

Segmentasi Pemain Bola Dengan Arsitektur U-Net

¹ Fajar Bima Laksono, ² Muhammad Hannan Isnaen, ³ Risky Okta Wijaya
, ⁴ Muhammad Munsarif

^{1,2,3,4} Program Studi Informatika, Universitas Muhammadiyah Semarang, Semarang, Indonesia

^{1,2,3,4} Jl. Kedungmundu No. 18, Semarang, Jawa Tengah, Indonesia

¹ fajarbima09@gmail.com, ² hannanisnaen540@gmail.com, ³ riskyokta50@gmail.com,
⁴ m.munsarif@unimus.ac.id

ABSTRACT

Segmentation is a technique in digital image processing that focuses on the division of objects into parts and the separation of regions (objects) from the background. In this context, there are three significant types of image characteristics, namely points, lines, and edges. Image segmentation itself can be grouped into three main categories, namely object identification, semantic identification, and instant identification. In this research, the focus is on segmenting soccer players using a deep learning approach, specifically with the Convolutional Neural Network (CNN) method and U-Net architecture. CNN is one of the neural network methods in deep learning and machine learning that is good in terms of accuracy in image recognition, while U-Net is commonly used in semantic type image segmentation. Semantic segmentation, is an image that is divided into object and non-object categories. The soccer player segmentation process involves the image encoder and decoder stages before being utilized in the model training process, the testing stage involves the application of the CNN-U-Net model to perform image classification, resulting in an output consisting of 11 classes namely Goal Bar, Referee, Advertisement, Field, Ball, Coach & Officials, Spectator, Goalkeeper A, Goalkeeper B, Team A, and Team B The output will be evaluated by calculating accuracy to ensure model performance.

Keywords – Player, Football, Segmentation, CNN, U-Net

ABSTRAK

Segmentasi merupakan teknik pada pengolahan citra digital yang memfokuskan pada pembagian objek ke dalam beberapa bagian dan pemisahan antara region (objek) dengan latar belakang. Dalam konteks ini, Ada tiga jenis karakteristik gambar yang signifikan, yaitu titik, garis, dan tepi. Segmentasi citra sendiri dapat dikelompokkan menjadi tiga kategori utama, yaitu identifikasi objek, identifikasi semantik, dan identifikasi instan.. Pada penelitian ini, fokusnya adalah pada segmentasi pemain bola menggunakan pendekatan deep learning, khususnya dengan metode Convolutional Neural Network (CNN) dan arsitektur U-Net. CNN merupakan salah satu metode neural network pada deep learning dan machine learning yang baik dalam hal akurasi pada pengenalan citra, sedangkan U-Net biasa Digunakan pada segmentasi citra yang berjenis semantik. Segmentasi semantik, adalah citra yang dibagi menjadi kategori objek dan bukan objek. Proses segmentasi pemain bola melibatkan tahap encoder dan decoder citra sebelum dimanfaatkan dalam proses pelatihan model, tahap pengujian melibatkan penerapan model CNN-U-Net untuk melakukan klasifikasi citra, menghasilkan. yang terdiri dari 11 kelas yaitu Bilah Gawang, Wasit, Iklan, Lapangan, Bola, Pelatih & Oficial, Penonton, Kiper A, Kiper B, Tim A, dan Tim B Output tersebut akan dievaluasi dengan menghitung akurasi untuk memastikan performa model.

Kata Kunci – Pemain, Sepak Bola, Segmentasi, CNN, U-Net

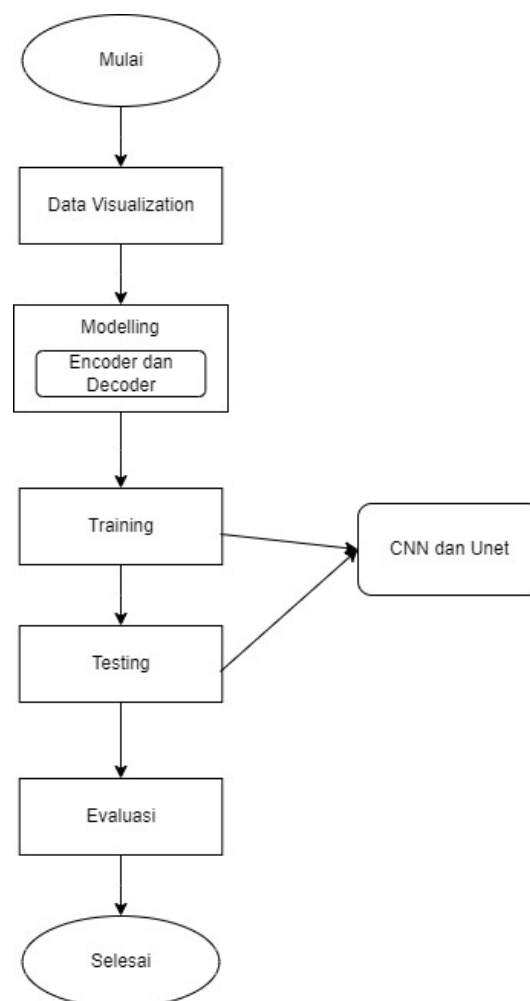
1. Introduction

Pemrosesan gambar digital telah menjadi bagian integral dari berbagai aplikasi, termasuk analisis video olahraga. Salah satu tantangan terbesar saat menganalisis video sepak bola adalah mengelompokkan pemain sepak bola berdasarkan latar belakang lapangannya. Segmentasi ini penting untuk memahami pergerakan pemain, mendukung analisis taktis, dan menghasilkan informasi berharga bagi pelatih dan analis olahraga. Segmentasi ini juga dapat digunakan untuk mendeteksi pelanggaran-pelanggaran yang terjadi di lapangan, dengan adanya deteksi tersebut maka dapat meminimalisir terjadinya kesalahan tuduhan pelanggaran oleh wasit.

Salah satu metode yang terbukti berhasil dalam tugas segmentasi citra semantik adalah arsitektur U-Net. U-Net adalah bentuk khusus Convolutional Neural Network (CNN) yang dirancang untuk menangani tugas segmentasi gambar dengan baik. U-Net adalah pilihan yang menjanjikan untuk mensegmentasi pemain sepak bola dalam konteks lapangan sepak bola yang dinamis karena kesadaran konteks spasial dan kemampuan untuk melestarikan detail lokal.

Penelitian ini mengusulkan untuk melakukan segmentasi pemain sepak bola dalam gambar sepak bola menggunakan arsitektur U-Net. Kami mengkaji proses segmentasi semantik yang membedakan pemain sepak bola dari elemen lain dalam sebuah gambar, seperti lapangan dan latar belakang. Melalui pendekatan deep learning ini diharapkan akurasi dan reliabilitas segmentasi pemain sepak bola dapat ditingkatkan.

2. Research Method



Gambar 1. Diagram alir metode penelitian.

Proses yang dilakukan dalam penelitian ini meliputi langkah-langkah berikut:

1) Data Visualization

Data visualization adalah representasi grafis dari data untuk membantu memahami pola, tren, dan informasi penting yang terkandung dalam dataset. Secara total dataset yang digunakan memiliki 100 Gambar dan 100 Label. Dataset yang digunakan adalah pertandingan Piala Super UEFA antara Real Madrid dan Manchester United pada tahun 2017 (Highlight).

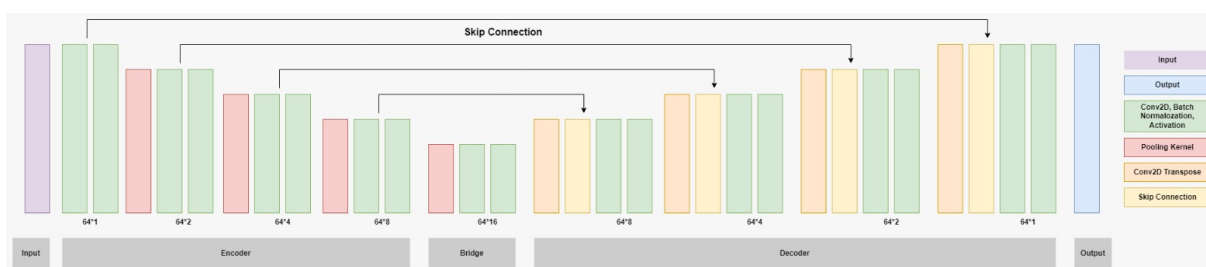
2) Modelling

Encoder dan Decoder, keduanya merupakan komponen utama dalam arsitektur U-Net atau dalam tugas segmentasi citra pada umumnya. Encoder dan decoder bekerja bersama untuk membuat arsitektur yang efektif dalam tugas segmentasi. Encoder mengambil citra input dan mengekstraksi fitur-fitur yang relevan, sedangkan decoder memproses informasi tersebut untuk menghasilkan output segmentasi yang sesuai.

3) Training dan Testing

Convolutional Neural Network (CNN) merupakan sebuah model jaringan saraf tiruan yang dikhususkan pada pengelolaan dan pemrosesan data gambar serta pola spasial lainnya. CNN terdiri dari berbagai lapisan yang tersusun secara hierarkis, termasuk lapisan konvolusi, pooling, dan pemahaman. Arsitektur CNN ini terinspirasi dari cara kerja visual manusia dan mampu secara efisien mengekstraksi fitur-fitur kunci dari citra untuk berbagai tugas seperti klasifikasi, deteksi objek, dan segmentasi. Dalam bidang pengenalan citra dan pengolahan citra komputer, CNN telah menjadi salah satu teknologi paling dominan dan efektif.

Arsitektur U-Net merupakan suatu struktur jaringan saraf tiruan yang dirancang khusus untuk melakukan segmentasi citra. U-Net terdiri dari dua komponen yaitu bagian encoder dan bagian decoder, yang saling terhubung. Bagian encoder bertugas untuk mengekstraksi fitur-fitur dari gambar asli, sedangkan bagian decoder bertanggung jawab untuk membangun kembali gambar dengan resolusi yang lebih tinggi berdasarkan fitur-fitur yang diekstraksi oleh bagian encoder. U-Net dikenal karena kemampuannya dalam menangani segmentasi citra dengan efisien, bahkan dengan menggunakan dataset yang terbatas.



Gambar 2. Model Arsitektur U-Net

4) Evaluasi

Dalam penelitian ini, dataset yang digunakan mencakup fitur `segmentation_mask`. Fitur ini menghasilkan gambar yang telah disegmentasi dengan akurat, dan digunakan untuk menghitung nilai loss serta *validation loss*. Nilai-nilai loss dan *validation loss* ini digunakan sebagai indikator untuk mengevaluasi efektivitas model dalam proses training. Terdapat tiga kemungkinan skenario:

1. Jika nilai *validation loss* meningkat dan *validation accuracy* menurun, ini menunjukkan kecenderungan underfitting.

2. Jika nilai *validation loss* meningkat dan *validation accuracy* juga meningkat, ini menunjukkan kecenderungan overfitting.
3. Jika nilai *validation loss* menurun dan *validation accuracy* meningkat, ini menunjukkan bahwa model berhasil belajar dengan baik.

3. Result and Analysis

Model yang dihasilkan oleh arsitektur U-Net adalah perbandingan citra yang dapat dilihat pada gambar (gambar 3)



Gambar 3. Perbandingan Original Image, Original Mask, Dan Predicted Mask

Berdasarkan hasil yang dapat dilihat pada gambar (Gambar 3) dimana pada kolom original image merupakan citra asli. Kolom Original Mask merupakan data dalam dataset yang digunakan untuk membandingkan hasil prediksi, telah terlihat pembagian menggunakan indikator warna seperti jingga, kuning, merah, abu, ungu, biru tua, dan biru muda. Warna jingga merupakan indikator Tim A dan Kiper A, warna kuning merupakan indikator Tim B kiper B, warna merah merupakan indikator bilah gawang, warna abu merupakan indikator lapangan bola, warna ungu merupakan indikator bola, warna biru tua merupakan indikator iklan, dan warna biru muda merupakan indikator Pelatih & Ofisial, Penonton . Hasil dari segmentasi yang menggunakan arsitektur U-Net biasa disebut Predicted Mask.

Di samping itu, terdapat evaluasi model yang melibatkan penghitungan nilai hasil, yang merupakan langkah untuk menilai seberapa tepatnya model yang telah dibangun dalam perbandingan hasil data

dengan data real. Hal ini bisa diamati di ilustrasi (Gambar 3), di mana proses evaluasi model dilakukan dengan menggunakan 50 epoch, sebagaimana terlihat dalam ilustrasi tersebut.

Table 1. Hasil Nilai Evaluasi

Epoch	loss	accuracy	val loss	val accuracy
1	0.7705	0.5322	60.3590	0.3578
2	0.6561	0.6170	175.1206	0.3549
3	0.6160	0.7308	35.7414	0.3561
4	0.6016	0.7496	47.2458	0.3566
5	0.5857	0.7786	16.8685	0.3513
6	0.5776	0.8295	3.5999	0.3459
7	0.5716	0.8330	2.9235	0.3573
8	0.5672	0.8236	2.3089	0.3584
9	0.5573	0.8325	0.8151	0.4354
10	0.5527	0.8501	0.6960	0.6806
11	0.5455	0.8560	0.7010	0.8639
12	0.5423	0.8614	0.6462	0.8364
13	0.5517	0.8561	0.6695	0.4546
14	0.5502	0.8725	0.6706	0.7232
15	0.5438	0.8551	0.7629	0.4253
16	0.5408	0.8823	0.7011	0.3123
17	0.5442	0.8875	0.9869	0.3152
18	0.5421	0.8969	0.7450	0.3106
19	0.5348	0.9058	0.6500	0.5296
20	0.5327	0.9049	0.6633	0.8325
21	0.5287	0.9127	0.6685	0.8833
22	0.5294	0.9173	0.6565	0.8863
23	0.5258	0.9243	0.6583	0.8872
24	0.5275	0.9226	0.6478	0.8871
25	0.5282	0.9244	0.6567	0.8919
26	0.5291	0.9301	0.6385	0.9007
27	0.5253	0.9320	0.6436	0.8948
28	0.5237	0.9369	0.6416	0.9077
29	0.5218	0.9435	0.6531	0.8928
30	0.5198	0.9421	0.6394	0.9044
31	0.5200	0.9463	0.6422	0.9039
32	0.5185	0.9481	0.6420	0.9063
33	0.5173	0.5173	0.6380	0.9107
34	0.5169	0.9496	0.6403	0.9005
35	0.5158	0.9536	0.6321	0.9132
36	0.5151	0.9564	0.6314	0.9149
37	0.5138	0.9592	0.6236	0.9261
38	0.5137	0.9607	0.6266	0.9201
39	0.5146	0.9611	0.6283	0.9190
40	0.5141	0.9594	0.6295	0.9201
41	0.5127	0.9598	0.6271	0.9218
42	0.5120	0.9637	0.6155	0.9337
43	0.5115	0.9651	0.6151	0.9318
44	0.5110	0.9657	0.6178	0.9349
45	0.5106	0.9683	0.6125	0.9388
46	0.5101	0.9686	0.6140	0.9382
47	0.5098	0.9672	0.6099	0.9412
48	0.5100	0.9694	0.6162	0.9374

49	0.5090	0.9708	0.6063	0.9456
50	0.5091	0.9711	0.6191	0.9356

Hasil perhitungan di Tabel 1 memperlihatkan bahwa terdapat perbedaan nilai antara loss, accuracy, validation loss, dan validation accuracy pada setiap epoch. Terlihat nilai loss dan validation loss cenderung menurun pada satu epoch ke epoch berikutnya, sementara nilai accuracy dan validation accuracy cenderung meningkat. Evaluasi pada Bab 2 menyimpulkan bahwa model segmentasi berhasil berjalan dengan baik, ditandai dengan penurunan nilai validation loss dan peningkatan nilai validation accuracy. Secara spesifik, nilai akhir validation loss mencapai 0,6191 (61,91%) dan nilai validation accuracy mencapai 0,9356 (93,56%), sedangkan nilai accuracy mencapai 0,9711 (97,11%).

4. Conclusion

Hasil uji segmentasi citra dengan arsitektur U-Net menyimpulkan:

1. Efektivitas arsitektur U-Net dalam mengerjakan segmentasi citra ditunjukkan pada tingkat akurasi mencapai 0.97, menggunakan 50 epoch untuk melatih model.
2. Evaluasi berdasarkan nilai validation loss dan validation accuracy mengindikasikan proses training model berjalan tanpa halangan.

References

- [1] RONNEBERGER, O., FISCHER, P. dan BROX, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).
- [2] VALUEVA, M. V., NAGORNOV, N.N., LYAKHOV, P.A., VALUEV, G. V. dan CHERVYAKOV, N.I., 2020. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. Mathematics and Computers in Simulation.
- [3] Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., ... & Ghayvat, H. (2021). CNN variants for computer vision: History, architecture, application, challenges and future scope. Electronics, 10(20), 2470.
- [4] Lui, M. S., Wijaya, E. K., & Hidayat, M. SEGMENTASI CITRA HEWAN DENGAN CONVOLUTIONAL NEURAL NETWORK ARSITEKTUR U-NET.
- [5] Mudzopar, I. M., & Wiharko, T. (2023). Pengembangan Sistem Deteksi Offside Berbasis Metode Yolo dalam Video Pertandingan Sepak Bola. Digital Transformation Technology, 3(2), 524-530.
- [6] Komorowski, J., Kurzejski, G., & Sarwas, G. (2019). Footandball: Integrated player and ball detector. arXiv preprint arXiv:1912.05445.
- [7] Annaffi, M. N., Putra, O. V., Harmini, T., & Trisnaningrum, N. (2022, November). Segmentasi Semantik pada Citra Hama Leafblast Menggunakan Unet dan Optimasi Hyperband. In Prosiding Seminar Sains Nasional dan Teknologi (Vol. 12, No. 1, pp. 453-459).
- [8] Aprilyanto, J., & Yohannes, Y. (2023, April). Implementasi Arsitektur VGG-Unet Dalam Melakukan Segmentasi Keretakan pada Citra Bangunan. In MDP Student Conference (Vol. 2, No. 1, pp. 257-264).
- [9] S. Roomy, "Football (semantic segmentation)," Kaggle, <https://www.kaggle.com/datasets/sadhliroomyprime/football-semantic-segmentation>
- [10] Michael Stephen Lui, M. D. (n.d.). ANALISIS DAMPAK TRANSFER LEARNING PADA SEGMENTASI SEMANTIK CITRA HEWAN MENGGUNAKAN U-NET.
- [11] Wahana, A., Firmansyah, E., Al Rosyid, H. I., Fuadi, R. S., & Maylawati, D. S. A. (2021). Fuzzy Tahani Method in the Recommendation System for Selecting Mountain Tourism Destinations in West Java.
- [12] Tamrin, M. A., Rizki, B., Nodas, A., Rahman, A., & Firmansyah, E. (2020). Perbandingan Penggunaan Metode Topsis dan Metode AHP dalam Penilaian Kinerja pada Karyawan (PT XYZ). Infoman's: Jurnal Ilmu-ilmu Informatika dan Manajemen, 14(1).