

APLIKASI PEMBUATAN LABEL GRAF BERBASIS MATRIX KETETANGGAAN PADA POTONGAN CITRA SPUTUM DAHAK

Mirza Tsabita W.- 13524114

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jalan Ganesha 10 Bandung

E-mail: mirzatsabita4@gmail.com , 13524114@std.stei.itb.ac.id

Abstract—Deep learning berbasis Graf NeuroNetwork, merupakan metode terbaru berbasis graf yang memungkinkan penggunaan AI model terbaru, namun menggunakan komputasi yang lebih ringan dan lebih cepat. Namun, aplikasi pembuatan label citra yang berbentuk graf masih sangat sedikit, berbeda dengan label citra yang tidak berbentuk graf seperti labelme dan SAM (Segment Anything Model). Potensi pembuatan label graf terutama yang unsupervised masih sangat besar. Kontribusi yang diberikan adalah (1) Membuat dataset patch baru khusus untuk pembuatan label graf, (2) Menghasilkan dataset label graf yang berikutnya berguna pada metode GNN. Model graf berbasis matrix adjacency. Proses pembuatannya dimulai dari melakukan transformasi dari citra warna ke citra biner, lalu dilanjutkan dengan pembuatan graf menggunakan matrix adjacency. Untuk studi perbandingan dipaparkan menggunakan graf RGB, graf OR, graf AND, graf XOR dan graf XNOR. Hasil perbandingan dianalisis menggunakan matrix conclusion, dari perbandingan tersebut didapatkan hasil bahwa metode graf AND adalah metode dengan tingkat akurasi terbaik yaitu 0,9772. Artinya, matrix adjacency dengan metode graf AND mampu mengenali pola bakteri dengan tepat.

Keywords—image process, label graph, matrix adjacency

I. PENDAHULUAN

1.2 Latar Belakang

Indonesia dengan jumlah penduduk terbesar ke-2 didunia (setelah Cina - 285 juta jiwa), ternyata menempati posisi yang sama untuk jumlah prevalensi TBC (setelah India), menurut WHO Tahun 2023. Hingga tahun 2024, jumlah prevalensi TBC mencapai 1:260 jiwa, artinya dari setiap 1.260 penduduk terdapat 1 penduduk yang terkena TBC. Screening TBC di Indonesia dilakukan menggunakan metode BTA (Bakteri Tahan Asam). BTA diawali dengan pengambilan dahak dari pasien, kemudian dilakukan apusan sputum dahak pada slide mikroskop yang berikutnya diwarnai dengan pewarnaan Ziehl-Neelsen (ZN)[1], [2], [3]. Hasil dari perwarnaan tersebut, bakteri akan berwarna merah dan dahak akan berwarna biru hingga tidak berwarna. Penentuan level TB dilakukan dengan menghitung jumlah bakteri dalam 300 layang pandang menurut UATLD[4], yang dihitung menggunakan arah zigzag[5], [6], [7], [8]. Level TB terdiri dari 5, yaitu Negatif, Scanty, +1, +2 dan +3.

Table 1. Level Bakteri Menurut UATLD

Level TB	Jumlah Bakteri	Jumlah layang pandang
Negatif	0 AFB/300FoV	300
Scanty	1-29 AFB/300FoV	300
+1	>30 AFB/300FoV	300
+2	1-10 AFB/FoV	50
+3	>10 AFB/FoV	20

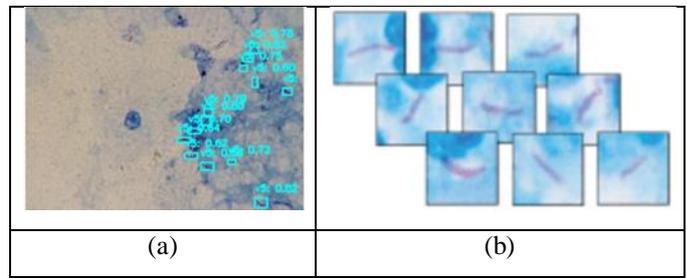
Permasalahan dilapangan adalah, masih banyak laboran yang melakukan human error dalam melakukan perhitungan BTA, hal ini karena banyak bakteri yang berkumpul dan bertumpuk, ada bakteri yang tertutup dengan dahak, dan karena dalam diri manusia sendiri, seperti aspek kelelahan mata, banyak pikiran ataupun karena kejenuhan. Manusia seharusnya hanya mampu menghitung dengan benar 4 slide dalam 1 hari. Namun karena banyaknya pasien, laboran bisa menghitung slide lebih dari 5 slide dalam sehari [9], [10]. Masalah yang lainnya juga karena laboran mempunyai kemampuan pewarnaan dan membuat apusan, sehingga banyak terjadi yang pewarnaan tebal ataupun apusan tebal. Apa yang terjadi jika terjadi kesalahan pada perhitungan, padahal jarak tiap level hanya terpaut 1 bakteri. Jika terjadi kesalahan hitung, maka akan terjadi juga kesalahan penentuan level TB. Artinya obat yang diberikan juga salah. Jika seharusnya level-nya adalah +2 namun diberikan obat +1, maka TB-nya tidak akan sembuh walaupun sudah berobat dalam jangka waktu lama. Namun sebaliknya jika seharusnya level-nya adalah +1 namun diberikan obat +2, maka akan terjadi TB RO (Resistan Obat), artinya adalah pasien ini tidak akan manjur lagi diberikan dosis obat yang lebih tinggi, sehingga pasien TB dengan RO akan menjalani pengobatan yang lebih berat dan lebih lama, serta ada kemungkinan berujung pada kematian. Salah satu metode terbaru yang digunakan Computer-Aided Diagnosis (CAD)[11], [12], [13]. Metode ini menggunakan kamera yang dikoneksikan dengan mikroskop. Citra mikroskopik dahak dari setiap layang pandang yang dihasilkan kemudian dimasukan kedalam pengolahan citra untuk kemudian dilakukan deteksi bakteri dan kemudian dilakukan perhitungan jumlah bakteri. Beberapa metode yang digunakan untuk deteksi bakteri diantaranya adalah menggunakan metode supervise [14], [15], unsupervised [16]

dan semi supervised [17], [18]. Pengolahan citra dilakukan dengan menggunakan kecerdasan buatan.

Kecerdasan buatan memungkinkan manusia menerapkan Machine Learning (ML), untuk mempelajari tingkah laku sistem. ML yang terbaru dan yang paling banyak diteliti pada penelitian terdahulu adalah penggunaan Deep Learning (DL). Berbeda dengan pembelajaran mesin konvensional yang menggunakan tahapan pembelajaran dengan ciri buatan tangan atau ciri geometris dari objek yang akan diteliti. DL menerapkan pembelajaran supervise dengan salah satu cirinya adalah menggunakan label sebagai penanda awal atau ground truth dari sistem. Hingga saat ini, metode DL masih menjadi metode pencetak sistem deteksi bakteri dengan Tingkat akurasi terbaik. Namun, dibalik tingginya tingkat akurasi, terdapat beberapa kelemahan dari DL, yaitu :

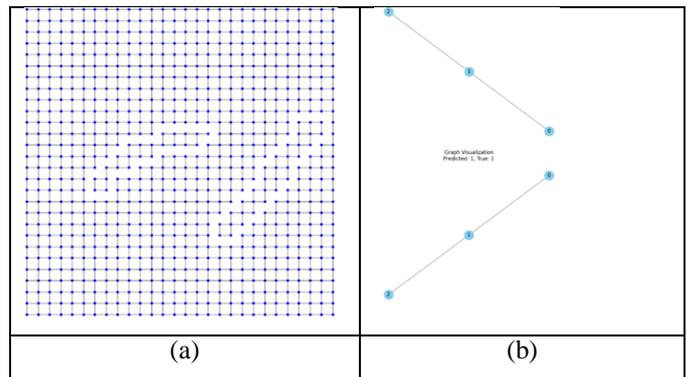
1. Pelabelan yang banyak memakan waktu. Label sekarang banyak dimudahkan dengan menggunakan tools yang sudah ada, seperti labelMe, dengan menggunakan polygon. Ketika hanya 1 bakteri masih bisa presisi, namun ketika banyak sekali bakteri, maka proses polygon akan tidak seragam. Mengapa polygon? Sebenarnya bisa juga menggunakan kotak (bondingbox) namun penggunaan kotak tidak disarankan untuk bakteri bertumpuk atau yang berkelompok.
2. Pada DL, semakin banyak citra pembelajaran dan semakin bervariasi, maka akurasi akan semakin tinggi. Artinya, citra yang digunakan untuk dataset harus besar, hal ini sebanding juga dengan jumlah label-nya. Artinya DL membutuhkan dataset besar (membutuhkan memori besar), dengan komputasi yang kompleks (membutuhkan memori proses yang besar). Hal ini juga berimbas pada waktu proses yang lama.

Dibutuhkan sebuah metode yang dapat digunakan sebagai label, yang dapat digunakan untuk proses DL namun dengan memori yang lebih kecil sehingga efisien dalam proses deteksi. Beberapa peneliti terdahulu menggunakan screening untuk melakukan deteksi bakteri dan dengan menggunakan teknik bounding box untuk citra pembelajaran. Namun deeplearning yang seperti ini, membutuhkan infrastruktur yang kompleks, termasuk penyediaan computer dengan komputasi tinggi . Ada cara lain dilakukan peneliti terdahulu adalah dengan melakukan pemrosesan berbasis patch (potongan kecil dari citra sputum dahak dalam 1 layang pandang). Citra sputum dahak dalam 1 layang pandang dapat dipotong-potong menjadi beberapa bagian, ada yang menggunakan ukuran 28x28 [9], [19], 32x32 [20] dan 64x64[21], baru kemudian dilakukan proses deteksi, jika ada citra dengan ukuran besar, maka akan dinormalisasi terlebih dahulu untuk mendapatkan kelipatan dari 28 atau 32 atau 64, kemudian dilakukan penambahan pixel "0" atau "255" pada bingkai citra lalu baru dilakukan proses pemotongan citra menjadi sejumlah patch [22]. Patch-patch ini kemudian satu persatu masuk ke bagian deteksi. Penggunaan patch kecil menurut peneliti lebih baik dibandingkan dengan citra utuh, karena variasi warna antara citra utuh dengan citra patch lebih kompleks citra utuh. Variasi warna sangat mempengaruhi hasil deteksi.



Gambar 1. Proses Deteksi (a) BoundingBox, (b) Patch

Label pada pengolahan citra, ada yang berbentuk citra biner (Semantic Segmentation), atau citra warna yang sudah tersegment menjadi beberapa warna (Instance Segmentation), dan ada juga yang menggunakan graf. Graf merupakan pilihan yang tepat untuk memperkecil kebutuhan memori untuk pelabelan[23], [24]. Transformasi dari citra ke graf terdiri dari 2, yaitu transformasi graf yang lossy (dari graf tidak bisa dikembalikan secara utuh menjadi citra) dan transformasi graf yang lossless (dari graf bisa dikembalikan secara utuh menjadi citra kembali). Beberapa peneliti menggunakan transformasi graf lossless untuk memudahkan dalam proses pengujiannya.



Gambar 2. Bentuk Graf untuk (a) Lossless, (b) Lossy

Graf digunakan sebagai label dalam GNN. (Graf Neuro Network)[25]. GNN merupakan metode deeplearning baru yg digunakan untuk menyederhanakan komputasi, memangkas waktu proses dan memperkecil kebutuhan memori. Artinya, jika bekerja menggunakan GNN[26], tidak membutuhkan komputer dengan spesifikasi komputasi tinggi, cukup dengan komputer standart dapat dilakukan proses deteksi bakteri, dengan kecepatan proses deteksi tetap tinggi, dan dengan akurasi yg tetap tinggi pula. Target penggunaannya adalah laboran, sehingga ketika projek ini diterapkan oleh pemerintah, tidak perlu menggunakan dana yg besar dalam pelaksanaannya. Kontribusi riset pada makalah ini adalah :

1. Penelitian ini menggunakan konsep baru dalam menghasilkan label graf untuk kebutuhan autolabeling pada gnn
2. Penelitian ini mengusung dataset patch baru yg digunakan sebagai dsta input dari sistem, didapatkan dari pemotongan citra pada dataset publik.

Tabel 1. Dataset Publik

Nama Dataset	Jumlah Citra	Link Url	Jumlah Sitasi Paper
NLM/NIH TB Dataset (USA)	31484	https://iee-dataport.org/documents/tbimages-smear-microscopy-image-dataset-support-development-automated-bacilli-detection	148
MIDTI (Microscopic Imaging Dataset for Tuberculosis Indonesia) dari ITB	371	https://universe.robotflow.com/suci-aulia/afb-test-y7	14
ODS1 Dataset (India)	90.200	https://plus.figshare.com/articles/dataset/_b_ODS1_-_Dataset_of_sputum_smear_microscopy_images_for_Tuberculosis_b_/27207987	9
Wellgen TB Smear Dataset (Asia)	75.087	https://docs.ngsci.org/datasets/tb-wellgen-smear/	8
DDS1 Patch Dataset (India)	5.272	https://figshare.com/articles/dataset/DDS1_-_Patch_Images_for_Bacilli_Detection/27209667	5
Kaggle Tuberculosis Image Dataset	2530	https://www.kaggle.com/datasets/saife245/tuberculosis-image-datasets	21
Hugging Face Tuberculosis Dataset	136	https://huggingface.co/datasets/moukaii/Tuberculosis_Dataset	2

1.3 Permasalahan

Permasalahan dari penelitian ini adalah bagaimana membuat label graf pada potongan citra sputum dahak menggunakan matriks ketetanggan (Adjacency) ?.

1.4 Tujuan

Tujuan dari penelitian ini adalah bagaimana membuat label graf pada potongan citra sputum dahak menggunakan matriks ketetanggan (Adjacency).

1.5 Manfaat

Manfaat dari penelitian ini adalah :

1. Menghasilkan label graf yang dapat digunakan untuk proses GNN.
2. Meningkatkan efisiensi sistem deteksi.

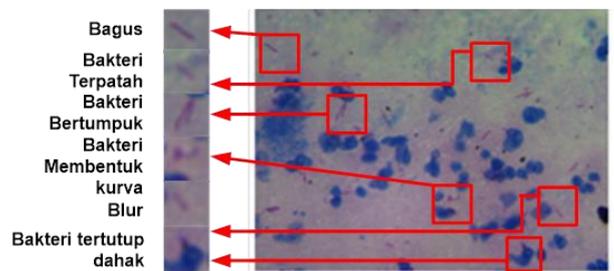
3. Mempercepat proses deteksi TB sehingga proses screening dapat meningkatkan akses Kesehatan.
4. Mengurangi akibat yang ditimbulkan oleh human error ketika melakukan proses deteksi.

II. TELAAH PUSTAKA

2.1 Dataset

Pada penelitian ini, kami mengumpulkan beberapa dataset yang dapat kami gunakan sebagai bahan citra input. Dataset yang digunakan pada penelitian ini adalah dataset publik yang diambil dari 1 data citra pada dataset dari beberapa dataset publik. Beberapa kumpulan dataset citra preparat dahak diperlihatkan pada Tabel 1.

Terdapat beberapa masalah yang sering dikeluhkan oleh peneliti terdahulu, menyangkut tentang turunnya nilai akurasi, karena beberapa citra bakteri yang tidak sempurna. Beberapa masalah yang sering dikeluhkan oleh peneliti terdahulu adalah masalah yang diungkapkan pada Gambar 3. Mulai dari citra bakteri yang terpatah kemungkinan karena slide yang sudah lama, sehingga bakterinya sudah menggelembung dan terpatah. Bakteri membentuk kurva, yang bisa terbentuk dari 1 bakteri yang meliuk, atau dari beberapa bakteri yang serial. Citra bakteri blur karena citra bakteri yang seharusnya adalah Binatang yang transparan, walaupun diwarnai, namun disorot di microscope dengan Cahaya dengan intensitas yang tinggi, sehingga hasilnya adalah citra blur.



Gambar 3. Citra patch dari Bakteri yang Normal dan yang Sering Dikeluhkan Peneliti Terdahulu.

2.2 Model Transformasi Citra ke Graf

Pada penelitian terdahulu, terdapat beberapa alternatif model transformasi dari Citra ke Graf, diantaranya diperlihatkan pada Tabel 2.

Tabel 2. Perbandingan Metode Transformasi Citra ke Graf

Metode	Cocok Untuk Citra Input
Adjacency Matix	Citra patch / yang dimensinya kecil
Superpixel Graf	Diawali dengan segmentasi citra yang lebih rumit
KNN (K-Nearest Neighbor)	Antara Objek dan latar belakang harus yang sangat jauh berbeda warnanya
Region Adjacency Graph (RAG)	Diawali dengan segmentasi SAM (Segmented Anything Method) kemudian dilanjutkan dengan Adjacency Matix

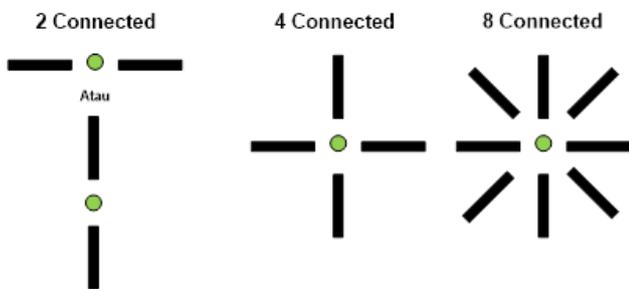
Pada penelitian ini digunakan metode Adjacency Matiks, karena metode ini banyak digunakan oleh peneliti terdahulu, dan terbukti juga masih relevan digunakan untuk penggunaan deep learning yaitu pada metode RAG (Tabel 2). Adjacency Matiks merupakan matriks yang merepresentasikan sebuah graf, yang terdiri dari himpunan simpul (node) dan himpunan sisi (edge). Persamaan dari Graf (G) diperlihatkan pada persamaan 1.

$$G = (V, E) \quad (1)$$

Dimana V merupakan himpunan nodes dan E merupakan himpunan edge, dengan $A \in \mathbb{R}^{n \times n}$, $n = |V|$. Untuk graf yang berasal dari Adjacency Matiks sendiri dilambangkan dengan $A_{(a,b)}$, dengan rumus yang diperlihatkan pada persamaan 2.

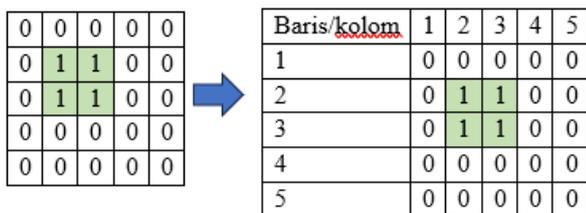
$$A_{(a,b)} = \begin{cases} 1, & \text{jika node } a \text{ terhubung ke node } b \\ 0, & \text{jika tidak terhubung} \end{cases} \quad (2)$$

Matriks Adjacency juga mempertimbangkan koneksi antara node yang satu dengan node yang lain melalui edge, bisa menggunakan 2 koneksi, 4 koneksi dan 8 koneksi. Untuk 2 koneksi bisa kearah kanan dan kiri atau ke arah atas dan bawah. Sedangkan untuk 4 koneksi (kanan, kiri, atas dan bawah), sedangkan untuk 8 koneksi bisa menggunakan diagonalnya. Representasi dari koneksi antara nodes dan edge diperlihatkan pada Gambar 4.



Gambar 4. Representasi Koneksi Antara Node (lingkaran hijau) dan Edge (Garis hitam)

Pada penelitian ini, dipilih 4 koneksi, karena untuk menyederhanakan proses, namun representasi graf yang dihasilkan masih mirip dengan citra aslinya. Sebuah contoh kasus, missal sebuah citra biner direpresntasikan pada sebuah matriks input dengan ukuran matriks 5x5 (Gambar 5).



Gambar 5. Matrix Input

Matriks input terdiri dari baris dan kolom, artinya 0 adalah pixel dengan warna hitam dan 1 atau 255 adalah pixel dengan warna putih. Artinya ada 4 node yang kemudian menjadi himpunan nodes (V) (Tabel 3).

Tabel 3. Himpunan Nodes (V)

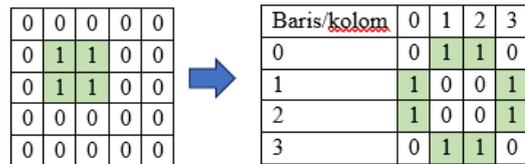
(Baris/Kolom)	Nama Node
(2,2)	0
(2,3)	1
(3,2)	2
(3,3)	3

Setelah itu dilakukan pembuatan himpunan edge (E), yaitu dengan menentukan simpul bertetangga antara node yang satu dengan node yang lain (Tabel 4). Node tetangga dilakukan dengan 4 koneksi.

Tabel 4. Himpunan Edge (E)

Nama Node	Edge (Node Tetangga)
0	(1,2) artinya adalah (0-1 dan 0-2)
1	(0,3) artinya adalah (1-0 dan 1-3)
2	(0,3) artinya adalah (2-0 dan 2-3)
3	(1,2) artinya adalah (3-1 dan 3-2)

Ketika matrix node (V) dan matrix adge (E) dijadikan 1 tabel, maka hasilnya adalah sebuah Adjacency matriks (A), yang diperlihatkan pada Gambar 6.

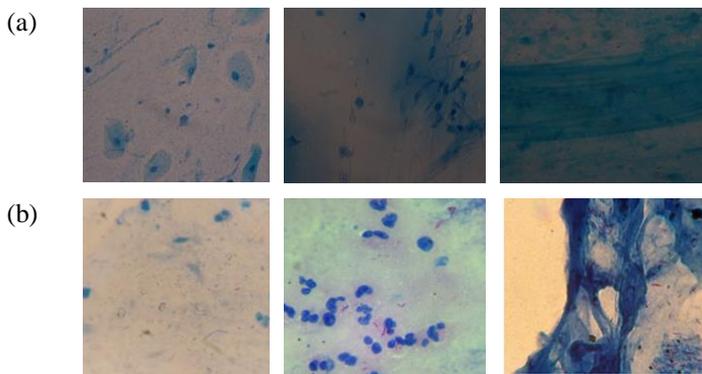


Gambar 6. Matrix Output

III. METODE

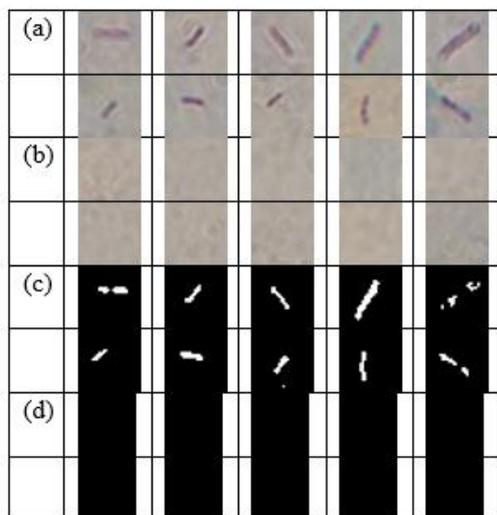
3.1 Dataset

Pada penelitian ini kami memilih menggunakan dataset dari NLM/NIH TB Dataset (USA) karena dilihat dari representasi citranya, citra bakterinya jelas (tidak blur, tidak bertumpuk dengan dahak, dan dahaknya tidak dominan warna biru), selain itu data MIDTI mempunyai kontras citra yang gelap (Gambar 7).



Gambar 7. Perbandingan Antara Citra (a) Dataset MIDTI dan (b) Dataset NLM/NIH TB

Dari dataset tersebut, diambil hanya 1 citra terbaik yang dipotong dengan dimensi 28x28. Citra yang dipotong hanya pada bagian yang ada bakterinya dan bagian kosong yang tidak ada bakterinya. Data patch yang diambil dari dataset NLM/NIH TB terdiri dari 10 citra patch bakteri dan 10 citra patch latar belakang, lengkap dengan citra labelnya yang diperlihatkan pada Gambar 8.



Gambar 8. Dataset Patch yang Digunakan (a) Citra bakteri, (b) Citra Latar Belakang, (c) Citra Label Bakteri, (d) Citra Label Latar Belakang

3.2 Matrix Adjacency

Untuk membentuk matrix adjacency dilakukan proses design software yang ditunjukkan dengan menggunakan Flowchart pada Gambar 8. Dari flowchart yang ditampilkan pada Gambar 8, dimulai dengan pengambilan citra biner dari dataset. Dimana citra biner bisa didapatkan dari citra label, ataupun citra input yang kemudian diubah menjadi citra biner, ataupun beberapa metode yang lain untuk mengubah citra warna menjadi citra biner. Berikutnya dilakukan proses pengambilan nilai koordinat pixel dan dilanjutkan dengan inisialisasi dari

matrix adjacency. Hasil tersebut digunakan sebagai input dari pembuatan jaringan koneksi 4 arah dan hasilkan kemudian ditampilkan dalam bentuk graf.



Gambar 8. Flowchart Matrix Adjacency

3.3 Eksperimen

Blok Diagram dari metode yang digunakan pada penelitian ini diperlihatkan pada Gambar 9. Dimana urutan-nya adalah preprocessing, transformasi dari citra warna ke citra biner. Lalu dilanjutkan dengan transformasi dari citra biner ke graf. Hasilnya merupakan output graf yang digunakan sebagai label untuk sistem pengenalan berikutnya.



Gambar 9. Metode Pengolahan Citra

3.3.1 Preprocessing

Pada tahap pre-processing, dilakukan proses resize atau perubahan ukuran/ dimensi citra, dimana pada penelitian ini telah dipilih ukuran paling kecil dari penelitian penelitian sebelumnya, yaitu 28x28. Peneliti terdahulu menyatakan bahwa pemilihan ukuran 28x28 didasarkan pada panjang bakteri dewasa maksimal adalah 25 pixel, sehingga kemudian disusun menjadi 28x28 untuk dapat menampilkan keseluruhan bakteri dalam 1 citra.

3.3.2 Transformasi Citra dari Citra Warna Ke Citra Biner

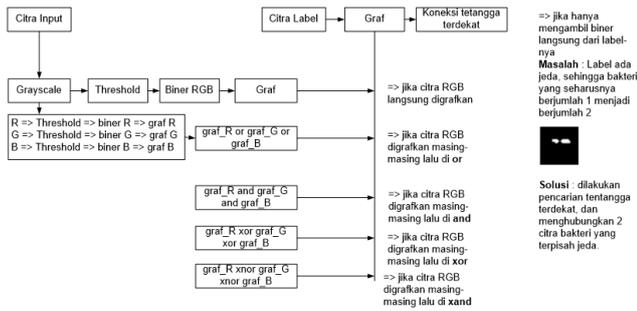
Proses citra warna ke citra biner dilakukan dengan menggunakan proses binerisasi. Citra biner didapatkan dari 2 macam, yaitu langsung menggunakan citra label, atau menggunakan citra warna, lalu dibuat ke grayscale dulu, lalu dithreshold, pada kasus ini digunakan threshold 127-255, sebagai objek-nya, selain itu sebagai latar belakang.

3.3.3 Transformasi Citra dari Citra Biner ke Graf

Transformasi dari citra biner ke graf, dilakukan dengan menggunakan matriks adjacency, Dimana proses perubahan dari citra biner menjadi graf dilakukan seperti pada bab 2.2.

3.4 Perbandingan Metode

Pada perbandingan metode dilakukan dengan melakukan beberapa kombinasi yang diperlihatkan pada Gambar 10.



Gambar 10. Perbandingan Metode

Terdapat 5 perbandingan metode. Yaitu Ketika RGB langsung digrafkan (Graf RGB), RGB digrafkan masing-masing kemudian di “or” (Graf OR), RGB digrafkan masing-masing kemudian di “and” (Graf AND), RGB digrafkan masing-masing kemudian di “xor” (Graf XOR), dan RGB digrafkan masing-masing kemudian di “xnor” (Graf XNOR).

3.5 Pengujian Performasi

Pengujian performasi dilakukan dengan menggunakan matrix confusion, Dimana dimatrix confusion terdapat 4 buah ruang yaitu pada Gambar 10.

		Prediksi	
		Objek	Latar Belakang
Label	Objek	True Positive	False Positive
	Latar Belakang	False Negative	True Negative

Gambar 10. Matrix Confusion

Pada matrix confisoion tersebut, true positif artinya adalah sebuah label objek terdeteksi sebagai objek, sedangkan untuk false positive merupakan sebuah label objek yang terdeteksi sebagai latar belakang. Untuk false negative adalah sebuah label latar belakang yang terdeteksi sebagai objek dan sebuah label latar belakang yang terdeketsi sebagai latar belakang.

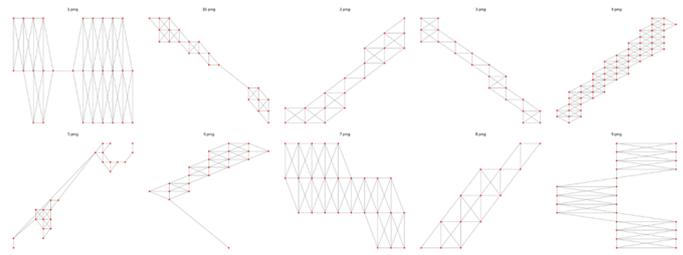
IV. HASIL DAN DISKUSI

4.1 Hasil Pengolahan Citra

4.1.1 Graf Citra Label

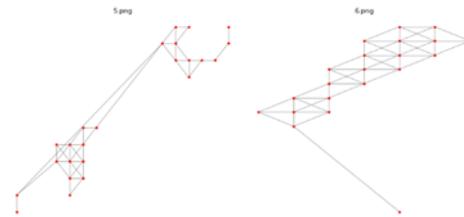
Citra label merupakan citra biner yang digunakan sebagai dasar dari graf yang lain. Citra biner ini kemudian langsung ditransformasikan ke graf sebagai dasar atau ground truth untuk digunakan sebagai acuan pada hasil pengujian. Hasil graf dari citra label diperlihatkan pada Gambar 11.

Kekurangan dari label yang ada didataset public adalah, label yang seharusnya hanya 1 bakteri terdapat jeda, sehingga kemudian terhitung menjadi 2 objek atau 2 bakteri. Jeda ini terjadi kemungkinan karena threshold yang tidak sesuai pada proses transformasi dari citra warna menjadi citra biner. Hal ini dapat diatasi dengan menghubungkan 2 graf yang berdekatan menggunakan jarak, Hasil dari graf yang sudah dihubungkan diperlihatkan pada Gambar 12.



Gambar 12. Graf dari Citra Label

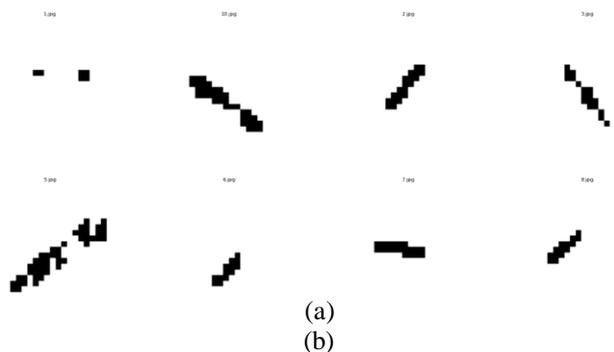
Gambar 12 memperlihatkan hasil dari penambahan koneksi antar graf. Setelah dilakukan pengamatan, dari 10 citra terdapat 2 citra yang mempunyai pola yang berbeda dengan citra asal (citra warna)-nya, yaitu pada citra 6 dan 7. Pada citra 6 dan citra 7, terdapat graf tambahan, Dimana seharusnya graf tersebut terjadi karena ada objek kecil yang bukan merupakan bagian dari bakteri. Node dari objek kecil tersebut hanya berjumlah 1-2 node, artinya seharusnya dapat ditambahkan filter untuk menghilangkan graf, dengan node dengan jumlah yang sedikit sehingga dapat mengurangi ketidaktepatan dalam proses deteksi. Graf yang salah beserta node jumlah kecil diperlihatkan pada Gambar 13.



Gambar 13. Graf yang salah beserta node jumlah kecil

4.1.2 Graf RGB

Graf RGB merupakan graf yang dihasilkan dari transformasi dari citra warna ke citra grayscale untuk kemudian dilakukan proses thresholding untuk menjadi citra biner. Citra biner yang dihasilkan kemudian digrafkan. Hasil dari citra biner dan graf RGB diperlihatkan pada Gambar 14.



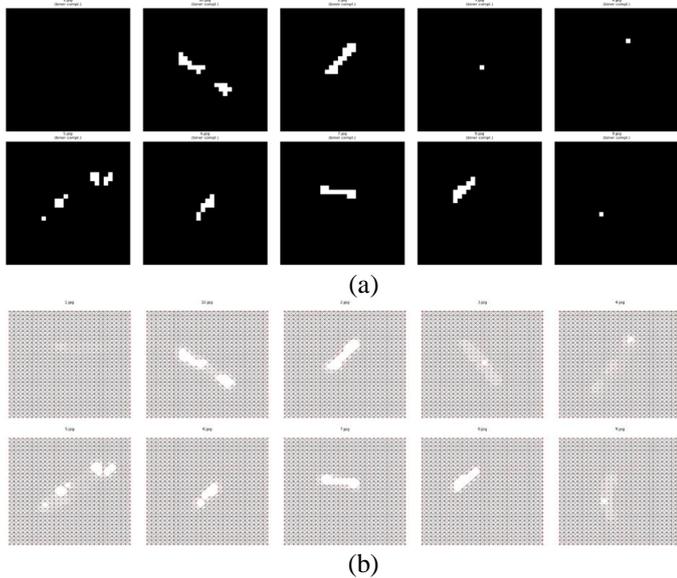
Gambar 14. Hasil (a) Citra Biner Graf RGB, (b) Graf RGB

4.1.3 Graf OR

Graf OR merupakan graf yang dihasilkan dari pemisahan warna dari citra warna RGB, sehingga menjadi 3 warna R-G dan B, yang masing-masing warna kemudian digrafkan, untuk

Identify applicable sponsor/s here. If no sponsors, delete this text box (sponsors).

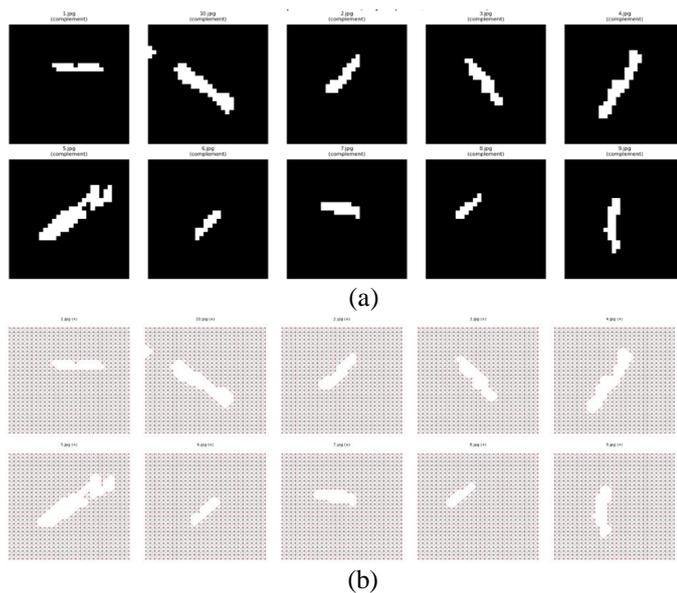
hasil dari graf masing-masing warna kemudian di "or" kan. Hasil dari citra biner dan graf yang di"or"kan diperlihatkan pada Gambar 15.



Gambar 15. Hasil (a) Citra Biner, (b) Graf yang salah di-"or"-kan

4.1.4 Graf AND

Graf AND merupakan graf yang dihasilkan dari pemisahan warna dari citra warna RGB, sehingga menjadi 3 warna R-G dan B, yang masing-masing warna kemudian digrafkan, untuk hasil dari graf masing-masing warna kemudian di "and" kan. Hasil dari citra biner dan graf yang di"and"kan diperlihatkan pada Gambar 16.

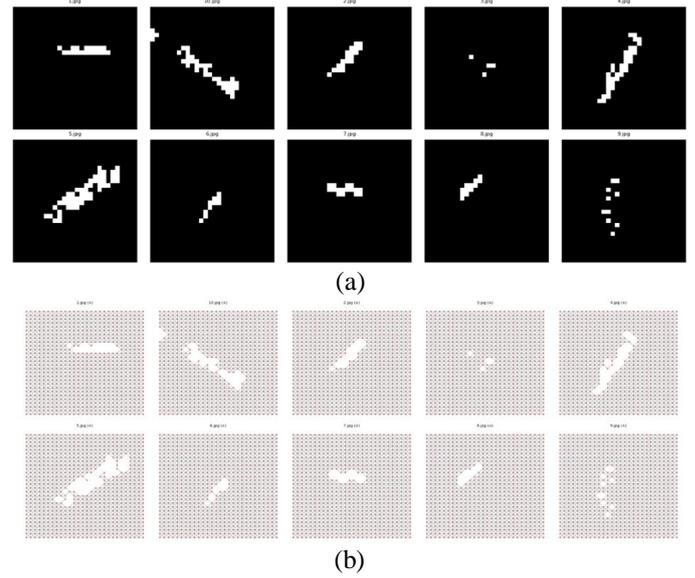


Gambar 16. Hasil (a) Citra Biner, (b) Graf yang salah di-"and"-kan

4.1.5 Graf XOR

Graf XOR merupakan graf yang dihasilkan dari pemisahan warna dari citra warna RGB, sehingga menjadi 3 warna R-G dan B, yang masing-masing warna kemudian digrafkan, untuk

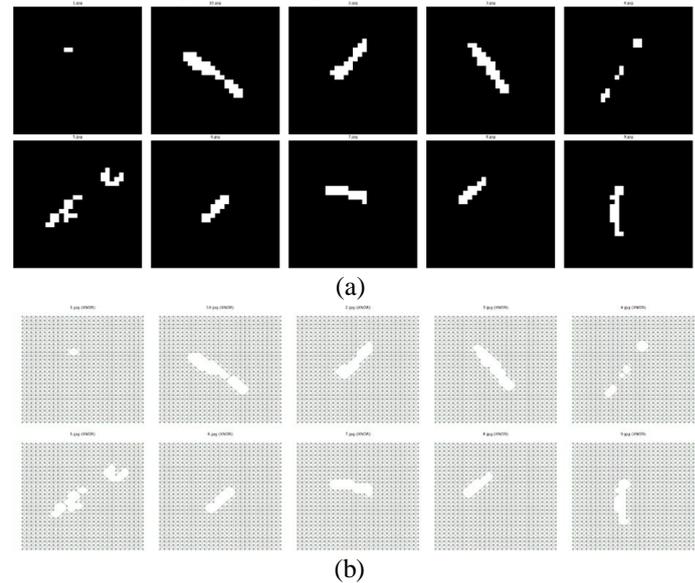
hasil dari graf masing-masing warna kemudian di "xor" kan. Hasil dari citra biner dan graf yang di"xor"kan diperlihatkan pada Gambar 17.



Gambar 17. Hasil (a) Citra Biner, (b) Graf yang salah di-"xor"-kan

4.1.6 Graf XNOR

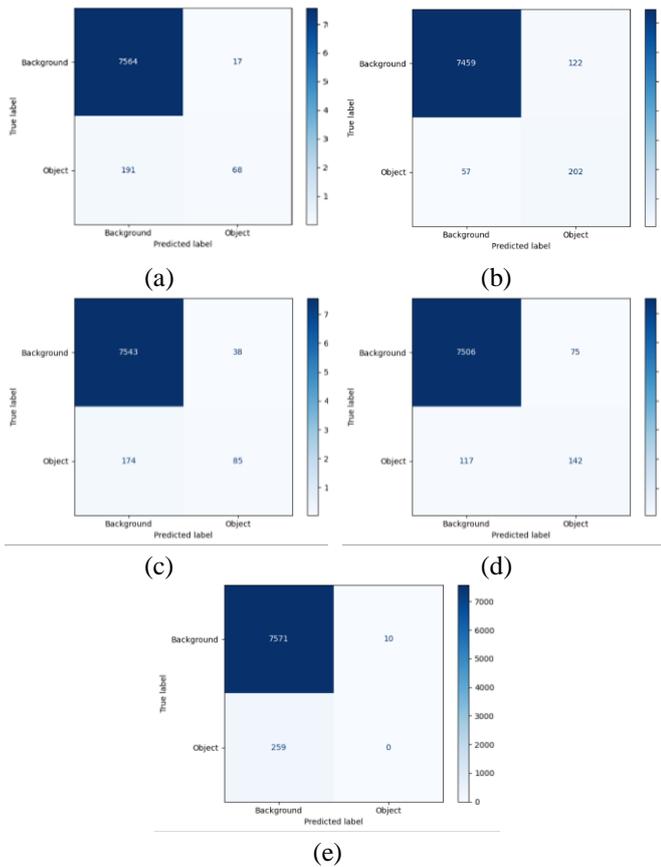
Graf XNOR merupakan graf yang dihasilkan dari pemisahan warna dari citra warna RGB, sehingga menjadi 3 warna R-G dan B, yang masing-masing warna kemudian digrafkan, untuk hasil dari graf masing-masing warna kemudian di "xnor" kan. Hasil dari citra biner dan graf yang di"xnor" kan diperlihatkan pada Gambar 18.



Gambar 18. Hasil (a) Citra Biner, (b) Graf yang salah di-"xand"-kan

4.2 Hasil Perbandingan

Hasil confusion matrix dari graf RGB, graf or, graf and, graf xor dan graf xnor diperlihatkan pada Gambar 19.



Gambar 19. Perhitungan (a) Graf RGB , (b) Graf OR, (c) Graf AND, (d) Graf XOR dan (e) Graf XNOR

4.3 Diskusi

Dari sekilas, hasil dari analisis qualitative menunjukkan bahwa graf yang dibentuk dari hasil AND antara setiap kanal warna dari citra aslinya. Namun demikian kita juga harus melihat parameter hasil dari confusion matrix yang ditunjukkan pada Tabel 4.

Tabel 4. Performasi Sistem

Model	Accuracy	Specificity	Sensitivity
Graf RGB : TN=7564, FP=17, FN=191, TP=68	0.9734	0.9978	0.2625
Graf OR TN = 7459 FP = 122 FN = 57 TP = 202	0.9730	0.9950	0.3282
Graf AND TN = 7543 FP = 38 FN = 174 TP = 85	0.9772	0.9839	0.7799
Graf XOR	0.9755	0.9901	0.5483

TN = 7506 FP = 75 FN = 117 TP = 142			
Graf XNOR TN = 7571 FP = 10 FN = 259 TP = 0	0.9657	0.9987	0.0000

Dari table 4 diperlihatkan bahwa model dengan graf and mempunyai akurasi yang paling besar dibandingkan dengan model yang lain. Artinya hal ini sesuai dengan analisis qualitative berdasarkan hasil graf. Pembuatan label graf ternyata juga sangat dipengaruhi oleh metode pembentukan citra biner. Jika pembentukan citra binernya mampu memetakan objek secara tepat, maka graf berbasis matrix adjacency-nya juga akan tepat.

4.4 Penelitian Berikutnya

Artikel ini menjelaskan tentang metode terbaik untuk pembuatan label graph unsupervise. Label ini berikutnya dapat digunakan untuk model yang berbasis GNN.

V. PENUTUP

5.1 Kesimpulan

Graf NeuroNetwork, merupakan metode terbaru berbasis graf yang memungkinkan penggunaan AI model terbaru, namun menggunakan komputasi yang lebih ringan dan lebih cepat. Namun, aplikasi pembuatan label citra yang berbentuk graf masih sangat sedikit, berbeda dengan label citra yang tidak berbentuk graf seperti labelme dan SAM (Segment Anything Model). Potensi pembuatan label graf terutama yang unsupervised masih sangat besar. Kontribusi yang diberikan adalah (1) Membuat dataset patch baru khusus untuk pembuatan label graf, (2) Menghasilkan dataset label graf yang berikutnya berguna pada metode GNN. Model graf berbasis matrix adjacency. Proses pembuatannya dimulai dari melakukan transformasi dari citra warna ke citra biner, lalu dilanjutkan dengan pembuatan graf menggunakan matrix adjacency. Untuk studi perbandingan dipaparkan menggunakan graf RGB, graf OR, graf AND, graf XOR dan graf XNOR. Hasil perbandingan dianalisis menggunakan matrix conclution, dari perbandingan tersebut didapatkan hasil bahwa metode graf AND adalah metode dengan tingkat akurasi terbaik yaitu 0,9772. Artinya, matrix adjacency dengan metode graf AND mampu mengenali pola bakteri dengan tepat.

5.2 Saran

Pembuatan label graf ternyata juga sangat dipengaruhi oleh metode pembentukan citra biner. Jika pembentukan citra binernya mampu memetakan objek secara tepat, maka graf berbasis matrix adjacency-nya juga akan tepat.

CODE LINK

https://colab.research.google.com/drive/1xcSY0TKLenDR_pXahiNHvNvyBHNYQ2em?usp=sharing

VIDEO LINK AT YOUTUBE

<https://drive.google.com/drive/folders/1jFr8UgWp-0eXVE05UHqa5q7JgGXlYlYY?usp=sharing>

DATASET LINK

<https://drive.google.com/drive/folders/1c3Du0qJKcFLIgw su7WoRigPIpmCIINB?usp=sharing>

ACKNOWLEDGMENT

Puji Syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas limpahan Rahmat dan karunia-Nya penulis dapat menyelesaikan makalah ini dengan baik. Ucapan terimakasih disampaikan kepada Bapak Arrival Dwi Sentosa, S.Kom., M.T. yang telah memberikan pengajaran dan pemahaman kepada mahasiswa dengan sangat baik.

REFERENCES

- [1] L. F. Sua *et al.*, "Detection of mycobacteria in paraffin-embedded Ziehl-Neelsen-Stained tissues using digital pathology," *Tuberculosis*, vol. 126, 2021, doi: 10.1016/j.tube.2020.102025.
- [2] K. Adi *et al.*, "Autothresholding Segmentation for Tuberculosis Bacteria Identification in the Ziehl-Neelsen Sputum," *The Proceedings of The 7th ICTS, Bali*, pp. 9–14, 2013.
- [3] P. Sadaphal, J. Rao, G. W. Comstock, and M. F. Beg, "Image processing techniques for identifying Mycobacterium tuberculosis in Ziehl-Neelsen stains."
- [4] 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. 51364–51380, 2024, doi: 10.1109/ACCESS.2024.3386208.
- [5] K. Adi *et al.*, "Autothresholding Segmentation for Tuberculosis Bacteria Identification in the Ziehl-Neelsen Sputum," *The Proceedings of The 7th ICTS, Bali*, pp. 9–14, 2013.
- [6] V. Shwetha, K. Prasad, C. Mukhopadhyay, B. Banerjee, and A. Chakrabarti, "Automatic Detection of Bacilli Bacteria from Ziehl-Neelsen Sputum Smear Images," in *Proceedings of the 2021 2nd International Conference on Communication, Computing and Industry 4.0, C2I4 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/C2I454156.2021.9689283.
- [7] R. A. A. Raof, M. Y., R. B., and S. S. M. Noor, "Image Segmentation of Ziehl-Neelsen Sputum Slide Images for Tubercle Bacilli Detection," in *Image Segmentation, InTech*, 2011. doi: 10.5772/15808.
- [8] C. del Carpio *et al.*, "An algorithm for detection of Tuberculosis bacilli in Ziehl-Neelsen sputum smear images," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 4, pp. 2968–2981, 2019, doi: 10.11591/ijece.v9i4.pp2968-2981.
- [9] J. L. Díaz-Huerta, A. del Carmen Téllez-Anguiano, M. Fraga-Aguilar, J. A. Gutiérrez-Gnecchi, and S. Arellano-Calderón, "Image processing for AFB segmentation in bacilloscopies of pulmonary tuberculosis diagnosis," *PLoS One*, vol. 14, no. 7, pp. 1–14, 2019, doi: 10.1371/journal.pone.0218861.
- [10] V. Khare, A. Agrawal, P. Gupta, and S. Saxena, "Artificial Intelligence based AFB microscopy for Pulmonary Tuberculosis in North India: A pilot study," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 9, no. 12, p. p9669, Dec. 2019, doi: 10.29322/ijsrp.9.12.2019.p9669.
- [11] "A Novel Architecture for Improving Tuberculosis Detection from Microscopic Sputum Smear Images." [Online]. Available: https://www.researchgate.net/publication/362322260_A_Novel_Architecture_for_Improving_Tuberculosis_Detection_from_Microscopic_Sputum
- [12] R. Vidal, N. Martín-Casabona, A. Juan, T. Falgueras, and M. Miravittles, "Incidence and significance of acid-fast bacilli in sputum smears at the end of antituberculous treatment," *Chest*, vol. 109, no. 6, pp. 1562–1565, 1996, doi: 10.1378/chest.109.6.1562.
- [13] C. M. Lo, Y. H. Wu, Y. C. (Jack) Li, and C. C. Lee, "Computer-aided bacillus detection in whole-slide pathological images using a deep convolutional neural network," *Applied Sciences (Switzerland)*, vol. 10, no. 12, pp. 1–12, 2020, doi: 10.3390/APP10124059.
- [14] C. Spahn *et al.*, "DeepBacs for multi-task bacterial image analysis using open-source deep learning approaches," 2022. [Online]. Available: <https://www.nature.com/articles/s42003-022-03634-z>
- [15] Y. Zaizen *et al.*, "Deep-Learning-Aided Detection of Mycobacteria in Pathology Specimens Increases the Sensitivity in Early Diagnosis of Pulmonary Tuberculosis Compared with Bacteriology Tests," *Diagnostics*, vol. 12, no. 3, Mar. 2022, doi: 10.3390/diagnostics12030709.
- [16] A. Fagg, S. Sridharan, S. Lucey, and P. Art, "Unsupervised Temporal Ensemble Alignment For Rapid Annotation," pp. 1–14.
- [17] F. Zhao, Z. Zeng, H. Liu, R. Lan, and J. Fan, "Semisupervised Approach to Surrogate-Assisted Multiobjective Kernel Intuitionistic Fuzzy Clustering Algorithm for Color Image Segmentation," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 6, pp. 1023–1034, Jun. 2020, doi: 10.1109/TFUZZ.2020.2973121.
- [18] S. A. Wulandari, I. K. Eddy Purnama, E. M. Yuniarno, and M. H. Purnomo, "Multicluster Kernel Intuitionistic

- Fuzzy C-Means and State Transition Algorithm: Framework for Low Light Segmentation Imaging of Tuberculosis Bacilli Base of Semisupervised Approach,” in *IEEE International Conference on Fuzzy Systems*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/FUZZ-IEEE60900.2024.10612071.
- [19] K. Swetha, B. Sankaragomathi, and J. B. Thangamalar, “Convolutional Neural Network based Automated Detection of Mycobacterium Bacillus from Sputum Images,” in *Proceedings of the 5th International Conference on Inventive Computation Technologies, ICICT 2020*, Institute of Electrical and Electronics Engineers Inc., Feb. 2020, pp. 293–300. doi: 10.1109/ICICT48043.2020.9112501.
- [20] J. L. Díaz-Huerta, A. del Carmen Téllez-Anguiano, M. Fraga-Aguilar, J. A. Gutiérrez-Gnecchi, and S. Arellano-Calderón, “Image processing for AFB segmentation in bacilloscopies of pulmonary tuberculosis diagnosis,” *PLoS One*, vol. 14, no. 7, Jul. 2019, doi: 10.1371/journal.pone.0218861.
- [21] C. M. Lo, Y. H. Wu, Y. C. (Jack) Li, and C. C. Lee, “Computer-aided bacillus detection in whole-slide pathological images using a deep convolutional neural network,” *Applied Sciences (Switzerland)*, vol. 10, no. 12, pp. 1–12, Jun. 2020, doi: 10.3390/APP10124059.
- [22] R. Rulaningtyas, A. B. Suksmono, T. L. R. Mengko, and P. Saptawati, “Nearest patch matching for color image segmentation supporting neural network classification in pulmonary tuberculosis identification,” in *AIP Conference Proceedings*, American Institute of Physics Inc., 2016. doi: 10.1063/1.4943354.
- [23] D. Gokay, E. Simsar, E. Atici, A. Ahmetoglu, A. E. Yuksel, and P. Yanardag, “Graph2Pix: A Graph-Based Image to Image Translation Framework.” [Online]. Available: <https://www.artbreeder.com/>
- [24] C. Gao, Y. Wang, J. Zhou, W. Ding, L. Shen, and Z. Lai, “Possibilistic Neighborhood Graph: A New Concept of Similarity Graph Learning,” *IEEE Trans Emerg Top Comput Intell*, vol. 7, no. 6, pp. 1636–1650, Dec. 2023, doi: 10.1109/TETCI.2022.3225173.
- [25] V. Vasudevan, M. Bassenne, M. T. Islam, and L. Xing, “Image classification using graph neural network and multiscale wavelet superpixels,” *Pattern Recognit Lett*, vol. 166, pp. 89–96, 2023, doi: 10.1016/j.patrec.2023.01.003.
- [26] Y. Liu, J. Zhang, C. Bharat, and J. Ren, “Actin Cytoskeleton Morphology Modeling Using Graph Embedding and Classification in Machine Learning,” in *IFAC-PapersOnLine*, Elsevier B.V., Nov. 2021, pp. 328–333. doi: 10.1016/j.ifacol.2021.11.195.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 1 Juni 2025



Mirza Tsabita W
(13524114)