

Klasifikasi Sel Serviks Menggunakan Analisis Fitur Nuclei pada Citra Pap Smear

Rahadian Kurniawan¹, Dinda Eling Kartikaning Sasmito², Fajar Suryani³

¹Jurusan Teknik Informatika, Universitas Islam Indonesia, Yogyakarta

rahadiankurniawan@gmail.com

², ³Magister Teknik Informatika, Universitas Islam Indonesia, Yogyakarta

² dindaelingsasmito@gmail.com

³ fajarsuryani@gmail.com

Abstract. Pada penelitian ini, kami mengevaluasi klasifikasi sel serviks dengan menggunakan analisis fitur yang diekstraksi hanya pada daerah nuclei dan tidak memperhitungkan fitur yang diekstraksi dari daerah sitoplasma. Penelitian ini dilatarbelakangi oleh kenyataan bahwa pada citra Pap Smear yang kompleks nuclei dapat diekstraksi secara otomatis dari citra Pap smear, berbeda dengan segmentasi sitoplasma yang susah diekstrak jika sel tumpang tindih. Kami melakukan seleksi fitur terhadap 25 fitur pada daerah nuclei untuk digunakan sebagai fitur dalam proses klasifikasi menggunakan *Fisher Criterion*. Selanjutnya, kami membandingkan hasil kinerja dua metode klasifikasi yaitu *K-Nearest Neighbor* dan *Backpropagation*. Hasil yang diperoleh menunjukkan bahwa analisis fitur dapat mencapai kinerja klasifikasi yang tinggi ketika kita menggunakan fitur hanya dari daerah nuclei.

Keywords: Kanker Serviks, Pap Smear, Fitur Nuclei, Fisher Criterion, Backpropagation, K-Nearest Neighbor.

1 Pendahuluan

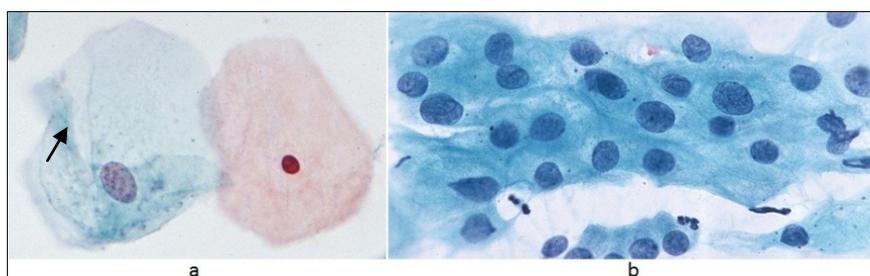
Keakuratan interpretasi tes Pap smear sangat dipengaruhi oleh bentuk umum pada daerah nuclei (inti sel). Nuclei adalah bagian penting dari sel yang menunjukkan perubahan signifikan ketika sel terkena penyakit. Perubahan bentuk secara umum ini merupakan kriteria awal untuk menemukan kelainan yang terjadi pada daerah leher rahim/ serviks oleh Dokter Spesialis Patologi Anatomi. Dokter Spesialis Patologi Anatomi membuat keputusan dengan inspeksi visual dari sejumlah besar sel dan fitur diagnostik seperti rasio ukuran nuclei dengan ukuran sitoplasma atau bentuk nuclei di setiap sel.

Pap Smear merupakan suatu metode pemeriksaan sel-sel yang diambil dari leher rahim dan kemudian diperiksa di bawah mikroskop. Interpretasi visual citra Pap smear secara manual memiliki banyak keterbatasan, membutuhkan waktu yang lama, dan rawan kesalahan prosedur. Oleh karena itu analisis secara otomatis berbantuan komputer sangat dibutuhkan untuk membantu memudahkan proses interpretasi tersebut.

Beberapa tahun terakhir ini, beberapa penelitian telah dilakukan untuk menginterpretasi citra Pap smear antara lain: melakukan deteksi letak nuclei dengan metode *Morphological Reconstruction* dan *Clustering* [1], *Fuzzy C-Means Clustering* [2], dan *Edge detectors* [3], [4]. Deteksi letak nuclei merupakan tahap awal dalam interpretasi citra Pap Smear. Proses selanjutnya adalah melakukan segmentasi pada daerah nuclei yang telah terdeteksi sebelumnya. Segmentasi adalah proses mengambil hanya bagian yang ingin dianalisis dari citra yang digunakan dan membuang daerah-daerah yang tidak diinginkan secara otomatis. Beberapa penelitian tentang segmentasi nuclei antara lain: *Deformable Templates* [5], *Pixel Classification Schemes* [6], *Morphological Operation and Watershed Transformation* [7]. Tahapan yang terakhir adalah melakukan klasifikasi terhadap sel yang tersegmentasi menggunakan fitur-fitur yang telah diekstrak sebelumnya. Beberapa teknik klasifikasi telah diusulkan untuk mengklasifikasikan

sel tertentu ke dalam kelas yang sesuai. Teknik berdasarkan logika fuzzy telah diusulkan [8], [9] untuk mengklasifikasikan sel-sel dalam citra Pap smear. Selanjutnya, Support Vector Machine (SVM) juga digunakan untuk klasifikasi sel Pap smear ke dalam kategori normal dan abnormal [10]. Perlu menjadi catatan bahwa sebagian besar metode ini menggunakan citra Pap Smear yang hanya berisi satu sel, sehingga hasil segmentasi pada daerah nuklei dan sitoplasma sangat layak untuk dilakukan proses interpretasi. Pada penelitian ini citra yang digunakan merupakan database citra yang dipakai juga pada penelitian [11].

Segmentasi daerah sitoplasma pada citra Pap Smear menjadi isu tersendiri karena dalam citra Pap Smear yang kompleks batas antara sitoplasma yang tumpang tindih sering kali tidak jelas. Gambar 1 menjelaskan bahwa pada citra Pap Smear yang kompleks untuk menentukan batas sitoplasma hampir tidak mungkin. Penelitian sebelumnya [11], [12] menunjukkan bahwa untuk melakukan proses klasifikasi sel normal atau abnormal, daerah sitoplasma digunakan sebagai fitur-fitur penting dalam proses klasifikasi.



Gambar 1. (a) Citra Pap smear dengan batas sitoplasma tiap nuklei yang jelas, (b) Citra Pap smear dengan batas sitoplasma tiap nuklei yang tidak jelas.

Pada penelitian ini, kami mengevaluasi klasifikasi sel serviks berbasis fitur hanya pada daerah nuklei, tanpa mengevaluasi daerah sitoplasma. Sebagai catatan, pada penelitian ini proses segmentasi telah dilakukan sebelumnya [7] dan telah melalui pemeriksaan secara random oleh Dokter spesialis Patologi Anatomi dengan tingkat akurasi yang sangat baik. Penelitian ini adalah langkah penting untuk interpretasi citra Pap smear, dimana sel tumpang tindih khususnya pada daerah sitoplasma merupakan fenomena yang sering ditemukan, dan segmentasi daerah sitoplasma tidak dapat dilakukan secara tepat, sehingga interpretasi berdasarkan rasio ukuran nuklei dengan ukuran sitoplasma tidak dapat dilakukan. Dalam tujuan ini, kami melakukan evaluasi terhadap 25 fitur pada daerah nuklei untuk digunakan sebagai fitur dalam proses klasifikasi. Selanjutnya, kami membandingkan hasil kinerja dua metode klasifikasi yaitu K-Nearest Neighbor dan Backpropagation. Dari penelitian ini diharapkan mendekati kesimpulan, apakah klasifikasi sel normal dan abnormal bisa dilakukan hanya dengan menggunakan fitur-fitur dari daerah nuklei, fitur apa saja yang signifikan, serta tingkat akurasi klasifikasi.

2 Materi dan Metode

Data yang digunakan pada penelitian ini merupakan koleksi data citra Pap Smear dalam [11]. Data yang digunakan terdiri dari 917 citra dengan masing-masing mengandung satu sel yang terbagi menjadi tujuh kelas. Pada penelitian ini tiga dari tujuh tipe sel yang digunakan dianggap sebagai kategori normal sedangkan empat dari tujuh tipe sel yang lainnya dianggap sebagai jenis sel abnormal. Penjelasan rinci tentang dataset yang digunakan terdapat pada Tabel 1.

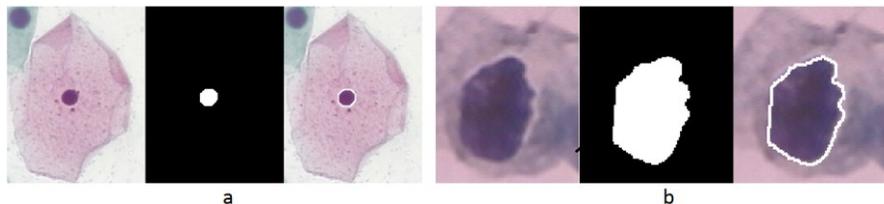
Table 1.Tabel Dataset

No	Tipe Sel	Kategori	Jumlah	Total
1	Superficial squamous epithelial	Normal	74	242 normal
2	Intermediate squamous epithelial	Normal	70	
3	Columnar epithelial	Normal	98	
4	Mild squamous non-keratinizing dysplasia	Abnormal	182	675 abnormal
5	Moderate squamous non-keratinizing dysplasia	Abnormal	146	
6	Severe squamous non-keratinizing dysplasia	Abnormal	197	
7	Squamous cell carcinoma in situ intermediate	Abnormal	150	

2.1 Ekstraksi Fitur

Ekstraksi fitur merupakan tahap mengambil karakter-karakter (fitur) penting dalam sel setelah proses segmentasi citra. Proses segmentasi citra pada penelitian ini telah dilakukan sebelumnya [6] dan telah melalui pemeriksaan secara random oleh Dokter spesialis Patologi Anatomi dengan tingkat akurasi yang sangat baik. Fitur yang akan diekstrak pada tahap ini terdiri dari 25 fitur yang berasal dari wilayah nuclei dan menggunakan beberapa kriteria fitur dari beberapa penelitian sebelumnya [11], [13], [14], [15], [16]. Fitur tersebut dikelompokkan menjadi 4 bagian berdasarkan ukuran, bentuk, intensitas dan tekstur. Tabel 2 menunjukkan 25 fitur yang digunakan pada penelitian ini. Adapun proses segmentasi dan ekstraksi fitur dilakukan secara otomatis menggunakan software matlab. Berikut penjelasan karakteristik fitur pada daerah nuclei.

Fitur Ukuran dan Fitur Bentuk.



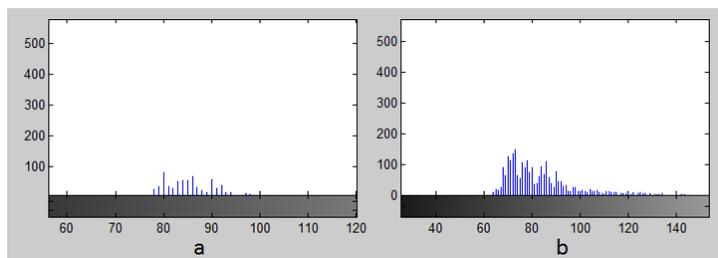
Gambar 2. Perbedaan hasil segmentasi, (a) sel normal, (b) sel abnormal

Sel epitel (nuclei) normal dan sel epitel abnormal memiliki bentuk yang hampir mirip. Akan tetapi, sel epitel abnormal memiliki bentuk yang lebih besar dan lebih tidak teratur dari pada sel epitel normal. Gambar 2 mengilustrasikan perbedaan bentuk dan ukuran antara sel epitel normal dan sel epitel abnormal. Pada Gambar 2(a) terlihat hasil segmentasi citra menunjukkan perbedaan antara sel epitel normal dan sel epitel abnormal dalam hal ukuran (diameter), maupun keteraturan bentuk.

Fitur Tekstur.

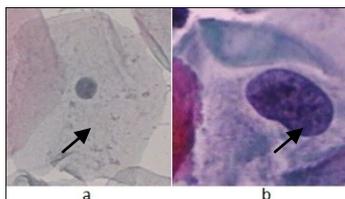
Analisis fitur tekstur dari hasil segmentasi dilakukan berdasarkan analisis statistik histogram intensitas pada 3 layer warna (red, green, blue). Sel epitel normal dan sel epitel abnormal memiliki tekstur yang cenderung berbeda. Pada Gambar 3, terlihat bahwa salah satu sample sel epitel abnormal memiliki persebaran intensitas warna yang agak lebar (antara 60 sampai 145 (Gambar 3(a)) daripada sel epitel

normal (antara 75 sampai 100 (Gambar 3(b)).Hal tersebut menunjukkan intensitas warna pada sel epitel normal lebih homogen daripada sel epitel abnormal yang dalam hal ini berarti sel epitel normal memiliki tekstur yang lebih halus daripada tekstur sel epitel abnormal.



Gambar 3. Perbedaan histogram sample sel epitel, (a) sel normal, (b) sel abnormal

Fitur Intensitas Warna.



Gambar 4. Perbedaan intensitas pada daerah nuclei, (a) sel normal, (b) sel abnormal

Fitur yang mencirikan intensitas pada masing-masing sel adalah nilai rata-rata intensitas semua piksel pada daerah nuclei. Perbedaan mendasar intensitas antara sel epitel normal dengan sel epitel abnormal yaitu sel epitel normal memiliki intensitas warna pada daerah nuclei yang cenderung lebih tinggi daripada sel epitel abnormal. Lebih lanjut, penelitian sebelumnya [17] menyebutkan bahwa intensitas warna pada sel kanker (abnormal) memiliki rentang antara 80-100, dan sel normal memiliki rentang intensitas antara 122-150. Gambar 4 menunjukkan perbedaan intensitas warna antara sel epitel normal dan sel epitel abnormal.

Tabel 2. Fitur yang digunakan

Fitur Ukuran		Fitur Bentuk	
Semi minor-Axis length	$b =$ jarak terpendek <i>centroid</i> dengan <i>boundary</i> nuclei yang tegak lurus dengan major axis (1)	Eccentricity	$c = \sqrt{a^2 - b^2}; e = c/a$ (6)
Semi major-Axis length	$a =$ jarak terpanjang <i>centroid</i> dengan <i>boundary</i> nuclei (2)	Sphericity	$\psi = \frac{\text{Min}_{Axis}}{\text{Max}_{Axis}}$ (7)
Average Radius	$Avr =$ rata-rata jarak <i>centroid</i> dengan <i>boundary</i> nuclei (3)	Circularity	$C = \frac{4\pi \times \text{Area}}{p^2}$ (8)
Equivalent Diameter	$ED = \frac{4 \times \text{Area}}{\pi}$ (4)	Compactness	$Cp = P^2 / \text{Area}$ (9)
Perimeter	$P =$ Jumlah pixel pada <i>boundary</i> nuclei (5)	Variance	$Eva = \frac{\sigma_r}{\mu_r}$ (10)
Fitur Tekstur		Radius	
		Moment of Inertia	$J = \frac{1}{4} \pi ab (a^2 + b^2)$ (11)

Mean (First Moment)	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	(14)	Elongation	$E = \frac{Min_{Axis}}{Max_{Axis}}$	(12)
Stdev (Second Moment)	$\sigma_{N-1} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$	(15)	Roundness	$Circle = \frac{\pi}{4 * (Max_{Axis})^2} \Rightarrow$ $Roundness = \frac{Area}{Circle}$	(13)
Skewness (Third Moment)	$skew = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)S_{N-1}^3}$	(16)	Fitur Intensitas		
Smoothness	$R = 1 - \frac{1}{1 + S_{N-1}^2}$	(17)	Mean intensity in red	$\bar{R} = \frac{1}{n} \sum_{i=1}^n x_i$ dimana n adalah banyaknya pixel dalam citra, dan x_i adalah nilai intensitas warna pada tiap pixel (red)	(22)
Entropy	$H(x) = - \sum_{i=1}^n p(z_i) \times \log_2 p(z_i)$	(18)	Mean intensity in green	$\bar{G} = \frac{1}{n} \sum_{i=1}^n x_i$ dimana n adalah banyaknya pixel dalam citra, dan x_i adalah nilai intensitas warna pada tiap pixel (green)	(23)
Uniformity	$U = \sum_{i=0}^n p^2(z_i)$ z_i adalah nilai intensitas pada setiap pixel, dan $p(z)$ adalah jumlah pixel untuk setiap intensitas pada histogram dibagi dengan jumlah pixel.	(19)	Mean intensity in blue	$\bar{B} = \frac{1}{n} \sum_{i=1}^n x_i$ dimana n adalah banyaknya pixel dalam citra, dan x_i adalah nilai intensitas warna pada tiap pixel (blue)	(24)
Sum of Square	$SS = \sum_{i=1}^n (x_i - \bar{x})^2$	(20)	Brightness	$Y = (0.299 * \bar{R})$ $+ (0.587 * \bar{G})$ $+ (0.114 * \bar{B})$	(25)
Contrast	$C = \frac{\sigma}{\left(\frac{\mu_4}{\sigma^4}\right)^{\frac{1}{4}}}$ μ_4 = kurtosis, σ = standard deviation	(21)	Nilai bobot merupakan standarkecerahan yang dirasakan oleh mata manusia		

2.2 Seleksi Fitur dan Klasifikasi

Pada tahap ini, hasil dari proses ekstraksi fitur kemudian dilakukan normalisasi data karena adanya perbedaan rentang nilai pada tiap fitur dengan menggunakan normalisasi min-max. Adapun untuk melakukan proses normalisasi data menggunakan min-max dapat dilakukan menggunakan persamaan berikut,

$$N(i) = \frac{F(i) - \min(F)}{\max(F) - \min(F)} \in D_F \quad (26)$$

Dengan F adalah index fitur yang di gunakan, dan i adalah index data pada dimensi (D_F).

Pada penelitian ini kami membandingkan kinerja hasil klasifikasi menggunakan dua metode seleksi fitur yaitu Fisher Criterion dan Mahalanobis Distance. Mahalanobis Distancedidapatkan dengan persamaan berikut,

$$D = \sqrt{(\mu_{-1} - \mu_{+1})^t \Sigma^{-1} (\mu_{-1} - \mu_{+1})} \quad (27)$$

Dimana $\mu_{j,+1}$ adalah nilai rata-rata yang termasuk pada kelas sel epitel normal, $\mu_{j,-1}$ adalah nilai rata-rata data pada kelas sel epitel abnormal, Σ^{-1} adalah inverse covariance matrix. Pada penelitian ini kami memilih Fisher Criterion daripada Mahalanobis Distance karena Fisher Criterion menghasilkan tingkat accuracy klasifikasi yang lebih tinggi. Nilai accuracy tertinggi dari Fisher Criterion adalah 92.26% menggunakan klasifikasi KNN dengan nilai $k=9$, sedangkan nilai accuracy tertinggi Mahalanobis distance adalah 91.71% menggunakan klasifikasi KNN dengan nilai $k=9$. Perbandingan metode seleksi fitur ini juga dibahas pada penelitian sebelumnya [18].

Pada penelitian ini kami menguji tingkat signifikansi masing-masing ke-25 fitur yang digunakan dengan menggunakan Fisher Criterion. Fisher Criterion adalah sebuah metode untuk menentukan fitur-fitur yang signifikan dari sekumpulan fitur. Hasil dari Fisher Criterion adalah ranking dari setiap fitur yang kemudian dilanjutkan dengan proses klasifikasi untuk melakukan proses seleksi fitur. Metode Fisher Criterion didapatkan dengan persamaan berikut,

$$F(x_j) = \frac{n_{+1}n_{-1} (\mu_{j,+1} - \mu_{j,-1})^2}{n_{+1}n_{-1} \sigma_{j,+1}^2 + n_{j,-1} \sigma_{j,-1}^2} \quad (28)$$

dimana n_{+1} adalah jumlah data pada fitur tertentu yang termasuk pada kelas epitel normal, n_{-1} adalah jumlah data pada fitur tertentu yang termasuk pada kelas sel epitel abnormal, $\mu_{j,+1}$ adalah rata-rata data pada fitur ke- j yang termasuk pada kelas epitel normal, $\mu_{j,-1}$ adalah rata-rata data pada fitur ke- j yang termasuk pada kelas sel epitel abnormal, $\sigma_{j,+1}^2$ adalah kuadrat dari standar deviasi data pada fitur ke- j yang termasuk pada kelas epitel normal, $\sigma_{j,-1}^2$ adalah kuadrat dari standar deviasi data pada fitur ke- j yang termasuk pada kelas sel epitel abnormal. Algoritma ini menghasilkan output berupa ranking setiap fitur.

Untuk mengetahui efektivitas dari pengurangan dimensi/ fitur, pada penelitian ini kami membandingkan dua metode klasifikasi. Lebih khusus, kami menggunakan *K-Nearest Neighbor* (KNN) dan *Backpropagation*. Pada pengujian pengurangan dimensi/ fitur menggunakan metode KNN, kami melakukan beberapa percobaan terhadap nilai k ($k=3$, $k=5$, $k=7$, dan $k=9$). Hasil klasifikasi dengan nilai k terbaik akan dibandingkan kinerjanya dengan metode *Backpropagation*. Beberapa percobaan dilakukan dan kinerja metode klasifikasi diukur dengan menggunakan pola peningkatan dimensi bervariasi dari 1 sampai 25 fitur. Pada penelitian ini metode *Backpropagation* menggunakan *hidden layer* yang berjumlah 13, dengan jumlah epoch sebanyak 1000.

Pada penelitian ini digunakan metode *k-fold cross validation* dengan nilai $k=10$ untuk memvalidasi metode klasifikasi yang digunakan. Data yang digunakan dibagi menjadi sepuluh subset secara acak yang masing-masing subset memiliki jumlah *instance* dan perbandingan jumlah kelas yang sama. Pembagian data ini digunakan pada proses iterasi klasifikasi. Iterasi dilakukan sebanyak sepuluh kali sesuai dengan nilai k yang digunakan. Pada setiap iterasi, satu subset digunakan untuk pengujian sedangkan subset-subset lainnya digunakan untuk pelatihan.

3 Hasil

Pada penelitian ini, kami menguji efektivitas skema pengurangan dimensi/ fitur di atas dalam kinerja klasifikasi. Output dari algoritma Fisher Criterion adalah berupa ranking fitur dimana semakin tinggi ranking sebuah fitur, maka tingkat signifikannya terhadap hasil klasifikasi semakin tinggi. Berikut Tabel 3 menunjukkan hasil ranking terhadap fitur yang digunakan menggunakan Fisher Criterion. Selanjutnya, kami menguji tingkat accuracy metode KNN terhadap hasil dari ranking fitur yang dihasilkan sebelumnya menggunakan Fisher Criterion. Adapun pengujian tersebut dilakukan terhadap nilai k yang berbeda-beda (k=3, k=5, k=7, dan k=9). Tabel 4. menunjukkan tingkat accuracy pengujian metode KNN terhadap ranking Fisher Criterion. Proses pengujian ini dilakukan dengan bantuan software Weka.

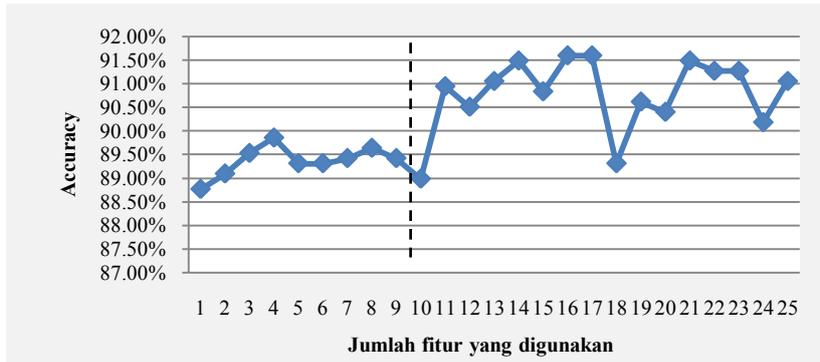
Tabel 3. Ranking Fitur dengan *Fisher Criterion*

Rank	Fitur	Rank	Fitur	Rank	Fitur
1	<i>Uniformity</i>	10	<i>Circularity</i>	19	<i>Mean Intensity Red</i>
2	<i>Semi Major</i>	11	<i>Skewwnes</i>	20	<i>Mean (First Moment)</i>
3	<i>Perimeter</i>	12	<i>Compactness</i>	21	<i>Contrast</i>
4	<i>Average Radius</i>	13	<i>Mean Intensity Green</i>	22	<i>Stdev (Second Moment)</i>
5	<i>Sum Of Square</i>	14	<i>Brightness</i>	23	<i>Smoothness</i>
6	<i>Semi Minor</i>	15	<i>Eccentricity</i>	24	<i>Variance Radius</i>
7	<i>Equivalent Diameter</i>	16	<i>Elongation</i>	25	<i>Entropy</i>
8	<i>Roundness</i>	17	<i>Mean Intensity Blue</i>		
9	<i>Moment of Inertia</i>	18	<i>Sphericity</i>		

Table 4. Accuracy dari metode KNN terhadap *Fisher Criterion*

k	MaxAccuracy	Jumlah Fitur yang digunakan
KNN-3	92.15%	21
KNN-5	91.71%	15
KNN-7	92.15%	14
KNN-9	92.26%	14

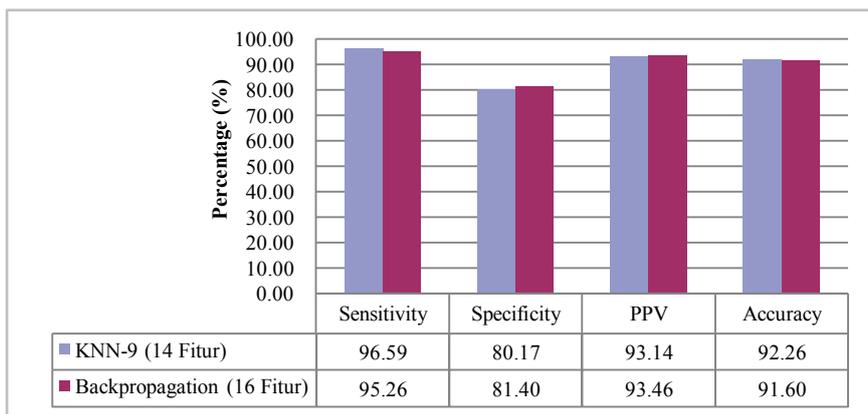
Dari Tabel 4. dapat dilihat bahwa nilai optimal klasifikasi adalah sebesar 92.26% dengan menggunakan nilai k=9 dan menggunakan 14 fitur. Selanjutnya, kami melakukan pengujian klasifikasi terhadap dataset yang digunakan menggunakan metode Backpropagation. Gambar 5. Menunjukkan tingkat accuracy pengujian metode Backpropagation terhadap ranking Fisher Criterion. Dari gambar tersebut menunjukkan nilai klasifikasi tertinggi dengan menggunakan metode Backpropagation adalah 91.60% dengan menggunakan 16 fitur. Pada pengujian klasifikasi menggunakan metode Backpropagation, terdapat kesamaan nilai tertinggi yakni saat menggunakan 16 fitur dan 17 fitur. Pada penelitian ini kami menggunakan metode Backpropagation dengan jumlah fitur 16 karena memiliki lebih kecil dimensi/ fitur sehingga komputasinya lebih efisien.



Gambar 5. Nilai Accuracy dari metode Backpropagation terhadap Fisher Criterion

Selanjutnya, kami melakukan pengujian kinerja klasifikasi antara metode KNN (dengan nilai $k=9$, dan menggunakan 14 fitur) dengan metode *Backpropagation* (menggunakan 16 fitur). Pengujian yang dilakukan adalah dengan mengukur tingkat akurasi, sensitivitas, spesifisitas dan nilai prediktif positif (PPV) masing-masing metode terhadap dataset yang digunakan. Perbandingan masing-masing jenis pengujian kinerja dapat dilihat pada Gambar 6.

Gambar 6 menunjukkan bahwa kinerja metode KNN memiliki tingkat akurasi klasifikasi yang lebih baik daripada metode *Backpropagation* saat diterapkan pada dataset yang kami gunakan. Dari hasil tersebut juga dapat disimpulkan bahwa proses seleksi fitur terbukti meningkatkan tingkat akurasi proses klasifikasi antara sel normal dan sel abnormal, karena jika menggunakan 25 fitur tingkat akurasinya hanya 91.06% sedangkan dengan 14 fitur tingkat akurasinya adalah 92.26%. Selanjutnya, proses klasifikasi antara sel normal dan abnormal dapat dilakukan menggunakan 14 dari 25 fitur yang di analisis yaitu: *Uniformity, Semi Major, Perimeter, Average Radius, Sum Of Square, Semi Minor, Equivalent Diameter, Roundness, Moment of Inertia, Circularity, Skewness, Compactness, Mean Intensity Green, dan Brightness*. Pada penelitian sebelumnya [16] dijelaskan bahwa hanya dari daerah nuclei dapat dilakukan ekstraksi fitur sebanyak 132 fitur yang meliputi: *Morphometry* (43 fitur), *Photometric* (5 fitur), *Discrete texture* (20 fitur), *Markovian texture* (7 fitur), *Non-Markovian texture* (5 fitur), *Fractal texture* (3 fitur), *Run length texture* (20 fitur), *Voronoi polygons* (18 fitur), *Delaunay graph* (4 fitur), dan *minimum spanning tree* (MST) (7 fitur). Penggunaan 132 fitur tersebut akan diinvestigasi apakah mampu meningkatkan nilai akurasi klasifikasi sel normal dan sel abnormal pada penelitian selanjutnya.



Gambar 6. Hasil pengujian kinerja

4 Kesimpulan

Dari daerah nucleikita dapat menghitung fitur-fitur penting yang berkontribusi dalam proses klasifikasi sel-sel normal dan abnormal. Dalam citra Pap smear yang kompleks, nuclei merupakan daerah penting yang bisa digali informasinya secara otomatis [2], [7], dan [19]. Pada penelitian ini terbukti bahwa klasifikasi terhadap fitur-fitur dari daerah nuclei yakni menggunakan 14 fitur tingkat akurasinya mencapai 92.26%. Hasil ini memiliki arti penting, karena hal ini menyiratkan bahwa kita dapat mencapai hasil kinerja klasifikasi yang baik hanya dengan penggunaan fitur pada daerah nuclei saja. Hal ini dapat berkontribusi dalam pengembangan metode otomatis untuk analisis citra mikroskopis sel serviks yang kompleks, melalui tahapan segmentasi pada daerah nuclei, mengekstrak fiturnya dan terakhir dilakukan proses klasifikasi. Selanjutnya, kami merasa perlu melakukan pengujian klasifikasi dengan membandingkan tingkat akurasi antara penggunaan fitur hanya pada nuclei dan penggunaan fitur pada daerah nuclei dan sitoplasma untuk menegaskan kesimpulan di atas.

5 Referensi

1. Plissiti, ME., Nikou, C., and Charchanti, A., "Automated detection of Cell Nuclei in Pap Smear Images Using Morphological Reconstruction and Clustering," *IEEE Trans on Information technology in Biomedicine.*, vol. 15, no 2, March. 2011, pp. 233-241.
2. Muhimmah, I., Kurniawan, R., Indrayanti, "Automatic Epithelial Cells Detection of Pap smears images using Fuzzy C-Means Clustering", *2012 4th International Conference on Bioinformatics and Biomedical Technology*, pp. 122-127, 2012.
3. Lin, C. H., Chan, Y. K., and Chen, C. C., "Detection and segmentation of cervical cell cytoplasm and nucleus," *Int. J. Imaging Syst. Technol.*, vol. 19, no. 3, pp. 260-270, 2009.
4. Yang-Mao, S. F., Chan, Y. K., and Chu, Y. P., "Edge enhancement nucleus and cytoplasm contour detector of cervical smear images," *IEEE Trans. Syst. Man Cybern. B, Cybern.*, vol. 38, no. 2, pp. 353-366, Apr. 2008.
5. Garrido, A., N. Pérez de la Blanca, "Applying deformable templates for cell image segmentation", *Pattern Recognition*, vol. 33, 2000, pp. 821-832
6. Bak, E., Najarian, K., and Brockway, J. P., "Efficient segmentation framework of cell images in noise environments," in *Proc. 26th Int. Conf. IEEE Eng. Med. Biol.*, Sep., 2004, vol. 1, pp. 1802-1805.
7. Muhimmah, I., Kurniawan, R., Indrayanti (2012) Automated Cervical Cell Nuclei Segmentation using Morphological Operation and Watershed Transformation. In *IEEE International Conference on Computational Intelligence and Cybernetics 2012*.
8. K. B. Kim, S. Kim, and K. B. Sim. Nucleus classification and recognition of uterine cervical Pap-smears using fuzzy art algorithm. In *Proceedings of 6th International Conference on Simulated Evolution and Learning, Lecture Notes in Computer Science*, volume 4247, pages 560-567, 2006.
9. K. B. Kim, D. H. Song, and Y. W. Woo. Nucleus segmentation and recognition of uterine cervical Pap smears. In *Proceedings of the 11th RSFDGrC 2007, Lecture Notes in Computer Science*, volume 4482, pages 153-160, 2007.
10. P. C. Huang, Y. K. Chan, P. C. Chan, Y. F. Chen, R. C. Chen, and Y. R. Huang. Quantitative assessment of Pap smear cells by PC-based cytopathologic image analysis system and support vector machine. In *Proceedings of the 1st International Conference on Medical Biometrics*, 2007.
11. Jantzen, J., Norup, J., Dounias, G., Bjerregaard, B.: Pap-smear benchmark data for pattern classification. In: *Proceedings of Nature inspired Smart Information Systems (NiSIS)*, pp. 1-9 (2005)
12. B. Rosidi, N. Jalil, N. M. Pista, L. H. Ismail, E. Supriyanto, and T. L. Mengko, "Classification of Cervical Cells Based on Labeled Colour Intensity Distribution," *International Journal of biology and biomedical engineering*, issue 2, vol. 5, 2011.
13. M. E. Plissiti, C. Nikou, and A. Charchanti, "Combining shape, texture and intensity features for cell nuclei extraction in Pap smear images," *Pattern Recognition Letters*, vol. 32, no. 6, pp. 838-853, 2011.
14. Drezek R, Guillaud M, Collier T, Boiko I, Malpica A, Macaulay C, Follen M, Richards-Kortum R, "Light scattering from cervical cells throughout neoplastic progression: influence of nuclear morphology, DNA content, and chromatin texture", *J. Biomed. Opt.* 8, 7 (2003);
15. L.A. West, R. Swartz, D. Cox, I.V. Boiko, A. Malpica, C. MacAulay, M. Follen, "Cytometric features of cell nuclei of adenocarcinoma in situ and invasive adenocarcinoma of the cervix," *Am J Obstet Gynecol*, vol.187, pp.1566-1573, 2002.

16. Guillaud, M., Cox, D., Adler-Storthz, K., Malpica, A., Staerckel, G., Maticic, J., van Niekerk, D., Poulin, N., Follen, M., and MacAulay, C., "Exploratory Analysis of Quantitative Histopathology of Cervical Intraepithelial Neoplasia: Objectivity, Reproducibility, Magnancy-Associated Changes, and Human Papillomavirus". *Cytometry Part A* vol. 60A, 2004, pp.8.
17. C. Li, C. Xu, C. Gui, and M. D. Fox., *Level Set Without Re-initialization: A New Variational Formulation*, Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005.
18. A. Jain and D. Zongker, "Feature Selection: Evaluation, Application, and Small Sample Performance," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, 153-158, Feb. 1997.
19. Muhimmah, I., Kurniawan, "Shape-based nuclei area of digitized pap smear images", *2012 4th International Conference on Digital Image Processing (ICDIP 2012)*, 83344J.