

MODEL DETEKSI MIKROALGA *SPIRULINA PLATENSIS* DAN *CHLORELLA VULGARIS* BERBASIS *CONVOLUTIONAL NEURAL NETWORK* YOLOV8

Zahra Cahya Ramadhani¹ dan Dimas Firmanda Al Riza^{2*}

^{1,2}Departemen Teknik Biosistem, Universitas Brawijaya, Jl. Veteran, Malang

¹Email: zahra.chvrmdn@gmail.com

^{2*}Email: dimasfirmanda@ub.ac.id

ABSTRAK

Mikroalga merupakan organisme mikroskopis bersel tunggal yang hidup di berbagai perairan. Mikroalga seperti *Spirulina platensis* dan *Chlorella vulgaris* berpotensi menjadi sumber bioenergi sehingga mulai banyak dikultivasi. Kultivasi ini umumnya masih melakukan pemantauan jumlah/kepadatan sel mikroalga secara manual menggunakan hemocytometer yang lebih lama dan rentan *human error*. Penelitian ini bertujuan untuk mengembangkan model deteksi mikroalga *Spirulina platensis* dan *Chlorella vulgaris* berbasis citra mikroskopis dan *Convolutional Neural Network* menggunakan YOLOv8. Metodologi penelitian mencakup persiapan sampel (pengenceran dan pengukuran *optical density*), penentuan kepadatan terbaik, akuisisi citra, anotasi citra, pembuatan *dataset* citra, pelatihan model YOLOv8, dan evaluasi kinerja model. Penentuan kepadatan terbaik bertujuan untuk mendapatkan citra mikroskopis yang baik. Akuisisi citra dilakukan menggunakan mikroskop binokuler dan menghasilkan 560 gambar yang kemudian dianotasi. Model YOLOv8n, YOLOv8s, dan YOLOv8m dilatih dengan *default hyperparameter* di Google *Colaboratory* untuk mengetahui pengaruh augmentasi terhadap akurasi model. Evaluasi kinerja model dilakukan pada model YOLOv8 terpilih dan dianalisis nilai mAP50. Hasil penelitian menunjukkan bahwa augmentasi (*crop*, *brightness*, dan *blur*) menghasilkan mAP *train* dan *test* tertinggi pada model YOLOv8m, yakni 0,945 dan 0,913. Model YOLOv8m ini dilatih kembali dengan variasi *hyperparameters* dan didapatkan konfigurasi terbaik pada *optimizer* SGD, *epoch* 50, dan *learning rate* 0,01 dengan mAP *train* dan *test* sebesar 0,934 dan 0,925. Namun, *training* 29 *epoch* dapat menghasilkan akurasi 0,8535 yang memperkecil *overfitting* serta pemborosan sumber daya. Kesimpulannya, penelitian ini dapat mempermudah peneliti maupun industri dalam melakukan penghitungan jumlah mikroalga secara otomatis dan lebih efisien.

Kata kunci: *deep learning*, deteksi objek, mikroalga, YOLOv8

ABSTRACT

Microalgae are unicellular microscopic organisms that live in various water. Microalgae such as *Spirulina platensis* and *Chlorella vulgaris* are grown due to their potential as bioenergy source. During cultivation, typically, hemocytometers are used to manually count the cells and that is time-consuming and prone to human error. This research aims to develop microalgae detection model based on microscopic images and Convolutional Neural Network using YOLOv8 architecture. The methodology includes sample preparation (dilution and optical density measurement), best density determination, image acquisition, annotation, creation of *datasets*, YOLOv8 model training, and model performance evaluation. Best density determines good microscopic images. Image acquisition was done using binocular microscope and acquired 560 images which were then annotated. The YOLOv8n, YOLOv8s, and YOLOv8m models were trained using default hyperparameters on Google Collaboratory to determine the augmentation effect on model accuracy. Model performance evaluation was done on selected YOLOv8 models. The results showed the augmentation (*crop*, *brightness*, *blur*) get the highest mAP *train* and *test* on YOLOv8m model, which are 0.945 and 0.913. The YOLOv8m model was retrained with various hyperparameters and it was found that the best configuration was SGD optimizer, epoch 50, and learning rate 0.01 with mAP *train* and *test* are 0.934 and 0.925. However, 29 epochs yielding a better model with accuracy of 0.8535, minimising *overfitting* and resource wastage. This research can facilitate the more efficient and automatic counting for microalgae-related research and industry.

Keywords: *deep learning*, object detection, microalgae, YOLOv8

1. PENDAHULUAN

Mikroalga merupakan organisme mikroskopis uniseluler yang hidup di perairan tawar maupun perairan laut [1]. Dalam bidang teknik pertanian, salah satu pemanfaatan mikroalga dapat difokuskan

*) Penulis Korespondensi

Dikirim : 02 April 2024

Diterima : 28 Mei 2024

Publikasi *Online* : 31 Oktober 2024

sebagai bahan baku produksi sumber energi alternatif (bioenergi) [2]. Biomassa mikroalga menghasilkan produk *biofuel* seperti metana, sin-gas, hidrogen, biogas, bioetanol, biodiesel, bahan bakar pesawat, biokarbon, *bio-oil*, dan lainnya. Kandungan energi *biofuel* dari mikroalga mencapai 35.800 kJ/kg untuk minyak mentah (*crude oil*), 38.100 kJ/kg untuk *bio-oil*, dan 39.900 kJ/m³ untuk biogas [3]. Mikroalga seperti *Spirulina platensis* memiliki kandungan karbohidrat yang cukup tinggi sehingga berpotensi menjadi bahan baku pembuatan bioetanol melalui proses sakarifikasi dan fermentasi [4]. Mikroalga lain seperti *Chlorella vulgaris* juga dapat dijadikan sebagai bahan baku bionergi seperti biodiesel karena mampu mengakumulasi 30,3% lipid (berdasarkan berat kering) dalam komposisinya [5]. Kedua mikroalga ini memiliki berbagai potensi pengembangan bioproduk sehingga keduanya paling banyak dikenal dan mulai dikultivasi secara luas oleh industri mikroalga.

Dalam kultivasi mikroalga, salah satu hal yang menjadi kegiatan rutin adalah penghitungan jumlah mikroalga. Penghitungan jumlah/kepadatan sel mikroalga dapat memberikan informasi terkait pertumbuhan dan kualitas mikroalga. Terdapat beberapa metode penghitungan jumlah mikroalga, mulai dari sistem manual hingga otomatis menggunakan *artificial intelligence*. Pada metode manual, pengamat menghitung jumlah mikroalga dengan hemositometer di mikroskop [6]. Namun, metode ini memakan waktu yang lama dan rentan *human error*. Metode lain berupa pengukuran *Optical Density* (OD) menggunakan spektrofotometer menawarkan waktu yang lebih cepat, tetapi belum mampu menghitung jumlah mikroalga dan sensitif terhadap partikel lain yang dapat mengganggu pengukuran. Oleh karena itu, metode otomatis mulai dikembangkan menggunakan pengolahan citra (*image processing*). Contohnya, penghitungan jumlah mikroalga dalam waktu singkat yang dilakukan menggunakan *image processing* dengan *software* ImageJ [7].

Teknologi *automatic counting* juga dapat direalisasikan menggunakan *deep learning* melalui pembuatan model deteksi objek, dimana citra digunakan sebagai masukan (*input*). Salah satu algoritma *deep learning* yang umum diterapkan ialah *Convolutional Neural Network* (CNN) yang dinilai memiliki kinerja sangat baik dalam deteksi objek [8]. Terdapat penelitian yang melakukan klasifikasi dan penghitungan lima spesies *cyanobacteria* menggunakan CNN dengan mengekstraksi morfologi spesies alga dari citra mikroskop [9]. Beberapa model arsitektur CNN yang populer antara lain *Region-based Convolutional Neural Network* (R-CNN), *Mask R-CNN*, *Single-Shot Detector* (SSD), *You Only Look Once* (YOLO), dan lain-lain. Namun, *Mask R-CNN* memiliki beban komputasi yang besar karena harus melalui dua tahap prediksi pada proses pembelajaran dan melakukan segmentasi objek citra yang lebih spesifik [10]. Di sisi lain, jenis model pendeteksi satu tahap seperti YOLO telah mengalami peningkatan kinerja yang pesat. Pada bulan Januari 2023, arsitektur YOLO generasi kedelapan (YOLOv8) telah dirilis dengan kinerja yang lebih baik dari generasi atau versi sebelumnya [11].

Penghitungan jumlah mikroalga secara otomatis menjadi salah satu penunjang yang dapat memberikan kemudahan dan efektivitas dalam pemantauan kultivasi mikroalga. Meskipun telah banyak metode yang dikembangkan, beberapa kendala masih muncul, seperti perbedaan bentuk dan ukuran sel mikroalga yang beragam. Oleh karena itu, teknologi *deep learning* berupa CNN dapat menjadi solusi efektif dalam pembuatan model deteksi objek mikroalga secara otomatis dengan akurasi yang lebih tinggi dan waktu yang lebih cepat. Pada penelitian ini, akan dilakukan pemodelan identifikasi dan penghitungan jumlah mikroalga *Spirulina platensis* dan *Chlorella vulgaris* secara otomatis berbasis citra mikroskop dan CNN. Penelitian ini diharapkan dapat memberikan kontribusi dalam kultivasi mikroalga dan membantu penelitian terkait mikroalga, serta menjadi dasar pengembangan sistem penghitungan otomatis mikroalga yang lebih baik di masa depan.

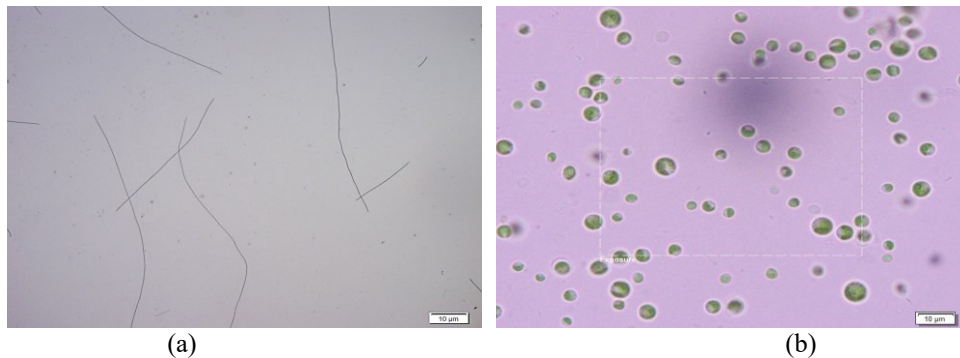
2. MATERI DAN METODE

Penelitian dilakukan dimulai dari bulan Juni hingga Desember 2023. Akuisisi data citra mikroalga dilakukan di Laboratorium Praktikum Dasar Biologi, Fakultas Teknologi Pertanian, Universitas Brawijaya. Adapun pembuatan dan evaluasi kinerja model deteksi mikroalga dilakukan secara fleksibel di rumah peneliti. Penelitian ini menggunakan metode penelitian eksperimental dan analisis citra dengan data latih (*training*), data validasi (*validation*), dan data uji (*testing*).

Akuisisi Data Citra

Mikroalga yang dijadikan sampel ialah *Spirulina platensis* dan *Chlorella vulgaris*. Perangkat yang digunakan dalam akuisisi citra adalah mikroskop binokuler yang terintegrasi dengan *software* cellSens Standard untuk membantu menangkap data gambar mikroalga *Chlorella vulgaris* pada perbesaran 1000x dan *Spirulina platensis* pada perbesaran 40x. *Exposure* pada kamera mikroskop diatur hingga berada di *region spot* 30% dan sensitivitas ISO 400. Masing-masing jenis mikroalga tersebut diambil gambarnya sebanyak 280 gambar sehingga diperoleh total *dataset* sebanyak 560 gambar yang akan digunakan sebagai

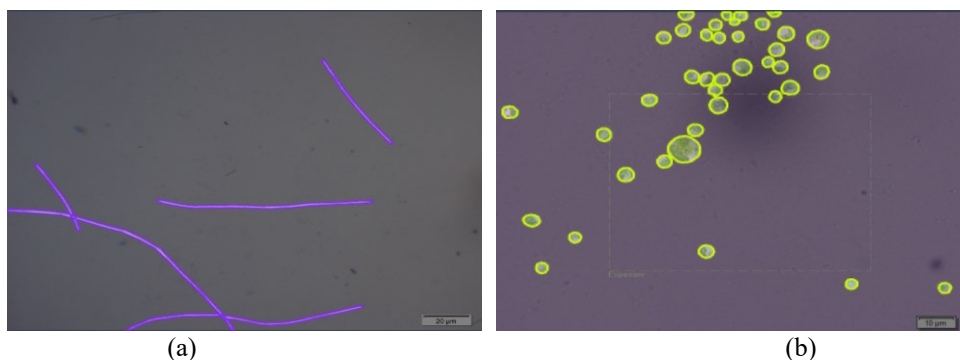
data primer pada pengumpulan data. Hasil akuisisi data citra mikroalga *Spirulina platensis* dan *Chlorella vulgaris* dapat dilihat pada Gambar 1.



Gambar 1. Hasil akuisisi citra pada mikroalga: (a) *Spirulina platensis*, (b) *Chlorella vulgaris*

Anotasi Data Citra

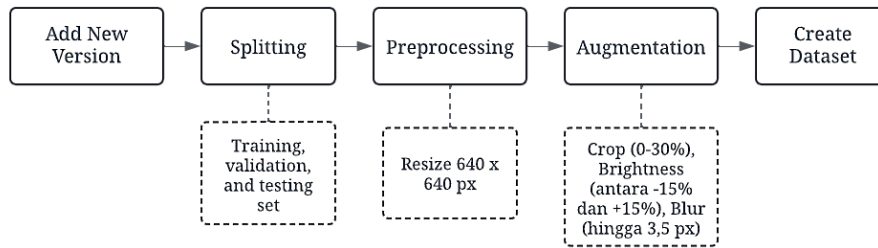
Anotasi data merupakan tahapan proses penambahan *label* berupa *mask* maupun *bounding box* pada objek mikroalga. Objek ditandai pada data gambar sesuai dengan morfologi mikroalga berupa bentuk dan warna yang kemudian diberi *class* berupa *Spirulina* dan *Chlorella*. Pada penelitian ini, jenis anotasi data yang digunakan ialah *instance segmentation* dengan bantuan *platform* Roboflow. Jenis anotasi tipe segmentasi menerapkan *polygon* untuk menandai permukaan mikroalga. Ketentuan anotasi yang baik adalah seluruh tepi objek mikroalga ditandai dengan *polygon* yang berjarak sangat dekat dengan tepi (mengikuti garis tepi objek). Dari tahap ini, dihasilkan sejumlah 560 gambar mikroalga yang telah dianotasi, yang kemudian akan diolah lebih lanjut untuk proses *training* menggunakan *framework* YOLOv8 segmentasi. Hasil anotasi data citra mikroalga *Spirulina platensis* dan *Chlorella vulgaris* dapat dilihat pada Gambar 2.



Gambar 2. Hasil anotasi citra pada mikroalga: (a) *Spirulina platensis*, (b) *Chlorella vulgaris*

Pembuatan Dataset Citra

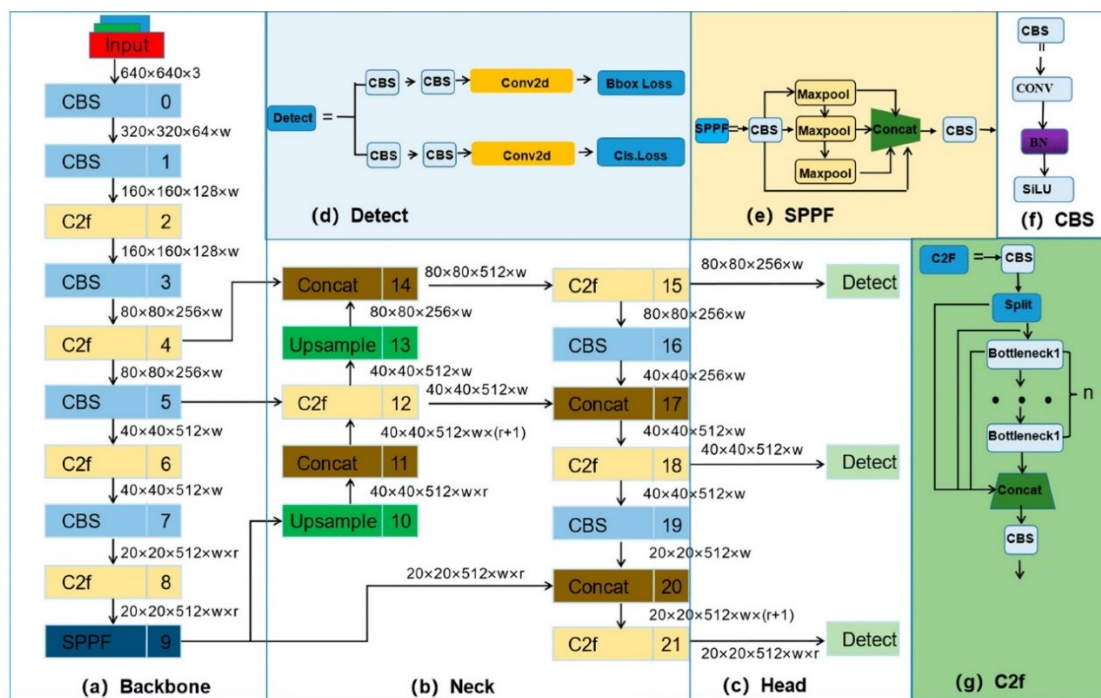
Pembuatan *dataset* citra dilakukan di *platform* Roboflow melalui fitur '*Generate Dataset*'. Seluruh data citra yang telah dianotasi akan dijadikan ke dalam sebuah *version dataset*. Dalam proses pembuatan *dataset*, terdapat beberapa tahap antara lain *train/test split*, *pre-processing (resize)*, dan augmentasi. *Splitting* dilakukan dengan proporsi 80:10:10 berturut-turut untuk data *training*, data validasi, dan data *testing*. *Resize* dilakukan dari 2448 x 1920 px menjadi 640 x 640 px dengan mempertimbangkan adanya objek kecil dalam gambar dan perlunya keseimbangan kinerja antara waktu dengan akurasi dalam proses deteksi. Ukuran ini memungkinkan model untuk digunakan pada *edge devices* tanpa kehilangan terlalu banyak informasi penting dari gambar [12]. Setelah itu, dilakukan augmentasi menggunakan tiga teknik, yakni *crop* (0-30%), *brightness* (antara -15% dan +15%), dan *blur* (hingga 3,5 px). Augmentasi ini hanya berlaku untuk data *training* sehingga jumlah *training dataset* menjadi 1344 gambar dan total seluruh *dataset* yang dihasilkan adalah 1456 gambar. *Dataset* yang dibuat dari Roboflow kemudian melakukan *import* melalui *download* code API Roboflow sehingga dapat terhubung oleh Google Colaboratory. Tahapan pembuatan *dataset* citra dapat dilihat pada Gambar 3.



Gambar 3. Tahapan pembuatan *dataset* citra

Pembuatan Model dan Evaluasi Kinerja

Dataset citra diolah menggunakan model segmentasi *deep learning* berupa *Convolutional Neural Network* YOLOv8. CNN terdiri dari beberapa lapisan, yakni lapisan konvolusi (*convolutional layer*), lapisan *sub-sampling* (*pooling layer*), dan *fully connected layer*. YOLOv8 terdiri dari empat komponen, yaitu *input*, *backbone*, *neck*, dan *output* (*head*). *Input* umumnya berupa gambar yang diaugmentasi mozaik ataupun penambahan *grayscale* adaptif [13]. Kemudian, *backbone* terdiri dari modul CBS (*Conv + BatchNorm + SiLU*), C2f (*Context to focus*), dan SPPF (*Spatial Pyramid Pooling Fusion*) yang secara keseluruhan akan mengekstrak fitur [14]. Modul CBS melakukan konvolusi pada *input*, menerapkan *batch normalization*, dan mengaktifkan aliran informasi dengan fungsi aktivasi SiLU. Modul C2f menggantikan modul C3 pada YOLOv5 untuk pembelajaran fitur residual, yang memperkaya aliran informasi dari jaringan ekstraksi fitur dengan bobot lebih ringan dari C3. Modul SPPF sama seperti pada YOLOv5, yang mengubah sembarang *feature map* menjadi vektor fitur berukuran tetap [15]. Sementara itu, *neck* mengadopsi gabungan antara FPN (*feature pyramid network*) dan PAN (*path aggregation network*) untuk menggabungkan fitur tingkat rendah menjadi representasi tingkat tinggi dengan skala bervariasi [13]. Terakhir, *output* (*head*) menunjukkan pemisahan antara deteksi dan klasifikasi. Setiap lapisan deteksi menghasilkan *predicted box* dan *confidence score* pada kategori yang sesuai [15]. Secara visual, struktur dari YOLOv8 [15] dapat dilihat di Gambar 4.



Gambar 4. Struktur YOLOv8

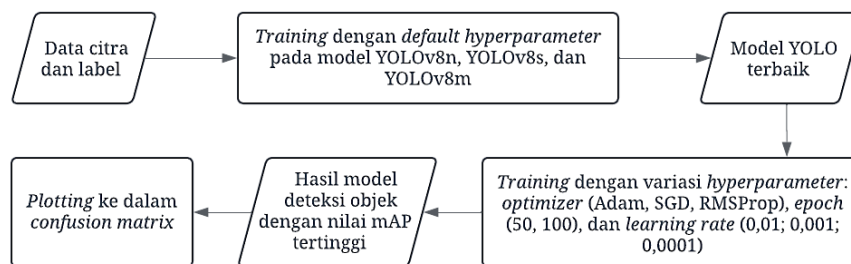
Proses pengolahan *dataset* dan analisis model dilakukan menggunakan Ultralytics YOLOv8 di Google Colaboratory. Model YOLOv8 dilatih menggunakan *framework* PyTorch. Sebelum *training*, dilakukan konfigurasi *hyperparameters* berupa *optimizer*, *epoch*, dan *learning rate*. Pada tahap awal, ketiga model YOLO yang berbeda (YOLOv8n, YOLOv8s, dan YOLOv8m) dilatih menggunakan *default hyperparameter* yaitu *optimizer* SGD, *epoch* 100, dan *learning rate* 0,01 yang diadaptasi dari penelitian mengenai deteksi beberapa mikroalga di perairan laut menggunakan YOLOv8 [16]. Model YOLOv8

dengan nilai *mean Average Precision* (mAP 50) tertinggi akan dipilih untuk selanjutnya dilatih menggunakan variasi *hyperparameter*. Variasi *hyperparameter* yang digunakan meliputi *optimizer* Adam, RMSProp, dan SGD dengan masing-masing variasi *epoch* adalah 50 dan 100 serta *learning rate* sebesar 0,01; 0,001; dan 0,0001. Hasil *training* dari tiap variasi *hyperparameter* akan dievaluasi melalui nilai mAP 50. Rumus untuk mencari nilai mAP 50 dapat dilihat pada persamaan 1 [11]:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (1)$$

dengan N = jumlah kelas atau kategori, dan AP_i = *average precision* pada kelas i.

Perbandingan difokuskan pada hasil mAP 50 *train*, *valid*, dan *test*. Dari evaluasi tersebut, akan didapatkan satu model terbaik dengan akurasi tertinggi. Model terbaik kemudian diproyeksikan ke *confusion matrix*. Secara ringkas, alur tahapan pembuatan model dan evaluasi kinerja dapat dilihat di Gambar 5.



Gambar 5. Tahapan pembuatan model dan evaluasi kinerja

3. HASIL DAN PEMBAHASAN

Pemilihan Jenis Model YOLOv8 Terbaik

Dataset teraugmentasi dilatih pada tiga model YOLO yang berbeda (YOLOv8n, YOLOv8s, YOLOv8m) menggunakan default *hyperparameter* dan dibandingkan nilai mAP 50 *train*, *validation*, dan *test*. Hasil akurasi tersebut dapat dilihat pada Tabel 1.

Tabel 1. Hasil akurasi ketiga model YOLOv8

Model	mAP Train	mAP Valid	mAP Test
YOLOv8n	0,901	0,841	0,894
YOLOv8s	0,936	0,844	0,912
YOLOv8m	0,945	0,866	0,913

Berdasarkan Tabel 1, model YOLOv8m menghasilkan nilai mAP *train* dan *test* tertinggi di antara model lainnya. Nilai mAP *train* tersebut juga lebih tinggi dibandingkan mAP *test*, di mana hal ini telah sesuai dengan pemahaman literatur yang menyatakan bahwa umumnya nilai *testing* lebih rendah dari nilai *training* karena model memprediksi objek menggunakan data dari dunia nyata. Hal ini didukung oleh sebuah penelitian yang menghasilkan model deteksi objek terbaik menggunakan model YOLOv8m, dibandingkan dengan YOLOv8n dan YOLOv8s [17].

Evaluasi Kinerja Model Terbaik Berdasarkan Konfigurasi *Hyperparameters*

Model YOLOv8m yang telah dilatih menggunakan default *hyperparameter* akan dilatih kembali menggunakan berbagai variasi *hyperparameter* untuk menentukan kinerja akhir terbaik dari model deteksi objek mikroalga. Hasil kinerja model dalam bentuk mAP berdasarkan variasi *hyperparameters* dapat dilihat pada Tabel 2.

Tabel 2. Hasil akurasi berdasarkan variasi *hyperparameters*

Optimizer	Epoch	Learning Rate	mAP Train	mAP Valid	mAP Test
Adam	50	0,01	0,835	0,822	0,842
		0,001	0,901	0,855	0,893
		0,0001	0,922	0,866	0,905
	100	0,01	0,863	0,841	0,859
		0,001	0,917	0,864	0,916
		0,0001	0,942	0,861	0,919
SGD	50	0,01	0,934	0,852	0,925

		0,001	0,880	0,85	0,900
		0,0001	0,840	0,833	0,843
	100	0,01	0,945	0,866	0,913
		0,001	0,919	0,859	0,920
		0,0001	0,855	0,843	0,871
RMSProp	50	0,01	0,420	0,425	0,418
		0,001	0,691	0,676	0,680
		0,0001	0,825	0,815	0,832
	100	0,01	0,427	0,431	0,427
		0,001	0,757	0,734	0,756
		0,0001	0,848	0,820	0,841

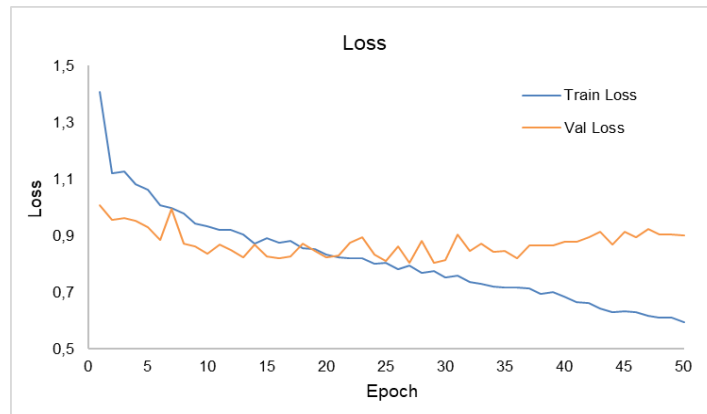
Berdasarkan Tabel 2, dapat dilihat adanya perbedaan nilai mAP di setiap variasi *hyperparameters*. Pada model yang dilatih menggunakan *Optimizer Adam*, nilai mAP cenderung naik seiring dengan semakin banyaknya *epoch* dan semakin kecilnya *learning rate*. Pada model yang dilatih menggunakan *Optimizer SGD*, nilai mAP cenderung menurun seiring dengan semakin kecilnya *learning rate*. Sementara itu, peningkatan *epoch* dari 50 menuju 100 mengakibatkan nilai mAP cenderung naik, kecuali nilai mAP *test* yang malah menurun. Pada model yang dilatih menggunakan *Optimizer RMSProp*, nilai mAP meningkat seiring dengan semakin banyaknya *epoch* dan semakin kecilnya *learning rate*.

Semakin banyak *epoch* yang digunakan, maka akurasi suatu model akan semakin tinggi [18]. Jika menggunakan *pre-trained model*, maka proses *training* menggunakan YOLOv8 dapat menghasilkan akurasi yang tinggi dengan jumlah *epoch* yang lebih sedikit [19]. Kemudian, semakin kecil nilai *learning rate*, maka akurasi model akan semakin tinggi, tetapi proses *training* akan menjadi semakin lama [20]. Selanjutnya, nilai mAP dianalisis berdasarkan kesesuaian tren nilai antara *training*, validasi, dan *testing*. Nilai mAP *train* umumnya paling tinggi, kemudian diikuti dengan mAP validasi, dan terakhir adalah mAP *test* yang memiliki nilai paling rendah.

Penentuan model deteksi objek mikroalga *Spirulina platensis* dan *Chlorella vulgaris* terbaik dilihat berdasarkan nilai mAP *train* dan mAP *test* tertinggi di antara model lainnya. Selain itu, model juga memiliki selisih nilai mAP *train* dan mAP *test* yang kecil. Berdasarkan Tabel 2, didapatkan bahwa setidaknya terdapat tiga model terbaik. Model pertama yaitu model yang dilatih dengan *hyperparameter* berupa *optimizer Adam*, *epoch* 100, dan *learning rate* 0,01. Model ini memperoleh mAP *train* sebesar 0,942 dan mAP *test* sebesar 0,919, dimana selisih nilai keduanya adalah 2,5%. Model kedua yaitu model yang dilatih dengan *optimizer SGD*, *epoch* 50, dan *learning rate* 0,01. Model ini memperoleh mAP *train* sebesar 0,934 dan mAP *test* sebesar 0,925, dimana selisih nilai keduanya adalah 0,97%. Model ketiga yaitu model yang dilatih dengan *optimizer SGD*, *epoch* 100, dan *learning rate* 0,01. Model ini memperoleh mAP *train* sebesar 0,945 dan mAP *test* sebesar 0,913, dimana selisih nilai keduanya adalah 3,5%.

Dengan demikian, di antara ketiga model tersebut, satu model YOLOv8m yang dipilih sebagai model terbaik adalah model yang dilatih menggunakan *optimizer SGD*, *epoch* 50, dan *learning rate* 0,01. Walaupun model ini memiliki nilai mAP *train* yang paling rendah di antara dua model lainnya, tetapi model ini memiliki nilai mAP *test* yang paling tinggi. Selain itu, selisih nilai mAP *train* dan mAP *test* pada model ini juga merupakan yang paling rendah, yakni hanya 0,97%. Nilai akurasi *training* dan validasi ataupun *testing* yang hampir sama menandakan bahwa model tersebut mampu melakukan generalisasi dengan baik, mampu mencocokkan hasil prediksinya dengan data *training* (yang telah dipelajari) maupun data validasi ataupun *testing* (yang belum dipelajari) [21]. Hal ini menunjukkan bahwa model deteksi objek YOLOv8m dalam penelitian ini berhasil mengatasi kompleksitas spesies mikroalga serta beradaptasi terhadap variasi fokus mikroskop.

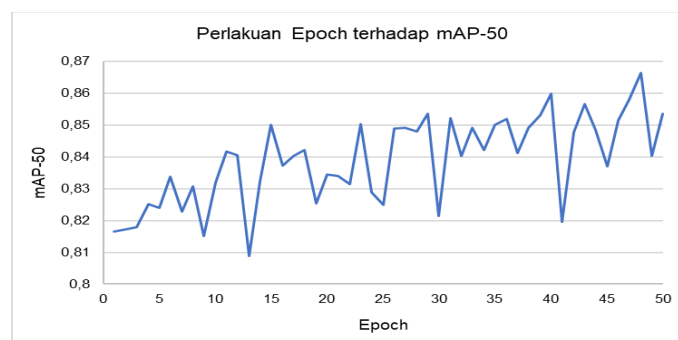
Selain nilai akurasi berupa mAP, hal penting yang perlu diperhatikan adalah nilai *train loss* dan *validation loss* yang diamati saat proses *training*. Proses pembelajaran dapat dikatakan selesai apabila kesalahan (*error* atau *loss*) antara nilai yang diprediksi (*predicted output*) dengan nilai yang sebenarnya (*actual output*) telah mencapai minimum [22]. Pada umumnya, *train loss* akan selalu turun untuk mencapai konvergensi. Sementara itu, *validation loss* juga mengalami penurunan, tetapi pada suatu titik tertentu, nilai *loss* akan cenderung meningkat hingga akhir. Peristiwa ini menunjukkan adanya *overfitting*. Grafik *train loss* dan *validation loss* pada model terbaik dapat dilihat pada Gambar 6.



Gambar 6. Grafik *train* dan *validation loss* pada model terbaik

Grafik *train* dan *validation loss* umumnya digunakan untuk melihat peristiwa *overfitting* yang mungkin saja terjadi pada model saat proses *training*. Jika *overfitting* mulai menunjukkan tanda-tandanya, maka proses *training* dapat dihentikan sebanyak *epoch* tertentu. Penentuan *epoch* yang optimal dapat dilakukan dengan memilih titik dimana garis *train loss* dan *validation loss* saling mendekati atau beririsan, tetapi tetap memperhatikan nilai *train loss* yang lebih rendah. Berdasarkan Gambar 6, dapat dilihat bahwa garis *train loss* semakin lama semakin menurun. Sementara itu, garis *validation loss* mulai mengalami peningkatan sejak *epoch* ke-27. Dengan mempertimbangkan nilai *train loss* yang lebih rendah, maka sebenarnya *epoch* yang optimal dapat ditentukan pada titik ke-29. Dengan demikian, proses *training* atau pembelajaran dapat diberhentikan pada saat memasuki *epoch* ke-29. Penambahan jumlah *epoch* hingga 50 hanya berdampak pada model yang semakin *overfitting* serta pemborosan waktu dan beban komputasi.

Untuk memastikan bahwa model benar-benar layak diselesaikan dalam 29 *epochs*, maka nilai akurasi berupa mAP perlu diperhatikan pula. Perlu diingat bahwa kriteria penentuan *epoch* yang optimal adalah tidak *overfitting* ataupun *underfitting*, memiliki nilai *train loss* yang rendah, dan ditunjang dengan nilai akurasi yang tinggi. Nilai mAP pada tiap *epoch* dapat dilihat melalui grafik. Grafik perubahan nilai mAP pada model terbaik seiring dengan peningkatan *epoch* dapat dilihat pada Gambar 7.



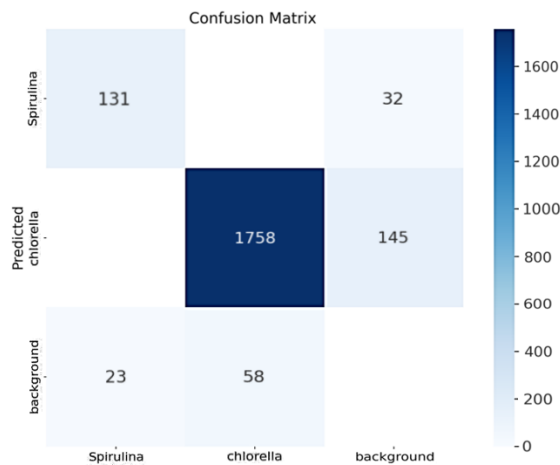
Gambar 7. Grafik perlakuan *epoch* terhadap mAP-50 pada model terbaik

Grafik yang ditampilkan pada Gambar 7 merupakan grafik mAP validasi. Umumnya, algoritma YOLO tidak menghitung dan mencatat nilai mAP *training* secara *default*. Hal ini disebabkan oleh kebutuhan komputasi yang besar, terutama saat melatih *dataset* besar sehingga dapat memperlambat proses *training* secara signifikan. Fokus utama selama *training* seringkali terletak pada nilai mAP set validasi untuk menilai kinerja model dalam melakukan generalisasi. Nilai mAP validasi menjadi indikator yang lebih diandalkan untuk menilai seberapa baik performa model pada data yang belum pernah dilihat. Namun, mAP pada *set training* dapat dihitung dengan cara memodifikasi *training script* atau *notebook* dengan komputasi yang mumpuni.

Berdasarkan Gambar 7, nilai mAP validasi pada *epoch* ke-29 bernilai 0,8535. Walaupun *epoch* ke-48 lebih menjanjikan dengan nilai sebesar 0,8664, tetapi model tetap mengalami *overfitting*. Selain itu, proses pembelajaran mulai dari *epoch* ke-29 menuju *epoch* ke-48 hanya menghasilkan selisih nilai mAP yang kecil, yakni 0,0129. Dengan demikian, *epoch* ke-29 lebih dipilih karena mampu mencegah *overfitting*, menghemat waktu pembelajaran dan beban komputasi, serta mampu menghasilkan nilai mAP yang cukup

baik. Hal ini sejalan dengan sebuah penelitian yang lebih memilih 1000 *epochs* dengan akurasi 74,59% dibandingkan dengan 3000 *epochs* dengan akurasi 79,58% [23]. Pemilihan ini didasarkan oleh kecilnya peningkatan mAP antara kedua nilai *epoch* tersebut, yakni hanya sebesar 5%. Dengan demikian, penelitian ini dapat menghasilkan model terbaik menggunakan model YOLOv8m dengan *optimizer* SGD, *epoch* 29, dan *learning rate* 0,01 yang mampu menghasilkan akurasi sebesar 0,8535.

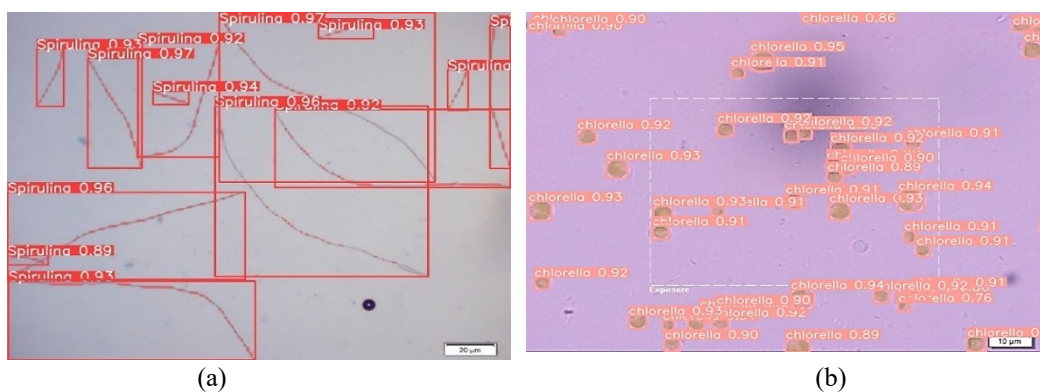
Hasil evaluasi model terbaik kemudian ditampilkan melalui *confusion matrix* yang dapat dilihat pada Gambar 8.



Gambar 8. *Confusion matrix* model terbaik

Hasil *confusion matrix* di Gambar 8 menunjukkan jumlah *instances* yang diprediksi oleh model terhadap *label* aslinya. Pada kategori *Spirulina*, model dapat memprediksi mikroalga *Spirulina platensis* dengan benar sebanyak 131 *instances* atau setara dengan 85%. Tidak ada objek *Spirulina platensis* yang terdeteksi sebagai *Chlorella vulgaris*. Namun, 23 *instances* (setara dengan 15%) tidak mampu diprediksi dengan benar dan terdeteksi sebagai *background*. Sementara itu, pada kategori *chlorella*, model dapat memprediksi objek mikroalga *Chlorella vulgaris* dengan benar sebanyak 1758 *instances* (97%). Tidak ada objek *Chlorella vulgaris* yang terdeteksi sebagai *Spirulina platensis*. Namun, sebanyak 58 *instances* (3%) tidak mampu diprediksi dengan benar dan terdeteksi sebagai *background*. Adanya fenomena ini dapat disebabkan oleh warna objek *Spirulina platensis* maupun *Chlorella vulgaris* yang hampir menyerupai warna *background* akibat kurangnya pengaturan titik fokus saat akuisisi citra menggunakan mikroskop. Selain itu, penyimpangan tersebut dapat pula disebabkan oleh proses anotasi (*labelling*) yang kurang rapi, seperti garis *polygon* yang masih jauh dari garis tepi objek sehingga masih terdapat celah warna *background* yang masuk dalam garis *polygon*. Hal ini pada akhirnya membuat model belum mampu mendeteksi objek tersebut dengan benar. Oleh karena itu, proses anotasi data menjadi tahapan paling penting dan perlu diperhatikan sebelum dilakukannya proses *training*.

Selain itu, telah dilakukan pula pengujian hasil deteksi menggunakan salah satu sampel data *test*. Hal ini bertujuan untuk mengetahui efektivitas hasil deteksi melalui banyaknya *bounding box* dan *mask* yang muncul sebagai penanda deteksi. Hasil deteksi mikroalga *Spirulina platensis* dan *Chlorella vulgaris* menggunakan model terbaik dapat dilihat pada Gambar 9.



Gambar 9. Hasil deteksi menggunakan model terbaik; (a) *Spirulina platensis*, (b) *Chlorella vulgaris*

Berdasarkan Gambar 9, dapat dilihat bahwa salah satu sampel data *test* berhasil dideteksi dengan baik menggunakan model terbaik. Seluruh mikroalga, baik *Spirulina platensis* maupun *Chlorella vulgaris* memiliki *bounding box*, *mask*, nama kelas (*Spirulina* ataupun *chlorella*), dan *confidence score*. *Bounding box* merupakan kotak yang melingkupi suatu objek. *Mask* merupakan *polygon* yang melingkupi suatu objek, dengan pola yang mengikuti bentuk objek tersebut. *Confidence score* menunjukkan probabilitas objek terdeteksi benar sesuai dengan nama kelas yang dipelajari saat proses *training*. Pada sampel data *test* untuk *Spirulina platensis*, model dapat melakukan penghitungan secara otomatis yang ditandai dengan keterangan jumlah *Spirulina* yang terdeteksi, yakni sebanyak 13 *Spirulina*. Sementara itu, sampel data *test* untuk *Chlorella vulgaris* juga berhasil dideteksi sejumlah 44 *chlorella*, dimana keterangan jumlah tersebut terdapat dalam model.

4. KESIMPULAN DAN SARAN

Model penghitungan otomatis pada mikroalga *Spirulina platensis* dan *Chlorella vulgaris* berbasis YOLOv8 dinilai memiliki kinerja terbaik apabila memiliki nilai mAP *train* dan mAP *test* tertinggi, selisih nilai mAP *train* dan mAP *test* terkecil, dan tidak *overfitting*. Model YOLOv8m terbaik berhasil memperoleh mAP *train* sebesar 0,934 dan mAP *test* sebesar 0,925, dimana selisih nilai keduanya hanyalah 0,97%. Akurasi ini dicapai melalui *hyperparameters* berupa *optimizer* SGD, *epoch* 50, dan *learning rate* 0,01. Namun, kinerja model lebih baik dilatih hanya dengan 29 *epoch* untuk mencegah *overfitting* serta pemborosan sumber daya. Akurasi yang dihasilkan masih cukup baik, yakni 0,8535. Penelitian ini masih dapat dikembangkan untuk penyempurnaan, terutama anotasi citra yang perlu dilakukan dengan teliti agar mampu mencapai potensi akurasi model yang lebih tinggi.

DAFTAR PUSTAKA

- [1] V. Dolganyuk, D. Belova, O. Babich, A. Prosekov, S. Ivanova, D. Katserov, N. Patyukov, and S. Sukhikh, "Microalgae: A promising source of valuable bioproducts," *Biomolecules*, vol. 10, no. 8, pp. 1-24, Aug. 2020, doi: <https://doi.org/10.3390/biom10081153>.
- [2] H. V. Mendonça, P. Assemany, M. Abreu, E. Couto, A. M. Maciel, R. L. Duarte, M. G. B. Santos, and A. Reis, "Microalgae in a global world: New solutions for old problems?," *Renewable Energy*, vol. 165, pp. 842-862, 2021, doi: [10.1016/j.renene.2020.11.014](https://doi.org/10.1016/j.renene.2020.11.014).
- [3] D. T. Zewdie and A. Y. Ali, "Cultivation of microalgae for biofuel production: Coupling with sugarcane-processing factories," *Energy, Sustainability and Society*, vol. 10, no. 1, pp. 1-16, 2020, doi: [10.1186/s13705-020-00262-5](https://doi.org/10.1186/s13705-020-00262-5).
- [4] V. S. Braga, J. B. Moreira, J. A. V. Costa, and M. G. Morais, "Enhancement of the carbohydrate content in *Spirulina* by applying CO₂, thermoelectric fly ashes and reduced nitrogen supply," *International Journal of Biological Macromolecules*, vol. 123, pp. 1241-1247, 2019, doi: <https://doi.org/10.1016/j.ijbiomac.2018.12.037>.
- [5] H. Chowdhury and B. R. Loganathan, "Third-generation biofuels from microalgae: A review," *Current Opinion in Green and Sustainable Chemistry*, vol. 20, pp. 39-44, 2019, doi: [10.1016/j.cogsc.2019.09.003](https://doi.org/10.1016/j.cogsc.2019.09.003).
- [6] T. Takahashi, "Potential of an automated- and image-based cell counter to accelerate microalgal research and applications," *Energies*, vol. 13, pp. 1-11, 2020, doi: <https://doi.org/10.3390/en13226019>.
- [7] V. E. Dökümcüoğlu and M. Yılmaz, "Assessment of cell counting method based on image processing for a microalga culture," *Mediterranean Fisheries and Aquaculture Research*, vol. 3, no. 1, pp. 75-81, 2020. [Online]. Available: <https://dergipark.org.tr/en/download/article-file/1128248>.
- [8] S. Patel and A. Patel, "Object detection with convolutional neural networks," *Machine Learning for Predictive Analysis Lecture Notes in Networks and Systems*, pp. 529-539, 2020. doi: https://doi.org/10.1007/978-981-15-7106-0_52.
- [9] S. S. Baek, J. C. Pyo, Y. Pachepskyb, Y. Park, M. Ligaray, C. Y. Ahn, Y. H. Kim, J. A. Chun, and K. H. Cho, "Identification and enumeration of *cyanobacteria* species using a deep neural network," *Ecological Indicators*, vol. 115, pp. 1-10, 2020. doi: <https://doi.org/10.1016/j.ecolind.2020.106395>.
- [10] X. Jiang, A. Hadid, Y. Pang, E. Granger, and X. Feng, *Deep learning in object detection and recognition*. Singapore: Springer Nature, 2019, doi: [http://dx.doi.org/10.1007/978-981-10-5152-4](https://doi.org/10.1007/978-981-10-5152-4).
- [11] E. Soylu and T. Soylu, "A performance comparison of YOLOv8 models for traffic sign detection in the Robotaxi-full scale autonomous vehicle competition," *Multimedia Tools and Applications*, pp. 1-31, 2023. doi: [10.1007/s11042-023-16451-1](https://doi.org/10.1007/s11042-023-16451-1).

- [12] X. Zhai, Z. Huang, T. Li, H. Liu, and S. Wang, "YOLO-drone: An optimized YOLOv8 network for tiny UAV object detection," *Electronics*, vol. 12, no. 17, pp. 2-21, 2023. doi: <https://doi.org/10.3390/electronics12173664>.
- [13] X. Wang, H. Gao, Z. Jia, and Z. Li, "BL-YOLOv8: An improved road defect detection model based on YOLOv8," *Sensors*, vol. 23, no. 20, pp. 1-22, 2023. doi: <https://doi.org/10.3390/s23208361>.
- [14] B. Luo, Z. Kou, C. Han, and J. Wu, "A "hardware-friendly" foreign object identification method for belt conveyors based on improved YOLOv8," *Applied Sciences*, vol. 13, pp. 1-24, 2023. doi: <https://doi.org/10.3390/app132011464>.
- [15] Y. Huang, W. Tan, L. Li, and L. Wu, "WFRE-YOLOv8s: A new type of defect detector for steel surfaces," *Coatings*, vol. 13, pp. 1-19, 2023. doi: <https://doi.org/10.3390/coatings13122011>.
- [16] S. Zhou, J. Jiang, X. Hong, P. Fu, and H. Yan, "Vision meets algae: A novel way for microalgae recognition and health monitor," *Front. Mar. Sci.*, pp. 1-10, 2023. doi: [10.3389/fmars.2023.1105545](https://doi.org/10.3389/fmars.2023.1105545).
- [17] T. Yang, S. Zhou, A. Xu, J. Ye, and J. Yin, "An approach for plant leaf image segmentation based on YOLOv8 and the improved DEEPLABv3+," *Plants*, vol. 12, no. 19, pp. 2-17, 2023. doi: <https://doi.org/10.3390/plants12193438>.
- [18] M. Wasil, H. Harianto, and F. Fathurrahman, "Pengaruh *epoch* pada akurasi menggunakan *convolutional neural network* untuk klasifikasi *fashion* dan *furniture*," *Infotek: Jurnal Informatika dan Teknologi*, vol. 5, no. 1, pp. 53-61, 2022. doi: <https://doi.org/10.29408/jit.v5i1.4393>.
- [19] L. Zhang, G. Ding, C. Li, and D. Li, "DCF-YOLOv8: An improved algorithm for aggregating low-level features to detect agricultural pests and diseases," *Agronomy*, vol. 13, no. 8, pp. 1-16, 2023. doi: <https://doi.org/10.3390/agronomy13082012>.
- [20] F. Khairati and H. Putra, "Prediksi kuantitas penggunaan obat pada layanan kesehatan menggunakan algoritma *backpropagation neural network*," *Jurnal Sistem Informasi dan Teknologi*, vol. 4, no. 3, pp. 128-135, 2022. doi: <https://doi.org/10.37034/jsisfotek.v4i3.158>.
- [21] E. C. Adjobo, A. T. S. Mahama, P. Gouton, and J. Tossa, "Automatic localization of five relevant dermoscopic structures based on YOLOv8 for diagnosis improvement," *J. Imaging*, vol. 9, no. 7, pp. 1-18, 2023. doi: <https://doi.org/10.3390/jimaging9070148>.
- [22] S. Dong, P. Wang, and K. Abbas, "A survey on deep learning and its applications," *Computer Science Review*, vol. 40, pp. 1-22, 2021. doi: <https://doi.org/10.1016/j.cosrev.2021.100379>.
- [23] A. Sarda, S. Dixit, A. Bhan, *Object detection for autonomous driving using YOLO [you only look once] algorithm*: Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), February 4-6, 2021, Tirunelveli, India. 2021, doi: <https://doi.org/10.1109/ICICV50876.2021.9388577>.