

# KLASIFIKASI KUALITAS PEWARNAAN ZIEHL NEELSEN PADA SAMPEL DAHAK TBC DENGAN MENGGUNAKAN MACHINE LEARNING

1<sup>st</sup> Faris Naufal Arif

*School of Electrical Engineering*

*Telkom University*

Bandung, Indonesia

farissun@student.telkomuniversity.ac.id

2<sup>nd</sup> Dilazapira

*School of Electrical Engineering*

*Telkom University*

Bandung, Indonesia

dilazapira@telkomuniversity.ac.id

3<sup>rd</sup> Siti Sadira Muntaz Dinar

*School of Electrical Engineering*

*Telkom University*

Bandung, Indonesia

sadiramuntaz@telkomuniversity.ac.id

4<sup>th</sup> Inung Wijayanto

*School of Electrical Engineering*

*Telkom University*

Bandung, Indonesia

iwijayanto@telkomuniversity.ac.id

5<sup>th</sup> Suci Aulia

*School of Applied Sciences*

*Telkom University*

Bandung, Indonesia

suciaulia@telkomuniversity.ac.id

**Abstrak**—Penyakit Tuberkulosis (TBC) masih menjadi masalah kesehatan global yang memerlukan diagnosis yang cepat dan akurat. Salah satu metode diagnosis yang umum digunakan adalah pemeriksaan mikroskopis dengan pewarnaan *Ziehl-Neelsen* (ZN) pada sampel dahak. Namun, kualitas pewarnaan yang kurang optimal seperti *Over Staining* atau *Less Staining* dapat memengaruhi akurasi pembacaan mikroskopis dan menyebabkan kesalahan diagnosis. Untuk mengatasi masalah tersebut, penelitian ini mengembangkan sebuah aplikasi *standalone* berbasis ML yang dapat melakukan klasifikasi kualitas pewarnaan ZN dan menghitung jumlah bakteri TBC secara otomatis. Sistem ini dibangun agar dapat berjalan secara lokal (*offline*) dan terintegrasi antara model klasifikasi berbasis CNN dan deteksi bakteri berbasis YOLOv11. Berdasarkan hasil pengujian terhadap *dataset* uji, sistem klasifikasi dengan model CNN mencapai akurasi, presisi, sensitivitas, spesifisitas, dan F1-skor sebesar 100% pada setiap kelas. Pada tahap deteksi bakteri, model YOLOv11 menunjukkan performa evaluasi dengan mAP, *precision*, *recall*, dan F1-skor lebih dari 80%. Waktu rata-rata pemrosesan gambar pada 30 laptop berbeda tercatat kurang dari 3 detik per gambar. Hasil ini menunjukkan bahwa sistem yang dikembangkan memiliki performa tinggi, fleksibel, dan siap diimplementasikan dalam skala operasional laboratorium maupun daerah dengan keterbatasan jaringan.

**Kata Kunci**—Aplikasi *standalone*, Klasifikasi gambar, Machine Learning, Tuberkulosis, *Ziehl Neelsen*.

## I. PENDAHULUAN

*Tuberculosis* atau biasa dikenal dengan singkatan TBC merupakan penyakit menular yang disebabkan oleh bakteri *Mycobacterium tuberculosis* (MTBC) [1]. Bakteri ini menyebar ketika pasien yang sudah terjangkit penyakit TBC mengeluarkan bakteri ke udara seperti batuk, bersin, atau membuang air liur sembarangan [2]. Menurut *World Health Organization* (WHO) Report 2023, Indonesia menempati peringkat dua di dunia dengan jumlah kasus TBC terbanyak

setelah India [2]. Kasus penyakit TBC di Indonesia mencapai 800 ribu pasien dan kasus kematian mencapai 23 ribu pasien yang meninggal dikarenakan TBC [1].

Pasien yang terjangkit bakteri TBC perlu waktu 6 bulan untuk tahap penyembuhan. Terdapat dua kemungkinan yang terjadi saat pasien berada pada tahap penyembuhan, yaitu sembuh atau *Multidrug-Resistant Tuberculosis* (MDR-TBC) [3]. MDR-TBC merupakan kondisi dimana pasien kebal terhadap antibiotik yang diberikan, sehingga dokter memberikan tingkat antibiotik yang lebih tinggi untuk diberikan kepada pasien [4]. MDR-TBC terjadi ketika pasien tidak menyelesaikan pengobatan sesuai dosis dan durasi yang ditentukan atau tertular oleh pasien yang memiliki bakteri TBC yang resisten terhadap antibiotik pada dosis yang diberikan [5].

Pada pengobatan TBC, screening secara sistematis sangat diperlukan untuk mendapatkan sampel. Hal tersebut dilakukan dengan menggunakan tes, pemeriksaan, atau prosedur lain yang dapat diterapkan dengan cepat [6]. Beberapa jenis screening utama yang dilakukan pada pengobatan TBC meliputi *Interferon-Gamma Release Assays* (IGRA), *Rontgen*, *Polymerase Chain Reaction* (PCR), dan konvensional mikroskopis.

Metode yang sesuai untuk diagnosis TBC di Indonesia adalah mikroskopis sputum smear karena efektif mengidentifikasi penderita TBC menular [7] dan menjadi pilihan utama di negara berpenghasilan rendah-menengah (LMIC). Metode ini efektif secara biaya untuk diagnosis, pemantauan pengobatan, dan konfirmasi kesembuhan [8]. Pewarnaan *Ziehl Neelsen* (ZN) dipilih karena murah dan perawatan alatnya mudah [9], menargetkan MTBC yang terlihat merah

pada latar biru karena sifatnya tahan asam [10].

Salah satu hambatan pemberantasan TBC adalah belum adanya diagnosis cepat dan tepat untuk mendeteksi bakteri MTBC dalam sputum [11], yang penting untuk pengobatan dan memutus rantai penularan [12]. Salah satu metode umum adalah *GeneXpert* dengan pewarnaan ZN, yang memberikan kontras basil merah pada latar biru cerah dan biaya reagen rendah [13]. Namun, metode ini memiliki keterbatasan interpretasi dan prosesnya dapat memakan waktu hingga sebulan [14]. Hasil pemeriksaan juga dipengaruhi keterampilan petugas dan jenis mikroskop [15]. Penelitian ini bertujuan mempercepat dan meningkatkan akurasi analisis, serta mengurangi kesalahan manusia.

## II. KAJIAN TEORI

### A. Convolutional Neural Network (CNN)

CNN adalah arsitektur deep learning yang terdiri dari tahap input, pemrosesan, dan output, direkomendasikan untuk mengenali pola spasial [16]. Tahap input menerima data mentah seperti gambar atau video, diproses melalui convolutional, pooling, dan fully connected layers untuk mengekstraksi fitur, mengurangi dimensi, dan menghasilkan prediksi [17]. CNN efektif mendeteksi pola, tekstur, tepi, dan bentuk, sehingga cocok untuk deteksi objek, segmentasi, dan klasifikasi citra.

Terdapat beberapa tahapan CNN yang meliputi operasi konvolusi dan penggabungan lapisan untuk memproses data. Terdapat 4 tahapan pada arsitektur CNN yang meliputi tahap masukan (input), tahap ekstraksi fitur, tahap klasifikasi, dan tahap keluaran (output) [17].

### B. Confusion Matrik

*Confusion matrix* merupakan alat yang sering digunakan untuk mengevaluasi kinerja model klasifikasi dalam ML. Matriks ini berbentuk tabel yang berisikan jumlah prediksi yang benar dan salah yang dibuat oleh model untuk setiap kelas dalam data uji. Pada *Confusion Matrix*, terdiri dari 4 komponen yaitu true positive (TP), true negative (TN), false positive (FP), dan false negative (FN) [18]. TP merupakan jumlah data yang benar diklasifikasikan ke kelas positif. TN merupakan jumlah data yang benar diklasifikasikan ke kelas negatif. FP merupakan jumlah data yang salah diklasifikasikan ke kelas positif. FN merupakan jumlah data yang salah di klasifikasikan ke kelas negatif.

Tabel I: Tabel Confusion Matrix

| Kelas Aktual | Kelas Prediksi |          |
|--------------|----------------|----------|
|              | Positive       | Negative |
| Positive     | TP             | FN       |
| Negative     | FP             | TN       |

Tabel I merupakan tabel posisi pada setiap komponen *confusion matrik*. Dengan 4 komponen tersebut dapat mengevaluasi aspek-aspek spesifik dari kinerja model, seperti akurasi, presisi, *recall* (Sensitivitas), spesifisitas, dan F1-Skor.

Berikut merupakan penjelasan mengenai Akurasi, Precision, Recall (Sensitivitas), Specificity, dan F1-skor:

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

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

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4)$$

$$\text{F1-skor} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

### C. YOLOv11

*Computer Vision* (CV) memungkinkan mesin memahami data visual, salah satunya melalui deteksi objek [19]. Deteksi objek melibatkan identifikasi dan lokalisasi objek secara tepat pada gambar atau video [20]. Kemajuan besar dicapai dengan hadirnya *You Only Look Once* (YOLO) pada 2015 [21], yang memandang deteksi objek sebagai masalah regresi dan menggunakan satu *convolutional network* untuk memprediksi batas dan probabilitas kelas sekaligus [22]. Versi terbarunya, YOLOv11 (2024), menghadirkan peningkatan arsitektur dan pelatihan sehingga lebih akurat, cepat, dan efisien [23].

YOLO merevolusi deteksi objek dengan arsitektur neural network yang menggabungkan regresi *bounding box* dan klasifikasi objek secara bersamaan [24], berbeda dari metode dua tahap tradisional. Arsitekturnya terdiri dari tiga komponen: *backbone* untuk ekstraksi fitur menjadi *multi-scale feature maps*, *neck* untuk mengagregasi dan meningkatkan representasi fitur, serta *head* untuk memprediksi lokalisasi dan klasifikasi objek [23].

### D. Roboflow

Roboflow adalah *platform* pengembangan model *computer vision* yang mencakup pelabelan data, pra-pemrosesan, pelatihan, hingga deployment [25]. Platform ini mendukung berbagai format gambar, menyediakan augmentasi data, terintegrasi dengan *framework* seperti TensorFlow, PyTorch, dan YOLO, serta memudahkan pengelolaan dataset untuk penerapan di dunia nyata.

### E. Random Search

*Random Search* adalah metode untuk memilih nilai-nilai hyperparameter secara acak dan mengevaluasi kinerjanya di setiap iterasi [26]. Setelah beberapa percobaan, kombinasi dengan performa terbaik akan dipilih. Metode ini cocok digunakan ketika jumlah hyperparameter banyak dan ruang pencariannya luas.

### F. Pretrained model

Model pra-trained adalah pendekatan yang menggunakan model yang sudah dilatih sebelumnya sebagai titik awal untuk menyelesaikan masalah baru. Dengan cara ini, parameter model hanya perlu diperbarui (*fine-tuned*) agar sesuai dengan tugas spesifik, sehingga proses pelatihan menjadi lebih cepat dan efisien dibanding melatih model dari awal [27].

### G. Adam Optimizer

Adam (*Adaptive Moment Estimation*) adalah algoritma optimasi yang menggabungkan RMSProp dan Momentum, menyesuaikan *learning rate* setiap parameter secara adaptif [28]. Optimizer ini populer di *deep learning* karena efisien untuk dataset besar dan parameter kompleks, serta mampu menyesuaikan *learning rate* secara dinamis selama pelatihan untuk hasil optimal tanpa penyesuaian manual

### H. Visual Studio Code

Visual Studio Code (VSCode) adalah teks editor multiplatform buatan Microsoft yang tersedia untuk Linux, Mac, dan Windows [29]. Dirilis pada 29 April 2015, VSCode mendukung berbagai bahasa pemrograman seperti C, C++, Java, dan Python, serta dilengkapi fitur seperti Git Integration, Debugging, dan ekstensi [30].

## III. METODE

### A. Spesifikasi Sistem

Untuk memastikan aplikasi TBClassify yang dirancang dapat memenuhi standar yang ditetapkan, diperlukan pendekatan yang cermat dalam pengukuran dan spesifikasi. Tujuan dari spesifikasi sistem adalah untuk menetapkan batasan yang diperlukan oleh sistem guna mencapai hasil yang optimal. Berikut merupakan tabel batasan dan spesifikasi utama yang telah ditetapkan untuk mengukur keberhasilan solusi yang dirancang.

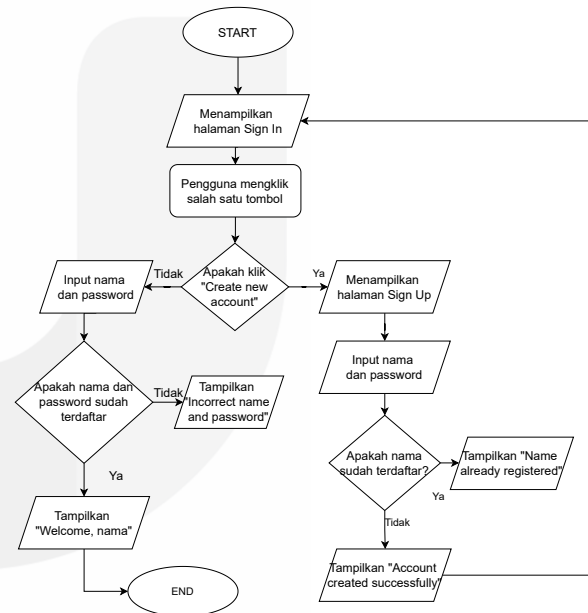
Berdasarkan Tabel II, spesifikasi dan batasan TBClassify menjadi acuan utama untuk mengukur keberhasilan pengembangan, meliputi akurasi klasifikasi, deteksi bakteri, validasi logika, dan kepuasan pengguna. Batasan ini memberikan kerangka kerja yang jelas untuk desain dan implementasi, sehingga aplikasi memenuhi kebutuhan teknis serta standar ilmiah, legal, dan etika yang berlaku.

### B. Desain Sistem

TBClassify adalah aplikasi *standalone* untuk klasifikasi kualitas pewarnaan ZN serta deteksi dan penghitungan bakteri TBC menggunakan ML. Pemrosesan dilakukan secara lokal, sehingga dapat digunakan di laboratorium atau fasilitas kesehatan dengan keterbatasan konektivitas. Model CNN mengklasifikasi kualitas pewarnaan ZN menjadi tiga kategori, sedangkan YOLO mendeteksi dan menghitung bakteri TBC.

Tabel II: Spesifikasi Sistem

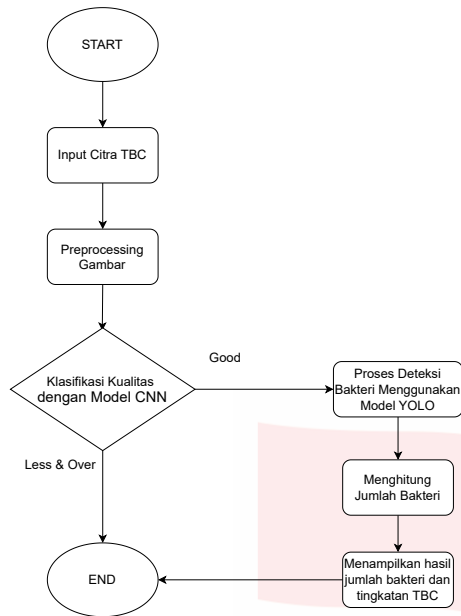
| Spesifikasi                       | Keterangan                                                                                                                                                                                                                                                                                          |
|-----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Klasifikasi kualitas pewarnaan ZN | Mengukur akurasi klasifikasi dan konsistensi model CNN terhadap variasi kualitas pewarnaan ZN ( <i>Good Staining</i> , <i>Less Staining</i> , <i>Over Staining</i> ) menggunakan dataset uji berlabel, dengan target semua metrik (akurasi, recall, specificity, precision, F1-skor) lebih dari 90% |
| Deteksi dan Perhitungan Bakteri   | Mengukur presisi <i>bounding box</i> , kemampuan menghitung jumlah bakteri TBC, dan ketahanan terhadap <i>false positive</i> hanya pada gambar dengan hasil klasifikasi <i>Good Staining</i> , menggunakan mAP dan metrik evaluasi lainnya dengan target lebih dari 80%                             |
| Validasi Logika Sistem            | Memastikan integrasi model CNN dan YOLO berjalan sesuai alur kerja, mulai dari Login, Input gambar, klasifikasi, deteksi, dengan logika yang benar, termasuk notifikasi untuk <i>Less Staining/Over Staining</i> dan eksekusi deteksi hanya pada <i>Good Staining</i>                               |
| Waktu Respon Sistem               | Sistem mampu memproses klasifikasi dan deteksi bakteri TBC dengan waktu respon total kurang dari 3 detik pada perangkat uji                                                                                                                                                                         |
| Kepuasan Pengguna                 | Berdasarkan kuisioner (skala 1–5), mengukur kemudahan penggunaan antarmuka dan keandalan sistem, dengan fokus pada konsistensi hasil dan kenyamanan pengguna                                                                                                                                        |



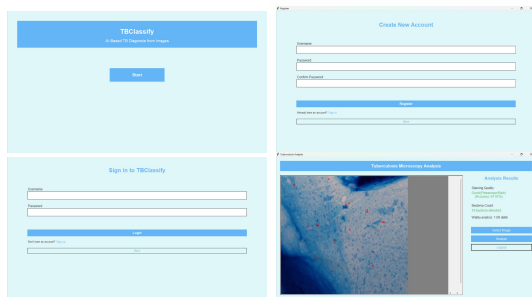
Gambar 1: Flowchart Registrasi Pengguna

1) *Flowchart*: Flowchart digunakan untuk memvisualisasikan proses kerja aplikasi standalone mulai dari penerimaan input citra hingga penyajian hasil analisis kepada pengguna. Flowchart dibagi menjadi dua, yaitu flowchart registrasi dan proses analisis. Gambar 1 adalah flowchart registrasi dan proses analisis dan Gambar 2 adalah Flowchart proses analisis input citra

2) *Desain Aplikasi*: Desain aplikasi berperan penting untuk memastikan bahwa seluruh komponen mulai dari input data, pemrosesan, hingga penyajian hasil analisis dapat



Gambar 2: Flowchart Proses Analisis Input Citra



Gambar 3: Desain Aplikasi

diintegrasikan dengan baik dan dioperasikan secara efisien. Dengan desain aplikasi yang matang, pengguna mendapatkan kenyamanan dan kemudahan untuk dapat menggunakan aplikasi dengan baik. Gambar 3 adalah desain dari aplikasi yang kami buat.

### C. Metode Pengukuran yang Sesuai

Setelah menentukan batasan dan spesifikasi, pengukuran solusi diperlukan untuk memastikan solusi memenuhi spesifikasi yang ditetapkan. Tujuan dilakukan pengukuran solusi adalah untuk menetapkan batasan yang diperlukan oleh sistem guna mencapai hasil yang optimal. Berikut merupakan rincian mengenai proses pengukuran dari setiap spesifikasi.

1) *Klasifikasi Kualitas Pewarnaan ZN*: Pengukuran ini bertujuan untuk memvalidasi kemampuan model CNN dalam mengklasifikasikan kualitas pewarnaan ZN pada sampel dahak ke dalam tiga kategori: *Good Staining*, *Less Staining*, dan *Over Staining*. Nilai reliabilitas pengujian modul ini diukur dari akurasi klasifikasi dan konsistensi model terhadap variasi kualitas pewarnaan ZN pada citra yang diuji.

Tabel III: Pengukuran Klasifikasi Kualitas Pewarnaan ZN

| Hal                  | Keterangan                                                                                                                                                                                                                                                                                              |
|----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Rincian yang Diukur  | Mengukur akurasi klasifikasi dan konsistensi model terhadap kualitas gambar [31]                                                                                                                                                                                                                        |
| Metode Pengukuran    | Evaluasi menggunakan dataset uji yang sudah dilabeli menjadi tiga kategori                                                                                                                                                                                                                              |
| Mekanisme Pengukuran | Menggunakan dataset uji yang tidak digunakan selama proses pelatihan dan menghitung metrik evaluasi mulai dari akurasi, recall, specificity, precision, F1-skor menggunakan <i>confusion matrix</i> [17], [22] dan semua metrik evaluasi mendapatkan nilai minimal 90% pada setiap kategori klasifikasi |

Tabel III merupakan mekanisme pengukuran klasifikasi kualitas pewarnaan ZN dengan *confusion matrix*. Dalam pengujian ini, model yang sudah dilatih diuji menggunakan dataset uji yang belum pernah digunakan selama proses pelatihan. Proses pengukuran menggunakan beberapa metrik evaluasi, yaitu akurasi, recall, specificity, precision, dan F1-skor. Agar sistem dikatakan berhasil, seluruh nilai metrik evaluasi pada setiap kategori klasifikasi mendapatkan nilai minimal 90%

2) *Deteksi dan Perhitungan Bakteri TBC* : Pengujian ini mengevaluasi performa model YOLO dalam mendeteksi dan menghitung bakteri TBC pada sampel yang telah di-klasifikasikan sebagai *Good Staining*. Nilai reliabilitas yang diuji meliputi *mAP*, *recall*, *specificity*, *precision*, dan F1-skor. Berikut merupakan tabel mekanisme pengukuran deteksi dan perhitungan bakteri TBC.

Tabel IV: Pengukuran Deteksi dan Perhitungan Bakteri TBC

| Hal                  | Keterangan                                                                                                                                                                                                                                                                                                   |
|----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Rincian yang Diukur  | Aspek yang diuji meliputi akurasi deteksi ( <i>precision bounding box</i> ) [21], kemampuan menghitung jumlah bakteri, serta ketahanan terhadap <i>false positive</i> (deteksi non-bakteri) [17].                                                                                                            |
| Metode Pengukuran    | Menggunakan dataset terpisah dengan anotasi <i>bounding box</i> bakteri TBC [31] dan hanya dijalankan ketika hasil klasifikasi memprediksi <i>Good Staining</i> .                                                                                                                                            |
| Mekanisme Pengukuran | Menggunakan dataset uji yang tidak digunakan selama proses pelatihan dan menghitung metrik evaluasi mulai dari <i>mAP</i> , recall, specificity, precision, F1-skor menggunakan <i>confusion matrix</i> [17], [22] dan semua metrik evaluasi mendapatkan nilai minimal 80% pada setiap kategori klasifikasi. |

Tabel IV menunjukkan mekanisme pengukuran klasifikasi kualitas pewarnaan ZN menggunakan *confusion matrix*. Dataset uji yang dianotasi dengan *bounding box* bakteri TBC (tidak digunakan saat pelatihan) dievaluasi menggunakan *mAP*, recall, specificity, precision, dan F1-score, dengan target minimal 80% per kelas. Pengujian juga mencakup ketahanan terhadap *false positive* (FP).

3) *Validasi Logika Sistem*: Pengujian validasi logika sistem dilakukan melalui simulasi alur kerja sistem secara menyeluruh, yang terdiri atas: proses sign in, input citra, klasifikasi kualitas pewarnaan ZN, deteksi dan perhitungan

bakteri TBC (jika hasil klasifikasi adalah *Good Staining*), dan analisis hasil akhir. Berikut merupakan tabel mekanisme pengukuran validasi logika sistem.

Tabel V: Pengukuran Validasi Logika Sistem

| Hal                  | Keterangan                                                                                                                                                                                                                             |
|----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Rincian yang Diukur  | Validasi logika sistem (notifikasi untuk hasil <i>Less/Over Staining</i> ) dan stabilitas integrasi antara komponen model CNN dengan YOLO dalam kondisi operasional yang nyata.                                                        |
| Metode Pengukuran    | Simulasi alur kerja sistem: <i>Sign in</i> → login → input gambar → klasifikasi → deteksi bakteri (Jika <i>Good Staining</i> ) → output analisis [31].                                                                                 |
| Mekanisme Pengukuran | Input citra <i>Less/Over Staining</i> , pastikan sistem menampilkan notifikasi tanpa menjalankan perhitungan bakteri [32]. Input citra <i>Good Staining</i> , pastikan sistem menjalankan deteksi dan menampilkan jumlah bakteri [17]. |

abel V menunjukkan validasi logika sistem melalui simulasi dengan dua skenario: (1) citra *Over Staining* memunculkan notifikasi tanpa deteksi bakteri TBC, dan (2) citra *Good Staining* menjalankan deteksi otomatis serta menampilkan jumlah bakteri TBC terdeteksi.

4) *Waktu Respon Sistem* : Waktu respon sistem dilakukan dengan menjalankan aplikasi pada 30 perangkat laptop atau pc yang berbeda untuk mengukur proses pada tahap klasifikasi hingga deteksi bakteri TBC. Berikut merupakan tabel mekanisme pengukuran waktu respon sistem.

Tabel VI: Pengukuran Waktu Respon Sistem

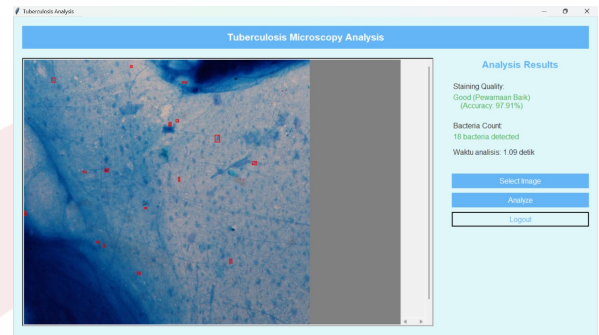
| Hal                  | Keterangan                                                                                                                |
|----------------------|---------------------------------------------------------------------------------------------------------------------------|
| Rincian yang Diukur  | Pengukuran waktu respon sistem untuk memastikan analisis berjalan dengan cepat.                                           |
| Metode Pengukuran    | Menjalankan aplikasi pada 30 perangkat laptop atau PC yang berbeda [23].                                                  |
| Mekanisme Pengukuran | Ketika menjalankan aplikasi, waktu inferensi per citra pada klasifikasi dan deteksi bakteri TBC kurang dari 3 detik [31]. |

Tabel VI merupakan mekanisme pengukuran waktu respon sistem dengan menjalankan aplikasi pada 30 laptop berbeda. Sistem dinyatakan memenuhi spesifikasi apabila total waktu inferensi per gambar (dari klasifikasi hingga deteksi) tidak melebihi 3 detik, sesuai target efisiensi minimum yang ditetapkan.

5) *Kepuasan Pengguna*: Pengujian sistem juga mencakup pengukuran kepuasan pengguna, yang diperoleh melalui penyebaran kuisioner dengan skala 1-5 yang merepresentasikan tingkat keandalan dan kenyamanan mereka menggunakan aplikasi. Berikut merupakan tabel mekanisme pengukuran kepuasan pengguna. Tabel VII menunjukkan hasil pengukuran kepuasan pengguna melalui kuisioner skala 1–5, mencakup keandalan sistem dan kemudahan antarmuka. Pengukuran ini bertujuan memastikan sistem berfungsi konsisten sekaligus memberikan pengalaman penggunaan (*user experience*) yang positif.

Tabel VII: Pengukuran Kepuasan Pengguna

| Hal                  | Keterangan                                                                                          |
|----------------------|-----------------------------------------------------------------------------------------------------|
| Rincian yang Diukur  | Uji kepuasan pengguna untuk menilai kemudahan dan keandalan sistem.                                 |
| Metode Pengukuran    | Pengisian kuisioner (Skala 1–5) terkait tingkat keandalan dan kenyamanan menggunakan aplikasi [25]. |
| Mekanisme Pengukuran | Kuisioner berisi pertanyaan terkait keandalan sistem [33] dan kemudahan penggunaan antarmuka [20].  |



Gambar 4: Fitur Utama TBClassify

## IV. HASIL DAN PEMBAHASAN

### A. Implementasi TBClassify

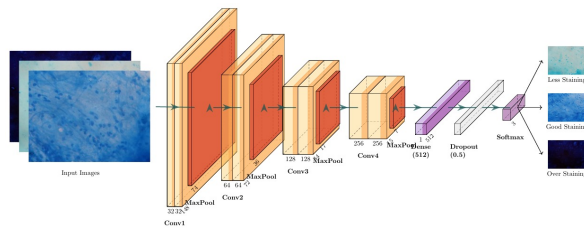
TBClassify merupakan aplikasi yang dirancang untuk mendukung proses diagnosis penyakit TBC melalui analisis citra bakteri. Dalam aplikasi TBClassify terdiri dari beberapa komponen utama:

- **Antarmuka Pengguna (GUI)**: Dibangun dengan menggunakan Tkinter dengan berbagai elemen yaitu Frame, Label, Button, Entry, dan Canvas
- **Database lokal**: Data pengguna disimpan dalam file JSON (*users.json*) dan data diproses langsung dari file lokal
- **Model ML**: Menggunakan 2 model yaitu model CNN yang digunakan untuk menentukan kualitas pewarnaan citra bakteri, dan model YOLO untuk mendeteksi bakteri pada citra bakteri.

Fitur utama TBClassify adalah perhitungan dan deteksi bakteri pada citra mikroskopis (Gambar 4). Proses diawali klasifikasi kualitas pewarnaan dengan model CNN; hanya gambar berlabel “Good” yang dilanjutkan ke tahap perhitungan bakteri, menampilkan jumlah dalam rentang serta posisi melalui *bounding box* merah.

### B. Implementasi Model CNN untuk Klasifikasi Kualitas Pewarnaan ZN

Model klasifikasi kualitas pewarnaan ZN merupakan hasil optimasi CNN sebelumnya, menggunakan *balancing dataset* dan teknik *callback* untuk meningkatkan akurasi pelatihan. Dataset tiga kelas pewarnaan terdiri dari *LessStaining* (40 sampel), *GoodStaining* (76), dan *OverStaining* (18). Ketidakseimbangan ini diatasi melalui *visual data augmentation*



Gambar 5: Arsitektur Model CNN

dan SMOTE, sehingga tiap kelas menjadi seimbang (70 sampel) dan model tidak bias terhadap kelas mayoritas.

Tabel VIII: Proses Optimasi Model CNN

| Tahapan                | Teknik Optimasi                                                                                                     | Catatan                       |
|------------------------|---------------------------------------------------------------------------------------------------------------------|-------------------------------|
| Model Awal             | Dataset Asli Tanpa regulasi overfitting                                                                             | Akurasi yang didapat 88%      |
| Model Optimasi Pertama | Menyeimbangkan dataset dengan teknik augmentasi<br>Menambahkan teknik <i>EarlyStopping</i> untuk target akurasi 95% | Akurasi meningkat menjadi 95% |
| Model Optimasi Kedua   | Target akurasi dinaikkan menjadi 97%                                                                                | Model stabil, Akurasi 100%    |

Tabel VIII menunjukkan optimasi CNN melalui tiga tahap: model awal, optimasi pertama, dan kedua. Fokus optimasi mencakup penyeimbangan dataset, penerapan regulasi seperti *EarlyStopping* untuk mencegah overfitting, serta peningkatan akurasi target di atas 97

Gambar 5 memperlihatkan rancangan model klasifikasi dengan *max pooling*, kernel 3×3, dan jumlah filter 32, 64, 128, serta 256 [28]. *Pooling layer* mengurangi ukuran spasial, sedangkan *max pooling* memilih nilai piksel terbesar. *Dense layer* memiliki 512 neuron (ReLU), *output layer* 3 neuron, dan optimizer Adam [28]. Dataset dibagi menjadi 70

### C. Implementasi Model YOLO untuk Deteksi dan Perhitungan Bakteri TBC

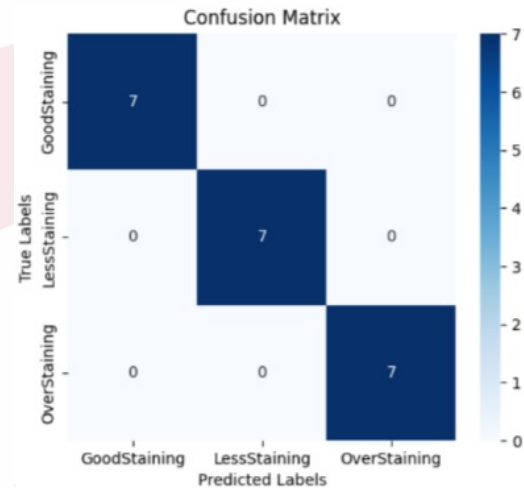
Pembuatan model deteksi bakteri meliputi pengumpulan data, *preprocessing*, *augmentation*, pelatihan, dan evaluasi. Tujuannya mengembangkan sistem deteksi objek andal untuk membedakan citra dengan atau tanpa bakteri. Pelatihan dilakukan menggunakan enam pendekatan (Tabel IX).

Tabel IX menjelaskan enam model: Model 1 dan 2 menggunakan *hyperparameter* manual dengan 50 dan 100 epoch. Model 3 dan 4 adalah *fine tuning* berurutan dari model sebelumnya dengan *hyperparameter* manual dan 100 epoch. Model 5 memakai *random search* (50 epoch), sedangkan Model 6 menggunakan *hyperparameter* terbaik dari Model 5 dengan 100 epoch dan *early stopping*.

Dataset terdiri dari 576 citra (238 positif, 238 negatif) sesuai standar IUATLD. Data dibagi menjadi 70% pelatihan, 20% validasi, dan 10% pengujian. *Augmentation* melalui Roboflow meningkatkan data pelatihan menjadi 974 citra sekaligus mempermudah proses ekspor dataset.

Tabel IX: Teknik yang Digunakan oleh Model YOLO

| Model | Teknik yang Digunakan                                                                             |
|-------|---------------------------------------------------------------------------------------------------|
| 1     | Pengaturan <i>hyperparameter</i> manual dengan 50 epoch                                           |
| 2     | Pengaturan <i>hyperparameter</i> manual dengan 100 epoch                                          |
| 3     | <i>Fine tuning</i> Model 2 dengan pengaturan <i>hyperparameter</i> manual dan 100 epoch           |
| 4     | <i>Fine tuning</i> Model 3 dengan pengaturan <i>hyperparameter</i> manual dan 100 epoch           |
| 5     | <i>Random search</i> dengan 50 epoch untuk setiap iterasi                                         |
| 6     | Menggunakan <i>hyperparameter</i> terbaik dari Model 5 dengan 100 epoch dan <i>early stopping</i> |



Gambar 6: Confusion Matrix menggunakan Data Uji

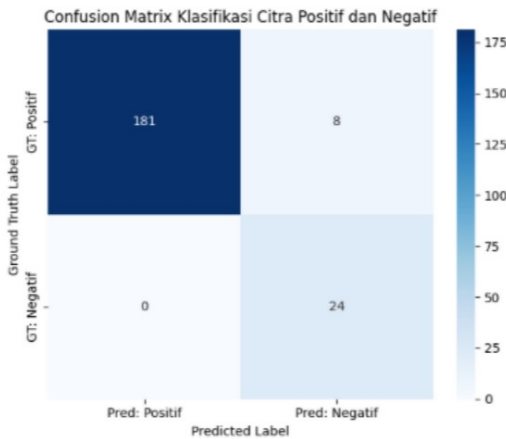
#### D. Hasil Pengujian

1) *Pengujian Model CNN* : Setelah optimasi tahap kedua, model dievaluasi menggunakan 10% data uji terpisah. Kinerja diukur melalui *confusion matrix* untuk menghitung akurasi, presisi, recall, spesifisitas, dan F1-score, memastikan prediksi seimbang antara positif dan negatif. Tahap selanjutnya adalah pengujian skenario deteksi dan perhitungan jumlah bakteri TBC secara otomatis untuk memastikan integrasi model pada aplikasi berjalan optimal. Gambar 6 menunjukkan *confusion matrix* dengan seluruh sampel tiap kelas terklasifikasi benar (nilai diagonal sempurna, 7 per kelas) tanpa kesalahan prediksi. Model mencapai akurasi 100% pada data uji untuk semua kelas: *GoodStaining*, *LessStaining*, dan *OverStaining*.

Gambar 7 menunjukkan metrik evaluasi (precision, recall/sensitivity, specificity, dan F1-skor) untuk setiap kelas. Pada *LessStaining*, *OverStaining*, dan *GoodStaining*, seluruh metrik mencapai 1.0 (100%), menandakan model mengidentifikasi semua kelas dengan presisi dan sensitivitas

| Class: LessStaining          | Class: OverStaining          | Class: GoodStaining          |
|------------------------------|------------------------------|------------------------------|
| Precision: 1.0000            | Precision: 1.0000            | Precision: 1.0000            |
| Recall (Sensitivity): 1.0000 | Recall (Sensitivity): 1.0000 | Recall (Sensitivity): 1.0000 |
| Specificity: 1.0000          | Specificity: 1.0000          | Specificity: 1.0000          |
| F1-Score: 1.0000             | F1-Score: 1.0000             | F1-Score: 1.0000             |

Gambar 7: Hasil Metrik Evaluasi



Gambar 8: Confusion matrix membedakan citra negatif dan positif pada Pengujian Menggunakan Data Uji

sempurna, tanpa false positive maupun false negative.

2) *Pengujian Model YOLO*: Semua model dievaluasi berdasarkan loss, recall, akurasi, F1-skor, mAP50, serta perbandingan kurva loss pada data train dan validasi. Model terbaik kemudian diuji untuk klasifikasi citra positif dan negatif (confusion matrix) serta kemampuan mendeteksi dan menghitung jumlah AFB per citra.

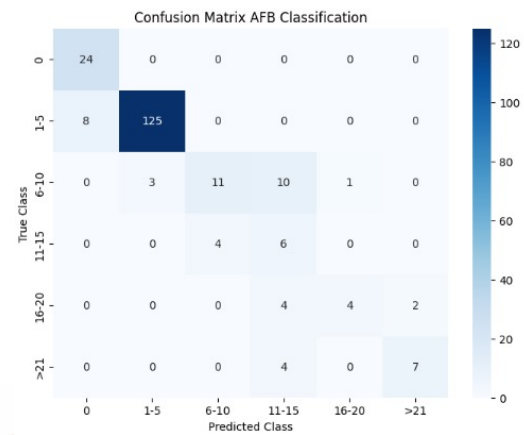
Berdasarkan Tabel X, model pertama dipilih sebagai yang terbaik dan diuji pada data testing (10% dataset) dengan 6 kategori jumlah AFB: 0, 1–5, 6–10, 11–15, 16–20, dan > 21, masing-masing berjumlah 24, 133, 25, 10, 10, dan 11 citra. Confusion Matrix berikut menunjukkan kemampuan model membedakan citra positif dan negatif.

Confusion matrik yang ada di Gambar 8 memperjelas seberapa baik model mampu mengklasifikasikan data secara akurat serta tidak ada bias terhadap kelas tertentu. Pada hasil analisis gambar diatas, model mampu memprediksi dengan benar sebanyak 181 citra positif dan 24 citra negatif. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang sempurna untuk mengklasifikasikan data yang diberikan.

Berdasarkan Tabel, model mencapai akurasi 96,2%, presisi 100%, sensitivitas 95,7%, spesifisitas 100%, dan F1-skor 97,8%. Model mendeteksi 181 citra positif dan 24 citra negatif, dengan 8 citra positif salah diklasifikasikan sebagai negatif. Pada tahap evaluasi perhitungan jumlah bakteri, akurasi prediksi model terhadap jumlah sebenarnya melebihi 80

Berdasarkan Tabel dan Gambar, model sangat andal mendeteksi citra dengan sedikit AFB maupun citra negatif. Namun, performa menurun pada kelas 6–10 dan 11–15 karena banyak AFB tidak terdeteksi atau salah klasifikasi. Pada kelas 16–20 dan >21, kinerja sedang dengan kecenderungan jarang mengklasifikasikan citra sebagai memiliki AFB dalam jumlah besar, sehingga sensitivitas pada kelas ini relatif rendah.

3) *Pengujian User Acceptance Testing (UAT)* : Pengujian aplikasi dilakukan menggunakan User Acceptance Testing



Gambar 9: Confusion Matrix dari Perhitungan Bakteri menggunakan Data Uji

(UAT) jenis Open Beta Testing dengan 30 responden yang mencoba aplikasi di perangkat masing-masing. Survei menggunakan skala Likert 1–5 (1 = sangat tidak nyaman, 5 = sangat nyaman) seperti pada tabel, dengan skor ideal dihitung menggunakan formula:

$$\text{Skor Ideal} = (\text{Jumlah Kategori Skala Likert}) \times (\text{Jumlah Responden}) \quad (6)$$

Untuk menghitung nilai persentase (P) yang didapat pada setiap jawaban responden adalah dengan formula berikut:

$$P = \frac{\text{Total Nilai}}{\text{Skor Ideal}} \times 100\% \quad (7)$$

Setelah seluruh data dianalisis, aplikasi TBClassify memperoleh skor sebesar 135 dari 150 atau 90%, yang termasuk dalam kategori penerimaan tinggi, menunjukkan tingkat kenyamanan dan usability yang baik serta dapat diterima oleh pengguna.

## V. KESIMPULAN

Klasifikasi kualitas pewarnaan ZN menggunakan model CNN mencapai performa sempurna pada seluruh label (*recall*, *specificity*, dan F1-skor = 1.0) tanpa kesalahan prediksi, meskipun validasi eksternal tetap diperlukan untuk menghindari bias distribusi data. Deteksi bakteri TBC dengan YOLO menunjukkan akurasi tinggi (96,2%; *precision* 100%; *recall* 95,7%; *specificity* 100%), namun menurun pada kategori menengah (11–15 dan 16–20 bakteri) akibat tantangan deteksi objek kecil dan tumpang tindih, yang dapat diatasi dengan augmentasi spasial, peningkatan resolusi, atau adaptasi arsitektur. Dari sisi pengujian fitur, *TBClassify* memperoleh nilai UAT 90,3%, dengan fitur utama berjalan baik di Windows namun belum mendukung MacOS dan Ubuntu, serta performa dipengaruhi kapasitas RAM.

## PUSTAKA

- [1] Kementerian Kesehatan Republik Indonesia, "Data Kondisi TBC," <https://www.tbindonesia.or.id/pustaka/tbc/data-kondisi-tbc/>, 2025, diakses pada 4 Agustus 2025.

Tabel X: Informasi dari Pelatihan Terakhir dari Masing-masing Model

| Model | Nama     | Precision | Recall  | mAP50   | mAP50-95 | Epoch   |
|-------|----------|-----------|---------|---------|----------|---------|
| 1     | Best.pt  | 0.67137   | 0.55492 | 0.59639 | 0.23822  | 50/50   |
| 2     | Best3.pt | 0.62248   | 0.58086 | 0.59501 | 0.23083  | 100/100 |
| 3     | Best4.pt | 0.67915   | 0.55114 | 0.59181 | 0.22468  | 100/100 |
| 4     | Best5.pt | 0.62895   | 0.58282 | 0.63720 | 0.23024  | 100/100 |
| 5     | Best6.pt | 0.54006   | 0.51817 | 0.49432 | 0.17496  | 50/50   |
| 6     | Best7.pt | 0.61482   | 0.53206 | 0.53497 | 0.18410  | 66/100  |

Tabel XI: Perhitungan Setiap Metrik Klasifikasi Citra Positif dan Negatif

| Model   | Akurasi | Presisi | Sensitivitas | Spesifisitas | F1-skor | TP  | FP | FN | TN |
|---------|---------|---------|--------------|--------------|---------|-----|----|----|----|
| Pertama | 0.92    | 1.00    | 0.95         | 1.00         | 0.97    | 181 | 8  | 0  | 24 |

Tabel XII: Perhitungan Mertrik Deteksi Bakteri TBC

| Label | Akurasi | Presisi | Sensitivitas | Spesifisitas | F1-Skor | TP  | FP | FN | TN  |
|-------|---------|---------|--------------|--------------|---------|-----|----|----|-----|
| 0     | 0.962   | 0.750   | 1.000        | 0.957        | 0.857   | 24  | 8  | 0  | 181 |
| 1–5   | 0.948   | 0.976   | 0.940        | 0.963        | 0.958   | 125 | 3  | 8  | 77  |
| 6–10  | 0.897   | 0.579   | 0.440        | 0.957        | 0.500   | 11  | 8  | 14 | 170 |
| 11–15 | 0.897   | 0.250   | 0.600        | 0.911        | 0.353   | 6   | 18 | 4  | 180 |
| 16–20 | 0.967   | 0.800   | 0.400        | 0.995        | 0.533   | 4   | 1  | 6  | 202 |
| >21   | 0.972   | 0.778   | 0.636        | 0.990        | 0.700   | 7   | 2  | 4  | 200 |

- [2] World Health Organization, *Global Tuberculosis Report 2023*. Geneva: World Health Organization, 2023.
- [3] Kementerian Kesehatan Republik Indonesia, *Petunjuk Teknis Penatalaksanaan Tuberkulosis Resistan Obat di Indonesia*. Jakarta: Kementerian Kesehatan RI, 2020.
- [4] J. G. Jang and J. H. Chung, "Diagnosis and treatment of multidrug-resistant tuberculosis," *Yeungnam University Journal of Medicine*, vol. 37, no. 4, pp. 277–285, Oct. 2020.
- [5] Kementerian Kesehatan Republik Indonesia, "Strategi nasional penanggulangan tuberkulosis di indonesia 2020–2024," <https://yankes.kemkes.go.id/>, 2020, diakses pada 4 Agustus 2025.
- [6] World Health Organization, *WHO Consolidated Guidelines on Tuberculosis. Module 2: Screening – Systematic Screening for Tuberculosis Disease*. Geneva: World Health Organization, 2021.
- [7] A. L. Garcia-Basteiro *et al.*, "Point of care diagnostics for tuberculosis," *Pulmonology*, Jan. 2018.
- [8] K. R. Steingart *et al.*, "Fluorescence versus conventional sputum smear microscopy for tuberculosis: A systematic review," *The Lancet Infectious Diseases*, 2006.
- [9] T. Hänscheid, "The future looks bright: Low-cost fluorescent microscopes for detection of *Mycobacterium tuberculosis* and coccidia," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 102, no. 6, pp. 520–521, Jun. 2008.
- [10] R. Yudianta, Z. Zulmansyah, and H. Garna, "Hubungan kepatuhan terapi obat anti tuberkulosis kombinasi dosis tetap (oat-kdt) dengan kesembuhan pasien tuberkulosis paru dewasa di puskesmas patokeb-usi subang," *Jurnal Integrasi Kesehatan & Sains*, vol. 4, no. 1, pp. 44–49, Jan. 2022.
- [11] S. U. Basundari, S. Harun, R. Ekowatiningsih, and E. Yuwarni, "Uji validitas teknik pcr (polymerase chain reaction) dan pemeriksaan mikroskopis bakteri tahan asam sebagai alat diagnosis penderita tb paru di rumah sakit persahabatan, jakarta," *Media Penelitian dan Pengembangan Kesehatan*, vol. 12, no. 3, 2002.
- [12] A. Tjokronegoro and A. Yusuf, *Tuberkulosis paru: pedoman penataan diagnostik dan terapi*. Fakultas Kedokteran Universitas Indonesia, Jakarta, 1985.
- [13] M. Misnarliah and M. Mudrika, "Pengaruh penundaan pewarnaan preparat bakteri tahan asam metode ziehl neelsen terhadap hasil pemeriksaan mikroskopik," *Jurnal Teknosains Kodepena*, vol. 1, no. 2, pp. 58–63, 2021.
- [14] H. R. da Silva and Y. Maulani, "Perbandingan hasil pemeriksaan mikroskopis sputum bta terhadap metode pcr (genexpert) pada pasien tuberkulosis paru," *Plenary Health: Jurnal Kesehatan Paripurna*, vol. 1, no. 3, pp. 455–460, 2024.
- [15] A. Achmadi, M. M. Mardiah, and S. Wahyu, "Penerapan pemantapan mutu internal terhadap kualitas sediaan pewarnaan ziehl nielsen untuk deteksi mycobacterium tb," *Jurnal Ilmiah Kesehatan (JIKA)*, vol. 3, no. 3, pp. 124–133, 2021.
- [16] IBM Team. (2025) Apa itu convolutional neural network? [Online]. Available: <https://www.ibm.com/id-id/think/topics/convolutional-neural-networks>
- [17] IONOS Editorial Team. (2025) Convolutional neural networks. IONOS. [Online]. Available: <https://www.ionos.com/digitalguide/websites/web-development/convolutional-neural-networks/>
- [18] GeeksforGeeks Editorial Team. (2025) Understanding the confusion matrix in machine learning. GeeksforGeeks. [Online]. Available: <https://www.geeksforgeeks.org/machine-learning/confusion-matrix-machine-learning/>
- [19] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis and machine vision*. Springer, 2013.
- [20] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 11, pp. 3212–3232, 2019.
- [21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.
- [22] J. Du, "Understanding of object detection based on cnn family and yolo," in *Journal of Physics: Conference Series*, vol. 1004. IOP Publishing, 2018, p. 012029.
- [23] R. Khanam and M. Hussain, "Yolov11: An overview of the key architectural enhancements. arxiv 2024," *arXiv preprint arXiv:2410.17725*, 2024.
- [24] R. Khanam, M. Hussain, R. Hill, and P. Allen, "A comprehensive review of convolutional neural networks for defect detection in industrial applications," *IEEE Access*, 2024.
- [25] Annisadev. (2024, November) Mengenal roboflow: Solusi lengkap untuk pengembangan model computer vision. Annisadev. [Online]. Available: <https://www.anisadev.com/news/read/1012/mengenal-roboflow-solusi-lengkap-untuk-pengembangan-model-computer-vision.html>
- [26] DQLab. (2024) Hyperparameter tuning dalam proses pemodelan data. DQLab. [Online]. Available: <https://www.dqlab.id/>
- [27] D. Kurniadi, R. M. Shidiq, and A. Mulyani, "Perbandingan penggunaan optimizer dalam klasifikasi sel darah putih menggunakan convolutional neural network," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 14, no. 1, pp. 77–86, 2025.

- [28] P. S. Foundation. (2024) Apa itu python? Python.org. [Online]. Available: <https://www.python.org/doc/essays/blurb/>
- [29] J. Salendah, P. Kalele, A. Tulenan, and J. S. R. Joshua, "Penentuan beasiswa dengan metode fuzzy tsukamoto berbasis web scholarship determination using web based fuzzy tsukamoto method," in *Seminar Nasional Ilmu Komputer (SNASIKOM)*, vol. 2, no. 1, 2022, pp. 81–90.
- [30] R. Ramadhan, E. Fitria, and R. Rosdiana, "Deteksi mycobacterium tuberculosis dengan pemeriksaan mikroskopis dan teknik pcr pada penderita tuberculosis paru di puskesmas darul imarah," *Sel Jurnal Penelitian Kesehatan*, vol. 4, no. 2, pp. 73–80, 2017.
- [31] S. Aulia, A. B. Suksmono, T. R. Mengko, and B. Alisjahbana, "A novel digitized microscopic images of zn-stained sputum smear and its classification based on iuatld grades," *IEEE Access*, vol. 12, pp. 51 364–51 380, 2024.
- [32] Trivusi. (2022, August) Apa bedanya epoch dan batch size pada deep learning? Trivusi. [Online]. Available: <https://www.trivusi.web.id/2022/08/epoch-dan-batch-size.html>
- [33] A. Sejati and L. Sofiana. (2015) Jurnal kesehatan masyarakat. [Online]. Available: <http://journal.unnes.ac.id/nju/index.php/kemas>

