

# Segmentasi Citra Sel Tunggal *Smear* Serviks Menggunakan *Radiating Component Normalized Generalized GVFS*

Nursuci Putri Husain<sup>1</sup>, Chastine Fatichah<sup>2</sup>

**Abstract**—Component Normalized Generalized Gradient Vector Flow Snake (CNGGVFS) method is the development of Gradient Vector Flow Snake (GVFS) method as an external force algorithm for active contour (snake) that can be used to get the contour of nucleus and cytoplasm of cervical smear image. However, CNGGVFS using a conventional calculation of edge map such as Sobel can not detect the nucleus area correctly in single cell cervical smear image segmentation. In this study, an external force algorithm in snake that uses Radiating Edge Map (REM) calculation to search the edge map in CNGGVFS, called as Radiating Component Normalized Generalized Gradient Vector Flow Snake (RCNGGVFS), is proposed. RCNGGVFS is used to get the contour of nucleus and cytoplasm of single cervical smear image. There are three main stages in this study, which are: pre-processing, initial segmentation, and contour segmentation. Experiments are conducted on Herlev data-set. The proposed method is compared with other methods in previous research in single cell cervical smear image segmentation. The experiment results show that the proposed method can detect the nucleus area correctly better than Radiating GVFS & Fuzzy C-Means (FCM) and Radiating GVFS & K-means. The average value of accuracy and Zijdenbos similarity index (ZSI) for nucleus segmentation is 95.34% and 88.06%. Then, the average value of accuracy and ZSI for cytoplasm segmentation is 83.48% and 87.16%. The evaluations show the proposed method can be used as a segmentation process of cervical smear image on automatic identification of cervical cancer.

**Intisari**—Metode *Component Normalized Generalized Gradient Vector Flow Snake* (CNGGVFS) merupakan pengembangan dari metode *Gradient Vector Flow Snake* (GVFS) sebagai algoritme pengganti energi eksternal *active contour* (snake). Namun, CNGGVFS menggunakan perhitungan *edge map* konvensional seperti *Sobel* yang belum mampu mendeteksi daerah nukleus dengan benar pada segmentasi kontur citra sel tunggal *smear* serviks. Makalah ini mengusulkan algoritme pengganti energi eksternal pada *snake* yang memanfaatkan perhitungan *Radiating Edge Map* (REM) untuk pencarian *edge map* pada CNGGVFS yang dinamakan *Radiating Component Normalized Generalized Gradient Vector Flow Snake* (RCNGGVFS). Metode RCNGGVFS digunakan untuk mendapatkan kontur nukleus dan sitoplasma pada citra sel tunggal *smear* serviks. Terdapat tiga tahapan utama untuk melakukan segmentasi kontur citra sel tunggal *smear* serviks pada makalah ini, yaitu: pra proses, segmentasi awal, dan segmentasi kontur. Uji coba dilakukan

dengan menggunakan *data set* Herlev. Pengujian dilakukan dengan membandingkan hasil segmentasi metode yang diusulkan dengan metode pada penelitian sebelumnya dalam melakukan segmentasi citra sel tunggal *smear* serviks. Hasil pengujian menunjukkan bahwa metode yang diusulkan mampu mendeteksi area nukleus lebih optimal dibanding metode *Radiating GVFS & Fuzzy C-Means* (FCM) dan *Radiating GVFS & K-means*. Nilai rata-rata akurasi dan *Zijdenbos similarity index* (ZSI) untuk segmentasi nukleus adalah 95,34% dan 88,06%. Kemudian, nilai rata-rata akurasi dan ZSI untuk segmentasi sitoplasma adalah 83,48% dan 87,16%. Dari hasil evaluasi tersebut, disimpulkan metode yang diusulkan dapat digunakan sebagai proses segmentasi citra *smear* serviks pada identifikasi kanker serviks secara otomatis.

**Kata Kunci**— *Radiating Component Normalized Generalized Gradient Vector Flow Snake*, sel tunggal *smear* serviks, ekstraksi kontur.

## I. PENDAHULUAN

Pemeriksaan *pap smear* merupakan prosedur penapisan manual yang digunakan untuk mendeteksi kanker serviks yang dikembangkan oleh Dr. Georges Papanicolaou pada Tahun 1940. 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.

Sebelumnya telah diusulkan sebuah sistem klasifikasi citra *smear* serviks menggunakan metode *Fuzzy k-Nearest Neighbor* dan memanfaatkan fitur bentuk serta fitur tekstur citra [1]. Hasil klasifikasi citra *smear* serviks yang diperoleh pada penelitian tersebut untuk seluruh kelas pada *data set* Herlev adalah 70,06 %. Hasil yang didapatkan berlum terlalu akurat, sehingga perlu dilakukan perbaikan pada tahapan segmentasi. Beberapa peneliti telah mengusulkan metode untuk segmentasi sel nukleus dan sitoplasma dari citra *smear* serviks. Metode berbasis *cluster* yaitu *Possibilistic Fuzzy C-Means* diusulkan untuk melakukan segmentasi pada citra *smear* serviks [2]. Kemudian, beberapa peneliti mengusulkan teknik *thresholding* untuk proses segmentasi sel nukleus dan sitoplasma dari citra *smear* serviks [3]. Ada juga yang menggunakan metode *watershed* [4]. Namun, metode-metode yang menggunakan teknik *thresholding* dan *watershed* tersebut belum mendapatkan hasil yang memuaskan karena

<sup>1</sup>Mahasiswa, Jurusan Teknik Informatika Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo Surabaya 60111 INDONESIA, (tlp: 031-5939214; fax: 031-5939363; e-mail: nursuci.husain15@mhs.if.its.ac.id)

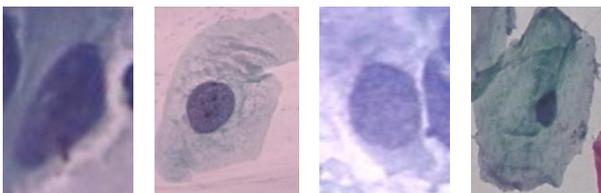
<sup>2</sup>Dosen, Jurusan Teknik Informatika Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo Surabaya 60111 INDONESIA (tlp: 031-5939214; fax: 031-5939363; e-mail: chastine@cs.its.ac.id)

kompleksitas citra *smear* serviks. Gbr. 1 merupakan contoh citra sel tunggal *smear* serviks dari *data set* Herlev.

Algoritme berbasis *active contour* atau *snake* merupakan salah satu algoritme yang dapat digunakan dalam proses segmentasi citra *smear* serviks untuk mendapatkan bagian nukleus, sitoplasma, dan *background* [5]. *Active contour* atau *snake* dapat dihitung dengan meminimalkan 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 [6]. Untuk mengatasi masalah tersebut *snake* dikembangkan menjadi *Gradient Vector Flow Snake* (GVFS) sebagai pengganti energi eksternal *snake* [6]. Energi eksternal ini digunakan untuk mengatasi kesulitan dalam mendapatkan batas kecekungan. Metode GVFS ini digunakan untuk segmentasi sel nukleus dan sitoplasma dari citra *smear* serviks [7]. Namun, GVFS memiliki konvergensi yang lemah terhadap batas kecekungan objek [6], [8].

Kemudian, diusulkan *Generalized Gradient Vector Flow Snake* (GGVFS) yang bertujuan untuk meningkatkan konvergensi GVFS [6]. 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 objek [8]. Kelemahan konvergensi terhadap batas kecekungan objek yang dihadapi GVFS dan GGVFS dikarenakan konsistensi komponen energi eksternal *snake* yang diperlukan untuk menarik *snake* ke arah objek secara signifikan lebih kecil dalam besaran (*magnitude*) sehingga dapat mengganggu proses difusi pada normalisasi vektor. Untuk mengatasi masalah tersebut, metode *Component Normalized Generalized Gradient Vector Flow Snake* (CNGGVFS) diusulkan. CNGGVFS didapatkan dengan melakukan normalisasi tiap komponen dari inialisasi vektor GGVFS dengan memperhatikan besaran [8]. Komponen yang lebih kecil harus ditingkatkan sehingga setiap komponen dapat memengaruhi proses pembentukan *snake*.

Dalam membentuk *snake*, terdapat fungsi *edge map*. *Edge map* dihitung dengan menggunakan operator deteksi tepi seperti *Sobel*. Namun, metode ini tidak dapat mendeteksi daerah nukleus dari citra *smear* serviks dengan benar [5]. Hal ini disebabkan area nukleus umumnya lebih gelap daripada area sitoplasma, dan area sitoplasma lebih gelap dibandingkan area *background*. Sebelumnya, diusulkan perhitungan *gradient edge map* pada metode GVFS menggunakan *Radiating Edge Map* (REM) [5]. 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.



Gbr. 1 Contoh citra sel tunggal *smear* serviks [9].

Metode CNGGVFS sebagai fungsi energi eksternal *snake* dapat digunakan untuk melakukan segmentasi kontur nukleus dan sitoplasma citra sel tunggal *smear* serviks. Namun, CNGGVFS menggunakan perhitungan *edge map* konvensional seperti *Sobel*. Perhitungan *edge map* menggunakan metode konvensional belum mampu mendeteksi daerah nukleus dengan benar [5].

Makalah ini mengusulkan metode baru untuk segmentasi citra sel tunggal *smear* serviks menggunakan *Radiating Component Normalized Generalized Gradient Vector Flow Snake* (RCNGGVFS). RCNGGVFS memanfaatkan perhitungan REM dalam pencarian *edge map* pada metode CNGGVFS. CNGGVFS sebagai fungsi energi eksternal *snake* didapatkan dengan melakukan normalisasi tiap komponen dari inialisasi vektor GGVFS dengan memperhatikan besaran. Adapun perhitungan REM dimanfaatkan untuk mendapatkan *edge map* citra *smear* serviks yang dilakukan di sepanjang garis pancaran. Terdapat tiga tahapan utama pada makalah ini, yaitu pra proses, segmentasi awal, dan segmentasi kontur. Pada tahap pra proses, citra *smear* serviks yang berada pada ruang warna RGB dikonversi ke dalam ruang warna CIELAB, dan pada kanal L dilakukan normalisasi *range* untuk mendapatkan citra *grayscale*. Proses segmentasi awal pada makalah ini menggunakan metode *Fuzzy C-Means Non Local Spatial* (FCM\_NLS) untuk mendapatkan tekstur dari citra sel tunggal *smear* serviks, yaitu nukleus, sitoplasma, dan *background*. Selanjutnya, tahap terakhir adalah tahap segmentasi kontur. Pada tahap ini, diusulkan metode RCNGGVFS sebagai fungsi pengganti energi eksternal *snake* yang bertujuan untuk mendapatkan kontur nukleus dan sitoplasma citra sel tunggal *smear* serviks.

Uji coba dilakukan dengan menggunakan *data set* Herlev, yang terdiri atas 917 citra sel tunggal *smear* serviks [9], [10]. Uji coba pertama adalah membandingkan nilai rata-rata akurasi dan *Zijdenbos similarity index* (ZSI) metode yang diusulkan dengan metode CNGGVFS dan menggunakan segmentasi awal yang sama yaitu FCM\_NLS. Kemudian, pengujian kedua adalah membandingkan nilai rata-rata akurasi dan ZSI metode yang diusulkan dengan metode pada penelitian sebelumnya. Dengan menggunakan metode yang diusulkan, diharapkan mampu diperoleh hasil segmentasi kontur nukleus dan sitoplasma pada citra sel tunggal *smear* serviks yang lebih optimal.

## II. SEGMENTASI

### A. *Generalized Gradient Vector Flow Snake*

*Active contour* atau *snake* merupakan kurva yang didefinisikan dalam domain citra yang bergerak di bawah pengaruh energi internal dari kurva itu sendiri dan energi eksternal yang dihitung dari data citra [6]. Energi internal dan eksternal dibuat sedemikian rupa sehingga *snake* akan menuju ke batas suatu objek atau fitur lain yang diinginkan.

Kelemahan *snake* adalah kontur awal harus ditempatkan dekat dengan objek dan konvergensi *snake* sangat lemah terhadap batas kecekungan objek. Untuk mengatasi masalah tersebut, *snake* dikembangkan menjadi GVFS sebagai

pengganti energi eksternal ( $-\nabla E_{\text{ext}}$ ) yang mengombinasikan aliran *optic* dan *gradient* keabuan.

GVFS memiliki konvergensi yang lemah terhadap batas kecekungan. Kelemahan ini disebabkan *smoothing* yang berlebihan pada area *boundary*, yang diatur oleh koefisien  $\mu$  [8]. Untuk mengatasi masalah ini, ditambahkan dua koefisien pembobot pada persamaan energi eksternal GVFS [8]. Persamaan energi eksternal baru tersebut dinamakan GGVFS. Persamaan energi dari GGVFS ada pada (1).

$$E_{GGVFS} = \iint \underbrace{g(|\nabla f|)|\nabla V|^2}_{\text{smoothness energy}} + \underbrace{h(|\nabla f|)|V - \nabla f|^2}_{\text{edge energy}} dx dy \quad (1)$$

dengan,

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

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

$$|\nabla f| = f_x + f_y$$

$$|V - \nabla f|^2 = (u - f_x)^2 + (v - f_y)^2$$

$$|\nabla V|^2 = u_x^2 + v_x^2 + u_y^2 + v_y^2.$$

Term pertama pada bagian kanan dari (1) merupakan term *smoothing* yang akan menghasilkan vektor, dengan  $g(|\nabla f|)$  merupakan pengganti dari  $\mu$  (koefisien *smoothing*) pada GVFS. Term kedua adalah term data atau disebut juga *edge energy* yang akan membawa vektor  $V$  dekat dengan *gradient edge map* ( $\nabla f$ ) dari citra, dengan  $h(|\nabla f|)$  merupakan pengganti dari  $|\nabla f|^2$  pada GVFS. Parameter  $k$  menentukan nilai koefisien pembobot untuk term *smoothing* dan data. Nilai dari  $k$  berhubungan dengan kekuatan derau dari sebuah citra. Semakin tinggi level derau, semakin banyak nilai  $k$ . Persamaan energi (1) dapat diselesaikan dengan menerapkan  $u$  dan  $v$  sebagai fungsi waktu  $t$ :

$$u_t(x, y, t) = g(|\nabla f|)\nabla^2 u(x, y, t) - h(|\nabla f|)[u(x, y, t) - f_x(x, y)]. \quad (2)$$

$$v_t(x, y, t) = g(|\nabla f|)\nabla^2 v(x, y, t) - h(|\nabla f|)[v(x, y, t) - f_y(x, y)]. \quad (3)$$

Pada (1) terdapat *gradient edge map* ( $\nabla f$ ) yang dapat dihitung menggunakan operator *gradient* apa pun. Jika term data pada (1) didefinisikan vektor  $V$  dekat dengan *gradient edge map* ( $\nabla f$ ) dari citra, maka parameter  $k$  harus  $> 0$ .

### B. Component Normalized Generalized Gradient Vector Flow Snake

Berdasarkan penelitian terhadap fungsi eksternal GGVFS, GVFS dan GGVFS memiliki kekuatan konvergensi yang hampir sama [8]. Secara umum, GGVFS masih sensitif terhadap derau jika bekerja pada lekukan yang tipis dan panjang. Masalah tersebut dapat diselesaikan dengan melakukan normalisasi pada vektor energi eksternal menggunakan dua komponen. Komponen yang lebih kecil

akan ditingkatkan sehingga setiap komponen dapat memengaruhi perubahan bentuk *snake*.

Energi eksternal baru yang disebut CNGGVFS memiliki vektor baru yaitu  $v_{cn}(x, y) = u_{cn}(x, y), v_{cn}(x, y)$ , dan komponen-komponen yang ditetapkan adalah:

$$u_{cn}(x, y) = \begin{cases} 1 & u(x, y) > 0 \\ 0 & u(x, y) = 0 \\ -1 & u(x, y) < 0 \end{cases} \quad (4)$$

$$v_{cn}(x, y) = \begin{cases} 1 & v(x, y) > 0 \\ 0 & v(x, y) = 0 \\ -1 & v(x, y) < 0 \end{cases} \quad (5)$$

dengan  $u(x, y)$  dan  $v(x, y)$  adalah komponen  $x$  dan  $y$  berturut-turut dari energi eksternal  $(x, y)$  pada vektor GGVFS.

### C. Radiating Component Normalized Generalized Gradient Vector Flow Snake

*Edge map* ( $\nabla f$ ) yang dihitung dengan menggunakan operator deteksi tepi seperti *Sobel*, ternyata tidak dapat menggunakan informasi *directional* sepenuhnya. Metode tersebut tidak dapat mendeteksi daerah nukleus dari citra *smear* serviks dengan benar [5]. Hal ini dikarenakan area nukleus umumnya lebih gelap daripada area sitoplasma, dan area sitoplasma lebih gelap daripada area *background*.

Oleh karena itu, pada makalah ini dimanfaatkan perhitungan REM untuk pencarian *edge map* pada metode CNGGVFS. REM merupakan proses perhitungan *gradient edge map* pada sepanjang garis pancaran yang dimulai dari *intensity-weighted area centroid* nukleus hasil segmentasi awal. Maka, area nukleus yang didapatkan lebih akurat dibanding menggunakan metode pencarian *edge map* konvensional. Tahap-tahap perhitungan REM dapat dijabarkan seperti berikut.

1. Tahap pertama pada perhitungan REM adalah mengekstraksi setiap garis pancaran.  $I(x, y)$  merupakan nilai piksel pada  $(x, y)$ , dan  $(x_c, y_c)$  merupakan 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  $x_b, y_b$ . Panjang dari garis  $L_{x_b, y_b}$  dapat dihitung menggunakan (6).

$$\text{len}_b = \lfloor \sqrt{(x_b - x_c)^2 + (y_b - y_c)^2} \rfloor. \quad (6)$$

2. Kemudian, koordinat-koordinat titik tidak selalu integer, sehingga nilai  $I(x, y)$  dihitung dengan interpolasi linear berdasarkan piksel citra. *Radiating difference*  $rd(x_i, y_i)$  pada titik  $(x_i, y_i)$  sepanjang  $L_{x_b, y_b}$  dapat dihitung menggunakan (7).

$$rd(x_i, y_i) = I_o(x_{i-1}, y_{i-1}) - I_o(x_i, y_i). \quad (7)$$

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 garis pancaran. *Radiating*

*difference* dapat diperbaiki menggunakan (8). Pada makalah ini parameter *stackrefine theta* dan *positive suppress* yang digunakan adalah 2 dan 0,3. Koefisien  $\gamma$  adalah bobot yang bernilai  $< 1$ .

$$F_s(x) = \begin{cases} x, & x < 0 \\ \gamma x, & x \geq 0 \end{cases} \quad (8)$$

RCNGGVFS suatu citra dihitung dengan cara meminimalkan fungsi energi GGVFS pada (1), dimulai dengan melakukan normalisasi untuk tiap komponen GGVFS. Komponen-komponen yang dinormalisasi tersebut terdapat pada (4) dan (5). Kemudian, RCNGGVFS pada proses pencarian *edge map* ( $\nabla f$ ) memanfaatkan perhitungan REM. Merujuk pada metode *Radiating Gradient Vector Flow Snake* (RGVFS) [5], yaitu  $\nabla f$  (*gradient edge map*) menggunakan deteksi tepi konvensional, diganti dengan  $\nabla REM$  yaitu pencarian *gradient edge map* menggunakan perhitungan REM, sehingga didapatkan fungsi energi RCNGGVFS seperti (9).

$$E_{RCNGGVFS} = \iint \underbrace{g(|\nabla REM|)|\nabla V|^2}_{\text{smoothness energy}} + \underbrace{h(|\nabla REM|)|V - \nabla REM|^2}_{\text{edge energy}} dx dy \quad (9)$$

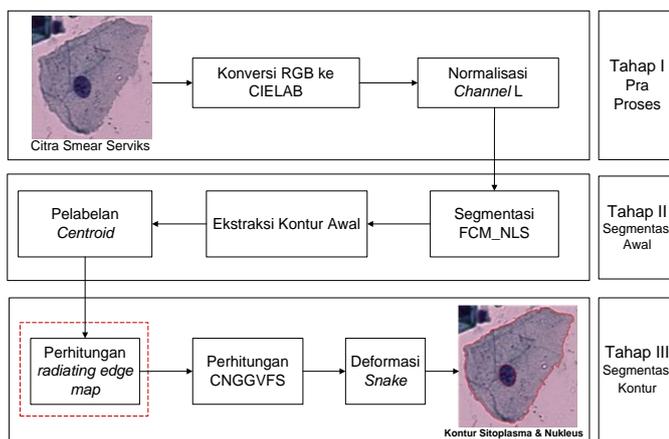
dengan,

$$g(|\nabla REM|) = e^{-|\nabla REM|/k}$$

$$h(|\nabla REM|) = 1 - g(|\nabla REM|).$$

### III. METODOLOGI

Segmentasi citra sel tunggal *smear* serviks pada makalah ini terdiri atas tiga tahapan utama yaitu tahap pra proses, tahap segmentasi awal, dan tahap segmentasi kontur. Adapun desain sistem yang diusulkan dan bagian kontribusi ditampilkan pada Gbr. 2.



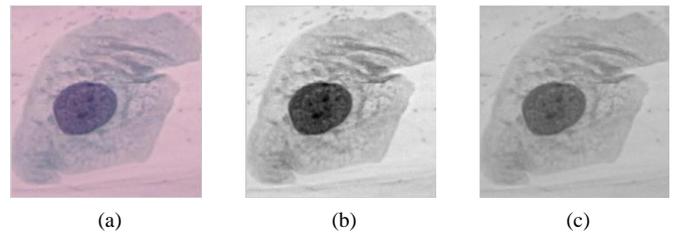
Gbr. 2 Desain sistem yang diusulkan.

#### A. Pra Proses

Citra *smear* serviks memiliki kontras yang rendah, pewarnaan yang tidak seragam, serta mengandung derau yang dapat menghambat proses segmentasi. Diperlukan adanya

perbaikan citra untuk meningkatkan akurasi proses segmentasi citra. Berdasarkan Gbr. 2, dapat dilihat ada beberapa tahapan pada tahap pra proses.

Pertama, citra *smear* serviks yang berada pada ruang warna RGB dikonversi ke dalam ruang warna CIELAB. Ruang warna ini terdiri atas tiga kanal. Kanal L merupakan dimensi pencahayaan, sedangkan kanal a\* dan b\* merupakan dimensi warna. CIELAB memisahkan kekontrasan dan warna penyusun citra. Kanal yang digunakan pada proses segmentasi ini adalah kanal L. Transformasi citra *smear* serviks dari ruang warna RGB ke ruang warna CIELAB dimulai dengan melakukan ekstraksi terhadap kanal L. Hasil ekstraksi dimensi L mirip dengan citra *grayscale*, tetapi hasil dari citra *grayscale* agak lebih gelap, seperti ditunjukkan pada Gbr. 3.



Gbr. 3 Perbandingan citra *smear* serviks (a) citra asli, (b) citra CIELAB kanal L, dan (c) citra *grayscale*.

Kemudian, kanal L dinormalisasi ke dalam *range* [0,255] untuk membentuk citra keabuan. Normalisasi dilakukan karena hasil ekstraksi kanal L tidak berada pada *range* [0,255]. Setelah dilakukan normalisasi *range*, maka akan didapatkan citra keabuan.

#### B. Segmentasi Awal

Tahap segmentasi awal terdiri atas tiga tahapan, yaitu segmentasi FCM\_NLS, ekstraksi kontur awal, dan pelabelan *centroid*.

1) *Segmentasi FCM\_NLS*: Pada tahap pra proses seharusnya dilakukan penghilang derau pada citra *smear* serviks, karena pada citra *smear* serviks terdapat derau *Gaussian* yang dapat menghambat proses segmentasi [5]. Namun, pada tahap segmentasi awal ini diusulkan metode segmentasi pengelompokan menggunakan FCM\_NLS [11]. Metode ini tegar (*robust*) terhadap citra yang memiliki derau dengan densitas tinggi, sehingga pada tahap pra proses tidak dilakukan penghilang derau. Hal ini dikarenakan pada metode FCM\_NLS telah terdapat proses filter derau dengan memanfaatkan informasi spasial non lokal atau biasa disebut dengan *non local means* [12].

Langkah pertama yang dilakukan adalah menentukan nilai parameter-parameter yang digunakan pada proses segmentasi. Parameter-parameter yang digunakan seperti  $T$ ,  $h$ ,  $c$ ,  $r$ , dan  $s$  untuk proses filter *non local means*. Parameter  $T$  merupakan jumlah maksimum iterasi,  $h$  merupakan parameter yang mengontrol fungsi bobot  $w_{ij}$ ,  $c$  merupakan jumlah *cluster*,  $r$  merupakan ukuran pencarian *window*, dan  $s$  merupakan ukuran dari ketetanggaan piksel yang digunakan pada proses filter *non local means*. Pada makalah ini, parameter  $T$  yang digunakan adalah 100,  $h=1$ ,  $c=3$ ,  $r=3$ , dan  $s=3$ . Nilai

parameter tersebut diperoleh setelah melalui uji coba dan didapatkan hasil segmentasi yang paling optimal. Setelah parameter yang dibutuhkan ditentukan, dilakukan proses filter *non local means*  $\bar{x}_i'$  untuk piksel ke- $i$  menggunakan (10).

$$\bar{x}_i' = \sum_{j \in W_i^r} w_{ij} x_j \quad (10)$$

dengan  $W_i^r$  merupakan *window* pusat yang berukuran  $rxr$  pada piksel ke- $i$ . Piksel pada domain ini digunakan untuk menghitung informasi spasial pada piksel ke- $i$ . Bobot  $w_{ij}$  ( $j \in W_i^r$ ) mengandalkan kesamaan antara piksel ke- $i$  dan piksel ke- $j$ , dengan  $0 \leq w_{ij} \leq 1$  dan  $\sum_{j \in W_i^r} w_{ij} = 1$ . Selanjutnya, ditentukan inisialisasi pusat kluster  $v_k$ . Setelah itu, fungsi keanggotaan  $u_{ki}$  diperbarui menggunakan (11).

$$u_{ki} = \frac{1}{\sum_{l=1}^c \left( \frac{\|x_i - v_k\|^2 + \beta \|x_i' - v_k\|^2}{\|x_i - v_k\|^2 + \beta \|x_i' - v_k\|^2} \right)^{1/(m-1)}} \quad (11)$$

Parameter  $m$  adalah eksponen pembobotan pada setiap keanggotaan *fuzzy* yang menentukan jumlah *fuzziness* dari kelompok yang dihasilkan. Kemudian, parameter  $\beta$  mengontrol efek akhir dari ketetanggaan spasial dan  $n$  merupakan jumlah piksel. Pusat *cluster*  $v_k$  dihitung menggunakan (12) hingga konvergen [11].

$$v_k = \frac{\sum_{i=1}^n u_{ki}^m |x_i - v_k|^2}{(1+\beta) \sum_{i=1}^n u_{ki}^m} \quad (12)$$

2) *Ekstraksi Kontur Awal*: Setelah mendapatkan citra *smear* serviks yang tersegmentasi, selanjutnya dilakukan proses ekstraksi kontur awal. Pada proses ini diasumsikan bahwa area yang memiliki kecekungan terluas disebut sitoplasma. Untuk kandidat nukleus dapat diasumsikan bahwa nukleus merupakan area yang terdapat di dalam sitoplasma dan area dengan *threshold* kurang dari nilai *threshold*  $\delta$  dari citra dianggap sebagai derau dan akan diabaikan. Sisanya merupakan kandidat area nukleus.

Ukuran nukleus berbeda-beda pada tiap citra *smear* serviks. Nilai *threshold*  $\delta$  ditentukan sebesar 0,5 berdasarkan analisis statistika pada [5], agar dapat menghilangkan area yang mengandung derau saat proses ekstraksi kontur dilakukan. Kontur awal diupayakan semakin dekat dengan objek, sehingga *snake* pada proses segmentasi kontur lebih mudah untuk mengikuti dan mendeteksi bentuk objek.

3) *Pelabelan Centroid*: Proses selanjutnya, setelah mendapatkan kontur awal, adalah melakukan pelabelan terhadap *centroid* dengan bobot intensitas. Proses ini penting untuk membentuk REM pada tahap segmentasi kontur.

$$\begin{aligned} x_c &= \frac{\sum_{i=1}^k x_i (255 - I(x_i - y_i))}{\sum_{i=1}^k (255 - I(x_i - y_i))} \\ y_c &= \frac{\sum_{i=1}^k y_i (255 - I(x_i - y_i))}{\sum_{i=1}^k (255 - I(x_i - y_i))} \end{aligned} \quad (13)$$

Dianggap area nukleus terdiri atas  $k$  titik  $(x_1, y_1)$ ,  $(x_2, y_2)$ , ...,  $(x_k, y_k)$ , kemudian,  $I(x, y)$  adalah intensitas dari

titik  $(x, y)$ , sehingga bobot *centroid*  $(x_c, y_c)$  dapat dihitung dengan (13).

### C. Segmentasi Kontur

Tahap segmentasi kontur menggunakan metode RCNGGVFS terdiri atas beberapa tahap yaitu perhitungan REM, perhitungan CNGGVFS, dan deformasi *snake*.

1) *Perhitungan REM*: Pada tahap ini, citra hasil pra proses dijadikan sebagai masukan untuk mendapatkan *gradient edge map*. Proses pertama adalah mengekstraksi setiap garis pancaran menggunakan (6). Kemudian, *radiating difference* sepanjang garis dihitung menggunakan (7). Proses terakhir pada tahap ini adalah *radiating difference* dapat diperbaiki menggunakan (8).

2) *Perhitungan CNGGVFS*: CNGGVFS suatu citra dihitung dengan cara meminimalkan fungsi energi GGVFS pada (1). Proses pertama pada tahap ini adalah melakukan normalisasi untuk tiap komponen GGVFS. Komponen-komponen vektor yang dinormalisasi tersebut terdapat pada (4) dan (5). Selanjutnya, fungsi energi eksternal GGVFS diminimalkan dengan menggunakan (9).

3) *Deformasi Snake*: *Active contour* atau *snake* dapat dihitung dengan meminimalkan fungsi energi eksternal dan internal yang dihitung dari data citra. Energi eksternal *snake* telah dihitung menggunakan metode RCNGGVFS. Energi internal *snake* dapat dihitung dengan langkah sebagai berikut.

- Menentukan kontur awal yang telah didapat dari proses segmentasi awal.
- Menentukan parameter-parameter *snake* yang dibutuhkan seperti *alpha* (*elasticity parameter*), *beta* (*rigidity parameter*), *gamma* (*viscosity parameter*), dan *kappa* (*external force weight*).
- Tiap titik *snake* akan bergerak secara berulang mencari lokasi yang memiliki energi *snake* minimum.

## IV. UJI COBA DAN ANALISIS HASIL

Setelah mengimplementasikan metode yang diusulkan, tahap selanjutnya adalah melakukan uji coba. Uji coba dilakukan untuk melihat unjuk kerja yang dihasilkan oleh metode yang diusulkan. Data masukan pada uji coba ini berupa citra mikroskopik sel tunggal *smear* serviks. Data yang digunakan diperoleh dari Departemen Patologi Rumah Sakit Universitas Herlev dan Departemen Automasi Universitas Teknik Denmark. *Data set* Herlev ini terdiri atas tujuh kelas yang berbeda dengan jumlah 917 citra sel tunggal *smear* serviks [9]. Data citra sel tunggal *smear* serviks ini dapat diakses di internet secara gratis [10]. Nama-nama kelas pada citra sel tunggal *smear* serviks adalah *Carcinoma in situ*, *Light Dysplastic*, *Moderate Dysplastic*, *Normal Columnar*, *Normal Intermediate*, *Normal Superficial*, dan *Severe Dysplastic*. Setiap satu citra sel tunggal *smear* serviks memiliki ukuran yang berbeda, berada dalam ruang warna RGB, dan berformat BMP. Terdapat dua skenario uji coba yang dilakukan pada makalah ini, yaitu sebagai berikut.

1. Membandingkan hasil segmentasi kontur citra sel tunggal *smear* serviks menggunakan metode yang diusulkan dengan hasil segmentasi menggunakan metode CNGGVFS dengan segmentasi awal yang sama yaitu FCM\_NLS.
2. Membandingkan hasil segmentasi kontur citra sel tunggal *smear* serviks menggunakan metode yang diusulkan dengan metode pada penelitian sebelumnya [5], [13].

Evaluasi uji coba dilakukan dengan cara membandingkan citra hasil segmentasi dengan citra segmentasi manual (*groundtruth*). Evaluasi dilakukan dengan dua cara, yaitu mencari nilai ZSI dan akurasi menggunakan *confusion matrix*. ZSI digunakan untuk menghitung seberapa besar rasio *overlap* antara dua citra. ZSI didefinisikan seperti pada (14).

$$ZSI = 2 \frac{|A_M \cap A_G|}{|A_M| + |A_G|} \tag{14}$$

$A_M$  merupakan citra hasil segmentasi otomatis menggunakan metode yang diusulkan, sedangkan  $A_G$  merupakan citra segmentasi manual atau *groundtruth*. *Confusion matrix* digunakan untuk mencari nilai akurasi dari hasil segmentasi. Tabel *confusion matrix* dapat dilihat pada Tabel I.

TABEL I  
CONFUSION MATRIX

Confusion matrix		Predicted	
		Negative	Positive
Actual	Negative	TN	FP
	Positive	FN	TP

Dengan menggunakan *true positive* (TP), *true negative* (TN), *false negative* (FN), dan *false positive* (FP), nilai akurasi segmentasi bisa didapatkan. Nilai akurasi segmentasi mendefinisikan seberapa besar yang telah diklasifikasi dengan benar secara keseluruhan. Nilai akurasi dapat dihitung dengan (15).

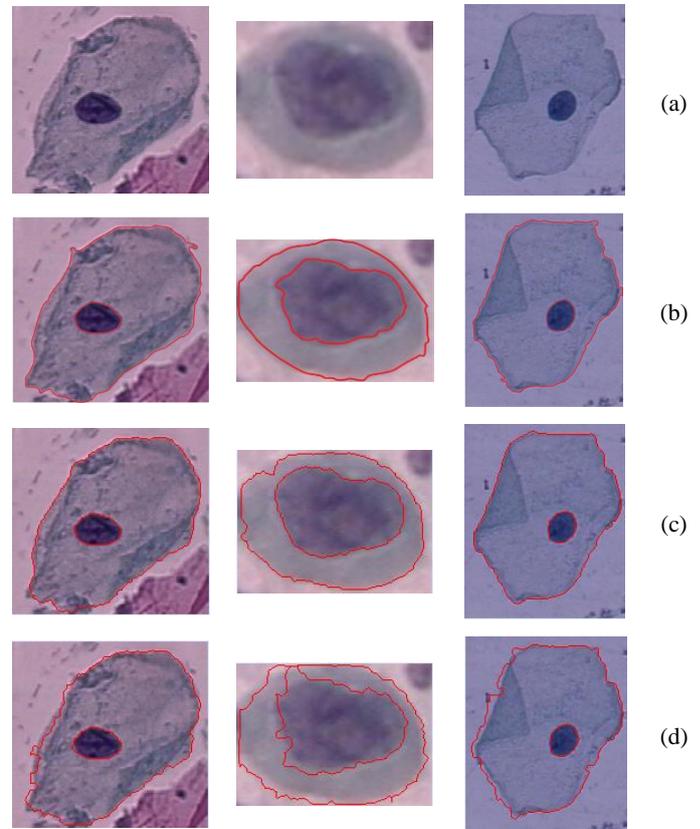
$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{15}$$

A. Hasil Uji Coba Skenario 1

Uji coba yang pertama adalah melakukan segmentasi dengan metode yang diusulkan, yaitu menggunakan metode fungsi energi eksternal RCNGGVFS yang memanfaatkan perhitungan REM untuk pencarian *edge map* pada metode CNGGVFS dan dengan segmentasi awal FCM\_NLS (RCNGGVFS & FCM\_NLS). Hasil segmentasi tersebut dibandingkan dengan hasil segmentasi metode CNGGVFS yang menggunakan perhitungan *edge map* konvensional (*Sobel*) dengan segmentasi awal FCM\_NLS (CNGGVFS & FCM\_NLS). Uji coba pertama ini dimaksudkan untuk mengetahui unjuk kerja metode yang menggunakan perhitungan REM pada pencarian *edge map*, dengan metode yang hanya menggunakan pencarian *edge map* konvensional (menggunakan *Sobel*).

Gbr. 4 menunjukkan hasil segmentasi pada skenario uji coba pertama. Garis merah pada Gbr. 4 menunjukkan area kontur nukleus dan sitoplasma citra sel tunggal *smear* serviks.

Dapat dilihat dari Gbr. 4(c), hasil segmentasi metode RCNGGVFS dengan segmentasi awal FCM\_NLS yang diusulkan lebih optimal dan mendekati *groundtruth* dibanding metode CNGGVFS yang menggunakan metode pencarian *edge map* konvensional (*Sobel*) dengan segmentasi awal FCM\_NLS.



Gbr. 4 Hasil segmentasi pada skenario uji coba pertama. (a) citra asli, (b) *groundtruth*, (c) Hasil segmentasi menggunakan RCNGGVFS dengan segmentasi awal FCM\_NLS, (d) Hasil segmentasi menggunakan CNGGVFS (menggunakan metode pencarian *edge map* konvensional) dengan segmentasi awal FCM\_NLS.

TABEL II  
NILAI RATA-RATA AKURASI & ZSI NUKLEUS PADA UJI COBA SKENARIO 1

Metode	Rata-rata akurasi nukleus (%)	Rata-rata ZSI nukleus (%)
RCNGGVFS & FCM_NLS (Metode usulan)	95,34	88,06
CNGGVFS & FCM_NLS ( <i>edge map</i> konvensional)	93,41	84,70

Untuk mengetahui unjuk kerja metode yang diusulkan, dilakukan perbandingan citra hasil segmentasi dengan citra *groundtruth*. Evaluasi dilakukan dengan dua cara, yaitu mencari nilai ZSI dan akurasi. Hasil evaluasi segmentasi nukleus dan sitoplasma citra sel tunggal *smear* serviks pada uji coba pertama ini dapat dilihat pada Tabel II dan Tabel III. Pada kedua tabel dapat dilihat hasil rata-rata akurasi dan ZSI dari segmentasi nukleus pada uji coba pertama ini, yaitu metode yang menggunakan perhitungan REM memiliki nilai

rata-rata akurasi dan ZSI yang lebih tinggi dibanding metode yang menggunakan pencarian *edge map* konvensional.

TABEL III  
NILAI RATA-RATA AKURASI & ZSI SITOPLASMA PADA UJI COBA SKENARIO 1

Metode	Rata-rata akurasi sitoplasma (%)	Rata-rata ZSI sitoplasma (%)
RCNNGVFS & FCM_NLS (Metode usulan)	83,48	87,16
CNGGVFS & FCM_NLS ( <i>edge map</i> konvensional)	82,55	86,68

B. Hasil Uji Coba Skenario 2

Skenario uji coba yang kedua adalah membandingkan metode yang diusulkan yaitu RCNNGVFS dengan segmentasi awal FCM\_NLS dengan metode pada penelitian sebelumnya yang melakukan segmentasi terhadap citra sel tunggal *smear* serviks. Metode tersebut adalah RGVFS dengan metode *K-means* sebagai segmentasi awal (RGVFS & *K-means* (tanpa *denoising*)) [5]. Metode yang kedua adalah metode RGVFS dengan metode FCM sebagai segmentasi awal (RGVFS & FCM (tanpa *denoising*)) [13]. RGVFS merupakan metode REM yang dimanfaatkan pada energi eksternal GVFS.

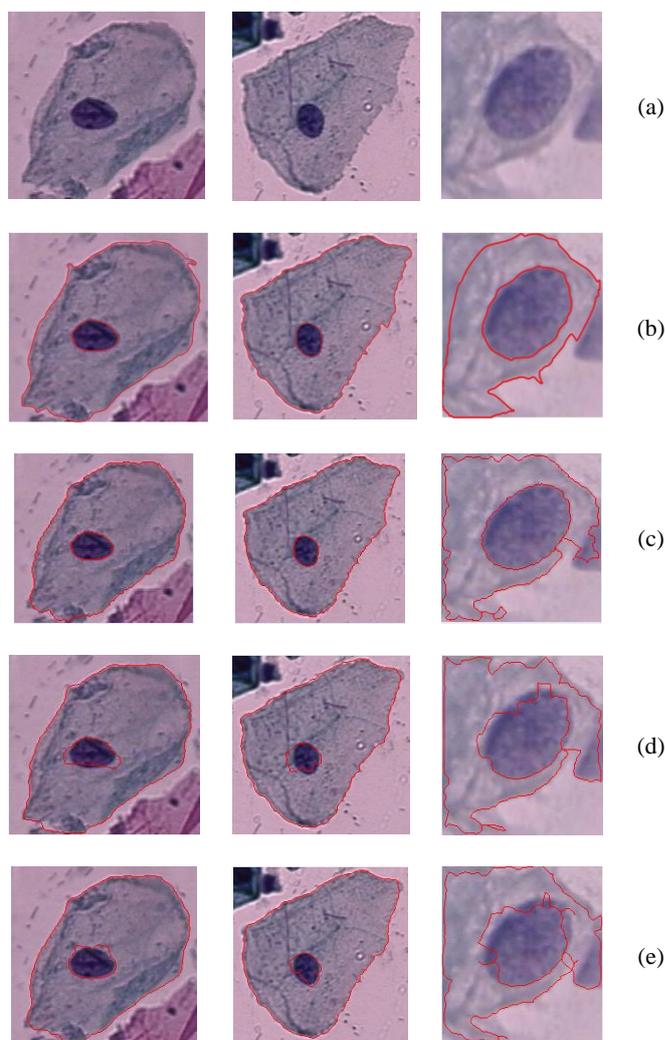
TABEL IV  
NILAI RATA-RATA AKURASI & ZSI NUKLEUS PADA UJI COBA SKENARIO 2

Metode	Rata-rata akurasi nukleus (%)	Rata-rata ZSI nukleus (%)
RCNNGVFS & FCM_NLS (Metode usulan)	95,34	88,06
RGVFS & FCM [13]	94,88	85,69
RGVFS & <i>K-means</i> [5]	94,56	86,66

TABEL V  
NILAI RATA-RATA AKURASI & ZSI SITOPLASMA PADA UJI COBA SKENARIO 2

Metode	Rata-rata akurasi sitoplasma (%)	Rata-rata ZSI sitoplasma (%)
RCNNGVFS & FCM_NLS (Metode usulan)	83,48	87,16
RGVFS & FCM [13]	81,83	86,00
RGVFS & <i>K-means</i> [5]	81,66	85,88

Pada Gbr. 5 dapat dilihat bahwa hasil segmentasi kontur yang dihasilkan menggunakan metode yang diusulkan lebih optimal dibanding metode lain. Sangat jelas terlihat bahwa hasil segmentasi kontur nukleus pada metode RGVFS & *K-means* dan RGVFS & FCM masih belum optimal. Ada area yang sebenarnya bukan area nukleus tapi dideteksi sebagai area nukleus. Tabel IV dan Tabel V menunjukkan hasil rata-rata akurasi dan ZSI dari segmentasi nukleus dan sitoplasma pada uji coba skenario kedua. Metode yang diusulkan masih memiliki rata-rata akurasi dan ZSI tertinggi.



Gbr. 5 Hasil segmentasi pada skenario uji coba kedua. (a) citra asli, (b) *groundtruth*, (c) metode RCNNGVFS dengan segmentasi awal FCM\_NLS, (d) RGVFS dengan segmentasi awal *K-means*, (e) RGVFS dengan segmentasi awal FCM.

C. Analisis Hasil Uji Coba

Pada makalah ini, data citra yang digunakan adalah citra mikroskopis sel tunggal *smear* serviks. *Data set* yang digunakan adalah *data set* Herlev yang terdiri atas 917 citra sel tunggal *smear* serviks yang diperoleh dari Departemen Patologi Rumah Sakit Universitas Herlev dan Departemen Automasi Universitas Teknik Denmark [9]. Pada *data set* citra sel tunggal *smear* serviks Herlev ini juga terdapat *groundtruth* atau hasil segmentasi manual.

Segmentasi awal yang digunakan pada makalah ini adalah FCM\_NLS. Segmentasi awal berfungsi mendapatkan kontur awal yang akan digunakan pada proses segmentasi kontur citra sel tunggal *smear* serviks untuk metode RCNNGVFS. Metode FCM\_NLS tegar terhadap citra yang memiliki derau densitas tinggi, sehingga pada tahap pra proses tidak dilakukan penghilangan derau. Hal ini dikarenakan pada metode FCM\_NLS telah terdapat proses filter derau dengan memanfaatkan informasi spasial non lokal atau biasa disebut

dengan *non local means*. Metode segmentasi awal yang baik akan membuat proses untuk mendapatkan kontur awal menjadi lebih mudah.

Makalah ini mengusulkan metode RCNGGVFS untuk segmentasi kontur citra sel tunggal *smear* serviks. Metode RCNGGVFS merupakan metode *active contour* yang menggunakan CNGGVFS sebagai fungsi energi eksternal dan memanfaatkan perhitungan REM untuk pencarian *edge map*. REM merupakan proses perhitungan *gradient edge map* pada sepanjang garis pancaran yang dimulai dari *intensity-weighted area centroid* nukleus hasil segmentasi awal. Untuk itu, dilakukan uji coba pertama untuk membuktikan bahwa metode REM menghasilkan segmentasi kontur nukleus yang lebih baik dibanding jika digunakan metode pencarian *edge map* konvensional. Pada Tabel II dan Tabel III dapat dilihat nilai rata-rata akurasi dan ZSI segmentasi nukleus menggunakan metode perhitungan REM adalah 95,34% & 88,06%. Sedangkan nilai rata-rata akurasi dan ZSI segmentasi nukleus jika menggunakan metode pencarian *edge map* konvensional adalah 93,41% dan 84,70%.

Selanjutnya, uji coba kedua dilakukan untuk membuktikan unjuk kerja metode RCNGGVFS dibanding dengan metode yang diusulkan oleh penelitian sebelumnya yang menggunakan RGVFS. Metode tersebut adalah RGVFS & *K-means* (tanpa *denoising*) [5] dan RGVFS & FCM (tanpa *denoising*) [13]. Pada Tabel IV dan Tabel V dapat dilihat bahwa metode yang diusulkan pada makalah ini menghasilkan nilai rata-rata akurasi dan ZSI dari semua kelas citra sel tunggal *smear* serviks lebih tinggi dibanding metode yang telah dibandingkan pada skenario uji coba, sedangkan metode yang diusulkan pada [5] dan [13] masih belum bisa mendapatkan hasil segmentasi yang baik terhadap area nukleus dari citra sel tunggal *smear* serviks. Pada penelitian tersebut, algoritme energi eksternal yang digunakan adalah RGVFS. Sebagaimana dijelaskan pada Bagian I, GVFS memiliki konvergensi yang lemah terhadap batas kecekungan objek, sehingga metode REM yang dimanfaatkan pada GVFS belum mampu mendeteksi area nukleus dengan benar.

Pada makalah ini, metode REM dimanfaatkan pada metode CNGGVFS yang memiliki rata-rata akurasi dan ZSI untuk hasil segmentasi kontur nukleus pada citra sel tunggal *smear* serviks lebih tinggi dibanding metode lain. Hal ini disebabkan oleh penggunaan REM pada CNGGVFS. Selain itu, CNGGVFS juga merupakan algoritme pengganti energi eksternal yang memiliki konvergensi yang lebih baik terhadap batas kecekungan objek, sehingga metode RCNGGVFS mampu mendapatkan hasil segmentasi kontur nukleus pada citra sel tunggal *smear* serviks lebih optimal.

#### V. KESIMPULAN DAN SARAN

Berdasarkan hasil uji coba dan analisis hasil pengujian segmentasi kontur sitoplasma dan nukleus pada citra sel tunggal *smear* serviks menggunakan metode RCNGGVFS, dapat disimpulkan bahwa metode yang diusulkan mendapatkan nilai rata-rata akurasi dan ZSI hasil segmentasi yang lebih optimal. Metode yang diusulkan lebih optimal dibanding dengan metode yang telah dibandingkan pada

skenario uji coba, khususnya hasil segmentasi kontur nukleus pada citra sel tunggal *smear* serviks. Rata-rata akurasi dan ZSI untuk segmentasi nukleus adalah 95,34% dan 88,06%. Kemudian, nilai rata-rata akurasi dan ZSI metode yang diusulkan untuk segmentasi sitoplasma adalah 83,48% dan 87,16%. Hasil ini lebih baik dibanding metode lainnya yaitu CNGGVFS yang menggunakan pencarian *edge map* konvensional, RGVFS & FCM (tanpa *denoising*), dan RGVFS & *K-means* (tanpa *denoising*). Hal ini disebabkan metode REM berhasil mendeteksi tepi nukleus lebih baik dibanding menggunakan *Sobel*, sehingga proses deformasi *snake* oleh metode CNGGVFS lebih mudah dilakukan dan menghasilkan segmentasi yang lebih optimal.

Pada penelitian ini dilakukan penggunaan *global threshold* pada proses ekstraksi kontur awal yang menyebabkan nukleus pada beberapa citra terdeteksi sebagai derau karena memiliki nilai *threshold* lebih kecil dari nilai *threshold* yang ditetapkan. Untuk penelitian selanjutnya, disarankan menggunakan *adaptive threshold* untuk proses ekstraksi kontur awal.

#### REFERENSI

- [1] Y.P.Pasrun, C.Fatihah & N.Suciati, "Penggabungan Fitur Bentuk dan Fitur Tekstur yang Invariant terhadap Rotasi untuk Klasifikasi Citra Pap Smear", *Jurnal Buana Informatika*, vol.7(1), hal.11-20, 2016.
- [2] K.A.Abuhasel, C.Fatihah, & A.M.Iliyasu, "A Bi-Stage Technique for Segmenting Cervical Smear Images Using Possibilistic Fuzzy C-Means and Mathematical Morphology", *Journal of Medical Imaging and Health Informatics*, vol.6 (7), hal.1663-1669, 2016.
- [3] H.S.Wu, J.Gil, J.Barbara, "Optimal segmentation of cell images, IEEE Proceedings of Vision", *Image and Signal Processing*, vol.145(1), hal.50-56, 1998.
- [4] P.Bamford & B.C.Lovell, "A water immersion algorithm for cytological image segmentation", *Proceedings of the APRS Image Segmentation Workshop*, Sydney, Australia, hal.75-79, 1996.
- [5] K.Li, Z.Lu, W.Liu & J.Yin, "Cytoplasm and nucleus segmentation in cervical smear images using Radiating GVF Snake", *Pattern Recognition*, vol.45(4), hal.1255-1264, 2012.
- [6] C.Xu, & J.L.Prince, "Generalized gradient vector flow external forces for active contours", *Signal Processing Elsevier*, vol.71, hal.131-139, 1998.
- [7] S.F.Yang-Mao, Y.K.Chan & Y.P.Chu, "Edge enhancement nucleus and cytoplasm contour detector of cervical smear images", *IEEE Transactions on Systems and Cybernetics*, vol.38(2), hal.353-366, 2008.
- [8] L.Qin, C.Zhu, Y.Zhao, H.Bai & H.Tian, "Generalized Gradient Vector Flow for Snakes: New Observations, Analysis, and Improvement", *IEEE Transactions on Circuits and Systems for Video Technology*, vol.23(5), hal.883-897, 2013.
- [9] J.Jantzen & G.Dounias, "Analysis of Papsmear image data", *Proceedings of the Nature-Inspired Smart Information Systems 2nd Annual Symposium NISIS*, 2006.
- [10] J.Jantzen & G.Dounias, *The Pap-Smear Benchmark* (2008) [Online] <http://mde-lab.aegean.gr/index.php/downloads>, tanggal akses: 12 Mei 2016.
- [11] F.Zhao, L.Jiao, & H.Liu, "Fuzzy c-means clustering with non local spatial information for noisy image segmentation", *Frontiers of Computer Science in China*, vol.5(1), hal.45-56, 2011.
- [12] A.Buades, B.Coll & J.M. Morel, "A non-local algorithm for image denoising", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol.2, hal.60-65, 2005
- [13] M.D.E.Susanti, H.Tjandrasa & C.Fatihah, "Segmentasi Nukleus dan Sitoplasma pada Citra Smear Serviks menggunakan Kombinasi Metode Fuzzy C- Means Clustering dan Radiating Gradient Vector Flow Snake", Tugas Akhir, Institut Teknologi Sepuluh Nopember, Surabaya, 2015.