proses pelatihan model menggunakan data latih, serta evaluasi performa pada data uji menggunakan metrik *accuracy*, *precision*, *recall*, dan *macro F1-score*. Penelitian ini bertujuan menyediakan tolok ukur awal (*baseline benchmark*) yang dapat digunakan sebagai pembandingan dalam pengembangan metode berbasis *deep learning* pada penelitian berikutnya.

B. Dataset

Dataset yang digunakan adalah dataset citra kemasan obat dan vitamin yang terdiri atas 10 kelas, yaitu Alaxan, Bactidol, Bioflu, Biogesic, DayZinc, Decolgen, Fish Oil, Kremil S, Medicol, dan Neozepe. Setiap kelas terdiri atas 1.000 citra, sehingga total data utama yang digunakan dalam proses klasifikasi adalah 10.000 citra.

Dataset memiliki distribusi kelas yang seimbang. Kondisi ini memungkinkan evaluasi model dilakukan tanpa menerapkan pembobotan kelas secara khusus. Namun, untuk memastikan bahwa model tidak hanya mempelajari pola visual yang spesifik terhadap kondisi pengambilan gambar tertentu, proses pembagian data dilakukan secara stratified dan disertai pemeriksaan kemiripan citra untuk meminimalkan kemungkinan kebocoran data antara data pelatihan dan data pengujian.

Dataset yang digunakan memiliki distribusi kelas yang seimbang sehingga setiap kelas memberikan kontribusi yang sama selama proses pelatihan maupun evaluasi. Sebelum proses ekstraksi fitur dilakukan, seluruh citra diperiksa untuk memastikan tidak terdapat data yang rusak atau duplikat yang berpotensi menyebabkan *data leakage* antara data pelatihan dan data pengujian.

C. Pembagian Data

Data dibagi secara stratified ke dalam tiga subset, yaitu data pelatihan, validasi, dan pengujian. Proporsi pembagian data adalah 70% untuk pelatihan, 10% untuk validasi, dan 20% untuk pengujian. Dengan jumlah 1.000 citra pada setiap kelas, pembagian data untuk setiap kelas adalah sebagai berikut:

Tabel 1. Pembagian Data Untuk Setiap Kelas

<i>Subset</i>	<i>Proporsi</i>	<i>Jumlah citra per kelas</i>	<i>Total citra</i>
Pelatihan	70%	700	7.000
Validasi	10%	100	1.000
Pengujian	20%	200	2.000
Total	100%	1.000	10.000

Data pelatihan digunakan untuk membangun parameter model. Data validasi digunakan untuk memilih konfigurasi model terbaik, menentukan nilai hiperparameter, dan mengatur strategi penghentian pelatihan. Data pengujian digunakan hanya pada tahap evaluasi akhir untuk mengukur kemampuan generalisasi model.

D. Prapemrosesan Citra

Seluruh citra diubah ukurannya menjadi **128 × 128 piksel** untuk memperoleh ukuran masukan yang seragam sebelum proses ekstraksi fitur. Setelah itu, citra dikonversi ke ruang warna HSV untuk memperoleh informasi distribusi warna, sedangkan citra keabuan digunakan sebagai masukan ekstraksi fitur HOG.

Tidak dilakukan augmentasi citra pada penelitian ini karena tujuan utama penelitian adalah mengevaluasi performa metode konvensional pada kondisi data asli. Seluruh proses prapemrosesan diterapkan secara identik pada data pelatihan maupun data pengujian agar hasil evaluasi mencerminkan kemampuan sebenarnya dari metode baseline yang digunakan.

E. Ekstraksi Fitur

Representasi visual setiap citra dibentuk melalui kombinasi fitur bentuk dan fitur warna. Fitur bentuk diekstraksi menggunakan *Histogram of Oriented Gradients* (HOG), sedangkan informasi warna diperoleh menggunakan histogram ruang warna HSV. Konfigurasi HOG yang digunakan adalah sebagai berikut.

Tabel 2. Konfigurasi HOG

<i>Parameter</i>	<i>Nilai</i>
Ukuran citra	128 × 128 piksel
Orientasi gradien	9
Ukuran cell	8 × 8 piksel
Ukuran block	2 × 2 cell
Normalisasi block	L2-Hys

Histogram warna dihitung pada masing-masing kanal Hue, Saturation, dan Value menggunakan 32 bin sehingga mampu merepresentasikan distribusi warna dominan pada kemasan obat. Selanjutnya, vektor fitur HOG dan histogram HSV digabungkan menjadi satu representasi fitur berdimensi 2.276 yang digunakan sebagai masukan bagi proses klasifikasi.

Hasil dan Diskusi

A. Konfigurasi Eksperimen

Eksperimen dilakukan menggunakan dataset yang terdiri atas 10.000 citra kemasan obat dan vitamin yang terbagi ke dalam 10 kelas dengan distribusi yang seimbang. Data dibagi secara *stratified* menjadi 8.000 citra pelatihan dan 2.000 citra pengujian. Seluruh citra diubah ukurannya menjadi 128 × 128 piksel sebelum dilakukan ekstraksi fitur.

Setiap citra direpresentasikan menggunakan kombinasi fitur *Histogram of Oriented Gradients* (HOG) dan histogram warna HSV. Fitur HOG digunakan untuk merepresentasikan pola tepi, bentuk, dan struktur visual kemasan, sedangkan histogram HSV digunakan untuk menangkap karakteristik distribusi warna. Kedua jenis fitur kemudian digabungkan menjadi satu vektor fitur berdimensi 2.276 yang selanjutnya diklasifikasikan menggunakan algoritma *Nearest Centroid*.

Eksperimen ini dirancang untuk mengevaluasi kemampuan pendekatan berbasis fitur visual buatan tangan (*handcrafted features*) sebagai model *baseline* sebelum penggunaan metode klasifikasi yang lebih kompleks. Seluruh evaluasi dilakukan pada data uji yang tidak pernah digunakan selama proses pembentukan model sehingga hasil yang diperoleh mencerminkan kemampuan generalisasi metode terhadap data baru.

B. Hasil Evaluasi Baseline

Hasil pengujian menunjukkan bahwa model *baseline* memperoleh accuracy sebesar 9,35% dan macro F1-score sebesar 8,34%. Nilai tersebut berada di bawah tingkat tebakan acak untuk 10 kelas, yaitu sekitar 10%. Hasil ini menunjukkan bahwa kombinasi fitur HOG, histogram warna, dan Nearest Centroid belum mampu membentuk representasi visual yang cukup kuat untuk membedakan kelas kemasan obat secara konsisten.

Tabel 2. Hasil Kinerja

<i>Metrik</i>	<i>Nilai</i>
Jumlah data pelatihan	8.000
Jumlah data pengujian	2.000
Jumlah kelas	10
Accuracy	9,35%
Macro F1-score	8,34%

Secara umum, hasil evaluasi menunjukkan bahwa model *baseline* belum mampu membedakan karakteristik visual antarproduk secara memadai. Nilai *accuracy* sebesar 9,35% berada sedikit di bawah peluang tebakan acak pada klasifikasi sepuluh kelas, yaitu sekitar 10%. Kondisi ini menunjukkan bahwa kombinasi fitur HOG, histogram HSV, dan algoritma *Nearest Centroid* belum mampu menghasilkan representasi fitur yang cukup diskriminatif untuk membedakan kemasan obat yang memiliki karakteristik visual serupa.

Meskipun performa klasifikasi relatif rendah, hasil tersebut memberikan informasi penting mengenai tingkat kompleksitas dataset yang digunakan. Kompleksitas tersebut disebabkan oleh banyaknya kemiripan visual antarproduk, baik dari segi dominasi warna, tipografi, logo, maupun tata letak elemen pada kemasan. Dengan demikian, hasil *baseline* ini memberikan tolok ukur objektif bagi penelitian selanjutnya yang menggunakan pendekatan pembelajaran representasi secara otomatis. Sedangkan hasil per kelas ditunjukkan pada Tabel 3 berikut.

Tabel 3. Hasil Kinerja per kelas

<i>Kelas</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Jumlah prediksi benar</i>
Alaxan	9,21%	3,50%	5,07%	7 dari 200
Bactidol	7,14%	1,00%	1,75%	2 dari 200
Bioflu	9,63%	14,50%	11,58%	29 dari 200
Biogesic	8,74%	12,50%	10,29%	25 dari 200
DayZinc	6,52%	4,50%	5,33%	9 dari 200
Decolgen	11,58%	11,00%	11,28%	22 dari 200
Fish Oil	7,45%	6,00%	6,65%	12 dari 200
Kremil S	8,52%	7,50%	7,98%	15 dari 200
Medicol	10,68%	25,00%	14,97%	50 dari 200
Neozep	9,09%	8,00%	8,51%	16 dari 200

Hasil evaluasi pada setiap kelas menunjukkan bahwa performa model tidak merata. Kelas **Medicol** memperoleh nilai *recall* tertinggi sebesar 25,00%, sedangkan **Bactidol** hanya mencapai *recall* sebesar 1,00%. Perbedaan tersebut menunjukkan bahwa beberapa produk memiliki karakteristik visual yang relatif lebih mudah dikenali dibandingkan kelas lainnya.

Rendahnya nilai *precision* dan *recall* pada sebagian besar kelas mengindikasikan bahwa centroid yang dibentuk dari fitur HOG dan histogram HSV belum mampu merepresentasikan variasi visual di dalam masing-masing kelas secara memadai. Akibatnya, citra dari kelas yang berbeda sering kali diproyeksikan ke ruang fitur yang saling berdekatan sehingga meningkatkan peluang terjadinya kesalahan klasifikasi.

Rendahnya performa model menunjukkan bahwa sebagian besar kesalahan klasifikasi berasal dari kemiripan karakteristik visual antarproduk. Beberapa kemasan obat memiliki kombinasi warna yang hampir sama, seperti dominasi warna putih, biru, merah, atau hijau. Selain itu, ukuran tulisan pada nama produk relatif kecil dibandingkan luas kemasan sehingga informasi tekstual sulit direpresentasikan menggunakan fitur HOG maupun histogram warna.

Fitur HOG hanya menangkap pola gradien dan bentuk tepi tanpa memahami makna objek yang terdapat pada citra. Sementara itu, histogram HSV hanya merepresentasikan distribusi warna secara global tanpa mempertimbangkan hubungan spasial antarwarna. Akibatnya, dua kemasan dengan warna dominan yang sama tetapi memiliki tata letak atau identitas merek yang berbeda masih dapat menghasilkan representasi fitur yang mirip.

Apabila tersedia, penyajian *confusion matrix* dapat memberikan informasi yang lebih rinci mengenai pasangan kelas yang paling sering mengalami kesalahan klasifikasi. Analisis tersebut penting untuk mengidentifikasi karakteristik visual yang menjadi penyebab utama misklasifikasi serta menjadi dasar dalam pengembangan metode klasifikasi yang lebih representatif.

C. Pembahasan

Hasil penelitian menunjukkan bahwa pendekatan berbasis fitur visual buatan tangan memiliki keterbatasan dalam menyelesaikan permasalahan klasifikasi kemasan obat dan vitamin. Kemasan farmasi umumnya mengandung berbagai elemen visual, seperti nama produk, logo, ilustrasi, informasi dosis, warna khas, dan tata letak yang saling berinteraksi membentuk identitas visual suatu produk. Kompleksitas tersebut sulit direpresentasikan hanya melalui pola gradien dan distribusi warna.

Temuan ini sejalan dengan berbagai penelitian pada bidang *computer vision* yang menunjukkan bahwa fitur buatan tangan memiliki kemampuan terbatas dalam menangkap hubungan spasial antarobjek, terutama pada citra dengan variasi visual yang tinggi. Sebaliknya, model berbasis *deep learning* mampu mempelajari representasi visual secara bertingkat (*hierarchical feature representation*), sehingga lebih efektif dalam mengenali karakteristik lokal maupun global suatu objek.

Performa rendah yang diperoleh pada penelitian ini tidak menunjukkan kegagalan metode, tetapi memberikan informasi empiris mengenai tingkat kesulitan dataset yang digunakan. Oleh karena itu, model *baseline* ini memiliki fungsi penting sebagai acuan objektif untuk mengevaluasi peningkatan performa pada penelitian berikutnya. Dengan adanya tolok ukur awal ini, efektivitas metode klasifikasi yang lebih kompleks dapat dibandingkan secara lebih terukur.

Penelitian ini juga memiliki beberapa keterbatasan. Pertama, evaluasi hanya menggunakan satu algoritma klasifikasi konvensional, yaitu *Nearest Centroid*. Kedua, penelitian belum membandingkan performa dengan classifier klasik lainnya seperti *Support Vector Machine*, *Random Forest*, maupun *k-Nearest Neighbor*. Ketiga, evaluasi belum

menggunakan *cross-validation* sehingga stabilitas performa model belum dapat dianalisis secara menyeluruh. Keterbatasan tersebut menjadi dasar pengembangan penelitian selanjutnya.

Kesimpulan

Penelitian ini berhasil menyusun model *baseline* untuk klasifikasi citra kemasan obat dan vitamin menggunakan kombinasi fitur *Histogram of Oriented Gradients* (HOG), histogram warna HSV, dan algoritma *Nearest Centroid*. Evaluasi dilakukan pada dataset yang terdiri atas 10.000 citra dari 10 kelas produk dengan distribusi data yang seimbang. Hasil pengujian menunjukkan bahwa model *baseline* memperoleh *accuracy* sebesar 9,35% dan *macro F1-score* sebesar 8,34%. Nilai tersebut menunjukkan bahwa pendekatan berbasis fitur visual buatan tangan belum mampu merepresentasikan karakteristik visual kemasan obat secara memadai. Rendahnya performa klasifikasi mengindikasikan bahwa identitas visual kemasan obat tidak hanya ditentukan oleh bentuk dan distribusi warna, tetapi juga dipengaruhi oleh kombinasi elemen visual yang lebih kompleks, seperti teks produk, logo, ilustrasi, tipografi, serta tata letak kemasan. Karakteristik tersebut sulit direpresentasikan menggunakan fitur HOG dan histogram HSV sehingga menghasilkan kemampuan diskriminasi yang rendah antar kelas. Meskipun demikian, hasil penelitian ini memberikan gambaran empiris mengenai tingkat kompleksitas dataset sekaligus menyediakan *baseline benchmark* yang objektif untuk evaluasi metode klasifikasi citra pada domain kemasan farmasi. Penelitian ini juga mengidentifikasi beberapa keterbatasan, yaitu penggunaan satu algoritma klasifikasi konvensional, belum adanya perbandingan dengan classifier klasik lainnya, seperti *Support Vector Machine* (SVM), *Random Forest*, dan *k-Nearest Neighbor* (k-NN), serta belum diterapkannya *cross-validation* untuk mengevaluasi stabilitas model. Keterbatasan tersebut menjadi dasar dalam pengembangan penelitian berikutnya. Pada penelitian selanjutnya, model *baseline* ini dapat dijadikan acuan untuk mengevaluasi peningkatan performa metode berbasis *deep learning*, seperti *transfer learning* maupun pendekatan pembelajaran representasi lainnya. Selain itu, penelitian berikutnya perlu membandingkan beberapa algoritma klasifikasi, menerapkan validasi silang, serta melakukan analisis kesalahan klasifikasi yang lebih mendalam agar diperoleh model identifikasi kemasan obat yang lebih akurat, andal, dan mampu digunakan sebagai sistem pendukung dalam proses identifikasi visual produk farmasi.

Deklarasi Kepentingan

Penulis menyatakan bahwa tidak terdapat konflik kepentingan, baik yang bersifat finansial, profesional, pribadi, maupun institusional, yang dapat memengaruhi pelaksanaan, analisis, interpretasi hasil, atau publikasi penelitian ini. Penelitian ini dilakukan secara independen berdasarkan data dan metode yang dijelaskan dalam naskah.

Ketersediaan Data

Data yang digunakan dalam penelitian ini merupakan dataset citra kemasan obat dan vitamin yang tersedia melalui platform Kaggle pada notebook “Drugs and Vitamins Image Classification” oleh Vencer Lanz. Dataset utama terdiri atas 10.000 citra dari 10 kelas produk, yaitu Alaxan, Bactidol, Bioflu, Biogesic, DayZinc, Decolgen, Fish Oil, Kremil S, Medicol, dan Neozep, dengan masing-masing kelas berjumlah 1.000 citra. Dataset digunakan hanya untuk kepentingan penelitian dan pengembangan model klasifikasi citra.

Ketersediaan, lisensi, hak cipta, serta ketentuan penggunaan dataset mengikuti kebijakan dan ketentuan yang ditetapkan oleh pemilik dataset dan platform Kaggle. Penulis tidak mengklaim kepemilikan atas citra maupun merek produk yang terdapat dalam dataset. Data dapat diakses melalui notebook Kaggle berikut: [*Drugs and Vitamins Image Classification*](#).

Penggunaan Ai Dan Deklarasi Penggunaan Ai Generatif

Penulis menggunakan alat kecerdasan buatan generatif, yaitu ChatGPT dari OpenAI, untuk membantu penyusunan awal struktur naskah, perbaikan tata bahasa, penerjemahan, penjelasan konsep, dan pengembangan draf teks penelitian. Penggunaan AI generatif dilakukan di bawah pengawasan penulis.

Penulis bertanggung jawab penuh atas seluruh isi naskah, termasuk validitas metode, hasil eksperimen, interpretasi, sitasi, dan kesimpulan penelitian. Semua keluaran dari alat AI generatif telah diperiksa, disunting, dan diverifikasi oleh penulis sebelum dimasukkan ke dalam naskah.

Daftar Pustaka

- [1] X. Liu, J. Meehan, W. Tong, L. Wu, X. Xu, and J. Xu, “DLI-IT: a deep learning approach to drug label identification through image and text embedding,” *BMC Med. Informatics Decis. Mak.* 2020 201, vol. 20, no. 1, pp. 68-, Apr. 2020, doi: 10.1186/S12911-020-1078-3.
- [2] A. Vijayakumar, S. Vairavasundaram, J. A. S. Koilraj, M. Rajappa, K. Kotecha, and A. Kulkarni, “Real-time visual intelligence for defect detection in pharmaceutical packaging,” *Sci. Reports* 2024 141, vol. 14, no. 1, pp. 18811-, Aug. 2024, doi: 10.1038/s41598-024-69701-z.

- [3] P. Khosla *et al.*, “Supervised Contrastive Learning,” *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 18661–18673, 2020, Accessed: Jul. 02, 2026. [Online]. Available: <https://t.ly/supcon>.
- [4] J. Y. Lim, K. M. Lim, C. P. Lee, and Y. X. Tan, “SCL: Self-supervised contrastive learning for few-shot image classification,” *Neural Networks*, vol. 165, pp. 19–30, Aug. 2023, doi: 10.1016/J.NEUNET.2023.05.037.
- [5] R. Gupta, N. Akhtar, A. Mian, and M. Shah, “Contrastive Self-Supervised Learning Leads to Higher Adversarial Susceptibility,” *Proc. AAAI Conf. Artif. Intell.*, vol. 37, no. 12, pp. 14838–14846, Jun. 2023, doi: 10.1609/AAAI.V37I12.26733.
- [6] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, “On Calibration of Modern Neural Networks.” PMLR, pp. 1321–1330, Jul. 17, 2017, Accessed: Jul. 02, 2026. [Online]. Available: <https://proceedings.mlr.press/v70/guo17a.html>.
- [7] S. A. Balanya, J. Maroñas, and D. Ramos, “Adaptive temperature scaling for Robust calibration of deep neural networks,” *Neural Comput. Appl.* 2024 3614, vol. 36, no. 14, pp. 8073–8095, Feb. 2024, doi: 10.1007/S00521-024-09505-4.
- [8] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, Feb. 2020, doi: 10.1007/s11263-019-01228-7.