

# Menerapkan K-Means Clustering untuk Segmentasi Gambar Database Berwarna

Aryo De Wibowo Muhammad Sidik  
Electrical Engineering  
Nusa Putra University  
Sukabumi, Indonesia  
aryo.dewibowo@nusaputra.ac.id

Edwinanto  
Electrical Engineering  
Nusa Putra University  
Sukabumi, Indonesia  
edwinanto@nusaputra.ac.id

Yufriana Imamulhak  
Electrical Engineering  
Nusa Putra University  
Sukabumi, Indonesia  
yufriana.imamulhak@nusaputra.ac.id

Ilman Himawan Kusumah  
Electrical Engineering  
Nusa Putra University  
Sukabumi, Indonesia  
Ilman.himawan@nusaputra.ac.id

Marina Artiyasa  
Electrical Engineering  
Nusa Putra University  
Sukabumi, Indonesia  
marina@nusaputra.ac.id

Yudha Putra  
Electrical Engineering  
Nusa Putra University  
Sukabumi, Indonesia  
yudha.putra@nusaputra.ac.id

Anang Suryana  
Electrical Engineering  
Nusa Putra University  
Sukabumi, Indonesia  
anang.suryana@nusaputra.ac.id

Anggy Pradiftha Junfithrana  
Electrical Engineering  
Nusa Putra University  
Sukabumi, Indonesia  
anggy.pradiftha@nusaputra.ac.id

**Abstrak**—Segmentasi gambar sangat penting dalam pendekatan analisis gambar untuk mempelajari gambar apa pun. Teknik pengelompokan K-means adalah algoritma yang banyak digunakan dalam sistem segmentasi citra. Karya ini memanfaatkan ruang warna  $L^*a^*b^*$  dan pengelompokan K-means untuk menawarkan segmentasi gambar berbasis warna. Penelitian ini menunjukkan segmentasi gambar database berdasarkan karakteristik warna menggunakan teknik K-means clustering unsupervised yang diimplementasikan dengan coding MATLAB. Seluruh pekerjaan dipisahkan menjadi dua fase. Pertama, augmentasi pemisahan warna dari gambar warna database ditingkatkan melalui peregangan de-korelasi. Kemudian keenam area citra database tersebut dikelompokkan menjadi tiga kelompok dengan menggunakan teknik K-means clustering. Dengan menggunakan ruang warna  $L^*a^*b^*$  dan metode pengelompokan k-means dalam gambar basis data warna, kita hanya dapat menunjukkan wilayah tengah dari gambar apa pun. Kami dapat mengisolasi daerah yang terkontaminasi dalam gambar berwarna database medis dengan pendekatan ini dan mengobati penyakitnya dengan cepat. Kami dapat menggunakan berbagai pendekatan seperti Particle swarm Optimization (PSO) untuk hasil yang lebih baik.

**Kata Kunci**—segmentasi citra, de-korelasi, k-means, basis data citra warna.

## I. PENDAHULUAN

Segmentasi gambar sering dicirikan sebagai "pencarian area homogen dalam sebuah gambar, diikuti dengan kategorisasinya." Selain itu, melibatkan membagi gambar menjadi bagian-bagian yang relevan menurut kriteria homogenitas atau heterogenitas. Pendekatan segmentasi berorientasi piksel, Berorientasi kontur, Berorientasi kawasan, Berorientasi model, Berorientasi warna, dan Hibrida dapat dibedakan dengan konsep dasar berikut: berorientasi kontur, berorientasi kawasan, berorientasi model, berorientasi warna, dan hibrida. Segmentasi warna gambar sangat penting untuk analisis gambar dan berbagai

visi komputer, interpretasi gambar, dan sistem pengenalan pola, dengan aplikasi dalam medis, penginderaan jauh, mikroskop, pengambilan gambar dan video berbasis konten, analisis dokumen, otomasi industri, dan kontrol kualitas bidang. Kinerja segmentasi warna secara substansial dapat mempengaruhi kualitas sistem pemahaman gambar. Tekstur, bentuk, intensitas tingkat keabuan, dan warna adalah karakteristik paling khas yang digunakan dalam segmentasi gambar. Masalah khas dalam segmentasi adalah pembentukan ruang data yang sesuai. Hal ini diperlukan untuk mencari karakteristik yang cukup menggambarkan proses yang sebenarnya untuk membuat pengklasifikasi realistis. Prosedur penilaian segmentasi umumnya terbagi menjadi dua jenis (diawasi dan tidak diawasi). Kategori pertama tidak berlaku untuk penginderaan jauh karena sulit untuk mendapatkan segmentasi yang optimal (segmentasi ground truth).

Selain itu, prosedur penilaian segmentasi untuk data penginderaan jauh belum diperiksa secara memadai. Akibatnya, dimungkinkan untuk melanjutkan proses klasifikasi untuk alasan perbandingan dan kemudian secara tidak langsung mengevaluasi proses segmentasi berdasarkan akurasi klasifikasi yang diperoleh. Untuk klasifikasi citra berbasis segmen, citra yang akan dikategorikan dibagi menjadi beberapa daerah homogen yang mengandung informasi spektrum yang sebanding. Kemudian, fitur segmen gambar diambil tergantung pada kriteria kategorisasi fitur tanah [1].

Homogenitas warna ditentukan oleh standar deviasi spektrum hue, sedangkan kekompakan dan kehalusan bentuk menentukan homogenitas bentuk. Ada dua ide yang terlibat dalam iterasi parameter: 1) Selain tingkat kehalusan yang diperlukan, kita harus memilih nilai skala yang sama pentingnya dengan kelayakan untuk membedakan di seluruh area; 2) Kriteria warna harus digunakan bila memungkinkan. Pentingnya informasi spektral dalam data

gambar akan mengurangi kualitas segmentasi jika kriteria bentuk memiliki bobot yang tinggi [2].

Studi ini memperkenalkan teknik segmentasi citra yang unik berdasarkan karakteristik warna gambar. Kami tidak menggunakan data pelatihan untuk ini, dan pekerjaan dipisahkan menjadi dua langkah. Pertama, pemisahan warna dari citra warna database dilakukan melalui peregangannya de-korelasi. Setelah melakukan teknik K-means clustering, area citra dibagi menjadi tiga kelompok berdasarkan warna yang ada. Namun, kami mendapatkan dua cluster ketika kami menerapkan pengelompokan K-means ke database MR [2], [3]. Dengan menggunakan metode dua langkah ini, adalah mungkin untuk memotong biaya komputasi dengan melewati perhitungan fitur untuk setiap piksel dalam gambar. Meskipun warna tidak sering digunakan untuk segmentasi gambar, namun memberikan tingkat diskriminasi yang tinggi pada area gambar.

Proyek ini disusun sebagai berikut: Bagian II menjelaskan peregangannya De-korelasi yang disarankan, teknik pengelompokan k-means dasar, dan segmentasi citra basis data warna menggunakan algoritma pengelompokan k-means. Bagian III menyajikan temuan eksperimental yang diperoleh dengan menggunakan strategi yang direkomendasikan. Bagian IV diakhiri dengan beberapa manfaat dan pengamatan. [7]

## II. ALGORITMA YANG DIUSULKAN

### A. Peregangannya De-korelasi

Peregangannya de-korelasi meningkatkan pemisahan warna dalam gambar dengan korelasi band-to-band yang cukup besar. Warna tinggi memfasilitasi persepsi visual dan memfasilitasi diferensiasi fitur. Fungsi peregangannya de-korelasi diimplementasikan. Tiga pita warna diekstraksi dari gambar. Namun, peregangannya de-korelasi dapat digunakan terlepas dari jumlah pita warna. Pertama, nilai warna asli gambar diterjemahkan ke rangkaian nilai baru dengan rentang yang lebih baik. Intensitas warna setiap piksel diterjemahkan ke dalam ruang eigen warna dari matriks korelasi, diregangkannya untuk menyeimbangkan varian pita, dan kemudian ditransformasikan kembali ke pita warna asli. Kami dapat menggunakan gambar asli lengkap atau subset tertentu untuk membuat statistik bijaksana pita dengan opsi subset.

### B. K-means clustering

K-means clustering adalah salah satu dari banyak teknik clustering data. Pengelompokan data mendeteksi kelompok atau klaster alami dalam data multidimensi menggunakan ukuran kesamaan (misalnya, jarak Euclidean) [Jain et al. 1999; Jain dkk. 2000]. Ini adalah langkah penting dalam kategorisasi gambar. Tiga bentuk pengelompokan data adalah pengelompokan hierarkis, pengelompokan partisi, dan pengelompokan spektral. Ada tiga bentuk clustering partisi. Teknik clustering meliputi K-means, Fuzzy C-means, dan quality threshold clustering. Hanya di sini k-means clustering akan dibahas. K-means clustering adalah teknik klasifikasi citra tanpa pengawasan. K-means clustering, juga dikenal sebagai C-means clustering, telah diterapkan di beberapa bidang, termasuk kompresi data gambar dan audio. Algoritma K-means meminimalkan jumlah jarak kuadrat antara setiap titik dan pusat cluster. K-

menandakan pengelompokan banyak vektor data menjadi sejumlah cluster yang telah ditentukan [9].

$$X_j = \text{Jumlah Vektor } (j = 1, 2, \dots, n)$$

$$G_i = \text{Jumlah Vektor } (i = 1, 2, \dots, c)$$

Dan menentukan pusat dari setiap cluster menggunakan persamaan fungsi equation.

$$J = \sum_{i=1}^c j_i = \sum_{i=1}^c \left( \sum_{k, xk \in G_i} \|X_i - c_i\|^2 \right) \quad (1)$$

Dimana  $X_k$  adalah vektor dari grup  $j$  dan  $c_i$  adalah pusat cluster dan Dimana,

$$j_i = \sum_{k, xk \in G_i} \|X_i - c_i\|^2 \text{ adalah fungsi biaya dalam grup } i.$$

Dengan demikian, nilai  $j_i$  tergantung pada sifat geometris  $G_i$  dan lokasi  $c_i$ .

Langkah 1: Inisialisasi pusat cluster  $c_i$ ,  $i = 1, \dots, c$ . Ini biasanya dicapai dengan memilih titik  $c$  secara acak dari semua titik data.

Langkah 2: Tentukan matriks keanggotaan  $U_{ij}$  dengan menggunakan Persamaan

$$U_{ij} = \{1 \text{ if } \|X_j - c_i\|^2 < \|X_j - c_k\|^2, \text{ untuk setiap } k \neq i \text{ jika tidak}\}$$

Langkah 3: Hitung fungsi biaya menurut Persamaan (1). Berhenti jika salah satu di bawah nilai toleransi tertentu peningkatannya atas iterasi sebelumnya di bawah ambang batas tertentu.

Langkah 4: Perbarui pusat cluster sesuai dengan Persamaan

$$c_i = \frac{1}{|G_i|} = \sum_{k, xk \in G_i} X_k$$

Dimana  $|G_i|$  adalah ukuran dari  $G_i$ , atau  $|G_i| = \sum_{i=1}^n U_{ij}$

Pergi ke langkah 2.

C. Segmentasi citra warna database mengimplementasikan k-means clustering

- 1) Baca database warna gambar satu per satu.
- 2) Untuk pemisahan warna gambar, terapkan peregangannya relasi Dekorasi.
- 3) Sekarang, konversikan satu per satu gambar warna database dari warna RGB ke warna  $L^*a^*b^*$ . Ruang  $L^*a^*b^*$  terdiri dari lapisan luminositas ' $L^*$ ', lapisan kromatisitas ' $a^*$ ' yang menunjukkan tempat jatuhnya warna di sepanjang sumbu merah-hijau, dan lapisan kromatisitas ' $b^*$ ' yang menunjukkan tempat jatuhnya warna bersama dengan sumbu biru-kuning. Semua informasi warna ada di lapisan ' $a^*$ ' dan ' $b^*$ '. Kita dapat mengukur perbedaan antara dua warna menggunakan metrik jarak Euclidean. Kemudian, ubah gambar menjadi ruang warna  $L^*a^*b^*$ .
- 4) Menerapkan pengelompokan k-means, berdasarkan jarak Euclidean, mengklasifikasikan warna setiap gambar dalam ' $a^*b^*$ '.

- Menggunakan k-means result melabeli setiap piksel pada citra. Kami akan mendapatkan tiga kelompok gambar tersegmentasi Hitam, Putih dan abu-abu pada output, di mana satu membuat gambar database tersegmentasi berdasarkan warna.

### III. HASIL DAN ANALISIS

Dalam penelitian ini, dengan memanfaatkan MATLAB v7.8 [10], kami menunjukkan area wajah dari gambar database warna. Foto-foto itu berbeda satu sama lain menurut sudut pandangnya. Seluruh metode dan hasil yang diprediksi dirinci di bawah ini [4], [5].

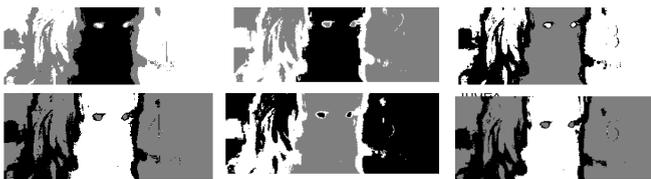


Gambar 1. Gambar Database Asli



Gambar 2. Gambar Database  $L^*a^*b^*$

Gambar diubah menjadi gambar ruang warna  $L^*a^*b^*$  sesuai dengan ilustrasi. Jika kita mengabaikan kecerahan gambar asli, kita dapat dengan cepat mengidentifikasi warna  $L^*a^*b^*$ , yaitu biru, merah muda, dan putih. Di sini, ruang  $L^*a^*b^*$  terdiri dari lapisan luminosititas ' $L^*$ ', lapisan kromatisitas ' $a^*$ ' yang menunjukkan letak warna di sepanjang sumbu merah-hijau, dan lapisan kromatisitas ' $b^*$ ' yang menunjukkan tempat warna terletak di sepanjang sumbu biru-kuning [2].



Gambar 3. Gambar Dilabeli oleh indeks cluster



Gambar 4. Tampilkan wilayah warna hitam dari gambar indeks cluster sebagai Cluster 1

Dari gambar di atas kita mengetahui bahwa gambar diberi label sesuai dengan piksel yang ada dalam gambar dan memberi label pada piksel tersebut dengan menggunakan algoritma pengelompokan K-means yang didasarkan pada konsep jarak Euclidean.



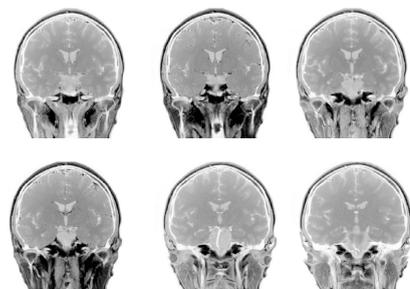
Gambar 5. Tampilkan wilayah warna hitam dari gambar indeks cluster sebagai Cluster 2



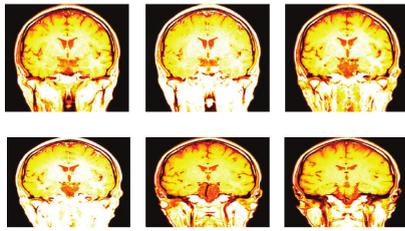
Gambar 6. Tampilkan wilayah warna hitam dari gambar indeks cluster sebagai Cluster 3

Gambar 4 hanya menampilkan area wajah gambar database 4 dan 6 karena hanya menampilkan bagian warna Hitam dari gambar indeks cluster berlabel dan menggunakan cluster1 sebagai referensi. Gambar 5 hanya menggambarkan area wajah gambar database 3 dan 5 karena hanya menampilkan bagian abu-abu dari citra indeks cluster berlabel dan mengklasifikasikannya sebagai cluster2. Gambar 6 hanya menampilkan area wajah gambar database 1 dan 2 karena hanya menampilkan bagian putih dari gambar indeks cluster berlabel dan menganggap ini sebagai cluster3.

Oleh karena itu, berdasarkan penjelasan hasil sebelumnya, kami telah menentukan bahwa gambar database telah diklasifikasikan menjadi tiga cluster: hitam, abu-abu, dan putih. Sesuai dengan gambar aslinya, gambar clustering menyorot area wajah dari gambar berwarna di database. Karena menggunakan algoritma pengelompokan K-means, hasil dari setiap eksekusi program berbeda. Dengan demikian, setiap siklus menghasilkan hasil yang unik. Sekarang, teknik clustering k-means telah diterapkan dalam bentuk database yang berbeda. Kami mendapatkan gambar Magnetic Resonance Imaging (MRI) abu-abu. Kami mengubah gambar abu-abu menjadi warna RGB dan mengekstrak area kanker dari gambar otak di database. Kami menerapkan teknik pengelompokan k-means pada gambar database warna abu-abu MR dan mengubahnya menjadi gambar berwarna karena gambar MR berwarna tidak sering digunakan.

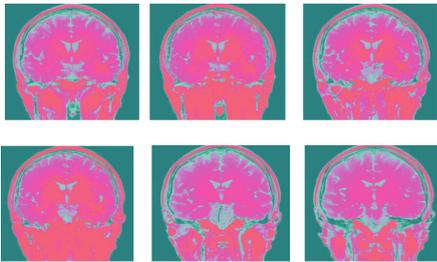


Gambar 7. Gambar otak MRI Database Asli

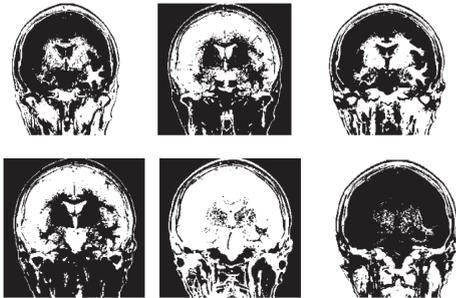


Gambar 8. Gambar yang dikonversi Warna

Dari gambar. 7 kami menemukan citra warna yang dikonversi dari basis data citra otak MRI.

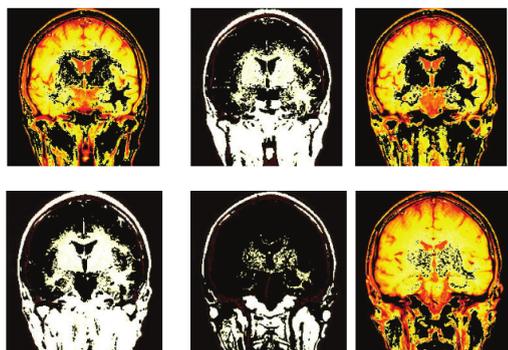


Gambar 9. L\*a\*b\* Database Citra Otak MRI

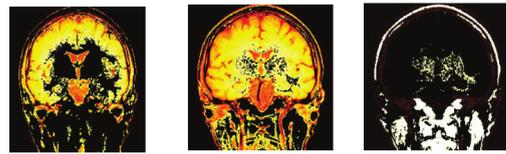
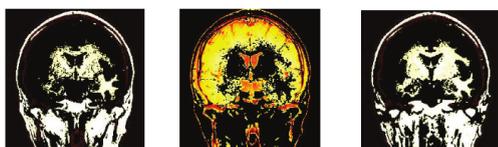


Gambar 10. Gambar Dilabeli oleh Indeks Cluster

Dari Gambar 9 diketahui bahwa citra diubah menjadi citra ruang warna L\*a\*b, ruang warna L\*a\*b sudah dibahas pada Gambar 1 (b). Dari Gambar 10, kita mengetahui bahwa gambar diberi label sesuai dengan piksel yang ada dalam gambar dan memberi label pada piksel tersebut dengan menggunakan algoritma pengelompokan K-means.



Gambar 11. Tampilkan wilayah warna hitam dari gambar indeks cluster sebagai Cluster 1



Gambar 12. Tampilkan wilayah warna hitam dari gambar indeks cluster sebagai Cluster 2

Gambar 11 terutama menggambarkan area sel kanker yang terinfeksi dalam gambar basis data 1,3, dan 6 karena hanya menampilkan bagian hitam dari gambar indeks klaster berlabel, yang ditafsirkan sebagai klaster1. Di sisi lain, karena hanya bagian putih dari gambar indeks klaster berlabel yang ditampilkan dan klaster2 diasumsikan, Gambar 12 terutama menggambarkan wilayah sel kanker yang terinfeksi dalam gambar basis data 2,4 dan 5.

Menurut uraian sebelumnya, dengan mengubah gambar MR skala abu-abu menjadi gambar berwarna, dimungkinkan untuk mengidentifikasi sel kanker otak dan daerah saraf kritis yang terkait dengan otak. Selanjutnya, jika kita mengambil seluruh database gambar otak, yang diambil dari berbagai sudut, akan lebih mudah untuk mengidentifikasi area otak di mana sel kanker telah menyebar.

#### IV. KESIMPULAN

Hal ini layak untuk memotong biaya komputasi menggunakan segmentasi gambar berbasis warna dengan melewati perhitungan fitur untuk setiap piksel gambar. Meskipun warna jarang digunakan untuk segmentasi gambar, mereka menyediakan area gambar dengan tingkat diskriminasi yang tinggi. Seperti yang dinyatakan sebelumnya, dengan melakukan teknik yang digambarkan pada gambar pertama, area wajah dari gambar database diidentifikasi. Ketika metode yang sama digunakan untuk gambar sel kanker database MRI, hasil ini memisahkan daerah yang terinfeksi dari sel kanker dari gambar sel kanker database warna dikonversi, yang fotonya berbeda dalam perspektif mereka dari sel kanker seseorang. Meskipun percobaan ini menggunakan perangkat lunak MATLAB untuk mengatasi tantangannya secara efisien, pekerjaan ini diselesaikan dengan biaya rendah. Untuk hasil yang lebih akurat, kita dapat menggunakan metode optimasi seperti PSO (Particle swarm Optimization) dengan algoritma clustering k-means, yang dikenal sebagai pendekatan clustering Hybrid K-means +PSO [3]. Pemindaian MR tidak memberikan gambaran yang baik tentang area sel kanker.

#### DAFTAR PUSTAKA

- [1] Image Segmentation for the Purpose Of Object-Based Classification, by Ahmed Darwish 2003 IEEE, pp 2039-2041
- [2] Colour-based image segmentation using k-means clustering by Anil Z chitade published in IJEST, vol. 2,2010.
- [3] Brain tumor detection of MR images based on colour converted hybrid PSO+Kmeans clustering segmentation published in European journal of scientific research vol.70(2012).
- [4] Database image classification using k-means clustering & PSO by Madhusmita Sahu published in IJSAA journal, vol.2, May 2012.
- [5] A colour image segmentation algorithm based on region growing by Jun Tang IEEE, vol 6, 2010, pp 634-637.
- [6] Evaluation of colour image segmentation hierarchies by Darren MacDonald Proceeding of the 3rd Canadian conference on computer and Robot Vision, IEEE, 2006.
- [7] Digital image processing by R.C.Gonzalez and R. E. Woods, chapter 10, paper 567-630 2nd edition.

- [8] Image Segmentation via Adaptive –Mean Clustering and Knowledge-Based Morphological Operations with Biomedical Applications by Chang Wen Chen, Jiebo Luo, and Kevin J. Parker, published in IEEE transaction on IMAGE PROCESSING, VOL. 7, NO. 12, DECEMBER 1998.
- [9] Neuro-Fuzzy and soft computing Textbook by jyh-shubg Roger jang, page no-450.
- [10] R. C. Gonzalez and R. E. Woods, “Digital Image Processing using MATLAB”, 2nd edition.
- [11] [www.imageprocessingplace.com](http://www.imageprocessingplace.com)
- [12] [www.mathworks.com](http://www.mathworks.com)
- [13] [www.wikipedia.org](http://www.wikipedia.org)
- [14] <http://ieeexplore.ieee.org/Xplore/guesthome.js>
- [15] <http://www.sciencedirect.com>