

Identifikasi Malaria Pada Citra Darah Dengan Convolutional Neural Network

Juanda Gilang Purnomo^{1)*}, Sigit Birowo²⁾, Muhammad Akbar Maulana³⁾

¹⁾²⁾³⁾Institut Bisnis dan Informatika Kwik Kian Gie
Jakarta Utara, Indonesia

¹⁾juandagilang7@gmail.com

²⁾sigit.birowo@kwikkiangie.ac.id

³⁾muhammad.maulana@kwikkiangie.ac.id

Article history:

Received 23 Nov 2024;
Revised 25 Nov 20xx;
Accepted 28 Nov 2024;
Available online 27 Des 2024

Keywords:

Computer Vision
ConvNeXtBase
Deep Learning
EfficientNetV2S
Malaria

Abstrak

Malaria adalah penyakit yang ditularkan oleh nyamuk betina Anopheles. Deteksi dini penyakit ini umumnya dilakukan menggunakan teknik mikroskopis pada sel darah merah. Penelitian ini bertujuan untuk mencari metode Convolutional Neural Network (CNN) berbasis arsitektur pretrained machine learning yang paling optimal untuk mendeteksi parasit malaria dalam citra darah mikroskopis. Kami membandingkan performa dua arsitektur CNN, yaitu EfficientNetV2S dan ConvNeXtBase, dengan menggunakan dataset yang terdiri dari 27.588 citra sel darah merah manusia, yang terbagi menjadi dua kategori: Parasitized dan Uninfected. Model dilatih dengan dua pengaturan learning rate: minimum (0,0001) dan maksimum (0,001). Evaluasi kinerja model dilakukan menggunakan confusion matrix untuk membandingkan performa dan stabilitas kedua metode dalam klasifikasi citra. Hasil penelitian menunjukkan bahwa EfficientNetV2S dengan learning rate minimum (0,0001) berhasil mencapai akurasi validasi tertinggi sebesar 96%, sedangkan ConvNeXtBase hanya mencapai akurasi maksimal sebesar 68%. Selain akurasi yang lebih tinggi, model EfficientNetV2S juga menunjukkan efisiensi komputasi yang lebih baik, sehingga menjadi pilihan yang lebih tepat untuk tugas deteksi malaria berbasis deep learning. Dengan demikian, penelitian ini menyimpulkan bahwa model arsitektur EfficientNetV2S dengan learning rate minimum merupakan pendekatan yang efektif untuk mendeteksi dini infeksi malaria dalam citra sel darah merah. Temuan ini menunjukkan bahwa teknologi deep learning memiliki potensi besar dalam membantu deteksi dini penyakit malaria. Penelitian lebih lanjut diperlukan untuk meningkatkan performa model atau mengeksplorasi metode alternatif untuk deteksi malaria.

I. PENDAHULUAN

Malaria adalah penyakit yang disebabkan oleh parasit protozoa dari genus Plasmodium dan ditularkan melalui gigitan nyamuk betina Anopheles [1]. Penyakit ini menjadi masalah kesehatan utama di wilayah tropis, termasuk Indonesia, dengan risiko besar terhadap masyarakat pedesaan yang memiliki akses terbatas ke layanan kesehatan [2]. Gejala yang ditimbulkan oleh penyakit ini meliputi demam, menggigil, keringat berlebihan, sakit kepala, dan kehilangan nafsu makan, yang dapat berkembang menjadi komplikasi serius jika tidak ditangani dengan baik [3]. Menurut laporan Organisasi Kesehatan Dunia (WHO), pada tahun 2022, sekitar 249 juta kasus malaria terjadi secara global, dengan angka kematian mencapai 608.000 jiwa [2]. Secara khusus, negara-negara Afrika Sub-Sahara seperti Nigeria, Republik Demokratik Kongo, dan Uganda menyumbang sebagian besar kasus ini [2].

Dalam dekade terakhir, pendekatan berbasis teknologi seperti Convolutional Neural Network (CNN) dapat mendeteksi adanya penyakit malaria dengan melakukan analisa gambar mikroskopis pada sel darah merah manusia, dengan teknologi ini diharapkan dapat mendeteksi dini penyakit malaria. Teknik ini tidak hanya mempercepat proses diagnosis, tetapi juga memungkinkan hasil yang lebih akurat dibandingkan metode manual seperti pemeriksaan mikroskopis tradisional [3]. Selain itu, penelitian terkini menunjukkan potensi tinggi pretrained model jaringan saraf tiruan (neural network) dalam mendeteksi keberadaan parasit pada citra darah manusia, memberikan harapan baru untuk penanganan penyakit ini [4].

* Corresponding author

Tujuan dari penelitian ini adalah untuk mencari arsitektur machine learning yang optimal untuk melakukan klasifikasi gambar. Lalu dihasilkan nya confusion matrix yang digunakan untuk mengevaluasi hasil dari arsitektur model machine learning.

II. TINJAUAN PUSTAKA

Malaria merupakan salah satu penyakit endemik yang telah lama menjadi fokus penelitian, khususnya di wilayah tropis. Berbagai metode telah dikembangkan untuk mendeteksi keberadaan parasit Plasmodium, mulai dari pendekatan manual hingga berbasis teknologi canggih seperti machine learning dan deep learning.

Aisyah dan Anraeni [1] memanfaatkan algoritma K-Nearest Neighbor (K-NN) untuk klasifikasi citra darah malaria, dengan akurasi sebesar 89%. Meskipun demikian, metode ini dinilai kurang fleksibel untuk dataset yang lebih kompleks. Penelitian dengan algoritma tradisional lainnya dilakukan oleh Banyal Nur et al. [5] yang menggunakan metode Support Vector Machine (SVM). Hasil penelitian menunjukkan akurasi sebesar 73,33%, namun memiliki kekurangan seperti jumlah data yang terlalu sedikit serta kesalahan deteksi akibat pewarnaan preparat yang kurang optimal.

Pendekatan berbasis deep learning juga menjadi perhatian utama dalam penelitian deteksi malaria. Shekar et al. [6] menggunakan Convolutional Neural Network (CNN) untuk mengekstraksi fitur dari citra apusan darah tipis dengan tiga pendekatan: basic CNN, frozen layer dari model VGG, dan fine-tune model VGG. Hasilnya, basic CNN mencapai akurasi sebesar 99,85%, frozen layer dari VGG mencapai akurasi 99,56%, sementara fine-tune model VGG mencatat akurasi tertinggi, yaitu 99,89%. Temuan ini menunjukkan bahwa metode fine-tune pada model machine learning memberikan hasil yang lebih baik dibandingkan pendekatan lainnya.

Penelitian lainnya oleh Yohanes et al. [7] menerapkan CNN berbasis saliency, menggunakan beberapa varian seperti Region Contrast Saliency (62,67% akurasi), Frequency-Tuned Saliency (90,32% akurasi), Histogram Contrast Saliency (79,06% akurasi), dan Spectral Residual Saliency (50% akurasi). Menariknya, model tanpa fitur saliency justru mencapai akurasi tertinggi, yaitu 95,06%. Hasil ini juga menunjukkan bahwa penghapusan fitur saliency dapat meningkatkan stabilitas performa model.

Pada penelitian Lydia et al. [8], arsitektur ResNet-34 digunakan untuk mendeteksi malaria, dengan akurasi sebesar 96,91%. Temuan ini mengindikasikan bahwa arsitektur ResNet-34 cukup andal untuk tugas deteksi malaria, meskipun penelitian ini hanya berfokus pada pengaturan learning rate tanpa mengoptimasi arsitektur untuk mendapatkan akurasi terbaik.

Model ringan berbasis deep learning juga diperkenalkan dalam beberapa penelitian, dengan tujuan mendukung implementasi di daerah dengan sumber daya terbatas [9]. Penelitian yang dilakukan oleh Bayu T. [10] menggunakan arsitektur seperti ResNet-101, AlexNet, dan VGG-19, di mana ResNet-101 mencapai akurasi 97,58% dengan penyesuaian learning rate. Zein [11] juga menggunakan arsitektur ResNet, yang mampu mencapai akurasi 97% pada data pengujian. Pendekatan ini memungkinkan deteksi parasit malaria dengan waktu pelatihan yang singkat tanpa mengorbankan tingkat akurasi [10].

Penelitian pada realtime object detection juga pernah dilakukan dengan menggunakan arsitektur CNN YOLOv4 dan YOLOv5, pada penelitian ini didapatkan akurasi dari YOLOv4 sebesar 90,51% dan YOLOv5 sebesar 60,7% [12][13]

III. METODE

Penelitian ini memanfaatkan berbagai model *pretrained machine learning* untuk mengklasifikasikan citra *Plasmodium* dalam mendeteksi malaria pada sel darah merah manusia. Tahapan penelitian meliputi pengumpulan data hingga klasifikasi pada data uji, seperti yang ditunjukkan dalam Gambar 1 [6].

A. Dataset

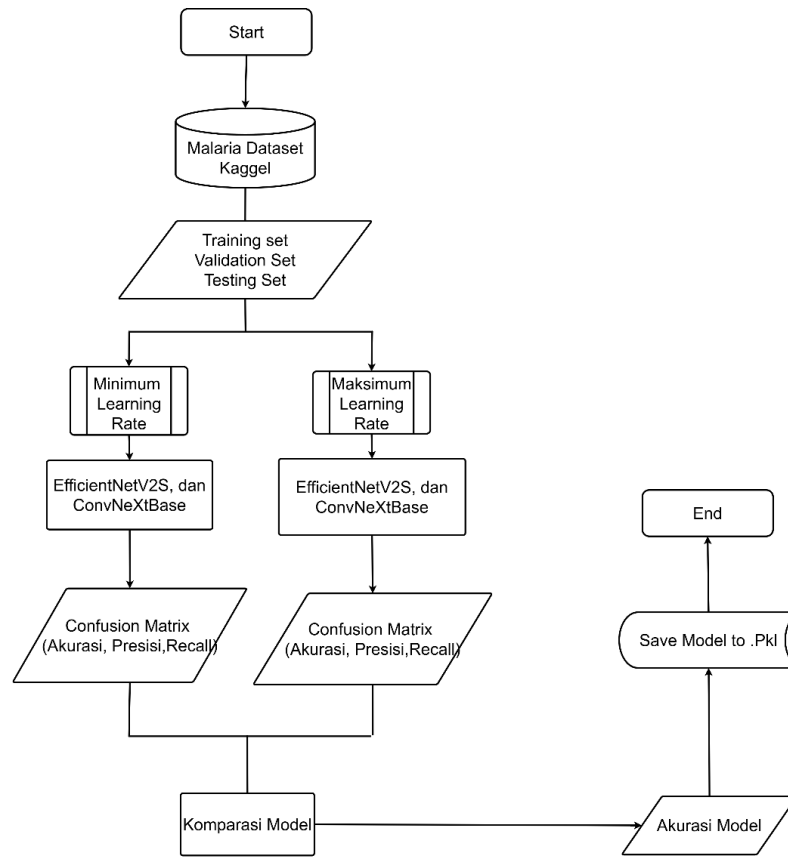
Tahap awal dari penelitian dimulai dari proses pengumpulan dataset. Dataset berupa hasil citra mikroskopis darah manusia yang terdiri dari 27.588 citra darah. Dataset citra microscopist ini diperoleh dari *National Library of Medicine* (NIH). Yang di sebarakan melalui platform kaggle. Kaggle adalah sebuah platform berbasis online yang memfasilitasi dan menyediakan akses dataset terbuka untuk belajar, berlatih, dan berkompetisi dalam berbagai masalah analisis data, machine learning dan ilmu data [14]. Lalu dataset ini di masukan kedalam folder Parasitized dan uninfected, masing - masing folder mempunyai 13.780 citra darah mikroskopis.

B. Training Set, Validation Set, & Testing Set

Training set yang peneliti gunakan adalah gabungan dari image, dan label dari nama folder. pada training set juga dibagi menjadi 80% sebagai training, 12% sebagai validation test, dan 8% sebagai *testing set*.

C. Minimum learning rate & Maksimum learning rate

Dalam penelitian ini kami mencoba menggunakan learning rate (LR) minimum (0.0001) dan maksimum LR (0.001) pada saat melakukan training model machine learning. Dimana LR tersebut adalah learning rate paling optimal yang dapat digunakan untuk metode CNN.



Gambar 1 Alur Penelitian

D. EfficientNetV2S & ConvNeXtBase

Pada penelitian ini kami mengajukan metode CNN dengan menggunakan arsitektur EfficientNetV2S dan ConvNextBase, arsitektur tersebut diajukan karena belum banyak di teliti dan digunakan dalam melakukan pendeteksian plasmodium dalam sel darah merah manusia [15][16].

E. Confusion Matrix

Confusion Matrix adalah matrik performa untuk mengevaluasi kinerja model *machine learning* [17]. *confusion matrix* berbentuk matrix 2 dimensi dengan 4 atribut yaitu *true positive* (TP), *false positive* (FP), *false negative* (FN), dan *true negative* (TN). *Confusion Matrix* digunakan untuk menghitung presisi, akurasi, *recall*, dan *F-measure* [3]. Contoh *Confusional Matrix* dapat dilihat pada Tabel 1 [3].

TABEL 1
 TABEL CONFUSION MATRIX

a)	b)	<i>Predicted Parasitised</i>	(1)	<i>Predicted Uninfected</i>
<i>Actual Parasitized</i>	<i>True Positive</i>		<i>False Negative</i>	
<i>Actual Uninfected</i>	<i>False Positive</i>		<i>True Negative</i>	

Sumber : Pratiwi N. et al., 2021

F. Komparasi Model & Akruasi Model

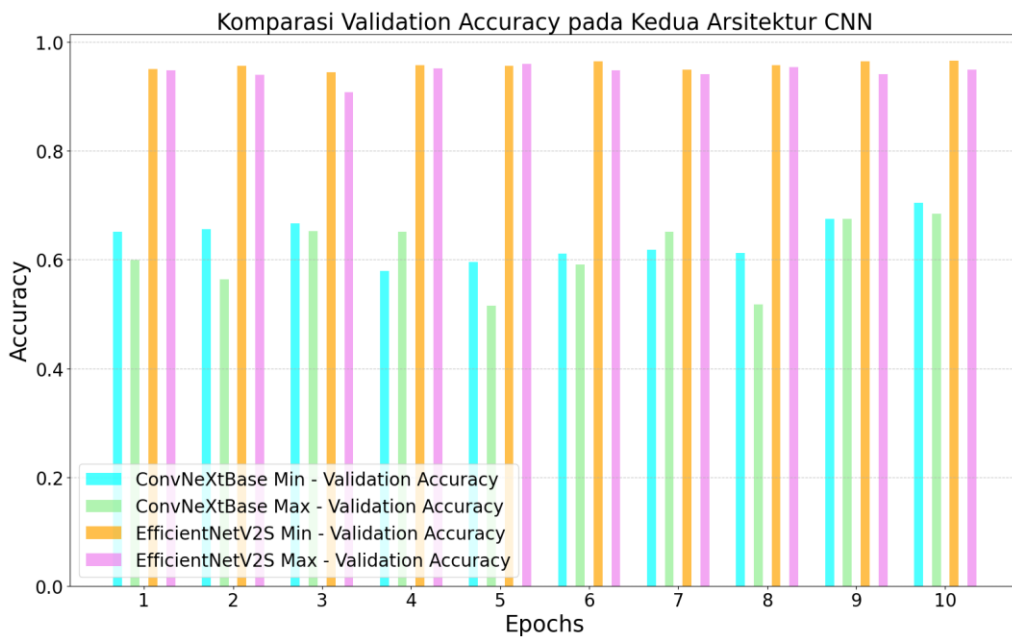
Kami melakukan komparasi dengan membandingkan hasil dari *confusion matrix* dari kedua arsitektur yang kami ajukan, lalu didapatkan akurasi terbaik dari arsitektur yang diajukan

G. Save model to .PKL

Hasil *training* model di simpan dalam bentuk ekstensi .pkl tujuannya untuk digunakan dalam penelitian selanjutnya.

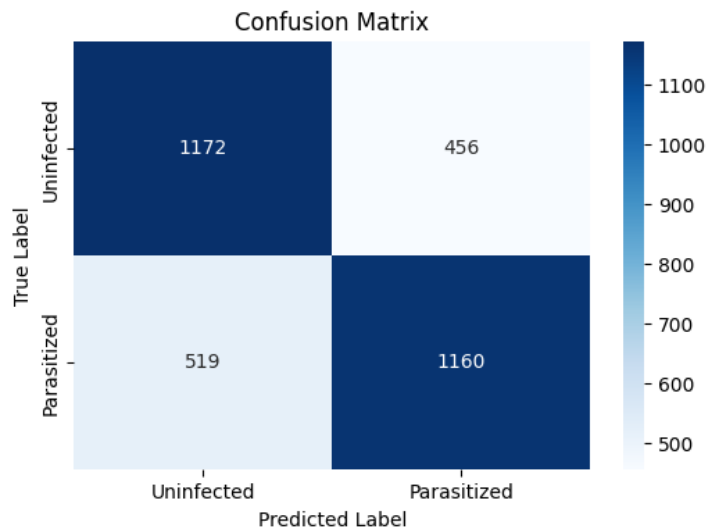
IV. HASIL

Hasil *training validation accuracy* pada model ConvNeXtBase dan EfficientNetV2S dengan *learning rate* minimum dan maksimum disajikan pada Gambar 2.

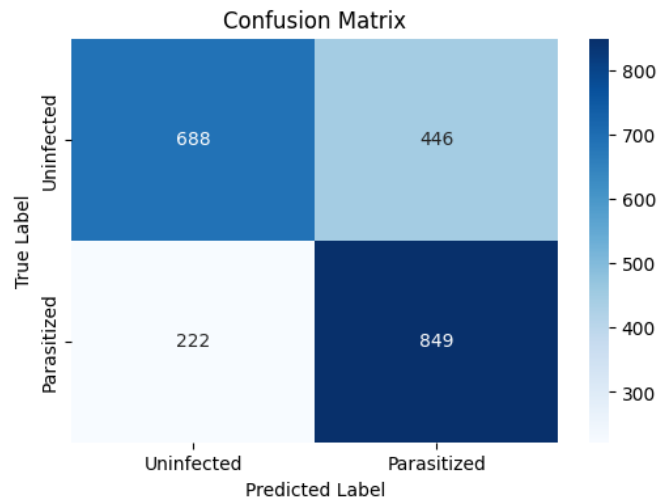


Gambar 2 Komparasi Validation Accuracy pada kedua arsitektur CNN

Selanjutnya dari *training* tersebut didapatkan confusion matrix, confusion matrix pada arsitektur ConvNeXtBase dapat dilihat pada Gambar 3 untuk minimal LR, dan Gambar 4 untuk maksimal LR.

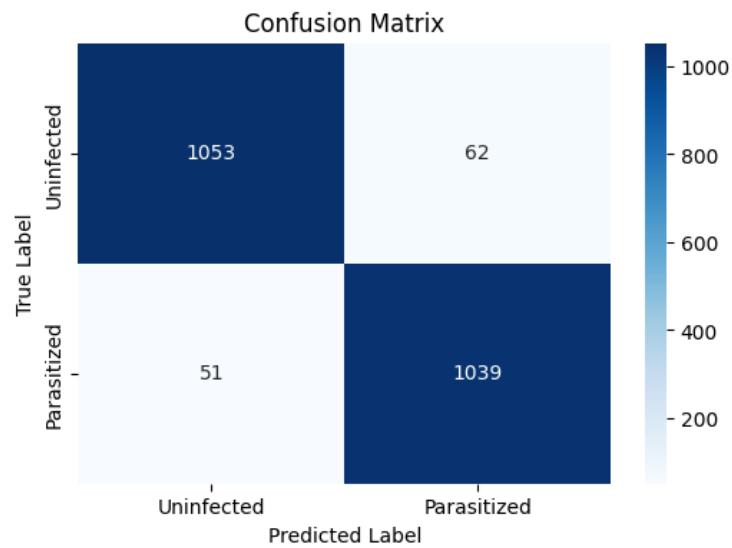


Gambar 3 ConvNeXtBase Confusion Matrix minimal LR

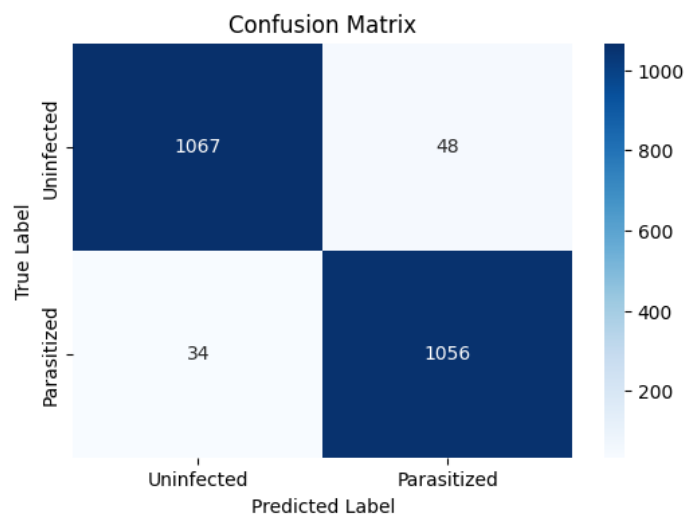


Gambar 4 ConvNeXt *Confusion Matrix* maksimal LR

Confusion matrix pada arsitektur EfficientNetV2S dapat dilihat pada gambar 5 untuk minimum LR dan Gambar 6 untuk Maksimum LR.



Gambar 5 EfficientNetV2S *Confusion Matrix* Minimum LR



Gambar 6 EfficientNetV2S *Confusion Matrix* Maksimum LR

V. PEMBAHASAN

Dari ke empat *confusion matrix*, peneliti membuat perbandingan pada *precision*, *recall*, dan *f1-score*. Perbandingan dapat dilihat pada Tabel 2.

TABEL 2
 PERBANDINGAN *PRECISION*, *RECALL*, DAN *F1-SCORE*

	<i>precision</i>	<i>recall</i>	<i>f1-score</i>
Min. LR ConvNeXtBase	0.75	0.70	0.70
Max. LR ConvNeXtBase	0.71	0.70	0.70
Min. LR EfficientNetV2S	0.94	0.94	0.96
Max. LR EfficientNetV2S	0.96	0.96	0.96

Dari Tabel 2, kita dapat mengetahui bahwa arsitektur CNN EfficientNetV2S mendapatkan *precision*, *recall*, dan *f1-score* tertinggi. Selanjutnya peneliti melakukan perbandingan pada akurasi dan *validation* akurasi terhadap kedua model CNN, perbandingan tersebut dapat dilihat pada Tabel 3.

TABEL 3
 PERBANDINGAN *PRECISION*, *RECALL*, DAN *F1-SCORE*

	<i>accuracy</i>	<i>Validation accuracy</i>
Min. LR ConvNeXtBase	0.68	0.70
Max. LR ConvNeXtBase	0.65	0.68
Min. LR EfficientNetV2S	0.94	0.96
Max. LR EfficientNetV2S	0.94	0.96

Dari penelitian metode CNN sebelumnya yaitu pada penelitian Yohanes et al. [7] mendapatkan akurasi sebesar 95.89% dengan metode CNN tanpa saliency, lalu pada penelitian Aisyah dan Anraeni [1] mendapatkan akurasi sebesar 89% , pada penelitian Banyal Nur et al., didapatkan akurasi sebesar 73,33%.

Model arsitektur yang kami kembangkan mendapatkan hasil akurasi yang lebih baik sebesar 94% untuk akurasi dan 96% untuk validasi akurasi.

VI. KESIMPULAN

Penelitian ini menunjukkan bahwa model *EfficientNetV2S* memberikan hasil akurasi tertinggi (94%) pada data citra darah malaria mikroskopis dibandingkan dengan *ConvNeXtBase* (68%). Model *EfficientNetV2S* juga mencapai nilai *precision*, *recall*, dan *f1-score* sebesar 96%. Selain itu, model ini memiliki efisiensi komputasi yang lebih baik, menjadikannya alternatif yang menjanjikan dalam sistem deteksi malaria berbasis *computer vision*.

Namun, penelitian ini memiliki beberapa batasan. Salah satunya adalah risiko *overfitting* dan *underfitting* yang masih ditemukan pada hasil pelatihan model. Hal ini mengindikasikan bahwa model belum mencapai performa optimal. Selain itu, penelitian ini menggunakan *dataset* yang relatif terbatas, sehingga hasilnya mungkin belum dapat digeneralisasi untuk berbagai kondisi klinis yang lebih kompleks.

Sebagai rekomendasi untuk penelitian selanjutnya, pengembangan dapat difokuskan pada eksplorasi model yang lebih kompleks, seperti arsitektur *transformer* atau model berbasis *self-supervised learning*. Penggunaan teknik *data augmentation* yang lebih canggih atau *transfer learning* juga dapat diterapkan untuk meningkatkan performa model. Selain itu, pengujian pada *dataset* yang lebih besar dan beragam perlu dilakukan guna meningkatkan validitas dan generalisasi hasil penelitian ini.

DAFTAR PUSTAKA

- [1] S. Anraeni, "Analisis penerapan metode K-Nearest Neighbor (K-NN) pada dataset citra penyakit malaria," *Indonesian Journal of Data and Science (IJODAS)*, vol. 3, no. 1, pp. 17–29, 2022.
- [2] World Health Organization, *World Malaria Report 2023*. World Health Organization, 2023. Accessed: Nov. 26, 2024. [Online]. Available: <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2023>
- [3] N. K. C. PRATIWI, N. IBRAHIM, Y. N. FU'ADAH, and S. RIZAL, "Deteksi Parasit Plasmodium pada Citra Mikroskopis Hapusan Darah dengan Metode Deep Learning," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 9, no. 2, p. 306, Apr. 2021, doi: 10.26760/elkomika.v9i2.306.
- [4] A. Sai Bharadwaj Reddy and D. Sujitha Juliet, "Transfer learning with RESNET-50 for malaria cell-image classification," in *Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing, ICCSP 2019*, Institute of Electrical and Electronics Engineers Inc., Apr. 2019, pp. 945–949. doi: 10.1109/ICCSP.2019.8697909.

- [5] N. Ain Banyaland A. Rachman Dayat, “KLASIFIKASI CITRA PLASMODIUM PENYEBAB PENYAKIT MALARIA DALAM SEL DARAH MERAH MANUSIA DENGAN MENGGUNAKAN METODE MULTI CLASS SUPPORT VECTOR MACHINE (SVM),” *ILKOM Jurnal Ilmiah*, vol. 8, no. 2, 2016.
- [6] G. Shekar, Revathy S., and E. Goud, *Malaria Detection using Deep Learning*. IEEE, 2022. doi: 10.1109/ICOEI48184.2020.9143023.
- [7] Yohannes, S. Devella, and K. Arianto, “Deteksi Penyakit Malaria Menggunakan Convolutional Neural Network Berbasis Saliency (Detection of Malaria Using Convolutional Neural Network Based on Saliency),” May 2020. doi: 10.30595/juita.v8i1.6671.
- [8] L. Lydia *et al.*, “Image classification using deep neural networks for malaria disease detection,” *International Journal on Emerging Technologies*, vol. 10, no. 4, pp. 66–70, 2019, [Online]. Available: www.researchtrend.net
- [9] K. M. F. Fuhad, J. F. Tuba, M. R. A. Sarker, S. Momen, N. Mohammed, and T. Rahman, “Deep learning based automatic malaria parasite detection from blood smear and its smartphone based application,” *Diagnostics*, vol. 10, no. 5, May 2020, doi: 10.3390/diagnostics10050329.
- [10] T. B. Sasongko and A. Yogyakarta, “Identifikasi Citra Plasmodium Guna Mendeteksi Parasit Penyakit Malaria dengan Menggunakan Algorithm Convolutional Neural Network Plasmodium Image Identification to Detect Malaria Parasites Using The Convolutional Neural Network Algorithm,” Aug. 2022. doi: <https://doi.org/10.30738/st.vol8.no2.a12796>.
- [11] A. Zein, J. Raya, P. Serpong, N. 10 Tangerang, and S. Banten, “Pendeteksian Penyakit Malaria Menggunakan Medical Images Analisis Dengan Deep Learning Python Detection of Malaria Using Medical Images Analysis with Deep Learning Python,” 2019.
- [12] D. Sukumarran *et al.*, “Automated Identification of Malaria-infected Cells and Classification of Human Malaria Parasites using a Two-stage Deep Learning Technique,” *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3459411.
- [13] A. Koirala *et al.*, “Deep Learning for Real-Time Malaria Parasite Detection and Counting Using YOLO-mp,” *IEEE Access*, vol. 10, pp. 102157–102172, 2022, doi: 10.1109/ACCESS.2022.3208270.
- [14] M. Sinaga, M. Albirra, and M. Sidiq, “Klasifikasi Gambar Pemandangan dengan Kecerdasan Buatan Berbasis CNN,” Apr. 2024, doi: <https://doi.org/10.35870/jtik.v8i2.1424>.
- [15] M. Tan and Q. V Le, “EfficientNetV2: Smaller Models and Faster Training,” 2021. [Online]. Available: <https://github.com/google/>
- [16] S. Woo *et al.*, “ConvNeXt V2: Co-Designing and Scaling ConvNets With Masked Autoencoders,” Jun. 2023. [Online]. Available: <https://github.com/facebookresearch/ConvNeXt-V2>
- [17] S. B. Boly and M. Akbar, “Jurnal Restikom : Riset Teknik Informatika dan Komputer Segmentasi Citra Sel Darah Menggunakan Convolutional Neural Network,” vol. 6, no. 2, pp. 390–398, 2024, [Online]. Available: <https://www.kaggle.com/jeetblahiri>