



PERBANDINGAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN) DAN ALGORITMA YOU ONLY LOOK ONCE (YOLO) UNTUK DETEKSI WAJAH

COMPARISON OF CONVOLUTIONAL NEURAL NETWORK (CNN) ALGORITHM AND YOU ONLY LOOK ONCE (YOLO) ALGORITHM FOR FACIAL DETECTION

Muhamad Abdurahman Kamil¹, Yan Mitha Djaksana².

Program Studi Teknik Informatika Universitas Pamulang

Email: abdurahmankamil00@gmail.com

Article Info

Article history :

Received : 02-11-2024

Revised : 04-11-2024

Accepted : 06-11-2024

Published : 09-11-2024

Abstract

This case of data loss when recording student attendance and spending time signing attendance if using a manual. Therefore, to minimize the case, his study aims to create system that can detect a human face object on face recognition using a laptop device. Mechanism system work is detecting students' faces and distinguishing between Convolutional Neural Network (CNN) and You Only Look Once (YOLO) algorithms. If a human object with another human is scanned by a laptop camera, then a green dividing box shaped line will appear around the object. The results of the system tool trial can detect human face objects with other humans according to the line around the face and detect objects that express the accuracy results between the two algorithms.

Keywords: *Computer Vision, Object Detection, Face recognition, CNN, YOLO*

Abstrak

Kasus kehilangan data pada saat merekapitulasi absensi siswa dan menghabiskan waktu pada saat menandatangani absen jika menggunakan manual. Sehingga untuk meminimalisir kasus tersebut, pada penelitian ini dibuat sebuah sistem yang dapat mendeteksi sebuah objek wajah manusia pada face recognition dengan menggunakan perangkat laptop. Mekanisme kerja sistem adalah mendeteksi wajah siswa dan Untuk membedakan algoritma Convolutional Neural Network (CNN) dan You Only Look Once (YOLO). Jika objek manusia dengan manusia yang lainnya di scan kamera laptop, maka akan muncul kotak pembatas berwarna hijau mengelilingi objek. Hasil dari uji coba alat sistem dapat mendeteksi objek wajah manusia dengan manusia lainnya sesuai kotak mengelilingi wajah dan mendeteksi objek yang berekspresi dan hasil akurasi antara kedua algoritma tersebut.

Kata kunci : *Computer Vision, Deteksi Wajah, Face Recognition, CNN, YOLO*

PENDAHULUAN

Artificial Intelligence (AI) saat ini sudah banyak digunakan, pada bidang teknologi khususnya. Teknologi AI dapat meringankan dan juga membantu pekerjaan manusia hampir disetiap bidang (Pradita et al., 2019; Anan et al. 2024). Teknologi Computer Vision adalah salah satu turunan dari kecerdasan buatan (AI). Saat ini dalam pengenalan suatu objek pada bidang computer vision merupakan hal yang sangat efektif digunakan dalam penelitian (Hasma & Silfianti, 2018), Selain pengenalan objek, pembelajaran mendalam (Deep Learning) juga dapat mengklasifikasikan sebuah gambar (Yohannes et al., 2020).



Mendeteksi objek manusia melalui kamera yang dipasang pada area public dengan menggunakan Convolutional Neural Network (CNN) dan You Only Look Once (YOLO) sebagai pendeteksi objek pada kotak pembatas dalam bingkai video. Objek disini adalah siswa, yang mana data seorang siswa yang berada pada ruang kelas dalam kotak pembatas akan terdeteksi dalam gambar. You Only Look Once (YOLO) adalah sebuah algoritma yang dikembangkan untuk mendeteksi objek secara real-time. Sistem pengenalan wajah merupakan hal menarik dalam dunia pendidikan, karena memadukan pengolahan citra dengan presensi siswa. Sistem presensi tersebut ialah sistem yang memadukan berbagai algoritma yang berkaitan dengan wajah seseorang, serta membutuhkan pemrosesan informasi secara kognitif yang berkaitan dengan karakter visual.

Berdasarkan pada uraian tersebut maka penelitian ini bertujuan untuk membangun sistem untuk Deteksi Wajah, dalam hal itu untuk mempermudah mengenali biodata siswa. Convolutional Neural Network (CNN) dan You Only Look Once (YOLO) menjadi algoritma yang digunakan pada penelitian ini, Dalam penelitian ini juga menggunakan library OpenCV yang merupakan salah satu pemrograman pada computer vision. Dimana akan dilakukan pencocokan objek melalui sebuah kamera yang menangkap gambar didalam ruangan, dengan data objek yang sudah tersimpan didalam dataset.

LANDASAN TEORI

Artificial Intelligence (AI)

Kecerdasan buatan atau *Artificial Intelligence (AI)* adalah dapat memahami bagaimana beberapa fakta saling berhubungan untuk membentuk pengetahuan dan merepresentasikan pengetahuan itu dalam bentuk mesin yang dapat dimengerti (Donald R. Franceschetti, 2018). Kecerdasan buatan mengacu pada pengembangan program yang menunjukkan kecerdasan tingkah laku atau bertindak seperti seorang ahli. Dalam konteks AI, terdapat beberapa konsep penting seperti machine learning (pembelajaran mesin), neural network (jaringan saraf tiruan), natural language processing (pemrosesan bahasa alami) dan banyak lagi. (Hau et al. 2020) Pengembangan AI telah memberikan dampak besar dalam berbagai bidang seperti pengenalan suara, pengenalan wajah, mobil otonom, pengobatan, dan masih banyak lagi (Géron, 2019; Eriana & Zein, 2021). Teknologi AI sudah banyak diterapkan di Indonesia mulai dari perkantoran, rumah sakit, hotel dan lainnya. *Artificial Intelligence* memiliki cabang ilmu atau turunan diantaranya pembelajaran mesin (*Machine Learning*) dengan kemampuan untuk belajar tanpa diprogram secara eksplisit dan pembelajaran mendalam (*Deep Learning*) yang merupakan sub cabang dari pembelajaran mesin. Ilmu cabang tersebut dapat diimplementasikan ke dalam perangkat komputer dan bisa membantu pekerjaan manusia.

Algoritma dan Machine Learning

Algoritma adalah sekumpulan langkah yang harus diikuti untuk menyelesaikan masalah tertentu yang menjelaskan metode untuk mencapai sebuah tujuan dengan membuat daftar setiap langkah yang harus dilakukan selama proses (Donald R. Franceschetti, 2018). Algoritma digunakan untuk memecahkan masalah komputasi yang ditentukan dengan baik.

Machine learning adalah cara membangun kecerdasan ke dalam mesin agar dapat belajar dari waktu ke waktu dan bekerja dengan baik menggunakan pengalamannya sendiri (Gollapudi, 2019). Hal itu berkaitan dengan pencarian pola yang menyaring detail yang relevan dari lingkungan. Algoritma ini dapat membangun kecerdasan yang bertujuan untuk menghasilkan hasil berupa suatu aturan yang akurat secara maksimal. Teknologi *machine learning* merupakan program yang dilatih dengan pemberian data yang

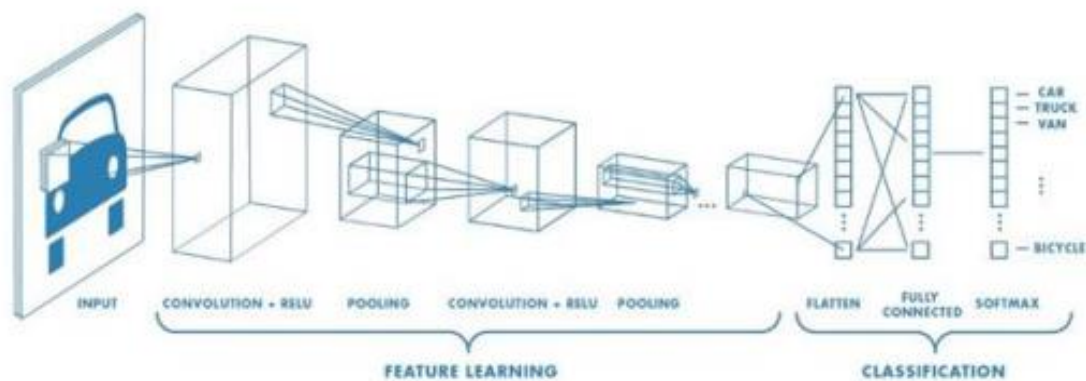


sudah disimpan ke dalam *dataset*. Sehingga kinerja mesin tersebut akan meningkat dan memiliki pengalaman dikarenakan banyaknya data yang menjadi sampel pelatihan pada program tersebut (Géron, 2019).

METODE PENELITIAN

Convolutional Neural Network (CNN)

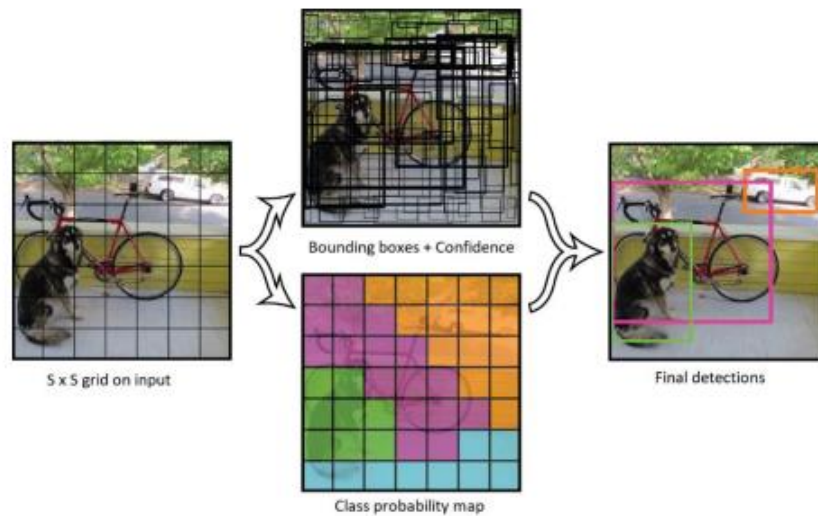
Algoritma *Convolutional Neural Network* (CNN) adalah jaringan neural yang digunakan untuk mengklasifikasikan sebuah citra dalam video dan juga dapat mendeteksi objek (Dadhich, 2018). CNN memasukan data atau informasi dalam bentuk yang lebih kecil dan menggabungkannya menggunakan jaringan dalam. Proses tersebut terjadi di beberapa lapisan CNN dan masing-masing lapisan terdiri dari neuron yang terhubung ke wilayah lokal dari lapisan sebelumnya (Fernando et al., 2019).



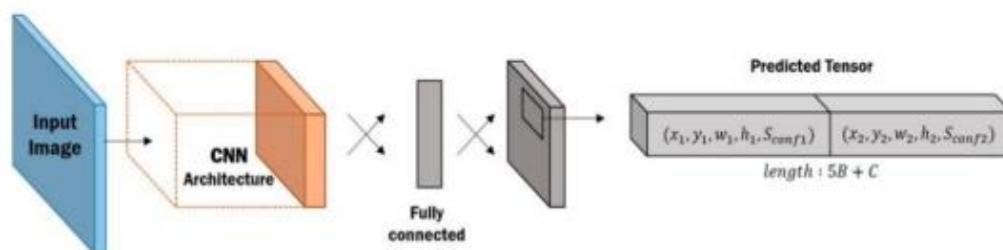
Gambar 1. Proses *Convolutional Neural Network* (CNN)

You Only Look Once (YOLO)

You Only Look Once (YOLO) adalah salah satu metode yang digunakan oleh jaringan saraf tiruan untuk mendeteksi sebuah objek yang berbeda dalam bentuk gambar ataupun video, baik berupa objek 3D ataupun 2D, seperti halnya sebuah kursi, meja, kendaraan mobil, motor, hewan, dll (Michelucci, 2019). Yolo merupakan jaringan yang hampir mirip dengan *Convolutional Neural Network* (CNN), karena di dalamnya terdapat proses konvolusi layer dan maxpooling, maka bisa dibilang yolo berbasis CNN yang dapat mendeteksi dimana objek berada, mengklasifikasikan beberapa objek dalam satu tahap secara realtimedengan menggunakan kotak pembatas atau bounding boxes. Yolo ditemukan pada tahun 2015 oleh Joseph Redmon, dkk dan memiliki beberapa versi diantaranya yolo v1, yolo v2, yolo v3, yolo v4, yolo v5 dan yang terbaru yolo v8.



Gambar 2. Ilustrasi Cara Kerja YOLO



Gambar 3. Arsitektur YOLO

Analisa Dan Perancangan

1. Analisa Sistem Berjalan

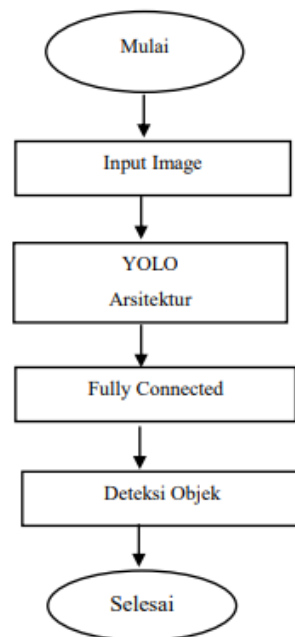
Pada tahap ini dilakukan analisa pada sistem yang berjalan sebelumnya yang diterapkan oleh SDN Salemban II Kabupaten Tangerang yaitu absensi siswa yang masih menggunakan cara manual.

2. Analisis Sistem Usulan

Pada tahap ini dilakukan analisa sistem yang nantinya akan dipakai untuk deteksi wajah, Langkah ini nantinya akan efektif jika dilakukan, mengingat tujuan dari penelitian ini untuk medeteksi wajah dan membandingkan antara algoritma *Convolutional Neural Network* (CNN) dan *You Only Look Once* (YOLO) dalam mendeteksi wajah.

3. Perancangan Sistem Identifikasi Obyek Menggunakan Algoritma YOLO

Perancangan sistem dilakukan untuk mengetahui bagaimana alur pada sistem yang akan dibuat, serta memberikan gambaran untuk *You Only Look Once* (YOLO) yang akan dituangkan secara sederhana dalam proses pendeteksian objek.



Gambar 4. Diagram Alir Sistem YOLO

4. Penerapan *You Only Look Once* (YOLO)

You Only Look Once (YOLO) digunakan untuk mendeteksi objek dengan cara mengkonvolusi beberapa layer, lalu menyisipkan layer pooling diantara lapisan konvolusi, kemudian masuk ke bagian peta baru atau feature map yang membentuk sebuah vector pada *fully conncted*. Yang kemudian dilakukan prediksi sebuah gambar setelah proses *fully connected*.

Tabel 1. Proses YOLO

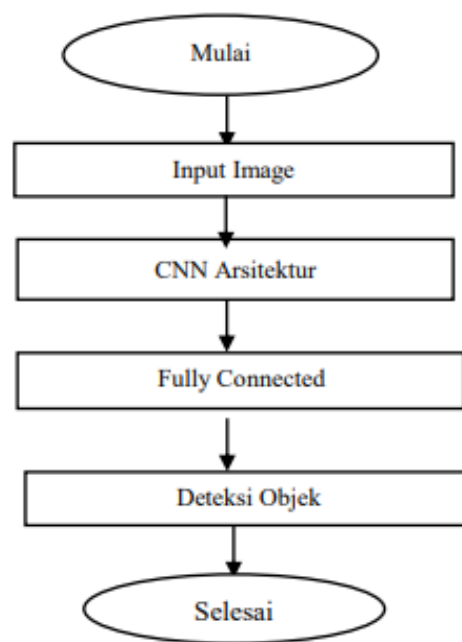
<i>Layer</i>	<i>Kernel</i>	<i>Stride</i>	<i>Output shape</i>
<i>input</i>			(416, 416, 3)
<i>Convolution</i>	3 x 3	1	(416, 416, 16)
<i>Max Pooling</i>	2 x 2	2	(208, 208, 16)
<i>Convolution</i>	3 x 3	1	(208, 208, 32)
<i>Max Pooling</i>	2 x 2	2	(104, 104, 32)
<i>Convolution</i>	3 x 3	1	(104, 104, 64)
<i>Max Pooling</i>	2 x 2	2	(52, 52, 64)
<i>Convolution</i>	3 x 3	1	(52, 52, 128)
<i>Max Pooling</i>	2 x 2	2	(26, 26, 128)
<i>Convolution</i>	3 x 3	1	(26, 26, 256)
<i>Max Pooling</i>	2 x 2	2	(13, 13, 256)
<i>Convolution</i>	3 x 3	1	(13, 13, 512)



<i>Max Pooling</i>	2 x 2	1	(13, 13, 512)
<i>Convolution</i>	3 x 3	1	(13, 13, 1024)
<i>Convolution</i>	3 x 3	1	(13, 13, 1024)
<i>Convolution</i>	1 x 1	1	(13, 13, 125)

5. Penerapan Sistem Identifikasi Obyek Menggunakan Algoritma CNN

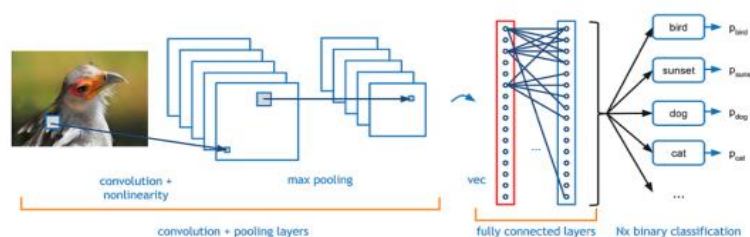
Perancangan sistem ini untuk mengetahui bagaimana alur pada sistem yang akan dibuat, serta memberikan gambaran untuk Algoritma CNN yang akan dituangkan secara sederhana dalam proses pendeteksian objek.



Gambar 5. Diagram Alir Sistem CNN

6. Penerapan CNN

Kemampuan utama CNN adalah arsitektur yang mampu mengenali informasi prediktif suatu objek (gambar, teks, potongan suara, dan sebagainya) walaupun objek tersebut dapat diposisikan dimana saja pada input. Kontribusi CNN di feature extraction layer adalah pada convolution dan pooling layer. Convolution bekerja dengan prinsip sliding window dan weight sharing (mengurangi kompleksitas perhitungan).



Gambar 6. Struktur CNN



7. Analisis Kebutuhan Sistem

Pada tahap ini mengenai analisis kebutuhan ini sangat penting tentunya untuk membantu berjalanya penelitian ini, kebutuhan apa saja yang diperlukan dalam pembuatan sistem akan dianalisis didalam kebutuhan perangkat keras (Hardware) dan kebutuhan perangkat lunak (Software).

Tabel 2. Kebutuhan Perangkat Keras

No	Perangkat Keras	Keterangan
1.	1 unit Laptop	Perangkat komputer yang digunakan dalam pembuatan sistem.
2.	1 unit camera laptop	Perangkat utama untuk deteksi scan wajah untuk sistem.

Tabel 3. Kebutuhan Perangkat Lunak

No	Perangkat Lunak	Keterangan
1.	<i>Phyton</i>	Sebagai <i>Bahasa pemrograman</i> yang digunakan untuk membuat <i>source code</i> .
2.	<i>Xampp, Visual Studio Code</i>	Software pelengkap.

Implementasi Dan Pengujian

1. Implementasi

Implementasi adalah proses penerapan rencana, kebijakan, atau desain ke dalam tindakan konkret dalam rangka mencapai tujuan tertentu. Dalam konteks proyek atau pengembangan perangkat lunak, implementasi mencakup pengkodean, pengujian, dan peluncuran sistem yang telah dirancang, serta memastikan bahwa semua komponen berfungsi sebagaimana mestinya dalam lingkungan operasional yang sebenarnya.

2. Pengumpulan Dataset

Dataset yang di kumpulkan secara manual yaitu mengambil gambar menggunakan kamera smartphone supaya hasil gambar nya jernih dan bagus.

**Tabel 4.** Pembagian Dataset

No	Dataset	Jumlah Data	Keterangan
1.	Data <i>Training</i>	604	JPG/JPEG
2.	Data Validasi	59	JPG/JPEG
3.	Data <i>Testing</i>	30	JPG/JPEG
Total		693	JPG/JPEG

3. Training Data

Pada tahap ini adalah proses dimana data yang sudah disiapkan pada web roboflow akan di training menggunakan Google Colab secara online atau harus terkoneksi pada internet. Google colab berkaitan dengan akun google dan untuk menggunakannya maka harus login terlebih dahulu. Pada google colab itu sangat terbatas limitnya yaitu memiliki waktu 12 jam saja dalam penggunaannya. Google colab memiliki salah satu kelebihan yaitu dengan adanya fitur Graphics Processing Unit (GPU) secara gratis, yang mana fitur ini dapat membantu proses training lebih cepat dari pada menggunakan Central Processing Unit (CPU) yang membutuhkan waktu cukup lama pada prosesnya. Semakin banyak data yang digunakan maka semakin lama pula prosesnya jika menggunakan CPU dalam training nya.

4. Testing Data

Pada proses ini adalah hasil prediksi pada sebuah gambar objek wajah yang menyertakan id, label, kotak pembatas dan nilai confidence nya dari hasil proses training data dan validasi data, yang masuk pada folder output.

**Gambar 7.** Hasil Gambar Testing

5. Rangkaian Alat

Pada proses ini yaitu rangkaian alat yang digunakan untuk melakukan pendeteksian pada objek wajah. Rangkaian alat di buat sebagai bentuk gambaran untuk komponen-komponen apa saja yang di butuhkan untuk sistem. Perangkat yang digunakan yaitu camera handphone dan Laptop untuk menangkap gambar nya.



6. Hasil Uji Coba

Pada uji coba yang dilakukan yaitu didalam dan diluar ruangan, Dimana objek wajah yang terdeteksi akan muncul kotak pembatas berwarna hijau, lalu objek wajah yang terdeteksi.. Hasil uji coba dilakukan dengan mendeteksi objek wajah manusia yang ditangkap pada kamera dengan objek benda lainnya, supaya mengetahui apakah sistem mendeteksi objek manusia saja atau tidak. Objek yang dideteksi tidak hanya wajah tegak lurus tetapi raut wajah tersenyum atau tidak tersenyum, setelah mendeteksi wajah kita juga bisa membedakan keakuratan antara algoritma *Convolutional Neural Network* (CNN) dan algoritma *You Only Look Once* (YOLO).

7. Hasil Akurasi algoritma YOLO

```
0: 320x320 1 0, 1 1, 57.0ms
Speed: 0.6ms preprocess, 57.0ms inference, 0.3ms postprocess pe
TP : 635.0 FP : 45.0 FN : 45.0
```

Gambar 8. Hasil *Confusion matrix* YOLO

Hasil dari pengujian training antara lain nilai True Positive (TP) sebesar 635, nilai False Negative (FN) sebesar 45, nilai False Positive (FP) sebesar 45 dan didapatkan hasil accuracy 0.8759, (87.59%). Dari beberapa nilai confusion matrix tersebut, dapat dihitung nilai accuracy dengan perhitungan matematika seperti berikut.

$$\begin{aligned} \text{Accuracy} &= (TP + TN) / (TP + FP + FN + TN) \\ &= (635 + 0) / (635 + 45 + 45 + 0) \\ &= 0.8759 \times 100 \\ &= 87.59\% \end{aligned}$$

8. Hasil Akurasi algoritma CNN

```
TP : 330.0 FP : 375.0 FN : 375.0
```

Gambar 9. Hasil *Confusion matrix* CNN

Hasil dari pengujian training antara lain nilai True Positive (TP) sebesar 330, nilai False Negative (FN) sebesar 375, nilai False Positive (FP) sebesar 375, dan didapatkan hasil accuracy 0.30 (30%). Dari beberapa nilai confusion matrix tersebut, dapat dihitung nilai accuracy dengan perhitungan matematika seperti berikut.

$$\begin{aligned} \text{Accuracy} &= (TP + TN) / (TP + FP + FN + TN) \\ &= (330 + 0) / (330 + 375 + 375 + 0) \\ &= 0.30 \times 100 \\ &= 30\% \end{aligned}$$



KESIMPULAN

Berdasarkan hasil penelitian untuk mendeteksi wajah menggunakan *You Only Look Once* (YOLO) dan *Convolutional Neural Network* (CNN) memerlukan pengumpulan dataset, dataset memiliki pembagian diantaranya Data Training, Data Validasi, dan Data Testing. Pada hasil uji coba yang telah dilakukan sistem ini memerlukan spesifikasi Hardware yang sangat tinggi sehingga dapat mendeteksi sebuah objek manusia sesuai dengan kotak pembatas dengan posisi penempatan kamera berada diatas. Dan perhitungan algoritma CNN dalam pengujian sistem ini sangat efektif dimana didukung oleh algoritma YOLO. Dan dari hasil pengujian ini dapat membantu sekolah SDN Salemban II dalam mendeteksi wajah siswa di sekolah. Perbedaan akurasi antara *You Only Look Once* (YOLO) dan *Convolutional Neural Network* (CNN) dapat dilihat dengan hasil nilai tersebut menggunakan confusion matrix, yang dapat dihitung nilai accuracy dengan perhitungan matematika sebagai berikut:

$$\text{Accuracy} = (TP + TN) / (TP + FP + FN + TN).$$

DAFTAR PUSTAKA

- Abadi, S., Jatiningrum, C., Riki. (2022). Integrating Analysis of Quality Management of Higher Education: Analytical Hierarchy Process and Multiple Regression Linear. *TEch-E.5(2)*, 150-162
- Adnan, M. H. M., Abadi, S. Jatiningrum, C. Hanafi, H. F, Mohd Zulkefli, N.A., Dewi, D. A. (2024). Multi-Criteria Decision-Making in SMEs Venture Capital Allocation: Analytic Hierarchy Process and Double Auction Approach. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 80-90.
- Dadhich, A. (2018). *Practical Computer Vision: Extract insightful information from images using TensorFlow, Keras, and OpenCV*. Packt Publishing.
- Donald R. Franceschetti, P. (2018). *Principles of Robotics & Artificial Intelligence* (P. Donald R. Franceschetti (ed.)). Grey House Publishing.
- Fernando, E., Andwiyani, D., Fitria Murad, D., Touriano, D., & Irsan, M. (2019). Face recognition system Using Deep Neural Network with Convolutional Neural Networks. *Journal of Physics: Conference Series*, 1235(1). <https://doi.org/10.1088/1742-6596/1235/1/012004>
- Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems* (N. Tache (ed.)). O'Reilly Media.
- Gollapudi, S. (2019). *Learn Computer Vision Using OpenCV With Deep Learning CNNs and RNNs*. Apress.
- Halfacree, G. (2019). *THE OFFICIAL Raspberry Pi Beginner's Guide How to use your new computer*. Raspberry Pi Trading Ltd.,
- Hau, Y. C., Baharuddin, M. Z., Yussof, S., & Dzulkifly, S. (2020). Social Distancing Detection with Deep Learning Model. *2020 8th International Conference on Information Technology and Multimedia, ICIMU*
- Hasma, Y. A., & Silfianti, W. (2018). Implementasi Deep Learning Menggunakan Framework Tensorflow Dengan Metode Faster Regional Convolutional Neural Network Untuk Pendeteksian Jerawat. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 23(2), 89–102. <https://doi.org/10.35760/tr.2018.v23i2.2459>



- Michelucci, U. (2019). *Advanced Applied Deep Learning Convolutional Neural Networks and Object Detection*. Apress. <https://doi.org/https://doi.org/10.1007/978-1-4842-4976-5>
- Pradita, F., Cahya, E., Nugroho, A. H., & Darmojo, H. S. (2019). *Sistem Computer Based Testing (CBT) Pegawai Di PT . Chingluh Indonesia Berbasis Web*. 7(1), 13–21.
- Yohannes, Y., Devella, S., & Arianto, K. (2020). Deteksi Penyakit Malaria Menggunakan Convolutional Neural Network Berbasis Saliency. *JUITA: Jurnal Informatika*, 8(1), 37. <https://doi.org/10.30595/juita.v8i1.6671>