

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Tabel 2.1 Penelitian Terdahulu

Referensi	Luo et al. (2024) [7]
Judul	<i>AI-based segmentation of renal enhanced CT images for quantitative evaluate of CKD</i>
Metode	3D nnU-Net
Ringkasan	Penelitian ini menyajikan pendekatan dua tahap yang menggabungkan segmentasi otomatis berbasis 3D nnU-Net dengan analisis kuantitatif fungsi ginjal pada pasien CKD. Data diperoleh dari 100 pasien CKD dan 90 individu sehat, dengan segmentasi dilakukan pada fase arteri dari CT ginjal. Model secara otomatis membedakan parenkim, korteks, dan medula ginjal. Volume serta nilai HU dari masing-masing bagian dikaitkan dengan nilai eGFR dan kreatinin serum (Scr). Korelasi tinggi ditemukan antara eGFR dan parameter volume maupun HU, serta korelasi negatif kuat antara Scr dan fitur-fitur ini. Hasil menunjukkan bahwa pendekatan ini dapat menggantikan pengukuran manual yang memakan waktu dalam menilai progresi CKD secara non-invasif dan terotomatisasi.
Hasil	Dice parenkim: 93.5%. Korelasi tinggi HU dan volume dengan eGFR ($p < 0.001$).
Fokus	Segmentasi otomatis ginjal dari CT untuk estimasi volume.
Referensi	Sharma et al. (2017) [8]
Judul	<i>Automatic Segmentation of Kidneys using Deep Learning for TKV Quantification in ADPKD</i>
Metode	Fully Convolutional Network (CNN)
Ringkasan	Fokus pada penyakit ginjal polistik (ADPKD), studi ini memanfaatkan deep learning untuk segmentasi otomatis ginjal dari citra CT aksial. Dataset mencakup 244 pasien dengan berbagai tingkat keparahan. CNN dilatih untuk melakukan segmentasi irisan demi irisan dan hasilnya dibandingkan dengan segmentasi manual oleh ahli. Tantangan utama adalah morfologi ginjal ADPKD yang sangat bervariasi karena kista. Studi ini menunjukkan bahwa metode otomatis dapat meniru akurasi pakar klinis secara konsisten dan mempercepat proses estimasi volume ginjal total (TKV), yang merupakan biomarker penting dalam menilai progresi penyakit.
Hasil	Dice 0.86 ± 0.07 , korelasi volume manual vs otomatis $\rho = 0.98$.
Fokus	Segmentasi ginjal untuk menghitung TKV secara efisien.
Referensi	Hussain et al. (2021) [9]
Judul	<i>Cascaded Regression Neural Nets for Kidney Volume Estimation</i>
Metode	Swin UNETR, ResU-Net, 3D U-Net
Ringkasan	Artikel ini memperkenalkan pendekatan segmentasi-bebas yang mengandalkan regresi untuk memperkirakan volume ginjal dari CT, tanpa membuat segmentasi eksplisit. Metode dimulai dengan estimasi jangkauan ginjal (inferior-superior) menggunakan CNN, diikuti oleh Mask-RCNN untuk bounding box, dan akhirnya FCN untuk estimasi volume langsung. Model diuji pada dataset VGH dan KiTS19. Hasil menunjukkan akurasi sebanding dengan pendekatan segmentasi modern, dengan arsitektur jauh lebih ringan. Kontribusi signifikan termasuk pengurangan beban komputasi dan waktu inferensi yang cepat.
Hasil	Error volume ~5%, error batas ginjal ~2.4 mm.

Fokus	Estimasi volume ginjal tanpa segmentasi eksplisit.
Referensi	Onthoni et al. (2020) [10]
Judul	<i>Deep Learning Assisted Localization of Polycystic Kidney on Contrast-Enhanced CT Images</i>
Metode	SSD-InceptionV2
Ringkasan	Deteksi otomatis ginjal pada pasien ADPKD dalam CT kontras dengan model Single Shot Detector (SSD) InceptionV2. Model dilatih pada 10.078 irisan dari 110 pasien, mengatasi tantangan segmentasi yang ditimbulkan oleh kista multipel, kista hati, dan densitas yang tumpang tindih. Meskipun bukan segmentasi penuh, model ini menyediakan lokalisasi yang dapat mempercepat tahap segmentasi berikutnya dan penghitungan TKV. Hasil sangat baik dalam metrik presisi rata-rata, menjadikannya langkah penting dalam sistem segmentasi ginjal berbasis AI.
Hasil	mAP 94% (image-wise), 82% (subject-wise).
Fokus	Deteksi lokasi ginjal ADPKD untuk mendukung segmentasi.
Referensi	Lin et al. (2021) [11]
Judul	<i>Automated segmentation of kidney and renal mass in CT urography using 3D U-Net</i>
Metode	3D U-Net
Ringkasan	Artikel ini menerapkan arsitektur 3D U-Net untuk mendeteksi dan mensegmentasi ginjal serta massa (tumor dan kista) dari CT urography (CTU). Data dari 882 ginjal digunakan untuk pelatihan dan evaluasi. Model dilatih dengan kombinasi augmentasi data dan validasi silang, dan menunjukkan hasil yang mendekati segmentasi manual. Keluaran ini dapat digunakan untuk analisis kuantitatif, pelacakan massa renal, atau mendukung sistem diagnosis berbantuan komputer (CAD).
Hasil	Dice 0.973 (ginjal), 0.844 (massa).
Fokus	Segmentasi ginjal dan massa renal dari CTU.
Referensi	Pandey and Gupta (2022) [4]
Judul	<i>Tumorous kidney segmentation in abdominal CT using active contour + 3D U-Net</i>
Metode	3D U-Net + Active Contour
Ringkasan	Menggabungkan pendekatan tradisional berbasis kontur aktif dengan 3D U-Net untuk segmentasi ginjal dengan tumor pada CT abdomen. Kontur aktif meningkatkan keakuratan batas awal, sedangkan U-Net menyempurnakan prediksi volume. Gabungan dua pendekatan ini memungkinkan peningkatan dalam mendeteksi morfologi abnormal. Model diuji pada data KiTS dan menunjukkan kinerja tinggi, relevan untuk perencanaan bedah dan evaluasi massa ginjal.
Hasil	Dice 97.6%, Jaccard 95.7%.
Fokus	Segmentasi ginjal dan tumor dari CT abdomen.
Referensi	Fu et al. (2021) [13]
Judul	<i>Deep Learning Based CT Imaging in Quantitative Evaluation of CKD</i>
Metode	RDA-U-Net (ResDualAttention U-Net)
Ringkasan	Studi ini memperkenalkan varian U-Net yang dilengkapi blok attention dan residual untuk meningkatkan akurasi segmentasi ginjal pada pasien CKD. Volume dan HU dari parenkim, korteks, dan medula dihitung dari CT dan dikorelasikan dengan stadium penyakit. Model ini mempercepat dan meningkatkan evaluasi kuantitatif pasien dengan CKD secara non-invasif, memberikan solusi alternatif terhadap pengukuran laboratorium konvensional.
Hasil	Dice ginjal kiri: 96.25%, kanan: 94.22%.
Fokus	Segmentasi ginjal dari CT untuk evaluasi kuantitatif CKD.

Referensi	Zhao et al. (2020) [14]
Judul	<i>MSS U-Net: 3D segmentation of kidneys and tumors from CT images</i>
Metode	MSS U-Net (multi-scale supervision)
Ringkasan	Makalah ini memperkenalkan arsitektur U-Net berbasis 3D dengan multi-scale supervision dan loss fungsi eksponensial-logaritmik untuk segmentasi ginjal dan tumor dari CT. Post-processing dilakukan untuk meningkatkan konsistensi spasial hasil prediksi. Diuji pada dataset KiTS19, model ini memberikan kinerja sangat tinggi baik dalam segmentasi organ normal maupun lesi.
Hasil	Dice ginjal: 0.969, tumor: 0.805.
Fokus	Segmentasi presisi tinggi ginjal dan tumor ginjal.
Referensi	Zang et al. (2022) [15]
Judul	<i>3D U-Net Based Semantic Segmentation of Kidneys and Renal Masses</i>
Metode	3D U-Net + deep supervision
Ringkasan	Fokus pada segmentasi semantik ginjal dan massa renal dalam dataset KiTS21, menggunakan modifikasi arsitektur 3D U-Net dengan pelatihan berlapis dan pengawasan mendalam. Studi ini menunjukkan kemampuan model untuk membedakan jaringan normal dari lesi kecil hingga besar, dan menyarankan penggunaannya dalam sistem analisis citra otomatis berbasis volume tinggi.
Hasil	Dice ginjal: 0.970, massa: 0.851, tumor: 0.819.
Fokus	Segmentasi semantik ginjal dan massa dari CT kontras.
Referensi	Hu et al. (2024) [16]
Judul	<i>STC-U-Net: Renal tumor segmentation based on enhanced feature extraction</i>
Metode	STC-U-Net (varian U-Net)
Ringkasan	STC-U-Net dirancang untuk meningkatkan segmentasi tumor ginjal dengan menambahkan blok ekstraksi fitur spasial dan lintas skala. Model ini mampu menangani ukuran dan bentuk tumor yang sangat bervariasi dalam data CT ginjal. Diuji pada dataset KiTS19, arsitektur ini menghasilkan prediksi presisi tinggi dan dinilai cocok untuk diterapkan dalam klinik dengan kompleksitas anatomi yang tinggi.
Hasil	Dice ~96.9%
Fokus	Segmentasi tumor ginjal menggunakan arsitektur U-Net terbaru.

Dalam beberapa tahun terakhir, pendekatan berbasis deep learning telah menjadi solusi utama dalam segmentasi ginjal dari citra CT, menggantikan metode manual yang memakan waktu. Penelitian [7] mengusulkan model 3D nnU-Net untuk segmentasi otomatis parenkim, korteks, dan medula ginjal pada pasien penyakit ginjal kronis (CKD). Hasilnya menunjukkan akurasi tinggi (Dice 93,5%) serta korelasi kuat antara parameter volume dan intensitas Hounsfield Unit (HU) dengan estimasi laju filtrasi glomerulus (eGFR). Studi serupa [13] memperkenalkan RDA-U-Net, varian U-Net yang diperkuat dengan mekanisme perhatian (attention), menghasilkan Dice score 96,25% untuk ginjal kiri. Kedua penelitian ini menekankan potensi segmentasi otomatis dalam evaluasi kuantitatif CKD.

Sementara itu, pada kasus penyakit ginjal polikistik autosomal dominan (ADPKD), [8] menggunakan Fully Convolutional Network (FCN) untuk segmentasi slice-by-slice, mencapai korelasi hampir sempurna ($\rho = 0,98$) antara volume otomatis dan manual. Namun, tantangan utama adalah variasi morfologi ekstrem akibat kista. Untuk mengatasinya, [10] mengembangkan pendekatan berbasis SSD-InceptionV2 guna mendeteksi lokasi ginjal ADPKD terlebih dahulu sebelum segmentasi, dengan presisi (mAP) mencapai 94%. Hal ini mempercepat proses estimasi Total Kidney Volume (TKV), biomarker kunci dalam pemantauan ADPKD.

Di sisi lain, beberapa penelitian berfokus pada optimasi komputasi dengan mengurangi kebutuhan segmentasi penuh. [9] mengusulkan pendekatan berbasis regresi langsung menggunakan Swin UNETR dan Mask R-CNN, yang mampu memperkirakan volume ginjal hanya dengan error $\sim 5\%$ tanpa melakukan segmentasi detail. Metode ini lebih cepat namun kurang presisi untuk deteksi lesi kecil. Sebaliknya, [11] dan [4] mempertahankan segmentasi 3D konvensional dengan 3D U-Net, yang dikombinasikan dengan active contour untuk meningkatkan akurasi batas tumor. Pendekatan ini menghasilkan Dice score sangat tinggi (97,6% untuk ginjal normal) dan berguna dalam perencanaan bedah.

Perkembangan terbaru dalam arsitektur segmentasi mencakup peningkatan ekstraksi fitur multi-skala. [14] memperkenalkan MSS U-Net dengan multi-scale supervision, menghasilkan Dice 0,969 untuk ginjal dan 0,805 untuk tumor. [15] juga menggunakan 3D U-Net dengan deep supervision, berfokus pada segmentasi semantik ginjal dan massa renal, sementara [16] mengusulkan STC-U-Net untuk ekstraksi fitur spasial yang lebih baik, mencapai Dice $\sim 96,9\%$ pada dataset KiTS19.

Secara keseluruhan, penelitian-penelitian ini menunjukkan bahwa segmentasi berbasis deep learning telah mencapai tingkat akurasi yang setara dengan ahli radiologi, terutama untuk ginjal normal. Namun, tantangan tetap ada dalam kasus ADPKD dan tumor kecil, di mana variasi bentuk dan intensitas memerlukan pendekatan yang lebih canggih. Ke depan, integrasi dengan radiomics dan

pembelajaran semi-supervised dapat menjadi solusi untuk meningkatkan generalisasi model dengan data terbatas.

Berbeda dari penelitian-penelitian sebelumnya yang umumnya mengusulkan arsitektur baru atau menyesuaikan model terhadap jenis penyakit ginjal tertentu, penelitian ini menghadirkan kebaruan dalam pendekatannya dengan melakukan perbandingan sistematis berbagai backbone dalam kerangka arsitektur U-Net yang sama. Tujuan dari pendekatan ini adalah untuk mengevaluasi secara objektif performa masing-masing backbone seperti EfficientNet, VGG, MobileNet, ResNet, dan DenseNet dalam konteks segmentasi ginjal menggunakan data citra CT. Dengan menjaga struktur decoder U-Net tetap konsisten, evaluasi dapat difokuskan secara spesifik pada efektivitas ekstraksi fitur dari masing-masing backbone. Hasil dari penelitian ini diharapkan dapat memberikan panduan yang lebih jelas mengenai backbone yang paling optimal untuk tugas segmentasi ginjal, sekaligus menjawab kebutuhan akan model yang akurat namun efisien untuk implementasi klinis. Pendekatan ini juga melengkapi studi-studi terdahulu dengan memberikan wawasan teknis mendalam mengenai kontribusi komponen encoder dalam sistem segmentasi berbasis U-Net.

2.2 Teori tentang Topik Penelitian

2.2.1 Chronic Kidney Disease (CKD)

Chronic Kidney Disease (CKD) merupakan kondisi gangguan fungsi ginjal yang bersifat progresif dan ireversibel, dengan durasi lebih dari 3 bulan, yang ditandai oleh penurunan fungsi ekskresi ginjal atau adanya kelainan struktural ginjal yang terdeteksi melalui pencitraan atau biomarker [17]. CKD telah menjadi masalah kesehatan global yang signifikan, dengan prevalensi yang terus meningkat, terutama akibat faktor risiko seperti diabetes melitus, hipertensi, dan penyakit kardiovaskular [18]. Menurut laporan dari Kidney Disease: Improving Global Outcomes (KDIGO) 2021, CKD diklasifikasikan berdasarkan dua parameter utama, yaitu laju filtrasi glomerulus (GFR) dan tingkat albuminuria, yang membagi penyakit ini menjadi lima stadium klinis [17].

Pada stadium awal CKD (stadium 1-2), seringkali tidak ditemukan gejala klinis yang spesifik, sehingga banyak kasus tidak terdiagnosis pada tahap ini. Namun, kerusakan ginjal dapat mulai terdeteksi melalui pemeriksaan laboratorium, seperti adanya proteinuria, atau melalui pencitraan medis yang menunjukkan kelainan struktural ginjal, seperti perubahan ukuran, ketidakteraturan korteks, atau munculnya lesi [19]. Deteksi dini pada stadium ini sangat krusial, mengingat intervensi yang tepat dapat memperlambat progresivitas penyakit dan menunda terjadinya gagal ginjal stadium akhir (End-Stage Renal Disease/ESRD).

Pada stadium lanjut (stadium 3-5), terjadi penurunan fungsi ginjal yang signifikan, dengan GFR menurun di bawah 60 mL/min/1.73m². Kondisi ini dapat disertai gangguan elektrolit, anemia, penumpukan toksin dalam darah, serta peningkatan risiko komplikasi sistemik seperti penyakit jantung dan gangguan metabolik [18]. Tanpa intervensi, CKD dapat berkembang menjadi ESRD, yang memerlukan terapi pengganti ginjal seperti dialisis atau transplantasi ginjal.

Salah satu tantangan utama dalam manajemen CKD adalah memantau perubahan struktur ginjal secara akurat dan konsisten. Dalam hal ini, pencitraan medis seperti Computed Tomography (CT) scan berperan penting dalam mengevaluasi morfologi ginjal, termasuk deteksi atrofi ginjal, penipisan korteks, perubahan volume ginjal, hingga identifikasi lesi atau tumor [19]. Namun, interpretasi citra CT secara manual sangat bergantung pada keahlian radiolog, memakan waktu, serta rentan terhadap variabilitas antar pengamat.

Oleh karena itu, pengembangan teknik segmentasi otomatis berbasis kecerdasan buatan, khususnya deep learning, menjadi solusi yang semakin relevan dalam praktik klinis. Segmentasi otomatis memungkinkan pemisahan struktur ginjal dan lesi secara objektif dan reproducible, sehingga mempermudah kuantifikasi volume ginjal, analisis perubahan struktural, serta deteksi dini abnormalitas seperti fibrosis atau massa tumor [20]. Dengan demikian, sistem segmentasi otomatis tidak hanya meningkatkan efisiensi dan akurasi diagnosis, tetapi juga menjadi bagian integral dalam sistem pendukung keputusan klinis untuk penanganan CKD secara komprehensif.

Lebih lanjut, penerapan segmentasi otomatis berbasis deep learning pada dataset pencitraan medis seperti KiTS19, yang berisi citra CT ginjal beserta label segmentasi manual dari para ahli, memberikan peluang besar untuk meningkatkan kapabilitas deteksi dan analisis struktural ginjal secara presisi. Hal ini diharapkan dapat mempercepat inovasi dalam diagnosis, monitoring progresi CKD, hingga perencanaan intervensi medis yang lebih tepat sasaran.

2.2.2 Deep Learning

Deep learning telah merevolusi bidang analisis citra medis melalui kemampuannya dalam mempelajari representasi fitur secara hierarkis langsung dari data mentah tanpa memerlukan proses ekstraksi fitur manual [21]. Berbeda dengan pendekatan machine learning tradisional yang mengandalkan fitur hand-crafted seperti tekstur, tepi, atau histogram intensitas, deep learning mampu secara otomatis mempelajari pola kompleks dari data dalam berbagai skala dan kedalaman [22]. Hal ini sangat penting dalam citra medis, yang umumnya memiliki karakteristik bervariasi antar individu, termasuk perbedaan ukuran organ, bentuk, struktur internal, hingga tingkat keparahan penyakit.

Keunggulan utama deep learning untuk aplikasi segmentasi medis terletak pada kemampuannya menangani variasi anatomi yang luas dan mengidentifikasi batas objek dengan presisi sub-piksel [23]. Berbagai studi telah membuktikan bahwa arsitektur deep learning, khususnya Convolutional Neural Networks (CNN), mampu mengatasi tantangan seperti noise, variasi kontras, maupun perbedaan orientasi citra yang sering dijumpai pada data medis seperti CT scan atau MRI.

Selain itu, deep learning memungkinkan penerapan transfer learning, yaitu pemanfaatan bobot prelatih dari dataset besar seperti ImageNet untuk mempercepat pelatihan dan meningkatkan performa pada dataset medis yang biasanya terbatas [24]. Transfer learning ini telah menjadi praktik umum dalam pengembangan model deep learning untuk segmentasi medis, termasuk dalam studi-studi terkait segmentasi ginjal dan tumor ginjal.

Arsitektur deep learning modern yang paling banyak digunakan untuk segmentasi medis adalah U-Net, yang menawarkan desain encoder-decoder simetris dengan mekanisme skip connection untuk mempertahankan informasi spasial resolusi tinggi [25]. Selain itu, arsitektur berbasis residual seperti ResNet atau dense connectivity seperti DenseNet semakin populer karena kemampuannya menjaga aliran informasi antar-lapisan serta meningkatkan efisiensi parameter. Tak hanya itu, inovasi terbaru seperti Transformer-based models juga mulai diadopsi dalam segmentasi medis, memanfaatkan kemampuan self-attention untuk menangkap dependensi global antar piksel, yang sebelumnya sulit dicapai dengan CNN konvensional [26].

Dalam konteks penyakit ginjal, segmentasi organ dan tumor secara presisi sangat krusial untuk diagnosis, perencanaan perawatan, hingga monitoring perkembangan penyakit seperti Chronic Kidney Disease (CKD) atau kanker ginjal [27]. Deep learning telah terbukti meningkatkan akurasi dalam mendeteksi dan memetakan struktur ginjal, korteks, medula, serta lesi tumor, bahkan pada kondisi dengan tingkat kerusakan struktural yang signifikan. Oleh karena itu, pengembangan model deep learning yang tepat, disertai pemilihan backbone arsitektur yang sesuai, menjadi kunci untuk mencapai segmentasi yang akurat dan dapat diandalkan dalam aplikasi klinis.

2.2.3 U-Net

U-Net merupakan salah satu arsitektur deep learning yang paling populer dan banyak digunakan untuk tugas segmentasi citra medis, khususnya segmentasi organ, jaringan, maupun kelainan patologis pada gambar biomedis [28]. Arsitektur ini pertama kali diperkenalkan dengan desain khusus untuk dapat bekerja efektif meskipun pada kondisi jumlah data yang terbatas, seperti yang umum terjadi pada dataset citra medis [29]. Keunggulan utama U-Net terletak pada struktur encoder-decoder simetris yang dilengkapi dengan mekanisme skip connection, yang menghubungkan fitur dari bagian encoder ke decoder pada resolusi yang sama. Skip connection ini berfungsi untuk mempertahankan informasi spasial beresolusi tinggi yang seringkali hilang akibat proses downsampling selama tahap encoding [29].

Pada bagian encoder, U-Net menggunakan serangkaian blok konvolusi dan operasi downsampling (biasanya max-pooling) untuk mengekstraksi fitur hierarkis dari citra masukan. Setiap blok akan memperdalam representasi fitur, namun mengurangi resolusi spasialnya. Sementara itu, bagian decoder bertujuan untuk merekonstruksi citra ke ukuran semula melalui upsampling, disertai dengan penggabungan fitur resolusi tinggi dari skip connection [30]. Dengan mekanisme ini, U-Net mampu mempertahankan detail spasial yang penting, terutama untuk segmentasi objek berukuran kecil seperti tumor, yang seringkali sulit dideteksi oleh arsitektur biasa.

Selain itu, salah satu keunggulan U-Net adalah kemampuannya beradaptasi dengan kondisi data terbatas melalui strategi data augmentation elastis, seperti rotasi, flipping, scaling, dan deformasi elastis. Teknik ini terbukti meningkatkan generalisasi model tanpa memerlukan dataset besar [30]. Oleh karena itu, U-Net sangat relevan untuk diterapkan dalam bidang segmentasi citra CT scan ginjal, yang sering kali memiliki keterbatasan jumlah data berlabel.

Perkembangan terbaru dari U-Net juga menunjukkan banyak modifikasi yang ditujukan untuk meningkatkan performa segmentasi, terutama pada area-area kritis yang kecil atau sulit dikenali. Salah satu varian yang terkenal adalah Attention U-Net, yang mengintegrasikan mekanisme attention untuk meningkatkan fokus model pada region of interest (ROI) dan mengurangi prediksi keliru pada area background [31]. Selain itu, adaptasi backbone encoder U-Net dengan arsitektur CNN modern seperti EfficientNet, ResNet, MobileNet, atau DenseNet, terbukti mampu meningkatkan kapabilitas ekstraksi fitur mendalam, memperbaiki akurasi segmentasi, dan mengoptimalkan efisiensi komputasi.

Dalam penelitian ini, U-Net digunakan sebagai kerangka utama untuk tugas segmentasi ginjal dan tumor ginjal berdasarkan citra CT scan dari dataset KiTS19. Berbagai pilihan backbone CNN diterapkan pada bagian encoder U-Net untuk mengevaluasi pengaruh arsitektur tersebut terhadap akurasi segmentasi. Pemilihan backbone mencakup arsitektur dengan karakteristik berbeda, mulai dari model ringan seperti MobileNetV2, model klasik seperti VGG19, hingga model canggih

seperti EfficientNetB7, ResNet50, dan DenseNet121. Dengan variasi ini, dapat dilakukan analisis menyeluruh terkait trade-off antara kompleksitas model, efisiensi komputasi, dan akurasi segmentasi dalam konteks citra medis ginjal.

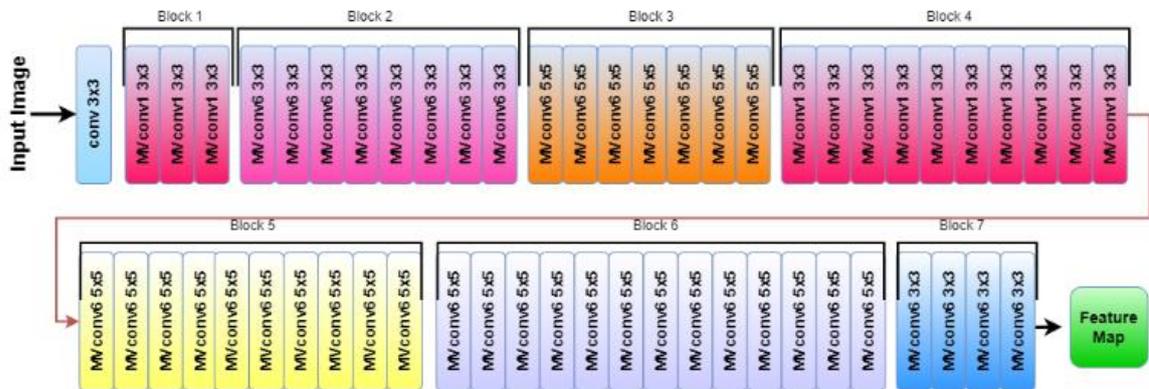
2.2.3.1 EfficientNet

EfficientNet merupakan arsitektur CNN yang dikembangkan dengan prinsip compound scaling, yang bertujuan untuk menyeimbangkan peningkatan skala pada depth (kedalaman), width (lebar), dan resolution (resolusi) secara proporsional dalam satu rumus sederhana [32]. Pendekatan ini menghindari peningkatan parameter secara berlebihan di salah satu aspek saja, sehingga menghasilkan model yang lebih efisien dan optimal. EfficientNet menggunakan blok Mobile Inverted Bottleneck Convolution (MBConv) yang dikombinasikan dengan Squeeze-and-Excitation (SE) block untuk meningkatkan selektivitas kanal dan mengurangi beban komputasi [33].

Dalam penelitian ini, digunakan EfficientNetB7, yaitu varian tertinggi dalam keluarga EfficientNet yang menawarkan kapasitas besar untuk menangkap detail fitur yang kompleks. Meskipun varian ini menuntut sumber daya komputasi yang cukup besar, pemilihannya didasarkan pada harapan dapat memberikan segmentasi tumor ginjal yang lebih akurat, terutama dalam mendeteksi area tumor yang kecil dan sulit dibedakan.

Selain itu, arsitektur ini dianggap mampu menjaga trade-off yang optimal antara akurasi dan efisiensi komputasi dalam konteks segmentasi medis [34]. Arsitektur EfficientNetB7 memungkinkan model untuk bekerja pada resolusi yang lebih tinggi dengan tetap mempertahankan efisiensi parameter, sehingga cocok untuk memproses citra CT scan dengan tingkat detail tinggi.

UNIVERSITAS
MULTIMEDIA
NUSANTARA



Gambar 2.1 Arsitektur EfficientNetB7 [40]

Berdasarkan pada gambar 2.1, arsitektur EfficientNetB7 merupakan salah satu model CNN modern yang dirancang dengan fokus pada efisiensi dan akurasi tinggi. Model ini mengadopsi pendekatan compound scaling, yaitu skema penskalaan yang seimbang antara kedalaman jaringan (*depth*), lebar jaringan (*width*), dan resolusi input, sehingga menghasilkan kinerja yang lebih baik dibandingkan pendekatan scaling tradisional yang hanya fokus pada satu aspek saja. Dalam struktur arsitekturnya, EfficientNetB7 menggunakan unit konvolusi yang disebut Mobile Inverted Bottleneck Convolution (MConv), yang menggabungkan beberapa teknik efisiensi modern seperti *depthwise separable convolution*, *expansion-projection layers*, dan *Squeeze-and-Excitation (SE) modules*.

Model diawali dengan sebuah lapisan konvolusi awal (Conv 3×3) yang bertugas mengekstraksi fitur dasar dari citra masukan. Setelah itu, input diproses melalui serangkaian blok yang terdiri dari berbagai kombinasi MConv dengan ukuran kernel yang berbeda (3×3 dan 5×5) dan faktor ekspansi yang bervariasi. Pada Block 1, digunakan MConv dengan kernel 3×3 dan faktor ekspansi 1, menandakan bahwa channel input tidak diperluas sebelum dilakukan konvolusi. Ini cocok untuk tahap awal karena menjaga efisiensi dan meminimalkan beban komputasi. Selanjutnya, pada Block 2 dan Block 3, digunakan MConv dengan faktor ekspansi 6, masing-masing dengan kernel 3×3 dan 5×5. Di sini, jumlah channel diperluas terlebih dahulu sebelum dilakukan *depthwise convolution*, kemudian diproyeksikan kembali ke dimensi yang lebih rendah. Strategi ini memungkinkan model

menangkap fitur yang lebih kompleks tanpa meningkatkan jumlah parameter secara drastis. Ukuran kernel 5×5 yang muncul pada Block 3 memungkinkan model untuk melihat area yang lebih luas dalam gambar, yang penting untuk memahami konteks spasial global.

Block 4 kembali menggunakan MBConv 3×3 , namun dengan jumlah lapisan yang lebih banyak. Yang menarik, terdapat panah koneksi shortcut dari Block 4 ke Block 5, menandakan adanya skip connection yang mengalirkan informasi langsung ke bagian lebih dalam dari jaringan. Ini berguna untuk mencegah hilangnya informasi penting selama pelatihan dan menjaga stabilitas aliran gradien. Block 5 dan Block 6 menggunakan MBConv 5×5 dan berperan dalam menangkap fitur tingkat tinggi dari representasi visual yang sudah lebih kompleks. Kedua blok ini memiliki jumlah unit yang banyak, menunjukkan bahwa di tahap ini model belajar memahami hubungan spasial yang lebih abstrak dan luas. Di bagian akhir, Block 7 kembali menggunakan MBConv dengan kernel 3×3 sebagai transisi menuju lapisan output. Pada tahap ini, dimensi spasial dari fitur map sudah kecil, namun kedalaman (jumlah channel) cukup besar, menyimpan informasi yang sangat kaya untuk menghasilkan prediksi akhir. Output dari jaringan berupa feature map, yaitu representasi fitur dari citra masukan yang bisa digunakan untuk berbagai tugas, seperti klasifikasi, segmentasi, atau deteksi objek.

Meskipun EfficientNetB7 menawarkan performa tinggi secara teoretis, kompleksitas arsitekturnya membuat model ini membutuhkan sumber daya komputasi besar dan jumlah data pelatihan yang cukup banyak agar dapat berfungsi secara optimal. Dalam kondisi di mana data terbatas atau perangkat keras tidak memadai, model ini bisa mengalami kesulitan konvergensi dan bahkan overfitting. Oleh karena itu, meskipun sangat kuat, penerapannya perlu dipertimbangkan dengan matang sesuai konteks penggunaan dan kemampuan sistem.

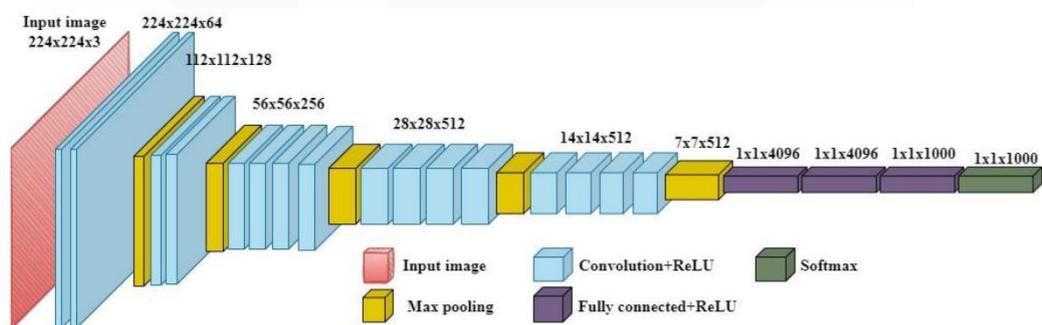
2.2.3.2 VGG (Visual Geometry Group Network)

VGG adalah salah satu arsitektur CNN paling berpengaruh yang dikembangkan oleh Visual Geometry Group, University of Oxford, dengan desain yang sederhana namun efektif. VGG memperkenalkan penggunaan konvolusi berukuran kecil

(3x3) secara berurutan dan mendalam, yang menghasilkan kemampuan ekstraksi fitur yang kuat [37]. VGG terdiri dari dua varian utama, yaitu VGG16 dan VGG19, yang berbeda pada jumlah lapisan konvolusional.

Dalam penelitian ini digunakan VGG19, yang memiliki 19 lapisan dengan pola stacking sederhana yang mempermudah analisis dan penerapan pada berbagai dataset, termasuk citra medis. VGG19 mampu mengekstraksi fitur spasial dan tekstural dari citra ginjal secara mendalam, meskipun tidak seefisien arsitektur modern dalam hal parameter dan komputasi. Pemilihan VGG19 bertujuan untuk menguji performa arsitektur konvensional dalam menangani segmentasi ginjal dan tumor, serta membandingkan kemampuannya dengan arsitektur yang lebih kompleks.

VGG19 tetap menjadi acuan penting dalam literatur computer vision karena stabilitasnya dan kemudahan penerapan pada berbagai jenis dataset.



Gambar 2.2 Arsitektur VGG19 [39]

Arsitektur VGG19, seperti ditunjukkan pada Gambar 2.2, merupakan salah satu model CNN klasik yang terkenal karena desainnya yang sederhana namun efektif dalam tugas ekstraksi fitur visual. Model ini menerima citra masukan berukuran 224×224×3 dan memprosesnya melalui serangkaian lapisan konvolusi kecil berukuran 3×3 yang diulang secara mendalam. Penggunaan kernel kecil ini memungkinkan model menangkap detail spasial dengan presisi tinggi tanpa meningkatkan jumlah parameter secara drastis. Struktur VGG19 terdiri dari lima blok utama, masing-masing diakhiri dengan operasi max pooling yang berfungsi mengurangi dimensi spasial secara progresif, yaitu dari 224×224 menjadi 112×112,

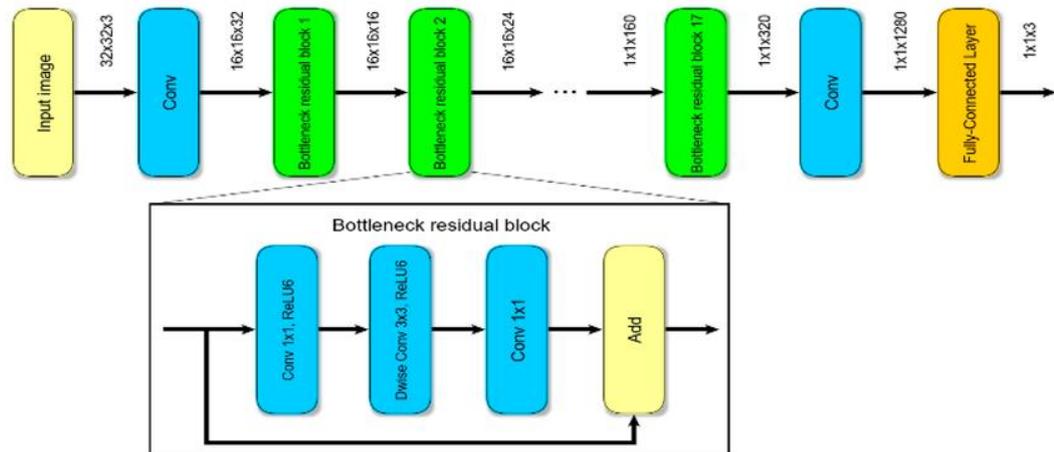
lalu 56×56 , 28×28 , 14×14 , hingga mencapai ukuran 7×7 dengan kedalaman fitur mencapai 512 channel. Setelah tahap ekstraksi fitur selesai, output diubah menjadi vektor satu dimensi melalui proses flattening, kemudian diproses oleh tiga lapisan fully connected, dua di antaranya memiliki 4096 neuron, dan lapisan akhir yang biasanya menyesuaikan jumlah kelas target. Aktivasi ReLU digunakan pada seluruh lapisan untuk meningkatkan non-linearitas, sedangkan Softmax diaplikasikan pada output untuk klasifikasi. Dalam penelitian ini, VGG19 diimplementasikan sebagai backbone U-Net untuk segmentasi ginjal dan tumor, dengan tujuan mengevaluasi performa arsitektur konvensional dalam mempelajari struktur anatomis ginjal yang kompleks. Meskipun VGG19 dikenal memiliki keterbatasan efisiensi dibandingkan model modern seperti ResNet atau DenseNet, arsitektur ini tetap menjadi referensi penting di bidang computer vision, khususnya dalam segmentasi medis, berkat kestabilannya dan kemampuan ekstraksi fitur mendalam yang sistematis.

2.2.3.3 MobileNet

MobileNet merupakan arsitektur CNN ringan yang dirancang untuk efisiensi tinggi dan optimalisasi penggunaan sumber daya [35]. MobileNet mengadopsi depthwise separable convolution, yaitu pemisahan proses konvolusi spasial dan konvolusi kanal, yang secara signifikan mengurangi jumlah parameter dan beban komputasi tanpa mengorbankan akurasi secara drastis.

Dalam penelitian ini, digunakan MobileNetV2 sebagai backbone U-Net, dengan pertimbangan bahwa model ini menawarkan keseimbangan antara akurasi dan efisiensi yang memungkinkan penerapan pada sistem dengan keterbatasan komputasi, seperti perangkat mobile atau sistem real-time. MobileNetV2 juga memperkenalkan inverted residual block yang mempercepat inferensi dan menjaga integritas fitur pada lapisan terdalam.

Pemilihan MobileNetV2 dimaksudkan untuk menguji sejauh mana model ringan mampu menghasilkan segmentasi ginjal dan tumor yang memadai, serta memberikan solusi yang dapat diintegrasikan ke sistem medis berbasis komputasi terbatas.



Gambar 2.3 Arsitektur MobileNetV2 [38]

Arsitektur MobileNetV2, sebagaimana ditunjukkan pada Gambar 2.3, dirancang untuk menghasilkan efisiensi tinggi dalam ekstraksi fitur visual, khususnya pada perangkat dengan keterbatasan sumber daya komputasi. Model ini menerima citra masukan berukuran $32 \times 32 \times 3$ yang kemudian diproses melalui lapisan konvolusi awal berukuran 3×3 untuk mengekstraksi fitur dasar. Struktur utama MobileNetV2 terdiri dari rangkaian blok yang disebut Bottleneck Residual Block, yang menggabungkan prinsip inverted residual dan depthwise separable convolution. Setiap Bottleneck Block diawali dengan lapisan konvolusi 1×1 yang berfungsi memperluas dimensi channel (expansion), diikuti depthwise convolution 3×3 yang memproses setiap channel secara independen, serta konvolusi 1×1 untuk proyeksi kembali ke dimensi awal. Ciri khas penting adalah adanya koneksi shortcut (residual connection) yang menghubungkan input dengan output blok, menjaga aliran informasi dan stabilitas gradien, sehingga mempermudah pelatihan model meskipun arsitektur dibuat ringan. Setelah melewati 17 bottleneck block, fitur diproses oleh lapisan konvolusi akhir dan dilanjutkan dengan lapisan fully connected yang menghasilkan output klasifikasi atau prediksi sesuai kebutuhan tugas. Dengan pendekatan ini, MobileNetV2 mampu mempertahankan keseimbangan optimal antara efisiensi komputasi dan akurasi, menjadikannya pilihan tepat untuk aplikasi seperti segmentasi medis di sistem dengan keterbatasan perangkat keras. Pada penelitian ini, MobileNetV2 digunakan sebagai backbone U-

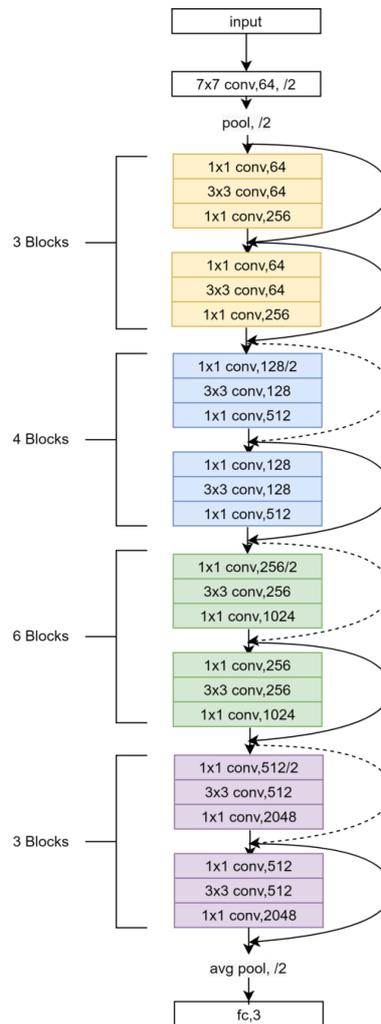
Net untuk segmentasi ginjal dan tumor, guna menguji sejauh mana model ringan dapat memberikan segmentasi presisi tinggi pada citra CT ginjal.

2.2.3.4 ResNet

ResNet atau Residual Network merupakan terobosan penting dalam pengembangan CNN yang mengatasi masalah vanishing gradient pada jaringan yang sangat dalam melalui penerapan residual (shortcut) connection [36]. Shortcut ini memungkinkan aliran gradien berjalan langsung ke lapisan awal tanpa tereduksi secara signifikan, sehingga memungkinkan arsitektur berisi ratusan lapisan tetap dapat dilatih dengan baik.

Dalam penelitian ini digunakan ResNet50, yang memiliki 50 lapisan dengan struktur residual block yang mendukung ekstraksi fitur secara progresif dan stabil. ResNet50 dipilih karena menawarkan kombinasi yang seimbang antara kedalaman, efisiensi komputasi, dan stabilitas pelatihan. ResNet efektif dalam mengenali pola kompleks pada citra medis, termasuk perbedaan halus antara jaringan ginjal dan tumor.

Pemanfaatan ResNet50 diharapkan mampu meningkatkan akurasi segmentasi dengan mempertahankan ketajaman fitur dan mendukung proses pembelajaran yang lebih optimal dibandingkan arsitektur yang lebih dangkal.



Gambar 2.4 Arsitektur ResNet50 [41]

Arsitektur deep convolutional neural network yang mengimplementasikan konsep residual learning untuk mengatasi masalah vanishing gradient yang sering terjadi pada jaringan yang sangat dalam. Model ini terdiri dari total 50 lapisan yang dibangun menggunakan blok residual, sebagaimana ditunjukkan dalam Gambar 2.4. Arsitektur diawali dengan lapisan konvolusi 7×7 dengan 64 filter dan stride 2, diikuti oleh operasi max pooling untuk mengurangi dimensi spasial.

Struktur utama ResNet50 terbagi menjadi empat tahap utama yang masing-masing terdiri dari beberapa residual block. Setiap block residual terdiri dari tiga lapisan konvolusi berurutan, yakni 1×1 convolution untuk penyesuaian dimensi channel, 3×3 convolution sebagai inti ekstraksi fitur, dan 1×1 convolution untuk

ekspansi atau penyesuaian dimensi output. Pada tahap awal (warna kuning), terdapat 3 blok dengan 64 filter, dilanjutkan tahap kedua (warna biru) dengan 4 blok berisi 128 filter, tahap ketiga (warna hijau) dengan 6 blok berisi 256 filter, dan tahap keempat (warna ungu) dengan 3 blok yang memiliki 512 filter.

Ciri khas utama arsitektur ini adalah adanya shortcut connection (garis putus-putus) yang menghubungkan input awal dengan output akhir dari setiap residual block. Shortcut ini memungkinkan informasi dan gradien mengalir secara langsung tanpa degradasi, sehingga mempercepat konvergensi dan meningkatkan stabilitas pelatihan.

Setelah melewati semua blok residual, dilakukan global average pooling untuk mereduksi dimensi spasial, sebelum akhirnya dilanjutkan ke fully connected layer sebagai tahap akhir klasifikasi atau prediksi. Dalam konteks penelitian ini, ResNet50 digunakan sebagai backbone U-Net untuk meningkatkan kemampuan ekstraksi fitur dalam tugas segmentasi ginjal dan tumor, memanfaatkan kedalaman arsitektur dan stabilitas aliran informasi yang ditawarkan oleh residual connection. ResNet50 dikenal efektif dalam memahami pola spasial yang kompleks, sehingga sangat sesuai untuk analisis citra medis seperti CT ginjal.

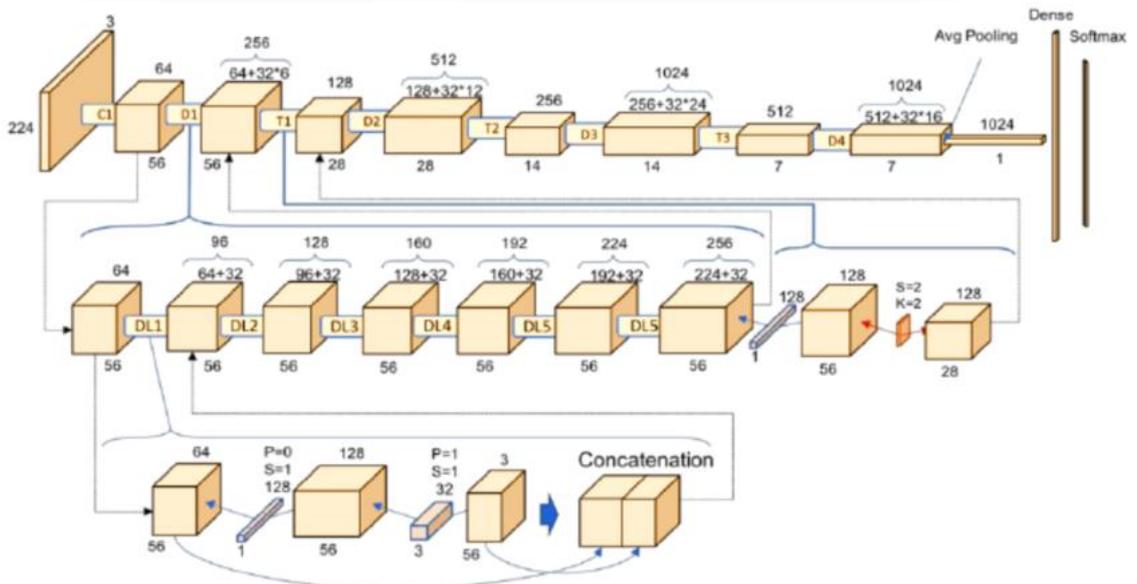
2.2.3.5 DenseNet

DenseNet adalah arsitektur CNN yang memanfaatkan konektivitas dense (padat), di mana setiap lapisan terhubung ke seluruh lapisan sebelumnya [35]. Pola koneksi ini memperkuat aliran informasi dan gradien, memungkinkan reuse fitur secara maksimal, dan mendorong efisiensi parameter. DenseNet dikenal memiliki performa yang kuat dalam mendeteksi pola halus dan menjaga detail spasial secara konsisten.

Dalam penelitian ini digunakan DenseNet121, yang dianggap memiliki struktur ringan namun tetap cukup dalam untuk mempelajari fitur penting pada citra ginjal dengan kerusakan akibat CKD. DenseNet121 dipilih sebagai backbone U-Net karena kemampuannya dalam mempertahankan informasi lokal dan global secara

efektif serta mendukung inferensi cepat dengan jumlah parameter yang relatif kecil dibandingkan arsitektur lain dengan kedalaman serupa.

DenseNet121 menjadi alternatif menarik dalam penelitian ini karena kemampuannya memberikan akurasi segmentasi yang baik dengan konsumsi memori yang lebih efisien, serta menghasilkan performa yang unggul dalam mendeteksi area tumor yang kecil dan sulit dibedakan.



Gambar 2.5 Arsitektur DenseNet121 [42]

Pada bagian awal arsitektur, citra masukan dengan resolusi $224 \times 224 \times 3$ diproses melalui lapisan konvolusi awal (C1) yang dilanjutkan dengan operasi downsampling (D1) untuk mereduksi dimensi spasial. Setelah itu, model memasuki serangkaian Dense Block (T1, T2, T3, T4) yang masing-masing berisi beberapa lapisan konvolusi yang saling terhubung secara dense. Di antara setiap Dense Block terdapat Transition Layer (D2, D3, D4) yang berfungsi mengontrol ukuran feature map dan jumlah channel, melalui proses konvolusi dan pooling.

Di bagian bawah arsitektur, terlihat detail struktur Dense Block, di mana setiap blok memiliki beberapa layer konvolusi kecil yang saling terhubung. Setiap layer tidak hanya menerima output dari layer sebelumnya, tetapi juga dari seluruh layer di blok tersebut. Dengan demikian, jaringan mampu memanfaatkan fitur dari

berbagai tingkat abstraksi sekaligus, meningkatkan akurasi tanpa harus menambah banyak parameter.

Di bagian akhir jaringan, dilakukan average pooling global, yang kemudian dilanjutkan dengan lapisan fully connected (dense) untuk menghasilkan output akhir berupa prediksi. Dalam DenseNet121, jumlah total lapisan mencapai 121, yang terdiri dari kombinasi konvolusi, pooling, dan dense layer.

Kelebihan utama DenseNet121 terletak pada efisiensi memori dan kecepatan inferensi, meskipun memiliki kedalaman jaringan yang besar. Selain itu, struktur dense connectivity sangat efektif untuk mempertahankan detail spasial, sehingga sangat bermanfaat dalam tugas-tugas segmentasi citra medis, termasuk segmentasi ginjal dan tumor, di mana informasi spasial yang detail sangat penting. DenseNet121 menawarkan keseimbangan yang baik antara kompleksitas model, jumlah parameter, dan akurasi, menjadikannya backbone yang ideal dalam implementasi U-Net untuk tugas segmentasi berbasis deep learning.

2.3 Teori tentang Tools dan Arsitektur yang digunakan

2.3.1 Python

Python merupakan bahasa pemrograman tingkat tinggi yang telah menjadi standar *de facto* dalam pengembangan aplikasi deep learning dan analisis citra medis dalam satu dekade terakhir. Popularitas Python dalam bidang ini tidak terlepas dari sintaksisnya yang sederhana, mudah dibaca, serta kemampuan integrasi yang kuat dengan berbagai pustaka dan platform komputasi modern. Dalam konteks penelitian segmentasi citra medis, seperti segmentasi ginjal dan tumor berbasis CT scan, Python menyediakan ekosistem yang sangat lengkap dan fleksibel untuk seluruh tahapan workflow, mulai dari akuisisi data, preprocessing, pembangunan model, hingga evaluasi hasil [25].

Salah satu keunggulan utama Python adalah ketersediaan pustaka komputasi numerik seperti NumPy dan SciPy, yang memungkinkan manipulasi array multidimensi serta operasi matematika tingkat lanjut secara efisien. Selain itu, library Matplotlib dan Seaborn mempermudah proses visualisasi data dan hasil

segmentasi, yang sangat penting dalam proses validasi dan interpretasi klinis. Untuk kebutuhan pemrosesan citra, Python menyediakan OpenCV dan PIL (Pillow) yang memungkinkan konversi format, filtering, resizing, augmentasi, hingga analisis karakteristik citra secara detail [26].

Lebih jauh lagi, dalam pengembangan model deep learning, Python mendukung berbagai framework canggih seperti TensorFlow, Keras, dan PyTorch. Framework ini mempermudah implementasi arsitektur kompleks seperti U-Net, serta integrasi dengan backbone CNN modern seperti EfficientNet, ResNet, MobileNet, VGG, hingga DenseNet. Kelebihan lain Python adalah dukungan penuh terhadap komputasi GPU melalui teknologi CUDA dan cuDNN, yang sangat krusial untuk mempercepat proses training, terutama pada dataset besar atau model 3D yang memiliki kompleksitas komputasi tinggi [27].

Python juga memiliki komunitas open-source yang sangat aktif, sehingga dokumentasi, modul tambahan, hingga solusi untuk berbagai permasalahan teknis dapat diakses dengan mudah. Selain itu, Python kompatibel dengan berbagai platform seperti Google Colab, yang menyediakan lingkungan komputasi gratis dengan akses GPU dan TPU, memungkinkan eksperimen deep learning dapat dijalankan tanpa memerlukan perangkat keras lokal berperforma tinggi.

Dalam penelitian ini, seluruh tahapan mulai dari konversi data .nii.gz, pra-proses citra, pembangunan model U-Net dengan berbagai backbone, evaluasi segmentasi, hingga visualisasi hasil dilakukan sepenuhnya menggunakan Python. Dengan fleksibilitas, efisiensi, serta dukungan pustaka yang lengkap, Python terbukti menjadi pilihan yang sangat tepat untuk tugas segmentasi citra medis berbasis deep learning, khususnya dalam upaya meningkatkan akurasi deteksi dan analisis penyakit ginjal kronis (CKD) secara otomatis.

2.3.1.1 Numpy

NumPy (Numerical Python) merupakan pustaka fundamental dalam bahasa pemrograman Python yang menyediakan dukungan untuk array multidimensi, fungsi-fungsi matematika tingkat lanjut, serta berbagai operasi linier dan statistik

yang sangat penting dalam komputasi ilmiah [43]. Dalam konteks segmentasi citra medis, NumPy memainkan peran sentral dalam mengelola dan memanipulasi data citra yang umumnya direpresentasikan dalam bentuk array 2D atau 3D. Operasi seperti slicing irisan CT scan, perhitungan statistik pixel (misalnya nilai mean dan standar deviasi), serta konversi dari format gambar ke tensor numerik semuanya dilakukan secara efisien menggunakan NumPy [44].

NumPy juga memungkinkan manipulasi array besar secara vektorized (tanpa perulangan eksplisit), yang menghasilkan waktu eksekusi yang jauh lebih cepat dibandingkan metode tradisional berbasis Python native. Dalam penelitian ini, NumPy digunakan secara luas mulai dari tahap preprocessing data (resizing, normalisasi, thresholding), pengolahan mask segmentasi, hingga proses evaluasi model (misalnya perhitungan Dice Coefficient). Fleksibilitas dan kecepatan eksekusi NumPy menjadikannya salah satu komponen tak tergantikan dalam pipeline deep learning untuk segmentasi citra ginjal dan tumor.

2.3.1.2 Matplotlib

Matplotlib adalah pustaka visualisasi data dua dimensi dalam Python yang digunakan untuk menghasilkan berbagai jenis grafik, termasuk visualisasi citra medis [45]. Dalam penelitian ini, Matplotlib digunakan untuk menampilkan citra asli, ground truth segmentasi, dan hasil prediksi model secara berdampingan, sehingga memudahkan proses evaluasi visual terhadap performa segmentasi [46]. Visualisasi tersebut sangat krusial dalam mengidentifikasi potensi kesalahan model, seperti prediksi berlebih (oversegmentation), kurang prediksi (undersegmentation), atau kegagalan dalam mendeteksi struktur ginjal atau tumor yang kecil.

Selain menampilkan citra, Matplotlib juga digunakan untuk memvisualisasikan tren pelatihan model, seperti kurva loss dan metrik akurasi selama epoch. Hal ini membantu dalam monitoring stabilitas model dan mendeteksi overfitting atau underfitting sejak dini. Matplotlib menyediakan berbagai fungsi tingkat tinggi seperti imshow(), plot(), hist(), dan subplots(), yang memungkinkan pengguna menghasilkan visualisasi yang informatif dengan kode yang ringkas dan mudah

dimodifikasi. Oleh karena itu, Matplotlib tidak hanya penting dari sisi dokumentasi hasil, tetapi juga sebagai alat eksplorasi dan debugging selama proses penelitian berlangsung.

2.3.1.3 Tensorflow

TensorFlow adalah framework open-source yang dikembangkan oleh Google, yang menyediakan infrastruktur komputasi numerik untuk pengembangan dan pelatihan model deep learning dalam skala besar [47]. Dalam penelitian ini, TensorFlow digunakan sebagai kerangka utama untuk membangun, melatih, dan mengevaluasi model segmentasi ginjal menggunakan arsitektur U-Net dengan berbagai backbone CNN. TensorFlow menyediakan API tingkat tinggi melalui modul `tf.keras`, yang memungkinkan desain model menjadi lebih modular dan mudah diintegrasikan dengan pipeline preprocessing, augmentasi, dan evaluasi.

Salah satu keunggulan utama TensorFlow adalah dukungannya terhadap eksekusi GPU (melalui CUDA dan cuDNN), yang mempercepat proses pelatihan model secara signifikan—terutama saat menangani data citra medis dalam jumlah besar [48]. Fitur seperti `tf.data.Dataset` digunakan untuk mengelola batch data secara efisien, sedangkan `tf.keras.callbacks` seperti `EarlyStopping` dan `ModelCheckpoint` dimanfaatkan untuk mengontrol pelatihan dan menyimpan model terbaik.

TensorFlow juga menyediakan fleksibilitas dalam merancang fungsi loss kustom, termasuk Dice Loss dan kombinasi dengan Binary Crossentropy, yang digunakan dalam penelitian ini untuk mengoptimalkan segmentasi pada kelas minoritas (tumor). Dengan dokumentasi yang kuat dan dukungan komunitas yang luas, TensorFlow menjadi fondasi yang solid bagi implementasi sistem segmentasi medis modern yang dapat diintegrasikan ke dalam pipeline klinis di masa depan.

2.3.1.4 PIL (Python Imaging Library)

Python Imaging Library (PIL), yang kini lebih dikenal melalui pengembangannya dalam pustaka Pillow, merupakan salah satu library utama di Python untuk memproses citra digital. Pillow dikembangkan sebagai penerus PIL

dengan kompatibilitas penuh terhadap kode lama, namun dilengkapi berbagai perbaikan, peningkatan kinerja, serta dukungan terhadap format citra yang lebih beragam [49]. Dalam penelitian ini, PIL digunakan untuk memfasilitasi berbagai tahapan penting dalam pengolahan citra sebelum data dimasukkan ke dalam model segmentasi. Salah satu peran utamanya adalah untuk membuka dan mengonversi citra hasil ekstraksi dari format medis .nii.gz menjadi format yang lebih mudah diolah seperti .png. Proses konversi ini penting, mengingat sebagian besar framework deep learning modern seperti TensorFlow dan PyTorch lebih kompatibel dengan format citra standar seperti PNG atau JPEG.

Selain itu, PIL berperan dalam proses resizing citra, yaitu menyamakan ukuran seluruh dataset ke dimensi tetap agar sesuai dengan input layer model U-Net dan backbone CNN yang digunakan. Konsistensi ukuran citra ini mutlak diperlukan untuk mencegah error saat pelatihan maupun inferensi model. PIL juga mempermudah manipulasi mode warna citra, misalnya mengubah citra grayscale menjadi RGB atau sebaliknya, agar seluruh input memiliki struktur kanal yang seragam sesuai kebutuhan arsitektur backbone. Pada tahap preprocessing mask segmentasi, PIL digunakan untuk memodifikasi nilai piksel agar label citra dapat dikenali oleh model, seperti membedakan antara background, ginjal, dan tumor berdasarkan warna atau intensitas piksel tertentu.

Seluruh citra yang telah diproses menggunakan PIL kemudian dikonversi ke dalam format array numerik melalui integrasi dengan NumPy. Format array inilah yang digunakan sebagai input akhir ke dalam pipeline pelatihan model deep learning. Dengan demikian, penggunaan PIL tidak hanya terbatas pada tahap visualisasi, namun juga memainkan peran penting dalam memastikan kesiapan data, keseragaman format, serta kelancaran proses pelatihan dan evaluasi model segmentasi ginjal dan tumor. Fleksibilitas serta kemudahan integrasi PIL dengan ekosistem Python menjadikannya alat yang sangat relevan dalam penelitian ini, terutama untuk menjembatani konversi data medis ke dalam format yang kompatibel dengan kebutuhan segmentasi berbasis deep learning.

2.3.2 Environment

Dalam pengembangan model deep learning, dua framework yang paling populer saat ini adalah PyTorch dan Keras. PyTorch dikenal dengan fleksibilitas tinggi dan sangat disukai di kalangan peneliti karena memberikan kontrol penuh terhadap arsitektur dan alur data. PyTorch cocok digunakan dalam eksplorasi eksperimen kompleks dan debugging secara langsung. Di sisi lain, Keras menawarkan antarmuka yang lebih sederhana dan intuitif, serta integrasi kuat dengan TensorFlow, yang memungkinkan implementasi model lebih cepat dan efisien.

Dalam konteks penelitian ini, Keras akan lebih diutamakan karena banyak backbone modern seperti EfficientNet, VGG, MobileNet, ResNet, dan DenseNet telah tersedia secara langsung melalui `keras.applications`, dengan dukungan pretrained weights dan kompatibilitas tinggi terhadap arsitektur U-Net. Hal ini mempermudah proses integrasi, pelatihan, serta evaluasi model, terutama ketika membandingkan berbagai jenis backbone secara konsisten. Sementara PyTorch tetap digunakan sebagai pembanding, Keras memberikan keunggulan praktis dalam pengujian berbagai varian model segmentasi ginjal secara efisien.

2.3.3 Google Colab

Google Colaboratory, atau lebih dikenal sebagai Google Colab, merupakan platform komputasi berbasis cloud yang disediakan oleh Google dan banyak dimanfaatkan untuk pengembangan serta eksperimen model machine learning dan deep learning. Google Colab menawarkan akses gratis ke lingkungan eksekusi Python dengan dukungan GPU (Graphics Processing Unit) dan TPU (Tensor Processing Unit), yang sangat berguna dalam mempercepat proses pelatihan model, khususnya untuk tugas-tugas komputasi intensif seperti segmentasi citra medis berbasis deep learning.

Keunggulan utama Google Colab terletak pada kemudahannya dalam integrasi dengan Google Drive, sehingga pengguna dapat menyimpan, mengakses, dan memproses dataset besar seperti citra CT scan atau hasil segmentasi tanpa memerlukan penyimpanan lokal. Dalam penelitian ini, Google Colab digunakan sebagai platform utama untuk seluruh tahapan eksperimen, mulai dari

preprocessing data citra medis, pelatihan model U-Net dengan berbagai backbone CNN, hingga evaluasi dan visualisasi hasil segmentasi.

Selain itu, Google Colab mendukung instalasi pustaka pihak ketiga secara fleksibel, seperti MONAI, PyTorch, TensorFlow, dan library pendukung lainnya yang dibutuhkan untuk pengolahan data medis dan pembangunan arsitektur segmentasi. Google Colab juga menyediakan notebook interaktif yang memungkinkan kombinasi kode, visualisasi grafik, dan dokumentasi teks dalam satu tempat, sehingga memudahkan proses analisis dan dokumentasi hasil penelitian.

Dengan adanya akses ke GPU Tesla T4 atau P100 secara gratis, Google Colab memungkinkan pelatihan model dilakukan dengan waktu yang relatif cepat meskipun model memiliki kompleksitas tinggi, seperti backbone EfficientNetB7 atau DenseNet121. Platform ini juga mendukung eksperimen berulang dengan parameter yang berbeda tanpa perlu investasi perangkat keras khusus, sehingga efisiensi pengembangan model menjadi lebih optimal. Dengan demikian, Google Colab berperan krusial dalam mendukung pelaksanaan penelitian ini, khususnya dalam konteks segmentasi ginjal dan tumor pada citra CT scan untuk mendukung diagnosis penyakit ginjal kronis (CKD) secara otomatis.

