

Sistem Deteksi Objek Pada Citra Menggunakan Metode YOLOv4

Dika Fajar Ramadan

Universitas Telkom

Kab Bandung, Jawa Barat

dikafajarramadan@student.telkomuniversity.ac.id

Abstrak

Citra merupakan suatu informasi yang mengandung banyak informasi penting didalamnya. Kita dapat mendeteksi objek dari sebuah citra yang diproses oleh teknologi komputer. Sistem deteksi objek merupakan sistem yang terkait dengan suatu citra yang diolah untuk proses deteksi objek. Pada penelitian ini dilakukan pendeteksian objek pada citra menggunakan YOLOv4 yang saat ini memiliki performa yang baik untuk mendeteksi suatu objek dalam citra. Penelitian menggunakan dataset coco data yang sudah dilatih sebelumnya sehingga dapat langsung digunakan untuk prediksi pada deteksi objek. Hasil dari penelitian ini sistem dapat mendeteksi objek dengan baik dan mendapatkan akurasi yang cukup tinggi. Pada eksperimen pertama dengan objek burung elang mendapatkan akurasi sebesar 97%, eksperimen kedua pada citra yang mengandung objek kuda sistem juga berhasil mendeteksi objek dengan akurasi rata-rata 90% dan Eksperimen terakhir yaitu dilakukan deteksi objek pada citra yang mengandung berbagai jenis objek. Hasil dari proses deteksi objek tersebut, sistem berhasil mendeteksi berbagai jenis objek yang ada pada citra dengan akurasi rata-rata diatas 90%.

Kata kunci: Sistem Deteksi, YOLOv4, Citra

1. Pendahuluan

Citra merupakan informasi berbentuk visual yang sangat penting[1]. Didalam sebuah citra mengandung informasi-informasi yang dapat memberikan pengetahuan kepada yang melihatnya. Di dalam citra terdiri dari beberapa bagian yaitu bagian *background* dan bagian objek. Manusia dapat secara langsung melihat objek di dalam citra tersebut. Namun terkadang objek yang sangat banyak dan bertumpuk membuat mata manusia menjadi lelah. Sistem deteksi objek merupakan suatu teknologi komputer yang berkaitan dengan citra digital dan pengolahan citra yang berfungsi untuk mendeteksi objek seperti Mobil, Manusia, Hewan, dan barang. Proses deteksi objek terdiri dari pencarian posisi dan ukuran setiap bentuk objek yang ada di dalam gambar. Tujuan deteksi objek yaitu menemukan semua bentuk objek dari satu atau beberapa golongan objek tertentu bagaimanapun ukuran, posisi, pose, sudut pandang pengambilan, dan kondisi pencahayaan objek tersebut. Pada penelitian ini digunakan metode YOLOv4 untuk mendeteksi objek pada suatu citra. YOLO merupakan teknologi kecerdasan buatan yang sangat berkembang saat ini dengan memiliki berbagai acuan parameter untuk dijalankan meliputi akurasi deteksi, kecepatan deteksi, dan ukuran model yang di hasilkan[2]. YOLOv4 merupakan metode deteksi objek yang menekankan konsep *one-stage object detection* dengan keunggulan pada waktu pemrosesan dan akurasi yang dihasilkan[2].

2. Penelitian Terkait

Telah dilakukan penelitian sebelumnya untuk mendeteksi objek menggunakan metode YOLOv4. Pada penelitian yang berjudul “YOLOv4 dan Mask R-CNN untuk Deteksi Kerusakan Pada Karung Komoditi” yaitu, mendeteksi kerusakan dan lubang pada karung komoditi yang ada pada suatu *frame* gambar. YOLOv4 merupakan metode deteksi objek yang menekankan konsep *one-stage object detection* dengan keunggulan pada waktu pemrosesan dan akurasi yang dihasilkan. Dataset yang digunakan merupakan hasil akuisisi kamera dengan jumlah yang disamakan pada masing-masing algoritma senilai 700 data training, 100 data validasi. Pengujian terhadap kedua algoritma tersebut dilakukan terhadap sejumlah 20 data uji dengan perbandingan terhadap kalkulasi secara manual oleh mata manusia. Hasilnya didapatkan bahwa model dari YOLOv4 dapat memberikan performa lebih baik dengan akurasi 96,8%, sedangkan model MaskR-CNN mengalami kinerja yang kurang dapat diandalkan dengan akurasi 65,78% pada data uji yang sama[2].

Selain itu terdapat juga penelitian lain yang mendeteksi citra menggunakan metode YOLOv4, yaitu penelitian yang berjudul “Perbandingan Algoritma YOLOv4 dan Scaled YOLOv4 untuk Deteksi Objek pada Citra Termal”, pada penelitian ini menggunakan data citra termal untuk dalam bentuk Citra RGB sebagai deteksi objek. Hasil dari penelitian ini deteksi didapatkan algoritma Scaled YOLOv4 CSP lebih unggul dengan nilai pengujian precision 94,3%, recall 83,8%, f1-Score 88,7%, dan mAP 86,9%. Hasil tersebut dipengaruhi oleh ukuran citra dan jumlah dataset dari citra *training*, citra *validation*, dan citra uji[3].

Penelitian lain yang mendeteksi objek menggunakan YOLO juga dilakukan untuk mendeteksi gulma, namun metode yang digunakan adalah YOLOv3, Data yang digunakan berupa video yang dipecah menjadi rasio 16:17:13, Hasil dari penelitian ini berhasil menunjukkan keefektifan sistem YOLO-WEED untuk

deteksi secara *real time* gulma UAV, dan memberikan kecepatan tinggi dan akurasi tinggi dalam deteksi[4].

3. Data

Dataset yang digunakan pada penelitian ini yaitu menggunakan coco dataset yang telah dilatih sebelumnya. Pada dataset tersebut terdapat 80 kelas yang dapat di prediksi secara langsung. Pada penelitian ini menggunakan gambar yang dapat mendeteksi objek-objek didalamnya, tidak terbatas hanya pada objek tertentu. Misalkan di dalam citra terdapat objek orang, mobil, dan kuda, maka seluruh objek tersebut akan terdeteksi oleh *bounding box* dan diberikan akurasinya yang berada diatas *boundingbox* tersebut.



Gambar 1. Contoh Dataset yang diambil dari COCO dataset.

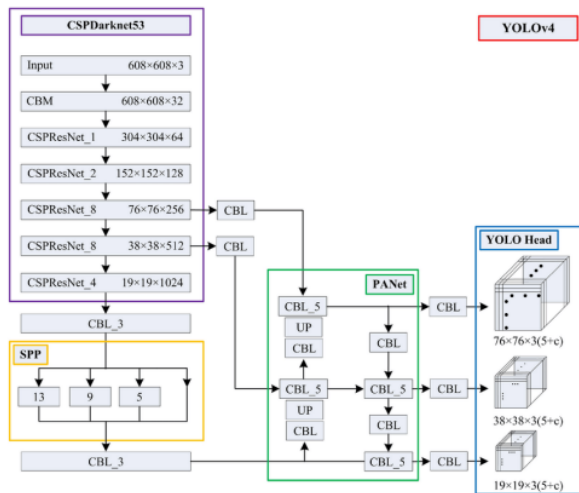
4. Metode

4.1. YOLOv4

Pada penelitian yang dilakukan oleh Redmon, YOLOv4 dapat mendeteksi objek secara *realtime* dengan kecepatan 45fps karena memiliki arsitektur yang sederhana dan mendapatkan akurasi yang cukup tinggi yaitu 88% pada *imagenet2012 validation*[5][3].

Arsitektur YOLOv4 menggunakan backbone: CSPDarknet53, neck: SPP dan PAN, dan head:

YOLOv3 seperti pada Gambar 5. Feature extraction yang bernama CSPDarknet53 ini memiliki 29 convolutional layers 3×3, 725×725 pada receptive field dan 27.6 M parameter. Berdasarkan nilai ini, CSPDarknet53 merupakan model yang optimal sebagai feature extraction atau backbone dari detektor. Beberapa pengaruh dari kemampuannya saat citra di-input yakni bisa melihat objek secara detail, dan meningkatkan jumlah kombinasi setiap titik di suatu citra. SPP pada neck membantu memisahkan fitur konteks secara signifikan dan tidak mengurangi kecepatan operasi jaringan. Sedangkan, PANet digunakan sebagai parameter penyatuan dari beberapa *level backbone*[3][6].



Gambar 2. Arsitektur YOLOv4

conv	32	3 x 3 / 1	608 x 608 x 3 -> 608 x 608 x 32 0.639 BF
1 conv	64	3 x 3 / 2	608 x 608 x 32 -> 304 x 304 x 64 3.407 BF
2 conv	64	1 x 1 / 1	304 x 304 x 64 -> 304 x 304 x 64 0.757 BF
3 route	1		-> 304 x 304 x 64
4 conv	64	1 x 1 / 1	304 x 304 x 64 -> 304 x 304 x 64 0.757 BF
5 conv	32	1 x 1 / 1	304 x 304 x 64 -> 304 x 304 x 32 0.379 BF
6 conv	64	3 x 3 / 1	304 x 304 x 32 -> 304 x 304 x 64 3.407 BF
7 Shortcut Layer: 4,	wt = 0, wn = 0,	outputs: 304 x 304 x 64 0.006 BF	
8 conv	64	1 x 1 / 1	304 x 304 x 64 -> 304 x 304 x 64 0.757 BF
9 route	8 2		-> 304 x 304 x 128
10 conv	64	1 x 1 / 1	304 x 304 x 128 -> 304 x 304 x 64 1.514 BF
11 conv	128	3 x 3 / 2	304 x 304 x 64 -> 152 x 152 x 128 3.407 BF
12 conv	64	1 x 1 / 1	152 x 152 x 128 -> 152 x 152 x 64 0.379 BF
13 route	11		-> 152 x 152 x 128
14 conv	64	1 x 1 / 1	152 x 152 x 128 -> 152 x 152 x 64 0.379 BF
15 conv	64	1 x 1 / 1	152 x 152 x 64 -> 152 x 152 x 64 0.189 BF
16 conv	64	3 x 3 / 1	152 x 152 x 64 -> 152 x 152 x 64 1.703 BF
17 Shortcut Layer: 14,	wt = 0, wn = 0,	outputs: 152 x 152 x 64 0.001 BF	
18 conv	64	1 x 1 / 1	152 x 152 x 64 -> 152 x 152 x 64 0.189 BF
19 conv	64	3 x 3 / 1	152 x 152 x 64 -> 152 x 152 x 64 1.703 BF
20 Shortcut Layer: 17,	wt = 0, wn = 0,	outputs: 152 x 152 x 64 0.001 BF	
21 conv	64	1 x 1 / 1	152 x 152 x 64 -> 152 x 152 x 64 0.189 BF
22 route	21 12		-> 152 x 152 x 128
23 conv	128	1 x 1 / 1	152 x 152 x 128 -> 152 x 152 x 128 0.757 BF
24 conv	256	3 x 3 / 2	152 x 152 x 128 -> 76 x 76 x 256 3.407 BF
25 conv	128	1 x 1 / 1	76 x 76 x 256 -> 76 x 76 x 128 0.379 BF
26 route	24		-> 76 x 76 x 256
27 conv	128	1 x 1 / 1	76 x 76 x 256 -> 76 x 76 x 128 0.379 BF
28 conv	128	1 x 1 / 1	76 x 76 x 128 -> 76 x 76 x 128 0.189 BF
29 conv	128	3 x 3 / 1	76 x 76 x 128 -> 76 x 76 x 128 1.703 BF
30 Shortcut Layer: 27,	wt = 0, wn = 0,	outputs: 76 x 76 x 128 0.001 BF	

Gambar 3. Arsitektur YOLOv4

4.2. Rancangan Penelitian

Adapun rancangan penelitian yang dilakukan yang dilakukan pada tugas besar ini adalah melakukan deteksi objek menggunakan data yang didapatkan dari coco dataset yang telah di latih sebelumnya sehingga dapat secara langsung diprediksi untuk proses deteksi menggunakan metode YOLOv4 dan darknet. YOLO pre-traned weight diunduh di website darknet kemudian dilakukan deteksi objek menggunakan darknet tersebut. Pengujian ini dilakukan menggunakan google colab dan di dijalankan menggunakan jenis *runtime* GPU. Hasil dari system deteksi ini akan menampilkan citra dengan objek yang ditandai dengan *boundingbox* dan akurasi. Berikut merupakan perhitungan akurasi untuk mendeteksi objek pada citra.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positives (TP)	False Negatives (FN)
	Negative	False Positives (FP)	True Negatives (TN)

Gambar 4. Tabel Confussion Matrix

dimana TP (*True positive*) adalah perkiraan kejadian benar dan hasil nyata juga benar, TN (*True Negative*) yaitu perkiraan kejadian benar namun hasilnya salah, FP (*False Positive*) yaitu perkiraan kejadian benar dan hasilnya salah, dan FN (*False Negative*) yaitu perkiraan kejadian salah namun hasilnya benar [7].

5. Eksperimen dan Hasil

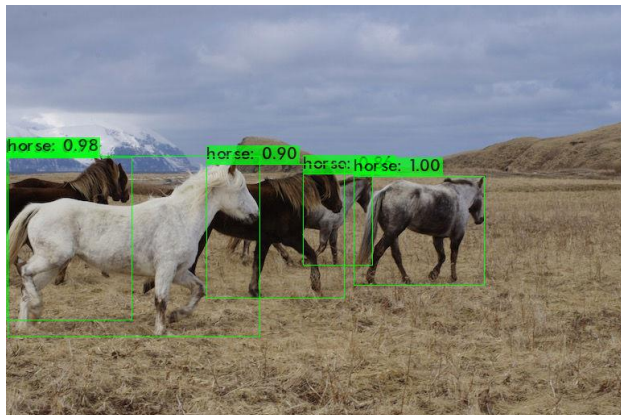
Pada tugas besar ini dilakukan 3 pengujian menggunakan dataset yang berbeda. Yang pertama menggunakan citra yang hanya

mengandung satu jenis objek, yang kedua menggunakan citra yang mengandung satu jenis namun banyak objeknya dan yang terakhir menggunakan citra yang terdapat beberapa jenis objek di dalamnya.



Gambar 2. Citra Burung Elang Setelah dideteksi

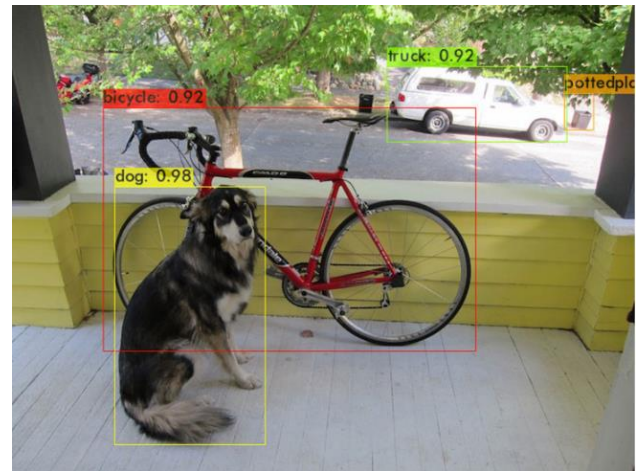
Eksperimen atau pengujian pertama yaitu dilakukan deteksi objek menggunakan citra yang terdapat objek burung elang di dalamnya, Dari gambar diatas dapat kita lihat bahwa system berhasil mendeteksi objek Burung Elang yang ditandai dengan *boundingbox* dan akurasinya sebesar 0.97 atau 97%.



Gambar 3. Citra Kuda Setelah dideteksi

Pada eksperimen kedua dilakukan deteksi objek pada citra yang mengandung beberapa objek kuda. Dari gambar diatas kita dapat melihat bahwa System berhasil mendeteksi beberapa kuda tersebut yang ditandai dengan *bounding box* dan sekaligus dengan hasil akurasi

yang didapatkan dalam proses deteksi. Akurasi yang didapatkan pada kuda-kuda tersebut yaitu 90-100%.



Gambar 4. Citra dengan Berbagai Objek

Eksperimen atau pengujian terakhir yaitu dilakukan deteksi objek pada citra yang mengandung berbagai jenis objek. Dari gambar diatas dapat kita lihat hasil dari proses deteksi objek tersebut, system berhasil mendeteksi berbagai jenis objek diantara seekor anjing, sepeda, truk, dan tong sampah yang ada pada citra dengan akurasi 92-98%.

6. Kesimpulan

Dari hasil pengujian atau eksperimen yang telah dilakukan dapat kita ketahui bahwa system deteksi objek menggunakan metode YOLOv4 berhasil dalam mendeteksi objek yang ada pada citra. Objek-objek yang dideteksi tidak terbatas pada satu objek saja namun dapat mendeteksi berbagai objek yang ada pada citra. Pada eksperimen pertama Citra yang mengandung objek burung elang berhasil di deteksi dengan akurasi 97%, sedangkan citra yang mengandung beberapa kuda mendapatkan akurasi 98-100% dan citra yang mengandung berbagai jenis objek mendapat akurasi 92-98%. ini menunjukkan bahwa metode YOLOv4 optimal dan efisien dalam mendeteksi suatu objek pada citra. Untuk penelitian selanjutnya akan membanding semua

jenis metode YOLO untuk dibandingkan hasil dan performanya dalam mendeteksi objek.

References

- [1] S. Sugiarti, “Peningkatan Kualitas Citra Dengan Metode Fuzzy Possibility Distribution,” *Ilk. J. Ilm.*, vol. 10, no. 1, pp. 100–104, 2018.
- [2] E. R. Setyaningsih and M. S. Edy, “YOLOv4 dan Mask R-CNN Untuk Deteksi Kerusakan Pada Karung Komoditi,” *Teknika*, vol. 11, no. 1, pp. 45–52, 2022.
- [3] A. A. RAHMAN, S. D. AGUSTIN, N. IBRAHIM, and N. C. KUMALASARI, “Perbandingan Algoritma YOLOv4 dan Scaled YOLOv4 untuk Deteksi Objek pada Citra Termal,” *MIND J.*, vol. 7, no. 1, pp. 61–71, 2022.
- [4] A. I. B. PARICO and T. AHAMED, “An Aerial Weed Detection System for Green Onion Crops Using the You Only Look Once (YOLOv3) Deep Learning Algorithm,” *Eng. Agric. Environ. Food*, vol. 13, no. 2, pp. 42–48, 2020.
- [5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection Joseph,” *ACM Int. Conf. Proceeding Ser.*, 2018.
- [6] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” 2020.
- [7] U. Ahmed, R. Mumtaz, H. Anwar, A. A. Shah, R. Irfan, and J. García-Nieto, “Efficient water quality prediction using supervised machine learning,” *Water (Switzerland)*, vol. 11, no. 11, pp. 1–14, 2019.