

## Artikel Penelitian

# FYU-Net: Jaringan Segmentasi Bertingkat untuk Pencitraan Medis Tumor Ginjal

Houwei Feng<sup>ID</sup>, Xupeng Kou, Zhan Tang, dan Lin Li<sup>ID</sup>

Sekolah Tinggi Teknik Informasi dan Elektro, Universitas Pertanian Cina, Beijing 100083, Cina

Korespondensi dapat dialamatkan ke Lin Li; [lilincan@126.com](mailto:lilincan@126.com)

Diterima 12 Agustus 2022; Diterima 13 September 2022; Dipublikasikan 18 Oktober 2022

Penyunting Akademik: Sathishkumar V E

Hak Cipta © 2022 Houwei Feng dkk. Ini adalah artikel akses terbuka yang didistribusikan di bawah Lisensi Atribusi Creative Commons, yang mengizinkan penggunaan, distribusi, dan reproduksi tanpa batas dalam media apa pun, asalkan karya aslinya dikutip dengan benar.

Segmentasi tumor ginjal secara otomatis sangat penting untuk evaluasi diagnostik kanker ginjal. Namun, volume tumor ginjal umumnya kecil dibandingkan dengan volume ginjal dan terdistribusi secara tidak teratur; selain itu, lokasi dan bentuk tumor ginjal sangat bervariasi, sehingga tugas segmentasi menjadi sangat menantang. Untuk mengatasi masalah yang disebutkan di atas, model segmentasi bertingkat (FYU-Net) untuk gambar computed tomography (CT) diusulkan dalam makalah ini untuk mencapai segmentasi tumor ginjal secara otomatis. Model yang diusulkan melibatkan dua langkah utama. Pada langkah pertama, pemindaian cepat data CT ginjal dilakukan dengan menggunakan jaringan lokalisasi untuk menemukan irisan yang mengandung tumor, dan segmentasi kasar dilakukan secara bersamaan. Pada langkah kedua, kerangka kerja segmentasi yang disematkan dengan modul jaringan piramida fitur digunakan untuk menyegmentasikan tumor ginjal secara halus. Dengan membangun struktur piramida fitur, target dengan ukuran yang berbeda didistribusikan untuk dideteksi pada lapisan fitur yang berbeda untuk mengekstrak informasi fitur yang lebih kaya. Selain itu, struktur top-down memungkinkan informasi dari peta fitur tingkat yang lebih tinggi ditransfer ke peta fitur tingkat yang lebih rendah, sehingga meningkatkan informasi semantik dari peta fitur tingkat yang lebih rendah. Eksperimen komparatif dilakukan pada dataset publik Kidney PARSing Challenge 2022; koefisien Jaccard rata-rata dan koefisien Dice rata-rata segmentasi struktur tumor masing-masing lebih dari 70,73% dan lebih dari 82,85%. Hasil penelitian ini menunjukkan keefektifan model yang diusulkan untuk segmentasi tumor ginjal.

## 1. Pendahuluan

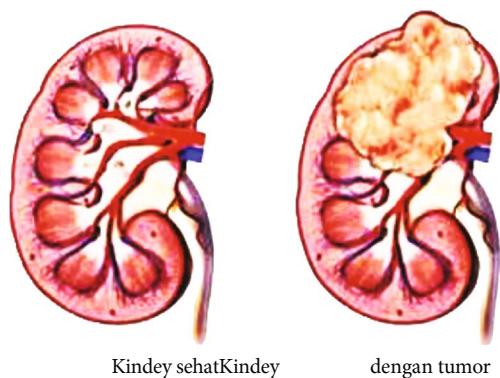
Segmentasi gambar digunakan dalam pemrosesan gambar digital dan visi komputer untuk menyegmentasikan gambar menjadi beberapa bagian atau wilayah berdasarkan fitur piksel dalam gambar [1]. Mengubah gambar menjadi kumpulan wilayah piksel yang direpresentasikan ulang oleh topeng melibatkan pemisahan latar depan dari latar belakang atau pengelompokan wilayah piksel berdasarkan kemiripan warna dan bentuk. Segmentasi gambar digunakan dalam pencitraan medis untuk mendeteksi dan melabeli wilayah gambar yang mewakili tumor di otak pasien dan organ lainnya. Tumor ginjal dapat bersifat jinak atau ganas. Tumor jinak termasuk kista, sedangkan tumor ganas terutama meliputi kanker ginjal dan kanker panggul [2]. Pemeriksaan rutin, deteksi tepat waktu, dan pengobatan yang tepat sangat penting untuk mengatasi tumor ginjal. Kanker ginjal dapat disembuhkan dengan pembedahan

untuk mengangkat sebagian atau seluruh organ yang terkikis oleh kanker [3]. Computed tomography (CT) dan magnetic resonance imaging

adalah modalitas pencitraan medis yang paling penting yang digunakan oleh ahli bedah untuk mendiagnosis dan menemukan tumor ginjal [4]. Lokalisasi tumor ginjal yang akurat sangat penting dalam pembedahan untuk pengobatan kanker ginjal. Oleh karena itu, segmentasi yang tepat dari area tumor ginjal memainkan peran penting. Jika berhasil, hal ini dapat membantu dokter bedah untuk menemukan area tumor ginjal dengan lebih akurat, sehingga meningkatkan tingkat keberhasilan operasi [5]. Struktur tumor ginjal ditunjukkan pada Gambar 1.

Dalam beberapa tahun terakhir, beberapa model segmentasi berbasis deep learning (DL) dengan kinerja yang baik telah dikembangkan [6]; Namun, segmentasi wilayah tumor organ dalam pencitraan medis masih menjadi tantangan, terutama karena dibandingkan dengan organ seperti otak atau jantung, segmentasi tumor dengan perbedaan bentuk dan tekstur yang besar antar individu lebih sulit. Di satu sisi, penilaian ahli pada dasarnya merupakan pertimbangan yang menentukan, dan bahkan dengan perkembangan pesat DL dan kecerdasan buatan





GAMBAR 1: Diagram skematis tumor ginjal.

kecerdasan, hasil diagnosis medis masih bergantung pada penilaian subjektif dokter. Namun, segmentasi manual oleh dokter memakan waktu dan dapat menyebabkan kelelahan mental pada dokter, yang dapat dengan mudah menyebabkan kesalahan operasional. Di sisi lain, tidak seperti segmentasi berbasis organ, bentuk dan tekstur tumor yang berbeda sangat bervariasi, dan sulit untuk menemukan kesamaan di antara keduanya dengan pencocokan langsung. Oleh karena itu, membedakan tumor dengan menggunakan metode segmentasi berorientasi kelas organ menjadi tantangan tersendiri. Meskipun metode DL telah menunjukkan potensi besar dalam menangani masalah segmentasi semantik citra alami, beberapa masalah ditemukan di bidang segmentasi citra medis, seperti kontras yang buruk antara jaringan dan organ serta otot dan lemak, lemahnya informasi batas jaringan target segmentasi, dan ekstraksi fitur yang tidak mencukupi karena informasi yang kaya akan dimensi citra medis [7].

Oleh karena itu, dalam makalah ini, jaringan segmentasi bertingkat untuk gambar medis, yang disebut model FYU-Net, diusulkan. Model yang diusulkan menggabungkan ide segmentasi bertingkat yang disempurnakan dari visi komputer DL. Eksperimen komparatif dilakukan pada dataset publik, dan hasilnya menunjukkan keefektifan model yang diusulkan untuk segmentasi tumor ginjal. Kontribusi utama dari makalah ini adalah sebagai berikut:

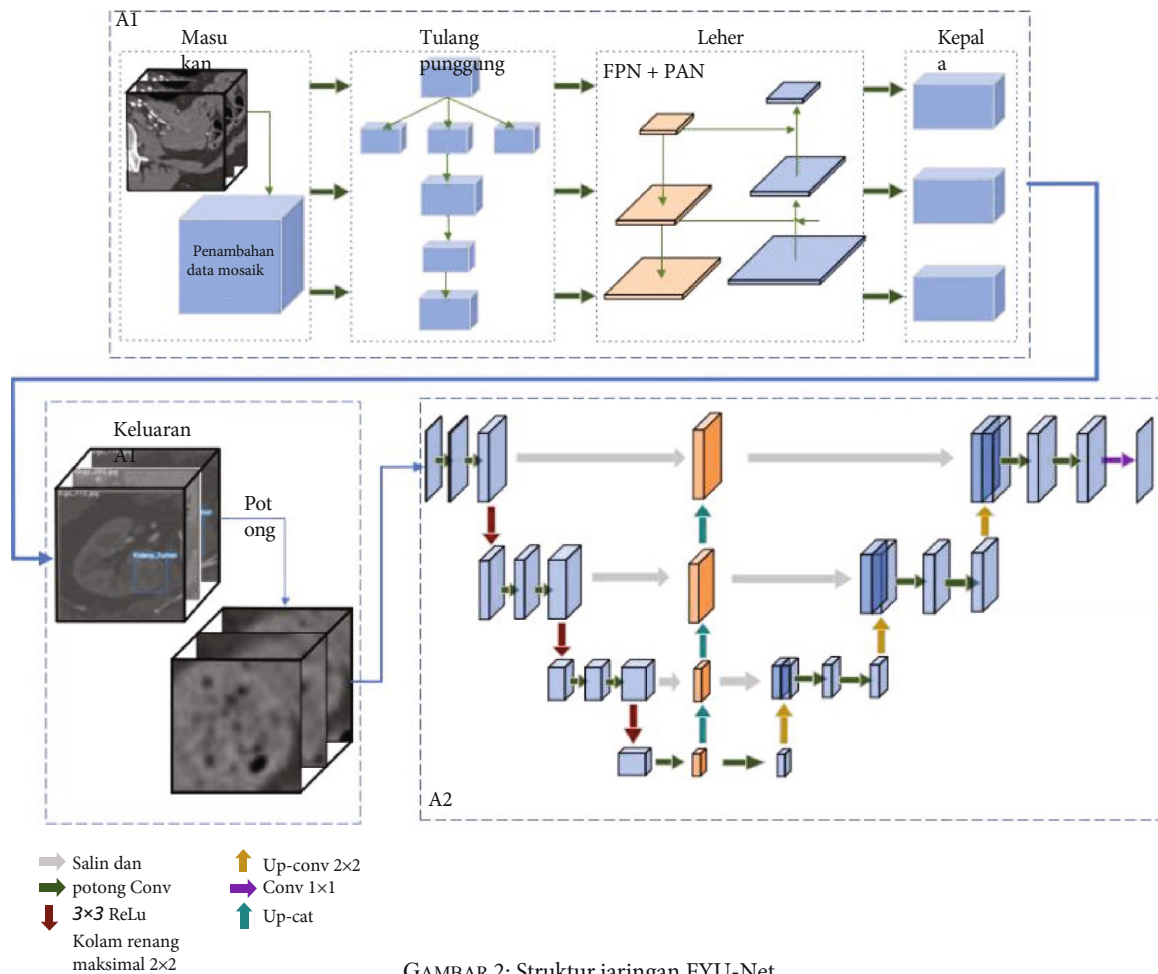
- (i) Model segmentasi kaskade tumor ginjal (FYU-Net) diusulkan untuk mencapai segmentasi yang akurat dari daerah tumor ginjal. Model jaringan lokalisasi digunakan untuk mengidentifikasi bagian dengan tumor dari data CT ginjal untuk mewujudkan lokalisasi cepat daerah tumor ginjal dan segmentasi kasar daerah tumor yang akan digunakan sebagai input jaringan segmentasi.
- (ii) Dalam model segmentasi tumor ginjal, modul jaringan piramida fitur (FPN) diperkenalkan untuk menggabungkan peta fitur resolusi rendah dengan informasi semantik yang kuat dan peta fitur resolusi tinggi dengan informasi semantik yang lemah tetapi informasi spasial yang kaya dengan upaya komputasi yang lebih sedikit. Terakhir, struktur **coder-decoder** klasik digunakan untuk menyusun pasangan informasi citra medis yang kaya untuk segmentasi tumor ginjal yang akurat.

## 2. Karya Terkait

Segmentasi gambar dianggap sebagai proses pencitraan medis yang paling penting di mana daerah yang diminati diekstraksi menggunakan proses semi-otomatis atau otomatis [8]. Gambar dibagi menjadi beberapa wilayah berdasarkan deskripsi tertentu, seperti segmentasi organ dan jaringan tubuh untuk deteksi batas, deteksi tumor, segmentasi, dan deteksi massa. Segmentasi citra medis melibatkan identifikasi area organ atau lesi dari citra medis untuk memberikan informasi penting kepada komunitas medis tentang bentuk dan volume jaringan dan organ tubuh dan merupakan salah satu tugas yang paling menantang dalam analisis citra medis [9]. Dalam beberapa tahun terakhir, segmentasi semantik telah dipelajari secara ekstensif di bidang pencitraan biomedis.

Jaringan Attention U-Net yang diusulkan oleh Oktay dkk. adalah model mekanisme perhatian baru untuk segmentasi pencitraan medis [10]. Model ini menggunakan mekanisme perhatian untuk menekan daerah yang tidak relevan pada gambar input dan fitur menonjol yang terang yang berguna untuk tugas tertentu dan dapat menggantikan perhatian yang keras dalam tugas klasifikasi dan modul pelokalan dalam tugas pelokalan organ. Jin dkk. mengusulkan RA-UNet, model segmentasi hibrida tiga dimensi (3D) yang didasarkan pada persepsi perhatian sisa [11]. Model ini secara akurat mengekstrak volume yang diminati dari hati dan mensegmentasi tumor hati darinya. Model ini memiliki struktur dasar 3D U-Net yang memungkinkan penggabungan pemetaan fitur yang mendasari dengan pemetaan fitur tingkat yang lebih tinggi untuk mengekstrak informasi kontekstual dan merupakan penggunaan pertama mekanisme perhatian residual untuk pemrosesan citra medis. Gu dkk. mengusulkan CE-Net, sebuah model jaringan penyandi konteks yang dapat menangkap informasi tingkat tinggi dan mempertahankan informasi spasial dalam segmentasi citra medis dua dimensi [12]. Model ini terdiri dari tiga modul utama: modul pengkodean fitur, modul ekstraksi konteks, dan modul penguraian fitur. Model  $U^2$ -Net yang diusulkan oleh Qin et al. adalah struktur bersarang berbentuk U dua tingkat yang menangkap informasi kontekstual dari skala yang berbeda dan menggunakan operasi penggabungan untuk meningkatkan kedalaman arsitektur secara keseluruhan tanpa secara signifikan meningkatkan overhead komputasi [13]. Model DeepLab V3 yang diusulkan oleh Chen et al. mengkaji ulang penerapan konvolusi Atrous dalam segmentasi semantik dan menggabungkan fitur-fitur tingkat gambar ke dalam modul ASPP, yang dapat memperoleh informasi kontekstual dengan menggunakan kerangka kerja modul kaskade dan piramida spasial [14]. DeepLab V3+ merupakan versi perbaikan dari DeepLab V3 yang berisi modul decoder tambahan untuk mengoreksi hasil segmentasi dan mengeksplorasi lebih lanjut kombinasi Xception dan deep separable convolution dengan modul ASPP dan decoder [15]. He dkk. mengusulkan model konteks piramida adaptif (APC-Net) yang multiskala, adaptif, dan memiliki afinitas lokal yang dipandu oleh global-guided local affinity (GLA) dan memecahkan masalah bagaimana membobotkan vektor konteks dan peta fitur asli [16]. Zhao dkk. mengembangkan modul pengumpulan piramida dan jaringan penguraian adegan

Metode Komputasi dan Matematika dalam Kadalsteran piramida (1.3.1 Net) untuk menangkap agregasi informasi kontekstual global berdasarkan kemampuan agregasi kontekstual dari berbagai wilayah, memberikan kerangka kerja yang unggul untuk prediksi tingkat piksel [17].



GAMBAR 2: Struktur jaringan FYU-Net.

Namun, model-model yang disebutkan di atas memiliki batasan tertentu dalam segmentasi tumor. Karena kebutuhan untuk menggunakan seluruh data CT sebagai input model dan ukuran tumor ginjal yang relatif kecil, hasil segmentasi dapat terganggu oleh daerah ginjal lainnya. Selain itu, metode-metode ini tidak mempertimbangkan fitur-fitur unik dari gambar medis selama ekstraksi fitur, yang mempengaruhi akurasi segmentasi. Dalam penelitian ini, kami mengembangkan jaringan segmentasi bertingkat (FYU-Net) yang terdiri dari dua langkah yang saling terkait, yaitu, lokalisasi dan segmentasi akurat wilayah tumor ginjal. Pertama, kami menggunakan jaringan lokalisasi target [18] untuk mengekstrak irisan yang mengandung daerah tumor dari data citra medis ginjal dan menyelesaikan operasi segmentasi kasar tumor dalam irisan tersebut. Pada jaringan pascasegmentasi, kami menambahkan modul FPN [19] berdasarkan U-Net [20] dan sepenuhnya mempertimbangkan fitur-fitur khusus dari citra medis saat mengekstraksi data CT ginjal, sehingga meningkatkan akurasi segmentasi.

### 3. Metode

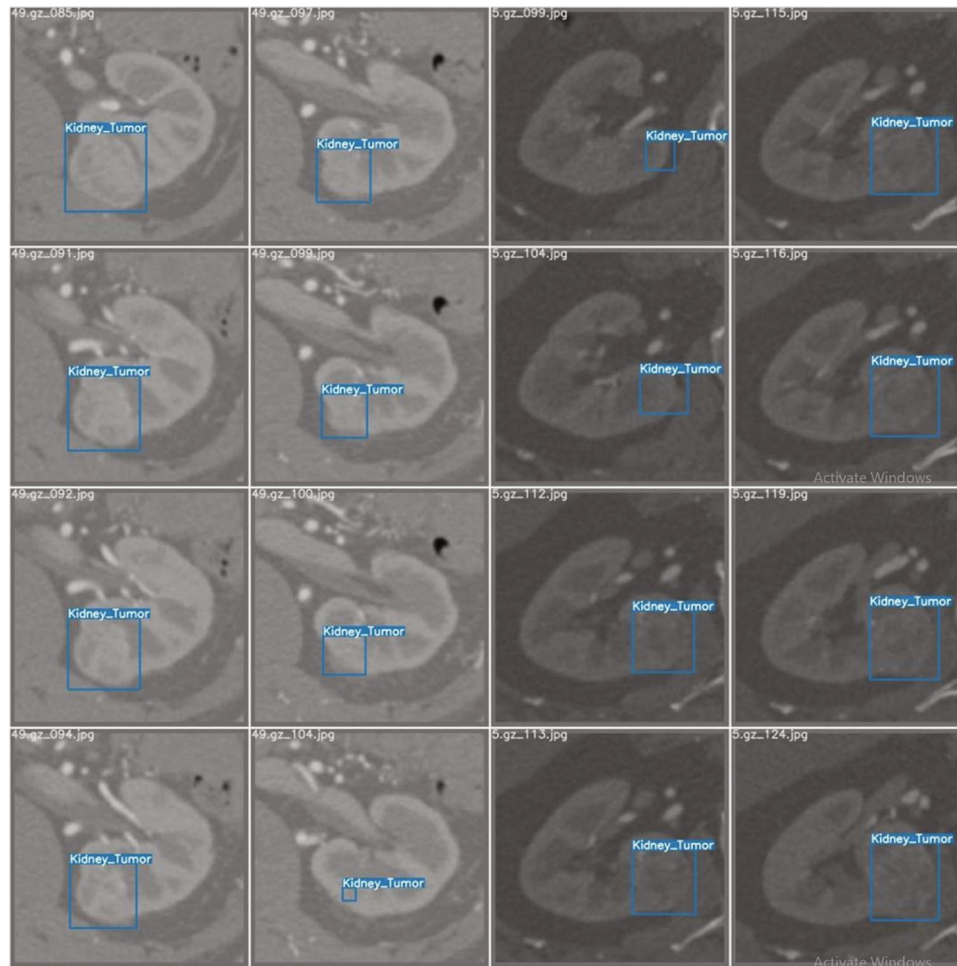
Kerangka kerja metode segmentasi bertingkat yang diusulkan ditunjukkan pada Gambar 2. Irisan data CT ginjal kaya akan informasi, sedangkan daerah tumor target relatif

berukuran kecil. Untuk alasan ini, dalam penelitian ini, wilayah tumor ginjal pertama kali diekstraksi dan diimplementasikan sebagai segmentasi kasar, dan kemudian, data yang diperoleh setelah segmentasi kasar digunakan sebagai jaringan segmentasi input untuk segmentasi halus. Tahap pertama dari kerangka kerja kaskade, yang dilambangkan sebagai A1, diimplementasikan dengan menggunakan model YOLO-V5 untuk menemukan irisan yang mengandung tumor dari total data CT ginjal untuk secara simultan mencapai lokalisasi tumor ginjal yang cepat dan otomatis serta segmentasi kasar tumor pada irisan ini. Pada tahap kedua dari jaringan kaskade (dilambangkan sebagai A2), model jaringan segmentasi yang tertanam dalam modul FPN digunakan untuk secara tepat menyegmentasikan irisan yang terdeteksi dari A1 dan daerah tumor yang sesuai berdasarkan irisan tersebut.

**3.1. Jaringan Lokasi.** Karena informasi dimensi yang kaya dari data CT ginjal asli, irisan yang mengandung informasi tumor hanya mencakup sebagian kecil dari jumlah total irisan ginjal. Daerah tumor dalam data CT ginjal secara otomatis berada pada tahap pertama dari kerangka kerja kaskade (dilambangkan sebagai A1), yang menyediakan data masukan untuk segmentasi kasar tumor ginjal pada tahap kedua dari kerangka kerja kaskade (dilambangkan sebagai A2). Tugas ini dapat dianggap sebagai masalah klasifikasi biner. Jaringan lokalisasi aliran utama termasuk







GAMBAR 3: Hasil jaringan pemosisian A1.

[21], SSD [22], dan jaringan Faster R-CNN [23]. Di antara mereka, YOLO-V5 memiliki kinerja terbaik. Oleh karena itu, dalam penelitian ini, kami menggunakan model jaringan YOLO-V5 untuk lokalisasi tumor. Gambar CT yang berisi wilayah tumor yang terdeteksi oleh jaringan A1 ditunjukkan pada Gambar 3. Setelah lokalisasi otomatis wilayah tumor ginjal, gambar dipotong dengan memperluas sepuluh titik piksel ke luar, sesuai dengan bingkai deteksi, untuk mencapai segmentasi kasar wilayah tumor sebagai data input A2.

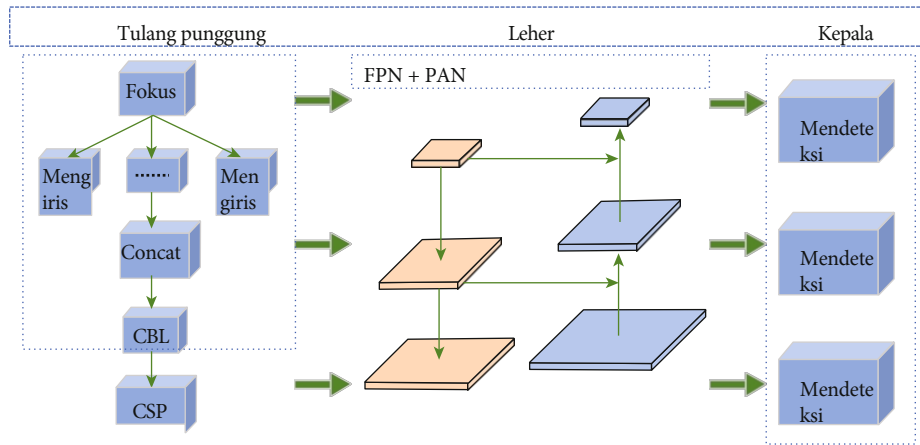
Pada tahap awal pelatihan jaringan pelokalan A1, operasi peningkatan data mosaik digunakan untuk meningkatkan kecepatan pelatihan dan akurasi model. Tulang punggungnya menggunakan modul fokus dengan struktur CSP untuk mengurangi pengulangan informasi gradien ketika menggunakan algoritme BP, sehingga mengurangi kompleksitas komputasi dan jumlah parameter serta meningkatkan kemampuan pembelajaran jaringan. Modul fokus digunakan untuk mengiris gambar sebelum memasuki tulang punggung, dan struktur CSP membagi input asli menjadi dua cabang, CSP1\_X dan CSP2\_X. CSP2\_X menggantikan Resunit dengan  $2 \times X$  CBL relatif terhadap CSP1\_X. Struktur CSP membantu mengurangi ukuran model sekaligus secara efektif mengurangi masalah hilangnya gradien. Struktur jaringan leher berisi FPN + PAN untuk menggabungkan fitur-fitur dengan dimensi yang berbeda. FPN menggunakan top-down

koneksi lateral untuk membangun peta fitur semantik tingkat tinggi dan struktur piramida fitur; namun, karena informasi target di lapisan bawah menjadi kabur setelah jaringan multilayer, PAN menambahkan rute dari bawah ke atas untuk mengkompilasi dan meningkatkan informasi lokalisasi. Struktur kepala mengeluarkan hasil deteksi target dan terdiri dari tiga detektor. Arsitektur jaringan A1 ditunjukkan pada Gambar 4.

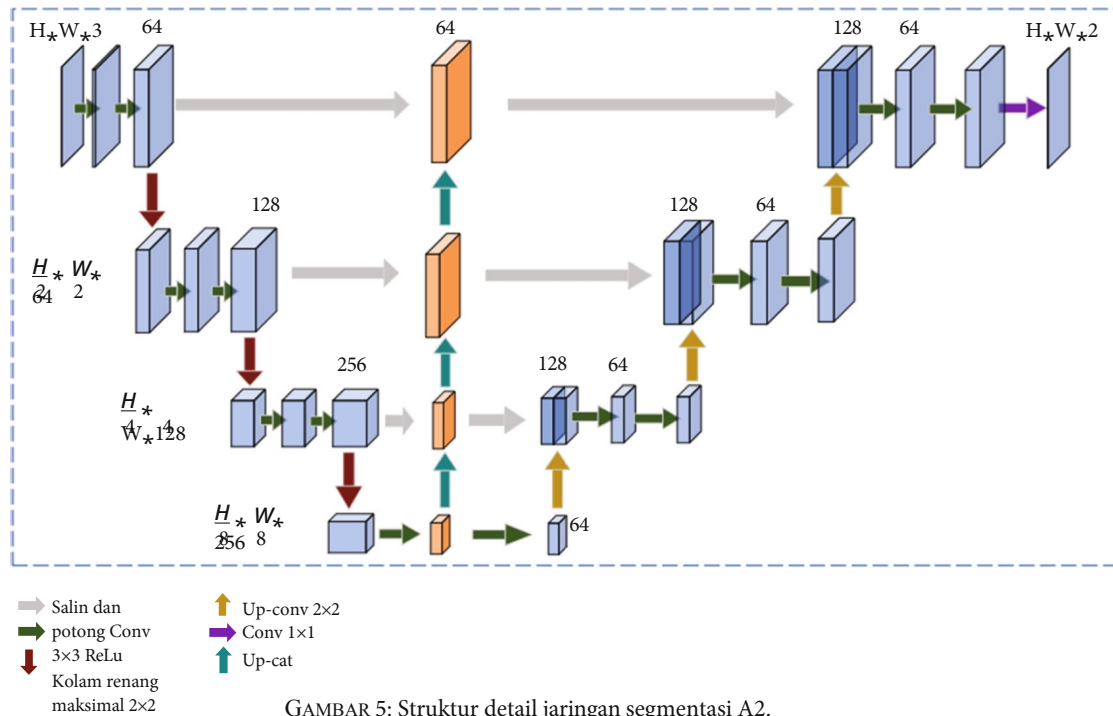
**3.2. Jaringan Segmentasi.** Pada tahap kedua dari kerangka kerja kaskade, struktur berbentuk U digunakan untuk segmentasi sebagai respons terhadap kekhususan fitur informasi gambar medis yang sulit diekstraksi. Lompatan suksesi memperkuat konduksi fitur dari seluruh jaringan dan mengurangi masalah hilangnya gradien sekaligus meningkatkan penggunaan kembali fitur dan meningkatkan kemampuan pembelajaran jaringan. Selain itu, kami menyematkan modul FPN pada tahap ini untuk segmentasi tumor ginjal yang akurat menurut wilayah lokal di A1. Arsitektur jaringan rinci dari A2 ditunjukkan pada Gambar 5.

Struktur encoding-decoding merupakan struktur utama jaringan segmentasi. Karena rendahnya resolusi tumor pada gambar ginjal, unit voxel berisi informasi yang jauh melampaui detail gambar alami





GAMBAR 4: Struktur detail jaringan pemosisian A1.

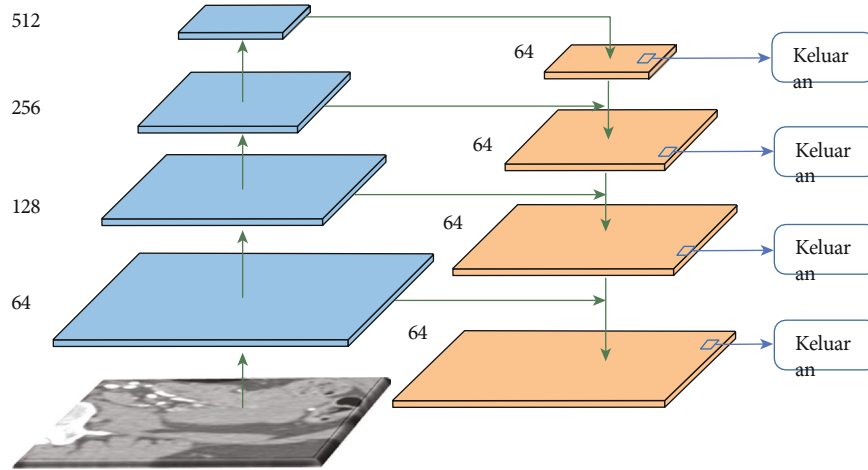


GAMBAR 5: Struktur detail jaringan segmentasi A2.

informasi; kapasitas voxel setidaknya 16 bit atau lebih. Oleh karena itu, kami menguranginya dengan satu lapisan di atas U-Net untuk mencegah model yang terlalu pas. Kami melakukan tiga operasi downsampling dan tiga operasi upsampling interpolasi linier. Ukuran input gambar adalah  $H \times W \times 3$ , dan ukuran output adalah  $H \times W \times 2$ , di mana  $H$  dan  $W$  masing-masing mewakili panjang dan lebar gambar. Jaringan encoder, yang ditunjukkan di sisi kiri pada Gambar 5, digunakan untuk melakukan serangkaian operasi downsampling melalui konvolusi dan max pooling, dan pada setiap operasi downsampling berikutnya, jumlah kanal fitur digandakan. Dua konvolusi  $3 \times 3$  dilakukan, masing-masing dengan fungsi aktivasi ReLU dan max pooling  $2 \times 2$  untuk downsampling dalam langkah 2. Jumlah peta fitur dikalikan dengan 2 setelah setiap downsampling.

plung; dengan demikian, terjadi perubahan ukuran peta fitur. Bagian decoding ditunjukkan di sisi kanan pada Gambar 5; setiap langkah termasuk peta fitur untuk upsampling, diikuti oleh konvolusi  $2 \times 2$  yang membagi dua jumlah saluran fitur dan dua konvolusi  $3 \times 3$ , sekali lagi diikuti oleh fungsi aktivasi ReLU. Pada lapisan terakhir, konvolusi  $1 \times 1$  digunakan untuk memetakan setiap vektor fitur 64 komponen ke jumlah kelas yang diinginkan.

Karena ukuran tumor ginjal yang bervariasi meningkatkan kesulitan segmentasi, sementara algoritme yang ada menggunakan konvolusi yang diperluas (ASPP) atau konvergensi piramida (PSPNet) untuk meningkatkan akurasi segmentasi, algoritme ini menderita efek grid dan hilangnya informasi lokasi tingkat piksel, yang masing-masing tidak kondusif untuk



GAMBAR 6: Diagram struktur jaringan piramida fitur.

konsistensi pemetaan fitur. Selain itu, modul penyatuan piramida yang digunakan dalam PSPNet kehilangan lokalisasi piksel dalam operasi penyatuan pada skala yang berbeda. Untuk memanfaatkan fitur secara penuh dan meningkatkan ekstraksi dan propagasi fitur, dalam penelitian ini, modul FPN dimasukkan karena dapat melakukan ekstraksi fitur untuk setiap skala gambar dan mampu menghasilkan representasi fitur multiskala, dan semua tingkat peta fitur memiliki informasi semantik yang kuat, termasuk beberapa peta fitur resolusi tinggi, yang pada akhirnya meningkatkan efisiensi segmentasi daerah tumor. Detail modul FPN ditunjukkan pada Gambar 6.

Struktur FPN terdiri dari jalur bottom-up, jalur top-down, dan koneksi lateral. Tulang punggung dari jalur bottom-up adalah perhitungan umpan balik maju dari ConvNet, yang menghitung struktur lapisan fitur yang terdiri dari peta fitur multiskala dalam langkah 2. Dalam penelitian ini, tingkat piramida juga ditentukan untuk setiap tahap, dan output dari lapisan terakhir dari setiap tahap dipilih sebagai set peta fitur referensi. Jalur top-down memperoleh fitur dengan resolusi yang lebih tinggi dengan melakukan upsampling terhadap peta fitur yang secara spasial lebih kasar tetapi secara semantik lebih kuat dari tingkat piramida yang lebih tinggi. Fitur-fitur ini kemudian ditambah dengan koneksi lateral dari jalur bottom-up. Setiap koneksi lateral menggabungkan peta fitur dengan ukuran spasial yang sama dari jalur bottom-up dan jalur top-down. Dalam penelitian ini, untuk peta fitur beresolusi lebih kasar, resolusi spasial ditingkatkan dengan faktor 2. Kemudian, peta yang telah ditingkatkan digabungkan dengan peta bottom-up yang sesuai dengan melakukan penambahan piksel per piksel. Proses ini diulang sampai peta dengan resolusi terbaik dihasilkan.

**3.3. Fungsi Kerugian.** Fungsi kerugian cross-entropi secara ekstensif digunakan untuk mengawasi sinyal model dalam segmentasi citra medis [24]. Fungsi kerugian cross-entropi menghindari masalah bahwa bentuk turunan dari fungsi tipe Sigmoid rentan terhadap kejenuhan. Selain itu, fungsi kerugian cross-entropi dapat menghindari dispersi gradien ketika melakukan perhitungan gradient descent, yang mengarah pada penurunan laju pembelajaran. Oleh karena itu, dalam penelitian ini, fungsi

Fungsi kerugian cross-entropi digunakan untuk menghitung kerugian segmentasi:

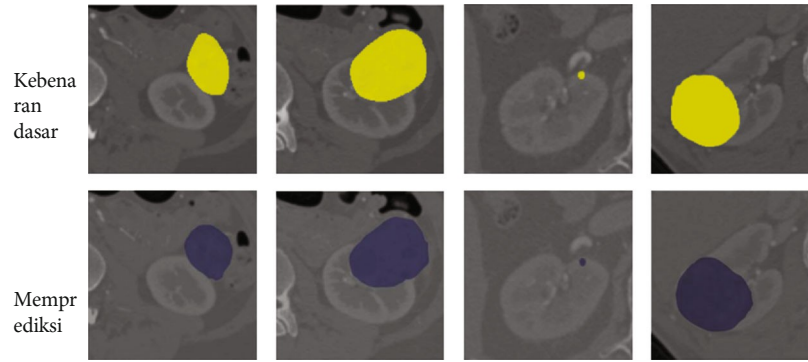
$$L_{CE}(p, q) = - \sum_{i=1}^N p(x_i) \log(q(x_i)), \quad (1)$$

di mana  $N$  menunjukkan jumlah kategori,  $p(x_i)$  menunjukkan distribusi sebenarnya dari sampel, dan  $q(x_i)$  menunjukkan distribusi yang diprediksi oleh model dan probabilitas bahwa sampel termasuk dalam kategori  $N$ .  $p(x_i) = 1$  ketika kategori yang diprediksi sama dengan kategori sampel; jika tidak,  $p(x_i) = 0$ .

#### 4. Detail dan Hasil Eksperimen

Untuk menggambarkan efektivitas dan kemampuan generalisasi model FYU-Net yang diusulkan, model ini dibandingkan dengan model arus utama pada dataset publik kompetisi. Hasil eksperimen dianalisis dari berbagai perspektif. Model yang diusulkan diimplementasikan menggunakan kerangka kerja PyTorch, dan lingkungan yang berjalan adalah GPU NVidia GeForce GTX 3080 dengan memori video 10 GB. Epoch ditetapkan sebagai 100 sehingga semua model dapat mencapai konvergensi. Laju pembelajaran ditetapkan sebagai 0,01, laju peluruhan bobot 0,0005, dan metode gradien stokastik batch kecil dengan nilai momentum 0,9 digunakan sebagai metode pengoptimalan. Pada tahap pertama jaringan A1, kami memasukkan ukuran gambar asli dalam format "jpg". Pada tahap kedua percobaan, hasil yang diperoleh pada model A1 dipotong untuk memasukkan irisan tumor ginjal dan kemudian dimasukkan ke dalam jaringan A2.

**4.1. Dataset.** Kami menggunakan dataset dari Kidney PARSing Challenge 2022 [25-28]. Dataset ini mencakup gambar CT ginjal yang tidak disempurnakan dari 70 pasien. Tumor adalah lesi multisubtype, dengan lima jenis subtumor dalam dataset dengan distribusi yang bervariasi, sehingga menghasilkan situasi yang lebih menantang. Pertama, kami menetapkan lebar jendela dan nilai bit yang sesuai untuk data gambar medis. Selanjutnya, kami mengonversi



GAMBAR 7: Hasil segmentasi model FYU-Net.

format gambar dari ".nii.gz" ke ".jpg" dan menggunakan gambar sebagai input dari model. Akhirnya, data tersebut dijalankan

domly dibagi menjadi set pelatihan dan pengujian dengan rasio 8 : 2.

**4.2. Indikator Evaluasi dan Pengaturan Parameter.** Untuk mengevaluasi kinerja segmentasi secara akurat dengan mengacu pada standar evaluasi utama saat ini, kami mengadopsi empat metrik, yaitu, koefisien Dice, koefisien Jaccard, Presisi, dan Recall:

$$\text{Dadu} = \frac{2 * |V_s \cap V_g|}{|V_s| + |V_g|}, \quad (2)$$

$$\text{Jaccard} = \frac{|V_s \cap V_g|}{|V_s \cup V_g|}, \quad (3)$$

$$\text{Presisi} = \frac{|V_s \cap V_g|}{|V_s|}, \quad (4)$$

$$\text{Recall} = \frac{|V_s \cap V_g|}{|V_g|}, \quad (5)$$

di mana  $V_s$  menunjukkan himpunan piksel yang secara otomatis memisahkan tumor ginjal dan  $V_g$  menunjukkan himpunan piksel yang sebenarnya di wilayah tumor ginjal.

**4.3. Analisis Hasil.** Untuk memverifikasi keefektifan model yang diusulkan, kami membandingkannya dengan model segmentasi arus utama, yaitu, APCNet, HRNet [29], FCN [30], dan U-Net. Parameter pelatihan model-model ini ditetapkan sesuai dengan parameter optimal mereka dalam literatur yang sesuai, dan hasil pengenalan direaksikan berdasarkan kode sumber yang disediakan oleh penulis. Evaluasi kualitatif dan perbandingan kuantitatif dengan model-model mainstream menunjukkan keefektifan model FYU-Net yang diusulkan. Selain itu, analisis konvergensi model menunjukkan stabilitas model yang diusulkan.

Seperti yang ditunjukkan pada Gambar 7, model FYU-Net yang diusulkan menunjukkan keunggulan visual yang sangat tinggi, yang menunjukkan bahwa model ini dapat sangat membantu untuk operasi reseksi tumor ginjal. Pada Gambar 7, baris pertama menunjukkan nilai sebenarnya

TABEL 1: Hasil segmentasi setiap model pada data tumor ginjal.

Model	Dadu	Jaccard	Presisi	Ingat
APCNet	0.6627	0.4956	0.7374	0.6017
HRNet	0.6957	0.5334	0.7389	0.6573
FCN	0.7313	0.5765	0.7490	0.7145
U-Net	0.7350	0.5810	0.8039	0.6770
Kami.	0.8285	0.7073	0.7877	0.8738

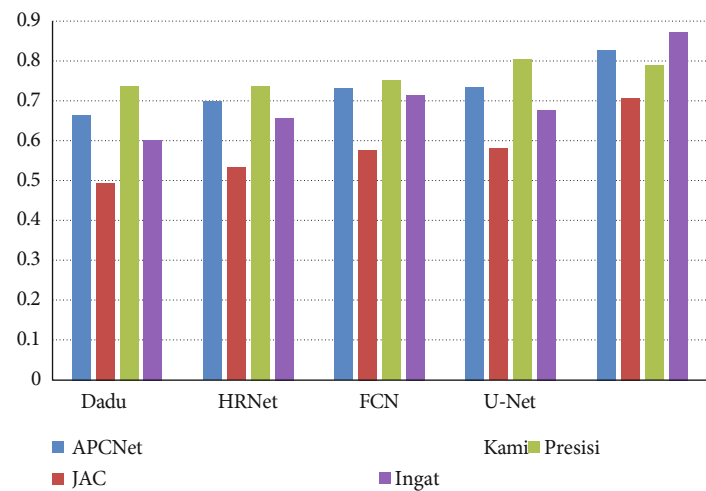
Model yang diajukan mencapai hasil segmentasi tumor yang baik secara visual.

Hasil eksperimen dari model yang diusulkan dan model utama pada dataset ginjal disajikan dalam Tabel 1. Model yang diusulkan mencapai koefisien Dice, koefisien Jaccard, dan nilai recall maksimum, yaitu 82,85%, 70,73%, dan 87,38%, dan kinerja pengenalnya jauh lebih baik daripada yang lain. Presisi mewakili probabilitas deteksi tumor dalam hasil prediksi; namun, nilai presisi yang tinggi saja tidak menunjukkan bahwa efek segmentasi model

dari tumor ginjal, dan baris kedua menunjukkan peta hasil segmentasi FYU-Net. Dapat diamati bahwa pro-

baik. Dalam model yang diusulkan, FCN digunakan untuk menggantikan lapisan yang terhubung penuh dengan lapisan konvolusi, dan dekonvolusi digunakan untuk upsampling. Selain itu, struktur jump-connected digunakan untuk menggabungkan informasi kasar dengan informasi halus untuk segmentasi. Dengan demikian, efek segmentasi tumor dari model yang diusulkan lebih baik daripada APCNet dan HRNet. Nilai koefisien Dice dan koefisien Jaccard untuk model yang diusulkan masing-masing adalah 73,13% dan 57,65%. Sebaliknya, U-Net berbeda dengan FCN; U-Net memperluas struktur jaringan FCN, dan upsampling-nya masih memiliki sejumlah besar saluran, yang memungkinkan jaringan untuk menyebarkan informasi kontekstual ke resolusi yang lebih tinggi, dan bagian upsampling menggabungkan output dari bagian ekstraksi fitur, yang menggabungkan fitur-fitur multiskala. Dengan demikian, U-Net mencapai hasil segmentasi yang lebih baik daripada FCN, dengan koefisien Dice, koefisien Jaccard, dan nilai pre-cision masing-masing sebesar 73,50%, 58,10%, dan 80,39%. Gambar 8 mengilustrasikan hasil perbandingan visualisasi.

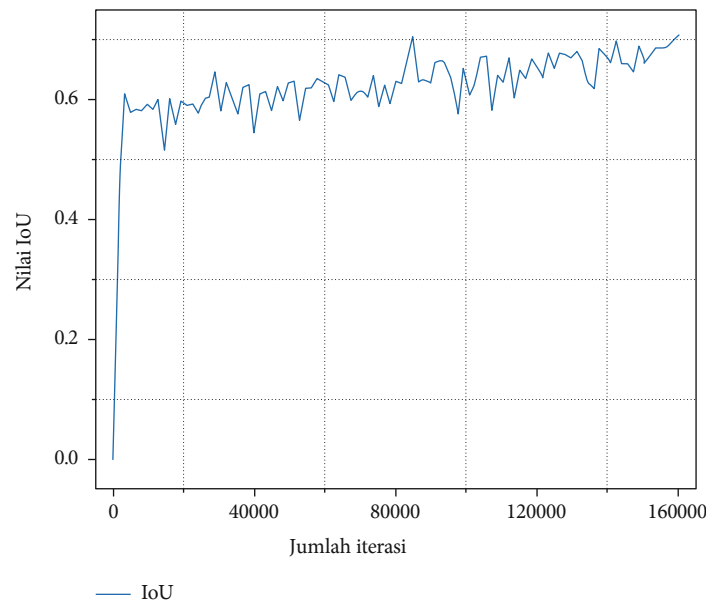
*4.4. Eksperimen Ablasi.* Untuk memvalidasi keefektifan model FYU-Net yang diusulkan dengan lebih baik, kami melakukan percobaan ablasi numerik. Seperti yang disajikan pada Tabel 2, kami membandingkan hasil segmentasi FYU-Net tanpa



GAMBAR 8: Visualisasi hasil perbandingan model.

TABEL 2: Hasil perbandingan eksperimen ablasi.

SHU-Net	LN	FPN	Dadu	Jaccard
-	-	-	0.7350	0.5810
-	-	✓	0.7513	0.6017
✓	-	✓	0.7649	0.6192
✓	✓	✓	0.8285	0.7073



GAMBAR 9: Tren konvergensi IoU dari model FYU-Net.

menambahkan berbagai strategi; SHU-Net menunjukkan U-Net dangkal, dan LN menunjukkan jaringan lokalisasi. Strategi yang kami tingkatkan menghasilkan peningkatan kinerja yang cukup besar. Nilai koefisien segmentasi Dice dan Jaccard dari U-Net asli pada data tumor ginjal masing-masing adalah 73,50% dan 58,10%. Koefisien Dice dan Jaccard

Nilai koefisien Dice dan Jaccard meningkat menjadi 75,13% dan 60,17%, masing-masing, ketika kami hanya menggunakan strategi FPN. Ketika struktur U-Net dangkal dan modul FPN digunakan, efek segmentasi meningkat, dan nilai koefisien Dice dan Jaccard meningkat menjadi 76,49% dan 61,92%. Ketika kami mengadopsi strategi penuh dan menggunakan FYU-

Net untuk segmentasi data CT tumor ginjal, efek segmentasi meningkat pesat, menghasilkan koefisien Dice rata-rata 82,85% dan koefisien Jaccard 70,73%, sehingga menunjukkan keefektifan skema perbaikan kami.

*4.5. Analisis Konvergensi Model.* Jumlah iterasi pelatihan model adalah salah satu faktor penting yang mempengaruhi struktur segmentasi tumor ginjal. Seperti yang ditunjukkan pada Gambar 9, IoU model yang diusulkan mulai meningkat secara tiba-tiba dari 0 ke 0,6, yang mencerminkan efektivitas FYU-Net. Seiring dengan bertambahnya jumlah iterasi, kurva IoU menunjukkan tren pertumbuhan secara bertahap dan akhirnya mencapai kondisi konvergen, dengan nilai IoU yang akhirnya stabil di sekitar 0,7. Karena memori video, kami menetapkan `batch_size` sebagai 2, sehingga menyebabkan kurva IoU masih memiliki osilasi kecil di akhir iterasi.

## 5. Diskusi

Dalam penelitian ini, kami mengembangkan jaringan segmentasi bertingkat untuk gambar medis untuk mewujudkan segmentasi wilayah tumor dalam data CT ginjal. Karena tumor berbentuk tidak beraturan dan volumenya relatif kecil dibandingkan dengan ginjal dan dapat muncul di lokasi yang berbeda di ginjal, kami pertama-tama menggunakan algoritma deteksi target berdasarkan implementasi YOLO-V5 m untuk melokalisasi daerah tumor dengan cepat dalam data CT ginjal mentah. Selanjutnya, kami menerapkannya pada daerah yang terlokalisasi untuk melakukan segmentasi tumor yang akurat. Hasil percobaan pada dataset CT ginjal yang tersedia untuk umum yang disediakan oleh Kidney PArasing Challenge 2022 menunjukkan bahwa model yang diusulkan lebih akurat dan kuat daripada metode yang ada. Nilai koefisien Dice dan koefisien Jaccard untuk model yang diusulkan masing-masing adalah 82,85% dan 70,73%. Lebih lanjut, model FYU-Net yang diusulkan dalam makalah ini menunjukkan efisiensi waktu yang baik, yang menunjukkan bahwa jaringan segmentasi bertingkat yang diusulkan sangat meningkatkan efisiensi segmentasi citra medis. Akhirnya, percobaan ablasi menunjukkan bahwa jaringan yang diusulkan menghasilkan kinerja segmentasi gambar medis yang unggul.

## Ketersediaan Data

Data yang digunakan dalam penelitian ini adalah dataset publik yang disediakan oleh Kidney PArasing Challenge 2022, yang dapat diperoleh dengan mendaftar ke kontes atau dengan mengakses data melalui situs web Github penulis (<https://github.com/L9uckin/kidney>).

## Poin Tambahan

*Ketersediaan Kode.* Kode yang digunakan untuk segmentasi tumor ginjal dari data CT tersedia di situs web Github penulis (<https://github.com/L9uckin/kidney>). Kode yang digunakan untuk analisis komputasi lainnya tersedia

berdasarkan permintaan kepada penulis.

Para penulis menyatakan tidak memiliki kepentingan yang bersaing.

## Ucapan terima kasih

Penelitian ini didukung oleh Program Riset dan Pengembangan Kunci Nasional Tiongkok (hibah no. 2021YFD1300100).

## Referensi

- [1] A. Ibrahim dan E. S. M. El-kenawy, "Metode segmentasi citra berdasarkan teknik superpixel: sebuah survei," *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 15, no. 3, pp. 1-11, 2020.
- [2] A. Hamroun, R. Lenain, F. Provôt, M. Maanaoui, dan A. Lionet, "Kasus yang tidak biasa dari tumor ginjal bilateral multipel," *Jurnal Nefrologi*, vol. 34, no. 5, pp. 1783-1784, 2021.
- [3] T. Kobayashi, A. Takeuchi, H. Nishiyama, dan M. Eto, "Status saat ini dan perspektif masa depan imunoterapi terhadap kanker urothelial dan kanker ginjal," *Japanese Journal of Clinical Oncology*, vol. 51, no. 10, pp. 1481-1492, 2021.
- [4] M. Z. F. Ho, K. M. Lim, dan E. C. P. Chua, "Memanfaatkan kecerdasan buatan (AI) untuk mengotomatiskan defacing pada hidung pada gambar computed tomography (CT) dan magnetic resonance imaging (MRI)," *Journal of Medical Imaging and Radiation Sciences*, vol. 53, no. 3, p. 8, 2022.
- [5] S. Campbell, R. G. Uzzo, M. E. Allaf dkk., "Massa ginjal dan kanker ginjal terlokalisasi: Pedoman AUA," *The Journal of Urology*, vol. 198, no. 3, pp. 520-529, 2017.
- [6] I. Ahmed, M. Ahmad, F. A. Khan, dan M. Asif, "Perbandingan model segmentasi berbasis deep-learning: menggunakan gambar orang tampak atas," *IEEE Access*, vol. 8, hal. 136361-136373, 2020.
- [7] M. H. Hesamian, W. Jia, X. He, dan P. Kennedy, "Teknik pembelajaran mendalam untuk segmentasi citra medis: pencapaian dan tantangan," *Journal of Digital Imaging*, vol. 32, no. 4, pp. 582-596, 2019.
- [8] P. Yin, R. Yuan, Y. Cheng, dan Q. Wu, "Deep guidance network untuk segmentasi citra biomedis," *IEEE Access*, vol. 8, hal. 116106-116116, 2020.
- [9] S. Pang, C. Pang, L. Zhao dkk., "SpineParseNet: penguraian tulang belakang untuk citra MR volumetrik dengan kerangka kerja segmentasi dua tahap dengan representasi citra semantik," *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 262-273, 2021.
- [10] O. Oktay, J. Schlemper, L. L. Folgoc dkk., "Perhatian u-net: mempelajari di mana harus mencari pankreas," <https://arxiv.org/abs/1804.03999>, 2018.
- [11] Q. Jin, Z. Meng, C. Sun, H. Cui, dan R. Su, "RA-UNet: jaringan hibrida yang sadar akan perhatian mendalam untuk mengekstrak hati dan tumor dalam pemindaian CT," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 1471, 2020.
- [12] Z. Gu, J. Cheng, H. Fu dkk., "CE-Net: jaringan penyandi konteks untuk segmentasi citra medis 2d," *IEEE Transactions on Medical Imaging*, vol. 38, no. 10, pp. 2281-2292, 2019.
- [13] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, dan

M. Jagersand, "U<sup>2</sup> -Net: melangkah lebih dalam dengan struktur U bersarang untuk deteksi objek yang menonjol," *Pengenalan Pola*, vol. 106, artikel ???, 2020.



- [14] L. C. Chen, G. Papandreou, F. Schroff, dan H. Adam, "Memikirkan kembali konvolusi atrofi untuk segmentasi citra semantik," <https://arxiv.org/abs/1706.05587>, 2017.
- [15] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, dan H. Adam, *Encoder-Decoder dengan Atrous Separable Convolution untuk Segmentasi Citra Semantik*, Prosiding konferensi Eropa tentang visi komputer (ECCV), 2018.
- [16] J. He, Z. Deng, L. Zhou, Y. Wang, dan Y. Qiao, *Jaringan Konteks Piramida Adaptif untuk Segmentasi Semantik*, Prosiding Konferensi IEEE / CVF tentang Visi Komputer dan Pengenalan Pola, 2019.
- [17] H. Zhao, J. Shi, X. Qi, X. Wang, dan J. Jia, *Pyramid Scene Parsing Network*, Prosiding konferensi IEEE tentang visi komputer dan pengenalan pola, 2017.
- [18] J. Aspnes, T. Eren, D. K. Goldenberg dkk., "Sebuah teori pelokalan jaringan," *IEEE Transactions on Mobile Computing*, vol. 5, no. 12, hal. 1663-1678, 2006.
- [19] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, dan S. Belongie, *Jaringan Piramida Fitur untuk Deteksi Objek*, Prosiding konferensi IEEE tentang visi komputer dan pengenalan pola, 2017.
- [20] O. Ronneberger, P. Fischer, dan T. Brox, "U-net: jaringan konvolusi untuk segmentasi citra biomedis," dalam *Konferensi Internasional tentang komputasi citra medis dan intervensi dengan bantuan komputer*, hal. 234-241, Springer, Cham, 2015.
- [21] A. Bochkovskiy, C. Y. Wang, dan H. Y. M. Liao, "Yolov4: kecepatan optikal dan akurasi pendeteksian objek," <https://arxiv.org/abs/2004.10934>, 2020.
- [22] W. Liu, D. Anguelov, D. Erhan dkk., "Ssd: single shot multibox detector," dalam *European Conference on Computer Vision*, pp. 21-37, Springer, Cham, 2016.
- [23] S. Ren, K. He, R. Girshick, dan J. Sun, "Faster r-cnn: menuju deteksi objek waktu nyata dengan jaringan usulan wilayah," *Advances in Neural Information Processing Systems*, vol. 28, 2015.
- [24] R. Das dan S. Chaudhuri, "Pada keterpisahan kelas dengan fungsi kerugian cross-entropi," <https://arxiv.org/abs/1909.06930>, 2019.
- [25] Y. He, G. Yang, J. Yang dkk., "Meta grayscale adaptive network untuk segmentasi struktur ginjal terintegrasi 3D," *Analisis Citra Medis*, vol. 71, artikel 102055, 2021.
- [26] Y. He, G. Yang, J. Yang dkk., "Jaringan bias padat dengan anatomi apriori mendalam dan adaptasi wilayah keras: pembelajaran semi-terpantau untuk segmentasi arteri ginjal halus," *Analisis Gambar Medis*, vol. 63, artikel 101722, 2020.
- [27] P. Shao, C. Qin, C. Yin dkk., "Nefrektomi parsial laparoskopi dengan penjepitan arteri ginjal segmental: teknik dan hasil klinis," *European Urology*, vol. 59, no. 5, pp. 849- 855, 2011.
- [28] P. Shao, L. Tang, P. Li dkk., "Penjepitan arteri ginjal segmental yang tepat di bawah panduan angiografi tomografi tomografi terkomputasi sumber ganda selama nefrektomi parsial laparoskopi," *Urologi Eropa*, vol. 62, no. 6, pp. 1001-1008, 2012.
- [29] J. Wang, K. Sun, T. Cheng dkk., "Pembelajaran represinya beresolusi tinggi yang mendalam untuk pengenalan visual," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 10, pp. 3349-3364, 2020.
- [30] J. Long, E. Shelhamer, dan T. Darrell, *Fully Convolutional Net- works for Semantic Segmentation*, Prosiding konferensi IEEE tentang visi komputer dan pengenalan pola, 2015.