

Bidang Penelitian: TIK

**LAPORAN AKHIR PENELITIAN UNGGULAN KOMPETITIF  
UNIVERSITAS SRIWIJAYA**

**PREDIKSI KEPERIBADIAN SESEORANG BERDASARKAN TULISAN  
TANGAN MENGGUNAKAN PENDEKATAN MULTIPLE CLASSIFIERS  
DAN CONVOLUTIONAL NEURAL NETWORKS**



Oleh:

**Ketua Penelitian: Samsuryadi, S.Si., M.Kom., Ph.D. (0004027101)**  
**Anggota Peneliti : 1. Julian Supardi, S.Pd., M.T. (0010077210)**  
**2. Dr. Ir. Sukemi, M.T. (0003126604)**

Dibiayai oleh:

Anggaran DIPA Badan Layanan Umum  
Universitas Sriwijaya Tahun Anggaran 2021  
Nomor SP DIPA-023.17.2.677515/2021, tanggal 23 November 2020  
Sesuai dengan SK Rektor  
Nomor : 0010/UN9/SK.LP2M.PT/2021  
Tanggal 28 April 2021

**PROGRAM STUDI ILMU KOMPUTER  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA  
Tahun Anggaran 2021**

**HALAMAN PENGESAHAN LAPORAN AKHIR  
SKEMA PENELITIAN UNGGULAN KOMPETITIF**

- |   |  |
|---|--|
| 1. Judul Penelitian   | : Prediksi Kepribadian Seseorang Berdasarkan Tulisan Tangan Menggunakan Multiple Classifiers dan Convolutional Neural Networks                                       |
| 2. Bidang Penelitian  | : Teknologi Informasi dan Komunikasi   |
| 3. Ketua Peneliti   |  |
| a. Nama Lengkap   | : Samsuryadi, S.Si., M.Kom., Ph.D.   |
| b. NIDN/NIDK  | : 0004027101   |
| c. Pangkat dan Golongan   | : Pembina, IV-A  |
| d. Fakultas/Jurusan/Prodi   | : Ilmu Komputer/Teknik Informatika/Ilmu Komputer   |
| e. Telepon/HP/E-mail  | : 08127814524  |
| 4. Jumlah Anggota Peneliti  | :  |
| a. Nama Anggota I   | : Julian Supardi, S.Pd., M.T.  |
| NIDN/NIDK   | : 0010077210   |
| b. Nama Anggota II  | : Dr. Ir. H. Sukemi, M.T.  |
| NIDN/NIDK   | : 0003126604   |
| 5. Jangka Waktu Penelitian  | : 8 (delapan) bulan  |
| 6. Jumlah Dana yang Disetujui   | : Rp. 55.000.000 (Lima puluh lima juta rupiah)   |
| 7. Target Luaran TKT  | : Level 3  |
| Nama, NIM dan Jurusan/<br>8. Program Studi/BKU<br>Mahasiswa yang Terlibat | 1. Rudi Kurniawan (03013681924008)/Teknik Informatika<br>2. Yusa Virginiawan Guntara (09012681923008)/Ilmu Komputer<br>3. Vincen (09021181722007)/Teknik Informatika |



Jauhan Jauhan, S.Pd., M.T.  
NIP. 197107212005011005

Indralaya, 23 November 2021  
Ketua Peneliti,

Samsuryadi, M.Kom., Ph.D.  
NIP. 197102041997021003

Indralaya, November 2021  
Sekretaris LPPM Universitas Sriwijaya,



Prof. Dr. Mulyadi Eko Purnomo, M.Pd.  
NIP. 195901171983031014

## KATA PENGANTAR

Alhamdulillah dipanjatkan syukur kepada Yang Maha segala-galanya sehingga laporan penelitian ini dapat terselesaikan dengan lancar dan baik. Penelitian pada tahun pertama ini berfokus untuk memprediksi kepribadian seseorang berdasarkan tulisan tangan menggunakan *K-Nearest Neighbor* (KNN), *Support Vector Machine* (SVM), dan *Decision Tree*. Penelitian ini membahas pendekatan untuk mengatasi masalah pada tahapan pemrosesan dalam memprediksi kepribadian seseorang dengan cara menghilangkan derau, mengubah citra abu-abu ke citra biner menggunakan *inverted thresholding*, dilasi, kontur, transformasi *affine* digunakan untuk merotasi teks tulisan tangan agar benar-benar *horizontal*, dan *deslanting technique* untuk mengetahui sudut kemiringan dari suatu teks. Analisis kepribadian seseorang menggunakan metode pengukuran psikologi *Five Factor Model* (FFM) dan klasifikasi menggunakan metode KNN, SVM, dan *Decision Tree*.

Tujuan penelitian ini adalah (1) mendesain suatu model baru dengan pendekatan pengolahan citra dan pembelajaran mesin untuk memprediksi kepribadian seseorang berdasarkan metode FFM, dan (2) menguji kehandalan model baru untuk memprediksi kepribadian seseorang berdasarkan citra tulisan tangan menggunakan pendekatan pengolahan citra dan pembelajaran mesin berdasarkan metode FFM.

Manfaat penelitian ini model yang dikembangkan dengan pendekatan pengolahan citra dan pembelajaran mesin dapat digunakan untuk memprediksi kepribadian seseorang berdasarkan metode FF, dan model ini juga dapat mempermudah Psikolog dalam memprediksi kepribadian seseorang berdasarkan citra tulisan tangan.

Kebaharuan (*novelty*) penelitian ini adalah menyajikan suatu model dan mekanisme klasifikasi psikologi kepribadian seseorang berdasarkan citra tulisan tangan dengan pendekatan pengolahan citra dan pembelajaran mesin. Setiap tahapan memiliki peranan penting dalam penelitian ini dan terintegrasi dalam

sebuah model. Namun, fokus utama dalam penelitian ini adalah mengkaji tahapan *preprocessing* dan algoritma ekstraksi fitur serta uji performansi model. Pengujian model melalui eksperimen dan performa serta melakukan komparasi dengan beberapa *classifiers*.

Target capaian luaran penelitian pada tahun pertama Skripsi dan Tesis, 1 (satu) draf Jurnal Internasional bereputasi dan TKT minimal 3.

Kami menyadari bahwa hasil penelitian masih belum sempurna dan masih ada kekurangannya. Oleh karena itu, kami akan melanjutkannya sesuai rencana seperti yang diajukan pada proposal pada tahun kedua (2022) untuk melengkapi dan menyempurnakan kelemahan tersebut. Walaupun demikian, penelitian yang dilakukan telah menghasilkan perpaduan teknik pengolahan citra dengan pembelajaran mesin (*decision tree* dan SVM) yang mampu menghasilkan prediksi kepribadian seseorang mencapai lebih dari 99%.

Salam Takzim

Tim Peneliti

## DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	<b>Error! Bookmark not defined.</b>
KATA PENGANTAR .....	iii
DAFTAR ISI.....	v
I. Ringkasan.....	1
II. Latar Belakang .....	2
III. Rumusan Masalah.....	3
IV. Tujuan Penelitian .....	4
V. Manfaat Penelitian .....	4
VI. Target Luaran.....	4
VII. Tinjauan Pustaka .....	4
VIII. Metode Penelitian.....	9
A. Kerangka Penelitian .....	9
B. Desain Penelitian.....	13
Hasil dan Pembahasan	
A. Hasil Pengujian Ekstraksi Tulisan Tangan .....	14
B. Uji Performansi .....	22
IX. Kesimpulan dan Saran .....	28
A. Kesimpulan .....	28
B. Saran.....	29
X. Daftar Pustaka .....	29

## I. Ringkasan

Kepribadian seseorang dapat ditentukan berdasarkan pada ciri-ciri dari suatu tulisan tangan. Ciri-ciri yang terdapat pada tulisan tangan seseorang dianalisis oleh grafologis untuk memprediksi kepribadian seseorang. Tulisan tangan memiliki pola tertentu berkaitan dengan mental, fisik, keadaan emosional, dan tingkah laku dari pembuatnya. Perkembangan komputerisasi saat ini membuat otomatisasi grafologi untuk menentukan kepribadian seseorang melalui tulisan tangan digital. Pemilihan teknik, model-model pendekatan, dan algoritma ekstraksi fitur yang tepat dalam setiap tahapan pemrosesan menjadi sangat krusial dan menjadi tantangan tersendiri untuk mendapatkan hasil yang maksimal dalam mengidentifikasi kepribadian seseorang. Salah satu pendekatan yang digunakan untuk mengatasi masalah tersebut adalah dengan teknik-teknik pengolahan citra, seperti teknik-teknik *bilateral filtering* untuk menghilangkan *noise*, *inverted thresholding* untuk mengubah citra abu-abu ke citra biner, *dilation*, *contour*, *affine transformation* untuk merotasi teks tulisan tangan sehingga benar-benar horizontal, dan *deslanting technique* untuk mengetahui sudut kemiringan dari suatu teks. Metode pengukuran psikologi *Five Factor Model* (FFM) digunakan untuk menganalisis kepribadian seseorang dan klasifikasi menggunakan KNN, SVM, dan *Decision Tree*.

Tujuan penelitian ini adalah (i) merancang suatu model baru menggunakan pendekatan pengolahan citra dan pembelajaran mesin untuk memprediksi kepribadian seseorang berdasarkan metode FFM, (ii) menguji model baru untuk memprediksi kepribadian seseorang berdasarkan citra tulisan tangan dengan pendekatan pengolahan citra dan pembelajaran mesin, sehingga dapat mengklasifikasi psikologi dan kepribadian seseorang berdasarkan metode FFM.

Manfaat penelitian ini adalah konsep pengolahan citra dan pembelajaran mesin untuk permasalahan grafologi dapat digunakan pada bidang ilmu psikologi dalam hal ini memprediksi kepribadian seseorang. Sedangkan manfaat teoritis penelitian ini dapat membantu memecahkan permasalahan grafologi yang berhubungan dengan metode FFM menggunakan konsep pengolahan citra dan pembelajaran mesin.

Kebaruan (*novelty*) penelitian ini adalah menyajikan suatu model dan mekanisme klasifikasi psikologi kepribadian seseorang berdasarkan citra tulisan tangan dengan pendekatan pengolahan citra dan pembelajaran mesin. Setiap tahapan memiliki peranan penting dalam penelitian ini yang terintegrasi dalam sebuah model. Namun, fokus utama dalam penelitian ini adalah teknik-teknik pengolahan citra pada tahapan *preprocessing* dan algoritma ekstraksi fitur serta uji performansi model. Selanjutnya, model diuji melalui eksperimen dan performa serta melakukan komparasi dengan beberapa *classifiers*.

Target capaian luaran penelitian pada tahun pertama Skripsi dan Tesis Mahasiswa, publikasi sebanyak dua artikel di Jurnal Nasional Terakreditasi minimal Sinta 3 atau Jurnal Internasional bereputasi dan TKT minimal 3.

## **II. Latar Belakang**

Penelitian pertama kali yang membahas otomatisasi *grafologi* dan disebut dengan *Computer-Aided Graphology* menggunakan prinsip-prinsip pengenalan pola yang terdiri dari tiga tahapan utama, antara lain: tahapan *pre-processing*, tahapan ekstraksi fitur-fitur tulisan tangan, dan tahapan analisis klasifikasi terhadap fitur-fitur yang telah diekstraksi [1]. Dari tahapan-tahapan tersebut, menjadi suatu model atau pendekatan yang tidak dapat terpisahkan dalam membangun suatu otomatisasi *grafologi*. Setelah itu, otomatisasi *grafologi* berkembang pesat dan menjadi area penelitian tersendiri untuk menentukan kepribadian seseorang melalui tulisan tangan.

Meskipun telah banyak dikembangkan berbagai model dan pendekatan dalam penelitian yang berhubungan dengan otomatisasi *grafologi*, namun masih terdapat banyak kendala yang dihadapi hingga saat ini seperti: pemilihan teknik-teknik *pre-processing* dalam mengolah data digital tulisan tangan, algoritma pengolahan citra untuk mengekstraksi fitur-fitur tulisan tangan, dan teknik-teknik klasifikasi yang tepat untuk mendapatkan akurasi yang maksimal dari setiap pemodelan yang dibangun.

Pada tahapan *pre-processing*, metode yang paling umum digunakan adalah menghilangkan derau (*noise*) hasil pemindaian dan proses binerisasi.

Penghilangan *noise* terutama *salt and pepper noise* yang ada pada citra tulisan tangan, beberapa peneliti menggunakan teknik median filter [2], [3]. Sedangkan untuk proses binerisasi digunakan otsu thresholding [4]. Pada tahapan ekstraksi fitur, beberapa penelitian menggunakan algoritma pengolahan citra seperti: skew normalization untuk menentukan jarak spasi dan baseline [5], dan *projection histogram method* untuk fitur *baseline* dan *thresholding* untuk fitur *pen pressure* [6]. Pada tahapan klasifikasi, beberapa pendekatan *machine learning* bisa digunakan antara lain: KNN [7], SVM [8], [9], dan CNN [10].

Berdasarkan uraian di atas, laporan penelitian ini menyajikan suatu model klasifikasi psikologi dari kepribadian seseorang dengan metode FFM melalui tulisan tangan dengan pendekatan pengolahan citra dan pembelajaran mesin. Arsitektur Model yang disajikan mulai dari tahapan *pre-processing* menggunakan *bilateral filter* sebagai *noise removal* [11], *inverted global thresholding* untuk binerisasi citra, dilasi, kontur, dan transformasi affine [12]. Selanjutnya dalam tahapan ekstraksi fitur tulisan tangan, menyajikan algoritma ekstraksi dengan menggunakan *OpenCV library* untuk menentukan 7 fitur, antara lain: *baseline*, *top margin*, *line spacing*, *word spacing*, *letter size*, *slant*, dan *pen pressure*. Tahapan klasifikasi menyajikan pengelompokan psikologi kepribadian seseorang berdasarkan hasil ekstraksi tulisan tangan dengan *multi-classifiers* sehingga didapat analisis terhadap performansi sistem yang maksimal dari masing-masing *classifier* tersebut.

### **III. Rumusan Masalah**

Rumusan masalah pada penelitian tahun pertama ini adalah bagaimana merancang dan menguji suatu model baru dengan pendekatan pengolahan citra dan pembelajaran mesin untuk memprediksi kepribadian seseorang berdasarkan metode FFM?



#### **IV. Tujuan Penelitian**

Tujuan penelitian ini adalah:

1. Mendesain suatu model baru dengan pendekatan pengolahan citra dan pembelajaran mesin untuk memprediksi kepribadian seseorang berdasarkan metode FFM;
2. Menguji kehandalan model baru untuk memprediksi kepribadian seseorang berdasarkan citra tulisan tangan menggunakan pendekatan pengolahan citra dan pembelajaran mesin berdasarkan metode FFM.

#### **V. Manfaat Penelitian**

Manfaat penelitian adalah sebagai berikut:

1. Model baru yang dikembangkan dengan pendekatan pengolahan citra dan pembelajaran mesin dapat digunakan untuk memprediksi kepribadian seseorang berdasarkan metode FFM;
2. Model baru yang dikembangkan dengan pendekatan pengolahan citra dan pembelajaran mesin dapat mempermudah Psikolog dalam memprediksi kepribadian seseorang berdasarkan citra tulisan tangan.

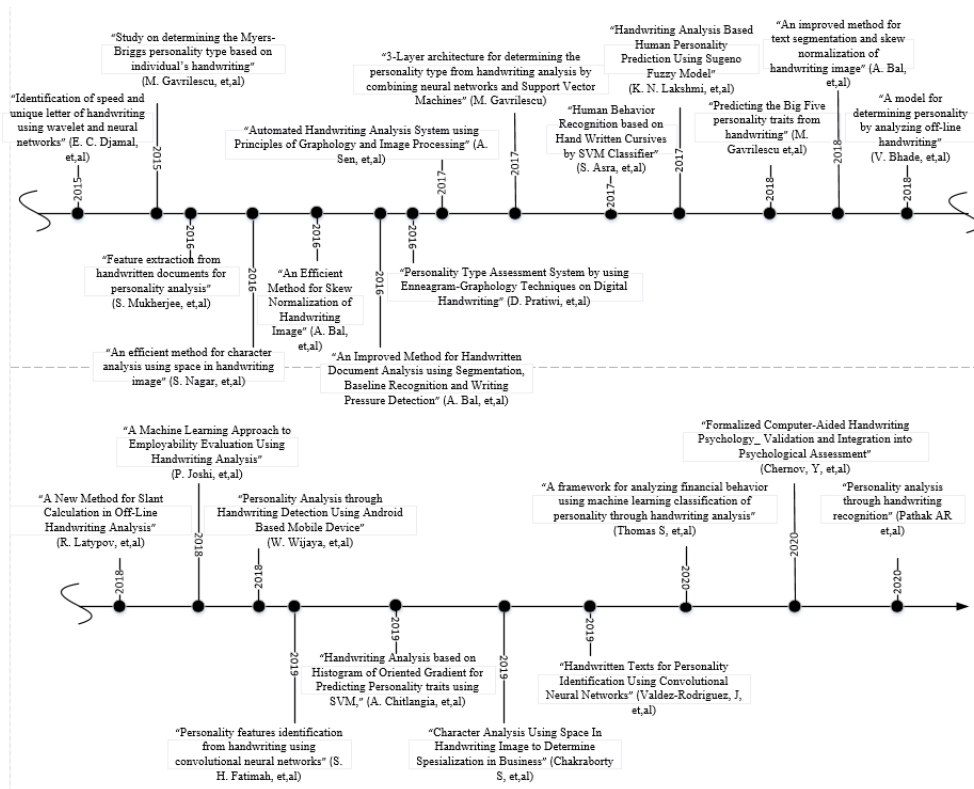
#### **VI. Target Luaran**

Target luaran penelitian pada tahun pertama Skripsi dan Tesis Mahasiswa, publikasi artikel di Jurnal Nasional Terakreditasi minimal Sinta 3 atau Jurnal Internasional bereputasi dan TKT minimal 3. Skripsi telah selesai dibuat oleh mahasiswa yang mengikuti penelitian ini dan sedang menunggu penjadwalan ujian komprehensif, sedangkan Tesis telah selesai sampai Bab III. Artikel untuk Jurnal Internasional bereputasi dalam bentuk draft. Capaian TKT 3 telah terpenuhi.

#### **VII. Tinjauan Pustaka**

Penelitian berkaitan otomatisasi grafologi telah menjadi topik yang menarik dan penting untuk dikaji. Gambar 1 menyajikan peta jalan (*road map*)

penelitian yang berhubungan dengan otomatisasi grafologi dalam kurun waktu 5 tahun terakhir.



Gambar 1. Road Map Penelitian Automatisasi Grafologi 5 Tahun Terakhir

Salah satu metode pengukuran psikologi dalam menentukan kepribadian seseorang seperti *Myer-Briggs Type Indicators* (MBTI) diterapkan penggunaannya dalam otomatisasi *graphology* dan membaginya ke dalam 3 (tiga) tingkatan arsitektur, yaitu: tingkatan dasar, tingkatan menengah, dan tingkatan atas seperti yang tersaji pada penelitian [13], Tingkatan dasar (*base level*) digunakan untuk menentukan fitur-fitur tulisan tangan dengan menerapkan normalisasi dan pemecahan karakter. Pada tingkatan menengah (*intermediate level*) terdiri dari 4 (empat) *Neural Network* (NN) yang merepresentasikan dikotomi-dikotomi dari MBTI, yaitu *Extrovert vs Introvert* (E/I), *Sensing vs Intuition* (S/N), *Thinking vs Feeling* (T/F), dan *Judging vs Perceiving* (J/P). Fitur-fitur yang diekstraksi pada tingkatan dasar antara lain: *baseline*, *pen pressure*, *slant*, *stroke*, huruf 't', dan huruf 'f'. Pada tingkatan atas, terdiri dari *Neural*

*Network* (NN) yang digunakan untuk melakukan pelatihan dan pengujian. Sebanyak 64 koresponden digunakan untuk mengambil sampel data selama 2 (dua) bulan dan dilakukan pengulangan setiap 2 (dua) minggu sekali. Pengujian dilakukan terhadap sampel tulisan tangan dan menghasilkan akurasi sebesar 86.7 %, sedangkan pengujian untuk data acak pada subjek yang sama menghasilkan akurasi sebesar 78.8%.

Teknik-teknik dalam pengolahan citra yang digunakan dalam *graphology* seperti segmentasi, *baseline*, dan *pen pressure* juga diusulkan dalam penelitian [6]. *Otsu thresholding* digunakan dalam teknik binerisasi, dan untuk menghilangkan derau dalam bentuk *salt and pepper noise* digunakan *median filter*. Dalam melakukan segmentasi, digunakan teknik *horizontal* dan *vertical projection histogram*, kemudian baru dilakukan teknik normalisasi dengan menggunakan *skew normalization*. Pengukuran deviasi standar digunakan untuk mengkalkulasikan tekanan pena yang dihasilkan pada citra digital tulisan tangan. Dengan menggunakan dataset publik IAM dataset, dihasilkan akurasi dalam proses segmentasi baris sebesar 95.65%, segmentasi kata sebesar 92.56% dengan menerapkan algoritma pengolahan citra. Total akurasi menggunakan normalisasi mencapai 96%.

Penelitian berbasis otomatisasi *graphology* dengan menerapkan metode *enneagram* dalam teknik pengukuran psikologinya telah dilakukan oleh [14], Data yang telah dipindai dengan *scanner* kemudian diubah ke dalam bentuk citra derajat keabuan dan kemudian dilakukan binerisasi. *Fuzzy C-Means* digunakan untuk mengklusterisasi fitur-fitur tulisan tangan seperti *baseline*, *slant*, dan *letter size*. Data yang digunakan sebanyak 49 sampel tulisan tangan dengan menghasilkan akurasi pada rentang 80-87%. Penelitian ini diungkapkan bahwa implementasi otomatisasi *graphology* menjadikan proses penilaian psikologi seseorang menjadi lebih cepat dibandingkan dengan menjawab sekumpulan pertanyaan tes psikologi.

Pada penelitian [9], melakukan pemetaan fitur-fitur analisis tulisan tangan menggunakan metode psikologi kepribadian *Five Factor Models* (FFM) yang terdiri dari: *baseline*, *slant*, *pen pressure*, *spacing*, huruf 't', dan huruf 'f' dan

dilakukan klasifikasi dengan teknik FFNN serta *Template Matching* menghasilkan akurasi dengan teknik sebesar 84.4%.

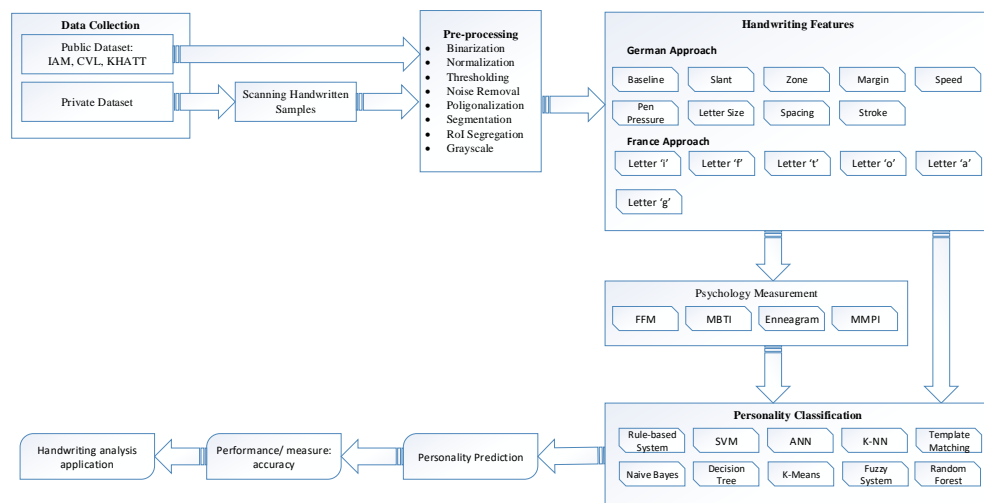
Untuk memprediksi kepribadian seseorang, dapat melakukan kombinasi fitur-fitur antara teknik Jerman dan teknik Prancis seperti yang dilakukan pada penelitian [10]. Pada tahapan pemrosesan awal, digunakan teknik pengolahan citra seperti: *resizing*, *grayscale*, dan *thresholding*. Proses ekstraksi fitur dibagi menjadi dua bagian, yaitu mengekstraksi fitur-fitur dengan teknik Jerman dengan menerapkan teknik segmentasi yang terdiri dari: margin, spasi, *slant*, dan zona. Bagian yang kedua dengan teknik Prancis meliputi fitur-fitur seperti: huruf 'a', huruf 'g', dan huruf 't'. Dataset terdiri dari 40 sampel tulisan tangan yang ditulis pada kertas A4. Proses pengklasifikasian pada teknik Jerman menerapkan *rule-based system* yang menghasilkan akurasi sebesar 82.5-100%, sedangkan proses pengujian pada teknik Prancis menggunakan *Convolutional Neural Network* (CNN) dengan akurasi sebesar 98%.

*Graphology* dapat juga digunakan untuk menentukan kepribadian seseorang dengan menganalisis tingkah laku dalam hal penggunaan finansial mereka seperti yang dilakukan dalam penelitian [15]. Dengan menggunakan metode FFM yang menjadi aspek pengukuran psikologi kepribadian, menyajikan sebuah pendekatan dalam bentuk kerangka kerja (*framework*). Pada tahapan pemrosesan awal, beberapa teknik-teknik digunakan antara lain: *noise reduction*, *resizing*, dan *thresholding*. Kemudian untuk mengekstraksi fitur-fitur seperti: margin, spasi, *slant* digunakan teknik segmentasi. Hasil klasifikasi menggunakan CNN dengan akurasi sebesar 97.27% dan data yang digunakan sebanyak 112 sampel data.

Meskipun pada penelitian-penelitian terdahulu telah diusulkan banyak metode dan teknik yang digunakan dalam otomatisasi grafologi, dan telah menunjukkan performa yang tinggi, namun demikian masih terdapat beberapa isu terkait model yang dibangun, dimulai dari tahapan *preprocessing*, ekstraksi fitur tulisan tangan, dan pengklasifikasian kepribadian melalui tulisan tangan. Dari hasil studi penelitian-penelitian yang telah dilakukan untuk menyelesaikan masalah tersebut, disimpulkan beberapa peluang di antaranya: (i) adanya celah penggunaan teknik filterisasi citra dengan *bilateral filtering*, binerisasi dengan

*inverted global thresholding*, segmentasi dan normalisasi menggunakan teknik dilasi, kontur, transformasi affine, dan teknik *deslanting*; (ii) ekstraksi fitur-fitur tulisan tangan menggunakan pendekatan pemrosesan citra; (iii) klasifikasi kepribadian masih terfokus pada klasifikasi masing-masing fitur; (iv) masih terdapat ruang eksplorasi terhadap hasil klasifikasi kepribadian dengan menggunakan metode pengukuran psikologi dengan pendekatan pembelajaran mesin. Mengacu pada peluang tersebut perlu dilakukan perancangan suatu model/*framework* yang mampu menganalisis psikologi dari kepribadian seseorang berdasarkan citra tulisan tangan dengan pendekatan pengolahan citra dan pembelajaran mesin.

Dari beberapa survei perkembangan otomatisasi *graphology* yang telah dijabarkan, dapat disimpulkan bahwa proses otomatisasi *graphology* terdiri dari beberapa proses, antara lain: (i) Proses pengumpulan data dan akuisisi data, (ii) Proses awal pengolahan citra digital, (iii) Proses ekstraksi fitur-fitur tulisan tangan berdasarkan prinsip-prinsip *graphology*, (iv) Proses klasifikasi ciri-ciri kepribadian, (v) Pengujian performansi yang berupa akurasi sistem. Gambar 2 menyajikan tahapan setiap proses yang dilalui dalam sistem otomatisasi *graphology*.

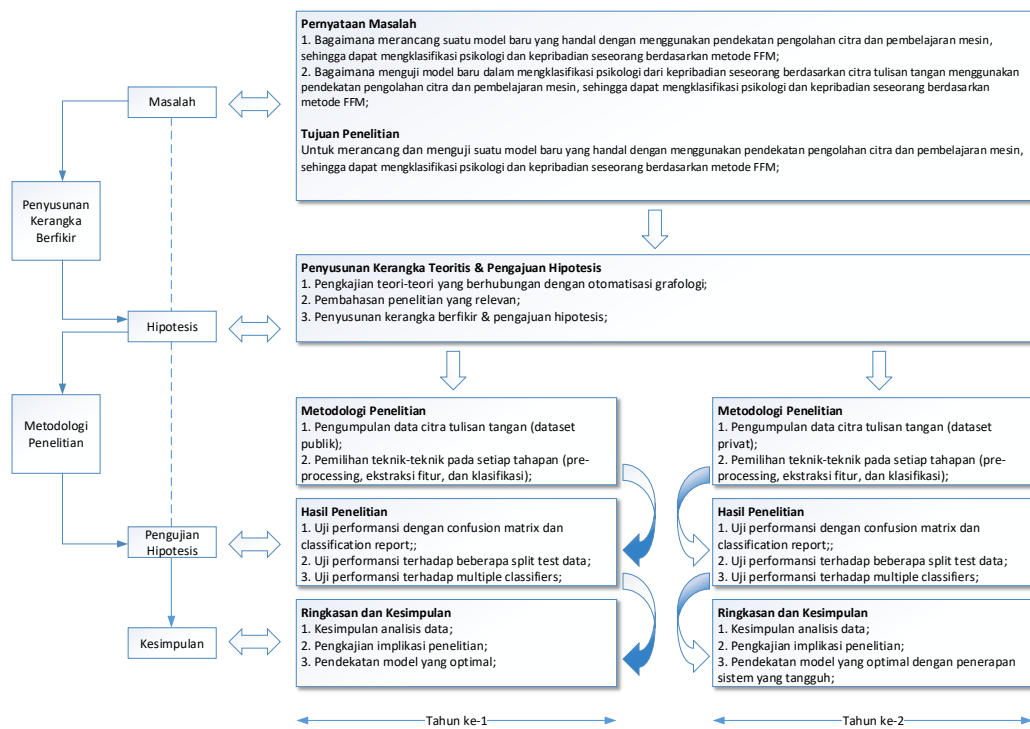


Gambar 2. Proses Automatisasi Grafologi

## VIII. Metode Penelitian

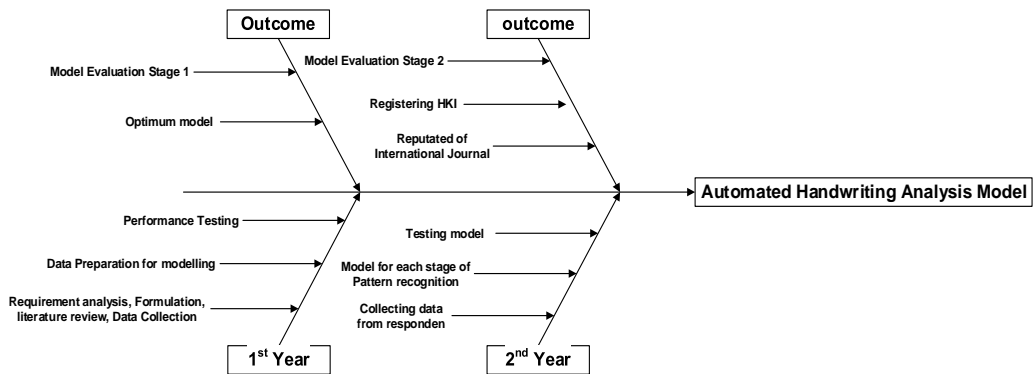
### A. Kerangka Penelitian

Kerangka penelitian menggambarkan rencana global penelitian dan luaran yang akan dicapai dari setiap kegiatan. Kerangka penelitian yang disusun pada Gambar 3 menggambarkan keterkaitan antara pertanyaan penelitian, tujuan penelitian, metodologi, pengukuran dan hasil yang ingin dicapai.



Gambar 3. Kerangka Penelitian

Adapun *roadmap* tahapan penelitian (*multi years*) dijabarkan pada Gambar 4 yang berupa diagram *fishbone*.



Gambar 4. Roadmap Tahapan Penelitian

Penjelasan dan rincian dari *road map* tersebut tersaji dalam Tabel 2.

Tabel 2. Rincian Kerangka Penelitian

<b>Prediksi Kepribadian Seseorang Menggunakan Tulisan Tangan dengan Pendekatan <i>Image Processing</i> dan Five Factor Model (FFM)</b>	
<b>Tahun Pertama (8 Bulan)</b>	<b>Tahun Kedua (8 Bulan)</b>
Perumusan masalah; 2. <i>Literature Review</i> : - <i>Handwriting &amp; Personality</i> ; - <i>Graphology</i> ; - <i>Pattern Recognition</i> ; - <i>Personality Identification Through Handwriting Analysis</i> ; 3. Pengumpulan Data : - Database publik tulisan tangan; - Normalisasi database; 4. Kebutuhan Perangkat yang akan digunakan ( <i>training machine</i> ); 5. Pembuatan data <i>train</i> dan dataset untuk <i>handwriting analysis</i> ; 6. Uji performansi model;	1. Pengumpulan dan pengolahan data responden tulisan tangan; 2. Pembuatan Model untuk setiap tahapan <i>Pattern Recognition</i> - Model untuk tahap <i>Preprocessing</i> : ( <i>noise removal</i> dengan bilateral filtering, binerisasi dengan inverted thresholding, normalisasi dengan teknik dilasi, kontur, transformasi affine, dan teknik <i>deslanting</i> ); - Model untuk Tahap <i>Feature Extraction</i> dengan menggunakan pendekatan <i>image processing</i> : (ekstraksi fitur yang digunakan, yaitu: <i>baseline, top margin, line spacing, word spacing, letter size, slant, dan pen pressure</i> ) - Model Untuk Tahap klasifikasi: <i>multiple classifiers</i> (KNN, SVM, Decision Tree); 3. Pengujian Keakuratan dan waktu yang dibutuhkan dalam Training data untuk <i>Handwriting Analisis</i> berdasarkan model yang telah dibangun;
Tingkat Kesiapterapan Sistem : - Prinsip dasar dari teknologi diteliti dan dilaporkan (TKT 1); - Formulasi konsep dan / atau aplikasi formulasi (TKT 2);	Tingkat Kesiapterapan Sistem : Pembuktian konsep fungsi dan / atau karakteristik penting secara analisis dan eksperimental (TKT 3);
Luaran :	Luaran :

- Evaluasi Sistem Tahap 1; - Model optimal;	- Evaluasi Sistem Tahap II; - Jurnal Internasional; - HKI sederhana;
--	--

a. Studi Pustaka

Pada tahap ini penulis melakukan pencarian sumber pustaka dan telaah pustaka. Pustaka yang dikumpulkan dan ditelaah berhubungan dengan teori-teori dan konsep bagaimana membangun suatu otomatisasi grafologi secara terperinci, di antaranya tentang: (i) perkembangan otomatisasi grafologi, (ii) data citra tulisan tangan yang digunakan, (iii) teknik-teknik pemrosesan awal berbasis image processing, (iv) fitur-fitur tulisan tangan yang digunakan, (v) metode pengukuran psikologi berdasarkan fitur-fitur tulisan tangan, (vi) metode-metode klasifikasi yang digunakan, (vii) pengukuran performansi dari sistem, (viii) dan teori-teori pendukung lainnya. Penelusuran referensi dilakukan melalui database IEEE, ScienceDirect, ACM Digital Library, Google Scholar, dan beberapa ebook yang berkaitan dengan penelitian sebagai penunjang. Untuk pengelolaan daftar referensi digunakan tools Mendeley. Melalui tahapan ini penulis mendapatkan wawasan tentang penelitian-penelitian yang berhubungan dengan otomatisasi grafologi (*state Of the art*), isu utama, tantangan dan teknik-teknik serta pendekatan yang digunakan untuk menyelesaikan permasalahan dalam otomatisasi grafologi.

b. Persiapan Data Tulisan Tangan

Pada penelitian ini, peneliti menggunakan *database* tulisan tangan publik dari *IAM- handwriting Database*. Setelah melakukan pengujian model dengan menggunakan database IAM, peneliti melanjutkan dengan menggunakan database private yang diambil dari responden.

c. Pemrosesan Awal

Desain pemrosesan awal dalam penelitian ini menggunakan pendekatan *image processing* antara lain: (i) *resizing* dan *cropping* (digunakan untuk menyamakan ukuran citra dan membuang atribut-



atribut yang tidak dibutuhkan dari sampel teks tulisan tangan), (ii) *smoothing* dengan filterisasi (*bilateral filtering*), (iii) segmentasi (*projection histogram*) untuk memisahkan setiap baris teks, (iv) binerisasi (*thresholding*) yang akan untuk mengekstraksi fitur tekanan pena, (v) normalisasi (*dilation, contour, dan affine transformation*) untuk mengekstraksi baris dasar, ukuran huruf, jarak spasi antar baris, jarak spasi antar kata, serta margin atas, dan *deslanting technique* untuk mengekstraksi fitur sudut kemiringan.

d. Ekstraksi fitur

Dalam penelitian ini, penulis menggunakan 7 fitur tulisan tangan dengan pendekatan Jerman, yaitu: baris dasar (*baseline*), margin atas (*top margin*), jarak spasi antar baris (*space between lines*), jarak spasi antar kata (*space between words*), ukuran huruf (*letter size*), sudut kemiringan (*slant*), dan tekanan Pena (*pen pressure*).

e. Klasifikasi

Klasifikasi fitur tulisan tangan dibagi menjadi 2 (dua) tahapan: tahapan pertama yaitu: dengan menggunakan *decision rule* untuk membagi fitur-fitur menjadi beberapa kelas sesuai dengan kaidah-kaidah *grafologi*, kemudian yang kedua yaitu: dilakukan *mapping* terhadap metode berdasarkan psikologi kepribadian *Big Five Model* (FFM). Setelah dilakukan *mapping* psikologi kepribadian, tahapan selanjutnya dilakukan pembelajaran dengan menggunakan beberapa pendekatan pembelajaran mesin.

f. Pengukuran Performansi

Pembelajaran yang dilakukan oleh mesin terhadap sistem yang dibangun harus dilakukan pengujian performansi. Dalam pengujian dan validasi sistem, penulis menggunakan *confusion matrix* dan *classification report*.

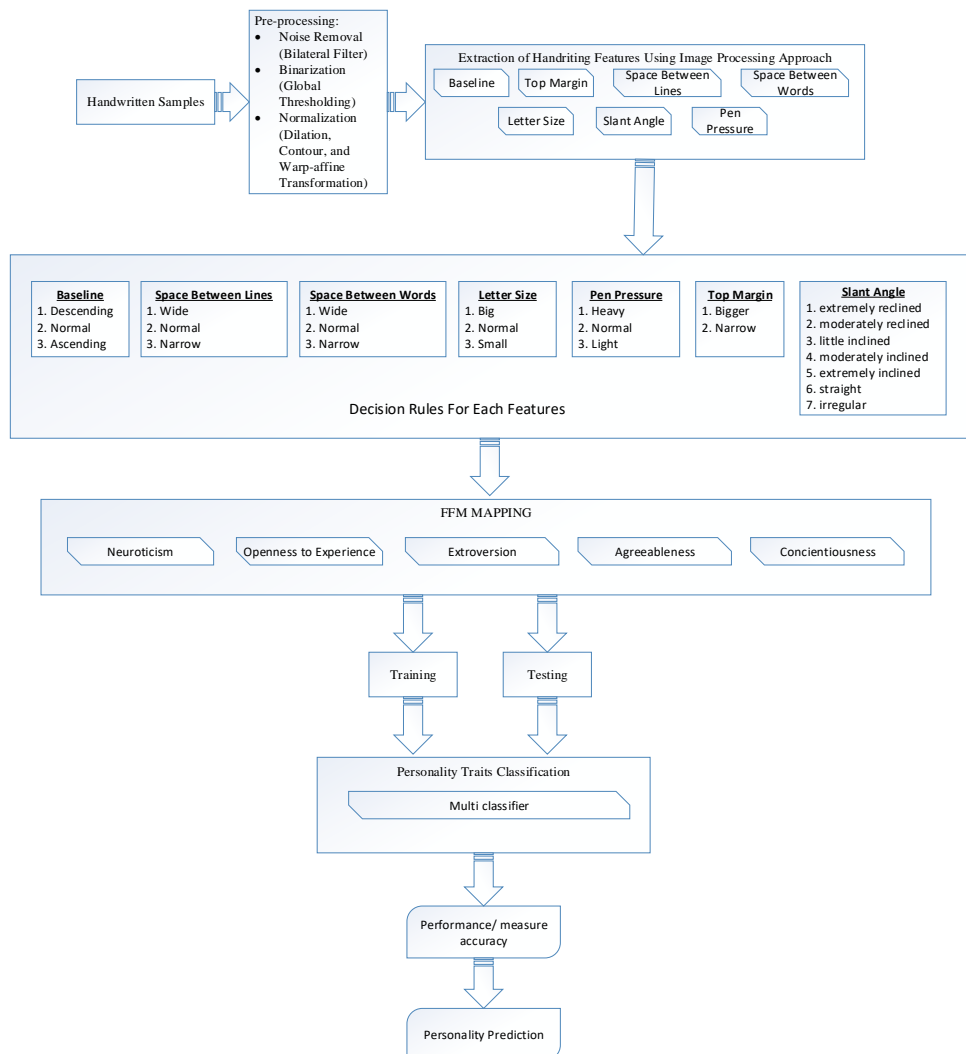
g. Penyusunan Laporan

Laporan penelitian menjadi bagian penting dalam penelitian ini, selain merupakan salah satu bentuk dokumentasi, laporan penelitian

merupakan bentuk pertanggungjawaban peneliti. Laporan penelitian ini juga menjadi sumber untuk menyusun publikasi. Terkait dengan publikasi, pada penelitian ini penulis merencanakan 1 (satu) publikasi melalui jurnal bereputasi internasional dan pencatatan HKI sederhana.

## B. Desain Penelitian

Agar penelitian ini dapat dilaksanakan secara sistematis sehingga tujuan penelitian dapat dicapai dan penelitian dapat terselesaikan tepat waktu, maka berdasarkan kerangka penelitian, peneliti mendesain suatu kerangka kerja (*framework*) berdasarkan sistem otomatisasi grafologi (Gambar 4).



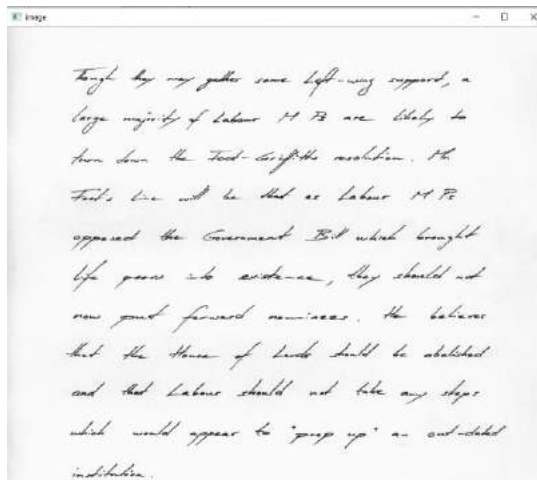
Gambar 4. Framework otomatisasi grafologi yang diusulkan

## IX. Hasil dan Pembahasan

Pada Subbab ini, disajikan hasil pengujian yang meliputi hasil ekstraksi fitur-fitur tulisan tangan dan uji performansi dari pemodelan yang diusulkan.

### A. Hasil Pengujian Ekstraksi Tulisan Tangan

Pengujian ini dilakukan untuk melihat hasil yang diekstraksi dari setiap fitur tulisan tangan yang dihasilkan melalui proses pengolahan citra. Citra tulisan tangan diinputkan kedalam sistem dengan perintah (hasil pengujian yang dibahas menggunakan contoh citra pada gambar 5):



**Gambar 5.** Citra Tulisan Tangan Yang Diuji

#### a) Pengujian dan Hasil Ekstraksi Baris Dasar (*Baseline*)

Pengujian dilakukan dengan membandingkan sudut yang dihasilkan dengan menggunakan transformasi *affine* pada setiap baris teks. Tabel 4 menyajikan hasil ekstraksi baris dasar pada masing-masing baris teks.

**Tabel 4.** Hasil Pengujian Terhadap Fitur Baris Dasar

Hasil Pengolahan Citra	Kategori	Raw
Teks baris ke-1	Menaik	0.687 <sup>0</sup>
Teks baris ke-2	Menaik	0.403 <sup>0</sup>
Teks baris ke-3	Menurun	-0.158 <sup>0</sup>
Teks baris ke-4	Normal	0.0 <sup>0</sup>
Teks baris ke-5	Normal	0.0 <sup>0</sup>
Teks baris ke-6	Normal	0.0 <sup>0</sup>
Teks baris ke-7	Normal	0.0 <sup>0</sup>
Teks baris ke-8	Menurun	-0.504 <sup>0</sup>
Teks baris ke-9	Normal	0.0 <sup>0</sup>

Teks baris ke-10	Menaik	0.492 <sup>0</sup>
<b>Kategori (Nilai rata-rata)</b>	<b>Normal</b>	<b>0.09<sup>0</sup></b>

Dari tabel 4 diatas, rata-rata sudut kemiringan di dalam dokumen tulisan tangan tersebut memiliki sudut 0.09<sup>0</sup> dan dikategorikan kedalam tulisan dengan baris dasar yang tergolong normal (-0.2<sup>0</sup> - 0.2<sup>0</sup>).

**b) Pengujian dan Hasil Ekstraksi Ukuran Huruf (*Letter Size*)**

Pengujian dilakukan dengan membandingkan besaran pixel yang dihasilkan pada masing-masing tinggi baris teks di zona tengah. Dari hasil pemindaian indeks *hplist* yang dalam *array python*, jumlah pixel yang mengandung ukuran huruf (zona tengah) sebesar 85 pixel yang tersebar dalam 10 baris teks. Tabel 5 menyajikan hasil ekstraksi ukuran huruf pada masing-masing baris teks.

**Tabel 5.** Hasil Pengujian Terhadap Fitur Ukuran Huruf

<b>Hasil Pengolahan Citra</b>	<b>Tinggi</b>	<b>Kategori</b>
Teks baris ke-1	8	Kecil (<12px)
Teks baris ke-2	9	Kecil (<12px)
Teks baris ke-3	9	Kecil (<12px)
Teks baris ke-4	8	Kecil (<12px)
Teks baris ke-5	9	Kecil (<12px)
Teks baris ke-6	7	Kecil (<12px)
Teks baris ke-7	7	Kecil (<12px)
Teks baris ke-8	9	Kecil (<12px)
Teks baris ke-9	9	Kecil (<12px)
Teks baris ke-10	10	Kecil (<12px)
<b>Kategori (Nilai rata-rata)</b>	<b>8.5</b>	<b>Kecil (&lt;12px)</b>

Dari hasil tabel 5 diatas, rata-rata ukuran baris teks di dalam dokumen tulisan tangan tersebut memiliki tinggi 8.5 pixel dan dikategorikan kedalam tulisan dengan ukuran huruf kecil (< dari 12 pixel atau setara dengan 2.24 mm).

**c) Pengujian dan Hasil Ekstraksi Jarak Spasi Antar Baris**

Spasi didapat dengan mengetahui banyaknya pixel yang ada di antara teks (zona putih (0)). Zona-zona ini meliputi: spasi pada teks yang paling atas, spasi di antara teks (*blank space*), dan spasi pada teks yang paling bawah yang disebut dengan spasi nol (*space zero*) serta zona atas dan zona bawah (zona yang tidak termasuk *midzone* karena mengandung ukuran huruf) yang disebut dengan spasi tidak nol (*space non zero*). Total dari keseluruhan spasi kemudian dibandingkan dengan ukuran huruf maka inilah nilai dari besarnya jarak spasi. Tabel 6 menyajikan hasil pengujian ekstraksi jarak spasi antar baris pada masing-masing baris teks.

**Tabel 6.** Hasil Pengujian Terhadap Fitur Spasi Antar Baris

<b>Spasi yang diukur</b>	<b><i>Space Zero (pixel)</i></b>	<b><i>Space Non Zero (pixel)</i></b>
Spasi atas teks	62	
Spasi antara baris 1 dan 2	14	
Spasi antara baris 2 dan 3	20	
Spasi antara baris 3 dan 4	19	
Spasi antara baris 4 dan 5	37	
Spasi antara baris 5 dan 6	16	
Spasi antara baris 6 dan 7	19	
Spasi antara baris 7 dan 8	14	
Spasi antara baris 8 dan 9	15	
Spasi antara baris 9 dan 10	30	
Spasi bawah teks	56	
Teks baris ke-1		21
Teks baris ke-2		35
Teks baris ke-3		34
Teks baris ke-4		17
Teks baris ke-5		30
Teks baris ke-6		43
Teks baris ke-7		38
Teks baris ke-8		35
Teks baris ke-9		30
Teks baris ke-10		30
<b>Total</b>	<b>302</b>	<b>313</b>

Dari tabel 6 didapat rata-rata spasi antar baris =  $(302+313)/10 = 61.5$ . Untuk mendapatkan nilai jarak spasi antar baris yang relatif terhadap

ukuran huruf, maka didapat jarak spasi antar baris =  $61.5/8.5 = 7.24$  x ukuran huruf (kategori *wide line spacing*).

**d) Pengujian dan Hasil Ekstraksi Fitur Jarak Antar Kata**

Pengujian untuk mengukur jarak spasi antar kata dimulai dengan melakukan pemindaian proyeksi vertikal pada masing-masing baris teks. Dari hasil pengujian, rata-rata jarak spasi antar kata yaitu perbandingan *pixel* pada ruang kosong (bernilai 0) dengan banyaknya kolom yang ada. Tabel 7 menampilkan hasil pengujian dari ekstraksi fitur jarak antara kata.

**Tabel 7.** Hasil Pengujian Terhadap Fitur Spasi Antar Kata

<b>Spasi yang diukur</b>	<b>Jarak Spasi (Pixel)</b>	<b>Jumlah Kolom</b>
Teks baris ke-1	21, 19, 13, 21, 21, 7, 22, 21	8
Teks baris ke-2	24, 9, 6, 27, 18, 26, 29, 27	8
Teks baris ke-3	21, 35, 25, 10, 25, 13, 23	7
Teks baris ke-4	29, 27, 23, 30, 15, 29, 7, 32, 9	9
Teks baris ke-5	6, 31, 29, 18, 11, 18	6
Teks baris ke-6	12, 36, 37, 5, 23, 21, 7	7
Teks baris ke-7	15, 6, 21, 5, 16, 39, 38	7
Teks baris ke-8	31, 33, 18, 18, 17, 23, 26	7
Teks baris ke-9	27, 21, 25, 26, 28, 23, 21	7
Teks baris ke-10	34, 9, 26, 31, 18, 13, 23, 33	8
<b>Total</b>	<b>1582</b>	<b>74</b>

Dari tabel diatas, maka didapat rata-rata spasi antar kata =  $1582/74 = 21.37$ . Untuk mendapatkan nilai jarak spasi antar kata yang relatif terhadap ukuran huruf, maka didapat jarak spasi antar baris =  $21.37/8.5 = 2.52$  x ukuran huruf (kategori *wide word spacing*).

**e) Pengujian dan Hasil Ekstraksi Fitur Margin Atas (Top Margin)**

Pengujian untuk fitur margin atas hanya melakukan pemindaian proyeksi horizontal terhadap bagian paling atas teks dokumen yang bernilai murni 0. Nilai pixel hasil pemindaian merupakan tinggi dari margin atas. Nilai margin atas =  $62/8.5 = 7.29$  x ukuran huruf (kategori *wide top margin*).

**Tabel 8.** Hasil Pengujian Terhadap Fitur Margin Atas

<b>Spasi yang diukur</b>	<b>Jarak Spasi (Pixel)</b>	<b>Jumlah Kolom</b>
Top Margin (Baris '0')	62	1

**f) Pengujian dan Hasil Ekstraksi Fitur Kemiringan (*Slant*)**

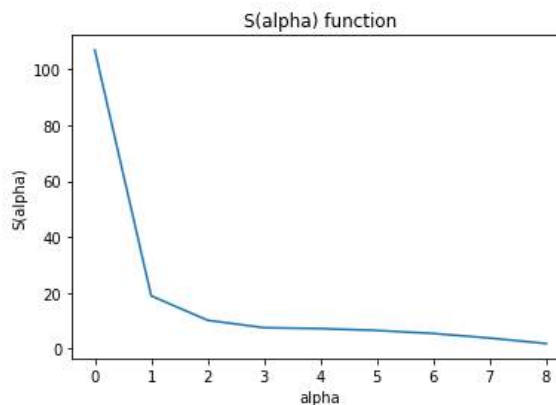
Pengujian untuk fitur kemiringan huruf dilakukan sesuai dengan penjelasan yang ada di penelitian. Tahapan-tahapan pengujian sebagai berikut:

- Menentukan setiap sudut ( $\alpha$ ) (-45, -30, -15, -5, 0, 5, 15, 30, 45 derajat), yang akan dilakukan normalisasi, sebelumnya nilai dari sudut-sudut ( $\alpha$ ) yang diinginkan dikonversi ke dalam bentuk radian.
- Melakukan pengindeksasian yang sesuai dengan nilai terbesar akan menjadi indeks sudut yang paling mungkin di ( $\alpha$ ).
- Melakukan penentuan titik koordinat suatu kata untuk menghitung tinggi ( $h$ ) dan lebar ( $w$ ) dari suatu kata, kemudian dilakukan pergeseran titik dalam kata tersebut dengan menggunakan transformasi *affine*.
- Menentukan jumlah piksel teks pada latar depan ( $jumlah\_piksel\_teks$ ),  $jumlah\_piksel\_teks$  dikatakan sebuah goresan (*stroke*)  $\geq 2/3$   $jumlah\_piksel\_teks$ .
- Lakukan perhitungan untuk menemukan jarak piksel paling atas dan paling bawah pada setiap kolom untuk menentukan  $\Delta y$ . Setelah menemukan  $\Delta y$  maka dapat ditentukan nilai  $h_{\alpha}(i) = (jumlah\_piksel\_teks)/\Delta y^2$ . Nilai akhir pada  $S(\alpha)$  mengindikasikan korelasi sudut yang tepat untuk menentukan fitur kemiringan dari suatu kata. Tabel 9 menampilkan hasil dari pengujian terhadap sudut kemiringan ( $\alpha$ ).

**Tabel 9.** Hasil Pengujian Terhadap Fitur Sudut Kemiringan

$\alpha$ (°)	$S(\alpha)$
45	<b>107.016</b>
30	18.948
15	10.194
5	7.539
0	7.211
-5	6.548
-15	5.467
-30	3.827
-45	1.856

Tabel 9 menunjukkan nilai terbesar dari  $S(\alpha)$  pada sudut pergeseran  $45^\circ$  (masuk kategori “sangat condong ke kanan”). Sebaran dari nilai  $S(\alpha)$  menunjukkan puncak tertinggi yang ada pada setiap citra *deslanted* pada masing-masing sudut yang sesuai.



**Gambar 6.** Plot Sebaran Nilai  $S(\alpha)$  terhadap  $\alpha$

**g) Pengujian dan Hasil Ekstraksi Fitur Tekanan Pena**

Pengujian fitur tekanan pena dilakukan dengan mengukur rata-rata intensitas *pixel* dengan menambahkan semua *pixel* bukan nol (intensitas *pixel* pada baris teks) dibagi dengan jumlah *pixel* yang dihitung setelah dilakukan proses binerisasi (Pers. 1).

$$\text{Intensitas Rata – rata} = \frac{\text{Intensitas Pixel}}{\text{Jumlah pixel}} \quad (1)$$

Persamaan (1) merupakan hasil hitung untuk masing-masing variabel disajikan dalam Tabel 10.

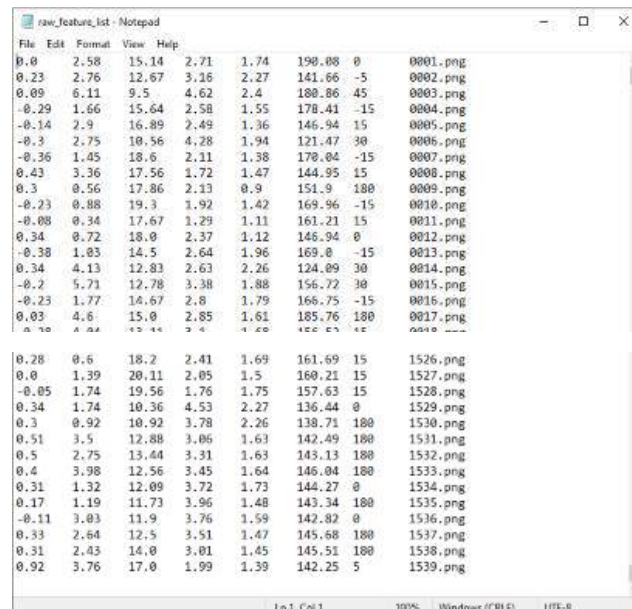


**Tabel 10.** Hasil Pengujian Terhadap Fitur Tekanan Pena

Nilai yang diuji	Total Intensitas	Jumlah Pixel	Intensitas Rata-rata
Rata-rata Intensitas <i>pixel</i> dalam teks	3104344	17164	<b>180.86</b>

Tabel 10 memperlihatkan nilai fitur tekanan pena (rata-rata intensitas *pixel* dalam teks) sebesar 180.86 (dikategorikan dengan tekanan yang berat atau *heavy pen pressure*).

Proses di atas dilakukan untuk semua *database* tulisan tangan yang ada di dokumen *IAM-Database*, kemudian hasil ekstraksi disimpan ke dalam satu *file* berformat *.file*. Proses ekstraksi berjalan kurang lebih selama 4 jam untuk keseluruhan dokumen citra yang ada di IAM. Gambar 7 menyajikan *file* hasil ekstraksi fitur-fitur.



**Gambar 7.** File Hasil Ekstraksi Fitur

Selanjutnya, dilakukan pengkategorian fitur-fitur untuk masing-masing dokumen citra tulisan tangan tersebut. Hasil kategori ini disimpan kembali ke dalam satu *file* berformat *.file*. Gambar 8 menyajikan *file* hasil pengkategorian fitur-fitur hasil ekstraksi.

File	Edit	Format	View	Help				
2	0	2	2	2	0	5	0001.png	
0	0	1	2	0	1	1	0002.png	
2	0	1	0	0	0	4	0003.png	
2	1	2	2	2	2	1	0004.png	
2	0	2	2	2	1	2	0005.png	
1	0	1	0	2	1	3	0006.png	
1	1	0	2	2	2	1	0007.png	
0	0	2	1	2	1	2	0008.png	
0	1	2	2	1	2	6	0009.png	
2	1	0	1	2	2	1	0010.png	
2	1	2	1	1	2	2	0011.png	
0	1	0	2	1	1	5	0012.png	
1	1	2	2	2	2	1	0013.png	
0	0	1	2	0	1	3	0014.png	
2	0	1	2	2	2	3	0015.png	
2	0	2	2	2	2	1	0016.png	
0	1	0	2	2	2	2	1526.png	
2	1	0	2	2	2	2	1527.png	
2	0	0	1	2	2	2	1528.png	
0	0	1	0	0	1	5	1529.png	
0	1	1	0	0	1	6	1530.png	
0	0	1	2	2	1	6	1531.png	
0	0	2	2	2	1	6	1532.png	
0	0	1	2	2	1	6	1533.png	
0	1	1	0	2	1	5	1534.png	
2	1	1	0	2	1	6	1535.png	
2	0	1	0	2	1	5	1536.png	
0	0	1	0	2	1	6	1537.png	
0	0	2	2	2	1	6	1538.png	
0	0	2	1	2	1	2	1539.png	

**Gambar 8.** File Hasil Kategori Fitur

Tahapan terakhir adalah melakukan pemetaan masing-masing model yang ada di *Big Five Model* (FFM) berdasarkan fitur-fitur tulisan tangan. Hasil pemetaan ini disimpan kembali ke dalam satu *file* berformat *.file*. *File mapping* FFM ini menjadi *file* data label untuk masing-masing dokumen citra tulisan tangan dan akan dilakukan pembelajaran mesin untuk pengujian performansi. Gambar 9 menyajikan *file* hasil mapping FFM berdasarkan fitur-fitur hasil ekstraksi.

File	Edit	Format	View	Help										
0.0	0.0	2.0	2.0	2.0	0.0	5.0	0	1	1	1	1	0001.png		
0.0	0.0	1.0	2.0	0.0	1.0	1.0	1	0	0	1	1	0002.png		
2.0	0.0	1.0	0.0	0.0	0.0	4.0	0	0	1	1	1	0003.png		
2.0	1.0	2.0	2.0	2.0	2.0	1.0	0	1	0	1	1	0004.png		
2.0	0.0	2.0	2.0	2.0	1.0	2.0	0	1	0	1	1	0005.png		
1.0	0.0	1.0	0.0	2.0	1.0	3.0	1	1	1	1	1	0006.png		
1.0	1.0	0.0	2.0	2.0	2.0	1.0	0	1	1	1	0	0007.png		
0.0	0.0	2.0	1.0	2.0	1.0	2.0	1	1	0	1	0	0008.png		
0.0	1.0	2.0	2.0	1.0	2.0	6.0	1	0	0	0	0	0009.png		
2.0	1.0	0.0	1.0	2.0	2.0	1.0	0	1	1	1	1	0010.png		
2.0	1.0	2.0	1.0	1.0	2.0	2.0	0	1	0	0	1	0011.png		
0.0	1.0	0.0	2.0	1.0	1.0	5.0	1	0	1	1	0	0012.png		
1.0	1.0	2.0	2.0	2.0	2.0	1.0	0	1	1	1	0	0013.png		
0.0	0.0	1.0	2.0	0.0	1.0	1.0	1	1	0	1	1	0014.png		
2.0	0.0	1.0	2.0	2.0	2.0	1.0	1	1	0	1	1	0015.png		
2.0	0.0	1.0	2.0	2.0	2.0	1.0	0	1	0	1	1	0016.png		
2.0	0.0	2.0	2.0	2.0	0.0	6.0	0	1	1	1	1	0017.png		
2.0	0.0	2.0	2.0	2.0	2.0	2.0	0	1	0	1	1	0018.png		
1.0	1.0	0.0	1.0	2.0	2.0	6.0	0	1	1	0	0	0019.png		
2.0	1.0	2.0	1.0	2.0	2.0	2.0	0	1	0	0	1	1521.png		
0.0	1.0	0.0	1.0	2.0	2.0	2.0	1	1	1	0	0	1522.png		
0.0	1.0	0.0	1.0	2.0	2.0	2.0	1	1	1	0	0	1523.png		
0.0	0.0	2.0	1.0	2.0	2.0	2.0	1	1	0	1	0	1524.png		
0.0	1.0	2.0	2.0	2.0	2.0	2.0	1	1	0	0	0	1525.png		
0.0	1.0	0.0	2.0	2.0	2.0	2.0	1	1	1	1	0	1526.png		
2.0	1.0	0.0	2.0	2.0	2.0	2.0	0	1	1	0	1	1527.png		
2.0	0.0	0.0	1.0	2.0	2.0	2.0	0	1	1	1	1	1528.png		
0.0	0.0	1.0	0.0	0.0	1.0	5.0	1	0	0	1	1	1529.png		
0.0	1.0	1.0	0.0	0.0	1.0	6.0	1	0	0	1	1	1530.png		
0.0	0.0	1.0	2.0	2.0	1.0	6.0	1	1	0	1	1	1531.png		
0.0	0.0	2.0	2.0	2.0	1.0	6.0	1	1	0	1	0	1532.png		
0.0	0.0	1.0	2.0	2.0	1.0	6.0	1	1	0	1	1	1533.png		
0.0	1.0	1.0	0.0	2.0	1.0	5.0	1	1	0	1	1	1534.png		
2.0	1.0	1.0	0.0	2.0	1.0	6.0	0	1	0	1	1	1535.png		
2.0	0.0	1.0	0.0	2.0	1.0	5.0	0	1	0	1	1	1536.png		
0.0	0.0	1.0	0.0	2.0	1.0	6.0	1	1	0	1	1	1537.png		
0.0	0.0	2.0	2.0	2.0	1.0	6.0	1	1	0	1	0	1538.png		
0.0	0.0	2.0	1.0	2.0	1.0	2.0	1	1	0	1	0	1539.png		

Gambar 9. File Hasil Mapping FFM Terhadap Hasil Ekstraksi Fitur

## B. Uji Performansi

Setelah melakukan pengujian ekstraksi fitur-fitur, tahapan pengujian selanjutnya adalah melakukan uji performansi dengan pembelajaran mesin (*machine learning*). Tolak ukur uji performansi yang digunakan yaitu: waktu proses, akurasi, komposisi data pelatihan dan pengujian, *confusion matrix*, dan *classification report*. Ada tiga variasi komposisi data yang dilakukan pengujian, yaitu 50:50, 35:65, dan 20:80. Pengujian dilakukan untuk kelima model yang ada di psikologi kepribadian *Big Five Model* (FFM) dengan menggunakan SVM (kernel RBF, kernel *linear*, kernel *polinomial*), K-NN, dan *decision tree* (DT).

Tabel 11 menyajikan *Summary* dari hasil pengujian yang dilakukan.

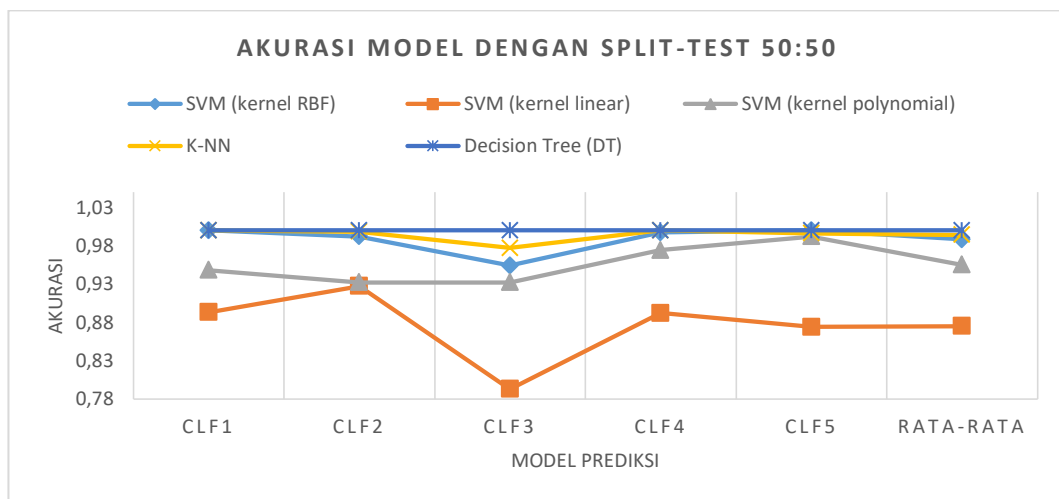
**Tabel 11. Summary Hasil Uji Performansi**

Note: clf1 (Neuroticism), clf2 (Openness to Experience), clf3 (extroversion), clf4 (Agreeableness), clf5 (Conscientiousness)

<b>Split-Test (50:50)</b>							
<b>Classifiers</b>	<b>Waktu Prediksi</b>	<b>Big Five Model (FFM)</b>					<b>Rata-rata</b>
		<b>Clf1</b>	<b>Clf2</b>	<b>Clf3</b>	<b>Clf4</b>	<b>Clf5</b>	
SVM (kernel RBF)	0.561 s	1.00	0.992	0.954	0.997	1.00	0.988
SVM (kernel linear)	0.572 s	0.893	0.927	0.793	0.892	0.874	0.875
SVM (kernel polinomial)	1.947 s	0.948	0.932	0.932	0.974	0.992	0.955
K-NN	36.82 s	1.00	0.998	0.977	1.00	0.996	0.994
Decision Tree (DT)	<b>0.507 s</b>	1.00	1.00	1.00	1.00	1.00	<b>1.00</b>
<b>Split Test (35:65)</b>							
SVM (kernel RBF)	0.551 s	1.00	0.998	0.992	1.00	1.00	0.998
SVM (kernel linear)	0.562 s	0.846	0.918	0.807	0.896	0.870	0.867
SVM (kernel polinomial)	1.920 s	0.953	0.920	0.931	0.981	0.990	0.955
K-NN	36.08 s	1.00	1.00	0.994	1.00	0.992	0.997
Decision Tree (DT)	<b>0.480 s</b>	1.00	1.00	1.00	1.00	1.00	<b>1.00</b>
<b>Split Test (20:80)</b>							
SVM (kernel RBF)	0.534 s	1.00	0.996	0.990	1.00	1.00	0.997
SVM (kernel linear)	0.652 s	0.873	0.909	0.811	0.899	0.853	0.869
SVM (kernel polinomial)	1.810 s	0.941	0.912	0.938	0.990	0.990	0.954
K-NN	32.98 s	1.00	1.00	0.996	1.00	0.993	0.997
Decision Tree (DT)	<b>0.430 s</b>	1.00	1.00	1.00	1.00	1.00	<b>1.00</b>

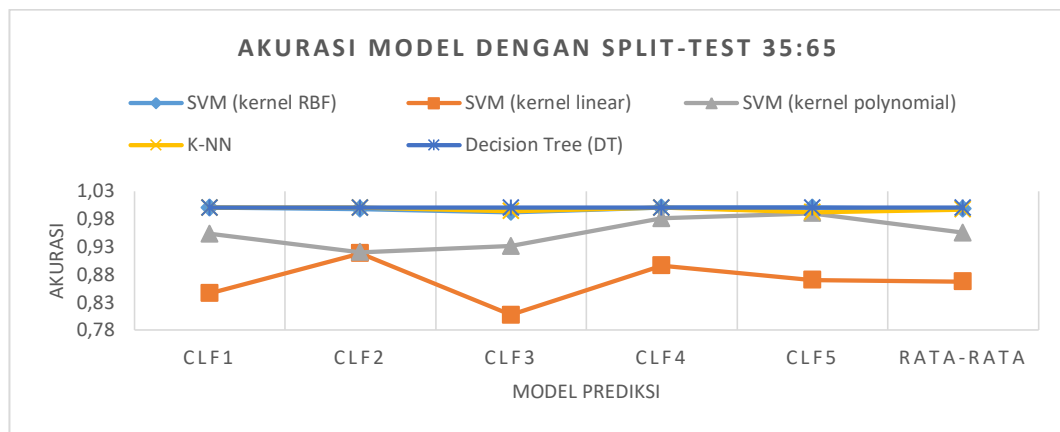
Tabel 11 memperlihatkan hasil eksperimen dimulai dengan mengklasifikasi kelima model dengan menggunakan prosedur *split test* 50 % pengujian dan 50 % pelatihan. Pertama-tama, dilakukan klasifikasi model dengan metode SVM. Untuk menemukan hasil pembelajaran yang maksimal dengan mengadopsi *kernel trick*. Beberapa kernel dilakukan uji coba yaitu kernel RBF, linear, dan polinomial. Hasil klasifikasi maksimal SVM didapat dengan kernel RBF memiliki akurasi rata-rata mencapai 99.8 %. Hasil ini jauh mengungguli dua kernel lainnya dengan hasil akurasi rata-rata 87.5 % untuk kernel *linear* dan 95.5

% untuk kernel polinomial. Waktu yang dibutuhkan untuk memprediksi SVM dengan kernel RBF lebih efisien dengan waktu sekitar 0.561 detik dan lebih cepat dibandingkan dengan kedua kernel lainnya, yaitu kernel linear dengan waktu konsumsi selama 0.572 detik dan kernel polinomial dengan waktu konsumsi selama 1.947 detik. Dari data yang dihasilkan, penulis berkesimpulan bahwa SVM dengan kernel RBF dalam hal akurasi dan waktu proses prediksi mengungguli kedua kernel lainnya. Selanjutnya untuk melihat hasil klasifikasi yang berbeda, penulis melakukan eksperimen dengan menggunakan K-NN *classifier*. Hasil yang didapat dengan menggunakan K-NN cukup memuaskan, yaitu sebesar 99.4 %. Yang menjadi persoalan adalah waktu konsumsi yang sangat tinggi dalam melakukan proses prediksi yaitu sekitar 36.82 detik. Hal ini disebabkan oleh K-NN melakukan pencarian untuk nilai K yang tepat sehingga membutuhkan banyak proses. Semakin banyak nilai K yang diuji, maka semakin besar waktu yang dibutuhkan untuk melakukan proses. Kemudian penulis juga melakukan uji coba dengan menggunakan metode klasifikasi *decision tree*. Hasil yang diperoleh sangat memuaskan dimana *decision tree* dapat mengklasifikasi model dengan sempurna untuk kelima model dengan akurasi sebesar 100 % dan waktu proses yang lebih cepat dari SVM maupun K-NN dengan proses konsumsi waktu prediksi selama 0.507 detik. Gambar 10 menyajikan grafik akurasi model pada masing-masing *classifiers* yang dilakukan uji coba dengan *split-test* 50:50.



**Gambar 10.** Grafik Perbandingan Akurasi *Classifiers* (Split-Test 50:50)

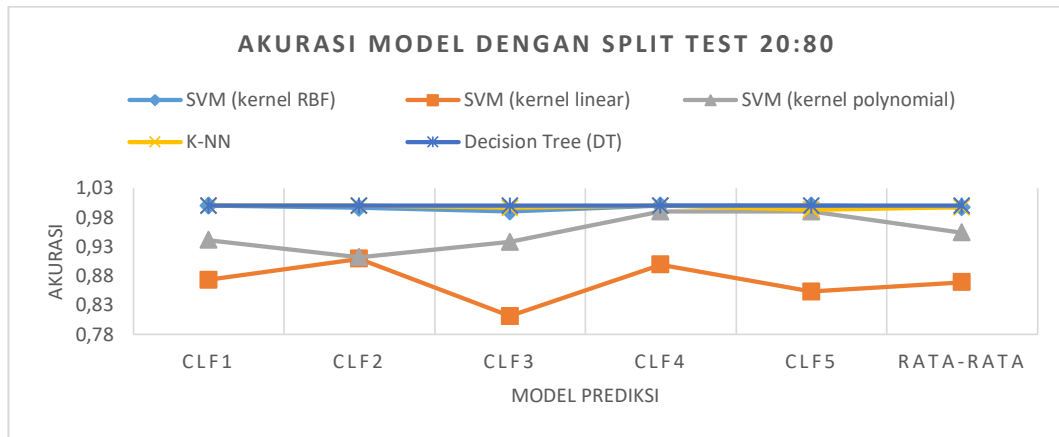
Untuk skenario eksperimen yang kedua, yaitu *split-test* terhadap data 35 % pengujian dan 65 % pelatihan, tidak ditemukan pengaruh yang signifikan terhadap hasil akurasi pada masing-masing metode klasifikasi yang diuji. Hanya SVM dengan kernel RBF yang mengalami peningkatan sebesar 1% menjadi 99.8 % dengan waktu prediksi yang relatif sama (selisih 0.01 detik lebih cepat), SVM kernel linear dengan akurasi sekitar 86.7 % (turun sebesar 0.8 % dengan selisih waktu 0.01 detik lebih cepat), SVM kernel polinomial dengan akurasi sekitar 95.5 % (akurasi tetap dengan selisih waktu 0.027 detik lebih cepat), KNN dengan akurasi sebesar 99.7 % (turun 0.1 % dengan selisih waktu 0.74 detik lebih cepat), dan *decision tree* dengan akurasi tetap 100 % (selisih waktu 0.027 detik lebih cepat). Gambar 11 menyajikan grafik akurasi model pada masing-masing *classifiers* yang dilakukan uji coba dengan *split-test* 35:65.



Gambar 11. Grafik Perbandingan Akurasi Classifiers (Split-Test 35:65)

Untuk skenario eksperimen yang ketiga, yaitu *split-test* terhadap data 20 % pengujian dan 65 % pelatihan, hasil akurasi yang didapat bisa dikatakan sama dengan hasil pada skenario kedua, hanya waktu proses konsumsi prediksi yang lebih cepat. SVM kernel RBF dengan akurasi sebesar 99.7 % (turun sebesar 0.1 % dengan selisih waktu 0.017 lebih cepat), SVM kernel linear dengan akurasi 86.9 % (naik 0.2 % dengan selisih waktu 0.09 detik lebih lambat), SVM kernel polinomial dengan akurasi 95.4 % (turun 0.1 % dengan selisih waktu 0.11 detik lebih cepat), K-NN dengan akurasi sama sebesar 99.7 % (selisih 3.1 detik lebih cepat), dan *decision tree* dengan akurasi 100 % (selisih waktu 0.05 detik lebih

cepat). Gambar 12 menyajikan grafik akurasi model pada masing-masing *classifiers* yang dilakukan uji coba dengan *split-test* 20:80.



Gambar 12. Grafik Perbandingan Akurasi Classifiers (Split-Test 20:80)

Pengujian menggunakan Arsitektur CNN:

Pengujian meliputi arsitektur sebagai berikut:

Tabel 12. Arsitektur CNN yang digunakan

	Input : 64x64
1	Conv3-32, relu
	MaxPooling 2x2
2	Conv3-64, relu
3	Conv3-64, relu
	MaxPooling 2x2
4	Conv3-128, relu
5	Conv3-128, relu
6	Conv3-128, relu
	MaxPooling 2x2
7	FC-512
8	Sigmoid

Tahapan *pre-processing* meliputi:

1. Melakukan resize pada setiap image ke ukuran 64x64;
2. Kemudian melakukan greyscaling (derajat keabuan) pada masing-masing image;

- Melakukan normalisasi dengan menggunakan proses binerisasi dimana foreground layer (tulisan tangan) ber-value =1, dan background layer bervalue =0;

Split dataset yang digunakan antara lain 80 % training dan 20 % testing, dengan proses hyperparameter antara lain:

- Epoch : 10
- Batch size : 32
- Learning rate : 0,001

Dari masing-masing sifat, berikut didapat akurasi antara lain:

Tabel 13. Hasil pengujian pada masing-masing sifat:

	Sifat	Akurasi (Proposed System)	CNN dengan Arsitektur PersonaNet [16]
1	Agreeableness	90%	75 %
2	Conscientiousness	75%	65 %
3	Extroversion	58%	66 %
4	Neuroticism	54%	60 %
5	Openness to Experience	75%	59 %
	<b>Rata-rata</b>	<b>70.4 %</b>	<b>65 %</b>

Hasil analisis menggunakan CNN:

Dari hasil yang didapatkan serta dibandingkan dengan penelitian serupa menggunakan arsitektur yang berbeda, dapat dikatakan lebih baik dengan rata-rata akurasi diatas 70 % (dibandingkan dengan penelitian serupa dengan menggunakan arsitektur personaNet menghasilkan rata-rata 65 %).

Testing dengan data agreeableness terlihat lebih baik daripada yang baik karena dapat membedakan yang mana tulisan tangan bersifat agreeableness. Tetapi pada hasil yang lain pada sifat ke 2-5, hasil akurasi tergantung jumlah data mana yang lebih banyak, misalnya: Pada data conscientiousness, jumlah train “detected” lebih banyak dibandingkan dengan data “not detected”, maka model akan selalu mendeteksi seluruh tulisan dengan output “detected”, begitu pula



sebaliknya, apabila data latih “not detected” lebih banyak, maka model akan selalu mendeteksi seluruh tulisan dengan output “not detected”. Anehnya pada data agreeableness jika dilihat juga memiliki imbalance data yaitu, jumlah data detected lebih banyak dibanding not detected, tetapi dapat dikatakan berhasil untuk mendeteksi tulisan lainnya. Padahal perilaku yang diberikan juga sama, hanya berbeda data saja. Pengujian dengan arsitektur lain juga dilakukan yaitu VGG16 dan ResNet50, tetapi hasil menunjukkan hasil yang sama. Pelatihan dengan menggunakan data lain juga telah dilakukan untuk menguji apakah arsitektur yang dibangun sudah benar apa belum, dan hasilnya cukup baik apabila menggunakan data lain (dogsvscats), berarti bisa dikatakan tidak ada kesalahan pada arsitektur.

## X. Kesimpulan dan Saran

### A. Kesimpulan

1. Untuk hasil pengujian sistem, disajikan pada subbab C secara keseluruhan. Pada eksperimen awal ini, penulis telah melakukan pengujian terhadap ekstraksi 7 fitur-fitur tulisan tangan antara lain: baris dasar (*baseline*), margin atas (*top margin*), jarak spasi antar baris (*line spacing*), jarak spasi antar kata (*word spacing*), ukuran huruf (*letter size*), sudut kemiringan (*slant*), dan tekanan pena (*pen pressure*). Berdasarkan eksperimen terhadap ekstraksi fitur, hasil ekstraksi berjalan dengan baik dalam mengkategorikan setiap atribut yang ada pada citra tulisan tangan sesuai dengan yang telah ditetapkan. Ada beberapa catatan yang penulis garis bawahi, khusus untuk ekstraksi fitur tekanan pena akan lebih baik jika pena yang digunakan sama (dalam IAM *handwriting database*, pena yang digunakan beragam), sehingga dapat memprediksi dengan baik sesuai dengan karakteristik penulisnya. Tentunya penggunaan bolpoin dengan pena tinta menghasilkan efek tekanan yang berbeda. Untuk *deslanting technique* juga berjalan dengan baik dimana dapat mengenali sudut yang dihasilkan berdasarkan nilai  $H(\alpha)$  yang maksimum;

2. Dalam hal klasifikasi, metode *decision tree* dan SVM dengan kernel RBF menjadi pilihan yang tepat karena mampu menghasilkan akurasi diatas 99 % dengan waktu proses kurang dari 1 detik (rata-rata 0.5 detik) dengan maksimal data uji sebesar 50 % dari keseluruhan dokumen tulisan tangan yang berjumlah 1539. K-NN juga dapat melakukan klasifikasi dengan baik dibuktikan dengan akurasi diatas 99 %, tetapi dalam hal konsumsi waktu prediksi, menjadi tidak efisien karena membutuhkan waktu yang lebih lama (rata-rata 32-36 detik). Dari hasil awal ini, menunjukkan *framework* yang diusulkan dapat diterapkan dengan baik untuk memprediksi psikologi kepribadian dengan metode FFM.

#### B. Saran

Adanya penelitian mengenai identifikasi kepribadian berdasarkan tulisan tangan ini, untuk penelitian serupa disarankan untuk menggunakan dataset privat dan juga mengembangkan aspek penelitian kearah metode psikologi yang lain, semisal MBTI.

### XI. Daftar Pustaka

- [1] G. Sheikholeslami, S. N. Srihari, and V. Govindaraju, "COMPUTER AIDED GRAPHOLOGY," State University of New York at Buffalo, 1995.
- [2] M. H. Suid, M. F. M. Jusof, and M. A. Ahmad, "Dual sliding statistics switching median filter for the removal of low level random-valued impulse noise," *J. Electr. Eng. Technol.*, vol. 13, no. 3, pp. 1383–1391, 2018.
- [3] B. Karthik, T. Krishna Kumar, S. P. Vijayaragavan, and M. Sriram, "Removal of high density salt and pepper noise in color image through modified cascaded filter," *J. Ambient Intell. Humaniz. Comput.*, pp. 1–8, 2020.
- [4] N. Paul and H. Tunga, "An Improved Method for Document Image Binarization," *Natl. Conf. Recent Innov. Comput. Sci. Commun.*, no. July, pp. 1–6, 2016.
- [5] S. Nagar, S. Chakraborty, A. Sengupta, J. Maji, and R. Saha, "An efficient method for character analysis using space in handwriting image," in *Proceedings - 2016 6th International Symposium on Embedded Computing and System Design, ISED 2016*, 2016, pp. 210–216.

- [6] A. Bal and R. Saha, "An Improved Method for Handwritten Document Analysis using Segmentation , Baseline Recognition and Writing Pressure Detection," in *Procedia Computer Science*, 2016, vol. 93, pp. 403–415.
- [7] P. Joshi, A. Agarwal, A. Dhavale, R. Suryavansi, and S. Kodolikar, "Handwriting Analysis for Detection of Personality Traits using Machine Learning Approach," *Int. J. Comput. Appl.*, vol. 130, no. 15, pp. 40–45, 2015.
- [8] M. Gavrilescu, "3-Layer architecture for determining the personality type from handwriting analysis by combining neural networks and Support Vector Machines," *UPB Sci. Bull. Ser. C Electr. Eng. Comput. Sci.*, vol. 79, no. 4, pp. 135–152, 2017.
- [9] M. Gavrilescu and N. Vizireanu, "Predicting the Big Five personality traits from handwriting," *EURASIP J. Image Video Process.*, 2018.
- [10] S. H. Fatimah, E. C. Djamal, R. Ilyas, and F. Renaldi, "Personality features identification from handwriting using convolutional neural networks," in *4th International Conference on Information Technology, Information Systems and Electrical Engineering, ICITISEE*, 2019, vol. 6, pp. 119–124, doi: 10.1109/ICITISEE48480.2019.9003855.
- [11] K. N. Chaudhury and S. D. Dabhade, "Fast and provably accurate bilateral filtering," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2519–2528, 2016, doi: 10.1109/TIP.2016.2548363.
- [12] C. Bauckhage, "NumPy / SciPy Recipes for Image Processing: Affine Image Warping," *researchgate.net*, no. November, 2018, doi: 10.13140/RG.2.2.29805.44004.
- [13] M. Gavrilescu, "Study on determining the Myers-Briggs personality type based on individual's handwriting," in *The 5th IEEE International Conference on E-Health and Bioengineering - EHB 2015*, 2015.
- [14] D. Pratiwi, G. Budi, and F. Hana, "Personality Type Assessment System by using Enneagram-Graphology Techniques on Digital Handwriting," *Int. J. Comput. Appl.*, vol. 147, no. 11, pp. 9–13, 2016.
- [15] S. Thomas, M. Goel, and D. Agrawal, "A framework for analyzing financial behavior using machine learning classification of personality through handwriting analysis," *J. Behav. Exp. Finance.*, vol. 26, p. 100315, 2020, doi: 10.1016/j.jbef.2020.100315.
- [16] A. A. Elngar, N. Jain, D. Sharma, H. Negi, A. Trehan, and A. Srivastava, "A deep learning based analysis of the big five personality traits from handwriting samples using image processing," *J. Inf. Technol. Manag.*, vol. 12, pp. 3–35, 2021, doi: 10.22059/JITM.2020.78884.