

## PENERAPAN RANDOM FOREST BERDASARKAN PERBAIKAN CITRA *CONTRACT LIMITED ADAPTIVE HISTOGRAM EQUALIZATION* (CLAHE) UNTUK PENGENALAN JENIS IKAN

R.A. Premunendar<sup>1</sup>, D.P.Prabowo<sup>2</sup>, F. Alzami<sup>3</sup>,R.A.Megantara<sup>4</sup>

<sup>1,2,3</sup>Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro  
Jl.Imam Bonjol No.207, Semarang

E-mail : ricardus.anggi@dsn.dinus.ac.id<sup>1</sup>, dwi.puji.prabowo@dsn.dinus.ac.id<sup>2</sup>,  
alzami@dsn.dinus.ac.id<sup>3</sup>, aria@dsn.dinus.ac.id<sup>4</sup>

### Abstrak

Salah satu kekayaan Indonesia yang sangat berlimpah pada sector perairan adalah biota ikan. Jenis ikan yang ada di perairan Indonesia sangat beragam, sehingga menyulitkan manusia untuk membedakan jenis-jenis ikan dan melindungi ikan yang terancam punah karena factor alam dan ulah manusia itu sendiri. Maka, perlu ditingkatkan kemampuan pengenalan ikan secara otomatis dengan bantuan computer. Telah ada penelitian sebelumnya untuk mengenali jenis-jenis ikan, namun tidak banyak yang mempertimbangkan adanya noise atau artefak-artefak yang timbul karena kondisi bawah air serta efek fitur-fitur ikan yang saling berkaitan. Oleh karena itu dalam penelitian ini, peneliti mengusulkan untuk melakukan analisis dampak pre-processing dari algoritma Contract Limited Adaptive Histogram Equalization (CLAHE). Pre-processing yang diberikan bertujuan untuk mengatasi artefak atau noise yang timbul pada citra bawah air dan mengatasi efek dari fitur-fitur keragaman jenis ikan. Klasifikasi dengan menggunakan Random Forest (RF) berbasis Contract Limited Adaptive Histogram Equalization (CLAHE), terbukti memberikan nilai akurasi rata-rata yang cukup tinggi yakni sebesar 98.54%, presisi 82.89%, dan recall 43.25%. Dilihat dari hasil penelitian ini, performa yang ditunjukkan oleh algoritma CLAHE sangat baik sehingga dianggap berhasil dan mampu mengatasi permasalahan adanya artefak atau noise yang timbul pada citra bawah air.

**Kata Kunci:** Clahe ,Jenis Ikan , Random florest

### I. PENDAHULUAN

Kekayaan laut yang dimiliki seperti ikan, udang, dan berbagai jenis hewan laut lainnya membuat perhatian masyarakat luar negeri menjadi meyakini hasil laut Indonesia. Hal ini tentu menjadi daya tarik tersendiri bagi Negara asing. Berbagai cara yang telah dilakukan masih memberikan hasil kinerja yang kurang tinggi jika diterapkan untuk mengelompokkan berbagai jenis biota ikan. Kemampuan mengenali biota perairan secara otomatis dengan bantuan komputer tidak mengandalkan keahlian atau pengalaman dari seorang ahli manusia adalah penyelesaian untuk permasalahan yang terjadi saat melakukan kegiatan yang banyak atau dilakukan secara berulang [1]. Selain itu, otomatisasi dalam memahami jenis biota ikan bermanfaat jika diterapkan pada data citra yang tidak terbatas jumlahnya dan beraneka ragam jenisnya.

Dengan kondisi seperti itu, peneliti ingin melakukan penelitian dalam hal pengenalan biota bawah laut khususnya ikan dengan cara mengekstraksi fitur citra dengan memperhatikan efek-efek yang timbul dari pengambilan citra bawah laut sehingga model baru tersebut diharapkan mampu untuk memberikan hasil yang lebih baik dari kinerja metode klasifikasi yang digunakan. Beberapa penelitian sebelumnya mengusulkan cara-cara otomatis untuk melakukan pengenalan jenis biota ikan, namun sangat jarang dalam penelitian yang mereka kerjakan memperhatikan dan mempertimbangkan efek yang ditimbulkan dari lingkungan bawah air dan efek lain dari fitur-fitur ikan yang dapat mempengaruhi gambar ikan yang diteliti. Dengan usulan yang seperti itu, peneliti akan memaparkan penelitian-penelitian terdahulu yang telah dilakukan oleh peneliti lain dengan tema yang serupa, metode, dan algoritma yang berbeda-beda pula. Banyak peneliti telah berupaya untuk menyelamatkan habitat air dengan cara mengurangi efek kepunahan jenis biota air, termasuk menerapkan teknologi *computer vision* dalam berbagai kegiatan pemantauan ekosistem air serta pengelompokan objek air[2]. Seperti yang dilakukan oleh Pujiono dkk [3] menggunakan *Contrast Limited Adaptive Histogram Equalization* (CLAHE) untuk memperbaiki kualitas citra dibawah air dan membandingkan 3 fungsi distribusi seperti distribusi *uniform*, *rayleigh*, dan *exponential*.

Menurut Andono, P. N., Purnama, I.K.E dan Hariadi, M. [4], register citra bawah air dengan menggunakan transformasi fitur skala-invariant (SIFT) sangat tergantung pada kualitas gambar. Registrasi dari SIFT *Contract Limited Adaptive Histogram Equalization* (CLAHE) dengan Rayleigh distribusi mampu meningkatkan citra bawah air kualitas menghasilkan 41% hasil yang lebih baik. Yusof, W.J.H.W., Hitam, M.S., Awaludin, E.A. dan Bachok, Z [5], menyatakan bahwa itu lebih efektif menggunakan kombinasi

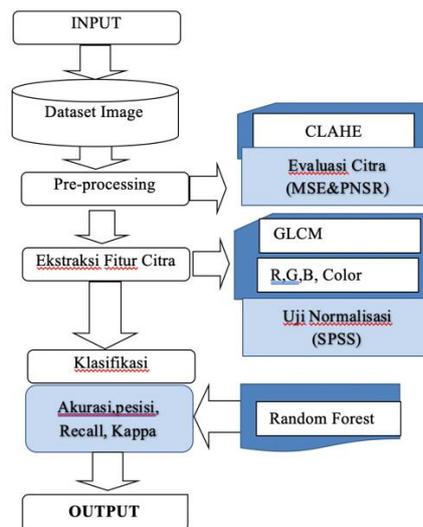
CLAHE dengan RGB dan HSV berbasis warna untuk peningkatan citra bawah air. Pengukuran dilakukan dengan menggunakan Mean Square Error (MSE) dan Rasio Sinyal ke Noise Puncak (PSNR), dimana nilai MSE lebih rendah dan PSNR lebih tinggi. Pada dasarnya kontras rendah citra bawah air dan perambatan cahaya adalah kendala utama dalam melakukan segmentasi citra bawah air.

Klasifikasi adalah proses akhir untuk mengetahui hasil kinerja dari implementasi algoritma-algoritma yang diterapkan pada pre-processing dan ekstraksi fitur pada image. Dalam penelitiannya, Breiman menunjukkan kelebihan Radom Forest yang dapat memberikan nilai error yang lebih rendah, memberikan hasil klasifikasi yang bagus dan dapat mengatasi data training dalam jumlah yang sangat besar secara efisien serta merupakan metode yang efektif untuk mengestimasi missing data. Fransiska,dkk mengimplementasikan gabungan metode RF dan CART mampu mengklasifikasikan kelas minor pada dataset KDD Cup 1999 dengan sedikit modifikasi teknik Balanced Random Forest (BRF) dengan cara menyeimbangkan jumlah kelas mayor dan kelas minor. Akurasi yang dicapai oleh model belum mencapai nilai maksimal, karena keterbatasan jumlah record pada data training setelah diimplementasikan BRF yang tidak dapat menggambarkan karakteristik kelas yang ada.

Dalam penelitian ini akan melakukan klasifikasi jenis ikan menggunakan metode random forest dengan melakukan proses preprocessing menggunakan algoritma Clahe untuk mendapatkan perbaikan citra dalam pengenalan jenis ikan.

## II. METODOLOGI PENELITIAN

### 2.1 Metode yang diusulkan



Gambar 2.1 metode yang diusulkan

Tahapan penelitian dimulai dari input, dataset image, preprocessing, ekstraksi citra, klasifikasi dan output. Proses utama pada penelitian ini ada di tahap pre-processing, ekstraksi citra dan klasifikasi.

### 2.2 Computer Vision (CV)

Computer vision merupakan sistem yang mempunyai kemampuan dalam menganalisa objek secara visual, seperti akuisisi citra, pengolahan citra, pengenalan dan membuat keputusan dengan data objek yang ada dimasukkan dalam bentuk citra (image). Computer vision mencoba meniru cara kerja sistem visual manusia (human vision) yang sesungguhnya sangat kompleks, manusia melihat objek dengan indera penglihatan (mata), lalu citra objek diteruskan ke otak untuk diinterpretasi sehingga manusia mengerti objek apa yang tampak dalam pandangan mata. Hasil dari interpretasi tersebut digunakan untuk pengambilan keputusan.

### 2.3 Pengenalan berdasarkan Citra Bawah Air

Pengenalan citra bawah air merupakan analisis konten dari visual bawah laut yang merupakan tugas menantang dalam topik CV. Hal ini dikarenakan hamburan dan penyerapan, citra bawah air mengalami masalah kontras rendah, kabur, dan distorsi warna [5]. Permasalahan tersebut dapat mempengaruhi konten visual di bawah laut. Kualitas dari konten tersebut mempengaruhi fitur-fitur yang digunakan dalam proses pengenalan.

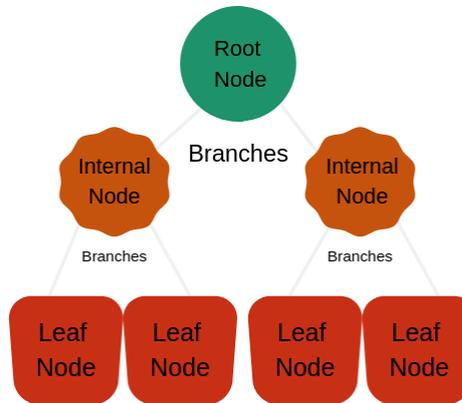
### 2.4 Clahe

Contrast limited adaptive histogram equalization (CLAHE) terbukti berhasil untuk mengintensifkan citra dengan kontras rendah. CLAHE adalah versi perbaikan adaptive histogram equalization (AHE). Kedua metode digunakan untuk mengatasi keterbatasan pemerataan histogram standar. Tepian bidang yang tajam dapat dipertahankan dengan intensifikasi selektif di dalam batas-batas bidang. Intensifikasi selektif dilakukan dengan mendeteksi terlebih dahulu tepi bidang pada

citra kemudian memproses daerah citra yang berada di dalam bidangnya. CLAHE membagi citra ke dalam bidang kontekstual dan menerapkan pemerataan histogram untuk masing-masing. Masing-masing bidang kontekstual dihitung dengan menggunakan histogram yang kemudian dipangkas dengan batas yang diinginkan untuk ekspansi kontras. Hal tersebut menunjukkan distribusi nilai abu-abu yang digunakan mampu membuat fitur tersembunyi dari citra dapat terlihat.

### 2.5 Random Forest

Algoritma klasifikasi Random Forest yakni sekumpulan klasifikasi yang terdiri dari banyak pohon keputusan dan melakukan klasifikasi berdasarkan output yang merupakan hasil klasifikasi dari setiap pohon keputusan anggota.



Gambar 2.2 Ilustrasi Metode Decision Tree

Metode Random Forest ini sangat tidak asing dalam dunia Machine Learning yang mana metode ini di implementasikan dalam penyelesaian masalah. Decision tree mengklasifikasikan suatu sampel data yang belum diketahui kelasnya kedalam kelas – kelas yang ada. Penggunaan decision tree agar dapat menghindari overfitting pada sebuah set data saat mencapai akurasi yang maksimum. Random Forest merupakan bagian dari metode Decision Tree. Decision Tree atau bias disebut pohon pengambil keputusan merupakan sebuah diagram alir yang berbentuk seperti pohon yang memiliki akar yang disebut dengan root node yang berfungsi untuk mengumpulkan data. Root node memiliki cabang akar yang disebut inner root yang berisi tentang data dan inner root ini memiliki cabang lagi disebut leaf node yang berfungsi sebagai pemecah masalah dan memberikan keputusan.

### 2.6 Validasi

Validasi hasil pengujian dilakukan berdasarkan tahapan proses pengenalan jenis ikan. Pada tahapan pengolahan data citra validasi dilakukan melalui hasil yang diperoleh dari evaluasi. Kesesuaian hasil pengujian yang diberikan metode *mean square error* (MSE), *peak signal to noise ratio* (PSNR), dari data yang telah mengalami proses perbaikan secara konsisten diartikan bahwa metode perbaikan citra tersebut sesuai dan layak. Pada tahapan pengolahan fitur, pengujian normalitas dilakukan untuk melihat distribusi data sebelum dilakukan proses klasifikasi, dan hasil dari proses klasifikasidiperlihatkan dalam hasil akurasi, presisi dan *recall*.

## III. HASIL DAN PEMBAHASAN

### 3.1 Pengumpulan data

Dalam penelitian ini, terdapat dataset yang tersedia secara publik dan ada beberapa sesuai dengan kriteria tersebut. Dataset tersebut adalah dataset *fish4knowledge*. Pada dataset *fish4knowledge* terdiri dari 23 jenis ikan dengan total citra sebanyak 27.370 citra ikan kemudian data citra ikan tersebut di gunakan sebagai data training dan testing.

### 3.2 Preprocessing

Pada tahap peningkatan citra dengan kontras limited adaptive histogram equalization, dilakukan analisis pemerataan histogram sehingga berpengaruh terhadap hasil akurasi CLAHE mampu mengurangi masalah noise dengan membatasi peningkatan kontras, terutama pada daerah homogen. CLAHE meningkatkan kontras citra dengan cara mengubah nilai intensitas pada citra.

Pada tahap pertama peningkatan citra dengan kontras limited adaptive histogram equalization (CLAHE), gambar dibagi menjadi beberapa daerah bagian dengan ukuran yang sama. Selanjutnya, CLAHE bertugas menghitung histogram di setiap wilayah bagian. Dalam proses perhitungan histogram, harus menentukan clip limit factor terlebih dulu untuk melakukan clipping histograms dengan menggunakan persamaan berikut:

$$\beta = \frac{M}{N} \left( 1 + \frac{\alpha}{100} ((S_{\max} - 1)) \right) \tag{1}$$

Dimana  $\alpha$  merupakan clip limit factor, M dan N adalah perbandingan luas wilayah bagian dan nilai grayscale (256). Selain itu,  $S_{\max}$  merupakan nilai maksimum dari setiap wilayah.

Setelah itu, histogram akan distribusikan kembali dengan menggunakan system distribusi yang tidak melewati batas clip limit factor. Pada tahap terakhir, hasil perhitungan kumulatif pada CLAHE digunakan untuk pemetaan citra grayscale.

### 3.3 Ekstraksi fitur

Pada tahap ini dilakukan normalisasi untuk menggunakan metode Gray Level Co-occurences Matrix (GLCM) sebagai ekstraksi tekstur, langkah untuk menghitung fitur GLCM adalah membuat matriks co-occurrence seperti berikut. Langkah pertama, Misal I adalah grayscale citra training Image1.jpg sudut 0° jarak  $d = 1$  dengan nilai piksel seperti diatas akan dinormalisasikan. Setelah perhitungan normalisasi dan mendapatkn hasil dari pembagian citra maka dilanjutkan dengan perhitungan *Angular Second Moment* (ASM), *contras*, *Inverse Defferent Moment* (IDM), Entropi, dan Korelasi.

### 3.4 Random florest

Setealah dilakukan ekstraksi fitur menggunakan GLCM, maka metode Random Forest akan berkerja lebih efisien pada jumlah data set training dan testing yang besar. Proses RF ini akan menghasilkan kumpulan pohon tunggal dengan ukuran dan bentuk yang berbeda-beda. Hasilnya pohon-pohon tunggal tersebut memiliki korelasi yang kecil antar pohon. Sehingga hasil yang didapatkan akn lebih akurat dalam pengenalan jenis ikan.

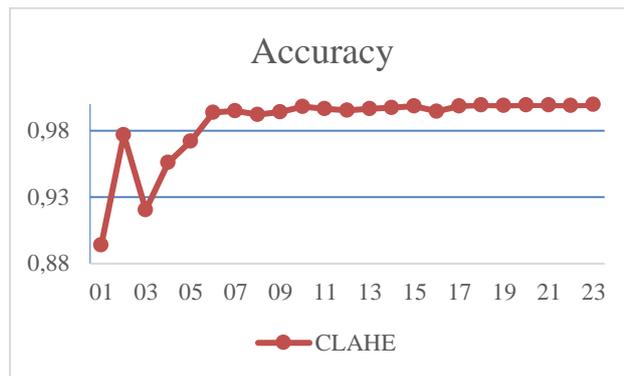
### 3.5 Evaluasi hasil

Sebelumnya, kami tampilkan tabel yang memuat nama kelompok ikan yang kami gunakan dalam eksperimen ini:

Tabel 3.1 nama kelompok ikan yang digunakan untuk eksperiment

01	02	03	04
Dascyllus reticulatus	Plectroglyphidodon dickii	Chromis chrysur	Amphiprion clarkia
05	06	07	08
Chaetodon lunulatus	Chaetodon trifascialis	Myripristis kuntee	Acanthurus nigrofuscus
09	10	11	12
Hemigymnus fasciatus	Neoniphon samara	Abudehduf vaigiensis	Canthigaster valentine
13	14	15	16
Pomacentrus moluccensis	Zebrasoma scopas	Hemigymnus melapterus	Lutjanus fulvus
17	18	19	20
Scolopsis bilineata	Scaridae	Pempheris vanicolensis	Zanclus cornutus
21	22	23	24
Neoglyphidodon nigroris	Balistapus undulates	Siganus fuscescens	-

Hasil akurasi dari Klasifikasi Random Forest dengan teknik peningkatan citra dengan clahe dapat ditunjukkan dengan Gambar 3.1 sebagai berikut:



**Gambar 3.1 Hasil Eksperimen Perhitungan Akurasi dari Clahe**

Dimana hasil klasifikasi dari Random Forest dengan teknik peningkatan citra dengan clahe memberikan nilai rata-rata 98.54%. Hal ini dapat dikatakan bahwa algoritma clahe mampu memberi akurasi yang baik.

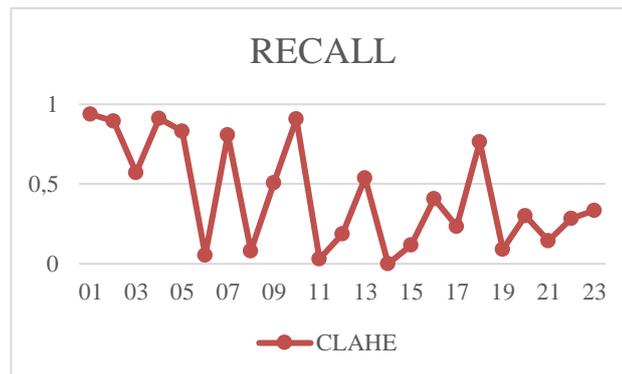
Pada hasil eksperimen, dilihat dari capaian angka presisi pada peningkatan citra clahe menunjukkan hasil perhitungan presisi seperti pada tabel 3.2.

**Tabel 3.2 Hasil Eksperimen Perhitungan Presisi dari clahe**

Kelp.Ikan	Clahe
01	0.84181
02	0.87378
03	0.7378
04	0.81408
05	0.86569
06	0.5
07	0.88112
08	0.75
09	0.62963
10	0.91753
11	1
12	0.69231
13	0.94444
14	None
15	1
16	0.75
17	1
18	0.86667
19	1
20	1
21	1
22	1
23	1

Dari data diatas memperlihatkan bahwa nilai presisi pada beberapa spesies tidak diketahui , hal ini disebabkan karena nilai presisi merupakan ukuran terdekat antar kumpulan data atau spesies ikan ikan yang sama. Hasil presisi dapat ditunjukkan dengan nilai 1 sebagai nilai presisi terbaik , hal ini menunjukkan bahwa karakter dan karakteristik data set sangatlah mirip. Nilai rata-rata presisi yang diperoleh dari pengujian yang dilakukan sebesar 82,89%

Selain dari nilai akurasi dan presisi, peningkatan citra dalam pengenalan spesies ikan dapat dilihat dari tingkat recall yang didapatkan pada saat pengujian. Nilai recall dalam pengujian ini ditunjukkan pada grafik 3.2 dibawah :



Gambar 3.2 Hasil Eksperimen Perhitungan Recall Dari Clahe

Data diatas menunjukkan nilai rata-rata recall yakni 43.25%, artinya nilai recall yang didapat masih cukup tinggi pada saat pengujian dengan data set yang telah digunakan

#### IV. KESIMPULAN

Berdasarkan hasil eksperimen klasifikasi citra spesies ikan menggunakan metode *Random Forest* berbasis peningkatan citra CLAHE dapat disimpulkan bahwa akurasi dalam proses klasifikasi adalah dipengaruhi oleh peningkatan gambar. Bahkan, disertai dengan penggunaan klasifikasi yang sesuai dengan parameter, kinerja klasifikasi dapat ditingkatkan. Oleh karena itu, masalah yang terjadi pada lingkungan bawah laut dan menciptakan hambatan dalam proses klasifikasi dapat diselesaikan dengan menerapkan peningkatan citra.

Klasifikasi dengan menggunakan *Random Forest (RF)* berbasis *Contract Limited Adaptive Histogram Equalization (CLAHE)*, terbukti memberikan nilai akurasi rata-rata yang cukup tinggi yakni sebesar 98.54%, presisi 82.89%, dan recall 43.25%. Dilihat dari hasil penelitian ini, performa yang ditunjukkan oleh algoritma CLAHE sangat baik sehingga dianggap berhasil dan mampu mengatasi permasalahan adanya artefak atau noise yang timbul pada citra bawah air

#### VI. REFERENSI

- [1] R. A. Pramunendar, C. Supriyanto, D. H. Novianto, I. N. Yuwono, G. F. Shidik, and P. N. Andono, "A classification method of coconut wood quality based on Gray Level Co-occurrence matrices," in 2013 International Conference on Robotics, Biomimetics, Intelligent Computational Systems, 2013, pp. 254–257.
- [2] Gu, Z. et al., 2016. Automatic searching of fish from underwater images via shape matching. In OCEAN 2016-Shanghai.IEEE,pp1-4. Available at: <http://ieeexplore.ieee.org/document/7485597/>.
- [3] Pujiono, E. M. Yuniarno, I. K. E. Purnama, M. H. Purnomo, and M. Hariadi, "Mathematical model for underwater color constancy based on polynomial equation," *Journal of Theoretical and Applied Information Technology*, vol. 87, no. 1, pp. 159–166, 2016.
- [4] Andono, P.N, Purnama, I.K.E. and Hariadi,M. "Underwater Image Enhancement Using Adaptive Filtering For Enhanced SIFT-Based Image Matching" , *Journal of Theoretical and Applied Information Technology*, vol 52, no. 3, 2013, pp. 273-280,
- [5] Yussof, W.N.J.H.W., Hitam, M.S., Awalludin, E.A. and Bachok, Z., "Performing Contrast Limited Adaptive Histogram Equalization Technique on Combine Color Models for Underwater Image Enhancement", *International Journal of Interactive Digital Media*, vol.1, April 2013