

Pemodelan Segmentasi Sel Epitel Serviks pada Citra Digital Pap Smear

(Modeling of Cervical Cell Segmentation on Pap Smear Digital Image)

Komariyuli Anwariyah^{[1]*}, Dedy Sofyan^[2]

^[1]Fakultas Sains Dan Teknologi
Universitas Qamarul huda Badaruddin - UNIQHBA, NTB, Indonesia
E-mail: yuliaanwariyah@gmail.com

^[2]Universitas Nahdlatul Ulama NTB - UNU, NTB, Indonesia
E-mail: dewapyan20017@gmail.com

KEYWORDS:

Cervical cancer, Segmentation, region minima, region maxima, watershed

ABSTRACT

Cervical cancer is a disease that is very deadly. The level of this disease in Indonesia is very high, the rate of prevalence in 2013 of about 0.8% with an estimated 92 692 absolute throughout Indonesia. Early detection of cancer can be done with a Pap cytology examination. Slide readings performed by a specialist in anatomical pathology where normal conditions at least there should be 8000-12000 cells in good condition. This reading is not easy and has some constraints, it is necessary to designed system capable of automatically reading slide. System automation is expected to reduce errors due to manual readings. One of the stages in the building automation system is the segmentation of the cell to separate the cells from background objects. Separation is done by implementing the method of minima and maxima region, as well as the implementation of a subset of the region area. From this study, it was found that the model is implemented showed that the implementation of the model can be used to do the segmentation process either cell cytoplasm and nucleus.

KATA KUNCI:

Kanker serviks, Segmentasi, watershed

ABSTRAK

Kanker serviks adalah salah satu penyakit yang sangat mematikan. Jumlah penderita penyakit ini di Indonesia cukup tinggi, dimana pada tahun 2013 prevelensi untuk penyakit ini sangat tinggi yaitu sekitar 0,8% dengan estimasi absolut 92.692 penderita. Deteksi awal kanker serviks dapat dilakukan dengan pemeriksaan cytology pap smear. Pembacaan slide pap smear dilakukan oleh spesialis patologi anatomi. Dimana untuk kondisi normal setidaknya harus terdapat 8000 – 12000 sel normal. Proses pembacaan ini tidak mudah dan memiliki beberapa kendala, untuk itu diperlukan adanya sebuah system yang mampu melakukan pembacaan slide pap smear secara otomatis. Otomasi sistem diharapkan dapat mengurangi kesalahan karena pembacaan manual. Salah satu tahapan dalam membangun sistem otomasi adalah segmentasi untuk memisahkan sel dari objek latar belakang. Pemisahan dilakukan dengan menerapkan metode minima dan maxima region, serta implementasi subset dari region region. Dari penelitian ini, ditemukan bahwa model yang diterapkan menunjukkan bahwa implementasi model dapat digunakan untuk melakukan proses segmentasi baik sitoplasma sel maupun inti.

I. PENDAHULUAN

Kanker serviks merupakan salah satu penyakit yang mematikan. Jumlah penderita penyakit ini di Indonesia terbilang cukup tinggi dimana pada tahun 2013 prevelensi untuk penyakit ini sangat tinggi yaitu sekitar 0.8% dengan estimasi absolut 92.692 penderita. [1]

Deteksi awal kanker serviks dapat dilakukan dengan pemeriksaan cytology pap smear. Pemeriksaan ini dilakukan dengan pembacaan

manual sel pada cairan serviks yang dilakukan oleh dokter spesialis patologi anatomi. Pemeriksaan ini terkadang mengalami kegagalan atau hasil yang dikenal dengan negative palsu atau positif palsu. Hal ini bias disebabkan oleh kurangnya kemampuan laboratorium atau kelompok penduduk yang diperiksa [2].

Selain itu interpretasi hasil *pap smear* bukan merupakan hal yang mudah, terdapat beberapa hambatan dalam melakukan proses ini yaitu

metode konvensional memberikan hasil layer yang tidak merata, berkerumun dan sel yang tumpang tindih, selain itu prosedur pewarnaan, adanya lendier dan darah yang berlebih dan pengeringan udara mengakibatkan pembacaan menjadi lebih sulit . [3].

Proses identifikasi sel diawali dengan proses segmentasi sel yang nantinya akan memisahkan object sel dari background. Pada penelitian ini akan dilakukan pemodelan segmentasi sel yang ada pada citra digital pap smear yang nantinya akan diidentifikasi sebagai object yang akan dikenali.

II. METODOLOGI

A. Teknik Pengumpulan Data

Dalam tahapan ini dilakukan pengumpulan sampel dengan mengambil sampel primer dari laboratorium patologi anatomi. Ada 3 tahap yang dilakukan dalam proses pengambilan sampel yaitu:

a. Tahap Pemilihan Preparat.

Pada tahap ini, preparat yang telah berisi sel epitel serviks diseleksi, seleksi dilakukan berdasarkan tipe preparat. Preparat yang dipilih diutamakan yang menggunakan tipe preparasi *liquid base preparation* (LBP) dan beberapa preparasi dengan tipe konvensional. Tipe *liquid base* dipilih karena hasil preparasi jauh lebih tipis dan baik sehingga proses identifikasi sel jauh lebih Mudah.

b. Tahap Identifikasi Area

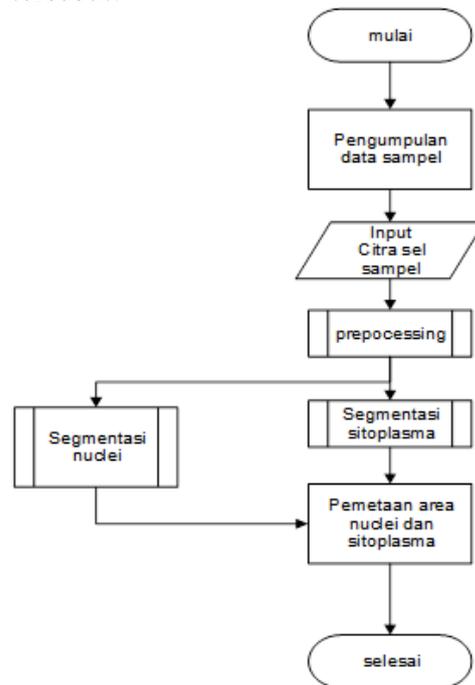
Proses identifikasi dilakukan untuk mengetahui area dan sel yang akan difoto sebagai sampel. Proses identifikasi dilakukan oleh peneliti dibawah mikroskop dengan tipe Nikon D.100. selanjutnya area yang akan difoto dilingkari.

c. Digitalisasi citra

Pada tahap ini area yang sudah ditandai akan difoto menggunakan mikroskop dengan tipe Nikon DN.100. Mikroskop ini merupakan mikroskop digital yang sudah dihubungkan network dan monitor untuk melihat area yang akan difoto. Hasil foto disimpan keserver dalam bentuk citra digital. Citra digital inilah yang nantinya diproses untuk disegmentasi dan dilakukan pengenalan.

B. Ekstraksi Citra Digital

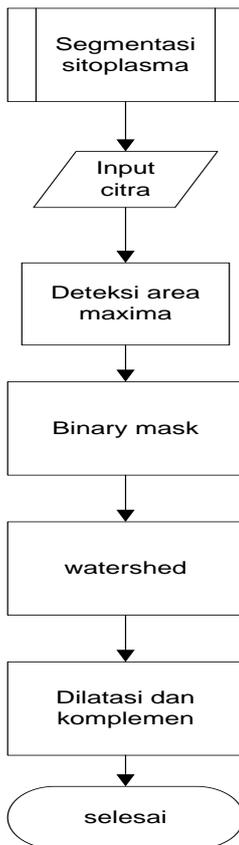
Dari Gambar 1 dapat dilihat bahwa terdapat tiga sub tahap utama yang dilakukan yaitu: *preprocessing*, segmentasi inti, dan segmentasi sitoplasma. Ketiga sub tahap ini dilanjutkan dengan evaluasi dan selanjutnya dilakukan pengenalan dari tiap sel yang ada pada *slide* citra tersebut.



Gbr 1. Flow Chart Penelitian



(a)



(b)

Gbr 2. (a) Detail proses segmentasi nuclei; (b) Detail proses sitoplasma

Dari diagram skema *flowchart* di atas terdapat beberapa metode yang diimplementasikan dalam tahapan *preprocessing*, dan segmentasi adalah:

a. Enhancement

Histogram citra sering digunakan dalam tahap perbaikan citra, ini didefinisikan sebagai sebuah vektor yang berisi perhitungan jumlah pixel pada setiap level keabuan. Untuk histogram, $h(i)$ didefinisikan sebagai [4].

$$h(i) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \delta(f(m,n) - i), i = 0, 1, \dots, P-1, (1)$$

Dimana :

$$\delta(w) = \begin{cases} 1, & w=0 \\ 0, & \text{otherwise} \end{cases}$$

Operasi perbaikan citra merupakan proses konvolusi menggunakan operator lokal yang dikenal dengan kernel. Proses konvolusi citra dengan kernel didefinisikan sebagai [4] :

$$g(m,n) = w(k,l) * f(m,n) = \sum_{k=-K}^K \sum_{l=-L}^L w(k,l).f(m-k,n-l); (2)$$

b. Rekonstruksi citra biner

Diberikan dua buah citra biner J dan I yang didefinisikan sebagai domain (D) diskrit yang sama dan dirumuskan sebagai :

$$\begin{aligned} J &\subseteq I, \text{ i.e., } \forall p \in D, \\ j(p) = 1 &\Rightarrow I(p) = 1 \end{aligned} (3)$$

Dimana I_1, I_2, \dots, I_n merupakan *connected component* dari I .

Sehingga dapat didefinisikan bahwa hasil rekonstruksi morfologi ($\square I(J)$) oleh citra mask (I) dari citra marker (J) adalah *union* dari *connected component* I yang mengandung pixel yang lebih kecil atau sama dengan J . rekonstruksi juga dapat dirumuskan sebagai *geodesic dilation* yang dirumuskan sebagai [5]:

$$\delta_X^{(n)}(Y) = \{p \in X \mid d_X(p, Y) \leq n\} (4)$$

Geodesic dilation dari ukuran n merupakan iterasi dari n elemen *dilatasi geodesic* yang dirumuskan sebagai [6]:

$$\delta_X^{(n)}(Y) = \underbrace{\delta_X^{(1)} \circ \delta_X^{(1)} \circ \dots \circ \delta_X^{(1)}}_{n \text{ times}}(Y). \quad (5)$$

c. Rekonstruksi Citra Grayscale

Grayscale reconstruction didefinisikan sebagai iterasi dilasi geodesic dari citra marker (J) di dalam citra mask (I) hingga mencapai intensitas yang stabil. Rekonstruksi pada citra grayscale dirumuskan sebagai [5].

$$\rho(I) = \lim_{n \rightarrow +\infty} \delta_I^{(n)}(J) = \lim_{n \rightarrow +\infty} \underbrace{\delta_I^{(1)} \circ \delta_I^{(1)} \circ \dots \circ \delta_I^{(1)}}_{n \text{ times}}(J) \quad (6)$$

d. Edge Detection

Tepi citra ialah posisi dimana intensitas pixel dari citra berubah dari nilai rendah ketinggian ataupun sebaliknya. Salah satu operator yang digunakan untuk deteksi tepi ialah Laplacian Of Gaussian (LOG). Fungsi gaussian inilah yang berfungsi untuk mengurangi derau yang ada pada citra. Fungsi LOG dirumuskan sebagai [6] :

$$LoG = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (7)$$

e. Region Filling

Pada dasarnya proses region filling berbasis pada himpunan dilasi, komplemen dan intersection. Untuk itu proses region filling dirumuskan sebagai [7]

$$X_k = (X_{k-1} \oplus B) \cap A^c \quad (8)$$

$k = 1, 2, 3, 4, \dots$

Pada Persamaan 8 diatas diasumsikan area bukan boundary dengan label sama dengan 0, selanjutnya diberikan nilai 1 hingga p untuk memulai. $X_0 = p$, dan B merupakan struktur elemen simetris. Algoritma diatas akan melakukan iterasi untuk step k jika $X_k = X_{k-1}$.

f. Metode Watershed

Metode watershed didefinisikan untuk sebuah kasus kontinu yang berdasarkan pada fungsi jarak. Transformasi watershed merupakan sebuah metode yang dipilih dalam segmentasi gambar dengan memanfaatkan konsep

morphology dalam matematika. Metode watershed dimodelkan sebagai [8]

$$Wshed(f) = D \cap \left(\bigcup_{i \in I} CB(m_i) \right)^c \quad (9)$$

Dimana

$$CB(m_i) = \{x \in D \mid \forall j \in I \setminus \{i\} : f(m_i) + T_f(x, m_i) < f(m_j) + T_f(x, m_j)\} \quad (10)$$

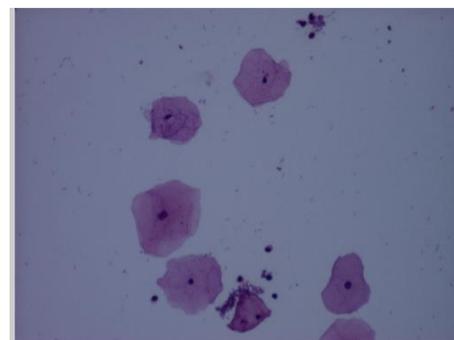
III. HASIL DAN PEMBAHASAN

A. Preprocessing

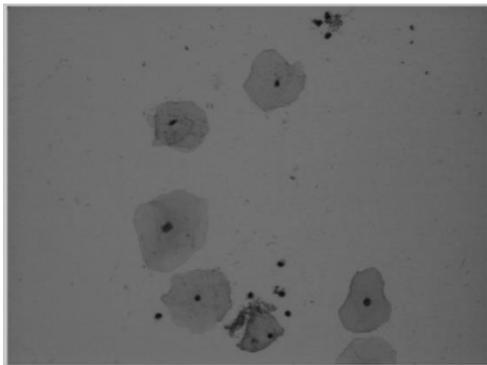
Tahap preprocessing dilakukan untuk memperbaiki kualitas gambar yang nantinya akan dilakukan proses segmentasi, ekstraksi dan selanjutnya menuju proses pengenalan sel tersebut. Dalam tahapan preprocessing ada beberapa proses yang dilakukan proses perbaikan dengan menggunakan konsep morphology citra.

a. Pemisahan layer gambar

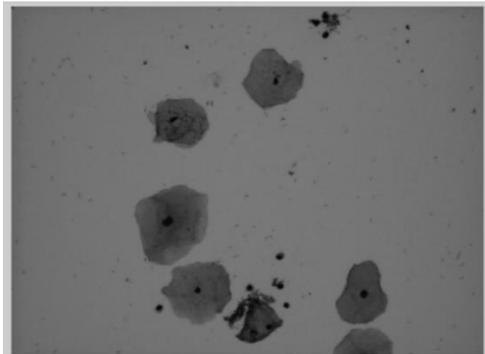
Citra input yang dibaca pada system merupakan gambar dengan format jpeg, proses pemisahan layer warna dilakukan untuk memudahkan proses pengolahan gambar. Proses pengolahan per layer menghasilkan yang lebih baik dari pada pengolahan dilakukan dengan melakukan konversi gambar ke grayscale secara langsung. Hasil pemisahan citra kedalam 3 layer Red, Green, dan Blue dapat dilihat pada Gambar 3 berikut :



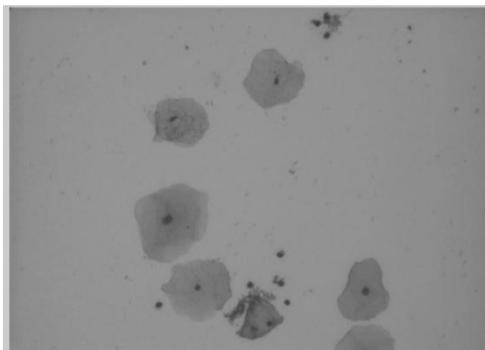
(a)



(b)



(c)



(d)

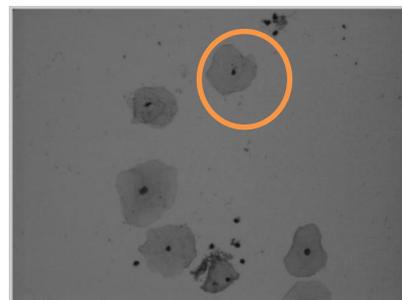
Gbr 3. Citra Input; (a) citra RGB input; (b) Citra Layer Merah; (c) Citra layer hijau; (d) citra layer biru

b. Erosi

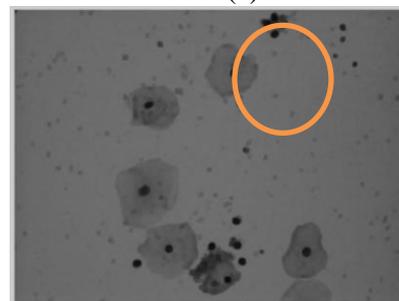
Erosi merupakan sebuah proses morfologi yang dilakukan dengan membandingkan setiap pxel dari nilai object dengan nilai pusat dari struktur elemen. Jika nilai setiap pixel sama dengan nilai pixel object maka pixel input akan diset nilainya dengan nilai pixel object. Bila tidak maka input pixel akan diberi nilai sesuai dengan background dari gambar tersebut [4]. Hal tersebut

mengakibatkan luasan object setelah dilakukan erosi akan menjadi lebih kecil jika dibandingkan dengan luasan saat sebelum dilakukan proses erosi.

Untuk citra *grayscale* proses erosi akan menurunkan intensitas warna dari citra input, sehingga gambar akan cenderung menjadi lebih gelap. Selain itu, proses erosi akan memperluas area gelap yang terdapat pada sekitar object sehingga dapat memperjelas pemisahan dua object yang berbeda. Pada Gambar 4 dapat dilihat perbedaan gambar setelah dilakukan proses erosi dengan sebelum dilakukan proses erosi.



(a)



(b)

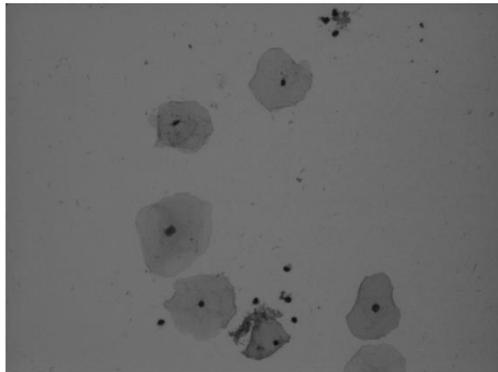
Gbr 4. (a) gambar sebelum erosi; (b) hasil setelah dilakukan erosi

Pada Gambar 4 (b) dapat dilihat pada contoh sel yang dilingkari, kandidat nuclei menjadi lebih jelas dan intensitas warnanya menjadi lebih baik. Hal ini memudahkan dalam proses indentifikasi selanjutnya.

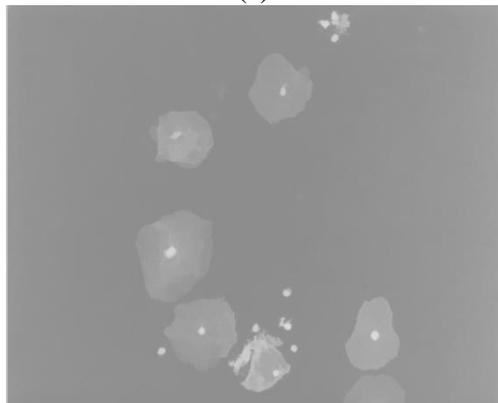
c. Rekonstruksi Citra

Proses rekonstruksi morfologi dilakukan untuk membantu proses segmentasi nantinya. Proses morfologi ini akan memperbaiki citra dengan tetap mempertahankan bentuk citra aslinya. Metode ini mengakibatkan hasil

segmentasi nantinya menjadi lebih tepat dan tidak berubah.



(a)

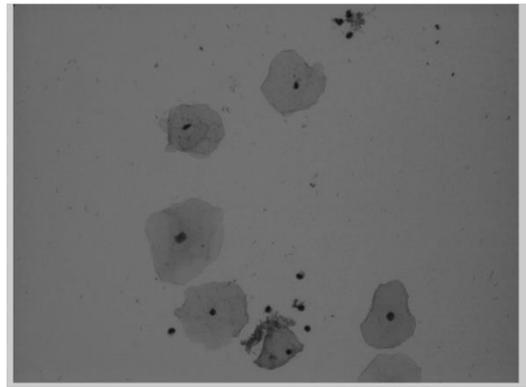


(b)

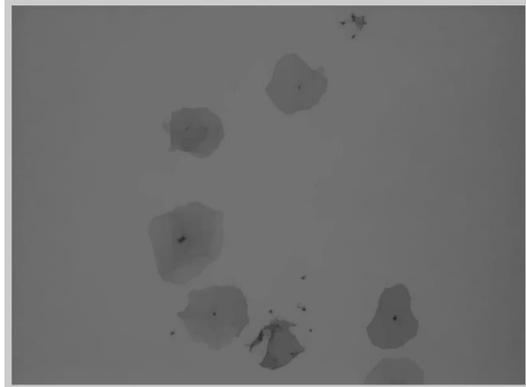
Gbr 5. (a) Citra sebelum dilakukan rekonstruksi; (b) citra setelah rekonstruksi

d. Dilasi

Berbeda dengan erosi yang akan memperkecil area object citra input, dilasi justru akan memperluas object dari citra input karena pada saat dilakukan hasil perbandingan nilai pixel object dengan struktur elemen akan mengubah nilai pixel yang tidak sama menjadi object. Pengolahan pada citra *grayscale* akan meningkatkan intensitas warna dari citra atau gambar input, hal ini dapat dilihat pada proses pengolahan berikut :



(a)

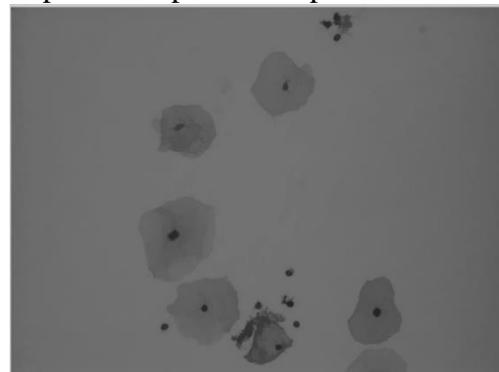


(b)

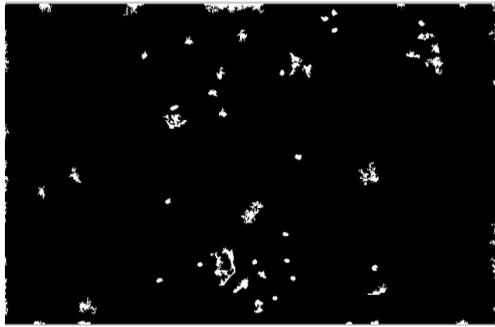
Gbr 6. (a) Gambar sebelum dilasi; (b) Gambar setelah dilasi

e. Seleksi Region

Seleksi region dilakukan untuk mengambil area. nucleus dan sitoplasma. Seleksi area dilakukan dengan metode seleksi region minima untuk nucleus dan seleksi region maxima untuk deteksi sitoplasma. Intensitas warna sitoplasma selalu lebih tinggi dibandingkan dengan nucleus dari sel tersebut. Hasil proses seleksi untuk nucleus dapat pada Gambar 7 dan hasil seleksi sitoplasma dapat dilihat pada Gambar 8



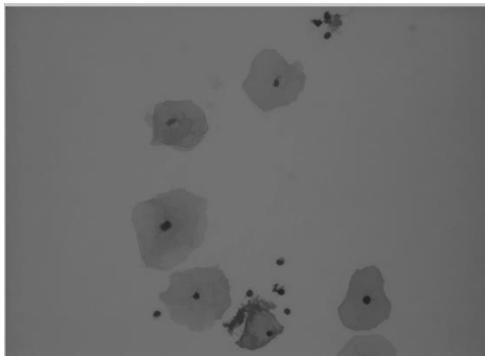
(a)



(b)

Gbr 7. (a) Gambar sebelum seleksi regionminima; (b) Gambar setelah seleksi region minima

Pada Gambar 7 (b) merupakan hasil seleksi region minima dimana pada proses ini diambil area dengan nilai intensitas minimum. Dimana area dengan intensitas minimum diberikan nilai 1 dan yang lainnya akan bernilai nol. Demikian pula dengan hasil seleksi region maksimum (Gambar 7 (b)), intensitas maksimum akan diberikan nilai 1 dan yang lainnya diberikan nilai 0. Sehingga hal ini memudahkan pada proses identifikasi dan segmentasi selanjutnya.



(a)



(b)

Gbr 8. (a) Gambar sebelum seleksi regionmaxima; (b) Gambar setelah seleksi region maxima.

B. Segmentasi

Proses segmentasi dilakukan untuk mendapatkan memisahkan area object dan background dari citra tersebut. Hasil segmentasi dibagi menjadi dua yaitu segmentasi nucleus dan segmentasi sitoplasma. Segmentasi awal dilakukan dengan menggabungkan proses seleksi minimum dan maksimum yang telah dilakukan pada setiap layer dengan persamaan berikut :

$$BwNuc = BwNuc1 \& BwNuc2 \& BwNuc3 \quad (11)$$

Dimana :

$BwNuc$ = citra biner hasil penggabungan

$BwNuc1$ = citra biner hasil seleksi region minimum layer R

$BwNuc2$ = citra biner hasil seleksi region minimum layer G

$BwNuc3$ = citra biner hasil seleksi region minimum layer B

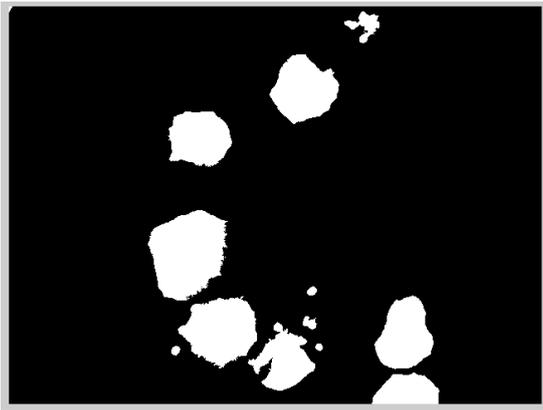
Sementara itu untuk segmentasi sitoplasma penggabungan menggunakan logika or seperti persamaan 3.2 berikut :

$$BwCty = (\sim BwCty1) | (\sim BwCty2) | (\sim BwCty3) \quad (12)$$

Dari Persamaan 11 dan Persamaan 12 didapatkan hasil citra segmentasi untuk nucleus dan sitoplasma seperti pada Gambar 9 sebagai berikut:



(a)



(b)

Gbr 9. (a) Hasil Segmentasi Nucleus; (b) Hasil segmentasi sitoplasma

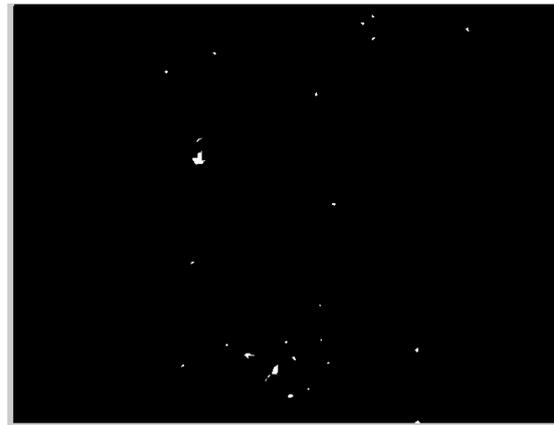
Setelah didapatkan hasil segmentasi awal seperti pada Gambar 9 selanjutnya dilakukan beberapa tahapan untuk membersihkan object pengganggu (noise) untuk mendapatkan hasil segmentasi yang lebih baik. Tahapan tersebut adalah sebagai berikut:

a. Seleksi intensitas

Setelah dilakukan proses segmentasi selanjutnya dilakukan proses seleksi nuclei berdasarkan intensitas warna, dalam hal ini nilai intensitas warna yang digunakan ialah lebih kecil sama dengan 80. Setiap kandidat nuclei dilakukan seleksi dengan mengecek nilai graylevel pada setiap kandidat, jika syarat terpenuhi maka kandidat diberi nilai 1 jika tidak diberi nilai 0 sehingga dengan model seleksi ini didapatkan sisa kandidat nuclei seperti pada Gambar 10



(a)



(b)

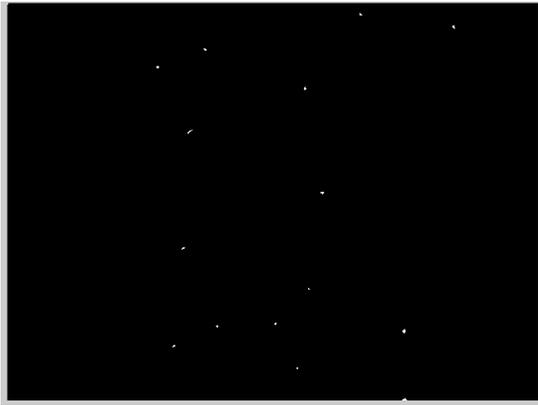
Gbr 10. (a) Gambar sebelum seleksi regionminima; (b) Gambar setelah seleksi region minima

b. Seleksi jarak

Seleksi jarak dilakukan untuk melakukan seleksi terhadap kandidat nuclei secara lebih optimal proses seleksi ini dilakukan dengan memberikan nilai jarak tertentu. Pada uji coba penelitian ini nilai jarak yang digunakan (R) ialah sebesar 100 pixel. Proses seleksi dilakukan dengan membandingkan jarak antar kandidat nuclei. Jika jarak kandidat lebih kecil dari jarak yang telah ditentukan maka kandidat akan dihapus atau diberi nilai nol. Seleksi terus dilakukan hingga semua kandidat *nuclei* terseleksi.

Seleksi dengan jarak ini mendapatkan hasil kandidat yang lebih optimal seperti pada Gambar 11 berikut:

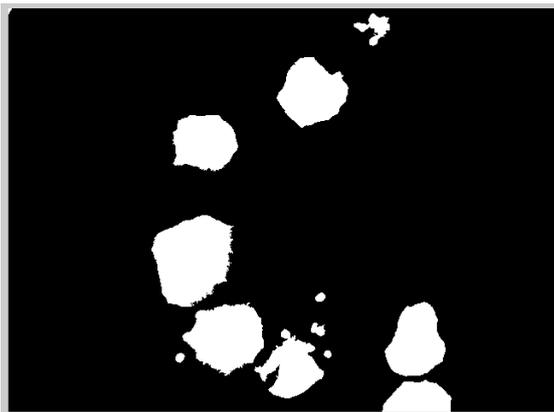




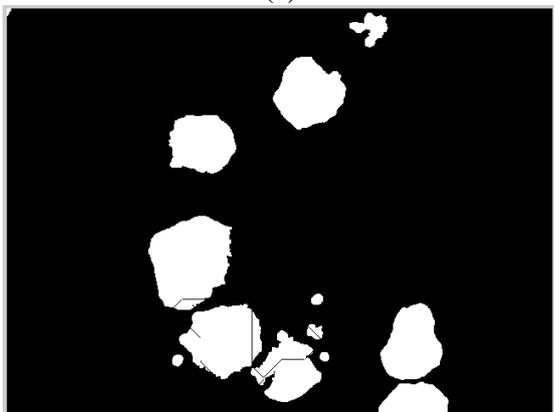
Gbr 11. (a) Kandidat nuclei sebelum dilakukan seleksi jarak (b) kandidat nuclei setelah dilakukan seleksi jarak

c. Segmentasi dengan watershed

Segmentasi dengan metode watershed dilakukan untuk membagi sel-sel yang overlapping atau saling bertindihan dengan asumsi setiap sel memiliki sebuah inti dan sitoplasma. Dengan metode ini didapatkan hasil segmentasi seperti terlihat pada gambar 3.10.



(a)



(b)

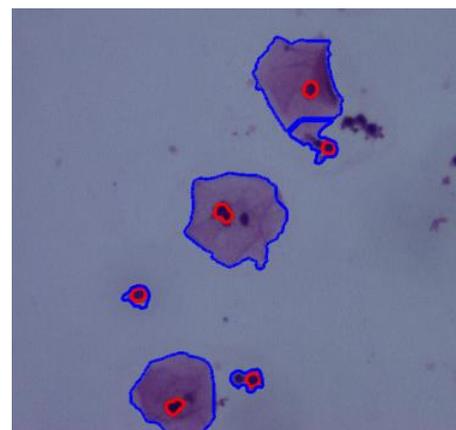
Gbr 12. (a) segmentasi sel sebelum implementasi metode watershed; (b) segmentasi sel setelah implementasi metode watershed

d. Pemetaan Nuclei Sitoplasma

Dari hasil seleksi ini dilakukan seleksi area sel untuk menghilangkan area bukan sel yang mungkin terdeteksi. Seleksi ini dilakukan dengan mengambil area yang hanya memiliki area inti karena setiap sel pasti memiliki inti. Seperti terlihat pada Persamaan 13 seleksi dilakukan dengan melihat anggota himpunan yang ada pada setiap inti dan sel. Dimana $f_n(x, y)$ dikatakan inti dari $f_s(x, y)$ jika semua anggota himpunan $f_n(x, y)$ merupakan anggota dari $f_s(x, y)$.

$$f_n(x, y) \rightarrow f_c(x, y) = \{f_n(x, y) \mid f_n(x, y) \subseteq f_c(x, y)\} \quad (13)$$

Dengan Persamaan 13 akan diperoleh pemetaan area inti ke satu area sitoplasma tertentu. Dan satu area sitoplasma hanya memiliki satu kandidat inti.



Gbr 13. Hasil Pemetaan Nuclei Sitoplasma

Pada Gambar 13 diatas dapat dilihat hasil pemetaan nucleus dan sitoplasma sel dimana lingkaran merah adalah nuclei sementara itu lingkaran biru adalah sitoplasma.

IV PENUTUP

A. Kesimpulan

Dari penelitian yang dilakukan dapat disimpulkan bahwa pemanfaatan model rekonstruksi citra dalam tahapan preprocessing membantu hasil segmentasi yang lebih baik,

dimana proses pengenalan kandidat nuclei dan sitoplasma dilakukan dengan menggunakan metode *region maxima* dan *minima* dan pemisahan sel akhir menggunakan metode *watershed*.

B. Saran

Saran dalam penelitian selanjutnya perlu dilakukan perbaikan model untuk mendapatkan hasil segmentasi yang lebih optimal. Model ini dirasa kurang efektif karena penentuan nilai *region maxima* menggunakan perkiraan nilai yang telah diuji coba sebelumnya.

REFERENSI

- [1] Anonym, 2015, Kementerian Kesehatan RI Pusat Data dan Informasi, Jakarta Selatan diakses 23/04/2015 pada www.depkes.go.id/download/infodatin.pdf website
- [2] Yatim,F, 2008. “ *Penyakit Kandungan*” , Pustaka Populer Obor, Jakarta. Hierarchical Neural Network” IJSSST, Vol. 9, No. 2, May 2008. Pp. 56-64.
- [3] Plissiti M E, Nikou C, Charcanti A, 2011, “*Automated Detection of Cell Nuclei in Pap Smear Images Using Morphological Reconstruction and Clustering*” IEEE Transactions On Information Technology In Biomedicine, Vol. 15, No. 2, March 2011.
- [4] Putra, 2010, *Pengolahan Citra Digital*, Andi Offset, Yogyakarta
- [5] Vincent L, 1993, *Morphological Grayscale Reconstruction in image Analysis: Applications and Efficient Algorithm*, Proc. IEEE Transaction on Image Processing Vol.2 No. 2, April 1993.
- [6] Muhimmah I, Kurniawan R, Indrayanti, 2012. *Automated Cervical Cell Nuclei Segmentation Using Morphological Operation and Watershed Transformation*.Cyberneticscom.
- [7] Gonzales, R.C. and Woods, R.E, 2002. *Digital Image Processing, 2nd ed.*,Prentice Hall. Upper Saddle River, NJ.
- [8] Roerdink Jos B.T.M and Meijster Arnold, 2001, *The Watershed Transform : Definition, Algorithm and Parallization Strategies*, Fundamentals Informaticae 41 Pp 187-228, IOS Press.
- [9] Lassouaoui N, Hammal L, Nouali N, 2005, *Morphological Description of Cervical Cell Images for the Pathological Recognition*, World Academy of Science, Engineering and Technology