

**KLASIFIKASI REMPAH MENGGUNAKAN EKSTRAKSI FITUR
HISTOGRAM OF ORIENTED GRADIENT (HOG) DENGAN
METODE *K-NEAREST NEIGHBOR* (K-NN)**

SKRIPSI

OLEH:

MELISAH

198160022



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
*HISTOGRAM OF ORIENTED GRADIENT (HOG) DENGAN
METODE K-NEAREST NEIGHBOR (K-NN)***

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana (S1) di Fakultas Teknik
Universitas Medan Area

OLEH:

MELISAH

198160022

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

LEMBAR PENGESAHAN

Judul Skripsi : Klasifikasi Rempah Menggunakan Ekstraksi Fitur *Histogram of Oriented Gradient (HOG)* dengan Metode *K-Nearest Neighbor (K-NN)*

Nama : Melisah
NPM : 198160022
Fakultas : Teknik
Prodi : Teknik Informatika

Disetujui Oleh



Muhathir, ST, M. Kom
Pembimbing



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



Rizki Melisah, S. Kom, M. Kom
NIDN : 01050588902

Tanggal Lulus : 04 Agustus 2023

iii

PALAMAN PERNYATAAN

Saya menyatakan bahwa skripsi yang saya susun, sebagai syarat memperoleh gelar sarjana 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 sanksisanksi lainnya dengan peraturan yang berlaku, apabila dikemudian hari ditemukan adanya plagiat dalam skripsi ini



iv

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

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


Nama : Melisah
NPM : 198160022
Program Studi : Teknik Informatika
Fakultas : Teknik
Jenis Karya : Skripsi

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

Klasifikasi Rempah Menggunakan Ekstraksi Fitur Histogram of Oriented Gradient (HOG) dengan Metode K-Nearest Neighbor (K- NN)

berserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti *Noneksklusif* ini Universitas Medan Area berhak menyimpan, mengalihmedia/formatkan, 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.

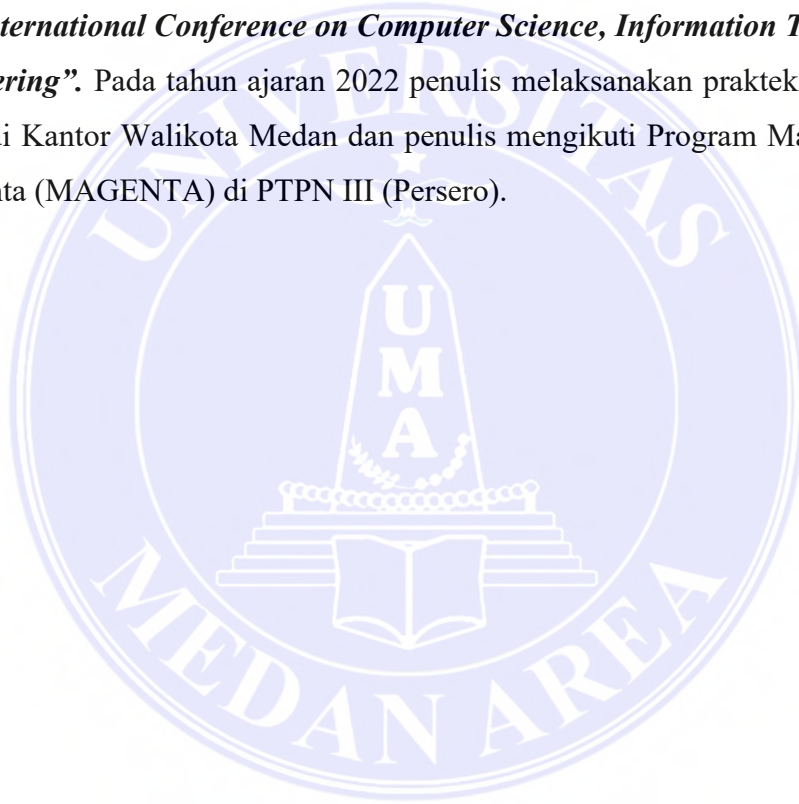
Dibuat di : Medan
Pada tanggal : 09 November 2023
Yang menyatakan


(Melisah)

RIWAYAT HIDUP

Penulis dilahirkan di Aek Nabara, Labuhanbatu Pada tanggal 05 Juni 2001 dari ayah Syahril dan ibu Supriatin Penulis merupakan putri kedua dari 2 bersaudara.

Tahun 2019 Penulis lulus dari SMK Citra Bangsa Al-Ittihad Aek Nabara dan pada tahun 2019 terdaftar sebagai mahasiswa Fakultas Teknik Jurusan Teknik Informatika Universitas Medan Area. Selama mengikuti perkuliahan penulis mengikuti ***“The International Conference on Computer Science, Information Technology and Engineering”***. Pada tahun ajaran 2022 penulis melaksanakan praktek kerja lapangan (PKL) di Kantor Walikota Medan dan penulis mengikuti Program Magang Generasi Bertalenta (MAGENTA) di PTPN III (Persero).



ABSTRAK

Rempah-rempah merupakan sumber daya hayati yang telah lama berperan sangat penting dalam kehidupan sehari-hari. Rempah rempah memiliki karakteristik, bentuk, dan warna yang hampir mirip dan sulit untuk membedakan antara rempah satu dengan rempah yang lainnya. Untuk membantu dalam mengenali karakteristik dari rempah-rempah yang ada maka penulis mencoba melakukan penelitian dengan judul. "Klasifikasi Rempah Menggunakan Metode *K-Nearest Neighbor* (K-NN) Dan Menggunakan Ekstraksi Fitur *Histogram oriented of Gradient*. Metode yang digunakan dalam penelitian ini adalah *K-Nearest Neighbor* (K-NN) dan menggunakan ekstraksi fitur *Histogram of Oriented Gradient* (HOG). Pada penelitian ini dataset yang digunakan sebanyak 2250 sampel citra dan dibagi menjadi dua kategori yaitu data training dan data testing dengan perbandingan 80% dan 20%. Hasil pengujian pada penelitian ini menghasilkan rata-rata akurasi sebesar 92%, presisi 92%, recall 92%, f1 score 92%, fbeta score 92%, dan jaccard score 85%. Berdasarkan pengujian yang telah dilakukan, penggunaan jarak *Manhattan* memberikan performa terbaik dengan akurasi sebesar 94%, sementara jarak jaccard memperoleh akurasi paling rendah dengan nilai sebesar 62%. Hasil ini menunjukkan bahwa ekstraksi fitur sangat berpengaruh terhadap jumlah jenis dalam mengekstraksi informasi. *Histogram of Oriented Gradient* (HOG) bekerja secara optimal jika jumlah jenis yang diekstraksi kecil dan tidak optimal jika digunakan dalam sejumlah besar jenis klasifikasi.

Keyword: Rempah, *K-Nearest Neighbor* (K-NN), *Histogram of Oriented Gradient* (HOG)

ABSTRACT

Spices are biological resources that have long played a very important role in everyday life. Spices have characteristics, shapes, and colors that are almost similar and it is difficult to distinguish one spice from another. To assist in recognizing the characteristics of existing spices, the author tries to do research with the title. "Spices Classification Using the K-Nearest Neighbor (K-NN) Method and Using Histogram Oriented Gradient Feature Extraction. The method used in this study is the K-Nearest Neighbor (K-NN) and uses the Histogram of Oriented Gradient (HOG) feature extraction. In this study, the dataset used was 2250 image samples and divided into two categories, namely training data and testing data with a ratio of 80%: 20%. The test results yield an average accuracy of 92%, 92% precision, 92% recall, 92% f1 score, 92% fbeta score, and 85% jaccard score. Based on testing, the use of the manhattan distance gives the best performance with an accuracy of 94%, while the jaccard distance has the lowest accuracy with a value of 58%. These results indicate that feature extraction greatly influences the number of types in extracting information, the Histogram of Oriented Gradient works optimally when the number of types extracted is small and not optimal when used in a large number of classification types.

Keyword : Spices, K-Nearest Neighbor (K-NN), Histogram Of Oriented Gradient (HOG).



KATA PENGANTAR

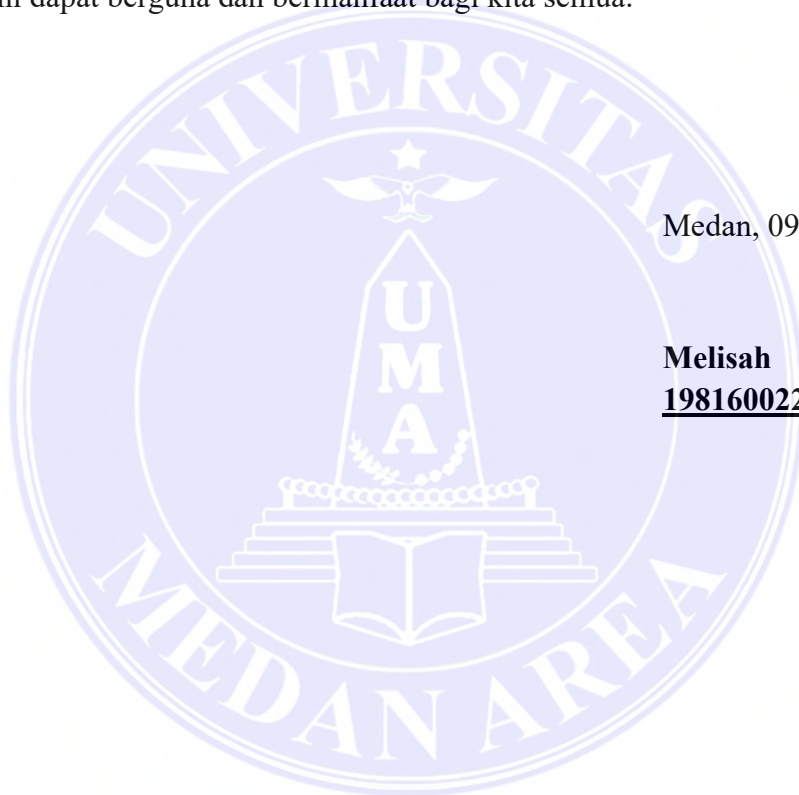
Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Kuasa atas segala karuniaNya, sehingga Tugas Akhir yang berjudul “**KLASIFIKASI REMPAH MENGGUNAKAN EKSTRAKSI FITUR *HISTOGRAM OF ORIENTED GRADIENT* (HOG) DENGAN MENGGUNAKAN METODE K- NEAREST NEIGHBOR (K-NN)**” dapat diselesaikan dengan baik dan tepat waktu. Adapun 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 dari beberapa pihak, oleh karena itu penulis hendak mengucapkan terimah kasih kepada:

1. Tuhan Yang Maha Esa, yang telah memberikan kesehatan dan keselamatan selama melakukan studi.
2. Kepada orang tua penulis ayah Syahril Siregar dan ibu Supriatin yang telah memberi semangat serta dukungannya agar penulis dapat menyelesaikan Pendidikan hingga ke perguruan tinggi.
3. Bapak Prof. Dr. Dadan Ramadhan, M. EnG, MScH selaku Rektor Universitas Medan Area
4. Bapak Rahmad Syah S. Kom, M. Kom selaku Dekan fakultas Teknik Universitas Medan Area
5. Ibu Susilawati S. Kom, M. Kom. Selaku Wakil Dekan Bidang Akademik Universitas Medan Area
6. Bapak Rizki Muliono S. Kom, M. Kom. Selaku Kaprodi Informatika Universitas Medan Area
7. Bapak Muhathir S. T, M. Kom Selaku Dosen Pembimbing yang telah memberikan arahan serta dukungan dalam penyusunan skripsi ini.
8. Seluruh Jajaran Dosen dan Staf pada Fakultas Teknik Universitas Medan Area
9. Ira safira yang telah berjuang bersama-sama dari Maba hingga selesai yang telah melewati suka maupun duka dalam menyelesaikan pendidikan ini.

10. Teman- teman Informatika Universitas Medan Area, yang selalu membantu memberikan dukungannya.
11. Seluruh pihak yang tidak dapat disebutkan satu persatu yang telah membantu dan memberikan dukungannya sehingga skripsi ini dapat di selesaikan dengan baik.

Penulis menyadari dalam penyusunan skripsi ini, masih terdapat banyak kekurangan dan jauh dari kata sempurna, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemanjuan penelitian selanjutnya. Semoga skripsi ini dapat berguna dan bermanfaat bagi kita semua.



Medan, 09 November 2023

Melisah
198160022

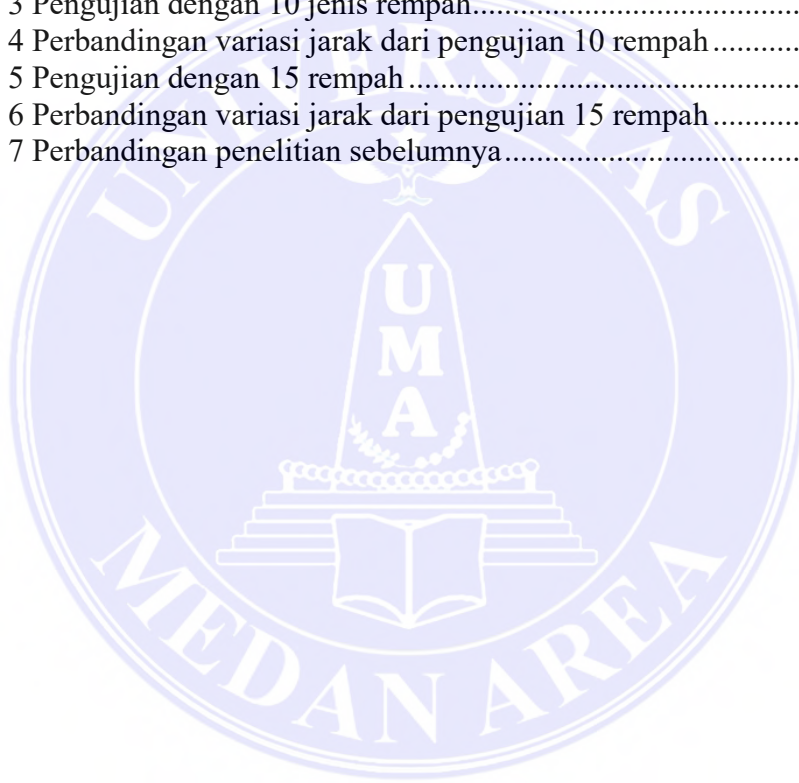
DAFTAR ISI

LEMBAR PENGESAHAN	iii
HALAMAN PERNYATAAN	iv
RIWAYAT HIDUP.....	vi
ABSTRAK.....	vii
<i>ABSTRACT</i>	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.....	5
1.3 Tujuan.....	6
1.4 Manfaat Penelitian.....	6
1.5 Batasan Masalah.....	6
BAB II LANDASAN TEORI.....	7
2.1 <i>K- Nearest Neighbor</i> (K-NN).....	7
2.2 <i>Histogram of Gradients</i> (HOG).....	10
2.3 Rempah - Rempah.....	10
2.4 Penelitian Terdahulu.....	11
BAB III METODOLOGI PENELITIAN.....	16
3.1 Alat dan Bahan Penelitian.....	16
3.1.1 Perangkat Keras.....	16
3.1.2 Perangkat Lunak.....	16
3.2 Metode Pengumpulan Data.....	17
3.2.1 Dataset.....	17
3.2.2 Data Analis.....	18
3.3 Diagram Alur Penelitian.....	19
3.3 Arsitektur Penelitian.....	20

3.4 Metode Evaluasi	21
3.4.1 Confusion Matrix	21
3.4.2 <i>Performa Measure</i>	22
BAB IV HASIL DAN PEMBAHASAN	25
4.1 Ekstraksi Fitur <i>Histogram of Oriented Gradients</i> (HOG)	25
4.1.1 Pengumpulan Data	25
4.1.2 Pengambilan Dataset	25
4.1.3 <i>Preprocessing</i> dan <i>Resizing</i> Citra	26
4.1.4 Preprosesan <i>Histogram of Oriented Gradient</i> (HOG)	26
4.1.5 Hasil CSV	27
4.2 Metode <i>K- Nearest Neighbor</i>	28
4.3 Pengujian	28
4.3.1 Pengujian dengan 5 jenis rempah	28
4.3.1.1 Penentuan Jarak Terbaik	28
4.3.1.2 Grafik Hasil Pengujian 5 Rempah	29
4.3.1.3 Evaluasi Performa Model	29
4.3.2 Pengujian dengan 10 jenis rempah	32
4.3.2.1 Penentuan Jarak Terbaik	32
4.3.2.2 Grafik Hasil Pengujian 10 Rempah	32
4.3.2.3 Evaluasi Performa Model	33
4.3.3 Pengujian dengan 15 rempah	35
4.3.3.1 Penentuan Jarak Terbaik	35
4.3.3.2 Grafik Hasil Pengujian 15 Rempah	36
4.3.3.3 Evaluasi Performa Model	37
4.4 Pembahasan	38
BAB V KESIMPULAN DAN SARAN	42
5.1 Kesimpulan	42
5.2 Saran	42
DAFTAR PUSTAKA	43
LAMPIRAN	50

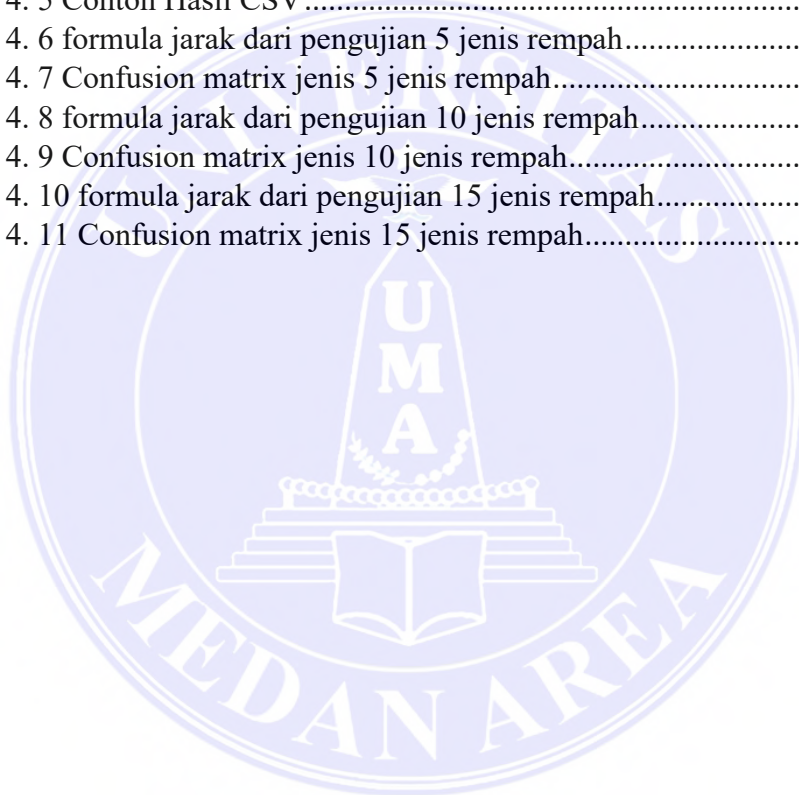
DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	11
Tabel 3. 1 Perangkat Keras yang digunakan.....	16
Tabel 3. 2 Perangkat Lunak yang digunakan.....	16
Tabel 3. 3 Pembagian data	18
Tabel 3. 4 Confusion Matrix	21
Tabel 4. 1 Pengujian 5 jenis rempah	30
Tabel 4. 2 Perbandingan variasi jarak dari pengujian 5 rempah.....	31
Tabel 4. 3 Pengujian dengan 10 jenis rempah.....	34
Tabel 4. 4 Perbandingan variasi jarak dari pengujian 10 rempah.....	34
Tabel 4. 5 Pengujian dengan 15 rempah	37
Tabel 4. 6 Perbandingan variasi jarak dari pengujian 15 rempah.....	38
Tabel 4. 7 Perbandingan penelitian sebelumnya.....	40



DAFTAR GAMBAR

Gambar 3. 1 Sampel Rempah.....	17
Gambar 3. 2 Diagram Alur Penelitian.....	19
Gambar 3. 3 Langkah Penelitian.....	20
Gambar 4. 1 Sample rempah sebelum di ekstraksi	25
Gambar 4. 2 Gambar Citra Rempah Yang Sudah Di Resizing	26
Gambar 4. 3 Gambar Citra Rempah Yang belum Di Resizing	26
Gambar 4. 4 Gambar Sampel Rempah Yang Sudah di Ekstraksi	27
Gambar 4. 5 Contoh Hasil CSV	27
Gambar 4. 6 formula jarak dari pengujian 5 jenis rempah.....	29
Gambar 4. 7 Confusion matrix jenis 5 jenis rempah.....	30
Gambar 4. 8 formula jarak dari pengujian 10 jenis rempah.....	32
Gambar 4. 9 Confusion matrix jenis 10 jenis rempah.....	33
Gambar 4. 10 formula jarak dari pengujian 15 jenis rempah.....	36
Gambar 4. 11 Confusion matrix jenis 15 jenis rempah.....	37



BAB I

PENDAHULUAN

1.1 Latar Belakang

Rempah-rempah adalah bahan aromatik tanaman yang digunakan dalam memasak dan obat herbal. Rempah-rempah dibagi menjadi dua kategori yaitu bumbu basah (dan bumbu kering (Hizaz Z Yahya, 2023). Rempah-rempah berfungsi sebagai penambah rasa dan pengawet makanan. Rempah-rempah memiliki sifat bioaktif dari rempah-rempah yang berbeda, yang menunjukkan bahwa mereka mungkin memiliki peran dalam mencegah penyakit kronis yang tidak menular (Opara, 2019).

Rempah-rempah dapat dijadikan sebagai pengobatan herbal untuk obat dan pengobatan/ teraphy terutama pada penyakit kronis karena rempah rempah memiliki kandungan antioksidan, anti-inflamasi, antitumorigenic, dan glukosa (Jiang, 2019). Rempah-rempah berasal dari bagian kering tanaman seperti kuncup, bunga (cengkeh, kunyit), kulit kayu (kayu manis), akar (jahe, kunyit), buah (cengkeh, cabai, lada hitam), atau biji-bijian (jintan) yang mengandung minyak volatil atau aroma dan rasa aromatic (Rosa Vázquez-Fresno, 2019).

Pada tanaman rempah di Indonesia terdapat berbagai tanaman rempah seperti, jahe, kunyit, kayu manis, secang dan andariman. Jahe mengandung senyawa antioksidan yaitu metil ester, asam oktadekanoat, dan selenium. Pada penelitian serbuk kunyit dan ekstrak asam jawa, ekstrak kunyit sebesar 55% dan ekstrak daun asam jawa sebesar 45%, sehingga menghasilkan nilai 4,844%. Kayu manis dapat menjadi sumber antioksidan karena mengandung banyak senyawa seperti *eugenol*, *safrrole*, *cinnamaldehyde*, *tanin*, dan *kalsium oksalat*. Bahan aktif yang terkandung dalam secang adalah *flavonoid*. Buah andaliman memiliki potensi sebagai antioksidan dan penghambat glukosidase. (Helmalia, 2019).

Banyak penelitian tentang klasifikasi rempah dengan metode yang beragam, Pada penelitian ini penulis ingin menggunakan metode *K-Nearest Neighbor* (K-NN)

dan ekstraksi fitur *Histogram of Oriented Gradient* (HOG) untuk mengklasifikasi rempah.

Metode K-NN pernah digunakan untuk mengklasifikasi penyakit *Disk Hernia* dan *Spondylolisthesis* pada tulang belakang pengujian menggunakan validasi silang k-fold (k-10), dengan *matriks confusion* 310 data terdiri dari 60 data pasien herni, 150 data pada pasien spondylolisthesis dan 100 normal data pasien memiliki akurasi pengklasifikasi K-NN sebesar 83%. Hal ini membuat K-NN dengan k=5 membutuhkan keputusan dari 5 tetangga terdekat yang mungkin tidak memiliki kedekatan tertinggi tetapi mungkin memiliki kelas yang paling sering dalam set solusi. Dengan memiliki nilai akurasi 83%. (Handayani, 2019).

Pada penelitian yang dilakukan oleh (R A Saputra, 2020) membahas tentang Klasifikasi Gambar Penyakit Daun Padi Menggunakan K-NN Berdasarkan Ekstraksi Fitur Glem, penelitian ini bertujuan untuk mendeteksi mekanisme untuk mengurangi risiko gagal panen dengan menggunakan deteksi klasifikasi penyakit daun padi yang terdiri dari tiga penyakit yaitu Hawar daun bakteri, bercak coklat, dan jelaga daun. Pada kasus ini peneliti menggunakan metode GLCM sebagai ekstraksi fitur untuk analisis text, dengan lima nilai fitur yang terdiri dari kontras, energi, entropi, homogenitas, dan korelasi. Metode K-NN digunakan untuk klasifikasi penyakit daun padi, dengan menemukan nilai k maksimal dari percobaan nilai k 1 sampai 20. Dengan hasil percobaan menunjukkan bahwa nilai k = 11 memiliki nilai akurasi tertinggi dibandingkan dengan nilai k lainnya sebesar 65,83% dan kappa 0,485.

Pada penelitian (Wiguna, 2018) membahas tentang deteksi deteksi katarak berbasis citra digital menggunakan nilai statistik ekstraksi tekstur (*Gray Level Co-Occurance Matrix*) dengan klasifikasi jarak minkowski hasil yang diperoleh sistem ini mampu untuk mendeteksi dan mengklasifikasi mata normal, katarak imatur dan katarak matur dengan akurasi 94,74% pada variasi nilai statistik tekstur dengan jarak pixel antar tetangga 5 pixel dengan klasifikasi jarak *minkowski* pada koefisien minkowski $\lambda = 1$, $\lambda = 2$ dan $\lambda = 3$. Pada penelitian (Istiqalaliyyah, 2019) membahas tentang Pencarian teks pada terjemahan ayat *Al-Qur'an* dengan melakukan

preprocessing teks dengan term-term yang dihasilkan oleh *pre-processing* jika dilakukan pembobotan dengan menggunakan TF-RF dan *Bray-Curtis separate* untuk mengukur jarak antara dokumen dan *inquiry*. Hasil pengujian variasi nilai k-rank menghasilkan bahwa saat k-rank 5 menghasilkan nilai outline terbaik sebesar 60,36% dibandingkan saat k-rank 10, 15 dan 20. Hasil pengujian dengan menggunakan dataset yang telah dilakukan proses cleaning, pembobotan term TF-RF, dan *Bray-Curtis remove* dapat sedikit meningkatkan nilai MAP pada k-rank 5 sebesar 61,80%. Hasil pengujian dengan menggunakan Kappa Measurement menghasilkan nilai kappa 0,8774.

Pada penelitian (Santosh, 2019) membahas tentang deteksi wajah dan mengenali enam ekspresi wajah universal seperti senang, marah, jijik, takut, terkejut, sedih dengan wajah netral menggunakan pengklasifikasi jarak canberra untuk mengklasifikasikan ekspresi ke dalam emosi yang berbeda dan hasil akurasi 90,5%. Pada penelitian (Brotosaputro, 2022) membahas tentang klasifikasi artikel hoax pada sosial media dengan menggunakan similarity Jacard dan algoritma stemming hasil pengujian mendapat nilai presisi sebesar 92%, recall 80 %, dan akurasi 87 %.

Pada Penelitian (Malik, 2018) membahas tentang sistem prediksi kelulusan dengan menggunakan metode K-NN dan menggunakan pengukuran jarak *manhattan distance*. uji coba menggunakan set data siswa dari 3 tahun terakhir yang berjumlah 124 jurusan IPA dan 270 jurusan IPS. dihasilkan rata-rata akurasi pada jurusan IPA 68,29% pada tahun ajaran 2013/2014, 82,52% pada tahun ajaran 2014/2015, 84,88% pada tahun ajaran 2015/2016, dan untuk jurusan IPS 74,71% pada tahun ajaran 2013/2014, 77,89% pada tahun ajaran 2014/2015, 70,23%

Histogram of Oriented Gradient (HOG) diperkenalkan pertama kali oleh Dalal dan Triggs guna mendeteksi tubuh manusia, *Histogram of Oriented Gradien* (HOG) adalah sistem dari local objek dengan nilai yang dipakai dari intensitas gradien untuk mengekstraksi fitur yang digunakan pada computer vision dan image processing. Kelebihan *Histogram of Oriented Gradient* mampu menangkap tepi atau struktur gradient yang sangat karakteristik dari bentuk sebenarnya (Yohannes, 2019). Teknik

ekstraksi fitur *Histogram of Oriented Gradient* (HOG) menggunakan distribusi arah gradien sebagai penting seperti gradien berorientasi diekstraksi menggunakan *Histogram of Oriented Gradient* (HOG). Gradien gambar bermanfaat karena besarnya gradien besar di sekitar tepi dan sudut karena hanya ada perubahan intensitas secara tiba-tiba (Pravin, 2019).

Metode ekstrasi *Histogram of Oriented Gradient* (HOG) pernah digunakan untuk Klasifikasi Tumor Otak Menggunakan Ekstraksi Fitur *Histogram of Oriented Gradient* (HOG) dan *Support Vector Machine* penelitian untuk mendeteksi jenis Tumor otak dengan menggunakan ekstraksi fitur *Histogram of Oriented Gradient* (HOG) dan *Support Vector Machine* (SVM). penelitian ini adalah untuk mendeteksi jenis Tumor otak dengan menggunakan ekstraksi fitur *Histogram of Oriented Gradient* (HOG) dan *Support Vector Machine* (SVM). Dalam penelitian ini, proses klasifikasi dilakukan dalam tiga langkah, *Pre-processing* yang mengubah ukuran citra, ekstraksi fitur yang mengambil informasi (fitur), kemudian diikuti dengan melatih data dan dilakukan pengujian pada SVM. Hasil dari penelitian ini adalah untuk mengklasifikasikan jenis tumor otak. Dari hasil pengujian klasifikasi dengan SVM, didapatkan tingkat akurasi yang cukup baik sebesar yaitu 91% (Azizah, 2022).

Pada penelitian oleh (Tunnisa, 2022) membahas tentang senyawa antioksidan dan anti diabetes dalam beberapa rempah di Indonesia yang kurang dimanfaatkan dengan mengidentifikasi senyawa dalam 12 rempah yang ditanam di Indonesia. Pada Penelitian (Kamila Kucharska-Ambrozej, 2020) membahas tentang Penerapan teknik spektroskopi dalam kombinasi dengan kemometrik untuk mendeteksi pemalsuan beberapa herbal dan rempah-rempah dikombinasikan dengan pendekatan kemometrik untuk penilaian kualitas produk herbal, rempah-rempah dan herbal disajikan teknik spektral dengan metode statistik canggih memenuhi semua persyaratan untuk menyelesaikan masalah yang terkait dengan evaluasi cepat sejumlah besar sampel.

Pada peneltian (Shalaka P.Deore, 2019) membahas tentang pengenalan tulisan tangan aksara devanagari salah satu aksara standar terkemuka di India yang digunakan untuk banyak bahasa menggunakan fitur *histogram of oriented gradient* dan juga

memeriksa efisiensi fitur histogram of oriented gradient di atasnya menggunakan beberapa sel deskriptor histogram of oriented gradient tulisan Tangan *Character Recognition System (HDCRS)* digunakan untuk mengenali karakter yang terisolasi fitur yang diekstraksi adalah diklasifikasikan menggunakan pengklasifikasi mencapai akurasi tertinggi: tingkat pengenalan 87,38%. + penelitian dapat membantu dalam digitalisasi dokumen devanagari tulisan tangan. Pada penelitian (Yohannes, 2019) membahas tentang Klasifikasi Wajah Hewan Mamalia tampak depan menggunakan metode K-NN dengan ekstraksi fitur *histogram of oriented gradient (HOG)*, dengan mengklasifikasikan jenis hewan mamalia, klasifikasi jenis hewan mamalia mendapatkan tingkat akurasi yang baik, yaitu mencapai 100%. Perhitungan presisi, recall, dan akurasi yang baik, yaitu pada $k=3$ mendapatkan hasil di atas 50%. Untuk perhitungan presisi, recall, dan akurasi yang paling rendah, yaitu pada $k=7$ rata-rata mendapatkan hasil di bawah 50%.

Rempah rempah memiliki karakteristik, bentuk, dan warna yang hampir mirip dan sulit untuk membedakan antara rempah satu dengan rempah yang lainnya. Untuk membantu dalam mengenali karakteristik dari rempah-rempah yang ada maka penulis mencoba melakukan penelitian dengan judul. “Klasifikasi Rempah Menggunakan Metode *K-Nearest Neighbor (K-NN)* Dan Menggunakan Ekstraksi Fitur *Histogram oriented of Gradient (HOG)*. Berdasarkan pada penelitian – penelitian sebelumnya, metode K-NN dan ekstraksi fitur *Histogram of oriented gradient (HOG)* memiliki tingkat akurasi yang baik untuk mengenali warna serta bentuk dan metode klasifikasi K-NN baik dalam mengklasifikasikan objek.

1.2 Rumusan Masalah

Berdasarkan Latar belakang diatas maka rumusan masalah dalam penelitian ini bagaimana menganalisis kinerja metode K-NN dengan variasi formula jarak *Minkowski*, jarak *Braycurtis*, Jarak *Canberra*, Jarak *Jaccard*, jarak *Euclidean*, dan jarak *Manhattan* dengan memanfaatkan ekstraksi fitur *Histogram of Oriented Gradient (HOG)* untuk mengklasifikasi jenis rempah.

1.3 Tujuan

Berdasarkan rumusan masalah diatas maka tujuan dari penelitian ini adalah untuk mengevaluasi metode K-NN dengan ekstraksi fitur *Histogram of Oriented Gradient* (HOG) sejauh mana dapat mengklasifikasi jenis rempah.

1.4 Manfaat Penelitian

Berdasarkan tujuan diatas maka manfaat dalam penelitian ini adalah untuk memperoleh capaian dari metode K-NN dengan ekstraksi fitur *Histogram of Oriented of Gradien* (HOG) dalam mengklasifikasi jenis rempah.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Pada penelitian ini, rempah yang digunakan adalah rempah- rempah yang memiliki bentuk dan ukuran yang berbeda beda yaitu Jahe, Kunyit, Lengkuas, Kencur, Temulawak, Pala, Kemiri, Jinten, Adas manis, Ketumbar, Daun Salam, Kapulaga, Sere, Merica, dan Cengkeh.
2. Citra yang di input adalah gambar rempah yang telah diambil menggunakan smartphone.
3. Format gambar yang digunakan adalah JPG.
4. Pengambilan citra tidak lebih dari 30 cm.
5. Jumlah sampel yang digunakan sebanyak 2250 citra rempah dengan 150 sampel per 15 jenis rempah.
6. Software yang digunakan dalam penelitian ini adalah *Google Collab*

BAB II

LANDASAN TEORI

2.1 *K- Nearest Neighbor (K-NN)*

Metode K-NN merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data yang jaraknya paling dekat dengan objek tersebut (Yeray Mezquita, 2020). Metode K-NN memiliki beberapa keunggulan, yaitu, pelatihannya sederhana, cepat, mudah dan dipahami.

1. Jarak *Minkowski*

Jarak *Minkowski* adalah jarak matrix umum yang dapat digunakan untuk menghitung tiga matrix jarak yang berbeda yaitu, *Manhattan distance*, jarak *Euclidean*, dan jarak *Chebyshev* (Castro, 2019)

$$D(x,y) = (\sum |x_i - y_i|^r)^{1/r} \quad \dots\dots\dots 2.1$$

Keterangan :

- D (x, y) = jarak *minkowski* antara dua titik x dan y
- x_i dan y_i = koordinat titik x dan y pada dimensi i
- r = parameter jarak yang mengontrol tingkat penekanan perbedaan antara dua titik
- || = menunjukkan nilai absolut
- \sum = menunjukkan penjumlahan dari hasil perhitungan pada setiap dimensi

Di mana x dan y adalah dua titik dalam ruang dimensi-p dan x_i dan y_i adalah koordinat titik-titik tersebut. Nilai p dapat disesuaikan untuk mengontrol jenis celah yang dibuat. Misalnya, jika $p=1$, rumus ini mengembalikan jarak *manhattan*, sedangkan jika $p=2$, rumus ini mengembalikan jarak *euclidean*.

2. Jarak *Braycurtis*

$$d(x, y) = \sum(|x_i - y_i|) / (\sum x_i + \sum y_i) \dots\dots\dots 2.2$$

Keterangan :

$D(x, y)$ = Jarak *braycurtis* antara dua sampel atau titik x dan y

x_i dan y_i = Proporsi atau kelimpahan spesies pada sampel x dan y pada dimensi i

$||$ = Menunjukkan nilai absolut

\sum = Menunjukkan penjumlahan dari hasil perhitungan pada setiap dimensi

3. Jarak *Canberra*

Jarak *canberra* adalah metrik yang digunakan untuk data yang tersebar sekitar asal. Jarak *canberra* d antara 2 vektor x dan y (Pambudi, 2018)

$$j(v_1, v_2) = \sum_{k=1}^N \frac{|v_{1k} - v_{2k}|}{|v_{1k} + v_{2k}|} \dots\dots\dots 2.3$$

Keterangan :

$j(v_1, v_2)$ = Jarak *canberra* antara dua vector v_1 dan v_2

k = Indeks dimensi vektor

v_{1k} dan v_{2k} = Nilai komponen ke -k dari vector v_1 dan v_2

$||$ = Nilai absolut

4. Jarak *Manhattan*

Ukuran jarak *Manhattan* menghitung jarak antara dua titik data sebagai jumlah dari perbedaan absolut koordinat kartesius mereka. Ketika titik data memiliki dimensi yang lebih tinggi, maka metrik jarak *manhattan* lebih disukai daripada metrik distance *euclidean*. Menetapkan nilai p sebagai 1 dalam rumus jarak *Minkowski*, memberi kita jarak *manhattan* antara vektor x dan y (M Faisal, 2019) diberikan oleh rumus :

$$D(x,y) = \sum_{i=1}^n |x_i - y_i| \dots\dots\dots 2.5$$

Keterangan:

$D(x,y)$ = jarak *manhattan* antara dua vector x dan y

x_i = Komponen ke- i dari vector x

y_i = Komponen ke- i dari vector y

$| |$ = Nilai absolut dari selisih antara dua komponen

5. Jarak *Euclidean*

Euclidean yaitu jarak sederhana atau tidak secara eksplisit dihitung dalam 1, 17, 18, dan 24. *Euclidean* adalah jarak terpendek dari sampel ke model (Batzelis, 2022)

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \dots\dots\dots 2.6$$

Keterangan :

x_i =Koordinat x untuk fasilitas i

y_i = Koordinat y untuk fasilitas i

d_{ij} = Jarak antar fasilitas I dan j

2.2 Histogram of Gradients (HOG)

Histogram Of Oriented Gradients (HOG) yaitu sebuah teknik ekstraksi fitur dalam pengolahan citra dengan mengelompokkan nilai gradien pixel menurut orientasi arah pada setiap bagian lokal dari citra (Dani Devito, 2019). Ekstraksi fitur *Histogram of Oriented Gradient* (HOG) merupakan sebuah metode untuk mendeteksi gradien vertikal dan horizontal pada objek sebuah objek (Nugraha, N. D., Utaminingrum, F., & Fitriyah, H. , 2021). *Histogram of Oriented Gradient* (HOG) mempunyai langkah-langkah untuk menghitung nilai gradien pada daerah tertentu. Nilai gradien yang dihasilkan dari perhitungan tiap piksel akan dibuat menjadi sebuah histogram. Langkah pertama yaitu menentukan cell size dan block size. Kemudian hitung nilai gradien dan orientasi gradien sesuai dengan cell yang telah ditentukan. Setiap cell tersebut akan dibuat histogram untuk mengetahui nilai dalam tiap cell. Langkah selanjutnya, menentukan bin orientasi dari nilai gradien yang berbeda untuk setiap cell (Anazmar, 2019).

Langkah pertama dalam implementasi *Histogram of Oriented Gradient* (HOG) ialah perhitungannya nilai gradien pada citra yang telah diterapkan filtering citra grayscale dengan menggunakan operasi central differences dengan matriks $[-1,0,1]$ yang diterapkan pada arah vertikal dan horizontal pada suatu piksel dengan dimana x adalah baris matriks dan y merupakan kolom matriks. Kemudian akan dicari besar magnitude ($|G|$) dan orientasi gradien (θ) (Nadya Sindi Safitri, 2019) dengan menggunakan rumus:

2.3 Rempah - Rempah

Rempah-rempah adalah sumber daya hayati yang sejak lama telah berperan penting dalam kehidupan manusia. Rempah-rempah merupakan bagian tumbuhan yang bersifat aromatik dan dapat digunakan sebagai bumbu, penguat cita rasa, pengharum, dan pengawet makanan yang digunakan secara terbatas. Penggunaan rempah-rempah

dalam seni kuliner telah diketahui secara luas yang berasal dari bagian batang, daun, kulit kayu, umbi, rimpang, akar, biji, bunga atau bagian-bagian tubuh tumbuhan lainnya (Yohanes Robi, 2019).

Tumbuhan rempah biasanya dapat digunakan dalam keadaan kering maupun basah sebagian besar rempah hadir dalam bentuk kering. Bagian dari tumbuhan yang digunakan sebagai rempah terdiri atas akar, batang, kulit kayu, daun, bunga, biji, umbi, dan rimpang (Pramesthi, 2020)

2.4 Penelitian Terdahulu

Tabel 2. 1 Penelitian Terdahulu

No	Penulis & Tahun	Judul	Keterangan
1.	(Wulandari, 2020)	Klasifikasi Citra Digital Bumbu Dan Rempah Dengan Algoritma <i>Convolutional Neural Network (CNN)</i>	Membahas tentang pengenalan bumbu rempah dengan menggunakan metode <i>Convolutional Neural Network</i> dengan mendapatkan hasil pengujian data baru dengan 3 citra untuk masing-masing kategori menghasilkan akurasi sebesar 88,89%.
2.	(Kaharruddin, 2019)	Klasifikasi jenis Rempah- Rempah Berdasarkan Fitur Warna <i>RGB</i> Dan Tekstur Menggunakan <i>K-NN</i>	Membahas tentang klasifikasi rempah-rempah berdasarkan warna <i>RGB</i> dan <i>Texture</i> menggunakan <i>K-NN</i> dan pengukuran jarak menggunakan <i>Euclidean Distance</i> , dari 30 kali percobaan pengujian didapatkan hasil bahwa tingkat kebenaran dengan $K = 1$ adalah 76%, $K = 3$ sama dengan 67% dan $K = 5$ sebesar 63%.

3.	(Tanuwijaya, 2021)	Modifikasi Arsitektur <i>VGG16</i> Untuk Klasifikasi Citra Digital Rempah- Rempah Indonesia	Membahas tentang klasifikasi rempah dengan menggunakan <i>Convolution Neural Network</i> dengan arsitektur <i>VGG 16</i> dan. mendapatkan hasil evaluasi rata-rata akurasi sebesar 81%, nilai recall sebesar 76%, dan nilai presisi sebesar 81%.
4.	Yijie Dang, 2018	<i>Image Classification Based on Quantum K-Nearest-Neighbor Algorithm</i>	Membahas tentang komputasi parallel kuantum untuk mengoptimalkan efisiensi dari klasifikasi gambar menggunakan metode <i>K-NN</i> dengan hasil akurasi klasifikasi 83,1% pada dataset <i>Graz-01</i> dan 78%.
5.	Rodrigues, 2018	<i>Combining Minkowski and Chebyshev: New Distance Proposal and Survey of Distance Metrics Using K-Nearest Neighbours Classifier</i>	Membahas tentang survei matriks menggunakan algoritma <i>K-NN</i> dengan menggabungkan <i>Minkowski</i> dan <i>Chebyshev</i> dengan Percobaan ekstensif mengenai 33 set data repositori UCI dilakukan untuk mengevaluasi efisiensi jarak yang diusulkan menggunakan <i>K-NN</i> . Jarak baru mengacu yang lebih baik daripada rata-rata lebih sering daripada (26 kasus dalam 33). Selanjutnya, ia memperoleh akurasi terbaik lebih sering (9 kasus dalam 33).

6.	T Mustaqim, 2020	<i>Twitter Text Mining for Sentiment Analysis on Government's Response to Forest Fires with Vader Lexicon Polarity Detection And K-Nearest Neighbor Algorithm</i>	Membahas tentang proses analisis sentimen terhadap respons pemerintah terhadap penanganan kebakaran hutan di Indonesia pada tahun 2019 dengan sumber data dari Twitter. Analisis dilakukan pada 6325 dataset yang ditulis di Twitter pada 20 September 2019, Hasil pengujian yang dibangun menggunakan alat rapidminer menunjukkan tingkat akurasi algoritma <i>K-NN</i> sebesar 79,45%.
7	Rico Andrian, 2020	<i>Butterfly Identification Using Gray Level Co-Occurrence Matrix (GlcM) Extraction Feature And K-Nearest Neighbor (K-NN) Classification</i>	Membahas tentang identifikasi kupu-kupu pada penangkaran Taman Gita Persada yang memiliki kurang lebih 211 spesies kupu-kupu.. Hasil penelitian ini menunjukkan bahwa kelas <i>Centhosia penthesilea</i> dan <i>Papilio nephelus</i> dapat diklasifikasikan dengan baik dibandingkan dengan 4 kelas lainnya. Akurasi tertinggi adalah 91,1% dengan nilai $\theta=5$ pada sudut 90° dan error rate=8,9%.
8.	Firman Tempola, 2018	Perbandingan Klasifikasi Antara <i>K-NN</i> dan <i>Naive Bayes</i>	Membahas tentang Penentuan Status Gunung Berapi menggunakan dua algoritma klasifikasi yaitu <i>K-NN</i> dan <i>Naive Bayes Classifier</i> pada aktivitas

		<p>pada Penentuan Status Gunung Berapi dengan <i>K-Fold Cross Validation</i></p>	<p>status gunung berapi yang ada di Indonesia. klasifikasi aktivitas gunung berapi menggunakan sistem kegunaan. Ada 5 kriteria yang digunakan dalam melakukan klasifikasi yaitu empat system kegunaan diantaranya gempa vulkanik dangkal, gempa tektonik jauh, gempa vulkanik dalam, gempa hembusan dan ditambah satu kriteria yaitu status sebelumnya. Ada 3 status yang di yang diklasifikasi yaitu normal, waspada dan siaga. Hasil penelitian klasifikasi perbandingan sistem memperoleh k-nn 63,68% dengan standar deviasi 7,47 %.</p>
9.	Yohannes, 2019	<p>Klasifikasi Mamalia Berdasarkan Bentuk Wajah Dengan <i>K-NN</i> Menggunakan Fitur <i>CAS</i> Dan <i>Histogram Of Oriented Gradient (HOG)</i></p>	<p>Klasifikasi Wajah Hewan Mamalia Tampak Depan Menggunakan <i>K-NN</i> Dengan Ekstraksi Fitur <i>Histogram Of Oriented Gradient (HOG)</i>, dengan mengklasifikasikan jenis hewan mamalia mendapatkan tingkat akurasi yang baik, yaitu mencapai 100%. Perhitungan presisi, recall, dan akurasi yang baik, yaitu pada k=3 mendapatkan hasil di atas 50%.</p>

<p>10.</p>	<p>R A Saputra, 2020</p>	<p><i>Rice Leaf Disease Image Classifications Using K-NN Based on GLCM Feature Extraction</i></p>	<p>Membahas tentang deteksi mekanisme untuk mengurangi risiko gagal panen dengan menggunakan deteksi klasifikasi penyakit daun padi yang terdiri dari tiga penyakit yaitu Hawar daun bakteri, bercak coklat, dan jelaga daun. Dengan menggunakan metode <i>GLCM</i> sebagai ekstraksi fitur untuk analisis text, dengan lima nilai fitur yang terdiri dari kontras, energi, entropi, homogenitas, dan korelasi. <i>K-NN</i> digunakan untuk klasifikasi penyakit daun padi, dengan menemukan nilai k maksimal dari percobaan nilai k 1 sampai 20. Dengan hasil percobaan menunjukkan bahwa nilai k = 11 memiliki nilai akurasi tertinggi dibandingkan dengan nilai k lainnya sebesar 65,83% dan kappa 0,485.</p>
------------	--------------------------	---	--

BAB III

METODOLOGI PENELITIAN

3.1 Alat dan Bahan Penelitian

Dalam penelitian ini, penulis menggunakan peralatan dan bahan pendukung untuk membantu penelitian ini agar dapat berjalan dengan lancar. Adapun perangkat keras dan perangkat lunak yang digunakan dalam penelitian ini adalah sebagai berikut:

3.1.1 Perangkat Keras

3.1.2 Perangkat Lunak

Tabel 3. 1 Perangkat Keras yang digunakan

No	Perangkat Keras	Keterangan
1.	Device	Laptop Asus VivoBook X509FAC_A509FA
2.	Processor	Intel(R) Core (TM) i3-10110U CPU @ 2.10GHz 2.50 GHz
3.	SSD	254 GB
4.	RAM	4.00 GB

Tabel 3. 2 Perangkat Lunak yang digunakan

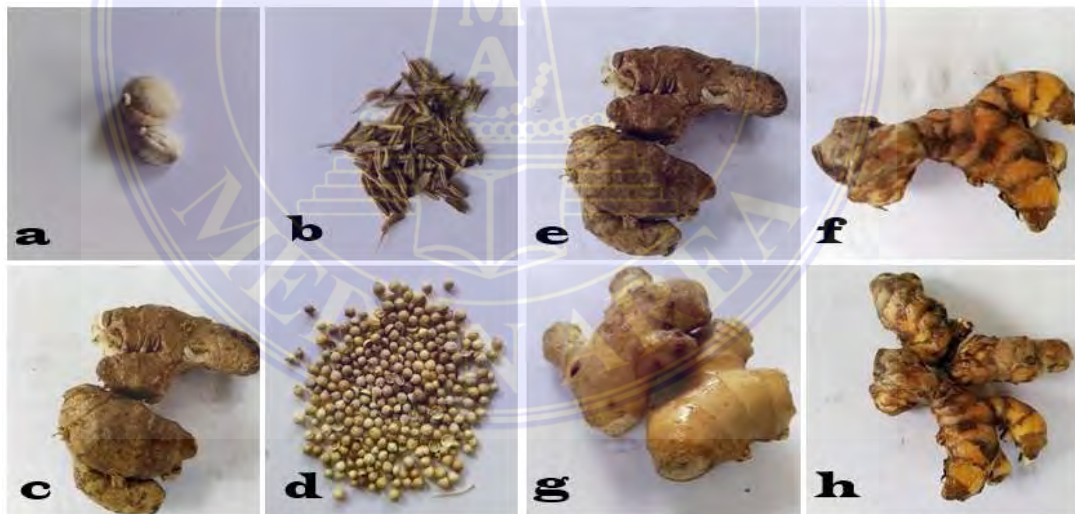
No	Perangkat Lunak	Keterangan
1.	Sistem Operasi	Windows 11 Home Single Language
2.	Goole Collab	- RAM: 12 GB - Disk: 107 GB
4.	RAM	4.00 GB

3.2 Metode Pengumpulan Data

3.2.1 Dataset

Data yang digunakan dalam penelitian ini berupa citra rempah atau gambar daun rempah-rempah. Peneliti melakukan riset atau pengambilan gambar citra rempah secara langsung di Pasar Lau Chi Medan, proses pengambilan gambar untuk keperluan dataset ini mencapai waktu selama satu bulan dan berjalan dengan lancar.

Pada saat pengambilan gambar, banyak ditemukan rempah rempah yang hampir sama seperti adas manis dengan jinten, ketumbar dan merica, kemiri dan buah pala dan lain lain. Dalam penelitian sampel citra yang digunakan dalam penelitian ini berjumlah 2250 yaitu Lengkuas, Kapulaga, Jahe, Merica, Ketumbar, Kencur, Jinten, Kunyit, Temulawak, Pala, Kemiri, Adas Manis, Daun Salam, Sere, dan Cengkeh. Gambar rempah dibagi menjadi 15 jenis, per tiap jenis berjumlah 150 gambar. Data yang diambil dalam penelitian ini yaitu klasifikasi rempah dengan hasil foto sampel rempah-rempah menggunakan kamera Hp Vivo Y35.



Gambar 3. 1 Sampel Rempah

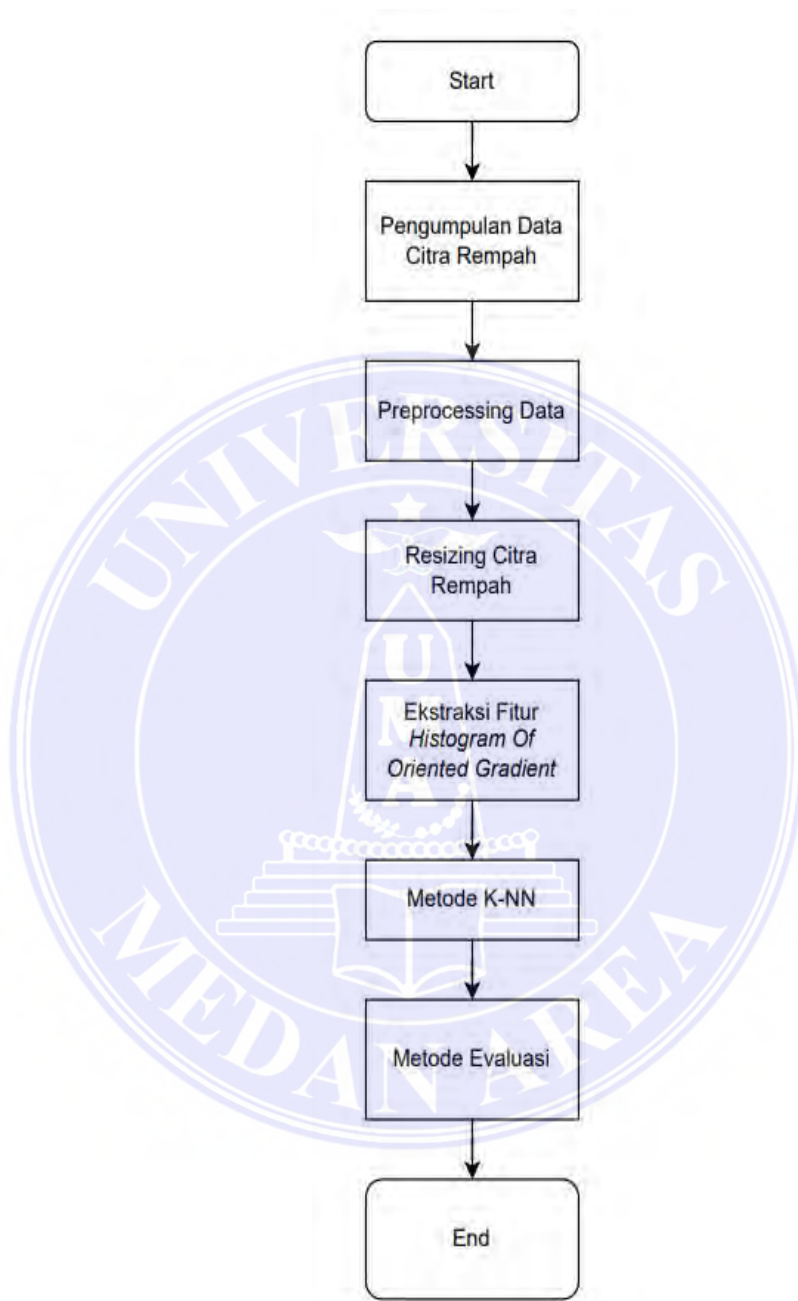
3.2.2 Data Analis

Sampel gambar yang digunakan sebagai model uji coba sebanyak 2250 sampel Model Evaluasi yang digunakan untuk penelitian ini dengan membagikan data menjadi dua kelompok yaitu training dan data testing, 80% data sampel data dijadikan sebagai data training sedangkan untuk 20 % sampel data dijadikan sebagai data testing. Gambar dibawah ini melampirkan sampel rempah yang digunakan sebagai data training dan testing.

Tabel 3. 3 Pembagian data

No	Jenis Rempah	Data Training	Data Testing
1.	Lengkuas	120	30
2.	Kapulaga	120	30
4.	Jahe	120	30
5.	Merica	120	30
6.	Ketumbar	120	30
7.	Merica	120	30
8.	Kencur	120	30
9.	Jinten	120	30
10.	Kunyit	120	30
11.	Temulawak	120	30
12.	Pala	120	30
13.	Kemiri	120	30
14.	Adas Manis	120	30
15.	Cengkeh	120	30

3.3 Diagram Alur Penelitian

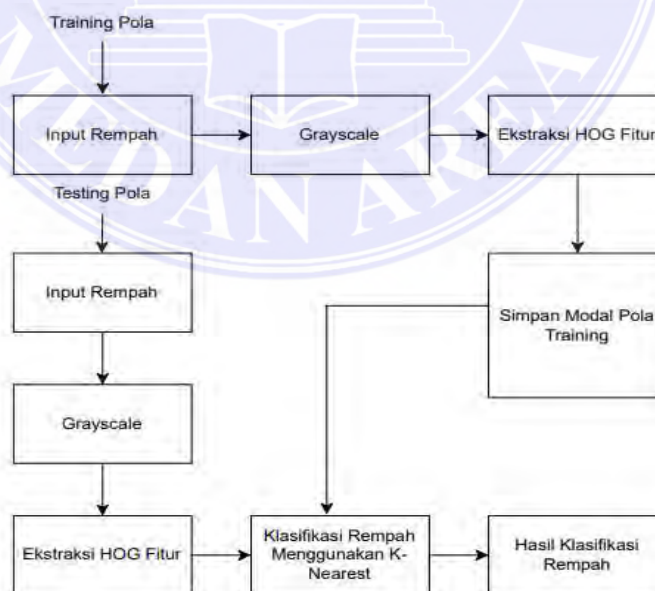


Gambar 3. 2 Diagram Alur Penelitian

Pada gambar 3.2 menjelaskan proses dimulai Langkah pertama yaitu "Pengumpulan Data Rempah", di mana data rempah dikumpulkan data gambar rempah. Setelah itu, data rempah melewati tahap "*Preprocessing Data*", di mana data tersebut diproses sebelum masuk ke langkah berikutnya. Langkah selanjutnya adalah "*Resizing Citra*", di mana citra-citra rempah diubah ukurannya agar memiliki ukuran yang seragam. Setelah itu, dilakukan "Ekstraksi Fitur HOG" dengan menggunakan metode *Histogram of Oriented Gradients*. Langkah ini bertujuan untuk mengekstraksi fitur-fitur penting dari citra rempah, seperti pola atau tekstur yang spesifik, langkah berikutnya adalah "Metode K-NN". Metode K-NN digunakan untuk melakukan klasifikasi atau pengenalan rempah berdasarkan fitur-fitur yang telah diekstraksi sebelumnya. Selanjutnya, hasil klasifikasi dievaluasi dengan menggunakan "Metode Evaluasi". Metrik evaluasi seperti *akurasi*, *presisi*, *recall*, atau *F1-score* dapat digunakan untuk mengukur kinerja model klasifikasi. Proses berakhir di langkah "*End*", yang menjadi titik akhir dari aliran proses pengolahan data rempah.

3.3 Arsitektur Penelitian

Langkah penelitian secara umum yang dirancang dalam penelitian ini sebagai berikut



Gambar 3. 3 Langkah Penelitian

Pada gambar terdapat dua proses yaitu proses training dan proses testing, pada proses training gambar masukan dilakukan dengan *pre-procecing* menggunakan *grayscale* dan dilanjutkan dengan mengekstraksi fitur dari citra rempah menggunakan K-NN dan ekstraksi fitur *Histogram of Oriented Gradient* (HOG). pada tahap testing, citra inputan dilakukan pre- processing dengan grayscale dan di ekstraksi menggunakan *Histogram of Oriented Gradient* (HOG) kemudian dilanjutkan ketahapan pengklasifikasian model pola menggunakan *K- Nearest Neighbor*, jika pola mirip atau mendekati pola training maka output dari klasifikasi yaitu hasil klasifikasi rempah.

3.4 Metode Evaluasi

3.4.1 Confusion Matrix

Confusion matrix adalah tabel yang digunakan untuk mengimplementasikan model dari klasifikasi pada data set pengujian yang nilai nya telah diketahui oleh *confusion matrix* itu sendiri cara kerjanya sederhana dan mudah di pahami, tetapi dapat membingungkan apabila tidak mengetahui *terminology* terkait (Chicco D. T., 2021)

Tabel 3. 4 *Coufusion Matrix*

	<i>Actually Positive (1)</i>	<i>Actually Negative (2)</i>
<i>Predicted Positive</i>	<i>True Negatives (TN)</i>	<i>False Positives (FP)</i>

<i>Predicted Negative</i>	<i>False Negatives (FN)</i>	<i>True Negatives (TN)</i>
---------------------------	-----------------------------	----------------------------

Dimana *Predicted Class* mewakili jumlah instan data yang diketahui berada di grup label benar dan di prediksi berada di prediksi label.

True Positives (TP): Ini adalah titik data yang hasil aktualnya positif dan algoritma mengidentifikasinya dengan benar sebagai positif.

True Negatives (TN): Ini adalah titik data yang hasil aktualnya negatif dan algoritma mengidentifikasinya dengan benar sebagai negatif.

False Positives (FP): Ini adalah titik data yang hasil aktualnya negatif tetapi algoritma salah mengidentifikasinya sebagai positif.

False Negatives (FN): Ini adalah titik data yang hasil aktualnya positif tetapi algoritma salah mengidentifikasinya sebagai negatif.

3.4.2 Performa Measure

Dalam pengklasifikasian ada 4 nilai yang digunakan untuk mengukur kemampuan klasifikasi yang akan di olah yaitu akurasi, presicion, recal, dan f1 score.

a. Akurasi

Nilai akurasi didapatkan dari jumlah data bernilai positif yang diprediksi positif dan data bernilai negatif yang diprediksi negatif dibagi dengan jumlah seluruh data di dalam dataset (Nasution, 2019). Rumus akurasi adalah

$$Akurasi = \frac{TN+TP}{TN+FP+TP+FN} \dots\dots\dots 3.1$$

b. Presisi

Presisi merupakan parameter penilaian yang menghitung nilai rata-rata precision dari data hasil klasifikasi yaitu jumlah data yang benar antara nilai sebenarnya (Yudianto, 2020). Rumus Presicion adalah

$$Presisi = \frac{TP}{TP+FP} \dots\dots\dots 3.2$$

c. *Recall*

Recall adalah ukuran kelengkapan dari sebuah model. Persamaan recall perbandingan antara true positive terhadap total contoh yang benar-benar positif. (Prabowo, 2021) Rumus Recall adalah

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

d. *F1 Score*

Nilai *F1-Score* atau dikenal juga dengan nama *F-Measure* didapatkan dari hasil Precision dan Recall antara kategori hasil prediksi dengan kategori sebenarnya (Chicco D. &, 2020). Rumus F1 Score adalah

$$F1\ Score = \frac{2*Presicion*Recall}{Presicion+Recall} \dots\dots\dots 3.4$$

e. *Fbeta score*

Fbeta score adalah kombinasi dari presisi dan recall, di mana beta dapat membuat salah satu dari dua pengukuran (*recall* dan presisi) lebih penting dalam hasilnya. Dengan beta 1, ada keseimbangan antara kedua pengukuran. Dengan beta lebih rendah dari 1, ada lebih banyak bobot pada presisi. Beta yang lebih tinggi dari 1 memberi bobot lebih pada penarikan. Skor akan berada di antara 0 dan 1. (Haas, 2021)

$$F\beta\ score = \frac{(1 + \beta) * precision * recall}{\beta * precision + recall} \dots\dots\dots 3.5$$

f. *Jaccard score*

Jaccard score adalah matrix pertama berdasarkan score antara dua jarak (Gu, 2021) dengan rumus:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \dots\dots\dots 3.6$$

J = Jarak *Jaccard*

A = Set 1

B = Set 2



BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dalam penelitian ini, telah dilakukan klasifikasi rempah menggunakan metode *K-NN* dengan variasi jarak *Minkowski*, *Jaccard*, *Manhattan*, *Braycurtis*, *Canberra*, dan *Euclidean*, serta ekstraksi fitur *Histogram of Oriented Gradient* (HOG). Hasil dari penelitian ini menunjukkan bahwa pengujian dengan 5 jenis rempah memberikan hasil yang tertinggi dengan $k=1$ akurasi rata-rata dan jenis jarak yang tertinggi yaitu jarak *manhattan* dengan akurasi 94%, sedangkan yang mendapatkan akurasi terendah yaitu jarak *Jaccard* dengan akurasi 62%. Selain itu, pengujian dengan jumlah rempah yang lebih besar menunjukkan hasil akurasi yang lebih rendah, hal ini menunjukkan semakin kompleks data yang diuji, semakin pentingnya teknik ekstraksi fitur dan jenis jarak yang dipilih untuk memperoleh hasil klasifikasi yang lebih baik.

Dari penelitian ini, dapat disimpulkan bahwa metode *K-NN* dengan menggunakan variasi jarak dan ekstraksi fitur *Histogram of Oriented Gradient* (HOG) dapat digunakan untuk melakukan klasifikasi rempah namun, pemilihan jenis jarak dan teknik ekstraksi fitur yang tepat sangat penting dalam meningkatkan akurasi klasifikasi, Terutama pada data yang kompleks. Oleh karena itu, penelitian selanjutnya dapat dilakukan untuk mencari teknik ekstraksi fitur dan jenis jarak yang lebih optimal untuk meningkatkan akurasi klasifikasi pada data yang lebih kompleks.

5.2 Saran

Pada penelitian ini jumlah kategori data yang besar seperti pada pengujian dengan 15 jenis rempah, dapat dilakukan pengembangan teknik ekstraksi fitur yang lebih baik untuk meningkatkan akurasi klasifikasi. Selain itu, dapat pula dilakukan penambahan jenis jenis rempah dan pengujian dengan variasi parameter k yang berbeda untuk melihat pengaruhnya terhadap akurasi klasifikasi. Selain itu, untuk memperkuat hasil penelitian ini, dapat dilakukan pengujian dengan metode klasifikasi lainnya, seperti *Deep learning*.

DAFTAR PUSTAKA

- Abdurrahman, M. H. (2019). Deteksi Kualitas Kemurnian Susu Sapi Melalui Pengolahan Citra Digital Menggunakan Metode Scale Invariant Feature Transform Dengan Klasifikasi K-nearest Neighbor. *eProceedings of Engineering*, 6(2). doi:10.34818/eoe.v6i2.10227
- Adriyendi, A. &. (2020). Klasifikasi Menggunakan Naïve Bayes Dan K-Nearest Neighbor Pada Manajemen Layanan Teknologi Informasi. *urnal Teknologi Dan Sistem Informasi Bisnis*, 2 (2), 99-107. doi:https://doi.org/10.47233/jteksis.v2i2.121
- Alexander Yashin, Y. Y. (2017). A Review. Antioxidant Activity of Spices and Their Impact on Human Health. doi:https://doi.org/10.3390/antiox6030070.
- Alromih, A. &. (2019). An energy-efficient gossiping protocol for wireless sensor networks using Chebyshev distance. *Procedia Computer Science*. 151, 1066-1071. doi:https://doi.org/10.1016/j.procs.2019.04.151
- Anazmar, H. (2019). Analisis performansi Sistem Pendeteksi Kualitas kayu Jati menggunakan Pengolahan Citra Dengan Metode Histogram Of Gradients (HOG) dan Support Vector Machine. 6 (2), 3485.
- Azizah, T. A. (2022). Jurnal Infotech. *Klasifikasi Tumor Otak Menggunakan Ekstraksi Fitur HOG dan Support Vector Machine*, Vol. 4, No. 1.
- Batzelis, E. B. (2022). Noise-Scaled Euclidean Distance: A Metric for Maximum Likelihood Estimation of the PV Model Parameters. *IEEE Journal of Photovoltaics*, 815- 826.
- Besta, M. K. (2020). Communication-efficient jaccard similarity. 1122-1132. Retrieved from arXiv:1911.04200v3

- Brotosaputro, G. W. (2022). Jurnal Sisfokom (Sistem Informasi dan Komputer). *Penentuan Hoax pada Artikel Politik Berbahasa Indonesia di Sosial Media dengan Similarity Jaccard dan Algoritma Stemming*, 11(1), 79-86.
- Castro, A. Y. (2019). The Bethe-Salpeter approach to bound states: from Euclidean to Minkowski. *OP Conf. Series: Journal of Physics: Conf. Series 1*, 1291 No 1. doi:10.1088/1742-6596/1291/1/012006
- Chicco, D. &. (2020). BMC genomics. *The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation*, 21(1), 1-13. doi:<https://doi.org/10.1186/s12864-019-6413-7>
- Chicco, D. T. (2021). *The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation.*, 14 (1), 1-22. doi:<https://doi.org/10.1186/s13040-021-00244-z>
- de Guzman, C. C. (1999). Plant Resources of south-East Asia. *Spices*.
- Devito, D. W. (2019). Ekstraksi Ciri Untuk Klasifikasi Gender Berbasis Citra Wajah Menggunakan Metode Histogram of Oriented Gradients. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3 (8), 8002-8011.
- Dinata, R. K. (2020). Igoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus. *ILKOM Jurnal Ilmiah*, 12(20), 104-111. doi:<https://doi.org/10.33096/ilkom.v12i2.539.104-111>
- Gu, A. C. (2021). Bedshift: perturbation of genomic. *Genome Biology*, 22-238. doi:10.1186/s13059-021-02440-w
- Haas, Y. J. (2021). *Data Science, master's level*.
- Handayani, I. (2019). Application of K-Nearest Neighbor Algorithm on Classification of Disk Hernia and Spondylolisthesis in Vertebral Column. *Indonesian Journal*

of Information Systems, 2 (1), 57-66.
doi:<https://doi.org/10.24002/ijis.v2i1.2352>

Helmalia, A. W. (2019). Potensi rempah-rempah tradisional sebagai sumber antioksidan alami untuk bahan baku pangan fungsional):(The potential of traditional spices as a source of natural antioxidants for functional food raw materials). *Canrea Journal: Food Technology, Nutritions, and Culinary Journal*, 02(01). doi:<https://doi.org/10.20956/canrea.v2i1.113>.

Hizaz Z Yahya, Y. (2023). Klasifikasi Bumbu Dapur Pasar Menggunakan Metode Deep Neural Network Berbasis Android. *Jurnal Nasional Komputasi dan Teknologi Informasi*, 6(1).

Ikhsan Nuh Atthalla, A. J. (2018). Computer Science and ICT. *Klasifikasi Penyakit Kanker Payudara, Vol.4 No.1* .

Inna Alvi Nikmatun, I. W. (2019). Jurnal Simetris. *Implementasi Data Mining Untuk Klasifikasi masa Studi Mahasiswa menggunakan Algoritma K- Nearest Neighbor*, 10.2, 2549-3108.

Isman, A. A. (2021). Perbandingan Metode KNN Dan LBPHPada Klasifikasi Daun Herbal. *Jurnal Resti*, 5 (3), 557-564.
doi:<https://doi.org/10.29207/resti.v5i3.3006>

Istiqlaliyyah, I. S. (2019). Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. *Pencarian Teks pada Terjemahan Ayat Al-Qur'an dengan Menggunakan TF-RF dan Bray-Curtis Distance.*, 3 (5), 4357-4363.

Jiang, T. A. (2019). Journal of AOAC International. *Health Benefits of Culinary Herbs and Spices*, Vol. 102. Issue 2, 1 pages 395-411.
doi:<https://doi.org/10.5740/jaoacint.18-0418>

- Kaharruddin, K. K. (2019). Informasi Interaktif Klasifikasi Jenis Rempah-Rempah Berdasarkan Fitur Warna RGB dan Tekstur Menggunakan Algoritma K-Nearest Neighbor. *Informasi Interaktif*, 4 (1), 17-22.
- Kamila Kucharska-Ambrożej, J. K. (2020). The application of spectroscopic techniques in combination with chemometrics for detection adulteration of some herbs and spices. *Microchemical Journal*, 153. doi:<https://doi.org/10.1016/j.microc.2019.104278>.
- Luchman Hakim, J. B. (2015). Etnobotani Rempah-Rempah di Dusun Kopen Dukuh. 06, 02.
- M Faisal, E. M. (2019). Comparative Analysis of Inter-Centroid K-Means. *Journal of Physics: Conference Series*, 1566. doi:10.1088/1742-6596/1566/1/012112
- Malik, A. M. (2018). SKANIKA. *Aplikasi Prediksi Kelulusan Ujian Nasional Menggunakan Algoritma K-Nearest Neighbor dengan Pengukuran Jarak Manhattan Distance*, 1 (2), 829-835.
- Mustaqim, T. U. (2020). Twitter text mining for sentiment analysis on government's International Conference on Mathematics,. *Science, and Education*, 1567. doi:10.1088/1742-6596/1567/3/032024
- Nasution, M. R. (2019). Jurnal Informatika. *Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter*, 226-235.
- Nugraha, N. D., Utaminingrum, F., & Fitriyah, H. . (2021). Alat Pendeteksi Uang untuk Tunanetra menggunakan Metode Histogram of Oriented Gradients dan K-Nearest Neighbor. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2548, 964X.
- Opara. (2019). urnal of AOAC International. *Health Benefits of Culinary Herbs and Spices*, 102(02), 395-411. doi:<https://doi.org/10.5740/jaoacint.18-0418>.

- Pambudi, E. A. (2018). Image segmentation analysis based on K-means PSO by using three distance measures. *ICTACT Journal on Image and Video Processing*, 9 1. doi:10.21917/ijivp.2018.0256
- Prabowo, T. T. (2021). Media Pustakawan. *Efektivitas Sistem Temu Kembali Informasi Perpustakaan Digital Institut Seni Indonesia (ISI) Yogyakarta dalam Tinjauan Recall dan Precision*, 37-48.
- Pramesthi, D. A. (2020). Potensi Tumbuhan Rempah dan Bumbu yang Digunakan dalam Masakan Lokal Buton sebagai Sumber Belajar:(Potential of Herbs and Spices Used in Local Buton Cuisine as Learning Resources). *BIODIK*, 6 (03), 225-232.
- Pravin, S. P. (2019). Histogram of Oriented Gradients Based Off-Line Handwritten Devanagari Characters Recognition Using SVM, K-NN and NN Classifiers. *IITEA* , Vol. 33, No. 6. doi:https://doi.org/10.18280/ria.330606.
- Putra, I. N. (2018). Pengenalan Wajah Berbasis Mobile Menggunakan Fisherface dan Distance Classifier. *JST (Jurnal Sains dan Teknol*, 7 (1), 135-145. doi:10.23887/jst-undiksha. v7i1
- R A Saputra, S. S. (2020). Conference Series. *Conference Series Rice Leaf Disease Image Classifications Using KNN Based On GLCM Feature Extraction*. doi:https://doi:10.1088/1742-6596/1641/1/012080
- Rodrigues, E. (2018). Combining Minkowski and Chebyshev: New distance proposal and survey of distance metrics using k-nearest neighbours classifier. *Pattern Recognition Letters*, 66-71. doi:https://doi.org/10.1016/j.patrec.2018.03.021
- Rosa Vázquez-Fresno, A. R.-O. (2019). Herbs and Spices- Biomarkers of Intake Based on Human Intervention Studies – A Systematic Review. *Genes & Nutrition*.

- Safitri, N. S. (2019). Klasifikasi Umur Berdasarkan Citra Dental Panoramic Radiograph Dengan Metode Histogram Of Oriented Gradient Dan Support Vector Machine. *6 (2)*.
- Santosh, M. &. (2019). Journal of Computational and Theoretical Nanoscience. *A proposed framework for emotion recognition using Canberra distance classifier, 16 (9), 3778-3782.*
- Shalaka P.Deore, A. P. (2019, 12). Artificielle Histogram of Oriented Gradients Based Off-Line Handwritten Devanagari Characters Recognition. *Revue d'Intelligence, 33 (6), 441-446.* doi:<https://doi.org/10.18280/ria.330606>
- Shayhan Ameen Chowdhury, M. M. (2018). Human detection utilizing adaptive background mixture models and improved histogram of oriented gradients. *ScienceDirect, 216-220.* doi:10.1016/j.ict.2017.11.016.
- Tanuwijaya, E. &. (2021). Modifikasi Arsitektur VGG16 untuk Klasifikasi Citra Digital Rempah Rempah Indonesia.
- Tunnisa, F. F. (2022). Antioxidant and antidiabetic compounds identification in several Indonesian underutilized Zingiberaceae spices using SPME-GC/MS-based volatilomics and in silico methods. *Food Chemistry: X, 14.*
- Wiguna, G. A. (2018). Seminar Ilmiah Nasional Teknologi, Sains, dan Sosial Humaniora (SINTESA). *SISTEM DETEKSI KATARAK BERDASARKAN TEKSTUR GRAY LEVEL COOCURRENCE MATRIX DENGAN KLASIFIKASI JARAK MINKOWSKI, 1.*
- Wijaya, N. (2019). Klasifikasi Jenis Buah Apel Dengan Metode K-Nearest Neighbors Dengan Ekstraksi Fitur HSV dan LBP. *Jurnal Sisfokom, 8 (01).*
- Wulandari, I. Y. (2020). Klasifikasi citra digital bumbu dan rempah dengan algoritma convolutional neural network (cnn). *Jurnal Gaussian, 9 (3), 273-282.*

Yeray Mezquita, R. S.-V. (2020). *A Review of k-NN Algorithm Based on Classical and Quantum Machine Learning*.

Yijie Dang, N. J. (2018). *Image Classification Based on Quantum KNN*.

Yohanes Robi, S. M. (2019). ETNOBOTANI REMPAH TRADISIONAL DI DESA EMPOTO. *Jurnal Hutan Lestari*, 7 (1), 130-142.

Yohannes, Y. P. (2019). Jurnal Teknik Informatika dan Sistem Informasi. *Klasifikasi Wajah Hewan Mamalia Tampak Depan Menggunakan k-Nearest Neighbor Dengan Ekstraksi Fitur HOG*, Vol. 5, No. 1. doi:<http://dx.doi.org/10.28932/jutisi.v5i1>

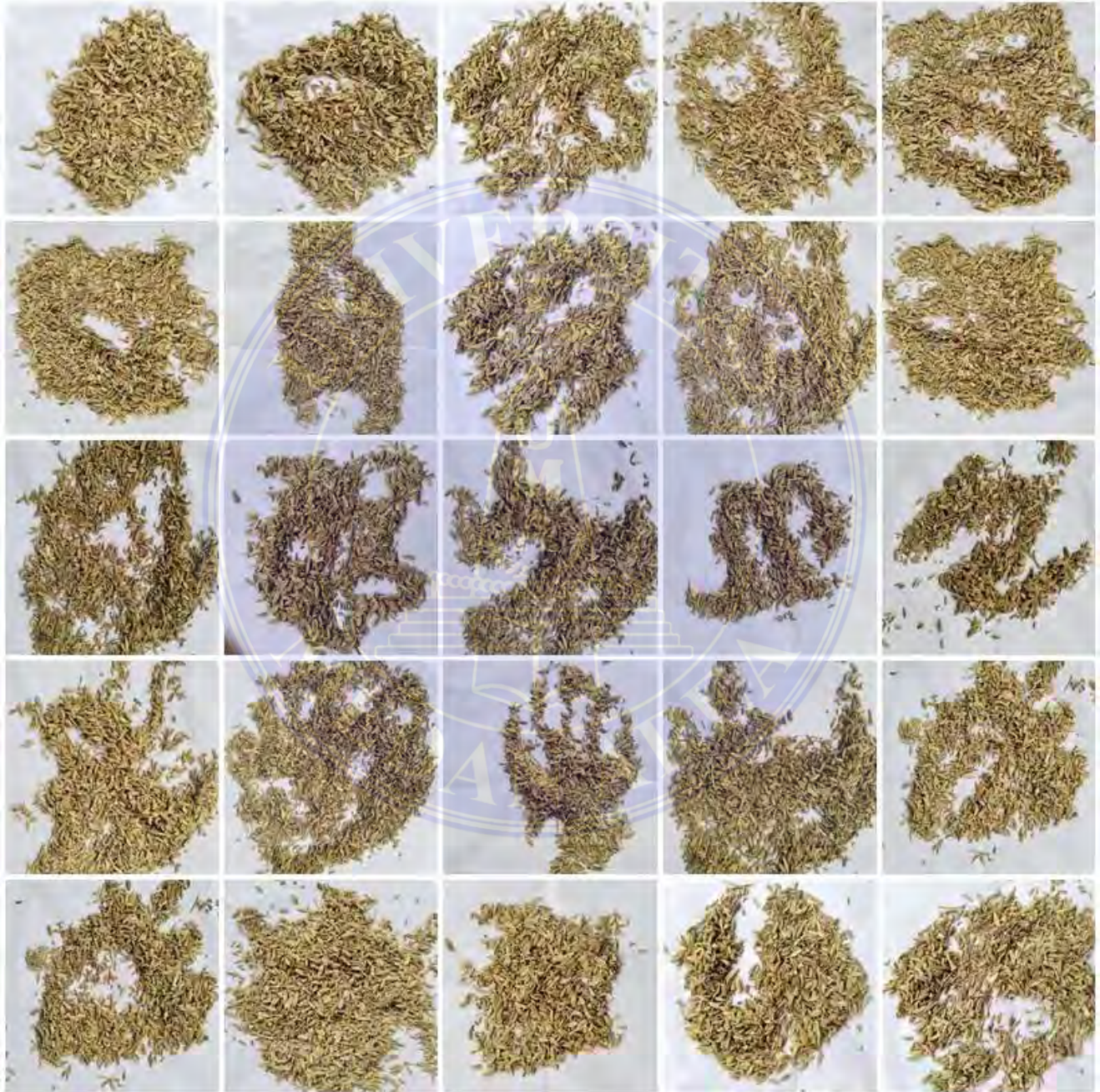
Yudianto, M. R. (2020). Analisis Pengaruh Tingkat Akurasi Klasifikasi Citra Wayang dengan Algoritma Convolutional Neural Network. (*JurTI*) *Jurnal Teknologi Informasi*, 4 (2), 182-191. doi:10.36294/jurti.v4i2.1319

Zailani, A. U. (2020). Algoritma Klasifikasi Random Forest Untuk Penentuan Kelayakan Pemberian Kredit Di Koperasi Mitra Sejahtera. *Infotech: Journal of Technology Information Penerapan*, 6 (1), 7-14.

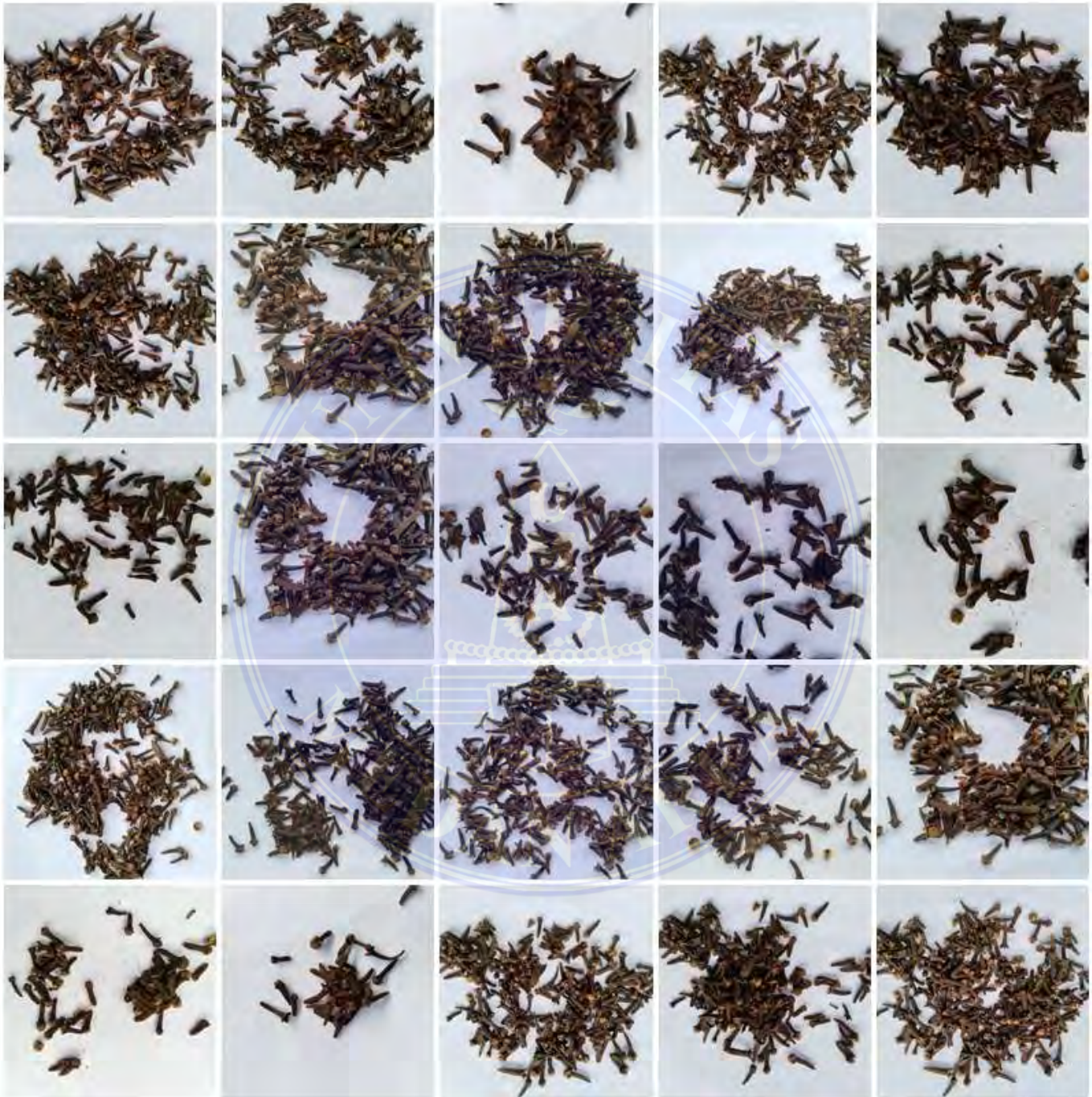
LAMPIRAN

a. Dataset

1. Adas Manis



2. Cengkeh



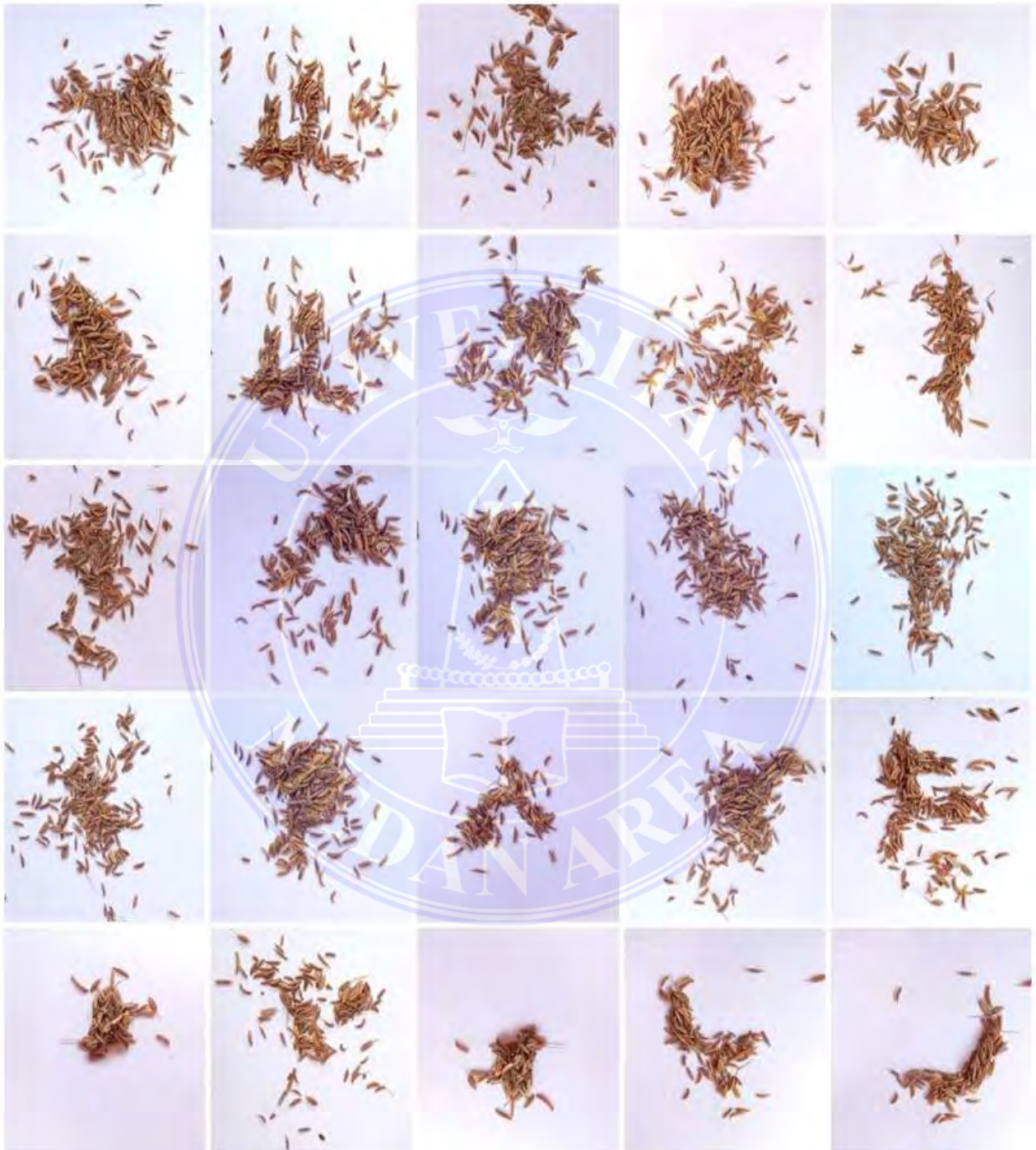
3. Daun Salam



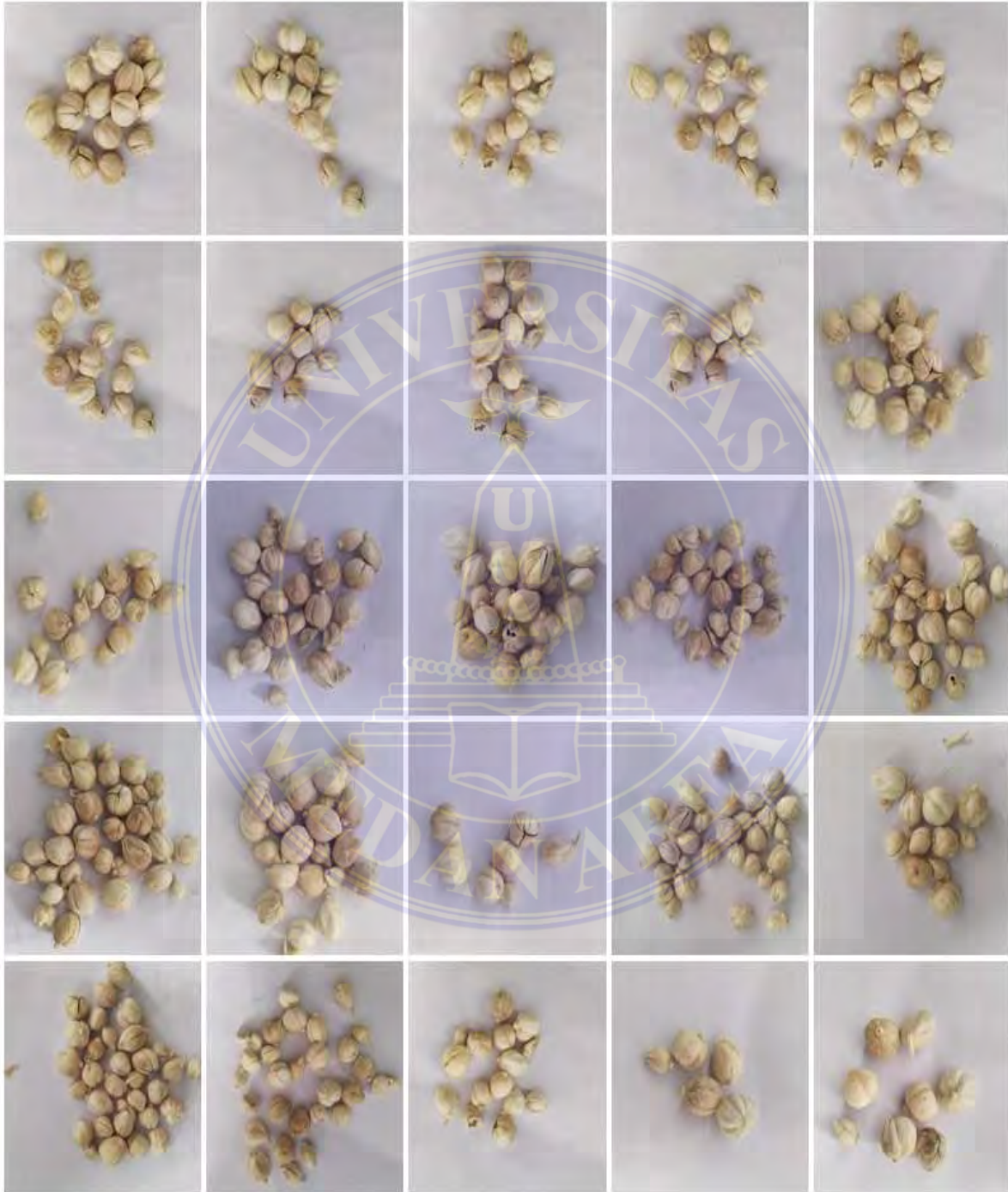
4. Jahe



5. Jinten



6. Kapulaga



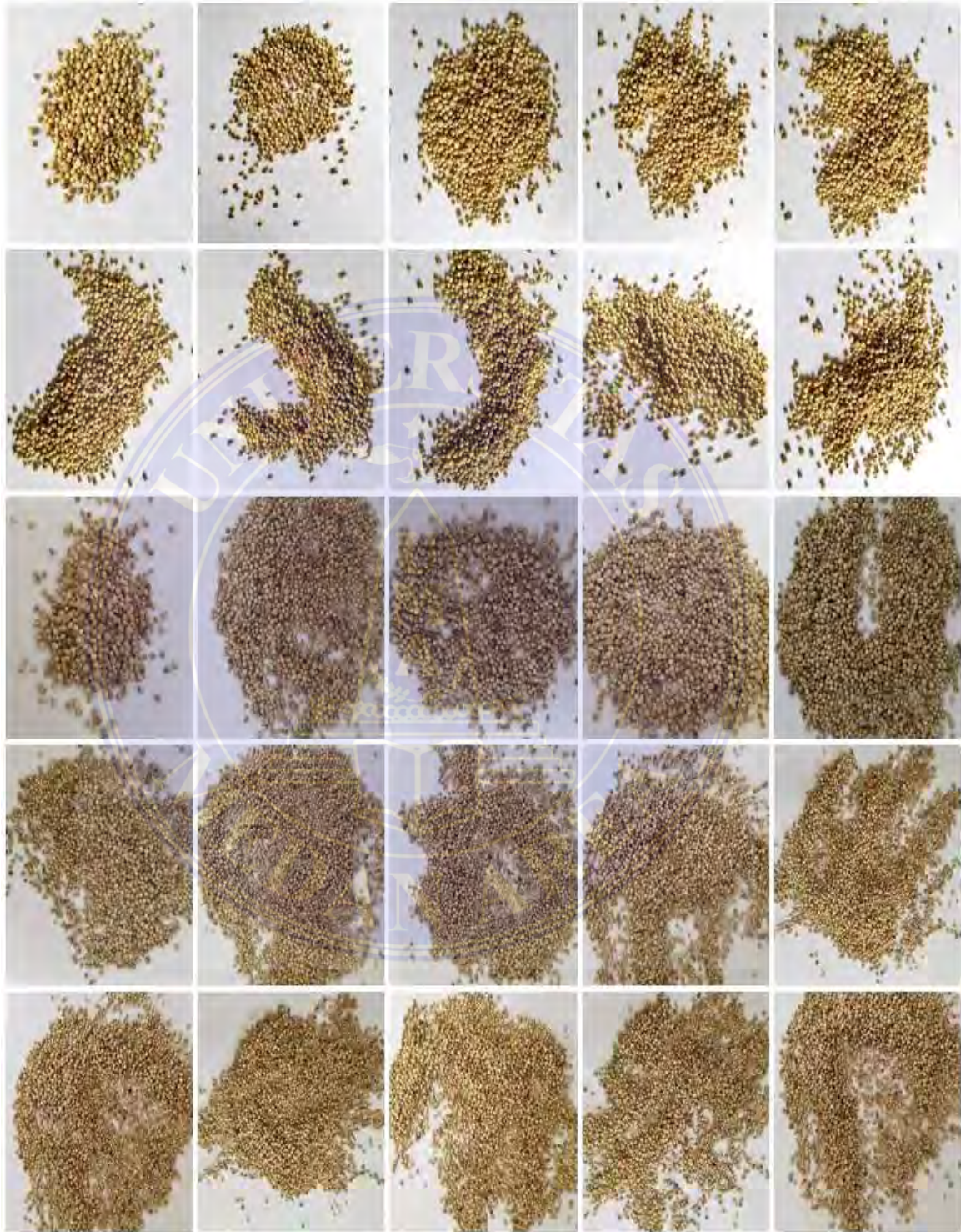
7. Kemiri



8. Kencur



9. Ketumbar



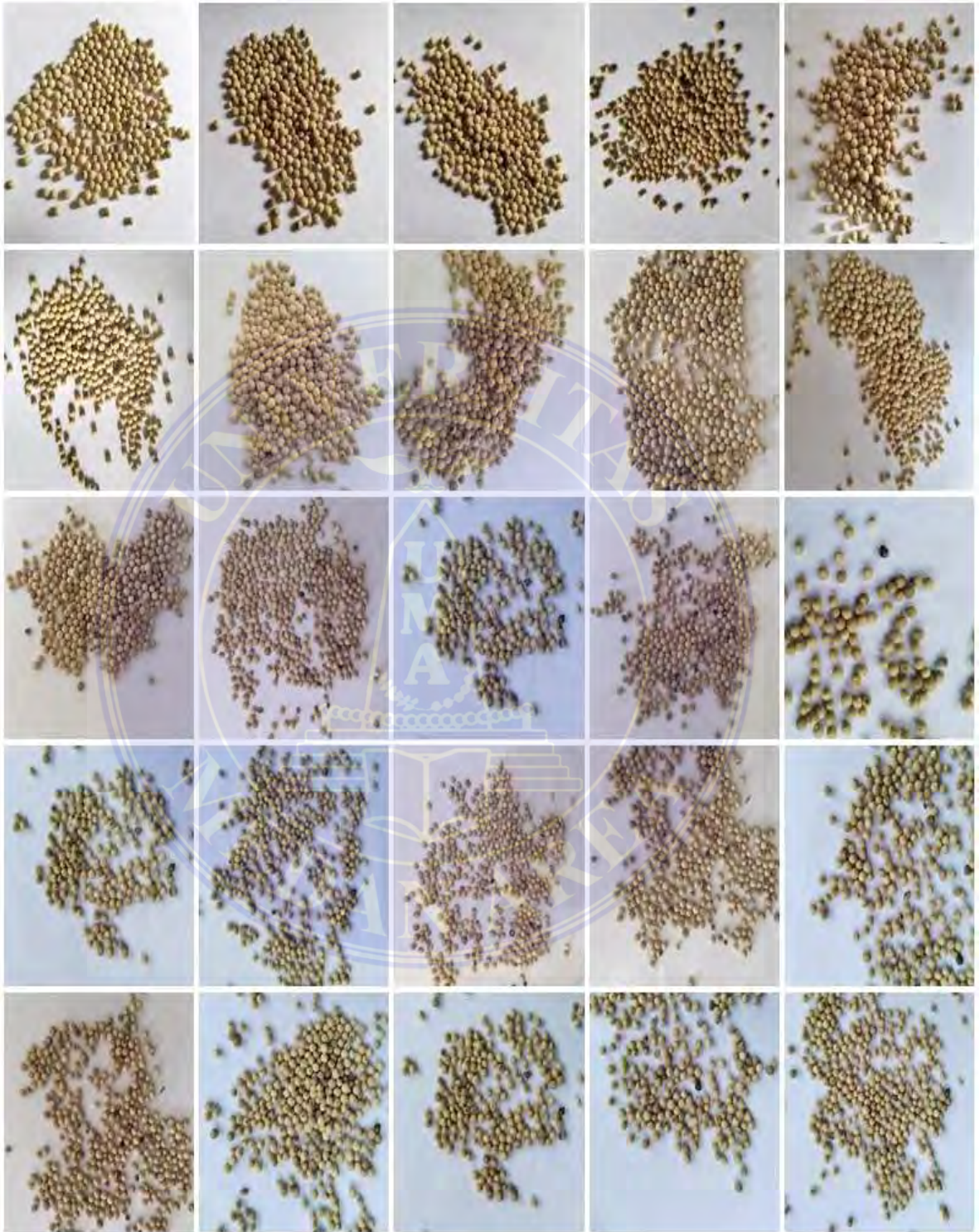
10. Kunyit



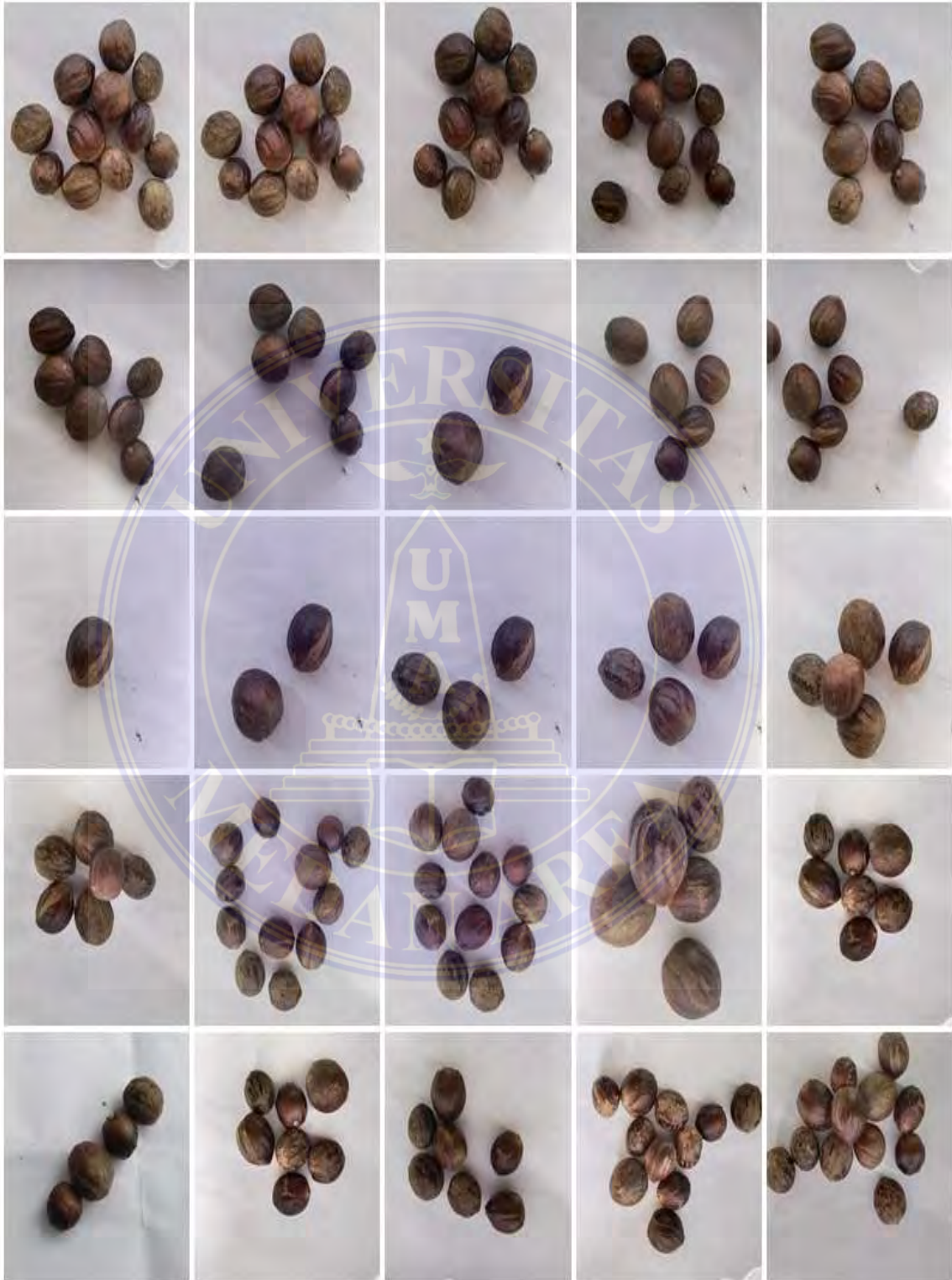
11. Lengkuas



12. Merica



13. Pala



14. Sere



b. Sourcecode

1. Import library

```
from skimage.io import imread
from skimage.transform import resize
from skimage.feature import hog
import matplotlib.pyplot as plt # mengimpor library pada python
import glob
import csv
```

2. Resize Image

```
for file in image_files:
    img = imread(file)
    resized_img = resize(img, (250, 280))
    plt.axis("off")
    plt.imshow(resized_img)
    plt.show()
    print(resized_img.shape)
```

3. Ekstraksi Fitur

```
with open('hog_features.csv', mode='w', newline='')
    writer = csv.writer(csv_file)
    for file in image_files:
        img = imread(file)
        resized_img = resize(img,(250,300))
        fd, hog_image = hog(resized_img, orientations=9, pixels_per_cell=(8, 8),
```

```
cells_per_block=(2, 2), visualize=True, multichannel=True)
writer.writerow(islice(fd, 20))

plt.axis("off")
plt.imshow(hog_image, cmap='BuGn')
plt.show()

print(fd.shape)
```

4. Pemanggilan Data

```
from google.colab import files

uploaded = files.upload()

df = pd.read_csv(io.BytesIO(uploaded['data10.csv']))
#print(df)
df.head
```

5. Pembagian Data Training dan Testing

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

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

from sklearn.neighbors import KNeighborsClassifier
```

```

k = KNeighborsClassifier(n_neighbors = 3, metric='manhattan') # euclidean -
chebyshev - minkowski - hamming - canberra - braycurtis - jaccard - manhattan

k.fit(X_train, y_train)

y_pred = k.predict(X_test)

confusion_matrix(y_test,y_pred)

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

```

6. Klasifikasi Metode K-NN

```

from sklearn.neighbors import KNeighborsClassifier

neighbors = [1,2,3,4,5] # menggunakan beberapa nilai k
minkowski_accuracy = np.empty(len(neighbors))
Jaccard_accuracy = np.empty(len(neighbors))
Manhattan_accuracy = np.empty(len(neighbors))
braycurtis_accuracy = np.empty(len(neighbors))
canberra_accuracy = np.empty(len(neighbors))
euclidean_accuracy = np.empty(len(neighbors))

for i,k in enumerate(neighbors):

    knn_minkowski = KNeighborsClassifier(n_neighbors=k, metric='minkowski')
    knn_Jaccard = KNeighborsClassifier(n_neighbors=k, metric='jaccard')

```

```

knn_Manhattan = KNeighborsClassifier(n_neighbors=k, metric='manhattan')
knn_braycurtis = KNeighborsClassifier(n_neighbors=k, metric='braycurtis')
knn_canberra = KNeighborsClassifier(n_neighbors=k, metric='canberra')
knn_euclidean = KNeighborsClassifier(n_neighbors=k, metric='euclidean')

knn_minkowski.fit(X_train, y_train)
knn_Jaccard.fit(X_train, y_train)
knn_Manhattan.fit(X_train, y_train)
knn_braycurtis.fit(X_train, y_train)
knn_canberra.fit(X_train, y_train)
knn_euclidean.fit(X_train, y_train)

minkowski_accuracy[i] = knn_minkowski.score(X_test, y_test)
Jaccard_accuracy[i] = knn_Jaccard.score(X_test, y_test)
Manhattan_accuracy[i] = knn_Manhattan.score(X_test, y_test)
braycurtis_accuracy[i] = knn_braycurtis.score(X_test, y_test)
canberra_accuracy[i] = knn_canberra.score(X_test, y_test)
euclidean_accuracy[i] = knn_euclidean.score(X_test, y_test)

import matplotlib.pyplot as plt

plt.title('K-NN Varying number of neighbors')
plt.plot(neighbors, minkowski_accuracy, label='K-NN - Minkowski')
plt.plot(neighbors, Jaccard_accuracy, label='K-NN - Jaccard')
plt.plot(neighbors, Manhattan_accuracy, label='K-NN - Manhattan')
plt.plot(neighbors, braycurtis_accuracy, label='K-NN - Braycurtis')
plt.plot(neighbors, canberra_accuracy, label='K-NN - Canberra')
plt.plot(neighbors, euclidean_accuracy, label='K-NN - Euclidean')

```


7. 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")

```

