

## Penggabungan Fitur Bentuk dan Fitur Tekstur yang *Invariant* terhadap Rotasi untuk Klasifikasi Citra *Pap Smear*

Yuwanda Purnamasari Pasrun<sup>1</sup>, Chastine Fatichah<sup>2</sup>, Nanik Suciati<sup>3</sup>

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember Surabaya

Kampus ITS Keputih, Sukolilo, Surabaya 60111, Jawa Timur

E-mail: <sup>1</sup>yuwanda13@mhs.if.its.ac.id, <sup>2</sup>chastine@if.its.ac.id, <sup>3</sup>nanik@if.its.ac.id

Masuk: 3 Juli 2015; Direvisi: 19 Juli 2015; Diterima: 27 Juli 2015

**Abstract.** *Pap test is a cervical cancer screening manually and requires a long time that it needs an exact cell classification system based computers. Features determination by observation in characteristic differences between the datasets visually between class will help a cell classification results which has relevant characteristics between classes. In addition, the change in orientation of the cells at the time of the acquisition will affect the value of the generated feature so extraction method that is rotation invariant is needed to overcome that problem. This research proposes the combination of simple shapes feature and the texture feature from extraction Local Binary Pattern Histogram Fourier (LBP-HF) that invariant to rotation as additional features to classify pap smear images. The result show that the proposed feature combination yield good performance with accuracy 92.44% for two category cell and 70.06% for seven class cell.*

**Keywords:** *classification, lbp-hf, pap smear image, shape feature.*

**Abstrak.** *Pap test adalah pemeriksaan kanker serviks secara manual yang membutuhkan waktu yang lama sehingga dibutuhkan sistem klasifikasi sel berbasis komputer yang tepat. Penentuan fitur melalui observasi pada perbedaan ciri antarkelas secara visual pada dataset akan membantu hasil klasifikasi sel untuk mendapatkan ciri yang relevan antarkelas. Selain itu, adanya perubahan orientasi sel pada saat akuisisi akan mempengaruhi nilai fitur yang dihasilkan sehingga dibutuhkan metode ekstraksi fitur yang invariant terhadap rotasi. Penelitian ini mengusulkan penggabungan fitur bentuk sederhana dan fitur tekstur dengan ekstraksi fitur Local Binary Pattern –Histogram Fourier yang invariant terhadap rotasi sebagai ciri tambahan dalam mengklasifikasikan citra pap smear. Hasilnya menunjukkan bahwa kombinasi fitur menghasilkan performa yang baik dengan akurasi 92,44% untuk dua kategori sel dan 70,06% untuk tujuh kelas sel.*

**Kata Kunci:** *klasifikasi, lbp-hf, citra pap smear, fitur bentuk.*

### 1. Pendahuluan

Kanker serviks terjadi karena adanya perkembangan sel-sel abnormal secara cepat dan tidak terkendali pada leher rahim. *Pap test* adalah langkah pertama dalam pemeriksaan kanker serviks (Njoroge, 2006). Pemeriksaan ini terdiri dari pengumpulan sampel sel dari daerah tertentu pada serviks. Sampel sel kemudian diamati melalui mikroskop dan didiagnosis oleh ahli sel secara manual (Erik, 2003). Klasifikasi dan diagnosis manual ini membutuhkan waktu yang lama sehingga dibutuhkan sistem klasifikasi sel yang tepat berbasis komputer.

Klasifikasi citra sel oleh ahli patologis dilakukan dengan mengamati fitur sel seperti warna sel, ukuran nukleus, sitoplasma, dan area sel. Kedua fitur ini direpresentasikan dalam fitur intensitas dan fitur bentuk. Fitur tersebut dapat mengindikasikan jika terdapat ciri abnormal pada sel serviks, karena sel terinfeksi *Human Papilloma Virus* (HPV) yang bisa berkembang menjadi tahap kanker akan menunjukkan perubahan biologis tertentu seperti ukuran nukleus yang membesar dan warna nukleus sel abnormal yang lebih gelap dibandingkan dengan sel normal (Yusoff, 2010). Selain fitur bentuk dan intensitas, fitur tekstur dapat direpresentasikan sebagai ciri sel untuk mengukur pola *chromatin* sel dalam citra *pap smear*. Adanya perubahan orientasi sel pada saat akuisisi akan mempengaruhi nilai fitur yang dihasilkan (Rodenacker, 2003). Tidak semua fitur bentuk dapat menghasilkan fitur yang sama jika citra sel yang sama mengalami perubahan posisi sehingga dibutuhkan metode ekstraksi fitur yang

*invariant* terhadap rotasi. *Dataset Herlev* adalah salah satu *dataset* citra *pap smear* satu sel (sel tunggal). Fitur intensitas dianggap kurang relevan sebagai fitur karakteristik kelas untuk *Dataset Herlev* karena tidak ada standarisasi pada metode pewarnaan dan zat warna (*dye/pap stain*) citra *pap smear* (Jantzen, 2006).

Chankong (2014) melakukan ekstraksi fitur bentuk sel berdasarkan deskriptor *regional*. Fitur bentuk dapat diaplikasikan jika ada kriteria yang dapat membedakan bentuk sel setiap kelas dan kriteria tersebut tidak dipengaruhi oleh orientasi sel pada saat akuisisi. Deskriptor *regional* seperti area dan perimeter adalah deskriptor *regional* sederhana yang tidak berpengaruh pada orientasi objek (Jahne, 1997). Sehingga untuk fitur bentuk akan digunakan fitur deskriptor *regional* serta seperti area, *compactness*, *major axis length*, *minor axis length*, aspek rasio lebar dan tinggi nukleus serta aspek rasio area nukleus dan sel. Akan tetapi pada *Dataset Herlev*, ada beberapa citra sel dari satu kelas yang termasuk pada kelas normal memiliki bentuk area yang mirip dengan kelas abnormal. Meskipun kemiripan tersebut hanya terdapat sebagian kecil pada citra sel dalam satu kelas normal, hal tersebut dapat mempengaruhi hasil klasifikasi. Sehingga perlu ada kombinasi dengan fitur lain yaitu fitur tekstur.

Klasifikasi citra *pap smear* yang hanya menggunakan fitur tekstur telah dilakukan dengan menggunakan ekstraksi fitur varian *Local Binary Pattern* (LBP) (Nanni, 2010). Hasilnya sudah dapat membedakan kelas abnormal dan normal, namun akurasinya masih di bawah 90% karena hanya melibatkan satu fitur saja. Hal ini menjadi dasar dalam mengombinasikan fitur bentuk dan tekstur untuk membedakan ciri antar kelas sel agar lebih spesifik. Plissiti (2011) menggunakan fitur bentuk dan fitur tekstur dengan perhitungan statistik histogram *uniform rotation invariant* LBP pada daerah nukleus yang digabungkan dengan fitur bentuk dan fitur intensitas untuk klasifikasi citra *pap smear*. *Uniform Rotation-Invariant* LBP merupakan operator hasil pengembangan LBP yang mampu mendeskripsikan tekstur yang *invariant* terhadap rotasi (Ojala, 2002) dengan cara membuat representasi tetap dengan mentransformasikan setiap deskriptor lokal yang *invariant* terhadap rotasi dan mendeskripsikan keseluruhan gambar dari hasil perhitungan deskriptor lokal sehingga metode ini dapat menghilangkan distribusi pola relatif (Ahonen, 2009). Chankong (2014) mengekstrak fitur bentuk sebagai ciri utama sel dan menambahkan perhitungan homogenitas melalui perhitungan *Gray Level Co-Occurrence Matrix* (GLCM), namun perhitungan homogenitas pada GLCM tidak *invariant* terhadap rotasi. Ahonen (2009) mengajukan *Local Binary Pattern Histogram Fourier* (LBP-HF) yang mempertahankan informasi distribusi relatif dari pola *uniform* LBP. LBP-HF menghitung histogram yang merepresentasikan keseluruhan *region* dan perhitungan *invariant* tidak harus diimplementasikan pada setiap lokasi piksel.

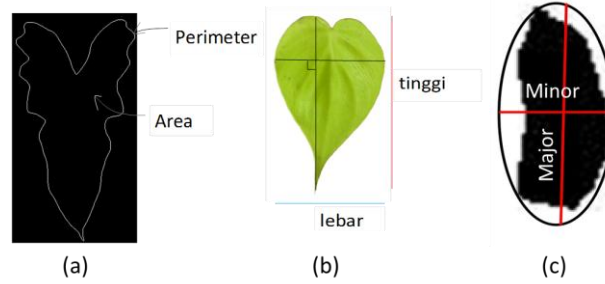
Penelitian ini mengusulkan penggabungan fitur bentuk berdasarkan perhitungan deskriptor *regional* bentuk sederhana dan fitur tekstur dan fitur LBP-HF yang *invariant* terhadap rotasi sebagai ciri tambahan dalam mengklasifikasikan citra *pap smear* sehingga menghasilkan fitur yang dapat merepresentasikan karakteristik dari masing-masing kelas meskipun dalam posisi *input* citra sel yang berbeda.

## 2. Tinjauan Pustaka

### 2.1. Ekstraksi Fitur Bentuk

Fitur bentuk diperoleh melalui perhitungan deskriptor bentuk *regional* sederhana yaitu perhitungan area, perimeter untuk *compactness*, *major axis length*, *minor axis length*, dan aspek rasio lebar dan tinggi. Pada citra biner digital, area didefinisikan sebagai jumlah piksel yang berada di dalam *region*. Area juga didefinisikan sebagai luas dari suatu objek (Kadir, 2013). Perimeter atau keliling menyatakan panjang tepi suatu objek. *Compactness* menyatakan perbandingan perimeter kuadrat dan area (Gonzales, 2002). *Major axis length* didefinisikan sebagai panjang (dalam piksel) *major axis* dari *ellips* yang mempunyai *second moment* yang sama sebagai *region*. Sedangkan *Minor axis length* didefinisikan sebagai panjang (dalam piksel) *minor axis* dari *ellips* yang mempunyai *second moment* yang sama sebagai *region* (Chankong, 2014).

Untuk memperoleh nilai aspek rasio lebar dan tinggi maka perlu dihitung besar diameter. Diameter adalah jarak terpanjang antara dua titik dalam tepi objek. Hal itu dapat dihitung dengan menggunakan metode *Brute Force* (Costa dan Cesar, 2001). Berdasarkan diameter, lebar objek dapat diperoleh. Lebar adalah garis terpanjang yang menghubungkan dua piksel di tepi objek yang tegak lurus terhadap panjang maksimum pada objek. Setelah dua titik dengan jarak terpanjang diperoleh, gradien garis yang melalui kedua piksel tersebut dihitung (Kadir & Susanto, 2013). Ilustrasi fitur bentuk ditunjukkan pada Gambar 1.



**Gambar 1. (a) Ilustrasi Perimeter dan Luas (b) Ilustrasi Lebar dan Tinggi (Kadir & Susanto, 2013) (c) Ilustrasi Major dan Minor Axis Length**

## 2.2. Ekstraksi Fitur Tekstur

*Local Binary Pattern Histogram Fourier* (LBP-HF) deskriptor dibentuk dengan menghitung histogram *uniform* LBP citra yang *non-invariant* terhadap rotasi dan membentuk fitur *invariant* dari histogram. Konsep asli *Local Binary Pattern* (LBP) adalah metode analisis tekstur yang menggunakan kombinasi model statistika dan struktur (Hamouchene & Hadjer, 2014). Pola spasial lokal didapatkan melalui perhitungan setiap piksel yang dibentuk dari hasil pengurangan delapan tetangganya. Berikut Persamaan (1) menunjukkan persamaan operator LBP dengan  $g_p$  adalah nilai piksel ketetanggaan,  $g_c$  adalah nilai piksel tengah serta Persamaan (2),  $s(z)$  yang menunjukkan fungsi *thresholding* (Ahonen, 2009).

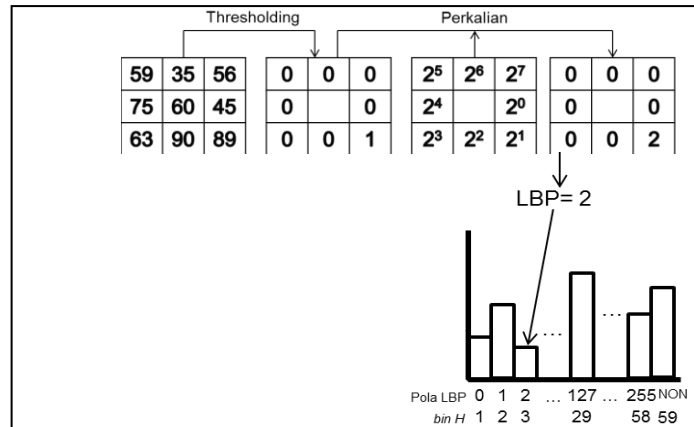
Setiap piksel tetangga dilabeli nilai 1 atau 0 menggunakan Persamaan 2. Jika nilai tetangganya lebih besar dari nilai tengahnya maka akan dilabeli 1 dan jika sebaliknya akan diberi nilai 0 (fungsi *thresholding*). Sehingga *threshold* ketetanggaan merepresentasikan kode biner yang menggantikan nilai tengah. Kode biner ini akan dikonversi ke angka desimal dengan cara mengalikan nilai biner dengan bobotnya dan nilai tengah tadi direpresentasikan dalam bentuk histogram.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (1)$$

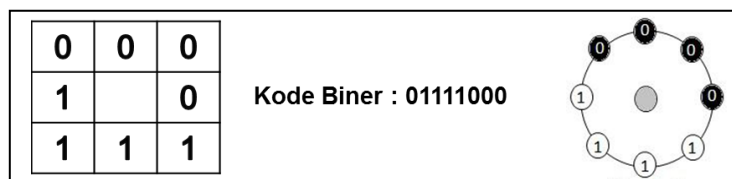
$$s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases} \quad (2)$$

Operator asli LBP dikembangkan menjadi *Uniform* LBP (Ojala, 2002). Jika pada LBP fitur yang dihasilkan berjumlah 256-*bin* histogram, maka fitur yang dihasilkan dari *Uniform* LBP adalah 59-*bin* histogram. Ilustrasi perhitungan LBP dan pembentukan histogram *uniform* LBP ditunjukkan pada Gambar 2.

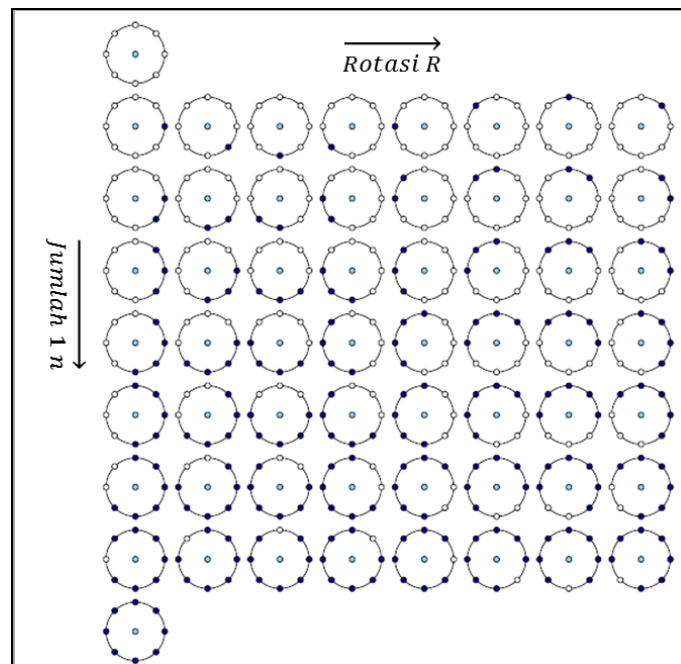
LBP *Uniform* hanya akan mengambil fitur yang memiliki nilai  $U$  maksimal 2. Maksud dari nilai  $U$  maksimal 2 ( $U \geq 2$ ) adalah bahwa pergantian pola bit 0/1 atau 1/0 pada deret kode hanya terjadi maksimal dua kali. Misalnya pada kode 00000000 pergantian piksel 0/1 atau 1/0 tidak terjadi sehingga nilai  $U$  untuk kode tersebut adalah 0. Namun pada deret kode 01111000 terjadi dua kali pergantian bit 0/1 atau 1/0 sehingga nilai  $U$  dari kode tersebut adalah 2. Contoh kode biner yang termasuk pola *uniform* dapat dilihat pada Gambar 3.



Gambar 2. Ilustrasi Perhitungan LBP dan Pembentukan Histogram *Uniform* LBP



Gambar 3. Kode Biner yang Termasuk Pola *Uniform*



Gambar 4. Pola LBP *Uniform* dengan  $(P,R)=(8,1)$  (Ahonen, 2009).

Pada histogram, pola *uniform* akan diletakkan pada *bin* yang berbeda dan pola yang tidak *uniform* akan digabung pada *bin* terakhir (*bin* 59). Pola LBP *Uniform* tertentu didefinisikan dengan  $U_p(n,r)$ . Pasangan  $(n,r)$  menunjukkan pola yang *uniform* sehingga  $n$  adalah jumlah 1-bit dalam pola (ditunjukkan nomor baris pada Gambar 4) dan  $r$  adalah rotasi pola (nomor kolom pada Gambar 4).

Berdasarkan properti yang menyatakan bahwa rotasi menginduksi pergeseran representasi polar  $(P,R)$  dari ketetangaan, diusulkan kelas fitur yang bersifat *invariant* terhadap rotasi gambar masukan. Fitur dihitung sepanjang baris histogram masukan, yang *invariant* terhadap perubahan siklis dengan menggunakan *Discrete Fourier Transform* (DFT).

$H(n,u)$  didefinisikan dari DFT baris ke-  $n$  dari histogram  $h_l(U_p(n,r))$  menggunakan persamaan (3). Fitur vektor akhir dari ekstraksi fitur LBP-HF ditunjukkan pada persamaan (4) merupakan gabungan nilai spektrum DFT, nilai histogram *all ones*, *all zeros* dan nilai histogram *non uniform* (Ahonen, 2009).

$$H(n,u) = \sum_{r=0}^{P-1} h_l(U_p(n,r)) e^{-i2\pi ur/P} \quad (3)$$

$$fv_{LBP-HF} = \left[ \left| H(1,0) \right|, \dots, \left| H\left(1, \frac{P}{2}\right) \right|, \left| H(P-1,0) \right|, \dots, \left| H\left(P-1, \frac{P}{2}\right) \right|, \right. \\ \left. h(U_p(0,0)), h(U_p(P,0)), h(U_p(P+1,0)) \right] \quad (4)$$

### 2.3. Fuzzy k-Nearest Neighbor

Untuk mengklasifikasikan fitur citra *pap smear* yang merupakan hasil penggabungan fitur bentuk dan fitur tekstur dari citra uji dan citra pada *database* digunakan metode *Fuzzy k-Nearest Neighbor* (*Fuzzy k-NN*). Algoritma *Fuzzy k-NN* diperkenalkan oleh Keller, dkk. (1985) dengan mengembangkan *k-NN* yang digabungkan dengan teori *Fuzzy* dalam menyampaikan definisi pemberian label kelas pada data uji yang diprediksi (Keller dkk., 1985). Seperti halnya pada teori *Fuzzy*, sebuah data mempunyai nilai keanggotaan pada setiap kelas, yang artinya sebuah data bisa dimiliki oleh kelas yang berbeda dengan nilai derajat keanggotaan dalam interval  $[0,1]$ . Pengukuran jarak (ketidakmiripan) dua data yang digunakan dalam *Fuzzy k-NN* digeneralisasi dengan persamaan (5) (Keller dkk., 1985) dan nilai keanggotaan ditunjukkan pada persamaan (6) (Li, 2007), dimana  $u(x,c_i)$  adalah nilai keanggotaan data kelas  $x$  ke kelas  $c_i$ ,  $k$  adalah jumlah tetangga terdekat yang digunakan.  $(x_k, c_i)$  adalah nilai keanggotaan data tetangga dalam  $k$  tetangga pada kelas  $c_i$ , nilainya 1 jika data latih  $x_k$  milik kelas  $c_i$  atau 0 jika bukan milik kelas  $c_i$ .  $d(x, x_k)$  adalah jarak dari data  $x$  ke data  $x_k$  dalam  $k$  tetangga terdekat.  $m$  adalah bobot pangkat (*weight exponent*) yang besarnya:  $m > 1$ .

$$d(x_i, x_j) = \left( \sum_{i=1}^N |x_{ij} - x_{ji}|^p \right)^{1/p} \quad (5)$$

$$u(x, c_i) = \frac{\sum_{k=1}^k u(x_k, c_i) * d(x, x_k)^{\frac{-2}{(m-1)}}}{\sum_{k=1}^k d(x, x_k)^{\frac{-2}{(m-1)}}} \quad (6)$$

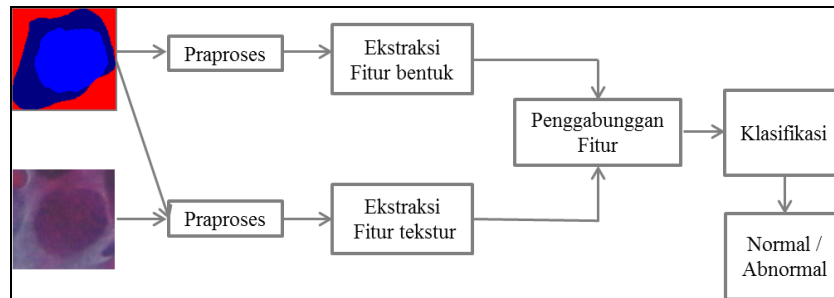
Nilai keanggotaan suatu data pada kelas sangat dipengaruhi oleh jarak data itu ke tetangga terdekatnya. Semakin dekat tetangganya, semakin besar nilai keanggotaan data tersebut pada kelas tetangganya, begitu pula sebaliknya. Algoritma *Fuzzy k-Nearest Neighbor* ditunjukkan pada Kode 1.

#### Kode 1. Algoritma prediksi FKNN (Prasetyo, 2012)

1. Normalisasikan data menggunakan nilai terbesar dan terkecil data pada setiap fitur.
2. Cari  $k$  tetangga terdekat untuk data uji  $x$  menggunakan persamaan 5.
3. Hitung nilai keanggotaan  $u(x, c_i)$  menggunakan persamaan 6 untuk setiap  $i$ , dengan  $1 \leq i \leq c$ .
4. Ambil nilai terbesar  $c = u(x, c_i)$  untuk semua  $1 \leq i \leq c$
5. Berikan label kelas  $c$  ke data uji  $x$ .

### 3. Metodologi Penelitian

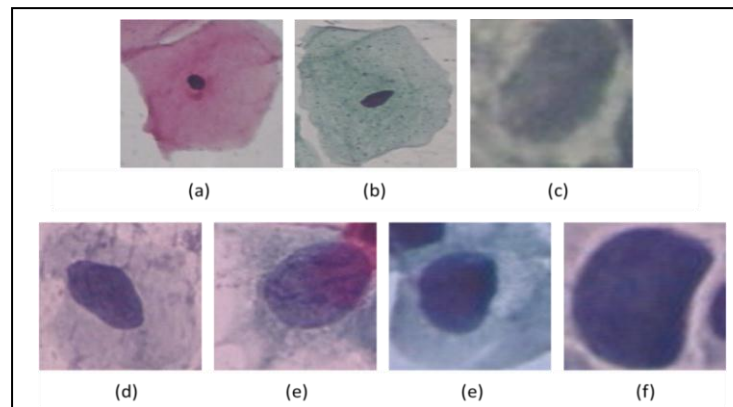
Rancangan sistem klasifikasi citra *pap smear* meliputi praproses, ekstraksi fitur bentuk, ekstraksi fitur tekstur, penggabungan fitur dan tahap klasifikasi. Penelitian ini menggabungkan fitur bentuk berdasarkan deskripsi *regional* dan fitur tekstur hasil ekstraksi *Local Binary Pattern Histogram Fourier* yang menghasilkan fitur yang *invariant* terhadap rotasi untuk klasifikasi citra *pap smear*. Gambaran umum dari rancangan sistem klasifikasi citra *pap smear* yang diusulkan dalam penelitian ini ditunjukkan pada Gambar 5.



Gambar 5. Rancangan Sistem Klasifikasi Citra *Pap Smear*.

#### 3.1. Dataset

*Dataset* yang digunakan merupakan *dataset Herlev* yang terdiri dari citra *pap smear* dalam bentuk tunggal atau satu sel (*single cell*) yang diakuisisi dan diidentifikasi oleh ahli sel terampil (*skilled cyto-technicians*). Gambar 6 menunjukkan contoh sel pada tujuh kelas *dataset Herlev*.



Gambar 6. *Dataset Herlev*: (a) *normal squamous*, (b) *intermediate squamous*, (c) *columnar*, (d) *mild dysplasia*, (e) *moderate dysplasia*, (f) *severe dysplasia*, dan (g) *carcinoma in situ* (Jayasingh & Stephen, 2015).

Setiap citra disegmentasi oleh ahli sel menggunakan *software CHAMP* untuk kepentingan ekstraksi fitur. Segmentasi menggunakan metode paten dengan warna 3D dibentuk untuk membedakan warna *cluster* bagian *background*, sitoplasma dan nukleus (Erik, 2003). Tabel 1 menunjukkan contoh distribusi jumlah sel *dataset Herlev*.

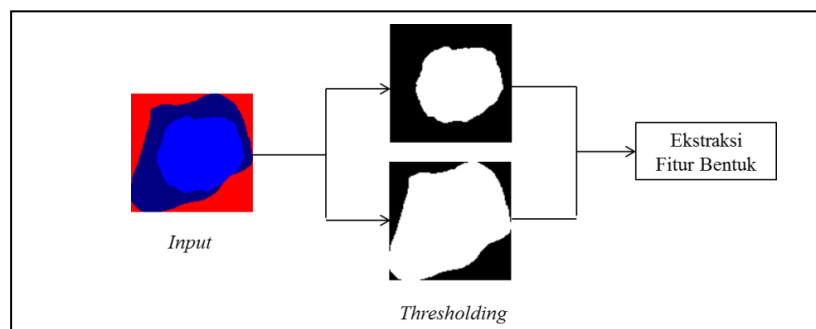
Tabel 1. Distribusi Jumlah Sel *Dataset Herlev*

Kelas	Kategori	Tipe sel	Jumlah Sel	Total
1	Normal	<i>Normal Squamous</i>	74	242
2	Normal	<i>Intermediate Squamous</i>	70	
3	Normal	<i>Columnar</i>	98	
4	Abnormal	<i>Mild Dysplasia</i>	182	675
5	Abnormal	<i>Moderate Dysplasia</i>	146	
6	Abnormal	<i>Severe Dysplasia</i>	197	
7	Abnormal	<i>Carcinoma in situ</i>	150	

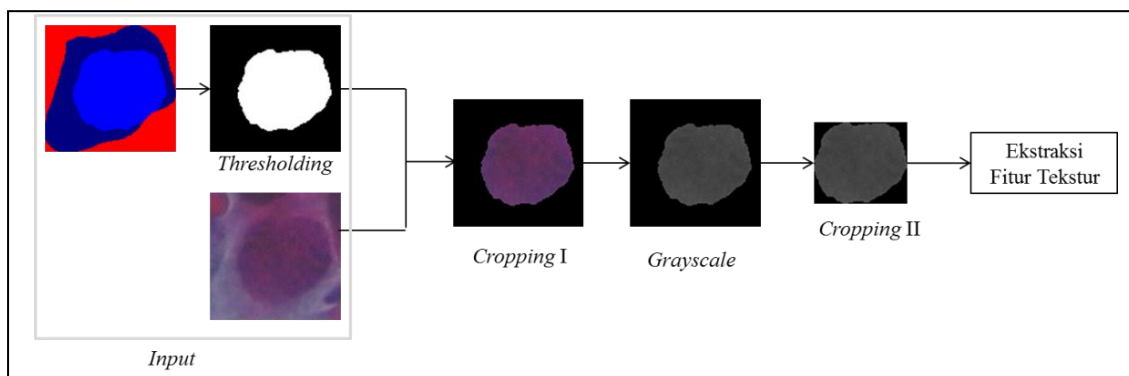
### 3.2. Sistem Klasifikasi Citra *Pap Smear*

Semua citra yang akan dijadikan data uji maupun data masukan awal sistem (*database*) akan melewati tahapan praproses yang sama. Setiap data memiliki dua citra yang berukuran sama. Citra pertama sebagai citra warna *pap smear* dan citra yang kedua adalah citra *ground truth*. Praproses untuk masukan ekstraksi fitur bentuk akan menggunakan citra *ground truth* yang sudah dilakukan proses *thresholding*. Pada citra biner hasil praproses akan dilakukan ekstraksi fitur bentuk berdasarkan perhitungan parameter bentuk sederhana. Fitur bentuk yang dihitung adalah nilai area, *compactness*, aspek rasio lebar dan tinggi, *major axis length*, *minor axis length* pada daerah nukleus sedangkan pada daerah keseluruhan sel akan dihitung area, rasio area antara nukleus dan sel, serta *compactness*. Gambar 7 menunjukkan praproses untuk masukan ekstraksi fitur bentuk.

Praproses untuk masukan fitur tekstur menggunakan citra *ground truth* untuk segmentasi daerah nukleus pada citra warna *pap smear* dan akan dilakukan operasi *grayscale*. Untuk memperoleh fitur tekstur nukleus, *input* citra berupa citra *gray*. Gambar 8 menunjukkan praproses untuk masukan ekstraksi fitur tekstur.



Gambar 7. Praproses Ekstraksi Fitur Bentuk



Gambar 8. Praproses Ekstraksi Fitur Tekstur

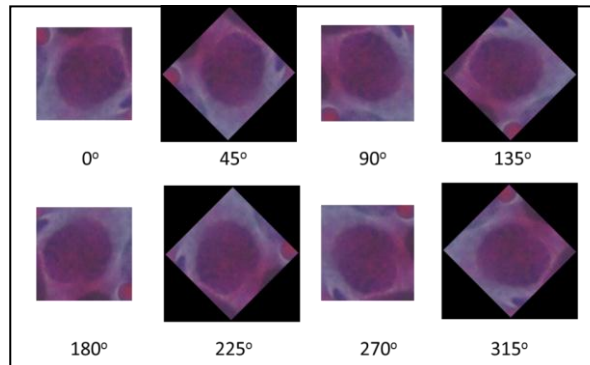
Ekstraksi fitur tekstur menggunakan metode *Local Binary Pattern Histogram Fourier* (LBP-HF) dengan tahapan pertama perhitungan menggunakan operator *Local Binary Pattern* (LBP). Hasil perhitungan LBP dari citra akan disimpan dalam bentuk histogram dengan menggunakan syarat metode *Uniform* LBP. Selanjutnya akan dihitung fitur LBP-HF berupa nilai spektrum dengan *Discrete Fourier Transform* dari baris  $n$  histogram *Uniform* LBP. Fitur akhir LBP-HF merupakan gabungan nilai histogram *all zeros*, *all ones*, *non-uniform* dan nilai spektrum *Fourier*.

Penggabungan fitur dilakukan dengan meletakkan posisi nilai fitur bentuk dan fitur tekstur dalam satu fitur vektor yang sama. Hasil penggabungan fitur ini menjadi masukan untuk proses klasifikasi data uji dan fitur ciri data yang akan disimpan dalam *database*. Klasifikasi citra *pap smear* akan dilakukan dengan cara mencocokkan fitur data uji dan fitur yang disimpan di dalam sistem (*database*) dengan menggunakan *Fuzzy k-Nearest Neighbor* (*Fuzzy k-NN*).

Fitur vektor hasil penggabungan berjumlah 38 yang terdiri dari 8 fitur bentuk dan 38 fitur tekstur.

#### 4. Pengujian dan Pembahasan

Untuk keperluan uji coba, setiap citra akan dirotasi dalam tujuh arah yaitu  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ ,  $180^\circ$ ,  $225^\circ$ ,  $270^\circ$ , dan  $315^\circ$ . Contoh citra uji yang telah dilakukan proses rotasi ditunjukkan pada Gambar 9.



**Gambar 9. Contoh Citra Uji yang Telah Dirotasi**

Pada penelitian ini, data yang digunakan merupakan *dataset Herlev* yang terdiri dari dua kategori dan tujuh kelas. Kategori normal terdiri dari tiga kelas dan kategori abnormal terdiri dari empat kelas. Masing-masing citra dilengkapi dengan citra *ground truth* hasil segmentasi manual oleh ahli sel. Untuk pengujian kategori pada *dataset* diasumsikan sebagai kelas sehingga pengujian yang akan dilakukan adalah klasifikasi dua kelas dan klasifikasi tujuh kelas.

Pengujian yang dilakukan terdiri dari klasifikasi berdasarkan fitur bentuk, fitur tekstur, dan klasifikasi menggunakan penggabungan fitur bentuk dan fitur tekstur. Hal ini bertujuan untuk melihat pengaruh masing-masing jenis fitur. Sehingga kinerja penggabungan dapat dibandingkan dengan fitur yang tidak dilakukan penggabungan. Data yang dijadikan data citra acuan (*database*) dalam klasifikasi berjumlah 917 citra. Sedangkan data uji merupakan data citra acuan yang telah dirotasi ke dalam tujuh arah ( $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ ,  $180^\circ$ ,  $225^\circ$ ,  $270^\circ$ , dan  $315^\circ$ ) dengan total 6419 citra.

Untuk memperoleh parameter  $k$  optimal pada klasifikasi *Fuzzy k-NN* maka akan dilakukan proses klasifikasi pada tujuh kelas fitur. Nilai  $k$  optimal akan diambil dari klasifikasi fitur gabungan dengan akurasi tertinggi dan akan digunakan untuk pengujian dua kelas. Hasil pengujian ditunjukkan pada Tabel 2.

**Tabel 2. Hasil Akurasi Klasifikasi Tujuh Kelas Berdasarkan Fitur Bentuk, Tekstur dan Gabungan**

$k$	Akurasi (%)		
	Bentuk	Tekstur	Bentuk+Tekstur
2	33,99	61,57	70,06
3	36,21	62,32	68,53
4	36,78	63,14	66,40
5	36,81	63,23	66,30
6	36,38	63,36	64,25
7	36,75	63,16	62,72

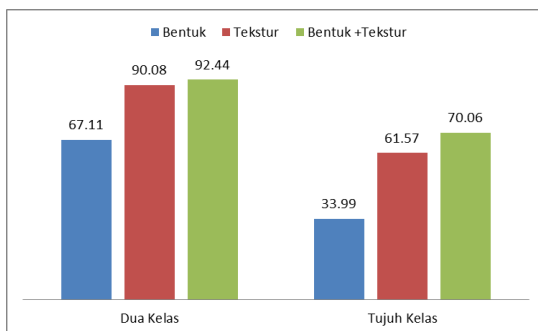
Dari hasil Tabel 2 diperoleh nilai akurasi klasifikasi tujuh kelas pada penggabungan fitur tertinggi (70,06%) diperoleh pada  $k=2$ . Hasilnya akurasi mengalami peningkatan jika dibandingkan dengan klasifikasi menggunakan fitur bentuk atau fitur tekstur saja. Sehingga dilakukan pengujian klasifikasi untuk dua kelas dengan menggunakan parameter klasifikasi *Fuzzy k-NN* yang sama yakni  $k=2$ . Hasil akurasi klasifikasi citra *pap smear* berdasarkan fitur



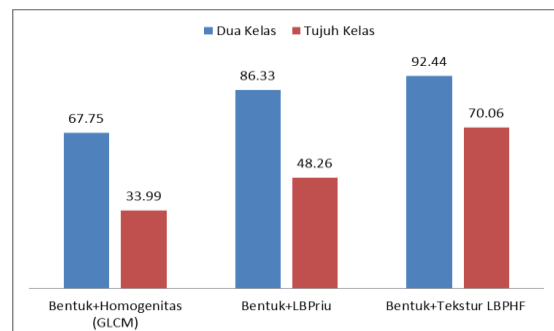
bentuk, fitur tekstur dan fitur gabungan bentuk dan tekstur untuk dua kelas dan tujuh kelas dan ditunjukkan pada Gambar 10.

Gambar 10 menunjukkan bahwa penggabungan fitur bentuk dan fitur tekstur dapat meningkatkan akurasi klasifikasi dibandingkan dengan menggunakan fitur bentuk *regional* atau fitur tekstur LBP-HF saja. Peningkatan ini dapat berlaku untuk klasifikasi dua kelas (berdasarkan kategori) dan tujuh kelas (berdasarkan kelas dalam kategori).

Metode sebelumnya yang telah digunakan dalam mengklasifikasikan citra *pap smear* tanpa memperhatikan adanya keadaan *invariant* terhadap rotasi adalah perhitungan nilai homogenitas metode *gray level co-occurrence matrix* (GLCM) (Chankong, 2014) dan empat nilai fitur GLCM (Walker, 1994) dan metode yang memiliki kemampuan ekstraksi fitur tekstur yang *invariant* terhadap rotasi adalah metode *uniform rotation invariant local binary pattern* (LBP<sup>riu</sup>) (Plissiti, 2011). Metode yang diusulkan dibandingkan dengan fitur tekstur yang digunakan dengan kedua penelitian sebelumnya. Fitur bentuk yang diujikan serupa dengan fitur bentuk usulan. Hasilnya dapat dilihat pada Gambar 11.



Gambar 10. Contoh citra uji yang telah dirotasi



Gambar 11. Contoh citra uji yang telah dirotasi

Perbedaan akurasi berdasarkan dua kelas sebesar 6,11% jika dibandingkan dengan LBP<sup>riu</sup> dan 24,69% jika dibandingkan dari nilai homogenitas metode GLCM. Sedangkan perbedaan akurasi berdasarkan klasifikasi tujuh kelas sebesar 21,8% jika dibandingkan dengan LBP<sup>riu</sup> dan 34% jika dibandingkan dari nilai homogenitas metode GLCM. Hal ini menunjukkan bahwa penggabungan fitur yang diusulkan memiliki kinerja yang cukup baik untuk digunakan pada sistem klasifikasi citra *pap smear*.

## 5. Kesimpulan

Berdasarkan hasil pengujian, hasil penggabungan beberapa fitur bentuk deskriptor *regional* sederhana dan fitur tekstur *Local Binary Pattern Histogram Fourier* (LBP-HF) dapat digunakan sebagai fitur yang *invariant* terhadap rotasi untuk klasifikasi *pap smear* pada *dataset Herlev*. Klasifikasi dapat dilakukan untuk tujuh kelas dan dua kelas citra *pap smear* dengan metode klasifikasi *Fuzzy k-NN* pada parameter  $k=2$ , yaitu dengan akurasi 90,44% dan 70,06%. Penggabungan fitur menghasilkan akurasi yang lebih tinggi jika dibandingkan dengan akurasi klasifikasi berdasarkan fitur bentuk atau fitur tekstur saja.

## Referensi

- Ahonen, T., Matas, J., He, C., dan Pietikäinen, M. 2009. Rotation Invariant Image Description With Local Binary Pattern Histogram Fourier Features. *Springer Berlin Heidelberg. In Image Analysis*, pp. 61-70.
- Chankong, T., Theera-Umpon, N., dan Auephanwiriyakul, S. 2014. Automatic Cervical Cell Segmentation and Classification in Pap Smears. *Computer methods and programs in biomedicine*, 113(2): 539-556.
- Costa, L.F. dan Cesar, R.M. 2001. *Shape Analysis and Classification Theory and Practice*. Florida: CRC Press LLC.

- Erik, M. 2003. *Pap-Smear Classification* (Doctoral dissertation, MSc Thesis, Technical University of Denmark, Oersted-DTU, Automation).
- Gonzales, R. C., & Woods, R. E. 2002. *Digital Image Processing Second Edition*. Beijing: Publishing House of Electronics Industry.
- Hamouchene, I., Aouat, S., & Lacheheb, H. 2014. Texture Segmentation and Matching Using LBP Operator and GLCM Matrix. In *Intelligent Systems for Science and Information*. Springer International Publishing. pp. 389-407.
- Jahne, B. 1997. *Digital image processing: Concepts, Algorithms, and Scientific Applications*. Berlin, Heidelberg: Springer-Verlag.
- Jantzen, J., & Dounias, G. 2006. Analysis of Pap-smear Image Data. In *Nature-Inspired Smart Information Systems 2nd Annual Symposium*.
- Kadir, A., & Susanto, A. 2013. *Pengolahan Citra Teori dan Aplikasi*. ANDI, Yogyakarta.
- Keller, J. M., Gray, M. R., & Givens, J. A. (1985). A Fuzzy K-Nearest Neighbor Algorithm. *Systems, Man and Cybernetics, IEEE Transactions on*, (4), 580-585.
- Li, D., Deogun, J. S., dan Wang, K. (2007, November). Gene Function Classification Using Fuzzy K-Nearest Neighbor Approach. In *Granular Computing*. GRC 2007. IEEE International Conference. pp. 644-644.
- Nanni, L., Lumini, A., & Brahnam, S. 2010. Local Binary Patterns Variants As Texture Descriptors for Medical Image Analysis. *Artificial intelligence in medicine*, 49(2), 117-125.
- Njoroge, E., Alty, S. R., Gani, M. R., & Alkatib, M. 2006. Classification of Cervical Cancer Cells Using FTIR Data. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE. IEEE*. pp. 5338-5341.
- Ojala, T., Pietikainen, M., dan Maenpaa, T. 2002. Multiresolution Gray-Scale and Rotation Invariant Texture Classification With Local Binary Patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7), 971-987
- Plissiti, M. E., Nikou, C., dan Charchanti, A. 2011. Combining Shape, Texture and Intensity Features for Cell Nuclei Extraction in Pap Smear Images. *Pattern Recognition Letters*, 32(6), 838-853.
- Prasetyo, E. 2012. *Data Mining Konsep Dan Aplikasi Menggunakan Matlab*. Yogyakarta, Andi.
- Rodenacker, K., & Bengtsson, E. 2003. A Feature Set for Cytometry on Digitized Microscopic Images. *Analytical Cellular Pathology*, 25(1), 1-36.
- Walker, R. F., Jackway, P., Lovell, B., & Longstaff, I. D. 1994. Classification of Cervical Cell Nuclei Using Morphological Segmentation and Textural Feature Extraction. In *Intelligent Information Systems, 1994. Proceedings of the 1994 Second Australian and New Zealand Conference*. IEEE. ISO 690. pp. 297-301.
- Yusoff, I. A., Isa, N. A. M., Othman, N. H., Sulaiman, S. N., & Jusman, Y. 2010. Performance of Neural Network Architectures: Cascaded MLP Versus Extreme Learning Machine on Cervical Cell Image Classification. In *Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference*. pp. 308-311.