

BAB 2

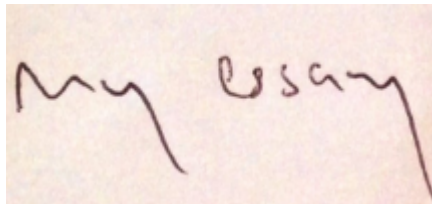
LANDASAN TEORI

2.1 Citra

Citra adalah sebuah istilah untuk gambar dalam multimedia yang berperan sebagai sarana komunikasi visual. Citra memiliki karakteristik yang tidak dimiliki oleh teks, yaitu kaya dengan informasi [14]. Citra juga dianggap sebagai representasi data yang memiliki ruang *layout* dan intensitas warna [15].

Secara harfiah, citra berarti gambar pada bidang dwimatra atau bidang dua dimensi. Ditinjau dari sudut pandang matematis, citra merupakan sebuah fungsi kontinu dari intensitas cahaya pada bidang dwimatra. Sumber cahaya menerangi objek, lalu objek memantulkan kembali sebagian dari berkas cahaya tersebut. Pantulan cahaya ini ditangkap oleh alat-alat optik, misalnya mata pada manusia, kamera, pemindai (*scanner*), dan sebagainya [14]. Citra umumnya terbagi menjadi 3 jenis yaitu citra warna, citra *Grayscale* dan citra biner.

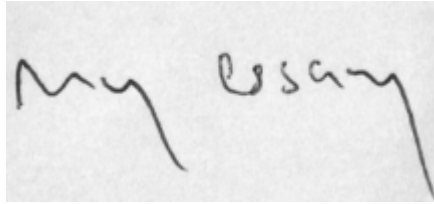
a. Citra warna



Gambar 2-1 Citra Warna

Citra warna adalah sebuah citra yang memiliki satu atau lebih jenis warna yang menentukan intensitas warna pada suatu *piksel* [15]. Contoh di atas adalah citra 8 bit dimana citra 8 bit ini memiliki kriteria di setiap *piksel* dari citra warna diwakili oleh 8 bit, jumlah warna maksimum 256 warna [16].

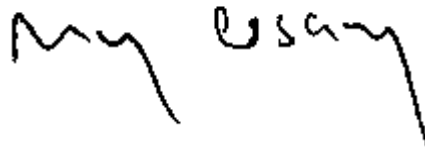
b. Citra *Grayscale*



Gambar 2-2 Citra *Grayscale*

Citra *Grayscale* adalah citra yang hanya memiliki salah satu nilai kanal dari sebuah warna pada setiap *piksel*-nya [16], artinya nilai dari Red = Green = Blue yang digunakan untuk menunjukkan intensitas warna.

c. Citra biner



Gambar 2-3 Citra Biner

Citra biner adalah citra yang hanya memiliki dua intensitas warna yaitu hitam dan putih (1 dan 0) [16]. Citra biner hanya membutuhkan 1 bit memori untuk setiap *piksel*-nya.

2.1.1 Pengolahan Citra Digital

Pengolahan Citra digital adalah ilmu yang mempelajari tentang teknik – teknik untuk mengolah citra[17]. Foto adalah salah satu contoh sebuah citra 2 dimensi yang dapat diolah. Setiap foto dalam bentuk digital dapat diolah dengan menggunakan perangkat lunak tertentu.

2.1.2 Thresholding

Thresholding adalah salah satu metode segmentasi citra dimana prosesnya didasarkan pada perbedaan derajat keabuan citra. Metode ini memisahkan suatu objek

dalam citra dengan latar belakangnya. Bagian citra yang cenderung gelap akan dibuat menjadi intensitas tergelap (0) dan bagian citra yang cenderung terang akan dibuat menjadi intensitas tercerah (1). Keluaran dari proses *thresholding* ini adalah citra biner atau citra hitam putih yang hanya memiliki intensitas 0 dan 1 sehingga akan diketahui mana bagian objek dan mana yang bagian latar belakangnya.

$$g(x,y) = \begin{cases} 1, & \text{jika } f(x,y) \geq T \\ 0, & \text{jika } f(x,y) < T \end{cases} \quad (01)$$

Dimana $g(x,y)$ adalah dari citra biner, $f(x,y)$ adalah nilai piksel pada citra *grayscale* dan T adalah nilai thresholdnya.

Namun proses *thresholding* dengan cara seperti di atas memiliki kekurangan, yaitu nilai ambang batasnya harus ditetapkan secara manual dan hasilnya belum tentu optimal untuk semua citra. Untuk mengatasi masalah ini maka proses *thresholding* menggunakan metode otsu yang dapat menentukan nilai ambang secara otomatis pada setiap citra yang diolah. Metode otsu bekerja dengan cara membagi histogram citra *grayscale* menjadi dua daerah yang berbeda secara otomatis tanpa ada bantuan dari pengguna untuk menetapkan nilai ambang [18]. Nilai ambang batas yang akan dicari dari citra *grayscale* dinyatakan dengan k yang bernilai antara 1 sampai dengan L atau 255. Probabilitas setiap piksel pada level ke i dapat dinyatakan dengan persamaan berikut :

$$p_i = n_i / N \quad (02)$$

Dengan n_i menyatakan jumlah piksel pada level ke i dan N menyatakan total jumlah piksel pada citra. Nilai jumlah kumulatif, rerata kumulatif, dan rerata intensitas global berturut-turut dapat dinyatakan dengan rumus sebagai berikut.

$$\omega(k) = \sum_{i=1}^k p_i \quad (03)$$

$$\mu(k) = \sum_{i=1}^k i \cdot p_i \quad (04)$$

$$\mu_T = \mu(L) = \sum_{i=1}^L i \cdot p_i \quad (05)$$

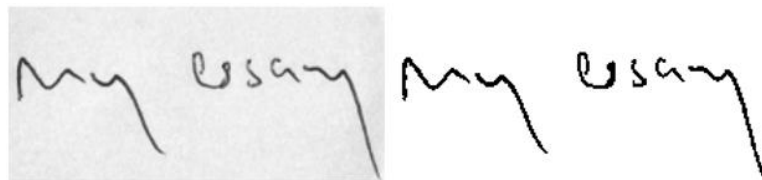
Dan nilai ambang k dapat ditentukan dengan menggunakan persamaan:

$$\sigma_B^2(K^*) = \max_{1 \leq k < L} \sigma_B^2(k) \quad (06)$$

dengan

$$\sigma_B^2(k) = \frac{[\mu_T \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (07)$$

Contoh hasil dari proses *thresholding* dengan metode otsu dapat dilihat pada Gambar 2-5 Proses Metode K Nearest[21]/Gambar 2-4 Citra Grayscale Dan Citra Biner.



Gambar 2-4 Citra Grayscale Dan Citra Biner

2.1.3 Ekstraksi Fitur

Ekstraksi fitur adalah cara mendapatkan sebuah informasi dalam suatu pola untuk membantu tugas klasifikasi menjadi lebih mudah [19]. Dalam *pattern recognition* dan *image processing*, ekstraksi fitur adalah bentuk khusus pengurangan dimensi. Tujuan utama ekstraksi fitur adalah untuk mendapatkan informasi yang paling

relevan dari data asli dan merepresentasikan informasi tersebut dalam ruang dimensi yang lebih kecil. [19].

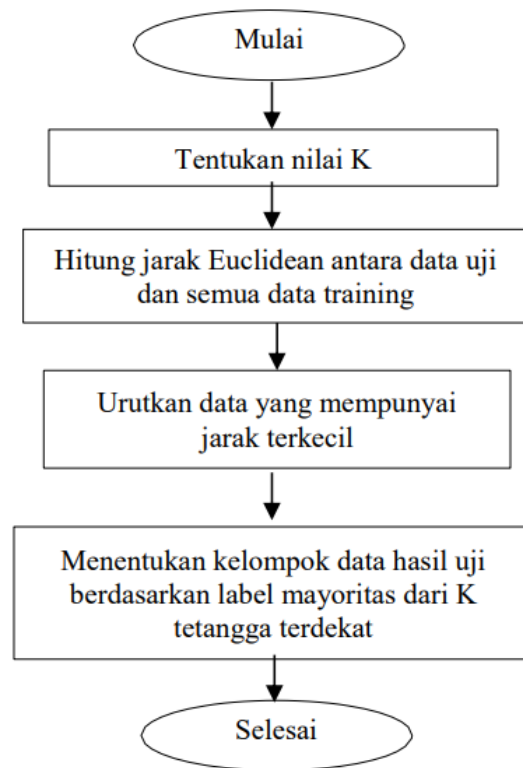
2.2 Dataset

Dataset adalah kumpulan dari sebuah data. Dalam kasus data tabular, satu dataset adalah satu atau lebih tabel database, di mana setiap kolom tabel mewakili variabel tertentu, dan setiap baris sesuai dengan catatan tertentu dari set data yang dimaksud. Dataset yang digunakan dalam penelitian ini adalah IAM Handwriting Top50 yang berjumlah 1126 buah gambar tulisan tangan.

2.3 K-Nearest Neighbor (KNN)

K nearest Neighbor atau KNN adalah sebuah algoritma *supervised learning* yang mengklasifikasikan berdasarkan tetangga terdekat [20]. KNN termasuk kelompok algoritma *instance-based learning*. Algoritma ini juga merupakan salah satu teknik *lazy learning*. KNN dilakukan dengan mencari kelompok k objek dalam data latih yang paling dekat (mirip) dengan objek pada data baru atau data testing. Sebelum melakukan perhitungan, algoritma KNN ini harus menentukan data latih dan data uji[21]. Kemudian akan dilakukan proses perhitungan untuk mencari jarak menggunakan Euclidean. Setelah itu, akan dilakukan tahapan perhitungan dengan metode KNN.

Algoritma *K-Nearest Neighbor* melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut[21]. Mirip dengan teknik *clustering*, yaitu mengelompokkan suatu data baru berdasarkan jarak data baru itu ke beberapa data/tetangga terdekat. Pertama sebelum mencari jarak data ke tetangga adalah menentukan nilai K tetangga (*neighbor*), lalu, untuk mendefinisikan jarak antara dua titik yaitu titik pada data latih dan titik pada data testing, maka digunakan rumus *Euclidean Distances*.



Gambar 2-5 Proses Metode K Nearest[21]

2.3.1 Elbow Method

Elbow method adalah metode yang sering dipakai untuk menentukan jumlah *cluster* yang akan digunakan. Metode ini akan membantu mencari nilai k yang terbaik untuk digunakan[22]. *Elbow method* ini bekerja dengan melihat persentase hasil perbandingan antara jumlah k yang akan membentuk siku pada suatu titik[23]. *Elbow method* ini memberi ide dengan cara memilih nilai k dan kemudian menambahkan nilai k tersebut untuk dijadikan sebuah model data dalam penentuan yang terbaik. Hasil dari persentase yang berbeda pada setiap k dapat ditunjukkan dengan menggunakan sebuah grafik sebagai sumber informasinya.

Algoritma Metode Elbow dalam menentukan nilai K :

1. Mulai
2. Inisialisasi awal nilai K
3. Naikkan nilai K
4. Hitung hasil sum of square error dari tiap nilai K

5. Melihat hasil sum of square error dari nilai K yang turun secara drastis
6. Tetapkan nilai K yang berbentuk siku
7. Selesai

2.3.2 Euclidean Distance

Euclidean Distance adalah cara perhitungan jarak dari 2 buah titik dalam *Euclidean Space*[24]. *Euclidean Distance* biasa diterapkan untuk membantu proses klasifikasi pada *data mining*. Salah satu algoritma yang sering menggunakan *Euclidean Distance* adalah KNN. *Euclidean Distance* digunakan pada penelitian ini karena menurut penelitian Chomboon dkk, pengukuran jarak yang menggunakan *Euclidean Distance* menghasilkan nilai akurasi yang tinggi [25].

Persamaan *Euclidean Distance*:

$$d(a, b) = \sum_{i=0}^n (X_i - Y_i)^2 \quad (08)$$

Keterangan:

d (a,b) : jarak Euclidean

x : data 1

y : data 2

i : fitur ke –

n : jumlah fitur

2.4 Huruf Latin

Alfabet Latin atau Abjad Rumi (dari kata Romawi) adalah alfabet yang pertama kalinya dipakai oleh orang Romawi untuk menuliskan bahasa Latin kira-kira sejak abad ke-7 Sebelum Masehi. Mereka belajar menulis dari orang-orang Etruria, sedangkan orang Etruria belajar dari orang Yunani. Alfabet Etruska merupakan adaptasi dari alfabet Yunani. Menurut hipotesis, semua aksara alfabetis tersebut berasal dari abjad Fenisia, dan abjad Fenisia berasal dari hieroglif Mesir [26].Alphabet

Menurut kamus Oxford alfabet adalah abjad atau huruf yang dimulai dari a sampai z yang biasa kita gunakan untuk menulis sebuah kata atau kalimat. Tabel 2-1 Alphabet di bawah merupakan kumpulan alfabet.

Tabel 2-1 Alphabet

Aa	Bb	Cc	Dd	Ee
Ff	Gg	Hh	Ii	Jj
Kk	Ll	Mm	Nn	Oo
Pp	Qq	Rr	Ss	Tt
Uu	Vv	Ww	Xx	Yy
Zz				

2.5 Pengertian Tulisan Tangan

Tulisan tangan adalah sebuah hasil atau sebuah cara menulis dengan tangan (bukan diketik) oleh seorang individu. Beberapa penelitian menunjukkan bahwa tulisan tangan setiap orang adalah unik dan dapat dikenali dengan perangkat lunak khusus. Kemampuan komputer untuk mengenali tulisan tangan dikenal dengan istilah pengenalan tulisan tangan. Pengenalan Tulisan Tangan.

2.5.1 Pengenalan Tulisan Tangan

Pengenalan tulisan tangan (*Handwriting Recognition*) adalah kemampuan komputer untuk menerima dan menafsirkan *input* tulisan tangan yang dapat dimengerti dari sumber seperti dokumen kertas, foto, layar sentuh dan perangkat lainnya. Gambar dari teks tertulis dapat digunakan secara luring dari selebar kertas oleh pemindai optik (rekognisi karakter optik). Selain itu, gerakan ujung pena dapat dimengerti secara daring, misalnya dengan menggunakan permukaan layar komputer berbasis pena. Pengenalan tulisan tangan memerlukan rekognisi karakter optik. Namun, sistem pengenalan tulisan tangan yang lengkap juga menangani format, melakukan segmentasi yang tepat ke dalam karakter dan menemukan kata yang paling masuk akal[27].

2.6 Kepribadian

Kepribadian adalah kombinasi dari berbagai macam karakteristik dan kualitas seseorang[28]. Sebuah karakteristik yang dinamis dan terorganisir yang berbeda beda yang dimiliki semua orang yang mempengaruhi motivasi, pengambilan keputusan dan perilaku seseorang. Kepribadian dipengaruhi oleh pertumbuhan, cara pandang terhadap suatu nilai dan atribut, hubungan dengan suatu komunitas, kebiasaan dan keterampilan[28]. Sebuah keputusan dan perilaku seseorang sangat dipengaruhi oleh kepribadiannya[28].

Kepribadian memiliki beberapa definisi menurut ahli. Banyak ahli yang merumuskan definisi kepribadian berdasarkan paradigma yang mereka yakini dan fokus analisis yang mereka kembangkan. Berikut ini dikemukakan beberapa ahli yang definisinya dapat dipakai acuan dalam mempelajari kepribadian :

1. Gordon W. W Allport

Menurut Allport kepribadian didefinisikan sebagai sebuah organisasi dinamis dalam individu sebagai sistem psikofisis yang menentukan caranya yang khas dalam menyesuaikan diri terhadap lingkungan.

2. Krech dan Crutchfield

Menurut David Krech DAN Richard S. Crutchfield dalam bukunya yang berjudul *Elements of Psychology* (1969) merumuskan sebuah kepribadian sebagai sebuah integrasi dari semua karakteristik individu ke dalam suatu kesatuan yang unik yang menentukan, dan yang dimodifikasi oleh usaha usahanya dalam menyesuaikan diri terhadap lingkungan yang berubah terus menerus.

3. Adolf Heuken, S.J. dkk.

Adolf Heuken S.J. dkk. dalam bukunya yang berjudul *Tantangan Membina Kepribadian* menyatakan bahwa kepribadian adalah pola menyeluruh semua kemampuan, perbuatan serta kebiasaan seseorang, baik yang jasmani, mental, rohani, emosional maupun yang sosial. Semuanya ini telah ditata-nya dalam caranya yang khas di bawah beraneka pengaruh dari luar. Pola ini terwujud dalam tingkah lakunya, dalam usahanya menjadi manusia sebagaimana dikehendakinya. Berdasarkan definisi para

ahli diatas dapat disimpulkan bahwa kepribadian adalah kesatuan yang kompleks, yang terdiri dari aspek psikis, seperti : inteligensi, sifat, sikap, minat, cita-cita, dst. serta aspek fisik, seperti : bentuk tubuh, kesehatan jasmani yang berinteraksi dengan lingkungannya yang mengalami perubahan secara terus-menerus dan terwujudlah pola tingkah laku yang khas atau unik yang bersifat dinamis.

2.6.1 Usaha-usaha Mempelajari Kepribadian

Usaha-usaha untuk mengerti perilaku dan kepribadian manusia sudah lama dilakukan dimulai dengan cara yang paling sederhana, yang tergolong pendekatan non-ilmiah, sampai dengan cara-cara modern atau pendekatan ilmiah.

Ada beberapa upaya mempelajari kepribadian seseorang diantaranya

1. Chirologi, yaitu pengetahuan yang berusaha mempelajari kepribadian manusia berdasarkan gurat-gurat tangan.
2. Astrologi, adalah pengetahuan yang berusaha menjelaskan kepribadian atas dasar dominasi benda-benda angkasa terhadap apa yang sedang sedang terjadi di alam, termasuk waktu kelahiran seseorang.
3. Grafologi, merupakan pengetahuan yang berusaha menjelaskan kepribadian atas dasar tulisan tangan.
4. Phisiognomi, adalah pengetahuan yang berusaha menjelaskan kepribadian atas dasar keadaan wajah.
5. Phrenologi, merupakan pengetahuan yang berusaha menjelaskan kepribadian berdasarkan keadaan tengkorak.
6. Onychology, pengetahuan yang berusaha menjelaskan kepribadian atas dasar keadaan kuku.

Pada penelitian ini cara yang digunakan adalah grafologi atau membaca kepribadian berdasarkan tulisan tangan dengan menggunakan fitur dominasi zona tulisan dan tekanan tulisan.

2.6.2 Big Five Personality

Pada dunia psikologi kepribadian memiliki berbagai macam teori menurut para ahli, dan salah satu yang paling disetujui oleh banyak ahli adalah teori *Big Five Personality Factor* yang di teliti oleh Norman pada tahun 1963 menggunakan riset dari Allport, Cattell dan lainnya sehingga menghasilkan konsep model lima faktor yang kemudian di teliti kembali oleh Lewis Goldberg yang kemudian dikenal sebagai *Big Five Personality Factor*. Banyak peneliti yang saat ini setuju bahwa perbedaan individual dapat diorganisir dalam 5 dimensi yang luas dan bipolar[29]. Oleh karena itu dinamakan *Big Five* karena memiliki keluasan dalam abstraksinya. 5 dimensi atau biasa disebut faktor tersebut yaitu *neuroticism*, *extraversion*, *openness*, *agreeableness* dan *conscientiousness* yang sering disingkat dengan OCEAN.

1. *Neuroticism*

Neuroticism mengacu pada kerentanan untuk mengalami sebuah emosi negatif, seperti kecemasan, kemarahan-permusuhan, mudah tersinggung, depresi, kesadaran diri, impulsif, dan kerentanan terhadap stres [29].

2. *Ekstraversi*

Ekstraversi yaitu sebuah pendekatan ke dunia sosial dengan energi dan aktivitas, dan mencakup sifat-sifat seperti kehangatan, ekspresi, ketegasan, pencarian kegembiraan, emosi positif, dan berorientasi pada orang [29].

3. *Openness*

Openness atau keterbukaan terhadap pengalaman yang mencakup keingintahuan intelektual, orisinalitas, kreativitas, dan apresiasi individu terhadap estetika dan sesuatu yang baru[29].

4. *Agreeableness*

Agreeableness, seperti *Ekstraversi* adalah sebuah kecenderungan untuk menjadi prososial, hangat, pemaaf, empati, terus terang, dan mudah berkomunikasi terhadap orang lain, dan mencakup sifat-sifat seperti sifat baik, kooperatif, kepercayaan, kesopanan, dan sifat lembut[29]

5. Conscientiousness

Conscientiousness adalah kecenderungan untuk termotivasi oleh tujuan, kepatuhan, terencana, dan menggambarkan kemampuan untuk mengikuti norma dan aturan yang ditentukan secara sosial, menahan dan mengendalikan emosi [29].

2.7 Grafologi

Grafologi atau analisis tulisan tangan adalah metode ilmiah mengidentifikasi, mengevaluasi dan pemahaman karakter melalui pola tulisan tangan [2]. Penulisan tiap orang berbeda beda, dari bentuk tulisan yang miring ataupun tegak hingga ketebalan penulisannya itu sendiri. Semuanya memiliki informasinya masing – masing. Tujuan mempelajari ilmu grafologi adalah untuk menilai kepribadian seseorang dari tulisan tangan[30]. Ilmu grafologi dapat menguak penggerak fisik dan mental, emosi, kepintaran, ciri kepribadian, perilaku sosial dan minat kejuruan[31]. Graphologist menginterpretasi tarikan, tekanan, dan perulangan yang akan mengungkapkan kekuatan, kelemahan dan motivasi penulis [31]. Fitur - fitur yang bisa digunakan dalam grafologi di antaranya zona, margin tulisan, *spacing*, *baseline*, ukuran huruf, tekanan, kemiringan huruf, jarak tulisan [31].

2.7.1 Zona

Zona berkaitan dengan aspek hidup seseorang. Zona tulisan dibagi menjadi 3 buah. *Upper Zone* atau Zona atas yang melambangkan aspek intelektual, imajinasi dan kecenderungan spiritual. Kode etik yang memerlukan disiplin yang diatur oleh kepala [3] seperti pada Gambar 2-6 Dominasi Zona Atas. *Middle Zone* atau zona tengah melambangkan cara berkomunikasi, kehidupan sosial, pengetahuan tentang diri sendiri dan kemampuan untuk adaptasi [3]. Contoh tulisan dengan dominasi zona tengah pada Gambar 2-7 Dominasi Zona Tengah. Terakhir adalah *Lower Zone* atau zona bawah. Zona bawah melambangkan ketidaksadaran diri, dorongan naluri, dorongan material

dan lainnya. Contoh tulisan dengan dominasi zona bawah pada Gambar 2-8 Dominasi Zona Bawah.



Gambar 2-6 Dominasi Zona Atas



Gambar 2-7 Dominasi Zona Tengah



Gambar 2-8 Dominasi Zona Bawah

2.7.2 Tekanan

Tekanan penulisan memiliki hubungan dengan tingkat emosi seseorang. Orang dengan tekanan penulisan yang kuat biasanya kuat, produktif dan ingin meninggalkan tanda mereka di masyarakat. Mereka memiliki tekad dan daya tahan yang kuat. Kedalaman emosi dan semangat agresif mereka berada di atas rata-rata [31], contoh penulisan dengan tekanan yang kuat pada Gambar 2-9 Penulisan dengan tekanan kuat. Orang yang menulis dengan tekanan sedang memiliki kedalaman emosi dan semangat yang agresif seperti orang yang menulis dengan tekanan kuat, hanya saja tidak dalam segala bidang, dapat beradaptasi dengan mudah, mudah memaafkan dan juga *easy-*

going[31], contoh penulisan dengan tekanan sedang pada Gambar 2-10 Penulisan dengan tekanan sedang. Orang yang memiliki tekanan penulisan yang ringan lebih memiliki mental energi daripada energi fisik dan juga lebih ke seorang pemikir daripada pelaku [31], contoh penulisan dengan tekanan ringan pada Gambar 2-11 Penulisan dengan tekanan ringan.

Gambar 2-9 Penulisan dengan tekanan kuat

Gambar 2-10 Penulisan dengan tekanan sedang

Gambar 2-11 Penulisan dengan tekanan ringan

2.8 Definisi Kepribadian dan Deteksi Kepribadian

Kepribadian berasal dari bahasa Inggris yaitu *personality*, Belanda (*personalita*), Prancis (*personalia*), Jerman (*personlichekesit*), Italia (*personalita*), dan Spanyol (*personalidad*). Sedangkan akar katanya berasal dari bahasa latin yaitu *persona* yang berarti topeng, maksudnya topeng yang dipakai oleh aktor. Kepribadian adalah organisasi dinamik dalam sistem psikofisik individu yang menentukan penyesuaian yang unik dengan lingkungannya. Suatu fenomena dinamik yang memiliki elemen psikologi dan fisiologi, yang berkembang dan berubah, yang memainkan peran aktif dalam berfungsinya individu[32].

Berbagai penelitian awal mengenai kepribadian berkisar untuk mengidentifikasi dan menamai karakteristik permanen yang menjelaskan perilaku individu seseorang. Karakteristik yang umumnya melekat dalam diri seorang individu adalah malu, agresif, patuh, malas, ambisius, setia, dan takut. Karakteristik-karakteristik tersebut jika ditunjukkan dalam berbagai situasi, disebut sifat-sifat kepribadian. Sifat kepribadian menjadi suatu hal yang mendapat perhatian cukup besar

karena para peneliti telah lama meyakini bahwa sifat-sifat kepribadian dapat membantu proses seleksi karyawan, menyesuaikan bidang pekerjaan dengan individu, dan memandu keputusan pengembangan karier. Terdapat sejumlah upaya awal untuk mengidentifikasi sifat-sifat utama yang mengatur perilaku. Seringnya, upaya ini sekadar menghasilkan daftar panjang sifat yang sulit untuk digeneralisasikan dan hanya memberikan sedikit bimbingan praktis bagi para pembuat keputusan organisasional. Dua pengecualian adalah Myers-Briggs Type Indikator dan Model Lima Besar. Selama 20 tahun hingga saat ini, dua pendekatan ini telah menjadi kerangka kerja yang dominan untuk mengidentifikasi dan mengklasifikasikan sifat-sifat seseorang.

2.9 Ekstraksi Fitur

Ekstraksi fitur adalah proses merubah data input menjadi fitur – fitur pembeda untuk data input tersebut [19]. Tujuan dari ekstraksi fitur adalah untuk mendapatkan informasi yang paling relevan dari data input dan merepresentasikan informasi tersebut ke satuan data yang lebih rendah[19]. Ciri atau karakteristik inilah yang digunakan sebagai parameter untuk menggambarkan sebuah objek. Nilai dari parameter inilah yang akan digunakan dalam proses klasifikasi citra. Ekstraksi fitur dilakukan sesudah citra input di proses (*preprocessing*) dan dapat dilakukan setelah tahapan segmentasi citra ataupun tanpa segmentasi citra. Fitur yang paling umum untuk di ekstraksi adalah kemiringan tulisan, *baseline*, zona dan tekanan tulisan.[33]

2.9.1 Ekstraksi Fitur Gray-level Distribution

Ekstraksi fitur untuk mendapat nilai dari fitur tekanan tulisan pada citra dapat dilakukan dengan algoritma yang dibuat oleh Mukherjee dan Ishita yaitu Gray-level distribution[33]. Fitur tekanan yang didapat menggunakan algoritma ini yaitu rata - rata intensitas piksel dari citra *grayscale* yang kurang dari 150, serta nilai persentase dari nilai – nilai tersebut. Berikut adalah algoritma yang digunakan untuk mendapatkan nilai ekstraksi ciri tekanan tulisan :

1. Input citra grayscale yang sudah melalui tahap median filter.

2. Ambil nilai baris dan kolom dari citra, misalkan $m = \text{jumlah baris}$ dan $n = \text{jumlah kolom}$
3. for $i = 1$ to m
 - for $j = 1$ to n
 - if ($\text{grayscale}(i,j) < 150$)
 - $\text{total_intensitas} = \text{total_intensitas} + \text{grayscale}(i,j)$
 - $\text{jumlah_piksel} = \text{jumlah_piksel} + 1$
 - end if
 - end for
- end for
4. $\text{rerata} = \text{total_intensitas} / \text{jumlah_piksel}$
5. $\text{persentase} = (\text{rerata} * 100) / 255$

Contoh hasil penerapan metode ekstraksi fitur tekanan tulisan dapat dilihat pada Tabel 2-2 Hasil Dari Metode Ekstraksi Fitur.

Tabel 2-2 Hasil Dari Metode Ekstraksi Fitur



Total Intensitas : 840278
 Jumlah Piksel : 11916
 Rata-Rata Intensitas : 70.52
 Persentase : 27.65

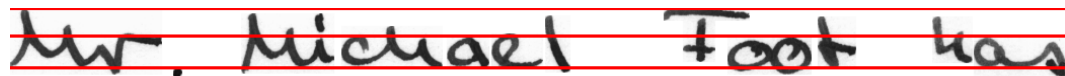
2.9.2 Ekstraksi Fitur Zoning

Ekstraksi fitur untuk mendapat nilai dari fitur zona tulisan pada citra dapat dilakukan dengan algoritma yang dibuat oleh Mukherjee dan Ishita yaitu *handwritten text feature extraction* [33]. Fitur dari dominasi zona tulisan yang digunakan pada proses klasifikasi adalah lebar piksel dari tiga zona tulisan yaitu zona atas, zona tengah dan zona bawah. Lebar zona didapat dengan mencari 4 titik koordinat y pada citra yang akan menjadi pembatas antar zona. Berikut adalah algoritma yang digunakan untuk mendapatkan nilai ekstraksi ciri dominasi zona tulisan :

1. Input citra biner yang sudah melalui tahap preprocessing.
2. Hitung jumlah piksel hitam di setiap baris.
3. Temukan nilai frekuensi horizontal tertinggi dari hasil perhitungan sebelumnya.
4. Hitung $1/3$ dari frekuensi horizontal maksimum. Nilai ini disebut T.
5. Cari baris pertama dari atas yang nilai frekuensinya tidak nol. Baris ini dinamakan top.
6. Cari baris pertama dari bawah yang nilai frekuensinya tidak nol. Baris ini dinamakan bottom.
7. Lintasi citra dari baris bawah ke baris atas dan periksa di mana jumlah piksel hitam dari baris yang melebihi nilai T. Nomor baris ini dinamakan r1.
8. Lintasi citra dari baris atas ke baris bawah dan periksa di mana jumlah piksel hitam dari baris yang melebihi nilai T. Nomor baris ini dinamakan r2.
9. Hitung lebar zona atas dengan $r2 - top$.
10. Hitung lebar zona tengah dengan $r1 - r2$.
11. Hitung lebar zona bawah dengan $bottom - r1$

Contoh hasil penerapan metode ekstraksi fitur tekanan tulisan dapat dilihat pada Tabel 2-3 Hasil Dari Metode Ekstraksi Fitur.

Tabel 2-3 Hasil Dari Metode Ekstraksi Fitur



Zona Atas: 26.0

Zona Tengah: 30.0

Zona Bawah: 11.0

2.10 Anaconda

Anaconda adalah aplikasi *open-source* terdistribusi untuk bahasa pemrograman Python dan R yang digunakan untuk komputasi ilmiah seperti *data science*, *machine*

learning, predictive analysis dan lainnya. Anaconda bertujuan untuk mempermudah *packaging* manajemen.

Anaconda Navigator adalah aplikasi *desktop* GUI yang berada di dalam paket instalasi Anaconda. Melalui Anaconda Navigator, pengguna dapat dengan mudah menjalankan dan mengelola Conda, *environment* dan kanal tanpa perlu menggunakan perintah *command prompt*. Anaconda Navigator dapat mencari paket – paket yang dibutuhkan pada *repository cloud* anaconda.