

Sistem Inspeksi Visual Objek Menggunakan YOLOv11n dan *Convolutional Autoencoder* untuk Deteksi Cacat pada Komponen Botol

¹Fitri Nailatul Khobibah, ²Suryadhi, ³Sinung Widiyanto

^{1,2,3}Teknik Elektro, Universitas Hang Tuah, Surabaya

¹nailatulfitri@gmail.com, ²suryadhi@hangtuah.ac.id, ³sinung.widiyanto@hangtuah.ac.id

Abstract – Visual inspection of bottles on production lines is generally performed using semi-manual procedures, making it susceptible to errors caused by human fatigue. In addition, the wide variation in visual defects, combined with the limited availability of defective samples, makes explicit defect classification challenging and complicates the data labeling process. Therefore, this study aims to develop a visual inspection system capable of detecting objects and analyzing their quality within a single integrated pipeline. This approach combines component localization and defect identification into a unified process, in contrast to previous studies that generally perform these two processes separately. The proposed system integrates the YOLO algorithm for localizing bottle components (bottle body, cap, and label) with a Convolutional Autoencoder (CAE) for anomaly detection. Anomalies are identified based on the Mean Squared Error (MSE) of the reconstruction error and subsequently classified using a predefined threshold value. Experimental results demonstrate that the proposed system achieves object classification accuracies of 88.2% for bottle bodies, 89.1% for caps, and 84.5% for labels. The system also attains very high precision, ranging from 94.5% to 98.5%, indicating strong reliability in minimizing false positives. However, the recall values are comparatively lower, ranging from 79.2% to 89.6%, primarily due to missed detections of micro-scale defects and defects located in low-contrast areas.

Keywords — *Deep Learning, YOLO, Convolutional Autoencoder, Inspection System, Bottle Inspection*

Abstrak – Proses inspeksi visual botol pada lini produksi umumnya masih dilakukan secara semi-manual, yang berpotensi menimbulkan kesalahan akibat faktor kelelahan manusia. Selain itu, variasi cacat visual seringkali sulit diklasifikasikan secara eksplisit dengan jumlah data cacat yang terbatas, sehingga menyulitkan pelabelan data. Oleh karena itu, penelitian ini bertujuan merancang sistem inspeksi visual yang mampu mendeteksi objek sekaligus menganalisis kualitasnya dalam satu *pipeline* yang terintegrasi. Pendekatan ini menggabungkan proses lokalisasi komponen dan identifikasi cacat dalam satu alur, berbeda dari penelitian sebelumnya yang umumnya melakukan kedua proses tersebut secara terpisah. Sistem ini mengintegrasikan algoritma YOLO untuk lokalisasi komponen botol (badan botol, tutup, dan label), serta *Convolutional Autoencoder* (CAE) untuk deteksi anomali berbasis nilai *Mean Squared Error* (MSE) dari *reconstruction error* yang hasilnya diklasifikasikan berdasarkan nilai *threshold* yang ditetapkan. Hasil pengujian menunjukkan bahwa sistem berhasil mengklasifikasikan objek dengan tingkat akurasi sebesar 88,2% untuk badan botol, 89,1% untuk tutup, dan 84,5% untuk label. Sistem mencapai nilai *precision* yang sangat tinggi, yaitu 94,5%–98,5%, membuktikan keandalannya dalam menekan *False Positive*. Namun, nilai *recall* yang dihasilkan lebih rendah berkisar 79,2%–89,6% yang diakibatkan oleh kegagalan deteksi pada cacat berdimensi mikro dan area berkontras rendah.

Kata Kunci — *Deep Learning, YOLO, Convolutional Autoencoder, Sistem Inspeksi, Inspeksi Botol*

I. PENDAHULUAN

Kualitas produk merupakan aspek penting dalam industri manufaktur karena berpengaruh terhadap kepuasan pelanggan dan efisiensi produksi. Salah satu tahapan utama dalam menjaga kualitas tersebut adalah inspeksi visual pada proses *quality control*. Namun, inspeksi masih banyak dilakukan secara manual sehingga rentan terhadap kesalahan akibat faktor manusia, seperti kelelahan, subjektivitas, dan inkonsistensi, yang dapat menyebabkan produk cacat lolos ke tahap distribusi [1].

Pada industri minuman, cacat visual dapat terjadi pada berbagai komponen botol, seperti badan botol, tutup, dan label. Di sisi lain, jumlah data cacat relatif lebih sedikit dibandingkan data normal sehingga menjadi tantangan bagi metode *supervised learning* yang membutuhkan data cacat berlabel dalam jumlah besar. Oleh karena itu, pendekatan *anomaly detection* menjadi alternatif karena hanya memerlukan data normal selama proses pelatihan [2].

Perkembangan *deep learning* telah mendorong berbagai penelitian dalam bidang inspeksi visual otomatis. Pendekatan deteksi objek, seperti YOLO, terbukti efektif dalam mendeteksi cacat secara *real-time* dan cocok untuk implementasi di lini produksi industri [3][4]. Selain itu, pendekatan *anomaly detection* menjadi alternatif untuk mengatasi keterbatasan data cacat dengan mempelajari karakteristik data normal dan mengidentifikasi cacat berdasarkan perbedaan rekonstruksi [5].

Pada proses inspeksi di industri, beberapa komponen produk diperiksa secara terpisah dalam satu alur inspeksi. Akan tetapi, integrasi antara deteksi objek dan identifikasi anomali pada setiap komponen masih relatif terbatas. Kondisi tersebut mengindikasikan perlunya suatu pipeline inspeksi yang mampu melokalisasi objek sekaligus menganalisis kualitas secara otomatis pada setiap komponen produk.

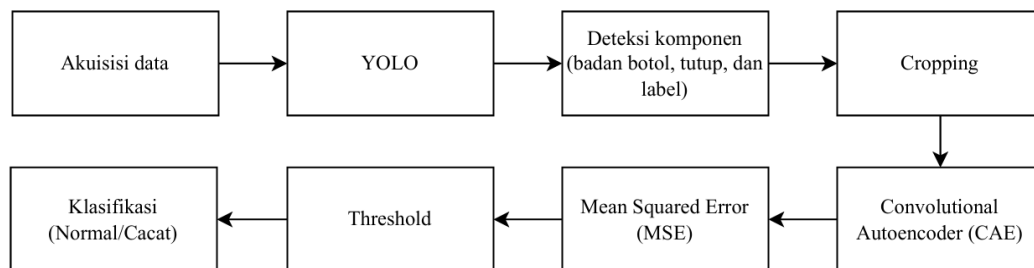
Berdasarkan permasalahan tersebut, penelitian ini mengusulkan *pipeline* inspeksi visual objek yang mengintegrasikan YOLO untuk mendeteksi dan melokalisasi komponen botol (badan botol, tutup, dan label), serta *Convolutional Autoencoder* (CAE) untuk mengidentifikasi cacat berdasarkan nilai *Mean Squared Error* (MSE). Penelitian ini juga mengevaluasi kemampuan *pipeline* dalam mengidentifikasi berbagai karakteristik cacat pada setiap komponen sehingga dapat diketahui kemampuan pendekatan *anomaly detection* pada inspeksi visual botol.

Kebaruan penelitian ini terletak pada integrasi antara tahap deteksi objek dan tahap analisis kualitas dalam satu *pipeline* inspeksi, di mana ketiga komponen botol (badan botol, tutup, dan label) masing-masing diinspeksi menggunakan model *Convolutional Autoencoder* (CAE) terpisah dan dikhususkan untuk karakteristik visual masing-masing komponen. Pendekatan ini berbeda dari penelitian sebelumnya yang hanya berfokus pada tahap lokalisasi objek [5][8], tanpa menggabungkan kedua tahap tersebut dalam satu alur kerja otomatis. Integrasi serupa juga pernah diterapkan pada objek berbeda, seperti pada inspeksi isolator saluran transmisi listrik [11], tetapi penerapan pada inspeksi komponen botol kemasan minuman dengan model CAE terpisah untuk setiap komponen masih terbatas. Dengan demikian, kontribusi utama penelitian ini adalah arsitektur *pipeline* gabungan YOLO dan CAE setiap komponen botol.

II. METODE PENELITIAN

A. Arsitektur Sistem

Pada Gambar 1 ditunjukkan diagram blok alur pemrosesan citra untuk deteksi objek dan identifikasi cacat menggunakan metode YOLO dan *Convolutional Autoencoder* (CAE).



Gambar 1. Diagram Blok Alur Pemrosesan Citra

Proses dimulai dengan *input* citra hasil akuisisi kamera yang diproses menggunakan model YOLO untuk mendeteksi dan melokalisasi komponen produk, yaitu badan botol, tutup, dan label. Hasil deteksi berupa *bounding box* pada setiap komponen yang selanjutnya digunakan untuk melakukan proses (*cropping*) sehingga setiap komponen dapat dianalisis secara terpisah. Tahap inspeksi ditentukan berdasarkan hasil deteksi YOLO, dengan kelas badan botol sebagai objek utama. Apabila ditemukan objek label di dalam area badan botol, maka proses masuk tahap inspeksi label. Jika label tidak ditemukan, tetapi terdapat objek tutup, maka botol diproses pada tahap inspeksi tutup. Sementara itu, apabila tidak ditemukan label maupun tutup, sistem hanya melakukan inspeksi pada bagian badan botol.

Citra hasil *cropping* kemudian diproses menggunakan tiga model *Convolutional Autoencoder* (CAE) yang masing-masing dikhususkan untuk inspeksi badan botol, tutup, dan label. Hasil rekonstruksi dari setiap model CAE dibandingkan dengan citra asli menggunakan nilai *Mean Squared Error* (MSE). Nilai MSE yang diperoleh selanjutnya dibandingkan dengan nilai *threshold* yang telah ditentukan untuk mengklasifikasikan setiap komponen ke dalam kondisi normal atau cacat. Melalui alur tersebut, sistem mampu melakukan inspeksi kualitas secara otomatis pada setiap komponen produk sesuai dengan tahapan inspeksi yang sedang berlangsung.

B. Akuisisi Data

Objek penelitian yang digunakan sebagai *dataset* dalam sistem ini adalah botol kemasan minuman Cimory berukuran 65 ml varian *blueberry*. Dataset terdiri atas citra badan botol, tutup, dan label yang diperoleh menggunakan *webcam*. Dataset tersebut digunakan sebagai data latih untuk model YOLO dan model *Convolutional Autoencoder* (CAE).

Model YOLO dilatih menggunakan total 215 citra yang terdiri atas sampel normal dan cacat, kemudian dibagi menjadi data *training*, *validation*, dan *testing*. Sementara itu, model CAE untuk badan botol menggunakan 627 citra, tutup menggunakan 701 citra, dan label menggunakan 705 citra. Masing-masing dataset CAE tersebut selanjutnya dibagi menjadi data *training* dan *validation*.

C. Deteksi Objek

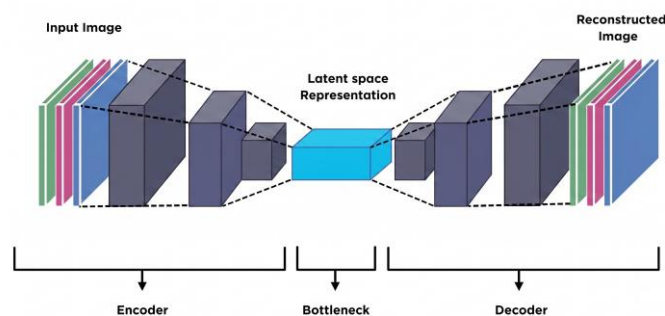
Deteksi objek dalam *computer vision* menggabungkan dua proses, yaitu klasifikasi dan lokalisasi. *Output* sistem deteksi objek umumnya berupa *bounding box* serta *confidence score* yang menunjukkan tingkat keyakinan model terhadap objek yang diprediksi [6]. *Bounding box* tersebut digunakan sebagai acuan dalam proses *cropping* sehingga setiap komponen dapat dianalisis secara individual oleh model *Convolutional Autoencoder*.

Pada penelitian ini, deteksi objek dilakukan menggunakan model YOLOv11 nano karena mampu mendeteksi objek secara cepat dengan kebutuhan komputasi yang rendah, sehingga sesuai untuk aplikasi inspeksi visual secara *real-time*. Model dilatih selama 300 *epoch* dengan ukuran citra masukan 640x640 piksel.

D. Identifikasi Anomali

Anomali merupakan data atau kejadian yang menyimpang dari pola umum sehingga dianggap tidak normal [7]. Salah satu metode *deep learning* untuk mendeteksi anomali adalah *Autoencoder*, yaitu jaringan saraf yang mempelajari representasi data melalui proses rekonstruksi untuk menghasilkan keluaran yang menyerupai data masukan [8].

Autoencoder terdiri atas tiga komponen utama, yaitu *encoder*, *bottleneck*, dan *decoder*. *Encoder* mengubah data masukan menjadi representasi berdimensi rendah, sedangkan *bottleneck* memadatkan informasi agar model mempelajari fitur-fitur penting dari pola normal data. Selanjutnya, *decoder* merekonstruksi representasi tersebut menjadi bentuk yang menyerupai data masukan. Tingkat kemiripan hasil rekonstruksi diukur menggunakan *reconstruction error*, yang banyak dimanfaatkan untuk deteksi cacat permukaan berbasis *unsupervised learning* [8].



Gambar 2. Arsitektur *Convolutional Autoencoder*

Gambar 2 menunjukkan arsitektur *Convolutional Autoencoder*. Citra masukan diproses melalui tiga lapisan konvolusi dengan aktivasi ReLU dan *MaxPooling* 2x2 untuk mengekstraksi serta mengompresi fitur ke dalam *bottleneck* (*latent space*). Selanjutnya, *decoder* merekonstruksi representasi laten menggunakan lapisan konvolusi dan *UpSampling2D* hingga menghasilkan citra dengan ukuran yang sama seperti citra masukan. Model CAE dilatih selama 250 *epoch* dengan *batch size* 32, *learning rate* 0,0005, fungsi *loss* MSE, dan menggunakan *optimizer* Adam.

Hasil deteksi anomali dinyatakan dalam nilai *Mean Squared Error* (MSE), di mana data anomali umumnya menghasilkan *reconstruction error* yang lebih tinggi dibandingkan data normal [9]. Nilai MSE tersebut kemudian dibandingkan dengan *threshold* untuk menentukan klasifikasi objek. Nilai *threshold* ditentukan secara empiris berdasarkan pengamatan nilai MSE yang dihasilkan dari pengujian citra botol normal secara *real-time*.

E. Evaluasi

Evaluasi model YOLO dilakukan untuk mengukur kemampuan model dalam mendeteksi dan melokalisasi objek pada citra. Metrik yang digunakan meliputi *precision*, *recall*, dan *mean Average Precision* (mAP). *Precision* menunjukkan ketepatan prediksi objek, *recall* mengukur kemampuan model mendeteksi seluruh objek yang sebenarnya, sedangkan mAP memberikan evaluasi menyeluruh terhadap performa deteksi pada berbagai ambang batas (*threshold*) [1].

Evaluasi model *Convolutional Autoencoder* (CAE) menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. *Accuracy* mengukur proporsi prediksi yang benar, *precision* menunjukkan ketepatan klasifikasi botol cacat, *recall*

mengukur kemampuan mendeteksi seluruh botol cacat, sedangkan *F1-score* digunakan untuk mengevaluasi keseimbangan antara *precision* dan *recall* [10].

III. HASIL DAN PEMBAHASAN

A. Dataset Pelatihan Model

Pengambilan *dataset* dilakukan dalam dua tahap. *Dataset* pertama digunakan untuk pelatihan model YOLO, terdiri atas citra yang memuat satu, dua, atau tiga objek dalam satu *frame*. Seluruh data kemudian dianotasi menggunakan Roboflow untuk membedakan tiga kelas objek yang digunakan dalam proses pelatihan YOLO, ditunjukkan pada Gambar 3 dan Gambar 4.



Gambar 3. Data Sebelum Dianotasi



Gambar 4. Data Setelah Dianotasi

Dataset kedua digunakan untuk pelatihan model *Convolutional Autoencoder* (CAE) dan hanya terdiri atas botol dalam kondisi normal, yaitu botol (dengan label dan tutup), botol tanpa label (dengan tutup), serta badan botol (tanpa label dan tanpa tutup). Sebelum digunakan untuk pelatihan, seluruh citra dipotong (*cropping*) berdasarkan *bounding box* hasil deteksi objek, seperti yang ditunjukkan pada Gambar 5, Gambar 6, dan Gambar 7.



Gambar 5. Data Pelatihan CAE Botol



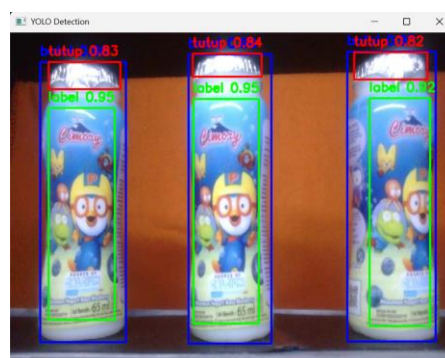
Gambar 6. Data Pelatihan CAE Label



Gambar 7. Data Pelatihan CAE Tutup

B. Evaluasi Performa Sistem

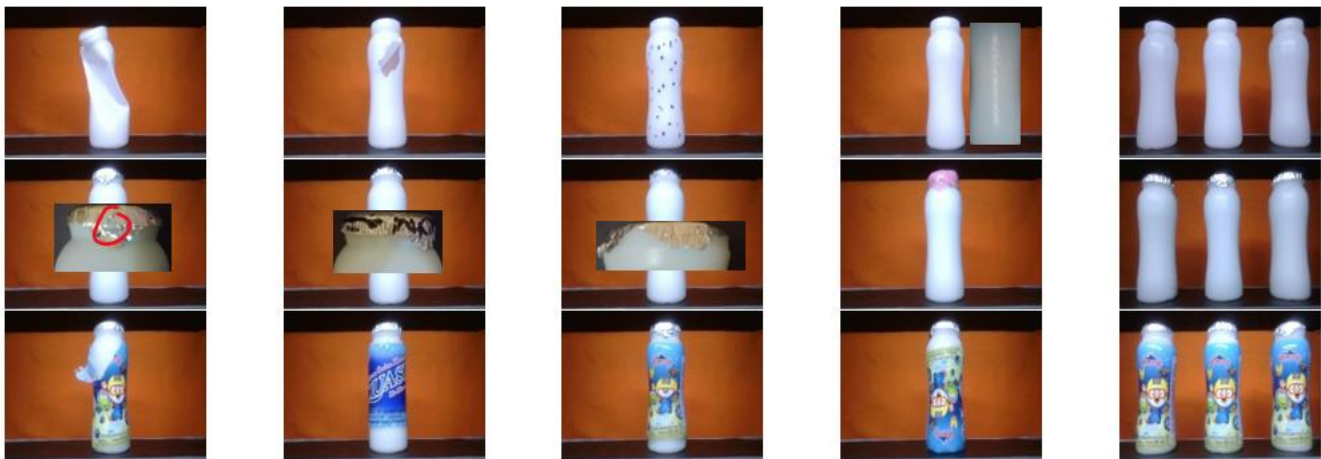
Model YOLO menunjukkan kemampuan deteksi dan lokalisasi objek yang akurat, sebagaimana ditunjukkan pada Gambar 8, dengan keberhasilan mencapai *precision* 99,6%, *recall* 99,7%, dan mAP@50 99,5%.



Gambar 8. Hasil Deteksi YOLO

Setelah model YOLO dipastikan mampu bekerja dengan baik, tahap selanjutnya adalah mengintegrasikan model tersebut dengan *Convolutional Autoencoder* (CAE) untuk melakukan analisis terhadap citra hasil *cropping* objek.

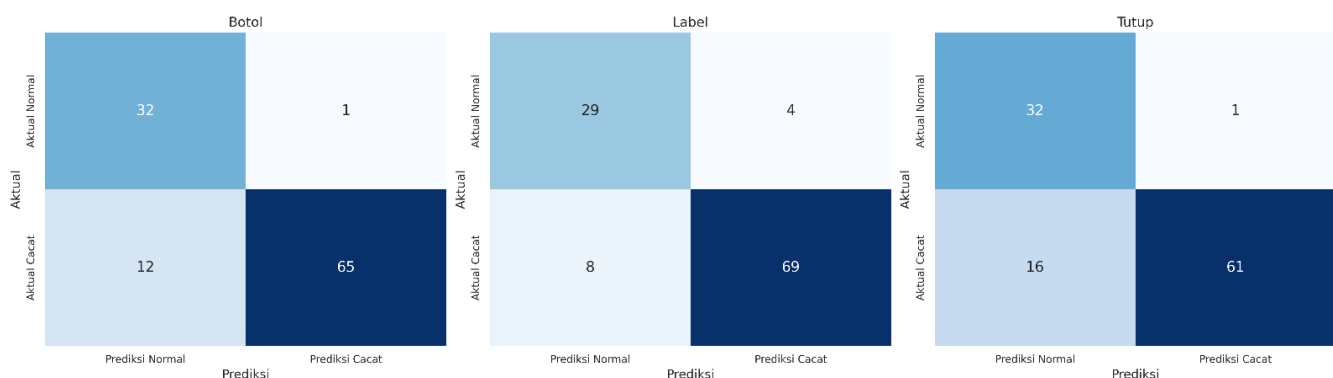
Pengujian dilakukan menggunakan 30 botol yang dibagi ke dalam tiga tahapan inspeksi, dengan masing-masing tahapan terdiri atas 10 botol, yaitu 3 botol normal dan 7 botol yang mewakili berbagai jenis cacat.



Gambar 9. Data Citra Pengujian Model

Gambar 9 menunjukkan contoh data yang digunakan dalam pengujian secara *real-time*. Secara berurutan terdapat badan botol penyok, sobek, berbintik, tergores, dan nor mal. Pada komponen tutup terdapat jenis cacat tutup berlubang, bernoda, sobek, berbeda warna, dan tutup normal. Selanjutnya, pada komponen label terdapat label sobek, label berbeda, terlipat, terbalik, dan label normal.

Selanjutnya, evaluasi deteksi anomali oleh CAE diukur dari perbandingan prediksi dengan kondisi aktual yang divisualisasikan dalam *confusion matrix* pada setiap komponen. Pengambilan keputusan cacat atau normal didasarkan pada distribusi nilai *Mean Squared Error* (MSE), di mana objek diklasifikasikan cacat apabila nilai MSE-nya melampaui *threshold* empiris yang ditetapkan: 0,0025 untuk badan botol, serta 0,0035 untuk tutup dan label. Hasil klasifikasi kemudian diringkas secara kuantitatif pada *confusion matrix*.



Gambar 10. *Confusion Matrix*

Gambar 12 menunjukkan hasil *confusion matrix* pada badan botol, label, dan tutup, di mana setiap komponen dilakukan pengujian sebanyak 110 kali. Pada badan botol, model berhasil mengklasifikasikan 32 badan botol normal sebagai normal (*True Negative*), 65 badan botol cacat sebagai cacat (*True Positive*), 1 badan botol normal sebagai cacat (*False Positive*), dan 12 badan botol cacat sebagai normal (*False Negative*). Selanjutnya pada label, model berhasil mengklasifikasikan 29 label normal sebagai normal (*True Negative*), 69 label cacat sebagai cacat (*True Positive*), 4 label normal sebagai cacat (*False Positive*), dan 8 label cacat sebagai normal (*False Negative*).

Kemudian pada tutup, model berhasil mengklasifikasikan 32 tutup normal sebagai normal (*True Negative*), 61 tutup cacat sebagai cacat (*True Positive*), 1 tutup normal sebagai cacat (*False Positive*), dan 16 tutup cacat sebagai normal (*False Negative*). Berdasarkan *confusion matrix* tersebut, dapat diperoleh nilai metrik evaluasi yang meliputi *accuracy*, *precision*, *recall*, dan *F1-score* untuk setiap komponen. Hasil evaluasi tersebut ditunjukkan pada Tabel 1.

Tabel 1. Metrik Evaluasi *Convolutional Autoencoder*

Komponen	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Badan Botol	88,2%	98,5%	84,4%	90,9 %
Tutup	84,5%	98,4%	79,2%	87,8%
Label	89,1%	94,5%	89,6%	92,0%

Model pada botol memperoleh akurasi sebesar 88,2%, *precision* 98,5%, *recall* 84,4%, dan *F1-score* 90,9%. Pada tutup diperoleh akurasi 84,5%, *precision* 98,4%, *recall* 79,2%, dan *F1-score* 87,8%. Sementara itu, pada label diperoleh akurasi 89,1%, *precision* 94,5%, *recall* 89,6%, dan *F1-score* 92,0%.

Nilai *precision* yang tinggi pada ketiga komponen menunjukkan bahwa sebagian besar objek yang diprediksi sebagai cacat memang merupakan objek cacat, sehingga kesalahan klasifikasi berupa *false positive* relatif rendah. Di sisi lain, nilai *recall* yang berkisar antara 79,2%–89,6% menunjukkan bahwa sebagian besar objek cacat berhasil dideteksi, meskipun masih terdapat beberapa objek cacat yang terklasifikasi sebagai normal (*false negative*).

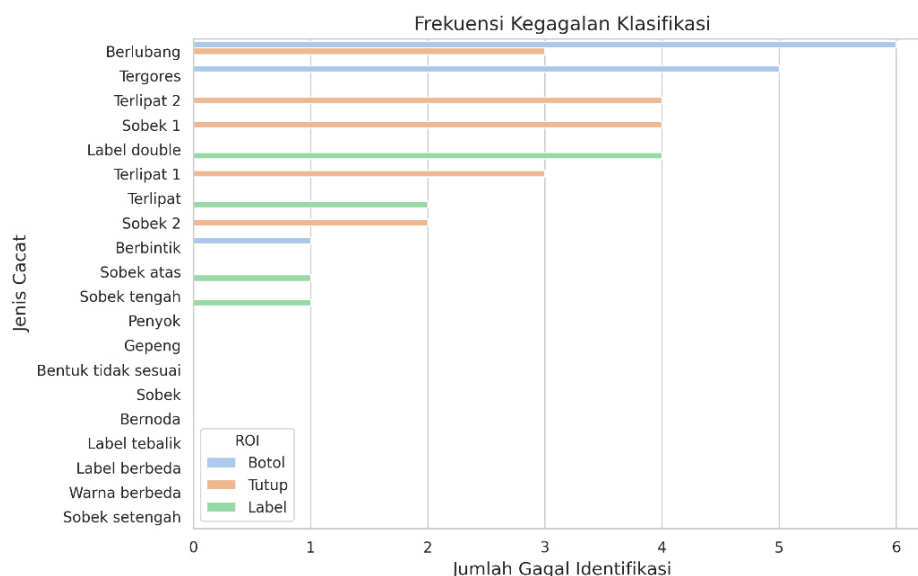
Keseimbangan antara *precision* dan *recall* ditunjukkan pada nilai *F1-score* yang berada pada rentang 87,8%–92,0%, sehingga menunjukkan bahwa model memiliki kemampuan klasifikasi yang baik tanpa terlalu menurunkan kinerja salah satu metrik.

Rendahnya nilai *recall* yang dihasilkan berkaitan dengan proses penentuan *threshold*. Pada awalnya, *threshold* ditentukan secara statistik menggunakan persentil ke-95 dari distribusi normal nilai MSE pada data *offline*. Namun, pendekatan tersebut tidak dapat diterapkan karena distribusi nilai MSE pada pengujian *real-time* berbeda dari data *offline*. Hal tersebut berkaitan dengan *domain shift*, yaitu perbedaan karakteristik data antara kondisi pelatihan dan pengujian yang dapat mempengaruhi kinerja *anomaly detection* pada kondisi nyata [12].

Pada penelitian ini, meskipun jarak dan posisi kamera telah dibuat konsisten serta area pengambilan citra diberi penutup untuk mengurangi pengaruh cahaya luar, variasi cahaya dan pergerakan objek pada kondisi *real-time* masih sulit dikendalikan sepenuhnya. Oleh karena itu, *threshold* akhir ditentukan berdasarkan pengamatan pada kondisi *real-time*. Namun, metode ini cenderung menghasilkan nilai yang konservatif, sehingga menghasilkan *precision* yang tinggi tetapi *recall* relatif lebih rendah. Kondisi ini perlu diperhatikan, karena dalam inspeksi kualitas, *false negative* berpotensi menyebabkan produk cacat lolos dari proses inspeksi.

C. Kegagalan Klasifikasi

Berdasarkan evaluasi performa sistem menunjukkan tingkat konsistensi yang fluktuatif dalam mengidentifikasi kualitas botol. Gambar 11 menunjukkan grafik frekuensi kegagalan klasifikasi yang menggambarkan seberapa sering sistem mengalami kesalahan dalam proses klasifikasi pada setiap pengujian.



Gambar 11. Frekuensi Kegagalan Klasifikasi Sistem

Berdasarkan grafik tersebut, model mampu mengklasifikasi jenis cacat yang menyebabkan perubahan struktur makro objek secara signifikan, seperti penyok, label terbalik, dan perbedaan warna, dengan akurasi mencapai 100%.

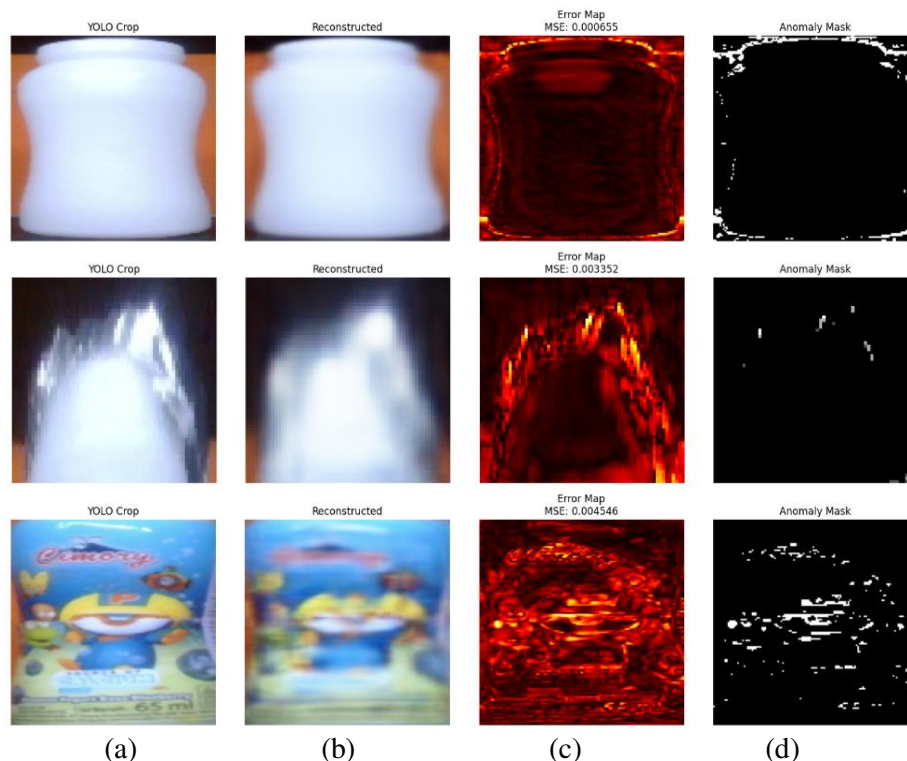
Sebaliknya, frekuensi kegagalan klasifikasi paling tinggi terjadi pada jenis cacat berdimensi mikro, seperti berlubang, tergores, dan tutup terlipat.

Kondisi tersebut disebabkan oleh model CAE mempelajari karakteristik citra normal dengan mengubahnya menjadi representasi yang lebih rendah (*latent space*), kemudian merekonstruksi kembali citra tersebut. Ketika objek dalam kondisi cacat yang mengakibatkan perubahan visual secara signifikan, representasi fitur citra berada di luar distribusi data normal yang telah dipelajari model. Akibatnya, model tidak mampu merekonstruksi citra dengan baik dan menghasilkan nilai MSE yang tinggi serta melebihi nilai *threshold*, sehingga objek dapat diklasifikasikan sebagai cacat.

Sebaliknya, pada cacat dengan perubahan visual yang relatif kecil seperti goresan pada badan botol, karakteristik anomali masih memiliki kemiripan dengan variasi citra normal, misalnya akibat pantulan cahaya, tekstur permukaan botol, atau garis kontur botol. Kemiripan tersebut menyebabkan model menganggapnya sebagai fitur normal dan tetap mampu merekonstruksi area yang mengalami goresan dengan baik. Hal tersebut menyebabkan selisih kesalahan menjadi kecil, nilai MSE tetap berada di bawah *threshold*, dan objek lolos diklasifikasi sebagai normal.

Fenomena serupa juga terjadi pada tutup. Jenis cacat seperti sobek dan terlipat cukup sering terdeteksi sebagai normal karena perbedaan karakteristik piksel antara area cacat dan kondisi normal relatif kecil. Warna tutup berbahan foil berwarna perak yang berdekatan dengan badan botol berwarna putih susu menghasilkan kontras yang rendah, sehingga perubahan akibat cacat tidak memberikan peningkatan nilai MSE yang signifikan. Dalam beberapa kasus, nilai MSE pada tutup yang mengalami cacat bahkan lebih rendah dibandingkan variasi citra normal, sehingga model gagal mengidentifikasinya sebagai anomali.

Visualisasi tersebut dapat dilihat pada Gambar 12 yang menampilkan alur inspeksi, yaitu hasil pemotongan (*cropping*) oleh YOLO sebagai citra *input*, citra hasil rekonstruksi CAE, *heatmap* yang menunjukkan gradasi selisih intensitas piksel, serta *anomaly mask* yang melokalisasi area anomali secara spesifik.



Gambar 12. (a) *Input* CAE, (b) Hasil Rekonstruksi, (c) *Anomaly Heatmap*, (d) *Anomaly Mask*

Baris pertama menunjukkan objek dengan cacat gores pada badan botol. Citra hasil rekonstruksi hampir identik dengan citra masukan karena model CAE masih mampu merekonstruksi goresan halus tersebut sebagai bagian dari karakteristik citra normal. Hal ini terlihat pada visualisasi *heatmap* dan *anomaly mask* yang hanya menampilkan area berintensitas rendah tanpa dominasi warna merah sebagai indikator anomali. Perbedaan spasial yang sangat kecil tersebut menyebabkan nilai *Mean Squared Error* (MSE) keseluruhan tidak cukup besar untuk melampaui nilai *threshold*, sehingga objek diklasifikasikan sebagai normal.

Baris kedua menunjukkan objek dengan cacat tutup terlipat pada tutup. Pada kasus ini, model mengalami kesulitan membedakan lipatan akibat rendahnya kontras pada permukaan tutup berbahan foil. Lipatan masih dapat direkonstruksi cukup baik sehingga *heatmap* hanya menampilkan sedikit perbedaan, terutama pada area pantulan cahaya di sekitar lipatan. Akibatnya, luas area yang teridentifikasi sebagai anomali pada *anomaly mask* relatif kecil, sehingga nilai MSE yang dihasilkan tidak cukup tinggi dan model sering gagal mengidentifikasinya sebagai cacat.

Baris ketiga menunjukkan objek dengan cacat label terlipat pada komponen label. Berbeda dengan dua kasus sebelumnya, area label memiliki pola, warna, dan elemen grafis yang lebih kompleks. Ketika label mengalami lipatan, terjadi perubahan struktur piksel yang cukup signifikan dibandingkan dengan representasi citra normal yang telah dipelajari model. Akibatnya, *heatmap* menampilkan area dengan intensitas perbedaan yang lebih tinggi dan lebih luas. Perbedaan tersebut kemudian ditandai secara jelas pada *anomaly mask*, sehingga menghasilkan nilai MSE yang lebih besar dan memungkinkan model mendeteksi anomali dengan lebih akurat.

IV. KESIMPULAN

Sistem inspeksi visual berhasil dibangun menggunakan *pipeline* gabungan, di mana YOLO memotong (*cropping*) objek sesuai komponen, yaitu badan botol, tutup, dan label secara otomatis, serta *Convolutional Autoencoder* (CAE) mengidentifikasi cacat berdasarkan selisih rekonstruksi piksel (*Mean Squared Error*). Keputusan anomali (cacat) didasarkan pada *threshold* yang telah ditentukan, yaitu 0,0025 untuk badan botol serta 0,0035 untuk tutup dan label. Kinerja sistem menunjukkan tingkat *precision* yang sangat tinggi berkisar 94,5% hingga 98,5%, membuktikan model andal dalam menekan *false positive*. Akurasi keseluruhan tercatat sebesar 88,2% untuk badan botol, 89,1% untuk tutup, dan 84,5% untuk label. Namun, nilai *recall* cenderung lebih rendah, yaitu 79,2% hingga 89,6% yang diakibatkan oleh tingginya kegagalan deteksi (*false negative*) pada cacat berdimensi mikro dan cacat di area berkontras rendah. Penelitian selanjutnya disarankan untuk melakukan kalibrasi *threshold* menggunakan data yang diperoleh langsung dari kondisi operasional *real-time* serta mempertimbangkan penggunaan kamera dengan kualitas dan stabilitas yang lebih baik. Upaya tersebut diharapkan dapat mengurangi kesenjangan antara kondisi pelatihan dan implementasi sehingga *recall* dapat ditingkatkan tanpa menurunkan *precision* yang telah dicapai.

V. DAFTAR PUSTAKA

- [1] R. Kühlechner, "Object Detection Survey for Industrial Applications with Focus on Quality Control," *Production Engineering*, vol. 19, no. 6, pp. 1271-1291, 2025.
- [2] S. Theodoropoulos, P. Zajec, J. M. Rožanec, D. Kyriazis, and P. Tsanakas, "On-the-fly Image-level Oversampling for Imbalanced Datasets of Manufacturing Defects," *Machine Learning*, vol. 113, no. 7, pp. 4013-4035, 2024.
- [3] T.-T.-H. Vu, D.-L. Pham, and T.-W. Chang, "A YOLO-based Real-time Packaging Defect Detection System," *Procedia Computer Science*, vol. 217, pp. 886-894, 2023.
- [4] M. Jamal, S. Faisal, D.S. Kusumaningrum, and T. Rohana, "Application of YOLOv8 for Product Defect Detection in Manufacturing Companies," *Jurnal Sistem Informasi dan Ilmu Komputer Prima*, vol. 8, no. 1, pp. 1-11, 2024.
- [5] C. Hu, J. Yao, W. Wu, W. Qiu, and L. Zhu, "A Lightweight Reconstruction Network for Surface Defect Inspection," *IEEE*, pp. 43-50, 2023.
- [6] M.L. Ali, and Z. Zhang, "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection," *Computers*, vol. 13, no. 336, pp. 1-37, 2024.
- [7] R. Foorthuis, "On the Nature and Types of Anomalies: A Review of Deviations in Data," *International Journal of Data Science and Analytics*, vol. 12, pp. 297-331, 2021.
- [8] T.G. Shiferaw, and L. Yao, "Autoencoder-Based Unsupervised Surface Defect Detection Using Two-Stage Training," *Journal of Imaging*, vol. 10, no. 111, pp. 1-17, 2024.
- [9] M. Thill, W. Konen, H. Wang, and T. Bäck, "Temporal Convolutional Autoencoder for Unsupervised Anomaly Detection in Time Series," *Applied Soft Computing*, vol. 112, 2021.
- [10] Z. Li, Y. Yan, X. Wang, Y. Ge, and L. Meng, "A Survey of Deep Learning for Industrial Visual Anomaly Detection," *Artificial Intelligence Review*, vol. 58, no. 9, pp. 1-82, 2025.
- [11] M. Allam, K. Shaban, and A. Hamdi, "Attention-Guided Autoencoder Fusion for Insulator Defect Detection Using UAV Transmission-Line Imaging," *arXiv preprint arXiv:2606.06536*, 2026.
- [12] Z. Zhang, Z. Zhao, X. Zhang, C. Sun, and X. Chen, "Industrial Anomaly Detection with Domain Shift: A Real-World Dataset and Masked Multi-Scale Reconstruction," in *Computers in Industry*, vol. 151, p. 103990, 2023.