



**TESIS – KI142502**

**SEGMENTASI DAN PEMISAHAN SEL DARAH PUTIH  
BERSENTUHAN MENGGUNAKAN METODE K-MEANS DAN  
HIERARCHICAL CLUSTERING ANALYSIS PADA CITRA  
LEUKEMIA MYELOID AKUT**

**ARYO HARTO**

**NRP. 5114201001**

**DOSEN PEMBIMBING**

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

**PROGRAM MAGISTER  
BIDANG KEAHLIAN KOMPUTASI CERDAS DAN VISI  
JURUSAN TEKNIK INFORMATIKA  
FAKULTAS TEKNOLOGI INFORMASI  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2017**



**THESIS – KI 142502**

**TOUCHED WHITE BLOOD CELLS SEGMENTATION AND  
SEPARATION USING K-MEANS METHOD AND HIERARCHICAL  
CLUSTERING ANALYSIS ON ACUTE MYELOID LEUKEMIA IMAGE**

**ARYO HARTO**

**NRP. 5114201001**

**SUPERVISOR**

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

**MASTER PROGRAM  
INTELLIGENT COMPUTING AND VISION  
INFORMATICS ENGINEERING DEPARTMENT  
FACULTY OF INFORMATION TECHNOLOGY  
SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY  
SURABAYA  
2017**

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar  
Magister Komputer (M.Kom.)

Di  
Institut Teknologi Sepuluh Nopember Surabaya

Oleh:  
Aryo Harto  
NRP. 5114201001

Dengan judul :  
Segmentasi dan Pemisahan Sel Darah Putih Bersentuhan Menggunakan Metode K-Means  
dan Hierarchical Clustering Analysis Pada Citra Leukemia Myeloid Akut

Tanggal Ujian : 3 - 1 - 2017

Periode Wisuda : 2016 Ganjil

Disetujui oleh:

Dr. Eng. Chastine Fatichah, S.Kom, M.Kom  
NIP. 197512202001122002

(Pembimbing 1)

Dr. Darlis Heru Murti, S.Kom, M.Kom  
NIP. 197712172003121001

(Penguji 1)

Bilqis Amaliah, S.Kom, M.Kom  
NIP. 197509172001122002

(Penguji 2)

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

(Penguji 3)

an. Direktur Program Pascasarjana  
Asisten Direktur



Prof. Dr. Ir. Tri Widjaja, M.Eng.  
NIP. 19611021198603 1 001

Direktur Program Pasca Sarjana,

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

*[Halaman ini sengaja dikosongkan]*

# Segmentasi dan Pemisahan Sel Darah Putih Bersentuhan Menggunakan Metode K-Means dan Hierarchical Clustering Analysis Pada Citra Leukemia Myeloid Akut

Nama Mahasiswa : Aryo Harto

NRP : 5114 201 001

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

## ABSTRAK

Keberhasilan identifikasi dan klasifikasi pada proses diagnosis penyakit *Acute Myeloid Leukemia* (AML) berbasis pengolahan citra sangat bergantung pada hasil segmentasi. Segmentasi pada citra mikroskopis sediaan apus darah tepi bertujuan untuk memisahkan *region* sel darah putih dari *region* lainnya. Untuk meningkatkan akurasi segmentasi pada citra AML beberapa hal seperti efek pencahayaan, kontras, *staining* yang beragam dan keberadaan sel yang bersentuhan harus diatasi.

Pada penelitian ini diusulkan metode segmentasi dan pemisahan sel darah putih bersentuhan pada citra AML menggunakan analisis *cluster* dengan metode K-Means dan *hierarchical clustering analysis* (HCA). Segmentasi dilakukan menggunakan analisis *cluster* dengan metode K-Means pada citra AML yang memiliki variasi *staining* beragam. Keberadaan sel yang bersentuhan dipisahkan dengan metode HCA secara otomatis dengan hasil *cluster* yang stabil. Segmentasi dan pemisahan sel akan dilakukan pada masing-masing *region* lokal atau sub-citra yang didapatkan dari *cropping* pada citra AML. *Initial ROI* sel darah putih didapatkan dengan metode *thresholding* Otsu pada citra hasil *Leukocytes Nucleus Enhancer* (LNE). Untuk mendeteksi sel yang bersentuhan digunakan fitur geometri berupa kebulatan dan luasan.

Dari hasil evaluasi yang dilakukan terhadap 40 citra pada dataset AML, metode yang diusulkan melakukan segmentasi sel darah putih dengan baik dan memisahkan *touching cell* menjadi sel-sel tunggal. Nilai rata-rata dari hasil segmentasi masing-masing adalah 0.977 untuk presisi, 0.885 untuk *recall* dan 0.928 untuk *Zijdenbos similarity index*(ZSI) pada *region* sel darah putih. Sedangkan pada *region* nukleus nilai rata-rata presisi adalah 0.975, *recall* adalah 0.924 dan ZSI adalah 0.948. Pada perhitungan jumlah sel, *error* yang dihasilkan juga cukup rendah yaitu sebesar 4.04%.

**Kata Kunci:** sel darah putih, AML, segmentasi, *clustering*, K-Means, *hierarchical clustering analysis*

*[Halaman ini sengaja dikosongkan]*

# **Touched White Blood Cells Segmentation and Separation Using K-Means Method and Hierarchical Clustering Analysis On Acute Myeloid Leukemia Image**

Student Name : Aryo Harto  
NRP : 5114 201 001  
Supervisor : Dr. Eng. Chastine Fatichah, S.Kom, M.Kom.

## **ABSTRACT**

The success of identification and classification on diagnosing acute myeloid leukemia (AML) diseases based on image processing relies heavily on segmentation result. Segmentation on peripheral blood smear images aims to separate the leukocytes region with others region. To increase the segmentation accuracy on AML images, a few things regarding lighting condition, contrast, staining variations and the existence of touching cells must be overcome.

In this study a method for leukocytes segmentation and separate the touching cell on AML images using cluster analysis with K-Means and hierarchical clustering analysis (HCA) is proposed. K-Means method is used to analyze the cluster for AML images segmentation. The AML image datasets with various staining variations is segmented using K-Means method. The existence of touching cells is separated automatically using HCA method which produce a stable clusters result. Segmentation and cell separation will be processed on local region or sub-image which is obtained from AML images cropping. To get leukocytes initial ROI, Otsu thresholding method is used with the image processed with leukocytes nucleus enhancer (LNE). The touching cell area is detected using two geometric feature, roundness and size.

From the evaluation results in 40 images of AML dataset, the proposed method is capable to properly segment the white blood cells region and separating the touching cell into a single cells. The average value of the segmentation results is 0.977 for precision, 0.885 for recall and 0.928 for Zijdenbos similarity index (ZSI) in white blood cell region. While in nucleus region the average value is 0.975 for precision, 0.924 for recall and 0.948 for ZSI. On cell counting the error rate is also low which is about 4.04%.

**Keywords:** *leukocytes, AML, segmentation, clustering, K-Means, hierarchical clustering analysis*

*[Halaman ini sengaja dikosongkan]*

## **KATA PENGANTAR**

Segala puji dan syukur kepada Allah SWT sehingga akhirnya penelitian serta dokumentasinya dalam buku tesis ini dapat diselesaikan dengan baik. Dokumentasi penelitian dalam buku tesis ini berhasil disusun berkat bantuan dan bimbingan dari berbagai pihak. Penulis ingin mengucapkan terima kasih kepada pihak-pihak yang telah membantu selama penelitian dan penulisan buku tesis ini sebagai berikut.

1. Kepada Ibu dan keluarga besar, yang telah memberikan motivasi, dukungan, bantuan dan doa yang tidak terbatas bagi Penulis selama menyelesaikan studi ini.
2. Kepada Sahabat yang telah memberikan semangat, saran dan doa untuk menyelesaikan studi ini.
3. Kepada Dosen Pembimbing, Ibu Dr. Eng. Chastine Fatichah, S.Kom. M.Kom. yang dengan sabar telah memberikan bimbingan kepada Penulis sehingga tesis ini dapat selesai dengan baik.
4. Kepada para Dosen Penguji, Bapak Dr. Darlis Heru Murti, S.Kom., M.Kom., Ibu Bilqis Amaliah, S.Kom, M.Kom. dan Bapak Arya Yudhi Wijaya, S.Kom, M.Kom. yang telah memberikan saran dan arahan yang baik untuk perbaikan penelitian ini dari masa proposal hingga tesis ini selesai.
5. Kepada Bapak Prof. Ir. Supeno Djanali, M.Sc., Ph.D selaku dosen wali dan Bapak Waskitho Wibisono, S.Kom., M.Eng., Ph.D selaku ketua program studi Pascasarjana Teknik Informatika ITS serta segenap dosen Teknik Informatika ITS yang telah memberikan ilmu dan membagikan pengalamannya.
6. Kepada Dr. Arlan Prabowo, Sp.PK dan staf di Laboratorium Pemerintah Daerah Kalimantan Selatan atas segala informasi dan bantuan mengenai dataset serta dasar pengetahuan untuk penelitian ini.
7. Kepada teman-teman di Pascasarjana, Pak Syaiful, Tio, Mas Benny, mbak Desmin, Hateem, mas Heru dan mas Eka, yang telah memberikan ide dan saran selama kegiatan penelitian serta teman seperjuangan dan seangkatan yang tidak dapat disebutkan satu persatu di sini, yang telah menjadi teman diskusi yang baik sekaligus sebagai tempat untuk mencari motivasi.

8. Kepada staf administrasi Program Pascasarjana Teknik Informatika atas pengertian dan kebijaksanaannya dalam proses pengurusan administrasi.
9. Kepada admin Laboratorium Pascasarjana Teknik Informatika, Mas Kunto, atas bantuan saat pengerjaan tesis di laboratorium Pascasarjana.
10. Kepada rekan-rekan lain yang tidak bisa Penulis sebutkan satu-persatu, terima kasih atas segala bantuan, baik berupa ide, gagasan, pemikiran, dan doa.

Akhirnya, Penulis berharap buku tesis ini dapat memberikan manfaat dan kontribusi ilmiah bagi dunia pendidikan, khususnya pada pengembangan riset di bidang Komputasi Cerdas dan Visi.

Surabaya, Januari 2017

Aryo Harto Harun

## DAFTAR ISI

<b>LEMBAR PENGESAHAN TESIS .....</b>	<b>i</b>
<b>ABSTRAK .....</b>	<b>iii</b>
<b>ABSTRACT .....</b>	<b>iii</b>
<b>KATA PENGANTAR.....</b>	<b>vii</b>
<b>DAFTAR ISI.....</b>	<b>ix</b>
<b>DAFTAR GAMBAR.....</b>	<b>xi</b>
<b>DAFTAR TABEL .....</b>	<b>xiii</b>
<b>BAB 1 PENDAHULUAN .....</b>	<b>1</b>
1.1. Latar Belakang .....	1
1.2. Perumusan Masalah .....	6
1.3. Batasan Masalah.....	7
1.4. Tujuan .....	7
1.5. Manfaat Penelitian .....	7
1.6. Kontribusi Penelitian.....	7
1.7. Sistematika Penulisan.....	7
<b>BAB 2 DASAR TEORI DAN KAJIAN PUSTAKA.....</b>	<b>9</b>
2.1 Sel Darah Putih .....	9
2.2 <i>Acute Myeloid Leukemia (AML)</i> .....	10
2.3 Ruang Warna HSV.....	12
2.4 <i>Cropping</i> Citra .....	12
2.5 <i>Gradient</i> Citra .....	13
2.6 Analisis Cluster <i>Hierarchical Clustering Anlaysia</i> dan K-Means .....	13
2.6.1 <i>Hierarchical Clustering Analysis</i> .....	14
2.6.2 Analisis K-Means Cluster .....	18
<b>BAB 3 METODOLOGI PENELITIAN.....</b>	<b>21</b>
3.1 Tahapan Penelitian .....	21
3.1.1 Studi Literatur .....	21
3.1.2 Analisis Masalah dan Perancangan Penelitian.....	22
3.1.3 Implementasi Metode.....	22

3.1.4 Uji Coba dan Analisis Hasil Penelitian.....	22
3.1.5 Dokumentasi .....	23
3.2 Dataset Citra AML.....	23
3.3 Rancangan Penelitian.....	23
3.3.1 Penentuan <i>Initial</i> ROI Sel Darah Putih.....	25
3.3.2 <i>Cropping</i> Citra AML Berdasarkan <i>Initial</i> ROI Sel Darah Putih .....	26
3.3.3 Segmentasi Sel Darah Putih Pada Sub-Citra AML.....	30
3.3.4 Pemisahan Sel Darah Putih yang Bersentuhan Pada Sub-Citra AML..	31
3.3.5 Pembersihan Citra ( <i>Image Cleaning</i> ).....	35
3.4 Perancangan Uji Coba.....	35
3.4.1 Skenario Uji Coba.....	36
3.4.2 Evaluasi Hasil Segmentasi .....	36
<b>BAB 4 UJI COBA DAN ANALISIS HASIL .....</b>	<b>39</b>
4.1 Lingkungan Uji Coba.....	39
4.2 Data Uji Coba.....	39
4.3 Hasil Uji Coba.....	40
4.3.1 Hasil Uji Coba Segmentasi Citra AML .....	40
4.3.2 Skenario Uji Coba Segmentasi Citra AML .....	46
4.3.3 Skenario Uji Coba Pemisahan Sel Darah Putih pada Citra AML.....	50
4.4 Analisis Hasil Uji Coba.....	52
<b>BAB 5 PENUTUP .....</b>	<b>57</b>
5.1 Kesimpulan .....	57
5.2 Saran.....	58
<b>DAFTAR PUSTAKA.....</b>	<b>61</b>
<b>LAMPIRAN I.....</b>	<b>63</b>
<b>LAMPIRAN II .....</b>	<b>69</b>
<b>LAMPIRAN III.....</b>	<b>71</b>
<b>BIODATA PENULIS.....</b>	<b>73</b>

## DAFTAR GAMBAR

Gambar 2. 1 Struktur sel darah putih .....	10
Gambar 2. 2 Perkembangan del darah .....	10
Gambar 2. 3 Visualisasi ruang warna HSV .....	12
Gambar 2. 4 <i>Dendogram</i> .....	15
Gambar 2. 5 Ilustrasi algoritma AHC .....	15
Gambar 2. 6 Ilustrasi pada jarak <i>single link</i> .....	16
Gambar 2. 7 Ilustrasi pada jarak <i>complete link</i> .....	17
Gambar 2. 8 Ilustrasi pada jarak <i>centroid link</i> .....	17
Gambar 2. 9 Ilustrasi pada jarak <i>average link</i> .....	18
Gambar 3. 1 Tahapan penelitian .....	21
Gambar 3. 2 Citra AML.....	23
Gambar 3. 3 Rancangan metode yang diusulkan.....	24
Gambar 3. 4 Proses penentuan <i>initial ROI</i> sel darah putih pada citra AML.....	26
Gambar 3. 5 Diagram proses <i>cropping</i> citra <i>initial ROI</i> sel darah putih .....	27
Gambar 3. 6 Ilustrasi <i>bounding box region initial ROI</i> sel darah putih.....	27
Gambar 3. 7 Alur penentuan <i>bounding box</i> awal.....	28
Gambar 3. 8 Ilustrasi <i>bounding box</i> final.....	29
Gambar 3.9 Proses segmentasi sel darah putih pada citra AML dengan K-Means ...	31
Gambar 3. 10 Proses pemisahan <i>touching cell</i> menggunakan HCA.....	32
Gambar 3. 11 Fitur ruang informasi spasial pada analisis <i>cluster</i> .....	33
Gambar 3. 12 Fitur <i>gradient magnitude</i> spasial pada analisis <i>cluster</i> .....	34
Gambar 4. 1 Hasil uji coba penentuan <i>initial ROI</i> sel darah putih .....	40
Gambar 4. 2 <i>Bounding box initial ROI</i> sel darah putih .....	41
Gambar 4. 3 Hasil sub-citra setelah proses <i>cropping</i> citra AML .....	41
Gambar 4.4 Hasil segmentasi sel darah putih dengan penggabungan cluster dari region nukleus dan sitoplasma.....	43
Gambar 4. 5 Hasil pemisahan dengan <i>hierarchical clustering analysis</i> .....	44
Gambar 4. 6 Hasil penggabungan sub-citra ke dimensi awal.....	45
Gambar 4. 7 Hasil akhir segmentasi sel darah putih citra AML.....	46
Gambar 4.8 Perbandingan hasil segmentasi citra AML dengan citra <i>ground-truth</i> ..	47

Gambar 4. 9 Evaluasi hasil segmentasi sel darah putih berdasarkan ZSI.....	49
Gambar 4. 10 Evaluasi hasil segmentasi nukleus berdasarkan ZSI.....	50
Gambar 4. 11 Jumlah sel darah putih teridentifikasi .....	52
Gambar 4. 12 Perbandingan hasil segmentasi pada <i>region</i> sitoplasma .....	54
Gambar 4. 13 Hasil pemisahan sel pada citra aml_31 di sub-citra 7.....	55
Gambar 4. 14 Hasil pemisahan sel pada citra aml_19 di sub-citra 8.....	55

## DAFTAR TABEL

Tabel 2. 1 Sub-tipe AML menurut sistem klasifikasi FAB .....	11
Tabel 4. 1 Proses segmentasi pada sub-citra AML.....	42
Tabel 4. 2 Hasil proses <i>image cleaning</i> pada citra AML.....	46
Tabel 4. 3 Evaluasi hasil segmentasi sel darah putih.....	48
Tabel 4. 4 Nilai rata-rata hasil segmentasi sel darah putih .....	48
Tabel 4. 5 Evaluasi hasil segmentasi <i>region</i> nukleus .....	49
Tabel 4. 6 Nilai rata-rata hasil segmentasi nukleus .....	50
Tabel 4.7 Perbandingan antara jumlah sel darah putih dengan perhitungan manual dan yang teridentifikasi menggunakan metode yang diusulkan.....	51

*[Halaman ini sengaja dikosongkan]*

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang

Pada tubuh manusia, darah merupakan cairan merah yang terdiri dari cairan *isotonic (plasma)* dimana berbagai sel (*hemocytes*) berada. Terdapat 3 bagian besar dari sel-sel tersebut dan semua sel darah tersebut dibuat di *bone marrow*, tumbuh dari sel yang dikenal sebagai *hematopoietic stem cell* (Ciesla, 2007). Sel darah merah (*Erythrocytes*) berisi hemoglobin yang membawa oksigen ke setiap jaringan, *platelets (Thrombocytes)* memiliki tugas membekukan darah dan sel darah putih (*Leukocytes*) merupakan sel yang berfungsi sebagai sistem imun yang melindungi tubuh dari infeksi penyakit dan material asing (Ciesla, 2007). Terdapat lima jenis sel darah putih yang membentuk respon untuk sistem kekebalan tubuh yang memiliki fungsi masing-masing yaitu *neutrophils*, *lymphocytes*, *monocytes*, *eosinophils*, dan *basophils*. Kelima tipe sel darah putih tersebut dibagi kedalam dua jenis yaitu *granulocytes* dan *agranulocytes*. Apusan darah umum digunakan pada laboratorium klinis untuk memberikan indikasi mengenai berbagai penyakit dan membantu diagnosa, mengamati dan mengevaluasi kasus pada seorang pasien (Madhloom et al, 2010).

Kelainan pada darah dianggap sebagai salah satu penyakit yang paling berbahaya yang dapat menyebabkan kematian dan banyak dari penyakit pada darah berhubungan dengan sel darah putih seperti Leukimia (Madhloom et al, 2010). Terdapat dua jenis leukemia yaitu akut dan kronis. Salah satu jenis leukemia akut menurut klasifikasi French-American-British (FAB) yang akan dibahas pada penelitian ini adalah *Acute Myeloid Leukemia (AML)*. AML juga dikenal sebagai *acute myelogenous leukemia* atau *acute nonlymphocytic leukemia (ANLL)*. AML merupakan jenis kanker yang ditandai dengan infiltrasi pada *bone marrow*, darah dan jaringan lainnya oleh sel proliferative, klonal dan tidak normal dari sistem *hematopoietic* (Dohner dkk, 2015). Pertumbuhan sel darah putih yang tidak normal tersebut mempengaruhi produksi sel darah normal pada *bone marrow*. AML terdiri dari beberapa jenis yang perawatan dan penyembuhannya berbagai macam pada

setiap jenis tersebut. Meskipun AML termasuk penyakit yang relatif jarang, di amerika serikat kasus ini memiliki persentase sebesar 1,2% untuk kasus kematian yang disebabkan oleh kanker dan diprediksi akan meningkat seiring dengan penambahan umur penduduk (Jemal A. dkk, 2002).

Observasi dari apusan darah melalui mikroskop memberikan informasi kualitatif dan kuantitatif yang penting untuk membantu diagnosa dari berbagai penyakit termasuk Leukimia (Theera-Umpon, 2005). Identifikasi dan klasifikasi dari sel darah putih ke beberapa sub-jenis sangat bernilai penting di pengujian klinis dan labolatorium, termasuk untuk diagnosis leukemia tipe AML. Akan tetapi observasi yang dilakukan oleh ahli secara manual dalam menghitung dan mengklasifikasikan sel darah putih menggunakan mikroskop sangat bergantung pada tingkat keahlian dan pengalaman. Pekerjaan tersebut juga menghabiskan cukup banyak waktu, seorang ahli memerlukan waktu sekitar 15 menit untuk mengevaluasi dan menghitung 100 sel pada setiap *slide* darah dan hal tersebut rentan terhadap kesalahan prosedur (Sabino dkk, 2004). Keterbatasan dari faktor teknis dan kelelahan tersebut menyebabkan efisiensi dan kualitas diagnosis kemungkinan menurun serta resiko terjadinya kesalahan analisis menjadi lebih tinggi. Oleh sebab itu, suatu sistem otomatis dapat membantu ahli dalam melakukan diagnosis dan mempersingkat waktu pemeriksaan (Rezatofighi dkk, 2011).

Pengenalan sel darah putih secara otomatis pada pengolahan citra hematologi pada umumnya terdiri dari 4 tahapan utama yaitu praproses, segmentasi, ekstraksi fitur dan klasifikasi. Pada penelitian ini tahapan segmentasi menjadi fokus utama penelitian. Tahap segmentasi menjadi sangat penting karena akurasi pada tahap ekstraksi fitur dan klasifikasi bergantung pada seberapa baik hasil segmentasi sel darah putih. Citra yang dianalisis adalah citra mikroskopis sediaan apus darah tepi (*peripheral blood smear images*) yang mengandung *myeloblasts* (Leukimia tipe AML). Secara umum pada citra mikroskopis sediaan apus darah tepi terdapat tiga bagian, yaitu sel darah merah, sel darah putih dan latar belakang. Pada sel darah putih sendiri terdapat dua wilayah (*region*) utama, yaitu nukleus atau inti sel dan sitoplasma (Faticah dkk, 2012). Pada tahap segmentasi citra *region* sel darah putih akan dipisahkan dengan *region* lainnya dan kemudian dilanjutkan dengan pemisahan antara *region* nukleus dan sitoplasma.

Penelitian yang fokus pada segmentasi sel darah putih dengan pendekatan atau metode yang berbeda-beda telah sering dilakukan sebelumnya. Beberapa penelitian menggunakan citra apus darah normal (Piuri dkk, 2004; Scotti, 2006; Wang dkk, 2009; Rezatofighi dkk, 2011; Huang dkk, 2012) dan juga citra apus darah tepi leukemia (Madhloom dkk, 2012; Mohapatra dkk, 2014; Putzu dkk, 2014).

Segmentasi dengan metode berbasis *thresholding* (nilai ambang) diusulkan oleh Huang, dkk (2012) dan Putzu, dkk (2014). Huang, dkk (2012) melakukan pengembangan pada salah satu metode *thresholding* yang populer digunakan yaitu metode Otsu untuk segmentasi pada *region* nukleus dengan mencari nilai ambang jamak. Pada penelitian Putzu, dkk (2014) segmentasi sel darah putih dilakukan dengan nilai ambang tunggal dari perhitungan algoritma *Zack* (*Zack Thresholding*). Pada metode berbasis *thresholding* segmentasi citra sangat bergantung pada beberapa faktor seperti tingkat intensitas dan saturasi serta perbandingan antara jumlah piksel sel darah putih dengan komponen lainnya sehingga variasi dari dataset yang beragam dapat menyebabkan hasil segmentasi yang didapatkan kurang maksimal.

Metode yang berbasis tepi (*boundary*) untuk segmentasi sel darah putih telah diperkenalkan oleh Piuri, dkk (2004). Metode yang diusulkan tersebut melakukan segmentasi pada sel darah putih dengan mendeteksi tepi citra dan kemudian dilanjutkan dengan operasi morfologi untuk memperbaiki hasil segmentasi. Permasalahan pada metode yang berbasis pada deteksi tepi objek terjadi saat hasil deteksi pada *region of interest* (ROI) terdapat garis tepi yang tidak terhubung, sehingga batas dari *region* tidak dapat diketahui.

Metode segmentasi lain yang berbasis wilayah (*region*) yaitu metode orthogonalisasi Gram-Schmidt dan *active contour* diperkenalkan oleh Rezatofighi, dkk (2011) dan pengembangan dari metode Gram-Schmidt yang dikombinasikan dengan *region growing* diusulkan oleh Abuhasel, dkk (2015). Metode Gram-Schmidt melakukan orthogonalisasi pada sekumpulan vector dalam *inner product space* dengan intensitas piksel citra RGB untuk mendapatkan area nukleus. Persentase keberhasilan metode Gram-Schmidt dalam melakukan segmentasi menunjukkan hasil yang baik, namun untuk citra dengan komposisi warna RGB yang

berbeda memerlukan inisialisasi ulang untuk menentukan vektor warna dari objek yang ingin disegmentasi. Oleh sebab itu metode Gram-Schmidt kurang efektif untuk segmentasi sel darah putih pada dataset yang memiliki variasi *staining* yang beragam.

Metode berbasis *clustering* untuk segmentasi sel darah putih diusulkan oleh Scotti (2006), Wang, dkk (2009) dan Mohapatra, dkk (2014). Pada penelitian Scotti (2006) dan Mohapatra, dkk (2014), ROI dari citra mikroskopis dibagi menjadi tiga *cluster* yaitu sel darah putih, sel darah merah dan latar belakang. Sementara itu Wang, dkk (2009) membagi ROI citra mikroskopis ke dalam enam *cluster* yaitu tiga *cluster* untuk variasi dari warna nukleus, satu *cluster* untuk sitoplasma dan dua *cluster* untuk variasi warna dari sel darah merah. Pada metode berbasis *clustering* informasi warna dan spasial yang digunakan dan penentuan jumlah kluster yang tepat sangat menentukan untuk mendapatkan hasil segmentasi sel darah putih yang baik, namun metode berbasis *clustering* merupakan metode yang efektif untuk dapat melakukan segmentasi pada sel darah putih dengan citra mikroskopis yang memiliki variasi *staining* yang beragam.

Permasalahan lain yang berpengaruh dalam proses analisis sel darah putih adalah keberadaan sel yang bersentuhan (*touching*) sehingga sel-sel tersebut perlu dideteksi dan dipisahkan (Putzu dkk, 2014). Proses identifikasi dan klasifikasi sel darah putih memerlukan informasi dari tiap sel (*single cell*) dan untuk mendapatkan informasi tersebut sel yang bersentuhan harus dipisahkan dengan baik agar dapat meningkatkan akurasi pada tahap tersebut. Selain itu perhitungan jumlah sel darah putih juga penting dalam proses diagnosis ataupun tindakan medis yang dilakukan, oleh karena itu pemisahan sel yang bersentuhan yang baik akan meningkatkan akurasi perhitungan sel darah putih.

Penelitian pada segmentasi sel bersentuhan atau bertumpuk masih berkembang dan menjadi tantangan bagi peneliti yang fokus pada bidang segmentasi sel (Jung, 2010). Salah satu metode yang sering digunakan untuk memisahkan sel yang bertumpuk adalah metode *watershed*. Penelitian untuk pemisahan sel dengan metode *watershed* diusulkan oleh Yang, dkk (2006), Cheng, dkk (2009) dan Putzu, dkk (2014).

Penelitian oleh Yang, dkk (2006) memperkenalkan metode *condition-erosion* (CE) *watershed* yang digunakan pada citra mikroskopik berflourensi dan merupakan pengembangan dari *marker controlled watershed*. Metode ini dimulai dengan proses erosi dengan batas nilai ambang (*threshold*) tertentu dan akan menjadi *extracted marker* sebagai pengganti region minimum pada *watershed*. Pada penelitian lain yang juga menggunakan *watershed*, Cheng, dkk (2009) mengusulkan metode *active contour* sebagai inisialisasi awal segmentasi dan *shape marker* sebagai *marker extraction* didapatkan melalui ekstraksi *adaptive H-minima transform*. Untuk memisahkan *region* pada *watershed* digunakan *Marking function* yang dibentuk dengan fungsi jarak *euclidean distance*. Namun pada kedua metode tersebut jika *marker* tidak merepresentasikan objek dengan baik akan mengakibatkan terjadinya *undersegmentation*.

Pada sel darah putih, penelitian yang menggunakan *watershed* untuk memisahkan sel yang bersentuhan diusulkan oleh Putzu, dkk (2014). Citra sel darah putih diidentifikasi untuk penentuan sel yang bersentuhan atau tidak menggunakan analisis pada kebulatan (*roundness*) sel darah putih. Sel darah putih yang memiliki nilai kebulatan di bawah *threshold* akan dianggap sebagai sel yang berkumpul atau bersentuhan dan kemudian dipisahkan menggunakan metode *watershed* dengan perhitungan *distance transform*. Kontur hasil pemisahan sel darah putih dihaluskan dengan *line refining* dari ekstraksi pada transformasi *watershed*. Hasil pemisahan sel menunjukkan hasil yang baik pada jumlah tertentu, namun pada citra sel darah putih yang memiliki sel bersentuhan yang cukup banyak terjadi penurunan pada akurasi segmentasi.

Pendekatan lain yang berbasis *clustering* untuk pemisahan sel yang bersentuhan diusulkan oleh Jung, dkk (2010) dan Effendy (2013). Pada penelitian Jung, dkk (2010) pemisahan dari sel yang bertumpuk pada citra sel serviks diformulasikan dengan analisis *cluster*. *Cluster validity index* digunakan sebagai parameter untuk mengetahui jumlah sel yang *overlap* yang kemudian cluster dianalisis menggunakan *Unsupervised Bayesian Classifier*. Sementara itu Effendy (2013) mengusulkan metode pemisahan pada citra sel darah merah menggunakan metode K-Means. Citra sel darah merah pada tahap awal di segmentasi menggunakan operasi morfologi dan untuk menentukan keberadaan dari sel darah

merah yang bertumpuk pada citra digunakan fitur geometri. Pada sel yang bertumpuk *cluster* akan dianalisis dan dipisahkan dengan metode *clustering* K-Means. Metode berbasis *clustering* yang diusulkan tersebut berhasil memisahkan sel dengan lebih baik dibanding metode lainnya (Effendy, 2013), namun jumlah *cluster* masih belum dapat ditentukan secara langsung dan harus ditentukan terlebih dahulu sebelum proses analisis *cluster*.

Pada tesis ini diusulkan metode segmentasi dan pemisahan sel darah putih bersentuhan pada citra AML menggunakan analisis *cluster* dengan metode K-Means dan *hierarchical clustering analysis*. Segmentasi dan pemisahan sel akan dilakukan pada masing-masing *region* lokal atau sub-citra. Sub-citra didapatkan dari *cropping* pada citra AML berdasarkan *initial* ROI sel darah putih menggunakan metode *thresholding* Otsu pada citra yang diperoleh dengan proses *Leukocytes Nucleus Enhancer* (LNE) (Huang dkk, 2012). Dalam analisis *cluster* digunakan informasi warna yaitu komponen S pada ruang warna HSV untuk mengatasi efek iluminasi akibat pencahayaan yang kurang baik. Segmentasi dilakukan menggunakan analisis *cluster* dengan metode K-Means untuk mengatasi variasi *staining* yang beragam pada citra AML. Untuk mendeteksi sel yang bersentuhan digunakan fitur geometri berupa kebulatan (*roundness*) dan luasan (*size*). Dari hasil segmentasi dan deteksi tersebut, sel yang bersentuhan dipisahkan dengan metode *hierarchical clustering analysis* secara otomatis tanpa penentuan jumlah *cluster* terlebih dahulu.

## 1.2. Perumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana menentukan *initial Region of Interest* sel darah putih serta membaginya menjadi sub-citra?
2. Bagaimana menentukan fitur geometri yang tepat untuk mengidentifikasi sel darah putih yang bersentuhan (*touching cells*)?
3. Bagaimana memisahkan sel yang bersentuhan (*touching cell*) menggunakan metode *Hierarchical Clustering Analysis*?

### **1.3. Batasan Masalah**

Permasalahan yang dibahas pada penelitian ini memiliki beberapa batasan sebagai berikut:

1. Segmentasi dilakukan pada citra mikroskopis sediaan apus darah tepi yang diambil dari Laboratorium Kesehatan Pemerintah Daerah Kalimantan Selatan.
2. Informasi yang digunakan dalam analisis *cluster* yaitu informasi warna dan informasi spasial.

### **1.4. Tujuan**

Tujuan dari penelitian ini adalah melakukan segmentasi dan pemisahan sel darah putih pada citra AML menggunakan metode K-Means dan *Hierarchical Clustering Analysis*.

### **1.5. Manfaat Penelitian**

Manfaat penelitian ini yaitu mendapatkan fitur *region* dari sel darah putih yang dapat digunakan untuk mengidentifikasi leukemia jenis AML secara otomatis sehingga membantu proses diagnosis penyakit AML di laboratorium.

### **1.6. Kontribusi Penelitian**

Kontribusi yang diharapkan dari penelitian ini adalah hasil partisi cluster yang optimal dari region sel darah putih termasuk nukleus dan sitoplasma dengan algoritma K-Means untuk segmentasi sel darah putih, serta hasil cluster sel darah putih tunggal yang stabil pada pemisahan touching cell menggunakan algoritma hierarchical clustering analysis untuk setiap sub-citra. Segmentasi dilakukan pada dataset citra AML yang memiliki variasi staining beragam. Sub-citra diperoleh dengan melakukan cropping pada citra AML berdasarkan initial ROI sel darah putih yang didapatkan dengan metode Leukocytes Nucleus Enhancer (LNE) dan Thresholding Otsu.

### **1.7. Sistematika Penulisan**

Sistematika penulisan proposal tesis ini adalah:

1. **BAB I PENDAHULUAN.** Bab ini berisi pendahuluan yang menjelaskan latar belakang, perumusan masalah, tujuan, batasan masalah, tujuan, manfaat, kontribusi penelitian dan sistematika penulisan terkait dengan penelitian.
2. **BAB II DASAR TEORI DAN TINJAUAN PUSTAKA.** Bab ini berisi dasar-dasar teori dan tinjauan pustaka terkait penelitian, diantaranya mengenai sel darah putih, ruang warna HSV, *cropping* citra dan pelabelan komponen, fitur geometri dan analisis *cluster* dengan metode K-means dan *hierarchical clustering analysis*.
3. **BAB III METODOLOGI PENELITIAN.** Bab ini berisi tahapan penelitian meliputi studi literatur, rancangan penelitian dan skenario pengujian terhadap penelitian ini.

## **BAB 2**

### **DASAR TEORI DAN KAJIAN PUSTAKA**

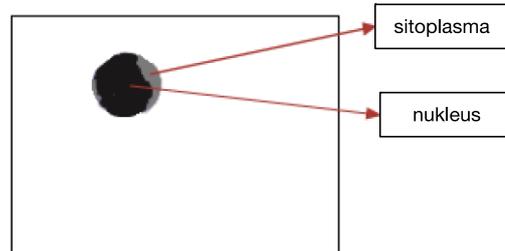
Pada bab ini dijelaskan mengenai dasar-dasar teori dan kajian pustaka yang berkaitan dengan penelitian.

#### **2.1 Sel Darah Putih**

Sel darah putih atau disebut juga leukosit merupakan sel darah yang bertugas sebagai sistem imun pada tubuh manusia. Sel darah putih yang diproduksi dalam sumsum tulang terdiri dari lima jenis sel. Lima jenis sel darah putih dibagi ke dalam dua tipe berdasarkan keberadaan granula yaitu granuler dan agranuler (Huang dkk, 2012). Sel darah putih tipe granuler adalah neutrofil, eosinofil, dan basofil, sedangkan tipe agranuler adalah limfosit dan monosit. Leukosit terdiri dari sitoplasma dan nukleus. Keseluruhan sel disebut membran. Gambar 2.1 menunjukkan sel darah putih beserta sitoplasma dan nukleusnya. Selain morfologinya, jumlah leukosit juga dapat digunakan sebagai referensi untuk mengidentifikasi penyakit leukemia. Ketika terjadi intrusi ke dalam tubuh manusia yang mengakibatkan infeksi, jumlah leukosit akan meningkat untuk mempertahankan daya tahan tubuh. Peningkatan atau penurunan jumlah sel darah putih merupakan indikator adanya infeksi dalam tubuh. Namun, infeksi bukan merupakan penyebab satu-satunya peningkatan jumlah leukosit, penyimpangan proliferasi sel agranuler mengakibatkan peningkatan jumlah sel darah putih terutama sel limfosit atau monosit dan penurunan jumlah sel darah putih jenis lainnya (Wartenberg dkk, 2008). Penyimpangan yang terjadi pada sel darah putih tipe agranuler tersebut disebut neoplasma ganas sel darah atau biasa disebut kanker darah (leukemia).

Pemeriksaan sel darah putih dilakukan menggunakan mikroskop pada sediaan apus darah (*blood smear*) yang telah dicat (*stained*). Pengecatan dilakukan karena leukosit bersifat transparan dan pucat dan tidak dapat diobservasi di mikroskop hingga sediaan apus diberi warna. Observasi sediaan apus darah oleh tenaga terlatih

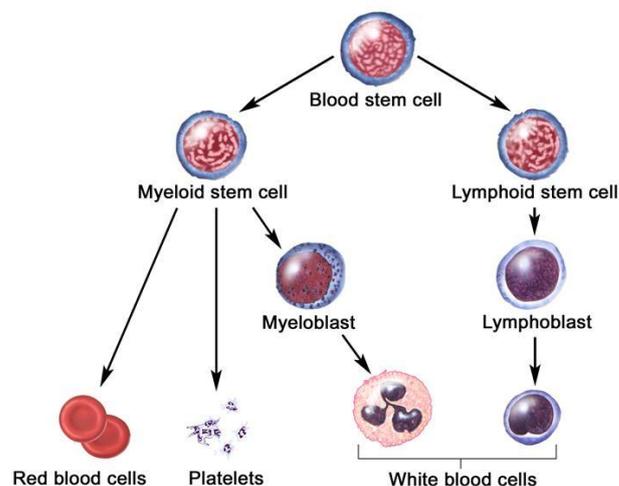
di laboratorium memerlukan waktu sekitar 15 menit untuk menghitung 100 sel pada tiap sediaan apus (Huang dkk, 2012).



Gambar 2.1. Struktur sel darah putih.

## 2.2 *Acute Myeloid Leukemia (AML)*

*Acute Myeloid Leukemia (AML)* yang juga dikenal sebagai *acute myelogenous leukemia* atau *acute nonlymphocytic leukemia (ANLL)* adalah kanker yang berasal dari garis sel darah myeloid, ditandai dengan pertumbuhan sel darah putih yang cepat yang terjadi di *bone marrow* dan mengganggu produksi dari sel darah yang normal. AML merupakan jenis kanker yang ditandai dengan infiltrasi pada bone marrow, darah dan jaringan lainnya oleh sel proliferative, klonal dan tidak normal dari sistem *hematopoietic* (Dohner dkk, 2015). Gejala yang muncul biasanya adalah mudah lelah, sesak nafas, mudah memar dan pendarahan, dan peningkatan terjadinya infeksi. Perkembangan dari sel darah digambarkan pada Gambar 2.2 dimana sel *myeloid* dapat berkembang menjadi beberapa jenis sel yaitu sel darah merah, sel darah putih dan *platelets* (National Cancer Institute, 2016).



Gambar 2.2. Perkembangan sel darah. (National Cancer Institute, 2016)

AML merupakan leukemia akut yang umumnya terjadi pada orang dewasa dan pengaruhnya meningkat seiring dengan umur penderita. meskipun AML termasuk penyakit yang jarang terjadi, namun menurut penelitian Jemal, dkk di tahun 2002 tentang statistik kanker AML memiliki persentase sekitar 1,2% sebagai penyebab terjadinya kematian di Amerika Serikat. Kejadiannya di perkirakan meningkat seiring dengan umur populasi. Beberapa faktor resiko dan kromosom yang tidak normal telah dapat diidentifikasi, namun penyebab yang lebih spesifik masih belum jelas. Sebagai leukemia tipe akut, AML berkembang secara cepat dan dapat berakibat fatal jika dalam jangka waktu beberapa minggu atau bulan tidak segera diobati.

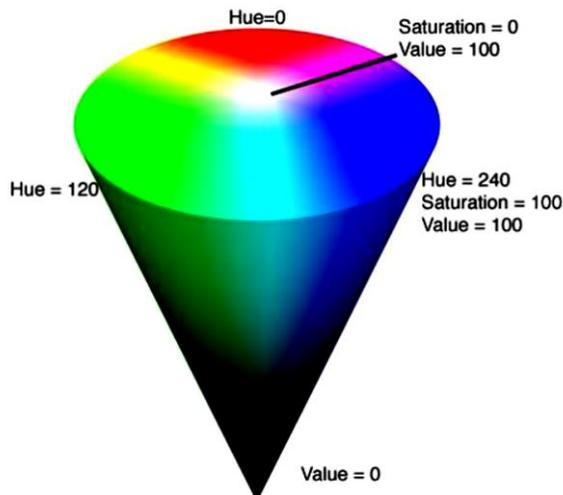
Menurut sistem klasifikasi FAB (*French-American-British*) AML memiliki beberapa sub-tipe yang ditunjukkan oleh tabel 2.1 dan memiliki pengobatan dan prognosa berbeda pada tiap sub-tipe. AML memiliki persentase kesembuhan sekitar 35%-40% pada penderita dengan umur di bawah 60 tahun dan 5-15% pada umur diatas 60 tahun. Penderita yang memiliki umur lebih tua yang tidak dapat menahan kemoterapi secara intensif memiliki peluang bertahan sekitar 5-10 bulan (Dohner dkk, 2015).

Tabel 2.1 Sub-tipe AML menurut sistem klasifikasi FAB.

Tipe	Nama Sub-jenis AML	Persentase Penderita (Dewasa)
M0	Acute myeloblastic leukemia ( <i>minimally differentiated</i> )	5%
M1	Acute myeloblastic leukemia ( <i>without maturation</i> )	15%
M2	Acute myeloblastic leukemia ( <i>granulocytic maturation</i> )	25%
M3	Acute promyelocytic leukemia (APL)	10%
M4	Acute myelomonocytic leukemia	20%
M4eo	Myelomonocytic (dengan <i>bone marrow eosinophilia</i> )	5%
M5	<ul style="list-style-type: none"> <li>▪ Acute monoblastic leukemia (M5a)</li> <li>▪ Acute monocytic leukemia (M5b)</li> </ul>	10%
M6	Acute erythroid leukemia <ul style="list-style-type: none"> <li>▪ Erythroleukemia (M6a)</li> <li>▪ Erythroid leukemia (M6b)</li> </ul>	5%
M7	Acute megakaryoblastic leukemia	5%

### 2.3 Ruang Warna HSV

HSV merupakan singkatan dari *Hue*, *Saturation*, *Value*. *Hue* adalah paduan warna, *saturation* adalah kepekatan warna, dan *value* adalah nilai kecemerlangan warna (Cardani, 2006). Ruang warna ini memperhatikan unsur cahaya yang dapat digambarkan sebagai suatu bentuk kerucut seperti yang terlihat pada Gambar 2.3.



Gambar 2.3. Visualisasi ruang warna HSV.

Pada Gambar 2.3, *hue* digambarkan sebagai lingkaran yang ada pada permukaan (kulit luar) kerucut dengan satuan derajat ( $^{\circ}$ ) dimana  $0^{\circ}$  berwarna merah,  $120^{\circ}$  berwarna hijau,  $240^{\circ}$  berwarna biru, dan seterusnya. *Saturation* dihitung dari titik tengah kerucut sampai pinggir kerucut yang memiliki rentang 0-100 dimana 0 berada di titik tengah kerucut dan 100 pada pinggiran kerucut. Sedangkan *value* dari bagian pangkal (permukaan lebar kerucut) yang bernilai 100 sampai ujung kerucut yang berbentuk runcing (bernilai 0) dimana bagian ujung berwarna hitam karena kecemerlangannya 0 dan permukaan lingkaran pada bagian pangkal yang berwarna putih karena kecemerlangannya maksimum, yaitu 100 (Cardani, 2006)

### 2.4 Cropping Citra

*Cropping* merupakan proses pemotongan citra berdasarkan *bounding box region* tertentu dalam citra. *Bounding box* adalah kotak terkecil yang melingkupi *region* komponen terhubung atau singkatnya batas *region* komponen terhubung.

*Cropping* dilakukan dengan menyusun citra baru dengan piksel-piksel penyusunnya adalah piksel yang terdapat dalam *bounding box region* tertentu pada citra asli.

## 2.5 Gradient Citra

*Gradient* citra merupakan informasi perubahan arah pada intensitas atau warna pada citra. Pada perangkat lunak grafis untuk *editing* citra digital, istilah *gradient* atau warna *gradient* digunakan untuk perubahan bertahap antara warna yang dapat diartikan juga sebagai gradasi dari nilai yang rendah ke tinggi.

Secara matematis *gradient* dari fungsi dua variabel yang pada citra adalah fungsi intensitas pada setiap titik citra merupakan vektor dua dimensi dengan komponen yang dibentuk dari arah *horizontal* dan *vertical*. Pada setiap titik citra, vektor *gradient* menuju ke arah perubahan intensitas terbesar dan panjang dari vektor *gradient* berhubungan dengan nilai perubahan dari arah tersebut.

Citra *gradient* dibuat dari citra asli yang secara umum meng-konvolusinya dengan suatu filter misalnya sobel filter atau prewitt filter. Setiap pixel pada citra *gradient* mengukur perubahan intensitas pada titik yang sama seperti di citra asli dengan arah tertentu. Untuk mendapatkan arah dari semua jarak, citra *gradient* pada arah x dan y dihitung. *Gradient* dari fungsi dua dimensi  $f(x,y)$  di definisikan oleh vektor pada persamaan 2.1 (Gonzalez, 2002).

$$\nabla f = \begin{pmatrix} G_x \\ G_y \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} \quad (2.1)$$

dengan  $G_x$  adalah *gradient* pada arah vektor x dan  $G_y$  pada arah vektor y. Sedangkan *magnitude* dari *gradient* pada citra dihitung dengan menggunakan persamaan 2.2.

$$\nabla f \approx |G_x| + |G_y| \quad (2.2)$$

## 2.6 Analisis Cluster *Hierarchical Clustering Analysis* dan K-Means

Pada sub-bab ini membahas tentang analisis *cluster* yaitu *hierarchical clustering analysis* dan K-Means.

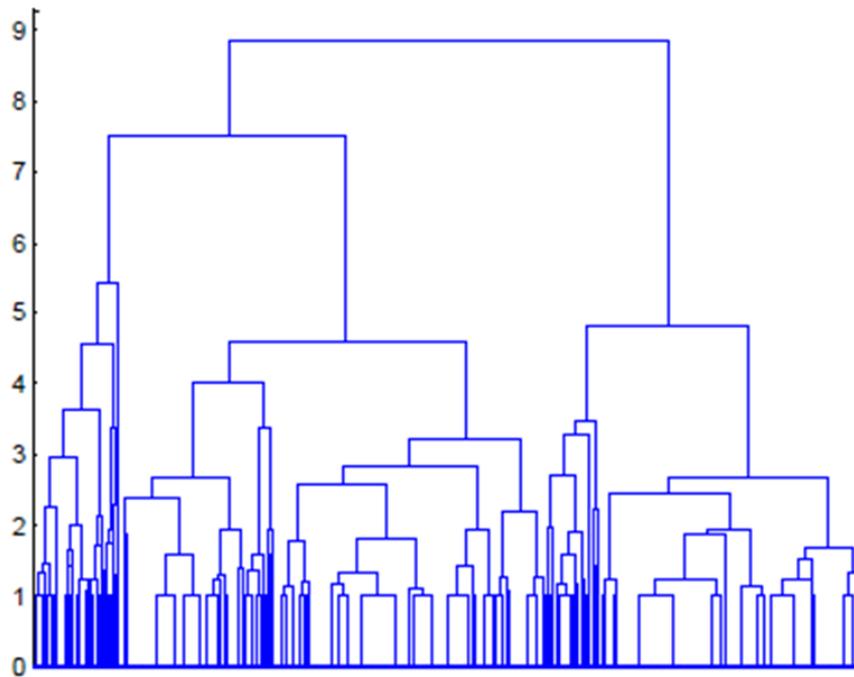
### 2.6.1 Hierarchical Clustering Analysis

Algoritma *hierarchical clustering* menghasilkan *cluster* yang bertingkat (*hierarki*) dengan satu buah partisi yang tersusun atas keseluruhan data terletak dibagian atas dan dengan kelompok-kelompok kecil yang tersusun atas satu data terletak dibagian bawah. Masing-masing tingkat perantara dapat dilihat sebagai kombinasi dua kelompok dari tingkat terbawah sebelumnya atau melakukan pemotongan kelompok dari tingkat teratas sebelumnya. Hasil dari algoritma *hierarchical* dapat digambarkan dalam bentuk *tree* yang disebut dengan *dendogram*. Gambar 2.4 menunjukkan bentuk *dendogram*.

Ada dua pendekatan umum dalam *hierarchical clustering analysis*, yaitu pendekatan *agglomerative* dan *divisive*. Pendekatan *agglomerative* dimulai dengan setiap satu data sebagai sebuah kelompok kemudian dalam setiap langkah dilakukan penggabungan terhadap kelompok-kelompok tersebut yang dinilai paling mirip atau dekat jaraknya. Pendekatan *divisive* berarti berawal dari satu kelompok besar yang tersusun atas keseluruhan data, kemudian dalam setiap langkah dilakukan pemotongan pada *hierarki* sampai tersusun kelompok-kelompok yang hanya beranggotakan satu data. Dalam hal ini artinya pada setiap langkah ada penentuan kelompok mana yang akan di partisi dan berdasarkan apa melakukan partisi pada *hierarki* tersebut dilakukan.

Berikut ini merupakan langkah-langkah dalam algoritma *hierarchical agglomerative clustering* secara umum (Adi, 2007):

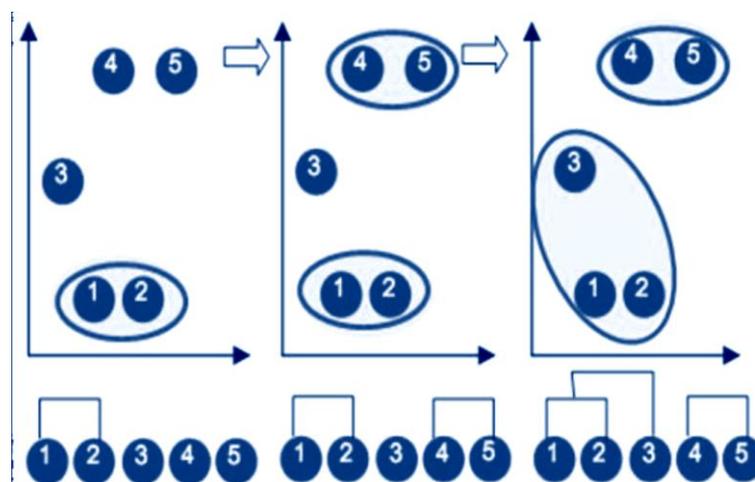
1. Menentukan  $k$  sebagai jumlah cluster yang ingin dibentuk.
2. Setiap data dianggap sebagai satu cluster (*trivial*). Jika terdapat  $N$  jumlah data, dan  $c$  jumlah cluster, maka jumlah  $c=N$ .
3. Menghitung jarak antar cluster pada *hierarki* dengan menggunakan salah satu perhitungan jarak, seperti *euclidean*, *manhattan*, *maximum* atau *mahalanobis distance*.
4. Tentukan dua buah *cluster* yang memiliki jarak paling dekat (*minimum*) dan gabungkan menjadi satu buah *cluster* dan jumlah *cluster*  $c$  menjadi  $c=c-1$ .
5. Bandingkan nilai  $c$  dengan  $k$ , jika nilai  $c>k$ , maka kembali ke-langkah tiga atau jika  $c=k$  maka hentikan proses pembagian *cluster*.



Gambar 2.4. Dendogram. (Koeanan, 2009)

Gambar 2.5 menunjukkan tahapan penggabungan *cluster* dari algoritma *agglomerative hierarchical clustering*. Parameter utama dalam algoritma *agglomerative hierarchical* adalah metode yang digunakan untuk menentukan pasangan kelompok mana yang akan digabungkan pada setiap langkah.

Terdapat beberapa skema yang dapat di gunakan dalam penghitungan jarak yang dilakukan pada *agglomerative hierarchical clustering*, yaitu *single link*, *complete link*, *centroid link*, dan *average link* (Lance, 1967).

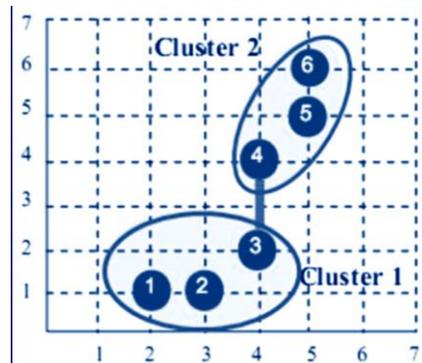


Gambar 2.5. Ilustrasi algoritma AHC. (Koeanan, 2009)

### 2.4.1.1 Single Link

*Single link* adalah proses pembagian *cluster* yang didasarkan pada jarak terdekat antar obyeknya (*minimum distance*) (Lance, 1967). Metode ini sangat baik digunakan untuk melakukan analisis pada setiap tahap pembentukan kelompok. Ilustrasi jarak pada *single link* ditunjukkan oleh Gambar 2.6. Algoritma *single link* sebagai berikut.

1. Diasumsikan setiap data adalah satu kelompok data.
2. Menghitung jarak dua kelompok dengan menggunakan salah satu pengukuran jarak.
3. Mencari dua kelompok yang mempunyai jarak antar kelompok paling minimal dan gabungkan.
4. Kembali kelangkah 3 dan ulangi sampai mencapai kelompok data yang diinginkan.



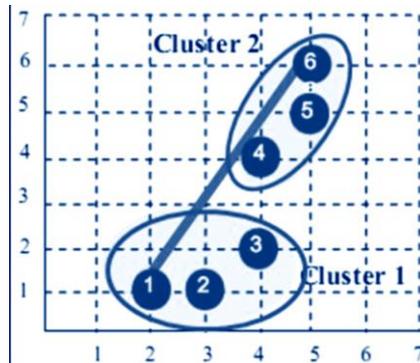
Gambar 2.6. Ilustrasi jarak pada *single link*. (Koeanan, 2009)

### 2.4.1.2 Complete Link

*Complete link* adalah proses *clustering* yang didasarkan pada jarak terjauh dari data antar kelompok (*maximum distance*) (Lance, 1967). Metode ini baik untuk kasus *clustering* dengan distribusi data normal. Akan tetapi tidak cocok untuk data yang mengandung *outlier* (pencilan). Ilustrasi jarak pada *complete link* ditunjukkan oleh Gambar 2.7. Algoritma skema *complete link* adalah sebagai berikut.

1. Diasumsikan setiap data adalah satu kelompok data.
2. Menghitung jarak dua kelompok dengan menggunakan salah satu pengukuran jarak.

3. Mencari 2 kelompok yang mempunyai jarak paling maksimal antar 2 kelompok dan gabungkan ke dalam satu kelompok baru.
4. Kembali kelangkah 3 dan ulangi sampai mencapai kelompok data yang diinginkan.

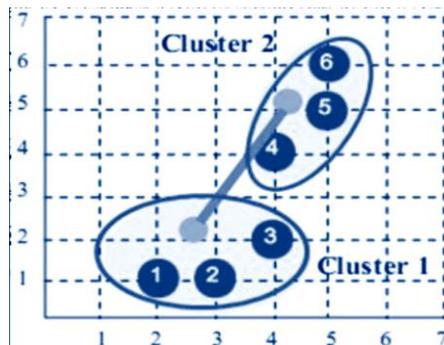


Gambar 2.7. Ilustrasi jarak pada *complete link*. (Koeanan, 2009)

#### 2.4.1.3 Centroid Link

*Centroid link* adalah proses *cluster* yang dilakukan berdasarkan jarak *centroid*-nya (Lance, 1967). Metode ini baik untuk data yang mengandung *outlier*. Ilustrasi jarak pada *centroid link* ditunjukkan oleh Gambar 2.8. Algoritma *centroid link* sebagai berikut.

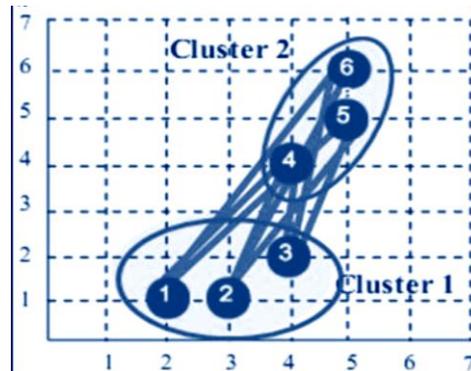
1. Diasumsikan setiap data adalah satu kelompok data.
2. Menghitung jarak dua kelompok dengan menggunakan salah satu pengukuran jarak.
3. Mencari 2 kelompok yang mempunyai jarak antar centroid paling minimal antar 2 kelompok dan gabungkan ke dalam satu kelompok baru.
4. Kembali kelangkah 3 dan ulangi sampai mencapai kelompok data yang diinginkan.



Gambar 2.8. Ilustrasi jarak pada *centroid link*. (Koeanan, 2009)

#### 2.4.1.4 Average Link

*Average link* adalah proses *clustering* yang didasarkan pada jarak rata-rata antar obyek (*average distance*) (Lance, 1967). Skema ini merupakan skema yang relatif paling baik dari skema yang lain, namun waktu komputasinya lebih tinggi. Ilustrasi jarak pada *average link* ditunjukkan oleh Gambar 2.9.



Gambar 2.9. Ilustrasi jarak pada *average link*. (Koeanan, 2009)

#### 2.6.2 Analisis K-Means Cluster

Analisis Cluster merupakan salah satu metode objek mining yang bersifat tanpa latihan (*unsupervised analysis*), sedangkan analisis K-Means *cluster* merupakan salah satu metode *cluster analysis* non-hirarki yang mempartisi data yang ada kedalam satu atau lebih *cluster* atau kelompok data berdasarkan karakteristiknya, sehingga data yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain. Tujuan pengelompokan adalah untuk meminimalkan *objective function* yang di set dalam proses *clustering*, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu *cluster* dan memaksimalkan variasi antar *cluster*.

Jika diberikan sekumpulan objek  $X = (x_1, x_2, \dots, x_n)$  maka algoritma *K-Means* akan mempartisi  $X$  dalam  $k$  buah *cluster*, setiap *cluster* memiliki *centroid* dari data-data dalam *cluster* tersebut. Pada tahap awal algoritma K-Means Cluster Analysis dipilih secara acak  $k$  buah data sebagai *centroid*, kemudian jarak antara data dengan *centroid* dihitung dengan menggunakan jarak *Eucliden*, data ditempatkan dalam *cluster* yang terdekat dihitung dari titik tengah *cluster*. *Centroid*

baru ditetapkan jika semua objek sudah ditempatkan dalam *cluster* terdekat. Proses penentuan *centroid* dan penempatan objek dalam *cluster* diulangi sampai nilai *centroid* konvergen. Secara umum metode K-Means Cluster Analysis menggunakan algoritma sebagai berikut (Agusta, 2007) :

1. Tentukan  $k$  sebagai jumlah *cluster* yang di bentuk dan pilih  $k$  data dari data set  $X$  sebagai *centroid*.
2. Pilih  $k$  *Centroid* (titik pusat *cluster*) awal secara *random*. Penentuan *centroid* awal dilakukan secara *random*/acak dari objek-objek yang tersedia sebanyak  $k$  *cluster*, kemudian untuk menghitung *centroid cluster ke-l* berikutnya, digunakan rumus sebagai berikut :

$$c_l = \frac{\sum_{i=1}^n x_i}{n}, i = 1, 2, \dots, n \quad (2.3)$$

$c_l$  adalah centroid pada cluster ke- $l$ ,  $x_i$  adalah data ke- $i$ , dan  $n$  adalah banyaknya data yang tergabung dalam sebuah cluster.

3. Hitung jarak setiap data ke masing-masing *centroid* dari masing-masing *cluster*. Pada penelitian ini, untuk menghitung jarak antara data dengan *centroid* menggunakan rumus Euclidean Distance.

$$d(x, c) = \|x - c\| = \sqrt{\sum_{i=1}^n (x_i - c_i)^2}, i = 1, 2, \dots, n \quad (2.4)$$

4. Alokasikan masing-masing data ke dalam *centroid* yang paling terdekat. Pada saat data sudah dihitung jarak terhadap setiap *centroid*, maka selanjutnya dipilih jarak yang paling kecil sebagai cluster yang akan diikuti sebagai relokasi data pada cluster di sebuah iterasi. Relokasi sebuah data dalam cluster yang diikuti dapat dinyatakan dengan nilai keanggotan  $a$  yang bernilai 1 atau 0. Nilai 0 jika tidak menjadi anggota sebuah cluster dan 1 jika menjadi anggota sebuah cluster. Karena *K-Means* mengelompokkan secara tegas data hanya pada satu *cluster*, maka nilai  $a$  sebuah data pada semua cluster, hanya satu yang bernilai 1, sedangkan lainnya 0 seperti yang dinyatakan oleh persamaan berikut:

$$a_{ij} = \begin{cases} 1 & \arg \min_j \{d(x_i, c_j)\} \\ 0 & \text{lainnya} \end{cases} \quad (2.5)$$

$d(x_i, c_j)$  menyatakan jarak dari data ke- $i$  ke cluster  $c_j$ .

5. Lakukan iterasi, kemudian tentukan posisi *centroid* baru dengan menggunakan Persamaan (2.3).
6. Ulangi langkah 3 jika posisi *centroid* baru tidak sama. Pengecekan konvergensi dilakukan dengan membandingkan matriks *group assignment* pada iterasi sebelumnya dengan matrik *group assignment* pada iterasi yang sedang berjalan. Jika hasilnya sama maka *algoritma k-means cluster analysis* sudah *konvergen*, tetapi jika berbeda maka belum konvergen sehingga perlu dilakukan iterasi berikutnya.

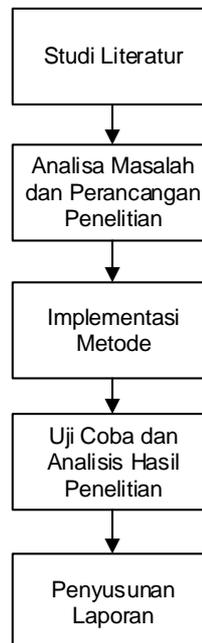
## BAB 3

### METODOLOGI PENELITIAN

Bab ini menjelaskan metode penelitian yang akan digunakan. Dalam bab ini dipaparkan tahapan penelitian, *dataset* yang akan digunakan, rancangan metode yang diusulkan dan skenario uji coba.

#### 3.1 Tahapan Penelitian

Penelitian ini memiliki beberapa tahapan yang terdiri dari: (1) studi literatur, (2) analisis masalah dan rancangan penelitian yang diusulkan, (3) implementasi metode yang diusulkan, (4) uji coba dan analisis hasil penelitian yang diusulkan, serta (5) penyusunan laporan. Alur tahapan penelitian tersebut ditunjukkan oleh Gambar 3.1.



Gambar 3. 1 Tahapan penelitian.

##### 3.1.1 Studi Literatur

Studi literatur bertujuan menggali informasi yang terdapat dalam pustaka. Pustaka dapat berupa buku atau publikasi ilmiah seperti jurnal, dokumentasi tesis,

atau prosiding. Informasi yang digali berguna sebagai literatur yang menunjang perancangan dan implementasi metode. Pada penelitian ini, literatur yang dipelajari meliputi: sel darah putih, ruang warna HSV, *cropping* citra dan pelabelan komponen, fitur geometri dan analisis *cluster* dengan metode K-means dan *hierarchical clustering analysis*.

### **3.1.2 Analisis Masalah dan Perancangan Penelitian**

Permasalahan yang akan diselesaikan pada penelitian ini yaitu segmentasi sel darah putih pada citra AML menggunakan analisis *cluster*. Analisis *cluster* yang digunakan adalah metode K-means dan *hierarchical clustering analysis*. Berdasarkan permasalahan tersebut dibangun rancangan metode. Rancangan metode yang akan dibangun meliputi *dataset* citra yang digunakan pada penelitian, proses pencarian *initial* ROI dari sel darah putih, proses pembagian citra AML menjadi citra lokal atau sub-citra, segmentasi dengan metode K-Means pada tiap sub-citra, deteksi sel yang bersentuhan pada masing-masing citra lokal dengan fitur geometri, pemisahan sel darah putih yang bersentuhan (*touching cell*) dari hasil segmentasi pada tiap sub-citra menggunakan metode *hierarchical clustering analysis*, uji coba sistem dan evaluasi akurasi hasil segmentasi dan pemisahan sel darah putih.

### **3.1.3 Implementasi Metode**

Tahapan ini bertujuan mengimplementasikan rancangan metode. Rancangan yang telah dibuat akan diimplementasikan menggunakan aplikasi Matlab versi 2013a.

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

Pada tahap ini dilakukan akan dilakukan pengujian terhadap sistem yang telah dibangun. Pengujian dilakukan sesuai dengan skenario yang telah dirancang. Analisis hasil pengujian dilakukan untuk mengetahui performa dari metode yang diusulkan.

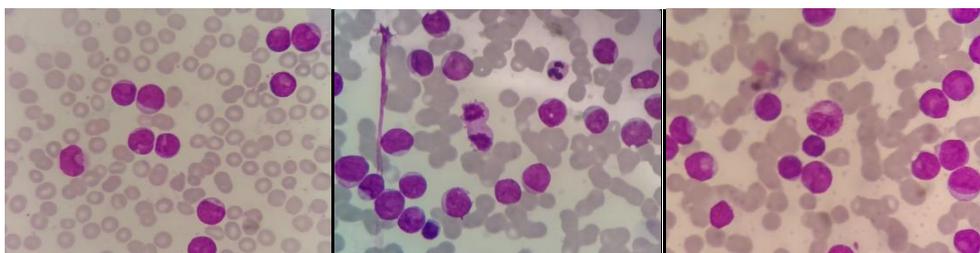
### 3.1.5 Dokumentasi

Pada tahap ini dilakukan pendokumentasian tahapan metode yang diusulkan beserta hasil implementasi dan uji coba. Dokumentasi berupa laporan penelitian.

### 3.2 Dataset Citra AML

*Dataset* citra yang digunakan dalam tesis ini adalah basis data citra *acute myeloid leukemia* (AML), yang diambil di Laboratorium Kesehatan Pemerintah Daerah Kalimantan Selatan. *Dataset* citra AML merupakan citra sampel sediaan apus darah tepi (*peripheral blood smear*) yang didalamnya terdapat sel darah putih normal dan abnormal (*blast*) dari pasien penderita AML dan yang telah mendapat izin untuk digunakan oleh ahli patologi dan pihak terkait untuk kepentingan ilmu pengetahuan dan pendidikan.

Sampel tersebut dikumpulkan oleh pakar pada pemeriksaan kesehatan di Laboratorium Kesehatan Pemerintah Daerah Kalimantan Selatan Banjarmasin, Indonesia. *Dataset* citra AML dapat digunakan untuk pengujian kemampuan algoritma segmentasi dan sistem klasifikasi karena terdiri dari citra dengan proses *staining*, resolusi, dan pencahayaan yang berbeda-beda. Gambar 3.2 merupakan sampel yang digunakan sebagai masukan pada sistem.



Gambar 3.2. Citra AML.

### 3.3 Rancangan Penelitian

Tujuan penelitian ini adalah merancang metode untuk segmentasi dan pemisahan sel darah putih yang bersentuhan pada citra AML yang memiliki variasi *staining* beragam menggunakan analisis *cluster* dengan metode K-Means dan *hierarchical clustering analysis*. Segmentasi dan pemisahan sel yang bersentuhan (*touching*) dilakukan tiap sub-citra. Sub-citra diperoleh dengan melakukan *cropping* pada citra AML. *Cropping* citra AML dilakukan berdasarkan *initial ROI*

sel darah putih yang diperoleh dengan metode *thresholding* Otsu dari citra hasil LNE. Segmentasi dilakukan dengan metode K-Means. Untuk sel darah putih yang bersentuhan, fitur geometri kebulatan (*roundness*) dan luasan (*size*) digunakan untuk proses deteksi. Dari hasil segmentasi dan deteksi tersebut pemisahan dilakukan pada sel darah putih yang bersentuhan. Rancangan metode yang akan dibangun ditunjukkan oleh Gambar 3.3. Metode segmentasi dan pemisahan sel darah putih bersentuhan (*touching*) dibagi menjadi lima tahap sebagai berikut:

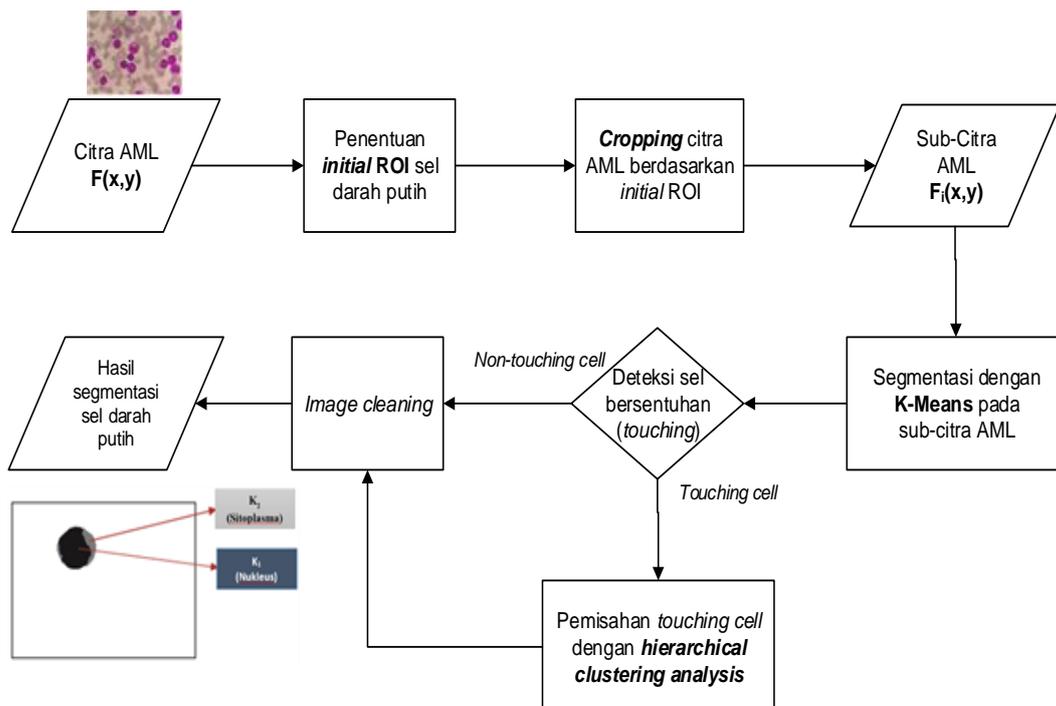
Tahap I: Penentuan *initial ROI* sel darah putih menggunakan metode *thresholding* Otsu pada citra hasil LNE.

Tahap II: *Cropping* citra AML berdasarkan *initial ROI* sel darah putih.

Tahap III: Segmentasi citra AML menggunakan metode K-Means pada sub-citra.

Tahap IV: Deteksi sel darah putih yang bersentuhan dengan fitur geometri kebulatan (*roundness*) dan luasan (*size*).

Tahap V: Pemisahan sel darah putih bersentuhan menggunakan algoritma *hierarchical clustering analysis*.



Gambar 3. 3. Rancangan metode yang diusulkan.

### 3.3.1 Penentuan *Initial ROI* Sel Darah Putih

Penentuan untuk lokasi awal ROI (*initial region of interest*) pada sel darah putih ditentukan dengan metode *thresholding* Otsu pada citra AML yang di pra-proses dengan *leukocytes nucleus enhancer* (LNE) (Huang dkk, 2012). LNE digunakan untuk memperkuat *region* nukleus yang umumnya memiliki intensitas warna piksel lebih gelap dibandingkan *region* lainnya sehingga menjadi dasar untuk perkiraan lokasi awal ROI.

Dua komponen digunakan pada pra-proses dengan LNE yaitu komponen G dari ruang warna RGB dan S pada ruang warna HSV. Tujuan pada pra-proses dengan LNE adalah untuk mengatasi warna leukosit dan eritrosit (sel darah merah) yang mirip dan efek pencahayaan. Secara umum perbedaan komponen S dan G pada sel darah merah kecil namun besar pada nukleus dari sel darah putih.

Pada LNE nilai dari komponen S dan G di normalisasi terlebih dahulu menggunakan persamaan 3.1.

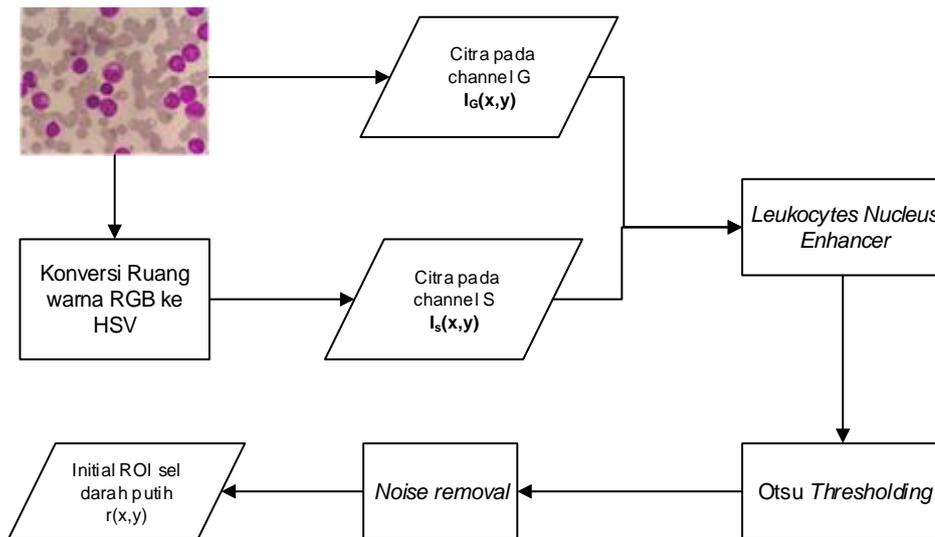
$$\begin{aligned} I_G &= (I_G - G_m) / (G_M - G_m) \\ I_S &= (I_S - S_m) / (S_M - S_m) \end{aligned} \quad (3.1)$$

dengan  $G_M$  dan  $G_m$  adalah nilai maksimum dan minimum dari komponen G ( $I_G$ ),  $S_M$  dan  $S_m$  merupakan nilai maksimum dari komponen S ( $I_S$ ). Kemudian citra LNE ( $I_E$ ) dapat dinyatakan melalui persamaan 3.2.

$$I_E = \frac{I_S}{I_G} \quad (3.2)$$

Hasil dari citra LNE merupakan citra *gray-level* dengan intensitas maksimum pada nukleus dan minimum pada *region* lainnya. Dari citra LNE untuk mendapatkan *initial ROI* dilakukan proses *thresholding* dengan metode Otsu yang membagi citra menjadi dua kelas yaitu objek dan *background* dengan nilai *threshold* yang optimal.

Pada citra biner sel darah putih hasil *thresholding* dengan metode Otsu kemungkinan terdapat sebagian kecil *region* sel darah merah yang ikut tersegmentasi sehingga perlu dilakukan dengan menghapus *region* yang memiliki ukuran lebih kecil dari rata-rata ukuran sel darah putih tunggal. Tahap penentuan *initial ROI* ditunjukkan paada Gambar 3.4.



Gambar 3.4. Proses penentuan *initial* ROI sel darah putih pada citra AML.

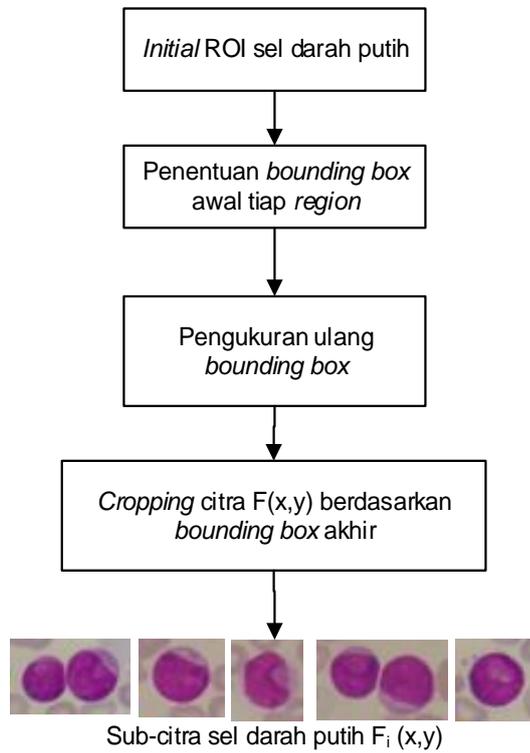
### 3.3.2 Cropping Citra AML Berdasarkan *Initial* ROI Sel Darah Putih

Setelah *initial* ROI sel darah putih disegmentasi, selanjutnya dilakukan *cropping* citra AML (Prakarsa, 2015). *Cropping* citra bertujuan membagi citra AML menjadi sub-sub citra atau citra lokal sel darah putih, sehingga segmentasi dengan metode K-Means dan pemisahan pada sel darah putih yang bersentuhan (*touching*) menggunakan *hierarchical clustering analysis* lebih mudah dan efektif karena variasi dan piksel yang di *cluster* pada sub-citra AML berukuran lebih kecil. Gambar 3.5 menunjukkan diagram proses *cropping* citra AML.

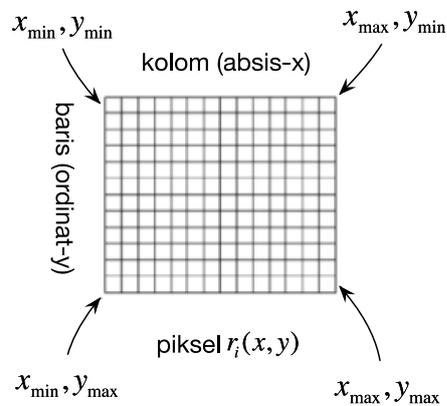
#### 3.3.2.1 Penentuan *Bounding Box* Awal

Bila citra biner *initial* ROI sel darah putih adalah  $r(x,y)$ , maka  $r_i(x,y)$  adalah komponen terhubung (*region-region*) *initial* ROI sel darah putih, dengan  $i=1, 2, \dots, n$ ;  $n$  adalah jumlah komponen terhubung *initial* ROI sel darah putih; dan  $r_i(x,y) \hat{\cap} r(x,y)$ . *Bounding box* adalah kotak imajiner yang mengelilingi  $r_i(x,y)$ . *Bounding box* awal tiap  $r_i(x,y)$  ditentukan berdasarkan koordinat spasial piksel *upper-left* (UL), *upper-right* (UR), *lower-left* (LL), dan *lower-right* (LR) dari  $r_i(x,y)$ . Ilustrasi *bounding box region initial* ROI sel darah putih ditunjukkan oleh

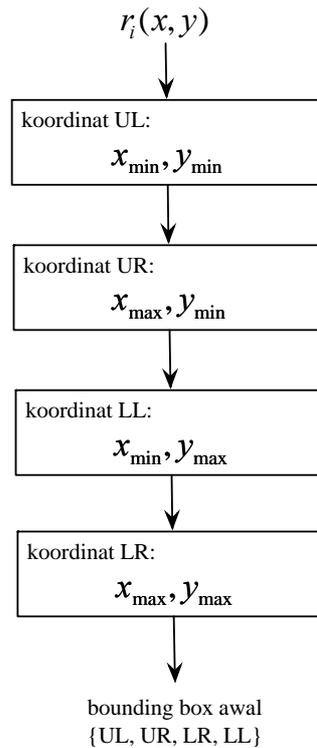
Gambar 3.6. Alur penentuan *bounding box* awal tiap  $r_i(x,y)$  ditunjukkan oleh Gambar 3.7.



Gambar 3.5. Diagram proses *cropping* citra *initial* ROI sel darah putih.



Gambar 3.6. Ilustrasi *bounding box* region *initial* ROI sel darah putih (Prakarsa, 2015).



Gambar 3.7. Alur penentuan *bounding box* awal (Prakarsa, 2015).

### 3.3.2.2 Resizing Bounding Box Awal dan Cropping Citra AML Berdasarkan Bounding Box Akhir

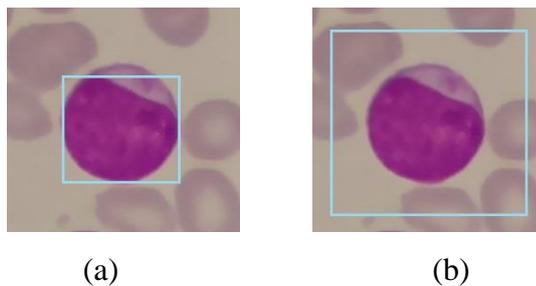
Bila *Bounding box* awal yang diperoleh tersebut hanya mengelilingi *region* nukleus tiap  $r_i(x, y)$ . Untuk memperoleh *region* membran sel darah putih, yaitu *region* nukleus dan sitoplasma, maka ukuran *bounding box* harus diatur kembali yaitu dengan memperlebar *bounding box* awal sehingga mencakup seluruh *region* sel darah putih (membran sel).

*Bounding box* awal diperlebar dengan menambahkan piksel sitoplasma. Setiap sudut *bounding box* awal (UL, UR, LL, LR) diperlebar dengan menambahkan  $k$  piksel. Nilai  $k=30$  diperoleh dari rata-rata lebar piksel sitoplasma. Misalkan ukuran citra  $F(x, y)$  adalah  $M \times N$ , dengan  $M$  adalah jumlah baris piksel dan  $N$  adalah jumlah kolom piksel, maka *bounding box* final dirumuskan ke dalam persamaan 3.3 (Prakarsa, 2015) sebagai berikut:

$$\begin{aligned}
UL_{final} &= \begin{cases} x_{\min} - k, y_{\min} - k ; & (x_{\min} - k) \geq 0, (y_{\min} - k) \geq 0 \\ x_{\min} - k, 1 ; & (y_{\min} - k) < 0, (x_{\min} - k) \geq 0 \\ 1, y_{\min} - k ; & (x_{\min} - k) < 0, (y_{\min} - k) \geq 0 \\ x_{\min}, y_{\min} ; & (x_{\min} - k) < 0, (y_{\min} - k) < 0 \end{cases} \\
UR_{final} &= \begin{cases} x_{\max} + k, y_{\min} - k ; & (x_{\max} + k) \in N, (y_{\min} - k) \geq 0 \\ N, y_{\min} - k ; & (x_{\max} + k) > N \\ x_{\max} + k, 1 ; & (y_{\min} - k) < 0 \\ x_{\min}, y_{\min} ; & (x_{\max} + k) > N, (y_{\min} - k) < 0 \end{cases} \\
LL_{final} &= \begin{cases} x_{\min} - k, y_{\max} + k ; & (x_{\min} - k) \geq 0, (y_{\max} + k) \in M \\ x_{\min} - k, M ; & (y_{\max} + k) > M \\ 1, y_{\max} + k ; & (x_{\min} - k) < 0 \\ x_{\min}, y_{\min} ; & (x_{\min} - k) < 0, (y_{\max} + k) > M \end{cases} \\
LR_{final} &= \begin{cases} x_{\max} + k, y_{\max} + k ; & (x_{\max} + k) \in N, (y_{\max} + k) \in M \\ x_{\max} + k, M ; & (y_{\max} + k) > M \\ N, y_{\max} + k ; & (x_{\max} + k) > N \\ x_{\min}, y_{\min} ; & (x_{\max} + k) > N, (y_{\max} + k) > M \end{cases}
\end{aligned} \tag{3.3}$$

bounding box final =  $\{UL_{final}, UR_{final}, LL_{final}, LR_{final}\}$

*Cropping* citra AML secara otomatis dilakukan menggunakan informasi *bounding box* final tiap  $r_i(x, y)$ . Citra AML  $F(x, y)$  yang telah dibagi menjadi sub-sub citra  $F_i(x, y)$ , dengan  $i=1, 2, \dots, N$ ;  $N$  adalah jumlah sub-citra,  $F_i(x, y) \hat{=} F(x, y)$ , kemudian sel darah putih disegmentasi dan dipisahkan menggunakan *hierarchical clustering analysis*. Gambar 3.8 menunjukkan ilustrasi *bounding box* final.



Gambar 3.8. Ilustrasi *bounding box* final.

(a) *bounding box* awal. (b) *bounding box* final setelah diperlebar ukurannya sebesar  $k$  piksel.

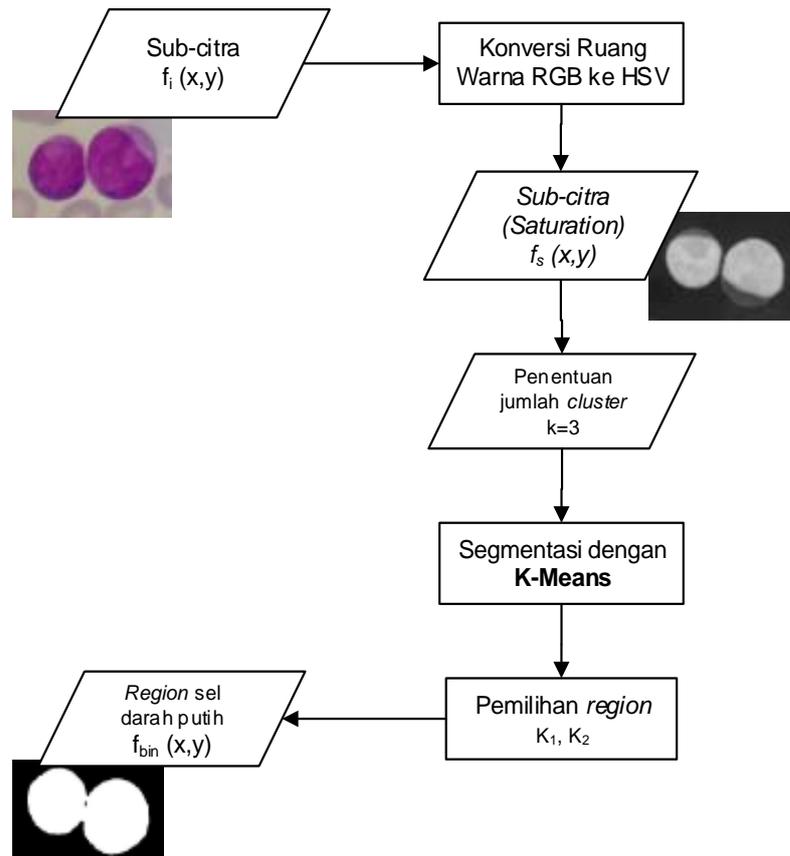
### 3.3.3 Segmentasi Sel Darah Putih Pada Sub-Citra AML

Dari sub-citra hasil *cropping* pada citra AML, segmentasi sel darah putih dilakukan dengan analisis *cluster* menggunakan metode K-means. Dataset citra AML yang digunakan pada penelitian ini memiliki variasi *staining* yang beragam sehingga analisis berbasis *cluster* yang bersifat adaptif terhadap variasi informasi piksel pada dataset citra AML lebih efektif untuk mendapatkan *region* dari sel darah putih dibanding metode lainnya seperti *thresholding* yang memerlukan nilai *threshold* yang berbeda.

Metode K-Means mengelompokkan data secara iteratif dengan melakukan partisi set data ke sejumlah K *cluster* yang telah ditentukan di awal (Prasetyo, 2014). Metode K-Means merupakan metode *clustering* yang sederhana dan relatif cepat serta efektif digunakan untuk segmentasi pada sub-citra AML dengan melakukan partisi pada informasi warna pada tiap piksel dan dapat mengatasi variasi informasi dataset yang memiliki variasi beragam. Data yang digunakan pada proses analisis *cluster* pada citra AML adalah komponen *saturation* (S) pada ruang warna HSV untuk mengatasi efek iluminasi akibat dari pencahayaan yang kurang baik.

Sub-citra AML,  $F'(x,y)$  pada analisis *cluster* menggunakan metode K-Means merupakan titik data  $X = \{x_i \mid i=1, \dots, N\}$  yang dipartisi ke dalam K *cluster*. Pada proses segmentasi, informasi warna pixel citra AML,  $F'(x,y)$ , dikelompokkan menjadi tiga *cluster* yaitu  $K_1$  merupakan *region* nukleus,  $K_2$  merupakan *region* sitoplasma dan  $K_3$  merupakan *region background* (sel darah merah dan *platelets*). Untuk mengukur ketidakmiripan (*dissimilarity*) atau jarak dalam penentuan keanggotaan data ke dalam satu *cluster* digunakan fungsi jarak *Euclidean distance* yang dituliskan pada persamaan 2.2.

Dari *cluster*  $K_1$  dan  $K_2$  yang merupakan *region* nukleus dan sitoplasma ditentukan sebagai *region* sel darah putih,  $f'_{bin}(x,y)$ . Pada sub-citra biner hasil segmentasi sel darah putih dengan metode K-Means akan menjadi dasar untuk pemisahan sel pada tahap selanjutnya. Proses segmentasi citra AML dengan metode K-Means ditunjukkan pada Gambar 3.9.

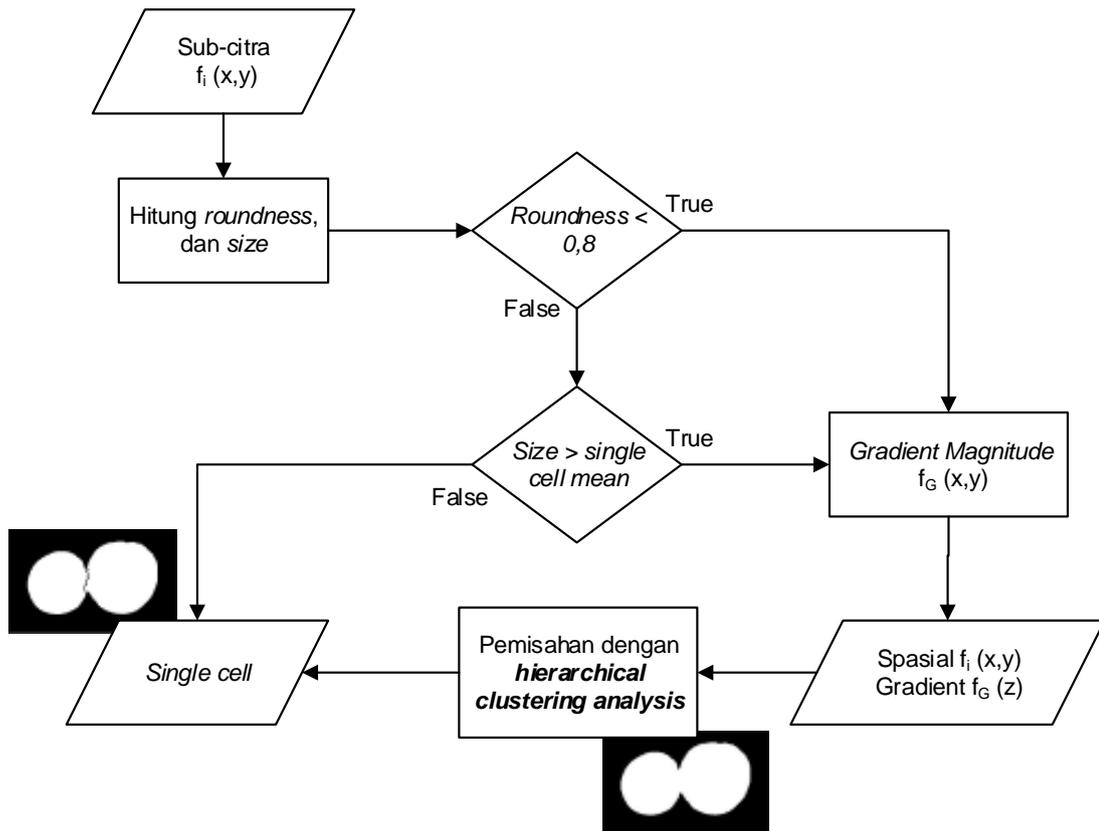


Gambar 3.9. Proses segmentasi sel darah putih pada citra AML dengan K-Means.

### 3.3.4 Pemisahan Sel Darah Putih yang Bersentuhan (*Touching Cells*) Pada Sub-Citra AML

Permasalahan pada citra leukemia yang mempengaruhi proses identifikasi adalah adanya sel darah putih yang bersentuhan (*touching cell*) atau sel yang berkumpul, oleh karena itu diperlukan pemisahan sel-sel tersebut. Adapun alur dari proses deteksi dan pemisahan pada segmentasi sel darah putih yang bertumpuk ditunjukkan pada Gambar 3.10.

Terdapat dua tahapan penting dalam segmentasi *touching cell* yaitu deteksi *region* sel darah putih yang bertumpuk dan pemisahan dengan analisis *cluster*. Kedua tahapan tersebut dilakukan pada tiap sub-citra hasil *cropping* ROI sel darah putih.



Gambar 3. 10 Proses pemisahan *touching cell* menggunakan *hierarchical clustering analysis*.

### 3.3.4.1 Deteksi pada Sel Darah Putih yang Bersentuhan

Dari setiap sub-citra hasil *cropping* pada citra AML dilakukan pendeteksian terlebih dahulu untuk menentukan sub citra dengan sel darah putih yang bersentuhan dengan sel darah putih tunggal. Untuk mendeteksi *region* yang termasuk *touching cell* digunakan fitur geometri yaitu kebulatan (*roundness*) dan luasan area (*size*). Pada analisis pada fitur geometri dengan menghitung kebulatan sel, secara umum sel darah putih tunggal memiliki bentuk lebih bundar dibanding sel yang bersentuhan. Perhitungan kebulatan dari *region* sel didapatkan dengan rumus 3.4.

$$roundness = \frac{4 \times \rho \times area}{perimeter^2} \quad (3.4)$$

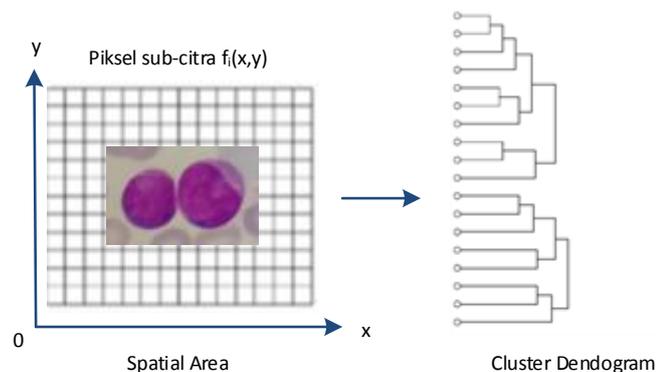
*Roundness* menyatakan *circularity* objek, yaitu rasio antara *area* dengan *perimeter*. *Roundness* bernilai satu untuk objek yang berbentuk bundar, sedangkan bernilai kurang dari satu untuk objek yang menjauhi bentuk bundar. Pada penelitian Putzu, dkk (2014) *roundness threshold* yang diusulkan sebesar 0.8.

Untuk deteksi dengan fitur *roundness* terdapat kemungkinan beberapa sel yang berkumpul mendekati bentuk bundar sehingga fitur luasan area (*size*) diperlukan untuk meningkatkan akurasi dari pendeteksian *touching cell*. Area atau jumlah piksel pada sel darah putih yang bersentuhan relatif berukuran lebih besar dibanding rata-rata ukuran sel tunggal sehingga sel dengan ukuran piksel diatas nilai *threshold* sel tunggal akan diklasifikasikan sebagai *touching cell*. *Threshold* untuk luasan area ditentukan menghitung rata-rata ukuran sel tunggal. Untuk nilai dari rata-rata luas dari sel darah putih tunggal diperoleh melalui perhitungan pada rumus 3.5.

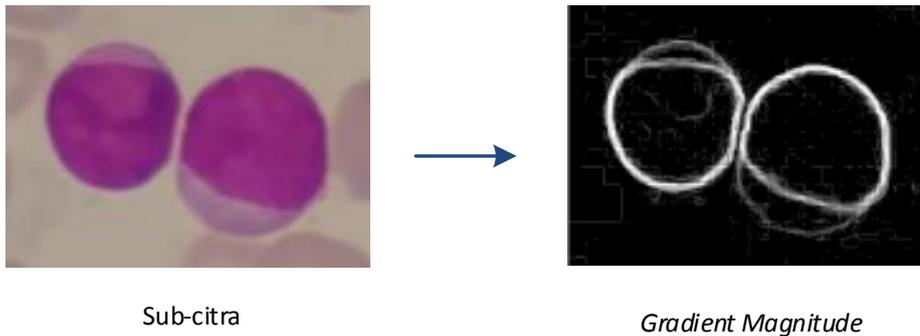
$$Single\ cell\ mean = mean(area_{wbc}) - (0,25 \cdot std(area_{wbc})) \quad (3.5)$$

### 3.3.4.2 Pemisahan Sel Darah Putih

Pada proses ini, sub-citra yang ditentukan sebagai *touching cell* pada perhitungan dengan fitur geometri sebelumnya akan dipisahkan dengan menggunakan metode *hierarchical clustering analysis*. Pemisahan sel darah putih yang bersentuhan dilakukan dengan proses klusterisasi dengan ruang informasi spasial (piksel) dan *gradient magnitude* dari citra AML yang ditunjukkan pada Gambar 3.11 dan 3.12.



Gambar 3. 11 Fitur ruang informasi spasial pada analisis *cluster*.



Gambar 3. 12 Fitur *Gradient Magnitude* pada analisis *cluster*.

Setiap koordinat piksel pada *region* sel darah putih merupakan titik yang merupakan informasi spasial dari sub-citra  $f_i(x,y)$ . Sedangkan informasi *gradient* yaitu sub-citra  $f_G(x,y)$  diperoleh melalui perhitungan dengan persamaan 2.1 dan 2.2 berfungsi untuk memberikan informasi seberapa besar (*magnitude*) perubahan intensitas yang terjadi pada sub-citra di *region* sel darah putih.

*Hierarchical clustering analysis* merupakan metode analisis *cluster* yang tidak memerlukan inisiasi jumlah kluster terlebih dahulu dan jumlah sel yang dipisahkan dapat diketahui dengan memisahkan hirarki kluster pada level tertentu serta memberikan hasil *cluster* yang lebih stabil karena tidak ada proses penentuan titik pusat *cluster* yang acak.

Metode pengelompokan pada hirarki *cluster* yang digunakan pada penelitian ini adalah aglomeratif yang memulai pengelompokan dari masing-masing data sebagai *cluster* kemudian secara rekursif mencari *cluster* terdekat sebagai pasangan untuk digabungkan sebagai satu *cluster* yang lebih besar. Kedekatan dua *cluster* pada penelitian ini dihitung dengan metode *centroid linkage* yaitu menggabung *cluster* berdasarkan jarak terdekat antara titik pusat dari dua *cluster*. Jarak dua *cluster* pada metode *centroid linkage* diukur dengan formula jarak tengah (*median proximity*) dengan persamaan 3.6.

$$d(U,V) = \| C_U - C_V \| ; d(U,V) \in D \quad (3.6)$$

dengan  $D$  menyatakan matriks kedekatan (jarak) antar *cluster*. Untuk parameter perhitungan jarak yang digunakan adalah *Euclidean distance* seperti

yang dituliskan pada persamaan 2.4. Penerapan *hierarchical clustering analysis* pada sub-citra juga mengoptimalkan proses analisis *cluster* karena pada citra dengan area piksel lebih kecil proses klusterisasi menjadi lebih cepat dibandingkan pada citra AML secara keseluruhan. Untuk kernel yang digunakan pada fitur *gradient magnitude* adalah operator *Roberts*.

### 3.3.5 Pembersihan Citra (*Image Cleaning*)

*Image cleaning* terdiri dari penghapusan semua region sel darah putih yang terdapat pada tepi citra dan semua komponen abnormal (*region* sel darah putih cacat pewarnaan), bertujuan untuk mengurangi *error* pada tahap klasifikasi. Penghapusan citra yang berada di tepi merupakan operasi sederhana, sedangkan penghapusan komponen abnormal merupakan proses yang lebih kompleks karena membutuhkan analisis geometri *region* sel darah putih.

Pertama, ukuran *area* dan *convex area* dihitung untuk masing-masing sel darah putih yang ada pada citra. Selanjutnya kombinasi *area* dan *convex area* digunakan untuk menghitung nilai *solidity* yang digunakan untuk memisahkan komponen abnormal. *Solidity* mengukur kepadatan objek. *Solidity* didefinisikan sebagai rasio antara objek *area* dengan area *convex hull* objek seperti pada rumus 3.7.

$$solidity = \frac{area}{convex\_area} \quad (3.7)$$

Nilai *solidity* 1 menandakan objek yang solid, dan nilai kurang dari 1 menandakan objek dengan *boundary irregular* (atau terdapat *holes*). Nilai ambang *solidity* yang digunakan pada penelitian ini berdasar pada metode usulan Putzu, dkk (2014) yaitu 0.9. Semua objek bernilai *solidity* di bawah nilai ambang (*threshold*) akan dihapus.

## 3.4 Perancangan Uji Coba

Pada sub bab ini menjelaskan tentang pembagian skenario uji coba dan evaluasi performa pada segmentasi dan pemisahan sel yang bersentuhan pada citra AML.

### 3.4.1 Skenario Uji Coba

Uji coba dilakukan untuk menganalisis performa dari sistem. Untuk mengetahui performa sistem, dilakukan uji coba dengan berbagai skenario sistem. Adapun beberapa skenario uji coba pada penelitian ini diantaranya:

1. Skenario pertama ujicoba perbandingan dengan segmentasi citra manual. Perbandingan dilakukan dengan citra *ground-truth* dataset citra AML yang telah divalidasi oleh ahli. Hal ini dilakukan untuk mengetahui seberapa besar segmentasi sel darah putih yang tersegmentasi dengan benar.
2. Skenario kedua ujicoba perbandingan hasil perhitungan sel darah putih dengan metode yg diusulkan dengan hasil perhitungan manual. Pengukuran berbasis region ini dilakukan untuk menguji performa metode yang diusulkan berdasarkan jumlah sel darah putih yang berhasil diidentifikasi. Dalam skenario ini, jumlah region sel darah putih, meliputi jumlah sel darah putih yang dihitung secara tepat (*correctly counted*), tidak terhitung (*not counted*) dan *over-counted* menggunakan metode yang diusulkan, dibandingkan dengan *ground truth* jumlah sel darah putih.
3. Skenario ketiga ujicoba proses segmentasi dengan analisis *cluster* pada citra global (citra AML) dengan sub-citra AML untuk mengetahui seberapa besar perbedaan akurasi segmentasi yang dihasilkan pada penerapan antara kedua jenis citra tersebut.

### 3.4.2 Evaluasi Hasil Segmentasi

Untuk hasil segmentasi metode yang diusulkan dan citra *ground truth* digunakan untuk mengklasifikasikan piksel dari ROI sel darah putih sebagai *true positive*, *false negative*, atau *false positive*, yang direpresentasikan ke dalam *Zijdenbos Similarity Index (ZSI)*, presisi, dan *recall*. ZSI didefinisikan sebagai rasio dua kali interseksi area antara hasil segmentasi dan citra *ground truth* dengan jumlah dari kedua *region* tersebut. Nilai dari ZSI dihitung menggunakan persamaan 3.8.

$$ZSI = \frac{2 \cdot |A \cap G|}{|A| + |G|} \quad (3.8)$$

dengan A merepresentasikan piksel hasil segmentasi metode yang diusulkan dan G merepresentasikan piksel citra *ground truth*. Dengan menggunakan *true positive* (TP), *false negative* (FN), dan *false positive* (FP), presisi (P) dan *recall* (R) dapat dihitung menggunakan persamaan 3.9.

$$\begin{aligned} P &= \frac{TP}{TP + FP} \\ R &= \frac{TP}{TP + FN} \end{aligned} \tag{3.9}$$

Nilai ZSI yang lebih besar dari 0,7 mengindikasikan similaritas tinggi yaitu kemiripan (*good match*) antar dua *region*. Presisi dapat merepresentasikan *over-segmentation*, sedangkan *recall* atau bisa disebut juga sensitivitas dapat merepresentasikan *under-segmentation*. Semakin tinggi nilai presisi maka kemungkinan *over-segmentation* semakin rendah, sedangkan semakin tinggi nilai *recall* maka kemungkinan *under-segmentation* juga semakin rendah. Nilai presisi yang tinggi dapat berarti *under-segmentation* jika nilai *recall*nya rendah, dan nilai *recall* yang tinggi dapat berarti *over-segmentation* jika nilai presisinya rendah.

*[Halaman ini sengaja dikosongkan]*

## **BAB 4**

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

Pada bab ini dipaparkan uji coba berdasarkan skenario yang telah dijelaskan pada sub-bab 3.4, kemudian dilakukan analisis berdasarkan hasil dari uji coba tersebut.

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

Lingkungan perangkat lunak yang digunakan dalam uji coba penelitian ini sebagai berikut:

- a. Sistem Operasi Microsoft Windows 8 64 Bit.
- b. Aplikasi MATLAB R2013a.

Lingkungan perangkat keras yang digunakan adalah komputer personal yang memiliki spesifikasi sebagai berikut:

- a. Prosesor Intel Core i3 2.7GHz.
- b. RAM 4GB.

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

Uji coba pada penelitian ini menggunakan data 36 citra AML dari basis data yang ada di Laboratorium Kesehatan Pemerintah Daerah Kalimantan Selatan. Citra uji tersebut merupakan citra sediaan apus darah tepi (*peripheral blood smear*) sel darah yang mengandung sel-sel AML (*blast*) dan sel-sel normal (*healthy cells*). Beberapa citra pada dataset terdapat sejumlah *region* sel yang berkelompok atau *touching cell*.

Hasil segmentasi dari metode yang diusulkan (analisis *cluster* dengan metode K-Means pada citra lokal) dibandingkan dengan citra yang telah disegmentasi secara manual. Citra yang disegmentasi secara manual tersebut dianggap sebagai hasil segmentasi yang benar (*ground truth*). Perhitungan manual jumlah sel darah putih pada data berdasar pada arahan yang diberikan oleh ahli/dokter patologi.

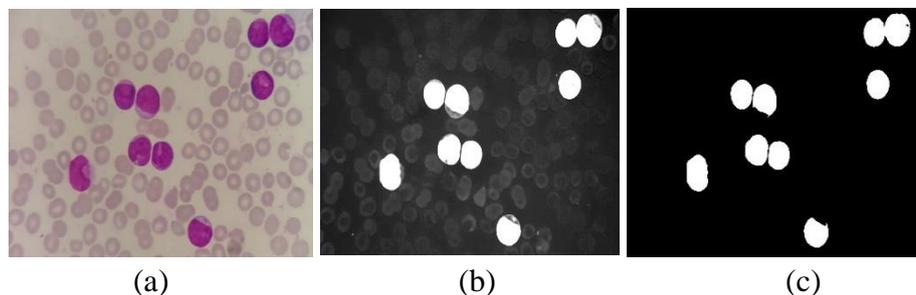
### 4.3 Hasil Uji Coba

Pada sub-bab ini dilakukan pembahasan hasil uji coba segmentasi citra AML menggunakan metode yang diusulkan berikut dengan hasil uji coba berdasarkan skenario uji cobanya.

#### 4.3.1 Hasil Uji Coba Segmentasi Citra AML

##### 4.3.1.1 Hasil Uji Coba Penentuan *Initial* ROI sel darah putih

Proses untuk mendapatkan *initial* ROI dari citra RGB sel darah merupakan kombinasi dari citra *Leukocytes Nucleus Enhancer* (LNE) dengan Otsu thresholding. ROI yang ingin didapatkan berupa citra *binner* bagian nukleus (inti sel) dari sel darah putih. Hasil uji coba pada penentuan *initial* ROI sel darah putih ditunjukkan pada Gambar 4.1.



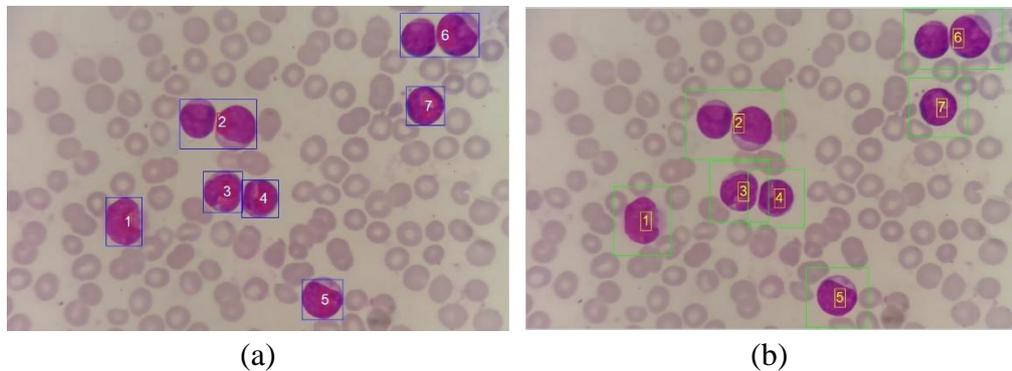
Gambar 4.1. Hasil uji coba penentuan *initial* ROI sel darah putih.

(a) citra RGB sel darah, (b) citra *grayscale* hasil proses *Leukocytes Nucleus Enhancer* (LNE), (c) citra biner *initial* ROI sel darah putih.

##### 4.3.1.2 Hasil Uji Coba *Cropping Initial* ROI Sel Darah Putih

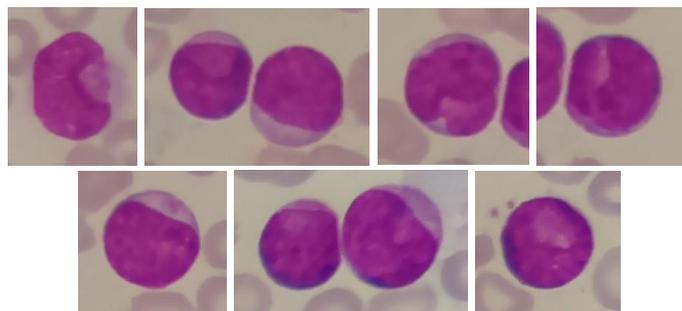
Pada tahap ini dilakukan proses untuk mendapatkan sub-citra yang merupakan ROI awal (*initial* ROI) dari sel darah putih. *Cropping* dilakukan berdasarkan *bounding box* citra *initial* ROI sel darah putih. Karena citra hasil dari LNE digunakan untuk memperkuat *region* nukleus dan memperlemah bagian lainnya sehingga *bounding box* awal citra *initial* ROI sel darah putih hanya melingkupi sebagian besar *region* nukleus (Gambar 4.2a), oleh karena itu *bounding box* awal perlu diperbesar wilayahnya agar meliputi seluruh area sel darah putih dengan menambahkan ukuran *bounding box* sebesar  $k$  piksel (Gambar 4.2b). Nilai

$k$  merupakan rata-rata lebar sitoplasma dari sel darah putih. Dalam uji coba, nilai  $k=30$  digunakan berdasarkan pengaturan yang diusulkan oleh Prakarsa (2015). Sub-citra setelah *dicrop* berdasarkan *bounding box final* (setelah *resize*) ditunjukkan pada Gambar 4.3.



Gambar 4.2. *Bounding box initial* ROI sel darah putih.

(a) lokasi awal *initial* ROI sel darah putih berdasarkan *bounding box* citra *initial* sel darah putih, dan (b) lokasi *initial* ROI sel darah putih setelah dilakukan *resize* pada *bounding box* citra *initial* sel darah putih.



Gambar 4.3. Hasil sub-citra setelah proses *cropping* citra AML.

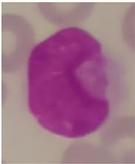
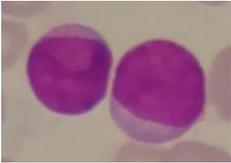
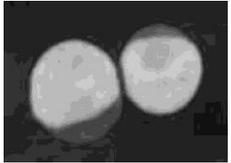
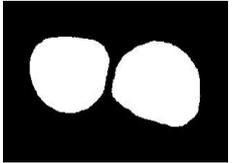
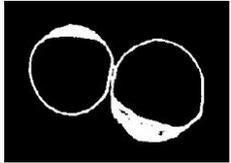
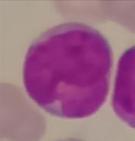
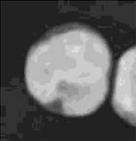
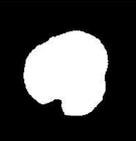
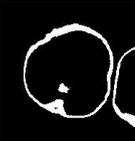
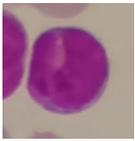
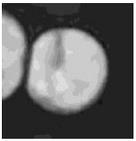
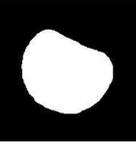
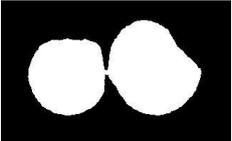
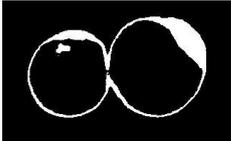
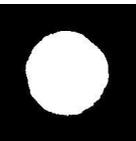
#### 4.3.1.3 Hasil Uji Coba Segmentasi Metode K-Means pada Citra Lokal

Segmentasi dilakukan dengan analisis *cluster* menggunakan metode K-Means pada setiap sub-citra AML. Data dipartisi ke dalam  $k$  *cluster*, dimana nilai  $k=3$  digunakan pada penelitian ini untuk membagi *region* dari sub-citra masing-masing menjadi *region* nukleus, sitoplasma dan *background* (sel darah merah dan

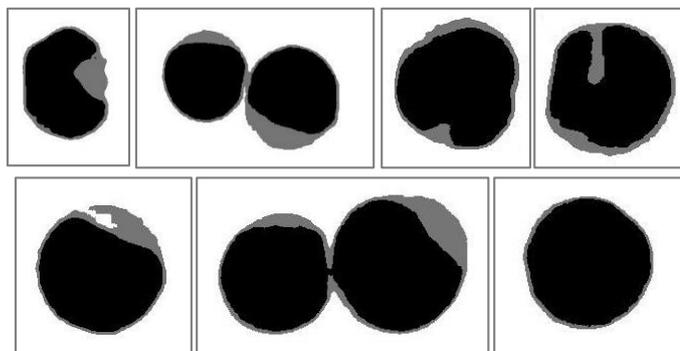
*platelets*). Data yang dipartisi adalah nilai *saturation* (S) dari sub-citra dalam ruang warna HSV.

Untuk menghindari *local optima*, proses pembagian *cluster* diulang (*replicates*) sejumlah  $r$  menggunakan posisi titik pusat *cluster* yang baru untuk mendapatkan hasil *cluster* yang *local minima* dengan memilih titik pusat (*centroid*) yang memiliki jarak total terkecil terhadap semua data pada setiap *cluster*. Pada uji coba digunakan nilai  $r=5$ .

Tabel 4.1. Proses segmentasi pada sub-citra AML.

Sub-citra RGB	Sub-citra S (HSV)	Nukleus	Sitoplasma
			
			
			
			
			
			
			

Input sub-citra RGB, sub-citra dari komponen S dari ruang warna HSV dan hasil *cluster* dari masing-masing *region* sel darah putih yaitu nukleus dan sitoplasma ditunjukkan pada Tabel 4.1. Hasil segmentasi sel darah putih yang merupakan penggabungan *cluster* dari *region* nukleus dan sitoplasma diperlihatkan pada Gambar 4.4.



Gambar 4.4. Hasil segmentasi sel darah putih dengan penggabungan *cluster* dari *region* nukleus dan sitoplasma.

Pada Gambar 4.4, *region* dari sitoplasma ditunjukkan oleh warna keabuan, sedangkan *region* nukleus ditunjukkan oleh warna hitam.

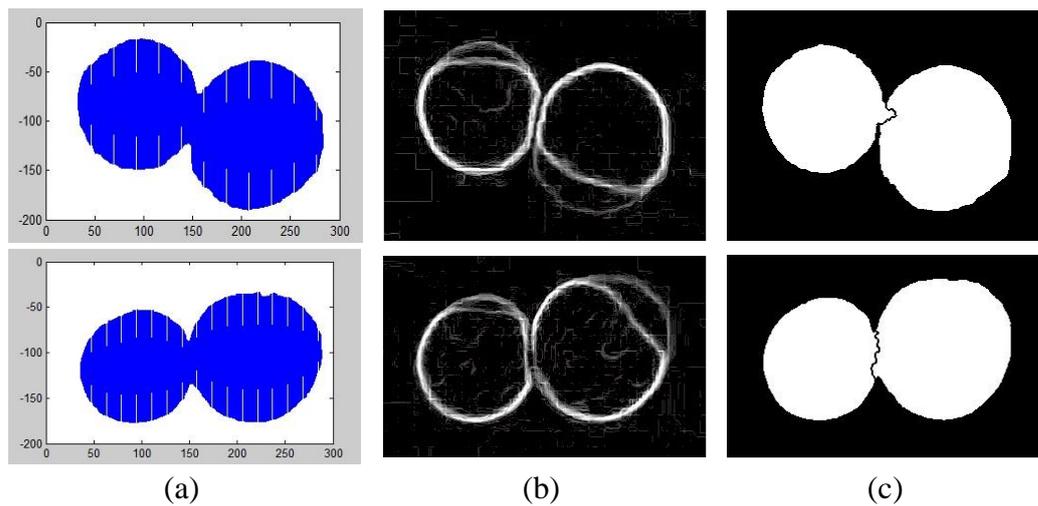
#### 4.3.1.4 Hasil Uji Coba Identifikasi dan Pemisahan *Touching Cell* dengan *Hierarchical Clustering Analysis (HCA)*

Proses identifikasi pada *touching cell* ditentukan menggunakan analisis fitur geometri dari sel darah putih pada sub-citra AML, yaitu fitur *roundness* dan rata-rata luasan dari sel tunggal. Untuk fitur *roundness* berdasarkan penelitian Putzu, dkk (2014), nilai ambang (*threshold*) yang digunakan adalah 0.8. *Region* pada citra yang memiliki nilai *roundness* kurang dari *threshold* diidentifikasi sebagai *touching cell*. Fitur kedua adalah luasan (*size*) sel tunggal pada citra yang dihitung dari rata-rata luasan setiap sel tunggal yang ada pada citra AML. *Region* dengan nilai luasan diatas nilai *threshold* tersebut juga diidentifikasi sebagai *region* dari *touching cell*.

Untuk pemisahan dari *touching cell* menjadi sel-sel tunggal digunakan metode *hierarchical clustering analysis*. Jumlah dari *cluster c* dihitung berdasarkan

nilai pembagian luas *region* dari *touching cell* dengan nilai luasan dari sel tunggal. Berdasarkan uji empiris metode *linkage* yang memberikan *cluster* terbaik adalah *centroid linkage*. Perhitungan jarak yang digunakan untuk mengukur kedekatan *cluster* (*similarity*) adalah transformasi jarak *Euclidean*. Untuk fitur yang digunakan untuk pembagian *cluster* adalah ruang informasi spasial  $f'(x,y)$  dari objek dan *gradient magnitude*  $g$  dari *touching cell*.

Hasil uji coba identifikasi dan pemisahan *touching cell* pada sub-citra beserta *gradient magnitude* ditunjukkan pada Gambar 4.5.



Gambar 4.5. Hasil pemisahan dengan *hierarchical clustering analysis*.

(a) Fitur Spasial, (b) *Gradient magnitude* dan (c) Hasil pemisahan sel

#### 4.3.1.5 Penggabungan Sub-Citra dan Pembersihan Citra (*Image Cleaning*)

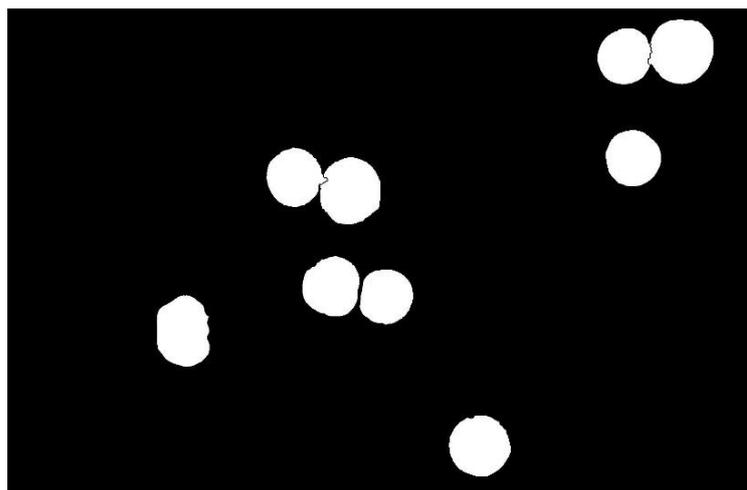
Sub-citra hasil proses segmentasi dengan *region touching cell* yang telah dipisahkan menjadi sel-sel tunggal digabungkan kembali menjadi satu citra yang memiliki dimensi yang sama dengan citra awal seperti yang ditunjukkan pada Gambar 4.6.

Pada dataset citra AML terdapat beberapa *region* yang cacat akibat proses pewarnaan sediaan apus (preparat) darah tepi, *noise* yang tersegmentasi karena memiliki kedekatan nilai *saturation* dengan sel darah putih dengan luas *region* yang kecil ataupun sel darah putih yang tidak sempurna karena terpotong di tepi citra

sehingga terdapat *region* yang hilang, oleh sebab itu diperlukan pembersihan pada *region* yang merupakan *non-object* tersebut.

Proses pembersihan citra (*image cleaning*) melakukan identifikasi kemudian menghilangkan *region non-object* dari citra. Identifikasi pada *region* yang cacat pewarnaan dilakukan berdasarkan nilai *solidity*. Nilai ambang *solidity* yang digunakan untuk membedakan *region* cacat atau normal adalah 0,9 sehingga *region* dengan nilai *solidity* di bawah nilai ambang tersebut dikelompokkan sebagai *region* cacat pewarnaan.

Untuk *region* yang merupakan *noise* diidentifikasi dengan perbandingan nilai luasan *region* dengan rata-rata luasan sel tunggal. *Region* dengan nilai luasan kurang dari setengah dari rata-rata luas sel darah putih tunggal dikelompokkan sebagai *noise* pada citra. Pada sel darah putih yang terdapat di tepi, semua *region* yang terhubung langsung dengan tepi dari citra akan dihilangkan. Hasil dari *image cleaning* pada beberapa citra yang memiliki *region non-object* ditunjukkan Pada tabel 4.2.

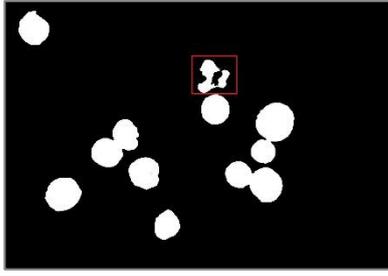
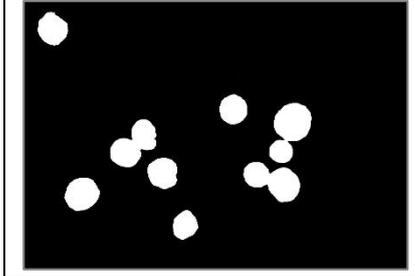
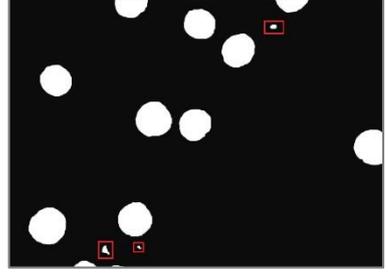
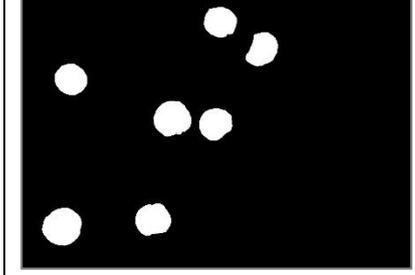
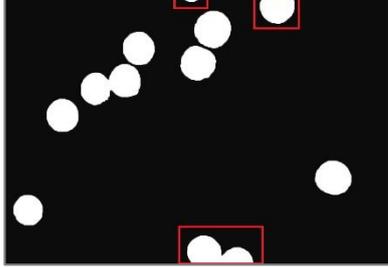
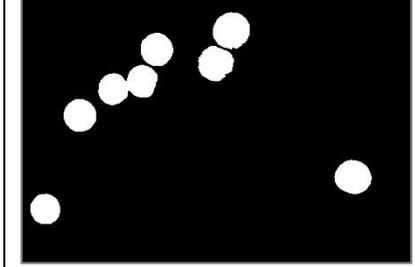


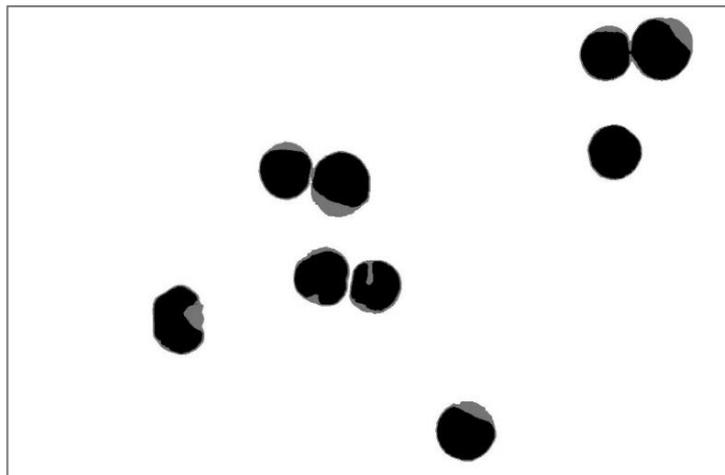
Gambar 4.6. Hasil penggabungan sub-citra kembali ke dimensi awal.

#### 4.3.1.6 Hasil Akhir Segmentasi Citra AML

Hasil akhir dari segmentasi adalah *region* yang merupakan *object*/sel darah putih. Hasil dari segmentasi *region* nukleus merupakan *region* yang berwarna hitam dan *region* sitoplasma yang berwarna keabuan seperti yang ditunjukkan Gambar 4.7.

Tabel 4.2 Hasil proses *image cleaning* pada citra AML.

<p>Pembersihan pada <i>non-object</i> karena cacat pewarnaan</p>		
<p>Pembersihan pada <i>non-object</i> yang merupakan <i>noise</i></p>		
<p>Pembersihan pada <i>non-object</i> di tepi citra</p>		



Gambar 4.7. Hasil akhir segmentasi sel darah putih citra AML.

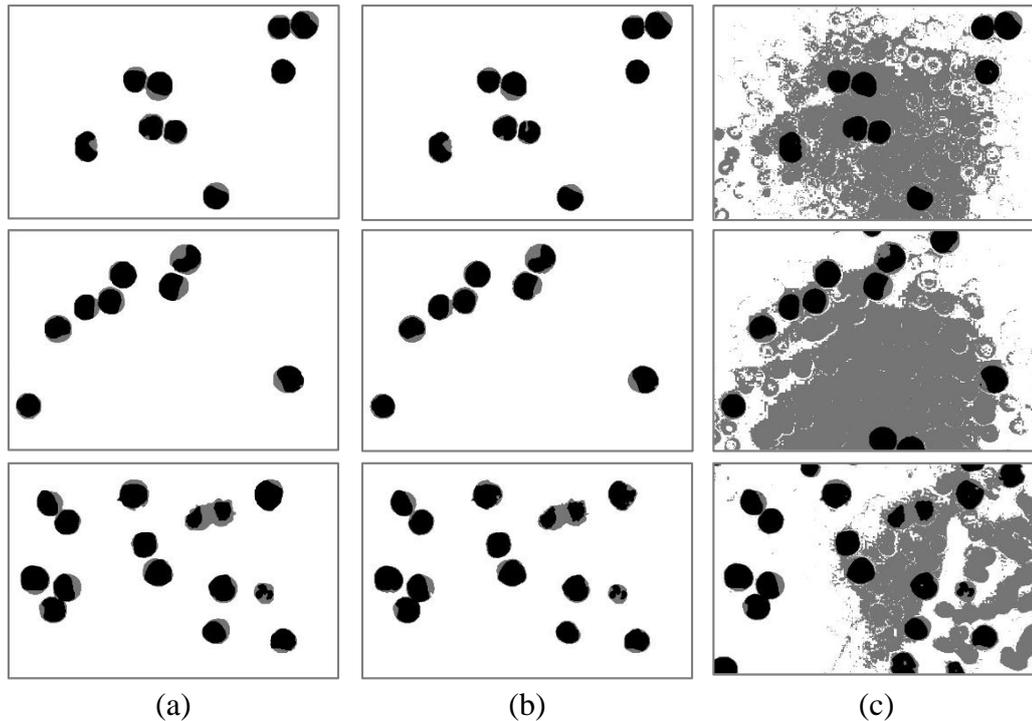
### 4.3.2 Skenario Uji Coba Segmentasi Citra AML

Pada tahap uji coba segmentasi terhadap *output* dari metode yang diusulkan, terdapat dua skenario uji coba yang dilakukan yaitu perbandingan hasil evaluasi

segmentasi antara metode yang diusulkan (LKM) dengan analisis *cluster* menggunakan metode K-Means pada citra global (GKM) terhadap *region* sel darah putih dan sitoplasma.

Untuk mengevaluasi hasil segmentasi citra AML digunakan tiga proses perhitungan yaitu presisi, *recall*, dan *Zijdenbos similarity index* (ZSI). Citra *ground truth* segmentasi manual dari citra AML dibandingkan dengan hasil segmentasi untuk mendapatkan nilai presisi, *recall*, dan ZSI.

LKM dan GKM melakukan analisis *cluster* terhadap nilai *saturation* (S) pada ruang warna *Hue Saturation Value* (HSV). Pada GKM, analisis *cluster* dilakukan pada nilai *saturation* dari keseluruhan citra. Pembagian jumlah *cluster* yang dilakukan pada GKM sama seperti LKM, yaitu hasil analisis *cluster* dibagi menjadi *region* nukleus, sitoplasma dan *background*. Gambar 4.8 menunjukkan perbandingan antara hasil segmentasi metode LKM dengan GKM. *Region* dari nukleus ditunjukkan dengan warna hitam, sedangkan sitoplasma dengan warna keabuan.



Gambar 4.8. Perbandingan hasil segmentasi citra AML dengan citra *ground-truth*.

(a) *Gradient magnitude*, (b) *Local K-Means* dan (c) *Global K-Means*

#### 4.3.2.1 Evaluasi Hasil Segmentasi Sitoplasma antara LKM dan GKM

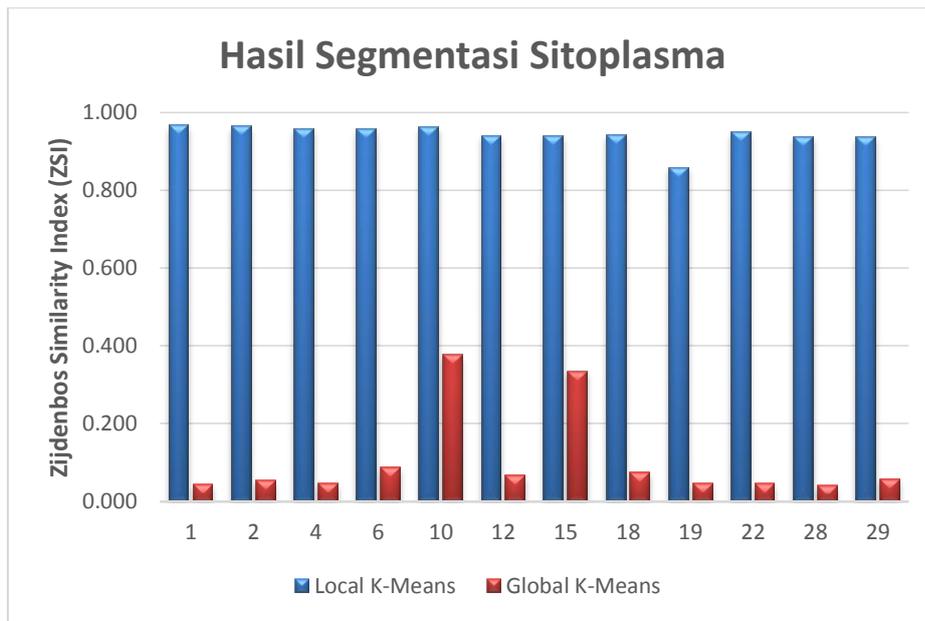
Pada tabel 4.3 ditunjukkan evaluasi hasil segmentasi pada *region* sitoplasma sel darah putih menggunakan LKM dan GKM sedangkan hasil perhitungan rata-rata nilai presisi (P), *recall* (R) dan *Zijdenbos similarity index* (ZSI) pada 40 citra uji dataset AML ditunjukkan oleh tabel 4.4. Untuk perbandingan hasil evaluasi segmentasi sel darah putih antara metode LKM dan GKM ditunjukkan oleh grafik yang ada pada Gambar 4.9.

Tabel 4.3. Evaluasi hasil segmentasi sitoplasma.

Citra	LKM			GKM		
	P	R	ZSI	P	R	ZSI
aml_01	0.997	0.938	0.967	0.023	0.675	0.045
aml_02	0.992	0.927	0.958	0.029	0.827	0.055
aml_03	0.998	0.935	0.966	0.018	0.739	0.036
aml_04	0.999	0.888	0.940	0.024	0.858	0.047
aml_05	0.998	0.890	0.941	0.033	0.774	0.064
aml_06	0.989	0.755	0.856	0.046	0.841	0.088
aml_07	0.994	0.909	0.949	0.451	0.721	0.555
aml_08	0.997	0.908	0.950	0.085	0.809	0.154
aml_09	1.000	0.855	0.922	0.723	0.501	0.592
aml_10	0.996	0.769	0.868	0.327	0.446	0.377
aml_12	0.959	0.779	0.860	0.036	0.710	0.068
aml_13	0.934	0.876	0.904	0.543	0.500	0.521
aml_14	0.905	0.863	0.883	0.313	0.451	0.370
aml_15	0.926	0.892	0.909	0.273	0.433	0.335
aml_16	0.930	0.856	0.891	0.375	0.387	0.381
aml_18	0.993	0.932	0.962	0.039	0.774	0.074
aml_19	0.999	0.917	0.956	0.025	0.652	0.047
aml_22	0.987	0.902	0.943	0.024	0.852	0.046
aml_28	0.996	0.934	0.964	0.021	0.705	0.040
aml_29	0.997	0.919	0.957	0.031	0.726	0.059

Tabel 4.4. Nilai rata-rata hasil segmentasi sitoplasma.

	Presisi	Recall	ZSI
LKM	0.977	0.885	0.926
GKM	0.162	0.673	0.189



Gambar 4.9. Evaluasi hasil segmentasi sel darah putih berdasarkan ZSI.

#### 4.3.2.2 Evaluasi Hasil Segmentasi Nukleus (Inti Sel) antara LKM dan GKM

Untuk *region* nukleus (inti sel), evaluasi hasil segmentasi antara LKM dan GKM ditunjukkan oleh Tabel 4.5 dan hasil perhitungan rata-rata nilai presisi (P), *recall* (R) dan *Zijdenbos similarity index* (ZSI) ditunjukkan pada Tabel 4.6. Untuk perbandingan dari hasil evaluasi segmentasi pada *region* nukleus diperlihatkan oleh grafik pada Gambar 4.10.

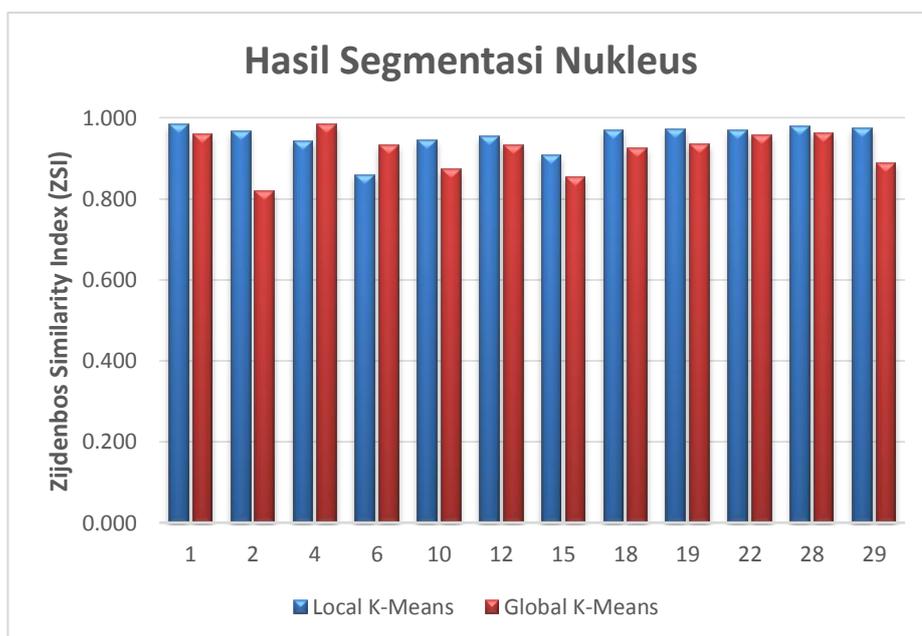
Tabel 4.5. Evaluasi hasil segmentasi *region* nukleus.

Citra	LKM			GKM		
	P	R	ZSI	P	R	ZSI
aml_01	0.991	0.976	0.983	0.924	0.999	0.960
aml_02	0.996	0.940	0.967	0.703	0.982	0.819
aml_03	0.998	0.924	0.960	0.805	0.995	0.890
aml_04	1.000	0.892	0.943	0.988	0.979	0.983
aml_05	0.999	0.856	0.922	0.749	0.964	0.843
aml_06	0.997	0.753	0.858	0.875	0.999	0.933
aml_07	0.996	0.902	0.946	0.789	0.914	0.847
aml_08	0.966	0.949	0.957	0.734	0.988	0.842
aml_09	0.998	0.939	0.968	0.995	0.947	0.971
aml_10	0.940	0.952	0.946	0.836	0.914	0.873
aml_12	0.971	0.937	0.954	0.955	0.912	0.933

Citra	LKM			GKM		
	P	R	ZSI	P	R	ZSI
aml_13	0.918	0.932	0.925	0.895	0.910	0.902
aml_14	0.819	0.932	0.872	0.613	0.912	0.734
aml_15	0.937	0.882	0.909	0.859	0.851	0.855
aml_16	0.921	0.890	0.905	0.840	0.832	0.836
aml_18	0.987	0.955	0.971	0.866	0.995	0.926
aml_19	0.996	0.951	0.973	0.887	0.991	0.936
aml_22	0.995	0.946	0.970	0.922	0.997	0.958
aml_28	0.992	0.969	0.980	0.927	1.000	0.962
aml_29	0.994	0.955	0.974	0.800	0.999	0.888

Tabel 4.6. Nilai rata-rata hasil segmentasi nukleus.

	Presisi	Recall	ZSI
LKM	0.975	0.924	0.948
GKM	0.833	0.958	0.887



Gambar 4.10. Evaluasi hasil segmentasi nukleus berdasarkan ZSI.

### 4.3.3 Skenario Uji Coba Pemisahan Sel Darah Putih pada Citra AML

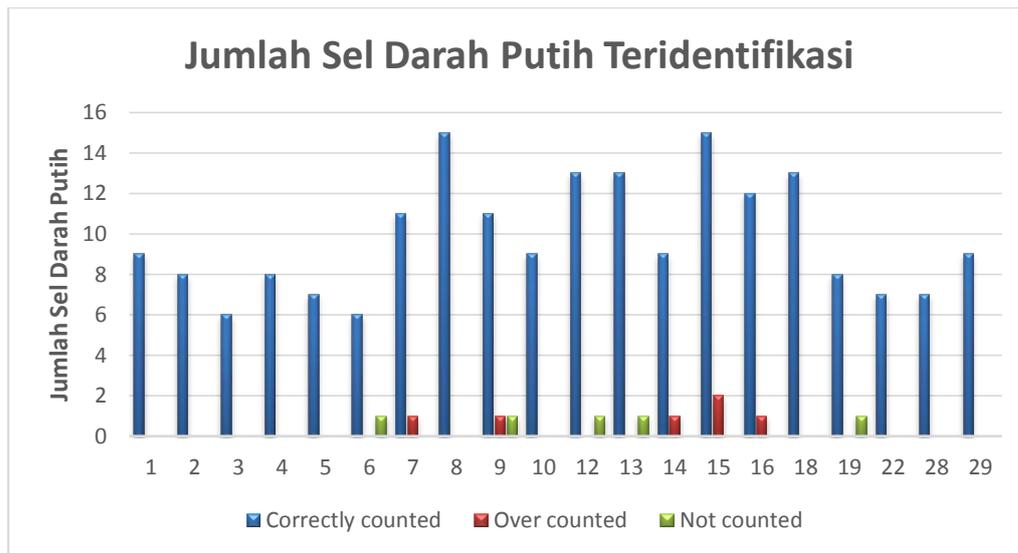
Pada skenario uji coba untuk mengevaluasi hasil pemisahan sel dilakukan perhitungan total terhadap jumlah sel darah putih tunggal yang ada pada citra,

termasuk *region touching cell* yang telah dipisahkan menggunakan metode yang diusulkan (*Hierarchical Clustering Analysis*). Hasil perhitungan jumlah sel dari metode yang diusulkan dibandingkan dengan perhitungan berdasarkan data *ground truth* yang ditunjukkan oleh Tabel 4.7.

Selain jumlah keseluruhan, perhitungan juga dilakukan berdasarkan beberapa kriteria yaitu jumlah *region* sel darah putih yang diidentifikasi benar yaitu *correctly counted*, jumlah *region* yang seharusnya tidak disertakan yaitu *over counted*, dan jumlah *region* yang merupakan objek tetapi tidak terhitung yaitu *not counted*. Hasil perhitungan dari Tabel 4.7 juga ditunjukkan dalam bentuk grafik oleh Gambar 4.11.

Tabel 4.7. Perbandingan antara jumlah sel darah putih dengan perhitungan manual dan yang teridentifikasi menggunakan metode yang diusulkan.

Citra	Ground truth	Metode yang diusulkan (HCA)			
		Counted	Correctly counted	Over counted	Not counted
aml_01	9	9	0	0	0
aml_02	8	8	0	0	0
aml_03	6	6	0	0	0
aml_04	8	8	0	0	0
aml_05	7	7	0	0	0
aml_06	7	6	6	0	1
aml_07	11	11	0	0	0
aml_08	15	15	0	0	0
aml_09	11	11	10	1	1
aml_10	9	9	9	0	0
aml_12	14	13	13	0	1
aml_13	14	13	13	0	1
aml_14	8	9	8	1	0
aml_15	13	15	13	2	0
aml_16	11	12	0	1	0
aml_18	13	13	0	0	0
aml_19	8	8	0	0	0
aml_22	7	7	0	0	0
aml_28	7	7	0	0	0
aml_29	9	9	0	0	0



Gambar 4.11. Jumlah sel darah putih teridentifikasi.

#### 4.4 Analisis Hasil Uji Coba

Pada hasil uji coba yang ditunjukkan oleh Tabel 4.3, hasil segmentasi sitoplasma sel darah putih menggunakan metode yang diusulkan memiliki nilai rata-rata presisi lebih tinggi dibandingkan dengan hasil segmentasi sitoplasma sel darah putih menggunakan analisis *cluster* dengan metode K-Means pada citra global (GKM), yaitu dengan nilai sebesar 0.977 sedangkan GKM hanya memiliki nilai rata-rata presisi sebesar 0.162.

Begitu juga dengan nilai rata-rata perhitungan *Zijdenbos Similarity Index* (ZSI), metode yang diusulkan pada penelitian ini memiliki nilai rata-rata yang lebih baik dibandingkan dengan GKM yaitu sebesar 0.926, lebih tinggi dibandingkan rata-rata nilai ZSI dari GKM yang hanya 0.189. Namun pada hasil perhitungan nilai rata-rata *recall*, hasil segmentasi dengan GKM memiliki nilai yang tidak jauh berbeda yaitu sebesar 0.673, sementara metode yang diusulkan memiliki nilai rata-rata 0.885.

Pada hasil segmentasi di *region* nukleus (inti sel) dari sel darah putih yang ditunjukkan oleh Tabel 4.5, metode yang diusulkan juga memiliki nilai rata-rata yang lebih baik pada hasil perhitungan presisi dan ZSI. Nilai rata-rata presisi dan ZSI masing-masing adalah 0.975 dan 0.948. Sedangkan untuk nilai rata-rata *recall*, hasil segmentasi oleh GKM memiliki nilai rata-rata yang sedikit lebih baik yaitu 0.958,

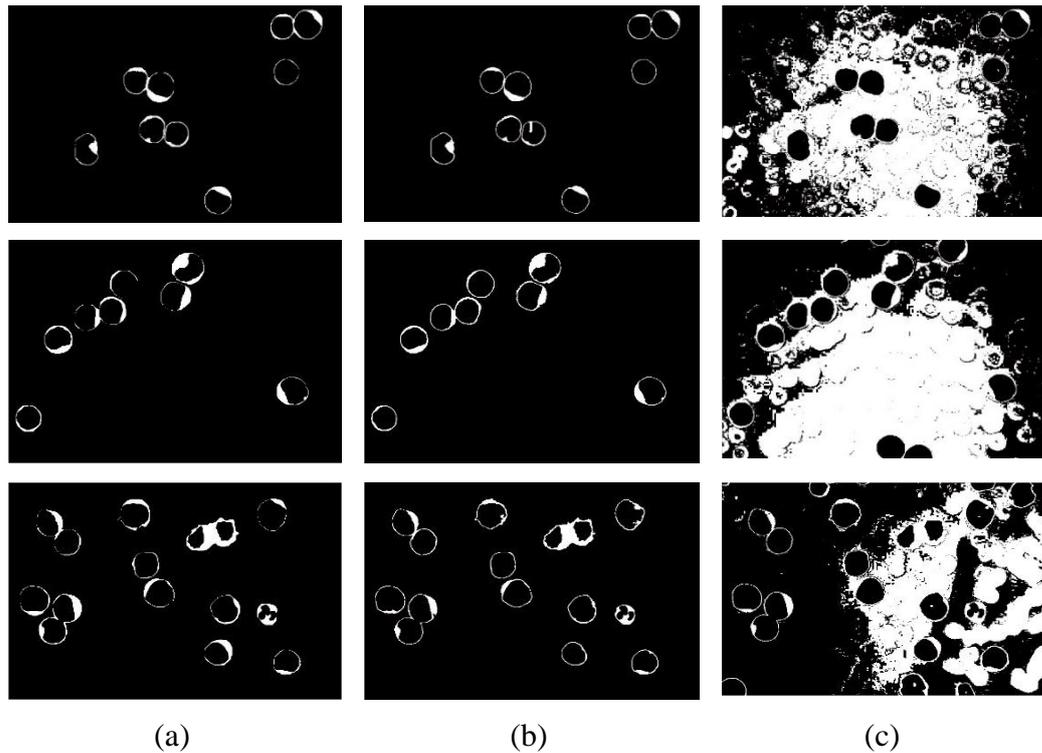
dibanding dengan metode yang diusulkan yang memiliki nilai rata-rata *recall* sebesar 0.924.

Secara umum hasil segmentasi pada metode yang diusulkan maupun GKM sama-sama memiliki hasil yang lebih baik pada *region* nukleus baik dari rata-rata nilai presisi, *recall* ataupun ZSI dibandingkan dengan hasil pada *region* sel darah putih. Namun di sisi lain perbedaan nilai rata-rata yang signifikan terdapat pada hasil perhitungan presisi dan ZSI di *region* sitoplasma sel darah putih dimana metode yang diusulkan memiliki nilai yang jauh lebih tinggi dengan perbedaan rata-rata sekitar 0.812 untuk presisi dan 0.737 untuk ZSI.

Pada rata-rata nilai *recall* dimana hasil segmentasi GKM memiliki nilai yang lebih baik pada *region* nukleus dibanding metode yang diusulkan, namun perbedaannya tidak terlalu signifikan dengan nilai perbedaan sekitar 0.033. Sedangkan pada *region* sitoplasma sel darah putih metode yang diusulkan memiliki nilai *recall* yang lebih baik dibanding GKM yaitu dengan perbedaan rata-rata sebesar 0.212. Dari hasil perhitungan rata-rata nilai *recall* juga dapat diketahui bahwa pada sebagian besar data citra AML yang disegmentasi memiliki nilai *false negative* yang cukup rendah terutama pada *region* nukleus, sehingga terjadinya *under-segmentation* juga kecil.

Akan tetapi pada untuk *region* sel darah putih rasio terjadinya *over-segmentation* cukup tinggi pada hasil segmentasi GKM karena nilai presisi yang rendah yaitu sebesar 0.162. Rasio *false positive* yang tinggi pada GKM disebabkan pada *cluster* sitoplasma terdapat banyak sel darah merah ataupun *background* yang ikut tersegmentasi, terutama pada citra yang memiliki rasio perbandingan piksel sel darah putih yang jumlahnya jauh dibawah piksel *region* sel darah merah dan *background*.

Dengan banyaknya *region* sel darah merah dan *background* yang ikut tersegmentasi sebagai objek juga menyebabkan nilai rata-rata ZSI pada hasil segmentasi sel darah putih menjadi sangat rendah pada GKM. Pada Gambar 4.12 terdapat beberapa perbandingan hasil segmentasi antara metode yang diusulkan dengan GKM yang menunjukkan terjadinya *over-segmentation* pada *region* sitoplasma dari sel darah putih.



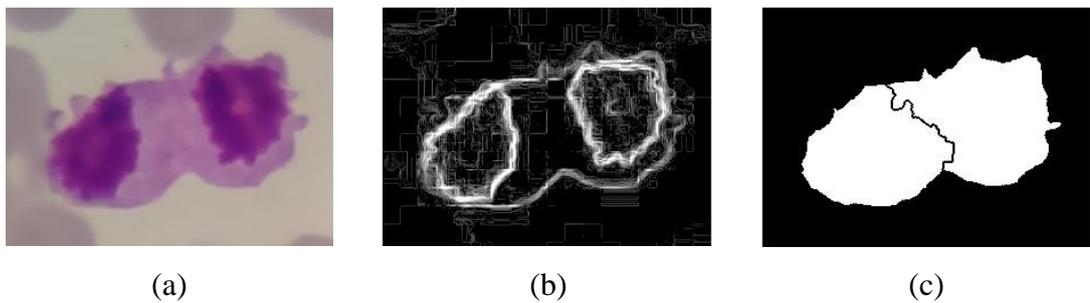
Gambar 4.12. Perbandingan hasil segmentasi pada *region* sitoplasma.

(a) *Gradient magnitude*, (b) *Local K-Means* dan (c) *Global K-Means*

Pada ujicoba identifikasi *touching cell*, secara umum fitur geometri yang digunakan yaitu fitur *roundness* dan rata-rata luasan sel tunggal dari sel darah putih telah berhasil menemukan *region* yang merupakan *touching cell* dan juga menentukan jumlah *cluster* yang akan digunakan pada proses pemisahan sel darah putih dengan HCA. Untuk pemisahan pada *touching cell* pada citra menjadi sel-sel tunggal, penggunaan fitur ruang spasial dan *gradient magnitude* dapat memisahkan sel-sel tersebut dengan identifikasi yang tepat untuk penentuan jumlah *cluster*. Fitur ruang spasial  $f'(x,y)$  merupakan fitur yang memberikan informasi letak koordinat dari sel darah putih dan fitur *gradient magnitude* memberikan informasi letak dari bagian tepi dari setiap sel tunggal. Antara *region* nukleus dan sitoplasma terdapat perbedaan intensitas yang cukup besar sehingga berguna sebagai batas pemisah.

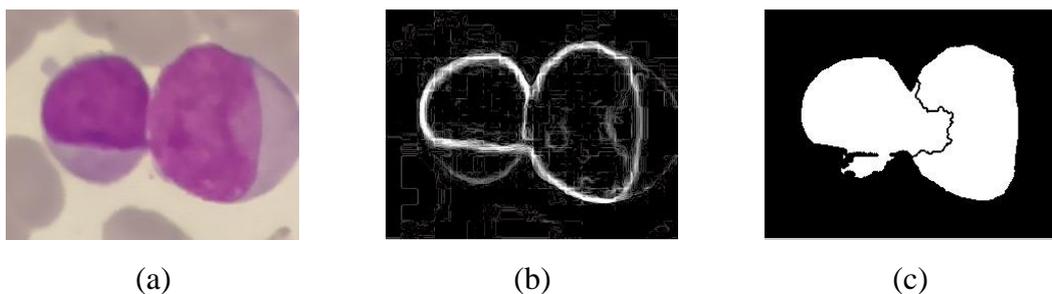
Namun karena pada fitur *gradient magnitude* secara umum *region* sitoplasma akan menjadi *region* yang lebih terang (karena perubahan intensitas dari *region* nukleus ke sitoplasma) yang kemudian menjadi pembatas untuk

memisahkan sel, pada beberapa sub-citra terdapat sel darah putih yang *region* sitoplasmanya memiliki rasio yang cukup besar atau hampir sebanding dengan nukleus sehingga perubahan intensitas *gradient* letaknya menjadi jauh dari *region* yang ideal sebagai *region* pemisahan sel. Gambar 4.13 menunjukkan sub-citra dengan sel darah putih yang memiliki luas *region* sitoplasma yang hampir sebanding dengan nukleus.



Gambar 4.13. Hasil pemisahan sel pada citra aml\_31 di sub-citra 7.  
(a) Sub-citra, (b) Gradient magnitude dan (c) hasil pemisahan sel

Selain kemungkinan luas *region* sitoplasma yang cukup besar sehingga menyebabkan hasil pemisahan tidak baik, letak nukleus yang saling berdekatan langsung juga dapat mengganggu proses pemisahan sel. Hal ini karena nukleus yang berdekatan memiliki intensitas yang mirip sehingga perubahan *gradient* menjadi kecil (memiliki nilai yang rendah) pada *region* yang seharusnya menjadi *region* ideal pemisahan sel. Karena terdapat *region* yang batasnya tidak jelas tersebut, saat proses analisis *cluster* dengan HCA dapat terjadi *over-cluster* seperti yang ditunjukkan pada Gambar 4.14.



Gambar 4.14. Hasil pemisahan sel pada citra aml\_19 di sub-citra 8.  
(a) Sub-citra, (b) Gradient magnitude dan (c) hasil pemisahan sel

*[Halaman ini sengaja dikosongkan]*

## **BAB 5**

### **PENUTUP**

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan beserta saran untuk pengembangan lebih lanjut.

#### **5.1 Kesimpulan**

Berdasarkan hasil uji coba yang dilakukan pada penelitian ini dapat disimpulkan beberapa hal.

- Metode yang diusulkan memiliki rata-rata nilai akurasi presisi 0.977, *recall* 0.885, dan *Zijdenbos similarity index* (ZSI) 0.928 pada segmentasi sitoplasma sel darah putih, sedangkan rata-rata nilai presisi 0.975, *recall* 0.924, dan ZSI 0.948 dicapai pada segmentasi *region* nukleus. Analisis *cluster* pada citra local (LKM) menghasilkan segmentasi sel darah putih dan nukleus lebih akurat dengan rata-rata ZSI 74% dan 6.1% lebih tinggi daripada metode segmentasi sitoplasma dan nukleus menggunakan analisis *cluster* metode K-Means pada citra global (GKM).
- Rata-rata nilai akurasi hasil segmentasi nukleus dari sisi presisi, *recall* dan ZSI memiliki nilai yang lebih tinggi pada hasil segmentasi nukleus dibandingkan dengan hasil segmentasi sel darah putih, baik itu metode yang diusulkan maupun pada GKM. Hal ini disebabkan karena intensitas warna nukleus yang lebih homogen dan memiliki perbedaan nilai *saturation* yang berbeda jauh dibandingkan dengan *region* lainnya seperti sitoplasma dan *background*.
- Metode yang diusulkan memiliki kelebihan dengan memiliki komposisi jumlah piksel *region* sel darah putih yang lebih seimbang antara *region* sel darah putih dengan *background* saat proses analisis *cluster* sehingga dapat melakukan segmentasi lebih baik terutama pada *region* sitoplasma dibandingkan dengan analisis *cluster* pada citra global.
- Pada proses identifikasi *touching cell* pada sub-citra, fitur geometri *roundness* dan rata-rata luasan sel tunggal secara umum telah dapat menentukan *region* dari *touching cell* dengan baik, hanya pada penentuan

jumlah *cluster* yang juga menggunakan perhitungan dari rata-rata luasan sel darah putih tunggal dapat terjadi *over-cluster* atau *under-cluster* saat memisahkan sel jika terdapat sel yang tidak tersegmentasi dengan baik pada proses di tahap sebelumnya sehingga luasan dari sel tersebut jauh di bawah ukuran sel normal.

- Pemisahan *touching cell* dengan metode *Hierarchical Clustering Analysis* (HCA) menggunakan fitur ruang spasial dan *gradient magnitude* dapat memisahkan sel darah putih menjadi sel-sel tunggal, akan tetapi terhadap *region* dari *touching cell* dengan rasio luasan *region* sitoplasma yang cukup besar dan letak nukleus yang bersebelahan langsung hasil pemisahan menjadi sangat jauh dari bentuk sel tunggal yang seharusnya.

## 5.2 Saran

Pada penelitian ini, metode yang diusulkan dengan melakukan analisis *cluster* pada citra lokal (sub-citra) secara umum telah dapat memberikan hasil segmentasi yang baik dan pemilihan ruang warna sangat menentukan keberhasilan proses segmentasi. Nilai *saturation* pada ruang warna HSV yang digunakan sebagai data analisis pada *cluster* telah mampu mendapatkan sebagian besar *region* sitoplasma yang merupakan salah satu masalah utama dalam penelitian di bidang pengolahan citra hematologi, namun untuk meningkatkan hasil segmentasi diperlukan komposisi ruang warna yang dapat memberikan informasi lebih baik pada *region* sitoplasma.

Untuk pemisahan dari *region* yang diidentifikasi sebagai *touching cell*, metode yang diusulkan dapat memisahkan *region* tersebut menjadi sel-sel tunggal. Akan tetapi hasil pemisahan masih kurang stabil sehingga sebagian sel yang telah dipisahkan memiliki bentuk yang kurang sesuai terhadap bentuk sel yang seharusnya, hal ini berkaitan kuat dengan fitur yang digunakan seperti fitur ruang spasial yang berbeda antara sub-citra dan *gradient magnitude* yang memerlukan sitoplasma agar dapat memberikan batas antar sel.

Oleh karena itu, penelitian selanjutnya diharapkan dapat mengembangkan metode ini dari kedua aspek tersebut dengan menemukan komposisi ruang warna yang mampu membuat *region* sitoplasma menjadi lebih berbeda dan memilih fitur

yang lebih baik untuk proses pemisahan *touching cell* dengan *Hierarchical Clustering Analysis* atau metode *post-processing* yang akan memperbaiki hasil pemisahan sel yang kurang baik.

*[Halaman ini sengaja dikosongkan]*

## DAFTAR PUSTAKA

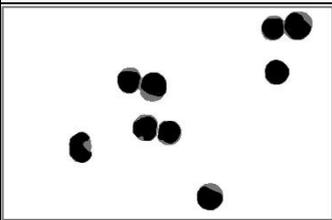
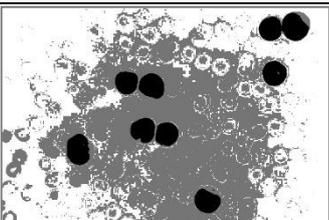
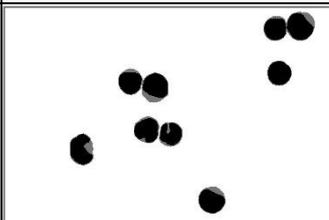
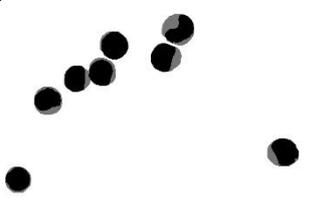
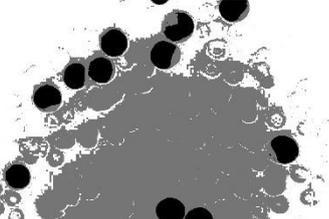
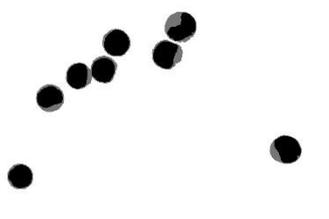
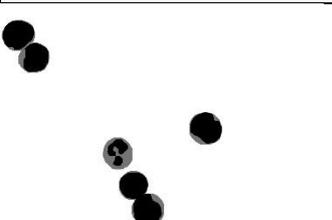
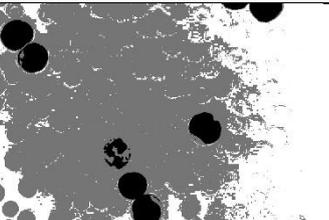
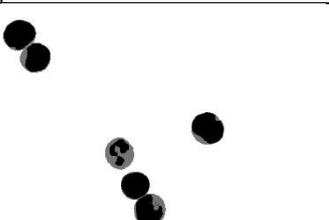
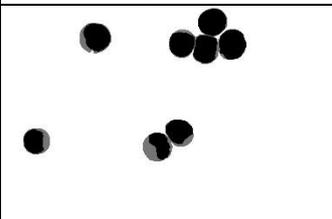
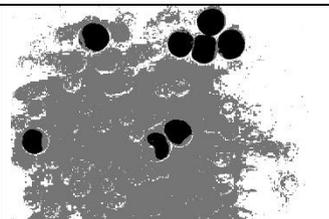
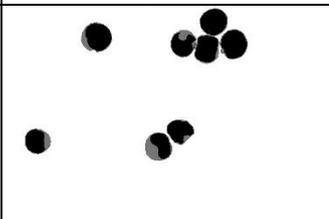
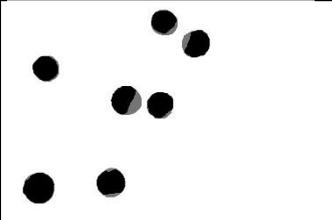
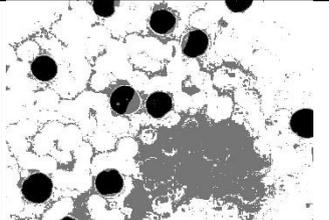
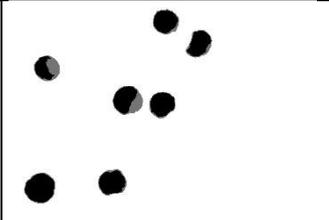
- Agusta, Y. (2007). K-Means - Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem Dan Informatika*, 3, 47–60.
- Döhner, H., Weisdorf, D.J., Bloomfield, C.D., (2015). "Acute Myeloid Leukemia". *The New England Journal of Medicine*. 373 (12): 1136–52.
- Fatichah, C., Tangel, M. L., Widyanto, M. R., Dong, F., dan Hirota, K. (2012), "Interest-Based Ordering for Fuzzy Morphology on White Blood Cell Image Segmentation", *JACIII*, Vol. 16, No. 1, hal. 76-86.
- Gonzalez, R. C., dan Woods, R. E, (2002), *Digital Image Processing*, 2<sup>nd</sup> edition, PrenticeHall, Upper Saddle River, New Jersey.
- Huang, D. C., Hung, K. D., dan Chan, Y. K. (2012), "A Computer Assisted Method for Leukocyte Nucleus Segmentation and Recognition in Blood smear Images", *Journal of Systems and Software*, Vol. 85, No. 9, hal. 2104-2118.
- Jemal, A., Thomas, A., Murray, T., Thun, M. (2002). "Cancer statistics, 2002". *CA Cancer J Clin*. 52 (1): 23–47.
- Labati, R. D., Piuri, V., dan Scotti, F. (2011), "All-IDB: The Acute Lymphoblastic Leukemia Image Database for Image Processing", *Proceedings of the 18th IEEE ICIP International Conference on Image Processing*, Eds: Macq, B., dan Schelkens, P., IEEE Signal Processing Society, Brussels, hal. 2045-2048.
- Madhloom, H. T., Kareem, S. A., dan Ariffin, H. (2012), "An Image Processing Application for the Localization and Segmentation of Lymphoblast Cell using Peripheral Blood Images", *Journal of medical systems*, Vol. 36, No. 4, hal. 2149-2158.
- Mandyarta, E.P., Fatichah, C., Three-level Local Thresholding Berbasis Metode Otsu untuk Segmentasi Leukosit pada Citra Leukemia Limfoblastik Akut. *JB I: Jurnal Buana Informatika*, vol. 7, No. 1, 2016.
- Mohapatra, S., Patra, D., dan Satpathy, S. (2014), "An Ensemble Classifier System for Early Diagnosis of Acute Lymphoblastic Leukemia in Blood Microscopic Images", *Neural Computing and Applications*, Vol. 24, No. 7-8, hal. 1887-1904.

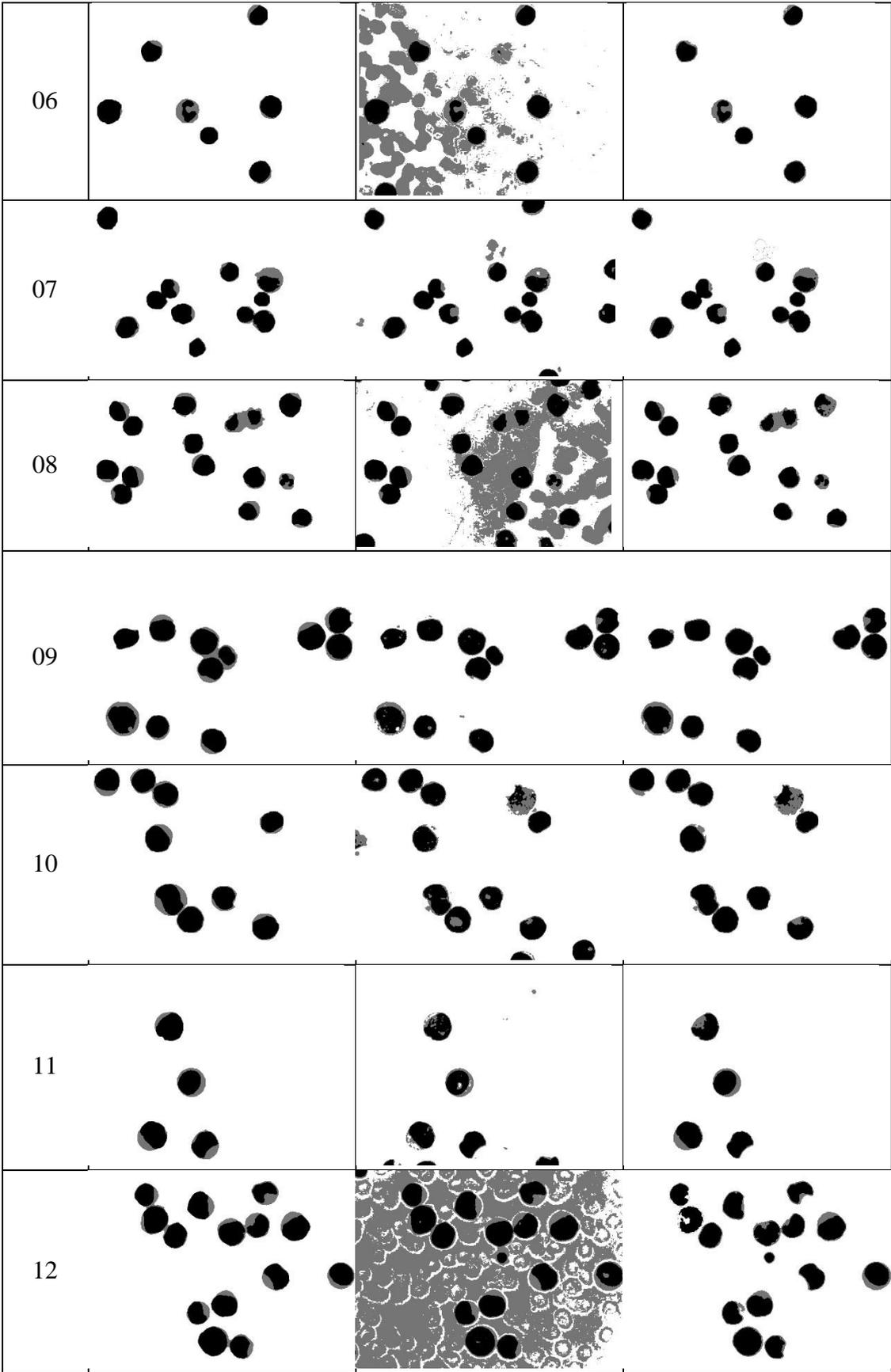
- Piuri, V., dan Scotti, F. (2004), "Morphological Classification of Blood Leucocytes by Microscope Images", *Proceedings of the 2004 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, Eds: Alippi, C. et al., IEEE, Boston, hal. 103-108.
- Prasetyo, E., (2011), *Pengolahan Citra Digital dan Aplikasinya Menggunakan Matlab*, Andi, Yogyakarta.
- Putzu, L., Caocci, G., dan Di Ruberto, C. (2014), "Leucocyte Classification for Leukaemia Detection using Image Processing Techniques", *Artificial Intelligence in Medicine*, Vol.62, No. 3, hal. 179-191.
- Rezatofighi, S. H., dan Soltanian-Zadeh, H. (2011), "Automatic Recognition of Five Types of White Blood Cells in Peripheral Blood", *Computerized Medical Imaging and Graphics*, Vol. 35, No. 4, hal. 333-343.
- Russ, J. C. (2011), *The Image Processing Handbook*, 6<sup>th</sup> edition, CRC press, Boca Raton, Florida.
- Saputra, R.A., Fatichah, C., Suciati, N. Penggabungan Fitur Tekstur yang Invariant terhadap Iluminasi dan Fitur Bentuk untuk Deteksi Acute Lymphoblastic Leukemia. *JBI: Jurnal Buana Informatika*, vol. 7, No. 1, 2016.
- Scotti, F. (2006), "Robust Segmentation and Measurements Techniques of White Cells in Blood Microscope Images", *Proceedings of the 2006 IEEE Instrumentation and Measurement Technology Conference*, Eds: Daponte, P. et al., IEEE Instrumentation and Measurement Society, Sorrento, hal. 43-48.
- Theera-Umpon, Wang, N, and Jin, Y. (Eds.) White Blood Cell Segmentation and Classification in Microscopic Bone Marrow Images Fuzzy Systems and Knowledge Discovery: *Second International Conference, FSKD 2005*, Changsha, China, August 27-29, 2005, Proceedings, Part II, Springer Berlin Heidelberg, 2005, 787-796.
- Wang, E. Y., Gou, Z., Miao, A. M., Peng, S. Q., Niu, Z. Y., dan Shi, X. L. (2009), "Recognition of Blood Cell Images based on Color Fuzzy Clustering", dalam *Fuzzy Information and Engineering Volume 2*, eds. Cao, B., Li, T., dan Zhang, C., Springer Berlin Heidelberg, New York, hal. 69-75.

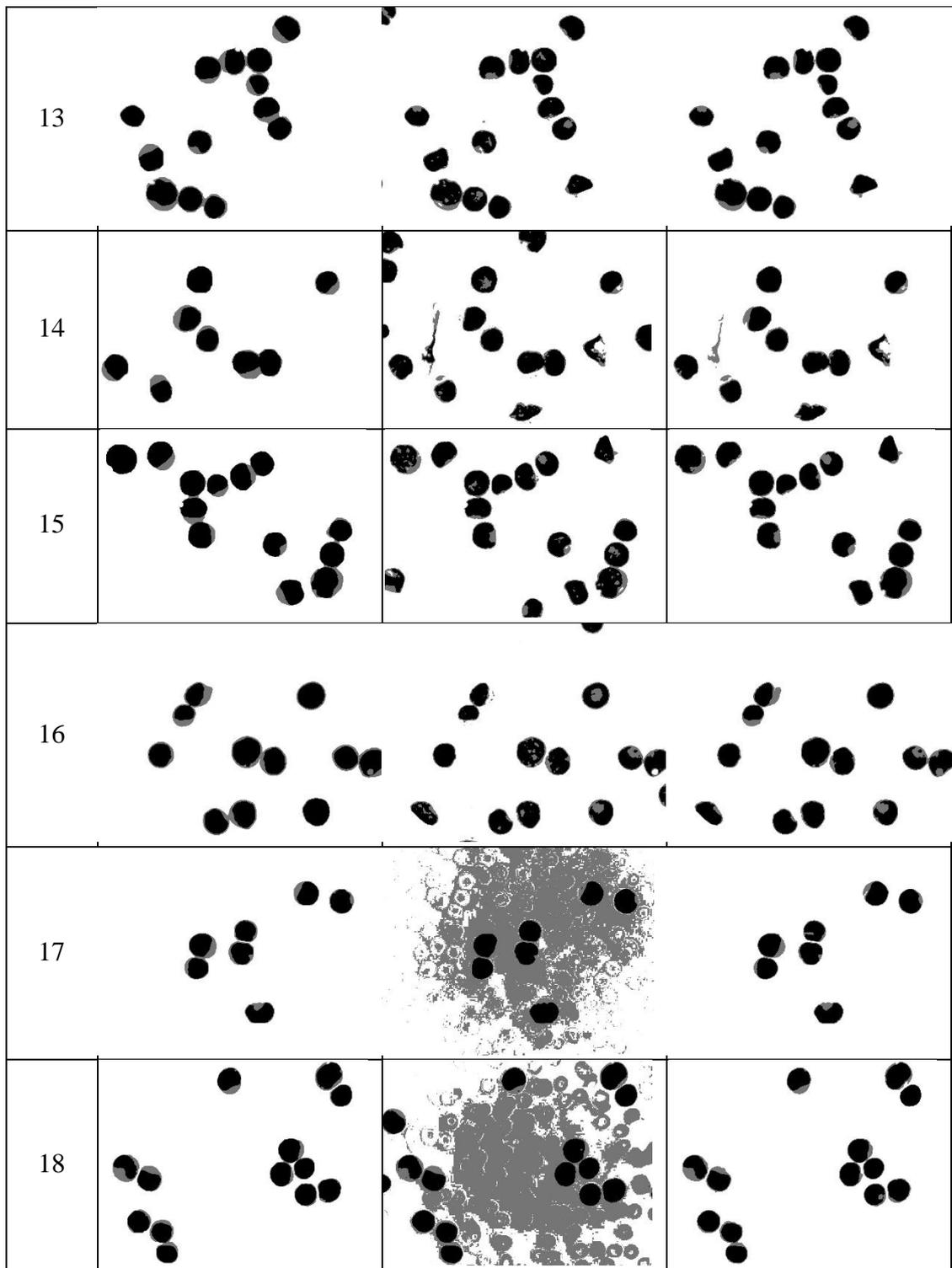
## LAMPIRAN 1

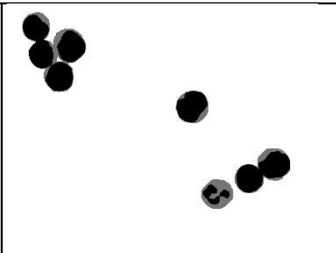
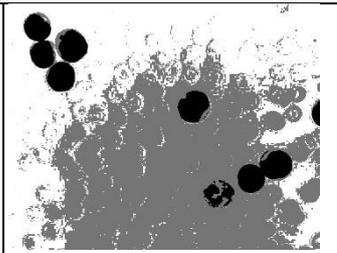
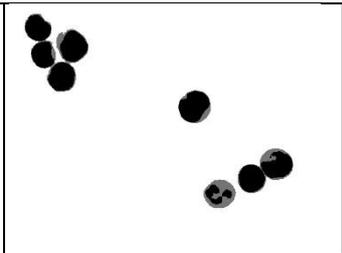
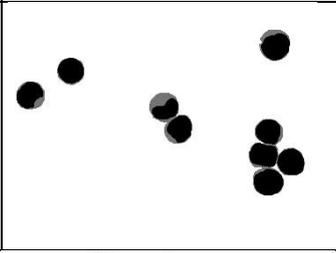
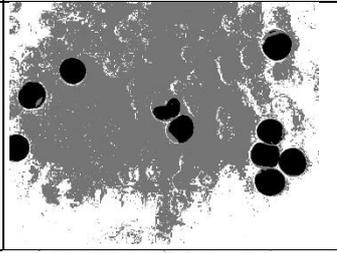
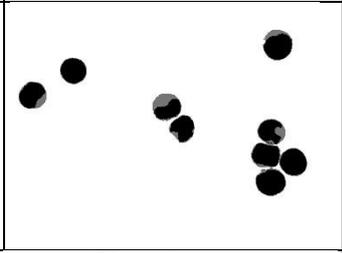
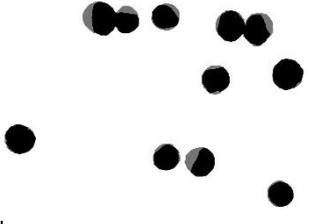
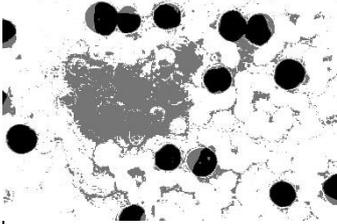
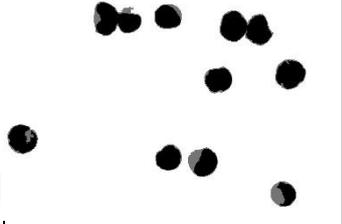
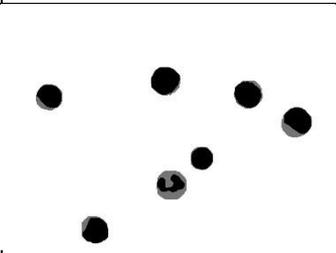
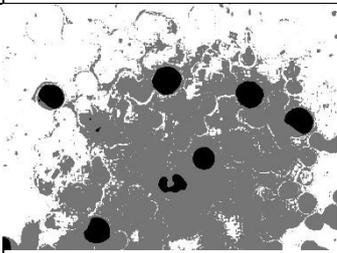
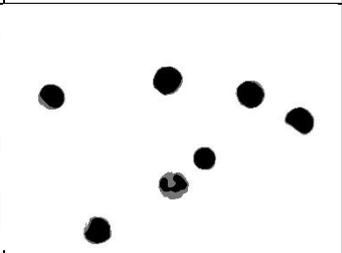
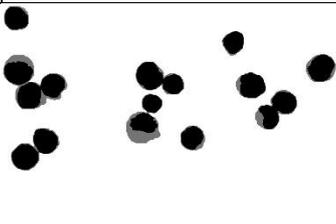
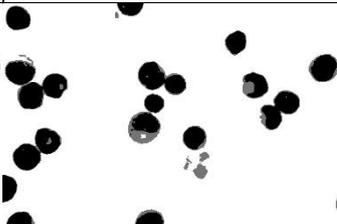
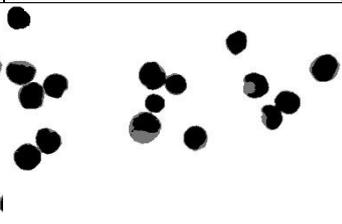
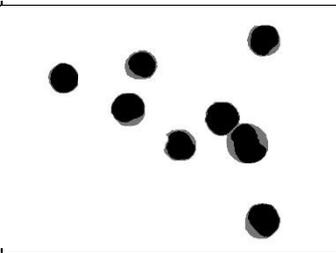
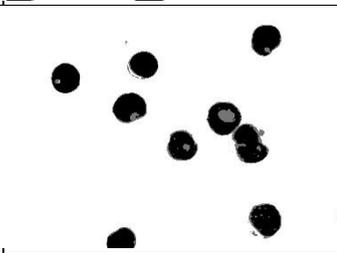
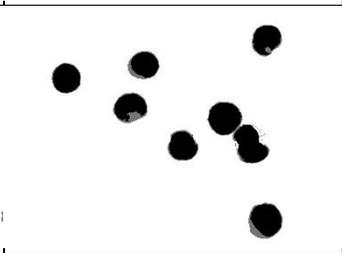
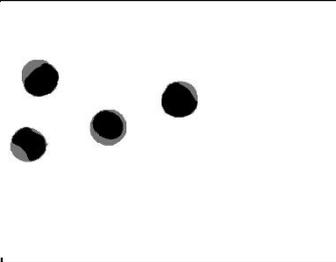
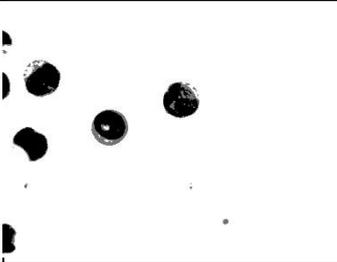
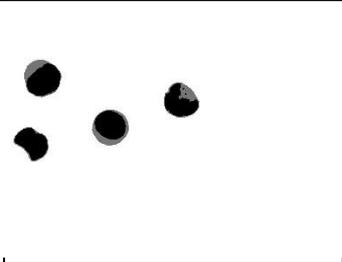
Lampiran ini berisi data dari hasil uji coba segmentasi sitoplasma dan nukleus dari sel darah putih pada 36 citra di dataset AML.

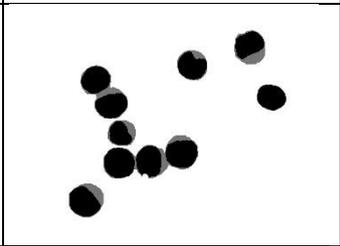
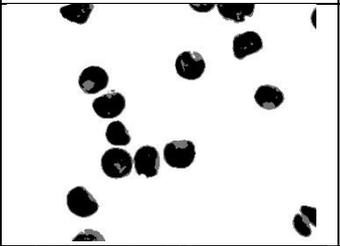
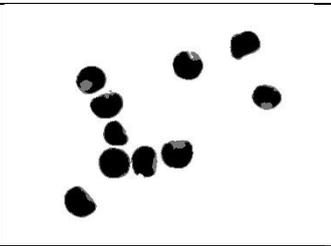
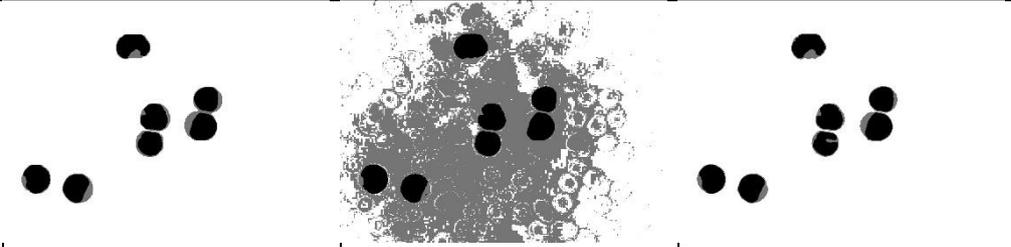
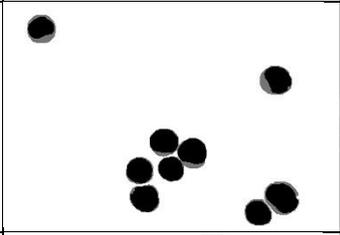
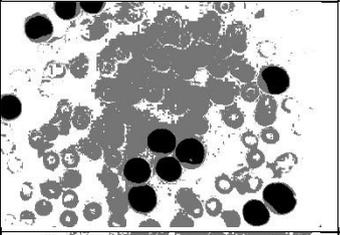
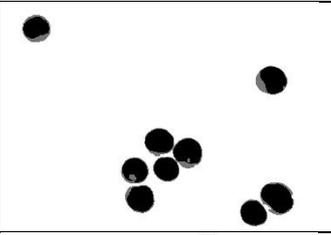
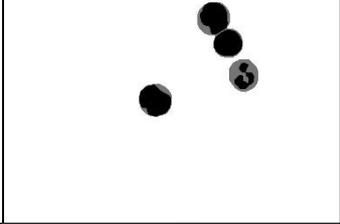
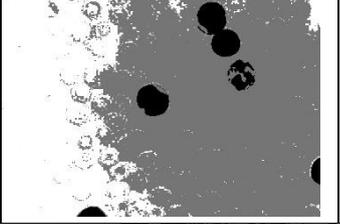
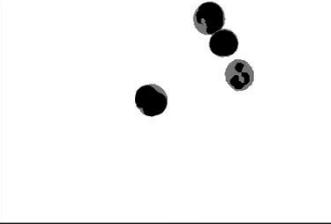
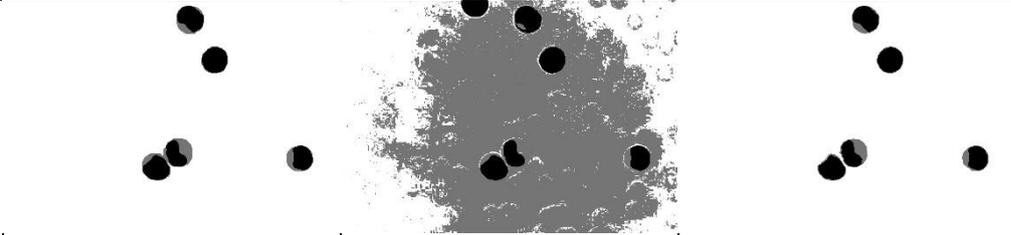
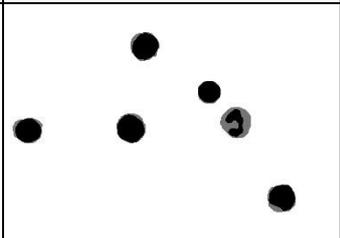
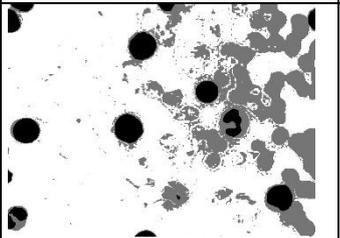
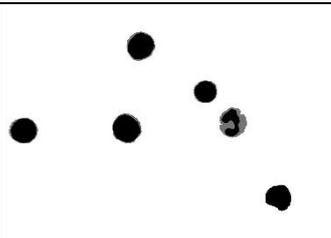
**Tabel 1.A.** Citra *ground-truth* beserta hasil segmentasi menggunakan analisis *cluster* pada citra lokal dan global.

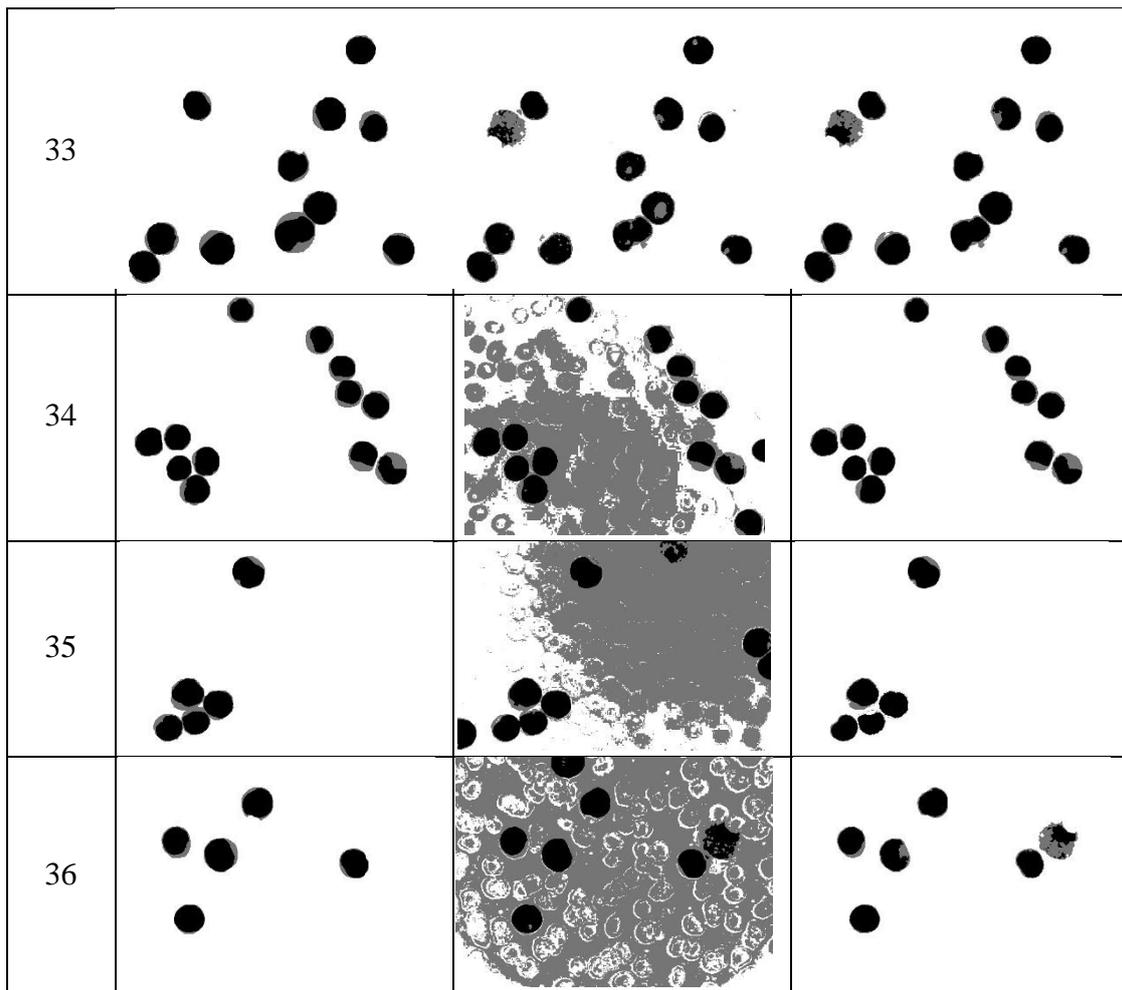
No. Citra (aml)	(a)	(b)	(c)
01			
02			
03			
04			
05			





19			
20			
21			
22			
23			
24			
25			

26			
27			
28			
29			
30			
31			
32			



Keterangan:

(a) - Ground-truth.

(b) - Hasil Segmentasi sel darah putih menggunakan K-Means pada citra global.

(c) - Hasil Segmentasi sel darah putih menggunakan K-Means pada citra lokal.

\* Warna hitam merupakan *region* nukleus dan warna keabuan merupakan *region* sitoplasma.

## LAMPIRAN 2

Lampiran ini berisi data dari hasil evaluasi segmentasi sitoplasma dan nukleus dari metode yang diusulkan (LKM) dan analisis cluster pada citra global (GKM) yang dibandingkan dengan citra *ground-truth* menggunakan nilai presisi, *recall* dan ZSI.

**Tabel 2.A.** Hasil segmentasi pada sitoplasma sel darah putih.

Citra	LKM			GKM		
	P	R	ZSI	P	R	ZSI
aml_01	0.997	0.938	0.967	0.023	0.675	0.045
aml_02	0.992	0.927	0.958	0.029	0.827	0.055
aml_03	0.998	0.935	0.966	0.018	0.739	0.036
aml_04	0.999	0.888	0.940	0.024	0.858	0.047
aml_05	0.998	0.890	0.941	0.033	0.774	0.064
aml_06	0.989	0.755	0.856	0.046	0.841	0.088
aml_07	0.994	0.909	0.949	0.451	0.721	0.555
aml_08	0.997	0.908	0.950	0.085	0.809	0.154
aml_09	1.000	0.855	0.922	0.723	0.501	0.592
aml_10	0.996	0.769	0.868	0.327	0.446	0.377
aml_11	0.999	0.881	0.937	0.025	0.807	0.048
aml_12	0.959	0.779	0.860	0.036	0.710	0.068
aml_13	0.934	0.876	0.904	0.543	0.500	0.521
aml_14	0.905	0.863	0.883	0.313	0.451	0.370
aml_15	0.926	0.892	0.909	0.273	0.433	0.335
aml_16	0.930	0.856	0.891	0.375	0.387	0.381
aml_17	0.996	0.935	0.964	0.020	0.650	0.040
aml_18	0.993	0.932	0.962	0.039	0.774	0.074
aml_19	0.999	0.917	0.956	0.025	0.652	0.047
aml_20	0.998	0.902	0.948	0.021	0.792	0.042
aml_21	0.998	0.872	0.931	0.070	0.829	0.129
aml_22	0.987	0.902	0.943	0.024	0.852	0.046
aml_23	0.997	0.854	0.920	0.565	0.680	0.618
aml_24	0.999	0.876	0.934	0.381	0.428	0.403
aml_25	0.999	0.884	0.938	0.580	0.534	0.556
aml_26	0.938	0.824	0.877	0.052	0.738	0.097
aml_27	0.999	0.839	0.912	0.261	0.340	0.295
aml_28	0.996	0.934	0.964	0.021	0.705	0.040
aml_29	0.997	0.919	0.957	0.031	0.726	0.059
aml_30	1.000	0.928	0.963	0.015	0.656	0.029
aml_31	0.998	0.904	0.949	0.015	0.844	0.029
aml_32	0.993	0.913	0.952	0.051	0.835	0.096

aml_33	0.897	0.884	0.891	0.280	0.390	0.326
aml_34	0.996	0.906	0.949	0.045	0.755	0.085
aml_35	0.997	0.907	0.950	0.013	0.734	0.026
aml_36	0.782	0.892	0.834	0.008	0.823	0.016

**Tabel 2.B.** Hasil segmentasi pada nukleus sel darah putih.

Citra	LKM			GKM		
	P	R	ZSI	P	R	ZSI
aml_01	0.991	0.976	0.983	0.924	0.999	0.960
aml_02	0.996	0.940	0.967	0.703	0.982	0.819
aml_03	0.998	0.924	0.960	0.805	0.995	0.890
aml_04	1.000	0.892	0.943	0.988	0.979	0.983
aml_05	0.999	0.856	0.922	0.749	0.964	0.843
aml_06	0.997	0.753	0.858	0.875	0.999	0.933
aml_07	0.996	0.902	0.946	0.789	0.914	0.847
aml_08	0.966	0.949	0.957	0.734	0.988	0.842
aml_09	0.998	0.939	0.968	0.995	0.947	0.971
aml_10	0.940	0.952	0.946	0.836	0.914	0.873
aml_11	0.997	0.897	0.945	0.823	0.983	0.896
aml_12	0.971	0.937	0.954	0.955	0.912	0.933
aml_13	0.918	0.932	0.925	0.895	0.910	0.902
aml_14	0.819	0.932	0.872	0.613	0.912	0.734
aml_15	0.937	0.882	0.909	0.859	0.851	0.855
aml_16	0.921	0.890	0.905	0.840	0.832	0.836
aml_17	0.982	0.982	0.982	0.907	1.000	0.951
aml_18	0.987	0.955	0.971	0.866	0.995	0.926
aml_19	0.996	0.951	0.973	0.887	0.991	0.936
aml_20	0.999	0.926	0.961	0.884	0.996	0.937
aml_21	1.000	0.862	0.926	0.872	0.979	0.923
aml_22	0.995	0.946	0.970	0.922	0.997	0.958
aml_23	0.995	0.933	0.963	0.833	0.948	0.887
aml_24	0.993	0.943	0.967	0.895	0.891	0.893
aml_25	0.987	0.911	0.947	0.833	0.907	0.869
aml_26	0.933	0.897	0.915	0.779	0.983	0.869
aml_27	0.998	0.860	0.924	0.769	0.854	0.810
aml_28	0.992	0.969	0.980	0.927	1.000	0.962
aml_29	0.994	0.955	0.974	0.800	0.999	0.888
aml_30	0.996	0.955	0.975	0.757	1.000	0.861
aml_31	0.998	0.966	0.981	0.824	0.997	0.902
aml_32	0.993	0.955	0.974	0.731	0.995	0.843
aml_33	0.951	0.916	0.933	0.955	0.879	0.916
aml_34	0.989	0.953	0.970	0.845	0.997	0.915
aml_35	0.997	0.962	0.979	0.625	0.997	0.768
aml_36	0.891	0.908	0.899	0.678	0.994	0.806

### LAMPIRAN 3

Lampiran ini berisi data dari hasil perhitungan (*counting*) jumlah sel darah putih pada dataset AML.

**Tabel 3.A.** Hasil perhitungan jumlah sel darah putih.

Citra	<i>Ground truth</i>	Metode yang diusulkan			
		<i>Counted</i>	<i>Correctly counted</i>	<i>Over counted</i>	<i>Not Counted</i>
aml_01	9	9	9	0	0
aml_02	8	8	8	0	0
aml_03	6	6	6	0	0
aml_04	8	8	8	0	0
aml_05	7	7	7	0	0
aml_06	7	6	6	0	1
aml_07	11	11	11	0	0
aml_08	15	15	15	0	0
aml_09	11	11	10	1	1
aml_10	9	9	9	0	0
aml_11	4	4	4	0	0
aml_12	14	13	13	0	1
aml_13	14	13	13	0	1
aml_14	8	9	8	1	0
aml_15	13	15	13	2	0
aml_16	11	12	11	1	0
aml_17	7	7	7	0	0
aml_18	13	13	13	0	0
aml_19	8	8	8	0	0
aml_20	9	9	9	0	0
aml_21	11	11	11	0	0
aml_22	7	7	7	0	0
aml_23	16	15	15	0	1
aml_24	8	9	8	1	0
aml_25	4	4	4	0	0
aml_26	14	12	12	0	2
aml_27	10	10	10	0	0
aml_28	7	7	7	0	0
aml_29	9	9	9	0	0

aml_30	4	4	4	0	0
aml_31	5	5	5	0	0
aml_32	6	6	6	0	0
aml_33	11	13	11	2	0
aml_34	12	12	12	0	0
aml_35	5	4	4	0	1
aml_36	5	6	5	1	0

## BIODATA PENULIS



Aryo Harto, S.ST. adalah putra tunggal dari pasangan Iswiyanto Adi Bimasakti dan Chairiah Harun. Lahir di Banjarmasin pada tahun 1989. Menempuh pendidikan tingkat Sekolah Dasar di SD Islam Sabilal Muhtadin Banjarmasin, Sekolah Menengah Pertama di SMP Islam Sabilal Muhtadin Banjarmasin dan Sekolah Menengah Atas di SMA Negeri 2 Banjarmasin.

Kemudian melanjutkan studi D-III pada tahun 2007 di Teknik Informatika Politeknik Negeri Banjarmasin hingga lulus pada tahun 2010. Pada tahun 2012 penulis kembali melanjutkan studi D-IV di Teknik Informatika Politeknik Elektronika Negeri Surabaya dan lulus pada tahun 2013.

Selama studi di tingkat diploma penulis aktif di beberapa organisasi kemahasiswaan seperti Badan Eksekutif Mahasiswa, Dewan Perwakilan Mahasiswa dan Lembaga Pers Mahasiswa. Sementara selama studi S2 di Teknik Informatika ITS penulis pernah berkontribusi pada kegiatan penelitian di bidang Komputasi Cerdas dan Rekayasa Perangkat Lunak.

Selain aktif di bidang pendidikan dan kemahasiswaan, penulis juga pernah aktif di perusahaan swasta di bidang Agro Industri dan *Plantation*. Penulis dapat dihubungi melalui email di alamat [ryouquinze@gmail.com](mailto:ryouquinze@gmail.com).