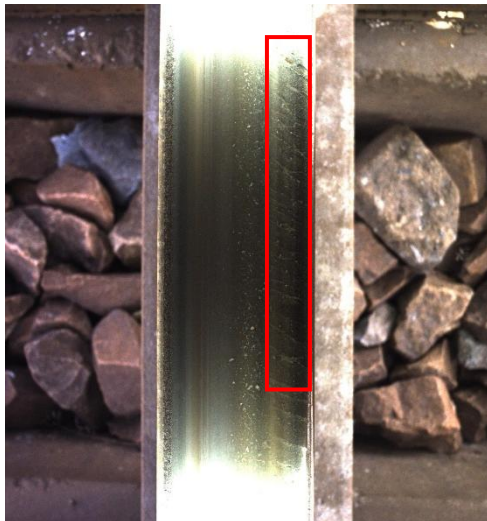


BAB II

LANDASAN TEORI

2.1 Cacat Pada Permukaan Rel Kereta Api

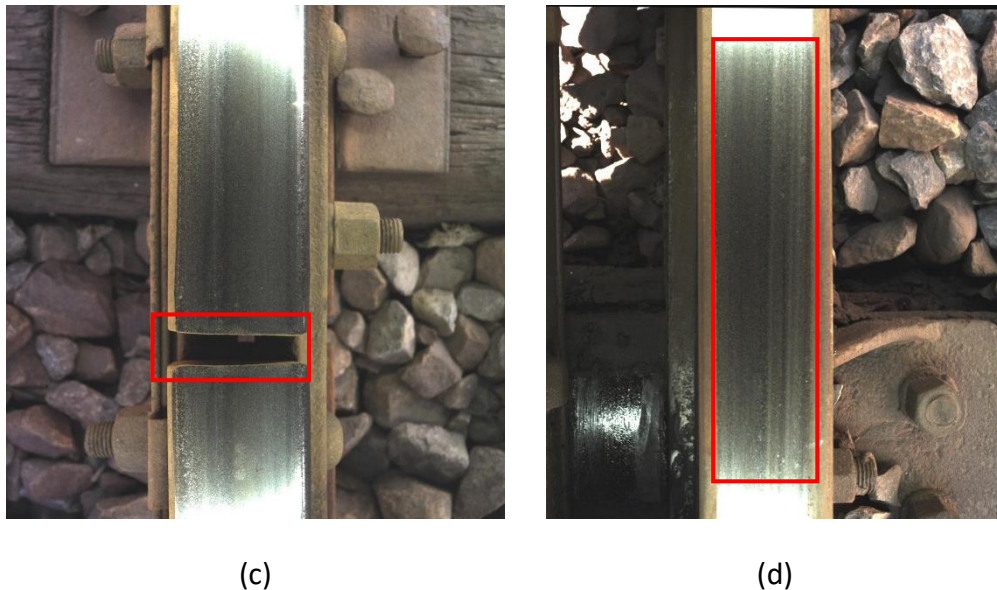
Rel kereta api merupakan salah satu infrastruktur penting dalam sistem transportasi modern yang harus memiliki keandalan tinggi, karena cacat pada rel menimbulkan risiko keselamatan yang serius, menyebabkan kecelakaan kereta api yang membahayakan keselamatan penumpang dan kru, merusak peralatan, serta mengganggu layanan transportasi (Kumar & Harsha, 2024). Cacat pada permukaan rel kereta api merupakan kerusakan yang muncul akibat berbagai faktor, baik dari proses manufaktur, pemakaian, maupun pengaruh lingkungan (Gong dkk., 2022). Jenis-jenis cacat permukaan rel ini sangat penting untuk dipahami karena dapat berdampak serius pada keamanan operasi kereta api dan memiliki karakteristik serta tingkat risiko yang bervariasi terhadap keselamatan perjalanan kereta api.



(a)



(b)



Gambar 2.1 (a) *Cracks*, (b) *Scars*, (c) *Breaks*, (d) *Lightbands*

Berdasarkan karakteristiknya, cacat permukaan rel dibedakan menjadi beberapa kategori, yaitu *cracks* (retakan), *scars* (goresan atau kerusakan akibat gesekan), *breaks*, dan *lightbands*. Cacat *cracks* atau retakan biasanya terjadi akibat kelelahan material yang terus-menerus menerima beban dinamis, terutama pada permukaan rel yang sering terpapar gaya geser dan beban berat dari roda kereta (Gong dkk., 2022). Jenis cacat *scars* memiliki tanda-tanda gesekan atau abrasi pada permukaan rel, yang dapat muncul dalam berbagai bentuk, seperti garis lurus, titik-titik, atau area yang luas (Wang dkk., 2024). *Breaks* merupakan cacat yang sangat berbahaya karena melibatkan kegagalan struktural pada rel berupa patahan yang memisahkan rel menjadi dua bagian (Wang dkk., 2024). Sedangkan *lightbands* merujuk pada tanda terang yang ditinggalkan pada tapak roda kereta api dan rel akibat keausan material selama proses gelinding dan geser (Qian dkk., 2025). Meskipun *lightband* pada umumnya tidak serta-merta menimbulkan ancaman

keselamatan langsung bagi transportasi kereta api, keberadaannya dapat menjadi indikator awal bagi cacat-cacat lain yang lebih parah (Wang dkk., 2024).

Kondisi rel yang optimal menjadi fondasi utama keandalan dan keselamatan operasional kereta api. Penggunaan rel secara intensif dan berulang dalam jangka waktu panjang secara alami akan menyebabkan kerusakan pada permukaan rel, yang pada akhirnya dapat mengganggu fungsi normal sistem transportasi (Arifianto, 2023). Dampak dari cacat permukaan rel sangat serius, tidak hanya menimbulkan risiko keselamatan yang serius bagi penumpang dan kru kereta, tetapi juga dapat menyebabkan kerusakan peralatan yang mahal dan mengganggu layanan transportasi secara keseluruhan (Kumar & Harsha, 2024). Mengingat potensi konsekuensi yang besar ini, deteksi dini terhadap cacat bukan lagi pilihan melainkan keharusan dalam menjaga keselamatan dan efisiensi transportasi kereta api (Zhao dkk., 2022).

Pentingnya deteksi cacat telah diakui secara luas, namun implementasi inspeksi rel dalam praktik menghadapi sejumlah tantangan teknis dan operasional yang kompleks. Industri transportasi kereta api telah lama bergantung pada pendekatan inspeksi konvensional yang memiliki keterbatasan signifikan (Xiong dkk., 2023). Pemeriksaan visual manual, sebagai metode yang paling umum digunakan, terbukti memerlukan investasi waktu yang substansial dan sangat bergantung pada keahlian serta pengalaman individu inspektur. Lebih problematik lagi, pendekatan ini rentan terhadap kesalahan manusia yang dapat berakibat fatal, sehingga tingkat efisiensi dan reliabilitasnya tidak memadai untuk memenuhi standar keselamatan modern (Kou, 2022; Xiong dkk., 2023). Tantangan operasional

juga menjadi hambatan serius karena proses ini tidak hanya membutuhkan waktu dan tenaga yang signifikan, tetapi juga sering kali mengharuskan penghentian sementara layanan kereta api, yang berdampak pada gangguan operasional dan kerugian ekonomi (Kumar & Harsha, 2024). Untuk mengatasi keterbatasan fundamental ini, industri telah mengadopsi berbagai teknologi pengujian non-destruktif (NDT) yang lebih canggih, termasuk pengujian ultrasonik (UT), pengujian elektromagnetik (ET), dan pengujian visual (VT) (Xiong dkk., 2023). Namun demikian, setiap teknologi NDT memiliki karakteristik dan keterbatasan spesifiknya sendiri. Sebagai contoh, meskipun inspeksi berbasis UT terbukti efektif untuk mendeteksi cacat internal pada struktur rel, teknologi ini kurang efektif untuk cacat permukaan yang seringkali menjadi indikator awal kerusakan yang lebih serius (Xiong dkk., 2023).

Evolusi teknologi dalam dekade terakhir telah membuka peluang baru untuk revolusi dalam metode inspeksi rel. Pengembangan metode inspeksi modern telah berhasil mengatasi banyak keterbatasan yang melekat pada pendekatan tradisional dan NDT konvensional. Inspeksi berbasis VT (*Visual Testing*) telah menunjukkan keunggulan yang signifikan, terbukti lebih cepat, efisien, dan sepenuhnya non-invasif dibandingkan teknik inspeksi lainnya, menjadikannya solusi yang ideal untuk mendeteksi cacat pada permukaan rel tanpa mengganggu operasi normal kereta api (Zhang dkk., 2023). Implementasi ini telah terbukti secara substansial meningkatkan efisiensi keseluruhan sistem deteksi, menawarkan solusi yang *scalable* dan *cost-effective* untuk industri transportasi kereta api modern (Kou, 2022).

2.2 Deteksi Objek

Computer vision memungkinkan komputer untuk memperoleh pemahaman tingkat tinggi dari gambar atau video digital. Tujuan utama dari *computer vision* dalam meniru persepsi visual manusia untuk berbagai tugas, mulai dari deteksi gambar, pengenalan objek, klasifikasi, hingga analisis video (Bhatt dkk., 2021; Khan & Al-Habsi, 2020). Kemampuan *computer vision* untuk memproses data visual dengan cepat dan akurat menjadikannya teknologi yang sangat valuable dalam meningkatkan efisiensi dan mengurangi ketergantungan pada inspeksi manual yang rentan terhadap *human error*.

Object detection merupakan salah satu tugas fundamental dalam *computer vision* yang bertujuan untuk mendeteksi dan melokalisasi objek tertentu dalam sebuah gambar atau video (Rodríguez-Abreo dkk., 2025). Proses *object detection* melibatkan dua tahap utama yaitu *feature extraction* untuk mengidentifikasi karakteristik visual objek dan *classification* untuk menentukan kategori objek yang terdeteksi (Diwan dkk., 2023). Teknologi ini memungkinkan sistem untuk mendeteksi *multiple objects* dalam satu gambar secara simultan, menjadikannya sangat cocok untuk aplikasi seperti deteksi cacat pada infrastruktur dimana *multiple defects* dapat muncul dalam satu area inspeksi.

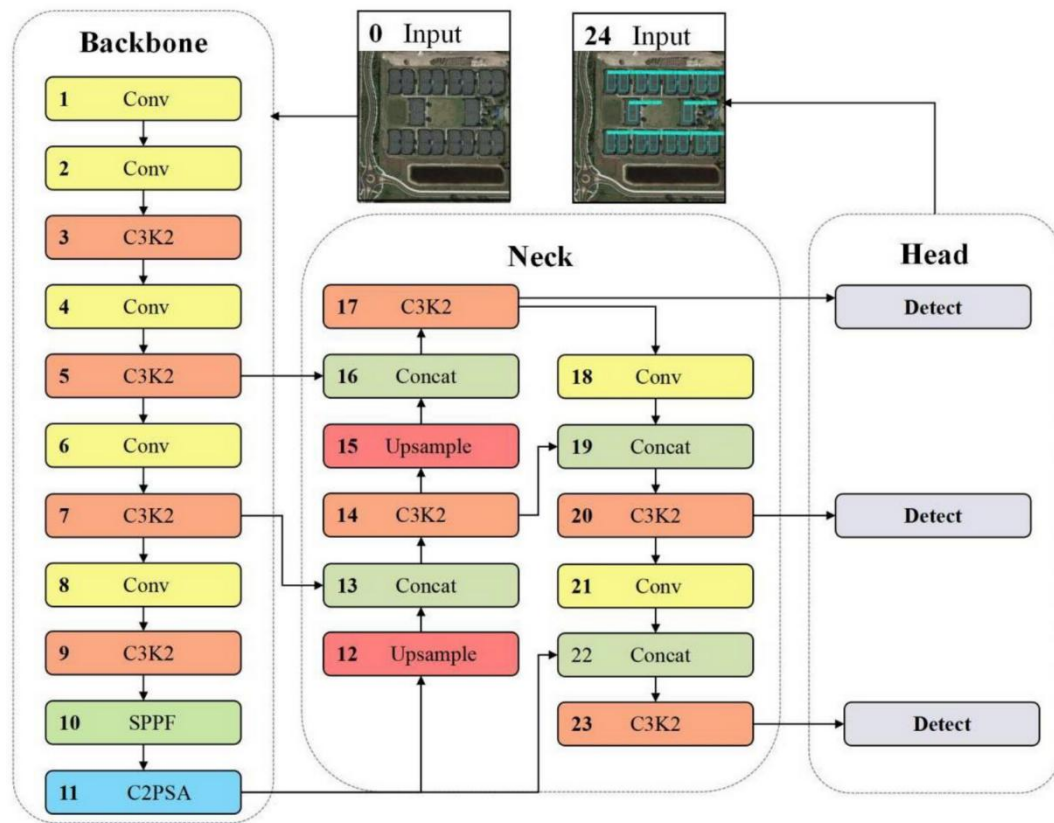
Dalam perkembangannya, algoritma deteksi objek telah mengalami evolusi signifikan dari pendekatan tradisional berbasis *handcrafted features* hingga *deep learning-based methods* yang lebih canggih. *Traditional methods* seperti *Histogram of Oriented Gradients* (HOG), *Scale-Invariant Feature Transform* (SIFT), dan *Support Vector Machine* (SVM) memiliki keterbatasan yang

substansial saat dihadapkan pada adegan kompleks, oklusi, atau variabilitas dalam tampilan objek, terutama karena ketergantungan mereka pada deskriptor fitur yang dibuat secara manual (Trigka & Dritsas, 2025). Sebaliknya, pendekatan *deep learning* khususnya *Convolutional Neural Networks* (CNN) mampu secara otomatis mempelajari representasi fitur yang optimal dari data *training*, menghasilkan performa yang jauh lebih superior. *Framework* deteksi objek yang modern seperti R-CNN, Fast R-CNN, Faster R-CNN, dan *single-stage detectors* seperti YOLO (*You Only Look Once*) dan SSD (*Single Shot MultiBox Detector*) telah merevolusi kemampuan deteksi objek dengan akurasi tinggi dan kecepatan inferensi yang *real-time* (Diwan dkk., 2023).

2.3 YOLO (*You Only Look Once*)

YOLO (*You Only Look Once*) merupakan salah satu algoritma deteksi objek *real-time* paling populer dalam bidang *computer vision*. Berbeda dengan pendekatan deteksi objek konvensional seperti R-CNN yang memisahkan proses region proposal dan klasifikasi, YOLO melakukan deteksi dalam satu kali *forward pass* pada jaringan saraf konvolusional, sehingga secara signifikan mempercepat proses inferensi (Terven dkk., 2023)

YOLO membagi gambar input ke dalam grid ukuran $S \times S$, dan untuk setiap grid, model memprediksi *bounding box*, *confidence score*, dan *class probability*. Konsep ini memungkinkan YOLO untuk mendeteksi berbagai objek dalam gambar dengan efisien. Versi-versi terbaru dari YOLO, seperti YOLOv3, YOLOv5, hingga YOLOv8, telah mengalami banyak peningkatan dalam hal arsitektur dan akurasi prediksi (Sohan dkk., 2024).



Gambar 2.3 Arsitektur Model YOLOv11 (L. H. He dkk., 2025)

Dalam penelitian ini, digunakan YOLOv11, yang merupakan versi pengembangan terbaru dari keluarga YOLO. YOLOv11 mengadopsi desain modular yang terdiri dari tiga bagian utama: *Backbone*, *Neck*, dan *Head* seperti yang ditunjukkan pada Gambar 2.3.

a. *Backbone*

Bagian ini bertanggung jawab untuk mengekstraksi fitur dari gambar input. Pada arsitektur YOLOv11, *backbone* terdiri dari beberapa lapisan konvolusi dan blok C3K2, yang digunakan untuk menangkap representasi spasial dan semantik dari gambar. Backbone berperan sebagai pengekstrak fitur utama yang mengubah citra input beresolusi tinggi menjadi serangkaian feature map bertingkat, sehingga

informasi spasial lokal dan konteks semantik global dapat direpresentasikan secara bertahap (Jegham dkk., 2024). Pada YOLOv11, backbone memanfaatkan blok C3k2 berbasis Cross-Stage Partial Network untuk meningkatkan efisiensi ekstraksi fitur dengan kompleksitas komputasi yang lebih rendah dibandingkan blok C2f pada YOLOv8 (Jegham et al., 2024; TriPerceptNet, 2025).

b. Neck

Neck bertugas untuk menggabungkan fitur dari berbagai level dalam backbone. Pada YOLOv11, neck terdiri dari proses *upsample*, *concatenate*, dan konvolusi tambahan, serta modul khusus seperti SPPF (*Spatial Pyramid Pooling - Fast*) dan C2PSA, yang meningkatkan kemampuan model dalam menangani objek dengan ukuran dan posisi yang beragam. Neck membantu menyatukan informasi global dan lokal dari fitur-fitur yang dihasilkan (L. He dkk., 2024).

c. Head

Head adalah bagian yang melakukan prediksi akhir, yaitu lokasi (*bounding box*), *confidence score*, dan kelas objek. YOLOv11 memiliki beberapa *head* deteksi untuk memprediksi objek pada berbagai skala resolusi (misalnya 80×80 , 40×40 , dan 20×20), yang meningkatkan performa pada objek kecil hingga besar (Jegham dkk., 2024).

2.4 Evaluasi Deteksi Objek Multi Kelas

Tahap evaluasi model melibatkan penilaian menyeluruh untuk memahami kualitas dan kinerja sistem yang telah dikembangkan. Evaluasi ini menggunakan metrik yang umum digunakan dalam tugas deteksi objek yaitu *precision*, *recall*,

mAP50, dan mAP50–95. Tantangan utama dalam deteksi objek multi kelas mencakup deteksi objek dengan ukuran yang sangat bervariasi *multi-scale objects*, deteksi objek berukuran sangat kecil *tiny objects*, dan penanganan ketidakseimbangan kelas pada dataset. Deteksi *tiny objects* (objek dengan ukuran di bawah 32×32 piksel) merupakan tantangan khusus karena informasi visual yang terbatas dan degradasi fitur akibat *downsampling* berturut-turut dalam arsitektur *deep learning* (Bosquet dkk., 2021). Studi pada skenario *tiny object detection* berbasis citra *aerial* menunjukkan bahwa peningkatan resolusi citra input dapat secara signifikan memperbaiki deteksi objek berskala kecil karena model mampu menangkap fitur yang lebih detail (Yan dkk., 2023). Sejalan dengan itu, YOLOv8 diintegrasikan dengan modul melaporkan peningkatan akurasi deteksi pada *benchmark tiny object* dibandingkan YOLOv8 asli, sehingga resolusi input tinggi dan *attention mechanism* merupakan strategi efektif untuk meningkatkan performa deteksi objek kecil (Kim dkk., 2025).

Evaluasi performa sistem deteksi objek menggunakan berbagai metrik yang mengukur akurasi dan efisiensi model. *Precision* mengukur proporsi *true positive detections* dari seluruh *positive predictions*, menunjukkan seberapa akurat model dalam mengidentifikasi objek target tanpa menghasilkan *false positives*. *Recall* (atau *sensitivity*) mengukur proporsi *true positive detections* dari seluruh *actual positive* anotasi, menggambarkan kemampuan model untuk mendeteksi semua objek target yang ada. *Mean Average Precision* (mAP) adalah metrik komprehensif yang menghitung rata-rata dari nilai *average precision* di semua kelas objek, sering

digunakan sebagai *standard benchmark* untuk membandingkan performa macam-macam model deteksi objek (Padilla et al., 2020.).

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$mAP_{50} = \frac{1}{N} \sum_{i=1}^N AP_{50,i} \quad (3)$$

$$mAP_{50-95} = \frac{1}{N} \sum_{i=1}^N \frac{1}{k} \sum_{j=1}^k \frac{1}{k} AP_{i,j} \quad (4)$$

2.5 Penelitian Terkait

Terdapat beberapa penelitian yang menggunakan berbagai pendekatan untuk meningkatkan kemampuan deteksi cacat pada permukaan rel kereta api. Temuan-temuan dari penelitian ini memberikan pemahaman penting terkait kekuatan dan keterbatasan masing-masing metode, serta menjadi dasar dalam perancangan solusi yang lebih efisien dan adaptif untuk mendeteksi berbagai jenis cacat pada permukaan rel. Namun, sebagian besar studi sebelumnya masih berfokus pada klasifikasi biner (cacat atau tidak cacat), sementara klasifikasi multi-kelas untuk jenis cacat seperti *scars* dan *cracks* masih jarang dijelajahi. Temuan-temuan tersebut dirangkum dalam Tabel 2.1.

Tabel 2.1 Penelitian Terkait

No	Penulis	Isi	Relevansi	Perbedaan
1	ECARRNet: An Efficient LSTM-Based Ensembled Deep Neural Network Architecture for Railway Fault Detection, (Eunus dkk., 2024)	Mengembangkan model hybrid untuk deteksi cacat pada rel dan memperoleh akurasi 93.28%.	Menunjukkan pendekatan alternatif dalam fokus penelitian yang sama yaitu deteksi cacat rel.	Menggunakan model hybrid dan berfokus pada klasifikasi biner, berbeda dengan pendekatan YOLOv11 untuk deteksi multi-kelas pada penelitian ini.
2	Fault Diagnosis in Railway Track using Efficient Net based CNN, (Mordia & Kumar Verma, 2022)	Mengembangkan model CNN berbasis EfficientNet untuk mendeteksi cacat rel, dengan akurasi 91% lebih dari 30 <i>epoch</i> .	Menunjukkan penggunaan CNN untuk deteksi cacat rel, relevan dengan fokus penelitian yang sama.	Menggunakan EfficientNet CNN, berbeda dengan pendekatan YOLOv11.
3	Automatic Detection of Railway Faults Using Neural Networks: A Comparative Study of Transfer Learning Models	Studi komparatif antara model transfer learning dan YOLOv11 untuk deteksi kesalahan rel, mencapai akurasi uji 92,64%	Menunjukan kinerja YOLOv11 dibandingkan dengan model lain.	Fokus pada perbandingan model dan deteksi cacat biner (cacat atau tidak cacat), bukan pada

	and YOLOv11, (Rodríguez-Abreo dkk., 2025)	dengan YOLOv11.		deteksi multi-kelas.
4	Railway Track Fault Detection with ResNet Deep Learning Models, (Saritas dkk., 2023)	Menggunakan model ResNet-18, ResNet-50, dan ResNet-101 untuk klasifikasi cacat rel, dengan akurasi tertinggi 81,74% pada ResNet-18.	Menunjukkan penggunaan CNN tradisional (ResNet) untuk fokus penelitian yang sama yaitu deteksi cacat rel.	Menggunakan ResNet, berbeda dengan pendekatan YOLOv11.
5	An Improved YOLOv8 Algorithm for Rail Surface Defect Detection, (Wang dkk., 2024)	Mengusulkan menambahkan E MA <i>attention</i> dan SPD-Conv di YOLOv8 ini mencapai presisi 93,9%, <i>recall</i> 93,7%, dan akurasi rata-rata 94,1%.	Menunjukkan efektivitas kinerja YOLO dengan versi berbeda yang dimodifikasi dalam fokus penelitian yang sama yaitu deteksi cacat rel	Menggunakan YOLOv8, berbeda dengan pendekatan YOLOv11.
6	Rail Surface Defect Detection Based on Image Enhancement and Improved	Meningkatkan YOLOX dengan <i>attention mechanism</i> dan <i>image enhancement</i> ,	Menunjukkan peningkatan performa model YOLO dengan <i>attention</i>	Menggunakan YOLOX, berbeda dengan pendekatan YOLOv11.

	YOLOX, (Zhang dkk., 2023)	meningkatkan mAP sebesar 2,42% menjadi 93,20%	<i>mechanism</i> , relevan sebagai metode peningkatan kinerja	
7	Scale-Balanced Real-Time Object Detection with Varying Input-Image Resolution (Yan dkk., 2023)	Mengusulkan deteksi objek real-time dengan variasi resolusi citra input dan menunjukkan bahwa peningkatan resolusi secara signifikan membantu deteksi objek kecil.	Menunjukkan bahwa penggunaan ukuran citra lebih besar lebih efektif untuk mendeteksi objek kecil.	Berfokus pada variasi resolusi untuk deteksi objek umum, berbeda dengan pendekatan YOLOv11 dan tidak berfokus pada cacat rel.
8	Enhanced Detection of Tiny Objects in Aerial Images (Kim dkk., 2025)	Mengintegrasikan modul <i>attention</i> SE dan CBAM pada YOLOv8 memperoleh akurasi deteksi <i>tiny objects</i> yang lebih tinggi daripada YOLOv8 asli.	Menunjukkan bahwa penambahan <i>attention mechanism</i> pada YOLO efektif	Menggunakan YOLOv8, berbeda dengan pendekatan YOLOv11.
9	Deteksi Kecacatan Permukaan Rel	Penelitian ini menggunakan <i>deep learning</i>	Relevan karena membahas deteksi cacat	Menggunakan <i>deep learning neural network</i>

	Menggunakan Metode Deep Learning Neural Network, (Arifianto, 2023)	<i>neural network</i> dengan akurasi tertinggi 80,41% (training) dan 79% (testing).	permukaan rel menggunakan deep learning.	dengan gabor filter, bukan YOLOv11. Fokus penelitian ini adalah klasifikasi biner, bukan deteksi multi-kelas.
10	YOLOv2 With Bifold Skip: A Deep Learning Model for Video Based Real Time Train Bogie Part Identification and Defect Detection, (Mohan Et Al., 2021)	Mengusulkan modifikasi YOLOv2 dengan <i>bifold skip architecture</i> untuk identifikasi bagian bogie kereta api dan deteksi cacat secara real-time dari video dan diperoleh akurasi 69.0%	Relevan karena menggunakan varian YOLO (YOLOv2) dan berfokus pada deteksi cacat.	Menggunakan YOLOv2 yang dimodifikasi, bukan YOLOv11. Fokus pada identifikasi bagian bogie selain deteksi cacat.
11	SED-YOLO Based Multi-Scale Attention for Small Object Detection in Remote Sensing (Wei dkk., 2025)	Menambahkan EMA pada YOLOv5s secara efektif meningkatkan akurasi deteksi objek kecil dalam citra <i>remote sensing</i>	Relevan karena mengintegrasikan EMA pada YOLO	Menggunakan YOLOv5s yang dimodifikasi, bukan YOLOv11.

12	Optimasi YOLOv11 melalui <i>Hyperparameter Tuning</i> dan Data Augmentasi untuk Meningkatkan Akurasi Deteksi Kendaraan pada Kondisi Malam Hari. (Alfath Zulkarnain & Kusri, 2025)	Mengoptimalkan YOLOv11 untuk deteksi kendaraan pada malam hari. Optimasi dilakukan dengan menyesuaikan <i>hyperparameter</i> serta menerapkan augmentasi data. Mencapai <i>precision</i> 0.97, <i>recall</i> 0.92, dan <i>mAP@0.5</i> sebesar 0.97, melampaui YOLOv8 dan YOLOv11 <i>baseline</i> .	Relevan karena menggunakan kombinasi <i>hyperparameter tuning</i> dan augmentasi data pada YOLOv11	Penelitian ini adalah deteksi kendaraan malam hari pada dataset UA-DETRAC, tidak berfokus pada cacat rel.
----	---	--	--	---

Pada tabel 2.1, terdapat beberapa penelitian yang relevan terkait deteksi cacat rel menggunakan berbagai arsitektur model deep learning, termasuk varian YOLO yang telah dimodifikasi dan metode lainnya. Salah satunya adalah penelitian oleh Eunus dkk. (2024) yang mengusulkan ECARRNet, sebuah model *hybrid* yang menggabungkan *convolutional autoencoder*, ResNet, dan *recurrent neural network*, serta berhasil mencapai akurasi 93,28%. Namun, penelitian ini masih menghadapi tantangan dalam mendeteksi cacat yang jarang terjadi dan fokus pada klasifikasi biner.

Penelitian lain oleh Mordia & Kumar Verma (2022) mengembangkan model CNN berbasis EfficientNet yang mencapai akurasi 91% lebih dari 30 *epoch* untuk deteksi cacat rel. Demikian pula, penelitian Arifianto (2023) yang menerapkan *deep learning neural network* dengan *gabor filter* juga berfokus pada klasifikasi biner, dengan akurasi *training* 80,41% dan akurasi *testing* 79%. Studi oleh Saritas et al. (2023) menunjukkan penggunaan model ResNet-18, ResNet-50, dan ResNet-101 untuk klasifikasi cacat rel, dengan ResNet-18 mencapai akurasi tertinggi 81,74%. Mohan dkk. (2021) mengusulkan modifikasi YOLOv2 dengan *bifold skip architecture* untuk identifikasi bagian bogie kereta api dan deteksi cacat secara *real-time* dari video dan diperoleh akurasi 69.0%. Performa model yang kurang optimal dapat disebabkan oleh penggunaan versi YOLO yang lebih lama, sehingga membatasi kemampuan deteksinya.

Beberapa penelitian lain juga menunjukkan upaya peningkatan performa model deteksi dengan berbagai modifikasi YOLO. Misalnya, Wang dkk. (2024) memodifikasi YOLOv8 untuk deteksi cacat rel permukaan, mencapai presisi 93,9%, *recall* 93,7%, dan akurasi rata-rata 94,1%. Zhang dkk. (2023) meningkatkan YOLOX dengan *attention mechanism* dan *image enhancement*, meningkatkan mAP sebesar 2,42% menjadi 93,20%. SED-YOLO memasukkan *Efficient Multi-Scale Attention* (EMA) ke dalam YOLOv5s dan menunjukkan peningkatan kinerja deteksi objek kecil pada citra remote sensing. Kim dkk. (2025) mengintegrasikan modul *attention* SE dan CBAM pada YOLOv8 (MoonNet) dan melaporkan peningkatan akurasi deteksi *tiny objects* dibandingkan YOLOv8 asli.

Selain penelitian yang secara langsung berfokus pada cacat rel, terdapat juga studi yang berkontribusi pada aspek teknis yang relevan dengan penelitian ini. Yandkk. (2023) menunjukkan bahwa peningkatan resolusi citra input secara signifikan dapat meningkatkan deteksi objek kecil pada detektor real-time berbasis YOLO karena fitur spasial yang lebih detail dapat dipertahankan. Alfath Zulkarnain & Kusri (2025) mengoptimasi YOLOv11 untuk deteksi kendaraan pada kondisi malam hari melalui kombinasi *hyperparameter tuning* (*learning rate*, *momentum*, *weight decay*) dan augmentasi data (penyesuaian HSV, translasi, skala, horizontal flip, mosaic, mixup), dan melaporkan peningkatan *precision* hingga 0,97 dan mAP@0,5 hingga 0,98 dibanding YOLOv8 dan YOLOv11 *baseline*.

Penelitian-penelitian di atas menunjukkan bahwa meskipun berbagai model *deep learning* telah diterapkan dalam deteksi cacat rel, sebagian besar penelitian masih berfokus pada klasifikasi biner (cacat atau tidak cacat) atau menggunakan versi YOLO sebelumnya (YOLOV5, YOLOv8, YOLOX, YOLOv2) atau pendekatan arsitektur lain seperti EfficientNet, ResNet, dan model *hybrid*. Walaupun Rodríguez-Abreo dkk. (2025) menggunakan YOLOv11 dan mencapai akurasi uji 92,64%, penelitian mereka tetap terbatas pada klasifikasi biner. Di sisi lain, penelitian terkait *small object detection*, strategi *resampling*, dan optimasi YOLOv11 melalui *hyperparameter tuning* dan data augmentasi memberikan landasan teoritis bagi konfigurasi model yang digunakan dalam penelitian ini. Oleh karena itu, penelitian ini dilakukan untuk mengeksplorasi potensi YOLOv11 dalam deteksi multi-kelas cacat rel, yang diharapkan mampu mengatasi keterbatasan

deteksi biner dan meningkatkan akurasi serta efisiensi deteksi berbagai jenis cacat pada permukaan rel kereta api.

Tabel 2.2 Matriks Penelitian

No	Peneliti (Tahun)	Multi Kelas	<i>Attention Mechanism</i>	Augmentasi Data	Resolusi
1	Eunus dkk. (2024)	-	-	✓	-
2	Mordia & Verma (2022)	-	-	-	-
3	Arifianto (2023)	-	-	-	-
4	Saritas dkk. (2023)	-	-	-	-
5	Mohan dkk. (2021)	-	-	-	-
6	Wang dkk. (2024)	✓	✓	✓	-
7	Yan dkk. (2023)	-	-	-	✓
8	Kim dkk. (2025)	-	✓	-	-
9	Luo dkk. (2023)	-	✓	✓	-
10	Rodríguez-Abreo dkk. (2025)	-	-	-	-
11	Wei dkk. (2025)	-	✓	-	-
12	Zulkarnain & Kusri (2025)	-	-	✓	-
13	Usulan Penelitian	✓	✓	✓	✓

Tabel 2.2 merangkum karakteristik utama penelitian terkait dari sisi metode yang digunakan, dukungan terhadap klasifikasi multi-kelas, pemanfaatan *attention mechanism*, strategi augmentasi data dan peningkatan resolusi. Sebagian besar

penelitian menggunakan arsitektur CNN konvensional atau varian YOLO generasi sebelumnya, dengan fokus dominan pada klasifikasi biner cacat atau tidak cacat.

Berdasarkan pemetaan pada Tabel 2.2 tersebut, dapat dilihat bahwa belum banyak penelitian yang menggabungkan YOLOv11 dengan skenario deteksi cacat permukaan rel pada dataset multi-kelas, sekaligus mengeksplorasi pengaruh resolusi input tinggi dan konfigurasi augmentasi yang disesuaikan dengan dataset. Hal ini menjadi celah utama yang diisi oleh penelitian ini.