

TESIS

**SEGMENTASI SEL PADA CITRA HISTOPATOLOGI KANKER
MENGUNAKAN ALGORITMA *DEEP LEARNING*
STUDI KASUS: CITRA PAYUDARA**

***NUCLEI SEGMENTATION ON CANCER HISTOPATHOLOGICAL
IMAGES USING DEEP LEARNING ALGORITHM
STUDI CASE : BREAST IMAGES***

**NISA MARDHATILLAH
D082192009**



**PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA
DEPARTEMEN TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2024**

PENGAJUAN TESIS

**SEGMENTASI SEL PADA CITRA HISTOPATOLOGI KANKER
MENGUNAKAN ALGORITMA *DEEP LEARNING*
STUDI KASUS: CITRA PAYUDARA**

Tesis
Sebagai Salah Satu Syarat untuk Mencapai Gelar Magister
Program Studi Teknik Informatika

Disusun dan diajukan oleh

ttd

**NISA MARDHATILLAH
D082192009**

Kepada

**FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2024**

TESIS

**SEGMENTASI SEL PADA CITRA HISTOPATOLOGI KANKER
MENGUNAKAN ALGORITMA *DEEP LEARNING*
STUDI KASUS: CITRA PAYUDARA****NISA MARDHATILLAH
D082192009**

Telah dipertahankan di hadapan Panitia Ujian Tesis yang dibentuk dalam rangka penyelesaian studi Program Magister Teknik Informatika Fakultas Teknik Universitas Hasanuddin pada tanggal 19 Juli 2024

Menyetujui,

Pembimbing Utama



Dr. Ir. Ingrid Nurtanio, MT
NIP. 19610813 198811 2 001

Pembimbing Pendamping



Prof. Dr. Eng. Ir. Syafaruddin, ST.,M.Eng, IPU
NIP. 19740530 199903 1 003

Dekan Fakultas Teknik
Universitas Hasanuddin

Prof. Dr.Eng. Ir. Muhammad Isran Ramli,
M.T.,IPM.,ASEAN.Eng.
NIP. 19730926 200012 1 002

Ketua Program Studi
S2 Teknik Informatika

Dr. Ir. Zahir Zainuddin, M.Sc
NIP. 19640427 198910 1 002

PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Yang bertanda tangan di bawah ini:

Nama : Nisa Mardhatillah
Nomor Mahasiswa : D082192009
Program Studi : Teknik Informatika

Dengan ini menyatakan, tesis berjudul “Segmentasi Sel Pada Citra Histopatologi Kanker Menggunakan Algoritma *Deep Learning*. Studi Kasus: Citra Payudara” adalah benar karya saya dengan arahan dari komisi pembimbing Dr. Ir. Ingrid Nurtanio, MT. dan Prof. Dr. Eng. Syafaruddin, ST., M.Eng., IPU. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apapun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Prosiding (International Seminar on Intelligent Technology and Its Applications (ISITIA) 2023, Halaman 581 s.d 586, DOI: 10.1109/ISITIA59021.2023.10221024) sebagai artikel dengan judul “*Nuclei Segmentation Using UNet on Breast Hematoxylin and Eosin Stained Histopathology Images*”.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Gowa, 19 Juli 2024

Yang menyatakan

A 10,000 Rupiah Indonesian postage stamp is shown with a signature over it. The stamp features the Garuda Pancasila emblem and the text 'SEKULUH RIBU RUPIAH', '10000', and 'METERAN'. The serial number '3299ALX25275862' is visible at the bottom.

Nisa Mardhatillah

KATA PENGANTAR

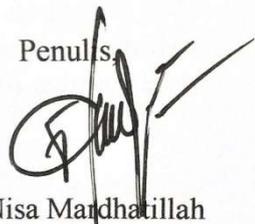
Alhamdulillah, segala puji dan syukur penulis panjatkan kepada Allah SWT. karena berkat rahmat dan karunia-Nya sehingga tesis dengan judul “Segmentasi Sel Pada Citra Histopatologi Kanker Menggunakan Algoritma *Deep Learning*. Studi Kasus: Citra Payudara” ini dapat diselesaikan sebagai salah satu syarat dalam menyelesaikan jenjang Strata-2 pada Departemen Teknik Informatika Fakultas Teknik Universitas Hasanuddin.

Penulis menyadari bahwa dalam penyusunan dan penulisan laporan tesis ini tidak lepas dari bantuan, bimbingan serta dukungan dari berbagai pihak, dari masa perkuliahan sampai dengan masa penyusunan tesis. Oleh karena itu, penulis dengan senang hati menyampaikan terima kasih kepada:

1. Allah SWT, Tuhan Yang Maha Esa atas rahmat, kasih sayang dan limpahan pertolongan-Nya yang tiada batas, yang diberikan kepada penulis dari awal perkuliahan dan penelitian hingga tesis ini bisa diselesaikan dengan sangat baik;
2. Bapak dan Ibu tercinta penulis, Bapak Drs. Teken, MM dan Ibu Kokamia Buchari, S.H serta adik-adik penulis yang selalu memberikan dukungan doa yang tidak terputus dan menjadi sumber kekuatan dan alasan penulis untuk terus semangat dalam menyelesaikan proses perkuliahan ini;
3. Ibu Dr. Ir. Ingrid Nurtanio, M.T. selaku pembimbing I yang telah dengan sangat sabar memberikan waktu, tenaga, pikiran, semangat serta perhatian untuk mengarahkan penulis dalam penyusunan tesis;
4. Bapak Prof. Dr.Eng.Ir. Syafaruddin,ST.,M.Eng,IPU. selaku pembimbing II yang juga telah dengan sangat sabar memberikan waktu, tenaga, pikiran, arahan dan kesempatan penulis untuk bimbingan selama proses penyusunan tesis;
5. Bapak Prof. Dr. Ir. Indrabayu,ST.,MT.,M.Bus.Sys.,IPM.,ASEAN.,Eng. Bapak Dr. Ir. Amil Ahmad Ilham, S.T., M.IT. dan Ibu Prof. Dr.Eng. Intan Sari Areni.ST.,MT. selaku dosen penguji yang telah memberikan kritik dan saran yang membangun dan bermanfaat untuk penulis sehingga laporan tesis ini dapat disusun menjadi lebih baik.
6. Bapak Dr. Ir. Zahir Zainuddin, M.Sc. selaku Ketua Program Studi Magister Teknik Informatika Fakultas Teknik Universitas Hasanuddin yang telah

memberikan banyak motivasi, bimbingan, dan semangatnya selama masa perkuliahan penulis;

7. Ibu Dr. dr. Fathul Djannah, Sp.PA dan Dr. dr. Najda Hidayah yang telah sangat banyak membantu penulis dalam proses penelitian, pengambilan data dan verifikasi data penelitian untuk tesis ini;
8. Ibu Yuanita S.Kom., M.Pd. serta segenap Staf Departemen Magister Teknik Informatika yang telah banyak membantu penulis selama pengurusan administrasi;
9. Teman-teman Lab. Animasi dan Multimedia yang selalu memberi dukungan dan semangat selama ini;
10. Teman-teman seperjuangan Magister Teknik Informatika atas dukungan dan semangat yang diberikan selama ini;
11. Terakhir, saya ingin berterima kasih pada diri saya sendiri yang selalu memilih untuk tidak menyerah dengan segala macam kesulitan-kesulitan yang dihadapi, yang tetap memilih bangkit walau sudah jatuh berkali-kali, yang tetap yakin pada kemampuan diri dan pertolongan Allah di setiap proses, dan tidak membandingkan diri dengan orang lain walau sesulit apa pun kondisinya. Yakin dan percaya bahwa Allah tidak akan menyia-nyiakan usaha hamba-Nya.

Penulis

Nisa Mandhatillah

ABSTRAK

NISA MARDHATILAH. *Segmentasi Sel Pada Citra Histopatologi Kanker Menggunakan Algoritma Deep Learning. Studi Kasus: Citra Payudara.* (dibimbing oleh **Ingrid Nurtanio and Syafaruddin**).

Penerapan teknologi visi komputer dalam analisis citra medis, khususnya dalam bidang histopatologi, menawarkan peluang signifikan untuk melakukan analisis sel secara kuantitatif. Namun, tantangan sering muncul ketika sel-sel berdekatan, berdempetan, atau saling tumpang tindih yang menyulitkan identifikasi objek secara individu sehingga mempengaruhi hasil perhitungan objek. Hal ini sering menimbulkan masalah *under segmentation* dan *over segmentation* dalam proses segmentasi sel. Penelitian ini bertujuan untuk meningkatkan kinerja segmentasi sel dengan menggunakan model U-Net yang memanfaatkan ResNet50 sebagai *backbone*. Penelitian ini melibatkan pembuatan dataset baru yang terdiri dari 106 citra dengan 3.383 sel serta *ground truth* yang sesuai. Model U-Net dengan ResNet50 dioptimalkan melalui penerapan teknik augmentasi data dan penyesuaian *hyperparameter*. Selain itu, penambahan tahapan modifikasi *ground truth* pada tahap pra-pemrosesan dilakukan untuk meningkatkan kinerja model. Hasil penelitian menunjukkan bahwa penambahan tahapan modifikasi *ground truth* berkontribusi dalam meningkatkan kinerja segmentasi sel. Modifikasi *ground truth* membantu model mengenali batas antar objek dengan lebih akurat, sehingga objek dapat dikenali secara individu yang mengurangi masalah *under segmentation*, dan meningkatkan pemahaman model terhadap morfologi objek yang berkontribusi pada pengurangan masalah *over segmentation*. Model tanpa modifikasi *ground truth* mendeteksi 515 objek merupakan jumlah yang jauh melebihi jumlah objek sebenarnya (327), sedangkan model dengan modifikasi GT mendeteksi 301 objek, mendekati jumlah objek sebenarnya. Peningkatan kinerja ini tercermin dari peningkatan nilai *Intersection over Union* (IoU) dari 0,7947 menjadi 0,8366, serta penurunan nilai rata-rata kesalahan deteksi (MAE) dari 0,0832 menjadi 0,0580. Perbaikan metrik-metrik ini menunjukkan bahwa penambahan tahapan modifikasi *ground truth* efektif dalam mengurangi kesalahan segmentasi dan meningkatkan akurasi segmentasi secara keseluruhan.

Kata Kunci: Kanker Payudara, Histopatologi, Segmentasi Sel, U-Net, ResNet.

ABSTRACT

NISA MARDHATILAH. *Nuclei Segmentation on Cancer Histopathological Images using Deep Learning Algorithm Study Case: Breast Images* (supervised by **Ingrid Nurtanio and Syafaruddin**).

The application of computer vision technology in medical image analysis, particularly in histopathology, offers significant opportunities for quantitative cell analysis. However, challenges frequently arise when cells are adjacent, clustered, or overlapping, complicating object identification and affecting object counting. This often leads to issues of under-segmentation and over-segmentation in the cell segmentation process. This study aims to enhance cell segmentation performance by utilizing the U-Net model with ResNet50 as the backbone. The study involves the creation of a new dataset comprising 106 images with 3,383 cells and their corresponding ground-truth. The U-Net model with ResNet50 was optimized through the application of data augmentation techniques and hyperparameter tuning. Additionally, a ground truth modification step was incorporated during pre-processing to further enhance the model's performance. The results indicate that the inclusion of ground truth modification significantly contributes to improved cell segmentation performance. This modification aids the model in more accurately recognizing the boundaries between objects, allowing individual object identification, which reduces under-segmentation issues. It also enhances the model's understanding of object morphology, contributing to the reduction of over-segmentation problems. The model without ground truth modification detected 515 objects, a number far exceeding the actual object count (327), whereas the model with ground truth modification detected 301 objects, closer to the actual count. This performance improvement is reflected in the increase of Intersection over Union (IoU) from 0.7947 to 0.8366 and the decrease in Mean Absolute Error (MAE) from 0.0832 to 0.0580. These metric improvements demonstrate that the inclusion of the ground truth modification step is effective in reducing segmentation errors and enhancing overall segmentation accuracy.

Keywords: Breast Cancer, Histopathology, Nuclei Segmentation, U-Net, ResNet.

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN PENGAJUAN TESIS	ii
LEMBAR PENGESAHAN TESIS	iii
PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	16
1.1 Latar Belakang.....	16
1.2 Rumusan Masalah.....	23
1.3 Tujuan Penelitian	23
1.4 Manfaat Penelitian	23
1.5 Batasan Masalah	23
1.5 Kontribusi Penelitian	24
BAB II TINJAUAN PUSTAKA	25
2.1 Kajian Pustaka	25
2.1.1 Kanker Payudara	25
2.1.2 Biopsy dan Histopatologi	28
2.1.3 Patologi Digital	33
2.1.4 <i>Computer-Aided Pathology</i>	38
2.1.5 <i>Artificial Intelligence (AI)</i>	44
2.1.6 Visi Komputer	45
2.1.7 <i>Convolutional Neural Network (CNN)</i>	47
2.1.8 U-Net	48
2.1.9 ResNet	50
2.1.10 <i>Morphological Operation</i>	51
2.1.11 <i>Connected Component Labeling</i>	53
2.1.12 Evaluasi Sistem	53

2.2 Metode Penyelesaian	60
2.2.1 State of The Art	60
2.2.2 Kerangka Pikir.....	68
BAB III METODOLOGI PENELITIAN	69
3.1 Jenis Penelitian	69
3.2 Tahapan Penelitian.....	69
3.2.1 Studi Literatur.....	70
3.2.2 Persiapan Data.....	71
3.2.3 Pengelolaan Data.....	71
3.2.4 Perancangan Sistem.....	71
3.2.5 Implementasi Sistem	72
3.2.6 Pengujian Sistem	72
3.3 Sumber Data.....	72
3.4 Waktu dan Lokasi Penelitian	73
3.5 Rancangan Sistem.....	73
3.6 Pengumpulan Data	74
3.7 Pengelolaan Data	76
3.7.1 Pembersihan Data (<i>Data Cleaning</i>)	76
3.7.2 <i>Patching</i>	77
3.7.3 <i>Data Annotation</i>	77
3.7.4 Normalisasi Warna (<i>Color Normalization</i>).....	79
3.7.5 Modifikasi <i>Ground truth</i>	80
3.8 <i>Training</i>	86
3.8.1 Augmentasi Data	87
3.8.2 <i>Encoding Label</i>	89
3.8.3 <i>Splitting Data</i>	89
3.8.4 Konfigurasi Model	90
3.8.5 Penyesuaian Hyperparameter (<i>hyperparameter tuning</i>)	93
3.8.6 Training	94
3.9 <i>Testing</i>	119
3.10 <i>Post-Processing</i>	119
3.11 Skenario Pengujian	120

3.12 Evaluasi Sistem.....	120
BAB IV HASIL DAN PEMBAHASAN	122
4.1 Hasil Anotasi	122
4.2 Hasil Pelatihan Model.....	123
4.1.1 Penyesuaian Hyperparameter	123
4.3 Pengujian Sistem	132
4.3.1 Pengujian Model Segmentasi	132
4.3.2 Evaluasi Metrik Kinerja	144
4.3.3 Evaluasi Perhitungan Objek	148
4.3.4 Perbandingan Metode Tanpa dan Dengan Modifikasi <i>Ground Truth</i>	154
BAB V KESIMPULAN DAN SARAN	162
5.1 Kesimpulan	162
5.2 Saran	163
DAFTAR PUSTAKA	164
LAMPIRAN.....	171

DAFTAR TABEL

Tabel 1 State of The Art	60
Tabel 2 Spesifikasi Perangkat	75
Tabel 3 Pembagian Awal Dataset	79
Tabel 4 Perbandingan jumlah piksel berdasarkan ukuran garis tepi	85
Tabel 5 Teknik Augmentasi	88
Tabel 6 Pembagian Data training dan validation	89
Tabel 7 Spesifikasi Eksperimen	90
Tabel 8 Konfigurasi Model	91
Tabel 9 Percobaan Hyperparameter	93
Tabel 10 Defenisi Hyperparameter	94
Tabel 11 Objek Teranotasi	122
Tabel 12 Evaluasi learning rate terhadap nilai loss dan IoU	124
Tabel 13 Evaluasi Batch Size terhadap nilai loss dan IoU	127
Tabel 14 Evaluasi Epochs terhadap nilai loss dan IoU	130
Tabel 15 Hasil Pengujian 1	132
Tabel 16 Hasil Pengujian 2	135
Tabel 17 Hasil Pengujian 3	137
Tabel 18 Hasil Pengujian 4	139
Tabel 19 Hasil Pengujian 5	142
Tabel 20 Tabel Confusion Matrix	144
Tabel 21 Hasil perhitungan objek set pengujian 1	149
Tabel 22 Hasil perhitungan objek set pengujian 2	149
Tabel 23 Hasil perhitungan objek set pengujian 3	150
Tabel 24 Hasil perhitungan objek set pengujian 4	151
Tabel 25 Hasil perhitungan objek set pengujian 5	152
Tabel 26 Hasil pengujian set 1 tanpa tahap modifikasi GT.....	155
Tabel 27 Hasil pengujian set 1 dengan tahap modifikasi GT.....	155

DAFTAR GAMBAR

Gambar 1 Peta sebaran penyakit kanker payudara dunia.....	16
Gambar 2 Persentase perbandingan jumlah temuan kasus penyakit kanker	17
Gambar 3 Ilustrasi prosedur pemeriksaan penyakit kanker payudara.....	29
Gambar 4 Pemeriksaan Payudara.....	30
Gambar 5 Ilustrasi proses <i>Fine Needle Aspiration Biopsy</i> (FNAB).....	30
Gambar 6 Ilustrasi pemrosesan jaringan (Mahbod et al., 2023).....	31
Gambar 7 Patologi Konvensional vs Patologi Digital.....	33
Gambar 8 Proses analisa preparat digital	34
Gambar 9 Tampilan mikroskop virtual (H&E Staining The Cell)	36
Gambar 10 Citra WSI asli (Kovalev, Kalinovsky and Liauchuk, 2016).....	36
Gambar 11 Citra WSI dan perbesarannya (Qaiser et al., 2019).	37
Gambar 12 Ilustrasi citra WSI dalam format piramida (Lehtonen, 2021).....	37
Gambar 13 Variasi bentuk dan ukuran sel (Naylor et al., 2017a)	40
Gambar 14 Batas yang halus antar sel (Naylor et al., 2017a)	40
Gambar 15 Penampakan sel	41
Gambar 16 Contoh under dan over-segmentation.....	42
Gambar 17 Contoh gambar histopatologi.....	43
Gambar 18 Variasi hasil pewarnaan Hematoxylin and Eosin (H&E).....	43
Gambar 19 Hubungan antara AI, ML, dan DL.....	44
Gambar 20 Cara kerja <i>Machine Learning</i> dan <i>Deep Learning</i>	45
Gambar 21 Sistem penglihatan manusia	46
Gambar 22 Komponen sistem visi komputer.....	47
Gambar 23 Arsitektur dasar CNN.....	48
Gambar 24 Arsitektur U-Net.....	49
Gambar 25 Arsitektur ResNet-50.....	51
Gambar 26 Contoh Operasi Dilasi.....	52
Gambar 27 Contoh operasi Erosi.....	53
Gambar 28 Contoh operasi <i>Connected Component Labeling</i>	53
Gambar 29 Penilaian performa segmentasi berdasarkan nilai IoU	56
Gambar 30 Penilaian performa segmentasi berdasarkan nilai <i>Dice Coefficient</i>	57
Gambar 31 Kerangka Pikir	68

Gambar 32	<i>Flowchart</i> Tahapan Penelitian.....	70
Gambar 33	Rancangan Sistem.....	73
Gambar 34	Perangkat Pengambilan Data.....	74
Gambar 35	Hasil penangkapan citra histopatologi.....	75
Gambar 36	Hasil proses <i>patching</i> citra	77
Gambar 37	<i>Flowchart</i> anotasi dan <i>labeling</i>	78
Gambar 38	Proses anotasi sel menggunakan Roboflow.....	78
Gambar 39	Proses export dataset	79
Gambar 40	Hasil dari normalisasi citra menggunakan metode.....	80
Gambar 41	Potongan kode untuk membuat kelas tepi	81
Gambar 42	Ilustrasi operasi erosi	82
Gambar 43	Hasil operasi erosi.....	82
Gambar 44	Ilustrasi operasi dilasi	83
Gambar 45	Hasil operasi dilasi.....	84
Gambar 46	Mengubah nilai piksel 255 menjadi 127.....	84
Gambar 47	Hasil pengembalian nilai objek	85
Gambar 48	Hasil penambahan garis pada tepi objek	86
Gambar 49	Visualisasi hasil augmentasi data	89
Gambar 50	Hasil pengecekan data	90
Gambar 51	Arsitektur Model.....	94
Gambar 52	Model ResNet-50.....	96
Gambar 53	Ilustrasi penerapan padding pada citra	98
Gambar 54	Ilustrasi penggunaan padding	98
Gambar 55	Ilustrasi <i>stride</i> 1, 2 dan 3	99
Gambar 56	Proses konvolusi <i>channel red</i> pada layer encoder.....	100
Gambar 57	Tahap lanjutan konvolusi pada <i>channel red</i>	101
Gambar 58	Proses konvolusi <i>channel green</i> pada layer encoder	102
Gambar 59	Tahap lanjutan konvolusi pada <i>channel green</i>	103
Gambar 60	Tahap lanjutan konvolusi pada <i>channel blue</i>	103
Gambar 61	Tahap lanjutan konvolusi pada <i>channel blue</i>	105
Gambar 62	Contoh peta fitur yang dihasilkan melalui proses konvolusi.....	105
Gambar 63	Proses <i>Max Pooling</i>	106

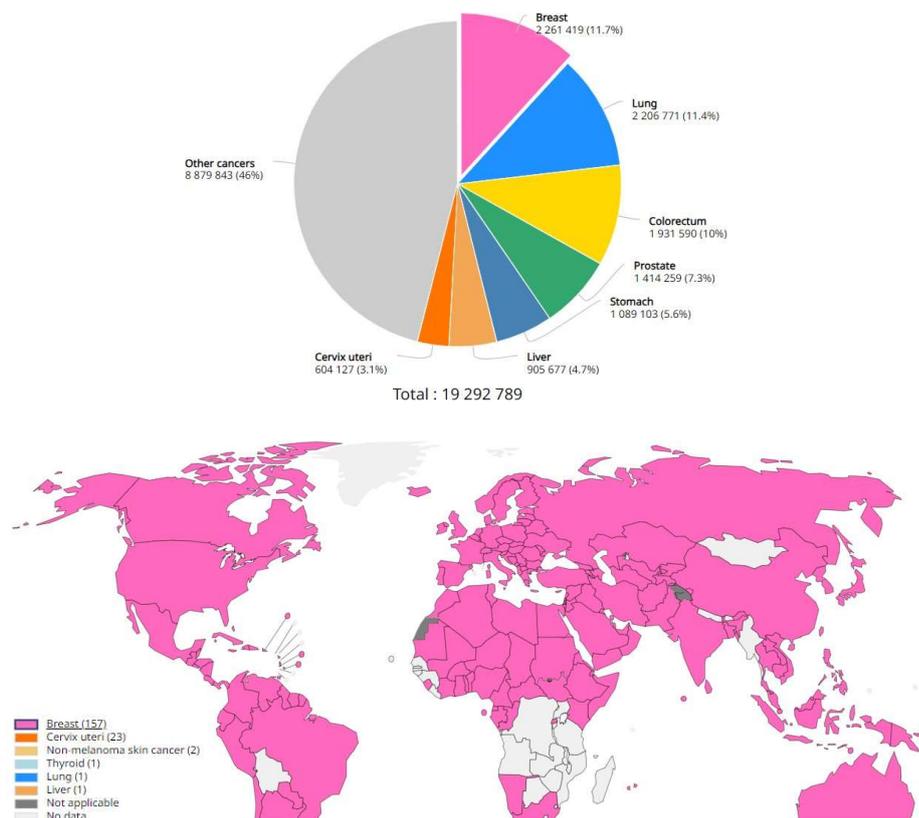
Gambar 64	Proses <i>Max Pooling</i> pada <i>Red Channel</i>	107
Gambar 65	Arsitektur Model (<i>decoder</i>)	108
Gambar 66	Sebelum dan setelah penyesuaian <i>stride</i>	109
Gambar 67	Ilustrasi proses konvolusi <i>transpose</i> pada <i>red channel</i>	110
Gambar 68	Ilustrasi proses lanjutan konvolusi <i>transpose</i> pada <i>channel red</i>	111
Gambar 69	Ilustrasi proses konvolusi <i>transpose</i> green channel	111
Gambar 70	Ilustrasi proses lanjutan konvolusi <i>transpose</i> pada green channel..	113
Gambar 71	Ilustrasi proses konvolusi <i>transpose</i> pada <i>blue channel</i>	113
Gambar 72	Ilustrasi proses lanjutan konvolusi <i>transpose</i> pada blue channel....	114
Gambar 73	Ilustrasi operasi <i>concatenate</i>	115
Gambar 74	Ilustrasi output hasil dari proses <i>upsampling</i>	117
Gambar 75	Ilustrasi hasil dari penerapan fungsi aktivasi <i>softmax</i>	119
Gambar 76	Contoh hasil segmentasi dari proses pengujian.....	119
Gambar 77	Hasil <i>Post-Processing</i>	120
Gambar 78	Citra dan <i>ground truth</i> hasil anotasi	122
Gambar 79	Evaluasi <i>learning rate</i> terhadap nilai <i>Loss</i>	125
Gambar 80	Evaluasi nilai dice loss dan IoU terhadap Batch Size.....	128
Gambar 81	Evaluasi nilai <i>Loss</i> dan IoU terhadap <i>Epoch</i>	131
Gambar 82	Hasil Segmentasi Sel pada pengujian 1	134
Gambar 83	Hasil Segmentasi Sel pada pengujian 2	136
Gambar 84	Hasil Segmentasi Sel pada pengujian 3	138
Gambar 85	Hasil Segmentasi Sel pada pengujian 4.....	140
Gambar 86	Hasil Segmentasi Sel pada pengujian 5	143
Gambar 87	Visualisasi confusion matrix sub tes 1 (pengujian 1).....	145
Gambar 88	Grafik perbandingan performa model pada setiap pengujian.....	147
Gambar 89	Grafik perbandingan kesalahan segmentasi vs IoU.....	153
Gambar 90	Grafik perbandingan performa antara dua metode	156
Gambar 91	Grafik perbandingan hasil perhitungan objek	157
Gambar 92	Hasil segmentasi tanpa dan dengan modifikasi GT.....	159

BAB I PENDAHULUAN

1.1 Latar Belakang

Tahun 2020, Organisasi Kesehatan Dunia (WHO) menyatakan bahwa kanker merupakan jenis penyakit paling mematikan yang menjadi salah satu penyebab kematian terbanyak hampir di seluruh dunia dengan 9,6 juta kasus kematian (*Cancer*, 2020). Payudara, paru-paru, kolorektal, prostat, perut dan hati adalah beberapa organ kanker yang paling umum ditemukan, namun penyakit kanker payudara merupakan jenis kanker dengan persentase jumlah penderita terbanyak yang ditemukan dengan jumlah kasus tercatat sebanyak 2.261.419 atau (11.7%) kasus pada tahun 2020 (*Cancer (IARC)*, 2020). Peta sebaran dan persentase penderita kanker tertinggi ditunjukkan pada gambar 1.

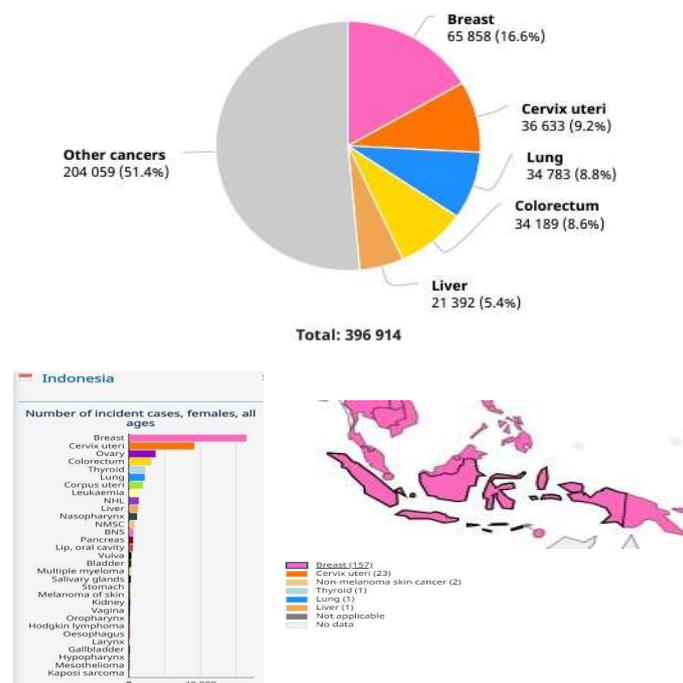
1



Gambar 1 Peta sebaran penyakit kanker payudara dunia.

Di negara berkembang khususnya Indonesia, kanker payudara juga merupakan jenis penyakit kanker dengan jumlah temuan kasus terbanyak. Berdasarkan data statistik yang dirilis oleh *Global Cancer Observatory* pada tahun 2020, jumlah kasus baru kanker payudara yang ditemukan di Indonesia merupakan kasus

terbanyak dengan jumlah kasus mencapai 65.858 atau 16,6% dari total 369.914 kasus kanker sementara itu kasus kematian akibat kanker payudara di tahun 2020 berjumlah 9.6% atau 22.430 kasus kematian (*Global Cancer Observatory (IARC), 2020*), berdasarkan data tersebut maka penyakit kanker payudara termasuk dalam daftar penyakit kanker dengan prevalensi tertinggi di Indonesia. Perbandingan persentase jumlah penderita kanker khususnya di Indonesia dapat dilihat pada gambar 2. Upaya yang dapat dilakukan untuk menurunkan tingkat prevalensi kematian akibat penyakit kanker payudara salah satunya dengan melakukan pemeriksaan dini untuk mengetahui kondisi penyakit lebih awal (Mahbod *et al.*, 2019). Pengobatan dapat dilakukan pada organ yang rentan terkena kanker, jika tumor ganas terdeteksi pada tahap awal, setelah kelainan pada organ terdeteksi, proses pemeriksaan biopsi dapat dilakukan, dan jaringan organ pasien akan diambil untuk menyelidiki lebih mendalam. Investigasi mendalam ini dilakukan oleh departemen patologi anatomi.



Gambar 2 Persentase perbandingan jumlah temuan kasus penyakit kanker di Indonesia

Metode pemeriksaan biopsi merupakan sebuah prosedur pemeriksaan histopatologi yang dilakukan oleh dokter ahli patologi anatomi untuk mendeteksi ketidaknormalan pada jaringan tubuh dengan menggunakan mikroskop. Dalam biopsi, sebagian kecil jaringan dari bagian tubuh yang diduga terinfeksi kanker

diangkat, jaringan yang telah diangkat kemudian akan dipotong, dikeringkan dan diwarnai untuk keperluan evaluasi lebih lanjut, jaringan yang sudah siap kemudian dilakukan analisa secara lebih detail. Pemeriksaan histopatologi telah dipandang sebagai pemeriksaan '*gold standart*' dalam mengenali hampir semua jenis kanker karena mampu menangkap struktur mikroskopis sel dan jaringan yang merupakan informasi yang sangat penting dalam menghasilkan hasil diagnosa yang akurat dan tindakan penanganan yang tepat (Rashmi *et al.*, 2020). Dokter ahli patologi anatomi melakukan evaluasi visual terhadap struktur jaringan dengan menggunakan mikroskop untuk mengklasifikasi dan menentukan tingkat keganasan kanker yang diderita pasien berdasarkan kuantitas dan fitur morfologi tertentu dari jumlah sel, sehingga perubahan morfologi yang terjadi pada sel dianggap sebagai sebuah isyarat penting dalam diagnosa penyakit, menilai tingkat keganasan kanker dan menentukan keputusan pengobatan (Mahbod *et al.*, 2022), (Gurcan *et al.*, 2009).

Perkembangan teknologi dalam bidang patologi digital dalam beberapa tahun terakhir memungkinkan tenaga medis untuk melakukan investigasi jaringan secara lebih mudah dengan memanfaatkan teknologi mikroskop dan alat pemindai digital yang digunakan oleh dokter ahli patologi untuk menganalisis jaringan dengan menggunakan *Wide Slide Image* (WSI) yang dihasilkan oleh alat pemindai digital yang dihubungkan dengan komputer sehingga memungkinkan dokter ahli patologi untuk melihat jaringan dan sel-sel dengan lebih jelas.

Analisis gambar histologi biasanya dilakukan pada *slide* dengan Teknik pewarnaan *Hematoxilyn and Eosin* (H&E) dimana sel di interpretasikan dengan warna ungu atau kebiruan sedangkan jaringan lain di visualisasikan dengan warna merah muda yang akan memberikan informasi berharga bagi ilmuwan medis yang mempelajari patofisiologi penyakit (Alom *et al.*, 2022), (Djuric *et al.*, 2017). Interpretasi gambar yang diwarnai dengan teknik H&E oleh para profesional seperti ahli patologi, dokter, atau ilmuwan kerap menghadapi kendala dalam analisis manual yang konvensional karena proses tersebut memerlukan waktu lama dan rentan terhadap varian interpretasi antar pengamat hal ini sangat dipengaruhi oleh pengalaman masing-masing pengamat (Mahbod *et al.*, 2022). Seiring kemajuan teknologi, pemindai slide berbasis mikroskop yang mampu mendigitalkan gambar histologi telah muncul. Sistem analisis gambar berbantuan komputer kini

diperkenalkan untuk meningkatkan keakuratan yang memungkinkan analisa secara kuantitatif dan objektif, serta membantu para profesional dalam mempercepat proses analisis (Wu *et al.*, 2022).

Penerapan teknik pengolahan citra dalam menganalisis citra mikroskopis sel kanker payudara pada komputasi patologi telah banyak dilakukan juga dikembangkan. Segmentasi citra merupakan permasalahan yang penting untuk diangkat dalam pengolahan citra histopatologi sel kanker payudara selain karena segmentasi merupakan proses yang paling mendasar dalam sistem pemrosesan citra untuk analisis citra histopatologi, proses ini juga harus menghasilkan *output* yang merepresentasikan citra asli. Selain itu proses segmentasi juga sangat berpengaruh terhadap beberapa proses analisa selanjutnya seperti perhitungan jumlah sel, klasifikasi tipe sel, penentuan stadium penyakit (Hassan *et al.*, 2020).

Segmentasi mengacu pada proses membagi piksel pada citra kedalam beberapa kategori sesuai dengan objek, hal ini bertujuan agar objek yang dimaksud terpisah dengan objek lainnya. Segmentasi citra histopatologi yang bertujuan untuk menetapkan label kelas pada setiap piksel pada suatu gambar merupakan tugas yang umum dilakukan dalam analisis citra histopatologi (Chan *et al.*, 2019). Segmentasi citra histopatologi pada dasarnya dapat dibagi menjadi dua kategori, yaitu segmentasi sel dan segmentasi jaringan. Tugas segmentasi sel berfokus pada eksplorasi ciri-ciri sel, seperti penampakan morfologi pada gambar histopatologi yang secara luas dikenal sebagai biomarker yang paling sering digunakan untuk diagnosis histologi kanker, sementara segmentasi jaringan dilakukan dengan mengambil gambar histopatologi sebagai masukan kemudian mengelompokkan jaringan yang terdiri dari sekelompok sel pada gambar masukan dengan karakteristik dan struktur tertentu (kelenjar, limfosit yang menginfiltrasi tumor, dan lain-lain). Sel dan jaringan yang diukur secara kuantitatif ini juga merupakan indikator penting untuk menentukan diagnosis dan prognosis kanker (Zhang *et al.*, 2021), (Sucher and Sucher, 2020).

Pola heterogen WSI pada segmentasi inti dan jaringan yang akurat pada gambar histopatologi menghadapi tantangan besar. Pertama, adanya variasi ukuran dan bentuk inti atau jaringan, sehingga memerlukan model segmentasi dengan kemampuan generalisasi yang kuat. Kedua, sel sering kali terlihat saling

berdekatan, berdempet bahkan mengelompok menjadi gumpalan menyebabkan sebagian mungkin tumpang tindih atau bersentuhan satu sama lain, sehingga gambar histopatologi tidak tersegmentasi dengan baik dan mengalami kondisi *over segmentation* dan *under segmentation*. Ketiga, pada beberapa kasus keganasan, seperti adenokarsinoma berdiferensiasi sedang dan buruk, struktur jaringan (misalnya kelenjar) mengalami degenerasi berat, sehingga sulit untuk didiskriminasi (Janowczyk and Madabhushi, 2016), (Chen *et al.*, 2017). Masalah inilah yang diangkat dalam penelitian-penelitian sebelumnya untuk melakukan segmentasi sel kanker. Beberapa penelitian terkait segmentasi sel pada citra histopatologi kanker telah dilakukan oleh beberapa peneliti bidang visi komputer.

Penelitian (Faticah *et al.*, 2015) menggunakan *Gram-Schmidt* untuk segmentasi sel kanker dengan vektor pembobotan dihitung untuk memperkuat vektor warna yang diinginkan dan melemahkan vektor warna yang tidak diinginkan. Penelitian ini menghasilkan tingkat MSE sebesar 0.08 yang lebih rendah dibandingkan dengan metode *Otsu's Thresholding* dengan nilai MSE sebesar 0.26, kekurangan penelitian ini belum mampu menangkap struktur sel dengan presisi dan terdapat pula kesalahan segmentasi dimana daerah yang harusnya tidak teridentifikasi sebagai sel, tetapi teridentifikasi sebagai sel. Sedangkan identifikasi sel menggunakan metode *Gram-Schmidt* juga dilakukan oleh (Tuwohingide and Faticah, 2017) untuk menangkap area sel kemudian area sel yang saling tumpang tindih di deteksi menggunakan fitur area geometri, untuk pemisahan sel bertumpuk penulis menggunakan metode *Fuzzy C-Means* untuk segmentasi area sel dan *Rapid Region Merging* untuk pemisah sel yang bertumpuk menghasilkan tingkat akurasi sebesar 78.41%. Namun Penelitian (Huang and Lai, 2010), (Faticah *et al.*, 2015) dan (Tuwohingide and Faticah, 2017) memiliki kekurangan yaitu penelitian-penelitian ini belum mampu melakukan segmentasi pada objek yang berkelompok dan tumpang tindih dengan optimal. (Rashmi *et al.*, 2020) menggunakan metode *color features*, *filterbank*, dan *gray level co-occurrence matrix* (GLCM) untuk melakukan ekstraksi fitur citra sel kanker payudara kemudian menggunakan dua algoritma untuk melakukan proses klasifikasi pixel dengan algoritma *Support Vector Machines* (SVM) dan *Multi Layer Perceptron* (MLP) dengan hasil MLP (LBP features) : 82% SVM (GLCM

features) : 82%. Kelebihan penelitian ini mampu melakukan segmentasi sel dengan baik pada sel yang berukuran lebih besar dan tunggal, namun kekurangannya penelitian ini belum optimal dalam melakukan segmentasi pada sel yang berukuran lebih kecil dan terlebih jika sel saling berkelompok atau tumpang tindih.

Secara berkelanjutan metode segmentasi dan pemisahan sel menggunakan pendekatan *deep learning* menjadi metode yang populer digunakan untuk mengatasi masalah ini. Metode *deep learning* yang memiliki kemampuan untuk mengekstrak fitur secara otomatis dan lebih mendalam telah digunakan untuk mengembangkan model yang memiliki performa yang lebih baik. Beberapa metode *deep learning* telah diaplikasikan untuk masalah segmentasi sel dan pemisahan sel tumpang tindih. Penelitian yang dilakukan oleh (Qingbo, Lao and Fevens, 2019) menggunakan metode *two-stage learning framework* dan *Deep Layer Aggregation* (DLA) dengan menggunakan dua tahapan yaitu dengan menumpuk model U-Net dimana U-Net tahapan pertama mengidentifikasi sel dan penentuan batas antar sel (secara kasar) kemudian U-Net tahap kedua melakukan segmentasi akhir dan memprediksi batas antar sel. Penelitian ini menghasilkan nilai Recall 83.3%, Precision 82.6%, F1 82.9% dan Aggregated Jaccard Index 0.611. Penelitian (Kong *et al.*, 2020) juga menerapkan konsep *two stage learning framework* yang menggunakan dua U-Net bertumpuk (Stacked UNets/SUNets). SUNets terdiri dari 4 layer encoder yang berbeda yang bekerja secara paralel yang kemudian *Attention Generation Model* (AGM) digunakan untuk mengakumulasikan hasil dari keempat layer encoder tersebut hingga menghasilkan segmentasi awal. Kemudian Proses tersebut diulang satu kali lagi untuk memisahkan sel yang bertumpuk, proses yang dilakukan sama seperti sebelumnya. Namun, yang berbeda adalah hasil dari segmentasi sebelumnya juga dimasukkan sebagai *input*. Penelitian ini menghasilkan nilai AJI 0.5965, F1 Score 82.47%, Recall 85.3%, Precision 79.2%, F1-Score 80.6%. Namun, penelitian (Qingbo, Lao and Fevens, 2019) dan (Kong *et al.*, 2020) ini memiliki kelemahan yaitu membutuhkan *computational resource* yang besar karena penelitian ini menggunakan mekanisme *two stage UNet* sehingga proses diulang lebih dari satu kali dan pada penelitian (Kong *et al.*, 2020) menggunakan dua tahapan U-Net dan menggunakan banyak empat layer *encoder* yang berbeda, hal ini menyebabkan semakin banyaknya sumber daya komputasi

yang dibutuhkan.

Selanjutnya (Wang, Li and Luo, 2020) dalam penelitiannya terkait segmentasi semantik sel kanker yang mengusulkan metode pengembangan UNet++ dengan menambahkan Inception-ResNet-V2 sebagai layer encoder yang memungkinkan jaringan melakukan pelatihan mendalam dengan kinerja yang sangat baik dalam melakukan pendeteksian objek. Penelitian ini menghasilkan nilai akurasi sebesar 95.05%, *precision* 55.81%, *recall* 60.35%, *dice coefficient* 52.07%. Penelitian (Zarbega and Gültepe, 2020) menerapkan algoritma CNN yang terdiri dari tiga lapisan utama, dimana setiap layer utama berisikan tiga sub layer yang akan menjadi lapisan *input* dan total sepuluh lapisan. Lapisan utama kedua terdiri dari *convolutional layer rectified unit* dan *transposed convolution layer*. Layer terakhir berisi *convolutional layer*, *softmax* dan *classification layer*. Kemudian menggunakan *Binarization and Segmentation* untuk menghasilkan segmentasi akhir. Penelitian ini menghasilkan nilai *precision* 84.4%, *recall* 83.2% dan *accuracy* 85.1%. Namun kedua penelitian ini memiliki kelemahan yaitu terdapat banyak area yang mengalami *over segmentation* disebabkan banyaknya objek yang mengalami oklusi dan kluster sel yang padat menyebabkan rentan terjadinya *over segmentation*.

Penelitian ini berfokus pada segmentasi sel pada citra histopatologi pemeriksaan payudara dengan warna, bentuk, ukuran sel yang bervariasi khususnya pada sel yang saling berdempet dan tumpang tindih.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah, rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana membangun model untuk melakukan segmentasi sel pada citra histopatologi payudara.
2. Bagaimana meminimalisir masalah *under segmentation* dan *over segmentation* yang sering ditemui pada segmentasi sel yang saling berdempet dan tumpang tindih.

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah

1. Membangun model untuk melakukan segmentasi sel pada citra histopatologi payudara,
2. Meminimalisir terjadinya masalah *under-segmentation* dan *over segmentation* yang sering dijumpai pada tugas segmentasi sel.

1.4 Manfaat Penelitian

Hasil dari penelitian ini mampu memberi kontribusi dalam meningkatkan kualitas segmentasi sel citra histopatologi pemeriksaan payudara dengan meminimalisir terjadinya masalah *under segmentation* dan *over segmentation*.

1.5 Batasan Masalah

Adapun batasan masalah dalam penelitian ini, yaitu:

1. Penelitian ini difokuskan kepada segmentasi citra sel secara umum pada *slide* pemeriksaan payudara tanpa mengklasifikasikan jenis selnya.
2. Citra sel yang akan diproses merupakan citra 2 dimensi (2D) dengan pembesaran 40X menggunakan mikroskop.
3. Dataset yang dikumpulkan berupa citra sel yang berasal dari *slide* pemeriksaan payudara.
4. Jenis pewarnaan pada citra yang digunakan adalah pewarnaan *Hematoxilyn and Eosin* (H&E).

1.5 Kontribusi Penelitian

Kontribusi penulis pada penelitian ini adalah dengan memanfaatkan metode operasi morfologi berupa erosi dan dilasi untuk penambahan kelas baru berupa garis tepi pada *ground truth* data latih pada tahapan modifikasi *ground truth*.

BAB II TINJAUAN PUSTAKA

2.1 Kajian Pustaka

2.1.1 Kanker Payudara

Kanker payudara merupakan salah satu jenis kanker yang banyak terjadi dan sering menjadi penyebab kematian khususnya untuk para wanita. Berdasarkan data statistik yang dirilis oleh *Global Cancer Observatory* pada tahun 2020, jumlah kasus baru kanker payudara yang ditemukan di Indonesia merupakan kasus terbanyak dengan jumlah kasus mencapai 65.858 atau 16,6% dari total 369.914 kasus kanker sementara itu kasus kematian akibat kanker payudara di tahun 2020 berjumlah 9.6% atau 22.430 kasus kematian (*Global Cancer Observatory (IARC), 2020*). Angka kejadian sampai saat ini masih terus meningkat pada wanita usia antara 40-45 tahun, diagnosis dini serta terapi yang dilakukan secara tepat dan tepat merupakan pendekatan utama dalam upaya yang bisa dilakukan untuk mengurangi risiko kematian akibat penyakit kanker payudara (Bray *et al.*, 2018).

Kanker Payudara adalah suatu jenis tumor ganas yang menyerang jaringan payudara. Jaringan payudara pada dasarnya terdiri dari kelenjar susu, saluran kelenjar, jaringan lemak, dan jaringan ikat. Secara normal, dalam jaringan payudara terjadi regenerasi sel, di mana sel-sel yang tua mengalami apoptosis (kematian sel) dan digantikan oleh sel-sel yang baru. Proses regenerasi ini penting untuk mempertahankan fungsi normal payudara. Dalam kasus kanker payudara, terjadi mutasi gen yang bertanggung jawab atas regulasi pertumbuhan sel. Mutasi gen ini menyebabkan sel-sel payudara berkembang secara tidak terkendali dan tidak teratur. Kondisi ini disebut sebagai kanker payudara (Mardiana, 2007).

Ada beberapa faktor yang memengaruhi kejadian kanker payudara, antara lain :

1. Umur: Wanita yang berusia di atas 50 tahun memiliki risiko lebih besar untuk mengembangkan kanker payudara.
2. Faktor hormonal: Riwayat menstruasi (menarche awal, menopause akhir) dan penggunaan hormon terkait dengan peningkatan risiko kanker payudara. Penggunaan hormon estrogen lebih dari 8-10 tahun dapat meningkatkan risiko.
3. Keturunan: Risiko kanker payudara meningkat sebesar 3 kali pada wanita yang

memiliki anggota keluarga (ibunya atau saudaranya) dengan riwayat kanker payudara. Risiko juga meningkat pada pria dengan sindrom Klinefelter dan keluarga dengan riwayat kanker endometrium, ovarium, atau kolorektal.

4. Gaya Hidup: Konsumsi diet tinggi lemak dan alkohol terkait dengan peningkatan risiko. Alkohol dapat meningkatkan sekresi estrogen dan mengurangi klerensi estrogen pada wanita. Kurangnya aktivitas fisik dan obesitas pasca menopause juga berkontribusi pada risiko kanker payudara. faktor-faktor ini dapat berinteraksi dan memengaruhi risiko seseorang untuk mengembangkan kanker payudara (Prayoga, 2019).

Adapun beberapa gejala klinis yang terjadi pada kanker payudara berupa (Zulkoni, 2010):

1. Benjolan pada payudara Benjolan pada payudara ini umumnya tidak terasa nyeri. Benjolan ini awalnya kecil, tetapi semakin lama akan semakin besar lalu melekat pada kulit atau menimbulkan perubahan pada kulit payudara atau pada puting susu.
2. Erosi atau eksema puting susu, kulit atau puting susu menjadi tertarik ke dalam (retraksi), berwarna merah muda, atau kecokelatan sampai menjadi bengkak hingga kulit kelihatan seperti kulit jeruk, mengerut atau timbul borok atau ulkus pada payudara. Ulkus ini semakin lama akan semakin besar dan mendalam sehingga dapat menghancurkan sebuah payudara, sering berbau busuk dan mudah berdarah. Pada gejala erosi atau eksema puting susu ini sering muncul gejala lain di antaranya:
 - a. Pendarahan pada puting susu.
 - b. Rasa sakit atau nyeri, pada umumnya baru timbul apabila tumor sudah besar, sudah timbul ulkus, atau bila sudah muncul metastate ke tulang-tulang.
 - c. Pembesaran kelenjar getah bening di ketiak, bengkak pada lengan, dan penyebaran kanker ke seluruh tubuh.
3. Keluarnya cairan (*Nipple discharge*)

Nipple discharge adalah keluarnya cairan dari puting susu secara spontan dan tidak normal. Seorang wanita harus waspada apabila dari puting susu keluar cairan berdarah, encer dengan warna merah atau coklat, keluar sendiri

tanpa harus memijit puting susu dan berlangsung secara terus menerus, hanya pada satu payudara dan cairannya ini bukan air susu.

Sebelum menetapkan diagnosis, dokter terlebih dahulu akan melakukan anamnesis (wawancara medis) seputar gejala, gaya hidup, riwayat kesehatan pasien, serta riwayat keluarga pasien. Setelah melakukan anamnesis, maka dokter akan melakukan serangkaian prosedur pemeriksaan untuk membantu menegakkan diagnosis, jenis pemeriksaan yang dimaksud berupa pemeriksaan fisik payudara sendiri (SADARI) ataupun secara klinis (SADANIS), termasuk didaerah leher hingga ketiak, pemeriksaan mamografi, pemeriksaan *ultrasonografi* (USG) payudara, biopsy dan pemeriksaan *Magnetic Resonance Imaging* (MRI).

Penanganan kanker payudara saat ini lebih cenderung mengadopsi pendekatan multidisipliner yang melibatkan kolaborasi antara berbagai disiplin ilmu medis. Pendekatan ini bertujuan untuk memastikan bahwa penanganan kanker payudara dapat disesuaikan secara lebih optimal dengan kondisi dan kebutuhan setiap pasien. Proses pengambilan keputusan, yang dimulai dari tahap diagnosis hingga pemilihan terapi yang paling tepat, melibatkan keterlibatan dan diskusi bersama antara tim medis yang terdiri dari ahli bedah onkologi, medis onkologi, radiologi onkologi, dan ahli *diagnostic imaging* (Prayoga, 2019).

Pendekatan multidisipliner dalam penanganan kanker payudara ini bertujuan untuk:

1. Menggabungkan Pengetahuan dan Keahlian: Dengan melibatkan berbagai spesialis, setiap aspek dari kasus kanker payudara dapat dievaluasi dengan teliti. Ahli bedah onkologi akan memberikan wawasan tentang opsi pembedahan, medis onkologi akan memberikan perspektif terkait terapi medis, radiologi onkologi akan memberikan informasi tentang gambaran radiologi yang mendukung diagnosis, dan ahli *diagnostic imaging* akan berperan penting dalam interpretasi citra dan pemantauan perkembangan penyakit.
2. Personalisasi Terapi: Dalam diskusi bersama ini, keputusan terkait jenis terapi yang akan diberikan kepada pasien dapat dipersonalisasi sesuai dengan

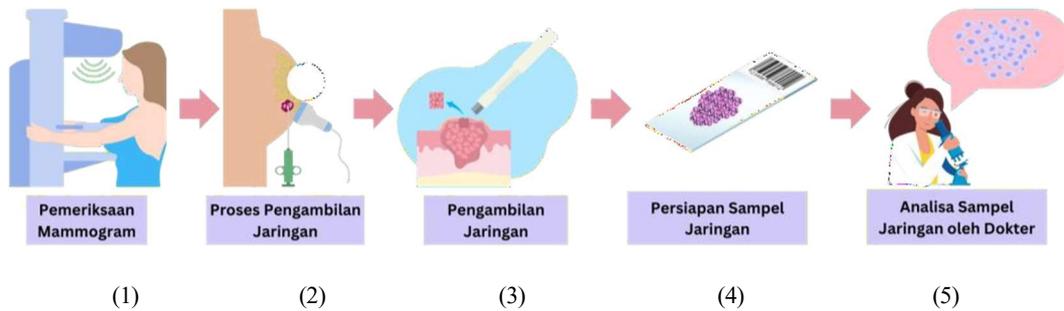
karakteristik unik dari kanker payudara dan kondisi kesehatan pasien. Ini memungkinkan pemilihan terapi yang paling sesuai dan efektif.

3. Mengurangi Risiko Kesalahan: Dengan melibatkan berbagai ahli, risiko kesalahan diagnosa dan terapi dapat dikurangi secara signifikan. Diskusi bersama dan kolaborasi memungkinkan evaluasi menyeluruh dan pemantauan yang lebih cermat.
4. Perawatan Holistik: Pendekatan multidisipliner juga memungkinkan perawatan holistik pasien dengan kanker payudara. Selain menangani aspek medis, perawatan psikologis dan kesejahteraan pasien juga menjadi perhatian. Dengan demikian, pendekatan multidisipliner dalam penanganan kanker payudara adalah pendekatan yang terfokus pada pasien, ilmiah, dan terarah untuk memaksimalkan efektivitas perawatan, meningkatkan prognosis, dan memastikan kualitas perawatan yang optimal bagi setiap pasien yang terkena kanker payudara.

2.1.2 Biopsy dan Histopatologi

Dalam proses menegakkan diagnosa penyakit kanker payudara, setelah proses wawancara pribadi atau anamnesis, dokter akan melakukan beberapa prosedur pemeriksaan untuk menentukan diagnosa penyakit dari pasien. Beberapa diantaranya adalah pemeriksaan mammogram, biopsy dan histopatologi. Seperti yang ditunjukkan pada gambar 3. Pemeriksaan awal dimulai dengan proses pemeriksaan mammogram (1) dimana proses ini pasien akan berdiri dihadapan alat pemindai mammogram kemudian akan dilakukan *scanning* pada payudara pasien dengan menggunakan sinar-x. Setelah pemeriksaan mammogram dan dinyatakan terdapat area yang dicurigai sebagai area tumor atau kanker, maka dilakukan pemeriksaan *ultrasonography* (USG) untuk melihat letak area yang dicurigai dengan lebih teliti kemudian dilakukan pembedahan kecil pada area tersebut untuk mengambil sampel kecil dari jaringan yang dicurigai sebagai tumor atau kanker, proses ini disebut dengan prosedur biopsy (2) dan (3). Setelah dilakukan pengambilan sampel, maka proses selanjutnya merupakan prosedur pemrosesan jaringan. Prosedur ini merupakan proses untuk menyiapkan sampel jaringan untuk pemeriksaan histopatologi (4). Setelah prosedur pemrosesan jaringan selesai, maka

selanjutnya sampel siap untuk diserahkan kepada dokter ahli patologi anatomi untuk dilakukan pengamatan dibawah mikroskop (5).



Gambar 3 Ilustrasi prosedur pemeriksaan penyakit kanker payudara.

1. Mammogram

Mammogram adalah proses pemeriksaan pencitraan medis menggunakan sinar-X rendah untuk menghasilkan proyeksi gambar payudara yang bertujuan untuk mendeteksi keadaan abnormal pada jaringan payudara, termasuk adanya massa yang dapat menjadi tanda-tanda kanker. Ilustrasi pemeriksaan mammogram dapat dilihat pada gambar 4 Proses pemeriksaan mammogram melibatkan kompresi payudara untuk mendapatkan gambar yang lebih jelas dan mengurangi dosis radiasi. Gambar hasil mammogram kemudian dievaluasi oleh radiolog untuk mencari tanda-tanda perubahan yang mencurigakan.

Pemeriksaan mammogram mampu membantu untuk melihat keberadaan ketidaknormalan pada jaringan payudara, namun suatu mammogram tidak bisa memastikan bahwa suatu daerah abnormal adalah kanker. Untuk mengonfirmasi bahwa jaringan abnormal tersebut merupakan kanker, sebagian kecil dari jaringan abnormal tersebut harus diambil dan diperiksa di bawah mikroskop. Prosedur investigasi ini disebut biopsi.



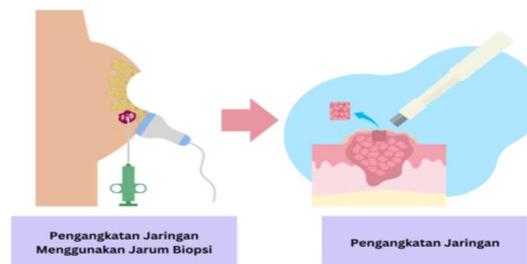
(a)



Gambar 4 Pemeriksaan Payudara (a) Ilustrasi proses pemeriksaan mammogram, (b) Gambar hasil pemeriksaan mammogram.

2. Biopsi

Biopsi merupakan tes yang dilakukan untuk mengangkat sel atau jaringan payudara untuk keperluan pemeriksaan oleh ahli patologi di bawah mikroskop dengan tujuan untuk memastikan keberadaan penyakit atau kondisi tertentu pada jaringan tubuh . Dalam beberapa kasus, biopsy dapat digunakan untuk menghilangkan kelainan kecil pada payudara sepenuhnya. Ada beberapa jenis biopsi yang dapat dilakukan untuk mengambil sampel jaringan payudara, antara lain biopsi jarum halus atau *Fine Needle Aspiration Biopsy (FNAB)*, *Cyst Aspiration for FNAB*, Biopsi inti atau *Core Needle Biopsy* dan Biopsi dengan bantuan vakum atau *Mammotome (Breast Cancer Surgery Melbourne, VIC , no date)*.



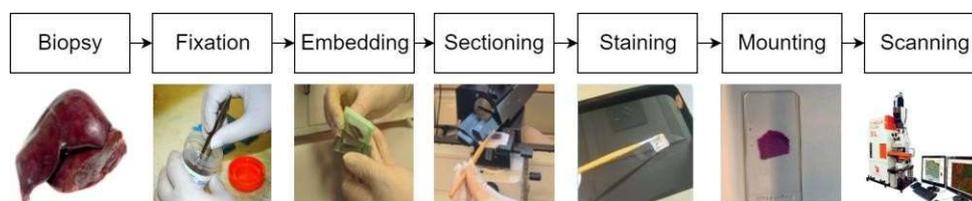
Gambar 5 Ilustrasi proses *Fine Needle Aspiration Biopsy (FNAB)*

Proses biopsy yang paling umum digunakan untuk pemeriksaan payudara adalah prosedur biopsy jarum halus atau *fine needle aspiration biopsy (FNAB)*. Prosedur ini dilakukan dengan bantuan alat USG atau CT scan yang digunakan untuk mencari posisi keberadaan jaringan abnormal (lihat Gambar 4), kemudian akan dilanjutkan dengan proses pengangkatan sampel jaringan menggunakan jarum biopsy yang disuntikkan ke tubuh pasien untuk mengambil

sampel jaringan (*Edward O. Uthman, MD , 2012*). Setelah proses pengangkatan jaringan selesai, selanjutnya sampel harus diproses terlebih dahulu sebelum dilakukan pengamat dibawah mikroskop.

3. Pemrosesan Jaringan (*Tissue Processing*)

Tahap-tahap pemrosesan jaringan (*tissue processing*) dalam pemeriksaan histopatologi adalah serangkaian langkah yang dilakukan dalam laboratorium patologi untuk mempersiapkan jaringan yang diambil selama biopsi atau operasi menjadi potongan tipis yang dapat dilihat di bawah mikroskop. Proses ini memungkinkan patologis atau ahli histopatologi untuk menganalisis jaringan dan membuat diagnosis yang ditunjukkan oleh Gambar 6.



Gambar 6 Ilustrasi pemrosesan jaringan (Mahbod et al., 2023).

- *Fixation* (Fiksasi)

Tahap pertama adalah fiksasi, di mana sampel jaringan yang diambil selama biopsi atau operasi diberi cairan fiksatif. Cairan ini biasanya berbasis formalin dan berfungsi untuk menjaga integritas struktural sel dan menghentikan perubahan seluler yang dapat terjadi setelah pengambilan sampel. Setelah fiksasi, sampel jaringan harus dihilangkan airnya. Ini dilakukan dengan melewati sampel melalui serangkaian larutan alkohol yang semakin tinggi konsentrasinya. Proses dehidrasi ini bertujuan untuk menggantikan air dalam jaringan dengan alkohol sehingga jaringan bisa diinfiltrasi oleh paraffin. Selanjutnya, sampel jaringan masih mengandung alkohol. Untuk menghilangkan alkohol dan membuat jaringan transparan, sampel direndam dalam zat pembening (*clearing agent*), seperti *xilol* atau *toluen*. Sampel jaringan yang telah dihilangkan airnya dan dibuat transparan selanjutnya diinfiltrasi dengan paraffin cair. Paraffin adalah bahan padat yang akan menggantikan zat pembening dan meresap ke dalam jaringan. Ini akan membantu menjaga struktur jaringan saat potongan tipis dibuat.

- *Embedding* (Pemasanan)

Setelah proses infiltrasi, sampel jaringan ditempatkan dalam cetakan (mold) yang berisi paraffin cair. Ketika paraffin mendingin dan mengeras, ini akan membentuk blok jaringan yang mengandung sampel.

- *Sectioning* (Pemotongan)

Blok jaringan yang telah dibuat kemudian dipotong menjadi potongan tipis menggunakan mikrotom. Potongan ini sangat tipis (biasanya sekitar 4-5 mikrometer) dan ditempatkan pada *slide* kaca.

- *Staining* (Pewarnaan)

Potongan jaringan pada slide kaca kemudian diwarnai dengan pewarna histologi tertentu, seperti *hematoxylin* dan *eosin* (H&E). Pewarnaan ini akan memberikan warna berbeda untuk berbagai komponen seluler, memungkinkan patologis untuk mengidentifikasi struktur dan perubahan patologis.

- *Mounting* (Pemasangan)

Setelah pewarnaan, potongan jaringan ditutupi dengan kaca penutup dan larutan *mounting* untuk melindungi dan melindungi sampel.

- *Microscopic Examination* (Pemeriksaan Mikroskopis)

Slide jaringan yang telah selesai diproses akan dianalisis di bawah mikroskop oleh ahli patologi atau ahli histopatologi. Mereka akan mengamati dan menganalisis struktur jaringan, mengidentifikasi perubahan patologis, dan membuat diagnosis (Mahbod *et al.*, 2023).

Tahap-tahap ini sangat penting dalam pemeriksaan histopatologi karena memastikan bahwa sampel jaringan yang diambil dapat dianalisis secara akurat di bawah mikroskop untuk membuat diagnosis yang tepat. Proses ini juga memungkinkan pengarsipan jaringan untuk konsultasi dan pemantauan jangka panjang.

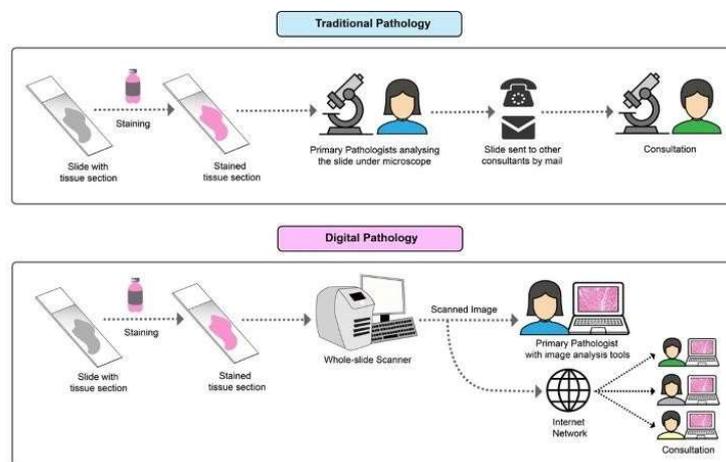
4. Histopatologi

Histopatologi merupakan cabang dari ilmu patologi yang berkaitan dengan studi mikroskopis jaringan biologis untuk mengidentifikasi perubahan patologis dalam jaringan dan organ manusia. Disiplin ilmu ini sangat penting dalam

membantu menentukan diagnosis, pemantauan penyakit bahkan untuk menentukan rencana penanganan dan perawatan pasien. Tujuan utama histopatologi adalah untuk mendiagnosis penyakit, mengidentifikasi perubahan patologis dalam jaringan, dan memahami dasar biologis penyakit.

2.1.3 Patologi Digital

Patologi digital adalah bidang ilmu dalam patologi yang menggunakan teknologi digital untuk mengambil, mengelola dan menganalisis gambar mikroskopis jaringan dan sel dalam konteks medis. Pendekatan patologi digital ini menggantikan pendekatan analisis konvensional patologi dimana potongan jaringan dianalisis secara manual dibawah mikroskop. Secara tradisional proses patologi dipelajari melalui penggunaan mikroskop untuk memeriksa jaringan dan organ pasien. Namun, kemajuan teknologi yang membawa teknologi pencitraan digital dalam bidang patologi membawa kemajuan histologi pada tingkat yang lebih tinggi, yang sekarang dikenal sebagai patologi digital (Arshad, 2022). Ilustrasi perbandingan patologi konvensional dan patologi digital ditunjukkan oleh gambar 7.



Gambar 7 Patologi Konvensional vs Patologi Digital (*Precision Oncology Ireland*).

Meskipun patologi telah menggunakan mikroskop sejak abad ke-17, teknologi terkini patologi digital memungkinkan para ahli patologi untuk melihat dan mendiagnosis slide dalam bentuk digital. Sistem patologi digital mengonversi preparat kaca yang biasanya digunakan untuk analisa jaringan dibawah mikroskop menjadi preparat digital beresolusi tinggi yang dilihat dan dianalisis menggunakan

komputer. Preparat-preparat digital beserta informasi yang menyertai dapat dibagikan dengan cepat kepada banyak orang, sehingga memudahkan ahli patologi untuk melakukan analisa secara lebih detail juga memudahkan mengirimkan sampel preparat dalam bentuk digital jika ingin melakukan konsultasi dengan dokter ahli lain terkait sebuah kasus yang dihadapi (Soenksen, 2009). Beberapa istilah populer dalam patologi digital antara lain pencitraan digital, mikroskop virtual, *whole slide images* (WSI) dan telepatologi.

1. Pencitraan Digital

Pencitraan digital merujuk pada proses penyimpanan informasi patologi anatomi dalam bentuk gambar digital, termasuk foto dari spesimen patologi makroskopis dan gambar dari preparat mikroskopis. Data gambar yang tersimpan dalam format digital memiliki berbagai aplikasi, termasuk penggunaan dalam edukasi medis, penciptaan pustaka medis, pengendalian kualitas laboratorium, diagnosis medis, dan pembuatan laporan pasien. Selama beberapa tahun terakhir, pencitraan digital telah muncul sebagai solusi yang sangat efektif untuk mengatasi keterbatasan penyimpanan informasi mikroskopis pada medium tradisional, seperti preparat kaca yang rentan terhadap kerusakan, pemudaran, serta memerlukan ruang penyimpanan yang signifikan.



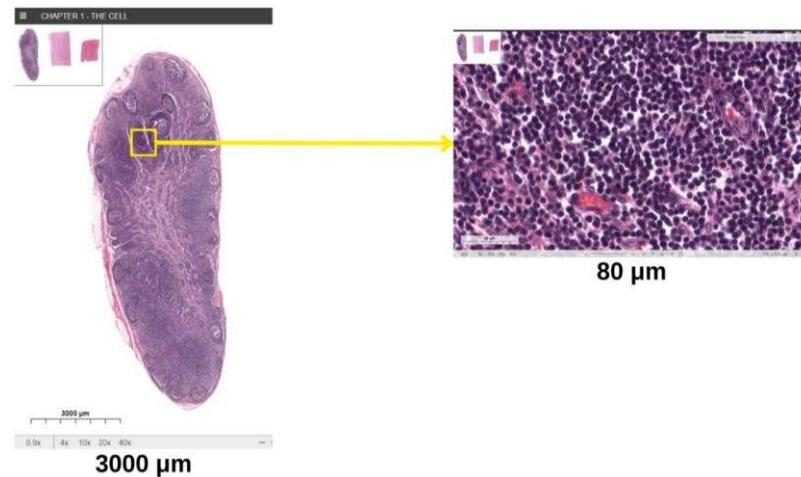
Gambar 8 Proses analisa preparat digital

Gambar 8 memperlihatkan contoh situasi seorang ahli patologi anatomi dalam menganalisis preparat digital. Pencitraan digital juga melibatkan penerapan algoritma-algoritma digital yang rumit untuk analisis dan kuantifikasi

gambar-gambar mikroskopis atau makroskopis yang disimpan dalam format digital. Algoritma-algoritma ini dirancang untuk memproses gambar secara otomatis atau semi-otomatis, sehingga memungkinkan analisa citra secara kuantitatif. Salah satu aplikasi utama dari perangkat lunak pencitraan digital adalah kuantifikasi morfometrik. Ini berarti perangkat lunak dapat mengukur berbagai parameter yang berkaitan dengan bentuk, ukuran, dan geometri objek dalam gambar. Contohnya termasuk menghitung ukuran sel, luas area tertentu, bentuk nukleus, dan sebagainya. Informasi morfometrik ini dapat memberikan wawasan yang mendalam tentang struktur jaringan dan sel, yang penting dalam penelitian ilmiah dan diagnosis medis (Słodkowska and Rojo, 2011).

2. Mikroskop Virtual

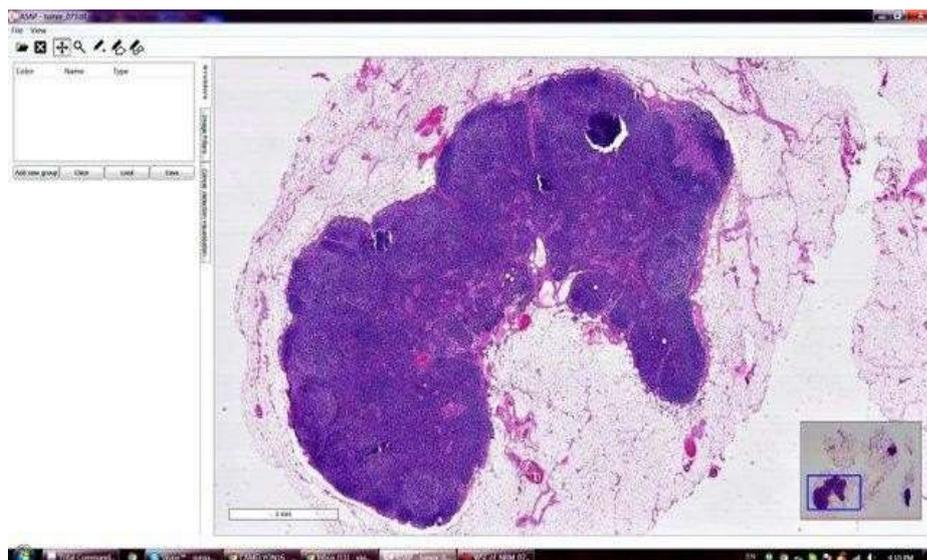
Mikroskop virtual adalah teknologi yang memungkinkan pengambilan citra digital dengan resolusi tinggi terhadap preparat kaca atau spesimen mikroskopis. Mikroskop virtual merupakan pendekatan yang lebih modern dalam pengamatan objek mikroskopis yang biasanya dilakukan dengan mikroskop konvensional. Mikroskopi virtual melibatkan pemindaian atau pengambilan gambar digital dari spesimen yang ditempatkan di bawah mikroskop. Hasil pemindaian ini menghasilkan citra beresolusi tinggi yang artinya memiliki tingkat detail yang sangat baik. Hal ini memungkinkan pengamatan objek mikroskopis dengan detail yang tinggi yang sering kali sangat penting dalam diagnostik medis dan penelitian ilmiah. Selain menghasilkan gambar dengan tingkat resolusi yang tinggi, mikroskop virtual juga memiliki kemampuan untuk memperbesar area yang ingin diamati, contoh mikroskop digital ditunjukkan pada gambar 9. Hal ini memungkinkan pengguna untuk dapat menjelajahi objek mikroskopis dengan cara yang sangat fleksibel dan mendalam, bahkan pada tingkat pembesaran yang berbeda. Dalam hal kemudahan akses, sistem mikroskopi virtual juga dapat diakses melalui internet, yang memungkinkan patolog atau peneliti dari lokasi yang berbeda untuk melihat dan berinteraksi dengan citra mikroskopis yang sama. Ini sangat bermanfaat untuk konsultasi jarak jauh, kolaborasi, dan edukasi (Gabril and Yousef, 2010).



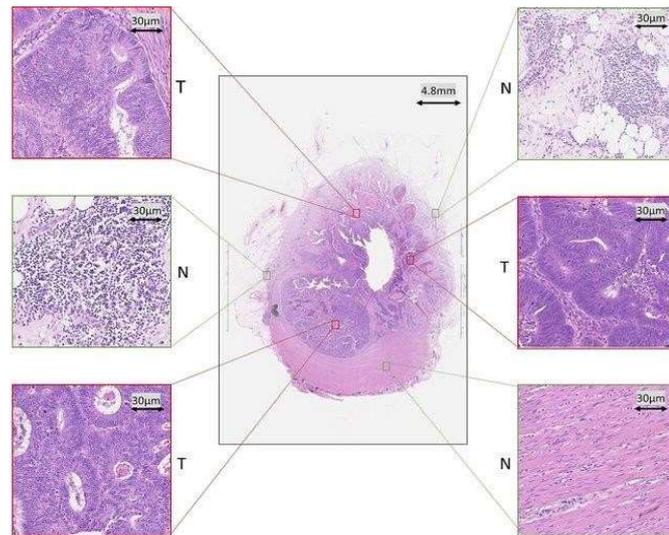
Gambar 9 Tampilan mikroskop virtual (H&E Staining | The Cell)

3. *Whole Slide Images (WSI)*

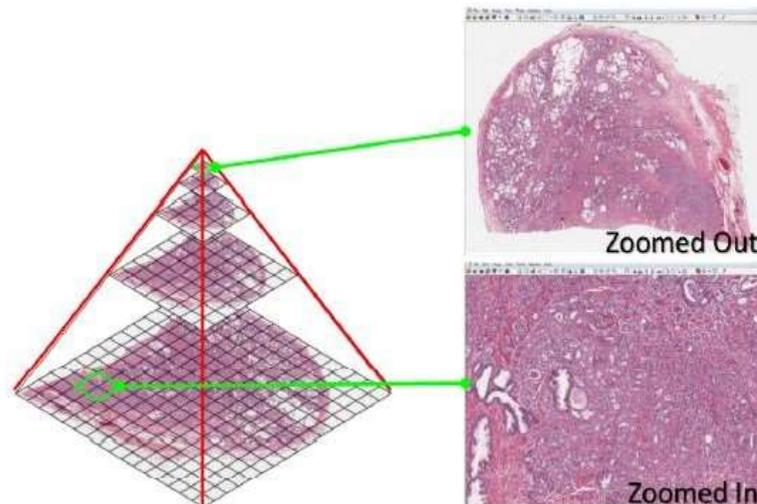
Whole Slide Images (WSI) dalam patologi digital merujuk pada gambar digital beresolusi tinggi yang mencakup seluruh spesimen mikroskopis yang telah dipindai atau difoto. Hal ini berarti bahwa bukan hanya bagian tertentu dari spesimen yang diambil gambar, tetapi seluruh permukaan spesimen yang dapat dilihat pada slide kaca dipindai menjadi gambar digital. Hal ini memungkinkan ahli patologi melakukan analisa dalam berbagai tingkatan pembesaran mulai dari 10X, 20X 30X dan 40X. Varian pembesaran mikroskop ditunjukkan oleh gambar 10, 11 dan 12.



Gambar 10 Citra WSI asli (Kovalev, Kalinovsky and Liauchuk, 2016).



Gambar 11 Citra WSI dan perbesarannya (Qaiser et al., 2019).



Gambar 12 Ilustrasi citra WSI dalam format piramida (Lehtonen, 2021).

4. Telepatologi

Telepatologi adalah bidang dalam patologi digital yang memanfaatkan teknologi komunikasi jarak jauh untuk memungkinkan patologis dan profesional medis lainnya untuk melakukan pemeriksaan dan konsultasi patologi tanpa harus berada di lokasi fisik yang sama dengan spesimen yang akan dianalisis. Ini melibatkan penggunaan komunikasi video, gambar digital, dan sistem manajemen data yang canggih untuk memfasilitasi kolaborasi antar patologis dan diagnosis jarak jauh (Wargasetia, 2015). Pada situasi patolog terpisah jauh dari preparat, patologi digital dapat meningkatkan kecepatan bekerja, mengedepankan kasus prioritas,

meningkatkan kolaborasi antara patolog, ahli bedah, dan teknologi laboratorium, distribusi beban kerja yang wajar, mengurangi isolasi dari patolog yang bekerja di wilayah yang jauh, dan meningkatkan akses untuk edukasi dan aktivitas jaminan mutu. Untuk sistem jaminan kesehatan dan pasien, patologi digital meningkatkan efisiensi, menurunkan kehilangan dan kerusakan selama transportasi preparat, menurunkan biaya perjalanan patolog dan/atau transportasi spesimen, mengurangi interupsi pelayanan yang disebabkan oleh ketiadaan patolog, menurunkan transfer pasien dan pembedahan dua langkah, dan meningkatkan rekrutmen ahli bedah lokal dan atau patolog dengan lokasi yang jauh. Singkatnya, patologi digital memfasilitasi peningkatan kualitas dan kelanjutan perawatan (Têtu and Evans, 2014).

2.1.4 Computer-Aided Pathology

Pemeriksaan patologi telah menjadi standar emas penentu diagnosis kanker. Pemeriksaan ini memegang peranan penting karena diagnosis dan penentuan stadium kanker akan membantu menentukan pilihan pengobatan. Namun, proses investigasi *slide* secara manual merupakan pekerjaan yang sangat melelahkan dan memakan waktu. Interpretasi yang salah mungkin saja terjadi karena kelelahan atau stres pada dokter spesialis. Selain itu, kekurangan ahli patologi terdaftar menjadi masalah di banyak negara (Hsu *et al.*, 2021). Oleh karena itu, sistem diagnosis berbantuan komputer (CAD) dikembangkan untuk membantu ahli patologi dalam penilaian slide dan berfungsi sebagai pertimbangan analisa untuk meringankan beban kerja ahli patologi (Hsu *et al.*, 2021).

Patologi berbantuan komputer (*Computer-Aided Pathology*) atau patologi komputasi adalah bidang dalam patologi digital yang menggabungkan teknologi komputer dan kecerdasan buatan untuk membantu patologis dalam mendiagnosis penyakit dan menganalisis gambar patologi digital. Berawal dari digitalisasi preparat kaca kedalam bentuk gambar digital berupa *whole slide images* (WSI) kini seiring dengan berkembangnya patologi komputasi dan teknologi visi komputer memungkinkan pendekatan berbasis kecerdasan buatan untuk melakukan deteksi, segmentasi, diagnosis dan analisa gambar digital secara kuantitatif (Nam *et al.*, 2020), (Bera *et al.*, 2019).

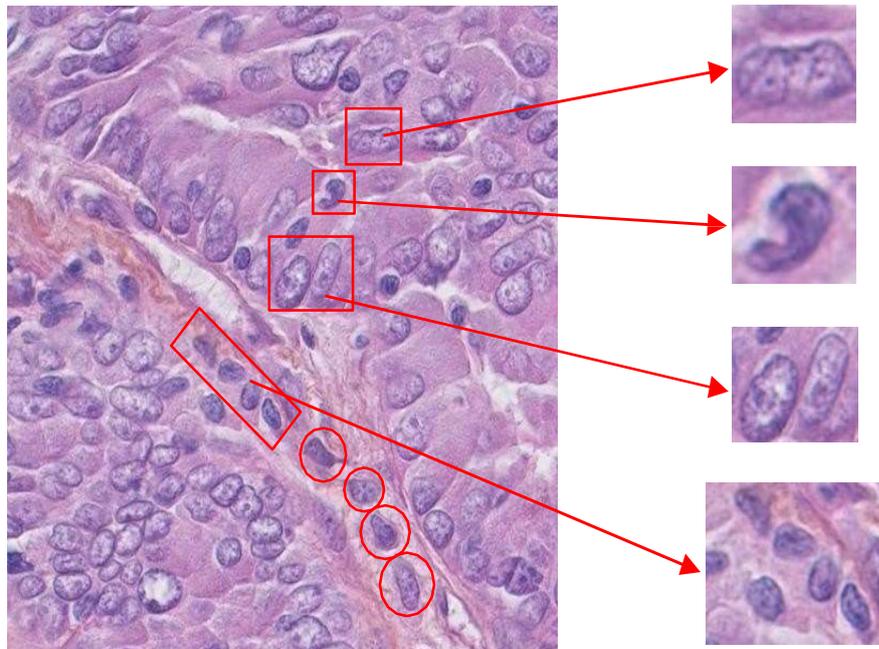
Beberapa tahap dasar dan awal dalam penerapan teknologi visi komputer dalam patologi berbantuan komputer adalah pada tugas segmentasi sel, deteksi sel melalui inti, sitoplasma dan struktur. Karena sel merupakan unit dasar gambar histopatologi, oleh karena itu identifikasi warna, intensitas, morfologi dan sitoplasma melalui segmentasi sel merupakan langkah pertama dan terpenting dalam analisa gambar histopatologi. Segmentasi sel telah diterapkan dalam analisis imunohistokimia (IHC) dan pewarnaan hematoksilin dan eosin (H&E) dan merupakan topik mendasar untuk analisis gambar komputasi untuk mencapai representasi histopatologi kuantitatif (Nam *et al.*, 2020).

1. Segmentasi Sel

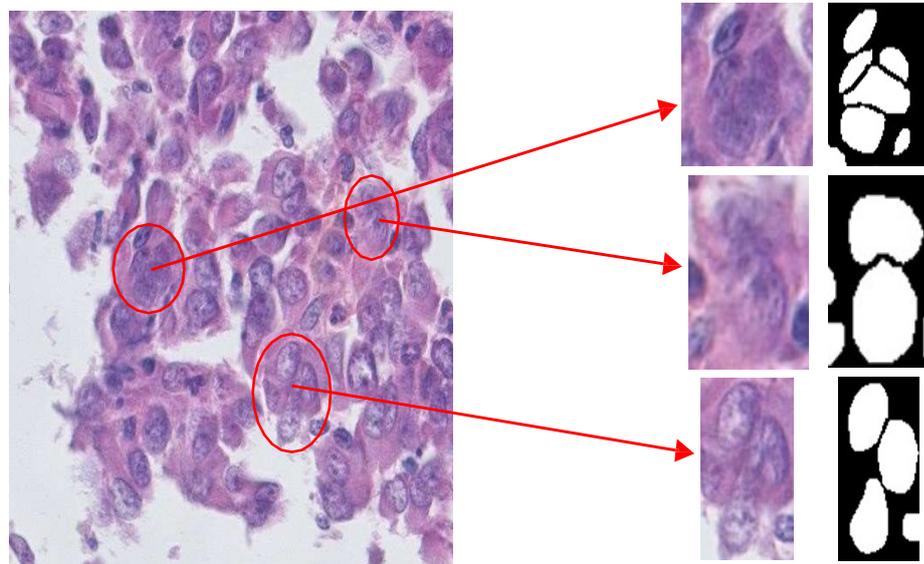
Penerapan teknik pengolahan citra dalam menganalisis citra mikroskopis sel kanker payudara pada komputasi patologi telah banyak dilakukan juga dikembangkan. Segmentasi sel merupakan permasalahan yang penting untuk diangkat dalam pengolahan citra histopatologi sel kanker payudara selain karena segmentasi merupakan proses yang paling mendasar dalam analisis citra histopatologi, proses ini juga sangat berpengaruh terhadap beberapa proses analisa jaringan lanjut seperti perhitungan jumlah sel, klasifikasi tipe sel, klasifikasi tipe jaringan, penentuan tingkat keganasan penyakit hingga rencana pengobatan pasien.

Proses segmentasi mengacu pada proses membagi piksel pada citra kedalam beberapa kategori sesuai dengan objek, hal ini bertujuan agar objek yang dimaksud terpisah dengan objek lainnya. Pada penerapannya menganalisis citra histopatologi merupakan pekerjaan yang menantang karena beberapa alasan seperti ketersediaan dataset citra histopatologi kanker payudara yang telah terannotasi yang terbatas, morfologi sel yang bervariasi ditunjukkan pada gambar 13, variasi warna pada citra histopatologi yang disebabkan oleh inkonsistensi pewarnaan (Rashmi *et al.*, 2020) dan batas yang tidak jelas antar sel (Hassan *et al.*, 2020) ditunjukkan pada gambar 14 yang menyebabkan objek tersegmentasi sebagai satu objek sehingga menyebabkan segmentasi sel kanker tidak maksimal. Selain itu adanya penampakan sel yang berkelompok yang berdempet dan tumpang

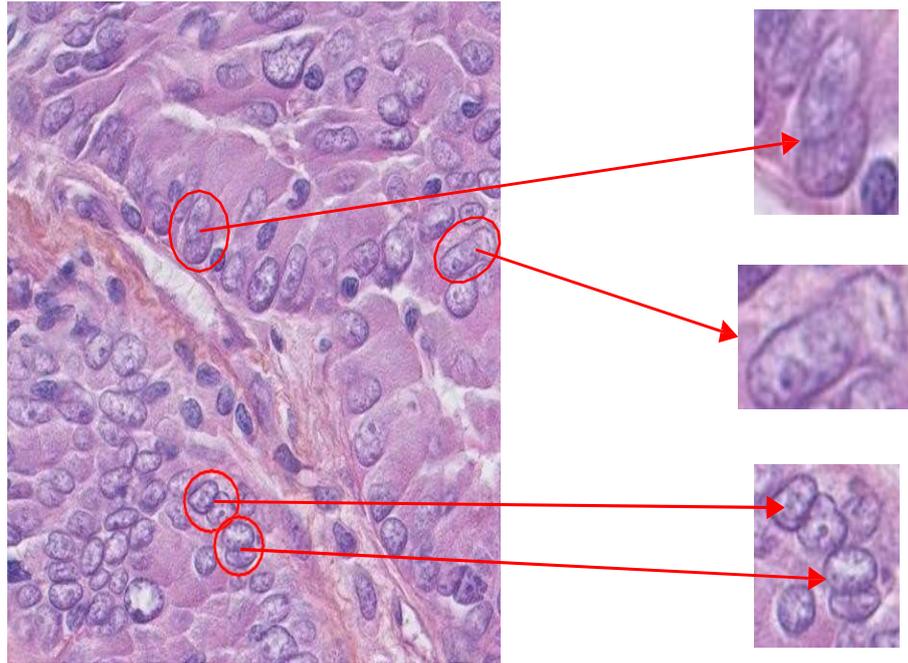
tindih pada sel kanker juga menjadi masalah selanjutnya dalam menganalisis citra histopatologi sehingga perlu dilakukan pemisahan sel yang berkelompok dan tumpang tindih untuk menggambarkan batas antar sel sehingga mampu menunjukkan struktur morfologi dari sel secara jelas. Gambaran sel berdempet dan tumpang tindih ditunjukkan pada gambar 15.



Gambar 13 Variasi bentuk dan ukuran sel (Naylor et al., 2017a)

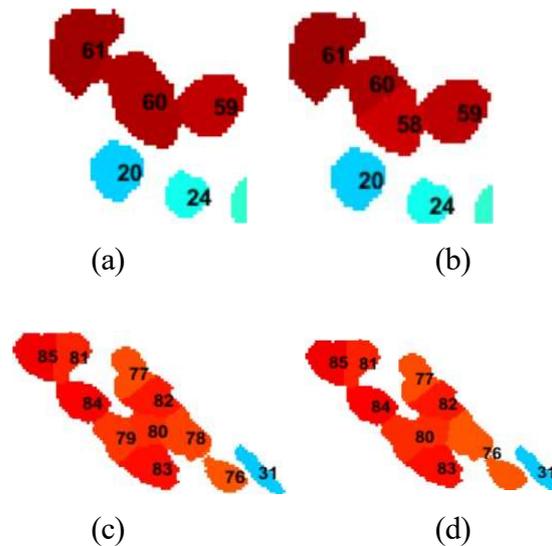


Gambar 14 Batas yang halus antar sel (Naylor et al., 2017a)



Gambar 15 Penampakan sel berkelompok yang berdekatan, berdempet dan tumpang tindih (Naylor et al., 2017a)

Proses segmentasi sel kanker merupakan proses yang sangat penting dan mendasar dalam analisis citra histopatologi karena akan berpengaruh pada proses selanjutnya seperti proses *grading* tingkat keganasan kanker, perhitungan jumlah sel kanker, dan klasifikasi tipe sel. Namun hal ini memiliki kesulitan terutama pada citra histopatologi yang meliputi keberadaan objek yang mengalami oklusi dan sebaran kluster sel yang luas dan saling berdekatan yang menyebabkan rentan terjadinya *under segmentation* dan *over segmentation* yang dapat menghambat proses pengidentifikasian jumlah sel yang akurat. Kedua, batas yang sangat halus antar sel dan variasi warna pada citra histopatologi sehingga menyebabkan kesulitan untuk melihat batas antar sel yang saling tumpang tindih. Ketiga, variabilitas dalam penampakan sel seperti, ukuran, bentuk dan kepadatan sel yang membutuhkan metode yang mampu melakukan identifikasi berbagai bentuk dan ukuran objek secara optimal. Selain itu, terbatasnya dataset citra histopatologi yang terannotasi juga menjadi tantangan dalam melakukan penelitian ini.

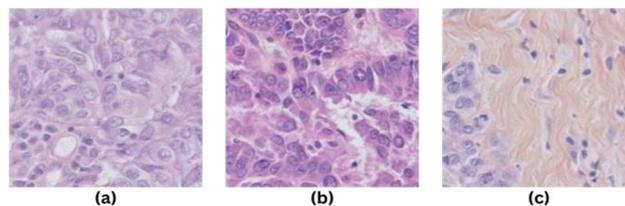


Gambar 16 Contoh under dan over-segmentation. (a) dan (c) citra *ground truth*, (b) *over-segmentation*, (d) *under-segmentation* (Tuwohingide and Fatichah, 2017).

Gambar 16 menunjukkan contoh kasus *under* dan *over segmentation* yang terjadi pada segmentasi sel pada citra sel payudara. Peristiwa *under segmentation* merupakan istilah yang sering digunakan dalam konteks pemrosesan citra, khususnya dalam segmentasi citra. Kedua istilah tersebut mengacu pada ketidakakuratan dalam memisahkan objek atau wilayah dalam citra. *Under segmentation* terjadi ketika beberapa objek yang seharusnya dipisahkan dalam citra dianggap sebagai satu objek atau wilayah tunggal oleh algoritma segmentasi. Dalam konteks ini algoritma gagal memisahkan objek atau wilayah yang berbeda dengan benar, yang mengakibatkan objek atau wilayah yang berbeda digabung menjadi satu, sedangkan *over segmentation* terjadi ketika satu objek atau wilayah dipecah menjadi beberapa bagian oleh algoritma segmentasi. Dalam konteks ini, algoritma mendeteksi terlalu banyak batas atau batas yang tidak perlu, yang mengakibatkan satu objek atau wilayah yang seharusnya kontinu terpecah menjadi beberapa bagian. Dalam banyak aplikasi pemrosesan citra, salah satu tantangan utama adalah menemukan keseimbangan yang tepat antara *under-segmentation* dan *over-segmentation*, untuk memastikan bahwa objek dan wilayah dalam citra dapat diidentifikasi dengan akurat.

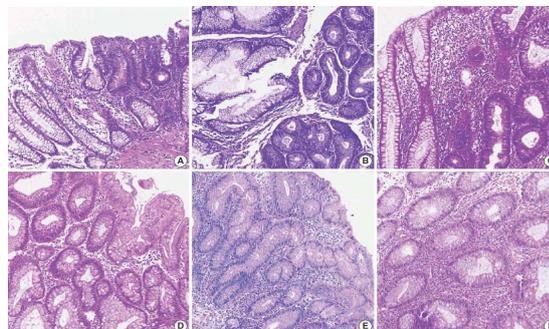
2. Pewarnaan *Hematoxylin and Eosin* (H&E)

Pewarnaan Hematoxylin dan Eosin (H&E) adalah teknik pewarnaan yang paling umum digunakan dalam patologi untuk evaluasi mikroskopik jaringan. Ini memberikan gambaran kontras tinggi dari berbagai komponen jaringan, yang memungkinkan patolog untuk menilai morfologi dan arsitektur jaringan. Hematoxylin digunakan sebagai pewarna dasar, ia mengikat dan mewarnai struktur yang kaya akan basa atau ion positif, seperti DNA dan RNA dan hasilnya nukleus sel biasanya tampak biru gelap atau ungu. Sementara Eosin digunakan sebagai pewarna asam, ia mengikat dan mewarnai komponen yang kaya akan ion negatif, seperti protein sitoplasma dan hasilnya berupa sitoplasma dan beberapa bagian ekstraseluler (seperti kolagen) tampak merah atau merah muda. Contoh citra histopatologi ditunjukkan oleh gambar 17.



Gambar 17 Contoh gambar histopatologi (Nam et al., 2020)

Pewarnaan H&E memberikan kontras yang baik antara berbagai komponen jaringan, memungkinkan ahli patologi untuk membedakan antara tipe sel yang berbeda, struktur jaringan, serta perubahan patologis. Teknik ini sangat berharga untuk mendiagnosis berbagai kondisi, termasuk inflamasi, infeksi, neoplasia, dan banyak kondisi patologis lainnya.



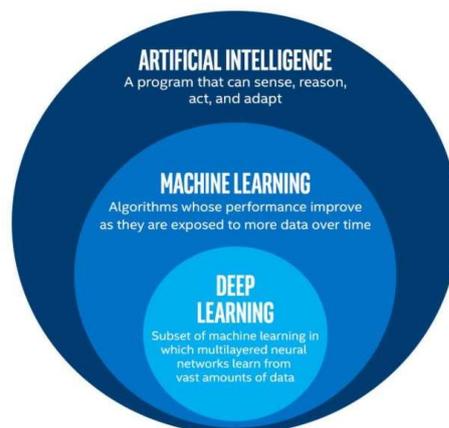
Gambar 18 Variasi hasil pewarnaan Hematoxylin and Eosin (H&E) (Nam et al., 2020)

Dalam pewarnaan sampel histopatologi kualitas kaca objek yang

diwarnai dengan H&E bervariasi menurut institusi dan dipengaruhi oleh konsentrasi pewarna, waktu pewarnaan, waktu fiksasi formalin, pembekuan, keterampilan memotong, jenis kaca objek, dan warna yang memudar setelah pewarnaan, *slide* yang diwarnai dengan teknik H&E menunjukkan keragaman padahal slide tersebut berasal dari institusi yang sama (Gambar 18A, dan D). Warna setiap slide juga dipengaruhi oleh pemindai dan pengaturan slide. Jika terdapat variabilitas pewarnaan yang berlebihan dalam kumpulan data, penerapan ambang batas dapat menghasilkan hasil yang berbeda karena protokol pewarnaan atau pencitraan yang berbeda, bukan karena karakteristik jaringan yang unik (Nam *et al.*, 2020).

2.1.5 Artificial Intelligence (AI)

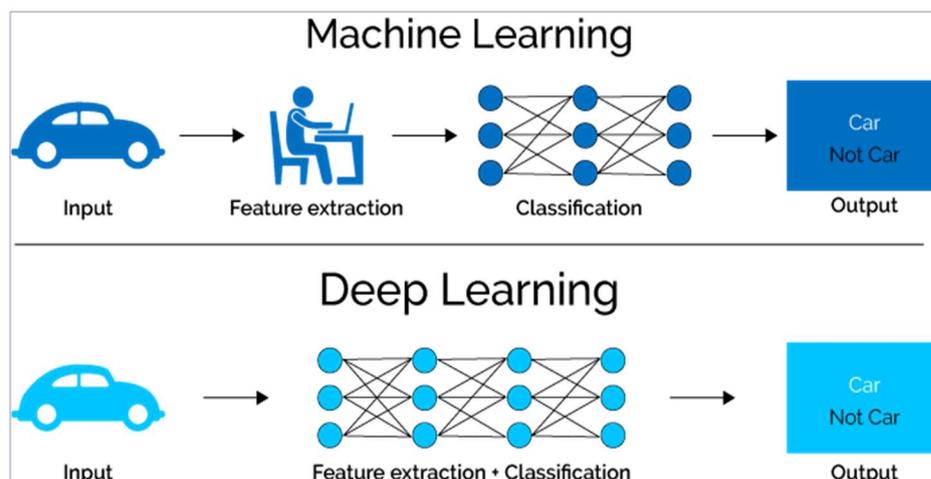
Artificial Intelligence (AI) atau Kecerdasan Buatan adalah cabang ilmu komputer yang berfokus pada pembuatan atau penerapan sistem yang mampu melakukan tugas yang biasanya memerlukan kecerdasan manusia. Ini mencakup tugas-tugas seperti pemahaman bahasa alami, pengenalan pola, pemecahan masalah, pembelajaran, dan pengambilan keputusan. *AI* adalah topik yang mencakup bidang ilmu *Machine Learning* (ML) dan *Deep Learning* (DL). Gambar 19 menunjukkan hubungan antara *Artificial Intelligence* (AI), *Machine Learning* (ML) dan *Deep Learning* (DL).



Gambar 19 Hubungan antara AI, ML, dan DL

Machine Learning (ML) adalah bagian dari *AI* yang menggunakan metode statistik. Dengan metode tersebut memungkinkan mesin (komputer) untuk mengambil keputusan berdasarkan data. Algoritma *machine learning* dirancang

agar dapat belajar dan kemampuannya meningkat seiring waktu ketika terdapat data baru, tanpa diprogram secara eksplisit (Singh, 2018). *Deep Learning (DL)* adalah jenis dari *machine learning* yang terinspirasi oleh fungsi sel otak manusia yang disebut neuron. Dari inspirasi itulah terdapat konsep yang dinamakan *Neural network (NN)*. Namun pada *Deep learning neural network*-nya lebih dari dua lapisan. Selain itu terdapat beberapa hal yang menjadi evolusi *neural network* pada *deep learning* lebih banyak neuron dari jaringan sebelumnya, cara yang lebih kompleks untuk menghubungkan lapisan/neuron, dan ekstraksi fitur otomatis (Vinet and Zhedanov, 2011). Terdapat distingsi fundamental antara pendekatan algoritma pembelajaran mesin (*Machine Learning, ML*) tradisional dan pendekatan berbasis jaringan saraf dalam (*Deep Learning, DL*) dalam konteks penglihatan komputer (*Computer Vision*). Dalam metodologi ML tradisional, proses ekstraksi fitur harus dilakukan secara eksplisit untuk menghasilkan vektor fitur, yang selanjutnya akan diklasifikasikan oleh model pembelajaran mesin tertentu. Sebagai kontras, dalam pendekatan DL, jaringan saraf dalam secara implisit bertanggung jawab baik untuk ekstraksi fitur maupun proses klasifikasi. Proses operasional keduanya dapat divisualisasikan dalam Gambar 20.

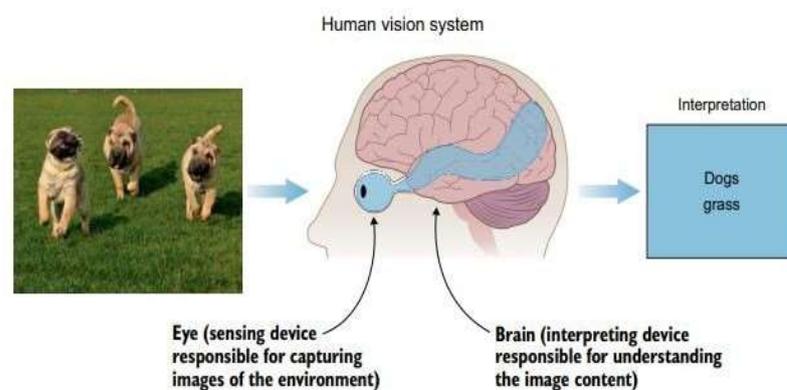


Gambar 20 Cara kerja *Machine Learning* dan *Deep Learning*

2.1.6 Visi Komputer

Visi komputer adalah bidang ilmu komputer yang memfokuskan pada pengembangan teknik-teknik yang memungkinkan mesin untuk mendapatkan informasi dari data visual seperti gambar atau urutan video. Jika cara manusia dalam mengenali suatu benda adalah dengan melihat benda tersebut dengan mata lalu

kemudian memproses informasi yang disaksikan kemudian menginterpretasikan objek yang dilihat seperti ilustrasi pada gambar 21, maka dalam visi komputer menggunakan perangkat seperti kamera atau sensor untuk merekam atau mengambil data masukan kemudian algoritma pada komputer memproses untuk mengenali data kemudian membuat prediksi objek tersebut. Tujuan teknologi visi komputer adalah untuk menyimulasikan kemampuan manusia dalam menginterpretasi dan memahami dunia visual melalui mata, tetapi menggunakan kamera dan komputer. Dengan kata lain, visi komputer berupaya untuk memahami dan memanfaatkan informasi yang terkandung dalam data visual.

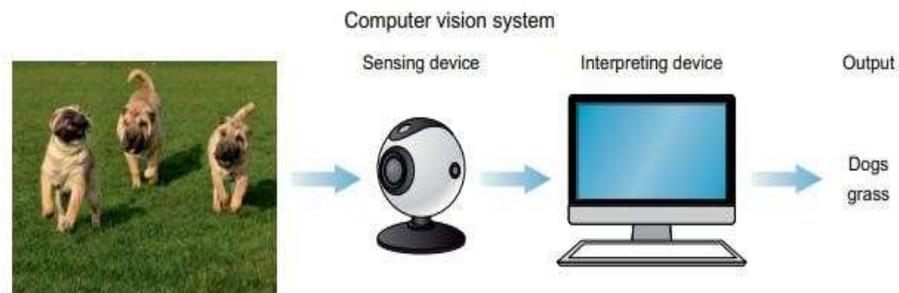


Gambar 21 Sistem penglihatan manusia (Elgendy, 2020)

Selama beberapa dekade terakhir, metode pemrosesan gambar konvensional sering kali disamakan dengan sistem visi komputer, tetapi persamaan tersebut tidak sepenuhnya tepat. Ada perbedaan mendasar antara mesin yang sekadar mengolah gambar dengan mesin yang benar-benar memahami isi dari gambar tersebut. Pemahaman ini melibatkan kompleksitas yang jauh lebih tinggi. Saat ini, pemrosesan gambar dianggap hanya sebagai salah satu komponen dalam kerangka kerja visi komputer yang lebih luas, yang fokus pada interpretasi kontekstual dari gambar. Pada tingkat tertinggi, sistem penglihatan pada manusia, hewan, serangga, dan sebagian besar organisme hidup hampir sama. Mereka terdiri dari sensor atau mata untuk menangkap gambar dan otak untuk memproses dan menafsirkan gambar. Sistem kemudian mengeluarkan prediksi komponen gambar berdasarkan data yang diekstraksi dari gambar tersebut. Visualisasi ditunjukkan oleh gambar 22.

Dalam upaya meniru kemampuan visual manusia, para peneliti telah memperoleh kemajuan signifikan dalam beberapa tahun terakhir. Untuk

menyimulasikan sistem penglihatan manusia dalam mesin, dua elemen krusial diperlukan: kamera dan sensor yang dapat menggantikan fungsi mata dan algoritma canggih yang dapat menirukan cara otak manusia menginterpretasi dan mengategorikan informasi visual (Elgandy, 2020).



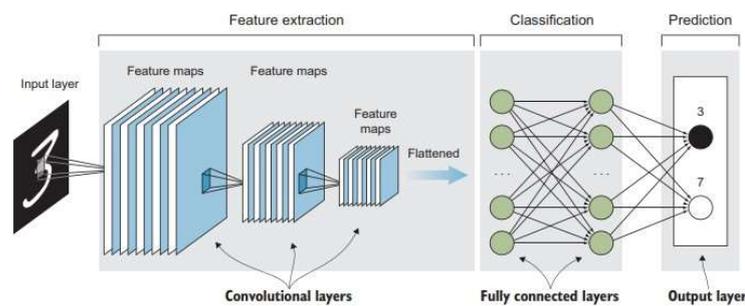
Gambar 22 Komponen sistem visi komputer (Elgandy, 2020).

2.1.7 Convolutional Neural Network (CNN)

Convolution Neural Network (CNN) merupakan bagian dari *deep learning* dan dikenal sebagai salah satu struktur jaringan saraf yang telah secara luas diaplikasikan untuk tugas-tugas deteksi, analisis dan klasifikasi gambar. Karakteristik khas CNN adalah kemampuannya dalam mendeteksi pola spasial dan temporal pada gambar, memberikan pemahaman konten gambar yang lebih holistik. Meskipun memiliki kesamaan dengan Jaringan Saraf Tiruan (ANN) dalam hal komponen dasar seperti *input*, bobot, bias, dan fungsi aktivasi, CNN dibedakan dengan arsitekturnya yang dioptimalkan untuk memproses data dengan dimensi besar, khususnya gambar. CNN bekerja dengan cara mempelajari fitur-fitur spesifik dari data masukan. Dalam konteks CNN mengenal konsep peta fitur. Peta fitur ini dihasilkan saat sebuah filter diterapkan ke lapisan gambar sebelumnya. Istilah peta fitur berasal dari fungsinya yang memetakan lokasi spesifik di mana fitur tertentu, seperti garis atau tepi, terdeteksi pada gambar. CNN secara sistematis memindai gambar untuk mengidentifikasi berbagai fitur. Misalnya, satu peta fitur mungkin fokus pada deteksi garis vertikal, sementara peta fitur lainnya mungkin fokus pada deteksi bentuk-bentuk melengkung. Dengan demikian, CNN mampu membedah gambar menjadi fitur-fitur dasar dan menafsirkan kompleksitasnya (Elgandy, 2020).

Dalam arsitektur *Convolutional Neural Network* (CNN), ada dua fase utama: fase ekstraksi fitur dan fase klasifikasi. Fase ekstraksi fitur, yang terdiri dari lapisan

konvolusi dan *pooling*, fokus pada pengenalan dan interpretasi fitur esensial dari data. Setelah fitur tersebut berhasil diambil, informasi tersebut dikirim ke fase klasifikasi. Fase ini melibatkan lapisan *fully connected* yang bertanggung jawab atas proses pengambilan keputusan berdasarkan fitur yang telah diekstraksi. Secara keseluruhan, CNN menawarkan pendekatan yang efisien dan adaptif dalam mengelola data beresolusi tinggi, dengan kemampuannya yang khas dalam mendalami pemahaman gambar, serta memberikan solusi canggih di ranah pengenalan dan analisis pola gambar. Arsitektur dasar CNN ditunjukkan pada gambar 23.



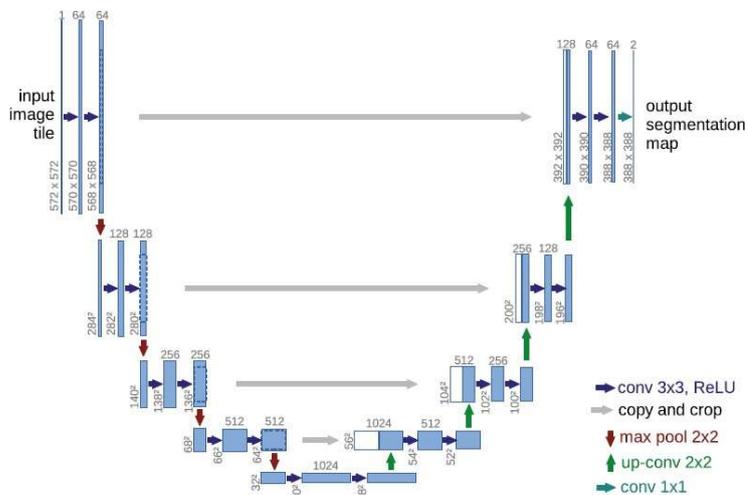
Gambar 23 Arsitektur dasar CNN ((Elgendy, 2020).

2.1.8 U-Net

U-Net adalah salah satu arsitektur jaringan saraf konvolusional yang dirancang khusus untuk segmentasi citra biomedis, namun kini telah digunakan dalam berbagai aplikasi segmentasi lainnya. Arsitektur ini mendapat namanya dari bentuknya yang menyerupai huruf "U" ketika digambarkan. Arsitektur U-Net pertama kali diperkenalkan oleh (Ronneberger, Fischer and Brox, 2015) untuk segmentasi struktur syaraf pada citra mikroskopis elektron pada *ISBI Challenge*. Arsitektur U-Net terdiri dari bagian *downsampling* (sisi kiri) dan bagian *up sampling* (sisi kanan) visualisasi mode U-Net ditunjukkan pada gambar 24. Bagian *downsampling* mengekstrak fitur hierarki dari gambar masukan, menangkap informasi tingkat rendah dan tingkat tinggi, sedangkan bagian *upsampling* mengambil fitur yang diekstraksi ini dari bagian *downsampling* dan melakukan operasi *upsampling* dan penggabungan untuk secara bertahap memulihkan resolusi spasial dan menyempurnakan keluaran segmentasi.

Downsampling path terdiri dari blok konvolusi berulang 3x3, masing-masing diikuti oleh unit linier yang diperbaiki (ReLU). Kemudian diakhiri dengan *max pool*

2x2 yang akan menjadi masukan untuk proses konvolusi selanjutnya pada blok berikutnya. Jalur ini berfungsi untuk mempelajari fitur-fitur yang terdapat pada gambar dengan merepresentasikan fitur-fitur pada berbagai tingkatan. Jalur downsampling mengurangi dimensi gambar masukan tetapi menggandakan jumlah peta fitur. Kemudian blok selanjutnya adalah jembatan yang menghubungkan jalur *downsampling* dan *upsampling*. *Bridge* terdiri dari blok konvolusi 3x3, tetapi *maxpool 2x2* tidak mengikuti bagian ini. Pada bagian ini ukuran gambarnya 32x32 dengan 1024 peta fitur. Selanjutnya jalur *upsampling* dimulai dengan proses konvolusi 2x2, namun jalur ini tidak lagi menggunakan *maxpool* seperti bagian *downsampling*. Operasi ini menggunakan empat blok konvolusi. Setiap blok konvolusi mengurangi jumlah peta fitur sebanyak dua kali lipat, sehingga jumlah peta fitur pada jalur ini akan seimbang dengan jumlah peta fitur sebelumnya. Tahap terakhir dalam arsitektur ini adalah proses lapisan konvolusi 1x1 yang diikuti dengan fungsi aktivasi sigmoid untuk memetakan setiap vektor fitur 64 komponen ke jumlah kelas yang diharapkan (Ronneberger, Fischer and Brox, 2015).



Gambar 24 Arsitektur U-Net (Ronneberger, Fischer and Brox, 2015).

U-Net belajar mengenali objek berdasarkan pola dan fitur yang dipelajari dari data pelatihan. U-Net belajar mengidentifikasi objek dengan menangkap fitur diskriminatif dan pola spasial yang menunjukkan objek berbeda pada data pelatihan berlabel. Selama proses pelatihan U-Net dilatih menggunakan gambar yang terannotasi di mana setiap gambar dipasangkan dengan masker segmentasi tingkat piksel yang sesuai. Jaringan belajar memetakan gambar masukan ke peta segmentasi dengan mengoptimalkan fungsi kerugian, yang membandingkan

keluaran segmentasi yang diprediksi dengan anotasi kebenaran dasar. Selain itu, U-Net juga terdiri dari struktur *downsampling* dan *upsampling* dengan operasi gabungan di mana bagian *downsampling* menangkap konteks dan mengekstrak fitur tingkat rendah dan tinggi dari gambar masukan, sedangkan bagian *upsampling* menghasilkan peta segmentasi dengan resolusi spasial sebagai masukan. Koneksi lewati dalam operasi gabungan antara dua bagian memungkinkan penggabungan fitur tingkat rendah dari bagian *downsampling* dengan fitur tingkat tinggi untuk menjaga informasi spasial selama proses *upsampling* koneksi lewati ini memungkinkan model menghasilkan segmentasi yang presisi.

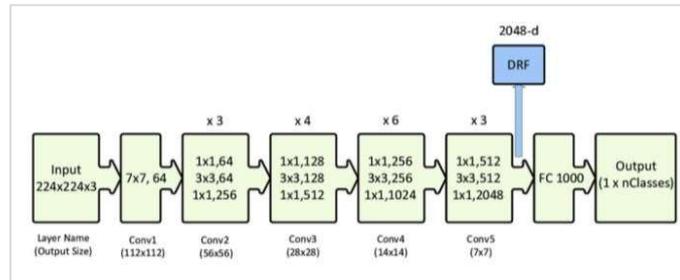
2.1.9 ResNet

Residual Neural Network atau ResNet adalah sebuah inovasi dalam dunia arsitektur *Convolutional Neural Network* (CNN) yang berasal dari tim peneliti di Microsoft Research Asia. Salah satu tantangan besar dalam melatih CNN dengan kedalaman yang signifikan adalah fenomena *vanishing gradient* dan *exploding gradient*. Dalam kondisi ini, selama proses pelatihan, gradien dapat menyusut menjadi sangat kecil atau melonjak menjadi sangat besar, menghambat kapasitas jaringan untuk memperbarui dan menyesuaikan bobotnya dengan efektif.

ResNet mengatasi masalah ini dengan memperkenalkan konsep *skip connections*. Esensi dari pendekatan ini adalah mengizinkan aktivasi dari satu lapisan untuk melewati satu atau lebih lapisan berikutnya, sebelum akhirnya dijumlahkan ke aktivasi dari lapisan *downstream*. Dengan cara ini, ResNet efektif dalam memelihara aliran informasi dan gradien melalui jaringan, memungkinkan pelatihan model dengan ratusan lapisan tanpa mengalami hambatan signifikan dalam konvergensi. Pendekatan ini telah terbukti revolusioner dan menghasilkan prestasi luar biasa dalam berbagai kompetisi pengenalan gambar, menegaskan keunggulan ResNet dalam menangani tantangan yang ditemui oleh jaringan saraf konvolusional yang dalam (He *et al.*, 2016).

ResNet memiliki beberapa varian berdasarkan jumlah lapisannya, seperti ResNet-18, ResNet-34, ResNet-50, ResNet-101 dan sebagainya. Semakin banyak lapisan pada arsitektur CNN biasanya akan memberikan kemampuan representasi fitur yang lebih kompleks namun juga dapat menimbulkan masalah seperti

overfitting pada dataset *training* yang relatif kecil. ResNet-50 merupakan salah satu varian dari ResNet dimana jumlah layernya 50 layer, visualisasi arsitektur ResNet-50 ditunjukkan pada gambar 25. Model ini cukup populer di kalangan peneliti visi komputer karena telah membuktikan diri mampu melakukan deteksi objek secara akurat serta efisien dalam waktu inferensi dibandingkan dengan menggunakan model lainnya.



Gambar 25 Arsitektur ResNet-50 (Malhotra et al., 2022)

2.1.10 Operasi Morfologi

Morfologi dalam konteks pengolahan citra berarti gambaran bentuk dan struktur objek dalam suatu citra. Operasi morfologi bekerja berdasarkan teori himpunan dan lebih mengandalkan pengurutan relatif piksel daripada nilai numeriknya. Karakteristik ini menjadikannya lebih berguna untuk pemrosesan gambar. Data masukan untuk operasi morfologi matematika adalah dua gambar: gambar mentah dan gambar primitif. Operasi morfologi didefinisikan dengan baik untuk gambar biner tetapi sama validnya dan berguna juga untuk gambar skala abu-abu (Srissha and Khan, 2013). Meskipun operasi morfologi didasarkan pada teori himpunan, banyak operasi morfologi pada dasarnya adalah operasi logis dan mudah digunakan. Operasi morfologi mendasar adalah erosi dan pelebaran. Pembahasan berikut menjelaskan beberapa operasi morfologi lainnya.

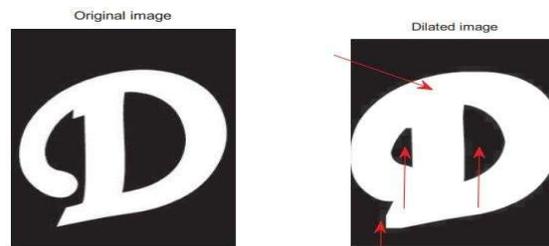
1. Dilasi

Operasi dilasi membuat suatu benda bertambah besar ukurannya. Sejauh mana pertumbuhannya bergantung pada sifat dan bentuk elemen penataan. Dilatasi suatu citra A (himpunan) oleh elemen penataan B didefinisikan sebagai:

$$A \oplus B = \{z | (\widehat{B})_z \cap A \neq \emptyset\} \quad (1)$$

Jika himpunan B dicerminkan terhadap titik asal dan digeser sebesar z, maka

dilatasi A oleh B adalah himpunan yang semua perpindahannya z sedemikian rupa sehingga B^z dan A mempunyai paling sedikit satu unsur persekutuan. Pelebaran yang dimaksudkan dengan menambahkan piksel pada elemen batas contoh operasi dilasi dapat dilihat pada gambar 26. Proses dilatasi memperbesar jumlah piksel yang bernilai satu (*foreground*) dan memperkecil jumlah piksel yang bernilai nol (*background*). Proses dilatasi juga pada dasarnya digunakan untuk mengisi lubang (piksel yang hilang) pada objek kontinu. Operasi dilatasi, karena menambahkan piksel pada batas objek, hal ini mempengaruhi intensitas di lokasi tersebut dan akibatnya efek keburaman dapat diamati (Srisha and Khan, 2013).



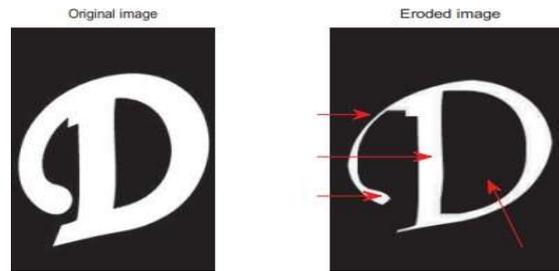
Gambar 26 Contoh Operasi Dilasi

2. Erosi

Operasi erosi merupakan pelengkap dari operasi dilatasi dalam konteks efek operasi. Artinya operasi erosi menyebabkan benda kehilangan ukurannya. Erosi suatu citra A oleh elemen penataan B didefinisikan sebagai:

$$A \ominus B = \{z | (\widehat{B})_z \subseteq A\} \quad (2)$$

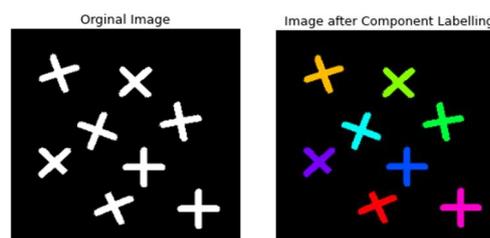
Pengikisan citra A oleh elemen penataan B merupakan himpunan semua titik z sehingga elemen penataan B diterjemahkan dengan z merupakan sub set citra. Operasi ini mengakibatkan hilangnya piksel batas objek. Proses erosi memperbesar jumlah piksel yang bernilai nol (*background*) dan memperkecil jumlah piksel yang bernilai satu (*foreground*). Operasi erosi menghilangkan struktur-struktur yang ukurannya lebih kecil dibandingkan dengan elemen struktur. Sehingga dapat digunakan untuk menghilangkan 'koneksi' yang berisik antara dua objek. Karena piksel yang tidak diinginkan 'terhapus', efek akhirnya adalah mempertajam objek dalam gambar (Srisha and Khan, 2013). Contoh operasi erosi ditunjukkan pada gambar 27.



Gambar 27 Contoh operasi Erosi

2.1.11 *Connected Component Labeling*

Connected Component Labeling juga dikenal sebagai "segmentasi objek" atau "*labeling blob*", adalah teknik yang digunakan dalam pengolahan citra dan analisis citra untuk mendeteksi dan memberi label setiap objek terhubung dalam citra biner. Sebuah objek dianggap terhubung jika terdapat sekumpulan piksel yang berdekatan satu sama lain dan memiliki nilai yang sama, biasanya piksel putih (nilai 1) pada latar belakang hitam (nilai 0). Tujuan dari operasi ini adalah untuk mengidentifikasi setiap area dalam citra biner yang terhubung dan memberi label berbeda untuk setiap area yang tidak terhubung. Operasi ini dimulai dari piksel pertama citra dan dilanjutkan sepanjang citra. Kemudian untuk setiap piksel yang ditemukan sebagai bagian dari objek (biasanya piksel dengan nilai 1), periksa piksel-piksel tetangganya. Jika piksel tetangga juga merupakan bagian dari objek dan belum diberi label, berikan label yang sama. Jika piksel tetangga adalah bagian dari objek lain, lanjutkan dengan label berikutnya.



Gambar 28 Contoh operasi *Connected Component Labeling* (Connected Component Labeling, 2020)

2.1.12 Evaluasi Sistem

Model U-Net adalah model yang digunakan untuk melakukan segmentasi khususnya pada kasus segmentasi citra medis. Arsitekturnya yang menggunakan mekanisme *encoder-decoder* yang memungkinkan ekstraksi fitur dilakukan pada

layer *encoder* lalu satukan kembali pada layer *decoder*, membuat *output* dari segmentasi ini memiliki ukuran yang sama dengan citra *input*, serta penggunaan mekanisme *skip connection* yang berada antara layer *encoder* dan *decoder* memungkinkan untuk menghasilkan segmentasi yang lebih presisi. Berikut beberapa metrik evaluasi yang akan digunakan untuk mengukur hasil percobaan pada penelitian ini:

1. Dice Loss

Dice Loss merupakan fungsi untuk menghitung kesalahan segmentasi yang dilakukan pada proses pelatihan, fungsi *dice loss* merupakan fungsi yang sering digunakan untuk mengukur sejauh mana hasil segmentasi model mendekati ground-truth atau segmen yang sebenarnya dalam citra. *Dice Loss* diperoleh dari metrik *Dice Coefficient*, yang merupakan perbandingan antara dua kali luas wilayah overlap antara hasil segmentasi dan ground-truth dengan jumlah area total dari kedua wilayah tersebut. Fungsi *Dice Loss* dapat dinyatakan sebagai berikut:

$$Dice\ loss = 1 - \frac{2X|X \cap Y|}{|X| + |Y|} \quad (3)$$

Dimana:

$|X \cap Y|$ merupakan jumlah piksel yang menjadi intersection antara prediksi dan ground-truth

$|X|$ merupakan jumlah piksel dalam prediksi

$|Y|$ merupakan jumlah piksel dalam ground-truth

Tujuan dari penggunaan *Dice Loss* adalah untuk meminimalkan perbedaan antara hasil segmentasi model dan ground-truth. Semakin kecil nilai *Dice Loss*, semakin baik hasil segmentasi model. *Dice Loss* sering digunakan dalam tugas segmentasi citra karena memiliki sensitivitas terhadap *overlap* antara segmen (Müller, Soto-Rey and Kramer, 2022).

2. Focal Loss

Focal Loss adalah fungsi kerugian yang diperkenalkan oleh (Lin *et al.*, 2017). *Focal Loss* dirancang untuk menangani masalah ketidakseimbangan

kelas dalam tugas-tugas pembelajaran mesin, terutama dalam deteksi objek dan segmentasi gambar. *Focal Loss* bertujuan untuk mengatasi masalah ketidakseimbangan dengan menyesuaikan kontribusi contoh yang mudah dan sulit dalam proses pembelajaran. Dengan memberikan lebih banyak bobot pada contoh yang sulit diklasifikasikan, *Focal Loss* membantu model untuk belajar lebih efektif dari contoh yang sulit dan meningkatkan performa keseluruhan pada dataset yang tidak seimbang.

$$FL(p_t) = -\alpha t(1 - p_t)^\gamma \log(p_t) \quad (4)$$

dimana:

p_t adalah probabilitas prediksi yang benar untuk kelas target,

αt adalah bobot kelas, digunakan untuk menangani ketidakseimbangan kelas,

γ adalah parameter *focusing* yang menentukan tingkat fokus pada kasus yang sulit.

3. F-Measure based Metrics

Berdasarkan ukuran-F, ada dua metrik yang populer digunakan di segmentasi citra medis. *Intersection-over-Union* (IoU), juga dikenal sebagai indeks *Jaccard* atau koefisien kesamaan *Jaccard*, dan *Dice Coefficient* atau Koefisien kesamaan Dice (DSC) juga indeks Sørensen-Dice (Müller, Soto-Rey and Kramer, 2022).

a. *Intersection Over Union* (IoU)

IoU adalah metrik yang sering digunakan dalam segmentasi objek untuk mengukur kesesuaian antara area prediksi dengan area ground truth. Ini dihitung dengan membagi area irisan (*intersection*) dari prediksi dan *ground truth* dengan area gabungan (*union*) dari kedua tersebut. Nilai IoU berkisar antara 0 hingga 1, di mana nilai yang mendekati 1 menandakan prediksi yang sangat akurat dan nilai mendekati 0 menandakan prediksi yang kurang akurat (Müller, Soto-Rey and Kramer, 2022).

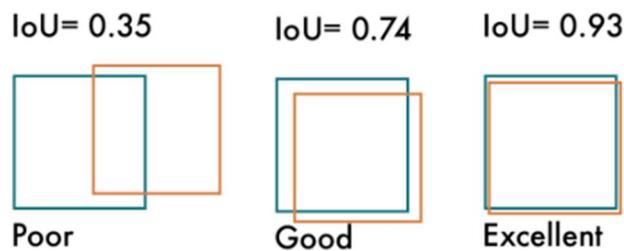
$$IoU = \frac{\text{Area of Intersection}}{\text{Area of Union}} \quad (5)$$

dimana :

Area of intersection: mengacu pada luas wilayah yang tumpang tindih antara dua set atau wilayah, seperti wilayah prediksi dan

wilayah *ground truth*.

Area of union: mengacu pada luas gabungan dari dua set atau wilayah tersebut, yaitu wilayah prediksi dan wilayah *ground truth*. Jadi, *area of union* adalah total luas wilayah yang dicakup oleh prediksi dan *ground truth*.



Gambar 29 Penilaian performa segmentasi berdasarkan nilai IoU

Gambar 29 menunjukkan penilaian performa segmentasi berdasarkan tiga nilai IoU yang berbeda. Nilai IoU sebesar 0,35 menunjukkan sangat sedikitnya area yang tumpang (sesuai) yang terjadi antara prediksi (kotak berwarna biru) dan *ground truth* (kotak berwarna coklat) yang menyatakan bahwa segmentasi yang dihasilkan tidak akurat karena tidak sesuai dengan *ground truth*. Nilai IoU = 0,74 menunjukkan area tumpang tindih yang lebih banyak antara prediksi dan *ground truth* yang mengindikasikan performa yang baik dalam segmentasi objek. Nilai IoU = 0,93 menunjukkan sebagian besar area tumpang tindih yang terdapat antara prediksi dan *ground truth* yang mengindikasikan performa yang sangat baik (unggul) dalam segmentasi objek (Terven *et al.*, 2023).

b. *Dice Coefficient*

Dice Coefficient juga dikenal sebagai *Sørensen–Dice index* atau *Dice's similarity coefficient* (DSC), adalah sebuah metrik yang digunakan untuk mengukur kesamaan antara dua sampel. Dalam konteks visi komputer, khususnya pada masalah segmentasi gambar, *Dice Coefficient* digunakan untuk mengukur seberapa baik piksel yang diprediksi oleh model segmentasi cocok dengan piksel yang sebenarnya (*ground truth*). Rumus untuk menghitung *dice coefficient* adalah sebagai berikut:

$$\text{Dice Coefficient} = \frac{2|X \cap Y|}{|X| + |Y|} \quad (6)$$

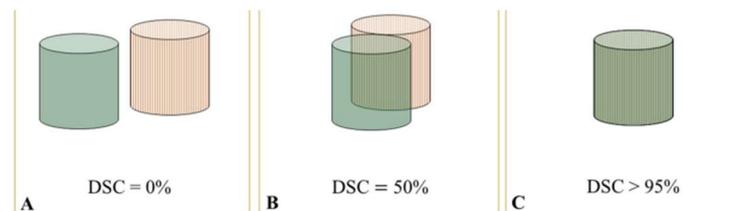
dimana:

$|X \cap Y|$ merupakan jumlah piksel yang menjadi intersection antara prediksi dan *ground truth*

$|X|$ merupakan jumlah piksel dalam prediksi

$|Y|$ merupakan jumlah piksel dalam *ground truth*

Dice Coefficient menghitung dua kali area *overlap* antara prediksi dan *ground truth*, dibagi dengan total jumlah piksel di prediksi ditambah jumlah piksel di *ground truth*. Nilai yang dihasilkan berkisar antara 0 sampai 1, di mana nilai yang lebih tinggi menunjukkan tingkat kesamaan yang lebih besar antara prediksi dan *ground truth* (Müller, Soto-Rey and Kramer, 2022).



Gambar 30 Penilaian performa segmentasi berdasarkan nilai *Dice Coefficient* (Barat *et al.*, 2021)

Gambar 30 menunjukkan penilaian kinerja model segmentasi berdasarkan nilai *dice coefficient*. Nilai *dice coefficient* = 0% menunjukkan bahwa tidak terdapat kemiripan antara prediksi dan *ground truth* yang mengindikasikan hasil segmentasi yang tidak akurat. Nilai *dice coefficient* = 50% menunjukkan bahwa setengah dari wilayah yang segmentasi mirip dengan *ground truth* yang mengindikasikan bahwa segmentasi yang dihasilkan lebih baik. Sementara nilai *dice coefficient* > 95% menunjukkan bahwa sebagian besar area prediksi sesuai dengan area *ground truth* dimana nilai > 95% mengindikasikan segmentasi yang semakin akurat (Barat *et al.*, 2021).

c. Recall (5)

Recall merupakan metrik evaluasi yang digunakan untuk mengukur kemampuan model dalam mendeteksi semua sampel positif yang relevan dalam dataset. *Recall* sering digunakan dalam konteks tugas klasifikasi,

terutama dalam situasi di mana penting untuk menemukan sebanyak mungkin sampel positif. *Recall* didefinisikan sebagai rasio antara jumlah *True Positives* (TP) dan jumlah total sampel positif yang sebenarnya (*True Positives + False Negatives*), dengan formula sebagai berikut:

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (7)$$

di mana:

True Positives (TP): Jumlah sampel positif yang diklasifikasikan dengan benar oleh model.

False Negatives (FN): Jumlah sampel positif yang salah diklasifikasikan sebagai negatif oleh model.

Nilai *Recall* yang tinggi menunjukkan bahwa model berhasil mendeteksi sebagian besar sampel positif yang ada dalam dataset, sedangkan nilai *recall* yang rendah menunjukkan bahwa model melewatkan banyak sampel positif, yang berarti banyak sampel positif yang salah diklasifikasikan sebagai negatif.

4. *Mean Absolute Error* (MAE)

Mean Absolute Error (MAE) adalah metrik evaluasi yang digunakan untuk mengukur kesalahan rata-rata absolut antara nilai yang diprediksi oleh model dan nilai aktual. MAE digunakan dalam tugas-tugas regresi untuk mengevaluasi seberapa dekat prediksi model dengan nilai sebenarnya. MAE didefinisikan sebagai rata-rata dari nilai absolut perbedaan antara nilai yang diprediksi (\hat{y}_i) dan nilai aktual (y_i):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (8)$$

di mana:

n adalah jumlah total observasi,

y_i adalah nilai aktual untuk observasi ke- i ,

\hat{y}_i adalah nilai prediksi untuk observasi ke- i .

Nilai MAE di interpretasi dengan, jika nilai $MAE = 0$ menunjukkan bahwa prediksi model sempurna, tanpa kesalahan. Jika nilai MAE tinggi menunjukkan bahwa ada kesalahan besar antara prediksi model dan nilai aktual, yang berarti model memiliki kinerja yang buruk dalam memprediksi nilai yang benar. Sedangkan jika nilai MAE lebih rendah menunjukkan bahwa model dekat dengan nilai aktual, yang berarti model memiliki kinerja yang baik.

2.2 Metode Penyelesaian

2.2.1 State of The Art

Tabel 1 State of The Art

No.	Judul Karya Ilmiah, Nama, Tahun Terbit dan Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja
1.	<p>Judul: Computer Aided Diagnosis of Breast Cancer based on Fine Needle Biopsy Microscopic Images</p> <p>Penulis: Kowal et al(Kowal <i>et al.</i>, 2013)</p> <p>Tahun: 2013</p> <p>Penerbit: Elsavier</p>	<p>Objek: Sel pada citra sitologi</p> <p>Permasalahan: Bagaimana melakukan segmentasi area sel kanker pada citra sitologi</p>	Adaptive thresholding, FCM, K-Means dan GMM	Gambar yang diproses setelah segmentasi menunjukkan bahwa metode yang diuji tidak dapat memisahkan sel yang berjarak dekat.
2.	<p>Judul: Nuclei segmentation of microscopic breast cancer image using Gram-</p>	<p>Objek: Sel kanker payudara</p>	Gram-schmidt, Bayesian-based clustering algorithm	Akurasi: 73%. Metode berbasis <i>clustering</i> yang diterapkan menunjukkan bahwa metode tersebut mampu membedakan antara

	<p>Schmidt and cluster validation algorithm</p> <p>Penulis: Fatichah et al (Fatichah <i>et al.</i>, 2015)</p> <p>Tahun: 2015</p> <p>Penerbit: IEEE</p>	<p>Permasalahan: Segmentasi sel kanker payudara dan pemisahan sel bertumpuk</p>		<p>objek dan latar belakang dengan baik, namun pada segmentasi antar objek belum mampu memisahkan sel yang saling berdekatan atau tumpang tindih, dan terdapat <i>under segmentation</i></p>
3.	<p>Judul: Pemisahan Sel Bertumpuk Citra Sel Kanker Payudara Menggunakan Metode Region-Based Active Contour dan Bayesian</p> <p>Penulis: Aini et al</p> <p>Tahun: 2015</p> <p>Penerbit: Jurnal Teknologi Informasi dan Komunikasi (JTIK)</p>	<p>Objek: Sel kanker payudara</p> <p>Permasalahan: Segmentasi sel dan pemisahan sel bertumpuk</p>	<p>Region based active contour dan bayesian clustering</p>	<p>Akurasi: 61,3%</p> <p>Hasil penelitian ini menunjukkan metode yang digunakan belum mampu melakukan pemisahan sel yang berdempet dan tumpang tindih. Pada hasil pemisahan cenderung terjadi <i>over segmentation</i> dimana sel yang seharusnya merupakan Tunggal, namun di definisikan sebagai sel tumpang tindih.</p>
4.	<p>Judul: Spatial Fuzzy C-Means dan Rapid Region Merging untuk Pemisahan Sel Kanker Payudara</p> <p>Penulis:</p>	<p>Objek: Sel kanker payudara pada citra histopatologi</p>	<p>Gram-schmidt, Fuzzy C-Mean, Rapid Region Merging</p>	<p>Akurasi: 78.41%</p> <p>Hasil penelitian ini menunjukkan bahwa metode yang diterapkan mampu mereduksi <i>over segmentation</i> pada segmentasi sel, namun metode ini menimbulkan banyak <i>under-</i></p>

	<p>Tuwongihide D et al Tahun: 2017 Penerbit: JNTETI</p>	<p>Permasalahan: Segmentasi sel bertumpuk</p>		<p><i>segmentation</i> hal ini disebabkan oleh kesalahan <i>clustering</i> sel sehingga sel yang sebenarnya tumpang tindih atau berdempet, dinyatakan sebagai sel Tunggal sehingga tidak dilakukan proses pemisahan sel.</p>
5.	<p>Judul: Nuclei Segmentation in Histopathology Images Using Deep Neural Networks Penulis: Naylor P et al (Naylor <i>et al.</i>, 2017b) Tahun: 2017 Penerbit: IEEE</p>	<p>Objek: Sel kanker payudara Permasalahan: Segmentasi sel dan segmentasi sel tumpang tindih pada citra histopatologi</p>	<p>PangNet, DeconvNet, FCN untuk segmentasi sel dan Watershed untuk segmentasi sel tumpang tindih</p>	<p>PangNet: IoU: 72.2% DeconvNet: IoU: 81,4% FCN IU: 78,2% Hasil penelitian ini menunjukkan bahwa DeconvNet dan watershed memiliki kinerja yang terbaik dalam segmentasi sel pada penelitian ini. Namun, terjadi banyak masalah <i>over segmentation</i> disebabkan oleh algoritma watershed yang belum mampu melakukan pemisahan sel berdempet dan tumpang tindih terhadap sel khususnya pada sel yang memiliki ‘lubang’ di dalam sel yang menyebabkan sel tunggal disegmentasi sebagai sel yang tumpang tindih.</p>
6.	<p>Judul: Nuclei Segmentation in Histopathological Images Using Two-Stage Learning</p>	<p>Objek: Sel kanker Permasalahan: Segmentasi sel Tunggal</p>	<p>Two-stage UNet, Deep Learning Aggregation (DLA)</p>	<p>F1: 82,9% IoU/AJI: 0,611 Penelitian ini melakukan tahap pelatihan dua kali melakukan U-Net bersusun.</p>

	<p>Penulis: Kang Q et al(Kang, Lao and Fevens, no date)</p> <p>Tahun: 2019</p> <p>Penerbit: MICCAI 2019</p>	dan segmentasi sel tumpang tindih		<p>Pelatihan yang pertama digunakan untuk memprediksikan batas antar sel sementara pelatihan kedua digunakan untuk memisahkan sel berdempet dan tumpang tindih. Namun, metode ini memerlukan komputasi yang lebih besar sebab metode ini menggunakan dua kali proses pelatihan dengan U-Net bersusun untuk menghasilkan batas antar objek kemudian menghasilkan segmentasi akhir.</p>
7.	<p>Judul: A Two-Stage U-Net Algorithm for Segmentation of Nuclei in H&E Stained Tissues</p> <p>Penulis: Mahbod A et al(Mahbod <i>et al.</i>, 2019)</p> <p>Tahun: 2019</p> <p>Penerbit: Springer</p>	<p>Objek: Sel</p> <p>Permasalahan: Segmentasi sel dan segmentasi sel bertumpuk</p>	Two-stage Unet, gaussian smoothing, watershed algorithm	<p>Dice Score: 79,32%</p> <p>IoU: 56,87%</p> <p>Penelitian ini menambahkan tahapan modifikasi mask dimana pada prosesnya menggunakan dua mask yaitu mask semantic dan mask instance. Mask instance dikonversi menjadi biner kemudian dilakukan erosi untuk memperkecil bentuk sel sehingga hal ini memperbesar jarak antar kedua sel (untuk kasus yang berdempet dan tumpang tindih) kemudian proses pelatihan di lakukan dengan memanfaatkan dua mask tersebut. Hasil penelitian menunjukkan bahwa metode ini berkerja dengan baik untuk objek yang berdekatan tapi tidak</p>

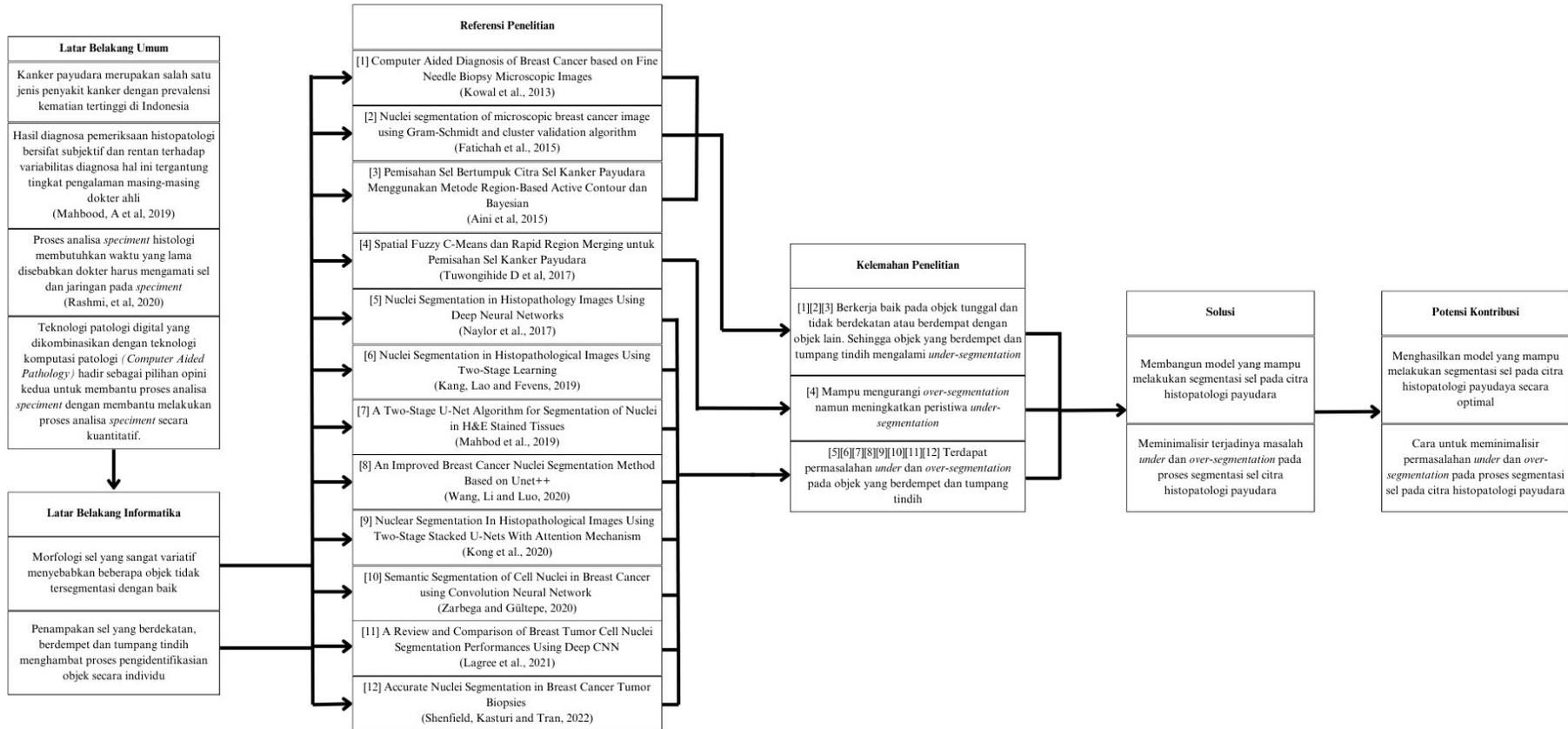
				berdempet. Sementara pada objek yang relatif lebih padat, metode ini banyak mengalami <i>over segmentation</i> yang disebabkan oleh model yang masih belum optimal dalam mengenali batas-batas objek.
8.	<p>Judul : An Improved Breast Cancer Nuclei Segmentation Method Based on Unet++</p> <p>Penulis: Wang H et al(Wang, Li and Luo, 2020)</p> <p>Tahun: 2020</p> <p>Penerbit: ICCAI 2020</p>	<p>Objek: Sel kanker payudara</p> <p>Permasalahan: Segmentasi area sel</p>	UNet++ dan Inception-ResNetV2 (as backbone)	<p>Acc: 95,05%</p> <p>Precision: 55,81%</p> <p>Recall: 60,35%</p> <p>Dice Coefficient: 52,07%</p> <p>Penelitian ini menggunakan U-Net++ dengan memanfaatkan ResNetV2 sebagai backbone untuk memperdalam ekstraksi fitur pada encoder. Namun dari hasil yang ditunjukkan, penelitian ini mengalami banyak <i>over-segmentation</i> dimana objek yang terdeteksi lebih banyak dari pada objek pada mask.</p>
9.	<p>Judul: Nuclear Segmentation In Histopathological Images Using Two-Stage Stacked U-Nets With Attention Mechanism</p> <p>Penulis: Kong Y et al(Kong <i>et al.</i>, 2020)</p> <p>Tahun:</p>	<p>Objek: Sel kanker payudara</p> <p>Permasalahan: Segmentasi sel bertumpuk</p>	Two-stage stacked UNet, Attention Generation Model (AGM), Focal loss	<p>Precision: 79,2%</p> <p>Recall: 85,3%</p> <p>F1: 82,47%</p> <p>AJI: 62,1%</p> <p>Penelitian ini menggunakan dua kali tahap pelatihan dimana pelatihan tahap pertama dilakukan dengan menggunakan berbagai jenis pembesaran dan menggunakan empat U-Net bertumpuk dengan yang masing-masingnya</p>

	2020 Penerbit: Frontiers			melakukan pelatihan pada resolusi data yang berbeda kemudian hasilnya akumulasikan dan dijadikan <i>input</i> untuk pelatihan tahap kedua menggunakan U-Net orisinal untuk memprediksi hasil segmentasi akhir. Metode yang digunakan sudah mampu mengidentifikasi sel yang berdempet namun belum maksimal dalam sel tumpang tindih disebabkan oleh data sel yang tumpang tindih lebih sedikit dibandingkan sel tunggal atau sel berdempet sehingga model belum optimal dalam mengenali kondisi sel tumpang tindih
10.	Judul: Semantic Segmentation of Cell Nuclei in Breast Cancer using Convolution Neural Network Penulis: Zarbega et al(Zarbega and Gültepe, 2020) Tahun: 2020 Penerbit: International Journal of Computer Application	Objek: Sel kanker payudara Permasalahan: Segmentasi sel berdempet dan tumpang tindih	CNN dan Binarization segmentation	Precision: 84,4% Recall: 83,2% Acc: 85,1% Penelitian ini melakukan pelatihan dengan menggunakan model CNN yang kemudian diikuti dengan menerapkan operasi morfologi pada tahap <i>post processing</i> untuk memisahkan sel yang berdempet dan tumpang tindih. Hasil penelitian ini menunjukkan bahwa model telah mampu mengenali objek sel dengan baik namun, terdapat banyak kondisi <i>over segmentation</i> dari hasil

				<p>pelatihan dimana piksel yang sebenarnya bukan objek, tapi di kelompokkan menjadi objek sehingga objek yang diprediksi jumlahnya melebihi jumlah objek yang sebenarnya. Selain itu, penggunaan operasi morfologi pada penelitian ini hanya menghilangkan piksel-piksel kecil hasil segmentasi pelatihan namun tidak memisahkan antara objek yang berdempet.</p>
11.	<p>Judul: A Review and Comparison of Breast Tumor Cell Nuclei Segmentation Performances Using Deep CNN</p> <p>Penulis: Lagree A et al (Lagree <i>et al.</i>, 2021)</p> <p>Tahun: 2019</p> <p>Penerbit: Nature</p>	<p>Objek: Sel kanker payudara</p> <p>Permasalahan: Segmentasi sel berdempet dan segmentasi sel tumpang tindih</p>	<p>GB-Unet Multiple U-Net (vgg19, denseNet121 dan ResNet101 backbone)</p>	<p>IoU/AJI: 0.5403 dan 0.53</p> <p>Metode ini menggunakan metode <i>gradient boosting network</i> yang menggunakan tiga U-Net bersusun dengan masing-masing U-Net memiliki layer encoder yang berbeda (vgg19, densenet121 dan resnet101) selanjutnya untuk mendapatkan segmentasi akhir peneliti menggunakan satu lagi U-Net tanpa mengubah encodernya. Hasil penelitian ini memperlihatkan nilai IoU yang masih rendah, dimana model mampu memisahkan objek dari latar belakang dan mampu mengenali dengan baik sel tunggalnya, namun pada citra yang terdiri dari banyak objek belum mampu memisahkan sel yang berdempet dan tumpang tindih sehingga</p>

				menyebabkan terjadinya kondisi <i>under-segmentation</i> . Selain itu, metode ini juga membutuhkan sumber daya komputasi yang sangat besar dan waktu yang lama dalam pelatihan karena menggunakan tiga susun U-Net dengan masing-masing encoder yang berbeda dan berbeda tingkat kedalaman mulai dari VGG, Densenet dan Resnet. Setelah itu, penggunaan U-Net terakhir untuk mendapat segmentasi akhir.
12.	<p>Judul: Accurate Nuclei Segmentation in Breast Cancer Tumor Biopsies</p> <p>Penulis: Kasturi S, et al (Shenfield, Kasturi and Tran, 2022)</p> <p>Tahun: 2022</p> <p>Penerbit: IEEE</p>	<p>Objek: Sel kanker Payudara</p> <p>Permasalahan: Segmentasi sel berdempet dan tumpang tindih</p>	Perbandingan model encoder Resnet50, Resnet101, Resnext50-4d, Resnest101-8d, DenseNet121 dan DenseNet201	<p>IoU/AJI: 0.669</p> <p>Penelitian ini membandingkan beberapa model dengan mengganti masing-masing layer encoder dari model U-Net asli. Penulis juga membandingkan kinerja <i>pre-train network</i> dengan model yang tidak menggunakan <i>pre-trained</i>. Berdasarkan hasil penelitian, model <i>pre-trained</i> dengan encoder DenseNet-121 dengan nilai IoU/AJI 0.669 dibandingkan semua model yang digunakan. Model telah berhasil mengenali objek yang berdekatan namun belum mampu melakukan pemisahan sel yang berdempet dan saling tumpang tindih.</p>

2.2.2 Kerangka Pikir



Gambar 31 Kerangka Pikir