

Optimasi Kinerja Algoritma SVM Dan LDA Pada *Detection* Citra Penyakit Mata Katarak

Finis Hermanto Laia¹

^aProgram Studi Ilmu Komputer, Fakultas Sains dan Teknologi Universitas Nias Raya
Jalan Pramuka, Nari-Nari, Kel. Pasar Telukdalam, Kab. Nias Selatan, Provinsi Sumatera Utara

¹finishermanto@email.com (Corresponding author)

Abstract

The eye is a visual organ that functions to capture light and convert it into signals that are processed by the brain to form visual perception. Cataracts are a condition in which the eye lens becomes cloudy, blocking light from entering, and causing visual impairment. Early detection of cataracts is essential to prevent or slow down vision loss. In this study, the performance optimization of the Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) algorithms was carried out in detecting cataracts from eye images using digital image processing techniques. Digital image processing is used to improve image quality, extract information, and perform classification with machine learning. LDA is used as a dimensionality reduction technique to improve classification efficiency, while SVM is used to find the optimal hyperplane that separates data with high accuracy. Several studies have combined LDA and SVM in the classification process to improve system performance. The results showed that SVM is superior to LDA in terms of accuracy, recall, and F1-score. The highest accuracy of SVM reached 95.98%, while LDA was only 90.20%. Both algorithms have 100% precision, but SVM recall is higher (92.0%) than LDA (79.3%). The F1-score of SVM is also better (90.9%) than LDA (84.0%), indicating an optimal balance between precision and recall. Thus, SVM is more recommended for cataract detection than LDA due to its higher accuracy and recall, while LDA is more suitable for classification tasks with lower complexity.

Keywords: *Cataract, Image Processing, Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Machine Learning*

Abstrak

Mata merupakan organ penglihatan yang berfungsi menangkap cahaya dan mengubahnya menjadi sinyal yang diproses oleh otak untuk membentuk persepsi visual. Katarak adalah kondisi di mana lensa mata mengalami kekeruhan menghambat cahaya masuk dan menyebabkan gangguan penglihatan. Deteksi dini katarak sangat penting untuk mencegah atau memperlambat kehilangan penglihatan. Dalam penelitian ini, dilakukan optimasi kinerja algoritma Support Vector Machine (SVM) dan Linear Discriminant Analysis (LDA) dalam mendeteksi citra katarak dari citra mata menggunakan teknik digital image processing. Digital image processing digunakan untuk meningkatkan kualitas citra mengekstrak informasi, dan melakukan klasifikasi identifikasi dengan machine learning. LDA digunakan sebagai teknik reduksi dimensi untuk meningkatkan efisiensi klasifikasi, sementara SVM digunakan untuk mencari hyperplane optimal yang memisahkan data dengan akurasi tinggi. Beberapa penelitian telah menggabungkan LDA dan SVM dalam proses klasifikasi untuk meningkatkan performa machine. Hasil penelitian menunjukkan bahwa SVM lebih unggul dibandingkan LDA dalam hal akurasi, recall, dan F1-score. Akurasi tertinggi SVM mencapai 95,98%, sedangkan LDA hanya 90,20%. Kedua algoritma memiliki presisi 100%, namun recall SVM lebih tinggi (92,0%) dibandingkan LDA (79,3%). F1-score SVM juga lebih baik (90,9%) dibandingkan LDA (84,0%), menunjukkan keseimbangan optimal antara presisi dan recall. Dengan demikian, SVM lebih direkomendasikan untuk deteksi katarak dibandingkan LDA karena akurasi dan recall yang lebih tinggi, sementara LDA lebih cocok untuk tugas klasifikasi dengan kompleksitas lebih rendah.

Kata Kunci: *Katarak, Image Processing, Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Machine Learning*

1. Pendahuluan

Mata adalah organ penglihatan yang berfungsi menangkap cahaya dan mengubahnya menjadi sinyal yang di proses oleh otak untuk membentuk persepsi visual. Katarak adalah kondisi dimana lensa mata mengalami kekeruhan sehingga menghambat cahaya masuk dan menyebabkan gangguan penglihatan. Algoritma *machine learning* telah menjadi sangat populer dan terus diteliti secara luas. Support Vector Machine (SVM) adalah salah satu teknik dari machine learning algoritma klasifikasi yang memiliki sifat supervised learning (pembelajaran) yang bekerja dengan mencari hyperplane (batas keputusan) optimal yang memisahkan jarak antar kelas, generalisasi dan akurasi klasifikasi yang stabil [5]. Linear Discriminant Analysis (LDA) adalah teknik reduksi dimensional yang dilakukan untuk tahap pengklasifikasian pola dan aplikasi machine learning. LDA digunakan untuk mendapatkan fitur citra dan memberikan jarak antar kelas yang lebih besar, sedangkan jarak antara data latih dalam suatu kelas menjadi lebih kecil [6]. Deteksi dini penyakit mata sangat penting, terutama bagi individu yang memiliki riwayat penyakit mata dalam keluarga orang yang berusia di atas 60 tahun atau kurang dari 60 tahun tergantung dari penyakit mata yang di alami oleh setiap individu. Mendeteksi penyakit mata sejak dini sangat penting untuk mencegah atau memperlambat perkembangan kehilangan penglihatan dan kebutaan [1].

Digital image processing adalah proses teknik komputasi untuk meningkatkan kualitas citra digital, mengekstrak informasi dari lensa mata yang dapat digunakan dalam identifikasi klasifikasi katarak. Citra dari sudut pandang matematis, ialah fungsi menerus ataupun continue dari intensitas sinar pada bidang 2 dimensi $f(x, y)$, dengan x serta y yakni koordinat spasial serta amplitude f pada pendamping koordinat (x, y) yang diartikan sebagai intensitas ataupun derajat keabuan citra pada titik [2]. Pada era teknologi yang saat ini semakin maju dalam bidang kesehatan termasuk penggunaan peralatan medis dan teknologi informasi dapat meningkatkan kualitas dan efisiensi layanan kesehatan seperti digital image processing digunakan dalam MRI, CT scan, dan deteksi penyakit berbasis citra, termasuk deteksi katarak dari citra mata menggunakan algoritma machine learning [3]. Algoritma machine learning telah diterapkan untuk menyelesaikan berbagai permasalahan dan banyak digunakan dalam bidang sains. Dalam implementasinya, terdapat berbagai algoritma yang dapat digunakan untuk tugas klasifikasi, di antaranya Support Vector Machine (SVM) dan Convolutional Neural Network (CNN) [4].

Beberapa peneliti secara umum menggabungkan penerapan SVM dengan LDA. Dimana LDA digunakan sebagai salah satu metode Feature Reduction untuk mengurangi kompleksitas pada SVM. Namun, beberapa peneliti seperti [5] telah menggunakan LDA sebagai metode pengolahan citra yang cukup baik dengan waktu komputasi yang baik pada pengenalan citra. Melakukan penelitian mengenai metode mana yang lebih baik di antara High Dimensional Features dan Low Dimensional Features dan memperoleh hasil bahwa metode mana yang lebih baik adalah tergantung kepada ukuran dataset yang ada dan juga ukuran data training dan data testing yang digunakan. Hal ini juga sama dengan penelitian mengenai penggunaan metode Linear Discriminant Analysis untuk pengenalan wajah dengan membandingkan dengan banyaknya data latih dan penelitian oleh [9] mengenai pengaruh ukuran dataset pelatihan untuk kinerja Support Vector Machine dan Decision Tree. Menarik untuk diamati bagaimana optimasi kinerja antara SVM dan LDA di dalam deteksi penyakit citra mata katarak dan mata normal khususnya dengan persentase jumlah data dan persentase data training dan data testing. Pengenalan penyakit mata citra digital dilakukan konversi dari rgb ke grayscale kemudian tahap preprocessing, segmentasi, ekstraksi ciri dan identifikasi deteksi penyakit mata serta manakah hasil akurasi tertinggi dari kedua algoritma machine. Oleh karena itu, peneliti di dalam penelitian ini tertarik untuk melakukan optimasi kinerja algoritma LDA dan SVM pada deteksi dini penyakit mata katarak khususnya terkait dengan persentase dataset dan juga persentase data training dan data testing.

2. Metode Penelitian

Pada penelitian ini dilakukan optimasi kinerja kedua dari algoritma identifikasi pada citra mata katarak normal dan yang terkena penyakit mata katarak. Penelitian ini akan menghasilkan manakah hasil akurasi tertinggi dalam mengenali identifikasi suatu citra dari hasil akurasi tertinggi dideteksi sebagai citra mata normal dan mata katarak, pada saat implementasi kedua algoritma LDA dan SVM. Gambar berikut Menunjukkan alur metodologi penelitian secara umum Figure 1.

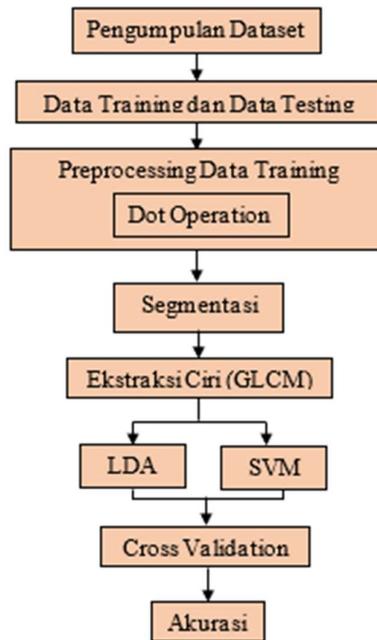


Figure 1. Metodologi Penelitian

A. Data Collection

Data yang digunakan dalam penelitian ini data sekunder bersumber dari dataset largest collection of open source computer vision datasets yang bisa ditemukan di alamat situs web kaggle.com Mata Katarak Image Dataset, penulis kumpulkan dengan data training berjumlah 150 dan data testing berjumlah 150 citra diantaranya, kelas penyakit mata katarak dan mata normal. Berikut gambar menunjukkan sampel satu dataset citra yang berpenyakit mata katarak dan mata normal

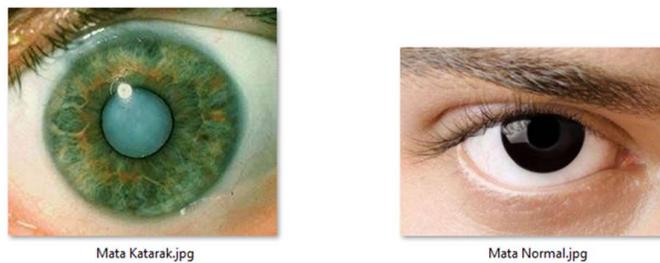


Figure 2. Citra Mata Katarak Dan Mata Normal

Adapun data lengkapnya dapat dilihat pada tabel 1 variasi pembuatan data training dan testing sebagai berikut:

Table 1. Persentase data Training Dan Data Testing

No	Variasi Dataset	Data Training	Data Testing	Total
1.	50% : 50%	150 citra	150 citra	300 citra
2.	60% : 40%	180 citra	120 citra	

B. Preprocessing Data Training

Tahap preprocessing citra pada pengolahan yaitu perbaikan kualitas citra (Image Enhancement) yang terdiri dari dot operation dan special operation dan proses resize dilakukan untuk menyeragamkan ukuran data set yang sama. Citra tersebut dilakukan perbaikan kualitas citra ke salah satu dot operation dengan teknik histogram Equalization. Berikut citra grayscale yang telah diperbaiki kualitas ke citra Intensity Adjustment Image ditunjukkan pada figure 3. tahap preprocessing dataset :

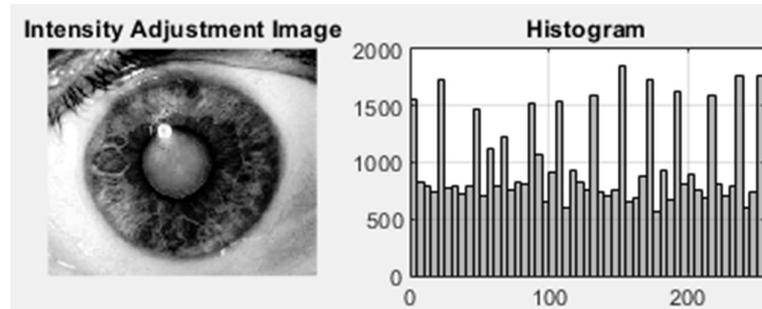


Figure 3. (a) Adjustment Image (b) Histogram

Pada figure 3 Perbaikan kualitas citra dari grayscale ke Intensity Adjustment Image (Histogram) bertujuan untuk meningkatkan perbedaan antara nilai intensitas pixel pada citra serta membantu memudahkan membedakan objek pada citra seperti mengurangi noise, meningkatkan visibilitas fitur, untuk proses Identifikasi. Pada tahap ini nilai intensitas pixel pada citra diubah menjadi nilai intensitas yang lebih tinggi sehingga objek pada citra lebih terlihat dan proses segmentasi menjadi lebih mudah.

C. Segmentasion

Pada tahap segmentasi memisahkan objek atau fitur pada citra, nilai intensitas pixel pada citra diubah menjadi threshold berukuran nilai ambang (128) dengan tujuan membagi citra yang dikategorikan menjadi objek dan latar belakang sebelum di konversi dari citra thresholding ke citra biner. Citra biner dihasilkan dari proses thresholding untuk mempermudah proses segmentasi. Citra biner memiliki dua warna yang kontras, sehingga objek pada citra lebih mudah untuk dipisahkan dari latar belakang. Berikut ini hasil citra mata katarak dari konversi nilai citra thresholding ke biner dapat dilihat pada gambar sebagai berikut:

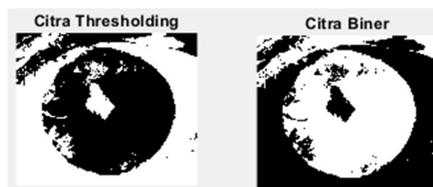


Figure 4. Citra Thresholding Konversi ke Biner

D. Ekstraksi Ciri Gray Level Co-occurrence Matrix (GLCM)

Proses ekstraksi ciri atau pengambilan ciri dari sebuah objek citra, hasil tersebut sebagai nilai masukan proses klasifikasi tahap selanjutnya. Hasil ekstraksi ciri yang digunakan dalam penelitian ini yaitu menggunakan ekstraksi ciri tekstur (GLCM) berdasarkan nilai contrast, correlation dan energy, untuk membedakan objek dengan tekstur tertentu dapat menggunakan nilai contrast yang digunakan agar dapat membedakan objek dengan tekstur halus dan kasar dan nilai correlation membedakan tekstur yang memiliki arah teratur dan tidak teratur pada pola, sedangkan nilai energy membedakan objek dengan tekstur berulang dan tidak berulang dari objek dalam suatu area.

Ciri Tekstur	Nilai
Contrast	0.042281
Correlation	0.91546
Energy	0.45956

Figure 5. Ekstraksi Ciri GLCM

Berikut ini hasil dari ekstraksi ciri tekstur GLCM Presentase data latih dan data testing sebanyak 300 citra, dapat dilihat pada Tabel dibawah ini split hasil nilai-nilai fitur pada GLCM:

Table 2. Ekstraksi Ciri Tekstur GLCM

Data Testing				
No.	Nama File	Contrast	Correlation	Energy
1.	'katarak1.jpg'	0.042281	0.91546	0.45956
2.	'katarak2.jpg'	0.35421	0.97702	0.65626
130.	'katarak130.jpg'	0.27252	0.87622	0.76542
150.	'katarak150.jpg'	0.19782	0.24311	0.42311
Data Training				
1.	'normal1.jpg'	0.87252	0.87622	0.76542
2.	'normal2.jpg'	0.09788	0.24311	0.42311
149.	'normal149.jpg'	0.042281	0.92546	0.45956
150.	'normal150.jpg'	0.35421	0.87702	0.15626

Pada table 2 hasil dari ekstraksi ciri tekstur menggunakan Gray Level Co-occurrence Matrix (GLCM) pada dua set data, data latih(training) dan data uji (testing). glcm digunakan untuk menganalisis tekstur dalam citra dengan memeriksa hubungan spasial antara intensitas piksel. Pada data testing, pola yang serupa dapat dilihat. Citra 'normal149.jpg' memiliki nilai kontras yang lebih rendah dan nilai korelasi yang lebih tinggi dibandingkan dengan citra 'katarak2.jpg'. Hasil yang didapat digunakan untuk melatih model klasifikasi untuk membedakan antara citra mata katarak dan mata normal berdasarkan ekstraksi ciri tekstur GLCM

3. Hasil dan Pembahasan

Implementasi software uji dan evaluasi kinerja terhadap penelitian yang dilakukan menggunakan bahasa pemrograman Matlab untuk identifikasi pada penyakit mata katarak dan mata normal. Penerapan algoritma dengan menggunakan LDA dan SVM dalam penelitian ini melakukan variasi persentase yang telah di tetapkan dengan dataset data latih dan data uji.

3.1. Tampilan Hasil Gui Aplikasi Matlab

Untuk optimasi model dari kedua algoritma yang telah dibangun, langkah pertama adalah menginput citra mata katarak dari data uji. Selanjutnya, citra tersebut melalui tahap preprocessing untuk meningkatkan kualitasnya sebelum dilakukan segmentasi. Setelah itu, fitur tekstur citra diekstraksi menggunakan metode GLCM dengan parameter contrast, correlation, dan energy. Langkah berikutnya adalah mengidentifikasi citra menggunakan metode lda dan svm. Berikut adalah hasil pengujian pada salah satu citra mata katarak.

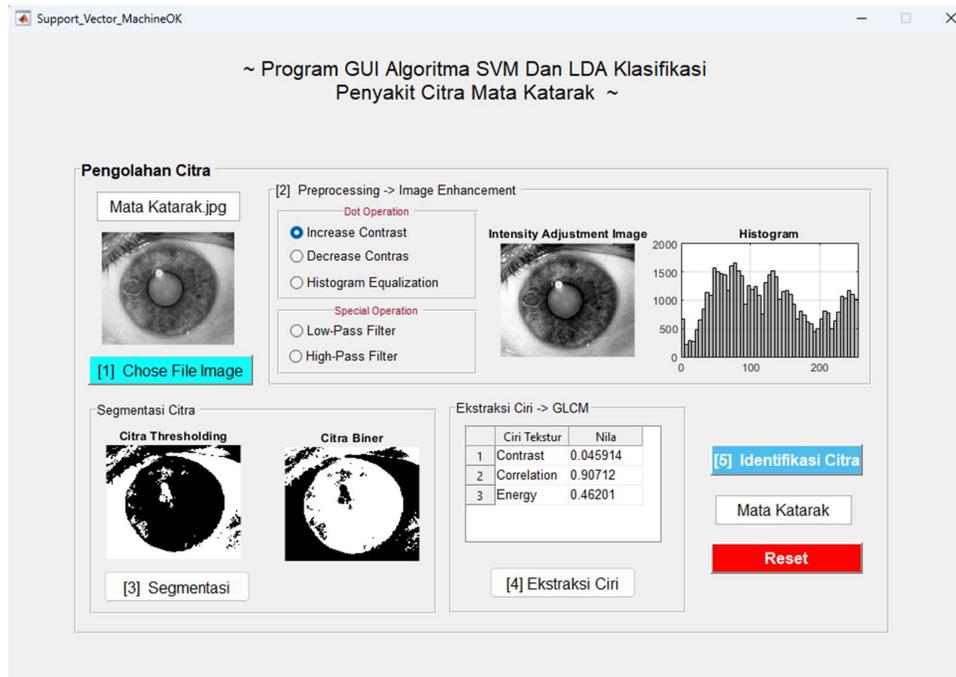


Figure 6. Gui Identifikasi Aplikasi Matlab

Pengujian dilakukan dengan menggunakan data uji sebanyak 150, 120 citra. Terdapat dua kelas, yaitu kelas mata katarak dan mata normal. Pada figure hasil identifikasi citra penyakit mata katarak. Semua citra uji akan dibandingkan dengan hasil detection untuk mengevaluasi performance suatu algoritma akurasi tertinggi.

3.1.1. Hasil Cross Validation

Algoritma yang digunakan untuk menguji validitas hasil akurasi adalah k=3 Cross Validation. Dataset dibagi menjadi 3 fold, dengan total 270 dari data testing. Pada setiap iterasi, pemilihan dataset uji dilakukan sesuai dengan urutan fold. Proses pelatihan (training) dilakukan sebanyak dua kali dengan persentase rasio data: 50:50, 60:40. Setelah setiap sesi pelatihan selesai, langsung dilakukan pengujian (testing) untuk memperoleh nilai prediksi. Tingkat akurasi kemudian dihitung sebagai rata-rata dari semua iterasi. Hasil pengujian Cross Validation dapat dilihat pada tabel berikut.

Table 3. Hasil Cross Validation K-3

Jenis Model ML	Iterasi Ke	Testing n-1%	Testing n-2%	Testing n-3%	Rata-Rata
Support Vector Machine (SVM)	1	88,67	98,0	89,0	91,89
	2	85,87	88,0	97,0	90,29
	3	88,78	97,0	88,0	91,26
	Rata-Rata	87,77	94,33	91,33	91,14
Jenis Model ML	Iterasi Ke	Testing n-1%	Testing n-2%	Testing n-3%	Rata-Rata
Linear Discriminant Analisis (LDA)	1	80,01	79,55	81,00	80,18
	2	78,32	88,59	76,03	80,98
	3	72,33	82,64	92,00	82,32

	Rata-Rata	76,88	83,59	83,01	81,16
--	-----------	-------	-------	-------	-------

Berdasarkan Table 3. hasil rata-rata pengujian menggunakan Linear Discriminant Analysis (LDA) menunjukkan bahwa prediksi tertinggi diperoleh pada iterasi ke-3 dengan persentase data uji 40%, mencapai 82,32%. Sementara itu, nilai rata-rata terendah ditemukan pada iterasi ke-1 dengan persentase data uji 50%, yaitu 80,18%. Secara keseluruhan, rata-rata akurasi prediksi sistem menggunakan LDA adalah 81,16%. Sedangkan untuk pengujian menggunakan Support Vector Machine (SVM), rata-rata prediksi tertinggi tercapai pada iterasi ke-1 dengan persentase data uji 40%, sebesar 91,89%. Nilai akurasi terendah ditemukan pada iterasi ke-2, yaitu 90,29%. Secara keseluruhan, rata-rata akurasi prediksi sistem menggunakan SVM adalah 91,14%, diperoleh berdasarkan hasil setiap rata-rata dari jumlah iterasi, berikut gambar grafik perbandingan akurasi data testing.

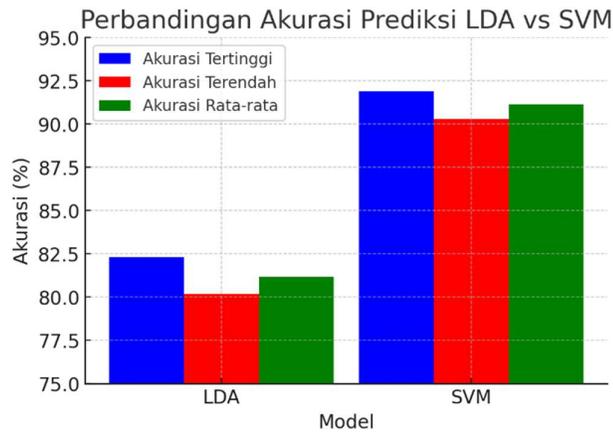


Figure 7. Grafik Perbandingan Akurasi Pengujian Metode K-Fold Cross

3.2. Hasil Akurasi

Evaluasi hasil dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score pada data latih dan uji untuk mengukur efektivitas serta performa kedua algoritma, dalam mengidentifikasi citra mata katarak dan mata normal. Berikut adalah hasil akurasi berdasarkan persentase data pengujian dari total dataset: Persentase pertama (50%) : 150 citra uji dari 300 total citra. Persentase kedua (40%) : 120 citra uji dari total 300 citra.

Table 4. Hasil Akurasi Data Testing Kedua Algoritma

Jenis Model ML	Variasi Persentase	Jumlah Data Testing	Akurasi	Precision	Recall	F1-Score
Linear Discriminant Analysis (LDA)	Ke-1	150	90,66%	100%	79,3%	84,0%
	Ke-2	120	65,78%	100%	80,0%	89,0%
Support Vector Machine (SVM)	Ke-1	150	95,98%	100%	92,0%	90,9%
	Ke-2	120	78,88%	100%	80,3%	90,0%

Pada data testing akurasi algoritma SVM lebih tinggi dari LDA pada variasi persentase ke-1 dan ke-2, LDA memberikan hasil pada variasi persentase ke-1 dengan akurasi 90,20%, precision 100%, recall 79,3%, dan F1-Score 84,0%. SVM memberikan hasil terbaik pada variasi persentase ke-1 dengan akurasi 95,98%, precision 100%, recall 92,0%, dan F1-

Score 90,9% dimana SVM cenderung lebih baik. Berikut gambar visualisasi grafik akurasi proses data testing kedua algoritma (LDA dan SVM) berdasarkan variasi persentase data, berikut gambar grafik dari kedua algoritma variasi persentase testing.

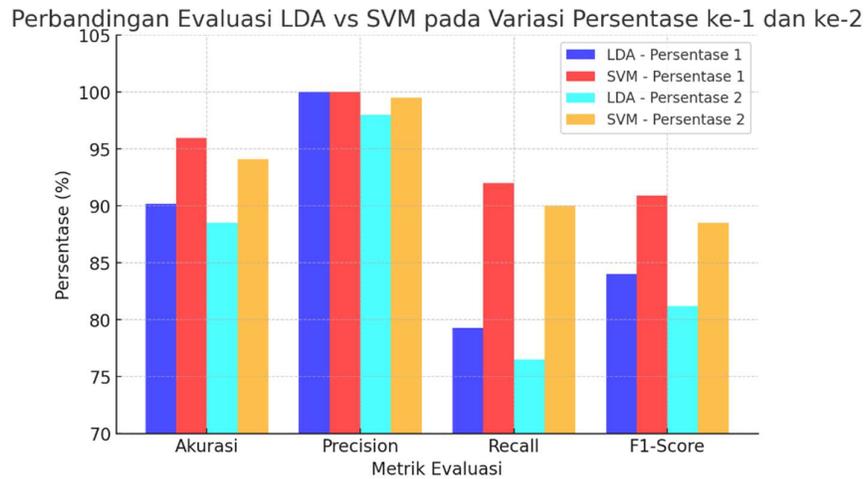


Figure 8. Grafik Variasi Persentase Testing LDA dan SVM

Dari hasil penelitian menunjukkan SVM lebih efektif dalam detection citra mata katarak atau mata normal berbanding dengan optimasi kinerja LDA. Hasil pada tahap proses testing data variasi persentase ke-1 (50%), variasi persentase ke-2 (40%), performa LDA menurun secara signifikan sementara SVM tetap relatif stabil, namun dengan nilai precision mencapai 100% kedua algoritma.

4. Kesimpulan

Dari hasil penelitian optimasi kinerja algoritma Support Vector Machine (SVM) dan Linear Discriminant Analysis (LDA) dalam mendeteksi katarak dari citra mata sebagai berikut:

1. Perbandingan Kinerja SVM lebih unggul dibandingkan LDA dalam hal akurasi, recall, dan F1-score. Akurasi tertinggi SVM mencapai 95,98%, sedangkan LDA hanya 90,20%.
2. Presisi Kedua algoritma mencapai presisi 100%, menunjukkan kemampuan dalam mengidentifikasi kasus katarak tanpa kesalahan positif.
3. Recall SVM memiliki recall lebih tinggi (92,0%) dibandingkan LDA (79,3%), sehingga lebih efektif dalam mendeteksi citra mata katarak.
4. F1-score SVM (90,9%) lebih baik dibandingkan LDA (84,0%), menunjukkan keseimbangan antara presisi dan recall yang lebih optimal.

Berdasarkan hasil optimasi kinerja dari matrik evaluasi SVM lebih baik untuk deteksi katarak karena akurasi dan recall yang lebih tinggi. LDA lebih sesuai untuk tugas klasifikasi dengan kompleksitas lebih rendah.

References

- [1] S. Yonal, "Indonesian Journal on Data Science Penerapan Algoritma Support Vector Machine (SVM)," vol. 2, no. 2, pp. 87–97, 2024.
- [2] A. Febriyanti and A. Alamsyah, "Peningkatan Akurasi Algoritma Support Vector Machine Menggunakan Dual-Tree Complex Wavelet Transform Pada Klasifikasi Citra Katarak dan Normal," Indones. J. Math. Nat. Sci., vol. 45, no. 2, pp. 51–58, 2022, doi: 10.15294/ijmns.v45i2.39723.
- [3] F. Adryansyah and A. Susilo, "Technology Information and Data Analytic Optimalisasi Algoritma YOLOv5 untuk," vol. 1, no. 2, pp. 63–68, 2024

- [4] A. Rusdy Prasetyo, Sussi, and B. Aditya, "Analisis Perbandingan Algoritma Support Vector Machine (Svm) Dan Convolutional Neural Network (Cnn) Untuk Sistem Deteksi Katarak," *J. Ilm. Tek. Mesin, Elektro dan Komput.*, vol. 3, no. 1, pp. 1–10, 2023, doi: 10.51903/juritek.v3i1.604
- [5] Z. He, M. Wu, X. Zhao, S. Zhang, and J. Tan, "Representative null space LDA for discriminative dimensionality reduction," *Pattern Recognit.*, vol. 111, p. 107664, 2021, doi: 10.1016/j.patcog.2020.107664
- [6] F. Last name, F. M. Last name and F. M. Last name, "Paper Title" *Journal name*, vol. Volume, no. Issue, p. Page Number, Year.
- [7] C. Li, Y. Shao, W. Chen, Z. Wang, and N. Deng, "Generalized two-dimensional linear discriminant analysis with regularization," *Neural Networks*, vol. 142, pp. 73–91, 2021, doi: 10.1016/j.neunet.2021.04.030.
- [8] R. D. Yunita, C. Rozikin, and M. Jajuli, "Implementasi Metode Linear Discriminan Analysis Untuk Klasifikasi Biji Kopi Abstrak," vol. 8, no. 1, pp. 27–39, 2022.
- [9] Y. Wei, K. Gu, and L. Tan, "A positioning method for maize seed laser-cutting slice using linear discriminant analysis based on isometric distance measurement," *Inf. Process. Agric.*, vol. 9, no. 2, pp. 224–232, 2022, doi: 10.1016/j.inpa.2021.05.002.