

Implementasi *Mask R-CNN* untuk Identifikasi Cacat pada Pengelasan *Shielded Metal Arc Welding* (SMAW)

Mubessirul Ummah¹, Agus Sasmito², Rima Tri Wahyuningrum³, Mohamad Imron Mustajib⁴,
Nandiko Rizal⁵, Irfan Eko Sandjaja⁶, Totok Triputrastyo Murwatono⁷, Andik Machfudin⁸,
Dian Purnama Sari⁹

^{1,2,5,6,7,8,9}Kelompok Riset Teknologi Rancang Bangun Perkapalan dan Bangunan Apung
Badan Riset dan Inovasi Nasional

Surabaya, Indonesia

^{1,3,4}Fakultas Teknik

Universitas Trunojoyo Madura

Bangkalan, Indonesia

e-mail: ¹mubarijojo.ummah11@gmail.com, ²agus107@brin.go.id,

³rimatriwahyuningrum@trunojoyo.ac.id, ⁴imronmustajib@trunojoyo.ac.id, ⁵nand006@brin.go.id,

⁶irfa002@brin.go.id, ⁷toto002@brin.go.id, ⁸andi014@brin.go.id, ⁹dian010@brin.go.id

Correspondence : e-mail: mubarijojo.ummah11@gmail.com

Diajukan: 06 Agustus 2024; Direvisi: 19 Agustus 2024; Diterima: 22 Agustus 2024

Abstrak

Penelitian ini mengembangkan metode pemrosesan citra digital untuk inspeksi visual cacat pengelasan baja, khususnya untuk mendeteksi cacat spatter dan slag inclusion. Penelitian ini menggunakan algoritma *Mask R-CNN* untuk melakukan deteksi dan segmentasi cacat las secara akurat pada dataset pengelasan busur listrik. Dataset yang digunakan terdiri dari 176 gambar spesimen hasil pengelasan SMAW, yang diambil di Politeknik Perkapalan Negeri Surabaya. Dataset diolah menggunakan platform Roboflow untuk pelabelan, prapemrosesan, dan augmentasi data. Setelah itu, model dilatih dengan menggunakan algoritma *Mask R-CNN*. Model ini kemudian diterapkan dalam aplikasi Streamlit untuk pengecekan cacat secara interaktif. Hasil eksperimen menunjukkan bahwa *Mask R-CNN* mampu mendeteksi cacat dengan akurasi yang bervariasi, diukur dengan metrik mean Average Precision (mAP) dan Average Recall (AR). Pada resolusi gambar 640×640 piksel, model mencapai mAP sebesar 65,8% untuk deteksi bounding box dan 70,2% untuk mask segmentasi pada threshold AP@0,5.

Kata kunci: *Mask R-CNN, Roboflow, Cacat Las, Deteksi Objek, Segmentasi.*

Abstract

This study develops a digital image processing method for the visual inspection of steel welding defects, specifically focusing on detecting spatter and slag inclusion defects. The research utilizes the *Mask R-CNN* algorithm for accurate defect detection and segmentation on an electric arc welding dataset. The dataset comprises 176 images of SMAW weld specimens, collected at the Politeknik Perkapalan Negeri Surabaya. The dataset is processed using the Roboflow platform for labeling, preprocessing, and data augmentation. Subsequently, the model is trained using the *Mask R-CNN* algorithm and integrated into a Streamlit application for interactive defect checking. Experimental results demonstrate that *Mask R-CNN* is capable of detecting defects with varying accuracy, measured by mean Average Precision (mAP) and Average Recall (AR) metrics. At an image resolution of 640×640 pixels, the model achieves an mAP of 65.8% for bounding box detection and 70.2% for mask segmentation at an AP@0.5 threshold.

Keywords: *Mask R-CNN, Roboflow, Welding Defects, Object Detection, Segmentation.*

1. Pendahuluan

Pengelasan adalah proses menyambungkan dua atau lebih bahan logam melalui pemanasan hingga mencapai titik leleh kemudian diikuti dengan pendinginan sehingga terbentuk sambungan yang kuat antara bahan-bahan tersebut [1]. Salah satu teknik pengelasan yang banyak digunakan dalam industri adalah *Shielded Metal Arc Welding* (SMAW). Pengelasan SMAW adalah teknik yang menggunakan busur listrik

dan elektroda berlapis fluks sebagai bahan pengisi, menawarkan kontrol yang lebih baik dan cocok untuk berbagai jenis logam serta posisi pengelasan dibandingkan dengan teknik pengelasan lainnya [2].

Namun, cacat pada hasil pengelasan tetap menjadi tantangan yang cukup berarti meskipun teknologi pengelasan telah berkembang pesat. Cacat pengelasan adalah ketidaksempurnaan yang dapat mengurangi kualitas dan kekuatan sambungan las. Selain itu dapat menyebabkan ketidaksesuaian terhadap standar industri seperti *ASME IX*, *AWS*, *API*, dan *ASTM*. Penyebab cacat ini bisa berasal dari prosedur pengelasan yang tidak tepat, persiapan yang kurang memadai, serta penggunaan peralatan dan bahan habis pakai yang tidak sesuai standar. Cacat pengelasan terbagi menjadi cacat internal dan cacat visual di mana cacat internal berada di dalam material las dan memerlukan pengujian khusus untuk mendeteksinya, sedangkan cacat visual dapat langsung terlihat pada permukaan [3].

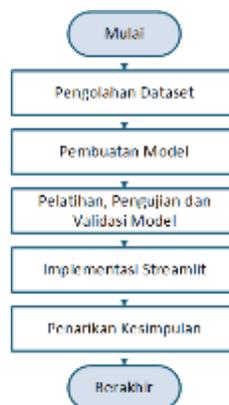
Dalam industri, deteksi dan evaluasi cacat las dilakukan melalui dua metode utama, yakni uji *Destructive Testing (DT)* dan *Non-Destructive Testing (NDT)*. *Destructive Testing* menguji kekuatan sambungan las dengan cara yang merusak spesimen dengan memberikan informasi langsung mengenai batasan tegangan-regangan maksimal. Namun, metode ini tidak cocok untuk produk yang siap didistribusikan karena proses pengujian merusak spesimen. Sebaliknya, *Non-Destructive Testing* digunakan untuk mendeteksi cacat tanpa merusak material dan menjadikannya sangat penting dalam proses pengelasan tradisional seperti *Shielded Metal Arc Welding (SMAW)* [1], [4].

Meskipun efektif, teknologi *NDT* ini sering kali membutuhkan waktu dan biaya yang cukup banyak, terutama ketika dihadapkan dengan banyaknya objek uji. Hal ini meningkatkan risiko kesalahan dan kelalaian, sehingga memerlukan pekerjaan tambahan. Salah satu solusi yang muncul adalah implementasi otomatisasi dalam proses deteksi cacat yang dapat mengurangi kesalahan manusia dan mempercepat analisis. *Deep Learning* adalah cabang dari *Machine Learning* telah terbukti efektif dalam berbagai aplikasi inspeksi visual otomatis melalui pemrosesan jaringan saraf tiruan yang dapat mempelajari pola dari data dalam jumlah besar [5]. Salah satu algoritma *Deep Learning* yang banyak digunakan untuk deteksi dan segmentasi objek adalah *Mask R-CNN*. *Mask R-CNN* merupakan perluasan dari *Faster R-CNN* dan dirancang untuk mendeteksi serta membatasi lokasi objek dalam gambar sekaligus menghasilkan mask segmentasi yang menunjukkan masing-masing objek [6].

Penelitian ini mengimplementasikan *Mask R-CNN* untuk identifikasi cacat pada pengelasan *Shielded Metal Arc Welding (SMAW)*, khususnya untuk mendeteksi cacat seperti *spatter* dan *slag inclusion*. *Slag inclusion* adalah ketidaksempurnaan pengelasan yang terjadi ketika partikel *slag (terak)* terjebak dalam logam las sedangkan *Spatter* adalah percikan kecil logam cair yang terlempar dari busur las dan menempel pada permukaan sekitar area las. Dengan menggunakan *dataset* pengelasan busur listrik yang telah dilabeli, diproses, dan diaugmentasi menggunakan platform *Roboflow*, algoritma *Mask R-CNN* dilatih dan diuji untuk mendeteksi serta melakukan segmentasi cacat las secara akurat. Model ini kemudian diterapkan dalam aplikasi *Streamlit* untuk pengecekan cacat secara interaktif serta diharapkan dapat meningkatkan efisiensi dan akurasi proses inspeksi las dalam industri, mengurangi biaya dan waktu.

2. Metode Penelitian

Metode penelitian yang digunakan terdiri dari 5 tahapan yaitu pengolahan *dataset*, pembuatan model deteksi, pelatihan pengujian dan validasi model, dan yang terakhir implementasi *streamlit*. Berdasarkan 5 tahapan itu peneliti akan memperjelas alur metodologi penelitian yang digunakan melalui Gambar 1.



Gambar 1 : Metode Penelitian

A. Pengolahan Dataset

Penelitian ini, tahap pertama adalah pengolahan *dataset* yang dimulai dengan akuisisi citra menggunakan kamera ponsel *Redmi Note 11* untuk menangkap gambar spesimen hasil pengelasan SMAW di Politeknik Perkapalan Negeri Surabaya. Pada awalnya, diperoleh 20 gambar yang kemudian dilabeli bersama inspektur las untuk menandai bagian yang mengalami cacat, seperti *spatter* dan *slag inclusion*, menggunakan warna berbeda. Proses pelabelan dilakukan dengan menandai cacat-cacat tersebut menggunakan poligon yang mengikuti kontur objek, dan hasilnya disimpan dalam format *JSON* menggunakan platform *Roboflow*.

Memastikan konsistensi data, gambar-gambar tersebut kemudian diproses melalui tahap *pre-processing* (*pre-processing*), yang meliputi *auto-orient* untuk memastikan orientasi gambar sesuai, perubahan ukuran menjadi 640x640 piksel, serta pemotongan gambar menjadi beberapa bagian atau *“tile”* yang terdiri dari 2 baris dan 2 kolom. Selanjutnya, untuk menghadapi keterbatasan jumlah gambar, maka dilakukan augmentasi data untuk meningkatkan kuantitas dan variasi gambar dalam *dataset*. Teknik augmentasi yang diterapkan meliputi *flip*, rotasi 90°, *shear*, konversi ke skala abu-abu, serta penyesuaian saturasi, kecerahan, eksposur, dan blur.

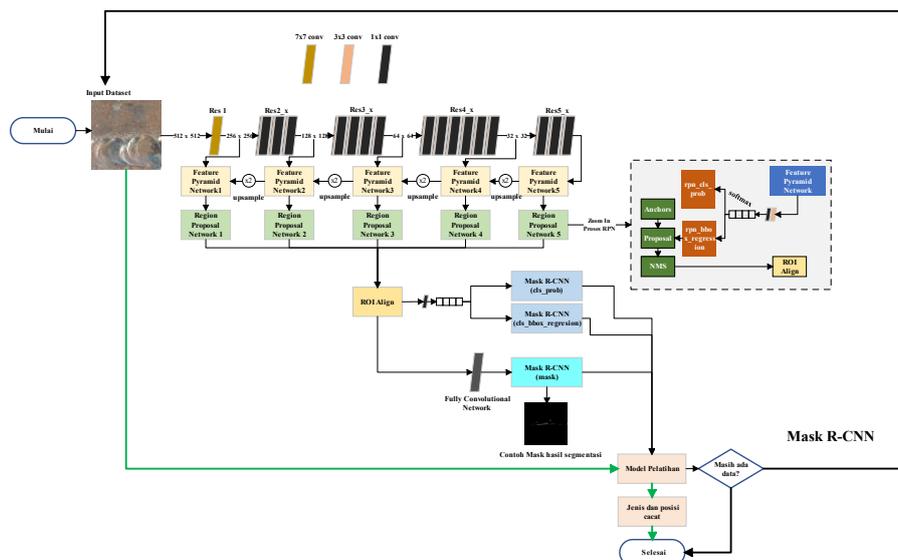
B. Pembuatan model deteksi

Pada tahap ini, dilakukan konfigurasi parameter yang digunakan dalam algoritma *Mask R-CNN* untuk mempersiapkan model deteksi cacat pengelasan. Konfigurasi ini mencakup pengaturan berbagai parameter yang mempengaruhi kinerja dan akurasi model. Pengaturan parameter tersebut dapat dilihat pada Tabel 1.

Tabel 1 : Konfigurasi Model *Mask R-CNN*

Parameter	Nilai
ARCHITECTURE	Resnet-101
MAX_ITER	10000
EVAL_PERIOD	300
BASE_LR	0,01
NUM_CLASSES	3
cfg.MODEL.ROI_HEADS.BATCH_SIZE_PER_IMAGE	64
cfg.SOLVER.IMS_PER_BATCH	2
cfg.DATALOADER.NUM_WORKERS	2
cfg.INPUT.MASK_FORMAT	bitmask
cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST	0,7
cfg.DATASETS.TRAIN	TRAIN_DATA_SET_NAME,
cfg.DATASETS.TEST	TEST_DATA_SET_NAME,

C. Pelatihan, Pengujian dan Validasi Model



Gambar 2 : Proses *Training* dan *Testing* Model *Mask R-CNN*

Proses pelatihan model menggunakan data latih bertujuan untuk membuat model dapat mendeteksi objek secara akurat. Pelatihan berlangsung selama MAX_ITER iterasi, dengan *hyperparameter* dan konfigurasi yang telah ditentukan untuk memastikan pelatihan yang optimal. Kinerja model dievaluasi secara berkala setiap $EVAL_PERIOD$ untuk memantau kemajuan selama pelatihan. Gambar 2 menunjukkan proses pelatihan dan pengujian *Mask R-CNN*. Panah hitam mewakili pelatihan, sedangkan panah hijau menunjukkan pengujian *dataset*.

Langkah pertama dalam pelatihan adalah memasukkan kumpulan data pelatihan gambar cacat pengelasan ke dalam model untuk diproses melalui *backbone ResNet-101*. *Backbone* ini berfungsi untuk mengekstrak fitur-fitur dasar dari gambar tersebut melalui serangkaian lapisan konvolusi yang mendalam. Fitur-fitur yang diekstraksi dari *backbone* kemudian diolah oleh *Feature Pyramid Network (FPN)*, yang berfungsi untuk mendeteksi objek dengan berbagai ukuran. *FPN* disini memastikan bahwa deteksi objek tetap konsisten meskipun objek berukuran kecil atau besar, dengan membangun hirarki fitur dari berbagai level resolusi [7].

Setelah fitur-fitur diekstraksi, gambar melalui *Region Proposal Network (RPN)*, yang menggunakan *anchor boxes* untuk memprediksi apakah area gambar (*anchor*) mengandung objek dan untuk menghitung *bounding box* yang paling sesuai dengan objek tersebut. *RPN* melakukan tugas ini dengan dua tahap utama, yakni klasifikasi untuk memutuskan apakah suatu *anchor* mengandung objek, dan regresi untuk menyempurnakan ukuran dan posisi *bounding box* [8]. Dalam proses ini, *Non-Maximum Suppression (NMS)* digunakan untuk menghilangkan *anchor boxes* yang tumpang tindih dengan menyisakan hanya kotak yang paling mungkin mengandung objek [9]. *RPN* memilih *anchor box* positif jika *Intersection over Union (IoU)* di atas 0,7 dan negatif jika di bawah 0,3 sesuai *hyperparameter* yang sudah diatur sebelumnya.

Setelah *RPN* mengusulkan kotak-kotak pembatas, *RoI Align* digunakan untuk mengambil fitur dari setiap kotak pembatas yang diusulkan. Tidak seperti metode sebelumnya yang menggunakan *RoI Pooling*, *RoI Align* menggunakan *interpolasi bilinear* untuk mengambil fitur dengan resolusi yang konsisten tanpa mengikis informasi spasial penting. Hal ini memungkinkan *Mask R-CNN* untuk mempertahankan lebih banyak detail dari gambar asli yang penting untuk segmentasi objek yang presisi. Fitur yang diekstrak kemudian diproses melalui *fully connected layer* untuk menghasilkan representasi akhir dari objek. *Mask R-CNN* memiliki tiga cabang output yang berfungsi untuk tiga tugas berbeda, yaitu klasifikasi untuk menentukan kelas dari objek yang dideteksi, lalu deteksi untuk menghasilkan koordinat kotak pembatas, dan segmentasi menggunakan *fully convolutional network* untuk menghasilkan *mask biner* dari objek dalam gambar [9].

Akhirnya, *Mask R-CNN* menghasilkan kotak-kotak pembatas, kelas objek, dan *mask* segmentasi untuk setiap objek yang dideteksi dalam gambar. Proses ini diulang hingga *dataset* dan jumlah iterasi pelatihan mencapai batas yang ditentukan. Pengujian dimulai dengan data *testing*, di mana model memprediksi kelas, *bounding box*, dan *mask segmentasi* tanpa melalui pelatihan kembali.

Model yang dilatih kemudian dievaluasi untuk mengetahui kinerjanya terhadap data yang tidak termasuk dalam *training set*. Evaluasi performa menggunakan metrik *Mean Average Precision (mAP)* dan *Average Recall (AR)*. *mAP* mengukur kemampuan model mendeteksi objek pada berbagai tingkat presisi, dengan menghitung *Average Precision (AP)* untuk setiap kelas. *AR* memberikan informasi tentang sejauh mana model mendeteksi *instance* objek dalam *dataset* [6]. Untuk menghitung *mAP* dapat menggunakan persamaan 1, dan untuk menghitung *AR* dapat menggunakan persamaan 2.

$$mAP@α = \frac{1}{n} \sum_{i=1}^n AP_i \quad \text{for } n \text{ class} \quad (1)$$

$$AR = 2 \int_{0.5}^1 recall(o) do \quad (2)$$

3. Hasil dan Pembahasan

Pada bagian ini dibahas hasil dari pengolahan *dataset*, serta hasil uji coba dan evaluasi model. Penjelasannya disajikan dalam beberapa poin berikut.

3.1. Hasil Pengolahan *Dataset*

Hasil dari proses pengolahan data ini adalah peningkatan jumlah *dataset* dari 20 menjadi 176 gambar, yang kemudian dibagi menjadi set pelatihan 140 gambar (80%), pengujian 18 gambar (10%), dan validasi 18 gambar (10%). Pengolahan *dataset* ini khususnya bagian augmentasi telah berhasil

meningkatkan ukuran *dataset* dan memperkenalkan variasi yang diperlukan untuk pelatihan model. Namun, perlu diperhatikan bahwa meskipun augmentasi menambah jumlah gambar, teknik ini tidak sepenuhnya menggantikan kebutuhan akan *dataset* yang lebih besar dan lebih beragam. Jumlah gambar awal yang terbatas tetap menjadi faktor yang dapat mempengaruhi kemampuan model menggeneralisasi pada data yang baru atau berbeda dari yang digunakan dalam pelatihan.

Tabel 2 : Contoh Sampel

Gambar	Keterangan
	Ukuran gambar 640 x 640 dengan 4 kelas spatter dan 0 kelas slag inclusion
	Ukuran gambar 640 x 640 dengan 0 kelas spatter dan 2 kelas slag inclusion
	Ukuran gambar 640 x 640 dengan 7 kelas spatter dan 5 kelas slag inclusion

Pada tabel 2 terlihat bahwa dalam satu *frame* gambar dapat terdapat banyak poligon yang mengelilingi berbagai objek, yang mana di setiap objek tersebut bisa merupakan kelas yang berbeda. Cacat dengan jenis *spatter* ditandai dengan poligon berwarna merah, sedangkan cacat dengan jenis *slag inclusion* ditandai dengan poligon berwarna biru.

3.2. Hasil Pelatihan Evaluasi Model Deteksi

Penelitian ini memanfaatkan *Mask R-CNN* yakni sebuah arsitektur *Convolutional Neural Network (CNN)* yang memanfaatkan teknik *transfer learning*. *Transfer learning* adalah metode di mana model yang telah dilatih pada besar seperti *COCO* digunakan sebagai dasar untuk tugas baru, yang memungkinkan penggunaan pengetahuan yang diperoleh dari pelatihan sebelumnya untuk meningkatkan performa pada yang lebih kecil. *Mask R-CNN* terdiri dari beberapa komponen utama seperti *backbone network* (seperti *ResNet*), *Feature Pyramid Network (FPN)*, *Region Proposal Network (RPN)*, serta tiga cabang fungsi utama yaitu klasifikasi, deteksi, dan segmentasi. Model ini diimplementasikan dengan menggunakan *Detectron2* yang memungkinkan akses ke model zoo yang berisi berbagai model *pre-trained* [10]. Selain *Detectron2*, digunakan juga library seperti *NumPy*, *Matplotlib*, *PIL*, *torch*, dan *roboflow*. *Training* dan *testing* model dilakukan di *Google Colab* karena mendukung penggunaan *GPU (Graphical Processing Unit)* dan memiliki spesifikasi yang tinggi. Kinerja model dievaluasi melalui proses validasi, dan hasil evaluasi tersebut dapat dilihat pada Tabel 3.

Tabel 3 : Hasil Evaluasi Model

Input	Metric	Bounding Box	Segmentation Mask
640 x 640 px	<i>AP@0.5</i>	0.658	0.702
	<i>AP@0.75</i>	0.422	0.469
	<i>AP@[0.5, 0.95]</i>	0.379	0.456
	<i>AR@1</i>	0.228	0.271
	<i>AR@10</i>	0.465	0.517
	<i>AR@100</i>	0.465	0.517

Model mencapai *mAP* sebesar 65,8% untuk deteksi *bounding box* dan 70,2% untuk *segmentation mask* pada threshold *AP@0,5*. Pada threshold *AP@0,75*, *mAP* mencapai 42,2% untuk *bounding box* dan

46,9% untuk *segmentation mask*. Secara keseluruhan, *mAP* pada rentang [0,5, 0,95] mencapai 37,9% untuk *bounding box* dan 45,6% untuk *segmentation mask*.

Penurunan performa pada $AP@0.75$ dan $AP@[0.5, 0.95]$ dibandingkan dengan $AP@0.5$ salah satu faktor penyebabnya adalah variasi ukuran dan bentuk objek. Pada *threshold* $AP@0.5$, model memiliki toleransi yang lebih besar terhadap kesalahan dalam deteksi selama terdapat tumpang tindih yang cukup antara prediksi dan *ground truth*. Namun, pada *threshold* yang lebih tinggi seperti $AP@0.75$, margin kesalahan yang diperbolehkan lebih kecil, sehingga model harus lebih akurat dalam menentukan ukuran dan posisi *bounding box* serta *segmentation mask*.

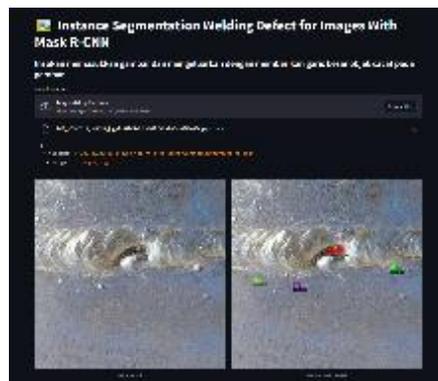
Hasil pengujian pada model dengan menggunakan *testing* dapat dilihat pada gambar 3. Pada gambar 3, cacat berhasil terdeteksi oleh model *Mask R-CNN*, di mana setiap cacat ditandai dengan *bounding box* serta *mask segmentasi* yang mengelilingi objek cacat. Selain itu, terdapat informasi mengenai akurasi prediksi untuk setiap cacat yang terdeteksi.



Gambar 3 : Hasil *Testing* Model

3.2.1. Hasil Implementasi *Streamlit*

Model deteksi yang telah dilatih kemudian diimplementasikan menggunakan *Streamlit*. Aplikasi ini digunakan untuk mendeteksi cacat pada hasil pengelasan dengan memproses gambar pengelasan dan menerapkan algoritma *Mask R-CNN*. Aplikasi ini memungkinkan identifikasi langsung terhadap jenis cacat seperti *spatter* dan *slag inclusion* pada gambar. Hasil dari implementasi model pada *Streamlit* dapat dilihat pada gambar 4. Pada gambar 4, ditampilkan tangkapan layar dari aplikasi *Streamlit* yang menunjukkan gambar asli bersama dengan gambar yang telah dilakukan deteksi.



Gambar 4 : Implementasi *Streamlit*

4. Kesimpulan

Penelitian ini berhasil mengimplementasikan algoritma *Mask R-CNN* untuk mendeteksi cacat pengelasan, khususnya *spatter* dan *slag inclusion* pada proses pengelasan SMAW. Dengan yang telah diolah dan dianotasi, model ini menunjukkan tingkat presisi tinggi dengan *mAP* sebesar 65,8% untuk *bounding box* dan 70,2% untuk *segmentation mask* pada *threshold* $AP@0,5$. Implementasi model dalam aplikasi *Streamlit* memungkinkan deteksi cacat secara *real-time* dan mempermudah inspeksi visual otomatis di industri pengelasan yang meningkatkan efisiensi, akurasi, dan mengurangi biaya, waktu, serta kesalahan manusia.

Penelitian selanjutnya, disarankan agar ukuran dan keragaman diperbesar sehingga model dapat lebih digeneralisasi dan diterapkan pada berbagai jenis pengelasan. Eksplorasi arsitektur model yang lebih kompleks dan penerapan teknik augmentasi data tambahan juga dapat dipertimbangkan untuk meningkatkan performa model. Selain itu, pengujian di lingkungan industri nyata akan memberikan gambaran lebih luas tentang efektivitas aplikasi ini dalam situasi nyata.

Daftar Pustaka

- [1] J. K. Park, W. H. An, and D. J. Kang, "Convolutional Neural Network Based Surface Inspection System for Non-patterned Welding Defects," *International Journal of Precision Engineering and Manufacturing*, vol. 20, no. 3, pp. 363–374, Mar. 2019, doi: 10.1007/s12541-019-00074-4.
- [2] H. Liu and Y. Hu, "Welding," in *Encyclopedia of Materials: Metals and Alloys*, F. G. Caballero, Ed., Oxford: Elsevier, 2022, pp. 39–65. doi: <https://doi.org/10.1016/B978-0-12-819726-4.00143-5>.
- [3] S. Rizvi and W. Alib, "Welding defects, Causes and their Remedies: A Review," *Teknomekanik*, vol. 2, pp. 39–47, Dec. 2019, doi: 10.24036/tm.v2i2.3272.
- [4] M. Shaloo, M. Schnall, T. Klein, N. Huber, and B. Reitingner, "A Review of Non-Destructive Testing (NDT) Techniques for Defect Detection: Application to Fusion Welding and Future Wire Arc Additive Manufacturing Processes," May 01, 2022, MDPI. doi: 10.3390/ma15103697.
- [5] S. J. Oh, M. J. Jung, C. Lim, and S. C. Shin, "Automatic detection of welding defects using faster R-CNN," *Applied Sciences (Switzerland)*, vol. 10, no. 23, pp. 1–10, Dec. 2020, doi: 10.3390/app10238629.
- [6] J.-N. Gou, X.-Y. Wu, and L. Liu, "Detection and Segmentation of Defects in Industrial CT Images Based on Mask R-CNN," *J Comput (Taipei)*, vol. 31, no. 6, pp. 141–154, 2020, doi: 10.3966/199115992020123106012.
- [7] K. Lin et al., "Face Detection and Segmentation Based on Improved Mask R-CNN," *Discrete Dyn Nat Soc*, vol. 2020, 2020, doi: 10.1155/2020/9242917.
- [8] L. Cai, T. Long, Y. Dai, and Y. Huang, "Mask R-CNN-Based Detection and Segmentation for Pulmonary Nodule 3D Visualization Diagnosis," *IEEE Access*, vol. 8, pp. 44400–44409, 2020, doi: 10.1109/ACCESS.2020.2976432.
- [9] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," Mar. 2017, [Online]. Available: <http://arxiv.org/abs/1703.06870>
- [10] A. V. Sai Abhishek and S. Kotni, "Detectron2 Object Detection & Manipulating Images using Cartoonization," *International Journal of Engineering and Technical Research*, vol. 10, Jul. 2022.