

PENGARUH RASIO PEMBAGIAN DATASET PADA KINERJA CNN KLASIFIKASI TUBERKULOSIS

Yovi Ibnu Nasikhin¹, Basuki Rahmat², Chrystia Aji Putra³

^{1,2,3} Program Studi Informatika, Universitas Pembangunan Nasional "Veteran" Jawa Timur

¹ korespondensi: 21081010128@student.upnjatim.ac.id

ABSTRACT

Tuberculosis remains a global health challenge with a high mortality rate. Chest X-rays are commonly used in diagnosis, but their interpretation is often subjective. Convolutional Neural Networks (CNN) can be one solution for automatic tuberculosis detection. However, CNN performance is greatly influenced by data quality, including data distribution. Therefore, this study aims to analyze the effect of dataset distribution ratio on the performance of the Convolutional Neural Network (CNN) model in classifying chest X-ray images for tuberculosis detection. The dataset used consists of 2,198 images from various public sources with two classes, namely Normal and Tuberculosis. Two dataset splitting scenarios were applied, namely 80:10:10 and 70:10:20, to evaluate the effect of data proportion on model performance. The CNN model was designed with three convolutional layers and trained using the Adam optimizer and binary cross-entropy loss function. The evaluation was conducted using accuracy, precision, recall, and F1-score metrics. The results showed that the model with an 80:10:10 ratio had the best performance with a testing accuracy of 85.97% and a loss of 0.3217, indicating better generalization capabilities compared to the 70:10:20 model. These findings confirm that selecting the appropriate dataset ratio can improve the stability and accuracy of CNN in classifying medical images. The results of this study are expected to support efforts to improve the quality of early detection of tuberculosis in order to accelerate diagnosis and contribute to reducing the mortality rate from tuberculosis in the community.

Keywords: Tuberculosis, Convolutional Neural Network, Dataset Splitting, Chest X-ray, Medical Image Classification

ABSTRAK

Tuberkulosis masih menjadi tantangan kesehatan global dengan angka kematian yang tinggi. Citra X-ray dada umum digunakan dalam diagnosis, tetapi interpretasinya sering subjektif. Convolutional Neural Network (CNN) dapat menjadi salah satu solusi untuk deteksi tuberkulosis secara otomatis. Namun, kinerja CNN sangat dipengaruhi oleh kualitas data, salah satunya pembagian dataset. Oleh karena itu penelitian ini bertujuan untuk menganalisis pengaruh rasio pembagian dataset terhadap kinerja model Convolutional Neural Network (CNN) dalam klasifikasi citra X-ray dada untuk deteksi tuberkulosis. Dataset yang digunakan terdiri atas 2.198 citra dari berbagai sumber publik dengan dua kelas, yaitu Normal dan Tuberculosis. Dua skenario pembagian dataset diterapkan, yakni 80:10:10 dan 70:10:20, untuk mengevaluasi pengaruh proporsi data terhadap performa model. Model CNN dirancang dengan tiga lapisan konvolusi dan dilatih menggunakan optimizer Adam serta fungsi loss binary cross-entropy. Evaluasi dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil menunjukkan bahwa model dengan rasio 80:10:10 memiliki kinerja terbaik dengan akurasi pengujian sebesar 85,97% dan loss 0,3217, menunjukkan kemampuan generalisasi yang lebih baik dibandingkan model 70:10:20. Temuan ini menegaskan bahwa pemilihan rasio dataset yang tepat dapat meningkatkan stabilitas dan akurasi CNN dalam mengklasifikasikan citra medis. Dari hasil penelitian ini diharapkan dapat mendukung upaya peningkatan kualitas deteksi dini tuberkulosis guna mempercepat diagnosis dan berkontribusi terhadap penurunan angka kematian akibat Tuberkulosis di masyarakat.

Kata Kunci: Tuberkulosis, Convolutional Neural Network, Pembagian Dataset, X-ray Dada, Klasifikasi Citra Medis

PENDAHULUAN

Tuberkulosis merupakan salah satu penyakit menular yang masih menjadi masalah kesehatan global. Menurut laporan World

Health Organization (WHO) tahun 2023, terdapat sekitar 10,8 juta kasus baru tuberkulosis di seluruh dunia dengan 1,25 juta kematian. Indonesia menempati posisi lima

besar negara dengan kasus tertinggi, menyumbang sekitar 10% dari total kasus global atau sekitar 1,08 juta penderita (1). Kondisi ini menunjukkan perlunya inovasi dalam penanganan tuberkulosis, termasuk melalui teknologi informasi untuk mendukung proses diagnosis yang lebih cepat dan akurat.

Salah satu metode utama deteksi tuberkulosis adalah pemeriksaan citra *X-ray* dada, yang menampilkan kondisi paru-paru secara visual. Sebuah penelitian mengungkapkan bahwa *X-ray* dada mampu mendeteksi antara 40-79 persen kasus tuberkulosis pada individu yang tidak mengalami gejala umum (2). Namun, interpretasinya sering terkendala oleh keterbatasan seseorang, seperti kelelahan dan subjektivitas tenaga medis (3). Oleh karena itu, pemanfaatan kecerdasan buatan, khususnya *Convolutional Neural Network* (CNN), menjadi solusi potensial. CNN mampu mengekstraksi dan mengenali pola visual kompleks secara otomatis, sehingga dapat meningkatkan kecepatan dan akurasi diagnosis tuberkulosis (4). Namun, kinerja CNN dalam tugas klasifikasi citra sangat dipengaruhi oleh kualitas *dataset* yang digunakan, terutama dalam tahap pra-pemrosesan dan pembagian data latih, validasi, serta data uji. Rasio pembagian *dataset* yang tidak tepat dapat menyebabkan model mengalami *overfitting* atau *underfitting*, sehingga mengurangi kemampuan model dalam melakukan generalisasi terhadap data baru (5).

Berbagai penelitian terdahulu menunjukkan bahwa CNN telah berhasil diterapkan dalam analisis citra medis, termasuk untuk mendeteksi penyakit tuberkulosis. Sebagai contoh, penelitian yang menggunakan data dari *National Library of Medicine* berhasil mengembangkan model CNN dengan akurasi klasifikasi mencapai 84% pada data uji (6). Penelitian lain membahas sistem *Computer-Aided Diagnosis (CAD)* berbasis CNN yang mampu mempercepat proses diagnosis tuberkulosis melalui analisis otomatis citra *X-ray* dada menggunakan berbagai arsitektur seperti *AlexNet*, *VGGNet*, dan *ResNet* (7). Meskipun demikian, sebagian besar penelitian tersebut berfokus pada pengembangan arsitektur jaringan, sementara aspek penting lain seperti pengaruh rasio pembagian *dataset* terhadap performa CNN masih jarang dianalisis secara mendalam. Padahal, proporsi data latih, validasi, dan uji sangat berperan dalam menentukan kemampuan model untuk melakukan generalisasi terhadap data baru. Berdasarkan hal tersebut, penelitian ini dilakukan untuk menganalisis pengaruh perbedaan rasio pembagian *dataset* terhadap performa model CNN dalam klasifikasi citra *X-ray* dada untuk deteksi tuberkulosis. Hasil penelitian ini diharapkan dapat memberikan gambaran bagi peneliti selanjutnya dalam menentukan strategi pembagian data yang efektif, sekaligus mendukung upaya peningkatan kualitas deteksi dini tuberkulosis guna mempercepat diagnosis dan berkontribusi terhadap penurunan angka kematian akibat Tuberkulosis di masyarakat.

METODE

Pengumpulan dan Pengolahan Data

Penelitian ini menggunakan data sekunder yang diperoleh dari beberapa public dataset di platform Kaggle, yaitu *DA and DB – TB Chest X-ray Datasets*, *Pulmonary Chest X-Ray Abnormalities*, dan *Nigeria Chest X-ray Dataset*. Dari kumpulan dataset tersebut, diperoleh total 2.378 citra *X-ray* dada, yang terdiri dari 1.179 citra Normal dan 1.169 citra Tuberkulosis. Penggunaan data dari berbagai sumber dimaksudkan untuk meningkatkan variasi citra. Selanjutnya seleksi citra dilakukan untuk menyeleksi yang tidak layak yang dapat memengaruhi proses pembelajaran. Citra yang tidak layak akan dihapus untuk memastikan seluruh citra yang ada berada dalam kondisi baik dan dapat digunakan. Hasil dari proses ini menghasilkan total 2.198 citra *X-ray* dada, terdiri atas 1.173 citra Normal dan 1.025 citra Tuberkulosis.

Citra-citra ini kemudian dibagi menjadi set pelatihan, validasi, dan pengujian dengan dua skenario pembagian yang berbeda, yaitu rasio 80:10:10 dan 70:10:20. Selanjutnya, citra dikonversi menjadi *grayscale*, diubah ukurannya menjadi 224×224 piksel agar seragam, dan dinormalisasikan. Untuk meningkatkan ketahanan model dan mengurangi risiko *overfitting*, diterapkan teknik augmentasi data seperti rotasi, pergeseran, dan pembesaran.

Perancangan Model CNN

Penelitian ini menggunakan *Convolutional Neural Network* (CNN) untuk melakukan klasifikasi citra *X-ray* dada ke dalam dua

kelas, yaitu Normal dan *Tuberculosis*. CNN dipilih karena kemampuannya dalam mengekstraksi pola spasial dan tekstur secara otomatis dari data citra tanpa memerlukan proses ekstraksi fitur manual (4). Dengan demikian, model dapat mengenali pola-pola visual khas pada paru-paru, seperti area opasitas, bercak, atau perbedaan densitas yang sering muncul pada pasien tuberkulosis (8). Model CNN ini dirancang khusus untuk kasus klasifikasi biner. Struktur model terdiri atas beberapa lapisan yaitu lapisan *input*, konvolusi, *pooling*, *flatten*, *fully connected*, dan *output*. Lapisan *input* menerima citra *X-ray* dada berukuran 224×224 piksel dalam format *grayscale*, yang kemudian menjadi masukan awal bagi proses ekstraksi fitur pada lapisan konvolusi. Kemudian tiga lapisan konvolusi dibuat dengan ukuran *filter* 3×3 dan jumlah *filter* berturut-turut 32, 64, dan 128, di mana setiap lapisan diikuti oleh lapisan *max pooling* berukuran 2×2 untuk mengekstraksi fitur secara bertahap sekaligus mengurangi dimensi spasial citra (9,10). Setelah konvolusi dan *pooling*, keluaran diratakan menggunakan lapisan *flatten* yang berfungsi untuk mengubah data berdimensi dua menjadi bentuk satu dimensi sehingga dapat diproses oleh lapisan berikutnya (11). Data hasil *flatten* kemudian diteruskan ke dua lapisan *fully connected (dense)* dengan masing-masing 64 dan 32 units. *Fully connected* ini berperan menggabungkan seluruh fitur hasil ekstraksi untuk membentuk pola hubungan nonlinier dan menghasilkan keputusan akhir pada proses klasifikasi (11). Untuk mengurangi risiko *overfitting*, ditambahkan lapisan

dropout dengan *rate* 0.4 setelah setiap lapisan *dense*, sesuai konsep *dropout* yang berfungsi meningkatkan kemampuan generalisasi model (12). Terakhir, digunakan lapisan output *sigmoid* dengan satu units untuk melakukan klasifikasi biner antara kelas Normal dan *Tuberculosis*.

Model dilatih selama 20 *epoch* menggunakan *optimizer Adam* dengan *learning rate* 0.001 serta fungsi *loss binary cross-entropy*. Mekanisme *early stopping* dengan *patience* tiga *epoch* diterapkan untuk mencegah *overfitting*.

Evaluasi Model

Setiap skenario pembagian *dataset* akan dilatih menggunakan model CNN yang telah dirancang, kemudian hasil keduanya dibandingkan berdasarkan performa yang diperoleh. Evaluasi performa model dilakukan menggunakan beberapa metrik, yaitu akurasi,

precision, *recall*, dan *F1-score*. Selain itu, digunakan juga *confusion matrix* untuk menganalisis distribusi hasil klasifikasi antara citra Normal dan *Tuberculosis*, sehingga dapat diketahui tingkat kesalahan prediksi pada masing-masing kelas.

Melalui analisis metrik dan hasil *confusion matrix* tersebut, dapat diidentifikasi rasio pembagian *dataset* yang memberikan kinerja terbaik bagi model CNN dalam melakukan klasifikasi citra *X-ray* dada.

HASIL DAN PEMBAHASAN

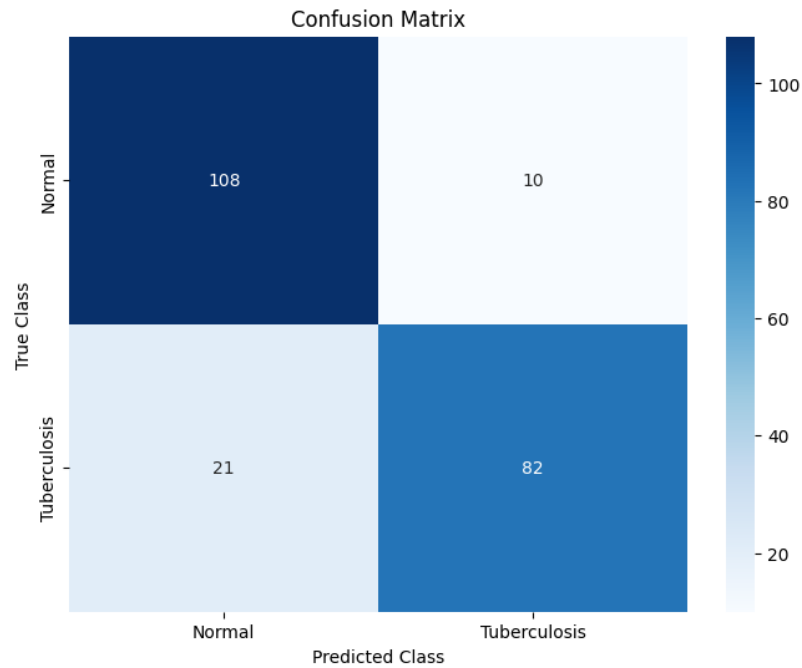
Proses pelatihan model dengan kedua skenario pembagian *dataset* menunjukkan kinerja yang berbeda. Model dengan pembagian 80:10:10 berhenti pada *epoch* ke-12 akibat *early stopping*, sedangkan model 70:10:20 berhenti pada *epoch* ke-17 meskipun sebenarnya model dilatih selama 20 *epoch*.

Tabel 1: Perbandingan Akurasi dan Loss Antar Model

Model	Train Acc	Train Loss	Val Acc	Val Loss	Test Acc	Test Loss
80:10:10	75,42%	0,4745	82,19%	0,4039	85,97%	0,3217
70:10:20	76,75%	0,4787	79,00%	0,4434	83,67%	0,3528

Berdasarkan Tabel 1, terlihat bahwa model dengan pembagian data 80:10:10 menunjukkan kinerja yang lebih baik pada data validasi dan pengujian walaupun memiliki akurasi pelatihan yang sedikit lebih rendah. Hal ini menunjukkan bahwa model 80:10:10 memiliki kemampuan generalisasi

yang lebih baik dibandingkan model 70:10:20. Nilai *loss* yang secara konsisten lebih rendah pada model 80:10:10 dibandingkan model 70:10:20 di semua tahap evaluasi mengindikasikan bahwa model ini lebih efisien dalam mengurangi kesalahan prediksi.



Gambar 1: Confusion Matrix CNN 80:10:10

Berdasarkan temuan tersebut, analisis lanjutan dilakukan menggunakan *confusion matrix* untuk melihat distribusi prediksi pada masing-masing kelas secara lebih rinci. Berdasarkan hasil model CNN 80:10:10 pada Gambar 1, dari 118 citra kelas Normal, sebanyak 108 citra berhasil diklasifikasikan dengan benar, sementara 10 citra lainnya salah terdeteksi sebagai *Tuberculosis*. Pada kelas *Tuberculosis*, dari 103 citra, sebanyak 82

terklasifikasi dengan benar, sedangkan 21 citra salah dikenali sebagai Normal.

Hasil ini menunjukkan bahwa model ini lebih akurat dalam mengenali citra Normal dibandingkan *Tuberculosis*. Hal tersebut kemungkinan disebabkan oleh pola visual pada citra Normal yang lebih seragam dan mudah dipelajari, sementara variasi tekstur pada citra *Tuberculosis* membuat model masih kesulitan membedakan fitur khas penyakit tersebut.

```
Classification Report:
              precision    recall  f1-score   support

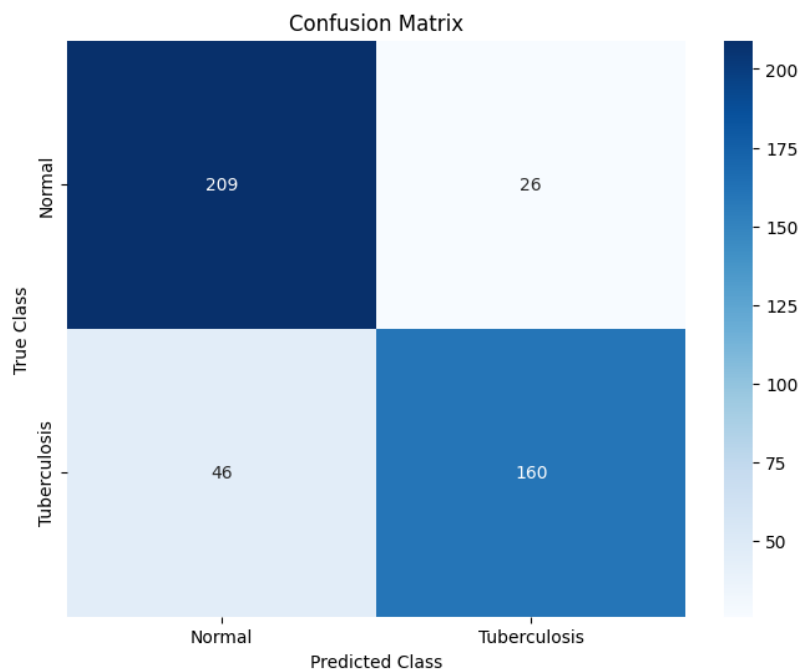
   Normal      0.84      0.92      0.87      118
 Tuberculosis  0.89      0.80      0.84      103

 accuracy              0.86      221
 macro avg             0.86      0.86      0.86      221
 weighted avg          0.86      0.86      0.86      221
```

Gambar 2: Classification Report CNN 80:10:10

Evaluasi kinerja model CNN 80:10:10 menggunakan *classification report* menunjukkan bahwa pada kelas Normal, model mencapai *precision* 0,84, *recall* 0,92, dan *F1-score* 0,87, menandakan sebagian besar citra Normal berhasil dikenali dengan baik. Sementara pada kelas *Tuberculosis*, model memperoleh *precision* 0,89, *recall* 0,80, dan *F1-score* 0,84. Nilai *precision* yang

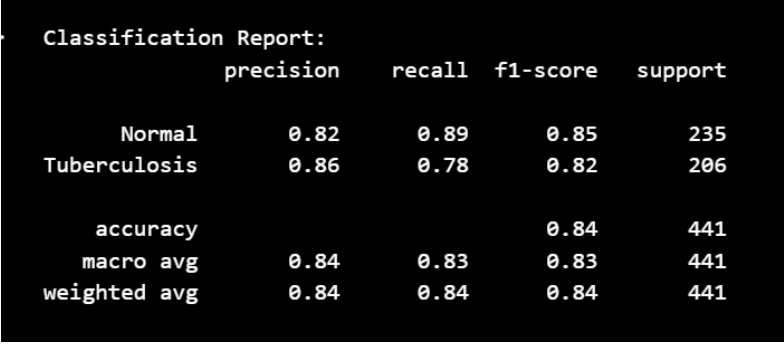
tinggi menunjukkan ketepatan prediksi tuberkulosis yang baik, tetapi *recall* yang lebih rendah mengindikasikan masih adanya citra tuberkulosis yang belum terdeteksi dengan optimal. Secara keseluruhan, hasil ini menggambarkan performa model yang baik, tetapi masih perlu peningkatan dalam mendeteksi citra *Tuberculosis*.



Gambar 3: Confusion Matrix CNN 70:10:20

Sementara itu, berdasarkan *confusion matrix* hasil model CNN 70:10:20 pada Gambar 3, dari 235 citra kelas Normal, sebanyak 209 citra berhasil diklasifikasikan dengan benar, sedangkan 26 citra lainnya salah terdeteksi sebagai *Tuberculosis*. Pada kelas *Tuberculosis*, dari 206 citra, sebanyak 160 citra dikenali dengan benar, sementara 46 citra lainnya salah diklasifikasikan sebagai

Normal. Hasil ini menunjukkan bahwa model memiliki kemampuan yang lebih baik dalam mengenali citra Normal dibandingkan *Tuberculosis*. Jumlah kesalahan yang lebih tinggi pada kelas *Tuberculosis* mengindikasikan bahwa model masih menghadapi kesulitan dalam membedakan pola visual tertentu yang menjadi ciri khas infeksi tuberkulosis pada citra *X-ray* dada.



Classification Report:				
	precision	recall	f1-score	support
Normal	0.82	0.89	0.85	235
Tuberculosis	0.86	0.78	0.82	206
accuracy			0.84	441
macro avg	0.84	0.83	0.83	441
weighted avg	0.84	0.84	0.84	441

Gambar 4: Classification Report CNN 70:10:20

Evaluasi menggunakan *classification report* model CNN 70:10:20 pada Gambar 4 menunjukkan bahwa pada kelas Normal, model menghasilkan *precision* 0,82, *recall* 0,89, dan *F1-score* 0,85. Nilai *recall* yang tinggi menunjukkan sebagian besar citra Normal berhasil dikenali dengan benar, meskipun *precision* yang sedikit lebih rendah menandakan adanya beberapa kesalahan klasifikasi. Sementara itu pada kelas *Tuberculosis*, model memperoleh *precision* 0,86, *recall* 0,78, dan *F1-score* 0,82. Nilai *precision* yang tinggi menunjukkan prediksi tuberkulosis cukup akurat, tetapi *recall* yang lebih rendah menandakan masih ada citra tuberkulosis yang belum terdeteksi dengan baik. Secara keseluruhan, nilai *F1-score* pada kedua kelas menggambarkan performa model yang baik, meskipun diperlukan peningkatan sensitivitas terhadap citra *Tuberculosis*.

Secara umum, hasil dari evaluasi memperlihatkan bahwa model yang menggunakan pembagian data 80:10:10 menunjukkan performa yang lebih baik dibandingkan dengan model yang menggunakan rasio 70:10:20. Model ini menghasilkan akurasi yang lebih tinggi dalam pengujian dan validasi serta memiliki *loss*

yang lebih rendah, yang menandakan kinerja yang lebih efisien dalam mengurangi kesalahan prediksi. Selisih akurasi pengujian sebesar 2,3% di antara keduanya juga cukup signifikan, mengingat peningkatan kecil pada metrik ini mencerminkan peningkatan kemampuan model dalam mendeteksi pola visual secara lebih konsisten pada data yang belum pernah dilihat sebelumnya. Peningkatan tersebut menunjukkan bahwa penambahan 10% pada bagian data latih dapat memperkaya variasi pola yang dipelajari oleh model, menghasilkan kemampuan generalisasi yang lebih baik.

Di samping itu, analisis terhadap hasil *confusion matrix* serta *classification report* pada kedua model menunjukkan pola yang serupa. Model dengan rasio 80:10:10 lebih adil dalam mengevaluasi kedua kelas, sementara model 70:10:20 menunjukkan sensitivitas yang lebih rendah terhadap citra *Tuberculosis*. Ini menunjukkan bahwa distribusi data latih yang lebih besar tidak hanya meningkatkan akurasi total, tetapi juga meningkatkan kestabilan model dalam mendeteksi variasi citra medis yang kompleks.

SIMPULAN

Penelitian ini membuktikan bahwa penerapan *Convolutional Neural Network* (CNN) efektif dalam mengklasifikasikan citra *X-ray* dada untuk mendeteksi penyakit tuberkulosis secara otomatis. Berdasarkan hasil evaluasi, model dengan pembagian data 80:10:10 menunjukkan performa terbaik dengan akurasi pengujian mencapai 85,97% dan nilai *loss* yang lebih rendah dibandingkan model 70:10:20. Hal ini menunjukkan bahwa model 80:10:10 memiliki kemampuan generalisasi yang lebih baik dan lebih efisien dalam meminimalkan kesalahan prediksi.

Analisis *classification report* memperlihatkan bahwa model 80:10:10 lebih baik dalam mengenali citra Normal, sedangkan pada citra *Tuberculosis*, baik model 80:10:10 maupun 70:10:20 masih menunjukkan kesulitan dalam mendeteksi seluruh variasi pola yang terdapat pada paru-paru penderita tuberkulosis. Hal ini menandakan bahwa meskipun performa keseluruhan sudah tergolong baik, kemampuan model dalam mengenali citra *Tuberculosis* masih perlu ditingkatkan agar sensitivitas terhadap karakteristik penyakit dapat lebih optimal.

Temuan ini menekankan pentingnya memilih rasio data yang seimbang dalam pembuatan model CNN, karena distribusi data yang tepat dapat meningkatkan kestabilan dan keandalan hasil pengklasifikasian. Namun, masih ada peluang untuk perbaikan, khususnya dalam memperluas representasi data tuberkulosis serta mengatur parameter model agar lebih responsif terhadap variasi visual yang rumit.

Secara umum, penelitian ini menunjukkan bahwa pendekatan CNN mampu menghasilkan kinerja yang baik dalam klasifikasi citra medis. Perbaikan dapat ditujukan pada proses prapemrosesan citra, augmentasi data, serta pengembangan struktur model yang lebih kompleks untuk meningkatkan ketepatan deteksi tuberkulosis. Oleh karena itu, penelitian ini diharapkan menjadi landasan untuk pengembangan sistem diagnosis berbasis kecerdasan buatan yang lebih akurat, adaptif, dan berguna dalam deteksi dini tuberkulosis.

DAFTAR PUSTAKA

1. World Health Organization. Global Tuberculosis Report 2024 [Internet]. 2024 [cited 2025 Jan 15]. Available from: <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2024>
2. Nyarko RO, Prakash A, Kumar N, Saha P, Kumar R. Tuberculosis a Globalized Disease: Review. *Asian Journal of Pharmaceutical Research and Development*. 2021;9 (1):198–201.
3. Moifo B, Pefura-Yone EW, Nguiefack-Tsague G, Gharingam ML, Tapouh JRM, Kengne AP, et al. Inter-Observer Variability in the Detection and Interpretation of Chest X-Ray Anomalies in Adults in an Endemic Tuberculosis Area. *Open Journal of Medical Imaging*. 2015;5 (3):143–9.
4. Saxena A. An Introduction to Convolutional Neural Networks. *Int J Res Appl Sci Eng Technol*. 2022;10 (12):943–7.
5. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of Deep Learning:

- Concepts, CNN Architectures, Challenges, Applications, Future Directions. *J Big Data*. 2021;8 (1).
6. Andika LA, Pratiwi H, Sulistijowati Handajani S. Convolutional Neural Network Modeling for Classification of Pulmonary Tuberculosis Disease. In: *Journal of Physics: Conference Series*. Institute of Physics Publishing; 2020.
 7. Puttagunta MK, Ravi S. Detection of Tuberculosis Based on Deep Learning Based Methods. In: *Journal of Physics: Conference Series*. IOP Publishing Ltd; 2021.
 8. Guo R, Passi K, Jain CK. Tuberculosis Diagnostics and Localization in Chest X-Rays via Deep Learning Models. *Front Artif Intell*. 2020;3.
 9. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Peter Campbell J. Introduction to Machine Learning, Neural Networks, and Deep Learning. *Transl Vis Sci Technol*. 2020;9 (2).
 10. Wu J. Introduction to Convolutional Neural Networks [Internet]. 2017 [cited 2025 Oct 17]. Available from: <https://cs.nju.edu.cn/wujx/paper/CNN.pdf>
 11. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional Neural Networks: An Overview and Application in Radiology. Vol. 9, *Insights Into Imaging*. Springer Verlag; 2018. p. 611–29.
 12. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*. 2014;15 (1):1929–58.