
SEGMENTASI PENYAKIT MATA KATARAK DENGAN METODE *K-MEANS*

Ayu Mestika Nurdy¹, Yusuf Ramadhan Nasution², Suhardi³

Universitas Islam Negeri Sumatera Utara. Medan

email: ayumestikanurdy@gmail.com¹, ramadhannst@uinsu.ac.id², suhardi@uinsu.ac.id³

Abstract: *Cataracts are one of the most common visual impairments, especially among the elderly. In Indonesia, the number of cataract patients is alarmingly high and poses a serious public health concern. It is estimated that around 50% of global blindness cases are caused by cataracts, with approximately 90% of patients residing in developing countries, including Indonesia. According to Nourkinan, Indonesia has the highest number of cataract cases in Southeast Asia, with approximately 2 million individuals affected and an additional 240,000 new cases each year at risk of blindness. Although cataracts are typically associated with individuals over the age of 55, recent trends indicate a growing number of cases among younger people aged between 35 and 40. In response to this issue, this research aims to design an image segmentation system to detect cataract disease using the K-Means algorithm. The system is expected to serve as an educational tool for students and learners to better understand how digital image analysis can be used to identify cataract symptoms. This study utilizes a dataset consisting of 330 retinal images. Based on testing using a confusion matrix, the segmentation process achieved an accuracy rate of 94.84%. To further improve the system's performance, increasing the number of training data is recommended to enhance segmentation accuracy.*

Keyword: *K-Means Clustering, Segmentation, Cataract Eye Disease*

Abstrak: Katarak merupakan salah satu gangguan penglihatan yang umum terjadi, terutama pada individu lanjut usia. Di Indonesia, jumlah penderita katarak tergolong tinggi dan menjadi masalah kesehatan yang cukup serius. Diperkirakan bahwa sekitar 50% kasus kebutaan di dunia disebabkan oleh katarak, dan sekitar 90% penderitanya berada di negara-negara berkembang, termasuk Indonesia. Berdasarkan pernyataan Nourkinan, Indonesia menempati posisi teratas di Asia Tenggara dengan jumlah penderita katarak mencapai sekitar 2 juta orang, dan setiap tahunnya terdapat penambahan sekitar 240 ribu kasus baru yang berisiko mengalami kebutaan. Walaupun umumnya diderita oleh masyarakat berusia di atas 55 tahun, kasus katarak juga mulai banyak ditemukan pada usia yang lebih muda, yakni antara 35 hingga 40 tahun. Maka dari itu, penulis bertujuan membuat sistem segmentasi penyakit mata katarak dengan menggunakan metode *K-Means* untuk membantu para mahasiswa dan pelajar untuk memahami lebih dalam tentang bagaimana citra mata dapat dianalisis secara teknis untuk mengenali indikasi penyakit mata katarak. Dataset yang digunakan penelitian ini berjumlah 330 dataset, berdasarkan pengujian yang telah dilakukan dalam mensegmentasi menggunakan *confussion matrix* tingkat keberhasilan 94.84%. Untuk memaksimalkan hasil penelitian dibutuhkan penambahan jumlah data latih untuk meningkatkan akurasi dari proses segmentasi.

Kata kunci: *K-Means Clustering, Segmentasi, Penyakit Mata Katarak*

PENDAHULUAN

Gangguan penglihatan atau penyakit pada mata merupakan kondisi medis yang dapat berdampak signifikan terhadap kualitas dan harapan hidup seseorang. Terdapat berbagai jenis penyakit mata, di antaranya katarak, glaukoma, dan gangguan retina, yang termasuk sebagai penyebab utama terjadinya kebutaan (Abdu et al., 2021). Oleh karena itu, deteksi dini terhadap penyakit atau kelainan pada mata sangatlah penting untuk mencegah terjadinya kebutaan di kemudian hari (Nugraha et al., 2024).

Katarak adalah penyakit mata yang sering ditemukan pada orang yang sudah lanjut usia. Di Indonesia, jumlah penderita katarak cukup memprihatinkan. Menurut Nourkinan, Indonesia memiliki jumlah penderita katarak tertinggi di Asia Tenggara, mencapai 2 juta orang, dengan tambahan 240 ribu orang setiap tahunnya yang berisiko mengalami kebutaan. Penderita katarak umumnya berusia 55 tahun ke atas, namun kini penyakit ini juga banyak ditemukan pada usia yang lebih muda, antara 35 sampai 40 tahun (Raflin Sinaga et al., 2023).

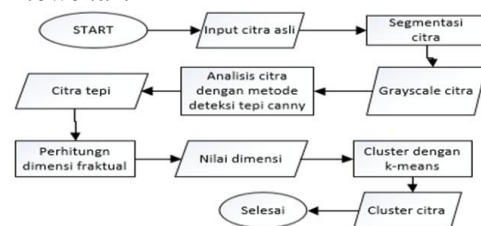
Teknologi memiliki peran penting dalam hampir setiap bidang, termasuk dalam bidang kesehatan. Hal ini mendorong para ahli untuk memanfaatkan pengolahan citra yang dapat menyimpan pengetahuan dalam mendiagnosis dan mengidentifikasi penyakit mata katarak pada manusia dengan hasil yang konsisten. Teknologi ini juga memungkinkan solusi yang tepat dan cepat untuk penyakit mata, membantu pakar atau ahli dalam mendiagnosis penyakit serta gejala-gejalanya (Raflin Sinaga & Dharma, n.d.).

Identifikasi pada citra telah lama dikembangkan, salah satunya melalui analisis tekstur. Tekstur citra dapat dibedakan berdasarkan beberapa faktor, seperti kerapatan, keseragaman, kekasaran, dan keteraturan. Komputer tidak bisa langsung membedakan tekstur seperti yang dilakukan oleh penglihatan

manusia, oleh karena itu diperlukan analisis tekstur untuk mengetahui pola dari citra digital. Analisis tekstur menghasilkan nilai dari ciri atau karakteristik tekstur yang kemudian dapat dikenali dan diproses oleh komputer untuk masuk ke tahap segmentasi. Segmentasi secara umum dapat diartikan sebagai proses pengelompokan, di mana segmentasi memisahkan objek yang berbeda.

METODE

Flowchart Sistem



Gambar 1 Flowchart Sistem

HASIL DAN PEMBAHASAN

Analisis Data

Dalam analisis data untuk segmentasi penyakit katarak berdasarkan kemiripan bentuk dan warna, langkah pertama adalah pengumpulan data, yaitu dataset yang berisi informasi retina mata untuk pengujian. Pengumpulan data dan analisis kebutuhan saling melengkapi untuk menciptakan sistem pengujian yang presisi dan akurat. Dataset retina mata yang digunakan akan diproses dengan algoritma K-Means dalam segmentasi penyakit katarak berdasarkan kemiripan bentuk dan warna (Sulistiyawati & Supriyanto, 2021).

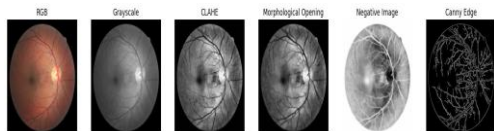
Representasi Data

Berikut adalah langkah-langkah representasi data citra (Suwanda & Juniati, 2022):

1. Gambar *input* dibaca dalam format RGB yang memiliki tiga kanal warna, yaitu merah (R), hijau (G), dan biru (B).
2. Ekstraksi nilai intensitas warna: Setiap piksel gambar terdiri dari tiga nilai intensitas warna R, G, dan B dengan rentang 0 hingga 255.

Hitung Kecerahan: Untuk setiap piksel, hitung nilai kecerahan dengan rumus tertentu. Ada beberapa rumus yang dapat digunakan, tetapi yang paling umum adalah rata-rata dari komponen R, G, dan B: $I = (R + G + B) / 3$.

3. Set Piksel *Grayscale*: Hasil perhitungan intensitas di setiap piksel menghasilkan gambar satu kanal yang hanya berisi nilai tingkat kecerahan.
4. Proses CLAHE (*Contrast Limited Adaptive Histogram Equalization*) adalah sebuah teknik pengolahan citra yang digunakan untuk meningkatkan kontras dan kualitas citra.
5. Proses *Morphological Opening* adalah sebuah operasi pengolahan citra yang digunakan untuk menghilangkan noise dan memperbaiki kualitas citra
6. Proses *negative image* untuk mengubah kontras citra, sehingga membuat objek yang sebelumnya tidak terlihat menjadi terlihat.
7. Proses deteksi dengan tepi canny: untuk mendeteksi tepi garis-garis pada citra retina. Pada proses ini ditentukan nilai threshold bawah yaitu 50 dan threshold atas 150. Hal ini bertujuan untuk menghilangkan noise pada citra retina mata dan menghasilkan garis-garis tepi yang lebih detaill.
8. Simpan Gambar: Setelah semua piksel diubah, simpan gambar hasil konversi. Pastikan untuk menggunakan format gambar yang mendukung gambar monokrom atau skala abu-abu, seperti format PNG atau JPEG.



Gambar 2 Proses Representasi Data Citra

Hasil Analisis Data

Dimensi Fraktal *Box Counting*

Data citra dari hasil metode Tepi Canny dalam format *jpg atau *jpeg selanjutnya diubah dalam bentuk nilai piksel *grayscale* dengan rentang nilai 0 - 255. Berikut langkah-langkah perhitungan nilai dimensi fractal dari sample dataset dalam bentuk pixel:

1. Menentukan ukuran kotak: karena ukuran data berukuran 16*16 sehingga ukuran kotak (r) adalah [1,2,4,8,16].
2. Menentukan jumlah setiap ukuran kotak (N(r)) yang memiliki nilai jumlah pixel lebih dari 0.

Retina 1

r = 1, Jumlah N(r) = 202, $\log(1/r) = \log(1/1) = 0.0$, $\log(N(r)) = 2.3054$

r = 2, Jumlah N(r) = 60, $\log(1/r) = \log(1/2) = -0.301$, $\log(N(r)) = 1.7782$

r = 4, Jumlah N(r) = 16, $\log(1/r) = \log(1/4) = -0.6021$, $\log(N(r)) = 1.2041$

r = 8, Jumlah N(r) = 4, $\log(1/r) = \log(1/8) = -0.9031$, $\log(N(r)) = 0.6021$

r = 16, Jumlah N(r) = 1, $\log(1/r) = \log(1/16) = -1.2041$, $\log(N(r)) = 0.0$

3. Menghitung nilai dimensi fraktal dengan rumus

$$\alpha = \frac{n(\sum_{k=1}^n xy) - ((\sum_{k=1}^n x)(\sum_{k=1}^n y))}{n(\sum_{k=1}^n x^2) - (\sum_{k=1}^n x)^2}$$

Dimana:

x = logaritma ukuran kotak $\log(1/r)$

y = logaritma jumlah kotak $\log(N(r))$

Retina 1

$$D = \frac{5 \cdot -1.80394 - (-3.01030 \cdot 5.88968)}{5 \cdot 2.71857 - (-3.01030)^2} = \frac{8.70999}{4.53095} = 1.92233$$

Hasil nilai dimensi fraktal diperoleh dari setiap pengolahan data citra diatas dengan menggunakan metode *box-counting* melalui aplikasi *python* (Fahmi et al., 2021):

Tabel 1 Nilai Dimensi Fraktal Citra Retina Mata Pada Sample Dataset

Nama Dataset	Tipe Penyakit Katarak	Nilai Dimensi Fraktal
Retina 1	Diabetik	1.92233

Retina 2	Diabetik	1.88823
Retina 3	Diabetik	1.91067
Retina 4	Kortikal	1.49823
Retina 5	Kortikal	1.58918
Retina 6	Kortikal	1.62951
Retina 7	Sklerosis	1.77507
Retina 8	Sklerosis	1.79103
Retina 9	Sklerosis	1.67024

Perancangan model K-Means untuk identifikasi penyakit mata katarak (Gustientiedina et al., 2019):

1. Dataset

Tabel 1. merupakan dataset dari tiga jenis penyakit retina mata, yaitu Diabetik, Kortikal, dan Sklerosis yang masing-masing diambil sampel sebanyak 3 gambar dan menghasilkan gambar sebanyak 9 gambar. Total keseluruhan data adalah 330 data citra retina yang terbagi atas 160 katarak diabetes, 106 katarak kortikal, dan 64 katarak sklerosis.

2. Penentuan jumlah cluster dan nilai centroid awal

Pada penelitian ini karena terdapat

3 penyakit sehingga jumlah cluster adalah 3 dengan ketentuan:

Cluster 1 (C1) : Sklerosis

Cluster 2 (C2) : Diabetes

Cluster 3 (C3) : Kortikal

Nilai centroid awal diperoleh dengan menghitung rata-rata nilai dimensi fraktal pada setiap jenis penyakit dan diperoleh nilai:

$$C1 = 1.77507 + 1.79103 + 1.67024 / 3 = 1,74544$$

$$C2 = 1.92233 + 1.88823 + 1.91067 / 3 = 1.90707$$

$$C3 = 1.49823 + 1.58918 + 1.62951 / 3 = 1.57230$$

3. Lakukan perhitungan jarak euclidean pada setiap data terhadap nilai centroid dan identifikasi setiap data berdasarkan jarak cluster terdekat. Contoh perhitungan seperti yang ditunjukkan pada Tabel 2.

Tabel 2 Contoh Perhitungan Jarak Euclidean dan Hasil Clustering Pada Sample Dataset Iterasi Ke 1

No	Dataset	C1 1,74544	C2 1.90707	C3 1.57230	Hasil
1.	Retina 1	$\frac{\sqrt{(1.74544 - 1.92233)^2}}{2} = 0.17689$	$\frac{\sqrt{(1.90707 - 1.92233)^2}}{2} = 0.01526$	$\frac{\sqrt{(1.57230 - 1.92233)^2}}{2} = 0.35003$	C2
2	Retina 2	$\frac{\sqrt{(1.74544 - 1.88823)^2}}{2} = 0.14279$	$\frac{\sqrt{(1.90707 - 1.88823)^2}}{2} = 0.01884$	$\frac{\sqrt{(1.57230 - 1.88823)^2}}{2} = 0.31593$	C2
3	Retina 3	$\frac{\sqrt{(1.74544 - 1.91067)^2}}{2} = 0.16523$	$\frac{\sqrt{(1.90707 - 1.91067)^2}}{2} = 0.0036$	$\frac{\sqrt{(1.57230 - 1.91067)^2}}{2} = 0.33837$	C2
4	Retina 4	$\frac{\sqrt{(1.74544 - 1.49823)^2}}{2} = 0.24721$	$\frac{\sqrt{(1.90707 - 1.49823)^2}}{2} = 0.40884$	$\frac{\sqrt{(1.57230 - 1.49823)^2}}{2} = 0,07407$	C3
5	Retina 5	$\frac{\sqrt{(1.74544 - 1.58918)^2}}{2} = 0.15626$	$\frac{\sqrt{(1.90707 - 1.58918)^2}}{2} = 0.3179$	$\frac{\sqrt{(1.57230 - 1.58918)^2}}{2} = 0.01688$	C3
6	Retina 6	$\frac{\sqrt{(1.74544 - 1.62951)^2}}{2} = 0.11593$	$\frac{\sqrt{(1.90707 - 1.62951)^2}}{2} = 0.27756$	$\frac{\sqrt{(1.57230 - 1.62951)^2}}{2} = 0.05721$	C3
7	Retina 7	$\frac{\sqrt{(1.74544 - 1.77507)^2}}{2} = 0.02963$	$\frac{\sqrt{(1.90707 - 1.77507)^2}}{2} = 0.132$	$\frac{\sqrt{(1.57230 - 1.77507)^2}}{2} = 0.20277$	C1
8	Retina 8	$\frac{\sqrt{(1.74544 - 1.79103)^2}}{2}$	$\frac{\sqrt{(1.90707 - 1.79103)^2}}{2}$	$\frac{\sqrt{(1.57230 - 1.79103)^2}}{2}$	C1

		$\frac{2}{= 0.04559}$	$\frac{2}{= 0.11604}$	$\frac{2}{= 0,21873}$	
9	Retina 9	$\sqrt[2]{(1.74544 - 1.67024)}$ $\frac{2}{= 0.0752}$	$\sqrt[2]{(1.90707 - 1.670)}$ $\frac{2}{= 0.23683}$	$\sqrt[2]{(1.57230 - 1.6)}$ $\frac{2}{= 0.09794}$	C1

Tentukan kembali titik pusat *cluster* yang baru berdasarkan rata-rata *cluster* baru tersebut didapat dari rumus: nilai hasil/banyak hasil.

$$C1 = 1.77507 + 1.79103 + 1.67024 / 3 = 1.74544$$

$$C2 = 1.92233 + 1.88823 + 1.91067 / 3 = 1.90707$$

$$C3 = 1.49823 + 1.58918 + 1.62951 / 3 = 1,57230$$

Lakukan kembali langkah ke 3 hingga titik pusat dari setiap *cluster* tidak berubah.

Tentukan kembali titik pusat *cluster* yang baru berdasarkan rata-rata *cluster* baru tersebut didapat dari rumus : nilai hasil/banyak hasil.

$$C1 = 1.77507 + 1.79103 + 1.67024 / 3 = 1.74544$$

$$C2 = 1.92233 + 1.88823 + 1.91067 / 3 = 1.90707$$

$$C3 = 1.49823 + 1.58918 + 1.62951 / 3 = 1,57230$$

Hasil dari proses iterasi ke-2 menunjukkan nilai centroid yang sama dengan iterasi ke-1 sehingga hasil sudah konvergen pada iterasi ke 2. Diperoleh hasil akhir untuk nilai centroid C1 adalah 1.74544, C2 adalah 1.90707, C3 adalah 1,57230.

Tampilan Awal Aplikasi

digunakan untuk melakukan segmentasi penyakit katarak berdasarkan kemiripan warna dan bentuk menggunakan algoritma K-Means. Antarmuka ini dilengkapi dengan tombol untuk menginput citra yang akan diproses menggunakan algoritma K-Means. Selain itu, terdapat tombol untuk melakukan prediksi dan menampilkan proses representasi data



Gambar 3 Tampilan Awal

Tampilan Hasil Prediksi

Gambar di bawah ini menunjukkan antarmuka aplikasi untuk menampilkan hasil prediksi penyakit katarak berdasarkan gambar yang dimasukkan dengan menggunakan algoritma k-means. Terdapat nama gambar yang dimasukkan serta hasil proses prediksi yang menunjukkan penyakit katarak apa gambar tersebut.



Gambar 4 Tampilan Hasil Prediksi

Hasil Pengujian sistem terhadap masing-masing citra retina mata dengan algoritma *K-MeAns*.

Tabel 3. menunjukkan hasil pengujian terhadap masing-masing sample data citra retina mata menggunakan k-means didapatkan secara otomatis dari sistem. Keseluruhan sample dataset menghasilkan identifikasi penyakit katarak yang sesuai dengan label pada dataset.

Tabel 3 Hasil Proses Prediksi Dengan K-Means Pada Sample Dataset

No	Nama Dataset	Hasil Identifikasi Proses K-Means
1	Retina 1	Diabetik (Berhasil)
2	Retina 2	Diabetik (Berhasil)
3	Retina 3	Diabetik (Berhasil)
4	Retina 4	Kortikal (Berhasil)
5	Retina 5	Kortikal (Berhasil)
6	Retina 6	Kortikal (Berhasil)
7	Retina 7	Sklerosis (Berhasil)
8	Retina 8	Sklerosis (Berhasil)
9	Retina 9	Sklerosis (Berhasil)

SIMPULAN

Berdasarkan hasil analisis pada penelitian ini, maka dapat ditarik kesimpulan sebagai berikut:

1. Menganalisis jenis warna penyakit mata katarak berdasarkan kemiripan warna dengan menggunakan proses segmentasi citra hasil pengolahan citra ditampilkan menggunakan program Python. Pertama, citra RGB diubah menjadi citra green channel, selanjutnya dilakukan perbaikan kontras dengan menggunakan CLAHE (Contrast Limited Adaptive Histogram Equalization). Kemudian dilanjutkan dengan proses morphological opening dengan tujuan menghaluskan garis-garis pembuluh darah. Selanjutnya, dilakukan operasi Complement Image (Negative Image) agar objek memiliki nilai 0 (berwarna hitam), sementara latar belakang memiliki nilai 1 (berwarna putih). Selanjutnya, garis-garis objek dideteksi menggunakan metode deteksi tepi Canny.
2. Identifikasi penyakit mata katarak berdasarkan hasil segmentasi citra diawali dengan mengekstraksi nilai indeks statistik citra dalam bentuk dimensi fraktal. Selanjutnya, nilai dimensi dilakukan pengelompokan (clustering) menggunakan algoritma K-Means dengan jumlah cluster sebanyak tiga. Hasil clustering menghasilkan nilai centroid untuk setiap cluster, yaitu nilai centroid C1 adalah 1.74544, C2 adalah 1.90707, C3 adalah 1.57230. Identifikasi dilakukan dengan menghitung jarak Euclidean terdekat antara nilai dimensi fraktal citra dengan nilai centroid masing-masing cluster.
3. Sistem identifikasi penyakit mata katarak dibangun menggunakan bahasa pemrograman Python dan HTML. Sistem memanfaatkan model K-Means yang dibuat sebelumnya sebagai fungsi utama dalam mengidentifikasi mata

katarak menggunakan data citra retina mata. Hasilnya menunjukkan sistem ini memiliki akurasi 94.84% dalam mengelompokkan dan melakukan segmentasi citra kedalam 3 cluster yaitu sklerosis, diabetes, dan kortika.

DAFTAR PUSTAKA

- Abdu, S., Saranga', J. L., Sulu, V., & Wahyuni, R. (2021). DAMPAK PENGGUNAAN GADGET TERHADAP PENURUNAN KETAJAMAN PENGLIHATAN. *Jurnal Keperawatan Florence Nightingale*, 4(1), 24–30. <https://doi.org/10.52774/jkfn.v4i1.59>
- Fahmi, R. N., Jajuli, M., Sulistiyowati, N., & Karawang, U. S. (2021). ANALISIS PEMETAAN TINGKAT KRIMINALITAS DI KABUPATEN KARAWANG MENGGUNAKAN ALGORITMA K-MEANS. *Journal of Information Technology and Computer Science (INTECOMS)*, 4(1). www.pasundanekspres.co
- Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 5(1), 17–24. <https://doi.org/10.25077/teknosi.v5i1.2019.17-24>
- Nugraha, M., Purwanto, T., Ai meri yulianti, & Nurlaili Dwi Hidayati. (2024). TINGKAT PENGETAHUAN PASIEN PRA-OPERASI KATARAK DI RSUD SINGAPARNA MEDIKA CITRAUTAMA. *JILOP (Journal of Indonesia Optometrists)*, 2, 1–11. <https://jurnaloptometris.org/index.php/JILOP>
- Raflin Sinaga, M. B., & Dharma, S. (n.d.). *Volume 2 Nomor 10 Oktober 2023 PENGARUH KENDALI GLUKOSA DARAH, HIPERTENSI, DAN DISLIPIDEMIA TERHADAP*

-
- KOMPLIKASI RETINOPATI DIABETIK PADA PENDERITA DIABETES MELITUS TIPE 2.*
<https://jmi.rivierapublishing.id/index.php/rp>
- Raflin Sinaga, M. B., Dharma, S., & Yensuari. (2023). PENGARUH KENDALI GLUKOSA DARAH, HIPERTENSI, DAN DISLIPIDEMIA TERHADAP KOMPLIKASI RETINOPATI DIABETIK PADA PENDERITA DIABETES MELITUS TIPE 2. *Jurnal Multidisiplin Indonesia*, 2(10), 3304–3319.
- <https://jmi.rivierapublishing.id/index.php/rp>
- Sulistiyawati, A., & Supriyanto, E. (2021). Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan. *Jurnal TEKNO KOMPAK*, 15(2), 25–36.
- Suwanda, A. E., & Juniati, D. (2022). Klasifikasi Penyakit Mata Berdasarkan Citra Fundus Retina Menggunakan Dimensi Fraktal Box Counting Dan Fuzzy K-Means. *Jurnal Penelitian Matematika Dan Pendidikan Matematika*, 5(1), 10–18.