PERBANDINGAN PERFORMA METODE SVM DAN KNN DALAM MENGKLASIFIKASIKAN CITRA INFEKSI TELINGA

Kuncoro Ariadi, Fetty Tri Anggraeny, Andreas Nugroho Sihananto
Informatika, Universitas Pembangunan Nasional Veteran Jawa Timur
Jalan Raya Rungkut Madya No. 1, Surabaya, Indonesia
kuncoroariadi@gmail.com

ABSTRAK

Infeksi telinga merupakan salah satu masalah kesehatan pada organ telinga yang sering terjadi pada manusia. Dengan perkembangan teknologi di bidang medis, metode klasifikasi citra menjadi salah satu pendekatan yang menjanjikan untuk membantu dalam mendeteksi infeksi telinga dengan cepat dan akurat. Penelitian ini bertujuan untuk membandingkan performa metode klasifikasi citra, yaitu *Support Vector Machine* (SVM) dan K-*Nearest Neighbors* (KNN) dalam mengklasifikasikan infeksi telinga pada citra telinga tengah. Citra telinga tengah yang digunakan merupakan citra otoskopi telinga yang terdiri dari beberapa kategori diantaranya adalah kategori Myringosklerosis, Telinga Normal, Sumbatan Kotoran Telinga, dan Otitis Media Kronik (COM). Data citra tersebut diolah dengan beberapa tahapan, yatu peningkatan kontras dan kualitas citra, ekstraksi *Region of Interest* (ROI) pada citra, ekstraksi fitur berbasis tekstur, dan *tuning hyperparameter* dari kedua metode SVM dan KNN. Berdasarkan hasil penelitian, metode SVM dengan kernel *Linear* berhasil mencapai akurasi yang lebih tinggi dibandingkan metode KNN dengan nilai *k* yang terbaik yaitu $k = 1$. Masing-masing metode dievaluasi menggunakan *confusion matrix* yang mencakup matriks akurasi, presisi, sensitivitas atau *recall*, dan *F1-score*. Dari tahapan pengujian yang dilakukan, didapatkan skor akurasi untuk SVM sebesar 87,50%, skor presisi sebesar 87,94%, skor sensitivitas sebesar 87,50%, dan *F1-Score* sebesar 87,50%, sedangkan untuk KNN didapatkan skor akurasi sebesar 72,50%, skor presisi sebesar 76,66%, skor sensitivitas sebesar 72,50%, dan *F1-Score* sebesar 71,80%.

Kata kunci : *Klasifikasi Citra, Support Vector Machine, K-Nearest Neighbors, Citra Otoskopi, Region of Interest, Ekstraksi Fitur*

1. PENDAHULUAN

Telinga adalah salah satu organ penting bagi kehidupan manusia. Fungsinya yang berperan sebagai alat pendengaran dan menjaga keseimbangan tubuh menjadikan organ tersebut sangat penting untuk diperhatikan kebersihan dan kesehatannya. Rendahnya kesadaran masyarakat saat ini menjadikan penyakit telinga umum terjadi diusia dewasa bahkan kanak-kanak. Data yang dimuat berdasarkan Organisasi Kesehatan Dunia atau WHO menyebutkan bahwa terdapat sekitar 360 juta orang mengalami masalah pendengaran dengan 32 juta diantaranya masih berada di usia kanak-kanak [1].

Salah satu masalah pendengaran yang umum dialami adalah infeksi telinga. Masalah pendengaran ini dapat mengganggu pendengaran disertai nyeri dan komplikasi yang serius jika tidak segera ditangani. Jenis infeksi telinga yang sering ditemukan pada penderita adalah otitis media. Infeksi ini menyerang pada membran timpani dan bagian telinga tengah, dengan beberapa jenis yaitu otitis media, otitis media efusi (OME), dan otitis media supuratif kronik (OMSK) [2].

Penyakit infeksi telinga masih umum terjadi meskipun dengan pesatnya perkembangan teknologi. Gejala infeksi telinga sering dilaksanakan oleh pasien, namun memerlukan proses identifikasi lebih lanjut untuk mendiagnosisnya. Maka dari itu, perlu nya dibuat sistem yang dapat melakukan pengolahan citra digital yang dapat mendeteksi kasus penyakit dengan cepat dan akurat. Penelitian sebelumnya menunjukkan penggunaan algoritma *Support Vector Machine* (SVM) dalam mengklasifikasi citra menghasilkan tingkat akurasi sebesar 90,10% dan penggunaan algoritma K-*Nearest Neighbors* (KNN) menghasilkan akurasi sebesar 81,31% [3].

Penelitian tentang klasifikasi citra telinga juga telah dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dan K-*Nearest Neighbors* (KNN). Penggunaan metode klasifikasi tersebut didukung dengan menggunakan ekstraksi fitur *filter bank* untuk meningkatkan performa kedua algoritma tersebut. Hasil penelitian ini menunjukkan performa SVM mengungguli performa KNN dengan tingkat akurasi validasi sebesar 99,03% sedangkan KNN sebesar 99,01%. Oleh karena itu, algoritma SVM dijadikan pilihan untuk dilakukan pengujian lebih lanjut sehingga menghasilkan tingkat akurasi sebesar 93,88% [4].

Penggunaan algoritma *Support Vector Machine* (SVM) dan K-*Nearest Neighbors* telah banyak diterapkan di berbagai klasifikasi citra. Penelitian sebelumnya yang membahas tentang klasifikasi otitis media efusi (OME) menunjukkan tingkat akurasi SVM sebesar 98,2% dan KNN mendapat skor akurasi sebesar 97,4% [5]. Selain itu, penelitian lain yang membahas mengenai klasifikasi citra penyakit hati memperlihatkan bahwa SVM mendapatkan skor akurasi 75,20% sedangkan KNN mendapatkan skor akurasi 70,90% [6].

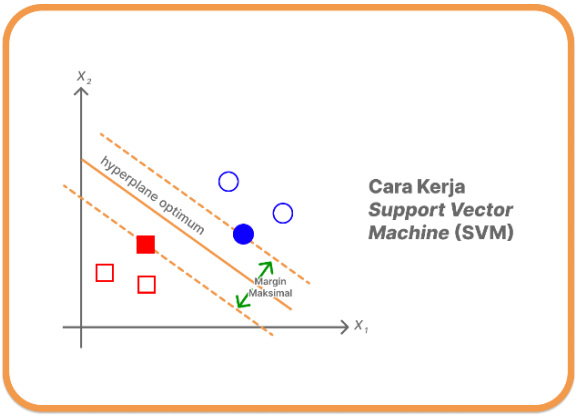
2. TINJAUAN PUSTAKA

Selama melakukan penelitian ini, tentunya dibutuhkan referensi dan dasar pengetahuan yang digunakan sebagai acuan untuk melakukan penelitian. Berbagai referensi yang relevan diperoleh dari penelitian terdahulu dengan tujuan agar dapat mengetahui *gap* antara penelitian ini dengan penelitian sebelumnya.

2.1. *Support Vector Machine*

Metode *Support Vector Machine* (SVM) adalah algoritma pembelajaran mesin yang sering dimanfaatkan untuk klasifikasi dan regresi data karena kemampuannya yang dapat melakukan teknik analisis data dan mengenali suatu pola atau batasan keputusan dalam suatu kumpulan data [7]. SVM diperkenalkan oleh Vapnik dan Chervonenkis yang pada awal pembuatannya metode ini hanya dapat megklasifikasi data linier dengan gambar yang disebut *hyperplane*. Pengembangan SVM dilanjutkan pada tahun 1992 oleh Vapnik, Boser, dan Guyon dengan membangun klasifikasi non-linier.

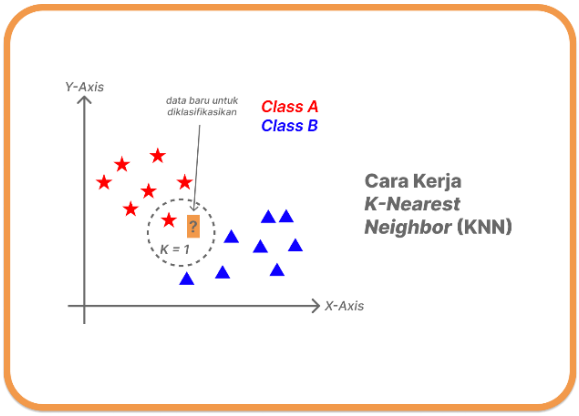
Cara kerja *Support Vector Machine* (SVM) menggunakan sebuah garis penentu yang disebut *hyperplane*. Penentuan *hyperplane* terbaik digunakan sebagai pembatas antara dua kelas dengan mengoptimalkan jarak antar kelas tersebut [8]. *Hyperplane* terbaik ditentukan dari kemampuannya dalam memisahkan data dengan *margin* terbesar. SVM bekerja dengan memisahkan kelas, seperti yang tertera pada Gambar 1 dengan lingkaran biru adalah kelas pertama dan kotak ungu adalah kelas kedua.



Gambar 1. Ilustrasi Cara Kerja SVM

2.2. K-*Nearest Neighbors*

Metode K-*Nearest Neighbors* (KNN) adalah salah satu algoritma pembelajaran mesin yang dapat digunakan untuk regresi dan klasifikasi. KNN bekerja dengan menentukan nilai atau label dari suatu data dengan mencari nilai *k* terdekat [9]. Algoritma ini bekerja dengan dua tahap, yang pertama adalah menganalisis nilai *k* tetangga terdekat dari data baru, yang kedua adalah menentukan kelas data baru tersebut berdasarkan kelas dari tetangga itu sendiri. Dalam menentukan nilai *k* tersebut umumnya diukur menggunakan *distance* bernama *euclidean distance* [10].



Gambar 2. Ilustrasi Cara Kerja KNN

Pada Gambar 2 diilustrasikan bagaimana K-*Nearest Neighbors* (KNN) dalam mengklasifikasi suatu kelas atau data. Data uji dibandingkan dengan keseluruhan vektor sampel untuk menghitung jarak nilai *k* yang terdekat sehingga *k* dapat memilih data atau kelas tersebut masuk kedalam kelas mana. Data baru yang digunakan untuk klasifikasi ditentukan oleh nilai *k* yang menjadi nilai terdekat [11]. Klasifikasi akhir pada KNN ini didasarkan pada kategori data atau label kelas yang paling sering muncul di antara nilai tetangga terdekat tersebut.

2.3. *Region of Interest (ROI)*

Ekstraksi *Region of Interest* (ROI) merupakan pengambilan area spesifik yang diambil pada suatu data yang akan dijadikan sebagai fokus utama untuk keperluan pemrosesan citra. Penggunaan ROI ini ditujukan untuk menghilangkan bagian yang tidak digunakan dan tidak relevan pada data citra dengan fokus penelitian [12]. Selain itu, ROI dapat membantu efisiensi analisis data citra dengan memusatkan perhatian pada bagian citra yang sangat penting. Teknik ROI ini sering digunakan pada penlitian yang menggunakan citra medis sebagai data utama nya, di mana ekstraksi ini cocok digunakan untuk memfokuskan deteksi tekstur citra medis.

2.4. *Maximum Response 8*

Ekstraksi fitur *Maximum Response 8* (MR8 Filter Bank) merupakan salah satu implementasi *filter bank* yang berfungsi sebagai ekstraksi dalam analisis tekstur pada data citra. Teknik ini menggunakan serangkaian filter anisotropik yang terdiri dari *filter edge* dan *filter bar* dengan menggunakan gabungan dari skala dan orientasi. Filter isotropik juga digunakan pada teknik ini yang terdiri dari *filter Gaussian* dan *filter Laplacian of Gaussian* (LoG) dengan menggunakan hal yang sama pada filter anisotropik sebelumnya. MR8 dapat menghasilkan 38 filter namun hanya delapan filter yang dapat digunakan dan dipertimbangkan dengan memilih respon maksimal dari beberapa orientasi yang ditentukan [13].

2.5. *Confusion Matrix*

Confusion matrix merupakan alat evaluasi kinerja model klasifikasi dalam pembelajaran mesin. Penggunaan *confusion matrix* umumnya digunakan dalam klasifikasi *multi-class* dengan setiap *instance* data hanya dapat termasuk dalam satu kelas. Namun, dalam penerapan nya, sebuah *instance* data tidak dapat dengan mudah untuk diklasifikasikan ke dalam satu kelas saja. Oleh karena itu, klasifikasi dengan *multi-labels* diperkenalkan dengan satu *instance* data dapat diberikan lebih dari satu label untuk menggambarkan setiap karakteristiknya [14]. *Confusion matrix* memiliki empat komponen utama yaitu *True Positive* atau TP, *True Negative* atau TN, *False Positive* atau FP, dan *False Negative* atau FN. Dari keempat komponen utama tersebut maka dapat dihitung beberapa metrik evaluasi yang diantaranya adalah sebagai berikut.

a. Akurasi atau proporsi prediksi benar dari keseluruhan prediksi model klasifikasi. Akurasi dapat dihitung menggunakan persamaan (1).

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

b. Presisi atau proporsi prediksi nilai positif yang benar dari seluruh prediksi positif yang diperoleh model klasifikasi. Nilai presisi dapat dihitung menggunakan persamaan (2).

$$Presisi = \frac{TP}{TP+FP} \quad (2)$$

c. Sensitivitas atau nilai *recall* merupakan proporsi prediksi positif yang benar dari seluruh data asli yang bernilai positif atau benar. Nilai *recall* dihitung dengan menggunakan persamaan (3)

$$Sensitivitas = \frac{TP}{TP+FN} \quad (3)$$

d. Nilai *F1-Score* merupakan rata rata yang didapatkan dari nilai presisi dan sensitivitas. Nilai ini berfungsi untuk menilai ketidakseimbangan kelas dalam data yang diklasifikasi oleh model. *F1-Score* dapat dihitung menggunakan persamaan (4) berikut.

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi+Recall} \quad (4)$$

e. Spesifisitas atau proporsi prediksi negatif yang benar dari keseluruhan data yang sebenarnya bernilai negatif. Spesifisitas dapat dihitung dengan menggunakan persamaan (5).

$$Spesifisitas = \frac{TN}{TN+FP} \quad (5)$$

3. METODE PENELITIAN

3.1. Pengumpulan Data

Tahapan pertama akan dilakukan dengan pengumpulan data citra infeksi telinga tengah. Dataset yang digunakan pada penelitian ini terdapat pada situs https://figshare.com/articles/dataset/Ear_imagery_database/11886630 . Di dalam situs tersebut terdapat dataset sebesar 880 data citra yang divalidasi oleh spesialis THT. Data citra tersebut terbagi menjadi 2 *folder* yaitu data train-validation sejumlah 720 data citra dan data testing sejumlah 160 data citra dengan masing-masing *folder* terdapat 4 kelas infeksi telinga tengah. Masing-masing kelas tersebut terdiri dari, citra telinga normal, citra penyakit myringosklerosis, citra penyakit otitis media kronik atau COM, dan citra kotoan sumbatan telinga. Seluruh data citra akan diproses terlebih dahulu pada tahap praproses data citra dan ekstraksi fitur. Data yang sudah melalui proses normalisasi fitur akan dilanjutkan ke tahapan pelatihan dan pengujian algoritma klasifikasi.

3.2. Praproses Data

Data citra yang telah diperoleh akan dilakukan praproses terlebih dahulu. Tahap pertama dilakukan dengan *resizing* citra dengan ukuran 64x64 sehingga dapat meningkatkan akurasi model dan menghemat waktu komputasi model. Data citra kemudian diubah menjadi *grayscale* untuk mengurangi tingkat kompleksitas data citra. Setelah itu, diterapkan metode *Contrast Limited Adaptive Histogram Equalization* (CLAHE) untuk meningkatkan intensitas kontras pada data citra sehingga *noise* pada gambar dapat dikurangi serta meningkatkan kinerja ekstraksi fitur untuk dapat mengekstraksi fitur dengan jumlah yang lebih banyak. Langkah selanjutnya yaitu ekstrasi *Region of Interest* (ROI) untuk mendapatkan area lingkaran pada citra otoskopi sehingga proses ekstraksi dan klasifikasi dapat terfokus pada satu objek yaitu membran timpani citra otoskopi. Data citra yang telah diekstraksi ROI kemudian dilakukan peningkatan kontras untuk membuat perbedaan dari objek yang diambil dengan latar belakangnya.

3.3. Ekstraksi Fitur

Tahapan selanjutnya akan diimplementasikan ekstraksi fitur menggunakan *Maximum Respon 8 Filter Bank*. Untuk filter yang digunakan pada penelitian ini terdapat 8 filter yang terdiri dari 2 filter bersifat isotropik, 3 untuk *filter bar*, dan 3 untuk *filter edge* yang masing-masing diterapkan skala yang sama kecuali filter isotropik. Masing-masing data citra yang telah didapatkan respon fitur nya akan distandarisasi untuk mendapatkan nilai *mean* dan standar deviasinya. Data citra yang telah distandarisasi tersebut kemudian dinormalisasi menggunakan *Principal Component Analysis* (PCA) untuk mereduksi jumlah fitur tanpa mengurangi informasi penting yang terdapat didalam masing-masing data citra.

3.4. Pelatihan dan Pengujian Model

Pada tahapan ini, masing-masing algoritma akan dilakukan pelatihan dengan *tuning hyperparameter*. Data train-validation dan data testing akan dibagi menjadi beberapa rasio *split data* dengan masing-masing rasio akan dilakukan pelatihan dan pengujian menggunakan parameter yang ditentukan pada masing-masing model.

Tabel 1. Tabel Rasio Pembagian Data

Iterasi	Latih	Validasi
1	80%	20%
2	75%	25%
3	70%	30%
4	65%	35%
5	60%	40%

Pada Tabel 1 dijelaskan mengenai rasio *split* data pada *folder* data train-validation. Set pelatihan dijadikan lebih besar ukurannya agar model dapat dilatih sebanyak mungkin sehingga dapat mengetahui secara akurat pada pengujian menggunakan data testing yang belum pernah dilihat sebelumnya.

3.5. Penerapan SVM

Pada penerapan *Support Vector Machine* (SVM) akan dilakukan pelatihan dan pengujian untuk menemukan parameter serta model terbaik. Terdapat beberapa paremeter yang digunakan yang diantaranya adalah kernel, nilai C, dan *gamma*.

Tabel 2. Parameter SVM

Kernel	Nilai C	Gamma
Linear	[0.1, 1, 10, 100]	-
Polynomial	[0.1, 1, 10, 100]	[‘scale’, ‘auto’, 0.1, 0.01, 0.001]
RBF	[0.1, 1, 10, 100]	[‘scale’, ‘auto’, 0.1, 0.01, 0.001]
Sigmoid	[0.1, 1, 10, 100]	[‘scale’, ‘auto’, 0.1, 0.01, 0.001]

Pada Tabel 2 dijelaskan masing-masing kernel yang digunakan, jumlah nilai C, dan jumlah nilai *gamma* yang akan digunakan sebagai *tuning hyperparameter*. Masing-masing rasio pembagian data akan ditemukan kernel terbaik yang akan dilatih menjadi model terbaik. Model tersebut akan diuji pada proses pengujian menggunakan data testing untuk dibandingkan dengan algoritma KNN

3.6. Penerapan KNN

Pada penerapan metode K-*Nearest Neighbors* (KNN) akan dilakukan pelatihan dan pengujian untuk menemukan nilai parameter yang terbaik. Nilai parameter yang akan digunakan pada penelitian ini terdiri dari nilai *k* dan parameter *weights*. Nilai *k* ditentukan dengan rentang nilai 1-10 dengan nilai weights *distance*, dan *uniform*. Perhitungan jarak *k* disamakan menggunakan *euclidean distance*.

Tabel 3. Parameter KNN

Weights	K
Uniform, Distance	1
	2
	3
	4
	5
	6
	7
	8
	9
	10

Pada Tabel 3 dijelaskan masing-masing parameter yang akan digunakan pada pelatihan dan pengujian. Model terbaik yang telah ditemukan dengan nilai *k* terbaik akan dievaluasi kembali kemudian akan diuji menggunakan data testing yang belum pernah dilihat sebelumnya. Hasil model terbaik pada pengujian tersebut akan dibandingkan dengan algoritma SVM.

3.7. Evaluasi Model

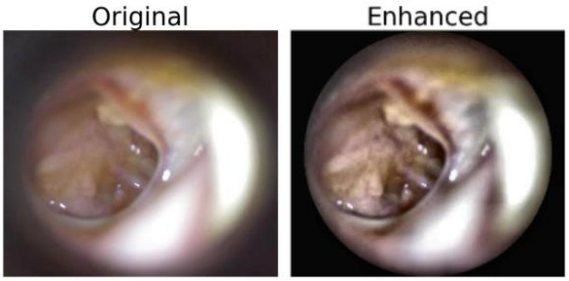
Untuk mengevaluasi model dengan parameter terbaik nya, model akan dievaluasi menggunakan *confusion matrix*. Evaluasi model akan dilakukan pada model yang telah melalui proses pengujian menggunakan data testing bukan terdapat pada pengujian model terbaik. Hal ini, bertujuan untuk mengetahui apakah model terbaik dari masing-masing algoritma dapat mengklasifikasi data yang belum pernah dibaca sebelumnya. Kinerja masing-masing model diukur dengan menggunakan nilai-nilai dari *confusion matrix* yaitu akurasi, sensitifitas, presisi, dan *F1-score*.

4. HASIL DAN PEMBAHASAN

Setelah melalui proses pada bab sebelumya, pada bab ini akan disajikan hasil penelitian dan pembahasan. Kedua performa model dievaluasi dan dibandingkan model mana yang memiliki performa lebih baik untuk memberikan rekomendasi penggunaan model dan peningkatan model pada penelitian selanjutnya.

4.1. Hasil Praproses

Praproses dilakukan untuk meningkatkan dan memaksimalkan citra agar dapat dibaca serta diekstraksi secara maksimal pada proses selanjutnya. Dari hasil praproses ini ditampilkan salah satu hasil dengan sampel dari kelas otitis media kronik (COM). Berikut merupakan hasil praproses yang telah dilakukan secara berurutan seperti yang telah dijelaskan pada bab sebelumnya.



Gambar 3. Sampel Kelas Otitis Media Kronik

Pada Gambar 3 merupakan sampel citra yang telah melalui praproses data. Gambar kiri menunjukkan citra asli yang masih terdapat *noise* dan latar belakang yang dapat mengganggu proses ekstraksi fitur. Gambar kanan merupakan citra *Region of Interest* (ROI) yang berhasil diekstrak serta berhasil diterapkan peningkatan kontras dan pengurangan *noise*.

4.2. Hasil Ekstraksi dan Normalisasi Fitur

Setelah melalui praproses, data citra kemudian dilakukan ekstraksi fitur untuk mendapatkan respon fitur. Fitur respon yang dihasilkan dihitung berdasarkan fitur MR8 dan ukuran citra. Berikut

merupakan hasil ekstraksi fitur yang dihasilkan yang menampilkan jumlah fitur pada setiap citra.

```
Shape of ROI images: (64, 64, 3)
Number of features per image: 98304
Training features: (720, 98304)
Testing features PCA: (160, 98304)
```

Gambar 4. Hasil Ekstraksi Fitur

Hasil ekstraksi fitur pada Gambar 4 ini terlalu besar apabila diproses secara langsung pada algoritma klasifikasi. Oleh karena itu, hasil tersebut distandarisasi untuk mendapatkan nilai *mean* = 0 dan nilai standar deviasi = 1. Proses standarisasi ini tidak merubah jumlah fitur yang telah dihasilkan.

```
Features per image after PCA: 474
Training features after PCA: (720, 474)
Testing features after PCA: (160, 474)
Holding Variance ratio: 0.9501823621922162
```

Gambar 5. Hasil Normalisasi Fitur

Pada Gambar 5 merupakan hasil yang didapatkan dari normalisasi fitur. Setelah dinormalisasi, fitur yang didapatkan sejumlah 474 dengan mempertahankan informasi fitur sebesar 95% dari total fitur.

4.3. Kinerja Model

Setelah melalui ekstraksi fitur, model dari masing-masing parameter *Support Vector Machine* (SVM) dan K-*Nearest Neighbor* (KNN) dilatih menggunakan data yang telah tersedia. Pada setiap rasio pembagian data, masing-masing model dilatih menggunakan parameter yang telah ditentukan kemudian diuji menggunakan set validasi. Dari hasil *tuning hyperparameter* ini, algoritma SVM dan KNN sama sama mendapatkan model terbaiknya pada rasio *split* data 75:25. Algoritma SVM mendapatkan model terbaik pada dengan kernel Linear dan mendapat akurasi sebesar 98,89%, sedangkan algoritma KNN mendapat nilai *k* terbaik dengan *k* = 1 dan akurasi sebesar 95,56%. Berikut pada Tabel 4 merupakan tabel *confusion matrix* untuk nilai model terbaik yang diperoleh dengan masing-masing algoritma.

Tabel 4. Hasil *Tuning Hyperparameter*

Metode	Akurasi	Presisi	Recall	F1-Score
SVM	98,89%	98,75%	98,75%	98,72%
KNN	95,56%	95,32%	95,81%	95,45%

Setelah mendapatkan model terbaik dari kedua algoritma, masing-masing model terbaik tersebut diuji menggunakan data testing. Model terbaik dari masing-masing algoritma klasifikasi dilatih ulang untuk memaksimalkan model dalam proses pelatihan data latih. Setelah proses pelatihan ulang, model tersebut diuji menggunakan data testing yang belum pernah dilihat oleh model klasifikasi sebelumnya. Dari hasil pengujian tersebut maka didapatkan hasil pengujian sebagai berikut.

Tabel 5. Hasil Final Kinerja Model

Metode	Akurasi	Presisi	Recall	F1-Score
SVM	87,50%	87,94%	87,50%	87,50%
KNN	72,50%	76,66%	72,50%	71,80%

Pada Tabel 5 merupakan performa model klasifikasi yang telah diuji menggunakan masing-masing model terbaik dengan data testing. Algoritma SVM mendapatkan akurasi sebesar 87,50% dan untuk algoritma KNN mendapatkan akurasi sebesar 72,50%. Model SVM menunjukkan tingkat akurasi yang lebih tinggi dibandingkan algoritma KNN dari segi akurasi, presisi, sensitivitas, dan *F1-Score*. Berikut pada Gambar 6 dan Gambar 7 merupakan laporan hasil klasifikasi dari kedua model algoritma.

SVM Classification Report:

	precision	recall	f1-score	support
Chronic otitis media	0.85	0.88	0.86	40
Earwax plug	0.97	0.93	0.95	40
Myringosclerosis	0.89	0.78	0.83	40
Normal	0.80	0.93	0.86	40
accuracy			0.88	160
macro avg	0.88	0.88	0.88	160
weighted avg	0.88	0.88	0.88	160

Gambar 6. Laporan Klasifikasi SVM

KNN Classification Report:

	precision	recall	f1-score	support
Chronic otitis media	0.70	0.75	0.72	40
Earwax plug	1.00	0.47	0.64	40
Myringosclerosis	0.67	0.88	0.76	40
Normal	0.70	0.80	0.74	40
accuracy			0.73	160
macro avg	0.77	0.73	0.72	160
weighted avg	0.77	0.72	0.72	160

Gambar 7. Laporan Klasifikasi KNN

5. KESIMPULAN DAN SARAN

Hasil pengujian yang telah dilakukan menunjukkan bahwa algoritma *Support Vector Machine* (SVM) mendapatkan skor akurasi 87,50% sedangkan algoritma K-*Nearest Neighbors* (KNN) mendapatkan skor akurasi 72,50%. Maka dari itu, dapat disimpulkan algoritma (SVM) menunjukkan tingkat keakuratan yang lebih tinggi dibandingkan algoritma (KNN). Implikasi hasil penelitian ini mengarahkan untuk penggunaan ekstraksi dan penambahan parameter pada masing-masing algoritma sehingga dapat meningkatkan performa kedua model. Dalam penggunaan praktik untuk mendeteksi infeksi telinga dapat direkomendasikan menggunakan algoritma SVM sebagai sistem yang mendukung citra otoskopi untuk mengetahui jenis infeksi yang dialami oleh pasien. Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi ekstraksi fitur, parameter masing-masing algoritma, dan skenario pengujian agar dapat mengetahui secara maksimal kedua algoritma dalam melakukan klasifikasi citra. Variasi kelas dan

dataset penyakit juga dapat menjadi pertimbangan untuk mendeteksi penyakit telinga lainnya sehingga sistem yang dibuat dapat bermanfaat secara keseluruhan dalam mendeteksi infeksi telinga.

DAFTAR PUSTAKA

[1] I. F. Martanegara, Wijana, dan S. Mahdiani, "Tingkat pengetahuan kesehatan telinga dan pendengaran siswa SMP di Kecamatan Muara Gembong Kabupaten Bekasi," *JSK J. Sist. Kesehat.*, vol. 5, no. 4, hal. 140–147, 2020, [Daring]. Tersedia pada: https://journal.unpad.ac.id/jsk_ikm/article/view/31281

[2] T. Tesfa *et al.*, "Bacterial otitis media in sub-Saharan Africa: A systematic review and meta-analysis," *BMC Infect. Dis.*, vol. 20, no. 1, hal. 1–12, 2020, doi: 10.1186/s12879-020-4950-y.

[3] D. A. Anggoro, "Comparison of Accuracy Level of Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) Algorithms in Predicting Heart Disease," *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 5, hal. 1689–1694, 2020, doi: 10.30534/ijeter/2020/32852020.

[4] M. Viscaino, J. C. Maass, P. H. Delano, M. Torrente, C. Stott, dan F. Auat Cheein, "Computer-aided diagnosis of external and middle ear conditions: A machine learning approach," *PLoS One*, vol. 15, no. 3, 2020, doi: 10.1371/journal.pone.0229226.

[5] H. Bingol, "Classification of OME with Eardrum Otoendoscopic Images Using Hybrid-Based Deep Models, NCA, and Gaussian Method," *Trait. du Signal*, vol. 39, no. 4, hal. 1295–1302, Agu 2022, doi: 10.18280/TS.390422.

[6] T. A. Assegie, "Support Vector Machine And K-Nearest Neighbor Based Liver Disease Classification Model," *Indones. J. Electron. Electromed. Eng. Med. informatics*, vol. 3, no. 1, hal. 9–14, 2021, doi: 10.35882/ijeeemi.v3i1.2.

[7] S. Ghosh, A. Dasgupta, dan A. Swetapadma, "A study on support vector machine based linear and non-linear pattern classification," *Proc. Int. Conf. Intell. Sustain. Syst. ICISS 2019*, no. Iciss, hal. 24–28, 2019, doi: 10.1109/ISS1.2019.8908018.

[8] D. A. Pisner dan D. M. Schnyer, *Support vector machine*. Elsevier Inc., 2019. doi: 10.1016/B978-0-12-815739-8.00006-7.

[9] K. Taunk, S. De, S. Verma, dan A. Swetapadma, "A Brief Review of Nearest Neighbor Algorithm for Learning and Classification," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, 2019, hal. 1255–1260. doi: 10.1109/ICCS45141.2019.9065747.

[10] H. Taufiqqurrahman, F. Tri Anggraeny, dan M. Muharrom Al Haromainy, "Perbandingan Algoritma Naïve Bayes Dan K-Nearest Neighbor Pada Analisis Sentimen Ulasan Aplikasi Mypertamina," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, hal. 3934–3939, 2024, doi: 10.36040/jati.v7i6.7801.

[11] A. PS, A. Nugroho Sihananto, dan D. Arman Prasetya, "Implementasi Metode K-NN dalam Klasterisasi Kasus Kesehatan Jantung," *ALINIER J. Artif. Intell. Appl.*, vol. 3, no. 2, hal. 18–21, 2022, doi: 10.36040/alinier.v3i2.5761.

[12] F. T. Anggraeny, M. S. Munir, dan I. Y. Purbasari, "Histogram Profil Proyeksi sebagai Metode Ekstraksi Fitur pada Pengenalan Karakter Tulisan Tangan," *Pros. Semin. Nas. Inform. Bela Negara*, vol. 1, hal. 164–168, 2020, doi: 10.33005/santika.v1i0.44.

[13] C. Bhowmick, P. K. Dutta, dan M. Mahadevappa, "Computer Aided Classification of Benign and Malignant Breast Lesions using Maximum Response 8 Filter Bank and Genetic Algorithm," in *2020 IEEE REGION 10 CONFERENCE (TENCON)*, 2020, hal. 43–46. doi: 10.1109/TENCON50793.2020.9293928.

[14] D. Krstinić, M. Braović, L. Šerić, dan D. Božić-Štulić, "Multi-label Classifier Performance Evaluation with Confusion Matrix," hal. 01–14, 2020, doi: 10.5121/csit.2020.100801.

.