

IMPLEMENTASI YOLOV11 UNTUK DETEKSI GERAKAN SANDI SEMAFOR DENGAN PENDEKATAN POSE ESTIMATION

Elvan Nasrul, Asriyanik, Agung Pambudi

Teknik Informatika, Universitas Muhammadiyah Sukabumi

Jl. R. Syamsudin, SH. No. 50 Kota Sukabumi Jawa Barat Indonesia 43113, Indonesia

elvan23nasrul@ummi.ac.id

ABSTRAK

Sandi semafor merupakan keterampilan komunikasi nonverbal dalam kegiatan kepramukaan yang menyandikan huruf alfabet melalui konfigurasi posisi kedua lengan menggunakan bendera. Evaluasi gerakan semafor yang masih dilakukan secara manual oleh instruktur menyebabkan penilaian cenderung subjektif dan tidak konsisten, terutama pada kondisi pencahayaan dan latar belakang yang bervariasi. Penelitian ini bertujuan mengembangkan sistem deteksi otomatis sandi semafor yang mampu memberikan penilaian lebih objektif serta mempertahankan konsistensi prediksi pada kondisi pencahayaan dan latar belakang yang bervariasi. Sistem dikembangkan menggunakan algoritma YOLOv11s-Pose dengan pendekatan *pose estimation*, mengacu pada kerangka kerja Microsoft Team Data Science Process (TDSP) yang mencakup tahapan *Business Understanding*, *Data Understanding*, *Modeling*, dan *Deployment*. Dataset terdiri dari 1.300 citra gabungan data publik dan data primer dengan variasi latar belakang dan intensitas cahaya, dianotasi menggunakan *keypoints* tubuh bagian atas, dan dilatih selama 150 epoch menggunakan optimizer AdamW. Hasil evaluasi menunjukkan nilai mAP@0.5 sebesar 0,984 dengan *precision* dan *recall* mendekati sempurna. Evaluasi melalui *confusion matrix* juga menunjukkan model mampu membedakan huruf dengan konfigurasi lengan yang mirip. Pengujian pada tiga kondisi visual berbeda membuktikan model tetap stabil dengan skor kepercayaan rata-rata 0,82, termasuk pada kondisi minim cahaya.

Kata kunci : *computer vision, pose estimation, sandi semafor, yolov11*

1. PENDAHULUAN

Sandi semafor merupakan sistem komunikasi nonverbal yang menyandikan huruf alfabet A hingga Z melalui konfigurasi posisi kedua lengan menggunakan bendera. Dalam kegiatan kepramukaan, semafor menjadi salah satu materi penting karena melatih koordinasi gerak, ketelitian, serta kemampuan komunikasi visual antarpeserta. Kegiatan evaluasi semafor umumnya dilaksanakan di lingkungan terbuka dengan kondisi pencahayaan alami dan latar belakang yang beragam, sehingga memberikan tantangan tersendiri bagi proses penilaian..

Permasalahan utama yang dihadapi dalam evaluasi gerakan semafor adalah proses penilaian yang masih dilakukan secara manual oleh pembina atau instruktur. Kondisi ini menyebabkan hasil penilaian cenderung bersifat subjektif dan tidak selalu konsisten antarpengamat. Upaya otomatisasi pengenalan semafor sebelumnya telah dilakukan dengan berbagai pendekatan. PoseNet mampu memanfaatkan *keypoints* tubuh untuk mengenali gerakan semafor, namun kinerjanya dipengaruhi oleh variasi pencahayaan dan posisi tubuh pengguna [1]. MediaPipe yang dikombinasikan dengan Multi-Layer Perceptron (MLP) mencapai akurasi 93,84 persen, tetapi masih mengalami kesalahan pada huruf dengan konfigurasi lengan yang mirip [2]. Pendekatan Decision Tree juga menunjukkan penurunan performa ketika kondisi pencahayaan melemah [3]. Temuan-temuan tersebut mengindikasikan bahwa tantangan utama bukan hanya

akurasi tinggi pada data terkontrol, melainkan juga konsistensi prediksi pada kondisi lingkungan yang dinamis.

Berdasarkan permasalahan dan gap penelitian tersebut, penelitian ini bertujuan mengembangkan sistem deteksi otomatis sandi semafor yang mampu memberikan penilaian lebih objektif serta mempertahankan konsistensi prediksi pada kondisi pencahayaan dan latar belakang yang bervariasi.

Untuk mencapai tujuan tersebut, penelitian ini mengusulkan pendekatan berbasis YOLOv11s-Pose dengan *pose estimation* sebagai solusi utama. YOLOv11 sebagai versi terbaru dari keluarga YOLO menekankan efisiensi arsitektur dan kualitas ekstraksi fitur yang kompetitif pada berbagai skenario dataset [4], [5]. Pendekatan YOLO-Pose dipilih karena mampu mengintegrasikan deteksi manusia dan prediksi *keypoints* dalam satu tahap inferensi, sehingga representasi pose lengan dapat diperoleh secara langsung [6]. Proses pengembangan mengikuti kerangka Microsoft Team Data Science Process (TDSP) agar tahapan penelitian terstruktur dan dapat direproduksi [7], mencakup *fine-tuning* menggunakan dataset kustom yang memuat variasi pencahayaan, latar belakang, jarak, dan sudut pengambilan. Evaluasi dilakukan menggunakan metrik mAP, *precision*, *recall*, serta *confusion matrix* untuk menganalisis performa dan pola kesalahan klasifikasi antar huruf [8], [9].

2. TINJAUAN PUSTAKA

2.1. Sandi Semafor

Sandi semafor merupakan sistem komunikasi visual yang menyandikan huruf alfabet melalui kombinasi posisi kedua lengan yang memegang bendera. Dalam kegiatan kepramukaan, semafor digunakan sebagai keterampilan dasar untuk melatih konsentrasi, koordinasi gerak, serta kemampuan komunikasi nonverbal. Setiap huruf direpresentasikan oleh orientasi lengan kiri dan kanan yang membentuk pola sudut tertentu, sehingga informasi utama dari semafor terletak pada konfigurasi geometris lengan [10]

Karena perbedaan antar huruf ditentukan oleh posisi persendian bahu, siku, dan pergelangan tangan, maka pengenalan semafor dapat dipandang sebagai permasalahan klasifikasi pose. Hal ini menjadikan pendekatan berbasis *pose estimation* lebih relevan dibandingkan pendekatan berbasis warna bendera atau latar belakang, karena fokus utama berada pada struktur gerakan tubuh [2].

2.2. Deep Learning dan Convolutional Neural Network (CNN)

Deep learning merupakan cabang lanjutan dari machine learning yang menggunakan jaringan saraf berlapis untuk membentuk representasi kompleks dari data dalam jumlah besar. Keunggulan utama deep learning terletak pada kemampuannya mengekstraksi fitur secara otomatis tanpa memerlukan rekayasa fitur manual, sehingga sangat sesuai untuk pengolahan citra dan pola visual [11].

Salah satu arsitektur *deep learning* yang paling banyak digunakan dalam *computer vision* adalah *Convolutional Neural Network* (CNN). CNN dirancang untuk mempelajari pola spasial dalam citra melalui operasi konvolusi bertingkat, sehingga mampu mengekstraksi fitur mulai dari bentuk sederhana hingga struktur kompleks [12]. Dalam konteks deteksi objek dan pengenalan gerakan, CNN menjadi fondasi utama bagi berbagai model modern termasuk YOLO dan pose [13].

2.3. Algoritma YOLO dan Perkembangannya

You Only Look Once (YOLO) merupakan algoritma deteksi objek berbasis *deep learning* yang memprediksi lokasi objek dan kelas dalam satu pipeline inferensi. Pendekatan ini menjadikan YOLO efisien karena deteksi dilakukan secara *end-to-end* tanpa memerlukan tahapan terpisah seperti proposal region [14].

Perkembangan YOLO mengalami inovasi arsitektural dari generasi awal hingga versi modern. Review komprehensif menunjukkan bahwa peningkatan versi YOLO berfokus pada efisiensi backbone, fusi fitur multi-skala, serta optimasi head untuk meningkatkan ketelitian deteksi pada berbagai kondisi visual [15], [16].

YOLOv11 merupakan salah satu versi terbaru yang menekankan efisiensi parameter serta

peningkatan kualitas ekstraksi fitur melalui modifikasi blok arsitektur seperti C3k2 dan modul perhatian spasial [4]. Studi benchmark evolusi YOLO juga menempatkan YOLOv11 sebagai model yang kompetitif untuk aplikasi deteksi modern dengan keseimbangan antara performa dan kompleksitas komputasi [17].

2.4. Pose Estimation dan YOLO-Pose

Pose estimation merupakan proses untuk menentukan posisi dan orientasi bagian tubuh manusia melalui identifikasi titik-titik kunci atau *keypoints*. Representasi pose memberikan struktur kerangka tubuh yang relatif lebih stabil terhadap perubahan latar belakang dan pencahayaan dibandingkan fitur warna [18].

YOLO-Pose merupakan varian dari YOLO yang mengintegrasikan deteksi manusia dan prediksi *keypoints* dalam satu tahap inferensi. Model ini tidak hanya menghasilkan *bounding box*, tetapi juga koordinat persendian tubuh yang dapat digunakan untuk analisis gerakan [19].

Dalam konteks sandi semafor, pose estimation menjadi pendekatan yang relevan karena klasifikasi huruf sepenuhnya ditentukan oleh orientasi kedua lengan. Penelitian sebelumnya menunjukkan bahwa penggunaan YOLO-Pose mampu memberikan performa tinggi dalam mendeteksi gerakan semafor, meskipun huruf dengan konfigurasi mirip masih menjadi tantangan [6].

2.5. Keypoints dalam Deteksi Pose

Keypoints adalah titik-titik penting dalam citra yang merepresentasikan fitur anatomis tubuh seperti bahu, siku, dan pergelangan tangan. Dalam pose estimation, *keypoints* digunakan untuk membangun struktur kerangka tubuh serta memahami hubungan spasial antar sendi [20].

Stabilitas prediksi *keypoints* sangat penting karena kesalahan kecil pada koordinat sendi dapat memengaruhi interpretasi gerakan. Oleh karena itu, pendekatan pose estimation yang baik harus mampu menghasilkan koordinat *keypoints* yang konsisten pada berbagai kondisi pengambilan gambar [21].

Dalam penelitian ini, *keypoints* difokuskan pada tubuh bagian atas, khususnya delapan titik utama pada kepala, leher, bahu, siku, dan pergelangan tangan, karena konfigurasi tersebut merupakan penentu utama huruf semafor.

2.6. Penelitian Terdahulu

Beberapa penelitian terdahulu telah mengeksplorasi pengenalan gerakan sandi semafor menggunakan berbagai pendekatan *computer vision* dan *machine learning*. Implementasi PoseNet dalam pengenalan gerakan semafor berbasis game menunjukkan bahwa pendekatan tersebut mampu mengidentifikasi *keypoints* tubuh untuk klasifikasi huruf semafor, namun performanya masih dipengaruhi

oleh variasi pencahayaan dan perubahan posisi tubuh pengguna [1].

Pendekatan kombinasi MediaPipe dan Multi-Layer Perceptron (MLP) berhasil mencapai akurasi sebesar 93,84 persen dalam mendeteksi pose semafor, namun masih mengalami kesalahan klasifikasi pada huruf-huruf yang memiliki konfigurasi lengan serupa [2]. Penelitian lain menggunakan algoritma Decision Tree untuk identifikasi gerakan tangan semafor secara *real-time* mengungkapkan keterbatasan serupa, yaitu penurunan performa yang signifikan ketika kondisi pencahayaan melemah [3].

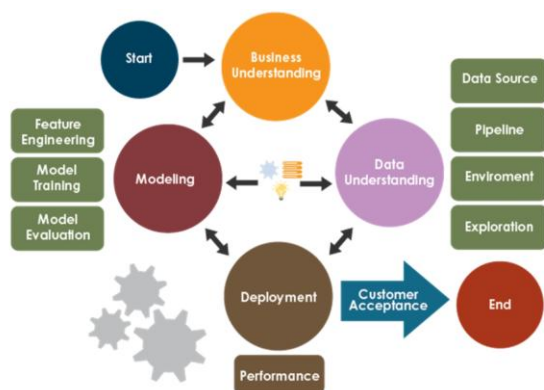
Implementasi YOLO-Pose sebagai media trainer kode semafor pramuka menunjukkan bahwa pendekatan ini mampu memberikan performa tinggi dalam mendeteksi gerakan semafor secara *end-to-end*, meskipun huruf dengan konfigurasi pose yang mirip masih menjadi tantangan utama dalam proses klasifikasi [6].

Berdasarkan kajian terhadap penelitian-penelitian terdahulu tersebut, terdapat gap yang belum terselesaikan, yaitu keterbatasan konsistensi prediksi pada kondisi pencahayaan yang bervariasi. Penelitian ini hadir untuk menjawab gap tersebut dengan memanfaatkan YOLOv11s-Pose yang memiliki arsitektur lebih mutakhir serta didukung dataset kustom dengan variasi kondisi lingkungan yang lebih beragam.

3. METODE PENELITIAN

Penelitian ini menggunakan kerangka kerja Microsoft Team Data Science Process (TDSP) sebagai metodologi utama dalam pengembangan sistem deteksi sandi semafor berbasis *pose estimation*. TDSP dipilih karena menyediakan alur kerja yang sistematis dan iteratif untuk membangun solusi *machine learning*, mulai dari perumusan masalah hingga implementasi model dalam bentuk aplikasi.

Tahapan penelitian dalam TDSP meliputi *Business Understanding*, *Data Understanding*, *Modeling*, dan *Deployment* [7]. Alur penelitian secara umum ditunjukkan pada Gambar 1.



Gambar 1. Metode TDSP

3.1. Business Understanding

Tahap *Business Understanding* dilakukan untuk mengidentifikasi permasalahan utama dalam evaluasi gerakan sandi semafor yang masih dilakukan secara manual dan berpotensi kurang konsisten pada kondisi lingkungan yang bervariasi. Berdasarkan permasalahan tersebut, penelitian ini diarahkan untuk mengembangkan sistem otomatis yang mampu mengenali gerakan sandi semafor huruf A sampai Z melalui pendekatan deteksi pose.

3.2. Data Understanding

Tahap *Data Understanding* bertujuan untuk memahami karakteristik data yang digunakan dalam penelitian. Dataset yang digunakan berupa citra pose sandi semafor yang mencakup seluruh huruf alfabet. Data diperoleh dari gabungan sumber publik dan pengambilan mandiri untuk meningkatkan keberagaman latar belakang, pencahayaan, serta variasi gaya pose. Seluruh data dianotasi dengan informasi *keypoints* lengan sebagai representasi utama gerakan semafor.

3.3. Modeling

Tahap *Modeling* merupakan proses pengembangan model deteksi pose untuk mengenali pola gerakan sandi semafor berdasarkan konfigurasi lengan. Model dilatih menggunakan dataset yang telah dipersiapkan pada tahap sebelumnya, sehingga mampu menghasilkan prediksi huruf semafor secara otomatis melalui estimasi *keypoints* tubuh bagian atas.

3.4. Deployment

Tahap *Deployment* dilakukan dengan mengimplementasikan model yang telah dikembangkan ke dalam bentuk sistem berbasis web. Sistem ini dirancang agar pengguna dapat memperoleh hasil deteksi berupa visualisasi pose dan label huruf yang dikenali. Tahap ini bertujuan memastikan model dapat digunakan sebagai media bantu evaluasi sandi semafor secara lebih objektif.

4. HASIL DAN PEMBAHASAN

Bagian ini membahas hasil implementasi dan evaluasi sistem deteksi sandi semafor berbasis YOLOv11s-Pose yang dikembangkan dalam penelitian ini. Pembahasan disusun mengikuti tahapan Microsoft Team Data Science Process (TDSP), yaitu *Business Understanding*, *Data Understanding*, *Modeling*, dan *Deployment*. Pada bagian ini ditampilkan proses pengolahan dataset, hasil pelatihan model, evaluasi performa, serta pengujian sistem pada beberapa kondisi pencahayaan dan jarak pengambilan citra.

4.1. Business Understanding

Tahap *Business Understanding* pada penelitian ini menegaskan kembali kebutuhan sistem otomatis dalam pengenalan sandi semafor. Evaluasi gerakan semafor dalam kegiatan kepramukaan masih

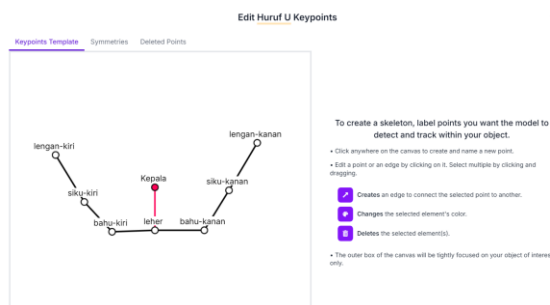
dilakukan secara manual melalui pengamatan instruktur, sehingga hasil penilaian dapat bersifat subjektif dan kurang konsisten. Selain itu, latihan semafor umumnya dilakukan pada lingkungan terbuka dengan variasi pencahayaan alami dan latar belakang yang kompleks, sehingga meningkatkan risiko kesalahan interpretasi pose.

Berdasarkan permasalahan tersebut, penelitian ini mengembangkan sistem deteksi sandi semafor huruf A sampai Z berbasis YOLOv11s-Pose. Sistem ini memanfaatkan pendekatan *pose estimation* untuk mengenali konfigurasi *keypoints* lengan sebagai representasi utama setiap huruf semafor. Target utama dari tahap ini adalah menghasilkan model yang mampu memberikan prediksi huruf secara lebih objektif serta tetap stabil pada kondisi lingkungan yang bervariasi.

4.2. Data Understanding

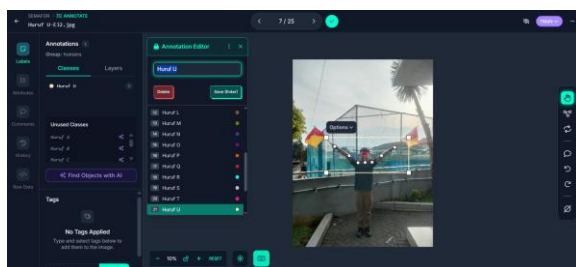
Tahap *Data Understanding* dilakukan untuk menganalisis karakteristik dataset yang digunakan dalam pelatihan model. Dataset penelitian ini merupakan gabungan antara dataset publik dan data primer yang dikumpulkan secara mandiri. Total dataset berjumlah 1.300 citra pose sandi semafor yang mencakup 26 huruf alfabet A sampai Z.

Seluruh citra dianotasi menggunakan Roboflow dengan format *bounding box* dan *keypoints* tubuh bagian atas. Fokus anotasi diarahkan pada konfigurasi lengan karena sandi semafor sepenuhnya ditentukan oleh posisi kedua lengan.



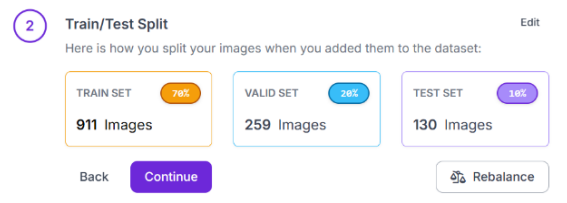
Gambar 2. Format *keypoints* tubuh bagian atas pada sandi semafor

Proses anotasi dilakukan secara manual untuk memastikan setiap pose memiliki label huruf yang sesuai.



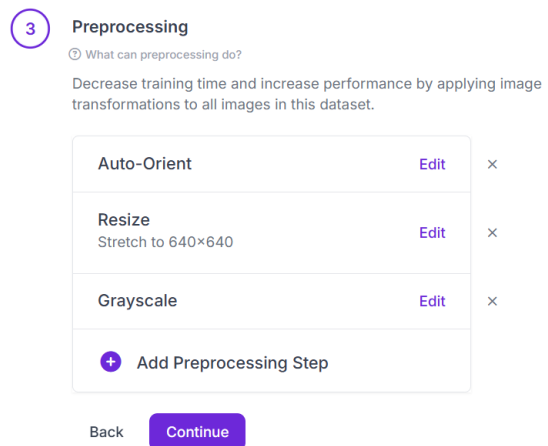
Gambar 3. Proses anotasi dataset menggunakan Roboflow

Dataset kemudian dibagi menjadi tiga subset, yaitu data pelatihan 70 persen, data validasi 20 persen, dan data pengujian 10 persen.



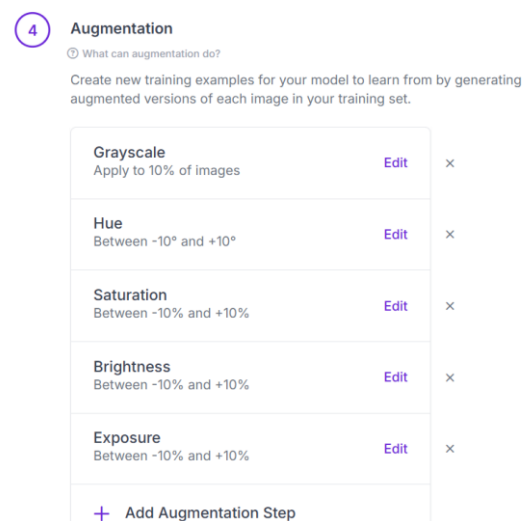
Gambar 4. Pembagian dataset

Sebelum pelatihan, dilakukan *preprocessing* berupa penyesuaian ukuran citra menjadi 640×640 piksel agar sesuai dengan arsitektur YOLOv11-Pose.



Gambar 5. Proses *preprocessing* dataset

Selain itu, augmentasi data diterapkan untuk meningkatkan variasi pencahayaan dan warna, sehingga model lebih tahan terhadap kondisi lingkungan yang berubah.



Gambar 6. Contoh augmentasi dataset sandi semafor

4.3. Modeling

Tahap *Modeling* dilakukan untuk melatih dan mengevaluasi model deteksi pose sandi semafor menggunakan arsitektur YOLOv11s-Pose. Model ini dipilih karena mampu mengintegrasikan deteksi objek manusia dan estimasi keypoints dalam satu proses *end-to-end*, sehingga representasi pose lengan dapat diperoleh secara langsung tanpa memerlukan pemrosesan tambahan. Fokus utama pada tahap ini adalah memastikan bahwa model dapat mengenali konfigurasi anatomis lengan dengan presisi tinggi sebagai dasar klasifikasi huruf sandi semafor secara objektif.

Pelatihan model dilakukan dengan memanfaatkan bobot pralatih *yolo11s-pose.pt* pada dataset sandi semafor yang telah dipersiapkan pada tahap sebelumnya. Proses training dijalankan selama 150 *epoch* dengan ukuran citra 640×640 piksel dan batch size 16. Optimizer AdamW digunakan dengan *learning rate* awal sebesar 0,001 untuk menjaga stabilitas konvergensi model selama proses *fine-tuning*. Seluruh eksperimen dilakukan menggunakan GPU pada lingkungan Google Colab agar proses pelatihan dapat berjalan secara efisien.

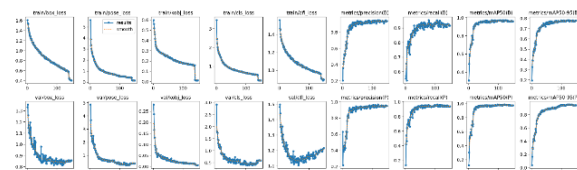
Kode utama yang digunakan dalam proses training ditampilkan pada Gambar 4.10 sebagai implementasi konfigurasi parameter pelatihan model.

```
# Load model YOLO11s-pose
model = YOLO('yolo11s-pose.pt')
# Training
results = model.train( data="/content/Semafor-5/data.yaml",
epochs=150,
imgsz=640,
batch=16,
optimizer='AdamW',
lr=0.001, device=0,
project='semaphore_training',
name='yolo11s_pose',
plots=True )
print("Training selesai!")
```

Gambar 7. Kode *training* model YOLOv11s-Pose

Setelah pelatihan selesai, sistem menghasilkan catatan performa yang mencakup perkembangan *loss* dan metrik evaluasi pada setiap *epoch*. Pemantauan ini penting untuk memastikan bahwa model mencapai konvergensi yang stabil serta tidak mengalami overfitting secara signifikan.

Hasil metrik pelatihan selama 150 *epoch* ditunjukkan pada Gambar 8. Grafik tersebut memperlihatkan penurunan nilai *loss* baik pada data pelatihan maupun validasi, yang menandakan bahwa model semakin akurat dalam mempelajari pola pose sandi semafor.

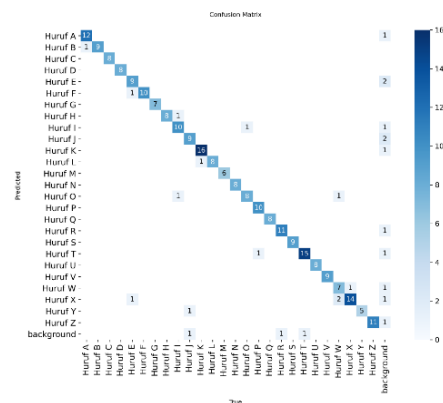


Gambar 8. Grafik metrik pelatihan model YOLOv11s-Pose

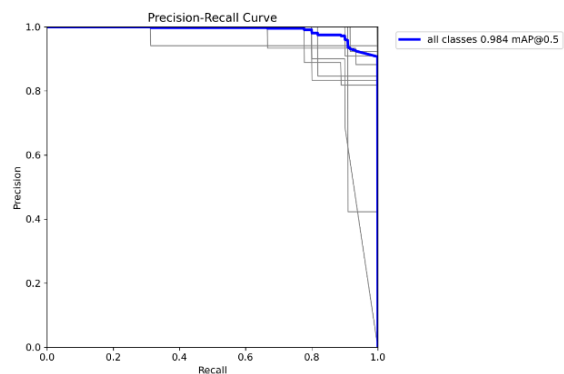
Secara lebih rinci, nilai *box loss* yang menurun menunjukkan peningkatan kemampuan model dalam melokalisasi tubuh subjek secara tepat. Selain itu, penurunan konsisten pada *pose loss* menjadi indikator utama bahwa model berhasil memprediksi koordinat keypoints lengan secara semakin presisi. Nilai *classification loss* yang semakin kecil juga mengonfirmasi bahwa model mampu membedakan kelas huruf A sampai Z dengan tingkat kesalahan yang rendah.

Metrik evaluasi seperti *precision* dan *recall* mencapai nilai tinggi mendekati 1, yang menunjukkan keseimbangan performa model dalam mendeteksi pose secara tepat tanpa banyak kehilangan data yang relevan. Nilai mAP50 dan mAP50–95 yang stabil pada level tinggi semakin memperkuat bahwa model memiliki ketelitian yang baik dalam mengenali pose sandi semafor pada berbagai ambang evaluasi.

Distribusi performa klasifikasi antar huruf divisualisasikan melalui *confusion matrix* pada Gambar 9. Matriks ini menunjukkan bahwa mayoritas prediksi berada pada diagonal utama, yang berarti sebagian besar huruf berhasil dikenali dengan benar. Beberapa kesalahan kecil masih muncul pada huruf dengan konfigurasi lengan yang sangat mirip, seperti huruf W yang sesekali tertukar dengan huruf X, serta huruf I yang memiliki peluang kecil terdeteksi sebagai huruf O akibat perbedaan sudut *keypoints* yang tipis.



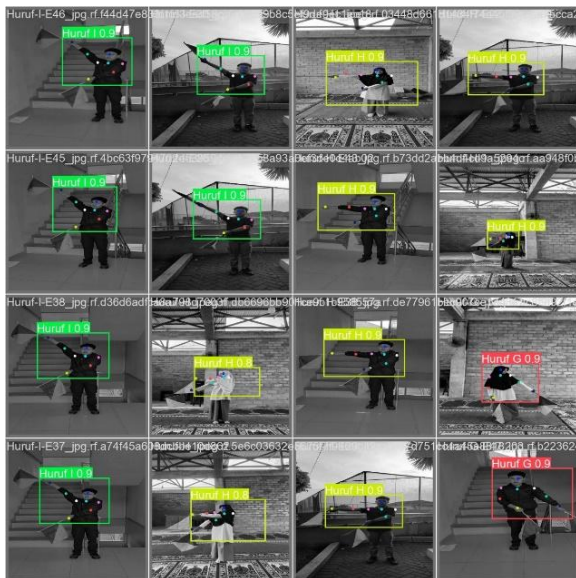
Gambar 9. *Confusion matrix* klasifikasi huruf semafor



Gambar 10. Kurva *Precision–Recall* model YOLOv11s-Pose

Konsistensi performa model juga terlihat pada kurva *Precision-Recall* yang disajikan pada Gambar 10 Hampir seluruh kelas alfabet memiliki area di bawah kurva yang mendekati nilai maksimum, dengan capaian rata-rata sebesar 0.984 pada $mAP@0.5$. Hal ini mengindikasikan bahwa model mampu mempertahankan presisi tinggi pada seluruh kelas tanpa adanya ketimpangan performa yang signifikan.

Secara visual, hasil inferensi pada satu batch data validasi ditunjukkan pada Gambar 11 Model mampu memetakan delapan keypoints utama tubuh bagian atas secara stabil, mencakup bahu, siku, hingga pergelangan tangan. Prediksi huruf yang dihasilkan juga disertai *confidence score* tinggi pada rentang 0.8 sampai 0.9, yang menunjukkan tingkat keyakinan model yang kuat. Selain itu, model tetap mampu mengenali pose dengan baik meskipun latar belakang dan pencahayaan bervariasi, sehingga menandakan *robustness* terhadap kondisi lingkungan yang dinamis.

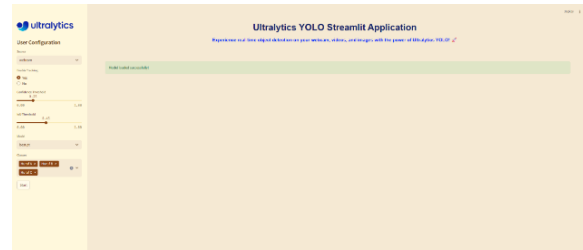


Gambar 11. Hasil inferensi model pada data validasi

Secara keseluruhan, evaluasi melalui confusion matrix, kurva precision-recall, serta hasil inferensi visual menunjukkan bahwa YOLOv11s-Pose mampu mengekstraksi fitur pose secara efektif dan memberikan klasifikasi huruf semafor yang akurat.

4.4. Deployment

Tahap terakhir dalam siklus penelitian ini adalah *deployment*, yaitu penerapan model terbaik hasil pelatihan (*best.pt*) ke dalam antarmuka sistem untuk memvalidasi performa deteksi sandi semafor pada citra nyata. Implementasi dilakukan menggunakan aplikasi berbasis Streamlit melalui antarmuka standar dari Ultralytics YOLO, sehingga model dapat dijalankan secara interaktif pada peramban web sebagai media evaluasi akhir sebelum digunakan lebih lanjut.



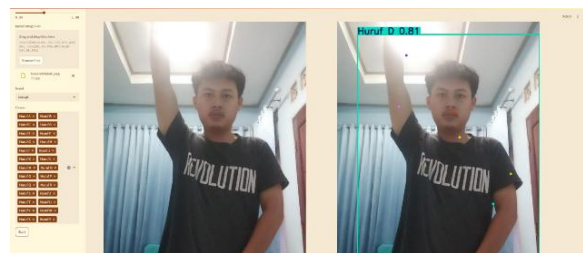
Gambar 12. Tampilan Utama Web

Antarmuka sistem yang digunakan menyediakan panel kendali pada sisi kiri untuk mengatur parameter utama deteksi, seperti ambang batas kepercayaan utama deteksi, seperti ambang batas kepercayaan (*confidence threshold*) serta ambang batas *Intersection over Union* (IoU). Pengaturan ini memungkinkan sistem menyesuaikan sensitivitas deteksi terhadap objek dan keypoints. Setelah sistem menerima input, model secara otomatis menghasilkan visualisasi pose berupa kerangka tubuh (*skeleton*) beserta label huruf sandi semafor yang dikenali. Untuk menilai keandalan model *best.pt* pada kondisi yang mendekati lingkungan latihan pramuka, pengujian dilakukan melalui tiga skema kondisi visual yang berbeda.



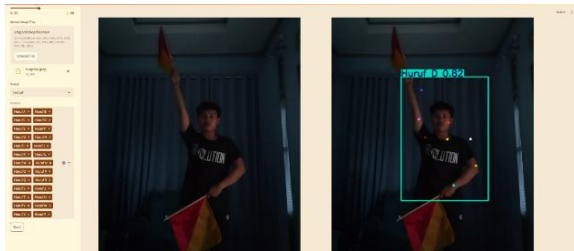
Gambar 13. Skema Pengujian Pertama

Skema pertama dilakukan pada kondisi pencahayaan cerah dengan jarak pengambilan yang ideal. Pada kondisi ini, model mampu melokalisasi *keypoints* tubuh bagian atas secara presisi, termasuk bahu, siku, dan pergelangan tangan yang menjadi penentu utama konfigurasi sandi semafor. Prediksi yang dihasilkan menunjukkan skor kepercayaan sebesar 0,82 untuk kelas “Huruf D”, dengan *bounding box* yang membungkus tubuh subjek secara proporsional. Keberhasilan deteksi pada kondisi optimal ini menegaskan bahwa model memiliki stabilitas tinggi dalam mengenali pola pose sesuai hasil evaluasi metrik pelatihan sebelumnya.



Gambar 14. Skema Pengujian Kedua

Skema kedua dilakukan pada kondisi jarak dekat (*close-up*), di mana sebagian area bendera berada di luar batas bingkai citra. Situasi ini menyebabkan informasi visual pose menjadi kurang lengkap, khususnya pada bagian atas lengan. Meskipun demikian, model tetap mampu mempertahankan prediksi yang benar dengan skor kepercayaan sebesar 0,81. Penurunan skor terjadi karena berkurangnya konteks visual pada citra, namun struktur anatomis lengan masih terdeteksi dengan baik sehingga klasifikasi huruf tetap dapat dilakukan secara tepat. Hal ini menunjukkan bahwa model tidak sepenuhnya bergantung pada keberadaan bendera, tetapi lebih fokus pada koordinat persendian tubuh.



Gambar 15. Skema Pengujian Ketiga

Skema ketiga dilakukan pada kondisi pencahayaan rendah untuk menguji batas kemampuan generalisasi model pada kualitas citra yang menurun. Walaupun detail visual pada citra tampak gelap dan fitur warna bendera tidak terlihat jelas, model tetap mampu membentuk struktur skeleton secara konsisten serta menghasilkan prediksi “Huruf D” dengan skor kepercayaan sebesar 0,82. Ketangguhan ini menunjukkan bahwa pendekatan *pose estimation* lebih menekankan pola geometri tubuh dibandingkan informasi warna, serta didukung oleh augmentasi data selama pelatihan seperti variasi kecerahan yang membantu model mengenali pose pada kondisi visual yang beragam.

4.5. Ringkasan Hasil Pengujian Deployment

Hasil dari ketiga skema pengujian tersebut dirangkum pada Tabel 1 berikut.

Tabel 1. Ringkasan Hasil Pengujian Deployment

No	Skema Pengujian	Skor Kepercayaan	Status Deteksi
1	Cahaya Cerah (Jarak Ideal)	0,82	Berhasil
2	Jarak Dekat (<i>Close-up</i>)	0,81	Berhasil
3	Kondisi Gelap (<i>Low Light</i>)	0,82	Berhasil

Berdasarkan hasil tersebut, model YOLOv11s-Pose yang telah di-*fine-tuning* mampu mempertahankan prediksi huruf sandi semafor secara stabil pada berbagai variasi kondisi pengambilan citra.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan sistem deteksi pose sandi semafor berbasis YOLOv11s-Pose untuk mengenali 26 huruf alfabet (A–Z) menggunakan dataset gabungan data publik dan data primer berjumlah 1.300 citra. Proses pelatihan selama 150 *epoch* dengan optimizer AdamW menghasilkan performa yang sangat baik dengan nilai mAP@0,5 sebesar 0,984 serta precision dan recall yang mendekati sempurna. Model juga mampu membedakan huruf-huruf semafor dengan konfigurasi lengan yang mirip, serta menunjukkan hasil prediksi yang stabil ketika diimplementasikan pada antarmuka web dan diuji pada variasi pencahayaan maupun jarak. Penelitian selanjutnya disarankan untuk memperluas dataset dengan menambahkan variasi subjek, atribut bendera, serta latar belakang yang lebih kompleks agar kemampuan generalisasi model semakin optimal pada kondisi lapangan yang beragam. Selain itu, pengembangan sistem dapat diarahkan pada penambahan fitur evaluasi gerakan secara *real-time*, sehingga sistem tidak hanya mengenali huruf semafor tetapi juga mampu memberikan umpan balik terhadap ketepatan pose pengguna sesuai standar gerakan yang benar.

DAFTAR PUSTAKA

- [1] A. M. Virginia and A. Cherid, “IMPLEMENTASI POSENET DALAM GAME SEMAPHORE UNTUK MENGENALI GERAKAN TUBUH PENGGUNA IMPLEMENTATION OF POSENET IN SEMAPHORE GAME FOR USER BODY MOVEMENT RECOGNITION,” *Bit (Fakultas Teknologi Informasi Universitas Budi Luhur)*, vol. 21, no. 1, Apr. 2024.
- [2] Milton Dapamudang, Fajar Hariadi, and Raynesta Mikaela Indri Malo, “Deteksi Pose Semaphore Berbasis Deep Learning Menggunakan Metode Multi-Layer Perceptron dan MediaPipe,” *Jurnal ilmiah Sistem Informasi dan Ilmu Komputer*, vol. 5, no. 2, pp. 405–423, Jul. 2025, doi: 10.55606/juisik.v5i2.1386.
- [3] A. S. P. Dwika, A. Abdullah, and S. P. A. Alkadri, “IDENTIFIKASI GERAKAN TANGAN PADA SANDI SEMAPHORE PRAMUKA SECARA REALTIME MENGGUNAKAN DECISION TREE,” *JUTECH: Journal Education and Technology*, vol. 5, Dec. 2024.
- [4] R. Khanam and M. Hussain, “YOLOv11: An Overview of the Key Architectural Enhancements,” Oct. 2024, [Online]. Available: <http://arxiv.org/abs/2410.17725>
- [5] V. Hendriko and D. Hermanto, “Performance Comparison of YOLOv10, YOLOv11, and YOLOv12 Models on Human Detection Datasets,” *Brilliance: Research of Artificial Intelligence*, vol. 5, no. 1, pp. 440–450, May 2025, doi: 10.47709/brilliance.v5i1.6447.

- [6] M. H. Khairy, Y. Hendriyani, A. Hadi, and H. Kurnia Saputra, "Implementasi Computer Vision pada Media Trainer Kode Semaphore Pramuka Berbasis YOLO-Pose," *Jurnal Pendidikan Tambusai*, vol. 9, no. 2, 2025.
- [7] M. Itqon and J. D. T. Purnomo, "Employee Attrition Prediction using Machine Learning in Rolling Stock Manufacturing Company," *Jurnal Teknobisnis*, vol. 8, no. 1, pp. 74–85, Jul. 2024, doi: 10.12962/j24609463.v8i1.941.
- [8] J. Erbani, P. É. Portier, E. Egyed-Zsigmond, and D. Nurbakova, "Confusion Matrices: A Unified Theory," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3507199.
- [9] I. Ahmad, F. Shang, and M. Faheem, "Enhanced YOLOv8 model for accurate and real-time remote sensing target detection," *Comput. Stand. Interfaces*, vol. 96, p. 104093, 2026, doi: <https://doi.org/10.1016/j.csi.2025.104093>.
- [10] L. S. A. Putra, F. T. P. Wigyarinto, E. Kusumawardhani, and V. A. Gunawan, "Semaphore letter code recognition system using wavelet method and back propagation neural network," *International Journal of Advanced Technology and Engineering Exploration*, vol. 11, no. 112, pp. 316–331, Mar. 2024, doi: 10.19101/IJATEE.2023.10102433.
- [11] F. M. Shiri, T. Perumal, N. Mustapha, and R. Mohamed, "A Comprehensive Overview and Comparative Analysis on Deep Learning Models," *Journal on Artificial Intelligence*, no. 1, 2024, doi: 10.32604/jai.2024.054314.
- [12] D. Bhatt *et al.*, "CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope," *Electronics (Basel)*, no. 20, Oct. 2021, doi: 10.3390/electronics10202470.
- [13] P. Gavallas, G. Stefanou, D. Savvas, C. Mattrand, and J.-M. Bourinet, "CNN-based prediction of microstructure-derived random property fields of composite materials," *Comput. Methods Appl. Mech. Eng.*, vol. 430, 2024, doi: 10.1016/j.cma.2024.117207.
- [14] M. G. Ragab *et al.*, "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," *IEEE Access*, vol. 12, pp. 57815–57836, 2024, doi: 10.1109/ACCESS.2024.3386826.
- [15] A. A. Murat and M. S. Kiran, "A comprehensive review on YOLO versions for object detection," Oct. 01, 2025, *Elsevier B.V.* doi: 10.1016/j.jestch.2025.102161.
- [16] R. Sapkota *et al.*, "YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series," *Artif. Intell. Rev.*, vol. 58, no. 9, Sep. 2025, doi: 10.1007/s10462-025-11253-3.
- [17] N. Jegham, C. Y. Koh, M. Abdelatti, and A. Hendawi, "YOLO Evolution: A Comprehensive Benchmark and Architectural Review of YOLOv12, YOLO11, and Their Previous Versions," Mar. 2025, [Online]. Available: <http://arxiv.org/abs/2411.00201>
- [18] Z. Gao, J. Chen, Y. Liu, Y. Jin, and D. Tian, "A systematic survey on human pose estimation: upstream and downstream tasks, approaches, lightweight models, and prospects," *Artif. Intell. Rev.*, vol. 58, no. 3, Jan. 2025, doi: 10.1007/s10462-024-11060-2.
- [19] D. Maji, S. Nagori, M. Mathew, and D. Poddar, "YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation Using Object Keypoint Similarity Loss," Apr. 2022, doi: 10.48550/arXiv.2204.06806.
- [20] F. C. Lin, H. H. Ngo, C. R. Dow, K. H. Lam, and H. L. Le, "Student behavior recognition system for the classroom environment based on skeleton pose estimation and person detection," *Sensors*, Aug. 2021, doi: 10.3390/s21165314.
- [21] A. Tharatipyakul, T. Srikaewsiew, and S. Pongnumkul, "Deep learning-based human body pose estimation in providing feedback for physical movement: A review," *Heliyon*, Aug. 2024, doi: 10.1016/j.heliyon.2024.e36589.