

Perbandingan Modul Atensi *Coordinate Attention* (CA) dan *Convolutional Block Attention Module* (CBAM) pada Arsitektur YOLOv8n untuk Deteksi Retak Jembatan

^{1*}Krisna Dzakyaa Darajat, ²Resty Wulanningrum, ³Risky Aswi Ramadhani

^{1,2,3} Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹krisnadzakyaa@gmail.com, ²restyw@unpkdr.ac.id,

³riskyaswiramadhani@unpkediri.ac.id

Penulis Korespondens : Krisna Dzakyaa Darajat

Abstrak— Pemeliharaan jembatan membutuhkan inspeksi retak rambut guna mencegah kegagalan struktural. Inspeksi manual memakan waktu dan sulit mengidentifikasi retakan tipis pada permukaan beton yang bising (*noisy background*). Penelitian ini bertujuan mengembangkan sistem deteksi otomatis dengan membandingkan efektivitas YOLOv8n standar terhadap model modifikasi bermekanisme atensi: *Coordinate Attention* (CA) dan *Convolutional Block Attention Module* (CBAM). Manfaat penelitian ini adalah menyediakan landasan sistem peringatan dini yang akurat, *real-time*, dan hemat biaya komputasi. Temuan ini berdampak pada efisiensi pengawasan karena membuktikan fitur dasar YOLOv8n sudah optimal untuk retakan tipis tanpa perlu arsitektur tambahan. Evaluasi dilakukan menggunakan dataset citra retak jembatan dengan mengukur metrik *Precision*, *Recall*, dan *mAP*. Hasil eksperimen menunjukkan model *baseline* YOLOv8n mencapai kinerja tertinggi dengan *Precision* 83%, *Recall* 72%, *mAP@50* 78%, dan *mAP@50-95* 49%. Sebaliknya, penambahan modul atensi menunjukkan sedikit penurunan performa; model CBAM menghasilkan *mAP@50* sebesar 77% dan model CA sebesar 76%.

Kata Kunci— Deteksi Retak, YOLOv8n, *Coordinate Attention*, CBAM, Infrastruktur, Jembatan

Abstract— Bridge maintenance requires hairline crack inspections to prevent structural failures. Manual inspections are time-consuming and struggle to identify thin cracks on noisy concrete surfaces. This study aims to develop an automated detection system by comparing the effectiveness of the standard YOLOv8n against modified models with attention mechanisms: *Coordinate Attention* (CA) dan *Convolutional Block Attention Module* (CBAM). This research provides a foundation for an accurate, *real-time*, and computationally efficient early warning system. These findings impact monitoring efficiency by proving that the *baseline* YOLOv8n features are already optimal for thin cracks without requiring additional architectures. The evaluation was conducted using a bridge crack image dataset by measuring *Precision*, *Recall*, and *mAP* metrics. Experimental results show the *baseline* YOLOv8n achieved the highest performance with 83% *Precision*, 72% *Recall*, 78% *mAP@50*, and 49% *mAP@50-95*. Conversely, adding attention modules showed a slight performance decrease; the CBAM model yielded an *mAP@50* of 77%, and the CA model scored 76%.

Keywords— Crack Detection, YOLOv8n, *Coordinate Attention*, CBAM, Infrastructure, Bridge

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Jembatan merupakan urat nadi infrastruktur transportasi yang memiliki peran krusial dalam menghubungkan wilayah, memfasilitasi mobilitas, dan mendorong pertumbuhan ekonomi. Namun, sebagai struktur fisik yang berada di ruang terbuka, jembatan terus-menerus menerima beban mekanis yang dinamis dari kendaraan, serta paparan cuaca ekstrem yang memicu penurunan kualitas material secara perlahan [1]. Tanpa adanya pemeliharaan yang memadai, penurunan fungsi material ini dapat berujung pada kerusakan fatal. Oleh karena itu, inspeksi dan pemeliharaan rutin menjadi langkah preventif yang mutlak diperlukan untuk memperpanjang usia pakai jembatan dan, yang terpenting, menjamin keselamatan nyawa para penggunanya [2].

Dalam siklus degradasi struktural tersebut, salah satu indikator awal yang paling sering muncul namun kerap diremehkan adalah fenomena retak rambut (*hairline cracks*) pada permukaan beton jembatan. Meski ukurannya sangat kecil dan sekilas tampak tidak membahayakan, retakan mikroskopis ini sebenarnya merupakan celah kerentanan yang serius. Retakan ini menjadi jalur masuk utama bagi air, oksigen, dan zat korosif seperti klorida menuju tulangan baja di dalam beton. Jika dibiarkan berlarut-larut, proses korosi akan menggerogoti struktur dari dalam, secara diam-diam melemahkan kekuatan jembatan, dan pada akhirnya dapat memicu kegagalan struktural secara tiba-tiba tanpa peringatan lanjutan [3].

Sayangnya, urgensi penanganan retak rambut ini sering kali terabaikan akibat sistem pengawasan konvensional yang masih sangat bergantung pada inspeksi visual secara manual oleh manusia. Metode ini tidak hanya memakan waktu dan biaya yang besar, tetapi juga sangat rentan terhadap subjektivitas dan kelelahan mata inspektur di lapangan. Keterbatasan visual manusia membuat retakan yang sangat tipis ini mudah sekali terlewatkan. Selain itu, kondisi permukaan beton jembatan yang sering kali kasar, dipenuhi noda, lumut, atau bayangan (*noisy background*) semakin menyamarkan bentuk retakan, sehingga menuntut tingkat ketelitian luar biasa yang sulit dicapai secara konsisten melalui metode pengecekan manual [3] [4].

Menjawab tantangan akurasi dan ketelitian tersebut, penelitian ini mengusulkan penerapan teknologi *computer vision* [3] berbasis *deep learning* [5], secara khusus menggunakan arsitektur deteksi objek satu tahap YOLOv8n (*You Only Look Once versi 8 Nano*) [6]. YOLOv8n dipilih karena keunggulannya dalam melakukan inferensi secara *real-time* dengan kebutuhan komputasi yang sangat ringan. Namun, untuk mengatasi kesulitan melacak fitur retak rambut yang tipis dan menyatu dengan latar belakang, penelitian ini mengintegrasikan mekanisme atensi (*attention mechanism*) [10]. Melalui penggunaan dataset citra retak jembatan dari *Roboflow Universe*, penelitian ini akan melakukan uji kinerja secara mendalam untuk membandingkan efektivitas modul *Coordinate Attention* (CA) [15] dan *Convolutional Block Attention Module* (CBAM) [16]. Evaluasi ini diarahkan untuk melihat metrik presisi dan akurasi dari masing-masing modul atensi ketika disematkan pada *backbone* YOLOv8n [6].

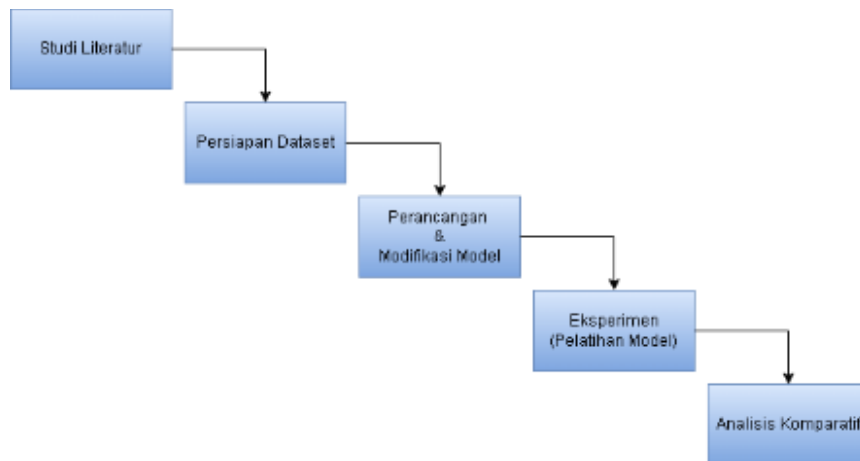
Diharapkan, dengan adanya sistem deteksi cerdas ini, proses identifikasi retak rambut pada infrastruktur jembatan dapat dilakukan secara otomatis, lebih cepat, dan jauh lebih objektif dibandingkan inspeksi manusia. Lebih jauh lagi, pemaparan komparatif dalam artikel ini akan memberikan wawasan baru mengenai arsitektur atensi mana yang paling optimal untuk mengurai kompleksitas visual retakan tipis. Pada akhirnya, inovasi sistem ini diharapkan dapat menjadi fondasi bagi sistem peringatan dini (*early warning system*) yang andal, membantu otoritas terkait

mengambil keputusan pemeliharaan dengan cepat, menghemat anggaran perbaikan berskala besar, dan mencegah potensi bencana keruntuhan di masa depan.

II. METODE

Penelitian ini menggunakan arsitektur YOLOv8n sebagai *baseline*. Modifikasi dilakukan secara struktural dengan menyisipkan modul *Coordinate Attention* (YOLOv8n+CA) dan modul *Convolutional Block Attention Module* (YOLOv8n+CBAM) pada *Backbone* setelah blok SPPF. Penilaian performa akhir dilakukan dengan menggunakan metrik pengujian standar dalam visi komputer, yaitu *Precision*, *Recall*, serta *mean Average Precision* (mAP) yang diukur pada ambang batas *Intersection over Union* (IoU) sebesar 0,5 (*mAP@50*) dan rentang rata-rata 0,5 hingga 0,95 (*mAP@50-95*).

Dimulai dengan persiapan dan anotasi dataset, diikuti oleh pembagian *train-validation-test* yang tetap dan digunakan secara bersama-sama di semua varian model. Setiap model dilatih menggunakan prosedur yang identik dan dievaluasi pada kumpulan data uji yang sama menggunakan metrik deteksi yang terstandarisasi.



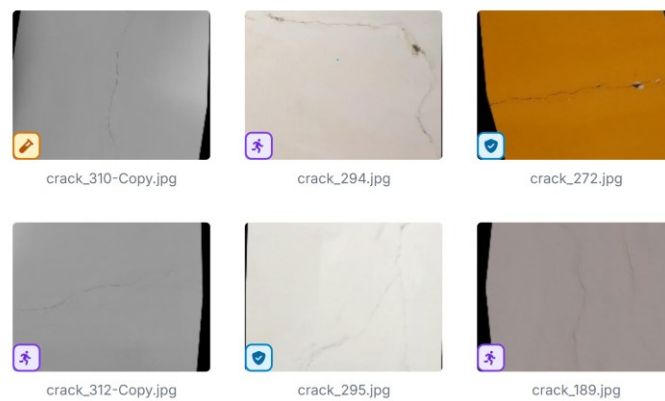
Gambar 1. Diagram Waterfall Prosedur Penelitian

1. Studi Literatur

Diawali dengan Studi Literatur, saya mempelajari teori YOLOv8, mekanisme *Coordinate Attention* (CA), *CBAM*, dan karakteristik kerusakan retak jembatan.

2. Persiapan Dataset

Penelitian ini menggunakan dataset sekunder publik yang bersumber dari *Roboflow Universe*, yaitu dataset "*Crack.v3i.yolov8*". Dataset ini dipilih karena memiliki karakteristik visual retak yang saya rasa cukup relevan dengan kasus jembatan.



Gambar 2. Dataset yang dipakai untuk penelitian ini

Spesifikasi datanya terdiri dari, Format Anotasi dengan *Object Detection*. Dengan jumlah kelas sebanyak 1 Kelas yaitu kelas "*Crack Detection 1 - v4 2023-04-04 2:42pm*". Lalu data tersebut terdapat total 4.295 citra, yang terdiri dari 3.017 citra latih (training set), 849 citra validasi (validation set) dan 429 citra tes. Serta karakteristik citra berisi variasi tekstur dinding dengan berbagai jenis retak (rambut, memanjang, dll) dalam kondisi pencahayaan yang beragam dan telah terotasi. Penggunaan dataset tersebut akan dibagi secara konsisten untuk ketiga skenario eksperimen untuk menjamin validitas perbandingan.

3. Perancangan & Modifikasi Model

Pada tahapan ini, saya menyiapkan kode konfigurasi model YOLOv8n standar (*Baseline*). Memodifikasi model melalui file konfigurasi YAML untuk menyisipkan modul *Coordinate Attention* (CA) pada *Backbone*. Memodifikasi model melalui file konfigurasi YAML untuk menyisipkan modul *CBAM* pada *Backbone*.

4. Eksperimen (Pelatihan Model)

Ketiga model akan dilatih sebanyak 100 *epochs*, *batch*=16, dan *image size*=256. Ini dilakukan untuk menjamin validitas perbandingan. Dan *image size* 256 dipilih karena menyesuaikan kemampuan hardware.

5. Analisis Komparatif

Saya Membandingkan hasil sesuai dengan metrik performa yang berlaku (*mAP@50*, *Precision*, *Recall*, *mAP@50-95*). Dan menggunakan parameter *hyperparameter* yang identik untuk menjaga konsistensi.

III. HASIL DAN PEMBAHASAN

Bagian ini menguraikan hasil dari serangkaian eksperimen yang telah dirancang pada bagian metode penelitian. Pembahasan dibagi menjadi tiga tahap utama yang selaras dengan metodologi, yaitu hasil pelatihan model, evaluasi metrik kinerja, dan analisis komparatif pengaruh modul atensi terhadap arsitektur dasar.

A. Hasil

Sesuai dengan skenario yang ditetapkan pada metode penelitian, ketiga model yaitu YOLOv8n (sebagai *baseline*), YOLOv8n dengan *Coordinate Attention* (CA), dan YOLOv8n dengan *Convolutional Block Attention Module* (CBAM) telah berhasil dilatih menggunakan dataset retak jembatan. Pelatihan dilakukan secara konsisten selama 100 epoch dengan ukuran citra masukan 256x256 piksel dan batch 16.

Penggunaan konfigurasi fail .yaml khusus pada model YOLOv8n+CA dan YOLOv8n+CBAM memastikan bahwa model dilatih dari awal tanpa pembobotan awal (*pre-trained weights*), sehingga murni mengevaluasi kemampuan arsitektur baru dalam mengekstraksi fitur retakan. Selama proses pelatihan, model baseline YOLOv8n menunjukkan tingkat konvergensi loss yang lebih cepat dibandingkan kedua model yang dimodifikasi.

B. Pembahasan

1. Evaluasi Kinerja Model

Berdasarkan pemasangan mekanisme atensi pada asitektur *Backbone* YOLOv8n.

```
1 # Ultralytics YOLO 🚀, AGPL-3.0 license
2 # YOLOv8n object detection model with P3-P5 outputs. For YOLOv8n.
3
4 # Parameters
5 nc: 1 # jumlah class (sesuaikan dengan dataset crack Anda)
6 scales: # model compound scaling constants, i.e. 'n', 's', 'm', 'l', 'x'
7 # [depth, width, max_channels]
8 n: [0.33, 0.25, 1024] # <--- PERUBAHAN DI SINI (width 0.25 untuk Nano)
9
10 # YOLOv8.0n backbone
11 backbone:
12 # [from, repeats, module, args]
13 - [-1, 1, Conv, [64, 3, 2]] # 0-P1/2
14 - [-1, 1, Conv, [128, 3, 2]] # 1-P2/4
15 - [-1, 3, C2f, [128, True]]
16 - [-1, 1, Conv, [256, 3, 2]] # 3-P3/8
17 - [-1, 6, C2f, [256, True]]
18 - [-1, 1, Conv, [512, 3, 2]] # 5-P4/16
19 - [-1, 6, C2f, [512, True]]
20 - [-1, 1, Conv, [1024, 3, 2]] # 7-P5/32
21 - [-1, 3, C2f, [1024, True]]
22 # --- MODIFIKASI (CoordAtt) pada layer 9 ---
23 - [-1, 1, CoordAtt, [1024]] # 9
24 # -----
25 - [-1, 1, SPPF, [1024, 5]] # 10
```

Gambar 3. Modifikasi model melalui file .yaml untuk menyisipkan modul (CA) pada *Backbone*

```
2 # Ultralytics YOLO 🚀, AGPL-3.0 license
3 # YOLOv8 object detection model with P3-P5 outputs. For YOLOv8
4
5 # Parameters
6 nc: 1 # Jumlah class (crack, mould)
7 scales: # model compound scaling constants, i.e. 'n'
8 n: [0.33, 0.25, 1024]
9
10 # YOLOv8.0n backbone
11 backbone:
12 # [from, repeats, module, args]
13 - [-1, 1, Conv, [64, 3, 2]] # 0-P1/2
14 - [-1, 1, Conv, [128, 3, 2]] # 1-P2/4
15 - [-1, 3, C2f, [128, True]]
16 - [-1, 1, Conv, [256, 3, 2]] # 3-P3/8
17 - [-1, 6, C2f, [256, True]]
18 - [-1, 1, Conv, [512, 3, 2]] # 5-P4/16
19 - [-1, 6, C2f, [512, True]]
20 - [-1, 1, Conv, [1024, 3, 2]] # 7-P5/32
21 - [-1, 3, C2f, [1024, True]]
22 - [-1, 1, CBAM, [1024]] # <--- Layer 9: Penambahan CBAM sebelum SPPF
23 - [-1, 1, SPPF, [1024, 5]] # 10
```

Gambar 4. Modifikasi model melalui file .yaml untuk menyisipkan modul (CBAM) pada *Backbone*

Kinerja dari ketiga model dievaluasi menggunakan metrik yang telah dijabarkan pada metode penelitian, yaitu *Precision*, *Recall*, *mAP@50*, dan *mAP@50-95*. Hasil pengujian pada data validasi dapat dilihat pada Tabel 1.

Tabel 1. Perbandingan Performa Model

Model	Precision	Recall	mAP@50	mAP@50-95
YOLOv8n	0.833	0.728	0.786	0.499
YOLOv8n+CA	0.811	0.691	0.765	0.486
YOLOv8n+CBAM	0.799	0.701	0.771	0.467

Berdasarkan Tabel 1, model YOLOv8n standar memperoleh kinerja tertinggi secara keseluruhan. Model ini mencatat tingkat presisi sebesar 0.833 dan $mAP@50$ sebesar 0.786. Sebaliknya, penambahan modul atensi justru menghasilkan penurunan kinerja. Model YOLOv8n+CBAM mengalami penurunan presisi terendah menjadi 0.799, sedangkan model YOLOv8n+CA mengalami penurunan *recall* terendah menjadi 0.691.

2. Analisis Pengaruh Modul Atensi pada Deteksi Retak

Hasil evaluasi menunjukkan adanya anomali terhadap hipotesis awal di mana penambahan mekanisme atensi diasumsikan dapat meningkatkan akurasi deteksi. Merujuk pada metode pengembangan arsitektur model, fenomena penurunan kinerja ini dianalisis berdasarkan dua faktor utama.

Pertama, karakteristik visual dataset retak jembatan yang sangat tipis dan bertekstur menyatu dengan latar belakang. Modul atensi, khususnya CBAM yang menerapkan operasi pemusatan spasial (*spatial pooling*), cenderung menurunkan detail piksel yang sangat halus. Akibatnya, informasi penting mengenai bentuk geometris retakan justru terabaikan (*blurred*) oleh lapisan jaringan tambahan tersebut, yang berujung pada tingginya angka *false positive* (penurunan nilai presisi).

Kedua, perbedaan status inisialisasi bobot. Model YOLOv8n standar memiliki struktur jaringan ringan yang sangat optimal dan stabil. Di sisi lain, penambahan lapisan CA dan CBAM ke dalam *backbone* (setelah blok SPPF) meningkatkan kompleksitas jaringan yang dilatih dari awal. Durasi pelatihan 100 epoch terbukti belum mencukupi bagi model modifikasi untuk menyeimbangkan penyesuaian bobot pada modul atensi yang baru, sehingga model mengalami kondisi *underfitting* relatif bila dibandingkan dengan model dasarnya.

IV. KESIMPULAN

Hasil eksperimen menunjukkan bahwa model *baseline* YOLOv8n ‘Nano’ mencapai nilai tertinggi pada *Precision* (0,833), *Recall* (0,728), $mAP@50$ (0,786), dan $mAP@50-95$ (0,499). Penambahan modul atensi pada dataset ini menunjukkan sedikit penurunan performa, dengan model CBAM mencapai $mAP@50$ sebesar 0,771 dan model CA mencapai $mAP@50$ sebesar 0,765. Hal ini mengindikasikan bahwa fitur ekstraksi dasar dari YOLOv8n sudah cukup optimal untuk karakteristik dataset retak jembatan yang digunakan, dan kompleksitas tambahan dari mekanisme atensi membutuhkan penyesuaian *hyperparameter* lebih lanjut.

DAFTAR PUSTAKA

- [1]. Y. Jeong, W. S. Kim, I. Lee, and J. Lee, “Bridge inspection practices and bridge management programs in China, Japan, Korea, and U.S.,” *J. Struct. Integr. Maint.*, vol. 3, no. 2, pp. 126–135, Apr. 2018, doi: 10.1080/24705314.2018.1461548.

- [2]. G. Xu and F. Azhari, "Bridge Maintenance Management Based on Routine Inspection Data: A Quantitative Approach," *J. Bridg. Eng.*, vol. 30, no. 2, p. 04024111, Feb. 2025, doi: 10.1061/JBENF2.BEENG-6685.
- [3]. K. Luo, X. Kong, J. Zhang, J. Hu, J. Li, and H. Tang, "Computer Vision-Based Bridge Inspection and Monitoring: A Review," Sep. 13, 2023, *Multidisciplinary Digital Publishing Institute*. doi: 10.3390/s23187863.
- [4]. F. Poli, M. F. Bado, A. Verzobio, and D. Zonta, "Bridge structural safety assessment: a novel solution to uncertainty in the inspection practice," *Struct. Infrastruct. Eng.*, vol. 21, no. 3, pp. 421–435, 2023, doi: 10.1080/15732479.2023.2211956;WGROU:STRING:PUBLICATION.
- [5]. Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/NATURE14539;SUBJMETA.
- [6]. H. Nie, H. Pang, M. Ma, and R. Zheng, "A Lightweight Remote Sensing Small Target Image Detection Algorithm Based on Improved YOLOv8," *Sensors*, vol. 24, no. 9, p. 2952, May 2024, doi: 10.3390/s24092952.
- [7]. Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2021, pp. 13708–13717. doi: 10.1109/CVPR46437.2021.01350.
- [8]. S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, pp. 3–19. doi: 10.1007/978-3-030-01234-2_1.
- [9]. J. Jia and Y. Li, "Deep Learning for Structural Health Monitoring: Data, Algorithms, Applications, Challenges, and Trends," *Sensors (Basel)*, vol. 23, no. 21, p. 8824, Oct. 2023, doi: 10.3390/S23218824/S1.
- [10]. M. Su, J. Wan, Q. Zhou, R. Wang, Y. Xie, and H. Peng, "Utilizing pretrained convolutional neural networks for crack detection and geometric feature recognition in concrete surface images," *J. Build. Eng.*, vol. 98, p. 111386, Dec. 2024, doi: 10.1016/J.JOBE.2024.111386.
- [11]. Y. Wang and J. He, "A Rapid Concrete Crack Detection Method Based on Improved YOLOv8," 2025, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2025.3555825.
- [12]. [12] P. F. Giordano, S. Quqa, and M. P. Limongelli, "The value of monitoring a structural health monitoring system," *Struct. Saf.*, vol. 100, p. 102280, Jan. 2023, doi: 10.1016/j.strusafe.2022.102280.
- [13]. [13] N. Avelina, T. Chang, and S. Chi, "Comparative Study of Bridge Inspection Practices in Indonesia and Foreign Countries," in *IEEE International Conference on Industrial Engineering and Engineering Management*, IEEE Computer Society, 2022, pp. 980–984. doi: 10.1109/IEEM55944.2022.9989924.
- [14]. A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep Learning for Computer Vision: A Brief Review," Jan. 01, 2018, *John Wiley & Sons, Ltd.* doi: 10.1155/2018/7068349.
- [15]. R. U. Khan and R. Kromanis, "Overview and Challenges of Computer Vision-Based Visual Inspection for the Assessment of Bridge Defects," pp. 336–344, Sep. 2025, doi: 10.3217/978-3-99161-057-1-052.

- [16]. A. A. Mustapha and M. S. Yoosuf, "Exploring the efficacy and comparative analysis of one-stage object detectors for computer vision: a review," *Multimed. Tools Appl.*, vol. 83, no. 20, pp. 59143–59168, Jun. 2024, doi: 10.1007/s11042-023-17751-2.
- [17]. Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/NATURE14539.
- [18]. H. Wang *et al.*, "Research on automatic pavement crack identification Based on improved YOLOv8," *Int. J. Interact. Des. Manuf.*, vol. 18, no. 6, pp. 3773–3783, Feb. 2024, doi: 10.1007/s12008-024-01769-3.
- [19]. L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, pp. 1–74, Mar. 2021, doi: 10.1186/s40537-021-00444-8.
- [20]. J. Li, "Improving the Application of YOLOv8 in Image Object Detection," in *2024 6th International Conference on Communications, Information System and Computer Engineering, CISCE 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 668–673. doi: 10.1109/CISCE62493.2024.10653313.
- [21]. E. S. Jun, H. J. Sim, and S. J. Moon, "Advancing YOLOv8-Based Wafer Notch-Angle Detection Using Oriented Bounding Boxes, Hyperparameter Tuning, Architecture Refinement, and Transfer Learning," *Appl. Sci.*, vol. 15, no. 21, Nov. 2025, doi: 10.3390/app152111507.
- [22]. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Dec. 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91.
- [23]. Z. J. Khaw, Y. F. Tan, H. A. Karim, and H. A. A. Rashid, "Improved YOLOv8 Model for a Comprehensive Approach to Object Detection and Distance Estimation," *IEEE Access*, vol. 12, no. May, pp. 63754–63767, 2024, doi: 10.1109/ACCESS.2024.3396224.