

# Penerapan Support Vector Machine untuk Klasifikasi Tuberkulosis Paru dari Citra Rontgen Dada

Aqilah Zahra<sup>1</sup>, Atifa Azzahirah<sup>2\*</sup>, Aulya Maharani Ahmad<sup>3</sup>  
<sup>1,2,3</sup>Program Studi Teknik Informatika, Universitas Muslim Indonesia, Kota Makassar

\*Email Korespondensi: [13020230065@student.umi.ac.id](mailto:13020230065@student.umi.ac.id)

Riwayat Artikel: Diterima: 12/02/2025; Direvisi: 17/04/2025; Disetujui: 02/05/2025

## ABSTRAK

Penelitian ini bertujuan untuk menerapkan algoritma Support Vector Machine (SVM) dalam mengklasifikasikan penyakit tuberkulosis paru menggunakan citra rontgen dada. Diagnosis tuberkulosis secara manual sering kali mengalami kendala, sehingga diperlukan sistem bantu keputusan berbasis komputer yang objektif. Metode penelitian ini meliputi pengumpulan *dataset* sekunder, tahap *preprocessing* citra yang mencakup *cropping*, *resizing*, dan konversi *grayscale*, serta pelatihan model klasifikasi menggunakan SVM. Berdasarkan hasil pengujian, model yang dikembangkan berhasil mencapai tingkat akurasi sebesar 90,11%. Hasil ini menunjukkan bahwa metode SVM efektif dalam membedakan antara paru-paru normal dan yang terinfeksi tuberkulosis. Meskipun demikian, performa model dalam mendeteksi kelas positif tuberkulosis masih lebih rendah dibandingkan kelas normal akibat ketidakseimbangan jumlah data. Secara keseluruhan, penelitian ini membuktikan bahwa pendekatan *machine learning* dapat diandalkan sebagai alat bantu diagnosis medis yang cepat dan konsisten.

**Kata Kunci:** citra rontgen dada; klasifikasi; pengolahan citra; support vector machine; tuberkulosis

## 1. PENDAHULUAN

Tuberkulosis (TBC) adalah masalah kesehatan serius global yang diagnosis konvensional terhambat oleh keterbatasan radiolog dan subjektivitas interpretasi citra Rontgen dada (CXR). Tuberkulosis (TBC) juga merupakan penyakit menular yang disebabkan oleh bakteri *Mycobacterium Tuberculosis*, dan sampai sekarang masih jadi masalah kesehatan global. Indonesia masuk dalam daftar negara dengan kasus tertinggi, dengan sekitar 824 ribu orang terkena dan 93 ribu meninggal dunia setiap tahunnya [1].

Kendala yang sering dihadapi tim medis adalah sulit untuk mengetahui daerah mana yang rawan untuk terjadi penyebaran penyakit sehingga pencegahan penyebaran sulit untuk dilakukan. Selama ini juga upaya untuk memantau penyakit masih dilakukan secara manual sehingga sulit untuk mendeteksi pola penyebaran secara akurat dan cepat [2]. Diagnosis TBC biasanya dilakukan lewat tes dahak atau analisis foto X-Ray dada, tapi cara ini butuh waktu lama dan sangat bergantung pada keahlian dokter, jadi biasa terjadi kesalahan [3]. Dengan perkembangan teknologi saat ini dunia medis dimudahkan dalam mempercepat proses diagnosis penyakit dengan memanfaatkan teknologi kecerdasan buatan [4].

Salah satu langkah awal yang krusial dalam penanganan TB adalah pemeriksaan radiologis atau sinar-X (*X-Ray*). Pemeriksaan ini memberikan gambaran apakah pasien terdiagnosis *positif* atau *negatif* tuberkulosis. Namun, pembacaan hasil radiologis secara manual memiliki tantangan tersendiri. Mengetahui ciri-ciri khusus yang terkandung pada citra sangat penting untuk memudahkan analisis dan memberikan gambaran hasil pengujian. Ekstraksi ciri merupakan bagian yang sangat penting dalam menggali informasi dari suatu objek yang terdapat pada citra diagnosis tuberkulosis [5]. Bagaimana membantu menganalisis citra rontgen paru untuk mendeteksi kasus tuberkulosis dengan cepat dan akurat. Pada proses diagnosis manual seringkali menimbulkan perbedaan hasil yang berbeda karena bergantung pada keahlian dan subjektivitas tenaga medis [4].

Kemajuan dalam pengolahan gambar dan *machine learning* telah mendorong riset di bidang kesehatan, terutama untuk mengklasifikasi penyakit paru-paru lewat foto rontgen dada. Meski begitu, mendiagnosis penyakit paru dari gambar X-Ray tetap sulit, bahkan untuk ahli radiologi, karena ada ketidakjelasan visual dan perbedaan subjektif dalam membaca hasilnya [6]. Oleh sebab itu, diperlukan sistem bantuan keputusan berbasis komputer yang bisa memproses gambar secara objektif dan konsisten untuk mendukung diagnosis

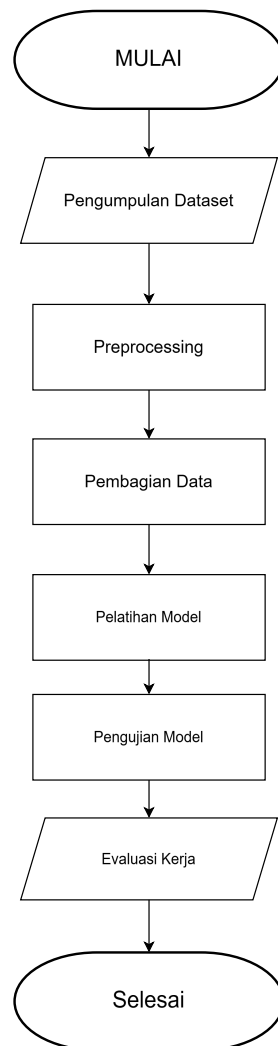
Salah satu algoritma *machine learning* yang bagus untuk klasifikasi adalah Support Vector Machine (SVM), karena bisa memisahkan data berdasarkan ciri visualnya dengan margin terbaik [7]. Metode ini terbukti efektif untuk membantu proses diagnosis TBC secara otomatis. Dengan menggunakan SVM dengan fungsi

Radial Basis Function (RBF) dapat mengklasifikasikan data kesehatan dengan akurasi hingga 98%, melampaui performa metode regresi logistik dalam menentukan pasien terinfeksi TBC [8].

Penelitian yang dilakukan oleh Teguh Muhammad Prasetyo dkk. Metode SVM digunakan untuk mengklasifikasikan citra paru-paru menjadi dua kategori, yaitu paru-paru normal dengan paru-paru pneumonia, yang menggunakan data dari kaggle. Proses klasifikasi dilakukan dengan mengekstraksi fitur pola bentuk seperti nilai *metric* dan *eccentricity* sebagai variabel masukan ke dalam algoritma SVM. Hasil penelitian menunjukkan bahwa metode SVM mampu menghasilkan tingkat akurasi tertinggi sebesar 79%, yang menunjukkan bahwa algoritma ini efektif dalam mengenali pola citra rontgen [9].

## 2. METODOLOGI PENELITIAN

Bagian Metodologi Penelitian menjelaskan secara sistematis tahapan, rancangan, dan pendekatan yang digunakan untuk mencapai tujuan penelitian.



Gambar 1. Desain Penelitian

### 2.1. Pengumpulan Dataset

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa citra radiografi dada (*Chest X-Ray*) yang diperoleh dari repositori publik Kaggle. *Dataset* yang digunakan adalah *Tuberculosis (TB) Chest X-ray Database*, yang dikumpulkan dan dipublikasikan oleh Tawsifur Rahman dkk. bekerja sama dengan peneliti dari Qatar University dan Dhaka University. *Dataset* ini digunakan dalam penelitian klasifikasi penyakit tuberkulosis paru menggunakan pendekatan *machine learning*. Akses *dataset* dapat diperoleh melalui tautan berikut: <https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>

*Dataset* ini berfokus pada klasifikasi penyakit paru-paru, khususnya Tuberkulosis (TBC), menggunakan metode *machine learning* seperti Support Vector Machine (SVM). Penyakit TBC disebabkan oleh infeksi

bakteri *Mycobacterium tuberculosis*. Data citra dikelompokkan ke dalam dua kelas utama yang digunakan dalam penelitian ini:

Tabel 1. Komposisi Data

Kelas Data	Deskripsi
Tuberculosis	Citra X-ray dada yang menunjukkan adanya infeksi TBC (positif TB) [6].
Normal	Citra X-ray dada dari individu sehat yang tidak menunjukkan kelainan paru-paru [4].

Total data yang tersedia bervariasi tergantung pada kebutuhan peneliti, namun umumnya mencakup ribuan citra untuk memastikan variasi data yang memadai [4].

## 2.2. Pembagian Dataset

Setelah melalui tahap *preprocessing*, *dataset* dibagi menjadi dua bagian, yaitu data latih dan data uji. Pembagian data dilakukan dengan proporsi 80% sebagai data latih dan 20% sebagai data uji. Data latih digunakan untuk melatih model Support Vector Machine, sedangkan data uji digunakan untuk mengukur kemampuan model dalam melakukan generalisasi terhadap data baru yang belum pernah dilihat sebelumnya. Proses pembagian data juga disertai dengan pengacakan (*shuffling*) untuk memastikan distribusi kelas yang merata dan menghindari bias urutan data. Selain itu, penggunaan nilai *random seed* diterapkan untuk menjaga konsistensi dan keterulangan hasil eksperimen.

## 2.3. Preprocessing Data

Pada *processing data* digunakan untuk memperbaiki serta menyiapkan data citra Rontgen Dada agar kualitasnya lebih optimal dan siap untuk di ekstraksi fitur serta dapat diklasifikasi dengan model Support Vector Machine (SVM). Berikut adalah tahapan *preprocessing data* citra untuk mengklasifikasikan Tuberculosis dengan menggunakan metode SVM:

### a. Cropping

*Cropping* adalah proses untuk membuang bagian yang tidak digunakan dan fokus citra hanya pada area paru-paru, proses ini dilakukan untuk mendapatkan citra yang lebih optimal dan membuang area-area yang diluar fokus seperti area diatas dan dibawah organ paru. Biasanya ukuran citra awalnya lebih besar dari 500 x 500 piksel.

### b. Resize

*Resize* adalah proses untuk mengubah ukuran citra Rontgen menjadi ukuran standart input data, proses ini dilakukan setelah citra di *crop*, setelah itu ukurannya akan diubah menjadi 128 x 128 piksel agar proses ekstraksi dan klasifikasinya menjadi lebih optimal.

### c. Grayscale

*Grayscale* adalah proses untuk menghilangkan warna warna pada citra, meskipun citra X-Ray pada dasarnya sudah berwarna abu-abu. Konversi ini memastikan citra diproses hanya berdasarkan intensitas piksel yang diperlukan untuk metode ekstraksi fitur seperti *Local Binary Pattern*.

### d. Ekstraksi Fitur

Ekstraksi Fitur adalah tahapan setelah *Processing Data* dengan menggunakan *Local Binary Pattern* untuk mengubah citra menjadi nilai ciri yang merepresentasikan tekstur pada citra, tekstur *grayscale* yang efektif dan tidak dipengaruhi oleh pencahayaan. Citra akan dipecah menjadi beberapa matrix wilayah dan nilai piksel diolah menjadi pola biner kemudian menghasilkan histogram ciri.

## 2.4. Pelatihan Model

Implementasi Support Vector Machine pada kasus klasifikasi citra rontgen paru-paru melibatkan serangkaian tahapan yang ketat:

- Pra-pemrosesan Data (*Preprocessing*): Citra rontgen dada perlu disiapkan agar dapat diolah secara seragam. Tahap ini umumnya mencakup penyesuaian dimensi citra (seperti *Resize* atau *Cropping*) untuk menyeragamkan format input model, suatu langkah penting dalam klasifikasi citra [10].
- Ekstraksi Fitur (*Feature Extraction*): Karena SVM bekerja dengan vector numerik, karakteristik citra harus diubah menjadi serangkaian nilai fitur. Untuk citra rontgen, ini bisa berupa fitur tekstur (misalnya, menggunakan metode *Local Binary Pattern* (LBP) yang efektif dalam mendeteksi pola kelainan, atau fitur bentuk (seperti Luas, Keliling, *Metric*, dan *Eccentricity* organ paru-paru) yang menjadi nilai masukan ( $x$ ) bagi model [10].

- c. Pelatihan Model dan Fungsi Kernel: Model dilatih untuk menemukan *hyperplane* optimal yang memisahkan kelas normal dan terinfeksi TB. Ketika data tidak terpisah secara linier (kasus non-linier), SVM menerapkan Fungsi Kernel atau dikenal sebagai *Kernel Trick*. Teknik ini secara implisit memetakan data ke ruang dimensi yang lebih tinggi agar dapat dipisahkan secara linier, tanpa perhitungan koordinat eksplisit. Kernel RBF (*Radial Basis Function*) adalah fungsi kernel non-linier yang sering digunakan dan terbukti memberikan hasil akurasi tinggi pada klasifikasi kasus TB [8].
- d. Evaluasi Model: Kinerja akhir model diukur dengan menggunakan data uji. Metrik yang digunakan untuk mengevaluasi ketepatan klasifikasi meliputi Akurasi, Presisi, dan *Recall*, yang dihitung berdasarkan matriks kebingungan (*Confusion Matrix*) [10].

**2.5. Pengujian Model**

Kinerja dan akurasi model SVM sangat bergantung pada penentuan *hyperparameter* yang optimal. Berikut adalah parameter kuncinya:

Tabel 2. Parameter Kunci SVM

Parameter Kunci	Nama Lengkap	Keterangan
Kernel	Fungsi Kernel	Digunakan untuk menangani data non-linier. Pilihan yang efektif untuk klasifikasi TB Adalah Radial Basis Function (RBF) [8].
C	Regularization Parameter	Mengontrol <i>trade-off</i> antara memaksimalkan <i>margin (generalization)</i> dan meminimalkan kesalahan klasifikasi pada data latih ( <i>training error</i> ) [8].
Gamma	Kernel Coefficient	Digunakan pada kernel RBF. Nilai Gamma menentukan radius pengaruh dari satu <i>Support Vector</i> Tunggal, nilai yang lebih kecil menghasilkan <i>decision boundary</i> yang lebih umum, sedangkan nilai yang lebih besar menghasilkan <i>boundary</i> yang lebih spesifik ( <i>overfitting</i> ) [8].

**2.6. Metode Evaluasi Kinerja Algoritma**

Evaluasi kinerja algoritma klasifikasi sangat penting untuk mengukur seberapa baik model yang dikembangkan dapat memprediksi kelas data baru yang belum pernah dilihat sebelumnya [4]. Dalam penelitian ini, metrik evaluasi kinerja model Support Vector Machine (SVM) diukur menggunakan pendekatan berbasis *Confusion Matrix*.

**2.6.1. Confusion Matrix**

*Confusion Matrix* (Matriks Kebingungan) adalah tabel yang digunakan untuk menggambarkan kinerja model klasifikasi pada sekumpulan data uji. Matriks ini menyajikan perbandingan antara hasil prediksi model dengan nilai *ground truth* (kelas yang sebenarnya) [4], [6]. Matriks ini mencakup empat komponen utama yang vital dalam mengukur kinerja klasifikasi biner (dua kelas, yaitu Normal dan TB):

Tabel 3. Teknik Evaluasi

	Diprediksi Normal	Diprediksi TB
Aktual Normal	<i>True Negative (TN)</i>	<i>False Positive (FP)</i>
Aktual TB	<i>False Negative (FN)</i>	<i>True Positive (TP)</i>

**2.6.2. Metrik Kinerja**

Berdasarkan komponen dari *Confusion Matrix*, metrik kinerja yang umum digunakan dalam klasifikasi citra medis, khususnya untuk TBC, meliputi akurasi (*accuracy*), presisi (*precision*), dan sensitivitas (*sensitivity*), yang dirumuskan sebagai berikut [6], [4]:

- a. Akurasi (*Accuracy*)

Akurasi mengukur proporsi total prediksi yang benar (baik TP maupun TN) terhadap seluruh jumlah data uji. Akurasi memberikan gambaran umum kinerja model secara keseluruhan.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

Di mana:

TP: jumlah data positif yang diprediksi benar.  
TN: jumlah data negatif yang diprediksi benar.  
FP: jumlah data negatif yang salah diprediksi sebagai positif.  
FN: jumlah data positif yang salah diprediksi sebagai negatif.

b. Presisi (*Precision*)

Presisi mengukur tingkat kebenaran hasil positif yang diprediksi model. Ini penting untuk mengetahui seberapa andal model dalam memprediksi suatu kasus benar-benar positif (TB).

$$precision = \frac{TP}{TP+FP} \quad (5)$$

Di mana:

Presisi membandingkan jumlah True Positive (TP), yaitu data positif yang diprediksi benar, dengan keseluruhan data yang diprediksi positif, yaitu TP ditambah False Positive (FP), di mana FP merupakan data negatif yang salah diprediksi sebagai positif.

c. Sensitivitas / *Recall* (*Sensitivity*)

Sensitivitas (sering disebut *Recall*) mengukur kemampuan model untuk mengidentifikasi semua kasus positif (TB) yang sebenarnya. Dalam konteks medis, Sensitivitas sangat penting, karena rendahnya nilai ini berarti model sering melewatkan pasien TB yang sebenarnya sakit.

$$sensitivity = \frac{TP}{TP+FN} \quad (6)$$

Di mana:

Sensitivitas membandingkan jumlah True Positive (TP), yaitu data positif yang berhasil diprediksi benar oleh model, dengan total data positif yang sebenarnya, yaitu TP ditambah False Negative (FN), di mana FN merupakan data positif yang gagal diidentifikasi oleh model.

### 3. HASIL DAN PEMBAHASAN (10 PT)

Hasil penelitian diuraikan dengan tahapan pengolahan data dan pengujian yang telah dilakukan, mulai dari *pre-processing* citra hingga evaluasi model. Pembahasan ini untuk menginterpretasikan performa model serta signifikansi teknis dari setiap tahapan pengolahan data.

#### 3.1. Hasil

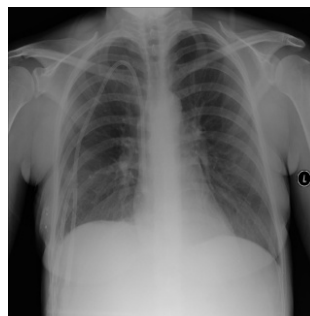
Pada bagian ini menyajikan hasil klasifikasi citra rontgen dada menggunakan algoritma Support Vector Machine (SVM) berdasarkan data uji yang telah disiapkan. Klasifikasi ini bertujuan untuk mengukur efektivitas model dalam membedakan kelas citra yaitu normal dan tuberkulosis.

##### 3.1.1. Hasil *Preprocessing* Citra

Sebelum melakukan klasifikasi, data citra mentah akan melalui tahap *preprocessing* sebelum masuk ke dalam algoritma klasifikasi. Digunakan untuk meminimalisir *noise*, menghilangkan informasi yang tidak dibutuhkan, dan membuat format data yang sama.

a. Citra Rontgen Dada Asli

Data awal yang masih mentah yang memiliki redolusi dan format warna. Citra ini masih mengandung banyak informasi latar belakang (*background*) serta artefak medis (tulang bahu) yang tidak diperlukan dalam proses klasifikasi.



Gambar 1. Citra Rontgen Dada Asli

b. Cropping

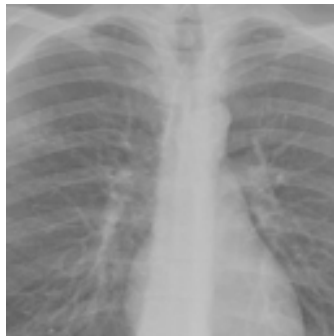
*Cropping* digunakan untuk menghilangkan bagian tepi citra yang tidak relevan dan memfokuskan tampilan hanya pada area rongga dada saja. Dengan menghilangkan area hitam di sekitar objek dan memfokuskan data pada area rongga dada maka model SVM dapat bekerja lebih efisien.



Gambar 2. Hasil *Cropping*

c. Grayscale

Citra asli dikonversi dari format RGB (3 *channel* warna) menjadi *grayscale* (1 *channel*). Langkah ini dilakukan untuk mereduksi dimensi data tanpa menghilangkan informasi tekstur dan intensitas piksel yang merepresentasikan struktur paru-paru.



Gambar 3. Hasil *Grayscale*

d. Resizing

Citra hasil *cropping* kemudian diubah ukurannya (*resize*) menjadi dimensi seragam, yaitu 128 x 128 piksel. Penyeragaman ini mutlak diperlukan karena algoritma SVM membutuhkan vektor fitur dengan panjang yang konsisten sebagai input.



Gambar 4. Hasil *Resizing*

### 3.1.2. Hasil Pembagian Data

Setelah melalui tahapan *preprocessing*, *dataset*  $x$  (fitur citra) dan  $y$  (label kelas) dibagi menjadi dua bagian. Proses ini bertujuan untuk memisahkan data yang digunakan untuk pembelajaran model dengan data yang digunakan untuk evaluasi kinerja. Berikut ini hasil dari pembagian data :

a. Proporsi Pembagian

Hal ini membagi *dataset* dengan proporsi 80% dialokasikan sebagai data latih untuk algoritma SVM, sedangkan 20% untuk data uji. Data uji digunakan untuk validasi eksternal untuk mengukur kemampuan generalisasi model terhadap data baru yang belum pernah dilihat selama proses pelatihan

b. Pengacakan Data (*Shuffling*)

Parameter *shuffle* diatur bernilai *True*. Pengaturan ini memastikan bahwa urutan data diacak terlebih dahulu sebelum dilakukan pemotongan (partisi). Pengacakan ini dilakukan untuk menghilangkan bias urutan input dan memastikan distribusi kelas (misalnya kelas normal dan tuberkulosis) tersebar secara representatif, baik pada data latih maupun data uji.

c. Konsistensi Eksperimen (*Reproducibility*)

Penentuan nilai *seed* ini bertujuan untuk menjamin *reproducibility* (keterulangan) eksperimen. Dengan pengaturan ini, proses pengacakan dan pembagian data akan menghasilkan subset indeks yang sama persis setiap kali program dijalankan, sehingga perbandingan performa model di masa mendatang tetap valid dan konsisten.

**3.1.3. Hasil Klasifikasi Menggunakan Support Machine Vector**

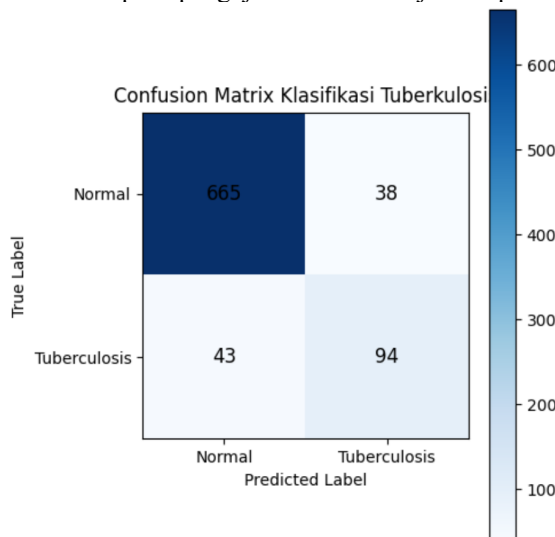
Model Support Vector Machine (SVM) digunakan untuk melakukan klasifikasi citra rontgen dada yang telah melalui tahap *preprocessing*. Model dilatih menggunakan data latih dan selanjutnya diuji menggunakan data uji untuk menghasilkan label prediksi pada setiap citra rontgen dada. Hasil klasifikasi berupa prediksi kelas normal atau tuberkulosis untuk setiap citra rontgen dada pada data uji, yang diperoleh berdasarkan hasil pemodelan SVM terhadap fitur citra yang digunakan.

**3.1.4. Hasil Evaluasi Model Klasifikasi**

Evaluasi performa model Support Vector Machine (SVM) dilakukan menggunakan *classification report* untuk menilai kinerja model pada masing-masing kelas. Kelas 0 merepresentasikan citra rontgen dada normal, sedangkan kelas 1 merepresentasikan citra rontgen dada dengan tuberkulosis. Hasil evaluasi menunjukkan bahwa model memperoleh nilai akurasi sebesar 90,11% pada data uji. Pada kelas normal, model mencapai nilai *precision* 93%, *recall* 95%, dan *F1-score* 94%, sedangkan pada kelas tuberkulosis diperoleh nilai *precision* 71%, *recall* 66%, dan *F1-score* 68%. Nilai *weighted average* menunjukkan performa keseluruhan model dengan mempertimbangkan proporsi jumlah data pada masing-masing kelas.

**3.1.5. Confusion Matrix**

*Confusion matrix* digunakan untuk menyajikan hasil klasifikasi model Support Vector Machine terhadap data uji secara terperinci berdasarkan kelas aktual dan kelas prediksi. *Confusion matrix* ini menunjukkan jumlah citra yang diklasifikasikan dengan benar dan salah untuk masing-masing kelas, yaitu paru normal dan paru tuberkulosis. Hasil *confusion matrix* pada pengujian model ditunjukkan pada gambar 6.



Gambar 6. Hasil *Confusion Matrix*

Berdasarkan hasil evaluasi menggunakan *confusion matrix*, model Support Vector Machine (SVM) menunjukkan kemampuan yang baik dalam mengklasifikasikan citra rontgen paru ke dalam kelas normal dan tuberkulosis. Dari total data uji, sebanyak 665 citra paru normal berhasil diklasifikasikan dengan benar sebagai normal, sedangkan 38 citra normal lainnya salah diklasifikasikan sebagai tuberkulosis. Pada kelas tuberkulosis, model berhasil mengenali 94 citra secara benar, namun masih terdapat 43 citra tuberkulosis yang salah diklasifikasikan sebagai normal.

### 3.2. Pembahasan

Hasil pengujian menunjukkan bahwa model Support Vector Machine (SVM) dapat mengklasifikasikan citra rontgen dada ke dalam dua kelas paru normal dan paru tuberkulosis dengan tingkat akurasi 90,11%. Nilai akurasi ini menunjukkan bahwa model dapat mengidentifikasi sebagian besar gambar pada data uji dengan benar. Dengan demikian, metode SVM dapat membedakan kedua kondisi paru berdasarkan fitur gambar yang digunakan dalam penelitian.

Berdasarkan hasil confusion matrix dan classification report, performa model pada kelas paru normal lebih tinggi dibandingkan dengan kelas paru tuberkulosis. Hal ini ditunjukkan oleh nilai precision dan recall yang lebih besar pada kelas normal. Kondisi tersebut dapat dipengaruhi oleh ketidakseimbangan jumlah data, di mana citra paru normal memiliki jumlah yang lebih banyak dibandingkan citra paru tuberkulosis. Selain itu, karakteristik visual pada citra tuberkulosis, khususnya pada kasus dengan tingkat keparahan ringan, sering kali memiliki pola yang menyerupai citra paru normal, sehingga menyulitkan model dalam membedakan kedua kelas tersebut secara optimal.

Kesalahan klasifikasi yang terjadi, terutama pada kelas tuberkulosis yang diprediksi sebagai paru normal, menunjukkan bahwa masih terdapat keterbatasan dalam proses ekstraksi fitur citra. Faktor lain yang turut memengaruhi hasil klasifikasi meliputi variasi kualitas citra rontgen, perbedaan tingkat kontras, serta keberadaan noise pada citra, yang dapat mengurangi kejelasan area lesi paru. Kondisi ini berdampak pada kemampuan model dalam mengenali ciri khas tuberkulosis secara konsisten.

Secara teoritis, Support Vector Machine dikenal sebagai metode yang efektif untuk permasalahan klasifikasi dua kelas dengan data berdimensi tinggi, termasuk pada bidang citra medis. Hasil penelitian ini sejalan dengan teori tersebut serta mendukung temuan penelitian terdahulu yang menyatakan bahwa SVM mampu memberikan performa yang cukup andal dalam klasifikasi citra rontgen dada. Dengan membentuk hyperplane pemisah yang optimal, SVM dapat memaksimalkan margin antar kelas, sehingga menghasilkan kinerja klasifikasi yang stabil.

Pada penelitian ini SVM memiliki kelebihan berupa struktur model yang sederhana dan kebutuhan komputasi yang relatif rendah, sehingga berpotensi diterapkan pada sistem dengan sumber daya terbatas. Namun, model ini masih memiliki keterbatasan dalam menangkap detail lokal pada area lesi paru karena penggunaan fitur citra yang bersifat global, yang berdampak pada masih terjadinya kesalahan klasifikasi pada kelas tuberkulosis. Secara keseluruhan, penerapan SVM menunjukkan performa yang cukup baik dalam klasifikasi tuberkulosis paru dari citra rontgen dada dan memiliki potensi untuk dikembangkan lebih lanjut melalui penambahan segmentasi paru, peningkatan teknik ekstraksi fitur, serta penyeimbangan data guna meningkatkan kemampuan deteksi kasus tuberkulosis.

## 2. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa algoritma Support Vector Machine (SVM) berhasil diterapkan untuk mengklasifikasikan tuberkulosis paru dari citra rontgen dada dengan tingkat akurasi mencapai 90,11%. Metode ini terbukti dapat membedakan antara citra paru normal dan citra paru tuberkulosis. Hasil menunjukkan kontribusi penelitian dalam membangun sistem deteksi penyakit alternatif yang tidak hanya memiliki model yang sederhana tetapi juga memiliki kinerja yang andal dan objektif dibandingkan dengan diagnosis manual.

Meskipun akurasinya tinggi, metode SVM yang digunakan masih sulit untuk mendeteksi kasus positif tuberkulosis dengan tepat karena nilai recallnya yang lebih rendah. Hal ini dipengaruhi oleh kemiripan karakteristik visual antara kasus infeksi ringan dan citra non-tuberkulosis, serta ketidakseimbangan jumlah data pada proses pelatihan model. Oleh karena itu, untuk penelitian selanjutnya, disarankan untuk mengembangkan dengan mengintegrasikan metode segmentasi paru-paru untuk mengurangi noise, menerapkan teknik ekstraksi fitur yang lebih mendetail, serta menggunakan metode penyeimbangan data guna meningkatkan sensitivitas model dalam mendeteksi kasus tuberkulosis.

## DAFTAR PUSTAKA

- [1] M. B. Sastramandala, R. N. Sari, and N. Novrina, "Pemanfaatan Deep Learning menggunakan Convolutional Neural Network Terhadap Penyakit Tuberkulosis melalui Citra Rontgen," *J. Minfo Polgan*, vol. 14, no. 1, pp. 631–642, 2025, doi: 10.33395/jmp.v14i1.14818.
- [2] A. Pathak *et al.*, "Application of Machine Learning K-Means Clustering and Linear Regression in Determining the Risk Level of Pulmonary Tuberculosis," *2024 IEEE Conf. Comput. Appl. Syst. COMPAS 2024*, vol. 12, no. 1, pp. 336–348, 2024, doi: 10.1109/COMPAS60761.2024.10796963.
- [3] A. A. Alimi, A. R. Adriansyah, and P. Prima, "Pengembangan Sistem Deteksi Tuberkulosis pada Citra X-Ray Menggunakan Metode Convolutional Neural Network (CNN) dengan Framework Laravel," *J. Inform. Terpadu*, vol. 10, no. 2, pp. 165–171, 2024, doi: 10.54914/jit.v10i2.1437.
- [4] A. Rasyid and L. Heryawan, "Klasifikasi Penyakit Tuberculosis (TB) Organ Paru Manusia Berdasarkan

- Citra Rontgen Thorax Menggunakan Metode Convolutional Neural Network (CNN),” *J. Manaj. Inf. Kesehat. Indones.*, vol. 11, no. 1, pp. 35–44, 2023, doi: 10.33560/jmiki.v11i1.484.
- [5] J. Lin and H. Irsyad, “Klasifikasi Pneumonia Pada Citra X-Rays Paru-Paru Menggunakan GLCM Dan LVQ,” *J. Algoritm.*, vol. 1, no. 2, pp. 184–194, 2021, doi: 10.35957/algoritme.v1i2.897.
- [6] Ningsiah, S. Yusuf Irianto, and J. Triloka, “Analisis Perbandingan Penyakit Paru-Paru Dari Citra X-Ray Dada Menggunakan Metode Convolutional Neural Network Dan Support Vector Machine,” *J. Teknol. dan Komputasi*, vol. 6, no. 1, pp. 80–100, 2024.
- [7] K. Sowjanya, G. Poojitha, C. K. Saran, B. Priyanka, and D. Ahalya, “Pulmonary Tuberculosis Detection from Chest X-Ray Images Using Machine Learning,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 4, pp. 1538–1542, 2023, doi: 10.22214/ijraset.2023.50382.
- [8] S. Rahayu and Y. Yamasari, “Klasifikasi Penyakit Stroke dengan Metode Support Vector Machine (SVM),” *J. Informatics Comput. Sci.*, vol. 5, no. 03, pp. 440–446, 2024, doi: 10.26740/jinacs.v5n03.p440-446.
- [9] T. M. Prasetyo, A. Amrullah, S. Syahrir, and B. N. Sari, “Implementasi Algoritma Svm (Support Vector Machine) Dalam Klasifikasi Penyakit Paru-Paru Berdasarkan Fitur Pola Bentuk,” *J. Teknol. Inf.*, vol. 6, no. 1, pp. 1-- 6, 2022.
- [10] D. Mariana, H. Irsyad, D. Mariana, and H. Irsyad, “Klasifikasi Pneumonia pada Chest X-Ray Paru-paru dengan Ekstraksi Fitur Local Binary Pattern Menggunakan Support Vector Machine,” *J. Ilm. BETRIK Besemah Teknol. Inf. dan Komput.*, vol. 12, no. 1, pp. 54–62, 2021.