



TESIS

**SEGMENTASI CITRA SEL TUNGGAL SMEAR SERVIKS
MENGUNAKAN METODE *RADIATING NORMALLY BIASED
GENERALIZED GRADIENT VECTOR FLOW SNAKE***

Martini Dwi Endah Susanti

NRP. 05111650010056

DOSEN PEMBIMBING

Prof. Ir. Handayani Tjandrasa, M.Sc., Ph.D.

NIP: 19490823 19760320 01

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

NIP: 19751220 20011220 02

PROGRAM MAGISTER

DEPARTEMEN INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI DAN KOMUNIKASI

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2018



THESIS

**CERVICAL SMEAR IMAGES SEGMENTATION USING RADIATING
NORMALLY BIASED GENERALIZED GRADIENT VECTOR FLOW
SNAKE**

Martini Dwi Endah Susanti

NRP. 05111650010056

SUPERVISOR

Prof. Ir. Handayani Tjandrasa, M.Sc., Ph.D.

NIP: 19490823 19760320 01

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

NIP: 19751220 20011220 02

MASTER PROGRAM

DEPARTMENT OF INFORMATICS

**FACULTY OF INFORMATION TECHNOLOGY AND
COMMUNICATION**

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2018

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

oleh:
Martini Dwi Endah Susanti
Nrp. 05111650010056

Dengan judul :
SEGMENTASI CITRA SEL TUNGGAL SMEAR SERVIKS
MENGUNAKAN METODE RADIATING NORMALLY BIASED
GENERALIZED GRADIENT VECTOR FLOW SNAKE

Tanggal Ujian : 23-7-2018
Periode Wisuda : 2018 Genap

Disetujui oleh:

Prof. Ir. Handayani Tjandrasa, M.Sc., Ph.D.
NIP. 194908231976032001


(Pembimbing 1)

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


(Pembimbing 2)

Dr. Agus Zainal Arifin, S.Kom., M.Kom.
NIP. 197208091995121001


(Penguji 1)

Tohari Ahmad, S.Kom., MIT., Ph.D.
NIP. 197505252003121002


(Penguji 2)

Dr.Eng. Darlis Herumurti, S.Kom., M.Kom.
NIP. 197712172003121001


(Penguji 3)

DEKAN FTIK - ITS



Dr. Agus Zaidi Arifin, S.Kom., M.Kom.
NIP. 197208091995121001

SEGMENTASI CITRA SEL TUNGGAL SMEAR SERVIKS MENGUNAKAN METODE *RADIATING NORMALLY BIASED GENERALIZED GRADIENT VECTOR FLOW SNAKE*

Nama : Martini Dwi Endah Susanti
NRP : 05111650010056
Pembimbing 1 : Prof. Ir. Handayani Tjandrasa, M.Sc., Ph.D.
Pembimbing 2 : Dr. Eng. Chastine Fatichah, S. Kom, M.Kom

ABSTRAK

Kanker serviks merupakan jenis kanker yang menyerang wanita yang mengakibatkan kematian setiap tahunnya. Pap smear merupakan sebuah langkah medis untuk mendeteksi ada atau tidaknya gangguan pada sel serviks wanita. Analisis hasil pemeriksaan pap smear secara manual memiliki kelemahan yaitu membutuhkan banyak tenaga ahli di bidang patologi, memakan banyak waktu, dan rawan terhadap kesalahan sehingga dibutuhkan sebuah sistem penyaringan otomatis dan sistem diagnosa yang akurat. Langkah yang paling utama dari sistem tersebut adalah proses segmentasi citra sel nukleus dan sitoplasma pada citra hasil pemeriksaan pap smear karena dapat memengaruhi keakuratan sistem.

Normally Biased Generalized Gradient Vector Flow Snake (NBGGVFS) merupakan sebuah algoritma pengganti gaya eksternal untuk *active contour (snake)* yang menggabungkan metode *Generalized Gradient Vector Flow Snake* (GGVFS) dan *Normally Biased Gradient Vector Flow Snake* (NBGVFS). Dalam memodelkan *snake* terdapat fungsi *edge map* yang biasanya dihitung dengan menggunakan operator deteksi tepi seperti *sobel*, namun metode tersebut tidak dapat mendeteksi daerah nukleus dari citra smear serviks dengan benar.

Penelitian ini mengusulkan metode baru untuk segmentasi citra sel tunggal smear serviks menggunakan metode *Radiating Normally Biased Generalized Gradient Vector Flow Snake* (RNBGGVFS). Penelitian ini memanfaatkan penggunaan *Radiating Edge Map* (REM) untuk menghitung *edge map* dari citra dengan metode NBGGVFS. Metode yang diusulkan terdiri atas 3 tahapan utama, yaitu tahap praproses, citra asli smear serviks RGB dikonversi ke dalam ruang warna CIELAB dan layer L^* dinormalisasi untuk mendapatkan citra keabuan kemudian dilakukan penghilangan *noise* dengan metode *median filtering*. Citra smear serviks kemudian dibagi ke dalam tiga bagian yaitu nukleus, sitoplasma dan area latar belakang menggunakan *mean shift clustering*. Selanjutnya tahap terakhir adalah segmentasi kontur dengan metode RNBGGVFS sebagai fungsi pengganti gaya eksternal *snake* yang bertujuan untuk mendapatkan kontur nukleus dan sitoplasma citra sel tunggal smear serviks.

Berdasarkan uji coba, nilai rata-rata *Zijdenbos Similarity Index* (ZSI) dan akurasi untuk segmentasi nukleus adalah 96,96% dan 90,68%. Nilai rata-rata ZSI dan akurasi untuk segmentasi sitoplasma adalah 86,78% dan 89,35%. Hasil pengujian menunjukkan bahwa metode yang diusulkan mampu mendeteksi area nukleus dan sitoplasma pada citra sel tunggal smear serviks yang lebih optimal dibanding metode sebelumnya.

Kata Kunci : Segmentasi citra, citra smear serviks, *mean shift clustering*, *Radiating Normally Biased Generalized Gradient Vector Flow Snake*

(Halaman ini sengaja dikosongkan)

CERVICAL SMEAR IMAGES SEGMENTATION USING RADIATING NORMALLY BIASED GENERALIZED GRADIENT VECTOR FLOW SNAKE

Name : Martini Dwi Endah Susanti
Student Identity Number : 05111650010056
Supervisor 1 : Prof. Ir. Handayani Tjandrasa, M.Sc., Ph.D.
Supervisor 2 : Dr. Eng. Chastine Fatichah, S. Kom, M.Kom

ABSTRACT

Cervical cancer is a type of cancer that affects women, resulting in women death every year. Pap smear is a medical step to detect the abnormality of cervical cells among women. The analysis of pap smear examination results manually has a disadvantage that requires a lot of experts in the field of pathology, time consuming, and prone to errors, so it needs an automatic filtration system and accurate diagnostic system. The most important step of the system is the segmentation process of nucleus and cytoplasm cells in pap smear examination image, because it can affect the accuracy system.

Normally Biased Generalized Gradient Vector Flow Snake (NBGGVFS) is an external force algorithm for active contour (snake) that combines Generalized Gradient Vector Flow Snake (GGVFS) and Normally Biased Gradient Vector Flow Snake (NBGVFS) methods. In snake models, there is an edge map function. Edge maps are typically calculated using edge detection operators such as sobel. However, this method can not detect the nucleus region of cervical smear image correctly.

This research proposes a new method for segmentation of single cell cervical smears images using the Radiating Normally Biased Generalized Gradient Vector Flow Snake (RNBGGVFS) method. This method uses Radiating Edge Map to calculate image edge map with NBGGVFS method. The proposed method consists of 3 main stages, namely the pre-process stage, the original RGB cervical smear image will be converted into the CIELAB color space and the L* channel is normalized to obtain a gray image and then remove noise using median filtering method. The cervical smear image is divided into three parts (nucleus, cytoplasm and background area) using mean shift clustering. The last stage is contour segmentation using RNBGGVFS method as a replacement function of external force snake which aims to get the contour nucleus and cytoplasm of single cell cervical smear image.

Based on the experimental result, the average value of Zijdenbos Similarity Index (ZSI) and accuracy for nucleus segmentation is 96,96% and 90,68%. Then, the average value of ZSI and accuracy for cytoplasm segmentation is 86,78% and 89,35%. The experimental result show that the proposed method can detect the nucleus and cytoplasm area of single cell cervical smear images optimally than the previous method.

Keywords : image segmentation, cervical smear images, mean shift clustering, Radiating Normally Biased Generalized Gradient Vector Flow Snake.

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Alhamdulillahirabbil'alamin. Segala puji dan syukur penulis panjatkan kehadiran Allah Swt. Atas berkat, rahmat dan hidayah-Nya, penyusunan tesis ini dapat diselesaikan. Tesis ini ditulis sebagai salah satu syarat dalam menyelesaikan Program Studi Magister di Teknik Informatika Institut Teknologi Sepuluh Nopember Surabaya, Penulis menyadari bahwa tesis ini dapat diselesaikan karena dukungan dari berbagai pihak, baik dalam bentuk dukungan moral dan material.

Melalui kesempatan ini dengan segenap kerendahan hati penulis mengucapkan terima kasih sebesar-besarnya kepada semua orang untuk segala bantuan yang telah diberikan, antara lain kepada :

1. Kedua orang tua, kakak dan adik dan seluruh keluarga tercinta untuk semua doa, dukungan dan motivasi kepada penulis dalam menyelesaikan pendidikan magister ini.
2. Ibu Prof. Ir. Handayani Tjandrasa, M.Sc., Ph.D. selaku pembimbing 1 yang telah membantu, membimbing dan memotivasi penulis dalam menyelesaikan tesis ini dengan sabar. Semoga Tuhan YME senantiasa merahmati Ibu dan keluarga.
3. Ibu Dr. Eng. Chastine Fatichah, S.Kom., M.Kom. selaku pembimbing 2 serta dosen wali penulis yang telah sabar membantu, membimbing dan memotivasi penulis dalam menyelesaikan tesis ini. Semoga Allah SWT senantiasa merahmati Ibu dan keluarga.
4. Bapak Dr. Agus Zainal Arifin, S.Kom., M.Kom., Dr. Eng. Darlis Herumurti, S.Kom., M.Kom., dan Tohari Ahmad, S.Kom., MIT, Ph.D. sebagai tim penguji tesis yang memberikan masukan dan kritik yang membangun untuk tesis ini.
5. Seluruh dosen S2 Teknik Informatika yang telah memberikan ilmu dan pengetahuan kepada penulis selama menempuh studi.
6. Teman seperjuangan S2 TC angkatan 2016 yang tidak dapat disebutkan satu persatu, terima kasih atas segala bantuan dan motivasi yang telah diberikan.

Akhirnya dengan segala kerendahan hati penulis menyadari masih banyak terdapat kekurangan pada tesis ini. Penulis berharap tesis ini dapat bermanfaat bagi pengembangan ilmu pengetahuan. Sehingga dengan kerendahan hati, penulis mengharap kritik dan saran yang membangun dari semua pihak untuk perbaikan ke depannya. Penulis berharap bahwa perbuatan baik dari semua orang yang dengan tulus memberikan kontribusi terhadap penyusunan tesis ini mendapatkan pahala dari Allah. *Aamiin Allahumma Aamiin.*

Surabaya, Juli 2018.

Martini Dwi Endah Susanti

(Halaman ini sengaja dikosongkan)

DAFTAR ISI

| | |
|---|------|
| LEMBAR PENGESAHAN TESIS..... | v |
| ABSTRAK..... | vii |
| ABSTRACT..... | ix |
| KATA PENGANTAR | xi |
| DAFTAR ISI..... | xiii |
| DAFTAR GAMBAR | xv |
| DAFTAR TABEL..... | xvii |
| BAB 1 PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Tujuan Penelitian..... | 4 |
| 1.3 Perumusan Masalah..... | 4 |
| 1.4 Batasan Masalah..... | 4 |
| 1.5 Manfaat Penelitian..... | 4 |
| 1.6 Kontribusi Penelitian..... | 4 |
| BAB 2 KAJIAN PUSTAKA | 7 |
| 2.1 Kanker Serviks | 7 |
| 2.2 Pemeriksaan Pap Smear | 7 |
| 2.3 Citra Digital..... | 8 |
| 2.3.1 Citra Biner | 9 |
| 2.3.2 Citra Keabuan | 9 |
| 2.3.3 Citra Berwarna..... | 9 |
| 2.3.4 Ruang warna CIELAB..... | 9 |
| 2.4 Pengolahan Citra Digital | 9 |
| 2.5 Median Filtering | 10 |
| 2.6 Segmentasi Citra | 10 |
| 2.7 Mean Shift Clustering | 11 |
| 2.8 Active Contour | 12 |
| 2.8.1 Gradient Vector Flow Snake | 14 |
| 2.8.3 Normally Biased Gradient Vector Flow Snake | 18 |
| 2.8.4 Normally Biased Generalized Gradient Vector Flow Snake..... | 19 |
| 2.9 Radiating Edge Map..... | 20 |

| | |
|--|----|
| BAB 3 METODE PENELITIAN | 23 |
| 3.1 Tahapan Penelitian | 23 |
| 3.1.1 Studi Literatur | 23 |
| 3.1.2 Analisis Masalah | 23 |
| 3.1.3 Desain dan Implementasi Metode | 24 |
| 3.1.4 Pengujian dan Analisis | 24 |
| 3.1.5 Penulisan Laporan | 24 |
| 3.2 Data Masukan | 25 |
| 3.3 Rancangan Metode | 25 |
| 3.3.1 Pra Proses | 25 |
| 3.3.2 Segmentasi Awal | 27 |
| 3.3.2.1 Segmentasi mean shift clustering | 27 |
| 3.3.2.2 Ekstraksi Kontur Awal | 28 |
| 3.3.2.3 Pelabelan centroid dengan bobot intensitas | 29 |
| 3.3.3 Segmentasi Kontur | 29 |
| 3.3.3.1 Perhitungan Radiating Edge Map | 30 |
| 3.3.3.2 Generalized Gradient Vector Flow Snake | 30 |
| 3.3.3.3 Normally Biased Generalized Gradient Vector Flow Snake | 31 |
| 3.3.3.4 Deformasi Snake | 31 |
| 3.4 Uji Coba | 32 |
| 3.5 Evaluasi | 32 |
| BAB 4 UJI COBA DAN ANALISIS | 35 |
| 4.1 Lingkungan Uji Coba | 35 |
| 4.2 Data Uji Coba | 35 |
| 4.3 Hasil Uji Coba | 35 |
| 4.3.1 Hasil Uji Coba Skenario 1 | 35 |
| 4.3.2 Hasil Uji Coba Skenario 2 | 38 |
| 4.4 Analisis Hasil Uji Coba | 41 |
| BAB 5 PENUTUP | 47 |
| 5.1 Kesimpulan | 47 |
| 5.2 Saran | 47 |
| DAFTAR PUSTAKA | 49 |
| LAMPIRAN 1 | 51 |
| LAMPIRAN 2 | 53 |
| BIODATA PENULIS | 57 |

DAFTAR GAMBAR

| | |
|---|----|
| Gambar 2.1 Citra sel tunggal smear serviks: (a) <i>carcinoma in situ</i> , (b) <i>light dysplastic</i> , (c) <i>moderate dysplastic</i> , (d) <i>normal columnar</i> , (e) <i>normal intermediate</i> , (f) <i>normal superficiel</i> , (g) <i>severe dysplastic</i> (Jantzen et al. 2006)..... | 8 |
| Gambar 2.2 Bentuk dasar <i>active contour</i> (Kass et al. 1988). | 13 |
| Gambar 2.3 Contoh <i>active contour</i> : (a) Inisial kontur; (b) Hasil dengan algoritma <i>snake</i> ; (c) Hasil dengan algoritma GVF <i>snake</i> (Xu et al. 2002)..... | 15 |
| Gambar 2.4 <i>Edge map</i> (Xu et al. 2002). | 16 |
| Gambar 2.5 Dengan normalisasi GVF (Xu et al. 2002). | 16 |
| Gambar 2.6 Citra smear serviks; (a) citra hasil preproses, (b) <i>edge map</i> dengan <i>sobel</i> , (c) daerah nukleus yang sebenarnya tidak terdeteksi dengan benar, (d) citra yang disegmentasi dengan GVF <i>snake</i> (Li et al. 2012). | 20 |
| Gambar 2.7 Ilustrasi <i>radiating line</i> (Li et al. 2012)..... | 20 |
| | |
| Gambar 3.1 Tahapan penelitian | 23 |
| Gambar 3.2 Contoh citra yang belum berhasil disegmentasi. | 24 |
| Gambar 3.3 Contoh citra sel tunggal smear serviks beserta citra segmentasi manual (Jantzen et al. 2006). | 25 |
| Gambar 3.4 Desain Sistem Usulan | 26 |
| Gambar 3.5 Diagram alir pra proses citra smear serviks. | 27 |
| Gambar 3.6 Diagram alir proses segmentasi awal..... | 28 |
| Gambar 3.7 Diagram alir segmentasi kontur. | 31 |
| | |
| Gambar 4.1 Contoh hasil segmentasi menggunakan metode RNBGGVFS dan <i>mean shift clustering</i> dengan beberapa variasi nilai <i>bandwidth</i> (radius). | 38 |
| Gambar 4.2 Contoh citra yang dapat dianalisis pada metode RNBGGVFS & <i>mean shift clustering</i> namun belum berhasil pada metode sebelumnya..... | 40 |
| Gambar 4.3 Contoh hasil segmentasi awal metode <i>mean shift clustering</i> | 40 |
| Gambar 4.4 Contoh hasil segmentasi pada beberapa metode..... | 41 |
| Gambar 4.5 Contoh citra yang tidak berhasil disegmentasi. | 44 |

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

| | |
|---|----|
| Tabel 3.1 Distribusi 917 sel tunggal citra smear <i>serviks</i> dalam dataset Herlev (Jantzen et al. 2006). | 27 |
| Tabel 3.2 Tabel <i>Confusion Matrix</i> | 32 |
| Tabel 4.1 Jumlah citra yang dapat dianalisis. | 36 |
| Tabel 4.2 Hasil uji coba dengan nilai <i>bandwidth</i> = 0,2. | 37 |
| Tabel 4.3 Hasil uji coba dengan nilai <i>bandwidth</i> = 0,25. | 37 |
| Tabel 4.4 Hasil uji coba dengan nilai <i>bandwidth</i> = 0,3. | 38 |
| Tabel 4.5 Jumlah citra yang dapat dianalisis pada uji coba skenario 2. | 39 |
| Tabel 4.6 Hasil evaluasi segmentasi nukleus citra sel tunggal smear <i>serviks</i> pada beberapa metode..... | 42 |
| Tabel 4.7 Hasil evaluasi segmentasi nukleus citra sel tunggal smear <i>serviks</i> pada beberapa metode..... | 42 |
| Tabel 4.8 Hasil evaluasi segmentasi sitoplasma citra sel tunggal smear <i>serviks</i> pada beberapa metode..... | 43 |
| Tabel 4.9 Hasil evaluasi segmentasi sitoplasma citra sel tunggal smear <i>serviks</i> pada beberapa metode..... | 44 |
| Tabel 4.10 Hasil evaluasi segmentasi nukleus pada semua kelas dataset Herlev..... | 45 |
| Tabel 4.11 Hasil evaluasi segmentasi nukleus pada semua kelas dataset Herlev..... | 45 |
| Tabel 4.12 Contoh data keluaran sistem. | 46 |

(Halaman ini sengaja dikosongkan)

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Menurut data yang terdapat pada World Health Organization (WHO), kanker serviks merupakan jenis kanker kedua yang menyerang wanita, yang mengakibatkan kematian. George Papanicolaou memperkenalkan sebuah teknik pemeriksaan bernama Pap tes, sebuah teknik yang sederhana, terjangkau dan efektif untuk mendiagnosa kanker serviks. Pemeriksaan pap smear juga dapat mendeteksi perubahan pra-kanker pada sel-sel serviks berdasarkan warna, bentuk, dan tekstur nukleus dan sitoplasma (Kale et al. 2010). Hasil pemeriksaan akan menunjukkan sel-sel serviks masih normal atau sudah terdapat gejala awal kanker serviks. Analisis hasil pemeriksaan pap smear secara manual memiliki kelemahan yaitu membutuhkan banyak tenaga ahli di bidang patologi, memakan banyak waktu, dan rawan terhadap kesalahan. Sebuah sistem penyaringan otomatis dan sistem diagnosa yang akurat untuk hasil pemeriksaan pap smear akan sangat bermanfaat untuk mengatasi kelemahan-kelemahan tersebut. Langkah yang paling utama dan paling penting dari sistem tersebut adalah proses segmentasi sel nukleus dan sitoplasma pada citra hasil pemeriksaan pap smear karena dapat memengaruhi keakuratan sistem.

Hingga kini, berbagai metode segmentasi telah diusulkan untuk citra sel serviks. Algoritma berbasis kontur aktif atau *snake* merupakan salah satu algoritma yang dapat digunakan dalam proses segmentasi citra smear serviks untuk mendapatkan bagian nukleus, sitoplasma dan area latar belakang (Kass et al. 1988). Kontur aktif atau *snake* dapat dihitung dengan meminimasi fungsi energi eksternal dan internal yang dihitung dari data citra. Kelemahan *snake* adalah kontur awal harus ditempatkan dekat dengan objek dan konvergensi *snake* sangat lemah terhadap batas kecekungan objek (Xu et al. 1998). Untuk mengatasi masalah tersebut, *snake* dikembangkan menjadi *Gradient Vector Flow Snake* (GVFS) sebagai pengganti energi eksternal *snake* (Xu et al. 1998). Fungsi energi eksternal ini digunakan untuk mengatasi kesulitan dalam mendapatkan batas kecekungan. Metode GVFS digunakan untuk segmentasi sel nukleus dan sitoplasma dari citra smear serviks (Yang-Mao et al. 2008). Namun, GVFS memiliki konvergensi yang lemah dan sensitif terhadap inisialisasi *snake* (Qin et al. 2013). Kelemahan ini disebabkan karena penghalusan yang berlebihan pada area batas objek sehingga belum bisa mendapatkan hasil segmentasi kontur sel nukleus dan sitoplasma dari citra

smear serviks secara akurat. Metode GVFS yang menggunakan operator deteksi tepi seperti *sobel* untuk melakukan pencarian *edge map* menyebabkan kurang optimalnya pendeteksian area nukleus citra dengan benar karena area nukleus umumnya lebih gelap dari pada area sitoplasma dan area sitoplasma lebih gelap dibanding dengan area latar belakang (Li et al. 2012).

Kemudian kombinasi metode *spatial k-means* dan *Radiating Gradient Vector Flow Snake (RGVFS)* diusulkan untuk melakukan segmentasi kontur sitoplasma dan nukleus citra smear serviks (Li et al. 2012). Hal yang diusulkan adalah perhitungan *edge map* pada metode GVFS menggunakan *Radiating Edge Map (REM)*, REM dapat dihitung di sepanjang *radiating line* yang dimulai dari bobot intensitas *centroid* nukleus. Penelitian tersebut menggunakan metode *spatial k-means* untuk segmentasi awal dan menyebabkan beberapa citra tidak dapat tersegmentasi dikarenakan ukuran nukleus yang berukuran sangat kecil dan terdapat banyak *noise*. Metode RGVFS juga digunakan pada penelitian sebelumnya dalam mendeteksi kontur sitoplasma dan nukleus citra smear serviks dengan dikombinasikan dengan metode *Fuzzy C-Means (FCM)* (Susanti et al. 2015). Namun metode tersebut ternyata juga belum berhasil untuk mendapatkan kontur sitoplasma dan nukleus secara akurat. Hal ini disebabkan karena metode perhitungan REM dimanfaatkan pada metode GVFS, dimana GVFS memiliki konvergensi yang lemah dan sensitif terhadap inisialisasi *snake* (Li et al. 2012). Metode ini juga belum berhasil melakukan segmentasi ke beberapa citra dikarenakan metode FCM sangat sensitif terhadap *noise* (Zhao et al. 2011).

Generalized Gradient Vector Flow Snake (GGVFS) diusulkan dengan tujuan untuk meningkatkan konvergensi GVFS (Xu et al. 1998). GGVFS didapat dengan mengganti koefisien *smoothing* pada GVFS dengan dua koefisien pembobot pada persamaan GGVFS. Namun, GGVFS sendiri ternyata belum dapat meningkatkan konvergensi batas kecekungan *snake* (Qin et al. 2013). Kelemahan yang dihadapi GVFS dan GGVFS adalah karena konsistensi komponen energi eksternal *snake* yang diperlukan untuk menarik *snake* ke arah objek secara signifikan lebih kecil dalam besaran (*magnitude*) yang dapat mengganggu proses normalisasi vektor, sehingga metode *Component Normalized Generalized Gradient Vector Flow Snake (CNGGVFS)* diusulkan. CNGGVFS didapatkan dengan melakukan normalisasi tiap komponen dari inisialisasi vektor GGVFS dengan memperhatikan *magnitudenya*. CNGGVFS menunjukkan kemampuan yang baik dalam meningkatkan konvergensi batas kecekungan (Qin et al. 2013).

Penelitian lain mengusulkan metode baru untuk segmentasi citra sel tunggal smear serviks menggunakan *Radiating Component Normalized Generalized Gradient Vector Flow Snake* (RCNGGVFS) dan dikombinasikan dengan metode *Fuzzy C-Means Non Local Spatial* (FCM-NLS) sebagai metode segmentasi awal (Husain et al. 2017). RCNGGVFS memanfaatkan perhitungan REM dalam pencarian *edge map* pada metode CNGGVFS. Metode CNGGVFS memberikan kemampuan yang baik dalam meningkatkan konvergensi batas kecekungan dan tidak sensitif terhadap inisialisasi *snake*. FCM-NLS mengatasi permasalahan *noise* pada citra smear serviks. Namun metode tersebut ternyata juga belum berhasil secara optimal pada beberapa citra sel serviks dikarenakan pemilihan nilai ambang batas yang kurang tepat pada proses segmentasi awal sehingga menyebabkan beberapa citra tidak dapat tersegmentasi dengan baik. Citra yang tidak dapat tersegmentasi umumnya memiliki warna sitoplasma yang hampir mirip dengan latar belakang, terdapat banyak *noise*, serta memiliki nukleus yang berukuran sangat kecil sehingga sering diabaikan dan dianggap sebagai *noise*. Oleh karena itu dibutuhkan metode segmentasi awal yang handal dalam menangani ukuran dan bentuk kluster data akibat perbedaan ukuran sitoplasma dan nukleus yang kontras, serta yang dapat menangani *noise* yang terdapat pada citra.

Penelitian ini mengusulkan metode baru untuk segmentasi citra sel tunggal smear serviks menggunakan kombinasi metode *mean shift clustering* dan *Radiating Normally Biased Generalized Gradient Vector Flow Snake* (RNBGGVFS). Metode ini menggunakan *Radiating Edge Map* untuk menghitung *edge map* dari citra dengan menggabungkan metode GGVFS dan *Normally Biased Gradient Vector Flow Snake* (NBGVFS) (Wang et al. 2009). Model GGVFS memperbesar jangkauan konvergensi kontur aktif, meningkatkan kinerja konvergensi dan lebih handal terhadap *noise*. Berdasarkan metode NGVF, yang memiliki efisiensi difusi yang lebih tinggi, NBGVF memberikan solusi untuk masalah perlindungan tepi yang tipis. Oleh karena itu, metode ini menggabungkan GGVF dan NBGVF untuk mengusulkan versi energi eksternal yang baru. Metode *mean shift clustering* digunakan sebagai metode segmentasi awal yang akan menghasilkan inisial kontur. *Mean shift clustering* merupakan metode *clustering* berbasis densitas dan *nonparametric* yang dapat menangani bentuk dan ukuran *cluster* data. Metode yang diusulkan terdiri atas 3 tahapan utama, yaitu tahap praproses, citra smear serviks RGB asli dikonversi ke dalam ruang warna CIELAB dan layer L^* dinormalisasi untuk mendapatkan citra keabuan dan dilakukan penghilangan *noise* dengan metode *median filtering*. Kemudian citra smear serviks dibagi ke dalam tiga bagian (nukleus, sitoplasma dan area latar belakang) menggunakan *mean shift clustering*. Selanjutnya tahap terakhir adalah segmentasi kontur. Pada tahap ini diusulkan metode RNBGGVFS sebagai fungsi pengganti energi eksternal *snake* untuk

mendapatkan kontur nukleus dan sitoplasma citra sel tunggal smear serviks. Metode yang diusulkan diharapkan dapat memberikan hasil segmentasi citra sel tunggal smear serviks yang akurat. Dengan segmentasi yang akurat akan dapat meningkatkan performa klasifikasi citra smear serviks.

1.2 Tujuan Penelitian

Penelitian ini bertujuan untuk mengusulkan sebuah metode segmentasi citra sel tunggal smear serviks dengan kombinasi metode *Radiating Normally Biased Generalized Gradient Vector Flow Snake* (RNBGGVFS).

1.3 Perumusan Masalah

Rumusan masalah dalam penelitian ini adalah:

1. Bagaimana melakukan tahap pra proses pada segmentasi citra sel tunggal smear serviks?
2. Bagaimana melakukan segmentasi awal menggunakan metode *mean shift clustering* pada citra sel tunggal smear serviks?
3. Bagaimana melakukan segmentasi akhir untuk mendapatkan kontur nukleus dan sitoplasma pada citra sel tunggal smear serviks menggunakan metode RNBGGVFS?

1.4 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Citra smear serviks diperoleh dari dataset Herlev.
2. Dataset Herlev terdiri dari 917 citra sel tunggal smear serviks.
3. Citra yang digunakan dalam ruang warna RGB dan berformat BMP.

1.5 Manfaat Penelitian

Metode yang diusulkan diharapkan dapat memberikan kontribusi pada bidang informatika khususnya citra biomedik dalam membantu mengetahui area sitoplasma dan nukleus pada citra sel tunggal serviks secara tepat, sehingga dapat memudahkan peneliti pada proses identifikasi kanker serviks secara otomatis.

1.6 Kontribusi Penelitian

Kontribusi pada penelitian ini adalah menggunakan metode *mean shift clustering* sebagai metode segmentasi awal untuk menghasilkan initial kontur yang akan digunakan pada proses segmentasi akhir, serta mengusulkan perhitungan *Radiating Edge Map* dalam pencarian *edge*

map pada kombinasi metode *Generalized Gradient Vector Flow Snake* dan *Normally Biased Gradient Vector Flow Snake* untuk mendapatkan kontur nukleus dan sitoplasma citra sel tunggal smear serviks secara lebih akurat.

(Halaman ini sengaja dikosongkan)

BAB 2

KAJIAN PUSTAKA

2.1 Kanker Serviks

Kanker serviks merupakan kanker ganas yang terbentuk dalam jaringan serviks (organ yang menghubungkan uterus dengan vagina). Ada beberapa tipe kanker serviks. Tipe yang paling umum dikenal adalah *squamous cell carcinoma* (SCC), yang merupakan 80 hingga 85 persen dari seluruh jenis kanker serviks. Infeksi *Human Papilloma Virus* (HPV) merupakan salah satu faktor utama tumbuhnya kanker jenis ini.

Tipe-tipe lain kanker serviks seperti *adenocarcinoma*, *small cell carcinoma*, *adenosquamous*, *adenosarcoma*, *melanoma* dan *lymphoma*, merupakan tipe kanker serviks yang langka yang tidak terkait dengan HPV. Beberapa tipe kanker yang telah disebutkan, tidak dapat ditanggulangi seperti SCC (Parkway Cancer Centre, 2014).

Beberapa vaksinasi telah dikembangkan dan secara efektif membunuh HPV yang menjadi penyebab dari 70 hingga 85 persen kanker serviks. Vaksin HPV ditujukan untuk anak perempuan dan wanita dewasa dari usia 9 hingga 26 tahun karena vaksin hanya dapat bekerja sebelum infeksi terjadi. Akan tetapi, vaksinasi masih dapat dilakukan pada wanita yang belum aktif secara seksual pada usia dewasa. Mahalnya harga vaksin ini menjadi penyebab kekhawatiran. Akan tetapi, karena vaksin ini hanya ditujukan untuk beberapa tipe kanker beresiko tinggi, wanita tetap harus melakukan pap smear, bahkan setelah vaksinasi.

2.2 Pemeriksaan Pap Smear

Pap smear adalah sebuah langkah pengujian medis untuk mendeteksi ada tidaknya gangguan pada leher rahim, biasa berkaitan perihal kanker serviks pada wanita. Adapun wilayah kewanitaan yang diperiksa meliputi sel-sel dari leher rahim hingga panggul.

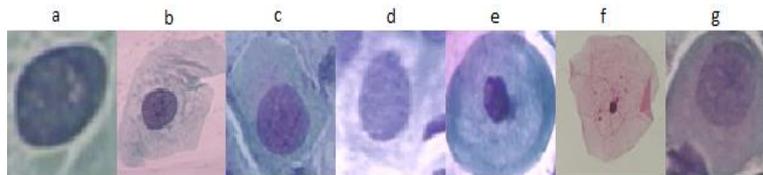
Langkah tes pap smear memberikan fakta medis keberadaan virus papiloma yang notabene adalah virus yang bertanggung jawab menyebabkan kanker serviks. Pap smear memberikan kesempatan untuk melakukan deteksi dini dan mengambil langkah yang dibutuhkan sebelum terlampaui parah. Sel tunggal citra pap smear ini dapat diklasifikasikan ke dalam 7 kelas, yaitu:

1. *Carcinoma in situ*,
2. *Light dysplasia*,
3. *Normal columnar*,
4. *Normal intermediate*,
5. *Normal Superficial*,

6. *Moderate dysplasia*,

7. *Severe dysplasia*.

Sel serviks abnormal yang telah mengalami perubahan pra-kanker disebut sel displastik, yang memiliki tiga fase utama. Fase pertama disebut displasia ringan, di mana inti menjadi lebih besar dan lebih terang dari yang normal. Fase kedua disebut displasia sedang, di mana nukleus menjadi lebih besar dan lebih gelap. Pada fase terakhir yang disebut displasia berat, baik inti dan sitoplasma berubah ukuran dan tekstur, nukleus lebih besar dan lebih gelap dengan bentuk aneh, dan sitoplasma biasanya lebih gelap dan lebih kecil. Oleh karena itu, sel pada kanker dan pra-kanker ditandai oleh banyak perubahan morfologi dan arsitektur, termasuk bentuk dan ukuran sitoplasma dan nukleus, rasio nuklir sitoplasma meningkat, dan lain-lain (Parkway Cancer Centre, 2014). Gambar 2.1 menunjukkan hasil pemeriksaan pap smear berupa gambar atau citra digital.



Gambar 2.1 Citra sel tunggal smear serviks: (a) *carcinoma in situ*, (b) *light dysplastic*, (c) *moderate dysplastic*, (d) *normal columnar*, (e) *normal intermediate*, (f) *normal superficial*, (g) *severe dysplastic* (Jantzen et al. 2006).

2.3 Citra Digital

Citra digital adalah gambar dua dimensi yang dapat ditampilkan pada layar monitor komputer sebagai himpunan berhingga (diskrit) nilai digital yang disebut piksel (*picture elements*). Piksel adalah elemen citra yang memiliki nilai yang menunjukkan intensitas warna. Banyaknya nilai yang dapat digunakan dalam kuantisasi citra bergantung kepada kedalaman piksel, yaitu banyaknya bit yang digunakan untuk merepresentasikan intensitas warna piksel. Kedalaman piksel sering disebut juga kedalaman warna. Citra digital yang memiliki kedalaman piksel n bit disebut juga citra n -bit (C. N. Shanti, 2011). Berdasarkan warna-warna penyusunnya, citra digital dapat dibagi menjadi tiga macam yaitu citra biner, citra keabuan dan citra berwarna. Selain itu juga terdapat pengembangan struktur warna seperti CIELAB, HSV, dan lain sebagainya. Dalam penelitian ini, yang akan digunakan adalah citra biner, citra keabuan, dan citra berwarna serta struktur warna CIELAB.

2.3.1 Citra Biner

Citra biner yaitu citra yang hanya terdiri atas dua warna, yaitu hitam dan putih. Oleh karena itu, setiap piksel pada citra biner cukup direpresentasikan dengan 1 bit. Citra biner hanya mempunyai dua nilai derajat keabuan yakni hitam dan putih (C. N. Shanti, 2011).

2.3.2 Citra Keabuan

Citra keabuan atau *grayscale* yaitu citra yang nilai pikselnya merepresentasikan derajat keabuan. Nilai intensitas paling rendah merepresentasikan warna hitam dan nilai intensitas paling tinggi merepresentasikan warna putih. Pada umumnya citra keabuan memiliki kedalaman piksel 8 bit (256 derajat keabuan), tetapi ada juga citra keabuan yang kedalaman pikselnya bukan 8 bit, misalnya 16 bit untuk penggunaan yang memerlukan ketelitian tinggi (C. N. Shanti, 2011).

2.3.3 Citra Berwarna

Citra berwarna yaitu citra yang nilai pikselnya merepresentasikan warna tertentu. Banyaknya warna yang mungkin digunakan bergantung kepada kedalaman piksel citra yang bersangkutan. Intensitas suatu citra pada citra berwarna merupakan kombinasi dari tiga intensitas yaitu merah, hijau, dan biru (C.N. Shanti, 2011).

2.3.4 Ruang warna CIELAB

CIELab telah banyak digunakan dalam aplikasi teknik dan aplikasi citra warna. Pada tahun 2010, dilakukan penerapan CIELab pada proses segmentasi dan deteksi tepi. CIELab adalah salah satu struktur warna yang didefinisikan CIE. Pada CIELab, besaran CIE_{L^*} untuk mendeskripsikan kecerahan warna, 0 untuk hitam dan $L^* = 100$ untuk putih. Dimensi CIE_{a^*} mendeskripsikan jenis warna hijau–merah, dimana angka negatif a^* mengindikasikan warna hijau dan sebaliknya CIE_{a^*} positif mengindikasikan warna merah. Dimensi CIE_{b^*} untuk jenis warna biru–kuning, dimana angka negatif b^* mengindikasikan warna biru dan sebaliknya CIE_{b^*} positif mengindikasikan warna kuning (G. Hoffmann, 2010).

2.4 Pengolahan Citra Digital

Pengolahan citra merupakan proses mengolah piksel-piksel di dalam citra digital untuk tujuan tertentu. Pada awalnya pengolahan citra dilakukan untuk memperbaiki kualitas citra, namun dengan berkembangnya dunia komputasi yang ditandai dengan semakin meningkatnya kapasitas dan kecepatan proses komputer serta munculnya ilmu-ilmu komputasi yang memungkinkan manusia dapat mengambil informasi dari suatu citra, maka pengolahan citra digital tidak dapat dilepaskan dengan bidang visi komputer (Noercholis et al. 2013).

2.5 Median Filtering

Metode *median filter* merupakan filter *non-linear* yang dikembangkan Tukey, yang berfungsi untuk menghaluskan dan mengurangi *noise* atau gangguan pada citra. Dikatakan *non-linear* karena cara kerja penapis ini tidak termasuk kedalam kategori operasi konvolusi. Operasi *non-linear* dihitung dengan mengurutkan nilai intensitas sekelompok piksel, kemudian menggantikan nilai piksel yang diproses dengan nilai tertentu. Pada *median filter* suatu *window* atau penapis yang memuat sejumlah piksel ganjil digeser titik per titik pada seluruh daerah citra. Nilai-nilai yang berada pada *window* diurutkan secara *ascending* untuk kemudian dihitung nilai *mediannya*. Nilai tersebut akan menggantikan nilai yang berada pada pusat bidang *window*. Jika suatu *window* ditempatkan pada suatu bidang citra, maka nilai piksel pada pusat bidang *window* dapat dihitung dengan mencari nilai *median* dari nilai intensitas sekelompok piksel yang telah diurutkan. Secara umum, dapat dituliskan pada persamaan:

$$f(x, y) = \text{median} \{f(x - i, y - i), (i, j) \in w\} \quad (2.1)$$

dimana $f(x,y)$ merupakan citra hasil *median filtering*, dengan w sebagai *window* yang ditempatkan pada bidang citra dan (i,j) merupakan elemen *window* tersebut (Ratman, 2013).

2.6 Segmentasi Citra

Segmentasi citra adalah proses membagi citra ke dalam area-area atau kategori yang mewakili objek yang berbeda atau bagian dari objek. Setiap piksel pada citra dialokasikan ke dalam satu kategori. Segmentasi yang baik adalah dimana piksel-piksel pada kategori yang sama mempunyai kemiripan tingkat keabuan dan membentuk area yang terhubung.

Segmentasi citra merupakan langkah yang penting dalam analisis citra dan pengenalan pola. Secara umum, ada tiga metode umum yang sering digunakan dalam segmentasi, yaitu metode *thresholding*, metode berbasis tepi, dan metode berbasis *region*.

Metode *thresholding* merupakan metode yang didasarkan pada pemisahan piksel ke dalam kelas yang berbeda tergantung pada tingkat keabuan masing-masing piksel. Metode *thresholding* tidak bisa diterapkan pada citra dengan tingkat keabuan yang berdekatan sehingga biasanya dikombinasikan dengan metode lain.

Pada metode berbasis tepi, tepi merupakan perubahan nilai intensitas derajat keabuan yang mendadak (besar) dalam jarak yang singkat. Tepi dapat diorientasikan dengan suatu arah, dan arah ini berbeda-beda, tergantung pada perubahan intensitas. Penggunaan metode deteksi tepi yang tidak tepat akan menghasilkan pendeteksian yang gagal.

Metode berbasis *region* seperti menggabungkan *thresholding* dengan kondisi konektivitas atau kriteria daerah homogenitas. Keberhasilan dari metode tersebut bergantung pada ketepatan

informasi anatomi untuk meletakkan baik satu maupun beberapa piksel untuk masing-masing daerah homogen. Kelemahan lain dari metode ini adalah metode tersebut hanya dapat bekerja dengan baik pada daerah yang homogen dan membutuhkan operator untuk menentukan daerah yang akan disegmentasi (Glasbey et al. 1995).

2.7 Mean Shift Clustering

Mean shift menganggap ruang fitur sebagai sebuah fungsi empiris probabilitas densitas. Jika data masukan berupa kumpulan titik, maka *mean shift* diasumsikan sebagai sampel dari fungsi kepadatan densitas yang mendasarinya. Jika daerah yang rapat atau kluster hadir di ruang fitur, maka akan didapatkan lokal maksima pada fungsi probabilitas densitas.

Untuk tiap data, *mean shift* menghubungkan dengan puncak terdekat dari fungsi kepadatan probabilitas dataset. *Mean shift* mendefinisikan sebuah window disekitarnya dan menghitung rata-rata titik data, kemudian pusat *window* bergeser ke rata-rata dan mengulangnya hingga konvergen. Tiap iterasi dapat dilihat bahwa *window* bergeser ke area yang lebih rapat dari dataset.

Kernel :

Kernel merupakan sebuah fungsi yang memenuhi persyaratan :

1. $\int R^d \phi(x) = 1.$
2. $\phi(x) \geq 0.$

Beberapa contoh kernel yaitu :

1. Rectangular $\rightarrow \phi(x) = \begin{cases} 1 & a \leq x \leq b \\ 0 & \text{else} \end{cases}$
2. Gaussian $\rightarrow \phi(x) = e^{-\frac{x^2}{2\sigma^2}}.$
3. Epanechnikov $\rightarrow \phi(x) = \begin{cases} \frac{3}{4}(1 - x^2) & \text{if } |x| \leq 1 \\ 0 & \text{else} \end{cases}$

Algoritma *mean shift* adalah metode pengkluster nonparameter yang tidak membutuhkan informasi jumlah kluster serta tidak ada batasan bentuk kluster. *Kernel density estimation* merupakan cara nonparameter yang digunakan untuk mengestimasi fungsi densitas dari sebuah variabel random. Cara ini disebut juga teknik Parzen window. Diberikan n titik data, $x_i, i = 1, \dots, n$ pada ruang dimensi R^d , perkiraan kerapatan kernel multivariat yang diperoleh dengan kernel $k(x)$ dan radius h adalah

$$f(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right). \quad (2.2)$$

Untuk kernel simetris radial, profil kernel $k(x)$

$$K(x) = c_{k,d} k(\|x\|^2), \quad (2.3)$$

dimana $c_{k,d}$ merupakan konstanta normalisasi yang menjamin $K(x)$ terintegrasi 1. Modus fungsi kerapatan terletak pada nol dari fungsi gradien $\nabla f(x) = 0$. Gradien densitas estimatornya adalah

$$\begin{aligned} \nabla f(x) &= \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (x_i - x) g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \\ &= \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \right] \left[\frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \right], \end{aligned} \quad (2.4)$$

dimana $g(s) = -k'(s)$. Persamaan pertama sebanding dengan estimasi kerapatan pada x yang

dihitung dengan kernel $G(x) = c_{g,d} g(\|x\|^2)$ dan persamaan kedua $m_h(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} -$

x adalah *mean shift* (Domanicius et al. 2002). *Mean shift* selalu mengarah ke arah peningkatan maksimum densitas. Prosedur pergeseran rata-rata diperoleh secara berturut-turut menghitung *mean shift* vektor $m_h(x^t)$, dan menggeser *window* $x^{t+1} = x^t + m_h(x^t)$, yang dijamin akan bertemu di titik dimana gradien fungsi kepadatan adalah nol. Algoritma *mean shift* adalah aplikasi praktis dari prosedur pencarian mode yang dimulai dari titik data, kemudian jalankan prosedur *mean shift* untuk menemukan titik stasioner dan fungsi kepadatan. Lalu, pangkas titik-titik ini dengan mempertahankan hanya lokal maksima.

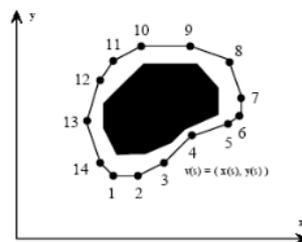
2.8 Active Contour

Active contour mula-mula diperkenalkan oleh Kass et al. dan diberi nama *snakes*. *Active contour* ini dapat bergerak melebar ataupun menyempit dengan cara meminimumkan energi yang terdapat padanya menggunakan tenaga eksternal, serta dipengaruhi juga oleh ciri-ciri suatu gambar seperti garis ataupun sisi (*edge*) (Kass et al., 1988).

Ada dua metode untuk *active contour* yaitu *edge-based* dan *region-based* (Mahabadi, K. Navid, 2008). *Active contour* yang menggunakan *edge-based* adalah metode yang menggunakan deteksi tepi yaitu tergantung pada gradien gambar untuk menghentikan kurva berkembang di perbatasan dari objek yang diinginkan. Sedangkan *region-based* bukan mencari tepi tetapi melakukan deteksi terhadap *foreground* dan *background* dari citra. Proses yang

dilakukan adalah mulai dengan *mask* awal (misalnya persegi *a*) direpresentasikan dalam bentuk kurva tertutup, dan secara iteratif memodifikasi *mask* yang mengakibatkan operasi menyusut atau mengembang sampai dengan objek yang diinginkan.

Istilah *active* pada *active contour models* mengacu pada deformabilitas atau fleksibilitas *snake* untuk memperkecil energi dinamis yang diadaptasi oleh suatu citra. Proses pengekstraksian *boundary* objek diinisialisasi oleh suatu kurva yang diperkirakan dekat dengan batas-batas objek atau *region* yang dituju pada citra. Bentuk dasar *active contour* ditunjukkan pada Gambar 2.2.



Gambar 2.2 Bentuk dasar *active contour* (Kass et al. 1988).

Secara matematis energi *snake* dirumuskan dari tiga jenis energi, yaitu E_{int} (*internal energy*), E_{con} (*external energy*), dan E_{image} (*image force*).

$$E_{snake} = \int E_{snake} (v(s)) ds. \quad (2.5)$$

$$E_{snake} = \int E_{int} (v(s)) ds + \int E_{image} (v(s)) d + \int E_{con} (v(s)) ds. \quad (2.6)$$

Sehubungan dengan spesifikasi algoritma *snake* yang digunakan, hasil akhir dari energi *snake* kemungkinan berbeda sesuai dengan fungsi energi-energi apa saja yang didefinisikan. *Internal energy* dirumuskan dalam fungsi

$$E_{int} = (\alpha(s)|v_s(s)|^2 + (\beta(s)|v_{ss}(s)|^2)/2, \quad (2.7)$$

dimana $\alpha(s)$ merupakan fungsi kesinambungan (*stretching*), $\beta(s)$ merupakan fungsi kelengkungan (*bending*). Fungsi $\alpha(s)$ dan $\beta(s)$ merupakan *internal energy* yang mempengaruhi performansi *snake* seperti halnya selaput dan wadah tipis. *Internal energy* merupakan energi yang akan membuat titik-titik pada kurva deformabilitas *snake* tetap terhubung dengan menciptakan kondisi yang membuat setiap titik selalu terletak berdekatan pada jarak yang relatif sama.

Image forces merupakan energi yang dapat diteruskan untuk melakukan deteksi objek dan menarik kurva deformabilitas *snake* menuju fitur citra yang diinginkan. Pada *image force* digunakan tiga jenis fungsi yaitu *line*, *edge*, dan *termination*.

$$E_{snake} = W_{line}E_{line} + W_{edge}E_{edge} + W_{term}E_{term}. \quad (2.8)$$

Line functional adalah jenis *image force* yang paling sederhana. *Line functional* berpusat pada intensitas (I) citra itu sendiri. *Line functional* dapat dirumuskan sebagai berikut:

$$E_{line} = I(x, y). \quad (2.9)$$

Line functional akan menentukan kemanakah kurva deformabilitas *snake* akan tertarik, jika nilai koefisiennya positif, maka kurva deformabilitas *snake* akan tertarik ke arah garis yang gelap. Jika koefisien bernilai negatif, maka kurva deformabilitas *snake* akan tertarik pada garis yang terang. *Edge functional* digunakan untuk menarik kurva deformabilitas *snake* ke arah sebuah sudut. *Edge functional* dapat dirumuskan sebagai berikut:

$$E_{edge} = -|\Delta I(x, y)|^2. \quad (2.10)$$

Efek dari *edge functional* adalah untuk membuat kurva deformabilitas *snake* tertarik pada sebuah kontur yang memiliki gradien yang tinggi (misal dari hitam [0] ke putih [255]) semakin tinggi kontras dari gradien tersebut, maka daya tarik *edge functional* akan semakin besar. *Termination functional* digunakan untuk mencari titik akhir dari sebuah segmen garis dan sudut. Dengan mengkombinasikan kedua fungsi *line functional* dan *edge functional* dengan *termination functional*, *snake* akan mampu bergerak menuju fitur citra yang diinginkan dengan akurasi yang tinggi.

External energy membatasi deformabilitas kurva *snake*, batasan yang diberikan oleh *external energy* dapat berbeda-beda. Pada hakekatnya biasanya *external energy* hampir tidak pernah digunakan. Setiap bentuk *external energy* memiliki tujuan masing-masing.

Agar *snake* menjadi dinamis, x dibuat sebagai fungsi waktu sebagai berikut:

$$x_t(s, t) = \alpha x''(s, t) - \beta x''''(s, t) - \kappa \nabla E_{ext}, \quad (2.11)$$

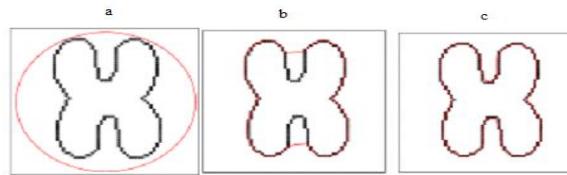
dimana α dan β adalah parameter pembobotan yang mengontrol tegangan dan kekakuan *snake*, x'' dan x'''' adalah turunan kedua dan keempat dari x terhadap s , κ merupakan *weighting parameter* dari energy eksternal, sedangkan $-\nabla E_{ext}$ adalah gaya eksternal yang didapatkan dari citra. Apabila nilai $x=0$, berarti *snake* telah minimum, maka solusi telah ditemukan (Kass et al. 1998).

2.8.1 Gradient Vector Flow Snake

Snake atau *active contour* adalah kurva yang didefinisikan dalam domain citra yang bergerak karena pengaruh gaya internal dari kurva itu sendiri dan gaya eksternal yang dihitung

dari data citra. Gaya internal dan eksternal dibuat sedemikian rupa sehingga *snake* akan menuju ke batas suatu objek atau fitur lain yang diinginkan. *Snake* digunakan secara luas dalam banyak aplikasi termasuk deteksi tepi, pemodelan bentuk, segmentasi, dan *motion tracking* (Xu et al. 2002).

Kelemahan *snake* adalah kontur awal harus ditempatkan dekat dengan objek dan menunjukkan konvergensi yang lemah terhadap batas kecekungan objek. Sehingga dikembangkan menjadi *Gradient Vector Flow (GVF) Snake* yang mengombinasikan aliran *optic* dan gradien keabuan. Perbandingan algoritma *snake* dan *GVF snake* ditunjukkan melalui contoh Gambar 2.3.



Gambar 2.3 Contoh *active contour*: (a) Inisial kontur; (b) Hasil dengan algoritma *snake*; (c) Hasil dengan algoritma *GVF snake* (Xu et al. 2002).

Untuk mendapatkan *GVF*, langkah awal adalah dengan mengekstraksi fungsi *edge map* $f(x, y)$ dari citra $I(x, y)$ untuk citra keabuan dengan persamaan:

$$f^{(3)}(x, y) = -|\nabla I(x, y)|^2, \quad (2.12)$$

dan

$$f^{(4)}(x, y) = -|\nabla[G_\sigma(x, y) * I(x, y)]|^2, \quad (2.13)$$

dimana $G_\sigma(x, y)$ adalah fungsi *Gaussian* dua dimensi dengan standart deviasi σ , dan ∇ adalah operator gradien. $\vec{v}(x, y) = (u(x, y), v(x, y))$ didefinisikan sebagai solusi yang menggantikan energi eksternal dengan *GVF* untuk meminimalkan fungsi energi:

$$E_{GVF} = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 \vec{v} - \nabla f|^2 dx dy, \quad (2.14)$$

dimana μ adalah bobot yang mengontrol *relative importance* pada term pertama dan kedua. Dan ∇f adalah *gradient edge map*. Contoh *edge map* suatu citra ditunjukkan pada Gambar 2.4. *Edge map* dengan normalisasi *GVF* ditunjukkan pada Gambar 2.5.



Gambar 2.4 *Edge map* (Xu et al. 2002).

Untuk mendapatkan nilai \rightarrow dengan menyelesaikan 2 persamaan *euler* berikut:

$$\mu \nabla^2 u - (u - f_x)(f_x^2 - f_y^2) = 0, \quad (2.15)$$

dan

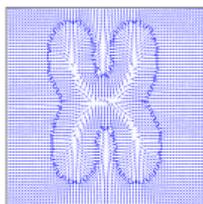
$$\mu \nabla^2 v - (v - f_x)(f_x^2 - f_y^2) = 0, \quad (2.16)$$

dimana ∇^2 adalah operator *Laplacian*. Kedua persamaan dapat diselesaikan dengan menerapkan u dan v sebagai fungsi waktu t dan menyelesaikan persamaan generalisasi difusi selanjutnya untuk $t \rightarrow \infty$.

$$U_t(x, y, t) = \mu \nabla^2(x, y, t) - (u(x, y, t) - f_x(x, y))(f_x^2(x, y) - f_y^2(x, y)) = 0, \quad (2.17)$$

dan

$$v_t(x, y, t) = \mu \nabla^2(x, y, t) - (v(x, y, t) - f_x(x, y))(f_x^2(x, y) - f_y^2(x, y)) = 0. \quad (2.18)$$



Gambar 2.5 Dengan normalisasi GVF (Xu et al. 2002).

2.8.2 Generalized Gradient Vector Flow Snake

GVFS memiliki konvergensi yang lemah dan sensitif terhadap inisialisasi *snake* (Qin et al. 2013). Kelemahan ini disebabkan karena *smoothing* yang berlebihan pada area *boundary*, yang diatur oleh koefisien μ pada Persamaan (2.14). Untuk mengatasi masalah ini, ditambahkan dua koefisien pembobot pada persamaan gaya eksternal GVFS (Qin et al. 2013). Persamaan gaya eksternal baru tersebut dinamakan *Generalized*

Gradient Vector Flow Snake (GGVFS). Fungsi turunan energi dari GGVFS seperti persamaan berikut

$$v_t(x, y, t) = \underbrace{g(|\nabla f| \nabla^2 v(x, y, t))}_{\text{smoothness energy}} + \underbrace{h(|\nabla f|) |v(x, y, t) - \nabla f|}_{\text{edge energy}}, \quad (2.19)$$

dimana,

$$g(|\nabla f|) = e^{-|\nabla f|/k}, \quad (2.20)$$

$$h(|\nabla f|) = 1 - g(|\nabla f|), \quad (2.21)$$

term pertama pada bagian kanan dari persamaan (2.19) merupakan term *smoothing*, yang akan menghasilkan vektor. Term kedua adalah term data atau disebut juga *edge energy*, yang akan membawa vektor v dekat dengan *gradient edge map* (∇f) dari citra. Parameter k menentukan nilai koefisien pembobot untuk term *smoothing* dan data. Nilai dari k berhubungan dengan kekuatan *noise* dari sebuah citra. Semakin tinggi level *noise*, maka semakin banyak nilai k (Qin et al. 2013).

Pada persamaan (2.19), $|\nabla f|$ dapat dihitung menggunakan operator gradien apapun. Jika term data pada persamaan mendefinisikan vektor v dekat dengan *gradient edge map* (∇f) dari citra, maka koefisien pembobot $h(|\nabla f|)$ harus bertambah. Berdasarkan persamaan, parameter k harus > 0 . Untuk mengatur solusi perulangan, anggap Δx dan Δy merepresentasikan ruang antar piksel, kemudian Δt adalah notasi waktu untuk tiap iterasi. Pada *grid diskrit*, operator *Laplacian* ∇^2 menjadi:

$$\nabla^2 v(x, y, t) = \frac{1}{\nabla x \nabla y} A * v(x, y, t), \quad (2.22)$$

dimana $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ adalah matriks *Laplacian*. Substitusi persamaan diatas ke persamaan

(2.19), memberikan solusi perulangan kepada vector GGVFS menjadi:

$$v(x, y, t + \Delta t) = v(x, y, t) + \frac{1}{4} g(|\nabla f|) A * v(x, y, t) - \Delta t h(|\nabla f|) [v(x, y, t) - \nabla f]. \quad (2.23)$$

Untuk mendapatkan *smoothing* yang baik dari vector GGVFS, ditetapkan $\Delta t = 1/4$.

2.8.3 Normally Biased Gradient Vector Flow Snake

Telah diketahui bahwa operator *Laplacian* dapat didekomposisi menjadi dua persamaan yaitu

$$\Delta u = u_{TT} + u_{NN}, \quad (2.24)$$

dimana u_{TT} dan u_{NN} adalah turunan kedua dari $u(x,y)$ pada arah normal dan tangensial. Sebagai operator interpolasi, u_{NN} memiliki performa terbaik, Δu memiliki performa kedua terbaik, dan yang terakhir adalah u_{TT} . Dengan mengambil proses difusi sebagai proses interpolasi, *Normal Gradient Vector Flow* (NGVF) diusulkan menggunakan interpolator terbagi sebagai berikut.

$$\begin{cases} u_t = \mu u_{NN} - (u - f_x)|\nabla f|^2 = 0 \\ v_t = \mu v_{NN} - (v - f_y)|\nabla f|^2 = 0 \end{cases} \quad (2.25)$$

Dekomposisi operator Laplacian telah diteliti pada konteks persamaan difusi. Difusi pada arah tangensial bertujuan untuk perlindungan *edge* yang tipis serta untuk mengurangi *noise*, sedangkan difusi pada arah normal bermaksud untuk memperhalus *edge*. Jika dilihat pada NGVF, difusi pada arah tangensial cenderung diabaikan dan hanya pada arah normal yang diperhatikan. Sebagai hasilnya, NGVF akan sensitif terhadap *noise* dan dapat menghapus *edge* yang tipis.

Kemudian diusulkan *Normally Biased Gradient Vector Flow* (NBGVF) sebagai energi eksternal yang akan tetap memperhatikan difusi arah tangensial dan dibiaskan pada arah normal. Bobot pembiasan merupakan pengurangan fungsi magnitude *edge map* pada citra. Persamaan difusi pada NBGVF adalah sebagai berikut.

$$\begin{cases} u_t = \mu(u_{TT} + g(|\nabla f|u_{NN}) - (u - f_x)|\nabla f|^2 = 0 \\ v_t = \mu(v_{TT} + g(|\nabla f|v_{NN}) - (v - f_y)|\nabla f|^2 = 0 \end{cases} \quad (2.26)$$

dimana μ merupakan bobot positif seperti pada persamaan (2.14), $g(\cdot)$ merupakan bobot bias, yaitu fungsi pengurangan dari ∇f . Jika ∇f besar, pada tepi citra, maka $g(\cdot) = 0$, pada kasus difusi arah normal menghilang dan pada arah tangensial dipertahankan, ini akan sangat membantu untuk mempertahankan *edge* yang tipis serta mengurangi *noise*. Jika $\nabla f=0$ seperti pada area homogen, $g(\cdot) = 1$, oleh karena itu NBGVF akan diubah ke GVF dimana dalam hal ini akan sama efektifnya dengan GVF dengan memperbesar jangkauan gambar. Maka $g(\cdot)$ didefinisikan sebagai berikut

$$g(\nabla f) = \exp\left(-\frac{|\nabla f|^2}{k^2}\right), \quad (2.27)$$

dimana k merupakan konstanta yang menentukan kontras *edge* untuk dipertahankan dan harus ditentukan untuk aplikasi tertentu. Jika *edge* yang harus dipertahankan lemah, nilai k harus kecil, dan sebaliknya (Wang et al. 2009). Sehingga persamaan energinya menjadi

$$v(x, y, t) = \mu(V_{NN}(x, y, t) + V_{TT}(x, y, t)) - \nabla f^2[V(x, y, t) - \nabla f]. \quad (2.28)$$

2.8.4 Normally Biased Generalized Gradient Vector Flow Snake

Model GGVSF memperbesar jangkauan konvergensi kontur aktif, meningkatkan kinerja konvergensi dan lebih handal terhadap *noise*. Berdasarkan metode NGVF, yang memiliki efisiensi difusi yang lebih tinggi, NBGVF memberikan solusi untuk masalah perlindungan *edge* yang tipis. Oleh karena itu, metode ini menggabungkan GGVF dan NBGVF untuk mengusulkan versi energi eksternal yang baru. Dan didefinisikan sebagai berikut:

$$E(V) = \iint g(x, y)(g_s(x, y)V_{NN} + h_s(x, y)V_{TT})dxdy + h(x, y)(v - \nabla f)dxdy. \quad (2.29)$$

dimana

$$g(|\nabla f|) = e^{-|\nabla f|/K}, \quad (2.30)$$

$$h(|\nabla f|) = 1 - e^{-\frac{|\nabla f|}{K}}, \quad (2.31)$$

$$h_s(f) = \begin{cases} 1 & (|e| \geq \tau) \\ \frac{-f^3}{8\tau^3} + \frac{5f}{8\tau} + \frac{1}{2} & (0 < |e| < \tau), \\ 0 & (|e| = 0) \end{cases} \quad (2.32)$$

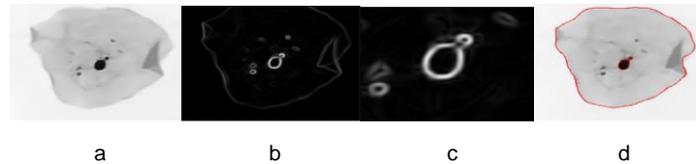
$$g_s(f) = 1 - h_s(f).$$

V_{NN} dan V_{TT} merupakan turunan kedua sepanjang arah normal dan tangensial. Sedangkan $g(|\nabla f|)$ dan $h(|\nabla f|)$ merupakan koefisien *smoothing* dan *data terms*. Seperti yang didefinisikan pada GGVF, nilai k bertambah dengan intensitas *noise* gambar, tetapi hal ini dapat menyebabkan tepi yang tipis menjadi terlalu halus. Parameter h_s dan g_s merupakan derajat difusi gambar pada arah normal dan tangensial. Tidak seperti koefisien operator difusi normal dan tangensial pada NBGVF, kedua koefisien secara langsung bergantung pada intensitas dibanding dengan gradien *edge*, sehingga mengurangi kompleksitas komputasi (Zhang et al. 2016). Sehingga, persamaan evolusi energi eksternal dapat dituliskan sebagai:

$$v(x, y, t) = g(|\nabla f|)(g_s(f)V_{NN}(x, y, t) + h_s(f)V_{TT}(x, y, t)) - h(|\nabla f|)[V(x, y, t) - \nabla f]. \quad (2.33)$$

2.9 Radiating Edge Map

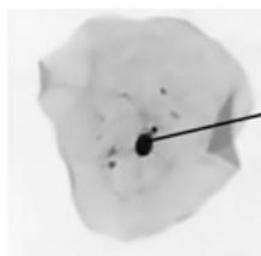
Edge map yang dihitung dengan menggunakan operator deteksi tepi seperti *sobel*, ternyata tidak dapat menggunakan informasi direksional sepenuhnya. Dampak terkuat terjadi pada area gelap di dekat nukleus yang sebenarnya seperti yang ditunjukkan pada Gambar 2.6.



Gambar 2.6 Citra smear serviks; (a) citra hasil preproses, (b) *edge map* dengan *sobel*, (c) daerah nukleus yang sebenarnya tidak terdeteksi dengan benar, (d) citra yang disegmentasi dengan *GVF snake* (Li et al. 2012).

Ketika penggunaan *GVF, snake* membentuk kontur nukleus seperti yang diinginkan. Setelah diselidiki, ditemukan bahwa area nukleus umumnya lebih gelap dari area sitoplasma, dan area sitoplasma lebih gelap dari area *background*. Berdasarkan fakta tersebut, dikembangkan metode untuk menghitung *edge map* pada *radiating line*, dimulai dari *centroid* pada nukleus yang dilakukan pembobotan intensitas. Gambar 2.7 menunjukkan ilustrasi *radiating line*. $I(x, y)$ adalah nilai piksel pada (x, y) dan (x_c, y_c) adalah bobot intensitas *centroid*. Tiap piksel (x_b, y_b) pada tepi citra, dapat ditarik garis L_{x_b, y_b} dari titik x_c, y_c ke x_b, y_b yang panjangnya

$$len_b = \lfloor \sqrt{(x_b - x_c)^2 + (y_b - y_c)^2} \rfloor. \quad (2.34)$$



Gambar 2.7 Ilustrasi *radiating line* (Li et al. 2012).

Koordinat titik tidak selalu integer, sehingga nilai $I_0(x, y)$ dihitung dengan interpolasi linear berdasarkan piksel citra. *Radiating difference* $rd(x_i, y_i)$ pada titik (x_i, y_i) sepanjang lx_b, y_b dapat dihitung dengan

$$rd(x_i, y_i) = I_0(x_{i-1}, y_{i-1}) - I_0(x_i, y_i). \quad (2.35)$$

Radiating gradient $rg(x_i, y_i)$ pada titik (x_i, y_i) dapat dihitung dengan

$$rg(x_i, y_i) = \frac{|rd(x_i, y_i)| + |rd(x_{i+1}, y_{i+1})|}{2}. \quad (2.36)$$

Biasanya, intensitas pada inti memiliki nilai terkecil dan nilai intensitas piksel akan semakin besar jika semakin jauh dari *centroid* di sepanjang *radiating line*. Oleh karena itu, *positive radiating difference* harus ditekan dengan fungsi:

$$F_s(x) = \begin{cases} x, & x < 0 \\ \gamma x, & x \geq 0 \end{cases} \quad (2.37)$$

dimana γ adalah bobot < 1 . Persamaan perhitungan $rg(x_i, y_i)$ menjadi

$$rg(x_i, y_i) = \frac{|Fs(rd(x_i, y_i))| + |Fs(rd(x_{i+1}, y_{i+1}))|}{2}. \quad (2.38)$$

Karena distribusi yang kompleks pada citra smear serviks, seperti ada bercak darah, kontaminasi dan interferensi sel tubuh lain, maka hanya dengan menekan nilai *radiating difference* yang positif saja tidak cukup. Gradien didekat interferensi objek biasanya lebih besar atau hampir sama dengan tepi objek yang sebenarnya, sehingga RGVF bisa saja mendeteksi posisi yang salah (Li et al. 2012). Untuk menghindari hal tersebut, dilakukan perbaikan pada *radiating difference* yang negatif dengan langkah-langkah sebagai berikut.

1. Memperkecil perubahan *radiating difference* (rd) dengan membatasi nilai rd ,
2. Membagi *radiating difference* menjadi beberapa segmen. Tiap segmen garis dapat bertanda positif dan negatif,
3. Perbaikan negatif *radiating difference*.

Dengan mengganti f pada persamaan dengan *Radiating Edge Map* (REM) yang telah diperoleh melalui proses diatas, *Radiating Normally Biased Generalized Gradient Vector Flow* citra didapatkan dengan meminimasi fungsi

$$E_{RNBBGVF} = \iint g(x, y)(gs(x, y)V_{NN} + hs(x, y)V_{TT})dxdy + h(x, y)(v - \nabla REM)dxdy. \quad (2.39)$$

(Halaman ini sengaja dikosongkan)

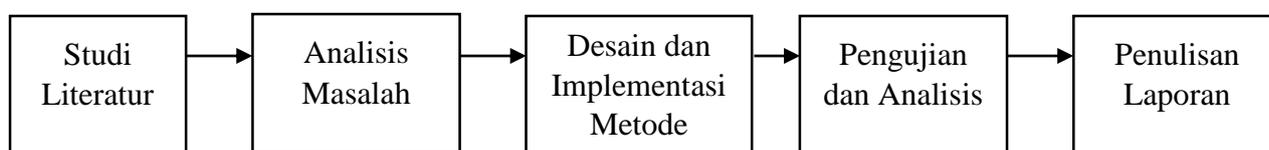
BAB 3

METODE PENELITIAN

Dalam bab ini akan dijelaskan penelitian, data input, rancangan metode, skenario uji coba, evaluasi hasil segmentasi dan jadwal penelitian yang akan digunakan.

3.1 Tahapan Penelitian

Tahapan-tahapan yang dilakukan pada penelitian ini antara lain studi literatur, analisis dan desain, implementasi metode, pengujian dan analisis, serta penulisan laporan penelitian. Secara lebih detail, penelitian ini dirancang dengan urutan tahapan seperti pada Gambar 3.1.



Gambar 3.1 Tahapan penelitian

3.1.1 Studi Literatur

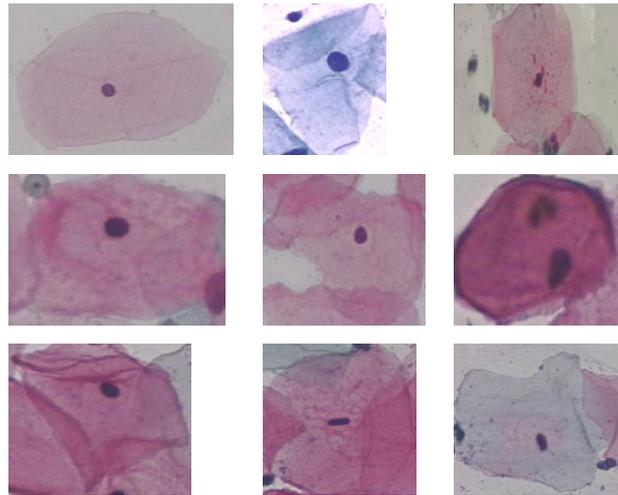
Pada tahap ini akan dipelajari tentang segala informasi dan sumber pustaka yang sesuai dengan konteks penelitian yang dilakukan. Dalam penelitian ini, literatur yang dikaji secara garis besar meliputi konsep-konsep dasar yang berkaitan dengan segmentasi citra sel tunggal smear serviks, metode penghilangan *noise* citra, metode pengklasteran pada segmentasi citra menggunakan algoritma *mean shift clustering*, metode segmentasi kontur menggunakan metode *snake* seperti GVF, GGVFS, NBGVFS dan penggabungannya, serta perhitungan *edge map* menggunakan metode REM.

3.1.2 Analisis Masalah

Tahap analisis ini dimaksudkan sebagai tahapan untuk melakukan analisis terhadap hasil penelitian-penelitian sejenis yang telah dilakukan untuk dapat dipakai sebagai referensi ataupun pembandingan terhadap metode yang diusulkan.

Pada penelitian sebelumnya, terdapat permasalahan, yakni beberapa citra tidak dapat tersegmentasi dengan baik. Citra tersebut kebanyakan berasal dari kelas *normal superficial*. Karakteristik citra pada kelas tersebut umumnya memiliki ukuran nukleus yang sangat kecil, juga objek yang warnanya hampir sama dengan area latar belakang, serta umumnya memiliki banyak *noise*. Citra yang memiliki ukuran nukleus sangat kecil tersebut, ketika telah melewati proses segmentasi awal, dengan menggunakan parameter nilai ambang batas yang digunakan

untuk mendapatkan kontur awal, nukleus tersebut dianggap sebagai *noise* dan diabaikan, sehingga nukleus citra tersebut tidak dapat dideteksi kontur awalnya dan proses selanjutnya tidak dapat dilanjutkan. Adanya perbedaan ukuran sitoplasma dan nukleus yang cukup kontras membutuhkan metode awal yang dapat menangani bentuk dan ukuran kluster data. Gambar 3.2 berikut adalah beberapa contoh gambar yang tidak dapat disegmentasi pada penelitian sebelumnya.



Gambar 3.2 Contoh citra yang belum berhasil disegmentasi.

3.1.3 Desain dan Implementasi Metode

Tahap desain sistem, pada bagian ini akan dipaparkan mengenai format data masukan dan keluaran serta penggambaran alur proses yang terjadi dalam metode untuk menghasilkan data keluaran. Penentuan desain model sistem ini akan memberikan gambaran mengenai apa dan bagaimana suatu penelitian dilaksanakan.

Tahap implementasi metode bertujuan untuk mengimplementasikan rancangan metode yang diusulkan. Pada penelitian ini rancangan metode yang diusulkan akan diimplementasikan menggunakan aplikasi Matlab versi 2014a.

3.1.4 Pengujian dan Analisis

Tahapan ini akan dilakukan pengujian dan analisis terhadap hasil dan performa metode yang diusulkan dalam melakukan segmentasi citra sel tunggal smear serviks.

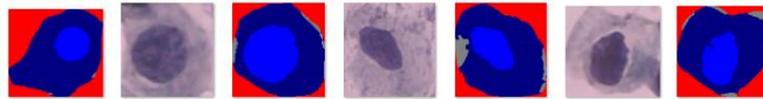
3.1.5 Penulisan Laporan

Tahap ini dilakukan penyusunan laporan terhadap penelitian yang telah dilakukan, mulai dari studi literatur, analisis masalah dan desain, hasil implementasi, hingga uji coba dan analisis.

3.2 Data Masukan

Data masukan dari penelitian ini berupa citra mikroskopis sel tunggal smear serviks. Dataset yang digunakan adalah dataset Herlev yang terdiri dari 917 citra sel tunggal smear serviks yang diperoleh dari Departemen Patologi Rumah Sakit Universitas Herlev dan Departemen Automasi Universitas Teknik Denmark. Dataset citra sel tunggal smear serviks dapat diakses dan diunduh melalui (<http://fuzzy.iau.dtu.dk/download/smear2005/>).

Dataset citra smear serviks dikumpulkan oleh ahli *cytology* menggunakan mikroskop yang terhubung ke kamera digital. Dataset citra smear serviks Herlev juga terdapat citra segmentasi manual (*ground truth*). Contoh citra smear serviks beserta segmentasi manualnya terdapat pada Gambar 3.3. Segmentasi manual citra smear serviks didapat dengan menggunakan perangkat lunak CHAMP dan tersegmentasi menjadi 3 bagian, yaitu sitoplasma, nukleus dan area latar belakang (Jantzen et al. 2006). Ahli *cytology* secara manual mengklasifikasikan dataset citra sel tunggal smear serviks menjadi 7 kelas, distribusi kelas citra sel tunggal smear serviks dapat dilihat pada Tabel 3.1.



Gambar 3.3 Contoh citra sel tunggal smear serviks beserta citra segmentasi manual (Jantzen et al. 2006).

3.3 Rancangan Metode

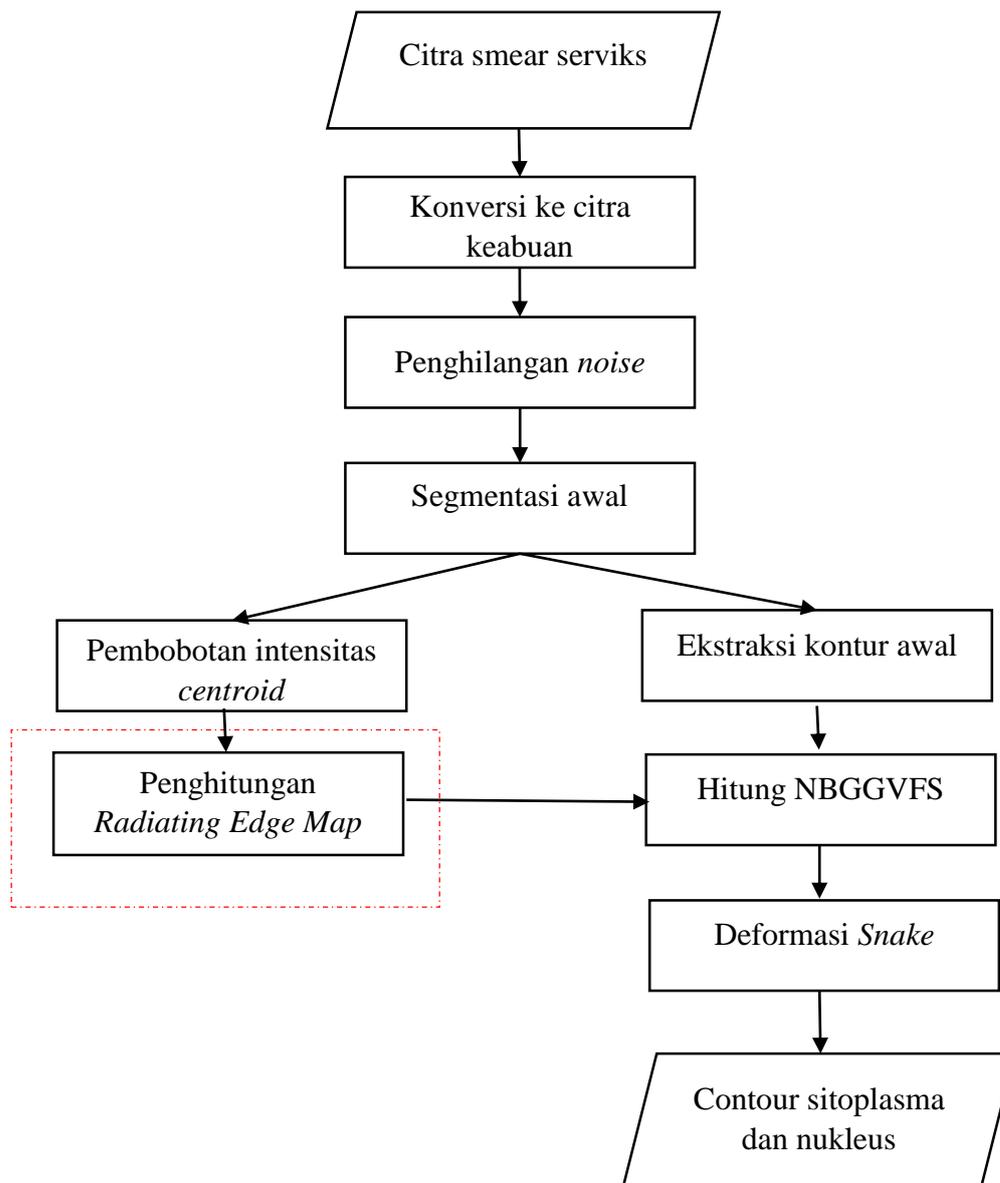
Tujuan penelitian ini adalah untuk melakukan segmentasi nukleus dan sitoplasma pada citra sel tunggal smear serviks menggunakan metode *Radiating Normally Biased Generalized Gradient Vector Flow Snake* (RNBGGVFS). Adapun rancangan metode yang diusulkan dan bagian kontribusi ditampilkan pada Gambar 3.4. Metode dalam penelitian ini terdiri atas tiga tahapan utama yaitu:

- Tahap 1: Pra proses.
- Tahap 2: Segmentasi awal, dan
- Tahap 3: Segmentasi kontur.

3.3.1 Pra Proses

Dalam citra smear serviks, terdapat kontras yang rendah, pewarnaan yang tidak seragam serta *noise* yang seringkali akan menghambat proses segmentasi. Diperlukan adanya perbaikan citra untuk meningkatkan akurasi proses segmentasi citra. Langkah utama yang dilakukan seperti yang terdapat pada diagram alir Gambar 3.5 adalah citra smear serviks (RGB) dikonversi kedalam ruang warna CIELAB. Ruang warna ini terdiri dari tiga dimensi, L^*

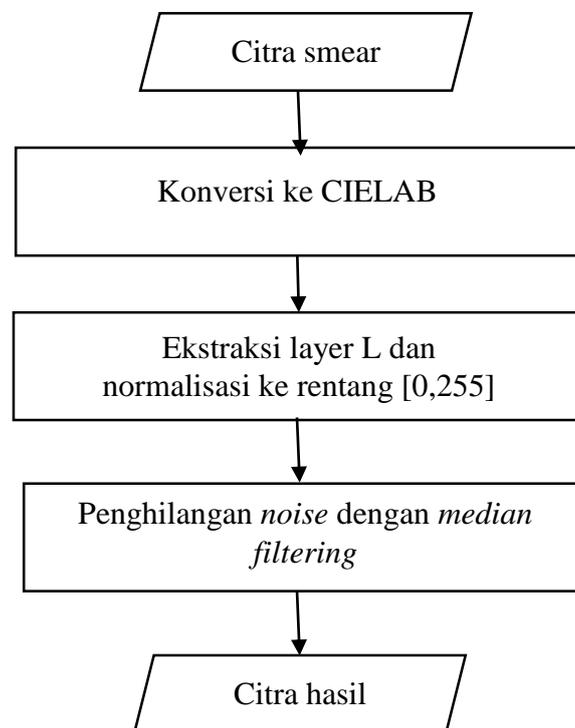
dimensi pencahayaan, a^* dan b^* dimensi warna. CIELAB berhasil memisahkan kekontrasan dari warnanya. Setelah kekontrasan dan warna penyusun citra telah terpisah, tahap selanjutnya adalah mengekstraksi dimensi L^* dan menormalisasi ke rentang $[0,255]$ untuk membentuk citra keabuan. Kemudian yang dilakukan selanjutnya adalah menghilangkan *noise* yang dapat menghambat proses segmentasi dengan metode *median filtering*. Metode tersebut dikenal efektif dalam melakukan *filtering* penghilangan *noise Gaussian* dan *impulse* pada citra smear serviks.



Gambar 3.4 Desain Sistem Usulan

Tabel 3.1 Distribusi 917 sel tunggal citra smear *serviks* dalam dataset Herlev (Jantzen et al. 2006).

| Kelas | Kategori | Tipe Sel | Jumlah Sel | Total |
|-------|--------------|---|------------|----------------------|
| 1. | Normal | <i>Superficial squamous epithelial</i> | 74 | 242 sel normal |
| 2. | Normal | <i>Intermediate squamous epithelial</i> | 70 | |
| 3. | Normal | <i>Columnar epithelial</i> | 98 | |
| 4. | Tidak Normal | <i>Mild squamous nonkeratinizing dysplasia</i> | 182 | 675 sel tidak normal |
| 5. | Tidak Normal | <i>Moderate squamous nonkeratinizing dysplasia</i> | 146 | |
| 6. | Tidak Normal | <i>Severe squamous nonkeratinizing dysplasia</i> | 197 | |
| 7. | Tidak Normal | <i>Squamous cell carcinoma in situ intermediate</i> | 150 | |



Gambar 3.5 Diagram alir pra proses citra smear serviks.

3.3.2 Segmentasi Awal

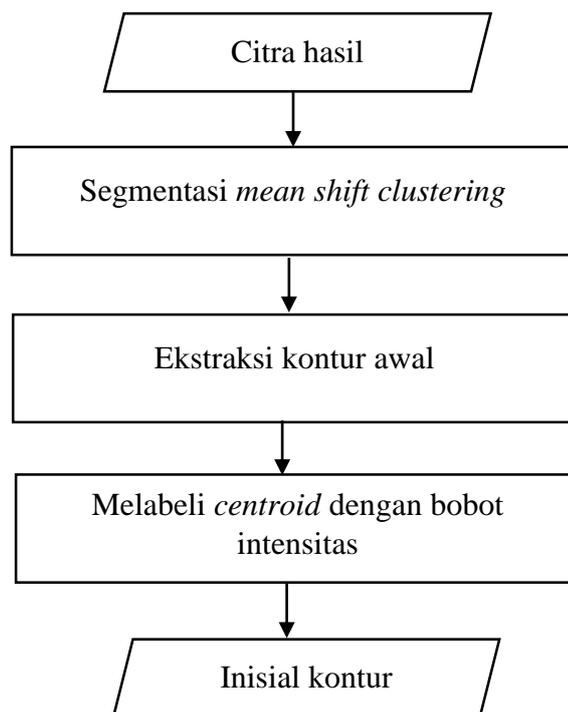
Setelah dilakukan tahap praproses, tahap selanjutnya adalah segmentasi awal dengan menggunakan metode *mean shift clustering*. Citra smear serviks dibagi ke dalam tiga bagian (nukleus, sitoplasma dan area latar belakang). Diagram alir proses segmentasi awal dapat dilihat pada Gambar 3.6.

3.3.2.1 Segmentasi *mean shift clustering*

Tahap pertama segmentasi awal adalah melakukan segmentasi citra hasil tahap praproses. Citra hasil praproses disegmentasi menggunakan metode *mean shift clustering* untuk memisahkan citra tersebut menjadi 3 bagian, nukleus, sitoplasma dan area latar belakang.

Hasil dari segmentasi awal selanjutnya akan digunakan sebagai inisial kontur tahap segmentasi akhir. Pertama-tama menentukan parameter *bandwidth* dan *threshold* konvergensi, selanjutnya memilih titik secara acak sebagai titik inisial, kemudian menetapkan nilai titik tersebut sebagai *mean* awal. Setelah itu melakukan penghitungan jarak antar piksel dan membandingkan dengan *bandwidth* (radius). Titik dengan jarak yang lebih kecil dari *bandwidth* selanjutnya digabung menjadi klaster yang sama dan dihitung *meannya*. Atau tahapan lebih jelasnya sebagai berikut.

1. Pilih titik data sebagai inisialisasi,
2. Letakkan kernel/*window* dengan radius tertentu di sekitar titik yang diinginkan,
3. Hitung nilai *mean*,
4. Pindahkan pusat kernel ke titik tengah,
5. Ulangi 3 & 4 hingga konvergen. Setiap iterasi akan bergerak menanjak pada gradien densitas distribusi data hingga mencapai puncak bukit (maksimum lokal),
6. Ulangi 1-5 untuk semua poin data. Poin yang menyatu dengan maksimum lokal yang sama adalah anggota dari kelompok yang sama. Jumlah klaster adalah jumlah maksima lokal.



Gambar 3.6 Diagram alir proses segmentasi awal.

3.3.2.2 Ekstraksi Kontur Awal

Setelah didapatkan citra yang tersegmentasi, berikutnya dilakukan ekstraksi kontur awal dengan asumsi bahwa area dengan kecekungan terluas disebut sitoplasma. Kandidat nukleus yang dapat memiliki beberapa kemungkinan diasumsikan dengan memastikan bahwa

nukleus merupakan area yang terdapat di dalam sitoplasma dan kandidat area dengan nilai ambang batas kurang dari nilai ambang batas $\beta\%$ dari citra dianggap sebagai *noise* dan diabaikan.

3.3.2.3 Pelabelan centroid dengan bobot intensitas

Kemudian dilakukan pelabelan terhadap *centroid* dengan bobot intensitas, proses ini penting untuk membentuk *Radiating Edge Map* pada tahap segmentasi. Anggap area nukleus terdiri dari k titik $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$. Dan $I(x, y)$ adalah intensitas titik (x, y) . Maka bobot intensitas *centroid* (x_c, y_c) dapat dilihat pada Persamaan (3.1).

$$\begin{cases} x_c = \frac{\sum_{i=1}^k x_i(255-I(x_i, y_i))}{\sum_{i=1}^k (255-I(x_i, y_i))} \\ y_c = \frac{\sum_{i=1}^k y_i(255-I(x_i, y_i))}{\sum_{i=1}^k (255-I(x_i, y_i))} \end{cases} \quad (3.1)$$

3.3.3 Segmentasi Kontur

Pada tahap sebelumnya telah didapatkan citra hasil segmentasi awal dan *centroid* yang telah dilabeli dengan bobot intensitas, selanjutnya melakukan segmentasi kontur dengan menggunakan metode *Radiating Normally Biased Generalized Gradient Vector Flow Snake* (RNBGGVFS). Metode perhitungan REM yang dimanfaatkan pada metode GVFS belum berhasil untuk mendapatkan kontur sitoplasma dan nukleus secara akurat. Hal ini disebabkan karena GVFS memiliki konvergensi yang lemah dan sensitif terhadap inisialisasi *snake* (Li et al. 2012). Maka dari itu diusulkan metode RNBGGVFS, metode RNBGGVFS merupakan metode yang memanfaatkan perhitungan REM dalam pencarian *edge map* pada kombinasi metode *Generalized Gradient Vector Flow Snake* (GGVFS) dan *Normally Biased Gradient Vector Flow Snake* (NBGVFS). Model GGVF memperbesar jangkauan konvergensi kontur aktif, meningkatkan kinerja konvergensi dan lebih handal terhadap *noise*. Berdasarkan metode NGVF, yang memiliki efisiensi difusi yang lebih tinggi, NBGVF memberikan solusi untuk masalah perlindungan tepi yang tipis. Oleh karena itu, metode ini menggabungkan GGVF dan NBGVF untuk mengusulkan versi energi eksternal yang baru. Dengan menggunakan perhitungan REM sebagai pengganti perhitungan konvensional *edge map* pada metode NBBGVFS akan dapat memberikan hasil segmentasi kontur yang akurat pada citra sel tunggal smear serviks. Proses segmentasi kontur dengan metode RNBGGVFS dapat dilihat pada Gambar 3.7. Citra hasil segmentasi awal dijadikan masukan pada proses

segmentasi kontur dengan metode RNBGGVFS untuk mendapatkan kontur sitoplasma dan nukleus.

3.3.3.1 Perhitungan Radiating Edge Map

Radiating Edge Map (REM) dapat dihitung dengan mengekstraksi setiap *radiating line* dan *radiating difference* sepanjang garis. Tahap-tahap perhitungan *Radiating Edge Map* (REM) dapat dijabarkan sebagai berikut:

1. Tahap pertama pada perhitungan REM adalah mengekstraksi setiap *radiating line* menggunakan persamaan

$$len_b = \lfloor \sqrt{(x_b - x_c)^2 + (y_b - y_c)^2} \rfloor. \quad (3.2)$$

2. Kemudian, menghitung *radiating difference* sepanjang garis menggunakan persamaan

$$rd(x_i, y_i) = I_0(x_{i-1}, y_{i-1}) - I_0(x_i, y_i). \quad (3.3)$$

Radiating gradient $rg(x_i, y_i)$ pada titik (x_i, y_i) dapat dihitung dengan

$$rg(x_i, y_i) = \frac{|rd(x_i, y_i)| + |rd(x_{i+1}, y_{i+1})|}{2}. \quad (3.4)$$

3. *Radiating difference* yang bernilai positif harus diperbaiki, karena intensitas area *centroid* memiliki nilai kecil dan akan semakin besar jika semakin jauh dari *centroid* di sepanjang *radiating line*. *Radiating difference* diperbaiki menggunakan persamaan

$$rg(x_i, y_i) = \frac{|Fs(rd(x_i, y_i))| + |Fs(rd(x_{i+1}, y_{i+1}))|}{2}. \quad (3.5)$$

3.3.3.2 Generalized Gradient Vector Flow Snake

GVFS memiliki konvergensi yang lemah dan sensitif terhadap inialisasi *snake* (Qin et al. 2013). Kelemahan ini disebabkan karena *smoothing* yang berlebihan pada area *boundary* yang diatur oleh koefisien μ pada Persamaan (2.14). Untuk mengatasi masalah ini, ditambahkan dua koefisien pembobot pada persamaan gaya eksternal GVFS (Qin et al. 2013). Persamaan gaya eksternal tersebut dinamakan *Generalized Gradient Vector Flow Snake* (GGVFS). Langkah untuk metode GGVFS adalah sebagai berikut:

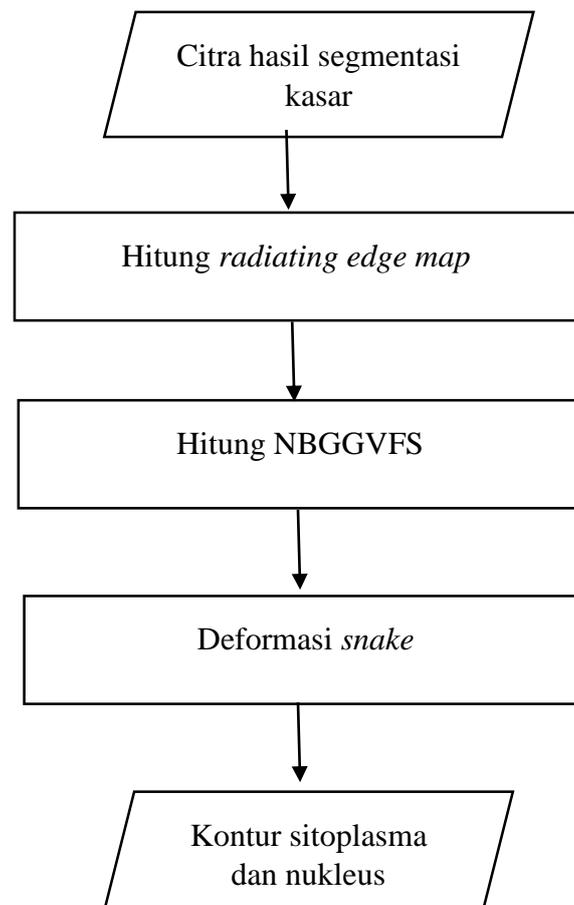
1. Normalisasi citra *edge map* yang telah didapat dari proses perhitungan REM ke dalam range $[0,1]$,
2. Update vektor GGVFS secara berulang dengan menggunakan persamaan :

$$v(x, y, t + \Delta t) = v(x, y, t) + \frac{1}{4}g(|\nabla REM|)A * v(x, y, t) - \frac{1}{4}h(|\nabla REM|)[v(x, y, t) - \nabla REM]. \quad (3.6)$$

3.3.3.3 Normally Biased Generalized Gradient Vector Flow Snake

Model GGVF memperbesar jangkauan konvergensi kontur aktif, meningkatkan kinerja konvergensi dan lebih handal terhadap *noise*. Berdasarkan metode NGVF, yang memiliki efisiensi difusi yang lebih tinggi, NBGVF memberikan solusi untuk masalah perlindungan tepi yang tipis. Oleh karena itu, metode ini menggabungkan GGVF dan NBGVF untuk mengusulkan versi energi eksternal yang baru, yang didefinisikan sebagai berikut:

$$v(x, y, t) = g(|\nabla REM|)(g_s(f)V_{NN}(x, y, t) + h_s(f)V_{TT}(x, y, t)) - h(|\nabla REM|)[V(x, y, t) - \nabla REM]. \quad (3.7)$$



Gambar 3.7 Diagram alir segmentasi kontur.

3.3.3.4 Deformasi Snake

Active contour atau *snake* dapat dihitung dengan meminimasi fungsi gaya eksternal dan internal yang dihitung dari data citra. Gaya eksternal *snake* telah dihitung dengan menggunakan metode NBGGVFS. Untuk gaya internal *snake* dihitung dengan langkah-langkah sebagai berikut:

1. Tentukan kontur awal yang telah didapat dari proses segmentasi awal,

2. Tentukan parameter *snake* yang dibutuhkan seperti *alpha* (parameter elastisitas), *beta* (parameter ketegaran), *gamma* (parameter kerekatan) dan *kappa* (bobot tekanan eksternal),
3. Tiap titik *snake* akan bergerak secara berulang mencari lokasi yang terdapat energi *snake* paling minimum.

3.4 Uji Coba

Setelah tahapan implementasi metode selesai, maka tahapan selanjutnya adalah melakukan uji coba terhadap sistem yang telah dibuat. Uji coba dimaksudkan untuk mengetahui apakah penelitian yang dilakukan telah dapat memenuhi tujuan penelitian sebagaimana yang telah direncanakan. Pada tahap ini akan dilakukan uji coba segmentasi citra sel tunggal smear serviks menggunakan metode usulan yang akan dibandingkan dengan beberapa metode segmentasi smear serviks yang sudah ada sebelumnya yaitu kombinasi *spatial k-means* dan RGVFS, FCM dan RGVFS, serta FCM_NLS dan RCNGGVFS.

3.5 Evaluasi

Evaluasi performa metode pada segmentasi kontur nukleus dan sitoplasma citra sel tunggal smear serviks dilakukan dengan cara membandingkan citra hasil segmentasi dengan citra hasil segmentasi manual. Evaluasi dilakukan dengan dua cara, diantaranya mencari nilai *Zijdenbos Similarity Index (ZSI)*, dan menggunakan *confusion matrix*. ZSI merupakan metode pengukuran untuk mengetahui performa metode segmentasi berbasis region. ZSI digunakan untuk menghitung seberapa besar rasio overlap diantara dua citra. ZSI didefinisikan sebagai berikut

$$ZSI = 2 \frac{|A_M \cap A_G|}{|A_M| + |A_G|}, \quad (3.8)$$

dimana A_M merupakan citra hasil segmentasi menggunakan metode yang diusulkan, sedangkan A_G merupakan citra segmentasi manual. *Confusion matrix* digunakan untuk mencari nilai akurasi dari hasil segmentasi seperti yang terlihat pada Tabel 3.2.

Tabel 3.2 Tabel *Confusion Matrix*

| <i>Confusion matrix</i> | | <i>Predicted</i> | |
|-------------------------|-----------------|------------------|-----------------|
| | | <i>Negative</i> | <i>Positive</i> |
| <i>Actual</i> | <i>Negative</i> | TN | FP |
| | <i>Positive</i> | FN | TP |

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}. \quad (3.9)$$

Dengan menggunakan *true positive* (TP), *true negative* (TN), *false negative* (FN), dan *false negative* (FN), nilai akurasi segmentasi bisa didapatkan. Nilai akurasi segmentasi mendefinisikan seberapa besar yang terklasifikasi dengan benar secara keseluruhan.

(Halaman ini sengaja dikosongkan)

BAB 4

UJI COBA DAN ANALISIS HASIL

4.1 Lingkungan Uji Coba

Data uji coba dalam penelitian ini diuji dengan menggunakan perangkat lunak dengan spesifikasi sebagai berikut:

- a. *Processor Intel(R) Core(TM) i5-3210M CPU @2.50GHz 2.50GHz*
- b. Kapasitas memori 4,00 GB
- c. Kapasitas *harddisk* 232 GB

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

- a. Sistem operasi *Windows* 8.1 64-bit
- b. Aplikasi *Matlab* R2014a

4.2 Data Uji Coba

Data masukan pada uji coba ini berupa citra mikroskopik sel tunggal smear serviks. Data yang digunakan diperoleh dari Departemen Patologi Rumah Sakit Universitas Teknik Denmark. Dataset Herlev ini terdiri dari 7 kelas yang berbeda dengan jumlah keseluruhan 917 citra sel tunggal smear serviks (Jantzen et al. 2006). Setiap satu citra sel tunggal smear serviks memiliki ukuran yang berbeda-beda, berada dalam ruang warna RGB dan berformat BMP. Contoh citra sel tunggal smear serviks yang digunakan pada penelitian ini dapat dilihat pada Gambar 3.3.

4.3 Hasil Uji Coba

4.3.1 Hasil Uji Coba Skenario 1

Uji coba skenario pertama yaitu melakukan segmentasi dengan metode usulan yaitu kombinasi metode RNBGGVFS dan metode segmentasi awal *mean shift clustering* (RNBGGVFS & *mean shift*). Uji coba skenario 1 ini membandingkan hasil segmentasi dengan menggunakan variasi nilai parameter pada segmentasi awal, yaitu metode *mean shift clustering*. Parameter yang digunakan adalah parameter *bandwidth* atau radius. Parameter *bandwidth* atau radius merupakan parameter jarak antar titik dalam sebuah kernel yang menentukan suatu titik merupakan bagian dari suatu kluster atau tidak. Seperti yang dijelaskan pada Sub Bab 3.3, penghitungan jarak antar piksel dibandingkan dengan *bandwidth* (radius),

dimana titik dengan jarak yang lebih kecil dari *bandwidth* selanjutnya digabung menjadi kluster yang sama dan dihitung *meannya*.

Penelitian ini tidak berhasil menganalisis semua citra pada dataset. Ada beberapa citra yang gagal disegmentasi dikarenakan beberapa hal. Citra pada dataset Herlev memiliki karakteristik yang berbeda, masing-masing kelas memiliki ukuran nukleus yang berbeda pula. Banyak *noise* yang terdapat pada citra serta kekontrasan yang rendah juga memengaruhi hasil segmentasi. Beberapa citra juga ditemukan memiliki jumlah sel lebih dari satu serta ada pula citra dengan sel yang overlap. Hal tersebut juga memengaruhi keberhasilan proses segmentasi. Tabel 4.1 menunjukkan banyaknya citra dari dataset Herlev yang dapat dianalisis.

Tabel 4.1 Jumlah citra yang dapat dianalisis.

| Nama Kelas | Jumlah data citra Herlev | Jumlah data yang berhasil dianalisis dengan metode RNBGGVFS & <i>mean shift</i> (metode usulan) |
|----------------------------|--------------------------|---|
| <i>Carcinoma in situ</i> | 150 | 116 |
| <i>Light Dysplastic</i> | 182 | 126 |
| <i>Moderate Dysplastic</i> | 146 | 100 |
| <i>Normal Columnar</i> | 98 | 67 |
| <i>Normal Intermediate</i> | 70 | 28 |
| <i>Normal Superficial</i> | 74 | 44 |
| <i>Severe Dysplastic</i> | 197 | 159 |

Percobaan skenario pertama yaitu uji coba dengan menggunakan variasi nilai *bandwidth*. Nilai yang diuji coba yaitu 0,2 , 0,25 dan 0,3. Tabel 4.2 merupakan hasil evaluasi uji coba nilai *bandwidth* = 0,2. Hasil tersebut meliputi nilai rata-rata akurasi dan ZSI pada nukleus dan sitoplasma citra sel tunggal smear serviks. Nilai akurasi didapatkan dengan membandingkan area putih dan area hitam pada citra *groundtruth* dan citra hasil segmentasi sistem. Nilai index ZSI merupakan seberapa besar area yang overlap antara citra hasil segmentasi sistem dan citra *groundtruth*. Tabel 4.2 menunjukkan bahwa nilai rata-rata akurasi segmentasi nukleus diperoleh sebesar 96,27%, dan rata-rata akurasi segmentasi sitoplasma citra sel tunggal smear serviks sebesar 83,56%. Sedangkan nilai rata-rata ZSI segmentasi nukleus didapatkan sebesar 89,20% dan nilai rata-rata ZSI segmentasi sitoplasma sebesar 87,24%.

Tabel 4.3 menunjukkan hasil evaluasi uji coba segmentasi citra sel tunggal smear serviks dengan variasi nilai *bandwidth* = 0,25. Nilai rata-rata akurasi segmentasi nukleus citra sel tunggal smear serviks pada keseluruhan kelas dataset diperoleh sebesar 96,24% dan rata-rata

akurasi sitoplasma diperoleh sebesar 83,11%. Sedangkan nilai rata-rata ZSI segmentasi nukleus diperoleh sebesar 89,10% dan nilai rata-rata ZSI segmentasi sitoplasma diperoleh sebesar 87,14%. Berdasarkan hasil pada tabel diketahui bahwa hasil segmentasi dengan variasi nilai *bandwidth* = 0,25 belum mendapatkan hasil yang lebih baik dari nilai *bandwidth* = 0,2 pada hasil rata-rata akurasi nukleus, rata-rata ZSI nukleus dan sitoplasma. Sementara pada hasil akurasi sitoplasma, nilai rata-rata yang diperoleh lebih baik dibandingkan pada *bandwidth* = 0,2.

Tabel 4.2 Hasil uji coba dengan nilai *bandwidth* = 0,2.

| Kelas | Rata-rata Akurasi | | Rata-rata ZSI | |
|----------------------------|-------------------|------------|---------------|------------|
| | Nukleus | Sitoplasma | Nukleus | Sitoplasma |
| <i>Carcinoma in situ</i> | 0,927 | 0,807 | 0,918 | 0,858 |
| <i>Light Dysplastic</i> | 0,976 | 0,849 | 0,957 | 0,875 |
| <i>Moderate Dysplastic</i> | 0,966 | 0,812 | 0,931 | 0,854 |
| <i>Normal Columnar</i> | 0,925 | 0,765 | 0,836 | 0,826 |
| <i>Normal Intermediate</i> | 0,996 | 0,918 | 0,867 | 0,935 |
| <i>Normal Superficial</i> | 0,998 | 0,875 | 0,818 | 0,902 |
| <i>Severe Dysplastic</i> | 0,945 | 0,823 | 0,918 | 0,857 |
| Rata-rata | 0,962 | 0,836 | 0,892 | 0,872 |

Tabel 4.3 Hasil uji coba dengan nilai *bandwidth* = 0,25.

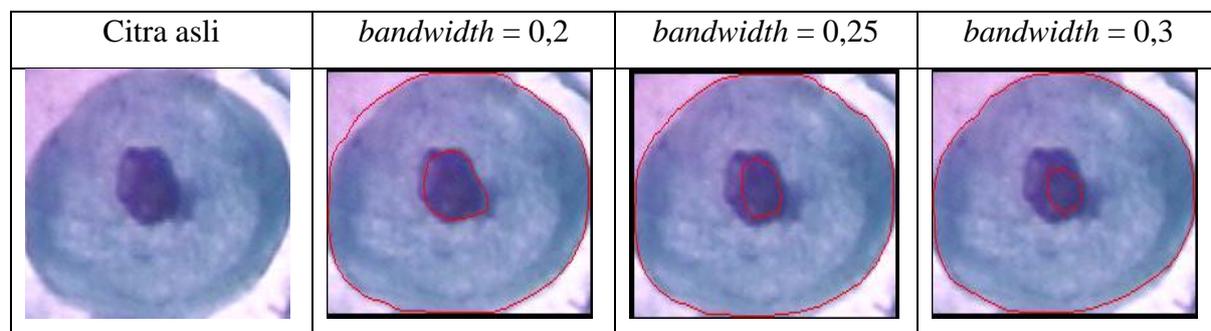
| Kelas | Rata-rata Akurasi | | Rata-rata ZSI | |
|----------------------------|-------------------|------------|---------------|------------|
| | Nukleus | Sitoplasma | Nukleus | Sitoplasma |
| <i>Carcinoma in situ</i> | 0,949 | 0,808 | 0,940 | 0,879 |
| <i>Light Dysplastic</i> | 0,973 | 0,827 | 0,930 | 0,910 |
| <i>Moderate Dysplastic</i> | 0,963 | 0,844 | 0,909 | 0,802 |
| <i>Normal Columnar</i> | 0,938 | 0,704 | 0,900 | 0,824 |
| <i>Normal Intermediate</i> | 0,982 | 0,919 | 0,826 | 0,923 |
| <i>Normal Superficial</i> | 0,995 | 0,919 | 0,813 | 0,912 |
| <i>Severe Dysplastic</i> | 0,938 | 0,796 | 0,919 | 0,849 |
| Rata-rata | 0,962 | 0,831 | 0,891 | 0,871 |

Tabel 4.4 menunjukkan hasil evaluasi segmentasi citra sel tunggal smear serviks dengan variasi nilai *bandwidth* pada segmentasi awal *mean shift clustering* yaitu 0,3. Pada uji coba kali ini diperoleh nilai rata-rata akurasi nukleus sebesar 95,9% dan rata-rata sitoplasma sebesar 86,3%. Sedangkan untuk nilai rata-rata ZSI nukleus didapatkan sebesar 89,19% dan rata-rata ZSI sitoplasma citra sel tunggal smear serviks didapatkan sebesar 87,18%. Berdasarkan uji coba pada ketiga nilai *bandwidth* tersebut, diketahui bahwa nilai rata-rata akurasi dan ZSI, baik pada nukleus maupun sitoplasma, cenderung lebih tinggi pada saat *bandwidth* atau radius bernilai 0,2 dibandingkan pada saat bernilai 0,25 dan 0,3, sehingga *bandwidth* bernilai 0,2 akan

digunakan pada uji coba skenario kedua. Gambar 4.3 menunjukkan contoh gambar hasil segmentasi citra sel tunggal smear serviks menggunakan metode RNBGGVFS dan *mean shift clustering* dengan beberapa variasi nilai *bandwidth* yaitu 0,2 , 0,25 dan 0,3.

Tabel 4.4 Hasil uji coba dengan nilai *bandwidth* = 0,3.

| Kelas | Rata-rata Akurasi | | Rata-rata ZSI | |
|----------------------------|-------------------|------------|---------------|------------|
| | Nukleus | Sitoplasma | Nukleus | Sitoplasma |
| <i>Carcinoma in situ</i> | 0,943 | 0,863 | 0,919 | 0,826 |
| <i>Light Dysplastic</i> | 0,942 | 0,894 | 0,929 | 0,910 |
| <i>Moderate Dysplastic</i> | 0,964 | 0,844 | 0,939 | 0,872 |
| <i>Normal Columnar</i> | 0,947 | 0,794 | 0,847 | 0,824 |
| <i>Normal Intermediate</i> | 0,989 | 0,909 | 0,827 | 0,920 |
| <i>Normal Superficiel</i> | 0,990 | 0,910 | 0,831 | 0,929 |
| <i>Severe Dysplastic</i> | 0,938 | 0,827 | 0,950 | 0,821 |
| Rata-rata | 0,959 | 0,863 | 0,891 | 0,872 |



Gambar 4.1 Contoh hasil segmentasi menggunakan metode RNBGGVFS dan *mean shift clustering* dengan beberapa variasi nilai *bandwidth* (radius).

4.3.2 Hasil Uji Coba Skenario 2

Uji coba skenario kedua dilakukan dengan membandingkan metode usulan yaitu kombinasi *mean shift clustering* dan RNBGGVFS dengan metode penelitian sebelumnya untuk segmentasi citra sel tunggal smear serviks. Metode tersebut diantaranya kombinasi *spatial k-means* dengan *Radiating Gradient Vector Flow Snake* (RGVFS) (Li et al. 2012), kombinasi FCM dan RGVFS (Susanti et al. 2015), serta kombinasi FCM_NLS dan RCNGGVFS (Husain et al. 2017). Tabel 4.5 menunjukkan banyaknya citra yang dapat dianalisis pada keempat metode yang diuji.

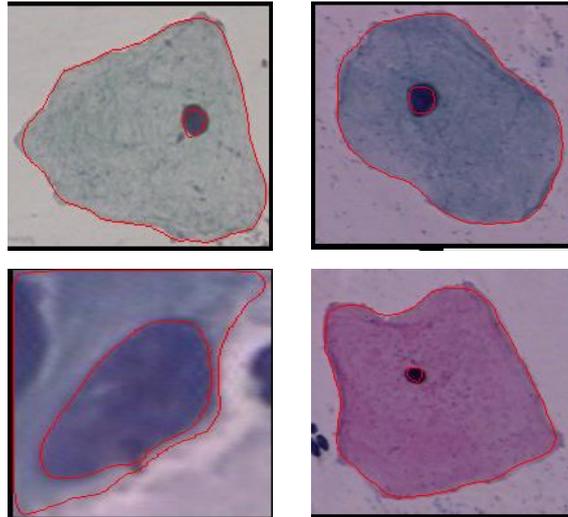
Uji coba skenario pertama telah diketahui jumlah citra yang dapat dianalisis dengan metode usulan (RNBGGVFS & *mean shift*) beserta alasan mengapa tidak semua data citra berhasil dianalisis. Pada metode penelitian sebelumnya, citra yang dapat dianalisis cenderung lebih sedikit dibanding dengan metode usulan. Pada metode RGVFS & FCM serta RGVFS & *spatial kmeans*, sedikitnya jumlah citra yang berhasil dianalisis dikarenakan salah satu

penyebabnya adalah metode *denoising* yang digunakan belum handal, sementara citra yang ada memiliki banyak *noise*. Hal tersebut berdampak pada proses segmentasi awal citra sel tunggal smear serviks. Pada tabel diketahui bahwa kelas dengan jumlah citra paling sedikit yang berhasil disegmentasi adalah kelas *Normal Superficiel*. Kelas tersebut memiliki karakteristik ukuran nukleus yang relatif kecil. Nukleus tersebut tidak berhasil disegmentasi pada proses segmentasi awal, sehingga proses pencarian kontur tidak dapat dilanjutkan ke proses selanjutnya yaitu segmentasi akhir. Kemudian pada metode RCNGGVFS & FCM_NLS jumlah citra yang dapat dianalisis meningkat, dikarenakan proses segmentasi awal yang memiliki metode penghilangan *noise* yang cukup handal didalamnya. Pada penelitian ini digunakan metode segmentasi awal yaitu *mean shift clustering*. Metode ini diketahui dapat menangani ukuran serta bentuk klaster data. Metode ini juga tidak sensitif terhadap *noise*, sehingga menghasilkan lebih banyak citra yang berhasil disegmentasi. Metode *mean shift* merupakan metode *clustering* berbasis densitas dan *nonparametric* yang mampu menangani ukuran serta bentuk data. Dengan metode ini diharapkan akan mampu meningkatkan jumlah citra yang dapat disegmentasi.

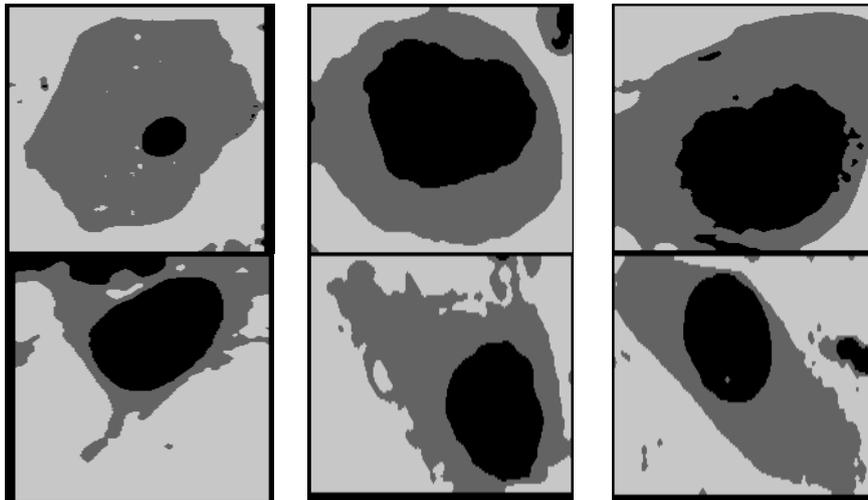
Tabel 4.5 Jumlah citra yang dapat dianalisis pada uji coba skenario 2.

| Nama Kelas | Jumlah data citra herlev | Jumlah data yang berhasil dianalisis | | | |
|----------------------------|--------------------------|--|--------------------|-------------|--------------------------------|
| | | RNBGGVFS & <i>mean shift</i> (metode usulan) | RCNGGVFS & FCM_NLS | RGVFS & FCM | RGVFS & <i>spatial k-means</i> |
| <i>Carcinoma in situ</i> | 150 | 116 | 67 | 19 | 20 |
| <i>Light Dysplastic</i> | 182 | 126 | 61 | 22 | 19 |
| <i>Moderate Dysplastic</i> | 146 | 100 | 51 | 20 | 19 |
| <i>Normal Columnar</i> | 98 | 67 | 33 | 20 | 20 |
| <i>Normal Intermediate</i> | 70 | 28 | 20 | 15 | 14 |
| <i>Normal Superficiel</i> | 74 | 44 | 3 | 3 | 3 |
| <i>Severe Dysplastic</i> | 197 | 159 | 88 | 17 | 17 |

Uji coba skenario 2 ini dilakukan evaluasi dengan membandingkan hasil segmentasi menggunakan metode usulan dengan beberapa metode segmentasi sebelumnya (RCNGGVFS & FCM_NLS, RGVFS & FCM dan RGVFS & *spatial kmeans*) dan dengan menggunakan data citra yang sama. Berdasarkan hasil uji coba skenario pertama, parameter *bandwidth* yang digunakan pada segmentasi awal adalah 0,2. Gambar 4.2 merupakan contoh citra hasil segmentasi dengan metode usulan (RNBGGVFS & *mean shift*) dan Gambar 4.3 adalah contoh hasil segmentasi awal menggunakan metode *mean shift clustering*.



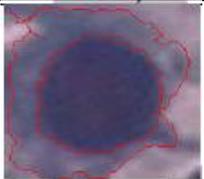
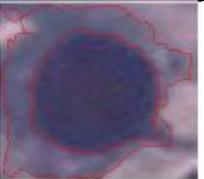
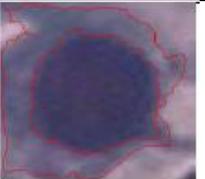
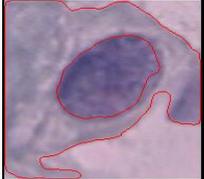
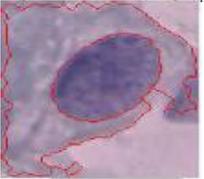
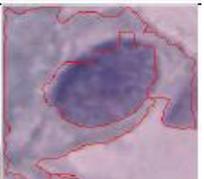
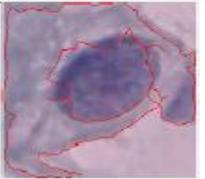
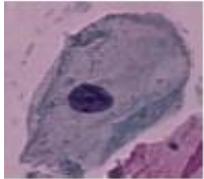
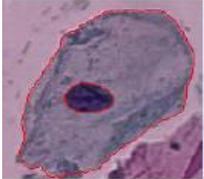
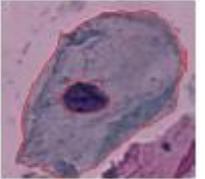
Gambar 4.2 Contoh citra yang dapat dianalisis pada metode RNBGGVFS & *mean shift clustering* namun belum berhasil pada metode sebelumnya.



Gambar 4.3 Contoh hasil segmentasi awal metode *mean shift clustering*.

Gambar 4.4 menunjukkan kontur sitoplasma dan nukleus citra sel tunggal smear serviks. Hasil percobaan dengan menggunakan metode usulan terlihat bahwa kontur yang dihasilkan semakin mendekati kontur pada citra asli. Sementara pada metode lain masih terlihat bahwa ada area bukan nukleus yang dideteksi sebagai area nukleus, begitu juga pada area sitoplasma. Tabel 4.6 dan Tabel 4.7 menunjukkan hasil evaluasi yaitu nilai akurasi dan ZSI pada semua kelas dataset citra Herlev area nukleus dengan menggunakan metode RNBGGVFS & *mean shift clustering* (metode usulan), RCNNGGVFS & FCM_NLS, RGVFS & FCM, dan RGVFS & *spatial k-means*. Nilai akurasi dan ZSI diperoleh dengan membandingkan suatu gambar dengan *groundtruth*-nya. Hasil tabel terlihat bahwa rata-rata akurasi nukleus pada metode usulan memiliki nilai tertinggi yaitu 96,96%, dan rata-rata ZSI nukleus pada metode usulan juga memiliki nilai tertinggi yaitu sebesar 90,68%. Namun pada kelas *normal intermediate*, nilai

rata-rata tertinggi diperoleh dengan metode RCNGGVFS & FCM_NLS. Metode RCNGGVFS & FCM_NLS memperoleh nilai rata-rata akurasi sebesar 95,34% dan rata-rata ZSI sebesar 88,06%. Metode RGVFS & FCM memperoleh nilai rata-rata akurasi sebesar 94,88% dan nilai rata-rata ZSI sebesar 85,69%. Sedangkan metode RGVFS & *spatial k-means* memperoleh nilai rata-rata akurasi sebesar 94,49% dan rata-rata ZSI sebesar 86,58%.

| Citra Asli | RNBGGVFS & <i>mean shift clustering</i> (metode usulan) | RCNGGVFS & FCM_NLS | RGVFS & FCM | RGVFS & <i>spatial k-means</i> |
|---|---|---|--|---|
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Gambar 4.4 Contoh hasil segmentasi pada beberapa metode.

Tabel 4.8 dan Tabel 4.9 menunjukkan hasil evaluasi segmentasi sitoplasma pada seluruh dataset citra Herlev dengan menggunakan metode RNBGGVFS & *mean shift* (metode usulan), RCNGGVFS & FCM_NLS, RGVFS & FCM dan RGVFS & *spatial k-means*. Tabel menunjukkan bahwa rata-rata akurasi dan ZSI sitoplasma pada metode usulan memiliki nilai yang lebih tinggi dibandingkan ketiga metode lain yaitu sebesar 86,78% dan 89,35%. Segmentasi sitoplasma citra kelas *normal intermediate* dan *normal columnar* diperoleh nilai rata-rata ZSI dan akurasi yang tertinggi pada metode RCNGGVFS & FCM_NLS.

4.4 Analisis Hasil Uji Coba

Dataset citra yang digunakan pada penelitian ini adalah citra mikroskopis sel tunggal smear serviks yang diperoleh dari dataset Herlev. Dataset Herlev terdiri dari 917 citra sel tunggal smear serviks yang diperoleh dari Departemen Patologi Rumah Sakit Universitas Herlev dan Departemen Automasi Universitas Teknik Denmark. Dataset citra tersebut dikumpulkan oleh ahli *cytology* menggunakan mikroskop yang terhubung ke kamera digital.

Ahli *cytology* mengklasifikasikan dataset citra sel tunggal smear serviks menjadi 7 kelas secara manual. Kelas tersebut terdiri dari kelas *Carcinoma in situ*, *Light Dysplastic*, *Moderate Dysplastic*, *Normal Columnar*, *Normal Intermediate*, *Normal Superficial*, dan *Severe Dysplastic*. Dataset Herlev juga terdapat *groundtruth* atau hasil segmentasi manual. *Groundtruth* citra sel tunggal smear serviks didapatkan dengan menggunakan *software* CHAMP dan tersegmentasi menjadi 3 bagian, yaitu sitoplasma, nukleus, dan *background* (Jantzen et al. 2006).

Tabel 4.6 Hasil evaluasi segmentasi nukleus citra sel tunggal smear *serviks* pada beberapa metode.

| Kelas | Rata-rata Akurasi | | | |
|----------------------------|--|-------------------|-------------|--------------------------------|
| | RNBGGVFS & <i>mean shift</i> (metode usulan) | RCNGGVF & FCM_NLS | RGVFS & FCM | RGVFS & <i>spatial k-means</i> |
| <i>Carcinoma in situ</i> | 0,953 | 0,918 | 0,915 | 0,910 |
| <i>Light Dysplastic</i> | 0,973 | 0,972 | 0,964 | 0,962 |
| <i>Moderate Dysplastic</i> | 0,973 | 0,959 | 0,955 | 0,954 |
| <i>Normal Columnar</i> | 0,952 | 0,897 | 0,895 | 0,881 |
| <i>Normal Intermediate</i> | 0,994 | 0,993 | 0,991 | 0,988 |
| <i>Normal Superficial</i> | 0,995 | 0,995 | 0,992 | 0,988 |
| <i>Severe Dysplastic</i> | 0,947 | 0,938 | 0,930 | 0,931 |
| Rata-rata | 0,969 | 0,953 | 0,948 | 0,945 |

Tabel 4.7 Hasil evaluasi segmentasi nukleus citra sel tunggal smear *serviks* pada beberapa metode.

| Kelas | Rata-rata ZSI | | | |
|----------------------------|--|--------------------|-------------|--------------------------------|
| | RNBGGVFS & <i>mean shift</i> (metode usulan) | RCNGGVFS & FCM_NLS | RGVFS & FCM | RGVFS & <i>spatial k-means</i> |
| <i>Carcinoma in situ</i> | 0,944 | 0,906 | 0,896 | 0,896 |
| <i>Light Dysplastic</i> | 0,942 | 0,918 | 0,906 | 0,916 |
| <i>Moderate Dysplastic</i> | 0,946 | 0,926 | 0,916 | 0,902 |
| <i>Normal Columnar</i> | 0,898 | 0,832 | 0,834 | 0,833 |
| <i>Normal Intermediate</i> | 0,856 | 0,869 | 0,803 | 0,791 |
| <i>Normal Superficial</i> | 0,823 | 0,797 | 0,738 | 0,817 |
| <i>Severe Dysplastic</i> | 0,938 | 0,916 | 0,905 | 0,905 |
| Rata-rata | 0,907 | 0,881 | 0,857 | 0,866 |

Dari semua kelas citra sel tunggal smear serviks yang berjumlah 917 citra, tidak semua citra berhasil dianalisis. Ada beberapa citra yang gagal disegmentasi dikarenakan beberapa hal. Citra pada dataset Herlev memiliki karakteristik yang berbeda, masing-masing kelas memiliki

ukuran nukleus yang berbeda pula. Banyak *noise* yang terdapat pada citra serta kekontrasan yang rendah juga memengaruhi hasil segmentasi. Beberapa citra juga ditemukan memiliki jumlah sel lebih dari satu serta ada pula citra dengan sel yang *overlap*. Hal tersebut juga memengaruhi keberhasilan proses segmentasi. Gambar 4.5 merupakan beberapa contoh citra yang belum berhasil disegmentasi menggunakan metode usulan pada penelitian ini.

Metode segmentasi awal yang digunakan pada penelitian ini adalah *mean shift clustering*. Segmentasi awal berfungsi untuk mendapatkan kontur awal yang akan digunakan sebagai data masukan pada proses segmentasi kontur menggunakan metode *Radiating Normally Biased Generalized Gradient Vector Flow Snake* (RNBGGVFS). Metode *mean shift clustering* merupakan metode *clustering* berbasis densitas dan *nonparametric* yang dapat menangani ukuran dan bentuk klaster data. Beberapa karakteristik citra dataset memiliki perbedaan ukuran data yang sangat mencolok, contohnya pada kelas *normal superficiel* yang memiliki ukuran nukleus yang sangat kecil, tidak berhasil dideteksi dengan metode segmentasi awal dari penelitian sebelumnya. Hal ini bermanfaat pada proses segmentasi awal untuk meningkatkan jumlah citra yang berhasil dianalisis dan dapat dilanjutkan ke tahap segmentasi berikutnya. Seperti yang ditunjukkan pada Tabel 4.5, jumlah citra yang dapat dianalisis dengan metode segmentasi awal *mean shift clustering* lebih tinggi dibanding metode penelitian sebelumnya.

Tabel 4.8 Hasil evaluasi segmentasi sitoplasma citra sel tunggal smear *serviks* pada beberapa metode.

| Kelas | Rata-rata Akurasi | | | |
|----------------------------|--|---------------------|-------------|--------------------------------|
| | RNBGGVFS & <i>mean shift</i> (metode usulan) | RCNBBGVFS & FCM_NLS | RGVFS & FCM | RGVFS & <i>spatial k-means</i> |
| <i>Carcinoma in situ</i> | 0,880 | 0,799 | 0,790 | 0,779 |
| <i>Light Dysplastic</i> | 0,898 | 0,835 | 0,835 | 0,834 |
| <i>Moderate Dysplastic</i> | 0,864 | 0,804 | 0,761 | 0,765 |
| <i>Normal Columnar</i> | 0,792 | 0,805 | 0,795 | 0,803 |
| <i>Normal Intermediate</i> | 0,916 | 0,937 | 0,931 | 0,935 |
| <i>Normal Superficiel</i> | 0,903 | 0,852 | 0,841 | 0,838 |
| <i>Severe Dysplastic</i> | 0,822 | 0,811 | 0,775 | 0,768 |
| Rata-rata | 0,868 | 0,835 | 0,818 | 0,817 |

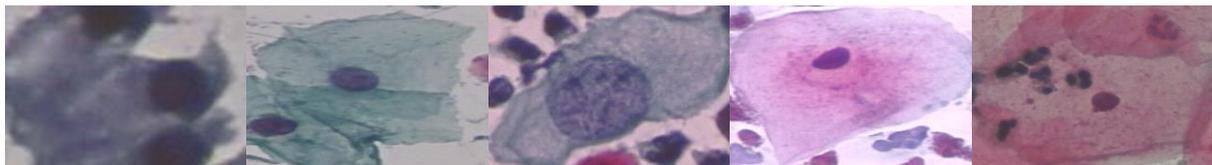
Edge map yang dihitung dengan menggunakan operator deteksi tepi seperti *sobel*, belum dapat mendeteksi daerah nukleus dari citra smear serviks dengan benar (Li et al. 2012). Hal ini disebabkan karena area nukleus umumnya lebih gelap daripada area sitoplasma dan area sitoplasma lebih gelap dibanding dengan area *background*. REM merupakan proses

perhitungan *gradient edge map* pada sepanjang garis pancaran atau *radiating line* yang dimulai dari *intensity-weighted area centroid* nukleus hasil segmentasi awal.

Tabel 4.9 Hasil evaluasi segmentasi sitoplasma citra sel tunggal smear *serviks* pada beberapa metode.

| Kelas | Rata-rata ZSI | | | |
|----------------------------|--|--------------------|-------------|--------------------------------|
| | RNBGGVFS & <i>mean shift</i> (metode usulan) | RCNGGVFS & FCM_NLS | RGVFS & FCM | RGVFS & <i>spatial k-means</i> |
| <i>Carcinoma in situ</i> | 0,905 | 0,849 | 0,845 | 0,835 |
| <i>Light Dysplastic</i> | 0,916 | 0,867 | 0,865 | 0,866 |
| <i>Moderate Dysplastic</i> | 0,886 | 0,847 | 0,809 | 0,812 |
| <i>Normal Columnar</i> | 0,826 | 0,835 | 0,830 | 0,829 |
| <i>Normal Intermediate</i> | 0,931 | 0,951 | 0,944 | 0,948 |
| <i>Normal Superficial</i> | 0,933 | 0,895 | 0,889 | 0,887 |
| <i>Severe Dysplastic</i> | 0,857 | 0,856 | 0,839 | 0,833 |
| Rata-rata | 0,893 | 0,871 | 0,860 | 0,858 |

Metode perhitungan REM yang dimanfaatkan pada metode GVFS belum berhasil untuk mendapatkan kontur sitoplasma dan nukleus secara akurat. Hal ini disebabkan karena GVFS memiliki konvergensi yang lemah dan sensitif terhadap inisialisasi *snake* (Li et al. 2012). Penelitian ini menggunakan metode RNBGGVFS, yaitu metode yang memanfaatkan perhitungan REM dalam pencarian *edge map* pada kombinasi metode *Generalized Gradient Vector Flow Snake* (GGVFS) dan *Normally Biased Gradient Vector Flow Snake* (NBGVFS). Model GGVFS memperbesar jangkauan konvergensi kontur aktif, meningkatkan kinerja konvergensi dan lebih handal terhadap *noise*. Berdasarkan metode NGVF, yang memiliki efisiensi difusi yang lebih tinggi, NBGVF memberikan solusi untuk masalah perlindungan tepi yang lemah. Oleh karena itu, metode ini menggabungkan GGVF dan NBGVF untuk mengusulkan versi energi eksternal yang baru. Dan dengan menggunakan perhitungan REM sebagai pengganti perhitungan konvensional *edge map* pada metode NBBGVFS dapat memberikan hasil segmentasi kontur yang akurat pada citra sel tunggal smear serviks.



Gambar 4.5 Contoh citra yang tidak berhasil disegmentasi.

Tabel 4.10 Hasil evaluasi segmentasi nukleus pada semua kelas dataset Herlev.

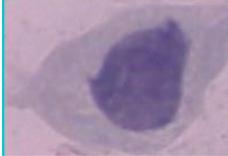
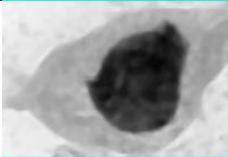
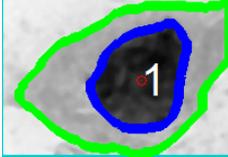
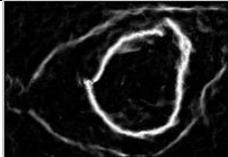
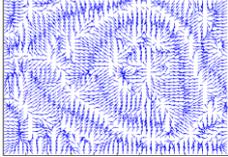
| Metode | Rata-rata akurasi nukleus | Rata-rata ZSI nukleus |
|------------------------------------|---------------------------|-----------------------|
| <i>RNBGGVFS & mean shift</i> | 0,969 | 0,907 |
| <i>RCNNGGVFS & FCM_NLS</i> | 0,953 | 0,881 |
| <i>RGVFS & FCM</i> | 0,948 | 0,857 |
| <i>RGVFS & spatial k-means</i> | 0,945 | 0,866 |

Tabel 4.11 Hasil evaluasi segmentasi nukleus pada semua kelas dataset Herlev.

| Metode | Rata-rata akurasi sitoplasma | Rata-rata ZSI sitoplasma |
|------------------------------------|------------------------------|--------------------------|
| <i>RNBGGVFS & mean shift</i> | 0,868 | 0,893 |
| <i>RCNNGGVFS & FCM_NLS</i> | 0,835 | 0,871 |
| <i>RGVFS & FCM</i> | 0,818 | 0,860 |
| <i>RGVFS & spatial k-means</i> | 0,817 | 0,858 |

Uji coba dilakukan untuk membuktikan performa kombinasi metode *RNBGGVFS & mean shift clustering* dibanding dengan metode pada penelitian sebelumnya, yaitu *RGVFS & spatial k-means* (Li et al. 2012), *RGVFS & FCM* (Susanti et al. 2015) dan *RCNNGGVFS & FCM_NLS* (Husain et al. 2017). Tabel 4.10 dan Tabel 4.11 menunjukkan bahwa metode yang diusulkan pada penelitian ini memiliki rata-rata akurasi dan ZSI pada semua kelas citra sel tunggal smear serviks yang lebih tinggi dibandingkan dengan metode pada penelitian sebelumnya. Rata-rata akurasi dan ZSI nukleus pada metode usulan dihasilkan sebesar 96,96% dan 90,68% sedangkan pada sitoplasma dihasilkan 86,78% dan 89,35%. Metode *RCNNGGVFS & FCM_NLS* menghasilkan rata-rata akurasi dan ZSI nukleus sebesar 95,34% dan 88,06% sedangkan pada sitoplasma dihasilkan 83,48% dan 87,14%. Kemudian metode *RGVFS & FCM* menghasilkan rata-rata akurasi dan ZSI nukleus sebesar 94,88% dan 85,69%, sedangkan pada sitoplasma dihasilkan 81,83% dan 86%. Metode *RGVFS & spatial k-means* menghasilkan rata-rata akurasi dan ZSI nukleus sebesar 94,49% dan 86,58% sedangkan untuk sitoplasma dihasilkan sebesar 81,76% dan 85,88%. Contoh hasil keluaran sistem segmentasi citra sel tunggal smear serviks secara lengkap dapat dilihat pada Tabel 4.12.

Tabel 4.12 Contoh data keluaran sistem.

| No | Tahap | Hasil | Keterangan |
|----|-------------------|---|--|
| 1. | Citra Masukan |  Citra Asli | Dataset Herlev berformat bmp. |
| 2. | Praproses |  Citra Hasil Praproses | Setelah proses konversi ke CIELAB dan citra keabuan. |
| 3. | Segmentasi Awal |  Citra Hasil Segmentasi Awal | Hasil proses <i>mean shift clustering</i> . |
| 4. | Segmentasi Awal |  Inisial kontur | Setelah proses ekstraksi kontur awal dan pelabelan <i>centroid</i> . |
| 5. | Segmentasi Kontur |  <i>Radiating Edge Map</i> | Edge map yang dihitung dari citra hasil segmentasi awal. |
| 6. | Segmentasi Kontur |  <i>Normally Biased Generalized Gradient Vector Flow Snake</i> | <i>Edge map</i> dengan normalisasi energi eksternal NBGGVF. |
| 7. | Segmentasi Kontur |  Kontur nukleus dan sitoplasma citra sel tunggal smear serviks | Hasil segmentasi dengan metode RNBGGVFS. |

BAB 5

PENUTUP

Berdasarkan uji coba dan analisis hasil pengujian pada metode segmentasi citra sel tunggal smear serviks, dapat diuraikan beberapa kesimpulan dan saran untuk pengembangan lebih lanjut.

5.1 Kesimpulan

- a. Segmentasi kontur citra sel tunggal smear serviks dengan metode *Radiating Normally Biased Generalized Gradient Vector Flow Snake* (RNBGGVFS) berhasil dengan nilai rata-rata akurasi dan ZSI dari semua kelas untuk segmentasi nukleus adalah 96,96% dan 90,68%. Nilai rata-rata akurasi dan ZSI dari semua kelas untuk segmentasi sitoplasma adalah 86,78% dan 89,35%.
- b. Tahap pra proses berhasil dilakukan untuk mendapatkan citra keabuan dengan cara konversi citra dari ruang warna RGB ke ruang warna CIELAB layer L.
- c. Tahap segmentasi awal dengan metode *mean shift clustering* berhasil untuk mendapatkan hasil segmentasi yang baik sebagai data masukan proses segmentasi kontur serta meningkatkan jumlah citra yang dapat dianalisis.
- d. Berdasarkan hasil uji coba, metode usulan lebih optimal dalam mendeteksi kontur nukleus dan sitoplasma citra sel tunggal smear serviks dibanding kombinasi metode RCNGGVFS & FCM_NLS, RGVFS & FCM dan RGVFS & *spatial k-means*.

5.2 Saran

Pada penelitian ini masih terdapat beberapa citra yang belum berhasil dianalisis dikarenakan karakteristik citra yang saling overlap, maupun terdapat banyak sel dalam satu gambar, sehingga untuk penelitian selanjutnya disarankan mencari metode yang dapat menangani permasalahan tersebut.

(Halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- "Parkway Cancer Centre," Parkway Cancer Centre, 2014. [Online]. Available: <http://www.parkwaycancercentre.com/>. [Accessed 14 April 2018].
- C. N. Shanti, "Mengubah Citra Berwarna menjadi Grayscale dan Citra Biner," *Jurnal Teknologi Informasi DINAMIK*, Vols. 16, No.1, pp. 14-19, 2011.
- G. Hoffmann, "CIE color space," 2010.
- A. Noercholis and M. A. M. d. Maftuch, "Ekstraksi Fitur Roundness untuk Menghitung Jumlah Leukosit dalam Citra Sel Darah Ikan," *Jurnal EECCIS*, vol. 7, Juni 2013.
- Ratman, "Perancangan Aplikasi Perbaikan Citra Berbasis Web dengan Menggunakan Metode Filtering," *Pelita Informatika Budi Darma*, vol. V, no. 2, 2013.
- C. Glasbey and G. Horgan., "Image Analysis for the Biological Sciences," Wiley , 1995.
- Chuanyun, Xu., Yang, Zhang., & Sen, Wang (2013). "Cell Segmentation in Cervical smear images using polar coordinates GVF Snake with Radiating Edge Map", *Journal of Multimedia*, 8(3), 213-219.
- Xu, C., & Prince, J. L. (1998), "Generalized gradient vector flow external forces for active contours", vol.71, hal.131-139.
- D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Machine Intell.*, 24:603-619, 2002.
- Ganesan, P. et al., (2010), "Segmentation and Edge Detection of Color Images Using CIELAB Color Space and Edge Detectors", *IEEE On Emerging Trends in Robotics and Communication Technologies (INTERACT)*, hal.393-397.
- Gonzalez, R.C. & Woodz, R.E., (2008), "Digital Image Processing Third Edition", *Pearson Education, Inc.* Jantzen, J & Dounias, G., (2006), "Analysis of Papsmear image data," *Proceedings of the Nature-Inspired Smart Information Systems 2nd Annual Symposium NISIS*.
- Jati, A. et al., (2014), "Automatic Leukocyte Nucleus Segmentation by Intuitionistic Fuzzy Divergence Based Thresholding", *Micron*, 58, hal.55-65.
- Jemal, A., Bray, F., & Ferlay, J. (2011), *Global Cancer Statistics*, vol.61(2), hal.69-90.
- Kale, A., & Aksoy, S. (2010), "Segmentation of Cervical Cell Images", *IEEE on International Conference Pattern Recognition (ICPR)*.

- Kass, M., Witkin, A., & Terzopoulos, D. (1987), "Snakes: Active contour models", *Int. J. Comput. Vis.*, vol. 1, hal. 321–331.
- Guan, Tao., Zhou, Dongxiang & Liu, Yunhui (2015). "Accurate Segmentation of Partially Overlapping Cervical Cells Based on Dynamic Sparse Contour Searching and GVF Snake Model". *IEEE Journal of Biomedical and Health Informatics*, 19(4), 1494-1504.
- Husain, Nursuci Putri., Faticah, C. (2016), Radiating Component Normalized Generalized Gradient Vector Flow Snake untuk Segmentasi Citra Sel Tunggal Smear Serviks, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Kumar, B.K. Shreyamsha (2013). "Image denoising based on non-local means filter and its method noise thresholding", *Signal image and Video Processing*, 7, 1211-1227.
- Qin, L., Zhu, C., Member, S., Zhao, Y., Member, S., Bai, H., & Tian, H. (2013), "Generalized Gradient Vector Flow for Snakes : New Observations, Analysis, and Improvement", vol.23(5), hal.883–897.
- Li, Kuan., Lu, Zhi., Liu Wenyin & Jianping Yin (2012). "Cytoplasm and nucleus segmentation in cervical smear images using Radiating GVF Snake", *Pattern Recognition*, 45, 1255-1264.
- Wang, Yuanquan (2009). "NBGVF:Normally Biased Gradient Vector Flow External Force for Active Contours", *IEEE Signal Processing Letters*, 17(10), 1-4.
- Xu, Chenyang., Prince, Jerry L.(1998). "Generalized Gradient vector flow external forces for active contours", *Signal Processing*, 71, 131-139.
- Zhang, Ling., Chen, Siping (2014). "Segmentation of Cytoplasm and Nuclei of Abnormal Cells in Cervical Cytology using Global and Local Graph Cuts", *Computerized medical imaging and graphics*, 38, 369-380.
- Zhang, Rui., Zhu, Shiping., & Zhou, Qin (2016). "A novel gradient vector flow snake model based on convex function for infrared image segmentation", *Sensors*.
- A. Zijdenbos, B. Dawant, R. Margolin and A. Palmer, "Morphometric analysis of white matter lesions in MR images: method and validation," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 716-724, 1994

LAMPIRAN 1

Rata-rata akurasi dan ZSI Nukleus dan sitoplasma citra sel tunggal smear serviks.

| Kelas | Metode | Total citra yang dievaluasi | Rata-rata Akurasi | | Rata-rata ZSI | |
|----------------------------|---------------------------------------|-----------------------------|-------------------|------------|---------------|------------|
| | | | Nukleus | Sitoplasma | Nukleus | Sitoplasma |
| <i>Carcinoma in situ</i> | RNBGGVFS & mean shift (metode usulan) | 116 | 0,927 | 0,807 | 0,918 | 0,858 |
| | RCNNGGVFS & FCM_NLS | 67 | 0,918 | 0,799 | 0,906 | 0,849 |
| | RGVFS & FCM | 19 | 0,915 | 0,790 | 0,896 | 0,845 |
| | RGVFS & spatial k-means | 20 | 0,910 | 0,779 | 0,896 | 0,835 |
| <i>Light Dysplastic</i> | RNBGGVFS & mean shift (metode usulan) | 126 | 0,976 | 0,848 | 0,957 | 0,875 |
| | RCNNGGVFS & FCM_NLS | 61 | 0,972 | 0,835 | 0,918 | 0,867 |
| | RGVFS & FCM | 22 | 0,964 | 0,835 | 0,906 | 0,865 |
| | RGVFS & spatial k-means | 19 | 0,967 | 0,834 | 0,916 | 0,866 |
| <i>Moderate Dysplastic</i> | RNBGGVFS & mean shift (metode usulan) | 100 | 0,966 | 0,812 | 0,931 | 0,854 |
| | RCNNGGVFS & FCM_NLS | 51 | 0,959 | 0,804 | 0,926 | 0,847 |
| | RGVFS & FCM | 20 | 0,955 | 0,761 | 0,916 | 0,808 |
| | RGVFS & spatial k-means | 19 | 0,954 | 0,765 | 0,907 | 0,812 |
| <i>Normal Columnar</i> | RNBGGVFS & mean shift (metode usulan) | 67 | 0,925 | 0,765 | 0,836 | 0,826 |
| | RCNNGGVFS & FCM_NLS | 33 | 0,897 | 0,805 | 0,832 | 0,836 |
| | RGVFS & FCM | 20 | 0,895 | 0,795 | 0,834 | 0,830 |
| | RGVFS & spatial k-means | 20 | 0,881 | 0,796 | 0,833 | 0,829 |

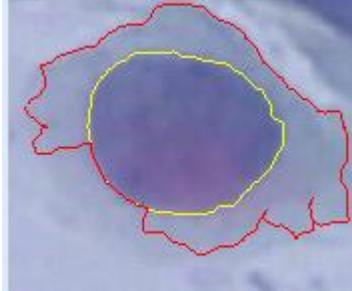
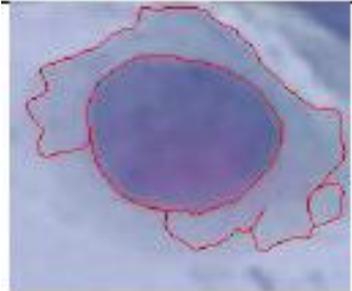
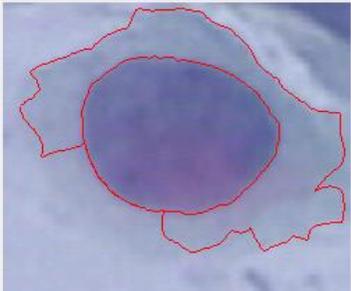
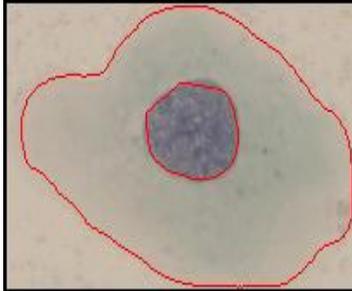
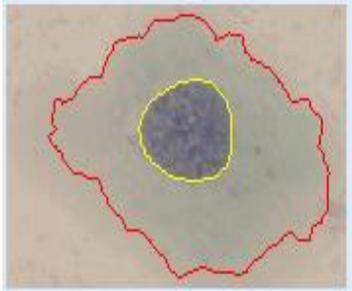
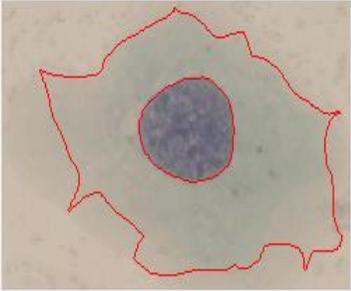
| Kelas | Metode | Total citra yang dievaluasi | Rata-rata Akurasi | | Rata-rata ZSI | |
|----------------------------|---------------------------------------|-----------------------------|-------------------|------------|---------------|------------|
| | | | Nukleus | Sitoplasma | Nukleus | Sitoplasma |
| <i>Normal Intermediate</i> | RNBGGVFS & mean shift (metode usulan) | 28 | 0,996 | 0,918 | 0,867 | 0,935 |
| | RCNNGVFS & FCM_NLS | 20 | 0,993 | 0,937 | 0,869 | 0,951 |
| | RGVFS & FCM | 15 | 0,991 | 0,931 | 0,803 | 0,944 |
| | RGVFS & spatial k-means | 14 | 0,988 | 0,935 | 0,791 | 0,948 |
| <i>Normal Superficiel</i> | RNBGGVFS & mean shift (metode usulan) | 44 | 0,998 | 0,875 | 0,817 | 0,902 |
| | RCNNGVFS & FCM_NLS | 3 | 0,995 | 0,852 | 0,797 | 0,895 |
| | RGVFS & FCM | 3 | 0,992 | 0,841 | 0,738 | 0,889 |
| | RGVFS & spatial k-means | 3 | 0,899 | 0,838 | 0,817 | 0,887 |
| <i>Severe Dysplastic</i> | RNBGGVFS & mean shift (metode usulan) | 159 | 0,945 | 0,823 | 0,918 | 0,857 |
| | RCNNGVFS & FCM_NLS | 88 | 0,939 | 0,811 | 0,916 | 0,856 |
| | RGVFS & FCM | 17 | 0,930 | 0,775 | 0,905 | 0,839 |
| | RGVFS & spatial k-means | 17 | 0,931 | 0,768 | 0,905 | 0,833 |

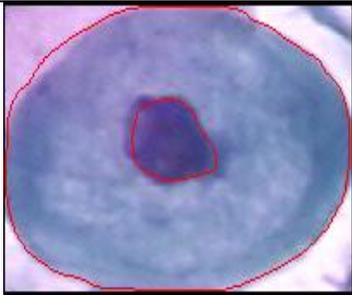
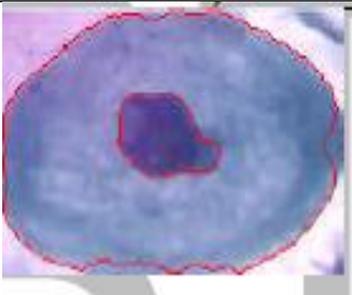
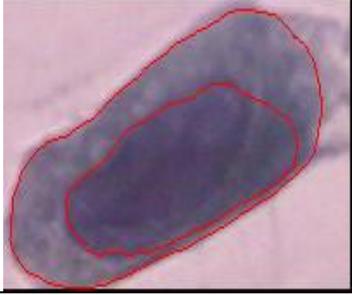
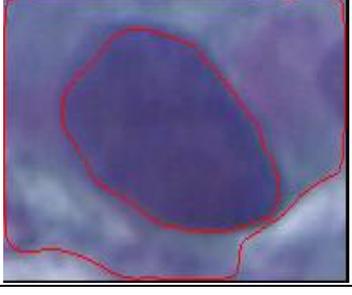
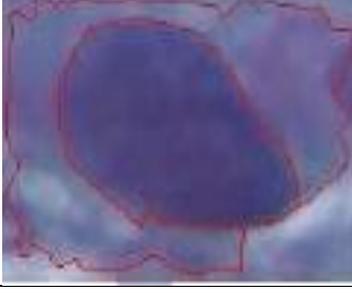
Nilai rata-rata Akurasi dan ZSI dari seluruh kelas pada tiap metode.

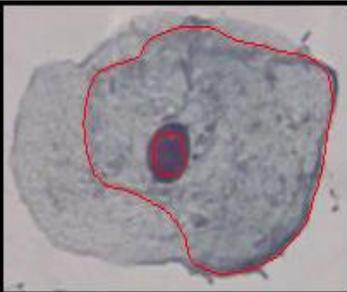
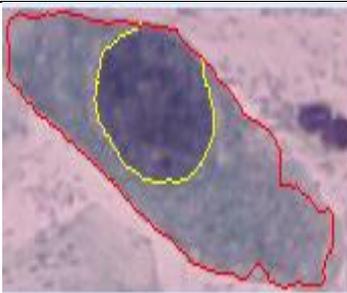
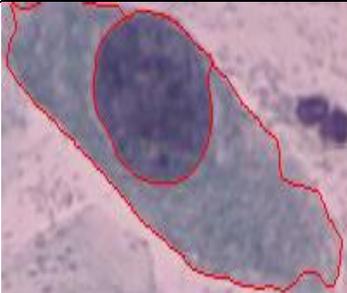
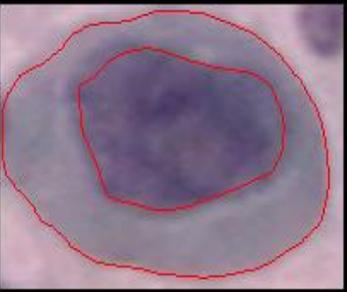
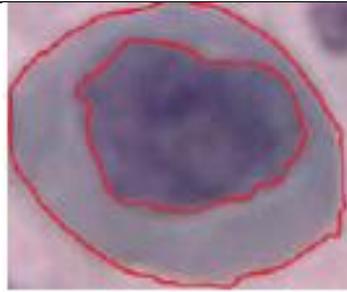
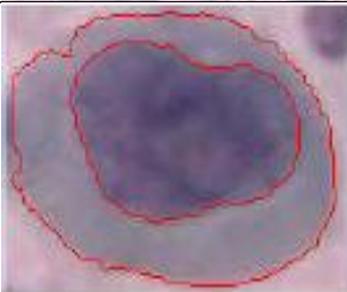
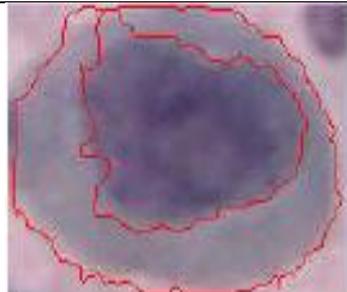
| Metode | Rata-rata akurasi nukleus | Rata-rata akurasi sitoplasma | Rata-rata ZSI nukleus | Rata-rata ZSI sitoplasma |
|------------------------------------|---------------------------|------------------------------|-----------------------|--------------------------|
| <i>RNBGGVFS & mean shift</i> | 0,969 | 0,86 | 0,907 | 0,893 |
| <i>RCNNGVFS & FCM_NLS</i> | 0,953 | 0,835 | 0,881 | 0,871 |
| <i>RGVFS & FCM</i> | 0,948 | 0,818 | 0,857 | 0,860 |
| <i>RGVFS & spatial k-means</i> | 0,945 | 0,817 | 0,866 | 0,858 |

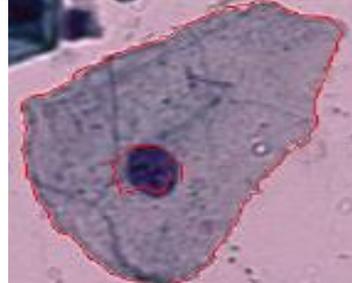
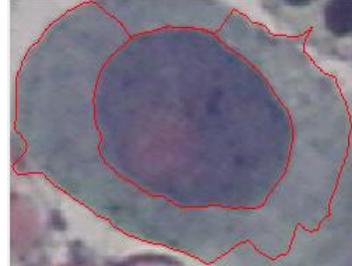
LAMPIRAN 2

Contoh hasil segmentasi nukleus dan sitoplasma citra sel tunggal smear serviks menggunakan metode yang diusulkan dan beberapa metode pada penelitian sebelumnya.

| No | Citra Asli | RNBGGVFS & mean shift (metode usulan) | RCNNGVFS & FCM_NLS | RGVFS & FCM | RGVFS & spatial k-means |
|----|--|---|--|---|--|
| 1 |  |  |  |  |  |
| 2 |  |  |  | GAGAL |  |

| No | Citra Asli | RNBGGVFS & mean shift (metode usulan) | RCNGGVFS & FCM_NLS | RGVFS & FCM | RGVFS & spatial k-means |
|----|--|---|--|---|-------------------------|
| 3 |  |  |  | GAGAL | GAGAL |
| 4 |  |  |  |  | GAGAL |
| 5 |  |  |  | GAGAL | GAGAL |

| No | Citra Asli | RNBGGVFS & mean shift (metode usulan) | RCNGGVFS & FCM_NLS | RGVFS & FCM | RGVFS & spatial k-means |
|----|--|---|--|--|--|
| 6 |  |  | GAGAL |  | GAGAL |
| 7 |  |  |  |  | GAGAL |
| 8 |  |  |  |  |  |

| No | Citra Asli | RNBGGVFS & mean shift (metode usulan) | RCNGGVFS & FCM_NLS | RGVFS & FCM | RGVFS & spatial k-means |
|----|---|---|---|---|---|
| 9 |  |  |  |  |  |
| 10 |  |  |  | GAGAL |  |

BIODATA PENULIS



Martini Dwi Endah Susanti. Anak kedua dari tiga bersaudara, lahir 16 Maret 1993 di kota Salatiga, Jawa Tengah. Penulis menempuh pendidikan mulai dari TK Tarbiyatul Banin 2 Salatiga (1998-1999), SD Negeri Mojo VIII/227 Surabaya (1999-2005), SMP Negeri 6 Surabaya (2005-2006), SMP Negeri 7 Medan (2006-2008), SMA Negeri 4 Medan (2008-2011) dan S1 Teknik Informatika ITS (2011-2015). Kemudian penulis melanjutkan pendidikan S2 jurusan Teknik Informatika di Institut Teknologi Sepuluh Nopember (ITS) Surabaya pada tahun 2016. Penulis dapat dihubungi melalui email: martini11@mhs.if.its.ac.id.

(Halaman ini sengaja dikosongkan)