

Optimasi Support Vector Machine Dengan PSO Untuk Klasifikasi Kelayakan Export Kerang Batik

Optimization of Support Vector Machine with PSO for Classification of Batik Shells Export Feasibility

Bagus Prindo Sugihartono Putro¹, M Arief Soeleman², Pujiono³
Magister Teknik Informatika, Universitas Dian Nuswantoro

E-mail: ¹bagusprindo96@gmail.com, ²m.arief.soeleman@dsn.dinus.ac.id,
³pujiono@dsn.dinus.ac.id

Abstrak

Kerang Batik (*Paphia undulata*) memiliki pola cangkang yang mirip batik, dengan warna dasar cangkang yang bervariasi dari kuning cerah hingga gelap. Sebagai komoditas ekspor Indonesia yang permintaannya terus meningkat, penting untuk menjaga standar kualitas tinggi agar kerang siap ekspor. Penelitian ini menyelidiki metode kontrol kualitas yang efektif untuk kerang batik yang layak ekspor dengan mengambil sampel dari perusahaan terkait. Setelah proses pra-pemrosesan citra, dilakukan ekstraksi fitur, termasuk fitur bentuk (eccentricity, metric) dan fitur tekstur (GLCM). Fitur-fitur ini digunakan dalam algoritma SVM (Support Vector Machine) dengan kernel RBF, yang dipilih karena kemampuannya menangani data non-linear, untuk mencapai akurasi optimal. Metode optimasi PSO (Particle Swarm Optimization) juga diterapkan untuk meningkatkan akurasi lebih lanjut. Penelitian menunjukkan bahwa SVM dengan kernel RBF mencapai akurasi tertinggi sebesar 96,43% pada sudut 45° dan 90°. Setelah dioptimalkan dengan PSO, akurasi meningkat menjadi 97,86% pada sudut 90°. Dengan demikian, penerapan PSO pada algoritma SVM secara signifikan meningkatkan akurasi klasifikasi.

Kata kunci: SVM, PSO, Kerang Batik, Kernel RBF, GLCM

Abstract

Batik Shell (Paphia undulata) features a shell pattern resembling batik, with a base color ranging from bright yellow to dark hues. As a growing Indonesian export commodity, maintaining high-quality standards for export is crucial due to increasing international demand. This study investigates an effective quality control method for export-worthy Batik shells by sampling from relevant companies. Image preprocessing is followed by feature extraction, including shape features (eccentricity, metric) and texture features (GLCM). These features are used in an SVM (Support Vector Machine) algorithm with an RBF kernel, chosen for its effectiveness with non-linear data, to achieve optimal accuracy. PSO (Particle Swarm Optimization) is additionally applied to further enhance accuracy. The study finds that SVM with RBF kernel achieved the highest accuracy of 96.43% at 45° and 90° angles. After PSO optimization, the accuracy improved to 97.86% at a 90° angle. Thus, integrating PSO into the SVM algorithm significantly enhances classification accuracy.

Keywords: SVM, PSO, Batik Shell, RBF Kernel, GLCM

1. PENDAHULUAN

Kerang merupakan sejenis moluska invertebrata, hewan lunak dengan daging tersembunyi di balik cangkang keras yang hidup di laut dan di pantai berpasir. Cangkang Bentuk tubuhnya simetris dan ukuran cangkang seimbang di kedua sisi. Kerang Batik (*Paphia undulata*) memiliki pola warna cangkang yang mirip dengan karakteristik menyerupai batik, yang warna cangkang dasarnya kuning muda dan relatif gelap, dengan panjang cangkang maksimum: 6,5 cm [1].

Pendeteksian kelayakan kerang batik dapat membantu perusahaan menghindari proses penyortiran yang memakan waktu lama serta meningkatkan kualitas hasil produksi. Oleh karena itu, diperlukan sistem deteksi kelayakan kerang batik yang akurat dan menggunakan teknologi yang tepat. Kelayakan kerang batik dapat ditentukan dengan mengidentifikasi ciri-ciri tertentu. Untuk menjaga kualitas produk, diperlukan seleksi antara kerang yang layak dan tidak layak untuk diekspor. Kerang batik dikategorikan layak ekspor jika memenuhi kriteria berikut: (1) cangkang kerang utuh, (2) cangkang kerang tidak terbuka, (3) ukuran kerang tidak terlalu kecil, dan (4) corak cangkang kerang terlihat. Sebaliknya, kerang dianggap tidak layak ekspor jika: (1) cangkang kerang pecah, (2) cangkang kerang terbuka, (3) ukuran kerang terlalu kecil, dan (4) corak cangkang kerang tidak terlihat. Meskipun banyak upaya yang dilakukan untuk menjaga kualitas kerang batik, deteksi dini tetap menjadi tantangan yang signifikan.

Pengolahan citra merupakan proses manipulasi dan analisis citra digital untuk memperoleh informasi atau kualitas gambar yang diinginkan termasuk meningkatkan kualitas gambar, mengekstraksi informasi, dan melakukan analisis lebih lanjut, seperti deteksi objek menggunakan model normalisasi warna RGB, meliputi pemrosesan warna, pengurangan noise, deteksi tepi, dan segmentasi [2].

Ekstraksi fitur GLCM (*Gray Level Co-Occurrence Matrix*) efektif untuk mendeteksi luas permukaan kerang berdasarkan tingkat keabuan. Metode ini digunakan untuk menganalisis rentang warna 1 hingga 255 dan mengukur kejadian piksel berdekatan dengan jarak dan orientasi tertentu. GLCM menggunakan empat sudut utama: 0° , 45° , 90° , dan 135° , serta menghasilkan fitur seperti Dissimilarity, Correlation, Homogeneity, Contrast, dan Energy [3].

Algoritma SVM menyelesaikan masalah klasifikasi dengan menggunakan hyperplane optimal untuk memisahkan kelas, yang mengandalkan contoh pelatihan yang disebut support vector. SVM awalnya merupakan classifier linear, namun dikembangkan untuk menangani masalah non-linier dengan trik kernel yang mengubah ruang dimensi tinggi [4].

Particle Swarm Optimization (PSO) adalah algoritma pencarian global yang didasarkan pada kerja sama kelompok, meniru perilaku mencari makan burung dan ikan. Algoritma ini menginisialisasi partikel secara acak, kemudian memperbarui setiap generasi partikel untuk mencari solusi optimal dengan bergerak menuju posisi terbaik secara individu dan global dalam ruang berdimensi tinggi [5].

Pada penelitian sebelumnya yang telah dilakukan oleh Manoj Kumar Behera et al yang membahas penerapan algoritma klasifikasi menggunakan *Support Vector Machine* (SVM). Penelitian ini menguji berbagai kernel SVM linier, polinomial, dan RBF untuk validasi dan kompresi model. Hasilnya menunjukkan bahwa SVM dengan kernel *Radial Basis Function* (RBF) memberikan performa terbaik dengan akurasi yang lebih tinggi mencapai 95,2% [6].

Pada penelitian selanjutnya yang telah dilakukan oleh Changlu Xu et al yang berjudul membahas metode deteksi otomatis cacat permukaan potongan tiang baterai lithium menggunakan fusi multifitur dan PSO-SVM. Proses dimulai dengan pengurangan dan penyesuaian kontras gambar cacat untuk mengurangi pengaruh area yang tidak rusak dan memperjelas fitur cacat. SVM yang dioptimalkan PSO digunakan untuk mengklasifikasikan gambar secara otomatis. Hasil menunjukkan metode ini efektif dalam mendeteksi berbagai jenis cacat permukaan dengan tingkat pengenalan rata-rata mencapai 98,3% [7].

Selanjutnya, penelitian oleh Mustikasari et al. yang mendeteksi penyakit pada gambar daun teh bertujuan untuk meningkatkan akurasi deteksi penyakit dengan mengoptimasi parameter algoritma menggunakan *Particle Swarm Optimization* (PSO) menunjukkan bahwa kinerja algoritma mengalami peningkatan akurasi sekitar 2-4% setelah dioptimasi menggunakan PSO. Akurasi yang diperoleh dengan standar SVM sebesar 87%, dan setelah optimasi meningkat menjadi 91.68%. Sementara itu, standar Gradient Boosting memperoleh akurasi sebesar 89%, dan setelah dioptimasi dengan PSO, akurasinya meningkat menjadi 91% [8].

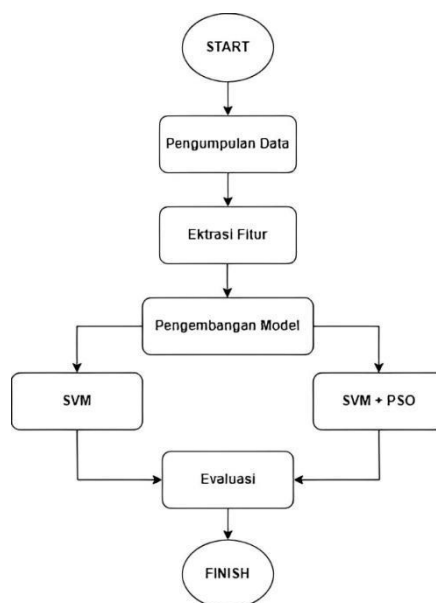
Namun, Algoritma SVM tidak selalu menghasilkan performa yang optimal, sehingga diperlukan metode optimasi untuk meningkatkan kinerjanya. Dalam penelitian ini, metode optimasi *Particle Swarm Optimization* (PSO) diterapkan untuk memperbaiki kinerja algoritma *Support Vector Machine* (SVM). Berdasarkan studi sebelumnya, PSO telah terbukti efektif dalam

meningkatkan akurasi SVM. Sehingga pada tahap identifikasi fitur tekstur dari gambar kemudian dimasukkan ke dalam pengklasifikasi SVM dioptimalkan dengan algoritma PSO untuk mencapai tujuan mengidentifikasi image kerang batik layak dan tidak layak tersebut.

Penelitian ini bertujuan menggunakan teknik terbaik untuk membantu mendeteksi kelayakan ekspor kerang batik, mengurangi potensi kualitas, serta meningkatkan produktivitas. Penggunaan teknik machine learning dan optimalisasi parameter telah terbukti efektif dalam meningkatkan akurasi performa. Melalui penelitian ini, diharapkan industri teh dapat memanfaatkan teknologi tersebut untuk meningkatkan kualitas dan hasil produksi, sekaligus menekan biaya pengobatan yang mahal sehingga dapat mencegah penurunan pendapatan perusahaan.

2. METODE PENELITIAN

Penelitian ini berfokus pada penerapan Optimasi *Support Vector Machine* dengan *Particle Swarm Optimization* untuk mendeteksi kelayakan kerang batik untuk export. Metode penelitian ini dirancang untuk memastikan pendekatan yang komprehensif dan sistematis dalam mengidentifikasi, menganalisis, dan mengevaluasi efektivitas teknik yang diusulkan.



Gambar 1. Metode Penelitian

2.1 Pengumpulan Data

Penelitian ini menggunakan data sekunder berupa gambar yang diperoleh dari data privat milik sebuah perusahaan eksportir kerang. Dataset yang digunakan bersifat privat dan diperoleh melalui metode observasi, yaitu pengumpulan informasi melalui pengamatan langsung, dengan jumlah dataset sebanyak 200 file citra.

Tabel 1 Komposisi Dataset Kerang Batik

Jenis data	Spesifikasi		Total
	Layak	Tidak Layak	
Data Training	70	70	140
Data Testing	30	30	60
Jumlah Data			200

Kerang Batik (*Paphia undulata*) memiliki pola warna cangkang yang mirip dengan karakteristik menyerupai batik, yang warna cangkang dasarnya kuning muda dan relatif gelap dengan panjang cangkang maksimum: 6,5 cm [1].



Gambar 2 Kerang Batik Kondisi Layak

Gambar 2 menunjukkan kerang batik yang memenuhi kriteria untuk diekspor. Cangkang kerang terlihat utuh dan tidak terbuka, dengan ukuran yang sesuai dan corak cangkang yang jelas terlihat. Kerang seperti ini dianggap memenuhi standar kualitas dan dapat diproses lebih lanjut untuk tujuan ekspor.



Gambar 3 Kerang Batik Kondisi Tidak Layak

Gambar 3 menunjukkan kerang batik yang tidak memenuhi kriteria ekspor. Cangkang kerang terlihat pecah atau terbuka, ukuran kerang terlalu kecil, dan corak cangkang tidak terlihat jelas. Kerang dengan kondisi seperti ini dianggap tidak layak untuk diekspor karena kualitasnya yang rendah dan tidak memenuhi standar yang ditetapkan.

2.2 Pengolahan Citra digital

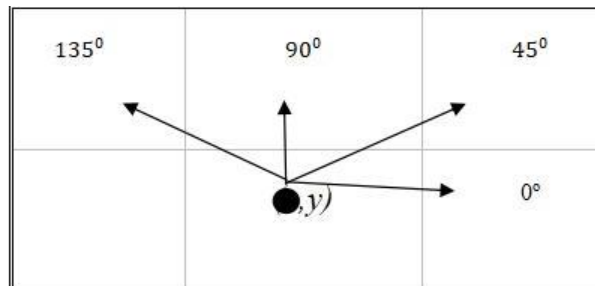
Pengolahan citra digital adalah proses manipulasi citra yang diwakili oleh array angka yang terbentuk dari persimpangan antara baris dan kolom pada setiap titik citra, membentuk piksel (pixel) sebagai unit terkecil. Setiap piksel memiliki dua informasi utama: Nomor Digital (DN) dan Nilai Kecerahan (Brightness Value/BV). Nilai DN dapat bervariasi, seperti 0 hingga 1063 untuk citra 10 bit, 0 hingga 255 untuk citra 8 bit, atau 0 hingga 63 untuk citra 6 bit, yang masing-masing mewakili rentang warna dalam piksel tersebut. Nilai kecerahan (BV) mengacu pada tingkat intensitas warna atau kecerahan yang dimiliki suatu piksel dalam citra.

Pengolahan citra digital memungkinkan ekstraksi informasi yang berguna, seperti pola, tekstur, dan fitur lainnya, yang berguna untuk aplikasi seperti klasifikasi, segmentasi, dan analisis visual. Salah satu teknik populer dalam analisis tekstur adalah Gray-Level Co-occurrence Matrix (GLCM), yang digunakan untuk menganalisis distribusi intensitas warna pada citra dan mendapatkan informasi mengenai pola tekstur [9].

2.4 Ekstraksi Fitur

Metode ekstraksi fitur menggunakan GLCM, GLCM (*Gray Level Co-occurrence Matrix*) merupakan ekstraksi fitur dengan menggunakan matriks co-occurrence yang merepresentasikan hubungan antar piksel berdasarkan jarak dan arah tertentu [10]. Matriks ini menghitung probabilitas terjadinya dua piksel dengan nilai tertentu yang berada pada jarak (d) dan sudut (θ) tertentu. Arah ditentukan oleh empat sudut pada interval 45° , yaitu 0° , 45° , 90° , dan 135° . serta menghasilkan fitur seperti Dissimilarity, Correlation, Homogeneity, Contrast, dan Energy [3].

Dalam hal ini, jeda antar piksel umumnya diatur ke 1 piksel [11]. Matriks cooccurrence adalah matriks persegi dengan total elemen sama dengan kuadrat dari jumlah tingkat intensitas piksel pada citra. Setiap titik (p, q) dalam matriks cooccurrence terarah berisi probabilitas piksel nilai p yang berdekatan dengan piksel nilai q pada jarak dan arah (180°). Properti yang digunakan pada penelitian ini [2]. Homogenitas, kontras, korelasi, dan energi. Rumus yang dipakai pada penelitian ini merupakan menjadi berikut, karena sifat-sifat tersebut memiliki sifat-sifat yang dipersyaratkan:



Gambar 4 Array Hubungan Ketetangaan pada GLCM

1) *Energy*

Energy menunjukkan tingkat keseragaman atau kesamaan suatu citra pada matriks. Ketika intensitas citra relatif seragam atau homogen, maka citra tersebut akan memiliki nilai *Energy* yang besar. *Energy* mencerminkan seberapa konsisten distribusi intensitas dalam citra [11].

2) *Homogeneity*

Homogeneity digunakan untuk menunjukkan sifat homogenitas suatu citra, yang mengacu pada seberapa mirip nilai intensitas di seluruh citra. Citra yang lebih homogen memiliki nilai *homogeneity* yang lebih tinggi. Nilai *homogeneity* dihitung dengan mengukur perubahan lokal pada tekstur gambar, dengan semakin sedikit perubahan, semakin besar nilai homogenitasnya [11].

3) *Contrast*

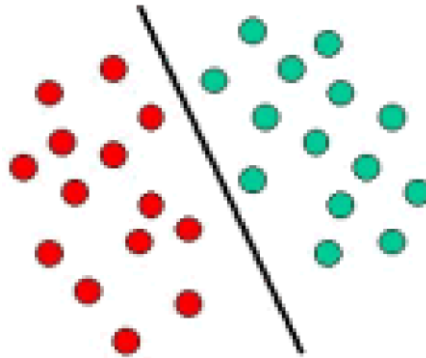
Contrast berfungsi untuk mendeteksi tingkat perbedaan atau kontras dalam suatu daerah pada citra. *Contrast* juga menunjukkan penyebaran intensitas nilai piksel dalam citra, yang dihitung dengan rumus tertentu. *Contrast* diagonal utama mengukur distribusi nilai dalam matriks, yang mencerminkan kejelasan gambar dan kedalaman bayangan, dan dikenal sebagai kontras dalam pengolahan citra. [11].

4) *Correlation*

Correlation (korelasi) digunakan untuk mengukur intensitas perbedaan antara piksel yang saling berdekatan. *Correlation* mengukur keteraturan tekstur gambar, di mana nilai korelasi yang rendah dihasilkan ketika nilai matriks sangat seragam, sementara nilai korelasi tinggi muncul ketika distribusi matriks sangat tidak merata [11].

2.5 SVM (*Support Vector Machine*)

Algoritma SVM memecahkan permasalahan klasifikasi menggunakan pemisah hyperplane terbaik antara kelas-kelas. Hyperplane mengandalkan contoh pelatihan yang menempatkannya di tepi deskriptor kelas yang disebut *support vector* [4]. *Hyperplane* dengan margin terbesar menghasilkan generalisasi yang lebih baik. SVM berusaha mencari *hyperplane* optimal yang memisahkan dua kelas data dalam ruang input, dengan margin yang diukur merupakan jarak antara hyperplane dan data paling dekat dari setiap kelas, yang dikenal sebagai *support vector*.



Gambar 5 Dua Kelas Terpisah Dengan Hyperplane

Pada gambar 5 adalah contoh lapisan pengklasifikasi atau diskriminator, sering disebut sebagai lapisan hyperplane 3, dengan dua kelas yang dipisahkan oleh linear hyperplane sebagai bentuk paling sederhana dari SVM. Referensinya adalah set *training D*, yang dapat dilambangkan dengan persamaan.

2.5.1 Kernel SVM

Apabila terdapat masalah dengan data yang tidak dipisahkan secara linier dalam rentang input, SVM soft-margin tidak dapat menemukan hyperplane yang efektif sebagai pemisah. yang dapat meminimalkan kesalahan klasifikasi data dan melakukan generalisasi dengan baik.. Oleh karena itu, kernel dapat digunakan untuk mengubah ruang fitur data menjadi ruang dimensi yang lebih tinggi yang dikenal sebagai ruang kernel di mana data dipisahkan secara linier [12].

1) Dot product Kernel

Dot kernel adalah salah satu jenis kernel yang dapat digunakan di SVM. Kernel ini merupakan hasil kali dari perkalian dot product input terhadap bobot. Dot kernel di definisikan dengan:

$$K(x_i, x_j) = x_i \cdot x_j \quad (8)$$

2) Radial Basis Function (RBF) Kernel

Kernel RBF adalah fungsi kernel yang biasanya digunakan untuk menganalisis situasi ketika data tidak dapat dipisahkan dengan cara linear. Kernel RBF memiliki dua parameter utama, yaitu Cost dan Gamma. Parameter gamma digunakan untuk mengetahui seberapa luas pengaruh *pattern training set* terhadap *training set*, nilai yang rendah menunjukkan bahwa luas dan nilai tinggi berarti tertutup. Ketika gamma kecil, masuk akal untuk memasukkan titik lebih jauh di sepanjang garis pemisah dalam perhitungan garis pemisah. Jika gamma tinggi, maka titik-titik di sekitar garis pemisah akan ikut dipertimbangkan dalam perhitungan. Berikut adalah persamaan untuk kernel RBF [6].

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (9)$$

3) Polynomial Kernel (derajat d)

Kernel Polinomial adalah jenis fungsi kernel yang diterapkan ketika data tidak dapat dipisahkan secara linear [13]. Polinomial kernel sangat cocok dengan masalah ketika semua data pelatihan dinormalisasi. Yang membedakan inti ini dengan inti linier adalah kita harus menentukan parameter lain yaitu d atau derajat, misalnya 2 atau 3 untuk membentuk persamaan kurva.

$$K(x_i, x_j) = (\alpha \cdot (x_i, x_j) + c)^d \quad (10)$$

2.6 PSO (*Particle Swam Optimization*)

Particle Swarm Optimization (PSO) adalah algoritma pencarian global yang meniru perilaku pencarian makan burung dan ikan. PSO bekerja dengan menginisialisasi partikel secara acak dan memperbarui posisi mereka di setiap generasi untuk menemukan solusi optimal, bergerak menuju individu terbaik dan posisi global terbaik dalam pencarian di ruang multi-dimensi [7].

2.7 *Confusion Matrix*

Mengukur efisiensi suatu sistem klasifikasi sangat penting karena dapat menggambarkan seberapa efektif dan akurat sistem tersebut dalam mengklasifikasikan data [14]. Dalam Confusion Matrix dikenal empat istilah diketahui yang mewakili hasil dari proses klasifikasi

Tabel 2 Confussion Matrix

	Aktual: Negatif	Aktual: Positif
Prediksi: Negatif	TN	FN
Prediksi: Positif	FP	TP

Pada Tabel 2, TN (*True Negative*) menunjukkan jumlah prediksi negatif yang benar, sementara TP (*True Positive*) menunjukkan jumlah prediksi positif yang benar. Dua jenis kesalahan yang ada adalah FN (*False Negative*), yang menunjukkan jumlah prediksi negatif yang tidak sesuai dengan aktual, dan FP (*False Positive*), yang menunjukkan jumlah prediksi positif yang salah.

- 1) Akurasi menggambarkan tingkat keakuratan model dalam melakukan klasifikasi dengan benar. Rumus dari akurasi dapat dilihat dari sebagai berikut:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

- 2) Recall menjelaskan sebuah model dalam menemukan kembali sebuah informasi. Recall dihitung menggunakan rumus berikut:

$$\text{Recall} = \frac{TP}{TP+FN} \tag{12}$$

- 3) Presisi atau nilai untuk menggambarkan keakuratan antar data dengan hasil prediksi. Rumus untuk menghitung presisi adalah sebagai berikut:

$$\text{Presisi} = \frac{TP}{TP+FP} \tag{13}$$

3. HASIL DAN PEMBAHASAN

3.1. Ekstrasi Fitur GLCM

Citra yang sudah melalui tahapan *preprocessing* kemudian akan di ekstrasi menggunakan GLCM dengan fitur (*dissimilarity, correlation, homogeneity, contrast* dan *energy* akan menghasilkan angka dari berbagai sudut 0°, 45°, 90°, 130°. Sebagai berikut

1) Hasil Ekstrasi Fitur Sudut 0°

Tabel 3 Ekstrasi fitur pada sudut 0°

No	Nama Image	Contrast	Homogeneity	Correlation	Energy	Kelas
1.	Image1	0.907571039	0.209515957	1498.289492	0.12571463	Layak
2.	Image2	0.921328062	0.165453205	785.1869195	0.061697603	Layak
3.	Image3	0.901163257	0.230852134	1552.511251	0.125596751	Layak
4.	Image4	0.800644374	0.147490884	2434.82093	0.061283158	Layak
5.	Image5	0.880470253	0.164350557	1383.264306	0.077445394	Layak
...
400	Image400	0.920096022	0.193234448	1143.880865	0.065227393	Tidak Layak

Tabel 3 menunjukkan hasil ekstraksi fitur pada sudut 0°, dengan lima fitur GLCM yang dihitung: Dissimilarity, Correlation, Homogeneity, Contrast, dan Energy. Rentang nilai pada tabel tersebut adalah sebagai berikut: Contrast berkisar antara 261.1361 hingga 3317.0440, Homogeneity antara 0.1206 hingga 0.3858, Correlation antara 0.7767 hingga 0.9627, dan Energy antara 0.0249 hingga 0.1829.

2) Hasil Ekstrasi Fitur Sudut 45°

Tabel 4 Ekstrasi fitur pada sudut 45°

No	Nama Image	Contrast	Homogeneity	Correlation	Energy	Kelas
1.	Image1	0.843725407	0.179218312	2453.635042	0.108390042	Layak
2.	Image2	0.891880507	0.153397306	1050.311111	0.057573854	Layak
3.	Image3	0.867627373	0.214834062	2039.185676	0.115009879	Layak
4.	Image4	0.766944366	0.133925136	2779.21653	0.05808582	Layak
5.	Image5	0.844849314	0.1498435	1747.817462	0.072659817	Layak
...
400	Image400	0.891601188	0.173509046	1510.815006	0.060479343	Tidak Layak

Tabel 4 menampilkan hasil ekstraksi fitur pada sudut 45°, dengan lima fitur GLCM yang dianalisis: Dissimilarity, Correlation, Homogeneity, Contrast, dan Energy. Rentang nilai pada tabel tersebut adalah sebagai berikut: Contrast berkisar antara 371.5897 hingga 3866.8932, Homogeneity antara 0.1037 hingga 0.3590, Correlation antara 0.7199 hingga 0.9384, dan Energy antara 0.0232 hingga 0.1691.

3) Hasil Ekstrasi Fitur Sudut 90°

Tabel 5 Ekstrasi fitur pada sudut 90°

No	Nama Image	Contrast	Homogeneity	Correlation	Energy	Kelas
1.	Image1	0.826127618	0.178065335	2756.123899	0.10635219	Layak
2.	Image2	0.886733557	0.151669326	1118.295067	0.059125778	Layak
3.	Image3	0.861678131	0.215754435	2146.201734	0.114106445	Layak
4.	Image4	0.779322917	0.136798691	2676.564034	0.059778371	Layak
5.	Image5	0.83512385	0.143016271	1884.632266	0.072570648	Layak
...
400	Image400	0.898740209	0.191217626	1438.840409	0.068348981	Tidak Layak

Tabel 5 menunjukkan hasil ekstraksi fitur pada sudut 90°, dengan lima fitur GLCM yang dianalisis: Dissimilarity, Correlation, Homogeneity, Contrast, dan Energy. Rentang nilai pada tabel tersebut adalah sebagai berikut: Contrast berkisar antara 300.5725 hingga 3570.7769, Homogeneity antara 0.1068 hingga 0.3640, Correlation antara 0.7230 hingga 0.9432, dan Energy antara 0.0246 hingga 0.1705.

4) Hasil Ekstraksi Fitur Sudut 135°

Tabel 6 Ekstraksi fitur pada sudut 135°

No	Nama Image	Contrast	Homogeneity	Correlation	Energy	Kelas
1.	Image1	0.84114059	0.181395101	2494.206012	0.109665814	Layak
2.	Image2	0.883976238	0.147620583	1127.095988	0.055543237	Layak
3.	Image3	0.862066986	0.209038646	2124.842819	0.113259539	Layak
4.	Image4	0.764661395	0.132805423	2806.461563	0.056618071	Layak
5.	Image5	0.827358945	0.139054195	1944.851543	0.068489472	Layak
...
400	Image400	0.898771392	0.177604845	1410.879864	0.064299526	Tidak Layak

Tabel 6 menunjukkan hasil ekstraksi fitur pada sudut 135°, dengan lima fitur GLCM yang dianalisis: Dissimilarity, Correlation, Homogeneity, Contrast, dan Energy. Rentang nilai pada tabel tersebut adalah sebagai berikut: Contrast berkisar antara 278.5247 hingga 3686.5242, Homogeneity antara 0.1061 hingga 0.3647, Correlation antara 0.7147 hingga 0.9452, dan Energy antara 0.0229 hingga 0.1720.

3.2. Evaluasi

Untuk mendapatkan prediksi, kumpulan data dibagi menjadi dua kategori data, yaitu data pelatihan dan data pengujian. Data pelatihan digunakan untuk membangun model, sedangkan data pengujian digunakan untuk melakukan prediksi Selanjutnya, kinerja model diukur dengan membandingkan label output yang diprediksi terhadap output sesungguhnya. Hasil evaluasi ini berupa tag yang menunjukkan apakah prediksi valid atau tidak.

Nilai akurasi, presisi, dan recall dihitung berdasarkan tabel confusion matrix ini. Berikut adalah contoh perhitungan menggunakan confusion matrix yang terbentuk:

Tabel 7 Sample Confussion Matrix Kernel RBF

	Aktual: layak	Aktual: tidaklayak	class precision
Prediksi: layak	TP: 68	FP: 3	95.77%
Prediksi: tidaklayak	FN: 2	TN: 67	97.10%
class recall	97.14%	95.71%	

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP+TN}{TP+TN+FP+FN} \\
 &= \frac{68+67}{68+67+3+2} \\
 &= \frac{135}{140} \\
 &= 0.9643 * 100\% \\
 &= 96.43\%
 \end{aligned}$$

1) Pengujian Kernel Dot

Pada tabel dibawah ini menunjukkan bahwa pengujian kernel Dot Product dengan parameter pada SVM berupa Cost (C). Range nilai C yang digunakan ditunjukkan pada tabel dibawah ini.

Tabel 8 Pengujian Kernel Dot Product

C	Performance Sudut 0	Performance Sudut 45	Performance Sudut 90	Performance Sudut 135
0.0	83.57%	85.71%	83.57%	83.57%
100	86.43%	85.71%	85.00%	85.71%
10	85.71%	85.00%	85.00%	85.71%
1.0	87.14%	85.00%	84.29%	84.29%
0.1	83.57%	84.29%	84.29%	83.57%
0.01	65.71%	67.14%	72.86%	70.71%

Tabel 8 menunjukkan bahwa nilai pengujian kernel Dot pada keempat sudut menghasilkan nilai terbaik sebesar 87.14%, dengan nilai C sebesar 1.

2) Pengujian Kernel RBF

Pada tabel dibawah ini menunjukkan bahwa pengujian kernel RBF dengan penambahan parameter pada SVM yaitu gamma. Berikut nilai range pada parameter C dan parameter Gamma.

Tabel 9 Pengujian Kernel RBF

Gamma	C	Performance Sudut 0	Performance Sudut 45	Performance Sudut 90	Performance Sudut 135
1.0	0.0	94.29%	96.43%	96.43%	95.71%
0.01	100.0	89.29%	85.71%	91.43%	82.14%
0.1	10.0	93.57%	92.14%	91.43%	92.14%
1.0	1.0	94.29%	96.43%	96.43%	95.71%
1.0	0.1	85.00%	83.57%	82.14%	83.57%
0.0001	0.1	72.14%	66.43%	69.29%	70.71%

Tabel 9 menunjukkan bahwa nilai pengujian kernel RBF pada keempat sudut menghasilkan nilai terbaik sebesar 96.43%, dengan nilai Gamma sebesar 1.0 dan nilai C sebesar 0.0.

3) Pengujian Kernel Polynomial

Pada tabel dibawah ini menunjukkan bahwa pengujian kernel polynomial terdapat penambahan parameter pada SVM berupa degree.

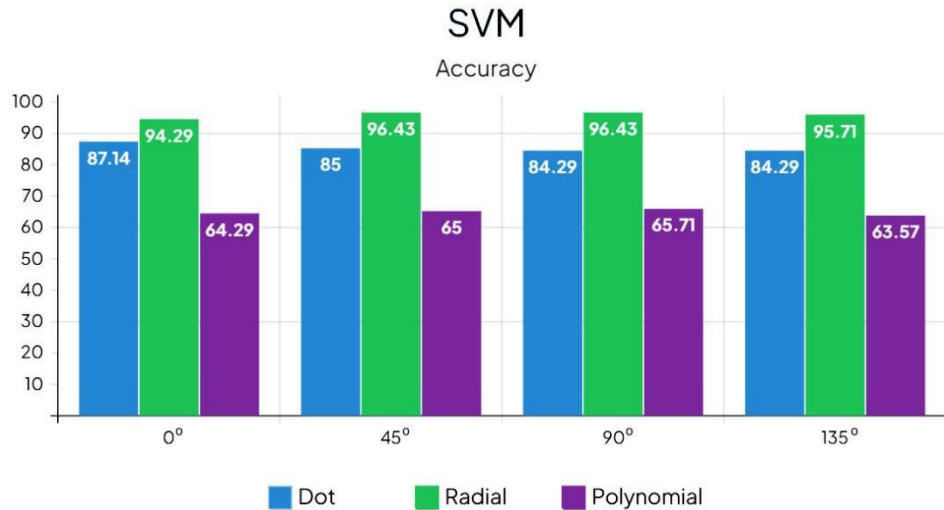
Tabel 10 Pengujian Kernel Polynomials

Kernel Degree	Performance Sudut 0	Performance Sudut 45	Performance Sudut 90	Performance Sudut 135
2.0	64.29%	65.00%	65.71%	63.57%
3.0	57.86%	57.14%	57.86%	57.86%
4.0	55.71%	55.71%	55.71%	55.71%
5.0	55.00%	55.00%	55.00%	55.00%
6.0	53.57%	55.00%	55.00%	54.29%

Tabel 10 menunjukkan bahwa nilai pengujian kernel Polynomial pada keempat sudut menghasilkan nilai terbaik sebesar 65.71%, dengan nilai degree sebesar 2.0.

3.3. Klasifikasi SVM

Pada eksperimen pertama, klasifikasi akan dilakukan menggunakan metode SVM dengan kernel Dot, RBF, dan Polynomial. Berikut adalah nilai akurasi yang dihasilkan oleh SVM. Pengujian akan dilaksanakan menggunakan RapidMiner Studio.

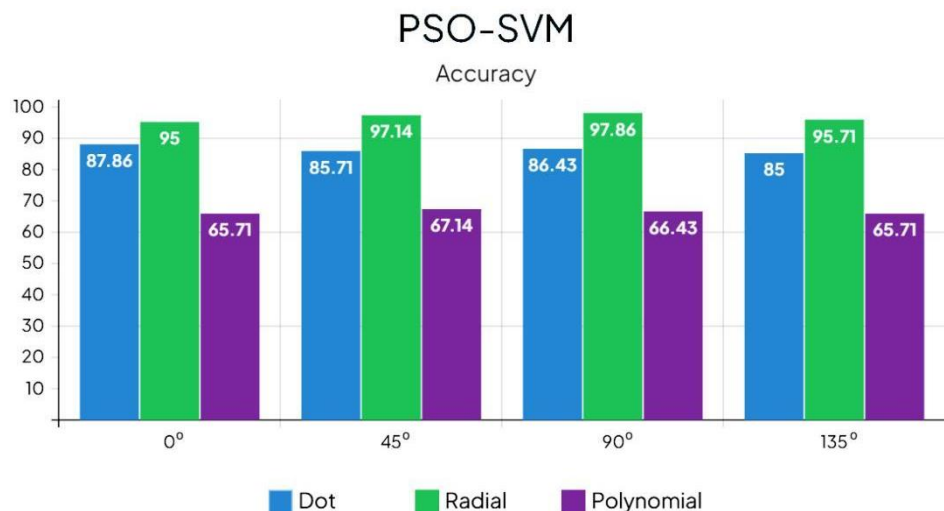


Gambar 6 Grafik Accuracy Kernel pada SVM

Pada Gambar 6, dari ketiga kernel SVM, terdapat dua sudut dengan nilai akurasi tertinggi pada kernel RBF, yaitu pada sudut 45° dan 90° dengan akurasi masing-masing mencapai 96.43%. Keempat sudut ini merupakan hasil klasifikasi menggunakan SVM tanpa adanya optimasi.

3.4. Optimasi PSO pada SVM

Pada eksperimen kedua, dengan menggunakan dataset dan algoritma yang sama, kali ini akan ditambahkan optimasi menggunakan *Particle Swarm Optimization* (PSO) untuk mengevaluasi apakah PSO dapat meningkatkan atau bahkan menurunkan nilai akurasi pada suatu metode. Pengujian akan dilakukan menggunakan RapidMiner Studio dengan parameter population size 5, maximum number of generations 30, dan inertia weight 1.0.



Gambar 7 Grafik Accuracy Kernel pada SVM dengan Optimasi PSO

Pada Gambar 7, dari ketiga kernel SVM yang telah dioptimasi, terdapat nilai akurasi tertinggi pada kernel RBF di sudut 90° dengan akurasi sebesar 97.86%. Hasil ini menunjukkan bahwa optimasi menggunakan PSO dapat meningkatkan akurasi, karena nilai akurasi yang dihasilkan lebih tinggi setelah dioptimasi.

4. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa ekstraksi fitur dapat meningkatkan kinerja model Machine Learning dengan menghilangkan noise atau atribut yang tidak penting, memungkinkan model fokus pada fitur utama. Normalisasi data juga dapat meningkatkan kinerja model Machine Learning secara signifikan, meskipun tidak berlaku secara umum untuk semua model.

Penelitian menunjukkan bahwa normalisasi data dapat meningkatkan akurasi metode SVM dengan kernel RBF, yang menghasilkan nilai tertinggi dibandingkan kernel Dot Product dan Polynomial. Partisi data untuk pelatihan dan pengujian mempengaruhi kinerja model, di mana semakin banyak data yang digunakan untuk pelatihan, semakin akurat model yang dihasilkan. Optimasi parameter menggunakan *Particle Swarm Optimization* (PSO) dapat meningkatkan akurasi algoritma Machine Learning sekitar 1-4%. Hasil terbaik diperoleh dengan SVM yang dioptimasi menggunakan PSO, menghasilkan akurasi 97.86% pada kernel RBF dengan sudut 90° .

Penerapan PSO pada SVM kernel RBF menghasilkan peningkatan akurasi yang signifikan. Model terbaik ini dapat diterapkan untuk deteksi kelayakan ekspor kerang batik, membantu meminimalkan beban kerja dan mempercepat proses penyortiran. Penelitian selanjutnya dapat fokus pada optimasi parameter menggunakan PSO pada metode Machine Learning lainnya.

DAFTAR PUSTAKA

- [1] Rangkuti, F. M. (2021). Penentuan Kadar Lemak pada Kerang Darah dan Kerang Batik dengan Metode Hidrolisis (Weibull).
- [2] Kusumanto, R., & Tompunu, N. (2011). Digital Image Processing To Detect Objects Using RGB Normalization Model Color Processing. In National Seminar on Applied Information and Communication Technology.
- [3] Raczko, E., & Zagajewski, B. (2017). Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *European Journal of Remote Sensing*, 50(1), 144-154.
- [4] Kusuma, S. F., Pawening, R. E., & Dijaya, R. (2017). Otomatisasi klasifikasi kematangan buah mengkudu berdasarkan warna dan tekstur. *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, 3(1), 17-23.
- [5] D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: an overview," *Soft Comput.*, vol. 22, pp. 387-408, 2018.
- [6] Behera, M. K., Chakravarty, S., Gourav, A., & Dash, S. (2020, November). Detection of Nuclear Cataract in Retinal Fundus Image using RadialBasis FunctionbasedSVM. In 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC) (pp. 278-281). IEEE.
- [7] Xu, C., Li, L., Li, J., & Wen, C. (2021). Surface defects detection and identification of lithium battery pole piece based on multi-feature fusion and PSO-SVM. *Ieee Access*, 9, 85232-85239.
- [8] Ramli, A. R., & Gibran, A. K. (2023). Analisis Kinerja Algoritma Machine Learning Untuk Deteksi Penyakit Daun Teh Dengan Particle Swarm Optimization. *Journal of Embedded Systems, Security and Intelligent Systems*, 4(2), 97-106.
- [9] Abbas Z., Rehman M.U., Najam S., Rizvi S.M.D., 2019 "An Efficient Gray-Level Co-Occurrence Matrix (GLCM) based Approach Towards Classification of Skin Lesion," *IEEE*, vol. 8, no. 2, pp. 317 - 321, 2019.

- [10] Zeeshan Abbas, Mobeen-ur-Rehman, Shahzaib Najam, S.M. Danish Rizvi, "An Efficient Gray-Level Co-Occurrence Matrix (GLCM) based Approach Towards Classification of Skin Lesion," IEEE, vol. 8, no. 2, pp. 317 - 321, 2019.
- [11] Aditya Budi Pradana, Prawito Prajitno, "A Portable Surface Roughness Measurement System Using Laser Speckle Imaging Based on GLCM," 2019 6th International Conference on Instrumentation, Control, and Automation (ICA), vol. 6, pp. 100 - 105, 31 July – 2 August 2019.
- [12] Ngadilan, M. A. A., Ismail, N., Rahiman, M. H. F., Taib, M. N., Ali, N. A. M., & Tajuddin, S. N. (2018, August). Radial Basis Function (RBF) tuned Kernel Parameter of Agarwood Oil Compound for Quality Classification using Support Vector Machine (SVM). In 2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC) (pp. 64-68). IEEE.
- [13] Prajapati, G. L., & Patle, A. (2010, November). On performing classification using SVM with radial basis and polynomial kernel functions. In 2010 3rd International Conference on Emerging Trends in Engineering and Technology (pp. 512-515). IEEE.
- [14] J. J. Remus and L. M. Collins, "Identifying impaired cochlear implant channels via speech-token confusion matrix analysis," ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc., vol. 4, pp. 741–744, 2007, doi: 10.1109/ICASSP.2007.367019.