

TESIS

**Identifikasi Kepiting Molting menggunakan Klasifikasi
Machine Learning**

Crab Molting Identification using Machine Learning Classifiers

Disusun dan diajukan oleh

RUNAL REZKIAWAN BAHARUDDIN

D032172014



**DEPARTEMEN TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
MAKASSAR
2021**

**Identifikasi Kepiting Molting menggunakan Klasifikasi
Machine Learning**

Tesis Ini Diajukan Sebagai Salah Satu Syarat Untuk Mencapai Gelar

Magister Program Studi Teknik Informatika

Departemen Teknik Elektro Fakultas Teknik

Universitas Hasanuddin

RUNAL REZKIAWAN BAHARUDDIN

D032172014

**DEPARTEMEN TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
MAKASSAR
2021**

LEMBAR PENGESAHANAN

PERNYATAAN KEASLIAN TESIS

Yang bertanda tangan di bawah ini :

Nama : Runal Rezkiawan Burhanuddin

NIM : D032172014

Program Studi : Teknik Elektro

Menyatakan dengan sebenarnya bahwa tesis yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilan tulisan atau pemikiran orang lain. Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan tesis ini hasil karya orang lain, saya bersedia menerima sanksi atas perbuatan tersebut.

Makassar, 15 Desember 2021

Yang menyatakan,

Runal Rezkiawan Baharuddin

ABSTRAK

Produksi kepiting cangkang lunak dilakukan dengan cara memilih kepiting sesaat sebelum molting dan menempatkannya di dalam box sampai terjadi molting. Sesaat setelah molting, cangkang kepiting yang baru masih lunak dan akan mengeras dalam beberapa jam kemudian. Petani harus memanen kepiting saat cangkang kepiting masih lunak. Penelitian ini bertujuan untuk mengidentifikasi molting kepiting menggunakan pengklasifikasi machine learning. Kami menggunakan tiga pengklasifikasi machine learning, yaitu KNN, SVM, dan RFC, dan membandingkannya untuk menentukan algoritma klasifikasi terbaik. Hasil percobaan menunjukkan bahwa KNN merupakan algoritma klasifikasi terbaik untuk identifikasi molting kepiting.

Keywords: identifikasi kepiting molting, algoritma klasifikasi, Machine Learning.

ABSTRACT

The production of soft-shell crabs done by selecting the crabs just prior to molting and placing them in a box until the molting occurs. Shortly after molting, the new crab shells are still soft and will be hardened in a few hours later. Farmer must harvest the crab while the crabs' shell is soft. This study aims to identify crab molting using machine learning classifiers. We use three machine learning classifiers, namely KNN, SVM, and RFC, and compared them to determine the best classification algorithm. The experimental results show that KNN is the best classification algorithm for crab molting identification.

Keywords: Crab Molting Identification, Classification Algorithms, Machine Learning.

KATA PENGANTAR

Puji syukur kami panjatkan kehadirat Allah SWT. atas berkah dan rahmat-Nya sehingga pengajuan proposal penelitian dengan judul “**Identifikasi Kepiting Molting menggunakan Klasifikasi Machine Learning**” dapat terselesaikan.

Pada penyusunan kali ini kami menyajikan beberapa hal yang menyangkut judul yang telah kami angkat dan telah melalui proses pencarian dari berbagai sumber baik jurnal penelitian, buku maupun dari situs-situs di internet, selain dari penelitian sebelumnya. Untuk itu melalui kesempatan ini kami ingin mengucapkan terima kasih kepada seluruh pihak yang telah membantu dalam penyusunan ini. Walaupun dalam penyusunannya, kami telah berusaha secara maksimal, namun jika masih ada kekurangan baik dari segi pengetikan maupun dari segi isi mohon kritik dan saran, demi penyusunan selanjutnya agar lebih baik lagi. Akhir kata saya selaku penyusun mengucapkan terima kasih.

Makassar, Desember 2021

Penulis

DAFTAR ISI

LEMBAR PENGESAHANAN	iii
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
BAB I PENDAHULUAN	1
A. LATAR BELAKANG	1
B. RUMUSAN MASALAH.....	3
C. TUJUAN PENELITIAN.....	3
D. MANFAAT PENELITIAN.....	3
E. BATASAN MASALAH	4
F. SISTEMATIKA PENULISAN	5
BAB II TINJAUAN PUSTAKA	7
A. LANDASAN TEORI.....	7
1. Budidaya / Aquaculture	7
2. Kepiting.....	8
3. Kepiting Lunak (Soft Shell Crab)	11
4. Artificial Intelligence	12
5. Machine Learning	14
6. Klasifikasi.....	16
7. K-Nearest Neighbours Classifier	19
8. Support Vector Machine	22
9. Random Forest Classifier	23
10. Pembagian Data	25
11. Confusion Matrix.....	26
B. PENELITIAN TERKAIT.....	29
C. STATE OF THE ART	31
D. KERANGKA PIKIR	34

BAB III	METODOLOGI PENELITIAN	35
A.	TAHAPAN PENELITIAN	35
B.	WAKTU DAN LOKASI PENELITIAN	36
C.	JENIS PENELITIAN.....	36
D.	PERANCANGAN SISTEM.....	37
E.	SUMBER DATA.....	38
F.	INSTRUMENTASI PENELITIAN.....	39
BAB IV	HASIL DAN PEMBAHASAN.....	40
A.	Pengumpulan Dataset	40
1.	Data Training.....	40
2.	Data Testing	41
B.	Preprocessing.....	41
1.	<i>Resize</i> Citra.....	42
2.	Transformasi Citra	46
3.	Ekstraksi fitur HOG.....	50
C.	Hasil Klasifikasi.....	54
1.	Performa Klasifikasi Algoritma KNN	55
2.	Performa Klasifikasi Algoritma SVM.....	56
3.	Performa Klasifikasi Algoritma RFC	57
D.	Perbandingan Performa Metode Klasifikasi	58
BAB V	PENUTUP	60
A.	Kesimpulan.....	60
B.	Saran	60
DAFTAR PUSTAKA.....		61

DAFTAR GAMBAR

Gambar 2. 1 Konsep Kecerdasan Buatan	13
Gambar 2. 2 Klasifikasi sebagai pemetaan sebuah himpunan atribut input x ke dalam label kelasnya.....	18
Gambar 2. 3 Ilustrasi Random Forest.....	25
Gambar 2. 4 Metode Pembagian Data Menjadi Dua Kelompok Data	26
Gambar 3. 1 Tahapan penelitian.....	35
Gambar 3. 2 Pecancangan Sistem	37
Gambar 4. 1 Akuisisi Data	40
Gambar 4. 2 Citra Kepiting Molting	41
Gambar 4. 3 Citra Kepiting non-molting.....	41
<i>Gambar 4. 4 data citra molting yang sebelum resize dan setelah resize</i> .	46
Gambar 4. 5 data citra non-molting yang sebelum resize dan setelah resize	46
Gambar 4. 6 data citra molting dan non-molting yang belum ditransformasi dan yang telah ditransformasi ke greyscale.....	49
Gambar 4. 7 Algoritma Histogram of Gradient (HOG)	50
Gambar 4. 8 Cell yang menyusun sebuah blok.....	52
Gambar 4. 9 HOG pada Kepiting Molting	53
Gambar 4. 10 HOG pada Kepiting Non-Molting.....	53
Gambar 4. 11 Hasil confusion matrix k-NN.....	56
Gambar 4. 12 Hasil Confusion Matrix SVM	57
Gambar 4. 13 Hasil Confusion Matrix RFC.....	58
Gambar 4. 14 Hasil Perbandingan Metode Klasifikasi	59

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	27
Tabel 2. 2 state of the art	31
Tabel 4. 1 Performa klasifikasi Algoritma KNN	55
Tabel 4. 2 Performa klasifikasi Algoritma SVM	57
Tabel 4. 3 Performa klasifikasi Algoritma RFC.....	58
Tabel 4. 4 Perbandingan Performa Metode Klasifikasi	59

BAB I
PENDAHULUAN
A. LATAR BELAKANG

Kepiting merupakan salah satu makanan laut (seafood) yang paling banyak diminati oleh masyarakat baik dari lingkungan menengah maupun kalangan keatas, kepiting adalah makanan laut yang lezat setelah melalui proses pengolahan yang baik maka akan mengandung protein dan vitamin yang dibutuhkan oleh tubuh manusia, salah satu kepiting yang menjadi incaran para penikmat kuliner *seafood* belakangan ini adalah kepiting soka/lunak (*Soft Shell Crab*).

Kepiting lunak (soft shell crab) adalah salah satu makanan laut (seafood) di dunia yang terkenal karena kelezatannya. Produk ini belum dikenal luas oleh masyarakat Indonesia meskipun banyak diproduksi di Indonesia. Hal ini terjadi karena kepiting lunak adalah produk ekspor yang mana permintaan luar negeri jauh lebih tinggi dibanding produksi. Komoditas ini diekspor ke Amerika, Cina, Jepang, Hongkong, Korea Selatan, Taiwan, Malaysia, dan sejumlah negara di kawasan Eropa. Produksi kepiting lunak dilakukan dengan memelihara kepiting secara individu dalam kotak (crabs box) yang ditempatkan di dalam tambak hingga molting. Molting adalah proses pergantian kulit secara alami, yakni melepaskan kulit lama yang keras untuk tujuan pertumbuhan. Sesaat setelah molting, kulit kepiting yang baru masih dalam kondisi sangat lunak dan akan mengeras kembali beberapa jam kemudian setelah terjadi

penyerapan air. Kepiting dengan kondisi lunak inilah yang dipanen sebagai kepiting lunak. (Fujaya, 2011)

Agar panen kepiting lunak dapat dilakukan lebih cepat maka ada beberapa perlakuan yang dapat diberikan yaitu salah satunya melalui teknik mutilasi, yakni dengan sengaja menanggalkan kaki-kaki kepiting untuk merangsang molting. Menurut Raden Ario dkk, dalam penelitiannya tentang Perbedaan Metode Mutilasi Terhadap Lama Waktu Molting *Scylla serrata* menunjukkan berpengaruh terhadap lama waktu molting dan menggunakan teknik tersebut, tidak berpengaruh terhadap pertumbuhan berat mutlak (Ario et al., 2019)

Dalam budidaya kepiting lunak, pengawasan sangat menentukan kualitas kepiting lunak yang berhasil. Hal tersebut dikarenakan kepiting yang dipanen sesaat setelah molting. Bila kepiting terlambat dipanen, maka kulit mereka segera mengeras kembali dan hal ini menyebabkan kualitas menurun. (Yushinta Fujaya, Siti Aslamyah, Letty Fudjaja, 2019)

Penelitian di atas membuktikan bahwa pentingnya pemantauan terhadap kepiting molting dalam budidaya kepiting cangkang lunak dengan menggunakan teknologi klasifikasi Machine Learning untuk deteksi lebih dini agar lebih cepat di tangani ketika kepiting lunak. Oleh karena itu, penelitian ini mengusulkan sistem identifikasi kepiting molting untuk memberikan kemudahan kepada pembudidaya dalam pemantauannya.

B. RUMUSAN MASALAH

Adapun rumusan masalah pada penelitian Identifikasi Molting Kepiting menggunakan Teknik klasifikasi *Machine Learning* ini adalah :

Bagaimana perbandingan akurasi algoritma klasifikasi dalam memonitoring molting pada kepiting menggunakan identifikasi citra.

C. TUJUAN PENELITIAN

Adapun tujuan penelitian yang ingin dicapai dari penelitian ini adalah:

Mengidentifikasi perbandingan akurasi algoritma klasifikasi dalam memonitoring molting pada kepiting menggunakan identifikasi citra

D. MANFAAT PENELITIAN

Adapun manfaat yang dapat diperoleh dari penelitian ini adalah :

1. Bagi industri, penelitian yang diusulkan memberikan kontribusi teknologi menuju revolusi Industri 4.0 guna meningkatkan budidaya kepiting cangkang lunak.
2. Bagi peneliti, penelitian yang diusulkan dapat menambah pengetahuan dan dapat memberikan kontribusi identifikasi

kepiting molting sehingga pembudidaya tidak perlu lagi mendatangi tambak untuk melakukan pengecekan kepiting molting, cukup memantau dari jarak jauh yang diberikan sistem yang sebelumnya dilakukan manual oleh petani dengan mengecek satu persatu kepiting yang molting, sehingga dengan system identifikasi molting ini dapat membantu dalam mamantau kondisi kepiting.

3. Bagi pelaku usaha budidaya kepiting cangkang lunak, dapat menggantikan sumber daya manusia untuk mengontrol tiap kepiting hendak molting sehingga memiliki kontribusi terhadap pengurangan penggunaan tenaga manusia.
4. Bagi institusi pendidikan magister Departemen Teknik Elektro konsentrasi Teknik Informatika, dapat digunakan sebagai referensi ilmiah dalam mengembangkan penelitian yang berhubungan dengan identifikasi kepiting molting menggunakan klasifikasi machine learning.

E. BATASAN MASALAH

Adapun batasan masalah yang dapat diperoleh dari penelitian ini adalah:

1. Sistem deteksi kepiting molting hanya bisa digunakan dalam ruangan (*Indoor*).

2. Menggunakan satu wadah untuk mendapatkan dataset.
3. Pengambilan citra diambil pada pagi-sore hari.
4. Kepiting yang dijadikan objek yaitu jenis bakau (*Scylla Serrata*)
5. Algoritma klasifikasi yang digunakan KNN, SVM dan RFC
6. Kepiting bakau yang menjadi objek merupakan kepiting yang digunakan petambak dalam budidayanya yaitu kepiting bakau yang telah di mutilasi kakinya.

F. SISTEMATIKA PENULISAN

Adapun sistematika penulisan pada penelitian pembuatan Deteksi Molting Kepiting menggunakan Teknik klasifikasi *Machine Learning* adalah:

Bab I Pendahuluan

Bab I berisi penjelasan tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah serta sistematika penulisan.

Bab II Landasan Teori dan Kerangka Pemikiran

Bab II berisi penjelasan tentang landasan teori, *state of the art*, dan kerangka pemikiran. Landasan teori dijelaskan tentang kepiting soka, pakan kepiting, dan sistem Otomasi yang akan digunakan. Dalam bab ini juga diuraikan *state of the art* yang berisi tentang penelitian terkait dan kerangka berfikir untuk memecahkan masalah yang sedang diteliti.

Bab III Metodologi Penelitian

Bab III berisi penjelasan tentang tahapan penelitian yang akan dilakukan dimulai dari proses awal hingga akhir penelitian. Diuraikan pula perancangan sistem yang diusulkan yaitu alur dan desain sistem. Desain sistem berupa gambaran umum dari sistem yang akan dibuat. Selain itu bab III menjelaskan sumber data, instrumen penelitian dan pengujian sistem yang akan digunakan.

Bab IV Hasil dan Pembahasan

Bab IV berisi penjelasan tentang hasil dan pembahasan penelitian serta implikasi dari penelitian yang dilakukan. Hasil merupakan penjelasan tentang data kuantitatif yang dikumpulkan sesuai dengan metodologi yang telah ditetapkan. Pembahasan berisi penjelasan tentang pengolahan data dan interpretasinya, baik dalam bentuk deskriptif ataupun penarikan inferensinya. Implikasi Penelitian merupakan suatu penjelasan tentang tindak lanjut Penelitian yang terkait dengan aspek sistem, maupun aspek Penelitian lanjutan.

Bab V Kesimpulan dan Saran

Bab V peneliti menjelaskan tentang kesimpulan dari hasil pembahasan dan saran untuk pengembangan penelitian di masa yang akan datang.

BAB II

TINJAUAN PUSTAKA

A. LANDASAN TEORI

1. Budidaya / Aquaculture

Budidaya perikanan adalah usaha pemeliharaan dan pengembangbiakan ikan atau organisme air lainnya. Budidaya perikanan disebut juga sebagai budidaya perairan atau akuakultur mengingat organisme air yang dibudidayakan bukan hanya dari jenis ikan saja tetapi juga organisme air lain seperti kerrang, udang maupun tumbuhan air. Istilah akuakultur yang diambil dari istilah dalam bahasa inggris Aquaculture. Berikut definisi akuakultur menurut beberapa sumber.

Akuakultur merupakan suatu proses pembiakan organisme perairan dari mulai proses produksi, penanganan hasil sampai pemasaran (Wheaton, 1977).

Akuakultur merupakan upaya produkso biota atau organisme perairan melalui penerapan Teknik domestika (membuat kondisi lingkungan yang mirip dengan habitat asli organisme yang dibudidayakan), penumbuhan hingga pengelolaan usaha yang berorientasiekonomi (Bardach, dkk., 1972).

Akuakultur merupakan proses pengaturan dan perbaikan organisme akuatik untuk kepentingan konsumsi manusia (Webster's Dictionary, 1990).

Akuakultur terbagi menjadi tiga bagian yaitu, akuakultur air tawar (freshwater aquaculture), air payau (brackishwater aquaculture) , dan air laut (mariculture).

Akuakultur adalah seni dan teknologi pemeliharaan atau budidaya biota (organisme) air dalam perairan yang terkontrol maupun semi-terkontrol. Hal tersebut tidak terlepas dari kenyataan bahwa tidak seluruh kegiatan budidaya dapat dikontrol manusia tetapi tergantung pada alam, seperti budidaya rumput laut, budidaya kerrang Mutiara, budidaya ranching, budidaya penangkaran dan sebagainya (Stickney, 1979).

Akuakultur adalah usaha memelihara ikan secara rasional termasuk mengontrol pertumbuhan dan perkembangbiakannya (produksi dan kualitas) (Huet, 1972).

2. Kepiting

Kepiting adalah hewan golongan krustasea yang termasuk ke dalam ordo Decapoda, Subordo Pleocyemata, dan infraordo Brachyura, yang umumnya dicirikan dengan adanya tonjolan “ekor” yang sangat pendek (Latin: brachys = pendek, oura = ekor), atau dengan ciri bagian abdomen yang mengecil yang seluruhnya terlindung di bawah thorax (Rusdi dan Hanafi, 2009).

Tubuh kepiting umumnya dilindungi oleh cangkang luar (eksoskeleton) yang tebal, dan memiliki sepasang senjata berupa cakar tunggal (chelae). Kepiting dapat ditemukan diseluruh lautan di duniam sedangkan kepiting yang hidup di air tawar atau darat, kebanyakan hidup di daerah tropis. Kepiting dapat ditemukan dalam berbagai ukuran, mulai dari kepiting kacang (pea crab) yang memiliki ukuran lebar hanya beberapa millimeter saja, sampai dengan kepiting laba-laba Jepang, yang memiliki rentangan kaki sampai dengan 4 meter.

Sekitar 850 spesies kepiting adalah kepiting air tawar, kepiting datar atau kepiting semi-terrestrial, mereka dapat ditemukan diseluruh wilayah tropis maupun sub-tropis. Fosil kepiting pertama yang pernah ditemukan diduga berasal dari masa Jurassic. Spesies Carboniferous *inocaris* yang meskipun hanya dikenali dari kerapasnya, diduga merupakan kepiting prasejarah. Radiasi fosil kepiting yang ditemukan berasal dari jaman bantuan kapur (Cretaceous) dan jaman sesudahnya diduga ada kaitannya dengan masa perpecahannya belahan bumi selatan (Gondwana) atau radiasi dari fosil ikan bertulang, yang merupakan predator utama kepiting.

Kepiting seringkali menunjukkan tanda-tanda seksual dimorfisme. Kepiting jantang seringkali memiliki cakar yang lebih besar, suatu kecenderungan yang sering terjadi pada kepiting *Uca* (Fiddler crab) genus *Ocypodidae*. Kepiting *Uca* jantan memiliki cakar yang tumbuh sangat besar, yang digunakan untuk berkomunikasi, khususnya untuk menarik

perhatian kepiting betina. Perbedaan lain yang mencolok adalah bentuk perut (pleon). Pada hampir semua kepiting *Uca* jantan memiliki pleon yang sempit dan berbentuk segitiga, sementara pada kepiting betina memiliki pleon yang lebih besar dan berbentuk bulat. Hal ini menunjukkan bahwa kepiting betina mengerami telur-telurnya yang telah dibuahi di pleopod.

Kepiting biasanya berjalan dengan arah menyamping. Hal ini dikarenakan sendi-sendi kaki kepiting yang hanya memungkinkan cara berjalan kepiting yang lebih efisien ke arah samping. Tetapi terdapat pula beberapajenis kepiting yang berjalan dengan arah maju maupun mundur. Tetapi terdapat pula beberapa jenis kepiting yang berjalan dengan arah maju ataupun mundur, contohnya beberapa kepiting yang termasuk family Rannidae: *Libinia emarginata* dan *Mictyris platycheles*. Beberapa jenis kepiting, seperti family Portunidae dan Matutidae, juga memiliki kemampuan berenang.

Hampir semua kepiting merupakan hewan yang aktif dengan pola tingkah laku yang kompleks. Kepiting-kepiting dapat berkomunikasi dengan cara memukul-mukul atau melambai-lambaikan capit mereka. Kepiting cenderung bersifat agresif terhadap kepiting-kepiting lainnya dan kepiting-kepiting jantan seringkali berkelahi hanya untuk memperebutkan lubang perlindungan, berupa daerah karang berbatu, gua-gua dan celah-celah di daerah lepas pantai.

Kepiting merupakan hewan omnivore. Makanan utama kepiting berupa ganggang, selain dapat pula memakan makanan lainnya berupa moluska, cacing, jamur, bakteri, detritus dan jenis krustasea lainnya, tergantung pada spesies kepiting dan ketersediaan makanan. Bagi kebanyakan kepiting, campuran makanan berupa nabati dan hewani menghasilkan pertumbuhan dan kesehatan yang paling baik.

Kepiting diketahui mempunyai sifat bekerja sama dalam menyediakan makanan dan perlindungan bagi keluarga, dan selama musim pemijahan kepiting-kepiting jantan akan mencari lokasi yang nyaman bagi kepiting-kepiting betina untuk melepaskan telur-telur mereka. Berdasarkan data yang diketahui bahwa kepiting memberikan kontribusi sebanyak 20% dari semua jenis krustasea laut, baik berupa hasil tangkapan, budidaya, maupun yang dikonsumsi di seluruh dunia. Volume tersebut mencapai sekitar 1 1/2 juta ton per tahun. Seperlima dari jumlah tersebut berasal dari spesies *Portunus Trituberculatus*. Jenis-jenis kepiting lainnya yang memiliki nilai ekonomis penting adalah *Portunus pelagicus*, beberapa spesies dalam genus *Chionoecetes*, kepiting biru (*Callinectes sapidus*), *Charybdis* spp., *Cancer Pagurus*, kepiting Dungeness (*Metacarcinus magister*) dan kepiting bakau (*Scylla serrata*), yang masing-masing jenis dapat menghasilkan produk lebih dari 20.000 ton setiap tahun.

3. Kepiting Lunak (Soft Shell Crab)

Kepiting soka atau kepiting lunak adalah sebutan kepada kepiting yang baru melewati tahap ganti kulit atau molting. Pada saat ganti kulit,

cangkang kepiting yang keras di tinggalkan dan muncul cangkang baru yang masih lunak. Cangkang baru yang lunak ini akan mengeras beberapa saat setelah garis kulit. Karena itu, setelah ganti kulit, kepiting harus di panen. Keterlambatan dalam mengangkat lebih dari 4 – 6 jam setelah molting mengakibatkan kulit kepiting mengeras kembali.

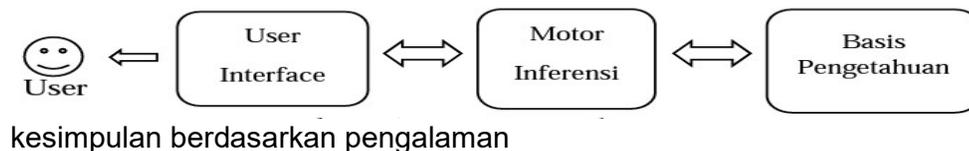
Pemberian hormon melalui pakan adalah mencampurkan hormone ke dalam pakan, sehingga hormone akan masuk ke dalam tubuh kepiting melalui pakan. Setelah proses perangsangan molting, kepiting disebar kedalam wadah pemeliharaan.

4. Artificial Intelligence

Artificial Intelligence (AI) atau yang biasa disebut Kecerdasan Buatan didefinisikan sebagai kecerdasan yang ditunjukkan oleh sebuah mesin atau software. Lebih spesifik lagi, menurut Kusumadewi (2003) menyatakan kecerdasan buatan merupakan salah satu bagian dalam ilmu komputer yang membuat agar mesin (komputer) dapat melakukan pekerjaan seperti dan sebaik yang dilakukan manusia. Kecerdasan diciptakan dan dimasukkan ke dalam suatu mesin (komputer) agar dapat melakukan pekerjaan seperti yang dapat dilakukan manusia. Beberapa macam bidang yang menggunakan kecerdasan buatan antara lain sistem pakar (expert system), permainan komputer (games), logika fuzzy, jaringan saraf tiruan (artificial neural system), robotika (robotics), pengolahan bahasa alami (natural language processing), pengenalan pola

(pattern recognition), dan pengenalan suara (speech recognition) (Simarmata, 2006).

Kecerdasan buatan bertujuan untuk membuat komputer menjadi lebih cerdas, bisa bernalar, dan berguna untuk manusia. Kecerdasan buatan dapat juga digunakan untuk membantu meringankan beban kerja. Kecerdasan buatan ditujukan dalam perancangan otomatisasi tingkah laku cerdas dalam sistem kecerdasan komputer. Pengaplikasian kecerdasan buatan terdiri dari 2 bagian utama yang sangat dibutuhkan, yaitu (Kusumadewi, 2003): a. Basis Pengetahuan (Knowledge Base), berisi fakta-fakta, teori, pemikiran dan hubungan antara satu dengan lainnya. b. Motor Inferensi (Inference Engine) yaitu kemampuan untuk menarik



kesimpulan berdasarkan pengalaman

Gambar 2. 1 Konsep Kecerdasan Buatan

Menurut Rusell dan Norvig (2010 : 2) terdapat empat macam pendekatan dalam AI, yaitu :

1. Thinking Humanly, yaitu sistem yang menangkap pemikiran psikologis, misalnya melalui eksperimen.
2. Acting Humanly, yaitu sistem dengan pendekatan menirukan tingkah laku manusia.

3. Thinking Rationally, yaitu sistem dengan penalaran komputasi.
4. Acting Rationally, yaitu sistem yang bertindak untuk mencapai hasil terbaik atau, ketika terdapat ketidakpastian, mengeluarkan hasil terbaik yang diharapkan.

5. Machine Learning

Machine Learning (ML) adalah cabang dari kecerdasan buatan (artificial intelligence) merupakan disiplin ilmu yang mencakup perancangan dan pengembangan algoritma yang memungkinkan komputer untuk mengembangkan perilaku yang didasarkan kepada data empiris. Machine Learning juga dapat diartikan sebagai disiplin ilmu yang menugaskan komputer untuk belajar dan bertindak seperti manusia, dan meningkatkan kemampuan belajarnya dari waktu ke waktu secara otomatis, dengan cara menyuplai data dan informasi sebagai bentuk pengalaman dan interaksi dunia nyata. Sedangkan Mitchell menjelaskan bahwa Machine Learning adalah kemampuan komputer untuk melakukan pembelajaran dari pengalaman (experience) E terhadap tugas (task) T yang dibebankan dengan kinerjanya (performance) P yang terukur (Mitchell, 1997). Machine Learning dapat dikelompokkan menjadi 3 kategori utama yaitu:

1. Pembelajaran terarah (supervised learning), merupakan suatu pembelajaran yang terawasi dimana jika output yang diharapkan telah diketahui sebelumnya. Biasanya pembelajaran ini dilakukan dengan

menggunakan data yang telah ada (Pustejovsky, 2012). Supervised learning dibagi menjadi 2 bagian, yaitu klasifikasi dan regresi.

2. Pembelajaran tidak terarah (unsupervised learning), merupakan pembelajaran yang tidak terawasi dimana tidak memerlukan target output. Pada metode ini tidak dapat ditentukan hasil seperti apa yang diharapkan selama proses pembelajaran, nilai bobot yang disusun dalam proses range tertentu tergantung pada nilai output yang diberikan. Tujuan metode unsupervised learning adalah agar kita dapat mengelompokkan unit-unit yang hamper sama dalam satu area tertentu. Pembelajaran ini biasanya sangat cocok untuk klasifikasi pola (Pustejovsky, 2012). Unsupervised learning terbagi menjadi 2 bagian yaitu asosiasi dan clustering.
3. Reinforcement learning bertujuan untuk menggunakan pengamatan dan mengumpulkan data melalui interaksi langsung dengan lingkungan untuk mengambil tindakan yang akan memaksimalkan reward dan meminimalkan resiko. Algoritma reinforcement learning terus belajar dari lingkungan secara iterative. Dalam prosesnya, agent belajar dari pengalamannya tentang lingkungan sampai mengeksplorasi keseluruhan kemungkinan keadaan. Dengan menggunakan algoritma ini, mesin dilatih untuk membuat keputusan yang lebih spesifik sehingga akan didapatkan keputusan yang akurat.

Ciri khas dari Machine Learning adalah adanya proses pelatihan, pembelajaran, atau training. Sehingga ML sangat membutuhkan data

untuk dapat dipelajari atau yang biasa disebut data training. Peter Harington (2012) memaparkan beberapa proses untuk membangun sebuah sistem Machine Learning, yaitu :

1. Mengumpulkan data (Collect data)
2. Mempersiapkan data masukan (Prepare input data)
3. Menganalisis data masukan (Analyse input data)
4. Mengikutsertakan keterlibatan manusia (Human involvement)
5. Melatih algoritma (Train algorithm)
6. Menguji algoritma (Test algorithm)
7. Menggunakan algoritma (Use it)

6. Klasifikasi

Pada penelitian ini, kategori utama unsupervised learning. Machine Learning yang digunakan adalah Hal itu karena sistem dibuat dengan memanfaatkan data yang sudah ada untuk melakukan prediksi. Karena prediksi pada sistem pendeteksi ini bersifat diskrit, maka proses yang digunakan adalah klasifikasi. Teknik klasifikasi telah banyak digunakan di berbagai permasalahan dalam suatu penelitian. training Klasifikasi merupakan suatu metode pengelompokan data yang akan mempelajari data dengan menggunakan algoritma pengklasifikasian.

Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang

tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu : pertama, Pembangunan model sebagai prototype untuk disimpan sebagai memori dan kedua, Penggunaan model tersebut untuk melakukan pengenalan/ klasifikasi/ prediksi pada suatu objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang mudah disimpan (Prasetyo, Eko. 2012).

Klasifikasi juga dapat diartikan sebagai suatu proses untuk mengelompokkan sejumlah data ke dalam kelas-kelas tertentu yang sudah diberikan berdasarkan kesamaan sifat dan pola yang terdapat dalam data-data tersebut. Secara umum, proses klasifikasi dimulai dengan diberikannya sejumlah data yang menjadi acuan untuk membuat aturan klasifikasi data. Data-data ini biasa disebut dengan training sets. Dari training sets tersebut kemudian dibuat suatu model untuk mengklasifikasikan data. Model tersebut kemudian digunakan sebagai acuan untuk mengklasifikasikan data-data yang belum diketahui kelasnya yang biasa dengan test sets (Rifqi, Maharani, & Shaufiah, 2011).

Data input untuk klasifikasi adalah koleksi dari record disebut . Setiap record dikenal sebagai instance atau contoh, yang ditentukan oleh sebuah tuple (x,y) , dimana x adalah himpunan atribut dan y adalah atribut tertentu, yang dinyatakan sebagai label kelas (juga dikenal sebagai kategori atau atribut target). Klasifikasi adalah tugas pembelajaran sebuah fungsi target f yang memetakan setiap

himpunan atribut x ke salah satu label kelas y yang telah didefinisikan sebelumnya. Fungsi target juga dikenal secara informal sebagai model klasifikasi. Model klasifikasi berguna untuk keperluan berikut :

1. Pemodelan Deskriptif

Model klasifikasi dapat bertindak sebagai alat penjelas untuk membedakan objek-objek dari kelas-kelas yang berbeda. Sebagai contoh untuk para ahli Biologi, model deskriptif yang meringkas data.

2. Pemodelan Prediktif

Model klasifikasi juga dapat digunakan untuk memprediksi label kelas dari record yang tidak diketahui. Seperti pada Gambar 1 tampak sebuah model klasifikasi dapat dipandang sebagai kotak hitam yang secara otomatis memberikan sebuah label ketika dipresentasikan dengan himpunan atribut dari record yang tidak diketahui.



Gambar 2. 2 Klasifikasi sebagai pemetaan sebuah himpunan atribut input x ke dalam label kelasnya.

Contoh aplikasi yang sering ditemui adalah pengklasifikasian jenis hewan, yang mempunyai sejumlah atribut. Dengan atribut tersebut, jika ada hewan baru, kelas hewannya bisa langsung diketahui. Contoh lain adalah bagaimana melakukan diagnosis penyakit kulit kanker melanoma (Prasetyo, Eko. 2012), yaitu dengan melakukan pembangunan model berdasarkan data latih yang ada, kemudian menggunakan model tersebut untuk mengidentifikasi penyakit pasien baru sehingga diketahui apakah pasien tersebut menderita kanker atau tidak. Ada beberapa algoritma yang dapat digunakan untuk melakukan klasifikasi. Algoritma klasifikasi akan menghasilkan pola atau aturan yang dapat digunakan untuk memprediksi kelas. Beberapa di antaranya yaitu, Naïve Bayes Classifier, Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbour (KNN), dan Random Forest Classifier.

7. K-Nearest Neighbours Classifier

KNearest Neighbors algoritma supervised Naïve Bayes Classifier (KNN) adalah suatu metode yang menggunakan dimana hasil dari query instance yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada KNN. Tujuan dari algoritma ini adalah mengklasifikasi objek baru berdasarkan atribut dan training sample . Classifier tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori.

Algoritma KNearest Neighbor (KNN) juga dapat diartikan sebagai metode yang digunakan untuk mengelompokkan objek

berdasarkan contoh pelatihan terdekat di ruang fitur. KNearest Neighbor merupakan jenis yang paling dasar dari contoh based learning atau lazy learning juga termasuk kelompok *instance-based* learning . KNearest Neighbor dilakukan dengan mencari kelompok objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing .

Algoritma metode KNN sangatlah sederhana, bekerja dengan berdasarkan pada jarak terpendek dari query instance ke training sample KNN nya. Setelah mengumpulkan KNN, kemudian d untuk menentukan diambil mayoritas dari KNN untuk dijadikan prediksi dari query instance. Data untuk algoritma KNN terdiri dari beberapa atribut multivariate X_i yang akan digunakan untuk mengklasifikasikan Y . Data dari KNN dapat dalam skala ukuran apapun, dari ordinal ke nominal.

Menurut Kusrini dan Emma (2009) algoritma K-Nearest Neighbor adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama dengan berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada yang memiliki kesamaan (similarity). Tujuan dari algoritma ini untuk mengklasifikasikan objek baru berdasarkan atribut dan training sample. Classifier tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Berikut ini adalah langkah-langkah menghitung K-Nearest Neighbor (kNN) (Sari, 2011) :

1. parameter k (jumlah tetangga terdekat)
2. Hitung jarak antara data yang masuk dan semua sampel training yang sudah ada dengan Euclidian Distance. Formula untuk mencari Euclidian Distance adalah sebagai berikut:
3. Tentukan K label data yang mempunyai jarak yang minimal
4. Klasifikasikan data baru ke dalam label data yang mayoritas

Pada fase training, algoritma ini hanya melakukan penyimpanan vector-vector fitur dan klasifikasi data training sample. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk testing data (yang klasifikasinya tidak diketahui). Jarak dari vektor baru yang ini terhadap seluruh vektor training sample dihitung dan sejumlah k buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut (Sikki, 2009).

Sebagai contoh, untuk mengestimasi $p(x)$ dari n training sample dapat memusatkan pada sebuah sel disekitar x dan membiarkannya tumbuh hingga meliputi k samples. Samples tersebut adalah KNN dari x . Jika densitasnya tinggi di dekat x , maka sel akan berukuran relatif kecil yang berarti memiliki resolusi yang baik. Jika densitas rendah, sel akan tumbuh lebih besar, tetapi akan berhenti setelah memasuki wilayah yang memiliki densitas tinggi (Sikki, 2009).

Nilai k yang terbaik untuk algoritma ini tergantung pada data. Secara umum, nilai k yang tinggi akan mengurangi efek noise pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi

semakin kabur. Nilai k yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan cross-validation . Kasus khusus dimana klasifikasi diprediksikan berdasarkan data training yang paling dekat (dengan kata lain, $k = 1$) disebut algoritma neighbor.

Ketepatan algoritma KNN sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Riset terhadap algoritma ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur agar performa klasifikasi menjadi lebih baik.

KNN memiliki beberapa kelebihan yaitu ketangguhan terhadap training data yang memiliki banyak noise dan efektif apabila data training besar. Sedangkan, kelemahan KNN adalah KNN perlu menentukan nilai dari parameter k (jumlah dari tetangga terdekat), training berdasarkan jarak tidak jelas mengenai jenis jarak apa yang harus digunakan dan atribut mana yang harus digunakan untuk mendapatkan hasil terbaik, dan biaya komputasi cukup tinggi karena diperlukan perhitungan jarak dari tiap query instance pada keseluruhan training sample.

8. Support Vector Machine

Support Vector Machine atau SVM adalah suatu sistem pembelajaran yang menggunakan ruang hipotesis dari suatu fungsi linear dalam suatu ruang dimensi berfitur tinggi yang dikembangkan oleh Boser, Guyon, Vapnik, dan pertama kali dipresentasikan pada

tahun 1992 di *12 Annual Workshop on Computational Learning Theory*.

SVM merupakan metode klasifikasi yang kini banyak dikembangkan dan diterapkan. Metode ini berasal dari teori pembelajaran statistik yang menjanjikan dan memberikan hasil yang lebih baik dibandingkan dengan metode yang lainnya. SVM bekerja sangat baik pada himpunan data berdimensi tinggi. SVM yang menggunakan teknik kernel harus memetakan data asli dari dimensi asalnya menjadi dimensi lain yang relatif lebih tinggi. Pada metode ANN (Artificial Neural Network), selama proses pelatihan semua data latih akan dipelajari. Berbeda dengan SVM yang hanya sejumlah data terpilih saja yang digunakan dalam klasifikasi yang akan dipelajari. Berbeda juga dengan Nearest Neighbour semua data training yang disimpan yang akan digunakan pada saat prediksi, SVM hanya menyimpan sebagian kecil data latih untuk digunakan pada saat prediksi. Hal inilah yang menjadi kelebihan SVM karena tidak semua data latih dilibatkan pada saat pelatihan (Prasetyo, 2014).

9. Random Forest Classifier

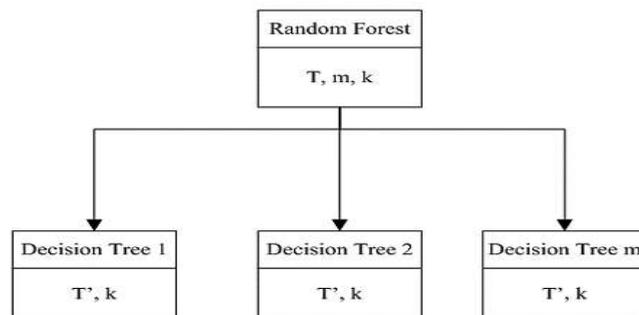
Random forest merupakan metode bagging yaitu metode yang membangkitkan sejumlah tree dari data sample dimana pembuatan satu tree pada saat training tidak bergantung pada tree sebelumnya kemudian keputusan diambil berdasarkan voting terbanyak (Wibowo, Saikhu, & Soelaiman, 2016).

Dua konsep yang menjadi dasar dari random forest adalah membangun ensemble dari tree via bagging dengan replacement dan penyeleksian fitur secara acak untuk tiap tree yang dibangun. Pertama, setiap sample yang diambil dari dataset untuk training tree bisa dipakai lagi untuk training tree yang lain. Kedua, fitur yang digunakan pada saat training untuk tiap tree merupakan subset dari fitur yang dimiliki oleh dataset (Wibowo et al., 2016).

Klasifikasi berbasis ensemble akan mempunyai performa yang maksimal jika antar basic learner mempunyai korelasi yang rendah. Sebuah ensemble harus membangun basic learner yang lemah, karena learner yang kuat kemungkinan besar akan mempunyai korelasi yang tinggi dan biasanya juga menyebabkan overfit, sedangkan random forest meminimalkan korelasi serta mempertahankan kekuatan klasifikasi dengan cara melakukan pengacakan pada proses training, yaitu dengan memilih sejumlah fitur secara acak dari semua fitur yang ada pada setiap melakukan training tree, kemudian menggunakannya menggunakan fitur-fitur yang terpilih untuk mendapatkan percabangan tree yang optimal. Berbeda dengan proses training tree pada decision tree biasa, proses training tree yang menjadi bagian dari random forest tidak menggunakan proses pruning akan tetapi percabangan akan terus dilakukan sampai ukuran batas leaf tercapai (Wibowo et al., 2016).

Random forest mempunyai dua parameter utama, yaitu: m jumlah tree yang akan dipakai dan k yaitu maksimal banyaknya fitur yang

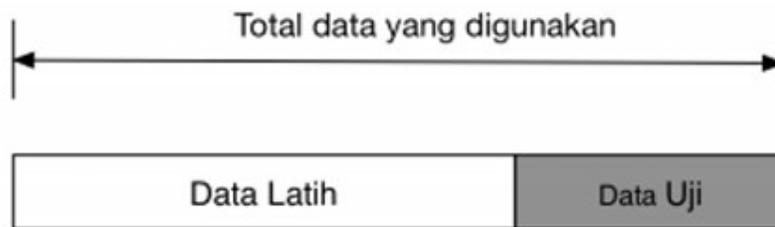
dipertimbangkan ketikan proses percabangan. Semakin banyak nilai m maka semakin bagus hasil klasifikasi, sedangkan untuk nilai k direkomendasikan sebesar akar kuadrat atau logaritma dari jumlah total fitur (Wibowo et al., 2016).



Gambar 2. 3 Ilustrasi Random Forest

10. Pembagian Data

Ada beberapa macam cara untuk melakukan pembagian data guna membangun model klasifikasi dari data yang telah dikumpulkan. Salah satu metode di antaranya adalah Data Splitting. Data splitting ini maksudnya adalah membagi data menjadi dua kelompok, yaitu kelompok data training dan kelompok data testing. Data training digunakan untuk melatih klasifikator, sedangkan data uji digunakan untuk memperkirakan error rate dari klasifikator yang telah dilatih. Gambar 2.4 mengilustrasikan bagaimana data dibagi menjadi dua kelompok.



Gambar 2. 4 Metode Pembagian Data Menjadi Dua Kelompok Data

Kekurangan dari metode ini terutama adalah

- a) untuk data yang jumlahnya sedikit, keleluasaan untuk menyisihkan data uji sangat sempit,
- b) karena percobaan data latihdata uji ini hanya sekali, kemungkinan er ror rate yang tinggi akan sangat mungkin, terlebih apabila pembagian data kurang memberikan hasil yang maksimal.

11. Confusion Matrix

Confusion matrix adalah salah satu metode pengukuran keputusan yang paling banyak digunakan dalam supervised machine learning. Confusion matrix memvisualisasikan nilai tingkat kebingungan dari algoritma pada setiap kelas yang berbeda dan tergantung pada algoritma klasifikasi. Tujuan dari confusion matrix adalah untuk melakukan perhitungan akurasi pada konsep data mining. Evaluasi dengan confusion matrix menghasilkan nilai akurasi, presisi dan recall. Nilai akurasi adalah persentase ketepatan record data yang diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. Presisi atau confidence

merupakan proporsi kasus yang diprediksi positif yang juga hasilnya positif benar pada data yang sebenarnya. Recall atau sensitivity adalah proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar (Mayadewi & Rosely, 2015).

Tabel 2. 1 Confusion Matrix

<i>Aktual</i>	<i>Prediksi</i>	
	+	-
+	<i>True positives (A)</i>	<i>False Negatives (B)</i>
-	<i>False Positives (C)</i>	<i>True negatives (D)</i>

Untuk dapat menghitung akurasi pada table *confusion matrix* dapat menggunakan rumus sebagai berikut :

$$Akurasi = \frac{(A + D)}{(A + B + C + D)} \quad (2. 1)$$

Presisi (*Precision*) merupakan rasio item relevan yang dipilih terhadap semua item yang terpilih. Sehingga presisi dapat diartikan sebagai kecocokan antara permintaan informasi dengan jawaban terhadap permintaan tersebut. Untuk dapat menghitung presisi dapat digunakan rumus sebagai berikut:

$$Precision = \frac{(A)}{(C + A)} \quad (2. 2)$$

Recall merupakan rasio dari item relevan yang dipilih terhadap total jumlah item relevan yang tersedia. *Recall* dapat dihitung dengan rumus sebagai berikut:

$$Recall = \frac{(A)}{(A + D)} \quad (2. 3)$$

Presisi dan *Recall* dapat diberi nilai dengan menggunakan perhitungan persentase (1-100%) atau dengan menggunakan bilangan antara 0-1.

F1 Score merupakan perbandingan rata-rata presisi dan recall (Powers & Ailab, 2011).

$$F1\ Score = \frac{precision \times recall}{precision + recall} \quad (2. 4)$$

Akurasi memiliki tingkat nilai diagnosa yaitu :

- a. Akurasi dengan nilai 0.90-1.00 dapat dikategorikan sebagai *excellent classification*
- b. Akurasi dengan nilai 0.80-0.90 dapat dikategorikan sebagai *good classification*

- c. Akurasi dengan nilai 0.70-0.80 dapat dikategorikan sebagai *fair classification*
- d. Akurasi dengan nilai 0.60-0.70 dapat dikategorikan sebagai *poor classification*
- e. Akurasi dengan nilai 0.50-0.0 dapat dikategorikan sebagai *failure*

B. PENELITIAN TERKAIT

Dalam penelitian terkait, Faisal Fahri Ferdiansyah dkk, melakukan penelitian tentang klasifikasi dan pengenalan objek ikan dengan menggunakan teknologi machine learning metode algoritma support vector machine(SVM). Dengan mengklasifikasi jumlah 5 ikan hias menggunakan algoritma SVM dengan 250 jumlah data uji citra. Hasil menunjukkan Nilai Accuracy 50%, Precision 90%, Recall 47% dan f1 score 63,94% (Ferdiansyah et al., 2020).

Peneliti sebelumnya mengusulkan dalam mendeteksi dan mengenali jenis objek ikan menggunakan metode ORB dan KNN. Dalam pengujiannya dataset yang digunakan berjumlah 165 citra, yang terdiri dari 40 data testing dan 125 data training, hasil pengujian menunjukkan bahwa metode yang diusulkan dapat mengklasifikasi 39 citra ikan dari 40 data testing dengan tingkat akurasi total hingga 97,5% adalah menggunakan metode k- nearest neighbor (k-NN) (Ramadhani & Murti, 2018).

Penelitian tentang klasifikasi ikan menggunakan teknologi machine learning juga pernah dilakukan oleh Mohamed Mostafa M. Fouad dkk,

dengan pengujiannya terhadap beberapa algoritma machine learning yaitu artificial neural networks (ANN) and k- nearest neighbor (k-NN) dan support vector machine(SVM) menunjukkan algoritma terbaik dengan akurasi 94\% adalah algoritma support vector machine(SVM)(Fouad et al., 2014).

Selain klasifikasi ikan, peneliti sebelumnya mengimplementasikan sistem klasifikasi penyakit udang menggunakan teknologi machine learning, dari beberapa algoritma machine learning yang di gunakan menunjukkan algoritma support vector machine(SVM) mendapatkan akurasi paling tinggi yakni 81.27\% (Quach et al., 2020).

Teknologi mechine learning dengan mengidentifikasi kepiting, pernah dilakukan oleh Rozniza Ali dkk, dalam penelitian nya peneliti melakukan identifikasi terhadap parasit kepiting menggunakan beberapa algoritma machine learning,yakni logistic regression (LR), k-nearest neighbors (kNN), Gaussian Naive Bayes (GNB), support vector machine (SVM), and linear discriminant analysis (LDA). hasilnya menunjukkan algoritma LDA dan GNBmendapatkan akurasi tertinggi. (Ali et al., 2021)

Dalam penelitian kami, kami mengembangkan identifikasi terhadap kepiting molting dengan menggunakan teknologi machine lerning untuk membantu petambak dalam memonitoring budidaya kepiting cangkang lunak.

C. STATE OF THE ART

Beberapa penelitian tentang sistem Otomasi pakan perikanan yang telah dilakukan dan yang akan diusulkan dapat dilihat pada Tabel 1 *state of the art*.

Tabel 2. 2 state of the art

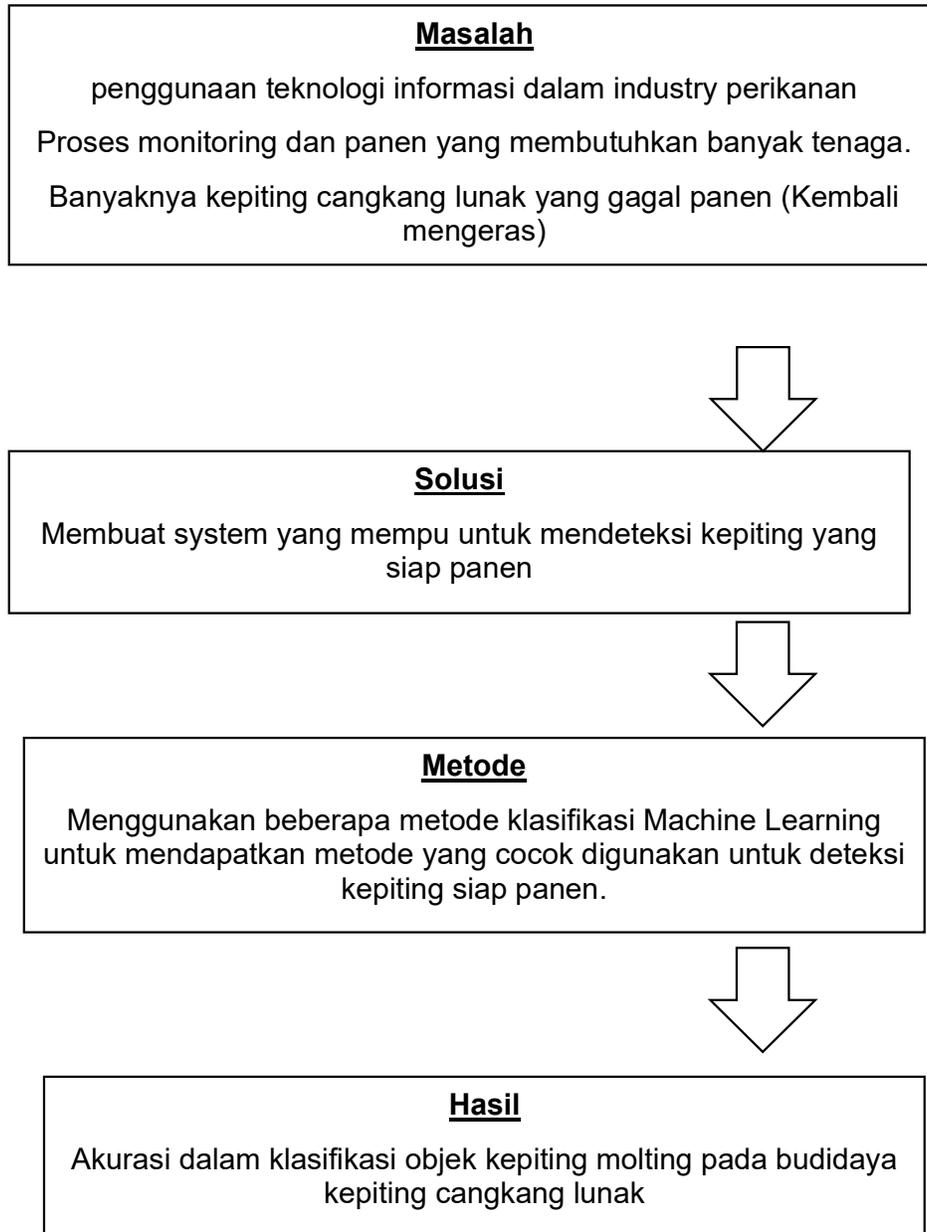
	NO	PENULIS / TAHUN	JUDUL	Metode Klasifikasi	Tinjauan Data	Objek	Akurasi
Penelitian yang telah di lakukan	1.	Faisal Fahri Ferdiansyah , Basuki Rahmat, Intan Yuniar / 2020	Klasifikasi Dan Pengenalan Objek Ikan Menggunakan Algoritma Support Vector Machine (Svm)	SUPPORT VECTOR MACHINE (SVM)	Klasifikasi citra	Ikan Hias	Akurasi 50 %, presisi 90%, recall 47% dan f1 score 63.94%
	2.	Mirza Ramadhani dan Darlis Heru Murti / 2018	Klasifikasi Ikan Menggunakan Oriented Fast And Rotated Brief (Orb) Dan K-Nearest	ORIENTED FAST AND ROTATED BRIEF (ORB) DAN K-NEAREST	Klasifikasi citra	Ikan	Akurasi 97.5%

	NO	PENULIS / TAHUN	JUDUL	Metode Klasifikasi	Tinjauan Data	Objek	Akurasi
			Neighbor (Knn)	NEIGHBOR (KNN)			
	3.	Mohamed Mostafa M. Fouad , Hossam M. Zawbaa , Nashwa El-Bendary , Aboul Ella Hassanien / 2013	Automatic Nile Tilapia Fish Classification Approach using Machine Learning Techniques Mohamed	artificial neural networks (ANN) dan k- nearest neighbor (k-NN) algorithms,	Klasifikasi citra	Ikan nila	Akurasi 94.4%
	4.	Luyi-Da Quach, Long Quach Hoang, Nghia Duong-Trung, Chi- Ngon Nguyen / 2019	Towards Machine Learning Approaches To Identify Shrimp Diseases Based On Description	Multilayer Perceptron, Random Forest, Logistic Regression, and Naïve Bayes.	Klasifikasi citra, penyakit udang	Udang	Akurasi 81.27%
	5.	Rozniza Ali, Muhamad Munawar Yusro, Muhammad	Machine learning with multistage classifiers for identification of of	logistic regression (LR), k-nearest neighbors (kNN),	Klasifikasi citra, penyakit kepiting bakau	Kepiting	Akurasi 100%

	NO	PENULIS / TAHUN	JUDUL	Metode Klasifikasi	Tinjauan Data	Objek	Akurasi
		Suzuri Hitam, Mhd Ikhwanuddin./2021	ectoparasite infected mud crab genus Scylla	Gaussian Naive Bayes (GNB), support vector machine (SVM), and linear discriminant analysis (LDA).			
	6.	R.A. Pramunendar, D.P.Prabowo, F.Alzami, R.A Megantara. / 2021	Penerapan Random Forest Untuk Pengenalan Jenis Ikan Berdasarkan Perbaikan Citra Clahe Dan Dark Channel Prior	Random Forest	Klasifikasi citra	Ikan	Akurasi 98.51%
Penelitian yang akan di lakukan	7.	Runal Rezkiawan/2021	Identifikasi Kepiting Molting menggunakan Klasifikasi Machine Learning	KNN, SVM dan RFC	Klasifikasi citra	Kepiting	

D. KERANGKA PIKIR

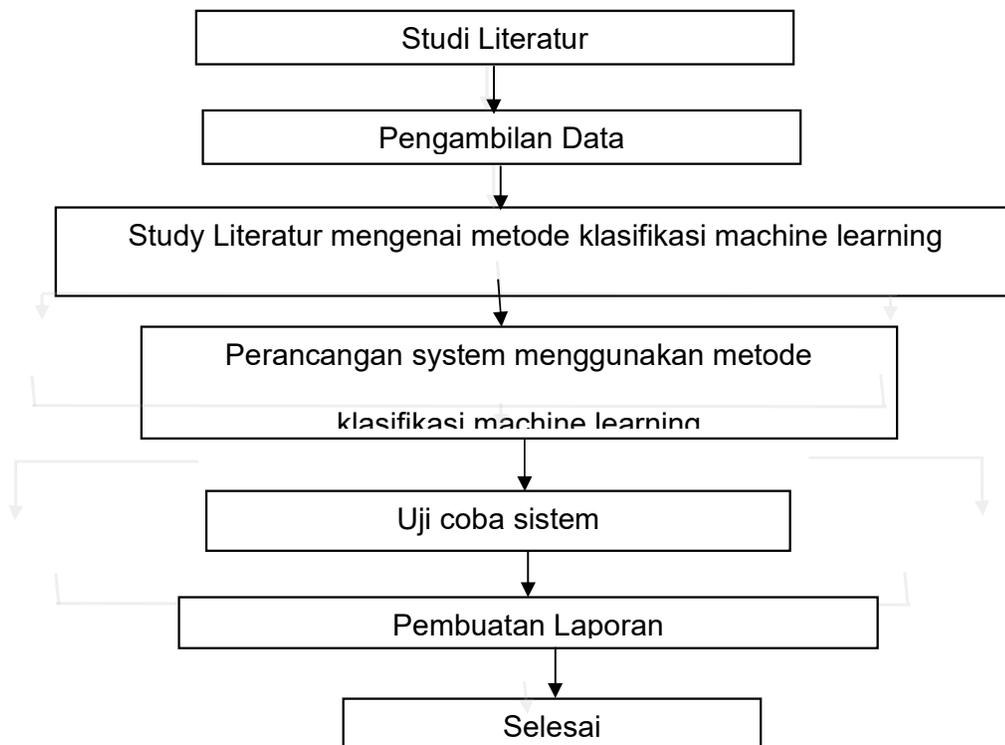
Adapun kerangka pikir pada penelitian ini ditunjukkan pada Gambar 3.



BAB III METODOLOGI PENELITIAN

A. TAHAPAN PENELITIAN

Penelitian ini bertujuan untuk mendapat akurasi terbaik dari beberapa algoritma machine learning yang di gunakan pada kepiting molting dan non-molting dengan jumlah data 1060. Pada tahapan ini dilakukan identifikasi terhadap kepiting molting dan non-molting menggunakan beberapa metode klasifikasi Algoritma Machine Learning. Setelah beberapa metode tersebut diujicoba dan di bandingkan menggunakan beberapa parameter dari masing-masing model klasifikasi.



Gambar 3. 1 Tahapan penelitian

B. WAKTU DAN LOKASI PENELITIAN

1. Waktu

Waktu penelitian akan dilaksanakan selama dari Juli 2021 sampai Oktober 2021.

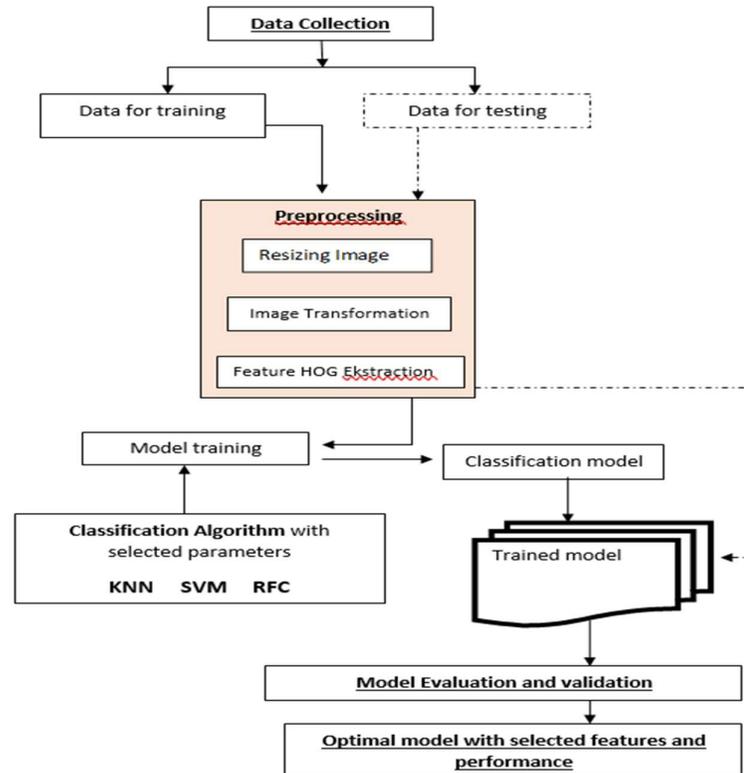
2. Lokasi

Pada penelitian ini mencakup pembuatan system identifikasi kepiting molting menggunakan klasifikasi machine learning yang akan dilakukan di laboratorium *Cloud Computing and Internet Engineering* Departemen Teknik Elektro Universitas Hasanuddin Gowa dan lokasi pengambilan data di Tambak kepiting di daerah barombong, Gowa.

C. JENIS PENELITIAN

Jenis penelitian ini adalah penelitian experimental yang bersifat analisis dengan penelusuran sumber-sumber tertulis (*library research*), dipadukan dengan pengumpulan data-data faktual di lapangan, dengan konsep dan data tersebut dikembangkan sistem identifikasi kepiting molting menggunakan klasifikasi machine learning yang akan digunakan peneliti untuk membuat keputusan algoritma machine learning terbaik pada budidaya kepiting cangkang lunak.

D. PERANCANGAN SISTEM



Gambar 3. 2 Pecancangan Sistem

Pada Gambar 3.1 memperlihatkan diagram perancangan sistem identifikasi kepiting molting di mana secara umum terdapat proses koleksi data atau pengambilan dataset, Preprocessing, identifikasi kepiting molting.

a. Koleksi gambar/Pengambilan data

Tahapan pertama dalam teknik image processing adalah pengambilan data. Pada penelitian ini, penulis

mengambil data dari tambak buudidaya kepiting cangkang lunak.

b. Pre-Processing

Pada tahapan ini di lakukan proses pre-processing .Tujuan dari proses ini yaitu untuk memudahkan proses pengolahan citra. Pada identifikasinya, data di cropping, grayscale.

c. Identifikasi kepiting molting

Pada tahap identifikasi dilakukan uji coba system dengan membandingkan beberapa algoritma klasifikasi machine learning untuk mendapat kan algoritma terbaik dalam mengidentifikasi kepiting molring.

E. SUMBER DATA

Tahapan pengumpulan data terdiri dari pengumpulan data primer dan pengumpulan data sekunder. Terdapat dua jenis data yang digunakan yaitu :

1. Data Primer

Pengambilan data primer berupa gambar yang telah diambil dari tempat penelitian kepiting. Jumlah citra yang digunakan yaitu sebanyak 1060 citra dimana 530 citra kepiting molting dan 530 citra kepiting non-

molting. Pada prosesnya 530 data citra dikelompokkan menjadi 2 kategori, yaitu data train dan data testing, di mana 500 data train dan 30 data uji.

2. Data Sekunder

Pengambilan data sekunder merupakan literatur yang berasal dari jurnal, prosiding, buku, maupun internet. Literature yang di kumpulkan berkaitan teori, konsep, dan metode klasifikasi Machine Learning pada sistem deteksi kepiting molting

F. INSTRUMENTASI PENELITIAN

1. Perangkat Keras

Perangkat keras terdiri dari yaitu

1. Laptop Legion
2. ESPCAM

2. Perangkat Lunak

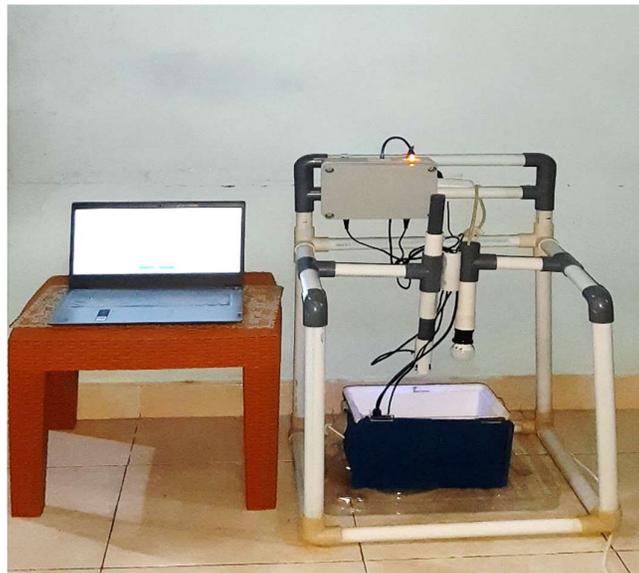
- a. Windows 10 Profesional 64 bit
- b. Python
- c. Jupyter Notebook

BAB IV

HASIL DAN PEMBAHASAN

A. Pengumpulan Dataset

Pada pengumpulan dataset kepiting, total dataset citra kepiting yaitu 1060 yang di kelompokkan menjadi 2 bagian sebagai berikut :



Gambar 4. 1 Akuisisi Data

1. Data Training

Pada tahapan ini dilakukan pembagian terhadap dataset citra kepiting. Sebelum melakukan training citra, dataset dibagi menjadi 2 kelas yakni kepiting molting dan non-molting. Untuk data training dari total citra kepiting yakni 500 dataset citra kepiting molting dan 500 dataset citra kepiting non-molting.

2. Data Testing

Tahap selanjutnya yakni membagi data testing menjadi 2 kelas yakni kepiting molting dan non-molting. Dari total dataset citra yang akan di training 1000, maka data testing yang digunakan yaitu 30 dataset citra untuk kepiting molting dan 30 dataset citra untuk kepiting non-molting.



Gambar 4. 2 Citra Kepiting Molting



Gambar 4. 3 Citra Kepiting non-molting

B. Preprocessing

Setelah melakukan pengelompokan data, maka tahap selanjutnya adalah melakukan pre-processing citra. Adapun tahap-tahap yang dilakukan yaitu:

1. *Resize Citra*

Tahap awal pra-pemrosesan data adalah resize image. Resize image diperlukan untuk menyeragamkan citra yang didapat sehingga mempermudah pengolahan tanpa menghilangkan kualitas citra.

Resize citra adalah proses mengubah ukuran piksel citra dari ukuran besar menjadi ukuran yang lebih kecil, agar lebih mudah untuk dilakukan proses

pendeteksian. Pada saat proses cropping, size yang dimiliki citra berbeda-beda dan masih berukuran besar sehingga dibutuhkan proses resize citra yang bertujuan untuk membuat ukuran matriks piksel setiap data citra dengan ukuran yang sama.

Berikut adalah ilustrasi proses resize baris dari citra grayscale yang berukuran 10 x 10 yang akan di-resize menjadi berukuran 5 x 10 dengan cara mencari nilai rata-rata dan nilai terdekat:

$$R = \begin{pmatrix} 160 & 161 & 164 & 81 & 74 & 156 & 166 & 165 & 168 & 168 \\ 160 & 163 & 155 & 59 & 50 & 121 & 166 & 171 & 173 & 175 \\ 163 & 166 & 102 & 76 & 61 & 62 & 135 & 174 & 186 & 178 \\ 166 & 156 & 88 & 51 & 56 & 95 & 77 & 135 & 195 & 177 \\ 168 & 116 & 54 & 95 & 61 & 76 & 56 & 150 & 200 & 185 \\ 167 & 148 & 66 & 64 & 62 & 74 & 137 & 194 & 200 & 192 \\ 152 & 83 & 67 & 56 & 58 & 135 & 187 & 182 & 197 & 195 \\ 152 & 83 & 50 & 56 & 64 & 127 & 80 & 72 & 89 & 158 \\ 189 & 191 & 161 & 83 & 40 & 96 & 94 & 69 & 97 & 161 \\ 199 & 198 & 200 & 169 & 90 & 99 & 184 & 154 & 173 & 159 \end{pmatrix}$$

Mencari rata-rata matriks R dari kolom pertama

$$\begin{array}{cccccc} 160 & 160 & 163 & 166 & 168 & 167 & 152 & 152 & 189 & 199 \\ \underbrace{\hspace{1.5em}} & \underbrace{\hspace{1.5em}} \\ 160 & & 164 & & 167 & & 152 & & 194 & \end{array}$$

Jadi, kolom pertama pada matriks R menjadi 160, 166, 167, 152, 199.

$$\begin{array}{cccccc} 161 & 163 & 166 & 156 & 166 & 156 & 83 & 83 & 191 & 198 \\ \underbrace{\hspace{1.5em}} & \underbrace{\hspace{1.5em}} \\ 162 & 161 & 132 & 83 & 194 & & & & & \end{array}$$

Mencari rata-rata matriks R dari kolom kedua

Jadi, kolom kedua pada matriks R menjadi 163, 166, 148, 83, 191.

$$\begin{array}{cccccc} 164 & 155 & 102 & 88 & 54 & 66 & 67 & 50 & 161 & 200 \\ \underbrace{\hspace{1.5em}} & \underbrace{\hspace{1.5em}} \\ 159 & 95 & 60 & 58 & 180 & & & & & \end{array}$$

Mencari rata-rata matriks R dari kolom ketiga

Jadi, kolom ketiga pada matriks R menjadi 155, 102, 66, 50, 161

$$\begin{array}{cccccc} 81 & 59 & 76 & 51 & 95 & 64 & 56 & 56 & 83 & 169 \\ \underbrace{\hspace{1.5em}} & \underbrace{\hspace{1.5em}} \\ 70 & 63 & 79 & 56 & 126 & & & & & \end{array}$$

Mencari rata-rata matriks R dari kolom keempat

Jadi, kolom keempat pada matriks R menjadi 81, 51, 64, 56, 169

$$\begin{array}{ccccc}
 74 & 50 & 61 & 56 & 61 & 62 & 58 & 64 & 40 & 90 \\
 \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & \\
 62 & & 58 & & 61 & & 61 & & 65 &
 \end{array}$$

Mencari rata-rata matriks R dari kolom kelima

Jadi, kolom kelima pada matriks R menjadi 74, 56, 61, 64, 90.

$$\begin{array}{ccccc}
 156 & 121 & 62 & 95 & 76 & 74 & 135 & 127 & 96 & 99 \\
 \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & \\
 138 & & 78 & & 75 & & 131 & & 97 &
 \end{array}$$

Mencari rata-rata matriks R dari kolom keenam

Jadi, kolom keenam pada matriks R menjadi 121, 62, 76, 135, 96.

$$\begin{array}{ccccc}
 166 & 166 & 135 & 77 & 56 & 137 & 187 & 80 & 94 & 184 \\
 \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & \\
 166 & & 106 & & 96 & & 133 & & 139 &
 \end{array}$$

Mencari rata-rata matriks R dari kolom ketujuh

Jadi, kolom ketujuh pada matriks R menjadi 166, 135, 56, 80, 184.

$$\begin{array}{ccccc}
 165 & 171 & 174 & 135 & 150 & 194 & 182 & 72 & 69 & 154 \\
 \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & & \underbrace{\hspace{1.5em}} & \\
 168 & & 154 & & 172 & & 127 & & 111 &
 \end{array}$$

Mencari rata-rata matriks R dari kolom kedelapan

Jadi, kolom kedelapan pada matriks R menjadi 171, 135, 194, 182, 69.

$$\begin{array}{ccccccccc} 168 & 173 & 186 & 195 & 200 & 200 & 197 & 89 & 97 & 173 \\ \underbrace{\hspace{1.5em}} & \underbrace{\hspace{1.5em}} \\ 170 & & 190 & & 200 & & 143 & & & 135 \end{array}$$

Mencari rata-rata matriks R dari kolom kesembilan

Jadi, kolom kesembilan pada matriks R menjadi 168, 186, 200, 197, 173.

$$\begin{array}{ccccccccc} 168 & 175 & 178 & 177 & 185 & 192 & 195 & 158 & 161 \\ \underbrace{\hspace{1.5em}} & \underbrace{\hspace{1.5em}} \\ 159 & & & & & & & & \end{array}$$

Mencari rata-rata matriks R dari kolom kesepuluh

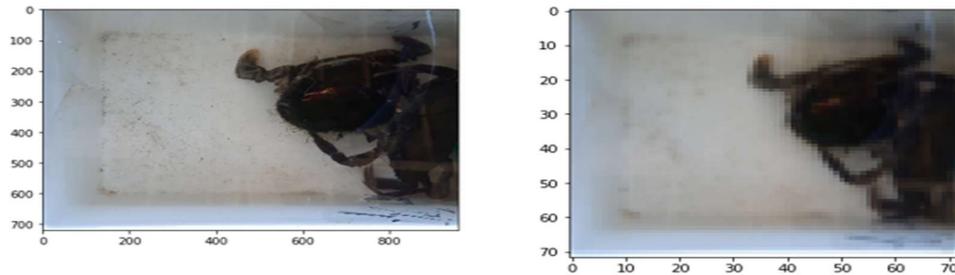
Jadi, kolom kesepuluh pada matriks R menjadi 168, 177, 185, 195, 161.

Sehingga matriks R yang diperoleh setelah proses resize menjadi matriks 5*10.:

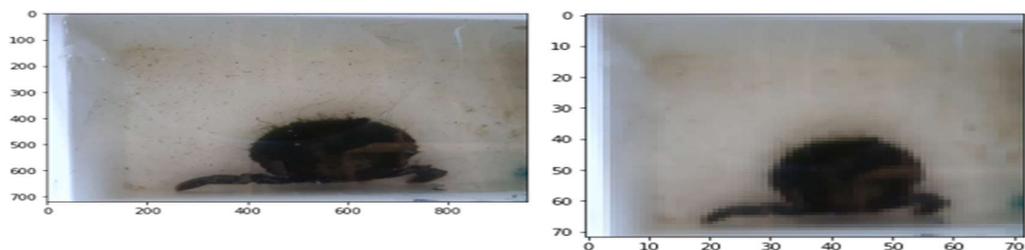
$$R = \begin{pmatrix} 160 & 163 & 155 & 81 & 74 & 121 & 166 & 171 & 168 & 168 \\ 166 & 166 & 102 & 51 & 56 & 62 & 135 & 135 & 186 & 177 \\ 167 & 148 & 66 & 64 & 61 & 76 & 56 & 194 & 200 & 185 \\ 152 & 83 & 50 & 56 & 64 & 135 & 135 & 182 & 197 & 195 \\ 199 & 191 & 161 & 169 & 90 & 96 & 184 & 69 & 173 & 161 \end{pmatrix}$$

Proses resize kolom dapat dilakukan melalui cara yang sama dengan resize baris, tetapi yang dicari adalah rata-rata antara dua piksel baris yang bertetangga dan mencari nilai terdekat.

Representasi resize citra dengan size 960 x 720 piksel dan 72 x 72 piksel tampak pada Gambar 4.4 dan 4.5.



Gambar 4. 4 data citra molting yang sebelum resize dan setelah resize



*Gambar 4. 5 data citra non-molting yang sebelum resize dan setelah
resize*

2. Transformasi Citra

Tahap selanjutnya adalah transformasi citra. Citra yang telah diresize akan dirubah warnanya menjadi grayscale. Hal ini dilakukan agar nilai piksel yang akan diambil tidak terlalu luas. Sehingga citra yang ditransformasi menjadi warna greyscale hanya memiliki rentang piksel dari 0 ke 1, 0 sebagai warna putih, 1 sebagai warna hitam, dan nilai diantaranya adalah warna abu. Setelah menjadi citra grayscale akan diambil nilai piksel dari warna tersebut.

Berikut adalah ilustrasi proses citra RGB to Grayscale yang digunakan pixel berukuran 10 x 10 x 3. citra tersebut kemudian diubah dari citra RGB menjadi citra grayscale dengan persamaan :

(143,155,182)	(144,156,183)	(147,159,186)	(73,90,80)	(67,85,72)	(140,152,176)	(152,161,187)	(152,161,184)	(156,164,184)	(158,164,183)
(144,156,182)	(146,158,185)	(140,152,174)	(54,74,50)	(44,67,41)	(106,121,137)	(153,163,184)	(159,167,188)	(163,169,189)	(167,171,189)
(146,159,185)	(150,162,188)	(94,110,104)	(68,87,75)	(55,75,53)	(54,74,59)	(119,135,152)	(171,182,201)	(174,184,202)	(170,174,192)
(149,162,189)	(141,154,174)	(80,99,85)	(47,68,40)	(54,75,40)	(81,101,105)	(69,89,75)	(153,168,184)	(183,193,209)	(169,173,191)
(152,164,189)	(106,121,121)	(49,69,44)	(84,102,100)	(58,77,48)	(68,88,73)	(52,77,40)	(135,149,168)	(190,198,212)	(176,182,198)
(152,164,185)	(137,149,159)	(63,81,54)	(58,77,57)	(62,83,43)	(70,94,60)	(128,141,144)	(186,192,206)	(192,198,210)	(184,191,202)
(140,153,165)	(77,95,79)	(63,81,59)	(53,70,46)	(58,79,37)	(128,142,135)	(182,188,192)	(173,182,191)	(188,195,208)	(188,193,205)
(126,136,150)	(80,96,73)	(48,67,36)	(54,68,47)	(63,79,52)	(124,139,119)	(78,99,63)	(66,88,62)	(83,102,83)	(145,154,175)
(182,187,199)	(187,190,197)	(158,166,161)	(80,98,71)	(37,60,25)	(91,112,86)	(90,114,78)	(64,88,57)	(87,108,96)	(143,156,185)
(193,197,208)	(192,197,207)	(196,199,207)	(165,175,167)	(87,109,74)	(91,114,92)	(174,186,193)	(144,157,163)	(157,168,195)	(141,154,184)

Berdasarkan persamaan tersebut maka diperoleh citra grayscale

sebagai berikut :

$$\text{Gray} = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$$

Untuk elemen pada baris pertama kolom matriks diperoleh matriks

Grayscale sebagai berikut :

$$\text{Grey} = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

$$(0.299 \times 143) + (0.587 \times 155) + (0.114 \times 182) = 154.49 = 154$$

Untuk elemen pada baris pertama kolom pertama matriks diperoleh

matriks Grayscale sebagai berikut :

$$\text{Grey} = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

$$(0.299 \times 143) + (0.587 \times 155) + (0.114 \times 182) = 154.49 = 154$$

Untuk elemen pada baris pertama kolom kedua matriks diperoleh matriks Grayscale sebagai berikut :

$$Grey = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

$$(0.229 \times 144) + (0.587 \times 156) + (0.114 \times 183) = 155.49 = 155$$

Untuk elemen pada baris pertama kolom ketiga matriks diperoleh matriks Grayscale sebagai berikut :

$$Grey = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

$$(0.229 \times 147) + (0.587 \times 159) + (0.114 \times 186) = 158.49 = 158$$

Untuk elemen pada baris kedua kolom pertama matriks diperoleh matriks Grayscale sebagai berikut :

$$Grey = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

$$(0.229 \times 144) + (0.587 \times 156) + (0.114 \times 182) = 155.376 = 155$$

Untuk elemen pada baris kedua kolom kedua matriks diperoleh matriks Grayscale sebagai berikut :

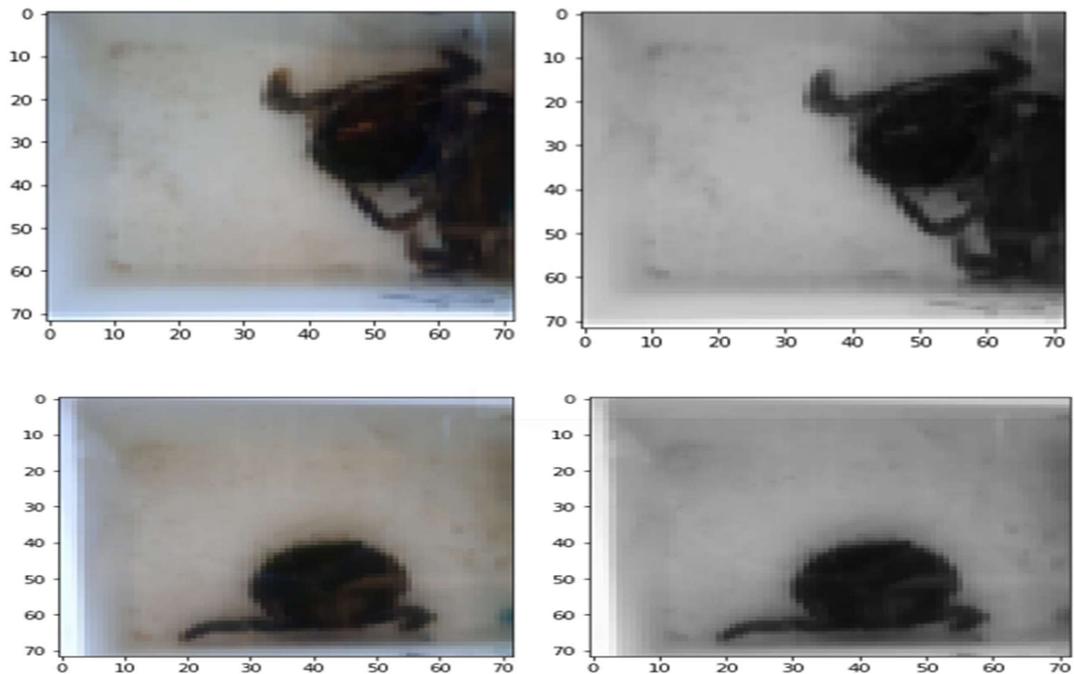
$$Grey = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

$$(0.229 \times 146) + (0.587 \times 158) + (0.114 \times 185) = 157.27 = 157$$

Untuk memperoleh elemen yang lain dari matriks *Gray* dihitung dengan menggunakan rumus yang sama. Elemen pada matriks *Gray* yang dihasilkan awalnya berupa bilangan pecahan, selanjutnya dilakukan proses pembulatan untuk mendapatkan bilangan bulat. Pembulatan ini

menggunakan fungsi *fix* yang prosesnya mengabaikan angka pecahan yang dimiliki meskipun angka tersebut mendekati nilai 1. Seluruh komponen *grayscale* dapat dilihat pada matriks *Gray* berikut ini :

$$Gray = \begin{pmatrix} 154 & 155 & 158 & 84 & 78 & 151 & 161 & 161 & 164 & 164 \\ 155 & 157 & 151 & 65 & 57 & 118 & 162 & 167 & 169 & 172 \\ 158 & 161 & 105 & 80 & 67 & 66 & 132 & 181 & 183 & 175 \\ 161 & 152 & 92 & 59 & 65 & 95 & 81 & 165 & 192 & 174 \\ 163 & 117 & 60 & 96 & 68 & 80 & 65 & 147 & 197 & 182 \\ 163 & 147 & 73 & 69 & 72 & 83 & 137 & 192 & 198 & 190 \\ 150 & 88 & 73 & 62 & 68 & 137 & 187 & 180 & 194 & 193 \\ 135 & 89 & 58 & 61 & 71 & 132 & 89 & 78 & 94 & 154 \\ 187 & 190 & 163 & 90 & 49 & 103 & 103 & 77 & 100 & 155 \\ 197 & 197 & 199 & 171 & 98 & 105 & 183 & 154 & 168 & 154 \end{pmatrix}$$



Gambar 4. 6 data citra molting dan non-molting yang belum ditransformasi dan yang telah ditransformasi ke greyscale

3. Ekstraksi fitur HOG

Pada tahapan ini dilakukan ekstraksi fitur HOG citra latih dan citra uji. Pada proses ini, sebelumnya telah dilakukan resize yang berukuran 72x72 pixel. Setelah melakukan resize data sebanyak 1060 citra, kemudian citra keping molting dan non molting pada data latih dan data uji diubah ke citra grayscale.

HOG merupakan bentuk dari local objek dan nilai yang digunakan dari intensitas gradien untuk mengekstraksi fitur yang digunakan pada computer vision dan image processing. HOG memiliki keunggulan, yaitu mampu menangkap tepi atau struktur gradient yang sangat karakteristik dari bentuk sebenarnya. Tiap image mempunyai karakteristik yang ditunjukkan oleh distribusi gradient. Karakteristik tersebut diperoleh dengan membagi image ke dalam daerah kecil yang disebut cell. Berikut algoritma dari algoritma HOG.



Gambar 4. 7 Algoritma Histogram of Gradient (HOG)

Dari Gambar 4.7, tahap awal dari metode HOG adalah menghitung nilai gradien citra dihitung menggunakan (4.1).

$$|G| = \sqrt{I_x^2 + I_y^2} \quad 4.1$$

Dimana I adalah citra greyscale. I_x merupakan matrik terhadap sumbu x dan I_y merupakan matrik terhadap sumbu y . I_x dan I_y dapat dihitung dengan (4.2).

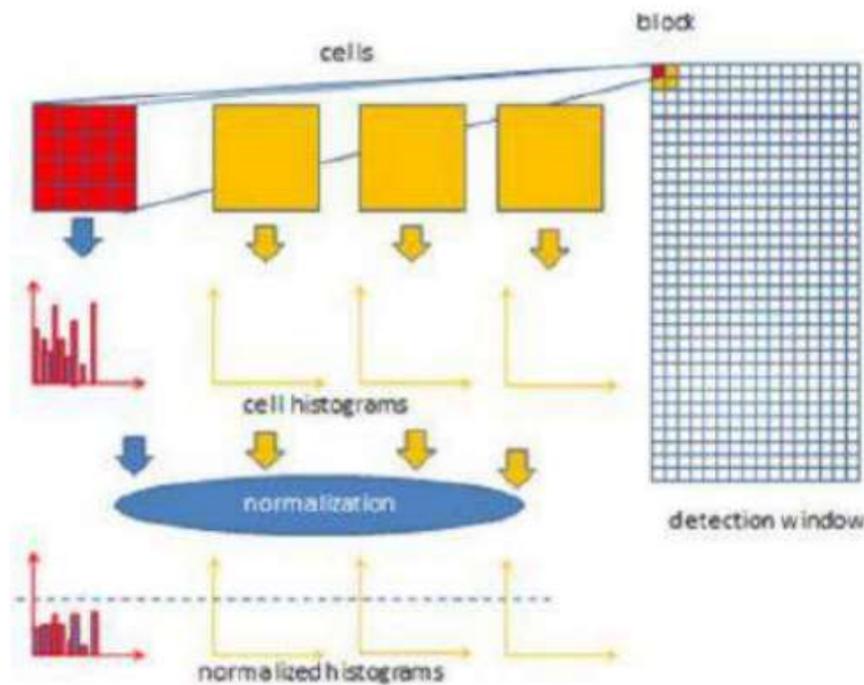
$$I_x = I * D_x, I_y = I * D_y \quad 4.2$$

D_x adalah mask $[-1 \ 0 \ 1]$, sedangkan D_y adalah mask $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ (4.2).

Masing-masing di hitung dengan cara konvolusi. Kemudian gradien di transformasi ke dalam koordinat sumbu dengan sumbu antara 0 sampai 1800 yang di sebut orientasi gradien. Orientasi gradien (θ) dapat di hitung dengan (4.3).

$$\theta = \arctan\left(\frac{I_x}{I_y}\right) \quad 4.3$$

Tahap selanjutnya adalah melakukan perhitungan histogram dari orientasi gradien tiap cell. Setiap piksel dalam sebuah cell mempunyai nilai histogram sendiri-sendiri berdasarkan nilai yang dihasilkan dalam perhitungan gradien yang kemudian dilakukan normalisasi pada setiap blok. Cell memiliki ukuran 8x8 piksel pada sebuah citra. Sedangkan blok memiliki ukuran 2x2 cell. Ilustrasi ditunjukkan pada Gambar 4.9.



Gambar 4. 8 Cell yang menyusun sebuah blok

Fitur blok dinormalisasi untuk mengurangi efek perubahan kecerahan obyek pada satu blok. Variabel b merupakan nilai blok fitur dan variabel e merupakan bilangan positif yang bernilai kecil untuk mencegah pembagian dengan 0.

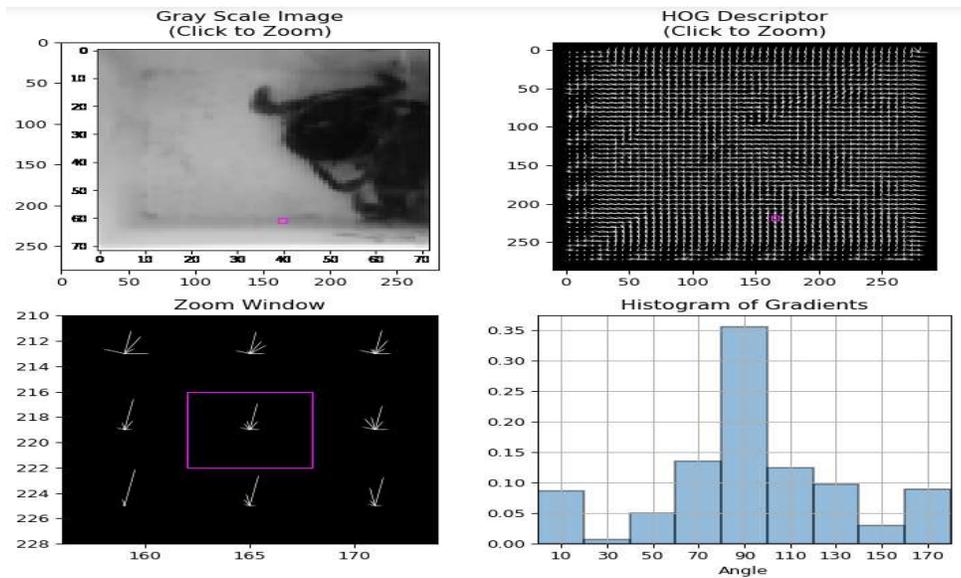
$$b = \frac{b}{\sqrt{b^2 + e}} \quad 4.4$$

Nilai normalisasi tiap blok digabungkan menjadi satu vektor menjadi fitur vektor HOG. Kemudian fitur vektor HOG dilakukan normalisasi. Normalisasi dilakukan melalui (4.5). Variabel h merupakan nilai fitur HOG dan variabel e merupakan bilangan positif yang bernilai kecil untuk mencegah pembagian dengan 0.

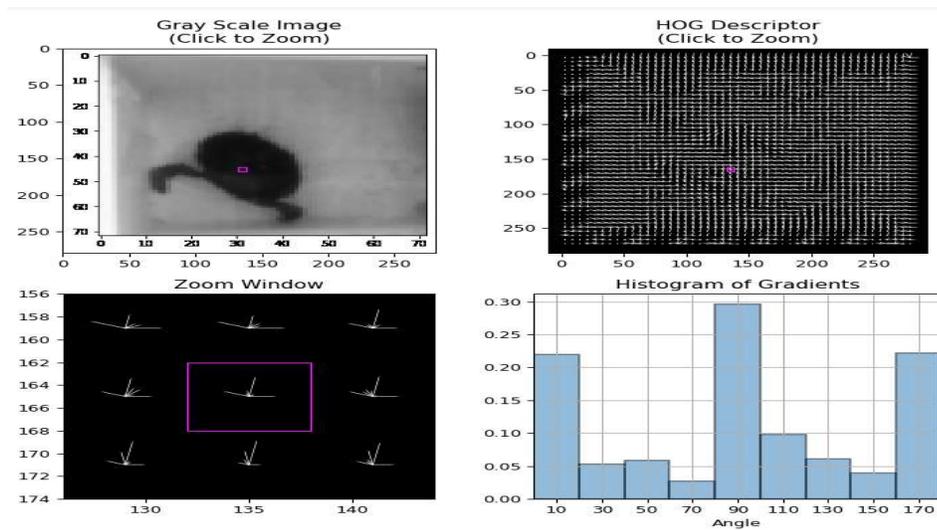
$$h = \frac{h}{\sqrt{\|h\|^2 + e}}$$

4. 5

Berikut hasil ekstraksi fitur HOG:



Gambar 4. 9 HOG pada Kepiting Molting



Gambar 4. 10 HOG pada Kepiting Non-Molting

C. Hasil Klasifikasi

Pada tahap ini merupakan hasil klasifikasi yang dilakukan dengan menggunakan nilai ekstraksi fitur yang telah dilakukan. Dalam penelitian ini, yang digunakan untuk mengklasifikasikan data yang berasal dari hasil ekstraksi ciri fitur yang dilakukan di Jupiter Notebook.

Data klasifikasi yang dilakukan menggunakan fitur bentuk HOG. Fitur mempunyai beberapa nilai ciri yang nantinya diolah dengan menggunakan metode klasifikasi KNN, SVM, RFC. Metode klasifikasi ini digunakan untuk menentukan dua kelas yang berbeda yaitu molting dan non molting.

Dari percobaan yang diuji nilai TP, TN, FP, dan FN merupakan hasil dari nilai penggabungan antara data pelatihan dan data pengujian. Tinggi nilai TP dan rendahnya nilai FN akan meningkatkan nilai sensitifitas, sedangkan tinggi nilai TN dan rendahnya nilai FP akan meningkatkan nilai spesifisitas. Pada performa klasifikasi morfologi mempunyai tingkat akurasi sensitifitas dan spesifisitas tinggi diantara semua performa yang telah diujikan. klasifikasi dengan menggunakan metode KNN memiliki tingkat akurasi yang tinggi dengan tingkat kesalahan yang rendah yang telah diuji dibandingkan dengan penggunaan metode SVM dan RFC.

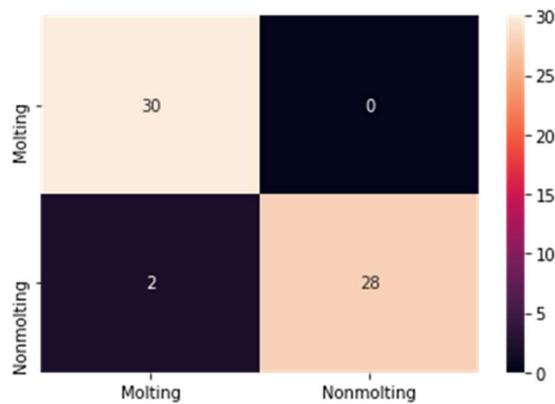
1. Performa Klasifikasi Algoritma KNN

Klasifikasi Algoritma KNN dengan tipe distance Euclidean menggunakan Jumlah Neighbors 3.

Tabel 4.1 menunjukkan hasil performa dari algoritma KNN pada tipe distance Euclidean. Matric adalah parameter penghitung jarak. Nilai yang di ambil adalah AVG atau rata-rata performa dari tiap kelas. Dapat dilihat pada Tabel 4.1 hasil klasifikasi dari metode KNN yang mempunyai tingkat nilai akurasi 96,66%. kemudian untuk nilai sensitifitas 100%. sedangkan tingkat spesifisitas pada metode KNN dengan nilai spesifisitas 96,66%. dengan klasifikasi KNN diketahui bahwa fitur bentuk mampu mendeteksi kepiting molting dideteksi molting (TP) 30 data dari 30 data molting dan untuk data kepiting dideteksi non molting (TN) 2 data dari 30 data non-molting.

Tabel 4. 1 Performa klasifikasi Algoritma KNN

Metode	Jml Data test	Matrix				AKURASI	RECALL	PRESISI	F1 SCORE
		TP	TN	FP	FN				
k-NN	60	30	28	0	2	96.6%	100%	100%	96.6%



Gambar 4. 11 Hasil confusion matrix k-NN

Berikut perhitungan hasil akurasi confusion matrix

$$\text{Akurasi} = \frac{58}{60} = 0,96 = 96\%$$

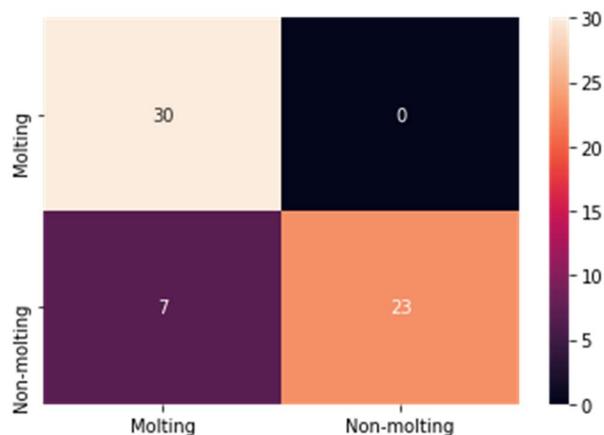
2. Performa Klasifikasi Algoritma SVM

Klasifikasi Algoritma SVM dengan menggunakan tipe kernel polynomial.

Tabel 4.2 menunjukkan hasil klasifikasi dari metode SVM yang mempunyai tingkat nilai akurasi 88.3%. kemudian untuk nilai sensitifitas 100%. sedangkan tingkat spesifisitas pada metode SVM dengan nilai spesifisitas 96,66%. dengan klasifikasi SVM diketahui bahwa fitur bentuk mampu mendeteksi kepiting molting dideteksi molting (TP) 30 data dari 30 data molting dan untuk data kepiting dideteksi non molting (TN) 23 data dari 30 data non-molting.

Tabel 4. 2 Performa klasifikasi Algoritma SVM

Metode	Jml Data test	Matrix				AKURASI	RECALL	PRESISI	F1 SCORE
		TP	TN	FP	FN				
SVM	60	30	23	0	7	88.3%	100%	100%	88.3%



Gambar 4. 12 Hasil Confusion Matrix SVM

Berikut perhitungan hasil akurasi confusion matrix

$$\text{Akurasi} = \frac{53}{60} = 0,88 = 88\%$$

3. Performa Klasifikasi Algoritma RFC

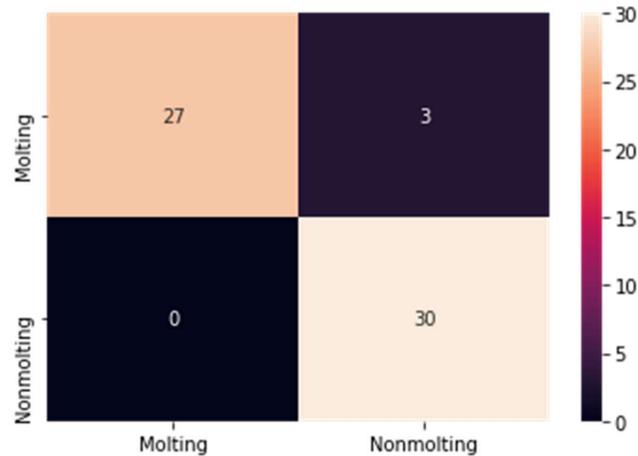
Klasifikasi Algoritma RFC dengan menggunakan tipe kernel estimator 30.

Tabel 4.3 menunjukkan hasil klasifikasi dari metode SVM yang mempunyai tingkat nilai akurasi 88.3%. kemudian untuk nilai sensitifitas 100%. sedangkan tingkat spesifisitas pada metode SVM dengan nilai spesifisitas 96,66%. dengan klasifikasi SVM diketahui bahwa fitur bentuk mampu mendeteksi kepiting molting dideteksi

molting (TP) 27 data dari 30 data molting dan untuk data kepiting dideteksi non molting (TN) 30 data dari 30 data non-molting.

Tabel 4. 3 Performa klasifikasi Algoritma RFC

Metode	Jml Data test	Matrix				AKURASI	RECALL	PRESISI	F1 SCORE
		TP	TN	FP	FN				
RFC	60	27	30	3	0	95%	90%	90%	95%



Gambar 4. 13 Hasil Confusion Matrix RFC

Berikut perhitungan hasil akurasi confusion matrix

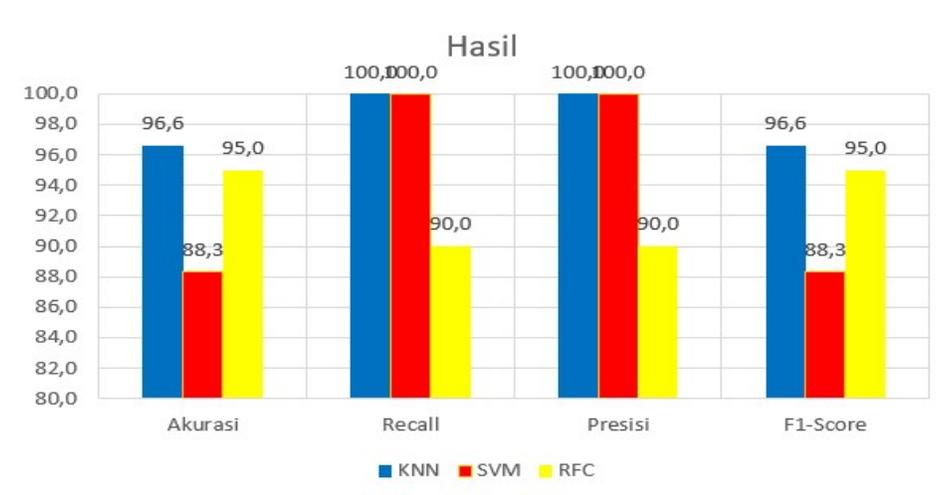
$$\text{Akurasi} = \frac{57}{60} = 0,95 = 95\%$$

D. Perbandingan Performa Metode Klasifikasi

Tabel 4.4 menunjukkan perbandingan performa antara algoritma KNN, SVM, dan RFC.

Tabel 4. 4 Perbandingan Performa Metode Klasifikasi

Metode	Jml Data test	Matrix				AKURASI	RECALL	PRESISI	F1 SCORE
		TP	TN	FP	FN				
k-NN	60	30	28	0	2	96%	100%	100%	96%
SVM	60	30	23	0	7	88%	100%	100%	88%
RFC	60	27	30	3	0	95%	90%	90%	95%



Gambar 4. 14 Hasil Perbandingan Metode Klasifikasi

Pada gambar diatas menunjukkan nilai akurasi tertinggi yaitu 96.66% dengan menggunakan algoritma KNN, dan selanjutnya akurasi untuk algoritma RFC yakni 95% dan SVM yakni 88,3%. Kemudian nilai presisi dan Recall tertinggi yakni 100% dengan menggunakan algoritma KNN dan SVM, dan selanjutnya nilai presisi untuk algoritma RFC 90%. Untuk nilai F1 Score tertinggi yakni 96,6% menggunakan algoritma KNN dan selanjutnya RFC dengan nilai 95%, lalu nilai 88,3% untuk SVM.

BAB V

PENUTUP

A. Kesimpulan

Kesimpulan dari penelitian ini yaitu :

Hasil eksperimen menunjukkan bahwa algoritma KNN memiliki kinerja yang lebih baik daripada SVM dan RFC dalam mengklasifikasikan kepiting molting dan non-molting. KNN memiliki akurasi 96,6% oleh karena itu KNN Classifier direkomendasikan untuk digunakan dalam identifikasi kepiting molting pada budidaya kepiting cangkang lunak.

B. Saran

Adapun saran yang dapat di sampaikan penulis terkait identifikasi kepiting molting menggunakan klasifikasi machine learning yaitu:

1. Penelitian mendatang yang mengembangkan penelitian ini dapat menggunakan system identifikasi kepiting molting menggunakan klasifikasi machine learning ini dapat digunakan untuk implementasi mengingat akurasi yang didapat sangat bagus.
2. Mengingat hanya beberapa algoritma klasifikasi yang digunakan, maka peneliti berharap selanjutnya yang ingin mengembangkan penelitian ini dapat mencoba algoritma klasifikasi lain.

DAFTAR PUSTAKA

- Ali, R., Yusro, M. M., Hitam, M. S., & Ikhwanuddin, M. (2021). Machine Learning With Multistage Classifiers For Identification Of Of Ectoparasite Infected Mud Crab Genus Scylla. *Telkomnika (Telecommunication Computing Electronics and Control)*, 19(2), 406–413. <https://doi.org/10.12928/TELKOMNIKA.v19i2.16724>
- Ario, R., Djunaedi, A., Pratikto, I., Subardjo, P., & Farida, F. (2019). Perbedaan Metode Mutilasi Terhadap Lama Waktu Molting Scylla serrata. *Buletin Oseanografi Marina*, 8(2), 103. <https://doi.org/10.14710/buloma.v8i2.24886>
- Ferdiansyah, F. F., Rahmat, B., & Yuniar, I. (2020). Klasifikasi Dan Pengenalan Objek Ikan Menggunakan Algoritma Support Vector Machine (Svm). *Jurnal Informatika dan Sistem Informasi (JIFoSI)*, 1(2), 522–528.
- Fouad, M. M. M., Zawbaa, H. M., El-Bendary, N., & Hassanien, A. E. (2014). Automatic Nile Tilapia fish classification approach using machine learning techniques. *13th International Conference on Hybrid Intelligent Systems, HIS 2013*, 173–178. <https://doi.org/10.1109/HIS.2013.6920477>
- Fujaya, Y. (2011). Pertumbuhan dan molting kepiting bakau yang diberi dosis vitomolot berbeda Growth and molting of mud crab

administered by different doses of vitomolt. *Journal of Indonesia Aquaculture*, 10(1), 24–28.

Quach, L.-D., Hoang, L. Q., Trung, N. D., & Nguyen, C. N. (2020). *Towards Machine Learning Approaches To Identify Shrimp Diseases Based on Description*. <https://doi.org/10.15625/vap.2019.00063>

Ramadhani, M., & Murti, D. H. (2018). Klasifikasi Ikan Menggunakan Oriented Fast and Rotated Brief (Orb) Dan K-Nearest Neighbor (Knn). *JUTI: Jurnal Ilmiah Teknologi Informasi*, 16(2), 115. <https://doi.org/10.12962/j24068535.v16i2.a711>

Yushinta Fujaya, Siti Aslamyah, Letty Fudjaja, N. A. (2019). *Budidaya dan Bisnis Kepiting Lunak: Stimulasi Molting Dengan Ekstrak Bayam* (Agus Wijaya (ed.)). Firstbox Media.