

IMPLEMENTASI MODEL *CONVOLUTIONAL NEURAL NETWORK* (CNN) UNTUK KLASIFIKASI PENYAKIT TBC BERBASIS DEKSTOP

¹ Suharni
² Eel Susilowati
³ Taufiq Muldan Hidayat

¹Universitas Gunadarma, harni@staff.gunadarma.ac.id

²Universitas Gunadarma, eel@staff.gunadarma.ac.id

³Universitas Gunadarma, muldantaufiq@gmail.com

ABSTRAK

Penyakit Tuberkulosis (TBC) merupakan salah satu penyakit yang masih menjadi masalah utama di negara berkembang seperti Indonesia. Tuberkulosis (TBC) adalah penyakit yang disebabkan oleh bakteri Mycobacterium Tuberculosis. Tuberkulosis menyerang paru-paru dan dapat menginfeksi orang lain. Penyakit ini bersifat menular. Radiologi termasuk salah satu pemeriksaan diagnosis yang dapat dilakukan pada penyakit TB-paru. Dengan menganalisa citra x-ray hasil rontgen pada dada tubuh seseorang, maka para dokter dapat melakukan proses diagnosa penyakit TBC. Permasalahan yang terjadi adalah diagnosa penyakit TBC masih dilakukan secara manual oleh dokter. Tujuan penelitian ini adalah mengimplementasikan model Convolutional Neural Network (CNN) untuk klasifikasi penyakit TBC berbasis dektop. Untuk mengoptimalkan model tersebut menggunakan optimizer Stochastic Gradient Descent (SGD). Model CNN optimal yang diimplementasikan bisa membantu dokter dalam proses identifikasi hasil rontgen menggunakan proses komputerisasi. Hasil identifikasi ini dapat mengklasifikasi apakah seseorang mempunyai penyakit tbc atau tidak. Dataset didapat dari website kaggle.com berjumlah 240 data yang terdiri dari 160 data latih dan atau 80 data uji. Citra yang digunakan berukuran 128x128 piksel. Pelatihan model menggunakan 50 epochs dan learning rate 0.01. Hasil pelatihan tersebut mendapatkan nilai akurasi sebesar 98.75%. Implementasi model CNN pada uji coba aplikasi menggunakan 15 citra x-ray yang terdiri dari 8 citra x-ray normal dan 7 citra x-ray TBC menunjukkan bahwa semua prediksi berdasarkan sistem dan aktual hasilnya sesuai.

Kata Kunci: TBC, Convolutional Neural Network, Citra X-ray.

PENDAHULUAN

Penyakit Tuberkulosis (TBC) merupakan salah satu penyakit yang masih menjadi masalah utama di negara berkembang seperti Indonesia. Tuberkulosis (TBC) adalah penyakit yang disebabkan oleh bakteri Mycobacterium Tuberculosis (Zimmermann dan Hammond, 2017). Tuberkulosis menyerang paru-paru dan dapat menginfeksi orang lain. Penyakit ini dapat menular melalui kontak langsung dengan penderita TBC, lalu

melalui udara seperti batuk dan bersin. Radiologi termasuk salah satu pemeriksaan diagnosis yang dapat dilakukan pada penyakit TB-paru. Dengan menganalisa citra x-ray hasil rontgen pada dada tubuh seseorang, maka para dokter dapat melakukan proses diagnosa penyakit TBC. Proses yang dibutuhkan harus cepat dan akurat. Selama ini dalam melakukan diagnosa hasil rontgen penyakit TBC masih dilakukan secara manual oleh dokter, sehingga kecermatan mata

dalam mendiagnosa hasil rontgen sangat diperlukan. Keterbatasan kemampuan manusia dalam menginterpretasikan hasil diagnosa secara visual bisa mengakibatkan kesalahan.

Penelitian-penelitian sebelumnya yang terkait dengan diagnosa penyakit paru sudah banyak dilakukan, namun masih terdapat keterbatasan dalam hasil tingkat akurasi. Penelitian yang dilakukan oleh Abdul Hamid (2019) yaitu “Klasifikasi Penyakit Tuberculosis Dan Pneumonia Pada Paru-Paru Manusia Berdasarkan Citra *Chest X-Ray* Menggunakan *Convolutional Neural Network*”. Pemilihan model terbaik dipilih berdasarkan nilai *AUC (Area Under The Curve)*. Hasil Penelitian menunjukkan model yang dibuat memperoleh tingkat akurasi mencapai 85.96% menggunakan epochs 50 dan skenario data 90:10.

Penelitian selanjutnya yaitu “Klasifikasi *Chest X-Ray Images* Berdasarkan Kriteria Gejala Covid-19 Menggunakan *Convolutional Neural Network*”. Penelitian ini dilakukan untuk melakukan klasifikasi penyakit covid-19. Penelitian ini menggunakan dataset terdiri dari 2 kelas yaitu kasus positif COVID-19 dengan jumlah citra 1200 data dan normal dengan jumlah citra sebanyak 1341 data. Dataset dibagi menjadi 80% data latih dan 20% data uji. Epochs yang digunakan sebanyak 50. Hasil evaluasi menunjukkan model yang digunakan dapat menghasilkan akurasi 96% (Vina dan Ida, 2021).

Dengan masih terdapat keterbatasan kemampuan manusia dalam mengklasifikasi hasil diagnosa penyakit paru dan dari hasil penelitian sebelumnya maka penting adanya sebuah sistem yang mampu untuk mengklasifikasi hasil diagnosa tersebut. Pada penelitian ini akan mengimplementasikan model

Convolutional Neural Network (CNN) berbasis *desktop* untuk mengklasifikasi penyakit TBC yang bertujuan untuk membantu dokter dalam proses identifikasi hasil rontgen. Hasil identifikasi ini dapat mengklasifikasi apakah seseorang mempunyai penyakit TBC atau tidak.

METODE PENELITIAN

Metode yang digunakan untuk melakukan klasifikasi penyakit TBC pada paru-paru manusia menggunakan teknologi *Deep Learning* dengan model *Convolutional Neural Network (CNN)*. Metode CNN merupakan salah satu metode yang sedang berkembang saat ini. Pada metode ini menggunakan masukan berupa gambar. Kemudian melalui lapisan konvolusi, gambar tersebut akan diolah berdasarkan filter yang ditentukan. Setiap lapisan ini menghasilkan pola dari beberapa bagian citra yang memudahkan proses klasifikasi (Li, W., Zhao, R., Xiao, T., dan Wang, X, 2014).

Pada bagian ini akan menjelaskan tahapan-tahapan dari proses klasifikasi penyakit TBC dengan metode CNN. Tahap pertama akuisisi citra atau pengumpulan dataset, tahap kedua preprocessing, tahap ketiga membuat arsitektur model *convolutional neural network (CNN)*, tahap keempat melakukan pelatihan dataset, dan tahap kelima melakukan testing program. Tahapan-tahapan tersebut disajikan pada gambar 1.

Tahap Pengambilan Citra

Tahap awal dimulai dengan proses pengambilan data. Data yang dikumpulkan berupa data citra paru *grayscale* yang terdiri dari 2 jenis yaitu, citra *paru yang normal* dan citra paru yang terkena TBC. Data citra paru diambil dari *website kaggle.com*. Seluruh data pada penelitian ini berjumlah 240 data dengan 120 data citra paru normal dan 120 data citra

paru terkena TBC. Data yang diperoleh disimpan pada penyimpanan lokal komputer dan diunggah pada google drive dan dimasukkan ke dalam folder dengan nama tbc-xray dan sub foldernya dengan nama folder normal dan tbc. Proses pengambilan citra paru disajikan pada gambar 2.

Tahap Preprocessing

Pada tahap ini, proses yang dilakukan adalah memisahkan data citra paru untuk data latih dan data uji. Sebelum memisahkan data, proses yang dilakukan adalah mengubah ukuran citra dan type data citra. Langkah-langkah pada tahap *Preprocessing* disajikan pada gambar 3.

Proses awal tahap *preprocessing* adalah membaca dataset yang telah disiapkan pada *google drive* dengan *library opencv*. Setelah dataset dibaca dilakukan *resize* gambar agar semua resolusi pada data gambar menjadi ukuran yang sama, dan mempercepat proses *training*. Resolusi yang digunakan adalah 128x128 piksel. Selanjutnya adalah membagi dataset yang digunakan untuk data latih sebanyak 80 dan 60 untuk data uji. Kemudian tipe data diubah menjadi float32 pada *x_train* dan *x_test*. Langkah terakhir pada tahap *preprocessing* adalah mengubah *y_train* dan *y_test* menjadi *categorical*.

Architectural Modeling

Tahap Arctitrcutral Model dibuat perancangan arsitektur *Convolutional Neural Network* (CNN) menggunakan *library Keras-Python*. Arsitektur dibuat secara bertahap menggunakan metode sekuensial.

Convolution layer berfungsi untuk mengekstrak filter dari citra *input* dan menghasilkan peta 2D atau peta fitur (*feature map*). Filter yang terkandung dalam lapisan lilitan memiliki panjang, tinggi, dan ketebalan sesuai dengan saluran *input* data. Proses *convolution*

menggunakan filter pada layer untuk menghasikan *feature map*. Pemrosesan layer pertama Conv2D dengan *kernel_size(3,3)*, fungsi aktivasi ReLu, dan *input_shape(128,128,3)*, kemudian akan di proses kembali layer kedua Conv2D dengan *kernel_size(3,3)*. Citra tersebut menggunakan 32 filter acak dari program. Filter digunakan untuk menentukan pola apa yang akan dilakukan selanjutnya, dimana nilai matriks yang ada bergantung pada pola tersebut. Tujuan dilakukannya konvolusi pada data citra adalah untuk meng-ekstraksi fitur dari citra input. Konvolusi akan menghasilkan transformasi linear dari data input sesuai informasi spasial pada data. Bobot pada layer tersebut menspesifikasikan kernel konvolusi yang digunakan, sehingga kernel konvolusi dapat dilatih berdasarkan input pada CNN (I Wayan Suartika E. P, Arya Yudhi Wijaya, dan Rully Soelaiman., 2016).

Pooling layer berfungsi untuk mengurangi masukan secara spasial atau mengurangi jumlah parameter. Pooling layer yang digunakan adalah *MaxPooling*. Max pooling membagi output dari convolution layer menjadi beberapa grid kecil lalu mengambil nilai maksimal dari setiap grid untuk menyusun matriks citra yang telah direduksi. Prosesnya dilakukan dengan mengambil nilai piksel terbesar dari dimensi yang ditentukan. *Max pooling* yang digunakan adalah 2x2 dan stride 2.

Dropout adalah teknik regulasi jaringan syaraf dimana beberapa neuron akan dipilih secara acak dan tidak dipakai selama pelatihan (I Wayan Suartika E. P, Arya Yudhi Wijaya, dan Rully Soelaiman., 2016). Dropout merupakan proses mencegah terjadinya *overfitting* dan juga mempercepat proses *learning* dengan menghapus *neuron* secara acak sebesar 0.25 pada proses pertama dan proses *dropout*

kedua menghapus *neuron* secara acak sebesar 0.5.

Flattening adalah merubah matriks yang ada di pooling layer menjadi vektor tunggal. Pooling masih berupa array 2 dimensi, kemudian diubah menjadi data satu dimensi single vector (I Wayan Suartika E. P, Arya Yudhi Wijaya, dan Rully Soelaiman., 2016).

Fully – Connected Layer adalah lapisan dimana semua neuron aktivasi dari lapisan sebelumnya terhubung semua dengan neuron di lapisan selanjutnya seperti halnya jaringan saraf tiruan biasa. Setiap aktivasi dari lapisan sebelumnya perlu diubah menjadi data satu dimensi sebelum dapat dihubungkan ke semua neuron di *fully connected layer* (I Wayan Suartika E. P, Arya Yudhi Wijaya, dan Rully Soelaiman, 2016). *Feature map* yang dihasilkan dari *feature extraction layer* masih berbentuk multidimensional array, sehingga harus melakukan “flatten” atau *reshape feature map* menjadi sebuah vektor agar bisa digunakan sebagai input dari *fully-connected layer*.

Softmax adalah sebuah fungsi aktivasi yang di gunakan untuk permasalahan klasifikasi, fungsi aktivasi ini digunakan pada output layer. Fungsi softmax biasa digunakan pada layer terakhir dari Neural Network untuk mengklasifikasi atau mengelompokkan lebih dari 2 kelas (Yin, X., Goudriaan, J., Lantinga, E. A, Vos, L, dan Spiertz, H. J, 2003).

Training Dataset

Training dataset adalah sebuah proses pembuatan model yang digunakan untuk membuat prediksi dan menjalankan fungsi dari algoritma *convolutional neural network*. Parameter yang digunakan untuk mengoptimalkan model dari algoritma tersebut adalah optimizer *Stochastic*

Gradient Descent (SGD), epochs 50, dan learning rate 0.01.

Testing Program

Testing program dilakukan setelah proses *training dataset*. *Testing program* ini berfungsi untuk mengimplementasikan ke dalam program unit prediksi dari hasil yang sudah didapatkan pada proses *training* serta mengevaluasi apakah model yang dibuat untuk mengklasifikasikan penyakit TBC pada paru-paru dapat berjalan dengan baik atau tidak. Akurasi yang didapatkan sebesar 98.75%. Nilai akurasi sudah sangat baik untuk dapat mengklasifikasikan antara penyakit TBC atau Normal.

Testing ini dilakukan dengan menggunakan metode *Confusion Matrix* untuk memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh sistem model dengan hasil klasifikasi sebenarnya. *Confusion Matrix* dapat diartikan sebagai suatu alat yang memiliki fungsi untuk melakukan analisis apakah classifier tersebut baik dalam mengenali tuple dari kelas yang berbeda. Pada Confusion Matrix terdapat 4 kombinasi nilai prediksi dan nilai aktual yang berbeda. Nilai dari *TruePositive* dan *True-Negative* memberikan informasi ketika classifier dalam melakukan klasifikasi data bernilai benar, sedangkan *FalsePositive* dan *False-Negative* memberikan informasi ketika classifier salah dalam melakukan klasifikasi data. Kombinasi nilai dari confusion Matrix dapat dilihat pada Tabel 1.

Terdapat 4 istilah sebagai representasi hasil proses klasifikasi pada confusion matrix yaitu:

True Positive (TP)

Merupakan data positif yang diprediksi benar. Contohnya, hasil rontgen pasien menderita TBC (class 0) dan dari model yang dibuat

memprediksi pasien tersebut menderita TBC (class 0).

True Negative (TN)

Merupakan data negatif yang diprediksi benar. Contohnya, hasil rontgen pasien tidak menderita TBC (class 1) dan dari model yang dibuat memprediksi pasien tersebut tidak menderita TBC (class 1).

False Postive (FP)

Merupakan data negatif namun diprediksi sebagai data positif. Contohnya, hasil rontgen pasien tidak menderita TBC (class 1) tetapi dari model yang telah memprediksi pasien tersebut menderita TBC(class 0).

False Negative (FN)

Merupakan data positif namun diprediksi sebagai data negatif. Contohnya, hasil rontgen pasien menderita TBC (class 0) tetapi dari model yang dibuat memprediksi pasien tersebut tidak menderita TBC (class 1). Berikut confusion matrix yang telah dibuat terdapat pada gambar 6.

Perancangan Tampilan Aplikasi

Tampilan aplikasi terdiri dari tampilan utama dan tampilan hasil Pada tampilan aplikasi terdapat beberapa komponen yang digunakan yaitu canvas yang berfungsi untuk menampilkan gambar yang telah dipilih untuk pendeteksian penyakit TBC atau Normal. Label berfungsi untuk menunjukkan hasil prediksi penyakit TBC. *Button Select & Process* berfungsi untuk memilih dan

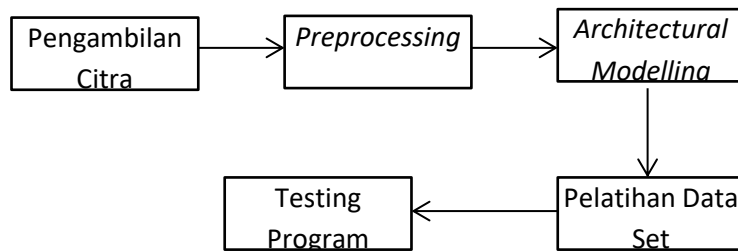
memproses gambar yang akan di prediksi, dan *button Clear* berfungsi untuk membersihkan canvas dan label hasil prediksi. Tampilan Utama aplikasi dapat dilihat pada gambar 7, sedangkan untuk tampilan hasil prediksi disajikan pada gambar 8.

HASIL DAN PEMBAHASAN

Implementasi Model *Convolutional Neural Network (CNN)* untuk Klasifikasi Penyakit TBC Berbasis Dekstop merupakan salah satu model yang digunakan untuk mengidentifikasi seseorang mempunyai penyakit TBC atau tidak. Untuk mengetahui hasil dari implementasi model tersebut perlu dilakukan uji coba aplikasi.

Uji Coba Aplikasi

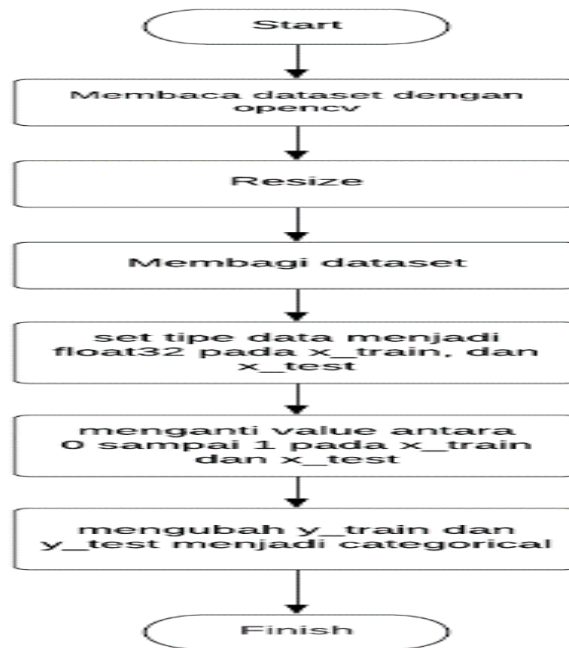
Uji coba aplikasi dari model CNN yang diimplementasikan menggunakan data aktual sebanyak 15 citra x-ray yang terdiri dari 8 citra x-ray normal dan 7 citra x-ray TBC. Hasil uji coba aplikasi disajikan pada tabel 2. Pada Tabel 2, kolom 2 adalah citra x-ray dari data aktual hasil rontgen paru-paru manusia. Pada kolom 3 adalah hasil dari sistem yang memprediksi dari citra x-ray dengan model CNN yang diimplementasikan. Sedangkan pada kolom 4 adalah hasil nyata (aktual) dari citra x-ray. Dari kolom 3 dan 4 bahwa secara sistem maupun secara aktual dengan jumlah data sebanyak 15, hasilnya adalah sesuai dan diterangkan pada kolom 5.



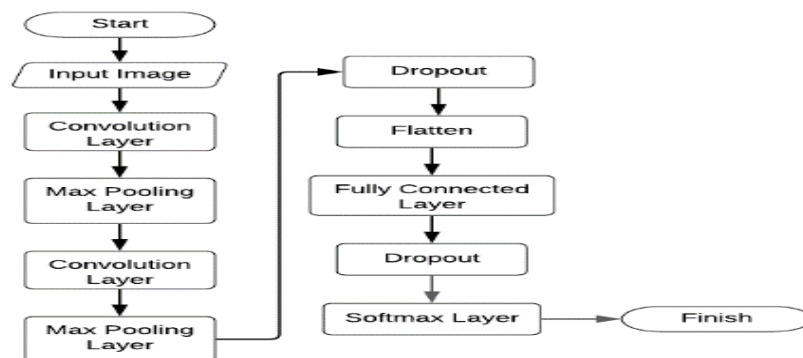
Gambar 1. Tahapan Umum Proses Klasifikasi Penyakit TBC



Gambar 2. Proses Pengambilan Citra Paru



Gambar 3. Tahap Preprocessing



Gambar 4. Arsitektur CNN

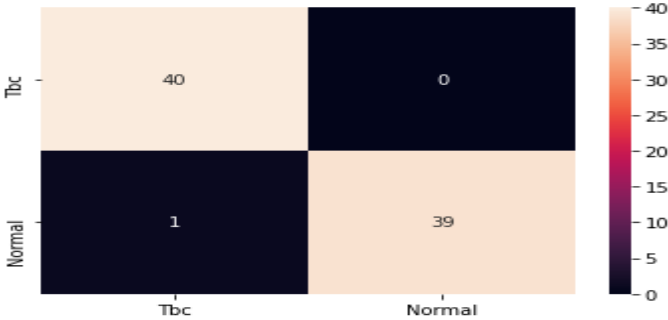
```
0.9875
array([[40, 0],
       [ 1, 39]])
```

Gambar 5. Hasil dari nilai akurasi

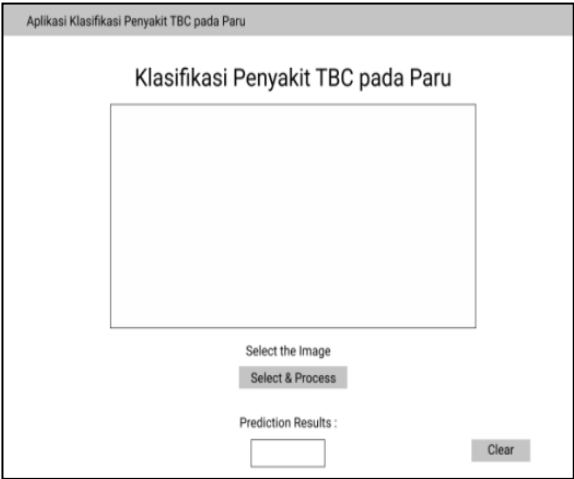
Tabel 1.
Model Confusion Matrix

Nilai prediksi	Nilai sebenarnya	
	True	False
True	TP (true positif)	FP (false negative)
False	FN (false negative)	TN (true negative)

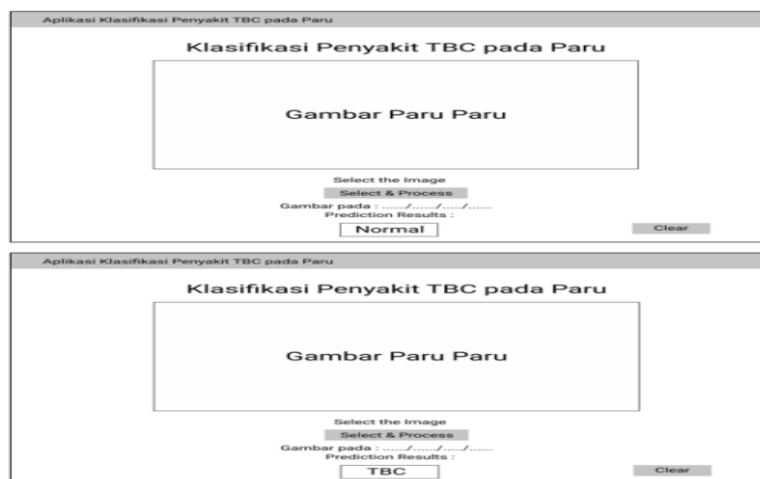
Sumber: J. Han dan M. Kamber (2006).



Gambar 6. Hasil dari Confusion Matrix Akurasi
















Gambar 7. Tampilan Awal Aplikasi





Gambar 8. Tampilan Hasil Prediksi Matrix Akurasi

**Tabel 2.
Hasil Uji Coba Deteksi Penyakit TBC**

No	Gambar	Sistem	Aktual	Keterangan
1		Normal	Normal	Sesuai
2		TBC	TBC	Sesuai
3		TBC	TBC	Sesuai
4		Normal	Normal	Sesuai
5		Normal	Normal	Sesuai
6		Normal	Normal	Sesuai
7		TBC	TBC	Sesuai
8		TBC	TBC	Sesuai
9		TBC	TBC	Sesuai
10		Normal	Normal	Sesuai
11		Normal	Normal	Sesuai
12		TBC	TBC	Sesuai
13		Normal	Normal	Sesuai

Tabel 2 lanjutan

14		Normal	Normal	Sesuai
15		TBC	TBC	Sesuai

KESIMPULAN DAN SARAN

Berdasarkan hasil dari penelitian yang telah dilakukan, implementasi model *Convolutional Neural Network* untuk klasifikasi penyakit TBC menunjukkan bahwa hasil pengujian model dengan menggunakan 240 data yang terdiri dari 160 data latih dan 80 data uji dengan 50 epochs mendapatkan akurasi hasil sebesar 98.75%. Uji coba aplikasi dari model CNN yang diimplementasikan menggunakan 15 citra x-ray yang terdiri dari 8 citra x-ray normal dan 7 citra x-ray TBC memperoleh hasil bahwa semua prediksi menunjukkan hasil yang sesuai. Dalam penelitian ini dataset yang digunakan masih sedikit.

Untuk itu disarankan pada pengembangan selanjutnya dapat menambahkan data set citra x-ray TBC dan citra x-ray Normal agar mendapatkan hasil yang lebih maksimal.

DAFTAR PUSTAKA

- Abdul Hamid. (2019). *Klasifikasi Penyakit Tuberculosis Dan Pneumonia Pada Paru-Paru Manusia Berdasarkan Citra Chest X-Ray Menggunakan Convolutional Neural Network*. Tugas Akhir, Program Studi Matematika, Fakultas Sains dan Teknologi, UIN Syarif Hidayatullah Jakarta.
- I Wayan Suartika E. P, Arya Yudhi Wijaya, dan Rully Soelaiman. (2016). *Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101*, Jurnal Teknik ITS Vol. 5, No. 1.
- J. Han and M. Kamber. (2006). *Data Mining: Concepts and Techniques* (2nd Edition). San Francisco: Elsevier.
- Li, W., Zhao, R., Xiao, T., dan Wang, X. (2014). *Deepreid: Deep filter pairing neural network for person re-identification*. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- P. G. Zimmermann dan B. B. Hammond, (2017), *Sheehy's Emergency and Disaster Nursing - 1st Indonesian Edition*, Elsevier Health Sciences.
- Vina A dan Ida N. (2021). *Klasifikasi Chest X-Ray Images Berdasarkan Kriteria Gejala Covid-19 Menggunakan Convolutional Neural Network*, JSAI: Journal Scientific and Applied Informatics Vol. 4, No. 2.
- Xinyou Yin., Jan Goudriaan, Egbert A Lantinga, Jan Vos, dan Huub J Spiertz. (2003). *A flexible sigmoid function of determinate growth*. *Annals of Botany (Lond)* , 2003 May;91(6):753. Doi: 10.1093/aob/mcg029.