

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kanker leher rahim (kanker serviks) adalah kanker yang terjadi pada servik *uterus*, suatu daerah pada organ reproduksi wanita yang merupakan pintu masuk ke arah rahim yang terletak antara rahim dengan liang senggama (*vagina*). Kanker ini biasanya terjadi pada wanita yang telah berumur, tetapi bukti statistik menunjukkan bahwa kanker leher rahim dapat juga menyerang wanita yang berumur antara 20 sampai 30 tahun (Riono, 1999).

Frekuensi kanker serviks banyak dijumpai di negara-negara berkembang seperti Indonesia, India, Bangladesh, Thailand, Vietnam dan Filipina. Frekuensi kanker serviks di Amerika Latin dan Afrika Selatan menempati posisi tertinggi dari penyakit keganasan yang ada (Tambunan, 1995). Insiden kanker serviks di Indonesia, diperkirakan ± 40.000 kasus pertahun dan masih merupakan kanker perempuan yang tersering. Mortalitas kanker serviks masih tinggi karena $\pm 90\%$ terdiagnosis pada stadium invasif, lanjut bahkan terminal (Suwiyoga, 2006).

Kendala yang dihadapi oleh Indonesia saat ini adalah jumlah Dokter Patologi Anatomi yang masih minim. Patologi Anatomi adalah ilmu mengenai struktur tubuh dan perubahan yang berkaitan dengan penyakit atau cedera (Sloane, 2003). Dikutip dari (Humas Universitas Indonesia, 2007), "Indonesia hanya memiliki 220 Dokter spesialis Patologi Anatomi." Penyebaran 220 Dokter spesialis Patologi Anatomi tersebut pun tidak merata. Daerah seperti Papua dan Ambon, mereka tidak memiliki Dokter spesialis Patologi Anatomi.

Menurut (Humas Universitas Indonesia, 2007) “Dokter Patologi Anatomi kurang diminati karena banyak yang harus dipelajari.” Dokter spesialis anak akan fokus mempelajari anak begitu juga Dokter spesialis jantung. Namun, Dokter Patologi Anatomi mempelajari semuanya. Salah satu hal yang menjadi tanggungjawab Dokter Patologi Anatomi adalah mendiagnosis sel serviks. Diagnosis sel serviks dapat dilakukan dengan menggunakan tes *Pap smear*. Seperti yang dikutip dari (Indman, 2000) *Pap smear* didefinisikan sebagai:

“Beberapa sel leher rahim akan dikerok selama *Pap smear* dan disusun berdasarkan *slide* yang dapat diperiksa di bawah mikroskop. Tujuan dari tes *Pap smear* adalah untuk mendeteksi perubahan yang dapat menyebabkan kanker serviks jauh sebelum kanker berkembang.”

Proses interpretasi hasil dari *Pap smear* sendiri masih mengalami beberapa hambatan. Seperti yang dikutip dari (Plissiti, et al., 2011) menyebutkan:

“Interpretasi visual citra *Pap smear* tersebut sangat rumit, membutuhkan waktu yang lama dan rawan kesalahan prosedur dalam banyak kasus. Hal ini disebabkan karena *Pap smear* konvensional memberikan hasil sel yang memiliki *layer* tidak merata, berkerumun, dan sel saling tumpang tindih. Masalah prosedur pewarnaan, terdapat darah yang berlebihan, lendir, bakteri dan pengeringan udara menyebabkan proses pengenalan sel menjadi sulit.

Pemeriksaan dengan tes *Pap smear* tidak jarang mengalami kegagalan dalam analisisnya. Kegagalan analisis disebabkan karena sel-sel *abnormal* yang mengindikasikan adanya pertumbuhan kanker tidak terambil sebagai sel-sel yang diamati. Kegagalan analisis tersebut sekarang sudah dapat diatasi dengan pemeriksaan sitologi berbasis cairan.

“Sitologi berbasis cairan (LBC) adalah sebuah metode baru dalam menyiapkan sampel untuk pemeriksaan sitologi serviks. Berbeda dengan persiapan konvensional

'Pap', LBC melibatkan pembuatan suspensi sel dari sampel dan ini digunakan untuk menghasilkan lapisan sel-sel tipis pada *slide*.”(Karnon, et al., 2004).

Pengambilan sampel pada pemeriksaan sitologi berbasis cairan ini, dilakukan dengan cara yang sama seperti pada pemeriksaan sitologi biasa. Sampel sel tersebut terlebih dahulu dicelupkan ke dalam cairan khusus dan diproses dengan alat otomatis sebelum sampel tersebut diusapkan pada kaca benda. Sample yang telah dicelupkan tersebut kemudian diwarnai dan dilihat di bawah mikroskop oleh seorang Dokter ahli Patologi Anatomi.

Hasil penelitian diharapkan dapat membantu Dokter ahli Patologi Anatomi dalam menghitung jumlah sel *epitel* serviks. Penghitungan sel *epitel* serviks dapat dilakukan dengan cara menghitung jumlah sel *nuclei* pada citra *Pap smear*. Penghitungan sel *epitel* serviks ini merupakan standar prosedur yang harus dilakukan dalam penentuan kelayakan *slide* preparat, namun dalam prakteknya tidak dilakukan karena adanya kesulitan dalam pembacaan manual melalui mikroskop seperti yang telah dijelaskan pada paragraf sebelumnya. Adanya kesulitan dalam pembacaan manual melalui mikroskop ini maka diperlukan suatu inovasi dalam menganalisis citra hasil dari tes *Pap smear* dengan bantuan komputer sehingga diharapkan proses diagnosis menjadi lebih cepat dan lebih mudah.

1.2 Rumusan Masalah

Secara lebih spesifik, rumusan masalah untuk penelitian ini adalah:

Bagaimana membuat sebuah model yang dapat menghitung jumlah sel epitel serviks pada citra *Pap Smear* secara otomatis, serta menghitung tingkat *sensitivity*, *specificity*, dan kinerja sistem dari model yang diusulkan?

1.3 Batasan Masalah

Batasan masalah penelitian terhadap penghitungan sel *epitel* serviks dari test *Pap-Smear* adalah sebagai berikut:

1. Penelitian yang dilakukan, hanya melakukan proses otomatisasi deteksi jumlah sel *epitel* serviks dengan cara mendeteksi jumlah *nuclei* pada citra *Pap smear*, sedangkan untuk penelitian persiapan dan pengolahan preparat sel serviks menjadi citra digital tidak dilakukan.
2. Citra yang digunakan untuk sampel merupakan koleksi data citra test *Pap-Smear* dari NCI *Bethesda System* yang dapat diunduh dari Internet (<http://nih.techriver.net/>) dengan tipe preparat *Liquid Based Preparation* (LBP)/ *ThinPrep*. Citra tersebut merupakan citra digital yang difoto melalui mikroskop dengan perbesaran *high* dan perbesaran *medium*

1.4 Tujuan Penelitian

Hal yang akan dicapai dalam penelitian ini yaitu mengusulkan sebuah model untuk menghitung jumlah *nuclei* secara otomatis pada citra *Pap Smear*.

1.5 Manfaat Penelitian

Beberapa manfaat yang diperoleh dari penelitian dapat dirasakan bagi objek penelitian, manfaat-manfaat tersebut antara lain:

1. Dapat membantu Dokter ahli Patologi Anatomi dalam menghitung jumlah sel *epitel* serviks sebagai standar prosedur yang harus dilakukan dalam penentuan kelayakan *slide* preparat.
2. Dapat menjadi prototipe dan dapat dikembangkan dengan bahasa pemrograman yang lainnya dari model yang diusulkan kedepannya.

1.6 Metodologi Penelitian

Dataset yang digunakan dalam penelitian ini berjumlah 13 sampel dari 13 *slide* preparasi berbeda yang merupakan koleksi data citra tes *Pap smear* dari NCI Bethesda System dengan tipe preparat *Liquid Based Preparation (LBP)/ ThinPrep*. Penggolongan berdasarkan perbesaran mikroskopnya terdiri dari 7 jumlah data dengan perbesaran *high* dan 6 data dengan perbesaran *medium*.

Model yang ditawarkan awalnya meniru penelitian yang dilakukan (Plissiti, et al., 2011) yaitu melakukan segmentasi sel serviks dengan *Morphological Reconstruction* dan *Clustering*. Penelitian tersebut merupakan referensi utama pada penelitian ini. Model yang diadopsi dan model yang menjadi hasil perbaikan yang dilakukan pada penelitian ini akan dibahas pada bab selanjutnya.

Penelitian ini dibantu dengan program yang dibuat dengan perangkat lunak MATLAB untuk melakukan penerapan model yang diusulkan untuk menganalisis citra sel epitel serviks. Keluaran yang diperoleh adalah jumlah sel *nuclei* yang terdeteksi dan citra hasil segmentasi yang dapat menunjukkan posisi sel *nuclei*.

Penelitian yang dilakukan terdapat 5 tahap utama untuk penghitungan jumlah sel *epitel* serviks, antara lain: 1) *Preprocessing*; 2) Deteksi *Centroids* Kandidat Sel *Nuclei*; 3) Perbaikan Posisi Kandidat Sel *Nuclei*; 4) Pengurangan Jumlah *Centroids* dengan Kriteria Jarak; dan 5) *Clustering*. Penjelasan lebih lanjut akan dijelaskan pada bab selanjutnya.

Model yang ditawarkan akan dibandingkan dengan hasil evaluasi dari dokter ahli Patologi Anatomi yang berupa hasil pengamatan secara manual. Hasil dari keluaran sistem akan dievaluasi oleh Dokter ahli Patologi Anatomi dengan menggunakan *Single Decision Threshold* untuk mencari tingkat *sensitivity*, *specificity*, dan kinerja sistem. Hasil perbandingan antara model yang ditawarkan dengan penilaian pakar diharapkan berupa evaluasi kuantitatif.

1.7 Sistematika Penulisan

Penulisan laporan penelitian tersusun atas beberapa bab dan masing-masing bab dibagi menjadi beberapa sub bab. Sistematika penulisannya adalah sebagai berikut:

BAB 1 Pendahuluan, merupakan pembahasan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metodologi penelitian, dan sistematika penulisan laporan pada penelitian ini.

BAB 2 Landasan Teori, merupakan pembahasan mengenai landasan teori yang mendukung penelitian. Bab ini dimulai dengan sekilas tentang sel serviks dan *Pap smear*, kemudian dilanjutkan dengan teori pengolahan citra secara umum dan penelitian terkait yang pernah dilakukan sebelumnya untuk proses deteksi dan *clustering*.

BAB 3 Metodologi, merupakan penjelasan tentang metode analisis yang digunakan, rancangan model dan algoritma yang telah dimodifikasi untuk digunakan dalam analisis citra hasil tes *Pap smear*.

BAB 4 Preprocessing, merupakan penjelasan tahap *preprocessing* model yang diusulkan, dan pembahasan pengujian metode *thresholding*.

BAB 5 Deteksi Centroids Kandidat Sel Nuclei, merupakan penjelasan tentang model yang digunakan, pembahasan parameter yang digunakan, dan analisis permasalahan pada *dataset* yang digunakan .

BAB 6 Perbaikan Posisi Kandidat Sel Nuclei, merupakan penjelasan tentang model yang digunakan dalam analisis citra *Pap smear*.

BAB 7 Pengurangan Jumlah Centroids dengan Kriteria Jarak, merupakan penjelasan tentang model yang digunakan, pembahasan parameter jarak, pembahasan penggunaan intensitas warna dan analisis permasalahan pada *dataset* yang digunakan.

BAB 8 *Clustering*, merupakan penjelasan tentang model yang digunakan, dan pembahasan kriteria jarak terdekat pada tahap *clustering*.

BAB 9 Pengujian dan Evaluasi, menjelaskan tentang pembahasan waktu eksekusi model yang diusulkan, pengujian parameter dan pembahasan pengurangan *centroids* pada tiap tahap model yang diusulkan. Bab ini juga membahas tentang perbesaran lensa, evaluasi tentang pembacaan manual yang dilakukan oleh dokter, evaluasi hasil model yang diusulkan dan pengujian model dengan tipe preparasi yang lain.

BAB 10 Kesimpulan dan Saran, merupakan kesimpulan berupa rangkuman-rangkuman hasil penelitian dan saran-saran mengenai keterbatasan yang ditemukan selama proses pembuatan laporan penelitian.

BAB II

LANDASAN TEORI

2.1 Kanker Serviks

2.1.1 Pengertian Kanker Serviks

“Kanker leher rahim (kanker serviks) adalah kanker yang terjadi pada servik *uterus*, suatu daerah pada organ reproduksi wanita yang merupakan pintu masuk ke arah rahim yang terletak antara rahim dengan liang senggama (*vagina*). Kanker ini biasanya terjadi pada wanita yang telah berumur, tetapi bukti statistik menunjukkan bahwa kanker leher rahim dapat juga menyerang wanita yang berumur antara 20 sampai 30 tahun” (Riono, 1999).

“Proses terjadinya kanker serviks sangat erat hubungannya dengan proses *metaplasia*. Masuknya *mutagen* atau bahan - bahan yang dapat mengubah perangai sel secara genetik pada saat fase aktif *metaplasia* dapat menimbulkan sel-sel yang berpotensi ganas.” (Sjamsuddin, 2001).

“Kanker serviks dimulai dengan adanya suatu perubahan dari sel serviks normal menjadi sel *abnormal* yang kemudian membelah diri tanpa terkendali. Sel serviks yang *abnormal* ini dapat berkumpul menjadi tumor yang bersifat jinak atau ganas.” (Rasjidi & Sulistiyanto, 2007).

2.1.2 Penyebab Kanker Serviks

Penyebab utama kanker servik adalah infeksi *Human Papilloma Virus* (HPV). Lebih dari 90% kanker serviks jenis *skuamosa* mengandung DNA virus HPV dan 50% kanker servik berhubungan dengan HPV tipe 16. Penyebaran Virus ini terutama melalui hubungan seksual (Aziz, et al., 2006). Terdapat beberapa faktor risiko selain

faktor penyebab utama dari infeksi HPV. Faktor risiko dari kanker serviks yaitu mulai melakukan hubungan seksual pada usia kurang dari 20 tahun, pemakaian kontrasepsi oral jangka panjang, merokok, dan defisiensi nutrisi yaitu kekurangan beta karoten, Vitamin C (Sjamsuddin, 2001).

2.1.3 Klasifikasi Kanker Serviks

The American Joint Committee on Cancer (AJCC) TNM dan the International Federation of Gynecology and Obstetrics (FIGO) membuat klasifikasi kanker serviks berdasarkan perkembangan secara klinis sesuai tabel 2.1 sebagai berikut (Boardman C.H, et al., 2011):

Tabel 2.1 Klasifikasi kanker serviks berdasarkan TNM dan FIGO

Primary tumor (T)		
TNM	FIGO	Surgical-Pathologic Findings
<i>Categories</i>	<i>Stages</i>	
TX		<i>Primary tumor cannot be assessed</i>
T0		<i>No evidence of primary tumor</i>
Tis		<i>Carcinoma in situ (preinvasive carcinoma)</i>
T1	I	<i>Cervical carcinoma confined to the cervix (disregard extension to the corpus)</i>
T1a	IA	<i>Invasive carcinoma diagnosed only by microscopy; stromal invasion with a maximum depth of 5.0 mm measured from the base of the epithelium and a horizontal spread of 7.0 mm or less; vascular space involvement, venous or lymphatic, does not affect classification</i>
T1a1	IA1	<i>Measured stromal invasion ≤ 3.0 mm in depth and ≤ 7.0 mm in horizontal spread</i>

T1a2	IA2	<i>Measured stromal invasion > 3.0 mm and ≤ 5.0 mm with a horizontal spread ≤ 7.0 mm</i>
T1b	IB	<i>Clinically visible lesion confined to the cervix or microscopic lesion greater than T1a/IA2</i>
T1b1	IB1	<i>Clinically visible lesion ≤ 4.0 cm in greatest dimension</i>
T1b2	IB2	<i>Clinically visible lesion > 4.0 cm in greatest dimension</i>
T2	II	<i>Cervical carcinoma invades beyond uterus but not to pelvic wall or to lower third of vagina</i>
T2a	IIA	<i>Tumor without parametrial invasion</i>
T2a1	IIA1	<i>Clinically visible lesion ≤ 4.0 cm in greatest dimension</i>
T2a2	IIA2	<i>Clinically visible lesion > 4.0 cm in greatest dimension</i>
T2b	IIB	<i>Tumor with parametrial invasion</i>
T3	III	<i>Tumor extends to pelvic wall and/or involves lower third of vagina and/or causes hydronephrosis or nonfunctional kidney</i>
T3a	IIIA	<i>Tumor involves lower third of vagina, no extension to pelvic wall</i>
T3b	IIIB	<i>Tumor extends to pelvic wall and/or causes hydronephrosis or nonfunctional kidney</i>
T4	IV	<i>Tumor invades mucosa of bladder or rectum and/or extends beyond true pelvis (bullous edema is not sufficient to classify a tumor as T4)</i>
T4a	IVA	<i>Tumor invades mucosa of bladder or rectum (bullous edema is not sufficient to classify a tumor as T4)</i>
T4b	IVB	<i>Tumor extends beyond true pelvis</i>
Regional lymph nodes (N)		
NX	<i>Regional lymph nodes cannot be assessed</i>	

N0	<i>No regional lymph node metastasis</i>
N1	<i>Regional lymph node metastasis</i>
<i>Distant metastasis (M)</i>	
MX	<i>No distant metastasis</i>
M0	<i>Distant metastasis (including peritoneal spread; involvement of supraclavicular, mediastinal, or para-aortic lymph nodes; and lung, liver, or bone)</i>

2.2 *Pap smear*

2.2.1. Definisi *Pap Smear*

Pap smear pertama kali diperkenalkan tahun 1928 oleh Dr. George Papanicolou dan Dr. Aurel Babel, namun mulai populer sejak tahun 1943 (Purwoto dan Nuranna, 2002). Tes *Pap smear* adalah pemeriksaan sitologi dari serviks dan *porsio* untuk melihat adanya perubahan atau keganasan pada *epitel* serviks atau *porsio* (displasia) sebagai tanda awal keganasan serviks atau prakanker (Rasjidi, 2008). *Pap smear* merupakan suatu metode pemeriksaan sel-sel yang diambil dari leher rahim dan kemudian diperiksa di bawah mikroskop. *Pap smear* merupakan tes yang aman dan murah dan telah dipakai bertahun-tahun lamanya untuk mendeteksi kelainan-kelainan yang terjadi pada sel-sel leher rahim (Diananda, 2009). Pemeriksaan ini mudah dikerjakan, cepat, dan tidak sakit, serta bisa dilakukan setiap saat, kecuali pada saat haid (Dalimartha, 2004).

2.2.2. Manfaat *Pap smear*

Dikutip dari (Nuranna, 2001), telah diakui bahwa pemeriksaan tes *Pap smear* mampu menurunkan kematian akibat kanker serviks di beberapa negara, walaupun tentu ada kekurangan. Sensitivitas tes *Pap smear* untuk mendeteksi NIS berkisar 50-98% sedang negatif palsu antara 8-30% untuk *lesi skuamosa* 40% untuk

adenomatosa. Spesifisitas tes *Pap smear* adalah 93%, nilai prediksi positif adalah 80.2% dan nilai prediksi negatif adalah 91.3%. Harus hati-hati justru pada *lesi* serviks invasif, karena negatif palsu dapat mencapai 50%, akibat tertutup darah, adanya radang dan jaringan *nekrotik*. Fakta ini menunjukkan bahwa pada lesi invasif kemampuan pemeriksa melihat serviks secara makroskopik sangat diperlukan.

Pemeriksaan *Pap smear* berguna sebagai pemeriksaan penyaring (skrining) dan pelacak adanya perubahan sel ke arah keganasan secara dini sehingga kelainan prakanker dapat terdeteksi serta pengobatannya menjadi lebih murah dan mudah (Dalimartha, 2004).

Pap smear mampu mendeteksi *lesi* prekursor pada stadium awal sehingga *lesi* dapat ditemukan saat terapi masih mungkin bersifat kuratif (Crum, et al., 2007). Manfaat *Pap smear* secara rinci dapat dijabarkan sebagai berikut (Manuaba, 2005):

a. Diagnosis dini keganasan

Pap smear berguna dalam mendeteksi dini kanker serviks, kanker *korpus endometrium*, keganasan *tuba fallopi*, dan mungkin keganasan *ovarium*.

b. Perawatan ikutan dari keganasan

Pap smear berguna sebagai perawatan ikutan setelah operasi dan setelah mendapat kemoterapi dan radiasi.

c. Interpretasi hormonal wanita

Pap smear bertujuan untuk mengikuti siklus menstruasi dengan ovulasi atau tanpa *ovulasi*, menentukan maturitas kehamilan, dan menentukan kemungkinan keguguran pada hamil muda.

- d. Menentukan proses peradangan

Pap smear berguna untuk menentukan proses peradangan pada berbagai infeksi bakteri dan jamur.

2.1.3. Interpretasi Hasil *Pap smear*

Terdapat banyak sistem dalam menginterpretasikan hasil pemeriksaan *Pap smear* yaitu: sistem Papanicolaou, sistem *Cervical Intraepithelial Neoplasia* (CIN), dan sistem Bethesda.

Klasifikasi Papanicolaou membagi hasil pemeriksaan menjadi 5 kelas (Saviano, 1993), yaitu:

- a. Kelas I: tidak ada sel *abnormal*.
- b. Kelas II: terdapat gambaran sitologi *atipik*, namun tidak ada indikasi adanya keganasan.
- c. Kelas III: gambaran sitologi yang dicurigai keganasan, *displasia* ringan sampai sedang.
- d. Kelas IV: gambaran sitologi dijumpai *displasia* berat.
- e. Kelas V: keganasan.

Klasifikasi ini banyak ditinggalkan karena (Kusuma dan Moegni, 2001):

1. Tidak mencerminkan pengertian *neoplasia* serviks/ *vagina*
2. Tidak memiliki padanan dengan terminologi histopatologi
3. Tidak mencantumkan diagnosis non kanker
4. Interpretasinya tidak seragam
5. Tidak menunjukkan pernyataan diagnosis

Pada sistem CIN, pengelompokan hasil uji *Pap smear* terdiri dari (Feig, 2001):

- a. CIN I merupakan *displasia* ringan dimana ditemukan sel *neoplasma* pada kurang dari sepertiga lapisan *epitelium*.
- b. CIN II merupakan *displasia* sedang dimana melibatkan dua pertiga *epitelium*.
- c. CIN III merupakan *displasia* berat atau *karsinoma in situ* yang dimana telah melibatkan sampai ke *basement membrane* dari *epitelium*.

Tahun 1988 dan 1991 pertemuan para ahli sitopatologi melahirkan sistem Bethesda sebagai sistem pelaporan sitopatologi baru yang bertujuan (Kusuma dan Moegni, 2001):

1. Menghilangkan kelas-kelas *Papanicolaou*
2. Menciptakan terminologi seragam memakai istilah diagnostik
3. Memasukkan pernyataan *adekuasi*
4. Membuat sitologi sebagai konsultasi medik antar ahli sitologi dan klinikus.

Selain ini sistem Bethesda juga mengandung unsur (Kusuma dan Moegni, 2001):

1. Komunikasi yang efektif antara ahli sitopatologi dan dokter yang merujuk
2. Mempermudah korelasi sitologi-histopatologi
3. Mempermudah penelitian epidemiologi, biologi dan patologi
4. Data yang dapat dipercaya untuk analisis statistik nasional dan internasional.

Kelebihan cara pelaporan *The Bethesda System* (TBS) adalah penyederhanaan terminologi dengan memakai terminologi diagnostik yang jelas untuk kategori umum (Kusuma dan Moegni, 2001):

1. Dalam batas normal
2. Perubahan seluler jinak
3. Abnormalitas sel *epitel*

Setelah melalui beberapa kali pembaharuan, maka saat ini digunakan klasifikasi Bethesda 2001. Klasifikasi Bethesda 2001 adalah sebagai berikut (Marquardt, 2002):

1. Sel *skuamosa*

- a. Atypical Squamous Cells Undetermined Significance (ASC-US)*
- b. Low Grade Squamous Intraepithelial Lesion (LSIL)*
- c. High Grade Squamous Intraepithelial Lesion (HSIL)*
- d. Squamous Cells Carcinoma*

2. Sel *glandular*

- a. Atypical Endocervical Cells*
- b. Atypical Endometrial Cells*
- c. Atypical Glandular Cells*
- d. Adenokarsinoma Endoservikal In situ*
- e. Adenokarsinoma Endoserviks*
- f. Adenokarsinoma Endometrium*
- g. Adenokarsinoma Ekstrauterin*
- h. Adenokarsinoma yang tidak dapat ditentukan asalnya (NOS)*

2.2.4 Penggolongan Citra *Pap smear*

Berikut ini beberapa penggolongan Citra *Pap Smear*

2.2.4.1 *Pap smear* Konvensional

Selama ini proses *Pap smear* konvensional meliputi peletakan sel-sel servik ke dalam slide, dengan sel-sel ini terlihat mucus, darah, dan peradangan sel-sel. Sel-sel ini sering mengering, pecah, dan menyusut yang menyebabkan sulitnya interpretasi morfologi *celluler* sehingga tidak cukupnya spesimen untuk diagnosis. Prosedur pengambilan sampel pada *Pap smear* konvensional kurang nyaman. *Cytologist* dan *Patologist* pun kurang nyaman dengan kualitas *slide* untuk diagnosis. *Pap smear* konvensional memiliki tingkat sensitivitas yang rendah yakni hanya 51% (*Confirmed by the Agency for Health Care Policy and Research (AHCPR) after reviewing over 600 publications*).

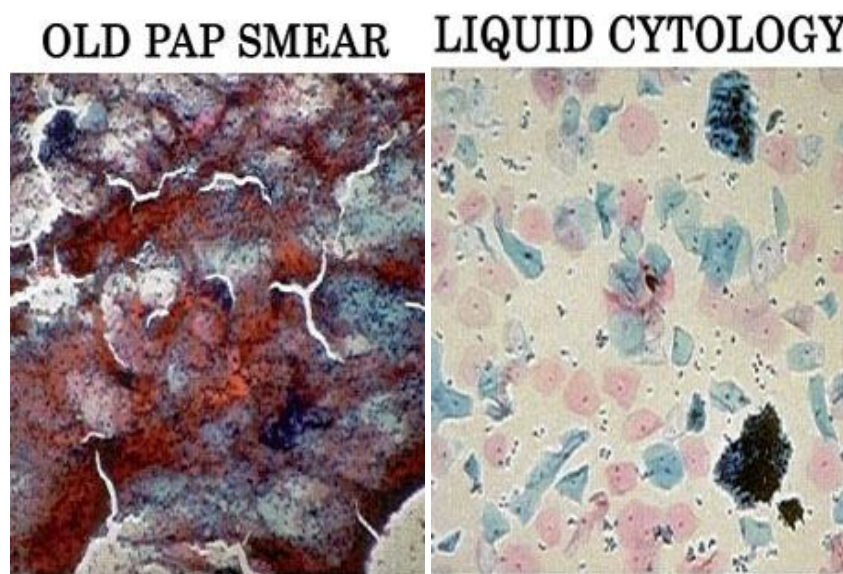
2.2.4.2 *Liquid-Based Cytology* (LBC)

“Sitologi berbasis cairan (LBC) adalah sebuah metode baru dalam menyiapkan sampel untuk pemeriksaan sitologi serviks. Berbeda dengan persiapan konvensional 'Pap', LBC melibatkan pembuatan suspensi sel dari sampel dan ini digunakan untuk menghasilkan lapisan sel-sel tipis pada *slide*.” (Karnon, et al., 2004).

Pengambilan sampel pada pemeriksaan sitologi berbasis cairan ini dilakukan dengan cara yang sama seperti pada pemeriksaan sitologi biasa. Sampel sel tersebut terlebih dahulu dicelupkan ke dalam cairan khusus dan diproses dengan alat otomatis sebelum sampel tersebut diusapkan pada kaca benda. Sample yang telah dicelupkan tersebut kemudian diwarnai dan dilihat di bawah mikroskop oleh seorang Dokter ahli Patologi Anatomi.

Pemeriksaan sitologi berbasis cairan diperkenalkan pada awal tahun 1990 oleh perusahaan Amerika yang bernama CYTYC, produk tes saring kanker serviks yang

dinamakan *ThinPrep*. Dikutip dari (Arbyn et al., 2008), saat ini sudah tersedia 2 sistem berbasis cairan secara komersial yaitu *ThinPrep* dan *SurePath* yang disetujui oleh *U.S. Food and Drug Administration*. LBC hadir untuk memperbaiki kualitas *slide* pada *Pap smear* konvensional sehingga memudahkan Dokter Patologi Anatomi untuk membaca dan menginterpretasi spesimen. Gambar 2.1 menunjukkan hasil dari Citra Mikroskopis antara *Pap smear* konvensional dan LBC.



Gambar 2.1. Citra *Pap smear* konvensional dan berbasis cairan

2.3 Konsep Dasar Citra Digital

Citra diskrit atau citra digital adalah gambar pada dwimatra atau dua dimensi yang merupakan informasi berbentuk visual dan dihasilkan melalui proses digitalisasi terhadap citra analog dua dimensi yang kontinu. Data digital direpresentasikan dalam komputer berbentuk kode seperti biner dan desimal.

Representasi citra digital terdiri dari 2 bagian yaitu:

1. Bitmap

Gambar bitmap direpresentasikan dalam bentuk matrik, atau dipetakan dengan menggunakan bilangan biner atau sistem bilangan lain, memiliki kelebihan untuk memanipulasi warna namun untuk merubah objek lebih sulit.

2. Grafik

Gambar grafik data tersimpan dalam bentuk vektor posisi, dimana yang tersimpan dalam bentuk vektor posisinya dengan bentuk sebuah fungsi, lebih sulit dalam merubah warna tetapi lebih mudah membentuk objek dengan cara merubah nilai.

2.3.1 Model Citra Digital

Citra merupakan fungsi menerus (*continue*) dari intensitas cahaya pada bidang dimatra. Secara matematis, fungsi intensitas cahaya pada bidang dua dimensi disimbolkan dengan $f(x,y)$, dimana:

(x,y) = koordinat pada bidang dimensi,

$F(x,y)$ = intensitas cahaya (*brightness*) pada titik (x,y) .

Citra digital yang berukuran $N \times M$ lazimnya dinyatakan dengan matriks berukuran N baris dan M kolom, dan masing-masing elemen pada citra digital disebut *pixel* (*picture element*)

$$f(x,y) = \begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0,M) \\ f(1,0) & f(1,1) & \dots & f(1,M) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{pmatrix} \dots\dots\dots(2.1)$$

2.3.2 Elemen-Elemen Dasar Citra Digital

Elemen-elemen dasar dari citra digital, yaitu:

1. Kecerahan (*brightness*)

Kecerahan merupakan intensitas cahaya rata-rata dari suatu area citra yang melingkupinya.

2. Kontras (*contrast*)

Sebaran terang (*lightness*) dan gelap (*darkness*) di dalam sebuah citra. Citra dengan kontras rendah komposisi citranya sebagian besar terang atau sebagian besar gelap. Citra dengan kontras yang baik, komposisi gelap dan terangnya tersebar merata.

3. Kontur (*contour*)

Kontur adalah keadaan yang ditimbulkan oleh perubahan intensitas pada piksel-*pixel* tetangga, sehingga dapat mendeteksi tepi objek di dalam citra.

4. Warna (*color*)

Warna merupakan persepsi yang dirasakan oleh sistem visual manusia terhadap panjang gelombang cahaya yang dipantulkan oleh objek. Warna-warna yang dapat ditangkap oleh mata manusia merupakan kombinasi cahaya dengan panjang berbeda.

Kombinasi yang memberikan rentang warna paling lebar adalah *red (R)*, *green (G)*, dan *blue (B)*.

5. Bentuk (*shape*)

Bentuk adalah properti intrinsik dari objek tiga dimensi, dengan pengertian bahwa bentuk merupakan properti intrinsik utama untuk visual manusia. Secara umum citra yang dibentuk oleh manusia merupakan dua dimensi (2D), sedangkan objek yang dilihat adalah tiga dimensi (3D).

6. Tekstur (*texture*)

Tekstur adalah distribusi spesial dari derajat keabuan di dalam sekumpulan *pixel* yang bertetangga.

2.3.3 Operasi Dasar Pengolahan Citra

Operasi-operasi yang dilakukan di dalam pengolahan citra banyak ragamnya. Operasi pengolahan citra secara umum dapat diklasifikasikan sebagai berikut:

1. Perbaikan kualitas citra (*image enhancement*)

Jenis operasi ini bertujuan untuk memperbaiki kualitas citra dengan cara memanipulasi parameter-parameter citra. Ciri-ciri khusus yang terdapat di dalam citra lebih ditonjolkan. Contoh-contoh operasi perbaikan citra:

- a. Perbaikan gelap/terang,
- b. perbaikan tepian objek (*edge enhancement*),
- c. penajaman (*sharpening*),
- d. pemberian warna semu (*pseudocoloring*), dan
- e. penapisan derau (*noise filtering*).

2. Pemugaran citra (*image restoration*)

Operasi ini bertujuan menghilangkan atau meminimumkan cacat pada citra. Tujuan pemugaran citra hampir sama dengan operasi perbaikan citra, bedanya pada pemugaran citra penyebab degradasi gambar diketahui. Contoh-contoh operasi pemugaran citra adalah penghilangan kesamaran (*deblurring*) dan penghilangan derau (*noise*).

3. Pemampatan citra (*image compression*)

Jenis operasi ini dilakukan agar citra dapat direpresentasikan dalam bentuk yang lebih kompak sehingga memerlukan memori yang lebih sedikit. Hal penting yang harus diperhatikan dalam pemampatan adalah citra yang telah dimampatkan harus tetap mempunyai kualitas gambar yang bagus. Contoh metode pemampatan citra adalah metode *JPEG*.

4. Segmentasi citra (*image segmentation*)

Jenis operasi ini bertujuan untuk memecah suatu citra ke dalam beberapa segmen dengan suatu kriteria tertentu. Jenis operasi ini berkaitan erat dengan pengenalan pola.

5. Analisis citra (*image analysis*)

Jenis operasi ini bertujuan menghitung besaran kuantitatif dari citra untuk menghasilkan deskripsinya. Teknik pengorakan citra mengekstraksi ciri-ciri tertentu yang membantu dalam identifikasi objek. Proses segmentasi kadang diperlukan untuk melokalisasi objek yang diinginkan dari sekelilingnya.

Contoh-contoh operasi pengorakan citra:

- a. Pendeteksian tepi objek (*edge detection*),
- b. Ekstraksi batas (*boundary*), dan
- c. Representasi daerah (*region*).

6. Rekonstruksi citra (*image recontruction*)

Jenis operasi ini bertujuan membentuk ulang objek dari beberapa citra hasil proyeksi. Pengolahan citra mempunyai aplikasi yang sangat luas dalam berbagai bidang kehidupan. Berikut beberapa aplikasi Pengolahan citra dalam berbagai bidang:

1. Bidang kedokteran

- (a) Pengolahan citra sinar X untuk mammografi (deteksi kanker payudara).
- (b) *NMR (Nuclear Magnetic Resonance)*.
- (c) Mendeteksi kelainan tubuh dari foto sinar X.
- (d) Rekonstruksi foto janin hasil *USG*.
- (e) Pendeteksian jenis penyakit melalui gambar mikroskopik. (contoh: *Pap smear*)

2. Bidang biologi

Pengenalan jenis kromosom melalui gambar mikroskopik.

2.3.4 Pewarnaan dalam Citra

Warna secara utuh bergantung pada sifat pantulan (*reflectance*) suatu objek. Warna yang dilihat merupakan yang dipantulkan sedangkan yang lainnya diserap. Sehingga sumber sinar perlu diperhitungkan begitu pula sifat alami sistem visual manusia ketika menangkap suatu warna.

Ada beberapa jenis citra pewarnaan, yaitu:

1. Citra monokrom (*monochrome image*)

Citra monokrom merupakan citra hitam-putih fungsi $f(x,y)$ sebagai fungsi tingkat keabuan, fungsi dua dimensi (2D) dengan x menyatakan variabel baris dan y variabel kolom.

2. Citra multispektural

Citra multispektural merupakan citra berwarna biasanya dinyatakan dalam tiga komponen *RGB* (*Red-Green-Blue*). Intesitas suatu titik pada citra berwarna merupakan kombinasi dari intensitas:

$$f(x, y) = \{f_{merah}(x, y), f_{hijau}(x, y), f_{biru}(x, y)\} \quad (1)$$

3. Derajat keabuan (*grey-level*)

Grey-level merupakan intensitas fungsi citra hitam-putih pada titik (x,y) . *Grey-level* bergerak dari hitam ke putih dan skala keabuan memiliki rentang:

$$I_{min} < f < I_{max} \text{ atau } [0, L] \quad (2)$$

Dimana, intensitas 0 menyatakan hitam dan L menyatakan putih.

Contoh: citra hitam-putih dengan 256 level, artinya mempunyai skala abu-abu dari 0 sampai 255 atau $[0,255]$, dalam hal ini nilai 0 menyatakan hitam dan 255 menyatakan putih, nilai antara 0 sampai 255 menyatakan warna keabuan yang terletak antara hitam dan putih.

2.4 Operasi Citra Digital

2.4.1 CLAHE (*Contrast Limited Adaptive Histogram Equalization*)

CLAHE awalnya dikembangkan untuk pencitraan medis dan telah terbukti berhasil untuk peningkatan dari gambar dengan kontras rendah seperti film portal. CLAHE merupakan versi perbaikan dari AHE atau *Adaptive Histogram Equalization*, keduanya mengatasi keterbatasan standar pemerataan *histogram*. Menurut Zuiderveld (1994), masalah *noise* pada AHE dapat dikurangi dengan membatasi peningkatan kontras khususnya di daerah yang *homogen*. Daerah ini dapat dicirikan sebagai puncak yang tinggi pada *histogram* terkait dengan daerah kontekstual oleh karena banyak piksel tergabung di dalam kisaran abu-abu yang sama. Lereng yang terkait dengan skema penempatan *gray-level* menjadi terbatas dengan CLAHE. Hal ini dapat dicapai dengan hanya memungkinkan jumlah maksimum dari *pixel* di setiap kelompok data yang terkait dengan *histogram* lokal.

2.4.2 Segmentasi Citra

Segmentasi adalah mendapatkan representasi kompak dari data gambar yang diinginkan dengan cara menekankan pada sifat-sifat yang dari gambar yang dimaksud (Forsyth dan Ponce, 2003). Menurut (Castleman, 1996) segmentasi citra sendiri merupakan suatu proses memecah suatu citra digital menjadi banyak segmen/ bagian daerah yang tidak saling bertabrakan (*non-overlapping*). Hasil segmentasi tersebut merupakan kelompok *pixel* yang bertetangga atau berhubungan. Segmentasi citra dapat dilakukan melalui beberapa pendekatan, menurut (Castleman, 1996) terdapat 3 macam pendekatan, antara lain:

- a. Pendekatan batas (*boundary approach*), pendekatan ini dilakukan untuk mendapatkan batas yang ada antar daerah.
- b. Pendekatan tepi (*edge approach*), pendekatan tepi dilakukan untuk mengidentifikasi *pixel* tepi dan menghubungkan *pixel* tersebut menjadi suatu batas yang diinginkan
- c. Pendekatan daerah (*region approach*), pendekatan daerah bertujuan untuk membagi citra dalam daerah-daerah sehingga didapatkan suatu daerah sesuai kriteria yang diinginkan.

Proses segmentasi digunakan dalam berbagai hal. Metode yang digunakan sangat bervariasi meskipun semuanya memiliki tujuan sama yaitu mendapatkan representasi sederhana yang berguna dari suatu citra. Pemilihan metode bergantung pada pendekatan yang akan digunakan dan fitur yang ingin diperoleh dari citra karena cukup sulit untuk menentukan metode yang komprehensif. Secara umum pendekatan segmentasi citra yang sering digunakan adalah melalui pendekatan intensitas, pendekatan warna dan pendekatan bentuk (Rujikietgumjorn, 2008).

2.4.3 Thresholding

Castleman (1996) mendefinisikan *thresholding* sebagai “*Thresholding is particularly useful region-approach technique for scenes containing solid objects resting upon a contrasting background*”. Sedangkan (Pal dan Pal, 1993) menyebutkan bahwa *thresholding* adalah teknik sederhana dan umum yang digunakan untuk segmentasi citra.

Threshold dengan Metode Otsu

Merupakan salah satu metode untuk mensegmentasi citra digital. *Threshold* dengan metode Otsu diperkenalkan oleh Nobuyuki Otsu dalam tulisannya “*A Threshold Selection Method from Gray-Level Histogram*”. Tujuan dari metode Otsu adalah membagi *histogram* citra *gray-level* kedalam dua daerah yang berbeda secara

otomatis tanpa membutuhkan bantuan *user* untuk memasukkan nilai ambang. Pendekatan yang dilakukan oleh metode Otsu adalah dengan melakukan analisis diskriminan yaitu menentukan suatu variabel yang dapat membedakan antara dua atau lebih kelompok yang muncul secara alami. Analisis Diskriminan akan memaksimumkan *variable* tersebut agar dapat membagi objek latar depan (*foreground*) dan latar belakang (*background*).

Penggunaan *global thresholding* menggunakan metode Otsu dalam penelitian ini adalah dengan menggunakan fungsi dari *Tools* Matlab yaitu *graythresh* yang merupakan fungsi dari *Image Processing Toolbox* pada Matlab.

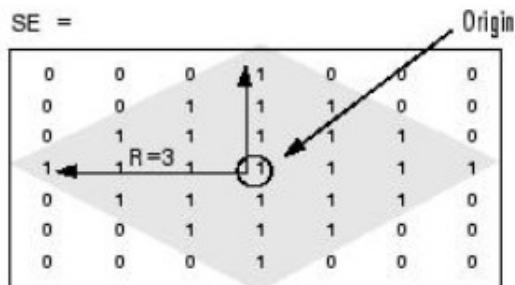
2.4.4 Morphology

Matematika Morfologi adalah alat untuk mengekstrak komponen-komponen citra yang berguna dalam representasi dan deskripsi dari suatu wilayah dalam citra (Sutoyo, et al., 2009).

Structuring element merupakan salah satu unsur utama pada operasi morfologi, *structuring element* digunakan sebagai cetakan untuk citra masukkan. *Structuring element* merupakan suatu matriks yang hanya berisi dua nilai yaitu 1 atau 0 yang disusun sedemikian rupa sehingga membentuk suatu formasi dan ukuran, *pixel* dengan nilai 1 pada *structuring element* biasa disebut dengan *neighborhood*. Jenis-jenis dari *structuring element* (SE) antara lain adalah:

a. *Diamond*

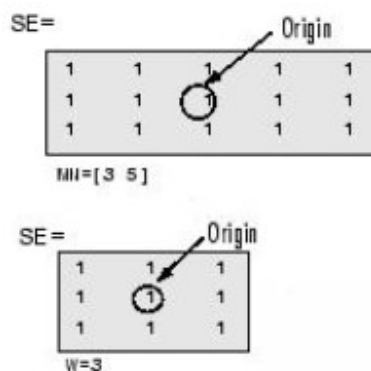
SE yang berbentuk *diamond* dengan R adalah jarak dari origin ke ujung/tepi dari SE diamond. Gambar 2.2 adalah merupakan gambar SE dengan jenis *diamond*.



Gambar 2.2. SE diamond

b. *Rectangle/Square*

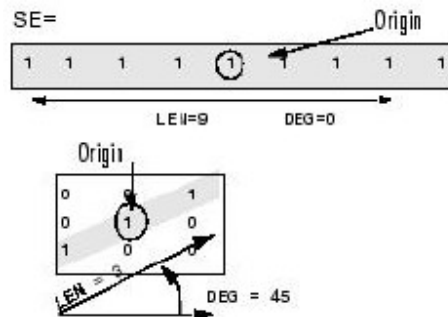
SE yang berbentuk persegi atau kotak. MN merepresentasikan ukuran dari SE. MN terdiri dari dua buah elemen *vector nonnegative integers*. M adalah ukuran untuk baris dan N adalah ukuran untuk kolom. Gambar 2.3 adalah merupakan Gambar SE dengan jenis *rectangle/ square*.



Gambar 2.3. SE rectangle/ square

c. *Line*

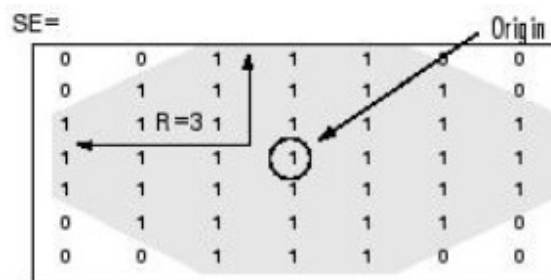
Sebuah SE yang *flat* dan linier, dimana LEN merepresentasikan panjang dan DEG merepresentasikan sudut (dalam derajat) *line* yang diukur dari arah sumbu horizontal. LEN dapat diartikan jarak dari titik ujung SE ke ujung SE lainnya. Gambar 2.4 adalah merupakan gambar SE dengan jenis *line*.



Gambar 2.4. *SE line*

d. *Octagon*

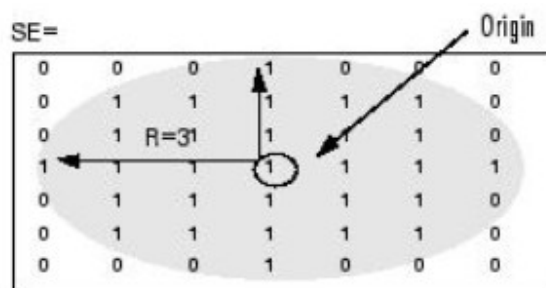
SE yang berbentuk segi-8, dimana R adalah jarak dari *origin* SE dengan tepian dari segi-8, diukur dari sumbu x dan sumbu y . Gambar 2.5 adalah merupakan gambar SE dengan jenis *octagon*.



Gambar 2.5. *SE octagon*

e. *Disk*

SE berbentuk lingkaran/ *disk*, dimana R adalah jari-jari yang diukur dari *origin* ke tepi dari lingkaran. Gambar 2.6 adalah merupakan gambar SE dengan jenis *disk*.



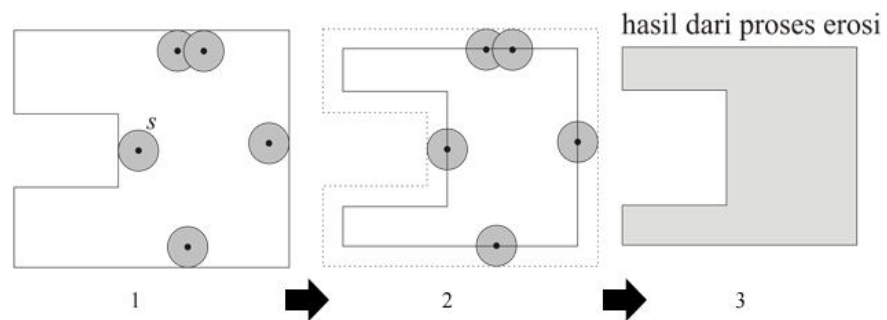
Gambar 2.6. *SE disk*

Adapun Menurut Castleman (1996) terdapat beberapa contoh dasar morfologi yakni:

a. Erosion

Erosion adalah proses mengurangi *pixel* pada batas dari objek. Jumlah *pixel* yang dikurangkan tergantung dari besar dan bentuk dari SE yang digunakan untuk mengolah citra. *Erosion* berguna untuk menghilangkan/ mengurangi objek-objek yang terlalu kecil dari *image* yang telah tersegmentasi. Gambar 2.7 adalah merupakan gambar *Erosion*. *Erosion* didefinisikan sebagai:

$$E = B \otimes S = \{x,y \mid S_{xy} \subseteq B\} \quad (2.4)$$

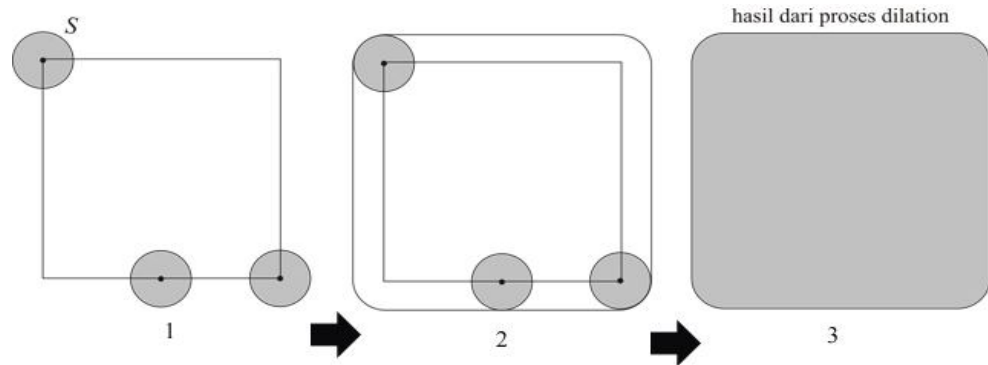


Gambar 2.7. Proses *erosion*

b. Dilation

Dilation adalah proses menambahkan *pixel* pada batas dari objek di sebuah *image*. Jumlah *pixel* yang ditambahkan tergantung dari besar dan bentuk dari SE yang digunakan untuk mengolah citra. *Dilation* berguna untuk mengisi lubang-lubang yang ada pada objek yang telah tersegmentasi. Gambar 2.8 adalah merupakan gambar *Dilation*. *Dilation* didefinisikan sebagai:

$$D = B \oplus S = \{x,y \mid S_{xy} \cap B \neq \emptyset\} \quad (2.5)$$



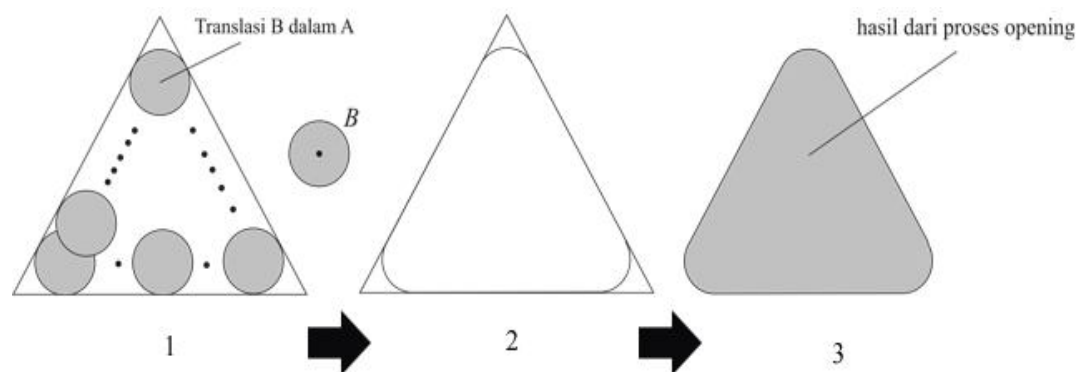
Gambar 2.8. Proses *dilation*

c. *Opening*

Opening adalah proses *erosion* yang diikuti dengan *dilation*. Proses ini menghilangkan objek-objek yang kecil dan tipis dan umumnya menghaluskan *boundary* objek yang lebih besar tanpa merubah areanya secara signifikan. Gambar 2.9 adalah merupakan gambar *Opening*. *Opening* didefinisikan menjadi:

$$A \circ B = (A \otimes B) \oplus B \quad (2.6)$$

Persamaan (2.6). menunjukkan rumusan dari operasi *opening*, terlihat bahwa set *A* mengalami operasi erosi terlebih dahulu, kemudian dilanjutkan dengan operasi dilasi.



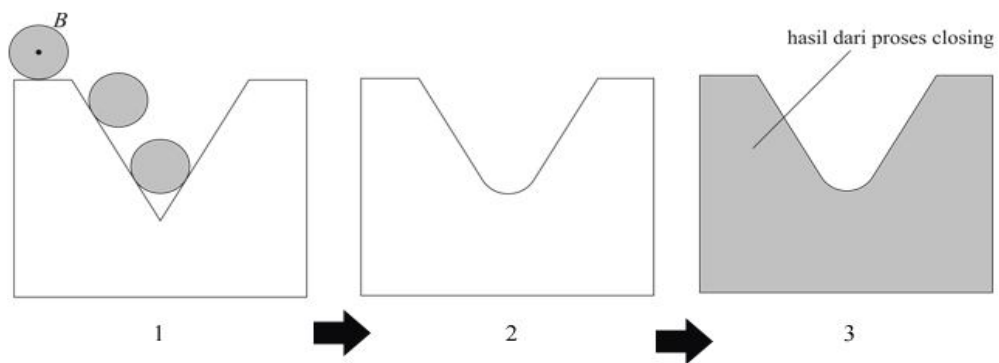
Gambar 2.9. *Opening*

d. Closing

Closing adalah proses *dilation* yang diikuti dengan proses *erosion*. Proses ini mengisi lubang-lubang kecil dan tipis dari objek dan umumnya menghaluskan *boundary* objek tanpa merubah areanya secara signifikan. Gambar 2.10 adalah merupakan gambar *Closing*. *Closing* didefinisikan menjadi:

$$A \bullet B = (A \oplus B) \otimes B \quad (2.7)$$

Persamaan (2.7) menunjukkan operasi *closing*, terlihat bahwa set *A* terlebih dahulu di dilasi oleh *B*, kemudian dilanjutkan dengan melakukan erosi ke proses tersebut.



Gambar 2.10. *Closing*

2.5 Fuzzy C -Means Clustering

Pengertian *Fuzzy C-Means* (FCM) seperti yang dikutip dari (Kusumadewi, et al., 2006) adalah suatu teknik pengclusteran data yang mana keberadaan tiap-tiap data dalam suatu *cluster* ditentukan oleh nilai keanggotaan. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981.

Algoritma FCM diberikan sebagai berikut:

1. Tentukan:

- Matriks X berukuran $n \times m$, dengan n = jumlah data yang akan *dicluster*; dan m = jumlah variabel (kriteria).
- Jumlah *cluster* yang akan dibentuk = $C (\geq 2)$.
- Pangkat (pembobot) = $w (> 1)$.
- Maksimum iterasi
- Kriteria penghentian = ζ (nilai positif yang sangat kecil)
- Iterasi awal, $t = 1$, dan $\Delta=1$;

2. Bentuk matriks partisi awal, U_0 , sebagai berikut:

$$U = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \dots & \mu_{1n}(x_n) \\ \mu_{21}(x_1) & \mu_{22}(x_2) & \dots & \mu_{2n}(x_n) \\ \dots & \dots & \dots & \dots \\ \mu_{c1}(x_1) & \mu_{c2}(x_2) & \dots & \mu_{cn}(x_n) \end{bmatrix} \quad (2.8)$$

3. Hitung pusat *cluster* V , untuk setiap *cluster*:

$$V_{ij} = \frac{\sum_{k=1}^n (\mu_{ik})^w X_{kj}}{\sum_{k=1}^n (\mu_{ik})^w} \quad (2.9)$$

4. Perbaiki derajat keanggotaan setiap data pada setiap *cluster* (perbaiki matriks partisi), sebagai berikut:

$$\mu_{ik} = \left[\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{(w-1)}} \right]^{-1} \quad (2.10)$$

Dengan

$$d_{ik} = d(X_k - V_i) = \left[\sum_{j=1}^m (X_{kj} - V_{ij})^2 \right]^{\frac{1}{2}} \quad (2.11)$$

5. Tentukan kriteria berhenti, yaitu perubahan matriks partisi pada iterasi sekarang dengan iterasi sebelumnya, sebagai berikut:

$$\Delta = \|U^t - U^{t-1}\| \quad (2.12)$$

Apabila $\Delta \leq \xi$, maka iterasi dihentikan, namun apabila $\Delta > \xi$, maka naikkan iterasi ($t = t+1$) dan kembali ke langkah-3. Pencarian nilai Δ dapat dilakukan dengan mengambil elemen terbesar dari nilai mutlak selisih antara $\mu_{ik}(t)$ dengan $\mu_{ik}(t-1)$.

Penggunaan FCM dalam penelitian ini adalah dengan menggunakan fungsi dari Tools Matlab yaitu *fcm* yang merupakan fungsi dari *Logic Toolbox* pada Matlab.

2.6 Evaluasi

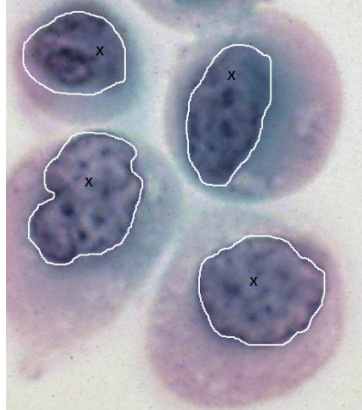
A. Single Decision Threshold

		Model keputusan		
		+	-	
Kenyataan	+	TP	FN	100%
	-	FP	TN	100%

Gambar 2.11. Diagram *single decision threshold*

a. TP (*True Positive*)

Hasil deteksi dikatakan *True Positive* apabila kenyataan dan sistem memiliki nilai yang positif atau bisa dikatakan jika kenyataan “sel *nuclei*”, maka sistem memutuskan “sel *nuclei*”. Gambar 2.12 adalah gambar yang menunjukkan *True Positive*.

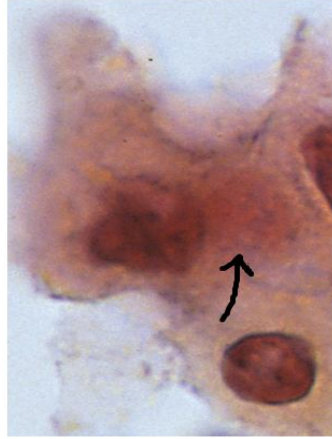


Gambar 2.12. *Nuclei* sel terdeteksi *true positive*

Gambar 2.12 menunjukkan area yang di kelilingi garis putih merupakan area *nuclei* sel, tanda (x) merupakan *centroids*. Gambar 2.12 *centroids* tepat berada dalam *nuclei*. *Centroids* yang berada dalam *nuclei* inilah yang disebut dengan *True Positive*.

b. TN (*True Negative*)

Hasil deteksi dikatakan *True Negative* apabila kenyataan dan sistem menghasilkan hasil yang negative atau bisa dikatakan jika kenyataan “bukan sel *nuclei*”, tetapi sistem memutuskan bukan “sel *nuclei*”. Gambar 2.13 adalah gambar yang menunjukkan *True Negative*.

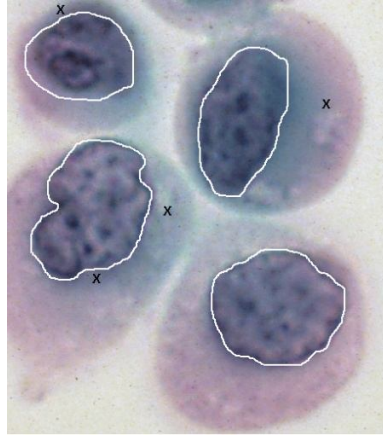


Gambar 2.13. *Nuclei* sel terdeteksi *true negative*

Gambar 2.13, tanda panah menunjukkan area yang memiliki intensitas rendah dan hampir menyerupai *nuclei*. Sistem berhasil tidak mendeteksi wilayah ini maka hal ini disebut dengan *True Negative*. Wilayah *cytoplasm*a bisa juga disebut dengan area *True Negative*.

c. FP (*False Positive*)

Hasil deteksi dikatakan *False Positive* apabila kenyataan negatif, tetapi sistem memutuskan positif atau bisa dikatakan kenyataan bukan “sel *nuclei*”, tetapi sistem memutuskan “sel *nuclei*”. Gambar 2.14 adalah gambar yang menunjukkan *False Positive*.

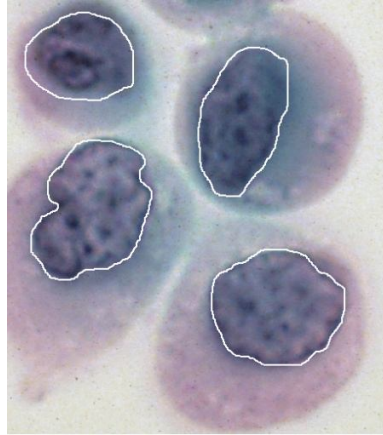


Gambar 2.14. *Nuclei sel terdeteksi false positive*

Gambar 2.14 menunjukkan area yang dikelilingi garis putih merupakan area *nuclei* sel. Gambar 2.14 menunjukkan *centroids* (tanda x) tidak berada dalam *nuclei*. *Centroids* mendeteksi *cytoplasma* sebagai *nuclei*. *Centroids* yang berada tidak di dalam *nuclei* inilah yang disebut dengan *False Positive*.

d. FN (*False Negative*)

Hasil deteksi dikatakan *False Negative* apabila kenyataan positif, tetapi sistem memutuskan negatif atau bisa dikatakan kenyataan “Sel *Nuclei*”, namun sistem memutuskan bukan “Sel *Nuclei*”. Gambar 2.15 adalah gambar yang menunjukkan *False Negative*.



Gambar 2.15. *Nuclei* sel terdeteksi *false negative*

Gambar 2.15 menunjukkan area yang dikelilingi garis putih merupakan area *nuclei* sel. Gambar 2.15 menunjukkan *nuclei* tidak terdeteksi, terbukti tidak adanya *centroids* yang berada di dalam area *nuclei*. *Centroids* yang tidak berada dalam *nuclei* inilah yang disebut dengan *False Negative*

B. *Sensitivity*

Merupakan ukuran yang merepresentasikan tingkat kemampuan model untuk berhasil menemukan *nuclei* sel. Semakin tinggi *sensitivity* maka *nuclei* yang tidak terdeteksi semakin sedikit. *Sensitivity* didapat menggunakan rumus (Bemmel,1997) berikut:

$$Sensitivity (\%) = \frac{TP}{TP+FN} \times 100\% \quad (2.13)$$

C. *Specificity*

Merupakan ukuran yang merepresentasikan tingkat kemampuan model untuk berhasil menghindari area-area yang bukan *nuclei* namun dianggap sebagai *nuclei*. Semakin tinggi *specificity* maka wilayah yang seharusnya bukan *nuclei* tetapi terdeteksi sebagai *nuclei* semakin sedikit. *Specificity* didapat menggunakan rumus (Bemmel,1997) berikut:

$$Specificity (\%) = \frac{TN}{TN+FP} \times 100\% \quad (2.14)$$

D. System Performance

Merupakan ukuran yang merepresentasikan tingkat kinerja model secara keseluruhan. Ukuran tingkat kinerja model digunakan rumus (Bemmel,1997) berikut:

$$System Performance (\%) = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2.15)$$

2.7 Penelitian Terdahulu

Bagian ini akan membahas penelitian - penelitian terdahulu yang terkait secara langsung ataupun tidak langsung dengan penelitian yang dilakukan dalam penelitian ini.

Penelitian terdahulu yang terkait secara langsung adalah: *Automated Detection of Cell Nuclei in Pap smear Images Using Morphological Reconstruction and Clustering* (Plissiti, et al., 2011). Pada Penelitian yang dilakukan (Plissiti, et al., 2011) membahas teknik segmentasi otomatis pada citra hasil *Pap smear*. Model yang diusulkan pada penelitian (Plissiti, et al., 2011) merupakan acuan utama dari penelitian ini. Penelitian (Plissiti, et al., 2011) terbagi menjadi 4 langkah utama, yakni: *Preprocessing, Detection of Candidate Cell Nuclei Centroids, Refinement of Candidate Cell Nuclei Centroids, Decision*.

Tahap *Preprocessing* yang dilakukan oleh (Plissiti, et al., 2011) bermaksud untuk mensegmentasi citra agar pencarian *nuclei* lebih terpusat dan komputasinya menjadi singkat. Langkah selanjutnya (Plissiti, et al., 2011) melakukan pemrosesan citra untuk mendeteksi *centroids* pada kandidat *nuclei* pada tahap *Detection of Candidate Cell Nuclei Centroids*. Proses pencarian *centroids* pada tahap ini dilakukan dengan cara menemukan intensitas rendah di tiap-tiap area pada citra. Tahap

selanjutnya, tahap *Refinement of Candidate Cell Nuclei Centroids* dilakukan untuk memperbaiki letak *centroids* yang terdeteksi dengan cara menghilangkan *centroids* yang berada pada tepi-tepi *nuclei*. Tahap yang terakhir yang dilakukan adalah *Decision*. Tahap ini (Plissiti, et al., 2011) melakukan perbaikan letak *nuclei* menggunakan kriteria jarak, yakni dengan cara menghilangkan *centroids* yang saling bedekatan. Setelah itu, (Plissiti, et al., 2011) melakukan *clustering* menggunakan *Fuzzy C-Means* dan *Support Vector Machine*. Pada tahap *clustering* inilah (Plissiti, et al., 2011) mendapatkan letak *centroids* yang diinginkan.

Pemrosesan citra yang dilakukan (Plissiti, et al., 2011) diperlukan untuk membedakan antara latar belakang citra dengan daerah yang dikehendaki (*region of interest*) pada tahap *Preprocessing* yang dilakukan pada penelitian ini. Proses yang dilakukan mencakup proses *contrast limited adaptive histogram equalization* (CLAHE) dengan tujuan untuk mendapatkan kontras yang lebih baik dan mempertajam bagian tepi dari obyek yang berada di dalam citra. Langkah selanjutnya dilanjutkan dengan menentukan *global threshold* untuk mendapatkan citra biner dengan metode yang diusulkan oleh Otsu, dan diakhiri dengan menghilangkan obyek-obyek yang terlalu kecil dan tidak signifikan. Proses *preprocessing* yang dilakukan oleh Plissiti merupakan acuan utama pada tahap *preprocessing* dalam penelitian ini. Proses deteksi tepi pada tahap *Refinement of Candidate Cell Nuclei Centroids* dan *Clustering* menggunakan *Fuzzy C-means* juga menjadi acuan utama dalam penelitian ini.

Penelitian ini melakukan modifikasi pada tahap *Detection of Candidate Cell Nuclei Centroids* dan *Refinement of Candidate Cell Nuclei Centroids* serta *clustering* yang dilakukan oleh (Plissiti, et al., 2011). Modifikasi ini dilakukan untuk mengusulkan model yang lebih mudah dan lebih singkat komputasinya, selain itu dengan komputasi yang lebih singkat tentu akan mempercepat waktu yang digunakan untuk mendeteksi posisi dan menghitung *nuclei* pada Citra *Pap smear* tersebut.

BAB III

METODOLOGI

3.1 Model yang Diusulkan

Penelitian yang dilakukan terdiri dari 5 tahap utama untuk mendeteksi sel *nuclei*, antara lain: 1) *preprocessing*; 2) deteksi *centroids* kandidat sel *nuclei*; 3) perbaikan posisi kandidat sel *nuclei*; 4) pengurangan jumlah *centroids* dengan kriteria jarak; dan 5) *clustering*. Berikut gambaran secara umum masing-masing tahap yang telah disebutkan sebelumnya:

3.1.1 *Preprocessing*

Pada tahap *preprocessing* langkah-langkah yang dilakukan antara lain:

1. Melakukan penajaman kontras dengan menggunakan CLAHE pada tiap-tiap *layer* warna.
2. Melakukan *global thresholding* untuk membuat citra biner dari citra *grayscale* hasil penajaman kontras dengan metode Otsu.
3. Melakukan morfologi dilasi untuk mendapatkan area yang lebih sesuai.
4. Menghilangkan area yang memiliki luasan kurang dari 500 *pixel*

Hasil yang diharapkan pada tahap ini adalah citra yang telah tersegmentasi untuk mengurangi area pencarian sehingga waktu eksekusinya menjadi lebih cepat.

3.1.2 Deteksi *Centroids* Kandidat Sel *Nuclei*

Tahap deteksi *centroids* kandidat sel *nuclei* terbagi menjadi dua sub tahap yaitu deteksi *region minima* dan deteksi tepi *nuclei*.

a. Deteksi *region minima*

Sub tahap ini langkah-langkah yang dilakukan antara lain:

1. Melakukan transformasi *h-minima* terhadap citra hasil segmentasi pada tahap *preprocessing* setiap *layer* warnanya.
2. Melakukan subtraksi terhadap citra hasil *h-minima* dengan nilai *h* yang sama dengan nilai pada transformasi *h-minima*.
3. Melakukan *morphological reconstruction* terhadap citra hasil subtraksi untuk mendapatkan area dengan intensitas yang tinggi.
4. Melakukan subtraksi terhadap citra hasil *morphological reconstruction* dengan citra yang tersegmentasi pada tahap *preprocessing* tiap *layer* warnanya.
5. Membuat citra biner dari hasil *union* dari masing-masing *layer* warna pada langkah sebelumnya untuk membuat area yang merupakan *regional maxima*.
6. Mencari *centroids* dari citra biner yang dihasilkan.

Pada sub tahap deteksi *region minima* hasil yang diharapkan adalah berupa citra dengan *centroids* yang berada pada area yang memiliki intensitas warna yang rendah.

b. Deteksi tepi *nuclei*

Sub tahap ini langkah-langkah yang dilakukan antara lain:

1. Membuat citra pertama dengan melakukan proses morfologi erosi untuk memperbesar luasan area dengan intensitas rendah.
2. Membuat citra kedua dengan melakukan *averaging filter* dari citra asli.
3. Melakukan morfologi *gradient* terhadap citra hasil proses erosi dengan citra hasil *averaging filter* untuk membuat batas-batas *nuclei*.
4. Memperjelas tepi *nuclei* dengan *contrast enhancement filter*.
5. Membuat citra biner dari proses normalisasi *min-max* terhadap citra hasil morfologi *gradient*

6. Menghilangkan centroids yang sudah terdeteksi pada sub tahap Deteksi *region minima* dengan menghilangkan centroids yang berada pada tepi *nuclei*.

Dari citra biner yang dihasilkan, tahap ini berguna untuk menghilangkan *centroids* yang dihasilkan pada sub tahap deteksi *region minima* yang berada pada perpotongan beberapa *nuclei*. Sub tahap ini juga digunakan untuk mereduksi jumlah *centroids* yang ditemukan.

3.1.3 Perbaikan Posisi Kandidat Sel *Nuclei*

Tahap perbaikan posisi kandidat sel *nuclei* langkah-langkah yang dilakukan antara lain:

1. Melakukan morfologi dilasi dari *pixel* tempat *centroids* berada sampai jumlah *centroids* awal dan hasil dilasi berbeda.
2. Memberi label dari area yang dibentuk dari proses morfologi dilasi.
3. Mencari *pixel* dengan intensitas terendah dari masing-masing area terlabelisasi dari langkah morfologi dilasi.
4. Menentukan *centroids* dari *pixel* dengan intensitas warna terendah dari area terlabelisasi dari langkah morfologi dilasi.

Hasil pada tahap ini diharapkan mampu mereduksi jumlah *centroids* yang ditemukan dari tahap sebelumnya.

3.1.4 Pengurangan Jumlah *Centroids* dengan Kriteria Jarak

Tahap pengurangan jumlah *centroids* dengan kriteria jarak langkah-langkah yang dilakukan adalah:

Masing-masing *centroids* yang dihasilkan pada tahap Perbaikan Posisi Kandidat Sel *Nuclei* dihitung jaraknya. Dua *centroids* yang memiliki jarak yang pendek akan dibandingkan intensitasnya. *Centroids* dengan intensitas rendah akan bertahan dan *centroids* dengan intensitas tinggi akan hilang.

Hasil dari tahap ini diharapkan mampu mereduksi jumlah *centroids* yang ditemukan secara signifikan.

3.1.5 *Clustering*

Tahap yang terakhir adalah tahap *clustering*. Langkah-langkah yang dilakukan antara lain:

1. Menentukan jarak terdekat dari masing-masing *centroids* sebagai salah satu kriteria pada tahap *clustering*.
2. Menentukan intensitas warna pada tiap *centroids* sebagai salah satu kriteria pada tahap *clustering*.
3. Melakukan normalisasi *min-max* dari proses 1 dan 2.
4. Melakukan proses *preprocessing* pada kriteria jarak terdekat.
5. Melakukan *clustering* dengan metode *Fuzzy C-Means* menggunakan dua kriteria yang disebutkan pada langkah sebelumnya dengan jumlah *cluster* 2 yaitu *cluster nuclei* dan *cluster* bukan *nuclei*.

Hasil yang diharapkan adalah berupa citra dengan *centroids* yang berada pada area yang memiliki intensitas warna yang rendah dan tepat berada pada *nuclei*. *Centroids* yang berada pada area dengan intensitas rendah adalah *centroids* yang merupakan himpunan dari *cluster nuclei*, sedangkan himpunan *cluster* yang bukan *nuclei* dihilangkan.

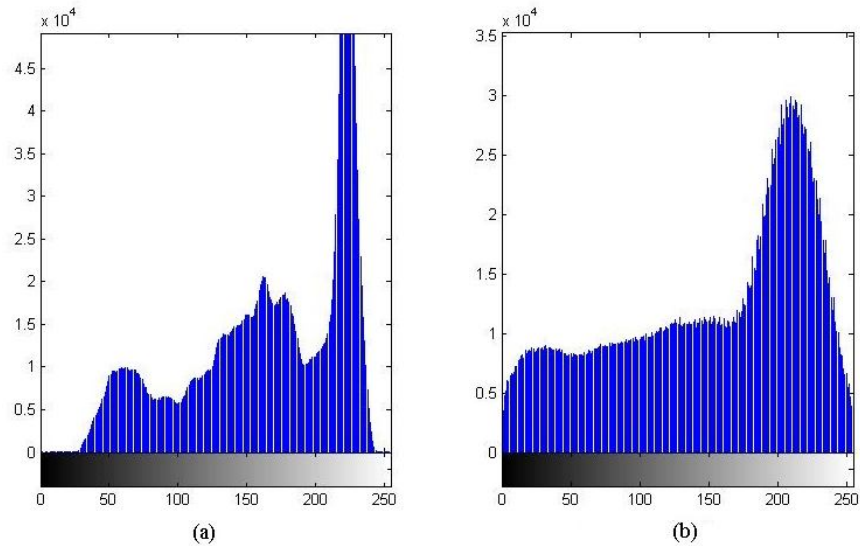
BAB IV

PREPROCESSING

4.1 Langkah yang Dilakukan

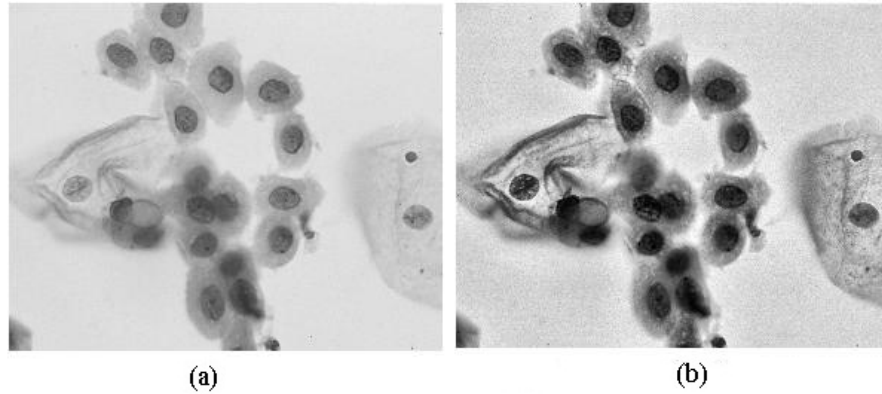
Interpretasi visual citra *Pap smear* merupakan hal yang sangat rumit. Hal ini disebabkan karena citra *Pap smear* memberikan hasil sel yang beragam mulai dari sel yang tidak merata, bekerumun, sampai sel yang saling tumpang tindih. Masalah pewarnaan, terdapat darah, lendir, bakteri dan pengeringan udara menyebabkan proses pengenalan sel *nuclei* menjadi sulit serta komputasinya menjadi panjang. Tahap *preprocessing* diperlukan untuk memperoleh area pencarian *nuclei* yang lebih fokus sehingga proses komputasinya menjadi lebih singkat. Tahap *preprocessing* perlu dilakukan untuk mengurangi area yang diinginkan pada gambar. Tahap ini melakukan proses pemisahan latar belakang citra dengan daerah yang diinginkan (*Region of Interest*). Penelitian yang dilakukan mengadopsi model *preprocessing* yang ditawarkan oleh Plissiti (Plissiti, et al., 2011).

Langkah pertama yang dilakukan adalah melakukan penajaman tepi dan peningkatan kontras menggunakan *Contrast Limited Adaptive Histogram Equalization* (CLAHE). Citra awal akan dipisahkan menjadi 3 *layer* warna, yaitu: merah, hijau dan biru. Ketiga *layer* warna tersebut dilakukan penajaman tepi dan peningkatan kontras menggunakan CLAHE. Gambar 4.1 adalah Gambar perbandingan *histogram* citra tanpa CLAHE dan *histogram* citra dengan CLAHE.



Gambar 4.1 (a) *Histogram* citra *Pap smear* pada *layer* warna merah, (b) *histogram* citra hasil CLAHE pada *layer* warna merah

Gambar 4.1(a) menjelaskan bahwa *histogram* pada *layer* warna merah citra *Pap smear* memiliki intensitas kontras yang rendah dan cenderung berkumpul di tengah. CLAHE akan meningkatkan kontras tiap area-area kecil (*tiles*) pada citra dan akan membatasi level kontrasnya sehingga derajat keabuannya menjadi lebih merata (Gambar. 4.1(b)). Gambar 4.2 menunjukkan bahwa citra yang telah mengalami peningkatan kontras (Gambar 4.2(b)) menjadi lebih jelas daripada citra aslinya (Gambar 4.2(a)) .

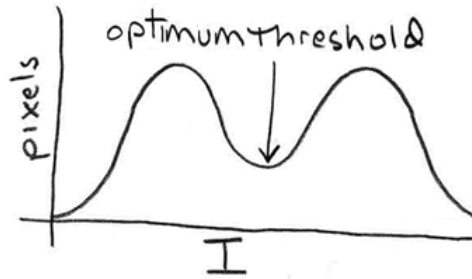


Gambar 4.2 (a) Citra *Pap smear* pada *layer* warna merah, (b) Citra hasil CLAHE pada *layer* warna merah

Langkah selanjutnya, dari citra *grayscale* pada tiap *layer* warna yang dihasilkan dari proses sebelumnya dilakukan transformasi biner melalui *global thresholding* menggunakan metode Otsu. Metode Otsu dipilih karena secara khusus, metode ini dianggap sebagai teknik yang terbaik dan paling umum digunakan untuk *global thresholding* (Dong, et al., 2008). Proses *thresholding* akan memproses citra yang memiliki nilai dengan intensitas kurang dari nilai ambang (T) akan di set menjadi 0 dan yang lebih dari atau sama dengan nilai ambang (T) akan di set menjadi 1. Operasi nilai ambang (*thresholding*) dilakukan dengan persamaan 4.1 berikut:

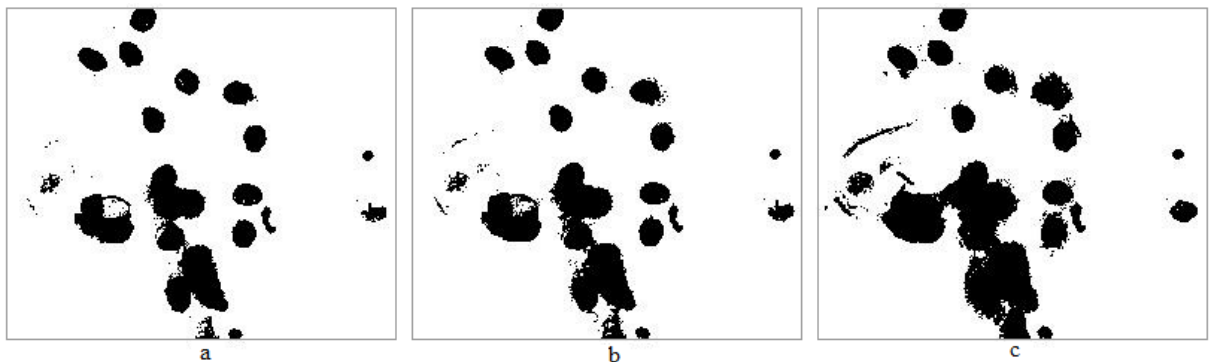
$$f_0(x, y) = \begin{cases} 0, & f_i(x, y) < T \\ 1, & f_i(x, y) \geq T \end{cases} \quad (4.1)$$

Metode Otsu akan memberikan nilai *threshold/* ambang (T) yang optimal secara otomatis. Nilai ambang ini sangat penting karena dari nilai ambang ini akan ditentukan batasan-batasan untuk membuat segmentasi pada citra. Nilai optimal yang dimaksud diilustrasikan pada Gambar 4.3.



Gambar 4.3 Ilustrasi *threshold* optimal

Nilai optimal tersebut berada diantara 2 derajat keabuan yang *homogen*. Bisa dikatakan bahwa dua derajat keabuan yang *homogen* itu adalah latar belakang dan daerah yang akan diekstrak informasinya. Gambar 4.4 adalah citra hasil pada tahap transformasi biner dengan *global thresholding* menggunakan metode Otsu.



Gambar 4.4 (a) Citra biner hasil *threshold* Otsu pada *layer* warna merah, (b) Gambar citra biner hasil *threshold* Otsu pada *layer* warna hijau, dan (c) Gambar citra biner hasil *threshold* Otsu pada *layer* warna biru

Gambar 4.4 merupakan hasil transformasi biner dengan *global thresholding* menggunakan metode Otsu pada 3 *layer* warna hasil proses CLAHE sebelumnya. Langkah selanjutnya, menggabungkan 3 citra biner tersebut menjadi sebuah citra biner dengan persamaan 4.2 berikut:

$$BW = \neg BW_1 \cup \neg BW_2 \cup \neg BW_3 \quad (4.2)$$

dengan BW_1 , BW_2 , dan BW_3 merupakan citra biner yang dihasilkan dari metode Otsu pada tiap *layer* warna (merah (Gambar 4.4(a)), hijau (Gambar 4.4(b)), dan biru (Gambar 4.4(c))) yang lebih dulu dilakukan proses $inverse(\neg)$ pada ketiga citra tersebut. Proses $inverse$ citra tersebut dilakukan sesuai dengan persamaan 4.3 berikut:

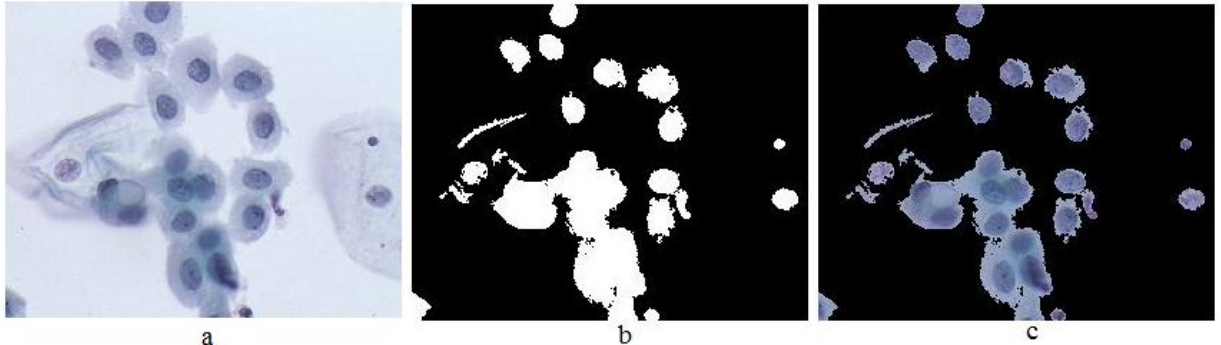
$$F_0(x, y) = f_{maximum} - f_i(x, y) \quad (4.3)$$

citra yang dihasilkan dari proses sebelumnya merupakan data dengan 1 bit maka nilai $f_{maximum} = 1$ dan f_i adalah intensitas warna pada tiap *pixel*-nya.

Morfologi dilasi dilakukan pada citra BW (citra hasil tahap sebelumnya/ citra hasil persamaan 4.2) untuk memperoleh daerah yang lebih tepat, dengan X (adalah 3×3 *flat structuring element* dengan jumlah tetangga 4) seperti persamaan 4.4 berikut:

$$BW = BW \oplus X \quad \dots\dots\dots (4.4)$$

Hasil dari proses ini masih terdapat beberapa area yang tidak dibutuhkan, karena itu perlu dihilangkan area yang memiliki luasan kurang dari 500 *pixel*. Citra biner yang dihasilkan pada proses inilah yang nantinya digunakan sebagai daerah yang akan dilakukan operasi-operasi selanjutnya. Gambar 4.5 adalah citra hasil akhir pada tahap *preprocessing*. Langkah selanjutnya, citra hasil akhir tahap *preprocessing* yang memiliki nilai 1 (Gambar 4.5(b)) akan diganti dengan nilai pada citra aslinya pada tiap *layer* warna. Ketiga *layer* warna tersebut apabila digabungkan akan menjadi seperti Gambar 4.5(c).

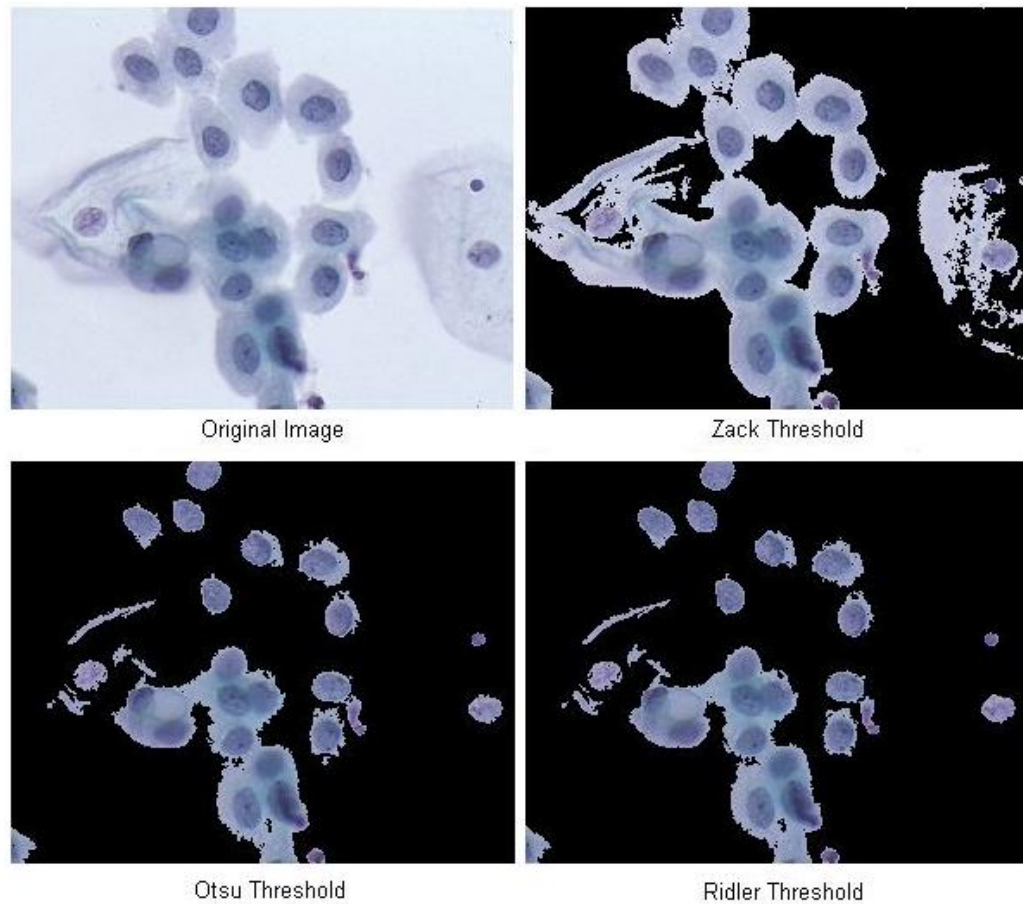


Gambar 4.5. (a) Citra *Pap smear* original, (b) Citra biner hasil akhir pada tahap *preprocessing*, dan (c) Hasil segmentasi citra

4.2 Pembahasan Penggunaan Metode untuk *Thresholding*

Proses *thresholding* merupakan tahap yang sangat penting dan dilakukan sebanyak 2 kali pada penelitian ini, yaitu pada tahap *preprocessing*, dan sub tahap deteksi tepi *nuclei* pada tahap deteksi *centroids* kandidat sel *nuclei*. Penelitian ini menggunakan metode Otsu sebagai metode *thresholding*. Metode Otsu dipilih karena secara khusus, metode ini dianggap sebagai teknik yang terbaik dan paling umum digunakan untuk *global thresholding* (Dong, et al., 2008). Sub bab ini akan membahas sedikit tentang perbandingan metode *thresholding*.

Penelitian yang dilakukan membandingkan metode *thresholding* yang diusulkan oleh (Zack et al., 1977), (Otsu, 1979) dan (Ridler dan Calvard, 1978). *Thresholding* dengan metode Zack dan Ridler tidak termasuk komponen yang terdapat pada Matlab, maka untuk metode Zack menggunakan *plug in* yang dibuat oleh (Panneton, 2010) dan metode Ridler menggunakan *plug in* yang dibuat oleh (Ramachandram, 2003). Penggunaan metode Otsu pada penelitian ini adalah merupakan komponen (*function*) yang terdapat pada *software* Matlab yaitu *graythresh*.



Gambar 4.6 Citra hasil perbandingan metode

Gambar 4.6 menunjukkan hasil dari masing-masing metode yang diusulkan. Hasil segmentasi pada citra dengan metode Zack masih banyak mengandung area yang tidak dibutuhkan, terlihat dari banyaknya *cytoplasm* yang masih terdapat pada citra hasil segmentasi. Metode Zack kurang baik jika digunakan pada *dataset* yang digunakan pada penelitian ini. Metode Ridler menghasilkan citra tersegmentasi yang hampir sama dengan metode Otsu. Penelitian dilanjutkan dengan melakukan pengujian terhadap waktu eksekusi metode Ridler dan metode Otsu karena hasil yang hampir sama. Tabel 4.1 menunjukkan hasil pengujian waktu eksekusi yang dibutuhkan oleh kedua metode tersebut. Pengujian dilakukan dengan Komputer i5 2.30 GHz dengan RAM 4 GB.

Tabel 4.1. Tabel pengujian waktu eksekusi metode Otsu dan Ridler

<i>file</i>	Otsu (detik)	Ridler(detik)
1	3,39	3,35
2	4,18	3,37
3	3,86	3,25
4	4,17	3,69
5	3,51	3,19
6	4,50	4,27
7	3,46	3,16
8	1,63	1,53
9	2,33	2,11
10	1,98	1,86
11	2,42	2,29
12	2,16	2,05
13	2,18	2,06
Rata2	3,06	2,78

Tabel 4.1 menunjukkan bukti bahwa metode yang diusulkan oleh Ridler memiliki waktu rata-rata eksekusi yang lebih cepat dari pada metode yang diusulkan oleh Otsu. Metode Ridler memang menghasilkan waktu eksekusi yang lebih cepat, namun demikian perbedaan waktu antara metode Ridler dan metode Otsu tidak terlalu signifikan. *Threshold* menggunakan metode Ridler bukan merupakan termasuk komponen yang disediakan oleh *software* Matlab R2008a dan belum teruji maka metode yang dipilih adalah metode dari Otsu.

Waktu eksekusi metode Ridler yang lebih cepat daripada metode yang diusulkan oleh Otsu merupakan sebuah temuan yang dapat diimplementasikan apabila telah terdapat penelitian yang lebih lanjut mengenai perbandingan kedua metode ini. Penelitian ini mengusulkan perlu adanya penelitian lebih lanjut untuk memodelkan metode yang lebih baik dengan komputasi yang lebih singkat menggunakan metode *thresholding* yang diusulkan oleh Ridler.

BAB V

DETEKSI CENTROIDS KANDIDAT SEL NUCLEI

5.1 Langkah yang Dilakukan

Tahap deteksi *centroids* kandidat sel *nuclei* ini dibagi menjadi dua sub tahap yaitu deteksi *region minima* dan deteksi tepi *nuclei*.

5.1.1 Deteksi *Region Minima*

Citra yang dihasilkan dari tahap *preprocessing* masih mengandung sel yang tumpang tindih dan terisolasi. Sel yang tumpang tindih dan terisolasi ini akan dipisahkan pada tahap kedua ini. Secara umum sel *nuclei* tampak lebih gelap daripada *cytoplasma* yang mengelilinginya, untuk itu diperlukan pencarian terhadap intensitas rendah/ *valley* pada tiap *layer* warna.

Beberapa langkah yang dilakukan pada tahap ini masih merupakan adopsi dari model yang diusulkan oleh (Plissiti, et al., 2011). Inti dari sub tahap deteksi *region minima* adalah mencari *region maxima* dari citra hasil *preprocessing* pada tahap sebelumnya. Ketika mendapat wilayah dengan intensitas tinggi (*region maxima*) tentu akan mendapati wilayah-wilayah dengan intensitas rendah (*region minima*) karena wilayah yang bukan *region maxima* sudah pasti adalah *region minima*.

Langkah pertama yang dilakukan untuk mendapatkan *region maxima* dari citra adalah melakukan transformasi *h-minima* terhadap citra hasil *preprocessing* (Gambar 5.1(a)) pada tiap *layer* warnanya (merah, hijau, biru) dengan nilai *h* tertentu. Pembahasan mengenai nilai dari tiap *variable* pada bab ini akan dibahas pada sub bab 5.2. Gambar 5.1(b) adalah citra hasil proses transformasi *h-minima* pada *layer* warna merah.

Proses transformasi *h-minima* akan mencari wilayah-wilayah yang memiliki intensitas kurang dari atau sama dengan nilai ambang, menghilangkan area-area dengan intensitas rendah namun lebih besar dari nilai ambang dan menambahkan nilai h pada wilayah yang memiliki intensitas rendah tersebut sebagai penanda. Penambahan nilai h pada tiap intensitas rendah itu menyebabkan intensitas yang tadinya sudah rendah menjadi tinggi. Gambar 5.1(b) menunjukkan bahwa warna latar belakang citra cenderung menjadi abu-abu karena intensitasnya naik oleh proses transformasi *h-minima*.

Langkah selanjutnya adalah melakukan subtraksi terhadap citra hasil transformasi *h-minima* pada tiap *layer* warnanya dengan nilai h sama dengan nilai h pada proses transformasi minima. Hal ini dilakukan untuk mengembalikan intensitas warna pada daerah dengan intensitas rendah agar kembali seperti semula. Persamaan yang digunakan untuk proses subtraksi ini adalah sebagai berikut:

$$G(p) = I(p) - h, \quad p \in D_I \quad (5.1)$$

dengan I adalah citra hasil transformasi *h-minima* pada tiap *layer* warna dan h adalah nilai ambang yang senilai dengan nilai h pada proses transformasi *h-minima*. Proses subtraksi akan mengurangi nilai dari masing-masing *pixel* pada dimensi citra I (D_I) dengan *threshold* h (Gambar 5.1 (c)).

Langkah selanjutnya, *grayscale reconstruction* dilakukan untuk mendapatkan intensitas yang tinggi dari citra hasil proses *Preprocessing* (Gambar 5.1(a)) pada tiap *layer* warna. Menurut (Vincent, 1993) *Grayscale reconstruction* dilakukan sesuai dengan persamaan berikut:

$$\delta_I^{(1)}(G) = (G \oplus B) \wedge I \quad (5.2)$$

dengan G adalah citra hasil transformasi *h-minima* (persamaan 5.1) pada tiap *layer* warna, B adalah *flat structuring element* dan I adalah citra hasil akhir *preprocessing*

(Gambar 5.1(a)) pada tiap *layer* warna. Tiap *pixel* pada citra G akan didilasi oleh *flat structuring element*, namun hasilnya tidak boleh melebihi intensitas pada citra I (*pointwise minimum* $\wedge$). Persamaan 5.2 dilakukan sebanyak n kali seperti persamaan berikut:

$$\delta_I^{(n)}(G) = \underbrace{\delta_I^{(1)} \left(\delta_I^{(1)} \left(\delta_I^{(1)} \dots \left(\delta_I^{(1)}(G) \right) \right) \right)}_{n \text{ times}} \quad (5.3)$$

Grayscale Reconstruction $\rho_I(G)$ dari citra *mask* I dan *marker* G dilakukan sampai tidak ada perubahan nilai intensitas warna pada citra G . Sehingga persamaan 5.3 menjadi seperti persamaan 5.4. Gambar 5.1(d) adalah citra hasil *grayscale reconstruction*.

$$\rho_I(G) = \lim_{n \rightarrow +\infty} \delta_I^{(n)}(G) \quad (5.4)$$

Penelitian ini mengusulkan beberapa langkah untuk memperbaiki model yang diusulkan Plissiti. Langkah perbaikan tersebut antara lain adalah melakukan substraksi dari citra hasil akhir tahap *preprocessing* (Gambar 5.1(a)) pada tiap *layer* warnanya dengan citra hasil persamaan 5.4 atau hasil *grayscale reconstruction* untuk mendapatkan wilayah yang memiliki intensitas yang tinggi. Proses ini dilakukan dengan persamaan berikut:

$$S(p) = I(p) - G(p), \quad p \in D_I \quad (5.5)$$

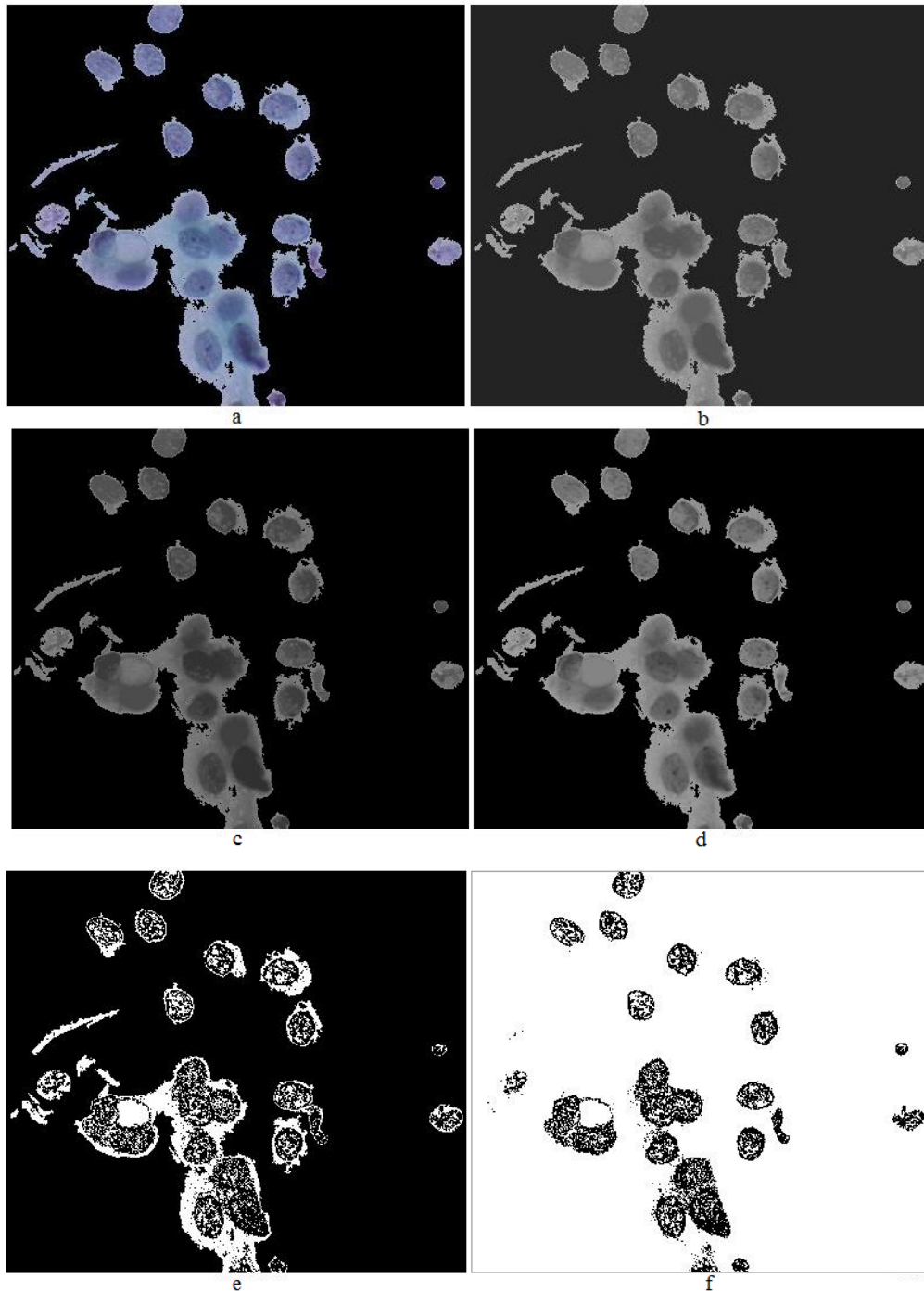
dengan I adalah citra hasil akhir tahap *preprocessing* (Gambar 5.1(a)) dan G adalah citra hasil *grayscale reconstruction*. Persamaan 5.5 dilakukan untuk setiap *layer* warna. Intensitas warna citra S pada tiap *pixel*-nya didapat dari proses pengurangan tiap *pixel* pada citra I oleh citra G dalam dimensi $I(D_I)$.

Langkah selanjutnya, citra hasil *grayscale reconstruction* pada *layer* warna merah (IM_1), *layer* warna hijau (IM_2) dan *layer* warna biru (IM_3) digabungkan menggunakan persamaan berikut:

$$IM = IM_1 \cup IM_2 \cup IM_3 \quad (5.6)$$

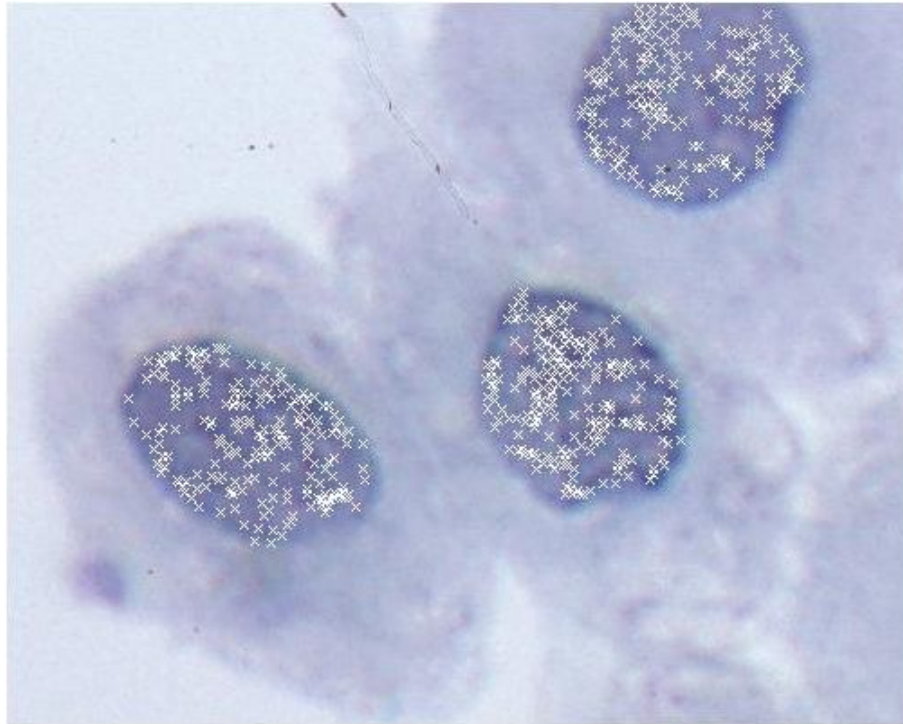
citra IM adalah citra yang merupakan area yang mengandung *region maxima* (Gambar 5.1(e)). langkah selanjutnya untuk mendapatkan *region minima* adalah mencari nilai minimum antara *inverse*($\neg$) citra IM (Gambar 5.1(e)) dengan citra biner hasil *preprocessing* (BM) (Gambar 4.5(b)). Hasil pencarian nilai minimum tersebut kemudian dilakukan proses *inverse* seperti persamaan 5.7. Gambar 5.1(f) adalah citra hasil persamaan 5.7 berikut:

$$G = \neg (\neg IM \cap BM) \quad (5.7)$$



Gambar 5.1. (a) Citra RGB hasil *preprocessing*, (b) Citra hasil transformasi *h-minima* pada *layer* merah, (c) Citra hasil subtraksi(b), (d) Citra hasil *grayscale reconstruction*, (e) Citra hasil persamaan 5.6, dan (f) Citra hasil persamaan 5.7

Langkah yang terakhir, dari citra yang dihasilkan pada persamaan 5.7 (Gambar 5.1(f)) kemudian dicari posisi *centroids*-nya. Gambar 5.2 adalah hasil akhir dari sub deteksi *region minima*. *Centroids* ditandai dengan tanda (x).



Gambar 5.2. Citra dengan *centroids* hasil akhir sub tahap deteksi *region minima*

5.1.2 Deteksi Tepi *Nuclei*

Sub tahap deteksi *region minima* telah mendeteksi *centroids* yang berada di dalam maupun di luar *nuclei*, akan tetapi *centroids* yang dihasilkan masih sangat banyak. Langkah selanjutnya jumlah *centroids* akan dikurangi dengan cara menghilangkan *centroids* berdasarkan letaknya.

Sub tahap deteksi tepi *nuclei* ini akan mendeteksi sel *nuclei* dengan mencarinya berdasarkan deteksi tepi dengan sedikit mengadopsi model yang ditawarkan oleh Plissiti (Plissiti, et al., 2011). Telah diketahui sebelumnya, sel *nuclei* umumnya

berbentuk seperti elips, maka proses pencarian tepi *nuclei* dapat dilakukan dengan memanfaatkan morfologi *gradient* dari dua citra yang berbeda. Sub tahap ini akan dibuat dua citra untuk melakukan proses morfologi *gradient*. Citra pertama dihasilkan dari proses erosi terhadap citra asli pada tiap *layer* warnanya sebagai berikut:

$$A = I \ominus X \quad (5.8)$$

dengan I adalah citra asli dan X adalah *flat disk-shaped structuring element* dengan radius 3. Erosi dilakukan untuk memperbesar area dengan intensitas rendah yang dalam hal ini adalah *nuclei* (Gambar 5.3(b)).

Proses morfologi erosi ini menyebabkan area *nuclei* menjadi lebih besar, akan tetapi proses ini juga memperbesar noise yang ada. Langkah selanjutnya untuk mengurangi *noise*, dibuat citra yang kedua dengan menggunakan 5×5 *averaging filter* terhadap citra asli (Gambar 5.3(c)). Berikut persamaan dari proses *averaging filter*:

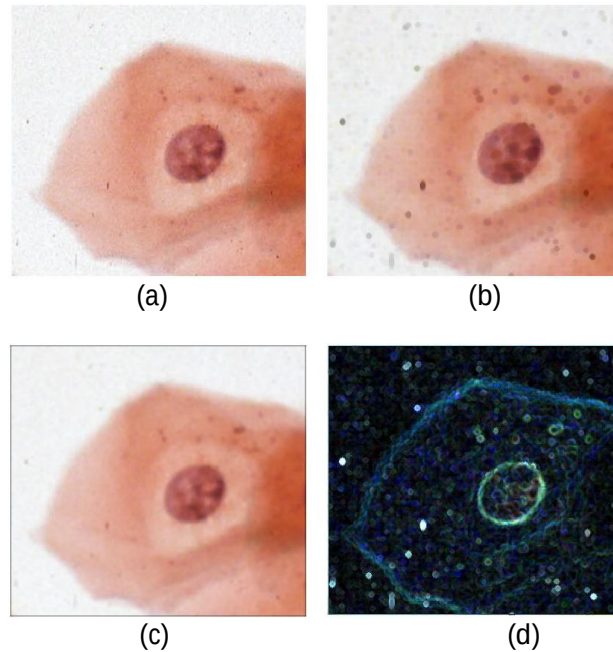
$$G(x, y) = \frac{1}{m.n}, 1 \leq x \leq m, 1 \leq y \leq n \quad (5.9)$$

dengan m dan n digunakan untuk menentukan jumlah tetangga yang dilibatkan. Nilai $m=5$, dan $n=5$ pada penelitian ini.

Morfologi *gradient* J terhadap citra I untuk mendapatkan tepi *nuclei* didapatkan dari persamaan berikut:

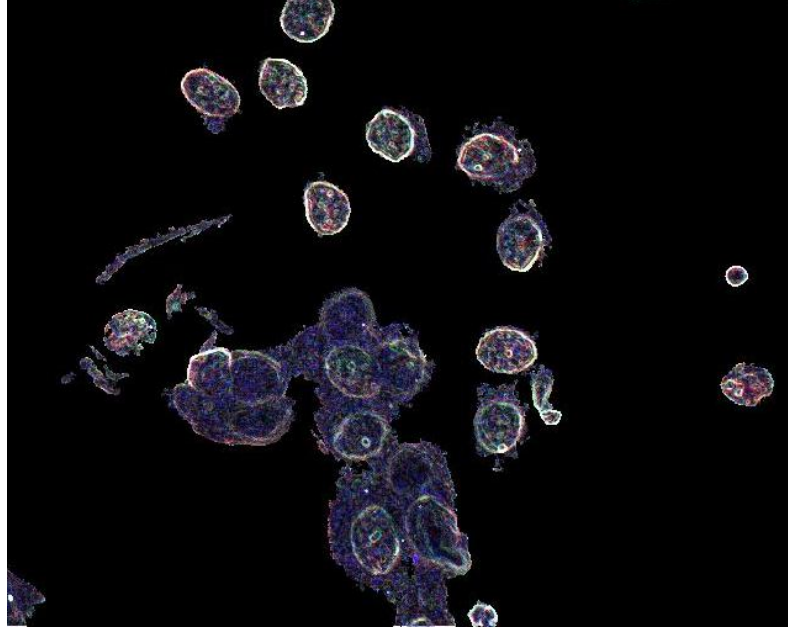
$$J(x, y) = |A(x, y) - B(x, y)| \quad (5.10)$$

dengan A adalah citra hasil proses erosi (persamaan 5.8) dan B adalah citra hasil 5×5 *filter* rata-rata (persamaan 5.9). *Contrast enhancement filter* dilakukan pada citra hasil morfologi *gradient* untuk menajamkan contrastnya dengan tingkat *saturates* 1% pada intensitas rendahnya dan 1% pada intensitas tingginya.



Gambar 5.3. (a) Citra asli, (b) Citra hasil erosi, (c) Citra hasil *averaging filter*, dan (d) Citra morfologi *gradient* dari citra (b) dan (c) dengan *filter* peningkatan contrast.

Penelitian ini mengusulkan beberapa langkah untuk memperbaiki model yang diusulkan Plissiti. Langkah pertama, untuk mengurangi hasil pencarian citra hasil morfologi *gradient* maka citra yang dihasilkan akan diekstrak latar belakangnya. Proses ekstraksi dilakukan dengan cara menggunakan citra hasil akhir tahap *preprocessing* (Gambar 4.5(b)) yang memiliki nilai 1 akan diganti dengan nilai pada citra hasil morfologi *gradient* pada tiap-tiap *layer* warna. Gambar 5.4 adalah hasil gabungan 3 *layer* warna dari proses pemisahan latar belakang citra hasil morfologi *gradient*.



Gambar 5.4 Citra hasil pemisahan latar belakang dari citra hasil morfologi *gradient*

Citra yang dihasilkan pada proses sebelumnya (Gambar 5.4) pada tiap *layer* warnanya dilakukan normalisasi *min-max*, hal ini diperlukan untuk melakukan *global thresholding* dengan menggunakan metode Otsu. *Global thresholding* menggunakan metode Otsu akan menghasilkan nilai antara 0 sampai 1 maka intensitas warna pada tiap *pixel* citra hasil morfologi *gradient* (Gambar 5.4) harus dilakukan normalisasi. Normalisasi *min-max* didapat dengan persamaan berikut:

$$N(p) = \frac{I(p) - \min(I)}{\max(I) - \min(I)}, \quad p \in D_I \quad (5.11)$$

dengan I adalah citra tahap sebelumnya (Gambar 5.4), dan p adalah setiap *pixel* pada dimensi (D_I). Hasil normalisasi *min-max* akan diambil nilai pada tiap *pixel* yang lebih besar daripada nilai *threshold* yang dihasilkan dengan metode Otsu, sehingga persamaan 5.11 dilanjutkan menjadi seperti berikut:

$$\begin{aligned} \text{If } N(p) > T, \quad p \in D_I \\ M(p) = 1 \end{aligned} \tag{5.12}$$

Else

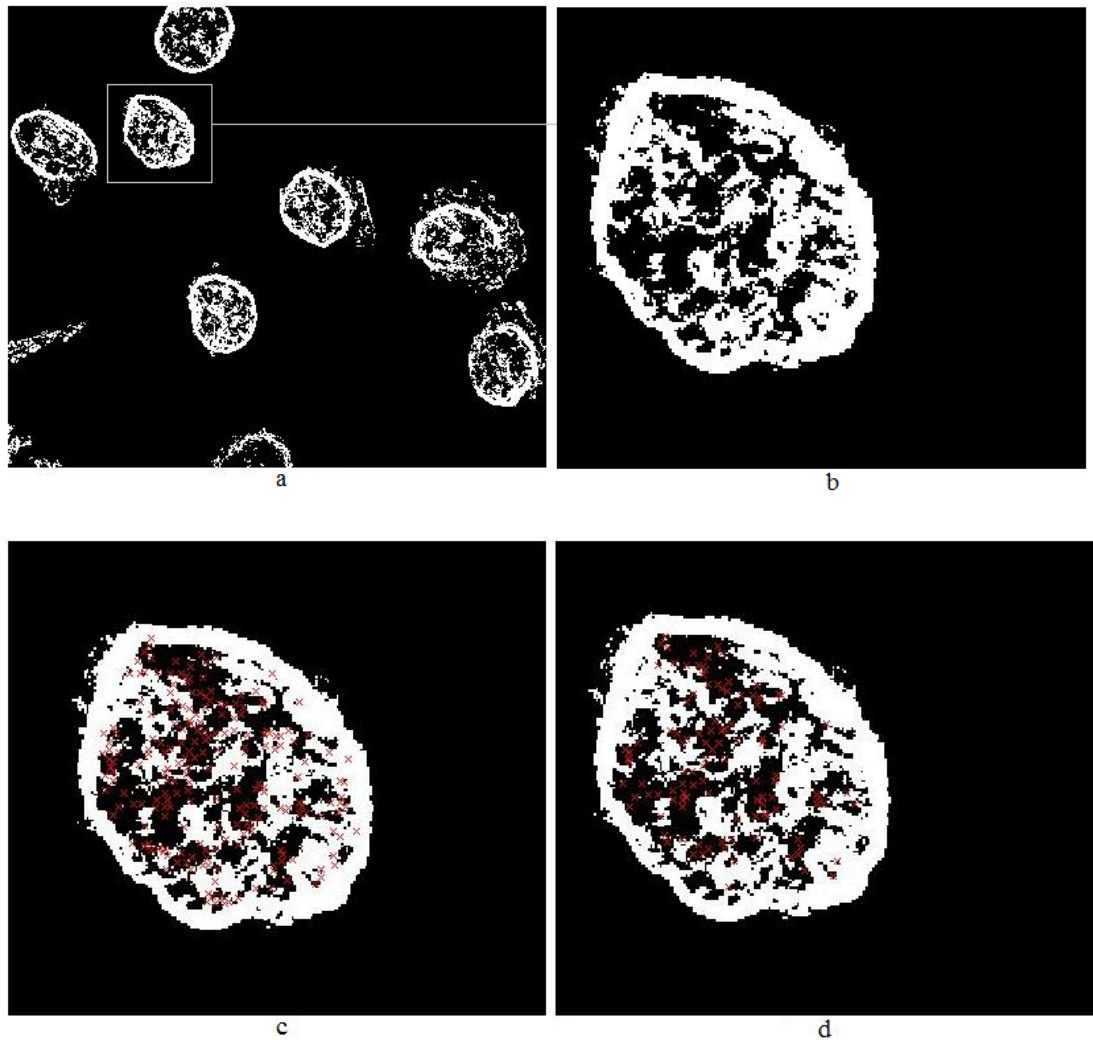
$$M(p) = 0$$

dengan T adalah nilai *threshold* dengan metode Otsu pada citra I . M adalah citra biner hasil *global thresholding* menggunakan metode Otsu dan p adalah setiap *pixel* pada dimensi $N(D_I)$. Persamaan 5.12 diatas dilakukan pada tiap-tiap *layer* warna.

Output dari tahap diatas digabungkan menggunakan *intersect* setiap *layer* warnanya sehingga mendapatkan citra biner dari persamaan berikut:

$$T = N_1 \cap N_2 \cap N_3 \tag{5.13}$$

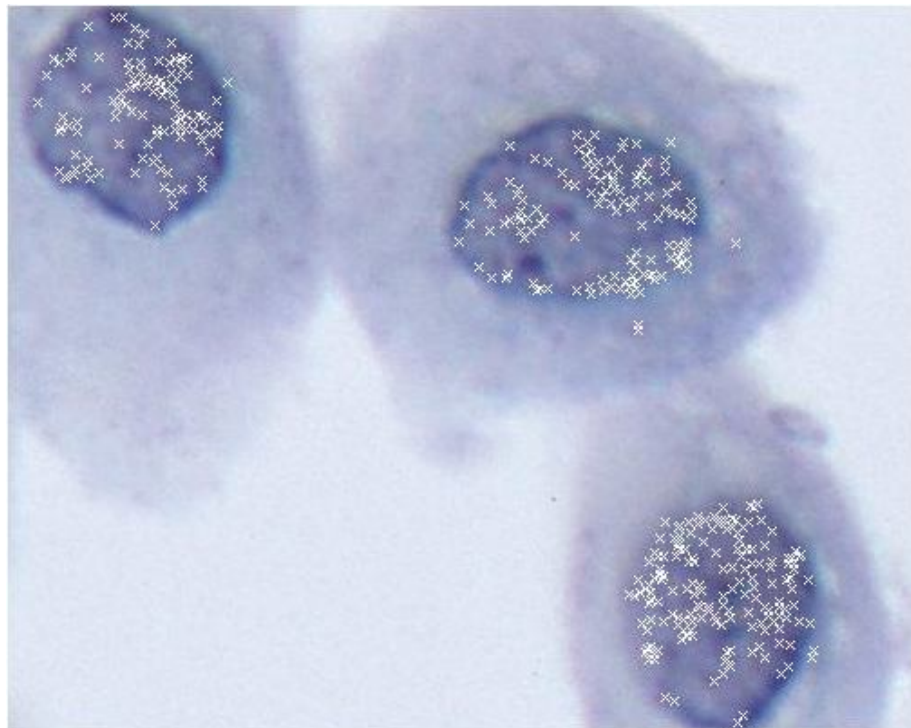
dengan (N_1) adalah citra hasil persamaan 5.12 pada *layer* warna merah, (N_2) citra hasil persamaan 5.12 pada *layer* warna hijau, dan (N_3) citra hasil persamaan 5.12 pada *layer* warna biru. Hasil dari proses ini akan menghasilkan area dengan bentuk seperti elips dimana bentuk ini dihasilkan dari tepi-tepi *nuclei* (Gambar 5.5(a)). Langkah selanjutnya adalah dari *centroids* yang ditemukan pada tahap sebelumnya (Gambar 5.2) akan dicari *centroids* yang tidak berada pada area yang bernilai 1 pada citra hasil persamaan 5.13.



Gambar 5.5. (a) Citra hasil persamaan 5.13, (b) Citra hasil perbesaran (a), (c) Citra (b) dengan *centroids*, dan (d) Citra dengan *centroids* dimana *centroids* berada di area tidak sama dengan 1.

Gambar 5.5(b) adalah perbesaran dari Gambar 5.5(a) dan Gambar 5.5(c) adalah *centroids* yang ditemukan pada tahap deteksi *region minima*. Gambar 5.5(c) menggambarkan banyak terdapat *centroids* yang berada di area putih/ area pada citra yang bernilai 1. Langkah yang terakhir, adalah menghilangkan *centroids* dimana *pixel* tempat *centroids* berada tidak memiliki nilai sama dengan 1 (area berwarna putih).

Citra yang dihasilkan pada langkah sebelumnya (Gambar 5.5(d)) menggambarkan bahwa *centroids* yang berada pada *pixel* yang bernilai 1 sudah hilang, hal ini dilakukan untuk meminimalisir jumlah *centroids* yang ditemukan sekaligus mengilangkan *centroids* yang berada di tepi-tepi *nuclei*. *Centroids* yang berada di tepi-tepi *nuclei* harus dihilangkan karena yang ingin dihasilkan adalah *centroids* yang berkumpul ditengah *nuclei* sehingga pada tahap selanjutnya *centroids* yang ditemukan dapat lebih diminimalisir dengan kriteria jarak. Gambar 5.6 adalah citra hasil akhir sub tahap deteksi tepi *nuclei*.

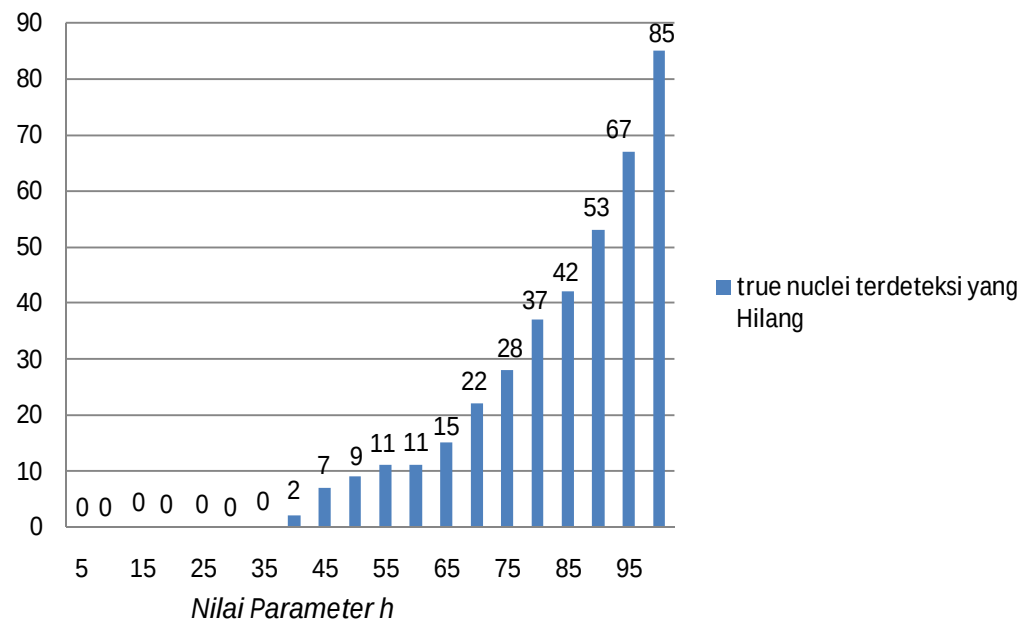


Gambar 5.6 Citra dengan *centroids* pada akhir tahap deteksi *centroids* kandidat sel *nuclei*

Gambar 5.6 terdapat tanda (x) menunjukkan *centroids* yang dihasilkan pada akhir tahap deteksi *centroids* kandidat sel *nuclei*.

5.2 Pembahasan Parameter h

Nilai parameter h pada sub tahap deteksi *region minima* ditentukan berdasarkan ujicoba yang dilakukan terhadap 13 citra dalam *dataset* penelitian ini. Gambar 5.7 adalah grafik hasil ujicoba yang dilakukan untuk menentukan nilai h .



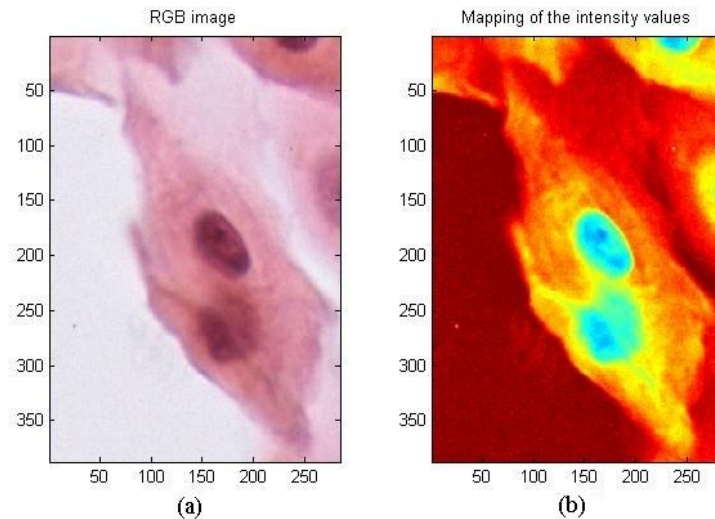
Gambar 5.7 Grafik ujicoba penentuan nilai parameter h pada sub tahap deteksi *region minima*

Gambar 5.7 menggambarkan bahwa semakin besar nilai h maka *true nuclei* terdeteksi yang hilang semakin besar, namun *false positive* yang ditemukan semakin sedikit. Berdasarkan hasil ujicoba yang dilakukan, nilai h yang dipakai dalam penelitian ini adalah 35. Nilai 35 diambil karena merupakan nilai maksimal dari hasil pengujian dengan tidak terdapat *true positive nuclei* yang hilang.

5.3 Analisis Permasalahan pada *Dataset*

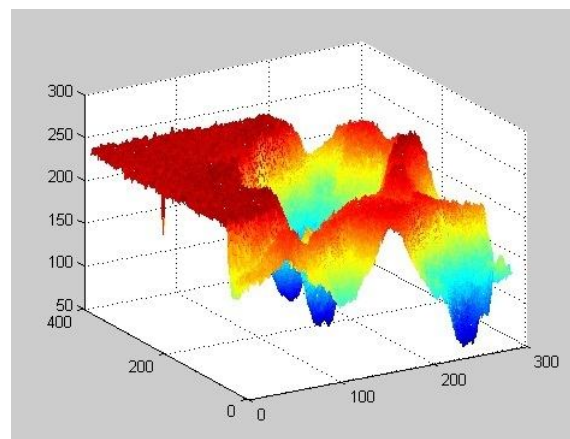
Secara umum, sel *nuclei* tampak lebih gelap daripada *cytoplasma* yang mengelilinginya. Gambar 5.8(b) menggambarkan bahwa daerah yang memiliki

intensitas warna rendah ditunjukkan dengan warna biru, sedangkan daerah yang memiliki intensitas tinggi ditunjukkan dengan warna merah. Daerah yang memiliki intensitas rendah inilah yang mungkin merupakan kandidat sel *nuclei*.



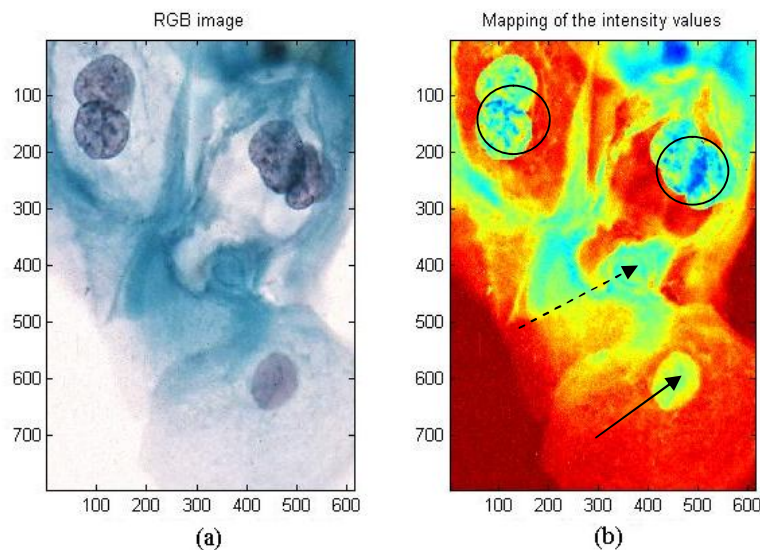
Gambar 5.8 Citra sel serviks dan peta intensitas warnanya

Gambar 5.9 menunjukkan bahwa yang dapat dilakukan untuk menghitung jumlah *nuclei* adalah dengan cara menghitung jumlah lembah yang terbentuk untuk melakukan perhitungan jumlah *nuclei* pada citra *Pap smear*. Gambar 5.9 adalah citra tiga dimensi yang menggambarkan intensitas warna yang dimiliki gambar 5.8 (a).



Gambar 5.9 Peta intensitas warna citra serviks dalam 3D

Tidak semua daerah yang memiliki intensitas warna yang rendah termasuk sel *nuclei*. Gambar 5.10(b) menunjukkan bahwa justru intensitas paling rendah merupakan daerah yang merupakan perpotongan antara dua *nuclei*. Daerah ini ditandai dengan lingkaran pada Gambar 5.10(b). Gambar 5.10(b) menunjukkan bahwa intensitas warna *nuclei* yang ditandai dengan tanda panah memiliki intensitas yang hampir sama besar dengan *cytoplasm*a yang berada di sekitarnya yang ditandai dengan anak panah dengan garis putus-putus.

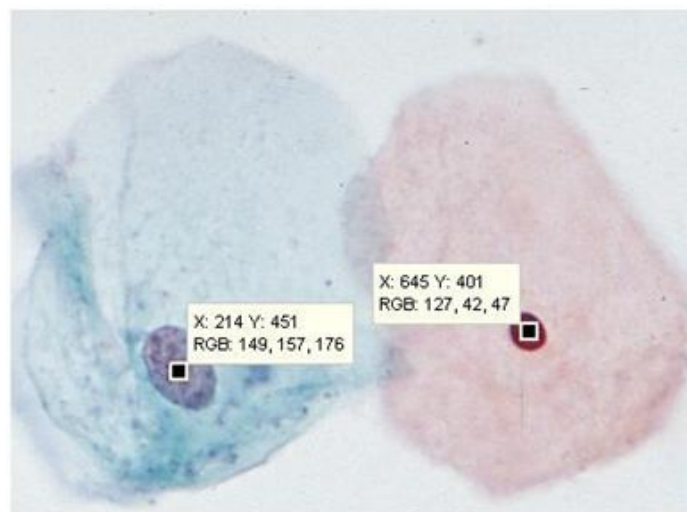


Gambar 5.10 Citra Sel Serviks dan Peta Intensitas Warnanya

Selain masalah yang telah disebutkan terdapat masalah lain yaitu perbedaan intensitas warna pada *nuclei* yang beragam. Perbedaan warna pada *nuclei* yang signifikan sangat berpengaruh pada tahap *clustering*. Perbedaan warna yang signifikan mengakibatkan *centroids* yang sudah tepat terletak pada daerah *nuclei* menjadi hilang apabila *nuclei* tersebut memiliki intensitas yang tinggi. Gambar 5.11 menunjukkan bahwa dalam satu citra sangat mungkin terjadi intensitas warna pada masing-masing sel *nuclei*nya beragam.

Gambar 5.11 menunjukkan bahwa *nuclei* pada sel serviks sebelah kanan (warna merah) memiliki intensitas pada *layer* warna merah = 127, *layer* warna hijau = 42

dan, *layer* warna biru = 47. Intensitas warna yang dimiliki pada *nuclei* ini tentu lebih rendah daripada sel *nuclei* yang berada di sebelah kiri, yaitu pada *layer* warna merah 149, *layer* warna hijau = 157 dan pada *layer* warna biru =176. Kedua *nuclei* tersebut memiliki intensitas yang tidak *homogen*, sangat besar kemungkinannya salah satu dari *nuclei* tersebut tidak teridentifikasi sebagai sel *nuclei* pada tahap *clustering*.



Gambar 5.11 Intensitas warna pada masing-masing *nuclei*

BAB VI

PERBAIKAN POSISI CENTROIDS KANDIDAT SEL NUCLEI

6.1 Langkah yang Dilakukan

Citra yang dihasilkan pada tahap deteksi *centroids* kandidat sel *nuclei* masih terdapat banyak *centroids* yang ditemukan. Penelitian ini mengusulkan beberapa langkah untuk memperbaiki model yang diusulkan Plissiti pada tahap perbaikan posisi *centroids* kandidat sel *nuclei* ini. *Centroids* yang telah terdeteksi ditahap sebelumnya akan direduksi dengan melakukan morfologi dilasi pada tiap *pixel* tempat *centroids* berada. Proses yang dilakukan adalah seperti algoritma berikut:

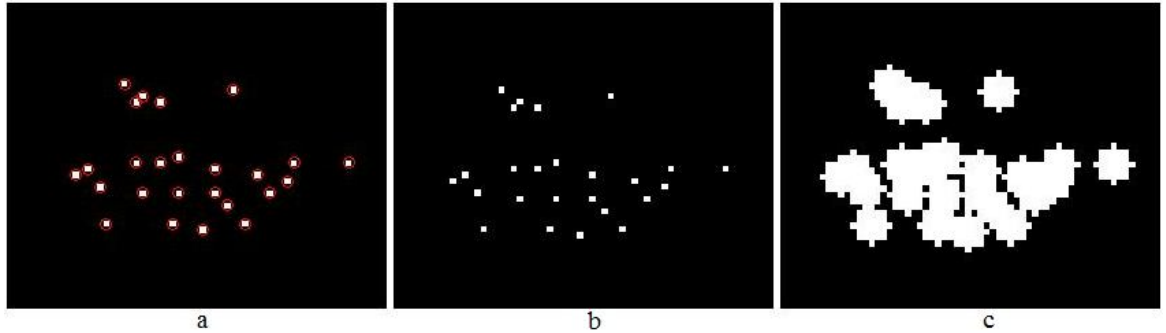
Repeat

$$h_{k+1} = h_k \oplus B$$

$$Lc' = total(h_{k+1});$$

Until $Lc \neq Lc'$

dengan B adalah *flat disk-shaped structuring element* dengan radius 3, h_k adalah *pixel* tempat *centroids* berada. Lc adalah jumlah *centroids* pada tahap deteksi *centroids* kandidat sel *nuclei* dan Lc' adalah jumlah *centroids* baru yang dihasilkan oleh proses morfologi dilasi sebelumnya. Gambar 6.1 berikut menggambarkan urutan citra pada tiap langkah algoritma yang telah disebutkan sebelumnya.



Gambar 6.1. (a) Posisi *centroids* pada tahap deteksi *centroids* kandidat sel *nuclei*, (b) *Pixel* letak *centroids* berada, (c) Morfologi dilasi terhadap citra (b)

Gambar 6.1(a) menggambarkan *centroids* yang ditandai dengan tanda “o”. Tiap *centroids* yang ditemukan diambil lokasi *pixel* tempat *centroids* berada (Gambar 6.1(b)). Algoritma sebelumnya akan memperbesar area tempat lokasi *centroids* berada hingga tampak seperti Gambar 6.1(c). Langkah selanjutnya, dari area yang dihasilkan (Gambar 6.1(c)) dicari *pixel* dengan intensitas yang paling rendah. Algoritma berikut dilakukan untuk mendapatkan area dengan intensitas yang rendah dari area yang telah dibentuk pada tahap sebelumnya Gambar 6.1(c),

$i = 1$

Repeat

$\forall p = (x, y) \in LA(i)$

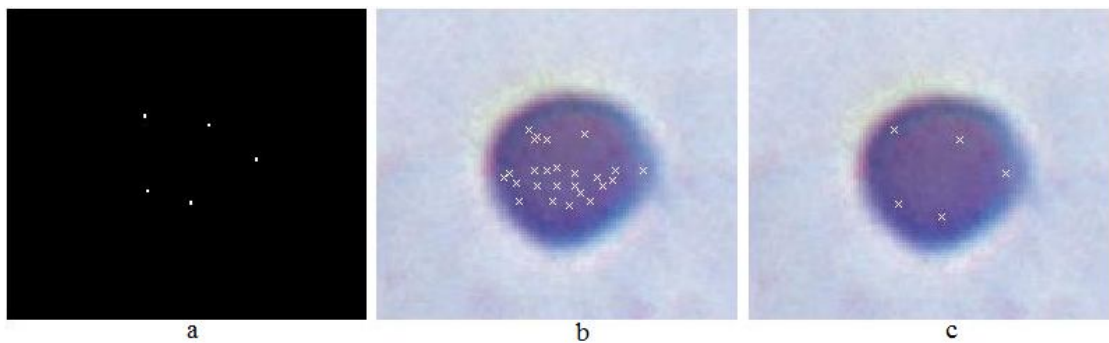
Select $r = \{p \mid \min(I(p))\}$

$i++$

Until $i = nL$

dengan LA adalah area terlabelisasi yang memiliki label sama dengan i , nL adalah jumlah area yang terlabelisasi, I adalah citra *grayscale* dari citra asli. Himpunan r pada algoritma di atas bernilai koordinat (x,y) pada citra yang memiliki nilai paling rendah dalam satu area yang telah terlabelisasi.

Tiap koordinat (x,y) atau *pixel* r diberi nilai 1 sebagai penanda dan semua area pada citra yang bukan termasuk r akan diset menjadi 0. Citra yang dihasilkan pada algoritma di atas (Gambar 6.2(a)) kemudian dicari *centroids*-nya. Gambar 6.2(c) adalah hasil pencarian *centroids* dari citra 6.2(a).



Gambar 6.2. (a) *Pixel* yang memiliki intensitas paling rendah tiap area terlabelisasi, (b) Citra dengan *centroids* pada akhir tahap deteksi *centroids* kandidat sel *nuclei*, (c) Citra dengan *centroids* pada akhir tahap perbaikan posisi *centroids* kandidat sel *nuclei*

Gambar 6.2(c) menunjukkan bahwa *centroids* yang ditemukan sudah berkurang secara signifikan. Tahap selanjutnya *centroids* yang ditemukan akan direduksi dengan kriteria jarak, artinya *centroids* yang memiliki jarak dengan *centroids* yang lain namun jaraknya kurang dari nilai ambang akan dipilih intensitas warna diantara 2 *centroids* tersebut.

BAB VII

PENGURANGAN JUMLAH CENTROIDS MENGGUNAKAN KRITERIA JARAK

7.1 Langkah yang Dilakukan

Tahap perbaikan posisi *centroids* kandidat sel *nuclei* telah mendapatkan citra dengan *centroids* yang berada pada *nuclei*-nya, akan tetapi *nuclei* yang dihasilkan masih cukup banyak. Penelitian ini mengusulkan langkah untuk memperbaiki model yang diusulkan Plissiti. Beberapa *centroids* yang saling berdekatan akan dihilangkan pada tahap ini untuk mengurangi *centroids* yang berkerumun. Penelitian ini mengusulkan algoritma seperti berikut:

```

temp = ∅
Repeat
    ∀ p = (x, y) ∈ Rc
        if exists q = {(xq,yq) | D(p, q) ≤ T}
            temp = {temp, p, q | max {I(p), I(q)}}
Until p evaluated

Utemp = unique(temp)
RcNew = Rc Δ Utemp

```

dengan Rc adalah kumpulan *centroids* yang ditemukan pada tahap perbaikan posisi *centroids* kandidat sel *nuclei*, D adalah jarak *Euclidean* antara dua *centroids*, T adalah nilai ambang minimum jarak, dan $I(p)$ adalah intensitas pada citra pada *centroids* p ,

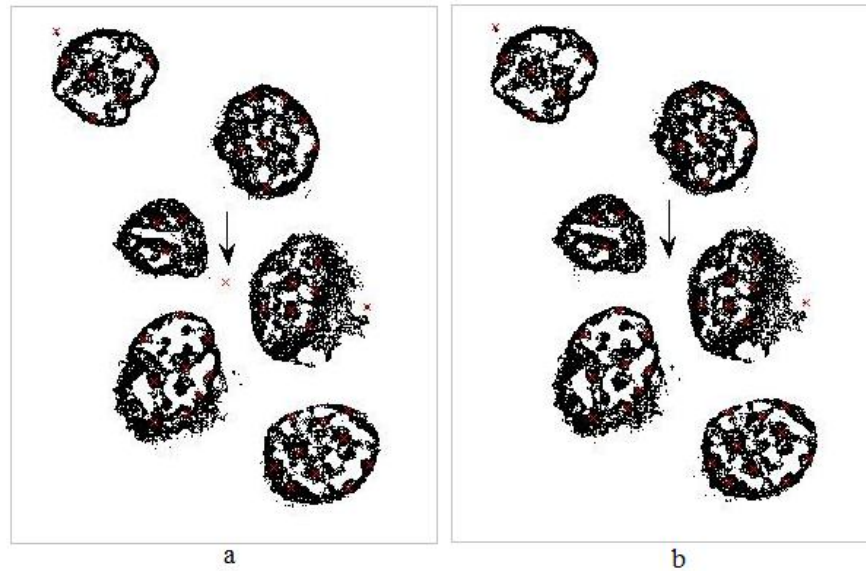
temp adalah *variable* penyimpanan sementara. *Variable temp* akan mengambil nilai yang tinggi diantara 2 intensitas warnanya. Hal ini berarti intensitas yang lebih rendah tetap bertahan, sedangkan intensitas yang lebih tinggi dimasukkan dalam *temp*. Nilai *T* didapat dari serangkaian uji coba yang akan dibahas pada sub bab 7.2. Pengurangan *centroids* akan minimal apabila *T* di set rendah, namun apabila nilai *T* di set tinggi dapat memungkinkan *centroids* yang sudah tepat di *nuclei* menjadi hilang.

Algoritma tersebut diulang sampai seluruh *p* terevaluasi. Adapun yang dimaksud dengan seluruh *p* terevaluasi yaitu seluruh *centroids* telah dievaluasi dengan kombinasi terhadap *q*. Jumlah seluruh kombinasi antara 2 *centroids* didapat dengan persamaan 7.1 berikut:

$${}_nC_k \equiv \binom{n}{k} \equiv \frac{n!}{k!(n-k)!} \dots\dots\dots (7.1)$$

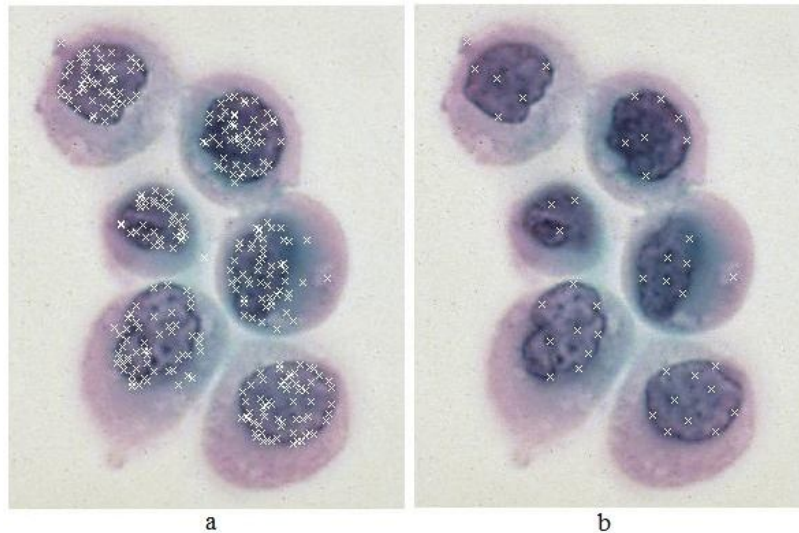
dengan *n* adalah jumlah *centroids* yang ditemukan sedangkan *k* = 2 karena hanya melibatkan kombinasi diantara 2 *centroids*.

Utemp adalah *variable* yang berisi nilai himpunan *temp* yang tidak redundan atau lebih dari satu. *RcNew* adalah selisih simetri (Δ) antra *Rc* dan *Utemp*. Nilai *RcNew* adalah koordinat *centroids* baru yang ditemukan. *Centroids* pada koordinat tersebut dicari intensitas warna tiap *pixel*-nya dari citra yang dihasilkan pada persamaan 5.7 dimana nilai intensitas tersebut sama dengan 0. Gambar 7.1(a) terdapat tanda panah yang menunjukkan *centroids* hilang pada Gambar 7.1(b) dimana *centroids* tersebut memiliki intensitas pada *pixel*-nya sama dengan 1 (berwarna putih). Gambar 7.1 adalah citra hasil persamaan 5.7 dengan *centroids* yang didapat dari algoritma diatas.



Gambar 7.1. (a) Citra hasil persamaan 5.7 dengan *centroids* pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak, (b) Citra hasil persamaan 5.7 dengan *centroids* pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak dengan intensitas citra hasil persamaan 5.7 sama dengan 0

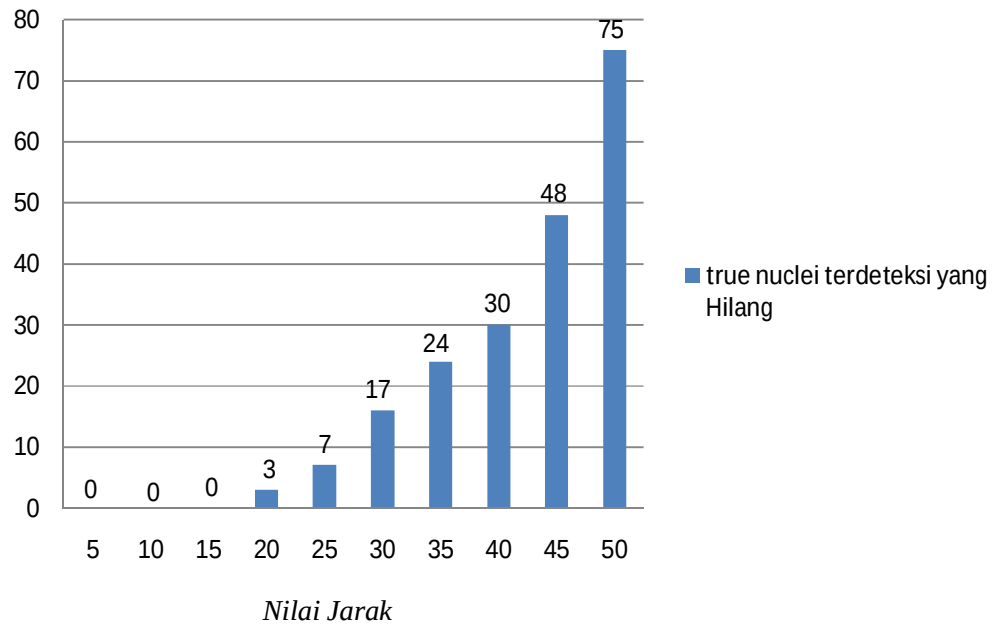
Aplikasi perbaikan letak *centroids* berdasarkan jarak ini mampu memperkecil jumlah *centroids* secara signifikan namun tetap relevan. Gambar 7.2 menunjukkan perbandingan antara hasil pada citra dengan *centroids* pada tahap perbaikan *centroids* kandidat sel *nuclei* dengan perbaikan citra berdasarkan jarak.



Gambar 7.2. (a) Citra dengan *centroids* pada tahap perbaikan *centroids* kandidat sel *nuclei*, (b) Citra dengan *centroids* pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak

7.2 Pembahasan Parameter Jarak

Tahap pengurangan jumlah *centroids* menggunakan kriteria jarak, nilai parameter jarak antara *vector* U dan V didapatkan dengan rumus $D^{Euc} = \sqrt{(u - v)^T(u - v)}$. Nilai ambang jarak dalam penelitian ini didapat melalui uji coba terhadap 13 citra dari *dataset* yang digunakan. Gambar 7.3 adalah grafik ujicoba yang dilakukan untuk menentukan nilai ambang jarak(*T*).



Gambar 7.3 Grafik ujicoba penentuan nilai ambang jarak pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak

Uji coba yang dilakukan pada penelitian ini mengambil nilai ambang jarak yaitu 30, hal ini dikarenakan jumlah *false positive* yang ditemukan pada tahap ini masih sangat banyak. Parameter ambang jarak (T) tidak menggunakan nilai 15 yang memiliki *True Nuclei Lost* (TNL) sebanyak 0, hal ini dikarenakan apabila nilai yang dipakai adalah 15 *false positive* yang ditemukan masih sangat banyak. Nilai 30 dipilih karena *true positive nuclei* yang hilang hanya terdapat pada citra tertentu. Tabel 7.1 berisi detail percobaan yang dilakukan pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak.

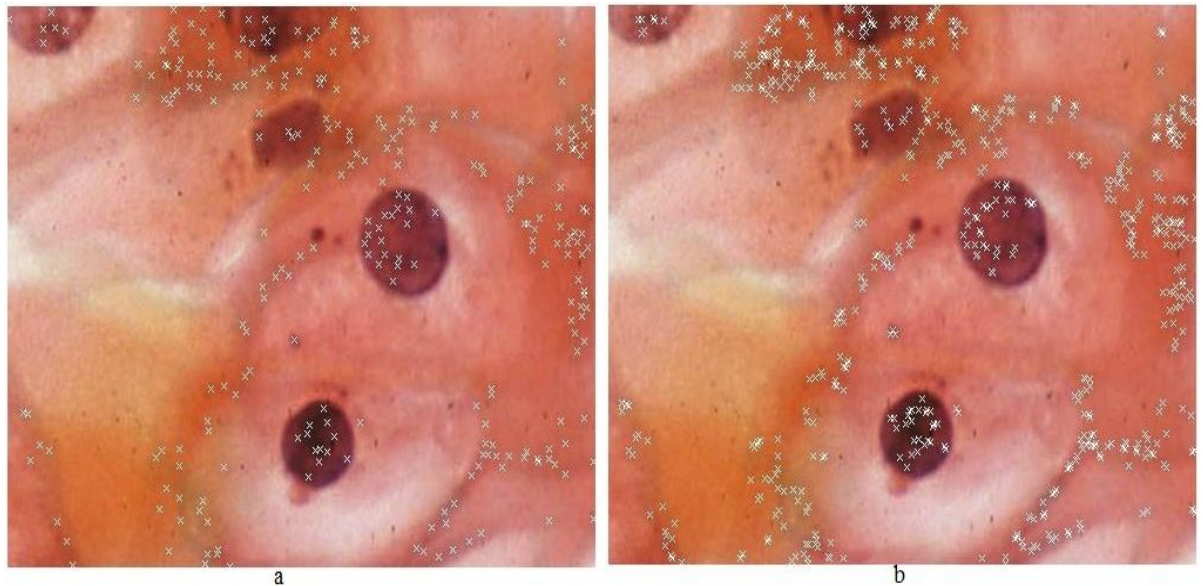
Tabel 7.1 Tabel ujicoba nilai ambang jarak

<i>File</i>	TNL (5)	TNL (10)	TNL (15)	TNL (20)	TNL (25)	TNL (30)	TNL (35)	TNL (40)	TNL (45)	TNL (50)
1	0	0	0	2	3	8	9	10	14	23
2	0	0	0	1	3	7	13	18	31	47
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	1	2	1	1	1	2
7	0	0	0	0	0	0	0	0	1	2
8	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	1	1	1	1
jml	0	0	0	3	7	17	24	30	48	75

Tabel 7.1 menunjukkan pada parameter ambang jarak (T) dengan nilai 30 mulai ditemukan *true positive nuclei* yang terdeteksi hilang sebanyak 17, namun pada sebagian besar citra tidak mengalami kehilangan *true positive nuclei*. Nilai ini dianggap masih relevan sebagai nilai ambang jarak karena pada sebagian besar citra yang digunakan sebagai *dataset* tidak kehilangan *true positive nuclei*.

7.3 Pembahasan Penggunaan Intensitas

Sub bab ini akan membahas mengenai penggunaan intensitas warna pada tiap *centroids*. Pengambilan intensitas warna pada masing-masing *centroids* dilakukan dengan memanfaatkan citra *grayscale* dari citra asli pada tahap perbaikan posisi *centroids* kandidat sel *nuclei*, pengurangan jumlah *centroids* menggunakan kriteria jarak, dan *clustering*. Berikut penjelasan mengenai pemanfaatan citra *grayscale* sebagai intensitas warna pada *centroids* berada. Gambar 7.4 menggambarkan perbandingan pemanfaatan citra *grayscale* dengan RGB pada intensitas rendah pada citra.

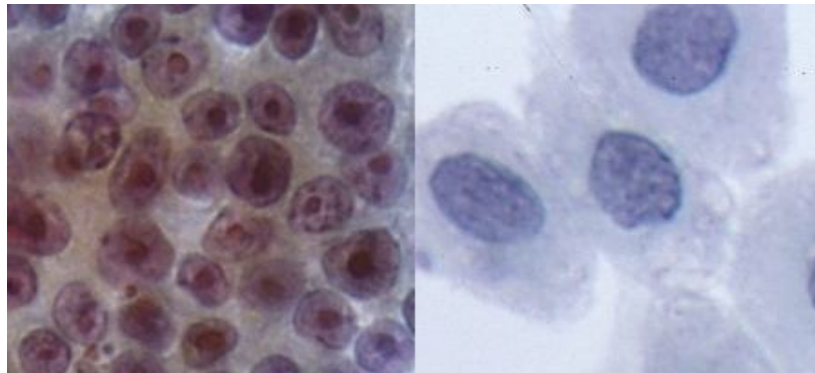


Gambar 7.4 (a) *Centroids* yang ditemukan menggunakan citra *grayscale* sebagai intensitas warna, dan (b) *Centroids* yang ditemukan menggunakan citra RGB sebagai intensitas warna,

Gambar 7.4(a) memiliki *centroids* yang lebih sedikit daripada gambar 7.4(b), hal ini dikarenakan penentuan intensitas terendah pada citra *grayscale* hanya menggunakan 1 citra, sedangkan pada citra RGB penentuan intensitas rendah diambil menggunakan 3 citra yaitu *Red*, *Green* dan *Blue*. Pemanfaatan RGB sebagai intensitas rendah pada tiap *centroids* menyebabkan *centroids* yang ditemukan menjadi lebih banyak dan menambah waktu komputasi, artinya waktu eksekusinya menjadi lebih panjang daripada menggunakan citra *grayscale*. Penelitian ini memanfaatkan citra *grayscale* sebagai pengambilan nilai intensitas warna pada tiap *centroids*.

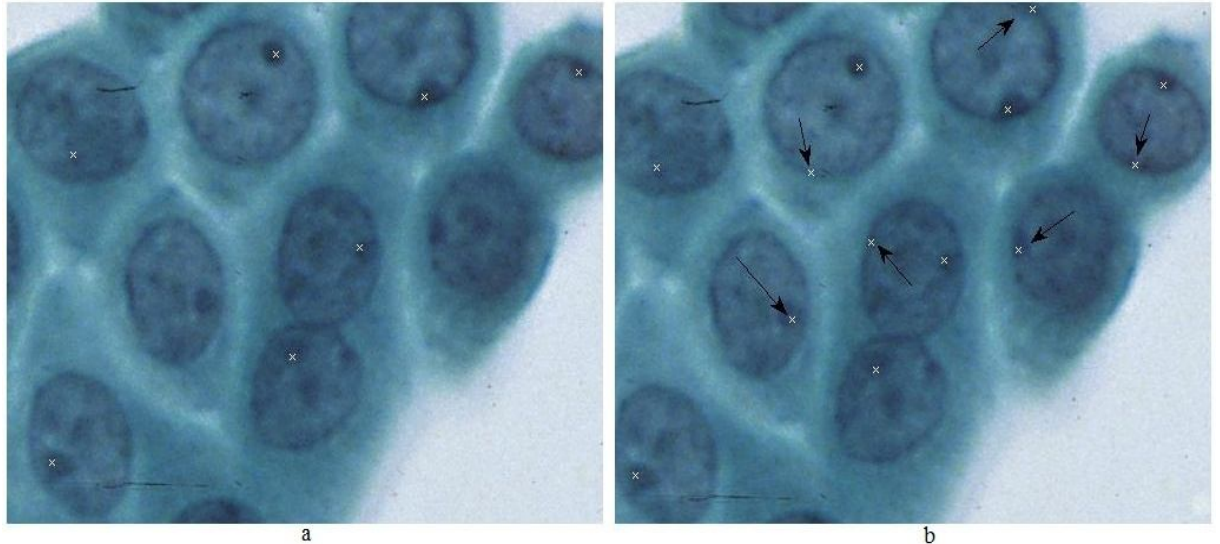
7.4 Analisis Permasalahan pada Dataset

Gambar 7.5 menunjukkan bahwa citra sebelah kiri memiliki sel *nuclei* yang jauh lebih kecil daripada *nuclei* pada citra sebelah kanan. Perbedaan ukuran *nuclei* yang signifikan ini mempengaruhi penghitungan nilai pada *variable* jarak terhadap *centroids* yang satu dan *centroids* yang lain pada tahap pengurangan jumlah *centroids* dengan kriteria jarak. Citra *Pap smear* dengan ukuran *nuclei* yang kecil, memiliki kemungkinan dalam satu *nuclei* terdapat satu *centroids* lebih besar daripada citra *Pap smear* dengan *nuclei* yang besar. Perbedaan perbesaran pada citra *Pap smear* sangat mempengaruhi nilai parameter jarak.



Gambar 7.5 Perbandingan besar *nuclei* pada dataset yang digunakan

Dataset yang digunakan pada penelitian ini sangat beragam. Citra yang satu dan yang lainnya memiliki keragaman dalam jarak antar *nuclei*-nya. Hal ini menyebabkan kesulitan tersendiri dalam menentukan jarak pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak. Gambar 7.6 menjelaskan bahwa apabila jarak pada tahap tersebut ditentukan tinggi maka banyak *nuclei* yang tidak terdeteksi. Namun apabila ditentukan jaraknya rendah maka dalam *nuclei* masih tersisa lebih dari 1 *centroids* dalam 1 *nuclei*.



Gambar 7.6 (a) Citra Jarak antar *nuclei* yang di set tinggi, dan (b) Citra Jarak antar *nuclei* yang di set rendah

Gambar 7.6(b) terdapat tanda panah yang menunjukkan perbedaan hasil *centroids* yang ditemukan terhadap Gambar 7.6(a). Penggunaan jarak pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak sangat mempengaruhi tujuan penggunaan model. Jarak yang tinggi sangat baik apabila tujuan penggunaan model adalah menghitung jumlah *nuclei* dari jumlah *centroids* yang ditemukan. Hal ini dikarenakan 1 *nuclei* diwakili oleh 1 *centroids*, akan tetapi hasil penggunaan jarak yang tinggi memiliki banyak *True Positive* yang hilang sehingga sensitivitas hasilnya lebih rendah daripada jika jaraknya rendah. Penggunaan jarak yang rendah sangat baik apabila digunakan untuk mendeteksi keberadaan *centroids*, karena pada penggunaan jarak yang rendah tidak banyak *true positive* yang terbuang. Penggunaan jarak yang rendah kurang baik apabila digunakan untuk tujuan menghitung jumlah *nuclei* dari *centroids* yang ditemukan, karena pada 1 *nuclei* terdapat lebih dari 1 *centroids* yang apabila dihitung tentu hasilnya tidak tepat.

BAB VIII

CLUSTERING

8.1 Langkah yang Dilakukan

Penelitian ini mengusulkan beberapa langkah *clustering* untuk memperbaiki model yang diusulkan oleh Plissiti. Metode yang digunakan pada tahap *clustering* adalah Fuzzy C-Means dengan kriteria yang digunakan dalam penelitian ini adalah jarak terdekat antara *centroids* yang satu dengan yang lain, dan intensitas warna pada tiap *pixel* tempat *centroids* berada. Algoritma berikut digunakan untuk mendapatkan jarak terdekat dari masing-masing *centroids*,

Repeat

$$\forall p = (x, y) \in C_n$$

$$\forall q = (x, y) \in C_n, \exists D(p, q)$$

$$temp = \{temp, D(p, q) \Delta 0\}$$

$$Ds = \min(temp)$$

Until *p* evaluated

dengan C_n adalah kumpulan *centroids* yang ditemukan pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak, D adalah jarak *Euclidean* antara dua *centroids*, $temp$ adalah selisih simetri (Δ) antara D dan 0 . Algoritma tersebut diulang sampai seluruh p terevaluasi. Dimaksud dengan seluruh p terevaluasi yaitu seluruh *centroids* telah dievaluasi dengan permutasi terhadap q . Jumlah seluruh permutasi antara 2 *centroids* didapat dengan persamaan 8.1 berikut:

$${}_nP_k \equiv \frac{n!}{(n-k)!} \quad (8.1)$$

dimana n adalah jumlah *centroids* yang ada sedangkan $k = 2$ karena hanya melibatkan 2 *centroids*.

Intensitas warna yang dimaksud dalam algoritma tersebut adalah intensitas citra *grayscale* tiap *centroids* yang ditemukan pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak. Langkah selanjutnya, intensitas warna dan jarak terdekat antara masing-masing *pixel*-nya dilakukan normalisasi *min-max* karena besar nilai pada kedua kriteria tersebut yang berbeda. Normalisasi *min-max* terhadap jarak terdekat dirumuskan pada persamaan berikut:

$$Dist(n) = \frac{Ds(n) - \min(Ds)}{\max(Ds) - \min(Ds)}, \quad n \in Ds \quad (8.2)$$

normalisasi *min-max* terhadap intensitas warna dirumuskan pada persamaan berikut:

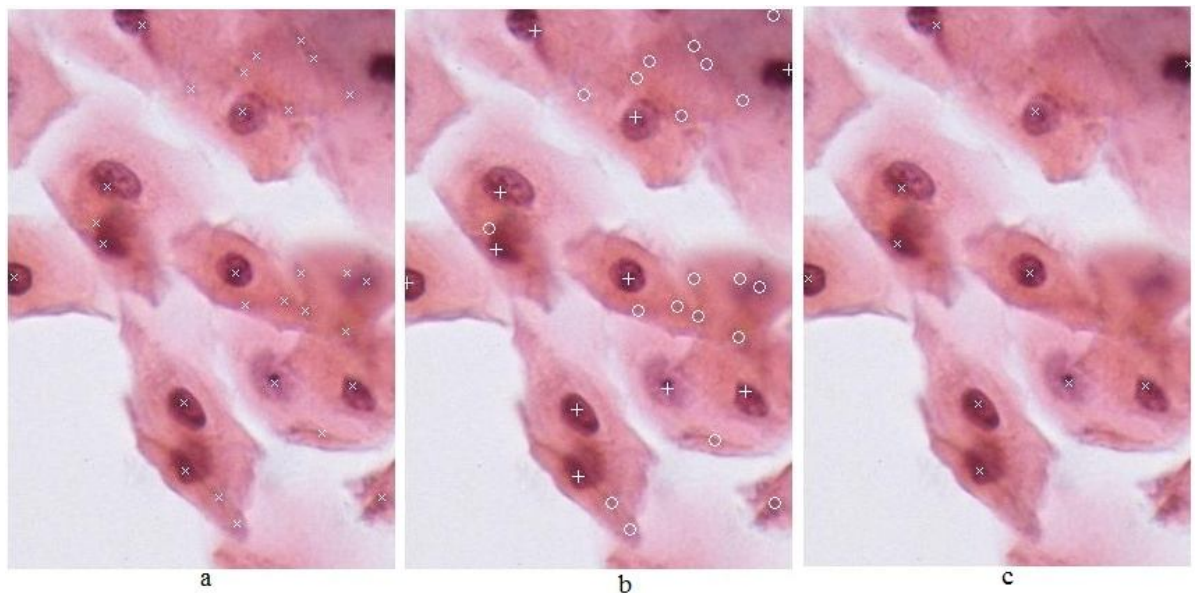
$$Int(p) = \frac{I(p) - \min(I)}{\max(I) - \min(I)}, \quad p \in D_I \quad (8.3)$$

dengan $Dist$ adalah himpunan dari normalisasi jarak terdekat, Ds adalah himpunan jarak terdekat dari algoritma penentuan jarak terdekat. I adalah intensitas citra *grayscale* dari citra asli pada tiap *pixel* tempat *centroids* berada dan p adalah setiap *pixel* pada dimensi(D_I). Pada penelitian ini intensitas warna adalah kriteria *cost* sedangkan jarak terpendek adalah kriteria *benefit* maka perlu adanya sebuah proses *preprocessing* seperti persamaan berikut:

$$N_{Dist} = 1 - Dist(n) \quad \dots\dots\dots (8.4)$$

dengan *Dist* adalah output dari persamaan 8.2. Setiap anggota dari himpunan *Dist* dikurangi dengan 1 pada persamaan 8.4. Pengurangan ini dilakukan untuk membalik nilai kriteria, dalam hal ini agar kriteria *benefit* dapat dibandingkan dengan kriteria *cost*.

FCM digunakan untuk membagi *centroids* yang ditemukan pada tahap sebelumnya menjadi 2 *cluster*, yaitu *cluster nuclei* dan *cluster* bukan *nuclei*. Hasil dari *clustering* menggunakan FCM adalah membagi himpunan jarak terdekat dan intensitas menjadi 2 *cluster*. Kedua buah *cluster* tersebut kemudian dibandingkan, intensitas warna rata-rata tiap anggota *cluster* yang bernilai rendah dianggap sebagai *positive class* atau *cluster centroids* sedangkan yang memiliki rata-rata tinggi dianggap sebagai *negative class* atau *cluster* bukan *nuclei*. Hasil dari proses *clustering* FCM dengan kriteria intensitas warna dan jarak terdekat dapat dilihat pada Gambar 8.1.

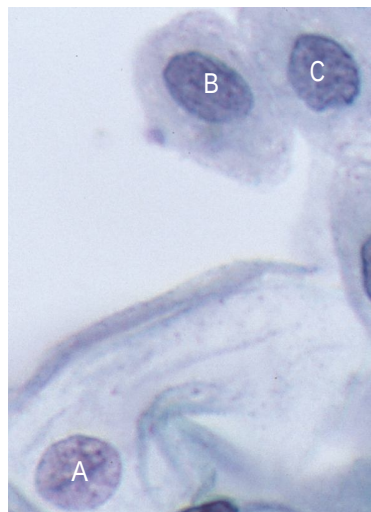


Gambar 8.1. (a) Citra dengan *centroids* pada tahap perbaikan *centroids* kandidat sel *nuclei*, (b) Hasil *clustering* FCM dengan *cluster nuclei* ditandai dengan “+” dan *cluster* bukan *nuclei* ditandai dengan “o”, dan (c) Hasil *clustering* FCM *cluster nuclei*.

Gambar 8.1(a) merupakan citra dengan *centroids* pada tahap perbaikan *centroids* kandidat sel *nuclei*. FCM akan membagi *centroids* yang ditemukan pada tahap tersebut menjadi dua *cluster* yaitu *cluster nuclei* dan *cluster* bukan *nuclei*. Gambar 8.1(b) terdapat tanda “o” yang berarti *cluster* bukan *nuclei* dan tanda “+” adalah *cluster nuclei*. Gambar 8.1(c) merupakan hasil dari proses penghilangan *cluster* bukan *nuclei* pada citra.

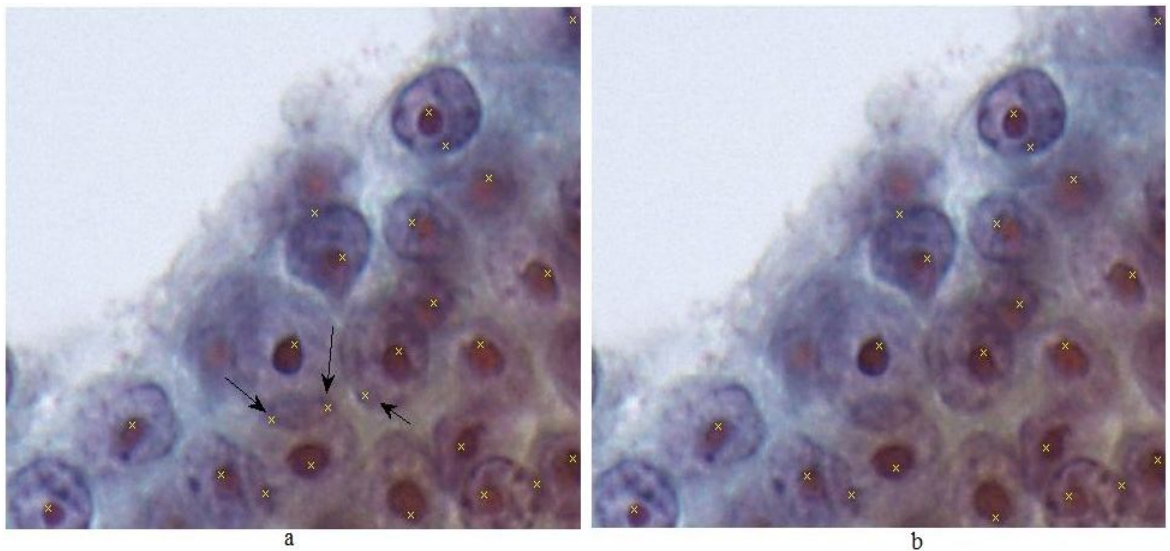
8.2 Pembahasan Kriteria Jarak Terdekat pada Tahap *Clustering*

Jumlah *cluster* yang akan dibentuk pada tahap *clustering* adalah 2, yaitu *cluster nuclei* dan bukan *nuclei*. Kriteria yang digunakan pada tahap ini adalah jarak terpendek antar *centroids* dan intensitas warna pada tiap *centroids*. Gambar 8.2 menjelaskan bahwa jika hanya intensitas warna yang digunakan untuk membuat *cluster* pada tahap *clustering* ini, maka *nuclei* A dengan intensitas warna yang relatif tinggi dibandingkan dengan *nuclei* B dan C bisa jadi tidak terdeteksi. *Nuclei* A memiliki jarak yang relatif jauh maka ada kemungkinan *nuclei* A terdeteksi sebagai *true positive*.



Gambar 8.2. Gambar penjelasan tahap *clustering*

Gambar 8.3 menggambarkan penggunaan jarak terdekat sebagai salah satu kriteria pada tahap *clustering*. Gambar 8.3(a) merupakan hasil *clustering* hanya dengan kriteria intensitas warna dan Gambar 8.3(b) merupakan hasil *clustering* menggunakan kriteria jarak terdekat dan intensitas warna. Gambar 8.3(a) menunjukkan bahwa *centroids* yang ditunjukkan dengan tanda panah, kemudian hilang pada Gambar 8.3(b). Hilangnya *centroids* pada Gambar 8.3(b) terjadi karena *centroids* yang ditunjukkan oleh tanda panah memiliki intensitas yang tidak terlalu rendah namun memiliki jarak yang relative yang dekat dengan *centroids* yang lain.



Gambar 8.3 (a) *Centroids* pada tahap *clustering* dengan kriteria intensitas warna, dan (b) *Centroids* pada tahap *clustering* dengan kriteria intensitas warna dan jarak terdekat

BAB IX

EVALUASI

9.1 Analisis Data

Dataset yang digunakan pada penelitian ini terdiri dari 13 citra dengan tipe preparasi *ThinPrep/ LBP*. *Dataset* yang digunakan untuk sampel merupakan koleksi data citra tes *Pap-Smear* dari NCI *Bethesda System* yang dapat diunduh dari Internet (<http://nih.techriver.net/>). Detail *dataset* yang digunakan dalam penelitian ini terdapat pada tabel 9.1.

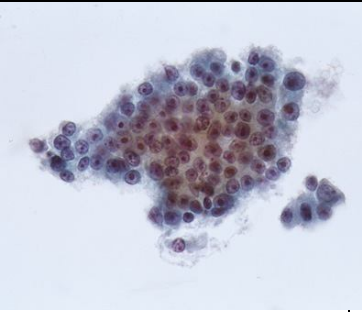
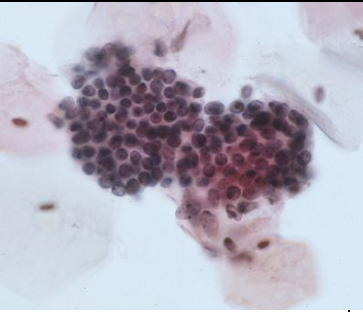
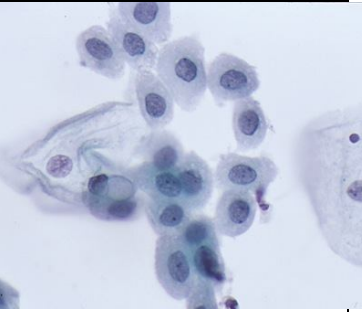
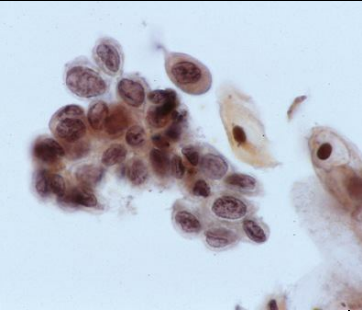
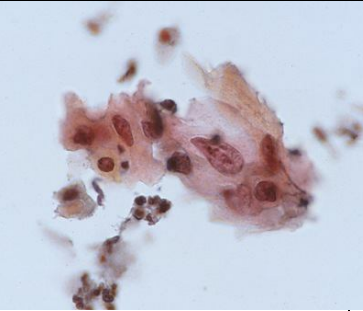
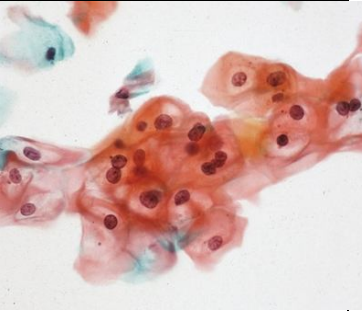
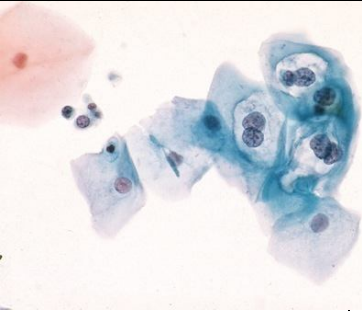
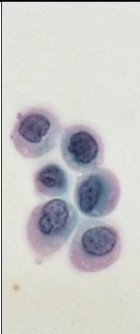
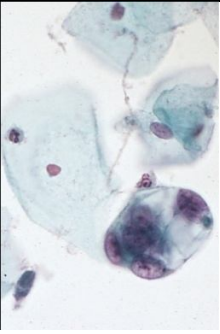
Tabel 9.1. Tabel detail *dataset* yang digunakan

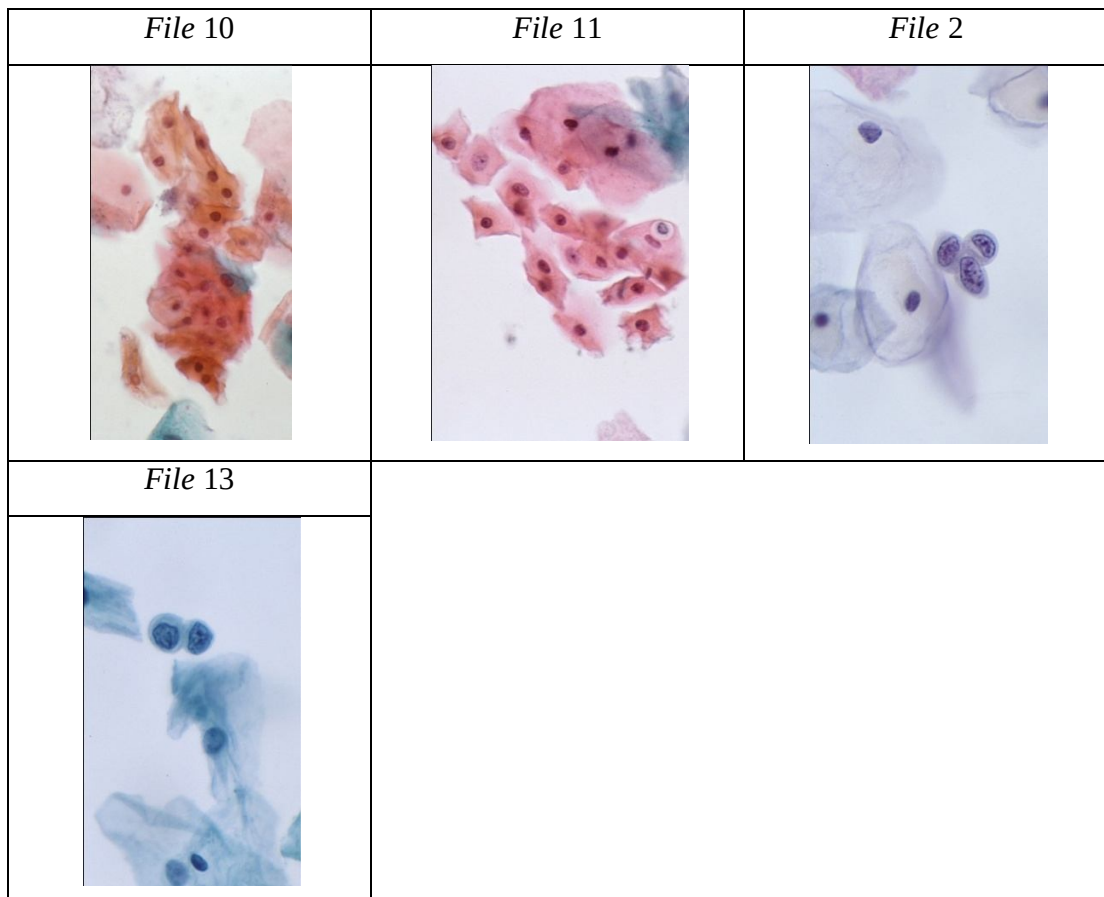
<i>File</i>	<i>Interpretation</i>	<i>Magnification</i>	<i>Clinical History</i>	<i>Size</i>
1	<i>Adenocarcinoma, Endocervical</i>	<i>Medium</i>	<i>39 year old female, Day 12 of cycle</i>	2000x1600
2	<i>Atypical endometrial cells</i>	<i>Medium</i>	<i>52 year old female on estrogenic hormone therapy.</i>	2000x1600
3	<i>NILM: Squamous metaplasia</i>	<i>Medium</i>	<i>Routine Pap Test, 28 year old female</i>	2000x1600
4	<i>HSIL</i>	<i>High</i>	<i>31 year old</i>	2000x1600
5	<i>HSIL, r/o invasion</i>	<i>High</i>	<i>42 year old woman</i>	2000x1600
6	<i>LSIL</i>	<i>Medium</i>	<i>43 year old man, HIV+. Anal-rectal sample</i>	2000x1600

7	<i>LSIL</i>	<i>High</i>	<i>32 year old, LMP: 20 days, routine screening</i>	2000x1600
8	<i>HSIL</i>	<i>High</i>	<i>27 year old woman</i>	672x1600
9	<i>Atypical endometrial cells</i>	<i>High</i>	<i>63 year old woman with postmenopausal bleeding</i>	1035x1550
10	<i>NILM: Keratotic cellular changes:"typical parakeratosis"</i>	<i>Medium</i>	<i>33 year old woman, routine smear</i>	864x1600
11	<i>NILM: Keratotic cellular changes:"typical parakeratosis"</i>	<i>Medium</i>	<i>33 year old woman, routine smear</i>	1107x1600
12	<i>HSIL</i>	<i>High</i>	<i>29 year old, from "high-risk" clinic</i>	1017x1600
13	<i>HSIL</i>	<i>High</i>	<i>29 year old, from "high-risk" clinic</i>	955x1600

Tabel 9.1 menunjukkan bahwa, citra yang digunakan terdiri dari 13 citra dari 13 *slide* preparat berbeda yang terdiri dari 6 perbesaran *medium* dan 7 perbesaran *high*. Ukuran perbesaran lensa dan ukuran *file* yang digunakan pada penelitian ini sangat beragam untuk menguji ketahanan model yang diusulkan terhadap berbagai perbesaran lensa dan ukuran *file* yang berbeda. Citra yang digunakan juga memiliki intensitas warna yang sangat beragam. *Dataset* yang digunakan mengandung 375 buah sel *nuclei* setelah dilakukan penghitungan manual oleh Dokter spesialis Patologi Anatomi. *Dataset* yang digunakan dalam penelitian ini terdapat pada Gambar 9.1.

Citra yang digunakan sebagai *dataset* pada penelitian ini adalah citra dengan format JPG.

File 1	File 2	File 3
		
File 4	File 5	File 6
		
File 7	File 8	File 9
		



Gambar 9.1 *Dataset* yang digunakan

Koleksi citra *Pap smear* pada <http://nih.techriver.net/> dengan preparasi *ThinPrep/ LBP* hanya terdapat 24 citra yang memiliki ukuran yang cukup besar untuk dilakukan pengujian. Citra *Pap smear* yang digunakan pada penelitian ini hanya 13 dari 24 citra yang tersedia karena 11 citra yang lain dianggap susah dalam pembacaan manual menurut seorang Dokter Patologi Anatomi, sehingga sulit untuk membandingkan hasilnya dengan hasil dari model yang diusulkan. Beberapa hal yang menyebabkan proses pembacaan manual menjadi sulit adalah proses pewarnaan yang kurang baik dalam proses persiapan *slide*. Proses pewarnaan yang kurang baik ini menyebabkan *nuclei* pada citra tidak terlalu jelas. Tiga belas citra yang digunakan

sudah dianggap cukup karena masing-masing citra sudah mewakili untuk tiap kasus yang sering ditemukan. Masing-masing citra memiliki karakteristik yang berbeda-beda baik dari ukuran, perbesaran lensa, kasus-kasus, maupun proses pewarnaan.

9.2 Pembahasan Waktu Eksekusi Model yang Diusulkan

Tahap ini berisi tentang pembahasan waktu eksekusi dari model yang diusulkan. Tabel 9.2 menunjukkan rata-rata waktu yang dibutuhkan untuk mengeksekusi model yang diusulkan pada penelitian ini. Masing-masing tahap dilakukan dengan *software* MATLAB R2008a dengan menggunakan komputer Pentium 2.66 GHz dan RAM 4 GB terhadap 13 citra pada *dataset* yang digunakan.

Tabel 9.2. Tabel *Execution Time*

Tahap pada model yang diusulkan	Rata-rata waktu yang dibutuhkan (detik)	Standard Deviasi
<i>Preprocessing</i>	3,65	1,40
Deteksi <i>Region Minima</i>	12,75	5,41
Deteksi Tepi <i>Nuclei</i>	7,81	3,17
Perbaikan Posisi <i>Centroids</i> Kandidat Sel <i>Nuclei</i>	3,18	1,35
Pengurangan Jumlah <i>Centroids</i> Menggunakan Kriteria Jarak	2,17	3,03

Tahap pertama yang dilakukan pada penelitian ini adalah *Preprocessing*. Tahap ini merupakan tahap yang cepat dengan rata-rata waktu yang dibutuhkan hanya 3,65 detik. Waktu yang dibutuhkan paling lama dari *dataset* yang digunakan pada tahap ini adalah 5,1 detik, sedangkan waktu yang tercepat adalah 1,6 detik. Perbedaan waktu eksekusi ini dipengaruhi adanya ukuran *file* yang berbeda dari *dataset* yang digunakan dan intensitas warna yang beragam, semakin *homogen* intensitas warna

pada citra dan semakin besar ukuran *file* pada *dataset* yang digunakan semakin besar waktu yang diperlukan untuk mengkomputasi citra.

Tahap selanjutnya adalah deteksi *centroids* kandidat sel *nuclei*. Tahap ini dibagi menjadi 2 sub tahap yaitu, deteksi *region minima* dan deteksi tepi *nuclei*. Rata-rata waktu yang diperlukan untuk proses komputasi pada sub tahap deteksi *region minima* adalah 12,75 detik. Waktu yang dibutuhkan paling lama dari *dataset* yang digunakan pada sub tahap ini adalah 20,5 detik, sedangkan waktu yang tercepat adalah 5,2 detik. Sub tahap ini adalah tahap yang paling lama dalam keseluruhan proses dari model yang diusulkan, hal ini karena komputasi pada sub tahap ini relatif lebih panjang daripada tahap-tahap yang lainnya.

Tahap selanjutnya masih merupakan tahap deteksi *centroids* kandidat sel *nuclei*, yaitu sub tahap deteksi tepi *nuclei*. Rata-rata waktu yang diperlukan untuk proses komputasi pada sub tahap ini adalah 7,81 detik. Waktu yang dibutuhkan paling lama dari *dataset* yang digunakan pada sub tahap ini adalah 11,3 detik, sedangkan waktu yang tercepat adalah 3,3 detik.

Perbaikan Posisi *Centroids* Kandidat Sel *Nuclei* merupakan tahap berikutnya pada penelitian ini. Rata-rata waktu yang diperlukan untuk proses komputasi pada tahap ini adalah 3,18 detik. Tahap ini termasuk tahap yang memiliki rata-rata waktu komputasi yang cepat karena tahap ini memiliki proses yang sedikit. Waktu yang dibutuhkan paling lama dari *dataset* yang digunakan pada tahap ini adalah 4,9 detik, sedangkan waktu yang tercepat adalah 1,4 detik.

Waktu rata-rata yang diperlukan untuk proses komputasi pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak yang selanjutnya dilakukan adalah 2,17 detik. Tahap pengurangan jumlah *centroids* menggunakan kriteria jarak adalah tahap yang paling cepat rata-rata waktu komputasinya dari keseluruhan tahap pada penelitian ini. Waktu yang dibutuhkan paling lama dari

dataset yang digunakan pada tahap ini adalah 8,9 detik, sedangkan waktu yang tercepat adalah 0,2 detik. Tahap ini memiliki waktu eksekusi yang sangat bervariasi, terlihat dari standard deviasi yang melebihi nilai rata-rata waktu pada tahap ini. Kecepatan waktu komputasi pada tahap ini tergantung dari jumlah *centroids* yang ditemukan. Semakin banyak *centroids* yang ditemukan pada tahap perbaikan posisi *centroids* kandidat sel *nuclei* maka waktu eksekusinya menjadi makin lama. Tabel 9.3 adalah tabel detail dari waktu yang dibutuhkan pada tiap tahap model yang diusulkan per *dataset* yang digunakan.

Tabel 9.3 Tabel detail *execution time*

<i>File</i>	I (detik)	II (detik)	III (detik)	IV (detik)	V (detik)
1	5,0	18,0	11,3	4,9	8,5
2	5,0	20,5	11,2	4,9	8,9
3	4,7	15,4	9,9	3,9	0,8
4	4,9	18,1	11,2	4,4	3,4
5	4,8	16,8	10,4	4,2	1,9
6	5,1	16,5	10,0	4,1	1,8
7	4,7	16,5	10,0	4,0	1,1
8	1,6	5,2	3,3	1,4	0,3
9	2,3	7,8	5,1	2,0	0,4
10	2,1	7,1	4,3	1,7	0,3
11	2,7	8,9	5,5	2,2	0,4
12	2,4	7,6	4,8	1,9	0,2
13	2,2	7,3	4,6	1,8	0,3

Keterangan Tabel:

- I. Tahap *Preprocessing*
- II. Sub Tahap Deteksi *Region Minima*
- III. Sub Tahap Deteksi Tepi *Nuclei*
- IV. Tahap Perbaikan Posisi *Centroids* Kandidat Sel *Nuclei*
- V. Pengurangan Jumlah *Centroids* Menggunakan Kriteria Jarak

9.3 Rangkuman Parameter

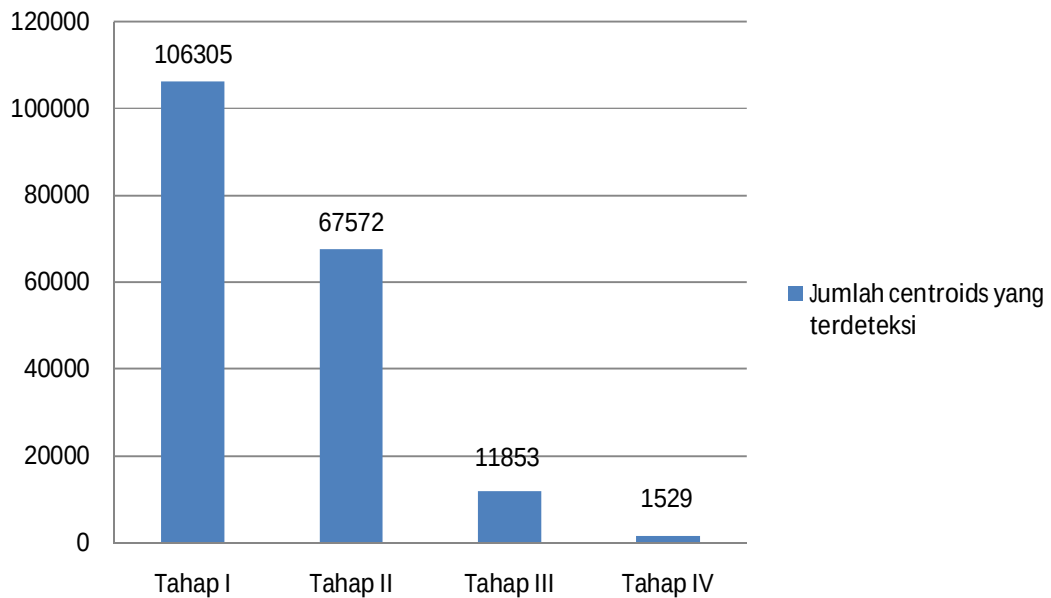
Model yang diusulkan ini sepenuhnya otomatis dan penerapannya dilakukan tanpa melakukan inputan apapun. Tabel 9.4 adalah rangkuman nilai dari parameter yang digunakan dalam penelitian ini. Parameter yang digunakan dalam penelitian ini sangat tergantung dengan *dataset*. Perlu investigasi lanjut mengenai perlu atau tidaknya pengujian ulang terhadap parameter yang digunakan apabila *dataset* berbeda dari *dataset* pada penelitian ini.

Tabel 9.4 Tabel Nilai pada setiap parameter dari model yang diusulkan

Tahap dari Model	Parameter	Nilai
<i>Preprocessing</i>	Nilai ambang area	500
Deteksi <i>Region Minima</i>	Intensitas ambang rendah (<i>h</i>)	35
Pengurangan Jumlah <i>Centroids</i> Menggunakan Kriteria Jarak	Nilai ambang jarak (<i>T</i>)	30
<i>Clustering</i>	Jumlah <i>cluster</i>	2

9.4 Pembahasan Pengurangan *Centroids* pada Tiap Tahap Model yang Diusulkan

Sub bab ini berisi tentang pembahasan pengurangan jumlah *centroids* yang ditemukan pada tiap tahap di penelitian ini. Gambar 9.2 adalah Jumlah *centroids* yang ditemukan pada tiap tahap dari penelitian yang dilakukan. Penilaian yang dilakukan oleh Dokter spesialis Patologi Anatomi, *dataset* yang digunakan mengandung 375 buah sel *nuclei*.



Gambar 9.2 Grafik jumlah *centroids* yang terdeteksi pada tiap tahap

Keterangan Gambar:

- I. Sub Tahap Deteksi *Region Minima*
- II. Sub Tahap Deteksi Tepi *Nuclei*
- III. Tahap Perbaikan Posisi *Centroids* Kandidat Sel *Nuclei*
- IV. Pengurangan Jumlah *Centroids* Menggunakan Kriteria Jarak

Jumlah *centroids* yang ditemukan pada sub tahap deteksi *region minima* sangat banyak, dengan total *centroids* sebanyak 106.305. Nilai rata-rata *true nuclei* yang terdeteksi pada sub tahap ini adalah 100%, artinya tidak ada *true* positif *centroids* yang hilang pada tahap ini. Sub tahap deteksi tepi *nuclei* mampu mereduksi jumlah *centroids* sebanyak 36,44% sehingga pada sub tahap deteksi tepi *nuclei centroids* yang ditemukan menjadi 67.572. Nilai rata-rata *true nuclei* yang terdeteksi pada sub tahap ini adalah 100%, artinya pada sub tahap ini tidak ada *true* positif *centroids* yang hilang.

Tahap selanjutnya adalah tahap perbaikan posisi *centroids* kandidat sel *nuclei*. Jumlah *centroids* pada sub tahap deteksi tepi *nuclei* direduksi secara signifikan pada tahap ini sebanyak 82,46%, sehingga *centroids* yang terdeteksi pada tahap ini adalah sebanyak 11.853. Tahap ini mampu mereduksi jumlah *centroids* secara signifikan namun memiliki nilai rata-rata *true nuclei* yang terdeteksi sebesar 100%, artinya sampai tahap ini belum ada *true positif centroids* yang hilang. Tahap selanjutnya yaitu tahap pengurangan jumlah *centroids* menggunakan kriteria jarak mampu mereduksi jumlah *centroids* secara signifikan sebanyak 87,10%, sehingga pada tahap ini *centroids* yang ditemukan adalah sebanyak 1529 *centroids*. Nilai rata-rata *true nuclei* yang terdeteksi pada tahap ini adalah sebesar 95% karena pada tahap ini terdapat *true positive centroids* yang hilang sebanyak 17.

Tabel 9.5 Tabel jumlah *centroids* pada tiap tahap

<i>FILE</i>	Tahap I	Tahap II	Tahap III	Tahap IV
1	13505	9718	2088	190
2	22828	8943	2033	201
3	7934	5586	708	122
4	15202	10233	1400	179
5	9943	6474	1110	174
6	7273	5511	1829	144
7	8497	5394	748	134
8	2306	1575	277	45
9	3143	2623	263	57
10	4813	3679	565	100
11	4514	3677	463	93
12	2266	1630	158	40
13	4081	2529	211	50
	106305	67572	11853	1529

Keterangan Tabel:

- I. Sub Tahap Deteksi *Region Minima*
- II. Sub Tahap Deteksi Tepi *Nuclei*
- III. Tahap Perbaikan Posisi *Centroids* Kandidat Sel *Nuclei*
- IV. Pengurangan Jumlah *Centroids* Menggunakan Kriteria Jarak

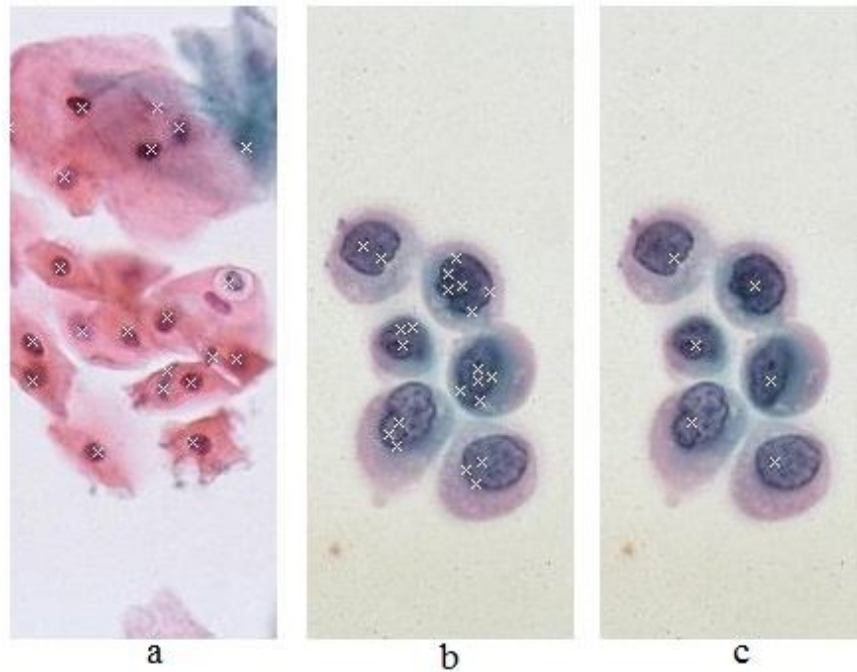
Tabel 9.5 adalah detail jumlah *centroids* pada tiap tahap model yang diusulkan terhadap *dataset* yang digunakan. Perbedaan jumlah *centroids* yang ditemukan dari *dataset* yang digunakan pada sub tahap deteksi *region minima*, sangat dipengaruhi oleh tingkat kontras dari intensitas warna pada *dataset*. Tabel 9.5 menunjukkan semakin besar jumlah *centroids* pada tahap sub tahap deteksi *region minima* maka semakin rendah tingkat kotras dari citra yang digunakan sebagai *dataset*, selain itu ukuran *file/* dimensi citra juga sangat berpengaruh pada besarnya *centroids* yang ditemukan.

Tingkat kontras citra sangat mempengaruhi jumlah *centroids* yang ditemukan pada sub tahap deteksi tepi *nuclei* maupun sub tahap deteksi *region minima*, semakin kontrasnya rendah semakin susah menentukan tepi *nuclei* sehingga *centroids* yang tereduksi semakin sedikit. Ukuran *file/* dimensi citra juga sangat berpengaruh pada besarnya *centroids* yang ditemukan. Besarnya jumlah *centroids* yang ditemukan pada tahap perbaikan posisi *centroids* kandidat sel *nuclei* dan pengurangan jumlah *centroids* menggunakan kriteria jarak sangat tergantung dari jarak *centroids* yang satu dengan *centroids* yang lain. Kumpulan *centroids* akan tereduksi sedikit apabila kumpulan *centroids* tersebut memiliki jarak yang jauh antar *centroids*-nya, namun semakin *centroids* memiliki jarak yang dekat dengan *centroids* yang lain maka kumpulan *centroids* akan tereduksi banyak.

9.5 Pembahasan Perbesaran Lensa

Dataset yang digunakan pada penelitian ini terdapat dua jenis perbesaran lensa, yaitu *high* dan *medium*. Perbesaran lensa ini mempengaruhi diameter *nuclei* pada citra. *Centroids* yang tepat di dalam *nuclei* sudah minim (mendekati 1) pada *dataset* dengan perbesaran *high* karena diameter *nuclei*-nya kecil, namun untuk perbesaran *medium* *centroids* yang terdapat di dalam *nuclei* masih banyak. Banyaknya *centroids* yang terdapat di dalam *nuclei* ini dikarenakan diameter *nuclei* pada *dataset* dengan perbesaran *high* lebih besar daripada diameter *nuclei* pada citra dengan perbesaran *medium*. Pengujian yang dilakukan terhadap parameter jarak masih *global*, artinya, pengujian jarak pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak dilakukan terhadap citra dengan perbesaran *medium* dan *high*.

Gambar 9.3(a) menunjukkan bahwa citra dengan perbesaran *medium* tersebut memiliki diameter *nuclei* yang kecil, sehingga *centroids* yang ditemukan pada *nuclei* tepat 1. Gambar 9.3(b) adalah citra dengan perbesaran *high* dimana diameter *nuclei*-nya lebih besar daripada Gambar 9.3(a). Gambar 9.3(b) menunjukkan bahwa *centroids* yang ditemukan dalam *nuclei* masih banyak karena diameternya lebih besar. Masih banyaknya *centroids* yang ditemukan dalam *nuclei* menyebabkan perlu adanya penambahan tahap pada citra dengan perbesaran *high* ini. Tahap pengurangan jumlah *centroids* menggunakan kriteria jarak dilakukan sekali lagi untuk memperbaiki *centroids* yang ditemukan pada citra dengan perbesaran *high* dengan nilai jarak yang lebih besar daripada nilai jarak yang semula. Besarnya parameter jarak pada tahap ini adalah 85. Nilai tersebut diambil berdasarkan serangkai percobaan yang terdapat pada tabel 9.6. Gambar 9.3(c) adalah hasil penerapan tahap pengurangan jumlah *centroids* menggunakan kriteria jarak terhadap *centroids* yang ditemukan pada tahap *clustering*.



Gambar 9.3 (a) Citra dengan perbesaran *medium* dan *centroids* yang ditemukan, dan (b) Citra dengan perbesaran *high* dan *centroids* yang ditemukan, dan (c) Hasil pengurangan jumlah *Centroids* menggunakan kriteria jarak terhadap citra (b)

Tabel 9.6 Tabel rata-rata *centroids* per *nuclei*

<i>File</i>	Rata2 (50)	Rata2 (55)	Rata2 (60)	Rata2 (65)	Rata2 (70)	Rata2 (75)	Rata2 (80)	Rata2 (85)	Rata2 (90)	Rata2 (95)	Rata2 (100)
4	3	3	2	2	2	2	2	1	1	1	1
5	2	2	2	2	2	2	2	1	1	1	1
7	2	2	2	2	2	1	1	1	1	1	1
8	2	2	2	2	2	2	1	1	1	1	1
9	2	2	2	2	2	2	1	1	1	1	1
12	2	2	2	2	2	1	1	1	1	1	1
13	2	2	2	2	1	1	1	1	1	1	1
	3	3	2	2	2	2	2	1	1	1	1

Tabel 9.6 menunjukkan bahwa rata-rata jarak yang optimal berada pada nilai 85. Nilai 85 diambil sebagai nilai parameter jarak pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak yang dilakukan setelah tahap *clustering* karena pada jarak ini jumlah *Centroids* pada tiap *Nuclei* tepat bernilai 1.

9.6 Evaluasi Pembacaan Manual

Sub bab ini akan membahas persentase konsistensi proses pembacaan manual yang dilakukan oleh Dokter Spesialis Patologi Anatomi. Evaluasi pembacaan manual dilakukan sebanyak 2 kali yaitu pada tanggal 20 September 2011 dan tanggal 28 September 2011. Tabel 9.7 menunjukkan hasil pembacaan manual oleh Dokter Patologi Anatomi.

Tabel 9.7 Tabel jumlah *nuclei* pada tahap pembacaan manual

<i>File</i>	Jumlah <i>nuclei</i> 20-09-2011	Jumlah <i>nuclei</i> 28-09-2011	Jumlah <i>Nuclei</i> yang konsisten	Jumlah <i>Nuclei</i> terbanyak
1	100	103	100	103
2	124	140	122	140
3	21	21	21	21
4	28	32	28	32
5	9	9	8	9
6	26	25	24	26
7	13	11	11	13
8	6	6	6	6
9	9	9	8	9
10	21	19	19	21
11	20	16	15	20
12	7	7	7	7
13	6	6	6	6
JML	390	404	375	413

Tabel 9.7 menunjukkan bahwa tingkat konsistensi pembacaan manual yang dilakukan Dokter Patologi Anatomi adalah sebesar 91%. Persentase tersebut didapat dari jumlah *nuclei* yang konsisten dibagi dengan jumlah *nuclei* terbanyak dari yang

terdeteksi dikali dengan 100%. Dari tingkat konsistensi pembacaan manual yang dilakukan Dokter Patologi Anatomi tersebut dapat disimpulkan bahwa *dataset* yang digunakan memiliki tingkat kesulitan dalam pembacaan yang cukup tinggi. Model yang diusulkan pada penelitian ini diharapkan mampu mendeteksi area-area yang dianggap sebagai *nuclei* yang konsisten pada tahap pembacaan manual oleh Dokter Patologi Anatomi.

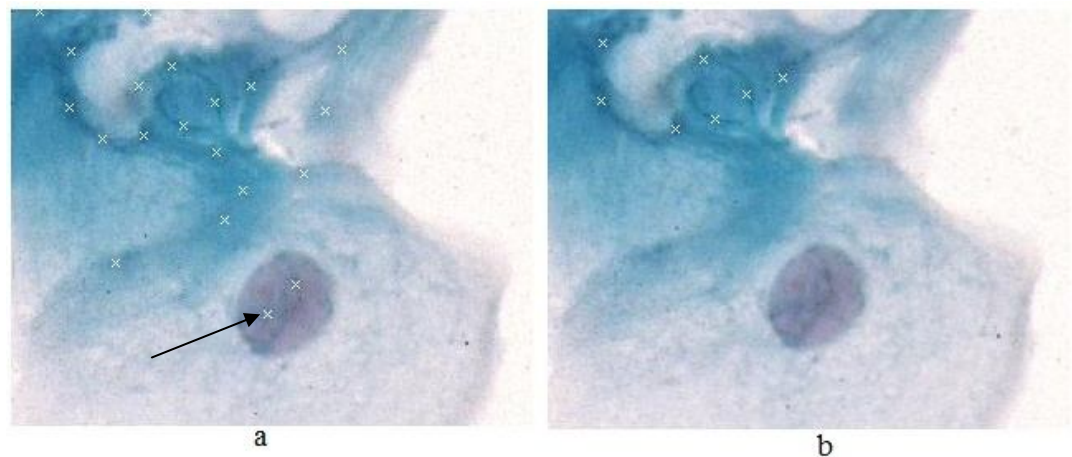
9.7 Evaluasi Hasil dari Model yang Diusulkan

Sub tahap ini akan membahas hasil evaluasi model yang diusulkan. Hasil dari evaluasi terhadap model yang diusulkan dan telah diuji oleh seorang Dokter Patologi Anatomi terdapat pada tabel 9.8.

Tabel 9.8 Tabel evaluasi hasil dari model yang diusulkan

<i>File</i>	<i>Jumlah Nuclei</i>	<i>True Positive</i>	<i>True Negative</i>	<i>False Positive</i>	<i>False Negative</i>	<i>Sensitivity</i>	<i>Specificity</i>
1	100	96	72	28	4	96%	72%
2	122	108	122	0	14	89%	100%
3	21	21	21	4	0	100%	84%
4	28	25	31	2	3	89%	94%
5	8	8	25	16	0	100%	61%
6	24	23	14	12	1	96%	54%
7	11	5	13	7	6	45%	65%
8	6	6	6	0	0	100%	100%
9	8	5	12	4	3	63%	75%
10	19	18	11	11	1	95%	50%
11	15	14	19	4	1	93%	83%
12	7	7	7	0	0	100%	100%
13	6	6	6	2	0	100%	75%
JML	375	342	359	90	33	91.20%	79.96%

Tabel 9.8 menunjukkan bahwa pada *file* 2, 4, 7, dan 9 memiliki tingkat *sensitivity* kurang dari 90%. Tingkat *sensitivity* sangat rendah pada *file* 7 yakni hanya 45%, hal ini dikarenakan proses pewarnaan yang tidak merata pada citra *Pap smear* tersebut sehingga sebagian besar *cytoplasma*-nya memiliki intensitas lebih rendah (lebih gelap) daripada intensitas warna pada *nuclei*. Rendahnya intensitas warna pada area *nuclei file* 7 mengakibatkan *centroids* yang terdapat pada *nuclei* hilang karena proses *clustering*.



Gambar 9.4 (a) Citra dengan *centroids* pada tahap pengurangan jumlah *centroids* dengan kriteria jarak, dan (b) Citra dengan *centroids* pada tahap *clustering*

Tanda panah dalam Gambar 9.4(a) menunjukkan *centroids* yang berada di *nuclei* pada tahap pengurangan jumlah *centroids* dengan kriteria jarak. Gambar 9.4(b) menunjukkan pada tahap *clustering centroids* tersebut hilang karena *cytoplasma*-nya memiliki intensitas yang jauh lebih rendah daripada *nuclei*-nya. Permasalahan pada *file* 7 hampir serupa dengan masalah yang terdapat pada *file* 9. Intensitas warna pada *nuclei* pada *file* 9 cenderung tidak *homogen* ditambah dengan terdapatnya sel radang yang cenderung lebih rendah intensitas warnanya sehingga pada tahap *clustering centroids* yang sudah tepat berada pada area *nuclei* menjadi hilang.

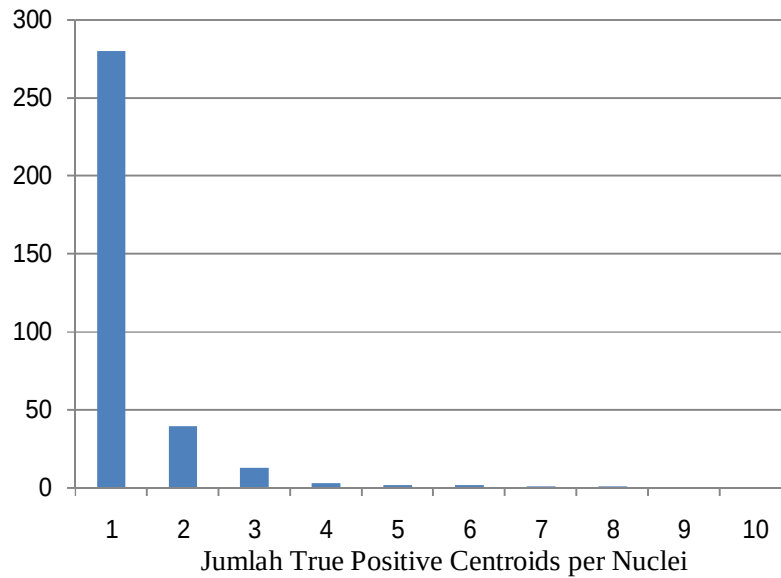


Gambar 9.5 Citra dengan *centroids* pada tahap *clustering*

Gambar 9.5 menunjukkan bahwa *centroids* yang bertahan adalah yang berada pada sel radang (tanda panah) sedangkan *centroids* pada *nuclei*-nya hilang karena intensitas warna pada *nuclei*-nya lebih tinggi daripada sel radang. Intensitas warna *nuclei* pada *file* 7 dan 9 cenderung tidak *homogen* dan cenderung tinggi sehingga kedua *file* tersebut memiliki *false negative* yang cukup banyak.

Tabel 9.8 menunjukkan bahwa model yang diusulkan memiliki tingkat *sensitivity* sebesar 91,20% dan *specificity* sebesar 79,96%. Evaluasi yang dilakukan tersebut menunjukkan bahwa model yang diusulkan memiliki tingkat kinerja sebesar 85%.

Penelitian ini masih menemukan jumlah *True Positive Centroids* yang lebih dari 1 per *nuclei*-nya pada beberapa kasus. Gambar 9.6 adalah grafik yang menunjukkan jumlah *True Positive Centroids* per *nuclei*.



Gambar 9.6 Grafik jumlah *true positive* pada sebuah *nuclei*

Gambar 9.6 menunjukkan bahwa jumlah *true positive centroids* yang tepat satu untuk mewakili satu *nuclei* sudah sebagian besar, namun pada Gambar 9.6 juga masih terdapat beberapa *true positive centroids* yang lebih dari satu untuk mewakili satu *nuclei*. Terdapatnya beberapa *true positive centroids* yang lebih dari satu untuk mewakili satu *nuclei* dikarenakan perbesaran lensa pada citra yang digunakan sebagai *dataset* sangat beragam. Tabel 9.9 adalah detail dari jumlah *true positive centroids* per *nuclei*.

Tabel 9.9 Tabel detail jumlah *true positive* pada sebuah *nuclei*

<i>File</i>	1	2	3	4	5	6	7	8	9	10
1	93	-	3	-	-	-	-	-	-	-
2	82	23	3	-	-	-	-	-	-	-
3	3	3	7	2	2	2	1	1	-	-
4	25	-	-	-	-	-	-	-	-	-
5	7	1	-	-	-	-	-	-	-	-
6	14	8	-	1	-	-	-	-	-	-
7	5	-	-	-	-	-	-	-	-	-
8	6	-	-	-	-	-	-	-	-	-
9	5	-	-	-	-	-	-	-	-	-
10	14	4	-	-	-	-	-	-	-	-
11	13	1	-	-	-	-	-	-	-	-
12	7	-	-	-	-	-	-	-	-	-
13	6	-	-	-	-	-	-	-	-	-
	280	40	13	3	2	2	1	1	-	-

9.8 Evaluasi Penerapan Model pada Tipe Preparasi yang Lain

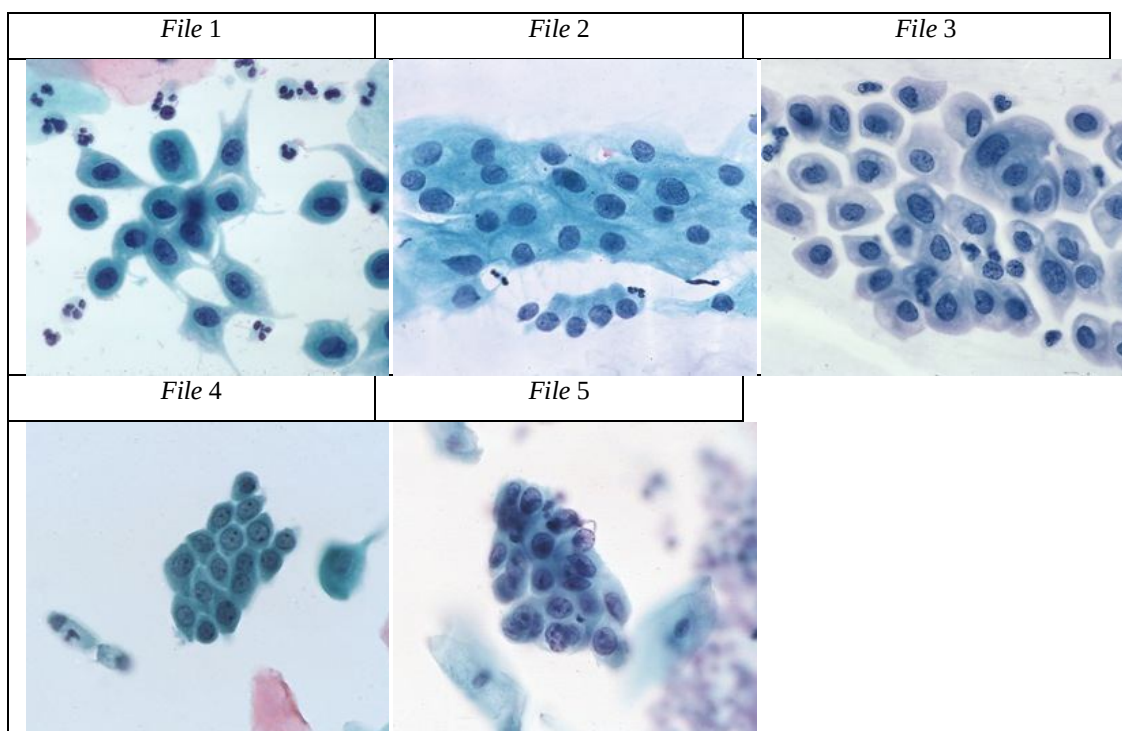
Penelitian ini melakukan percobaan apabila diterapkan pada tipe preparasi yang lain, model ini diujikan pada 5 citra dengan tipe preparasi lain yang terdiri dari 2 citra dengan preparasi *Surepath/ LBP* dan 3 citra dengan preparasi konvensional. Tabel 9.10 berikut adalah *dataset* yang digunakan sebagai percobaan penerapan model yang diusulkan terhadap tipe preparasi yang lain.

Tabel 9.10. Tabel detail *dataset* yang digunakan pada tahap percobaan

<i>File</i>	<i>Interpretation</i>	<i>Magnification</i>	<i>Clinical History</i>	<i>Type of Preparation</i>
1	<i>NILM: Squamous metaplasia</i>	<i>High</i>	<i>27 year old woman, routine exam, LMP 18 days</i>	Conventional
2	<i>NILM vs ASC-US</i>	<i>High</i>	<i>Premenopausal woman</i>	Conventional

3	<i>ASC-H</i>	<i>Medium</i>	-	Conventional
4	<i>NILM: Endocervical cells</i>	<i>High</i>	<i>32 year old woman</i>	SurePath/ LBP
5	<i>Adenocarcinoma, endocervical</i>	<i>High</i>	<i>46 year old woman</i>	SurePath/ LBP

Seluruh citra yang digunakan juga termasuk koleksi citra dari NCI *Bethesda System* yang dapat diunduh dari Internet (<http://nih.techriver.net/>) dengan *size* 2000x1600. *Dataset* yang digunakan dalam tahap ujicoba ini dapat dilihat pada Gambar 9.7.



Gambar 9.7 *Dataset* yang digunakan

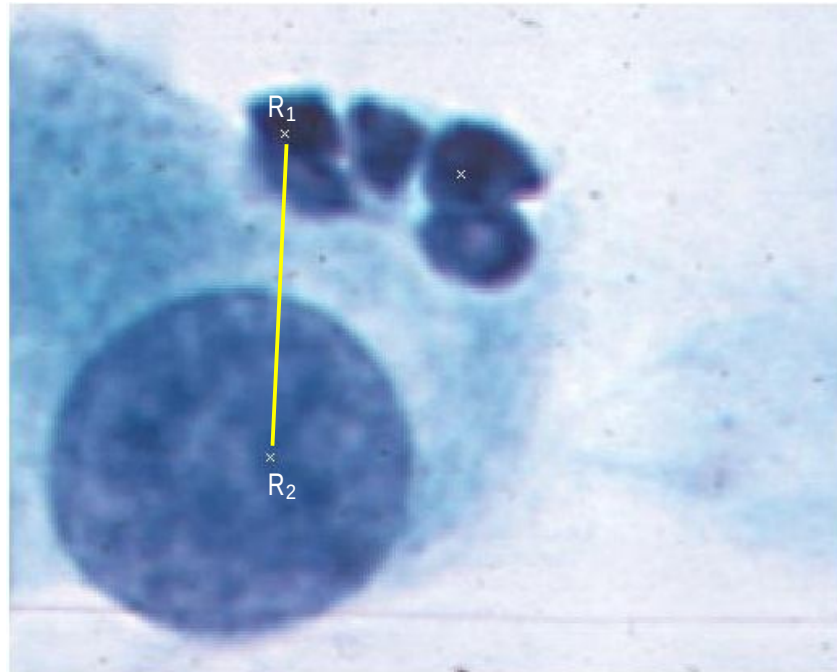
Hasil evaluasi citra *Pap smear* dengan tipe preparasi konvensional dan *Surepath* yang telah diujikan terhadap seorang Dokter Patologi Anatomi dari model yang diusulkan terdapat pada tabel 9.11.

Tabel 9.11 Tabel hasil percobaan

<i>File</i>	Jumlah <i>Nuclei</i>	<i>True Positive</i>	<i>True Negative</i>	<i>False Positive</i>	<i>False Negative</i>	<i>Sensitivity</i>	<i>Specificity</i>
1	21	21	44	14	0	100%	76%
2	31	29	34	6	2	94%	85%
3	40	39	40	12	1	98%	77%
4	21	19	21	0	2	90%	100%
5	23	17	29	3	6	74%	90%

Model yang diusulkan mampu mendeteksi semua *nuclei* yang terdapat pada *file* 1 sehingga tingkat sensitivitasnya adalah 100%. *False positive* pada *file* 1 terdapat sebanyak 14, hal ini dikarenakan banyak terdapatnya sel radang pada citra tersebut. Sel radang memiliki intensitas warna yang cenderung lebih gelap daripada *nuclei* sehingga sel radang memiliki kemungkinan terdeteksi sebagai *nuclei*. Tingkat spesifikasi pada *file* 1 adalah 76%.

Model yang diusulkan mampu mendeteksi sebagian besar *nuclei* yang terdapat pada *file* 2 dengan tingkat sensitivitasnya sebesar 94%. *File* 2 mengandung *False Negative* sebanyak dua yang artinya terdapat dua *nuclei* yang tidak terdeteksi. *Nuclei* yang tidak terdeteksi ini diakibatkan karena memiliki intensitas yang cukup tinggi sehingga pada tahap *clustering nuclei* tersebut dianggap bukan sebagai *nuclei*. Posisi *nuclei* yang berdekatan dengan sel radang juga menjadi alasan tidak terdeteksinya *nuclei* tersebut karena karakteristik sel radang yang memiliki intensitas warna yang sangat rendah. *Centroids* yang terletak pada *nuclei* menjadi hilang karena *centroids* tersebut akan dibandingkan dengan *centroids* pada sel radang dengan intensitas yg lebih rendah pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak. *File* 2 memiliki tingkat spesifikasinya adalah sebesar 85%.



Gambar 9.8 Jarak antar *nuclei* yang di set tinggi pada *file 2*

Gambar 9.8 menunjukkan sebuah *nuclei* dan beberapa sel radang. Sel radang memiliki intensitas yang cenderung lebih rendah daripada *nucleinya*. *Centroids* yang sudah tepat pada *nuclei* (R_2) memiliki kemungkinan menjadi hilang ditahap pengurangan jumlah *centroids* menggunakan kriteria jarak pada citra yang memiliki *nuclei* yang berdekatan dengan sel radang. Gambar 9.8 mengilustrasikan bahwa jika terdapat *centroids* yang saling berdekatan maka akan diambil yang intensitasnya rendah (diantara R_1 dan R_2). *Centroids* pada R_2 akan hilang dan *centroids* pada R_1 akan bertahan karena *centroids* pada *nuclei* lebih rendah intensitas warnanya daripada intensitas warna pada sel radang.

Nuclei yang tidak terdeteksi pada *file 3* adalah sebanyak 1 karena posisi *nuclei* itu terdapat pada sudut citra sehingga bentuk *nuclei* tersebut terpotong menjadi sebagian kecil yang tampak. Tahap ini memiliki sensitivitas yang cukup tinggi yaitu 98%. *False Positif* pada *file* ini juga masih terdapat sebanyak 12 hal ini dikarenakan

terdapat banyak sel radang pada citra tersebut. Persentase spesifikasi pada *file 3* ini adalah 77%.

Hasil dari ketiga citra *Pap smear* konvensional tersebut dapat disimpulkan bahwa sel radang dapat membuat kesalahan yang cukup signifikan dalam proses perhitungan jumlah *nuclei*. Citra konvensional yang digunakan pada *dataset* penelitian ini memiliki sel radang yang cukup tinggi, akan tetapi model yang diusulkan belum terdapat pemodelan untuk menangani masalah akibat terdapatnya sel radang. Jumlah sel radang yang cukup tinggi ini menyebabkan perlu adanya penelitian lebih lanjut untuk memodelkan model yang lebih baik dan mampu membedakan sel radang dengan sel *nuclei* pada citra.

File 4 dan *5* menunjukkan bahwa *nuclei* yang ada cenderung menggerombol dan tumpang tindih. *Nuclei* yang cenderung tumpang tindih ini sangat mempengaruhi tingkat sensitifitas hasil model yang diusulkan karena pada citra yang memiliki *nuclei* yang menggerombol tentu nilai parameter jarak pada tahap pengurangan jumlah *centroids* menggunakan kriteria jarak seharusnya lebih kecil daripada citra *Pap smear* dengan *nuclei* yang merata. Tingkat sensitifitas pada *file 4* dan *5* model yang diusulkan cukup rendah yakni hanya memiliki 90% dan 74%. Sensifitas dari model yang diusulkan dapat lebih tinggi jika parameter jaraknya diset lebih rendah, namun ada kemungkinan spesifikasinya menjadi lebih kecil. Tingkat spesifikasi pada *file 4* dan *5* adalah 100% dan 90%.

BAB X

KESIMPULAN DAN SARAN

10.1 Kesimpulan

Berdasarkan hasil ujicoba yang diperoleh dalam penelitian ini, maka dapat disimpulkan beberapa hal berikut:

1. Tingkat *sensitivity* model yang diusulkan adalah sebesar 91,20%.
2. Tingkat *specifity* model yang diusulkan adalah sebesar 79,96%.
3. Tingkat *system Performance* model yang diusulkan adalah sebesar 85%.

10.2 Saran

Berdasarkan hasil ujicoba yang diperoleh dalam penelitian ini, maka untuk penelitian selanjutnya diperlukan hal-hal berikut:

1. Perlu adanya tahapan model yang lebih baik lagi dengan adanya tahap pengenalan sel radang.
2. Perlu adanya pengujian yang lebih lanjut mengenai *thresholding* menggunakan metode Ridler yang pada tahap pengujian penelitian ini terbukti lebih cepat waktu eksekusinya.
3. Perlu adanya pengujian kriteria yang lain sebagai kriteria pada tahap *Clustering*.
4. Perlu adanya penelitian penggunaan metode *clustering* yang lain.
5. Perlu adanya standar perbesaran lensa yang digunakan.
6. Perlu adanya penelitian parameter apabila menggunakan *dataset* yang berbeda dengan *dataset* yang digunakan pada penelitian ini.

DAFTAR PUSTAKA

- Arbyn, M., Bergeron, C., Klinkhamer, P., Hirsch, P. M., Siebers, A. G., Bulten, J. (2008). "Liquid Compared With Conventional Cervical Cytology: A Systematic Review and Meta-analysis," *American College of Obstetricians and Gynecologist.*, vol. 111, no. 1.
- Aziz, M.F., Andrijono, Saifuddin, A.B. (2006). *Buku acuan nasional onkologi ginekologi*. Jakarta: Yayasan Bina Pustaka Sarwono Prawirohardjo.
- Boardman, C. H, Haris, J. E. (2011). TNM and FIGO Classifications for Cervical Cancer. Diakses tanggal: 21 Juni 2011.
- <http://emedicine.medscape.com/article/2006486-overview>
- Castleman K.R. (1996). *Digital image processing*. New Jersey: Prentice Hall.
- Crum, C.P., Lester, S.C., Cotran, R.S. (2007). Sistem Genitalia Perempuan dan Payudara. In: Hartanto, H., et al., ed. *Buku Ajar Patologi* (vol. 2), 7th ed. Jakarta: EGC. 767-770.
- Dalimartha, S. (2004). *Deteksi Dini Kanker & Simulasi Antikanker*. Jakarta: Penebar Swadaya.
- Diananda, R. (2009). *Mengenal Seluk-Beluk Kanker*. Yogyakarta: Katahari.
- Dong, L., Yu, G., Ogunbona, P., Li, W. (2008). "An efficient iterative algorithm for image thresholding", *Pattern Recognition Letters* 29 (2008) 1311–1316.
- Feig, R.L., Johnson, N. C., Stead, L. (2001). *First Aid For The Obstetrics & Gynecology Clerkship*. US: McGraw-Hill.

Forsyth, D.A., Ponce, J. (2003). *Computer vision: A modern approach*. New Jersey: Prentice Hall.

Humas Universitas Indonesia. (2007). *Ilmu Patologi Anatomi Kurang Diminati* . Diakses tanggal: 19 Juni 2011.

http://www.ui.ac.id/download/kliping/200407/Ilmu_Patologi_Anatomi_Kurang_Diminati.pdf

Indman, P.D. (2000). *Abnormal Pap*. California. Diakses tanggal: 19 Juni 2011. <http://www.gynalternatives.com/abnpap.htm>

Karnon J, Peters, J., Platt, J., Chilcott, J., McGoogan, E., Brewer, N. (2004). Liquid-based cytology in cervical screening: an updated rapid and systematic review and economic analysis. *Health Technology Assessment*; Vol. 8: No. 20. Diakses tanggal: 07 Juli 2011

<http://www.hta.ac.uk/fullmono/mon820.pdf>

Kusuma F, Moegni E. M. (2001). Penatalaksanaan Tes Pap Abnormal. *Cermin Dunia Kedokteran*; 133:19-22. Diakses tanggal: 07 Juli 2011.

http://www.kalbe.co.id/files/cdk/files/08_PenatalaksanaanTesPapAbnormal.pdf/08_PenatalaksanaanTesPapAbnormal.pdf

Kusumadewi, S. Hartati, S., Wardoyo, R., dan Harjoko, A. (2006). *Fuzzy Multi-Attribute Decision Making*. Yogyakarta: Graha Ilmu.

Manuaba, I.B.G. (2005). *Dasar-Dasar Teknik Operasi Ginekologi*. Jakarta: EGC.

Marquardt, N., 2002. Cervical Neoplasma and Carcinoma. In: Marquardt, N., ed. *Obstetrics and Gynecology*, 4th ed. Philadelphia: Lippincott Williams & Wilkins, 547-565.

Nayar, R., Solomon, D. (2001). diakses *tanggal akses 19 Juni 2011*.

<http://nih.techriver.net>

Nuranna, L. (2001). Skrining Kanker Serviks dengan Metode Skrining Alternatif: IVA. *Cermin Dunia Kedokteran*, No 133, hal.22-25. Diakses tanggal: 07 Juli 2011.

http://www.kalbefarma.com/files/cdk/files/cdk_133_obstetri_dan_ginekologi.pdf

Otsu, Noboyuki. (1979). A Threshold Selection Method from gray Level histogram, *IEEE Transantions on Systems, MAN Cybernetics*, Vol. 9, issue 1, page.62 – 66

Pal, N. R., Pal, S. K. (1993). A review on image segmentation techniques. *Pattern Recognition* 26 (9), 1277–1294.

Panneton, B. (2010). Gray image thresholding using the Triangle Method, Diakses tanggal 15-9-2011.

<http://www.mathworks.com/matlabcentral/fileexchange/28047-gray-image-thresholding-using-the-triangle-method>

Plissiti, M.E., Nikou, C., Charchanti, 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, page.233 – 241

Purwoto, G. & Nuranna, L. (2002). Metode Skrining Alternatif pada Kanker Serviks. In: Ramli, H.M., et al., eds. *Deteksi Dini Kanker*. Jakarta: Balai Penerbit FKUI, 142-143.

Ramachandram, D. (2003). Automatic Thresholding, Diakses tanggal 15-9-2011.

<http://www.mathworks.com/matlabcentral/fileexchange/3195>

Rasjidi, I., Irwanto, Y., Sulistyanto, H. (2008). *Manual Prakanker Serviks*. Jakarta: Sagung Seto.

Rasjidi, I., Sulistiyanto, H. (2007). *Vaksin HPV dan eradikasi kanker mulut rahim*. Jakarta: Sagung Seto.

Ridler T.W., Calvard, S. (1978). Picture thresholding using an iterative selection method, *IEEE Trans. System, Man and Cybernetics*, SMC-8 (1978) 630-632.

Riono, Y. (1999). Dept of Surgery Holywood Hospital. Diakses tanggal: 20 Juni 2011.

http://1zkie.student.umm.ac.id/download-as-doc/student_blog_article_27.doc

Rujikietgumjorn, S. (2008). Segmentation methods for multiple body parts, Project in lieu of Thesis, University of Tennessee, Knoxville.

Saviano, E.C. (1993). Papanicolaou Smear & Cervical Intraepithelial Neoplasia. In: Brown, J.S., Crombleholme, W.R., eds. *Handbook of Gynecology & Obstetrics*. Stamford: Appleton & Lange, 36-40.

Sjamsuddin. S. (2001). Pencegahan dan deteksi dini kanker serviks. *Cermin Dunia Kedokteran*, No 133, hal.8-13

Sloane, E. (2003). *Anatomi dan fisiologi untuk pemula*. Jakarta. EGC.

Sutoyo, T., Mulyanto, E., Suhartono, V., Nurhayanti, O. D., Wijanarto. (2009). *Teori Pengolahan Citra Digital*. Yogyakarta: Penerbit ANDI.

- Suwiyoga, I.K. (2006). Tes Human Papilloma virus sebagai skrining alternative kanker serviks. *Cermin Dunia Kedokteran*, No 151, hal.29-32
- Tambunan, G. W. (1995). *Sepuluh jenis kanker terbanyak di Indonesia: diagnosis dan tatalaksana*. Jakarta: EGC
- Van Bommel, J.H., Musen. M.A. (1997). *Handbook of Medical Informatics*. Diegem: Bohn Stafleu Van Loghum, 240 – 244.
- Vincent, L. (1993). “Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms,” *IEEE Trans. Image Process.*, vol. 2, no. 2, pp. 176–201, Apr. 1993.
- Zack, G. W., Rogers, W. E. and Latt, S. A. (1977). Automatic Measurement of Sister Chromatid Exchange Frequency, *Journal of Histochemistry and Cytochemistry* 25 (7), pp. 741-753.
- Zuiderveld, K. (1994). “Contrast limited adaptive histogram equalization,” in *Graphics Gems IV*. San Diego, CA: Academic, 1994, pp. 474–485.

