

Deteksi Kepribadian Melalui Margin Pada Tulisan Tangan Menggunakan *Random Forest*

Agung Gumilang¹, Soffiana Agustin²

^{1,2} Universitas Muhammadiyah Gresik, Jl. Sumatera No.101, Gn. Malang, Gresik, Indonesia
E-mail: agunggumilang716@gmail.com¹, soffiana@umg.ac.id²

Abstract - Graphology in handwriting is a technique for assessing a person's personality by examining various aspects of their handwriting. Each individual's handwriting is unique and has its own characteristics, making it analyzable through a grapho-test. Graphology is used in various fields such as talent identification and employee placement. The challenges faced include the grapho-test being time-consuming, costly, and finding efficient ways to detect character through handwriting while achieving accurate results. Research applying artificial intelligence in the field of graphology has been conducted by several researchers, but most have used multiple characteristics such as margin, slant, text size, and letter shape without classifying them into a single characteristic, resulting in less accurate outcomes. Therefore, this study presents a model for applying computer vision in graphology, focusing on a single characteristic, namely margin, and classifying it into four classes: wide, narrow, widening, and narrowing. Personality is classified using the *random forest* method, and based on the test results, it is known that the *random forest* method provides 95% accuracy, demonstrating better performance and yielding accurate handwriting detection results without the need for a psychology expert.

Keywords - handwriting, personality, computer vision, random forest

Intisari - Grafologi pada tulisan tangan merupakan teknik menilai kepribadian seseorang dilihat dari berbagai aspek seperti tulisan tangan. Tulisan tangan setiap orang bersifat unik dan memiliki karakteristiknya, sehingga dapat dianalisis dengan melakukan *grafo-test*. Grafologi digunakan dalam berbagai bidang seperti penentuan minat bakat dan penempatan karyawan. Permasalahan yang terjadi adalah *grafo-test* membutuhkan waktu yang lama, biaya yang tidak murah dan bagaimana cara untuk mendeteksi karakter lewat tulisan tangan dengan waktu yang efisien dan mendapatkan hasil yang akurat. Penelitian yang menerapkan kecerdasan buatan dibidang grafologi sudah dilakukan beberapa peneliti, akan tetapi sebagian besar menggunakan beberapa karakteristik seperti margin, kemiringan, ukuran tulisan, bentuk huruf dan tidak diklasifikasikan kedalam satu karakteristik dengan hasil yang kurang akurat. Oleh karena itu, pada penelitian ini memberikan model penerapan *computer vision* dalam bidang grafologi yang berfokus pada satu karakteristik yaitu margin dan diklasifikasikan kedalam empat kelas seperti lebar, sempit, melebar menyempit. Kepribadian diklasifikasikan menggunakan metode *random forest*, berdasarkan hasil pengujian diketahui metode *random forest* memberikan akurasi 95% memiliki kinerja yang lebih baik dan memberikan hasil dari pendeteksian tulisan tangan yang akurat tanpa perlunya ahli psikologi.

Kata Kunci - tulisan tangan, kepribadian, visi komputer, random forest

I. PENDAHULUAN

Grafologi merupakan cabang ilmu pengetahuan dalam menginterpretasikan karakter dan kepribadian seseorang melalui tulisan tangan. Karakter atau kepribadian seseorang merujuk pada kumpulan sifat, perilaku, sikap, dan pola pikir yang membentuk identitas unik individu tersebut. Ini adalah kombinasi dari berbagai faktor internal dan eksternal yang mempengaruhi cara seseorang berinteraksi dengan dunia sekitarnya dengan cara bereaksi terhadap situasi yang berbeda. Tes kepribadian merupakan sebuah metode uji untuk mengukur sifat unik yang ada dalam diri seseorang [1]. Hal ini menjadi salah satu alasan mengapa penting dalam memahami kepribadian seseorang dengan melakukan tes kepribadian lewat tulisan tangan, manfaat yang

diperoleh dari tes kepribadian antara lain dapat menyesuaikan diri dengan lingkungan kerja, kecocokan penempatan pegawai, dan dapat berinteraksi dalam hubungan sosial. Permasalahan yang terjadi adalah tes kepribadian membutuhkan waktu yang lama dan biaya yang tidak murah, Salah satu tes psikologi yang digunakan untuk mengetahui kepribadian adalah *grafo-test* yaitu dengan menganalisis tulisan tangan [2].

Tulisan tangan merupakan hasil dari pikiran sadar dan alam bawah sadar manusia yang menggambarkan atau mencerminkan karakter kepribadian manusia. Terdapat beberapa aspek dalam mengidentifikasi kepribadian seseorang lewat tulisan tangan salah satunya margin [3]. *Margin* merupakan sebagai penanda bagaimana seseorang menulis naskah pada halamannya menunjukkan pendekatan terhadap dunia. *Margin* menggambarkan pikiran alam bawah sadar dalam mengontrol diri dalam kehidupan. *Margin* merupakan jarak yang diambil oleh penulis dari tepi kertas dan dibedakan menjadi margin atas, kanan dan kiri. Dalam prakteknya, *margin* tulisan tangan yang terbentuk ada yang lurus ada yang semakin melebar atau menyempit pada *margin* kiri, kanan dan ada juga yang tidak beraturan. *Margin* Kiri dalam *grafolog* identik dengan masa lalu, sedangkan margin kanan identik dengan masa depan sedangkan *margin* atas menggambarkan diri seseorang dan orang lain [4].

Penelitian untuk mengenali kepribadian melalui aspek *margin* telah dilakukan [5, 6, 7, 8, 9, 10] dimana margin dapat menunjukkan *respon* penulis pada masa lalu dan masa yang akan datang. Selain itu penggunaan margin dalam penentuan kepribadian dijadikan patokan untuk mengukur sifat introvert, ekstrovert, kepemimpinan dan lainnya.

Penelitian sebelumnya menyatakan bahwa *grafo-test* masih menggunakan pengukuran secara manual dengan menggunakan penggaris, hal ini tentu tidak efisien waktu serta keakuratannya tidak dapat dipastikan dalam mengidentifikasi karakter atau kepribadian seseorang

Penelitian sebelumnya menyatakan bahwa *grafo-test* masih menggunakan pengukuran secara manual dengan menggunakan penggaris. Metode ini melibatkan pengukuran berbagai aspek tulisan tangan seperti margin, tinggi huruf, kemiringan, dan jarak antar kata secara manual. Teknik ini memiliki beberapa batasan spesifik seperti pengukuran manual memerlukan waktu yang cukup lama untuk setiap sampel tulisan tangan terutama ketika menganalisis sejumlah data besar, setiap aspek tulisan tangan harus diukur satu per satu dengan penggaris, yang memakan waktu dan tenaga, untuk mendapatkan hasil yang lengkap, analisis manual memerlukan berbagai langkah tambahan seperti pencatatan dan perhitungan, yang semakin memperlambat proses, pengukuran manual rentan terhadap kesalahan manusia, baik dalam membaca skala penggaris maupun dalam menentukan titik awal dan akhir pengukuran, interpretasi hasil pengukuran bisa sangat subjektif, tergantung pada siapa yang melakukan pengukuran, dua *grafolog* mungkin mendapatkan hasil yang berbeda dari sampel tulisan tangan yang sama. Batasan spesifik ini menunjukkan bahwa metode *grafo-test* manual tidak hanya tidak efisien waktu tetapi juga kurang dapat diandalkan dalam mengidentifikasi karakter atau kepribadian seseorang. Dengan perkembangan teknologi, metode otomatis berbasis kecerdasan buatan dan visi komputer dapat menawarkan solusi yang lebih efisien dan akurat untuk analisis tulisan tangan, mengurangi atau menghilangkan banyak dari batasan yang ada dalam metode manual tradisional. Sehingga dilakukan perancangan sistem identifikasi kecenderungan perilaku dengan menggunakan metode *Optical Character Recognition* dan klasifikasi. [6].

Dalam Visi Komputer, tulisan tangan akan diubah menjadi matriks yang berisikan pembangkitan nilai piksel dari tulisan tangn. Deteksi *margin* dilakukan dengan menghitung jumlah piksel dari *starting point* dengan awal kalimat. Namun, penelitian ini hanya melihat margin kiri pada tulisan tangan berupa paragraf dalam lembaran kosong [5].

Penelitian ini dibuat untuk menjadi solusi agar tes kepribadian bisa tetap berjalan dengan formula dasar dan bisa mendapatkan hasil yang lebih cepat. Adapun tujuan khusus dari penelitian ini adalah membangun aplikasi pengenalan tipe kepribadian melalui tulisan tangan

berbasis *Computer Aided Grafology*, memberikan alternatif teknologi dalam bidang ilmu psikologi. Dengan melihat latar belakang tersebut, penelitian ini mengusulkan sebuah teknik visi komputer dalam mendeteksi jenis margin menggunakan *Random Forest*.

II. SIGNIFIKANSI STUDI

Penelitian ini berkontribusi signifikan terhadap pengetahuan yang ada dengan memperkenalkan aplikasi visi komputer dalam analisis grafologi, suatu pendekatan yang belum banyak diimplementasikan secara luas dalam studi sebelumnya. Keunikan penelitian ini terletak pada fokusnya yang spesifik terhadap satu karakteristik tulisan tangan, yaitu margin, yang kemudian diklasifikasikan ke dalam empat kategori: lebar, sempit, melebar, dan menyempit. Sebelumnya, sebagian besar penelitian grafologi menggunakan berbagai karakteristik seperti margin, kemiringan, ukuran tulisan, dan bentuk huruf tanpa fokus pada satu aspek spesifik, yang sering menghasilkan hasil yang kurang akurat. Dalam penelitian ini, metode *random forest* digunakan untuk klasifikasi kepribadian berdasarkan analisis margin, dan hasil uji menunjukkan bahwa metode ini memberikan akurasi hingga 95%. Hal ini menunjukkan bahwa pendekatan baru ini tidak hanya meningkatkan efisiensi waktu tetapi juga keakuratan dalam mendeteksi kepribadian, tanpa memerlukan intervensi ahli psikologi. Penggunaan teknologi visi komputer dan metode *random forest* dalam analisis grafologi merupakan inovasi yang belum pernah diuji secara menyeluruh sebelumnya, menjadikan penelitian ini unik dan membuka jalan bagi aplikasi lebih lanjut dari teknologi canggih dalam bidang grafologi.

A. *Personality*

Personality atau kepribadian menurut psikologi merupakan suatu organisasi yang dinamis dari sistem psikofisik individu yang menentukan tingkah laku dan pemikiran yang khas dari individu tersebut. Kepribadian juga merupakan suatu bidang studi empiris, dan bukan sebagai dasar untuk melakukan penilaian baik atau buruk suatu perilaku. Namun kepribadian juga dalam pengertian populer dapat dikategorikan menjadi dua yaitu: keterampilan atau kecakapan sosial dan kesan yang paling menonjol yang dimiliki seseorang terhadap orang lain [8, 7]. *Personality* dapat dianalisis dengan menggunakan media tulisan tangan. Analisis tulisan tangan sejak tahun 1871 adalah ilmu berdasarkan psikologi dan statistik yang mengkaji ciri - ciri konkrit tulisan tangan untuk mengetahui profil psikologi penulis. Analisis tulisan tangan menjelaskan bahwa penulisan adalah *refleks* otak, sehingga dapat memperlihatkan kepribadian dan kerangka pikiran seseorang. Analisis tulisan tangan telah terbukti dalam berbagai aplikasi, seperti profil kerja.

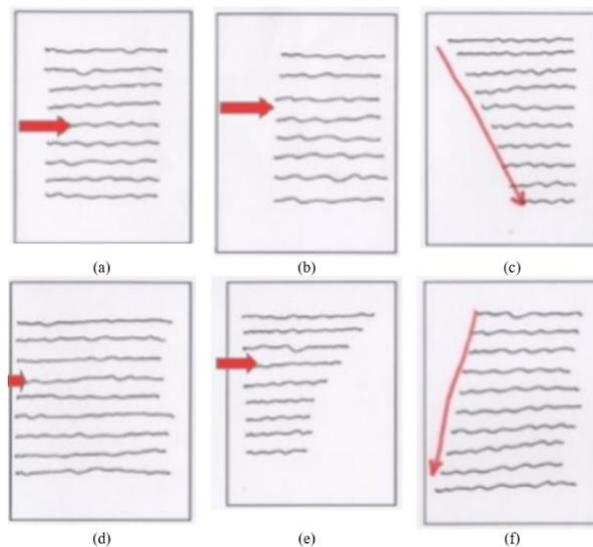
Tubuh, mental, dan emosi seseorang berkontribusi pada proses menulis. Otak manusia bertanggung jawab atas fungsi tubuh yang sadar dan tidak sadar. Otak juga bertanggung jawab atas gaya tulisan tangan seseorang. Grafologi adalah bidang yang mempelajari karakter dan kepribadian seseorang melalui gaya tulisan tangan dan tanda tangan (San, 2016). Ini tidak melihat kualitas tulisan atau isi tulisan (San, 2016) [11].

Teknik-teknik dalam grafologi digunakan untuk mendapatkan gambaran kepribadian seseorang, antara lain tentang pola berfikir, rasa percaya diri, produktifitas, emosi, dan tanda-tanda lainnya termasuk factor-faktor yang dapat menghambat perkembangan seseorang sehingga bisa dilakukan pencegahan sedini mungkin. Selain digunakan untuk diri sendiri tes kepribadian, juga bisa digunakan untuk memahami orang lain. Dengan mengenal kepribadian seseorang dapat lebih menggali potensi dan bakat.

B. *Margin*

Margin adalah sejumlah ruang yang ditinggalkan oleh penulis di sisi kiri atau kanan halaman, atau bagian atas atau bawah halaman. Margin merupakan titik mulai menulis di lembar kertas kosong, dengan mengasumsikan sejumlah ruang pada lembar kertas. Margin dapat mendefinisikan masa lalu dan masa depan, penyesuaian, kecerdasan, kecepatan, dan

kejujuran. Margin memiliki berbagai jenis, yaitu margin atas, bawah, kiri, kanan. Penelitian ini berfokus pada pengukuran margin kiri yang dibentuk. *Margin* kiri adalah tepi bagian kiri atas tempat penulis memulai tulisannya. Ketidaksadaran diri dalam memilih titik awal penulisan menunjukkan kepribadian yang berbeda sesuai kategori jenis margin yang ditampilkan. Ragam jenis margin ditampilkan pada gambar 1. Adapun makna grafologi dari jenis margin ditampilkan pada tabel I. Salah satu penggunaan teknologi dalam mendeteksi kepribadian melalui margin adalah dengan menggunakan teknik visi computer [12, 13].



Gambar 1. Jenis Margin pada Tulisan Tangan (a) Normal (b) Lebar (c) Melebar (d) Tidak ada Margin (e) Sempit (f) Menyempit

TABEL I
MAKNA UKURAN MARGIN YANG TERBENTUK PADA TULISAN TANGAN

Jenis Margin	Indikasi
Normal (15% dari lebar kertas)	Penulis dapat menyikapi masa lalunya dengan wajar, jadi bahan evaluasi, dan jadi penyemangat untuk hidup yang sekarang. Menunjukkan keseimbangan, rasa estetika yang baik, kehidupan normal dan sederhana, teratur, harmonis. Pengendalian diri berlebihan
Lebar (antara 15% - 25% dari lebar kertas)	Dermawan, ekstroverssi, ramah, inisatif, mudah mengambil Keputusan, boros, suka pamer
Melebar	Penulis berusaha meninggalkan masa lalunya terutama pengalaman - pengalaman buruk agar tidak memengaruhi dirinya sekarang dan nanti. Menunjukkan kedermawanan, keberanian, inisiatif, ekstroversi, optimis, antusias Sisi negatifnya: buang - buang uang, waktu dan tenaga tidak sabaran, mudah gugup, impulsive, sembrono
Tidak ada Margin	Urgensi ekonomi, rasa malu, introversi, kurang sentuhan social, kurang selera dan rasa estetika. Ketidak percayaan, Tabungan kehati-hatian. Egois, kejam, kepribadian tidak terlalu ramah
Sempit (10% dari lebar kertas)	Penulis punya kemauan dan usaha untuk meninggalkan masa lalu, upaya bergerak maju untuk masa depan lebih baik, tetapi terkadang masih teringat masa lalu. Hemat, hati-hati, rasa malu, introversi, keterikatan pada keluarga tinggi
Menyempit	Penulis mempunyai masa lalu yang membebani dirinya sampai kini, bercermin pada masa lalu dalam bertindak. Kurang sentuhan sosial, kurangnya selera dan rasa estetika, tidak mudah percaya, egosentris, depresi, kehati - hatian, terlalu ramah. Takut akan masa depan, cenderung hidup hemat

Penelitian Akbar (2020) dengan judul “Perancangan Sistem Identifikasi Kecenderungan Emosional Optimis Dan Pesimis Individu Berbasis Pengolahan Citra Digital” mendapat kesimpulan sistem yang bertujuan untuk mengidentifikasi pola tulisan tangan melalui margin untuk mengetahui kecenderungan emosional optimis dan pesimis individu. Suatu sistem tentu memiliki kekurangan dan kelebihan. Pada sistem ini memiliki kelebihan yaitu dapat mendeteksi pola tulisan tangan dengan akurasi sebesar 90,47%. Sedangkan, kekurangannya adalah sistem masih belum bisa membaca garis dasar tulisan tangan sesuai dengan data sampel yang asli hal ini di pengaruhi pada nilai sudut dan pengambilan data melalui kamera *handphone*.

C. Binary

Penelitian ini mengubah citra rgb menjadi biner untuk mengurangi beban komputasi. Proses ini dilakukan menggunakan *otsu* yang berbasis pada histogram nilai intensitas keabuan citra. Dalam pengolahan citra digital, metode Otsu digunakan mengubah citra menjadi bentuk yang sederhana yang terdiri dari dua nilai. Nilai 1 (putih) sebagai latar depan (objek) dan 0 (hitam) untuk latar belakang (*background*) dengan menggunakan suatu nilai ambang (*threshold*) secara otomatis. Piksel yang memiliki derajat keabuan lebih kecil dari nilai ambang yang ditentukan akan diberikan nilai 0, sementara piksel yang memiliki derajat keabuan yang lebih besar dari batas akan diubah menjadi bernilai 1 [15].

D. Komplemen

Komplemen dalam pengolahan citra merujuk pada operasi yang mengubah nilai intensitas piksel menjadi nilai yang berlawanan. Pada citra biner, piksel dengan nilai 0 diubah menjadi 1 dan sebaliknya. Pada citra grayscale atau citra berwarna, setiap nilai intensitas piksel diubah menjadi nilai maksimum minus nilai intensitas tersebut [16]. Adapun rumus dari *complement* sebagai berikut [17]:

1. Komplemen citra biner :

Setiap piksel dengan nilai 0 diubah menjadi 1, dan sebaliknya.

$$B'(x, y) = 1 - B(x, y)$$

di mana $B(x, y)$ adalah nilai piksel pada koordinat (x, y) dalam citra biner.

2. Komplemen citra grayscale:

Setiap nilai intensitas piksel $I(x, y)$ diubah menjadi nilai maksimum minus nilai intensitas tersebut. Untuk citra dengan intensitas 8-bit, nilai maksimum adalah 255.

$$I'(x, y) = 255 - I(x, y)$$

di mana $I(x, y)$ adalah nilai intensitas pada koordinat (x, y) dalam citra grayscale.

3. Komplemen Citra Berwarna (RGB)

Setiap kanal warna (R, G, B) diubah secara independen dengan menggunakan rumus yang sama seperti pada citra grayscale.

$$R'(x, y) = 255 - R(x, y)$$

$$G'(x, y) = 255 - G(x, y)$$

$$B'(x, y) = 255 - B(x, y)$$

di mana $R(x,y)$, $G(x,y)$ dan $B(x,y)$ adalah nilai intensitas pada kanal merah, hijau, dan biru pada koordinat (x,y) .

E. Morfologi

Inti operasi morfologi melibatkan dua larik piksel. Larik pertama berupa citra yang akan dikenai operasi morfologi, sedangkan larik kedua dinamakan sebagai kernel atau *structuring element* [18, 19]. Operasi morfologi yang sering digunakan dalam adalah [20]:

1. Dilasi

Operasi dilasi biasa dipakai untuk mendapatkan efek pelebaran terhadap piksel yang bernilai 1. Operasi ini dirumuskan seperti berikut:

$$A \oplus B = \{z | [(B)_z \cap A] \subseteq A\}$$

Dalam hal ini,

$$\hat{B} = \{w | w = -b, \text{ untuk } b \in B\}$$

$$(B)_z = \{c | c = a + z, \text{ untuk } a \in A\}$$

$$z = (z1, z2)$$

2. Erosi

Operasi erosi mempunyai efek memperkecil struktur citra. Operasi ini dirumuskan seperti berikut:

$$A \ominus B = \{z | (B)_z \subseteq A\}$$

3. Opening

Operasi opening adalah operasi erosi yang diikuti dengan dilasi dengan menggunakan elemen penstruktur yang sama. Operasi ini berguna untuk menghaluskan kontur objek dan menghilangkan seluruh piksel di area yang terlalu kecil untuk ditempati oleh elemen penstruktur. Dengan kata lain, semua struktur latardepan yang berukuran lebih kecil daripada elemen penstruktur akan tereliminasi oleh erosi dan kemudian penghalusan dilakukan melalui dilasi. Definisi operasi opening seperti berikut:

$$A \circ B = (A \ominus B) \oplus B$$

4. Closing

Operasi closing berguna untuk menghaluskan kontur dan menghilangkan lubang-lubang kecil. Definisinya seperti berikut:

$$A \bullet B = (A \oplus B) \ominus B$$

Jadi, operasi closing dilaksanakan dengan melakukan operasi dilasi terlebih dahulu dan kemudian diikuti dengan operasi erosi.

F. Ekstraksi Fitur Geometri

Ekstraksi Fitur Geometri merupakan cabang matematika yang berhubungan dengan posisi sebuah benda dalam ruang bangun. Penelitian ini mengambil nilai koordinat dari bounding box yang terbentuk untuk dijadikan fitur dalam mengenali jenis margin. Geometri koordinat disebut juga sebagai geometri analitis. Fitru berbasis geometri telah digunakan dalam mendeteksi huruf braille dan memberikan hasil pengenalan yang cukup baik, yaitu sebesar 75% [21].

G. *Random Forest*

Random Forest adalah metode pembelajaran mesin yang kuat dan serbaguna yang digunakan untuk tugas-tugas klasifikasi dan regresi. Metode ini termasuk dalam kategori *ensemble learning*, yang menggabungkan prediksi dari beberapa model untuk meningkatkan akurasi dan stabilitas hasil. Pada intinya, *Random Forest* terdiri dari sejumlah besar pohon keputusan (*decision trees*) yang dilatih pada sub-sampel acak dari dataset asli dan menggunakan subset acak dari fitur-fitur yang tersedia. Teknik ini memanfaatkan dua konsep utama, yaitu bootstrap sampling dan pemilihan fitur acak, untuk menciptakan pohon-pohon yang saling tidak berkorelasi dan mengurangi risiko overfitting dengan menggabungkan prediksi dari banyak pohon. Adapun rumus *random forest* sebagai berikut [22] :

1. *Gini Impurity* (untuk klasifikasi) :

Salah satu cara untuk menentukan split terbaik dalam pohon keputusan adalah dengan meminimalkan Gini impurity. Rumus *Gini impurity* untuk node t adalah:

$$Gini(t) = 1 - \sum_{i=1}^c p_i^2$$

Di mana C adalah jumlah kelas dan p_i adalah proporsi sampel yang termasuk dalam kelas i pada node t .

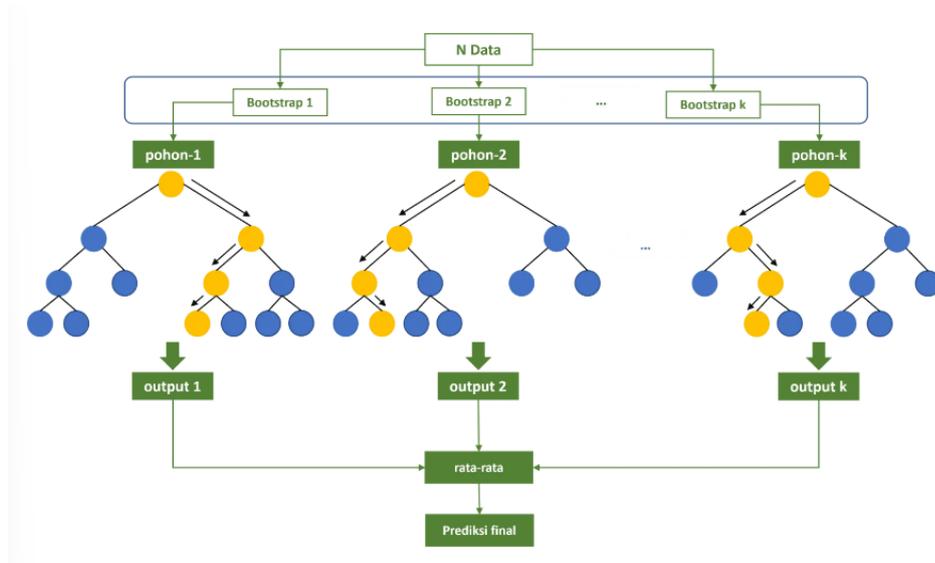
2. *Reduction in Variance* (untuk Regresi):

Untuk regresi, split terbaik sering kali dipilih dengan meminimalkan varian dari nilai-nilai yang dihasilkan. Varian di node t dapat dihitung dengan:

$$Var(t) = \frac{1}{N_t} \sum_{i=1}^{N_t} (y_i - y)^2$$

Di mana N_t adalah jumlah sampel di node t , y_i adalah nilai target untuk sampel i , dan y adalah rata-rata dari nilai target di node tersebut.

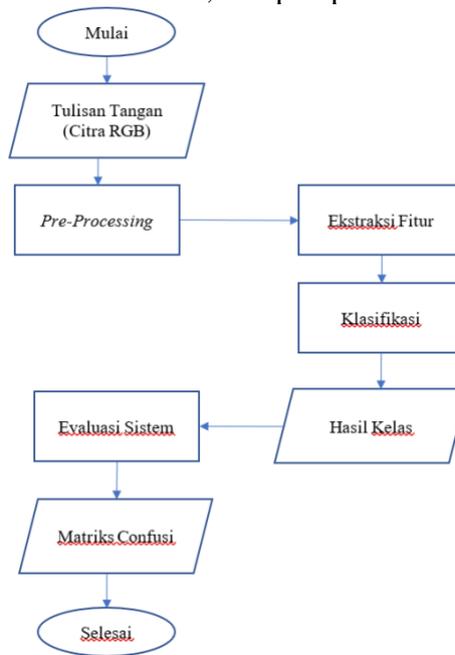
Random forest melakukan klasifikasi dengan membangkitkan beberapa pohon keputusan secara acak. Setiap pohon Keputusan akan melakukan klasifikasi dan menghasilkan kelas prediksi. Dari hasil prediksi tersebut kemudian dilakukan voting (hasil kelas terbanyak) dalam memutuskan memberikan hasil *random forest*. Proses klasifikasi *random forest* ditampilkan pada gambar 2. Model yang dihasilkan *Random Forest* memberikan hasil yang lebih baik dibandingkan dengan penggunaan mesin pembelajaran lain [23, 24].



Gambar 1. Ilustrasi Model *Random Forest*

H. Metode Penelitian

Penelitian dilakukan dengan menggunakan dataset dari *IAM-handwriting dataset* yang telah disesuaikan, yaitu dengan mengambil 100 data tulisan tangan yang dikelompokkan dalam 4 kelas, yaitu lebar, melebar, sempit dan menyempit. Pemilahan data latih dan data uji dilakukan menggunakan *10 k-validation*. Secara umum, tahapan penelitian ditampilkan pada gambar 1.



Gambar 2. Tahapan Penelitian

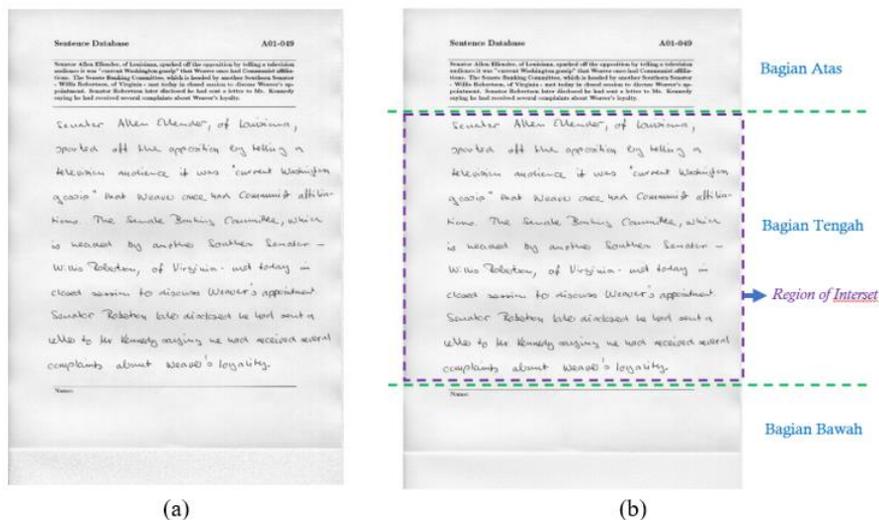
Penelitian dimulai dengan menganalisis kondisi citra (*pre-processing*) bertujuan untuk mengubah citra *rgb* menjadi *biner* dan memperbaiki kualitas citra agar proses analisis berjalan dengan baik. Selain itu, tujuan dari langkah *pre-processing* adalah untuk mendapatkan objek tulisan tangan agar dapat diekstraksi ciri untuk kemudian dilakukan pengenalan. Ekstraksi fitur yang digunakan dalam penelitian ini berbasis geometri, yaitu mengacu pada posisi objek secara spasial. Setelah didapat fitur dari objek tulisan tangan, kemudian dilakukan klasifikasi menggunakan *Random Forest* (RF) untuk mendapatkan kelas prediksi. Untuk mengukur ketepatan hasil prediksi digunakan alat uji akurasi menggunakan matrik konfusi.

III. HASIL DAN PEMBAHASAN

Margin dapat dikenali dari jarak antara tepi kertas hingga awal goresan tulisan tangan. Hal ini memudahkan pendeteksian margin dengan menggunakan pendekatan visual berbasis geometri. Untuk itu penelitian ini menggunakan perhitungan pada citra di bidang spasial. Berikut hasil dan pembahasan pada tiap tahapan penelitian yang dilakukan:

A. Pre-processing

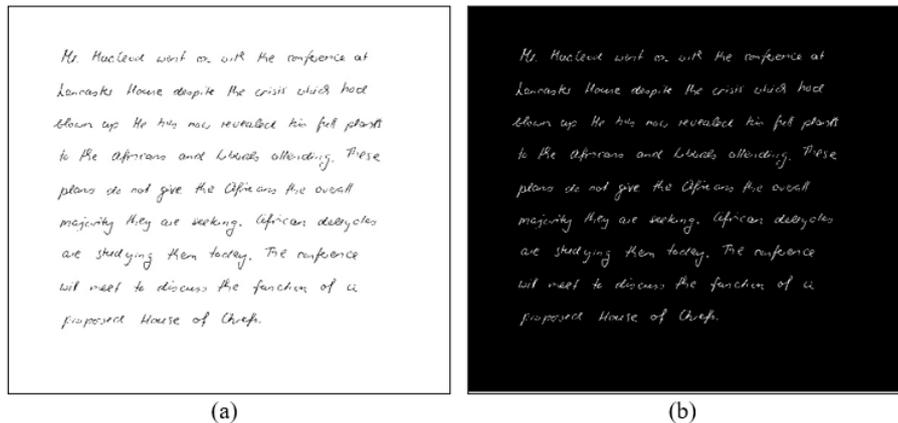
Citra pada dataset IAM handwriting berupa citra rgb dengan ukuran 2479x3542 piksel. Citra ini terdiri dari bagian atas, tengah dan bawah. Bagian atas berisikan identitas lembar kerja dan teks kalimat yang akan disalin. Bagian tengah berisi tulisan tangan dan bagian bawah berisikan tulisan Note. Dataset IAM ditampilkan pada Gambar 4 (a). Penelitian ini mengambil 25 citra yang mempunyai keragaman margin berbeda, sehingga didapat 100 citra secara keseluruhan. Kemudian untuk mengurangi kompleksitas komputasi, dilakukan pemotongan pada bagian tengah yang hanya berisi tulisan tangan seperti pada gambar 2 (b).



Gambar 3. Pengambilan dataset Citra (a) citra asli (b) area pemotongan dataset

Selain memotong citra, pada tahap *pre-processing* dilakukan perubahan citra rgb kedalam *biner*. Hal ini dilakukan untuk menghemat komputasi. Hasil dari konversi rgb ke biner ditampilkan pada gambar 5 (a). Citra biner hanya terdiri dari 0 dan 1, dimana matlab mengenali nilai 1 (putih) sebagai objek sedangkan 0 (hitam) sebagai *background*. Operasi perubahan citra rgb menjadi biner dilakukan menggunakan metode otsu.

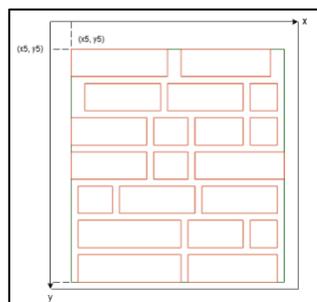
Pada citra dataset yang digunakan tulisan dilakukan menggunakan alat tulis berwarna gelap pada kertas putih, sehingga ketika dilakukan perubahan ke biner, objek akan berwarna hitam dan background berwarna putih. Hal ini menyebabkan kondisi terbalik pada analisis citra, sehingga diperlukan proses pembalikan dengan melakukan komplemen. Hasil komplemen citra ditampilkan pada gambar 5 (b).



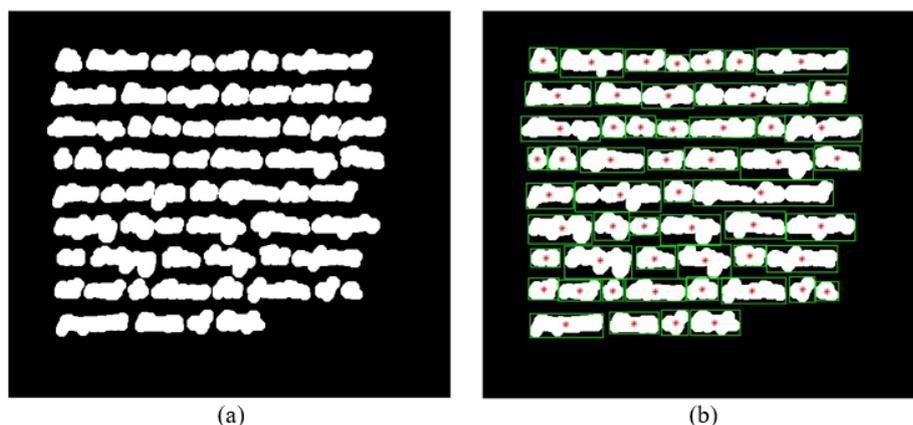
Gambar 4. Hasil Pre-Processing (a) Hasil Biner (b) Hasil Komplemen

B. Ekstrasi Fitur

Setelah dilakukan komplemen, objek yang berwarna putih merupakan representasi dari huruf-huruf yang ditulis tangan. Margin merupakan jarak antara tepi kertas dengan awal permulaan penulisan. Tahapan penelitian ini fitur yang digunakan adalah margin kiri. Fitur yang digunakan diilustrasikan pada gambar 4. Fitur akan diambil dari nilai posisi x dan y pembentuk *bounding box*. Untuk itu sebelum mengekstraksi fitur berbasis geometris dilakukan penggabungan antar huruf dalam kata atau dalam satu baris kalimat. Untuk itu dilakukan proses morfologi dilasi. Pada penelitian ini dilasi dilakukan menggunakan structuring element berupa disk dengan ukuran besar. Hasil dilasi ditunjukkan pada gambar 7 (a). setelah itu dilakukan ekstraksi fitur dengan menghitung centroid yang merupakan titik pusat objek. Dari centroid ini kemudian dibangkitkan *bounding box* yang mencakup satu objek dalam bentuk persegi empat, seperti pada gambar 7 (b).



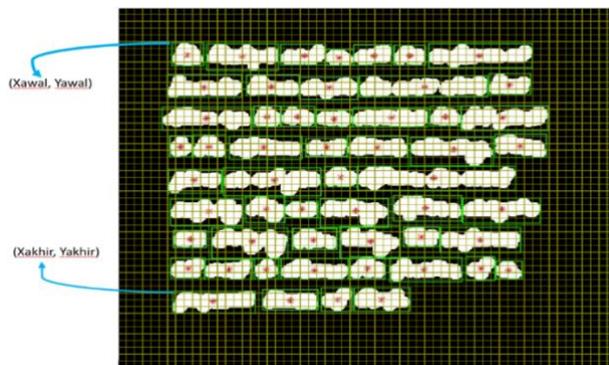
Gambar 5. Ilustrasi fitur berbasis geometris yang akan dilakukan



Gambar 6. Ekstraksi Fitur berbasis Geometri (a) hasil dilasi (b) penentuan *centroid* dan *bounding box*

Bounding box memberikan efek kotak pembatas yang mengelilingi objek (dalam hal ini tulisan tangan) yang mewakili koordinat atas-kiri (*top-left*), koordinat bawah-kanan (*bottom-right*) atau dapat mewakili oleh pusat, lebar, tinggi kotak.

Setelah didapat *bounding box* tiap objek, ekstraksi fitur melakukan pemilihan untuk mengambil nilai (x,y) pada baris awal dan baris akhir saja yang akan disimpan sebagai *feature-set*, seperti ditampilkan pada gambar 6. *Feature-set* inilah yang digunakan pada tahapan klasifikasi untuk menentukan jenis margin.



Gambar 7. Ilustrasi pengambilan *Feature-set*

C. *Klasifikasi*

Proses klasifikasi dilakukan menggunakan *random forest*, dimana seluruh citra dilakukan ekstraksi fitur dan diberi label. *Random forest* dilakukan dengan membagi data latih dan data uji menggunakan *10 k-validation*. Pembagian kelas dilakukan berdasarkan ciri geometris yang didapat seperti pada gambar 7. Pada penelitian ini citra dibagi dalam 4 kategori jenis margin, yaitu: sempit, menyempit, lebar dan melebar. *Random forest* dibangun dengan 3 *random features* hasil *random fores* memberikan ketepatan prediksi sebanyak 95 data dari 100 data yang digunakan. Hasil akurasi dihitung dengan membandingkan jumlah data yang prediksi kelasnya sesuai dengan kelas aktual (benar) terhadap jumlah seluruh data.

$$akurasi = \frac{Total\ nilai\ prediksi\ benar}{Total\ jumlah\ data} \times 100\%$$

$$akurasi = \frac{95}{100} \times 100\% = 95\%$$

Secara keseluruhan didapat akurasi sebesar 95%. Hasil akurasi pada tiap k (10 k-validation) ditampilkan pada tabel II.

TABEL III
HASIL AKURASI TIAP K

Nilai k (k-validation)	Akurasi
1	9
2	10
3	10
4	10
5	9
6	10
7	9
8	10
9	8
10	10
TOTAL	95

D. Evaluasi

Tahapan penelitian selanjutnya adalah mengevaluasi sistem dengan cara mengukur ketepatan hasil prediksi yang dikerjakan pada tahap klasifikasi menggunakan *random forest*. Dalam mengukur kinerja sistem digunakan matriks konfusi 4 kelas yang dapat memberikan gambaran keakuratan hasil klasifikasi di tiap kelas seperti ditampilkan pada tabel III.

TABEL III
MARIKS KONFUSI

Prediksi \ Kelas	Lebar	Melebar	Menyempit	Sempit
Lebar	24	0	0	1
Melebar	1	25	0	0
Menyempit	0	0	24	0
Sempit	2	0	1	22

Matriks konfusi adalah matrik yang menggambarkan hasil prediksi terhadap masing-masing kelas. Dari matriks konfusi dapat diketahui komponen akurasi tiap kelas seperti pada tabel 4.

TABEL IVV
PERHITUNGAN KOMPONEN MARIK KONFUSI

Komponen \ Kelas	1	2	3	4	Rata-rata
TruePositif (TP)	24	25	24	22	23.75
False Positif (FP)	3	0	1	1	1.25
False Negatif (FN)	1	1	0	3	1.25
True Negatif (TN)	72	74	75	74	73.75

Dari matrik konfusi tabel IV, dapat dihitung nilai precision dan recall dari pengujian yang telah dilakukan. Precision menunjukkan tingkat keakuratan dari model klasifikasi pada tiap kelas yang terdeteksi benar pada masing-masing kelas. Perhitungan Nilai Precision dilakukan menggunakan rumus berikut:

$$Precision = \frac{Jumlah\ True\ Positive}{Jumlah\ True\ Positive + Jumlah\ False\ Positive} \times 100\%$$

Recall, disebut juga sebagai Sensitivitas, menunjukkan tingkat keakuratan dari model klasifikasi. Perhitungan Nilai Recall atau Sensitivitas dilakukan menggunakan rumus berikut:

$$Recall = Sensitivitas = \frac{Jumlah\ True\ Positive}{Jumlah\ True\ Positive + Jumlah\ False\ Negatif} \times 100\%$$

Selain itu, dari matriks konfusi ini juga dapat diketahui nilai spesifisitas. Nilai Spesifisitas menunjukkan ketidak adaan hasil prediksi pada kelas tertentu. Spesifisitas juga dikenal sebagai rerata True Negatif, dihitung menggunakan rumus:

$$Spesifisitas = \frac{Jumlah\ True\ Negatif}{Jumlah\ True\ Negatif + Jumlah\ False\ Positive} \times 100\%$$

Untuk menghitung hasil evaluasi sistem (performance) dari 4 kelas tipe margin pada penelitian ini, dihitung dahulu pada masing-masing kelas kemudian dirata-rata. Hasil performance sistem ditampilkan pada tabel V.

TABEL V
MAKNA UKURAN MARGIN YANG TERBENTUK PADA TULISAN TANGAN

Komponen Matrik Konfusi	Kelas	1	2	3	4	Rata-rata
Precision = $TP/(TP + FP)$		0.89	1	0.96	0.96	0.951353
Recall = Sensitivitas = $TP/(TP+FN)$		0.96	0.96	1	0.88	0.950385
Spesivitas = $(TN)/(TN + FP)$		0.96	1	0.99	0.99	0.983377

IV. KESIMPULAN

Penelitian ini mengusulkan ekstraksi fitur berbasis geometri dan klasifikasi *random forest* untuk mendeteksi margin. Penelitian ini dilakukan dengan menggunakan 100 dataset dari *IAM-handwriting dataset* dengan membagi data latih dan data uji menggunakan *10 k-validation*. Penelitian ini mengenali margin dalam 4 kelas, yaitu lebar, melebar, menyempit dan sempit. Secara umum, penelitian ini memberikan akurasi sebesar 95% dengan nilai *Precision* = 95,135%, *Recall* dan *Sensitivitas* sebesar 95% dan *Spesivitas* sebesar 98,34%. Hal ini menunjukkan bahwa metode ekstraksi fitur berbasis geometri dapat mengenali jenis margin dengan baik dan *random forest* mampu mengklasifikasikan jenis margin dengan akurat. Meskipun hasil penelitian ini menunjukkan performa yang sangat baik, terdapat beberapa batasan yang perlu dipertimbangkan untuk memberikan gambaran yang lebih seimbang dan kredibel seperti penggunaan dataset yang lebih besar dan lebih beragam bisa membantu meningkatkan keakuratan dan reliabilitas model, faktor lain yang dapat mempengaruhi akurasi metode adalah variasi dalam tulisan tangan yang disebabkan oleh kondisi psikologis dan fisik penulis, jenis alat tulis, dan permukaan tempat menulis. Penelitian ini belum mempertimbangkan variabel-variabel tersebut, yang bisa memberikan pengaruh signifikan pada hasil analisis, dibandingkan dengan studi sebelumnya yang menggunakan berbagai karakteristik tulisan tangan tanpa fokus khusus, penelitian ini menunjukkan peningkatan akurasi dan efisiensi dengan fokus pada satu karakteristik spesifik, yaitu margin. Metode *random forest* yang diterapkan dalam penelitian ini memberikan hasil yang lebih akurat dibandingkan dengan metode yang digunakan dalam penelitian sebelumnya yang seringkali memberikan hasil yang kurang andal. Berdasarkan temuan ini, ada beberapa saran untuk penelitian selanjutnya seperti : menggunakan dataset yang lebih besar dan lebih beragam, memasukkan variabel tambahan seperti kondisi psikologis dan fisik penulis, jenis alat tulis, dan permukaan tempat menulis, menjelajahi metode klasifikasi lain seperti deep learning yang mungkin memberikan performa lebih baik dalam analisis yang lebih kompleks. Mengembangkan dan menguji metode ekstraksi fitur yang lebih canggih yang bisa mengidentifikasi lebih banyak karakteristik tulisan tangan secara otomatis dan lebih akurat.

REFERENSI

- [1] W. H. P. L. B. M. D. D. J. J. H. M. H. C. J. .. & R. B. Bleidorn, "The policy relevance of personality traits.," *American Psychologist*, , p. 74(9), 2019.
- [2] K. A. & W. R. Safitri, "Aplikasi Pengenalan Pola Tulisan Tangan Menggunakan Metode Support Vector Machine," *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, pp. 201-206, 2020.

- [3] S. G. N. K. & K. M. Singh, "Feature extraction and classification techniques for handwritten Devanagari text recognition: a survey," *Multimedia Tools and Applications*, 82(1), pp. 747-775, 2023.
- [4] S. K. T. K. P. H. M. & A. S. A. Yadav, "A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions.," *Knowledge-Based Systems*, 223, 106970., 2021.
- [5] M. G. Nugrapratama, "Pendeteksi Kepribadian Berdasarkan Pengenalan Tulisan Tangan Menggunakan Support Vector Machine," *Doctoral dissertation, Universitas Komputer Indonesia*, 2020.
- [6] T. B. P. H. F. T. S. & A. P. Ekasanti, "Perancangan Sistem Identifikasi Kecenderungan Perilaku Self And Confident Dan Inferiority Dengan Pengolahan Citra Digital Menggunakan Metode Optical Character Recognition.," dalam *e-Proceeding of Engineering*, Bandung, 2020.
- [7] P. D. D. & S. M. Conia, "Gambaran Pemahaman Teori Psikoanalisis Dan Implikasinya Dalam Pendidikan Pada Mata Kuliah Karakteristik Dan Kompetensi Usia Dewasa Pada Mahasiswa Bimbingan Dan Konseling Fkip Untirta.," *In National Conference on Applied Business, Education, & Technology (NCABET)*, pp. 547-555, 2021.
- [8] A. Magfiroh, "Pengenalan Kepribadian Seseorang Melalui Bentuk Tulisan Tangan Menggunakan Metode Radial Basis Function Neural Network (RBFNN)," *Zeta-Math Journal*, 7(1), pp. 34-41, 2022.
- [9] Mariana, "Handwriting & Graphology," 20 June 2022. [Online]. Available: <https://handwritinggraphology.com/graphology-analysis-margins-page/>.
- [10] S. H. D. E. C. I. R. & R. F. Fatimah, "Personality features identification from handwriting using convolutional neural networks.," *International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, pp. 119-124, 2019.
- [11] D. S. Prasetyono, *Bedah Lengkap Grafologi*, VI penyunt., Yogyakarta: DIVA Press, 2012.
- [12] W. T. H. & U. F. Wijaya, "Metode Grid-Double Block Untuk Deteksi Margin Kiri Tulisan Tangan Pada Aplikasi Grafologi.," *Jurnal Teknologi Informasi dan Ilmu Komputer*, p. 69, 2018.
- [13] U. & R. N. Rosyidah, "Analisis Kepribadian Melalui Tulisan Tangan Menggunakan Metode Support Vector Machine," *Journal of Informatics and Computer Science (JINACS)*, 2019.
- [14] Admin, "Handwriting & Graphology," *Handwriting & Graphology Analysis*, 20 June 2016. [Online]. Available: <https://www.handwriting-graphology.com/graphology-analysis-margins-page/>. [Diakses 10 February 2021].
- [15] I. D. & K. A. Kurniawati, "Implementasi Algoritma Canny Dalam Pengenalan Wajah Menggunakan Antarmuka GUI Matlab," *ResearchGate, Surabaya.*, 2017.
- [16] D. P. & G. A. B. Adi, "Deteksi Emosi Wicara pada Media On-Demand menggunakan SVM dan LSTM.," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 2020.
- [17] R. L. & H. M. Hasanah, "Deteksi Lesi Acne Vulgaris pada Citra Jerawat Wajah Menggunakan Metode K-Means Clustering.," *Indonesian Journal on Software Engineering (IJSE)*, 2022.
- [18] A. Kadir dan A. Susanto, *Teori dan Aplikasi PengolahanCitra*, Yogyakarta, Indonesia: Andi Publisher, 2013.

- [19] R. C. Gonzalez dan R. E. Woods, *Digital Image Processing*, 2 penyunt., New Jersey: Prentice Hall, 2002.
- [20] A. Susanto, “enerapan operasi morfologi matematika citra digital untuk ekstraksi area plat nomor kendaraan bermotor.,” *Pseudocode*, pp. 49-57, 2019.
- [21] S. S. A. & E. E. Agustin, “Pengenalan Huruf Braille menggunakan Radially Average Power Spectrum dan Geometri.,” *Jurnal Inovtek Polbeng Seri Informatika*, 8(1), pp. 25-36, 2023.
- [22] Tedyyana, Agus & Ghazali, Osman & Purbo, Onno. (2024). Model Design of Intrusion Detection System on Web Server Using Machine Learning Based. 10.4108/eai.21-9-2023.2342879..
- [23] M. A. M. V. F. & A. E. E. Amrustian, “Studi Komparasi Metode Machine Learning untuk Klasifikasi Citra Huruf Vokal Hiragana.,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, pp. 905-912, 2021.
- [24] Y. W. S. P. & M. S. Sitorus, “Analisis Deteksi Malware Android menggunakan metode Support Vector Machine & Random Forest.,” *eProceedings of Engineering*, 2021.