



TESIS - KI142502

**SEGMENTASI SEL BERTUMPUK PADA CITRA MIKROSKOPIS SEL  
KANKER PAYUDARA MENGGUNAKAN *SPATIAL FUZZY C-MEANS  
CLUSTERING* DAN *RAPID REGION MERGING***

Desmin Tuwohingide

NRP. 5114201049

DOSEN PEMBIMBING

Dr.Eng. Chastine Fatichah, S.Kom., M.Kom.

PROGRAM MAGISTER

BIDANG KEAHLIAN KOMPUTASI CERDAS DAN VISUAL

JURUSAN TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2017

*[Halaman ini sengaja dikosongkan]*



THESIS - KI142502

**OVERLAPPING CELL SEGMENTATION ON MICROSCOPIS BREAST  
CANCER IMAGE USING *SPATIAL FUZZY C-MEANS* AND  
*RAPID REGION MERGING***

DESMIN TUWOHINGIDE

NRP. 5114201049

SUPERVISOR

Dr.Eng. Chastine Fatichah, S.Kom., M.Kom.

MASTER PROGRAM

THE EXPERTISE FIELD OF INTELLIGENT COMPUTING AND VISION

DEPARTMENT OF INFORMATICS

FACULTY OF INFORMATION TECHNOLOGY

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2017

*[Halaman ini sengaja dikosongkan]*

Telah disusun untuk memenuhi salah satu syarat memperoleh gelar  
Magister Komputer (M.Kom.)  
di  
Institut Teknologi Sepuluh Nopember Surabaya

Oleh:

DESMIN TUWOHINGIDE

Nrp. 5114201049

Dengan judul:

Segmentasi Sel Bertumpuk Pada Citra Mikroskopis Sel Kanker Payudara  
Menggunakan Spatial Fuzzy C-Means Clustering dan Rapid Region Merging

Tanggal Ujian : 11-1-2017  
Periode Wisuda : 2016 Gasal

Disetujui oleh:

Dr. Eng. Chastine Fatichah, S.Kom, M.Kom  
NIP. 19751220200112002



(Pembimbing 1)

Dr. Eng. Nanik Suciati, S.Kom, M.Kom  
NIP. 197104281994122001



(Penguji 1)

Bilqis Amaliah, S.Kom, M.Kom  
NIP. 197509172001122001



(Penguji 2)

Arya Yudhi Wijaya, S.Kom, M.Kom  
NIP. 198409042010121002



(Penguji 3)

an. Direktur Program Pascasarjana



Prof. Dr. Ir. Tri Widjaja, M.Eng.  
NIP. 196110211986031001

Direktur Program Pasca Sarjana,

Prof. Ir. Djauhar Manfaat, M.Sc., Ph.D  
NIP. 196012021987011001

*[Halaman ini sengaja dikosongkan]*

## **Segmentasi Sel Bertumpuk pada Citra Mikroskopis Sel Kanker Payudara Menggunakan *Spatial Fuzzy C-Means Clustering* dan *Rapid Region Merging***

Nama Mahasiswa : Desmin Tuwohingide

NRP : 5114201049

Pembimbing : Dr.Eng Chastine Fatichah, S.Kom., M.Kom.

### **ABSTRAK**

Penerapan teknik pengolahan citra untuk menganalisis citra mikroskopis sel kanker payudara dilakukan untuk mempermudah diagnosis penyakit kanker payudara. Proses pemisahan sel kanker payudara bertumpuk dianggap penting karena hasil pemisahan sel kanker bertumpuk akan mempengaruhi akurasi perhitungan jumlah sel. Keberhasilan proses pemisahan sel bertumpuk juga dipengaruhi oleh proses identifikasi sel, proses deteksi sel bertumpuk dan penanganan masalah *oversegmentation*. Pemisahan sel kanker menggunakan algoritma *clustering* pada citra mikroskopis sel darah putih menghasilkan nilai akurasi yang cukup baik.

Kombinasi metode *Spatial Fuzzy C-means Clustering* (SFCM) dan *Rapid Region Merging* (RRM) untuk pemisahan sel kanker bertumpuk dan penanganan masalah *oversegmentation* dipaparkan pada penelitian ini. Citra masukan yang digunakan pada tahapan pemisahan sel bertumpuk adalah citra hasil identifikasi sel kanker payudara berdasarkan metode *Gram-Schmidt*, sedangkan sel kanker yang diproses pada tahapan pemisahan sel kanker bertumpuk adalah sel kanker yang dideteksi bertumpuk berdasarkan informasi fitur geometri area.

Berdasarkan hasil pengujian dilakukan terhadap 40 citra mikroskopis jenis *benign* dan *malignant*, kombinasi metode SFCM dan RRM memberikan hasil paling baik berdasarkan perolehan nilai rata-rata *MSE* sebesar 0,07 pada tahapan identifikasi sel dan nilai akurasi pemisahan sel bertumpuk sebesar 78.41%

**Kata Kunci:** *Segmentasi sel kanker, Segmentasi Sel Bertumpuk, Kanker Payudara, Spatial Fuzzy C-Means, Rapid Region Merging*

*[Halaman ini sengaja dikosongkan]*

## **Overlapping Cell Segmentation on Microscopic Breast Cancer Image Using Spatial Fuzzy C-Means and Rapid Region Merging**

Student Name : Desmin Tuwohingide  
NRP : 5114201049  
Supervisor : Dr.Eng Chastine Fatichah, S.Kom., M.Kom.

### **ABSTRACT**

One of data information that used for the breast cancer diagnosis is a cancer cell data in microscopic images which taken from the patient's body by using a special camera. The application techniques for image processing in analyzing the microscopic image of the breast cancer cells have been widely applied and developed. Segmentation and separation for overlapping cell is an important step in processing the microscopic image of breast cancer cells due to the accuracy of separation overlapping cell results will determine the accuracy of quantization of breast cancer cells. Quantitation of breast cancer cells is one of the factor that being consider in doctor to take action against patients.

Separation overlapping cell by using a clustering algorithm on the microscopic image of blood cells produces a good value. Spatial Fuzzy C-Means Clustering is a method that used to image segmentation based on the spatial information and adjacency probability so that besides robust to noise, the method also reduces misclassification. In addition to avoid the occurrence undersegmentation it must be done overlapping cell detection. Feature geometry is the most common information that used for overlapping cell detection, but the use of it on the microscopic image of breast cancer cells have constraints because of the breast cancer cells size are diverse, resulting susceptible to oversegmentation.

Therefore, this study proposes a method for segmentation of overlapping cells on microscopic image of breast cancer cells with Spatial Fuzzy C-Means Clustering algorithm and rapid region algorithm.

**Kata Kunci:** *Segmentation of cancer cells, segmentation of overlapped Cell, Breast Cancer, Spatil Fuzzy C-Means, Rapid Region Merging*

*[Halaman ini sengaja dikosongkan]*

## KATA PENGANTAR

Segala puji syukur dipanjatkan ke hadirat Tuhan yang mahakuasa, karena hanya atas kemurahan-Nya saja buku tesis ini dapat selesai ditulis. terselesainya penulisan buku tesis ini adalah berkat bimbingan dan bantuan berbagai pihak. Sehingga pada kesempatan ini, penulis ingin mengucapkan terima kasih kepada pihak-pihak sebagai berikut:

1. Kepada orang tua, bapak Herson Tuwohingide dan ibu Feronika Supit, serta adik Juldrike Tuwohingide, atas segala doa, dukungan dan motivasi yang diberikan selama masa studi di Institut Teknologi Sepuluh Nopember.
2. Kepada Dr. Eng. Chastine Fatichah, S.Kom, M.Kom, selaku dosen pembimbing dan dosen wali yang selalu memberikan arahan dan bimbingan dengan sabar sampai buku tesis ini bisa selesai ditulis.
3. Kepada tim dosen penguji, ibu Dr. Eng. Nanik Suciati, S.Kom, M.Kom, ibu Bilqis Amaliah, S.Kom, M.Kom dan bapak Arya Yudhi Wijaya, S.Kom M.Kom yang telah meluangkan waktu untuk memberikan arahan serta masukan.
4. Kepada seluruh civitas akademik Politeknik Negeri Nusa Utara yang telah memberikan kesempatan kepada penulis untuk bisa melanjutkan pendidikan ke jenjang S2.
5. Kepada teman-teman seperjuangan dilaboratorium, aryo dan heru dan pak Saiful Bahri Musa atas bimbingan dan teman diskusi selama implementasi penelitian, yang selalu memberi arahan dan masukan.
6. Kepada teman-teman seangkatan PraS2-Saintek Informatika tahun 2013 dan teman-teman pascasarjana informatika angkatan 2014, yang selalu memberikan motivasi
7. Kepada teman-teman KTB Timothy, kak Fitria, kak Ana, kak Zhelvyanie, kak Wati, kak Rocky, kak Ishak, abang Eri, Bela, Eun, Ann dan Pingkan atas doa dan dukungan kalian disetiap kesempatan.
8. Kepada teman-teman mba Lilia, Ningsi, ka Ida, Sri yang selalu jadi tempat mengeluh dan membagi duka.
9. Kepada teman-teman pelayanan di GKI Manyar, terutama kepada keluarga kak Anang dan kak Nike atas perhatiannya terhadap anak timur dan kepada Elin yang selalu siap siaga dengan pisang goreng.

10. Kepada staf LPMI khususnya kak Kesia atas doa, semangat, perhatian yang selalu diberikan
11. Kepada teman-teman pemuda jemaat GMIST “Kalama Darat” atas dukungan dan doa kalian.
12. Kepada keluarga besar Tuwohingide-Supit di Kabupaten Kepulauan Sangihe untuk setiap doa dan dukungan finansial yang diberikan.
13. Kepada semua pihak yang telah membantu penulis selama masa studi di Insitut Teknologi Sepuluh Nopember yang tidak bisa disebutkan satu persatu.

Surabaya, 11 Januari 2017

Desmin Tuwohingide

## DAFTAR ISI

ABSTRAK .....	vii
ABSTRACT .....	ix
KATA PENGANTAR .....	xi
DAFTAR ISI.....	xiii
DAFTAR TABEL.....	xvii
DAFTAR GAMBAR .....	xix
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Permasalahan.....	5
1.3 Batasan Masalah.....	6
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
1.6 Kontribusi .....	6
1.7 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA.....	9
2.1 Teori Kanker Payudara.....	9
2.2 Teori Dasar Segmentasi Citra.....	10
2.2.1 Deteksi Diskontinuitas .....	10
2.2.2 <i>Thresholding</i> .....	12
2.2.3 Segmentasi Berbasis Region .....	13
2.3 Metode Segmentasi Citra Mikroskopis Sel Kanker Payudara .....	13
2.3.1 Model Warna RGB .....	13
2.3.2 <i>Gram-Schmidt</i> .....	15
2.3.3 Operasi Morfologi.....	15
2.4 Metode Pemisahan Sel Bertumpuk Pada Citra Mikroskopis Sel Kanker Payudara .....	16
2.4.1 <i>Fuzzy C-Means Clustering</i> .....	16
2.4.2 <i>Spatial Fuzzy C-Means</i> .....	18
2.4.3 <i>Rapid Region Merging</i> .....	19
BAB III METODE PENELITIAN.....	23
3.1 Tahapan Penelitian .....	23

3.1.1	Tahapan Studi Literatur .....	23
3.1.2	Tahapan Analisis dan Desain Sistem.....	24
3.1.3	Tahapan Implementasi.....	24
3.1.4	Tahapan Uji Coba dan Analisis Hasil.....	24
3.1.5	Tahapan Penulisan Laporan.....	24
3.2	Analisis dan Desain Sistem.....	25
3.2.1	Dataset .....	25
3.2.2	Segmentasi Citra Mikroskopis Sel Kanker Payudara.....	26
3.2.3	Pemisahan Sel Bertumpuk Pada Citra Mikroskopis Sel Kanker Payudara .....	27
3.3	Skenario Uji Coba.....	33
3.4	Analisis Hasil .....	34
<b>BAB IV UJI COBA DAN ANALISIS HASIL .....</b>		<b>38</b>
4.1	Lingkungan Uji Coba.....	38
4.2	Data Uji Coba .....	38
4.3	Skenario Uji Coba.....	38
4.3.1	Hasil Uji Coba Skenario Pertama.....	39
4.3.2	Hasil Uji Coba Skenario Kedua.....	45
4.4	Analisis Hasil Uji Coba .....	52
4.4.1	Analisi Hasil Uji Coba Skenario Pertama .....	52
4.4.2	Analisis Hasil Uji Coba Skenario Kedua .....	53
<b>BAB V PENUTUP .....</b>		<b>58</b>
5.1	Kesimpulan .....	58
5.2	Saran .....	59
<b>DAFTAR PUSTAKA.....</b>		<b>62</b>
<b>Lampiran 1.....</b>		<b>64</b>
<b>Lampiran 2.....</b>		<b>67</b>
<b>BIODATA .....</b>		<b>69</b>



*[ halaman ini sengaja dikosongkan ]*

## DAFTAR TABEL

Tabel 3. 1 Vektor Warna RGB pada Gram-Schmidt .....	27
Tabel 4. 1 Perhitungan Nilai Mean Square Error (MSE).....	44
Tabel 4. 2 Perhitungan Nilai Akurasi Kombinasi Metode Spatial Fuzzy C-Means dan Rapid Region Merging .....	51
Tabel 4. 3 Tabel Perbandingan Metode Sebelumnya dengan Metode yang Diusulkan .....	54
Tabel 4. 4 Perhitungan uji-t untuk variabel hasil SFCM dan Metode yang Diusulkan pada taraf signifikan 5%.....	55
Tabel 4. 5 Perhitungan uji-t untuk variabel hasil SFCM dan Metode yang Diusulkan pada taraf signifikan 1%.....	55
Tabel 4. 6 Perhitungan uji-t untuk variabel hasil <i>K-Means</i> dan Metode yang Diusulkan pada taraf signifikan 5% .....	56
Tabel 4. 7 Perhitungan uji-t untuk variabel hasil K-Means dan Metode yang Diusulkan pada taraf signifikan 1% .....	56

[ *halaman ini sengaja dikosongkan* ]

## DAFTAR GAMBAR

Gambar 2. 1 Mask 3x3 (Prasetyo, 2011).....	10
Gambar 2. 2 Mask Untuk Deteksi Garis (Prasetyo, 2011).....	11
Gambar 2. 3 Skema Warna RGB Dalam Kubus (Prasetyo, 2011).....	14
Gambar 2. 4 Contoh Dua Region yang adjacent dengan 8-Adjacency (Gonzalez & Woods, 2002) .....	21
Gambar 3. 1 Tahapan Penelitian .....	23
Gambar 3. 2 Citra RGB Berupa Citra Mikroskopis Sel Kanker Payudara .....	25
Gambar 3. 3 Gambaran Umum Desain Sistem Segmentasi Sel Bertumpuk Pada Citra Mikroskopis Sel Kanker Payudara.....	26
Gambar 3. 4 Flowchart Identifikasi Sel .....	28
Gambar 3. 5 FLOWchart Pemisahan Sel Kanker Bertumpuk .....	29
Gambar 3. 6 Aturan Penggabungan Daerah Menggunakan Metode Rapid Region Merging (Chen & Chen, 2014) .....	33
Gambar 4. 1 Penentuan nilai thresholding untuk segmentasi sel kanker.....	41
Gambar 4. 2 Citra yang Dihasilkan pada Setiap Tahapan Segmentasi .....	42
Gambar 4. 3 Perbandingan Citra Hasil Segmentasi Dengan Citra Asli dan Citra Groundtruth.....	43
Gambar 4. 4 Hasil Percobaan Penentuan Nilai Threshold untuk Citra Bertumpuk .....	46
Gambar 4. 5 Uji Coba Penentuan Nilai c .....	47
Gambar 4. 6 Hasil Citra Berdasarkan Tahapan-Tahapan Pemisahan Sel Bertumpuk .....	48
Gambar 4. 7 Perbandingan Citra Hasil Pemisahan Sel Bertumpuk.....	49
Gambar 4. 8 (a)(c) Citra hasil SFCM yang Oversegmentasi (b) Citra Hasil RRM yang Sesuai dengan Citra Asli (d) Citra Hasil RRM yang Undersegmentasi.....	50
Gambar 4. 9 Citra yang Menghasilkan Nilai Akurasi Terendah.....	53
Gambar 4. 10 Citra Yang Menghasilkan Nilai Akurasi Paling Tinggi .....	53

*[ halaman ini sengaja dikosongkan ]*

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Salah satu tumor ganas yang paling banyak diderita oleh perempuan adalah kanker payudara. Kanker payudara merupakan kanker jaringan payudara yang didefinisikan sebagai suatu penyakit neoplasma ganas yang berasal dari *parenchyma*. Menurut data yang dirilis oleh *World Health Organization* (WHO) pada *website* <https://who.int>, di Indonesia angka kematian perempuan paling tinggi diakibatkan oleh kasus kanker payudara, bahkan pada tahun 2015 Pemerintah Indonesia melalui Kementerian Kesehatan melaporkan bahwa menurut data GLOBOCAN (IARC) tahun 2012, kanker payudara merupakan penyakit kanker dengan jumlah kasus baru dan penyebab kematian tertinggi. Secara nasional prevalensi penyakit kanker pada penduduk semua umur di Indonesia tahun 2013 sebesar 1,4% atau sekitar 347.792 orang, sehingga penyakit kanker payudara masuk dalam daftar penyakit kanker dengan prevalensi tertinggi di Indonesia.

Pada umumnya, kanker payudara didiagnosis melalui pemeriksaan rutin atau ketika pasien menyadari gejala-gejala tertentu, tetapi pemeriksaan fisik saja tidak cukup untuk mengonfirmasi diagnosis kanker payudara sehingga dilakukan berbagai prosedur pemeriksaan seperti mamografi, biopsi dan lain sebagainya untuk memperoleh informasi terkait konfirmasi diagnosis dan kebutuhan terapi pasien. Salah satu data yang informasinya digunakan untuk diagnosis penyakit kanker payudara adalah data sel kanker berupa citra mikroskopis yang diambil dari tubuh pasien menggunakan kamera khusus. Citra mikroskopis tersebut yang kemudian akan dianalisis oleh tim medis untuk mengonfirmasi kesehatan pasien dengan melakukan identifikasi sel kanker dan perhitungan sel kanker payudara.

Penerapan teknik pengolahan citra untuk menganalisis citra mikroskopis sel kanker payudara telah banyak dilakukan dan dikembangkan. Segmentasi citra adalah salah satu masalah paling penting yang diangkat dalam pengolahan citra

mikroskopis sel kanker payudara karena proses segmentasi harus menghasilkan *output* yang merepresentasikan citra asli. Segmentasi citra bertujuan membagi citra menjadi beberapa objek atau region. Hal ini dilakukan agar objek yang diinginkan dalam suatu citra, terpisahkan dari objek lainnya. Segmentasi citra umumnya didasarkan pada beberapa hal seperti intensitas yang mencolok dan kemiripan region. Pada penerapannya, proses segmentasi citra mikroskopis terdapat berbagai permasalahan, yaitu teknik segmentasi tradisional tidak bisa secara maksimal melakukan segmentasi dan pemisahan sel bertumpuk pada citra mikroskopis yang memiliki beragam karakteristik dari segi warna, bentuk, ukuran dan tekstur.

Keberagaman inilah yang menjadi tantangan yang diangkat dalam penelitian sebelumnya untuk melakukan identifikasi sel kanker. Pada penelitian (Mouelhi, Sayadi, & Fnaiech, 2013) Identifikasi sel kanker payudara menggunakan *Color Active Contour* dan *Multilayer Neural Network* berdasarkan informasi warna sel menggunakan *Fisher Linear Discriminant* menghasilkan nilai akurasi segmentasi sebesar 95,50%. Identifikasi sel kanker yang dilakukan pada penelitian (Aini, 2015) berdasarkan *Region-Based Active Contour* menghasilkan nilai MSE dalam rentan nilai 0,05 sampai 0,17 dengan rata-rata nilai MSE sebesar 0,08. Kedua penelitian ini menggunakan dataset yang sama dengan jumlah dataset sebanyak 15 citra dan metode yang diusulkan dari kedua penelitian ini menghasilkan nilai MSE yang tidak jauh berbeda. Dalam penelitian (Fatichah & Suciati, 2015) dipaparkan mengenai identifikasi sel kanker payudara pada citra mikroskopis sel kanker payudara dengan menggunakan metode *Gram Schmidt* menghasilkan nilai *Mean Square Error* (MSE) dalam rentan nilai dari 0,05 sampai 0,11 dan menghasilkan nilai rata-rata MSE sebesar 0,08 dibandingkan dengan metode *Otsu* yang menghasilkan nilai MSE dalam rentan nilai 0,11 sampai 0,57 dengan rata-rata nilai MSE 0,26.

Dalam pengolahan citra mikroskopis sel kanker payudara, masalah deteksi sel bertumpuk dan pemisahan sel juga menjadi masalah yang banyak diangkat dalam penelitian karena hasil pemisahan sel bertumpuk sangat menentukan akurasi perhitungan jumlah sel kanker payudara. Beberapa penelitian berikut

mengangkat masalah deteksi dan pemisahan sel bertumpuk dalam penelitiannya, (Phukpattaranont & Boonyaphiphat, 2007) menjelaskan mengenai pentingnya deteksi sel kanker payudara bertumpuk sebelum proses pemisahan sel bertumpuk. Keberhasilan deteksi sel bertumpuk dalam segmentasi dan pemisahan sel bertumpuk yang dipaparkan dalam penelitian ini adalah dengan menerapkan deteksi sel bertumpuk berdasarkan fitur geometri area sel kanker yang dibedakan ke dalam tiga kategori yaitu *small* yang dianggap *noise*, *medium* dianggap sel tunggal dan *large* dianggap sel bertumpuk, sehingga untuk proses pemisahan sel bertumpuk hanya sel yang dianggap bertumpuk yang diproses menggunakan metode *Marker-Controlled Watershed*. Deteksi sel kanker bertumpuk dinilai mampu memperkecil kemungkinan *undersegmentation* dan *oversegmentation*. Penelitian (Mouelhi, Sayadi, Fnaiech, & Mrad, 2013) menerapkan deteksi sel kanker payudara bertumpuk menggunakan deteksi *concave point*, karena bentuk sel kanker payudara yang beragam sehingga dilakukan deteksi titik yang memiliki *concavity* yang tinggi pada daerah *boundaries*. Titik-titik *concavity* ditentukan dengan teknik pendeteksian sudut berdasarkan informasi *global* dan *local curvature* dan metode *Watershed* untuk pemisahan sel bertumpuk. Hasil nilai akurasi yang diperoleh pada penelitian ini adalah 89,20% untuk dataset 1 dan 92,70% untuk dataset 2.

Secara berkelanjutan (Mouelhi, Sayadi, & Fnaiech 2013) menerapkan *Watershed* untuk pemisahan sel bertumpuk berdasarkan *Fuzzy Active Contour Model* dan deteksi *concave point* menggunakan konstruksi *concave vertex graph* dan pemilihan jalur garis terpendek untuk inialisasi garis pemisah, dan menghasilkan nilai akurasi sebesar 96,23%. Pada penelitian (Faried, 2013) deteksi dan pemisahan sel bertumpuk pada citra mikroskopis sel darah merah menggunakan informasi fitur geometri dan analisis kluster *K-Means Clustering* untuk pemisahan sel kanker bertumpuk dengan memperoleh hasil yang menunjukkan bahwa metode ini memiliki rata-rata akurasi yang lebih tinggi yaitu 96,46% dan nilai *error* yang lebih rendah dibandingkan dengan metode *Classical Watershed* yang memperoleh rata-rata nilai akurasi sebesar 79,87% dan metode *Condition Erosion Watershed* yang menghasilkan rata-rata nilai akurasi sebesar

87,22%. Dalam penelitian (Aini, 2015) kombinasi metode *Region-Based Active Contour* dan *Bayesian* untuk pemisahan sel bertumpuk menghasilkan nilai rata-rata akurasi sebesar 61,30% sementara kombinasi metode *Region-Based Active Contour* dan *Watershed* menghasilkan rata-rata nilai akurasi sebesar 65%.

Berdasarkan penelitian-penelitian terdahulu, beberapa hal yang sangat perlu untuk dipertimbangkan guna meningkatkan hasil pemisahan sel kanker payudara bertumpuk. Pertama, identifikasi sel kanker payudara harus dilakukan dengan memperhatikan karakteristik citra mikroskopis guna mendapatkan sel kanker yang sesuai dengan citra *groundtruth* agar menghasilkan citra segmentasi yang tingkat kesalahannya kecil. Kedua, deteksi sel kanker payudara bertumpuk perlu untuk dilakukan untuk mengurangi kemungkinan *oversegmentation* terhadap sel tunggal dan *undersegmentation* terhadap sel bertumpuk seperti yang dipaparkan dalam penelitian (Phukpattaranont & Boonyaphiphat, 2007) berdasarkan informasi fitur geometri area dengan memberikan asumsi luasan sel kanker tunggal dan luasan sel kanker bertumpuk. Untuk pemisahan sel bertumpuk, Berdasarkan hasil penelitian (Faried, 2013) bahwa metode *clustering* menghasilkan pemisahan sel darah bertumpuk yang baik dan memperoleh nilai akurasi yang tinggi dibandingkan dengan metode *classical watershed* dan metode *condition erosion watershed*. Pada penelitian (Li & Shen, 2010)(Beevi, Sathik, Senthamaraiannan, & Yasmin, 2010) dipaparkan metode Spatial Fuzzy C-Means Clustering (SFCM) yang merupakan pengembangan metode *Fuzzy C-Means Clustering* berdasarkan informasi spasial dan probabilitas ketetanggaan, karena informasi spasial dan probabilitas ketetanggaan memiliki korelasi yang tinggi dan sangat penting dalam pengelompokan citra. Metode ini dianggap memiliki kelebihan yang tidak hanya *robust* terhadap *noise* tapi juga mengurangi kesalahan pengelompokan.

Selain beberapa masalah di atas, *oversegmentation* menjadi salah satu masalah utama yang harus ditangani dalam proses pemisahan sel bertumpuk. Metode *Region merging* adalah metode yang diimplementasikan untuk mengurangi masalah *oversegmentation*, salah satunya adalah *Rapid Region Merging* yang merupakan metode penggabungan daerah berdasarkan kemiripan

region yang mampu meningkatkan nilai akurasi segmentasi, karena mampu mengatasi masalah *oversegmentation* yang dihasilkan oleh metode *watershed* (Chen & Chen, 2014).

Berdasarkan uraian di atas, penelitian ini mengusulkan metode segmentasi sel kanker bertumpuk pada citra mikroskopis sel kanker payudara dengan *Spatial Fuzzy C-Means Clustering* (SFCM) dan *Rapid Region Merging* (RRM). Tahapan pertama pada penelitian ini adalah melakukan segmentasi sel kanker pada citra mikroskopis sel kanker payudara berdasarkan identifikasi sel menggunakan metode *Gram-Schmidt*, kemudian dilakukan deteksi sel kanker bertumpuk berdasarkan informasi fitur geometri. Selanjutnya adalah pemisahan sel kanker bertumpuk menggunakan *Spatial Fuzzy C-Means Clustering* dan metode *Rapid Region Merging* mengatasi masalah *oversegmentation*.

## **1.2 Permasalahan**

Berdasarkan latar belakang di atas permasalahan yang diangkat pada penelitian ini adalah :

1. Bagaimana melakukan identifikasi sel kanker payudara berdasarkan metode *Gram-Schmidt*?
2. Bagaimana melakukan pemisahan sel pada citra mikroskopis sel kanker payudara menggunakan algoritma *Spatial Fuzzy C-Means Clustering* berdasarkan informasi fitur geometri area untuk deteksi sel kanker bertumpuk?
3. Bagaimana melakukan penggabungan daerah *oversegmentation* menggunakan metode *Rapid Region Merging*?
4. Bagaimana menguji performa metode yang diusulkan berdasarkan hasil yang diperoleh pada tahapan identifikasi sel kanker dan tahapan pemisahan sel bertumpuk?

### **1.3 Batasan Masalah**

Permasalahan yang dibahas dalam penelitian ini memiliki beberapa batasan masalah sebagai berikut:

1. Aplikasi dirancang menggunakan MATLAB R2015a
2. Dataset yang digunakan adalah citra mikroskopis sel kanker payudara yang diunduh di [www.bioimage.ucsb.edu](http://www.bioimage.ucsb.edu)

### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah melakukan identifikasi sel dan pemisahan sel bertumpuk pada citra mikroskopis sel kanker payudara menggunakan algoritma *Spatial Fuzzy C-Means Clustering* dan *Rapid Region Merging*.

### **1.5 Manfaat Penelitian**

Manfaat dari penelitian ini adalah dihasilkannya sistem yang digunakan untuk segmentasi sel bertumpuk pada citra mikroskopis sel kanker payudara menggunakan kombinasi algoritma *Spatial Fuzzy C-Means* dan *Rapid Region Merging* yang diharapkan memberikan hasil perhitungan sel kanker payudara yang lebih akurat.

### **1.6 Kontribusi**

Kontribusi pada penelitian ini adalah mengusulkan metode *Spatial Fuzzy C-Means Clustering* dan *Rapid Region Merging* untuk pemisahan sel bertumpuk pada citra mikroskopis sel kanker payudara berdasarkan fitur geometri area dengan harapan dapat mengurangi permasalahan *oversegmentation* dan meningkatkan akurasi perhitungan jumlah sel kanker payudara.

### **1.7 Sistematika Penulisan**

Laporan metodologi penelitian ini dibagi dalam beberapa bab, sebagai berikut:

- a. Bab I Pendahuluan, berisi Latar Belakang, Permasalahan, Batasan Masalah, Tujuan, Manfaat, Kontribusi, dan Sistematika Penulisan.

- b. Bab II Tinjauan pustaka, akan dibahas dasar-dasar teori yang mendukung penelitian yang diusulkan.
- c. Bab III Metodologi, akan membahas rancangan sistem, dataset, skenario uji coba hingga jadwal penelitian.
- d. Bab IV Uji coba dan Analisis hasil, akan membahas hasil implementasi dari rancangan sistem yang dibuat, uji coba sistem serta analisis hasil.
- e. Bab V Kesimpulan, akan membahas kesimpulan dari penelitian yang diusulkan.

*[Halaman ini sengaja dikosongkan]*

## **BAB II**

### **TINJAUAN PUSTAKA**

Pada Bab 2 ini akan dijelaskan mengenai tinjauan pustaka yang terkait dengan penelitian yang diusulkan. Beberapa dasar teori yang akan di bahas pada bab ini, antara lain: teori kanker payudara, segmentasi citra, metode segmentasi citra, *Spatial Fuzzy C-Means Clustering*, dan metode *Rapid Region Merging*.

#### **2.1 Teori Kanker Payudara**

Penyakit kanker adalah penyakit yang timbul akibat pertumbuhan tidak normal sel jaringan tubuh yang berubah menjadi sel kanker. Penyakit kanker merupakan salah satu penyebab kematian utama di seluruh dunia. Pada tahun 2012, sekitar 8,2 juta kematian disebabkan oleh kanker. Faktor-faktor risiko penyakit kanker antara lain, faktor genetik, faktor karsinogen di antaranya yaitu zat kimia, radiasi, virus, hormon, dan iritasi kronis, dan faktor perilaku atau gaya hidup di antaranya yaitu merokok, pola makan yang tidak sehat, konsumsi alkohol dan kurang aktivitas fisik. Lebih dari 30% angka kematian akibat kanker disebabkan karena faktor risiko perilaku atau gaya hidup, dan merokok merupakan faktor risiko utama kanker yang menyebabkan terjadinya lebih dari 20% angka kematian kanker di dunia. Kanker paru, hati, perut, kolorektal dan kanker payudara adalah penyebab terbesar kematian akibat kanker setiap tahunnya (Pusat data dan informasi Kementerian Kesehatan RI, [www.depkes.go.id](http://www.depkes.go.id)).

Kanker payudara merupakan kanker jaringan payudara yang didefinisikan sebagai suatu penyakit neoplasma ganas yang berasal dari *parenchyma* dan merupakan tumor ganas yang paling umum diderita oleh perempuan. Melalui Kementerian Kesehatan Republik Indonesia, dilaporkan bahwa menurut data GLOBOCAN (IARC), tahun 2012 diketahui bahwa kanker payudara merupakan penyakit kanker dengan jumlah kasus baru tertinggi, yaitu dengan jumlah sebesar 43,3%, dan jumlah kematian akibat kanker payudara adalah sebesar 12,9%. Walaupun jumlah penderita kanker payudara bertambah banyak setiap tahunnya,

namun angka kematian akibat penyakit kanker payudara lebih rendah dibandingkan dengan jumlah kasus baru, sehingga jika kanker payudara tersebut dapat dideteksi dan ditangani sejak dini maka kemungkinan sembuh akan lebih tinggi (Pusat data dan informasi Kementerian Kesehatan RI, [www.depkes.go.id](http://www.depkes.go.id)).

## 2.2 Teori Dasar Segmentasi Citra

Tujuan utama dari segmentasi citra adalah membagi citra menjadi beberapa region untuk mendapatkan objek yang diinginkan dalam suatu citra, dengan kata lain objek yang diinginkan terpisah dengan objek lain yang dianggap *noise*. Maka, segmentasi seharusnya berhenti ketika objek yang diinginkan dalam aplikasi telah terisolasi. Algoritma segmentasi citra umumnya didasarkan satu dari dua properti nilai intensitas, yaitu diskontinuitas dan kemiripan. Pada diskontinuitas, pendekatan yang digunakan adalah memecah/memilah citra berdasarkan perubahan kasar dalam intensitas, seperti tepi dalam citra. Pendekatan yang digunakan pada kemiripan berdasarkan pemecahan citra ke dalam region yang sama menurut sejumlah kriteria yang didefinisikan, seperti *thresholding*, *region growing*, *region splitting*, dan *merging* (Prasetyo, 2011).

### 2.2.1 Deteksi Diskontinuitas

Cara paling umum untuk mencari diskontinuitas adalah menjalankan *mask* sepanjang citra dengan prosedur komputasi jumlah dari perkalian koefisien dengan nilai *gray level* yang mengisi region yang ditempati oleh *mask* pada gambar 2.1. Maka hasil *masking* pada sembarang titik dalam citra adalah *gray level* piksel  $z_i$  diasosiasikan dengan koefisien mask  $w_i$ , mengikuti Persamaan 2.1 dan Gambar 2.1(a) (Prasetyo, 2011).

$$R = w_1z_1 + w_2z_2 + \dots + w_9z_9 \quad 2.1$$

$w_1$	$w_2$	$w_3$
$w_4$	$w_5$	$w_6$
$w_7$	$w_8$	$w_9$

Umum

-1	-1	-1
-1	8	-1
-1	-1	-1

(b) Mask Laplacian

(a) Mask

Gambar 2. 1 Mask 3x3 (Prasetyo, 2011)

- Deteksi Titik

Untuk deteksi titik-titik terisolasi pada citra menggunakan *mask* pada Gambar 2.1(a). Titik deteksi pada lokasi di mana *mask* dipusatkan. Jika mutlak hasil *masking*  $R$  lebih besar atau sama dengan  $R$  *threshold* non-negatif (Persamaan 2.2). Pada dasarnya formula ini digunakan untuk mengukur perbedaan bobot antara titik pusat *mask* dan tetangganya. Ide bahwa *isolating point* (sebuah titik di mana gray levelnya secara signifikan berbeda dari *background* dan ditempatkan dalam daerah *homogenous* atau dekat dengan *homogenous*) akan berbeda dari sekitarnya dan dengan mudah dideteksi oleh *mask* ini (Prasetyo, 2011).

$$|R| \geq T \tag{2.2}$$

- Deteksi Garis

Jika  $R_1, R_2, R_3,$  dan  $R_4$  menyatakan hasil *mask* pada Gambar 2.2 dari kiri ke kanan, dan nilai  $R$  diberikan oleh Persamaan 2.2 dan jika empat *mask* dijalankan secara individu terhadap citra, untuk titik tertentu dalam citra,  $|R_i| > |R_j|$  untuk semua  $j \neq i$ , titik tersebut akan dikatakan lebih berasosiasi dengan garis dalam arah *mask*  $i$ . Alternatifnya, deteksi garis dapat dilakukan dalam arah tertentu. Dalam kasus ini, *mask* yang digunakan berasosiasi dengan arah dan output *threshold* (Prasetyo, 2011).

-1	-1	-1
2	2	2
-1	-1	-1

(a) Horizontal

2	-1	-1
-1	2	-1
-1	-1	2

(b)  $+45^\circ$

-1	2	-1
-1	2	-1
-1	2	-1

(b) Vertikal

-1	-1	2
-1	2	-1
2	-1	-1

(d)  $-45^\circ$

Gambar 2. 2 *Mask* Untuk Deteksi Garis (Prasetyo, 2011)

Dengan kata lain, jika ingin mendeteksi semua garis dalam citra dengan arah yang telah didefinisikan *mask*, maka cukup menjalankan *mask* sepanjang bidang citra dan melakukan *threshold* nilai absolut dari hasil. Titik-titik yang tersisa adalah hasil yang terkuat, di mana garis dengan ketebalan satu piksel berhubungan dengan arah yang didefinisikan oleh *mask*.

- Deteksi Tepi

Deteksi tepi merupakan pendekatan yang paling umum untuk pendeteksian diskontinuitas nilai intensitas, seperti diskontinuitas yang dideteksi oleh penggunaan turunan pertama dan kedua. Pilihan turunan pertama dalam pengolahan citra adalah Gradien. Gradien fungsi 2-D  $f(x, y)$  didefinisikan sebagai vektor pada Persamaan 2.3. Turunan kedua dalam pengolahan citra umumnya dihitung menggunakan *Laplacian*, sehingga fungsi *Laplacian* 2-D  $f(x, y)$  dibentuk dari turunan kedua mengikuti Persamaan 2.4 (Prasetyo, 2011).

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (2.3)$$

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \quad (2.4)$$

*Laplacian* jarang digunakan sendirian tanpa deteksi tepi, karena sebagai turunan kedua, *Laplacian* tidak sensitif terhadap *noise*. Jaraknya menghasilkan tepi *double* dan tidak bisa mendeteksi arah tepi tetapi *Laplacian* bisa menjadi pelengkap yang *powerfull* ketika digunakan dalam kombinasi dengan teknik deteksi tepi yang lain.

Ide dasar deteksi tepi adalah untuk mencari tempat di dalam citra di mana intensitas berubah secara cepat, menggunakan satu dari dua kriteria umum berikut:

1. Mencari tempat di mana turunan pertama dari intensitas lebih besar dalam jarak daripada *threshold* yang ditetapkan.
2. Mencari tempat di mana turunan kedua dari intensitas mempunyai *zero crossing*.

### 2.2.2 *Thresholding*

*Thresholding* citra menjadi titik pusat dalam aplikasi segmentasi citra karena properti dan kesederhanaannya dalam implementasi. Jika sebuah citra

$f(x, y)$ , terdiri dari objek terang pada *background* gelap, maka piksel objek dan *background* mempunyai level intensitas yang dikelompokkan ke dalam dua mode dominan. Satu cara yang jelas untuk mengekstrak objek dari *background* adalah dengan memilih *threshold*  $T$  yang membagi mode-mode ini. Kemudian sembarang titik  $(x, y)$  untuk dimana  $f(x, y) \geq T$  disebut *object point*, sedangkan yang lain disebut *backgroundpoint*. Citra yang di *threshold*  $g(x, y)$  didefinisikan pada Persamaan 2.5

$$g(x, y) = \begin{cases} 1 & \text{jika } f(x, y) \geq T \\ 0 & \text{jika } f(x, y) < T \end{cases} \quad (2.5)$$

Piksel yang diberi nilai 1 berkaitan dengan objek sedangkan piksel yang diberi nilai 0 berkaitan dengan *background*. Ketika  $T$  adalah konstanta, pendekatan ini disebut *global thresholding* (Prasetyo, 2011).

### 2.2.3 Segmentasi Berbasis Region

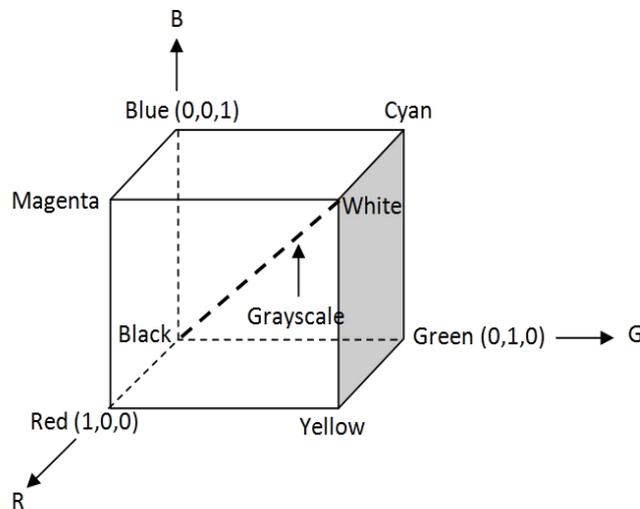
Jika  $R$  merepresentasikan semua daerah citra, segmentasi adalah proses di mana pemecahan  $R$  ke dalam sub-bagian berdasarkan beberapa syarat. Syarat (1) mengindikasikan bahwa segmentasi harus lengkap; bahwa setiap piksel berada dalam region. Syarat (2) membutuhkan titik-titik dalam region yang dihubungkan dalam beberapa cara yang didefinisikan dalam 4 atau 8-*connected*. Syarat (3) mengindikasikan bahwa region harus *disjoint*. Syarat (4) bekerja dengan properti yang harus dicapai oleh piksel dalam region yang tersegmentasi, misalnya  $P(R_i) = TRUE$  jika semua piksel dalam  $R_i$  mempunyai gray level yang sama. Syarat (5) mengindikasikan bahwa region tetangga  $R_i$  dan  $R_j$  adalah berbeda dalam predikat  $P$  (Prasetyo, 2011).

## 2.3 Metode Segmentasi Citra Mikroskopis Sel Kanker Payudara

### 2.3.1 Model Warna RGB

Model warna (*color space* atau *color system*) bertujuan memfasilitasi spesifikasi warna dalam beberapa standar. Model warna merepresentasikan sistem koordinat dan *sub-space* di dalam sistem di mana setiap warna direpresentasikan oleh titik tunggal. Dalam model warna RGB, setiap warna memperlihatkan

komponen *spectral primary red, green, dan blue*, yang didasarkan pada sistem koordinat kartesius.



Gambar 2. 3 Skema Warna RGB Dalam Kubus (Prasetyo, 2011)

*Sub-space* warna yang dicari adalah kubus yang asumsikan bahwa semua nilai warna dinormalisasikan, maka nilai R,G,B diasumsikan dalam *range* [0,1] seperti yang ditunjukkan pada Gambar 2.2, nilai RGB berada pada tiga sudut dan pada tiga sudut lainnya adalah *cyan, magenta,* dan *yellow* sedangkan hitam adalah *origin* dan putih adalah titik paling jauh dari *origin*. Pada model warna ini, *grayscale* diperluas dari hitam ke putih, sepanjang garis gabungan dua titik. Perbedaan warna dalam model ini adalah titik-titik yang berada dalam kubus dan didefinisikan oleh penyebaran vektor dari *origin*.

Citra yang direpresentasikan dalam model warna RGB terdiri dari tiga komponen citra, masing-masing untuk setiap warna primer R,G,B. Jumlah bit yang digunakan untuk merepresentasikan setiap piksel dalam *space* RGB disebut *pixel depth*. Setiap citra *red, green, dan blue* adalah citra 8-bit. Dalam kondisi setiap warna piksel memiliki kedalaman 24-bit (3 lapis citra dengan jumlah bit per lapis) (Prasetyo, 2011).

### 2.3.2 Gram-Schmidt

Sebuah himpunan orthogonal dimana setiap vektor mempunyai norm 1, dinamakan himpunan orthonormal. Suatu himpunan yang bukan orthonormal dapat diubah menjadi himpunan orthonormal dengan menggunakan proses *gram-schmidt*. Misalkan  $S = \{v_1, v_2, \dots, v_n\}$  dan menghasilkan sebuah basis orthogonal  $S = \{u_1, u_2, \dots, u_n\}$  yang memiliki subruang sama dengan  $S$  yang diperoleh berdasarkan proses *Gram Schmidt*. Proyeksi *Gram Schmidt* dijelaskan pada persamaan 2.6(Rezatofighi, 2009).

$$Proj_u v = \frac{\langle u, v \rangle}{\langle u, u \rangle} u = \langle u, v \rangle \frac{u}{\langle u, u \rangle} \quad (2.6)$$

Dimana  $\langle u, v \rangle$  adalah ruang hasil kali dalam dan vektor  $u$  dan vektor  $v$ . Proses *Gram Schmidt* dapat dilihat pada persamaan berikut:

$$\begin{aligned} u_1 &= v_1, & e_1 &= \frac{u_1}{\|u_1\|} \\ u_2 &= v_2 - proj_{u_1} v_2 & e_2 &= \frac{u_2}{\|u_2\|} \\ u_3 &= v_3 - proj_{u_1} v_3 - proj_{u_2} v_3 & e_3 &= \frac{u_3}{\|u_3\|} \\ u_k &= v_k - \sum_{j=1}^{k-1} proj_{u_j} v_k & e_k &= \frac{u_k}{\|u_k\|} \end{aligned} \quad (2.7)$$

$u_1, \dots, u_k$  diperlukan untuk vektor orthogonal, dan vektor normal  $e_1, \dots, e_k$ , membentuk himpunan orthonormal. Berdasarkan metode tersebut, untuk  $S = \{v_1, v_2, \dots, v_n\}$ , dapat ditemukan vektor  $v_k$  (yang diinginkan) yang memiliki *orthogonality* maksimum dan *orthogonality* minimum untuk vektor yang lain sesuai  $N$  dimensi subruang (). Vektor  $w_k$  dapat dihitung dengan persamaan 2.8 dan hasil *inner product* dalam set  $S$  oleh  $w_k$  sesuai dengan persamaan 2.9 (Rezatofighi, 2009).

$$w_k = v_k - \sum_{j=1}^{k-1} proj_{v_j} v_k \quad (2.8)$$

$$\begin{cases} \langle v_j, w_k \rangle = 0 & j \in 1, \dots, n \text{ dan } j \notin k \\ \langle v_k, w_k \rangle = K & K \neq 0 \end{cases} \quad (2.9)$$

### 2.3.3 Operasi Morfologi

Dalam pengolahan citra konsep matematika morfologi digunakan untuk mendapatkan atau mengekstrak komponen dari suatu gambar yang berguna untuk

merepresentasikan dan mendeskripsikan suatu bentuk, seperti *boundaries*, *skeletons* dan *convex hull*. Teknik morfologi juga digunakan untuk *preprocessing* dan *postprocessing*, seperti morfologi *filtering*, *thinning* dan *pruning*, selain itu teknik morfologi, biasanya diterapkan untuk menghapus *noise* selama proses segmentasi. Dua operasi morfologi yang paling dasar adalah dilasi dan erosi, yang kemudian dikembangkan menjadi *opening* dan *closing*.

Proses dalam dilasi adalah penambahan dan penebalan dalam citra biner. Penebalan dalam citra dikontrol oleh bentuk *structuring elemen* (*strel*) yang digunakan. Dilasi sangat berguna ketika diterapkan dalam objek-objek yang terputus akibat hasil pengambilan citra yang terganggu oleh *noise*, kerusakan objek yang dijadikan citra atau disebabkan oleh resolusi yang jelek. Erosi adalah proses pengecilan atau menipiskan objek pada citra biner. Jumlah piksel yang dicecilkan atau ditipiskan dari batas objek pada citra tergantung pada ukuran dan bentuk dari *strel* yang digunakan. Erosi dapat dianggap sebagai operasi morfologi *filtering* karena detail citra yang lebih kecil dari *strel* akan dihilangkan dari citra (Abbas, Amel H; Kareem, Aryan A; Kamil, 2015).

Dua morfologi lain yang penting adalah *opening* dan *closing*. *Opening* secara umum menghaluskan garis-garis bentuk objek, menghilangkan bagian-bagian yang sempit, dan menghilangkan penonjolan-penonjolan yang tipis. *Closing* cenderung menghaluskan garis-garis bentuk tetapi kebalikan dari *opening*, secara umum *closing* menolak pecahan-pecahan sempit dan teluk yang panjang dan tipis, menghilangkan lubang kecil dan mengisi gap pada garis-garis bentuk (*contour*).

## **2.4 Metode Pemisahan Sel Bertumpuk Pada Citra Mikroskopis Sel Kanker Payudara**

### **2.4.1 Fuzzy C-Means Clustering**

*Clustering* adalah proses pengelompokan dataset yang tidak memiliki label berdasarkan pengukuran kemiripan antara pola-pola data. Sebuah kluster berisi pola-pola yang sama yang ditempatkan bersama. Teknik pengelompokan *fuzzy* menghasilkan partisi *fuzzy* dari data. Pola-pola data yang mungkin menjadi anggota beberapa *cluster* memiliki nilai *membership* yang berbeda dengan *cluster* lainnya.

Nilai *membership* dari sebuah pola data ke *cluster* menunjukkan kemiripan antara pola data yang diberikan dengan *cluster* (Beevi et al., 2010).

*Clustering* dengan metode *Fuzzy C-Means* (FCM) didasarkan pada teori logika *fuzzy*. Teori ini pertama kali diperkenalkan oleh Lofti Zadeh (1965) dengan nama himpunan (*fuzzy set*). Dalam teori *fuzzy*, keanggotaan sebuah data tidak diberikan secara tegas dengan nilai 1 (anggota) dan 0 (bukan anggota), melainkan dengan suatu nilai derajat keanggotaan yang jangkauan nilainya 0 sampai 1. Semakin tinggi nilai keanggotaannya maka semakin tinggi derajat keanggotaannya, dan semakin kecil nilai keanggotaannya maka semakin rendah derajat keanggotaannya. FCM merupakan versi *fuzzy* dari *K-Means* dengan beberapa modifikasi yang membedakannya (Prasetyo, 2014).

Jika diasumsikan sejumlah data dalam himpunan data  $X$  yang berisi  $n$  data yang dinotasikan  $X = \{x_1, x_2, \dots, x_n\}$ , dimana setiap data mempunyai fitur  $r$  dimensi  $x_{i1}, x_{i2}, \dots, x_{ir}$ , yang dinotasikan  $x_i = \{x_{i1}, x_{i2}, \dots, x_{ir}\}$ . Ada sejumlah *cluster*  $C$  dengan *centroid*  $c_1, c_2, \dots, c_k$ , yang mana  $k$  adalah jumlah *cluster*. Setiap data mempunyai derajat keanggotaan pada setiap cluster, dinyatakan dengan  $u_{ij}$ , pada persamaan 2.12,  $i$  menyatakan data  $x_i$  dan  $j$  menyatakan  $c_j$ . Jumlah nilai derajat keanggotaan setiap data  $x_i$  selalu sama dengan 1 seperti pada persamaan 2.10 (Prasetyo, 2014).

$$\sum_{j=1}^k u_{ij} = 1 \quad (2.10)$$

Untuk *cluster*  $c_j$ , setiap cluster berisi paling sedikit satu data dengan nilai keanggotaan tidak nol, namun tidak berisi derajat satu pada semua data. *Cluster*  $c_j$  dapat diformulasikan pada persamaan 2.11

$$0 < \sum_{i=1}^n u_{ij} < n \quad (2.11)$$

$$u_{ij} = \frac{D(x_i, c_j)^{\frac{-2}{w-1}}}{\sum_{l=1}^k D(x_i, c_l)^{\frac{-2}{w-1}}} \quad (2.12)$$

Parameter  $c_j$  adalah *centroid cluster* ke- $j$ ,  $D()$  adalah jarak data dan *centroid*, sedangkan  $w$  adalah parameter bobot pangkat (*weighting exponent*).  $w$  tidak memiliki nilai ketetapan, biasanya nilai  $w > 1$  dan umumnya diberi nilai 2. Untuk menghitung *centroid* pada *cluster*  $c_l$  pada fitur  $j$ , digunakan persamaan 2.13,

dengan parameter  $N$  adalah jumlah data,  $w$  adalah bobot pangkat, dan  $u_{ij}$  adalah derajat keanggotaan data  $x_i$  ke cluster  $c_l$ . Sementara, fungsi objektif menggunakan persamaan 2.14(Prasetyo, 2014).

$$c_{lj} = \frac{\sum_{i=1}^N (u_{il})^w x_{ij}}{\sum_{i=1}^N (u_{il})^w} \quad (2.13)$$

$$J = \sum_{i=1}^N \sum_{l=1}^k (u_{ij})^w D(x_i, c_l)^2 \quad (2.14)$$

Secara prinsip, algoritma FCM memiliki kesamaan dengan *K-Means*. Seperti yang ditunjukkan pada beberapa tahapan algoritma FCM berikut (Prasetyo, 2014):

1. Inisialisasi : menentukan jumlah *cluster* ( $k \geq 2$ ), menentukan bobot pangkat ( $w \geq 1$ ), menentukan jumlah maksimal iterasi, menentukan nilai ambang batas perubahan fungsi objektif
2. Berikan nilai awal pada matriks *fuzzy pseudo-partition*, dengan syarat seperti pada persamaan 2.10.
3. Lakukan langkah 4 sampai 5 selama syarat masih terpenuhi: (1) apabila perubahan objektif masih di atas nilai ambang batas yang ditentukan; atau (2) iterasi maksimal belum tercapai.
4. Hitung nilai *centroid* dari masing-masing *cluster* menggunakan persamaan 2.13
5. Hitung kembali matriks *fuzzy pseudo-partition* (derajat keanggotaan setiap data pada setiap *cluster*) menggunakan persamaan 2.12

#### 2.4.2 *Spatial Fuzzy C-Means*

Salah satu karakteristik penting sebuah citra adalah ketetanggaan piksel yang berkorelasi tinggi. Dengan kata lain, ketetanggaan piksel memiliki kemiripan nilai fitur, dan kemungkinan mereka berada dalam *cluster* yang sama sangat besar. Hubungan spasial adalah penting dalam *clustering*, tetapi hal ini tidak digunakan pada FCM standar. Sehingga algoritma FCM standar sangat rentan terhadap *noise*. Untuk mengatasi kelemahan tersebut (Li & Shen, 2010) memanfaatkan informasi spasial yang dibentuk menggunakan distribusi statistika dari ketetanggaan piksel dan *prior probability* untuk memperoleh sebuah *membership function* baru yang

gunakan pada *clustering* (dijelaskan pada Bab selanjutnya). Metode ini tidak hanya efektif untuk menghapus *noise* tetapi juga untuk mengurangi kesalahan pengelompokan piksel.

Algoritma SFCM yang dipaparkan dalam penelitian (Li & Shen, 2010) dapat dinotasikan ke dalam tahapan-tahapan berikut:

- 1) Menentukan nilai  $c, m, \varepsilon$  dan mengatur  $l = 0$ , melakukan inisialisasi nilai  $U^0$  dan  $V^0$
- 2) Menghitung *spatial function* dan *Prior Probability* sesuai Persamaan 3.1 dan Persamaan 3.2
- 3) Memperbaharui keanggotaan matriks  $U^{*(l)}$  dan matriks *centroid*  $V^{*(l)}$  menggunakan Persamaan 3.3 dan Persamaan 2.13
- 4) Memperbaharui perhitungan, jika  $\|v^{*(l+1)} - v^{*(l)}\| < \varepsilon$ , maka berhenti, jika tidak maka  $l = l + 1$  dan lanjutkan ke tahapan nomor 2.

### 2.4.3 *Rapid Region Merging*

Metode *Rapid Region Merging* merupakan metode penggabungan daerah yang pada citra yang memiliki kemiripan maksimum. Penggabungan daerah *oversegmentation*, dilakukan berdasarkan perhitungan *maximum similarity* (Chen & Chen, 2014).

Tahapan penggabungan daerah menggunakan algoritma *rapid region merging* dimulai dengan menentukan terlebih dahulu daerah yang *pseudo-blob* dan daerah yang *real-blob*. Selanjutnya daerah yang dianggap sebagai daerah *pseudo-blob* atau daerah yang terbentuk akibat *oversegmentation* akan diproses dengan menghitung nilai kemiripannya dengan daerah *real-blob* yang merupakan tetangganya atau yang memiliki *region adjacent* dengan *pseudo-blob*. Penggabungan daerah *pseudo-blob* dan *real-blob* berdasarkan nilai *maximum similar region*. Daerah tetangga yang memiliki nilai *maximum similarity* akan ditandai sebagai *winner block P*. Berdasarkan Gambar 3.6, *winner block P*, ditandai berdasarkan Persamaan 2.15 dan parameter yang digunakan untuk menghitung kemiripan antara *pseudo-blob* dan *real-blob* adalah *gray level contrasts*  $G(a, b)$ , kemiripan *ridge*  $E(a, b)$  dan *region adjacency*  $A(a, b)$  (Chen & Chen, 2014).

$$\text{Block } P = \text{Max}(\text{Sim}(a, b), \text{Sim}(a, c), \text{Sim}(a, d)) \quad (2.15)$$

Penggabungan daerah menggunakan metode *Rapid Region Merging* dilakukan dengan cara menghitung kemiripan daerah yang dianggap sebagai daerah *oversegmentation* atau *pseudo-blob* dengan sel tetangganya yang dianggap *real-blob*. Untuk menghitung kemiripan antara *cluster*, parameter yang digunakan adalah *gray level contrasts*  $G(a, b)$ , kemiripan *ridge*  $E(a, b)$  dan *region adjacency*  $A(a, b)$  mengikuti Persamaan 3.9. *Gray level contrasts*  $G(a, b)$  dihitung berdasarkan jumlah piksel di region a  $R_a$  dan region b  $R_b$  serta rata-rata keabuan pada region a  $\mu_a$  dan region b  $\mu_b$  mengikuti Persamaan 3.6. Kemiripan *ridge*  $E(a, b)$  dihitung berdasarkan  $\sigma$  rata-rata keabuan piksel di daerah *boundary* pada kedua sisi *ridge* dan  $\pi$  variasi keabuan piksel di daerah *boundary* pada kedua sisi. Jika *ridge* adalah *boundary*, *gray level* piksel pada kedua sisi akan menunjukkan perubahan yang signifikan,  $|\pi_a - \pi_b|$  akan menjadi lebih besar sedangkan  $\sigma_a$  dan  $\sigma_b$  akan menjadi lebih kecil. Sehingga, untuk perhitungan ditambahkan 1 sebagai pembilang dan penyebut untuk menghindari  $|\pi_a - \pi_b|$  dan  $\sigma_a + \sigma_b$  jika salah satu bernilai 0 seperti pada Persamaan 3.7. Untuk *region adjacency*  $A(a, b)$  mengikuti Persamaan 3.8 (Chen & Chen, 2014).

- Ketetanggaan Piksel dan *Adjacency*

Sebuah piksel pada koordinat  $(x, y)$ , memiliki empat ketetanggaan vertikal dan horisontal dimana koordinatnya didapat dari :

$$(x + 1, y), (x - 1, y), (x, y + 1), (x, y - 1) \quad (2.16)$$

Himpunan piksel ini disebut empat ketetanggaan dari  $p$ , yang di notasikan dengan  $N_4(p)$ . Setiap piksel adalah sebuah *unit distance* dari  $(x, y)$  dan beberapa lokasi ketetanggaan dari  $p$  berada di luar citra jika  $(x, y)$  berada pada batas gambar.

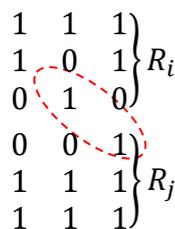
Empat ketetanggaan diagonal dari  $p$  memiliki koordinat mengikuti Persamaan 2.17 yang dinotasikan dengan  $N_D(p)$ . Pada ketetanggaan ini, bersama dengan ketetanggaan empat disebut ketetanggaan 8 dari  $p$ , atau dinotasikan  $N_8(p)$ . Sama seperti sebelumnya, beberapa lokasi ketetanggaan pada  $N_D(p)$  dan  $N_8(p)$  berada di luar citra jika  $(x, y)$  berada pada batas gambar (Gonzalez & Woods, 2002).

$$(x + 1, y + 1), (x + 1, y - 1), (x - 1, y + 1), (x - 1, y - 1) \quad (2.17)$$

Piksel-piksel di dalam region disebut *connected* (memiliki konektivitas) bila mematuhi aturan *adjacency* atau kedekatan piksel. Aturan kedekatan piksel ini memanfaatkan sifat ketetanggaan piksel, sehingga piksel-piksel yang dikatakan *connected* pada dasarnya memiliki sidat *adjacency* satu sama yang lain karena setiap piksel masih memiliki hubungan ketetanggaan. Jika  $V$  adalah himpunan dari nilai-nilai intensitas yang digunakan untuk menyatakan *adjacency*. Pada citra biner,  $V = \{1\}$  jika mengacu untuk *Adjacency* dari piksel-piksel dengan nilai 1. Menurut (Gonzalez & Woods, 2002), terdapat tiga jenis *adjacency*, yaitu:

- 1) 4-*adjacency*, dua piksel  $p$  dan  $q$  dengan nilai dari  $V$  adalah 4-*adjacent* jika  $q$  adalah anggota dari himpunan  $N_4(p)$ .
- 2) 8-*adjacency*, dua piksel  $p$  dan  $q$  dengan nilai dari  $V$  adalah 8-*adjacent* jika  $q$  adalah anggota dari himpunan  $N_8(p)$ .
- 3) m-*adjacency* (*mixed adjacency*), piksel  $p$  dan  $q$  dengan nilai dari  $V$  adalah m-*adjacency* jika
  - (i)  $q$  adalah anggota himpunan  $N_4(p)$ , atau
  - (ii)  $q$  adalah anggota himpunan  $N_D(p)$ , dan himpunan  $q$  adalah anggota himpunan  $N_4(p) \cap N_4(q)$  tidak memiliki piksel yang bernilai  $V$

Jika  $R$  adalah sebuah subset dari piksel-piksel dari sebuah citra,  $R$  disebut sebuah *region* dari citra jika  $R$  adalah himpunan *connected*. Dua *region*,  $R_i$  dan  $R_j$  dikatakan *adjacent* jika piksel-piksel mereka merupakan himpunan *connected*. dua *region* yang tidak *adjacent* disebut *disjoint*. Untuk dua *region* yang dikatakan *adjacency* bisa dilihat pada Gambar 2.4. Berdasarkan tiga definisi *adjacency*, maka contoh Gambar 2.5 hanya dikatakan *adjacent* jika mengikuti sifat 8-*adjacency*.



Gambar 2. 4 Contoh Dua Region yang adjacent dengan 8-Adjacency (Gonzalez & Woods, 2002)

*[Halaman ini sengaja dikosongkan]*

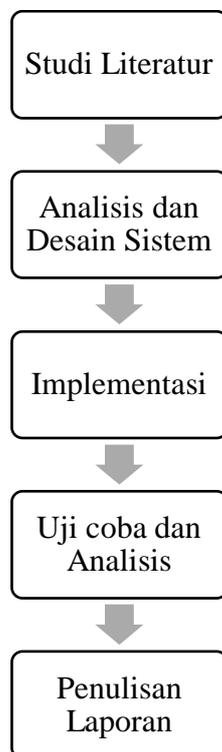
## **BAB III**

### **METODE PENELITIAN**

Pada bab ini akan dijelaskan metode penelitian yang digunakan pada penelitian ini. Beberapa pokok bahasan yang akan dipaparkan adalah meliputi tahapan penelitian, dataset yang digunakan, rancangan sistem, skenario uji coba dan implementasi.

#### **3.1 Tahapan Penelitian**

Secara umum tahapan penelitian ini yang akan dilakukan pada penelitian ini dijelaskan pada Gambar 3.1



Gambar 3. 1 Tahapan Penelitian

##### **3.1.1 Tahapan Studi Literatur**

Tahapan studi literatur adalah tahapan pencarian informasi dan referensi penelitian-penelitian sebelumnya yang terkait dengan informasi diagnosis kanker payudara, pengolahan citra secara umum, metode-metode segmentasi dan

pemisahan sel kanker payudara pada citra mikroskopis sel kanker payudara serta kajian tentang algoritma *Spatial Fuzzy C-Means* dalam melakukan pemisahan sel bertumpuk pada citra mikroskopis kanker payudara dan *Rapid Region Merging* dalam penanganan masalah *oversegmentation*.

### **3.1.2 Tahapan Analisis dan Desain Sistem**

Pada tahap analisis dilakukan perumusan masalah, batasan-batasan masalah dan bagaimana menyelesaikan masalah tersebut. Permasalahan yang diangkat pada penelitian ini adalah bagaimana melakukan segmentasi dan pemisahan sel bertumpuk pada citra mikroskopis sel kanker payudara. Pada tahap desain sistem dilakukan perancangan sistem untuk proses segmentasi dan pemisahan sel bertumpuk pada citra mikroskopis sel kanker payudara berdasarkan hasil tahap analisis. Pada tahapan desain sistem ini, dijelaskan secara lebih detail mengenai data masukan atau dataset, tahapan proses segmentasi dan tahapan pemisahan sel bertumpuk, dan skenario uji coba yang akan dilakukan.

### **3.1.3 Tahapan Implementasi**

Tahapan implementasi adalah tahapan penerapan tahapan desain sistem dengan harapan tahapan ini menghasilkan keluaran yang sesuai dengan tahap analisis. Pada penelitian ini, sistem akan diimplementasikan menggunakan Matlab R2015a.

### **3.1.4 Tahapan Uji Coba dan Analisis Hasil**

Sistem yang telah diimplementasikan sesuai dengan desain sistem akan diuji coba pada tahapan ini. Uji coba dilakukan sesuai dengan skenario uji coba yang dirancang pada tahap desain sistem. Berdasarkan hasil uji coba tersebut sistem nantinya akan dianalisis untuk mengetahui performa dari sistem yang dibangun.

### **3.1.5 Tahapan Penulisan Laporan**

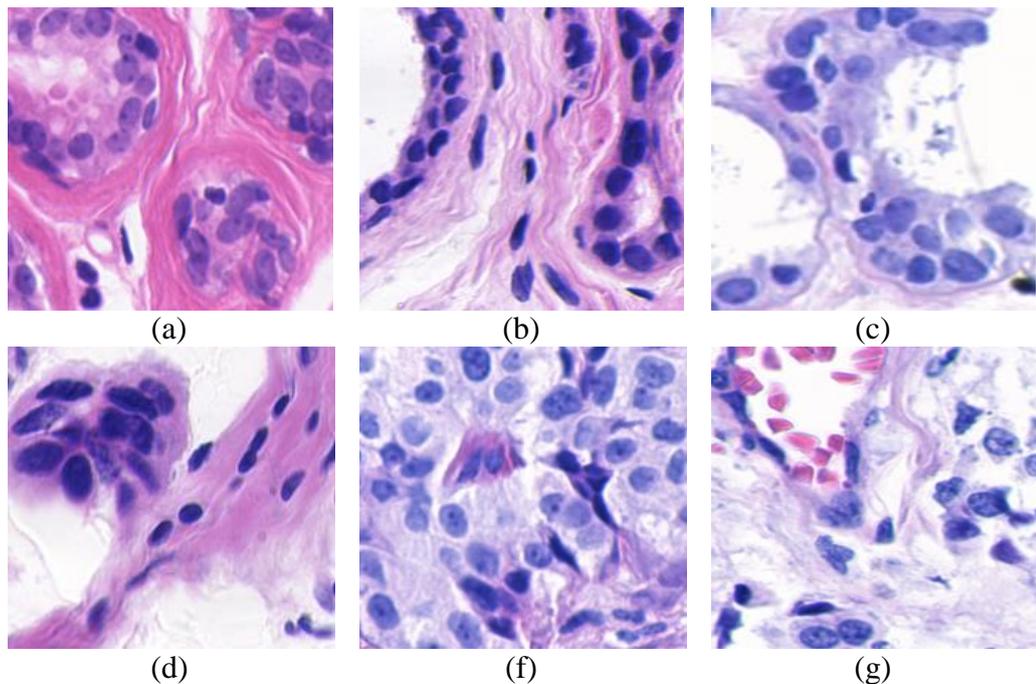
Pada tahapan ini merupakan penulisan laporan dari keseluruhan tahapan yang dilakukan pada penelitian ini. Setiap tahapan akan didokumentasikan secara tertulis sebagai bukti dari hasil penelitian yang dilakukan.

### 3.2 Analisis dan Desain Sistem

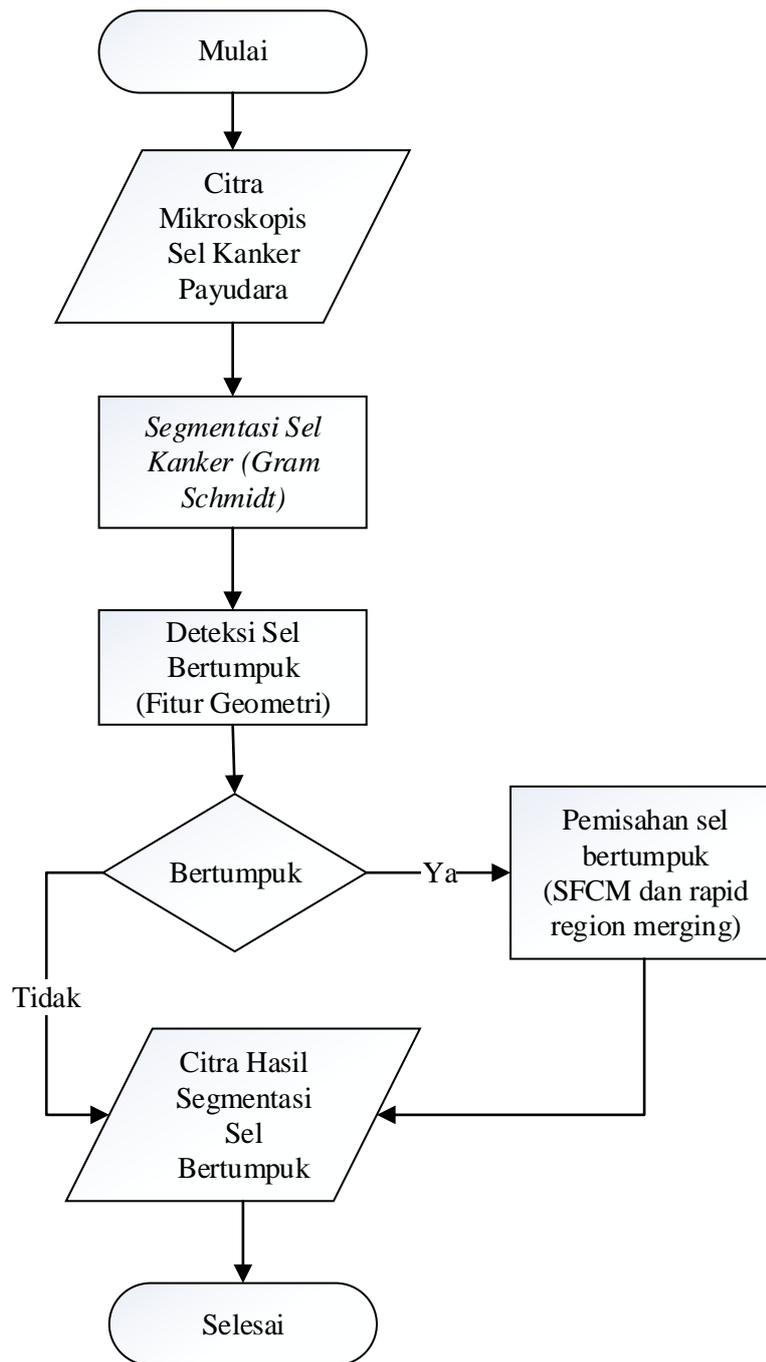
Secara garis besar, proses segmentasi sel bertumpuk pada citra mikroskopis sel kanker payudara terbagi atas dua tahapan penting, yaitu tahapan segmentasi untuk identifikasi atau mengambil objek sel kanker dan tahapan pemisahan sel bertumpuk. Citra yang digunakan pada penelitian ini adalah citra RGB berupa citra mikroskopis sel kanker payudara. Identifikasi sel kanker payudara dilakukan berdasarkan metode *Gram-Schmidt* dan Pemisahan sel kanker bertumpuk diusulkan kombinasi metode *Spatial Fuzzy C-Means* dan *Rapid Region Merging* untuk mengurangi masalah *oversegmentation*. Pada gambar 3.3 digambarkan secara umum desain sistem dan letak kontribusi pada penelitian ini.

#### 3.2.1 Dataset

Data uji coba yang digunakan pada penelitian ini adalah citra mikroskopis sel kanker payudara yang terbagi dalam dua jenis data yaitu data *benign* dan *malignant*, yang berukuran  $896 \times 768$  pixel, seperti pada Gambar 3.2. Dataset yang digunakan dapat diunduh di [www.bioimage.ucsb.edu](http://www.bioimage.ucsb.edu).



Gambar 3. 2 Citra RGB Berupa Citra Mikroskopis Sel Kanker Payudara (a)-(c) Citra Data Benign (d)-(g) Citra Data Malignant



Gambar 3. 3 Gambaran Umum Desain Sistem Segmentasi Sel Bertumpuk Pada Citra Mikroskopis Sel Kanker Payudara

### 3.2.2 Segmentasi Citra Mikroskopis Sel Kanker Payudara

Segmentasi citra mikroskopis sel kanker payudara bertujuan untuk memisahkan sel kanker payudara dengan objek yang bukan sel kanker payudara.

Pada Gambar 3.5 diperlihatkan *flowchart* segmentasi citra mikroskopis sel kanker payudara.

Pada metode *Gram Schmidt*, sebuah vektor bobot dihitung untuk memperkuat vektor warna dari sel kanker dan melemahkan vektor-vektor warna yang tidak diinginkan, seperti *background*. Pada citra mikroskopis sel kanker payudara, warna sel kanker cenderung lebih gelap dari objek lainnya. Nilai vektor warna yang digunakan pada penelitian ini ditetapkan berdasarkan perhitungan rata-rata vektor warna dari tiga dataset yang dipilih secara *random*, yang dapat dilihat pada tabel 3.1. Di mana vektor warna 1 adalah vektor warna yang diperkuat atau diidentifikasi sebagai sel kanker dan vektor warna 2 dan vektor warna 3 adalah vektor warna yang ditentukan untuk melemahkan objek yang dianggap derau dan *background*.

Setelah identifikasi sel kanker maka akan dilakukan segmentasi sel kanker payudara berdasarkan nilai *thresholding*. Nilai ini ditentukan berdasarkan nilai uji coba terbaik yang dilakukan secara *random* terhadap dataset dengan memasukkan nilai dalam rentan 0,01 sampai 0,08 (dapat dilihat pada Gambar 4.1). Tahapan terakhir untuk skema segmentasi sel kanker payudara adalah menerapkan operasi morfologi pada citra hasil *thresholding* untuk menghapus derau.

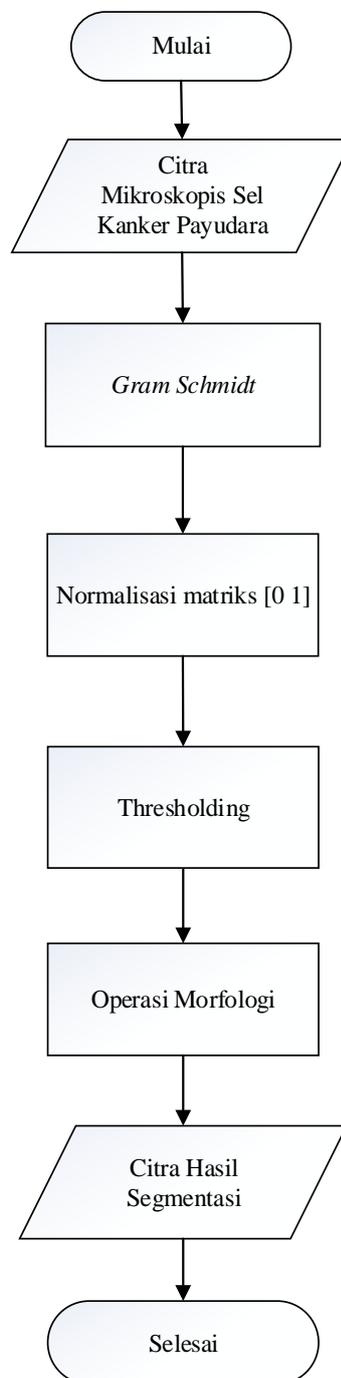
Tabel 3. 1 Vektor Warna RGB pada *Gram-Schmidt*

	Vektor Warna RGB
Vektor Warna 1	[97 65 172]
Vektor Warna 2	[197 145 217]
Vektor Warna 3	[236 182 231]

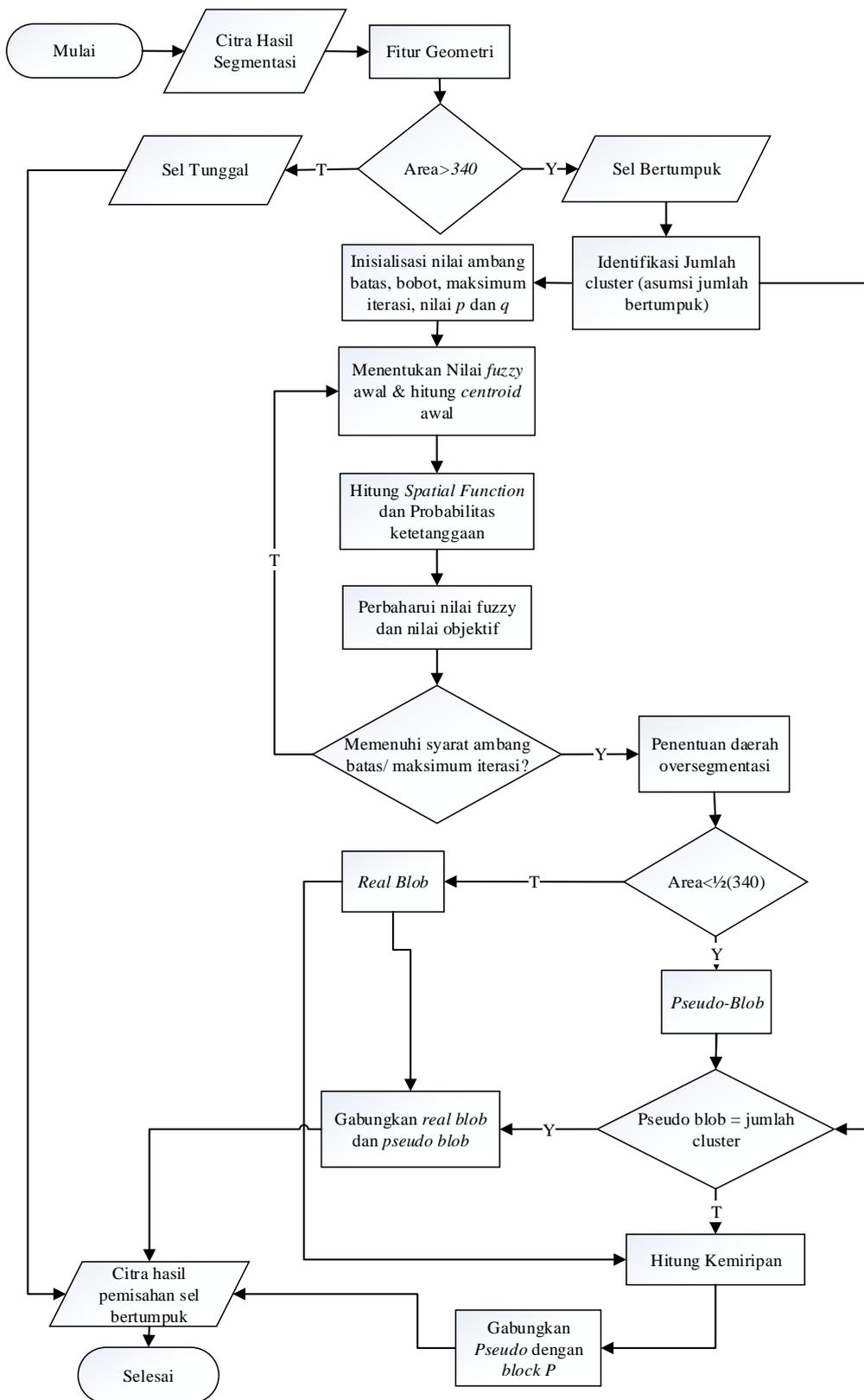
### 3.2.3 Pemisahan Sel Bertumpuk Pada Citra Mikroskopis Sel Kanker Payudara

Citra masukan pada proses pemisahan sel bertumpuk adalah citra hasil segmentasi pada tahap sebelumnya. Langkah awal untuk melakukan pemisahan sel bertumpuk adalah dengan mendeteksi sel kanker payudara bertumpuk. Pendeteksian sel kanker bertumpuk pada citra mikroskopis sel kanker payudara dilakukan berdasarkan informasi fitur geometri sel kanker. Algoritma SFCM

*clustering* digunakan untuk memisahkan sel kanker bertumpuk. Selanjutnya, diterapkan metode *Rapid Region Merging* untuk penggabungan daerah yang *oversegmentation*. Gambar 3.6 menampilkan tahapan dalam proses pemisahan sel kanker bertumpuk menggunakan algoritma SFCM dan *Rapid Region Merging*.



Gambar 3. 4 *Flowchart* Identifikasi Sel



Gambar 3. 5 FLOWchart Pemisahan Sel Kanker Bertumpuk

### 3.2.3.1 Deteksi Sel Bertumpuk

Deteksi sel bertumpuk pada penelitian ini dilakukan berdasarkan informasi fitur geometri. Fitur geometri yang digunakan adalah informasi area atau luas objek yang merupakan jumlah piksel dari objek yang menunjukkan ukuran dari sel kanker payudara. Pada penelitian ini, penentuan sel kanker bertumpuk adalah dengan menentukan nilai *threshold*. Setelah dilakukan analisis fitur geometri area terhadap seluruh dataset dan dilakukan uji coba pada beberapa dataset maka ditentukan fitur geometri area yang digunakan sebagai nilai *threshold* adalah nilai fitur geometri area yang ditentukan berdasarkan uji coba. Setiap objek yang memiliki nilai fitur area di atas dari nilai *threshold* akan dianggap sebagai sel kanker bertumpuk dan yang memiliki nilai sama dengan atau kurang dari fitur area adalah sel tunggal.

### 3.2.3.2 Pemisahan Sel Bertumpuk

Algoritma *Spatial Fuzzy C-Means Clustering* (SFCM) merupakan algoritma yang dikembangkan dari *fuzzy c-means clustering* tradisional. Algoritma SFCM mengembangkan informasi spasial yang dibentuk menggunakan distribusi statistika dari ketetanggaan piksel berdasarkan perhitungan probabilitas yang tidak hanya *robust* terhadap *noise* tapi juga mengurangi kesalahan pengelompokan piksel. *Spatial function* didefinisikan seperti pada Persamaan 3.1, Di mana  $x_j$  merepresentasikan nilai fitur,  $\Omega$  merepresentasikan sebuah jendela persegi yang titik tengahnya berada pada piksel  $x_j$ .  $h_{ij}$  merepresentasikan probabilitas piksel  $x_j$  menjadi anggota sebuah klaster. Pada *spatial function*,  $c$  merupakan jumlah *cluster* yang diinginkan,  $P_{ij}$  adalah faktor kontribusi dari tetangga  $x_t$  dan bobot  $u_{it}$  adalah *membership pattern*  $x_t$ . Secara umum, *prior probability* dari  $p_{ij}$  merupakan hasil perbandingan dari  $N_j^i$  jumlah ketetanggaan piksel yang menjadi anggota *cluster* dan  $N_j$  jumlah ketetanggaan piksel pada  $x_j$ , yang dinotasikan kedalam Persamaan 3.2 (Li & Shen, 2010).

$$h_{ij} = 1 + \left( \frac{\sum_{t \in \Omega_j} u_{it} p_{ij}}{\sum_{k=1}^c \sum_{t \in \Omega_j} u_{kt} p_{kj}} \right) \quad (3.1)$$

$$p_{ij} = \frac{N_j^i}{N_j} \quad (3.2)$$

Untuk *fuzzy membership* dihitung mengikuti Persamaan 3.3. di mana  $p$  dan  $q$  adalah parameter untuk mengontrol *relative importance* dari kedua *function*. Pada penelitian ini nilai  $p$  dan  $q$  adalah 1. Untuk penentuan titik pusat atau *centroid* ditetapkan berdasarkan pada Persamaan 2.13 dan untuk *cost function* dihitung menggunakan persamaan 3.4 (Li & Shen, 2010).

Penentuan jumlah *cluster* pada penelitian ini adalah berdasarkan perhitungan sesuai dengan Persamaan 3.5, di mana  $a$  adalah fitur geometri area bertumpuk,  $b$  adalah nilai *threshold* bertumpuk dan  $c$  adalah nilai *threshold* sebesar. Nilai  $c$  akan ditentukan berdasarkan hasil uji coba terhadap dataset.

$$u_{ij}^* = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^c u_{kj}^p h_{kj}^q} \quad (3.3)$$

$$J_m^*(U, V) = \sum_{i=1}^c \sum_{j=1}^n (U_{ij}^*)^m (d_{ij}^*)^2 \quad (3.4)$$

$$k = a / \left(\frac{b}{c}\right) \quad (3.5)$$

Tahapan dalam metode *Spatial Fuzzy C-Means* yang diterapkan pada penelitian ini, dapat disimpulkan ke dalam tahapan-tahapan berikut.

1. Inisialisasi : Menentukan jumlah *cluster* berdasarkan asumsi banyaknya sel tunggal yang menumpuk pada daerah yang dianggap sel bertumpuk dengan mengikuti perhitungan seperti pada persamaan 3.5, menentukan bobot pangkat ( $w \geq 1$ ), menentukan jumlah maksimal iterasi, menentukan nilai ambang batas perubahan fungsi objektif
2. Memberikan nilai awal pada matriks *fuzzy pseudo-partition*, dengan syarat seperti pada persamaan 2.10. dan hitung nilai *centroid* awal menggunakan persamaan 2.13
3. Menghitung *spatial function* dan *prior probability* ketetangaan menggunakan persamaan 3.1 dan 3.2
4. Menghitung kembali keanggotaan matriks *fuzzy pseudo-partition* menggunakan persamaan 3.3 dan memperbaharui nilai *centroid*

5. Memperbaharui perhitungan, jika memenuhi syarat (1) maksimum iterasi, atau (2) perubahan nilai objektif di bawah nilai ambang batas, maka berhenti. Jika tidak kembali ke tahapan ke 2.

### 3.2.3.3 Penggabungan Daerah

Penggabungan daerah menggunakan metode *Rapid Region Merging* dilakukan dengan cara menghitung kemiripan daerah yang dianggap sebagai daerah *oversegmentation* atau *pseudo-blob* dengan sel tetangganya. Pada penelitian ini, semua sel bertumpuk yang telah dikelompokkan akan diproses tahapan penggabungan daerah. Langkah awal pada metode ini adalah penentuan daerah *pseudo-blob*, daerah *pseudo-blob* adalah daerah yang dihasilkan akibat *oversegmentation* dari hasil pemisahan sel bertumpuk berdasarkan *Spatial Fuzzy C-Means*. Daerah *pseudo-blob* ditentukan berdasarkan nilai *threshold*, semua sel yang memiliki nilai fitur luas di bawah dari nilai *threshold* ditandai sebagai *pseudo-blob*. Selanjutnya, akan dicari daerah pemenang atau *Block P* yang merupakan daerah tetangga yang memiliki kemiripan paling tinggi dengan *pseudo-blob* dibandingkan dengan anggota tetangganya yang lain. Kemudian daerah yang ditandai sebagai *pseudo-blob* akan digabungkan dengan *block P*.

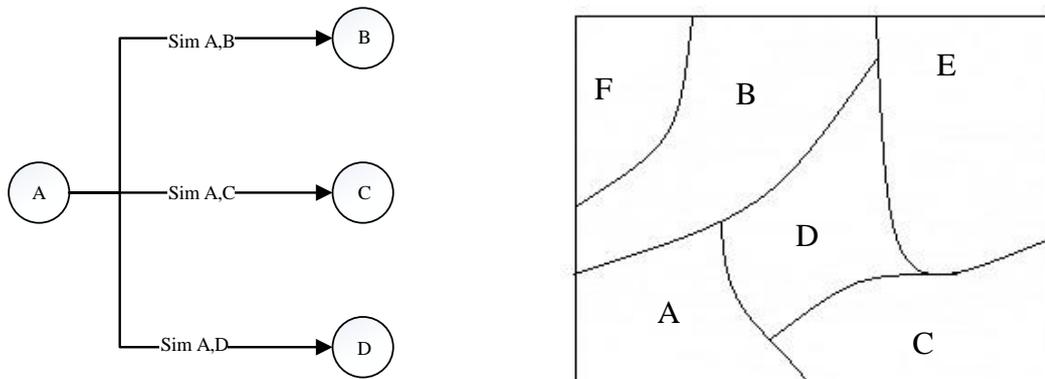
Pada penelitian ini, ditambahkan sebuah parameter baru untuk penggabungan daerah, yaitu jika jumlah *pseudo-blob* pada daerah sel bertumpuk sama banyaknya dengan jumlah *cluster*, maka daerah sel bertumpuk tersebut adalah sel tunggal. Sehingga semua daerah tersebut akan digabungkan. Contoh pada Gambar 3.6, di ilustrasikan jika A adalah sebuah daerah yang ditandai sebagai *pseudo-blob* maka A akan dibandingkan kemiripannya dengan daerah ketetanggaannya yaitu daerah B, C dan D. Region yang mendapatkan *Maximum similarity* atau kemiripan paling maksimum akan menjadi *winner block P* dan akan digabungkan daerahnya dengan daerah A. Jika daerah A,B,C,D,E dan F adalah daerah yang nilai fitur areanya di bawah dari nilai *threshold* maka daerah A,B,C,D,E dan F adalah sel tunggal dan digabungkan menjadi satu daerah.

Untuk menghitung kemiripan antara *cluster*, parameter yang digunakan adalah *gray level contrasts*  $G(a, b)$ , kemiripan *ridge*  $E(a, b)$  dan *region adjacency*  $A(a, b)$  mengikuti Persamaan 3.9. *Gray level contrasts*  $G(a, b)$  dihitung

berdasarkan jumlah piksel di region a  $R_a$  dan region b  $R_b$  serta rata-rata keabuan pada region a  $\mu_a$  dan region b  $\mu_b$  mengikuti Persamaan 3.6. Kemiripan *ridge*  $E(a, b)$  dihitung berdasarkan  $\sigma$  rata-rata keabuan piksel di daerah *boundary* pada kedua sisi *ridge* dan  $\pi$  variasi keabuan piksel di daerah *boundary* pada kedua sisi. Jika *ridge* adalah *boundary*, *gray level* piksel pada kedua sisi akan menunjukkan perubahan yang signifikan,  $|\pi_a - \pi_b|$  akan menjadi lebih besar sedangkan  $\sigma_a$  dan  $\sigma_b$  akan menjadi lebih kecil. Sehingga, untuk perhitungan ditambahkan 1 sebagai pembilang dan penyebut untuk menghindari  $|\pi_a - \pi_b|$  dan  $\sigma_a + \sigma_b$  jika salah satu bernilai 0 seperti pada Persamaan 3.7. Untuk *region adjacency*  $A(a, b)$  mengikuti Persamaan 3.8 (Chen & Chen, 2014).

$$G(a, b) = \frac{R_a - R_b}{R_a + R_b} (\mu_a - \mu_b)^2 \quad (3.6)$$

$$E(a, b) = \frac{|\pi_a - \pi_b| + 1}{\sigma_a + \sigma_b + 1} \quad (3.7)$$



Gambar 3. 6 Aturan Penggabungan Daerah Menggunakan Metode Rapid Region Merging (Chen & Chen, 2014)

$$A(a, b) = \begin{cases} 1 & a \text{ dan } b \text{ adjacent} \\ \infty & a \text{ dan } b \text{ tidak adjacent} \end{cases} \quad (3.8)$$

$$Sim(a, b) = \frac{1}{G(a, b) \cdot E(a, b) \cdot A(a, b)} \quad (3.9)$$

### 3.3 Skenario Uji Coba

Pada penelitian ini dilakukan beberapa skenario uji coba, antara lain :

1. Mengukur performa segmentasi metode yang diusulkan dengan menghitung nilai MSE berdasarkan *groundtruth*
2. Membandingkan metode yang diusulkan yaitu kombinasi metode *Spatial Fuzzy C-Means Clustering* dan *Rapid Region Merging* dengan 2 metode lain, yaitu metode *Spatial Fuzzy C-Means* dan metode *K-Means Clustering* untuk mengukur performa pemisahan sel bertumpuk pada citra mikroskopis sel kanker payudara berdasarkan perhitungan jumlah sel kanker payudara secara manual. Untuk mengukur seberapa signifikan peningkatan performa metode yang diusulkan dilakukan perhitungan *uji – t* dalam pengujian hipotesis rata-rata dua sampel yang berpasangan.

### 3.4 Analisis Hasil

Berdasarkan hasil uji coba skenario pertama, maka akan dilakukan pengukuran performa metode yang diusulkan dengan menghitung nilai *Mean Square Error* (MSE) atau rata-rata *error*. Nilai MSE merupakan hasil perbandingan nilai selisih piksel-piksel pada citra asal  $f(x, y)$  dengan piksel-piksel pada citra hasil  $g(x, y)$  di posisi piksel yang sama. Semakin kecil nilai MSE yang diperoleh maka semakin tinggi akurasi segmentasi dan pemisahan sel bertumpuk pada citra mikroskopis sel kanker payudara. Secara matematis perhitungan nilai MSE dapat ditulis sebagai berikut :

$$MSE = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [f(x, y) - g(x, y)]^2 \quad (3.9)$$

Analisis hasil untuk uji coba skenario kedua adalah membandingkan rata-rata nilai akurasi dari metode yang diusulkan dengan metode SFCM dan metode *K-Means*. Nilai akurasi dihitung berdasarkan rasio jumlah sel kanker yang diidentifikasi dengan benar oleh metode yang diusulkan dengan jumlah sel kanker pada citra asli. Sel kanker yang diidentifikasi benar adalah sel kanker yang terpisah sesuai dengan jumlah sel bertumpuk pada citra asli dengan memperhatikan garis pemisah antar *cluster*. Nilai *oversegmentation* dihitung berdasarkan jumlah *cluster* yang lebih banyak dari jumlah sel pada citra asli dan berdasarkan kondisi garis pemisah antar *cluster* sedangkan *undersegmentation* dihitung berdasarkan

jumlah *cluster* yang kurang dari jumlah sel bertumpuknya yang sebenarnya, seperti yang dijelaskan pada Gambar 4.6.

Untuk mengetahui ada tidaknya perbedaan yang signifikan antara rata-rata akurasi metode yang diusulkan dengan 2 metode pembandingan, dilakukan perhitungan uji statistik *uji - t*. *uji - t* dikembangkan oleh William Seely Gosset yang adalah seorang konsultan statistika Irlandia pada tahun 1915. Penggunaan *uji - t* dapat dibedakan menjadi (Hartono, 2004) :

- 1) *uji - t* untuk sampel kecil dan besar yang berkorelasi
- 2) *uji - t* untuk sampel kecil dan besar yang tidak berkorelasi

Besarnya koefisien komparatif dengan menggunakan *uji - t* diberi simbol *uji - t* diberi simbol  $t_0$  ( $t_{observasi}$ ) dihitung mengikuti persamaan 3.10, angkanya dapat bertanda positif dan negatif. Namun tanda negatif bukanlah tanda aljabar, sehingga tanda positif dan negatif memiliki arti yang sama.

Cara memberikan interpretasi terhadap  $t_0$  adalah dengan merumuskan hipotesis alternatif ( $H_a$ ) yang menyatakan ada perbedaan dan hipotesis nol ( $H_0$ ) menyatakan tidak ada perbedaan. Pada penelitian dirumuskan hipotesis berikut ini; hipotesis nol ( $H_0$ ) adalah  $t_0 = t_{tabel}$  yang menyatakan tidak ada perbedaan antara variabel pertama dan variabel kedua dan hipotesis alternatif ( $H_a$ ) adalah  $t_0 \neq t_{tabel}$  yang menyatakan ada perbedaan yang signifikan antara variabel pertama dan variabel kedua.

Setelah mencari  $df$  dan  $db$ , lalu dengan besarnya  $df$  dan  $db$  tersebut berkonsultasi pada tabel nilai  $t$  yang hasilnya disebut  $t_{tabel}(t_t)$ . Selanjutnya bandingkan  $t_0$  dan  $t_{tabel}$  dengan ketentuan (Hartono, 2004):

- 1) Bila  $t_0$  sama dengan atau lebih besar dari  $t_{tabel}$  maka hipotesis nol ( $H_0$ ) ditolak, yang berarti ada perbedaan yang signifikan.
- 2) Bila  $t_0$  lebih kecil dari  $t_{tabel}$  maka hipotesis ( $H_0$ ) diterima, yang berarti tidak ada perbedaan yang signifikan.

Teknik analisis *uji - t* yang digunakan pada penelitian ini adalah *uji - t* untuk sampel besar dengan nilai  $N \geq 30$  yang berkorelasi. Sampel-sampel yang berkorelasi yang dimaksud pada analisis ini adalah nilai atau skor dari kedua sampel

diambil dari subjek yang sama. Proses analisis statistik *uji – t* dapat dijelaskan sebagai berikut (Hartono, 2004):

- 1) Menghitung harga  $t_0$  sesuai persamaan 3.10, dengan mencari nilai *mean* variabel x  $M_x$  dan nilai *mean*  $M_y$ , dan standar deviasi perbedaan skor dari kedua variabel

$$t_0 = \frac{M_x - M_y}{\sqrt{\left(\frac{SD_x}{\sqrt{N-1}}\right)^2 + \left(\frac{SD_y}{\sqrt{N-1}}\right)^2 - [2 r_{xy}] \left[\frac{SD_x}{\sqrt{N-1}}\right] \left[\frac{SD_y}{\sqrt{N-1}}\right]}} \quad (3.10)$$

- 2) Memberi interpretasi terhadap  $t_0$  dengan mencari nilai *df* dengan perhitungan banyaknya sampel  $N - 1$  kemudian dikorelasikan dengan harga kritik tabel atau  $t_{tabel}$  pada taraf signifikan 5% dan pada taraf signifikan 1%
- 3) Membandingkan  $t_0$  dan  $t_{tabel}$  dengan ketentuan yang dijelaskan sebelumnya
- 4) Memberikan Kesimpulan apakah ada perbedaan yang signifikan antara kedua metode yang dibandingkan berdasarkan hasil perbandingan pada proses sebelumnya.

*[Halaman ini sengaja dikosongkan]*

## **BAB IV**

### **UJI COBA DAN ANALISIS HASIL**

Pada bab ini akan dijelaskan setiap tahapan pengujian sesuai dengan skenario uji coba dan analisis hasil dari penelitian yang telah dilakukan. Tahapan implementasi metode segmentasi sel kanker bertumpuk dibagi atas dua tahapan penting, yaitu segmentasi sel kanker berdasarkan identifikasi sel menggunakan metode *Gram-Schmidt* dan pemisahan sel bertumpuk berdasarkan kombinasi metode *Spatial Fuzzy C-Means* dan *Rapid Region Merging* yang sesuai dengan desain sistem pada Gambar 3.3. Uji coba akan dilakukan sesuai dengan skenario yang telah dijelaskan pada sub bab 3.3

#### **4.1 Lingkungan Uji Coba**

Pada penelitian ini data diuji dengan menggunakan komputer Acer dengan spesifikasi *processor Inter® Core™ i3-2130 CPU @ 3.40 GHz* 3.40 GHz, kapasitas *memory* (RAM) 4.0. Perangkat lunak yang digunakan pada penelitian ini adalah Windows 7 Professional dan *tools* program matlab R2015a.

#### **4.2 Data Uji Coba**

Dataset yang digunakan pada penelitian ini adalah citra RGB yang berupa citra mikroskopis sel kanker payudara, sejumlah 40 citra yang diunduh di [www.bioimage.ucsb.edu](http://www.bioimage.ucsb.edu). Dataset yang tersedia terbagi dalam data *benign* dan data *malignant*. Kemudian, untuk mempermudah proses identifikasi sel kanker payudara dilakukan *cropping* manual pada setiap citra sesuai dengan citra *groundtruth* yang didapat, menggunakan *tools* Adobe Photoshop CS3 (dapat dilihat pada Gambar 3.2).

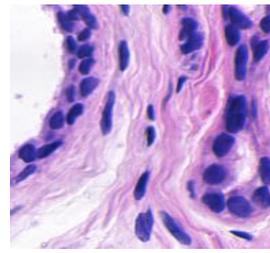
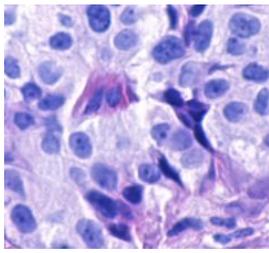
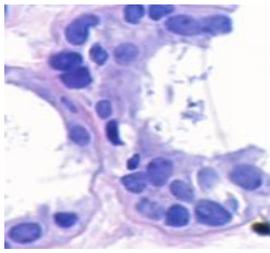
#### **4.3 Skenario Uji Coba**

Untuk mengukur performa metode yang diusulkan maka akan dilakukan beberapa uji coba. Pada penelitian ini, terdapat dua skenario uji coba yang diterapkan. Pertama, melakukan uji performa hasil segmentasi terhadap citra *groundtruth* berdasarkan perhitungan nilai *Mean Square Error* (MSE) untuk menghitung nilai rata-rata kesalahan, artinya semakin kecil nilai MSE yang

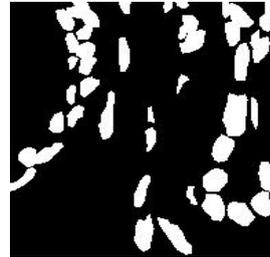
diperoleh maka semakin baik hasil segmentasi yang dihasilkan. Kedua, melakukan uji performa terhadap hasil pemisahan sel bertumpuk, yaitu dengan melakukan perhitungan jumlah sel secara manual terhadap citra hasil pemisahan sel bertumpuk. selanjutnya akan dibandingkan hasil pemisahan sel bertumpuk antara metode yang diusulkan Kombinasi metode *Spatial Fuzzy C-Means* dan *Rapid Region Merging* dengan dua metode pembanding yaitu *Spatial Fuzzy C-Means* dan *K-Means Clustering*. Untuk metode pembanding *K-Means Clustering*, penentuan jumlah *cluster* disesuaikan dengan metode yang diusulkan. Setelah hasil perbandingan didapatkan, akan dilakukan perhitungan statistik uji-t terhadap kedua nilai hasil perbandingan untuk mengetahui apakah ada perbedaan yang signifikan antara hasil metode yang diusulkan dengan metode pembanding.

#### **4.3.1 Hasil Uji Coba Skenario Pertama**

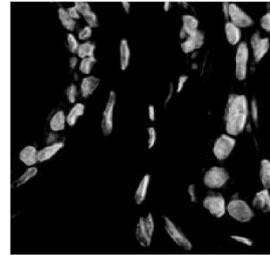
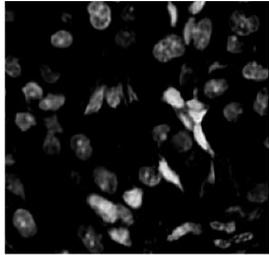
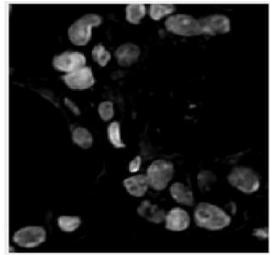
Uji coba pada skenario pertama digunakan untuk mengukur performa metode yang digunakan sesuai dengan tahapan yang dijelaskan pada bab sebelumnya (Gambar 3.5). Citra RGB berupa citra mikroskopis sel kanker payudara adalah citra masukan. Sel kanker yang akan diidentifikasi adalah *blue nuclei* atau sel yang berwarna biru dan biru keunguan yang dijelaskan dalam Gambar 4.2(a). Identifikasi sel kanker payudara diperoleh berdasarkan metode *Gram Schmidt*. Berdasarkan penelitian sebelumnya, pada penelitian ini digunakan nilai vektor warna yang lebih gelap untuk memperkuat atau mendapatkan sel kanker dan menggunakan nilai vektor yang cenderung terang untuk melemahkan derau. Nilai vektor warna yang digunakan pada penelitian ini ditentukan secara manual dan dapat dilihat pada tabel 3.1 . Untuk mendapatkan citra segmentasi, dilakukan *thresholding* dengan melakukan normalisasi data dari *range* 0 sampai 1 terlebih dahulu dan menerapkan operasi morfologi setelah *thresholding*. Berdasarkan uji coba terhadap dataset yang dipilih secara acak, nilai 0,03 digunakan sebagai nilai *thresholding* sesuai dengan hasil yang disajikan pada gambar 4.1, yang mana gambar 4.1(f) mampu memberikan representasi citra asli yang paling baik. Hasil citra yang diperoleh pada tahapan-tahapan identifikasi sel dapat dilihat pada Gambar 4.2.



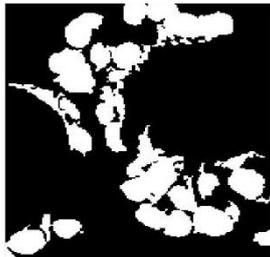
(a)



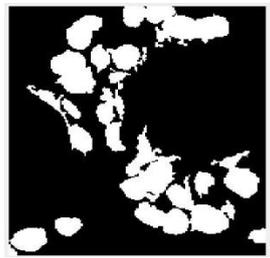
(b)



(c)



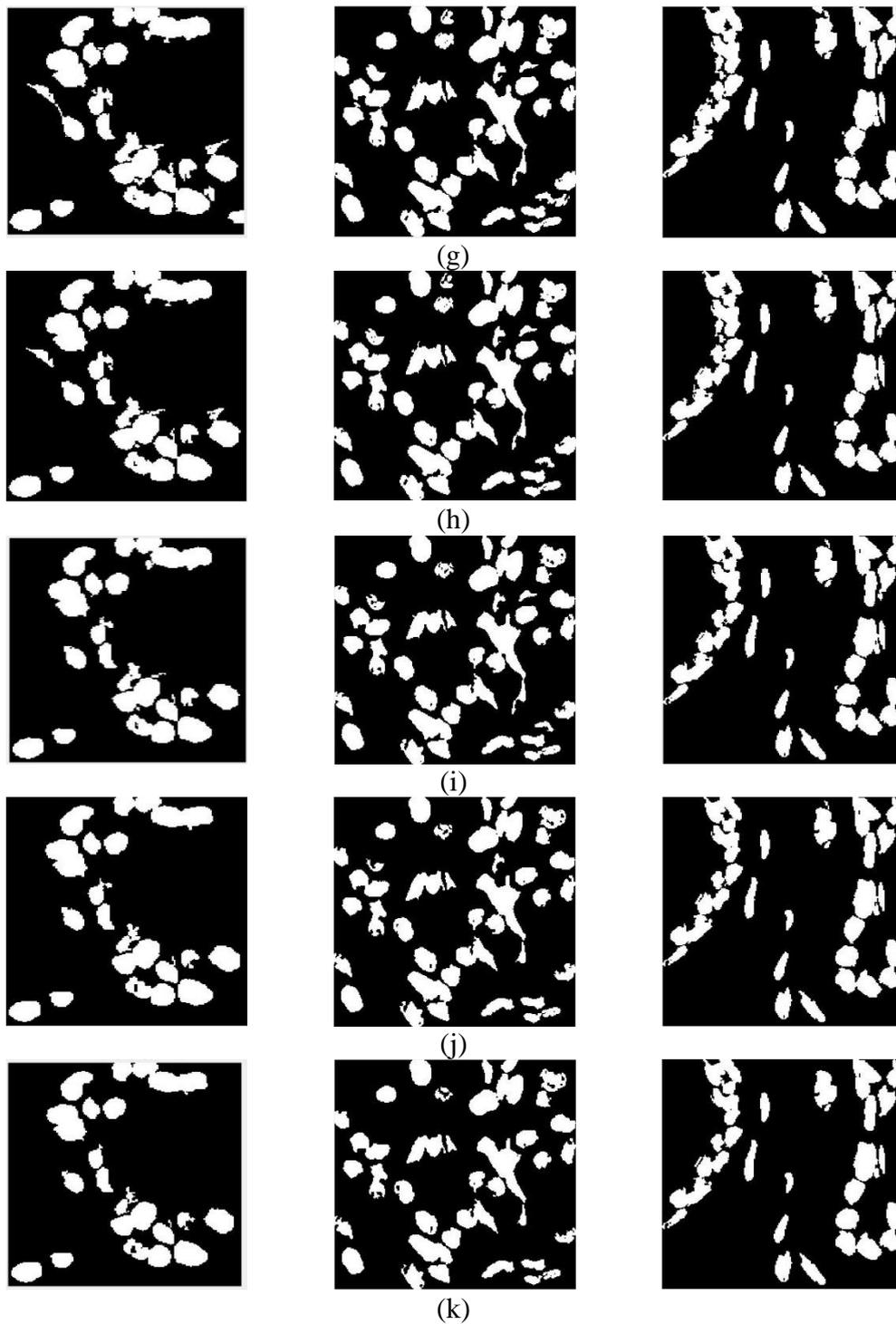
(d)



(e)

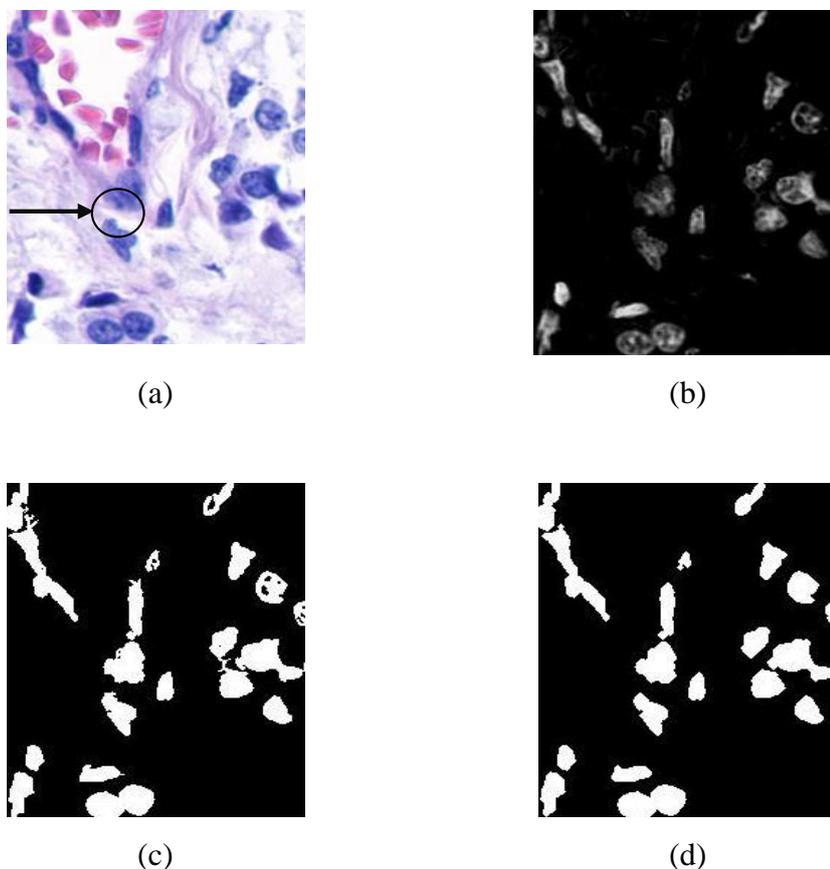


(f)

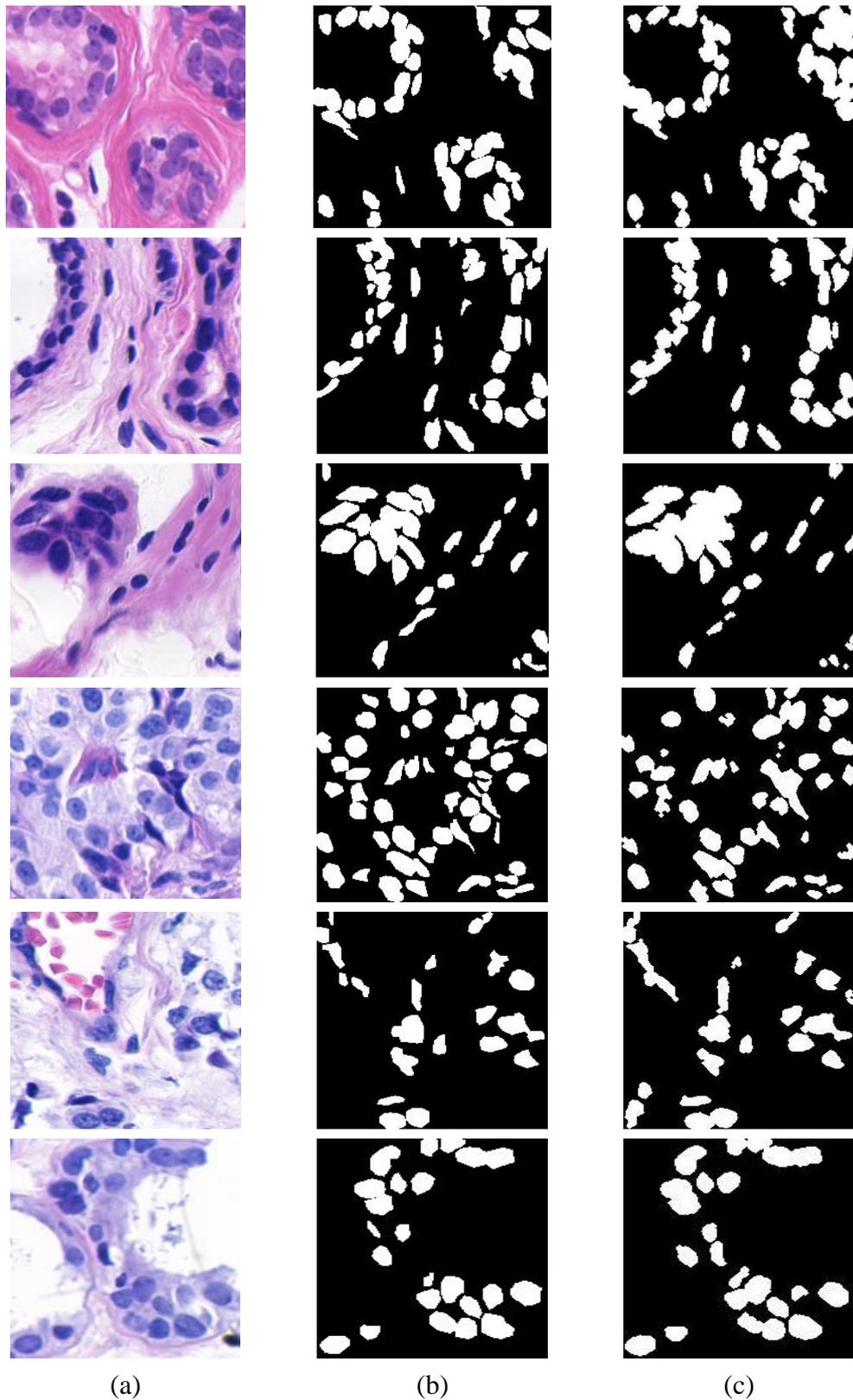


Gambar 4. 1 Penentuan nilai *thresholding* untuk segmentasi sel kanker (a) Citra Asli (b) Citra Groundtruth (c) Identifikasi Sel Kanker Berdasarkan *Gram-Schmidt* (d) Citra hasil *thresholding* 0,01 (e) Citra hasil *thresholding* 0,02 (f) Citra hasil *thresholding* 0,03 (g) Citra Hasil *thresholding* 0,04 (h) Citra hasil *thresholding* 0,05 (i) Citra hasil *thresholding* 0,06 (j) Citra hasil *thresholding* 0,07 (k) Citra hasil *thresholding* 0,08

Berdasarkan uji coba yang dilakukan terhadap 40 citra mikroskopis sel kanker payudara, hasil uji coba skenario pertama ditunjukkan pada Gambar 4.3 dengan hasil performa segmentasi yang diukur menggunakan perhitungan *Mean Square Error* (MSE). Nilai MSE merupakan hasil perbandingan nilai selisih piksel-piksel pada citra *groundtruth* dengan piksel-piksel pada citra hasil segmentasi di posisi piksel yang sama. Semakin kecil nilai MSE yang diperoleh maka semakin tinggi akurasi segmentasi sel pada citra mikroskopis sel kanker payudara. Pada Tabel 4.1 disajikan hasil perhitungan nilai MSE hasil segmentasi dengan metode *Gram-Schmidt* dan hasil segmentasi citra menggunakan metode *Region Based Active Contour*.



Gambar 4. 2 Citra yang Dihasilkan pada Setiap Tahapan Segmentasi (a) Citra RGB sebagai citra masukan dan Blue Nuclei yang harus diidentifikasi sebagai sel kanker payudara (b) Identifikasi Blue Nuclei Berdasarkan Gram-Schmidt (c)Citra Hasil thresholding (d) Citra Hasil Thresholding setelah Operasi Morfologi



(a) (b) (c)  
 Gambar 4. 3 Perbandingan Citra Hasil Segmentasi Dengan Citra Asli dan Citra Groundtruth (a) Citra Asli (b) Citra Groundtruth (c) Citra Hasil Segmentasi

Tabel 4. 1 Perhitungan Nilai Mean Square Error (MSE)

Citra Ke	Nilai MSE	
	<i>Gram-Schmidt</i>	<i>Region Based Active Contour</i>
1	0,0622	0,8855
2	0,0719	0,0816
3	0,0822	0,8646
4	0,1054	0,9002
5	0,0989	0,9158
6	0,0850	0,9119
7	0,0458	0,0610
8	0,0990	0,9045
9	0,0552	0,9346
10	0,0461	0,9578
11	0,0398	0,9177
12	0,0541	0,0728
13	0,0844	0,1384
14	0,0827	0,1706
15	0,0604	0,1383
16	0,0854	0,1552
17	0,0879	0,8385
18	0,0418	0,9061
19	0,1035	0,8658
20	0,0859	0,7977
21	0,0851	0,8573
22	0,0552	0,1959
23	0,0899	0,1626
24	0,0642	0,2253
25	0,0794	0,8363
26	0,0589	0,8411
27	0,0585	0,8836
28	0,0960	0,7324
29	0,0409	0,7505
30	0,0662	0,8508
31	0,0782	0,9105
32	0,0785	0,0991
33	0,1080	0,8859
34	0,0368	0,8955
35	0,0785	0,9123
36	0,0647	0,8650
37	0,0958	0,1352
38	0,0770	0,8879
39	0,1225	0,8311
40	0,0335	0,8428
Rata-rata=	0,0736	0,6505

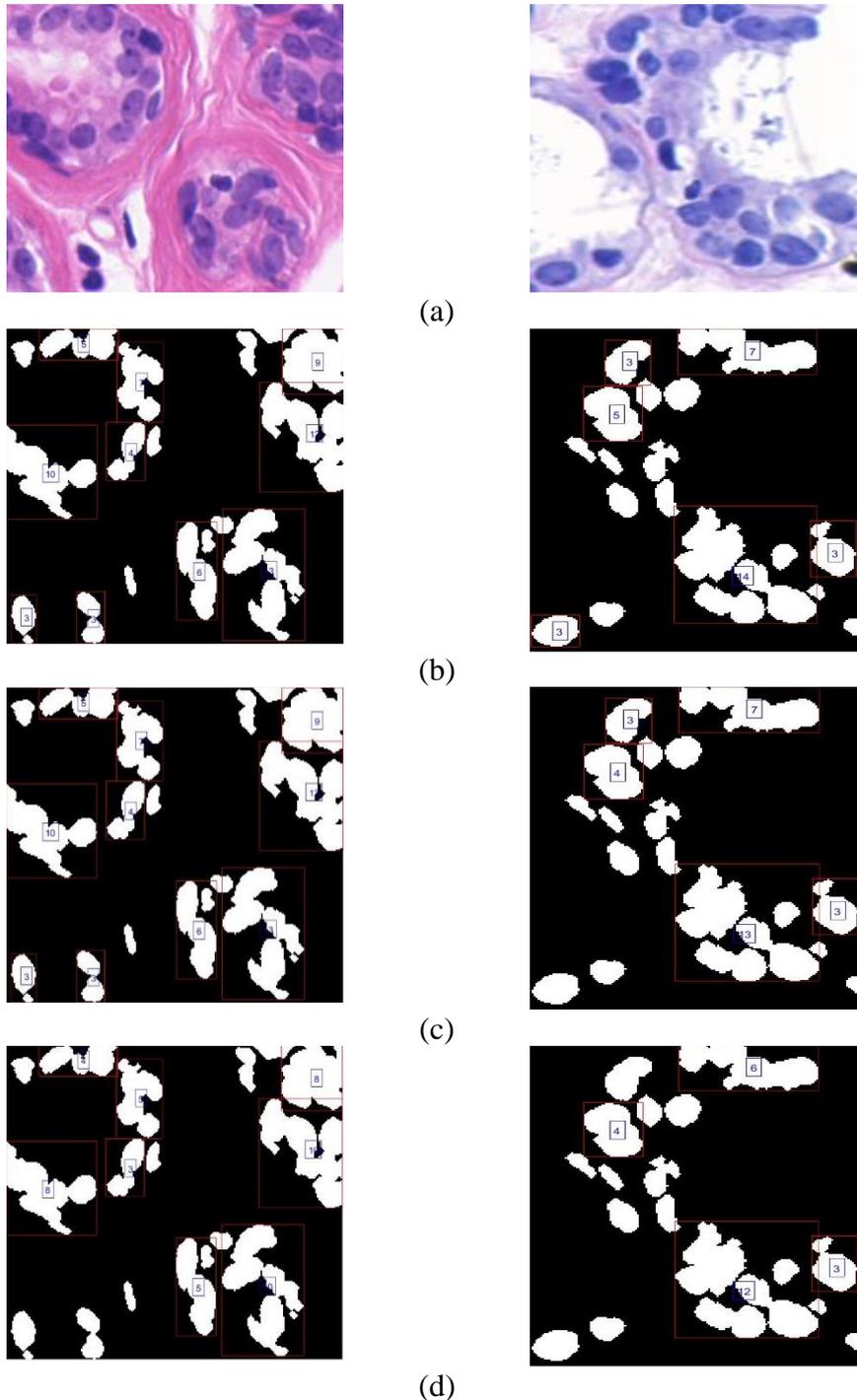
### 4.3.2 Hasil Uji Coba Skenario Kedua

Pengujian skenario kedua dilakukan untuk mengukur performa metode yang diusulkan dalam melakukan pemisahan sel kanker payudara bertumpuk. Proses pemisahan sel kanker payudara bertumpuk telah dijelaskan pada bab sebelumnya (Gambar 3.5) dan hasil dari setiap prosesnya dapat dilihat pada Gambar 4.4. Data masukan yang digunakan pada tahapan ini adalah citra hasil segmentasi berupa citra biner yang diperoleh dari tahapan sebelumnya yaitu tahapan identifikasi dan segmentasi sel kanker.

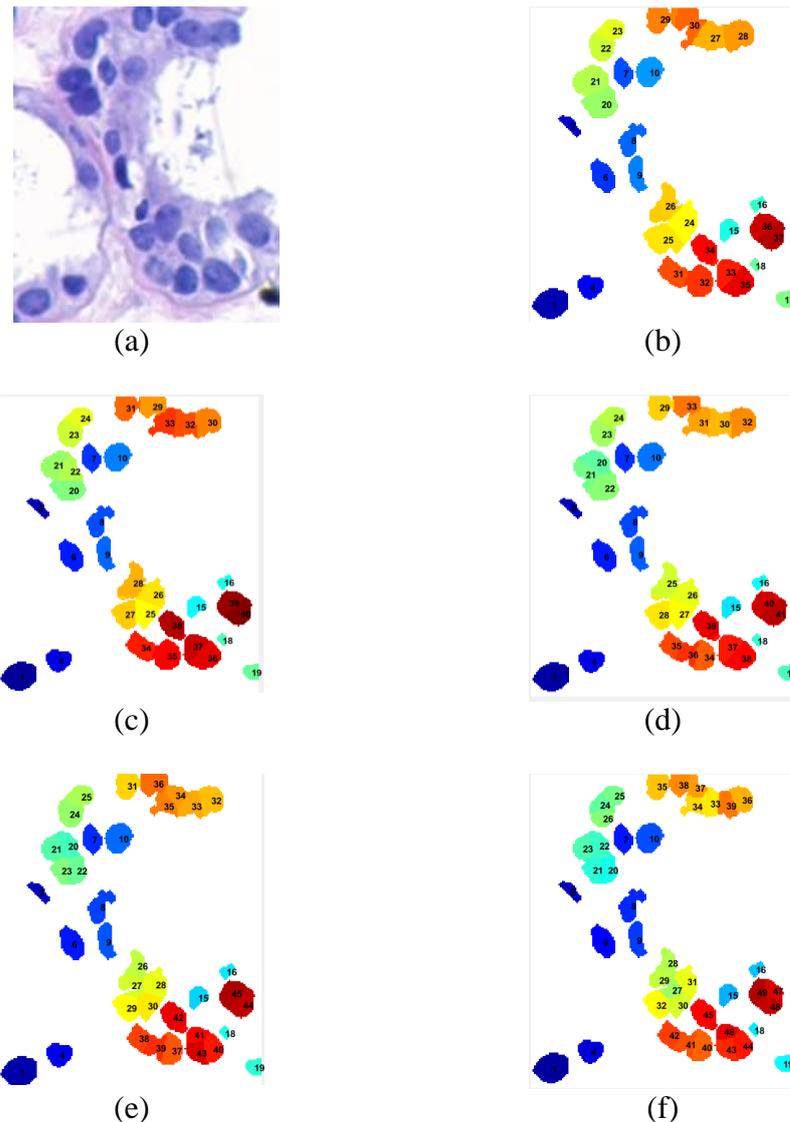
Langkah pertama yang dilakukan pada tahapan pemisahan sel bertumpuk adalah melakukan deteksi sel bertumpuk menggunakan fitur geometri area. Penentuan luasan daerah yang menjadi nilai *threshold* bertumpuk dilakukan berdasarkan uji coba terhadap 20 dataset yang dipilih secara acak. Berdasarkan uji coba yang dilakukan terhadap dataset, nilai *threshold* 340 menghasilkan deteksi bertumpuk paling baik di mana dari 20 dataset hanya menghasilkan kesalahan deteksi bertumpuk pada 1 dataset. Pada Gambar 4.4 dijelaskan hasil percobaan pada citra dengan memberikan nilai *threshold* yang berbeda-beda, yang memberikan hasil bahwa nilai *threshold* di bawah nilai 340 (Gambar 4.4(b)) menghasilkan deteksi bertumpuk yang berlebihan di mana banyak sel tunggal yang dideteksi bertumpuk sementara deteksi sel bertumpuk menggunakan nilai *threshold* di atas dari nilai 340 (Gambar 4.4(d)) menghasilkan kesalahan deteksi sel bertumpuk dimana sel bertumpuk dianggap sebagai sel tunggal, sehingga diasumsikan nilai *threshold* 340 (Gambar 4.4(c)) adalah nilai *threshold* terbaik. Untuk langkah berikutnya, hanya sel kanker yang fitur geometri area di atas nilai *threshold* atau dianggap bertumpuk yang akan diproses.

Sel kanker bertumpuk kemudian akan dipisahkan menggunakan *Spatial Fuzzy C-Means Clustering*. Pemisahan sel kanker bertumpuk pada tahapan ini adalah berdasarkan asumsi bahwa jumlah sel bertumpuk sama dengan jumlah *cluster* dengan melakukan perhitungan berdasarkan pada persamaan 3.5 di mana nilai *c* terbaik adalah 1,5. Nilai *c* didapatkan berdasarkan uji coba terhadap seluruh dataset dengan rentang nilai percobaan yang diberikan adalah 1,1 sampai 2,0 seperti yang ditampilkan pada Gambar 4.5. Data yang akan dikelompokkan atau data

masuk pada tahapan ini adalah berupa nilai koordinat piksel daerah yang diidentifikasi sebagai sel kanker bertumpuk.



Gambar 4. 4 Hasil Percobaan Penentuan Nilai Threshold untuk Citra Bertumpuk (a) Citra Asli (b) Citra Hasil Deteksi Sel Bertumpuk dengan Nilai threshold 300 (c) Citra Hasil Deteksi Sel Bertumpuk dengan Nilai threshold 340 (d) Citra Hasil Deteksi Sel Bertumpuk dengan Nilai threshold 400

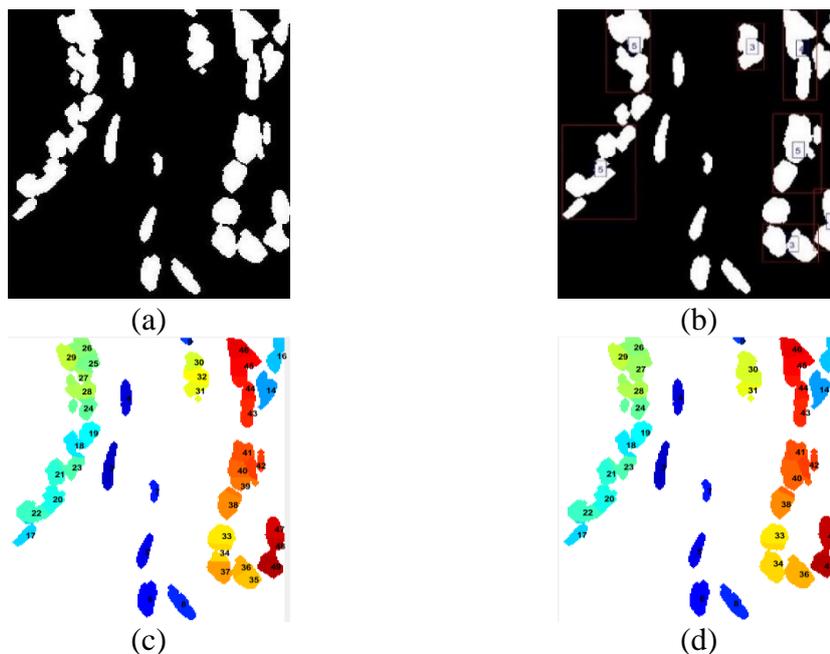


Gambar 4. 5 Uji Coba Penentuan Nilai  $c$  (a) citra asli (b) citra hasil pemisahan bertumpuk  $c=1,1$  (c) citra hasil pemisahan bertumpuk  $c=1,3$  (d) citra hasil pemisahan bertumpuk  $c=1,5$  (e) citra hasil pemisahan bertumpuk  $c=1,7$  (f) citra hasil pemisahan bertumpuk  $c=2$

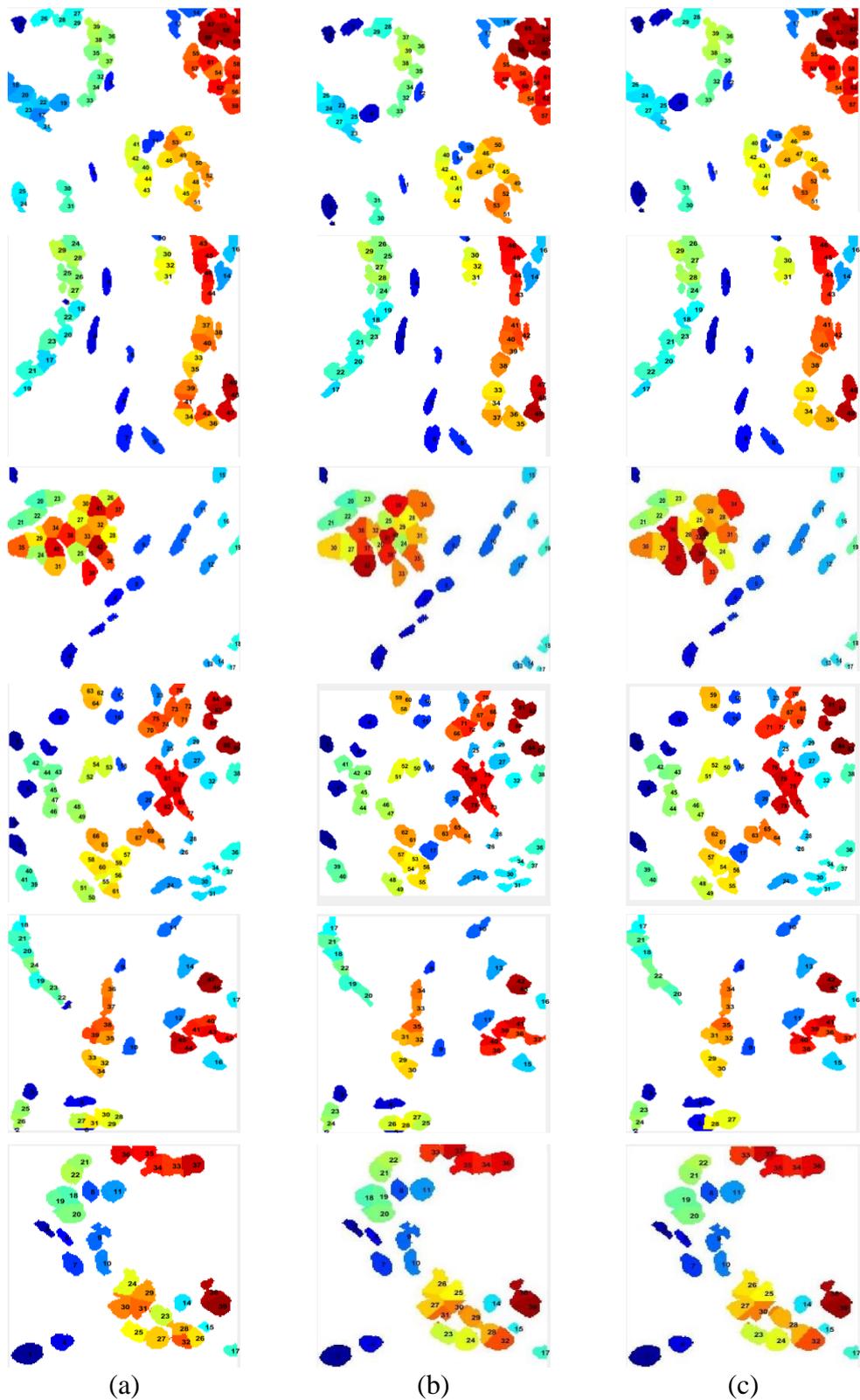
Semua sel kanker bertumpuk yang telah terpisah berdasarkan metode SFCM akan diproses ke tahapan penggabungan daerah menggunakan metode *Rapid Region Merging* untuk mengatasi masalah *oversegmentation*. Penggabungan daerah dilakukan berdasarkan penentuan daerah palsu atau *pseudo-blob* terlebih dahulu. Pada penelitian ini daerah yang dianggap *pseudo-blob* adalah daerah yang memiliki fitur geometri area di bawah dari nilai  $\frac{1}{2}$  *threshlod bertumpuk* dan daerah yang memiliki fitur geometri area yang lebih tinggi ditandai sebagai *real-*

*blob*. Kemudian *pseudo-blob* akan dihitung kemiripannya dengan *real-blob* yang memiliki *region adjacency* dengannya. Penggabungan daerah *pseudo-blob* adalah berdasarkan *maximum similarity*, di mana *real-blob* yang memiliki kemiripan paling tinggi dengan *pseudo-blob* akan dinyatakan sebagai daerah pemenang atau *winner block*. Daerah yang ditentukan sebagai *pseudo-blob* akan digabungkan ke daerah pemenang atau menjadi bagian dari daerah pemenang. Pada tahapan penggabungan daerah ini juga, dilakukan penambahan parameter di mana jika jumlah *pseudo-blob* berjumlah sama dengan jumlah *cluster* pada sebuah daerah sel bertumpuk maka diasumsikan daerah tersebut adalah sel tunggal, sehingga semua daerah akan digabungkan.

Hasil citra yang diperoleh pada setiap proses dalam tahapan pemisahan sel bertumpuk ini dimulai dari citra masukan, deteksi bertumpuk, hasil pemisahan sel bertumpuk dapat dilihat pada Gambar 4.6. Pada skenario uji coba ini juga dilakukan perbandingan hasil pemisahan sel bertumpuk menggunakan metode *K-Means*, metode SFCM dan kombinasi metode SFCM dan RRM yang ditampilkan pada Gambar 4.7.



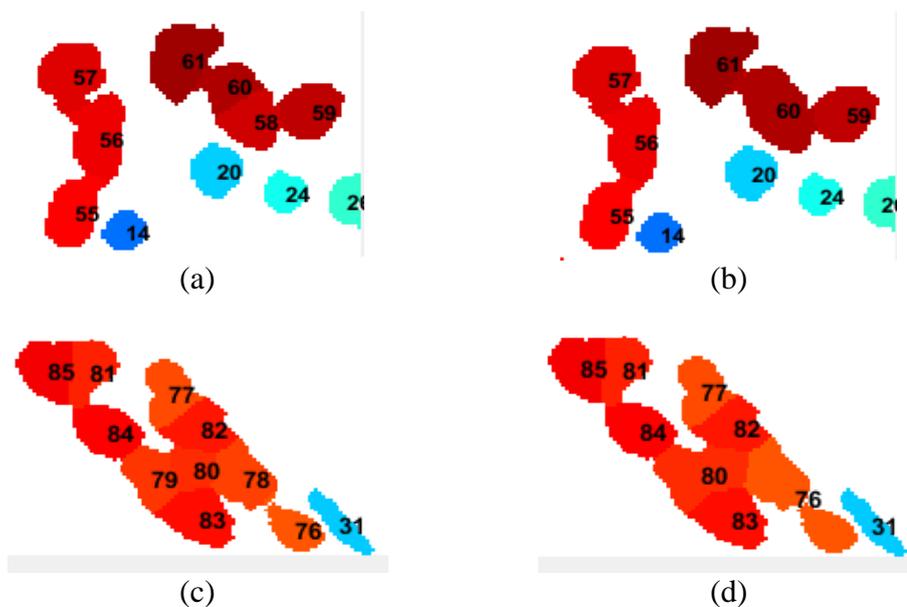
Gambar 4. 6 Hasil Citra Berdasarkan Tahapan-Tahapan Pemisahan Sel Bertumpuk (a)Citra Masukan pada Tahapan Clustering (b) Identifikasi Sel Kanker Bertumpuk dan Identifikasi Jumlah Cluster (c)Hasil Pemisahan Sel Bertumpuk Menggunakan SFCM (d)Citra Hasil Penggabungan Daerah Oversegmentasi Menggunakan RRM



Gambar 4. 7 Perbandingan Citra Hasil Pemisahan Sel Bertumpuk (a) Citra Masukan Pada Proses *Clustering* (b) Citra Hasil Pemisahan Sel Bertumpuk Menggunakan SFCM (c) Citra Hasil Pemisahan Sel Bertumpuk Menggunakan Kombinasi SFCM dan RRM

Berdasarkan perbandingan hasil pemisahan sel bertumpuk ini di atas, maka dilakukan perhitungan sel kanker yang dipisahkan secara benar, sel kanker yang *oversegmentation* dan sel kanker yang *undersegmentation*. Sel kanker benar dihitung berdasarkan ketepatan inisialisasi jumlah *cluster* yang sama dengan jumlah sel bertumpuk dengan memperhatikan garis pemisah, seperti pada Gambar 4.8(b). Sel kanker *oversegmentation* dihitung berdasarkan jumlah *cluster* yang lebih banyak dibandingkan dengan jumlah sel sebenarnya, seperti pada Gambar 4.8(a) dan Gambar 4.8(c) sedangkan sel kanker *undersegmentation* dihitung berdasarkan jumlah *cluster* yang lebih sedikit dibandingkan dengan jumlah sel yang sebenarnya, seperti pada Gambar 4.8(d).

Perhitungan manual jumlah sel pada hasil pemisahan sel kanker bertumpuk menggunakan metode *K-Means*, metode SFCM, kombinasi metode SFCM dan RRM berdasarkan pada citra asli dan *groundtruth* dapat dilihat pada Tabel 4.2, Tabel pada Lampiran 1 dan Lampiran 2. Selain itu, untuk mengetahui apakah terdapat perubahan yang signifikan dari metode usulan hasil dengan metode pembanding, dilakukan perhitungan uji-t yang disajikan pada tabel 4.4, tabel 4.5, tabel 4.6 dan tabel 4.7



Gambar 4. 8 (a)(c) Citra hasil SFCM yang Oversegmentasi (b) Citra Hasil RRM yang Sesuai dengan Citra Asli (d) Citra Hasil RRM yang Undersegmentasi

Tabel 4. 2 Perhitungan Nilai Akurasi Kombinasi Metode Spatial Fuzzy C-Means dan Rapid Region Merging

citra ke	Manual	SFCM & Rapid Region Merging			Akurasi (%)
		Benar	Over	Under	
1	47	33	10	3	70,21
2	44	33	8	1	75,00
3	34	29	5	0	85,29
4	46	32	10	0	69,57
5	67	54	1	5	80,60
6	57	45	1	1	78,95
7	42	32	2	1	76,19
8	27	15	13	0	55,56
9	27	19	7	0	70,37
10	28	19	8	1	67,86
11	32	27	3	1	84,38
12	47	42	2	1	89,36
13	62	52	2	2	83,87
14	19	8	9	0	42,11
15	27	19	8	2	70,37
16	40	34	6	0	85,00
17	74	64	0	5	86,49
18	36	31	1	3	86,11
19	71	56	0	4	78,87
20	24	23	1	0	95,83
21	82	68	1	7	82,93
22	42	34	2	2	80,95
23	31	20	10	1	64,52
24	22	18	4	0	81,82
25	46	41	3	0	89,13
26	40	33	4	3	82,50
27	50	37	12	1	74,00
28	61	50	9	1	81,97
29	43	40	0	2	93,02
30	39	32	2	4	82,05
31	74	62	6	6	83,78
32	89	75	0	4	84,27
33	53	37	12	0	69,81
34	26	18	6	0	69,23
35	47	36	3	1	76,60
36	43	36	3	1	83,72
37	58	49	9	0	84,48
38	25	17	8	0	68,00
39	108	94	1	9	87,04
40	26	22	4	0	84,62
Rata-Rata Akurasi =					78,41

#### 4.4 Analisis Hasil Uji Coba

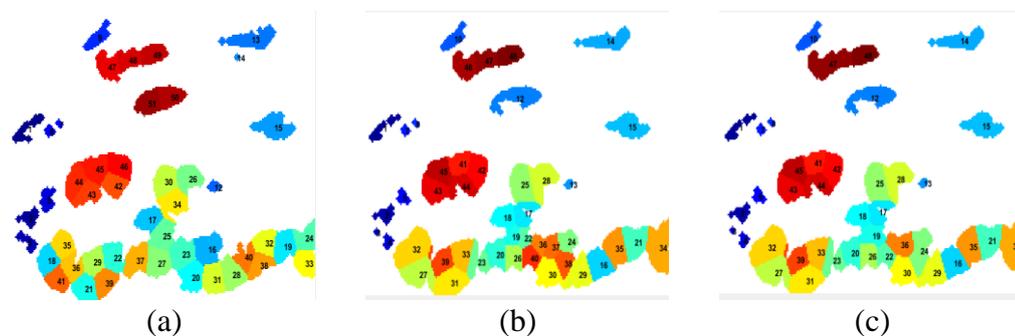
Hasil uji coba yang diperoleh dari setiap skenario uji coba akan dianalisis untuk mengetahui tingkat keberhasilan metode yang diusulkan. Analisis dilakukan berdasarkan perhitungan *Mean Square Error* (MSE) untuk skenario pertama, yaitu segmentasi sel kanker payudara dan perhitungan nilai akurasi untuk skenario kedua, yaitu pemisahan sel kanker bertumpuk. Nilai MSE merupakan nilai performa yang menggambarkan tingkat kesalahan segmentasi. Nilai akurasi digunakan untuk menilai performa metode yang diusulkan dalam melakukan pemisahan sel bertumpuk. Nilai akurasi didapat berdasarkan perhitungan manual rasio jumlah sel yang diidentifikasi atau jumlah sel yang terpisah dengan benar oleh metode yang diusulkan terhadap jumlah sel pada citra asli. Nilai akurasi ini yang akan menjadi nilai pembanding antara metode *K-means*, metode *Spatial Fuzzy C-Means* (SFCM) tanpa metode penggabungan dengan metode yang diusulkan yaitu kombinasi *Spatial Fuzzy C-Means* dan *Rapid Region Merging* untuk mengetahui apakah kinerja metode usulan memberi dampak yang baik dan menghasilkan perubahan yang signifikan.

##### 4.4.1 Analisa Hasil Uji Coba Skenario Pertama

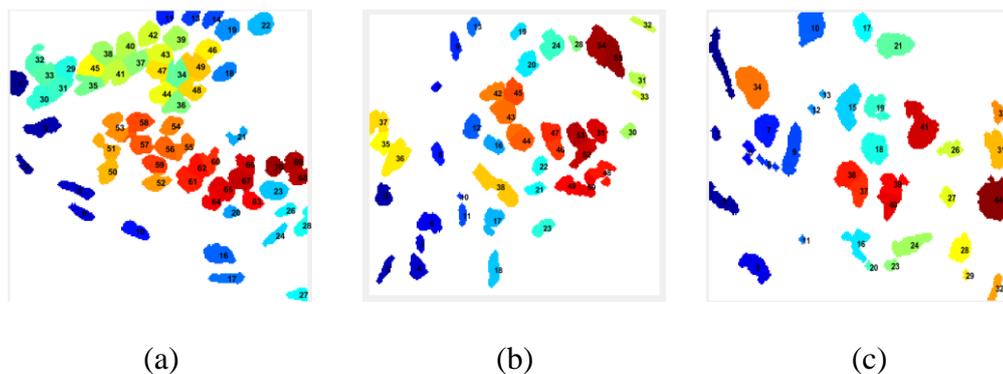
Berdasarkan hasil uji coba terhadap 40 data uji berdasarkan skenario pertama yang ditampilkan pada Tabel 4.1, nilai MSE yang dihasilkan metode *Gram-Schmidt* berada dalam rentang nilai 0,0394 sampai 0,1225 dan nilai rata-rata kesalahan segmentasi sebesar 0,0745, di mana nilai MSE terendah dihasilkan pada citra ke-11 dan nilai MSE tertinggi dihasilkan oleh citra ke-19. Sementara, nilai MSE yang dihasilkan oleh metode *Region Based Active Contour* adalah dalam rentang nilai 0,0728 pada citra ke-12 sampai 0,9578 pada citra ke-10 dan menghasilkan rata-rata MSE sebesar 0,6505. Hal ini menunjukkan bahwa tingkat kesalahan segmentasi pada penelitian adalah rendah dan citra hasil segmentasi dapat dijadikan citra masukan pada tahap selanjutnya. Pada Hasil Uji Coba Skenario Pertama, nilai *threshold* ditentukan berdasarkan uji coba nilai *threshold* dari angka 0,1 sampai 0,8 yang menghasilkan nilai MSE terbaik adalah nilai *threshold* 0,3.

#### 4.4.2 Analisis Hasil Uji Coba Skenario Kedua

Berdasarkan hasil uji coba skenario kedua yang ditampilkan pada tabel 4.6 dan tabel terlampir dalam lampiran 1 dan lampiran 2, nilai akurasi pemisahan sel bertumpuk menggunakan metode *K-Means* menghasilkan nilai akurasi terendah sebesar 21,05%, akurasi tertinggi sebesar 86,57% dan rata-rata akurasi mencapai 68,73%. Pemisahan sel bertumpuk menggunakan SFCM menghasilkan nilai akurasi terendah sebesar 36,84% dan nilai akurasi tertinggi sebesar 88,37% dengan rata-rata nilai akurasi sebesar 71,35%. Nilai akurasi pemisahan sel bertumpuk menggunakan kombinasi metode SFCM dan RRM menghasilkan nilai akurasi terendah sebesar 42,11% dan nilai akurasi tertinggi sebesar 95,83% dengan rata-rata nilai akurasi sebesar 78,41%. Pada Gambar 4.9 dan Gambar 4.10 disajikan citra hasil pemisahan sel bertumpuk yang menghasilkan nilai akurasi terendah dan nilai akurasi tertinggi dari metode yang diusulkan dan metode pembanding.



Gambar 4. 9 Citra yang Menghasilkan Nilai Akurasi Terendah (a) K-means (b) SFCM (c) SFCM & RRM



Gambar 4. 10 Citra Yang Menghasilkan Nilai Akurasi Paling Tinggi (a) K-Means (b) SFCM (c) SFCM & RRM

Pada tabel 4.3 dapat dilihat perbedaan jumlah sel yang diidentifikasi benar, jumlah sel yang *oversegmentasi*, jumlah sel yang *undersegmentasi* dan rata-rata akurasi berdasarkan citra hasil metode *K-Means*, metode SFCM dan kombinasi metode SFCM dan RRM yang dihitung secara manual. Hasil ini menunjukkan bahwa jumlah sel benar, jumlah sel yang *undersegmentation* dan nilai akurasi pemisahan sel bertumpuk menggunakan kombinasi metode SFCM dan RRM mengalami peningkatan dibandingkan dengan metode *K-Means* dan SFCM. Sementara jumlah sel yang *oversegmentation* mengalami penurunan.

Tabel 4. 3 Tabel Perbandingan Metode Sebelumnya dengan Metode yang Diusulkan

	Jumlah Benar	Jumlah <i>Oversegmentation</i>	Jumlah <i>Undersegmentation</i>	Rata-Rata Akurasi (%)
K-Means	1333	382	35	68,73
SFCM	1402	314	39	73,11
SFCM dan RRM	1486	197	72	78,41

Untuk mengetahui apakah peningkatan nilai akurasi metode yang diusulkan adalah signifikan, dilakukan perhitungan uji  $t$  sampel besar  $N \geq 30$  yang berkorelasi, dihitung pada taraf signifikan 1% dan taraf signifikan 5%. Uji  $t$  terhadap variabel data hasil SFCM dan metode yang diusulkan disajikan pada tabel 4.4 dan tabel 4.5 dan untuk uji  $t$  terhadap variabel data hasil *K-means* dan metode yang diusulkan disajikan pada tabel 4.6 dan tabel 4.7. *Mean* menunjukkan rata-rata akurasi, *variance* adalah varians data yang digunakan pada penelitian ini, *observations* adalah jumlah data, *df* adalah derajat bebas, *t Stat* adalah nilai  $t_0$  dan *t-critical two-tail* adalah nilai  $t_{tabel}$ .

Hasil perhitungan uji  $t$  terhadap variabel data hasil SFCM dan metode yang diusulkan menghasilkan nilai  $t_0$  atau dalam tabel di tulis *t Stat* adalah sebesar -6,4229 dimana telah dibahas pada bab sebelumnya bahwa tanda negatif bukanlah merupakan tanda aljabar sehingga nilai  $t_0$  sama dengan 6,4229. Nilai  $t_{tabel}$  pengujian dua arah pada taraf signifikan 5% dengan nilai  $df = 39$  sebesar 2,0226.

Nilai  $t_{tabel}$  pengujian dua arah pada taraf signifikan 1% dengan nilai  $df = 39$  sebesar 2,7079. Diperolehnya nilai  $t_0$  sebesar 6,7820 dan nilai  $t_{tabel}$  pada taraf signifikan 5% sebesar 2,0226 dan  $t_{tabel}$  pada taraf signifikan 1% sebesar 2,7079 sehingga nilai  $t_0$  lebih besar dari nilai  $t_{tabel}$  pada taraf signifikan 5% maupun taraf signifikan 1% atau dapat ditulis bahwa  $(2,022 < 6,7820 > 2,7079)$ . Dengan demikian, berdasarkan ketentuan perbandingan  $t_0$  dan  $t_{tabel}$  maka hipotesis nol  $H_0$  ditolak dan yang berarti ada perbedaan yang signifikan antara hasil pemisahan bertumpuk dengan menggunakan metode SFCM dan kombinasi metode SFCM dan RRM.

Tabel 4. 4 Perhitungan uji-t untuk variabel hasil SFCM dan Metode yang Diusulkan pada taraf signifikan 5%

	SFCM	SFCM & RRM
Mean	73,1132697	78,41026504
Variance	123,5220251	104,7615607
Observations	40	40
Pearson Correlation	0,883817577	
Hypothesized Mean Difference	0	
df	39	
t Stat	-6,422959482	
P(T<=t) one-tail	6,70181E-08	
t Critical one-tail	1,684875122	
P(T<=t) two-tail	1,34036E-07	
t Critical two-tail	2,02269092	

Tabel 4. 5 Perhitungan uji-t untuk variabel hasil SFCM dan Metode yang Diusulkan pada taraf signifikan 1%

	SFCM	SFCM & RRM
Mean	73,1132697	78,41026504
Variance	123,5220251	104,7615607
Observations	40	40
Pearson Correlation	0,883817577	
Hypothesized Mean Difference	0	
df	39	
t Stat	-6,422959482	
P(T<=t) one-tail	6,70181E-08	
t Critical one-tail	2,42584141	
P(T<=t) two-tail	1,34036E-07	
t Critical two-tail	2,707913184	

Tabel 4. 6 Perhitungan uji-t untuk variabel hasil *K-Means* dan Metode yang Diusulkan pada taraf signifikan 5%

	<i>K-Means</i>	SFCM & RRM
Mean	68,72965787	78,41026504
Variance	170,5114121	104,7615607
Observations	40	40
Pearson Correlation	0,831066084	
Hypothesized Mean Difference	0	
df	39	
t Stat	-8,400110806	
P(T<=t) one-tail	1,39592E-10	
t Critical one-tail	1,684875122	
P(T<=t) two-tail	2,79185E-10	
t Critical two-tail	2,02269092	

Tabel 4. 7 Perhitungan uji-t untuk variabel hasil *K-Means* dan Metode yang Diusulkan pada taraf signifikan 1%

	<i>K-Means</i>	SFCM & RRM
Mean	68,72965787	78,41026504
Variance	170,5114121	104,7615607
Observations	40	40
Pearson Correlation	0,831066084	
Hypothesized Mean Difference	0	
df	39	
t Stat	-8,400110806	
P(T<=t) one-tail	1,39592E-10	
t Critical one-tail	2,42584141	
P(T<=t) two-tail	2,79185E-10	
t Critical two-tail	2,707913184	

Hasil perhitungan uji – *t* terhadap variabel data hasil *K-Means* dan metode yang diusulkan menghasilkan nilai  $t_0$  atau dalam tabel di tulis *t Stat* adalah sebesar 8,4001. Diperolehnya nilai  $t_0$  sebesar 8,4001 dan nilai  $t_{tabel}$  pada taraf signifikan 5% sebesar 2,0226 dan  $t_{tabel}$  pada taraf signifikan 1% sebesar 2,7079 sehingga nilai  $t_0$  lebih besar dari nilai  $t_{tabel}$  pada taraf signifikan 5% maupun taraf signifikan 1% atau dapat ditulis bahwa  $(2,022 < 8,4001 > 2,7079)$ . Dengan demikian, berdasarkan ketentuan perbandingan  $t_0$  dan  $t_{tabel}$  maka hipotesis nol  $H_0$  ditolak

dan yang berarti ada perbedaan yang signifikan antara hasil pemisahan bertumpuk dengan menggunakan metode *K-Means* dan kombinasi metode SFCM dan RRM.

## BAB V

### PENUTUP

Pada bab ini akan dijelaskan mengenai kesimpulan dari penelitian yang telah dilakukan dan saran untuk pengembangan penelitian lebih lanjut yang berhubungan dengan segmentasi dan pemisahan sel bertumpuk pada citra mikroskopis sel kanker payudara.

#### 5.1 Kesimpulan

Berdasarkan hasil uji coba dan analisis hasil uji coba yang dijelaskan pada bab sebelumnya, dapat disimpulkan beberapa hal sebagai berikut:

- 1) Identifikasi sel kanker payudara metode *Gram-Schmidt* yang dilakukan pada penelitian ini, memberikan hasil yang baik dibandingkan dengan metode *Region Based Active Contour*. Hal ini ditunjukkan dengan diperolehnya rata-rata nilai MSE sebesar 0,07 menggunakan metode *Gram-Schmidt* dengan nilai *threshold* 0,03.
- 2) Pemisahan sel bertumpuk menggunakan metode *Spatial Fuzzy C-Means* dengan perhitungan *spatial* dan probabilitas ketetanggaan berdasarkan koordinat atau *pixellist* sel bertumpuk menghasilkan rata-rata akurasi yang lebih tinggi yaitu 73,11% dibandingkan dengan metode *K-Means* yang menghasilkan nilai akurasi sebesar 68,73%. Meningkatnya akurasi metode SFCM dikarenakan oleh beberapa hal, yaitu mampu meminimalisir jumlah sel *oversegmentation* yang diakibatkan oleh kesalahan pengelompokan, keberhasilan deteksi sel kanker bertumpuk menggunakan nilai *threshold* 340 berdasarkan fitur geometri area sehingga sel yang di proses adalah sel yang diidentifikasi bertumpuk dan keberhasilan inisialisasi jumlah *cluster* karena sel kanker akan terpisah sesuai dengan jumlah masukan yang didapat dari proses inisialisasi jumlah *cluster* yang artinya bahwa banyaknya jumlah sel yang *oversegmentation* bergantung pada inisialisasi jumlah *cluster*. Pada penelitian ini, inisialisasi jumlah *cluster* terbaik dihasilkan menggunakan nilai  $c = 1,5$ .
- 3) Metode *Rapid Region Merging* untuk menggabungkan sel yang *oversegmentation* mampu meningkatkan hasil akurasi pemisahan sel

bertumpuk. Hal ini ditunjukkan dengan berkurangnya jumlah sel yang *oversegmentation* sehingga dihasilkannya nilai rata-rata akurasi sebesar 78,41%. Namun penerapan metode *Rapid Region Merging* memiliki kelemahan yaitu meningkatnya jumlah sel yang *undersegmentation* akibat kesalahan penggabungan daerah. Kesalahan penggabungan daerah terjadi karena kesalahan deteksi *pseudo blob* dan akibat nilai *similarity* yang sama.

- 4) Hasil perhitungan *uji - t* yang dilakukan menghasilkan kesimpulan bahwa kombinasi metode SFCM dan RRM memberikan peningkatan nilai akurasi yang signifikan dari metode SFCM dan metode *K-Means*.

## 5.2 Saran

- 1) Penelitian selanjutnya dapat melakukan pengembangan metode identifikasi sel kanker payudara khususnya pada proses penentuan vektor-vektor warna yang digunakan pada metode *Gram-Schmidt* yang masih dilakukan secara manual.
- 2) Penentuan jumlah *cluster* sangat berpengaruh pada hasil pemisahan sel bertumpuk menggunakan SFCM sehingga untuk meningkatkan nilai akurasi pemisahan sel bertumpuk perlu dikembangkannya metode untuk menentukan atau inisialisasi jumlah *cluster* yang tidak hanya berdasarkan informasi fitur geometri area saja, tapi menambahkan fitur-fitur yang bisa merepresentasikan banyaknya jumlah sel kanker bertumpuk.
- 3) Meningkatnya jumlah sel kanker payudara yang *undersegmentation* pada hasil pemisahan sel kanker payudara menggunakan metode yang diusulkan diakibatkan karena semua sel bertumpuk yang telah dipisahkan menggunakan SFCM akan dilanjutkan ke tahapan penggabungan daerah tanpa identifikasi daerah *oversegmentation* terlebih dahulu sehingga diperlukan pengembangan metode untuk identifikasi daerah *oversegmentation* sebelum sel diproses pada tahap penggabungan daerah.
- 4) Selain identifikasi sel kanker yang *oversegmentation*, deteksi *pseudo blob* yang berdasarkan informasi fitur geometri area sangat rentan terhadap *error* karena karakteristik citra mikroskopis kanker payudara yang memiliki beragam ukuran

dan bentuk sehingga perlu diteliti lebih lanjut mengenai parameter penentuan *pseudo-blob* untuk memperkecil kesalahan penggabungan.

*[Halaman ini sengaja dikosongkan]*

## DAFTAR PUSTAKA

- <https://depkes.go.id>, Pusat Data dan Informasi Kementerian Kesehatan RI, di akses Februari 2016.
- <https://who.int>, di akses Februari 2016.
- Abbas, Amel H; Kareem, Aryan A; Kamil, M. Y. (2015). Breast Cancer Image Segmentation Using Morphological Operations. *International Journal Of Electronics And Communication Engineering & Technology (IJECEET)*, 6.
- Aini, N. (2015). *Metode Hybrid Region-Based Active Contour dan Bayesian untuk Segmentasi dan Pemisahan Sel Bertumpuk pada Citra Mikroskopis Sel Kanker Payudara*.
- Beevi, S. Z., Sathik, M. M., Senthamaraikannan, K., & Yasmin, J. H. J. (2010). A RobustT Fuzzy Clustering Technique With Spatial Neighborhood Information for Effective Noisy Medical Image Segmentation.
- Chen, Y., & Chen, J. (2014). A Watershed Segmentation Algorithm Based on Ridge Detection and Rapid Region Merging, 420–424.
- Fariied, E. (2013). Segmentasi Sel Darah Merah Bertumpuk Berdasarkan Fitur Geometri Pada Perhitungan Sel Darah Merah.
- Fatichah, C., & Suciati, N. (2015). Nuclei Segmentation of Microscopic Breast Cancer Image using Gram-Schmidt and Cluster Validation Algorithm, (November), 27–29.
- Gonzalez, R., & Woods, R. (2002). *Digital image processing. Prentice Hall*.
- Hadi, S. (1997). Statistik, Penerbit Andi.
- Li, Y., & Shen, Y. (2010). Fuzzy c-means clustering based on spatial neighborhood information for image segmentation, *21(2)*, 323–328.
- Mouelhi, A., Sayadi, M., & Fnaiech, F. (2013). A Supervised Segmentation Scheme Based on Multilayer Neural Network and Color Active Contour Model for Breast Cancer Nuclei Detection, (1).

- Mouelhi, A., Sayadi, M., Fnaiech, F., & Member, S. (2013). Hybrid Segmentation of Breast Cancer Cell Images Using a New Fuzzy Active Contour Model and an Enhanced Watershed Method, (1), 382–387.
- Mouelhi, A., Sayadi, M., Fnaiech, F., & Mrad, K. (2013). Biomedical Signal Processing and Control Automatic image segmentation of nuclear stained breast tissue sections using color active contour model and an improved watershed method. *Biomedical Signal Processing and Control*, 8(5), 421–436.
- Phukpattaranont, P., & Boonyaphiphat, P. (2007). Color Based Segmentation of Nuclear Stained Breast Cancer Cell Images. *Communications*, 5(2), 158–164.
- Rezatofighi, S. H. (2009). International Conference on Digital Image Processing A New Approach to White Blood Cell Nucleus Segmentation Based on Gram-Schmidt Orthogonalization, 107–111.
- Prasetyo, E. (2011). *Pengolahan Citra Digital dan Aplikasinya Menggunakan MATLAB*, Penerbit Andi.
- Prasetyo, E. (2014). *Data Mining Mengolah Data Menjadi Informasi Menggunakan MATLAB*, Penerbit Andi.

*[Halaman ini sengaja dikosongkan]*

## Lampiran 1

Tabel Perhitungan Nilai Akurasi Metode Spatial Fuzzy C-Means

citra ke	Manual	Spatial Fuzzy C-Means (SFCM)			Akurasi (%)
		Benar	Over	Under	
1	47	32	12	2	68,09
2	44	28	14	0	63,64
3	34	26	8	0	76,47
4	46	30	12	0	65,22
5	67	56	1	3	83,58
6	57	43	3	1	75,44
7	42	28	6	1	66,67
8	27	15	13	0	55,56
9	27	18	8	0	66,67
10	28	18	9	0	64,29
11	32	27	3	1	84,38
12	47	40	4	1	85,11
13	62	49	6	1	79,03
14	19	7	10	0	36,84
15	27	18	9	2	66,67
16	40	29	11	0	72,50
17	74	60	3	6	81,08
18	36	30	3	2	83,33
19	71	56	1	3	78,87
20	24	20	4	0	83,33
21	82	71	2	3	86,59
22	42	30	7	1	71,43
23	31	18	13	0	58,06
24	22	17	5	0	77,27
25	46	31	13	0	67,39
26	40	33	6	1	82,50
27	50	35	16	0	70,00
28	61	47	13	0	77,05
29	43	38	4	0	88,37
30	39	33	5	0	84,62
31	74	61	9	4	82,43
32	89	76	1	2	85,39
33	53	35	14	0	66,04
34	26	15	9	0	57,69
35	47	32	8	0	68,09
36	43	33	6	1	76,74
37	58	44	14	0	75,86
38	25	13	12	0	52,00
39	108	90	10	4	83,33
40	26	20	6	0	76,92
Rata-Rata Akurasi =					73,11

*[Halaman ini sengaja dikosongkan]*

## Lampiran 2

Tabel Perhitungan Nilai Akurasi Metode K-means

citra ke	Manual	K-Means			Akurasi (%)
		Benar	Over	Under	
1	47	28	17	1	59,57
2	44	25	15	2	56,82
3	34	24	10	0	70,59
4	46	28	13	0	60,87
5	67	58	0	2	86,57
6	57	42	4	1	73,68
7	42	27	7	1	64,29
8	27	14	14	0	51,85
9	27	16	10	0	59,26
10	28	17	9	0	60,71
11	32	22	8	0	68,75
12	47	35	8	0	74,47
13	62	43	12	1	69,35
14	19	4	13	0	21,05
15	27	17	10	0	62,96
16	40	27	13	0	67,50
17	74	61	3	5	82,43
18	36	30	3	2	83,33
19	71	55	0	5	77,46
20	24	18	6	0	75,00
21	82	70	3	3	85,37
22	42	27	11	0	64,29
23	31	18	13	0	58,06
24	22	16	6	0	72,73
25	46	30	14	0	65,22
26	40	31	7	2	77,50
27	50	33	17	0	66,00
28	61	47	13	0	77,05
29	43	35	7	0	81,40
30	39	32	6	0	82,05
31	74	61	9	4	82,43
32	89	72	4	3	80,90
33	53	32	17	0	60,38
34	26	14	10	0	53,85
35	47	33	7	0	70,21
36	43	30	8	2	69,77
37	58	43	15	0	74,14
38	25	10	15	0	40,00
39	108	87	16	1	80,56
40	26	21	5	0	80,77
Rata-Rata Akurasi =					68,73

*[Halaman ini sengaja dikosongkan]*

## BIODATA



Desmin Tuwohingide, lahir di Tahuna pada tanggal 12 Desember 1990 merupakan anak pertama dari pasangan bapak Herson Tuwohingide dan ibu Feronika Supit. Penulis mulai menempuh pendidikan formal pada tahun 1996 di Sekolah Dasar Negeri (SDN) Inpres Pananaru, pada tahun 2002-2005 menempuh pendidikan di Sekolah Menengah Pertama (SMP) Negeri 1 Manganitu Selatan dan melanjutkan ke Sekolah Menengah Kejuruan (SMK) Negeri Satu Tahuna, dengan pilihan program studi Teknik Komputer dan Jaringan pada tahun 2005-2008. Pada tahun 2008 penulis melanjutkan pendidikan ke jenjang perguruan tinggi di Universitas Sariputra Indonesia Tomohon (UNSRIT), Jurusan Teknik, Program Studi Teknik Informatika dan memperoleh gelar Sarjana Teknik pada tahun 2012. Pada bulan februari tahun 2013, penulis mulai mengabdikan sebagai dosen tetap di Jurusan Teknik Komputer dan Komunikasi, Politeknik Negeri Nusa Utara (Polnustar). Pada bulan Oktober 2013, penulis menerima beasiswa DIKTI PraS2 Saintek Daerah 3T di Jurusan Matematika, Insitut Teknologi Sepuluh Nopember Surabaya (ITS) dan pada tahun 2014 melanjutkan pendidikan di Program Magister Jurusan Teknik Informatika, Insitut Teknologi Sepuluh Nopember Surabaya (ITS) dan memperoleh gelar Master Komputer (M.Kom) Pada tahun 2017.