

Identifikasi Infeksi Penyakit Malaria Berdasarkan Citra Darah Menggunakan Convolutional Neural Network

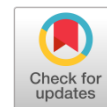
Adelia Fitriawati Zakiyah^{a,1,*}

^aTeknik Informatika, Universitas Ahmad Dahlan, Yogyakarta 55191, Indonesia

¹adeliafitriawati29@gmail.com

ABSTRAK

Malaria merupakan penyakit yang disebabkan oleh nyamuk *anopheles* yang teridentifikasi parasit *plasmodium*. Penyebaran kasus malaria yang semakin meningkat dapat dihentikan apabila terdapat sistem atau metode diagnosa secara akurat dan cepat. Tujuan penelitian ini adalah untuk mendeteksi sel darah merah yang terinfeksi malaria, untuk membantu tim kesehatan atau tim medis dalam mendiagnosa malaria dengan waktu yang lebih singkat serta menekan biaya peralatan medis serta mengurangi human error. Penelitian ini menggunakan *convolutional neural network* untuk mengklasifikasi citra sampel sel darah merah pasien positif malaria dan sampel sel darah normal. Tahapan penelitian dimulai dari identifikasi masalah, perumusan masalah, studi literatur, pengumpulan data di peroleh dari Kaggle.com dan NIH menggunakan kurang lebih 2000 data citra, spesifikasi kebutuhan, pengolahan data, perancangan model menggunakan arsitektur alexnet yang telah melalui proses reduksi dengan cara optimalisasi arsitektur CNN dan mencari kombinasi parameter yang menghasilkan nilai akurasi terbaik, implementasi menggunakan python, dan GUI *flask* serta pengujian performa sistem menggunakan akurasi dari *confusion matrix*. Hasil penelitian membuktikan pasien positif malaria dan pasien negatif malaria, menggunakan 80% data training serta 20% data testing dengan total dataset sebanyak 2000 data. Hasil akurasi sebesar 98% dan hasil uji validitas model menggunakan data baru diluar data *training* dan *testing* didapatkan hasil akurasi sebesar 100%.



Kata Kunci

Convolutional neural network
Citra
Preprocessing
Malaria



This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

1. Pendahuluan

Malaria adalah penyakit yang disebabkan oleh nyamuk *anopheles* yang teridentifikasi parasit *plasmodium* [1]. WHO mencatat penyebaran malaria diseluruh dunia menjangkau 228 juta kasus dengan total kematian sekitar 405.000. Kematian akibat malaria di dominasi oleh anak usia dini (di bawah 5 tahun) mereka menyumbang 67% dari total kematian malaria di seluruh dunia pada tahun 2018. Menurut data statistik pada gambar 1.1 yang didapatkan dari WHO menjelaskan mengenai tingkat kejadian kasus malaria menurut wilayah WHO dari tahun 2010-2018 mengalami penurunan kasus dari 71,1 pada tahun 2010 menjadi 57,4 pada tahun 2018. Akan tetapi pada tahun 2014-2018 [3]. Malaria dapat dicegah, dikendalikan dan disembuhkan jika tersedia metode diagnostik yang akurat dan efisien. Metode standar untuk deteksi malaria biasanya adalah dengan pemeriksaan mikroskopis. Tetapi metode ini tidak efisien dan kualitas diagnosanya bergantung pada pengalaman dan pengetahuan ahli mikroskop. Pengecekan malaria secara mikroskopis tidak selalu memberikan hasil yang akurat hal tersebut dapat dipicu oleh tenaga ahli yang kurang berpengalaman, peralatan medis dan bahan reagen tidak sesuai standar operasional serta jumlah ketersediaan yang dapat diperiksa oleh tenaga ahli melebihi bobot kerja. Tes diagnostik cepat memang dapat digunakan, tetapi tes tersebut lebih mahal dan memberikan informasi yang lebih sedikit ketimbang pengujian menggunakan mikroskop [4].

Perkembangan teknologi telah membantu untuk menyelesaikan permasalahan Kesehatan. Diagnosa yang salah dapat dicegah dengan mempelajari pola dari sel darah merah pasien dan dapat diklasifikasikan menggunakan computer [7]. Sebuah penelitian mengatakan bahwa metode deep learning merupakan metode klasifikasi yang paling akurat dibandingkan dengan *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN). Alasan deep learning lebih unggul karena letak dimensi perwakilan yang diakumulasi

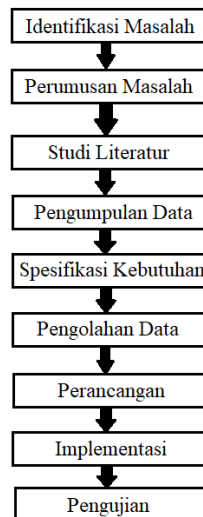
lebih bagus untuk merepresentasikan matriks kecil. Pada parameter *deep learning* memiliki beberapa komponen yaitu jumlah *filter*, *layer* dan *neuron*. *Epoch* yang digunakan pada *deep learning* sangat kompleks yaitu bobot dan bias. Sehingga pendekatan yang dilakukan oleh *machine learning* lebih presisi dari hasil neuron yang telah diekstraksi[12].

Salah satu penelitian mengenai klasifikasi malaria telah dilakukan oleh Trsyne pada penelitiannya yang berjudul “Implementasi Metode *Support Vector Machine* Untuk Klasifikasi Jenis Penyakit Malaria” menghasilkan nilai akurasi sebesar 72% [8]. Penelitian selanjutnya dilakukan oleh zhaohui liang pada penelitiannya yang berjudul “*CNN-based image analysis for malaria diagnosis*” menghasilkan nilai akurasi sebesar 97,37% [7]. Pembelajaran mesin telah digunakan untuk mendeteksi parasit *plasmodium vivax* menggunakan metode *haar cascade* dengan memanfaatkan segmentasi citra untuk pendeteksian sel parasit *plasmodium*. Dengan penelitian yang berjudul “Segmentasi Citra Parasit Malaria *Plasmodium Vivax* Dengan Menggunakan Metode *Haar Cascade*” dan menghasilkan akurasi sebesar 95% [10]. Selanjutnya pendeteksian parasit *plasmodium falciparum* menggunakan metode *haar cascade*. Dengan penelitian yang berjudul “Pengaruh *Image Engagement* pada Aplikasi Parasit Malaria” *image Engagement* disini digunakan untuk mengatur kecerahan citra gambar agar mudah untuk di Analisa. Penelitian ini menghasilkan nilai akurasi sebesar 85% [6]. Dari perbandingan sumber data tersebut yang sampai saat ini memiliki hasil paling relevan dalam identifikasi citra adalah *Convolutional Neural Network* (CNN) karena hasil akurasi yang dapat menyaingi manusia pada dataset tertentu [9]. Hal itu dapat terjadi karena algoritma CNN dapat meniru sistem identifikasi citra pada *visual cortex* manusia yang berfungsi untuk mengolah informasi [5]. Dengan munculnya kemampuan komputasi *high-end*, CNN baru-baru ini menjadi standar *de-facto* untuk klasifikasi dan telah memberikan hasil yang dapat diandalkan dalam domain medis [11]. Sedangkan penggunaan segmentasi warna pada citra serta penggunaan *image engagement* bersifat eksperimental, subjektif serta bergantung pada tujuan yang akan dicapai oleh peneliti. Salah satu kelebihan dari penggunaan CNN adalah memberikan akurasi yang tinggi pada dataset gambar objek benda hal itu karena pengolahan konvolusi pada CNN menggunakan matriks ukuran piksel sehingga deteksinya bisa lebih detail. Ukuran *convolutional matrix* pada CNN sangat mempengaruhi kecepatan durasi komputasi dari hasil akurasi, sehingga untuk mengatasi masalah tersebut peneliti dapat mengurangi jumlah konvolusi untuk mempersingkat waktu komputasi. Pada penelitian oleh simon, et al mengusulkan untuk mereduksi jumlah layer *fullyconnected* menggunakan arsitektur alexnet. Hal tersebut dilakukan guna mengurangi *overfitting* karena jumlah varians kelas yang digunakan hanya sedikit. Kemudian Rui, et al mengusulkan untuk memotong jumlah neuron dan layer *fullyconnected* pada arsitektur alexnet, tetapi jumlah layer konvolusi yang digunakan tetap menggunakan alexnet dengan 5 layer. Mereduksi layer *fullyconnected* akan berdampak pada kecepatan komputasi tetapi dapat menurunkan nilai akurasi. Selain itu jumlah filter pada layer konvolusi juga dapat menyebabkan *overfitting* karena jumlah varian data nya sedikit [12].

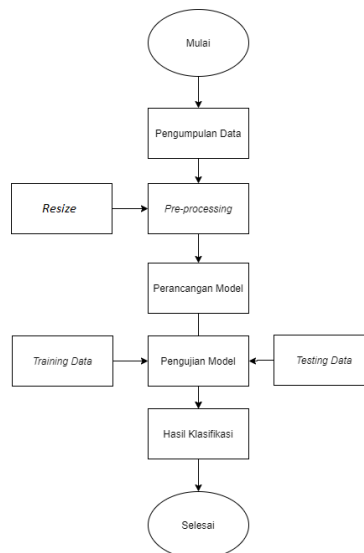
Setelah melihat permasalahan yang ada dan membaca penelitian yang telah dilakukan sebelumnya, maka pada penelitian ini akan mengusulkan metode *Convolutional Neural Network* (CNN) untuk identifikasi citra darah pasien malaria. Untuk mendapatkan waktu komputasi yang cepat dan akurasi tinggi maka akan dilakukan efisiensi pada kompleksitas model alexnet dengan mengurangi jumlah layer *konvolusi* serta memotong layer *fullyconnected*. Pengembangan tersebut diimplementasikan pada penelitian tentang Identifikasi Infeksi Penyakit Malaria Berdasarkan Citra Darah Menggunakan *Convolutional Neural Network* guna membantu mengenali citra darah yang terinfeksi dan tidak terinfeksi malaria.

2. Metode

Penelitian ini akan menggunakan algoritma *convolutional neural network* sebagai *classifier* untuk mengklasifikasi citra sampel sel darah merah pasien positif malaria dan sampel sel darah pasien normal. Dataset yang digunakan berjumlah 2000 data dan akan di klasifikasikan menjadi 2 kelas yaitu positif malaria dan negative malaria. Data tersebut berasal dari Kaggle.com yang diperoleh dari *National Intitute of Health* (NIH). Setelah dataset didapatkan selanjutnya adalah perancangan sistem, implementasi menggunakan python dan GUI *flask* serta pengujian performa sistem menggunakan akurasi dari *confusion matrix*. Tahap penelitian dapat dilihat pada gambar 1.



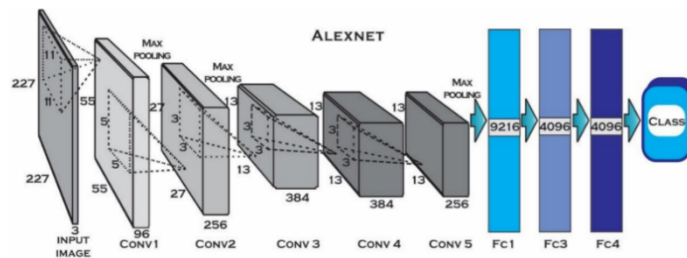
Gambar 1. Tahapan Penelitian



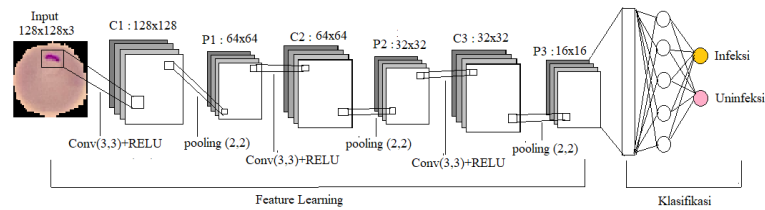
Gambar 2. Diagram Perancangan Sistem

Pada Gambar 2 merupakan diagram dari perancangan sistem yang meliputi :

- 1) Pengumpulan data
Data yang digunakan pada penelitian ini berupa citra sel darah pasien positif malaria dan citra sel darah pasien normal, data tersebut kemudian disimpan dalam bentuk format **png*. Data tersebut di peroleh dari Kaggle.com yang bersumber dari NIH.
- 2) Pre-processing
Pada proses ini citra inputan akan melalui proses *resize* untuk memperkecil ukuran data citra guna mempercepat proses komputasi dalam penelitian, citra inputan akan di *resize* menjadi ukuran 128x128 pixel.
- 3) Perancangan model
Selanjutnya adalah perancangan model CNN, pada proses ini akan menggunakan alexnet yang melalui proses reduksi layer. Gambar 3 merupakan arsitektur alexnet dan gambar 4 merupakan arsitektur usulan model CNN. Pada rancangan usulan disini akan menggunakan input shape berukuran 128x128, 3 layer *convolution* serta 2 layer *fullyconnected*.



Gambar 3. Model CNN Alexnet



Gambar 4. Rancangan Usulan Model Arsitektur CNN

Tabel 1. Usulan Rancangan Model Arsitektur CNN

	Layer	Size and Future	Kernel Size	Stride	Activation
	Input	Image	128x128x3	-	-
1	Conv1	128x128x16	3x3	1	Relu
	Maxpool	64x64x16	2x2	1	Relu
2	Conv2	64x64x32	3x3	1	Relu
	Maxpool	32x32x32	2x2	1	Relu
3	Conv3	32x32x64	3x3	1	Relu
	Maxpool	16x16x64	2x2	1	Relu
4	Dropout	0.2	-	-	-
5	flatten	16384	-	-	-
6	Fullyconnected	64	-	-	relu
output	Fullyconnected	2	-	-	Softmax

- Pada proses konvolusi pertama menggunakan kernel berukuran 3x3 dengan jumlah filter sebanyak 16 filter. Proses konvolusi ini merupakan proses kombinasi antara dua buah matriks untuk menghasilkan nilai matriks yang baru. Setelah proses konvolusi selesai selanjutnya ditambahkan fungsi aktivasi ReLu yang bertujuan untuk mengubah nilai negatif menjadi nol. Hasil dari konvolusi ini adalah 448. Untuk menghitung input kedalaman konvolusi menggunakan rumus :

$$(\text{input_size} + (2 * \text{padding}) - (\text{filter_size} - 1)) \dots \dots \dots (1)$$
 untuk menghitung parameter menggunakan rumus :

$$((\text{kernel_size} * \text{dimensi_sblmnya}) + 1) * \text{filter_size} \dots \dots \dots (2)$$
- Selanjutnya adalah proses pooling, pada lapisan ini akan melakukan pengurangan ukuran matriks menggunakan operasi pooling. Berdasarkan hasil pooling menghasilkan matriks baru berukuran 64x64 menggunakan kernel pooling 2x2.
- Proses konvolusi ke 2 yaitu dengan meneruskan hasil output dari pooling layer pertama. Dengan Input matriks berukuran 64x64 serta jumlah filter 32 dengan ukuran kernel matriks 3x3.
- Proses pooling ke 2, berdasarkan hasil pooling pertama kemudian menghasilkan matriks baru berukuran 32x32 dengan menggunakan kernel pooling berukuran 2x2.

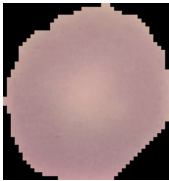
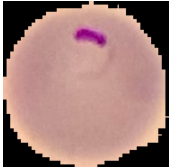
- e) Proses konvolusi ke 3 yaitu dengan meneruskan hasil output dari pooling layer kedua. Input matriks tersebut berukuran 32x32 menggunakan jumlah filter 64 serta ukuran kernel 3x3.
 - f) Proses pooling ke 3, berdasarkan hasil pooling kedua menghasilkan matriks baru yang berukuran 16x16 menggunakan kernel pooling berukuran 2x2.
 - g) Dropout adalah Teknik regulasi jaringan syaraf yang digunakan untuk menentukan beberapa neuron secara acak untuk dibuang.
 - h) Pada tahap flatten hanya menggunakan satu hidden layer pada jaringan MLP. Dengan mengubah output dari pooling layer menjadi sebuah vector.
- 4) Pengujian model
Pengujian ini akan melalui dua tahap yaitu tahap training dan yang kedua adalah tahap testing. Pada tahap training model akan diuji menggunakan data gambar yang telah disediakan sebanyak 2000, dengan jumlah gambar perkelas sebanyak 1000 data. Kemudian data tersebut dibagi menjadi dua yaitu 1600 untuk data training serta 400 untuk data testing. Terakhir adalah data yang akan digunakan untuk evaluasi menggunakan 10 data citra uji diluar data training serta data testing untuk menguji apakah model tersebut sudah menghasilkan performa yang baik dalam melakukan klasifikasi sebuah gambar.
- 5) Hasil klasifikasi
Hasil klasifikasi ini akan membuktikan pasien positif malaria dan pasien negatif malaria, dengan jumlah klasifikasi sebanyak 2 kelas.

3. Hasil dan Pembahasan

3.1. Pengumpulan Data

Data yang digunakan pada penelitian “klasifikasi citra untuk identifikasi infeksi pada penyakit malaria menggunakan *convolutional neural network*” akan menggunakan citra hapusan darah yang diperoleh dari National Institutes of Health (NIH). Total data dalam dataset yang digunakan untuk pelatihan ialah 2000 data citra hapusan darah. Data tersebut dibagi menjadi 80% data training, dan 20% data testing sehingga didapatkan 1600 data digunakan sebagai data training dan 400 sebagai data validation. Dengan struktur kelas sebanyak 2 kelas.

Tabel 2. Indikator Variabel Dataset Citra

Gambar sel darah	Variabel	Definisi Variabel
	Sel darah normal	Citra RGB dengan bentuk bulat polos.
	Sel darah berparasit	Citra RGB dengan bentuk bulat memiliki pola noda berbentuk lonjong atau bulat.

3.2. Pre-processing

Setiap inputan gambar yang masuk kedalam sistem akan mengalami proses resize guna mengatur besar ukuran pixel citra tersebut dan mengubahnya menjadi ukuran 128x128 piksel. Pada kebutuhan processing memiliki beberapa kebutuhan seperti sistem dapat memproses data citra dari setiap inputan dan menampilkan hasil dari setiap inputan tersebut.

3.3. Perancangan Model CNN

Untuk mendapatkan model terbaik dibutuhkan parameter yang terbaik juga. Parameter disini dapat dipengaruhi oleh jumlah epoch, pengaruh ukuran input gambar, pengaruh jumlah data training, pengaruh skenario pembagian data serta pengaruh learning rate. Penentuan parameter ini bertujuan untuk membandingkan model yang paling terbaik dengan memperhatikan nilai dari parameter tersebut.

1) Pengaruh Input Image

Pada penelitian ini akan melakukan percobaan terhadap Input image yang akan digunakan. Input yang digunakan adalah 64x64, 128x128. Dengan jumlah epoch yang digunakan adalah 30 epoch.

Tabel 3. Hasil Akurasi Pengaruh Input Image

<i>Input Shape Image</i>	<i>Accuracy Validation</i>	<i>Loss Validation</i>	<i>Time (second)</i>
64 x 64	97%	0.1	160 m/s
128 x 128	97%	0.2	157 m/s

Berdasarkan percobaan perbedaan input shape, didapatkan hasil seperti yang terlihat pada tabel 3. tingkat akurasi yang didapatkan dari model tersebut keduanya menghasilkan akurasi 97% dan tidak memberikan hasil yang signifikan sehingga tidak ada pengaruh yang signifikan jika memilih input shape berukuran 64x64 atau 128x128. Akan tetapi pada penelitian ini akan memilih input shape berukuran 128x128 karena jumlah time nya lebih kecil.

2) Pengaruh Jumlah Epoch

Epoch merupakan seluruh dataset yang telah melalui proses training model pada neural network sampai dikembalikan ke awal dalam satu putaran. Nilai epoch yang akan diujikan disini adalah kelipatan 10 dengan nilai terkecil 10 dan nilai terbesar adalah 110. Berikut hasil perbandingan epoch dari hasil training.

Tabel 4. Hasil Akurasi Pengaruh Jumlah Epoch

<i>Epoch</i>	<i>Accuracy Validation</i>	<i>Loss Validation</i>	<i>Time (second)</i>
10	93%	0.24	388 m/s
20	93%	0.20	221 m/s
30	97%	0.20	157 m/s
40	98%	0.1	390 m/s
50	97%	0.24	182 m/s
60	97%	0.17	162 m/s
70	97%	0.39	151 m/s
80	98%	0.22	146 m/s
90	96%	0.22	142 m/s
100	98%	0.33	150 m/s
110	97%	0.16	351 m/s

Berdasarkan tabel 4 diatas menggunakan nilai learning rate 0.001 didapatkan nilai akurasi yang cukup tinggi pada epoch ke 40, 80, 100 dengan nilai sebesar 98%. Jika dilihat pada tabel tersebut semakin tinggi epoch yang ditambahkan nilai loss function semakin tinggi, hal ini dapat disebabkan karena jumlah epoch yang terlalu banyak sehingga pada penelitian ini akan menggunakan nilai epoch sebanyak 40 dengan nilai loss validation 0.1.

3) Pengaruh Jumlah Konvolusi Layer

Layer konvolusi digunakan sebagai proses ekstraksi fitur pada citra gambar. Jumlah layer konvolusi dapat mempengaruhi tingkat akurasi dari model yang telah dibuat.

Tabel 5. Hasil Akurasi Pengaruh Jumlah Konvolusi Layer

<i>Jumlah Konvolusi</i>	<i>Accuracy Validation</i>	<i>Loss Validation</i>	<i>Time (second)</i>
3	97%	0.1	150 m/s
4	96%	0.1	180 m/s
5	50%	0.6	249 m/s

Berdasarkan tabel 5 menunjukkan bahwa semakin banyak layer konvolusi yang digunakan akan membuat model kehilangan keseimbangan sehingga menurunkan performa akurasi dari proses pelatihan model. Hal tersebut dapat dipicu karena banyaknya tahap ekstraksi fitur yang harus dilakukan oleh komputer, dan dapat menyebabkan overfitting karena model yang terlalu kompleks. Sehingga layer konvolusi yang akan digunakan sebanyak 3 layer dengan nilai akurasi 97% dan nilai loss validasi 0.1.

4) Pengaruh Pooling Layer

Pooling layer adalah proses yang dilakukan untuk mengurangi ukuran matriks dari hasil proses konvolusi. Pooling layer memiliki dua metode yang dapat digunakan, yaitu max-pooling dan average-pooling. Pada pengujian ini akan menggunakan jumlah epoch 40. Berikut adalah perbandingan keduanya :

Tabel 6. Hasil Akurasi Pengaruh Pooling Layer

<i>Pooling layer</i>	<i>Accuracy Validation</i>	<i>Loss Validation</i>	<i>Time (second)</i>
<i>max-pooling</i>	97%	0.1	349 m/s
<i>average-pooling</i>	97%	0.4	412 m/s

Berdasarkan tabel 6 merupakan hasil dari percobaan pooling layer yang telah diujikan. Pada tabel tersebut dapat dilihat bahwa tingkat akurasi menggunakan max-pooling dan average-pooling sama-sama menghasilkan akurasi 97%. Akan tetapi nilai loss nya lebih baik saat menggunakan max-pooling.

5) Pengaruh Skenario Data

Pada penelitian ini akan melakukan uji coba pada skenario jumlah data training dan testing pada model CNN yang akan digunakan. Model akan diuji menggunakan 40 epoch dengan learning rate 0.001. Jumlah dataset yang digunakan adalah 2000 dengan pembagian data sebagai berikut:

Tabel 7. Hasil Akurasi Pengaruh Skenario Data

<i>Skenario (Train:Test)</i>	<i>Accuracy Validation</i>	<i>Loss Validation</i>	<i>Time (second)</i>
70% : 30%	97%	0.2	201 m/s
80% : 20%	98%	0.1	192 m/s
90% : 10%	98%	0.3	178 m/s

Berdasarkan uji coba yang dilakukan pada tabel 7 menunjukkan bahwa semakin banyak data training yang digunakan pada penelitian maka hasilnya akan semakin baik. Hal tersebut karena semakin banyak model tersebut belajar maka akan semakin baik dalam mengenali pola citra gambar dengan akurat. akan tetapi pada scenario pembagian 90% data training dan 10% data testing mengalami kenaikan loss function, sehingga scenario yang akan digunakan pada penelitian ini adalah 80%:20%.

6) Pengaruh Learning Rate (40 epoch)

Tabel 8. Hasil Akurasi Pengaruh Learning Rate

<i>Learning Rate</i>	<i>Accuracy Validation</i>	<i>Loss Validation</i>	<i>Time (second)</i>
0.01	50%	0.6	356 m/s
0.001	98%	0.1	148 m/s
0.0001	96%	0.1	476 m/s

Berdasarkan tabel 8 menunjukkan bahwa nilai learning rate 0.01 mendapatkan akurasi yang kurang optimal, dengan tingkat akurasi sebesar 50%. Kemudian pada learning rate 0.001 mendapatkan akurasi yang optimal dengan akurasi 98%. Sedangkan pada nilai learning rate 0.0001 mendapatkan tingkat akurasi sebesar 96%. Nilai loss pada learning rate 0.01 cukup tinggi yaitu 0.6 hal tersebut dapat disebabkan karena adanya overfitting saat pelatihan model. Sedangkan pada nilai learning rate 0.0001 nilai loss nya tidak setinggi pada learning rate 0.01, tetapi penggunaan learning rate ini cukup lambat untuk memperkecil nilai loss.

Selain parameter model jenis dari inputan citra juga akan diuji berdasarkan pengaruh tipe data serta pengaruh warna citra.

1) Pengaruh Tipe Data

Citra input yang akan diuji disini adalah citra sel darah pasien dengan tipe data *png, dan *jpg. Citra input yang digunakan adalah 5 citra uji diluar data testing dan data training. Yang meliputi 3 data pasien positif malaria dan 2 data pasien normal. Maka didapatkan hasil sebagai berikut :

Tabel 9. Rata-rata akurasi pengaruh tipe data untuk citra uji

Tipe Data	Citra Uji	Akurasi	Rata-rata Akurasi
*png	Citra 1	96.5	96.3
	Citra 2	96.0	
	Citra 3	96.5	
	Citra 4	96.0	
	Citra 5	96.5	
*jpg	Citra 1	96.0	96.1
	Citra 2	96.0	
	Citra 3	96.25	
	Citra 4	95.75	
	Citra 5	96.5	

Berdasarkan uji coba yang dilakukan pada tabel 4.8 tipe data *png menghasilkan rata-rata akurasi sebesar 96,3% dan tipe data *jpg menghasilkan rata-rata akurasi sebesar 96,1%. Sehingga tipe data yang akan digunakan pada penelitian ini adalah *png.

2) Pengaruh Warna Citra

Warna pada citra juga akan mempengaruhi hasil akhir dari akurasi. Pada percobaan ini akan menguji citra sel darah pasien rgb dan grayscale. Citra input yang digunakan adalah 5 citra uji diluar data testing dan data training. Yang meliputi 3 data pasien positif malaria dan 2 data pasien normal. Maka didapatkan hasil sebagai berikut :

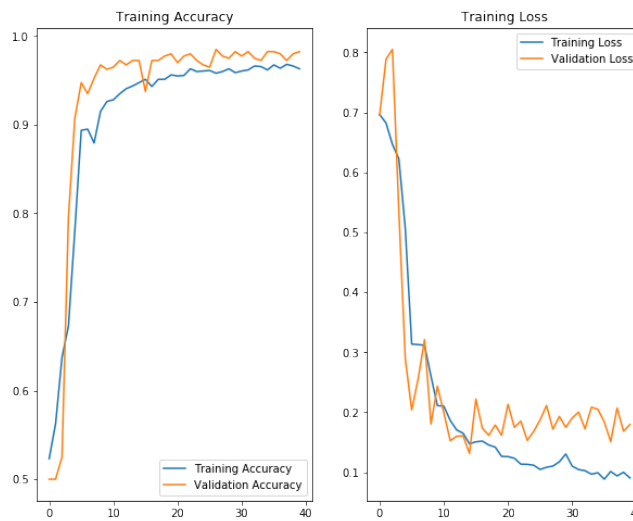
Tabel 10. Rata-rata akurasi pengaruh warna citra untuk citra uji

Warna Citra	Citra Uji	Akurasi	Rata-rata Akurasi
RGB	Citra 1	96.5	96.30
	Citra 2	96.0	
	Citra 3	96.5	
	Citra 4	96.0	
	Citra 5	96.5	
Grayscale	Citra 1	96.5	96.15
	Citra 2	96.0	
	Citra 3	96.25	
	Citra 4	95,75	
	Citra 5	96,25	

Berdasarkan uji coba yang dilakukan pada tabel 10 warna citra RGB menghasilkan rata-rata akurasi sebesar 96,30% dan warna citra grayscale menghasilkan rata-rata akurasi sebesar 96,15%. Sehingga performa saat menggunakan citra RGB lebih baik jika dibandingkan menggunakan citra grayscale.

3.4. Training dan Testing

Setelah menentukan parameter yang akan digunakan pada model, kemudian akan melihat bagaimana performa model tersebut pada proses training dan testing. Dataset berjumlah 2000 data dengan pembagian 80% data testing dan 20% data training, dengan total kelas 2. Pada proses training akan menggunakan 1600 data citra sel darah merah yang berformat *.png, serta 400 data citra untuk testing yang berformat *.png. Pelatihan model akan menggunakan CNN dengan jumlah epoch 40 dan nilai learning rate 0.001. Berikut adalah grafik dan hasil training :

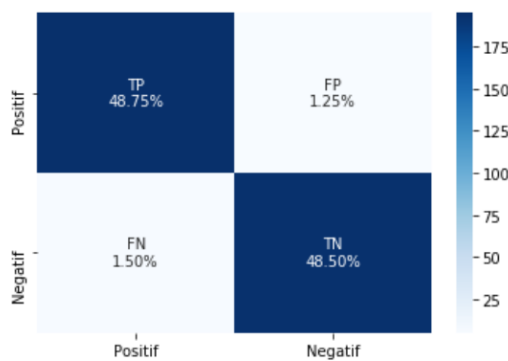


Gambar 5. Grafik Training Model

Berdasarkan gambar 5 akurasi yang didapatkan pada training model sebesar 97% dengan nilai loss function sebesar 0.1073. kemudian nilai akurasi pada data validasi sebesar 98% dengan nilai loss function sebesar 0.1879. proses training ini menggunakan input shape berukuran 128x128, dengan jumlah epoch 40, serta menggunakan learning rate sebesar 0.001.

3.5. Pengujian Model

setelah model selesai di running kemudian akan melalui tahap pengujian model menggunakan *confusion matrix*. Gambar 6 merupakan hasil *confusion matrix* pada tahap pelatihan model.



Gambar 6. Confusion Matrix pada tahap pelatihan

Kemudian untuk menguji performa dari model tersebut akan dilakukan proses uji evaluasi menggunakan data baru diluar data training dan data testing sebanyak 10 data uji, untuk setiap kelas citra sebanyak 5 gambar. Input gambar yang digunakan berukuran 128x128. Berikut adalah hasil confusion matriks nya.

Berdasarkan tabel 11 hasil prediksi model terhadap data testing baru menunjukkan hasil yang sangat baik. Prediksi terhadap pasien positif di klasifikasikan sebagai pasien positif yang artinya hasil klasifikasi tersebut benar. Prediksi pasien negatif di klasifikasikan sebagai pasien negatif yang artinya hasil klasifikasi tersebut benar. Berikut adalah perhitungan matriksnya :

Tabel 11. Confusion Matrix pada tahap uji evaluasi model

<i>Matriks</i>	<i>Predict Class</i>	
	POSITIF	NEGATIF
POSITIF	10	0
NEGATIF	0	10

$$\begin{aligned} \text{Akurasi} &= (TP+TN)/(P+N) \times 100\% \\ &= (10+10)/20 \times 100\% = 100\% \end{aligned}$$

Jadi akurasi yang dihasilkan oleh model CNN menggunakan data uji berupa citra sel darah merah sebesar 100%.

Pada kebutuhan hasil memiliki beberapa kebutuhan yang akan digunakan pada sistem seperti sistem dapat mengklasifikasi citra darah dari pasien positif malaria dan pasien normal menjadi positif dan negatif malaria. Kemudian sistem dapat menampilkan hasil akurasi dari identifikasi pasien malaria menggunakan model CNN yang telah digunakan.

4. Kesimpulan

Berdasarkan hasil analisis yang telah dilakukan kemudian didapatkan kesimpulan sebagai berikut:

1. Model CNN yang digunakan pada penelitian ini menggunakan *input shape* berukuran 128x128, dengan jumlah *epoch* 40, serta menggunakan *learning rate* sebesar 0.001. Jumlah *dataset* yang digunakan sebanyak 2000 data dengan pembagian 80% data *training* dan 20% data *testing*.
2. Akurasi pada model CNN ini menghasilkan 97% untuk proses *training* serta 98% untuk proses *testing*. Sehingga dapat disimpulkan bahwa penerapan metode CNN untuk identifikasi infeksi penyakit malaria berdasarkan citra darah dapat dikatakan baik.
3. Akurasi model CNN menggunakan data uji baru menghasilkan akurasi sebesar 100%. Pengujian ini menggunakan 20 data citra dengan masing-masing kelas sebanyak 10 data.
4. Model CNN yang diusulkan dapat digunakan untuk klasifikasi citra darah pada pasien malaria.
5. Penelitian ini telah melalui tahap *trial* dan *error*, berikut hasil analisis dari beberapa *parameter* model yang telah melalui proses *trial* dan *error*:
 - a. Eksperimen terhadap *input shape* berukuran 64x64 dan 128x128 sama-sama menghasilkan nilai akurasi 97%.
 - b. Eksperimen terhadap jumlah *epoch* mendapatkan hasil akurasi tertinggi sebesar 98% pada *epoch* ke 40.
 - c. Eksperimen terhadap penggunaan *layer konvolusi* mendapatkan hasil akurasi terbaik saat menggunakan 3 *layer konvolusi*, dengan akurasi 97%.
 - d. Eksperimen terhadap penggunaan *pooling layer* dibagi menjadi dua yaitu *maxpooling* dan *average pooling*. Pada hasil eksperimen di dapatkan bahwa nilai akurasi yang dihasilkan sama-sama menghasilkan akurasi sebesar 97%, akan tetapi nilai *loss function* pada *average pooling* cukup tinggi sehingga performa menurun.
 - e. Eksperimen terhadap skenario data dibagi menjadi 3, yaitu pembagian data *train* dan *test* sebesar 70:30, 80:20, dan 90:10. Kemudian didapatkan hasil eksperimen bahwa nilai akurasi dengan pembagian data *train* dan *test* dengan perbandingan 80%:20% menghasilkan nilai akurasi yang cukup baik yaitu 98%.

- f. Eksperimen terhadap penggunaan *learning rate* mendapatkan hasil akurasi tertinggi pada nilai 0.001 dengan akurasi sebesar 98%.
- g. Eksperimen terhadap penggunaan citra input berdasarkan tipe data menguji citra *.png dan *.jpg masing-masing mendapatkan hasil rata-rata akumulasi sebesar 96,3% dan 96,1%. Dari rata-rata akumulasi tersebut dapat disimpulkan bahwa penggunaan png dan jpg sama sama menghasilkan nilai akurasi yang baik, sehingga kedua tipe data tersebut dapat digunakan. Tetapi lebih diutamakan apabila menggunakan png.
- h. Eksperimen terhadap penggunaan citra input berdasarkan warna citra menguji citra RGB dan Grayscale masing-masing mendapatkan hasil rata-rata akumulasi sebesar 96,30% dan 96,15%. Sehingga jenis warna citra yang akan digunakan pada penelitian yaitu citra RGB.

Daftar Pustaka

- [1] Tangpukdee, N., Duangdee, C., Wilairatana, P., & Krudsood, S. (2009). *Malaria Diagnosis: A Brief Review. The Korean Journal of Parasitology*, 47(2), 93. doi:10.3347/kjp.2009.47.2.93
- [2] Rinawati, W., & Henrika, F. (2019). Diagnosis Laboratorium Malaria. *J Indon Med Assoc*, 69(10), 327–335.
- [3] World Health Organization (WHO). 2019. World Malaria Report 2019. Technical report. World Health Organization Available at <https://www.who.int/publications/i/item/world-malaria-report-2019>.
- [4] Kusuma, W., Lestari, A. A. W., Herawati, S., Putu, I. W., & Yasa, S. (2000). Pemeriksaan Mikroskop Dan Tes Diagnostik Cepat Dalam Mengakkan Diagnosis Malaria. *Work*, 4(021), 57946053. <http://www.jstor.org/stable/2883974> %5Cn<http://www.jstor.org/stable/2883974>
- [5] Putra, S. R. (2015). *Implementasi Convolutional Neural Network Untuk Klasifikasi Obyek Pada Citra*. http://repository.its.ac.id/71292/1/5111100076-Undergraduate_Thesis.pdf
- [6] Sabir, M. I. (2019). Pengaruh Image Engagement pada Aplikasi Parasit Malaria. *Jurnal Penelitian Enjiniring*, 22(1), 34–37. <https://doi.org/10.25042/jpe.052018.06>
- [7] Liang, Z., Powell, A., Ersoy, I., Poostchi, M., Silamut, K., Palaniappan, K., ... Thoma, G. (2016). *CNN-based image analysis for malaria diagnosis. 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. doi:10.1109/bibm.2016.7822567
- [8] Biantong, T. R., Furqon, M. T., & Soebroto, A. A. (2019). Implementasi Metode Support Vector Machine Untuk Klasifikasi Jenis Penyakit Malaria. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya*, 3(2), 1215–1224.
- [9] Cirean, D., Meier, U., & Schmidhuber, J. (2012). *Multi-column deep neural networks for image classification. 2012 IEEE Conference on Computer Vision and Pattern Recognition*. doi:10.1109/cvpr.2012.6248110
- [10] Astini, N. K. S. (2018). Segmentasi Citra Parasit Malaria Plasmodium Vivax Dengan Menggunakan Metode Haar Cascade. *Jl. Udayana Kampus Tengah*, 0362, 27213. <http://pti.undiksha.ac.id/senapati>
- [11] Gopakumar, G. P., Swetha, M., Sai Siva, G., & Sai Subrahmanyam, G. R. K. (2018). Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner. *Journal of Biophotonics*, 11(3). <https://doi.org/10.1002/jbio.201700003>
- [12] Sabilla, I. A. (2020). *Arsitektur Convolutional Neural Network (Cnn) Untuk Klasifikasi Jenis Dan Kesehatan Buah Pada Neraca Buah. Tesis, 201510370311144*, 1–119. https://repository.its.ac.id/73567/1/05111850010020-Master_Thesis.pdf
- [13] Putri, O. N. (2020). Implementasi metode cnn dalam klasifikasi gambar jamur pada analisis image processing.

-
- [14] NURHIKMAT, T. (2018). Implementasi Deep Learning Untuk Image Classification Menggunakan Algoritma Convolutional Neural Network (Cnn) Pada Citra Wayang Golek. *Universitas Islam Indonesia, 10*(2), 1–15. <https://doi.org/10.13140/RG.2.2.10880.53768>
- [15] Andika, L. A., Pratiwi, H., & Handajani, S. S. (2019). Lingga Aji Andika 1 , Hasih Pratiwi 2 , and Sri Sulistijowati Handajani 3 1. *Indonesian Journal of Statistics and Its Applications, 3*(3), 331–340
- [16] Faiz, F. A. (2019). *Klasifikasi jenis kulit hewan asli dengan kulit imitasi berdasarkan citra tekstur permukaan samak kulit menggunakan convolutional neural network.*
- [17] Sutrimah. (2017). Validitas Diagnosis Suspek Malaria Pada Pasien Demam Di Puskesmas Kaligesing Kabupaten Purworejo. 9–26. <https://repository.unimus.ac.id>