

## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Tinjauan Pustaka

Pada tahap awal penelitian, peneliti melakukan kajian mendalam terhadap berbagai studi terdahulu yang membahas penerapan metode deep learning dalam segmentasi ginjal, tumor, dan kista pada citra CT scan. Tinjauan ini bertujuan untuk memahami perkembangan terkini arsitektur segmentasi citra medis, mulai dari pendekatan konvensional berbasis U-Net hingga model yang lebih kompleks seperti ResUNet, Attention U-Net, dan variasi arsitektur berbasis backbone pretrained. Melalui kajian literatur tersebut, peneliti mengidentifikasi tantangan utama dalam segmentasi multi-kelas, khususnya pada objek tumor dan kista yang memiliki ukuran kecil, batas yang tidak jelas, serta ketidakseimbangan distribusi kelas pada dataset.

Selain itu, studi terdahulu juga menunjukkan keterbatasan beberapa model eksisting, seperti hilangnya detail spasial pada proses downsampling, kesulitan dalam mempertahankan konteks global dan lokal secara bersamaan, serta performa yang kurang optimal pada kelas minoritas. Dengan memahami kelebihan dan kekurangan pendekatan yang telah ada, peneliti memperoleh landasan teoritis yang kuat dalam merancang strategi segmentasi yang lebih efektif. Hasil tinjauan ini kemudian menjadi dasar dalam pemilihan arsitektur CE-ResUNet50-D, penentuan teknik augmentasi data, serta penggunaan fungsi loss yang adaptif, sehingga metode yang diusulkan diharapkan mampu mengatasi permasalahan yang ditemukan pada penelitian sebelumnya dan meningkatkan kualitas segmentasi ginjal, tumor, dan kista pada citra CT secara lebih optimal.

### 2.1.1 Penelitian 1

Tabel 2.1 Penelitian 1

Penelitian 1 [16]	
Title	Squeeze-and-Excitation Encoder-Decoder Network for Kidney and Kidney Tumor Segmentation in CT Images
Year	2021
Datasets	KiTS21
Classes	4 classes: Kidney, Tumor, Cyst, Background
Author	Jianhui Wen, Zhaopei Li, Zhiqiang Shen, Yaoyong Zheng, Shaohua Zheng
Research Gap	<ul style="list-style-type: none"><li>- Model U-Net dan variannya belum mampu menangkap long-range dependencies antar channel karena keterbatasan CNN.</li><li>- Segmentasi tumor ginjal sulit karena variasi bentuk, ukuran, dan posisi.</li><li>- Dataset KiTS21 tidak selalu memiliki label lengkap (beberapa hanya punya 3–4 kelas).</li><li>- Performa model sebelumnya pada lesi kecil (mass &amp; tumor) masih rendah.</li></ul>
Model	SeResUNet (U-Net encoder-decoder) dengan: <ul style="list-style-type: none"><li>- Encoder menggunakan SeResNet50</li><li>- Skip connection seperti U-Net</li><li>- Squeeze-and-Excitation modules untuk menangkap channel dependencies</li><li>- Deep Supervision di decoder</li><li>- Weighted Cross Entropy loss</li></ul>
Result	Dice Score (mean, 5-fold cross-validation): <ul style="list-style-type: none"><li>• Kidney: 91.60% , Masses: 58.80%, Tumor: 54.16%</li></ul> Surface Dice: <ul style="list-style-type: none"><li>• Kidney: 84.62%, Masses: 37.91%, Tumor: 37.59%</li></ul>

Penelitian pertama [16] melakukan segmentasi ginjal tumor dan kista menggunakan dataset Kits21 sebanyak 300 kasus. Studi ini tercipta karena keterbatasan arsitektur *U-Net* dan varian yang belum dapat menangkap *long-range channel dependencies* sehingga performa rendah pada lesi kecil seperti tumor dan

kista. Dengan memanfaatkan *SeResNet50* sebagai encoder, skip-connection ala *U-Net*, modul *Squeeze-and-Excitation*, serta *deep supervision*, model ini dapat melakukan peningkatan akurasi pada segmentasi *multiclass* meskipun hasil dice untuk tumor dan mass masih rendah.

### 2.1.2 Penelitian 2

Tabel 2.2 Penelitian 2

Penelitian 2 [17]	
Title	Kidney and Kidney Tumor Segmentation Using a Two-Stage Cascade Framework
Year	2021
Datasets	KiTS21
Classes	4 kelas: Background, Kidney, Tumor, Cyst
Author	Chaonan Lin, Rongda Fu, Shaohua Zheng
Research Gap	<ul style="list-style-type: none"> <li>- Segmentasi ginjal &amp; tumor sulit karena kontras rendah, batas organ tidak jelas, dan bentuk yang tidak konsisten.</li> <li>- Lesi kecil (tumor &amp; kista) memiliki variasi bentuk/ukuran sehingga model satu tahap sering gagal.</li> <li>- Ketidakseimbangan data antara organ besar (kidney) dan lesi kecil (tumor/kista).</li> <li>- Model sebelumnya (U-Net, cascaded networks) belum efektif menangani variasi spasial dan kontekstual secara bersamaan.</li> </ul>
Model	<p>Two-stage cascade framework:</p> <ol style="list-style-type: none"> <li>1) Kidney-Net → segmentasi ginjal awal (3D U-Net modifikasi).</li> <li>2) Masses-Net → segmentasi tumor &amp; kista menggunakan crop hasil tahap 1.</li> </ol> <p>+ Multi-Dimension Feature (MDF) module (3D + anisotropic conv)  + CBAM (Convolutional Block Attention Module)  + Deep Supervision pada decoder</p>
Result	<ul style="list-style-type: none"> <li>• Kidney: 0.9304</li> <li>• Tumor: 0.563</li> <li>• Masses (tumor+kista): 0.5729</li> </ul>

Penelitian kedua [17] menggunakan dataset KiTS21 dengan 300 kasus dan mengusulkan *two-stage cascade framework* untuk meningkatkan akurasi segmentasi pada lesi kecil. Tahap pertama, *Kidney-Net*, mengekstraksi ginjal, kemudian *Masses-Net* melakukan segmentasi tumor dan kista berdasarkan hasil crop tahap awal. Dengan tambahan *Multi-Dimension Feature module*, *CBAM*, dan *deep supervision*, model ini mampu meningkatkan akurasi segmentasi pada tumor dan kista, meskipun nilai Dice untuk keduanya masih sedang karena variasi spasial lesi yang tinggi.

### 2.1.3 Penelitian 3

Tabel 2.3 Penelitian 3

Penelitian 3[18]	
Title	Segmentation of kidney mass using AgDenseU-Net 2.5D model
Year	2022
Datasets	KiTS21
Classes	Kidney, Tumor, Cyst (Mass), dan Background
Author	P. Sun, Z. Mo, F. Hu, dan kolega.
Research Gap	<ul style="list-style-type: none"> <li>- Model 3D CNN terlalu berat secara komputasi dan membutuhkan sumber daya besar.</li> <li>- Banyak metode segmentasi belum efisien dalam hal konsumsi memori/GPU, padahal dataset seperti KiTS21 menuntut segmentasi ginjal + tumor + kista (multi-kelas).</li> <li>- Butuh metode yang seimbang antara efisiensi komputasi dan akurasi segmentasi multi-kelas.</li> </ul>
Model	AgDenseU-Net 2.5D — model hybrid: menggunakan arsitektur DenseU-Net yang diperkuat dengan attention gate, serta preprocess dengan automatic down-sampling 3D images untuk mengurangi ukuran volume CT sebelum segmentasi 2.5D
Result	<ul style="list-style-type: none"> <li>– Dice (HEC Kidney): 0.971</li> <li>– Dice (HEC Mass: tumor + cyst): 0.883</li> <li>– Dice (Tumor only): 0.815</li> </ul>

Penelitian ketiga [18] menggunakan model *AgDenseU-Net 2.5D* sebagai solusi untuk mengatasi akurasi yang rendah pada lesi yang kecil. Model ini memanfaatkan DenseU-Net yang diperkaya *attention gate* serta strategi *automatic down-sampling* untuk menekan konsumsi memori GPU. Pendekatan model ini memanfaatkan volumetric tanpa beban komputasi yang tinggi seperti model *3D CNN*. Hasil dari penelitian ini menunjukkan performa yang kompetitif, dengan nilai Dice tinggi pada *kidney* dan *masses* serta performa yang solid pada segmentasi tumor.

#### 2.1.4 Penelitian 4

Tabel 2.4 Penelitian 4

Penelitian 4 [19]	
Title	2.5D MFFAU-Net: a convolutional neural network for kidney segmentation
Year	2023
Datasets	KiTS19 (210 CT scans) dan KiTS21 (300 CT scans)
Classes	<ul style="list-style-type: none"> <li>- KiTS19: Kidney, Tumor</li> <li>- KiTS21: Kidney, Mass, Tumor, Cyst</li> </ul>
Author	Peng Sun, Zengnan Mo, Fangrong Hu, Xin Song, Taiping Mo, Bonan Yu, Yewei Zhang, Zhencheng Chen
Research Gap	<ul style="list-style-type: none"> <li>- 2D CNN tidak mampu menangkap fitur antar-slice (inter-slice).</li> <li>- 3D CNN membutuhkan GPU memory besar &amp; biaya komputasi tinggi.</li> <li>- Dibutuhkan solusi yang menggabungkan keunggulan 2D (ringan) dan 3D (memiliki informasi spasial).</li> <li>- Kebutuhan model yang mampu segmentasi kidney, tumor, dan cyst sekaligus secara efisien.</li> </ul>
Model	<p>2.5D MFFAU-Net:</p> <ul style="list-style-type: none"> <li>- Multi-level Feature Fusion Attention U-Net</li> <li>- Menggunakan ResConv (residual convolution)</li> <li>- Cascade model (coarse → fine segmentation)</li> <li>- Attention Gate pada skip-connection</li> <li>- Input multi-slice (2.5D)</li> </ul>
Result	<p>Hasil KiTS19:</p> <ul style="list-style-type: none"> <li>- Kidney Dice: 0.975, Tumor Dice: 0.872</li> </ul> <p>Hasil KiTS21:</p> <ul style="list-style-type: none"> <li>- Kidney Dice/SD: 0.973 / 0.941</li> <li>- Mass Dice/SD: 0.887 / 0.788</li> </ul>

	<ul style="list-style-type: none"> <li>- Tumor Dice/SD: 0.873 / 0.769</li> <li>- Cyst Dice/SD: 0.765 / 0.678</li> </ul>
--	---

Penelitian terakhir [19] menggunakan model 2.5D *MFFAU-Net* yang menggabungkan keunggulan 2D *CNN* dan 3D informasi spasial. Dengan memanfaatkan *multi-level feature fusion*, *ResConv*, *attention gate*, serta pendekatan *coarse-to-fine*, model ini dapat meningkatkan akurasi segementasi pada dataset Kits 19 dan Kits21. Penggunaan input multi slice membuat model ini dapat menangkap hubungan antar slice tanpa menggunakan 3D *CNN* untuk komputasi. Model ini menghasilkan dice tertinggi dibandingkan dengan penelitian lainnya terutama untuk segmentasi tumor, mass dan kista.

Secara keseluruhan, keempat penelitian sebelumnya menunjukkan perkembangan bertahap dalam meningkatkan performas segmentasi ginjal, tumor, kista pada citra CT. Penelitian 1 mengangkat pentingnya relasi antar channel, sementara penelitian 2 menakan penanganan spasial dan lesi kecil melalui arsitektur bertingkat. Penelitian 3 dan 4 menggali potensi pendekatan 2.5D yang lebih efisien namun tetap dapat menangkap informasi volumetric penting. Secara keseluruhan, semua penelitian menunjukkan bahwa penggunaan *attention*, model bertahap, serta kombinasi fitur yang lebih kaya mampu membantu meningkatkan kualitas segmentasi, terutama untuk lesi kecil seperti tumor dan kista. Pendekatan-pendekatan ini terbukti lebih efektif dalam menangani variasi bentuk dan ukuran organ, sehingga hasil segmentasi menjadi lebih stabil dan akurat.

## 2.2 Teori tentang topik penelitian

### 2.2.1 Chronic Kidney Disease (CKD)

Choronic Kidney Diases (CKD) adalah sebuah kondisi dimana kondisi ginjal menurun secara perlahan selama lebih dari tiga bulan. Ginjal yang seharusnya menyaring kotoran dan menjaga keseimbangan cairan tubuh menjadi tidak bekerja dengan baik, sehingga kotoran dan racun menumpuk dalam darah [20]. CKD umumnya terjadi karena diabetes, hipertensi infeksi

ginjal kronis dan faktor keturunan. Masalahnya, CKD sering kali tidak menunjukkan gejala pada tahap awal, sehingga banyak sekali pasien baru mengetahui ketika sudah memasuki tahap yang lebih lanjut dan sudah harus ditangani secara serius seperti spasial atau transplantasi ginjal [21].

Kelainan pada ginjal biasanya dapat berupa Tumor dan Kista. Tumor ginjal adalah pertumbuhan sel yang tidak normal pada jaringan ginjal. Tumor bisa bersifat jinak seperti angiomyolipoma atau onkositoma yang biasanya tumbuh secara lambat dan tidak berbahaya. Namun juga ada tumor ganas seperti *Renal Cell Carcinoma* (RCC) yang dapat menyebar secara cepat ke organ lain. Tumor sering kali sulit dideteksi pada tahap awal karena jarang sekali memperlihatkan gejala. Baru ketika ukuran tumor membesar, pasien bisa merasakan nyeri pinggang atau melihat darah dalam urin. Karena bentuk dan ukurannya bisa bervariasi [22].

Kista ginjal adalah kantung yang berisi cairan yang letaknya didalam ginjal. Kista sederhana biasanya tidak berbahaya dan sering kali ditemukan pada orang dewasa terutama usia lanjut. Namun terdapat kista yang kompleks memiliki struktur yang tidak teratur dan berpotensi menjadi ganas. Ada juga penyakit ginjal seperti polikistik (PKD) yaitu kondisi genetik yang menjadi pemicu kista tumbuh dan merusak jaringan ginjal yang dapat menyebabkan risiko gagal ginjal. Dalam pemrosesan citra medis, kista sulit sekali dibedakan dengan tumor karena keduanya berukuran kecil, berkontras rendah dan memiliki bentuk yang tidak seragam bahkan tidak jarang mirip [23].

### 2.2.2 Machine Learning

Machine learning telah membawa kemajuan besar dalam analisis citra medis karena kemampuannya dapat mempelajari representasi fitur secara bertingkat langsung dari data mentah tanpa perlu untuk ekstraksi fitur secara manual [24]. Pendekatan ini berbeda dengan *machine learning tradisional* yang sangat bergantung pada fitur buatan seperti tekstur atau histogram intensitas. Dengan adanya arsitektur jaringan yang lebih mendalam, machine learning

dapat memahami pola visual yang kompleks pada berbagai level, hal ini sangat dibuthkan dalam citra medis mengingat setiap pasien dapat memiliki bentuk, ukuran organ serta kondisi yang berbeda beda [25].

Keunggulan machine learning dalam tugas segmentasi medis adalah kemampuannya yang dapat mengangani variasi anatomi yang luas dan dapat mengidentifikasi batas objek secara presisi meskipun citra Ct Scan memiliki noise, perbedaan kontras serta variasi orientasi. Arsitektur berbasis *Convolutional Neural Networks* (CNN) sudah terbukti efektif untuk CT Scan serta MRI. Selain itu, penggunaan *transfer learning* semakin memperkuat performa model karena bobot pralatih seperti imangenet dapat mempercepat proses pelatihan dan meningkatkan akurasi terutama ketika dataset medis terbatas [26].

Dalam pengembangan model modern, berbagai arsitektur sudah digunakan untuk meningkatkan kualitas segmentasi termasuk model berbasis residual dan mekanisme anttention. Penelitian ini menggunakan arsitektur *CE-ResUNet50D*, dengan memanfaatkan decoder *ResNet-50* sebagai *backbone*. Desain ini menggabungkan kekuatasn antara residual learning untuk menjaga aliran infromasi antar lapisan, serta *channel enhancement* yang dapat membantu mempekuat fitur penting pada setiap tahap pemrosesan [27]. Dengan mekanisme ini, model dapat mempertahankan detail spasial, menangani variasi bentuk organ serta meningkatkan identifikasi lesi seperti tumor maupun kista. Pendekatan seperti ini sangat penting untuk segmentasi ginjal karena struktur ginjal yang kompleks serta adanya lesi yang sulit terlihat oleh CT Scan [28]. Penggunaan arsitektur *CE-ResUNet50D* menjadikan proses segmentasi lebih stabil, akurat, dan dapat diandalkan dalam aplikasi medis.

### 2.2.3 U-Net

U-Net adalah salah satu arsitektur deep learning yang paling berpengaruh dalam bidang segmentasi citra medis karena khsusu dirancang untuk mempertahankan detail spasial penting, bahkan ketika jumlah data pelatihan

terbatas [26]. Keunggulan utama U-Net terletak pada struktur encoder yang simetris, dimana bagian ini bertugas untuk mengecilkan resolusi citra agar dapat menangkap konteks global, sementara decoder bertugas untuk memulihkan redolusi dan menghasilkan segmentasi akhir [26]. Proses ini diperkuat dengan *skip connection* yang menghubungkan fitur dari encoder ke decoder pada tingkat resolusi yang sama [26]. Mekanisme penghubung ini sangat krusial karene berfungsi untuk mencegah hilangnya informasi spasial yang diakibatkan oleh *downsampling*, dengan itu model tetap dapat mengenali struktur kecil seperti tumor ataupun kista pada citra CT yang memiliki bentuk tidak beraturan dan berkontras rendah.

Pada bagian encoder, U-Net mengolah citra melalui beberapa rangkaian konvolusi berlapis dan operasi *downsampling* yang secara bertahap memperkecil ukuran fitur namun memperluas kedalaman representasi [14]. Tahapan ini membuat model dapat menangkap pola visual dan struktur anatomi yang lebih dalam termasuk memberdakan tekstur antara jaringan sehat dan jaringan patologis. Sementara itu, bagian decoder melakukan proses kebalikan dengan *upsampling* dan konvolusi untuk mengambilikan citra ke resolusi aslinya. Fitur beresolusi tinggi dari *skip connection* digabungkan kembali pada tahap ini agar model mampu mempertahankan tepi objek, batas organ serta fitur halus yang sering kali hilang jika hanya mengandalkan encoder [14]. Kombinasi ini membuat U-Net sangat efektif dalam segmentasi organ, jaringan maupun lesi kecil pada citra medis.

Seiring perkembangan teknologi, U-Net telah mengalami banyak sekali modifikasi untuk meningkatkan performas segmentasi, terutama pada area kompleks atau berukuran kecil. Salah satu pengembangan U-Net adalah CE-ResUNet50D, yang digunakan pada penelitian ini. Arsitektur ini tetap mengikuti kerangka kerja dasar dari U-Net, namun pada bagian encoder diperkuat menggunakan backbone *ResNet-50*, sebuah arsitektur residual yang terkenal cukup stabil dan dapat mempertahankan aliran informasi melalui *skip*

*connection internal* pada setiap blok [29]. Pendekatan residual ini mengatasi masalah *vanishing gradient* dan memungkinkan jaringan mempelajari fitur lebih dalam tanpa kehilangan detail penting. Selain itu, CE-ResUNet50D juga menambahkan modul Channel Enhancement (CE) atau *context encoder* yang berfungsi memperkuat fitur-fitur penting dan mengurangi informasi yang tidak relevan. Mekanisme ini membantu model lebih fokus pada area ginjal, tumor, serta kista, yang seringkali memiliki perbedaan tekstur halus dan batas yang tidak jelas pada CT scan [29].

Keunggulan lain dari U-Net adalah kemampuannya mempertahankan performa yang baik meskipun jumlah data terbatas. Hal ini dimungkinkan melalui penerapan *data augmentation* seperti rotasi, flipping, perubahan skala, hingga deformasi elastis yang dapat memperkaya variasi data tanpa perlu menambah dataset baru. Dalam konteks segmentasi ginjal, tumor, dan kista, teknik ini sangat membantu karena dataset medis umumnya memiliki variasi anatomi antar pasien yang tinggi, sementara jumlah sampel berlabel sering kali sedikit. Dengan memadukan arsitektur encoder-decoder U-Net, backbone ResNet-50, serta modul Channel Enhancement, CE-ResUNet50D mampu menghasilkan segmentasi yang lebih akurat, stabil, dan konsisten pada citra CT ginjal, meskipun menghadapi kondisi citra yang kompleks seperti noise, artefak, atau batas lesi yang samar.

#### 2.2.4 CeresUnet

CeresUnet pada dasarnya adalah sebuah arsitektur yang berfungsi untuk melakukan segmentasi citra medis berbasis U-Net yang sudah dikembangkan lebih lanjut melalui integrasi berbagai komponen modern seperti backbone residual dan modul peningkatan fitur (*feature enhancement*) [30]. Meskipun CeresUnet sangat jarang sekali ditemukan dipublikasi akademik sebagai arsitektur resmi, istilah ini umumnya merujuk pada model hibrida yang memadukan prinsip U-Net dengan teknik ekstraksi fitur yang mendalam serta

mekanisme peningkatan konteks [313]. Model dengan konsep serupa sudah banyak diteliti dalam penelitian modern seperti CE-Net, ResNet++ dan berbagai varian U-Net berbasis backbone CNN. Tujuan utama pendekatan ini adalah untuk meningkatkan keakuratan segmentasi pada citra medis yang memiliki karakteristik kompleks yang mencakup struktur organ yang tidak beraturan, batas lesi yang tidak jelas serta variasi intensitas yang sering kali muncul pada citra CT ginjal, tumor dan kista [313].

Secara Arsitektural, *CeresUnet* tetap berpegang pada prinsip seperti inti *U-Net* yaitu konfigurasi *encoder* yang simetris dengan *skip connection* pada setiap tingkat resolusi [11]. *Encoder* ini bertugas untuk mengekstrasi fitur bertahap melalui operasi konvolusi berlapis serta *downsampling*, sedangkan *decoder* melakukan *upsampling* untuk memulihkan resolusi spasial sambil menggabungkan fitur resolusi tinggi yang berasal dari *skip connection* [11]. Pada model klasik, bagian *encoder* *U-Net* menggunakan konvolusi standar, namun pada pengembangan model seperti *CeresUnet*, *encoder* sudah diganti dengan *backbone residual* yang lebih kuat seperti *Resnet-50*. Integrasi *backbone ResNet-50* memberikan keunggulan yang sangat signifikan dikarenakan *residual block* memiliki banyak sekali *short cut connection internal* yang dapat mempertahankan aliran gradien, sehingga membuat jaringan dapat dilatih lebih mendalam tanpa harus mengalami degradasi performa. Dengan demikian, bagian *encoder* pada *CereUnet* dapat membaca pola visual yang lebih kompleks dan mendalam termasuk pada tekstur halus yang dapat membedakan jaringan sehat maupun jaringan patologis [28].

Selain perbaikan pada bagian *encoder*, *CeresUnet* juga dilengkapi oleh modul tambahan seperti *Channel Enhancement* atau *Context Encoding*. Modul ini teradopsi dari arsitektur Ce-Net yang memperkenalkan mekanisme *context extractor*, seperti *Dense Atrous Convolution (DAC)* dan *Residual Multi-Kernel Pooling (RMP)* untuk memperluas jangkauan konteks serta memperkaya fitur [30]. Modul *enhancement* berfungsi untuk memperkuat fitur yang penting serta relevan dengan objek target untuk menekan fitur yang tidak memiliki kontribusi

signifikan terhadap segmentasi. Dengan adanya modul ini, model dapat mempertahankan informasi global serta lokal secara seimbang yang membuat model lebih efektif dalam mendeteksi struktur kompleks maupun lesi yang berukuran kecil [32].

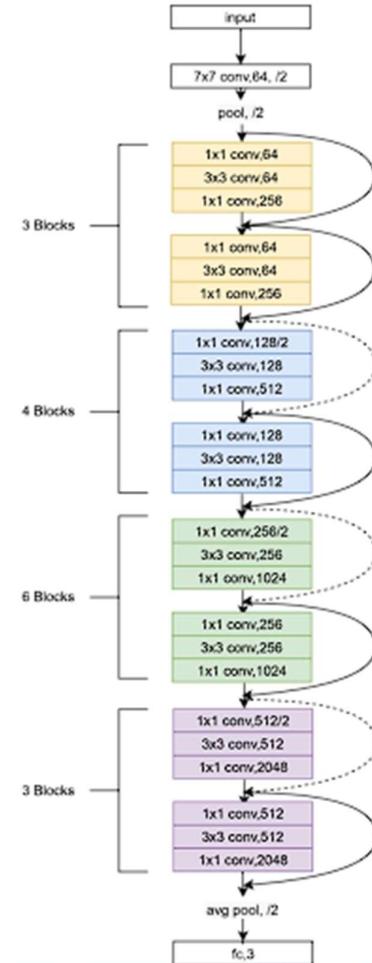
Secara keseluruhan, CeresUnet dapat dipahami sebagai pendekatan modern U-Net yang bertujuan untuk meningkatkan kualitas segmentasi melalui pemanfaatan backbone berkapasitas besar, mekanisme residual serta modul peningkatan konteks. Meskipun bukan arsitektur yang resmi dipublikasikan secara spesifik, konsep model ini selaras dengan berbagai penelitian terkini yang menekankan pentingnya penggabungan kemampuan ekstraksi fitur mendalam dan representasi konteks yang kaya untuk meningkatkan performa segmentasi pada dataset medis kompleks seperti CT scan ginjal, tumor dan kista. Dengan fondasi teori yang kuat dari CE-Net, ResNet, dan varian U-Net modern lainnya, pemahaman mengenai CeresUnet dapat dijelaskan secara ilmiah sebagai arsitektur U-Net yang telah diperkuat secara substansial untuk mencapai akurasi segmentasi yang lebih stabil, efisien, dan konsisten.

### 2.2.5 ResNet

ResNet adalah salah satu arsitektur *deep convolutional neural network* yang berperan penting dalam mengatasi permasalahan *vanishing gradient* pada jaringan yang sangat mendalam melalui penggunaan *residual* dan *shortcut connection* [33]. Mekanisme shortcut ini memungkinkan informasi dan gradien mengalir lebih stabil ke lapisan awal tanpa mengalami degradasi, sehingga model dapat mencapai kedalaman yang jauh lebih besar tanpa mengalami kesulitan dalam proses latihan [34]. Dalam arsitektur *CE-ResUNet50D* yang digunakan pada penelitian ini, backbone *ResNet-50* dipilih sebagai dasar encoder karena terdiri dari 50 lapisan konvolusi yang tersusun dalam bentuk *residual block*, menjadikannya mampu mengekstraksi fitur secara lebih kaya, mendalam, dan stabil [34].

Pada *CE-ResUNet50D*, *backbone ResNet-50* digunakan untuk memperkuat bagian encoder dari sisi struktur U-Net yang membuat proses ekstraksi fitur menjadi lebih efektif, terutama dalam membedakan tekstur dan pola halus pada citra CT ginjal [35]. *ResNet-50* terdiri dari empat tahap utama *residual block* yang masing masing mempunyai jumlah lapisan dan jumlah filter berbeda. Blok awal bertugas untuk menangkap fitur dasar seperti tepi dan tekstur sederhana, sementara blok yang lebih dalam berfungsi untuk mempelajari fitur tingkat tinggi seperti kontur organ, batas lesi dan pola patologis yang kompleks. Keberadaan shortcut connection pada setiap residual block membantu menjaga kestabilan aliran informasi, membuat model dapat mempelajari fitur dengan lebih akurat tanpa harus kehilangan detail penting [36].





Gambar 2.1 Arsitektur ResNet50 [36]

Gambar 2.1 menunjukkan alur kerja arsitektur ResNet-50 yang digunakan sebagai backbone utama dalam proses ekstraksi fitur pada model CE-ResUNet50D. ResNet-50 dipilih karena kemampuannya dalam mengekstraksi fitur secara efektif melalui mekanisme residual learning, sehingga pelatihan jaringan yang dalam dapat dilakukan dengan lebih stabil. Dalam arsitektur ini, ResNet-50 berperan sebagai encoder yang dikombinasikan dengan modul Channel Enhancement (CE), yang berfungsi untuk menekankan fitur penting serta menyaring informasi yang kurang relevan [11]. Modul CE memungkinkan model memperkuat konteks global dan lokal secara simultan, sehingga mampu

menangkap struktur anatomi ginjal serta mendeteksi lesi seperti tumor dan kista yang memiliki batas tidak jelas dan berukuran kecil [30].

Fitur-fitur yang telah diperkaya selanjutnya diteruskan ke bagian decoder U-Net melalui mekanisme skip connection, yang bertujuan untuk mempertahankan informasi spasial beresolusi tinggi selama proses rekonstruksi citra segmentasi [11]. Dengan mengombinasikan backbone ResNet-50, modul Channel Enhancement, dan arsitektur U-Net, CE-ResUNet50D menjadi model yang kuat untuk segmentasi ginjal, tumor, dan kista pada citra CT. Model ini mampu menghasilkan segmentasi yang lebih presisi dan konsisten, serta tetap andal meskipun dihadapkan pada tantangan seperti noise, artefak pencitraan, dan variasi anatomi antar pasien.

## 2.3 Teori tentang Tools dan Arsitektur yang digunakan

### 2.3.1 Python

Python merupakan bahasa pemrograman tingkat tinggi yang kini menjadi standar umum dalam pengembangan deep learning dan analisis citra medis. Dominasi Python tidak lepas dari sintaksnya yang ringkas, mudah dipahami, serta kemampuannya berintegrasi dengan berbagai pustaka dan platform komputasi modern. Dalam penelitian segmentasi citra medis, termasuk segmentasi ginjal dan tumor berbasis CT scan, Python menawarkan ekosistem yang sangat lengkap untuk mendukung seluruh alur kerja, mulai dari pengolahan data awal, tahap praproses, pembangunan model, hingga proses evaluasi akhir [37].

Salah satu alasan utama Python banyak digunakan adalah keberadaan pustaka komputasi numerik seperti NumPy dan SciPy yang mampu menangani manipulasi array multidimensi serta berbagai operasi matematika kompleks secara cepat dan efisien. Untuk kebutuhan visualisasi, library seperti Matplotlib dan Seaborn memberikan kemudahan dalam menampilkan grafik, citra, serta hasil segmentasi yang menjadi bagian penting dalam proses analisis

maupun verifikasi. Sedangkan untuk pengolahan citra, Python mempunyai OpenCV dan PIL atau Pillow yang memungkinkan pengguna melakukan konversi format, filtering, perubahan ukuran, augmentasi, dan analisis detail karakteristik citra medis.

Dalam pengembangan model deep learning, Python juga didukung oleh berbagai framework tingkat lanjut seperti TensorFlow, Keras, dan PyTorch. Ketiga framework ini mempermudah implementasi arsitektur kompleks seperti U-Net, serta integrasi dengan berbagai backbone CNN modern seperti EfficientNet, ResNet, MobileNet, VGG, dan DenseNet. Python juga didesain untuk bekerja optimal dengan komputasi GPU melalui teknologi CUDA dan cuDNN, yang sangat penting untuk mempercepat proses pelatihan model terutama ketika menangani dataset besar atau model tiga dimensi yang membutuhkan sumber daya komputasi tinggi [38].

Python turut didukung oleh komunitas open-source yang sangat aktif, sehingga dokumentasi, pembaruan pustaka, modul tambahan, dan solusi atas berbagai masalah teknis dapat ditemukan dengan mudah. Dukungan terhadap berbagai platform juga sangat luas, termasuk Google Colab yang menyediakan lingkungan komputasi gratis dengan akses GPU dan TPU, sehingga eksperimen deep learning dapat dijalankan tanpa memerlukan perangkat keras berperforma tinggi di sisi pengguna [39].

Dalam penelitian ini, seluruh proses mulai dari konversi data .nii.gz, tahap praproses citra, pembangunan model U-Net dengan berbagai backbone, evaluasi segmentasi, hingga visualisasi hasil, dilakukan sepenuhnya menggunakan Python. Dengan fleksibilitas, kemudahan penggunaan, serta dukungan pustaka yang sangat komprehensif, Python menjadi pilihan ideal untuk pengembangan sistem segmentasi citra medis berbasis deep learning, khususnya dalam upaya meningkatkan ketepatan deteksi dan analisis otomatis pada penyakit ginjal kronis

### 2.3.2 Numpy

NumPy atau Numerical Python adalah salah satu pustaka dasar yang sangat penting dalam ekosistem Python karena menyediakan dukungan array multidimensi serta berbagai fungsi matematika dan operasi linier yang dibutuhkan dalam komputasi ilmiah. Dalam penerapan segmentasi citra medis, NumPy memegang peran utama sebagai fondasi pengelolaan data citra, terutama karena sebagian besar data medis tersimpan dalam format array dua dimensi maupun tiga dimensi. Berbagai proses seperti mengambil irisan tertentu dari citra CT, menghitung statistik piksel seperti nilai rata-rata atau standar deviasi, hingga mengubah citra menjadi tensor numerik dilakukan dengan cepat dan efisien menggunakan NumPy [40].

Selain itu, NumPy memungkinkan manipulasi array berukuran besar melalui metode vektorisasi, yaitu pemrosesan tanpa perlu membuat perulangan secara eksplisit. Pendekatan ini membuat waktu eksekusi menjadi jauh lebih cepat dibandingkan pemrosesan berbasis Python murni. Dalam penelitian ini, NumPy digunakan secara intensif pada tahap prapemrosesan data seperti mengubah ukuran citra, melakukan normalisasi, menerapkan threshold, serta memproses mask segmentasi. NumPy juga digunakan pada tahap evaluasi model, termasuk perhitungan metrik seperti Dice Coefficient. Kecepatan dan fleksibilitas yang ditawarkan menjadikan NumPy sebagai komponen yang sangat penting dalam seluruh alur kerja pengembangan model deep learning untuk segmentasi ginjal, tumor dan kista [41].

### 2.3.3 Matplotlib

Matplotlib merupakan pustaka visualisasi dua dimensi dalam Python yang banyak digunakan untuk membuat berbagai jenis grafik, termasuk untuk menampilkan citra medis. Pada penelitian ini, Matplotlib dimanfaatkan untuk menampilkan citra asli, mask ground truth, serta hasil segmentasi yang diprediksi model dalam satu tampilan, sehingga proses evaluasi visual terhadap

kualitas segmentasi dapat dilakukan dengan lebih jelas. Visualisasi seperti ini sangat penting untuk melihat langsung kesalahan model, misalnya area yang diprediksi terlalu luas, kurang terdeteksi, atau kegagalan model dalam mengenali struktur ginjal maupun tumor yang berukuran kecil [42].

Selain menampilkan citra, Matplotlib juga digunakan untuk menggambarkan perkembangan proses pelatihan model, seperti grafik perubahan loss dan metrik akurasi pada setiap epoch. Informasi tersebut membantu peneliti dalam memantau kestabilan proses training serta mendeteksi munculnya overfitting atau underfitting sejak awal. Berbagai fungsi siap pakai seperti `imshow()`, `plot()`, `hist()`, dan `subplots()` membuat Matplotlib mampu menghasilkan visualisasi yang informatif dengan kode yang ringkas dan mudah disesuaikan. Karena itu, Matplotlib berperan penting tidak hanya sebagai alat dokumentasi hasil akhir, tetapi juga sebagai media eksplorasi dan debugging selama proses pengembangan model berlangsung [43].

#### 2.3.4 Tensorflow

TensorFlow merupakan framework open-source yang dikembangkan oleh Google dan banyak digunakan dalam pengembangan model deep learning berskala besar. Framework ini menyediakan berbagai fasilitas komputasi numerik yang sangat mendukung proses pembangunan, pelatihan, dan evaluasi model kecerdasan buatan, termasuk pada tugas segmentasi citra medis. Dalam penelitian ini, TensorFlow dimanfaatkan sebagai fondasi utama untuk merancang dan melatih arsitektur U-Net dengan beragam backbone convolutional neural network. Melalui modul `tf.keras`, TensorFlow menawarkan antarmuka tingkat tinggi yang membuat proses perancangan model menjadi lebih terstruktur dan mudah diintegrasikan dengan tahapan preprocessing, augmentasi, hingga evaluasi performa [44].

Salah satu keunggulan penting TensorFlow adalah kemampuannya menjalankan proses komputasi pada GPU melalui teknologi CUDA dan cuDNN. Hal ini memungkinkan pelatihan model berlangsung jauh lebih cepat, terutama ketika bekerja dengan dataset citra medis yang berukuran besar dan

memiliki dimensi kompleks. Selain itu, fitur seperti `tf.data.Dataset` membantu mengatur aliran data secara efisien, sedangkan komponen seperti `EarlyStopping` dan `ModelCheckpoint` digunakan untuk menghentikan pelatihan ketika performa berhenti meningkat serta menyimpan bobot model terbaik selama proses training.

TensorFlow juga memberikan fleksibilitas dalam pembuatan fungsi loss kustom, termasuk penerapan Dice Loss maupun kombinasi dengan Binary Crossentropy, yang sangat membantu dalam meningkatkan kualitas segmentasi pada kelas yang jarang muncul seperti tumor. Dengan dukungan dokumentasi yang komprehensif dan komunitas pengguna yang sangat aktif, TensorFlow menjadi platform yang kuat dan stabil untuk pengembangan sistem segmentasi medis modern, serta berpotensi untuk diadaptasi lebih lanjut dalam aplikasi klinis di masa mendatang [45].

### 2.3.5 Deep Learning (DL)

Deep learning merupakan salah satu cabang dari kecerdasan buatan yang memanfaatkan jaringan saraf tiruan dengan banyak lapisan (deep neural networks) untuk mempelajari pola, representasi, dan karakteristik data secara otomatis tanpa memerlukan proses rekayasa fitur secara manual. Pendekatan ini bekerja dengan mengekstraksi informasi secara hierarkis melalui serangkaian lapisan non-linear, di mana setiap lapisan bertugas mempelajari representasi fitur dari tingkat sederhana hingga kompleks. Kemampuan tersebut menjadikan deep learning sangat efektif dalam menangani data berskala besar dan berdimensi tinggi, termasuk citra medis yang memiliki variasi intensitas piksel, bentuk organ yang kompleks, serta perbedaan tekstur yang halus antar jaringan.

Dalam konteks segmentasi citra medis, deep learning, khususnya model berbasis Convolutional Neural Network (CNN), telah menunjukkan performa yang unggul dibandingkan metode konvensional. CNN mampu menangkap

informasi spasial dan kontekstual pada berbagai skala melalui operasi konvolusi dan pooling, sehingga sangat sesuai untuk mengidentifikasi batas organ dan lesi secara presisi. Salah satu arsitektur yang paling banyak digunakan dalam tugas segmentasi citra medis adalah U-Net, yang mengadopsi struktur encoder-decoder dengan skip connection. Mekanisme ini memungkinkan penggabungan fitur beresolusi tinggi dari encoder dengan fitur semantik dari decoder, sehingga detail spasial tetap terjaga selama proses segmentasi [24]. Dengan kemampuan pembelajaran end-to-end serta tingkat akurasi yang tinggi, deep learning kini menjadi pendekatan utama dalam pengembangan sistem segmentasi otomatis pada citra medis modern dan banyak diterapkan dalam berbagai penelitian klinis.

### 2.3.6 Environment

Framework deep learning yang paling banyak digunakan saat ini adalah PyTorch dan Keras. PyTorch dikenal memiliki fleksibilitas tinggi dan sering menjadi pilihan para peneliti karena memberikan kendali penuh terhadap desain arsitektur serta alur pemrosesan data. Framework ini sangat cocok untuk eksperimen yang sifatnya eksploratif dan untuk proses debugging yang memerlukan interaksi langsung. Sementara itu, Keras menawarkan antarmuka yang lebih sederhana dan mudah dipahami, sekaligus memiliki integrasi kuat dengan TensorFlow sehingga proses pembangunan dan pelatihan model dapat dilakukan dengan lebih cepat dan efisien [46].

Di sisi lain, Keras menawarkan antarmuka pemrograman tingkat tinggi yang lebih sederhana dan mudah dipahami, sehingga memungkinkan pengembangan model deep learning dilakukan dengan lebih cepat dan efisien. Keras memiliki integrasi yang sangat kuat dengan TensorFlow, yang menyediakan dukungan komputasi berskala besar, optimasi performa, serta kompatibilitas dengan berbagai perangkat keras seperti GPU dan TPU [46]. Pendekatan ini menjadikan Keras sangat sesuai untuk pengembangan model yang terstruktur dan berorientasi pada produktivitas.

Dalam penelitian ini, Keras dipilih sebagai framework utama karena menyediakan berbagai arsitektur dan backbone modern secara langsung melalui modul *keras.applications*, seperti U-Net, ResNet, dan CeresUnet. Setiap backbone tersebut telah dilengkapi dengan bobot pralatin yang memungkinkan proses *transfer learning* dilakukan secara efektif, sehingga mempercepat konvergensi model dan meningkatkan performa segmentasi. Selain itu, kompatibilitas backbone tersebut dengan arsitektur U-Net memudahkan proses integrasi, pelatihan, serta evaluasi model secara konsisten. Kemampuan Keras dalam mendukung pengujian dan perbandingan berbagai varian backbone secara sistematis menjadikannya pilihan yang tepat untuk penelitian segmentasi ginjal. Meskipun PyTorch tetap digunakan sebagai opsi pembanding, Keras memberikan keunggulan praktis dalam hal efisiensi pengembangan, stabilitas pelatihan, dan kemudahan eksperimen pada berbagai konfigurasi model.

### 2.3.7 Google Colab

Google Colaboratory, atau yang lebih dikenal sebagai Google Colab, merupakan platform komputasi berbasis cloud yang banyak dimanfaatkan dalam penelitian dan pengembangan model deep learning karena menyediakan lingkungan pemrograman Python yang siap digunakan tanpa memerlukan instalasi perangkat lunak secara lokal. Platform ini dilengkapi dengan dukungan perangkat keras berupa GPU dan TPU yang dapat diakses secara gratis, sehingga sangat membantu dalam menangani proses komputasi yang kompleks dan membutuhkan sumber daya besar. Dukungan ini menjadi sangat penting pada penelitian segmentasi citra medis, khususnya pada pemrosesan citra CT scan dan pelatihan model deep learning yang melibatkan jaringan saraf dengan jumlah parameter yang besar [47].

Dalam penelitian ini, Google Colab digunakan sebagai platform utama untuk seluruh tahapan eksperimen, mulai dari tahap prapemrosesan data,

visualisasi citra CT scan, pemilihan slice terbaik, hingga pembangunan dan pelatihan arsitektur CE-ResUNet50D dengan backbone ResNet-50. Selain itu, proses evaluasi model, perhitungan metrik performa seperti Dice Score dan IoU, serta visualisasi hasil segmentasi juga dilakukan secara langsung di lingkungan Colab. Integrasi yang erat antara Google Colab dan Google Drive memberikan kemudahan dalam pengelolaan dataset berukuran besar, memungkinkan penyimpanan, pembacaan, serta pembaruan data dilakukan secara efisien tanpa keterbatasan kapasitas penyimpanan lokal.

Keunggulan lain dari Google Colab adalah kemampuannya dalam mendukung pemasangan dan penggunaan berbagai pustaka pemrograman yang dibutuhkan dalam penelitian ini, seperti TensorFlow, PyTorch, MONAI, OpenCV, Albumentations, dan library pendukung lainnya. Dengan tersedianya GPU seperti Tesla T4 atau P100, proses pelatihan model CE-ResUNet50D dapat diselesaikan dalam waktu yang relatif lebih singkat, meskipun arsitektur yang digunakan bersifat kompleks dan melibatkan mekanisme attention. Selain itu, fleksibilitas Colab dalam menjalankan eksperimen berulang dengan konfigurasi yang berbeda tanpa memerlukan perangkat keras khusus menjadikannya platform yang sangat efektif, ekonomis, dan efisien dalam mendukung pengembangan serta pengujian model segmentasi ginjal, tumor, dan kista pada citra CT scan dalam penelitian ini.