Pengaruh *Underwater Image Enhancement* terhadap Peningkatan mAP50-95 YOLOv7 pada dataset Lobster Aruna

1st Farhan Nur Hakim
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia
farhannurhakim@student.telkomuniversity.ac.id

2nd Ledya Novamizanti
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia
ledyaldn@telkomuniversity.ac.id

3rd Fityanul Akhyar
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia
fityanul@telkomuniversity.ac.id

***Abstrak* — Pada proyek utama dari jurnal ini, yaitu pembuatan *prototype* alat untuk monitoring lobster air laut di Aruna, tentu memiliki beberapa permasalahan, salah satunya adalah tidak jelasnya tampilan video di dalam laut, diantaranya disebabkan oleh *noise* dari air laut, kurangnya pencahayaan, dan dominasi warna tertentu dalam laut yang bisa menyebabkan hasil mAP dari model YOLOv7 kurang maksimal, hal itu dikarenakan penangkaran lobster air laut disimpan dalam kedalaman sekitar 15 meter, sehingga memunculkan permasalahan yang telah disebutkan tadi. Berdasarkan permasalahan tersebut penulis berinisiatif untuk menambahkan *underwater image enhancement* pada sistem monitoring lobster ini agar dapat memaksimalkan hasil mAP dari model YOLOv7 yang telah di *train*. Metode *Image Enhancement* yang diusulkan untuk dilakukan tahap pengujian adalah *Deep Learning and Image Formation Model* dan *Water-Net*. Dari kedua *image enhancement* tersebut dibandingkan hasil mAP-nya pada proses *testing* dengan hasil tanpa *image enhancement* serta sesama metode *image enhancement*-nya. Dari hasil pengujian tersebut didapatkan mAP50-95 yang terbaik adalah menggunakan metode *image enhancement Deep Learning and Image Formation Model* dengan perolehan mAP50-95 0.461. Maka dari itu metode terbaik ini yang dipakai untuk proses *image enhancement prototype* sistem monitoring lobster pada Tambak Aruna.**

***Kata kunci* — image enhancement, YOLOv7, Deep Learning and Image Formation Model, Water-Net, lobster**

I. PENDAHULUAN

Selama beberapa tahun terakhir, telah ada perhatian yang signifikan difokuskan pada *image enhancement* dalam pemrosesan gambar bawah air [11], [12]. Memperbaiki gambar bawah air merupakan hal yang menantang karena kompleksitas lingkungan bawah air dan kondisi pencahayaan. Biasanya, gambar-gambar bawah air mengalami degradasi yang disebabkan oleh *absorption* dan *scattering* yang tergantung pada panjang gelombang, termasuk *forward and backward scattering* [6]. Selain itu, partikel-partikel yang ada di dalam laut seperti plankton dan bakteri bisa menambahkan *noise* dan memperburuk *scattering effect*. Faktor-faktor yang tidak diinginkan ini menyebabkan visibilitas berkurang, kontras berkurang, dan bahkan timbulnya warna yang tidak wajar, sehingga dapat mengurangi kualitas gambar-gambar bawah air dari video dalam berbagai bidang seperti biologi dan arkeologi kelautan [13], ekologi kelautan [14], dan lain-lain [15]. Terdapat beberapa penelitian juga yang sangat membutuhkan *image enhancement* seperti klasifikasi ikan dalam air [17] dan sistem inspeksi cacat pada permukaan kayu [16], [18]. Hal ini menunjukan betapa pentingnya *image enhancement* dalam *image processing* dan *computer vision*.

Dalam proses monitoring lobster menggunakan kamera yang ditanam pada kandang lobster 15 meter di bawah laut, tentu memiliki beberapa permasalahan di bawah air yang telah disebutkan, sehingga *output* yang berupa video secara visual akan berkurang kualitasnya dan bisa mengakibatkan model YOLOv7 yang telah di train menggunakan data lobster di bawah laut tersebut memiliki mAP yang tidak maksimal. Berdasarkan permasalahan yang telah disebutkan tadi maka *underwater image enhancement* ini sangat diperlukan untuk meningkatkan mAP dari model YOLOv7 yang telah di-*train* dan juga untuk meningkatkan kualitas gambar secara visual.

II. KAJIAN TEORI

A. *Underwater Image Enhancement Deep Learning and Image Formation Model*

Image Enhancement Deep Learning and Image Formation Model merupakan *physical model-based methods*. Maka dari itu metode ini menggunakan formula dari *image formation model*. Formula *underwater image formation model* yang mendekati dapat diperoleh dari *image formation model* gambar berembun [5].

$$I(x) = D(x)t(x) + B\big(1 - t(x)\big) \quad (1)$$

Pada rumus persamaan (1), x menyatakan koordinat pixel, $I(x)$ menyatakan gambar yang diperoleh oleh kamera dalam lingkungan bawah air, $D(x)$ menyatakan radiasi dari objek itu sendiri, yang dapat dipahami sebagai gambar bawah air setelah mengeliminasi pengaruh faktor lingkungan bawah air, $t(x)$ adalah pemetaan transmisi langsung, B adalah cahaya lingkungan [1]. $D(x)t(x)$ berkorespondensi dengan

cahaya yang ditransmisikan secara langsung, dan $B(1 - t(x))$ berkorespondensi dengan *background* cahaya terdispersi.

Cahaya yang ditransmisikan secara langsung akan mengalami efek penurunan intensitas, besarnya ditentukan oleh *attenuation coefficient* β dan jarak transmisi d, seperti yang dijelaskan dalam persamaan (2).

$$t(x) = e^{-\beta d} \quad (2)$$

Metode-metode yang ada untuk *underwater image enhancement* biasanya memperlakukan *t(x)* dari tiga saluran gambar RGB sebagai satu pemetaan untuk menyederhanakan perhitungan. Namun, pendekatan ini mengabaikan perbedaan *attenuation coefficient* β dalam *t(x)* antara saluran yang berbeda dan gambar yang di-*enhance* masih memiliki bias warna [6], dilakukan sejumlah besar eksperimen bawah air dan diusulkan model yang direvisi [6], seperti yang ditunjukkan dalam persamaan (3). Model yang direvisi mengubah koefisien penurunan berdasarkan karakteristik optik lingkungan bawah air [1].

$$I_c(x) = D_c(x)e^{-\beta_c^D(v_D)^d} + B_c\left(1 - e^{-\beta_c^B(v_B)^d}\right) \quad (3)$$

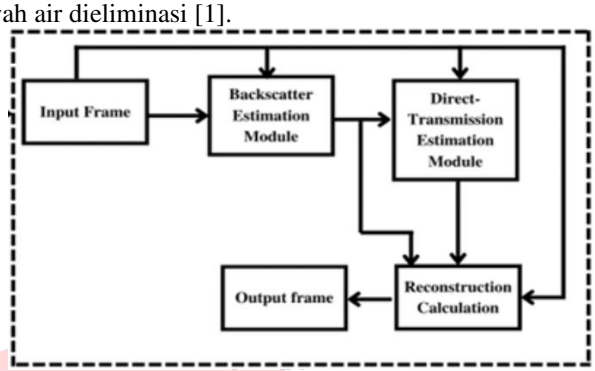
Pada persamaan (3) dijelaskan bahwa c ∈ {R, G, B} menyatakan *channel* warna, β_c^D dan β_c^B menyatakan *attenuation coefficients* untuk *direct transmission* dan *backscatter*, dan d menyatakan jarak transmisi [1]. *Attenuation coefficients* untuk *direct-transmission* dan *backscatter* tergantung pada V_c^d = [d, ρ , E, S_c, a, b] dan V_c^b = [E, S_c, a, b], di mana d adalah jarak transmisi, ρ adalah spektrum pantulan, E adalah *irradiance*, S_c adalah respon spektral kamera, a dan b adalah koefisien penyerapan dan hamburan [1]. Dengan menyusun ulang persamaan (3), dihasilkan persamaan (4) sebagai berikut:

$$D_c(x) = (I_c(x) - B_c)e^{\beta_c^D d} + B_c\, e^{(\beta_c^D - \beta_c^B)d} \quad (4)$$

Berdasarkan eksperimen di paper [7], nilai β_c^D-β_c^B sangat kecil dan dapat dianggap sebagai 0 pada jarak lebih dari 3 meter. Sehingga, persamaan (4) dapat diformulasikan kembali sebagai bentuk persamaan (5).

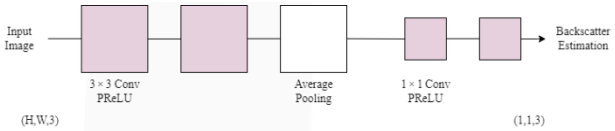
$$D_c(x) = (I_c(x) - B_c)e^{\beta_c^D d} + B_c \quad (5)$$

Inti dari metode *image enhancement Deep Learning and Image Formation Model* adalah menggunakan *convolutional neural networks* untuk menyesuaikan beberapa komponen dalam *image formation model* [1]. Seperti yang ditunjukkan pada Gambar 1, metode yang diusulkan terdiri dari tiga bagian. Bagian pertama mengestimasi *ambient light* dari proses *backscatter* menggunakan gambar masukan, yaitu B_c dalam persamaan (5). Bagian kedua menggunakan *ambient light* yang diestimasi dan gambar masukan untuk menghitung *direct transmission map*, yaitu $e^{\beta_c^D d}$ dalam persamaan (5). Di sini, berbeda dengan sebagian besar metode yang ada untuk memperkirakan $t(x)$, $e^{\beta_c^D d}$ sebenarnya merupakan invers dari $t(x)$. Keuntungan dari pengaturan ini adalah bahwa operasi rekonstruksi akhir bersifat multiplikatif, yang memudahkan *backpropagation of the loss function* dalam *training deep learning* [1]. Bagian ketiga menggunakan estimasi B_c dan $e^{\beta_c^D d}$ yang digabungkan dengan persamaan (5) untuk melakukan operasi rekonstruksi guna memperoleh gambar yang di-*enhance*, sehingga pengaruh lingkungan bawah air dieliminasi [1].



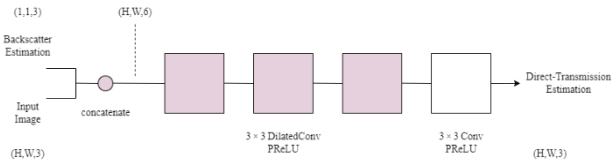
GAMBAR 1.
Arsitektur *Image Enhancement Deep Learning and Image Formation Model*

Backscatter Estimation Module mencakup dua kelompok kernel konvolusi 3x3, satu *global mean pooling layer*, dan dua kelompok kernel konvolusi 1x1. Jumlah setiap kelompok kernel konvolusi adalah 3 [1]. Detail struktur ditunjukkan pada Gambar 2 [1]. *Parametric Rectified Linear Unit* (PReLU) [18] dipilih sebagai fungsi aktivasi dari operasi konvolusi [1].



GAMBAR 2.
Backscatter Estimation Module [1]

Direct Transmission Estimation Module pertama-tama menggabungkan *Backscatter Estimation* dengan *input frame*, dan operasi selanjutnya meliputi tiga kelompok *dilated convolutional kernels* berukuran 3x3 dan satu kelompok kernel konvolusi normal berukuran 3x3 [1]. Jumlah setiap kelompok *dilated convolutional kernels* adalah 8, dan jumlah kelompok kernel konvolusi normal terakhir adalah 3 [1]. Detail dari strukturnya ditampilkan pada Gambar 3 [1]. *Dilated convolution* dapat memperluas *receptive field* dengan mengisi nol tambahan pada kernel konvolusi tanpa menambahkan lebih banyak parameter. Untuk menghindari *gridding effect* [8], dapat mengadopsi metode desain *hybrid dilated convolution*, yaitu memilih tingkat dilasi yang berbeda untuk operasi *dilated convolution* berurutan. Tingkat dilasi dari tiga kelompok *dilated convolution kernels* di modul ini adalah masing-masing (1, 2, 5) [1].



GAMBAR 3.
Direct-transmission Estimation Module [1]

B. *Underwater Image Enhancement Water-Net*

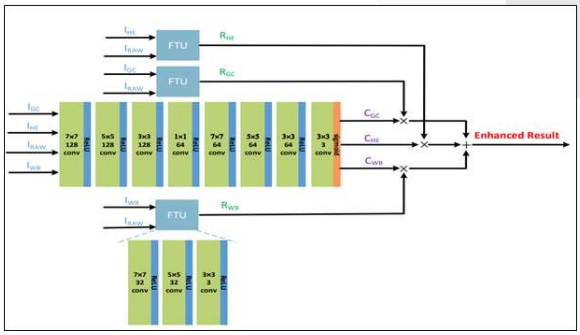
Water-Net merupakan *gated fusion network* yang menggabungkan *input* dengan *predicted confidence maps* agar mencapai hasil gambar yang ter-*enhanced* [2]. Berdasarkan karakteristik degradasi gambar bawah air, dihasilkan tiga input dengan menerapkan masing-masing algoritma *White Balance* (WB), *Histogram Equalization* (HE), dan *Gamma Correction* (GC) pada gambar bawah air [2]. Secara khusus, algoritma WB digunakan untuk mengoreksi pergeseran warna, sementara algoritma HE dan GC bertujuan untuk meningkatkan kontras dan mencerahkan daerah-daerah gelap, secara berturut-turut. Digunakan algoritma WB yang diusulkan pada [9], yang efektivitasnya telah terbukti. Untuk algoritma HE, diterapkan fungsi *adapthisteq* [10] yang disediakan oleh MATLAB pada komponen L dalam ruang warna Lab, dan kemudian mengubahnya kembali ke ruang warna RGB. Ditetapkan nilai Gamma dari algoritma GC menjadi 0,7 secara empiris [2].

Pada Gambar 4 [2] *Water-Net* menggunakan arsitektur *gated fusion network* untuk mempelajari tiga *confidence maps* yang akan digunakan untuk menggabungkan tiga gambar input menjadi hasil yang ter-*enhanced* [2]. *Learned confidence maps* menentukan fitur-fitur paling signifikan dari input yang tersisa dalam hasil akhir [2].

Sebagai model dasar, *Water-Net* adalah sebuah *Fully* CNN sederhana [2]. Dimasukkan tiga input yang dihasilkan dan *original* input ke *Water-Net* untuk memprediksi *confidence maps*. Sebelum melakukan fusi, dilakukan penjumlahan tiga *Feature Transformation Units* (FTU) untuk menyempurnakan tiga input tersebut [2]. Tujuan dari FTU adalah untuk mengurangi warna yang salah dan artefak yang diakibatkan oleh algoritma WB, HE, dan GC [2]. Akhirnya, ketiga input yang telah disempurnakan dikalikan dengan tiga *learned confidence maps* untuk mencapai hasil yang ter-*enhanced* [2].

$$I_{en} = R_{WB} \bullet C_{WB} + R_{HE} \bullet C_{HE} + R_{GC} \bullet C_{GC}, \quad (6)$$

Persamaan (6) merupakan formulasi untuk mendapatkan hasil gambar yang ter-*enhance,* dengan I_{en} merupakan hasil gambar yang telah di-*enhance,* simbol • menunjukkan perkalian elemen dari matriks. R_{WB}, R_{HE}, dan R_{GC} adalah *refined results* dari input setelah diproses oleh algoritma WB, HE, dan GC. C_{WB}, C_{HE}, dan C_{GC} adalah *learned confidence maps*.



GAMBAR 4.
Arsitektur *Water-Net* [2]

C. *Mean Average Precision* (mAP)

mAP atau *mean Average Precision* adalah metrik evaluasi yang banyak digunakan dalam tugas deteksi objek, termasuk algoritma YOLO (You Only Look Once) [4]. Metrik ini mengukur akurasi dari model deteksi objek dengan mempertimbangkan presisi dan recall untuk berbagai kategori objek [4]. Presisi adalah rasio prediksi positif benar terhadap total prediksi positif, sementara recall adalah rasio prediksi positif benar terhadap total jumlah contoh positif sebenarnya dalam dataset. AP (Average Precision) untuk setiap kategori objek menyimpulkan kurva presisi-recall [4]. mAP diperoleh dengan menghitung rata-rata skor AP dari semua kategori objek. Ini memberikan penilaian menyeluruh tentang performa model, mempertimbangkan kemampuannya untuk mendeteksi objek dari berbagai kelas secara akurat [4].

mAP50 adalah singkatan dari *mean Average Precision at IoU (Intersection over Union)* 50% [19]. mAP50 merupakan varian khusus dari metrik *Mean Average Precision* (mAP) yang digunakan untuk mengukur kinerja deteksi objek dalam tugas *image processing* atau *computer vision*. Metrik ini menggabungkan *precision* dan *recall* untuk mengevaluasi seberapa baik sistem *object detection* dapat mengidentifikasi dan menempatkan *bounding boxes* secara akurat pada objek-objek yang ada dalam gambar. Pada mAP50, IoU diatur pada 50%, yang berarti tumpang tindih antara *bounding boxes* yang dihasilkan oleh sistem deteksi dengan *bounding boxes* yang sebenarnya pada objek harus minimal 50% untuk dianggap sebagai deteksi yang benar [19]. Dengan kata lain, mAP50 mengukur sejauh mana sistem deteksi objek mampu menghasilkan *bounding boxes* yang memiliki tumpang tindih yang cukup besar dengan *bounding boxes* yang sebenarnya pada objek-objek dalam gambar. Semakin tinggi nilai mAP50, semakin baik kinerja sistem deteksi objek tersebut.

mAP50-95 adalah singkatan dari *mean Average Precision at IoU (Intersection over Union) range* 50-95% [19]. mAP50-95 merupakan metrik yang digunakan untuk mengukur kinerja deteksi objek dalam tugas *image processing* atau *computer vision*. Metrik ini menggabungkan *precision* dan *recall* untuk mengevaluasi seberapa baik sistem deteksi objek dapat mengidentifikasi dan menempatkan *bounding boxes* secara akurat pada objek-objek yang ada dalam gambar. Jangkauan IoU dari 50% hingga 95% mengacu pada rentang tumpang tindih atau *overlap* antara *bounding boxes* yang dihasilkan oleh sistem deteksi dengan *bounding boxes* yang benar-benar ada pada objek sebenarnya dalam gambar [19]. Ini mengukur sejauh mana *bounding boxes* yang dihasilkan oleh sistem memiliki tumpang tindih yang signifikan dengan *bounding boxes* yang sebenarnya. Secara keseluruhan, mAP50-95 memberikan gambaran tentang kinerja deteksi objek dengan mempertimbangkan variasi tumpang tindih yang berbeda. Semakin tinggi nilai mAP50-95, semakin baik sistem deteksi objek dalam memposisikan *bounding boxes* secara akurat pada berbagai objek yang ada dalam gambar.

III. METODE

Dalam proses pengujian ini terdapat dua metode *image enhancement* yang akan diuji, yaitu *Underwater Image*

Enhancement based on Deep Learning and Image Formation Model [1] dan *Water-Net* [2]. Penelitian ini menggunakan metode kuantitatif, hal ini dikarenakan hasil dari proses pengujian ini akan menampilkan hasil mAP YOLOv7 [3] dari masing-masing *image enhancement* yang digunakan dan membandingkannya dengan hasil tanpa *image enhancement* serta sesama metode *image enhancement* juga. *Pre-trained* model YOLOv7 yang digunakan dalam pengujian ini untuk proses *training* adalah yolov7.pt [3]. Berikut prosedur atau langkah untuk pengujian masing-masing *image enhancement*.

A. Langkah Pengujian *Underwater Image Enhancement Deep Learning and Image Formation Model*

Untuk pengujian *image enhancement* ini kita menggunakan data dari video Aruna yang diekstrak menjadi beberapa gambar. Kemudian gambar-gambar tersebut diproses menggunakan *image enhancement* yang telah dipilih, yaitu menggunakan *Deep learning and Image Formation Model*, setelah itu data untuk *train* dilakukan proses *training* menggunakan YOLOv7. Hasil dari *training* tersebut menghasilkan model yang akan dilakukan *testing* pada dataset *test* dari video yang pihak Aruna berikan, hasil mAP dari *testing* tersebut yang akan dijadikan objek untuk perbandingan. Untuk hasil pengujian bisa dilihat pada Gambar 5, Tabel 1, dan Tabel 2.

Terlihat cara kerja dari *Underwater Image Enhancement based on Deep Learning and Image Formation Model* pada Gambar 1 bagian *Image Enhancement*, pertama *input frame* dimasukan pada proses *Backscatter Estimation Module*, setelah itu hasilnya diteruskan ke *Direct-Transmission Estimation Module* bersama *input frame* sebelumnya. Hasil akhirnya dilakukan *Reconstruction Calculation* menggunakan masukan dari *input frame*, *Backscatter Estimation Module*, dan *Direct Transmission Module* [1].

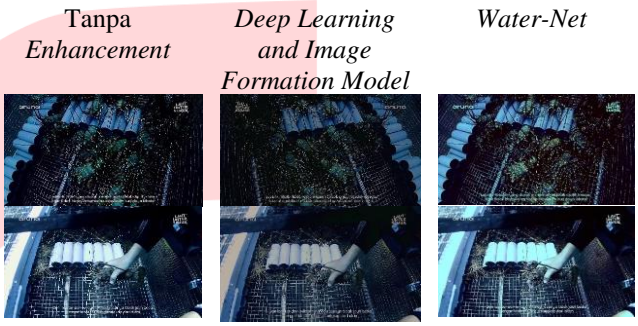
B. Langkah Pengujian *Underwater Image Enhancement Water-Net*

Untuk pengujian *Water-Net* memiliki langkah yang sama yaitu menggunakan data dari video Aruna lalu video tersebut diekstrak menjadi *frame-frame*. Kemudian *frame* tersebut diproses menggunakan *Water-Net*, setelah itu data untuk *train* dilakukan proses *training* menggunakan YOLOv7. Hasil dari *training* tersebut menghasilkan model yang akan dilakukan *testing* pada dataset *test* dari video yang pihak Aruna berikan, hasil mAP dari *testing* tersebut yang akan dijadikan objek untuk perbandingan. Untuk hasil dari pengujian *Water-Net* dapat dilihat pada Gambar 5, Tabel 1, dan Tabel 2.

Pada Gambar 4, pertama *input* gambar yang terdiri dari *raw image* (Iraw), *gamma correction image* (Igc), *white balance image* (Iwb), dan *histogram equalization image* (Ihe) di masukan ke *layer* konvolusi lalu hasilnya menghasilkan *learned confidence map* Cgc, Che, dan Cwb, setelah itu masing-masing dari *confidence map* tersebut dikalikan dengan hasil *refined result*-nya dari keluaran *Feature Transformation Unit* (FTU) yaitu Rhe, Rgc, dan Rwb, masing-masing dari hasilnya digabungkan sehingga menghasilkan *enhanced result* [2].

IV. HASIL DAN PEMBAHASAN

Pada Gambar 5 terlihat perbedaan visualisasi gambar yang sama namun dengan kualitas pencahayaan yang berbeda. Dengan diterapkannya sistem *image enhancement* berupa *Deep Learning and Image Formation Model*, kualitas pencahayaan pada visualisasi data meningkat. Hal ini akan mendukung untuk proses selanjutnya yaitu pada pendeteksian objek karena lobster akan semakin jelas terlihat.



GAMBAR 5.
Perbandingan Hasil Pengujian Gambar Orisinil dengan *Image Enhancement*; (Atas) detik ke 1; (Bawah) detik ke 3

TABEL 1.
Perbandingan mAP

Skenario	**mAP50**	**mAP50-95**
Tanpa Enhancement	0.982	0.388
Water-Net	0.926	0.437
Deep Learning and Image Formation Model	0.942	0.461

Berdasarkan Gambar 5 hasil pengujian secara visual menunjukkan bahwa gambar yang menggunakan *image enhancement Deep Learning and Image Formation Model* terlihat lebih jelas dan terang, sehingga lobster lebih ter-highlight. Pada data Tabel 1 nilai mAP 50-95 yang terbaik pun dimiliki oleh *image enhancement Deep Learning and Image Formation Model*.

Untuk pemrosesan per-*frame*-nya dapat dilihat pada Tabel 2, rata-rata *Water-Net* memproses 1 *frame* adalah sekitar 1.9 detik sedangkan untuk *Deep Learning and Image Formation Model* hanya membutuhkan rata-rata waktu 0.2 detik per *frame*. Dari hasil tersebut diputuskan untuk menggunakan *image enhancement Deep Learning and Image Formation Model* pada proyek utama.

TABEL 2.
Sampel Perbandingan Proses per *Frame* dalam detik

***Deep Learning and Image Formation Model* (detik)**	***Water-Net* (detik)**
0.11	1.60
0.13	1.77
0.37	0.74
0.12	0.92
0.48	4.61

V. KESIMPULAN

Penelitian ini membahas pengujian dua metode *underwater image enhancement*, yaitu *Underwater Image Enhancement based on Deep Learning and Image Formation Model* [1] dan *Water-Net* [2]. Metode ini dievaluasi secara kuantitatif menggunakan metrik mAP YOLOv7 [3] untuk masing-masing metode *image enhancement*, dan hasilnya dibandingkan dengan hasil tanpa *image enhancement* serta dengan metode *image enhancement* lainnya. Pengujian *Underwater Image Enhancement based on Deep Learning and Image Formation Model* melibatkan pemrosesan data video Aruna menjadi gambar-gambar individu. Gambar-gambar tersebut ditingkatkan menggunakan metode Deep Learning and Image Formation Model, dan hasilnya diuji dengan menggunakan model YOLOv7 yang telah di-*training* pada data tersebut. Hasil mAP dari pengujian ini digunakan untuk perbandingan. Selain itu, skema kerja metode ini juga diilustrasikan dalam gambar-gambar yang relevan. Pengujian *Water-Net* ini melibatkan langkah serupa dengan metode *Underwater Image Enhancement Deep Learning and Image Formation Model*, yaitu mengolah data video Aruna menjadi serangkaian *frame* yang kemudian ditingkatkan menggunakan metode *Water-Net*. Seperti pada pengujian sebelumnya, hasil mAP dari pengujian ini digunakan untuk perbandingan. Hasil dari pengujian *Water-Net* juga disajikan dalam gambar dan tabel yang relevan. Hasil akhirnya mengarah pada gambar atau *frame* yang telah di-*enhance*. Secara keseluruhan, penelitian ini mengidentifikasi dan menguji dua metode *image enhancement* yang berbeda, memberikan langkah-langkah pengujian yang terperinci, dan membandingkan hasil kinerja keduanya. Hasil yang didapatkan menyatakan bahwa dengan menerapkan sistem *image enhancement* menggunakan *Deep Learning and Image Formation Model*, kualitas pencahayaan pada visualisasi data meningkat. Hal ini mendukung proses selanjutnya, yaitu pendeteksian objek, karena objek lobster menjadi lebih jelas terlihat. Hasil pengujian secara visual menunjukkan bahwa gambar yang telah di-*enhance* menggunakan *image enhancement Deep Learning and Image Formation Model* tampak lebih jelas dan terang, dengan lobster menjadi lebih ter-*highlight*. Selain itu, nilai mAP 50-95 *image enhancement Deep Learning and Image Formation Model* memiliki kinerja terbaik dengan nilai 0.461 dibandingkan dengan *Water-Net* dan tanpa *enhancement*. Dalam pemrosesan per-*frame*-nya, model *Deep Learning and Image Formation Model* lebih cepat dibandingkan dengan *Water-Net*, dengan rata-rata waktu sekitar 0.2 detik per *frame* dibandingkan dengan *Water-Net* yang memperoleh rata-rata waktu 1.9 detik per *frame*. Berdasarkan hasil tersebut, diputuskan untuk *image enhancement* pada sistem monitoring lobster ini menggunakan *image enhancement Deep Learning and Image Formation Model*.

REFERENSI

[1] X. Chen, P. Zhang, L. Quan, C. Yi, and C. Lu, "Underwater Image Enhancement based on Deep Learning and Image Formation Model," *IEEE Access*, Jan. 2021, Accessed: Jun. 22, 2023. [Online]. Available: https://arxiv.org/abs/2101.00991

[2] C. Li et al., "An Underwater Image Enhancement Benchmark Dataset and beyond," IEEE Transactions on Image Processing, vol. 29, 2020, doi: 10.1109/TIP.2019.2955241.

[3] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," Jul. 2022, [Online]. Available: http://arxiv.org/abs/2207.02696.

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.

[5] S. G. Narasimhan and S. K. Nayar, "Vision and the Atmosphere," *Int J Comput Vis*, vol. 48, no. 3, pp. 233–254, 2002, doi: 10.1023/A:1016328200723.

[6] D. Akkaynak and T. Treibitz, "A Revised Underwater Image Formation Model," pp. 6723–6732, 2018, Accessed: Jul. 20, 2023. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/papers/Akkaynak_A_Revised_Underwater_CVPR_2018_paper.pdf

[7] C. Li, J. Guo, and C. Guo, "Emerging From Water: Underwater Image Color Correction Based on Weakly Supervised Color Transfer," *IEEE Signal Process Lett*, vol. 25, no. 3, pp. 323–327, Mar. 2018, doi: 10.1109/LSP.2018.2792050.

[8] Li C, Guo C, Ren W, et al. An Underwater Image Enhancement Benchmark Dataset and Beyond[J]. IEEE Transactions on Image Processing, 2019, 29:4376-4389.

[9] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2012, pp. 81–88. doi: 10.1109/CVPR.2012.6247661.

[10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788, 2016, [Online]. Available: http://pjreddie.com/yolo/

[11] J. Jaffe, "Underwater optical imaging: The past, the present, and the prospects," IEEE J. Oceanic Eng., vol. 40, no. 3, pp. 683-700, 2015. 1

[12] M. Sheinin and Y. Schechner, "The next best underwater view," in Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR), 2016, pp. 3764- 3773. 1

[13] M. Ludvigsen, B. Sortland, G. Johnsen, et al., "Applications of georeferenced underwater photo mosaics in marine biology and archaeology," J. Oceanography, vol. 20, no. 4, pp. 140-149, 2007. 1

[14] N. Strachan, "Recognition of fish species by colour and shape," J. Image Vis. Comput., vol. 11, no. 1, pp. 2-10, 1993. 1

[15] M. Yang, J. Hu, C. Li, et al., "An in-depth survey of underwater image enhancement and restoration," IEEE Access, vol. 7, pp. 123638-123657, 2019. 1

[16] F. Akhyar, L. Novamizanti & T. Riantiarni, "Sistem Inspeksi Cacat pada Permukaan Kayu menggunakan Model Deteksi Obyek YOLOv5", ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika, vol. 10, no. 4, pp. 990, 2022.

[17] H. M. Lathifah, L. Novamizanti, and S. Rizal, "Fast and Accurate Fish Classification from Underwater Video using You only Look Once," in IOP Conference Series:

Materials Science and Engineering, IOP Publishing Ltd, Dec. 2020. doi: 10.1088/1757- 899X/982/1/012003.

[18] F. Akhyar, L. Novamizanti, T. Putra, E. N. Furqon, M. C. Chang & C. Y. Lin, Lightning YOLOv4 for a surface defect detection system for sawn lumber, In 2022 IEEE 5th International Conference on Multimedia Information Processing and Retrieval (MIPR), pp. 184-189, 2022.

[19] X. Liu, J. Hu, H. Wang, Z. Zhang, X. Lu, C. Sheng, Shibin Song, Jun Nie, "Gaussian-IoU loss: Better learning for bounding box regression on PCB component detection, Expert Systems with Applications", Volume 190, 2022, 116178, ISSN 0957-4174, Available: https://doi.org/10.1016/j.eswa.2021.116178.