

## SEGMENTASI CYTOPLASMA DENGAN METODE GRADIENT DAN DETEKSI TEPI PADA CITRA PAP SMEAR

Komariyuli Anwariyah, Izzati Muhimmah\*

Magister Teknik Informatika, Fakultas Teknik Industri, Universitas Islam Indonesia.

Jl. Kaliurang Km. 14 Yogyakarta 55501

Telp. (0274) 895287 ext. 122, Faks. (0274) 895007ext. 148

Jl. Yogyakarta Km 40 Jogja 55558

\*E-mail: izzati@uii.ac.id

### ABSTRAK

Salah satu tujuan penelitian pembacaan citra digital pap smears adalah untuk mendapatkan satu model yang paling sesuai guna dapat mengidentifikasi sel yang ada pada citra tersebut, baik itu sel normal maupun sel abnormal. Hal pertama yang harus dilakukan adalah mencari fitur atau ciri untuk pengelompokan berdasarkan tipe sel dari epithelial serviks. Salah satu cara identifikasi tipe sel serviks dapat dilakukan dengan meninjau karakteristik luas sel. Untuk itu perlu dilakukan segmentasi tidak hanya pada nucleus tetapi juga cytoplasm yang ada pada sel tersebut. Pada paper ini, metode gradient dan deteksi tepi citra diusulkan sebagai metode segmentasi sitoplasma. Hasil segmentasi sitoplasma kemudian digunakan sebagai dasar perhitungan luasan sel. Pada tahap ini, evaluasi metode yang diusulkan dilakukan dengan melalui penilaian secara visual dan kalkulasi area melalui persentase cytoplasm dibanding nilai rujukan. Dari penelitian ini dapat ditunjukkan bahwa identifikasi sel normal atau abnormal dapat dilakukan secara kuantitatif.

Kata kunci: pap smear, cytoplasm, gradient, deteksi tepi,

### 1. PENDAHULUAN

#### 1.1 Latar Belakang

Frekuensi kanker serviks menempati posisi paling tinggi dari pada kanker yang lain. Kanker serviks 99,7% disebabkan oleh *human papilloma virus* (HPV) *onkogenik* yang menyerang leher rahim. Hal ini juga diungkapkan Dr. Elfahmia Noor Azis bahwa penyebab kanker ini ialah *Human papilloma virus* (HPV) (Astana, 2009).

Kanker serviks stadium awal dapat didiagnosa dengan melakukan tes cytology Pap smear. Pap smear dilakukan untuk mengetahui jumlah sel yang ada, pada kondisi normal untuk metode konvensional setidaknya harus ada sekitar 8000 – 12000 sel dalam kondisi baik yang terhitung. Sementara itu untuk metode liquid base preparation (LBP) setidaknya harus ada 5000 sel normal yang terhitung (Hoda, 2007).

Sel serviks memiliki beberapa tipe diantaranya *superficial cell*, *Intermediate cell* dan *Parabasal cell*. Tiap sel ini memiliki persentase cytoplasm yang berbeda. Perbedaan karakteristik ini dapat digunakan untuk mengelompokkan tiap tipe sel tersebut dan dapat digunakan untuk mengetahui kondisi sel baik normal ataupun abnormal. Untuk itu perlu dilakukan deteksi sel tidak hanya terbatas pada nucleusnya tetapi juga terhadap cytoplasm sel tersebut.

Pada penelitian (Muhimmah, 2012), telah dilakukan deteksi cells dengan metode Fuzzy C-Means Clustering. Deteksi sel yang dilakukan berorientasi pada deteksi nuclei sel. Penelitian lain juga dilakukan (Tsai, M. Et al, 2008), yang terkait dengan Cytoplasm nuclei countour. Penelitian ini

menggunakan beberapa model untuk mendeteksi cytoplasm dan nuclei untuk masing - masing sel

Pada penelitian ini dilakukan deteksi dan segmentasi cytoplasm dengan menggunakan gradient dan deteksi tepi citra. Sehingga dapat dilakukan perhitungan persentase cytoplasm untuk tiap sel yang tersegment. Dari penelitian ini diharapkan didapatkan model segmentasi cytoplasm dengan metode gradient dan deteksi tepi.

Paper ini terdiri atas beberapa bagian. Bagian material dan metode pada bagian 2, hasil dan pembahasan bagian 3 dan terakhir kesimpulan pada bagian 4.

### 2. MATERIAL DAN METODE

#### 2.1 Material

Data yang digunakan pada penelitian ini berasal dari data database NCI Bethesda system dengan mengambil dua tipe perbesaran yaitu medium dan high. Penelitian ini menggunakan 7 slide data, dimana setiap slide terdiri atas minimal 6 sel. Sampel yang diambil menggunakan berbagai macam tipe preparasi baik itu konvensional, Liquid base preparation (LBP) ataupun thin prep.

#### 2.2 Deteksi Nuclei

Deteksi nuclei dilakukan dengan mengambil satu batas rentan intensitas tertentu. Sebelumnya citra input yang memiliki layer warna RGB (3 layer) ditransformasikan kedalam bentuk citra grayscale dengan persamaan (Solomon Chris, Breckon Toby, 2011) :

$$I_{grayscale} = \alpha I_{colour}(m,n,r) + \beta I_{colour}(m,n,g) + \alpha I_{colour}(m,n,b) \quad (1)$$

Selanjutnya kandidat nuclei dideteksi dengan rentan intensitas [0–150], menurut (Rosidi, et, all, 2011) dikatakan bahwa “level Intensitas warna dari sel kanker berada antara 80 – 100 sementara itu sel normal,berada antara 122 – 150.” Sehingga dapat dirumuskan sebagai berikut :

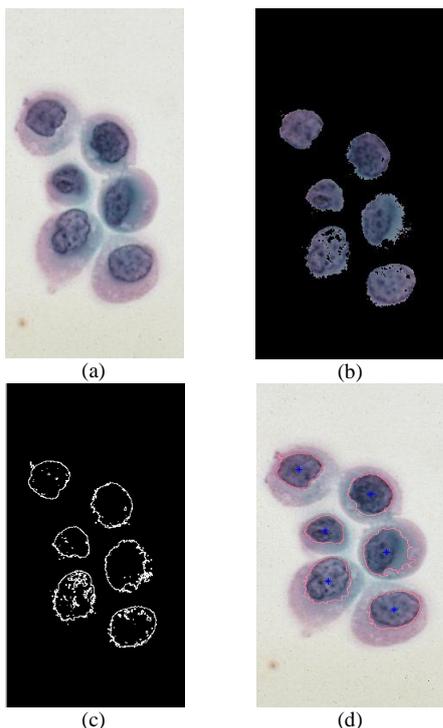
$$h(x, y) = \begin{cases} I(x, y), & bw(x, y) = 0 \text{ \& } I(x, y, c) \leq 150 \\ 0, & \text{else} \end{cases} \quad (2)$$

Dimana  $h(x,y)$  adalah citra kandidat nuclei,  $I(x,y)$  intensitas warna pada titik  $(x,y)$ ,  $bw(x,y)$  citra biner,  $I(x,y,c)$  intensitas pada setiap layer. Sehingga didapat kandedat nuclei untuk setiap sel (gambar 1).

Selanjutnya dilakukan deteksi tepi dengan filtering Laplacian of Gaussian (Log) . filtering Log dirumuskan sebagai :

$$LoG = -\frac{1}{\pi\sigma^4} \left[ 1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (3)$$

Untuk menghubungkan garis yang terputus dilakukan proses dilatasi dengan struktur elemen garis dan sudut mulai dari 0, 45, 90, 135 (supriyanto, 2011). Hasilnya berupa deteksi tepi citra (gambar 1).



Gambar 1. (a) Citra input, (b) hasil ekstraksi persamaan (2); (c) hasil deteksi tepi filter LoG; (d) hasil deteksi nuclei

### 2.3 Deteksi Cytoplasma

Deteksi Cytoplasma dilakukan dengan menggunakan metode gradient. Gradient image dari  $f$  pada  $(x,y)$  dirumuskan sebagai vector dua dimensi (Gonzales, 2002, ) :

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (4)$$

Gradient magnitude dirumuskan sebagai :

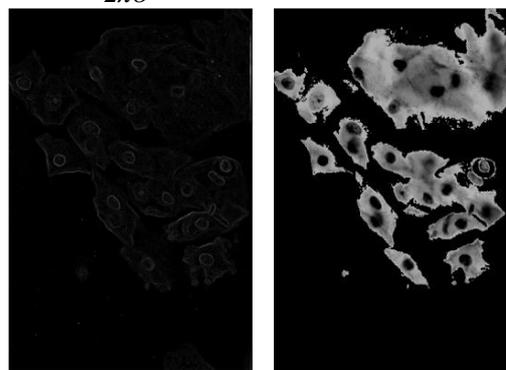
$$\nabla f = \left( G_x^2 + G_y^2 \right)^{\frac{1}{2}} \quad (5)$$

Konvolusi dengan filter Gaussian

$$f(x, y) = I(x, y) * G(x, y) \quad (6)$$

Dimana  $I(x,y)$  merupakan gambar input dan  $G(x,y)$  ialah filter Gaussian yang dirumuskan sebagai :

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (7)$$



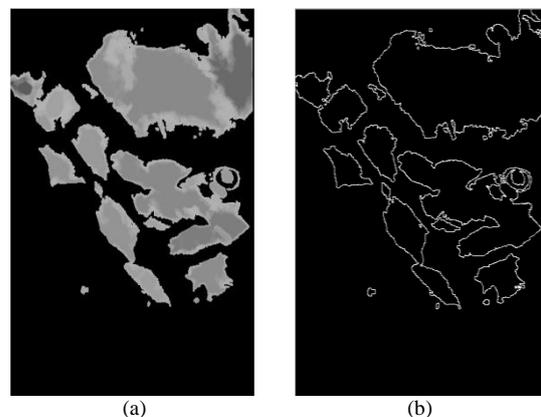
Gambar 2. (a) gradient magnitude; (b) hasil deteksi cytoplasma sel

Untuk menghilangkan area nuclei dilakukan proses region filling yang dirumuskan sebagai (Gonzales, 2002 ):

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

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

Deteksi tepi dilakukan dengan menggunakan filtering Laplacian Of Gaussian (Log) (persamaan 2) yang dilanjutkan proses dilatasi dengan struktur elemen garis. Seperti terlihat pada gambar 3



Gambar 2. (a) Hasil Region filling pers.(8); (b) hasil deteksi tepi cytoplasma

## 2.4 Spesifikasi Sel Epithel Pap Smear

Setiap sel terdiri atas cytoplasma dan nuclei. Sel epitel serviks pada citra pap smear memiliki komposisi luas sel dan nuclei yang berbeda untuk setiap tipe selnya. Ini dapat digunakan untuk melakukan pengelompokkan sel, selain itu criteria ini juga dapat digunakan untuk mengetahui normal tidaknya sel epitel serviks yang ada pada citra pap smear tersebut. Komposisi luas untuk setiap tipe sel dapat dilihat pada table 1 dibawah ini (Hoda,Hoda, 2007) :

Tabel 1 Comparative sizes dari sel Pap Smear

Tipe sel	Luas sel ( $\mu\text{m}^2$ )	Luas nuclear ( $\mu\text{m}^2$ )
Superficial	1500	20
Intermediate	1500	35
Parabasal	300	50
Endocervical	200	50
Endometrial	175	30
Reserve cells	200	50

Dari data pada tabel diatas didapatkan persentase cytoplasma untuk setiap tipe sel sebagai berikut (table 2) :

Tabel 2. Persentase Cytoplasma sel

Tipe sel	Cytoplasma (%)
Superficial	98.66
Intermediate	97.66
Parabasal	83.33
Endocervical	75
Endometrial	82.85
Reserve cells	75

Proses spesifikasi nuclei diperlukan untuk mengetahui area nuclei untuk setiap cytoplasma. Dapat dikatakan bahwa A merupakan nuclei B jika A merupakan himpunan bagian (*subset*) B.

## 2.5 Metode Evaluasi

Evaluasi hasil pada penelitian ini belum melibatkan pakar dalam hal ini dokter spesialis patologi anatomi. Evaluasi hanya didasarkan atas perbandingan persentase cytoplasma sel dalam kondisi normal (tabel 2) dengan persentase yang dihasilkan melalui perhitungan system (tabel 3). Penilaian juga didasarkan pada interpretasi data sampel yang diambil dari database NCI Bethesda system.

## 3. HASIL DAN PEMBAHASAN

Dari metode diatas diperoleh segmentasi untuk setiap sel baik segmentasi nucleus ataupun cytoplasmanya. Pada tahap deteksi nuclei diperoleh beberapa kandidat nuclei (gambar 1) yang nantinya akan diseleksi kembali sesuai dengan area cytoplasmanya.

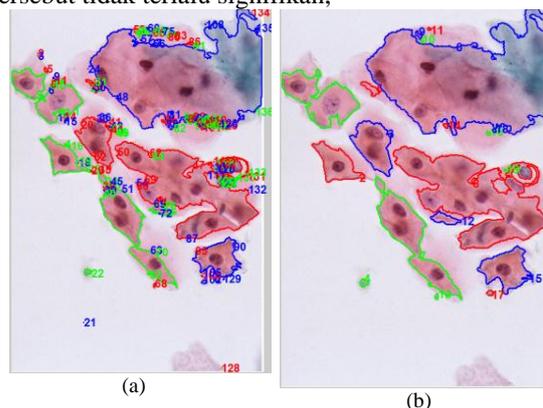
Pada tahap deteksi cytoplasma, pemrosesan nilai gradient menghasilkan suatu gradient magnitude yang menunjukkan tingkat perbedaan intensitas pada tiap pixel. Nilai ini diakumulasikan dan diambil ambang batas yang paling sesuai untuk kemudian dijadikan acuan deteksi cytoplasma. (gambar 2).

Pada gambar tersebut dapat dilihat bahwa sel sudah dapat disegmentasikan meskipun masih terdapat area hitam yang merupakan nuclei. hal ini tentunya mengganggu proses segmentasi cytoplasma. Untuk itu dilakukan proses region filling, dimana area hitam tersebut akan diisikan dengan nilai yang sama dengan area dominan di sekitarnya, yang kemudian dilanjutkan dengan proses deteksi tepi. Proses deteksi tepi dilakukan untuk mengetahui batasan area untuk setiap cytoplasma. (gambar 3) dan untuk mendapatkan hasil segmentasi yang lebih spesifik.

Pada tahap awal (gambar 4(a)) diperoleh sekitar 144 area cytoplasma yang kemudian pada tahap akhir menyusut hingga tersisa hanya 19 area cytoplasma.(gambar 4(b)).

Pada tahap awal (pers.5 dan pers. 6) proses segmentasi masih belum sempurna karena masih terdapat area – area kecil yang terdeteksi sebagai area cytoplasma. Untuk itu dilakukan proses erosi untuk menghilangkan area – area tersebut. Kemudian dilanjutkan dengan region filling (pers 7). Region filling bertujuan untuk mengisi area yang ada didalam boundary sehingga menjadi satu area yang sama. Tahapan ini dilanjutkan dengan deteksi tepi menggunakan filtering LoG(pers.3) sehingga segmentasi yang dihasilkan dapat menjadi lebih baik dan area yang pada proses sebelumnya (gambar 4(a)) terdeteksi sebagai cytoplasma dapat dihilangkan.

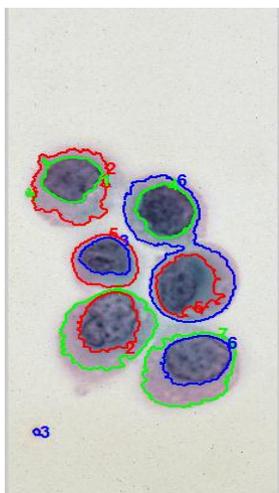
Hasil segmentasi ini sudah cukup baik meskipun dapat dikatakan belum sempurna karena seperti terlihat pada gambar 4(b) masih terdapat area cytoplasma yang tidak terdeteksi secara utuh dan untuk sel yang berkerumun, belum dapat dipisahkan secara sempurna hal ini disebabkan karena perbedaan nilai gradient magnitude pada area tersebut tidak terlalu signifikan,



Gambar 4. (a) Segmentasi cytoplasma tahap awal, pers.(5), (6); (b) segmentasi cytoplasma tahap akhir pers.(8)

Dari hasil segmentasi ini dilakukan proses spesifikasi nuclei untuk masing – masing cytoplasma (gambar 5). Hal ini dilakukan untuk mengetahui area nuclei untuk setiap segmen cytoplasma. Pada tahap ini dilakukan pengecekan anggota himpunan pada setiap area nuclei. seperti

pada gambar 5, nuclei dengan label 1 dikatakan nuclei dari segment cytoplasma dengan label 2 hanya jika seluruh anggota himpunan label 1 merupakan anggota himpunan label 2. Dengan kata lain label 1 merupakan himpunan bagian dari label 2. Dari hasil spesifikasi ini kemudian akan dapat dihitung persentase cytoplasma tiap sel tersebut.



Gambar 5. Spesifikasi nuclei untuk tiap cytoplasma

Pada gambar diatas dapat dilihat bahwa setiap sel yang tersegment memiliki area nuclei dan cytoplasma. Sampel diatas (gambar 5) merupakan sampel dengan tipe preparasi *thin prep*. Hasil interpretasi Bethesda menunjukkan slide diatas berada pada kondisi *ASC-H versus HSIL* (High squamous Intra epithelial Lesion). Kondisi ini menunjukkan kondisi pasien telah memasuki CIN II (*carcinoma in situ*). CIN merupakan gugusan kanker pra kanker ganas yang masih berada pada tempat awalnya dan belum berpindah menyerang organ vital yang lainnya.

Dari hasil spesifikasi yang telah dilakukan maka dapat dilakukan perhitungan persentase cytoplasma untuk setiap sel yang ada pada slide preparasi citra pap smear. Seperti diperlihatkan pada table 3. Table ini merupakan perhitungan persentase cytoplasma untuk sampel pada gambar 5.

Tabel 3. Perhitungan luas salah satu sampel

No	Luas sel (*)	Luas Nuclei(*)	% cytoplasm
1	3.3613	15043	55
2	42519	20921	51
3	20939	10956	48
5	35283	17856	49
6	35283	21370	39
7	41237	19897	52

(\*)Luas pada area ini dalam pixel

Dari hasil perhitungan diatas dapat dilihat bahwa persentase cytoplasma untuk setiap sel jauh lebih kecil dari pada persentase cytoplasma pada kondisi normal (table 2). Pada kondisi normal persentase terkecil cytoplasma terkecil dimiliki oleh tipe sel *Endocervical* dan *Reserve Cells* yaitu hanya 75%.

Sementara itu pada table 3 persentase terkecil mencapai 39%.

Jika dilihat berdasarkan hasil interpretasi NCI Bethesda system yang menunjukkan bahwa terdapat kelainan (*abnormal sel*) pada citra pap smear tersebut. Selain itu jumlah persentasi cytoplasma pada sel *abnormal* lebih sedikit dibandingkan dengan sel normal (Hoda, 2007). Sehingga dapat dikatakan bahwa kecilnya persentase cytoplasma pada sebuah sel dapat menjadi indikasi adanya kelainan pada sel tersebut.

#### 4. KESIMPULAN

Dari pembahasan diatas dapat disimpulkan bahwa metode gradient dan deteksi tepi merupakan salah satu model segmentasi yang dapat digunakan dalam proses segmentasi cytoplasma. Meskipun secara visual masih diperlukan beberapa perbaikan terutama dalam estimasi segmentasi untuk sel berkerumun.

Dari hasil kalkulasi area untuk beberapa sel diperoleh persentase cytoplasma yang jauh lebih kecil dari persentase untuk sel normal yang dapat menjadi indikasi adanya kelainan.

#### PUSTAKA

- Astana,M. 2009. *Bersahabat dengan kanker*. Araska. Bantul, Yogyakarta.
- Gonzales, R.C. and Woods, R.E, 2002. *Digital Image Processing, 2<sup>nd</sup> ed.*,Prentice Hall. Upper Saddle River, NJ.
- Hoda, R.S dan Hoda, S,A, 2007. *Fundamentals of Pap Test Cytology*, Humana Press, Totowa, New Jersey.
- Muhimmah, et al, 2012. *Automatic Epithelial Cells Detection of Pap smears images using Fuzzy C-Means Clustering*. International Conference on Bioinformatics and Biomedical Technology IPCBEE vol.29 (2012), IACSIT Press, Singapore
- Plissiti,M.E, et al. 2010. *Accurate localization of cell nuclei in pap smear images using gradient vector flow deformable models.*, BIOSIGNALS 2010 - International Conference on Bio-inspired Systems and Signal Processing.
- Rosidi, B., et al, 2011, *Classification of cervical cells based on labeled colour intensity distribution*, International journal of biology and biomedical engineering.
- Shah, S., 2007. *Automatic Cell Image Segmentation usingaShape ClassificationModel*.
- Solomon, C. and Breckon, T. 2011, “*Fundamentals of Digital Image Processing*” Wiley-Blackwell, west Sussex, UK.
- Supriyanto, E., et al. *Automatic Detection of Cervical Cancer Cells Using Color Intensity Clasfication*, Recent Researches in Computer Science.
- Tsai,M. et all. 2008. *Nucleus and cytoplasm contour detector of cervical smear image*. Pattern Recognition Letters 29 (2008) 1441–1453