

# **Metode *Gram-Schmidt* dan *Watershed Distance Transform* untuk Pemisahan Sel Kanker Payudara**

## **(Gram-Schmidt Method and Watershed Distance Transform for Breast Cancer Cell Separation)**

**Desmin Tuwohingide, Miske Silangen, dan Christian Koloay**

Jurusan Teknik Komputer dan Komunikasi, Politeknik Negeri Nusa Utara

Email: desmin.tuwohingide@gmail.com; miske.silangen@yahoo.com; christiankoloay@gmail.com

**Abstrak:** Penerapan pengolahan citra dalam membantu melakukan analisis citra mikroskopis semakin berkembang, salah satunya adalah penelitian tentang pemisahan sel kanker payudara bertumpuk. Hasil pemisahan sel kanker payudara akan sangat mempengaruhi akurasi perhitungan jumlah sel kanker. Metode watershed adalah metode segmentasi yang paling umum digunakan dan dikembangkan dalam pemisahan sel kanker bertumpuk pada penelitian sebelumnya. Berdasarkan penelitian sebelumnya juga diketahui bahwa metode Gram-Schmidt mampu menghasilkan nilai Mean Square Error yang relatif lebih rendah dibandingkan metode lainnya sehingga pada makalah ini diusulkan Pemisahan sel kanker payudara menggunakan metode Watershed Distance Transform berdasarkan metode Gram-Schmidt. Dataset yang digunakan pada penelitian ini adalah citra mikroskopis sel kanker payudara jenis benign dan malignant. Kombinasi metode yang digunakan pada penelitian ini menghasilkan nilai MSE 0,09041 dan nilai akurasi pemisahan sel kanker sebesar 71,02%.

**Kata Kunci:** pengolahan citra, citra mikroskopis sel kanker payudara, gram-schmidt, watershed distance transform

*Abstract: Implementation of image processing in helping perform microscopic image analysis is growing, one of them is research about separation of stacked breast cancer cells. Breast cancer cell separation results will greatly affect the accuracy of calculating the number of cancer cells. The watershed method is the most commonly used segmentation method and developed in the stacked cancer cell segregation in previous studies. Based on previous research also known that Gram-Schmidt method is able to produce Mean Square Error value is relatively lower than other methods so that this paper proposed Separation of breast cancer cells using Watershed Distance Transform method based on Gram-Schmidt method. The dataset used in this study is a microscopic image of benign and malignant breast cancer cells. Combination method used in this research yield value of MSE 0,09041 and accuracy value of cancer cell separation equal to 71,02%.*

**Keywords:** image processing, microscopic image of breast cancer cells, gram-schmidt, watershed distance transform

Berdasarkan data statistik *World Health Organization* (WHO) dalam *Cancer Country Profiles* tahun 2014 untuk Indonesia, kanker payudara menempati urutan pertama penyebab kematian perempuan yaitu 21,4% kematian dari 92.200 kasus kematian bahkan jumlah kasus perempuan terdiagnosis kanker payudara mencapai 48.998

kasus dan menjadi kasus kanker terbanyak di Indonesia.

Penerapan pengolahan citra untuk menganalisis citra biomedik telah banyak dilakukan, salah satunya adalah segmentasi citra sel kanker payudara. Tahapan segmentasi citra sel kanker ini merupakan salah satu tahapan pengolahan citra yang efektif

dan efisien untuk mengamati jumlah sel kanker. Identifikasi sel kanker pada citra mikroskopis sel kanker payudara menjadi sangat menantang karena keberagaman karakteristik sel kanker mulai dari warna, tekstur dan ukuran sehingga penggunaan metode tradisional dinilai belum memberikan hasil yang maksimal. Hal ini dibuktikan dalam penelitian (Aini, 2015), dimana rata-rata nilai *Mean Square Error* (MSE) yang dihasilkan oleh metode tradisional adalah 0.26 sedangkan metode *Region-Based Active Contour* menghasilkan rata-rata 0.0870. Pada penelitian (Tuwohingide, 2017) dalam identifikasi sel kanker metode *Gram-Schmidt* menghasilkan nilai rata-rata MSE yang lebih baik dibandingkan dengan metode *Region-Based Active Contour*, yaitu 0.0736 sedangkan untuk *Region-Based Active Contour* adalah 0,6505.

Segmentasi dan pemisahan sel bertumpuk dengan algoritma marker controlled *watershed* berdasarkan kanal warna RGB dan deteksi sel kanker payudara menggunakan perhitungan luas area sel kanker yang dibedakan kedalam tiga kategori yaitu *small* yang dianggap *noise*, *medium* dianggap sel tunggal dan *large* dianggap sel bertumpuk (Phukpattaranont & Boonyaphiphat, 2007). Penelitian dalam bidang medis khususnya untuk segmentasi citra mikroskopis sel kanker payudara telah di teliti secara berkelanjutan oleh Mouelhi, dkk. Pada tahun 2011, dilakukan penelitian dengan fokus penelitian untuk memisahkan sel kanker payudara yang bertumpuk menggunakan algoritma *watershed* berdasarkan *concave point detection* dan *selecting separating path* untuk pemilihan garis pemisah sel bertumpuk yang lebih akurat (Mouelhi, Sayadi, & Fnaiech, 2011) selanjutnya pada tahun 2013 melakukan segmentasi citra sel kanker payudara, menggunakan klasifikasi *multilayer neural network* dan *color active contour* (Mouelhi, Sayadi, & Fnaiech, 2013) dan pada tahun yang sama juga dillakukan penelitian mengenai segmentasi dan pemisahan sel bertumpuk pada citra mikroskopis sel kanker payudara dengan *color active contour* dan algoritma *watershed* berdasarkan *concave point detection* dan pemilihan *optimum path* untuk pemilihan garis pemisah sel (Mouelhi, Sayadi, Fnaiech, & Mrad, 2013). Metode *Watershed* tradisional menjadi metode yang sangat umum dan dianggap sangat tepat diterapkan untuk pemisahan sel bertumpuk, namun metode ini memiliki kelemahan yaitu *oversegmentation*, sehingga banyak penelitian-penelitian yang mengembangkan

metode ini. Seperti dalam penelitian (Hari, Prasad, Rao, 2014) yang menerapkan *Watershed Distance Transform* untuk pemisahan sel kanker darah berdasarkan fitur geometri dengan menerapkan perhitungan jarak *euclidean*, *quasi euclidean*, *city block* dan *chessboard*. Berdasarkan perhitungan jumlah sel kanker darah menghasilkan nilai *oversegmentation* yang sangat kecil.

Berdasarkan penelitian terdahulu dan uraian masalah diatas, maka pada makalah ini diusulkan penelitian untuk pemisahan sel kanker bertumpuk pada citra mikroskopis sel kanker payudara menggunakan metode *Watershed Distance Transform* berdasarkan identifikasi sel menggunakan metode *Gram-Schmidt* yang diharapkan mampu memberikan hasil perhitungan sel kanker yang lebih akurat.

## METODE PENELITIAN

Secara umum proses pemisahan sel kanker bertumpuk pada citra mikroskopis sel kanker payudara dapat dilihat pada Gbr 2.2. Pada proses pertama dilakukan *pre-processing* berupa *cropping* dataset agar sesuai dengan *groundtruth* dan penghapusan *noise* pada citra untuk memudahkan identifikasi atau segmentasi sel kanker. Proses selanjutnya adalah identifikasi sel kanker atau *cells segmentation* dengan menggunakan metode *Gram-Schmidt*. Kemudian setelah sel kanker teridentifikasi dilakukan pemisahan sel kanker yang bertumpuk atau *cells separation* menggunakan metode *Watershed Distance Transform* berdasarkan ukuran piksel citra.

### *Gram-Schmidt*

Sebuah himpunan orthogonal di mana setiap vektor mempunyai norm 1, dinamakan himpunan orthonormal. Suatu himpunan yang bukan orthonormal dapat diubah menjadi himpunan orthonormal dengan menggunakan proses *gram-schmidt*. Misalkan  $S = \{v_1, v_2, \dots, v_n\}$  dan menghasilkan sebuah basis orthogonal  $S = \{u_1, u_2, \dots, u_n\}$  yang memiliki subruang sama dengan  $S$  yang diperoleh berdasarkan proses *Gram Schmidt*. Proyeksi *Gram Schmidt* dijelaskan pada persamaan 2.1 (Rezatofighi, 2009).

$$Proj_u v = \frac{\langle u, v \rangle}{\langle u, u \rangle} u = \langle u, v \rangle \frac{u}{\langle u, u \rangle} \quad (2.1)$$

Di mana  $\langle u, v \rangle$  adalah ruang hasil kali dalam dan vektor  $u$  dan vektor  $v$ . Proses *Gram Schmidt*

dapat dilihat pada persamaan berikut:

$$u_1 = v_1, \quad e_1 = \frac{u_1}{\|u_1\|} \quad (2.2)$$

$$u_2 = v_2 - \text{proj}_{u_1} v_2 \quad e_2 = \frac{u_2}{\|u_2\|}$$

$$u_3 = v_3 - \text{proj}_{u_1} v_3 - \text{proj}_{u_2} v_3 \quad e_3 = \frac{u_3}{\|u_3\|}$$

$$u_k = v_k - \sum_{j=1}^{k-1} \text{proj}_{u_j} v_k \quad e_k = \frac{u_k}{\|u_k\|}$$

$u_1, \dots, u_k$  diperlukan untuk vektor orthogonal, dan vektor normal  $e_1, \dots, e_k$ , membentuk himpunan orthonormal. Berdasarkan metode tersebut, untuk , dapat ditemukan vektor (yang diinginkan) yang memiliki *orthogonality* maksimum dan *orthogonality* minimum untuk vektor yang lain sesuai dimensi subruang (). Vektor dapat dihitung dengan persamaan 2.3 dan hasil *inner product* dalam set S oleh  $w_k$  sesuai dengan persamaan 2.4 (Rezatofighi, 2009).

$$w_k = v_k - \sum_{j=1}^{k-1} \text{proj}_{v_j} v_k \quad (2.3)$$

$$\begin{cases} \langle v_j, w_k \rangle = 0 & j \in 1, \dots, n \text{ dan } j \neq k \\ \langle v_k, w_k \rangle = K & K \neq 0 \end{cases} \quad (2.4)$$

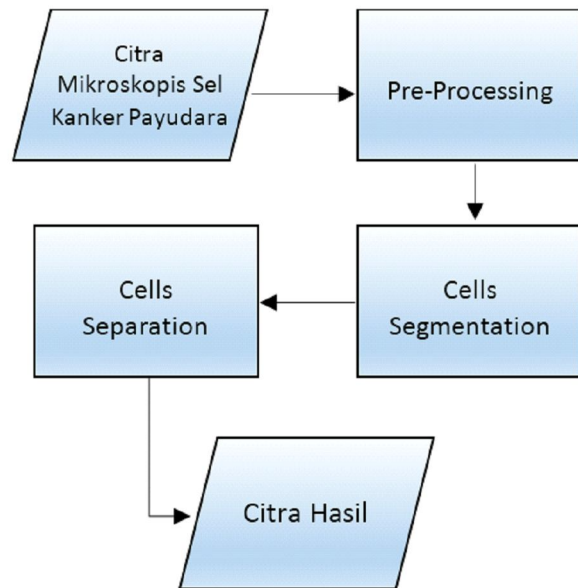
### Watershed Distance Transform

Algoritma *watershed* adalah salah satu algoritma segmentasi citra yang berdasarkan pada morfologi matematika. Algoritma *watershed* memiliki konsep dasar memvisualisasikan sebuah gambar pada tiga dimensi, dua ruang koordinat versus level gray seperti sebuah interpretasi topographic, yang ditentukan dalam tiga jenis titik:

- Titik yang merupakan sebuah regional minimum
- Titik dimana jika tetesan air ditempatkan atau jatuh pada lokasi ini, akan menjadi sebuah minimum tertentu.
- Titik dimana air akan seperti jatuh ke lebih dari satu posisi minimum (kemungkinan akan jatuh ke salah satu titik minimum tetapi dapat jatuh ke titik minimum tertentu atau titik minimum lainnya)

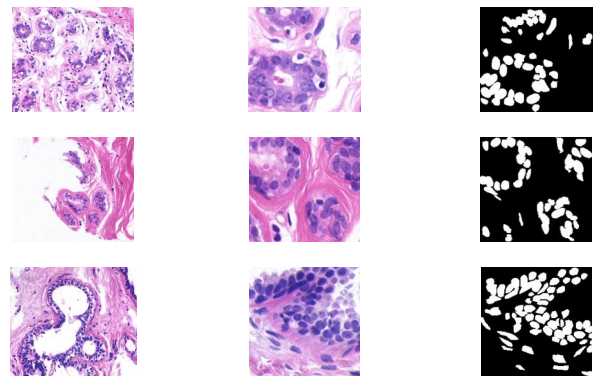
*Distance transform* adalah *tools* yang secara umum digunakan dalam konjungsi dengan transformasi *watershed* untuk segmentasi. Transformasi jarak dari citra biner adalah konsep yang relatif sederhana; jarak dari setiap piksel ke nilai piksel *non-zero* terdekat (Prasetyo, 2011).

## HASIL DAN PEMBAHASAN



Gambar 1. Gambaran Umum Sistem

Sesuai dengan tahapan penelitian yang dijelaskan pada gambar 1. Citra dataset pada penelitian ini adalah citra mikroskopis sel kanker payudara. Pada tahapan *pre-processing* Citra mikroskopis RGB di *cropping* mengikuti citra *groundtruth* seperti yang diperlihatkan pada gambar 2.



Gambar 2. Citra Dataset (a) Citra Mikroskopis (b) Citra Mikroskopis Hasil Cropping (c) Citra Groundtruth

### Dataset

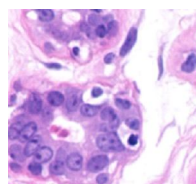
Dataset atau data uji coba yang digunakan pada penelitian ini adalah berupa citra mikroskopis sel kanker payudara, yang berukuran 896 x 768 piksel, seperti pada gambar 2. Citra mikroskopis sel kanker payudara yang digunakan sebagai dataset diunduh dari laman [www.bioimage.ucsb.edu](http://www.bioimage.ucsb.edu).

## Cells Segmentation

*Cells Segmentation* adalah tahapan segmentasi untuk melakukan identifikasi sel kanker. Sel kanker yang diidentifikasi adalah *blue nuclei* atau sel yang berwarna biru dan biru keunguan. Nilai-nilai vektor yang digunakan pada penelitian ini ditentukan berdasarkan penelitian sebelumnya (Tuwohingide, 2017).

## Cells Separation

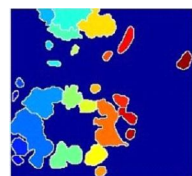
Pada proses ini, citra masukan merupakan citra hasil dari proses *cells segmentation*, kemudian diimplementasikan metode *watershed distance transform* untuk *cells separation*. Citra hasil yang peroleh pada tahapan *cells segmentation* yang digunakan sebagai citra masukan pada tahapan ini dan hasil yang diperoleh pada tahapan *cells separation* disajikan pada gambar 3 Kemudian dilakukan perhitungan manual jumlah sel pada citra hasil dan dibandingkan dengan jumlah sel pada citra asli, yang disajikan pada Tabel 1.



(a)



(b)



(c)

**Gambar 3. Citra Hasil Cells Separation (a) Citra Asli (b) Citra Masukan (c) Citra Hasil**

**Tabel 1. Hasil Perhitungan Manual Jumlah Sel**

Jumlah Sel Citra Asli	Jumlah Sel Citra Hasil			Akurasi
	Under	Over	Benar	
1391	400	3	988	71,02%

## KESIMPULAN

Berdasarkan hasil implementasi dan pengujian sistem yang dijelaskan pada bab sebelumnya, maka dapat diambil kesimpulan sebagai berikut:

- 1) Pemisahan sel kanker bertumpuk pada citra mikroskopis sel kanker payudara menggunakan kombinasi metode *Gram-Schmidt* dan *Watershed Distance Transform* menghasilkan nilai rata-rata MSE 0,09041 dan nilai akurasi pemisahan sel kanker sebesar 71,02%
- 2) Kombinasi metode yang digunakan mampu mengurangi masalah *oversegmentation* namun menghasilkan jumlah sel *undersegmentati* lebih tinggi dibandingkan dengan metode *Spatial Fuzzy C-Means (SFCM)*.
- 3) Mengingatnya jumlah sel *undersegmentation* terjadi karena setelah proses *cells segmentation* tidak dilakukan deteksi sel bertumpuk hal ini dikarenakan karakteristik citra mikroskopis yang berbeda dari segi bentuk, ukuran dan warna sehingga diperlukan pengembangan lebih lanjut mengenai deteksi sel kanker bertumpuk pada citra mikroskopis sel kanker payudara.

## DAFTAR RUJUKAN

- <https://who.int>, di akses April 2017.
- Aini, N. 2015. *Metode Hybrid Region-Based Active Contour dan Bayesian untuk Segmentasi dan Pemisahan Sel Bertumpuk pada Citra Mikroskopis Sel Kanker Payudara*.
- J. Hari., A. San Prasad, S. Koteswara Rao. 2014. Separation and Counting of Blood Cells Using Geometrical Features and Distance Transformed Watershed. *2014 2nd International Conference On.Devices, Circuits and System (ICDCS)*.
- Mouelhi, A., Sayadi, M., & Fnaiech, F. 2011. Automatic segmentation of clustered breast cancer cells using watershed and concave vertex graph. *2011 International Conference on Communications, Computing and Control Applications, CCCA (1), 2–7*.
- Mouelhi, A., Sayadi, M., & Fnaiech, F. 2013. A Supervised Segmentation Scheme Based on Multilayer Neural Network and Color Active Contour Model for Breast Cancer Nuclei Detection, (1).

- Mouelhi, A., Sayadi, M., Fnaiech, F., & Mrad, K. 2013. Biomedical Signal Processing and Control Automatic image segmentation of nuclear stained breast tissue sections using color active contour model and an improved watershed method. *Biomedical Signal Processing and Control*, 8(5), 421–436.
- Phukpattaranont, P., & Boonyaphiphat, P. 2007. Color Based Segmentation of Nuclear Stained Breast Cancer Cell Images. *Communications*, 5(2), 158–164.
- Prasetyo, E. 2011. *Pengolahan Citra Digital Dan Aplikasinya Menggunakan MATLAB*, Penerbit Andi.
- Rezatofighi, S.H. 2009. International Conference on Digital Image Processing A New Approach to White Blood Cell Nucleus Segmentation Based on Gram-Schmidt Orthogonalization, 107–111.
- Tuwohingide, D., Faticah, C. 2017. Spatial Fuzzy C-Means dan Rapid Region Merging untuk Pemisahan Sel Kanker Bertumpuk, *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*.