



Perbandingan Versi Terbaik YOLO Dalam Mendeteksi Jarak Spasi Antar Baris Tulisan Tangan

Comparison of the Best Version of YOLO in Detecting Spaces Between Lines of Handwriting

Zaid Romegar Mair^{1*}, Muhamad Ardi Rahmada²

^{1,2}Teknik Informatika, Ilmu Komputer dan Sains, Universitas Indo Global Mandiri, Kota Palembang, Indonesia

¹zaidromegar@uigm.ac.id, ²ardierchman@gmail.com

Abstract

Handwriting still holds significant value in education and psychology, especially in evaluating students' writing structure skills. One important indicator is the line spacing, which reflects the orderliness and neatness of the writing. Manual assessment of this aspect is prone to subjectivity and is less efficient. Therefore, this study aims to compare the performance of three versions of the You Only Look Once (YOLO) model, namely YOLOv5, YOLOv8, and YOLO 11 in detecting the distance between handwriting lines. A total of 90 samples of students' handwriting were collected, scanned, and annotated into two categories: "narrow" and "wide." The training process was conducted with uniform parameters and tested using evaluation metrics such as precision, recall, and mean Average Precision (mAP). The research results show performance differences among the three models, providing insights into the effectiveness of each YOLO version in detecting spatial elements of handwriting. YOLOv5x consistently has the highest performance across the three metrics, especially in Precision, with a value of around 0.75. YOLOv8x and YOLO 11x show comparable performance on mAP@0.5 and Recall, but lower than YOLOv5x. The YOLOv5x model has better capabilities for detecting and classifying handwritten objects than the other two models.

Keywords: YOLO; handwriting; line spacing; object detection; innovation; automatic evaluation

Abstrak

Tulisan tangan tetap memiliki nilai penting dalam pendidikan dan psikologi, khususnya dalam mengevaluasi kemampuan struktur penulisan siswa. Salah satu indikator penting adalah jarak spasi antar baris, yang mencerminkan keteraturan dan kerapihan tulisan. Penilaian manual terhadap aspek ini rentan terhadap subjektivitas dan kurang efisien. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja tiga versi model You Only Look Once (YOLO), yaitu YOLOv5, YOLOv8, dan YOLO 11 dalam mendeteksi jarak antar baris tulisan tangan. Sebanyak 90 sampel tulisan tangan mahasiswa dikumpulkan, dipindai, dan dianotasi menjadi dua kategori: "sempit" dan "lebar". Proses pelatihan dilakukan dengan parameter yang seragam dan diuji menggunakan metrik evaluasi seperti *precision*, *recall*, dan *mean Average Precision* (mAP). Hasil penelitian menunjukkan perbedaan performa di antara ketiga model, yang memberikan wawasan tentang efektivitas masing-masing versi YOLO dalam mendeteksi elemen spasial tulisan tangan. YOLOv5x secara konsisten memiliki performa tertinggi pada ketiga metrik, terutama pada *Precision*, dengan nilai sekitar 0.75. YOLOv8x dan YOLO 11x menunjukkan performa yang sebanding pada nilai mAP@0.5 dan *Recall*, tetapi lebih rendah daripada YOLOv5x. Model YOLOv5x memiliki kemampuan yang lebih baik untuk mendeteksi dan mengklasifikasikan objek tulisan tangan dibandingkan dengan dua model lainnya.

Kata kunci: YOLO; tulisan tangan; spasi antar baris; deteksi objek, inovasi; evaluasi otomatis

1. Pendahuluan

salah satu bentuk komunikasi penting dalam kehidupan sehari-hari [1]. Tulisan tangan menunjukkan kemampuan kognitif seseorang serta informasi tentang kepribadian penulisnya [2]. Keterampilan menulis tangan dalam pendidikan, dinilai bukan hanya dari kualitas estetika tetapi juga dari kemampuan untuk menyampaikan ide secara

sistematis [3]. Akibatnya, evaluasi tulisan tangan masih relevan dan penting untuk menilai kompetensi siswa secara objektif. Namun, penilaian manual terhadap tulisan tangan menghadapi berbagai tantangan, terutama dari segi subjektivitas dan efisiensi. Kualitas tulisan sering kali dinilai berdasarkan faktor seperti bentuk huruf, kelurusinan garis dasar, ukuran, dan jarak antar baris. Semua aspek ini dapat bervariasi tergantung pada kondisi

emosional, fisik, dan lingkungan penulis [4]. Evaluasi manual dalam skala besar, seperti ujian atau survei kepribadian, juga memerlukan waktu dan tenaga yang tidak sedikit [5]. Oleh karena itu, dibutuhkan pendekatan otomatis berbasis kecerdasan buatan untuk meningkatkan konsistensi dan efisiensi proses evaluasi.

Kemajuan teknologi pemrosesan citra dan pembelajaran mesin, khususnya *deep learning*, telah membuka peluang besar dalam analisis visual tulisan tangan. Teknik seperti segmentasi teks dan *Optical Character Recognition* (OCR) memungkinkan komputer mengenali struktur tulisan secara otomatis [6]. Namun, metode konvensional seperti *thresholding* atau *edge detection* masih memiliki keterbatasan dalam menangani variasi tulisan tangan yang kompleks dan tidak terstruktur [4].

Sebagai solusi, *Convolutional Neural Network* (CNN) telah terbukti efektif dalam mengenali pola tulisan tangan secara langsung dari data citra tanpa perlu ekstraksi fitur manual [7,13]. CNN mampu mendekripsi variasi bentuk huruf, ketebalan garis, serta spasi antar baris dengan tingkat akurasi tinggi [8]. CNN juga telah berhasil digunakan untuk pengenalan kepribadian melalui tulisan tangan tangan [9], klasifikasi aksara lokal [8], serta segmentasi teks dalam dokumen tidak berlabel [6]. Dalam implementasinya, CNN sering digunakan melalui *framework* deteksi objek seperti *You Only Look Once* (YOLO), yang dirancang untuk mengenali objek dalam citra secara cepat dan akurat. Beberapa versi YOLO seperti YOLOv5, YOLOv8, hingga YOLO 11 telah dikembangkan dengan berbagai peningkatan arsitektur dan performa. Meskipun banyak digunakan dalam deteksi objek umum, studi mengenai efektivitas masing-masing versi YOLO dalam mendekripsi elemen spesifik dalam tulisan tangan, seperti jarak spasi antar baris, masih terbatas.

Penelitian sebelumnya juga menunjukkan bahwa pendekatan berbasis jaringan saraf seperti *Radial Basis Function Neural Network* (RBFNN) [10] dan *Support Vector Machine* (SVM) [11] dapat digunakan untuk menganalisis fitur tulisan tangan. Meski efektif, metode-metode ini tetap memiliki keterbatasan dalam mengenali pola spasial kompleks secara langsung dibandingkan CNN. Selain kompleksitas bentuk tulisan, deteksi jarak antar baris juga menghadapi tantangan tambahan seperti garis dasar yang tidak lurus, tinta yang pudar, atau adanya noise pada kertas [12,14]. Hal ini mengharuskan model deteksi memiliki kemampuan adaptif serta didukung oleh teknik pra-pemrosesan yang memadai untuk meningkatkan akurasi [6,15]. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membandingkan performa beberapa versi YOLO dalam mendekripsi jarak spasi antar baris tulisan tangan. Evaluasi dilakukan dengan membandingkan metrik seperti *precision*, *recall*, dan *mean Average Precision* (mAP).

Dengan mengetahui versi YOLO yang paling optimal untuk tugas ini, diharapkan hasil penelitian dapat berkontribusi pada pengembangan sistem evaluasi otomatis tulisan tangan, khususnya di bidang pendidikan dan psikologi. Beberapa penelitian lainnya telah menunjukkan efektivitas metode deep learning [16,17, 18,19] dalam analisis tulisan tangan.

2. Metode Penelitian

Penelitian ini dilakukan untuk mengembangkan model deteksi jarak garis tulisan tangan menggunakan metode Convolutional Neural Network (CNN) yang diimplementasikan melalui framework YOLO [20]. Dimulai dengan mengidentifikasi masalah dan study literatur dalam melakukan pendekatan menggunakan metode yang relevan. Selanjutnya, dilakukan persiapan objek penelitian, media, dan alat yang diperlukan untuk memastikan sumber daya siap untuk digunakan. [15] Data tulisan tangan yang dikumpulkan dari mahasiswa semester satu Universitas Indo Global Mandiri kemudian diproses dengan memastikan pengambilan gambar tersebut konsisten dan memiliki kualitas tinggi. Label diberikan pada gambar selama tahap pemrosesan, yang memungkinkan model untuk mengidentifikasi, data akan dibagi menjadi dua bagian yaitu gambar latih dan gambar uji.

Data latihan digunakan untuk melatih model YOLO untuk menemukan jarak antar baris dengan benar. Setelah pelatihan selesai, model diuji dengan data uji untuk mengetahui seberapa baik dia dapat menemukan jarak antar baris. Hasil pengujian dianalisis untuk mengetahui seberapa baik model CNN dapat menemukan jarak antar baris dengan tepat, berikut ini gambaran alur yang akan dilakukan.



Gambar 1. Kerangka Kerja Penelitian

Alur kerja penelitian ini [12] bertujuan untuk membandingkan kinerja berbagai versi model You Only Look Once (YOLO), yaitu YOLOv5, YOLOv8, dan YOLO 11 dalam menentukan jarak antar baris tulisan tangan, seperti yang ditunjukkan pada Gambar 1. Fokus utama penelitian ini adalah mengevaluasi keandalan dan akurasi masing-masing versi YOLO dalam mengidentifikasi spasi antar baris tulisan tangan yang berbeda, yang disebabkan oleh perbedaan gaya dan fitur individu. Dengan menggunakan desain eksperimen, penelitian ini menggunakan pendekatan



kuantitatif. Data dikumpulkan dari sembilan puluh mahasiswa semester satu Universitas Indo Global Mandiri, yang diminta untuk menyalin paragraf teks ke dalam lembar yang telah disiapkan. Untuk memastikan bahwa proses pelatihan dan pengujian model berjalan dengan benar, data tulisan tangan yang diperoleh dibagi menjadi dua kelompok: dua puluh persen digunakan sebagai data latih dan dua puluh persen digunakan sebagai data uji. Pengumpulan data dilakukan secara standar dengan menggunakan pulpen Faber-Castell K7 berukuran 0,7 mm dan kertas HVS A4 merek Sidu dengan berat 80 gsm. Papan ujian kayu digunakan untuk menjaga stabilitas tulisan, dan pengambilan data dilakukan pada suhu dan pencahayaan yang cukup untuk mengurangi dampak tulisan tangan dari luar.



Gambar 2. Alat Yang Digunakan Dalam Pengambilan Data

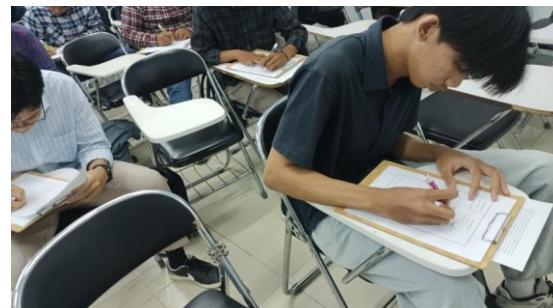
Selanjutnya, data yang ditulis dipindai dengan resolusi 300 dpi menggunakan printer multifungsi Epson L3250 untuk menghasilkan gambar dengan detail yang cukup tinggi. Tabel 1 menunjukkan spesifikasi teknis printer ini yang meningkatkan kualitas pemindaian.

Tabel 1. Spesifikasi Printer Epson L3250

Spesifikasi	Detail
Jenis Pemindai	<i>Flatbed colour image scanner</i>
Jenis Sensor	CIS (<i>Contact Image Sensor</i>)
Resolusi Optik	1200 x 2400 dpi
Area Pindai Maksimum	216 x 297 mm
Kedalaman Bit Pemindai (Warna)	48-bit input, 24-bit output
Kedalaman Bit Pemindai (Grayscale)	16-bit input, 8-bit output
Kedalaman Bit Pemindai (Hitam & Putih)	16-bit input, 1-bit output

Peserta akan diminta menulis dalam gaya dan kecepatan yang paling mereka sukai. Tujuannya adalah untuk menciptakan lingkungan penulisan yang alami sehingga data yang dihasilkan dapat benar-benar mencerminkan kebiasaan menulis setiap orang. pencahayaan lampu buatan untuk menjamin distribusi cahaya yang merata. Memastikan peserta tetap nyaman selama proses penulisan, suhu ruangan diatur oleh

pendingin udara (AC) pada 16 derajat Celsius. Kertas HVS berukuran A4 dibagikan ke peserta, dengan desain kotak pembatas dalam kertas, untuk memastikan tulisan tetap berada dalam batas yang ditentukan. Selain itu, setiap peserta akan diberikan lembar teks yang harus ditulis ulang. Sebelum proses penulisan dimulai, peneliti akan memberikan penjelasan menyeluruh tentang prosedur dan petunjuk pengisian kertas. Ini termasuk petunjuk tentang posisi duduk dan teknik menulis yang disarankan untuk konsistensi.



Gambar 3. Teknik Pengumpulan Data

3. Hasil dan Pembahasan

Penelitian ini menghasilkan model deteksi jarak antar baris tulisan tangan dengan menerapkan arsitektur *Convolutional Neural Network* (CNN) dalam framework YOLO, yang dibandingkan antara beberapa versi seperti YOLOv5, YOLOv8, dan YOLO 11. Perbandingan dilakukan untuk mengevaluasi keandalan masing-masing versi dalam menghadapi variasi gaya tulisan tangan. Data yang digunakan terdiri dari 80 citra tulisan tangan mahasiswa tingkat awal Universitas Indo Global Mandiri. Contoh sampel data yang digunakan dalam pelatihan dan pengujian model ditampilkan pada Gambar 4.

3.1. Pemilihan Thresholding untuk Crop Data

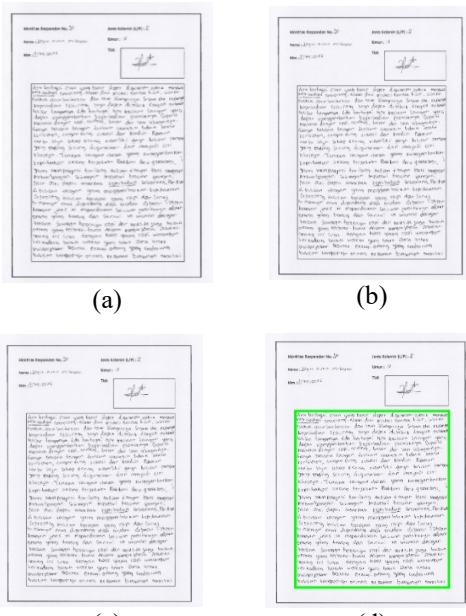
Teknik dalam mendapatkan area tulisan yang sesuai, maka dilakukan pemilihan *thresholding* dengan nilai awal 0.1 yang kemudian dinaikkan bertahap. Proses ini menggunakan pendekatan segmentasi berbasis intensitas piksel, di mana nilai *threshold* menentukan ambang batas untuk memisahkan piksel *foreground* (tulisan) dari *background* (kertas). Teknik *thresholding* ini sangat penting karena akurasi deteksi objek tulisan bergantung pada sejauh mana sistem



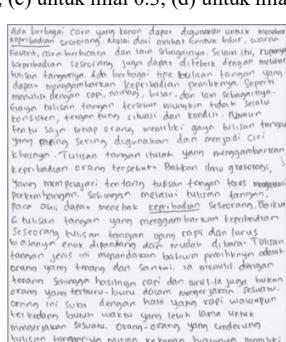
mampu mengekstraksi kontur tulisan secara utuh dan bersih. Selain itu, proses *cropping* dilakukan berdasarkan *bounding box* hasil segmentasi thresholding yang telah disesuaikan. Tujuan adanya *bounding box* ini adalah untuk memastikan bahwa hanya *area* yang relevan, yaitu tulisan tangan mahasiswa, yang diambil dan digunakan dalam pelatihan model. Dengan demikian, noise dari margin luar atau elemen tidak penting lainnya dapat diminimalkan.



Gambar 4. Sampel Data Yang Dipakai



Gambar 5. Hasil Thresholding: (a) untuk nilai 0.1, (b) untuk nilai 0.2, (c) untuk nilai 0.3, (d) untuk nilai 0.4



Gambar 6. Hasil Cropping Menggunakan Thresholding Dengan Nilai 0.4

Gambar 5 menunjukkan hasil visualisasi dari proses thresholding dengan berbagai nilai. Pada nilai threshold 0.1 hingga 0.3, *bounding box* yang dihasilkan belum mampu mendeteksi dan memisahkan secara jelas area tulisan mahasiswa yang ingin di-crop. Pada nilai 0.4, *bounding box* mampu mendeteksi dan memotong area tulisan dengan lebih presisi dan konsisten. Oleh karena itu, nilai threshold 0.4 dipilih sebagai parameter optimal untuk proses cropping data. Gambar 6 memperlihatkan hasil cropping akhir setelah penerapan thresholding dengan nilai tersebut.

3.2. Pemilihan Thresholding untuk Ukuran Tulisan dan Jarak Spasi

Terdapat tiga zona utama dalam pengambilan data tulisa tangan, yang terdiri dari: atas, tengah, dan bawah. Hanya zona tengah yang digunakan untuk mengukur jarak dan ukuran tulisan. Untuk mengukur sejauh mana nilai ambang (*t*) dapat memisahkan zona tengah yang sesuai, nilainya dimulai dengan nilai 0,1 dan kemudian dinaikkan secara bertahap. Tabel 2 menunjukkan hasil uji coba nilai ambang.

Tabel 2. Uji Coba Perbandingan Nilai Thresholding Ukuran Tulisan Dan Jarak Spasi

<i>t</i>	Hasil
0.1	Khasnya . Tulisan tangan itulah yang menggambarkan keribadian orang tersebut. Bahkan ilmu grafologi,
0.2	Khasnya . Tulisan tangan itulah yang menggambarkan keribadian orang tersebut. Bahkan ilmu grafologi,
0.3	Khasnya . Tulisan tangan itulah yang menggambarkan keribadian orang tersebut. Bahkan ilmu grafologi,
0.4	Khasnya . Tulisan tangan itulah yang menggambarkan keribadian orang tersebut. Bahkan ilmu grafologi,

Nilai threshold 0.2 dipilih sebagai parameter terbaik karena mampu memisahkan zona tengah tanpa mengikutsertakan bagian atas dan bawah.

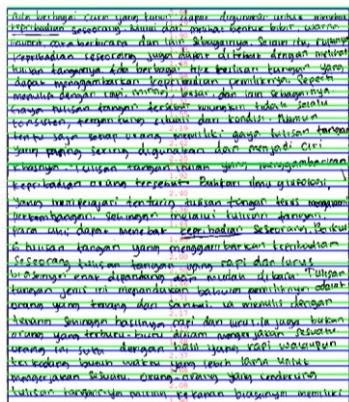
3.3. Penentuan Nilai Rasio

Nilai rasio diperoleh dengan membagi rata-rata jarak spasi antar baris dengan rata-rata ukuran tulisan dalam piksel. Nilai rasio ini digunakan untuk menentukan kategori jarak antar baris, di mana rasio di atas 2 dikategorikan sebagai "Lebar" dan di bawah 1.5 sebagai "Sempit". Gambar 7 menampilkan contoh hasil perhitungan rasio dengan MATLAB.

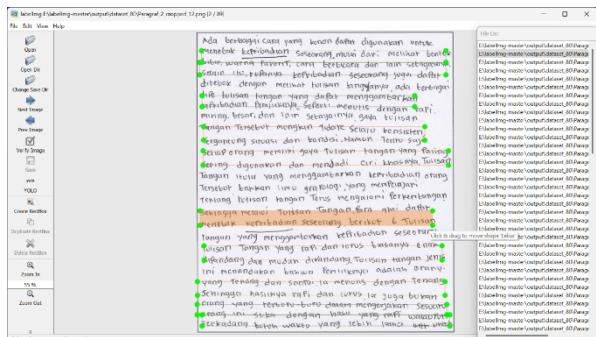
3.4. Proses Labelling Menggunakan LabelImg

Proses pelabelan pada tahap ini dilakukan menggunakan pustaka *LabelImg*, dengan penentuan *bounding box* didasarkan pada hasil perhitungan rasio antara ukuran tulisan dan jarak antar baris yang telah diperoleh melalui proses thresholding sebelumnya. Gambar 7 menyajikan proses pemberian label menggunakan *labelimg*.





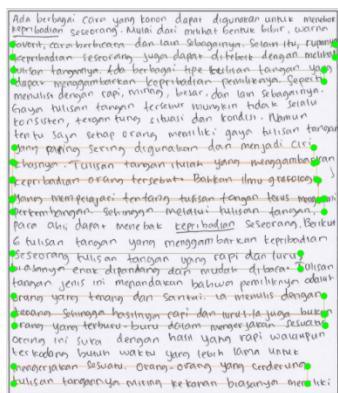
Gambar 7. Hasil Perhitungan Rasio Menggunakan MATLAB



Gambar 8. Proses Labeling Menggunakan LabelImg

3.5. Bounding Box Berdasarkan Perhitungan Rasio

Bounding box dibuat berdasarkan nilai rasio yang diperoleh melalui perhitungan pada MATLAB. Gambar 9 menyajikan contoh *bounding box* berdasarkan rasio.



Gambar 9. Contoh Bounding Box Berdasarkan Perhitungan Rasio

Area dengan *bounding box* berwarna oranye yang merepresentasikan tulisan dengan jarak spasi lebar, sedangkan *bounding box* abu-abu menunjukkan tulisan dengan jarak spasi sempit. Proses anotasi citra dilakukan menggunakan *library LabelImg* dalam format anotasi YOLO. Setelah anotasi selesai, data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian.

3.6. Pelatihan Data Menggunakan YOLO

Model yolo yang dipakai adalah YOLOv5x, YOLOv8x dan YOLO 11x dilatih dengan menggunakan 64 citra latih dengan total 484 objek *bounding box*. Rasio data latih 80%, percobaan dilakukan sebanyak 1 kali pada masing-masing model YOLO tersebut. Proses pelatihan model YOLO dimulai dengan dataset yang telah dianotasi dengan format yang sesuai dengan spesifikasi YOLO, termasuk file konfigurasi dan label *bounding box*. Dataset kemudian dibagi menjadi data latihan dan data validasi untuk memastikan bahwa model dapat digeneralisasi. Setelah itu, file konfigurasi YOLO diubah untuk menyesuaikan arsitektur jaringan, jumlah kelas, dan parameter pelatihan seperti kecepatan belajar, ukuran batch, dan gambar. Metode pelatihan menggunakan struktur seperti *Darknet* atau *PyTorch*, di mana model membaca dataset, menggunakan jaringan *convolutional* untuk memproses fitur, dan menggunakan algoritma optimisasi seperti SGD (*Stochastic Gradient Descent*) atau Adam untuk memperbarui bobot. Model menilai kinerjanya dengan metrik seperti fungsi loss, dan hasilnya divalidasi secara berkala menggunakan data validasi untuk mengawasi kemampuan mendekripsi spasi sebagai spasi dengan kategori sempit atau lebar. Hasil dari berbagai percobaan disajikan dalam Tabel 3 dan Gambar 10 sebagai grafik perbandingan dari pelatihan dari YOLOv5, YOLOv8 dan YOLO 11.

Tabel 3. Perbandingan Hasil Train YOLO

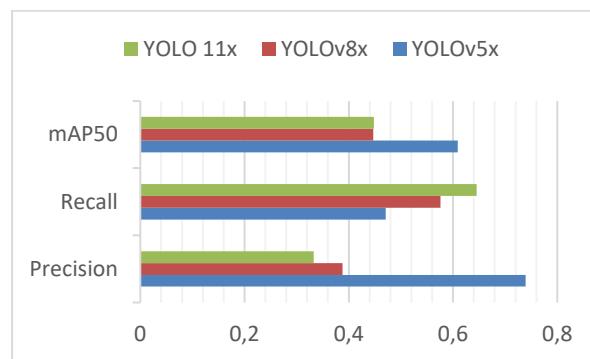
Model	Precision	Recall	mAP50
YOLOv5x	0.739	0.471	0.609
YOLOv8x	0.388	0.576	0.447
YOLO 11x	0.333	0.645	0.448

Dari perbandingan hasil YOLO pada tabel 3 diatas terlihat bahwa nilai presisi tertinggi terdapat pada YOLOv5x, sedangkan persentase tinggi Recall yaitu YOLO 11x, dan mAP50 adalah YOLOv5X. Grafik perbandingan bisa terlihat pada gambar 10. Dari hasil tabel diatas bahwa YOLOv5x memiliki nilai precision tertinggi (0.739) dibandingkan dua model lainnya, yaitu karena arsitektur backbone CSPDarknet53 dan head model yang lebih stabil dalam mengklasifikasikan objek. sedangkan YOLOv11x memiliki recall tinggi (0.645), namun precisionnya rendah (0.333), yaitu karena kemampuan deteksi yang sensitif tetapi yang memiliki kecenderungan overfitting terhadap *background*.

Gambar 10 menunjukkan perbandingan performa tiga varian model YOLO, yaitu YOLOv5x, YOLOv8x, dan YOLO 11x terhadap tiga metrik evaluasi: mAP@0.5, Precision, dan Recall. Metrik yang digunakan untuk mengevaluasi model ini adalah *mean Average Precision (mAP)*, *Precision*, *Recall*. *mAP* mengukur rata-rata akurasi deteksi objek dengan



mempertimbangkan *precision* dan *recall* pada setiap kelas, sehingga mencerminkan kemampuan keseluruhan model dalam mendeteksi dan mengklasifikasikan kategori jarak spasi. *Precision* mengukur sejauh mana prediksi model sesuai dengan kategori yang benar, menunjukkan proporsi deteksi yang akurat dibandingkan dengan jumlah total prediksi. Sementara itu, *Recall* mengukur sejauh mana model berhasil mendeteksi semua objek yang relevan, menunjukkan sensitivitas model terhadap kategori yang ada. Dari grafik terlihat bahwa YOLOv5x secara konsisten memberikan performa tertinggi pada ketiga metrik, terutama pada Precision yang mencapai nilai mendekati 0.75. Di sisi lain, YOLOv8x dan YOLO 11x menunjukkan performa yang cukup seimbang pada nilai mAP@0.5 dan Recall, namun berada di bawah YOLOv5x. Hal ini menunjukkan bahwa model YOLOv5x lebih mampu dalam mendeteksi dan mengklasifikasi objek tulisan tangan secara lebih presisi dibandingkan dua model lainnya.



Gambar 10. Grafik Perbandingan Pelatihan Model YOLOv5, YOLOv8 dan YOLO 11

3.7. Pengujian Data Menggunakan YOLO

Pengujian data menggunakan YOLO bertujuan untuk mengevaluasi kinerja model YOLO yang dibangun berdasarkan variasi jarak antara baris tulisan tangan. Pengujian dilakukan secara terpisah untuk setiap model. Pengujian menggunakan 20% dari total data, untuk memastikan bahwa setiap model diuji pada data yang memiliki karakteristik yang sama dengan data pelatihannya, tetapi tanpa label. Pengujian data dilakukan menggunakan tiga versi model YOLO, yaitu YOLOv5, YOLOv8, dan YOLOv11, untuk membandingkan performa masing-masing dalam mendeteksi jarak antar baris tulisan tangan. Setiap versi diuji menggunakan data uji yang telah dipisahkan sebelumnya, dengan fokus pada evaluasi akurasi deteksi terhadap variasi gaya tulisan tangan mahasiswa. Hasil pengujian dari ketiga versi model tersebut disajikan pada Tabel 4 dan Tabel 5.

Tabel 4. Hasil Pengujian Berbagai Versi YOLO

Kelas	Objek	YOLOv5	YOLOv8	YOLO11
Lebar	103	0.592	0.591	0.588
Sempit	101	0.351	0.302	0.305

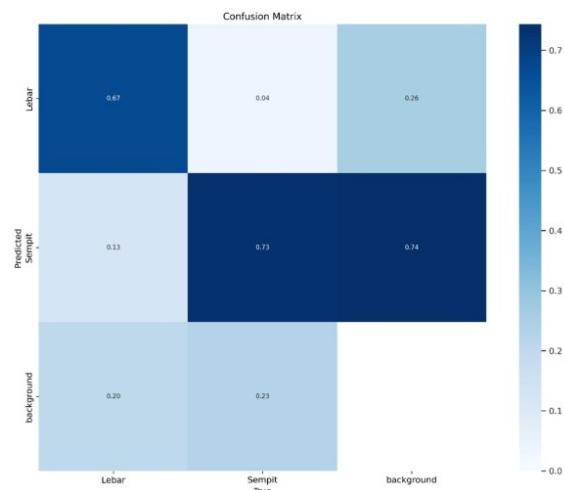
Rata-rata Total	204	0.472	0.447	0.446
-----------------	-----	-------	-------	-------

Pada tabel 4 diatas terdapat 2 jenis spasi berupa lebar dan sempit dengan jumlah masing-masing objek yaitu 103 untuk kelas lebar dan 101 untuk kelas sempit. Total keseluruhan objek yaitu sebanyak 204 jarak spasi. Dengan adanya pelatihan dan pengujian pada masing-masing versi YOLO, maka dapat dihasilkan nilai *precision* dan *recall* pada masing-masing model sebagaimana ditampilkan pada tabel 5 dibawah ini.

Tabel 5. Perbandingan nilai Precision dan Recall pada masing-masing kelas antara model YOLOv5, YOLOv8, dan YOLO 11

Kelas	YOLOv5 (P/R)	YOLOv8 (P/R)	YOLO11 (P/R)
Lebar	0.533 / 0.592	0.506 / 0.577	0.397 / 0.747
Sempit	0.301 / 0.515	0.276 / 0.574	0.264 / 0.535
Rata-Rata	0.417 / 0.554	0.391 / 0.576	0.330 / 0.646

Gambar 11 menampilkan *confusion matrix* dari model YOLOv5, YOLOv8, dan YOLOv11. Hasil menunjukkan bahwa model versi 5 memiliki kemampuan klasifikasi yang cukup baik dalam mendeteksi kelas Lebar dengan nilai prediksi benar sebesar 0.67, namun masih terdapat kekeliruan dalam memprediksi kelas Sempit dengan background yang cukup tinggi, dengan nilai masing-masing 0.26 dan 0.04 dari kelas Lebar.

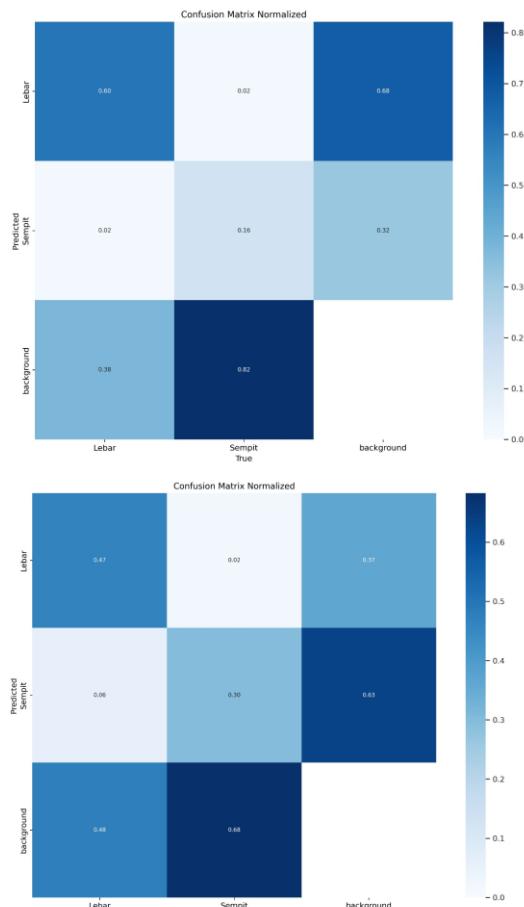


Gambar 11. Confusion Matrix YOLOv5

Sedangkan pada model versi 8 cenderung mendeteksi objek *background* dengan lebih dominan (prediksi *background* tinggi pada ketiga kelas), namun kurang akurat dalam membedakan kelas Lebar dan Sempit, dengan kesalahan prediksi silang yang signifikan, seperti 0.37 dari Sempit diprediksi sebagai Lebar. Pada gambar 11 juga memperlihatkan *confusion matrix* dari YOLO 11x, yang menunjukkan performa kurang stabil. Terlihat bahwa hampir semua kelas memiliki tingkat kesalahan yang tinggi dalam klasifikasi silang, seperti kelas Sempit yang justru lebih banyak



terdeteksi sebagai *background* (0.72). Hal ini menunjukkan model belum mampu membedakan fitur antar kelas secara konsisten.



Gambar 12. Confusion Matrix YOLOv8 dan YOLOv11

Gambar 12 menunjukkan confusion matrix dari model YOLOv5, YOLOv8, dan YOLO 11x. YOLOv5 menunjukkan performa terbaik dalam mendeteksi kelas Lebar dengan akurasi 0.67, meskipun masih terjadi kesalahan pada kelas *background*. Sementara itu, YOLOv8 cenderung mendeteksi *background* secara dominan namun kurang akurat membedakan kelas Lebar dan Sempit. Adapun YOLO 11x menunjukkan hasil paling rendah, dengan banyak prediksi salah terutama pada kelas Sempit yang sebagian besar diklasifikasikan sebagai *background*.

4. Kesimpulan

Berdasarkan hasil pelatihan dan pengujian, model YOLOv5 menunjukkan performa paling seimbang dengan nilai precision dan mAP yang relatif tinggi baik pada pelatihan maupun pengujian, yaitu precision 0,417 dan mAP@0.5 sebesar 0,472. Sementara itu, YOLOv8 dan YOLO 11 memiliki nilai recall yang lebih tinggi, terutama YOLO 11 dengan recall sebesar 0,646, namun precision-nya lebih rendah sehingga menghasilkan lebih banyak deteksi yang kurang tepat. Hal ini menunjukkan bahwa YOLOv5 lebih cocok

digunakan untuk deteksi jarak antar baris tulisan tangan dengan hasil yang stabil dan akurat, sedangkan YOLOv11 lebih sensitif dalam mendeteksi objek namun kurang akurat dalam prediksinya. Kontribusi spesifik penelitian ini adalah evaluasi komparatif ketiga model YOLO (v5, v8, dan v11) dalam konteks deteksi spasial tulisan tangan, dengan menggunakan data primer yang belum banyak diteliti. Sebagai saran dan lanjutan peneliti berikutnya dilakukan pengembangan ke dalam tiga kategori ("sempit", "tepat/pas", dan "lebar").

Reference

- [1] W. Wulansari and R. I. Khan, "Evaluasi Pelaksanaan Pelatihan Menulis Kreatif Berbentuk Dongeng Anak," *Edukatif J. Ilmu Pendidik.*, vol. 4, no. 3, pp. 4490–4498, 2022, doi: 10.31004/edukatif.v4i3.2711.
- [2] A. Gumilang and S. Agustin, "Deteksi Kepribadian Melalui Margin Pada Tulisan Tangan Menggunakan Random Forest," *INOVTEK Polbeng - Seri Inform.*, vol. 9, no. 1, pp. 311–325, 2024, doi: 10.35314/isi.v9i1.4176.
- [3] A. Anwar, "Grafologi Sebagai Alat Intervensi Praktik Pekerjaan Sosial Mikro Di SMKN 15 Kota Bandung," *J. Abdinas Kartika Wijayakusuma*, vol. 5, no. 2, pp. 431–439, 2024, doi: 10.26874/jakw.v5i2.443.
- [4] U. Rosyidah and N. Rochmawati, "Analisis Kepribadian Melalui Tulisan Tangan Menggunakan Metode Support Vector Machine," *J. Informatics Comput. Sci.*, vol. 1, no. 02, pp. 91–96, 2020, doi: 10.26740/jinacs.v1n02.p91-96.
- [5] Siswanto, *Menyngkap Kepribadian Lewat Tulisan Tangan*, 1st ed. Jakarta: Libri, 2010.
- [6] A. Drobny, B. Kurar Barakat, R. Saabni, R. Alaasam, B. Madi, and J. El-Sana, "Understanding Unsupervised Deep Learning for Text Line Segmentation," *Appl. Sci.*, vol. 12, no. 19, p. 9528, 2022, doi: 10.3390/app12199528.
- [7] G. Renton, Y. Soullard, C. Chatelain, S. Adam, C. Kermorvant, and T. Paquet, "Fully convolutional network with dilated convolutions for handwritten text line segmentation," *Int. J. Doc. Anal. Recognit.*, vol. 21, no. 3, pp. 177–186, 2018, doi: 10.1007/s10032-018-0304-3.
- [8] R. Aryanto, M. Alfan Rosid, and S. Busono, "Penerapan Deep Learning untuk Pengenalan Tulisan Tangan Bahasa Aksara Lota Ende dengan Menggunakan Metode Convolutional Neural Networks," *J. Inf. dan Teknol.*, vol. 5, no. 1, pp. 258–264, 2023, doi: 10.37034/jidt.v5i1.313.
- [9] Y. V. Guntara, Syamsuryadi, and Sukemi, "Pengenalan Kepribadian Melalui Tulisan Tangan Menggunakan Convolutional Neural Network Dengan LS Classifiers," *Digit. Zo. J. Teknol. Inf. dan Komun.*, vol. 14, no. 2, pp. 151–167, 2023, doi: 10.31849/digitalzone.v14i2.15193.
- [10] A. Magfiroh, "Pengenalan Kepribadian Seseorang Melalui Bentuk Tulisan Tangan Menggunakan Metode Radial Basis Function Neural Network (RBFNN)," *Zeta - Math J.*, vol. 7, no. 1, pp. 34–41, 2022, doi: 10.31102/zeta.2022.7.1.34-41.
- [11] T. Faizal et al., "Implementasi Metode Support Vector Machine Dalam Mendeteksi Kepribadian," vol. 4, no. 2, pp. 196–203, 2022.
- [12] Mair, Z. R., Cholil, W., Yulianti, E., & Marcelina, D. (2023). *Convolutional Neural Network Analysis on Handwriting Patterns and Its Relationship to Personality: A Systematic Review*. Paper presented at the 2023 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), Jakarta Selatan, Indonesia, 308–312.
- [13] Mair, Z. R., & Irfani, M. H. (2023). *Permainan INGBAS (gunting, batu, kertas) menggunakan arsitektur convolutional neural network*. *Jurnal Teknik Informatika dan Sistem Informasi*, 10 (1), 4322. ISSN: 2407-XXXX. Diterbitkan oleh JATISI.
- [14] Mair, Z. R. (2022). *Kajian Bakat Anak Melalui Pola Tulisan Tangan dengan Algoritma Convolutional Neural Network*.



- Jurnal Komputasi, 10(2), 66–71. <https://doi.org/10.23960/komputasi.v10i2.3157>
- [15] N. P. V. V. Prashanti, I. G. Santi Astawa, A. A. I. Ngurah Eka Karyawati, I. W. Santiyasa, I. B. Gede Dwidasmara, and I. W. Supriana, “Segmentasi Baris Aksara Bali Pada Citra Lontar,” JELIKU (Jurnal Elektron. Ilmu Komput. Udayana), vol. 10, no. 2, p. 205, 2022, doi: 10.24843/jlk.2021.v10.i02.p03.
- [16] Dixit, R., Kushwah, R., & Pashine, S. (2020). Handwritten Digit Recognition using Machine and Deep Learning Algorithms. In International Journal of Computer Applications (Vol. 176, Issue 42).
- [17] Ibrahim, R. T., & Ramo, F. M. (2023). Hybrid Intelligent Technique with Deep Learning to Classify Personality Traits. International Journal of Computing and Digital Systems, 13(1), 231–244. <https://doi.org/10.12785/ijcds/130119>
- [18] Masruroh, S. U., Syahid, M. F., Munthaha, F., Muhamram, A. T., & Putri, R. A. (2023). Deep Convolutional Neural Networks Transfer Learning Comparison on Arabic Handwriting Recognition System. JOIV : Int. J. Inform. Visualization, 7(2), 330–337. www.joiv.org/index.php/joiv
- [19] Mondal, R., Malakar, S., Barney Smith, E. H., & Sarkar, R. (2022). Handwritten English word recognition using a deep learning based object detection architecture. Multimedia Tools and Applications, 81(1), 975–1000. <https://doi.org/10.1007/s11042-021-11425-7>
- [20] Alqoyum, M. A., Wibowo, A., & Sarwoko, E. A. (2023). YOLOv4 algorithm implementation based on darknet and optical character recognition on vehicle license plate detection. 050023.<https://doi.org/10.1063/5.0124890>

