

BAB II TINJAUAN PUSTAKA

Bab ini menyajikan landasan teoretis yang komprehensif, tinjauan empiris yang mendalam, serta kerangka konseptual yang menjadi fondasi utama bagi penelitian ini. Pembahasan disusun secara sistematis dan hierarkis, dimulai dari pemetaan posisi penelitian ini dalam khazanah keilmuan (*state of the art*) melalui tinjauan kritis terhadap penelitian terdahulu yang relevan. Selanjutnya, dibahas secara mendetail aspek klinis yang meliputi anatomi ginjal, patofisiologi tumor ginjal, serta spesifikasi teknis dan karakteristik dataset KiTS23 (*The 2023 Kidney and Kidney Tumor Segmentation Challenge*) yang menjadi objek material utama penelitian ini. Bagian inti dari bab ini mengeksplorasi landasan matematis dan arsitektural dari *Deep Learning*, *Convolutional Neural Networks* (CNN), arsitektur U-Net, serta mekanisme tingkat lanjut berupa *Residual Learning* dan *Attention Gates*. Bab ini diakhiri dengan sintesis kerangka pemikiran yang merangkum alur logika penyelesaian masalah.

2.1 Landasan Medis: Anatomi Ginjal dan Patologi Tumor

Pemahaman mendalam mengenai anatomi organ dan karakteristik patologi tumor merupakan prasyarat mutlak dalam pengembangan algoritma segmentasi medis. Algoritma tidak hanya memproses nilai piksel, tetapi harus mampu mengenali representasi digital dari struktur biologis (Li, 2025).

2.1.1 Anatomi Fisiologis Ginjal

Ginjal (Ren) adalah sepasang organ ekskresi berbentuk seperti kacang yang terletak di ruang retroperitoneal pada dinding posterior abdomen, terlindung di balik peritoneum parietal. Secara topografi, ginjal terbentang dari vertebra torakal 12 hingga lumbal 3 (Dalley & Agur, 2023). Struktur internal ginjal yang relevan dengan citra CT Scan terdiri dari:

1. Parenkim Ginjal003A
 - a Korteks: Bagian terluar yang mengandung glomerulus. Pada CT Scan fase arteri, korteks memiliki vaskularisasi yang sangat tinggi sehingga tampak

hiperdens (sangat terang). Kolumna Bertin adalah bagian korteks yang menjorok ke dalam medula.

- b Medula: Bagian dalam yang berisi piramida ginjal dan tubulus. Medula memiliki vaskularisasi yang lebih sedikit dibanding korteks, sehingga tampak lebih hipodens (gelap) pada fase arteri awal, namun akan menyangat pada fase nefro grafik (parenkimal).
- 2. Sinus Ginjal: Rongga yang berisi lemak (sangat hipodens), pembuluh darah, saraf, dan sistem pelviokalis.
- 3. Sistem Pelviokalis: Terdiri dari kaliks minor, kaliks mayor, dan pelvis renalis yang berfungsi mengalirkan urin ke ureter. Struktur ini berisi cairan dan biasanya memiliki densitas mendekati air (0-10 HU) (DenOtter & Schubert, Hounsfield Unit, 2025).

2.1.2 Karsinoma Sel Ginjal (*Renal Cell Carcinoma*)

Karsinoma Sel Ginjal (RCC) adalah keganasan ginjal primer yang paling umum, mencakup sekitar 90% kasus tumor ginjal ganas. Variabilitas visual tumor pada citra CT sangat dipengaruhi oleh sub tipe histologisnya:

- 1. Clear Cell RCC (ccRCC): Sub tipe paling dominan (70-80%). Karakteristik utamanya adalah hipervaskularitas. Pada fase arteri, ccRCC menyangat sangat kuat dan heterogen karena adanya area nekrosis, perdarahan (*hemorrhage*), atau kista di dalamnya. Heterogenitas ini menjadi tantangan bagi segmentasi otomatis karena teksturnya yang tidak seragam.
- 2. Papillary RCC (pRCC): Sub tipe kedua terbanyak (10-15%). Bersifat hipovaskular, artinya menyangat lebih lemah dibandingkan parenkim ginjal normal. Pada CT Scan, pRCC tampak sebagai massa hipodens yang seringkali sulit dibedakan dari kista ginjal hiperdens. Chromophobe RCC: Sub tipe yang lebih jarang, biasanya tampak homogen (Schieda, 2025).

2.1.3 Perbedaan Tumor dan Kista Ginjal

Salah satu tantangan terbesar dalam dataset KiTS23 adalah perbedaan antara Tumor (Label 2) dan Kista (Label 3).

1. Kista Ginjal (Cyst): Lesi jinak berisi cairan. Kista sederhana (*Simple Cyst*) memiliki dinding tipis, isinya homogen (densitas air, ~ 0 HU), dan tidak menyangat pasca-pemberian kontras. Kista kompleks dapat memiliki septa (sekat), dinding tebal, atau kalsifikasi, yang membuatnya menyerupai tumor kistik maligna (Schieda & al., 2025).
2. Tumor (Neoplasma): Massa jaringan padat yang menyangat (*enhancing*) setelah pemberian kontras. Peningkatan nilai HU ($>115-20$ HU) setelah injeksi kontras adalah tanda kunci keganasan (DenOtter & Schubert, 2025).

Sistem segmentasi otomatis harus mampu mempelajari fitur halus ini—tekstur homogen vs heterogen, dan pola penyangatan dinding—untuk menghindari misklasifikasi kista sebagai tumor (*false positive*) yang dapat memicu intervensi bedah yang tidak perlu.

2.1.4 Karakteristik Citra CT Scan dan *Hounsfield Unit*

Computed Tomography (CT) merepresentasikan densitas jaringan menggunakan skala kuantitatif *Hounsfield Unit* (HU) (DenOtter & Schubert, 2025). Skala ini linier, di mana:

1. Udara = -1000 HU (Hitam pekat)
2. Lemak = -100 hingga -50 HU (Abu-abu gelap)
3. Air = 0 HU (Referensi)
4. Jaringan Lunak = 20 hingga 80 HU (Abu-abu)
5. Tulang Kortikal = >1000 HU (Putih terang)

Dalam konteks ginjal, nilai HU sangat dinamis bergantung pada fase pemindaian (LITFL Radiology Library, n.d):

1. Fase Non-Kontras: Ginjal ($\sim 30-50$ HU), Kista ($\sim 0-20$ HU), Batu Ginjal (>400 HU).
2. Fase Kortikomeduler (Arteri Lanjut): 25-70 detik pasca injeksi. Korteks ginjal menyangat maksimal, medula belum. Diferensiasi korteks-medula sangat jelas.
3. Fase Nefro Grafik (Parenkimal): 80-180 detik pasca injeksi. Kontras merata di seluruh parenkim ginjal. Ginjal tampak homogen terang. Ini adalah fase

terbaik untuk mendeteksi tumor, karena tumor (yang seringkali memiliki vaskularisasi berbeda) akan tampak kontras terhadap latar belakang ginjal yang homogen. Dataset KiTS23 secara spesifik memasukkan banyak kasus dari fase ini (Heller, Isensee, Maier-Hein, & al., 2021).

2.2 Dataset Penelitian: KiTS23

Data penelitian ini bersumber dari repositori publik KiTS23 (*The 2023 Kidney and Kidney Tumor Segmentation Challenge*) (Heller, Isensee, Trofimova, Tejpau, & al., 2023). Dataset ini merupakan kumpulan data CT Scan abdomen multisenter yang dikurasi dengan standar kualitas tinggi untuk keperluan pengembangan algoritma kecerdasan buatan.

2.2.1 Spesifikasi dan Statistik Data

Dataset KiTS23 memiliki karakteristik yang jauh lebih kompleks dan representatif dibandingkan pendahulunya (KiTS19). Berikut adalah rincian spesifikasi dataset:

Tabel 2.1 Spesifikasi Statistik Dataset KiTS23 (Heller, Isensee, Trofimova, Tejpau, & al., 2023)

Atribut	Detail Statistik	Keterangan
Total Kasus	599 Kasus	489 Data Latih (<i>Training</i>), 110 Data Uji (<i>Testing</i>)
Format Data	NIfTI (.nii.gz)	Format standar neuroimaging/medis 3D
Resolusi Spasial	Bervariasi	Diperlukan <i>resampling</i> ke isometrik (misal 1.5mm ³)
Fase Kontras	Campuran	Meliputi fase <i>Late Arterial</i> (seperti KiTS19) dan <i>Nephrogenic</i>

Jumlah Ginjal	489 kasus (100% Coverage)	Mencakup ginjal normal dan anomali
Jumlah Tumor	489 kasus (100% Coverage)	Termasuk tumor kecil hingga massa raksasa
Jumlah Kista	248 kasus (50.72% Coverage)	Hanya separuh pasien yang memiliki kista
Volume Tumor	Min: 0.2 cm ³ , Max: 2590.96 cm ³	Variasi ukuran ekstrem (>10.000x lipat)
Volume Kista	Min: 0.04 cm ³ , Max: 937.4 cm ³	Rentang ukuran yang sangat luas

Berdasarkan spesifikasi resmi pada Tabel 2.1, terdapat 489 data latih yang tersedia. Namun, dalam implementasi eksperimen ini, jumlah kasus yang diolah adalah 488 pasien. Hal ini disebabkan oleh adanya satu kasus (ID: case_00448) yang mengalami kerusakan data (*corrupt*) pada saat proses ekstraksi dan pemuatan ke dalam *environment* komputasi, sehingga dilakukan eksklusi data untuk memastikan stabilitas proses pelatihan.

Penambahan data fase *nephrogenic* dan *portal venous* pada KiTS23 sangat krusial. Pada fase ini, ginjal normal tampak homogen, sehingga setiap ketidakteraturan (tumor atau kista) lebih mudah dideteksi, namun juga menuntut model untuk lebih presisi dalam membedakan jenis ketidakteraturan tersebut (Frank, Krishna, Schieda, & McGrath, 2025).

2.2.2 Hierarki Label dan Kelas Evaluasi

Sistem pelabelan pada KiTS23 berbeda dengan segmentasi biner biasa. Terdapat tiga label semantik utama dalam mask anotasi (*ground truth*):

1. Label 0: Latar Belakang (*Background*)
2. Label 1: Ginjal (*Kidney*) - mencakup parenkim normal.
4. Label 2: Tumor (*Tumor*) - massa neoplasma ganas.

5. Label 3: Kista (*Cyst*) - lesi cairan jinak.

Dalam proses evaluasi kompetisi, KiTS23 menggunakan konsep *Hierarchical Evaluation Classes* (HECs) untuk mengukur kinerja klinis yang lebih bermakna (Heller, Isensee, Trofimova, Tejpau, & al., 2023):

1. **Kidney and Masses:** Gabungan wilayah Ginjal + Tumor + Kista. Mengukur kemampuan model mendeteksi batas organ ginjal secara keseluruhan.
2. **Kidney Mass:** Gabungan wilayah Tumor + Kista. Mengukur kemampuan model mendeteksi adanya kelainan/lesi pada ginjal, tanpa mempedulikan jenisnya.
3. **Tumor Only:** Label Tumor saja. Ini adalah metrik yang paling kritis dan sulit, mengukur kemampuan model membedakan keganasan dari jaringan sehat dan kista jinak.

2.2.3 Protokol Pra-pemrosesan Data KiTS23

Mengingat data berasal dari berbagai institusi dengan pemindai (CT Scanner) yang berbeda, data mentah memiliki heterogenitas tinggi dalam hal resolusi spasial (*voxel spacing*) dan intensitas. Pra-pemrosesan yang ketat diperlukan untuk standarisasi input model:

1. **Resampling Spasial:** Data CT asli memiliki ketebalan irisan (*slice thickness*) yang bervariasi antara 1mm hingga 5mm. Model *Deep Learning* 3D memerlukan representasi spasial yang konsisten. Oleh karena itu, seluruh volume citra di-*resample* menjadi resolusi isotropik (misalnya $1.0 \times 1.0 \times 1.0$ mm atau sesuai median dataset) menggunakan interpolasi *spline* orde-3 untuk citra dan *nearest neighbor* untuk label (Isensee, Jaeger, Kohl, Petersen, & Maier-Hein, 2021).
2. **Windowing dan Clipping Intensitas:** Rentang HU pada data mentah sangat luas (-1000 hingga >3000). Untuk fokus pada jaringan ginjal, diterapkan *intensity clipping* pada rentang jaringan lunak, umumnya antara -150 HU (batas bawah lemak perirenal) hingga 250 HU (batas atas penyangatan kontras). Nilai di luar rentang ini dipotong, karena tidak mengandung informasi relevan untuk segmentasi tumor (Isensee F. , Jaeger, Kohl, Petersen, & Maier-Hein, 2021).

3. Normalisasi Z-Score: Setelah *clipping*, nilai intensitas dinormalisasi agar memiliki rata-rata 0 dan standar deviasi 1. Rumus normalisasi per citra (x) adalah:

$$x_{norm} = \frac{x - \mu}{\sigma}$$

Langkah ini krusial untuk mempercepat konvergensi algoritma optimasi (seperti Adam) selama pelatihan (Isensee F. , Jaeger, Kohl, Petersen, & Maier-Hein, 2021).

2.3 Landasan Teori: *Deep Learning* dan CNN

Sub-bab ini membahas fondasi matematis dari metode komputasi yang digunakan. *Deep Learning* adalah sub-bidang *Machine Learning* yang berbasis pada *Artificial Neural Networks* (ANN) dengan banyak lapisan tersembunyi (*hidden layers*), yang mampu mempelajari representasi fitur secara hierarkis langsung dari data mentah (*representation learning*) (Li & al., 2025).

Secara struktural, arsitektur dasar jaringan *Deep Learning* terdiri dari tiga bagian utama:

1. Lapisan Input (*Input Layer*): Lapisan awal yang menerima representasi data mentah. Pada penelitian ini, input berupa matriks nilai intensitas piksel dari citra CT Scan.
2. Lapisan Tersembunyi (*Hidden Layers*): Lapisan-lapisan di mana komputasi matematika utama terjadi. Semakin banyak lapisan ini, jaringan dianggap semakin "dalam" (*deep*), yang memungkinkannya mengenali pola abstrak dan kompleks dari data.
3. Lapisan Output (*Output Layer*): Lapisan akhir yang menghasilkan prediksi atau peta segmentasi.

Mekanisme Pengolahan Data (*Training Process*) Cara jaringan "belajar" dan mengolah data berlangsung melalui dua fase siklus matematis yang terus berulang:

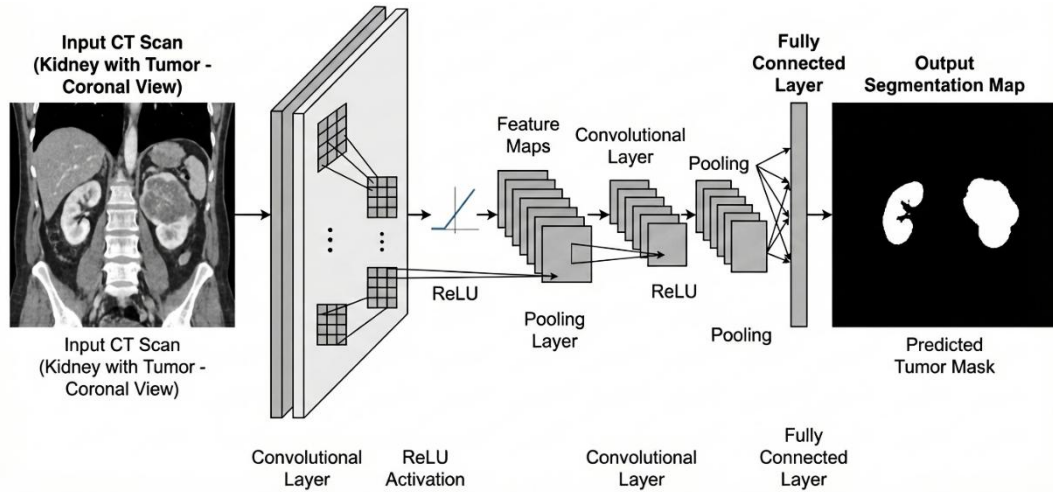
1. *Forward Propagation* (Perambatan Maju): Data input dialirkan masuk ke dalam jaringan. Pada setiap neuron, data dikalikan dengan matriks bobot (*weights*), ditambahkan dengan nilai bias, dan dilewatkan melalui fungsi aktivasi untuk menghasilkan output sementara. Hasil di lapisan output kemudian dievaluasi menggunakan fungsi kerugian (*Loss Function*) untuk

menghitung tingkat kesalahan (*error*) tebakan model dibandingkan dengan label asli (*Ground Truth*).

2. *Backpropagation* (Perambatan Balik): Setelah nilai *error* didapatkan, jaringan menghitung turunan parsial (gradien) dari *error* tersebut terhadap setiap parameter di dalam jaringan. Sinyal *error* ini dikirim kembali dari lapisan output mundur ke lapisan input. Algoritma optimasi kemudian memperbarui nilai bobot dan bias secara perlahan untuk meminimalkan *error* pada iterasi berikutnya.

2.3.1 Convolutional Neural Networks (CNN)

Untuk data berstruktur grid seperti citra medis, arsitektur *Multi-Layer Perceptron* (MLP) tidak efisien karena mengabaikan korelasi spasial antar piksel dan memiliki jumlah parameter yang meledak. CNN mengatasi hal ini dengan tiga prinsip utama: koneksi lokal (*local connectivity*), pembagian bobot (*parameter sharing*), dan ekuivarian terhadap translasi (Li & al., 2025).



Gambar 2.1 Ilustrasi Arsitektur *Convolutional Neural Network*

Inti dari CNN adalah operasi konvolusi. Secara matematis, untuk citra input 2D I dan kernel (filter) K berukuran $(2h + 1) \times (2w + 1)$, nilai output peta fitur (*feature map*) S pada posisi (i, j) dihitung sebagai konvolusi diskrit:

$$S(i, j) = (I * K)(i, j) = \sum_{m=-h}^h \sum_{n=-w}^w I(i - m, j - n) \cdot K(m, n)$$

Dalam konteks 3D (seperti pada KiTS23), kernel memiliki dimensi $D \times H \times W$, memungkinkan jaringan mempelajari fitur volumetrik antar-irisan (*inter-slice features*) (Li & al., 2025).

2.3.2 Fungsi Aktivasi Non-Linear

Kombinasi linier dari operasi konvolusi tidak cukup untuk memodelkan data kompleks. Fungsi aktivasi non-linear diperlukan untuk memberikan kapasitas aproksimasi fungsi universal pada jaringan. Fungsi yang paling dominan digunakan adalah ReLU (*Rectified Linear Unit*) (Li & al., 2025):

$$f(x) = \max(0, x)$$

ReLU memiliki keunggulan komputasi yang efisien dan mengurangi masalah *vanishing gradient* dibandingkan fungsi Sigmoid atau Tanh, karena gradiennya bernilai 1 untuk semua input positif. Varian lain seperti Leaky ReLU sering digunakan untuk mencegah masalah "dead neurons" dengan memberikan gradien kecil pada input negatif (Li & al., 2025).

2.4 Arsitektur U-Net

Arsitektur U-Net, yang diperkenalkan oleh (Ronneberger, Fischer, & Brox, 2015), dirancang khusus untuk segmentasi biomedis dengan data terbatas. Bentuknya yang menyerupai huruf "U" terdiri dari dua jalur simetris:

1. Jalur Kontraksi (*Encoder*): Berfungsi sebagai ekstraktor fitur. Terdiri dari blok konvolusi diikuti oleh *Max Pooling* untuk mengurangi resolusi spasial dan meningkatkan *receptive field*. Pada setiap langkah *pooling*, jumlah kanal fitur (*channels*) biasanya dilipatgandakan.
2. Jalur Ekspansi (*Decoder*): Berfungsi untuk lokalisasi presisi. Menggunakan operasi *Upsampling* (seperti *Transposed Convolution*) untuk mengembalikan resolusi spasial. Fitur dari decoder digabungkan dengan fitur dari encoder.
3. Skip Connections: Komponen kritis yang menghubungkan lapisan encoder dengan lapisan decoder pada resolusi yang sama melalui operasi konkatenasi (*concatenation*).

$$Input_{decoder} = [Feature_{encoder}, Upsampled(Feature_{decoder})]$$

Skip connection memungkinkan transfer informasi detail resolusi tinggi (seperti tepi tumor) yang biasanya hilang di bagian bottleneck jaringan, sehingga batas objek hasil segmentasi menjadi lebih presisi (Siddique, Paheding, Elkin, & Devabhaktuni, 2021).

Keterbatasan U-Net Standar pada KiTS23: Meskipun handal, U-Net standar memiliki kelemahan mendasar saat menghadapi dataset kompleks seperti KiTS23. Pertama, *skip connection* mentransfer fitur secara naif, termasuk fitur latar belakang (*noise*) yang tidak relevan. Kedua, kedalaman jaringan yang terbatas membuat U-Net kesulitan menangkap fitur semantik abstrak dari tumor heterogen yang berukuran besar (Azad, et al., 2024).

2.5 Mekanisme *Residual Learning*

Untuk meningkatkan kemampuan ekstraksi fitur tanpa terkendala masalah pelatihan pada jaringan yang dalam, penelitian ini mengadopsi konsep Residual Learning dari ResNet (Jin, Meng, Sun, Cui, & Su, 2020).

2.5.1 Masalah Degradasi dan Solusi Residual

Pada jaringan yang sangat dalam, akurasi seringkali jenuh dan kemudian menurun (*degradation problem*). (He, Zhang, Ren, & Sun, 2016) menunjukkan bahwa ini bukan disebabkan oleh *overfitting*, melainkan kesulitan optimasi. Mereka mengusulkan Residual Block yang menggunakan koneksi pintas (*shortcut connection*) identitas.

Alih-alih mempelajari pemetaan fungsi asli $H(x)$, lapisan dipaksa mempelajari fungsi residual $F(x) := H(x) - x$. Maka, fungsi asli direkonstruksi sebagai:

$$H(x) = F(x) + x$$

Operasi penjumlahan elemen $+x$ ini tidak menambah parameter baru namun secara signifikan mempermudah aliran gradien (Jin, et al., 2021).

2.5.2 Analisis Aliran Gradien (*Gradient Flow*)

Keunggulan matematis utama dari blok residual terletak pada propagasi balik (*backpropagation*). Gradien fungsi kerugian $\mathcal{L}$ terhadap input x_l dapat ditulis sebagai:

$$\frac{\partial \mathcal{L}}{\partial x_l} = \frac{\partial \mathcal{L}}{\partial x_L} \cdot \left(1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i) \right)$$

Angka 1 dalam persamaan di atas bertindak sebagai "jalan tol" yang menjamin gradien dapat mengalir dari lapisan akhir ke lapisan awal tanpa mengalami penyusutan (*decay*) akibat perkalian berulang. Hal ini memungkinkan pembangunan encoder yang jauh lebih dalam (misalnya ResNet-50 atau ResNet-101) untuk menangkap variasi morfologi tumor ginjal yang ekstrem di KiTS23 (Jin, et al., 2021).

2.6 Mekanisme *Attention Gates* (AG)

Untuk mengatasi masalah redundansi fitur pada skip connection, digunakan mekanisme Attention Gate (Jin et al., 2021). Pada citra CT abdomen, organ seperti hati, limpa, dan usus sering memiliki intensitas yang mirip dengan ginjal. *Skip connection* standar akan mengirimkan fitur organ-organ ini ke decoder, menyebabkan *false positive* (Jin Q. , 2021).

2.6.1 Prinsip Kerja *Attention Gate*

Attention Gate berfungsi sebagai filter adaptif yang menekan respons fitur pada area latar belakang dan memperkuat respons pada area target (tumor/ginjal) sebelum fitur digabungkan. Mekanisme ini meniru sistem visual manusia yang memfokuskan perhatian hanya pada bagian relevan dari pandangan (Oktay, 2018).

2.6.2 Formulasi Matematis

AG menerima dua input: vektor fitur dari encoder (x_l) dan vektor gating dari decoder (g). Sinyal gating g berasal dari lapisan yang lebih dalam, sehingga ia memiliki informasi kontekstual yang lebih baik tentang "di mana" objek berada, meskipun resolusinya rendah. Proses perhitungan koefisien atensi α melibatkan langkah-langkah berikut (Jin Q. , 2021):

1. Transformasi Linier: Input x_l dan g diproyeksikan ke ruang dimensi yang sama menggunakan konvolusi 1×1 (W_x dan W_g).
2. Kombinasi Aditif dan Aktivasi: Kedua sinyal dijumlahkan dan dilewatkan fungsi ReLU (σ_1):

$$q_{att} = \psi^T \left(\sigma_1(W_x x_l + W_g g + b_g) \right) + b_\psi$$

3. Normalisasi Sigmoid: Nilai dipetakan ke rentang $[0, 1]$ menggunakan Sigmoid (σ_2) untuk menghasilkan peta bobot atensi:

$$\alpha = \sigma_2(q_{att})$$

4. Pemberian Bobot (*Gating*): Fitur awal dikalikan dengan koefisien atensi secara *element-wise*:

$$\hat{x}_l = \alpha \odot x_l$$

Hasilnya ($\hat{x}_l$) adalah peta fitur yang "bersih", di mana area latar belakang memiliki nilai mendekati nol. Integrasi AG pada U-Net terbukti meningkatkan sensitivitas model tanpa menambah beban komputasi yang signifikan (Jin Q. , 2021).

2.7 Arsitektur Usulan: *Hybrid Residual-Attention U-Net*

Penelitian ini mengusulkan arsitektur Hybrid Residual-Attention U-Net yang mensinergikan ketiga konsep di atas untuk menangani tantangan spesifik dataset KiTS23.

2.7.1 Desain Arsitektur

1. Backbone Encoder: Menggunakan blok Residual (seperti ResNet-34) sebagai pengganti blok konvolusi standar. Ini memberikan kedalaman yang cukup untuk membedakan tekstur halus antara kista kompleks dan tumor nekrotik.
2. Modul Atensi: Setiap *skip connection* antara encoder dan decoder dilengkapi dengan *Attention Gate*. Ini memastikan decoder hanya menerima informasi spasial dari area ginjal, mengurangi risiko *over-segmentation* pada organ abdomen lain.
3. Decoder: Menggunakan *Transposed Convolution* untuk *upsampling*, diikuti oleh konvolusi standar untuk rekonstruksi mask segmentasi akhir (3 kanal: Background, Ginjal, Tumor).

Hipotesis utama adalah bahwa kombinasi ini akan memberikan keseimbangan optimal: Residual meningkatkan Recall (menemukan tumor), sementara Attention meningkatkan Precision (mengurangi alarm palsu) (Luo, 2022).

2.8 Metrik Evaluasi Kinerja

Mengingat ketidakseimbangan kelas (*class imbalance*) yang ekstrem pada dataset KiTS23 (di mana volume tumor seringkali $< 1\%$ dari total volume tubuh), penggunaan metrik akurasi standar menjadi tidak relevan. Oleh karena itu, evaluasi kinerja model dilakukan menggunakan tiga metrik utama berbasis tumpang tindih (*overlap-based*) dan analisis piksel (Heller, 2021).

2.8.1 Dice Similarity Coefficient (DSC)

DSC adalah metrik standar emas dalam segmentasi citra medis yang mengukur nilai rata-rata harmonik antara *precision* dan *recall* (Azad, 2025). Metrik ini menghitung seberapa besar area prediksi model tumpang tindih dengan *ground truth*. Rumusnya didefinisikan sebagai:

$$\text{DSC} = \frac{2 \times |A \cap B|}{|A| + |B|}$$

Di mana A adalah area prediksi model dan B adalah area ground truth. Nilai DSC berkisar antara 0 (gagal total) hingga 1 (segmentasi sempurna).

2.8.2 Intersection over Union (IoU)

Dikenal juga sebagai Jaccard Index, IoU mengukur rasio antara irisan (intersection) dan gabungan (union) dari area prediksi dan ground truth (Azad, 2025). IoU memberikan penalti yang lebih ketat terhadap kesalahan prediksi dibandingkan DSC.

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}$$

Hubungan matematisnya adalah $\text{IoU} = \text{DSC} / (2 - \text{DSC})$. Metrik ini sangat sensitif dalam mendeteksi kesalahan segmentasi pada batas objek yang kompleks.

2.8.3 Confusion Matrix (Pixel-Level)

Untuk melakukan analisis kesalahan yang lebih mendalam pada tingkat piksel, digunakan matriks kebingungan yang memetakan setiap piksel ke dalam empat kategori (Azad, 2025):

1. True Positive (TP): Piksel tumor yang diprediksi dengan benar sebagai tumor.

2. True Negative (TN): Piksel latar belakang yang diprediksi dengan benar sebagai latar belakang.
3. False Positive (FP): Piksel latar belakang yang salah diprediksi sebagai tumor (*over-segmentation*).
4. False Negative (FN): Piksel tumor yang salah diprediksi sebagai latar belakang (*under-segmentation*).

2.9 Kerangka Pemikiran

Berdasarkan tinjauan pustaka di atas, kerangka pemikiran penelitian ini dapat dirumuskan sebagai berikut:

1. Masalah: Segmentasi tumor ginjal pada CT Scan sulit karena variasi bentuk, tekstur, dan kemiripan dengan kista/organ lain (Schieda, 2025). Dataset terbaru KiTS23 menambah kompleksitas dengan variasi fase kontras dan label kista.
2. Gap: Model U-Net standar kurang dalam (fitur terbatas) dan *skip connection*-nya berisik (banyak *false positive*) karena mentransfer fitur latar belakang yang tidak relevan (Jin Q. , 2021). Selain itu, model canggih lain seringkali terlalu berat secara komputasi untuk implementasi praktis (Isensee F. , Jaeger, Kohl, Petersen, & Maier-Hein, 2021).
3. Solusi: Mengembangkan *Hybrid Residual-Attention U-Net*.
 - a *Residual Block* → Memperdalam jaringan → Fitur semantik lebih kuat (mengatasi variasi tekstur tumor) (He, Zhang, Ren, & Sun, 2016).
 - b *Attention Gate* → Memfilter fitur *skip connection* → Fokus spasial pada ROI (mengatasi *noise* latar belakang dan kista) (Oktay, 2018).
4. Validasi: Melatih dan menguji model pada dataset KiTS23 dengan metrik Dice Score dan IoU, serta membandingkannya dengan *baseline* U-Net standar untuk membuktikan efektivitas pendekatan hibrida tersebut (Heller, 2021).

Kerangka ini memberikan landasan logis yang kokoh bagi metodologi penelitian yang akan dijalankan, memastikan bahwa setiap komponen arsitektur memiliki justifikasi teoretis dan klinis yang kuat.