

Penerapan Arsitektur U-Net untuk Segmentasi Semantik Citra Buah dengan Latar Belakang Kompleks

Muh. Jamil^{*1}, Asiah Wati², Sitti Rahmah³, Aldi Bastiatul Fawait⁴

^{1,3,4} Fakultas Teknik dan Informatika, Program Studi Ilmu Komputer, Universitas Widya Gama Mahakam Samarinda, Samarinda 75243, Indonesia.

² Fakultas Pertanian, Program Studi Agroteknologi, Universitas Widya Gama Mahakam Samarinda, Samarinda 75243, Indonesia.

e-mail: ^{*1}jamil@uwgm.ac.id, ²asiahwati88@uwgm.ac.id, ³sitti.rahmah@uwgm.ac.id,

⁴aldi.bas.fawait@uwgm.ac.id

Abstrak

Perkembangan ilmu Computer Vision telah berhasil membuat terobosan baru di berbagai bidang, salah satunya adalah di bidang pertanian. Hal ini semakin membuka peluang bagi pertanian cerdas berbasis IoT untuk dapat diterapkan guna mengoptimalkan pekerjaan petani. Salah satu langkah awal yang bisa diterapkan adalah segmentasi citra, tujuannya adalah untuk mendeteksi area tertentu seperti buah, daun dan objek lain yang memiliki oklusi dan latar belakang kompleks pada lingkungan perkebunan secara langsung. Penelitian yang dikerjakan mengusulkan metode segmentasi semantik berbasis Deep Learning yaitu dengan memadukan ketangguhan dari U-Net dan juga VGG16 dalam melakukan proses segmentasi buah melon pada lingkungan perkebunan secara langsung. Model segmentasi yang dirancang terbukti mampu melakukan proses segmentasi dengan sangat optimal dengan score loss minimum sebesar 0.0157 dan score IoU maksimum sebesar 0.9922 pada proses train dan testing model, selain itu model yang dirancang mampu bekerja secara optimal dengan score IoU sebesar 0.9893 pada validasi akhir model. Penelitian yang dikerjakan bukan hanya berkontribusi sebagai rujukan akademis terkait dengan penerapan computer vision pada bidang pertanian cerdas, tetapi juga dapat berkontribusi menciptakan dataset baru bagi penelitian lanjut di masa mendatang.

Kata kunci— U-Net, Deep Learning, Segmentasi, Melon, IoT.

1. PENDAHULUAN

Segmentasi citra adalah teknik pengolahan citra digital yang dilakukan sebagai langkah awal untuk memisahkan objek dengan objek yang lain pada citra [1]. Tujuan pemisahan objek dapat berbeda-beda tergantung dari kebutuhan dan hasil akhir yang diinginkan. Pada penelitian terdahulu teknik segmentasi banyak digunakan sebelum dilakukan proses ekstraksi fitur [2] [3], [4] tujuannya adalah agar fitur yang diekstrak dapat lebih spesifik sehingga dapat mempermudah proses klasifikasi citra. Metode segmentasi yang dilakukan bisa bersifat manual [5] ataupun otomatis menggunakan algoritma tradisional [6] hingga ke algoritma yang lebih cerdas berbasis machine learning [7].

Teknik segmentasi citra tidak hanya digunakan untuk *improvement* hasil klasifikasi [8], [9], namun dapat juga digunakan secara spesifik untuk tugas deteksi objek pada citra medis [10], citra tanaman [11] hingga ke tugas yang lebih kompleks seperti segmentasi objek oleh mobil *autonomous* [12]. Hal ini menunjukkan bahwa proses segmentasi citra merupakan salah satu tahap yang penting dalam menyelesaikan permasalahan tertentu pada bidang pengolahan citra digital ataupun pada bidang *computer vision*.

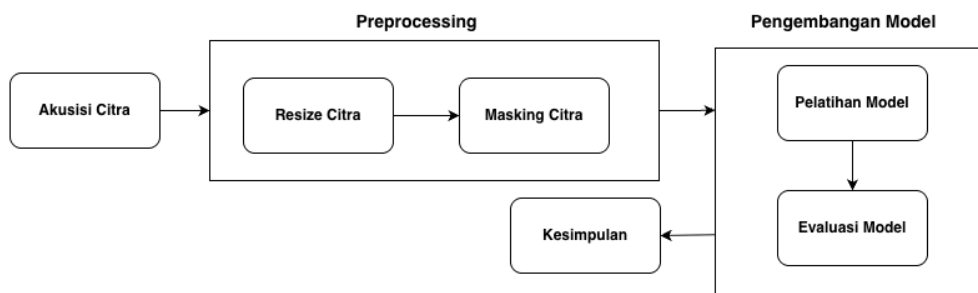
Dengan semakin populernya penggunaan teknologi *Internet of Thing (IoT)* pada bidang pertanian. Maka penelitian terkait teknik segmentasi banyak digunakan sebagai langkah awal dalam mendeteksi objek tanaman seperti pada penelitian [13] menggunakan teknik segmentasi berbasis *Deep Learning* untuk membuang objek yang tidak dibutuhkan sebelum melakukan klasifikasi bibit tanaman. Ada juga penelitian [14], [15] yang telah memodifikasi arsitektur dari *U-Net* untuk mendapatkan akurasi segmentasi yang optimal pada bidang pertanian. Selain itu *review* penelitian mengenai segmentasi objek pada bidang pertanian juga sudah pernah dilakukan di antaranya adalah penelitian [16], [17]. Menurut penelitian tersebut metode *Deep Learning* dapat melakukan lokalisasi yang lebih baik pada kondisi citra yang memiliki kontras, warna, oklusi dan latar belakang yang lebih bervariasi dibandingkan dengan metode segmentasi tradisional.

Meskipun belakangan ini terdapat metode yang lebih baru untuk melakukan segmentasi dan deteksi objek dengan menggunakan algoritma *You Look Only Once (YOLO)* [18]. Namun berdasarkan hasil penelitian yang dilakukan oleh [19] menunjukkan bahwa *U-Net* lebih unggul untuk tugas segmentasi objek yang lebih presisi pada area piksel. Bahkan penelitian yang dikerjakan oleh [20],[21] tetap menggunakan *U-Net* untuk melakukan segmentasi akhir setelah area sekitar *foreground* berhasil diberi *bounding box* oleh algoritma *YOLO*.

Maka dari itu penelitian yang dikerjakan mengusulkan metode *Deep Learning* untuk segmentasi objek buah melon di lapangan. Hal ini dilakukan sebagai langkah awal untuk memulai penelitian pertanian cerdas berbasis *IoT* yang dilaksanakan di Balai Pengembangan SDM Pertanian, Provinsi Kalimantan Timur (BPSDMP Kaltim) [22]. Metode yang diusulkan adalah *U-Net* yang merupakan metode *semantic* segmentasi berbasis *Deep Learning* [23]. Model *U-Net* yang dirancang menggunakan *pretrain VGG16* sebagai *encoder* agar model dapat memiliki kemampuan belajar yang optimal ketika mendeteksi peta fitur pada citra. Perancangan model *U-Net* yang dikerjakan terinspirasi dari penelitian sebelumnya yang dikerjakan oleh [24], [25], [26] di mana lapisan *encoder* pada *U-Net* dimodifikasi menggunakan *pretrain model VGG16*.

2. METODE PENELITIAN

Penelitian yang dikerjakan menggunakan beberapa tahapan yang ditunjukkan pada Gambar 1.



Gambar 1. Alur Metode Penelitian

2.1. Metode Akuisisi Citra

Akuisisi citra adalah proses mengambil Dataset citra buah melon di kebun *hydroponic* BPSDMP Kaltim. Citra buah melon yang diambil adalah dengan kondisi buah masih berada di pohon menggunakan kamera ponsel cerdas. Tahap akuisisi citra dikerjakan mulai dari bulan Mei 2024 sampai dengan September 2025.

2.2. Metode Resize Citra

Resize adalah proses mengubah ukuran citra ke ukuran yang lebih kecil. Hal ini dimaksudkan untuk mengurangi beban komputasi yang dibutuhkan oleh komputer ketika

melakukan proses pelatihan model. Agar detail citra tidak hilang namun tetap memiliki dimensi yang kecil, maka resolusi citra yang digunakan adalah 256 x 256 piksel.

2.3. Metode Masking Citra

Masking citra adalah proses menandai area tertentu pada citra. Fungsi citra *masking* adalah sebagai *groundtruth* atau label piksel, sehingga nilai setiap piksel citra *rgb* adalah input dan nilai setiap piksel citra *masking* adalah target. Karena model segmentasi yang akan dirancang adalah *binary segmentation* maka *masking* yang dibuat hanya memiliki 2 warna yaitu warna hitam dengan nilai piksel = 0 untuk *background* dan oklusi sementara warna putih dengan nilai piksel = 1 untuk *foreground* (objek buah melon). Citra *masking* dibuat menggunakan perangkat lunak *Adobe Photoshop* karena memiliki presisi piksel yang baik.

2.4. Metode Perancangan Model

Perancangan model yang digunakan pada arsitektur *U-Net* menggunakan *pretrain/backbone VGG16* sebagai lapisan *encoder*. Lapisan *encoder* adalah bagian pada arsitektur *U-Net* yang berfungsi untuk melakukan ekstraksi fitur serta, *downsampling* dan menangkap setiap detail fitur pada citra [27]. Setelah detail citra ditangkap oleh *encoder* maka selanjutnya informasi pada citra akan masuk ke bagian *skip connection*. *Skip connection* adalah jembatan antara bagian *encoder* dan *decoder*, sehingga *skip connection* akan mengambil setiap detail fitur pada citra yang berasal dari *encoder* untuk dimasukkan ke dalam blok *decoder* [28]. Pada bagian *decoder* inilah proses *up sampling* dan pemetaan fitur pada citra dilakukan.

2.5. Metode Pelatihan Model

Skenario pelatihan model yang dikerjakan akan menggunakan *optimizer Adam* [29] dengan *learning rate* = 0.0001, *loss* = *binary_crossentropy*. Selanjutnya nilai *epoch* akan diatur sebanyak 50 kali namun dengan mekanisme *early stop* = 3, *monitor*=*loss_validation*, *save_best*=*true*, artinya jika tidak ada *improvement* pada *loss validation* sebanyak 3 kali secara berturut turut maka proses pelatihan akan otomatis berhenti dan bobot terbaik akan disimpan sehingga memperkecil kemungkinan *overfit* model. Selanjutnya dataset yang telah diproses pada tahap sebelumnya dipisah menjadi 3 bagian yang berbeda yang ditunjukkan pada Tabel 1.

Tabel 1. Proporsi Pembagian Dataset

Data	Input Image	Groundtruth Image
Train	428 <i>rgb images</i>	428 <i>binary images</i>
Test	107 <i>rgb images</i>	107 <i>binary images</i>
Validation	15 <i>rgb images</i>	15 <i>binary images</i>

Berdasarkan data yang disajikan pada Tabel 1 menunjukkan bahwa *U-Net* membutuhkan 2 jenis input citra yaitu dalam bentuk *rgb* ataupun boleh dalam bentuk *grayscale*. Sementara untuk citra labelnya membutuhkan image dalam bentuk *binary* ataupun boleh *grayscale* jika segmentasi yang dirancang adalah *multiclass segmentation*. Selanjutnya untuk memantau kinerja proses pelatihan model akan digunakan 2 metrik yaitu *Intersection Over Union (IoU)* dan juga *Binary Cross Entropy (BCE)*.

IoU bekerja dengan cara membandingkan rasio kesamaan nilai piksel antara citra *masking* (*groundtruth*) dengan citra hasil prediksi [30]. Secara matematis *IoU* akan melakukan proses pembagian antara luas area yang tumpang tindih dengan total luas area gabungan [31] sehingga rumus dari metrik *IoU* ditunjukkan pada Persamaan 1.

$$IoU = \frac{TP}{TP + FP + FN} \quad (1)$$

Selain menggunakan metrik *IoU* digunakan juga metrik lain untuk menghitung *loss* ketika proses pelatihan model dilakukan. Metrik *loss function* yang digunakan adalah *BCE* (*Binary Cross Entropy*), metode ini cukup populer digunakan pada kasus *binary segmentation*. *BCE* akan mengukur seberapa besar perbedaan antara probabilitas prediksi model dengan label sebenarnya (0 atau 1) [32]. *BCE* menggunakan rumus yang ditunjukkan pada Persamaan 2.

$$L = -(y \log(p) + (1 - y) \log(1 - p)) \quad (2)$$

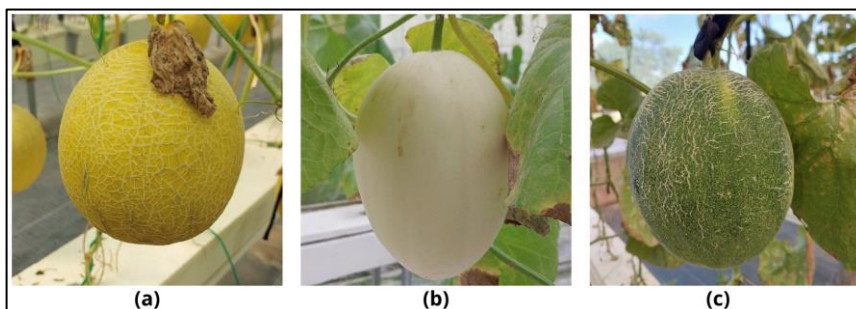
2.6. Metode Evaluasi Kinerja Model

Tahap evaluasi akan menggunakan model sebelumnya yang telah dilatih dan diuji untuk memprediksi data *validation*. Terdapat 15 data *validation* yang telah disediakan dalam bentuk *rgb image* dan juga *binary image* sebagai *masking*. Kinerja segmentasi pada data *validation* akan diamati menggunakan metrik *IoU*.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Akuisisi Citra

Pengambilan dataset citra melon yang telah dikerjakan di BPSDMP Kaltim menghasilkan 550 dataset citra buah melon yang terdiri dari 3 varietas melon yaitu Inthanon sebanyak 200 citra, Kirani sebanyak 200 citra dan The Blues sebanyak 150 citra. Sampel citra melon yang telah dikumpulkan disajikan pada Gambar 2.



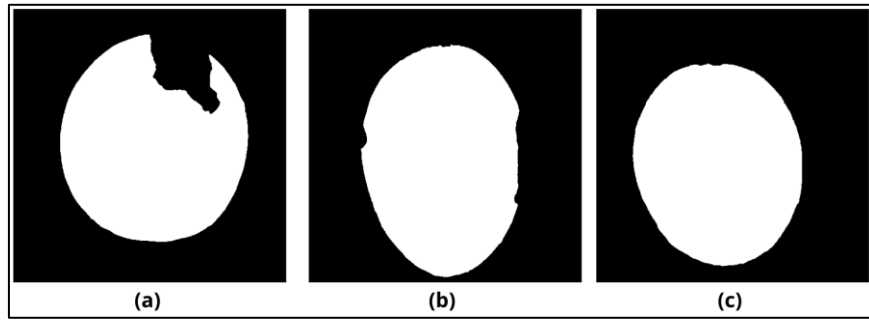
Gambar 2. Sampel dataset (a). Inthanon, (b). Kirani, (c). The Blues

3.2 Hasil Resize Citra

Dataset citra melon yang telah berhasil dikumpulkan masih memiliki dimensi yang sangat besar yaitu 2992x2992 piksel dengan ukuran data rata-rata sebesar 2,4 *Megabyte*. *Resize* citra dikerjakan menggunakan pustaka *Open CV Python*, hasil akhir dari proses *resize* adalah citra dengan resolusi 256x256 piksel dengan ukuran data rata-rata sebesar 70 *Kilobyte*.

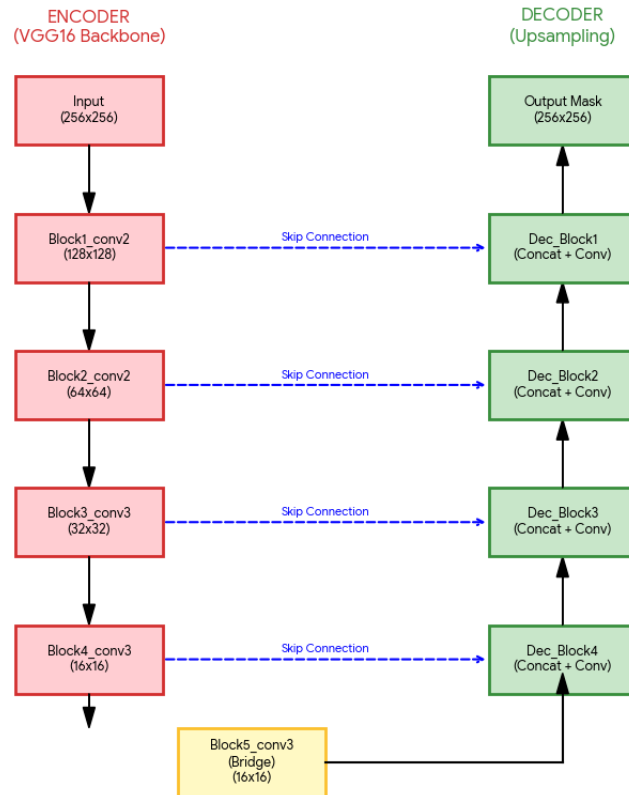
3.3 Hasil Masking Citra

Setelah proses *resize* citra selesai dikerjakan maka setiap dataset citra *rgb* diolah menggunakan *software Adobe Photoshop*. Setiap objek buah melon pada citra diseleksi lalu diberi *masking* berwarna putih, lalu untuk bagian *background* dan oklusi diberi *masking* berwarna hitam. Hasil akhir dari 550 citra *rgb* menghasilkan 550 citra *masking*, sehingga total keseluruhan citra *rgb* dan *masking* adalah 1100 citra. Sampel hasil *masking* citra ditunjukkan pada Gambar 3.

Gambar 3. Sampel dataset *masking* (a). Inthanon, (b). Kirani, (c). The Blues

3.4 Hasil Perancangan Model

Perancangan model *U-Net* yang dikerjakan menggunakan *VGG16* sebagai *block encoder* ditampilkan pada Gambar 4.

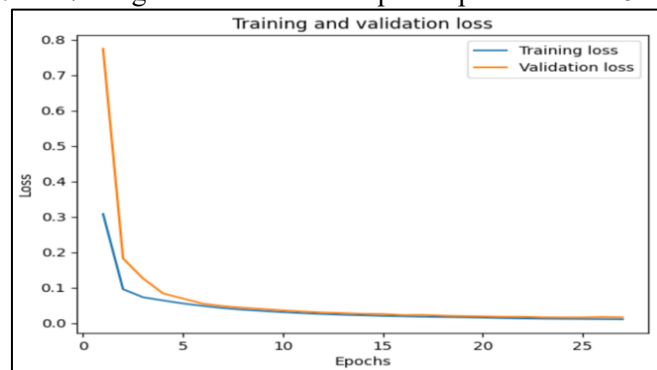
Gambar 4. Hasil Rancangan Arsitektur *U-Net*.

Berdasarkan desain pada Gambar 4 menunjukkan bahwa jalur *encoder* terdiri dari 5 blok ekstraksi fitur. Hasil keluaran dari empat blok pertama (*block1_conv2*, *block2_conv2*, *block3_conv3*, dan *block4_conv3*) dihubungkan dengan *skip connection* ke *decoder*. Sementara blok ke lima (*block5_conv3*) berfungsi sebagai jembatan atau *bottleneck*. Karakteristik utama dari *VGG16* adalah adanya *maxpooling* pada akhir setiap blok untuk memperkecil dimensi spasial citra secara bertahap.

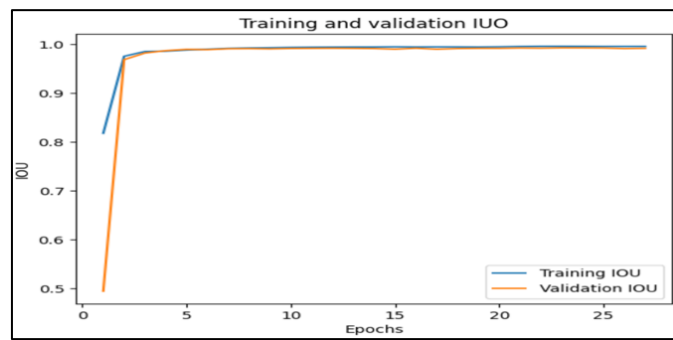
3.5 Hasil Pelatihan Model

Proses pelatihan model dikerjakan menggunakan *notebook* dari *Kaggle* [33]. Alasan penggunaan *notebook* ini adalah karena memiliki akselerasi yang cepat dengan dukungan *GPU*

P100 namun dapat diakses secara gratis. Dari 50 *epoch* yang telah ditentukan, proses pelatihan berhenti pada *epoch* ke 27 dengan visualisasi ditampilkan pada Gambar 5 dan Gambar 6.



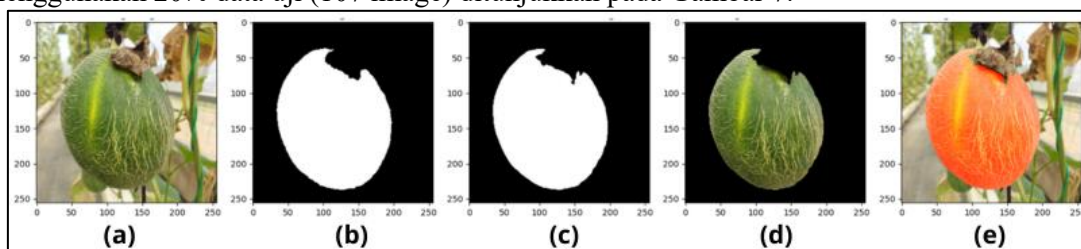
Gambar 5. Train dan Validation Loss.



Gambar 6. Train dan Validation IoU

Berdasarkan visualisasi pada Gambar 5 menunjukkan nilai *train loss* dan *validation loss* dari proses pelatihan. Data yang digunakan adalah data *train* (80%) dan *test* (20%), berdasarkan visualisasi tersebut dapat diamati bahwa dari *epoch* 1 sampai dengan 24 model terus mengalami *improvement* secara perlahan. Akan tetapi pada *epoch* 25 sampai 27 nilai *loss* sudah tidak lebih rendah dari sebelumnya secara berulang sebanyak 3 kali sehingga model berhenti melakukan proses pelatihan. Visualisasi pada Gambar 5 juga menunjukkan bahwa nilai *train loss* dan *validation loss* turun secara linear dengan nilai yang sangat dekat sehingga menghasilkan *validation loss* terbaik di angka 0.0157.

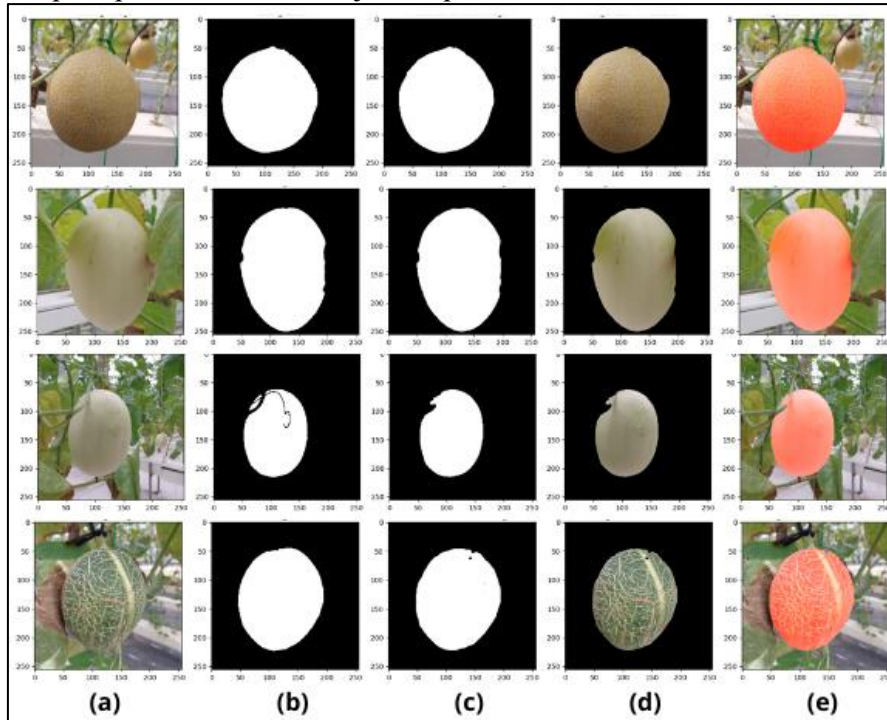
Metrik lain yang menjadi perhatian adalah *IoU* yang disajikan pada Gambar 6. Berdasarkan visualisasi tersebut *train IoU* dan *validation IoU* mengalami kenaikan yang sangat signifikan pada epoch 1 sampai dengan 3, namun pada *epoch* selanjutnya *score IoU* pada model mengalami kenaikan secara perlahan namun konsisten. Hingga pada akhirnya proses *train* berhenti pada *epoch* ke 27 karena nilai *loss* pada Gambar 5 sudah tidak mengalami *improvement*. Sehingga model akan menyimpan bobot terbaik selama proses pelatihan dengan nilai *validation IoU* terbaik di angka 0.9922. Sampel hasil prediksi model segmentasi pada proses *train* menggunakan 20% data uji (107 image) ditunjukkan pada Gambar 7.



Gambar 7. Sampel Hasil Prediksi Model (a). citra input, (b). citra label, (c). hasil prediksi, (d). overlay citra input dan prediksi, (e). overlay citra input dan masking prediksi.

3.6 Hasil Evaluasi Model

Model yang telah dilatih pada proses *train* dan *testing* kemudian diuji kembali menggunakan data *validation* sebanyak 15 citra. Berdasarkan hasil validasi model yang dilakukan, maka model mampu melakukan prediksi dengan *score IoU* sebesar 0.9893 yang artinya model masih mampu melakukan prediksi segmentasi citra dengan sangat baik. Sampel hasil prediksi pada proses validasi ditunjukkan pada Gambar 8.



Gambar 8. Sampel Hasil Uji Validasi Model (a). citra input, (b). citra label, (c). hasil prediksi, (d). *overlay* citra input dan prediksi, (e). *overlay* citra input dan *masking* prediksi.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dikerjakan maka dapat ditarik beberapa Kesimpulan sebagai berikut:

1. Pengumpulan dataset yang dikerjakan di BPSDMP Kaltim telah menghasilkan dataset citra buah melon. Dataset yang telah dikumpulkan dan diolah tentunya dapat berkontribusi pada penelitian selanjutnya dengan topik yang sama.
2. Penelitian yang dikerjakan adalah sebuah langkah awal dalam memulai penelitian berbasis *IoT* di lingkungan *in-situ* dengan pendekatan berupa segmentasi citra berbasis *Deep Learning*.
3. Implementasi model *U-Net* yang dipadukan dengan *backbone VGG16* pada lapisan *encoder* mampu bekerja dengan sangat optimal untuk melakukan proses segmentasi citra buah melon dengan kondisi latar belakang yang kompleks disertai dengan beberapa oklusi dengan *score validation loss* sebesar 0.0157 dan *score IoU* optimal sebesar 0.9922. Selain itu metrik performansi yang digunakan mampu mengukur kemampuan model *U-Net* yang dibuktikan dengan *score IoU* sebesar 0.9893 pada proses validasi akhir model.

4. Penelitian yang dikerjakan berkontribusi secara akademis dan dapat menjadi rujukan terkait penerapan ilmu *Computer Vision* pada Ilmu Pertanian berbasis teknologi *IoT* bagi peneliti selanjutnya.

5. SARAN

Meskipun penelitian yang dikerjakan telah mampu menghasilkan model yang mampu mendeteksi buah melon pada lingkungan kompleks namun masih terdapat beberapa cela untuk dapat dilakukan penelitian lanjutan yaitu:

1. *U-Net* hanyalah salah satu dari sekian banyak model segmentasi yang populer digunakan. Sehingga masih banyak pilihan lain yang dapat diuji oleh peneliti selanjutnya dengan studi kasus yang sama.
2. Segmentasi hanyalah proses awal untuk melakukan pemisahan objek, sehingga masih dapat dilakukan penelitian lanjutan berupa klasifikasi citra ataupun *object detection* menggunakan algoritma yang lebih *robust*.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada pihak UWGM (Universitas Widya Gama Mahakam Samarinda) atas dukungan finansial pada penelitian yang telah dikerjakan. Ucapan terimakasih juga disampaikan kepada pihak BPSDMP Kaltim atas kontribusi dataset yang mendukung terlaksananya proses penelitian dan juga kepada rekan dosen di UWGM atas dukungan ilmu dan moral yang telah diberikan.

DAFTAR PUSTAKA

- [1] F. A. Ahmad Khairul Anuar, J. Jone, R. F. A. Raja Azhar, and A. K. Jumaat, "Hybrid deep learning and active contour for segmenting hazy images," *IAES International Journal of Robotics and Automation (IJRA)*, vol. 14, no. 3, p. 429, Sep. 2025, doi: 10.11591/ijra.v14i3.pp429-437.
- [2] H. Haviluddin, M. Wati, R. Alfred, A. E. Burhandenny, and A. A. Pratama, "An extraction of shapes and support vector machine methods for identification of decorative wall 'Lamin' motifs of the Dayak Kenyah Pampang tribe," *International Journal of Artificial Intelligence Research*, vol. 7, no. 1, p. 1, Jun. 2023, doi: 10.29099/ijair.v7i1.475.
- [3] A. R. Khan, S. Khan, M. Harouni, R. Abbasi, S. Iqbal, and Z. Mehmood, "Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification," *Microsc. Res. Tech.*, vol. 84, no. 7, pp. 1389–1399, Jul. 2021, doi: 10.1002/jemt.23694.
- [4] M. Shahzad *et al.*, "Blood cell image segmentation and classification: a systematic review," *PeerJ Comput. Sci.*, vol. 10, pp. 1–35, 2024, doi: 10.7717/peerj-cs.1813.
- [5] M. M. Philip *et al.*, "Comparison of semi-automatic and manual segmentation methods for tumor delineation on head and neck squamous cell carcinoma (HNSCC) positron emission tomography (PET) images," *Phys. Med. Biol.*, vol. 69, no. 9, May 2024, doi: 10.1088/1361-6560/ad37ea.
- [6] A. Septiarini, H. Hamdani, H. R. Hatta, and K. Anwar, "Automatic image segmentation of oil palm fruits by applying the contour-based approach," *Sci. Hortic.*, vol. 261, p. 108939, Feb. 2020, doi: 10.1016/J.SCIENTA.2019.108939.

- [7] A. Yang, Y. Bai, H. Liu, K. Jin, T. Xue, and W. Ma, "Application of SVM and its Improved Model in Image Segmentation," *Mobile Networks and Applications*, vol. 27, no. 3, pp. 851–861, Jun. 2022, doi: 10.1007/S11036-021-01817-2/METRICS.
- [8] T. Agrawal and P. Choudhary, "Segmentation and classification on chest radiography: a systematic survey," *Visual Computer*, vol. 39, no. 3, pp. 875–913, Mar. 2023, doi: 10.1007/S00371-021-02352-7/METRICS.
- [9] T. H. Nguyen, T. N. Nguyen, and B. V. Ngo, "A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease," *AgriEngineering*, vol. 4, no. 4, pp. 871–887, Dec. 2022, doi: 10.3390/agriengineering4040056.
- [10] Z. Barzegar and M. Jamzad, "WLFS: Weighted label fusion learning framework for glioma tumor segmentation in brain MRI," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, doi: 10.1016/j.bspc.2021.102617.
- [11] X. Ni, C. Li, H. Jiang, and F. Takeda, "Deep learning image segmentation and extraction of blueberry fruit traits associated with harvestability and yield," *Hortic. Res.*, vol. 7, no. 1, Dec. 2020, doi: 10.1038/s41438-020-0323-3.
- [12] S. Abdigapporov, S. Miraliev, V. Kakani, and H. Kim, "Joint Multiclass Object Detection and Semantic Segmentation for Autonomous Driving," *IEEE Access*, vol. 11, pp. 37637–37649, 2023, doi: 10.1109/ACCESS.2023.3266284.
- [13] S. Kamal *et al.*, "IOT Automation with Segmentation Techniques for Detection of Plant Seedlings in Agriculture," *Wirel. Commun. Mob. Comput.*, vol. 2022, 2022, doi: 10.1155/2022/6466555.
- [14] H. Wang *et al.*, "MFBP-UNet: A Network for Pear Leaf Disease Segmentation in Natural Agricultural Environments," *Plants*, vol. 12, no. 18, Sep. 2023, doi: 10.3390/plants12183209.
- [15] Changgeng Yu, D. Lin, and C. He, "ASE-UNet: An Orange Fruit Segmentation Model in an Agricultural Environment Based on Deep Learning," *Optical Memory and Neural Networks (Information Optics)*, vol. 32, no. 4, pp. 247–257, Dec. 2023, doi: 10.3103/S1060992X23040045/METRICS.
- [16] L. Lei, Q. Yang, L. Yang, T. Shen, R. Wang, and C. Fu, "Deep learning implementation of image segmentation in agricultural applications: a comprehensive review," *Artif. Intell. Rev.*, vol. 57, no. 6, Jun. 2024, doi: 10.1007/s10462-024-10775-6.
- [17] I.-S. Oh, "Review of Fruit Tree Image Segmentation," Dec. 2024. Accessed: Oct. 27, 2025. [Online]. Available: <https://arxiv.org/pdf/2412.14631>
- [18] C. H. Kang and S. Y. Kim, "Real-time object detection and segmentation technology: an analysis of the YOLO algorithm," *JMST Advances 2023 5:2*, vol. 5, no. 2, pp. 69–76, Sep. 2023, doi: 10.1007/S42791-023-00049-7.
- [19] C. Beldek, E. Sariyildiz, and G. Alici, "A Deployment-Oriented Benchmarking of You Look Only Once (YOLO) Models for Orange Detection and Segmentation in Agricultural Robotics," *Agriculture*, vol. 15, no. 20, p. 2170, Oct. 2025, doi: 10.3390/agriculture15202170.
- [20] B. Bai, J. Tian, S. Luo, and T. Wang, "YUSEG: Yolo and Unet is all you need for cell instance segmentation."

- [21] N. Iriawan *et al.*, “YOLO-UNet Architecture for Detecting and Segmenting the Localized MRI Brain Tumor Image,” *Applied Computational Intelligence and Soft Computing*, vol. 2024, 2024, doi: 10.1155/2024/3819801.
- [22] UPTD BPPSDMP Provinsi Kalimantan Timur, “UPTD BPPSDMP Provinsi Kalimantan Timur.”
- [23] M. A. Bhatti *et al.*, “Utilizing convolutional neural networks (CNN) and U-Net architecture for precise crop and weed segmentation in agricultural imagery: A deep learning approach,” *Big Data Research*, vol. 36, p. 100465, May 2024, doi: 10.1016/J.BDR.2024.100465.
- [24] X. Yi *et al.*, “U-Net with Coordinate Attention and VGGNet: A Grape Image Segmentation Algorithm Based on Fusion Pyramid Pooling and the Dual-Attention Mechanism,” *Agronomy*, vol. 14, no. 5, May 2024, doi: 10.3390/agronomy14050925.
- [25] X. Li, J. Fang, and Y. Zhao, “A Multi-Target Identification and Positioning System Method for Tomato Plants Based on VGG16-UNet Model,” *Applied Sciences (Switzerland)*, vol. 14, no. 7, Apr. 2024, doi: 10.3390/app14072804.
- [26] D. Riana, M. Jamil, S. Hadianiti, J. Na’am, H. Sutanto, and R. Sukwadi, “Model of Watershed Segmentation in Deep Learning Method to Improve Identification of Cervical Cancer at Overlay Cells,” *TEM Journal*, vol. 12, no. 2, pp. 813–819, May 2023, doi: 10.18421/TEM122-26.
- [27] X.-C. Tai, H. Liu, R. H. Chan, and L. Li, “A Mathematical Explanation of UNet,” *ArXiv*, Oct. 2024, doi: <https://doi.org/10.48550/arXiv.2410.04434>.
- [28] S. T. Tran, C. H. Cheng, T. T. Nguyen, M. H. Le, and D. G. Liu, “Tmd-unet: Triple-unet with multi-scale input features and dense skip connection for medical image segmentation,” *Healthcare (Switzerland)*, vol. 9, no. 1, Jan. 2021, doi: 10.3390/healthcare9010054.
- [29] “Adam.” Accessed: Apr. 04, 2026. [Online]. Available: <https://keras.io/api/optimizers/adam/>
- [30] A. N. Gajjar and J. Jethva, “Intersection over Union based analysis of Image detection/segmentation using CNN model,” *ICPC2T 2022 - 2nd International Conference on Power, Control and Computing Technologies, Proceedings*, 2022, doi: 10.1109/ICPC2T53885.2022.9776896.
- [31] R. Wang, M. Cai, Z. Xia, and Z. Zhou, “Remote Sensing Image Road Segmentation Method Integrating CNN-Transformer and UNet,” *IEEE Access*, vol. 11, pp. 144446–144455, 2023, doi: 10.1109/ACCESS.2023.3344797.
- [32] M. S. Neyestanak *et al.*, “A Quantitative Comparison between Focal Loss and Binary Cross-Entropy Loss in Brain Tumor Auto-Segmentation Using U-Net,” *Journal of Biostatistics and Epidemiology*, vol. 11, no. 1, pp. 15–35, 2025, doi: 10.18502/jbe.v11i1.19315.
- [33] “Kaggle Notebooks.” Accessed: Apr. 06, 2026. [Online]. Available: <https://www.kaggle.com/code>