

Deteksi Huruf pada Tulisan Tangan Latin Menggunakan You Only Look Once (YOLO)

Yoyon Kusnendar Suprpto¹, Aisyah Nurul Hidayah¹

¹Dept. Teknik Komputer Institut Teknologi Sepuluh Nopember Surabaya, Indonesia

Email: yoyonsuprpto@ee.its.ac.id, aisyah16@mhs.te.its.ac.id

Abstrak

Tulisan tangan hingga saat ini masih digunakan dalam berbagai bidang, khususnya tulisan tangan Latin. Namun, terdapat beberapa tantangan dalam pengenalan tulisan tangan Latin. Pertama, tulisan tangan memiliki banyak variasi. Variasi yang dimaksud meliputi ukuran huruf, kemiringan huruf, garis dasar atau baseline penulisan, serta perbedaan jarak antarhuruf dan antar kata [1]. Kedua, setiap huruf dalam tulisan tangan Latin saling terhubung. Tantangan ini menyebabkan waktu yang lebih lama dalam proses pengenalan tulisan tangan Latin. You Only Look Once (YOLO) digunakan untuk mendeteksi huruf pada tulisan tangan Latin sebagai upaya untuk mempercepat proses pengenalan tulisan tangan Latin. Eksperimen ini dilakukan dengan mengumpulkan citra sampel dari empat jenis tulisan tangan yang berbeda. Hasil eksperimen menggunakan GPU NVIDIA GeForce RTX 1050 menunjukkan bahwa waktu pemrosesan per citra untuk sistem deteksi pada citra satu kata adalah sebesar 0,0776 detik dengan menggunakan nilai ambang (threshold score) sebesar 0,3.

Keyword: *Latin handwriting, You only look once, yolo*

Diterima Redaksi: 3-11-2023 Selesai Revisi: 14-11-2023 Diterbitkan Online: 15-12-2023
DOI: <https://doi.org/10.59378/jcenim.v1i3.41>

I. LATAR BELAKANG

Tulisan tangan merupakan salah satu bentuk komunikasi. Dengan tulisan tangan, manusia dapat menyampaikan ide tanpa harus berada di tempat yang sama dengan pembaca. Dari sini, perkembangan tulisan tangan semakin beragam dan memiliki keunikan tersendiri pada setiap budaya di dunia.

Sejarah perkembangan tulisan sendiri dimulai dari masyarakat primitif, di mana mereka mulai membuat coretan-coretan pada dinding gua, kayu, maupun batu. Seiring berjalannya waktu, coretan-coretan tersebut berkembang menjadi simbol dan pada akhirnya menjadi alfabet atau huruf. Huruf-huruf alfabet tersebut kemudian menyebar ke berbagai negara sehingga memunculkan berbagai jenis alfabet yang berbeda, misalnya alfabet Yunani, alfabet Rusia, dan alfabet Latin. Secara khusus di Eropa, tulisan tangan Latin dikenal sebagai tulisan kursif, yang kemudian menyebar ke seluruh dunia.

Tulisan tangan hingga saat ini masih digunakan dalam berbagai bidang, salah satunya adalah layanan pengiriman. Menurut Badan Pusat Statistik (BPS), pengiriman surat oleh PT Pos Indonesia Divisi Regional Jawa Timur pada tahun 2015 mencapai 22.200.517 surat [2] dan paket sebanyak 627.922 [3]. Angka ini diperkirakan akan terus meningkat pada tahun-tahun berikutnya.

Tulisan tangan, khususnya tulisan tangan Latin, memiliki beberapa tantangan. Pertama, tulisan tangan memiliki variasi yang sangat besar. Variasi tersebut meliputi ukuran huruf, kemiringan huruf dan kata, garis dasar penulisan, tekanan tulisan, hingga jarak antarhuruf dan antar kata [1]. Kedua, setiap huruf dalam tulisan tangan Latin saling tersambung menjadi satu.

Kedua tantangan tersebut menyebabkan proses pengenalan tulisan tangan Latin menjadi lebih lama. Selain itu, Optical Character Recognition (OCR) yang ada saat ini belum mendukung pengenalan tulisan tangan Latin. Oleh karena itu, diperlukan suatu sistem yang dapat mendeteksi huruf pada tulisan tangan Latin sebagai upaya untuk mempercepat proses pengenalan. Tugas akhir ini bertujuan untuk membuat sistem pendeteksian huruf pada tulisan tangan Latin menggunakan You Only Look Once (YOLO). You Only Look Once (YOLO) digunakan karena mampu mendeteksi objek dengan cepat.

II. PENELITIAN TERKAIT

Beberapa penelitian terkait tulisan tangan dan pendeteksian objek menggunakan YOLO telah dilakukan oleh berbagai pihak. Penelitian pertama berjudul “Javanese Script Handwriting Recognition Rules Become Alphabet Letters On Paper Media Using The Convolutional Neural Network (CNN) Method With You Only Look Once (YOLO) Architecture” [4]. Penelitian ini menggunakan YOLO untuk mendeteksi tulisan tangan dari 5 judul dan penulis yang berbeda dalam aksara Jawa Mardikawi. Pengujian menggunakan 33 jenis karakter Jawa Mardikawi menghasilkan rata-rata Intersection over Union (IoU) sebesar 46%, sedangkan pengujian menggunakan 44 jenis karakter Jawa Mardikawi menghasilkan nilai IoU sebesar 40% dengan menggunakan dataset berukuran kertas asli.

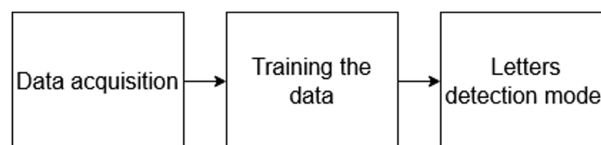
Penelitian kedua berjudul “TextCaps: Handwritten Character Recognition with Very Small Datasets” [5]. Penelitian ini membahas pengenalan tulisan tangan dengan menggunakan dataset yang relatif kecil. Sistem ini bermanfaat untuk pengenalan karakter pada bahasa daerah yang tidak memiliki banyak dataset. Dataset tersebut dilakukan augmentasi sehingga diperoleh lebih banyak variasi. Augmentasi yang dilakukan bersifat realistis dan mencerminkan variasi tulisan tangan manusia yang sebenarnya dengan menambahkan noise terkontrol secara acak. Pengenalan karakter menggunakan arsitektur yang terdiri dari capsule network dan network decoder, di mana hasil yang diperoleh melampaui hasil pengenalan karakter pada dataset EMNIST-letter.

Penelitian ketiga berjudul “Real-Time Face Detection based on YOLO” [6]. Penelitian ini membahas pendeteksian wajah menggunakan jaringan YOLOv3. Pengujian menggunakan dataset WIDER FACE menunjukkan bahwa pendeteksian wajah menjadi lebih cepat dengan rata-rata waktu deteksi sebesar 0,028 detik. Selain lebih cepat, hasil yang diperoleh juga lebih akurat. Kecepatan deteksi tersebut mampu memenuhi kebutuhan sistem real-time dan menghasilkan performa yang baik.

Penelitian keempat berjudul “Thermal Object Detection in Difficult Weather Conditions Using YOLO” [7] yang bertujuan untuk mendeteksi objek pada citra dan video termal yang diambil dalam kondisi cuaca yang berbeda (cuaca cerah, hujan lebat, dan kabut) selama proses pengawasan area yang dilindungi dengan menggunakan pendeteksi objek real-time. Pendeteksi objek yang digunakan dalam pengujian meliputi Faster RCNN, SSD, FCOS, Cascade RCNN, dan YOLOv3. Berdasarkan hasil pengujian, YOLOv3 memiliki kecepatan pemrosesan tertinggi, yaitu sebesar 27.472 FPS dengan nilai Average Precision (AP) sebesar 97,93%.

III. METODOLOGI

Tugas akhir ini merupakan penelitian di bidang visi komputer yang bertujuan untuk mendeteksi huruf pada tulisan tangan Latin menggunakan You Only Look Once (YOLO). Diagram alir pada Gambar 1 merupakan gambaran umum dari sistem yang akan dibuat.

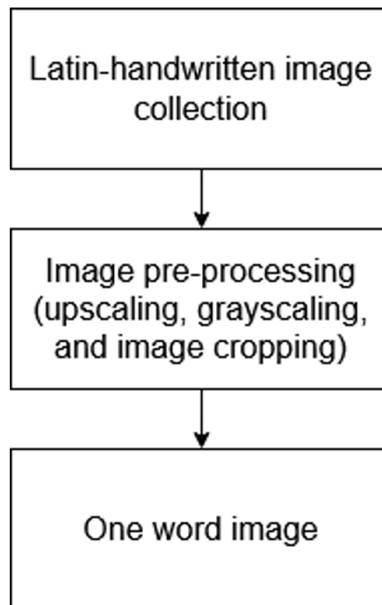


Gambar 1: Desain sistem secara keseluruhan

Tahap awal yang dilakukan adalah akuisisi data dengan mengumpulkan berbagai kata tulisan tangan Latin dari berbagai penulis pada selembar kertas. Selanjutnya, kertas tersebut difoto menggunakan kamera sehingga diperoleh citra kertas. Citra kertas yang diperoleh kemudian diproses dalam bentuk peningkatan ukuran dan kualitas citra (upscaling) serta konversi ke citra keabuan (grayscale). Citra yang berisi tulisan tangan Latin kemudian dipotong per kata sehingga satu citra hanya berisi satu kata. Citra kata tersebut kemudian dibagi menjadi data latih dan data uji serta dianotasi sesuai dengan huruf yang ada. Terdapat 26 jenis huruf yang digunakan, mulai dari A hingga Z. Setelah dataset siap, proses pelatihan dilakukan menggunakan Darkflow. Hasil pelatihan berupa sebuah model (bobot keluaran) yang terdiri dari file meta dan file pb untuk mendeteksi huruf pada data uji. Setiap hasil deteksi huruf kemudian digabungkan sehingga kata pada citra dapat dikenali. Akurasi model kemudian dihitung untuk mengevaluasi kinerja model.

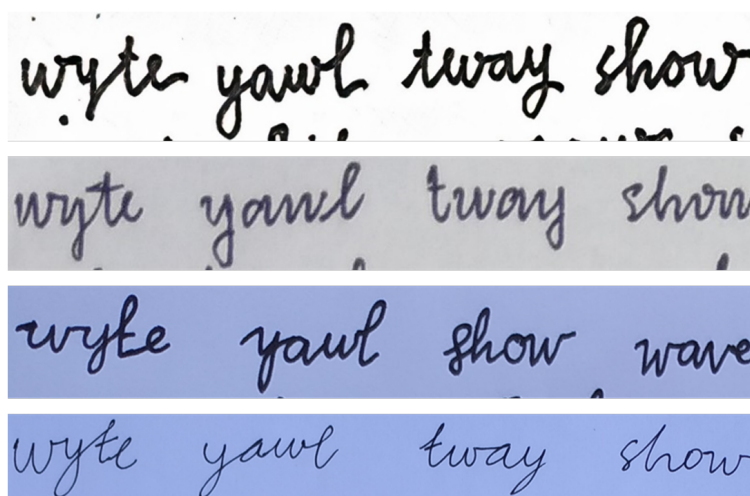
A. Akuisisi Data

Pada tugas akhir ini, akuisisi data dilakukan dengan memotret teks pada kertas menggunakan kamera. Pada Gambar 2 dijelaskan tahapan-tahapan dalam proses akuisisi data.



Gambar 2: Alur akuisisi data

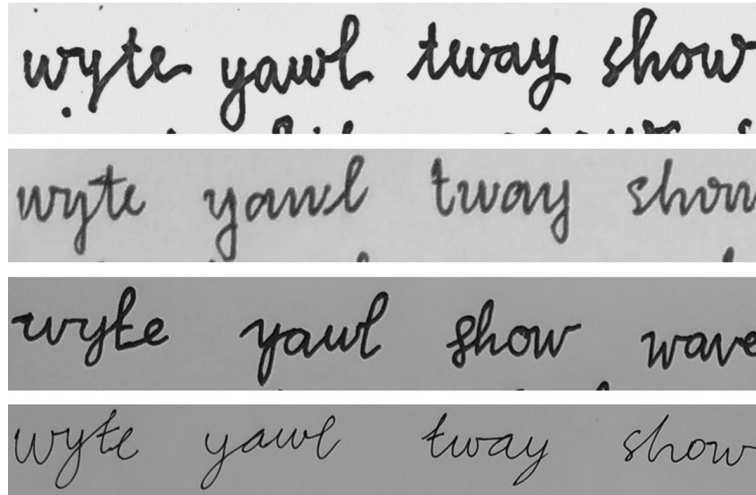
Proses pengumpulan tulisan tangan Latin diawali dengan menyiapkan kombinasi kata yang akan dituliskan. Kombinasi kata yang tepat mempengaruhi variasi posisi huruf sehingga setiap huruf dapat berada pada posisi awal, tengah, atau akhir kata. Setiap orang menuliskan kembali kombinasi kata yang digunakan sebagai dataset pada selembar kertas. Tulisan tangan Latin dikumpulkan dari beberapa orang untuk memperoleh beberapa lembar kertas dari penulis yang berbeda. Kertas yang berisi tulisan tangan Latin tersebut kemudian diakuisisi dengan cara difoto menggunakan kamera atau dipindai menggunakan mesin pemindai, sehingga diperoleh citra kertas. Hasil dari proses pengumpulan citra tulisan tangan ditunjukkan pada Gambar 3.



Gambar 3: Hasil proses pengumpulan citra tulisan tangan

Pra-pemrosesan citra terdiri dari tiga tahap, yaitu peningkatan kualitas dan ukuran citra (upscaling), konversi citra ke dalam bentuk keabuan (grayscale), serta pemotongan citra menjadi citra yang hanya berisi satu kata. Citra yang memiliki ukuran di bawah 2000 piksel akan diperbesar sehingga ukurannya menjadi di atas 4000 piksel. Setelah proses upscaling, citra kemudian dikonversi ke dalam bentuk keabuan. Tujuan dari proses grayscale ini adalah untuk menyeragamkan warna tinta yang digunakan pada

kertas sehingga fitur dan latar belakang dapat dibedakan. Proses upscaling dan grayscaling dilakukan menggunakan perangkat lunak pengolah foto seperti Adobe Photoshop. Hasil dari proses grayscaling citra ditunjukkan pada Gambar 4.



Gambar 4: Hasil citra setelah proses grayscaling

Citra tulisan tangan Latin kemudian dipotong sehingga satu citra hanya berisi satu kata. Proses pemotongan citra dilakukan menggunakan perangkat lunak Microsoft Photos. Citra kata tersebut kemudian dibagi menjadi dua bagian, yaitu data latih dan data uji dengan perbandingan masing-masing sebesar 70% dan 30%. Hasil pemotongan citra menjadi citra satu kata ditunjukkan pada Gambar 5.



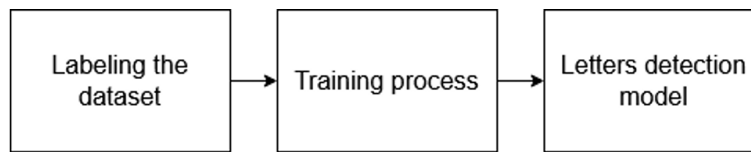
Gambar 5: Hasil pemotongan citra menjadi citra satu kata

B. Pelatihan Data

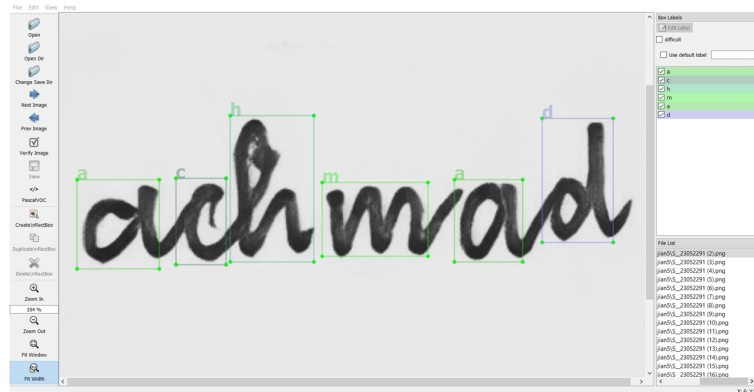
Setelah proses pengumpulan citra dan pra-pemrosesan citra selesai, tahap selanjutnya adalah pelatihan data. Terdapat beberapa tahapan dalam proses ini yang dijelaskan dalam bentuk diagram blok pada Gambar 6.

Citra pada dataset baik data latih maupun data uji kemudian diberikan anotasi atau pelabelan. Pelabelan merupakan proses pemberian label pada citra dengan menentukan bounding box pada objek serta nama kelasnya. Tujuan dari proses pelabelan ini adalah untuk memperoleh ground-truth bounding box. Proses pelabelan dilakukan menggunakan LabelImg Tools [8]. Gambar 7 menunjukkan proses pelabelan pada citra satu kata.

Hasil dari proses pelabelan berupa sebuah berkas yang berisi informasi koordinat bounding box objek pada citra. Berkas ini disimpan dalam bentuk .xml atau .txt sehingga setiap citra memiliki satu berkas



Gambar 6: Alur pelatihan data



Gambar 7: Proses pelabelan pada citra satu kata

anotasi. Dataset yang dibutuhkan kemudian disiapkan dari proses ini, yang terdiri dari citra dan anotasi citra tersebut. Hasil dari proses pelabelan ditunjukkan pada Gambar 8.

```

<object>
  <name>w</name>
  <pose>Unspecified</pose>
  <truncated>0</truncated>
  <difficult>0</difficult>
  <bndbox>
    <xmin>16</xmin>
    <ymin>126</ymin>
    <xmax>76</xmax>
    <ymax>194</ymax>
  </bndbox>
</object>
  
```

Gambar 8: Hasil dari proses pelabelan

Proses pelatihan membutuhkan konfigurasi layer dalam bentuk berkas cfg serta bobot awal (pre-trained weight). Selain itu, versi jaringan yang digunakan harus sama dengan versi bobot awal yang digunakan. Pada tugas akhir ini, versi YOLO yang digunakan adalah YOLOv1. Sama seperti pada proses pelabelan, berkas labels.txt pada Darkflow [9] diubah sesuai dengan kelas yang akan digunakan, yaitu huruf Latin dari a hingga z. Jaringan perlu dikonfigurasi terlebih dahulu sebelum memulai proses pelatihan. Konfigurasi jaringan disimpan dalam berkas .cfg. Beberapa parameter yang diatur dalam berkas cfg antara lain:

- 1) Mengubah jumlah kelas pada layer region sesuai dengan jumlah kelas yang akan dilatih, ukuran lebar dan tinggi (width & height), learning rate, serta jumlah filter pada layer konvolusional yang terletak pada dua layer terakhir, sesuai dengan Persamaan 1.

$$\text{filters} = \text{num} \times (\text{classes} + 5) \quad (1)$$

dengan num = 5 dan classes merupakan jumlah kelas yang digunakan.

- 2) Mengaktifkan baris batch dan subdivisions pada bagian training dengan menghapus tanda #.

Rincian pengaturan parameter jaringan untuk tugas akhir ini ditunjukkan pada Tabel 1.

Tabel 1: Rincian Pengaturan Parameter Jaringan

Parameter	Nilai
Lebar & Tinggi	416x416
Batch & Subdivision	64 & 32
Epoch	1000
Learning rate	0.001
Decay	0.0005
Momentum	0.9
Kelas	26
Filter	155

Proses pelatihan menggunakan framework Darkflow dan Google Colaboratory. Selama proses pelatihan, Darkflow akan menyimpan hasil sementara dalam bentuk Tensorflow checkpoint. Hasil akhir bobot pelatihan akan disimpan dalam berkas .meta dan .pb. Bobot akhir tersebut digunakan sebagai bobot pada model pendeteksian huruf. Pada proses pelatihan ini, hasil pelatihan dari checkpoint terakhir disimpan sebagai bobot akhir.

Selain itu, proses pembacaan tulisan tangan juga diperlukan. Pembacaan yang dibahas adalah pembacaan kata. Proses pembacaan berasal dari kelas yang muncul pada hasil prediksi. Kelas-kelas yang muncul tersebut kemudian dapat digabungkan sehingga dapat terbaca sebagai sebuah kata. Hasil pembacaan ditampilkan pada pojok kiri atas setiap citra. Algoritma pembacaan diawali dengan mengambil nilai xmin dan label dari setiap bounding box yang muncul setelah proses prediksi, kemudian menyimpannya ke dalam sebuah array dua dimensi. Isi array tersebut kemudian diurutkan berdasarkan nilai xmin untuk setiap bounding box. Nilai xmin diurutkan dari yang terkecil hingga terbesar. Selanjutnya, label akan dikembalikan sesuai dengan urutan pasangan nilai xmin tersebut. Algoritma pembacaan ini diilustrasikan secara jelas dalam diagram alir seperti pada Gambar 9.

IV. PENGUJIAN DAN ANALISIS

Pada tugas akhir ini, terdapat tiga pengujian model pendeteksian yang dibagi menjadi tiga bagian sebagai berikut:

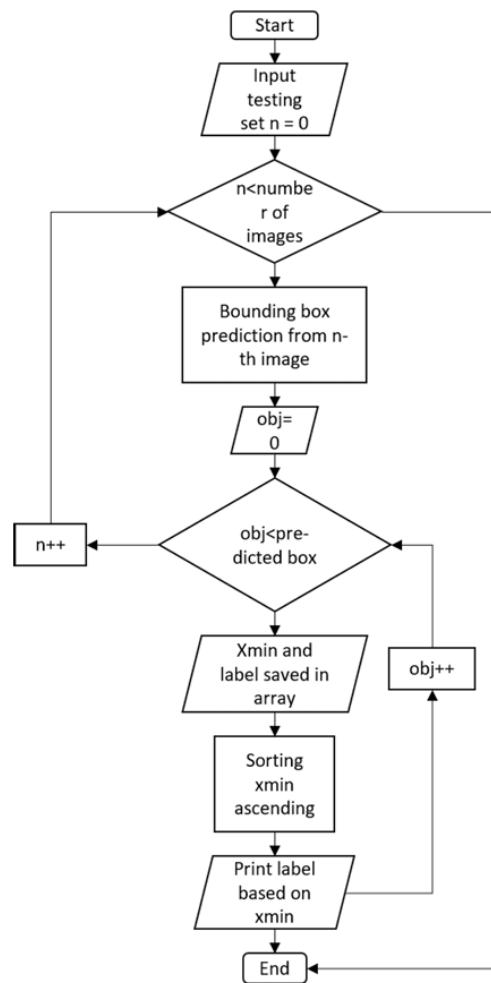
- 1) Pengujian pendeteksian huruf dengan variasi ukuran tulisan tangan
- 2) Pengujian pendeteksian huruf pada citra satu kata
- 3) Pembacaan tulisan tangan pada citra satu kata

A. Pengujian Pendeteksian Huruf dengan Variasi Ukuran Tulisan Tangan

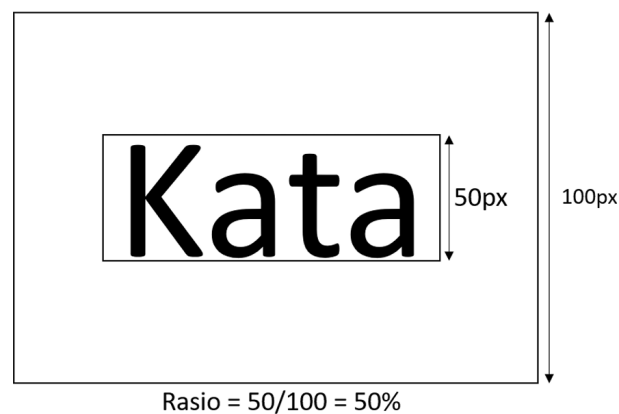
Pengujian ini dilakukan untuk mengetahui pengaruh variasi ukuran tulisan terhadap hasil pendeteksian huruf. Ukuran tulisan tangan merupakan perbandingan antara tinggi tulisan terhadap tinggi keseluruhan citra dalam satuan piksel. Ilustrasi ukuran tulisan tangan ditunjukkan pada Gambar 10. Pada pengujian ini terdapat tiga percobaan pelatihan dengan ukuran tulisan tangan yang berbeda. Eksperimen 1 merupakan percobaan pelatihan menggunakan dataset dengan rasio 50%. Eksperimen 2 merupakan percobaan pelatihan menggunakan dataset dengan rasio 20%. Eksperimen 3 merupakan percobaan pelatihan menggunakan dataset gabungan antara percobaan pelatihan 1 dan 2. Hasil dari ketiga eksperimen tersebut berupa model yang diuji pada dataset dengan berbagai ukuran tulisan tangan, yaitu dengan rasio 50%, 40%, 20%, dan 10%.

Pada eksperimen 1, model tidak dapat mendeteksi tulisan dengan rasio 10%. Untuk ukuran tulisan dengan rasio 40% dan 20%, model cukup kesulitan dalam mendeteksi huruf. Meskipun demikian, bounding box yang dihasilkan cukup presisi dalam mendeteksi huruf, khususnya pada rasio 50%. Hasil dari eksperimen 1 ditunjukkan pada Gambar 11.

Pada eksperimen 2, terlihat bahwa model dapat mendeteksi huruf pada seluruh variasi ukuran tulisan. Namun, hasil pendeteksian pada tulisan dengan rasio 50% tidak sebaik pada tulisan dengan rasio 40%.



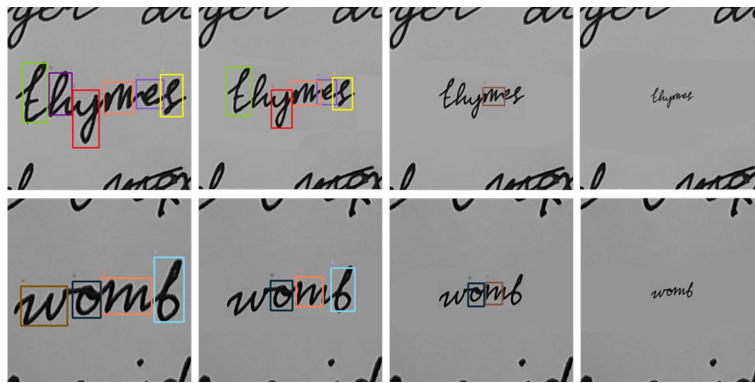
Gambar 9: Diagram alir pembacaan tulisan tangan Latin



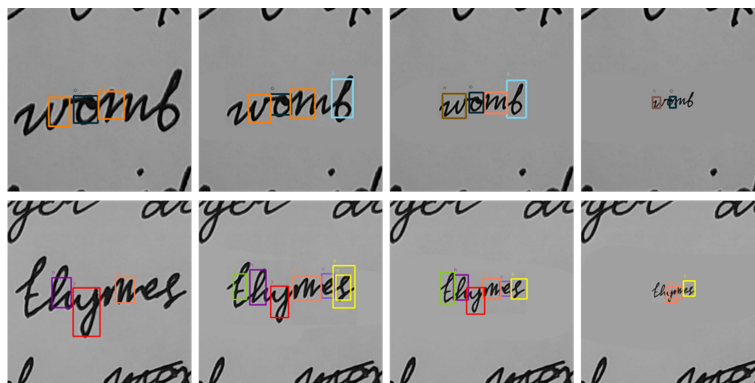
Gambar 10: Ilustrasi ukuran tulisan tangan

Bounding box yang presisi hanya dihasilkan pada tulisan dengan rasio 20%. Hasil dari eksperimen 2 ditunjukkan pada Gambar 12.

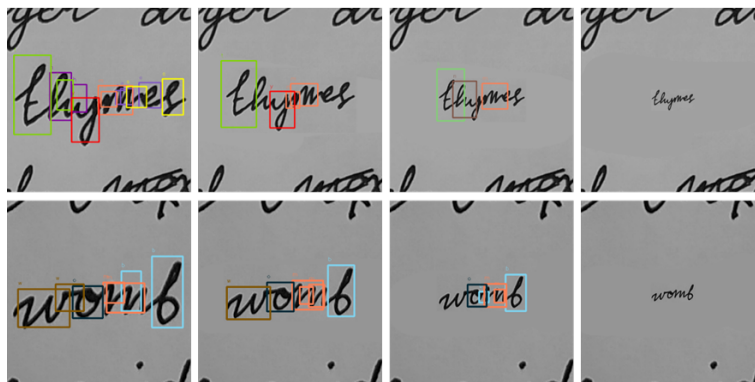
Pada eksperimen 3, terlihat bahwa model tidak dapat mendeteksi huruf pada rasio 10%. Selain itu, banyak bounding box yang tidak akurat atau merupakan false positive. Hal ini terjadi karena model menghasilkan cukup banyak prediksi yang salah. Meskipun demikian, masih terdapat cukup banyak bounding box yang presisi. Hasil dari eksperimen 3 ditunjukkan pada Gambar 13.



Gambar 11: Hasil eksperimen 1



Gambar 12: Hasil eksperimen 2



Gambar 13: Hasil eksperimen 3

B. Pengujian Pendeteksian Huruf pada Citra Satu Kata

Model yang telah dibuat harus mampu mendeteksi huruf pada citra tulisan tangan. Pengujian ini bertujuan untuk mengetahui hasil pendeteksian huruf dan waktu pemrosesan pada citra satu kata untuk setiap nilai ambang kepercayaan (threshold confidence score), sehingga dapat diketahui nilai ambang kepercayaan yang optimal. Proses pendeteksian dilakukan pada data uji dengan total 56 kata atau 1268 huruf yang ditulis oleh empat penulis yang berbeda. Hasil pendeteksian ditunjukkan pada Gambar 14.

Berdasarkan Gambar 14, terdapat sepuluh huruf yang dituliskan oleh penulis. Model berhasil mendeteksi huruf pada citra tersebut ketika nilai ambang kepercayaan adalah 0.3. Hal ini disebabkan oleh warna tulisan yang sangat kontras dengan latar belakang sehingga huruf dapat terdeteksi dengan baik. Tabel 2 menunjukkan bahwa ketika nilai ambang kepercayaan sama dengan 1, tidak ada kotak prediksi yang muncul. Hal ini terjadi karena tidak ada kotak prediksi yang memiliki nilai kepercayaan sama dengan 1. Nilai ambang akan menghilangkan kotak prediksi jika nilai kepercayaannya lebih ren-



Gambar 14: Hasil pendeteksian huruf pada citra satu kata (threshold 0.3)

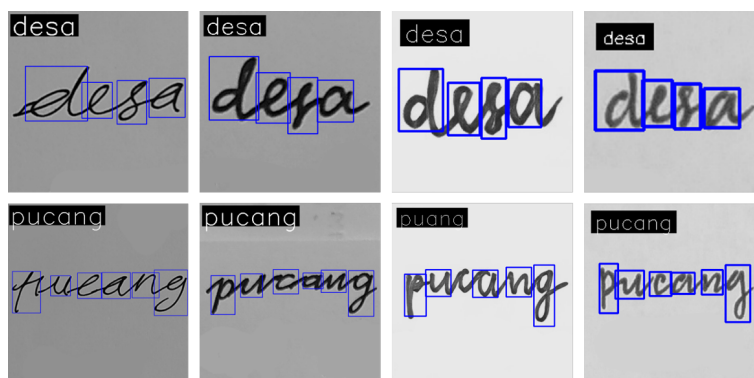
dah dari nilai ambang yang ditentukan. Hasil prediksi kemudian dianalisis untuk mengevaluasi kinerja model. Analisis yang digunakan adalah analisis confusion matrix dan mean average precision, seperti ditunjukkan pada Tabel 2.

Tabel 2: Hasil Pendeteksian dan Pembacaan Huruf pada Citra Satu Kata

Threshold	$\hat{B}$	Waktu pemrosesan	Confusion Matrix				Performa		Kata benar
			TP	FP	FN	Precision	Recall	mAP	
0.1	1463	0.075	1129	334	139	0.70	0.81	84.70	76
0.2	1315	0.080	1107	208	161	0.76	0.79	83.16	110
0.3	1241	0.078	1089	152	179	0.77	0.78	81.93	114
0.4	1167	0.076	1053	114	215	0.79	0.75	79.14	96
0.5	1074	0.077	999	75	269	0.82	0.71	74.68	79
0.6	933	0.075	886	47	382	0.80	0.62	65.81	46
0.7	660	0.077	634	26	634	0.77	0.43	45.97	9
0.8	244	0.075	240	4	1028	0.77	0.15	15.95	0
0.9	33	0.074	33	0	1235	0.27	0.02	2.44	0
1.0	0	0.074	0	0	1268	0.00	0.00	0.00	0

C. Pembacaan Tulisan Tangan pada Citra Satu Kata

Pengujian ini bertujuan untuk mengetahui hasil pembacaan citra. Pembacaan tulisan tangan dilakukan berdasarkan hasil prediksi yang diperoleh sebelumnya. Pengujian ini menggunakan 56 kata yang ditulis oleh empat penulis yang berbeda. Hasil pengujian sistem pembacaan pada citra satu kata ditunjukkan pada Tabel 2 dan Gambar 15.



Gambar 15: Hasil pembacaan tulisan tangan pada citra satu kata (threshold 0.3)

Jumlah teks yang terbaca paling banyak terjadi ketika nilai ambang kepercayaan adalah 0.3, yaitu sebanyak 114 kata terbaca dengan benar dari total 224 kata. Ketika nilai ambang dinaikkan, jumlah

kata yang terbaca dengan benar mengalami penurunan. Hal ini disebabkan oleh jumlah bounding box yang muncul semakin sedikit.

V. KESIMPULAN

Berdasarkan hasil pengujian yang telah dilakukan, dapat ditarik beberapa kesimpulan sebagai berikut:

- 1) Berdasarkan hasil pengujian pendeteksian huruf dengan variasi ukuran tulisan, model pelatihan yang paling sesuai menggunakan dataset dengan rasio tinggi tulisan sebesar 50%.
- 2) Berdasarkan hasil pengujian pendeteksian huruf dan pembacaan pada citra satu kata, nilai ambang kepercayaan (threshold confidence score) yang ideal untuk digunakan adalah 0.3, dengan waktu pemrosesan per citra sebesar 0.0776 detik, nilai precision sebesar 76.92%, nilai recall sebesar 78.27%, serta nilai mean average precision (mAP) sebesar 81.93%.
- 3) Sistem pembacaan tulisan tangan telah mampu membaca huruf hasil pendeteksian secara berurutan dari kiri ke kanan.

Daftar Pustaka

- [1] J. Ryu and H. I. Koo, "Word segmentation method for handwritten documents based on structured learning," *IEEE Signal Processing Letters*, vol. 22, no. 8, pp. 1161–1165, 2015.
- [2] "Surat pos dalam dan luar negeri yang dikirim dan diterima, 2015 (update terakhir 18 Oct 2017)," Badan Pusat Statistik Provinsi Jawa Timur. [Online]. Tersedia: <https://jatim.bps.go.id/statictable/2017/10/18/701/surat-pos-dalam-dan-luar-negeri-yang-dikirim-dan-diterima-di-provinsi-jawa-timur-2015.html>. Diakses pada 16 April 2020.
- [3] "Paket pos dalam dan luar negeri yang dikirim dan diterima, 2015 (update terakhir 18 Oct 2017)," Badan Pusat Statistik Provinsi Jawa Timur. [Online]. Tersedia: <https://jatim.bps.go.id/statictable/2017/10/18/700/paket-pos-dalam-dan-luar-negeri-yang-dikirim-dan-diterima-di-provinsi-jawa-timur-2015.html>. Diakses pada 16 April 2020.
- [4] L. E. Damayanti, "Pengenalan tulisan tangan naskah aksara Jawa aturan penulisan naskah Kawi menjadi huruf alfabet pada media kertas menggunakan metode Convolutional Neural Network (CNN) arsitektur You Only Look Once," 2020.
- [5] V. Jayasundara, S. Jayasekara, H. Jayasekara, J. Rajasegaran, S. Seneviratne, and R. Rodrigo, "TextCaps: Handwritten character recognition with very small datasets," *arXiv preprint arXiv:1904.08095*, 2019.
- [6] W. Yang and Z. Jiachu, "Real-time face detection based on YOLO," in *Proceedings of the 2018 1st International Conference on Control and Communication Engineering*, 2018, pp. 221–224.
- [7] M. Kristo, M. Ivasic-Kos, and M. Pobar, "Thermal object detection in difficult weather conditions using YOLO," *IEEE Access*, vol. 8, pp. 125459–125476, 2020.
- [8] Tzutalin, "LabelImg," GitHub repository. [Online]. Tersedia: <https://github.com/tzutalin/labelImg>, 2015. Diakses pada 4 April 2020.
- [9] T. H. Trieu, "Darkflow," GitHub repository. [Online]. Tersedia: <https://github.com/thtrieu/darkflow>, 2018. Diakses pada 18 November 2019.