

Klasifikasi Fase Penyakit Malaria *Falciparum* Berdasarkan Citra Sel Darah Merah Menggunakan *Extraction Texture* dan *Morphological* Menggunakan 1D-Convolutional Neural Network

1st Rheanov Bija Adzlan
Fakultas Informatika
Universitas Telkom
Bandung

rheanov@students.telkomuniversity.ac.id

2nd Febryanti Sthevanie
Fakultas Informatika
Universitas Telkom
Bandung

sthevanie@telkomuniversity.ac.id

3rd Kurniawan Nur Ramadhani
Fakultas Informatika
Universitas Telkom
Bandung

kurniawanr@telkomuniversity.ac.id

Abstrak — Malaria merupakan salah satu penyakit menular yang menginfeksi sel darah merah. Pada pasien terduga terjangkit malaria, pemeriksaan mikroskopik hapusan darah masih menjadi standar yang dilakukan untuk mendiagnosis malaria. Hal ini menimbulkan permasalahan dalam diagnosis malaria, karena dalam melakukan diagnosis tersebut membutuhkan pelatihan khusus dan keahlian yang cukup pada sumber daya yang ada. Sensitivitas sebesar 57% dan masih terbilang rendah menjadi kendala pemeriksaan mikroskopik, sehingga berdasarkan hal tersebut diperlukan cara identifikasi/klasifikasi secara tepat dan akurat terhadap penyakit malaria dengan memanfaatkan perkembangan teknologi yang telah maju salah satunya dengan menggunakan *Machine Learning*. Pada penelitian ini menggunakan 1D-Convolutional Neural Network, dan menggunakan *feature extraction texture* dan *morphological*. Dataset yang digunakan dari website Kaggle yaitu *Malaria Parasite Image (Different Malaria Species)* dengan *species* yang digunakan *Falciparum* dan label terdiri dari fase *Gametocytes (G)*, *Rings (R)*, *Trophozoite (T)*, *Schizont (S)*. Hasil dari penelitian ini mendapatkan nilai *confusion matrix* tertinggi pada proporsi 80:20 dengan nilai *accuracy* sebesar 82%.

Kata kunci— Citra, *Extraction Texture*, *Falciparum*, *Morphological*, 1D-CNN

I. PENDAHULUAN

Malaria merupakan salah satu penyakit menular yang menginfeksi sel darah merah. Penyebab penyakit malaria disebabkan oleh parasit plasmodium yang ditularkan melalui gigitan nyamuk *Anopheles betina* yang terinfeksi dan menginfeksi sel darah merah (RBC) manusia [1]. Pada pasien terduga terjangkit malaria, pemeriksaan mikroskopik hapusan darah masih menjadi standar yang dilakukan untuk mendiagnosis malaria. Hal ini menimbulkan permasalahan dalam diagnosis malaria, karena dalam melakukan diagnosis tersebut membutuhkan pelatihan khusus dan keahlian yang cukup pada sumber daya yang ada. Sensitivitas sebesar 57% dan masih terbilang rendah menjadi kendala pemeriksaan mikroskopik [2], sehingga berdasarkan hal tersebut diperlukan cara identifikasi/klasifikasi secara tepat dan akurat terhadap penyakit malaria dengan memanfaatkan

perkembangan teknologi yang telah maju salah satunya dengan menggunakan *Machine Learning* [3].

Terdapat beberapa penelitian sebelumnya yang menerapkan teknik *machine learning* dan membahas mengenai identifikasi dan klasifikasi citra terhadap penyakit malaria, salah satu nya penelitian yang dilakukan oleh Setyawan, Wardoyo, Wibowo dan Murhandarwati (2022) [4]. Penelitian tersebut membahas mengenai penerapan *Neural Network* pada klasifikasi *plasmodium falciparum* dengan melakukan beberapa tahapan seperti tahap *preprocessing*, segmentasi, *feature extraction* (kombinasi *texture* dan *Morphological*) dengan menghasilkan akurasi 82.67% yang dimana akurasi ini lebih tinggi jika hanya mengandalkan *texture* yaitu sebesar 72% atau hanya mengandalkan *Morphological* sebesar 76%. Penelitian lainnya yang dilakukan Nurriski dan Alamsyah (2023) [5] membahas mengenai deteksi aritmia melalui sinyal EKG. Permasalahan yang melatarbelakangi penelitian tersebut adalah untuk melakukan analisa sinyal EKG yang memerlukan keahlian dan tidak jarang kurang praktis, karena setiap detak jantung perlu dilakukan analisis mendetail, karena setiap detak jantung harus dianalisis, sehingga penelitian tersebut melakukan deteksi aritmia melalui sinyal EKG menggunakan Conv1D dan menghasilkan tingkat akurasi 98.88%.

Berdasarkan hal tersebut, penelitian ini akan menggunakan sebuah algoritma 1D-CNN. Pemilihan 1D-CNN didasari dari penelitian yang dilakukan oleh Nurriski dan Alamsyah (2023) [5] yang menghasilkan tingkat akurasi 98.88% pada deteksi aritmia melalui sinyal EKG. Dalam melakukan klasifikasi penyakit malaria penelitian ini menggunakan *feature extraction texture* dan *morphological*. Penggunaan *Feature Extraction* pada penelitian ini dikarenakan pada penelitian Setyawan, Wardoyo, Wibowo dan Murhandarwati, 2022 [4] memberikan kenaikan akurasi yang cukup signifikan sekitar 10%, sehingga dengan dilakukan penelitian ini diharapkan dapat membantu mengidentifikasi citra malaria dan mengetahui kinerja 1D-CNN pada klasifikasi citra malaria dengan menggunakan *feature extraction*. Penggunaan 1D CNN dalam penelitian ini meskipun data input adalah citra 2D dapat dibenarkan karena

fokus utama bukan pada seluruh informasi spasial citra, melainkan pada ekstraksi fitur spektral yang relevan. Menurut Florensia, Samsuryadi dan Saparudin [6], dengan menggunakan 1D, CNN dapat menangkap fitur spektral intrinsik sepanjang satu dimensi dari citra, yang diubah atau diproyeksikan menjadi bentuk 1D untuk tujuan spektral. Ini memungkinkan identifikasi pola spektral yang lebih efisien tanpa perlu memproses seluruh ruang dua dimensi citra secara langsung. Pendekatan ini memaksimalkan efektivitas dalam menangkap konten spektral yang penting dalam domain masalah yang dihadapi

Topik yang akan dibahas pada penelitian ini adalah klasifikasi citra sel darah merah untuk fase penyakit Malaria menggunakan 1D-CNN dengan menerapkan kombinasi *feature extraction* yaitu *texture extraction* dan *morphological extraction*. Batasan pada penelitian ini adalah *dataset* yang digunakan bersumber dari *Kaggle Malaria Parasite Image (Different Malaria Species)*. *Species* yang digunakan dalam penelitian ini yaitu *falciparum* yang terdiri dari 4 label yaitu *Gametocytes (G)*, *Rings (R)*, *Trophozoites (T)*, dan *Schizonts (S)* dengan jumlah *dataset* keseluruhan sejumlah 194 data citra.

Tujuan dari penelitian ini adalah melakukan klasifikasi citra sel darah merah untuk fase penyakit Malaria pada *species falciparum* menggunakan 1D-CNN dengan menerapkan kombinasi *feature extraction* yaitu *texture extraction* dan *morphological extraction* serta untuk mengetahui hasil kinerja dari 1D-CNN pada klasifikasi citra sel darah merah pada *species falciparum* dengan menggunakan *feature extraction*.

Setelah dilakukan penjabaran mengenai bab pendahuluan, selanjutnya akan dijabarkan mengenai bab selanjutnya yaitu pada bab 2 merupakan studi literatur yang akan membahas mengenai penelitian terdahulu, bab 3 merupakan pembahasan untuk sistem yang dibangun, pada bab 4 merupakan penjabaran mengenai evaluasi dari model yang digunakan, dan bab 5 merupakan kesimpulan

II. STUDI TERKAIT

Terdapat beberapa penelitian terkait yang telah dilakukan, seperti penelitian yang dilakukan oleh Dewi dan Ismawan (2021) [7] yang membahas mengenai pengenalan wajah. Permasalahan yang melatarbelakangi penelitian ini adalah untuk pengenalan wajah cepat dilakukan oleh manusia, namun butuh waktu yang lama bagi komputer. Sehingga dengan dilakukannya penelitian ini maka dilakukan deteksi pengenalan wajah menggunakan Conv1. Hasil dari penelitian ini adalah mendapatkan nilai *precision* 98.4%, *recall* 98%, akurasi 99.84%.

Penelitian kedua yang dilakukan oleh Devassy, dan George (2020) [8] yang membahas mengenai klasifikasi *strawberry* menggunakan *hyperspectral imaging* tanpa kontak. Hal yang melatarbelakangi penelitian ini adalah estimasi kualitas buah tanpa kontak dengan cepat merupakan faktor penting untuk pemrosesan makanan agar lebih efisien.

Sehingga dalam penelitian tersebut memanfaatkan *hyperspectral imaging* dan 1D-CNN. Hasil dari penelitian ini adalah 1D dapat mencapai akurasi sebesar 96.6%.

Penelitian ketiga yang dilakukan oleh Fuadah, Saidah, Sy, Magdalena, dan Ubaidullah (2022) [9] yang membahas mengenai klasifikasi *glaucoma* berdasarkan pemrosesan citra fundus. Permasalahan yang melatarbelakangi penelitian ini adalah kelambatan diagnosa dan penanganan masalah kerusakan saraf optik akibat glaukoma dapat mengakibatkan kebutaan secara permanen. Sehingga dari permasalahan tersebut diperlukan sistem deteksi dini glaukoma berbasis pengolahan citra dengan memanfaatkan *machine learning*. Pada penelitian ini menggunakan 1D-CNN, dan menghasilkan penelitian dengan akurasi terbaik sebesar 99%.

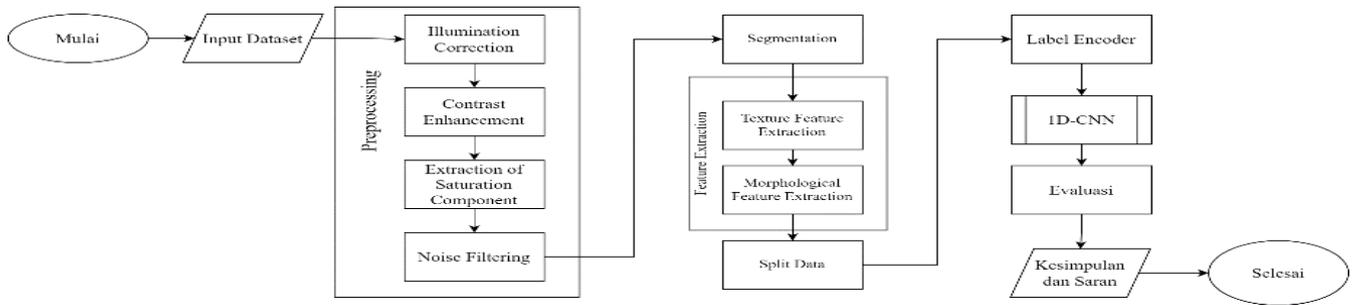
Penelitian selanjutnya dilakukan oleh Magdalena, et al (2022) [10] yang membahas mengenai deteksi anemia berdasarkan citra konjungtiva palpebra. Hal yang melatarbelakangi penelitian ini adalah pengecekan berkala pada kadar hemoglobin ibu hamil sangat penting untuk diterapkan. Sehingga diperlukan sistem deteksi anemia menggunakan CNN. Pada penelitian tersebut model 1D-CNN yang diajukan dilatih dengan menggunakan beragam metode optimasi dan *learning rate* yang bervariasi. Hasil dari penelitian ini adalah dengan menghasilkan akurasi terbaik sebesar 94%.

Penelitian lain yang dilakukan oleh Pramudito, Fu'adah, Magdalena, Rizal, dan Taliningsih (2022) [11] yang membahas mengenai identifikasi sinyal *congestive heart failure*. Permasalahan yang melatarbelakangi penelitian ini adalah dalam membaca sinyal EKG terdapat hal yang perlu diperhatikan, jika hasil yang didapatkan dalam kondisi abnormal tidak selalu menunjukkan bahwa terdapat kelainan pada jantung, atau sebaliknya. Sehingga diperlukan identifikasi sinyal *congestive heart failure* menggunakan 1D-CNN. Hasil dari penelitian ini adalah dapat memberikan akurasi 100%.

Penelitian selanjutnya yang dilakukan oleh Hsieh dan Kiang (2020) [12] yang membahas mengenai klasifikasi citra hiperspektral di lahan pertanian. Hal yang melatarbelakangi penelitian ini yaitu untuk mengetahui kinerja CNN dalam mengeksploitasi fitur yang telah ada untuk klasifikasi citra hiperspektral. Dalam penelitian ini menggunakan 1D-CNN untuk klasifikasi dan penelitian tersebut memperoleh akurasi tertinggi sebesar 99,8% dan 98,1%, masing-masing, dicapai dengan menerapkan 1D-CNN dengan vektor input yang ditambah, yang berisi fitur spektral dan spasial yang tertanam dalam data HSI.

III. SISTEM YANG DIBANGUN

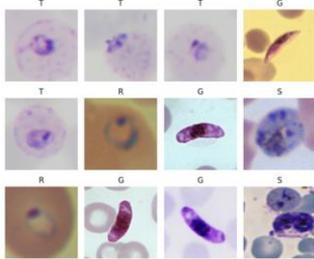
Berikut pada Gambar 1 merupakan alur pemodelan yang digunakan dalam penelitian ini dan menggambarkan sistem yang dibangun.



Gambar 1. Alur Pemodelan

3.1. Dataset

Dataset yang digunakan bersumber dari *Kaggle* yaitu *Malaria Parasite Image (Different Malaria Species)*, *species* yang digunakan adalah *falciparum* yang terdiri dari 4 label yaitu *Gametocytes (G)* sebanyak 33 *images*, *Rings (R)* sebanyak 67 *images*, *Trophozoites (T)* sebanyak 61 *images*, dan *Schizonts (S)* sebanyak 33 *images* dengan jumlah dataset keseluruhan sejumlah 194 data citra. Gambar 1 merupakan dari dataset yang digunakan.



Gambar 2. Dataset

3.2. Preprocessing

Pada *preprocessing* ini dilakukan proses *illumination correction*, *contrast enhancement*, *extraction of saturation component*, *noise filtering*.

3.2.1. Illumination Correction

Illumination correction bertujuan untuk mengatasi ketidakseragaman iluminasi pada gambar. Pada fungsi ini menerapkan *illumination correction* dengan prinsip *gray world*. Proses dalam *illumination correction* yaitu menghitung rata-rata warna dari seluruh piksel dalam gambar pada sumbu RGB dan nantinya rata-rata tersebut dibagi oleh array ['128','128','128'] mewakili nilai netral abu-abu. Terakhir, hasil dikonversi menjadi data *unsigned integer* 8-bit karena nilai piksel gambar umumnya berada dalam rentang 0-255. Persamaan rumus (1) merupakan rumus yang digunakan untuk menghitung *illumination correction* pada bagian (r).

$$F_{ravg} = \frac{1}{MN} \sum_{X=1}^M \sum_{y=1}^N f_r(x, y) \quad (1)$$

Dalam teknik ini, nilai rata-rata saluran hijau digunakan untuk menghitung penguatan bagian saluran merah menggunakan persamaan rumus (2).

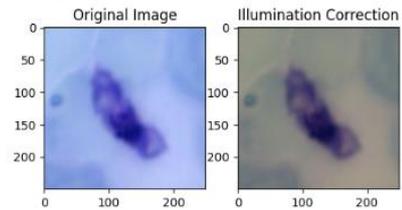
$$g_r = \frac{F_{gavg}}{F_{ravg}} \quad (2)$$

Selanjutnya, g_r digunakan untuk menghitung saluran merah yang telah disesuaikan dengan menggunakan persamaan rumus (3).

$$f_{radj}(x, y) = g_r x f_r(x, y) \quad (3)$$

Gambar yang dihasilkan dari asumsi dunia abu-abu diperoleh dengan menggabungkan saluran $f_{radj}(x, y)$, $f_g(x, y)$, dan $f_{badj}(x, y)$.

Gambar 3 merupakan hasil perbandingan dari *original image* dan *illumination correction*.



Gambar 3. Hasil *Illumination Correction*

3.2.2. Contrast Enhancement

Pada tahap ini dilakukan *contrast enhancement* untuk meningkatkan kontras gambar sehingga objek parasit plasmodium terlihat lebih jelas. Fungsi *contrast enhancement* dengan proses *amplification factor map*, *shadow compensation* dan *recursive filtering* agar kontras yang ditampilkan dapat menampilkan hasil yang cukup bagus. *Amplification factor* diterapkan sesuai dengan *journal* dimana p dan α yang bernilai 0.125. Sedangkan pada *shadow compensation* melibatkan konversi gambar ke ruang warna LAB dengan menggunakan fungsi 'calculate_amplification_factor' dan selanjutnya gambar dikembalikan ke ruang warna BGR. Sementara *recursive filtering* terdapat proses *loop* yang akan diulang dua kali untuk meningkatkan efek peningkatan kontras. Persamaan rumus (4) merupakan rumus yang digunakan untuk proses *contrast enhancement* pada bagian (r).

$$\underline{R}(i, j) = \beta \underline{R}(i, j - 1) + (1 - \beta) I(i, j, 1) \quad (4)$$

Kemudian, nilai koreksi untuk setiap piksel pada komponen warna merah dihitung dengan menggunakan persamaan rumus (5).

$$YR(i, j) = \frac{\underline{g}(i, j)}{\underline{R}(i, j)} \left(\frac{(1-\alpha)\underline{R}(i, j)+255\alpha}{(1-\alpha)\underline{g}(i, j)+255\alpha} \right) \quad (5)$$

Hasil akhir nilai piksel penyesuaian $F(i, j)$ pada bagian (r) dengan menggunakan kompensasi white balance dikomputasikan dengan menggunakan persamaan rumus (6).

$$F(i, j, 1) = YR^{l(i,j,1)} \quad (6)$$

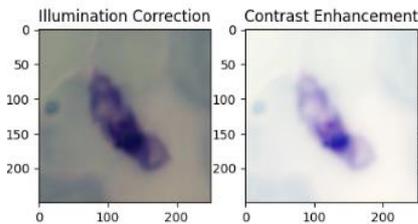
Selanjutnya menghitung faktor amplifikasi pada tiap piksel adaptif terhadap intensitas rata-rata piksel di sekitarnya menggunakan persamaan rumus (7).

$$H(i, j) = P.H(i, j - 1) + (1 - p) \left((1 - \alpha) + \frac{255\alpha}{Y(i, j)} \right) \quad (7)$$

Langkah terakhir adalah penguatan piksel menggunakan pemrosesan gambar logaritmik (LIP), seperti yang disajikan dalam persamaan (8).

$$F_{out}(i, j, k) = M - M \left(1 - \frac{F(i, j, k)}{M} \right)^{H(i, j)} \quad (8)$$

M adalah nilai maksimum setiap komponen dalam ruang warna RGB atau YCbCr. Dalam kasus gambar RGB, nilai M adalah 255. k menunjukkan indeks komponen warna, 1 untuk merah, 2 untuk hijau, dan 3 untuk biru. Berikut pada Gambar 4 merupakan hasil perbandingan dari *illumination correction* dan *contrast enhancement*.



Gambar 4. Hasil Contrast Enhancement

3.2.3. Extraction of Saturation Component

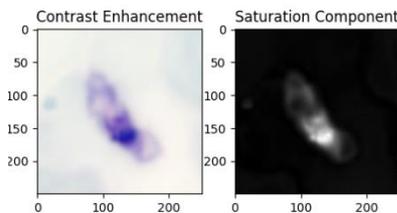
Pada proses *extraction of saturation component* hanya mengambil *index* [:, :, 1] dari *HSV image* dimana (0 = Hue, 1= saturation, 2= value). Berikut merupakan persamaan untuk proses *extraction of saturation component*.

$$H = \tan \left(\frac{3(G - B)}{(R - G) + (R - B)} \right) \quad (9)$$

$$S = 1 - \frac{\min(R, G, B)}{V} \quad (10)$$

$$V = \frac{R + G + B}{3} \quad (11)$$

Berikut pada Gambar 5 merupakan hasil perbandingan dari *contrast enhancement* dan *extraction of saturation component*.

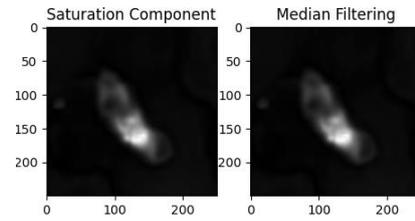


Gambar 5. Hasil Saturation Component

3.2.4. Noise Filtering

Noise pada gambar dapat membuat kualitas gambar menjadi rendah, yang juga akan berdampak pada akurasi segmentasi bentuk dalam sel darah merah. Teknik *filtering*

dapat digunakan untuk menghilangkan noise pada citra. Pada penelitian ini menggunakan *filter_size* yang diatur ke 3 dan *filter_size* juga digunakan untuk menggantikan nilai setiap piksel dengan nilai median. Berikut pada Gambar 6 merupakan hasil perbandingan dari *extraction of saturation component* dan *noise filtering/median filtering*.



Gambar 6. Hasil Media Filtering

3.3. Segmentation

Segmentasi digunakan untuk memisahkan citra ke dalam beberapa daerah berdasarkan kemiripan atribut yang dimilikinya. Pada proses ini menggunakan *otsu thresholding* dan *morphological operation*. *Thresholding* mengubah citra skala keabuan menjadi citra biner. Sementara *morphological operation* digunakan untuk menyederhanakan struktur citra.

3.4. Feature Extraction

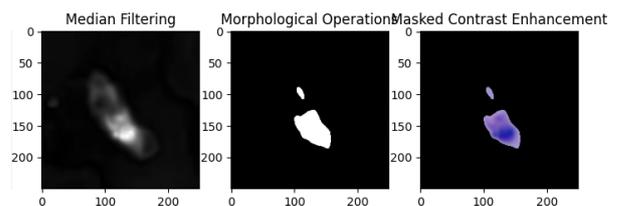
Pada proses *feature extraction*, dalam penelitian ini dilakukan *texture feature extraction* dan *morphological feature extraction*. Berikut merupakan penjabaran terkait kedua proses dari *feature extraction*.

3.4.1. Texture Feature Extraction

Pada penelitian ini untuk melakukan proses *texture feature extraction* yaitu dengan menghitung nilai *mean* untuk menghitung nilai rata-rata dari gambar *grayscale*, *standard deviation* untuk menghitung standar deviasi dari gambar *grayscale*, *skewness* menghitung sejauh mana distribusi piksel miring ke kiri atau kanan, menghitung *energy*, *entropy*, dan *smoothness*. Untuk menghitung *entropy* diperlukan untuk mengimport 'shannon_entropy'.

3.4.2. Morphological Feature Extraction

Morphological dapat disebut sebagai wujud dan struktur sebuah objek atau dalam penjelasan lainnya sebagai susunan dan hubungan antar komponen pada sebuah objek [13]. Pada proses ini dilakukan inialisasi untuk menghitung *area ratio* dan *eccentricity*. *Area ratio* dapat dihitung menggunakan *countour area* dan dibagi dengan total area. Sedangkan *eccentricity* untuk mengukur seberapa jauh suatu bentuk dari bentuk bulat sempurna. Terdapat nilai 1 dan 0, untuk nilai 1 menunjukkan bentuk yang lebih ellipse, sedangkan 0 menunjukkan bentuk yang lebih mendekati lingkaran. Berikut pada Gambar 7 merupakan hasil *feature extraction*.



Gambar 7. Hasil Feature Extraction

3.5. Split Data

Setelah itu, dilakukan *split data*. Proses ini digunakan untuk membagi *data train* dan *data test*. Pada penelitian ini *split data* menggunakan proporsi 80%:20%, 60%:40%, 90:10%, dan 70%:30%.

3.6. Label Encoder

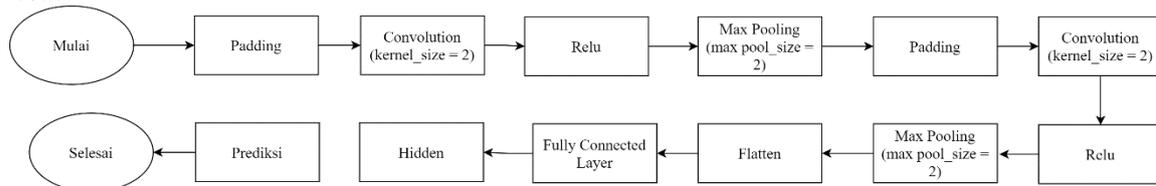
Pada tahap ini dilakukan *label encoder* untuk mengubah data kategorikal nomor berurutan sebagai inputan label pada *machine learning* [14]. Proses ini digunakan untuk mengonversi label pada *class* menjadi nilai numerik yang berasal dari *data train* dan *data test*. Tabel 1 merupakan hasil dari *label encoder*.

Tabel 1. Label Encoder

| <i>Gemotocytes</i> (G) | <i>Rings</i> (R) | <i>Schizonts</i> (S) | <i>Trophozoites</i> (T) |
|---------------------------|---------------------|-------------------------|----------------------------|
| 0 | 1 | 2 | 3 |

3.7. 1D-CNN

Selanjutnya akan dilakukan proses klasifikasi citra malaria menggunakan 1D-CNN dengan proses seperti pada Gambar 8.



Gambar 8. Flowchart 1D-CNN

Proses pertama yaitu memilih fitur dari dataset yang digunakan dalam model. Fitur tersebut seperti *mean*, *std*, *skewness*, *energy*, *entropy*, *smoothness*, *area ratio*, dan *eccentricity* yang berasal dari *data train* dan *data test*. Selanjutnya akan ditampilkan dalam bentuk (*shape*) dan nilai dari fitur pertama pada *data pengujian*. Selanjutnya dilakukan *reshape* fitur-fitur dari *array* dua dimensi menjadi *array* tiga dimensi yang biasanya terjadi pada 1D-CNN. Setelah itu, dilakukan representasi kategorikal dari data sebagai vektor biner.

Setelah mengimpor *library* dari TensorFlow, pembuatan model *neural network* dilakukan menggunakan objek *Sequential*. Model ini terdiri dari beberapa lapisan (*layers*) yang membentuk *Convolutional Neural Network (CNN) 1D*. Lapisan pertama adalah *Conv1D* dengan 32 *filter*, ukuran *kernel (kernel_size)* sebesar 2, fungsi aktivasi '*relu*', dan *input shape* yang sesuai dengan dimensi data *input*. Setelah lapisan konvolusi, *BatchNormalization* digunakan untuk menormalkan *output*, menstabilkan distribusi aktivasi dalam jaringan, dan membantu mempercepat proses pelatihan. Kemudian, *MaxPooling1D* dengan *pool_size=2* digunakan untuk mengurangi dimensi data secara spasial, tanpa menambah parameter baru. Selanjutnya, terdapat lapisan *Dropout* dengan rasio 0.25 yang berfungsi untuk mencegah *overfitting* dengan mengabaikan sejumlah neuron secara acak selama pelatihan; lapisan ini tidak menambah

parameter. Data kemudian diubah menjadi bentuk satu dimensi dengan menggunakan lapisan *Flatten*, yang juga tidak menambah parameter. Lapisan berikutnya adalah *Dense* dengan 64 neuron, menggunakan fungsi aktivasi '*relu*' dan regularisasi L2 sebesar 0.65, diikuti oleh *BatchNormalization* untuk normalisasi lebih lanjut. Lapisan *Dropout* lain ditambahkan dengan rasio 0.5 untuk mencegah *overfitting*, tanpa menambah parameter. Model ini kemudian dilanjutkan dengan lapisan *Dense* kedua yang memiliki 32 neuron, juga menggunakan fungsi aktivasi '*relu*' dan regularisasi L2 sebesar 0.65. Lapisan akhir adalah *Dense* dengan 4 neuron yang menggunakan fungsi aktivasi '*softmax*', yang digunakan untuk klasifikasi multikelas.

Selanjutnya, dilakukan proses konfigurasi dan pelatihan model dengan mendefinisikan fungsi *loss weighted_categorical_crossentropy*, yang dirancang untuk menangani masalah klasifikasi multikelas dengan memberikan bobot khusus berdasarkan *class weights*. Setelah menghitung *class weights*, model dikompilasi menggunakan *optimizer Adam* dengan *learning rate* sebesar 0.001, fungsi *loss*, dan metrik akurasi. Untuk memonitor proses pelatihan, beberapa fungsi *callbacks* digunakan, seperti *ModelCheckpoint* untuk menyimpan model terbaik, *ReduceLROnPlateau* untuk menurunkan laju pembelajaran secara dinamis, dan *EarlyStopping* untuk menghentikan pelatihan lebih awal saat tidak ada peningkatan pada validasi

```
47. Combination: ('Std', 'Energy', 'Entropy', 'Smoothness', 'Eccentricity')
/media/mz/BA96E8A396E860FD/Project/jupyter/env/lib/python3.10/site-packages/k
super().__init__(activity_regularizer=activity_regularizer, **kwargs)
Model: "sequential_46"
```

| Layer (type) | Output Shape | Param # |
|---|---------------|---------|
| conv1d_46 (Conv1D) | (None, 4, 32) | 96 |
| batch_normalization_92 (BatchNormalization) | (None, 4, 32) | 128 |
| max_pooling1d_46 (MaxPooling1D) | (None, 2, 32) | 0 |
| dropout_92 (Dropout) | (None, 2, 32) | 0 |
| flatten_46 (Flatten) | (None, 64) | 0 |
| dense_138 (Dense) | (None, 64) | 4,160 |
| batch_normalization_93 (BatchNormalization) | (None, 64) | 256 |
| dropout_93 (Dropout) | (None, 64) | 0 |
| dense_139 (Dense) | (None, 32) | 2,080 |
| dense_140 (Dense) | (None, 4) | 132 |

Total params: 6,852 (26.77 KB)
Trainable params: 6,660 (26.02 KB)
Non-trainable params: 192 (768.00 B)

Gambar 9. Total Parameter

Kemudian dilanjutkan dengan proses pelatihan model neural network menggunakan data pelatihan dengan berbagai kombinasi variabel dari 3 hingga 8 variabel, serta epoch yang bervariasi. Setelah setiap proses pelatihan, bobot model yang memberikan performa terbaik pada data pelatihan disimpan. Model yang digunakan untuk evaluasi dimuat dari file yang sesuai untuk memastikan performa terbaik selama pelatihan. Selanjutnya, dilakukan visualisasi hasil pelatihan model untuk masing-masing kombinasi variabel dan epoch yang digunakan..

3.8. Evaluasi

Evaluasi model dilakukan untuk mengetahui performa model yang digunakan dalam melakukan klasifikasi citra malaria. Evaluasi model dilakukan dengan perhitungan *confusion matrix*. Hasil *confusion matrix* berupa nilai akurasi, *precision*, *recall*, *f1-score*. Dasar perhitungan *confusion matrix* adalah dengan melakukan perbandingan dari total data yang dikelompokkan pada kelas yang benar dengan total dari keseluruhan data yang terdapat pada *matrix* [15]. Berikut pada persamaan rumus (12) hingga persamaan rumus (15) untuk perhitungan *confusion matrix*.

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Akurasi = \frac{TP + TN + FP + FN}{TP + TN + FP + FN} \quad (14)$$

$$F1\ Score = 2 \frac{Precision * Recall}{Precision + Recall} \quad (15)$$

Untuk analisis keseluruhan dari sebuah pengklasifikasi, dapat menggunakan *weighted average of precision, recall, and Fscore* [16]. Berikut pada persamaan (16) untuk perhitungan *weighted average*.

$$F_{weighted}^{\beta} Score = \frac{\sum_{c=0}^{q-1} F_c^{\beta} S_c}{\sum_{c=0}^{q-1} S_c} \quad (16)$$

dimana q adalah jumlah kelas, F_c^{β} adalah skor F^{β} kelas C dan S_c (kolom Support) adalah jumlah instance yang telah diberi label sebagai kelas C.

3.9. Kesimpulan dan Saran

Pada tahap ini akan ditarik kesimpulan terhadap penelitian ini dari hasil yang telah didapatkan. Selain itu, dilakukan pemberian saran yang bermanfaat sehingga dapat digunakan untuk penelitian selanjutnya sebagai bahan masukan.

IV. EVALUASI

4.1. Hasil Pengujian

Setelah melakukan kombinasi fitur yang berbeda dan menguji model menggunakan proporsi data 80%:20%, 70%:30%, 60%:40%, 90%:10%, diperoleh hasil evaluasi *confusion matrix* untuk beberapa kombinasi fitur. Evaluasi dilakukan menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score*, dengan rata-rata *weighted* untuk setiap kelas, yang memberikan bobot lebih besar pada kelas dengan jumlah data yang lebih banyak. Penggunaan rata-rata *weighted* dalam evaluasi ini memberikan bobot proporsional sesuai dengan jumlah sampel dalam setiap kelas, sehingga metrik yang dihasilkan mencerminkan performa model secara lebih akurat untuk kelas-kelas yang memiliki distribusi data tidak seimbang. Hasil pengujian tersebut meliputi kombinasi tiga hingga delapan fitur yang dievaluasi berdasarkan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Berikut adalah penjelasan untuk hasil yang diperoleh:

A. 80:20

1. Top Combination 3 Variable

Kombinasi fitur terbaik dengan tiga variabel adalah Mean, Std, dan Energy. Kombinasi ini menghasilkan *Accuracy* sebesar 79.49%, *Precision* 79.70%, *Recall* 79.49%, dan *F1-Score* 79.49%. Hasil ini menunjukkan bahwa kombinasi ini

memberikan performa yang lebih baik dibandingkan dengan kombinasi lain seperti Mean, Std, dan Entropy, yang memiliki Accuracy sebesar 76.92%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|---------------------------------|-------------|----------|-----------|--------|-------|
| 1 | 1 ('Mean', 'Std', 'Energy') | 281 | 79.49 | 79.70 | 79.49 | 79.49 |
| 2 | 2 ('Mean', 'Std', 'Entropy') | 250 | 76.92 | 77.84 | 76.92 | 76.68 |
| 3 | 3 ('Mean', 'Std', 'Smoothness') | 226 | 76.92 | 81.83 | 76.92 | 76.32 |

Gambar 10. Hasil Evaluasi Kombinasi Fitur dengan Tiga Variabel

2. Top Combination 4 Variable

Untuk kombinasi empat variabel, kombinasi terbaik adalah Std, Energy, Area Ratio, dan Eccentricity dengan Accuracy

sebesar 79.49% dan F1-Score sebesar 79.69%. Kombinasi ini berhasil menangkap variasi fitur dengan lebih baik, sehingga meningkatkan ketepatan klasifikasi.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|--|-------------|----------|-----------|--------|-------|
| 50 | 50 ('Std', 'Energy', 'Area Ratio', 'Eccentricity') | 232 | 79.49 | 80.09 | 79.49 | 79.69 |
| 49 | 49 ('Std', 'Energy', 'Smoothness', 'Eccentricity') | 214 | 79.49 | 80.12 | 79.49 | 79.58 |
| 46 | 46 ('Std', 'Energy', 'Entropy', 'Area Ratio') | 206 | 79.49 | 80.91 | 79.49 | 79.84 |

Gambar 11. Hasil Evaluasi Kombinasi Fitur dengan Empat Variabel

3. Top Combination 5 Variable

Pada kombinasi lima variabel, kombinasi Std, Energy, Entropy, Smoothness, dan Eccentricity memberikan

performa terbaik dengan Accuracy sebesar **79.49%** dan Precision **83.27%**. Penambahan variabel Smoothness dan Eccentricity pada kombinasi ini memperbaiki presisi tanpa mengorbankan akurasi..

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|--|-------------|----------|-----------|--------|-------|
| 46 | 46 ('Std', 'Energy', 'Entropy', 'Smoothness', 'Ec...') | 197 | 79.49 | 83.27 | 79.49 | 79.43 |
| 11 | 11 ('Mean', 'Std', 'Energy', 'Entropy', 'Area Rat...') | 296 | 79.49 | 81.29 | 79.49 | 79.46 |
| 30 | 30 ('Mean', 'Energy', 'Entropy', 'Smoothness', 'A...') | 215 | 79.49 | 80.09 | 79.49 | 79.69 |

Gambar 12. Hasil Evaluasi Kombinasi Fitur dengan Lima Variabel

4. Top Combination 6 Variable

Kombinasi terbaik dengan enam variabel adalah Mean, Energy, Entropy, Smoothness, Area Ratio, dan

Eccentricity, yang menghasilkan Accuracy 79.49%, Precision 80.09%, Recall 79.49%, Precision 80.09%, Recall 79.49%, dan F1-Score 79.69%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|--|-------------|----------|-----------|--------|-------|
| 20 | 20 ('Mean', 'Energy', 'Entropy', 'Smoothness', 'A...') | 262 | 79.49 | 80.09 | 79.49 | 79.69 |
| 12 | 12 ('Mean', 'Std', 'Energy', 'Entropy', 'Area Rat...') | 352 | 79.49 | 81.39 | 79.49 | 79.23 |
| 11 | 11 ('Mean', 'Std', 'Energy', 'Entropy', 'Smoothne...') | 323 | 79.49 | 83.76 | 79.49 | 79.46 |

Gambar 13. Hasil Evaluasi Kombinasi Fitur dengan Enam Variabel

5. Top Combination 7 Variable

Kombinasi tujuh variabel yaitu Mean, Std, Energy, Entropy, Smoothness, Area Ratio, dan Eccentricity,

menunjukkan penurunan performa dengan Accuracy 76.92% dan F1-Score 76.67%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|---|-------------|----------|-----------|--------|-------|
| 5 | 5 ('Mean', 'Std', 'Energy', 'Entropy', 'Smoothne...') | 250 | 76.92 | 79.40 | 76.92 | 76.67 |
| 0 | 0 ('Mean', 'Std', 'Skewness', 'Energy', 'Entropy...') | 227 | 74.36 | 77.33 | 74.36 | 75.35 |
| 1 | 1 ('Mean', 'Std', 'Skewness', 'Energy', 'Entropy...') | 239 | 74.36 | 80.60 | 74.36 | 72.68 |

Gambar 14. Hasil Evaluasi Kombinasi Fitur dengan Tujuh Variabel

6. Top Combination 8 Variable

Untuk kombinasi delapan variabel, Mean, Std, Skewness, Energy, Entropy, Smoothness, Area Ratio, dan

Eccentricity menghasilkan Accuracy 74.36%, Precision 75.75%, Recall 74.36%, dan F1-Score 74.48%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|---|-------------|----------|-----------|--------|-------|
| 0 | 0 ('Mean', 'Std', 'Skewness', 'Energy', 'Entropy... | 400 | 74.36 | 75.75 | 74.36 | 74.48 |

Gambar 15. Hasil Evaluasi Kombinasi Fitur dengan Delapan Variabel

B. 70:30

1. Top Combination 3 Variable

Kombinasi fitur terbaik dengan tiga variabel adalah Std, Energy, dan Entropy. Kombinasi ini menghasilkan Accuracy

sebesar 77.97%, Precision 80.95%, Recall 77.97%, dan F1-Score 78.59%. Kombinasi ini menunjukkan performa yang lebih unggul dibandingkan dengan kombinasi seperti Std, Energy, dan Area Ratio, yang memiliki Accuracy sebesar 76.27%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|------------------------------------|-------------|----------|-----------|--------|-------|
| 26 | 26 ('Std', 'Energy', 'Entropy') | 223 | 77.97 | 80.95 | 77.97 | 78.59 |
| 28 | 28 ('Std', 'Energy', 'Area Ratio') | 213 | 76.27 | 82.45 | 76.27 | 76.92 |
| 2 | 2 ('Mean', 'Std', 'Entropy') | 400 | 76.27 | 77.93 | 76.27 | 76.04 |

Gambar 16. Hasil Evaluasi Kombinasi Fitur dengan Tiga Variabel

2. Top Combination 4 Variable

Untuk kombinasi empat variabel, kombinasi terbaik adalah Std, Energy, Entropy, dan Eccentricity dengan Accuracy sebesar 77.97% dan F1-Score sebesar 78.24%. Kombinasi ini

mampu meningkatkan ketepatan klasifikasi dibandingkan kombinasi lainnya seperti Mean, Std, Entropy, dan Area Ratio.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|--|-------------|----------|-----------|--------|-------|
| 47 | 47 ('Std', 'Energy', 'Entropy', 'Eccentricity') | 209 | 77.97 | 80.00 | 77.97 | 78.24 |
| 10 | 10 ('Mean', 'Std', 'Entropy', 'Area Ratio') | 283 | 77.97 | 79.89 | 77.97 | 77.58 |
| 13 | 13 ('Mean', 'Std', 'Smoothness', 'Eccentricity') | 295 | 76.27 | 78.32 | 76.27 | 76.37 |

Gambar 17. Hasil Evaluasi Kombinasi Fitur dengan Empat Variabel

3. Top Combination 5 Variable

Kombinasi lima variabel terbaik adalah Std, Energy, Entropy, Area Ratio, dan Eccentricity, yang memberikan Accuracy sebesar 77.97% dan Precision 80.64%.

Penambahan variabel Area Ratio dan Eccentricity membantu meningkatkan presisi model tanpa mengurangi akurasi.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|--|-------------|----------|-----------|--------|-------|
| 47 | 47 ('Std', 'Energy', 'Entropy', 'Area Ratio', 'Ec... | 274 | 77.97 | 80.64 | 77.97 | 77.76 |
| 0 | 0 ('Mean', 'Std', 'Skewness', 'Energy', 'Entropy') | 400 | 76.27 | 78.95 | 76.27 | 76.07 |
| 13 | 13 ('Mean', 'Std', 'Energy', 'Smoothness', 'Area ... | 389 | 76.27 | 79.58 | 76.27 | 76.19 |

Gambar 18. Hasil Evaluasi Kombinasi Fitur dengan Lima Variabel

4. Top Combination 6 Variable

Kombinasi terbaik dengan enam variabel adalah Std, Energy, Entropy, Smoothness, Area Ratio, dan

Eccentricity. Kombinasi ini menghasilkan Accuracy tertinggi sebesar 79.66%, Precision 80.75%, Recall 79.66%, dan F1-Score 79.74%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|--|-------------|----------|-----------|--------|-------|
| 26 | 26 ('Std', 'Energy', 'Entropy', 'Smoothness', 'Ar... | 357 | 79.66 | 80.75 | 79.66 | 79.74 |
| 20 | 20 ('Mean', 'Energy', 'Entropy', 'Smoothness', 'A... | 400 | 77.97 | 77.84 | 77.97 | 77.68 |
| 5 | 5 ('Mean', 'Std', 'Skewness', 'Energy', 'Area Ra... | 384 | 76.27 | 77.54 | 76.27 | 76.25 |

Gambar 19. Hasil Evaluasi Kombinasi Fitur dengan Enam Variabel

5. Top Combination 7 Variable

Kombinasi tujuh variabel yaitu Mean, Std, Skewness, Energy, Smoothness, Area Ratio, dan Eccentricity menunjukkan penurunan performa dengan Accuracy 76.27% dan F1-Score 75.94%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|---|-------------|----------|-----------|--------|-------|
| 3 | 3 ('Mean', 'Std', 'Skewness', 'Energy', 'Smoothn... | 287 | 76.27 | 77.39 | 76.27 | 75.94 |
| 5 | 5 ('Mean', 'Std', 'Energy', 'Entropy', 'Smoothne... | 400 | 76.27 | 78.00 | 76.27 | 76.15 |
| 7 | 7 ('Std', 'Skewness', 'Energy', 'Entropy', 'Smoo... | 400 | 76.27 | 77.56 | 76.27 | 76.54 |

Gambar 20. Hasil Evaluasi Kombinasi Fitur dengan Tujuh Variabel

6. Top Combination 8 Variable

Untuk kombinasi delapan variabel, *Mean*, *Std*, *Skewness*, *Energy*, *Entropy*, *Smoothness*, *Area Ratio*, dan *Eccentricity*

menghasilkan *Accuracy* 71.19%, *Precision* 81.36%, *Recall* 71.19%, dan *F1-Score* 70.23%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|---|-------------|----------|-----------|--------|-------|
| 0 | 0 ('Mean', 'Std', 'Skewness', 'Energy', 'Entropy... | 390 | 71.19 | 81.36 | 71.19 | 70.23 |

Gambar 21. Hasil Evaluasi Kombinasi Fitur dengan Delapan Variabel

C. 60:40

1. Top Combination 3 Variable

Kombinasi fitur terbaik dengan tiga variabel adalah *Mean*, *Energy*, dan *Entropy*. Kombinasi ini menghasilkan *Accuracy* sebesar 75.64%, *Precision* 77.61%, *Recall* 75.64%, dan *F1-*

Score 74.20%. Hasil ini menunjukkan bahwa kombinasi ini memberikan performa yang lebih baik dibandingkan dengan kombinasi lain seperti *Std*, *Energy*, dan *Eccentricity*, yang memiliki *Accuracy* sebesar 71.79%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|---------------------------------------|-------------|----------|-----------|--------|-------|
| 11 | 11 ('Mean', 'Energy', 'Entropy') | 400 | 75.64 | 77.61 | 75.64 | 74.20 |
| 29 | 29 ('Std', 'Energy', 'Eccentricity') | 260 | 71.79 | 79.91 | 71.79 | 72.05 |
| 32 | 32 ('Std', 'Entropy', 'Eccentricity') | 292 | 71.79 | 79.31 | 71.79 | 72.40 |

Gambar 16. Hasil Evaluasi Kombinasi Fitur dengan Tiga Variabel

2. Top Combination 4 Variable

Untuk kombinasi empat variabel, kombinasi terbaik adalah *Mean*, *Energy*, *Entropy*, dan *Eccentricity* dengan *Accuracy*

sebesar 75.64% dan *F1-Score* sebesar 75.54%. Kombinasi ini berhasil menangkap variasi fitur dengan lebih baik, sehingga meningkatkan ketepatan klasifikasi.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|--|-------------|----------|-----------|--------|-------|
| 27 | 27 ('Mean', 'Energy', 'Entropy', 'Eccentricity') | 400 | 75.64 | 76.59 | 75.64 | 75.54 |
| 6 | 6 ('Mean', 'Std', 'Energy', 'Smoothness') | 357 | 74.36 | 76.18 | 74.36 | 74.90 |
| 11 | 11 ('Mean', 'Std', 'Entropy', 'Eccentricity') | 260 | 73.08 | 75.41 | 73.08 | 73.69 |

Gambar 17. Hasil Evaluasi Kombinasi Fitur dengan Empat Variabel

3. Top Combination 5 Variable

Pada kombinasi lima variabel, kombinasi *Mean*, *Std*, *Energy*, *Entropy*, dan *Eccentricity* memberikan performa terbaik

dengan *Accuracy* sebesar 74.36% dan *Precision* 76.84%. Penambahan variabel *Entropy* dan *Eccentricity* pada kombinasi ini memperbaiki presisi tanpa mengorbankan akurasi.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|--|-------------|----------|-----------|--------|-------|
| 12 | 12 ('Mean', 'Std', 'Energy', 'Entropy', 'Eccentri... | 400 | 74.36 | 76.84 | 74.36 | 74.80 |
| 30 | 30 ('Mean', 'Energy', 'Entropy', 'Smoothness', 'A... | 400 | 73.08 | 73.16 | 73.08 | 73.05 |
| 13 | 13 ('Mean', 'Std', 'Energy', 'Smoothness', 'Area ... | 400 | 71.79 | 74.97 | 71.79 | 72.36 |

Gambar 18. Hasil Evaluasi Kombinasi Fitur dengan Lima Variabel

4. Top Combination 6 Variable

Kombinasi terbaik dengan enam variabel adalah *Mean*, *Std*, *Energy*, *Entropy*, *Area Ratio*, dan *Smoothness*, yang

menghasilkan *Accuracy* 74.36%, *Precision* 76.18%, *Recall* 74.36%, dan *F1-Score* 74.90%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|--|-------------|----------|-----------|--------|-------|
| 12 | 12 ('Mean', 'Std', 'Energy', 'Entropy', 'Area Rat... | 360 | 74.36 | 76.18 | 74.36 | 74.90 |
| 10 | 10 ('Mean', 'Std', 'Energy', 'Entropy', 'Smoothne... | 325 | 74.36 | 76.25 | 74.36 | 74.60 |
| 14 | 14 ('Mean', 'Std', 'Entropy', 'Smoothness', 'Area... | 400 | 71.79 | 74.12 | 71.79 | 72.04 |

Gambar 19. Hasil Evaluasi Kombinasi Fitur dengan Enam Variabel

5. Top Combination 7 Variable

Kombinasi tujuh variabel yaitu *Mean*, *Std*, *Energy*, *Entropy*, *Smoothness*, *Area Ratio*, dan *Eccentricity*, menunjukkan

penurunan performa dengan *Accuracy* 73.08% dan *F1-Score* 72.50%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|---|-------------|----------|-----------|--------|-------|
| 5 | 5 ('Mean', 'Std', 'Energy', 'Entropy', 'Smoothne... | 293 | 73.08 | 73.96 | 73.08 | 72.50 |
| 1 | 1 ('Mean', 'Std', 'Skewness', 'Energy', 'Entropy... | 400 | 69.23 | 74.94 | 69.23 | 68.67 |
| 2 | 2 ('Mean', 'Std', 'Skewness', 'Energy', 'Entropy... | 400 | 69.23 | 73.06 | 69.23 | 69.94 |

Gambar 20. Hasil Evaluasi Kombinasi Fitur dengan Tujuh Variabel

6. Top Combination 8 Variable

Untuk kombinasi delapan variabel, *Mean*, *Std*, *Skewness*, *Energy*, *Entropy*, *Smoothness*, *Area Ratio*, dan *Eccentricity*

menghasilkan *Accuracy* 67.95%, *Precision* 70.66%, *Recall* 67.95%, dan *F1-Score* 68.42%.

| Unnamed: 0 | Combinations | Final Epoch | Accuracy | Precision | Recall | F1 |
|------------|---|-------------|----------|-----------|--------|-------|
| 0 | 0 ('Mean', 'Std', 'Skewness', 'Energy', 'Entropy... | 400 | 67.95 | 70.66 | 67.95 | 68.42 |

Gambar 21. Hasil Evaluasi Kombinasi Fitur dengan Delapan Variabel

D. 90:10

1. Top Combination 3 Variable

Kombinasi fitur terbaik dengan tiga variabel adalah *Mean*, *Std*, dan *Energy*. Kombinasi ini menghasilkan *Accuracy* sebesar 80.0%, *Precision* 84.62%, *Recall* 80.0%, dan *F1-Score* 80.41%. Hasil ini menunjukkan bahwa kombinasi ini memberikan performa terbaik dibandingkan dengan kombinasi lain, seperti *Mean*, *Std*, dan *Skewness*, yang memiliki *Accuracy* sebesar 75.0%.

kombinasi ini memperbaiki presisi tanpa mengorbankan akurasi.

2. Top Combination 4 Variable

Untuk kombinasi empat variabel, kombinasi terbaik adalah *Mean*, *Std*, *Energy*, dan *Entropy* dengan *Accuracy* sebesar 80.0% dan *F1-Score* sebesar 80.20%. Kombinasi ini berhasil menangkap variasi fitur dengan lebih baik, sehingga meningkatkan ketepatan klasifikasi dibandingkan dengan kombinasi lain seperti *Mean*, *Std*, *Skewness*, dan *Eccentricity*, yang memiliki *Accuracy* sebesar 75.0%.

4. Top Combination 6 Variable

Kombinasi terbaik dengan enam variabel adalah *Mean*, *Energy*, *Entropy*, *Smoothness*, *Area Ratio*, dan *Eccentricity*, yang menghasilkan *Accuracy* 80.0%, *Precision* 80.83%, *Recall* 80.0%, dan *F1-Score* 79.95%. Kombinasi ini menunjukkan performa yang sangat baik dalam menangkap variasi fitur dan meningkatkan ketepatan klasifikasi.

3. Top Combination 5 Variable

Pada kombinasi lima variabel, kombinasi *Mean*, *Std*, *Energy*, *Entropy*, dan *Eccentricity* memberikan performa terbaik dengan *Accuracy* sebesar 80.0% dan *Precision* 80.36%. Penambahan variabel *Entropy* dan *Eccentricity* pada

5. Top Combination 7 Variable

Kombinasi tujuh variabel yaitu *Mean*, *Std*, *Energy*, *Entropy*, *Smoothness*, *Area Ratio*, dan *Eccentricity*, menunjukkan performa dengan *Accuracy* 75.0% dan *F1-Score* 74.85%. Penambahan variabel ini tidak memberikan peningkatan yang signifikan dibandingkan dengan kombinasi enam variabel.

Setelah tahap pembuatan model dilakukan, selanjutnya model tersebut dilakukan evaluasi. Evaluasi pada penelitian ini menggunakan *confusion matrix* dari proporsi data yang ada seperti pada Tabel 2. Berikut merupakan salah satu perhitungan *confusion matrix* yang menggunakan proporsi 80%:20% berdasarkan persamaan rumus (12) hingga (15).

$$Recall = \frac{\frac{4}{7} + \frac{9}{11} + \frac{10}{10} + \frac{8}{11}}{4} = \frac{0.57 + 0.82 + 1.00 + 0.73}{4} = 0.79$$

$$Precision = \frac{\frac{4}{4+0+0+0} + \frac{9}{9+0+0+0} + \frac{10}{0+0+10+0} + \frac{8}{2+2+0+8}}{4} = \frac{1.00 + 1.00 + 0.71 + 0.67}{4} = 0.83$$

$$Akurasi = \frac{4 + 9 + 10 + 8}{39} = \frac{31}{39} = 0.79$$

$$F1 - Score = 2 \times \frac{0.73 + 0.90 + 0.83 + 0.70}{4} = 0.79$$

Dari hasil evaluasi perhitungan diatas dapat dilihat hasil *confusion matrix* dari keseluruhan proporsi split data yang digunakan pada Tabel 2.

Tabel 2. Hasil Confusion Matrix

| Proporsi | Confusion Matrix | | | |
|----------|------------------|-----------|--------|----------|
| | Accuracy | Precision | Recall | F1-Score |
| 60%:40% | 0.75 | 0.76 | 0.75 | 0.75 |
| 70%:30% | 0.78 | 0.81 | 0.78 | 0.78 |
| 80%:20% | 0.79 | 0.83 | 0.79 | 0.79 |
| 90%:10% | 0.80 | 0.84 | 0.80 | 0.80 |

4.2. Analisis Hasil Pengujian

Analisis hasil pengujian model berdasarkan *sensitivity* menunjukkan pentingnya memilih kombinasi fitur yang tepat untuk menjaga kemampuan model dalam mendeteksi *instance* positif. Berikut adalah ringkasan analisis performa berdasarkan sensitivitas untuk berbagai proporsi data

a. 80:20

Pada proporsi ini, kombinasi fitur terbaik untuk tiga variabel adalah '*Mean*', '*Std*', dan '*Energy*', dengan recall mencapai 79.49%. Kombinasi ini menunjukkan performa deteksi positif yang kuat dibandingkan dengan kombinasi lainnya. Untuk empat variabel, '*Std*', '*Energy*', '*Area Ratio*', dan '*Eccentricity*' juga mencapai recall 79.49%. Meskipun penambahan fitur pada lima variabel tidak meningkatkan recall, tetapi memperbaiki presisi. Namun, kombinasi dengan tujuh dan delapan variabel menunjukkan penurunan recall, mengindikasikan bahwa terlalu banyak fitur dapat menurunkan kemampuan model dalam mendeteksi *instance* positif.

b. 70:30

Pada proporsi ini, kombinasi '*Std*', '*Energy*', dan '*Entropy*' menghasilkan recall 77.97%, menunjukkan performa lebih baik dibandingkan kombinasi lain. Untuk empat variabel, '*Std*', '*Energy*', '*Entropy*', dan '*Eccentricity*' juga mencapai recall 77.97%. Peningkatan jumlah fitur pada enam variabel menunjukkan recall tertinggi sebesar 79.66%. Namun, kombinasi dengan tujuh dan delapan variabel menunjukkan penurunan recall, menyiratkan bahwa penambahan fitur yang berlebihan tidak selalu meningkatkan performa deteksi.

c. 60:40

Kombinasi terbaik untuk tiga variabel adalah '*Mean*', '*Energy*', dan '*Entropy*', dengan recall 75.64%. Kombinasi ini memberikan performa yang lebih baik dibandingkan dengan kombinasi lainnya. Untuk empat variabel, '*Mean*', '*Energy*', '*Entropy*', dan '*Eccentricity*' juga menunjukkan recall 75.64%. Peningkatan fitur pada enam variabel menghasilkan recall yang stabil pada 74.36%, meskipun presisi meningkat. Namun, penambahan fitur pada tujuh dan delapan variabel

menunjukkan penurunan recall, mengindikasikan kesulitan dalam mempertahankan tingkat recall dengan banyak fitur.

d. 90:10

Kombinasi fitur terbaik pada proporsi ini adalah '*Mean*', '*Std*', dan '*Energy*', dengan recall 80.0%, yang merupakan performa tertinggi dibandingkan kombinasi lainnya. Kombinasi dengan empat dan lima variabel juga mencapai recall 80.0%. Meskipun penambahan fitur pada enam variabel tetap menghasilkan recall 80.0%, kombinasi dengan tujuh variabel menunjukkan penurunan menjadi 75.0%, tanpa peningkatan signifikan.

Secara keseluruhan, hasil analisis dari keempat proporsi data menunjukkan bahwa kombinasi fitur tertentu dapat secara konsisten mempertahankan sensitivitas yang tinggi. Pada proporsi 80%:20%, sensitivitas stabil pada kombinasi tiga hingga lima variabel, namun menurun pada tujuh hingga delapan variabel. Pada proporsi 70%:30%, sensitivitas tertinggi dicapai dengan enam variabel, tetapi penambahan lebih lanjut menurunkan performa. Di proporsi 60:40, sensitivitas cenderung stabil dengan tiga hingga enam variabel, namun menurun dengan tujuh hingga delapan variabel. Sementara itu, pada proporsi 90:10, sensitivitas tertinggi dicapai pada kombinasi tiga hingga enam variabel, namun menurun pada tujuh variabel. Kesimpulannya, penambahan fitur yang berlebihan cenderung menurunkan sensitivitas, dan memilih jumlah fitur yang optimal dan sangat penting untuk mempertahankan performa model dalam mendeteksi *instance* positif

V. KESIMPULAN

Berdasarkan hasil penelitian pada klasifikasi citra sel darah merah untuk fase penyakit Malaria *falciparum* menggunakan 1D-CNN dengan kombinasi fitur ekstraksi tekstur dan morfologi, penelitian ini berhasil diimplementasikan menggunakan bahasa pemrograman *Python* dengan beberapa *library* yang mendukung proses klasifikasi. Penggunaan teknik ekstraksi fitur seperti perhitungan mean, standard deviation, skewness, energy, entropy, dan smoothness pada ekstraksi tekstur, serta area ratio pada ekstraksi morfologi, memungkinkan model untuk memiliki representasi data yang lebih baik. Dari hasil evaluasi, proporsi data 90%:10% menunjukkan performa terbaik dengan akurasi 80%, precision 84%, recall 80%, dan F1-Score 80%. Ini menjadikannya proporsi data yang paling efektif dalam meningkatkan kinerja model dibandingkan dengan proporsi lainnya. Diikuti oleh proporsi data 80%:20% dengan akurasi 79%, precision 83%, recall 79%, dan F1-Score 79%. Meskipun hasilnya sangat baik, proporsi ini tidak sebaik proporsi 90%:10% dalam hal metrik kinerja keseluruhan. Sebagai perbandingan, proporsi data 70%:30% memberikan akurasi 78%, precision 81%, recall 78%, dan F1-Score 78%, sedangkan proporsi 60%:40% menghasilkan akurasi 75%, precision 76%, recall 75%, dan F1-Score 75%. Hasil ini menunjukkan bahwa peningkatan proporsi data untuk pelatihan umumnya meningkatkan kinerja model.

Berdasarkan hasil penelitian ini, saran untuk penelitian selanjutnya adalah melakukan perbandingan dengan menggunakan *layer* konvolusi lain seperti 2D CNN untuk memperluas pemahaman mengenai hasil dari berbagai jenis konvolusi CNN pada data yang sama. Hal ini diharapkan dapat memberikan *insight* tambahan tentang bagaimana struktur konvolusi mempengaruhi performa model dalam klasifikasi citra.

REFERENSI

- [1] I. Budiningsih, J. M. Middeldorp, Y. P. Dachlan and U. Hadi, "Epstein-Barr Virus and Malaria Interactions: Immunology Perspective," *Hayati Journal of Biosciences*, pp. 824-833, 2022.
- [2] A. W. Setiawan, Y. A. Rahman, A. Faisal, M. Siburian, N. Resfita, M. W. Gifari and R. Setiawan, "DETEKSI MALARIA BERBASIS SEGMENTASI WARNA CITRA DAN PEMBELAJARAN MESIN," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, pp. 769-776, 2021.
- [3] D. S. Yohannes and K. Arianto, "Deteksi Penyakit Malaria Menggunakan Convolutional Neural Network Berbasis Saliency (Detection of Malaria Using Convolutional Neural Network Based on Saliency)," *JUITA: Jurnal Informatika*, pp. 37-44, 2020.
- [4] D. Setyawan, R. Wardoyo, M. E. Wibowo and E. H. Murhandarwati, "Classification of plasmodium falciparum based on textural and morphological features," *International Journal of Electrical and Computer Engineering (IJECE)*, pp. 5036-5048, 2022.
- [5] Y. J. Nurriski and Alamsyah, "Optimasi Deep Convolutional Neural Network (Deep CNN) untuk Deteksi Aritmia Melalui Sinyal EKG Menggunakan Arsitektur Conv1D," *Indonesian Journal of Mathematics and Natural Sciences*, vol. 46, pp. 10-21, 2023.
- [6] Y. Florensia, Samsuryadi and Saparudin, "Klasifikasi Citra Hiperspektral Pada Kasus Tutupan Lahan Menggunakan Metode Convolutional Neural Network," *Prosiding Annual Research Seminar 2019*, vol. 5, no. 1, pp. 262-266, 2020.
- [7] N. Dewi and F. Ismawan, "Implementasi Deep Learning Menggunakan Convolutional Neural Network untuk Sistem Pengenalan Wajah," *Journal homepage*, vol. 14, pp. 34-43, 2021.
- [8] B. M. Devassy and S. George, "Contactless Classification of Strawberry Using Hyperspectral Imaging," *CEUR Workshop Proceedings*, pp. 1-12, 2020.
- [9] Y. N. Fuadah, S. Saidah, N. K. Sy, R. Magdalena and I. D. Ubaidullah, "GLAUCOMA CLASSIFICATION BASED ON FUNDUS IMAGES PROCESSING WITH CONVOLUTIONAL NEURAL NETWORK," *Jurnal Teknik Informatika (JUTIF)*, pp. 717-722, 2022.
- [10] R. Magdalena, S. Saidah, I. D. S. Ubaidah, Y. N. Fuadah, N. Herman and N. Ibrahim, "CONVOLUTIONAL NEURAL NETWORK FOR ANEMIA DETECTION BASED ON CONJUNCTIVA PALPEBRAL IMAGES," *Jurnal Teknik Informatika (JUTIF)*, pp. 349-354, 2022.
- [11] M. A. Pramudito, Y. N. Fu'adah, R. Magdalena, A. Rizal and F. F. Taliningsih, "Identifikasi Sinyal Congestive Heart Failure dengan Metode Convolutional Neural Network 1D," *MIND (Multimedia Artificial Intelligent Networking Database) Journal*, pp. 11-20, 2022.
- [12] T.-H. Hsieh and J.-F. Kiang, "Comparison of CNN Algorithms on Hyperspectral Image Classification in Agricultural Lands," *sensors*, pp. 1-17, 2020.
- [13] A. Fadjeri, A. Setyanto and M. P. Kurniawan, "Pengolahan Citra Digital Untuk Menghitung Ekstraksi Ciri Greenbean Kopi Robusta Dan Arabika (Studi Kasus: Kopi Temanggung)A," *Jurnal TIKomSiN*, pp. 8-14, 2020.
- [14] A. Wibowo and H. Syahputra, "Sistem Deteksi Konten Negatif pada Teks Website Menggunakan Metode Random Forest," *Journal of Informatics and Data Science (J-IDS)*, vol. 1, pp. 1-5, 2022.
- [15] D. Sartika and I. Saluza, "Penerapan Metode Principal Component Analysis (PCA) Pada Klasifikasi Status Kredit Nasabah Bank Sumsel Babel Cabang KM 12 Palembang Menggunakan Metode Decision Tree," *GENERIC Jurnal Ilmu Komputer dan Teknologi Informasi*, vol. 14, pp. 45-49, 2022.
- [16] T. F. Basar, D. E. Ratnawati and I. Arwani, "Analisis Sentimen Pengguna Twitter terhadap Pembayaran Cashless menggunakan ShopeePAY dengan Algoritma Random Forest," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, pp. 1426-1433, 2022.