

Optimasi Deteksi Objek pada Video dengan Kompresi Region of Interest menggunakan Model YOLOv8

Azhryl Akbar Assagaf¹, Muis Muhtadi^{1,*}

¹ Program Studi Teknik Informatika, Universitas Negeri Malang, Indonesia

* Correspondence: muis.muhtadi.ft@um.ac.id

Copyright: © 2025 by the authors

Received: 13 April 2025 | Revised: 29 April 2025 | Accepted: 20 Mei 2025 | Published: 9 Agustus 2025

Abstrak

Kebutuhan akan sistem deteksi objek *real-time*, seperti pada pengawasan video dan kendaraan otonom, mendorong efisiensi dalam penyimpanan dan transmisi data tanpa mengorbankan akurasi. Salah satu pendekatan yang menjanjikan adalah kompresi video berbasis *Region of Interest* (ROI), yang mempertahankan kualitas visual pada area penting. Penelitian ini bertujuan mengevaluasi pengaruh kompresi video terhadap akurasi deteksi objek menggunakan model YOLOv8 melalui pengujian *Analysis of Variance* (ANOVA), serta membandingkan efektivitas kompresi seragam dan berbasis ROI. Video dari VIRAT Video Dataset dikompresi menggunakan parameter *Constant Rate Factor* (CRF) dengan dua metode dan dievaluasi berdasarkan mAP_{50} , $mAP_{50_{95}}$, serta ukuran file. Hasil uji ANOVA menunjukkan tidak terdapat perbedaan signifikan secara statistik antara kedua metode. Pada CRF 50, ukuran file dapat berkurang lebih dari 60%, namun akurasi mAP_{50} menurun hingga di bawah 50% karena degradasi kualitas pada area non-ROI mengganggu konteks spasial yang dibutuhkan model. Penelitian ini berkontribusi dengan menguji batas toleransi kompresi pada YOLOv8 dan menunjukkan bahwa kualitas visual menyeluruh lebih menentukan performa model dibanding hanya mempertahankan kualitas objek. Implikasi temuan ini penting untuk aplikasi *real-time*, seperti CCTV dan kendaraan otonom, yang membutuhkan keseimbangan antara efisiensi kompresi dan akurasi deteksi. Penelitian selanjutnya dapat mengeksplorasi pendekatan ROI adaptif yang mempertimbangkan pergerakan objek secara dinamis.

Kata kunci: deteksi objek; evaluasi map kompresi video; *region of interest*; yolov8

Abstract

The demand for real-time object detection systems, such as those used in video surveillance and autonomous vehicles, drives the need for efficient data storage and transmission without compromising accuracy. One promising approach is Region of Interest (ROI)-based video compression, which preserves visual quality in important areas. This study aims to evaluate the impact of video compression on object detection accuracy using the YOLOv8 model through statistical analysis using Analysis of Variance (ANOVA), and to compare the effectiveness of uniform and ROI-based compression methods. Videos from the VIRAT Video Dataset were compressed using the Constant Rate Factor (CRF) parameter and evaluated based on mAP_{50} , $mAP_{50_{95}}$, and file size. ANOVA results indicate no statistically significant differences between the two methods. At CRF 50, file size can be reduced by over 60%, but mAP_{50} accuracy drops below 50% due to quality degradation in non-ROI areas, which disrupts the spatial context required by the model. This study contributes by examining the compression tolerance limits of YOLOv8 and reveals that overall visual quality, rather than just object-focused quality, plays a crucial role in model performance. These findings have important implications for real-time applications such as CCTV and autonomous vehicles, where a balance between compression efficiency and detection accuracy is critical. Future studies may explore adaptive ROI approaches that consider dynamic object movement.



Keywords: *object detection; video compression; region of interest; yolov8; map evaluation*

PENDAHULUAN

Video telah menjadi media utama dalam kehidupan digital modern, digunakan secara luas dalam berbagai bidang seperti pendidikan (*e-learning*), hiburan (*streaming* film), keamanan (CCTV), dan bisnis (pemasaran video). Popularitas video meningkat seiring perkembangan teknologi digital dan perluasan akses internet. Pertumbuhan trafik data seluler sebagian besar dipengaruhi oleh konsumsi layanan berbasis data seperti video, yang diharapkan menjadi pendorong utama peningkatan trafik data global pada beberapa tahun mendatang (Ariyanti et al., 2021). Platform seperti YouTube, TikTok, dan layanan konferensi seperti Zoom dan Microsoft Teams memperkuat posisi video sebagai alat komunikasi utama. Dalam konteks pendidikan dan kerja jarak jauh, video menawarkan komunikasi waktu nyata dan interaktif (Lakhan & Verma, 2023). Di bidang keamanan, video membantu mengurangi kejahatan dan meningkatkan keselamatan, namun juga menghadirkan tantangan privasi dan keamanan data (Yu et al., 2021).

Pertumbuhan konsumsi video yang pesat menimbulkan tantangan besar dalam penyimpanan dan transmisi data. Penyimpanan video yang tidak efisien dapat menyebabkan peningkatan biaya operasional dan beban infrastruktur. Menurut Haidous et al. (2022), ukuran file video yang besar serta kebutuhan akses yang tinggi menjadi hambatan utama dalam sistem penyimpanan hemat daya. Setiap hari, jutaan video diunggah di berbagai platform sosial media (Al, 2021), dan resolusi tinggi seperti 4K/8K menambah tekanan terhadap *bandwidth* dan kapasitas penyimpanan (Achlisson et al., 2023).

Sebagai upaya mengatasi tantangan yang ditimbulkan oleh ukuran file video yang besar dan kebutuhan *bandwidth* yang tinggi dalam penyimpanan dan transmisi data, kompresi video menjadi solusi kunci dalam mengurangi ukuran file tanpa menurunkan kualitas secara signifikan. Kompresi seperti H.264 dan H.265 digunakan luas karena efisiensi dan kemampuannya mengurangi redundansi (Khadir et al., 2024). H.265 menghasilkan ukuran file yang lebih kecil dibandingkan H.264 (Tubagus et al., 2021), meskipun dalam beberapa kasus kualitas visual H.264 lebih baik (Alnoor & Mustafa, 2022). Namun, kompresi tidak hanya berdampak pada efisiensi penyimpanan, tetapi juga dapat menurunkan akurasi sistem deteksi objek dalam visi komputer. Artefak kompresi seperti *blur*, *blocking*, dan *noise* dapat mengganggu proses inferensi model deteksi objek berbasis deep learning, terutama pada tingkat kompresi tinggi (Kwon et al., 2022). (Liu et al., 2023) mencatat bahwa performa deteksi sangat dipengaruhi oleh kualitas visual input. Meskipun beberapa *codec* modern lebih tahan terhadap *noise* (Hanooman et al., 2021), model seperti YOLO tetap memerlukan input berkualitas untuk hasil optimal (Rajasekhar, 2024).

Salah satu pendekatan yang menjanjikan untuk menjaga akurasi deteksi sambil tetap menghemat ukuran file adalah kompresi berbasis *Region of Interest* (ROI). Pendekatan ini menjaga kualitas tinggi hanya pada area penting (seperti objek manusia atau kendaraan), dan mengompresi latar belakang secara agresif. Hal ini sejalan dengan prinsip *perceptual coding* yang mengutamakan bagian paling relevan secara visual. Yao et al. (2022) menunjukkan bahwa mempertahankan kualitas area penting meningkatkan akurasi deteksi. ROI berbasis *deep learning* juga memungkinkan pengkodean adaptif yang efisien (Li et al., 2023). semenantara itu, Ma et al. (2021) membuktikan bahwa teknik ROI mereka unggul dalam metrik kualitas visual seperti LPIPS dan FID pada *bitrate* rendah.

Penelitian ini bersifat komprehensif karena mengkaji pengaruh kompresi ROI terhadap akurasi deteksi objek menggunakan model YOLOv8. YOLOv8 merupakan model terbaru dalam seri YOLO yang dikembangkan oleh Ultralytics, dirancang untuk deteksi objek *real-time* dengan kecepatan dan akurasi tinggi serta kemudahan penggunaan yang lebih baik dibandingkan versi sebelumnya (Terven et al., 2023). Sebagian besar penelitian sebelumnya

berfokus pada kompresi seragam (O’Byrne et al., 2022; Gandor & Nalepa, 2022), atau mengevaluasi ROI dengan model lain seperti *Single Shot Multibox Detector* (SSD) atau CenterNet (Shinohara et al., 2020). Beberapa penelitian hanya menilai efisiensi *bandwidth* tanpa mengevaluasi akurasi deteksi objek (Achlisson et al., 2023), yang berisiko menimbulkan kesimpulan keliru karena kompresi yang efisien secara ukuran belum tentu menghasilkan performa deteksi yang andal. Tanpa pengukuran akurasi seperti mAP, performa sistem deteksi dapat menurun drastis, terutama pada aplikasi nyata seperti CCTV atau kendaraan otonom yang menuntut ketepatan tinggi. Oleh karena itu, penelitian ini mengisi kekosongan tersebut dengan menilai akurasi deteksi secara menyeluruh pada kompresi ROI menggunakan YOLOv8, untuk memastikan efisiensi kompresi tidak mengorbankan kemampuan sistem dalam mengenali objek secara tepat.

Penelitian ini bertujuan untuk mengevaluasi pengaruh kompresi video berbasis ROI terhadap akurasi deteksi objek pada model YOLOv8. Secara khusus, penelitian ini membandingkan akurasi deteksi antara metode kompresi seragam dan ROI, serta mengidentifikasi batas *Constant Rate Factor* (CRF) maksimum yang masih dapat mempertahankan akurasi deteksi yang dapat diterima. Temuan dari penelitian ini diharapkan dapat memberikan wawasan baru mengenai penerapan kompresi ROI dalam sistem deteksi objek dan memberikan kontribusi dalam pengembangan teknik kompresi video yang lebih efisien tanpa mengorbankan akurasi deteksi.

METODE

Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan desain *true experiment*, yang memungkinkan kontrol penuh terhadap variabel-variabel yang terlibat. Desain ini dipilih karena dapat memberikan perbandingan yang tepat antara dua metode kompresi, yaitu kompresi seragam dan kompresi berbasis ROI, serta mengendalikan variabel lainnya seperti tingkat kompresi dan jenis dataset. Pendekatan ini memungkinkan pembuktian hubungan sebab-akibat antara variabel bebas (metode kompresi) dan variabel terikat (akurasi deteksi objek dan efisiensi penyimpanan).

Dataset yang digunakan adalah VIRAT Video Dataset, yang terdiri dari tiga video berdurasi 15 detik, masing-masing berisi objek manusia yang direkam dalam kondisi pengawasan dengan kualitas video rendah dan beberapa *noise* latar belakang. Variasi ini memberikan tantangan tambahan dalam deteksi objek dan memungkinkan evaluasi kompresi video dalam kondisi dunia nyata yang kompleks.

Proses kompresi video dilakukan menggunakan perangkat FFmpeg, perangkat lunak *open source* untuk pemrosesan video. Dua metode kompresi diterapkan dalam penelitian ini, yaitu kompresi seragam dan berbasis ROI. Pada metode kompresi seragam, seluruh *frame* video dikompresi menggunakan CRF yang berkisar antara 22 hingga 60. Nilai CRF yang lebih rendah menghasilkan kompresi yang lebih agresif, mengurangi ukuran file secara signifikan, namun menyebabkan penurunan kualitas visual. Sebaliknya, nilai CRF yang lebih tinggi menghasilkan kualitas video yang lebih baik, namun dengan ukuran file yang lebih besar. Penelitian ini memilih CRF karena parameter ini telah terbukti mempengaruhi kualitas video dan performa deteksi objek pada penelitian sebelumnya.

Pada metode kompresi berbasis ROI, FFmpeg dikonfigurasi untuk memprioritaskan kualitas tinggi pada area terdeteksi sebagai ROI, sementara area non-ROI dikompresi lebih agresif. ROI sendiri ditentukan berdasarkan deteksi objek awal yang dilakukan oleh model YOLOv8 pada video yang belum dikompresi. Teknik ini memungkinkan penghematan ukuran file pada bagian latar belakang yang tidak mengandung objek penting, sementara kualitas visual pada area yang mengandung objek utama dipertahankan.

Proses deteksi objek pertama dilakukan pada video yang belum dikompresi menggunakan model YOLOv8 *pretrained*, yang dipilih karena kemampuannya mendeteksi

objek dengan akurasi tinggi dalam waktu relatif singkat. Anotasi objek dilakukan menggunakan *Computer Vision Annotation Tool* (CVAT) untuk menghasilkan *ground truth* berupa *bounding box* yang mengelilingi objek manusia. Setelah proses kompresi, deteksi objek dilakukan kembali pada video yang telah dikompresi, dan hasilnya dibandingkan dengan *ground truth* yang telah dihasilkan.

Analisis statistik dilakukan menggunakan ANOVA untuk menguji apakah ada perbedaan signifikan dalam mAP₅₀ antara kedua metode kompresi pada berbagai nilai CRF. Uji ANOVA ini bertujuan untuk mengevaluasi apakah perbedaan dalam metode kompresi dan tingkat CRF yang berbeda memberikan dampak signifikan terhadap akurasi deteksi objek yang dihasilkan oleh model YOLOv8. Hasil uji statistik ini akan memberikan gambaran mengenai apakah metode ROI memberikan keuntungan dalam mengurangi ukuran file tanpa mengorbankan akurasi deteksi objek.

HASIL DAN PEMBAHASAN

Hasil

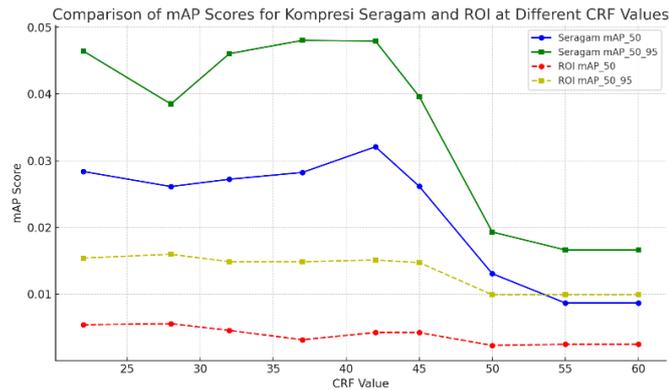
Hasil penelitian menunjukkan bahwa meskipun kompresi seragam cenderung memberikan akurasi deteksi objek yang lebih baik dibandingkan dengan kompresi ROI, perbedaan ini tidak signifikan secara statistik menurut hasil uji ANOVA. Meskipun ROI menghasilkan ukuran file yang lebih kecil, penurunan akurasi deteksi objek pada ROI lebih terlihat pada tingkat CRF yang lebih tinggi. Sementara itu, kompresi seragam mempertahankan kualitas deteksi yang lebih baik, terutama pada CRF 32, yang memberikan *trade-off* terbaik antara ukuran file dan akurasi deteksi objek, meskipun perbedaan antara kedua metode tetap tidak signifikan. Selain itu, meskipun ukuran file pada ROI lebih kecil pada nilai CRF tertentu, kualitas deteksi objek mengalami penurunan yang lebih besar dibandingkan kompresi seragam, yang menekankan pentingnya keseimbangan antara efisiensi kompresi dan akurasi deteksi.

Tabel 1. Hasil kompresi seragam

CRF	mAP ₅₀	mAP _{50_95}	Ukuran file (Mb)	Deteksi
22	0,028379	0,046434	7,333879	110
28	0,026120	0,038486	2,057688	111
32	0,027216	0,046041	0,790911	112
37	0,028236	0,048025	0,298460	90
42	0,032081	0,047947	0,047947	95
45	0,026147	0,039604	0,131265	103
50	0,013064	0,019307	0,099494	52
55	0,008669	0,016620	0,097104	33
60	0,008669	0,016620	0,097104	33

Tabel 2. Hasil kompresi menggunakan roi

CRF	mAP ₅₀	mAP _{50_95}	Ukuran file (Mb)	Deteksi
22	0,005402	0,015402	5,961729	106
28	0,005557	0,015969	2,554031	103
32	0,004559	0,014851	1,334572	107
37	0,003157	0,014851	0,620703	100
42	0,004249	0,015103	0,348492	112
45	0,004229	0,014731	0,264114	116
50	0,002310	0,009901	0,193444	87
55	0,002474	0,009901	0,186893	91
60	0,002474	0,009901	0,186893	91



Gambar 1. Visualisasi nilai mAP_50 dan mAP_50_95 untuk metode kompresi seragam dan roi

Gambar 1 menunjukkan hubungan antara CRF (sumbu X) dan mAP_50 serta mAP_50_95 (sumbu Y) pada kedua metode kompresi. Grafik ini menggunakan warna atau jenis garis yang berbeda untuk membedakan antara kompresi seragam dan ROI. Grafik ini memperlihatkan bahwa pada tingkat CRF yang lebih tinggi, baik seragam maupun ROI mengalami penurunan akurasi deteksi objek yang cukup signifikan, meskipun kompresi seragam cenderung mempertahankan akurasi yang lebih tinggi pada sebagian besar nilai CRF.

Pada masing-masing nilai CRF, analisis dilakukan untuk membandingkan hasil mAP_50, mAP_50_95, ukuran file, dan jumlah deteksi objek antara kedua metode kompresi. Berdasarkan data yang diperoleh, analisis menunjukkan beberapa temuan. Pada nilai CRF 22, kompresi seragam menghasilkan mAP_50 yang lebih tinggi (0,028379) dibandingkan dengan ROI (0,005402), meskipun ukuran file keduanya hampir serupa. Ini menunjukkan bahwa meskipun ROI seharusnya mempertahankan kualitas objek yang terdeteksi, ROI justru mengalami penurunan akurasi deteksi objek yang lebih besar. Pada CRF 28, meskipun ukuran file pada ROI sedikit lebih kecil (2,55 MB) dibandingkan seragam (2,06 MB), mAP_50 pada ROI (0,0056) lebih rendah dibandingkan seragam (0,026120). Ini mengindikasikan bahwa meskipun ROI dapat mengurangi ukuran file lebih signifikan, pengurangan kualitas visual pada ROI menyebabkan penurunan akurasi deteksi objek yang lebih besar. Pada CRF 32, kompresi seragam menunjukkan mAP_50 sebesar 0,027216, yang lebih tinggi dibandingkan dengan ROI yang hanya mencapai 0,004559. Ini menunjukkan bahwa kompresi seragam memberikan *trade-off* terbaik antara ukuran file dan akurasi deteksi objek, sedangkan ROI, meskipun dirancang untuk mempertahankan objek penting, gagal mempertahankan kualitas deteksi objek secara optimal. Pada nilai CRF yang lebih tinggi, baik seragam maupun ROI mengalami penurunan mAP_50 yang signifikan. Kompresi seragam pada nilai CRF 45 memberikan mAP_50 sebesar 0,026147, sementara ROI hanya mencapai 0,004229. Penurunan ini menunjukkan bahwa pada CRF yang lebih tinggi, kualitas video menjadi lebih terdegradasi, yang mempengaruhi hasil deteksi objek.

Pada konteks deteksi objek, jumlah deteksi objek merujuk pada total prediksi yang benar atau *true positives* yang dihasilkan oleh model deteksi objek. Jumlah deteksi ini dihitung berdasarkan objek yang berhasil dideteksi oleh model YOLOv8 dan sesuai dengan *ground truth* yang telah dianotasi secara manual. Metrik ini sangat penting karena menunjukkan seberapa efektif model dalam mengidentifikasi objek yang relevan dalam video yang telah dikompresi.

Seiring dengan peningkatan nilai CRF, perubahan dalam jumlah deteksi objek menunjukkan adanya korelasi dengan perubahan nilai mAP. Sebagai contoh, pada CRF 22, meskipun jumlah deteksi objek pada metode ROI (106 deteksi) sedikit lebih rendah dibandingkan seragam (110 deteksi), mAP_50 pada seragam lebih tinggi. Hal ini

mengindikasikan bahwa meskipun jumlah deteksi objek hampir sama, kualitas deteksi (diukur dengan mAP) pada seragam lebih baik. Sebaliknya, pada CRF yang lebih tinggi, penurunan jumlah deteksi objek diikuti dengan penurunan nilai mAP₅₀ yang signifikan, yang menunjukkan bahwa akurasi deteksi sangat dipengaruhi oleh tingkat kompresi yang lebih tinggi.

Beberapa anomali dalam data, seperti fluktuasi pada mAP₅₀ dan mAP_{50_95}, terdeteksi di beberapa titik CRF, terutama pada CRF yang lebih tinggi. Hal ini mungkin disebabkan oleh *noise* yang ada dalam video pengawasan, yang mengganggu deteksi objek pada ROI. Pada video pengawasan dengan kualitas yang kurang dan adanya *noise* latar belakang, ROI yang dirancang untuk mempertahankan objek dengan kualitas tinggi justru menghadapi kesulitan dalam meningkatkan akurasi deteksi, karena *noise* latar belakang dapat membingungkan model dalam mengidentifikasi objek yang relevan.

Analisis statistik menggunakan ANOVA dilakukan untuk mengevaluasi apakah perbedaan mAP₅₀ antara kedua metode kompresi dan berbagai tingkat CRF dapat dianggap signifikan atau hanya merupakan variasi acak. Hasil uji ANOVA menunjukkan *p-value* sebesar 0,8436, yang lebih besar dari 0,05, yang berarti kompresi dengan menggunakan metode ROI memiliki pengaruh terhadap akurasi deteksi objek meskipun kurang signifikan.

Tabel 3. CRF optimal

Metode	CRF Optimal	mAP ₅₀	Ukuran file (MB)	Jumlah Deteksi
Seragam	32	0,027	0,79	106
ROI	28	0,0056	2,55	103

Berdasarkan hasil analisis tabel 1 dan 2, CRF 32 pada metode seragam menunjukkan *trade-off* terbaik antara ukuran file dan akurasi deteksi objek. Pada tabel 3 ditunjukkan CRF yang optimal untuk masing-masing metode kompresi dan CRF 32 pada metode seragam mendapatkan hasil yang lebih baik. Pada nilai CRF ini, kualitas video masih dapat dipertahankan dengan ukuran file yang relatif kecil, sementara akurasi deteksi objek tetap berada pada tingkat yang dapat diterima.

Pembahasan

Hasil penelitian ini menunjukkan bahwa semakin tinggi nilai CRF, semakin kecil ukuran file yang dihasilkan. Penurunan ukuran file ini diiringi dengan penurunan akurasi deteksi objek, yang mencerminkan *trade-off* antara efisiensi penyimpanan dan kualitas deteksi. Pada metode kompresi seragam, mAP₅₀ tetap relatif stabil hingga CRF 45, setelah itu terjadi penurunan performa yang signifikan. Hal ini mengindikasikan bahwa pada tingkat CRF yang lebih tinggi, kompresi video mulai menghilangkan detail penting yang diperlukan untuk deteksi objek yang akurat. Penurunan ini terlihat pada mAP₅₀ yang semakin menurun seiring dengan meningkatnya CRF. Meskipun demikian, kompresi seragam tetap menunjukkan hasil yang lebih baik dibandingkan metode ROI pada nilai CRF tinggi, meskipun ukuran file lebih besar.

Sementara itu, pada metode ROI, meskipun dirancang untuk mempertahankan kualitas objek penting dalam video, hasil penelitian ini menunjukkan bahwa ROI tidak memberikan hasil yang lebih baik dibandingkan metode seragam. Salah satu penyebabnya adalah degradasi kualitas pada area non-ROI yang justru memengaruhi akurasi model. YOLOv8, sebagai model deteksi objek modern, mengandalkan informasi spasial menyeluruh dari seluruh *frame*. Ketika bagian non-ROI mengalami degradasi, konteks spasial terganggu dan lapisan awal model yang bertugas mengekstraksi fitur visual tidak dapat bekerja secara optimal. Akibatnya, meskipun objek utama berada dalam ROI, kualitas deteksi tetap menurun.

Hasil uji statistik menggunakan ANOVA menunjukkan bahwa perbedaan akurasi antara kedua metode tidak signifikan secara statistik. Namun secara praktis, penurunan kinerja deteksi yang konsisten pada ROI tetap relevan dalam konteks penerapan sistem nyata seperti pengawasan video atau kendaraan otonom, yang sangat bergantung pada akurasi deteksi yang konsisten.

Temuan ini selaras dengan hasil dari O'Byrne et al. (2022) dan Gandor & Nalepa (2022), yang mencatat penurunan akurasi seiring peningkatan tingkat kompresi. Begitu pula dengan Shinohara et al. (2020), yang menunjukkan bahwa meskipun ROI mampu mengurangi ukuran file, penurunan performa deteksi tetap menjadi konsekuensi signifikan. Namun, berbeda dengan penelitian sebelumnya yang lebih banyak menggunakan model seperti SSD atau hanya mengevaluasi efisiensi kompresi seragam, penelitian ini memanfaatkan YOLOv8 dan mengevaluasi pengaruh kompresi ROI secara sistematis melalui pengukuran akurasi mAP serta pengujian statistik.

Penelitian ini juga memiliki kontribusi metodologis dan teknis yang penting. Sebelumnya, Gong et al. (2021) mengusulkan *Temporal ROI Align* untuk meningkatkan akurasi deteksi dalam video dengan memanfaatkan informasi temporal, namun belum menguji pendekatan tersebut pada model YOLOv8. Liu et al. (2022) mengembangkan metode kompresi berbasis *tensor train decomposition* khusus untuk YOLOv8, tetapi tidak membahas aspek ROI. Penelitian ini mengisi celah tersebut dengan mengevaluasi secara langsung efektivitas ROI pada YOLOv8, model yang lebih efisien dan akurat, serta memperkenalkan analisis batas toleransi CRF dalam konteks ROI, yang sebelumnya belum dibahas secara rinci dalam literatur.

Temuan ini memberikan dasar bagi pengembangan pendekatan ROI yang lebih adaptif. Salah satu kelemahan utama dari pendekatan yang digunakan dalam penelitian ini adalah penentuan ROI yang statis, sehingga rentan terhadap kondisi video dengan *noise* atau latar yang kompleks. Solusi yang dapat dikembangkan antara lain pelacakan ROI secara temporal atau agregasi *multi-frame*, yang dapat menjaga kualitas area penting secara dinamis. Adaptasi ROI berdasarkan pergerakan atau kepadatan objek juga dapat meningkatkan efisiensi tanpa mengorbankan performa deteksi. Dan Secara keseluruhan, hasil penelitian dapat diterapkan pada sistem deteksi objek berbasis video di dunia nyata, khususnya dalam aplikasi *real-time* seperti sistem pengawasan CCTV dan kendaraan otonom, yang menuntut keseimbangan antara efisiensi kompresi dan akurasi deteksi.

SIMPULAN

Penelitian ini menganalisis pengaruh kompresi video terhadap akurasi deteksi objek menggunakan model YOLOv8. Hasil penelitian menunjukkan bahwa kompresi seragam memberikan akurasi deteksi yang lebih baik pada tingkat kompresi CRF 32, dengan mAP sekitar 2,7% lebih tinggi dibandingkan dengan ROI yang hanya mencapai 0,46% pada tingkat kompresi yang sama. Meskipun ROI dapat mengurangi ukuran file video lebih banyak, penurunan akurasi deteksi objek pada tingkat kompresi yang lebih tinggi (CRF 45 dan 50) lebih signifikan, dengan mAP menurun lebih dari 70%. Berdasarkan temuan ini, kompresi seragam lebih disarankan untuk aplikasi yang mengutamakan akurasi tinggi, seperti sistem pengawasan CCTV dan kendaraan otonom. Penelitian selanjutnya disarankan untuk mengembangkan ROI adaptif, yang dapat menyesuaikan dengan pergerakan objek dan mengurangi penurunan kualitas deteksi objek akibat degradasi pada area non-ROI, sehingga efisiensi kompresi tetap terjaga tanpa mengorbankan akurasi.

REFERENSI

Achlison, U., Santoso, N. J. T., Rozikin, N. K., & Diapoldo, N. F. (2023). Analisis Latensi Video Streaming Antara Jaringan Berbasis Local Area Network dan Web. *Pixel Jurnal Ilmiah Komputer Grafis*, 15(2), 473–477. <https://doi.org/10.51903/pixel.v15i2.1037>

- Al, G. M. E. (2021). State-of-the-Art in Video Processing: Compression, Optimization, and Retrieval. *Türk Bilgisayar Ve Matematik Eğitimi Dergisi*, 12(5), 1256–1272. <https://doi.org/10.17762/turcomat.v12i5.1793>
- Alnoor, J. A. A., & Mustafa, A. B. A. N. (2022). Comparative Objective Analysis of Video Quality between H.265/HEVC and H.264/AVC. *University of Khartoum Engineering Journal*, 10(1), 1–10. <https://doi.org/10.53332/kuej.v10i1.914>
- Ariyanti, S., Setiawan, A. S., & Munandar, J. M. (2021). Study of Mobile Operator Readiness Measurement in Indonesia for 5G Technology Deployment. *Buletin Pos dan Telekomunikasi*, 19(2), 105–118. <https://doi.org/10.17933/bpostel.2021.190203>
- Gandor, T., & Nalepa, J. (2022). First Gradually, Then Suddenly: Understanding the Impact of Image Compression on Object Detection Using Deep Learning. *Sensors*, 22(3), 1104. <https://doi.org/10.3390/s22031104>
- Gong, T., Chen, K., Wang, X., Chu, Q., Zhu, F., Lin, D., Yu, N., & Feng, H. (2021). Temporal ROI Align for Video Object Recognition. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(2), 1442–1450. Vancouver, Kanada: Association for the Advancement of Artificial Intelligence. <https://doi.org/10.1609/aaai.v35i2.16234>
- Haidous, A., Oswald, W., Das, H., & Gong, N. (2022). Content-Adaptable ROI-Aware Video Storage for Power-Quality Scalable Mobile Streaming. *IEEE Access*, 10, 26830–26848. <https://doi.org/10.1109/access.2022.3156274>
- Hanooman, V., Kokaram, A. C., Su, Y., Birkbeck, N., & Adsumilli, B. (2021). The Effect of Degradation on Compressibility of Video. *SPIE Optical Engineering*, 35. <https://doi.org/10.1117/12.2593916>
- Khadir, M., Hashmi, M. F., Kotambkar, D. M., & Gupta, A. (2024). Innovative Insights: A Review of Deep Learning Methods for Enhanced Video Compression. *IEEE Access*, 12, 125706–125725. <https://doi.org/10.1109/access.2024.3450814>
- Kwon, M. J., Nam, S. H., Yu, I. J., Lee, H. K., & Kim, C. (2022). Learning JPEG Compression Artifacts for Image Manipulation Detection and Localization. *International Journal of Computer Vision*, 130(8), 1875–1895. <https://doi.org/10.1007/s11263-022-01617-5>
- Lakhan, N. R., & Verma, N. K. S. (2023). Mastering the Art of Video Conferencing: Remote Learning and Virtual Conferences. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, 5, 554–561. <https://doi.org/10.32628/cseit2390272>
- Li, B., Liang, J., Fu, H., & Han, J. (2023). ROI-Based Deep Image Compression With Swin Transformers. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5. Singapura: IEEE. <https://doi.org/10.1109/icassp49357.2023.10094674>
- Liu, H., Jin, F., Zeng, H., Pu, H., & Fan, B. (2023). Image Enhancement Guided Object Detection in Visually Degraded Scenes. *IEEE Transactions on Neural Networks and Learning Systems*, 35(10), 14164–14177. <https://doi.org/10.1109/tnnls.2023.3274926>
- Ma, Y., Zhai, Y., Yang, C., Yang, J., Wang, R., Zhou, J., Li, K., Chen, Y., & Wang, R. (2021). Variable Rate ROI Image Compression Optimized for Visual Quality. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1936–1940. New Orleans, USA: IEEE. <https://doi.org/10.1109/cvprw53098.2021.00221>
- O’Byrne, M., Sugrue, M., Vibhoothi, N., & Kokaram, A. (2022). Impact of Video Compression on the Performance of Object Detection Systems for Surveillance Applications. *2022 19th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 1–8. Singapore: IEEE. <https://doi.org/10.1109/avss56176.2022.9959476>
- Rajasekhar, J. (2024). Understanding YOLO: Real-Time Object Detection Explained. *International Journal of Scientific Research in Engineering and Management*, 8(7), 1–9. <https://doi.org/10.55041/ijssrem36359>

- Shinohara, Y., Itsumi, H., Florian, B., & Iwai, T. (2020). Video Compression Estimating Recognition Accuracy for Remote Site Object Detection. *2020 International Wireless Communications and Mobile Computing Conference (IWCMC)*, 285–290. Limassol, Cyprus: IEEE. <https://doi.org/10.1109/iwcmc48107.2020.9148347>
- Tubagus, A. S., Mahdi, R. S., Rizal, A., & Suharso, A. (2021). Analisis Perbandingan Teknik Video Codec H.264/AVC, H.265/HEVC, VP9 dan AV1. *Edumatic: Jurnal Pendidikan Informatika*, 5(2), 187–195. <https://doi.org/10.29408/edumatic.v5i2.3850>
- Terven, J., Córdova-Esparza, D.-M., & Romero-González, J.-A. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*, 5(4), 1680–1716. <https://doi.org/10.3390/make5040083>
- Yao, F., Jia, X., Wang, J., & Deng, K. (2022). EWF-Based Unequal Error Protection Scheme for ROI-Coded Images. In *Proceedings of the 2022 5th International Conference on Telecommunications and Communication Engineering (ICTCE '22)*, 238–243. Chengdu, China: IEEE. <https://doi.org/10.1145/3577065.3577108>
- Yu, J., Kim, Y., & Kim, Y. (2021). Intelligent Video Data Security: A Survey and Open Challenges. *IEEE Access*, 9, 26948–26967. <https://doi.org/10.1109/access.2021.3057605>