

Reduksi Duplikasi Bounding Box Deteksi Wajah Menggunakan Algoritma Greedy dan Intersection over Union

Julius Arthur - 13523030
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jalan Ganesha 10 Bandung
E-mail: juliusarthur2005@gmail.com , 13523030@std.stei.itb.ac.id

Abstract— Deteksi wajah merupakan teknologi penting dalam sistem keamanan, pengawasan, hingga presensi otomatis. Salah satu tantangan dalam deteksi wajah adalah munculnya beberapa bounding box yang merujuk pada satu wajah yang sama, seperti saat menggunakan model deteksi YOLOv5. Makalah ini mengusulkan metode penyaringan bounding box menggunakan algoritma greedy yang dikombinasikan dengan perhitungan Intersection over Union (IoU) untuk mengatasi masalah duplikasi tersebut. Algoritma penyaringan ini akan memilih bounding box dengan confidence tertinggi secara berurutan dan menghapus box lain yang memiliki tingkat tumpang tindih (IoU) melebihi ambang batas tertentu. Pengujian dilakukan pada berbagai citra dengan kualitas rendah dan pencahayaan tidak merata. Dengan efisiensi dan kesederhanaannya, pendekatan ini layak diterapkan pada sistem real-time berbasis deteksi wajah.

Keywords: Deteksi wajah, algoritma greedy, Intersection over Union, bounding box.

I. INTRODUCTION (HEADING 1)

Dalam perkembangan teknologi yang pesat, pengenalan wajah menjadi sebuah teknologi yang marak digunakan terutama pada bidang pengawasan, keamanan, presensi, hingga otorisasi akses. Teknologi ini secara otomatis mampu tidak hanya mampu mendeteksi wajah manusia dari citra langsung, tetapi juga mengidentifikasi wajah tersebut dengan efisien dan keakuratan yang tinggi. Kemampuan ini membuat tingginya penggunaan deteksi wajah sebagai sarana keamanan.

Salah satu metode deteksi yang digunakan adalah YOLO[2][3], yang mampu mendeteksi berbagai objek termasuk wajah. Salah satu keluaran dari algoritma ini merupakan sebuah area yang merupakan objek hasil deteksi, disebut sebagai *bounding box*. Namun, dalam berbagai situasi, model deteksi wajah dapat menghasilkan lebih dari satu bounding untuk satu wajah yang sama. Hal ini dapat dipengaruhi oleh pencahayaan, posisi wajah, orientasi wajah, hingga kualitas gambar yang rendah.

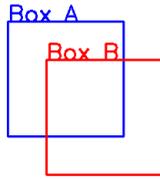
Keluaran yang memiliki beberapa *bounding box* untuk objek yang sama tentu berdampak negative pada keakuratan

metode dan algoritma deteksi. Selain itu, tergantung pada penggunaannya, sebuah program dapat mencatat satu buah objek/ wajah lebih dari sekali, sehingga meningkatkan komputasi, penggunaan penyimpanan, hingga kesalahan identifikasi. Oleh karena itu, perlu dilakukan pengolahan lebih lanjut dari hasil dari deteksi wajah, sehingga didapat hanya satu *bounding box* terbaik untuk satu buah wajah. Salah satu pendekatan yang penulis gunakan adalah algoritma greedy berdasarkan perhitungan Intersection over Union (IoU). Pendekatan membandingkan seluruh *bounding box* berdasarkan banyaknya area yang tumpang tindih dengan *bounding box* lain. *Bounding box* yang memiliki Tingkat tumpang tindih dengan *bounding box* lain, lebih tinggi dari nilai threshold, akan dihapus. Melalui pendekatan sederhana ini, diharapkan sebuah keluaran deteksi wajah yang lebih akurat, terutama pada gambar dengan pencahayaan yang tidak merata, serta posisi serta orientasi yang beragam.

Algoritma Greedy merupakan sebuah algoritma yang selalu mengambil langsung solusi (lokal) optimal yang ada[5]. Algoritma ini selalu mengambil tindakan langsung, tanpa bisa melakukan langkah mundur (backtracking) dan hanya memperhatikan solusi lokal saja. Oleh karena itu, algoritma greedy dapat menghasilkan solusi yang tidak optimal dalam berbagai masalah. Dalam masalah penyaringan *bounding box* hasil deteksi wajah, algoritma greedy akan memproses seluruh *bounding box* yang ada, diurutkan dari tingkat *confidence*-nya, mengambil keputusan (membuang atau menyimpan *bounding box*), tanpa bisa kembali ke kondisi sebelumnya (backtrack). Dengan metode ini, proses penyaringan dilakukan dengan waktu dan sumber daya yang efisien. Kecepatan dan kesederhanaan komputasi algoritma greedy memungkinkannya untuk diterapkan pada sistem langsung (*real time*), seperti kamera keamanan.

Dalam deteksi objek, Intersection over Union merupakan sebuah metrik yang digunakan untuk mengevaluasi tingginya tingkat tumpang tindih dua buah *bounding box*[4]. Nilai ini merupakan hasil rasio antara luas area irisan (*intersection*) dan luas area gabungan (*union*) dari dua buah *bounding box*.

IoU: 0.2882



Nilai IoU dengan langsung memberikan tingkat tumpang tindih dua buah *bounding box*. Semakin tinggi nilai IoU tersebut, semakin tinggi pula kemungkinan kedua *bounding box* menunjuk objek yang sama. Sebaliknya, semakin rendah nilai IoU, semakin tinggi kemungkinan kedua *bounding box* merujuk pada dua objek yang berbeda/ terpisah.

Dalam pendekatan yang penulis gunakan, setiap *bounding box* akan diputuskan untuk diambil atau tidak secara *greedy* melalui nilai IoU nya dengan *bounding box* lain yang sudah dipilih sebelumnya. Diperlukan sebuah nilai *threshold* IoU yang menuntukan kedua *bounding box* dianggap merujuk pada objek yang sama atau berbeda. Proses ini diulangi pada setiap *bounding box* yang ada.

Kombinasi antara algoritma *greedy* dan perhitungan IoU ini akan diuji untuk menyaring keluaran deteksi wajah berbasis YOLOV5. Diharapkan, pendekatan ini mampu meningkatkan keakuratan deteksi wajah, yang kemudian dapat mengurangi kompleksitas komputasi yang berdasarkan keluaran deteksi wajah tersebut. Dengan begitu, akan didapatkan hasil deteksi yang lebih akurat, bersih, terutama pada pencahayaan yang tidak merata dan *noise* yang tinggi

bounding box yang merujuk pada kepala atau tubuh, akan diinterpretasikan sebagai wajah pada pengujian makalah ini.

C. Seleksi Bounding Bounding Box dengan Algoritma Greedy

Seluruh *bounding box* hasil deteksi akan disortir berdasarkan nilai *confidence* secara menurun. *Bounding box* yang terpilih akan dibandingkan dengan seluruh *bounding box* lain yang sudah terpilih sebelumnya. Pada setiap tahap, akan ditentukan bagi *bounding box* terpilih untuk dimasukkan ke larik *bounding box* terpilih atau tidak. Keputusan ini diambil berdasarkan besarnya area yang saling tumpang tindih.

D. Penyaringan Bounding Box dengan Intersection over Union

Setiap *bounding box* yang dievaluasi akan dihitung nilai IoU-nya dengan seluruh *bounding box* lain yang telah dipilih sebelumnya. Jika nilai IoU ini melebihi sebuah ambang batas, maka *bounding box* tersebut dianggap merujuk pada objek yang sama dengan objek yang telah dirujuk sebelumnya pada larik *bounding box* terpilih. Pada penelitian ini, nilai ambang batas (*threshold*) IoU ditetapkan sebesar 0.5. Nilai ini ditujukan untuk menyeimbangkan sensitivitas terhadap duplikasi dengan toleransi terhadap *bounding box* yang berdekatan. Akibatnya, *bounding box* tersebut tidak akan dimasukkan ke larik *bounding box* terpilih. Sebaliknya, jika nilai IoU dibawah nilai ambang batas, maka *bounding box* yang sedang diproses dianggap merujuk ke objek yang berbeda. Proses perhitungan IoU ini dilakukan pada *bounding box* terpilih pada setiap *bounding box* di dalam larik. Proses ini memastikan bahwa setiap *bounding box* yang ada pada larik merujuk pada objek unik, menghindari adanya *bounding box* duplikat pada sebuah wajah.

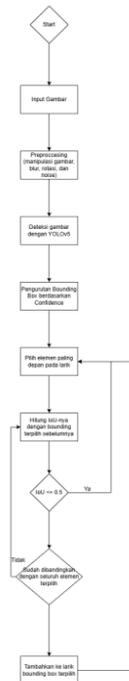
II. METODOLOGI

A. Akusisi Gambar

Gambar yang akan diproses dapat diperoleh dari berbagai sumber, termasuk video dan gambar. Gambar- gambar yang dipilih pada pengujian ini merupakan gambar yang memicu adanya deteksi ganda pada model deteksi wajah YOLOV5. Kriteria umum dari gambar- gambar ini adalah gambar dengan pencahayaan yang tidak merata, orientasi wajah yang miring, posisi wajah terpotong, serta *noise* yang tinggi. Selain itu, penulis juga akan melakukan manipulasi gambar seperti blur, rotasi, hingga *noise*, untuk meningkatkan kemungkinan terjadinya *bounding box* duplikat pada sebuah wajah. Hal ini juga pendekatan yang dapat mensimulasikan kondisi visual pada dunia nyata, seperti pada sistem keamanan.

B. Deteksi Wajah

Langkah selanjutnya merupakan proses deteksi wajah. Proses ini menggunakan algoritma YOLOv5. Model ini merupakan model yang marak digunakan dalam program deteksi objek. Selain itu, YOLOv5 mampu melakukan deteksi objek secara tepat dan akurat. Keluaran dari model ini merupakan kumpulan *bounding box* yang disertai dengan nilai *confidence*. Model ini tidak dilatih hanya untuk melakukan deteksi wajah, tetapi "orang" secara keseluruhan, dari dataset COCO (Common Object in Context). Oleh karena itu,



Gambar 1 Diagram alir program penyaringan *bounding box*

III. HASIL DAN ANALISIS

Pengujian pada makalah ini menggunakan beberapa nilai threshold IoU. Model YOLOv5, yang digunakan akan menghasilkan *bounding box* yang memiliki nilai IoU hingga 0.9 dengan *bounding box* lainnya. Hal ini dipilih untuk meningkatkan kemungkinan terjadinya duplikasi *bounding box* untuk wajah yang sama.

A. Percobaan 1

IoU threshold: 0.5



Gambar 2.1.1 Hasil deteksi wajah dari model YoloV5.



Gambar 2.1.2 Hasil penyaringan *bounding box*.

Waktu penyaringan: 0.10ms

B. Percobaan 2

IoU threshold: 0.5



Gambar 2.2.1 Hasil deteksi wajah dari model YoloV5.



Gambar 2.2.2 Hasil penyaringan *bounding box*.

Waktu penyaringan: 0.44ms

C. Percobaan 3

IoU threshold: 0.5



Gambar 2.3.1 Hasil deteksi wajah dari model YoloV5.

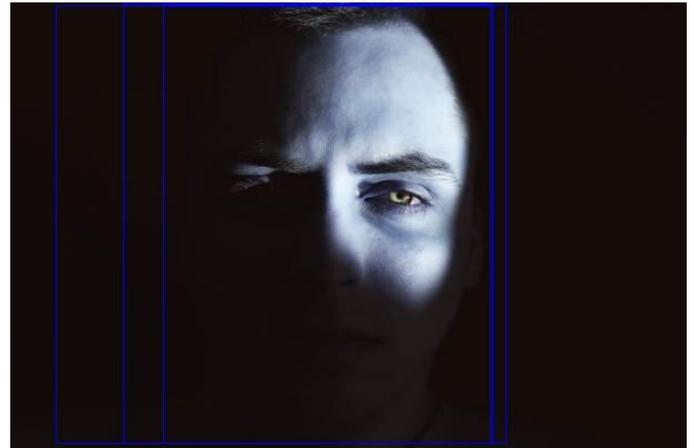


Gambar 2.3.2 Hasil penyaringan *bounding box*.

Waktu penyaringan: 0.09ms

E. Percobaan 5

IoU threshold: 0.75



Gambar 2.5.1 Hasil deteksi wajah dari model YoloV5.

D. Percobaan 4

IoU threshold: 0.5



Gambar 2.4.1 Hasil deteksi wajah dari model YoloV5.



Gambar 2.5.2 Hasil penyaringan *bounding box*.

Waktu penyaringan: 0.1ms



Gambar 2.4.2 Hasil penyaringan *bounding box*.

Waktu penyaringan : 0.88ms

F. Percobaan 6

IoU threshold: 0.75



Gambar 2.6.1 Hasil deteksi wajah dari model YoloV5.



Gambar 2.6.7 Hasil penyaringan *bounding box*.

Waktu penyaringan: 0.59ms.

G. Analisis Percobaan

Pada percobaan 1 hingga percobaan 6, gambar yang dipilih memiliki pencahayaan yang rendah ataupun kualitas yang cenderung rendah pula.

Pada percobaan 1, dapat terlihat jelas bahwa keluaran mentah deteksi wajah menghasilkan 3 buah *bounding box* pada satu wajah. Setelah dilakukan penyaringan, algoritma mampu mereduksinya menjadi satu buah *bounding box* saja. Algoritma pada pengujian mampu meningkatkan akurasi deteksi dengan baik dan benar.

Pada percobaan 2, algoritma mampu mereduksi *bounding box* hasil deteksi gambar. Algoritma penyaringan mampu mereduksi duplikasi *bounding box* hingga 40%. Penyaringan juga mampu dilakukan dengan benar dan tepat.

Pada percobaan 3, gambar yang digunakan merupakan gambar dengan pencahayaan yang rendah dan kualitas yang rendah pula. Keluaran model deteksi memiliki *bounding box* duplikat pada objek paling kiri. Algoritma penyaringan mampu mengurangi duplikasi ini dengan baik dan benar, sehingga hanya terdapat dua *bounding box* untuk dua objek.

Pada percobaan 4, gambar yang digunakan merupakan gambar dengan kualitas rendah dan terdapat banyak objek bertumpukan yang dapat terdeteksi model deteksi. Dari hasil model deteksi, dapat dilihat bahwa terdapat beberapa objek yang memiliki *bounding box* lebih dari satu. Algoritma penyaringan mampu menghilangkan duplikasi ini dengan tepat.

Pada percobaan 5, digunakan nilai ambang batas IoU sebesar 0.75. Hal ini berdampak pada keoptimalan algoritma penyaringan. Pada hasil deteksi, terdapat tiga *bounding box* untuk satu objek. Namun, algoritma penyaringan gagal menghilangkan seluruh duplikasi, dengan masih menunjukkan dua *bounding box* untuk satu objek.

Pada percobaan 6, masih digunakan nilai ambang batas IoU sebesar 0.75. Dapat dilihat, algoritma penyaringan masih gagal menghasilkan solusi yang optimal. Masih terdapat sebuah objek yang memiliki duplikasi *bounding box*.

Berikut adalah perbandingan jumlah *bounding box* sebelum dan setelah dilakukan penyaringan serta persentase reduksinya.

Percobaan	Jumlah bounding box sebelum penyaringan	Jumlah bounding box setelah penyaringan	Persentase reduksi
1	3	1	66%
2	20	12	40%
3	3	2	33%
4	30	16	46%
5	3	2	33%
6	21	12	42%

Tabel 1 tabel perbandingan jumlah *bounding box* pada setiap percobaan, sebelum dan setelah penyaringan.

Dari percobaan, didapatkan bahwa, secara rata-rata, algoritma penyaringan menggunakan algoritma greedy dan perhitungan Intersection over Union mampu mereduksi jumlah *bounding box* sebesar 43.8%.

Algoritma greedy memiliki karakteristik untuk mengambil keputusan lokal setiap langkahnya, tanpa mempertimbangkan keputusan yang akan datang yang mungkin menghasilkan solusi global yang lebih baik. Dalam penyaringan *bounding box*, perbandingan dilakukan satu per satu berdasarkan Intersection over Union. Dengan strategi ini, program akan secara langsung membuang *bounding box* yang tumpang tindih dengan *bounding box* terpilih sebelumnya tanpa menggunakannya sebagai kombinasi dengan *bounding box* lainnya yang mungkin mendapatkan hasil lebih optimal.

Secara keuntungan, pendekatan ini memiliki efisiensi komputasi yang cukup tinggi. Dari percobaan yang ada, kalkulasi penyaringan hanya membutuhkan waktu 0.09 ms hingga 0.88 ms. Algoritma ini terbukti mampu melakukan reduksi duplikasi *bounding box* dengan cepat dan ringan. Oleh karena itu, algoritma penyaringan ini cocok digunakan pada sistem real time, seperti kamera keamanan. Meski demikian, algoritma ini akan butuh waktu yang lebih banyak untuk citra yang memiliki objek terpisah yang lebih banyak. Setiap *bounding box* akan dibandingkan dengan seluruh *bounding box* terpilih sebelumnya, tanpa ada pembuangan lebih awal. Maka, algoritma greedy pada kasus ini memiliki kompleksitas waktu $O(nk)$ dengan k merupakan jumlah *bounding box* yang disimpan.

Namun, karakteristik algoritma greedy juga membawa kelemahan. Algoritma ini hanya mencari solusi lokal, tanpa menjamin solusi global. Salah satu kasus kesalahan yang muncul adalah pemilihan *bounding box* yang kurang ideal (bukan *bounding box* yang mengelilingi objek paling benar secara spasial dan posisi). Saat menemukan *bounding box* yang tidak tumpang tindih, algoritma ini langsung menyimpannya pada larik hasil. Jika selanjutnya ditemukan *bounding box* lain pada objek yang sama dengan posisi yang lebih baik, *bounding box* ini akan tetap dibuang. Pengambilan dan pembuangan *bounding box* selalu bersifat final pada tiap tahap, sehingga tidak bisa dilakukan kombinasi dari *bounding box* yang saling tumpang tindih, untuk menghasilkan solusi optimal global. Akibatnya, algoritma penyaringan dapat menghasilkan

bounding box yang bukan merupakan *bounding box* optimal untuk tiap objeknya.

Selain karakteristik algoritma greedy, nilai ambang batas IoU juga menjadi bagian penting dalam menentukan solusi yang optimal. Dapat dilihat dari percobaan 1 dan percobaan 5, nilai IoU yang terlalu tinggi dapat menghasilkan solusi yang kurang optimal. Algoritma greedy gagal untuk membuang *bounding box* duplikat karena dianggap sebagai *bounding box* yang merujuk objek yang berbeda. Namun, ambang batas IoU yang terlalu rendah juga dapat menghasilkan solusi yang kurang optimal. Dengan ambang batas IoU yang rendah, objek berbeda yang berdekatan akan dianggap sebagai objek yang sama dan dibuang oleh algoritma greedy.

Meski demikian, pada banyak percobaan, algoritma greedy mampu menunjukkan kinerja yang efektif dan berhasil menyaring duplikasi dengan akurat dan presisi.

IV. KESIMPULAN

Percobaan yang penulis lakukan mengusulkan sebuah metode penyaringan *bounding box* dari deteksi wajah menggunakan algoritma greedy berdasarkan perhitungan Intersection over Union (IoU). Metode ini bertujuan untuk mengatasi masalah duplikasi *bounding box* yang dapat dihasilkan keluaran model deteksi YOLOv5, terutama pada gambar dengan pencahayaan tidak merata atau rendah, orientasi wajah yang tidak lurus terhadap kamera, kualitas gambar rendah, hingga *noise* yang tinggi. Strategi ini akan mempertahankan hanya satu *bounding box* saja untuk setiap wajah yang terdeteksi, sehingga mengurangi kesalahan deteksi.

Dari percobaan yang telah dilakukan, pendekatan ini mampu mereduksi duplikasi *bounding box* hingga, rata-rata, 43.8%. Selain itu, proses penyaringan pada setiap percobaan dilakukan hanya dibawah 1 detik saja. Hasil ini menunjukkan bahwa algoritma greedy mampu mengolah berbagai data secara efektif dan efisien. Algoritma greedy mampu menyaring *bounding box* yang tumpang tindih secara tepat dan efisien, berdasarkan nilai IoU. Namun, algoritma greedy tidak selalu menghasilkan solusi optimal secara global. Dengan hanya memperhatikan nilai confidence, *bounding box* yang lebih tepat secara spasial dapat dibuang, tanpa kemungkinan masuk ke kombinasi larik hasil sama sekali dan tidak diperhitungkan sama sekali.

Kombinasi algoritma greedy dan IoU terbukti mampu secara ringan meningkatkan kualitas deteksi wajah dari model YOLOv5. Metode ini dapat diterapkan pada sistem real time seperti kamera pengawasan hingga sistem presensi berbasis wajah tanpa perlu melakukan modifikasi kode model deteksi wajah.

V. SARAN PENGEMBANGAN

Saat ini, pengurutan *bounding box* hanya berdasarkan nilai *confidence*. Untuk meningkatkan keoptimalan solusi, perlu

ditambahkan heuristik spasial yang menganalisis posisi terbaik dari beberapa *bounding box* yang ada pada sebuah objek. Dengan begitu, *bounding box* yang dihasilkan merupakan *bounding box* yang merepresentasikan objek lebih baik lagi.

Selain itu, perlu dilakukan percobaan lebih lanjut untuk menentukan nilai ambang batas IoU yang terbaik, berdasarkan kualitas citra dan jumlah objek yang terdeteksi untuk mengurangi kegagalan penyaringan. Penggunaan model deteksi yang lebih akurat juga dapat menghasilkan *bounding box* yang lebih akurat. Pengembangan juga perlu dilakukan untuk mengevaluasi performa algoritma penyaringan ini pada sistem real time seperti video.

Saran terakhir adalah dengan mengganti algoritma greedy dengan algoritma yang mampu melakukan penyaringan dengan menghasilkan solusi optimal global, seperti dynamic programming yang tidak langsung membuang *bounding box* seperti pada algoritma greedy. Dengan ini program akan mempertimbangkan berbagai kombinasi *bounding box* untuk menghasilkan solusi yang benar-benar optimal secara global.

REFERENCES

- [1] <https://xlinux.nist.gov/dads/HTML/greedyalgo.html>
- [2] A. Bochkovskiy, C. Y. Wang, H.-Y. M. Liao, YOLOv4: Optimal Speed and Accuracy of Object Detection:2020.
- [3] J. Redmon, S Divvala, R. Girshick, A. Fahadi, You Only Look Once: Unified, Real-Time Object Detection
- [4] H. Rezatofghi et al., Generalized Intersection over Union: A Metric and a Loss for Bounding Box Regression.
- [5] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, Introduction to Algorithms 3rd ed., MIT Press, 2009, pp. 414-450

Tautan kode sumber:

<https://github.com/julius123123/Reduksi-Duplikasi-Bounding-Box>

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 24 Juni 2025



Julius Arthur 13523030