

IJCIT

(Indonesian Journal on Computer and Information Technology)

Journal Homepage: <http://ejournal.bsi.ac.id/ejurnal/index.php/ijcit>

Diagnosa Tuberculosis Paru Berbasis Citra X-ray Menggunakan Convolutional Neural Network

Saeful Bahri¹, Rusda Wajhillah², Miftah Farid Adiwisastro³

¹Sistem Informasi, Universitas Nusa Mandiri
Jakarta, Indonesia
e-mail: saeful.sel@nusamandiri.ac.id

²Sistem Informasi Akuntansi Kampus Kota Sukabumi, Universitas Bina Sarana Informatika
Sukabumi, Indonesia
e-mail: rusda.rwh@bsi.ac.id

³ Sistem Informasi Kampus Kota Tasikmalaya, Universitas Bina Sarana Informatika
Tasikmalaya, Indonesia
e-mail: miftah.mow@bsi.ac.id

ABSTRAK

Tuberkulosis merupakan sebuah penyakit yang disebabkan oleh Mycobacterium tuberculosis. Penyakit ini dapat menyerang darah, tulang otak dan paru-paru. Diagnosa yang cepat dan akurat sangat diperlukan agar dapat dilakukan pengobatan yang tepat. Diagnosa biasanya dilakukan dengan cara melihat hasil citra x-ray thorax dan hasil test BTA pada pasien. Penelitian ini bertujuan untuk mempercepat identifikasi dalam proses diagnosa dari citra paru yang terinfeksi bakteri tuberkulosis. Diagnosa dilakukan dengan bantuan machine learning berdasarkan hasil citra x-ray menggunakan Algoritma CNN, dengan cara mengklasifikasikan citra x-ray normal dan citra x-ray tuberkulosis. Dalam penelitian ini, dibahas tentang penggunaan citra x-ray paru atau citra thorax untuk mendeteksi diagnosa penyakit paru yang disebabkan oleh microbacterium tuberculosis. Hasil klasifikasi dengan CNN yang didapat setelah proses evaluasi model menunjukkan nilai yang cukup baik yaitu untuk nilai akurasi di kisaran 89%, sementara untuk nilai f1-score 0,89.

Kata Kunci: algoritma CNN, machine learning, tuberkulosis

ABSTRACTS

Tuberculosis is a disease caused by microbacterium tuberculosis. This disease can attack the blood, bones, barain and lungs. A fast and accurate diagnosis is nedded so that appropriate treatment can be carried out. In case tuberculosis, diagnostic is usually done by the result of the chest x-ray image, and the diagnostic result by laboratories BTA test on the patient. This research aims to accelerated the identification in the diagnostic process of lungs image infeted tuberculosis bactria diagnosis is done with the help of machine learning based on the results of x-ray images using the CNN Algorithm, by classifying normal x-ray images and tuberculosis x-ray images. In this study, we discussed the use of the thorax image for diagnostic disease lungs caused by microbacterium tuberculosis. The classification results with CNN obtained after the model evaluation process showed a fairly good value, namely for the accuracy value in the range of 89%, while for the f1-score value of 0,89.

Keywords: CNN algorithm, machine learning, tuberculosis



1. PENDAHULUAN

Tuberculosis (TBC) merupakan salah satu penyakit yang paling berbahaya dan mematikan diseluruh dunia (Oloko-Oba & Viriri, 2020), tbc juga menjadi permasalahan utama khususnya di Indonesia. Secara global, penderita TBC di seluruh dunia mencapai angka 10,4 juta kasus penderita, sedangkan di Indonesia sendiri terdapat hampir sebanyak 550 ribu penderita dan menempatkan Indonesia di urutan ketiga sebagai negara di dunia yang mengidap tuberculosis terbanyak (Kementerian Kesehatan RI, 2018). Tuberculosis merupakan penyakit yang disebabkan oleh mikrobakterium tuberculosis, bakteri ini dapat menyerang darah, tulang, otak dan paru-paru, namun kebanyakan dari bakteri ini menyerang organ paru-paru (Ladumor et al., 2021).

Diagnosa penderita tuberculosis berdasarkan citra x-ray perlu didiagnosa secara akurat sangat penting dalam pencegahan penyakit tersebut (Sathitratanacheewin et al., 2020). Metode-metode dalam klasifikasi citra x-ray paru-paru dalam pengklasifikasian penderita TBC dengan metode machine learning telah banyak digunakan oleh para peneliti di dunia seperti penelitian yang dilakukan oleh (Alawi et al., 2021) (Syafiq et al., 2021) (Lopez-Garnier et al., 2019) yang menerapkan CNN untuk melakukan screening pada penyakit tuberculosis kemudian penelitian yang dilakukan oleh (Higashiguchi et al., 2021) yang melakukan penelitian tentang klasifikasi citra penderita tuberculosis paru dengan model CNN, kemudian penelitian yang dilakukan oleh (Tavolara et al., 2020), yang mengklasifikasikan citra x-ray paru dan mengklasifikasikan jenis penyakit TB yang diderita.

Dibandingkan dengan metode Neural Network, pada umumnya CNN memiliki kelebihan yaitu dapat mengolah data lebih dari satu dimensi sehingga objek tidak akan kehilangan objek spasialnya (Zhong et al., 2021).

Pada penelitian ini akan diterapkan model CNN untuk melakukan klasifikasi citra x-ray paru-paru pada penderita tuberculosis.

Beberapa penelitian terkait penggunaan metode klasifikasi berbasis citra x-ray menggunakan Convolutional Neural Network dan atau penelitian yang berkaitan tentang klasifikasi citra penderita tuberculosis paru yang pernah dilakukan sebelumnya yaitu:

Metode CNN dengan Citra x-ray Grayscale 112x112 pixel dengan hasil dari prediksi dengan

model CNN ditemukan nilai aktual dengan metode korelasi pearson sebesar 0,392, $p=0,002$ rata-rata error 18,0. Dari nilai tersebut CNN mampu mengenali beberapa temuan yang terdapat dalam hasil photo rontgen dada yang dapat digunakan untuk memprediksi durasi yang dibutuhkan untuk mencapai TB negatif. Walaupun hasil dari penelitian tersebut tidak memuaskan namun CNN terbukti dapat digunakan memprediksi durasi pengobatan Tuberculosis (Higashiguchi et al., 2021).

Metode machine learning dan algoritma Convolutional Neural Network dengan hasil diagnosa menggunakan Metode deep learning dan algoritma convolutional neural network terbukti memiliki akurasi yang cukup baik dalam mengenali citra hasil radiologi pada paru yang terinfeksi *Mycobacterium tuberculosis* (Tavolara et al., 2020).

Metode deep learning dengan algoritma convolutional neural network dengan hasil Melakukan deteksi pencemaran air dari citra molekul air, menggunakan 121 neuron yang menghasilkan nilai RMSE pada data testing sebesar 0.30-0.35, dalam klasifikasian citra molekul air (Zhong et al., 2021)

2. METODE PENELITIAN

Penelitian yang kami lakukan adalah dengan model penelitian eksperimental. Model penelitian ini, menekankan pada proses manipulasi yang bertujuan untuk mengetahui akibat dari perlakuan yang dilakukan oleh (Wang et al., 2017), Gambar 1 menyajikan kerangka penelitian yang dilakukan dalam melakukan penelitian ini.

a. Collecting Data

Dalam penelitian ini, data yang digunakan adalah data citra x-ray penderita dengan hasil negatif dan positif tb yang diambil dari kaggle.com. Terdapat sebanyak 3500 total data untuk citra normal 3500 untuk citra yang diklasifikasikan tuberculosis. Data tersebut dibagi kedalam 2 kelompok dengan jumlah pembagian diambil menggunakan metode random sampling, yaitu sebanyak 2000 citra normal dan 2000 citra tuberculosis sebagai data training dan sisanya digunakan sebagai data testing.

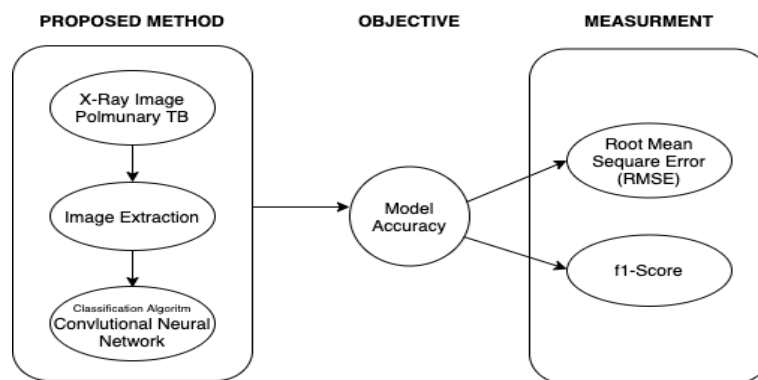
b. Preprocessing Data Citra

Cara CNN dalam melakukan pemrosesan gambar relatif sama dengan dengan MLP, perbedaannya terletak pada presentasi neuron, yang dipresentasikan ke dalam 2 dimensi arsitektur MLP dapat dilihat pada gambar 2.

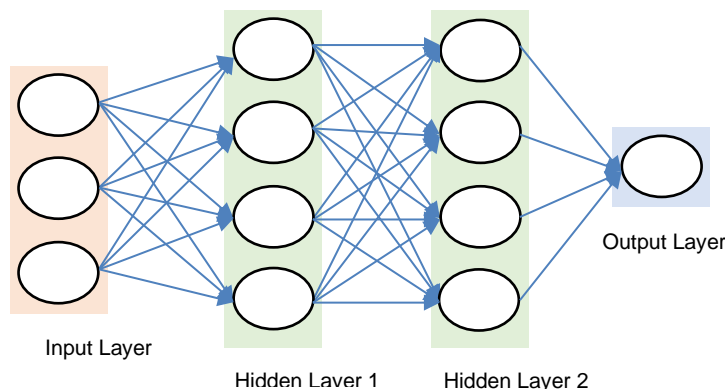
Pada gambar 2, arsitektur MLP terlihat bahwa MLP memiliki 3 layer (kotak merah dan hijau) dengan setiap layer berisi 4 neuron (lingkaran putih). Pada MLP input data yang diterima hanya 1 dimensi kemudian data tersebut dipropagasikan pada jaringan sehingga menghasilkan output layer. Hubungan antar neuron di setiap layer yang berdampingan memiliki bobot satu dimensi yang menemukan kualitas mode. Pada layer input data akan dilakukan operasi linear dengan bobot yang ada, selanjutnya hasil komputasi ditransformasikan menggunakan operasi non-linear yang biasa disebut fungsi aktivasi.

CNN data yang dipropagasi pada jaringan merupakan data dua dimensi, jadi bobot parameter operasi linear antara CNN dan MLP berbeda. CNN operasi linearnya menggunakan operasi konvolusi, sehingga bobot yang digunakan tidak hanya satu dimensi saja, pada gambar 2 akan digambarkan kumpulan kernel konvolusi pada CNN.

Sifat konvolusi maka CNN hanya dapat digunakan pada data yang memiliki struktur dua dimensi seperti citra dan suara, ilustrasi dari konvolusi dapat dilihat pada gambar 3.



Gambar 1. Proposed Method

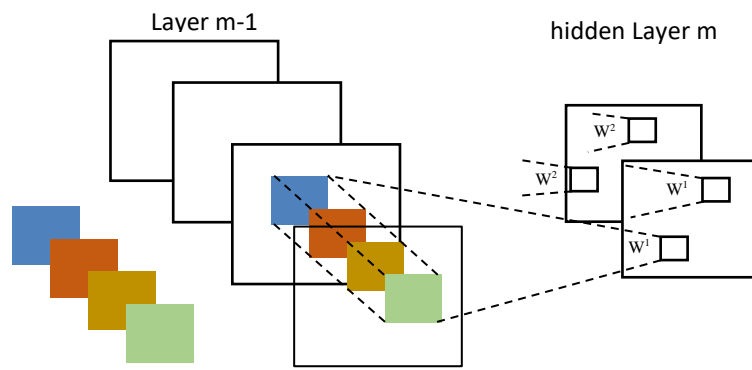


Gambar 2. Arsitektur MLP

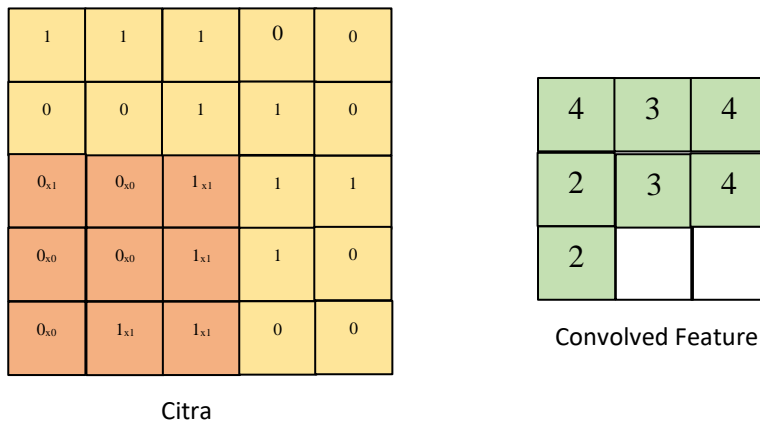
Konvolusi bertujuan untuk melakukan ekstraksi fitur dari citra, hasil dari ekstraksi berupa transformasi linear dari data input yang sesuai informasi data, sehingga kernel konvolusi dapat dilatih oleh CNN ilustrasi dapat dilihat pada gambar 4.

Semua citra akan diubah kedalam ukuran gambar 225 kemudian gambar diubah kedalam dimensi warna grayscale. Lalu gambar dipertajam dengan meningkatkan nilai sharpness, selanjutnya citra diekstraksi ke dalam bentuk

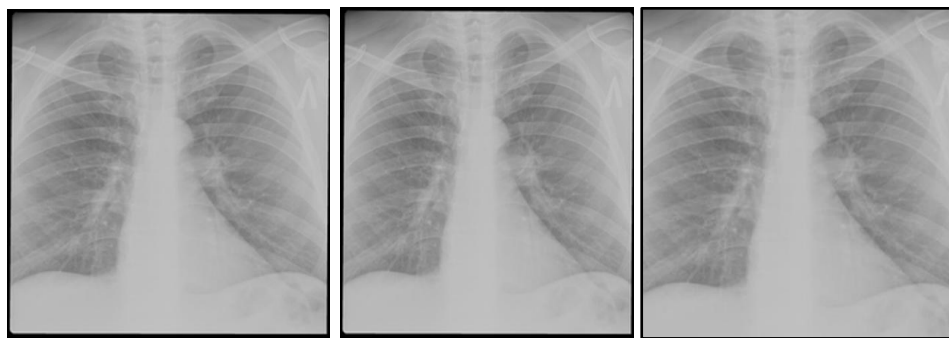
matriks dengan size 24x24 pixel -> 12x12 pixel -> 8x8 pixel -> 4x4 pixel sampai dengan diketahui 1 pixel, selanjutnya setelah gambar selesai diekstraksi maka dilakukan proses klasifikasi selanjutnya dievaluasi, presisi, recall f1-score dan accuracy dari model klasifikasinya. Pada gambar 5 dapat dilihat citra paru yang telah diubah ke dalam format grayscale untuk selanjutnya akan dilakukan prses training dan modeling data menggunakan algoritma *convlutional Neural Network*.



Gambar 3. Ilustrasi Konvolusi



Gambar 4. Ekstraksi fitur dari citra



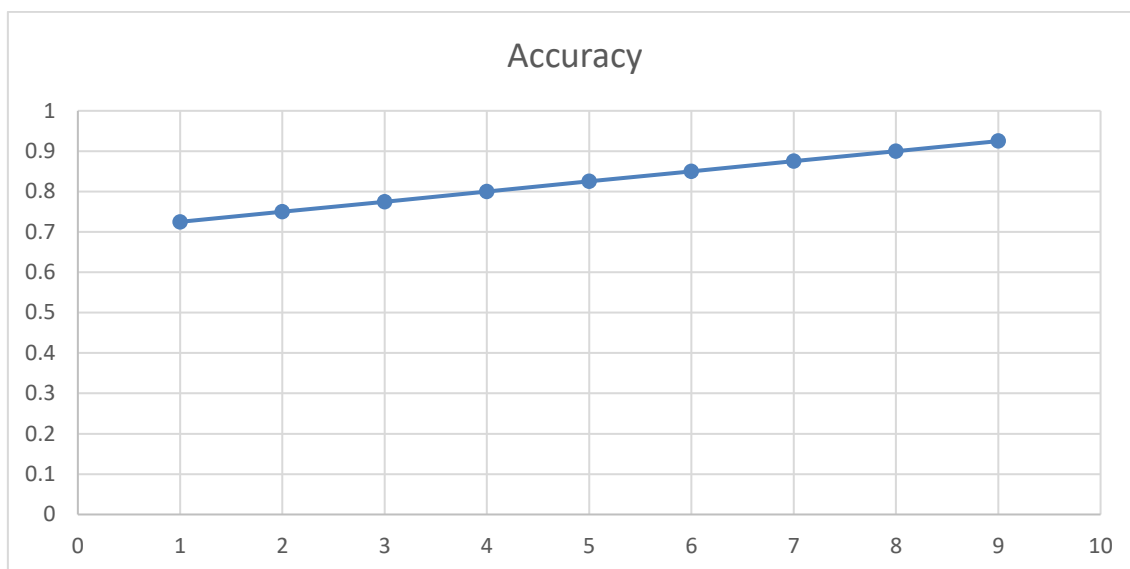
Gambar 5. Citra Dimensi Warna Grayscale

3. HASIL DAN PEMBAHASAN

Berdasarkan data hasil uji yang telah dilakukan, tingkat akurasi dari metode convolution neural network dalam mengklasifikasi penyakit tuberculosis paru dapat mencapai 92,5%. Nilai tersebut diperoleh pada pengulangan ke-9 dari 30 kali uji coba. Hasil akurasi dapat dilihat pada tabel 1. Pada gambar 6 terlihat bahwa dari 9 kali percobaan pengujian didapat akurasi terbaik pada iterasi ke 9 dengan nilai akurasi sebesar 0,925

Tabel 1. Hasil Akurasi

Epoch	Accuracy
1	0,725
2	0,75
3	0,775
4	0,8
5	0,825
6	0,85
7	0,875
8	0,9
9	0,925



Gambar 6. Grafik Akurasi Citra X-Ray CNN Pada TB Paru

Tabel 2. Tingkat Akurasi CNN Klasifikasi TB Paru berbasis Citra X-Ray

Hasil Diagnosa	Precision	Recall	F1-score	support
Normal	0,91	0,92	0,92	244
Tuberculosis	0,85	0,91	0,89	203
Confusion Matrix				
Normal	0,91	0,02		
Tuberculosis	0,89	0,02		

Pada tabel 2 dapat dilihat nilai confusion matrix yang didapat nilai precision dan nilai recall dari hasil normal dan recall tuberculosis lebih tinggi maka dapat dilihat hasil dari klasifikasi dengan CNN cukup baik. Hal itu terbukti dari cukup baiknya nilai F1 score sebesar 0,92 yang mengindikasikan bahwa model memiliki precision dan recall yang baik.

4. KESIMPULAN

Dari penelitian experimental tentang model CNN untuk melakukan klasifikasi citra x-ray paru-paru pada penderita tuberculosis dihasilkan bahwa CNN mampu meningkatkan hasil klasifikasi dengan tingkat akurasi dapat mencapai 92,5%. Nilai tersebut diperoleh pada pengulangan ke-9 dari 30 kali uji coba.

5. REFERENSI

Alawi, A. E. B., Al-Basser, A., Sallam, A., Al-Sabaei, A., & Al-Khateeb, H. (2021). Convolutional Neural Networks Model for Screening Tuberculosis Disease. *2021 International Conference of Technology, Science and Administration, ICTSA 2021*.

<https://doi.org/10.1109/ICTSA52017.2021.9406520>

Higashiguchi, M., Nishioka, K., Kimura, H., & Matsumoto, T. (2021). Prediction of the duration needed to achieve culture negativity in patients with active pulmonary tuberculosis using convolutional neural networks and chest radiography. *Respiratory Investigation*, *59*(4), 421–427. <https://doi.org/10.1016/j.resinv.2021.01.004>

Kementerian Kesehatan RI. (2018). Info Datin TBC. In *Kemendes RI* (Vol. 67, Issue 1).

Ladumor, H., Al-Mohannadi, S., Ameerudeen, F. S., Ladumor, S., & Fadl, S. (2021). TB or not TB: A comprehensive review of imaging manifestations of abdominal tuberculosis and its mimics. *Clinical Imaging*, *76*, 130–143. <https://doi.org/10.1016/j.clinimag.2021.02.012>

Lopez-Garnier, S., Sheen, P., & Zimic, M. (2019). Automatic diagnostics of tuberculosis using convolutional neural networks analysis of MODS digital images. *PLoS*

- ONE, 14(2).
<https://doi.org/10.1371/journal.pone.0212094>
- Oloko-Oba, M., & Viriri, S. (2020). Diagnosing tuberculosis using deep convolutional neural network. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12119 LNCS, 151–161. https://doi.org/10.1007/978-3-030-51935-3_16
- Sathitratanacheewin, S., Sunanta, P., & Pongpirul, K. (2020). Deep learning for automated classification of tuberculosis-related chest X-Ray: dataset distribution shift limits diagnostic performance generalizability. *Heliyon*, 6(8). <https://doi.org/10.1016/j.heliyon.2020.e04614>
- Syafiq, M., Saaidi, I., Yusoff, M., Sadikin, A., Afendi, M., & Hassan, A. M. (2021). *Tuberculosis X-Ray Images Classification based Dynamic Update Particle Swarm Optimization with CNN*. 48(9).
- Tavolara, T. E., Niazi, M. K. K., Ginese, M., Piedra-Mora, C., Gatti, D. M., Beamer, G., & Gurcan, M. N. (2020). Automatic discovery of clinically interpretable imaging biomarkers for Mycobacterium tuberculosis supersusceptibility using deep learning. *EBioMedicine*, 62. <https://doi.org/10.1016/j.ebiom.2020.103094>
- Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 3462–3471. <https://doi.org/10.1109/CVPR.2017.369>
- Zhong, S., Hu, J., Yu, X., & Zhang, H. (2021). Molecular image-convolutional neural network (CNN) assisted QSAR models for predicting contaminant reactivity toward OH radicals: Transfer learning, data augmentation and model interpretation. *Chemical Engineering Journal*, 408. <https://doi.org/10.1016/j.cej.2020.127998>