

**PENGENALAN KARAKTER TULISAN TANGAN AKSARA LAMPUNG  
MENGUNAKAN EKSTRAKSI FITUR *MOMENT INVARIANT* DAN  
*PIXEL DENSITY***

**(Skripsi)**

**Oleh**

**REVI NOVITA**



**JURUSAN ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
2019**

## **ABSTRACT**

### **RECOGNITION OF HANDWRITTEN CHARACTER LAMPUNG SCRIPT BY USING MOMENT INVARIANT AND PIXEL DENSITY**

**By**

**REVI NOVITA**

Nowadays many ethnics live together in Lampung Province. They communicate to others by Indonesia language for verbal communication and roman script for written communication. Due to the massive use of Roman script, Lampung script is now neglected. This condition shifts the use of the script to Roman script and may trigger its extinction in the future. Therefore, an effort should be initiated to avoid this calamity. One endeavor to preserve Lampung script is by conducting research on the script. This research is an extension of the Lampung handwritten character recognition research by introducing the features of moment invariant and pixel density. These features are extracted from 32 by 32 pixel size of a binary image. The moment invariant feature is extracted from binary skeleton image while pixel density feature is directly extracted from actual binary image. Then, the support vector machine (SVM) is used to classify images into 18 character classes. The accuracy of moment invariant feature is 49.94%, pixel density feature is 91.83%, and the concatenation of both features is 91.90%. Most

of misclassification occur within the moment invariant feature due to the counter impact of rotation, translation and scaling invariant of the feature.

**Keywords:** *moment invariant, pixel density, handwritten character recognition,*

*Lampung script*

## **ABSTRAK**

### **Pengenalan Karakter Tulisan Tangan Aksara Lampung Menggunakan Ekstraksi Fitur *Moment Invariant* dan *Pixel Density***

**Oleh**

**REVI NOVITA**

Saat ini banyak etnis yang tinggal di provinsi Lampung. Mereka berkomunikasi satu sama lain dengan bahasa Indonesia untuk komunikasi verbal dan aksara latin untuk komunikasi tertulis. Karena penggunaan aksara latin yang intens, Aksara Lampung sekarang diabaikan. Kondisi ini menggeser penggunaan Aksara Lampung ke Aksara Latin dan dapat memicu kepunahan aksara tersebut di masa depan. Oleh karena itu, suatu upaya seharusnya diinisiasi untuk menghindari bencana ini. Salah satu usaha untuk melestarikan aksara ini adalah dengan melakukan penelitian terhadap aksara tersebut. Penelitian ini merupakan pengembangan penelitian pengenalan tulisan tangan Aksara Lampung yang telah dilakukan sebelumnya dengan memperkenalkan fitur-fitur *moment invariant* dan *pixel density*. Fitur-fitur ini diekstraksi dari citra biner berukuran 32 x 32 piksel. Fitur *moment invariant* diekstraksi dari citra biner kerangka (*skeleton*) sedangkan fitur *pixel density* diekstraksi langsung dari citra biner yang utuh. Selanjutnya,

*support vector machine* (SVM) digunakan untuk klasifikasi citra ke dalam 18 kelas aksara. Tingkat akurasi pengenalan dengan fitur *moment invariant* adalah 49,94%, fitur *pixel density* adalah 91.83%, dan fitur gabungan keduanya adalah 91.90%. Banyak kesalahan klasifikasi terjadi pada fitur *moment invariant* akibat dampak negatif dari sifat fitur tersebut yaitu *rotation*, *translation* dan *scaling invariant*.

**Kata kunci :** *moment invariant*, *pixel density*, pengenalan karakter tulisan tangan, aksara Lampung.

**PENGENALAN KARAKTER TULISAN TANGAN AKSARA LAMPUNG  
MENGUNAKAN EKSTRAKSI FITUR *MOMENT INVARIANT* DAN  
*PIXEL DENSITY***

Oleh

**REVI NOVITA**

**Skripsi**

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar  
SARJANA KOMPUTER**

**Pada**

**Jurusan Ilmu Komputer  
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKUTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2019**

Judul Skripsi : **PENGENALAN KARAKTER TULISAN  
TANGAN AKSARA LAMPUNG  
MENGUNAKAN EKSTRAKSI FITUR  
MOMENT INVARIANT DAN PIXEL DENSITY**

Nama Mahasiswa : *Revi Novita*

Nomor Pokok Mahasiswa : 1517051034

Jurusan : Ilmu Komputer

Fakultas : Matematika dan Ilmu Pengetahuan Alam



2. Mengetahui

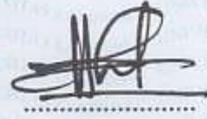
Ketua Jurusan Ilmu Komputer  
FMIPA Universitas Lampung

**Dr. Ir. Kurnia Muludi, M.S.Sc**  
NIP. 19640616 198902 1 001

**MENGESAHKAN**

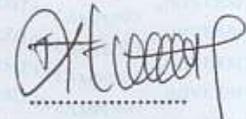
**1. Tim Penguji**

**Ketua : Dr. Rer. Nat. Akmal Junaidi, M.Sc**

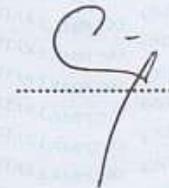


**Penguji**

**Bukan Pembimbing : 1. Tristiyanto, S.Kom., M.I.S., Ph.D**



**2. Rico Andrian, S.Si, M.Kom.,**



**2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam**



**Dr. Suratman, M.Sc.**

**N.P. 19640604 199003 1 002**



**Tanggal Lulus Ujian Skripsi : 30 September 2019**

## PERNYATAAN

Saya yang bertanda tangan di bawah ini, menyatakan bahwa skripsi saya yang berjudul "Pengenalan Karakter Tulisan Tangan Aksara Lampung Menggunakan Ekstraksi Fitur *Moment Invariant* dan *Pixel Density*" merupakan hasil kerja saya sendiri dan bukan karya orang lain. Semua tulisan yang tertuang di skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila dikemudian hari terbukti skripsi saya merupakan hasil penjiplakan atau dibuat orang lain, maka saya bersedia menerima sanksi berupa pencabutan gelar yang telah saya terima.

Bandar Lampung, 30 September 2019



**REVI NOVITA**  
NPM. 1517051034

## RIWAYAT HIDUP



Penulis dilahirkan pada tanggal 11 Desember 1998 di Way Kanan, dengan Ibu bernama Siti Harmini dan Ayah bernama Efendi. Penulis menyelesaikan pendidikan formal pertama kali Sekolah Dasar di SDN Negeri Mulyo pada tahun 2009, menyelesaikan Sekolah Menengah Pertama (SMP) di SMP Negeri 4 Bukit Kemuning pada tahun 2012, kemudian melanjutkan jenjang Sekolah Menengah Atas (SMA) di SMA Negeri 1 Bukit Kemuning dengan mengambil jurusan Ilmu Pengetahuan Alam (IPA) dan lulus pada tahun 2015.

Pada tahun 2015, penulis terdaftar sebagai mahasiswa Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SNMPTN. Selama menjadi mahasiswa, penulis pernah bergabung dalam Badan Eksekutif Mahasiswa (BEM) serta menjadi anggota Bidang Keilmuan (2015&2016) Himpunan Mahasiswa Ilmu Komputer (HIMAKOM). Penulis pernah menjadi asisten praktikum untuk matakuliah Basis Data (2017). Bulan Januari – Februari 2018, penulis melakukan kerja praktik di Kantor Kecamatan Gunung Labuhan dengan penempatan pada Kasi Kesejahteraan Masyarakat. Bulan Juli – Agustus 2018, penulis melakukan Kuliah Kerja Nyata (KKN) di Desa Muara Gading Mas, Kecamatan Labuhan Maringgai, Kabupaten Lampung Timur.

## **PERSEMBAHAN**

*Puji dan syukur kepada Allah SWT atas segala limpahan berkah-Nya sehingga skripsi ini dapat terselesaikan dengan tepat waktu.*

*Teruntuk kedua orang tua tercinta Bapak Efendi dan Ibu Siti Harmini yang senantiasa mengiringi langkah ini dengan do'a dan segala jerih payah demi tercapai segala harapan dan cita-citaku.*

*Kakak dan adik tercinta Rani Ervina, Risa Novera, dan M.Refky, kesayanganku.*

*Keluarga Ilmu Komputer 2015.*

*Serta Almamater tercinta,  
Universitas Lampung.*

## **MOTTO**

*"Man Jadda Wa Jadda- Siapa yang  
bersungguh-sungguh pasti akan berhasil"  
(Al-Hadits)*

*"Lebih baik membuat kesalahan daripada  
tidak melakukan apapun"  
(Seth Godin)*

## SANWACANA

*Alhamdulillahirobbil 'alamin.*

Puji syukur kehadirat Allah SWT, atas limpahan rahmat dan karunia-Nya penulis dapat menyelesaikan penulisan skripsi yang berjudul “Pengenalan Karakter Tulisan Tangan Aksara Lampung Menggunakan Ekstraksi Fitur *Moment Invariant* dan *Pixel Density*” dengan baik. terselesaikannya laporan ini tidak terlepas dari bantuan banyak pihak, sehingga pada kesempatan ini dengan ketulusan hati dan penuh rasa hormat penulis menghaturkan terimakasih kepada semua pihak yang telah membantu dan berperan besar dalam proses penyusunan skripsi ini. Ungkapan terimakasih penulis haturkan kepada :

1. Kedua orangtua tercinta, sumber motivasi terbesar, Bapak Efendi dan Ibu Siti Harmini yang selalu memberikan doa, kasih sayang, dukungan dan semangat yang tak terhingga serta memfasilitasi segala kebutuhan dalam menyelesaikan skripsi ini.
2. Kakak dan adik tersayang, Rani Ervina, Risa Novera dan M.Refky yang telah menjadi penyemangat dalam menyelesaikan perkuliahan selama ini.
3. Bapak dr.rer.nat.Akmal Junaidi, M.Sc sebagai pembimbing utama yang telah memberikan tema penelitian, membimbing, memotivasi serta memberikan kritik dan saran selama masa perkuliahan dan penyusunan skripsi sehingga penulis dapat sampai di tahap ini.
4. Bapak Tristiyanto, S.Kom., M.I.S., Ph.D sebagai pembahas yang telah memberikan kritik dan saran yang bermanfaat untuk perbaikan dalam penyusunan skripsi ini.
5. Bapak Rico Andrian, S.Si., M.Kom sebagai pembahas yang telah memberikan kritik dan saran yang bermanfaat untuk perbaikan dalam penyusunan skripsi ini.
6. Bapak Kurnia Muludi, Dr, Ir, M.S.Sc. selaku Ketua Jurusan Ilmu Komputer Universitas Lampung.
7. Bapak Didik Kurniawan, S.Si, M.T. selaku Sekretaris Jurusan Ilmu Komputer Universitas Lampung.

8. Bapak Febi Eka Febriansyah, S.T selaku Pembimbing Akademik.
9. Bapak dan Ibu Dosen Jurusan Ilmu Komputer yang telah memberikan ilmu dan pengalaman yang luar biasa kepada penulis selama menjadi mahasiswa.
10. Ibu Nora Ade Maela selaku Staff Administrasi Jurusan Ilmu Komputer.
11. Keluarga seataap Eni Ayu Lestari dan Siti Rosdiana yang selalu membagikan kerecehan-kerecehan sebagai hiburan di sela kepenatan dalam menyusun skripsi.
12. Keluarga Lambe Julid Indrianti, Titik Zulaiha dan Siti Rosdiana yang selalu menampung keluh kesah dan memberikan semangat.
13. Keluarga Antimo Indrianti, Zuliana Nurfadlillah, Nur Fitriana, Nadya Wiandhini, Nanda Pertiwi, Wahyuda Makhuta, Ardian Hermawansyah, Dana Febri Setiawan, Inas Firhan, Akmal Jahidi, Jaka Al Akbar dan Nanda Adi Pratama yang telah menjadi teman selama masa perkuliahan serta memberikan semangat dan dukungan selama menyelesaikan skripsi ini.
14. Keluarga Ilmu Komputer 2015 yang tidak bisa disebutkan satu per satu, terima kasih banyak atas kebersamaannya.
15. Almamater tercinta, Universitas Lampung.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, akan tetapi besar harapan semoga skripsi ini dapat bermanfaat bagi perkembangan ilmu pengetahuan terutama bagi teman-teman Ilmu Komputer serta semua pihak yang membacanya.

Bandar Lampung, 30 September 2019

Revi Novita

# DAFTAR ISI

	Halaman
<b>DAFTAR ISI</b> .....	<b>xv</b>
<b>DAFTAR GAMBAR</b> .....	<b>xvii</b>
<b>DAFTAR TABEL</b> .....	<b>xviii</b>
<b>I. PENDAHULUAN</b>	
<b>A. Latar Belakang</b> .....	1
<b>B. Rumusan Masalah</b> .....	3
<b>C. Batasan Masalah</b> .....	3
<b>D. Tujuan Penelitian</b> .....	3
<b>E. Manfaat Penelitian</b> .....	4
<b>II. TINJAUAN PUSTAKA</b>	
<b>A. Penelitian Terdahulu</b> .....	5
<b>C. Pengenalan Pola</b> .....	9
<b>D. <i>Handwritten Character Recognition (HWCR)</i></b> .....	11
<b>E. Fitur</b> .....	12
<b>F. <i>Moment Invariant</i></b> .....	13
<b>G. <i>Pixel Density</i></b> .....	16
<b>H. <i>Support Vector Machine</i></b> .....	16
<b>III. METODOLOGI PENELITIAN</b>	
<b>B. Alat Pendukung</b> .....	22
<b>C. <i>Dataset</i></b> .....	23
<b>D. Tahapan Penelitian</b> .....	23
<b>1. Tahap 1 – Studi Literatur</b> .....	24
<b>2. Tahap 2 – Persiapan Bahan</b> .....	24
<b>3. Tahap 3 – <i>Preprocessing</i></b> .....	25
<b>4. Tahap 4 - Ekstraksi Fitur</b> .....	25
<b>5. Tahap 5 – Klasifikasi</b> .....	26
<b>6. Tahap 6 – Hasil</b> .....	26
<b>IV. HASIL DAN PEMBAHASAN</b>	

A. <i>Preprocessing</i> .....	28
B. <i>Pembagian Dataset</i> .....	29
C. <i>Feature Extraction</i> .....	30
1. <i>Ekstraksi Fitur Moment Invariant</i> .....	31
2. <i>Ekstraksi Fitur Pixel Density</i> .....	36
3. <i>Ekstraksi Fitur Gabungan Moment Invariant dan Pixel Density</i> .....	41
D. <i>Klasifikasi Fitur</i> .....	42
E. <i>Hasil</i> .....	45
F. <i>Analisis Kesalahan Klasifikasi</i> .....	49
1. <i>Kesalahan Klasifikasi Fitur Moment Invariant</i> .....	49
2. <i>Kesalahan Klasifikasi Fitur Pixel Density</i> .....	53
3. <i>Kesalahan Klasifikasi Fitur Moment Invariant dan Pixel Density</i> .....	53
<b>V. KESIMPULAN DAN SARAN</b>	
A. <i>Kesimpulan</i> .....	59
B. <i>Saran</i> .....	59
<b>DAFTAR PUSTAKA</b> .....	<b>61</b>
<b>LAMPIRAN</b>	

## DAFTAR GAMBAR

Gambar	Halaman
1. Induk Aksara Lampung.....	6
2. Diakritik Atas.....	7
3. Diakritik Bawah.....	7
4. Diakritik Kanan.....	8
5. Tanda Baca Aksara Lampung.....	8
6. Konsep <i>Support Vector Machine</i> .....	18
7. Tahapan Penelitian.....	24
8. Hasil Normalisasi Ukuran dan Binerisasi Citra Aksara A.....	28
9. Karakter Ba dan Ma.....	31
10. <i>Screenshot textfile</i> hasil ekstraksi fitur moment invariant.....	36
11. Pembagian blok citra karakter aksara Nga.....	37
12. <i>Screenshot textfile</i> hasil ekstraksi fitur <i>pixel density</i> .....	41
13. <i>Screenshot textfile</i> hasil ekstraksi fitur gabungan <i>moment invariant</i> dan <i>pixel density</i> .....	42
14. <i>Screenshot textfile</i> hasil ekstraksi fitur gabungan.....	50
15. Perbedaan (a) aksara Ka dan (b) aksara A.....	51

## DAFTAR TABEL

<b>Tabel</b>	<b>Halaman</b>
1. Rincian Pembagian Kelompok <i>Dataset</i> .....	29
2. Contoh Hasil Perhitungan Moment Karakter ‘ba’, ‘ga’, dan ‘ya’ .....	35
3. Sampel hasil ekstraksi fitur <i>pixel density</i> .....	38
4. Hasil Akurasi Pengujian Data <i>Validation</i> dan Data <i>Testing</i> .....	45
5. <i>Confusion Matrix</i> Fitur <i>Moment Invariant</i> .....	46
6. <i>Confusion Matrix</i> Hasil Fitur <i>Pixel Density</i> .....	47
7. <i>Confusion Matrix</i> Fitur Gabungan <i>Moment Invariant</i> dan <i>Pixel Density</i> .....	47
8. Jumlah data terbanyak kesalahan klasifikasi dengan fitur <i>moment invariant</i> . .....	49
9. Perbandingan Nilai Fitur Aksara Da dan Ga .....	54
10. Perbandingan nilai fitur <i>pixel density</i> Aksara Da dan Ga .....	56
11. Perbandingan nilai fitur <i>moment invariant</i> Aksara Da dan Ga .....	58

## I. PENDAHULUAN

### A. Latar Belakang

Indonesia merupakan negara yang kaya akan warisan budaya. Salah satu warisan tersebut adalah aksara tradisional. Dokumentasi-dokumentasi terdahulu tersimpan dalam bentuk tulisan aksara tradisional sehingga menjadi bukti sejarah dan alat dalam ilmu pengetahuan yang dapat dipelajari oleh generasi saat ini dan yang akan datang. Lampung merupakan salah satu daerah yang memiliki warisan budaya berupa aksara tradisional yang digunakan untuk berkomunikasi melalui tulisan. Aksara Lampung telah banyak ditemukan di beberapa dokumen kuno yang kini tersimpan di museum. Kondisi dokumen tersebut umumnya tertulis di atas kertas yang rapuh terhadap kerusakan. Dengan fakta tersebut, melakukan pengarsipan bagi dokumen-dokumen penting seperti dokumen kuno bertuliskan aksara Lampung itu sangat perlu untuk dilakukan. Salah satu cara tersebut yaitu dengan cara memanfaatkan teknologi komputer untuk mengenali masukan gambar berupa tulisan tangan aksara Lampung. Bidang ilmu yang membahas hal tersebut adalah pengenalan pola, khususnya pengenalan pola tulisan tangan atau disebut juga *Handwritten Character Recognition* (HWCR).

Salah satu tahapan penting yang terdapat dalam proses pengenalan pola tulisan tangan adalah *feature extraction*. Tahapan ini bertujuan untuk mengekstrak ciri khusus dari suatu karakter sehingga dapat dijadikan sebagai pembeda/ciri dengan karakter lain.

Selama berabad-abad, aksara Lampung kuno telah berevolusi dari keturunannya, naskah Devnagari. Proses evolusi ini telah mengubah skrip hingga batas tertentu hingga bentuk terbaru. Skrip saat ini jauh lebih sederhana daripada yang kuno. Naskah lampau Lampung terdiri dari 19 huruf, sedangkan skrip terakhir memiliki satu huruf lagi. Beberapa huruf mengandung satu atau dua zig-zag, sementara yang lain memiliki struktur yang lebih kompleks. Selain itu, beberapa dari mereka memiliki garis lurus yang terhubung ke zig-zag. Huruf-huruf utama itu disebut "*kelabai sughat*" (Junaidi *et al*, 2011).

Penelitian ini merupakan lanjutan dari penelitian sebelumnya tentang pengenalan karakter tulisan tangan. Pengembangan dilakukan dari segi penggunaan fitur untuk pengenalan karakter aksara. Salah satu penelitian pada pengenalan Aksara Lampung, sudah dilakukan oleh Akmal Junaidi *et al*. pada tahun 2011 dengan menggunakan 82 dokumen tulisan tangan beraksara Lampung. Penelitian ini berhasil mengenali tulisan tangan aksara lampung dengan tingkat akurasi sebesar 94,27% (Junaidi *et al*, 2011).

Berdasarkan fakta-fakta tersebut, penulis akan melakukan penelitian lanjutan tentang pengenalan karakter tulisan tangan aksara Lampung dengan

mengembangkan fitur yang belum pernah di implementasikan sebelumnya. Salah satu fitur yang akan di implementasikan dalam penelitian ini yaitu fitur *moment invariant*. *Moment Invariant* sering digunakan dalam pemrosesan gambar sebagai fitur dalam proses pengenalan pola. *Moment Invariant* memiliki properti yang dapat digunakan sebagai invarian dalam penskalaan gambar (Arafah *et al*, 2017).

## **B. Rumusan Masalah**

Berdasarkan latar belakang yang telah dipaparkan, masalah pokok yang perlu diselesaikan dan menjadi dasar dalam penelitian ini sebagai berikut.

1. Bagaimana cara mengekstraksi fitur *moment invariant* dan *pixel density* dari *image* tulisan tangan aksara Lampung untuk mengenali karakter tulisan tangan aksara Lampung;
2. Bagaimana tingkat akurasi penggunaan fitur *moment invariant* dan *pixel density* dalam pengenalan karakter tulisan tangan aksara Lampung.

## **C. Batasan Masalah**

Dalam penelitian ini diberikan batasan masalah diantaranya sebagai berikut.

1. *Dataset* yang diolah dalam penelitian ini merupakan hasil pemrosesan gambar tulisan tangan aksara Lampung dari penelitian yang telah dilakukan sebelumnya oleh Akmal Junaidi dengan jumlah sebanyak 32140 *dataset* (Junaidi *et al*, 2011);

2. Proses klasifikasi aksara Lampung menggunakan pendekatan SVM (*Support Vector Machine*) dengan memanfaatkan *tools* LibSVM.

#### **D. Tujuan Penelitian**

Tujuan dari penelitian ini adalah mengeksplorasi dan menganalisis fitur *moment invariant* dan *pixel density* dalam mengenali tulisan tangan aksara Lampung serta mengklasifikasikannya menggunakan SVM (*Support Vector Machine*).

#### **E. Manfaat Penelitian**

Manfaat yang diperoleh dari penelitian ini adalah sebagai berikut.

##### 1. Manfaat Praktis

- a) Mengembangkan penggunaan fitur-fitur dalam pengenalan pola tulisan tangan aksara Lampung dengan menerapkan fitur *moment invariant* dan *pixel density*;
- b) Mengetahui tingkat akurasi penggunaan fitur *moment invariant* dan *pixel density* dalam mengenali tulisan tangan aksara Lampung.

##### 2. Manfaat Akademis

- a) Mengaplikasikan ilmu yang telah dipelajari selama masa perkuliahan;
- b) Sebagai acuan untuk penelitian atau pengembangan sistem yang terkait dengan pengenalan karakter tulisan tangan aksara Lampung.

## II. TINJAUAN PUSTAKA

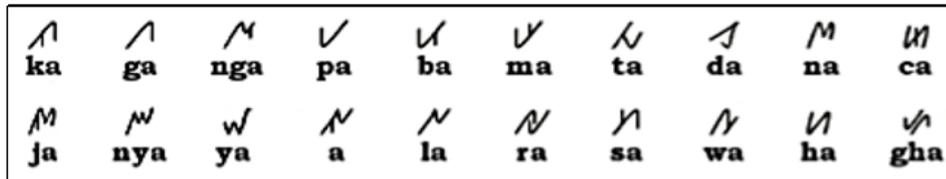
### A. Penelitian Terdahulu

Penelitian pada bidang HWCR telah banyak dilakukan oleh para peneliti. Salah satunya yaitu pengenalan pola tanda tangan menggunakan fitur *moment invariant* (Jariah *et al*, 2011). *Moment Invariant* digunakan untuk mengekstraksi citra tanda tangan menjadi suatu vektor *input* yang merepresentasikan setiap citra tanda tangan. Satu citra tanda tangan diwakili oleh 7 *moment invariant*. Hasil klasifikasi tanda tangan mencapai tingkat akurasi sebesar 80%. Penelitian lainnya yaitu penggunaan fitur *moment invariant* untuk pengenalan pola tanda tangan menggunakan metode klasifikasi KNN (Sany *et al*, 2017). Pada penelitian ini, Metode ekstraksi yang digunakan adalah metode *moment invariant* sebagai fitur untuk mengenali pola tanda tangan. Berdasarkan penelitian yang dilakukan, diperoleh hasil bahwa penggunaan *moment invariant* sebagai ekstraksi fitur mempunyai tingkat keberhasilan 72%.

### B. Aksara Lampung

Aksara Lampung berasal dari kelompok aksara Brahmi, yaitu aksara kuno yang berasal dari wilayah India Selatan. Aksara Lampung hanya terdiri dari 20 karakter

dasar. Bentuk utama semua karakter adalah lengkungan atau kurva. Dengan kata lain, setiap karakter mengandung setidaknya satu kurva yang bisa menghadap ke atas atau satu kurva ke bawah. Aksara Lampung terdiri dari huruf induk, anak huruf, juga terdapat lambang, angka dan tanda baca. Bentuk-bentuk huruf induk aksara Lampung ditunjukkan pada Gambar 2.1 berikut.



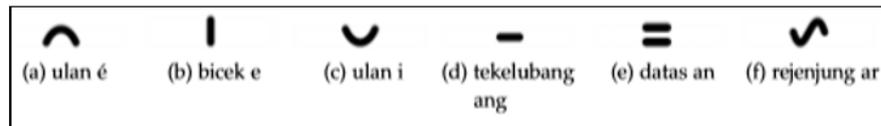
**Gambar 2.1** Induk Aksara Lampung

Aksara Lampung juga disebut dengan istilah Kaganga ditulis dan dibaca dari kiri ke kanan. Aksara Lampung yang dikenal saat ini telah mengalami perkembangan dan perubahan yang signifikan. Sebelumnya aksara Lampung kuno jauh lebih kompleks dengan bentuk yang lebih banyak memiliki lengkungan/kurva. Setiap huruf induk aksara Lampung dibaca dengan komposisi bunyi yang dihasilkan dari huruf konsonan dengan akhiran huruf vokal “a”. Untuk mengubah bunyi huruf vokal yang melekat pada huruf induk menjadi huruf vokal lain seperti i, u, e dan o, sistem aksara Lampung menggunakan diakritik (anak huruf). Sistem penulisan aksara Lampung mempunyai beberapa tanda baca diakritik (anak huruf). Diakritik tersebut dapat diletakkan di atas, bawah atau sebelah kanan aksara dasar.

Kelompok diakritik tersebut adalah diakritik atas, diakritik bawah dan diakritik kanan. Berikut merupakan penjelasan dari ketiga kelompok tersebut.

## 1. Diakritik Atas

Diakritik atas terdiri dari enam anak huruf yang memiliki nama atau istilah masing-masing. Keenam anak huruf tersebut ditunjukkan pada Gambar 2.2 berikut.

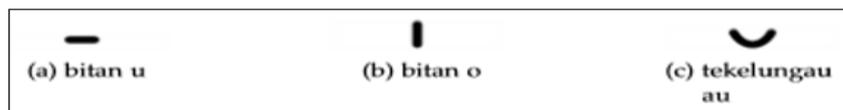


**Gambar 2. 2** Diakritik Atas

Diakritik ulan é dan bicek dapat menggantikan huruf vokal yang melekat pada induk huruf menjadi huruf vokal e. Perbedaan keduanya dapat terlihat dari segi pengucapan misalnya ulan é untuk pengucapan kata “pesta” dan bicek untuk pengucapan kata “pertama”. Diakritik ulan i menggantikan huruf vokal yang melekat pada induk huruf menjadi huruf vokal i. Diakritik tekelubang, datasan dan rejunjung dapat menambahkan ekspansi huruf konsonan “ng”, “n”, dan “r” diakhir vokal pada induk huruf.

## 2. Diakritik Bawah

Diakritik bawah terdiri dari tiga anak huruf dengan nama atau istilah masing-masing. Bentuk ketiga anak huruf tersebut ditunjukkan pada Gambar 2.3 berikut.



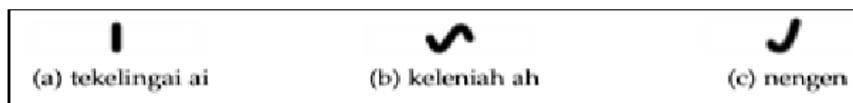
**Gambar 2. 3** Diakritik Bawah.

Diakritik bitan u menggantikan huruf vokal yang melekat pada huruf induk menjadi huruf vokal u. Diakritik bitan o menggantikan huruf vokal yang melekat

pada huruf induk menjadi vokal o. Diakritik tekelungau au mengubah huruf vokal yang melekat pada huruf induk menjadi gabungan dua huruf vokal (diftong) “au”.

### 3. Diakritik Kanan

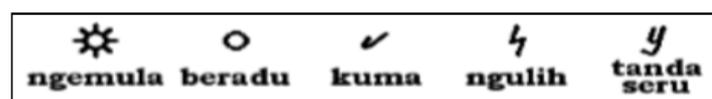
Diakritik kanan terdiri dari tiga anak huruf dengan nama atau istilah yang berbeda. Bentuk dari masing-masing anak huruf tersebut ditunjukkan pada Gambar 2.4 berikut.



**Gambar 2.4** Diakritik Kanan

Diakritik kanan merupakan diftong yang merubah huruf vokal pada huruf induk menjadi dua huruf vokal vokal “ai”. Diakritik keleniah dapat menimpa huruf vokal yang melekat pada huruf induk dan menambahkan ekspansi huruf konsonan “h” di akhir vokal. Diakritik nengen digunakan sebagai *vocal silencer* untuk huruf induk sehingga yang tersisa hanyalah konsonan dari huruf induk tersebut.

Selain huruf dan diakritik, dalam penulisan aksara Lampung juga terdapat beberapa tanda baca. Jumlah tanda baca yang terdapat pada aksara Lampung yaitu sebanyak lima tanda baca. Bentuk-bentuk dari kelima tanda baca tersebut beserta namanya masing-masing ditunjukkan pada Gambar 2.5 berikut.



**Gambar 2.5** Tanda Baca Aksara Lampung

### C. Pengenalan Pola

Pengenalan pola merupakan suatu ilmu untuk mengklasifikasikan atau menggambarkan pengukuran kuantitatif fitur (ciri) atau sifat utama dari suatu obyek. Pola sendiri merupakan suatu entitas yang terdefinisi dan dapat diidentifikasi serta diberi label atau nama (Putra, 2010). Pengenalan pola memiliki beragam jenis penerapannya antara lain *face recognition*, *biometric*, *character recognition*, *object recognition*, dan lain-lain. Secara sederhana, pengenalan pola adalah proses pengelompokan suatu objek ke dalam kelas-kelas berdasarkan ciri dan parameter yang diekstrak dari objek (Naufal, 2007).

Proses *Character Recognition* terdiri dari beberapa fase. Urutan fase-fase tersebut adalah sebagai berikut.

#### 1. *Image Acquisition*

*Image Acquisition* adalah tahap pertama dalam *Character Recognition* untuk mendapatkan *image* melalui kamera atau pemindai. Format penyimpanan *image* dapat dipilih dari beberapa format standar yang tidak dikompresi. Hal ini diperoleh untuk mempertahankan sifat asli objek yang akan dikenali.

#### 2. *Pre-Processing*

Tahap *preprocessing* melibatkan berbagai tahap operasi untuk meningkatkan kualitas citra agar dapat diproses lebih lanjut. *Preprocessing* melibatkan pengurangan *noise*, binerisasi, deteksi tepi dan *thresholding*.

### 3. *Segmentation*

Proses pensesegmentasian terdiri dari segmentasi baris dan segmentasi karakter. Tujuan segmentasi adalah membagi serangkaian karakter menjadi karakter-karakter yang terpisah (Sharma *et al*, 2017).

### 4. *Feature Extraction*

*Feature Extraction* atau ekstraksi fitur merupakan tahap dalam pengenalan pola yang bertujuan untuk memperoleh informasi yang lebih jelas mengenai data dalam sebuah citra (Yodha *et al*, 2014).

### 5. *Classification and Recognition*

Klasifikasi merupakan proses *training* yang dapat digunakan untuk memprediksi kelas dari objek yang label kelasnya belum diketahui. Algoritma klasifikasi yang banyak digunakan secara luas, yaitu *Decision* atau *Classification Trees*, *Bayesian Classifiers* atau *Naive Bayes Classifiers*, *Neural Networks*, *Analisa Statistik*, *Algoritma Genetika*, *Rough Sets*, *K-Nearest Neighbor*, *Metode Rule Based*, *Memory Based Reasoning*, dan *Support Vector Machines (SVM)* (Leidiyana, 2013). Pada fase ini ketika *input* gambar dalam bentuk *feature vector* dimasukkan ke dalam *classifier* sebagai sistem pengenalan karakter yang sudah terlatih. Perbandingan fitur *input* dengan pola tersimpan kemudian dilakukan untuk menemukan pencocokan kelas yang sesuai dengan *input* gambar (Sharma *et al*, 2017).

#### **D. Handwritten Character Recognition (HWCR)**

*Character Recognition* merupakan proses mendeteksi dan mengenali karakter dari suatu *input* objek atau gambar karakter. Berdasarkan cara pembacaan *input*, *character recognition* dibedakan menjadi dua macam yaitu *offline* dan *online character recognition*. Sistem *offline* mengolah data atau dokumen yang sebelumnya memerlukan proses digitalisasi untuk menjadi *file* gambar digital yang dapat dibaca oleh komputer. Sedangkan sistem *online* mengolah data atau dokumen yang dibuat langsung menggunakan perangkat digital sehingga dapat diproses secara *real time* (Priyanka, 2019). Berdasarkan jenis *input* tulisan tangan, *character recognition* dibedakan menjadi jenis aksara cetak atau *printed character recognition* dan *handwritten character recognition*. Akurasi pada HWCR biasanya masih rendah dibandingkan *printed character recognition* dikarenakan besarnya perbedaan bentuk dan tipe tulisan dari setiap individu (Rao *et al*, 2016).

*Handwritten Character Recognition* (HWCR) merupakan salah satu cabang dari *Optical Character Recognition* (OCR) yang berfokus untuk memaksimalkan kemampuan komputer dalam hal pengenalan karakter pada tulisan tangan. Penelitian di bidang HWCR yang dilakukan dengan berbagai bahasa menjadi salah satu topik yang jumlahnya terus meningkat karena semakin dianggap penting di dunia dan salah satu aplikasinya adalah deteksi otomatis pengirim surat pada dinas pos atau identifikasi tulisan dan tanda tangan pada cek perbankan (Mohammadpoor, 2015).

Gaya tulisan pada setiap orang memiliki variasi yang sangat beragam. Oleh karena itu, pengenalan karakter tulisan tangan secara otomatis menggunakan komputer menjadi suatu tantangan yang sangat menarik. Penelitian di bidang HWCR terus dikembangkan dengan harapan tersedianya sistem yang mampu menafsirkan tulisan tangan tanpa memaksa pengguna untuk menyesuaikan gaya tulisan sesuai dengan yang disediakan oleh perangkat (Napper, 2011).

Pengenalan karakter tulisan tangan (HWCR) didefinisikan sebagai konversi teks tulisan tangan menjadi format yang dapat diproses mesin. HWCR juga memiliki kapasitas komputer untuk mendapatkan dan menguraikan input tulisan tangan yang dapat dipahami dari sumber seperti laporan kertas, foto, layar sentuh dan perangkat lainnya. Sistem pengenalan tulisan tangan yang lengkap juga menangani pemformatan, melakukan segmentasi yang benar ke dalam karakter dan menemukan kata-kata yang paling mungkin.

### **E. Fitur**

Ekstraksi fitur merupakan tahapan yang penting dalam pembuatan *machine learning* untuk klasifikasi atau pengenalan pola. Dengan ekstraksi fitur, informasi ciri gambar bisa didapatkan (Ardiansjah *et al*, 2012). Fitur atau atribut merupakan karakteristik ciri/pembeda dari suatu objek yang diperoleh melalui pengukuran dan dapat berupa simbol, numerik, kategori, dan lain sebagainya. Fitur dapat direpresentasikan dalam bentuk variabel kontinu, diskrit, atau diskrit biner. Untuk

memperoleh fitur dapat menggunakan berbagai metode pencarian fitur atau disebut juga metode ekstraksi fitur (Putra, 2010).

Proses ekstraksi ciri merupakan bagian dari teknik pengenalan pola (*pattern recognition*). Ekstraksi ciri bertujuan untuk mengambil atau mengekstraksi nilai-nilai unik dari suatu objek yang membedakan suatu objek dengan objek yang lain. Ekstraksi ciri secara umum dianggap sebagai sebuah proses pemetaan objek *input* awal menjadi ciri yang mempresentasikan objek secara kuantitatif. Proses ekstraksi ciri pada pengolahan citra digital dilakukan setelah proses segmentasi sering terbentur pada permasalahan posisi objek, rotasi objek, dan perubahan skala dari objek. Posisi objek yang bergeser, berputar dan berubah ukurannya menyebabkan degradasi informasi yang sudah dimiliki sebagai *knowledge* dapat menyebabkan kesalahan dalam pengenalan/identifikasi objek tersebut (Sunnyoto, 2013). Fakta-fakta tersebut menunjukkan bahwa proses-proses ekstraksi ciri/fitur harus dilakukan dengan hati-hati. Selain itu, penambahan fitur juga harus mempertimbangkan kesesuaian/ketepatan fitur dengan objeknya sehingga pengenalan objek dapat berlangsung dengan optimal.

#### **F. *Moment Invariant***

*Hu* memperkenalkan *moment invariant* sebagai fitur untuk pengenalan pola. *Moment* dapat menggambarkan suatu objek dalam hal area, posisi, dan orientasi. Fitur ini menggunakan dasar *moment* pusat yang ternormalisasi. *Moment* yang dihasilkan dapat digunakan untuk menangani translasi, penyekalaan, dan rotasi

gambar (Kadir, 2013). Persamaan dasar dari *moment* suatu objek didefinisikan sebagai berikut.

$$\sum_{x=1}^M \sum_{y=1}^N x^i y^j I(x, y) \quad (\text{persamaan 1})$$

Keterangan:

$i, j = 0, 1, 2, \dots$ , dengan  $ij$  menyatakan orde momen;

M menyatakan jumlah kolom pada citra;

N menyatakan jumlah baris pada citra;

$x$  adalah ordinat piksel;

$y$  adalah absis piksel;

$I(x,y)$  menyatakan intensitas piksel pada posisi  $(x,y)$ .

*Moment* order ke-0 dan ke-1 (*zero and first-order moments*) didefinisikan sebagai berikut.

$$M_{00} = \sum_{xy} a_{xy} \quad (\text{persamaan 2})$$

$$M_{10} = \sum_x \sum_y x \cdot a_{xy} \quad (\text{persamaan 3})$$

$$M_{01} = \sum_x \sum_y y \cdot a_{xy} \quad (\text{persamaan 4})$$

Pada citra biner yang mana  $a_{xy}$  akan bernilai 0 atau 1, *moment* order ke-0 ( $m_{00}$ ) adalah sama dengan area dari objek yang dapat dicari dengan persamaan (1).

Pusat dari area atau massa (*centroid*) adalah parameter yang baik untuk menyatakan lokasi dari objek. Pusat massa dari objek didefinisikan sebagai berikut.

$$x' = \frac{M_{10}}{M_{00}}, y' = \frac{M_{01}}{M_{00}} \quad (\text{persamaan 5})$$

Objek  $(x', y')$  merupakan pusat koordinat dari objek yang dihitung setelah mengetahui  $m_{00}, m_{01}, m_{10}$ . *Moment* pusat (*central moment*) adalah *moment* yang bersesuaian dengan pusat massa yang didefinisikan melalui persamaan:

$$\mu_{ij} = \sum_{x=1}^M \sum_{y=1}^N (x - x')^i (y - y')^j I(x, y) \quad (\text{persamaan 6})$$

*Moment* pusat yang ternormalisasi dapat dihitung dengan persamaan berikut:

$$\eta_{ij} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}, \gamma = \frac{i+j+2}{2} \quad (\text{persamaan 7})$$

*Moment* tingkat ke-1 adalah selalu *invariant*. Dari persamaan (7) yaitu *moment* ternormalisasi, sekumpulan *moment invariant* dapat didefinisikan. *Moment-moment* ini sangat berguna dalam membuat vektor ciri untuk pengenalan objek.

Persamaan dari *moment-moment invariant* adalah sebagai berikut.

$$\phi_1 = \eta_{20} + \eta_{02} \quad (\text{persamaan 8})$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (\text{persamaan 9})$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} + \eta_{03})^2 \quad (\text{persamaan 10})$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (\text{persamaan 11})$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[(3\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (\text{persamaan 12})$$

$$\begin{aligned} \phi_6 = & (\eta_{20} - 3\eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ & 4\eta_{11}(\eta_{30} - \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned} \quad (\text{persamaan 13})$$

$$\begin{aligned} \phi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (\text{persamaan 14})$$

Ketujuh persamaan *moment invariant*, yaitu persamaan (8) sampai dengan (14) menunjukkan bahwa setiap objek memiliki tujuh ciri (Septiarini, 2012).

### **G. Pixel Density**

*Pixel Density* merupakan fitur lokal yang didapatkan dengan cara menghitung jumlah piksel obyek (Sharma *et al*, 2011). Secara matematis, perhitungan *pixel density* dapat dimodelkan sesuai dengan persamaan berikut. (Wai *et al*, 2013).

$$PD = \frac{A}{n}$$

Dengan:

PD = *pixel density*

A = area, luas citra (satuan pixel)

n = jumlah piksel obyek

### **H. Support Vector Machine**

*Support Vector Machine* (SVM) pertama kali diperkenalkan oleh Vladimir Vapnik pada tahun 1992. SVM adalah sekumpulan metode pembelajaran terawasi terkait yang digunakan untuk klasifikasi dan regresi. SVM secara bersamaan meminimalkan kesalahan klasifikasi empiris dan memaksimalkan margin geometrik. Jadi SVM disebut *Maximum Margin Classifiers* (Srivastava, 2005).

SVM berada dalam satu kelas dengan ANN dalam hal fungsi dan kondisi permasalahan yang biasa diselesaikan. Keduanya masuk dalam kelas *supervised learning*. Para ilmuwan ataupun praktisi telah banyak menerapkan teknik ini

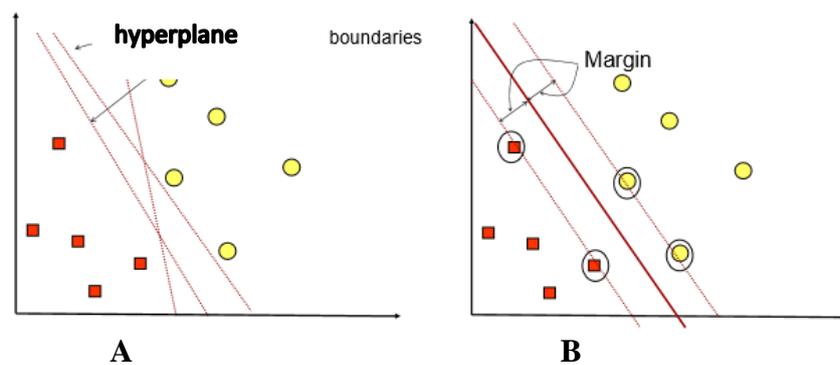
dalam menyelesaikan masalah-masalah nyata dalam kehidupan sehari-hari, baik dalam masalah *gene exception analysis*, *finacial*, cuaca hingga pada kedokteran. Dalam banyak implementasi, SVM memberikan hasil yang lebih baik dari ANN, terutama dalam solusi yang dicapai. ANN menemukan solusi berupa *local optimal* sedangkan SVM menemukan solusi yang *global optimal* (Lukman, 2016).

Model *linear classification* dibatasi menjadi *linear perceptron* dan *linear SVM*. Proses training pada kedua *classifier* pada prinsipnya bertujuan mencari “batas” pemisah antara dua buah *class/category* di ruang vektor. Yang membedakan kedua buah model di atas adalah prinsip yang dipakai untuk menemukan batas pemisah (*separating hyperplane*) tersebut. Pada *perceptron*, fungsi klasifikasi dinyatakan dalam *primal form*, yaitu persamaan yang melibatkan *weights/bobot* (tepatnya *dot product* antara data dan *weight*). *Weights* ini sebenarnya tidak lain merepresentasikan posisi *hyperplane* itu pada ruang vektor. Proses *training* pada *perceptron* berjalan dengan mengoreksi nilai *weights* secara berulang, sedemikian hingga kedua buah kelas dapat dipisahkan secara sempurna oleh *hyperplane* itu. *Weight* ini adalah hasil akhir dari proses pembelajaran pada *perceptron*.

Untuk *non-linear problem*, *perceptron* harus menambahkan satu *layer* lagi (*hidden layer*) agar bisa menemukan *hyperplane* tersebut secara *non-linear*. Dalam hal ini proses pembelajaran yang populer antara lain *backpropagation*. SVM pada prinsipnya adalah *linear classifier*. Tetapi SVM bisa juga dipakai untuk *non-linear problem*, dengan dua tahap. Pertama-tama data diproyeksikan ke ruang vektor baru berdimensi tinggi sedemikian hingga data itu dapat terpisah

secara *linear*. Baru selanjutnya, SVM mencari *hyperplane* optimal dengan cara yang sama sebagaimana di atas, yaitu bekerja sebagai *linear machine*. Ini didukung oleh teori *Cover*, yang menyatakan bahwa suatu ruang vektor dapat ditransformasikan ke ruang vektor baru yang pada probabilitas tinggi dapat terpisah secara *linear*, jika memenuhi dua syarat yaitu transformasi itu *non-linear* dan dimensi ruang vektor yang baru itu cukup tinggi.

Fungsi klasifikasi pada SVM dinyatakan sebagai *dual form*, yang melibatkan *dot product* antara data pada *training set*. Dengan demikian, proses pencarian *hyperplane* optimal melibatkan *dot product* dari data yang sudah diproyeksikan tersebut. Di sinilah muncul konsep “*kernel trick*”, yaitu menghitung nilai *dot product* dua buah data itu secara implisit, yang artinya dapat dihitung tanpa perlu mengetahui bagaimana wujud dari fungsi transformasi itu (Nugroho *et al*, 2003).



**Gambar 2. 6** Konsep Support Vector Machine (Nugroho *et al*, 2003).

Berdasarkan pada Gambar 2.6, klasifikasi SVM dapat didefinisikan sebagai usaha menemukan *hyperplane* yang memisahkan antara kedua *class*. *Hyperplane* pemisah terbaik antara kedua *class* dapat ditemukan dengan mengukur titik

maksimal margin dari *hyperplane*. Margin adalah jarak antara *hyperplane* dengan *pattern* terdekat dari masing-masing *class*. *Pattern* yang paling dekat dengan *hyperplane* ini disebut sebagai *support vector*.

Garis solid pada gambar B menunjukkan *hyperplane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua *class*, sedangkan titik merah dan kuning yang berada dalam lingkaran hitam adalah *support vector*. Usaha untuk mencari lokasi *hyperplane* ini merupakan inti dari proses pembelajaran pada SVM. Data yang tersedia dinotasikan  $\bar{x}_i \in \mathfrak{R}^d$  sedangkan label masing-masing dinotasikan  $y_i \in \{-1,+1\}$  untuk  $i= 1,2,\dots,l$ , yang mana  $l$  adalah banyaknya data. Diasumsikan kedua class  $-1$  dan  $+1$  dapat terpisah secara sempurna oleh *hyperplane* berdimensi  $d$ , yang didefinisikan

$$\bar{w} \cdot \bar{x} + b = 0 \quad (\text{Persamaan 15})$$

*Pattern*  $\bar{x}_i$  yang termasuk class  $-1$  (sampel negatif) dapat dirumuskan sebagai *pattern* yang memenuhi pertidaksamaan

$$\bar{w} \cdot \bar{x}_i + b \leq -1 \quad (\text{Persamaan 16})$$

sedangkan *pattern*  $\bar{x}_i$  yang termasuk class  $+1$  (sampel positif)

$$\bar{w} \cdot \bar{x}_i + b \geq +1 \quad (\text{Persamaan 17})$$

Margin terbesar dapat ditemukan dengan memaksimalkan nilai jarak antara *hyperplane* dan titik terdekatnya, yaitu  $1/\|W\|$ . Hal ini dapat dirumuskan sebagai *Quadratic Programming (QP) problem*, yaitu mencari titik minimal persamaan (4), dengan memperhatikan *constraint* persamaan (5).

$$\min_{\vec{w}} \tau(\vec{w}) = \frac{1}{2} \|\vec{w}\|^2 \quad (\text{Persamaan 18})$$

$$y_i(\vec{x}_i \cdot \vec{w} + b) - 1 \geq 0, \quad \forall i \quad (\text{Persamaan 19})$$

*Problem* ini dapat dipecahkan dengan berbagai teknik komputasi, di antaranya *Lagrange Multiplier*.

$$L(\vec{w}, b, \alpha) = \frac{1}{2} \|\vec{w}\|^2 - \sum_{i=1}^l \alpha_i (y_i (\vec{x}_i \cdot \vec{w} + b) - 1)$$

$$(i = 1, 2, \dots, l) \quad (\text{Persamaan 20})$$

$\alpha_i$  adalah *Lagrange multipliers*, yang bernilai nol atau positif ( $\alpha_i \geq 0$ ). Nilai optimal dari persamaan (6) dapat dihitung dengan meminimalkan  $L$  terhadap  $w$  dan  $b$ , dan memaksimalkan  $L$  terhadap  $\alpha_i$ . Dengan memperhatikan sifat bahwa pada titik optimal *gradient*  $L=0$ , persamaan (6) dapat dimodifikasi sebagai maksimalisasi *problem* yang hanya mengandung saja  $\alpha_i$ , sebagaimana persamaan (7) di bawah.

Maximize:

$$\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \vec{x}_i \cdot \vec{x}_j \quad (\text{Persamaan 21})$$

*Subject to:*

$$\alpha_i \geq 0 \quad (i = 1, 2, \dots, l) \quad \sum_{i=1}^l \alpha_i y_i = 0 \quad (\text{Persamaan 22})$$

Dari hasil dari perhitungan ini  $\alpha_i$  yang kebanyakan bernilai positif. Data yang berkorelasi dengan  $\alpha_i$  yang positif inilah yang disebut sebagai *support vector*.

Seperti yang sudah dijelaskan sebelumnya, SVM tidak hanya dirancang untuk dapat menyelesaikan permasalahan *linear* tetapi juga permasalahan *non-linear*. Untuk menyelesaikan masalah *non-linear* menggunakan SVM, data  $\bar{x}$  dipetakan oleh suatu fungsi ke ruang vektor yang berdimensi lebih tinggi. Fungsi yang memetakan suatu permasalahan *non-linear* ke dalam ruang ciri yang lebih tinggi disebut dengan fungsi *kernel*. Metode SVM memiliki kelebihan menggunakan *Empirical Risk Management* (ERM) yaitu metode *machine learning* yang difokuskan pada usaha untuk meminimalkan *error* pada *training set*, dan dalam SVM diwujudkan dengan memilih *hyperplane* dengan *margin* terbesar. Berbagai studi empiris menunjukkan bahwa pendekatan ERM pada SVM memberikan *error* generalisasi yang lebih kecil dibandingkan yang diperoleh dari strategi ERM pada *neural network* maupun metode yang lain (Nugroho, 2003).

### **III. METODOLOGI PENELITIAN**

#### **A. Waktu dan Tempat Pelaksanaan**

Penelitian ini dilaksanakan pada periode genap tahun ajaran 2018/2019, di Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

#### **B. Alat Pendukung**

Pengerjaan penelitian ini menggunakan peralatan pendukung sebagai berikut.

##### 1. Perangkat Keras (*Hardware*)

Perangkat keras yang digunakan dalam penelitian ini berupa notebook dengan spesifikasi sebagai berikut.

- *Processor* : Intel® CPU N2840 @ 2.16GHz
- *Installed RAM* : 2,00 GB
- *System Type* : 64-bit operating system, x64-based processor

##### 2. Perangkat Lunak (*Software*)

Perangkat lunak yang digunakan dalam penelitian ini adalah sebagai berikut.

- Sistem Operasi *Windows 10 Home Single Language* 64-bit
- Octave versi 4.4.1

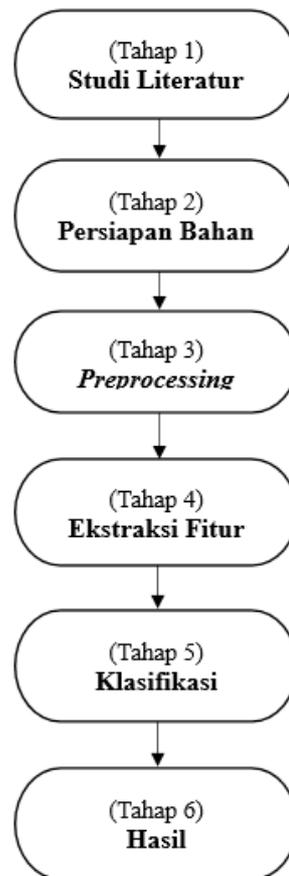
### **C. Dataset**

*Dataset* yang digunakan dalam penelitian ini bersumber dari *Lampung Handwritten Characters Datasets* yang tersedia di <https://patrecs.cs.tudortmind.de/cms/en/home/-Resources/index.html> (Junaidi *et al*, 2011).

*Dataset Lampung Handwritten Characters* terdiri dari beberapa file yaitu gambar dokumen tulisan tangan asli (*raw image*), *connected component* dalam bentuk *graycale*, dan *file* anotasi citra aksara Lampung. Total dokumen *raw image* adalah 82 *file* dengan format gambar ppm dan terletak pada direktori *raw\_images* dalam *file dataset*. Total *connected component grayscale* adalah 32140 *file* gambar yang terdistribusi dalam 82 sub direktori dengan format gambar pgm dan terletak pada direktori *char\_css* dalam *file dataset*. Informasi pada *file* anotasi terdiri dari enam kolom. Kolom pertama merupakan nama *file* dari komponen yang diekstraksi. Kolom kedua sampai kelima merupakan koordinat *bounding box* dari *connected component* pada dokumen *raw image*. Kolom keenam merupakan keterangan karakter yang merupakan label aksara Lampung dengan 18 kelas.

### **D. Tahapan Penelitian**

Penelitian ini terdiri dari beberapa tahapan. Tahapan-tahapan tersebut ditunjukkan pada Gambar 3.1.



**Gambar 3.1** Tahapan Penelitian

### **1. Tahap 1 – Studi Literatur**

Tahap awal pada penelitian ini yaitu studi literatur tentang isu-isu penting dalam penelitian ini. Isu-isu tersebut antara lain adalah konsep dasar tentang pengenalan pola, HWCR, *Moment Invariant* serta pendekatan-pendekatan teknis yang digunakan untuk mengimplementasikannya pada aksara Lampung.

### **2. Tahap 2 – Persiapan Bahan**

Sebelum masuk ke tahapan inti pengenalan pola, tahap kedua yang dilakukan adalah mempersiapkan bahan yang akan dipergunakan sebagai objek penelitian.

Bahan penelitian tersebut yaitu berupa *dataset* yang diperoleh dari situs *web Pattern Recognition Group di TU Dortmund University*. *Dataset* nantinya akan dibagi menjadi 3 bagian, yaitu *dataset* untuk *training*, *dataset* untuk validasi dan *dataset* untuk *testing*. Data *training* terdiri dari 52 dokumen, data validasi terdiri dari 10 dokumen dan data *testing* terdiri dari 20 dokumen. *Dataset* yang telah dipakai selanjutnya disimpan pada direktori yang berbeda dengan direktori yang menyimpan gambar asli.

### **3. Tahap 3 – Preprocessing**

Tahap selanjutnya adalah memproses gambar agar sesuai dengan kebutuhan penelitian. *Dataset* merupakan gambar bertipe *grayscale* namun belum memiliki ukuran yang seragam. Gambar pada *dataset* yang ukurannya masih bervariasi akan dilakukan proses *resizing* atau penyamaan ukuran terlebih dahulu sekaligus dibinerisasi. Ukuran yang digunakan yaitu 32x32 piksel serta menggunakan metode interpolasi *bilinear*.

Proses transformasi gambar menjadi lebih kecil bertujuan untuk memperkecil nilai yang harus diproses saat proses klasifikasi.

### **4. Tahap 4 - Ekstraksi Fitur**

Tahap selanjutnya adalah tahapan ekstraksi fitur. Ekstraksi fitur bertujuan untuk memperoleh (mengeksktraksi) nilai-nilai yang unik dalam objek gambar. Nilai-nilai unik atau yang dikenal dengan istilah “fitur” inilah yang menjadi ciri dari setiap

kelas sehingga membedakan antara kelas yang satu dengan yang lainnya. Tahapan yang menjadi fokus bahasan pada penelitian ini menggunakan fitur *moment invariant* yang pertama kali diperkenalkan oleh Timo Ojala dkk (Ojala *et al*, 2002). Pada tahapan ekstraksi fitur *moment invariant*, dihasilkan sebuah berkas baru yang menyimpan karakteristik setiap objek *dataset* Aksara Lampung. Karakteristik ini yang dipelajari oleh SVM untuk dapat dikenali.

### **5. Tahap 5 – Klasifikasi**

Tahapan selanjutnya yang dilakukan setelah nilai fitur didapatkan adalah tahapan klasifikasi. Tahapan klasifikasi pada penelitian ini menggunakan *Support Vector Machine* (SVM). SVM dikenal sebagai alat klasifikasi di berbagai bidang penelitian dengan hasil yang memuaskan. Teknik ini diperkenalkan oleh Vapnik pada tahun 1992 dan berhasil terus dikembangkan oleh para peneliti (Nugroho, 2003). Pengklasifikasi pada SVM membuat keputusan dengan memaksimalkan margin antara dua kelas. Batas keputusan dengan margin yang lebih besar biasanya akan menghasilkan prediksi yang lebih baik. Hasil klasifikasi ini yang selanjutnya akan dievaluasi untuk mendapatkan tingkat akurasi pengenalan Aksara Lampung.

### **6. Tahap 6 – Hasil**

Tahap terakhir pada penelitian ini adalah mendapatkan nilai akurasi pengenalan Aksara Lampung dengan menggunakan fitur *moment invariant*. Proses ini menjadi salah satu tolak ukur keberhasilan pekerjaan. Hasil klasifikasi diuji

dengan menggunakan *confusion matrix* yang kemudian diukur tingkat akurasinya menggunakan persamaan sebagai berikut :

$$\text{Detection Rate} = \frac{TP}{TP + TN} \times 100\%$$

Keterangan:

*TP (True Positive)* = Data yang diprediksi benar

*TN (True Negative)* = Data yang diprediksi salah

## V. KESIMPULAN DAN SARAN

### A. Kesimpulan

Kesimpulan yang diperoleh berdasarkan penelitian yang telah dilakukan adalah sebagai berikut.

1. Fitur *moment invariant* dan *pixel density* telah berhasil diimplementasikan pada pengenalan karakter tulisan tangan Aksara Lampung.
2. Penggunaan fitur *moment invariant* belum cukup baik dalam pengenalan karakter tulisan tangan Aksara Lampung karena sifat fitur *moment invariant* yaitu *rotation*, *translation* dan *scaling invariant*.
3. Akurasi tertinggi didapatkan pada penggunaan kernel linear dengan nilai mencapai 49,94% untuk fitur *moment invariant*, 91,83% untuk fitur *pixel density*, dan 91,90% untuk fitur gabungan.

### B. Saran

Saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai berikut.

1. Menggunakan fitur lain selain *moment invariant* dan *pixel density* agar perbandingan kinerja dapat dievaluasi.

2. Mengaplikasikan dan membandingkan penggunaan metode klasifikasi lain agar didapat hasil terbaik pada penggunaan fitur *moment invariant* dalam proses pengenalan tulisan tangan Aksara Lampung.
3. Meningkatkan kualitas citra yang digunakan sebagai dataset pada proses *preprocessing*.

## DAFTAR PUSTAKA

- Ardiansjah, Y., Nanik Suciati, Darlis Herumurti. 2012. *Pengenalan Spesies Tanaman Berdasarkan Bentuk Daun Menggunakan Metode Klasifikasi Move Median Center (Mmc) Hypersphere*. no. Mmc: 1–7.
- Heutte, L., T. Paquet a, J.V. Moreau b, Y. Lecourtier a, C. Olivier. 1998. *A Structural/Statistical Feature Based Vector for Handwritten Character Recognition*. Pattern Recognition Letters, Volume 19, Issue 7.
- Jariah, A., Mohammad Isa Irawan, Imam Mukhlash. 2011. *Pengenalan Pola Tanda Tangan Menggunakan Metode Moment Invariant Dan Jaringan Syaraf Radial Basis Function (Rbf)*. Prosiding Seminar Nasional Penelitian. Universitas Yogyakarta.
- Junaidi, A. 2016. *Lampung Handwritten Character Recognition*. Disertasi. Program Doctoral Dortmund University. Dortmund. German.
- Junaidi, A., Szilard Vajda, Gernot A. Fink. 2011. *Lampung- A New Handwritten Character Benchmark: Database, Labeling and Recognition*. Proceedings of International Workshop on Multilingual OCR, MOCR. Page 105112.
- Kadir, A. dan Adhi Susanto. 2013. *Teori Dan Aplikasi Pengolahan Citra*. Yogyakarta: ANDI.
- Leidiyana, H. 2013. *Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bemotor*. Jurnal Penelitian Ilmu Komputer, System Embedded & Logic. Vol. 1, No. 1.
- Lukman. 2016. *Penerapan Algoritma Support Vector Machine (SVM) Dalam Pemilihan Beasiswa : Studi Kasus Smk Yapimda*. Faktor Exacta, Volume 9, Issue 1.
- Mohammadpoor, R. 2015. *Recognition of Persian Handwritten Numbers Using LBP-HOG Descriptor*. Bull. Env.Pharmacol. Life Sci., Vol 4.
- Napper, J. 2011. *Handwritten Character Recognition System*. U.S Patent Documents No. 8.280.168 B2.

- Naufal, M. 2017. *Implementasi Metode Klasifikasi K-Nearest Neighbor (K-Nn) Untuk Pengenalan Pola Batik Motif Lampung*. Bandar Lampung. Perpustakaan Universitas Lampung.
- Nugroho, A. Satriyo., Arief Budi Witarto, Dwi Handoko. 2003. *Support Vector Machine Teori Dan Aplikasinya*. Kuliah Umum IlmuKomputer.com
- Nurajizah, S. 2013. *Penerapan Metode Support Vector Machine Berbasis Particle*. Jurnal Techno Nusa Mandiri, Volume X, Issue 1.
- Ojala, Timo., Pietikainen, Matti., Maenpaä, Topi. 2002. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24, No.7.
- Patel, S. (2017, Mei 3). *Chapter 2 : SVM (Support Vector Machine)—Theory*. <https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72> (Diakses tanggal 7 September 2019).
- Prasetyo, E. 2011. *Pengolahan Citra Digital Dan Aplikasinya Menggunakan Matlab*. Yogyakarta: ANDI.
- Priyanka, M., S. Ramakrishnan, N. R. Raajan. 2019. *Electronic Handwriting Character Recognition (E-HWCR)*. International Journal of Engineering and Technology (IJET), Vol.24, No.7.
- Putri, E., Diah Puspitaningrum and Andre Mirfen. 2015. *Identifikasi Tanda Tangan Dengan Pendekatan Support Vector Machine*. Jurnal Sains, Teknologi dan Industri, Volume 12, Issue 2.
- Putra, Darma. 2010. *Pengolahan Citra Digital*. Yogyakarta: ANDI.
- Rahman, M. F., Darmawidjaja, M. I. & Alamsah, D. 2017. *Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN)*. Pp 36-45 In : Jurnal Informatika. Universitas Ahmad Dahlan, Yogyakarta. Volume 11. No 1. 1978-0524. DOI 10.26555/jifo.v11i1.a5452.
- Rao, N.Venkata., Dr.A.S.C.S.Sastry, A.S.N.Chakravarthy, Kalyanchakravarthi P. 2016. *Optical Character Recognition Technique Algorithm*. Journal of Theoretical and Applied Information Technology, Volume 83, No.2.

- Sany, J. S., Agus Wahyu Widodo, & Chandra Dewi. 2017. *Penggunaan Ciri Geometric Invariant Moment pada Pengenalan Tanda Tangan*. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. Vol. 1, No.9 ( 859-867)
- Saputra, H., Fitri Arnia, Fardian. 2016. *Penggunaan Fitur Momen Zernike Untuk Pengenalan Karakter Jawi Cetak*. Jurnal Online Teknik Elektro. Universitas Syiah Kuala, Volume 1, No. 3 (15-20).
- Septiarini, A. 2012. *Pengenalan Pola Pada Citra Digital Dengan Fitur Momen Invariant*. Jurnal Informatika Mulawarman, Volume 7, No. 1.
- Sharma, A., Soumil Khare, Sachin Chavan. 2017. *A Review on Handwritten Character Recognition*. International Journal of Computer Science and Technology, Volume 8, Issue 1.
- Sharma, R. & S. M., 2011. *An Offline Signature Verification System Using Neural Network Based on Angle Feature and Energy Density*. International Journal on Emerging Technologies, 2(2), pp. 84-89
- Sunyoto, A. 2013. *Analisa Metode Moment Invariant Untuk Mendeteksi Obyek Yang Telah Mengalami Transformasi*. Jurnal Dasi, Volume 14, Issue 1.
- Sutoyo, T., Edy Mulyanto, Vincent Suhartono, Oky Dwi Nurhayati. 2009. *Teori Pengolahan Citra Digital*. Yogyakarta: ANDI.
- Wai, H. H. & A. S. L., 2013. *Feature Extraction for Offline Signature Verification System*. IJCCER International Journal of Computer & Communication Engineering Research, I(3), pp. 84-87.
- Wijaya, I Gede Pasek Suta et al. 2004. *Pengenalan Citra Sidik Jari berbasis Transformasi Wavelet dan Jaringan Syaraf Tiruan*. Jurnal Teknik Elektro Vol. 4, No. 1.
- Yodha, J., Achmad Wahid Kurniawan. 2014. *Pengenalan Motif Batik Menggunakan Deteksi Tepi Canny dan K-Nearest Neighbor*. Techno.COM, Vol. 13, No. 4.