

# Prediksi Parasit Plasmodium pada Citra Mikroskopis Sel Darah Merah dengan Convolutional Neural Networks

*by Jullend Gatc*

---

**Submission date:** 06-Aug-2022 06:48PM (UTC+0700)

**Submission ID:** 1879430134

**File name:** 4\_5007\_Publikasi.pdf (817.87K)

**Word count:** 4577

**Character count:** 28437

## Prediksi Parasit Plasmodium pada Citra Mikroskopis Sel Darah Merah dengan *Convolutional Neural Networks*

Jullend Gate<sup>1</sup>, Febri Maspiyanti<sup>2</sup>

<sup>1,2</sup>Sistem Informasi, Fakultas Ilmu Komputer dan Desain, Kalbis Institute  
Jalan Pulomas Selatan Kav.22, Jakarta, Indonesia

Email: <sup>1</sup>jullend.gate@kalbis.ac.id, <sup>2</sup>febri.maspiyanti@univpancasila.ac.id

**Abstract. Prediction of Plasmodium Parasites on Microscopic Image of Red Blood Cells with Convolutional Neural Networks.** Malaria is a deadly disease that attacks human blood cells caused by the plasmodium parasite. The need for fast and accurate detection in the diagnosis of malaria is certainly necessary to reduce the mortality rate from this disease. The technique for detecting the presence of Plasmodium parasites that has been widely used in routine examinations is using a microscope. With the presence of experienced medical experts, it is easy to detect the presence of Plasmodium in the blood. However, the weakness of this technique is that it relies heavily on the presence and competence of medical experts because the accuracy of microscopic examination results can decrease from 64% to 95%. The purpose of this study was to build a predictive model to classify Plasmodium parasites on red blood cell images with a good degree of accuracy with the Convolutional Neural Network algorithm. The test results show good accuracy results, namely the model of the CNN Algorithm gives an accuracy result of 97.96% and a loss of 0.06 with an average computation time of about 121 seconds/epoch.

**Keywords:** malaria, optimization, Plasmodium, neural network, CNN.

**Abstrak.** Malaria merupakan salah satu penyakit mematikan yang menyerang sel darah manusia yang disebabkan parasit plasmodium. Kebutuhan deteksi cepat dan akurat dalam diagnosis malaria tentunya sangat diperlukan untuk menekan angka kematian dari penyakit ini. Teknik deteksi keberadaan parasit Plasmodium yang telah banyak digunakan dalam pemeriksaan rutin adalah dengan menggunakan mikroskop. Dengan adanya tenaga ahli medis yang berpengalaman, pendeteksian keberadaan Plasmodium dalam darah pun dengan mudah dilakukan. Namun kelemahan dari teknik ini adalah sangat bergantung pada keberadaan dan kompetensi dari tenaga ahli medis karena akurasi dari hasil pemeriksaan mikroskop dapat menurun 64% sampai dengan 95%. Tujuan penelitian ini adalah membangun sebuah model prediksi untuk mengklasifikasikan parasit plasmodium pada citra sel darah merah dengan tingkat akurasi yang baik dengan algoritma Convolutional Neural Network. Hasil pengujian memperlihatkan hasil akurasi yang baik yaitu model dari Algoritma CNN ini memberikan hasil akurasi yaitu 97,96% dan loss 0,06 dengan rata-rata waktu komputasi sekitar 121 detik/epoch.

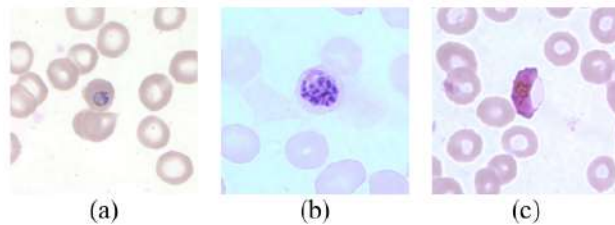
**Kata kunci:** malaria, deep learning, Plasmodium, neural network, CNN.

### 1. Pendahuluan

Kebutuhan deteksi cepat dan akurat dalam diagnosis malaria tentunya sangat diperlukan untuk menekan angka kematian dari penyakit ini. Beberapa penelitian terkait di bidang biomedis dan ilmu komputer sudah banyak dilakukan dalam pengembangan suatu metode pendagnosa penyakit malaria yang sifatnya sensitif pada gejala-gejala klinis sangatlah diperlukan. Beberapa metode diagnostik konvensional yang digunakan di lab kesehatan terkait penyakit malaria ini telah dilakukan seperti pengujian biokimia (*parasitic lactate dehydrogenase LDH*), metode fluorometrik (*Fluorocrom Hoechi 33458-AND of Plasmodium falciparum*), atau penggabungan radio berlabel precursor (*Hypoxanthine triiated, actual gold standard in malaria drug research*). Pemeriksaan dengan teknik ini merupakan salah satu teknik yang akurat, namun pemeriksaan dengan metode ini dikenal mahal dan tidak ditemukan di fasilitas kesehatan pada umumnya [1]. Teknik lain untuk mendeteksi keberadaan parasit

Plasmodium adalah dengan menggunakan *Rapid Diagnostic Test* (RDT) yang dijual secara komersial, tetapi RDT tidak mampu mendeteksi semua jenis Plasmodium. Kebanyakan hanya yang berjenis Plasmodium falciparum [4]. Selain itu, luaran yang dihasilkan dengan pemeriksaan ini hanya sebatas positif dan negatif, belum dapat memberikan informasi terkait jenis parasit ataupun fase *life-cycle* dari parasit tersebut. Teknik lainnya yang juga sering digunakan pada fasilitas kesehatan dalam pemeriksaan malaria adalah pemeriksaan mikroskopis. Pemeriksaan ini dikenal sebagai salah satu pemeriksaan yang akurat dan murah, namun memerlukan tenaga ahli dalam proses pemeriksaan [2]. Dengan adanya tenaga ahli medis yang berpengalaman, pendeteksian keberadaan Plasmodium dalam darah pun dengan mudah dilakukan. Namun kelemahan teknik ini adalah sangat bergantung pada keberadaan dan kompetensi tenaga ahli medis karena akurasi hasil pemeriksaan mikroskop dapat menurun 64% sampai dengan 95% bila jumlah parasit dalam darah menurun atau pada kasus infeksi campuran dimana hanya satu parasit yang terdeteksi [3].

Deteksi malaria yang cepat dan akurat sebenarnya menjadi salah satu isu pada pemerintahan beberapa tahun terakhir. Hal ini dapat dilihat dari beberapa program dan kegiatan pemerintah untuk menekan permasalahan malaria, khususnya pada kelompok rentan seperti ibu hamil dan bayi [4], terlebih di wilayah yang memiliki angka kasus terbanyak dan fasilitas dan tenaga kesehatan yang minim. Beberapa alat dan algoritma sudah banyak dikembangkan bahkan sampai level produksi, namun beberapa peralatan ini tentu tidak dimiliki oleh fasilitas kesehatan, terlebih pada level puskesmas. Hal ini menjadi motivasi bagi tim peneliti untuk mengembangkan sistem yang cepat dan akurat dalam melakukan deteksi malaria dengan pendekatan komputasional, serta sistem yang adaptif dengan kondisi di lapangan yang memiliki beberapa permasalahan seperti data *slide* darah yang terkadang dibuat dengan tidak standar dan menyebabkan gambar yang dilihat pada mikroskop menjadi tidak jelas [5]. Selain itu, sistem komputasional yang dapat membedakan RBC yang terinfeksi parasit plasmodium dan *parasite stages* seperti *Trophozoite*, *Schizont*, dan *Gametocyte* (lihat Gambar 1) dengan akurasi yang baik tetap menjadi tantangan [6] dan juga secara tipe, parasit plasmodium terdiri dari 5 jenis seperti plasmodium *falciparum*, *vivax*, *ovale*, *malariae* dan *knowlesi* yaitu varian endemik yang terdapat di pulau Kalimantan.



**Gambar 1. Plasmodium Falciparum (a) Trophozoite, (b) Schizont, (c) Gametocyte [1].**

Peningkatan algoritma sistem deteksi untuk mendapatkan akurasi yang baik tentunya merupakan tantangan bagi peneliti, khususnya di bidang biomedis. Salah satu hal yang harus diperhatikan adalah proses segmentasi. Segmentasi merupakan tugas yang krusial dalam pengolahan citra medis, karena proses segmentasi secara langsung berdampak kepada proses *post-processing*, seperti *image analysis* dan *feature extraction* [7]. Dalam hal ini, peningkatan algoritma menggunakan teknik *Deep Learning* dapat dijadikan referensi untuk peningkatan akurasi deteksi yang lebih baik. Penelitian ini bertujuan meningkatkan akurasi sistem deteksi malaria menggunakan pendekatan algoritma *Convolutional Neural Network* (CNN). Eksekusi CNN juga dianggap kompleks dan memerlukan infrastruktur komputasi yang baik agar dapat memberikan hasil yang cepat dan akurat. Penelitian ini menggunakan *library* TensorFlow Keras dimana proses training data dapat dilakukan secara *cloud* dengan bantuan komputasi menggunakan GPU. Penelitian ini diharapkan dapat meningkatkan akurasi dan mampu

diimplementasikan untuk pendeteksian dini penyakit malaria secara cepat, tepat, dan akurat di Indonesia.

## 2. Tinjauan Pustaka

Pada penelitian sebelumnya [8], penelitian deteksi malaria berhasil dengan baik dan dapat melakukan deteksi terhadap parasit Plasmodium pada citra *Red Blood Cell* menggunakan algoritma *double thresholding* untuk memisahkan parasit dan *background* serta informasi lain yang tidak diperlukan. Penelitian ini [8] dibuktikan dengan perolehan *Predictive Positive Value* (PPV) dan tingkat Sensitivity sebesar 92,85% dan 85,52% secara berurutan. Namun pada penelitian tersebut ada beberapa kesalahan deteksi pada sistem diakibatkan oleh beberapa penampakan sel darah putih menyerupai sel darah merah yang mengandung parasit malaria. Akhirnya, ditemukan bahwa angka *false positif* dan *negatif* diperoleh dari keberadaan sel darah putih di citra RBC. Berdasarkan kesalahan tersebut, penelitian selanjutnya mencoba menggunakan metode kombinasi antara *Double Thresholding* dan *Binary Large Objecy* (BLOB) [9]. Keberadaan sel darah putih dapat dibedakan secara ukuran sel karena sel darah putih akan relatif lebih besar dibandingkan ukuran sel darah merah. Penelitian tersebut menggunakan fitur *shape* dan luas area dengan BLOB untuk melakukan segmentasi sel darah merah dan sel darah putih dengan baik.

Penelitian selanjutnya memfokuskan untuk dapat membuat model deteksi sistem yang dapat mengklasifikasikan fase hidup parasit [10]. Jumlah data yang digunakan lebih banyak dan didapatkan dari fase hidup parasit yang berbeda, yaitu fase *Trhopozoite*, *Gametocyte* dan *Schizont*. Kegiatan penelitian tersebut juga dirasa penting untuk mengetahui fase parasit pada tubuh manusia agar penanganan kesehatan yang lebih presisi ke depannya. Namun peningkatan akurasi deteksi tetap merupakan sesuatu yang penting untuk diteliti, mengingat tidak semua data sampel darah yang diperoleh dengan kondisi baik karena beberapa faktor seperti pengalaman tenaga medis serta ketersediaan *slide* darah yang seringkali tidak dibuat dengan standar [11].

Berdasarkan penjelasan di atas maka peningkatan presisi dan akurasi merupakan salah satu langkah penting yang harus dilakukan dalam penelitian ini. Peningkatan akurasi dari hasil prediksi dan klasifikasi Parasit *Plasmodium Falcifarum* dilakukan dengan menggunakan algoritma *Convolutional Neural Network* (CNN) pada citra *red blood cell*, khususnya pada riil yang sudah dikumpulkan langsung dari Dinas Kesehatan Provinsi Kalimantan Tengah. Tujuan penelitian ini adalah membangun sebuah model prediksi untuk mengklasifikasikan parasit plasmodium falcifarum pada citra sel darah merah dengan tingkat akurasi yang baik. Selain itu juga penelitian ini juga diharapkan dapat mengidentifikasi jenis parasit serta fase dari parasit malaria tersebut. Output penelitian ini juga bertujuan menjadi salah satu acuan peneliti lainnya dalam penelitian malaria dari aspek ilmu komputer dan biomedis dalam pembelajaran parasit plasmodium.

## 3. Metodologi Penelitian

Penelitian ini bersifat eksperimental dan akan dilakukan dalam beberapa tahapan seperti studi literatur, pengumpulan data, membangun model sistem deteksi, uji coba model, evaluasi, eksperimen dan komparasi algoritma, dan pencatatan hasil serta penarikan kesimpulan untuk sistem hasil dan kesimpulan akan dilakukan pada tahap akhir. Langkah-langkah penelitian yang dilakukan tercantum pada Gambar 2.

### 3.1. Akuisisi Data

Tahapan *data acquisition* menggunakan *Personal Computer* (PC) yang digunakan untuk melakukan eksperimen serta membangun sistem deteksi (lihat Gambar 3). Mikroskop yang digunakan pada penelitian ini adalah mikroskop monokuler dengan menggunakan Lensa okuler dengan pembesaran 10 kali. Selain itu lensa obyektif dengan pembesaran 10, 40, dan 100 kali. Kamera tambahan *Eyepiece Cam* digunakan sebagai alat pengamat pada mikroskop yang dihubungkan dengan *computer* atau *notebook* menggunakan kabel *Universal Serial Bus* (USB). Kamera ini memiliki fungsi layaknya mata manusia yang mengamati obyek pada *slide* sediaan

yang diamati. Spesifikasi kamera *Eyepice Cam* yang digunakan yaitu: sensor: Aptina (color); Sensor Type: CMOS; Sensor Size: 4.73x3.52mm (1/3.2" class); Pixel Size: 2.8; Resolution: 2MP; Frame Rate: 20fps @ 800x600, 7.5fps @ 1280x960, 5fps @ 1600x1200



Gambar 3. Akusisi Data Citra dari eyepiece mikroskop

### 3.2. Data

Data *slide* darah yang yang digunakan berasal dari berbagai Puskesmas dan Dinas Kesehatan di wilayah Kalimantan Tengah, seperti Teluk Bogam, Murung Raya, dan Lamandau. Data yang diperoleh berupa *slide* darah yang dikemas dalam kotak penyimpanan seperti terlihat pada Gambar 4. Kondisi darah pada masing-masing *slide* tersebut berbeda-beda tergantung dari teknik pembuatan *slide* darah yang dilakukan oleh petugas medis pada masing-masing wilayah (ada yang tipis, ada yang tebal). Tidak semua memiliki kualitas baik sehingga tidak layak untuk digunakan dalam penelitian.

Penelitian menggunakan sekitar 100 *slide* darah yang diambil secara random dan diberikan label (*positive* dan *negative*) oleh pakar, yaitu oleh tenaga Lab Kesehatan Daerah Kalimantan Tengah. Masing-masing *slide* darah dapat menghasilkan lebih dari 10 citra yang diambil dari beberapa *capture perspective* dan sesuai dengan arahan pakar dalam proses *capture* citra. Citra ini digunakan sebagai data *training* untuk membentuk model CNN dan *testing*. Seluruh data diambil dari *eyepiece camera* yang memiliki resolusi citra 800x600. Sumber data lain sebagai pembandingan dan data latih ialah sumber data terbuka Kaggle, Malaria Dataset (*National Library of Medicine*), serta data publik dari komunitas peneliti *Lister Hill National Center for Biomedical Communications* (LHNCBC) sebagai acuan pembentukan model prediksi [12]. Seluruh data dikumpulkan dalam *folder* yang sama dengan penamaan *file* mengikuti format dari LHNCBC agar komputasi program dapat dieksekusi dengan baik.



Gambar 4a. Box Data



Gambar 4b. Slide Darah

Gambar 4. Data *Slide* Darah Penelitian.

Total data yang digunakan adalah 27.558 citra darah yang terdiri dari citra yang terinfeksi parasit dan citra yang tidak terinfeksi (*uninfected*). Label parasit (*Ground-Truth*) darah diperoleh dari LHNCBC [12]. Kami juga menyertakan file CSV yang berisi ID Pasien ke pemetaan sel untuk kelas yang diparasit dan tidak terinfeksi [11]. File CSV untuk kelas parasit berisi 151 entri ID pasien. Gambar *slide* untuk ID pasien parasit "*C47P8thinOriginal*" dibaca

dari dua model mikroskop yang berbeda. File CSV untuk kelas yang tidak terinfeksi berisi 201 entri karena sel normal dari *slide* pasien yang terinfeksi juga masuk ke kategori sel normal.

### 3.3. Algoritma CNN Pada Citra RBC

Seluruh data masukan akan dimasukkan ke dalam model *training* dengan algoritma *Convolutional Neural Network* (CNN) yang merupakan salah satu algoritma *Deep Learning* yang dapat mengambil gambar input, menetapkan kepentingan (bobot dan bias yang dapat dipelajari) untuk berbagai aspek/ objek dalam gambar dan dapat membedakan satu dari yang lain [13]. *Preprocessing* yang diperlukan dalam CNN jauh lebih rendah dibandingkan dengan algoritma klasifikasi lainnya. Sementara dalam metode primitif, proses rekayasa dilakukan dengan pelatihan yang cukup. CNN memiliki kemampuan untuk mempelajari filter/karakteristik ini [14].

Kernel merupakan elemen kunci pada beberapa algoritma *computer vision*, ini tentunya tidak hanya digunakan pada CNN saja namun pada algoritma lainnya yang menggunakan data citra. Kernel *Convolutional* akan membuat sebuah matriks kecil dari data citra (bisa menggunakan istilah filter atau kernel), dimana nilai matriks baru ini merupakan hasil transformasi dari filter pada citra sebelumnya.

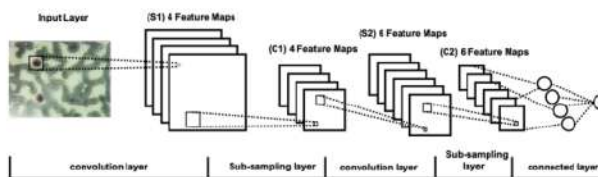
$$G[m, n] = (f * h)[m, n] = \sum_i \sum_k h[j, k] f[m - j, n - k]$$

Simbol  $f$  merupakan citra masukan dan  $h$  merupakan kernel. Sedangkan  $m$  dan  $n$  merupakan indeks matriks kolom dan baris secara berurutan. Masing-masing data akan disimpan pada *feature map* yang menggunakan beberapa model filter, seperti 3x3, 5x5, dan 9x9.

Ketika membuat arsitektur desain dari CNN, ukuran dimensi spasial ditentukan dari *feature map* yang digunakan. Penentuan output matrik menggunakan perhitungan sebagai berikut:

$$n_{out} = \left\lfloor \frac{n_m + 2p - f}{s} + 1 \right\rfloor$$

Simbol  $p$  merupakan *padding* atau posisi piksel dan  $f$  merupakan dimensi filter yang digunakan. Secara singkat, arsitektur serta *layer key* dalam *model* CNN termasuk lapisan konvolusi dan penyatuan, seperti yang ditunjukkan pada Gambar 5 berikut.



Gambar 5. Arsitektur CNN pada Data Citra RBC

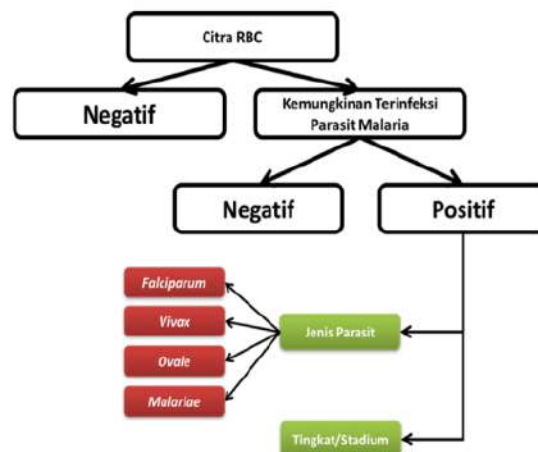
Berdasarkan arsitektur dalam kode ini, model CNN memiliki empat lapisan konvolusi dan penyatuan, diikuti oleh dua lapisan padat, dan *dropout* untuk regularisasi. Penelitian ini menerapkan dua strategi paling populer untuk pembentukan model dengan menggunakan CNN, yaitu *pre-trained* model sebagai *feature extractor* dan *pre-trained model with fine-tuning*.

Model *CNN VGG-19* yang digunakan telah dilatih sebelumnya dan dikembangkan oleh *Visual Geometry Group (VGG)* dari Universitas Oxford untuk eksperimen yang akan dilakukan. Model *VGG-19* memang sering kali digunakan untuk dataset citra yang besar (*ImageNet*) dengan kategori gambar yang beragam [15].

Model CNN yang diusulkan di-*training* dan dievaluasi menggunakan Google Colab, yaitu *tools* Python berbasis *cloud* yang tersedia untuk akses gratis. Colab menyediakan sistem yang telah dikonfigurasi sebelumnya untuk melatih dan mengevaluasi algoritma *deep learning* dan memberikan akses ke Graphical Processing Unit (GPU) dengan performa yang tinggi. Saat ini, *library* ini memberikan akses ke GPU NVIDIA Tesla P100 16 GB. Seluruh *library* diinstal dengan paket *pip* (PIP) di Python 3 dan menggunakan *library* Keras API dan TensorFlow di *backend* [16].

#### 4. Hasil dan Diskusi

Penelitian ini bersifat eksperimental dan dilakukan secara bertahap dengan melakukan beberapa kali percobaan dengan kombinasi jumlah *Epoch Batch Size* dan jumlah iterasi. Oleh karena itu sistem akan dievaluasi terus menerus dan diharapkan nantinya akan memperoleh presisi dan akurasi yang tinggi dalam mengidentifikasi parasit penyebab penyakit malaria, serta memberikan informasi baik dari jenis dan stadium parasit berdasarkan data sampel yang digunakan. Perhatikan Gambar 6.



Gambar 6. Klasifikasi Model Sistem

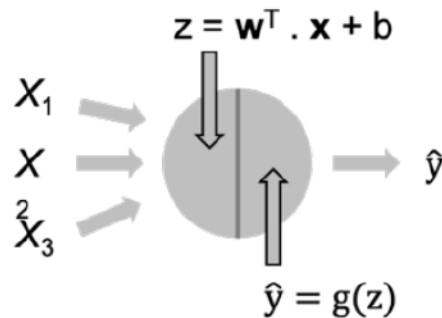
Tujuan dari penelitian ini yaitu membangun model yang dapat membedakan parasit plasmodium ke dalam empat jenis parasit, yaitu *plasmodium falciparum*, *plasmodium vivax*, *plasmodium malariae*, dan *plasmodium ovale*, serta stadium dari masing-masing data. Selain menggunakan data dari Dinas Kesehatan Provinsi Kalimantan Tengah yang sudah diberikan label, penelitian ini juga menggunakan data malaria dari sumber data terbuka dan kredibel lainnya, seperti official NIH [17] ([website: https://ceb.nlm.nih.gov/](https://ceb.nlm.nih.gov/)) yang terdiri lebih dari 20.000 data yang terbagi menjadi data *image* yang terinfeksi (*infected*) dan data yang tidak terinfeksi (*uninfected*).

Hasil klasifikasi model CNN yang diajukan juga dapat menunjukkan kategori jenis parasit ke dalam empat kelompok parasit: *plasmodium falciparum*, *vivax*, *ovale*, dan *malariae* sesuai data latih. Pada penelitian ini, model masih tidak dapat mengelompokkan jenis *plasmodium knowlesi* (varian plasmodium endemik di pulau Kalimantan) karena keterbatasan data latih. Oleh sebab itu, penelitian tidak menggunakan kelompok ini pada model yang digunakan. Model yang dibangun masih memiliki kekurangan untuk klasifikasi tingkat stadium dari jenis parasit *plasmodium falciparum*. Terdapat beberapa *false positif* dari data yang salah klasifikasi, seperti stadium *tropozoit*. Secara umum, fase awal *tropozoit* memiliki satu titik kromatin dengan sitoplasma parasit berbentuk cincin, dan vakuola terlihat jelas. Namun, tahap akhir *tropozoit* memiliki dua titik kromatin yang melekat pada cincin di dalam satu vakuola

parasit. Dengan demikian kondisinya mirip dengan tahap pembelahan yang memiliki titik kromatin antara 2 sampai 32. Perbedaannya adalah setiap titik kromatin dalam skizon memiliki satu vakuola sendiri (setiap titik kromatin dipisahkan dari titik kromatin lainnya oleh vakuola). Vakuola ini memiliki intensitas tingkat abu-abu yang rendah dan mudah dibedakan dengan metode penelitian ini, sehingga salah mengklasifikasikan tahap akhir *tropozoit* menjadi tahap skizon.

#### 4.1. Pembentukan Model CNN

Tahapan konvolusi atas volume serta *feature* ap merupakan konsep yang sangat penting pada CNN, dimana ini tidak hanya bekerja pada citra RGB namun juga dapat menerapkan beberapa filter pada satu *layer*. Pada *library* tensor flow yang digunakan, setiap neuron akan menerima 1 set nilai  $x$  sebagai input (berisi data *training*) dan menghitung nilai  $y$ -hat sebagai nilai yang akan diprediksi [16]. Setiap unit akan memiliki parameternya sendiri-sendiri yaitu  $w$  yaitu bobot vektor pada kolom dan  $b$  (bias) yang kemungkinan akan selalu berubah selama proses iterasi *learning*. Dalam setiap iterasi, neuron akan melakukan perhitungan *weighted average* dari nilai vektor  $x$ . Perhitungan ini akan terus dilakukan fungsi linier  $g$  (Lihat Gambar 7.)



Gambar 7. Ilustrasi proses *training* dengan Single Neuron

Pada penelitian ini, khususnya pada proses pembentukan model deteksi menggunakan CNN, perlu menggunakan data latih (*data training*) dan pengujian kinerja model pada data yang belum diberikan label. Pembentukan model kali ini menggunakan *split* 60:10:30 untuk masing-masing set data *training*, data validasi, dan data *testing*. Set data *training* dan validasi selama *training* serta pemeriksaan kinerja model pada dataset uji dilakukan pada penelitian ini.

Berdasarkan Gambar 8, terlihat beberapa perbedaan kecil antara malaria dan gambar sel yang sehat. Selanjutnya, dibuat model CNN untuk mempelajari pola-pola ini selama *training* model.

Dalam fase *training* model, akan dibangun empat model, lalu proses *train* data tersebut dengan data *training*, dan membandingkan hasil *training* dengan menggunakan data valid (*Gold Standard*). Model ini akan disimpan, lalu pada tahap evaluasi model akan digunakan model CNN VGG-19 yang telah dilatih sebelumnya dan dikembangkan oleh *Visual Geometry Group* (VGG) dari Universitas Oxford untuk eksperimen yang akan dilakukan. Model VGG-19 memang sering kali digunakan untuk dataset besar (*ImageNet*) dengan kategori gambar yang beragam [18].

#### 4.2. Setting Parameter

Hasil akurasi yang baik merupakan tujuan dari penelitian ini agar dapat bermanfaat jika diterapkan pada kasus nyata deteksi malaria dengan menggunakan model deteksi [19]. Upaya yang ingin dilakukan ke depannya adalah melakukan eksperimen tambahan dengan mencoba melakukan modifikasi dari arsitektur model CNN yang dibuat sesuai dengan layer dan



parameter yang sudah ditentukan sebelumnya. Total parameter yang digunakan untuk proses *training* ini mencapai 238.811 dan semuanya sukses digunakan sebagai parameter *training*. Tabel 1 menunjukkan parameter CNN.

**Tabel 1. CNN Parameter**

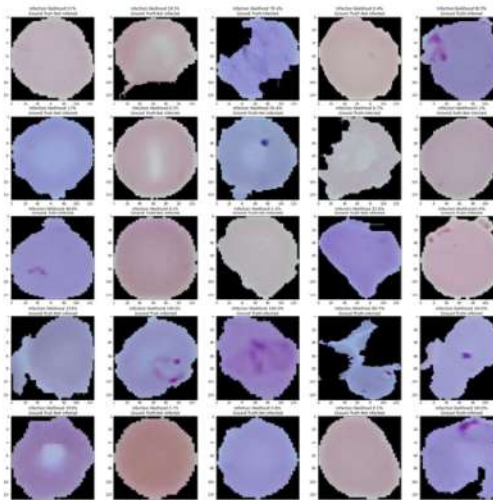
| Layer (type)               | Output Shape                   | Param # |
|----------------------------|--------------------------------|---------|
| conv2d_28 (Conv2D)         | (None, 128, 128, 10)           | 760     |
| max_pooling2d_28           | (MaxPooling (None, 64, 64, 10) | 0       |
| conv2d_29 (Conv2D)         | (None, 64, 64, 20)             | 1820    |
| max_pooling2d_29           | (MaxPooling (None, 32, 32, 20) | 0       |
| conv2d_30 (Conv2D)         | (None, 32, 32, 30)             | 5430    |
| max_pooling2d_30           | (MaxPooling (None, 16, 16, 30) | 0       |
| flatten_10 (Flatten)       | (None, 7680)                   | 0       |
| dense_37 (Dense)           | (None, 30)                     | 230430  |
| dropout_19 (Dropout)       | (None, 30)                     | 0       |
| dense_38 (Dense)           | (None, 10)                     | 310     |
| dropout_20 (Dropout)       | (None, 10)                     | 0       |
| dense_39 (Dense)           | (None, 5)                      | 55      |
| dense_40 (Dense)           | (None, 1)                      | 6       |
| activation_10 (Activation) | (None, 1)                      | 0       |

**4.3. Evaluasi Model CNN**

Penelitian ini memperoleh tingkat akurasi sistem sebesar 95% dari sekitar 20 epoch. Hasil model ini juga menunjukkan sekitar 145 data yang masih salah prediksi. Akurasi untuk evaluasi model CNN digunakan dalam penelitian ini. Saat melakukan *training* model [20], penelitian menggunakan *K-fold cross validation* untuk memastikan model klasifikasi yang digunakan dapat berjalan secara netral dan tidak menimbulkan bias. Penelitian ini menggunakan fungsi *K-fold* pada *library* scikit-learn (Python) yang dieksekusi menggunakan Tensorflow Keras [21]. Tabel 2 menunjukkan evaluasi CNN.

**Tabel 2. Evaluasi CNN**

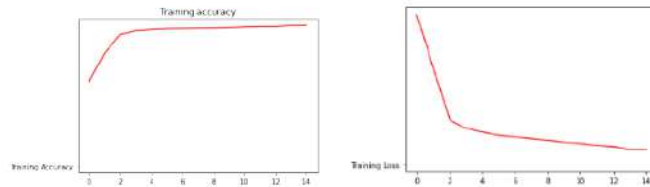
| precision    | recall | f1-score | support |      |
|--------------|--------|----------|---------|------|
| 0            | 0.93   | 0.97     | 0.95    | 1407 |
| 1            | 0.97   | 0.92     | 0.94    | 1349 |
| accuracy     |        |          | 0.95    | 2756 |
| macro avg    | 0.95   | 0.95     | 0.95    | 2756 |
| Weighted avg | 0.95   | 0.95     | 0.95    | 2756 |



### Gambar 8. Hasil Prediksi CNN

Penelitian ini menggunakan beberapa kali eksperimen, khususnya melakukan kombinasi *test* pada *Epoch Batch size* dengan beberapa *setting* parameter dan iterasi yang beragam agar hasil *training* dan *testing* diharapkan dapat menjadi baik. Hasil *testing* dapat dilihat pada Gambar 8.

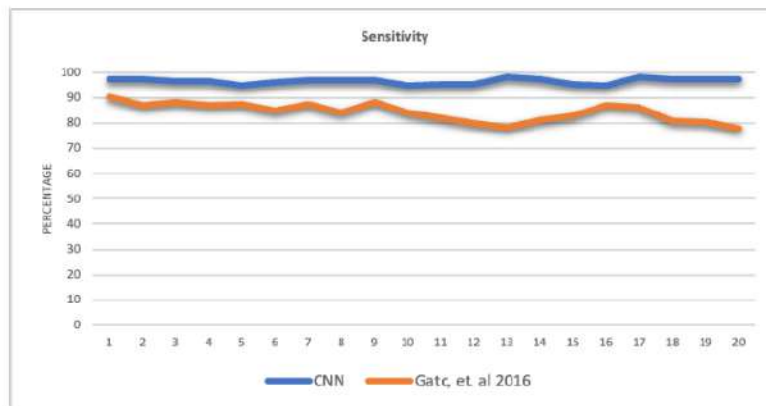
Selain itu juga digunakan evaluasi pada *training loss* dan *training accuracy* untuk model yang dibangun, dimana *loss* merupakan suatu ukuran dari sebuah *error* yang dibuat oleh network [19] [20], dan tujuannya adalah mengetahui tren pada kegiatan *training* untuk meminimalisirnya. Standar fungsi *loss* untuk mengkategorisasikan masalah tersebut adalah *cross-entropy loss*. Hasil dari evaluasi ini dapat dilihat pada Gambar 8.



Gambar 9. Hasil Eksperimen Akurasi dan Loss

Dari grafik pada Gambar 9, dapat diketahui bahwa hasil dari *Training Loss* dari data yang digunakan menunjukkan tren menurun dari setiap iterasi *epoch* diatas 20. Sementara, *training loss* menunjukkan nilai hampir mencapai 0. Sebaliknya, akurasi menunjukkan tren yang semakin baik tiap iterasinya dan mencapai angka tertinggi yaitu 97% *training* akurasi, dan ini merupakan hasil terbaik sejauh ini dibandingkan penelitian sebelumnya yaitu hanya diangka 90% [9].

Pengujian *Sensitivity* dan *PPV* dilakukan pada model yang dibangun dengan menggunakan CNN seperti penelitian sebelumnya [9] dengan menggunakan 20 data yang berbeda secara random. Penelitian ini tidak menggunakan perhitungan akurasi karena beberapa data yang diambil sebagai pembanding belum memiliki label khususnya untuk *True Negatif* namun data positif sudah diberikan label (lihat Gambar 10).



Gambar 10. Perbandingan CNN dengan Penelitian Gatc, et.al 2016

## 5. Kesimpulan dan Saran

Salah satu fase yang terpenting dalam identifikasi parasit plasmodium pada *slide* darah yaitu fase segmentasi. Fase ini digunakan untuk memisahkan sel darah dan *background*.

Penelitian ini juga menggunakan segmentasi untuk memisahkan informasi sel darah merah dan sel darah putih. Selain itu pembentukan model merupakan salah satu langkah yang penting yang harus dilakukan agar hasil deteksi dapat mendapatkan hasil yang baik. Berdasarkan penelitian sebelumnya, sudah dilakukan penyempurnaan algoritma *double thresholding* agar segmentasi dapat memberikan hasil yang baik.

Penelitian ini menggunakan data dari sumber terbuka Kaggle untuk membentuk model deteksi (*data training*) dan menggunakan data riil sebagai *data testing* dari *capture* menggunakan mikroskop dengan perangkat *Digital Camera Eyepiece* beresolusi 2 MP. Hasil pengujian memperlihatkan hasil akurasi yang baik yaitu model dari Algoritma CNN ini memberikan hasil akurasi yaitu 97,96% dan *loss* 0,06 dengan rata-rata waktu komputasi sekitar 121 detik/*epoch*. Setelah dilihat lebih jauh, penggunaan data riil sepertinya masih menjadi tantangan untuk penelitian selanjutnya dimana hasil *capture* data yang diperoleh dari data lapangan masih cenderung memberikan hasil yang kurang baik.

Isu terkait peningkatan akurasi yang baik, kemudahan deteksi, dan kemudahan penggunaan oleh tenaga kesehatan di berbagai tempat tentunya merupakan tujuan utama dari penelitian ini. Saat ini teknik mikroskopis untuk deteksi malaria tetap menjadi salah satu teknik deteksi yang memiliki tingkat kredibilitas yang baik, tentu kedepannya dengan perkembangan perangkat keras seperti mikroskop digital terbaru yang dapat digunakan pada *smartphone* dengan pembesaran dari 60x - 100x bisa menjadi penelitian lanjutan untuk pengembangan sistem deteksi dengan menggunakan *smartphone*. Tentu ini juga bisa menjadi lompatan besar dalam proses pemeriksaan malaria yang lebih mudah dan praktis. Selain itu, teknologi komputasi dan penyimpanan berbasis *cloud* dapat menjadi saran penelitian selanjutnya dimana sistem dapat dieksekusi dengan *cloud computing* serta hasil *training* dan deteksi parasit dapat tersentralisasi dan dapat menjadi *bank data* yang dapat berguna bagi banyak pihak.

## 6. Ucapan Terima Kasih

Terima kasih kami ucapkan kepada Bapak Vendyan Sirius G. atas supervisinya selama proses pengerjaan penelitian ini.

## Referensi

- [1] World Health Organization. Malaria Rapid Diagnostic Test Performance: Executive Summary. WHO. 2020.
- [2] World Health Organization, Basic Malaria Microscope. Part 1. Learner's Guide, 2nd ed., Geneva: WHO. 2010. ISBN: 9241547820. Pp. 7-12.
- [3] World Health Organization. Diagnosis of malaria. Guidelines for the treatment of malaria. Second edition. World Health Organization 2012. ISBN 978 92 4 154792 5. pp. 10-12.
- [4] Massamba, J.E., Djontu, J.C., Vouvougui, C.J. et al. Plasmodium falciparum multiplicity of infection and pregnancy outcomes in Congolese women from southern Brazzaville, Republic of Congo. *Malar J* 21, 114 (2022). <https://doi.org/10.1186/s12936-022-04105-w>
- [5] DIMI, B., ADAM, A., ALIM, A. Prevalensi Malaria Berdasarkan Karakteristik Sosio Demografi. *Jurnal Ilmiah Kesehatan*. Vol. 19 No. 1. 2020. p-ISSN: 1412-2804. DOI: 10.33221/jikes.v19i01.399.
- [6] Centers for Disease Control and Prevention. Malaria. CDC. Web. 10 May 2022. <<http://www.cdc.gov/dpdx/malaria/index.html>>
- [7] LIU, D., JIANYU. Otsu Method and K-means. International Conference on Hybrid Intelligent Systems. Department of Computer Science Beijing Jiaotong University Beijing. IEEE. China. 2009. ISBN 978-0-7695-3745-0. pp. 344-349.
- [8] GATC, J., et al. Plasmodium Parasite Detection on Red Blood Cell Image for the Diagnosis of Malaria Using Double Thresholding. ICACSI: IEEE. 2013. pp. 236-240
- [9] GATC, J., MASPIYANTI, F. Red blood cell and white blood cell classification using double thresholding and BLOB analysis. IEEE 4th International Conference on

- Information and Communication Technology (ICoICT). 2016. ISBN: 978-1-4673-9879-4. pp. 1-5.
- [10] MASPIYANTY, F., GATC, J. Plasmodium Falciparum Stages Classification on Red Blood Cell Image using Region Property. IEEE 1st International Conference on Information Technology, Information Systems and Electrical Engineering. 2016. ISBN 978-1-5090-1567-2. pp. 105-109.
- [11] ANGGRAINI, D., et.al. Automatic Status Identification of Microscopic Images Obtained from Malaria Thin Blood Smears. Proc. of 3rd International Conference on Electrical Engineering and Informatics (ICEEI 2011), CDROM A3-2. Institut Teknologi Bandung. Bandung, Indonesia. 2011. pp. 347-352.
- [12] JAEGER, S., Malaria Dataset: NIAID. National Library of Medicine. Lister Hill National Center for Biomedical Communication. 2019. <https://doi.org/10.7717/peerj.6977>.
- [13] BORAY, F., et. al. Parasite detection and identification for automated thin blood film malaria diagnosis. Computer Vision and Image Understanding. Elsevier. 2011. pp. 89-95.
- [14] ROMANI, L., PANE, S., SEVERINI, C. et al. Challenging diagnosis of congenital malaria in non-endemic areas. Malar J 17, 470 (2018). <https://doi.org/10.1186/s12936-018-2614-9>.
- [15] RAHNAMATAN, Z., MOHAMAND, S., Tissue Segmentation in Medical Images Based on Image Processing Chain Optimization. International Workshop on Real Time Measurement, Instrumentation & Control (RTMIC), IEEE NPSS. Toronto. 2010. pp. 697-703.
- [16] Mehedi Masud, Hesham Alhomyani, Sultan S. Alshamrani, Omar Cheikhrouhou, Saleh Ibrahim, Ghulam Muhammad, M. Shamim Hossain, Mohammad Shorfuzzaman, "Leveraging Deep Learning Techniques for Malaria Parasite Detection Using Mobile Application", Wireless Communications and Mobile Computing, vol. 2020, Article ID 8895429, 15 pages, 2020. <https://doi.org/10.1155/2020/8895429>
- [17] ZHANG, Y., LOBO-MUELLER, E.M., KARANICOLAS, P. et al. CNN-based survival model for pancreatic ductal adenocarcinoma in medical imaging. BMC Med Imaging 20, 11 (2020). <https://doi.org/10.1186/s12880-020-0418-1>
- [18] REUTERSWÄRD, P., BERGSTRÖM, S., ORIKHIRIZA, J. et al. Levels of human proteins in plasma associated with acute paediatric malaria. Malar J 17, 426 (2018). <https://doi.org/10.1186/s12936-018-2576-y>
- [19] QUAKYI, I.A., ADJEI, G.O., SULLIVAN, D.J. et al. Diagnostic capacity, and predictive values of rapid diagnostic tests for accurate diagnosis of Plasmodium falciparum in febrile children in Asante-Akim, Ghana. Malar J 17, 468 (2018). <https://doi.org/10.1186/s12936-018-2613-x>
- [20] TANNER, M., et.al. Malaria eradication: benefits, future scenarios and feasibility: Executive summary. World Health Organization. 2019.
- [21] Fuhad KMF, Tuba JF, Sarker MRA, Momen S, Mohammed N, Rahman T. Deep Learning Based Automatic Malaria Parasite Detection from Blood Smear and its Smartphone Based Application. Diagnostics (Basel). 2020;10(5):329. Published 2020 May 20. doi:10.3390/diagnostics10050329

# Prediksi Parasit Plasmodium pada Citra Mikroskopis Sel Darah Merah dengan Convolutional Neural Networks

---

## ORIGINALITY REPORT

---

12%

SIMILARITY INDEX

9%

INTERNET SOURCES

2%

PUBLICATIONS

4%

STUDENT PAPERS

---

## MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

---

2%

★ [media.neliti.com](http://media.neliti.com)

Internet Source

---

Exclude quotes  On

Exclude matches  < 10 words

Exclude bibliography  On