

USULAN TUGAS AKHIR

1. IDENTITAS PENGUSUL

NAMA : **HARDIKA KHUSNULIAWATI**
NRP : **5110100206**
DOSEN WALI : **Ir. R. V. Hari Ginardi, M.Sc., Ph.D.**
DOSEN PEMBIMBING : **1. Arya Yudhi Wijaya, S.Kom., M.Kom.**
2. Rully Soelaiman, S.Kom., M.Kom.

2. JUDUL TUGAS AKHIR

“Implementasi *Fuzzy C-Means* dengan Estimasi *Robust Spatial* pada Segmentasi Citra dengan *Noise*”

3. LATAR BELAKANG

Segmentasi citra merupakan salah satu kunci penting untuk proses analisis citra dan pengembangan berbagai aplikasi yang melibatkan pengolahan citra seperti pengenalan pola, deteksi objek, aplikasi citra medis, dan penginderaan jauh atau *remote sensing* [1]. *Fuzzy C-Means* merupakan metode yang sering digunakan untuk segmentasi citra karena termasuk metode yang cukup sederhana, mudah diimplementasikan dan memiliki kemampuan mengumpulkan lebih banyak informasi dibandingkan metode pembagian secara kasar. Akan tetapi algoritma *Fuzzy C-Means* biasa tidak melibatkan informasi spasial pada citra sehingga menyebabkan algoritma ini kurang sensitif terhadap *noise* atau *outliers data*. Padahal citra dengan *noise* merupakan permasalahan yang akan sering dijumpai dalam analisis dan pengolahan citra di dunia nyata. Permasalahan untuk mengatasi *noise* pada citra menjadi hal yang penting karena kualitas dari hasil segmentasi citra bergantung pada kualitas inputan citra. Semakin bebas suatu citra dari *noise* maka hasil segmentasi yang diperoleh pun akan semakin baik. Oleh karena itu, dibutuhkan suatu modifikasi pada algoritma *Fuzzy C-Means* sehingga tetap dapat memberikan hasil segmentasi yang baik meskipun inputan citra yang digunakan merupakan inputan citra dengan *noise*.

Estimasi *Robust Spatial* merupakan metode yang digunakan untuk mengumpulkan informasi spasial dari citra sehingga dapat menyaring data dari *noise* dengan pendekatan statistik [2]. Estimasi yang bersifat *robust* memiliki kesensitifan yang tinggi terhadap suatu penyimpangan seperti adanya *outlier* yang menyebabkan *noise*. Estimasi *Robust Spatial* yang diusulkan merupakan kombinasi dari *RM-Estimator* dengan *L-Filter* yang kemudian disebut dengan *RM-L-*

Estimator. Dengan *RM-L-Estimator* yang ditambahkan pada algoritma *Fuzzy C-Means* diharapkan mampu memberikan hasil segmentasi yang baik terhadap inputan citra dengan *noise*.

4. RUMUSAN MASALAH

Rumusan masalah untuk tugas akhir ini adalah sebagai berikut.

1. Merancang metode estimasi *Robust Spatial* untuk meningkatkan kemampuan dari algoritma *Fuzzy C-Means* dalam segmentasi citra dengan *noise*.
2. Mengimplementasikan estimasi *Robust Spatial* pada algoritma *Fuzzy C-Means* untuk meningkatkan kemampuan dalam segmentasi citra dengan *noise*.
3. Melakukan segmentasi citra menggunakan algoritma *Fuzzy C-Means* yang dimodifikasi dengan penambahan metode estimasi *Robust Spatial* untuk meningkatkan kemampuan dalam segmentasi citra dengan *noise*.

5. BATASAN MASALAH

Permasalahan yang dibahas dalam tugas akhir ini memiliki beberapa batasan, yaitu sebagai berikut.

1. Sistem perangkat lunak dibangun dengan menggunakan perangkat lunak MATLAB.
2. Segmentasi dilakukan pada kumpulan gambar dari *Barkeley image segmentation* dengan simulasi kondisi *noise* meliputi *salt and pepper*, *Gaussian*, dan *speckle noise*.
3. Segmentasi dilakukan pada gambar pada ruang warna *RGB* dan *CIELAB*.
4. Kinerja hasil uji coba dihitung dengan *Probabilistic Rand Index (PRI)* dan *Variation of Information (VOI)*.

6. TUJUAN PEMBUATAN TUGAS AKHIR

Tujuan dari tugas akhir ini adalah sebagai berikut.

1. Memahami konsep estimasi *Robust Spatial* yang merupakan metode untuk meningkatkan kemampuan dari algoritma *Fuzzy C-Means* dalam segmentasi citra dengan *noise*.
2. Mengimplementasikan algoritma *Fuzzy C-Means* dengan estimasi *Robust Spatial* yang dapat digunakan dalam segmentasi citra dengan *noise*.
3. Mengetahui kinerja algoritma *Fuzzy C-Means* dengan estimasi *Robust Spatial* dalam melakukan segmentasi citra dengan *noise* melalui uji coba dan evaluasi.

7. MANFAAT TUGAS AKHIR

Tugas akhir ini dibuat dengan harapan dapat memberikan manfaat di bidang informatika dalam memberikan metode gabungan yang memiliki ketahanan dan efektif untuk segmentasi citra dengan *noise* sehingga hasil dari analisis citra dan aplikasi yang melibatkan pengolahan citra yang dihasilkan dapat lebih baik.

8. TINJAUAN PUSTAKA

Fuzzy C-Means merupakan salah satu algoritma *clustering* yang sering digunakan dalam segmentasi citra. Perbedaan mendasar dari algoritma *Fuzzy C-Means* dengan algoritma *clustering* lainnya yaitu nilai keanggotaan yang diukur dalam rentang nilai $[0,1]$ sehingga suatu data dapat berada pada lebih dari satu *cluster*. Melalui pengukuran keanggotaan seperti itu, algoritma tersebut memberikan model data yang lebih detail dimana nilai ambigu suatu data yang *overlapping* terhadap *cluster* yang ada dapat diketahui [3]. Jika diberikan sekumpulan data $X = \{x_1, \dots, x_n\}$ dengan jumlah data n dimana setiap data x_k direpresentasikan dalam sejumlah p fitur. Pembagian X ke dalam c *cluster* adalah kumpulan dari *mutually disjoint subset* X_i dari X dimana $X_i \cup \dots \cup X_c = X$ (gabungan dari X_i hingga X_c membentuk X) dan $X_i \cap X_j = \emptyset$ dimana setiap $i \neq j$ (tidak ada anggota dari X_i dan X_j yang saling beririsan).

Pembagian data dapat direpresentasikan ke dalam *hard partition matrices* U dengan ukuran $(c \times n)$ dimana c menyatakan jumlah *cluster* dan n menyatakan jumlah data [4]. Untuk nilai keanggotaan dari data x_k dinyatakan dengan $u_{ik} = 1$ jika $x_k \in X_i$ dan $u_{ik} = 0$ jika x_k bukan anggota X_i . Sedangkan untuk algoritma *Fuzzy C-means*, nilai u_{ik} berada pada rentang $[0,1]$ ($u_{ik} \in [0,1]$). Pada algoritma *Fuzzy C-Means*, suatu titik data memiliki derajat keanggotaan yang semakin besar apabila jarak titik data tersebut semakin dekat dengan pusat *cluster* sehingga Persamaan 1 yang dinyatakan dengan fungsi J menjadi optimum jika hasil yang diperoleh minimum.

$$J = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m d_{ik}^2 \quad (1)$$

Parameter m disebut sebagai *fuzzifier* atau *weighting exponent* dengan nilai $m > 1$. Semakin besar nilai m maka batas pembeda antar *cluster* menjadi semakin tidak terlihat, sebaliknya semakin rendah nilai m maka batas pembeda antar *cluster* semakin jelas. Variabel d menunjukkan jarak dari data x_k terhadap pusat *cluster* c_i dengan persamaan penghitungan jarak yang digunakan yaitu *Euclidean distance*, $d_{ik} = ||x_k - c_i||^2$. Untuk memperoleh pembagian data x_k ke dalam *cluster* yang optimum maka dilakukan perulangan dimana setiap kali perulangan nilai u_{ik} dan c_i diperbarui hingga fungsi optimum J diperoleh. Persamaan 2 merupakan fungsi untuk memperbarui derajat keanggotaan u_{ik} sedangkan Persamaan 3 merupakan fungsi untuk memperbarui pusat *cluster* c_i .

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}^2}{d_{jk}^2} \right)^{\frac{1}{m-1}}} = \frac{d_{ik}^{\frac{-2}{m-1}}}{\sum_{j=1}^c d_{jk}^{\frac{-2}{m-1}}} \quad (2)$$

$$c_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m} \quad (3)$$

Kelemahan algoritma *Fuzzy C-Means* yang kurang sensitif terhadap *noise* karena tidak melibatkan informasi spasial dari citra dapat diperbaiki dengan estimasi *Robust Spatial* yaitu metode *RM-L-Estimator* [2]. Estimasi *Robust Spatial* dengan *RM-L-Estimator* merupakan kombinasi dari *RM-Estimator* dan *L-Filter* dimana *RM-Estimator* dan *L-Filter* adalah jenis dari metode penyaringan data dari *noise*. *RM-Estimator* diperoleh dari *M-Estimator* yang dikombinasikan dengan penghitungan nilai tengah untuk memperoleh standar yang cukup dalam menolak *noise* [5]. Metode penyaringan untuk menghapus *noise* terbagi menjadi beberapa kategori yaitu *Linear Filter*, *Nonlinear Filter*, dan *Rank Based Filter*. *Linear Filter* efisien untuk menghapus *Gaussian Noise* akan tetapi sering mendistorsi tepi dan lemah dalam mengatasi *noise* yang bersifat impulsif. *Nonlinear Filter* memiliki kemampuan yang lebih baik dalam mengatasi *noise* dengan kondisi berbeda termasuk *noise* yang bersifat impulsif. *Rank Based Filter* mendapatkan perhatian lebih karena kemampuan yang baik dalam menolak *outlier* yang menyebabkan *noise* tanpa mempengaruhi sifat-sifat asli dari data. *R-estimator* dengan *median filter* dan *L-filter* merupakan contoh dari *Rank Based Filter*. Metode penyaringan tersebut memiliki kinerja yang baik dan dapat digunakan dalam berbagai aplikasi pemrosesan citra.

Citra hasil segmentasi diuji nilai akurasi dengan *Probabilistic Rand Index* (PRI) yang merupakan metode perluasan dari *Rand Index* [6]. PRI memungkinkan perbandingan hasil segmentasi dari citra yang diuji dengan beberapa *ground truth* melalui pembobotan *nonuniform* sederhana dari pasangan piksel sebagai fungsi untuk variabilitas *ground truth*. Diketahui himpunan hasil segmentasi secara manual (*ground truth*) dari gambar $X = \{x_1, \dots, x_N\}$ dengan sejumlah N piksel sebagai $\{S_1, \dots, S_K\}$. Sedangkan S_{TEST} sebagai hasil segmentasi dari citra yang diuji. Setiap piksel x_i diberi label $l_i^{S_{TEST}}$ pada S_{TEST} dan label $l_i^{S_K}$ pada S_K . Untuk memperoleh nilai PRI dibentuk persamaan sebagaimana Persamaan 4. Nilai c_{ij} merepresentasikan nilai biner piksel x_i dan x_j untuk nilai label $l_i^{S_{TEST}}$ dan $l_j^{S_{TEST}}$. Sedangkan nilai p_{ij} merupakan himpunan bilangan random mengikuti distribusi *bernoulli* yang menyatakan persepsi segmentasi secara benar dari citra X .

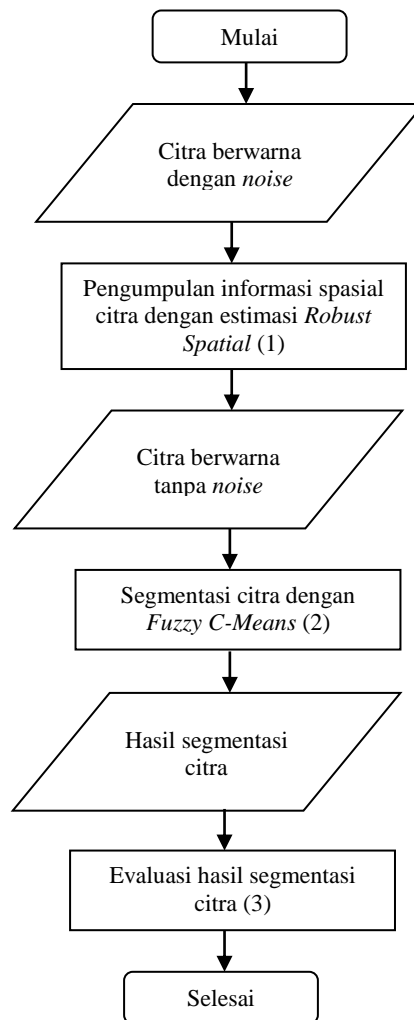
$$PR(S_{TEST}, \{S_K\}) = \frac{1}{\binom{N}{2}} \sum_{i < j} [p_{ij}^{c_{ij}} (1 - p_{ij})^{1-c_{ij}}] \quad (4)$$

Hasil dari PRI bernilai $[0,1]$ dimana nilai 0. menyatakan bahwa hasil segmentasi $\{S_1, \dots, S_K\}$ dan S_{TEST} tidak memiliki kesamaan sedangkan nilai 1 menyatakan keduanya identik. Metode pengujian lain yang digunakan adalah *Variation of Information* (VOI) yang dihitung berdasarkan jarak antara dua *cluster* C dan C' dalam hal perbedaan informasi antara keduanya [7]. VOI dinyatakan dalam Persamaan 5. Nilai H menyatakan nilai entropy dan nilai I menyatakan informasi yang saling beririsan antara dua *cluster* tersebut. Semakin kecil nilai VOI maka hasil segmnetasi semakin mendekati kebenaran

$$d_{VI}(C, C') = H(C) + H(C') - 2I(C, C') \quad (5)$$

9. RINGKASAN ISI TUGAS AKHIR

Langkah-langkah untuk segmentasi citra dengan *noise* menggunakan algoritma *Fuzzy C-Means* dengan estimasi *Robust Spatial* meliputi beberapa tahap yang ditampilkan seperti pada Gambar 1. Tahap pertama mengumpulkan informasi spasial dari citra yang memiliki *noise* menggunakan estimasi *Robust Spatial RM-L-Estimator*. Dari tahap pertama diperoleh setiap piksel citra yang telah tersaring atau terbebas dari *noise*. Tahap kedua melakukan segmentasi citra menggunakan algoritma *Fuzzy C-Means* dimana *dataset* yang akan dikelompokkan merupakan piksel citra yang telah tersaring atau terbebas dari *noise*. Tahap terakhir merupakan tahap evaluasi dengan penghitungan kinerja menggunakan *Probabilistic Rand Index* (PRI) dan *Variation of Information* (VOI). *Probabilistic Rand Index* (PRI) digunakan sebagai evaluasi dari akurasi segmentasi citra yang diuji sedangkan *Variation of Information* (VOI) digunakan sebagai validasi dari hasil pengelompokan.



Gambar 1. Diagram Alur Keseluruhan Sistem

10. METODOLOGI

Metode yang akan dilakukan dalam pengerjaan tugas akhir ini memiliki beberapa tahapan, yaitu sebagai berikut.

1. Penyusunan Proposal Tugas Akhir
Tahap awal yang harus dilakukan untuk memulai pengerjaan tugas akhir adalah penyusunan proposal tugas akhir. Dalam proposal tugas akhir yang dibuat, penulis mengajukan gagasan segmentasi terhadap citra dengan *noise* menggunakan algoritma *Fuzzy C-Means* dengan estimasi *robust spatial*.
2. Studi Literatur
Pada tahap studi literatur, penulis melakukan pencarian informasi dan studi literatur yang terkait dengan algoritma yang digunakan. Informasi dan literatur yang digunakan dapat diperoleh dari berbagai sumber, misalnya buku acuan, *internet*, paper utama, dan paper pendukung.
3. Analisis dan Desain Perangkat Lunak
Analisis dan desain merupakan tahapan yang penting dilakukan sebelum pembangunan perangkat lunak (algoritma) dilakukan. Pada tahap ini, penulis menganalisis serta mendesain perangkat lunak (algoritma) yang akan dibangun dengan mengacu pada proposal yang telah dibuat dan hasil studi literatur yang telah dilakukan sebelumnya.
4. Implementasi Perangkat Lunak
Implementasi merupakan tahapan pembangunan perangkat lunak (algoritma) sesuai dengan hasil analisis dan desain yang telah dilakukan sebelumnya. Implementasi dilakukan dengan menggunakan kaskas bantu MATLAB.
5. Pengujian dan Evaluasi
Pada tahap pengujian dan evaluasi, penulis melakukan beberapa percobaan kemudian mengevaluasi hasil percobaan tersebut. Evaluasi dilakukan dengan melihat hasil dari empat indeks kuantitatif yaitu *Probabilistic Rand Index* (PRI) dan *Variation of Information* (VOI).
6. Penyusunan Buku Tugas Akhir
Pada tahap ini dilakukan penyusunan buku tugas akhir. Buku ini berisi dasar teori dan metode yang digunakan dalam tugas akhir serta hasil dari implementasi aplikasi perangkat lunak yang telah dibuat. Sistematika penulisan buku tugas akhir secara garis besar adalah sebagai berikut.
 1. Pendahuluan
 - 1.1 Latar Belakang
 - 1.2 Rumusan Masalah
 - 1.3 Batasan Tugas Akhir
 - 1.4 Tujuan
 - 1.5 Metodologi
 - 1.6 Sistematika Penulisan
 2. Tinjauan Pustaka
 3. Desain dan Implementasi
 4. Pengujian dan Evaluasi
 5. Kesimpulan dan Saran
 6. Daftar Pustaka

11. JADWAL KEGIATAN

Jadwal kegiatan dalam pengerjaan tugas akhir dapat dilihat pada Tabel 1 sebagai berikut.

Tabel 1. Jadwal Kegiatan Pengerjaan Tugas Akhir

No	Tahapan	Bulan															
		1				2				3				4			
1	Penyusunan Proposal	■	■														
2	Analisa Kebutuhan dan Studi Literatur			■	■	■											
3	Implementasi					■	■	■	■	■	■	■					
4	Pengujian dan Evaluasi											■	■	■	■		
5	Penyusunan Buku Tugas Akhir											■	■	■	■	■	■

12. DAFTAR PUSTAKA

- [1] D. Mujica-Fargas, F. J. Gallegos-Funes and R. Cruz-Santiago, "Segmentation of Noisy Images Using the Rank M-Type L-Filter and The Fuzzy C-Means Clustering Algorithm," *Springer-Verlag Berlin Heidelberg*, 2011.
- [2] D. Mujica-Vargas, F. J. Gallegos-Funes and A. J. Rosales-Silva, "A Fuzzy Clustering Algorithm with Spatial Robust Estimation Constraint for Noisy Color Image Segmentation," *Elsevier: Pattern Recognition Letter*, 2013.
- [3] J. Valente de Oliveira and Pedrycz, *Advances in Fuzzy Clustering and Its Application*, John Wiley & Son, Ltd., 2007.
- [4] H. Le Capitaine and C. Frelicot, "A Fast Fuzzy C-Means Algorithm for Color Image Segmentation," *Laboratoire Mathematics, Image et Applications*, 2011.
- [5] F. J. Gallegos-Funes and V. I. Ponomaryov, "Real-time image filtering scheme based on robust estimators in presence of impulsive noise," *Elsevier*, no. Real Time Imaging, 2004.
- [6] R. Unnikrishnan, C. Pantofaru and M. Hebert, "Toward Objective Evaluation of Image Segmentation Algorithms," *IEEE*, vol. 29, no. Transactions On Pattern Analysis And Machine Intelligence, 2007.
- [7] M. Meila, "Comparing Clusterings - An Axiomatic View," *Proceedings of the 22nd International Conference on Machine Learning*, no. Machine Learning, 2005.