

Klasifikasi Citra Sel Darah Untuk Penyakit Malaria Dengan Metode CNN

Achmad Fauzi Saksenata^{*1}, Agus Eko Minarno², Yufis Azhar³

^{1,2,3}Universitas Muhammadiyah Malang

fsaksenata@gmail.com^{*1}, aguseko@umm.ac.id², yufis@umm.ac.id³

Abstrak

Semakin berkembangnya teknologi mengakibatkan pola hidup masyarakat ikut berubah, karena kebutuhan masyarakat bergantung pada teknologi yang ada. Perubahan ini dapat dilihat dalam penggunaan komputer dengan dibantu berbagai alat yang telah dibuat dapat mendeteksi penyakit pada seseorang. Dengan adanya alat tersebut dapat membantu masyarakat agar lebih cepat dan mudah mengetahui penyakit yang diderita. Oleh karena itu, diperlukan sistem yang dapat menganalisa, mengenali, secara sensitive, akurat dan otomatis mendiagnosa manusia terkena penyakit malaria atau tidak. Metode yang disulkan untuk menyelesaikan permasalahan tersebut Deep Learning (DL) yaitu Convolutional Neural Network (CNN) yang unggul untuk klasifikasi. Dalam penelitian ini, mengusulkan penggunaan CNN untuk membantu dalam mengklasifikasikan penyakit malaria. Dataset terdiri dari 27558 gambar sel darah. Model yang diusulkan mencapai kinerja dengan akurasi terbaik 96%. Pengujiannya berhasil serta berjalan dengan baik.

Kata Kunci: CNN, Deep Learning, Dataset, Malaria, Image Processing

Abstract

Development of technology results in changes in people's lifestyle, because people's needs existing technology. This change can be use of computers with help of tools that have been made to detect disease in a person. With this tool, it can help people to find out the disease more quickly and easily. Therefore, we need a system that can analyze, recognize, sensitively, accurately and automatically diagnose humans with infected malaria or not. Method proposed for solving problems is Deep Learning (DL) we use Convolutional Neural Network (CNN), the model is excels for classification. In this research, propose a method using CNN model for help to classification in malaria. Dataset in this research consist of 27558 blood cell image. Model we proposed reach the best performance with 96% accuracy. Test successful and went well.

Keywords: CNN, Deep Learning, Dataset, Malaria, Image Processing

1. Pendahuluan

Semakin berkembangnya teknologi mengakibatkan pola hidup masyarakat ikut berubah, karena kebutuhan masyarakat bergantung pada teknologi yang ada. Perubahan ini dapat dilihat dalam penggunaan komputer dengan dibantu berbagai alat yang telah dibuat dapat mendeteksi penyakit pada seseorang. Pada beberapa tahun sebelumnya hasil yang ditunjukkan oleh World Health Organization (WHO) terdeteksi sampai 216 juta kasus yang terkena penyakit malaria[1]. Kondisi sosial ekonomi yang buruk meyebabkan akses untuk perawatan dan pencegahan penyakit menjadi sulit [2].

Penyakit malaria yang menyerang manusia terdapat tiga tahapan trophozoite, schizont, gametocyte. Manusia yang terinfeksi akan menunjukkan berbagai gejala dari ringan sampai parah yang bisa menyebabkan kematian[3]. Untuk memberikan diagnosa yang sesuai membutuhkan sumber daya manusia yang terlatih dan pengalaman, sedangkan sumber daya ini tidak tersedia di daerah yang tertinggal[4].

Prosedur mendiagnosa dilakukan secara manual oleh patologi sangat tidak konsisten dan dapat menimbulkan perawatan yang salah[5]. Perawatan yang salah dapat menyebabkan masalah ringan seperti efek samping mual, muntah, sakit perut dan masalah berat dapat menyebabkan penyakit menjadi malaria berat yang dapat berujung kematian[6].

Malaria merupakan penyakit yang banyak menjangkit para wisatawan yang kembali dari daerah tropis[7]. Sangatlah penting untuk menghindari resiko terinfeksi perlu untuk mengembangkan diagnosa terhadap penyakit malaria yang sensitive, murah dan akurat[8].

Penelitian yang dilakukan oleh Khatami, dkk bertujuan untuk mencari gambar yang mirip dengan dataset yaitu citra x-ray. Hasil dari penelitian tersebut menggunakan metode Convolutional Neural Network (CNN) memiliki hasil terbaik yaitu dengan akurasi 90,30% dibandingkan dengan metode lain yang disebutkan dalam paper tersebut [9]. Penelitian yang dilakukan oleh Vijayalakshmi, dkk bertujuan untuk mengidentifikasi infeksi oleh. Dengan menggunakan metode Visual Geometry Group (VGG) network dan Support Vector Machine (SVM). Dengan model Convolutional Neural Network (CNN) dapat diketahui bahwa metode yang digunakan tingkat akurasinya merupakan yang paling tinggi yaitu 93.13%[1]. Penelitian yang dilakukan oleh Ahmedelmubarak Bashir, dkk bertujuan untuk mengidentifikasi penyakit malaria dengan menggunakan gambar mikroskopis. Dengan menggunakan metode Artificial Neural Network (ANN) dengan hasil yang didapatkan dengan tingkat akurasi 99%[2].

Penelitian yang dilakukan oleh Hanung Adi Nugroho, dkk dengan tujuan untuk ekstraksi fitur dan klasifikasi penyakit malaria. Menggunakan metode ANN dengan tingkat akurasi 90.8%[3]. Dalam penelitian yang dilakukan oleh Yuhang dong, dkk meneliti penggunaan metode CNN untuk sel yang terinfeksi malaria dengan menggunakan beberapa model yang dilakukan dalam percobaan tersebut mencapai hasil akurasi LeNet, AlexNet, GoogLeNet, diatas 95% dan SVM dengan akurasi dibawah 92%[4]. Penelitian yang dilakukan oleh Dhanya Bibin, dkk memiliki hasil yang tinggi 96.21% dalam ekstraksi fitur dan klasifikasi untuk menentukan sel darah yang terinfeksi malaria. Dengan menggunakan metode Deep Belief Network (DBN) [5]. Penelitian yang dilakukan oleh Zhaohui Liang, dkk bertujuan untuk klasifikasi untuk penyakit malaria yang menggunakan citra sel darah merah yang mikroskopis. Dengan menggunakan otomatisasi pengenalan citra yang berbasis machine learning. Dengan menggunakan metode CNN mendapatkan hasil dengan akurasi 97.37%[10]. Penelitian yang dilakukan Frid Adar, dkk mengklasifikasikan lesi pada liver menggunakan metode Generative Adversarial Network (GANs) yang digunakan untuk meningkatkan metode pengklasifikasian dengan CNN medical augmentasi gambar. Hasil dari penelitian tersebut mendapatkan metode yang digunakan dapat meningkatkan akurasi dari 78.6% menjadi 85.7%[11]. Penelitian D.H. Kim, dkk tentang mengidentifikasi transfer learning dari CNN Hasil didapatkan berdasarkan tes dengan nilai 0,954 mengatur pada ambang batas diagnosa untuk memaksimalkan sensitivitas dan spesifitas dengan nilai 0,9 dan 0,8[12].

Penelitian yang dilakukan oleh Sivaramakrishnan Rajaraman, dkk yaitu meneliti segmentasi sel darah dengan metode CNN untuk penyakit malaria. Agar menghasilkan dari banyak sel darah yang bertumpuk – tumpuk ketika dilihat pada dataset yang digunakan, menjadi satu sel yang terpisah. Kemudian dilakukan klasifikasi tiap sel darah dengan beberapa model dengan hasil yang didapat dengan tingkat akurasi yang 98% menggunakan model ResNet-50[13].

Pada penelitian ini, peneliti mengusulkan untuk meneliti tentang klasifikasi citra sel darah untuk penyakit malaria dengan metode CNN.

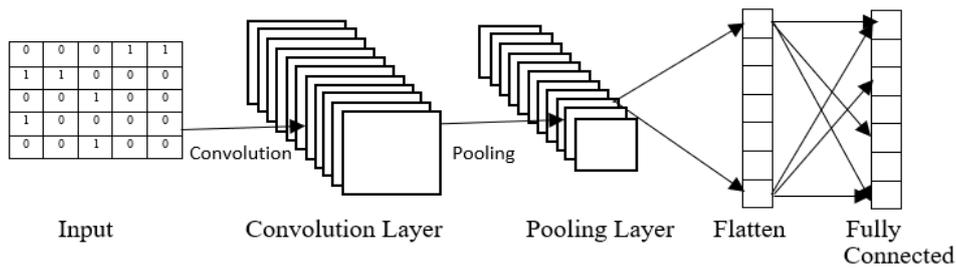
2. Metode Penelitian

2.1 Deep Learning

Termasuk ke dalam sub bidang dari machine learning yang memiliki salah satu jenis algoritma jaringan saraf tiruan (JST) menggunakan metadata yang diproses menggunakan hidden layer, perubahan non linier dari data untuk menghitung nilai keluaran[14]. Dalam algoritma deep learning memiliki fitur unik yang dapat mengekstraksi data secara otomatis. Algoritma tersebut digunakan untuk pemecahan masalah. Deep learning berlaku dalam banyak domain seperti sains, bisnis, pemerintahan dan medis. Teknik pembelajaran mesin ini dapat melakukan pengenalan gambar, pengenalan ucapan, memprediksi aktifitas, menganalisis data[15].

2.2 Algoritma CNN

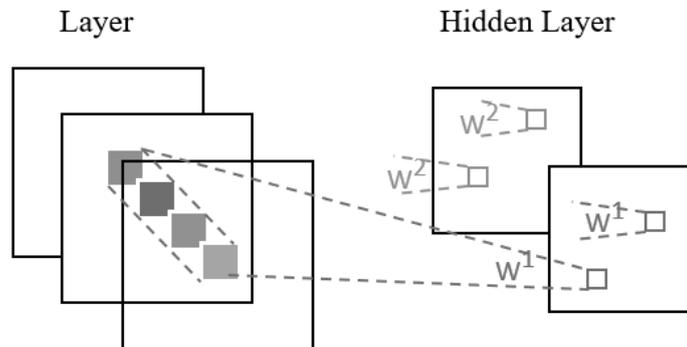
CNN merupakan Teknik yang terinspirasi dari cara manusia menghasilkan perspektif visual. Konteks visual manusia tersusun dari beberapa bentuk kecil yang berfokus pada karakteristik tertentu. Dengan kata lain gambar tersebut tersusun dari banyak ekstraksi fitur yang akan membentuk gambar secara keseluruhan, seperti pada Gambar 1[16].



Gambar 1. Arsitektur Convolutional Neural Network

2.3 Convolutional Layer

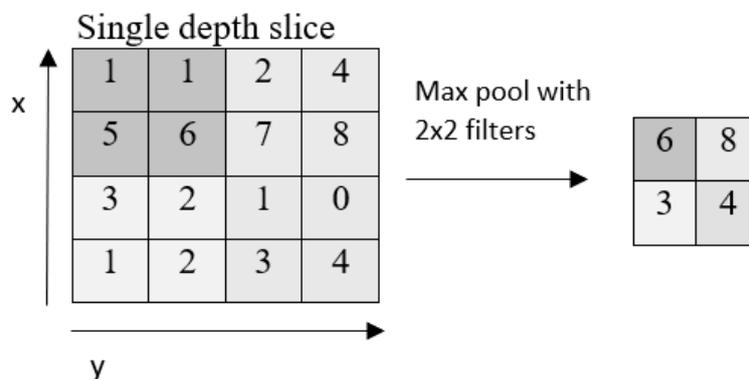
Pada Gambar 2, proses utama yang berada dalam CNN adalah Convolutional Layer yang melakukan proses konvolusi pada keluaran dari layer yang telah digunakan[17]. Setiap lapisan yang dikonvolusi memodifikasi beberapa filter ke bagian di dalam data masukan dan menghasilkan activation map atau feature map[18].



Gambar 2. Convolutional Layer

2.4 Pooling Layer

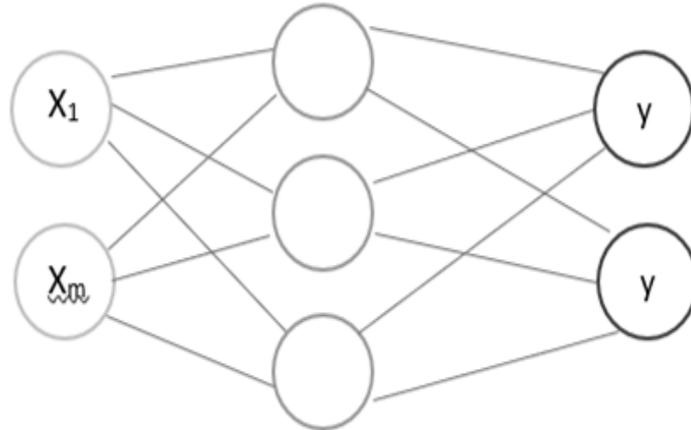
Gambar 3 merupakan tahapan setelah convolutional layer yang dalam tahap ini dilakukan pengukuran matriks. Terdiri dari beberapa filter dan stride yang setiap adanya pergeseran tergantung oleh jumlah stride yang dipeindahkan pada area *feature map* atau *activation map*[18]. Penggunaan *feature map* merupakan tahapan yang penting dalam CNN. Dengan penggunaan paling umum dengan menggunakan filter 2x2 diaplikasikan dengan langkah sebanyak 2 yang beroperasi pada setiap irisan dari input. *Max pooling* membagi output dari *convolutional layer* menjadi beberapa kotak kecil yang diambil nilai maksimal dari nilai yang berada dalam wilayah kotak tersebut digunakan untuk menyusun matriks citra yang telah direduksi[19].



Gambar 3. Pooling Layer

2.5 Fully Connected Layer

Gambar 4 pada tahapan ini digunakan untuk mengolah data untuk diklasifikasikan. Pada tahap sebelumnya dihasilkan *feature map* yang diproses dengan *flatten* atau *reshape* yang dapat menghasilkan angka - angka yang berupa vektor digunakan sebagai masukan dalam tahap *fully connected layer*[18].



Input Fully Connected Layer Output

Gambar 4. Fully Connected Layer

2.6 Data Penelitian

Data yang digunakan dalam penelitian ini diambil dari National Institutes of Health (NIH). Penelitian ini membagi data menjadi 90% sebagai data latih, dan 10% sebagai data uji. Gambar 5 berikut merupakan contoh dari dataset.



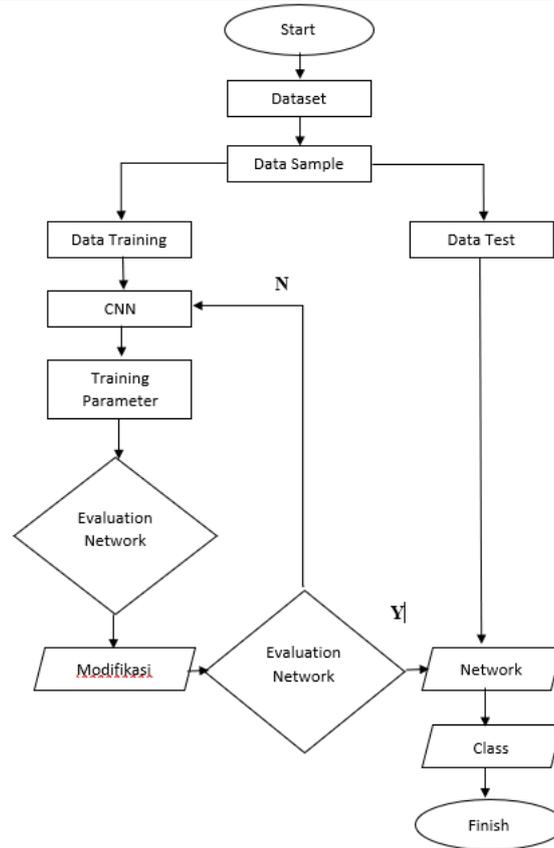
Gambar 5. Dataset sel darah terinfeksi



Gambar 6. Dataset sel darah tidak terinfeksi

2.7 Rancangan Arsitektur Sistem

Rancangan arsitektur sistem yang digunakan dalam pengujian sistem proses klasifikasi citra seperti pada Gambar 7.



Gambar 7. Proses Klasifikasi Sel Darah

2.8 Pengujian Klasifikasi

Tabel 1 merupakan tahapan untuk membuat model yang cocok kemudian dilakukan pengujian klasifikasi yang merupakan tahapan untuk mengukur nilai performa dari sistem. Pada hasil penelitian ini dilakukan pengujian terhadap algoritma yang menggunakan confusion matrix. Berikut ilustrasi dari confusion matrix.

Tabel 1. Confusion Matrix

		Nilai Sebenarnya	
		TRUE	FALSE
Nilai Prediksi	TRUE	TP (True Positive)	FP (False Positive)
	FALSE	FN (False Negative)	TN (True Negative)

Berdasarkan confusion matrix, pengukuran performa menggunakan rumus precision, accuracy, dan recall.

1. Precision, mengevaluasi seberapa baik ketepatan jumlah prediksi berhasil diprediksi dengan baik dengan hasil dari pengujian. Rumus perhitungan precision ditulis dengan Persamaan 1.

$$precision = \frac{TP}{TP + FP} \tag{1}$$

2. Accuracy adalah mengukur tingkat kedekatan kuantitas dengan nilai yang sebenarnya. Rumus perhitungan accuracy ditulis dengan Persamaan 2.

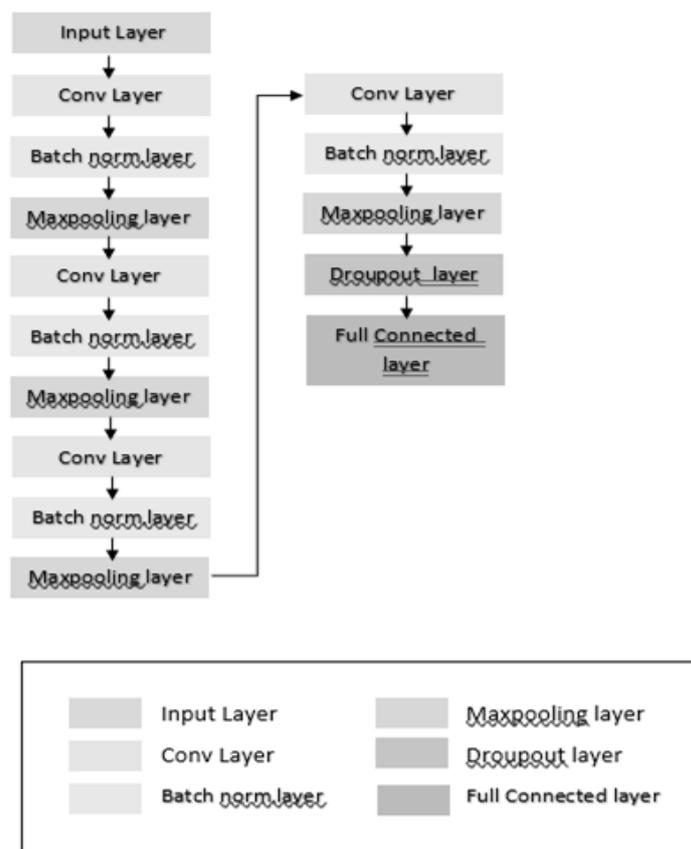
$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

3. Recall, mengevaluasi seberapa baik ketepatan jumlah prediksi berhasil diprediksi dengan baik dari jumlah data yang memiliki label kelas tersebut. Rumus perhitungan recall ditulis dengan Persamaan 3.

$$recall = \frac{TP}{TP + FN} \quad (3)$$

3. Hasil Penelitian dan Pembahasan

Perancangan model pada Gambar 8 dibuat menggunakan model CNN ini melibatkan 4 convolutional layer dengan jumlah kernel 64(Conv 1 dan Conv 2), 128(Conv 3 dan Conv 4) dan ukuran kernel 3 x 3, dengan setiap convolutional layer menggunakan activation relu, batchnormalization, maxpooling layer, kemudian diikuti dengan full connected layer dengan 128 menggunakan aktivasi softmax.



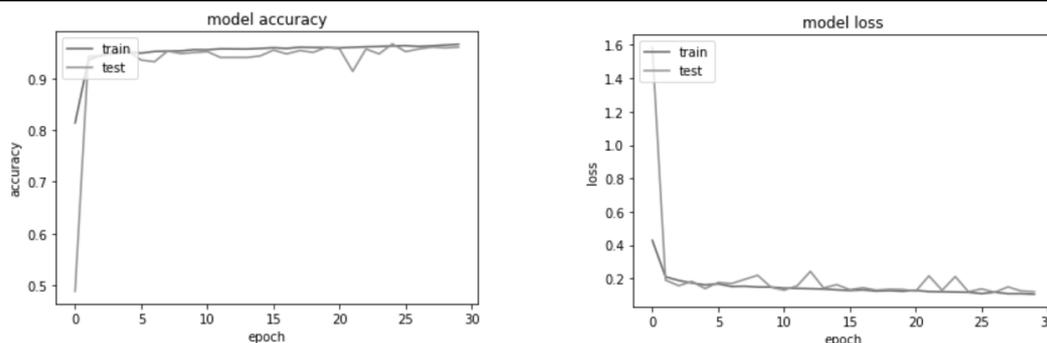
Gambar 8. Arsitektur Model

3.1 Perancangan Sistem

Pada pengujian ini data yang digunakan sebanyak 27558 sel darah. Detail skenario dapat dilihat pada Tabel 2. Akurasi yang didapatkan dari skenario adalah 96%, precision 96%, recall 96% dan f1-score 96%. Detail hasil skenario dapat dilihat pada Tabel 2. Grafik model akurasi dan loss dari data train dan validasi dapat dilihat pada Gambar 9.

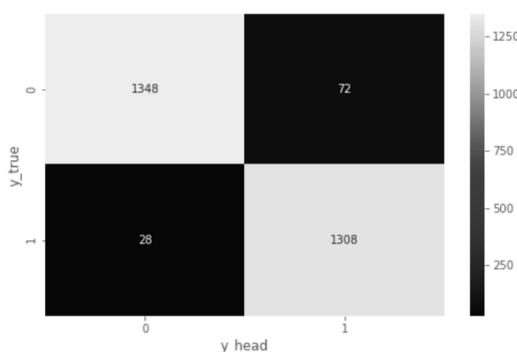
Tabel 2. Klasifikasi performa model

Method	Accuracy	Precision	Recall	F1-Score
CNN	96%	96%	96%	96%



Gambar 9. Akurasi dan Loss Pada Model

Hasil pengujian menggunakan confusion matrix menunjukkan performa yang cukup baik dalam model yang dibuat peneliti. Hal ini dapat dilihat pada Gambar 10. Mencapai 1348 dari data keseluruhan 2756.



Gambar 10. Confusion Matrix

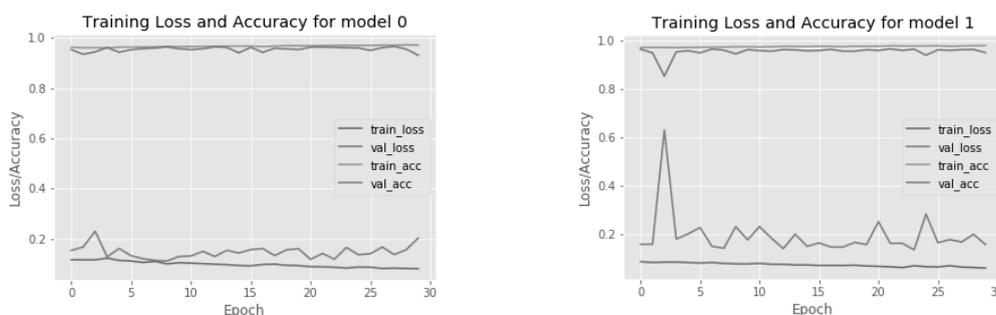
3.2 Ensemble Method

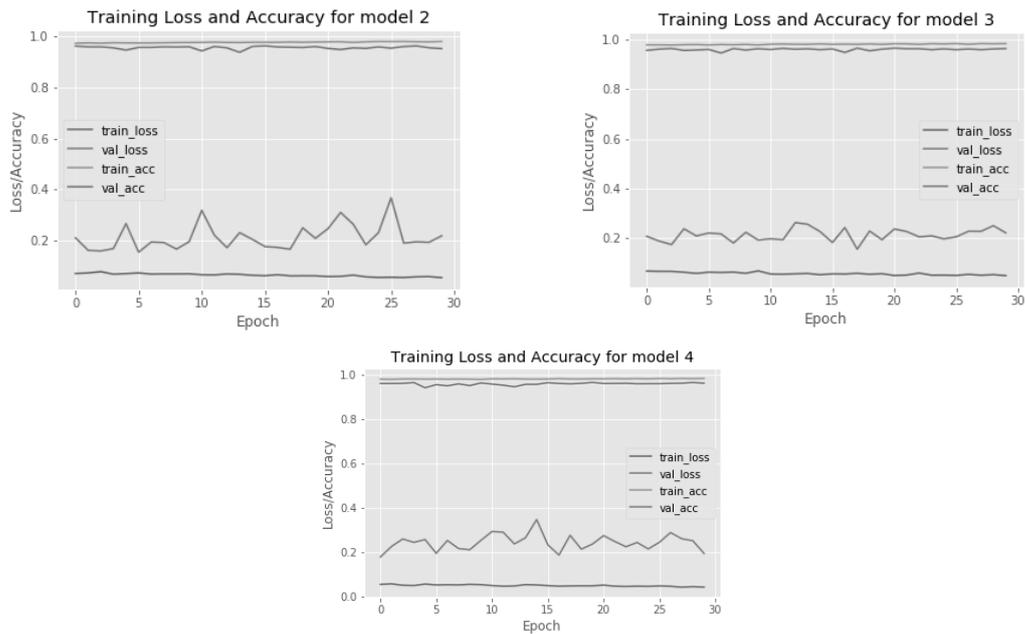
Merupakan suatu Teknik dimana mengulang iterasi dari awal model yang sudah ada untuk mendapatkan performa yang lebih baik. Pada tahap ini model yang telah diuji peneliti akan digunakan lagi sebanyak 5 kali. Detail hasil dapat dilihat pada Tabel 3.

Tabel 3. Klasifikasi performa model

Model	Accuracy	Precision	Recall	F1-Score
0	93%	94%	93%	93%
1	95%	95%	95%	95%
2	95%	95%	95%	95%
3	96%	96%	96%	96%
4	96%	96%	96%	96%

Grafik hasil dari 5 model tersebut pada Gambar 11.





Gambar 11. Grafik ensemble dari model 0 – 4

Pada Tabel 4 merupakan detail hyperparameter yang digunakan pada model CNN diantaranya jumlah convolutional layer dan relu layers, jumlah batch normalization, jumlah dropout layers, epoch, jumlah fully connected layers, kernel pada convolutional layer, pooling layer, optimizer, batch size, step per epoch dan dropout rate.

Tabel 4. Hyperparameter model CNN

Factor(s)	Values
Jumlah Convolutional + ReLU layers	1,2,3,4
Jumlah Batch Normalization	1,2,3,4
Jumlah dropout layers	1,2
Epoch	30
Jumlah fully connected layers	1,2
Kernel convolutional layer	64,64,128,128
Pooling layers	Max pooling
Optimizers	Adam
Batch size	64
Step per epoch	388
Dropout rate	0.25, 0.50

Pada Tabel 5 menyajikan perbandingan performa model CNN yang diajukan dengan penelitian sebelumnya yang menggunakan data sama dalam mengklasifikasi jenis penyakit malaria.

Tabel 5. Perbandingan Dengan Penelitian Sebelumnya

Performa (%)	Proposed Structure	S. Rajaraman et al [13]
Accuracy	96%	98%
Specificity / Precision	96%	99%

Recall / Sensitivity	96%	98%
F1-Score	96%	98%

Pada Tabel 5 merupakan perbandingan model dari S. Rajaraman *et al* [13] dan peneliti. Pada model S. Rajaraman *et al* [13] menggunakan model ResNet-50.

Model yang diusulkan menggunakan 4 *convolutional layer* yaitu 64,64,128,128 menggunakan maxpooling 2 x 2 dengan batchnormalization pada setiap *convolutional layer*, serta pembagian dataset antara data training dan data uji sebesar 90 : 10.

Perbedaan model penelitian sebelumnya dengan penelitian terkait adalah pada jumlah layer konvolusi, ukuran pooling layer, penambahan batchnormalization, dan dropout. Pada penelitian sebelumnya tidak disebutkan perbandingan antara data latih dengan data uji. Dari perbedaan tersebut model yang dibangun performanya masih kurang baik daripada penelitian S. Rajaraman *et al* [13].

skenario runtime. Oleh karena itu, untuk menggambarkan hal tersebut maka harus diketahui objek - objek yang terlibat dalam sebuah use case beserta metode yang dimiliki secara sequential (berurutan). Sequence diagram digunakan pada tahapan berikut

4. Kesimpulan

Berdasarkan pada pengujian yang telah dilakukan banyak data tidak selalu mempengaruhi akurasi yang dihasilkan didukung dengan penggunaan model CNN yang baik untuk mengklasifikasikan penyakit malaria dengan sel darah. Pada pengujian dengan menggunakan dataset dengan pembagian 24802 sebagai data latih dan 2756 sebagai data uji yang mampu menghasilkan akurasi 96% tetapi masih kurang baik daripada penelitian sebelumnya.

Referensi

- [1] A. Vijayalakshmi and R. K. B, "Deep learning approach to detect malaria from microscopic images," 2019.
- [2] A. Bashir, Z. A. Mustafa, I. Abdelhameid, and R. Ibrahim, "Detection of Malaria Parasites Using Digital Image Processing," no. c, 2017.
- [3] E. Engineering, I. Technology, F. Engineering, and U. G. Mada, "Feature Extraction and Classification for Detection Malaria Parasites in Thin Blood Smear," vol. 1, no. c, pp. 197–201, 2015.
- [4] Y. Dong *et al.*, "Evaluations of Deep Convolutional Neural Networks for Automatic Identification of Malaria Infected Cells," *2017 IEEE EMBS Int. Conf. Biomed. Heal. Informatics*, pp. 101–104, 2017.
- [5] D. Bibin, M. S. Nair, and P. Punitha, "Malaria Parasite Detection from Peripheral Blood Smear Images Using Deep Belief Networks," *IEEE Access*, vol. 5, pp. 9099–9108, 2017.
- [6] M. Poostchi, K. Silamut, R. Maude, S. Jaeger, and G. Thoma, "Image Analysis and Machine Learning for Detecting Malaria," *Transl. Res.*, no. 2018, 2017.
- [7] R. Pcr, "crossm Detection of Plasmodium Infection by the illumi gene Malaria Assay," vol. 55, no. 10, pp. 3037–3045, 2017.
- [8] G. Figueroa-Miranda *et al.*, "Aptamer-based electrochemical biosensor for highly sensitive and selective malaria detection with adjustable dynamic response range and reusability," *Sensors Actuators, B Chem.*, vol. 255, pp. 235–243, 2018.
- [9] A. Khatami, M. Babaie, H. R. Tizhoosh, A. Khosravi, T. Nguyen, and S. Nahavandi, "A sequential search-space shrinking using CNN transfer learning and a Radon projection pool for medical image retrieval," *Expert Syst. Appl.*, vol. 100, pp. 224–233, 2018.
- [10] Z. Liang *et al.*, "CNN-Based Image Analysis for Malaria Diagnosis," pp. 493–496, 2016.
- [11] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, 2018.
- [12] D. H. Kim and T. MacKinnon, "Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks," *Clin. Radiol.*, vol. 73, no. 5, pp. 439–445, 2018.
- [13] S. Rajaraman *et al.*, "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images," pp. 1–17, 2018.
- [14] J. Heaton, "Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning," *Genet. Program. Evolvable Mach.*, vol. 19, no. 1–2, pp. 305–307, 2018.

-
- [15] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [16] H. Lee and H. Kwon, "Going Deeper with Contextual CNN for Hyperspectral Image Classification," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4843–4855, 2017.
- [17] E. N. Arrofiqoh and H. Harintaka, "Implementasi Metode Convolutional Neural Network Untuk Klasifikasi Tanaman Pada Citra Resolusi Tinggi," *Geomatika*, vol. 24, no. 2, p. 61, 2018.
- [18] A. Santoso and G. Ariyanto, "Implementasi Deep Learning Berbasis Keras Untuk Pengenalan Wajah," *Emit. J. Tek. Elektro*, vol. 18, no. 01, pp. 15–21, 2018.
- [19] W. S. Eka Putra, "Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101," *J. Tek. ITS*, vol. 5, no. 1, 2016.