3Terbit online pada laman web jurnal: **http://journal.isas.or.id/index.php/JACOST**

JOURNAL OF APPLIED COMPUTER SCIENCE AND TECHNOLOGY (JACOST)

Vol. 6 No. 2 (2025) 149 – 161 | ISSN: 2723-1453 (Media Online)

Pengembangan Sistem Anti-Spoofing Berbasis Face Recognition Menggunakan Arsitektur YOLOv8n

Carmelita Angeline Tanujaya[1], Nur Fajri Azhar[2], Bowo Nugroho[3]
[1,2,3]Jurusan Teknik Elektro, Informatika dan Bisnis, Fakultas Sains dan Teknologi Informasi, Institut Teknologi Kalimantan
[1]carmlt.tan@gmail.com, [2]fajri@lecturer.itk.ac.id, [3]bowo.nugroho@lecturer.itk.ac.id

Abstract

Face spoofing poses a major threat to facial recognition–based authentication systems, especially in web-based environments that require lightweight and real-time verification. This study develops a real-time anti-spoofing system that integrates YOLOv8n for classifying four facial categories (real, printed, digital, and mask), combined with blink-based liveness verification using the Eye Aspect Ratio (EAR). Using 400,800 images and 18 videos, two training strategies—pretrained and from scratch—were evaluated. The pretrained model achieved a precision of 99.5%, recall of 98.6%, mAP50 of 99.4%, and mAP50–95 of 90.4%, slightly outperforming the from-scratch model. EAR threshold evaluation showed that a value of 0.17 yielded the best performance with 99.02% accuracy, 100% recall, a FAR of 16.11%, and an FRR of 0%. The proposed integration of YOLOv8n and EAR represents a practical novelty for lightweight, web-based anti-spoofing, providing fast inference and stable real-time performance suitable for modern facial authentication systems.

Keywords: anti-spoofing, eye aspect ratio, facial landmark, face recognition, real-time, YOLOv8n

Abstrak

Face spoofing merupakan ancaman bagi sistem autentikasi berbasis wajah, terutama pada aplikasi web yang membutuhkan proses verifikasi yang ringan dan *real-time*. Penelitian ini mengembangkan sistem *anti-spoofing real-time* yang mengintegrasikan YOLOv8n untuk mengklasifikasikan empat kategori wajah (*real, printed, digital,* dan *mask*) serta verifikasi *liveness* melalui deteksi kedipan berbasis *Eye Aspect Ratio* (EAR). Dengan menggunakan 400.800 citra dan 18 video, dua strategi pelatihan—*pretrained* dan *from scratch*—dievaluasi. Model pretrained mencapai precision 99,5%, recall 98,6%, mAP50 99,4%, dan mAP50–95 90,4%, sedikit lebih tinggi dibandingkan model from scratch. Evaluasi ambang EAR menunjukkan bahwa nilai 0,17 menghasilkan performa terbaik dengan akurasi 99,02%, recall 100%, FAR 16,11%, dan FRR 0%. Integrasi YOLOv8n dan EAR ini menjadi kontribusi praktis yang menawarkan solusi anti-spoofing berbasis web yang ringan dengan inferensi cepat dan performa *real-time* yang stabil untuk kebutuhan autentikasi wajah modern.

Kata kunci: Face Recognition, Anti-Spoofing, YOLOv8n, Eye Aspect Ratio, Facial Landmark, Real-Time

1. Pendahuluan

Teknologi pengenalan wajah banyak digunakan dalam sistem keamanan biometrik, seperti pengendalian akses, pembayaran digital, dan identifikasi pengguna. Namun, meningkatnya penggunaan teknologi ini disertai dengan bertambahnya serangan spoofing, di mana penyerang memalsukan identitas menggunakan foto, video, atau masker sehingga sistem salah mengenali objek palsu sebagai wajah asli [1]. Serangan seperti ini berpotensi dimanfaatkan untuk tujuan jahat dan dapat membahayakan keamanan sistem. Oleh karena itu, Face Anti-Spoofing (FAS) menjadi sangat penting untuk memastikan bahwa wajah yang dikenali oleh sistem benar-benar berasal dari individu secara langsung.

Sejumlah penelitian telah dilakukan untuk meningkatkan akurasi deteksi spoofing. Face-Fake-Net, misalnya, mencapai akurasi 99,7% pada CASIA-SURF dan 99,4% pada CelebA-Spoof [2]. Pendekatan lain yang mengombinasikan Histogram of Oriented Gradients (HOG), Support Vector Machine (SVM), dan deteksi kedipan mata dilaporkan mampu mencapai akurasi 92,68% [3]. Pendekatan berbasis YOLOv3–YOLOv5 juga menunjukkan hasil yang kompetitif, dengan akurasi mencapai 98,2% pada dataset CASIA-FASD [4]. Di sisi lain, beberapa metode FAS memanfaatkan sensor tambahan seperti depth atau NIR untuk meningkatkan akurasi, tetapi solusi ini membutuhkan perangkat keras khusus yang relatif mahal dan kurang praktis untuk aplikasi web maupun perangkat edge dengan keterbatasan komputasi [1], [5].

Performa deteksi pada edge device sangat bervariasi. Varian ringan YOLOv3-tiny yang direduksi mencapai sekitar 25.9 FPS pada CPU laptop, menunjukkan bahwa optimisasi arsitektur dapat meningkatkan efisiensi

inferensi pada perangkat terbatas [6]. Namun, pada perangkat dengan sumber daya yang lebih terbatas, pendekatan tradisional yang tidak dioptimalkan masih jauh dari real-time. Selain itu, model CNN konvensional cenderung sensitif terhadap perubahan pencahayaan, artefak visual, dan domain shift, sehingga performanya menurun ketika diuji pada dataset atau kondisi yang berbeda dari data pelatihan [7], [8]. Pendekatan berbasis video-level yang memanfaatkan isyarat temporal seperti gerakan bibir sering kali memiliki kompleksitas komputasi yang tinggi [9].

Perkembangan arsitektur YOLOv8 menghadirkan detektor anchor-free yang lebih efisien, cepat, dan akurat dibandingkan generasi sebelumnya [10] - [14]. Dari berbagai variannya, YOLOv8n merupakan model paling ringan sehingga ideal untuk implementasi real-time pada perangkat dengan keterbatasan komputasi. Di sisi lain, metode Eye Aspect Ratio (EAR) menawarkan pendekatan liveness yang sederhana dan praktis dengan memanfaatkan sinyal fisiologis berupa kedipan mata, tanpa memerlukan sensor tambahan.

Namun, penelitian sebelumnya umumnya masih berfokus pada arsitektur YOLO generasi terdahulu (misalnya YOLOv3–YOLOv5) atau pada skenario single-modal yang tidak terintegrasi dengan verifikasi liveness berbasis kedipan. Selain itu, pemanfaatan YOLOv8n secara khusus untuk Face Anti-Spoofing masih sangat terbatas, terutama dalam konteks deteksi multi-kategori yang digabungkan dengan liveness berbasis EAR pada lingkungan real-time berbasis web. Dengan demikian, terdapat gap penelitian dalam pengembangan sistem FAS yang mengombinasikan detektor objek ringan seperti YOLOv8n dengan metode liveness sederhana berbasis EAR untuk kebutuhan autentikasi wajah modern di aplikasi web.

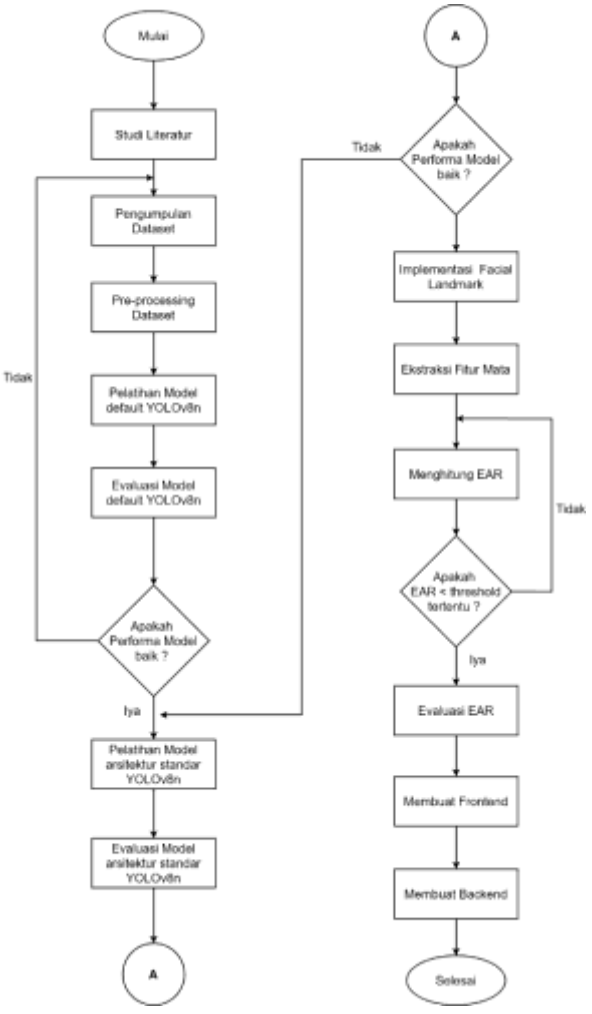
Selain aspek keamanan, kebutuhan aksesibilitas juga menjadi pertimbangan penting. Individu dengan keterbatasan fisik tertentu tidak selalu dapat memanfaatkan biometrik lain seperti sidik jari atau suara. Dalam kasus tersebut, autentikasi berbasis wajah sering menjadi pilihan yang paling realistis.

Penelitian ini bertujuan mengevaluasi performa YOLOv8n dalam mendeteksi empat kategori spoofing menggunakan dataset besar dan beragam, sekaligus membandingkan dua strategi pelatihan yaitu pretrained dan dari awal. Selain itu, penelitian ini mengintegrasikan YOLOv8n dengan metode liveness berbasis EAR untuk menghasilkan sistem anti-spoofing real-time berbasis web yang tidak memerlukan sensor khusus. Penelitian ini diharapkan dapat menghadirkan solusi anti-spoofing yang adaptif, praktis, dan sesuai dengan kebutuhan autentikasi wajah modern.

2. Metode Penelitian

2.1. Alur Penelitian

Penelitian ini dilaksanakan melalui beberapa tahapan sistematis yang dirancang untuk mengembangkan dan mengevaluasi sistem Face Anti-Spoofing berbasis YOLOv8n dengan deteksi liveness menggunakan Eye Aspect Ratio (EAR). Tahap awal meliputi studi literatur, pengumpulan dataset, dan praproses data. Tahap berikutnya adalah perancangan desain eksperimen dan skenario pelatihan, dilanjutkan dengan pelatihan model dan evaluasi performa deteksi spoofing. Setelah itu, dilakukan evaluasi liveness berbasis EAR dan integrasi keduanya ke dalam sistem berbasis web yang menggabungkan komponen frontend dan backend. Seluruh tahapan dirancang agar sistem yang dihasilkan ringan, efisien, dan mampu berjalan secara real-time melalui browser. Alur keseluruhan penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

2.2. YOLOv8

YOLO (You Only Look Once) merupakan algoritma deep learning untuk melakukan deteksi objek secara real time dengan memanfaatkan algoritma Convolutional

DOI : https://doi.org/10.52158/jacost.v6i2.1362

Neural Network (CNN) [15]. YOLOv8 mengadopsi pendekatan anchor-free, yang menggantikan sistem anchor boxes pada versi sebelumnya, sehingga menyederhanakan proses deteksi dan mengurangi kompleksitas komputasi.

Arsitektur YOLOv8 terdiri dari tiga komponen utama, yaitu Backbone untuk ekstraksi fitur menggunakan blok C2f, Neck untuk penggabungan fitur multi-skala melalui SPPF dan upsampling, serta Head yang menghasilkan prediksi bounding box dan klasifikasi objek. YOLOv8 hadir dalam beberapa varian, di mana YOLOv8n merupakan model paling ringan dengan jumlah parameter paling kecil, sehingga cocok untuk kebutuhan inferensi cepat di perangkat dengan daya komputasi terbatas.

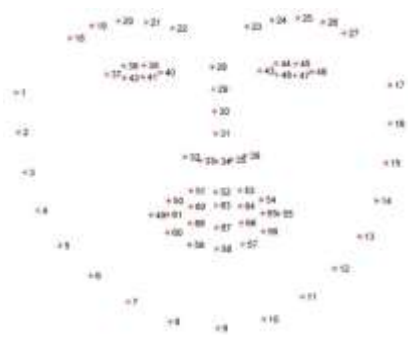
Penggunaan YOLOv8n pada penelitian ini didasarkan pada efisiensinya dalam memproses citra secara real-time serta peningkatan akurasi dan stabilitas yang telah dilaporkan dibandingkan YOLO generasi sebelumnya. Sifatnya yang ringan menjadikannya sesuai untuk diintegrasikan dalam sistem anti-spoofing berbasis web yang membutuhkan performa cepat tanpa beban komputasi tinggi. Perbandingan performa masing-masing varian dapat dilihat pada Gambar 2 [14].

Model	size (pixels)	mAP$^{val}_{50-95}$	Speed CPU ONNX (ms)	Speed A100 TensorRT (ms)	params (M)	FLOPs (B)
YOLOv8n	640	37.3	80.4	0.99	3.2	8.7
YOLOv8s	640	44.9	128.4	1.20	11.2	28.6
YOLOv8m	640	50.2	234.7	1.83	25.9	78.9
YOLOv8l	640	52.9	375.2	2.39	43.7	165.2
YOLOv8x	640	53.9	479.1	3.53	68.2	257.8

Gambar 2. Varian YOLOv8

2.3. Facial Landmark

Facial landmark adalah metode lokalisasi titik-titik yang menonjol pada wajah. Sebelum gambar diolah, gambar harus dideteksi apakah terdapat wajah [16]. Titik-titik ini digunakan untuk mengidentifikasi posisi mata sebagai dasar perhitungan Eye Aspect Ratio (EAR). Penggunaan facial landmark diperlukan karena EAR bergantung pada jarak vertikal dan horizontal kelopak mata, sehingga lokalisasi titik mata yang akurat menjadi langkah penting dalam proses deteksi kedipan. Ilustrasi penyebaran titik-titik landmark wajah dapat dilihat pada Gambar 3 [17].



Gambar 3. Lokasi 68 Titik Landmark Wajah

2.4. Eye Aspect Ratio (EAR)

Eye Aspect Ratio (EAR) adalah nilai skalar yang digunakan untuk membedakan kondisi mata terbuka dan tertutup berdasarkan perubahan jarak vertikal dan horizontal kelopak mata [18]. Nilai EAR cenderung stabil ketika mata terbuka dan menurun signifikan ketika mata tertutup, sehingga indikator ini efektif untuk mendeteksi kedipan secara real-time. EAR dihitung menggunakan enam titik landmark pada area mata, seperti ditunjukkan pada Persamaan 1.

$$EAR = \frac{||P2-P6||+||P3-P5||}{2\,||P1-P4||} \quad (1)$$

Titik P1–P4 menggambarkan jarak horizontal mata, sedangkan P2–P6 dan P3–P5 mewakili perubahan vertikal kelopak mata. Ilustrasi struktur titik-titik tersebut ditampilkan pada Gambar 4 [19].

Gambar 4. Eye Aspect Ratio

EAR kemudian dibandingkan dengan beberapa nilai ambang (threshold) untuk menentukan kondisi mata. Evaluasi performa liveness detection dilakukan menggunakan tiga metrik utama sebagaimana pada Persamaan 2–4: Accuracy, False Acceptance Rate (FAR), dan False Rejection Rate (FRR).

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$FAR = \frac{FP}{FP+TN} \quad (3)$$

$$FRR = \frac{FN}{FN+TP} \quad (4)$$

Accuracy mengukur tingkat keseluruhan prediksi yang benar. FAR (False Acceptance Rate) menunjukkan seberapa sering mata tertutup salah diklasifikasikan sebagai terbuka, sedangkan FRR (False Rejection Rate) menunjukkan seberapa sering mata terbuka salah diklasifikasikan sebagai tertutup. Ketiga metrik tersebut digunakan untuk menentukan threshold EAR yang paling stabil dan akurat pada skenario real-time.

2.5. Frontend

Frontend berfungsi sebagai antarmuka utama yang menangani pengambilan video dari kamera perangkat menggunakan API *getUserMedia* dan menampilkan hasil deteksi secara real-time. Frame video yang diambil dikonversi ke format Base64 dan dikirim ke backend melalui permintaan HTTP POST secara *asynchronous* agar proses deteksi tidak mengganggu tampilan video.

HTML, CSS, dan JavaScript digunakan karena kompatibel dengan seluruh browser modern serta mendukung integrasi langsung dengan kamera. JavaScript mengatur pengiriman frame, menampilkan bounding box dan label klasifikasi, serta memperbarui status liveness berdasarkan respons JSON dari backend.

DOI : https://doi.org/10.52158/jacost.v6i2.1362

Desain ini memastikan frontend tetap ringan dan responsif pada berbagai perangkat.

Melalui desain antarmuka yang sederhana dan interaktif, frontend memungkinkan pengguna melihat hasil deteksi spoofing dan kedipan mata secara real-time. Dengan demikian, frontend tidak hanya berperan sebagai tampilan visual, tetapi juga sebagai komponen penting dalam memastikan alur autentikasi berjalan cepat dan konsisten pada aplikasi berbasis web.

2.6. Backend

Backend berfungsi sebagai pusat pemrosesan utama yang menjalankan inferensi deteksi spoofing dan liveness. Setiap frame yang dikirimkan dari frontend diterima melalui endpoint REST API, kemudian melalui proses prapemrosesan sebelum diolah oleh model YOLOv8n. Model menghasilkan deteksi wajah beserta kategorinya, yaitu real, printed, digital, atau mask. Jika wajah terdeteksi sebagai "real", backend mengekstraksi titik facial landmark pada area mata dan menghitung Eye Aspect Ratio (EAR) untuk menilai kondisi kedipan sebagai verifikasi liveness.

Backend dikembangkan menggunakan framework Flask karena sifatnya yang ringan, fleksibel, dan efisien untuk memenuhi kebutuhan inferensi real-time pada aplikasi web. Flask memungkinkan pengaturan routing yang sederhana, integrasi pustaka deep learning berbasis Python, serta latensi rendah dalam pemrosesan citra. Seluruh proses inferensi dilakukan di sisi server untuk menjaga keamanan model dan memastikan performa konsisten pada berbagai perangkat pengguna.

Backend mengirimkan hasil deteksi kepada frontend dalam bentuk respons JSON yang memuat label klasifikasi, nilai *confidence*, dan status kedipan mata. Dengan demikian, backend berperan penting dalam memastikan alur autentikasi berjalan cepat, akurat, dan stabil pada sistem anti-spoofing berbasis web.

2.7. Dataset and Preprocessing

Dataset dalam penelitian ini terdiri dari 400.800 citra dan 18 video yang mencakup empat kategori spoofing, yaitu real, printed, digital, dan mask. Seluruh citra diperoleh dari kombinasi dataset terbuka seperti CelebA-Spoof, iBeta, dan beberapa dataset Kaggle [20]-[39] yang menyediakan keragaman etnis, kondisi pencahayaan, pose wajah, serta variasi teknik serangan. Selain itu, data video direkam secara mandiri untuk menyediakan sampel kedipan yang diperlukan dalam evaluasi liveness berbasis EAR. Keberagaman ini memastikan bahwa model dapat belajar dan melakukan generalisasi pada berbagai kondisi autentikasi di dunia nyata.

Tahap preprocessing dilakukan untuk memastikan kualitas dan konsistensi data yang masuk ke model YOLOv8n. Wajah terlebih dahulu diekstraksi dari setiap frame menggunakan OpenCV dan modul deteksi wajah. Citra kemudian dipotong dengan margin tertentu untuk mempertahankan fitur penting pada area wajah. Deteksi blur dilakukan menggunakan metode variansi Laplacian untuk menghapus citra yang tidak fokus. Setiap citra yang lolos seleksi diberi anotasi dalam format YOLO, termasuk class ID dan koordinat bounding box yang ter-normalisasi. Contoh proses deteksi wajah ditampilkan pada Gambar 5.

Gambar 5. Proses Deteksi Wajah

Dataset dibagi menjadi tiga subset, yaitu 70% untuk pelatihan, 10% untuk validasi, dan 20% untuk pengujian. Pembagian ini dilakukan secara proporsional pada setiap kategori spoofing untuk mencegah bias kelas dan memastikan evaluasi model lebih representatif. Struktur dataset difinalisasi dalam berkas konfigurasi YAML agar dapat langsung terintegrasi dengan pipeline pelatihan YOLOv8n.

Secara keseluruhan, proses preprocessing dirancang untuk menghasilkan data yang bersih, terstruktur, dan konsisten sehingga model dapat belajar pola spoofing secara efektif serta mempertahankan performa yang stabil pada tahap pengujian.

2.8. Desain Eksperimen

Desain eksperimen penelitian ini mencakup dua komponen utama, yaitu evaluasi deteksi spoofing menggunakan YOLOv8n dan evaluasi liveness detection berbasis Eye Aspect Ratio (EAR). Seluruh konfigurasi pelatihan dan pemilihan parameter ditetapkan berdasarkan pertimbangan teknis serta acuan penelitian terdahulu.

Pada deteksi spoofing, dua skenario pelatihan digunakan: *fine-tuning pretrained weights* dan *training from scratch*. Pendekatan pretrained dipilih untuk memperoleh konvergensi yang lebih cepat dan stabil karena model telah memiliki representasi awal dari dataset COCO. Sebaliknya, pelatihan dari awal digunakan untuk menilai kemampuan model mempelajari pola spoofing secara murni tanpa bias dari dataset umum. Perbandingan kedua skenario ini digunakan untuk mengetahui efektivitas transfer learning terhadap performa akhir model.

Parameter pelatihan ditetapkan agar seimbang antara akurasi dan efisiensi. Resolusi input 640×640 digunakan karena merupakan konfigurasi default YOLOv8 yang efektif pada varian ringan. Pelatihan dilakukan selama

DOI : https://doi.org/10.52158/jacost.v6i2.1362

100 epoch berdasarkan hasil percobaan awal yang menunjukkan konvergensi pada rentang epoch 70–90, sedangkan batch size 16 dipilih untuk menjaga stabilitas pembaruan bobot pada GPU RTX 3060. Pemilihan varian YOLOv8n didasarkan pada jumlah parameternya yang kecil sehingga mampu mendukung inferensi real-time pada sistem berbasis web.

Evaluasi model dilakukan menggunakan metrik Precision, Recall, mAP50, dan mAP50–95 untuk menggambarkan kualitas deteksi pada berbagai tingkat akurasi bounding box. Selain evaluasi berbasis dataset, model juga diuji secara real-time menggunakan kamera untuk menilai ketahanannya terhadap variasi pencahayaan, pose, dan aksesori wajah

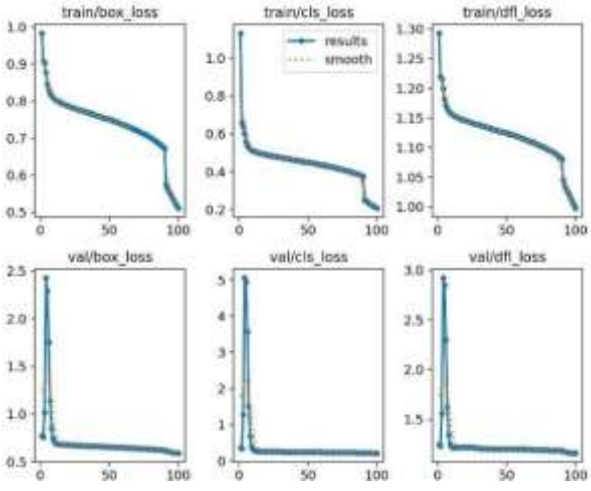
Komponen liveness detection diuji menggunakan empat nilai threshold EAR, yaitu 0.15, 0.17, 0.20, dan 0.25. Rentang threshold ini dipilih berdasarkan temuan penelitian terdahulu [18], [40], yang menunjukkan bahwa nilai EAR mata terbuka umumnya berada pada rentang 0.20–0.30, sedangkan EAR mata tertutup berada di bawah 0.18. Setiap threshold diuji menggunakan metrik Accuracy, False Acceptance Rate (FAR), dan False Rejection Rate (FRR) untuk menentukan ambang yang paling stabil dalam membedakan kondisi mata pada penggunaan real-time.

Secara keseluruhan, desain eksperimen ini memastikan bahwa kedua komponen, yaitu deteksi spoofing dan liveness detection, diuji secara sistematis dan berbasis justifikasi ilmiah sehingga valid untuk implementasi sistem anti-spoofing berbasis web.

3. Hasil dan Pembahasan

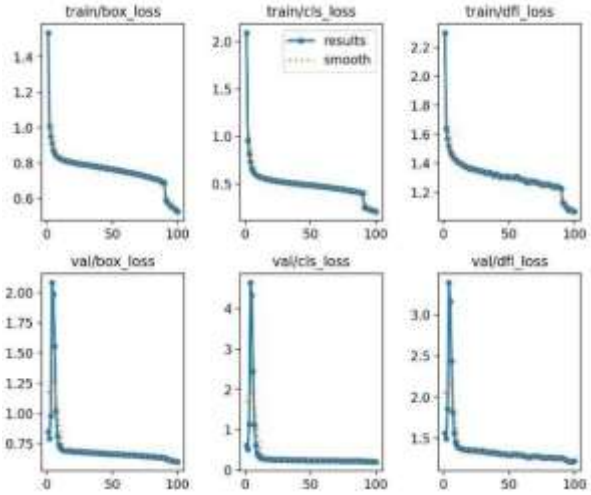
3.1. Pelatihan Model YOLOv8n

Pelatihan dilakukan menggunakan dua skenario, yaitu fine-tuning bobot pre-trained COCO dan pelatihan dari awal tanpa bobot awal. Kurva hasil pelatihan model dengan pre-trained dapat dilihat pada Gambar 6 dan Gambar 7.



Gambar 6. Kurva Loss Pelatihan Model YOLOv8n Pre-trained

Gambar 6 menyajikan kurva loss hasil pelatihan model YOLOv8n dengan bobot pre-trained, yang terdiri dari box_loss, cls_loss, dan dfl_loss pada data pelatihan dan validasi. Kurva-kurva tersebut menunjukkan penurunan tajam pada 10–15 epoch pertama sebelum mendatar pada nilai yang rendah hingga epoch ke-100. Jarak antara loss pelatihan dan validasi relatif kecil dan tidak terlihat peningkatan loss yang berkepanjangan pada data validasi. Pola ini menunjukkan bahwa proses optimisasi berjalan stabil, model tidak mengalami overfitting yang signifikan, dan posisi bounding box serta distribusinya dapat dipelajari dengan baik. Lonjakan kecil pada val/box_loss dan val/cls_loss di awal epoch mencerminkan fase adaptasi terhadap distribusi data spoofing yang lebih beragam dibandingkan COCO, namun setelah itu kurva cepat turun dan stabil pada nilai yang konsisten.



Gambar 7. Kurva Loss Pelatihan Model YOLOv8n Tanpa Pre-trained

Gambar 7 menampilkan kurva loss untuk box_loss, cls_loss, dan dfl_loss pada data pelatihan dan validasi. Secara umum, bentuk kurva mirip dengan model pre-trained, tetapi fase penurunan awal berlangsung sedikit lebih lama. Hal ini wajar karena model tidak membawa pengetahuan awal dari COCO dan harus membangun representasi fitur dari nol. Meskipun demikian, loss pada data pelatihan dan validasi tetap turun hingga mencapai nilai rendah yang stabil, dengan gap yang kecil di antara keduanya. Pola ini menunjukkan bahwa arsitektur YOLOv8n yang anchor-free dengan blok C2f dan modul SPPF mampu memfasilitasi pembelajaran fitur secara efisien bahkan tanpa bobot awal, selama didukung oleh dataset yang cukup besar dan beragam.

Rata-rata mAP50 pada Tabel 1 mencapai 99,5% baik pada data pelatihan maupun pengujian, sedangkan mAP50–95 berada di kisaran 91,2% untuk pelatihan dan 90,4% untuk pengujian. Kelas real menunjukkan kinerja terbaik dengan mAP50–95 sekitar 95–96%, yang mengindikasikan bahwa pola tekstur dan kontur wajah manusia asli relatif konsisten dan mudah dipelajari oleh model. Sebaliknya, kelas digital mencatat mAP50–95

DOI : https://doi.org/10.52158/jacost.v6i2.1362

terendah, yaitu sekitar 88–89%. Perbedaan ini dapat dijelaskan oleh karakteristik tampilan layar yang sering memunculkan artefak visual seperti pantulan cahaya, pola moiré, dan variasi brightness yang ekstrem. Artefak tersebut cenderung mengganggu pola gradien dan kontur tepi wajah sehingga batas objek pada layar menjadi kurang jelas, membuat detektor berbasis bounding box seperti YOLOv8n lebih sulit memetakan objek dengan presisi tinggi.

Tabel 1. Hasil Pelatihan Model YOLOv8n Pre-trained Model

Kelas	mAP50 Train	mAP50-95 Train	mAP50 Test	mAP50-95 Test
real	99,5%	96,2%	99,5%	95,8%
printed	99,4%	89,0%	99,4%	88,1%
digital	99,5%	88,5%	99,4%	87,4%
mask	99,5%	91,1%	99,5%	90,0%
Rata-rata	99,5%	91,2%	99,4%	90,4%

Tabel 2 merangkum hasil pelatihan untuk skenario tanpa bobot pre-trained. Rata-rata mAP50 yang diperoleh sebesar 99,4% baik pada data pelatihan maupun pengujian, sedangkan mAP50–95 berada di kisaran 90,8% untuk pelatihan dan 90,1% untuk pengujian. Nilai-nilai ini hanya sedikit lebih rendah dibandingkan model pre-trained dan tetap menunjukkan ketelitian spasial yang tinggi dalam memprediksi posisi dan ukuran wajah. Pola per kelas juga konsisten dengan skenario pre-trained: kelas real kembali mencatat mAP50–95 tertinggi sekitar 95–96%, sedangkan kelas digital menjadi yang terendah sekitar 87–88%. Konsistensi tren ini menguatkan dugaan bahwa tantangan utama pada kelas digital berasal dari karakteristik intrinsik media tampilan, bukan dari perbedaan strategi pelatihan. Selain itu, kedekatan nilai mAP antara data pelatihan dan pengujian pada kedua skenario menunjukkan bahwa model mampu melakukan generalisasi dengan baik dan tidak terjebak pada overfitting yang berat.

Tabel 2. Hasil Pelatihan Model YOLOv8n Tanpa Pre-trained Model

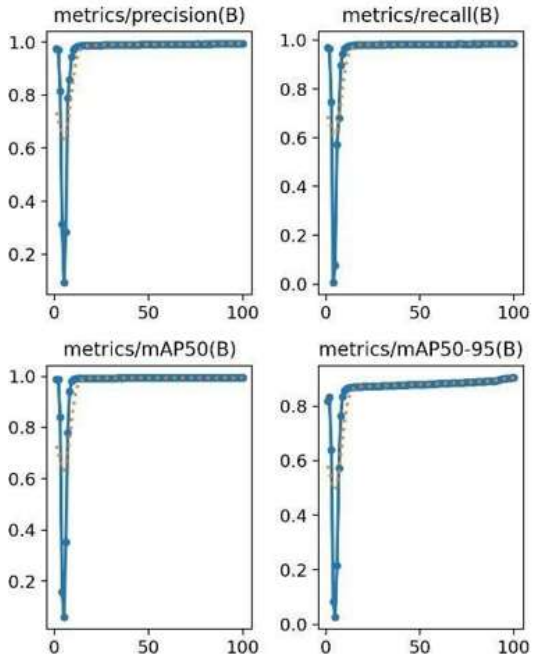
Kelas	mAP50 Train	mAP50-95 Train	mAP50 Test	mAP50-95 Test
real	99,5%	96,0%	99,5%	95,7%
printed	99,3%	88,6%	99,3%	87,7%
digital	99,4%	87,9%	99,4%	87,1%
mask	99,5%	90,6%	99,5%	89,8%
Rata-rata	99,4%	90,8%	99,4%	90,1%

Secara keseluruhan, hasil pada pelatihan model menunjukkan bahwa kedua skenario pelatihan menghasilkan performa yang sangat tinggi dan stabil. Penggunaan bobot pre-trained memberikan keuntungan utama berupa konvergensi yang lebih cepat, tetapi perbedaan mAP50 dan mAP50–95 akhir antara kedua model relatif kecil. Perbedaan kinerja antar kelas lebih banyak dipengaruhi oleh sifat tekstur dan artefak visual masing-masing kategori, khususnya pada kelas digital yang dipengaruhi pantulan dan pola layar, dibandingkan oleh konfigurasi pelatihan. Temuan ini mengindikasikan bahwa YOLOv8n memiliki kapasitas representasi yang memadai untuk tugas face anti-spoofing multi-kategori dan layak dijadikan basis untuk analisis evaluasi model dan perbandingan dengan studi terdahulu pada subbab berikutnya.

3.2. Evaluasi Model

Evaluasi dilakukan pada dua model YOLOv8n—pre-trained dan from-scratch—menggunakan 40.080 citra data uji (10% dari total dataset) yang mencakup empat kategori spoofing: real, printed, digital, dan mask. Penilaian performa dilakukan menggunakan precision, recall, mAP50, dan mAP50–95 untuk menilai kemampuan klasifikasi model.

Gambar 8 menampilkan evolusi metrik precision, recall, mAP50, dan mAP50–95 selama pelatihan untuk model dengan bobot pre-trained. Kurva pada Gambar 8 menunjukkan bahwa seluruh metrik naik sangat cepat pada awal pelatihan dan kemudian berada pada plateau tinggi hingga epoch ke-100. Precision dan recall mendekati nilai 1,0, menandakan bahwa jumlah false positive dan false negative sangat kecil. Nilai mAP50 mencapai level tinggi sejak awal, sedangkan mAP50–95 meningkat lebih bertahap sebelum stabil di kisaran 0,90. Pola ini konsisten dengan teori transfer learning, di mana bobot COCO menyediakan fitur generik yang kuat sehingga model hanya perlu menyesuaikan ke karakteristik spoofing wajah.

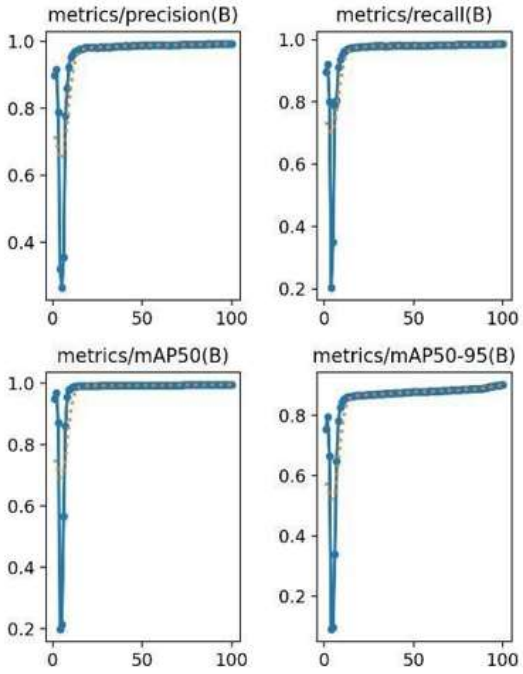


Gambar 8. Hasil Evaluasi Pelatihan Model YOLOv8n Pre-trained

Gambar 9 menyajikan metrik yang sama untuk model yang dilatih tanpa bobot pre-trained. Bentuk kurva pada Gambar 9 mirip dengan Gambar 7, tetapi fase kenaikan awal sedikit lebih panjang. Precision dan recall baru benar-benar stabil mendekati 0,99 setelah sekitar dua puluh epoch, sementara mAP50–95 membutuhkan

DOI : https://doi.org/10.52158/jacost.v6i2.1362

beberapa epoch tambahan sebelum mencapai nilai mendekati 0,90. Perbedaan pola ini menunjukkan bahwa model from-scratch membutuhkan waktu lebih lama untuk membangun representasi fitur, tetapi pada akhirnya mampu menyamai kualitas model pre-trained karena terbantu oleh ukuran dan keragaman dataset.



Gambar 9. Hasil Evaluasi Pelatihan Model YOLOv8n Tanpa Pre-trained

Hasil evaluasi pada data validasi untuk model pre-trained ditampilkan pada Tabel 3. Tabel 3 menyajikan nilai precision, recall, mAP50, dan mAP50–95 per kelas. Nilai rata-rata precision mencapai 99,5% dan recall 98,6%, dengan mAP50 sebesar 99,4% dan mAP50–95 sebesar 90,4%. Kelas real memiliki kombinasi nilai tertinggi, yaitu precision dan recall 99,9% serta mAP50–95 sekitar 95,9%. Hal ini menunjukkan bahwa pola tekstur dan kontur wajah asli relatif konsisten dan mudah dipisahkan dari serangan spoofing. Kelas mask juga mencatat performa tinggi dengan mAP50–95 sekitar 90,2%, karena topeng memiliki pola tepi dan tekstur material yang berbeda dari kulit manusia sehingga lebih mudah dikenali detector.

Tabel 3. Hasil Evaluasi Model YOLOv8n Dengan Pre-trained Model

Kelas	Precision	Recall	mAP50	mAP50-95
real	99,9%	99,9%	99,5%	95,9%
printed	98,7%	98,4%	99,4%	87,9%
digital	99,5%	97,6%	99,4%	87,6%
mask	99,8%	98,6%	99,5%	90,2%
Rata-rata	99,5%	98,6%	99,4%	90,4%

Tabel 4 menyajikan hasil evaluasi untuk model YOLOv8n yang dilatih tanpa bobot pre-trained. Secara rata-rata, precision sebesar 99,2% dan recall 98,7%, sedangkan mAP50 dan mAP50–95 masing-masing berada pada 99,4% dan 90,4%. Performa rata-rata ini hampir identik dengan model pre-trained. Kelas real kembali menjadi yang paling tinggi dengan mAP50–95 sekitar 95,7%, sementara kelas digital kembali menjadi yang terendah dengan mAP50–95 sekitar 87,3–87,7%. Konsistensi pola ini menunjukkan bahwa perbedaan utama antar kelas lebih ditentukan oleh karakteristik visual tiap kategori, bukan oleh skenario pelatihan yang digunakan.

Tabel 4. Hasil Evaluasi Model YOLOv8n Tanpa Pre-trained Model

Kelas	Precision	Recall	mAP50	mAP50-95
real	99,8%	99,8%	99,5%	95,7%
printed	98,1%	98,5%	99,3%	87,7%
digital	99,2%	97,9%	99,4%	87,3%
mask	99,7%	98,7%	99,5%	90,0%
Rata-rata	99,2%	98,7%	99,4%	90,4%

Performa kelas digital yang secara konsisten sedikit lebih rendah dibanding kelas lain dapat dijelaskan dari sisi karakteristik data. Wajah digital direkam dari tampilan layar ponsel atau monitor yang menghasilkan berbagai artefak, seperti pantulan cahaya, pola moiré akibat interaksi piksel layar dengan sensor kamera, serta variasi brightness yang ekstrem. Artefak ini membuat batas kontur wajah dan gradasi tekstur kulit menjadi kurang jelas dibanding wajah asli atau printed, sehingga prediksi bounding box dan klasifikasi menjadi sedikit lebih sulit. Selain itu, beberapa sampel digital memiliki resolusi efektif lebih rendah dibanding kelas lainnya, karena wajah yang ditampilkan di layar sering hanya menempati sebagian kecil area frame. Kondisi ini secara teoritis menurunkan rasio signal-to-noise fitur lokal yang dibutuhkan oleh head deteksi YOLOv8n.

Tabel 5 menunjukkan bahwa model YOLOv8n dalam penelitian ini mencapai recall yang lebih tinggi dan precision yang sebanding dengan hasil YOLOv3–YOLOv5 pada penelitian Vardhan et al. (2025). Pada model pre-trained, nilai precision berada pada tingkat yang sama dengan Vardhan, namun recall meningkat cukup signifikan, mengindikasikan kemampuan deteksi spoof yang lebih baik. Sementara itu, model from-scratch menghasilkan precision yang sedikit lebih rendah, tetapi tetap memberikan recall yang lebih tinggi. Perbaikan recall ini sangat relevan dalam konteks anti-spoofing, karena recall yang tinggi membantu meminimalkan kelolosan serangan.

Tabel 5. Perbandingan Kinerja Model Penelitian Ini dengan R. Vishnu Sai Vardhan dkk. (2025)

Model/Studi	Precision	Recall
Proposed YOLOv8n (Pre-trained)	99,5%	98,6%
Proposed YOLOv8n (From Scratch)	99,2%	98,7%
YOLOv3–YOLOv5 (Vardhan et al., 2025)	99,5%	97,6%

DOI : https://doi.org/10.52158/jacost.v6i2.1362

Kemampuan model dalam skenario penggunaan nyata dievaluasi melalui pengujian real-time berbasis kamera, yang hasil visualnya dirangkum pada Tabel 6.

Tabel 6. Hasil Pengujian Model YOLOv8n

Real	Printed	Digital	Mask

Pengujian dilakukan terhadap 10 partisipan dengan karakteristik yang beragam (jenis kelamin, etnis, dan kondisi pencahayaan berbeda). Meskipun jumlah partisipan relatif terbatas, uji ini bertujuan untuk mengevaluasi performa sistem secara praktis pada kondisi dunia nyata, bukan untuk analisis statistik populasi. Validasi generalisasi model telah dilakukan sebelumnya melalui dataset pelatihan yang besar dan beragam, sehingga jumlah partisipan dianggap memadai untuk tahap verifikasi sistem real-time.

Tabel 6 menampilkan contoh keluaran sistem untuk sepuluh partisipan pada empat kondisi, yaitu wajah asli, printed, digital, dan mask, baik untuk model dengan maupun tanpa bobot pre-trained. Pada sebagian besar contoh pada Tabel 6, bounding box dan label kelas yang dihasilkan sudah konsisten dengan kondisi sebenarnya dengan confidence di atas 0,8. Rata-rata confidence untuk kelas real berada di kisaran 94–95%, printed dan mask sekitar 92–93%, sedangkan digital berkisar 86–87%. Pola ini sejalan dengan temuan kuantitatif pada Tabel 3 dan Tabel 4, yaitu bahwa kelas digital merupakan kasus paling menantang, sementara kelas real dan mask relatif lebih mudah diidentifikasi.

Selain variasi kondisi wajah dan serangan spoofing, pengujian juga mencakup variasi pencahayaan, termasuk kondisi *low-light* dan pencahayaan tidak merata. Hasil pengujian menunjukkan bahwa model YOLOv8n tetap mampu mendeteksi wajah dengan confidence di atas 90% pada kondisi pencahayaan redup, sehingga dapat disimpulkan bahwa sistem memiliki ketahanan yang baik terhadap variasi intensitas cahaya pada proses autentikasi real-time.

Sistem juga diuji pada partisipan yang menggunakan kacamata untuk mengetahui pengaruh *occlusion* terhadap performa deteksi. Hasil pengujian menunjukkan bahwa keberadaan kacamata tidak memberikan dampak signifikan terhadap klasifikasi wajah, di mana wajah asli tetap terdeteksi sebagai kategori REAL dengan confidence di atas 93%. Hal ini mengindikasikan bahwa model YOLOv8n mampu mengekstraksi fitur struktural wajah secara konsisten meskipun sebagian area mata tertutup oleh kacamata

Selama proses pengujian ditemukan beberapa kasus kesalahan klasifikasi. Salah satu contohnya adalah ketika wajah asli yang ditampilkan dalam posisi miring tidak berhasil dikenali oleh sistem dan memunculkan prediksi yang tidak tepat. Selain itu, terdapat pula kondisi di mana wajah palsu, seperti gambar pada media cetak (printed), justru diklasifikasikan sebagai REAL. Kesalahan tersebut bersifat sesaat dan tidak terjadi secara terus-menerus, karena pada frame-frame berikutnya prediksi model kembali sesuai dengan label yang seharusnya. Temuan ini menunjukkan bahwa meskipun model secara umum memberikan hasil yang baik, tetap terdapat potensi misclassifications yang perlu diperhatikan, terutama pada kondisi input yang tidak

DOI : https://doi.org/10.52158/jacost.v6i2.1362

ideal atau menyerupai karakteristik wajah asli secara visual.

Meskipun performa deteksi spoofing dan liveness pada penelitian ini menunjukkan hasil yang sangat baik, terdapat beberapa batasan yang perlu diperhatikan untuk interpretasi dan penerapan di dunia nyata.

Dataset yang digunakan sudah cukup besar dan beragam, namun tetap berpotensi mengandung bias tertentu terkait distribusi etnis, variasi perangkat perekam, dan kondisi pencahayaan yang mungkin belum sepenuhnya mewakili seluruh populasi pengguna. Selain itu, penelitian ini belum secara eksplisit mengevaluasi ketahanan model terhadap serangan berbasis deepfake yang memanfaatkan manipulasi temporal yang lebih kompleks, sehingga generalisasi sistem pada kategori serangan tersebut perlu diuji lebih lanjut. Kinerja pada perangkat berbeda, khususnya perangkat mobile atau edge dengan kapasitas komputasi sangat terbatas, juga berpotensi bervariasi meskipun YOLOv8n dirancang ringan. Oleh karena itu, pengujian lanjutan pada berbagai konfigurasi perangkat, kondisi lingkungan, dan tipe serangan yang lebih modern menjadi langkah penting untuk memperkuat validitas dan ketangguhan sistem.

3.3. EAR

Eye Aspect Ratio (EAR) digunakan untuk membedakan kondisi mata terbuka dan tertutup berdasarkan jarak vertikal dan horizontal kelopak mata. Dalam setiap frame, sistem mendeteksi wajah menggunakan YOLOv8n, kemudian mengekstrak landmark mata kiri dan kanan. Enam titik landmark digunakan untuk menghitung EAR, sehingga perubahan geometri mata dapat dilacak secara konsisten. Ketika mata mulai menutup, jarak vertikal antar kelopak mata menurun jauh lebih cepat dibanding jarak horizontalnya, sehingga nilai EAR mengikuti pola penurunan yang signifikan. Mekanisme ini membuat EAR menjadi indikator fisiologis yang relevan untuk sistem anti-spoofing modern.

Pada penelitian ini, 18 video menghasilkan 16.630 frame yang kemudian diberi pseudo-label otomatis berdasarkan ambang awal EAR sebesar 0,18, mengacu pada nilai umum yang digunakan dalam literatur deteksi kedipan [18]. Frame dengan EAR < 0,18 dikategorikan sebagai mata tertutup, sedangkan nilai ≥ 0,18 ditetapkan sebagai mata terbuka. Pseudo-label ini digunakan sebagai acuan evaluasi untuk beberapa variasi threshold prediksi (0.15, 0.17, 0.20, 0.25). Hasil pengujian ditampilkan pada Tabel 7.

Tabel 7. Hasil Evaluasi EAR Berdasarkan Variasi Threshold

Threshold	Accuracy	Precision	Recall	FAR	FRR
0.15	97,52%	97,42%	100%	40,81%	0%
0.17	99,02%	98,97%	100%	16,11%	0%
0.20	97,87%	100%	97,73%	0%	2,27%
0.25	87,65%	100%	86,85%	0%	13,15%

Hasil pada Tabel 7 menunjukkan bahwa setiap threshold memberikan karakteristik performa yang berbeda. Pada threshold 0.15, nilai recall mencapai 100%, tetapi FAR sangat tinggi (40,81%). Kondisi ini terjadi karena ambang batas yang terlalu rendah membuat sistem terlalu sensitif, sehingga penurunan EAR kecil akibat bayangan, pantulan cahaya, atau perubahan pose dianggap sebagai kedipan. Akibatnya, banyak mata terbuka salah terdeteksi sebagai tertutup.

Saat threshold dinaikkan menjadi 0.25, terjadi situasi sebaliknya. Sistem menjadi lebih ketat sehingga hanya kedipan dengan penurunan EAR yang jelas yang dapat dikenali. Meskipun FAR menjadi 0%, FRR naik hingga 13,15%, menunjukkan bahwa sebagian kedipan nyata tidak terdeteksi. Hal ini dapat disebabkan oleh variasi fisiologis kedipan pengguna, misalnya sebagian orang memiliki kedipan yang lebih kecil atau tidak menutup mata secara penuh.

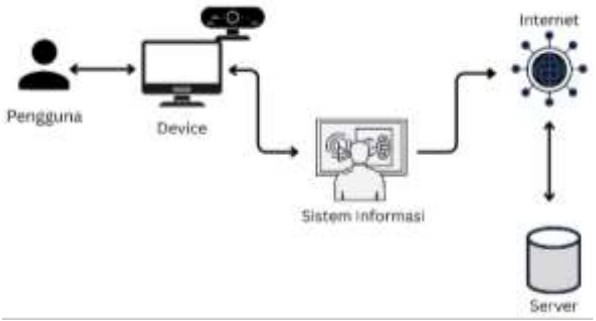
Threshold 0.20 menunjukkan kinerja yang lebih seimbang, ditandai dengan FAR dan FRR yang rendah. Namun, recall menurun menjadi 97,73%, yang berarti sebagian kecil kedipan masih terlewat. Sementara itu, threshold 0.17 memberikan performa paling stabil dan konsisten di semua metrik, ditunjukkan oleh accuracy tertinggi sebesar 99,02%, precision sebesar 98,97%, recall tetap sempurna pada 100%, serta FAR yang jauh lebih rendah dibandingkan threshold 0.15. FRR juga berada pada nilai 0%, sehingga tidak ada kedipan asli yang diabaikan oleh sistem.

Secara matematis, pemilihan threshold 0.17 dapat dijustifikasi dari pola distribusi EAR. Nilai EAR untuk mata terbuka pada sebagian besar pengguna berada pada kisaran 0.20–0.30, sedangkan nilai EAR untuk mata tertutup biasanya berada di bawah 0.18. Dengan meletakkan threshold pada 0.17, sistem berada tepat pada area pemisah antara dua distribusi tersebut sehingga tumpang tindih (overlap) menjadi minimal. Pemisahan ini mengurangi peluang kedua kelas saling salah diklasifikasikan. Hasil tersebut menunjukkan bahwa threshold 0.17 merupakan titik optimal yang meminimalkan trade-off sensitivitas (recall) dan ketelitian (precision), sekaligus menjaga stabilitas deteksi pada kondisi nyata yang dipengaruhi noise kamera dan variasi fisiologi pengguna.

DOI : https://doi.org/10.52158/jacost.v6i2.1362

3.4. Pembangunan Sistem Berbasis Web

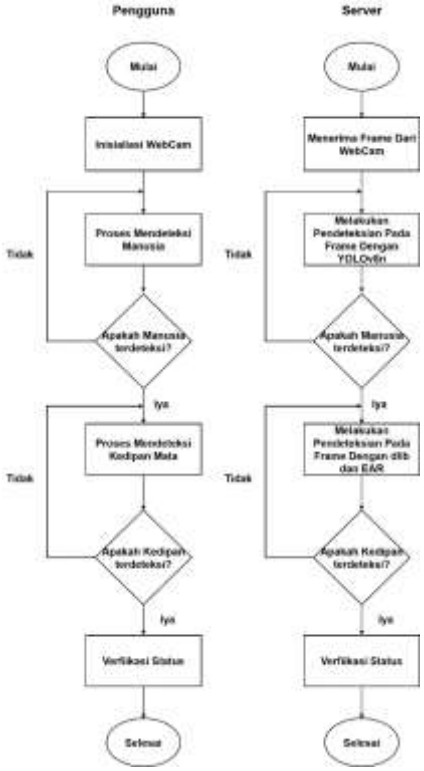
Pembangunan Sistem Anti-Spoofing berbasis web berfungsi untuk melakukan autentikasi pengguna secara real-time dengan mendeteksi keberadaan manusia serta memastikan adanya kedipan mata sebagai indikator liveness. Sistem dibangun menggunakan pendekatan client-server, di mana pengguna akan mengakses antarmuka berbasis web melalui perangkat dengan kamera yang terintegrasi. Kamera akan menangkap citra wajah secara langsung yang kemudian dikirimkan ke sistem informasi untuk diproses. Proses deteksi pertama menggunakan model YOLOv8n, yang bertujuan untuk mengidentifikasi apakah wajah yang tertangkap merupakan manusia asli atau objek palsu seperti gambar cetak, tampilan digital, atau topeng. Setelah wajah dikategorikan sebagai asli, sistem melanjutkan ke proses kedua yaitu pendeteksian kedipan mata menggunakan metode facial landmark dari Dlib dan perhitungan Eye Aspect Ratio (EAR) untuk menentukan liveness pengguna. Alur proses ini dapat dilihat pada Gambar 10 yang menggambarkan infrastruktur sistem secara keseluruhan.



Gambar 10. Infrastruktur Sistem

Gambar 10 menggambarkan bahwa proses dimulai ketika pengguna memberikan izin akses kamera melalui browser. Frame video dikirimkan secara berkala menuju server melalui koneksi HTTP. Server kemudian menjalankan tahap deteksi pertama menggunakan model YOLOv8n untuk menentukan apakah objek pada frame merupakan wajah dan mengklasifikasikannya ke dalam kategori real, printed, digital, atau mask. Hanya wajah yang terdeteksi sebagai real yang diteruskan ke tahap liveness detection. Desain dua tahap ini diterapkan untuk mengurangi beban komputasi pada proses EAR karena sistem tidak perlu menghitung landmark mata apabila wajah sebelumnya telah diklasifikasikan sebagai spoofing pada tahap YOLOv8n. Mekanisme interaksi antara pengguna dan server ditunjukkan pada Gambar 11.

Gambar 11 menunjukkan mekanisme sistem yang terbagi menjadi dua yakni, sisi pengguna (client) dan server (backend). Alur proses dimulai dari pengguna yang mengakses sistem melalui browser. Setelah halaman dimuat, sistem akan secara otomatis meminta izin untuk mengakses kamera pada perangkat pengguna. Jika izin diberikan, maka webcam akan diinisialisasi untuk mulai menangkap video secara langsung. Sistem informasi pada sisi frontend akan melakukan proses ekstraksi frame dari video, kemudian mengirimkan frame tersebut ke server menggunakan protokol HTTP. Setelah frame diterima oleh server, proses dilanjutkan dengan tahapan deteksi manusia menggunakan model YOLOv8n. Model ini akan memproses frame untuk mendeteksi wajah dan mengklasifikasikan apakah wajah tersebut tergolong real (asli) atau termasuk dalam kategori spoofing seperti printed, digital, atau mask. Jika hasil klasifikasi menunjukkan bahwa wajah merupakan wajah asli, sistem akan melanjutkan proses ke tahap deteksi liveness melalui analisis kedipan mata.



Gambar 11. Mekanisme Sistem

Tahap deteksi kedipan dilakukan menggunakan pustaka dlib untuk mengekstraksi 68 titik landmark wajah, khususnya pada area mata. Berdasarkan titik- titik tersebut, sistem menghitung nilai Eye Aspect Ratio (EAR) untuk setiap mata. Jika nilai EAR mengalami penurunan signifikan dalam waktu singkat, maka dianggap sebagai kedipan. Apabila kedipan berhasil terdeteksi, maka sistem akan mengembalikan status verifikasi sebagai berhasil, menandakan bahwa pengguna adalah manusia asli yang hidup. Sebaliknya, jika tidak ditemukan kedipan, atau wajah dikategorikan sebagai palsu, maka status verifikasi ditolak. Proses ini berlangsung secara cepat dan efisien karena hanya menganalisis wajah yang telah dinyatakan valid oleh YOLOv8n.

Gambar 12 menunjukkan halaman sistem web pendeteksian yang dapat menampilkan hasil dari proses

DOI : https://doi.org/10.52158/jacost.v6i2.1362

autentikasi secara real-time. Antarmuka ini menampilkan area video yang menayangkan citra langsung baik dari kamera internal atau eksternal pengguna. Terdapat instruksi yang mengarahkan pengguna untuk memperlihatkan wajah dan melakukan kedipan mata sebagai bagian dari proses verifikasi. Sistem akan menjalankan proses pendeteksian secara otomatis setelah pengguna menekan tombol "Mulai Verifikasi".

Gambar 12. Tampilan Sistem Web

Hasil dari dua proses utama, yaitu pendeteksian manusia dan pendeteksian kedipan mata, ditampilkan secara eksplisit melalui indikator teks di bawah video. Label "Manusia" akan menunjukkan apakah sistem berhasil mendeteksi bahwa wajah pengguna tergolong sebagai manusia asli (real), sedangkan label "Kedip" akan memperlihatkan apakah sistem berhasil mendeteksi liveness melalui aktivitas kedipan mata. Masing-masing status disajikan secara visual dalam bentuk simbol centang (✓) atau silang (✗) untuk memudahkan pemahaman pengguna. Apabila kedua indikator menyatakan deteksi berhasil, maka pengguna dianggap telah terverifikasi secara sah sebagai manusia hidup

3.5. Pengujian Waktu Pendeteksian Dengan Perangkat Laptop

Pengujian waktu pendeteksian bertujuan untuk mengetahui sejauh mana sistem dapat bekerja secara responsif ketika dijalankan pada perangkat laptop biasa tanpa dukungan perangkat keras khusus. Dua komponen utama yang diuji adalah waktu deteksi wajah menggunakan YOLOv8n dan waktu validasi liveness menggunakan EAR. Pengujian dilakukan dengan webcam 720p pada model YOLOv8n pre-trained (100 epoch) dan threshold EAR sebesar 0.17, berdasarkan hasil evaluasi optimal pada pengujian sebelumnya.

Pengambilan waktu dalam pengujian ini dibagi menjadi dua tahap terpisah yang mencerminkan alur kerja sistem secara berurutan. Tahap pertama merupakan proses deteksi wajah menggunakan model YOLOv8n, di mana penghitungan waktu dimulai sejak tombol "Mulai Deteksi" ditekan oleh pengguna, dan dihentikan saat sistem berhasil mendeteksi wajah serta menampilkan status klasifikasi real. Setelah status dari YOLOv8n muncul, sistem secara otomatis masuk ke tahap kedua, yaitu proses validasi liveness melalui deteksi kedipan mata menggunakan metode EAR. Pada tahap ini, waktu dihitung kembali mulai dari nol detik, dan diakhiri saat sistem berhasil memverifikasi kedipan mata.

Hasil pengujian dapat dilihat pada Tabel 8 yang menunjukkan bahwa waktu yang dibutuhkan oleh model YOLOv8n untuk mendeteksi wajah berkisar antara 0,45 detik hingga 5,63 detik, dengan rata-rata sekitar 1,65 detik. Sementara itu, waktu yang diperlukan oleh sistem untuk melakukan validasi kedipan mata menggunakan metode EAR memiliki rentang yang lebih bervariasi, yakni antara 0,46 detik hingga 8,65 detik, dengan rata-rata sekitar 2,89 detik. Perbedaan ini disebabkan oleh sifat alami proses deteksi kedipan, yang sangat tergantung pada interaksi pengguna. Dalam beberapa kasus, sistem membutuhkan waktu lebih lama untuk mendeteksi kedipan, terutama jika pengguna tidak segera melakukan gerakan mata atau berada dalam posisi yang kurang ideal terhadap kamera.

Waktu total yang dibutuhkan sistem untuk menyelesaikan proses deteksi dan validasi berkisar antara dua hingga sepuluh detik. Rentang waktu ini masih dapat dikategorikan sebagai responsif untuk penggunaan real-time dalam konteks autentikasi berbasis wajah. Dengan performa tersebut, sistem dapat diandalkan untuk digunakan dalam aplikasi berbasis web pada perangkat laptop tanpa memerlukan perangkat keras tambahan. Hasil ini juga menunjukkan bahwa kombinasi antara model YOLOv8n dan metode EAR cukup efektif dalam mendukung proses verifikasi identitas pengguna dengan mempertimbangkan aspek keamanan dan kenyamanan.

Tabel 8. Hasil Pengujian Waktu Pendeteksian

Pengujian ke-	YOLOv8n (detik)	EAR (detik)
1	05.63	03.83
2	01.35	03.28
3	01.59	04.70
4	01.84	03.38
5	01.71	01.83
6	00.45	01.26
7	01.83	01.09
8	01.65	01.73
9	00.64	01.56
10	01.78	01.66
11	01.21	02.66
12	01.66	00.46
13	01.65	01.16
14	01.19	01.21
15	01.16	04.29
16	01.26	06.41
17	01.68	08.65
18	01.96	02.35
19	01.71	01.18
20	01.23	05.15

4. Kesimpulan

Penelitian ini menghasilkan sistem anti-spoofing wajah real-time yang mengintegrasikan YOLOv8n untuk

DOI : https://doi.org/10.52158/jacost.v6i2.1362

deteksi spoofing empat kelas (real, printed, digital, mask) serta *Eye Aspect Ratio* (EAR) untuk verifikasi liveness melalui kedipan mata. Evaluasi menunjukkan bahwa kedua strategi pelatihan—baik pretrained maupun from scratch—memberikan performa tinggi dan stabil dengan rata-rata precision 99%, recall 98%, mAP50 99.4%, dan mAP50–95 sekitar 90%. Pengujian real-time pada berbagai kondisi pencahayaan, variasi etnis, serta pengguna berkacamata turut mengonfirmasi ketahanan sistem dengan rata-rata confidence di atas 90%.

Komponen liveness EAR menunjukkan bahwa threshold 0.17 memberikan hasil terbaik dengan akurasi 99.02%, recall 100%, serta FRR 0%, menandakan keseimbangan optimal antara sensitivitas dan ketelitian dalam mendeteksi kedipan mata.

Novelty penelitian ini terletak pada integrasi model YOLOv8n—yang masih minim eksplorasi dalam domain face anti-spoofing—dengan mekanisme liveness EAR dalam satu pipeline inferensi real-time berbasis web. Kontribusi ini memberikan solusi yang ringan, cepat, dan dapat diimplementasikan pada perangkat dengan sumber daya terbatas, termasuk laptop dan perangkat mobile.

Secara keseluruhan, hasil penelitian menunjukkan bahwa kombinasi YOLOv8n dan EAR mampu menghasilkan sistem anti-spoofing berbasis web yang ringan dengan inferensi cepat dan performa *real-time* yang stabil untuk kebutuhan autentikasi wajah modern.

Sebagai arahan pengembangan selanjutnya, penelitian ini perlu diperluas melalui pengujian terhadap serangan berbasis deepfake video, evaluasi pada berbagai model kamera dan perangkat mobile, serta validasi pada lingkungan non-terkontrol untuk memastikan ketangguhan sistem pada skenario nyata yang lebih beragam.

Daftar Rujukan

[1] R. Budiarto Hadiprakoso and I. K. S. Buana, "Deteksi Serangan Spoofing Wajah Menggunakan Convolutional Neural Network," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 3, Dec. 2021, doi: 10.28932/jutisi.v7i3.4001.

[2] M. Alshaikhli, O. Elharrouss, S. Al-Maadeed, and A. Bouridane, "Face-Fake-Net: The Deep Learning Method for Image Face Anti-Spoofing Detection : 45," in *Proceedings - European Workshop on Visual Information Processing, EUVIP*, Institute of Electrical and Electronics Engineers Inc., Jun. 2021. doi: 10.1109/EUVIP50544.2021.9484023.

[3] R. Ganjoo and A. Purohit, "Anti-Spoofing Door Lock Using Face Recognition and Blink Detection," in *Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021*, Institute of Electrical and Electronics Engineers Inc., Jan. 2021, pp. 1090–1096. doi: 10.1109/ICICT50816.2021.9358795.

[4] R. Vishnu, S. Vardhan, S. Varun, and N. R. Krishnamoorthy, "Detection of Anti-Spoofing Face Using Yolo," vol. 7, no. 2, 2025, [Online]. Available: www.ijfmr.com

[5] H. Xing, S. Y. Tan, F. Qamar, and Y. Jiao, "Face Anti-Spoofing Based on Deep Learning: A Comprehensive Survey," Jun. 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/app15126891.

[6] A. Ali-Gombe, E. Elyan, C. Francisco Moreno-García, and J. Zwiegelaar, "Face detection with YOLO on edge," vol. 3, pp. 284–292, 2021, doi: https://doi.org/10.1007/978-3-030-80568-5_24.

[7] Z. Yu *et al.*, "Searching Central Difference Convolutional Networks for Face Anti-Spoofing." [Online]. Available: https://github.com/ZitongYu/CDCN.

[8] A. Malik, M. Kuribayashi, S. M. Abdullahi, and A. N. Khan, "DeepFake Detection for Human Face Images and Videos: A Survey," 2022, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2022.3151186.

[9] A. Haliassos, K. Vougioukas, S. Petridis, and M. Pantic, "Lips Don't Lie: A Generalisable and Robust Approach to Face Forgery Detection."

[10] W. Lee, M. H. Kang, J. Song, and K. Hwang, "The design of preventive automated driving systems based on convolutional neural network," *Electronics (Switzerland)*, vol. 10, no. 14, Jul. 2021, doi: 10.3390/electronics10141737.

[11] A. Abdullah, G. A. Amran, S. M. A. Tahmid, A. Alabrah, A. A. AL-Bakhrani, and A. Ali, "A Deep-Learning-Based Model for the Detection of Diseased Tomato Leaves," *Agronomy*, vol. 14, no. 7, Jul. 2024, doi: 10.3390/agronomy14071593.

[12] M. L. Hoang, "Smart Drone Surveillance System Based on AI and on IoT Communication in Case of Intrusion and Fire Accident," *Drones*, vol. 7, no. 12, Dec. 2023, doi: 10.3390/drones7120694.

[13] "YOLOv5 vs. YOLOv8: A Detailed Comparison - Ultralytics YOLO Docs." Accessed: Aug. 26, 2025. [Online]. Available: https://docs.ultralytics.com/compare/yolov5-vs-yolov8/

[14] "Explore Ultralytics YOLOv8 - Ultralytics YOLO Docs." Accessed: Aug. 01, 2025. [Online]. Available: https://docs.ultralytics.com/models/yolov8/#supported-tasks-and-modes

[15] Husnan, C. Fatichah, and R. Dikairono, "Deteksi Objek Menggunakan Metode Yolo dan Implementasinya pada Robot Bawah Air," vol. 12(3), Dec. 2023, doi: 10.12962/j23373539.v12i3.122326.

[16] W. Hutamaputra and F. Utaminingrum, "Implementasi Facial Landmark dalam Pengenalan Wajah pada Sistem Pembayaran Elektronik," vol. 5, no. 5, pp. 2058–2064, May 2021, [Online]. Available: http://j-ptiik.ub.ac.id

[17] G. Amato, F. Falchi, C. Gennaro, and C. Vairo, "A Comparison of Face Verification with Facial Landmarks and Deep Features," Apr. 2018. [Online]. Available: https://www.researchgate.net/publication/338048224

[18] C. Dewi, R. C. Chen, C. W. Chang, S. H. Wu, X. Jiang, and H. Yu, "Eye Aspect Ratio for Real-Time Drowsiness Detection to Improve Driver Safety," *Electronics (Switzerland)*, vol. 11, no. 19, Oct. 2022, doi: 10.3390/electronics11193183.

[19] C. Dewi, R. C. Chen, X. Jiang, and H. Yu, "Adjusting eye aspect ratio for strong eye blink detection based on facial landmarks," *PeerJ Comput Sci*, vol. 8, 2022, doi: 10.7717/peerj-cs.943.

[20] "CelebA Spoof For Face AntiSpoofing." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/attentionlayer241/celeba-spoof-for-face-antispoofing

[21] "CelebA Spoof Dataset - Real People." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/celeba-spoof-dataset

[22] "iBeta 1 - 42,280 Liveness Detection Dataset." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/ibeta-level-1-liveness-detection-dataset-part-1

[23] "iBeta Level 2 Dataset, 33 000 attacks." Accessed: Aug. 26, 2025. [Online]. Available:

https://www.kaggle.com/datasets/tapakah68/pad-ibeta-level-2

[24] "Hispanic People - Liveness Detection Video Dataset." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/hispanic-people-liveness-detection-video-dataset

[25] "Black People - Liveness Detection Video Dataset." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/black-people-liveness-detection-video-dataset

[26] "Caucasian People - Liveness Detection Dataset." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/caucasian-people-liveness-detection-dataset

[27] "Anti Spoofing Real Dataset - 98,000 files." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/anti-spoofing-live

[28] "Silicone Mask Attack Dataset for Anti-Spoofing." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/axondata/silicone-mask-biometric-attack-dataset

[29] "Web Camera Face Liveness Detection - Face Dataset." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/web-camera-face-liveness-detection

[30] "On-Device Face Liveness Detection - Face Dataset." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/on-device-face-liveness-detection

[31] "Real VS Fake - Liveness Detection Dataset." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/real-vs-fake-anti-spoofing-video-classification

[32] "Silicone Masks Biometric Attacks Dataset." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/silicone-masks-biometric-attacks

[33] "Attacks with 2D Printed Masks of Indian People." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/attacks-with-2d-printed-masks-of-indian-people

[34] "Anti-Spoofing Dataset, 30,000 sets." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/tapakah68/anti-spoofing

[35] "Selfie and Video Dataset - Back Camera." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/tapakah68/selfie-and-video-on-back-camera

[36] "Full HD Videos - Liveness Detection Dataset." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/full-hd-webcam-live-attacks

[37] "2D Masks Presentation Attack Detection." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/trainingdatapro/real-people-and-attacks-with-2d-masks

[38] "Fake-Vs-Real-Faces (Hard)." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/hamzaboulahia/hardfakevsrealfaces

[39] "IDR&D_train_set." Accessed: Aug. 26, 2025. [Online]. Available: https://www.kaggle.com/datasets/nurmukhammed7/idrd-train-set

[40] E. Essel, F. Lacy, W. Elmedany, F. Albalooshi, and Y. Ismail, "Driver Drowsiness Detection Using Fixed and Dynamic Thresholding," in *2022 International Conference on Data Analytics for Business and Industry, ICDABI 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 552–557. doi: 10.1109/ICDABI56818.2022.10041670.

DOI : https://doi.org/10.52158/jacost.v6i2.1362