

**PENERAPAN CONVOLUTIONAL NEURAL NETWORK
UNTUK DIAGNOSA PENYAKIT TBC**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

SIDIQ AFRIADI PUJANTORO

20.11.3692

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**PENERAPAN CONVOLUTIONAL NEURAL NETWORK
UNTUK DIAGNOSA PENYAKIT TBC**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

SIDIQ AFRIADI PUJANTORO

20.11.3692

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA**

YOGYAKARTA

2024

HALAMAN PERSETUJUAN


SKRIPSI

**PENERAPAN CONVOLUTIONAL NEURAL NETWORK UNTUK
DIAGNOSA PENYAKIT TBC**

yang disusun dan diajukan oleh
SIDIQ AFRIADI PUJANTORO
20.11.3692

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 September 2024

Dosen Pembimbing,


Mulia Sulistivono, S.Kom., M. Kom
NIK. 190302248

HALAMAN PENGESAHAN

SKRIPSI

**PENERAPAN CONVOLUTIONAL NEURAL NETWORK UNTUK
DIAGNOSA PENYAKIT TBC**

yang disusun dan diajukan oleh
SIDIQ AFRIADI PUJANTORO
20.11.3692

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 September 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Nuri Cahyono, S.Kom.,M.Kom
NIK. 190302278

Dina Maulina, S.Kom.,M.Kom
NIK. 190302250

Mulia Sulistiyono, S.Kom.,M.Kom
NIK. 190302248

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 September 2024

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Sidiq Afriadi Pujantoro
NIM : 20.11.3692

Menyatakan bahwa Skripsi dengan judul berikut:

PENERAPAN CONVOLUTIONAL NEURAL NETWORK UNTUK DIAGNOSA PENYAKIT TBC

Dosen Pembimbing: Mulia Sulistiyono, S.Kom.,M. Kom

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 23 September 2024

Yang Menyatakan,



Sidiq Afriadi Pujantoro

KATA PENGANTAR

Segala puji bagi Allah SWT atas nikmat, rahmat, dan karunia- Nya yang telah memperlancar peneliti menyelesaikan penyusunan skripsi ini dengan judul "PENERAPAN CONVOLUTIONAL NEURAL NETWORK UNTUK DIAGNOSA PENYAKIT TBC ". Skripsi ini merupakan salah satu tahapan penting untuk memenuhi persyaratan kelulusan program Studi Informatika Fakultas Ilmu Komputer di Universitas Amikom Yogyakarta. Penyusunan skripsi ini tidak lepas dari doa, support dan bimbingan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terimakasih yang sebesar-besarnya kepada:

1. Keluarga tersayang Bapak, Ibu, dan kedua adik penulis, yang senantiasa memberikan doa terbaik, dukungan moral, materil, serta semangat dalam perjalanan penyusunan skripsi ini.
2. Bapak Mulia Sulistiyono, S.Kom.,M. Kom, selaku dosen pembimbing dan dewan penguji, atas dukungan, arahan, serta masukan yang telah membimbing dalam penyusunan skripsi ini.
3. Bapak Nuri Cahyono, S.Kom.,M.Kom dan Ibu Dina Maulina, S.Kom.,M.Kom selaku dewan penguji yang telah meluangkan waktu nya.
4. Tiara Dwina Lestari yang selalu mendoakan, memberikan dukungan, nasihat serta menemani dan membantu penulis dalam mengerjakan skripsi maupun dalam keadaan apapun.
5. Sahabat BAROKAH, yang telah berbagi ilmu, pengalaman, dan inspirasi selama mengemban pendidikan di Universitas Amikom Yogyakarta.
6. Semua individu yang turut menyumbangkan gagasan, pandangan, dan sokongan dalam berbagai bentuk, yang telah membantu kelancaran perjalanan penyelesaian skripsi ini.

Yogyakarta, 23 September 2024

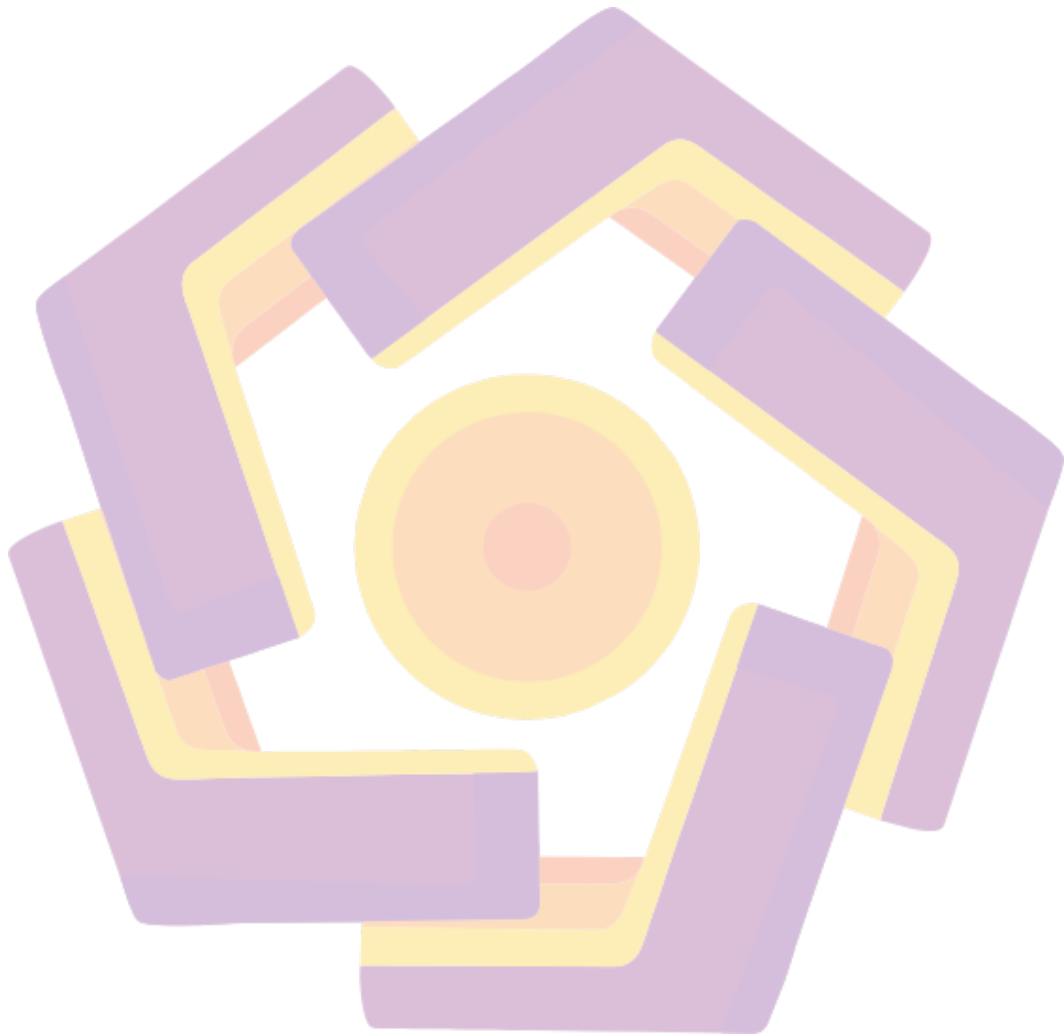
Penulis

DAFTAR ISI

COVER	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	vii
DAFTAR GAMBAR	viii
DAFTAR LAMBANG DAN SINGKATAN	ix
DAFTAR ISTILAH	x
INTISARI	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori.....	13
BAB III METODE PENELITIAN	17
3.1 Objek Penelitian.....	17
3.2 Metode Penelitian	17
BAB IV HASIL DAN PEMBAHASAN	20
4.1 Pengumpulan Dataset.....	20
4.2 Preprocessing Data.....	20
4.3 Segmentasi	25
4.4 Identifikasi	26
4.5 Uji Peforma.....	29
4.6 Interpretasi Hasil.....	40
BAB V PENUUP	44
5.1 Kesimpulan	44
5.2 Saran	44
REFERENSI	

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	10
Tabel 4.1 Hasil Preprocessing Data.....	24
Tabel 4.2 Hasil dari pengujian Dataset	40

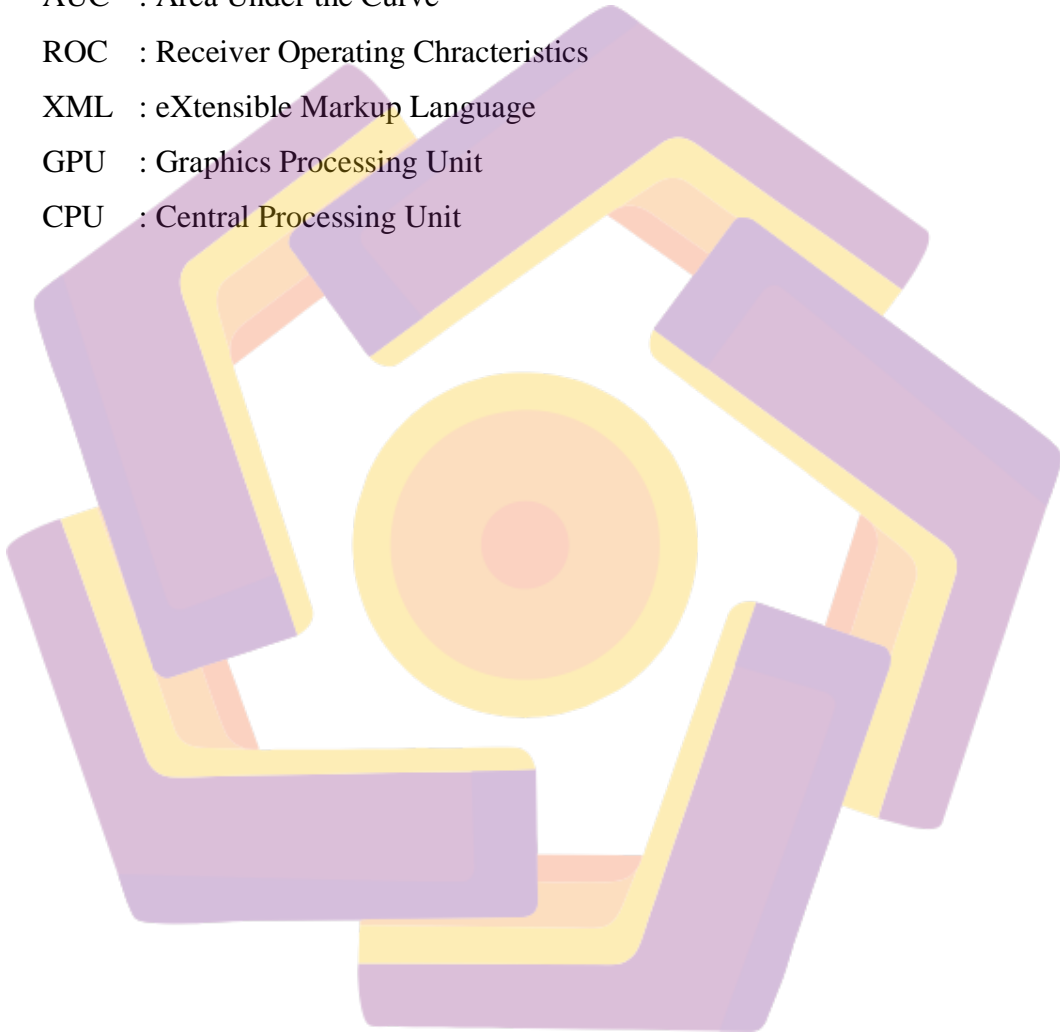


DAFTAR GAMBAR

Gambar 2.1 Cara Kerja Algoritma CNN	14
Gambar 2.2 Confusion Matrix	15
Gambar 3.1 Alur Penelitian	18
Gambar 3.2 Tampilan Dataset	19
Gambar 4.1 Gambar Dataset Tuberculosis	20
Gambar 4.2 Proses Input Dataset	20
Gambar 4.3 Proses Visualisasi Dataset	22
Gambar 4.4 Proses Pemisahan Gambar File xml	23
Gambar 4.5 Proses Ekstraksi Bounding Box	24
Gambar 4.6 Proses Penyusunan Anotasi	25
Gambar 4.7 Hasil dari Proses Segmentasi	26
Gambar 4.8 Pemrosesan Gambar dan Anotasi xml	26
Gambar 4.9 Hasil Dari Pemrosesan Gambar xml	28
Gambar 4.10 Proses Pemanggilan df.head	28
Gambar 4.11 Hasil Dari Proses Pemanggilan Df.head	29
Gambar 4.12 Proses mencetak gambar	29
Gambar 4.13 Proses Pembagian Dataset	29
Gambar 4.14 Proses Memodifikasi model deteksi objek	29
Gambar 4.15 Hasil dari Memodifikasi model deteksi objek	30
Gambar 4.16 Proses Mencetak Semua Komponen	30
Gambar 4.17 Proses Modifikasi Objek dan Peatihan Parameter	31
Gambar 4.18 Proses Menghitung Akurasi Mode Deteksi Objek	32
Gambar 4.19 Proses Fitur Penghentian Awal dan Pemantauan Kerugian	34
Gambar 4.20 Membuat Model yang Telah Disimpan	37
Gambar 4.21 Proses Menampilkan Gambar Dengan Overlay	37
Gambar 4.22 Proses Memuat Gambar dan Anotasi Kotak Pembatas	38
Gambar 4.23 Proses Menghitung Akurasi Akhir	40
Gambar 4.24 Hasil Akurasi	41
Gambar 4.25 Hasil Interpertasi Dataset	41
Gambar 4.26 Gambar Confusion Matrix	42
Gambar 4.27 Gambar Layer Epoch	43

DAFTAR LAMBANG DAN SINGKATAN

- TB : Tuberculosis
ANN : Artificial Neural Network
CNN : Convolutional Neural Network
AUC : Area Under the Curve
ROC : Receiver Operating Characteristics
XML : eXtensible Markup Language
GPU : Graphics Processing Unit
CPU : Central Processing Unit



DAFTAR ISTILAH

Tuberculosis	: Penyakit menular yang terutama menyerang paru-paru, tetapi dapat menyebar ke organ lain. Penyakit ini disebabkan oleh bakteri <i>Mycobacterium tuberculosis</i> .
Convolutional Neural Network	: Jenis arsitektur jaringan saraf tiruan yang dirancang khusus untuk memproses data yang memiliki pola grid, seperti gambar.
Kaggle	: Sebuah platform online yang berfokus pada kompetisi ilmu data (data science) dan pembelajaran mesin (machine learning).
Chest X-Ray	: Prosedur pencitraan medis yang menggunakan sinar-X untuk mengambil gambar bagian dalam dada, terutama paru-paru dan jantung.
DenseNET121	: Salah satu varian dari arsitektur DenseNet (Densely Connected Convolutional Networks) yang dirancang untuk memperbaiki efisiensi jaringan saraf konvolusional (CNN).
Airborne Disease	: Jenis penyakit yang disebarkan melalui partikel-partikel kecil di udara, seperti droplet atau aerosol.
Computer Vision	: Cabang ilmu komputer yang bertujuan untuk memungkinkan komputer memahami dan menafsirkan informasi visual dari dunia nyata.
Artificial Inteligent	: Bidang ilmu komputer yang berfokus pada pengembangan sistem atau mesin yang mampu melakukan tugas-tugas yang biasanya memerlukan kecerdasan manusia.
Confusion Matrix	: Sebuah tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan prediksi model terhadap label yang sebenarnya.
Symlink	: Jenis file di sistem operasi Unix-like yang berfungsi sebagai shortcut atau referensi ke file atau direktori lain.
Framework Pytorch	: Sebuah framework open-source untuk machine learning dan deep learning yang dikembangkan oleh Facebook's AI Research lab (FAIR).
Pandas	: Pustaka open-source untuk bahasa pemrograman Python yang digunakan untuk analisis data dan manipulasi data.

Dataframe	: Struktur data dua dimensi yang disediakan oleh pustaka seperti Pandas di Python, yang digunakan untuk menyimpan dan mengelola data dalam bentuk tabel.
Open CV	: Pustaka open-source yang digunakan untuk pemrosesan citra dan visi computer.
Shuffle Split	: Teknik untuk membagi dataset menjadi subset yang berbeda secara acak, yang sering digunakan dalam evaluasi dan pelatihan model machine learning.
Scikit Learn	: Pustaka open-source untuk machine learning dan analisis data di Python.
backbone ResNet50	: Sebuah arsitektur jaringan neural deep learning yang merupakan salah satu varian dari model Residual Networks (ResNet).
Ground Truth	: Istilah yang digunakan untuk merujuk pada data atau informasi yang dianggap sebagai referensi yang benar dan akurat dalam konteks evaluasi dan pelatihan model machine learning atau analisis data.
calculate_accuracy	: Istilah yang merujuk pada proses atau fungsi yang digunakan untuk mengukur seberapa akurat prediksi yang dibuat oleh sebuah model dibandingkan dengan nilai sebenarnya atau ground truth.
iou_threshold	: Parameter yang sering digunakan dalam konteks deteksi objek dan segmentasi gambar untuk menilai seberapa baik prediksi bounding box (kotak pembatas) dari model cocok dengan ground truth (kebenaran) dalam data.

INTISARI

Tuberculosis (TB) adalah penyakit yang disebabkan oleh *Mycobacterium tuberculosis*. Tuberculosis bisa menyerang bagian paru-paru dan dapat menyerang semua bagian tubuh. Tuberculosis adalah penyakit menular langsung, sebagian besar kuman TB menyerang paru-paru, tetapi dapat juga mengenai organ tubuh lainnya. Tuberculosis adalah penyakit infeksius kronis dan berulang biasanya mengenai organ paru yang disebabkan oleh *Mycobacterium tuberculosis*. Identifikasi penyakit TBC tidak mudah maka dari itu diperlukan suatu aplikasi untuk membantu mengidentifikasi penyakit TBC berdasarkan citra sampel dahak. Rumusan masalah dari penelitian ini adalah mengetahui bagaimana tingkat akurasi CNN dalam melakukan klasifikasi penyakit TBC. Dalam bidang kecerdasan buatan, salah satu algoritma terbaik yang dapat digunakan untuk melakukan klasifikasi citra adalah Convolutional Neural Network (CNN). Hasil pengujian menunjukkan bahwa model yang telah dilatih memiliki performa yang sangat baik, dengan akurasi pelatihan akhir sebesar 96.15% dan akurasi pengujian akhir sebesar 95.61%. Dengan akurasi yang hampir sama pada kedua dataset, dapat disimpulkan bahwa model ini stabil dan efektif dalam tugas klasifikasi yang dihadapinya.

Kata kunci: Tuberculosis (TB), *Mycobacterium tuberculosis*, Citra sampel dahak, Convolutional Neural Network (CNN), Kecerdasan buatan.

ABSTRACT

Tuberculosis (TB) is a disease caused by Mycobacterium tuberculosis. Tuberculosis can attack the lungs and can attack all parts of the body. Tuberculosis is a directly infectious disease, most TB germs attack the lungs, but can also affect other body organs. Tuberculosis is a chronic and recurrent infectious disease usually affecting the lungs caused by Mycobacterium tuberculosis. Identifying TB disease is not easy, therefore an application is needed to help identify TB disease based on images of sputum samples. The problem formulation of this research is to find out the level of accuracy of CNN in classifying TB disease. In the field of artificial intelligence, one of the best algorithms that can be used to classify images is Convolutional Neural Network (CNN). The test results show that the trained model has very good performance, with a final training accuracy of 96.15% and a final testing accuracy of 95.61%. With almost the same accuracy on both datasets, it can be concluded that this model is stable and effective in the classification tasks it faces.

Keyword: *Tuberculosis (TB), Mycobacterium tuberculosis, Image of sputum sample, Convolutional Neural Network (CNN), Artificial intelligence.*

