

Deteksi dan tracking objek secara *real-time* berbasis computer vision menggunakan metode YOLO V3

¹Iwan Virgiawan, ²Fadel Maulana, ³Muhammad Abimanyu Putra, ⁴Diffa Dwi Kurnia, ⁵Estu Sinduningrum
^{1,2,3,4,5}Universitas Muhammadiyah Prof. Dr. Hamka, Indonesia

INFO ARTIKEL

Sejarah artikel:

Diterima
27 April 2024
Disetujui
8 Mei 2024
Diterbitkan
25 Mei 2024

Penulis Korespondensi*:

Iwan Virgiawan
Universitas Muhammadiyah
Prof. Dr. Hamka, Indonesia
iwanvi02@gmail.com



©2023 Penulis. Diterbitkan oleh Program Studi Akuntansi, Institut Koperasi Indonesia. Ini adalah artikel akses terbuka di bawah lisensi CC BY NC (<https://creativecommons.org/licenses/by/4.0>)

ABSTRAK

Tujuan dari penelitian ini bermaksud untuk mengimplementasikan pendekatan YOLOv3 dalam sistem deteksi objek *real-time*. Metode yang digunakan dalam penelitian ini adalah YOLO V3 (You Only Look Once). Metode ini dikombinasikan dengan algoritma tracking, untuk melacak objek secara berkelanjutan. Hasil dari penelitian ini menunjukkan bahwa metode YOLO V3 mampu melakukan deteksi objek secara cepat dan akurat dalam berbagai kondisi lingkungan. Kecepatan deteksi dan tracking objek yang dihasilkan memenuhi persyaratan *real-time*, sehingga dapat digunakan dalam aplikasi yang membutuhkan respons cepat, seperti pengawasan keamanan, pengenalan wajah, dan mobil otonom. Hasil dari penelitian ini adalah Sistem ini memiliki jangkauan deteksi hingga 100 cm saat objek dekat dengan kamera, dan hingga 150 cd/m² saat objek berada dalam cahaya paling terang. Sistem memiliki waktu deteksi kurang dari 3 detik dan akurasi deteksi 100% untuk banyak item dalam satu gambar. Sistem dapat mendeteksi gambar secara akurat dan mendeteksi hewan secara akurat. Sistem dapat mendeteksi objek dalam keadaan gelap yang sudah ditentukan ketika objek tersebut memiliki kontur yang sama dengan object yang ada di dataset meski mengalami penurunan akurasi yang cukup signifikan dan waktu deteksi.

Kata Kunci : Deteksi Objek; Tracking Objek; *Real-time*; Computer Vision; YOLO V3.

ABSTRACT

The aim of this research is to implement the YOLOv3 approach in a real-time object detection system. The method used in this research is YOLO V3 (You Only Look Once). This method is combined with a tracking algorithm, to track objects continuously. The results of this research show that the YOLO V3 method is able to detect objects quickly and accurately in various environmental conditions. The resulting object detection and tracking speed meets real-time requirements, so it can be used in applications that require fast response, such as security surveillance, facial recognition, and autonomous cars. The results of this research are that this system has a detection range of up to 100 cm when the object is close to the camera, and up to 150 cd/m² when the object is in the brightest light. The system has a detection time of less than 3 seconds and 100% detection accuracy for multiple items in a single image. The system can detect images accurately and detect animals accurately. The system can detect objects in predetermined dark conditions when the object has the same contour as the object in the data set, although it experiences a significant decrease in accuracy and detection time.

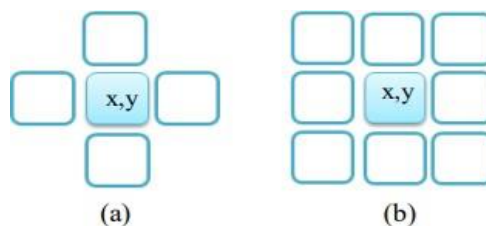
Keywords : Object Detection; Object Tracking; Real-time; Computer Vision; YOLO V3.

PENDAHULUAN

Seiring dengan berkembangnya penelitian tentang kecerdasan buatan salah satunya tentang pendeteksian objek, teknologi ini dapat membantu kita mengenali objek pada sebuah gambar. Pendeteksian objek adalah salah satu diantara bidang-bidang pada computer vision. Computer vision merupakan suatu ilmu yang mempelajari tentang seperti apa komputer dapat menganalisis dan melihat pada objek di dalam gambar (Dadang et al. 2022).

OpenCV adalah library fungsi pemrograman yang terutama digunakan untuk pengolahan citra. Ini tersedia secara bebas di sumber terbuka Lisensi Distribusi Perangkat Lunak Berkely. Itu dimulai sebagai proyek penelitian oleh Intel. OpenCV berisi berbagai alat untuk memecahkan masalah penglihatan komputer. Ini berisi gambar tingkat rendah fungsi pemrosesan dan algoritma tingkat tinggi untuk deteksi wajah, pencocokan fitur, dan pelacakan (M. Naveenkumar and A. Vadivel, 2015).

Schwenk dan Huber (2015) menyatakan bahwa saat membuat tag, digunakan 4-konektivitas atau 8-konektivitas. Piksel pusat dalam elemen bertetangga 4-terhubung ditampilkan dalam empat piksel yang berdekatan baik secara horizontal maupun vertikal. Jika dua piksel yang berdekatan memiliki label model warna yang sama, piksel tengah diberi label dari dua piksel yang berdekatan. Piksel pusat dilambangkan dengan 8 piksel melintasi dan ke bawah diagonal, berlawanan dengan 4 piksel untuk masing-masing komponen koneksi yang berdekatan. Jika dua piksel yang berdekatan memiliki label model warna yang sama, piksel tengah diberi label dari dua piksel yang berdekatan. Gambar 1 menggambarkan bentuk jenis koneksi.



Gambar 1. Bentuk tipe connectivity (a) Tipe 4-connectivity; (b) Tipe 8-connectivity

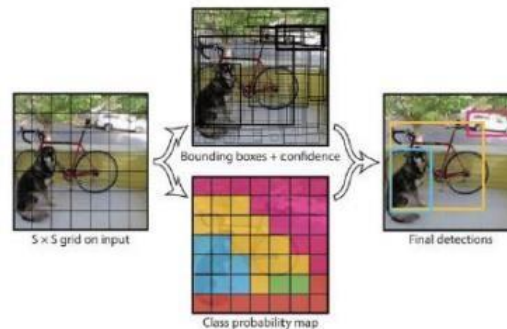
Deteksi objek secara real-time telah menjadi topik penelitian yang penting dalam bidang computer vision. Kemampuan untuk mengenali dan membedakan objek dalam gambar atau video dengan cepat dan akurat memiliki berbagai aplikasi praktis, (Hanchinamani et al., 2016) termasuk pengawasan keamanan, pengenalan wajah, kendaraan otonom, dan banyak lagi. Dalam pendeteksian objek secara real-time, kecepatan sangat penting dalam pendeteksian objek, karena berbeda dengan sebuah gambar, sebuah video dapat memproses lebih dari 24 frame per detik (FPS) atau 24 frame per detik. Jika proses pendeteksian objek terlalu lama maka video yang dihasilkan tidak bagus, akan terjadi delay pada setiap frame sehingga video menjadi pecah-pecah (Aini et al. 2021)

Dalam beberapa tahun terakhir, teknik pembelajaran mendalam telah diterapkan dalam deteksi objek (Zhou et al., 2017). Deep learning menggunakan fitur tingkat rendah untuk membentuk fitur tingkat tinggi yang lebih abstrak, dan untuk merepresentasikan data secara hierarki guna meningkatkan deteksi objek dengan salah satu metode yang paling populer adalah YOLO (You Only Look Once) v3.

Tidak seperti kebanyakan teknik deteksi pembelajaran mendalam lainnya, algoritme YOLO menggunakan jaringan saraf tunggal untuk menganalisis keseluruhan gambar. YOLO membuat grid $S \times S$ dari gambar yang diberikan. Ukuran setiap sel grid ditentukan oleh besarnya data yang dimasukkan ke dalam desain. Jika ukuran input 416×416 , ukuran grid yang sesuai di YOLOv3 adalah 13×13 , 26×26 , dan 52×52 . Tugas setiap sel adalah menebak apa yang ada di dalamnya berdasarkan kotak pembatas dan nilai probabilitas keberadaan objek di sana. Ini telah terbukti menjadi kasusnya (you only look).

Selanjutnya, YOLO memprediksi kelas item yang ada di kotak pembatas dan probabilitasnya, menghasilkan peta probabilitas untuk kelas tersebut, berdasarkan tingkat kepercayaan yang diperoleh. Untuk meningkatkan kemungkinan mengidentifikasi kotak pembatas dan kelas fitur dengan benar, hanya kotak pembatas dengan probabilitas lebih

besar dari ambang batas yang digunakan. Non-Max Suppression (NMS) bertanggung jawab untuk menekan duplikat jika ada tumpang tindih di kotak pembatas. Metode YOLO ditunjukkan pada Gambar 2 (Asyhar et al., 2020).



Gambar 2. Ilustrasi algoritma YOLO

Berfokus pada pengenalan objek video, penelitian ini bermaksud untuk mengimplementasikan pendekatan YOLOv3 dalam sistem deteksi objek real-time. Untuk meningkatkan performa pendeteksian objek secara real-time, penelitian ini berangkat dari sebelumnya dengan menggunakan pendekatan YOLOv3 yang lebih cepat dan akurat untuk mendeteksi objek.

Tujuan utama dari perangkat lunak ini adalah deteksi objek secara real-time, oleh karena itu pertama-tama akan dibuat sistem yang dapat melakukan deteksi objek dengan kecepatan tinggi secara real time. Pengenalan objek menggunakan YOLOv3 telah menjadi subjek dari sejumlah penelitian sebelumnya, meskipun analisis ini kurang memperhatikan kinerja waktu nyata. Berbagai jenis penelitian telah dilakukan di dua sebelumnya pendekatan untuk mengatasi masalah deteksi objek, seperti menggunakan deteksi objek dengan algoritma Haar, HOG, dan SVM. Namun, metode YOLOv3 memiliki kekurangan dalam keterbatasan dataset yang di pakai.

METODE PENELITIAN

1. Perancangan sistem

Untuk mengambil gambar secara real time, sistem deteksi objek menggunakan webcam yang terpasang pada komputer pribadi (PC). Setelah database citra disusun, citra diproses menggunakan pendekatan image fitting berbasis image processing dengan membandingkan citra hasil olahan dengan contoh dunia nyata. Sekarang mungkin saat yang tepat untuk mengambil beberapa bidikan. sedemikian rupa sehingga dapat diidentifikasi oleh komputer dan diikuti.

2. Perancangan perangkat keras dan perangkat lunak

Sumber daya dan peralatan yang akan digunakan dalam pembuatan program ini diuraikan dalam bagian ini.

Perangkat keras:

- a. Laptop lenovo ideapad3i dengan spesifikasi RAM 16384MB RAM, AMD Ryzen 5 4600H@ 3,0GHz
- b. Webcam lenovo ideapad3i , 0,92 Mega Pixel
- c. SSD 512GB 20Gbps
- d. Layar Anti glare IPS LED backlight 14 inci resolusi Full HD (1920 x 1080) pixel
- e. Grafis Radeon grafik vega7 dan Nvidia Geforce GTX 1650TI 4GB
- f. DDR6

Perangkat Lunak :

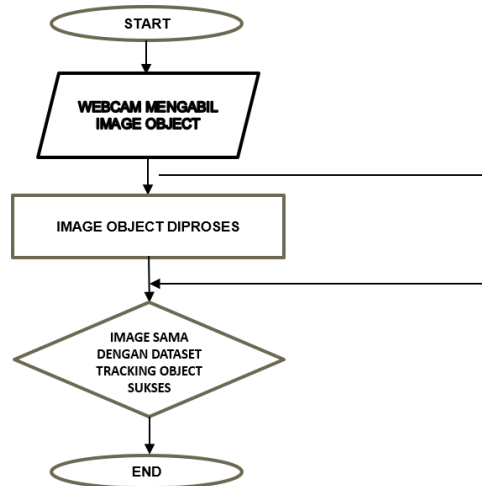
- a. Python 3.11 (64-bit)

Deteksi dan tracking objek secara real time berbasis computer vision menggunakan metode YOLO V3

- b. Library OpenCV
- c. Jupyter
- d. Visual studio code
- e. Numpy

3. Perencanaan dan implementasi sistem secara keseluruhan

Prosedur YOLOv3 untuk membuat sistem objek memerlukan sejumlah tahapan. Flowchart sistem ini adalah sebagai berikut (Aprilino & Amin, 2022).



Gambar 3. Flowchart deteksi object

Ada beberapa langkah yang terlibat dalam proses YOLOv3 untuk mengembangkan sistem objek. Bagan alir sistem ini terlihat seperti ini: (Aprilino & Amin, 2022).



Gambar 4. Sistem mendeteksi adanya objek berupa orang yang ditampilkan dalam output.

HASIL DAN PEMBAHASAN

1. Hasil

Sistem identifikasi objek berbasis Python dan OpenCV ditunjukkan pada Gambar 4. Komputer akan mencari sesuatu dalam berbagai gambar dan video. Dimulai dengan ukuran minimum 4x4 yang diulang berkali-kali dengan perbesaran 1,1, gambar masukan dipindai ke jendela tengah atau dari item yang pertama kali ditampilkan di kamera. Ulangi prosedur ini dengan menggeser x dan y ke kanan. Setiap subwindow yang dipindai memiliki fungsi

YOLO yang diterapkan padanya. Karena Grid dengan ukuran yang telah ditentukan digunakan untuk mempartisi gambar. Jika item berada di dalam sel kisi tertentu, sel tersebut bertanggung jawab untuk melakukan pendeteksian. Beberapa prediksi kotak pembatas, masing-masing menunjukkan kemungkinan lokasi dan ukuran item di dalam sel kisi tersebut, dibuat untuk setiap sel. Selain itu, probabilitas kelas untuk setiap item yang teridentifikasi ditautkan ke kotak pembatas yang sesuai. Menurut penelitian terbaru

Evaluasi dilakukan pada ruangan dengan pencahayaan alami dan pencahayaan buatan dengan menggunakan laptop yang dilengkapi kamera beresolusi 0,92 MP. Pencahayaan, jarak, banyak objek, dan gambar atau binatang hanyalah beberapa faktor yang dapat diuji.

Pengujian pencahayaan tampak juga ditunjukkan pada Gambar 3. Jika item menghadap ke sistem, item tersebut akan terdeteksi dengan benar. Akurat menyiratkan tidak ada positif atau negatif palsu dalam proses deteksi gambar wajah. Beberapa percobaan pada kualitas objek yang dapat dideteksi ditunjukkan pada Gambar 5 sampai 8.

Tabel 1. Hasil Pengujian

No	Pengajuan	Hasil
1	Berdasarkan pencahayaan	70 nilai mean untuk maksimal gelap, 150 nilai mean untuk maksimal terang, pencahayaan harus merata pada objek
2	Jarak	100 cm maksimal terjauh 20 cm maksimal terdekat
3	Multiple Object	Sistem mampu mendeteksi adanya beberapa objek
4	Gambar	Sistem mampu mendeteksi gambar dengan akurat sesuai dataset yang ada
5	Hewan	Sistem mampu deteksi hewan secara akurat

Dalam tabel 1 diberitahukan bahwa dari beberapa pengajuan yang di ujikan dalam bentuk berdasarkan pencahayaan, berdasarkan jarak, berdasarkan multiple object, gambar dan hewan menunjukan hasil yang konsisten

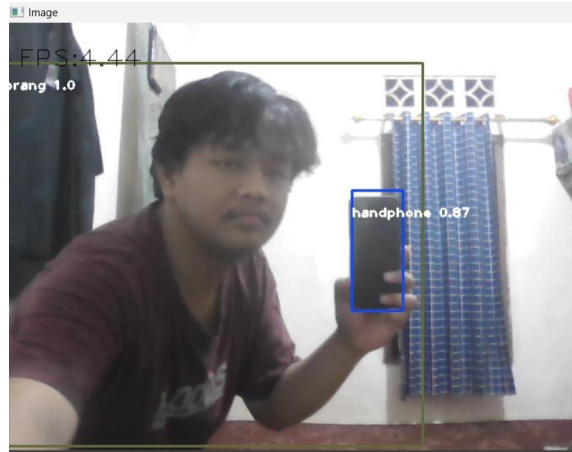
2. Pembahasan

Pada penelitian ini kita dapat mendeteksi objek secara real time dan gambar sedangkan pada penelitian terdahulu kebanyakan hanya dapat mendeteksi gambar saja, tidak secara real time, dan penelitian terdahulu hanya dapat mendeteksi satu objek seperti buah-buahan, pakaian, dan juga helm, oleh karena itu kita membuat penelitian ini untuk memperluas pendeteksi dalam mendeteksi, dalam penelitian ini kita dapat mendeteksi banyak hal, seperti orang, benda, buah-buahan maupun hewan.

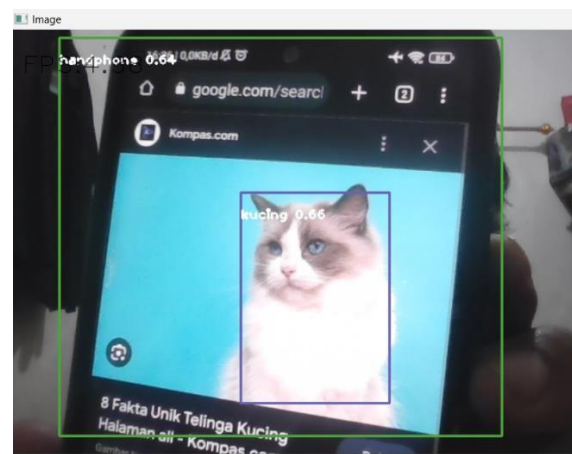
Karena dalam penelitian yang berisikan tiga kelas yang akan dideteksi yaitu apple, banana dan orange. Setelah itu ada `images_path` untuk mengetahui gambar apa yang akan kita deteksi dan hasilnya itu akan mengeluarkan bounding box prediksinya J. Jabez, G. Shanmugam, V. S, J. A. Mayan, and S.Srinivasulu (2019), dalam penelitiannya hanya menampilkan foto yang berobjek buah untuk mendeteksi citra pada foto buah tersebut begitu pula pada penelitian yang dilakukan oleh Seong, H., Choi, H., Cho, H., Lee, S., Son, H., & Kim, C. (2017). Deteksi alat pelindung diri dalam beberapa tahun terakhir telah mendapatkan perhatian yang cukup signifikan. Hal ini dianggap penting karena sangat diperlukan dalam manajemen keselamatan dan produktivitas para pekerja konstruksi hanya mendeteksi objek secara banyak dalam konteks tidak bisa mendeteksi helm, rompi keselamatan dan orang secara bersamaan.

Berikut ini adalah contoh dari hasil penelitian ini yang dapat mendeteksi berbagai objek :

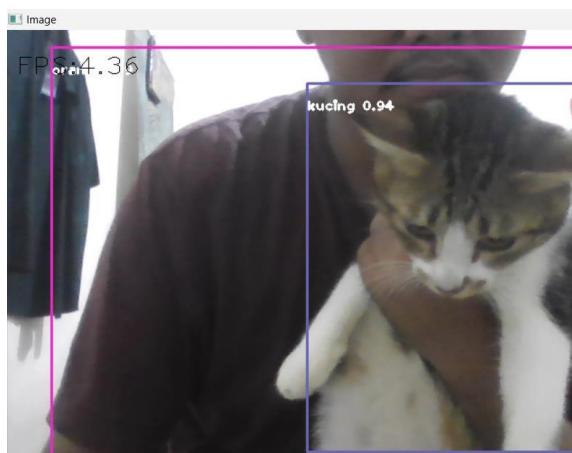
Deteksi dan tracking objek secara real time berbasis computer vision menggunakan metode YOLO V3



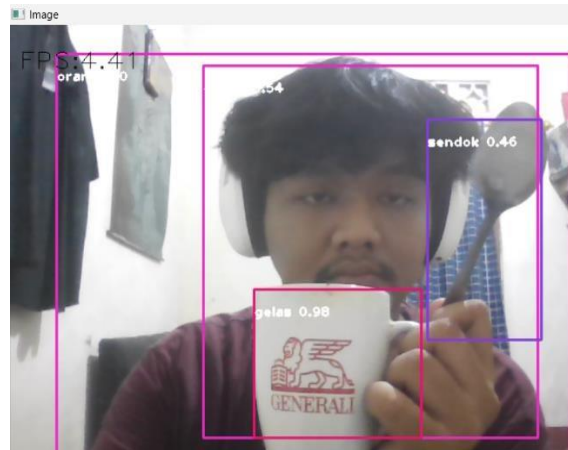
Gambar 5. Dalam gambar 5 sistem mendeteksi adanya objek berupa handphone dengan kondisi jarak tertentu dari kamera.



Gambar 6 . Dalam gambar 6 sistem mendeteksi gambar dan hewan apa yang ada di gambar tersebut.



Gambar 7. Dalam gambar 7 sistem mendeteksi adanya objek dengan output hewan berupa kucing.



Gambar 8. Dalam gambar 8 sistem mampu mendeteksi adanya objek dalam jumlah lebih dari 3 yang ditampilkan dalam output.

Kekurangan dalam penelitian ini adalah penelitian ini tidak dapat mendeteksi atau mengenali profil dari orang yang terdeteksi dan juga penelitian ini memiliki batasan yang tidak menimbulkan suara ketika orang atau benda terdeteksi sebagai tanda, oleh karena itu peneliti selanjutnya dapat mengembangkan penelitian ini dengan menambahkan pengenalan profil orang yang terdetek dan juga dapat menimbulkan suara ketika ada barang yang terdeteksi agar lebih mudah untuk mengenali objek tersebut.

KESIMPULAN

Sistem ini memiliki jangkauan deteksi hingga 100 cm saat objek dekat dengan kamera, dan hingga 150 cd/m² saat objek berada dalam cahaya paling terang. Sistem memiliki waktu deteksi kurang dari 3 detik dan akurasi deteksi 100% untuk banyak item dalam satu gambar. Sistem dapat mendeteksi gambar secara akurat sesuai dataset yang ada. Sistem dapat mendeteksi hewan secara akurat. Sistem dapat mendeteksi objek dalam keadaan gelap yang sudah ditentukan dalam poin 2 ketika objek tersebut memiliki kontur yang sama dengan object yang ada di dataset meski mengalami penurunan akurasi yang cukup signifikan dan waktu deteksi.

DAFTAR PUSTAKA

- Aini, Q., Lutfiani, N., Kusumah, H., & Zahran, M. S. (2021). Deteksi dan Pengenalan Objek Dengan Model Machine Learning: Model Yolo. *CESS (Journal of Computer Engineering, System and Science)*, 6(2), 192.
- Al Asyhar, H. H., Wibowo, S. A., & Budiman, G. (2020). Implementasi dan analisis performansi metode you only look once (yolo) sebagai sensor pornografi pada video. *eProceedings of Engineering*, 7(2).
- Hanchinamani, S. R., Sarkar, S., & Bhairannawar, S. S. (2016). Design and implementation of high speed background subtraction algorithm for moving object detection. *Procedia computer science*, 93, 367-374. <https://doi.org/10.1016/j.procs.2016.07.222>
- Jabez, J., Gowri, S., Vigneshwari, S., Albert Mayan, J., & Srinivasulu, S. (2019). Anomaly detection by using CFS subset and neural network with WEKA tools. In *Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2018, Volume 2* (pp. 675-682). Springer Singapore.
- Li, Y., Rong, L., Li, R., & Xu, Y. (2022, April). Fire Object Detection Algorithm Based on Improved YOLOv3-tiny. In *2022 7th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)* (pp. 264-269). IEEE. <https://doi.org/10.1109/ICCCBDA55098.2022.9778892>
- Mulyana, D. I., & Rofik, M. A. (2022). Implementasi Deteksi Real Time Klasifikasi Jenis Kendaraan Di Indonesia Menggunakan Metode YOLOV5. *Jurnal Pendidikan*

Tambusai, 6(3), 13971-13982.

- Naveenkumar, M., & Vadivel, A. (2015, March). OpenCV for computer vision applications. In *Proceedings of national conference on big data and cloud computing (NCBDC'15)* (pp. 52-56). <https://www.researchgate.net/publication/301590571>
- Nurjanah, T. S., & Insanudin, E. (2016). Hack Database Website Menggunakan Python dan Sqlmap Pada Windows. <https://www.researchgate.net/publication/303372599>
- Plastiras, G., Kyrkou, C., & Theocharides, T. (2018, September). Efficient convnet-based object detection for unmanned aerial vehicles by selective tile processing. In *Proceedings of the 12th international conference on distributed smart cameras* (pp. 1-6). <https://doi.org/10.1145/3243394.3243692>
- Prabowo, D. A., & Abdullah, D. (2018). Deteksi dan perhitungan objek berdasarkan warna menggunakan Color Object Tracking. *Pseudocode*, 5(2), 85-91. www.ejournal.unib.ac.id/index.php/pseudocode
- Pratama, Y. B., & Dalimunthe, N. P. (2022). Implementasi Teknik Computer Vision Untuk Deteksi Viridiplantae Pada Lahan Pasca Tambang. *Bulletin of Computer Science Research*, 3(1), 64-72. <https://doi.org/10.47065/bulletincsr.v3i1.193>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788). <http://pjreddie.com/yolo/>
- Schwenk, K., & Huber, F. (2015, October). Connected component labeling algorithm for very complex and high-resolution images on an FPGA platform. In *High-Performance Computing in Remote Sensing V* (Vol. 9646, pp. 9-22). SPIE. <https://doi.org/10.1117/12.2194101>
- Seong, H., Choi, H., Cho, H., Lee, S., Son, H., & Kim, C. (2017). Vision-based safety vest detection in a construction scene. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction* (Vol. 34). IAARC Publications..
- Wibowo, A. P. W. (2016). Implementasi Teknik Computer Vision Dengan Metode Colored Markers Trajectory Secara Real Time. *Jurnal Teknik Informatika*, 8(1), 38-42.
- Awan Aprilino, & Imam Husni Al Amin. (2022). Implementasi Algoritma Yolo Dan Tesseract Ocr Pada Sistem Deteksi Plat Nomor Otomatis. *Jurnal Teknoinfo*.
- Zhou, X., Gong, W., Fu, W., & Du, F. (2017, May). Application of deep learning in object detection. In *2017 IEEE/ACIS 16th international conference on computer and information science (ICIS)* (pp. 631-634). IEEE.