

Evaluasi Komparatif Arsitektur Convolutional Neural Network untuk Deteksi Kanker Prostat pada Citra Magnetic Resonance Imaging

Fikri Yoma Rosyidan¹, Rimuljo Hendradi², Eto Wuryanto³

^{1,2,3} Universitas Airlangga, Jalan Dharmawangsa Selatan 2A, Surabaya, 60286, Indonesia

Info Artikel

Riwayat Artikel:

Received 2025-01-31

Revised 2025-06-22

Accepted 2025-06-28

Abstract – Prostate cancer is a malignant tumor originating from the prostate gland, an essential component of the male reproductive system. With the increasing prevalence of prostate cancer, accurate early detection methods are crucial. This study focuses on the utilization of deep learning, specifically Convolutional Neural Networks (CNN), to diagnose prostate cancer from MRI images. Research was conducted to evaluate three CNN architectures: U-Net, nnU-Net, and nnDetection, aiming to determine the best-performing architecture. The study employed secondary data comprising 1294 MRI images from The PI-CAI Challenge "Artificial Intelligence & Radiologists Prostate Cancer Detection in MRI" conducted in 2022. The data underwent pre-processing steps, including pixel intensity normalization, data augmentation techniques such as rotation and scaling, and image cropping to eliminate irrelevant regions. Subsequently, the processed data were used to train models based on each of the three architectures. Model performance was evaluated using the Area Under the Receiver Operating Characteristic Curve (AUROC) and Average Precision (AP) metrics. The evaluation results indicated that the U-Net architecture achieved an AUROC of 89.94% and AP of 51.22%, the nnU-Net architecture reached an AUROC of 97.75% and AP of 86.67%, while the nnDetection architecture had an AUROC of 83.66% and AP of 49.91%. Comparison of these architectures revealed that nnU-Net yielded the best performance, achieving an AUROC of 97.75% and an AP of 86.67%. This research highlights the potential of CNN in prostate cancer diagnosis using MRI images, emphasizing the importance of selecting an appropriate architecture for deep learning applications in medical imaging.

Keywords: AUROC, AP, Convolutional Neural Network (CNN), Prostate Cancer.

Corresponding Author:

Fikri Yoma Rosyidan

Email:

fikriyomarosyidan01@gmail.com



This is an open access article under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.

Abstrak – Kanker prostat adalah tumor ganas yang berada dari kelenjar prostat, yang merupakan bagian penting dari sistem reproduksi pria. Adanya peningkatan prevalensi kanker prostat maka diperlukan deteksi dini yang akurat. Penelitian ini memfokuskan pada pemanfaatan deep learning, khususnya metode Convolutional Neural Network (CNN) untuk mendiagnosis kanker prostat melalui citra MRI. Diperlukan penelitian untuk mengkaji tiga arsitektur CNN: U-Net, nnU-Net, dan nnDetection agar didapatkan arsitektur yang terbaik. Data penelitian ini menggunakan data sekunder sejumlah 1294 citra MRI dari The PI-CAI Challenge "Artificial Intelligence & Radiologists Prostate Cancer Detection in MRI" tahun 2022. Data tersebut menjalani proses pre-processing, termasuk normalisasi intensitas piksel, augmentasi data seperti rotasi dan scaling, serta pemotongan gambar untuk menghilangkan area yang tidak relevan. Proses selanjutnya data tersebut akan masuk ke tahap pelatihan model dengan menggunakan ketiga arsitektur. Hasil dari pelatihan tersebut akan dievaluasi kinerja modelnya dengan menggunakan metrik Area Under the Receiver Operating Characteristic Curve (AUROC) dan Average Precision (AP). Hasil evaluasi menunjukkan bahwa arsitektur U-Net mencapai AUROC 89,94% dan AP 51,22%, arsitektur nnU-Net mencapai AUROC 97,75% dan AP 86,67%. dan arsitektur nnDetection mencapai AUROC 83,66% serta AP 49,91%. Dari perbandingan hasil ketiga arsitektur maka didapatkan hasil terbaik adalah arsitektur nnU-Net dengan capaian AUROC 97,75% dan AP 86,67%. Penelitian ini menunjukkan potensi penggunaan CNN dalam diagnosis kanker prostat melalui citra MRI. Temuan penelitian menegaskan pentingnya pemilihan arsitektur yang tepat dalam aplikasi deep learning untuk citra medis.

Kata Kunci: AP, AUROC, Convolutional Neural Network (CNN), Kanker Prostat.

I. PENDAHULUAN

Kanker prostat adalah tumor ganas yang berasal dari kelenjar prostat, suatu bagian penting dari sistem reproduksi pria. Kelenjar prostat memiliki peran vital dalam menghasilkan cairan mani yang membantu dalam pengangkutan dan perlindungan sperma [1]. Kanker prostat berkembang ketika sel-sel pada kelenjar prostat tumbuh secara tidak terkendali, membentuk tumor yang berpotensi menyebar ke organ atau jaringan tubuh lainnya [2]. Pertumbuhan tumor yang tidak terdeteksi dini ini dapat memperburuk prognosis pasien dan menambah tantangan dalam pengobatan [3]. Dengan kejadian lebih dari 1,4 juta kasus baru dan 375.304 kematian pada tahun 2020, kanker prostat menjadi jenis kanker kedua yang paling sering didiagnosis pada pria

di seluruh dunia [4]. Tingkat kejadian kanker prostat bervariasi di berbagai wilayah, tetapi yang tertinggi ditemukan di Eropa Utara sebesar 83,4 kasus per 100.000 orang dan di Eropa Barat sebesar 77,6 kasus per 100.000 orang [4]. Di Indonesia, kanker prostat juga menjadi salah satu perhatian utama dalam penanggulangan penyakit kanker. Berdasarkan data *Global Cancer Observatory* (Globocan) tahun 2020, kanker prostat merupakan jenis kanker ketujuh terbanyak di Indonesia dengan 13.563 kasus baru yang didiagnosis pada tahun tersebut dan menempati peringkat ke-11 dalam jumlah kasus baru di Indonesia. Selain itu, pada tahun yang sama, kanker prostat menyebabkan sekitar 4.863 kematian, menjadikannya kanker dengan angka kematian ke-13 tertinggi di Indonesia [3].

Pemeriksaan kanker prostat melibatkan *Digital Rectal Examination* (DRE) dan pengujian *prostate-specific antigen* (PSA). Namun, metode-metode ini memiliki keterbatasan dalam akurasi dan dapat mengakibatkan *overdiagnosis* serta *overtreatment* [5]. Biopsi dilakukan ketika ada indikasi mencurigakan. Namun, prosedur ini invasif dan berpotensi menyebabkan infeksi dan perdarahan [6]. Teknik pencitraan seperti *Magnetic Resonance Imaging* (MRI) mulai digunakan untuk meningkatkan akurasi diagnostik [7]. Dalam dekade terakhir, MRI *multiparametrik* diakui sebagai standar non-invasif untuk menilai lesi prostat secara anatomis dan fungsional [8]. Namun, akurasi interpretasi citra MRI bergantung kuat pada keahlian radiolog sehingga memicu kebutuhan akan sistem bantu diagnosis terotomatisasi [9].

Kemajuan *deep learning*, khususnya *Convolutional Neural Networks* (CNN), menawarkan peluang meningkatkan akurasi sekaligus konsistensi interpretasi. Beberapa penelitian menunjukkan bahwa model CNN mampu menyamai bahkan melampaui kinerja radiolog dalam mendeteksi kanker prostat signifikan klinis [10]. Di antara berbagai arsitektur, *U-Net* [11] masih menjadi *baseline* dominan berkat kesederhanaan, jumlah parameter yang relatif kecil, dan fleksibilitas adaptasinya, bahkan kompetisi PI-CAI merilis *U-Net semi-supervised* sebagai *baseline* resmi pada 2023 [12]. Sejumlah studi mutakhir memodifikasi *U-Net* dengan modul *lightweight*, konteks multi-skala, ataupun *backbone* hibrida CNN–Mamba sehingga kebutuhan memori dapat ditekan tanpa mengorbankan presisi [13]. Sebagai pengembangan lebih lanjut, *nnU-Net* [14] hadir sebagai kerangka *self-configuring* yang secara konsisten memuncaki berbagai kompetisi segmentasi, termasuk pada data prostat. Kendati demikian, strategi *ensembling* multi-skala serta pelatihan berbasis *patch* besar membuat *nnU-Net* relatif haus sumber daya komputasi sehingga kurang ideal untuk rumah sakit dengan GPU terbatas [15]. Di sisi lain, *nnDetection* [16] memperluas konsep *self-configuration* ke tugas deteksi volumetrik 3D, memungkinkan lokalisasi lesi tanpa memerlukan anotasi *voxel* penuh, tetapi perbandingan langsung antara model segmentasi (*U-Net* dan *nnU-Net*) dan model deteksi murni (*nnDetection*) pada dataset PI-CAI belum pernah dipublikasikan.

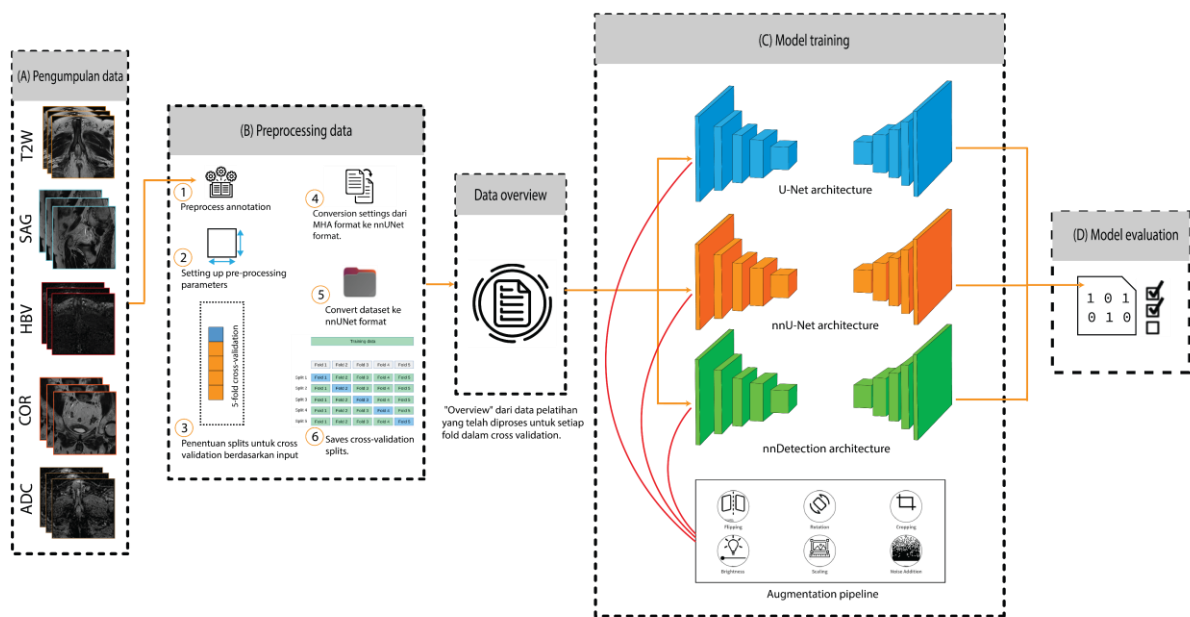
Penelitian ini bertujuan mengisi celah tersebut dengan melakukan *benchmark* komprehensif atas ketiga arsitektur pada dataset PI-CAI 2022 yang mencakup 1.294 pemindaian [17]. Seluruh citra diproses menggunakan *pipeline* pra-proses yang seragam, dan evaluasi dilakukan melalui validasi silang lima lipatan. Dengan skema itu, studi ini mengkaji secara kritis pertukaran antara akurasi tinggi *nnU-Net* dan efisiensi komputasi *U-Net*, sekaligus menilai kinerja *nnDetection* pada skenario yang sama. Selain menggunakan metrik AUROC untuk mengukur kemampuan diskriminasi [17], penelitian ini juga menambahkan *Average Precision* (AP) guna menangkap performa pada distribusi data yang sangat tidak seimbang [18]. Hasil yang diperoleh diharapkan memberikan landasan teknis bagi pengembangan sistem *Computer-Aided Diagnosis* (CAD) kanker prostat berbasis CNN di lingkungan rumah sakit dengan sumber daya beragam.

U-Net yang diperkenalkan pada 2015 telah melahirkan ragam turunan seperti *Attention-U-Net* [19], *MAS-U-Net* [20], dan *MM-U-Net* berbasis *Mamba* [21], masing-masing menambahkan mekanisme perhatian, konteks spasial multi-skala, atau *backbone* sekuensial baru untuk menekan *false positive* tanpa menambah kompleksitas berlebihan. Terlepas dari dominasi *nnU-Net* dalam berbagai kompetisi, *U-Net* tetap relevan sebagai *baseline* ringan untuk keperluan *edge deployment* dan validasi lintas-studi. Hal ini tercermin dalam rilis *open-source baseline semi-supervised* PI-CAI tahun 2023. *nnU-Net* memformalkan pendekatan “*no-code required*” dengan menyesuaikan normalisasi, strategi *cropping*, dan konfigurasi jaringan secara otomatis terhadap distribusi data [22]. Meta-analisis terbaru menunjukkan peningkatan rerata 2–5 poin *Dice* pada segmentasi MRI prostat bila dibandingkan model *ad-hoc*, meski biaya GPU dan waktu inferensi meningkat signifikan [15]. Berbeda dari kedua arsitektur tersebut, *nnDetection* menggeneralisasi filosofi *self-configuration* untuk tugas deteksi objek 3D sehingga dapat melokalisasi lesi tanpa membutuhkan segmentasi *voxel* demi *voxel* [16]. Pendekatan ini cocok ketika label segmentasi *granular* sukar diperoleh, tetapi performa relatifnya terhadap *nnU-Net* pada dataset besar PI-CAI belum pernah dibahas secara sistematis. Di luar ranah CNN, tren mutakhir mengarah pada model fondasi medis berbasis *transformer*, semisal *Medical-GPT* atau *MedSAM* [23], yang memadukan *backbone U-Net* dengan blok *self-attention* untuk meningkatkan *robustness* antarpusat data. Kendati menjanjikan, model-model tersebut cenderung membutuhkan memori GPU besar dan belum sepenuhnya *open-source*, sehingga tolok ukur berbasis CNN konvensional masih menjadi rujukan utama.

Berdasarkan telaah literatur, belum ada kajian yang secara bersamaan membandingkan paradigma segmentasi dasar, segmentasi *self-configuring*, dan deteksi *self-configuring* pada dataset PI-CAI dengan *pipeline* pra-proses identik serta evaluasi menggunakan kombinasi AUROC dan AP. Penelitian ini dirancang untuk menutup kesenjangan tersebut sekaligus menyediakan panduan praktis pemilihan arsitektur. Melalui pemahaman keunggulan dan keterbatasan masing-masing arsitektur CNN, para peneliti dan praktisi medis dapat membuat keputusan yang lebih tepat dan efektif dalam memanfaatkan teknologi *deep learning* untuk diagnosis kanker prostat.

II. METODE

Penelitian ini menggunakan desain eksperimental dengan tujuan untuk membandingkan efektivitas tiga arsitektur *Convolutional Neural Network* (CNN), yaitu *U-Net*, *nnU-Net*, dan *nnDetection*, dalam mendeteksi kanker prostat menggunakan citra *Magnetic Resonance Imaging* (MRI). Tahapan penelitian yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Diagram alir penelitian

A. Pengumpulan Data

Pada tahap ini, dilakukan pengambilan data untuk menjadi masukan dalam sistem. Data yang digunakan merupakan data sekunder yang didapat dari situs web resmi perlombaan *The PI-CAI Challenge "Artificial Intelligence & Radiologists Prostate Cancer Detection in MRI"* tahun 2022. Dataset ini berjumlah 1294 pemindaian MRI *biparametrik* prostat anonim, yang semuanya didiagnosa mengidap penyakit kanker prostat dan telah dianotasi untuk membedakan antara jaringan prostat yang normal dan yang terkena kanker. Pemindaian ini diperoleh menggunakan tujuh pemindai yang berbeda, dari dua vendor, di tiga pusat yang berbeda, sehingga menghasilkan variasi signifikan dalam resolusi spasial gambar antar pasien. Selain itu, semua gambar MRI ini ditampilkan dalam *grayscale* yang membutuhkan peninjauan cermat terhadap intensitas piksel untuk identifikasi yang akurat. Setiap kasus pasien dalam dataset ini memiliki setidaknya tiga urutan pemindaian: T2W *axial*, DWI *axial*, dan pemindaian ADC *axial*, yang masing-masing ditandai dengan ekstensi file *_t2w.mha*, *_hbv.mha*, dan *_adc.mha*. Selain itu, ada kemungkinan adanya urutan pemindaian opsional, seperti T2W *sagittal* dan *coronal*, yang ditandai dengan ekstensi *_sag.mha* dan *_cor.mha*, yang bisa ada pada sebagian kasus, keduanya, atau tidak sama sekali.

B. Pre-processing Data

Serangkaian langkah *pre-processing* dijalankan untuk mempersiapkan dan menyatukan dataset sebelum pelatihan model dimulai. Tujuan utamanya adalah memastikan data yang masuk ke model memiliki bentuk dan format yang konsisten. Proses anotasi yang detail merupakan langkah awal yang berperan penting dalam menandai lesi kanker prostat klinis signifikan (csPCa) pada gambar MRI. Anotasi ini dilakukan pada tingkat *voxel* menggunakan *ITK-SNAP v3.80*, dengan mengintegrasikan berbagai sumber informasi termasuk gambar MRI, laporan diagnostik, dan, bila tersedia spesimen *prostatektomi* penuh. Dataset yang terdiri dari 1294 kasus,

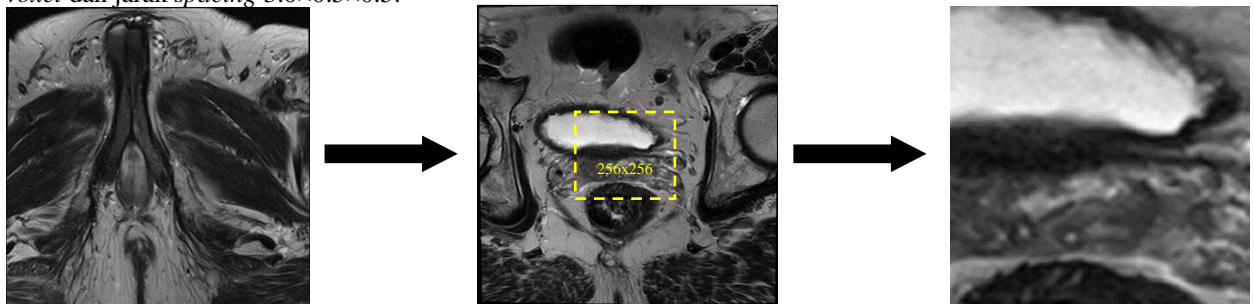
menunjukkan 1075 kasus tanpa csPCa atau hanya dengan PCa tidak agresif, yang direpresentasikan dengan anotasi 'kosong' atau nilai-nilai 0. Sebaliknya, 219 kasus diketahui memiliki csPCa.

Untuk dilakukan analisis, data anotasi *granular* dikonversi menjadi format biner dengan mengubah label menjadi dua kategori: positif untuk kanker prostat (csPCa) dan negatif (tidak ada csPCa). Pada proses konversi, label granular (nilai yang lebih tinggi dari 0 atau 1) dikonversi menjadi label biner dilakukan dengan menetapkan semua nilai label di atas ambang batas tertentu ke label positif (1) dan semua yang lain ke label negatif (0). Nilai ditetapkan pada ambang batas 2 yang berfokus pada deteksi kanker prostat yang klinis signifikan (csPCa). Selanjutnya label *granular* menunjukkan semua area dengan skor ISUP *grade* ≥ 2 dikonversi menjadi label positif (1) menandakan keberadaan csPCa, sementara semua lesi dengan skor ISUP *grade* ≤ 1 dikonversi menjadi negatif (0) mengindikasikan ketiadaan kanker prostat klinis signifikan. Informasi konversi format biner dapat dikodekan ke dalam angka dengan keterangan seperti pada Tabel 1.

TABEL 1
KONVERSI LABEL KE FORMAT BINER

Diagnosis				
Label	Anotasi Ahli Radiologi	Jumlah Data	Kode Numerik	Arti
0	ISUP <i>grade</i> ≤ 1	1075	0	negatif (tidak ada csPCa)
2	ISUP <i>grade</i> 2			
3	ISUP <i>grade</i> 3			
4	ISUP <i>grade</i> 4	219	1	positif (csPCa)
5	ISUP <i>grade</i> 5			

Dua jenis format anotasi disediakan: anotasi asli dan yang disesuaikan. Anotasi asli, berdasarkan urutan T2W, DWI, dan ADC, mungkin menampilkan variasi dalam orientasi dan resolusi. Sementara itu, anotasi yang disesuaikan di-resample untuk mencapai dimensi dan resolusi yang seragam dengan gambar T2W, memastikan konsistensi dalam dataset dan memudahkan integrasi data untuk pelatihan model. Selanjutnya, penyesuaian parameter *pre-processing* seperti *spacing* dan *matrix_size* dilakukan, memungkinkan untuk mengontrol cara data diproses sebelum masuk ke model. Karena dataset melibatkan pemindaian MRI dari tujuh pemindai yang berbeda, dari dua vendor, di tiga pusat, *resampling* resolusi spasial menjadi penting untuk men-standarisasi resolusi spasial gambar. Pemotongan ke *Region-of-Interest* (ROI), terutama area kelenjar prostat, dilakukan untuk mengurangi dimensi data dan menghilangkan segmen yang tidak informatif atau bisa mempengaruhi kinerja model. Peneliti mengambil pemotongan tengah dari setiap pemindaian, dengan ukuran $20 \times 256 \times 256$ voxel dan jarak *spacing* $3.0 \times 0.5 \times 0.5$.



Gambar 2. Pemotongan *Region-of-Interest* (ROI) pada MRI prostat

Selanjutnya dilakukan proses pembagian data menjadi set pelatihan (*training*) dan validasi (*validation*) menggunakan metode *5-fold cross-validation* secara alami mencakup konsep rasio 80/20. Rasio 80% *training* dan 20% *validation* dilakukan melalui penggunaan file konfigurasi JSON dengan pendekatan *cross-validation*. *Cross-validation* adalah teknik statistik yang digunakan untuk memvalidasi keefektifan model dengan membagi data menjadi beberapa bagian atau *fold* [24]. Data dibagi menjadi lima *fold* yang berbeda, dengan setiap *fold* memiliki set pelatihan dan validasi yang unik. Setiap file JSON ini berisi daftar subjek atau kasus yang akan digunakan dalam pelatihan dan validasi untuk *fold* tertentu. Pendekatan ini memastikan bahwa model diuji pada berbagai subset data untuk meningkatkan kemampuan generalisasi dan mengurangi risiko *overfitting*.

Hasil dari proses pembagian data tersebut kemudian diekspor kembali ke disk dalam bentuk konfigurasi set data. Tujuan utama dari langkah ini adalah untuk memastikan dokumentasi yang jelas dan *reproducibility*, yang memungkinkan peneliti untuk secara akurat mencatat bagaimana data dibagi selama proses pelatihan dan

validasi. Informasi terkait subjek disimpan dalam file JSON yang terorganisir berdasarkan masing-masing *fold*. Langkah ini dilakukan untuk memastikan bahwa informasi tentang pembagian data tersimpan secara sistematis dan bisa diakses untuk keperluan analisis terhadap tiga arsitektur yang berbeda.

Name	Date modified	Size
PI-CAI_train-fold-0.json	22/09/2023 16:30	544 KB
PI-CAI_train-fold-1.json	22/09/2023 16:30	542 KB
PI-CAI_train-fold-2.json	22/09/2023 16:30	547 KB
PI-CAI_train-fold-3.json	22/09/2023 16:30	541 KB
PI-CAI_train-fold-4.json	22/09/2023 16:30	548 KB
PI-CAI_val-fold-0.json	22/09/2023 16:30	137 KB
PI-CAI_val-fold-1.json	22/09/2023 16:30	139 KB
PI-CAI_val-fold-2.json	22/09/2023 16:30	134 KB
PI-CAI_val-fold-3.json	22/09/2023 16:30	140 KB
PI-CAI_val-fold-4.json	22/09/2023 16:30	132 KB

Gambar 3. File disk hasil konfigurasi *cross-validation*

C. Pelatihan Model

Pada tahap ini, pelatihan model dilakukan dengan memuat direktori hasil konfigurasi set data yang telah diorganisir menjadi file-file data train dan validation untuk memfasilitasi proses validasi silang dengan metode *5-fold cross-validation*. Direktori ini kemudian berisi serangkaian file JSON yang merepresentasikan setiap *fold*, dengan file-file bernama PI-CAI_train-fold-0.json, PI-CAI_train-fold-1.json, dan seterusnya untuk data train, serta PI-CAI_val-fold-0.json, PI-CAI_val-fold-1.json, dan seterusnya untuk data validasi.

Menggunakan metode *5-fold cross-validation*, keseluruhan dataset dibagi menjadi lima bagian (*fold*) yang seimbang. Setiap *fold* secara bergantian digunakan sebagai set validasi, sementara sisanya digunakan sebagai set pelatihan dengan rasio 80% *training* dan 20% *validation*. Cara ini memastikan bahwa setiap bagian dari dataset digunakan baik sebagai set pelatihan maupun validasi.

D. Evaluasi Model

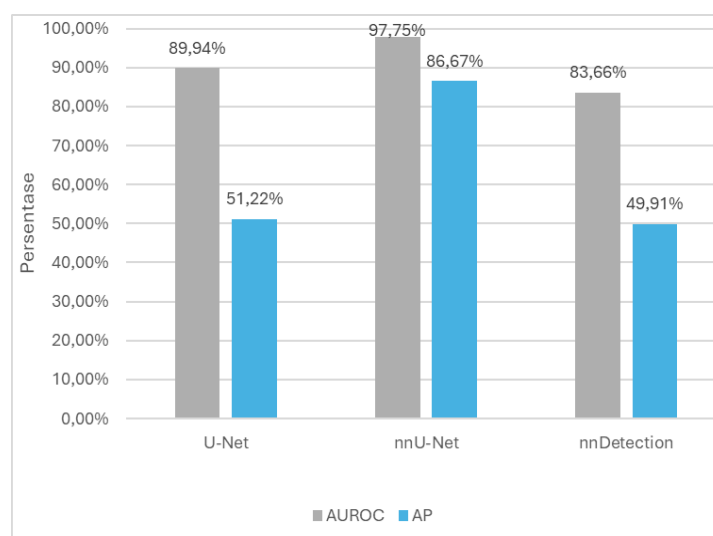
Model selanjutnya dievaluasi menggunakan metrik evaluasi yang meliputi *Area Under the Receiver Operating Characteristic Curve* (AUROC) dan *Average Precision* (AP). AUROC digunakan untuk mengukur sejauh mana model mampu mengklasifikasikan data dengan benar, terutama dalam kasus klasifikasi yang *sensitive* [25]. Secara umum semakin tinggi nilai AUROC maka semakin baik model tersebut dalam memprediksi data sebagai kanker atau bukan kanker. Sedangkan AP digunakan untuk mengevaluasi kinerja model klasifikasi biner terutama kelas positif yang memiliki kepentingan lebih tinggi atau ketika data menunjukkan ketidakseimbangan kelas yang signifikan [18].

III. HASIL DAN PEMBAHASAN

Tujuan dari penelitian ini adalah membandingkan tiga arsitektur, yaitu *U-Net*, *nnU-Net*, dan *nnDetection* dalam mendeteksi kanker prostat. Tujuan ini diharapkan dapat mengidentifikasi model dengan arsitektur paling efektif untuk diagnosis kanker prostat. Berdasarkan penelitian yang telah dilakukan, nilai AUROC dan AP yang tertinggi adalah arsitektur *nnU-Net* dengan nilai AUROC sebesar 97,75% dan AP sebesar 86,67%. Selanjutnya di urutan kedua adalah model dengan arsitektur *U-Net* yang memiliki nilai AUROC sebesar 89,94% dan AP sebesar 51,22%. Terakhir model dengan arsitektur *nnDetection* menempati urutan ketiga dengan nilai AUROC sebesar 83,66% dan AP sebesar 49,91%. Perbandingan selengkapnya terkait performa ketiga model arsitektur pada *fold* terbaik ditunjukkan pada Tabel 2 dan Gambar 4.

TABEL 2
PERBANDINGAN ARSITEKTUR PADA FOLD TERBAIK

No	Arsitektur	AUROC (%)	AP (%)	Keterangan
1	<i>nnU-Net</i>	97,75	86,67	<i>Fold-5</i>
2	<i>U-Net</i>	89,94	51,22	<i>Fold-5</i>
3	<i>nnDetection</i>	83,66	49,91	<i>Fold-5</i>



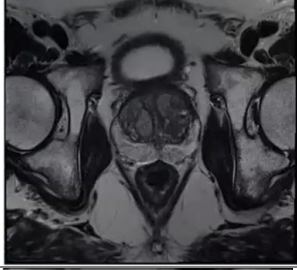
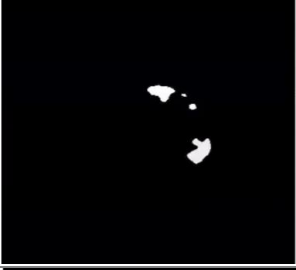
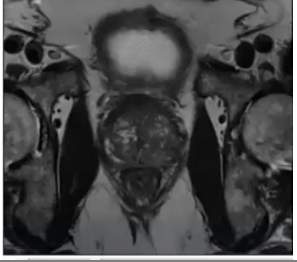

Gambar 4. Grafik hasil perbandingan arsitektur dengan nilai AUROC dan AP terbaik

nnU-Net muncul sebagai arsitektur paling unggul dengan AUROC 97,75 % dan AP 86,67 %, disusul *U-Net* dan *nnDetection*. Keunggulan ini dapat dipahami dari karakteristik desain masing-masing model. *nnU-Net* menerapkan rangkaian prosedur *self-configuration* mulai dari normalisasi, penentuan ukuran *patch*, hingga skema *loss* yang disesuaikan otomatis terhadap distribusi spasial dan intensitas citra MRI [14]. Proses penyesuaian data-driven tersebut memungkinkan jaringan beroperasi pada resolusi optimal dan meminimalkan ketidaksesuaian skala antarpusat pemindaian, suatu aspek krusial dalam dataset PI-CAI yang berasal dari tujuh pemindai berbeda. Selain itu, *nnU-Net* memanfaatkan *deep supervision* multi-resolusi sehingga gradien propagasi ke lapisan awal lebih stabil. Mekanisme inilah yang sering dilaporkan meningkatkan sensitivitas deteksi lesi kecil bila dibandingkan *U-Net* standar [26]. Di sisi lain, *U-Net* tetap kompetitif berkat arsitektur *encoder-decoder* simetris yang efisien parameter [27], tetapi ketiadaan modul adaptif membuatnya kurang mampu menangani variasi *voxel-size* ekstrem, sementara *nnDetection* meskipun dirancang untuk lokalisasi objek 3-D, hasilnya bergantung pada kotak pembatas (*bounding box*) dan tidak memodelkan konteks antar *voxel* secara sehalus pendekatan segmentasi [16], sehingga presisi keseluruhan lebih rendah.

Tingginya nilai AUROC *nnU-Net* menandakan kapasitas diskriminatif model untuk memisahkan distribusi skor antara kasus positif dan negatif di seluruh rentang ambang (*threshold*). Secara teoritis, AUROC merepresentasikan probabilitas bahwa model memberi skor lebih tinggi pada sampel positif acak dibandingkan sampel negatif acak [28]. Metrik ini bersifat bebas ambang dan invarian terhadap ketidakseimbangan kelas, sehingga cocok menilai kinerja dasar klasifikator pada data medis yang biasanya condong ke kelas negatif. Sementara itu, AP yang mengintegrasikan presisi dan sensitivitas lebih sensitif terhadap penurunan presisi pada *cut-off* tinggi [17]. Hal ini menjelaskan mengapa *U-Net* mempertahankan AUROC mendekati 90% tetapi AP-nya turun drastis karena masih menghasilkan sejumlah *false positive* ketika model dipaksa meningkatkan sensitivitas. Lebih jauh, perbedaan 35% AP antara *nnU-Net* dan *U-Net* mengindikasikan bahwa *calibration* internal *nnU-Net* (melalui *batch instance normalization* adaptif dan strategi penyeimbangan kelas) efektif mereduksi prediksi keliru pada kelas positif yang langka. Selain itu, perbandingan dengan *nnDetection* menunjukkan bahwa, untuk deteksi kanker prostat berbasis MRI, representasi segmentasi *voxel-level* masih lebih kaya informasi daripada hanya kotak pembatas, terutama ketika lesi berukuran kecil dan berbentuk ireguler. Temuan ini sejalan dengan literatur terkini yang menunjukkan model segmentasi adaptif cenderung mengungguli pendekatan deteksi murni pada struktur anatomi kompleks.

Dengan demikian, performa superior *nnU-Net* yang didapatkan merupakan hasil kombinasi konfigurasi otomatis, supervisi multi-skala, dan normalisasi kontekstual yang secara teoritis meningkatkan *feature alignment* antar *voxel* serta stabilitas gradien. Penjelasan teoritis ini selaras dengan observasi kuantitatif, yakni korelasi kuat antara skor prediksi tinggi dan label positif maupun skor rendah dan label negatif pada uji kasus individual, yang mengilustrasikan konsistensi diskriminatif *nnU-Net* dalam praktik. Berikut ini adalah hasil penerapan prediksi kasus kanker prostat menggunakan model terbaik *nnU-Net* dengan perbandingan hasil label kasus yang sebenarnya yang ditampilkan pada Tabel 3.

TABEL 3
 HASIL PREDIKSI KASUS KANKER PROSTAT MENGGUNAKAN MODEL nnU-NET

Input MRI prostat	csPCa map deteksi	Skor prediksi kasus	Label sebenarnya
		0,03131103515625	0 (negatif)
		0,607421875	1 (positif csPCa)

Tabel tersebut menggambarkan hasil prediksi kasus kanker prostat menggunakan model *nnU-Net* terbaik dengan perbandingan terhadap label kasus yang sebenarnya. Pada baris pertama, ditampilkan gambar MRI prostat seorang pasien yang diinterpretasikan oleh model. Dalam hasil deteksi, terdapat area kecil yang diidentifikasi sebagai kemungkinan kanker prostat signifikan klinis (csPCa). Model ini memberikan skor prediksi sebesar 0,0313 yang menunjukkan kemungkinan rendah adanya csPCa. Label sebenarnya untuk kasus ini adalah 0 (negatif) yang berarti tidak ditemukan csPCa dalam kenyataannya.

Pada baris kedua, ditampilkan gambar MRI dari pasien lain. Model mendeteksi area yang lebih besar, yang diduga sebagai csPCa, dan memberikan skor prediksi sebesar 0,6074 yang mengindikasikan kemungkinan tinggi adanya csPCa. Label sebenarnya untuk kasus ini adalah 1 (positif csPCa) yang menunjukkan bahwa pasien ini memang memiliki kanker prostat signifikan klinis. Hasil ini memperlihatkan bahwa model *nnU-Net* mampu membedakan antara kasus positif dan negatif dengan baik, terlihat dari korelasi antara skor prediksi yang tinggi dengan label positif, serta skor prediksi yang rendah dengan label negatif. Hal tersebut sejalan dengan kajian teoritis yang telah dipaparkan sebelumnya.

IV. SIMPULAN

Pendeteksian kanker prostat pada data citra MRI dapat dilakukan dengan menggunakan teknik *Convolutional Neural Networks* (CNN), khususnya melalui penerapan arsitektur *U-Net*, *nnU-Net*, dan *nnDetection*. Proses ini melibatkan tahapan *pre-processing* data untuk menstandarisasi citra MRI, diikuti oleh pelatihan model CNN menggunakan ketiga arsitektur tersebut. Arsitektur *U-Net*, *nnU-Net*, dan *nnDetection* masing-masing diuji kemampuannya dalam mengidentifikasi karakteristik kanker prostat dengan menggunakan citra MRI. Setelah proses pelatihan, evaluasi model dilakukan untuk menentukan efektivitas setiap arsitektur dalam mendeteksi kanker prostat dengan menggunakan metrik evaluasi *Area Under the Receiver Operating Characteristic Curve* (AUROC) dan *Average Precision* (AP). Berdasarkan keseluruhan rangkaian pra-proses, pelatihan lima-fold, dan analisis kinerja, *nnU-Net* terbukti paling unggul dalam mendeteksi kanker prostat pada citra MRI, sebagaimana ditunjukkan oleh nilai AUROC 97,75% dan AP 86,67%. Keunggulan ini konsisten dengan pembahasan sebelumnya, yaitu konfigurasi otomatis *nnU-Net* berhasil menyesuaikan resolusi *voxel*, menyeimbangkan kelas, dan menerapkan *deep supervision* multi-skala sehingga dapat membedakan lesi klinis signifikan secara lebih presisi dibandingkan *U-Net* yang ringan namun kurang adaptif serta *nnDetection* yang tidak memodelkan detail *voxel-level*. Dengan demikian, *nnU-Net* direkomendasikan sebagai arsitektur pilihan paling optimal untuk implementasi aplikasi medis deteksi kanker prostat. Ke depan, riset dapat diperluas dengan memvalidasi model pada dataset lintas institusi serta modalitas MRI lain untuk menakar generalisasi, mengoptimalkan inferensi *real-time*, eksperimen penggabungan keluaran *nnU-Net* dengan data klinis seperti skor PSA atau genomik dalam kerangka *multimodal* demi meningkatkan akurasi diagnosis.

DAFTAR PUSTAKA

- [1] A. Villers, J. Haffner, and S. Bouye, "What is prostate cancer?," *Bull. Acad. Natl. Med.*, vol. 192, no. 5, 2008, doi: 10.1016/s0001-4079(19)32740-2.
- [2] J. S. Wefel, C. J. Ryan, J. Van, J. C. Jackson, and A. K. Morgans, "Assessment and Management of Cognitive Function in Patients with Prostate Cancer Treated with Second-Generation Androgen Receptor Pathway Inhibitors," *CNS Drugs*, vol. 36, no. 5, 2022. doi: 10.1007/s40263-022-00913-5.
- [3] H. Sung *et al.*, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA. Cancer J. Clin.*, vol. 71, no. 3, 2021, doi: 10.3322/caac.21660.
- [4] J. Ferlay *et al.*, "Cancer statistics for the year 2020: An overview," *Int. J. Cancer*, vol. 149, no. 4, 2021, doi: 10.1002/ijc.33588.
- [5] M. F. Leitzmann and S. Rohrmann, "Risk factors for the onset of prostatic cancer: Age, location, and behavioral correlates," *Clinical Epidemiology*, vol. 4, no. 1, 2012. doi: 10.2147/CLEP.S16747.
- [6] H. U. Ahmed *et al.*, "Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study," *Lancet*, vol. 389, no. 10071, 2017, doi: 10.1016/S0140-6736(16)32401-1.
- [7] J. C. Weinreb *et al.*, "PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2," *Eur. Urol.*, vol. 69, no. 1, 2016, doi: 10.1016/j.eururo.2015.08.052.
- [8] E. ALTINTAŞ *et al.*, "Comparison of Post Biopsy Pathology and Post Radical Prostatectomy Pathology in Patients with Prostate Cancer Detected After Fusion Biopsy," *Genel Tip Derg.*, vol. 33, no. 5, 2023, doi: 10.54005/geneltip.1314371.
- [9] Y. Li *et al.*, "MRI-based prostate cancer classification using 3D efficient capsule network," *Med. Phys.*, vol. 51, no. 7, 2024, doi: 10.1002/mp.16975.
- [10] T. Syer *et al.*, "Artificial intelligence compared to radiologists for the initial diagnosis of prostate cancer on magnetic resonance imaging: A systematic review and recommendations for future studies," *Cancers*, vol. 13, no. 13, 2021. doi: 10.3390/cancers13133318.
- [11] A. Comelli *et al.*, "Deep learning-based methods for prostate segmentation in magnetic resonance imaging," *Appl. Sci.*, vol. 11, no. 2, 2021, doi: 10.3390/app11020782.
- [12] M. He *et al.*, "Research progress on deep learning in magnetic resonance imaging-based diagnosis and treatment of prostate cancer: a review on the current status and perspectives," *Frontiers in Oncology*, vol. 13, 2023. doi: 10.3389/fonc.2023.1189370.
- [13] N. Ibtihaz and M. S. Rahman, "MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation," *Neural Networks*, vol. 121, 2020, doi: 10.1016/j.neunet.2019.08.025.
- [14] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," *Nat. Methods*, vol. 18, no. 2, 2021, doi: 10.1038/s41592-020-01008-z.
- [15] M. T. King *et al.*, "Characterizing the Performance of a Deep-Learning Segmentation Algorithm for Outlining Dominant Intraprostatic Lesions on Prostate Multi-Parametric MRI," *Int. J. Radiat. Oncol.*, vol. 114, no. 3, 2022, doi: 10.1016/j.ijrobp.2022.07.1177.
- [16] M. Baumgartner, P. F. Jäger, F. Isensee, and K. H. Maier-Hein, "nnDetection: A Self-configuring Method for Medical Object Detection," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2021. doi: 10.1007/978-3-030-87240-3_51.
- [17] G. Cacciamani *et al.*, "Assessment of a novel bpMRI-based machine learning framework to automate the detection of clinically significant prostate cancer using the PI-CAI (Prostate Imaging: Cancer AI) challenge dataset," *Eur. Urol.*, vol. 83, 2023, doi: 10.1016/s0302-2838(23)01017-5.
- [18] R. Viola, L. Gautheron, A. Habrard, and M. Sebban, "MetaAP: A meta-tree-based ranking algorithm optimizing the average precision from imbalanced data," *Pattern Recognit. Lett.*, vol. 161, 2022, doi: 10.1016/j.patrec.2022.07.019.
- [19] Y. Zhang, J. Chen, X. Ma, G. Wang, U. A. Bhatti, and M. Huang, "Interactive medical image annotation using improved Attention U-net with compound geodesic distance," *Expert Syst. Appl.*, vol. 237, 2024, doi: 10.1016/j.eswa.2023.121282.
- [20] C. Lee *et al.*, "Placental MRI segmentation based on multi-receptive field and mixed attention separation mechanism," *Comput. Methods Programs Biomed.*, vol. 242, 2023, doi: 10.1016/j.cmpb.2023.107699.
- [21] B. Andersen, A. Godfrey, J. Hou, and D. Kropaczek, "APPLICATION OF DEEP LEARNING NETWORKS TO SURROGATE MODELING OF CRUD DEPOSITION," in *Proceedings of the International Conference on Mathematics and Computational Methods Applied to Nuclear Science and Engineering, M and C 2021*, 2021. doi: 10.13182/M&C21-33930.
- [22] F. Isensee *et al.*, "nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation," in *Informatik aktuell*, 2019. doi: 10.1007/978-3-658-25326-4_7.
- [23] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment anything in medical images," *Nat. Commun.*, vol. 15, no. 1, 2024, doi: 10.1038/s41467-024-44824-z.
- [24] F. Provost, T. Fawcett, and R. Kohavi, "The case against accuracy estimation for comparing induction algorithms," in *International Conference on Machine Learning*, 1998.
- [25] S. Lonang, A. Yudhana, and M. K. Biddinika, "Rancangan Sistem Klasifikasi Kekurangan Gizi Balita Dengan Metode K-Nearest Neighbor," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 5, no. 1, 2023, doi: 10.36499/jinrpl.v5i1.7834.
- [26] O. J. Pellicer-Valero *et al.*, "Deep learning for fully automatic detection, segmentation, and Gleason grade estimation of prostate cancer in multiparametric magnetic resonance images," *Sci. Rep.*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-06730-6.
- [27] X. Huang *et al.*, "Application of U-Net Based Multiparameter Magnetic Resonance Image Fusion in the Diagnosis of Prostate Cancer," *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3061078.
- [28] A. Karagoz *et al.*, "Anatomically guided self-adapting deep neural network for clinically significant prostate cancer detection on bi-parametric MRI: a multi-center study," *Insights Imaging*, vol. 14, no. 1, 2023, doi: 10.1186/s13244-023-01439-0.