

SEGMENTASI DENGAN METODE ACTIVE CONTOUR UNTUK PENINGKATAN AKURASI KLASIFIKASI CITRA USG KANKER PAYUDARA MENGGUNAKAN K-NEAREST NEIGHBOR (KNN)

Muhammad Ihza Alhafiz^{*1)}, Wirasno²⁾, Achmad Solichin³⁾

1. Magister Ilmu Komputer, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia
2. Magister Ilmu Komputer, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia
3. Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Article Info

Kata Kunci: Dataset; GLCM; Kanker Payudara; Klasifikasi; Segmentasi

Keywords: *Breast Cancer; Classification; Datasets; GLCM; Segmentation*

Article history:

Received 17 November 2024

Revised 15 Desember 2024

Accepted 1 Maret 2025

Available online 1 Februari 2025

DOI :

<https://doi.org/10.29100/jupi.v10i1.5681>

* Corresponding author.

Corresponding Author

E-mail address:

Zaafiz9@gmail.com

ABSTRAK

Kanker payudara, sebuah penyakit yang mengancam, ditandai oleh pertumbuhan sel yang tak terkendali yang dapat menyerang jaringan sehat dalam payudara. Segmentasi citra digital menjadi penting untuk memisahkan gambaran kanker dari latar belakangnya, memungkinkan analisis yang lebih mendalam. Dalam penelitian ini, metode *Grayscale*, *Median Filtering*, dan Aktif Kontur diadopsi untuk mengidentifikasi morfologi kanker dalam citra, sementara proses klasifikasi menggunakan metode *K-Nearest Neighbor* dengan ekstraksi fitur melalui *Gray Level Co-Occurrence Matrix* untuk menganalisis tekstur kanker. Temuan dari penelitian ini menegaskan bahwa proses aktif kontur digunakan untuk melihat morfologi kanker, diikuti dengan pemotongan citra asli sesuai dengan kontur yang dihasilkan. Selain itu, hasil pengujian *K-Nearest Neighbor* pada $K=18$ dengan pembagian data 90% untuk Data Training dan 10% untuk Data Uji menggunakan *K-Fold Cross Validation*, menunjukkan akurasi algoritma sebesar 68% yang diperoleh melalui *Confusion Matrix*.

ABSTRACT

Breast cancer, a threatening disease, is characterized by the uncontrollable growth of cells that can invade healthy breast tissue. Digital image segmentation is crucial to separate the depiction of cancer from its background, enabling deeper analysis. In this study, methods such as Grayscale, Median Filtering, and Active Contour were adopted to identify the morphology of cancer in images, while the classification process employed the K-Nearest Neighbor method with feature extraction through the Gray Level Co-Occurrence Matrix to analyze cancer texture. Findings from this research confirm that the active contour process is utilized to visualize cancer morphology, followed by cropping the original image according to the generated contour. Additionally, the testing results of K-Nearest Neighbor with $K=18$, dividing data into 90% for Training Data and 10% for Test Data using K-Fold Cross Validation, indicate an algorithm accuracy of 68% obtained through Confusion Matrix.

I. PENDAHULUAN

KANKER payudara merupakan jenis kanker yang timbul di dalam jaringan payudara. Kondisi ini terjadi ketika sel-sel di dalam jaringan payudara tumbuh secara tidak terkendali dan menggantikan jaringan payudara yang sehat serta sekitarnya. Kanker payudara bisa muncul di dua tempat utama, yaitu di kelenjar yang menghasilkan susu (lobulus) atau di saluran yang membawa air susu dari kelenjar ke puting payudara (duktus). Selain itu,

kanker payudara juga bisa terbentuk di jaringan lemak atau jaringan ikat dalam payudara. Meskipun lebih umum terjadi pada wanita, pria juga berisiko mengalami kanker payudara[1][2].

Berdasarkan data Global Burden Cancer, Amerika Serikat mencatat 231.840 kasus baru kanker payudara pada tahun 2015, dengan perkiraan 40.290 wanita meninggal dunia. Angka kasus baru meningkat menjadi 246.660 pada tahun 2016, dengan jumlah kematian sekitar 40.450. Kanker payudara juga menempati urutan pertama penyakit pada wanita di Asia, dengan estimasi insidensi sebesar 650.983 kasus pada tahun 2012 [3].

Menurut data Globocan, pada tahun 2020 jumlah kasus kanker payudara baru di Indonesia mencapai 68.858 (16,6%), terhitung 396.914 dari total kasus kanker baru di Indonesia. Sementara itu, jumlah kematian telah melampaui 22.000. Selain angka kematian yang relatif tinggi, pengobatan terminal pasien kanker mengakibatkan beban ekonomi yang semakin meningkat, dengan pengeluaran BPJS mencapai Rp 7,6 triliun selama 2019-2020. Sekitar 43% kematian akibat kanker dapat dicegah jika pasien secara teratur melakukan deteksi dini dan menghindari faktor risiko yang menyebabkan kanker[4].

Strategi Nasional Pencegahan Kanker Payudara Indonesia memuat tiga pilar yaitu promosi kesehatan, deteksi dini, dan tatalaksana kasus, dengan target deteksi dini pada 80% wanita berusia 30-50 tahun. Untuk mencapai target ini, Kementerian Kesehatan tidak bekerja sendiri, melainkan turut dibantu oleh berbagai pihak seperti Yayasan Kanker Payudara Indonesia (YKPI), yang berhasil menjangkau lebih dari 150.000 peserta baik secara daring dan luring pada 2016-2021[4].

Program skrining kanker payudara bisa dilakukan oleh masyarakat dengan periksa payudara sendiri (SADARI) ataupun periksa payudara secara klinis (SADARNIS) yang dilakukan oleh tenaga kesehatan sebagai ujung tombak pelayanan kesehatan untuk masyarakat. Deteksi dini dimulai dengan meningkatkan kesadaran masyarakat tentang perubahan bentuk atau adanya kelainan pada payudara. Para wanita, baik yang sudah dewasa maupun yang masih remaja, masih kurang memahami dan merasa tidak nyaman melakukan SADARI (Pemeriksaan Payudara Sendiri), sehingga hanya sedikit yang melakukannya. Teknik SADARI sebenarnya mudah dilakukan namun banyak perempuan belum melakukannya karena kurang informasi dan motivasi untuk mendapatkan informasi mengenai pencegahan dan deteksi dini kanker payudara[5].

Data yang dikumpulkan adalah gambar USG payudara pada wanita usia antara 25 sampai 75 tahun, dikumpulkan dari Baheya Hospital For Early Detection and Treatment of Women's Cancer di Cairo, Mesir pada tahun 2018, terdiri dari 788 gambar dengan rata-rata ukuran 500×500 piksel dan dikategorikan menjadi tiga kelas[4].

Model CV merupakan varian *active-contour* fundamental yang menginspirasi banyak turunan lain untuk kepentingan deteksi dan segmentasi objek. Model ini bekerja secara global, dengan prinsip seperti *clustering* yang menginvestigasi keseluruhan objek disuatu bidang citra. Model-model turunan CV unggul karena tidak sensitif terhadap inisialisasi dan tahan terhadap gradien tepi objek yang samar[6].

Salah satu dari penelitian ini menggunakan metode KNN yang dapat digunakan untuk mengklasifikasi penyakit kanker payudara. Metode KNN memiliki beberapa keunggulan, yaitu pelatihan secara sederhana, cepat, mudah dimengerti, dan efektif apabila ukuran data pelatihan besar. Namun, KNN juga memiliki kekurangan yaitu nilai K yang berbeda. Hasil klasifikasi dapat menghasilkan suatu prediksi penentuan jenis sel kanker payudara bersifat ganas atau jinak atau normal[7].

Peneliti menggunakan metode *grayscale preprocessing* dan *median filtering* untuk pengolahan citra USG untuk memperbaiki citra USG kanker payudara. Proses segmentasi menggunakan metode active contour untuk meningkatkan akurasi pada citra, berdasarkan Grey Level Co-Occurance Matrix (GLCM) dengan kedekatan 1 piksel dan sudut 0° , 45° , 90° , 135° . Setelah melakukan ekstrak fitur tekstur dari data citra kanker yang diketahui, langkah selanjutnya adalah klasifikasi. Mencakup definisi gambar uji untuk grup atau level yang ditentukan sebelumnya. Gambar ultrasound kanker payudara dicat dan fitur tekstur diekstraksi dari data gambar masing-masing penulis menggunakan metode KNearest Neighbor.

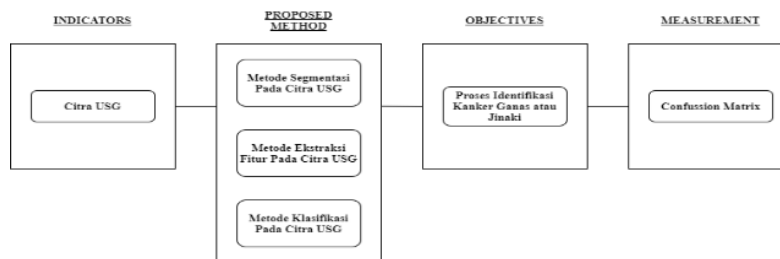
Tujuan dari penelitian ini adalah untuk mengetahui bagaimana identifikasi bentuk kanker payudara menggunakan metode segmentasi active contour, serta melakukan proses klasifikasi menggunakan metode K-Nearest Neighbors (KNN). Penelitian ini bertujuan untuk meningkatkan akurasi deteksi dan klasifikasi kanker payudara, sehingga dapat membantu dalam diagnosis dini dan penentuan langkah-langkah penanganan yang lebih tepat.

II. METODOLOGI PENELITIAN

Metode Penelitian ini bertujuan untuk menjelaskan langkah-langkah dan pendekatan yang digunakan dalam penelitian ini untuk menganalisis citra ultrasonografi (USG) payudara guna deteksi dan klasifikasi kanker. Diharapkan penjelasan yang diberikan tentang kerangka pemikiran penelitian, proses pengumpulan data, serta penerapan teori dan metode analisis citra digital yang telah dipilih akan memberikan gambaran komprehensif tentang pendekatan penelitian yang diadopsi untuk mencapai tujuan penelitian yang telah ditetapkan.

A. Kerangka Pemikiran

Kerangka pemikiran adalah suatu struktur konseptual yang memberikan pandangan yang terperinci dan sistematis mengenai proses analisis data, mulai dari langkah-langkah awal seperti pengumpulan dan pengolahan data hingga tahap akhir yaitu pencapaian kesimpulan dari data yang telah dianalisis. Secara visual tahapan-tahapan yang tercakup dalam kerangka pemikiran dalam penelitian ini, yang membantu dalam memandu jalannya proses analisis dengan lebih terarah dan efektif.



Gambar 1. Kerangka Pemikiran

B. Pengumpulan Data

Data untuk penelitian dikumpulkan melalui dua metode utama. Pertama, metode observasi digunakan dengan mengambil data dari database publik yang berasal dari sumber di <https://scholar.cu.edu.eg/?q=afahmy/pages/dataset>, yang dikenal sebagai "Dataset of breast ultrasound image" [8]. Data ini diperoleh dari dataset gambar USG kanker payudara yang tersedia di database tersebut. Kedua, studi literatur dilakukan dengan mencari dan mengumpulkan informasi dari berbagai sumber seperti buku, jurnal, dan modul yang relevan dengan objek penelitian. Proses studi literatur ini bertujuan untuk mengumpulkan informasi dan penelitian terdahulu yang terkait dengan topik yang sedang diteliti.

Data yang dikumpulkan adalah gambar USG payudara pada wanita usia antara 25 sampai 75 tahun, dikumpulkan dari Baheya Hospital For Early Detection and Treatment of Women's Cancer di Cairo, Mesir pada tahun 2018, terdiri dari 788 gambar dengan rata-rata ukuran 500×500 piksel dan dikategorikan menjadi tiga kelas. Data kanker jinak (benign) sebanyak 445 data, data kanker ganas (malignant) sebanyak 210 data, dan data kanker normal sebanyak 133 data.

C. Data Mining

Data mining merupakan suatu kegiatan untuk mengidentifikasi berbagai jenis data. Mulai dari banyaknya data yang tersimpan di *database*, data-data tersebut diidentifikasi dari kemungkinan adanya pola ataupun lainnya yang dianggap berpotensi untuk menghasilkan sesuatu yang dipakai oleh perusahaan atau jenis organisasi pemilik *database* tersebut [9].

D. Preprocessing

Data *preprocessing* merupakan teknik untuk menggali informasi dari data sebelum pembuatan model. Adapun tahapan-tahapan yang dilakukan pada saat preprocessing adalah :

1. Data Cleaning : Merupakan proses menemukan, mengoreksi, dan menghapus kesalahan, ketidakakuratan atau anomali data. Tujuan dari data cleaning adalah memastikan bahwa data yang digunakan untuk analisis atau pemrosesan lebih lanjut adalah akurat, lengkap, konsisten, dan relevan.
2. Features Selection : Proses modifikasi fitur sebelum pembuatan model. Teknik yang digunakan pada penelitian ini menggunakan korelasi spearman.
3. Standartisasi Data : Proses mentransformasi nilai-nilai dalam dataset sehingga memiliki rata-rata 0 deviasi standar 1. Tujuannya adalah membuat distribusi data lebih mudah diinterpretasikan dan diproses oleh algoritma pembelajaran mesin [10].

E. Seleksi Data

Seleksi data adalah proses pembersihan data dari atribut-atribut yang tidak diperlukan dan mengeliminasi data jika memiliki *missing value/duplicate value* yang tidak diperlukan dalam pengolahan data mining [11]. Seleksi data menggunakan K-Cross Validation yang berfungsi untuk menilai kinerja proses sebuah algoritma dengan membagi sampel data secara acak dan mengelompokkan data tersebut sebanyak nilai K-Fold. Kemudian salah satu kelompok K-Fold tersebut akan dijadikan sebagai data uji sedangkan sisa kelompok yang lainnya akan dijadikan sebagai data latih.

Dalam penyusunan penelitian ini, diperlukan sejumlah data citra kanker payudara yang bersifat jinak maupun ganas, yang diperoleh dari sumber database publik. Penulis mengambil citra dari data training hasil USG dengan ukuran rata-rata citra adalah 500x500 pixel dan format penyimpanannya adalah PNG. Untuk membagi data, penulis menggunakan metode K-Fold Cross Validation dengan rasio pembagian 90% untuk data training dan 10% untuk data uji [7].

F. Grayscale

Citra grayscale merupakan citra yang menampilkan hanya satu warna abu-abu, karena membutuhkan informasi yang lebih sedikit per pikselnya dibandingkan dengan citra berwarna. Warna abu-abu pada citra grayscale memiliki intensitas merah (R), 14 hijau (G), dan biru (B) yang sama. Oleh karena itu, citra grayscale hanya memerlukan satu nilai intensitas per piksel, berbeda dengan citra berwarna yang memerlukan tiga nilai intensitas per piksel [12]. Proses grayscale dilakukan dikarenakan citra USG tersebut masih memiliki citra RGB meskipun data terlihat seperti warna hitam putih, tetapi didalamnya masih ada nilai citra RGB, maka dilakukanlah konversi kedalam proses grayscale dan menjadi citra hitam putih.

G. Median Filtering

Filter median sebagai sebuah jendela yang berisi sejumlah piksel ganjil. Jendela ini digeser titik demi titik pada seluruh daerah citra. Setiap pergeseran menghasilkan jendela baru, dan nilai piksel tengah dalam jendela ini diubah menjadi nilai median dari piksel-piksel dalam jendela tersebut. Filter median berfungsi untuk menggantikan nilai suatu piksel dengan nilai median dari tingkat keabuan piksel-piksel tetangganya (dengan tetap menggunakan nilai asli piksel saat menghitung median) [13]. Penggunaan metode median filtering bertujuan untuk menghilangkan *noise-noise* yang ada pada gambar citra USG yang telah dikonversi ke proses grayscale sebelumnya, maka proses median filtering juga dapat meningkatkan kualitas citra gambar yang mana sebelumnya mempunyai *noise-noise* yang bisa mengganggu proses selanjutnya.

H. Active Contour

Active contour merupakan sebuah proses di mana energi minimal dari suatu batasan objek mencapai nilai minimum dari fungsi energi. Fungsi energi ini tidak hanya berdasarkan informasi mengenai tepi (edge) dari objek, tetapi juga mencakup properti yang mengontrol pergerakan kurva itu sendiri. Active contour merepresentasikan 15 energi minimum yang memenuhi properti dari kurva active contour dan properti dari gambar digital [6].

Penerapan metode active contour pada gambar USG sangat efektif untuk segmentasi dan identifikasi struktur internal seperti tumor atau jaringan abnormal. Gambar USG sering kali memiliki tingkat noise yang tinggi dan kontras yang rendah, yang membuat deteksi tepi menjadi sulit. Dengan menggunakan metode active contour, proses segmentasi dapat lebih akurat karena kurva dapat menyesuaikan diri secara dinamis dengan kontur objek yang ingin diidentifikasi, bahkan dalam kondisi pencitraan yang kurang ideal [14].

Dalam aplikasi ini, kurva inisial diletakkan di sekitar area target pada gambar USG. Kemudian, kurva ini akan bergerak menuju batas objek dengan meminimalkan fungsi energi yang telah ditentukan. Fungsi energi ini mencakup informasi tepi dari gambar USG serta properti elastisitas dan kekakuan kurva, yang membantu kurva tetap halus dan menghindari perangkap noise. Dengan demikian, metode active contour memungkinkan identifikasi dan segmentasi yang lebih presisi dari massa atau lesi dalam gambar USG, yang sangat penting dalam deteksi dini dan diagnosis kanker payudara.

Setelah proses segmentasi selesai, hasilnya dapat digunakan sebagai input untuk proses klasifikasi menggunakan metode K-Nearest Neighbors (KNN). KNN akan mengklasifikasikan area yang telah di-segmentasi berdasarkan fitur-fitur yang diekstrak dari gambar, seperti tekstur, bentuk, dan intensitas piksel. Kombinasi antara segmentasi active contour dan klasifikasi KNN diharapkan dapat memberikan hasil yang lebih akurat dan andal dalam mendeteksi dan mengidentifikasi kanker payudara pada gambar USG.

I. Grey Level Co-Occurance Matrix

Menurut Kulkarni (1994) dalam (Susanto, 2013), Gray Level Co-Occurrence Matrix pertama kali diusulkan oleh Haralick pada tahun 1973 dengan 28 fitur untuk menggambarkan pola spasial. Metode ini merupakan salah satu pendekatan statistik yang dapat digunakan untuk menganalisis tekstur dalam citra medis yang seringkali memiliki tekstur yang tidak teratur. Ekstraksi fitur menggunakan Grey Level CoOccurrence Matrix untuk mengambil karakteristik tekstur [15].

J. K-Nearest Neighbor

Algoritma *K-Nearest Neighbor* (KNN) adalah metode klasifikasi untuk menentukan target baru berdasarkan jarak terdekat antara data training dan data testing. Algoritma KNN ini banyak digunakan dalam klasifikasi karena mudah diimplementasikan dan memberikan akurasi yang baik [16]. Dalam bidang kecerdasan buatan, teknik machine learning diperkenalkan untuk membantu meningkatkan kemampuan pendeteksian secara otomatis. Metode klasifikasi machine learning salah satu teknik data mining dapat membantu untuk mendeteksi penyakit kanker. Salah satu metode terbaik untuk masalah klasifikasi adalah K-Nearest Neighbor.

Salah satu metode klasifikasi yang umum digunakan adalah metode K-Nearest Neighbor (K-NN). Tujuan penggunaan metode ini adalah untuk mengklasifikasikan objek baru berdasarkan data latih yang paling dekat dengannya. Algoritma K-Nearest Neighbor relatif mudah diimplementasikan. Dalam metode ini, jumlah tetangga 19 terdekat yang akan digunakan ditentukan oleh pengguna, dan nilai ini dikenal sebagai nilai k [17].

KNN dipilih karena beberapa KNN adalah algoritma non-parametrik yang tidak membuat asumsi tertentu tentang distribusi data, sehingga lebih fleksibel dan dapat digunakan pada berbagai jenis dataset. KNN juga sangat efektif dalam menangani data dengan banyak kelas dan fitur yang tidak linier, karena klasifikasi didasarkan pada kedekatan langsung dalam ruang fitur. Selain itu KNN memiliki performa yang baik pada dataset yang berukuran kecil hingga menengah, di mana algoritma lain mungkin memerlukan waktu pelatihan yang lebih lama atau lebih kompleks [18].

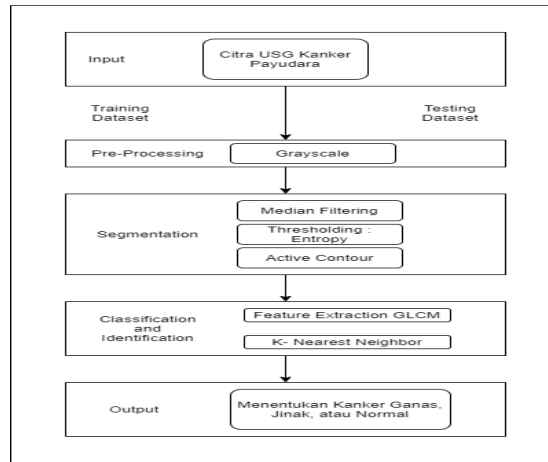
Kemudian, KNN juga memiliki kemampuan adaptasi yang baik dalam kasus di mana data latih diperbarui atau ditambahkan secara berkala, karena proses klasifikasi tidak memerlukan pelatihan ulang seluruh model. Kemudahan dalam implementasi dan interpretasi hasil juga menjadi keunggulan KNN, karena keputusan klasifikasi dapat dijelaskan secara intuitif dengan melihat tetangga terdekat yang mempengaruhi hasilnya.

III. HASIL DAN PEMBAHASAN

A. Analisis Rancangan Sistem

Rancangan sistem yang dimulai dengan implementasi analisis dari data citra asli yang belum melalui proses segmentasi. Tahapan pertama adalah pre-processing untuk mengubah citra menjadi citra skala abu-abu (grayscale), diikuti dengan proses segmentasi menggunakan beberapa metode yang telah ditentukan untuk memisahkan objek dan memperbaiki citra, yang kemudian diubah menjadi citra biner. Selanjutnya, tahap klasifikasi memanfaatkan ekstraksi fitur dari Grey Level Co-Occurrence Matrix untuk mendapatkan fitur-fitur yang akan diproses untuk identifikasi

menggunakan metode K-Nearest Neighbor. Tujuan akhir dari proses ini adalah identifikasi bentuk kanker yang dapat dikategorikan sebagai ganas, jinak, atau normal.



Gambar 2. Analisis Rancangan Sistem

B. Alat Penelitian

Alat penelitian yang digunakan sangat penting untuk memastikan keberhasilan dan akurasi dalam mengumpulkan data serta menganalisis informasi yang diperlukan. Dalam penelitian ini, alat-alat yang digunakan mencakup berbagai perangkat keras dan perangkat lunak seperti yang tersaji pada Tabel 1 dibawah ini :

TABEL I.
ALAT PENELITIAN

Hardware	Software
Spesifikasi Laptop :	Windows 10 64-Bit
Procesor Inter ® Core ™ i5-8265U CPU @ 1.60Ghz (8CPUs)	Microsoft Word
	Microsoft Excel
RAM 4GB	Google Chrome
	Snipping Tool
	Draw.io
SSD 512	Google Collab

C. Analisis Segmentasi

Analisis data dilakukan oleh peneliti untuk menjawab pokok permasalahan yang ada melalui data dan informasi yang diperoleh dalam penelitian. Terknik analisis ini membahas mengenai pengolahan gambar dengan menggunakan metode Grayscale, Median Filtering, dan Aktif Kontur untuk mendapatkan hasil segmentasi.

C.1 Proses Grayscale

Grayscale merupakan tahapan dimana warna citra RGB menjadi warna citra keabuan-abuan. Sampel nilai RGB pada citra “Benign 10.png” membentuk matrix 5x5, dimana nilai RGB diambil pada koordinat pixel X 184 sampai dengan 188 dan Y 258 sampai dengan 262. Berikut ini merupakan nilai RGB dalam bentuk matrix yang akan diolah menjadi citra grayscale.

57	65	67	69	69	57	65	67	69	69	57	65	67	69	69
62	65	67	69	69	62	65	67	69	69	62	65	67	69	69
60	62	63	64	63	60	62	63	64	63	60	62	63	64	63
67	72	74	76	77	67	72	74	76	77	67	72	74	76	77
77	84	89	94	92	77	84	89	94	92	77	84	89	94	92

Gambar 3. Matrix 5x5 Nilai RGB

Dalam proses grayscale, rumus operasi konversi Grayscale yang sesuai untuk menjadikan nilai citra menjadi keabu-abuan, dengan menggunakan persamaan 1

$$\text{Grayscale}(X, Y) = (R * W_r) + (G * W_g) + (B * W_b) \quad (1)$$

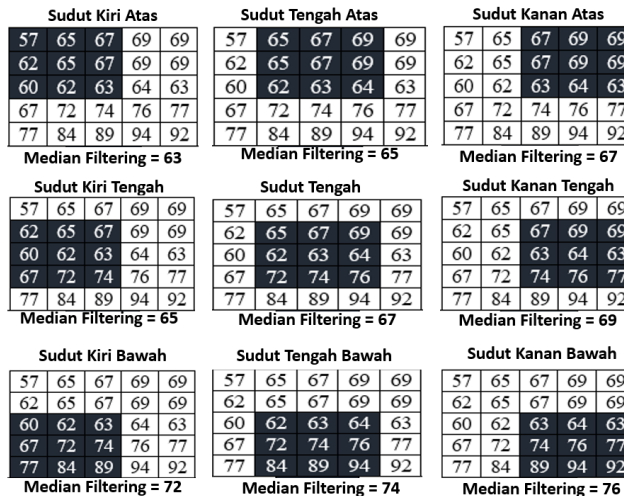
Dimana (X, Y) adalah Koordinat Titik Pixel, R adalah Nilai Red, G adalah Nilai Green, dan B adalah Nilai Blue, serta berdasarkan NTSC (National Television System Committee) yang menetapkan Nilai $W_r = 0.299$, Nilai $W_g = 0.587$ dan Nilai $W_b = 0.114$ maka didapat matrix baru 5x5 dari proses grayscale diatas, sebagai pada Tabels 2 berikut ini :

TABEL II.
 MATRIX 5X5 GRAY SCALE

57	65	67	69	69
62	65	67	69	69
60	62	63	64	63
67	72	74	76	77
77	84	89	94	92

C.2 Proses Median Filtering

Setelah mendapatkan matriks baru dari proses grayscale, proses median filtering dilakukan dengan mengambil matriks 3x3 dari setiap sudut untuk diolah menjadi median filtering. Proses ini diilustrasikan pada Gambar 3, yang menunjukkan proses median filtering dari matriks 5x5 grayscale.



Gambar 4. Matrix 3x3 pada setiap sudut

C.3 Proses Thresholding : Entropy

Proses ini menentukan nilai ambang batas dari matrix 5x5 median filtering, dengan persamaan 2 sebagai berikut :

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq T \\ 0 & \text{if } f(x, y) < T \end{cases} \quad (2)$$

Dimana (x, y) : Nilai Citra Biner, f(x, y) : Nilai Citra Median Filtering, T: Nilai Ambang Thresholding : Entropy

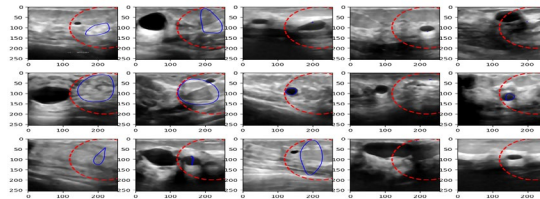
Dalam konteks nilai ambang batas sebesar 67, yang dianggap mencukupi untuk memahami pola pada kanker, syarat perubahan nilai biner dari matriks median filtering adalah sebagai berikut: jika nilai median filtering melebihi nilai ambang Thresholding, yang dalam hal ini adalah Entropy sebesar 67, maka nilai piksel akan diubah menjadi 1. Sebaliknya, jika nilai median filtering lebih rendah daripada nilai ambang Thresholding, yaitu Entropy sebesar 67, maka nilai piksel akan diubah menjadi 0. Sebagai hasilnya, kita mendapatkan matriks 5x5 dari proses Thresholding: Entropy, seperti yang terdokumentasikan dalam Tabel 3 di bawah ini:

TABEL III.
 MATRIX 5X5 THRESHOLDING : ENTROPY

0	0	1	1	1
0	0	0	1	1
0	0	1	1	0
1	1	1	1	1
1	1	1	1	1

C.4 Proses Active Contour

Pada proses segmentasi citra USG kanker payudara menggunakan metode Active Contour, terdapat energi kurva internal dan eksternal yang memengaruhi kurva active contour. Active contour membentuk batas sendiri untuk segmentasi objek[19]. Dalam penelitian ini, model balon digunakan untuk mensegmentasi wilayah objek pada citra USG kanker payudara. Hasil dari *active contour Snake* sebagai berikut :



Gambar 5. Hasil Active Contour Snake

hasil dari active contour snake, garis merah dan biru mengelilingi object kanker pada citra usg kanker payudara.

D. Analisis Klasifikasi

Analisis data dilakukan oleh peneliti untuk menjawab pertanyaan-pertanyaan yang penting yang ada melalui data dan informasi yang diperoleh dalam penelitian. Teknik analisis ini membahas mengenai perhitungan angka dengan proses Grey Level Co-Occurance Matrix, Metode Fitur Dispersi, dan untuk klasifikasinya menggunakan metode K-Nearest Neighbor.

D.1 Proses Grey Level Co-Occurance Matrix (GLCM)

Dalam proses Grey Level Co-Occurrence Matrix, citra diubah menjadi citra skala abu-abu, kemudian matrix kerangka kerja pada GLCM dihitung berdasarkan ketergantungan spasial yang terkait dengan berbagai sudut jarak antara pixel. Dalam salah satu contoh pengujian, setelah proses preprocessing, hasil dari metode Grey Level Co-Occurrence Matrix didapatkan nilai matrix dari citra asli yang berukuran 10x10. Contoh perhitungan nilai matrix tersebut dapat dilihat pada Tabel 4 di bawah ini, yang dihasilkan melalui pengujian menggunakan MATLAB[20].

TABEL IV.
 MATRIX 10X10 PRE-PROCESSING

95	93	89	87	83	77	71	62	66	60
85	80	76	74	71	67	63	60	59	57
66	61	57	53	51	50	49	50	51	52
55	52	48	46	45	45	46	48	51	53
47	46	43	44	43	44	45	47	50	51
53	54	53	51	50	49	48	48	49	50
64	63	61	57	55	53	51	51	51	52
67	66	65	61	58	55	53	53	54	54
61	61	61	58	55	52	52	53	56	56
59	57	56	52	51	50	52	55	58	60

Tabel 4 di atas merupakan representasi visual dari matriks 10x10 yang digunakan dalam perhitungan GLCM. Setiap nilai dalam matriks mencerminkan intensitas piksel pada lokasi tertentu dalam citra, yang kemudian digunakan dalam perhitungan GLCM untuk mengekstraksi fitur tekstur. Ini akan memberikan pemahaman yang lebih baik tentang bagaimana GLCM dihitung dan bagaimana nilai-nilai tersebut merepresentasikan tekstur citra.

Setelah memperoleh matriks 10x10 dari citra yang telah melalui proses preprocessing, langkah selanjutnya adalah melakukan konversi nilai dalam matriks tersebut ke dalam rentang level tertentu. Rentang level ini dibagi menjadi 10 kategori yang berbeda, yang masing-masing berkorelasi dengan sejumlah nilai piksel dalam matriks GLCM. Rentang nilai untuk masing-masing level ditentukan pada tabel 5 sebagai berikut :

Tabel 5. Rentang Nilai untuk Matrik GCLM

Level	Rentang Nilai
0	0 – 25
1	26 – 51
2	52 - 77
3	78 – 103
4	104 – 129
5	130 – 155
6	156 – 181
7	182 – 206
8	207 – 232
9	233 – 255

Dalam tabel ini, setiap level memiliki rentang nilai yang berbeda yang mencakup sejumlah nilai piksel dalam matriks GLCM. Misalnya, level 0 memiliki rentang nilai antara 0 hingga 25, sedangkan level 9 memiliki rentang nilai antara 233 hingga 255. Proses ini memungkinkan untuk mengelompokkan nilai-nilai piksel ke dalam kategori yang lebih terdefinisi, yang akan digunakan dalam perhitungan lanjutan pada analisis GLCM.

Proses Grey Level Co-Occurrence Matrix (GLCM) dibangun menggunakan histogram tingkat dua, yang menghasilkan matriks berukuran LxL, di mana L mewakili jumlah tingkat keabuan yang ada dalam citra. Elemen-elemen matriks GLCM, yang direpresentasikan sebagai P(X1, X2), mencerminkan distribusi probabilitas bersama antara dua tingkat keabuan yang berbeda. Dengan adanya GLCM, proses ekstraksi fitur menjadi lebih sederhana karena dapat menganalisis hubungan spasial antara piksel dalam citra. Rentang nilai yang digunakan untuk GLCM adalah dari 0 hingga 9, yang mencerminkan pembagian matriks ke dalam 10 kategori yang berbeda. Proses konversi matriks 10x10 ke rentang nilai mengubah nilai piksel dalam matriks ke dalam kategori yang sesuai dengan rentang nilai yang telah ditentukan. Hasil proses konversi matriks 10x10 ke rentang nilai diperlihatkan dalam tabel 6 berikut:

TABEL VI.
 HASIL KONVERSI MATRIX 10X10 KE RENTANG NILAI

3	3	3	3	3	2	2	2	2	2
3	3	2	2	2	2	2	2	2	2
2	2	2	2	1	1	1	1	1	2
2	2	1	1	1	1	1	1	1	2
1	1	1	1	1	1	1	1	1	1
2	2	2	1	1	1	1	1	1	1
2	2	2	2	2	2	1	1	1	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2
2	2	2	2	1	1	2	2	2	2

Dalam tabel ini, setiap angka mewakili tingkat keabuan yang telah dikonversi ke dalam rentang nilai yang sesuai. Proses konversi ini memungkinkan untuk menganalisis distribusi probabilitas piksel-piksel dalam citra dan mengekstraksi fitur-fitur tekstur yang relevan untuk analisis lebih lanjut

Selanjutnya proses perhitungan sesuai arah berdasarkan ketergantungan spasial yang terkait dengan berbagai sudut jarak antara pixel dengan menggunakan rumus Kontras pada persamaan 3, rumus Energi pada persamaan 4, rumus Homogenitas pada persamaan 5 dan rumus Korelasi pada persamaan 6

Persamaan Kontras :

$$\sum_i \sum_j (i^2 - j^2) P[i, j] \tag{3}$$

Persamaan Energi :

$$\sum_i \sum_j P^2[i, j] \tag{4}$$

Persamaan Homogenitas :

$$\sum_i \sum_j \frac{P[i, j]}{1+[i-j]} \tag{5}$$

Persamaan Korelasi :

$$\sum_i \sum_j \frac{I(i, j)(i-\mu_i)(i-\mu_j)}{\sigma_i \sigma_j} \tag{6}$$

Dalam persamaan tersebut, p mewakili probabilitas yang bernilai antara 0 dan 1, yang merupakan elemen dari matriks. Indeks (i, j) menjelaskan probabilitas p pada baris ke- i dan kolom ke- j . μ_i dan μ_j merupakan rata-rata dari nilai pada baris i dan kolom j , sementara σ_i dan σ_j adalah standar deviasi dari nilai pada baris i dan kolom j .

D.2 Proses K-Nearest Neighbor

Setelah proses Grey Level Co-Occurance Matrix selesai dan diketahui hasil dari masing-masing fitur dari data uji dan data training, selanjutnya proses klasifikasi data uji dan data training menggunakan metode K-Nearest Neighbor (KNN), pada sistem yang akan dibuat ini terjadi pada proses identifikasi. Langkah-langkah dalam metode K-Nearest Neighbor (KNN), sebagai berikut :

1. Menghitung jarak Euclidean, dengan persamaan 7 dibawah ini :

$$D(X_i, X_j) = \sqrt{\sum_{r=1}^n ((a_r(x_i) - a_r(x_j))^2)} \tag{7}$$

Keterangan :

$D(X_i, X_j)$ = Jarak kedekatan angka euclidean

X_i = Rekaman ke- i

X_j = Rekaman ke- j

a_r = Data ke- r

2. Mengurutkan berdasarkan nilai Euclidean Distance
3. Menentukan nilai K klasifikasi dengan ketetanggan terdekat
4. Target output dari sistem adalah menentukan kelas mayoritas dari data uji. Tahapan yang diperlukan untuk mengklasifikasi data uji dan data training, sehingga dapat mengetahui jenis dari data uji adalah sebagai berikut: Pertama, hasil dari setiap fitur pada Grey Level Co-Occurrence Matrix dari data uji dan data training diperoleh, kemudian dilakukan pengklasifikasian data uji dan data training untuk mengidentifikasi jenis kanker pada data uji. Selanjutnya, informasi nilai parameter yang diperoleh dari proses Grey Level Co-Occurrence Matrix digunakan untuk melakukan pengklasifikasian K-Nearest Neighbor (KNN). Setelah itu, dilakukan klasifikasi masing-masing fitur Grey Level Co-Occurrence Matrix data training dengan parameter fitur Grey Level Co-Occurrence Matrix data uji. Hasil minimum dari klasifikasi ini menentukan jenis dari data uji. Akhirnya, keputusan akhir terhadap data uji dapat diperoleh untuk menentukan bentuk kanker pada data tersebut.
- 5.

TABEL VII.
 DATA TRAINING SUDUT 0°

NO.	Kontras 0°	Energi 0°	Homogenitas 0°	Korelasi 0°
1	5.46743	0.03549	0.53273	0.99916
2	4.60064	0.18408	0.63790	0.99911
3	6.55433	0.04202	0.54718	0.99850
4	13.1303	0.04676	0.55186	0.99669
5	4.31116	0.04053	0.60183	0.99879

6	7.31321	0.03334	0.52271	0.99840
7	8.44620	0.03239	0.49541	0.99766
8	6.82522	0.04036	0.56298	0.99764
9	6.02787	0.05260	0.60141	0.99865
10	5.23537	0.06480	0.59869	0.99773

Tabel 7 menampilkan contoh data pelatihan untuk sudut 0° , yang mencakup nilai kontras, energi, homogenitas, dan korelasi untuk setiap sampel data. Data ini digunakan dalam proses pelatihan model untuk klasifikasi citra. Setiap baris dalam tabel mewakili satu sampel data pelatihan, dengan kolom-kolom yang menunjukkan fitur-fitur yang diekstraksi dari citra untuk analisis lebih lanjut. Misalnya, pada baris pertama, nilai kontras adalah 5.46743, energi adalah 0.03549, homogenitas adalah 0.53273, dan korelasi adalah 0.99916. Setiap nilai ini merupakan representasi dari karakteristik tekstur citra pada sudut 0° dan akan digunakan untuk melatih model klasifikasi. Data ini merupakan bagian dari proses pelatihan yang penting dalam pengembangan model untuk mengklasifikasikan citra kanker payudara. Dengan menggunakan fitur-fitur ini, model dapat belajar untuk membedakan antara citra kanker payudara yang ganas, jinak, dan normal berdasarkan

TABEL VIII.
 DATA TRAINING SUDUT 45°

NO.	Kontras 45°	Energi 45°	Homogenitas 45°	Korelasi 45°
1	19.00641	0.02912	0.40212	0.99708
2	19.41865	0.17937	0.48682	0.99625
3	26.50393	0.03349	0.39583	0.99390
4	29.98247	0.03985	0.41722	0.99238
5	12.30342	0.03221	0.46130	0.99651
6	17.56168	0.02752	0.41646	0.99617
7	33.07594	0.02493	0.35255	0.99080
8	24.54273	0.02957	0.38840	0.99150
9	22.50933	0.04080	0.43820	0.99491
10	20.06364	0.05773	0.44992	0.99130

Tabel 8 menyajikan data contoh pelatihan untuk sudut 45° , yang mencakup nilai kontras, energi, homogenitas, dan korelasi untuk setiap sampel. Data ini diperlukan dalam proses pelatihan model untuk mengembangkan kemampuan klasifikasi citra kanker payudara. Setiap baris dalam tabel mewakili satu sampel pelatihan, menyediakan informasi tentang karakteristik tekstur citra pada sudut 45° . Misalnya, pada baris pertama, nilai kontras adalah 19.00641, energi adalah 0.02912, homogenitas adalah 0.40212, dan korelasi adalah 0.99708. Data ini menjadi landasan penting bagi model untuk memahami variasi dalam tekstur citra dan membuat prediksi yang lebih akurat saat diberikan citra baru untuk diklasifikasikan.

TABEL IX.
 HASIL DATA UJI DARI BERBAGAI SUDUT

SUDUT	0°	45°	90°	135°
Kontras	11	21	24	27
Energi	2720	1780	1667	1686
Homogenitas	87	77	78	79
Korelasi	5,507	13,347	15,996	21,925

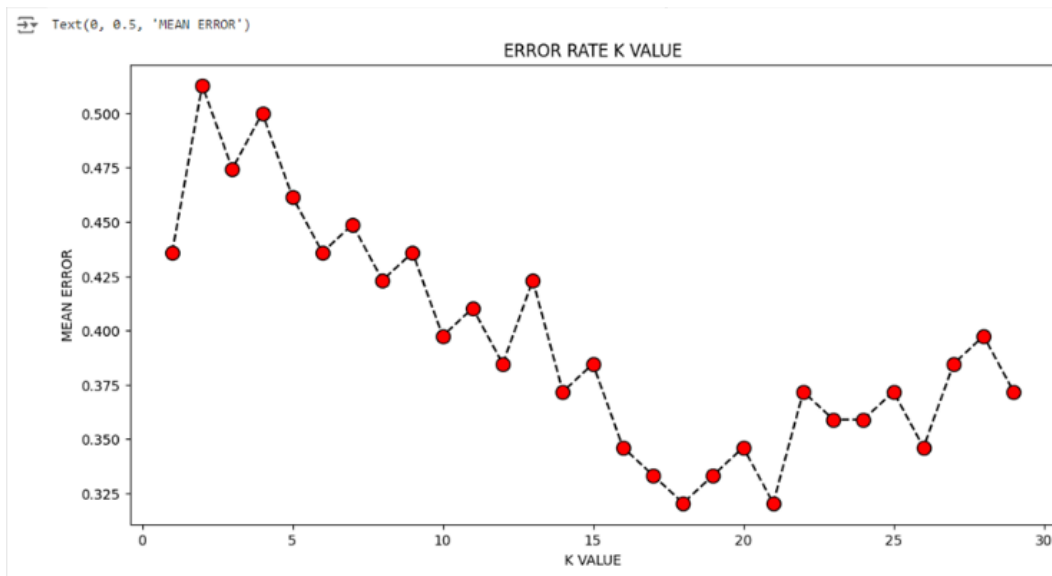
Tabel 9 hingga 12 menyajikan hasil dari data uji pada berbagai sudut (0° , 45° , 90° , dan 135°) dalam konteks analisis tekstur citra kanker payudara. Setiap tabel memberikan informasi mengenai nilai kontras, energi, homogenitas, dan korelasi yang diekstraksi dari citra yang diuji pada sudut yang sesuai. Misalnya, pada Tabel 9, untuk sudut 0° , diperoleh nilai kontras sebesar 11, energi sebesar 2720, homogenitas sebesar 87, dan korelasi sebesar 5.507. Data ini memberikan pemahaman tentang karakteristik tekstur citra pada sudut tertentu, yang penting untuk proses klasifikasi citra kanker payudara berdasarkan fitur tekstur yang diekstraksi.

Setelah proses perhitungan menggunakan KNN selesai, maka contoh hasil yang didapat untuk prediksi data uji tersebut disajikan pada Tabel 13 sebagai berikut :

TABEL X.
 CONTOH HASIL DATA UJI DAN DATA TRAINING SETELAH PROSES KNN SUDUT 0°

NO.	Kontras 0°	Energi 0°	Homogenitas 0°	Korelasi 0°	Label
1	10.14524	0.03376	0.49809	0.99715	benign
2	9.83646	0.03656	0.54694	0.99825	benign
3	3.46386	0.06360	0.64312	0.99753	normal
4	7.95089	0.05089	0.58690	0.99884	benign
5	5.77388	0.04759	0.58335	0.99864	malignant
6	9.24637	0.03355	0.52108	0.99816	benign
7	8.82770	0.03565	0.52781	0.99867	normal
8	6.07181	0.03884	0.57583	0.99900	benign
9	5.95544	0.04494	0.57943	0.99891	normal
10	4.41060	0.05649	0.61496	0.99825	normal
11	8.08013	0.04275	0.54294	0.99780	benign
12	7.72947	0.06624	0.60272	0.99799	benign
13	6.95539	0.03540	0.52256	0.99763	normal
14	10.87701	0.03211	0.48707	0.99794	benign
15	10.51248	0.03173	0.50382	0.99764	benign
16	4.93932	0.21924	0.64694	0.99769	malignant
17	4.37816	0.04177	0.59551	0.99870	benign
18	8.17765	0.04822	0.58108	0.99868	benign
19	17.95415	0.09161	0.54722	0.99818	benign
20	8.37877	0.07917	0.58715	0.99851	benign

Setelah semua data uji dan data trainign diproses dengan KNN maka hasilnya didapatkan, langkah selanjutnya adalah memisahkan label yang sesuai dengan ketentuan nilai K. Nilai K yang diambil dalam konteks ini adalah 18, yang disajikan dalam gambar sebagai berikut :



Gambar 8. Akurasi Nilai Error K

Berdasarkan gambar diatas yang merupakan grafik akurasi nilai error K dapat dilihat bahwa akurasi K=18 nilainya tertinggi yaitu 75% daripada yang lainnya. Grafik akurasi nilai error K menunjukkan bahwa K=18 merupakan nilai optimal, karena memberikan kombinasi terbaik antara generalisasi dan sensitivitas terhadap data latih.

Setelah pemilihan nilai K optimal, model KNN dapat digunakan untuk melakukan klasifikasi terhadap data uji. Dalam penelitian ini, mayoritas prediksi dari hasil jarak nilai K terkecil menunjukkan kelas kanker benign (jinak), yang menjadi hasil klasifikasi akhir untuk data uji.

TABEL XI.
 HASIL DATA UJI DAN DATA TRAINIG

Nilai K	Label
K1	Benign
K2	Malignant
K3	Benign
K4	Benign
K5	Normal
K6	Benign
K7	Benign
K8	Benign
K9	Malignant
K10	Benign
K11	Malignant
K12	Normal
K13	Normal
K14	Malignant
K15	Normal
K16	Normal
K17	Benign
K18	Malignant

Dari hasil tabel nilai K diatas menampilkan nilai K terkecil dari seluruh hasil data uji dan training yang sudah diproses dengan K-Nearest Neighbor, sehingga hasil klasifikasi untuk data uji menunjukkan Kanker Benign (Kanker Jinak) karena pada tabel Kanker Benign adalah mayoritas dari hasil jarak nilai K terkecil. Setelah hasil identifikasi sudah diketahui, maka berakhir proses pengolahan data dan perhitungan data training. Pada tahap ini akan dilakukan beberapa kali percobaan klasifikasi deteksi penyakit kanker payudara dengan menggunakan algoritma KNN. Dataset yang sudah ada sebelumnya akan dibagi menjadi dua bagian. Satu bagian menjadi data training dan bagian yang lain akan menjadi data testing. Pemilihan nilai parameter k akan ditentukan dengan membandingkan akurasi dari beberapa percobaan yang dilakukan. Nilai parameter k paling optimal didapatkan dengan mengambil hasil akurasi yang paling baik.

E. Proses Confussion Matrix

Hasil dari penggunaan confusion matrix untuk menghitung performa sistem. Matrix adalah tabel yang terdiri dari jumlah baris jenis klasifikasi yang diprediksi oleh sistem klasifikasi sebagai benar dan salah untuk menentukan kinerja sistem. Berikut adalah tabel confusion matrix berdasarkan klasifikasi menggunakan metode K-Nearest Neighbor pada sistem :

TABEL XII.
 HASIL PREDIKSI DATA UJI

Actual	Predicted		
	Benign	Malignant	Normal
Benign	37	3	0
Malignant	14	15	0
Normal	7	1	1

Dari 78 data uji jenis kanker benign, kanker malignant, dan kanker normal menggunakan sistem, maka diperoleh hasil Benign to Benign sebanyak 37 data, Benign to Malignant sebanyak 3 data, Malignant to Benign sebanyak 14 data, Malignant to Malignant sebanyak 15 data, Normal to Benign sebanyak 7 data, Normal to malignant sebanyak 1 data, dan Normal to Normal sebanyak 1 data. Setelah diketahui hasil maka langkah selanjutnya menghitung performa Akurasi, sebagai berikut :

$$Akurasi = \frac{37 + 15 + 1}{37 + 3 + 14 + 15 + 7 + 1 + 1}$$

$$\begin{aligned} Akurasi &= \frac{53}{78} \\ Akurasi &= 0,68 \times 100\% \\ Akurasi &= 68\% \end{aligned}$$

Dengan akurasi sebesar 68%, hasil ini menunjukkan bahwa sistem klasifikasi yang digunakan memiliki tingkat keberhasilan yang cukup baik dalam mengidentifikasi jenis kanker berdasarkan data uji yang diberikan. Akurasi 68% berarti bahwa 68% dari prediksi yang dilakukan oleh sistem adalah benar, sesuai dengan kelas asli dari data uji.

Namun, untuk mendapatkan gambaran yang lebih komprehensif mengenai performa sistem, perlu juga dipertimbangkan metrik evaluasi lainnya seperti sensitivitas (recall), spesifisitas, dan presisi. Sensitivitas mengukur kemampuan sistem dalam mengidentifikasi kasus positif dengan benar, sedangkan spesifisitas mengukur kemampuan sistem dalam mengidentifikasi kasus negatif dengan benar. Presisi, di sisi lain, mengukur berapa banyak dari prediksi positif yang benar-benar positif. Dengan menganalisis metrik-metrik ini, kita dapat lebih memahami kekuatan dan kelemahan sistem klasifikasi yang digunakan.

Selain itu, analisis kesalahan atau error analysis penting dilakukan untuk mengetahui penyebab utama dari kesalahan klasifikasi yang terjadi, seperti kasus Benign to Malignant sebanyak 3 data dan Malignant to Benign sebanyak 14 data. Dengan memahami pola kesalahan ini, perbaikan dapat dilakukan pada model atau fitur yang digunakan dalam klasifikasi, yang pada akhirnya diharapkan dapat meningkatkan akurasi dan performa keseluruhan sistem.

F. Interpretasi hasil penelitian

Hasil penelitian ini memberikan pemahaman yang mendalam tentang potensi metode segmentasi dan klasifikasi dalam deteksi kanker payudara melalui analisis citra digital. Dalam konteks penggunaan metode Grayscale, Median Filtering, dan Aktif Kontur, penelitian ini menunjukkan bahwa proses aktif kontur membantu mengidentifikasi morfologi kanker dengan tepat, memungkinkan segmentasi yang akurat dari gambaran kanker dalam citra. Sementara itu, penggunaan metode K-Nearest Neighbor (KNN) untuk klasifikasi citra memperkuat temuan sebelumnya dengan menunjukkan kemampuan untuk membedakan antara gambar kanker dan non-kanker dengan akurasi yang signifikan. Meskipun akurasi sebesar 68% yang diperoleh dari pengujian KNN pada data uji menyoroti tantangan dalam klasifikasi, hasil ini tetap memberikan dasar penting bagi pengembangan algoritma yang lebih canggih di masa depan. Implikasinya, terutama dalam konteks rumusan masalah terkait evaluasi akurasi klasifikasi kanker payudara, menunjukkan perlunya pengembangan dan pembaruan terus-menerus dalam metode analisis citra untuk meningkatkan kehandalan proses deteksi. Namun meskipun proses segmentasi dan klasifikasi menunjukkan kemajuan signifikan dalam deteksi kanker payudara, masih ada ruang untuk perbaikan. Interpretasi hasil menyoroti pentingnya untuk terus mengembangkan dan menguji metode analisis citra yang lebih kompleks dan akurat guna meningkatkan deteksi dini dan diagnosis kanker payudara. Selain itu, penekanan pada penggunaan teknologi untuk mendukung diagnosis dan perawatan kanker payudara memunculkan pertanyaan tentang bagaimana integrasi teknologi ini dapat ditingkatkan untuk memberikan solusi yang lebih efektif dalam penanganan kanker payudara secara keseluruhan.

IV. KESIMPULAN

Dari hasil penelitian di atas, dapat disimpulkan bahwa proses aktif kontur memiliki peran penting dalam identifikasi dan segmentasi objek kanker dalam citra. Dengan membatasi fokus pada objek kanker, proses GLCM kemudian dapat diaplikasikan secara lebih efisien, menghasilkan pemrosesan yang lebih tepat dan akurat terhadap area yang relevan. Pendekatan ini membantu meningkatkan kualitas analisis tekstur citra dan memfasilitasi proses klasifikasi selanjutnya.

Selanjutnya, hasil pengujian menggunakan metode K-Nearest Neighbor (KNN) menunjukkan bahwa nilai $K=18$ memberikan hasil yang optimal dalam klasifikasi citra kanker payudara. Dengan pembagian data training dan uji yang seimbang serta penerapan teknik K-Fold Cross Validation, model KNN mampu memberikan performa yang baik dalam mengklasifikasikan citra. Hal ini menegaskan bahwa penggunaan teknik pembelajaran mesin seperti KNN memiliki potensi yang signifikan dalam analisis citra medis.

Algoritma KNN dapat digunakan untuk mengidentifikasi dan mengklasifikasikan kanker payudara dengan akurasi yang memadai. Hasil penelitian menunjukkan bahwa metode ini memang dapat diterapkan, namun dengan akurasi yang saat ini mencapai 68%, ada indikasi bahwa hasil ini dapat ditingkatkan lebih lanjut.

Jika dibandingkan dengan metode lain seperti SVM (Support Vector Machine), Random Forest, atau CNN (Convolutional Neural Networks), KNN cenderung lebih sederhana namun bisa kurang efektif pada dataset yang kompleks atau tidak seimbang. Oleh karena itu, hasil yang diperoleh dengan KNN mungkin lebih rendah dibandingkan dengan metode klasifikasi tersebut di atas. Namun, kelebihan KNN adalah kemudahan implementasi dan interpretasi, yang membuatnya masih relevan untuk studi awal atau sebagai baseline.

Oleh karena itu, sebagai saran untuk penelitian mendatang, dapat dieksplorasi penggunaan metode lain atau penyesuaian parameter yang lebih baik guna meningkatkan akurasi klasifikasi. Selain itu, integrasi dengan teknologi lain seperti deep learning atau penggunaan fitur ekstraksi yang lebih baik juga dapat menjadi langkah selanjutnya dalam meningkatkan efektivitas deteksi dan diagnosis kanker payudara melalui citra medis.

DAFTAR PUSTAKA

- [1] D. Pittara, "Pengertian Kanker Payudara," *www.alodokter.com*, 2023. <https://www.alodokter.com/kanker-payudara> (accessed Jan. 09, 2023).
- [2] K. K. Payudara and B. F. Gabor, "Penelitian ini bertujuan untuk mengambil nilai fitur dari citra termal payudara melalui ekstraksi fitur Filter Gabor, dengan fokus pada," vol. 9, no. 4, 2024.
- [3] H. Sung *et al.*, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.
- [4] Kemenkes, "Kanker Payudara Paling Banyak di Indonesia, Kemenkes Targetkan Pemerataan Layanan Kesehatan," *www.sehatnegeriku.kemkes.go.id*, 2022. <https://sehatnegeriku.kemkes.go.id/baca/umum/20220202/1639254/kanker-payudara-paling-banyak-di-indonesia-kemenkes-targetkan-pemerataan-layanan-kesehatan/> (accessed Feb. 02, 2022).
- [5] E. Marfianti, "Peningkatan Pengetahuan Kanker Payudara dan Ketrampilan Periksa Payudara Sendiri (SADARI) untuk Deteksi Dini Kanker Payudara di Semutan Jatimulyo Dlingo," *J. Abdimas Madani dan Lestari*, vol. 3, no. 1, pp. 25–31, 2021, doi: 10.20885/jamali.vol3.iss1.art4.
- [6] A. Nugroho, A. Fauzi, B. Sunarko, H. Wibawanto, and N. Iksan, "Simplifikasi Model Cv Teregularisasi Berpadu Operasi," *J. Inform. Polinema*, vol. 8, no. 2, pp. 49–56, 2022, [Online]. Available: <http://jip.polinema.ac.id/ojs3/index.php/jip/article/view/913>
- [7] W. Ramdhani, D. Bona, R. B. Musyaffa, and C. Rozikin, "Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma K-Nearest Neighbor," *J. Ilm. Wahana Pendidik*, vol. 8, no. 12, pp. 445–452, 2022.
- [8] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Dataset of breast ultrasound images," *Data Br.*, vol. 28, p. 104863, 2020, doi: 10.1016/j.dib.2019.104863.
- [9] N. Meilani and O. Nurdiawan, "Data Mining untuk Klasifikasi Penderita Kanker Payudara Menggunakan Algoritma K-Nearest Neighbor," *J. Wahana Inform.*, vol. 2, no. 1, pp. 177–187, 2023, [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer>.
- [10] N. Afiatuddin, M. T. Wicaksono, V. R. Akbar, and D. Wulandari, "Komparasi Algoritma Machine Learning dalam Klasifikasi Kanker Payudara," vol. 8, no. April, pp. 889–899, 2024, doi: 10.30865/mib.v8i2.7457.
- [11] Yudha, "Penerapan algoritma," vol. V, pp. 217–224, 2014.
- [12] R. Nanda, S. W. Dari, and A. Ihsan, "Segmentasi Citra Medis untuk Deteksi Objek FAM pada Payudara Menggunakan Metode Sobel," *J. Media Inform. Budidarma*, vol. 3, no. 4, p. 248, 2019, doi: 10.30865/mib.v3i4.1232.
- [13] K. Djunaidi, N. Lanang, A. Syahdu, and D. Kuswardani, "Penerapan Metode Thresholding : Entropy dan Mathematical Morphology Pada Segmentasi Citra USG Kanker Payudara," vol. 11, no. 2, pp. 1–7, 2022, [Online]. Available: <https://doi.org/10.33322/kilat.v11i2.1503>
- [14] M. Rahmawaty, "Pendekatan Neutrosophic Set Untuk Segmentasi Citra Ultrasonografi Payudara Berdasarkan Metode Otsu," *JST (Jurnal Sains dan Teknol.)*, vol. 9, no. 1, pp. 11–19, 2020, doi: 10.23887/jstundiksha.v9i1.21153.
- [15] L. M. Wisudawati, "Klasifikasi Tumor Jinak Dan Tumor Ganas Pada Citra Mammogram Menggunakan Gray Level Co-Occurrence Matrix (GlcM) Dan Support Vector Machine (Svm)," *J. Ilm. Inform. Komput.*, vol. 26, no. 2, pp. 176–186, 2021, doi: 10.35760/ik.2021.v26i2.4897.
- [16] I. N. Atthalla, A. Jovandy, and H. Habibie, "Klasifikasi Penyakit Kanker Payudara Menggunakan Metode K Nearest Neighbor," *Pros. Annu. Res. Semin.*, vol. 4, no. 1, pp. 148–151, 2018, doi: 10.14421/fourier.2023.122.79-85.
- [17] M. I. P. Putra, D. T. Mardiansyah, and A. Aditsania, "Implementasi Algoritma Modified K-Nearest Neighbor (MKNN) untuk Klasifikasi Penyakit Kanker Payudara Tugas Akhir diajukan untuk memenuhi salah satu syarat memperoleh gelar sarjana dari Program Studi S1 Ilmu Komputasi M Ikhsan Perdana Putra Program Studi," *E-proceeding Eng.*, vol. 6, no. 1, pp. 2431–2441, 2019.
- [18] I. N. Atthalla, A. Jovandy, and H. Habibie, "Klasifikasi Penyakit Kanker Payudara Menggunakan Metode K Nearest Neighbor," *Pros. Annu. Res. Semin.*, vol. 4, no. 1, pp. 148–151, 2018.
- [19] S. Ahdan, A. Thyo Priandika, F. Andhika, and F. Shely Amalia, "Perancangan Media Pembelajaran Teknik Dasar Bola Voli Menggunakan Teknologi Augmented Reality Berbasis Android Learning Media for Basic Techniques of Volleyball Using Android-Based Augmented Reality Technology," *Education*, vol. 8, no. 3, pp. 1–16, 2020, [Online]. Available: <http://journalbalitbangdalampung.org>
- [20] L. Marlina, K. Munadi, and F. Arnia, "Klasifikasi Kanker Payudara Menggunakan Teknik Support Vector Machine (Svm) Pada Citra Termografi," *J. Komputer, Inf. Teknol. dan Elektro*, vol. 5, no. 2, pp. 11–13, 2020, doi: 10.24815/kitektro.v5i2.15571.