

**KLASIFIKASI REMPAH MENGGUNAKAN EKSTRAKSI  
FITUR *SPEEDED-UP ROBUST FEATURES* (SURF) DENGAN  
METODE *NAÏVE BAYES***

**SKRIPSI**

**OLEH :**

**IRA SAFIRA**

**198160002**



**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS MEDAN AREA**

**2023**

**UNIVERSITAS MEDAN AREA**

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 16/1/24

Access From (repository.uma.ac.id)16/1/24

**KLASIFIKASI REMPAH MENGGUNAKAN EKSTRAKSI  
FITUR *SPEEDED-UP ROBUST FEATURES* (SURF) DENGAN  
METODE *NAÏVE BAYES***

**SKRIPSI**

Diajukan sebagai salah satu syarat untuk memperoleh

Gelar Sarjana (S1) di Fakultas Teknik

Universitas Medan Area

**OLEH :**

**IRA SAFIRA**

**198160002**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS MEDAN AREA**

**2023**

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

ii

Document Accepted 16/1/24

### LEMBAR PENGESAHAN

Judul Skripsi : Klasifikasi Rempah Menggunakan Ekstraksi Fitur *Speeded-Up Robust Features* (SURF) dengan Metode *Naïve Bayes*

Nama : IRA SAFIRA  
NPM : 198160002  
Fakultas : Teknik  
Prodi : Teknik Informatika

Disetujui Oleh

Komisi Pembimbing



Muhathir, S.T, M.Kom

Mengetahui



Dr. Rahmad Syah, S.Kom, M.Kom  
NIDN : 0105058804



R. M. Mulibno, S.Kom, M.Kom  
NIDN : 0109038902

Tanggal Lulus : 09 Agustus 2023

### HALAMAN PERNYATAAN

Saya menyatakan bahwa skripsi yang saya susun, sebagai syarat memperoleh gelar serjana merupakan hasil karya tulis saya sendiri. Adapun bagian-bagian tertentu dalam penulisan skripsi ini yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi-sanksi lainnya dengan peraturan yang berlaku, apabila di kemudian hari ditemukan adanya plagiat dalam skripsi ini.



**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS  
AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Medan Area, saya yang bertanda tangan di bawah ini:

Nama : Ira Safira  
NPM : 198160002  
Program Studi : Teknik Informatika  
Fakultas : Teknik  
Jenis Karya : Skripsi

demikian pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Medan Area Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul :

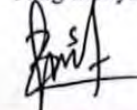
***Klasifikasi Rempah Menggunakan Ekstraksi Fitur Speeded-Up Robust Features dan Metode Naïve Bayes***

berserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-eksklusif ini Universitas Medan Area berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir/skripsi/tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Medan

Pada tanggal : 11 Agustus 2023

Yang menyatakan



( Ira Safira )

## ABSTRAK

Indonesia merupakan salah satu penghasil rempah-rempah berkualitas tinggi di dunia. Beraneka ragam rempah membuat masyarakat sulit untuk mengenali jenis rempah-rempah tersebut. Hal ini disebabkan rempah-rempah memiliki bentuk dan warna yang hampir sama. Penelitian ini berfokus pada klasifikasi citra rempah menggunakan ekstraksi fitur *Speeded-Up Robust Features* (SURF) dengan metode variasi *Naïve Bayes*. Tujuan dari penelitian ini adalah mengklasifikasikan citra rempah menggunakan ekstraksi fitur *Speeded-Up Robust Features* dengan metode variasi *Naïve Bayes* untuk memperoleh hasil klasifikasi yang baik. Data yang digunakan pada penelitian ini sebanyak 2250 sampel. Berdasarkan 3 pengujian dengan menggunakan jenis rempah yang berbeda pada penelitian ini. Dapat diketahui bahwa pengujian dengan 5 jenis rempah menggunakan Metode *Gaussian Naïve Bayes* mendapatkan hasil akurasi yang lebih baik dengan nilai akurasi sebesar 0.82, presisi sebesar 0.817, recall sebesar 0.815, f1-score sebesar 0.814, f-beta score sebesar 0.814, dan jaccard score sebesar 0.711. Dari hasil tersebut dapat disimpulkan bahwa ekstraksi fitur *Speeded-up Robust Features* akan bekerja optimal jika jumlah jenis rempah yang diklasifikasi sedikit dan hasil klasifikasi kurang optimal jika jumlah jenis rempah yang diklasifikasi besar. Untuk mendapatkan hasil klasifikasi yang lebih akurat dapat menggunakan metode lain dan ekstraksi fitur lainnya.

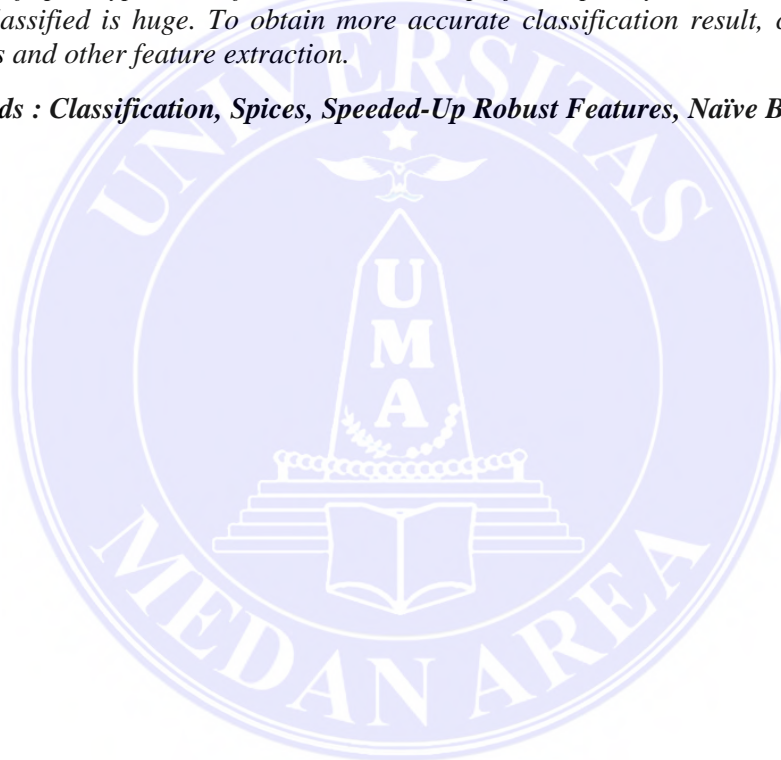
**Kata Kunci :** *Klasifikasi, Rempah, Speeded-Up Robust Features, Naïve Bayes*



## ABSTRACT

*Indonesia is one of the producers of high-quality spices in the world. The variety of spices drives a difficulty for a community to recognize the type of spice. This is because spices have almost the same shape and color. This study focuses on the classification of spice images using the extraction of Speeded-Up Robust Features (SURF) features with the Naïve Bayes variation method. The purpose of this research is to classification spice images using Speeded-Up Robust Features feature extraction with the Naïve Bayes variation method to obtain good classification results. The data used in this study was 2250 samples. Based on 3 tests using different types of spices in this study, it can be seen that testing 5 types of spices using Gaussian Nave Bayes method obtained the best acuuracy performance with an accuracy of 0.82, precision of 0.817, recall of 0.815, f1-score of 0.814, f-beta score of 0.814, and Jaccard score of 0.711. Based on these findings, it can be inferred that the Speeded-up Robust Features feature extraction works best when the number of spice types identified is small, and it performs poorly when the number of spice types classified is huge. To obtain more accurate classification result, can using other methods and other feature extraction.*

**Keywords :** *Classification, Spices, Speeded-Up Robust Features, Naïve Bayes*



## RIWAYAT HIDUP

Penulis lahir di Desa Bandar Khalipah pada tanggal 18 Maret 2002 dari Ayah Asrik dan Ibu Yanti. Penulis adalah anak ke-2 (kedua) dari 4 (empat) bersaudara.

Penulis pertama kali mengenyam Pendidikan dibangku SD Al-Mukmin pada tahun 2007-2013, meneruskan Pendidikan di Sekolah Menengah Pertama Negeri 29 Medan diselesaikan pada tahun 2013-2016, meneruskan Pendidikan SMK Swasta Teladan Medan pada tahun 2016-2019.

Pada tahun 2019 penulis lulus dari SMK Swasta Teladan Medan dan pada tahun 2019 terdaftar sebagai mahasiswa Fakultas Teknik Prodi Teknik Informatika Universitas Medan Area. Selama duduk dibangku perkuliahan penulis mengikuti berbagai kegiatan seperti Program Kreativitas Mahasiswa (PKM) dan penulis juga meraih beasiswa berprestasi dari Bank Indonesia tahun 2022. Penulis juga melaksanakan kerja praktek di Kantor Walikota Medan pada tahun 2022.

Medan, 11 November 2023

Ira Safira



## KATA PENGANTAR

Dengan mengucapkan puji dan syukur kepada Allah SWT atas segala karunia dan Hidayah-Nya, sehingga Tugas Akhir saya yang berjudul “Klasifikasi Rempah Menggunakan Ekstraksi Fitur *Speeded-Up Robust Features* (SURF) dan Metode Naïve Bayes” dapat diselesaikan dengan baik dan tepat waktu. Tujuan penulisan laporan ini adalah untuk memenuhi persyaratan dalam menyelesaikan Tugas Akhir Program Studi Teknik Informatika Universitas Medan Area.

Penyusunan laporan ini tidak terlepas dari bantuan beberapa pihak, oleh karena itu penulis hendak mengucapkan terima kasih sebesar-besarnya atas dukungan dan motivasi kepada :

1. Allah SWT yang telah memberikan saya kesehatan dan keselamatan selama melakukan studi.
2. Kedua Orang tua, Abang, Adik, Nenek serta Saudara yang telah memberikan doa, kasih sayang, semangat, dan dukungan yang tidak henti-hentinya merupakan anugerah terbesar dalam hidup
3. Bapak Prof. Dr. Dadan Ramdan, M.Eng, M.Sc, selaku Rektor Universitas Medan Area
4. Bapak Dr. Rahmad Syah, S.Kom, M.Kom, selaku Dekan Fakultas Teknik Universitas Medan Area
5. Ibu Susilawati, S.Kom, M.Kom Selaku Wakil Dekan Bidang Pendidikan Penelitian & Pengabdian Kepada Masyarakat Universitas Medan Area
6. Bapak Rizki Muliono, S.Kom, M.Kom, selaku Kepala Program Studi Teknik Informatika
7. Bapak Muhathir, S.T, M.Kom, selaku Dosen Pembimbing Tugas Akhir

8. Pasar Induk Lau Chi Medan yang telah mengizinkan saya untuk melakukan riset
9. Teman-teman seperjuangan yang telah memberikan semangat dan dukungannya
10. Semua Pihak yang tidak dapat disebutkan satu persatu yang terlibat dalam penyusunan Tugas Akhir ini sehingga dapat selesai dengan baik
11. Dan yang terakhir saya ucapkan terima kasih kepada diri sendiri yang sudah berjuang sejauh ini dalam menyelesaikan tugas akhir, terimakasih sudah bisa melewati titik ini, terima kasih tidak pernah berhenti berusaha, terima kasih atas semua yang telah diupayakan. Terimakasih *my self*.

*Tidak ada yang tidak mungkin selama kita ikhlas berdoa dan tak menyerah untuk berusaha.*

Penulis sangat menyadari bahwa dalam penyusunan Tugas Akhir ini masih jauh dari kata sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan agar kedepannya penulis bisa lebih baik lagi dalam Menyusun laporan ilmiah.

Medan, 11 November 2023

Ira Safira

NPM 198160002

## DAFTAR ISI

ABSTRAK .....	vi
ABSTRACT .....	vii
RIWAYAT HIDUP .....	viii
KATA PENGANTAR .....	ix
DAFTAR ISI .....	xi
DAFTAR TABEL .....	xiii
DAFTAR GAMBAR .....	xiv
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan .....	4
1.4 Batasan Masalah .....	4
1.5 Manfaat Penelitian .....	5
BAB II TINJAUAN PUSTAKA .....	7
2.1 Klasifikasi .....	7
2.2 Rempah .....	7
2.3 <i>Speeded-Up Robust Features (SURF)</i> .....	8
2.3.1 <i>Integral Image</i> .....	8
2.3.2 <i>Interest Point Detection</i> .....	8
2.3.3 <i>Feature Description</i> .....	9
2.3.4 <i>Feature Matching</i> .....	10
2.4 <i>Naïve Bayes</i> .....	10
2.5 Penelitian Terdahulu .....	11
BAB III METODOLOGI PENELITIAN .....	17
3.1 Alat dan Bahan Penelitian .....	17
3.1.1 <i>Hardware</i> .....	17
3.1.2 <i>Software</i> .....	18
3.2 Pengumpulan Data .....	18
3.3 Dataset .....	18
3.4 Data Analisis .....	19
3.5 Diagram Alur Penelitian .....	20
3.6 Arsitektur Penelitian .....	21

3.7 Metode Evaluasi .....	22
3.7.1 <i>Confusion Matrix</i> .....	22
3.7.2 <i>Performa Measure</i> .....	23
BAB IV HASIL DAN PEMBAHASAN .....	26
4.1 Hasil.....	26
4.1.1 Sampel Rempah.....	26
4.1.2 <i>Resizing</i> Citra .....	27
4.1.3 Ekstraksi Fitur <i>Speeded-up Robust Features</i> (SURF).....	27
4.1.4 Pra- Pemrosesan .....	28
4.1.5 Pengujian Metode <i>Naïve Bayes</i> .....	28
4.2 Pembahasan .....	35
BAB V KESIMPULAN DAN SARAN.....	38
5.1 Kesimpulan.....	38
5.2 Saran .....	39
DAFTAR PUSTAKA .....	40
LAMPIRAN.....	44

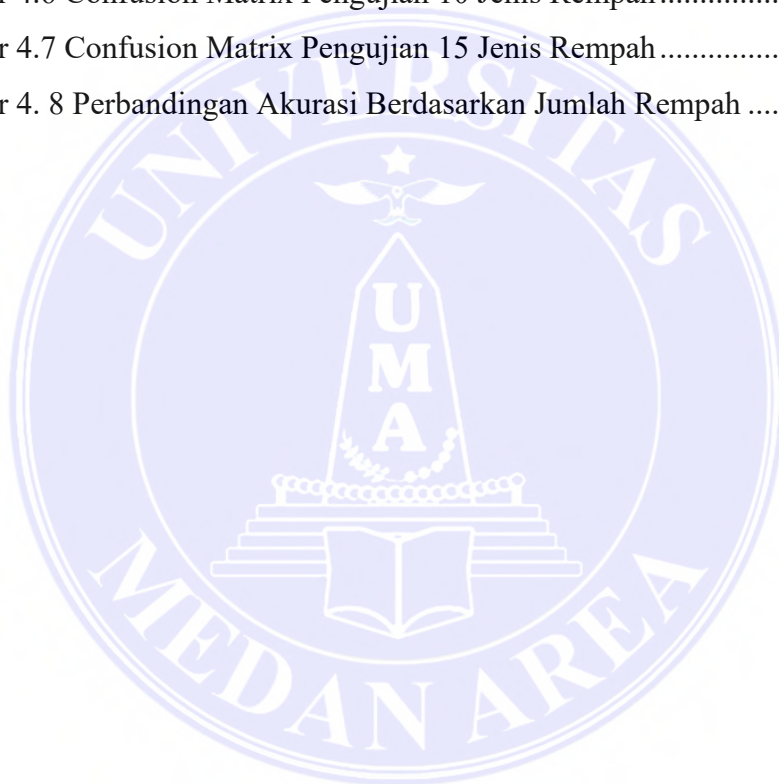
## DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu .....	11
Tabel 3. 1 <i>Hardware</i> .....	17
Tabel 3. 2 <i>Software</i> .....	18
Tabel 3. 3 Perbandingan Data .....	19
Tabel 3. 4 <i>Confusion Matrix</i> .....	22
Tabel 4. 1 Hasil Klasifikasi Pengujian dengan 5 jenis Rempah .....	30
Tabel 4. 2 Hasil Klasifikasi Pengujian dengan 10 jenis rempah.....	31
Tabel 4. 3 Hasil Klasifikasi Pengujian dengan 15 Jenis Rempah .....	33
Tabel 4. 4 Perbandingan penelitian ini dengan penelitian terdahulu .....	36



## DAFTAR GAMBAR

Gambar 3. 1 Diagram Alur Penelitian.....	20
Gambar 3. 2 Arsitektur Penelitian.....	21
Gambar 4.1 Sampel Rempah.....	26
Gambar 4.2 Citra Sebelum di resizing, (b) Citra Sesudah di resizing .....	27
Gambar 4.3 Hasil Ekstraksi Fitur SURF.....	28
Gambar 4.4 Hasil Citra ekstraksi ke array .....	28
Gambar 4.5 Confusion Matrix Pengujian 5 Jenis Rempah.....	29
Gambar 4.6 Confusion Matrix Pengujian 10 Jenis Rempah .....	31
Gambar 4.7 Confusion Matrix Pengujian 15 Jenis Rempah.....	32
Gambar 4. 8 Perbandingan Akurasi Berdasarkan Jumlah Rempah .....	34



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Indonesia merupakan salah satu penghasil rempah-rempah berkualitas tinggi di dunia. Rempah-rempah Indonesia memiliki nilai jual yang tinggi di pasar Eropa karena rasa, aroma dan kelezatannya (Hardoyono dkk., 2019). Rempah-rempah Indonesia menguasai 21,06% pasar rempah dunia di tahun 2013 (Yana dkk., 2018). Rempah-rempah telah digunakan dalam makanan selama bertahun-tahun sebagai penyedap, pewarna dan pengawet makanan (Tshabalala dkk., 2021).

Rempah-rempah merupakan hasil tanaman dengan aroma khas dan beragam manfaat kesehatan. Rempah adalah bagian tumbuhan yang mengandung berbagai bahan berkhasiat yang berasal dari biji, akar, umbi, rimpang, daun, batang, atau bagian tumbuhan lainnya (Sabila dkk., 2021). Rempah yang menggunakan rimpang antara lain kunyit, jahe, lengkuas, dan temu kunci, sedangkan yang menggunakan biji antara lain jintan, pala, merica, adas, dan ketumbar. Daun salam, daun kunyit, kemangi, daun jeruk, daun seledri, dan daun pandan adalah beberapa daun yang digunakan (Susiarti dkk., 2021).

Pada penelitian (Wulandari dkk., 2020) yang berjudul “Klasifikasi Citra Digital Bumbu Dan Rempah Dengan Algoritma *Convolutional Neural Network* (CNN)”, Sekitar 47% siswa di SMKN 9 Bandung belum mengenal bumbu dan rempah-rempah. Saat ini, perkembangan teknologi pemrosesan citra digital, memungkinkan untuk mengkategorikan bumbu dan rempah secara otomatis.

Dengan menggunakan klasifikasi citra, masalah tersebut dapat diatasi. Klasifikasi citra bertujuan untuk meniru pemahaman manusia terhadap informasi citra digital, sehingga komputer dapat mengklasifikasikan objek berupa citra dengan cara yang sama seperti yang dilakukan oleh manusia.

Beraneka ragam rempah membuat masyarakat sulit untuk mengenali jenis rempah-rempah. Hal ini disebabkan rempah-rempah memiliki bentuk dan warna yang hampir sama. Oleh sebab itu, untuk menyelesaikan masalah tersebut diperlukan suatu sistem untuk mengklasifikasikan jenis rempah, salah satunya adalah dengan memanfaatkan ekstraksi fitur SURF (*Speeded-Up Robust Features*) serta metode klasifikasi *Naïve Bayes*.

Pada penelitian yang dilakukan oleh (Batubara Nadya dkk., 2020) dengan judul “Klasifikasi Rempah Rimpang Berdasarkan Ciri Warna RGB dan Tekstur GLCM Menggunakan Algoritma Naïve Bayes”. Data yang digunakan sebanyak 80 citra dengan masing-masing rempah yaitu Kunyit 20 citra, Jahe 20 citra, Lengkuas 20 citra, dan Temulawak 20 citra. Teknik *Naive Bayes* selanjutnya digunakan untuk mengklasifikasikan hasil normalisasi data dengan *k-fold cross-validation*  $K=10$ . Berdasarkan hasil penelitian yang dilakukan diperoleh tingkat akurasi sebesar 52%.

SURF digunakan untuk mendeteksi objek, memasukkan gambar, klasifikasi, dan pembaharuan 3D (Wang dkk., 2019). Selain itu, *Speeded-up Robust Features* digunakan untuk mengekstraksi fitur gambar (Alwan & Ku-Mahamud, 2020). Ekstraksi Fitur SURF memiliki keunggulan yaitu mampu mendeteksi fitur suatu citra dengan handal dan cepat (Baresi Ariel & Dwi Atmaja, 2016). Selain itu, ada dua tahap dalam proses pembentukan fitur SURF yaitu tahap pertama



mendeteksi titik elemen yang meliputi gambar integral, pengenalan elemen berdasarkan *Matriks Hessian*, skala representasi ruang, dan lokasi titik elemen. Tahap kedua, yaitu mendeskripsikan titik elemen yang meliputi tugas orientasi dan *descriptor* berdasarkan jumlah respon Haar wavelet 2D (Yohannes dkk., 2020).

Ekstraksi Fitur SURF pernah digunakan dalam Klasifikasi Citra Sel Darah Putih dengan metode *Support Vector Machine*. Dalam penelitian ini, dataset yang digunakan yaitu dataset citra sel darah putih yang dilakukan dalam beberapa tahap yaitu tahap processing yang terdiri dari *crop*, *resize*, dan *saliency*. Metode Ekstraksi Fitur SURF dalam penelitian ini dapat memberikan *keypoint* untuk digunakan *Support Vector Machine* dalam mengenali jenis sel darah putih (Devella & Adi Putra, 2021).

Klasifikasi *Naïve Bayes* adalah metode yang menerapkan teorema bayes dengan menggabungkan informasi lama dan baru (Wibawa dkk., 2019). *Naïve Bayes* digunakan untuk klasifikasi teks karena *naïve bayes* merupakan klasifikasi sederhana dan efektif (Chen & Fu, 2018). Data dibuat dan diimplementasikan menggunakan teknik klasifikasi *Naïve Bayes* untuk menemukan dan menyusun informasi yang diperlukan untuk prediksi (Barik dkk., 2020). Metode *Naïve Bayes* dibagi menjadi beberapa tipe yaitu : *Multinomial Naïve Bayes*, *Bernoulli Naïve Bayes*, dan *Gaussian Naïve Bayes*.

Pada penelitian yang dilakukan oleh (Muhathir dkk., 2020) dengan judul “*Analysis Naïve Bayes in Classifying Fruit by Utilizing Hog Feature Extraction*”. Hasil penelitian tersebut mendapatkan hasil bahwa klasifikasi buah dengan

menggunakan metode *Naïve Bayes* dan ekstraksi fitur Hog belum memperoleh hasil klasifikasi yang maksimal dengan akurasi sebesar 56,52%.

Berdasarkan kelebihan dari ekstraksi fitur SURF yang mampu mendeteksi fitur dari suatu citra dengan handal dan cepat, serta metode klasifikasi *naïve bayes* yang mampu mengambil informasi baru dengan cepat dalam berbagai masalah klasifikasi. Oleh karena itu, penulis mencoba untuk melakukan penelitian tentang Klasifikasi Rempah Menggunakan Ekstraksi Fitur SURF dan Metode *Naïve Bayes*. Dalam penelitian ini, menerapkan metode ekstraksi fitur SURF sebagai ekstraksi fitur untuk memperoleh fitur dari citra, dan metode *Naïve Bayes* sebagai metode klasifikasi untuk mengklasifikasikan jenis rempah.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang sudah disebutkan diatas, maka dapat dirumuskan suatu permasalahan yaitu bagaimana mengklasifikasikan citra rempah menggunakan ekstraksi fitur SURF dengan variasi metode *Naïve Bayes* untuk memperoleh hasil klasifikasi yang baik.

## 1.3 Tujuan

Tujuan dari penelitian ini adalah mengklasifikasikan citra rempah menggunakan ekstraksi fitur SURF dengan variasi metode *Naïve Bayes* untuk memperoleh hasil klasifikasi yang baik.

## 1.4 Batasan Masalah

Adapun batasan masalah yang terdapat dalam penelitian ini adalah sebagai berikut :

1. Bahasa pemrograman yang digunakan dalam sistem yaitu Bahasa Pemrograman *Python*.
2. Citra yang digunakan adalah rempah jahe, lengkuas, kunyit, temulawak, kencur, kemiri, pala, adas manis, ketumbar, jinten, temu kunci, asam gelugur, kayu manis, bunga lawang, dan daun jeruk
3. Dataset yang digunakan sebanyak 2250 sampel dengan melakukan riset secara langsung di Pasar Induk Lau Chi, Kelurahan Kemenangan Tani, Kecamatan Medan Tuntungan, Kota Medan, Provinsi Sumatera Utara.
4. Data yang digunakan dikelompokkan menjadi dua data yaitu Data *Training* dan Data *Testing* . Perbandingan data *training* dan data *testing* adalah 80:20.
5. Metode yang digunakan adalah *Gaussian Naïve Bayes*, *Bernoulli Naïve Bayes* dan *Multinomial Naïve Bayes*.
6. Ekstraksi fitur yang digunakan adalah SURF.
7. Format sampel citra yang digunakan .JPG.

### 1.5 Manfaat Penelitian

Adapun manfaat penelitian yang akan diperoleh dari penelitian adalah sebagai berikut :

1. Untuk memberikan ilmu pengetahuan mengenai cara implementasi metode *Naïve Bayes* dan Ekstraksi Fitur SURF dalam mengklasifikasikan jenis rempah.

2. Untuk mengetahui tingkat akurasi yang didapat menggunakan variasi *Naïve Bayes* dan *Ekstraksi Fitur SURF* dalam mengklasifikasikan jenis rempah.
3. Sebagai bahan pembelajaran dan referensi dalam pengembangan ilmu pengetahuan yang berkaitan dengan metode *Naïve Bayes* dan *Ekstraksi Fitur SURF*.



## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Klasifikasi

Klasifikasi adalah suatu metode yang memiliki fungsi untuk menentukan setiap atribut dari sekumpulan objek. Klasifikasi merupakan sebuah proses mengkategorikan data (Putry Naisah & Sari Betha, 2022).

Model kemudian diterapkan untuk mengkategorikan atribut yang kelasnya tidak diketahui dengan tujuan mengidentifikasi model dari training set yang membedakan kelas yang tepat. Ada beberapa metode klasifikasi yang digunakan dalam data mining. Pada klasifikasi ada target variable kategori (Firdaus dkk., 2022).

#### 2.2 Rempah

Indonesia merupakan salah satu negara dengan sumber daya alam terbesar kedua setelah Brazil. Salah satu sumber daya alam yang paling populer di Indonesia adalah rempah-rempahnya yang melimpah (Hajriansyah, 2023). Rempah-rempah merupakan sumber daya alam hayati yang telah lama dimanfaatkan dalam kehidupan manusia. Rempah-rempah dapat digunakan dalam masakan sebagai bumbu, penguat rasa, pengharum, dan pengawet. Rempah yang berasal dari batang, daun, kulit kayu, umbi, rimpang, akar, biji, bunga, atau bagian tanaman lainnya sudah terkenal di dunia kuliner. (Robi dkk., 2019).

Rempah-rempah yang berasal dari rimpang termasuk jahe, kunyit, lengkuas, temulawak dan kencur. Daunnya merupakan bagian tanaman yang paling sering

digunakan sebagai bumbu masakan, terutama untuk menambah cita rasa dan aroma pada makanan. Daun yang biasa digunakan antara lain jeruk nipis, daun salam, seledri, dan daun pandan. (Wandira dkk., 2019).

### **2.3 Speeded-Up Robust Features (SURF)**

Peneliti dari ETH Zurich, Herbert Bay, menerbitkan algoritma SURF untuk pertama kalinya pada tahun 2006. Selain itu, dua rekannya, Tinne Tuytelaars dari Katholieke University Leuven dan Luc Van Gool, membantu Herbert Bay dalam pengembangannya. *Speeded-up Robust Features* (SURF) dapat dengan cepat dan handal mengidentifikasi fitur. *Scale Invariant Feature Transform* (SIFT), yang pertama kali muncul pada tahun 1999, berfungsi sebagai sumber inspirasi untuk algoritma ini, khususnya selama tahap penskalaan representasi ruang. (Feldi & Yuswanita, 2021).

Implementasi algoritma SURF dibagi menjadi beberapa tahapan yaitu *integral image*, *interest point detection*, *feature description*, dan *feature matching*:

#### **2.3.1 Integral Image**

*Integral image* adalah langkah pertama dalam algoritma SURF. Gambar dalam format skala abu-abu (*grayscale*) adalah citra input. Gambar integral yang mewakili hasil kamera dibuat pada titik ini, menghasilkan representasi gambar (Adi dkk., 2017).

#### **2.3.2 Interest Point Detection**

Pada citra digital, *point of interest detection* digunakan untuk memilih titik yang stabil terhadap gangguan lokal atau global dan mengandung banyak informasi.

Algoritma SURF memilih *blob detection point of interest detector* dengan sifat

*scale-invariant*. *Blob detection* merupakan teknik lain untuk mendeteksi bentuk . Teknik ini digunakan untuk melakukan pendeteksian dari kumpulan pixel yang mempunyai warna berbeda (lebih terang atau lebih gelap) (Yusuf, 2021). Sebuah blob citra digital adalah area yang tetap sama atau berubah dalam rentang tertentu. Penentu Hessian ( $D_oH$ ) gambar digunakan dalam persamaan 2.1 di bawah ini untuk melakukan pemrosesan deteksi *blob* :

$$H(x, \sigma) = \begin{pmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{pmatrix} \dots\dots\dots (2.1)$$

Posisi dan skala deskriptor dipilih menggunakan determinan matriks Hessian. Untuk titik  $x=(x, y)$  pada citra  $I$  , matriks Hessian  $H(\sigma, x)$  didefinisikan pada persamaan (2.1). Determinan matriks Hessian menunjukkan berapa banyak respon yang ada dan berapa banyak perubahan lokal yang ada.  $L_{xx}(\sigma, x) = \frac{\partial^2 g(\sigma)}{\partial^2 x}$ ,  $L_{xy}(\sigma, x)$  dan  $L_{yy}(\sigma, x)$  merepresentasikan nilai konvolusi citra  $I$  pada titik  $x$  dengan turunan Gaussian orde kedua.

### 2.3.3 Feature Description

Bagian-bagian dari suatu citra yang mengandung banyak informasi dikenal sebagai fitur, dan enkripsi objek dimulai dengan fitur tersebut. Proses pendeteksian fitur ini bertujuan untuk memberikan gambaran tentang fitur-fitur citra yang diamati. Langkah pertama yang dilakukan adalah menentukan orientasi dominan citra di tempat menarik. Selanjutnya, membuat area pengambilan dan mencari fitur korespondensi perbandingan gambar (Adi dkk., 2017).

### 2.3.4 *Feature Matching*

Pada tahap ini, adalah membandingkan fitur dari hasil perhitungan cara sebelumnya tetapi apabila ada perbedaan kontras, perbedaan ini dapat dideteksi dengan tanda *trace matriks Hessian* (Pratama dkk., 2021).

## 2.4 *Naïve Bayes*

*Naive Bayes* adalah klasifikasi yang dibuat menggunakan metode probabilitas dan statistik yang ditemukan oleh ilmuwan Inggris Thomas Bayes. Dengan asumsi bahwa teorema Bayes kurang berlaku untuk aplikasi dunia nyata dan semua atribut bersifat independen mengingat nilai variabel kelas ini, karakterisasi disediakan untuk kenyamanan, namun algoritma *Naïve Bayes* ini biasanya bekerja dengan baik dan mengambil informasi baru dengan cepat dalam berbagai masalah klasifikasi (Muhathir dkk., 2020).

*Naive Bayes* merupakan salah satu algoritma klasifikasi data. Di bidang data mining, algoritma ini menempati peringkat 10 besar. Klasifikasi probabilitas statistik sederhana dengan algoritma *Naive Bayes* menggunakan frekuensi dan kombinasi nilai dalam kumpulan data tertentu, algoritma ini menghitung sekumpulan probabilitas. Menghitung frekuensi setiap nilai fitur di kelas dari kumpulan data pelatihan menghasilkan probabilitas fitur tertentu dalam data, yang direpresentasikan sebagai anggota urutan probabilitas. Kumpulan data yang digunakan untuk melatih klasifikasi algoritma adalah kumpulan data pelatihan. Untuk memprediksi nilai yang tidak diketahui, proses pelatihan menggunakan nilai yang diketahui sebelumnya (Wibawa dkk., 2019).



Metode *Naive Bayes* merupakan salah satu algoritma kecerdasan buatan yang digunakan untuk klasifikasi. Algoritma *Naive Bayes* ini dikenal pada implementasi *machine learning* yang menganggap bahwa dengan setiap nilai atribut yang mempengaruhi nilai kelas tidak tergantung pada nilai atribut yang mempengaruhi nilai kelas lainnya (Hayami dkk., 2022). Persamaan teorema bayes sebagai berikut:

$$P = (H|X) = \frac{P(X|H)P(H)}{P(X)} \dots\dots\dots (2.2)$$

Keterangan :

X = Evidensi dengan kelas yang belum diketahui

H = Kelas spesifik suatu hipotesis dari data X

P(H|X) = Probabilitas Hipotesis H pada evidensi X

P(H) = Probabilitas Hipotesis H

P(X) = Probabilitas Evidensi X

### 2.5 Penelitian Terdahulu

Penelitian terdahulu adalah salah satu acuan mendasar saat melakukan penelitian. Ide yang akan digunakan dalam penelitian akan diperluas dan diperdalam sebagai hasil penelitian sebelumnya. Penelitian terdahulu dapat dilihat pada Tabel 2.1.

Tabel 2.1 Penelitian Terdahulu

No	Judul	Keterangan
1.	Analisis Pengaruh Data Augmentasi Pada Klasifikasi	Penelitian ini untuk membandingkan suatu arsitektur dengan menerapkan

	Bumbu Dapur Menggunakan <i>Convolutional Neural Network</i> (D & Kusumaningtyas, 2022)	augmentasi dimana untuk mengklasifikasikan jenis bumbu dapur dan rempah yang memiliki ragam yang hampir sama. Penelitian ini menggunakan metode <i>Convolutional Neural Network</i> (CNN) dengan pengujian 80 citra data testing. Hasil yang diperoleh dari penelitian ini adalah dengan tidak menerapkan augmentasi mendapatkan akurasi sebesar 54% dan menerapkan augmentasi mendapatkan akurasi sebesar 80% .
2	Klasifikasi Jenis Rempah-rempah Berdasarkan Fitur Warna RGB dan Tekstur Menggunakan Algoritma <i>K-Nearest Neighbor</i> (Kaharuddin dkk., 2019)	Penelitian ini bertujuan untuk mengklasifikasikan rempah-rempah berdasarkan warna RGB dan Tekstur menggunakan Algoritma <i>K-Nearest Neighbor</i> (KNN) dan pengukuran jarak menggunakan <i>Euclidean Distance</i> . Penelitian ini melakukan pengujian sebanyak 30 kali dengan tingkat kebenaran dengan K=1 sebesar 76%, K=3 sebesar 67%, dan K=5 sebesar 63%.
3	Komparasi Klasifikasi Jenis Tanaman Rimpang	untuk mengklasifikasikan citra rimpang jahe, kencur, Kunci, kunyit dan Laos. hasil

	Menggunakan <i>Principal Component Analysis, Support Vector Machine, K-Nearest Neighbor</i> dan <i>Decision Tree</i> (Mayasari dkk., 2022)	pengujian multiclass yang menggunakan teknik <i>Decision Tree, K-Nearest Neighbor</i> (KNN), dan <i>Support Vector Machine</i> (SVM). Terbukti akurasi klasifikasi multiclass yang dihasilkan adalah 87,3% untuk <i>Decision Tree</i> , 90,02% untuk <i>K-Nearest Neighbor</i> (KNN), dan 98,2% untuk <i>Support Vector Machine</i> (SVM) (Linear SVM)
4	<i>Comparison of SURF and HOG extraction in classifying the blood image of malaria parasites using SVM</i> (Muhathir dkk., 2019)	Penelitian ini digunakan sebagai pengidentifikasi pola untuk parasit dan sel yang tidak terinfeksi dan <i>Support Vector Machine</i> (SVM) digunakan sebagai metode klasifikasi. Hasil yang diperoleh pada penelitian ini dengan menggunakan ekstraksi fitur <i>Speeded-up Robust Features</i> (SURF) diperoleh akurasi sebesar 85,83% sedangkan dengan menggunakan ekstraksi fitur <i>Histogram of Gradient Oriented</i> (HOG) diperoleh akurasi sebesar 83,50%.
5	Klasifikasi Tanaman Obat Herbal Menggunakan Metode	klasifikasi tanaman obat herbal menggunakan metode <i>Support Vector Machine</i> dilakukan dengan pengambilan

	<p><i>Support Vector Machine</i> (Arifin dkk., 2021)</p>	<p>parameter pembeda yang digunakan yaitu warna menggunakan metode <i>Color Histogram</i>, bentuk menggunakan metode Sobel serta tekstur menggunakan metode <i>Gray Level Co-occurrence Matrix</i>. Data yang digunakan sebanyak 430 untuk masing-masing rempah yang digunakan yaitu jahe, kunyit, lengkuas dan temulawak berjumlah 1720 citra. Data tersebut dibagi menjadi dua yaitu data training 80% dan data testing 20%. Hasil akurasi yang diperoleh sebesar (93.3%, 73%, 60%, dan 53.3%)</p>
8	<p>Klasifikasi Rempah Rimpang Berdasarkan Ciri Warna RGB dan Tekstur GLCM Menggunakan Algoritma <i>Naive Bayes</i> (Batubara Nadya dkk., 2020)</p>	<p>Penelitian ini membahas Pengklasifikasian beberapa jenis rempah-rempah dengan menggunakan metode <i>Naive Bayes</i>. Penelitian ini menggunakan ekstraksi fitur tekstur GLCM dan warna RGB. Data yang digunakan sebanyak 80 citra dengan masing-masing rempah yaitu kunyit 20 citra, jahe 20 citra, lengkuas 20 citra, dan temulawak 20 citra. Teknik <i>Naive Bayes</i> selanjutnya digunakan untuk mengklasifikasikan hasil normalisasi data</p>

		dengan <i>k-fold cross-validation</i> $K=10$ untuk mendapatkan akurasi sebesar 52%. Sebagai hasil dari penggunaan 4 objek, akurasi adalah 52%.
9	<i>Performance Comparison of Boosting Algorithms in Spices Classification Using Histogram of Oriented Gradient Feature Extraction</i> (Muhathir dkk., 2023)	Penelitian ini membahas tentang klasifikasi rempah-rempah menggunakan metode ekstraksi fitur <i>Histogram of Oriented Gradient</i> (HOG) dan algoritma boosting. Data yang digunakan sebanyak 750 dataset dengan 5 jenis rempah berbeda (adas manis, cengkeh, jinten, kapulaga, dan kemiri). Hasil pengujian diperoleh bahwa model <i>XGB Classifier</i> mencapai performa terbaik, dengan presisi 0.811, recall 0.809, dan F1-score 0.809, sedangkan model <i>Adaboost Classifier</i> memiliki performa terendah, dengan presisi 0.709, recall 0.689, dan F1-score 0.682.
10	<i>Compares the effectiveness of the bagging method in classifying spices using the histogram of oriented</i>	Penelitian ini berfokus pada klasifikasi rempah menggunakan fitur ekstraksi <i>Histogram of Oriented Gradient</i> (HOG) dan metode bagging. Data yang digunakan sebanyak 750 data. Data dibagi menjadi

<p><i>gradient feature extraction technique</i> (Muhathir, 2023)</p>	<p>data <i>training</i> dan data <i>testing</i> dengan pembagian data 80:20. Penelitian ini menggunakan 5 jenis rempah berbeda yaitu adas manis, cengkeh, jinten, kapulaga, dan kemiri. Hasil penelitian menunjukkan bahwa model <i>Random Forest</i> mencapai performa terbaik, dengan presisi sebesar 0.861, recall sebesar 0.8633, F1-Skor 0.8587, F2-Skor sebesar 0.8607, Jaccard sebesar 0.7694 , dan nilai akurasi sebesar 0.8733, sedangkan dengan menggunakan <i>Extra Tree Classifier</i> mendapatkan kinerja terendah dengan presisi sebesar 0.7034, recall sebesar 0.7958 , F1-Skor sebesar 0.7037, F2-Skor sebesar 0.7047, Jaccard skor sebesar 0.5635, dan nilai akurasi sebesar 0.72</p>
--	--

## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Alat dan Bahan Penelitian

Dalam melakukan penelitian dengan judul “Klasifikasi Rempah menggunakan Ekstraksi Fitur *Speeded-Up Robust Features* dengan Metode *Naïve Bayes*” ini, peneliti menggunakan peralatan dan bahan yang mendukung proses kerja penelitian ini supaya penelitian ini berjalan dengan lancar. Peralatan dan bahan pendukung yang digunakan adalah sebagai berikut :

##### 3.1.1 Hardware

*Hardware* atau perangkat keras yang digunakan pada penelitian ini memiliki spesifikasi ditunjukkan pada Tabel 3.1 sebagai berikut :

Tabel 3.1 *Hardware*

No	Hardware	Deskripsi
1	Device	Laptop Asus-INMG2J46
2	Processor	Intel(R) Celeron(R) N4020 CPU @ 1.10GHz 1.10 GHz
3	SSD	239 GB
4	RAM	4.00 GB
5	Kamera	Smartphone Redmi 9

### 3.1.2 Software

*Software* atau perangkat lunak yang digunakan pada penelitian ini memiliki spesifikasi ditunjukkan pada Tabel 3.2 sebagai berikut :

Tabel 3.2 *Software*

No	Software	Deskripsi
1	Sistem Operasi	Windows 11 Home Single Language
2	Open CV	Version 3.4.2
3	Python	Bahasa Pemrograman
4	Kamera	13 MP

### 3.2 Pengumpulan Data

Pada penelitian ini data yang diperlukan berupa citra rempah. Data yang diperlukan pada peneliti dilakukan dengan mengumpulkan data dengan melakukan riset secara langsung di Pasar Induk Lau Chi, Kelurahan Kemenangan Tani, Kecamatan Medan Tuntungan, Kota Medan, Provinsi Sumatera Utara, peneliti melakukan proses pengambilan gambar untuk kebutuhan *dataset* dengan membutuhkan waktu selama kurang lebih 1 bulan.

### 3.3 Dataset

Jenis data yang digunakan pada penelitian ini adalah jenis data primer. Citra rempah yang digunakan yaitu citra rempah jahe, lengkuas, kencur, kunyit, ketumbar, kemiri, temulawak, pala, jinten, adas manis, bunga lawang, temu kunci, kayu manis, asam glugur, dan daun jeruk. Citra rempah diambil dengan menggunakan kamera smartphone Redmi 9, dengan format .JPG.



### 3.4 Data Analisis

Setelah melalui proses pengumpulan data, diperoleh 2250 data citra rempah. Masing-masing sampel yang digunakan sebanyak 150 jenis rempah.

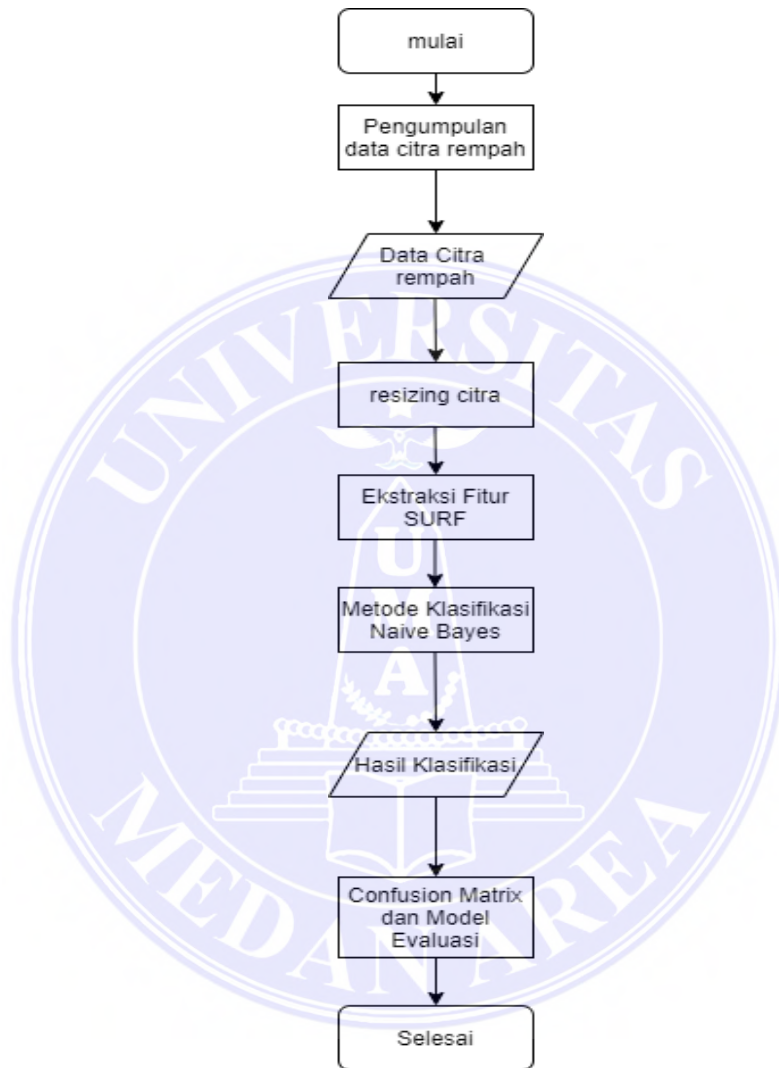
Data dalam penelitian ini dibagi menjadi dua data, data pertama adalah data *training* dan data kedua adalah data *testing*. Perbandingan data *training* dan data *testing* adalah 80:20. Tabel 3.3 menunjukkan kelengkapan informasi perbandingan data.

Tabel 3.3 Perbandingan Data

Jenis Rempah	Data <i>Training</i>	Data <i>Testing</i>
Jahe	120	30
Lengkuas	120	30
Kencur	120	30
Kunyit	120	30
Ketumbar	120	30
Kemiri	120	30
Temulawak	120	30
Pala	120	30
Jinten	120	30
Adas Manis	120	30
Bunga Lawang	120	30
Temu Kunci	120	30
Kayu Manis	120	30
Asam Glugur	120	30
Daun Jeruk	120	30

### 3.5 Diagram Alur Penelitian

Diagram alur penelitian yang di modelkan di dalam penelitian ini diilustrasikan pada Gambar 3.1.



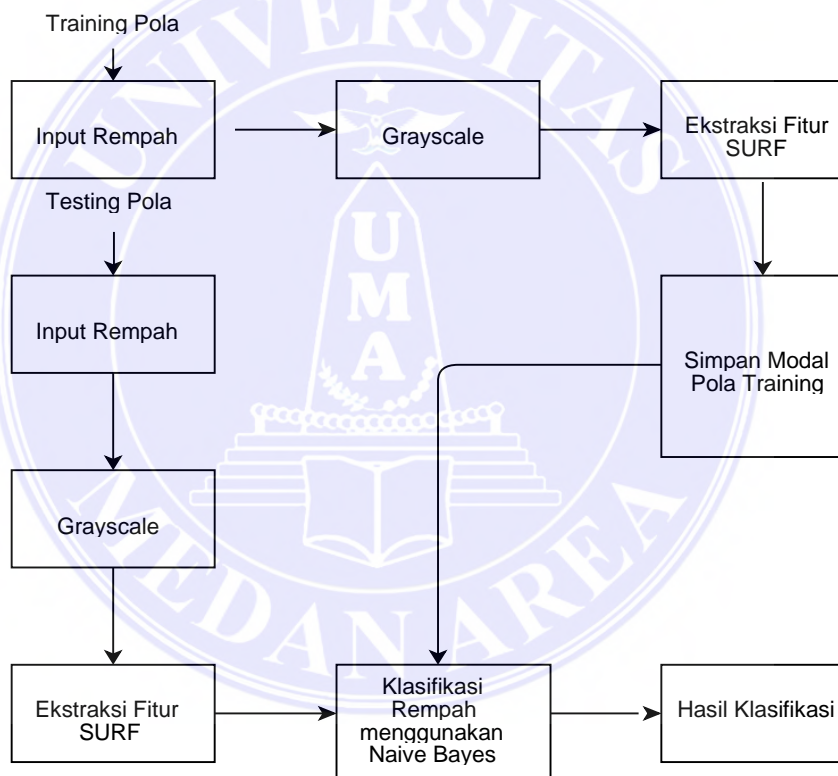
Gambar 3.1 Diagram Alur Penelitian

Pada Gambar 3.1 merupakan proses alur penelitian ini, pada proses pertama melakukan pengumpulan data citra rempah, Selanjutnya, setelah data dikumpulkan proses berikutnya menginputkan citra rempah kemudian data citra rempah tersebut akan dilakukan proses resizing agar ukuran pixel pada citra tersebut tidak terlalu besar Selanjutnya citra rempah tersebut akan dilakukan proses ekstraksi fitur

menggunakan SURF. Setelah itu, akan masuk ke proses klasifikasi dengan menggunakan metode *Naïve Bayes*, setelah proses klasifikasi selesai maka akan memperoleh hasil klasifikasi confusion matriks dan model evaluasi yang merupakan proses akhir dari alur penelitian.

### 3.6 Arsitektur Penelitian

Arsitektur penelitian yang dimodelkan dalam penelitian ini diilustrasikan pada Gambar 3.2.



Gambar 3.2 Arsitektur Penelitian

Arsitektur penelitian yang akan dilakukan menggunakan dua proses ditunjukkan pada Gambar 3.2. Pada proses *training* data, terlebih dahulu akan menginputkan citra rempah, setelah itu citra rempah akan masuk ke dalam proses *grayscale* atau keabuan, Kemudian, selanjutnya citra rempah akan diekstraksi menggunakan fitur SURF, dan informasi yang diambil kemudian disimpan sebagai

model pola informasi untuk digunakan dalam proses *testing*, sedangkan pada proses *testing* citra rempah yang telah dimasukkan dan telah di proses ke dalam *grayscale* atau keabuan selanjutnya citra rempah akan di ekstraksi menggunakan fitur SURF kemudian akan masuk kedalam proses pencocokan model dengan memanfaatkan metode *Naive Bayes* sebagai klasifikasi. Setelah menguji semua data hasil klasifikasi *Naive Bayes*, Persamaan (3.1), (3.2), (3.3), (3.4), (3.5), (3.6) digunakan untuk memproses kinerja algoritma.

### 3.7 Metode Evaluasi

Pada penelitian ini metode evaluasi yang digunakan dibagi menjadi dua bagian yaitu *Confusion Matrix* dan *Performa Measure*.

#### 3.7.1 Confusion Matrix

*Confusion Matrix* adalah Pengklasifikasi jumlah data yang lulus uji dan jumlah data yang tidak lulus uji ditunjukkan dalam tabel *confusion matrix* (Normawati & Prayogi, 2021). *Confusion Matrix* adalah sebuah teknik yang digunakan untuk mengembangkan matrik evaluasi yang dinilai pada tingkat yang dihasilkan oleh data pengujian (Mulyana & Wibowo, 2023) .

Tabel 3.4 *Confusion Matrix*

		Kelas Sebenarnya	
		Positif	Negatif
Kelas Prediksi	Positif	True Positive	False Positive
	Negatif	False Negative	True Negative

Jumlah data positif yang diklasifikasikan dengan benar oleh sistem disebut sebagai TP, atau *True Positive*. Jumlah data negatif yang diklasifikasikan dengan

benar oleh sistem disebut sebagai TN, atau *True Negative*. *False Negative* atau FN adalah jumlah data negatif yang dianggap salah oleh sistem. Jumlah data positif yang salah diklasifikasikan oleh sistem sebagai FP atau *False Positive* (Pratiwi dkk., 2020).

### 3.7.2 Performa Measure

Untuk mengukur kemampuan klasifikasi ada 6 (enam) nilai dalam klasifikasi yaitu nilai *accuracy*, *precision*, *recall*, *F-1 Score*, *F-Beta Score*, dan *Jaccard Score*.

#### 3.7.2.1 Accuracy

Sejauh mana sistem mampu mengklasifikasikan data dengan benar ditunjukkan oleh nilai akurasi. Sebaliknya, nilai akurasi adalah perbandingan data yang diklasifikasikan dengan benar dengan semua data. Persamaan berikut dapat digunakan untuk menentukan akurasi skor.

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN} \dots\dots\dots (3.1)$$

#### 3.7.2.2 Precision

Nilai presisi menunjukkan seberapa banyak data kategori positif yang diklasifikasikan dengan benar dengan membagi jumlah data kategori positif. Persamaan berikut dapat digunakan untuk mencapai presisi.

$$Precision = \frac{TP}{TP + FP} \dots\dots\dots (3.2)$$

#### 3.7.2.3 Recall

Nilai Recall menunjukkan persentase data kategori positif yang diklasifikasikan dengan benar oleh sistem ditunjukkan oleh nilai perolehan

kembali. Persamaan berikut dapat digunakan untuk mendapatkan nilai recall (Proboningrum & Sidauruk, 2021).

$$Recall = \frac{TP}{TP + FN} \dots\dots\dots (3.3)$$

#### 3.7.2.4 F1-Score

F1-Score menunjukkan nilai perbandingan rata-rata precision dan recall. Persamaan berikut dapat digunakan untuk mendapatkan nilai F1-Score.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \dots\dots\dots (3.4)$$

#### 3.7.2.5 F-Beta Score

Rata-rata nilai presisi dan nilai recall adalah skor F-beta, dengan nilai 1 mewakili presisi dan recall dan nilai 0 mewakili skor terburuk. Persamaan F-Beta Score adalah sebagai berikut (Masum dkk., 2022)

$$F_{\beta} = (1 + \beta^2) \frac{Precision \times Recall}{(\beta^2 \times Precision) + Recall} \dots\dots\dots (3.5)$$

Parameter F-beta menentukan bobot yang diberikan pada presisi dan recall ketika  $\beta = 1$ . Jika kita ingin memberikan bobot lebih pada presisi, kita dapat memilih 1, sedangkan nilai  $> 1$  memberikan bobot lebih pada recall.

#### 3.7.2.6 Jaccard Score

Jaccard Score adalah salah satu alat yang digunakan untuk mencari tahu seberapa mirip dua objek. Jaccard adalah alat yang berguna untuk pengelompokan dokumen. Sering kali, Jaccard digunakan untuk membandingkan dokumen dan mencari tahu seberapa mirip dua hal atau dokumen (Widiandana dkk., 2020) .

Persamaan Jaccard adalah sebagai berikut.

$$Jaccard (X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \dots\dots\dots (3.6)$$

Keterangan :

TP = *True Positive*

TN = *True Negative*

FP = *False Positive*

FN = *False Negative*



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan penelitian yang dilakukan, maka dapat disimpulkan sebagai berikut :

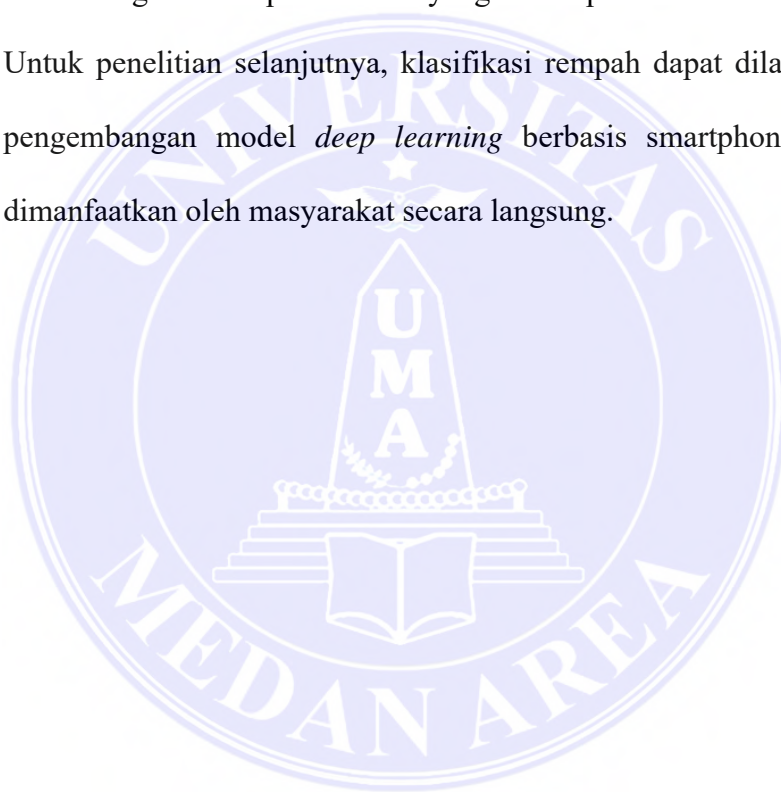
1. Ekstraksi Fitur SURF dapat bekerja dengan baik jika jumlah objek yang diklasifikasi dalam kategori sedikit akan tetapi jika diklasifikasi dengan objek dalam kategori banyak tidak dapat memberikan kontribusi yang maksimal.
2. Hasil klasifikasi rempah pada pengujian 5 jenis rempah dengan menggunakan metode *Gaussian Naïve Bayes* memperoleh tingkat akurasi yang lebih baik dengan akurasi sebesar 82%, pada pengujian 10 jenis rempah dengan menggunakan metode *Gaussian Naïve Bayes* memperoleh tingkat akurasi yang lebih baik dengan akurasi sebesar 67%, dan pada pengujian dengan 15 jenis rempah dengan metode *Gaussian Naïve Bayes* memperoleh tingkat akurasi yang lebih baik dengan akurasi sebesar 50,4%. Sedangkan dengan menggunakan metode Bernoulli Naïve Bayes memperoleh hasil yang lebih rendah.
3. Hasil penelitian ini menunjukkan tingkat kinerja yang cukup baik dalam mengklasifikasikan rempah-rempah menggunakan ekstraksi fitur SURF dan metode Naïve Bayes.



## 5.2 Saran

Berdasarkan hasil penelitian yang dilakukan, peneliti memiliki beberapa saran sebagai berikut ;

1. Dapat menambahkan jenis rempah yang ada dibelahan dunia seperti kluwak, saffron, garam masala, kauri methi, paprika, vanili dan lainnya. serta menggunakan variasi dari ekstraksi fitur lainnya seperti *Sift*, *Hog* dan lain-lain agar mendapatkan hasil yang lebih optimal.
2. Untuk penelitian selanjutnya, klasifikasi rempah dapat dilakukan dengan pengembangan model *deep learning* berbasis smartphone, agar dapat dimanfaatkan oleh masyarakat secara langsung.



## DAFTAR PUSTAKA

- Adi, F. F., Ichwan, M., & Miftahuddin, Y. (2017). Implementasi Algoritma Speeded Up Robust Features (SURF) Pada Pengenalan Rambu-Rambu Lalu Lintas. *Jurnal Teknik Informatika dan Sistem Informasi*, 3(3), 575–587.
- Alwan, H. B., & Ku-Mahamud, K. R. (2020). Cancellable face template algorithm based on speeded-up robust features and winner-takes-all. *Multimedia Tools and Applications*, 79(39–40). <https://doi.org/10.1007/s11042-020-09319-1>
- Arifin, Hendryli, J., & Dyah Erny, H. (2021). Klasifikasi Tanaman Obat Herbal Menggunakan Metode Support Vector Machine. *Journal of Computer Science and Information Systems*, 5(1), 25–35. <https://drive.google.com/drive/folders/17Dk6zNFN7WgpI0WvcUA5MeE4xXKaZp75?usp=sha>
- Baresi Ariel, M., & Dwi Atmaja, R. (2016). *Implementasi Metode Speed Up Robust Feature dan Scale Invariant Feature Transform untuk Identifikasi Telapak Kaki Individu* (Vol. 3, Nomor 4).
- Barik, L., Alrababah, A. A., & Al-Otaibi, Y. D. (2020). Enhancing Educational Data Mining based ICT Competency among e-Learning Tutors using Statistical Classifier. *IJACSA International Journal of Advanced Computer Science and Applications*, 11(3), 561–568. [www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)
- Batubara Nadya, P., Widiyanto, D., & Chamidah, N. (2020). Klasifikasi Rempah Rimpang Berdasarkan Ciri Warna RGB Dan Tekstur GLCM Menggunakan Algoritma Naive Bayes. *Jurnal Informatika Edisi ke-16*, 16(3), 156–163.
- Chen, H., & Fu, D. (2018). An Improved Naïve Bayes Classifier for Large Scale Text. In *2018 2nd International Conference on Artificial Intelligence: Technologies and Applications (ICAITA 2018)*, 146(Atlantis press), 33–36.
- D, W. M. P., & Kusumaningtyas, A. P. (2022). Analisis Pengaruh Data Augmentasi Pada Klasifikasi Bumbu Dapur Menggunakan Convolutional Neural Network. *Jurnal Media Informatika Budidarma*, 6(4), 2022–2031. <https://doi.org/10.30865/mib.v6i4.4201>
- Devella, S., & Adi Putra, C. (2021). Penggunaan Fitur Saliency-SURF Untuk Klasifikasi Citra Sel Darah Putih Dengan Metode SVM. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(4). <http://jurnal.mdp.ac.id>
- Feldi, M., & Yuswanita, L. (2021). Implementasi Content Based Video Retrieval Menggunakan Speeded-Up Robust Features (Surf). *Jurnal Ilmiah Pendidikan*, 1(1), 57–75.
- Firdaus, A., Walid, M., & Anwari. (2022). Klasifikasi Kasus Covid-19 Menggunakan Model Naive Bayes Classifier. *Jurnal Mahasiswa Teknik Informatika (JATI)*, 6(2), 583–588.

- Hajriansyah. (2023). Identifikasi Jenis Rempah Menggunakan Metode CNN Berbasis Android. *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK)*, 8(1), 223–232.
- Hardoyono, F., Windhani, K., Sambodo, H., & Pudjianto, H. (2019). Analysis of volatile compounds of spices grown in Banyumas District, Jawa Tengah, Indonesia using solid phase microextraction-gas chromatography mass spectrometry. *Malaysian Journal of Analytical Sciences*, 23(6). <https://doi.org/10.17576/mjas-2019-2306-05>
- Hayami, R., Soni, & Gunawan, I. (2022). Klasifikasi Jamur Menggunakan Algoritma Naïve Bayes. *Jurnal CoSciTech (Computer Science and Information Technology)*, 3(1), 28–33. <https://doi.org/10.37859/coscitech.v3i1.3685>
- Kaharuddin, Kusri, & Luthfi, E. T. (2019). Klasifikasi Jenis Rempah-rempah Berdasarkan Fitur Warna Rgb dan Tektur Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Informasi Interaktif*, 4(1), 17–22. <http://e-journal.janabadra.ac.id/>
- Masum, M., Faruk, M. J. H., Shahriar, H., Qian, K., Lo, D., & Adnan, M. I. (2022). Ransomware Classification and Detection with Machine Learning Algorithms. *2022 IEEE 12th Annual Computing and Communication Workshop and Conference, CCWC 2022*, 316–322. <https://doi.org/10.1109/CCWC54503.2022.9720869>
- Mayasari, M., Mulyana, D. I., & Yel, M. B. (2022). Komparasi Klasifikasi Jenis Tanaman Rimpang Menggunakan Principal Component Analisis, Support Vector Machine, K-Nearest Neighbor dan Decision Tree. *Jurnal Teknik Informatika Kaputama (JTİK)*, 6(2).
- Muhathir. (2023). Compares the effectiveness of the bagging method in classifying spices using the histogram of oriented gradient feature extraction technique. *Jurnal Teknik Informatika C.IT Medicom*, 15(1), 48–57.
- Muhathir, Pangestu, R. T., Safira, I., & Melisah. (2023). Performance Comparison of Boosting Algorithms in Spices Classification Using Histogram of Oriented Gradient Feature Extraction. *Journal of Computer Science, Information Technology and Telecommunication Engineering (JCoSITTE)*, 4(1), 342–349. <https://doi.org/10.30596/jcositte.v4i1.13710>
- Muhathir, Rizal, R. A., Sihotang, J. S., & Gultom, R. (2019). Comparison of SURF and HOG extraction in classifying the blood image of malaria parasites using SVM. *2019 International Conference of Computer Science and Information Technology, ICoSNIKOM*, 1–6. <https://doi.org/10.1109/ICoSNIKOM48755.2019.9111647>
- Muhathir, Santoso, M. H., & Muliono, R. (2020). Analysis Naïve Bayes In Classifying Fruit by Utilizing Hog Feature Extraction. *Journal of Informatics*

and *Telecommunication Engineering (JITE)*, 4(1), 151–160.  
<https://doi.org/10.31289/jite.v4i1.3860>

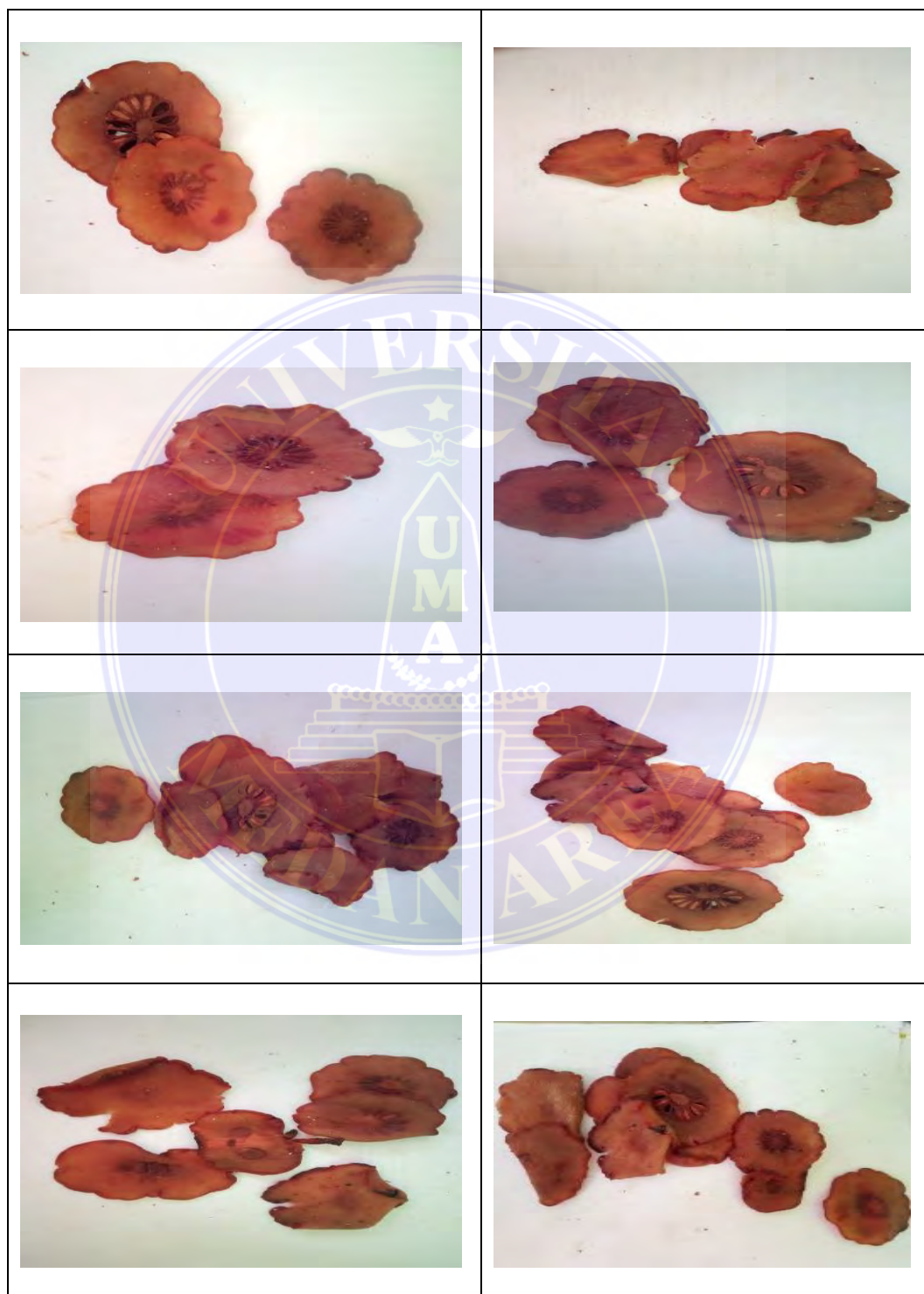
- Mulyana, D. I., & Wibowo, D. R. (2023). Implementasi Tingkat Kematangan Buah Monk Dengan Menggunakan Ekstraksi Gray-Level Co-Occurrence Matrix (GLCM) dan Support Vector Machine (SVM). *JINTEKS (Jurnal Informatika Teknologi dan Sains)*, 5(3), 334–339.
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Pratama, Y., Virgono, A., & Dirgantara, F. M. (2021). Implementasi Advanced Driver Assistance System Menggunakan Metode Surf dalam Mengenali Rambu-Rambu Lalu Lintas di Indonesia. *Jurnal MULTINETICS*, 7(2), 186–195.
- Pratiwi, B. P., Handayani, A. S., & Sarjana. (2020). Pengukuran Kinerja Sistem Kualitas Udara dengan Teknologi WSN menggunakan Confusion Matrix. *Jurnal Informatika UPGRIS*, 6(2), 66–75.
- Proboningrum, S., & Sidauruk, A. (2021). Sistem Pendukung Keputusan Pemilihan Supplier Kain dengan Metode Moora. *Jurnal Sistem Informasi*, 8(1), 43–48.
- Putry Naisah, M., & Sari Betha, N. (2022). Komparasi Algoritma KNN dan Naive Bayes Untuk Klasifikasi Diagnosis Penyakit Diabetes Melitus. *Jurnal Sains dan Manajemen*, 10(1), 45–57.
- Robi, Y., Kartikawati, S. M., & Muflihati. (2019). Etnobotani Rempah Tradisional Di Desa Empoto Kabupaten Sanggau Kalimantan Barat. *Jurnal Hutan Lestari*, 7(1), 130–142.
- Sabila, L., Muzammil, A. R., & Syahrani, A. (2021). Leksikon Rempah-rempah dalam Masakan Melayu Sambas. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa (JPPK)*, 10(11), 1–8.
- Susiarti, S., Rahayu, M., Ningsih, D. Q. W., Arifa, N., & Setiawan, M. (2021). Tanaman Rempah dan Masakan Tradisional di Kelurahan Nanggawer Mekar, Cibinong, Kabupaten Bogor. *Jurnal Masyarakat dan Budaya*, 23(3).  
<https://doi.org/10.14203/jmb.v23i3.1434>
- Tshabalala, R., Kabelinde, A., KAPTCHOUANG TCHATCHOUANG, C. D. O. N. A. L. D., Ateba, C. N., & Manganyi, M. C. (2021). Effect of Clove (*Syzygium aromaticum*) spice as microbial inhibitor of resistant bacteria and Organoleptic Quality of meat. *Saudi Journal of Biological Sciences*, 28(7).  
<https://doi.org/10.1016/j.sjbs.2021.03.052>
- Wandira, S., Noverita, D., & Ilmu Budaya Universitas Andalas, F. (2019). Leksikon Rempah-Rempah di Minangkabau. Dalam *Jurnal Elektronik WACANA ETNIK* (Nomor 2).

- Wang, R., Shi, Y., & Cao, W. (2019). GA-SURF: A new Speeded-Up robust feature extraction algorithm for multispectral images based on geometric algebra. *Pattern Recognition Letters*, 127. <https://doi.org/10.1016/j.patrec.2018.11.001>
- Wibawa, A. P., Kurniawan, A. C., Murti, D. M. P., Adiperkasa, R. P., Putra, S. M., Kurniawan, S. A., & Nugraha, Y. R. (2019a). Naïve Bayes Classifier for Journal Quartile Classification. *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, 7(2), 91. <https://doi.org/10.3991/ijes.v7i2.10659>
- Wibawa, A. P., Kurniawan, A. C., Murti, D. M. P., Adiperkasa, R. P., Putra, S. M., Kurniawan, S. A., & Nugraha, Y. R. (2019b). Naïve Bayes Classifier for Journal Quartile Classification. *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, 7(2), 91. <https://doi.org/10.3991/ijes.v7i2.10659>
- Widiandana, P., Riadi, I., & Sunardi. (2020). Implementasi Metode Jaccard pada Analisis Investigasi Cyberbullying WhatsApp Messenger Menggunakan Kerangka Kerja National Institute of Standards and Technology. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(6), 1046–1051.
- Wulandari, I., Yasin, H., & Widiharih, T. (2020). Klasifikasi Citra Digital Bumbu Dan Rempah Dengan Algoritma Convolutional Neural Network (Cnn). *Jurnal Gaussian*, 9(3), 273–282. <https://doi.org/10.14710/j.gauss.v9i3.27416>
- Yana, T., Malik, A., & Kurniawan, F. (2018). Study Jenis Rempah-Rempah dan Pemanfaatannya di Pasar Tradisional Angso Duo. (*Doctoral dissertation, UIN SULTHAN THAHA SAIFUDDIN JAMBI*), 1–125.
- Yohannes, Y., Devella, S., & Pandrean, A. H. (2020). Penerapan Speeded-Up Robust Feature pada Random Forest Untuk Klasifikasi Motif Songket Palembang. *Jurnal Teknik Informatika dan Sistem Informasi*, 5(3). <https://doi.org/10.28932/jutisi.v5i3.1978>
- Yusuf, A. (2021). Deteksi Objek Dengan Metode Color Filtering Hsv Dan Blob Detection Pada Robot Vertical Take Off And Landing (Studi Kasus Kontes Robot Terbang Indonesia 2018). *Journal of Engineering, Computer Science and Information Technology (JECSIT)*, 1(1), 1–24.

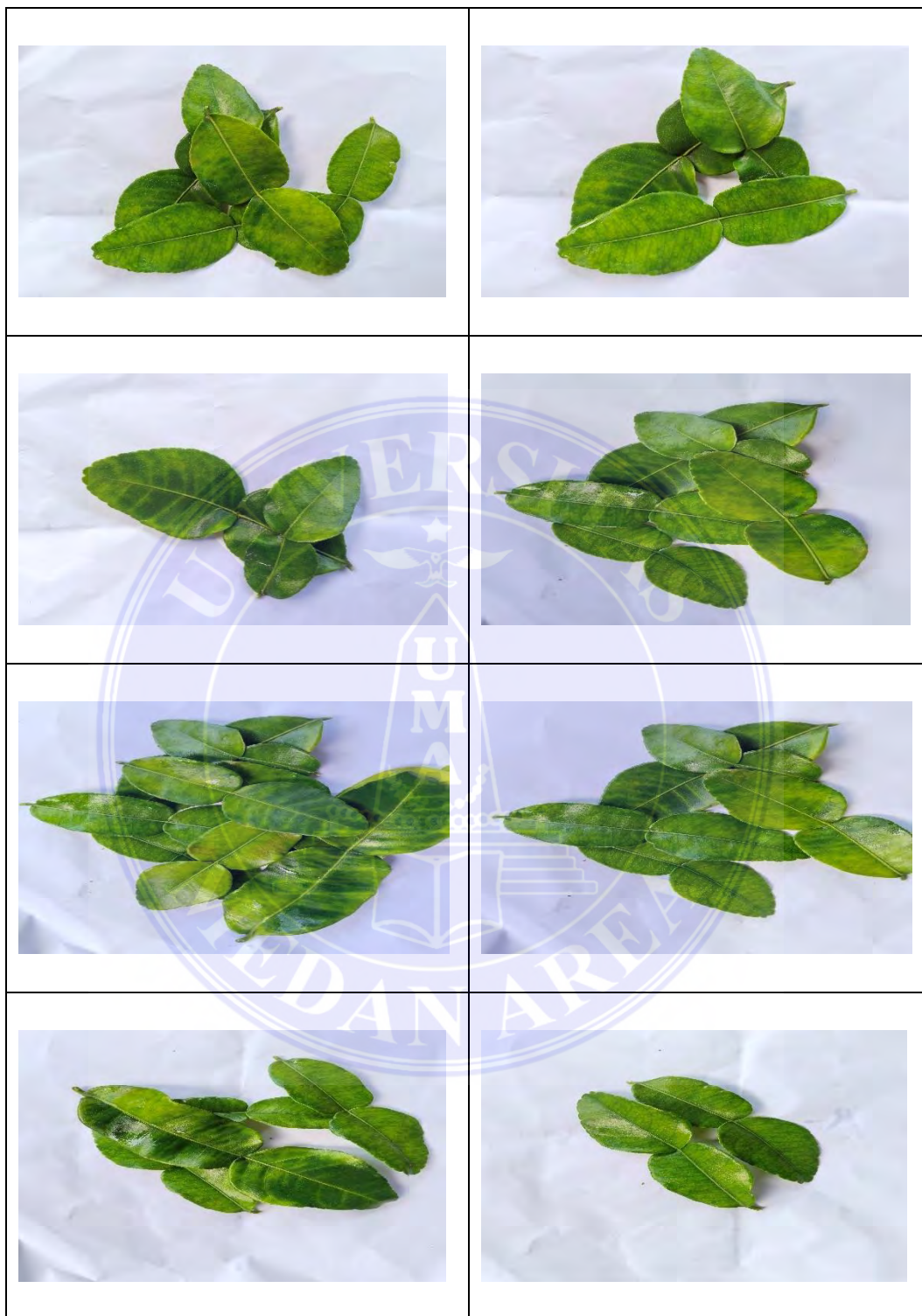
## LAMPIRAN

### Lampiran 1 : Dataset

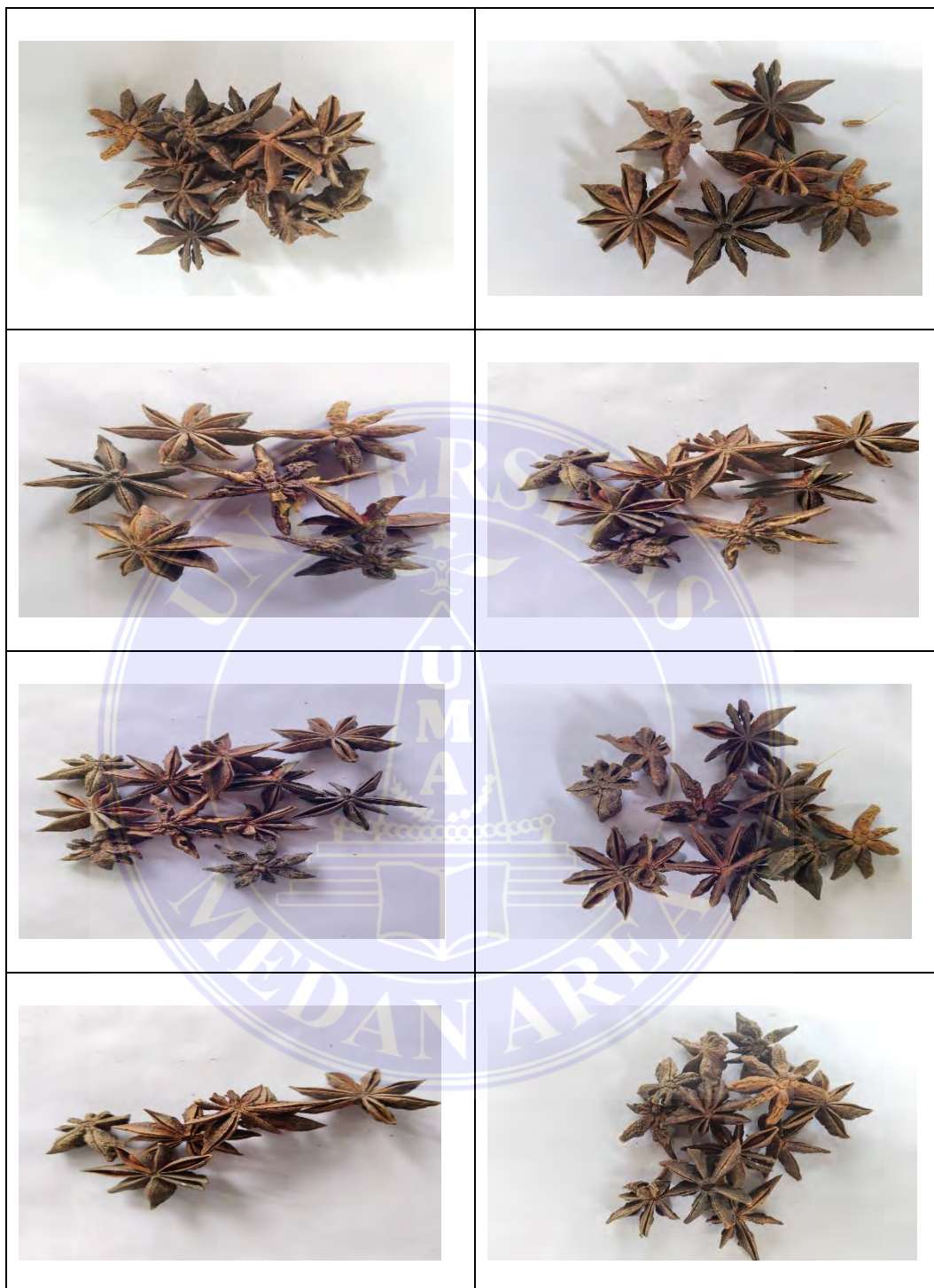
#### A. Asam Glugur



## B. Daun Jeruk



### C. Bunga Lawang

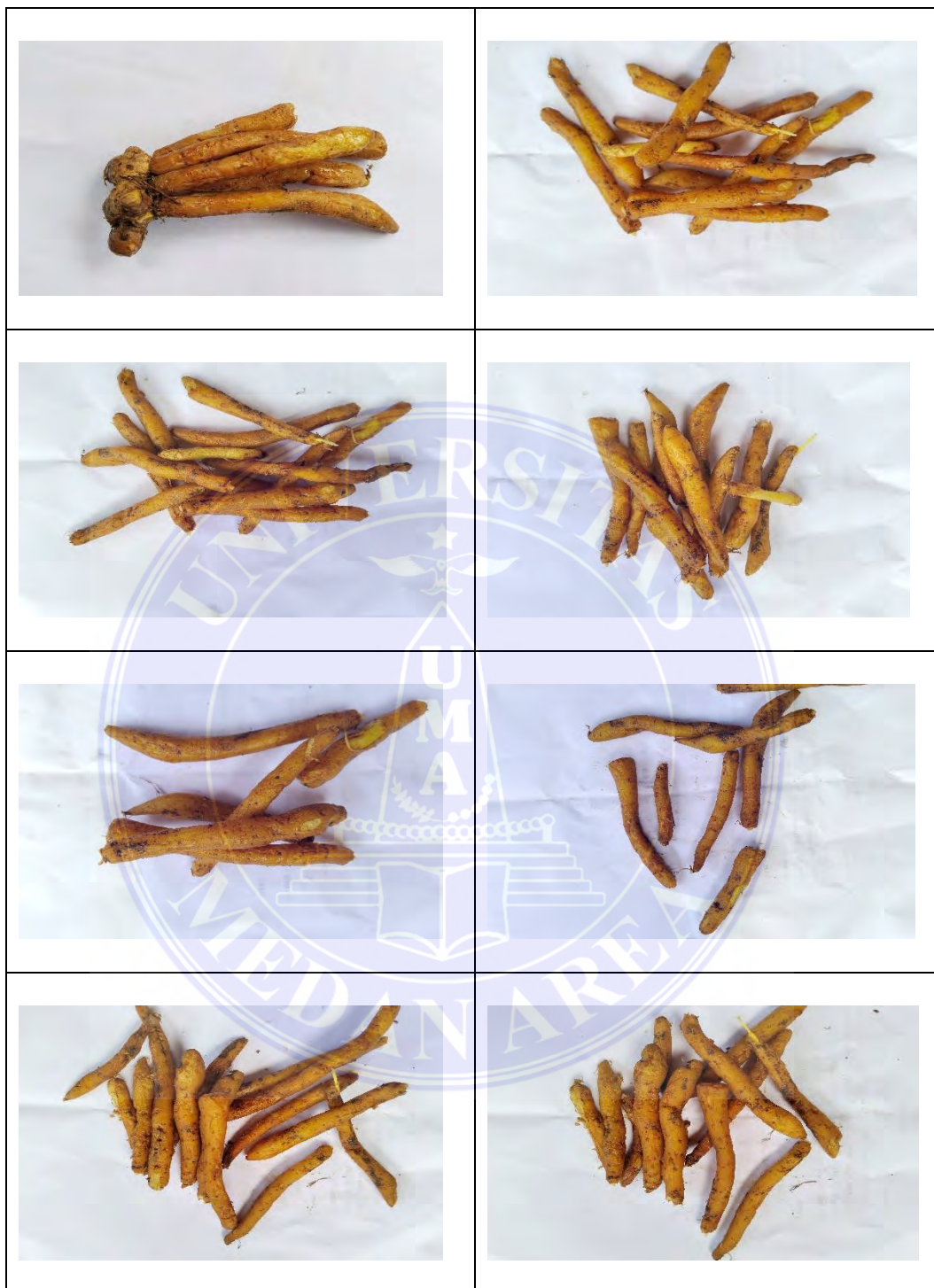




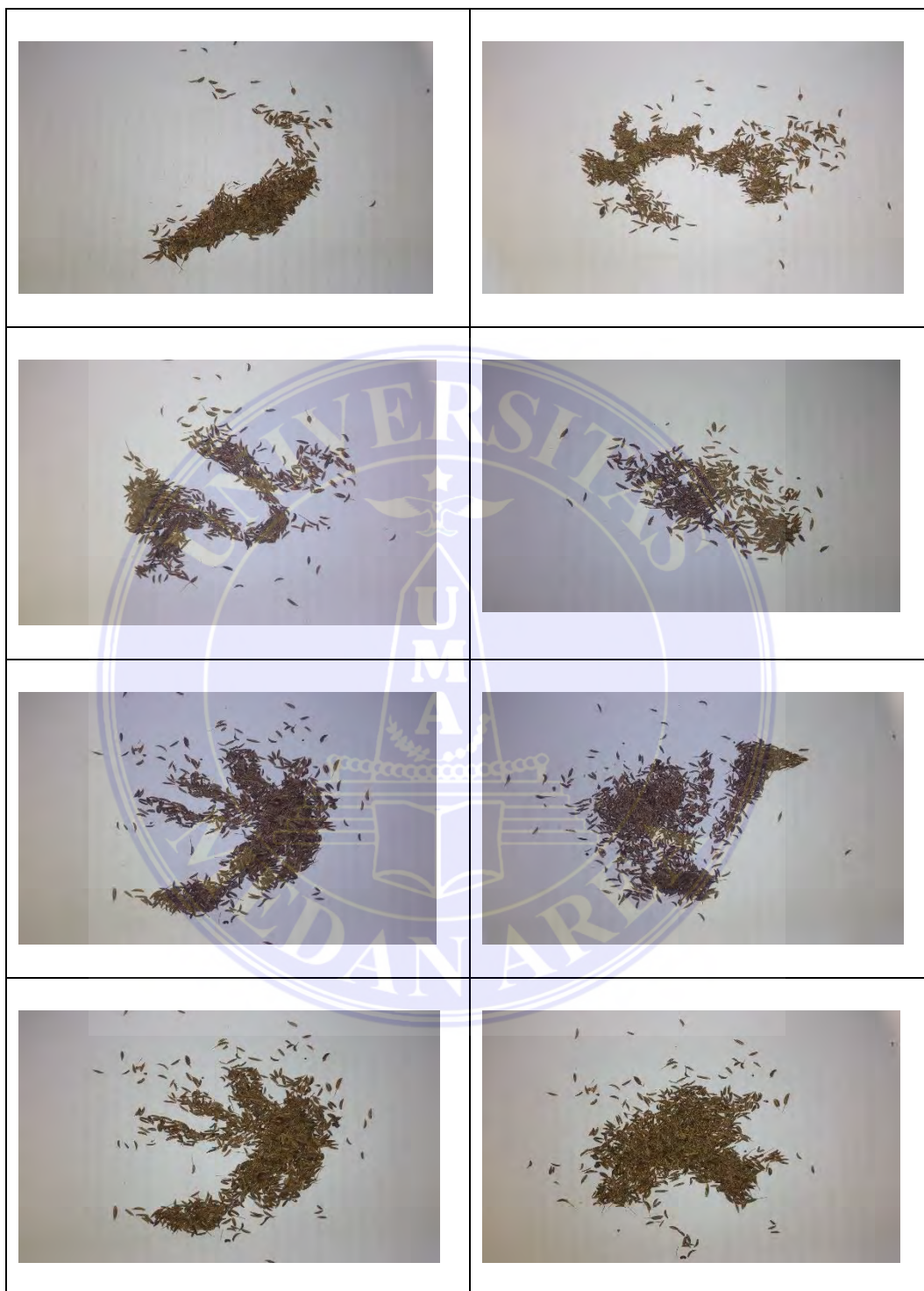
### D. Kayu Manis



### E. Temu Kunci



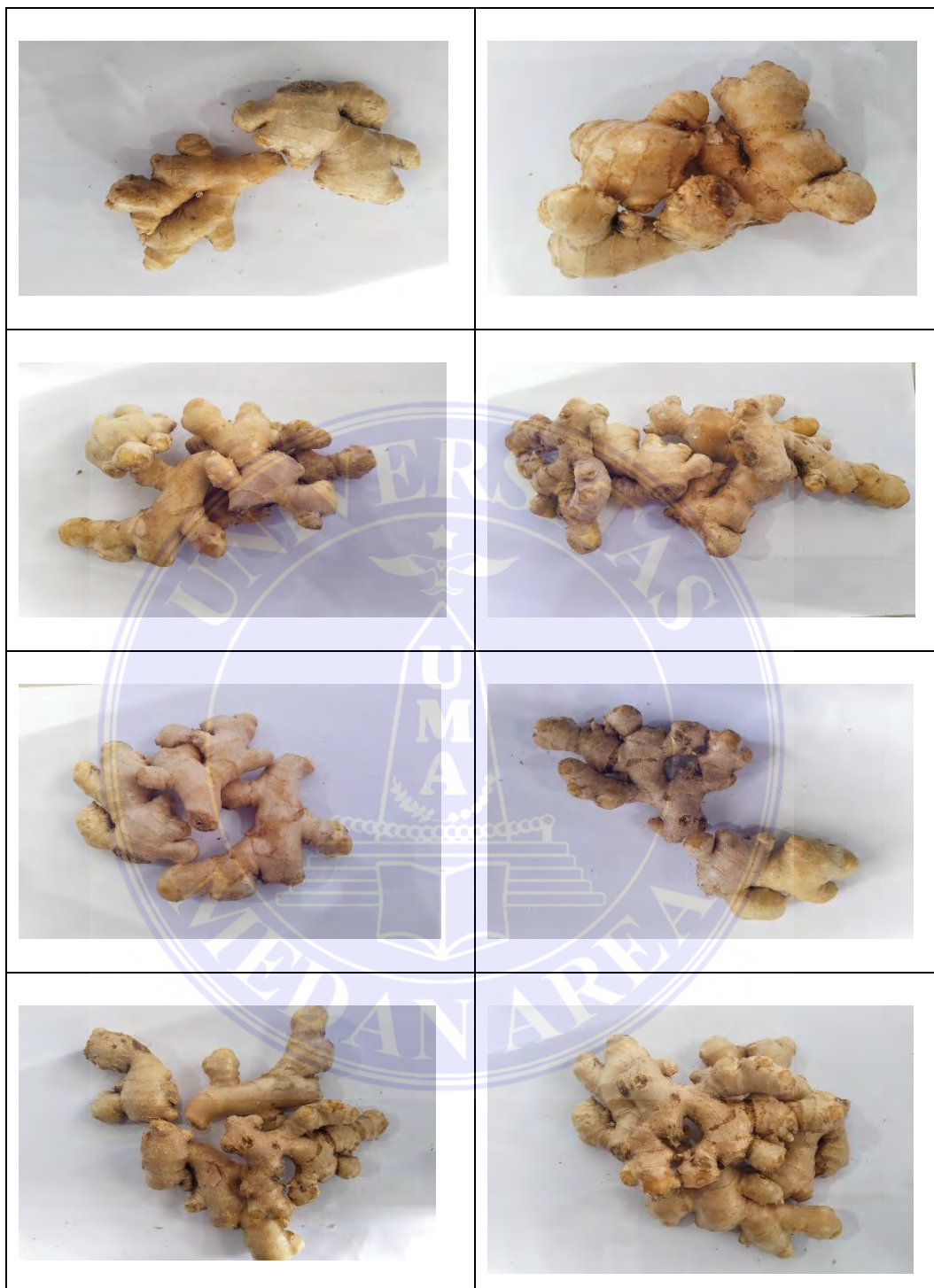
## F. Jinten



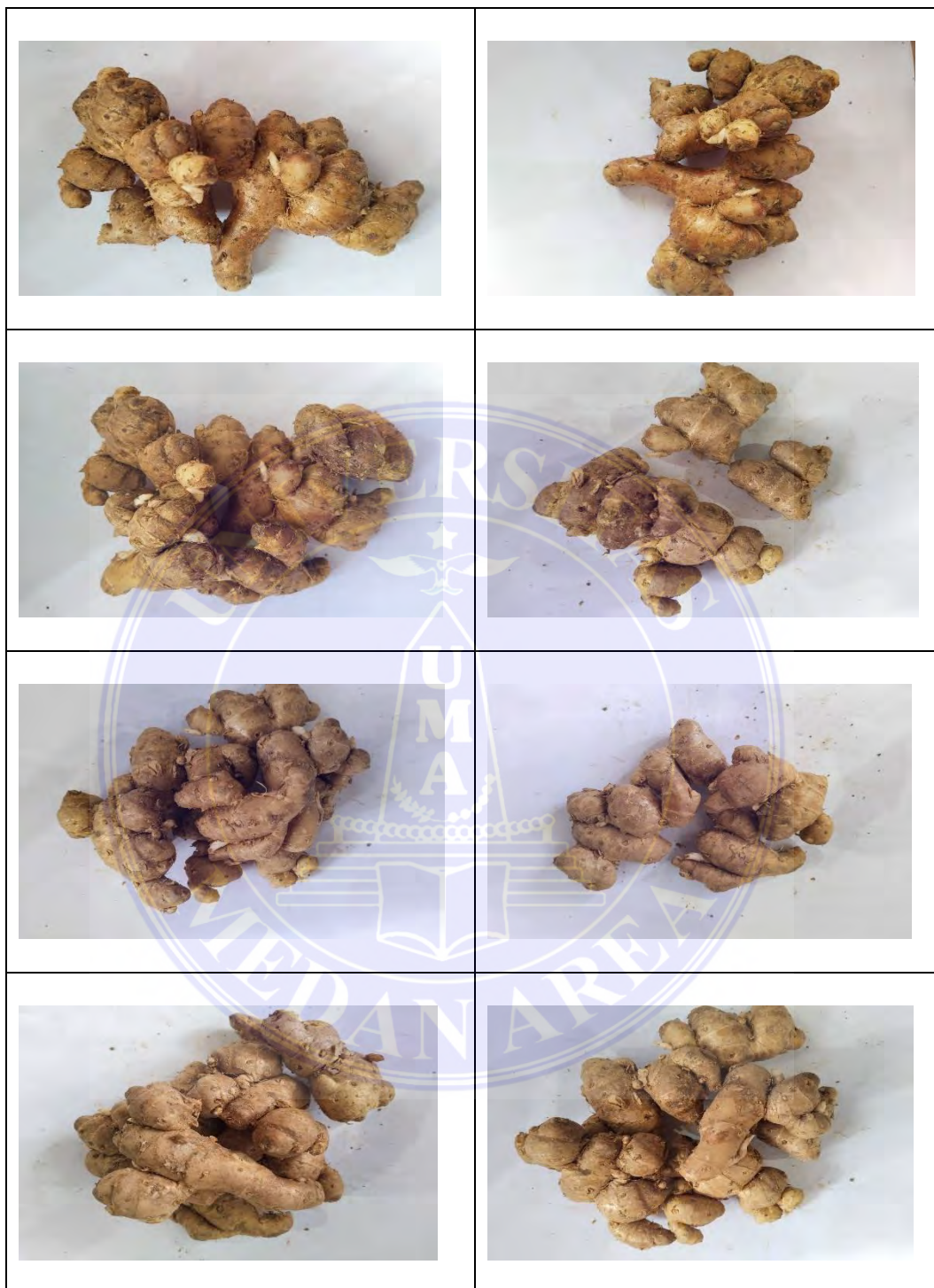
### G. Temulawak



## H. Jahe



## I. Kencur



## J. Kuyit



## K. Lengkuas

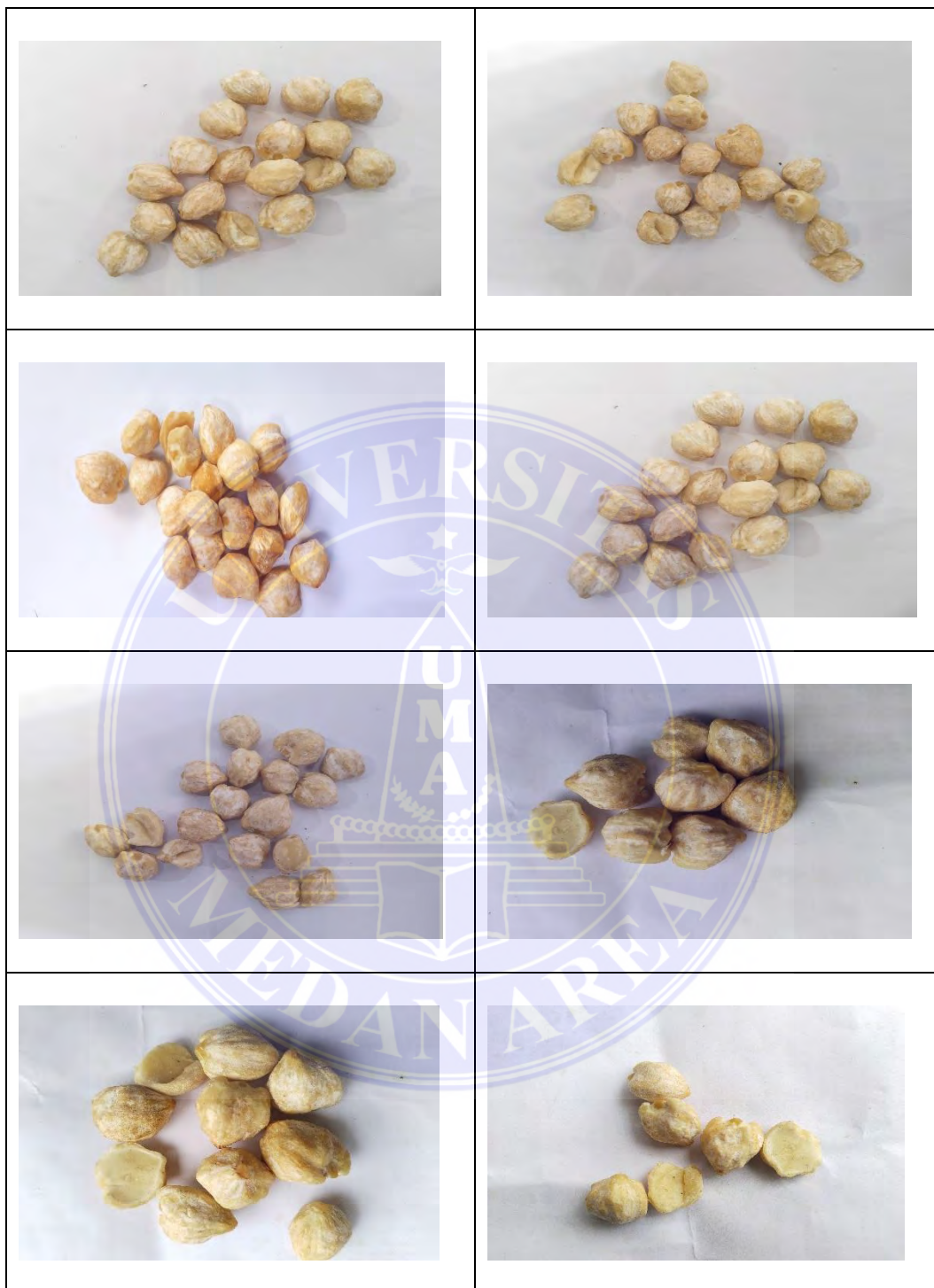




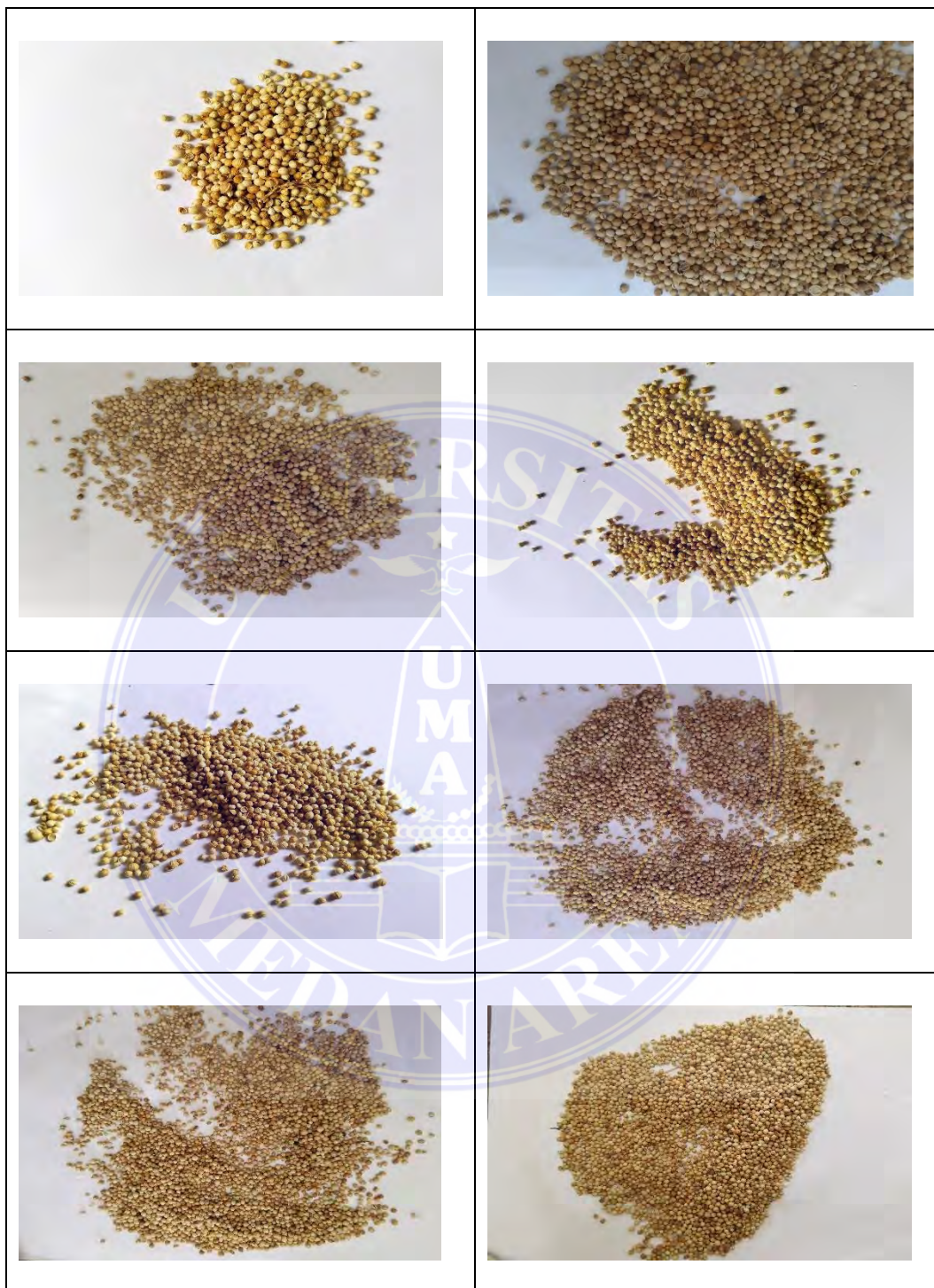
## L. Pala



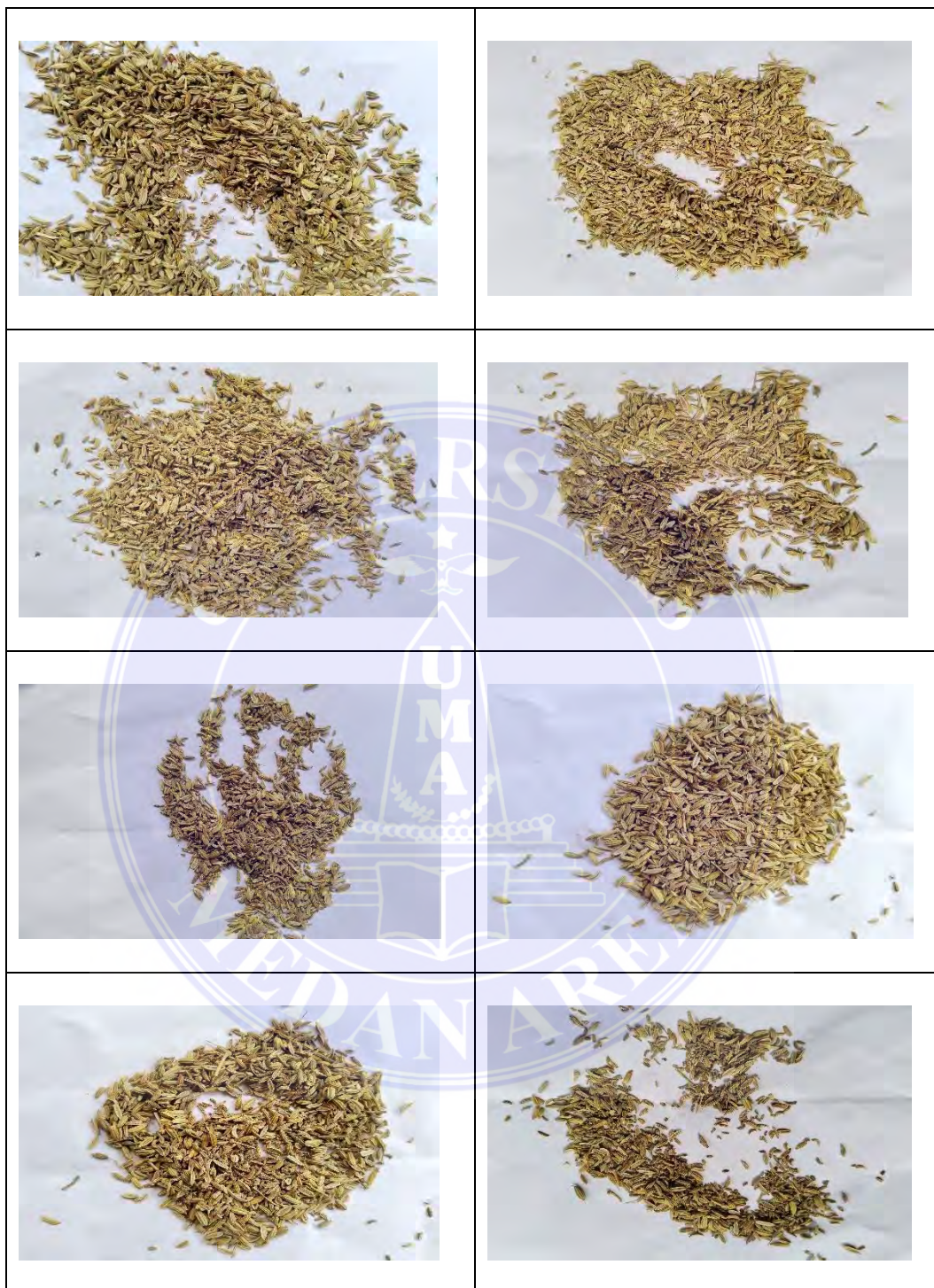
### M. Kemiri



## N. Ketumbar



### O. Adas Manis



## Lampiran 2 : Source Code

**A. Import Library**

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import io

from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix,
classification_report
```

**B. Resize Image**

```
image_folder = 'kunyit'
resized_folder = 'resize_kunyit'
files = os.listdir(image_folder)
kunyit = [file for file in files if file.endswith((''.jpg',
'.png'))]

if not os.path.exists(resized_folder):
    os.makedirs(resized_folder)

for image in kunyit:
    img = cv2.imread(os.path.join(image_folder, image))
    resized = cv2.resize(img, (379, 504))
    cv2.imwrite(os.path.join(resized_folder, image),
resized)
```

**C. Ekstraksi Fitur SURF**

```
descriptor_folder = 'hasil_ekstraksi_kunyit_3000'
if not os.path.exists(descriptor_folder):
    os.makedirs(descriptor_folder)

for image in resize_kunyit:
    img = cv2.imread(os.path.join(image_folder, image))
    gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

    surf =
cv2.xfeatures2d.SURF_create(hessianThreshold=3000)
    kp, des = surf.detectAndCompute(img, None)
```

```
img = cv2.drawKeypoints(img, kp, None, (0,255,0),4)
cv2.imwrite(os.path.join(descriptor_folder, image +
'_keypoints.jpg'), img)
```

#### D. Data Gambar menjadi Array

```
image_folder = 'hasil_ekstraksi_kunyit_3000'
files = os.listdir(image_folder)
hasil_ekstraksi_kunyit_3000 = [file for file in files if
file.endswith(('.jpg', '.png'))]

descriptors = []

for image in hasil_ekstraksi_kunyit_3000:
    img = cv2.imread(os.path.join(image_folder, image))
    gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

    surf = cv2.xfeatures2d.SURF_create()
    _, des = surf.detectAndCompute(gray, None)

    descriptors.append(des)

df = pd.DataFrame(np.concatenate(descriptors))
df = df.iloc[:150, :20]
df.to_csv('kunyit 3000.csv', index=False)
```

#### E. Pemanggilan Data

```
df = pd.read_csv('surf_5_rempah.csv')
#print(df)
print(df)

df.shape
X = df.iloc[:, :20].values
y = df['rempah'].values
```

#### F. Pembagian Data Training dan Testing

```
X_train, X_test, y_train, y_test = train_test_split(X, y ,
test_size=0.2, random_state=0)
```

#### G. Klasifikasi Naïve Bayes

```
#Naive Bayes
#from sklearn.naive_bayes import BernoulliNB
```

```

from sklearn.naive_bayes import GaussianNB
#from sklearn.naive_bayes import MultinomialNB

NB = GaussianNB()
#NB = BernoulliNB()
#NB = MultinomialNB()

X_train = np.clip(X_train, -0, None)

y_pred = NB.fit(X_train, y_train).predict(X_test)
confusion_matrix(y_test, y_pred)

pd.crosstab(y_test, y_pred, rownames=['True'],
            colnames=['Predicted'], margins=True)
print(y_pred)
print(pd.crosstab)

```

## H. Confusion Matrix

```

import seaborn as sns
import matplotlib.pyplot as plt
f, ax = plt.subplots(figsize=(8,5))

sns.heatmap(confusion_matrix(y_test, y_pred), annot=True,
            cmap="YlGnBu", ax=ax, annot_kws={"size": 18})
ax.xaxis.set_ticklabels(['bl', 'tl', 'ky', 'km', 'ag']);
ax.yaxis.set_ticklabels(['bl', 'tl', 'ky', 'km', 'ag']);
plt.yticks(fontsize=10)
plt.xticks(fontsize=10)
plt.xlabel("Predict label", fontsize=20)
plt.ylabel("True Label", fontsize=20)
plt.show()

```

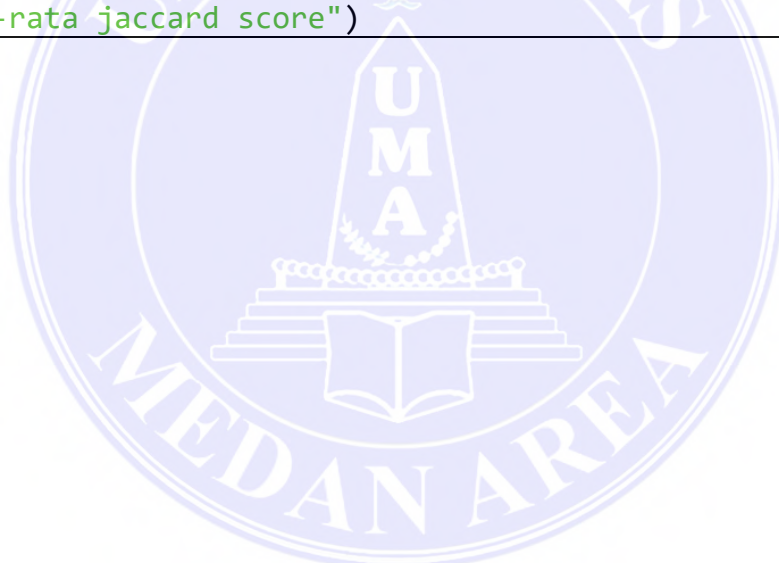
## I. Model Evaluasi

```

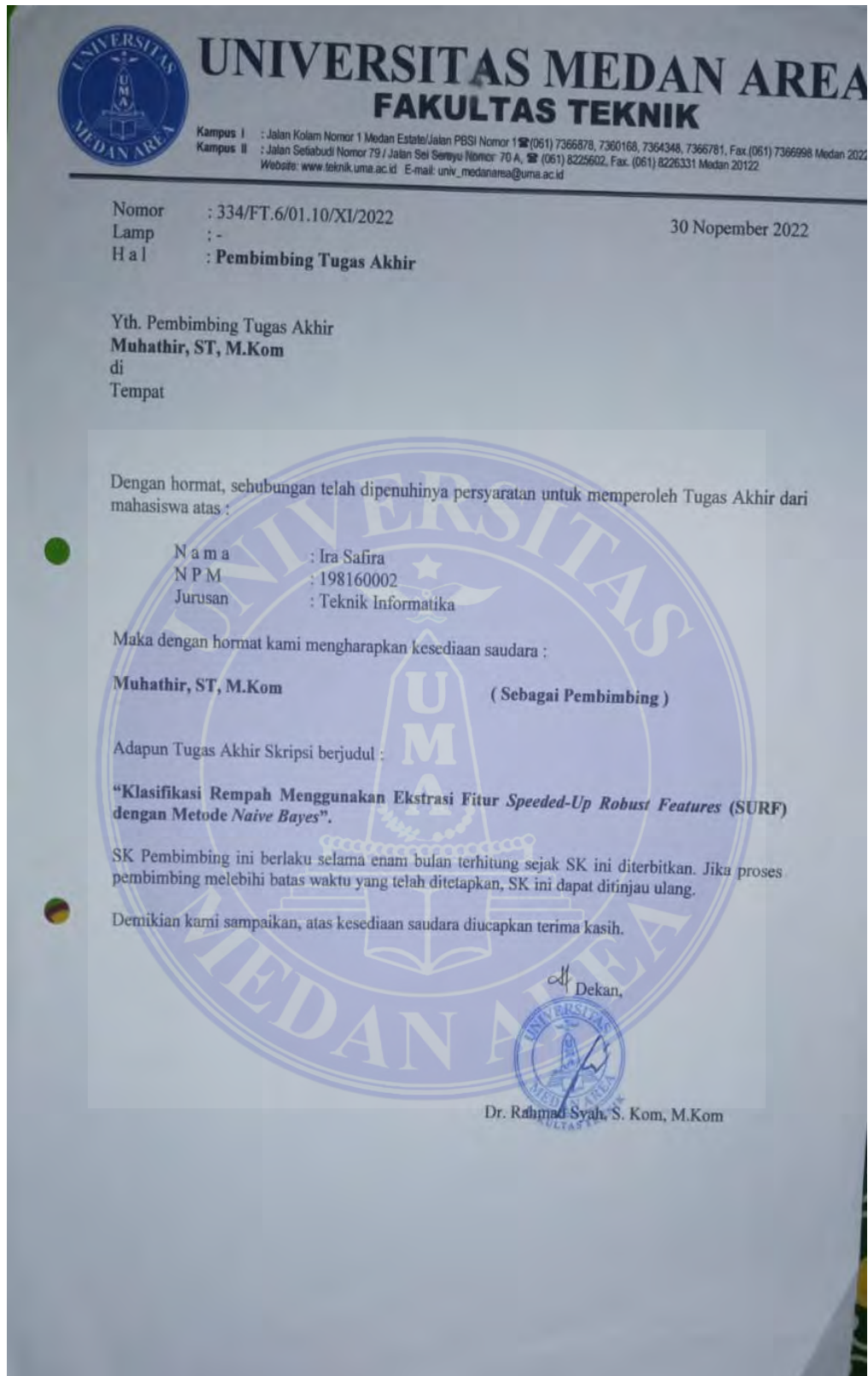
print(confusion_matrix(y_test, y_pred))
from sklearn.metrics import accuracy_score
print(accuracy_score(y_test, y_pred), ": is the accuracy score")
from sklearn.metrics import precision_score
print(precision_score(y_test, y_pred, average=None), ": is the precision score")
from sklearn.metrics import recall_score
print(recall_score(y_test, y_pred, average=None), ": is the recall score")
from sklearn.metrics import f1_score

```

```
print(f1_score(y_test, y_pred, average=None), ": is the f1 score")
from sklearn.metrics import fbeta_score
print(fbeta_score(y_test, y_pred, beta=2.0, average=None), ": is the fbeta score")
from sklearn.metrics import jaccard_score
print(jaccard_score(y_test, y_pred, average=None), ": is the jaccard score")
print(accuracy_score(y_test, y_pred), ": Rata-rata Accuracy")
print(precision_score(y_test, y_pred, average='macro'), ": Rata-rata precision score")
print(recall_score(y_test, y_pred, average='macro'), ": Rata-rata recall score")
print(f1_score(y_test, y_pred, average='macro'), ": Rata-rata f1 score")
print(fbeta_score(y_test, y_pred, beta=2.0, average='macro'), ": Rata-rata fbeta score")
print(jaccard_score(y_test, y_pred, average='macro'), ": Rata-rata jaccard score")
```







**turnitin** Similarity Report ID: oid:29477-35927701

PAPER NAME	AUTHOR
SKRIPSI IRA BAB 1-BAB 5 (1).pdf	IRA SAFIRA
WORD COUNT	CHARACTER COUNT
5680 Words	34050 Characters
PAGE COUNT	FILE SIZE
40 Pages	615.6KB
SUBMISSION DATE	REPORT DATE
May 22, 2023 7:48 PM GMT+7	May 22, 2023 7:49 PM GMT+7

**21% Overall Similarity**  
The combined total of all matches, including overlapping sources, for each database.

- 16% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 14% Submitted Works database

**Excluded from Similarity Report**

- Small Matches (Less than 10 words)

UNIVERSITAS MEDAN AREA

Summary