



Klasifikasi Tinggi Awan dari Citra Satelit Himawari-8 Menggunakan Fuzzy C-Means, GLCM, dan KNN.

Togi Siholmarito Simarmata^{1*}, Elsa Lolita Anggraini², Abdul Manab³,
Muhammad Khalis Fikri⁴, Andicho Haryus Wirasapta⁵, Elsi Alfionita Syawal⁶

^{1*-6} Teknik Elektro, Universitas Jambi, Indonesia

Email: togisiholmaritosimarmata@unja.ac.id

*Penulis Korespondensi : togisiholmaritosimarmata@unja.ac.id

Abstract. This project aims to develop an automatic categorization method of cloud height (low, medium and high) based on Himawari-8 satellite imagery from BMKG Cengkareng Meteorological Station. The methodology includes four phases: pre-processing, Fuzzy C-Means (FCM) segmentation, GLCM texture feature extraction (9 parameters), and K-Nearest Neighbor (K-NN) classification. The four scenarios evaluated have the test data volumes of 20, 30, 40 and 50. Results show that the method is rather accurate, even perfect for some parameters, in the recognition of low clouds. However, the classification of medium and high clouds is still inconsistent, due to the overlapping of the texture feature values of the classes in some circumstances, which makes it difficult for the K-NN method to find out the proper decision boundaries. In order to compensate this classification bias and improve the overall accuracy, the article suggests improving the feature extraction step, utilizing more nonlinear classification algorithms and developing a more flexible integration between the segmentation and feature extraction stage

Keywords: Cloud Height; Himawari-8; Fuzzy C-Means; GLCM; K-Nearest Neighbor

Abstrak. Monitoring ketinggian awan merupakan parameter krusial dalam observasi meteorologi yang berperan strategis dalam analisis dinamika atmosfer serta prediksi cuaca ekstrem. Penelitian ini bertujuan mengembangkan sistem klasifikasi otomatis untuk mengkategorikan awan rendah, awan menengah, dan awan tinggi menggunakan citra satelit Himawari-8. Metodologi penelitian mencakup empat tahapan utama: pre-processing untuk standarisasi data, segmentasi menggunakan metode Fuzzy C-Means (FCM) untuk memisahkan objek awan, ekstraksi fitur tekstur berbasis Gray Level Co-occurrence Matrix (GLCM) dengan sembilan parameter, serta klasifikasi menggunakan algoritma K-Nearest Neighbor (K-NN). Data penelitian bersumber dari BMKG Stasiun Meteorologi Cengkareng, Tangerang. Pengujian sistem dilakukan melalui empat skenario eksperimen dengan volume data uji sebanyak 20, 30, 40, dan 50. Hasil penelitian menunjukkan bahwa sistem mampu mendeteksi kategori awan rendah dengan tingkat akurasi yang sangat tinggi, bahkan mencapai performa sempurna pada beberapa parameter efektivitas. Namun, sistem mengalami keterbatasan signifikan dalam mengklasifikasikan kategori awan menengah dan awan tinggi secara konsisten. Fenomena tumpang tindih (overlapping) pada nilai parameter fitur tekstur antar kelas menyebabkan algoritma K-NN sulit menentukan batas keputusan yang optimal. Ketidakseimbangan performa ini mengindikasikan bahwa meskipun kombinasi FCM dan GLCM cukup efektif untuk pola visual awan rendah, diperlukan optimasi lebih lanjut pada tahap ekstraksi fitur atau pemilihan metode klasifikasi yang lebih non-linear. Penelitian ini merekomendasikan adanya integrasi yang lebih adaptif antara tahap segmentasi dan ekstraksi ciri guna meminimalisir bias klasifikasi pada satu kelas tertentu dan meningkatkan akurasi sistem secara menyeluruh pada seluruh tingkatan awan.

Kata kunci: Tinggi Awan; Himawari-8; Fuzzy C-Means; GLCM; K-Nearest Neighbor

1. LATAR BELAKANG

Awan adalah salah satu parameter penting dalam meteorologi karena berperan dalam proses radiasi, sirkulasi atmosfer, dan pembentukan curah hujan serta pemanfaatan untuk peramalan cuaca. Pemanfaatan klasifikasi ketinggian awan bukan hanya untuk

pemodelan cuaca tetapi dapat juga untuk meningkatkan prediksi peramalan cuaca (Lin et al., 2023) (Huang et al., 2024). Setiap ketinggian awan memiliki karakteristik yang berbeda-beda. Hal ini dapat mempengaruhi interpretasi data satelit dan kualitas prakiraan cuaca secara keseluruhan (Luo et al., 2023). Dalam beberapa tahun terakhir, penggunaan citra satelit untuk kegunaan klasifikasi awan berdasarkan jenis maupun ketinggian sangat berkembang dengan pesat. Satelit Himawari menawarkan resolusi temporal dan spektral yang tinggi sehingga dapat dimanfaatkan untuk pemantauan fenomena meteorologi berkelanjutan. Pada penelitian terdahulu telah menunjukkan bahwa pemanfaatan citra satelit untuk klasifikasi awan dapat memberikan informasi penting tentang awan dan atmosfer serta hal ini mencakup dalam wilayah yang luas (Li et al., 2025). Beberapa penelitian sudah pernah dilakukan tentang klasifikasi awan berbasis citra satelit. Penelitian yang dilakukan oleh Wang dkk. memanfaatkan metode *random forest* untuk mengklasifikasikan awan dari data citra satelit Himawari-8 menghasilkan akurasi tinggi dalam membedakan kelas awan. Namun penggunaan metode ini masih bergantung pada ketersediaan data ground truth berkualitas tinggi yang terbatas (Wang et al., 2023). Penelitian lain dari Yashuai Fu dkk. mengembangkan sistem klasifikasi berbasis inframerah Himawari-8 dengan mengacu pada data CloudSat dan CALIOP. Penelitian ini meraih akurasi sebesar 86,22 % dalam mengenali jenis awan seperti cirrus hingga nimbostratus, fokus penelitian ini lebih ke jenis awan secara umum, bukan pada pengelompokan spesifik berdasarkan tingkatan ketinggian awan (Fu et al., 2023).

Sementara itu penelitian lainnya dilakukan Kunstantin kunger dkk. Penelitian ini membahas bagaimana performa dari beberapa satelit cuaca dalam melakukan klasifikasi ketinggian awan yaitu awan rendah, menengah, dan tinggi. Serta menjelaskan tentang kelemahan sensor pasif (imager satelit geostasioner) dalam mendeteksi batas ketinggian yang jelas. Akan tetapi dalam penelitian ini belum menggunakan ekstraksi ciri dalam penentuan ciri tekstur (Krüger et al., 2026). Selanjutnya ada penelitian dari Mehmet Guzel dkk. penelitian ini membahas tentang klasifikasi awan dalam 10 jenis awan dan 3 tingkatan ketinggian menggunakan deep learning. Arsitektur deep learning yang digunakan mencakup MobileNetV2, Inception V3, EfficientNetV2L, VGG-16, Xception, ConvNeXtSmall, dan ResNet-152 V2. Hasil dari penelitian ini mencapai tingkat akurasi tertinggi dengan model Xception sebesar 97,66% (Guzel et al., 2024). Pada penelitian ini dibutuhkan jumlah data yang besar agar hasil akurasi baik. Penelitian lainnya adalah

penelitian yang dilakukan oleh Marrisaeka Mawarni dkk, dalam penelitian ini membahas tentang sistem klasifikasi langit cerah dan berawan menggunakan gray level co-occurrence matrix dan k-nearest neighbor. Penelitian ini berfokus pada klasifikasi cerah dan berawan yang diekstraksi ciri menggunakan glcm dan klasifikasi dengan knn. Penelitian ini tidak mengklasifikasikan berdasar pada ketinggian awan tetapi menggunakan metode ekstraksi ciri dan klasifikasi yang sama dengan penelitian yang dilakukan (Mawarni et al., 2021).

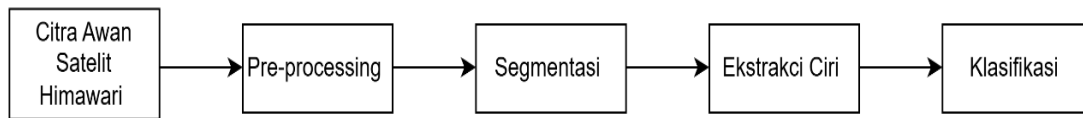
Ahendyarti dkk menggunakan satelit NOAA 18 dan 19 untuk klasifikasi awan serta bantuan perangkat RTL-SDR dan antena *Helix Quadrivial*. Pada penelitian ini membandingkan metode *Fuzzy C-Means* (FCM) dan pendekatan *multilevel thresholding* untuk segmentasi awan rendah, awan tengah, dan awan tinggi. Hasil dari segmentasi dilanjutkan ke tahap klasifikasi menggunakan algoritma *Learning Vector Quantization* (LVQ). Akurasi yang didapatkan sebesar 83,33% untuk FCM dan 72,22% untuk *multilevel thresholding* (Ahendyarti et al., 2020). Penelitian ini merupakan penelitian pembandingan dengan penelitian yang dilakukan. Terdapat beberapa persamaan yaitu meneliti tentang klasifikasi ketinggian awan dan menggunakan metode segmentasi yang sama yaitu FCM. Akan tetapi memiliki perbedaan pada data awan dan metode klasifikasi. Untuk data awan yang digunakan dalam penelitian ini adalah data awan satelit Himawari sedangkan untuk penelitian sebelumnya menggunakan data satelit NOAA. Sedangkan untuk metode yang digunakan dalam penelitian ini adalah KNN dan penelitian sebelumnya menggunakan metode LVQ.

Melihat dari penelitian yang telah dilakukan sebelumnya, ada celah riset yang dapat dilakukan yaitu penggunaan algoritma machine learning konvensional yang dipadukan dengan ekstraksi ciri serta penggunaan data awan dari satelit yang berbeda. Dalam pemrosesan citra digital diperlukan langkah lain sebelum klasifikasi, yaitu ekstraksi ciri. GLCM (*Gray Level Co-Occurance Matrix*) adalah metode ekstraksi ciri yang berfokus pada ciri tekstur. Ada beberapa ciri yang didapatkan seperti *contrast*, *homogeneity*, dan *correlation*. Penggunaan metode KNN dan GLCM telah banyak ditemui dalam berbagai penelitian untuk pengenalan tekstur objek (Mudaffarsyah et al., 2024). Penelitian ini mengkombinasikan 3 buah metode yaitu FCM, GLCM dan KNN dilakukan karena efisiensi komputasi yang tinggi dan diharapkan untuk menghasilkan model yang baik. Berbeda dengan metode deep learning yang membutuhkan sumber daya komputasi yang besar pendekatan dengan metode yang digunakan dalam penelitian ini lebih

sederhana. Penelitian ini akan membahas bagaimana mengklasifikasikan ketinggian awan menggunakan data dari satelit Himawari-8 dengan pendekatan GLCM untuk ekstraksi fitur tekstur dan KNN untuk proses klasifikasi. Diharapkan dengan metode ini dapat menghasilkan sistem yang akurat dan andal untuk membedakan awan rendah, menengah, dan tinggi berdasarkan tekstur.

2. METODE PENELITIAN

Metode penelitian ini dirancang menggambarkan setiap proses yang ada pada citra satelit himawari secara sistematis, jelas, dan terstruktur. Berikut pada Gambar 1 merupakan blok diagram dari tahapan awal hingga akhir dari penelitian ini.



Gambar 1. Blok Diagram

2.1 Pre-processing

Pre-processing merupakan tahap awal yang memiliki tujuan untuk menyamaratakan citra sebelum dilanjutkan ke tahap berikutnya (Hou et al., 2025). Pada tahap ini terdiri dari 3 langkah yaitu *resizing*, *grayscale*, dan *filtering*. *Resizing* merupakan tahap mengubah ukuran citra digital menjadi sama. Proses ini dimaksudkan untuk mempermudah pengolahan citra dalam jumlah yang besar tanpa mengurangi kualitas dari citra [14]. Ukuran citra yang diubah pada penelitian ini yaitu dari 1066 x 800 piksel menjadi 1000 x 750 piksel. Tahapan selanjutnya adalah konversi *grayscale*, tahap ini dilakukan dengan mengubah data citra ke dalam bentuk citra dengan instensitas skala keabuan (Jain et al., 2024). Proses ini dilakukan agar seluruh citra memiliki kesamaan dan dapat menghasilkan keluaran yang lebih baik untuk tahap selanjutnya. Tahap terakhir adalah *filtering* yang berfungsi untuk menghilangkan noise yang ada pada citra satelit (Pratidina et al., 2024).

2.2 Segmentasi

Tahapan segmentasi adalah tahap membagi objek menjadi beberapa bagian berdasarkan pada ciri atau karakteristik yang sama (Lukač et al., 2024). Teknik *threshoding* diterapkan untuk menghasilkan citra biner. Pada penelitian ini menggunakan metode *Fuzzy C-Means* (FCM) sebagai metode segmentasi berbasis pada *clustering* data atau *non-hierarchical clustering*, yang mampu mengelompokkan piksel ke dalam

beberapa kluster secara otomatis (Setiawan & Budiman, 2025). Algoritma FCM dalam penelitian ini bekerja secara otomatis berdasar pada fitur intensitas spektral, yaitu nilai intensitas abu-abu. Fungsi keanggotaan *fuzzy* yang bernilai kontinu dalam rentang $[0,1]$ merupakan parameter utama dalam menentukan penempatan piksel untuk kluster tertentu (Janusonis et al., 2024). Hal ini dapat memungkinkan pembedaan yang lebih akurat antara latar belakang (*background*) dengan kelas awan yaitu rendah, menengah, dan tinggi (Upadhyay & Gupta, n.d.). Adanya inisiasi kluster dapat merepresentasikan posisi rata-rata setiap kluster beserta piksel-piksel dengan nilai intensitas yang sama (Hallur & Gavade, 2025).

2.3 Ekstraksi Fitur

Ekstaksi fitur merupakan tahapan untuk mengambil karakteristik atau ciri dari suatu bentuk ataupun objek. Nilai yang diekstraksi kemudian akan dianalisis ke tahap selanjutnya yaitu klasifikasi (Setiawan & Budiman, 2025) (Silva & Rezendes, 2025). Pada penelitian ini menggunakan *Gray Level Co-Occurance Matrix* untuk ekstraksi fitur. GLCM merupakan salah satu metode ekstraksi ciri yang mengukur granulatitas, kekasaran, dan kontras antar piksel pada gambar berdasarkan pada tingkat keabuan atau yang dikenal dengan histogram (Yunianto et al., 2024). Ada 2 jenis ekstraksi ciri yaitu ekstraksi ciri orde satu yang menggunakan histogram citra, dan ekstraksi ciri orde dua menggunakan matriks ko-okurensi, yaitu matriks perantara penghubung antar piksel di berbagai dan jarak spasial (Domino et al., 2022). Penelitian ini menggunakan 9 parameter ciri tekstur diantaranya adalah *mean*, *variance*, *skewnes*, *kurtosis*, *entropy*, *energy*, *contrast*, *correlation*, *homogeneity* (Mawarni et al., 2021)(Wan & Du, 2020).

2.4 K-NN (K-Nearest Neighbor)

Algoritma KNN dikenal sebagai metode yang sederhana dan bergantung pada seberapa efektif KNN dalam menentukan tetangga terdekat. Metode ini merupakan bagian dari *supervised learning* atau yang dikenal sebagai metode pembelajaran terawasi (Aprilia Langnegara et al., 2026). Perhitungan jarak dengan metode KNN dilakukan dengan menghitung jarak dari satu data yang di uji dengan seluruh data latih menggunakan euclidean distance. Kelas target akan memiliki jarak rata-rata minimum dari data uji yang dipilih (Nurhidayat & Dewi, 2023).

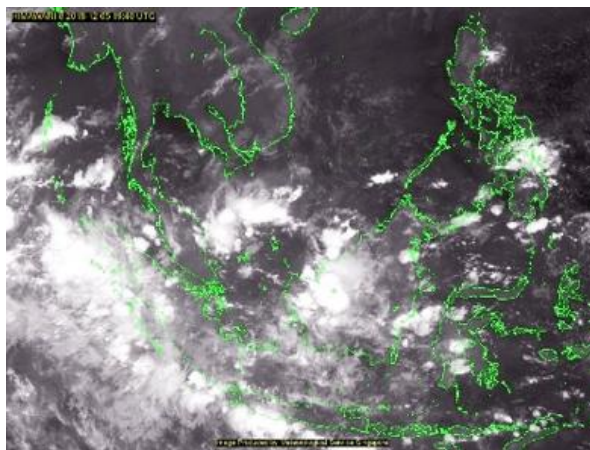
2.5 Confussion Matrix

Confussion matrix adalah salah satu model yang digunakan untuk mengetahui keakuratan sistem apakah data diklasifikasikan dengan benar atau salah. Penggunaan Confussion Matrix bertujuan dapat melihat secara detail kinerja suatu sistem klasifikasi dan mengidentifikasi di mana terjadi kesalahan klasifikasi. Confussion matrix merupakan sebuah teknik yang mudah dalam mengukur kinerja sebuah sistem klasifikasi (Nurhidayat & Dewi, 2023). Penelitian ini terdapat 3 nilai yang dibandingkan yaitu *sensitivity*, *specificity*, dan *accuracy* (Vanacore et al., 2024).

3. HASIL DAN PEMBAHASAN

3.1 Citra Awan Satelit Himawari

Data input penelitian ini terdiri dari 140 citra satelit Himawari 8 yang diperoleh dari kantor BMKG Cengkareng, Tangerang, Banten. Data citra dibagi menjadi 2 yaitu 90 data citra untuk pelatihan dan 50 citra untuk pengujian, dan dilakukan klasifikasi ketinggian awan dengan kategori awan rendah, awan menengah, dan awan tinggi. Berikut pada Gambar 2 merupakan salah satu contoh data citra satelit himawari.

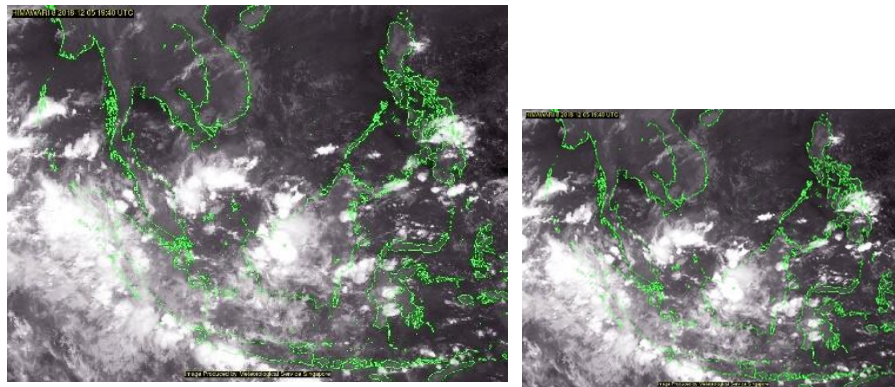


Gambar 2 Data citra satelit himawari.

3.2 Pre-Processing

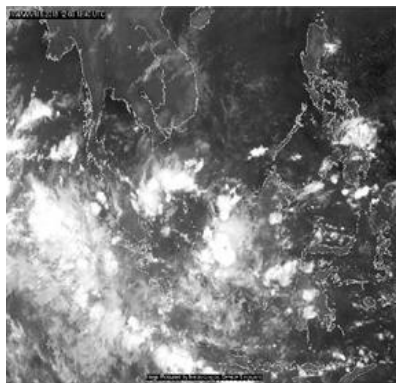
Proses pertama dalam pengolahan citra Himawari adalah tahap *pre-processing*, pada tahap ini memiliki tujuan untuk menyamaratakan atau membuat sama setiap citra yang menjadi *input* dalam penelitian ini. Tahap *preprocessing* terbagi menjadi 3 bagian yaitu *resize*, *grayscale*, dan *filtering*. Tahap pertama untuk *preprocessing* adalah *resize*. Citra awan yang menjadi masukan untuk proses ini terlalu besar ukuran pikselnya, untuk

itu perlu di *resize* untuk mengecilkan ukuran citra menjadi 1000 x 750 piksel dari ukuran citra masukan.



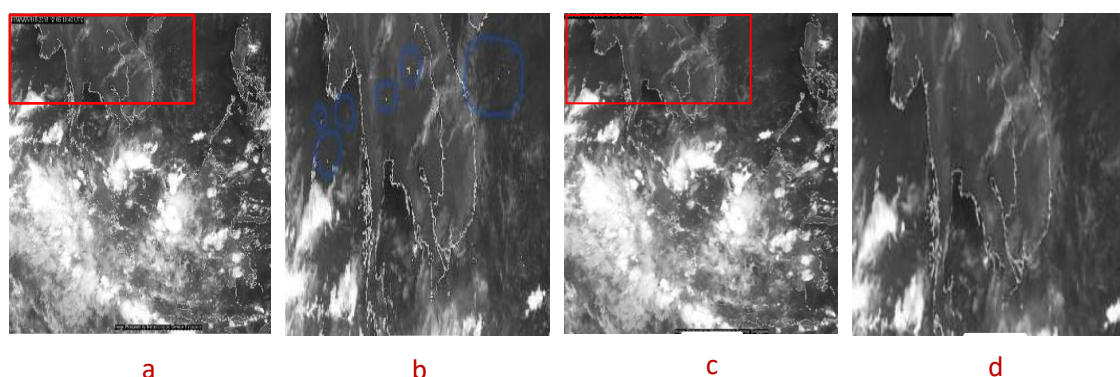
Gambar 3 *Preprocessing* (a) Ukuran Citra Asli, (b) Ukuran Citra *resize*.

Gambar 3 menunjukkan ukuran citra yang telah di *resize*. Proses ini bertujuan untuk menyamaratakan ukuran citra dan dapat mempersingkat waktu pemrosesan dan meminimalkan penggunaan memori. Tahap yang kedua adalah *grayscale*, pada tahap *grayscale* ini memiliki tujuan untuk menyeragamkan format citra masukan dan dapat mempermudah dalam pengujian tahap selanjutnya.



Gambar 4 Citra *resize* menjadi citra *grayscale*

Gambar 4 menunjukkan hasil *grayscale* yang sudah diproses setelah proses *resize*. *Filtering* merupakan proses untuk membuang atau membersihkan *noise-noise* yang ada dalam citra awan yang sudah di *grayscale*. *Noise-noise* yang dihilangkan fungsinya untuk mendapatkan informasi citra yang lebih baik.

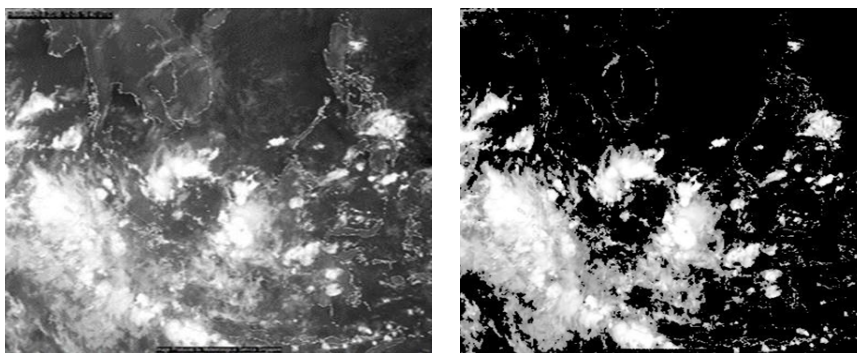


Gambar 5 Hasil Proses *Grayscale*; (a) Citra sebelum *noise* dihilangkan (b) Hasil perbesaran bagian *noise* (c) Hasil *filtering* (d) Perbesaran hasil sesudah *noise* dihilangkan.

Hasil dari proses menghilangkan *noise* citra awan dapat dilihat pada gambar 5. *Noise* yang ada pada gambar 5 dapat dihilangkan.

3.3 Segmentasi Citra

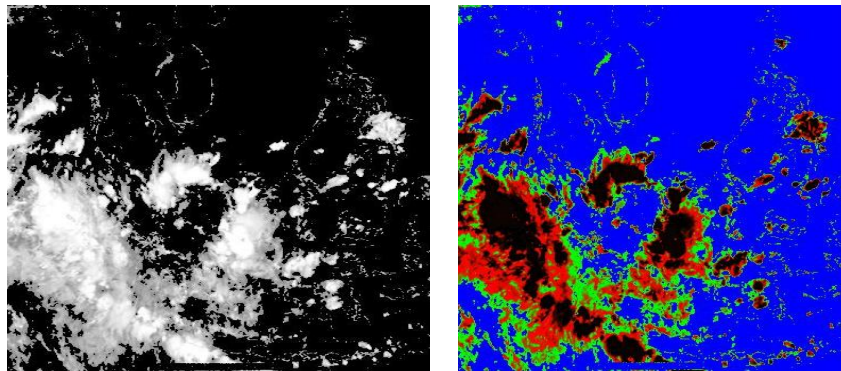
Segmentasi citra memiliki tujuan untuk memisahkan bagian awan dengan yang bukan awan. *Input* citra yang digunakan adalah citra hasil *filtering* dari hasil *preprocessing*. Operasi *thresholding* dilakukan pada tahap ini sehingga dapat diperoleh hasil citra *biner*, citra *biner* berbentuk hitam dan putih. Warna putih untuk piksel angka 1, kemudian warna hitam untuk piksel angka 0.



Gambar 6 Hasil Segmentasi dari citra *filtering* menjadi citra biner

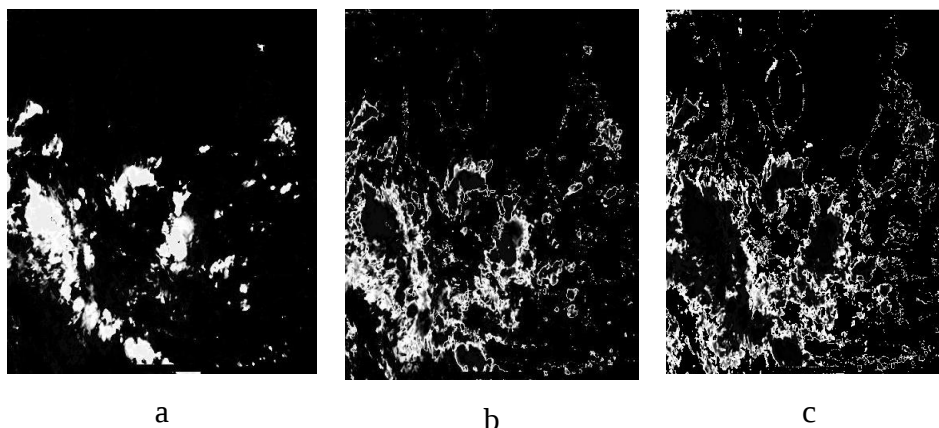
Proses *thresholding* merupakan proses yang didasarkan pada perbedaan derajat keabuan citra. Tujuan dari proses ini adalah untuk memisahkan antara objek dengan *background*. Hasil dari proses *thresholding* seperti pada Gambar 6 digunakan untuk proses *masking and cropping*, fungsi dari *masking and cropping* yaitu untuk memperoleh bagian awan yang diinginkan dan bagian yang tidak diinginkan dihilangkan. Hasil dari

tahap segmentasi citra yang sudah memisahkan bagian objek dan bukan objek, kemudian diproses menggunakan segmentasi awan. Proses segmentasi awan yang dilakukan dalam penelitian ini yaitu menggunakan metode *Fuzzy C-means*. Metode *Fuzzy C-means* dilakukan untuk menghasilkan bagian-bagian awan yang terdiri dari awan rendah, awan tengah, dan awan tinggi.



Gambar 7 Hasil Segmentasi Citra dari Citra Biner menjadi Citra FCM

Hasil dari pengolahan *fuzzy c-means* dapat dilihat pada gambar 7. Hasil yang diperoleh dari proses *fuzzy c-means* menghasilkan empat citra yang telah tersegmentasi dan dibedakan setiap bagian awannya dengan warna. Tiap warna mempresentasikan awan yang berbeda-beda. Awan rendah direpresentasikan dengan warna hitam, awan tengah dengan warna merah, awan tinggi dengan warna hijau, dan bagian yang bukan awan dengan warna biru. Pada gambar 8 memperlihatkan perbedaan *cluster* antar awan yang dihasilkan. Proses *Fuzzy C-means* menghasilkan citra *image* yang bercampur antar kelas awan karena saling berdekatan, ini disebabkan karena *Fuzzy C-means* menggunakan *fuzzy set theory* yang menetapkan anggotanya berdasarkan pada nilai pikselnya.



Gambar 8 Hasil Klaterisasi FCM (a) awan rendah, (b) awan tengah, (c) awan tinggi

3.4 Ekstraksi Ciri

Ekstraksi ciri memiliki tujuan mendapatkan karakteristik atau ciri dari suatu bentuk citra. Parameter yang dinilai pada proses ini ada *mean*, *variance*, *skewnes*, *kurtosis*, *entrophy*, *energy*, *contrast*, *correlation*, *homogeneity*. Nilai parameter tersebut dapat dilihat dalam nilai rata-rata Tabel 4.1 di bawah ini.

Tabel 1 Hasil Ekstraksi Fitur

Ekstraksi Ciri	Awan Rendah		Awan Tengah		Awan Tinggi	
	<i>Max</i>	<i>Min</i>	<i>Max</i>	<i>Min</i>	<i>Max</i>	<i>Min</i>
<i>Mean</i>	141,3	121,6	153,8	123,9	148,3	5,37
<i>Variance</i>	15625,5	15545,2	15137,5	14866,7	15395,9	15139,7
<i>Skewnes</i>	0,95	-0,21	0,64	-0,03	0,43	-0,107
<i>Kurtosis</i>	-1,93	-1,97	-1,83	-1,95	-1,94	-1,97
<i>Entrophy</i>	3,01	2,65	3,51	0,15	3,40	3,25
<i>Energy</i>	2,28	0,02	2,59	0,18	2,30	0,18
<i>Contrast</i>	100,9	680,1	233,1	138,7	441,07	263,61
<i>Correlation</i>	6,81	0,06	6,64	0,66	6,40	0,006
<i>Homogeneity</i>	5,67	0,05	5,42	0,47	5,14	0,46

Nilai yang ada pada Tabel 1 , merupakan nilai parameter ekstraksi ciri. Hasil dari nilai rata-rata ekstraksi ciri memiliki nilai yang saling tumpang tindih antar kelas awan. Salah satu nya adalah nilai dari *variance* dan *kurtosis* memiliki nilai yang berdekatan antar kelas awan. Hasil yang tumpang tindih juga merata pada semua jenis kelas awan. Hasil ini dapat mempengaruhi hasil klasifikasi, dan hasil dari ekstraksi ciri digunakan untuk parameter masukkan dalam proses klasifikasi yang mewakili setiap tipe awan.

3.5 KNN (K-Nearest Neighbor)

Proses klasifikasi menggunakan metode KNN (*K Nearest Neighbor*) yang dapat mengklasifikasikan tipe awan yang diuji terdiri dari awan rendah, awan tengah, dan awan tinggi. Pengujian berdasarkan pada hasil ekstraksi ciri yang telah dilakukan, yaitu nilai parameter yang didapat dari proses sebelumnya. Selain menentukan nilai parameter pada ekstraksi ciri, proses klasifikasi KNN menentukan nilai parameter *k*. Parameter *k* adalah

parameter menentukan nilai ketetanggaan terdekat dari objek yang diuji. Nilai parameter k diambil trial dan error.

a. Pengujian 20 Data

Pengujian dengan 20 data per kelas menghasilkan nilai parameter k optimal sebesar $k = 14$.

Tabel 2 Pengujian 20 Data

Kelas awan	Output pengujian 20 data				Efektivitas Sistem		
	Rendah	Tengah	Tinggi	Total	<i>Sensitivity</i> (%)	<i>Specificity</i> (%)	<i>Accuracy</i> (%)
Rendah	20	0	0	20	100	82,5	100
Tengah	5	13	2	20	65	100	65
Tinggi	2	0	18	20	90	95	90
Total	60						

Pada Tabel 2 terdapat 51 data yang diklasifikasikan dengan benar. 20 data terklasifikasi dengan benar sebagai awan rendah, 13 data terklasifikasi sebagai awan tengah, dan 18 data terklasifikasi sebagai awan tinggi. Presentase keberhasilan dari total jumlah data yang benar adalah 85%. Kinerja dari sistem pengklasifikasi citra awan rendah yang ada pada Tabel 2 di atas dapat diketahui menggunakan efisiensi sistem dengan menghitung nilai *sensitivity*, *specificity*, dan *accuracy*. Pengujian pertama dengan 20 data didapatkan bahwa nilai *sensitivity* pada kelas awan rendah sebesar 100 %, hal ini terjadi karena data awan rendah mampu mengklasifikasikan data sampel dengan benar sesuai kelasnya. Hasil pengujian pada Tabel 2 semua data awan rendah benar terklasifikasi sesuai data yang diujikan. Nilai *specificity* pada kelas awan rendah sebesar 82,5 %, hal ini terjadi karena citra awan rendah cukup mampu mengenali data sampel yang bukan kelasnya dengan baik. Tabel 4.2 menunjukkan bahwa ada 5 data kelas awan tengah dan 2 data kelas awan tinggi yang masuk dalam kelas awan rendah. Nilai *accuracy* sebesar 100 %, hal ini karena kelas awan rendah mampu mengidentifikasi data sampel sesuai dengan kelasnya saat pengujian. Klasifikasi citra awan tengah yang terdapat pada Tabel 2, nilai *sensitivity* pada kelas awan tengah sebesar 65 %, hal ini terjadi karena kelas awan tengah kurang mampu mengklasifikasikan dengan benar sesuai kelasnya. Tabel 2 di atas terdapat 13 data citra yang terklasifikasi dengan benar sesuai dengan kelasnya. Sedangkan

5 data terklasifikasi dalam bagian kelas awan rendah dan 2 data terklasifikasi dalam bagian kelas awan tinggi. Nilai *specificity* pada kelas awan tengah sebesar 100 %, hal ini terjadi karena citra awan tengah baik dalam mengenali sampel yang bukan kelasnya. Nilai *accuracy* sebesar 65 %, hal ini karena kelas awan tengah cukup mampu mengidentifikasi dengan benar data sesuai kelasnya. Dari 20 data citra yang diujikan, awan tengah hanya dapat mengklasifikasikan dan mengidentifikasi 13 data dengan benar.

Klasifikasi citra awan selanjutnya adalah awan tinggi sesuai pada Tabel 2, nilai *sensitivity* pada kelas awan tinggi sebesar 90 %, hal ini dikarenakan data citra awan tinggi cukup mampu dalam mengklasifikasikan data sesuai dengan kelasnya. Hasil pengujian kelas awan tinggi pada Tabel 2 dengan 20 data yang diujikan, 18 data terklasifikasi dengan benar sesuai dengan kelasnya, dan 2 data tidak terklasifikasi dengan benar dan masuk dalam bagian kelas awan rendah. Nilai *specificity* pada kelas awan tinggi sebesar 95 %, hal ini karena kelas awan rendah mampu mengenali sampel yang bukan kelasnya. Nilai *accuracy* pada kelas awan tinggi sebesar 90 %, hal ini karena kelas awan tinggi cukup mampu mengidentifikasi dengan benar data sampel sesuai kelasnya. Terdapat 18 data yang benar diklasifikasikan dan diidentifikasi dari 20 data citra yang diuji.

b. Pengujian 30 Data

Pengujian dengan 30 data per kelas juga menghasilkan nilai k optimal sebesar k = 14.

Tabel 3 Pengujian 30 data.

Output pengujian 30 data					Efektivitas Sistem		
Kelas awan	Rendah	Tengah	Tinggi	Total	<i>Sensitifity</i> (%)	<i>Spesificity</i> (%)	<i>Accuracy</i> (%)
Rendah	30	0	0	30	100	88,3	100
Tengah	5	22	3	30	73,3	100	73,3
Tinggi	2	0	28	30	93,3	95	93,3
Total	90						

Hasil pengujian 30 data seperti terlihat pada Tabel 3. Terdapat 80 data yang diklasifikasikan dengan benar. 30 data terklasifikasi dengan benar sebagai awan rendah, 22 data terklasifikasi sebagai awan tengah, dan 28 data terklasifikasi sebagai awan tinggi.

Presentase keberhasilan dari total jumlah data yang benar adalah 88,8%. Hasil efektivitas sistem Tabel 3 citra awan rendah dengan pengujian 30 data didapatkan nilai *sensitivity* pada kelas awan rendah sebesar 100 %, hal ini terjadi karena data awan rendah mampu mengklasifikasikan dengan benar sesuai kelasnya. Jika dilihat dari Tabel 3, dari 30 data yang diujikan semua data dapat diklasifikasikan dengan benar. Nilai *specificity* pada kelas awan rendah sebesar 88,3 %, hal ini terjadi karena citra awan rendah cukup mampu mengenali data sampel yang bukan kelasnya dengan baik. Hasil pengujian 5 data awan tengah dan 2 data awan tinggi terklasifikasi sebagai bagian dari awan rendah. Nilai *accuracy* sebesar 100 %, hal ini karena kelas awan rendah masih mampu mengidentifikasi dengan benar sesuai dengan kelasnya.

Efektivitas sistem pada Tabel 3 citra awan tengah dapat dihitung nilai *sensitivity* pada kelas awan tengah sebesar 73,3 %, hal ini terjadi karena kelas awan tengah kurang mampu mengklasifikasikan dengan benar sesuai kelasnya. Dari Tabel 3 di atas 30 data yang diujikan terdapat 22 data dapat diklasifikasikan dengan benar dan 5 data terklasifikasi sebagai awan rendah dan 3 data terklasifikasi sebagai bagian awan tinggi. Nilai *specificity* pada kelas awan tengah sebesar 100 %, hal ini terjadi karena citra awan tengah baik dalam mengenali sampel yang bukan kelasnya, terbukti bahwa data kelas awan rendah dan tinggi tidak ada yang masuk dalam bagian kelas awan tengah. Nilai *accuracy* sebesar 73,3 %, hal ini karena kelas awan tengah kurang mampu mengidentifikasi dengan benar data sampel sesuai kelasnya. Hasil pengujian yang dilakukan 22 data dari 30 data sampel yang diujikan dapat diklasifikasikan dan diidentifikasi dengan benar.

Hasil efektivitas sistem citra awan tinggi dengan nilai *sensitivity* kelas awan tinggi sebesar 93,3 %, karena data citra awan tinggi cukup mampu mengklasifikasikan data sesuai dengan kelasnya. Dari 30 data yang diujikan ada 28 data yang diklasifikasikan dengan benar dan 2 data tidak terklasifikasi dengan benar karena masuk dalam bagian awan rendah. Nilai *specificity* pada kelas awan tinggi sebesar 95 %, hal ini karena kelas awan tinggi baik dalam mengenali sampel yang bukan kelasnya. Terdapat 3 data awan tengah yang masuk dalam bagian kelas awan tinggi. Nilai *accuracy* pada kelas awan tinggi sebesar 93,3 %, hal ini karena kelas awan tinggi cukup mampu dalam mengidentifikasikan data sampel dengan benar.

c. Pengujian 40 Data

Pengujian dengan 40 data per kelas menghasilkan k optimal sebesar $k = 14$.

Tabel 4 Pengujian 40 data

Output pengujian 40 data					Efektivitas Sistem		
Kelas Awan	Rendah	Tengah	Tinggi	Total	<i>Sensitivity</i> (%)	<i>Spesificity</i> (%)	<i>Accuracy</i> (%)
Rendah	40	0	0	40	100	90	100
Tengah	5	31	4	40	77,5	100	77,5
Tinggi	3	0	37	40	92,5	95	92,5
Total	120						

Pada Tabel 4 menunjukkan hasil pengujian 40 data uji. Terdapat 108 data yang diklasifikasikan dengan benar. 40 data terklasifikasi dengan benar sebagai awan rendah, 31 data terklasifikasi sebagai awan tengah, dan 37 data terklasifikasi sebagai awan tinggi. Presentase keberhasilan dari total jumlah data yang benar adalah 90%. Klasifikasi citra awan rendah yang ada pada Tabel 4 dengan hasil efektivitas sistem pengujian 40 data pengujian didapatkan bahwa nilai *sensitivity* pada kelas awan rendah sebesar 100 %, hal ini terjadi karena data awan rendah mampu mengklasifikasikan dengan benar sesuai kelasnya. Tabel 4 pengujian 40 data di atas dapat dilihat bahwa tidak ada data yang terklasifikasi sebagai bagian dari kelas awan tengah dan kelas awan tinggi. Nilai *specificity* pada kelas awan rendah sebesar 90 %, hal ini terjadi karena citra awan rendah cukup mampu mengenali data sampel yang bukan kelasnya dengan baik. Nilai *accuracy* sebesar 100 %, hal ini karena kelas awan rendah mampu mengidentifikasi dengan benar sesuai dengan kelasnya.

Hasil analisis klasifikasi citra awan tengah, nilai *sensitivity* pada kelas awan tengah sebesar 77,5 %, hal ini terjadi karena kelas awan tengah kurang mampu mengklasifikasikan dengan benar sesuai kelasnya. Hasil pengolahan data pada Tabel 4 dengan 31 data dapat terklasifikasi dengan benar, 5 data masuk dalam klasifikasi awan rendah dan 4 data masuk dalam klasifikasi awan tinggi. Nilai *specificity* pada kelas awan tengah sebesar 100 %, hal ini terjadi karena citra awan tengah mampu dalam mengenali sampel yang bukan kelasnya dengan baik, terbukti bahwa tidak ada data kelas awan rendah dan tinggi masuk dalam bagian kelas awan tengah. Nilai *accuracy* sebesar 77,5

%, hal ini karena kelas awan tengah kurang mampu mengidentifikasi data dengan benar sesuai kelasnya.

Klasifikasi citra awan tinggi dari proses pengolahan data yang dilakukan, hasil *sensitivity* pada kelas awan tinggi sebesar 92,5 %, karena data citra awan tinggi cukup mampu dalam mengklasifikasikan data sesuai dengan kelasnya. Hasil pengujian pada Tabel 4 dapat dilihat bahwa 37 data dapat terklasifikasi dengan benar sebagai kelas awan tinggi sedangkan 3 data tidak terklasifikasi sesuai dengan kelas awan tinggi, tetapi terklasifikasi ke dalam kelas awan tengah. Nilai *specificity* pada kelas awan tinggi sebesar 95 %, hal ini karena kelas awan tinggi cukup baik dalam mengenali sampel yang bukan kelasnya. Nilai *accuracy* pada kelas awan tinggi sebesar 92,5 %, hal ini karena kelas awan tinggi cukup mampu dalam mengidentifikasi data sampel dengan benar.

d. Pengujian 50 Data

Pengujian dengan 50 data per kelas menghasilkan k optimal sebesar k = 14.

Tabel 5 Pengujian 50 data

Output pengujian 50 data					Efektivitas Sistem		
Kelas Awan	Rendah	Tengah	Tinggi	Total	<i>Sensitivity</i> (%)	<i>Specificity</i> (%)	<i>Accuracy</i> (%)
Rendah	50	0	0	50	100	86	100
Tengah	9	37	4	50	74	99	74
Tinggi	5	1	44	50	88	96	88
Total	150						

Hasil pada Tabel 5 merupakan hasil pengujian 50 data. Sejumlah 131 data yang terklasifikasi dengan benar. 50 data terkalsifikasi dengan benar sebagai kelas awan rendah, 37 data terklasifikasi dengan benar sebagai kelas awan tengah, dan 44 data terklasifikasi dengan benar sebagai kelas awan tinggi. Presentase keberhasilan dari total jumlah data yang benar adalah 87,3%. Nilai efektivitas sistem awan rendah yang dilakukan pada Tabel 5, didapatkan nilai *sensitivity* sebesar 100 %. Tabel 5 pengujian 50 data dihasilkan bahwa tidak ada data yang terklasifikasi sebagai bagian dari kelas awan tengah dan kelas awan tinggi. Nilai *specificity* sebesar 86 %, hal ini terjadi karena kelas awan rendah cukup mampu mengenali data yang bukan bagian dari kelasnya dengan baik.

Terbukti dari Tabel 5 ada 9 data kelas awan tengah dan 5 awan tinggi yang terklasifikasi ke dalam kelas awan rendah. Nilai *accuracy* sebesar 100 % karena mampu mengidentifikasi data sampel yang diujikan dengan baik. Hasil pengujian data awan tengah pada efektivitas sistem Tabel 5 menghasilkan *sensitivity* sebesar 74 %. Hal ini terjadi karena kelas awan rendah kurang mampu mengklasifikasikan data sampel yang diberikan dengan benar. Tabel 5 menunjukkan dari 50 data uji terdapat 37 data terklasifikasi dengan benar sesuai dengan kelasnya, 9 data terklasifikasi sebagai bagian dari kelas awan rendah, dan 4 data terklasifikasi sebagai bagian dari kelas awan tinggi. Nilai *specificity* sebesar 99 %, nilai tersebut didapatkan karena awan tengah mampu mengenali data sampel yang bukan kelasnya dengan baik. Terbukti ada 1 data awan tinggi yang terklasifikasi dalam bagian kelas awan tengah. Nilai *accuracy* sebesar 74 %, hal ini karena kelas awan tengah kurang mampu mengidentifikasi data sampel yang diujikan dengan benar sesuai kelasnya.

Klasifikasi citra awan tinggi dari proses pengolahan data yang dilakukan, nilai *sensitivity* pada kelas awan tinggi sebesar 88 %, karena data citra awan tinggi cukup mampu dalam mengklasifikasikan data sesuai dengan kelasnya. Hasil pengujian pada Tabel 5 bahwa 44 data dapat terklasifikasi dengan benar sebagai kelas awan tinggi sedangkan 1 data tidak terklasifikasi sesuai dengan kelas awan tinggi, namun terklasifikasi pada kelas lain yaitu masuk dalam kelas awan tengah dan 5 data terklasifikasi dalam kelas awan rendah. Nilai *specificity* pada kelas awan tinggi sebesar 96 %, hal ini karena kelas awan tinggi cukup baik dalam mengenali sampel yang bukan kelasnya. Dari Tabel 5 didapatkan 4 awan tengah terklasifikasi dalam kelas awan tinggi. Nilai *accuracy* pada kelas awan tinggi sebesar 88 %, hal ini karena kelas awan tinggi cukup mampu dalam mengidentifikasi data sampel dengan benar.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian klasifikasi ketinggian awan menggunakan citra satelit Himawari-8 dan algoritma K-NN, tahap pre-processing berhasil diselesaikan dengan baik melalui tiga langkah: Resize, konversi grayscale, dan filtering. Kualitas citra meningkat dibandingkan citra asli setelah melalui tahap ini. Teknik segmentasi citra berhasil memisahkan *foreground* (objek awan) dari *background*, yang kemudian memungkinkan ekstraksi fitur menggunakan GLCM dan pengelompokan awan ke dalam

tiga kategori menggunakan metode FCM. Proses klasifikasi KNN berhasil diselesaikan dengan nilai parameter k optimal sebesar $k = 14$. Eksperimen dilakukan empat kali menggunakan dataset uji sebanyak 20, 30, 40, dan 50 data. Setiap percobaan menambahkan 10 data baru dari percobaan sebelumnya untuk dievaluasi ulang. Secara keseluruhan dari 3 ketinggian awan yang dilakukan, awan rendah dapat diklasifikasikan dengan baik dibanding dengan awan tengah dan awan tinggi. Dari 4 percobaan hasil akurasi untuk awan rendah adalah 100%. Sedangkan untuk awan tengah akurasi yang dihasilkan 72,45% dan awan tinggi 90,95%.

DAFTAR REFERENSI

- Ahendyarti, C., Wiryadinata, R., Rohana, N., & Muhammad, F. (2020). Cloud Classification from NOAA Satellite Image Using Learning Vector Quantization Method. *Proceeding - 2020 2nd International Conference on Industrial Electrical and Electronics, ICIEE 2020*, 97–100. <https://doi.org/10.1109/ICIEE49813.2020.9277269>
- Aprilia Langnegara, D., Misriati, T., & Nawawi, I. (2026). Optimalisasi Rasio Data pada K-Nearest Neighbor untuk Klasifikasi Multikelas Tingkat Obesitas Populasi Dewasa. *Jurnal Sains Dan Teknologi Informasi*, 5(2), 83–90. <https://doi.org/10.47065/jussi.v5i2.9853>
- Domino, M., Borowska, M., Kozłowska, N., Trojakowska, A., Zdrojkowski, Ł., Jasiński, T., Smyth, G., & Maśko, M. (2022). Selection of Image Texture Analysis and Color Model in the Advanced Image Processing of Thermal Images of Horses following Exercise. *Animals*, 12(4). <https://doi.org/10.3390/ani12040444>
- Fu, Y., Mi, X., Han, Z., Zhang, W., Liu, Q., Gu, X., & Yu, T. (2023). A Machine-Learning-Based Study on All-Day Cloud Classification Using Himawari-8 Infrared Data. *Remote Sensing*, 15(24). <https://doi.org/10.3390/rs15245630>
- Guzel, M., Kalkan, M., Bostanci, E., Acici, K., & Asuroglu, T. (2024). Cloud type classification using deep learning with cloud images. *PeerJ Computer Science*, 10. <https://doi.org/10.7717/peerj-cs.1779>
- Hallur, S., & Gavade, A. (2025). Image feature extraction techniques: A comprehensive review. In *Franklin Open* (Vol. 12). Elsevier B.V. <https://doi.org/10.1016/j.fraope.2025.100366>
- Hou, Y., Guo, K., & Bi, X. (2025). Remote sensing image fusion based on real time image smoothing and image similarity. *Systems and Soft Computing*, 7. <https://doi.org/10.1016/j.sasc.2025.200226>

- Huang, Y., Kim, D., Zhou, T., & Shi, X. (2024). The Role of Cloud-Radiative Interaction in Tropical Circulation and the Madden–Julian Oscillation. *Journal of Climate*, 37(17), 4559–4576. <https://doi.org/10.1175/JCLI-D-23-0736.1>
- Jain, T., Mhaske, H., Chilveri, S., Chaudhar, A., & Doshi, C. (2024). Cloudy Weather Prediction Using CNN Models and Satellite Images. *International Journal for Research in Applied Science and Engineering Technology*, 12(4), 1168–1175. <https://doi.org/10.22214/ijraset.2024.59342>
- Janusonis, E., Kazakeviciute-Januskeviciene, G., & Bausys, R. (2024). Selection of Optimal Segmentation Algorithm for Satellite Images by Intuitionistic Fuzzy PROMETHEE Method. *Applied Sciences (Switzerland)*, 14(2). <https://doi.org/10.3390/app14020644>
- Krüger, K., Wirth, M., Floutsi, A. A., Donovan, D. P., van Zadelhoff, G.-J., & Groß, S. M. (2026). Validation of cloud macrophysical properties from the ATLID Level 2a products A-TC, A-FM, A-CTH using airborne lidar observations during the HALO missions PERCUSION, ASCCI and NAWDIC. <https://doi.org/10.5194/egusphere-2026-3051>
- Li, Y., Wu, Y., Li, J., Sun, A., Zhang, N., & Liang, Y. (2025). A Machine Learning Algorithm Using Texture Features for Nighttime Cloud Detection from FY-3D MERSI L1 Imagery. *Remote Sensing*, 17(6). <https://doi.org/10.3390/rs17061083>
- Lin, J., Bao, Y., Petropoulos, G. P., Mehraban, A., Pang, F., & Liu, W. (2023). Cloud-Type Classification for Southeast China Based on Geostationary Orbit EO Datasets and the LighGBM Model. *Remote Sensing*, 15(24). <https://doi.org/10.3390/rs15245660>
- Lukač, L., Nerat, A., Strnad, D., Horvat, Š., & Žalik, B. (2024). Region Segmentation of Images Based on a Raster-Scan Paradigm. *Journal of Sensor and Actuator Networks*, 13(6). <https://doi.org/10.3390/jsan13060080>
- Luo, H., Quaas, J., & Han, Y. (2023). Examining cloud vertical structure and radiative effects from satellite retrievals and evaluation of CMIP6 scenarios. *Atmospheric Chemistry and Physics*, 23(14), 8169–8186. <https://doi.org/10.5194/acp-23-8169-2023>
- Mudaffarsyah, M., Azza Al Kausar, M., Luter Sihombing, O., Putra Ash Sidiq, H., Putri Fercia, K., Hayaty, N., Teknik Informatika, J., Maritim Raja Ali Haji, U., & Politeknik Senggarang, J. (2024). Klasifikasi Jenis Lamun Menggunakan Ekstraksi Fitur GLCM dan Algoritma K-Nearest Neighbor (KNN). 13(02), 17–22.
- Nurhidayat, R., & Dewi, K. E. (2023). KOMPUTA : Jurnal Ilmiah Komputer dan Informatika Penerapan Algoritma K-Nearest Neighbor Dan Fitur Ekstraksi N-Gram

Dalam Analisis Sentimen Berbasis Aspek. 12(1).
<https://www.kaggle.com/datasets/hafidahmusthaanah/skincare-review?select=00.+Review.csv>.

Pratidina, C., Safira, D., Gelar, T., Permana, H., Suprihanto, S., Syakrani, N., & Fauzi, C. (2024). Morphological Grayscale Pre-processing to SAR Images for Reducing Noise in Ship Detection Based on YOLOv8. *Journal of Software Engineering, Information and Communication Technology (SEICT)*, 5(2), 95–106. <https://doi.org/10.17509/seict.v5i2.75970>

Setiawan, A., & Budiman, A. A. (2025). Optimization of Gray Level Co-occurrence Matrix (GLCM) Texture Feature Parameters in Determining Rice Seed Quality. *EMITTER International Journal of Engineering Technology*, 13(1), 110–123. <https://doi.org/10.24003/emitter.v13i1.928>

Silva, H., & Rezendes, C. (2025). Texture Analysis of Near-Infrared Vein Images During Reactive Hyperemia in Healthy Subjects. *Applied Sciences (Switzerland)*, 15(10). <https://doi.org/10.3390/app15105702>

Upadhyay, U., & Gupta, S. (n.d.). A Survey on Image Feature Extraction Techniques. In *International Journal of Scientific Research & Engineering Trends* (Vol. 10, Number 3).

Vanacore, A., Pellegrino, M. S., & Ciardiello, A. (2024). Fair evaluation of classifier predictive performance based on binary confusion matrix. *Computational Statistics*, 39(1), 363–383. <https://doi.org/10.1007/s00180-022-01301-9>

Wan, X., & Du, J. (2020). Cloud classification for ground-based sky image using random forest. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43(B3), 835–842. <https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-835-2020>

Wang, Y., Hu, C., Ding, Z., Wang, Z., & Tang, X. (2023). All-Day Cloud Classification via a Random Forest Algorithm Based on Satellite Data from CloudSat and Himawari-8. *Atmosphere*, 14(9). <https://doi.org/10.3390/atmos14091410>

Yunianto, M., Meilina, R. D., & Suryani, E. (2024). Using Decision Tree With First and Second-Order Statistical Feature Extraction for Classification of Lung Cancer. *INDONESIAN JOURNAL OF APPLIED PHYSICS*, 14(2), 339. <https://doi.org/10.13057/ijap.v14i2.87676>