

**PENAMBAHBAIKAN PROSES PENGEKSTRAKAN CIRI DAN
PENGKLASIFIKASIAN SEL KANSER PANGKAL RAHIM
UNTUK SISTEM NEURALPAP**

Oleh

SITI NORAINI SULAIMAN

**Tesis yang diserahkan untuk memenuhi
keperluan bagi Ijazah Doktor Falsafah**

Februari 2012

PENGHARGAAN

Dengan nama Allah yang maha pemurah lagi maha penyayang. Alhamdulillah...syukur ke hadrat Ilahi kerana limpah kurnia dan izinNya, saya berjaya menyiapkan penyelidikan dan tesis ini sebagai memenuhi keperluan pengijazahan peringkat doktor falsafah. Setelah 4 tahun bertungkus lumus dengan bantuan penyelia, rakan-rakan serta juruteknik-juruteknik akhirnya penyelidikan ini berjaya direalisasikan.

Di kesempatan ini, saya ingin merakamkan setinggi-tinggi penghargaan dan terima kasih kepada penyelia, Prof. Madya Dr. Nor Ashidi Mat Isa yang telah banyak membantu dalam pelbagai aspek. Dengan sokongan, tunjuk ajar, idea, nasihat dan kepercayaan yang diberikan oleh beliau telah mendorong saya bagi meneruskan dan seterusnya menyiapkan penyelidikan ini. Tidak dilupa kepada pemeriksa-pemeriksa yang berpengalaman, Prof. Dr. Mohd Yusoff Mashor, Prof. Madya Dr. Umi Kalthum Ngah, dan Prof Madya Dr. Kamal Zuhairi Zamli atas tunjuk ajar, komen dan idea-idea yang bernas dan membina bagi mengukuhkan lagi persembahan tesis ini.

Saya juga ingin mengucapkan ribuan terima kasih kepada majikan; Universiti Teknologi MARA cawangan Pulau Pinang yang telah meluluskan permohonan cuti belajar serta Kementerian Pengajian Tinggi Malaysia yang menganugerahkan biasiswa bagi membolehkan saya meneruskan pengajian di peringkat ini. Tidak lupa kepada Majlis Kanser Nasional (MAKNA) yang telah membiaya penyelidikan ini melalui geran penyelidikan yang bertajuk "Development of an Intelligent Screening System for Cervical Cancer".

Istimewa buat ibu yang sedang sakit, Pn. Siti Mariyah dan ayah, En. Sulaiman yang sentiasa menjadi pendorong agar saya meneruskan pengajian sehingga ke peringkat tertinggi. Tidak lupa buat suami tercinta, Roy Razlan dan anak-anak yang disayangi, Nur Alya Adriana dan Muhammad Afieq yang sanggup bersabar sepanjang saya meneruskan pengajian ini serta seluruh keluarga yang sangat memahami. Doa, dorongan, semangat dan kesabaran semua merupakan pendorong kepada saya bagi berusaha mencapai impian ini.

Akhir sekali, sekalung penghargaan dan ucapan terima kasih saya tujukan kepada Intan Aidha dan En. Sazali serta rakan-rakan seperjuangan di ISRT, warga Jabatan Patologi dan Jabatan Anatomi, HUSM terutama Prof. Nor Hayati Bt Othman, Prof Madya Dr Nor Aliza Bt Abdul Ghaffar, En. Ismail, En. Mazlan, dan En. Harissal, Prof. Madya Dr Rohana Bt Adnan dari Pusat Pengajian Sains Kimia, warga IPPT di bangunan Eureka USM Minden terutama Pn. Salwa dan seluruh warga Pusat Pengajian Kejuruteraan Elektrik dan Elektronik, USM serta warga Fakulti Kejuruteraan Elektrik, UiTM Pulau Pinang yang telah membantu sama ada secara langsung atau tidak langsung. Segala budi dan pertolongan yang dihulurkan, akan dikenang sehingga akhir hayat dan hanya Allah S.W.T sahaja yang mampu membalasnya. InsyaAllah.

Sekian, wassalam.

KANDUNGAN

	Muka Surat
PENGHARGAAN	ii
KANDUNGAN	iv
SENARAI RAJAH	viii
SENARAI JADUAL	xii
TERJEMAHAN ISTILAH	xiv
SINGKATAN ISTILAH	xix
ABSTRAK	xxi
ABSTRACT	xxiii
BAB 1 PENGENALAN	
1.1 Pendahuluan	1
1.2 Pengklasifikasian Imej Kanser Pangkal Rahim	2
1.3 Pernyataan Masalah	3
1.4 Objektif Penyelidikan	5
1.5 Skop Penyelidikan	6
1.6 Garis Panduan Tesis	7
BAB 2 KAJIAN ILMIAH	
2.1 Pendahuluan	9
2.2 Kanser Pangkal Rahim	11
2.2.1 Jenis-jenis dan Peringkat-peringkat Kanser Pangkal Rahim	13
2.2.2 Ujian Saringan Kanser Pangkal Rahim	17
2.2.3 Sistem Sokongan	20
2.3 Sistem Sokongan Pengklasifikasian	21
2.3.1 Sistem Pengklasifikasian Berdasarkan Teknik Pemprosesan Imej	22
2.3.2 Sistem Pengklasifikasian Berdasarkan Kecerdikan Buatan	45

2.3.3	Sistem Pengklasifikasian Berdasarkan Gabungan Pemrosesan Imej dan Kecerdikan Buatan	54
2.4	Sistem NeuralPap	60
2.4.1	Ciri-ciri Sistem NeuralPap	61
2.4.1(a)	Pengekstrakan Ciri dengan Teknik Pemrosesan Imej	65
2.4.1(b)	Pengklasifikasian Sel Kanser Menggunakan Rangkaian Neural	67
2.4.2	Limitasi Sistem NeuralPap	67
2.5	Ringkasan	70
BAB 3 METODOLOGI		
3.1	Pendahuluan	71
3.2	Pengumpulan Data dan Perolehan Imej	74
3.3	Peruasan Imej Palitan Pap	75
3.3.1	Peruasan Imej Tanpa Ciri Penyingkiran Hingar	75
3.3.2	Peruasan Imej dengan Ciri Penyingkiran Hingar Terbenam	79
3.3.2 (a)	Algoritma Pensuisan Kelompok	84
3.3.2 (b)	Algoritma Pengelompokan Berasaskan Nyah-hingar	88
3.4	Pengekstrakan Ciri-ciri Imej Palitan Pap	90
3.4.1	Algoritma Pengekstrakan Ciri Warna Pseudo Manual	92
3.4.2	Algoritma Pengekstrakan Ciri Warna Pseudo Separa-Automatik	97
3.5	Penambahbaikan Pengklasifikasian Menggunakan Rangkaian Neural	101
3.6	Ringkasan	104
BAB 4 KEPUTUSAN		
4.1	Pendahuluan	106
4.2	Keputusan Peruasan Imej Tanpa Ciri Penyingkiran Hingar	107
4.2.1	Keputusan Algoritma Pengelompokan Terhadap Imej Piawai	109
4.2.2	Keputusan Algoritma Pengelompokan Terhadap Imej Palitan Pap	120

4.3	Keputusan Perluasan Imej dengan Ciri Penyingkiran Hingar Terbenam	125
4.3.1	Keputusan Algoritma Pengelompokan dengan Ciri Penyingkiran Hingar Terbenam Terhadap Imej Piawai	127
4.3.2	Keputusan Algoritma Pengelompokan dengan Ciri Penyingkiran Hingar Terbenam Terhadap Imej Palitan Pap	139
4.4	Keputusan Pengekstrakan Ciri-ciri Imej Palitan Pap	145
4.4.1	Keputusan Algoritma Pengekstrakan Ciri Imej Sel Tunggal	146
4.4.2	Keputusan Algoritma Pengekstrakan Ciri Imej Sel Bertindih	158
4.5	Keputusan Penambahbaikan Pengklasifikasian Menggunakan Rangkaian Neural	164
4.5.1	Analisa Algoritma Penetapan Pusat Rangkaian	165
4.5.2	Analisa Prestasi Pengklasifikasian oleh Rangkaian Neural	166
4.6	Kesimpulan	169
BAB 5 KESIMPULAN DAN CADANGAN		
5.1	Kesimpulan	172
5.2	Cadangan Kajian Masa Depan	174
RUJUKAN		176
LAMPIRAN		
LAMPIRAN A: Kesan pemilihan nilai maksimum dan minimum kepada pembetulan nilai piksel 'hingar'		194
LAMPIRAN B: Keputusan pemerhatian secara visual pengaplikasian algoritma pengelompokan tanpa hingar ke atas imej piawai		196
LAMPIRAN C: Keputusan pemerhatian secara visual pengaplikasian algoritma pengelompokan tanpa hingar ke atas imej sel pangkal rahim		202
LAMPIRAN D: Keputusan pemerhatian secara visual pengaplikasian algoritma pengelompokan dengan ciri penyingkiran hingar terbenam ke atas imej piawai		204

LAMPIRAN E: Keputusan pemerhatian secara visual pengaplikasian algoritma pengelompokan dengan ciri penyingkiran hingar terbenam ke atas imej sel pangkal rahim	208
LAMPIRAN F: Keputusan pemerhatian visual pengekstrakan ciri imej sel pangkal rahim tunggal	210
LAMPIRAN G: Keputusan pemerhatian visual pengekstrakan ciri imej sel pangkal rahim bertindih	216
SENARAI PENERBITAN	219

SENARAI RAJAH

	Muka Surat
Rajah 2.1: Anatomi sistem pembiakan wanita	12
Rajah 2.2: Perubahan keadaan pangkal rahim mengikut peringkat kanser (QIAGEN, 2011)	15
Rajah 2.3: Pemprosesan imej dan algoritma-algoritma masing-masing	25
Rajah 2.4: Senibina rangkaian neural	46
Rajah 2.5: Ringkasan konsep asas empat cabang kecerdikan buatan	48
Rajah 2.6: Paparan antaramuka sub-menu <i>Diagnosis Manual</i> sistem NeuralPap	62
Rajah 2.7: Paparan antaramuka sub-menu <i>Diagnosis Automatic</i> sistem NeuralPap	62
Rajah 2.8: Paparan antaramuka sub-menu <i>Diagnosis Clinical</i> sistem NeuralPap.	62
Rajah 2.9: Contoh paparan antaramuka untuk fasiliti perbandingan prestasi sistem NeuralPap	65
Rajah 2.10: Contoh hasil dari setiap proses oleh sistem NeuralPap (a) Imej asal (b) Keputusan peruasan imej menggunakan teknik algoritma ASBRG2 (c) Penandaan titik awalan (d) Imej keluaran dari proses pengekstrakan ciri menggunakan teknik RGBFE (e) Paparan empat ciri-ciri imej yang diekstrak	66
Rajah 2.11: Keputusan pemprosesan imej tanpa hingar dan imej bercampur hingar oleh sistem NeuralPap. (a) Imej asal tanpa hingar. (b) Keputusan peruasan untuk imej tanpa hingar. (c) Imej bercampur hingar. (d) Keputusan peruasan untuk imej bercampur hingar	69

Rajah 2.12:	Keputusan peruasan imej oleh sistem NeuralPap ke atas imej yang bertindih. (a) Imej asal (b) Imej yang diruas	69
Rajah 3.1:	Gambarajah blok untuk proses peruasan imej tercemar hingar melalui (a) proses konvensional (b) proses peruasan imej dengan ciri penyingkiran hingar terbenam.	81
Rajah 3.2:	Histogram dan imej piawai 'AirForce' yang bersaiz 256×256 (a) Imej asal tanpa hingar, (b) Histogram imej asal tanpa hingar (c) Imej yang dicemari hingar garam-dan-lada sebanyak 10% (d) Histogram imej yang dicemari hingar	83
Rajah 3.3:	Contoh pengesanan hingar dan pengiraan nilai median	86
Rajah 3.4:	Carta Alir algoritma Pensuisan Kelompok	87
Rajah 3.5:	Konsep fuzi yang digunakan bagi menentukan maklumat tempatan	90
Rajah 3.6:	Gambarajah blok pelaksanaan algoritma PCFE	93
Rajah 3.7:	Cara pertumbuhan titik benih (a) 4 jiran yang bersambung (b) 4 jiran pepenjuru (c) 8 jiran tersambung	95
Rajah 3.8:	Carta alir pelaksanaan PCFE	96
Rajah 3.9:	Gambarajah blok proses pemisahan imej sel bertindih	98
Rajah 3.10:	Analogi setiap proses yang terlibat dalam pemisahan imej sel bertindih (a) Imej asal (b) Hasil imej setelah proses peningkatan kesan visual imej pada satah A (c) Hasil imej setelah proses peningkatan kesan visual imej pada satah B (d) Hasil imej proses pengesanan pinggir untuk sel 1 pada satah A (e) Hasil imej proses pengesanan pinggir untuk sel 2 pada satah B (f) Hasil proses pewarnaan pseudo terbenam untuk sel 1 pada satah A (g) Hasil proses pewarnaan pseudo terbenam untuk sel 2 pada satah B	99

Rajah 3.11:	Carta alir pelaksanaan algoritma PCFE separa-automatik bagi mengekstrak ciri imej palitan Pap	101
Rajah 4.1:	Imej piawai asal	108
Rajah 4.2:	Imej asal sel pangkal rahim	108
Rajah 4.3:	Keputusan imej yang telah diruas kepada 3 kluster	110
Rajah 4.4:	Keputusan imej yang telah diruas kepada 4 kluster	111
Rajah 4.5:	Keputusan imej yang telah diruas kepada 5 kluster	112
Rajah 4.6:	Keputusan peruasan (3 kluster) untuk imej potongan 'AirForce'	113
Rajah 4.7:	Keputusan peruasan (4 kluster) untuk imej potongan 'Elaine'	114
Rajah 4.8:	Keputusan peruasan (5 kluster) untuk imej potongan 'Golden Gate'	115
Rajah 4.9:	Keputusan peruasan imej CC1	121
Rajah 4.10:	Keputusan peruasan imej CC2	122
Rajah 4.11:	Keputusan peruasan imej CC3	122
Rajah 4.12:	Keputusan peruasan imej CC4	123
Rajah 4.13:	Keputusan peruasan imej CC5	123
Rajah 4.14:	Keputusan peruasan imej CC6	124
Rajah 4.15:	Imej piawai asal	126
Rajah 4.16:	Imej asal sel pangkal ahim	126
Rajah 4.17:	Keputusan peruasan untuk imej <i>Car2</i> yang telah dicemari dengan hingar garam-dan-lada pada ketumpatan 10%	129
Rajah 4.18:	Keputusan peruasan untuk imej <i>Parthenon</i> yang telah dicemari dengan hingar garam-dan-lada pada ketumpatan 20%	130
Rajah 4.19:	Keputusan peruasan untuk imej <i>Car3</i> yang telah dicemari dengan hingar garam-dan-lada pada ketumpatan 30%	132
Rajah 4.20:	Keputusan peruasan untuk imej <i>Baboon</i> yang telah dicemari dengan hingar garam-dan-lada pada ketumpatan 40%	133

Rajah 4.21:	Keputusan peruasan untuk imej <i>Army-Trucks</i> yang telah dicemari dengan hingar garam-dan-lada pada ketumpatan 50%	135
Rajah 4.22:	Keputusan peruasan untuk imej CC7	140
Rajah 4.23:	Keputusan peruasan untuk imej CC8	141
Rajah 4.24:	Keputusan peruasan untuk imej CC9	142
Rajah 4.25:	Keputusan peruasan untuk imej CC10	143
Rajah 4.26:	Keputusan peruasan untuk imej CC11	144
Rajah 4.27:	Keputusan pengekstrakan ciri untuk imej CC1	147
Rajah 4.28:	Keputusan pengekstrakan ciri untuk imej CC2	148
Rajah 4.29:	Keputusan pengekstrakan ciri untuk imej CC3	149
Rajah 4.30:	Keputusan pengekstrakan ciri untuk imej CC4	151
Rajah 4.31:	Keputusan pengekstrakan ciri untuk imej CC5	152
Rajah 4.32:	Keputusan pengekstrakan ciri untuk imej CC6	153
Rajah 4.33:	Keputusan PCFE Separa-Automatik (dengan FMKM) untuk imej CC12	159
Rajah 4.34:	Keputusan PCFE Separa-Automatik (dengan FMKM) untuk imej CC13	160
Rajah 4.35:	Keputusan PCFE Separa-Automatik (dengan FMKM) untuk imej CC14	161

SENARAI JADUAL

	Muka Surat
Jadual 2.1: Sistem Bethesda	16
Jadual 2.2: Penyelidikan-penyelidikan terdahulu yang mengaplikasikan teknik pemprosesan ke atas imej sel pangkal rahim	43
Jadual 2.3: Penyelidikan-penyelidikan terdahulu yang mengaplikasikan kecerdikan buatan ke atas kanser pangkal rahim	53
Jadual 4.1: Penilaian kuantitatif MSE untuk peruasan imej-imej yang terpilih	117
Jadual 4.2: Penilaian kuantitatif $F(I)$ untuk peruasan imej-imej yang terpilih	118
Jadual 4.3: Penilaian kuantitatif $F'(I)$ untuk peruasan imej-imej yang terpilih	118
Jadual 4.4: Penilaian kuantitatif $Q(I)$ untuk peruasan imej-imej yang terpilih	119
Jadual 4.5: Purata fungsi-fungsi penilaian peruasan untuk 121 imej	120
Jadual 4.6: Penilaian kuantitatif untuk peruasan imej <i>Car2</i> pada Rajah 4.17	137
Jadual 4.7: Penilaian kuantitatif untuk peruasan imej <i>Parthenon</i> pada Rajah 4.18	137
Jadual 4.8: Penilaian kuantitatif untuk peruasan imej <i>Car3</i> pada Rajah 4.19	138
Jadual 4.9: Penilaian kuantitatif untuk peruasan imej <i>Baboon</i> pada Rajah 4.20	138
Jadual 4.10: Penilaian kuantitatif untuk peruasan imej <i>Army-Trucks</i> pada Rajah 4.21	139
Jadual 4.11: Perbandingan keputusan pengekstrakan ciri menggunakan algoritma pengekstrakan ciri ke atas imej sel tunggal	154
Jadual 4.12: Keputusan ujian korelasi untuk 758 imej sel pangkal rahim tunggal	157
Jadual 4.13: Keputusan pengekstrakan ciri menggunakan algoritma PCFE Separa-Automatik ke atas imej sel bertindih	162
Jadual 4.14: Keputusan ujian korelasi untuk 41 imej sel pangkal rahim bertindih	163
Jadual 4.15: Ciri-ciri set data piawai	165
Jadual 4.16: Keputusan analisa prestasi rangkaian RBF	166

Jadual 4.17:	Keputusan analisa prestasi rangkaian HRBF	166
Jadual 4.18:	Ciri-ciri set data sel pangkal rahim	167
Jadual 4.19:	Perbandingan prestasi diagnosis di antara rangkaian neural	168
Jadual 4.20:	Keputusan perbandingan prestasi sistem NeuralPap asal dan NeuralPap Ubahsuai yang dicadangkan (purata dari kaedah Keesahan Silang Lipatan ke-5)	169

TERJEMAHAN ISTILAH

Bahasa Inggeris	Bahasa Melayu
5- fold cross-validation set	Set keesahan silang lipatan ke-5
Absolute luminance	Lar mutlak
Accuracy	Kejituan
Actual training	Latihan sebenar
Adapt	Penyesuaian
Adenocarcinoma	Adenokarsinoma
Arabic character recognition	Pengecaman aksara Arab
Automatic Seed Based Region Growing	Pertumbuhan Kawasan Secara Titik Benih Automatik
Back propagation	Perambatan balik
Belongingness	Kepunyaan
Carcinoma	Karsinoma
Cardiotocograph	Kardiotograf
Case Based Reasoning	Penaakulan Berasaskan Kes
Cervical cancer in-situ	Kanser pangkal rahim in-situ
Cervical intra-epithelial neoplasia	Neoplasia intra-epitelial serviks
Cervix	Pangkal rahim
Classifier	Pengkelas
Colposcopy	Kolposkopi
Competitive learning	Pembelajaran kompetitif
Competitive learning networks	Rangkaian pembelajaran kompetitif
Conditional random field	Julat rawak bersyarat
Correlation test	Ujian korelasi
Cytology	Sitologi
De-noising based clustering	Pengelompokan berasaskan nyah-hingar

Diagnosis	Diagnosis / pengesanan
Dissimilarity	Ketaksamaan
Dysplasia	Displasia
Embedded pseudo-color	Pewarnaan pseudo terbenam
Error free	Tanpa ralat
Expert system	Sistem pintar
False acceptance rate	Kadar penerimaan salah
False negative	Salah negatif
False positive	Salah positif
False rejection rate	Kadar penolakan salah
Fast generalized fuzzy <i>c</i> -means	Purata- <i>c</i> fuzi teritlak pantas
Filtering	Penurasan
Fitness	Padanan
Fully connected	Sambungan penuh
Fuzziness	Kefuzian
Fuzzy <i>c</i> -mean	Purata- <i>c</i> kabur
Fuzzy hough transform	Penjelmaan hough fuzi
Fuzzy membership- <i>k</i> -means	Purata- <i>k</i> -keahlian fuzi
Fuzzy partitioning	Pembahagian fuzi
Generalization	Pengitlakan
Genetic algorithm	Algoritma genetik
Genetic algorithm-artificial neural network	Algoritma genetik-rangkaian neural buatan
Givens Least Square	Kuasa Dua Terkecil Givens
Gray-level co-occurrence matrices	Matriks berlaku-serentak paras-kelabu
Growing method	Kaedah pertumbuhan
Hierarchical hybrid multilayered perceptron	Perseptron berbilang lapisan hibrid berhirarki

Hierarchical hybrid radial basis function	Fungsi asas jejarian hibrid berhirarki
Hierarchical multilayered perceptron	Perseptron berbilang lapisan berhirarki
Hierarchical radial basis function	Fungsi asas jejarian berhirarki
High grade squamous intraepithelial lesion	Lesi intraepitelial skuamus gred tinggi
High-end	Tertinggi
Histogram equalization	Penyeragaman histogram
Human papilloma virus	Virus papilloma manusia
Hybrid multilayered perceptron	Perseptron berbilang lapisan hibrid
Hybrid radial basis function	Fungsi asas jejarian hibrid
Image analyzer	Penganalisa imej
Image compression	Pemampatan imej
Image enhancement	Peningkatan imej
Image restoration	Pemulihan imej
Image segmentation	Peruasan imej
Image synthesis	Sintesis imej
Impulse noise	Hingar denyutan
Inadequate	Tidak cukup
Interface	Antaramuka
Intermediate	Pertengahan
<i>k</i> -mean	Purata- <i>k</i>
Knowledge based	Pengetahuan asas
Learning	Pembelajaran
Learning vector quantization	Pengkuantuman vektor pembelajaran
Lesion	Lesi
Local minima	Minima setempat
Local spatial	Ruang tempatan
Lossy	Kehilangan

Low grade squamous intraepithelial lesion	Lesi intraepitelial skuamus gred rendah
Low-end	Terendah
Magnetic resonance image	Imej salunan magnet
Malignant	Malignan
Mean square error	Ralat kuasa dua min
Memetic pareto artificial neural networks	Rangkaian neural buatan memetic pareto
Moderate displasia	Displasia pertengahan
Modified Recursive Prediction Error	Ralat Ramalan Jadi Semula Ubahsuai
Monolayer	Satu lapis
Moving k -mean	Purata- k boleh gerak
Multilayered perceptron	Perseptron berbilang lapisan
Multiresolution-based signature subspace classifier	Pengkelas sub-ruang penanda berasaskan bebilang-resolusi
Neural network	Rangkaian neural
Outliers	Terpencil
Overfitting	Padanan lampau
Pattern recognitionn	Pengecaman corak
Pseudo-colour feature extraction	Pengekstrakan ciri-ciri pewarnaan pseudo
Radial basis function	Fungsi asas jejarian
Ranking-order case-based reasoning	Penaakulan berasaskan kes susunan teratur
Region growing based feature extraction	Pengekstrakan ciri secara pertumbuhan kawasan
Region of interest	Kawasan dikehendaki
Reliable	Dipercayai
Remotely operated vehicle	Kenderaan terkendali jauh
Retain	Penyimpanan
Retreive	Perolehan

Reuse	Penggunaan semula
Revise	Pengolahan
Salt-and-pepper noise	Hingar garam-dan-lada
Sensitivity	Sensitiviti
Skewed distributions	Taburan tidak lurus
Specificity	Spesifisiti
Squamous cell carcinoma	Karsinoma sel skuamus
Supervised learning	Pembelajaran terselia
Support vector machine	Mesin vektor sokongan
Switching cluster	Pensuisan kelompok
Testing phase	Fasa ujian
Threshold	Ambang
Training phase	Fasa latihan
Transition zone	Zon transisi
Uniformity	Keseragaman
Unsupervised	Tak terselia
Validation	Validasi
Watersheds	Legeh

SINGKATAN ISTILAH

ACR	-	Arabic Character Recognition
ASBRG	-	Automatic seed based region growing
BP	-	Back Propagation
CBR	-	Case Base Reasoning
CIN	-	Cervical Intraepithelial Neoplasia
CIS	-	Cervical cancer in-Situ
CLN	-	Competitive Learning Networks
CRF	-	Conditional Random Field
CTG	-	Cardiotocograph
DB	-	De-noising based
ECBR	-	Euclidean Case-Based Reasoning
FAR	-	False Acceptance Rate
FCM	-	Fuzzy <i>c</i> -means
FDA	-	Food and Drug Administration
FGFCM	-	Fast Generalized Fuzzy <i>c</i> -means
FMKM	-	Fuzzy membership- <i>k</i> -means
FRR	-	False Rejection Rate
GA	-	Genetic Algorithm
GA-ANN	-	Genetic Algorithm-Artificial Neural Network
GCM	-	Gray-level Co-occurrence Matrices
GLS	-	Givens Least Square
GUI	-	Graphic User Interface
H ² MLP	-	Hierarchical Hybrid Multilayered Perceptron
H ² RBF	-	Hierarchical Hybrid Radial Basis Function
HiMLP	-	Hierarchical Multilayered Perceptron
HiRBF	-	Hierarchical Radial Basis Function

HMLP	-	Hybrid Multilayered Perceptron
HPV	-	Human Papilloma Virus
HRBF	-	Hybrid Radial Basis Function
HSIL	-	High-grade Squamous Intraepithelial Lesions
HUSM	-	Hospital Universiti Sains Malaysia
ICBR	-	Inductive Case-Based Reasoning
KM	-	<i>k</i> -means
LSIL	-	Low-grade Squamous Intraepithelial Lesions
LVQ	-	Learning Vector Quantization
MCBR	-	Manhattan Case-Based Reasoning
MKM	-	Moving <i>k</i> -means
MLP	-	Multilayered Perceptron,
MPANN	-	Memetic Pareto Artificial Neural Networks
MRI	-	Magnetic Resonance Image
MRPE	-	Modified Recursive Prediction Error
MSSC	-	Multiresolution-Based Signature Subspace Classifier
NASA	-	National Aeronautics and space Administration
PCFE	-	Pseudo-Colour Feature Extraction
RBF	-	Radial Basis Function
RGBFE	-	Region Growing Based Feature Extraction
ROCBR	-	Ranking-Order Case-Based Reasoning
ROV	-	Remotely Operated Vehicle
SBRG	-	Seed Based Region Growing
SC	-	Switching Cluster
SIL	-	Squamous Intraepithelial Lesions
SVM	-	Support Vector Machine
WHO	-	World Health Organization

**PENAMBAHBAIKAN PROSES PENGEKSTRAKAN CIRI DAN
PENGKLASIFIKASIAN SEL KANSER PANGKAL RAHIM UNTUK SISTEM
NEURALPAP**

Abstrak

Kanser pangkal rahim telah menyebabkan banyak kes kematian setiap tahun. Ujian saringan seperti ujian palitan Pap bagi pengesanan peringkat prakanser pangkal rahim mampu mengelak berlakunya kes kanser. Walau bagaimanapun, ujian palitan Pap mempunyai beberapa kelemahan seperti penyediaan slaid yang kurang efektif dan kesilapan manusia. Oleh itu, sistem diagnosis berbantu-komputer diperkenalkan sebagai penyelesaian kepada permasalahan tersebut. Salah satu sistem diagnosis yang telah dibina ialah NeuralPap. Walau bagaimanapun, prestasi NeuralPap dibatasi oleh beberapa kekangan. Penyelidikan ini mencadangkan beberapa algoritma pemprosesan imej baru bagi mengurangkan kekangan-kekangan tersebut. Algoritma pengelompokan Purata-k-Keahlian Fuzi (FMKM) dicadangkan sebagai pengganti algoritma pengelompokan Purata-k Boleh Gerak (MKM) bagi meruas imej palitan Pap kepada kawasan nukleus, sitoplasma dan latar belakang. Algoritma Pensuisan Kelompok (SC) dan Algoritma Pengelompokan Berasaskan Nyah-hingar (DB) pula dicadangkan bagi meruas sel pangkal rahim yang dicemari hingar. Seterusnya, algoritma pengekstrakan ciri berasaskan perwarnaan pseudo yang dinamakan Pengekstrakan Ciri Pewarnaan Pseudo (PCFE) manual dan PCFE Separa-Automatik direkabentuk bagi menggantikan Pengekstrakan Ciri Secara Pertumbuhan Kawasan (RGBFE) yang menggunakan imej monokrom. Penyelidikan ini telah melangkah setapak ke hadapan berbanding sistem NeuralPap dengan mencadangkan algoritma pengekstrakan ciri bagi sel-sel bertindih iaitu dengan menggabungkan konsep pewarnaan ruang dengan algoritma PCFE Separa-Automatik. Selain itu, penyelidikan ini juga telah mencadangkan penggunaan algoritma FMKM dalam proses penetapan pusat rangkaian Fungsi Asas Jejarian (RBF) dan RBF Hibrid (HRBF) bagi menggantikan algoritma MKM yang digunakan sebelum ini. Semua algoritma yang dicadangkan telah terbukti menghasilkan prestasi yang

lebih baik berbanding algoritma sepadan yang digunakan dalam sistem NeuralPap. Selain itu, penggabungan semua algoritma tersebut berjaya meningkatkan kejituan pengklasifikasian kanser pangkal rahim oleh sistem NeuralPap iaitu dengan peratusan 76.35% berbanding 73.40% yang diperolehi dari sistem NeuralPap sebelumnya, setelah diuji dengan 799 sel pangkal rahim.

IMPROVEMENT OF FEATURES EXTRACTION PROCESS AND CLASSIFICATION OF CERVICAL CANCER FOR THE NEURALPAP SYSTEM

Abstract

Cervical cancer has caused many deaths each year. Screening tests, such as Pap smear test used for the detection of the precancerous stage are able to avoid the occurrence of cervical cancer. However, the Pap smear test has several disadvantages such as less effective slides preparation and human error. Therefore, a computer-aided diagnosis system is introduced as a solution to the problem. One of the diagnostic systems that has been built is NeuralPap. However, the NeuralPap performance is limited by several constraints. This research suggests several new image processing algorithms to reduce these constraints. The Fuzzy membership-k-means (FMKM) clustering algorithm is proposed to replace the Moving k -Means (MKM) to segment Pap smear images into the nucleus, cytoplasm and background regions. The Switching Cluster algorithm and De-noising Based Clustering algorithm are proposed to segment noise contaminated cervical cell images. Next, the feature extraction algorithm based on pseudo colouring called the Pseudo Colour Feature Extraction (PCFE) manual and Semi-Automatic PCFE are designed to replace the Region Growing Based Feature Extraction (RGBFE) which uses monochromatic images. This research is a step forward compared with the NeuralPap system by proposing the feature extraction algorithm for overlapping cells by combining the concept of colour space with Semi-Automatic PCFE algorithm. In addition, this research has also suggested the FMKM algorithm as a new centre positioning algorithm for the Radial Basis Function (RBF) and Hybrid RBF (HRBF) networks replacing the MKM algorithm. The entire proposed algorithm has been proven to produce better performance than the corresponding algorithm used in the NeuralPap. In addition, the combination of all algorithms has managed to increase the accuracy of the classification of cervical cancer by NeuralPap system to 76.35%, compared with 73.40% which is obtained from the previous NeuralPap system, tested with 799 cervical cells.

BAB 1

PENGENALAN

1.1 Pendahuluan

Kanser pangkal rahim yang juga dikenali sebagai kanser serviks adalah penyakit yang digeruni oleh wanita. Menurut laporan dari Pertubuhan Kesihatan Dunia (*World Health Organization, WHO*), kira-kira 530,000 kes baru kanser pangkal rahim telah dikesan di serata dunia pada tahun 2008 dan telah meragut sehingga 275,000 nyawa wanita setahun (IARC, 2012). Pada tahun 2007, lebih 700 wanita Malaysia mati akibat kanser pangkal rahim setiap tahun (Jabatan Penerangan Malaysia, 2010). Kanser ini merupakan kanser ketiga paling kerap dihidapi wanita selepas kanser payudara dan kanser kolorektum (Rosly, 2010; Farush-Khan, 2011). Kerajaan Malaysia dengan kerjasama Persatuan Kanser Kebangsaan Malaysia telah berusaha meningkatkan kesedaran di kalangan wanita Malaysia mengenai bahaya penyakit kanser ini melalui pelbagai program kesedaran. Program-program ini adalah bagi mendidik dan menggalakkan wanita menambahkan pengetahuan mengenai kanser ini. Selain itu, kempen bagi menggalakkan wanita mendapat suntikan pelalian virus papilloma manusia (*Human Papilloma Virus, HPV*) dan menjalani pemeriksaan saringan turut dilancarkan sebagai usaha mengekang masalah ini.

Punca berlakunya kanser masih tidak diketahui menyebabkan langkah pencegahan bagi mengelak seseorang dari dihidapi kanser merupakan sesuatu yang mustahil. Namun, pengesanan pada peringkat awal mampu mengelak penyakit kanser daripada merebak dan menjadi semakin parah. Peringkat prakanser pangkal rahim boleh dikesan melalui ujian-ujian saringan dan ujian palitan Pap merupakan ujian saringan yang sering digunakan di Malaysia. Ujian palitan Pap digunakan bagi mengesan kewujudan sel-sel tidak normal pada pangkal rahim. Sel-sel tidak normal ini merupakan tanda-tanda awal seseorang berkemungkinan

menghidapi kanser pangkal rahim. Menurut perangkaan, kadar risiko menghidapi kanser pangkal rahim untuk wanita yang tidak menjalani pemeriksaan palitan Pap adalah lima kali ganda lebih tinggi berbanding wanita yang menjalani ujian tersebut (Jabatan Penerangan Malaysia, 2010).

1.2 Pengklasifikasian Imej Kanser Pangkal Rahim

Kanser pangkal rahim biasanya mengambil masa 10 hingga 15 tahun bagi bertukar dari tahap tidak bahaya (prakanser) kepada tahap bahaya (QIAGEN, 2011). Oleh itu, teknik atau ujian saringan seperti ujian palitan Pap perlu dilaksanakan bagi mengesan kanser pangkal rahim pada peringkat awal. Walau bagaimanapun, ujian palitan Pap mempunyai beberapa kelemahan sama ada dari segi penyediaan slaid mahupun kelemahan yang berpunca dari kesilapan manusia. Seajar dengan perkembangan dunia sains dan teknologi, kelemahan-kelemahan ujian palitan Pap ini telah mendorong kepada pembinaan sistem sokongan yang bertujuan bagi meningkatkan prestasi diagnosis ujian palitan Pap dan juga menyediakan peralatan sokongan kepada pengamal perubatan. Sistem-sistem yang dibangunkan melibatkan pengaplikasian teknik pemprosesan imej, kecerdikan buatan dan/atau gabungan kedua-dua teknik pemprosesan imej dan kecerdikan buatan.

Pengaplikasian teknik pemprosesan imej dalam permasalahan pengklasifikasian kanser pangkal rahim khususnya pengklasifikasian sel imej palitan Pap telah banyak dijalankan oleh penyelidik-penyelidik (Ghafar *et al.*, 2002; Mat-Isa *et al.*, 2002a; Mustafa *et al.*, 2007; Kale & Aksoy, 2010). Pelbagai teknik pemprosesan imej telah diaplikasikan yang menjurus kepada satu tujuan iaitu memudahkan proses pemeriksaan/pentafsiran sel pangkal rahim yang selama ini dilakukan secara manual oleh sitoteknologis di bawah pengawasan ahli patologi.

Kecerdikan buatan juga telah banyak digunakan untuk tujuan mengklasifikasi sel kanser pangkal rahim (Mitra *et al.*, 2000; Ramli *et al.*, 2004; Mustafa *et al.*, 2008). Setiap sistem yang dibangun berasaskan kecerdikan buatan ini telah menunjukkan prestasi ketepatan pengklasifikasian yang baik. Hal ini jelas menunjukkan keberkesanan kecerdikan buatan dalam memudahkan dan meringkaskan kerja-kerja dalam aplikasi yang digunakan.

Pengaplikasian gabungan teknik pemprosesan imej dan kecerdikan buatan turut mendapat perhatian para penyelidik terdahulu bagi membentuk satu sistem diagnosis yang lengkap (Li & Najarian, 2001; Srinivasan *et al.*, 2006; Schilling *et al.*, 2007; Mat-Isa *et al.*, 2008). Dalam kebanyakan penyelidikan, teknik pemprosesan imej telah digunakan sebagai alat prapemprosesan untuk tujuan meningkat kualiti imej bagi memudahkan proses perolehan maklumat. Kemudian, kecerdikan buatan digunakan sebagai agen pengkelasan. Data-data yang diperoleh dari teknik pemprosesan imej dijadikan sebagai masukan kepada struktur senibina kecerdikan buatan bagi melatih dan seterusnya membuat pengkelasan. Berdasarkan prestasi baik yang ditunjukkan dalam penyelidikan-penyelidikan tersebut, justeru, dalam penyelidikan ini gabungan kedua-dua teknik (iaitu pemprosesan imej dan kecerdikan buatan khususnya rangkaian neural) telah diaplikasikan dalam pembangunan sistem diagnosis peringkat prakanser pangkal rahim.

1.3 Pernyataan Masalah

Sehingga kini, proses pengklasifikasian imej terutamanya dalam bidang perubatan masih dilakukan secara manual. Pentafsiran dan pengklasifikasian imej pangkal rahim dilakukan oleh sitoteknologis bertauliah dengan diawasi oleh ahli patologi yang berpengalaman. Berdasarkan jumlah populasi ahli patologi yang semakin berkurangan, sebaliknya kes patologi pula semakin meningkat (disebabkan kempen penyaringan yang semakin giat dijalankan oleh kerajaan), bebanan yang dipikul oleh sitoteknologis dan ahli patologi adalah sangat berat. Hal ini boleh mewujudkan tekanan kerja yang boleh

mendorong kepada kebarangkalian yang tinggi bagi melakukan kesilapan akibat bekerja dalam suasana tertekan. Selain itu, disebabkan bilangan ahli patologi yang tidak seimbang dengan kes patologi yang diterima, proses mentafsir imej akan mengambil masa yang lama kerana seseorang ahli patologi dihadkan menyaring bilangan tertentu slaid dalam sehari bagi mengelakkan keletihan mata dan secara tidak langsung menjaga kejituan ujian. Hal ini menyebabkan pesakit perlu menunggu lama bagi mendapat keputusan. Jika pesakit didiagnos positif menghidapi kanser, kelewatan ini menyebabkan pesakit gagal menerima rawatan awal yang boleh mengurangkan risiko kematian.

Oleh itu, suatu sistem diagnosis secara automatik diperlukan sebagai penyelesaian kepada permasalahan tersebut. Pada masa kini terdapat banyak sistem mengklasifikasi atau mendiagnos peringkat prakanser pangkal rahim yang telah dibangunkan dan salah satu daripadanya ialah sistem NeuralPap yang dibangunkan oleh penyelidik Mat-Isa (2003) dan Mat-Isa *et al.* (2008). Sistem ini mengandungi dua bahagian iaitu pengekstrakan ciri secara automatik dan diagnosis pintar. Pada bahagian pertama, empat ciri-ciri sel pangkal rahim diekstrak menggunakan algoritma pengekstrakan ciri yang dibangunkan. Empat ciri ini dijadikan sebagai masukan kepada rangkaian neural (iaitu bahagian diagnosis pintar di bahagian kedua) bagi mengklasifikasi sel pangkal rahim kepada tiga peringkat iaitu normal, Lesi Intra-Epitelial Skuamus Gred Rendah (*Low-grade Squamous Intraepithelial Lesions, LSIL*) dan Lesi Intra-Epitelial Skuamus Gred Tinggi (*High-grade Squamous Intraepithelial Lesions, HSIL*). Walau bagaimanapun, sistem NeuralPap tersebut memproses imej pada skala kelabu sedangkan pemprosesan imej secara pewarnaan pseudo semakin digemari dalam aplikasi pengimejan perubatan (Rusev, 2003; Li *et al.*, 2007a; Jang & Ra, 2008; Lu *et al.*, 2009a; Zhanli *et al.*, 2009; Selvarasu *et al.*, 2010). Hal ini kerana sistem visual manusia adalah lebih sensitif kepada warna daripada imej monokrom.

Selain itu, algoritma peruasan yang digunakan dalam sistem NeuralPap adalah berdasarkan algoritma pengelompokan yang sering digunakan dalam literatur. Walaupun telah dibuktikan oleh penyelidik terdahulu (Mat-Isa *et al.*, 2002a; 2002b; 2002c) yang

algoritma ini berkesan bagi meruaskan imej palitan Pap dengan baik, namun masih terdapat ruang bagi diperbaiki. Hal ini kerana peruasan merupakan satu langkah penting untuk sistem diagnosis. Jika peruasan tidak dapat dilakukan dengan tepat, ciri-ciri sel pangkal rahim juga tidak dapat diekstrak dengan tepat dan seterusnya diagnosis yang tepat tidak akan dapat diperolehi.

Seterusnya, sistem NeuralPap juga dilengkapi dengan teknik penyingkiran hingar sebagai salah satu fasiliti sokongannya. Proses penyingkiran hingar ini tidak dimasukkan sebagai teknik prapemprosesan dalam sistem utama. Algoritma penapisan yang disediakan sebagai pilihan bagi menyingkir hingar imej palitan Pap sebelum proses selanjutnya diimplemen, tidak banyak membantu pengguna yang menggunakan sistem NeuralPap. Jika imej tercemar dengan hingar perlu diproses dan disingkirkan hingarnya, ia perlu dilakukan secara berasingan. Pengguna perlu mempunyai asas pemprosesan imej bagi menentukan sama ada proses penyingkiran hingar perlu dilakukan terlebih dahulu atau sebaliknya. Jika imej terus diproses tanpa disingkirkan hingarnya terlebih dahulu, ia akan menyumbang kepada ketidakjitian diagnosis.

Akhir sekali, sistem NeuralPap dibangunkan khas bagi mendiagnos sel pangkal rahim dalam keadaan terpalit secara lapisan tunggal sahaja. Jika imej yang ingin diperiksa itu terdiri daripada sel-sel yang berlapis atau bertindih, sistem NeuralPap tidak dapat meruaskan setiap sel dengan tepat. Ini akan menyumbang kepada kesilapan pengklasifikasian yang secara tidak langsung akan mengurangkan sensitiviti sistem.

1.4 Objektif Penyelidikan

Objektif utama penyelidikan ini ialah bagi membangunkan satu sistem diagnosis peringkat prakanser pangkal rahim secara berbantu-komputer berasaskan teknik pemprosesan imej dan rangkaian neural. Sistem ini bertujuan bagi mengurangkan kekangan yang terdapat dalam sistem NeuralPap melalui penambahbaikan kepada proses

pengekstrakan ciri dan pengklasifikasiannya bagi meningkatkan prestasi sistem. Objektif utama ini dipecahkan kepada beberapa objektif sokongan seperti berikut iaitu untuk:

- i. Mengkaji algoritma peruasan baru yang mampu melaksanakan peruasan pada imej palitan Pap dengan lebih baik.
- ii. Mengkaji algoritma peruasan dengan ciri penyingkiran hingar terbenam yang mampu meruaskan imej palitan Pap yang tercemar dengan hingar tanpa perlu melalui proses penurasan hingar.
- iii. Mengkaji algoritma pengekstrakan ciri secara pewarnaan pseudo bagi meningkatkan kesan visual.
- iv. Mengkaji kaedah bagi mengekstrak ciri sel-sel yang bertindih.
- v. Mengkaji algoritma penetapan pusat rangkaian untuk Rangkaian Fungsi Asas (*Radial Basis Function, RBF*) dan Fungsi Asas Jejarian Hibrid (*Hybrid Radial Basis Function, HRBF*).

1.5 Skop Penyelidikan

Seperti yang telah dinyatakan, tujuan utama penyelidikan ini ialah bagi membina satu sistem yang berupaya mendiagnos peringkat prakanser pangkal rahim secara berbantu-komputer berasaskan teknik pemprosesan imej dan rangkaian neural. Penyelidikan ini terhad kepada pemprosesan ke atas imej sel pangkal rahim yang diperolehi dari slaid palitan Pap. Slaid palitan Pap diperolehi dari Jabatan Patologi, Hospital Universiti Sains Malaysia (HUSM), Kota Bharu, Kelantan, yang terdiri daripada kes normal, LSIL dan HSIL. Imej sel pangkal rahim dirakam oleh sitoteknologis yang bertaualiah dengan diawasi dan disahkan oleh pakar patologi. Sebanyak 799 sel pangkal rahim dari ketiga-tiga kes tersebut telah dirakamkan. Imej dirakam dalam bentuk imej skala kelabu dan disimpan dalam format *bmp*. Pengklasifikasian sel-sel pangkal rahim terhad kepada tiga kelas sahaja iaitu normal, LSIL dan HSIL.

1.6 Garis Panduan Tesis

Tesis ini mengandungi 5 bab termasuk bab pengenalan. Setiap bab akan dimulakan dengan pendahuluan yang akan menerangkan secara ringkas objektif setiap bab. Ringkasan atau kesimpulan akan dibuat pada akhir setiap bab bagi menyimpulkan keseluruhan informasi yang dibentangkan dalam bab tersebut. Penyusunan bab-bab ini adalah mengikut langkah pelaksanaan penyelidikan ini.

Tesis ini bermula dengan Bab 1 iaitu pengenalan yang menerangkan secara ringkas tentang penyelidikan ini. Bab ini akan membantu pembaca supaya mendapat idea kasar atau gambaran awal mengenai penyelidikan ini. Objektif dan skop penyelidikan serta garis panduan tesis dimuatkan dalam bab ini.

Seterusnya Bab 2 iaitu kajian ilmiah, akan menerangkan tentang jenis-jenis dan peringkat kanser pangkal rahim, klasifikasi sel-sel pangkal rahim mengikut sistem piawai serta ujian palitan Pap yang sering digunakan sebagai ujian saringan. Kelemahan dan batasan ujian palitan Pap, yang mendorong kepada pembangunan sistem sokongan sama ada bagi meningkatkan kualiti imej (bagi memudahkan pentafsiran secara visual) atau meningkatkan prestasi pengklasifikasian (bagi mengautomasikan pengklasifikasian) turut dibentangkan dalam bab ini. Dalam bab ini, diterangkan juga tentang sistem NeuralPap yang menjadi tunjang kepada penyelidikan ini memandangkan objektif penyelidikan ini ialah membuat penambahbaikan kepada sistem NeuralPap.

Bab 3 pula beralih kepada metodologi pelaksanaan penyelidikan ini. Dalam bab ini, algoritma-algoritma baru untuk peruasan imej, teknik baru pengekstrakan imej serta penambahbaikan pengklasifikasian menggunakan rangkaian neural yang dicadangkan dalam penyelidikan ini akan diterangkan. Setiap algoritma yang dibincangkan, diletak dalam sub-sub bahagian yang diwakili olehnya. Persamaan matematik dan carta alir yang menyokong algoritma tersebut juga akan dibentangkan.

Bab 4 merupakan bab untuk keputusan. Bab ini akan membentangkan keputusan untuk imej sel pangkal rahim yang diproses menggunakan algoritma-algoritma yang telah dibincangkan dalam Bab 3. Sebelum algoritma-algoritma tersebut diuji dengan imej sel pangkal rahim, prestasi setiap algoritma tersebut diuji terlebih dahulu dengan imej piawai dan keputusan pemprosesan imej terhadap kedua-dua jenis imej akan dibentangkan. Seterusnya, keputusan pengekstrakan ciri dan prestasi pengklasifikasian akan turut dibentangkan.

Tesis ini diakhiri dengan Bab 5 yang merupakan bab penutup. Ia merangkumi kesimpulan yang dapat dibuat daripada penyelidikan yang telah dijalankan, perkara-perkara yang perlu ditangani sepanjang pelaksanaan penyelidikan dan juga cadangan untuk penyelidikan pada masa akan datang.

BAB 2

KAJIAN ILMIAH

2.1 Pendahuluan

Analisis imej merupakan aplikasi yang lazim digunakan di dalam bidang perubatan terutamanya dalam bidang sitologi yang mana pemeriksaan sel-sel dan tisu-tisu dilakukan secara mikroskopik. Pentafsiran imej mikroskopik secara visual adalah sangat rumit dan kerap kali mendorong kepada kecenderungan untuk berlakunya kesilapan. Sejalan dengan perkembangan dunia sains dan teknologi komputer, terdapat pelbagai percubaan yang telah dilakukan bagi menggantikan pemeriksaan visual manusia dengan analisis berkomputer atau sekadar sebagai bantuan tambahan kepada teknik pemeriksaan yang sedia ada (Mat-Noor *et al.*, 2008; Lu *et al.*, 2009b; Wang *et al.*, 2009; Park *et al.*, 2011). Penyelidikan-penyelidikan tersebut bertujuan bagi menjadikan tugas-tugas penyaringan secara visual bersifat automatik dan meminimumkan kebergantungan kepada pakar patologi dan juruteknik sitologi bagi mentafsir imej-imej mikroskopik dan seterusnya membuat keputusan.

Kanser pangkal rahim yang juga dikenali sebagai kanser serviks adalah penyakit yang digeruni oleh wanita. Setiap tahun, beribu-ribu wanita menderita akibat penyakit ini. Di Malaysia, lebih 1,100 orang wanita telah didiagnos menghidap kanser ini setiap tahun (Persatuan Kebangsaan Kanser Malaysia, 2011). Risiko bagi menghidap penyakit ini meningkat apabila wanita mencecah umur 30 tahun dan berisiko maksima apabila wanita mencecah umur antara 60 hingga 69 tahun (Persatuan Kebangsaan Kanser Malaysia, 2011). Menyedari risiko yang agak tinggi dan kadar kematian wanita akibat kanser pangkal rahim yang semakin meningkat dari hari ke hari, pelbagai kajian telah dilakukan bagi membantu mengurangkan kadar kematian. Ujian saringan palitan Pap telah terbukti berjaya mengurangkan kadar kematian secara signifikan sejak ia diperkenalkan sekitar tahun 1940-

an. Sejak itu, beberapa penyelidikan telah dijalankan yang memfokus kepada pencarian teknik penyaringan termasuk teknik penyediaan slaid yang lebih berkualiti dan berkesan supaya kanser dapat dikesan dari peringkat yang lebih awal agar pesakit masih dapat diselamatkan melalui penyediaan rawatan yang bersesuaian (Abulafia *et al.*, 2003; Stein, 2003; El-Tawil *et al.*, 2008). Pada masa yang sama, penyelidik-penyelidik yang lain memfokus kepada cara bagi mempertingkatkan dan memperbaiki kualiti imej-imej palitan Pap melalui pengaplikasian teknik-teknik pemprosesan imej agar proses penyaringan dapat dipermudahkan (Ghafar *et al.*, 2002; Mat-Isa *et al.*, 2002c; Suryatenggara *et al.*, 2009). Selain itu, terdapat juga penyelidik yang memfokus kepada pembangunan sistem sokongan bagi mendiagnos kanser pangkal rahim supaya penyaringan dapat dijalankan secara automatik (Nemec *et al.*, 2002; Wang *et al.*, 2009; Lei He *et al.*, 2010). Kaedah ini dikenali sebagai kaedah berbantu-komputer. Melalui kaedah ini, sistem dibangunkan agar mempunyai pengetahuan mentafsir dan menilai data-data yang menyerupai pengetahuan ahli patologi dan berupaya membuat keputusan yang setanding dengan ahli patologi.

Berlandaskan pernyataan-pernyataan berkaitan keberkesanan teknik-teknik berasaskan pemprosesan imej dan berbantu-komputer dalam membantu ahli patologi menyaring kanser pangkal rahim, secara umumnya Bab 2 ini akan membentangkan kajian ilmiah berkenaannya. Bab ini dimulakan dengan perbincangan mengenai kanser pangkal rahim yang menyentuh tentang penerangan ringkas berkenaan statistik, tahap serta rawatan kanser ini. Kemudian, perbincangan mengenai sistem-sistem sokongan yang diaplikasikan dalam bidang perubatan bagi membantu ahli-ahli perubatan akan diberikan. Ia turut merangkumi penerangan berkaitan pengekstrakan, pengklasifikasian dan kecerdikan buatan yang menjadi asas kepada pembangunan sesuatu sistem sokongan. Seterusnya, Bab 2 ini akan menyentuh secara khusus sistem NeuralPap yang merupakan tunjang kepada penyelidikan yang dilakukan. Pada akhir bab ini, ringkasan keseluruhan Bab 2 akan diberikan.

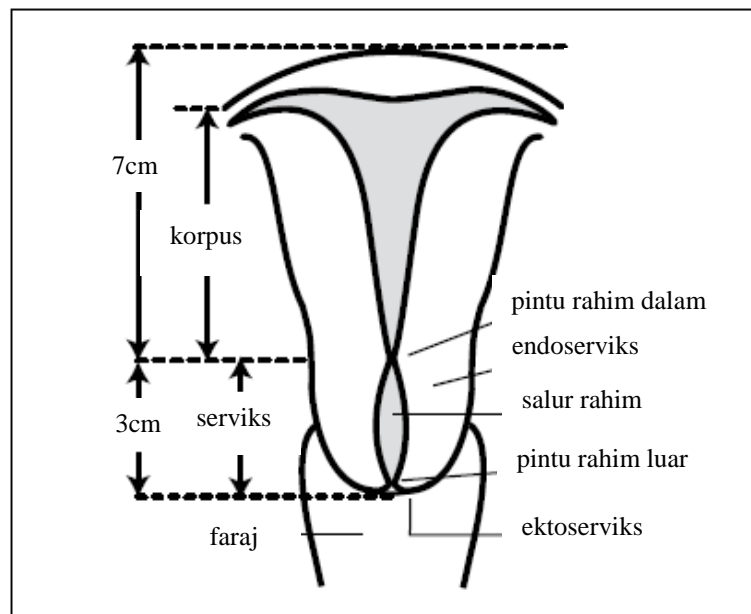
2.2 **Kanser Pangkal Rahim**

Kanser pangkal rahim atau kanser serviks adalah penyakit yang digeruni oleh kaum wanita. Ia merupakan kanser ketiga paling ketara di kalangan wanita selepas kanser payudara dan kanser kolorektum (Borsotti *et al.*, 1998). Di serata dunia, kira-kira 12% daripada semua kes kanser yang menyerang wanita ialah kanser serviks (Park *et al.*, 2011). Kanser ini meragut 270,000 nyawa wanita setahun di mana ini bersamaan dengan kematian seorang wanita untuk setiap dua minit (Jabatan Penerangan Malaysia, 2010). Statistik dari Pertubuhan Kesihatan Dunia, WHO juga menunjukkan pada setiap tahun, kadar kematian akibat kanser serviks ini adalah lebih tinggi berbanding kadar kematian wanita semasa melahirkan anak.

Di Malaysia, statistik menunjukkan 12.2 daripada setiap 100,000 populasi wanita, telah didiagnos menghidap kanser serviks (Jabatan Penerangan Malaysia, 2010). Menurut laporan WHO pada 2007, lebih 700 wanita Malaysia mati kerana kanser serviks setiap tahun (Jabatan Penerangan Malaysia, 2010). Menurut anggaran Persatuan Kanser Kebangsaan Malaysia, terdapat 1,500 pesakit kanser serviks atau empat hingga lima wanita Malaysia yang menghidap kanser ini sehari. Anggaran ini termasuk bilangan wanita yang baru disahkan menghidap penyakit dan sedang menjalani rawatan (Kamalludeen, 2010). Angka-angka ini menunjukkan kejadian kanser pangkal rahim di negara ini masih tinggi berbanding sesetengah negara maju seperti Amerika Syarikat. Contohnya, dari perangkaan tahun 2002, Malaysia mencatatkan 17.8 kes berbanding Amerika Syarikat yang hanya mencatatkan 4.5 kes untuk setiap 100,000 penduduk (Mohd-Yusof, 2011). Paling menyedihkan, daripada bilangan itu, 50 peratus pesakit ini langsung tidak mempunyai peluang bagi meneruskan kehidupan kerana kanser dikesan pada peringkat yang sudah parah.

Kanser ini dikenali sebagai kanser pangkal rahim kerana ketidaknormalan sel-sel berlaku di bahagian pangkal rahim. Pangkal rahim ialah bahagian sempit yang terletak di bawah rahim. Kanser pangkal rahim berkembang apabila sel-sel tidak normal di lapisan

epitelium pangkal rahim mula membahagi dengan tidak terkawal dan membentuk lesi-lesi prakanser. Jika tidak dikesan, lesi-lesi ini akan membentuk tumor dan merebak ke tisu-tisu berdekatan lalu membentuk kanser (Guard Against Cervical cancer, 2011). Kanser pangkal rahim masih boleh dirawat jika ia dikesan pada peringkat awal. Hal ini kerana, sebelum sampai ke tahap kanser, kanser pangkal rahim melalui fasa prakanser untuk tempoh yang lama iaitu selalunya antara 10 hingga 15 tahun (QIAGEN, 2011). Malah, organ serviks atau pangkal rahim, seperti yang ditunjukkan dalam Rajah 2.1, merupakan organ yang mudah bagi dipantau atau diperiksa (Mohd-Zain, 2011).



Rajah 2.1: Anatomi sistem pembiakan wanita

Jangkitan dari jenis tertentu virus papilloma manusia, *HPV* dianggap sebagai faktor utama penyebab kanser serviks. Sekurang-kurangnya sekali sepanjang hayat, seseorang wanita mungkin dijangkiti dengan salah satu dari dua jenis virus HPV yang telah dikenalpasti sebagai virus berisiko tinggi iaitu jenis 16 dan 18 yang boleh menyebabkan kanser. Kedua-dua jenis virus ini menyumbang kira-kira 70% daripada keseluruhan kejadian kanser pangkal rahim yang disebabkan oleh virus HPV (Guard Against Cervical cancer, 2011; Mohd-Yusof, 2011). Selebihnya adalah disebabkan oleh virus HPV jenis lain. Pada kebiasaannya, sistem pertahanan semulajadi badan berupaya bagi melawan jangkitan

tersebut menyebabkan kebanyakan wanita tidak mengalami sebarang masalah atau komplikasi. Namun, untuk sesetengah wanita jangkitan itu tidak sembuh dan menyebabkan virus tersebut kekal aktif di dalam badan. Jika virus tersebut kekal aktif di dalam badan untuk tempoh yang lama, ia boleh menyebabkan berlakunya perubahan pada sel-sel dan tisu-tisu pangkal rahim dan meningkatkan risiko untuk tisu bertukar menjadi kanser (Mohd-Yusof, 2011; QIAGEN, 2011). Perkembangan kanser ini mungkin terjadi dalam tempoh sebulan atau setahun selepas jangkitan virus (Motoyama *et al.*, 2004).

Kunci kepada pencegahan kanser ini ialah melalui pengesanan awal kehadiran sel-sel prakanser dalam pangkal rahim. Salah satu cara bagi mengesan kehadiran sel-sel ini ialah melalui ujian palitan Papanicolaou atau ringkasnya Pap. Ujian palitan Pap pada pangkal rahim ini perlu dilakukan setiap 2 hingga 3 tahun sekali bagi mengesan kehadiran sel prakanser. Sekiranya terdapat sel prakanser, rawatan setempat pada pangkal rahim boleh dilakukan dan pesakit masih mempunyai peluang yang sangat tinggi bagi sembuh sepenuhnya. Palitan Pap merupakan kaedah yang berkesan dan merupakan alat pengesanan awal yang penting bagi mengesan sebarang perubahan sel-sel di bahagian pangkal rahim. Walaupun kaedah ini tidak berupaya bagi mengesan semua perubahan prakanser atau kanser tetapi kadar risiko menghadapi kanser pangkal rahim dapat dikurangkan lima kali ganda berbanding wanita yang langsung tidak menjalani pemeriksaan palitan Pap (Jabatan Penerangan Malaysia, 2010).

2.2.1 Jenis-jenis dan Peringkat-peringkat Kanser Pangkal Rahim

Ketidaknormalan sel pangkal rahim terbahagi kepada dua iaitu Atipia dan Displasia. Ketidaknormalan jenis Atipia merujuk kepada sedikit perubahan pada sel dan ia biasanya boleh dikesan melalui ujian palitan Pap. Perubahan pada sel ini mungkin menjadi bertambah teruk atau mungkin boleh pulih seperti biasa. Jika keputusan ujian palitan Pap menunjukkan terdapatnya perubahan atipikal, wanita itu akan dinasihatkan supaya mengulangi ujian

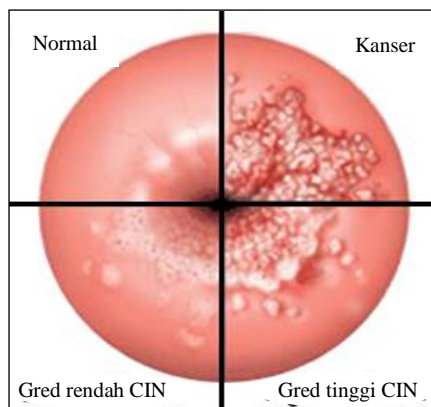
palitan Pap dalam masa enam bulan lagi bagi melihat sama ada perubahan sel atipikal masih wujud atau sudah tiada. Manakala ketidaknormalan jenis Displasia pula ialah sejenis perubahan luar biasa dalam sel di pangkal rahim. Dalam erti kata lain, terma ini merujuk kepada fasa awal pertumbuhan tidak normal sel pangkal rahim yang mungkin boleh berkembang menjadi kanser. Ia juga dikenali sebagai Neoplasia Intraepitelial Serviks (*Cervical Intraepithelial Neoplasia, CIN*) dan Lesi Intraepitelial Skuamus (*Squamous Intraepithelial Lesions, SIL*) (MyHealth, 2011).

Terdapat dua jenis kanser serviks utama yang dinamakan berdasarkan jenis sel tempat kanser bermula (MyHealth, 2011; Persatuan Kebangsaan Kanser Malaysia, 2011).

- i. Karsinoma sel skuamus (*Squamous cell carcinoma*): kanser ini bermula di dalam skuamus atau sel seperti lapisan kulit kepada serviks. Kanser ini merupakan kanser yang paling lazim didiagnos iaitu kira-kira 80% daripada semua kes kanser serviks.
- ii. Adenokarsinoma (*Adenocarcinoma*): kanser ini bermula pada bahagian atas serviks dan merupakan kanser yang sukar didiagnos dengan ujian palitan Pap. Ia berkembang dari sel-sel glandular dan merupakan kanser yang kurang lazim didiagnos.

Seperti yang dijelaskan sebelum ini, displasia pangkal rahim adalah fasa awal pertumbuhan tidak normal sel. Sel-sel displastik kelihatan seperti sel kanser namun ia tidak dianggap sebagai sel malignan kerana kebiasaannya lesi (*lesion*) awal yang dikenali sebagai Neoplasia Intraepitelial Serviks (*CIN*) ini kekal di lapisan atas serviks dan tidak merebak ke kawasan tisu yang sihat. Keadaan ini juga dikenali sebagai keadaan prakanser iaitu keadaan di mana sel-sel tidak normal hanya berada pada lapisan epitalial sahaja. CIN akan menunjukkan perkembangan apabila bilangan sel-sel tidak normal meningkat. CIN boleh bertukar menjadi kanser malignan apabila sel-sel neoplastik telah berkembang melampaui dasar membran epitelium, memasuki tisu-tisu berdekatan dan merebak ke tisu-tisu serta organ-organ yang lain (Wang *et al.*, 2009).

Diagnosis dan pengredan CIN adalah berdasarkan interpretasi secara visual ciri-ciri sitologi sel pangkal rahim. Pada peringkat awal, sistem WHO telah digunakan bagi mengelaskan ciri-ciri sitologi sel pangkal rahim semasa proses diagnosis (Nemec *et al.*, 2002). Sistem ini membahagikan displasia pangkal rahim kepada peringkat permulaan atau ringan, peringkat pertengahan dan peringkat parah. Manakala Kanser Pangkal Rahim in-Situ (*Cervical Cancer in-Situ, CIS*) telah dikelaskan kepada kategori yang berasingan. Pada tahun 1980-an, sistem Richard telah digunakan bagi mengelaskan lesi-lesi pangkal rahim. Sistem ini mengelaskan lesi-lesi pangkal rahim kepada tiga gred CIN iaitu, CIN I, CIN II dan CIN III. CIN I dan CIN II masing-masing adalah displasia permulaan dan pertengahan manakala CIN III pula menggabungkan displasia parah dan CIS. Sistem ini masih lagi digunakan untuk klasifikasi histologikal. Rajah 2.2 menunjukkan keadaan pangkal rahim yang berubah dari peringkat normal sehingga peringkat kanser.

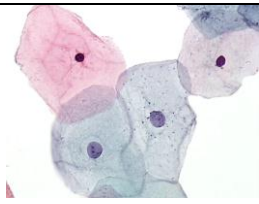

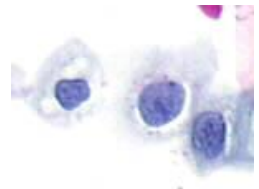
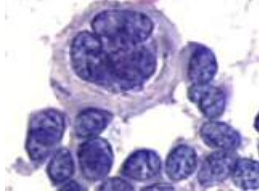


Rajah 2.2: Perubahan keadaan pangkal rahim mengikut peringkat kanser (QIAGEN, 2011)

Pada tahun 1994, satu sistem baru telah dirangka bagi memudahkan diagnosis sitologi. Sistem ini dikenali sebagai sistem Bethesda. Dalam sistem Bethesda, hanya dua kategori untuk lesi intra-epitelial skuamus, SIL dipertimbangkan iaitu Lesi Intra-Epitelial Skuamus Gred Rendah (*Low-grade Squamous Intraepithelial Lesions, LSIL*) dan Lesi Intra-Epitelial Skuamus Gred Tinggi (*High-grade Squamous Intraepithelial Lesions, HSIL*).

Kategori LSIL merujuk kepada kategori CIN I dan juga merangkumi perubahan sel-sel akibat dari jangkitan HPV. Kategori HSIL pula merujuk kepada gabungan CIN II dan CIN III. Walaupun ketiga-tiga sistem ini mengelaskan setiap ciri secara berbeza, namun untuk setiap kumpulan yang digabungkan dalam kategori tersebut, prognosis yang diberikan adalah tetap sama. Jadual 2.1 menunjukkan klasifikasi ujian palitan Pap berdasarkan sistem Bethesda serta contoh imej sel yang dikategorikan dalam kelas tersebut.

Jadual 2.1: Sistem Bethesda

Klasifikasi	Keterangan	Imej
Negatif	Normal	
LSIL	Ujian palitan Pap menunjukkan displasia sedikit. Ketidaknormalan sel terhad kepada 1/3 lapisan epitelium.	
HSIL	Ujian palitan Pap menunjukkan displasia sederhana dan teruk. Sel tidak normal sudah melangkaui 1/2 lapisan epitelium sehingga keseluruhan lapisan epitelium terdiri daripada sel-sel tidak normal namun, sel masih belum merebak ke lapisan bawah.	
Kanser	Ujian palitan Pap menunjukkan kanser. Sel kanser sudah merebak ke lapisan bawah pangkal rahim.	

Dalam menganalisa sel-sel pangkal rahim, ahli sitopatologi membezakan antara sel-sel normal dan tidak normal dengan melihat kepada ciri-ciri sel-sel tersebut berdasarkan beberapa morfologi. Antara morfologi-morfologi itu ialah:

- i. Sel-sel yang berubah dari keadaan normal kepada kanser akan mempunyai nukleus yang besar. Jadi, nisbah luas kawasan nukleus kepada luas kawasan sitoplasma untuk sel-sel kanser adalah tinggi berbanding dengan nisbah luas kawasan nukleus kepada luas kawasan sitoplasma untuk sel-sel normal (Govan *et al.*, 1986; Crum, 1994).
- ii. warna nukleus untuk sel-sel tidak normal (LSIL atau HSIL) adalah lebih gelap berbanding dengan warna nukleus untuk sel-sel normal (WebMD, 2011).

Dalam situasi tertentu, penentuan sel pangkal rahim tidak normal kadangkala boleh terabai. Hal ini kerana lazimnya ketepatan ujian palitan Pap sangat bergantung kepada kualiti sampel palitan yang diperolehi. Ada beberapa faktor yang menyumbang kepada kualiti sampel palitan Pap yang rendah seperti pendedahan berlebihan atau kurang pendedahan kepada cahaya mikroskop semasa proses penyediaan slaid atau sampel itu sendiri yang berlumuran darah atau keputihan dan lain-lain yang boleh menyebabkan kesulitan dalam penyediaan slaid. Oleh itu, ahli sitopatologi mungkin menghadapi kesulitan dalam mengekstrak morfologi penting dari sampel.

2.2.2 Ujian Saringan Kanser Pangkal Rahim

Ujian palitan Pap ialah ujian saringan yang popular bagi mengesan kewujudan sel-sel kanser pada pangkal rahim. Ujian ini menjadi pilihan kerana kosnya yang rendah serta tidak membebankan, ringkas, cepat dan tidak menyakitkan. Malah, sejak dari awal pengenalannya, ujian ini dilihat berkesan menurunkan kadar kematian wanita. Sebagai contoh, di Amerika Syarikat kadar penurunan direkodkan antara 70% hingga 80% manakala

di negara-negara mundur yang tidak menjadikan ujian palitan Pap sebagai ujian saringan asas telah mencatatkan peningkatan sehingga 80% kes kanser pangkal rahim (Stein, 2003). Ini membuktikan ujian palitan Pap merupakan satu kaedah yang efektif bagi menyaring kanser pangkal rahim.

Walaupun ujian ini telah menjadi pilihan pengamal perubatan sejak lebih dari separuh abad, terdapat beberapa kelemahan ujian ini yang kadangkala menyebabkan ujian tersebut menjadi kurang efektif sehingga dikhuatiri mampu mempengaruhi keputusan ujian saringan. Antara kelemahan ujian palitan Pap yang paling mendapat perhatian ialah isu ketepatannya. Walaupun dilaporkan sensitiviti ujian palitan Pap ialah 89% dan spesifisiti 98% (Stein, 2003), namun, salah negatif untuk ujian palitan Pap ini begitu tinggi iaitu sehingga 50% (Sprenger *et al.*, 1996). Isu salah negatif ini telah dikenalpasti berpunca dari kesalahan pensampelan termasuk semasa proses penyediaan sampel, kehadiran bendasing seperti darah dan mukus, penyaringan dan kesalahan pentafsiran (Abulafia *et al.*, 2003).

Berdasarkan kelemahan ujian palitan Pap konvensional, ujian saringan alternatif yang menggunakan teknologi berasaskan cecair telah diperkenalkan. Tujuan ujian alternatif ini ialah bagi memperbaiki kualiti spesimen pangkal rahim dan meningkatkan pengesanan ketidaknormalan serviks atau dalam kata lain mengurangkan kadar salah negatif ujian palitan Pap konvensional (Abulafia *et al.*, 2003). Dalam ujian palitan Pap konvensional, sampel diperoleh dari zon transisi dengan menggunakan alat pensampelan (iaitu spatula kayu Ayre atau berus endo-serviks) dan kemudian sampel sel dipalitkan ke atas slaid kaca mikroskop. Kemudian, alat pensampelan akan dilupuskan. Sebahagian sampel sel pesakit mungkin telah hilang semasa alat pensampelan dilupuskan dan bendasing seperti darah dan mukus turut sama dipalitkan di atas slaid dan ini akan merencatkan proses diagnosis (GroupHealth, 2010).

Berbeza dengan ujian yang menggunakan teknologi berasaskan cecair di mana sampel diperoleh menggunakan berus sito. Hujung berus yang mengandungi sampel diasingkan dan

diletakkan ke dalam bekas yang mengandungi agen pengekal. Kemudian sampel dibawa ke makmal untuk disingkirkan bendasing yang tidak diperlukan. Akhir sekali, dari sampel yang telah diasingkan semua bahan yang tidak diperlukan, hanya sel pangkal rahim lapisan tunggal yang homogen diletakkan di atas slaid kaca. Dengan ini proses pentafsiran dapat dilaksanakan dengan tepat tanpa direncatkan oleh kehadiran bendasing-bendasing yang mungkin menutupi maklumat penting yang terkandung di dalam slaid. Antara ujian saringan yang menggunakan teknologi ini ialah ThinPrep dan AutoCyte. ThinPrep telah diluluskan oleh Pentadbiran Makanan dan Dadah Amerika Syarikat (*United State of America Food and Drug Administration, FDA*) pada tahun 1996 manakala AutoCyte telah diluluskan pada tahun 1999 (GroupHealth, 2010).

Sejak itu, beberapa penyelidikan telah dijalankan bagi meninjau sejauh mana keberkesanan teknologi berasaskan cecair ini. Tidak dinafikan yang teknologi ini menyediakan slaid sampel pangkal rahim yang lebih baik berbanding teknik konvensional yang dapat memperbaiki proses penilaian dan pentafsiran. Beberapa kajian telah membuktikan yang ThinPrep adalah lebih baik dari ujian palitan Pap konvensional berdasarkan peningkatan sehingga kira-kira 140% kadar pengesanan LSIL (Bolick & Hellman, 1998; Díaz-Rosario & Kabawat, 1999; Guidos & Selvaggi, 1999), sensitiviti ujian ThinPrep dianggarkan kira-kira 80% (Hutchinson *et al.*, 1994; Ferenczy *et al.*, 1996) serta pengurangan salah negatif ThinPrep sebanyak tiga hingga empat kali berbanding ujian palitan Pap konvensional (Hutchinson *et al.*, 1994; Ferenczy *et al.*, 1996).

Walau bagaimanapun, kajian terkini mendapati tiada bukti yang menunjukkan teknologi berasaskan cecair ini berupaya mengurangkan kadar kematian akibat kanser pangkal rahim dan juga tiada bukti yang menunjukkan teknologi berasaskan cecair ini telah mendiagnos lebih banyak kanser pangkal rahim (Albertus *et al.*, 2009; GroupHealth, 2010). Malah, kajian Albertus *et al.* (2009) telah menyimpulkan bahawa ujian penyaringan dengan teknologi berasaskan cecair tidak dapat dibuktikan lebih sensitif dan juga lebih spesifik berbanding ujian palitan Pap konvensional dalam mengesan CIN atau kanser. Beberapa

kajian terkini juga turut mengesahkan bahawa teknologi berasaskan cecair memberikan sensitiviti dan spesifisiti yang sama dengan ujian palitan Pap konvensional (Ronco *et al.*, 2007; Arbyn *et al.*, 2008; Wright *et al.*, 2010).

2.2.3 Sistem Sokongan

Kelemahan-kelemahan ujian palitan Pap telah mendorong kepada pengenalan ujian saringan alternatif seperti yang telah dibincangkan dalam Bahagian 2.2.2. Namun, telah dikenalpasti oleh kajian-kajian terkini bahawa ujian-ujian ini tidak memberi peningkatan yang signifikan dari segi sensitiviti dan spesifisiti mahupun kemampuan bagi mendiagnos lebih banyak kes kanser pangkal rahim. Hal ini telah mendorong kepada penyelidikan bagi membina sistem sokongan yang dapat meningkatkan ketepatan dan kejituan ujian palitan Pap di samping bertujuan bagi mengurangkan kos ujian palitan Pap melalui pengurangan kebergantungan kepada ahli patologi dan mempercepatkan proses penyaringan.

Ujian palitan Pap konvensional sangat bergantung kepada khidmat ahli patologi. Kepakaran ahli patologi terlatih dan berpengalaman sangat diperlukan bagi mengesahkan sebarang penemuan sel tidak normal oleh sitoteknologis. Kebergantungan ini menyebabkan peningkatan kepada kos ujian palitan Pap ditambah pula dengan senario semasa di Malaysia yang kekurangan ahli patologi dan bilangan ujian yang tinggi yang perlu dilaksanakan di makmal patologi. Menurut statistik pada tahun 2007, bilangan ahli patologi di Malaysia ialah hanya seramai 300 orang (Utusan Malaysia, 2007) dengan nisbah patologi kepada populasi penduduk sebanyak 1:75,000 (PathWest Laboratory Medicine WA, 2008) manakala bilangan ujian yang dilaksanakan di makmal patologi adalah sebanyak 240 juta iaitu purata kira-kira 650,000 ujian sehari (Utusan Malaysia, 2007). Keadaan menjadi lebih meruncing ketika kempen penyaringan secara besar-besaran dijalankan bagi menggalakkan wanita menjalani ujian saringan. Ini akan menambahkan lagi bebanan yang sedia wujud di makmal patologi. Dengan kekurangan ahli patologi berpengalaman, pertambahan bilangan ujian palitan Pap

tidak mampu bagi ditangani menyebabkan keputusan diagnosis akan mengambil masa yang lebih lama.

Maka, penggunaan sistem sokongan untuk diagnosis kanser pangkal rahim dilihat sebagai satu cara bagi meningkatkan produktiviti di makmal sitologi dan menjadi jalan penyelesaian yang berpotensi bagi mengatasi masalah bebanan yang dihadapi ahli patologi dan sitoteknologis ketika melakukan penyaringan secara manual. Antara contoh sistem saringan automatik berbantuan teknologi komputer yang digunakan dalam beberapa penyelidikan ialah AutoPap Primary Screening System dari *TriPath Imaging*, ThinPrep® 2000 dari *Cytoc Corporation*, AutoCyte PREP® dari *AutoCyte Inc.* dan PapNet® dari *neuromedical Systems* (Lee *et al.*, 1998; Rosenthal, 1998; Wilbur *et al.*, 1998; Chang *et al.*, 2002). Penyelidik-penyelidik ini mendapati bahawa penggunaan sistem-sistem sokongan dalam melakukan penyaringan secara automatik dengan berbantuan teknologi komputer telah berjaya meningkatkan ketepatan proses saringan pangkal rahim dengan merekodkan salah negatif 40 kali lebih rendah berbanding proses manual oleh sitoteknologis (Lee *et al.*, 1998) serta terdapat peningkatan yang signifikan kepada produktiviti di makmal sitologi (Lee *et al.*, 1998; Rosenthal, 1998; Wilbur *et al.*, 1998; Chang *et al.*, 2002). Selain dari sistem-sistem sokongan yang telah dibincangkan, terdapat pelbagai lagi penyelidikan yang menggunakan kecerdikan buatan dan teknik pemprosesan imej sebagai sistem sokongan bagi mendiagnos kanser pangkal rahim pada peringkat awal. Kajian ilmiah berkenaan sistem-sistem ini akan dibincangkan secara terperinci di dalam Bahagian 2.3.

2.3 Sistem Sokongan Pengklasifikasian

Peranan komputer dalam kehidupan seharian semakin berkembang pesat. Hal ini kerana komputer telah terbukti mampu membantu memperkemaskan pengurusan dalam banyak urusan. Kemajuan dalam teknologi berasaskan komputer akhir-akhir ini telah banyak mempengaruhi perkembangan beberapa bidang lain, di mana fenomena ini turut

mempengaruhi bidang kesihatan dan perubatan. Penggunaan komputer di hospital dan klinik seperti sistem inventori dan belian ubat-ubatan, serta sistem penyimpanan rekod pesakit telah banyak menyelesaikan masalah pengurusan maklumat di hospital dan klinik (Azra, 2011). Selain itu, pengaplikasian sistem sokongan bagi mendiagnos atau mengklasifikasikan penyakit juga semakin diberi perhatian. Sistem sokongan diagnosis ialah program komputer yang dibangunkan bagi membantu pakar perubatan dalam menentukan penyakit yang dihidapi pesakit atau membentuk pelan diagnosis dan rawatan yang patut diberikan kepada pesakit. Sistem ini semakin mendapat perhatian berdasarkan kemampuannya yang seringkali dapat memberikan keputusan diagnosis yang tepat dan cepat. Ternyata, perkembangan sistem berasaskan komputer dalam bidang perubatan berupaya memberi penyelesaian kepada masalah-masalah pengurusan yang remeh dan rumit malah melalui perkembangan ini juga kualiti hidup manusia dapat diperbaiki di samping dapat menjimatkan masa, kos dan tenaga.

Pelbagai penyelidikan yang mengaplikasikan pelbagai kaedah, algoritma mahupun teknik telah dijalankan bagi menyediakan sistem sokongan diagnosis yang boleh digunakan dalam bidang perubatan. Terdapat pelbagai cabang penyelidikan bagi membangunkan sistem sokongan diagnosis, antaranya sistem diagnosis berdasarkan teknik pemrosesan imej, sistem diagnosis berdasarkan kecerdikan buatan dan juga sistem diagnosis berdasarkan gabungan kedua-dua teknik kecerdikan buatan dan pemrosesan imej. Penerangan terperinci untuk setiap cabang penyelidikan ini akan dibincangkan dengan terperinci dalam bahagian yang berikutnya.

2.3.1 Sistem Pengklasifikasian Berdasarkan Teknik Pemrosesan Imej

Pemrosesan imej digital ialah satu subset domain elektronik di mana imej ditukar kepada satu tatasusun integer-integer kecil yang dikenali sebagai piksel. Setiap satu piksel akan membawa ciri tertentu pada kedudukannya di dalam imej. Nilai integer yang dibawa adalah kekuatan isyarat yang mewakili kecerahan (Myler & Weeks, 1993). Antara fungsi

utama pemrosesan imej ialah bagi menghasilkan satu imej baru dengan mengolah data supaya kawasan yang dikehendaki dapat ditingkatkan dan kesan hingar dapat dikurangkan ataupun dihapuskan (Galbiati, 1990).

Pemrosesan imej melibatkan perubahan sifat sesuatu imej sama ada bagi meningkatkan maklumat bergambarnya untuk tafsiran manusia atau menjadikan ia lebih sesuai untuk tanggapan mesin berautonomi (Alasdair, 2011). Pemrosesan imej digital dibahagikan kepada lima bahagian asas iaitu peningkatan imej (*image enhancement*), pemulihan imej (*image restoration*), peruasan imej (*image segmentation*), pemampatan imej (*image compression*) dan sintesis imej (*image synthesis*) (Baxes, 1994; Alasdair, 2011).

Peningkatan imej merujuk kepada pemrosesan sesuatu imej agar menghasilkan imej yang lebih jelas dan menarik serta sesuai digunakan untuk sesuatu aplikasi. Imej diperbaiki dengan cara meningkatkan kontras imej tersebut, menyingkir atau mengurangkan kesan hingar, menajam atau menyah-kabur imej yang tidak terfokus dan menonjolkan pinggir. Pemulihan imej pula digunakan bagi memulihkan semula imej yang telah rosak akibat herotan geometrik, hingar yang berulang, pergerakan kamera atau pemfokusan yang tidak tepat. Seperti juga peningkatan imej, pemulihan imej bertujuan bagi memperbaiki kualiti imej namun proses pembaikannya adalah berdasarkan model matematik atau kebarangkalian degradasi imej (Gonzalez & Woods, 2002).

Peruasan imej pula melibatkan proses membahagikan sesuatu imej kepada bahagian-bahagian jujuk melalui pencarian garisan, bulatan atau bentuk-bentuk tertentu di dalam imej (Baxes, 1994; Gonzalez & Woods, 2002; Alasdair, 2011). Peruasan imej tidak menghasilkan keputusan bergambar tetapi keputusan yang dihasilkan adalah dalam bentuk sifat-sifat imej asal. Kaedah ini biasanya digunakan dalam penglihatan mesin. Pemampatan imej pula dilaksanakan bertujuan bagi mengurangkan saiz data bagi memperbaiki kadar penghantaran data dan mengurangkan saiz penyimpanan data (Baxes, 1994; Gonzalez & Woods, 2002). Ini kerana walaupun teknologi penyimpanan data pada masa kini telah berjaya mengatasi

masalah penyimpanan data, namun keupayaan penghantaran terutamanya melalui internet masih berdepan dengan masalah jika melibatkan penghantaran data yang bersaiz besar. Ketika proses penghantaran imej, jika kuantiti data yang terlibat adalah dalam saiz yang besar, maka pemampatan imej diperlukan.

Sintesis imej pula merupakan proses pembinaan imej dari imej lain atau dari data bukan imej (Baxes, 1994). Ia digunakan apabila imej yang dikehendaki tidak boleh diperolehi secara fizikal atau tidak praktikal bagi diperolehi atau langsung tidak wujud secara fizikal contohnya teknik tomografi komputer. Dalam penyelidikan ini, pemprosesan imej jenis peningkatan imej dan peruasan imej telah digunakan.

Antara algoritma-algoritma yang terlibat dalam setiap bahagian pemprosesan imej yang telah dinyatakan, ditunjukkan dalam Rajah 2.3. Walau bagaimanapun, setiap algoritma ini bukanlah milik eksklusif setiap bahagian pemprosesan imej tersebut tetapi ia bergantung kepada tujuan aplikasi itu dijalankan. Contohnya, algoritma pengelompokan yang sering digunakan untuk tujuan meruaskan imej (Chen *et al.*, 1998; Balafar *et al.*, 2008b), boleh juga digunakan bagi meningkatkan kualiti imej (Mat-Isa *et al.*, 2003b).

Tidak dapat dinafikan bahawa pemprosesan imej digital memainkan peranan penting dalam menganalisa dan memproses imej bagi mendapatkan maklumat yang sukar dilihat dengan mata kasar. Pemprosesan imej menjadi pilihan kerana ia mempunyai potensi yang baik dari segi ketepatan dan kecekapannya berbanding kaedah interpretasi imej menggunakan mata kasar. Pada peringkat awal, pemprosesan imej banyak digunakan dalam kajian angkasa lepas oleh *National Aeronautics and Space Administration* (NASA) di Amerika Syarikat. Namun, berasaskan ketepatan dan keberkesananannya memberikan keputusan yang memberangsangkan, penggunaan pemprosesan imej digital ini telah dikembangkan ke beberapa cabang sains yang lain, contohnya bidang kejuruteraan awam (Maerz, 1998; Pang & Liu, 2001; Lee, 2011), robotik (Mussa *et al.*, 1992; Maerz, 1998; Chung-Chang *et al.*, 2004), perubatan (de Pablo *et al.*, 2005; Khayati *et al.*, 2008) dan