BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

Mengidentifikasi studi dan penelitian terdahulu dalam pengembangan model deteksi metastasis kanker prostat berbasis metode *transformer* dan citra MRI merupakan langkah penting untuk menghindari duplikasi penelitian. Langkah ini perlu dilakukan untuk memastikan bahwa penelitian yang dilakukan memberikan kontribusi di bidangnya. Dengan melakukan kajian terhadap studi sebelumnya, temuan-temuan dari penelitian terdahulu dapat digunakan sebagai referensi penting dalam pengembangan model diagnosis yang lebih baik dan efisien. Pada tabel 2.1 ditampilkan rangkuman dari penelitian sebelumnya yang dicari menggunakan keyword "*prostate cancer metastasis* + *transformer*", "*image processing* + mri + *transformer*" yang berkaitan dengan penelitian ini.

Tabel 2.1 Penelitian Terdahulu

No	**Artikel Penelitian**	
1.	Judul	*Whole Slide Image* Analysis and Detection of Prostate Cancer using *Vision Transformer*s [13]
	Penulis	Kobiljon Ikromjanov, Rashadul Islam Sumon, Subrata Bhattacharjee, Hee-Cheol Kim, Yeong-Byn Hwang, Heung-Kook Choi
	Tahun	2022
	Nama Jurnal	ICAIIC Conference Proceedings
	Algoritma yang Digunakan	*Vision Transformer*s (ViT)
	Dataset yang Digunakan	10.616 gambar histopatologis *Whole Slide Image* (Radboud University Medical Center dan Karolinska Institute, PANDA Challenge).
	Hasil Penelitian	Mencapai rata-rata akurasi 80.4% dengan kinerja yang tinggi dalam klasifikasi benign dan stroma.
	Saran/Limitasi Penelitian	Penelitian masa depan yang akan dilakukan peneliti ini melibatkan penggunaan arsitektur Mask-RCNN dan peningkatan ukuran *dataset*.

2.	Judul	Research on Prostate Cancer Pathological Image Classification Method Based on *Vision Transformer* [18]
	Penulis	Zhixian Tang, Xinrong Yang, Yuhe Wei, Ling Wei, Wanran Zhang, Xue Wang
	Tahun	2024
	Nama Jurnal	International Conference on Image, Vision and Intelligent Systems (ICIVIS)
	Algoritma yang Digunakan	*Vision Transformer*s (ViT); Teknik pra-pemrosesan (crop, normalisasi) dan augmentasi data
	Dataset yang Digunakan	*Dataset* gambar patologis prostat dengan penyeimbangan melalui augmentasi
	Hasil Penelitian	Akurasi klasifikasi 76.7%, lebih baik dibandingkan model klasik
	Saran/Limitasi Penelitian	Perlu peningkatan interpretabilitas model dan *dataset* yang lebih besar
3.	Judul	Detecting Adverse Pathology of Prostate Cancer With a Deep Learning Approach Based on a 3D Swin-*Transformer* Model and Biparametric MRI: A Multicenter Retrospective Study [19]
	Penulis	Litao Zhao, Jie Bao, Ximing Wang, Xiaomeng Qiao, et al.
	Tahun	2024
	Nama Jurnal	Journal of *Magnetic Resonance Imaging*
	Algoritma yang Digunakan	3D Swin-*Transformer*; model TransNet dan TransCL
	Dataset yang Digunakan	616 pasien dari 6 institusi, data MRI biparametrik
	Hasil Penelitian	AUC model TransCL yang merupakan gabungan dari model Swin-*Transformer* network dan clinical characteristics mencapai akurasi 81.3%
	Saran/Limitasi Penelitian	Penelitian mendatang untuk meningkatkan interpretabilitas model dan *dataset* yang lebih besar

4.	Judul	On the Effectiveness of 3D *Vision Transformer*s for the Prediction of Prostate Cancer Aggressiveness [20]
	Penulis	Eva Pachetti, Sara Colantonio, Maria Antonietta Pascali
	Tahun	2022
	Nama Jurnal	ICIAP 2022
	Algoritma yang Digunakan	*Vision Transformer*s 3D, CNN untuk perbandingan
	Dataset yang Digunakan	*Dataset* ProstateX-2
	Hasil Penelitian	*Vision Transformer*s menunjukkan kinerja lebih baik dibandingkan CNN dengan optimasi minimal
	Saran/Limitasi Penelitian	Studi lebih lanjut diperlukan untuk optimasi lanjutan dan pengujian pada *dataset* yang lebih besar
5.	Judul	*Transformer*s Improve Breast Cancer Diagnosis from Unregistered Multi-View Mammograms [21]
	Penulis	Xuxin Chen, Ke Zhang, Neman Abdoli, et al.
	Tahun	2022
	Nama Jurnal	Diagnostics
	Algoritma yang Digunakan	DeiT (*Data-efficient Image Transformer*s), CNN-*based* multi-view models
	Dataset yang Digunakan	949 set mammogram (470 malignant, 479 benign)
	Hasil Penelitian	Two-view DeiT-tiny mencapai akurasi 77.0%, Two-view DeiT-small 76.3%, mengungguli CNN dengan feature concatenation 73.9% dan View-wise CNN model 71.7%
	Saran/Limitasi Penelitian	Penelitian masa depan untuk meningkatkan interpretabilitas *transformer* dan optimasi arsitektur untuk gambar medis multi-view
6.	Judul	Efficient Swin *Transformer* Model for Accurate Detection of Lung Cancer [22]
	Penulis	Ishak Pacal

	Tahun	2024
	Nama Jurnal	4th International Artificial Intelligence and Data Science Congress
	Algoritma yang Digunakan	Swin *Transformer* (optimasi dengan depth adjustment dan window size)
	Dataset yang Digunakan	IQ-OTH/NCCD CT *dataset*
	Hasil Penelitian	Mencapai akurasi 97,58% dan skor *F1* 95,42% dalam deteksi kanker paru-paru
	Saran/Limitasi Penelitian	Penelitian di masa depan dapat menyelidiki kinerja model pada set data yang beragam dan jenis kanker lainnya
7.	Judul	Vision *transformer*s for glioma classification using T1 magnetic resonance imaging [23]
	Penulis	Wickramasinghe et al.
	Tahun	2024
	Nama Jurnal	Artificial Intelligence in Health
	Algoritma yang Digunakan	*Vision Transformer* (ViT)
	Dataset yang Digunakan	*Dataset* BraTS 2015 yang mencakup gambar MRI pasien dengan glioma derajat tinggi (HGG) dan rendah (LGG), serta gambar tanpa tumor.
	Hasil Penelitian	Akurasi klasifikasi untuk dua kelas (tumor vs. non-tumor) mencapai 81,8%, sementara untuk tiga kelas (HGG, LGG, non-tumor) lebih rendah (63,2%) karena ketidakseimbangan kelas.
	Saran/Limitasi Penelitian	Penelitian ini terbatas oleh *dataset* kecil dan tidak seimbang, kebutuhan komputasi tinggi, serta rendahnya performa klasifikasi multikelas, sehingga disarankan untuk menggunakan *dataset* yang lebih besar dan seimbang, mengoptimalkan *preprocessing*, mengadopsi pendekatan hybrid, serta mengeksplorasi teknik explainable AI dan pengujian pada *dataset* terbaru.

8.	Judul	Brain Tumor Classification from MRI Using Image Enhancement and Convolutional Neural Network Techniques [15]
	Penulis	Rasheed et al.
	Tahun	2023
	Nama Jurnal	Brain Science
	Algoritma yang Digunakan	VGG16, ResNet50, VGG19, InceptionV3, MobileNetV2 dengan optimization Adam *optimizer*, ReduceLROnPlateau callback. Menggunakan teknik *image processing Gaussian-blur-based sharpening filter* dan CLAHE
	Dataset yang Digunakan	Gabungan 3 *dataset* publik dari Kaggle, Figshare, SARTAJ, BR35H, dengan 7023 gambar MRI grayscale
	Hasil Penelitian	Akurasi CNN dengan *image processing* gaussian + CLAHE: 97.84%, VGG16: 95%, ResNet50: 94.75%, VGG19: 94.83%, InceptionV3: 88.15%, MobileNetV2: 91.73%
	Saran/Limitasi Penelitian	Penelitian ini memiliki limitasi terkait ketergantungan pada *dataset* MRI yang spesifik, potensi pengaruh negatif teknik *image processing* terhadap akurasi klasifikasi, serta belum diuji secara menyeluruh pada berbagai jenis modalitas pencitraan dan skenario klinis yang berbeda.
9.	Judul	Diagnostic efficiency of multi-modal MRI *based deep learning* with *Sobel operator* in differentiating benign and malignant breast mass lesions—a retrospective study [16]
	Penulis	Tang W et al.
	Tahun	2023
	Nama Jurnal	PeerJ Computer Science
	Algoritma yang Digunakan	*Convolutional Neural Networks* (CNN), dengan layer yang sudah dikostumisasi dan menggunakan teknik *operator Sobel*.
	Dataset yang Digunakan	Data retrospektif dari 203 pasien (dengan total 207 lesi) yang menjalani MRI payudara.

	Hasil Penelitian	1. Model *deep learning* berbasis single-modal memiliki AUC > 0.75, menunjukkan performa yang baik. 2. Model single-modal terbaik (DCE-MRI fase post-kontras pertama) memiliki AUC sebesar 0.827 dengan akurasi 78.5%. 3. Model multi-modal dengan *operator Sobel* mencapai AUC 0.887, sensitivitas 79.8%, spesifisitas 86.1%, dan akurasi 82.9%, secara signifikan lebih baik daripada model single-modal. 4. *Operator Sobel* berhasil meningkatkan informasi tepi lesi, mengurangi redundansi, dan meningkatkan akurasi klasifikasi.
	Saran/Limitasi Penelitian	Penelitian ini memiliki keterbatasan berupa ukuran sampel kecil, tanpa validasi eksternal, dan hanya mengevaluasi lesi massa payudara, sehingga diperlukan studi lebih lanjut dengan ukuran sampel besar dan validasi multi-pusat untuk meningkatkan generalisasi model.
10.	Judul	Computer-aided diagnostic network for brain tumor classification employing modulated *Gabor filter* banks [17]
	Penulis	Singh R et al.
	Tahun	2021
	Nama Jurnal	The Visual Computer
	Algoritma yang Digunakan	Arsitektur CNN yang dikostumisasi, dengan *filter Gabor* yang digunakan untuk memodulasi *filter* convolutional dalam CNN.
	Dataset yang Digunakan	*dataset* BRaTS 2017, yang terdiri dari gambar MRI dari 210 pasien HGG dan 75 pasien LGG.
	Hasil Penelitian	Arsitektur yang diusulkan mencapai akurasi 98,68%, lebih tinggi dibandingkan metode *deep learning* lainnya, seperti Alex-Net (84,2%), VGG 19 (86,8%), dan Res-Net 34 (95,6%). Sensitivitas mencapai 99,39%, dan spesifisitas mencapai 96,77%.
	Saran/Limitasi Penelitian	Penelitian ini terbatas pada klasifikasi dua kelas tumor otak (HGG dan LGG) dengan *dataset* BRaTS 2017, sehingga disarankan untuk memperluas

		cakupan *dataset*, mengoptimalkan model untuk aplikasi klinis nyata, serta mengembangkan kemampuan deteksi multiklasifikasi dan integrasi waktu nyata.

(Sumber olahan peneliti, 2024)

Berbagai penelitian terbaru telah mengeksplorasi penggunaan model *transformer* untuk deteksi berbagai jenis kanker. *Vision Transformers* (ViT) telah diterapkan untuk mendeteksi dan menganalisis kanker prostat melalui gambar slide utuh (*Whole Slide Image*/WSI) [13]. Kanker prostat adalah salah satu penyakit yang paling sering terjadi pada pria, dan biasanya didiagnosis dengan sistem penilaian Gleason yang melihat struktur tumor dalam gambar jaringan tubuh. Namun, cara tradisional dalam mendiagnosis kanker ini bisa memakan waktu dan seringkali hasilnya berbeda-beda antar ahli patologi. Untuk mengatasi masalah ini, penelitian ini menggunakan metode otomatis berbasis kecerdasan buatan (AI) dan pembelajaran mendalam (*deep learning*) melalui model ViT untuk mengklasifikasikan gambar biopsi prostat sesuai dengan skor Gleason. *Dataset* penelitian ini berasal dari *dataset* publik Kaggle PANDA, yang mencakup lebih dari 10.016 citra WSI dan beserta mask-nya. Penelitian ini menggunakan teknik patching untuk memecah citra besar ini menjadi patch yang lebih kecil dan dapat dikelola, untuk dijadikan input ke dalam model ViT untuk klasifikasi. Model ini menghasilkan skor Gleason berdasarkan pola sel kanker dalam citra, membantu membedakan antara jaringan jinak dan jaringan kanker dengan berbagai tingkatan (skor 3, 4, 5). Selanjutnya, kinerja model ViT dievaluasi menggunakan metrik akurasi presisi, *recall*, dan *F1-score*, dan menghasilkan akurasi 80.4% dengan performa tinggi dalam klasifikasi benign dan stroma. Hasil penelitian ini menunjukkan bahwa model tersebut efektif dalam mengidentifikasi dan memberi penilaian tingkat keparahan kanker prostat, serta memiliki potensi untuk membantu ahli patologi dalam membuat diagnosis yang lebih cepat dan akurat. Pendekatan otomatis ini dapat secara signifikan mengurangi beban kerja dan variabilitas dalam diagnosis kanker prostat, serta menyoroti sinergi antara sistem AI dan keahlian manusia. Pengembangan lebih lanjut yang disarankan penelitian ini mencakup

peningkatan *dataset* dan penggunaan model yang lebih canggih seperti Mask-RCNN untuk meningkatkan akurasi.

Model *Vision Transformers* (ViT) juga telah digunakan untuk mengklasifikasikan gambar patologis dalam diagnosis kanker prostat [18]. Metode yang digunakan melibatkan teknik prapemrosesan data seperti pemotongan gambar dan normalisasi grayscale, serta augmentasi data untuk menangani ketidakseimbangan *dataset*. Model ViT dilatih dengan gambar patologis yang telah dinormalisasi dan diuji dengan data yang baru, dan menghasilkan akurasi klasifikasi sebesar 76.7%, yang melebihi model *deep learning* klasik atau tradisional. Hasil ini menunjukkan bahwa model ViT memiliki potensi untuk meningkatkan akurasi dan efisiensi diagnosis kanker prostat, sehingga membantu meningkatkan manajemen klinis pasien. Penelitian ini menyarankan untuk meningkatkan interpretabilitas model dan menggunakan *dataset* yang lebih besar pada penelitian kedepannya.

Pengembangan model *deep learning* dengan menggunakan *Swin Transformer* juga telah digunakan pada deteksi *adverse pathology* (AP) pada pasien kanker prostat [19]. Berdasarkan penelitian ini, metode penilaian sebelumnya yang dilakukan oleh radiolog dengan menggunakan karakteristik klinis menunjukkan kinerja yang buruk dalam mendeteksi keberadaan AP. Maka dari itu, penelitian ini mengusulkan penggunaan metode *deep learning* untuk meningkatkan akurasi dalam pendeteksian AP. Model *deep learning* yang digunakan adalah TransNet, yang merupakan model berbasis jaringan Swin-*Transformer*. Selain itu, model TransCL juga digunakan, dimana model ini menggabungkan model TransNet dengan karakteristik klinis, seperti skor Gleason, skor Prostate Imaging Reporting and Data System (PI-RADS), kadar antigen spesifik prostat, nilai ADC, dan beberapa skor medis lainnya. Penelitian ini menggunakan *dataset* biparametrik MRI dengan teknik pencitraan T2W dari 616 pria yang menjalani prostatektomi radikal, yang dibagi menjadi data pelatihan sebanyak 508 pasien dan data validasi eksternal sebanyak 108 pasien. Kinerja model dan interpretasi radiologis dinilai atau dievaluasi menggunakan *area under the receiver operating characteristic curve* (AUC), sensitivitas, dan spesifisitas. Uji statistik juga digunakan untuk

mengevaluasi perbedaan dalam AUC, dengan nilai P < 0,05 dianggap signifikan. Hasil penelitian menunjukkan bahwa model TransCL memiliki kemampuan deteksi patologi merugikan (AP) dengan nilai AUC sebesar 0,813, yang lebih tinggi dibandingkan model TransNet yang memiliki nilai AUC 0,791. Selain itu, nilai AUC TransCL juga secara signifikan lebih tinggi dibandingkan model klinis atau *Clinical Model* (CM) yang memiliki nilai 0,749 dan interpretasi radiologis atau *Radiologists Interpretation* (RI) yang hanya 0,664. Ini menunjukkan bahwa TransCL lebih efektif dalam mendeteksi keberadaan AP pada pasien kanker prostat. Kesimpulan dari penelitian ini adalah bahwa model TransNet dan TransCL memiliki potensi untuk membantu dalam mendeteksi adanya patologi yang berbahaya (AP). Hasil penelitian ini dapat berkontribusi pada peningkatan akurasi dan efisiensi dalam aplikasi klinis, sehingga berguna dalam diagnosis dan pengobatan kanker prostat. Penelitian ini juga menyarankan untuk meningkatkan interpretabilitas model dan *dataset* yang lebih besar pada penelitian mendatang.

Penggunaan *Vision Transformers* (ViTs) 3D juga telah diterapkan untuk memprediksi tingkat keagresifan kanker prostat berdasarkan data pencitraan dari magnetic resonance imaging (MRI) multiparametrik [20]. Penelitian ini didasari oleh banyaknya jumlah kanker prostat hingga menjadi jenis kanker paling umum pada pria di Eropa, dan prosedur diagnosis saat ini sebagian besar bergantung pada biopsi, yang bersifat invasif dan tidak nyaman. Sebelum biopsi dilakukan, pasien biasanya menjalani pemeriksaan MRI untuk membantu radiolog dalam membuat penilaian awal terhadap tumor. Pencitraan ini mencakup gambar T2-weighted yang memberikan detail anatomi serta peta pencitraan berbasis difusi (DWI) yang juga menghasilkan informasi fungsional tentang jaringan. Penilaian awal ini seringkali dipengaruhi oleh pengalaman radiolog dan protokol pencitraan yang digunakan, sehingga bisa menyebabkan overdiagnosis atau salah klasifikasi keparahan tumor. Maka, penelitian ini mengusulkan alternatif non-invasif untuk biopsi dengan menggunakan sistem otomatis yang memanfaatkan model pembelajaran mendalam (*deep learning*) yang dilatih menggunakan data pencitraan. Penelitian ini menggunakan model ViTs 3D, yang pada penelitian terbaru saat ini menunjukkan kemampuan dari model ini dalam mengungguli *Convolutional Neural Networks*

(CNNs) dalam berbagai tugas penglihatan komputer, seperti klasifikasi, deteksi, dan segmentasi. Penelitian ini berbeda dengan banyak penelitian lainnya yang menggunakan CNN atau data 2D, dimana penelitian ini menggunakan ViT 3D yang dapat memproses data volumetrik, mencerminkan struktur tiga dimensi dari gambar MRI. Arsitektur ViT 3D dimodifikasi dari ViT standar untuk menangani input 3D dengan mengganti lapisan konvolusi 2D dengan lapisan konvolusi 3D, sehingga memungkinkan model untuk memproses patch volumetrik. Model ViT 3D dilatih dari awal (*train from scratch*) menggunakan *dataset* ProstateX-2, yang terdiri dari 112 citra dari 99 pasien, dan dikategorikan berdasarkan tingkat keagresifan kanker prostat (Gleason *score* $\leq$ 3+4 untuk kanker tingkat rendah dan Gleason *score* $\geq$ 4+3 untuk kanker tingkat tinggi). Berdasarkan penelitian ini, penggunaan *dataset* ini cukup menantang karena ukuran sampel yang kecil serta kompleksitas dalam mengklasifikasikan keagresifan kanker prostat. Untuk membandingkan performa ViT, penelitian ini juga mengembangkan model CNN 3D yang dilatih dengan kondisi yang sama. Kedua model dilatih dari nol tanpa menggunakan *transfer learning*, sehingga pengujian ini menjadi pengujian yang ketat terhadap kapabilitas model dengan data yang terbatas. Sebelum dilakukan pemodelan, persiapan data yang dilakukan mencakup pemilihan irisan gambar yang berguna, mengubah ukuran gambar menjadi standar, melakukan *cropping* agar hanya fokus pada kelenjar prostat dan mengurangi gambar di sekitarnya yang tidak relevan. Setelah itu, teknik augmentasi data seperti *flipping* dan rotasi juga digunakan untuk mengatasi ketidakseimbangan antara jumlah kanker tingkat rendah dan tingkat tinggi. Setelah langkah ini, data akhir yang digunakan terdiri dari irisan gambar berukuran 128x128 piksel, dengan lima irisan per kanker. Selanjutnya, penelitian ini juga melakukan *grid search* untuk mengoptimalkan parameter arsitektur model seperti jumlah lapisan, jumlah attention heads, dan ukuran *multi-layer perceptron* (MLP). Setelah menemukan konfigurasi terbaik, model dilatih menggunakan pendekatan 5 fold cross-*validation* untuk mengurangi kemungkinan *overfitting*. Hasil penelitian ini menunjukkan bahwa dari segi performa, ViT 3D secara signifikan mengungguli CNN 3D dalam mengklasifikasikan tingkat keagresifan kanker prostat. ViT berhasil mencapai akurasi *area under the receiver operating*

characteristic curve (AUROC) sebesar 0,927 pada *dataset* uji, dibandingkan dengan 0,585 untuk CNN. ViT juga menunjukkan sensitivitas, spesifisitas, dan skor *F1* yang lebih baik, sehingga dapat mengidentifikasi kanker tingkat rendah dan tinggi dengan lebih akurat. Hasil ini cukup bagus, mengingat ukuran *dataset* yang digunakan cukup kecil dan tidak menggunakan metode *transfer learning*, yang biasanya diperlukan untuk model ViT agar bekerja dengan baik karena tuntutannya akan data yang besar. Keterbatasan pada penelitian ini berada terutama pada ukuran *dataset* yang kecil dan penggunaan gambar MRI T2-weighted saja, tanpa memanfaatkan modalitas pencitraan lain seperti ADC, yang bisa meningkatkan akurasi model dalam membedakan tumor agresif. Maka, penelitian lebih lanjut yang disarankan dari penelitian ini adalah dapat menggunakan *dataset* yang lebih besar dan integrasi data pencitraan tambahan untuk meningkatkan akurasi dan generalisasi model ViT.

Selain kanker prostat, *Vision Transformers* (ViTs) juga telah digunakan untuk meningkatkan diagnosis kanker payudara dengan menggunakan citra *mamogram multi-view* [21]. Dalam praktik medis, kanker payudara biasanya didiagnosis menggunakan mamografi, dan meskipun teknik ini telah terbukti membantu dalam deteksi dini, namun interpretasi mamografi sering kali sulit bagi ahli radiologi, mengakibatkan tingkat kesalahan yang tinggi dan variabilitas antar pembaca hasil mamogram. Untuk mengatasi masalah ini, berbagai skema diagnosis berbantuan komputer atau *Computer-Aided Diagnosis* (CAD) telah dikembangkan. Salah satunya adalah *deep learning*, dimana sejak 2015, metode *deep learning* menjadi pendekatan utama dalam pengembangan CAD berbasis mamografi karena kemampuannya dalam menganalisis gambar secara keseluruhan tanpa perlu mengekstraksi fitur secara manual. Namun, sebagian besar CAD berbasis mamografi yang ada hanya menggunakan satu gambar mamografi untuk melakukan analisis, padahal setiap pasien dalam skrining mamografi biasanya memiliki empat gambar dari dua sudut pandang yang berbeda, yaitu *craniocaudal* (CC) dan *mediolateral oblique* (MLO) dari kedua payudara (kiri dan kanan). Ahli radiologi biasanya membaca keempat gambar ini secara bersamaan untuk membuat diagnosis yang lebih komprehensif. Dalam penelitian ini, pendekatan menggunakan model

ViT digunakan untuk menganalisis empat gambar sekaligus, dengan fokus pada pengenalan hubungan jangka panjang antar-patch gambar. Model yang diusulkan menggunakan blok *transformer* lokal untuk mempelajari hubungan dalam setiap gambar individu, dan blok *transformer* global untuk memahami hubungan antar-gambar. *Dataset* yang digunakan terdiri dari 949 set mamogram, termasuk 470 kasus ganas dan 479 kasus jinak atau normal. Model dilatih dan diuji menggunakan metode 5 fold cross *validation*. Hasil penelitian menunjukkan bahwa tanpa perlu melakukan langkah pra-pemrosesan yang kompleks, seperti pemotongan gambar atau registrasi gambar, model berbasis *transformer* ini mencapai tingkat akurasi klasifikasi kasus yang lebih tinggi dengan AUC sebesar 0,818, dibandingkan dengan model CNN berbasis *multi-view* dengan AUC sebesar 0,784. Penelitian ini menunjukkan potensi ViT dalam pengembangan sistem CAD yang lebih akurat dan efisien untuk diagnosis kanker payudara. Penelitian ini menyarankan pada penelitian selanjutnya untuk meningkatkan interpretabilitas *transformer* dan optimasi arsitektur untuk gambar medis multi-view.

Model *transformer* seperti *Swin Transformer* juga telah digunakan untuk mendeteksi kanker paru-paru dari gambar *computed tomography* (CT) [22]. Penelitian ini didasari oleh permasalahan kanker paru-paru yang menjadi salah satu penyebab utama kematian terkait kanker di seluruh dunia, dan keberhasilan pengobatan sangat bergantung pada diagnosis dini, dimana deteksi kanker paru-paru sering kali menghasilkan banyak positif palsu dan negatif palsu, yang meningkatkan beban bagi pasien dan tenaga medis. Oleh karena itu, pengembangan sistem diagnosis berbantuan komputer atau computer aided diagnosis (CAD) berbasis kecerdasan buatan (AI) dapat membantu mengurangi kesalahan dan mempercepat proses diagnosis. Penelitian ini berfokus pada arsitektur Swin *Transformer* sebagai pendekatan baru untuk analisis citra medis, khususnya deteksi kanker paru-paru. Arsitektur ini memanfaatkan mekanisme *self-attention* yang efektif dalam menangkap hubungan jangka panjang antar-patch gambar, baik pada tingkat lokal maupun global. Model *Swin Transformer* yang dirancang pada penelitian ini menggunakan empat tahap yang dirancang secara hierarkis untuk memproses gambar, dengan masing-masing tahap memperhalus fitur yang

diekstraksi dari data input. Pada tahap pertama, gambar CT dibagi menjadi patch kecil, dan patch ini diproses untuk mengekstraksi pola dasar yang penting. Tahap kedua dan ketiga memperdalam analisis fitur yang lebih kompleks terkait deteksi kanker, sementara tahap terakhir menyatukan semua informasi untuk menghasilkan prediksi. Dalam penerapan ini, dilakukan modifikasi pada ukuran jendela (*window size*) dari 7 menjadi 6, serta pengurangan jumlah lapisan dari 18 menjadi 12 untuk meningkatkan efisiensi dan akurasi deteksi tumor yang lebih kecil. Model Swin *Transformer* dioptimalkan menggunakan teknik *transfer learning* dan augmentasi data untuk meningkatkan generalisasi dan akurasi model. *Dataset* yang digunakan adalah IQ-OTH/NCCD, yang berisi gambar CT dari pasien yang didiagnosis dengan berbagai tahap kanker paru-paru, serta beberapa pasien sehat. *Dataset* ini mencakup 1.097 gambar yang dikelompokkan ke dalam tiga kategori yaitu normal, jinak, dan ganas. Hasil penelitian menunjukkan bahwa model Swin *Transformer* yang diusulkan mencapai akurasi 97,58% dengan *F1 score* sebesar 95,42%, yang mengungguli berbagai model *deep learning* lainnya seperti ResNet50, DenseNet121, dan Xception, serta model *transformer* lainnya seperti *Vision Transformer* (ViT) dan DeiT. Swin *Transformer* terbukti lebih efektif dalam menangkap fitur penting dari gambar kanker paru-paru, baik dalam skala yang kecil hingga besar, dan dapat digunakan pada peningkatan kinerja diagnostik untuk kanker. Penelitian ini menyarankan untuk penelitian di masa depan dapat menyelidiki kinerja model pada set data yang beragam dan jenis kanker lainnya.

Klasifikasi penyakit glioma menggunakan citra MRI T1 juga dapat dilakukan dengan memanfaatkan arsitektur *Vision Transformers* (ViT) [23]. Dalam konteks klasifikasi tumor otak, ViT menawarkan pendekatan yang berbeda dari *Convolutional Neural Networks* (CNN), yang sering kali terbatas dalam menangkap hubungan jarak jauh antar fitur dalam citra medis. Penelitian ini memanfaatkan *dataset* BraTS 2015 yang mencakup gambar MRI glioma derajat tinggi (HGG), glioma derajat rendah (LGG), dan gambar tanpa tumor, dengan total 42.196 gambar hasil *preprocessing*. Untuk mengatasi keterbatasan data medis yang kecil, model ViT ini menggunakan *transfer learning* dengan pre*training* pada *dataset* CIFAR-10 dan *fine-tuning* menggunakan BraTS 2015. *Preprocessing* data dilakukan

dengan teknik *intensity uniformization* untuk menyesuaikan rentang piksel menjadi skala grayscale 8-bit, sehingga mengurangi kompleksitas data. Model ViT dioptimalkan dengan Adam *optimizer*, menggunakan resolusi patch 4×4 untuk memaksimalkan akurasi klasifikasi. Penelitian ini membandingkan performa ViT dengan CNN pada tugas klasifikasi tumor otak dan menunjukkan bahwa ViT unggul dalam tugas klasifikasi dua kelas (tumor dan non-tumor) dengan akurasi 81,8%. Namun, untuk klasifikasi tiga kelas (HGG, LGG, non-tumor), performa menurun akibat ketidakseimbangan kelas. Penelitian ini menyimpulkan bahwa ViT memiliki potensi besar dalam analisis citra medis berkat kemampuannya menangkap hubungan global, meskipun tetap membutuhkan *dataset* yang lebih besar dan teknik augmentasi data untuk meningkatkan generalisasi.

Teknik *image processing* juga sudah diterapkan bersamaan dengan algoritma *deep learning* untuk meningkat akurasi model. Salah satunya, klasifikasi tumor otak menggunakan gambar MRI sudah dilakukan dengan menggabungkan teknik peningkatan citra atau *image enhancement*, yang merupakan bagian dari *image processing*, dengan model *Convolutional Neural Networks* (CNN) [15]. Penelitian ini menggunakan *Gaussian-blur-based sharpening filter* dan *Contrast-Limited Adaptive Histogram Equalization* (CLAHE) untuk membantu meningkatkan kejelasan dan detail dalam gambar MRI otak yang cukup kompleks. *Image processing* perlu diterapkan pada gambar MRI pada konteks penelitian ini dikarenakan penggunaan data MRI membutuhkan pemrosesan yang akurat pada sifat jaringan otak yang rumit. Tujuan dari metode *image processing* ini meningkatkan kontras dan ketajaman gambar MRI, sehingga area tumor menjadi lebih terlihat. Penelitian ini membuat model CNN dengan beberapa lapisan konvolusi dan pooling, dikombinasikan dengan *skip connection* dan lapisan *dense* untuk meningkatkan ekstraksi fitur. Model ini dioptimalkan menggunakan Adam *optimizer* dengan *dynamic learning rate adjustment* menggunakan *callback* ReduceLROnPlateau. Untuk mencegah *overfitting*, model CNN ini menggunakan regularisasi L1 dan L2, dan metode dropout. *Dataset* yang digunakan berisi 7023 gambar MRI grayscale yang diambil dari database publik seperti Kaggle, Figshare, SARTAJ, dan BR35H, yang mencakup kasus glioma, meningioma, tumor pituitari,

serta gambar tanpa tumor. Hasil penelitian ini menunjukkan bahwa metode yang diusulkan (proposed method) yang menggunakan teknik *image processing* dapat mencapai akurasi sebesar 97,84% dengan presisi, *recall*, dan *F1-score* yang lebih baik serta waktu *testing* yang lebih singkat, sehingga dapat melampaui model pre-*train*ed seperti VGG16, ResNet50, dan InceptionV3 yang memiliki tingkat akurasi lebih rendah. Penelitian ini menunjukkan bahwa penggabungan *image processing* dengan teknik *deep learning* dapat secara signifikan meningkatkan akurasi klasifikasi tumor otak dari gambar MRI, sehingga dapat membantu diagnosis medis dengan tingkat akurasi yang baik.

Penerapan teknik *image processing* lainnya sudah dilakukan pada penelitian yang mengklasifikasikan lesi jinak dan ganas pada kanker payudara dengan menggunakan gambar MRI, yang menggabungkan teknik pengolahan citra *Sobel operator* dalam model *deep learning multi-modal* [16]. Penelitian ini menggunakan beberapa modalitas MRI, termasuk T2WI, non-fs T1WI, pre-kontras T1WI, dan post-kontras T1WI, yang digabungkan dalam model *deep learning* multi-modal untuk meningkatkan akurasi diagnosis. Penggunaan *Sobel operator* bertujuan untuk mengekstraksi informasi tepi lesi, sehingga menonjolkan perbedaan antara lesi jinak dan ganas dengan lebih jelas. Teknik ini membantu mengurangi redundansi informasi pada bounding cube minimum yang digunakan untuk segmentasi lesi. Penelitian ini memanfaatkan model *deep learning* berbasis *Convolutional Neural Networks* (CNN) yang mencakup *residual block, Convolutional Block Attention Module* (CBAM), dan *down-sampling module* untuk meningkatkan kemampuan model dalam mengekstraksi fitur utama dari gambar MRI. *Operator Sobel* diaplikasikan pada setiap modalitas gambar untuk menghasilkan citra gradien tepi yang lebih informatif, yang kemudian digabungkan ke dalam kanal input model multi-modal. *Dataset* penelitian terdiri dari 207 lesi dari 203 pasien, dengan proporsi data yang dibagi menjadi *training* set, *validation* set, dan *test* set. Hasil penelitian menunjukkan bahwa model multi-modal dengan *Sobel operator* memiliki performa superior dibandingkan model single-modal, dengan nilai AUC sebesar 0.887, akurasi 82.9%, sensitivitas 79.8%, dan spesifisitas 86.1%. Model ini secara signifikan lebih baik daripada model single-modal terbaik (DCE-MRI post-

kontras T1WI fase pertama) yang hanya memiliki AUC sebesar 0.827 dan akurasi 78.5%. Penelitian ini menunjukkan bahwa penggabungan informasi dari berbagai modalitas MRI dengan teknik *Sobel operator* dapat meningkatkan kemampuan model *deep learning* dalam mengklasifikasikan lesi kanker payudara. Temuan ini memberikan kontribusi penting dalam mendukung diagnosis kanker payudara yang lebih akurat, terutama untuk dokter dengan pengalaman terbatas dalam membaca citra MRI.

Penerapan *image processing* lainnya juga diterapkan pada klasifikasi tumor otak menggunakan gambar MRI yang dilakukan dengan mengintegrasikan *Gabor filter* dalam model *Convolutional Neural Networks* (CNN) yang dimodifikasi [17]. *Gabor filter* digunakan untuk memodulasi *filter* konvolusi, karena memiliki kemampuan unik dalam menangkap orientasi dan skala, sehingga meningkatkan kemampuan jaringan dalam mengenali fitur visual yang penting pada gambar MRI. Teknik ini diterapkan untuk menangani tantangan analisis citra MRI yang memiliki intensitas heterogen, bentuk tidak beraturan, dan batasan tumor yang sulit dikenali. Model CNN yang diusulkan terinspirasi oleh Res-Net, dengan penambahan *skip connections* untuk menghindari masalah gradien yang menghilang, serta menggunakan modulasi *Gabor filter* untuk memperkaya fitur yang diekstraksi di setiap lapisan konvolusi. *Dataset* BRaTS 2017 digunakan, terdiri dari 285 kasus tumor otak yang mencakup *High-Grade Glioma* (HGG) dan *Low-Grade Glioma* (LGG). Validasi dilakukan menggunakan metode *Leave-One-Patient-Out* (LOPO) untuk memastikan kinerja model pada data pasien yang belum pernah dilihat sebelumnya. Model ini dioptimalkan menggunakan Adam *optimizer* dengan fungsi *loss binary cross-entropy*, serta dilengkapi dengan regularisasi L2 untuk mencegah *overfitting*. Hasil penelitian menunjukkan bahwa metode yang diusulkan dapat mencapai akurasi 98,68% dengan sensitivitas 99,39% dan spesifisitas 96,77%, melampaui model pre-*train*ed seperti Alex-Net, VGG19, ResNet34, dan InceptionNetV1. Penelitian ini membuktikan bahwa integrasi *Gabor filter* dalam arsitektur CNN tidak hanya meningkatkan akurasi klasifikasi tumor otak tetapi juga mengurangi kebutuhan parameter jaringan, sehingga lebih efisien dan cocok untuk diterapkan dalam analisis citra medis.

2.2 Tinjauan Teori

2.2.1 Klasifikasi Prostate Cancer Staging

Sistem *staging* yang paling umum digunakan dalam penentuan stadium kanker prostat adalah sistem penentuan stadium TNM, yang mengevaluasi tiga komponen utama: Tumor (T), Nodul (N), dan Metastasis (M). Penjelasan klasifikasi di dalam *staging* TNM dijabarkan pada gambar 2.1.

CLASSIFICATION	DEFINITION
TUMOR	
Tx	Tumor cannot be evaluated (due to lack of information)
T0	No evidence of a primary tumor
T1*	Tumor was not detected during a digital rectal exam (DRE) and cannot be seen on imaging studies (tumor may be discoverd during surgery for a reason other cancer)*
T2 T2a T2b T2c	Tumor can be detected during a DRE but is present in the prostate only Tumor is in half or less than one side (lobe) of the prostate Tumor is in more than half of one prostate lobe, but has not yet invaded the other lobe Tumor is in both prostate lobes
T3 T3a T3b	Tumor extends outside of the prostate Tumor extends outside the prostate on one or both sides Tumor has spread to the seminal vesicles (the glands on each side of the bladder)
T4	Tumor has spread to tissues near the prostate other than the seminal vesicles, such as the bladder or wall of the plevis
Nearby (regional) lymph nodes (N)	
Nx	Nearby lymph nodes are not evaluated
N0	No cancer cells are found in nearby lymph nodes
N1	Cancer cells are found in nearby lymph nodes
Distant Metastasis (M)	
M0 M1 M1a M1b M1c	Cancer has not spread beyond the prostate Cancer has spread beyond the prostate Cancer has spread to distant lymph nodes Cancer has spread to bone Cancer has spread to another organ or site, with or without bone disease

Gambar 2.1 Staging TNM

Sumber: [24]

Gambar 2.1 menunjukkan klasifikasi *staging* TNM untuk kanker prostat. Staging kanker prostat melibatkan beberapa faktor penting yang membantu dokter menilai seberapa jauh penyakit ini berkembang dan menentukan pilihan pengobatan.

Kategori Tumor (T) mengukur ukuran dan penyebaran tumor utama. Tumor T1 kecil dan tidak dapat terdeteksi melalui pemeriksaan rektal atau pencitraan, tetapi bisa ditemukan melalui biopsi. Tumor T2 terlokalisasi di dalam prostat, sementara tumor T3 telah menyebar di luar prostat tetapi belum ke organ terdekat. Tumor T4 menandakan bahwa tumor telah tumbuh ke organ di sekitar, seperti kandung kemih atau rektum, yang menunjukkan kanker yang lebih lanjut.

Kategori Node (N) berguna untuk melihat apakah kanker telah menyebar ke kelenjar getah bening terdekat. Jika tidak ada kelenjar getah bening yang terpengaruh oleh tumor, maka diklasifikasikan sebagai N0. Jika kanker telah menyebar ke kelenjar getah bening regional, itu dikategorikan sebagai N1. Informasi N penting karena jika tumor telah menyebar ke kelenjar getah bening maka hal ini menandakan kemungkinan besar kanker sudah menyebar ke bagian tubuh lain.

Kategori Metastasis (M) menggambarkan apakah kanker telah menyebar ke bagian tubuh yang jauh. Jika tidak ada bukti penyebaran, maka diklasifikasikan sebagai M0. Jika kanker telah metastasis ke organ jauh, seperti tulang, hati, atau paru-paru, maka diklasifikasikan sebagai M1. Klasifikasi lebih lanjut seperti M1a, M1b, dan M1c menunjukkan lokasi spesifik metastasis, di mana M1b berarti penyebaran ke tulang dan M1c berarti keterlibatan organ lebih lanjut.

Selain sistem TNM, dokter juga memeriksa level PSA (Prostate Specific Antigen), yaitu protein yang diproduksi oleh sel-sel prostat, baik yang sehat maupun yang kanker. Tingginya level PSA bisa mengindikasikan kanker prostat. Terakhir, Skor Gleason mengevaluasi agresivitas kanker prostat berdasarkan seberapa abnormal penampilan sel-sel kanker di bawah hasil mikroskop. Skor ini berkisar dari 6 hingga 10, dengan skor yang lebih rendah menunjukkan kanker yang

kurang agresif dan skor yang lebih tinggi menunjukkan kanker yang lebih agresif dan cepat tumbuh. Skor Gleason sangat penting dalam menentukan pilihan pengobatan, karena kanker yang lebih agresif mungkin memerlukan perawatan yang lebih intensif. Maka dari itu, Level PSA, klasifikasi TNM, dan skor Gleason digunakan untuk membantu memprediksi tingkat keparahan penyakit dan kemungkinan penyebarannya, karena faktor-faktor ini dapat memberikan gambaran menyeluruh tentang tahap dan agresivitas kanker prostat untuk membantu dokter dalam mengambil keputusan pengobatan dan memprediksi hasilnya.

Dalam penelitian ini, klasifikasi *staging* yang akan diterapkan terbatas pada kategori Metastasis (M). Pemilihan ini didasarkan pada fokus penelitian yang hanya menganalisis data gambar MRI pelvis, yang bertujuan untuk mendeteksi *staging* kanker berdasarkan kategori Metastasis yang digunakan untuk menilai sejauh mana penyebaran kanker ke organ lain di luar prostat. Prosedur ini tidak mencakup analisis dengan kategori Tumor (T) dan Nodul (N), karena klasifikasi Tumor dan Nodul berfungsi untuk menilai tingkat keparahan kanker prostat itu sendiri. Oleh karena itu, untuk tujuan deteksi metastasis yang lebih akurat dan spesifik, hanya parameter M yang akan digunakan dalam penelitian ini.

2.2.2 Deep Learning

Deep learning atau pembelajaran mendalam adalah bagian dari pembelajaran mesin, yang merupakan bagian dari kecerdasan buatan. Pembelajaran ini berfokus pada algoritme yang terinspirasi dari struktur dan fungsi otak, yang dikenal sebagai jaringan saraf tiruan [25]. Jaringan ini dirancang untuk meniru cara otak manusia memproses informasi, menggunakan beberapa lapisan untuk mengekstrak fitur tingkat yang lebih tinggi secara progresif dari data masukan mentah [26].

Deep learning memiliki beberapa karakteristik utama yang membedakannya dari pendekatan pembelajaran mesin lainnya. Salah satunya adalah penggunaan jaringan saraf tiruan (artificial neural networks) yang terdiri dari lapisan-lapisan neuron yang saling terhubung. Jaringan saraf ini memproses data dengan cara yang menyerupai cara kerja otak manusia [27]. Model *deep learning* juga memanfaatkan banyak lapisan tersembunyi (hidden layers) untuk melakukan transformasi dan

abstraksi data yang kompleks, sehingga dapat mempelajari pola-pola dan representasi yang rumit [28]. Selain itu, *deep learning* juga menerapkan transformasi non-linear pada data, yang memungkinkan pemodelan hubungan-hubungan kompleks antara berbagai fitur dalam data tersebut [28].

Deep learning sudah banyak diterapkan dalam berbagai bidang. Dalam computer vision, *deep learning* dapat digunakan untuk melakukan klasifikasi gambar, deteksi objek, dan pengenalan wajah dengan akurasi yang tinggi [29]. Dalam konteks computer vision, *deep learning* bahkan dapat digunakan hingga ke ranah politik untuk sistem pemungutan suara online [30]. Di bidang *Natural Language Processing* (NLP), *deep learning* digunakan untuk berbagai tugas, seperti terjemahan bahasa otomatis dan analisis sentimen, yang memungkinkan pemahaman dan pengolahan bahasa alami lebih efektif [29]. Di dunia medis, *deep learning* juga memberikan kontribusi besar, terutama dalam penemuan obat, analisis genom, dan pencitraan medis, yang berpotensi meningkatkan diagnosis dan pengobatan penyakit [29].

2.2.3 *Transformer*

Transformer adalah jenis model *deep learning* yang didasarkan pada mekanisme self attention. Pada awalnya *transformer* dikembangkan untuk pemrosesan bahasa alami (NLP), namun saat ini telah diadaptasi untuk berbagai bidang, termasuk computer vision dan pembelajaran multimodal [31]. Fitur utama dari *transformer* adalah self attention mechanisme atau mekanisme perhatian diri, dimana mekanisme itu merupakan komponen inti dari *transformer*, yang memungkinkan model untuk menimbang pentingnya berbagai bagian dari data input, memungkinkannya untuk menangkap ketergantungan dan hubungan yang kompleks di dalam data [31]. Selain itu, berbeda dari jaringan saraf tiruan (RNN), *transformer* memproses data secara paralel, yang secara signifikan dapat mempercepat waktu pelatihan dan inferensi. Pengembangan dari model *transformer* untuk tugas computer vision dinamakan Vision *transformer* (ViT). ViT telah dikembangkan untuk menerapkan arsitektur *transformer* pada data gambar,

yang sering kali mengungguli jaringan saraf convolutional (CNN) tradisional dalam berbagai tugas visual.

2.2.4 *Gaussian Filter*

Filter Gaussian adalah teknik yang sering digunakan dalam pemrosesan sinyal dan analisis gambar, terutama untuk mengurangi noise dan menghaluskan data. Teknik ini bekerja dengan menerapkan fungsi *Gaussian*, yang memiliki bentuk kurva lonceng, pada data input. *Filter Gaussian* efektif dalam mengurangi noise dengan cara menghaluskan perubahan intensitas yang cepat, namun tetap mempertahankan komponen frekuensi rendah dalam data [32]. *Filter Gaussian* banyak diterapkan dalam pemrosesan gambar untuk tugas-tugas seperti penghalusan gambar dan pengurangan noise, dalam pemrosesan sinyal untuk menyaring noise dan meningkatkan akurasi data sensor, bahkan juga dapat digunakan dalam komputasi kuantum untuk memproyeksikan keadaan kuantum dengan lebih efisien [33]. Maka dari itu, *Filter Gaussian* merupakan teknik pemrosesan citra yang sangat penting dalam bidang komputasi baik klasik dan modern, karena dapat memberikan solusi yang tangguh untuk pengurangan noise dan penghalusan data.

2.2.5 *Sobel Operator*

Operator Sobel adalah metode yang sering digunakan untuk mendeteksi tepi dalam pemrosesan gambar dan computer vision, dan dirancang untuk mengidentifikasi gradien intensitas gambar pada setiap piksel [34]. Teknik ini bekerja dengan melingkarkan gambar dengan sepasang kernel 3x3, satu untuk mendeteksi tepi horizontal dan yang lainnya untuk tepi vertikal [35]. *Operator* ini disukai karena kesederhanaan dan efisiensinya, sehingga cocok untuk aplikasi real-time. Kemajuan terbaru telah memperluas kemampuannya, yang mengarah ke implementasi multi-arah yang meningkatkan akurasi pendeteksian tepi. *Operator Sobel* tidak hanya digunakan untuk deteksi tepi, tetapi juga dalam aplikasi tingkat lanjut seperti interpolasi warna dan fusi gambar medis, di mana *operator* ini membantu meningkatkan kualitas gambar dan pelestarian fitur [36].

2.2.6 *Gabor Filter*

Filter Gabor adalah *filter* linier yang digunakan untuk deteksi tepi dan analisis tekstur karena kemampuannya menangkap informasi spasial dan frekuensi secara efektif. *Filter* ini sangat menguntungkan untuk tugas-tugas yang melibatkan analisis tekstur, ekstraksi fitur relevan, dan segmentasi gambar, sehingga membuatnya penting dalam berbagai aplikasi, termasuk visi komputer dan pencitraan biomedis sebelum diproses lebih lanjut dengan teknik pembelajaran mesin [37].

Filter Gabor memiliki dua fitur utama yang membuatnya sangat berguna dalam analisis sinyal seperti gambar. Pertama, *filter* ini dapat menggabungkan bentuk gelombang sinus dengan fungsi *Gaussian*, yang memungkinkannya untuk menganalisis gambar dalam dua aspek sekaligus, yaitu dalam waktu dan frekuensi. Hal ini membuatnya sangat efektif untuk memahami detail gambar secara lebih lengkap. Kedua, *filter Gabor* sangat fleksibel karena pengguna dapat menyesuaikan berbagai parameter, seperti ukuran, arah, dan frekuensi gelombang. Dengan kemampuan untuk menyesuaikan parameter-parameter ini, *filter Gabor* dapat digunakan untuk menganalisis berbagai jenis gambar sesuai kebutuhan [38].

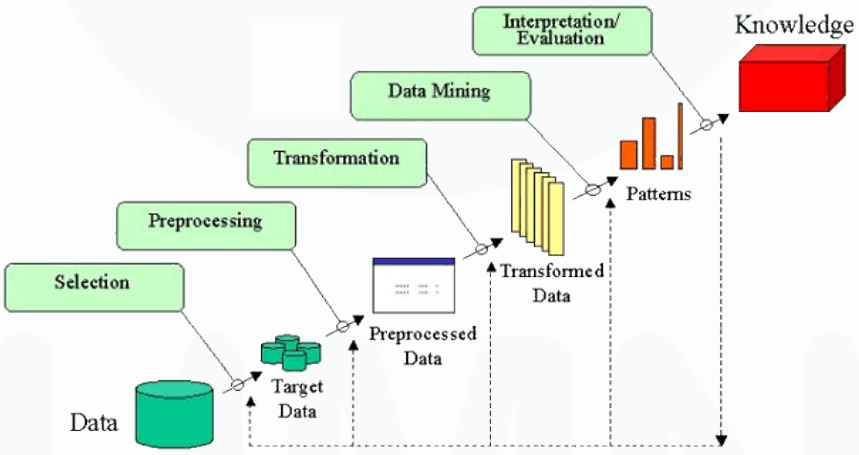
2.2.7 PyTorch

PyTorch adalah kerangka kerja *deep learning* open-source yang populer, yang menyediakan platform fleksibel dan efisien untuk membangun serta menerapkan model machine learning. PyTorch banyak digunakan dalam bidang penelitian dan industri karena kemudahan penggunaannya serta kemampuannya yang kuat. Salah satu fitur utama PyTorch adalah integrasinya dengan Python, yang memungkinkan pengguna mendefinisikan fungsi matematika dan menghitung gradien dengan mudah, hal ini sangat penting dalam berbagai tugas *deep learning* [39]. Selain itu, PyTorch juga mendukung penggunaan GPU, yang mempercepat pelatihan model dengan memanfaatkan kemampuan pemrosesan paralel dari GPU [39]. PyTorch juga menyediakan modul pelatihan terdistribusi yang memungkinkan pelatihan model besar secara efisien di beberapa GPU, dengan teknik optimasi seperti pembagian gradien dan penggabungan komputasi dengan komunikasi [40].

Fleksibilitas PyTorch juga menjadi keunggulan, karena pengguna dapat membangun berbagai model untuk beragam aplikasi, mulai dari klasifikasi gambar hingga penerjemahan mesin [41].

2.2.8 *Knowledge discovery in databases* (KDD)

Knowledge discovery in databases (KDD) adalah proses sistematik untuk penemuan pola dan pemodelan prediktif dalam kumpulan data yang besar, dengan memanfaatkan metode penambangan data dan teknik pembelajaran mesin untuk mengekstrak wawasan yang bermakna [42]. Proses KDD biasanya melibatkan beberapa tahap, termasuk pemilihan data, prapemrosesan, transformasi, penambangan, evaluasi, dan interpretasi, yang secara kolektif meningkatkan pengambilan keputusan di berbagai domain, seperti dalam konteks kesehatan dan bisnis.



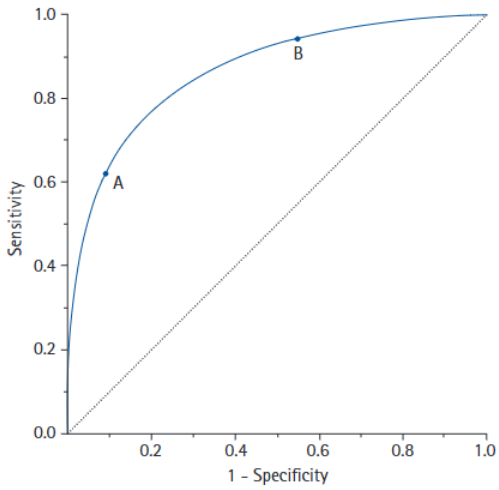
Tahapan seleksi mencakup kegiatan pengumpulan data yang relevan. Setelah target *dataset* ditentukan, langkah selanjutnya adalah tahap *preprocessing*, yang mencakup pembersihan data dan proses lainnya untuk memastikan kualitas dan relevansi data. Tahapan berikutnya adalah transformasi, di mana data dimodifikasi menggunakan berbagai metode, seperti pengurangan dimensi, agar lebih mudah diproses pada tahap selanjutnya. Setelah itu, dilakukan tahap *data mining*, yang bertujuan untuk menemukan pola-pola yang relevan sesuai dengan tujuan yang telah ditetapkan sebelumnya. Terakhir, tahapan evaluasi berfokus pada interpretasi dari pola-pola yang ditemukan [43].

2.2.9 *ROC Curve*

Receiver Operating Characteristic (ROC) *Curve* atau kurva ROC adalah representasi grafis yang digunakan untuk mengevaluasi performa metode klasifikasi diagnostik biner. Kurva ini memetakan hubungan antara *sensitivity* atau tingkat positif sebenarnya (*True Positive* Rate, TPR) pada sumbu y, dan *1-specificity* atau tingkat positif palsu (*False Positive* Rate, FPR) pada sumbu x untuk setiap nilai *cut-off* yang dipilih [44].

Dalam konteks medis, ROC curve digunakan untuk mengklasifikasikan kondisi pasien, seperti keberadaan atau ketiadaan penyakit, berdasarkan hasil tes diagnostik yang memiliki nilai kontinu. Semakin dekat kurva ROC ke pojok kiri atas grafik, semakin baik performa metode diagnostik dalam membedakan kondisi. Sebaliknya, jika kurva sejajar dengan garis diagonal (klasifikasi acak), metode tersebut tidak memiliki nilai diagnostik yang signifikan [44].

Salah satu kelebihan ROC curve adalah kemampuannya untuk memberikan gambaran menyeluruh tentang kinerja diagnostik tanpa terpengaruh oleh prevalensi penyakit. Hal ini membuat ROC curve sangat berguna, terutama dalam situasi dengan distribusi data yang tidak seimbang [44].

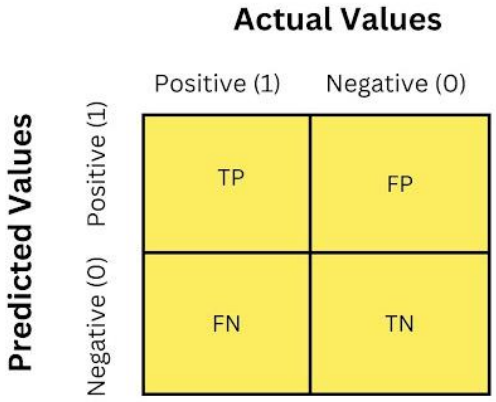


Area Under the Curve (AUC) pada kurva ROC merupakan ukuran numerik yang digunakan untuk menilai akurasi keseluruhan dari suatu metode diagnostik. Nilai AUC berkisar antara 0 hingga 1, di mana AUC = 1 menunjukkan performa

yang sempurna, sedangkan AUC = 0.5 menunjukkan performa yang setara dengan tebakan acak. Secara umum, nilai AUC yang lebih besar dari atau sama dengan 0.8 dianggap menunjukkan akurasi yang baik. Dengan kata lain, semakin tinggi nilai AUC, semakin baik kemampuan metode diagnostik dalam mendeteksi kondisi yang sebenarnya dengan benar pada berbagai nilai cut-off [44].

2.2.10 *Confusion Matrix*

Confusion matrix adalah tabel yang digunakan untuk menilai kinerja sebuah model klasifikasi. Tabel ini menunjukkan perbandingan antara hasil yang diprediksi oleh model dan hasil yang sebenarnya (aktual). Dalam klasifikasi biner, ada empat kategori utama yang ditampilkan dalam matriks ini yaitu *True Positive*, *False Positive*, *True Negative*, *False Negative* [45].



True Positive (TP) merujuk pada jumlah kasus yang sebenarnya positif dan juga diprediksi sebagai positif oleh model. *False Positive* (FP) adalah jumlah kasus yang sebenarnya negatif, tetapi oleh model diprediksi sebagai positif. Sementara itu, *True Negative* (TN) adalah jumlah kasus yang benar-benar negatif dan diprediksi dengan benar sebagai negatif. Terakhir, *False Negative* (FN) menggambarkan jumlah kasus yang sebenarnya positif, namun diprediksi sebagai negatif oleh model. Keempat kategori ini membantu kita untuk mengevaluasi seberapa baik model dalam membuat prediksi yang tepat [45].

Aplikasi *confusion matrix* dapat dibagi menjadi dua kategori utama. Pertama, klasifikasi biner yang digunakan untuk mengevaluasi model yang memprediksi dua

hasil yang mungkin, yaitu dua kategori yang saling eksklusif. Kedua, klasifikasi multi-kelas yang diperluas untuk masalah dengan lebih dari dua kelas, di mana matriks yang digunakan berbentuk persegi dengan dimensi yang sesuai dengan jumlah kelas yang ada. Dalam konteks ini, setiap nilai dalam matriks menggambarkan hubungan antara kelas yang diprediksi dan kelas yang sebenarnya, memungkinkan evaluasi performa model pada masalah yang lebih kompleks [45].