

Perbandingan Algoritma K-Nearest Neighbor dan Support Vector Machine Pada Pengenalan Pola Tanda Tangan Digital

Yuli Yadin*, Dyah Ayu Megawaty

Fakultas Teknik dan Ilmu Komputer, Informatika, Universitas Teknokrat Indonesia, Bandar Lampung, Indonesia

Email: ^{1,*}yuli_yadin@teknokrat.ac.id, ²dyahayumegawaty@teknokrat.ac.id

Email Penulis Korespondensi: yuli_yadin@teknokrat.ac.id

Submitted: 13/02/2025; Accepted: 28/03/2025; Published: 28/03/2025

Abstrak—Di era digital yang serba cepat, keamanan identitas menjadi krusial, dan tanda tangan digital berperan penting dalam verifikasi serta autentikasi. Penelitian ini berfokus pada analisis dan perbandingan kinerja algoritma K-Nearest Neighbors (KNN) dan Support Vector Machine (SVM) dalam pengenalan pola tanda tangan digital. Kedua algoritma ini banyak digunakan dalam tugas klasifikasi, dan penelitian ini bertujuan mengidentifikasi algoritma mana yang paling efektif dalam mengenali serta mengklasifikasikan tanda tangan digital dengan tingkat akurasi tertinggi. Data tanda tangan digital dikumpulkan dari berbagai sumber, termasuk dataset publik dan data yang dikumpulkan secara langsung. Fitur-fitur utama diekstraksi menggunakan metode Gray-Level Co-occurrence Matrix (GLCM), yang efektif dalam menggambarkan tekstur serta pola tanda tangan. Fitur-fitur ini digunakan untuk melatih model klasifikasi KNN dan SVM. Kinerja kedua algoritma dievaluasi berdasarkan metrik akurasi, presisi, dan recall. Hasil penelitian menunjukkan bahwa KNN dengan nilai $k = 3$ mencapai akurasi sebesar 91,42%, sementara SVM dengan kernel linear unggul dengan akurasi 97,06%. Selain itu, SVM juga lebih stabil dalam menangani tanda tangan yang kompleks serta memiliki presisi dan recall lebih tinggi dibandingkan KNN, masing-masing sebesar 97,52% dan 97,06%. Di sisi lain, KNN lebih cepat dalam proses pelatihan dan memiliki implementasi yang lebih sederhana. Penelitian ini memberikan wawasan berharga tentang pemilihan algoritma klasifikasi yang optimal untuk aplikasi pengenalan tanda tangan digital. Hasil penelitian ini dapat menjadi panduan bagi pengembang sistem keamanan dan autentikasi dalam memilih metode yang paling efektif guna melindungi identitas serta mencegah pemalsuan tanda tangan.

Kata Kunci: Tanda Tangan Digital; KNN; SVM; GLCM; Klasifikasi

Abstract—In the fast-paced digital era, identity security has become crucial, and digital signatures play an important role in verification and authentication. This study focuses on the analysis and comparison of the performance of the K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) algorithms in digital signature pattern recognition. Both algorithms are widely used in classification tasks, and this study aims to identify which algorithm is most effective in recognizing and classifying digital signatures with the highest accuracy. Digital signature data was collected from various sources, including public datasets and directly collected data. Key features were extracted using the Gray-Level Co-occurrence Matrix (GLCM) method, which is effective in describing the texture and pattern of the signature. These features were used to train the KNN and SVM classification models. The performance of both algorithms was evaluated based on accuracy, precision, and recall metrics. The results showed that KNN with a value of $k = 3$ achieved an accuracy of 91.42%, while SVM with a linear kernel excelled with an accuracy of 97.06%. In addition, SVM is also more stable in handling complex signatures and has higher precision and recall than KNN, at 97.52% and 97.06%, respectively. On the other hand, KNN is faster in the training process and has a simpler implementation. This study provides valuable insights into the selection of optimal classification algorithms for digital signature recognition applications. The results of this study can be a guide for security and authentication system developers in choosing the most effective method to protect identity and prevent signature forgery.

Keywords: Digital Signature; KNN; SVM; GLCM; Classification

1. PENDAHULUAN

Kemajuan teknologi yang terus berkembang pesat menjadikan aspek keamanan sebagai salah satu kebutuhan dasar dalam kehidupan manusia. Keamanan ini meliputi perlindungan terhadap individu serta hal-hal yang berhubungan dengan identitas pribadi, termasuk dokumen-dokumen penting dan legal. Salah satu metode keamanan yang banyak diterapkan adalah sistem pengenalan identitas, khususnya melalui teknologi biometrika. Biometrika itu sendiri terbagi dalam dua kategori utama yaitu biometrika fisik dan biometrika perilaku. Tanda tangan, misalnya, adalah bentuk biometrika perilaku yang paling umum digunakan[1]. Sistem biometrika memiliki dua fungsi utama, yakni verifikasi dan identifikasi. Verifikasi tanda tangan berfungsi untuk memastikan bahwa tanda tangan yang diajukan benar milik orang yang terdaftar serta untuk memverifikasi keaslian tanda tangan tersebut[2]. Sementara itu, identifikasi bertujuan untuk memastikan identitas seseorang berdasarkan tanda tangan yang diberikan [3]. Tanda tangan digital, sebagai salah satu jenis tanda tangan, digunakan sebagai metode autentikasi untuk memvalidasi pesan atau dokumen[4]. Tanda tangan adalah simbol yang mewakili identitas individu dan, selain untuk verifikasi identitas, juga berfungsi sebagai bukti sah persetujuan terhadap isi dokumen. Biasanya, tanda tangan dilakukan dengan pena biasa, yang dikenal sebagai tanda tangan basah. Dalam konteks yang lebih luas, tanda tangan berfungsi sebagai sistem verifikasi identitas, dengan setiap individu memiliki ciri khas dalam cara menandatangani. Oleh karena itu, tanda tangan pada suatu dokumen menandakan bahwa pemiliknya telah memahami dan menyetujui isi dokumen tersebut. Namun, meskipun tanda tangan sering digunakan sebagai metode verifikasi identitas, terdapat variasi yang mungkin terjadi pada tanda tangan yang dibuat oleh orang yang sama di kesempatan yang berbeda[5]. Variasi ini dapat mempersulit sistem verifikasi. Tanda tangan memiliki peran penting dalam proses verifikasi dan legalisasi dokumen, seperti Memorandum of Understanding (MoU), surat pernyataan, surat pengantar resmi, dan berbagai dokumen hukum lainnya. Dalam

konteks hukum, tanda tangan digunakan sebagai alat bukti dalam pembuktian perdata. Akta hanya dapat dianggap sah jika sengaja dibuat sebagai bukti suatu peristiwa yang relevan dan ditandatangani oleh pihak-pihak yang bersangkutan[6]. Beberapa tanda tangan dapat terbaca dengan jelas, namun ada juga yang sulit dibaca. Selain itu, tanda tangan bisa dianggap sebagai karya seni yang unik dan artistik[7]. Dalam hal ini, pengenalan pola tanda tangan dapat dilakukan dengan menggunakan algoritma klasifikasi untuk meningkatkan akurasi pengenalan pola tersebut.

K-Nearest Neighbors (KNN) adalah salah satu algoritma dalam machine learning yang digunakan untuk klasifikasi dan regresi. KNN bekerja dengan prinsip bahwa data yang serupa akan cenderung berdekatan dalam ruang fitur. Salah satu keunggulan KNN adalah kemampuannya menangani klasifikasi non-linear dan mengidentifikasi pola yang kompleks. Pemilihan jumlah tetangga terdekat (K) sangat mempengaruhi akurasi model. Jika K terlalu besar, model cenderung terlalu umum dan kesulitan menangkap variasi data, sementara jika K terlalu kecil, model lebih sensitif terhadap noise dan berisiko mengalami overfitting[8]. Algoritma Support Vector Machine (SVM) merupakan metode pembelajaran yang menggunakan ruang hipotesis linear dalam ruang fitur berdimensi tinggi. SVM memiliki hubungan erat dengan penambahan data dan pembelajaran mesin, karena mampu mengklasifikasikan data baru dengan akurasi tinggi. Proses pembelajaran SVM bertujuan menemukan bidang pemisah optimal yang memaksimalkan margin antara berbagai kelas, meningkatkan kemampuan generalisasi model. Meskipun SVM awalnya dirancang untuk klasifikasi linear, algoritma ini dapat disesuaikan untuk masalah non-linear melalui teknik kernel trick, yang memetakan data ke ruang berdimensi lebih tinggi untuk menemukan pemisahan yang lebih efektif[9].

SVM dikenal memiliki tingkat akurasi yang sangat baik dalam pengenalan pola, dan proses pelatihannya relatif cepat. Namun, salah satu kelemahan utama SVM adalah kesulitannya dalam mengelola data berukuran besar[10]. Penelitian ini bertujuan untuk membandingkan dua algoritma, KNN dan SVM, untuk memperoleh wawasan yang lebih mendalam tentang kelebihan, kekurangan, dan kinerja masing-masing algoritma. Diharapkan hasil perbandingan ini dapat memberikan gambaran yang jelas mengenai algoritma mana yang lebih efisien dan efektif dalam pengenalan pola tanda tangan. Sebelumnya, penelitian serupa tentang pengenalan pola tanda tangan telah dilakukan pada tahun 2020.

Penelitian sebelumnya dilakukan pada tahun 2020 [11], Penelitian tersebut menguji efektivitas KNN dalam proses klasifikasi tanda tangan berdasarkan pola yang telah diekstraksi. Hasil pengujian menunjukkan bahwa algoritma KNN memiliki keunggulan dalam mendeteksi pola tanda tangan dengan tingkat akurasi yang jauh lebih tinggi dibandingkan metode SFET dan MAP, yaitu sebesar 98.25%.

Keunggulan utama KNN terletak pada pendekatan berbasis jarak yang mampu mengelompokkan tanda tangan dengan tingkat kemiripan tinggi. Namun, meskipun KNN menghasilkan akurasi yang lebih baik, metode ini memiliki keterbatasan dalam hal efisiensi komputasi, terutama ketika jumlah data yang digunakan semakin besar.

Selain itu, metode Support Vector Machine (SVM) juga telah diterapkan dalam penelitian lain guna meningkatkan kinerja klasifikasi tanda tangan. Dalam penelitian tersebut, citra tanda tangan dibagi ke dalam beberapa zona untuk mengekstraksi karakteristik unik dari setiap bagian tanda tangan sebelum diklasifikasikan menggunakan SVM. Hasil penelitian ini menunjukkan bahwa metode SVM mampu mencapai tingkat akurasi sebesar 99,43%[12], yang lebih tinggi dibandingkan SFET dan MAP, tetapi masih sedikit lebih rendah dibandingkan KNN. Keunggulan utama dari metode ini adalah kemampuannya dalam menangani variasi pola tanda tangan dengan lebih baik dibandingkan metode berbasis ciri bentuk sederhana. Namun, tantangan utama dalam penerapan SVM adalah pemilihan parameter kernel yang optimal serta waktu komputasi yang relatif lebih tinggi dibandingkan KNN dalam beberapa kondisi.

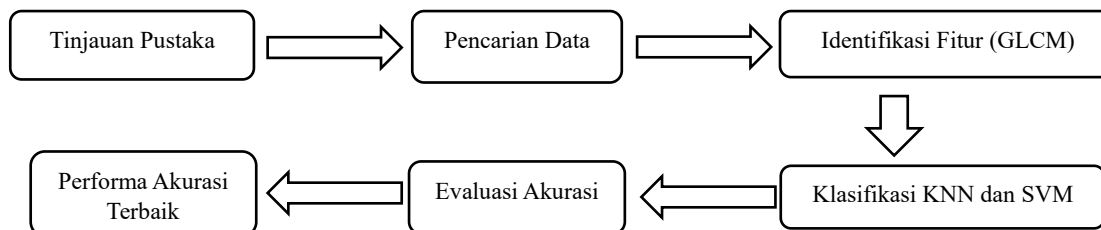
Berdasarkan perbedaan hasil dari penelitian sebelumnya, penelitian ini berupaya mengisi kesenjangan dengan melakukan perbandingan mendalam terhadap dua algoritma utama, yaitu KNN dan SVM, dalam konteks pengenalan pola tanda tangan digital. Salah satu aspek utama yang menjadi perhatian adalah sejauh mana kedua algoritma tersebut mampu mengatasi tantangan yang disebabkan oleh variasi tanda tangan, seperti perbedaan gaya tulisan, tekanan pena, ketebalan garis, atau perubahan akibat faktor manusia. Kontribusi utama dari penelitian ini adalah memberikan analisis sistematis mengenai kelebihan dan kekurangan masing-masing algoritma dalam pengenalan pola tanda tangan digital. Penelitian ini tidak hanya berfokus pada tingkat akurasi, tetapi juga mempertimbangkan efisiensi komputasi, ketahanan terhadap variasi tanda tangan, serta kemampuannya dalam mengklasifikasikan tanda tangan baru yang tidak terdapat dalam data latih. Dengan adanya perbandingan ini, penelitian ini diharapkan dapat memberikan rekomendasi strategis bagi pengembang sistem verifikasi tanda tangan digital berbasis kecerdasan buatan, sehingga dapat menjadi dasar dalam pemilihan metode klasifikasi yang lebih optimal sesuai dengan kebutuhan implementasi di dunia nyata.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Secara garis besar, tahapan penelitian merupakan serangkaian langkah sistematis yang dilakukan untuk memperoleh data dan informasi yang valid serta dapat dipertanggungjawabkan. Tahapan ini bertujuan agar penelitian berlangsung secara terstruktur dan sesuai dengan metode ilmiah [13]. Beberapa langkah utama dalam penelitian meliputi perencanaan, pengumpulan data, analisis data, interpretasi hasil, dan penyusunan laporan [14]. Dalam

pelaksanaannya, terdapat beberapa tahapan yang harus dilakukan agar penelitian berjalan secara efektif dan terorganisir. Gambar 1 di bawah ini menunjukkan tahapan-tahapan yang diterapkan dalam penelitian ini :



Gambar 1. Diagram Tahapan Penelitian

Pada Gambar 1 diagram tahapan penelitian ini menggambarkan proses penelitian yang diawali dengan Tinjauan Pustaka yang komprehensif, dilakukan dengan menelusuri sumber dari berbagai jurnal dan artikel, baik nasional maupun internasional. Data yang diperoleh berupa tanda tangan, yang selanjutnya diproses melalui Identifikasi fitur menggunakan metode Gray Level Co-occurrence Matrix (GLCM). Setelah fitur berhasil diekstraksi, tahap berikutnya adalah pemodelan pengenalan pola atau klasifikasi yang menggunakan algoritma K-Nearest Neighbors (KNN) dan Support Vector Machine (SVM). Pada tahap akhir, dilakukan evaluasi model guna menentukan akurasi terbaik, serta membandingkan kinerja kedua algoritma dalam mencapai tingkat akurasi yang optimal.

2.2 Tinjauan Pustaka

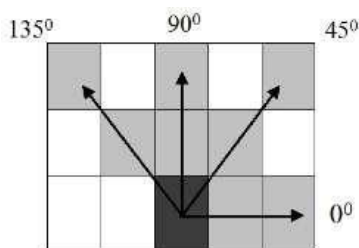
Tinjauan pustaka dalam penelitian ini bertujuan untuk mengumpulkan berbagai referensi terkait pengenalan pola tanda tangan. Pengenalan pola tanda tangan merupakan salah satu aplikasi biometrik yang berperan penting dalam verifikasi identitas dan autentikasi tanda tangan. Teknologi ini terus berkembang seiring dengan penerapan algoritma pembelajaran mesin guna meningkatkan akurasi dan efisiensi dalam proses pengenalan tanda tangan. Dua algoritma yang umum digunakan dalam bidang ini adalah K-Nearest Neighbors (KNN) dan Support Vector Machine (SVM).

2.3 Pencarian Data

Pencarian data dilakukan dengan memanfaatkan dataset publik dari Kaggle serta data yang diperoleh secara langsung. Dataset ini terdiri dari 1.050 gambar tanda tangan, yang akan digunakan sebagai materi utama dalam tahap analisis dan pengujian algoritma.

2.4 Identifikasi Fitur (GLCM)

Pada tahun 1973, Robert M. Haralick memperkenalkan konsep Gray Level Co-occurrence Matrix (GLCM) sebagai metode untuk menganalisis pola dalam citra. Berbeda dengan analisis tekstur tingkat pertama yang hanya mempertimbangkan distribusi intensitas piksel, GLCM melakukan analisis tekstur tingkat kedua dengan mengevaluasi hubungan antar piksel dalam suatu citra[15]. Matriks GLCM berfungsi untuk merekam informasi mengenai posisi pasangan piksel, jarak antar piksel, serta orientasi sudut dalam gambar, sehingga memungkinkan ekstraksi fitur tekstur yang lebih mendetail.



Gambar 2. Sudut GLCM

Gambar 2 sudut GLCM yang digunakan untuk menghitung celah. Selain 0°, 45°, 90°, dan 135°, orientasi sudut diukur dalam derajat, dan ruang yang digunakan biasanya 1 piksel[16]. Rumus (1) hingga (4) di bawah ini menggambarkan perbedaan nyata, hubungan, energi, dan homogenitas yang ada di antara fitur tekstur yang diperiksa menggunakan GLCM [17] :

a. Kontras (*Contrast*)

perbedaan tingkat kecerahan atau intensitas antara bagian terang dan gelap dalam gambar. Semakin besar perbedaan intensitas antara dua area, semakin tinggi kontrasnya. Kontras dihitung dengan mengukur perbedaan intensitas piksel yang ada dalam citra, yang berfungsi untuk meningkatkan ketajaman dan detail visual. Rumus buat menghitung kontras suatu gambar ialah dari persamaan berikut:

$$\text{Contrast} = \sum_{i,j=0}^{N-1} P_{i,j} (i - j)^2 \quad (1)$$

Rumus ini, $(P(i, j))$ merupakan probabilitas kemunculan pasangan piksel dengan nilai intensitas i dan j dalam citra. Variabel i dan j mewakili indeks baris dan kolom dalam matriks GLCM, sementara N menunjukkan jumlah tingkat keabuan dalam citra. Bagian $(i - j)^2$ merepresentasikan selisih kuadrat antara intensitas piksel dalam pasangan tertentu. Rumus ini digunakan untuk menghitung kontras, yang mengukur tingkat perbedaan lokal dalam suatu citra. Semakin tinggi nilai kontras, semakin besar variasi intensitas piksel dalam gambar, yang berhubungan dengan tingkat kekasaran atau detail tekstur dalam citra.

b. Korelasi (*Correlation*)

konteks citra grayscale merujuk pada pengukuran sejauh mana dua citra memiliki pola atau struktur yang serupa. Korelasi dihitung dengan membandingkan nilai intensitas piksel di citra satu dengan nilai intensitas piksel di citra lainnya. Rumus buat menghitung korelasi bisa dicermati pada persamaan berikut :

$$Correlation = \sum_{i,j=0}^{N-1} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right] \quad (2)$$

Rumus ini digunakan untuk menghitung korelasi dalam Gray Level Co-occurrence Matrix (GLCM), yang mengukur sejauh mana suatu piksel berkorelasi dengan piksel di sekitarnya dalam suatu citra. Dalam rumus ini, i dan j mewakili indeks baris dan kolom dalam matriks GLCM. Nilai μ_i dan μ_j masing-masing merupakan rata-rata intensitas piksel untuk baris i dan kolom j dalam GLCM. Sementara itu, σ_i^2 dan σ_j^2 adalah varians dari intensitas piksel pada baris dan kolom yang bersangkutan. Rumus ini digunakan untuk menilai hubungan linear antara pasangan piksel dalam suatu gambar. Nilai korelasi yang tinggi menunjukkan bahwa intensitas piksel di suatu area memiliki pola yang lebih seragam, sedangkan nilai yang rendah menunjukkan adanya variasi intensitas yang lebih besar.

c. Energi (*Energy*)

Energi merupakan konteks analisis gambar atau pemrosesan citra, merujuk pada ukuran homogenitas atau keseragaman distribusi intensitas piksel (sering disebut intensitas abu-abu) dalam suatu gambar. Rumus energi ditunjukkan dalam persamaan berikut:

$$Energy = \sqrt{\sum_{i,j=0}^{N-1} P_{i,j}^2} \quad (3)$$

Rumus ini digunakan untuk menghitung energi dalam Gray Level Co-occurrence Matrix (GLCM), yang menunjukkan tingkat keseragaman pola tekstur dalam suatu citra. Dalam rumus ini, $(P(i, j))$ merupakan probabilitas kemunculan pasangan piksel dengan nilai intensitas i dan j dalam matriks GLCM, sedangkan N adalah jumlah tingkat keabuan dalam citra. Energi mengukur keberagaman distribusi nilai piksel dalam suatu gambar. Nilai energi yang tinggi menunjukkan bahwa citra memiliki tekstur yang lebih seragam dengan sedikit variasi intensitas, sedangkan nilai yang rendah menunjukkan adanya banyak variasi dalam distribusi intensitas piksel.

d. Homogenitas (*Homogeneity*)

Homogenitas adalah konteks citra digital merujuk pada tingkat keseragaman atau kesamaan dalam distribusi intensitas piksel suatu citra. Citra yang homogen memiliki intensitas piksel yang hampir seragam di seluruh area, artinya perbedaan antara nilai intensitas piksel satu dengan yang lainnya sangat kecil. Dalam citra yang tidak homogen, akan terdapat variasi yang lebih besar antara piksel-piksel yang ada. Persamaan di bawah ini menunjukkan rumus untuk menghitung homogenitas:

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (4)$$

Rumus ini digunakan untuk menghitung homogenitas dalam Gray Level Co-occurrence Matrix (GLCM), yang mengukur kedekatan nilai piksel dalam suatu citra. Dalam rumus ini, $(P(i, j))$ adalah probabilitas kemunculan pasangan piksel dengan intensitas i dan j dalam matriks GLCM, sementara N menunjukkan jumlah tingkat keabuan dalam citra. Pembilang $(P(i, j))$ menunjukkan distribusi probabilitas pasangan piksel, sedangkan penyebut $1 + (i - j)^2$ berfungsi untuk menekan nilai kontribusi pasangan piksel yang memiliki perbedaan intensitas yang besar. Homogenitas bernilai tinggi ketika variasi intensitas antar piksel dalam citra rendah, menunjukkan bahwa piksel-piksel dalam gambar cenderung memiliki nilai yang mirip atau seragam. Sebaliknya, jika nilai homogenitas rendah, ini menandakan adanya perbedaan intensitas yang signifikan dalam citra, yang mencerminkan tingkat kompleksitas tekstur yang lebih tinggi.

2.5 Klasifikasi

Klasifikasi merupakan proses mengelompokkan objek atau info ke dalam kategori atau grup eksklusif berdasarkan karakteristik yang sudah ditentukan berasal sejumlah kategori yg tersedia [18]. Meskipun mencapai akurasi 100% tidak selalu memungkinkan, sistem klasifikasi wajib mampu menyajikan seluruh isu secara efektif. Estimasi penerapan klasifikasi beberapa kali menggunakan kisi-kisi yang tidak teratur, dengan tujuan melatih kemampuan pengenalan terhadap pola yang terdapat di data serta mengklasifikasikan objek kepada kategori yang benar [19]. Setelah analisis

relevansi, langkah selanjutnya dalam proses klasifikasi adalah menemukan atribut yang relevan. Tabel 1 menunjukkan jumlah akurasi prediksi hasil klasifikasi[20].

Tabel 1. Confusion Matrix

	Prediksi (+)	Prediksi (-)
Aktual (+)	TP (True Positives)	FN (False Negatives)
Aktual (-)	FP (False Positives)	TN (True Negatives)

Tabel 1 merupakan tabel Confusion Matrix, digunakan sebagai mengevaluasi kinerja model klasifikasi. Confusion matrix terdiri dari empat elemen utama, yang masing-masing merepresentasikan jumlah prediksi dalam setiap kategori. TP (Benar Positif) adalah jumlah data positif yang diidentifikasi dengan benar sebagai data positif, FP (Positif Palsu) adalah jumlah data negatif yang salah diklasifikasikan sebagai data positif, TN (Benar Negatif) adalah jumlah data negatif yang diidentifikasi dengan benar sebagai data negatif, dan FN (Negatif Palsu) adalah jumlah data positif yang salah diklasifikasikan sebagai data negatif.

a. Accuracy

Akurasi adalah suatu parameter yang digunakan untuk mengukur tingkat ketepatan hasil prediksi dalam suatu proses klasifikasi terhadap kategori yang sesuai. Perhitungan akurasi klasifikasi bertujuan untuk mengevaluasi sejauh mana model dapat mengklasifikasikan data dengan benar ke dalam kelas yang telah ditentukan, rumus yang dapat digunakan:

$$Accuracy = \frac{\text{jumlah klasifikasi benar}}{\text{jumlah total data}} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

Akurasi dalam evaluasi model klasifikasi dihitung dengan rumus $\frac{TP+TN}{TP+TN+FP+FN}$, dimana TP (True Positive) adalah jumlah sampel positif yang diklasifikasikan dengan benar, TN (True Negative) adalah jumlah sampel negatif yang diklasifikasikan dengan benar, FP (False Positive) adalah jumlah sampel negatif yang salah diklasifikasikan sebagai positif, dan FN (False Negative) adalah jumlah sampel positif yang salah diklasifikasikan sebagai negatif, sehingga akurasi menggambarkan proporsi prediksi yang benar terhadap keseluruhan data, dengan nilai tinggi menunjukkan performa model yang baik, sementara nilai rendah menunjukkan kesalahan klasifikasi yang signifikan.

b. Precision

Presisi adalah metrik yang digunakan untuk menilai sejauh mana prediksi model tepat dalam mengklasifikasikan suatu kategori tertentu. Berikut ini adalah rumus yang digunakan untuk menghitung presisi.:

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

Presisi dalam evaluasi model klasifikasi dihitung dengan $\frac{TP}{TP+FP}$, di mana TP (True Positive) adalah jumlah sampel positif yang diklasifikasikan dengan benar, sedangkan FP (False Positive) adalah jumlah sampel negatif yang salah diklasifikasikan sebagai positif, sehingga presisi mengukur seberapa akurat model dalam mengklasifikasikan suatu sampel sebagai positif, dengan nilai tinggi menunjukkan bahwa model jarang memberikan prediksi positif yang salah, sedangkan nilai rendah menunjukkan banyaknya kesalahan dalam mengidentifikasi sampel positif.

c. Recall

Recall adalah parameter yang mengukur seberapa banyak data dari suatu kelas dapat diprediksi secara akurat. Berikut rumus untuk menghitung recall:

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

Recall atau sensitivitas dalam evaluasi model klasifikasi dihitung dengan rumus $\frac{TP}{TP+FN}$, di mana TP (True Positive) adalah jumlah sampel positif yang diklasifikasikan dengan benar, sedangkan FN (False Negative) adalah jumlah sampel positif yang salah diklasifikasikan sebagai negatif, sehingga recall mengukur sejauh mana model mampu mendeteksi sampel positif dengan benar, dengan nilai tinggi menunjukkan bahwa model jarang melewatkan sampel positif, sementara nilai rendah menunjukkan banyaknya kesalahan dalam mengidentifikasi sampel positif yang sebenarnya.

2.6 Klasifikasi Dengan K-Nearest Neighbor

Pada Penelitian ini, data diklasifikasikan menggunakan algoritma K-Nearest Neighbor (KNN). KNN berfungsi mengklasifikasikan objek berdasarkan seberapa dekatnya objek tersebut dengan data pelatihan. Karena ide dasar di balik pendekatan ini mirip sebuah teknik pengelompokan yaitu, mengklasifikasikan data baru ke dalam kelompok berdasarkan masing-masing data yang terdekat, pendekatan ini sangat mudah dan sederhana untuk dijalankan. Rumus berikut dapat digunakan untuk menentukan jarak dalam algoritma KNN menggunakan Euclidean Distance[21] - [23].

$$Dx,y = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (8)$$

Dalam penjelasan rumus di atas, terdapat data pelatihan (x) dan data pengujian (y), yang masing-masing memiliki sejumlah atribut individu yang diberi label dari 1 hingga n . Setiap atribut dapat diidentifikasi menggunakan indeks i , yang berkisar dari 1 hingga n . Jarak Kedekatan (D) berfungsi sebagai ukuran untuk menilai tingkat kedekatan antara dua titik data.

2.7 Klasifikasi Dengan Support Vector Machine

Membandingkan teknik penambangan statistik, khususnya sistem vektor bantuan (SVM), prosedur solusi pembelajaran terbimbing klasik yang dirancang untuk memecahkan dilema klasifikasi, digunakan dalam proses klasifikasi. Ide utama di balik SVM adalah menemukan hiperbidang terdepan yang dapat memisahkan dua kelas dalam ruang masukan. Untuk mengatasi masalah non-linier, SVM disempurnakan menggunakan teknik kernelisasi, yang menggunakan fungsi kernel untuk mengganti representasi informasi dari dimensi rendah ke dimensi yang lebih tinggi secara lebih efisien. SVM memiliki berbagai jenis kernel, termasuk kernel linier, polinomial, karakter basis Radial Gaussian (RBF), dan kernel sigmoid. Tujuan SVM adalah untuk menemukan hyperplane terbaik yang berfungsi sebagai pemisah antara dua kelas data, yaitu kelas positif dan kelas negatif [24] - [26]. rumus dibawah ini merupakan persamaan SVM.:

$$wT \cdot x + b = 0 \quad (9)$$

Pada rumus di atas, w adalah vektor normal yang menentukan orientasi bidang hiper, x merupakan vektor data dalam ruang fitur, dan b adalah bias yang mengatur posisi bidang tanpa mengubah arahnya. Kombinasi ketiganya digunakan dalam perhitungan untuk memisahkan kelas secara optimal dalam model klasifikasi berbasis Support Vector Machine (SVM).

3. HASIL DAN PEMBAHASAN

3.1 Dataset Tanda Tangan

Penelitian ini menggunakan kumpulan data gambar tanda tangan yang diperoleh melalui pengambilan tanda tangan langsung, selain data dari platform publik Kaggle, yang terbuka agar bisa diakses dan dilihat oleh peneliti yang lain. Setiap gambar dalam kumpulan data memiliki resolusi standar, yang memudahkan ekstraksi dan klasifikasi fitur. Kumpulan data dibagi menjadi 21 kelas, yang masing-masing memiliki 50 gambar tanda tangan dari orang yang berbeda. Gambar 3 di bawah ini menunjukkan contoh gambar yang digunakan :



Gambar 3. Dataset Tanda Tangan

Gambar 3 memperlihatkan data yang dipakai untuk membuat model klasifikasi, dan data pengujian digunakan untuk mengevaluasi kinerja model secara objektif adapun sejumlah sampel dari kumpulan data yang digunakan dalam penelitian, yang dipisahkan yaitu 80% data yang dilatih dan 20% data yang akan diuji.

3.2 Hasil Ekstraksi Fitur GLCM

Penelitian ini menerapkan metode GLCM menggunakan dalam penelitian ini dalam mengekstraksi fitur dari foto tanda tangan. GLCM menghitung frekuensi pasangan piksel dengan nilai intensitas Untuk menaikkan akurasi contoh pembagian terstruktur mengenai dengan menyediakan isu yg lebih komprehensif tentang pola tekstur pada gambar, metode GLCM menggunakan dalam penelitian ini untuk mengekstraksi fitur dari foto tanda tangan. GLCM menentukan frekuensi pasangan piksel dengan nilai intensitas tertentu yang muncul pada jarak dan orientasi tertentu dalam gambar. Metode ini memetakan variasi intensitas piksel di sekitarnya, sehingga memungkinkan analisis yang lebih menyeluruh. Fitur yang diekstraksi dalam penelitian ini meliputi kontras, korelasi, energi, dan homogenitas. Selain itu, seluruh kumpulan data, baik data pelatihan maupun data pengujian, dianalisis secara konsisten untuk memastikan hasil yang valid dan andal.

Hasil ekstraksi fitur masing-masing gambar (tanda tangan) pada setiap fitur mempunyai nilai unik, bukan hanya memperkaya info tekstur yang tersedia tetapi jua menaikkan kemampuan model buat mengenali dan mengklasifikasikan tanda tangan berdasarkan fitur yang lebih halus dan kompleks. Pengenalan pola tulisan tangan menggunakan metode deskripsi KNN dan SVM menggunakan hasil ekstraksi fitur. Sementara SVM mencari hyperplane yang membagi kelas data dengan margin yang besar, KNN menemukan tetangga terdekat dari data uji dalam ruang fitur yang diekstraksi. Tabel 2 menunjukkan output ekstraksi fitur:

Tabel 2. Hasil Ekstraksi Fitur GLCM dari Dataset Tanda Tangan

No	Contrast	Correlation	Energy	Homogeneity	Label
0	102.300940	0.933228	0.355205	0.543965	dataset_ttd 20

No	Contrast	Correlation	Energy	Homogeneity	Label
1	96.803217	0.933584	0.222479	0.450331	dataset_ttd 20
2	113.238521	0.924984	0.304344	0.506785	dataset_ttd 20
3	112.750686	0.927567	0.259419	0.478066	dataset_ttd 20
4	105.759261	0.927385	0.745551	0.797383	dataset_ttd 20
....
1045	115.231537	0.969620	0.856383	0.883089	haidar
3402	115.231537	0.969620	0.856383	0.883089	haidar
3403	115.231537	0.969620	0.856383	0.883089	haidar
3404	115.231537	0.969620	0.856383	0.883089	haidar
3410	115.231537	0.969620	0.856383	0.883089	haidar

Tabel 2 Hasil Ekstraksi Fitur GLCM dari Dataset Tanda Tangan adalah menunjukkan hasil dari kumpulan data untuk digunakan dalam prosedur klasifikasi. Prosedur ekstraksi fitur dari setiap gambar dalam koleksi ini menghasilkan nilai yang berbeda untuk setiap fitur GLCM, yang menunjukkan variasi dalam sifat tekstur setiap gambar tanda tangan. Variasi ini membantu proses klasifikasi membedakan tingkatan yang berbeda.

3.3 Hasil Klasifikasi KNN

Setelah melakukan Metode GLCM untuk memperoleh fitur, yang kemudian disimpan dan digunakan sebagai input untuk pelatihan dan pengujian model. Algoritma KNN kemudian digunakan untuk mengklasifikasikan data dalam tiga skenario, mengevaluasi bagaimana kinerja dipengaruhi oleh variasi dalam pembagian dataset dan nilai k.

Dalam skenario pertama, hasil ini menunjukkan bahwa penggunaan nilai k yang lebih kecil cenderung memberikan akurasi prediksi yang lebih baik, mungkin karena model menjadi lebih sensitif terhadap struktur yang terkandung dalam data. Dalam skenario pertama, distribusi himpunan data ditetapkan sebesar 60% untuk pelatihan dan 40% untuk pengujian. Nilai k yang diuji adalah 3, 5, dan 7. Akurasi terbaik dalam skenario ini adalah pada k=3, dengan persentase mencapai 82.14%.

Skenario kedua, dengan nilai k yang sama seperti skenario pertama (3, 5, dan 7), kumpulan data dibagi ke dalam proporsi yang berbeda dalam skenario kedua, dengan 70% digunakan untuk pelatihan dan 30% untuk pengujian. Akurasi tertinggi yang dicapai dalam skenario ini adalah 86.03% pada k = 3, yang mungkin disebabkan oleh kumpulan data pelatihan yang lebih besar, yang memungkinkan model untuk mengidentifikasi pola dengan lebih tepat dan menghasilkan prediksi yang lebih baik.

Pada Skenario terakhir, Dengan menggunakan nilai k yang sama seperti dua skenario pertama, kumpulan data dibagi dalam skenario akhir dengan 80% untuk data pelatihan dan 20% untuk pengujian. Hasilnya menunjukkan bahwa akurasi terbaik dalam skenario ini adalah 91.42% pada k=3. Alasan peningkatan akurasi ini adalah karena model memiliki akses ke lebih banyak data pelatihan, yang memungkinkan jumlah pola yang dipelajari meningkat dan meningkatkan daya prediksi dibandingkan dengan skenario pertama.

Nilai k=3 secara konsisten memberikan akurasi tertinggi untuk setiap pengujian pemisahan himpunan data. menunjukkan bahwa nilai k yang kecil signifikan dalam teknik KNN dan bahwa pemisahan himpunan data yang lebih besar untuk pelatihan dapat menghasilkan model yang lebih kuat. Ini menampilkan nilai akurasi, presisi, dan perolehan kembali untuk setiap skenario, pada Tabel 3 berikut:

Tabel 3. Hasil Klasifikasi KNN

Pembagian Dataset	Nilai k	Akurasi	Presisi	Recall
60 : 40	3	82.14%	83.19%	82.14%
	5	77.38%	78.81%	77.38%
	7	75.47%	76.87%	75.47%
70 : 30	3	86.03%	88.16%	86.03%
	5	78.73%	80.10%	78.73%
	7	76.19%	77.96%	76.19%
80 : 20	3	91.42%	93.05%	91.42%
	5	82.38%	83.76%	82.38%
	7	75.71%	77.71%	75.71%

Tabel 3. Hasil Klasifikasi KNN menyajikan perbandingan yang cukup luas terkait dengan kinerja model dalam berbagai skenario, mencakup pengaruh Keragaman pemisahan kumpulan data dan nilai k terhadap kinerja algoritma KNN. Dari data yang ditampilkan, terlihat jelas bagaimana perubahan nilai k yang kecil memengaruhi kinerja KNN dalam proses klasifikasi. Hal ini berdampak pada tingkat akurasi, yang berperan penting dalam optimasi analisis citra digital menggunakan metode GLCM.

3.4 Hasil Klasifikasi SVM

Klasifikasi diterapkan dengan menggunakan metode KNN, sementara proses klasifikasi lebih lanjut dilakukan dengan ekstraksi fitur GLCM yang diproses menggunakan algoritma SVM dengan kernel linear. Pemilihan SVM dengan kernel linear dipilih karena kemampuannya dalam menangani dataset berdimensi tinggi secara efisien dan cepat.

Selain itu, SVM dengan kernel linear memiliki keunggulan dalam menghasilkan hyperplane yang memaksimalkan margin antar kelas yang berbeda, dengan cara membaginya secara efektif dalam ruang fitur.

Pengujian dengan SVM menunjukkan tingkat akurasi yang tinggi, mencapai 97,70%. Hasil ini menegaskan bahwa SVM mampu mengenali pola tulisan tangan dengan ketelitian yang baik. Selain akurasi, evaluasi juga dilakukan menggunakan metrik lain, seperti presisi dan recall. Presisi tercatat sebesar 98,73%, yang mengindikasikan bahwa sebagian besar Prediksi yang masuk akal oleh model memang benar. Untuk sementara, recall sebesar 97,70% menunjukkan bahwa model berhasil mengidentifikasi sebagian besar data positif secara akurat.

Keseluruhan, Penerapan SVM yang menggunakan kernel linear memberikan performa dengan sangat baik prediksi pola tanda tangan. Keunggulan dibantu karena akurasi yang tinggi serta keseimbangan antara presisi serta recall, dimana menandakan bahwa model tidak hanya mampu mengidentifikasi pola dengan tepat, tetapi juga dapat mendeteksi sebagian besar pola yang terdapat dalam data.

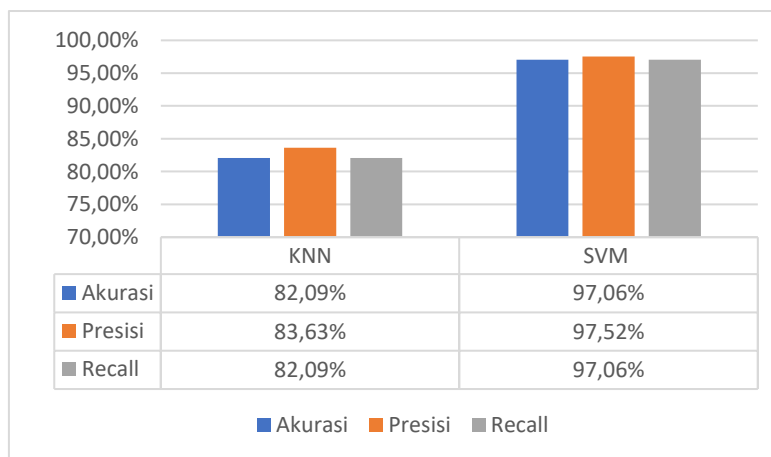
3.5 Hasil Klasifikasi SVM

Uji perbandingan antara KNN dan SVM dengan kernel linear menunjukkan bahwa SVM memiliki performa yang lebih unggul dalam mengenali pola tulisan tangan. SVM berhasil mencapai akurasi tertinggi sebesar 97,70%, sementara KNN hanya memperoleh akurasi 85,22%. Meskipun perbedaan tersebut tidak terlalu signifikan, hasil ini mengindikasikan bahwa SVM lebih efektif dalam menangani dataset yang lebih rumit.

Selain akurasi, SVM juga memiliki keunggulan dalam hal kecepatan dan efisiensi pada proses pelatihan dan prediksi, terutama ketika jumlah data latih semakin besar. Perbandingan detail mengenai akurasi, presisi, dan recall dapat ditemukan pada Tabel 4 dan Gambar 4.

Tabel 4. Perbandingan Akurasi, Presisi, dan Recall

No	Algoritma	Akurasi	Presisi	Recall
1	KNN	82,09%	83,63%	82,09%
2	SVM	97,06%	97,52%	97,06%



Gambar 4. Perbandingan Akurasi, Presisi, serta Recall

Pada Tabel 4 perbandingan akurasi, Presisi dan Recall adalah menampilkan keseluruhan tentang perbandingan algoritma KNN dan SVM dalam mengidentifikasi pola tanda tangan digital, menggunakan tiga metrik utama: akurasi, presisi, dan recall.

Sementara itu, Gambar 4. Perbandingan Akurasi, Presisi dan Recall adalah memvisualisasikan perbandingan ini dalam bentuk diagram batang, yang bertujuan untuk memberikan gambaran lebih jelas mengenai perbedaan kinerja antara algoritma KNN dan SVM.

4. KESIMPULAN

Penelitian ini membahas efektivitas dua algoritma klasifikasi, yaitu K-Nearest Neighbors (KNN) dan Support Vector Machine (SVM), dalam mengenali pola tanda tangan digital. Proses ekstraksi fitur tekstur tanda tangan dilakukan menggunakan metode Gray-Level Co-occurrence Matrix (GLCM), yang menghasilkan atribut utama seperti kontras, korelasi, energi, dan homogenitas untuk membedakan setiap tanda tangan berdasarkan pola uniknya. Hasil eksperimen menunjukkan bahwa KNN dengan nilai $k = 3$ mencapai akurasi tertinggi sebesar 91,42% saat menggunakan 80% data latih dan 20% data uji. Sementara itu, SVM dengan kernel linear menghasilkan akurasi yang lebih tinggi, yaitu 97,06%, menunjukkan keunggulannya dalam menangani dataset dengan efisiensi komputasi yang lebih baik. Selain itu, SVM juga mencatat performa optimal pada metrik presisi dan recall, masing-masing sebesar

97,52% dan 97,06%, yang mencerminkan kemampuannya dalam mengenali serta mengklasifikasikan tanda tangan dengan lebih akurat. Dari hasil perbandingan, SVM dengan kernel linear tidak hanya lebih akurat tetapi juga lebih stabil dibandingkan KNN. SVM lebih mampu menangani kompleksitas serta variasi tanda tangan, menjadikannya lebih efektif untuk sistem pengenalan tanda tangan digital. Teknologi ini dapat diterapkan dalam berbagai bidang, termasuk perbankan, administrasi pemerintahan, dan sistem keamanan berbasis biometrik. Dengan akurasi tinggi, sistem ini mampu membantu mencegah pemalsuan tanda tangan serta meningkatkan keamanan transaksi digital. Namun, penelitian ini masih terbatas pada dataset yang relatif kecil. Oleh karena itu, penelitian selanjutnya disarankan untuk memperluas cakupan dengan menggunakan dataset yang lebih besar, mempertimbangkan faktor seperti kualitas gambar tanda tangan, serta mengeksplorasi algoritma klasifikasi lain guna meningkatkan akurasi dan efisiensi sistem secara keseluruhan.

REFERENCES

- [1] M. D. R. Priyanto, E. Setiawan, and H. Fitriyah, "Sistem Biometrik Gerakan Tanda Tangan Menggunakan Sensor MPU6050 dengan Metode Backpropagation," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 4, no. 8, pp. 2338–2348, Aug. 2020, Accessed: May 03, 2025. [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/7623>
- [2] E. Wahyuni, S. Rahman, and A. Risma, "Keabsahan Digital Signature/Tanda tangan Elektronik Dinjau Dalam Perspektif Hukum Perdata dan UU ITE," *Journal of Lex Generalis (JLG)*, vol. 3, no. 5, pp. 1082–1098, May 2022, Accessed: May 03, 2025. [Online]. Available: <https://mail.pasca-umi.ac.id/index.php/jlg/article/view/886>
- [3] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "DeepSign: Deep On-Line Signature Verification," *IEEE Trans Biom Behav Identity Sci*, vol. 3, no. 2, pp. 229–239, Apr. 2021, doi: 10.1109/TBIOM.2021.3054533.
- [4] L. Amaludin and A. Rahmatulloh, "Penerapan ECDSA dan BLAKE2B Untuk Membentuk Tanda Tangan Digital Sebagai Autentikasi Dokumen," *Jurnal Informatika dan Multimedia*, vol. 16, no. 2, pp. 20–26, Dec. 2024, doi: 10.33795/JTMM.V16I2.6599.
- [5] S. Jain, M. Khanna, and A. Singh, "Comparison among different CNN Architectures for Signature Forgery Detection using Siamese Neural Network," *Proceedings - IEEE 2021 International Conference on Computing, Communication, and Intelligent Systems, ICCIS 2021*, pp. 481–486, Feb. 2021, doi: 10.1109/ICCIS51004.2021.9397114.
- [6] H. F. Lumentut, M. S. Borman, and N. Handayati, "Kekuatan Pembuktian Dokumen Elektronik Dengan Tanda Tangan Elektronik Di Dalam Hukum Acara Perdata," *Jurnal Penelitian Ilmiah Multidisiplin*, vol. 8, no. 10, pp. 2118–2451, Oct. 2024, Accessed: May 14, 2025. [Online]. Available: <https://oaj.jurnalhst.com/index.php/jpim/article/view/5159>
- [7] R. Jannah, M. Walid, and H. Hoiriyah, "Sistem Pengenalan Citra Dokumen Tanda Tangan Menggunakan Metode CNN (Convolutional Neural Network)," *Energy : Jurnal Ilmiah Ilmu-ilmu Teknik*, vol. 12, no. 2, pp. 54–61, Dec. 2022, doi: 10.51747/ENERGY.V12I2.1116.
- [8] A. C. Vidyanti, I. Riati, and A. Ramadhanu, "Identification of Signature Authenticity Using Binary Extraction and K-nearest Neighbor Feature Methods," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 13, no. 2, pp. 274–279, Jun. 2024, doi: 10.32736/SISFOKOM.V13I2.2063.
- [9] K. A. Safitri and R. Wulanningrum, "Aplikasi Pengenalan Pola Tulisan Tangan Menggunakan Metode Support Vector Machine," *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, vol. 4, no. 1, pp. 201–206, Dec. 2020, doi: 10.29407/INOTEK.V4I1.197.
- [10] C. R. (Chyntia) Widiawati and S. (Suliswaningsih) Suliswaningsih, "Analisa Hasil Perbandingan Poly Kernel dan Normalisasi Poly Kernel pada Support Vector Machine sebagai Metode Klasifikasi Citra Tanda Tangan," *Jurnal Khatulistiwa Informatika*, vol. 9, no. 1, pp. 71–77, Apr. 2022, doi: 10.31294/INF.V9I1.11288.
- [11] M. Kurniawan, N. Saidatin, D. H. Nugroho, I. T. Adhi, and T. Surabaya, "Implementasi Shape Feature dan K-Nearest Neighbor untuk Klasifikasi Tanda Tangan," *Prosiding Seminar Nasional Sains dan Teknologi Terapan*, vol. 1, no. 1, pp. 155–162, Sep. 2020, Accessed: May 24, 2025. [Online]. Available: <https://ejurnal.itats.ac.id/sntekpan/article/view/1230>
- [12] R. Pujiyanto *et al.*, "Pengolahan Citra Dan Metode Support Vector Machine (SVM) Dalam Pengenalan Pola Tanda Tangan," *Jurnal Rekayasa Komputasi Terapan*, vol. 1, no. 01, pp. 2776–5873, Mar. 2021, doi: 10.30998/JRKT.V1I01.4048.
- [13] M. Septiani, "Pengenalan Pola Batik Lampung Menggunakan Metode Principal Component Analysis," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 2, no. 4, pp. 552–558, Feb. 2021, doi: 10.33365/JATIKA.V2I4.1612.
- [14] A. Q. Maharani and T. Ardiansah, "Kombinasi Metode Multi-Attribute Utility Theory dan Pivot Pairwise Relative Criteria Importance Assessment Dalam Penentuan Lulusan Terbaik," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 4, p. 2074, Oct. 2023, doi: 10.30865/MIB.V7I4.6884.
- [15] G. D. Angel and R. Wulanningrum, "Machine Learning untuk Identifikasi Tanda Tangan Menggunakan GLCM dan Euclidean Distance," *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, vol. 4, no. 1, pp. 297–301, Dec. 2020, doi: 10.29407/INOTEK.V4I1.213.
- [16] W. K. Oktalao, I. M. B. Atmaja Darmawan, I. W. Santiyasa, I. P. G. Hendra Suputra, and I. G. N. Anom Cahyadi Putra, "Klasifikasi Motif Kain Tradisional Cepuk Menggunakan GLCM dan KNN," *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, vol. 11, no. 3, p. 545, Jul. 2022, doi: 10.24843/JLK.2023.V11.I03.P10.
- [17] L. Hakim, S. P. Kristanto, D. Yusuf, and F. N. Afia, "Pengenalan Motif Batik Banyuwangi Berdasarkan Fitur Grey Level Co-Occurrence Matrix," *Jurnal Teknoinfo*, vol. 16, no. 1, p. 1, Jan. 2022, doi: 10.33365/JTI.V16I1.1320.
- [18] N. Widiastuti, A. Hermawan, and D. Avianto, "Komparasi Algoritma Klasifikasi Datamining Untuk Prediksi Minat Pencari Kerja," *Jurnal Teknoinfo*, vol. 17, no. 1, pp. 219–227, Jan. 2023, doi: 10.25126/JTIK.202073080.
- [19] E. Fitriani, "Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan," *SISTEMASI*, vol. 9, no. 1, p. 103, Jan. 2020, doi: 10.32520/STMSI.V9I1.596.
- [20] A. H. Yunial, "Analisis Optimasi Algoritma Klasifikasi Support Vector Machine, Decision Trees, dan Neural Network Menggunakan Adaboost dan Bagging," *Jurnal Informatika Universitas Pamulang*, vol. 5, no. 3, pp. 247–260, Sep. 2020, doi: 10.32493/INFORMATIKA.V5I3.6609.



- [21] Y. M. Y. MZ, R. M. Fadilla, and I. Pratama, "Implementasi Metode K-Nearst Neighbor Berbasis Euclidean Distance Untuk Klasifikasi Penerimaan Vaksin Covid-19," *Informasi Interaktif*, vol. 6, no. 3, pp. 111–117, Sep. 2021, Accessed: May 24, 2025. [Online]. Available: <https://e-journal.janabadra.ac.id/index.php/informasiinteraktif/article/view/1426>
- [22] A. R. D. Nugraha, K. Auliasari, and Y. A. Pranoto, "Implementasi Metode K-Nearest Neighbor (Knn) Untuk Seleksi Calon Karyawan Baru (Studi Kasus : BFI Finance Surabaya)," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 4, no. 2, pp. 14–20, Oct. 2020, doi: 10.36040/JATI.V4I2.2656.
- [23] A. Purnamawati, M. N. Winnarto, and M. Mailasari, "Analisis Cart (Classification And Regression Trees) Untuk Prediksi Pengguna Sepeda Berdasarkan Cuaca," *Jurnal Teknoinfo*, vol. 16, no. 1, pp. 14–19, Jan. 2022, doi: 10.33365/JTI.V16I1.1478.
- [24] A. (Atang) Saepudin, R. (Riska) Aryanti, E. (Eka) Fitriani, and D. (Dahlia) Dahlia, "Optimasi Algoritma SVM dan K-NN Berbasis Particle Swarm Optimization pada Analisis Sentimen Fenomena Tagar #2019GantiPresiden," *Jurnal Khatulistiwa Informatika*, vol. 6, no. 1, pp. 95–102, Jan. 2020, doi: 10.31294/JTK.V6I1.6828.
- [25] A. Liani, U. Enri, and Y. Umaidah, "Analisis Perbandingan Kernel Algoritma Support Vector Machine dalam Mengklasifikasikan Skripsi Teknik Informatika berdasarkan Abstrak," *JOINS (Journal of Information System)*, vol. 5, no. 2, pp. 240–249, Nov. 2020, doi: 10.33633/JOINS.V5I2.3715.
- [26] D. Arisandi, T. Sutrisno, and I. Kurniawan, "Klasifikasi Opini Masyarakat Di Twitter Tentang Kebocoran Data Yang Terjadi Di Indonesia Menggunakan Algoritma Svm," *Jurnal Teknika*, vol. 15, no. 2, pp. 75–80, Sep. 2023, doi: 10.30736/JT.V15I2.993.