

Pengenalan dan Alih Aksara Jawa Pada Citra Menggunakan YOLOv11

1st Afif Faisal Alamsyah

Informatika
Telkom University

Surabaya, Indonesia

afiffaisal@student.telkomuniversity.ac.id

2nd Ardian Yusuf Wicaksono, S.Kom.,

M.Kom.

Informatika

Telkom University

Surabaya, Indonesia

ardiany@telkomuniversity.ac.id

3rd Pima Hani Safitri, S.Kom., M.Kom.

Informatika

Telkom University

Surabaya, Indonesia

phanisafitri@telkomuniversity.ac.id

Abstrak — Aksara Jawa sebagai salah satu warisan budaya Indonesia menghadapi ancaman kepunahan akibat rendahnya tingkat literasi serta dominasi alfabet Latin dalam kehidupan modern. Kompleksitas bentuk visual aksara, disertai kemiripan karakter antaraksara, turut menyulitkan proses identifikasi secara manual. Penelitian ini mengembangkan sistem deteksi dan alih aksara Jawa otomatis berbasis teknologi deep learning. Sistem dibangun menggunakan dua varian model YOLOv11, yaitu YOLOv11n dan YOLOv11x, dengan memanfaatkan dataset gabungan yang terdiri atas data primer sebanyak 49 kelas dan data sekunder sebanyak 20 kelas aksara dasar. Dataset melalui tahap preprocessing dan augmentasi data berupa blur dan noise sebelum dilakukan pelatihan model selama 100 epoch dengan konfigurasi optimal. Evaluasi kinerja model dilakukan menggunakan metrik precision, recall, F1-Score, dan mAP50. Hasil pengujian menunjukkan bahwa YOLOv11n memperoleh nilai precision 0,9473, recall 0,9847, F1-Score 0,9656, dan mAP50 0,9818, sedangkan YOLOv11x mencapai precision 0,9850, recall 0,9925, F1-Score 0,9887, dan mAP50 0,9944. Kedua model mampu mendeteksi aksara dasar, pasangan, dan sandhangan secara akurat serta mendukung proses alih aksara ke huruf Latin, sehingga berpotensi menjadi solusi berbasis teknologi dalam upaya pelestarian aksara Jawa melalui pengolahan citra.

Kata kunci — alih aksara, deep learning, pengenalan aksara Jawa, pelestarian budaya, pengolahan citra, YOLOv11.

I. PENDAHULUAN

Meningkatnya kebutuhan pelestarian budaya mendorong pemanfaatan teknologi untuk mengatasi keterbatasan dalam mengenali aksara tradisional. Aksara Jawa sebagai warisan budaya Indonesia mengalami ancaman kepunahan karena rendahnya tingkat literasi dan penggunaan dalam kehidupan modern, terutama di kalangan generasi muda [1]. Kondisi ini diperburuk oleh dominasi penggunaan alfabet Latin dalam berbagai aspek kehidupan sehari-hari, menyebabkan berkurangnya eksistensi aksara tradisional dalam komunikasi tulis [2]. Selain itu, perubahan pola pendidikan dan minimnya media pembelajaran berbasis teknologi turut mempercepat penurunan keterampilan membaca dan menulis Aksara Jawa [3]. Penggunaan media digital dalam pembelajaran Aksara Jawa terbukti mampu meningkatkan literasi budaya di

kalangan pelajar dan menjadi sarana adaptif dalam era digitalisasi pendidikan [4].

Aksara Jawa memiliki bentuk karakter visual yang kompleks dengan banyak lekukan dan kemiripan antar aksara, menjadikannya sulit untuk dibaca maupun diidentifikasi secara manual [5]. Tantangan ini mendorong kebutuhan akan solusi berbasis teknologi yang mampu mendeteksi Aksara Jawa secara otomatis guna mendukung pelestarian budaya [6]. Kemajuan teknologi kecerdasan buatan, khususnya dalam bidang pengolahan citra, membuka peluang besar dalam pengenalan aksara secara otomatis.

Dalam menghadapi tantangan tersebut, diperlukan pendekatan teknologi yang mampu mengenali karakter visual kompleks dari Aksara Jawa secara otomatis dan efisien. Salah satu metode yang berkembang pesat dalam deteksi objek adalah YOLO (You Only Look Once), yang memungkinkan deteksi cepat dalam satu tahap. YOLO adalah salah satu arsitektur CNN. Pendekatan berbasis deep learning menawarkan potensi signifikan untuk mengatasi keterbatasan ini, di mana model CNN kustom terbukti mampu meningkatkan akurasi pengenalan Aksara Jawa secara efisien [7]. Selain itu, YOLOv11 juga telah berhasil diterapkan untuk mendeteksi aksara Sunda secara otomatis, menunjukkan efektivitasnya dalam pengenalan aksara tradisional lain dengan hasil akurasi yang tinggi [8]. Dengan YOLOv11 sebagai generasi terbaru menunjukkan peningkatan signifikan dalam akurasi dan kecepatan dibandingkan versi sebelumnya [9]. YOLOv11 juga telah diterapkan dalam bidang medis seperti deteksi tumor dari citra medis, menunjukkan akurasi tinggi dan efisiensi waktu dalam proses diagnosis otomatis [10].

Sistem deteksi Aksara Jawa yang dikembangkan menggunakan YOLOv11n dan YOLOv11x menunjukkan hasil yang stabil dalam mengenali berbagai bentuk aksara, di mana YOLOv11n mampu bekerja secara efisien dengan performa yang ringan, sedangkan YOLOv11x memberikan hasil deteksi yang lebih kuat dan detail terhadap karakter yang kompleks. Kedua model sama-sama berhasil melakukan proses identifikasi aksara dasar, pasangan, dan sandhangan

secara otomatis serta mendukung alih aksara ke huruf Latin dengan konsistensi yang baik. Hasil ini menunjukkan bahwa baik model ringan maupun model berkapasitas besar dapat memberikan kontribusi dalam pengembangan sistem pelestarian Aksara Jawa berbasis teknologi, dengan keunggulan masing-masing sesuai kebutuhan penggunaan.

II. KAJIAN TEORI

Bab ini membahas landasan teori yang digunakan dalam penelitian, meliputi konsep Aksara Jawa, metode deep learning berbasis CNN dan YOLOv11 untuk deteksi serta klasifikasi objek, serta metrik evaluasi kinerja model seperti mAP, precision, recall, dan F1 score sebagai dasar analisis hasil penelitian.

A. Aksara Jawa

Aksara Jawa merupakan salah satu warisan budaya Nusantara yang lahir dari perkembangan aksara Brahmi India, yang kemudian mengalami akulturasi dengan budaya lokal Jawa [11]. Secara struktur, Aksara Jawa berbentuk silabis di mana satu karakter mewakili satu suku kata, memperlihatkan perbedaan mendasar dengan alfabet Latin yang berbasis fonem tunggal [12]. Aksara Jawa terdiri dari huruf carakan, pasangan, dan tanda baca khusus yang memperkaya nuansa penulisan dan pelafalan [13].

B. Aliran Gas pada Sistem Pipa

Aliran gas dalam sistem pipa dipengaruhi oleh perbedaan tekanan, suhu, laju aliran, serta sifat fisik dan komposisi gas. Perubahan parameter tersebut dapat memengaruhi stabilitas aliran dan performa distribusi gas menuju fasilitas hilir. Data tekanan, temperatur, dan komposisi gas sering dimanfaatkan sebagai input analisis prediktif berbasis data (Putra).

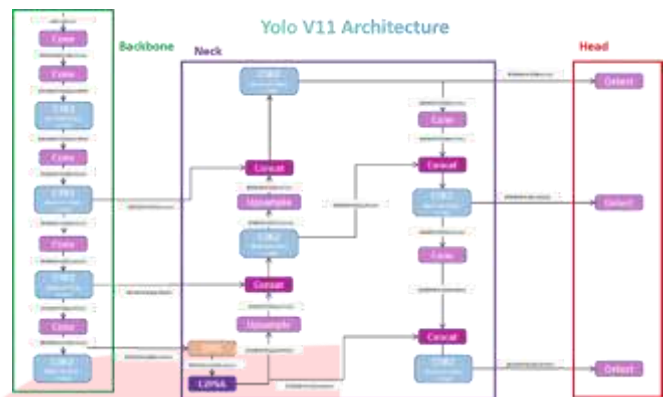
C. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan jenis jaringan saraf tiruan yang dirancang untuk memproses data yang berbentuk grid, seperti gambar, dengan menggunakan operasi konvolusi untuk mengekstraksi fitur penting [14]. CNN bekerja dengan prinsip memanfaatkan lapisan konvolusi dan pooling secara bergantian untuk mengurangi dimensi data sekaligus mempertahankan informasi penting yang ada [15]. Setiap lapisan konvolusi dalam CNN berfungsi untuk mengidentifikasi pola lokal, seperti tepi, tekstur, atau bentuk, yang nantinya digabungkan untuk membentuk pemahaman yang lebih kompleks tentang objek [16].

D. You Only Look Once (YOLO)

YOLO (You Only Look Once) adalah algoritma deteksi objek berbasis deep learning yang bekerja dalam satu tahap (single-stage) dengan cara mengklasifikasi dan memprediksi lokasi bounding box dalam satu proses inferensi yang efisien [17]. Sejak versi pertamanya, YOLO telah mengalami perkembangan pesat dari YOLOv1 hingga YOLOv11 dengan berbagai peningkatan pada arsitektur, akurasi, dan efisiensi komputasi [18]. Perubahan penting seperti anchor boxes di YOLOv2, multi-scale prediction pada YOLOv3, dan Cross-Stage Partial Networks di YOLOv4 meningkatkan kemampuan model dalam mengenali objek beragam ukuran

dan bentuk [19]. Hingga YOLOv10, setiap generasi YOLO berfokus pada peningkatan kecepatan real-time dan stabilitas deteksi dalam kondisi kompleks [20].



GAMBAR 1
(ARSITEKTUR YOLOV11)

YOLO (You Only Look Once) adalah algoritma deteksi objek berbasis deep learning yang bekerja dalam satu tahap (single-stage) dengan cara mengklasifikasi dan memprediksi lokasi bounding box dalam satu proses inferensi yang efisien [21]. Sejak versi pertamanya, YOLO telah mengalami perkembangan pesat dari YOLOv1 hingga YOLOv11 dengan berbagai peningkatan pada arsitektur, akurasi, dan efisiensi komputasi [18]. Perubahan penting seperti anchor boxes di YOLOv2, multi-scale prediction pada YOLOv3, dan Cross-Stage Partial Networks di YOLOv4 meningkatkan kemampuan model dalam mengenali objek beragam ukuran dan bentuk [19]. Hingga YOLOv10, setiap generasi YOLO berfokus pada peningkatan kecepatan real-time dan stabilitas deteksi dalam kondisi kompleks [20].

E. mean Average Precision (mAP)

Mean Average Precision (mAP) merupakan metrik evaluasi utama yang digunakan untuk menilai performa model deteksi objek secara keseluruhan. Nilai mAP dihitung dari rata-rata Average Precision (AP) pada setiap kelas objek, di mana AP merepresentasikan luas area di bawah kurva hubungan antara precision dan recall (Padilla et al., 2021).

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (II - 1)$$

Keterangan:

- C: jumlah total kelas objek yang dievaluasi.
- AP_i: nilai Average Precision untuk kelas ke-i, yaitu area di bawah kurva precision-recall untuk kelas tersebut.

F. Precision, Recall, dan F1 score

Dalam evaluasi kinerja model klasifikasi berbasis confusion matrix, beberapa metrik yang umum digunakan untuk mengukur kemampuan model dalam melakukan prediksi adalah precision, recall, dan F1 score (Amin, 2022). Rumus untuk menghitung precision ditunjukkan pada persamaan II-2.

$$Precision = \frac{TP}{TP + FP} \quad (II - 2)$$

Penjelasan:

- TP = jumlah sampel kelas positif yang diklasifikasikan secara benar.
- FP = jumlah sampel kelas negatif yang keliru diprediksi sebagai kelas positif.
- TP+FP = menunjukkan total sampel yang diprediksi sebagai kelas positif oleh model.

Rumus untuk menghitung recall ditunjukkan pada persamaan II-3.

$$Recall = \frac{TP}{TP + FN} \quad (II - 3)$$

Penjelasan:

- TP = jumlah sampel kelas positif yang diklasifikasikan secara benar.
- FN = jumlah sampel kelas positif yang salah diklasifikasikan sebagai kelas negatif.
- TP+FP = merepresentasikan keseluruhan sampel yang benar-benar berasal dari kelas positif.

Rumus untuk menghitung F1 Score ditunjukkan pada persamaan II-4.

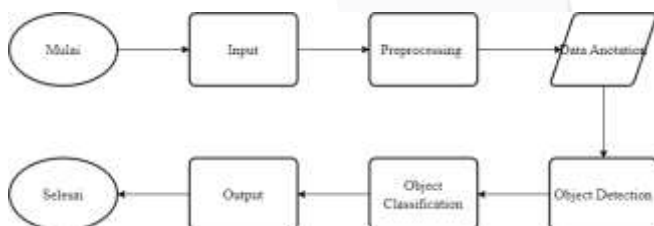
$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (II - 4)$$

Penjelasan:

- Precision = merepresentasikan tingkat ketepatan prediksi positif yang dihasilkan oleh model.
- Recall = merepresentasikan kemampuan model dalam mengidentifikasi seluruh sampel kelas positif.

III. METODE

Memberikan gambaran alur kerja sistem pengenalan Aksara Jawa yang mencakup tahapan input data, preprocessing, anotasi, deteksi objek, klasifikasi, hingga menghasilkan output berupa hasil pengenalan karakter. Alur proses sistem secara keseluruhan ditunjukkan pada gambar III.1.



GAMBAR 2
(ALUR PROSES SYSTEM)

Tahap input merupakan langkah awal dalam sistem pengolahan citra Aksara Jawa, di mana data yang digunakan berupa citra Aksara Jawa hasil tulisan tangan dan penulisan digital. Data primer terdiri dari 20 citra yang mencakup 49 kelas karakter, meliputi aksara dasar, pasangan, dan sandhangan. Selain itu, digunakan data sekunder sebanyak 20 citra dari Kaggle yang hanya mencakup aksara dasar untuk menambah variasi bentuk karakter. Seluruh citra disimpan dalam format JPG atau PNG dan digunakan sebagai masukan utama dalam proses pelatihan dan evaluasi model.

Pada tahap preprocessing, citra dilakukan penyesuaian ukuran (resize) agar sesuai dengan kebutuhan input model

serta penerapan augmentasi berupa blur dan noise. Proses ini bertujuan untuk meningkatkan variasi data dan ketahanan model terhadap kondisi citra yang beragam. Seluruh citra kemudian dianotasi secara manual dengan pemberian label kelas sesuai karakter Aksara Jawa yang terdapat pada gambar. Hasil anotasi ini menjadi dasar pembelajaran model YOLOv11 dalam mengenali dan membedakan karakter yang memiliki kemiripan bentuk.

Setelah proses pelatihan, sistem melakukan object detection untuk mendeteksi posisi huruf Aksara Jawa dalam citra menggunakan bounding box. Objek yang terdeteksi selanjutnya diklasifikasikan ke dalam kelas karakter yang sesuai berdasarkan ciri visualnya. Tahap ini menghasilkan identifikasi huruf Aksara Jawa secara otomatis. Output akhir ditampilkan dalam bentuk visual berupa citra dengan bounding box dan label huruf sebagai hasil deteksi dan klasifikasi sistem.

IV. HASIL DAN PEMBAHASAN

Sebelum dilakukan evaluasi lebih lanjut, terlebih dahulu ditampilkan hasil training dari kedua model untuk melihat performa awal selama proses pembelajaran, seperti yang ditunjukkan pada Tabel IV.1.

TABEL 1
(HASIL TRAINING MODEL YOLOV11N DAN YOLOV11X)

Model	Train Precision (P)	Train Recall (R)	Train mAP50
YOLOv11n	0.936	0.981	0.981
YOLOv11x	0.989	0.993	0.995

Hasil training menunjukkan bahwa YOLOv11n memiliki performa yang sangat baik meskipun menggunakan jumlah parameter yang lebih sedikit. Nilai precision 0.936 dan recall 0.981 menandakan kemampuan model dalam mempelajari karakter aksara Jawa secara konsisten. Nilai mAP50 yang mencapai 0.981 menunjukkan akurasi tinggi dalam memprediksi bounding box selama pelatihan. Secara keseluruhan, YOLOv11n menawarkan efisiensi komputasi dengan performa yang sangat memuaskan.

Sebaliknya, YOLOv11x menunjukkan performa training yang lebih unggul di seluruh metrik yang diuji. Dengan nilai precision 0.989 dan recall 0.993, model ini terbukti belajar lebih efektif terhadap kompleksitas bentuk aksara. Nilai mAP50 yang mencapai 0.995 menunjukkan akurasi sangat tinggi dalam memprediksi bounding box pada berbagai variasi data. Hal ini mengindikasikan bahwa YOLOv11x memiliki kapasitas yang lebih besar untuk menangkap detail visual yang kompleks dari karakter aksara Jawa.

Tahap validasi digunakan untuk mengevaluasi kemampuan model dalam mengenali data baru yang tidak digunakan selama pelatihan. Hasil lengkap proses validasi ditunjukkan pada Tabel IV.2.

TABEL 2
(HASIL VALIDASI MODEL YOLOV11N DAN YOLOV11X)

Model	Validation Precision	Validation Recall (R)	Validation F1-Score	Validation mAP50
YOLOv11n	0.9366	0.9805	0.9580	0.9809
YOLOv11x	0.9890	0.9932	0.9911	0.9949

Berdasarkan hasil validasi, YOLOv11n menunjukkan kemampuan generalisasi yang sangat baik dengan nilai precision 0.9366 dan recall 0.9805. Nilai mAP50 sebesar 0.9809 juga mengindikasikan bahwa prediksi bounding box masih berada pada tingkat akurasi yang sangat tinggi. F1-Score sebesar 0.9580 menunjukkan keseimbangan yang baik antara precision dan recall, membuktikan model mampu mempertahankan performa yang konsisten pada data baru.

YOLOv11x kembali menunjukkan performa terbaik dengan nilai precision dan recall yang lebih tinggi dibanding YOLOv11n. Nilai mAP50 yang mencapai 0.9949 dan F1-Score 0.9911 menandakan bahwa model mampu mempertahankan akurasi hampir sempurna bahkan pada data validasi. Hal ini menunjukkan bahwa model memiliki kemampuan generalisasi yang luar biasa terhadap variasi bentuk aksara Jawa. Selisih performa antara training dan validation yang sangat kecil pada YOLOv11x (kurang dari 0.001) menunjukkan bahwa model tidak mengalami overfitting signifikan.

Tahap testing dilakukan untuk mengetahui performa akhir model. Hasil lengkap pengujian dapat dilihat pada Tabel IV.3. yang menunjukkan perbandingan performa antara YOLOv11n dan YOLOv11x.

TABEL 2
(HASIL TESTING MODEL YOLOV11N DAN YOLOV11X)

Model	Testing Precision	Testing Recall (R)	Testing F1-Score	Testing mAP50
YOLOv11n	0.9473	0.9847	0.9656	0.9818
YOLOv11x	0.9850	0.9925	0.9887	0.9944

Pada tahap testing, YOLOv11n menunjukkan akurasi yang sangat tinggi dengan F1-Score sebesar 0.9656. Nilai precision 0.9473 dan recall 0.9847 menunjukkan performa yang stabil selama pengujian. Nilai mAP50 sebesar 0.9818 mengindikasikan kemampuan model dalam mendeteksi objek dengan akurasi tinggi. Hal ini membuktikan bahwa YOLOv11n memiliki kemampuan generalisasi yang sangat baik dan dapat diandalkan untuk aplikasi deteksi aksara Jawa pada perangkat dengan sumber daya terbatas.

Sementara itu, YOLOv11x menghasilkan performa yang hampir sempurna pada tahap testing. Dengan F1-Score mencapai 0.9887, model ini menunjukkan konsistensi tinggi dalam mengenali berbagai karakter aksara Jawa. Precision 0.9850 dan recall 0.9925 membuktikan bahwa model mampu memberikan prediksi yang sangat akurat dengan tingkat kesalahan yang sangat rendah. Nilai mAP50 sebesar 0.9944 menunjukkan kemampuan deteksi yang nyaris sempurna, menjadikan YOLOv11x pilihan ideal untuk aplikasi yang memprioritaskan akurasi tinggi tanpa mempertimbangkan batasan komputasi.

V. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan mengenai pengenalan dan alih aksara Jawa pada citra berbasis YOLOv11, dapat disimpulkan sebagai berikut:

1. Sistem deteksi aksara Jawa pada citra berhasil dirancang dan diimplementasikan menggunakan dua varian model YOLOv11, yaitu YOLOv11n dan YOLOv11x. Sistem ini dikembangkan meliputi tahap preprocessing data, augmentasi citra berupa blur, noise, serta kombinasi keduanya, proses pelatihan model, hingga visualisasi hasil deteksi yang disajikan melalui GUI yang informatif dan mudah dipahami.
2. Kinerja kedua varian model YOLOv11 dalam mendeteksi aksara Jawa menunjukkan hasil yang sangat baik berdasarkan metrik evaluasi yang digunakan. Model YOLOv11x memperoleh nilai precision sebesar 0,9850, recall 0,9925, F1-Score 0,9887, dan mAP50 0,9944, sedangkan YOLOv11n mencapai nilai precision 0,9473, recall 0,9847, F1-Score 0,9656, dan mAP50 0,9818. Hasil tersebut membuktikan bahwa penerapan teknologi deep learning berbasis YOLOv11 efektif dalam mendukung pelestarian aksara Jawa melalui sistem deteksi otomatis yang akurat dan efisien.

REFERENSI

- [1] M. Matsuri *et al.*, "Analysis of Elementary School Students' Literacy Ability of Javanese Script Through the Use of Script Card Reviewed Based on Critical Thinking Skills," *AJJP*, vol. 15, no. 4, Dec. 2023, doi: 10.35445/alishlah.v15i4.4519.
- [2] S. A. Ori and H. W. Susianti, "Preserving Javanese Language and Culture in The Digital Age: Challenge and Future Prospects," *lacultour*, vol. 2, no. 2, pp. 79–88, Nov. 2023, doi: 10.52352/lacultour.v2i2.1212.
- [3] H. O. Fajrin, "THE INFLUENCE OF TECHNOLOGY-BASED MEDIA OF DIGITAL JAVANESE SCRIPT AMONG JUNIOR HIGH SCHOOL STUDENTS," *Jurnal Ilmiah WUNY*, vol. 5, no. 1, May 2023, doi: 10.21831/jwuny.v5i1.58374.
- [4] D. Iskandar, S. Hidayat, U. Jamaludin, and S. Mukti Leksono, "Javanese script digitalization and its utilization as learning media: an etnopedagogical approach," *J.IJMSEd*, vol. 1, no. 1, pp. 21–30, June 2023, doi: 10.59965/ijmsed.v1i1.24.
- [5] A. A. Rahmawati, A. Muhaimin, and D. A. Prasetya, "CLASSIFICATION OF JAVANESE

- NGLEGENA SCRIPT USING COMPLEXVALUED NEURAL NETWORK,” *jiko*, vol. 7, no. 1, pp. 30–35, Apr. 2024, doi: 10.33387/jiko.v7i1.7808.
- [6] M. A. Rasyidi, T. Bariyah, Y. I. Riskajaya, and A. D. Septyani, “Classification of handwritten Javanese script using random forest algorithm,” *Bulletin EEI*, vol. 10, no. 3, pp. 1308–1315, June 2021, doi: 10.11591/eei.v10i3.3036.
- [7] A. Susanto, I. U. W. Mulyono, C. A. Sari, E. H. Rachmawanto, D. R. I. M. Setiadi, and M. K. Sarker, “Improved Javanese script recognition using custom model of convolution neural network,” *IJECE*, vol. 13, no. 6, p. 6629, Dec. 2023, doi: 10.11591/ijece.v13i6.pp6629-6636.
- [8] R. Rosalina, N. Afriliana, W. H. Utomo, and G. Sahuri, “Deep learning utilization in Sundanese script recognition for cultural preservation,” *IJECS*, vol. 36, no. 3, p. 1759, Dec. 2024, doi: 10.11591/ijeecs.v36.i3.pp1759-1768.
- [9] R. Sapkota, Z. Meng, M. Churuvija, X. Du, Z. Ma, and M. Karkee, “Comprehensive Performance Evaluation of YOLOv12, YOLO11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments,” Feb. 25, 2025, *arXiv*: arXiv:2407.12040. doi: 10.48550/arXiv.2407.12040.
- [10] T. Saidani, “Efficient Tumor Detection in Medical Imaging Using Advanced Object Detection Model: A Deep Learning Approach,” *ijacsa*, vol. 16, no. 1, 2025, doi: 10.14569/ijacsa.2025.01601109.
- [11] N. Nabilah and F. Nikmah, “The RELATIONSHIP OF AKSARA JAWA AS LOCAL FOLKLORE WITH MODERATE AND PROGRESSIVE ISLAMIC EDUCATION,” *AICOIES*, vol. 1, no. 1, June 2022, doi: 10.18326/aicoies.v1i1.257.
- [12] H. A. Ikasanti, N. K. Umam, and I. Marzuki, “PENGEMBANGAN VIDEO PEMBELAJARAN UNTUK MEMBACA AKSARA JAWA PADA PESERTA DIDIK KELAS IV SEKOLAH DASAR,” *jtice*, vol. 5, no. 1, p. 1, May 2021, doi: 10.30587/jtice.v5i1.2786.
- [13] H. A. Nugroho, D. Sulaksono, and R. Rahmat, “Analisis Kesalahan Penulisan Aksara Jawa Pada Papan Nama Jalan di Kota Surakarta,” *sabpbj*, vol. 7, no. 1, p. 26, Mar. 2023, doi: 10.20961/sabpbj.v7i1.61367.
- [14] R. R. Karim, “IMPLEMENTASI KLASIFIKASI SENJATA TRADISIONAL JAWA BARAT MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) DENGAN METODE TRANSFER LEARNING,” *JITET*, vol. 12, no. 2, Apr. 2024, doi: 10.23960/jitet.v12i2.4166.
- [15] A. T. Akbar, S. Saifullah, and H. Prapcoyo, “Klasifikasi Ekspresi Wajah Menggunakan Covolutional Neural Network,” *JTIK*, vol. 11, no. 6, pp. 1399–1412, Dec. 2024, doi: 10.25126/jtik.1168888.
- [16] Y. N. Fuadah, I. D. Ubaidullah, N. Ibrahim, F. F. Taliningsing, N. K. Sy, and M. A. Pramuditho, “Optimasi Convolutional Neural Network dan K-Fold Cross Validation pada Sistem Klasifikasi Glaukoma,” *ELKOMIKA*, vol. 10, no. 3, p. 728, July 2022, doi: 10.26760/elkomika.v10i3.728.
- [17] R. Sapkota *et al.*, “YOLOv12 to Its Genesis: A Decadal and Comprehensive Review of The You Only Look Once (YOLO) Series,” Feb. 21, 2025, *arXiv*: arXiv:2406.19407. doi: 10.48550/arXiv.2406.19407.
- [18] Y. Luo, “The Evolution of YOLO: from YOLOv1 to YOLOv11 with a Focus on YOLOv7’s Innovations in Object Detection,” *TNS*, vol. 87, no. 1, pp. 82–90, Jan. 2025, doi: 10.54254/2753-8818/2025.20335.
- [19] R. Sapkota and M. Karkee, “Improved YOLOv12 with LLM-Generated Synthetic Data for Enhanced Apple Detection and Benchmarking Against YOLOv11 and YOLOv10,” Mar. 19, 2025, *arXiv*: arXiv:2503.00057. doi: 10.48550/arXiv.2503.00057.
- [20] R. Sapkota and M. Karkee, “Comparing YOLOv11 and YOLOv8 for instance segmentation of occluded and non-occluded immature green fruits in complex orchard environment,” Jan. 26, 2025, *arXiv*: arXiv:2410.19869. doi: 10.48550/arXiv.2410.19869.
- [21] R. Sapkota, Z. Meng, M. Churuvija, X. Du, Z. Ma, and M. Karkee, “Comprehensive Performance Evaluation of YOLOv12, YOLO11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments,” Feb. 25, 2025, *arXiv*: arXiv:2407.12040. doi: 10.48550/arXiv.2407.12040.