

Klasifikasi Ikan Cupang Menggunakan Support Vector Machine

Irsyad Rafi Diesta¹, Wikky Fawwaz Al Maki²

^{1,2}Universitas Telkom, Bandung

¹irsyadr@students.telkomuniversity.ac.id, ²wikkyfawwaz@telkomuniversity.ac.id,

Abstrak

Ikan cupang merupakan ikan hias yang umum dijumpai di wilayah Indonesia. Dibandingkan dengan ikan secara umum ikan cupang memiliki bentuk, ukuran, serta warna yang jauh lebih menarik. Minimnya riset terkait ikan laut maupun ikan hias dan sulitnya membedakan jenis ikan cupang yang unik. Dengan menerapkan ilmu pengolahan citra, visi komputer, dan pembelajaran mesin dapat dibuat model yang dapat mengidentifikasi jenis ikan cupang berdasarkan ciri-ciri yang dimiliki ikan cupang. Beberapa penelitian terkait ikan masih meneliti berdasarkan beberapa bagian ikan saja seperti kepala, kulit atau sisik, ekor atau sirip saja. Dalam Tugas Akhir (TA) ini ikan cupang diteliti berdasarkan bentuk sirip, ekor, bentuk tubuh, dan warna sisik yang beragam dan lebih kompleks dibanding ikan secara umum. Dataset yang digunakan sebanyak 3,082 citra ikan cupang yang telah diaugmentasi menjadi 12.328 yang diterapkan beberapa teknik pengolahan citra, lalu metode fitur ekstraksi PHOG dan Color Moments, dan diklasifikasi menggunakan SVM dengan beberapa jenis kernel. Dari beberapa model yang telah dibuat, dihasilkan akurasi paling besar 87% sedangkan akurasi terkecil 73%. Dengan model yang berhasil dibuat diharapkan penelitian terkait ikan cupang yang memiliki bentuk dan warna kompleks dapat menginspirasi para peneliti lain untuk meneliti terhadap ikan hias maupun ikan laut yang memiliki informasi yang lebih kompleks lagi.

Kata kunci : ikan cupang, color moments, PHOG, Klasifikasi SVM, pengolahan citra

Abstract

Betta fish are ornamental fish that are commonly found in Indonesia. Compared to fish in general, betta fish have a much more attractive shape, size, and color. The lack of research related to marine fish and ornamental fish and the difficulty of distinguishing the unique types of betta fish. By applying the science of image processing, computer vision, and machine learning, a model can be made that can identify the type of betta fish based on the characteristics of the betta fish. Some research related to fish is still researching based on only a few parts of the fish such as the head, skin or scales, tail or fins only. In this Final Project betta fish were studied based on the shape of the fins, tail, body shape, and color of the scales which are diverse and more complex than fish in general. The dataset used is 3,082 betta fish images that have been augmented to 12,328 which applied several image processing techniques, then the PHOG and Color Moments feature extraction methods, and classified using SVM with several types of kernels. From several models that have been made, the highest accuracy is 87% while the smallest accuracy is 73%. With the successful model, it is hoped that research related to ornamental betta fish that have complex shapes and colors can inspire other researchers to research ornamental fish and marine fish that have more complex information.

Keywords: betta fish, color moments, PHOG, SVM Classification, image processing

1. Pendahuluan

Latar Belakang

Ikan cupang merupakan ikan hias yang umum dijumpai di wilayah Indonesia. Secara umum ikan cupang memiliki kepala, badan, sirip, dan ekor sama seperti ikan lainnya. Jika dibandingkan dengan ikan secara umum, ikan cupang memiliki bentuk, ukuran, serta warna yang jauh lebih menarik. Ikan cupang memiliki ukuran tubuh yang kecil, warna dan corak yang bervariasi, serta bentuk ekor dan sirip yang beragam. Karakteristik ikan cupang ini juga membuat banyak orang tertarik untuk memelihara dan membudidayakan ikan cupang. Dengan berbagai informasi yang dimiliki, proses identifikasi jenis ikan cupang menjadi lebih kompleks. Bagi para pembudidaya ikan cupang untuk mengenali jenis ikan cupang dapat dilakukan secara mudah karena memiliki pengalaman di bidang ikan cupang. Bagi para pemula dan penghobi yang ingin mengenali berbagai jenis ikan cupang akan lebih sulit.

Dengan berkembangnya ilmu pengetahuan, berbagai model telah dibuat dengan tujuan identifikasi jenis tertentu. Beberapa penelitian pembuatan model untuk proses identifikasi telah dilakukan dengan menerapkan pembelajaran mesin dalam berbagai bidang. Pada bidang flora seperti yang dilakukan Pantazi [13] berhasil membuat

model deteksi penyakit pada daun menggunakan klasifikasi OCSVM. Ebrahimi juga dalam risetnya [6] berhasil membuat model deteksi hama berbasis vision menggunakan klasifikasi SVM. Pada bidang fauna Hu pada risetnya [9] membuat model klasifikasi spesies ikan berdasarkan warna dan tekstur kulitnya menggunakan multi-class SVM. Ogunlana juga pada risetnya [12] membuat model klasifikasi ikan menggunakan SVM. Pada bidang lain juga seperti kebudayaan yang dilakukan oleh Kumar pada riset [11] dan pada bidang kedokteran dilakukan oleh Hamiane pada risetnya [7] Dengan memanfaatkan ilmu visi komputer dan pembelajaran mesin dapat dibuat sebuah model yang dapat mengidentifikasi jenis ikan cupang dengan informasi yang kompleks.

Penelitian ini berfokus untuk membuat model yang dapat mengidentifikasi 6 jenis ikan cupang dengan menggunakan ilmu pengetahuan di bidang visi komputer dan pembelajaran mesin. Model yang telah dibuat menggunakan SVM dengan berbagai kernel ini dapat mengenali 6 jenis ikan cupang yang telah diklasifikasikan secara optimal. Model ini dibuat dengan memperhatikan karakteristik ikan cupang yang memiliki bentuk dan warna yang bervariasi.

Topik dan Batasannya

Dengan adanya batasan dalam tugas akhir ini, pelaksanaan riset akan lebih terfokus dan terarah. Pada penelitian ini proses identifikasi ikan cupang menerapkan pembelajaran mesin dengan metode klasifikasi SVM (*Support Vector Machine*). Data yang digunakan merupakan kumpulan citra ikan cupang dengan format JPG dengan rasio citra berukuran 1:1.

Jenis ikan cupang yang digunakan dalam dataset dibatasi menjadi enam jenis yaitu HMSD (*Halfmoon Solid*), HMMC (*Halfmoon Multicolor*), PKSD (*Plakat Solid*), PKMC (*Plakat Multicolor*), CTSD (*Crowntail Solid*), dan CTMC (*Crowntail Multicolor*). Model yang dibuat berfokus dalam mengambil informasi citra ikan cupang dari segi bentuk dan warna yang selanjutnya diklasifikasikan menggunakan SVM dan dievaluasi untuk mendapatkan hasil performa setiap model.

Tujuan

Tujuan utama dibuatnya tugas akhir ini adalah membangun suatu model dengan performa optimal yang dapat mengidentifikasi 6 jenis ikan cupang berdasarkan ciri bentuk dan warna.

2. Studi Terkait

Studi Literatur

Pada tahun 2011 Alsmadi dkk. dalam risetnya menerapkan metode pengolahan citra cropping dengan fitur ekstraksi yaitu color extraction, RGB histogram, dan histogram of co-occurring greyscale. Hasil dari penelitian pengenalan 20 jenis famili ikan berdasarkan warna menghasilkan akurasi dengan nilai rentang dari 81% hingga 92%. Akurasi tersebut dihasilkan dari metode neural network dengan back-propagation classifier (BPC) [1].

Pada tahun 2012 Hu dkk. dalam risetnya menerapkan pre-processing edge detection dan cropping dengan fitur ekstraksi yaitu hue saturation color, grayscale histogram, histogram of co-occurring greyscale dan wavelet based texture features. Pada riset ini dilakukan klasifikasi terhadap enam jenis ikan yang umum dijumpai di China yaitu grass crap, silver crap, big head crap, snake head murrel, Wuchang bream, dan red-bellied pacu dengan dataset berjumlah 540 citra (90 citra per jenis ikan). Hasil pada riset ini memiliki nilai akurasi yang sangat tinggi dengan menggunakan metode library for support vector machine (LIBSVM) yaitu 96.30% dengan proses pengolahan wavelet filter dan 94% dengan proses penolahan grayscale dan gray level co-occurrence metrics (GLCM) [9].

Pada tahun 2015 Ogunlana dkk. dalam risetnya menerapkan pre-processing principal component analysis dengan fitur ekstraksi yaitu optimal separating hyperplane dan margin separation berhasil mengklasifikasikan citra ikan yang terdiri dari 150 dataset (76 data train, 74 data test) yang menghasilkan akurasi tertinggi 78.59% menggunakan metode SVM. Riset ini juga menggunakan beberapa metode lain sebagai perbandingan terhadap metode SVM seperti KNN dengan akurasi 52.69%, ANN dengan akurasi 60.01%, dan KMeans dengan akurasi 50.97% [12].

Pada tahun 2017 Rathi dkk. dalam risetnya menerapkan pengolahan citra otsu thresholding, dan noise removal (dilation, erosion, subtraction). Penggunaan 7 metode CNN dengan nilai aktivasi yang berbeda yaitu ReLu, softmax, dan tanh menghasilkan nilai akurasi yang berbeda. Pada aktivasi ReLu nilai akurasi yang dihasilkan ialah 96.29%, sedangkan pada aktivasi softmax nilai akurasi yang dihasilkan ialah 61.91%, dan pada aktivasi tanh nilai akurasi yang dihasilkan ialah 72.62%. Dataset yang digunakan berjumlah 27,132 citra ikan yang dibagi menjadi 21 kelas [15].

Dari beberapa riset yang telah dikemukakan, perkembangan penelitian terhadap berbagai spesies ikan dilakukan hanya berdasarkan beberapa bagian fisik saja seperti bagian kulit atau sisik ikan, pola sisik pada ikan, maupun ukuran ikan yang menentukan sifat pada ikan [1, 9, 12, 15]. Dalam Tugas Akhir (TA) ini dilakukan penelitian terhadap ikan cupang berdasarkan bentuk keseluruhan ikan (sirip, ekor, bentuk tubuh, dan warna) dengan menerapkan ilmu pengolahan citra, visi komputer, dan pembelajaran mesin.

Dataset

Citra ikan cupang yang digunakan sebagai dataset diambil secara manual dan juga secara daring dengan kerjasama bersama pembudidaya maupun penghobi ikan cupang. Tiap citra terdiri dari satu jenis ikan cupang di dalam wadah yang diambil dari samping. Dataset yang digunakan terdiri dari enam jenis ikan cupang yaitu HMSD, HMMC, PKSD, PKMC, CTSD, dan CTMC. Untuk gambar jenis ikan cupang yang digunakan sebagai dataset dapat dilihat pada Gambar 1.



Gambar 1. Dataset Original (PKMC, PKSD, HMMC, HMSD, CTMC, CTSD)

Segmentasi U²-Net

Segmentasi memiliki manfaat untuk simplifikasi citra dengan tujuan utama yaitu membagi citra menjadi beberapa bagian, sehingga memudahkan citra untuk dianalisis. Pada Tugas Akhir (TA) ini informasi pada citra yang akan digunakan merupakan objek ikan yang berada pada foreground. Informasi yang terkandung pada background citra dapat menyebabkan pengolahan informasi tidak maksimal, sehingga background pada citra dapat dipisah ataupun dihilangkan.

Segmentasi memiliki berbagai metode untuk menghilangkan background citra dengan hasil yang berbagai macam. Pada jurnal yang dibuat oleh Xuebin segmentasi U²-Net memiliki kelebihan dalam mencari informasi kontekstual berdasarkan perbedaan skala yang diterapkan menggunakan arsitektur deep network yaitu two-level nested U-structure [14]. U²-Net dalam segmentasinya menerapkan deep learning untuk menemukan objek salient pada citra yang kemudian akan dibuat citra masking sebagai dasar pemisah segmentasi pada citra. Untuk hasil gambar segmentasi dari data original dapat dilihat pada Gambar 2.



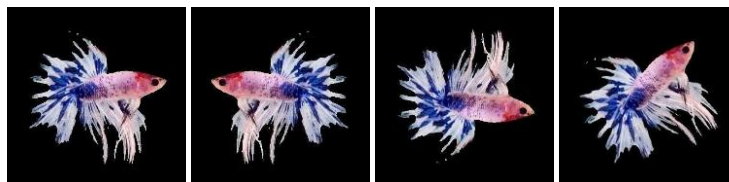
Gambar 2. Dataset Segmentasi (PKMC, PKSD, HMMC, HMSD, CTMC, CTSD)

Data Augmentasi

Dengan data yang terbatas, informasi citra dapat diambil secara maksimal dengan memperbanyak variasi data. Penerapan augmentasi dilakukan pada dataset yang telah disegmentasi sebelumnya dengan mengubah posisi citra. Komputer akan melihat citra yang telah diaugmentasi sebagai suatu matriks yang baru, sehingga informasi yang didapat akan bertambah berdasarkan metode augmentasi yang telah diterapkan. Meskipun informasi yang didapat menggunakan augmentasi lebih banyak, namun penerapan augmentasi juga dibatasi sehingga informasi yang diambil tidak jenuh. Dalam implementasinya tiga metode augmentasi telah diterapkan untuk memperkaya dataset citra yaitu horizontal flip, vertical flip, dan rotasi 90°. Untuk gambar hasil augmentasi dapat dilihat pada Gambar 3.

Pengolahan Citra

Kualitas citra berhubungan erat dengan informasi yang dimiliki suatu citra. Informasi citra salah satunya dipengaruhi oleh kualitas citra. Image preprocessing sendiri memiliki peran dalam meningkatkan kualitas citra. Dengan meningkatkan kualitas citra menggunakan metode yang terdapat pada image preprocessing, informasi



Gambar 3. Dataset Augmentasi (Tanpa Augmentasi, Horizontal Flip, Vertical Flip, Rotasi 90°)

citra yang didapat akan lebih detail. Penerapan preprocessing seperti yang dilakukan Hernandez dan Jimenez pada risetnya telah memberikan hasil yang bagus pada citra yang telah diolah sehingga model yang dibuat dapat mengenali dan membedakan jenis ikan, tumbuhan dan kupu-kupu [8]. Dalam Tugas Akhir (TA) beberapa metode pengolahan citra telah diterapkan untuk meningkatkan kualitas citra dari dataset original dan dataset augmentasi.

Resize/Image Scaling

Ukuran citra yang berbeda-beda akan mempengaruhi beban dan waktu yang dijalankan pada mesin. Tiap citra pada dataset diubah menjadi ukuran piksel 256x256. Pengubah ukuran citra digunakan untuk menstandarisasi tiap citra pada dataset, selain itu perubahan ukuran citra digunakan agar kompatibel terhadap metode ekstraksi fitur yang digunakan.

Image Sharpening

Peningkatan kualitas citra juga dilakukan dengan image sharpening untuk meningkatkan detail tekstur dan warna pada gambar. Dalam implementasinya citra dikonvolusi menggunakan filter kernel buatan. Dengan menerapkan image sharpening detail informasi dari bentuk dan warna pada citra dapat diambil secara maksimal.

Color Space Conversion

Setiap color space dalam citra memiliki informasi tersendiri yang dapat digunakan oleh mesin. Dengan mengubah color space citra menjadi beberapa color space lainnya, informasi yang didapat akan lebih banyak. Seperti yang dilakukan Cernadas dalam risetnya dilakukan klasifikasi berdasarkan color space dan menerapkan lima color space yang berbeda dan hasil akurasi yang berbeda pada tiap color space [3]. Pada eksperimen telah diterapkan tiga color space yang berbeda secara bersamaan dari tiap citra yaitu RGB, HSV, dan CMYK untuk mendapat informasi yang lebih detail.

Color Moments

Pada tahun 2005 Noah menjelaskan color moments sebagai salah satu cara mengukur yang dapat diterapkan pada citra berdasarkan fitur warna. Pengukuran ini dapat digunakan sebagai perhitungan kesamaan antar citra berdasarkan nilai yang telah diekstrak menggunakan perhitungan matematis yang terdapat pada color moments [10]. Penggunaan color moment telah diterapkan untuk ekstraksi fitur citra berbasis warna. Nilai color moments yang digunakan menggunakan nilai dari mean, standard deviation, dan skewness pada kumpulan piksel yang terdapat pada matriks citra. Perhitungan color moments dikalkulasikan pada tiap color space yang sebelumnya telah dibuat pada image preprocessing pada citra, sehingga didapatkan karakteristik informasi citra yang lebih spesifik.

Dalam penerapannya color moment mengimplementasikan rumus matematika yang ditampilkan pada rumus 1 moment mean, rumus 2 moment standar deviasi, dan rumus 3 untuk moment skewnes. Penggunaan color moment untuk ekstraksi citra berdasarkan basis warna dikalkulasikan pada tiap chanel warna pada tiap color space citra sehingga terciptanya karakteristik citra yang bervariasi. Dalam implementasi color moment diterapkan pada tiap chanel warna ke-i pada piksel citra ke-j sebagai P_{ij} yang ditunjukkan pada persamaan 1, 2, dan 3.

$$Mean = E = \frac{1}{N} \sum_{j=1}^N P_{ij} \tag{1}$$

$$StandardDeviation = \sigma_i = \sqrt{\frac{1}{N} \sum_{j=1}^N (P_{ij} - E_i)^2} \tag{2}$$

$$Skewness = S = \frac{1}{N} \sum_{i=1}^N (E_i)^3 \tag{3}$$

PHOG (Pyramid Histogram of Oriented Gradients)

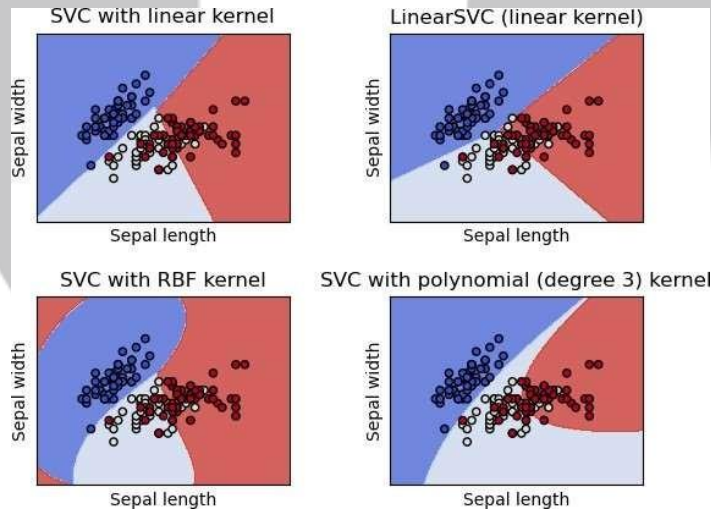
PHOG merupakan varian dari HOG yang diperkenalkan oleh Anna Bosch pada tahun 2007, ekstraksi fitur PHOG dikembangkan berdasarkan HOG dengan membagi citra menjadi subregion dalam tiap level skala piksel [16, 2]. Sedangkan HOG sudah banyak diterapkan sejak 2005 seperti yang diawali oleh Dalal dan Triggs pada [5] dan berkembang variannya menjadi PHOG yang sudah diterapkan pada beberapa eksperimen yang lain. Seperti Zhang pada [17] berhasil menerapkan PHOG dan secara efektif dapat mengenali penyakit pada ketimun.

PHOG yang diterapkan bertujuan untuk mendapatkan informasi bentuk gambar dalam citra berdasarkan local shape dan spatial layout kemudian disusun menjadi beberapa tingkatan resolusi. Dari kumpulan tingkatan atau level resolusi tersebut dideskripsikan menjadi histogram gradient berbentuk piramid. Penggunaan PHOG dengan level yang berbeda digunakan pada eksperimen ini, dimulai dari level 1, level 2, dan level 3. Penerapan PHOG dalam eksperimen ini menggunakan citra dataset dalam bentuk grayscale sehingga berfokus untuk mencari informasi berdasarkan bentuk yang dimiliki pada citra.

Klasifikasi SVM

Metode pembelajaran mesin SVM diterapkan untuk mengklasifikasikan enam jenis ikan cupang yang telah dilabeli sebelumnya. SVM yang diterapkan adalah SVM Multiclass dengan model One Vs One, dimana SVM akan membandingkan tiap antar label untuk mencari hyper-plane yang optimal yang dapat memisahkan kluster label pada vektor. Seperti Ebrahimi yang menerapkan SVM karena performa yang baik dalam klasifikasi dengan berbagai jenis data, seperti text hingga data genomic bahkan dalam bentuk data yang kompleks seperti graph, sequences, dan data relational dengan menggunakan kernel tertentu [6]. Selain itu beberapa peneliti lain juga menerapkan SVM untuk proses klasifikasi dan menghasilkan performa yang baik seperti pada [7, 9, 11, 12, 13].

Pada eksperimen ini tiga jenis kernel SVM diterapkan untuk mencari performansi terbaik untuk dataset ikan cupang. Kernel yang digunakan yaitu Linear, Polynomial, dan RBF. Selain itu parameter yang digunakan merupakan hasil dari tuning hyper-parameter GridSearchCV untuk memaksimalkan hasil performansi model SVM yang telah dibuat. Contoh gambar pembagian hyperplane pada tiap kernel dapat dilihat pada Gambar 4.



Gambar 4. Contoh multiclass SVM kernel hyper-plane (sumber: <https://scikit-learn.org/stable/modules/svm.html>)

Evaluasi performa Model

Metode evaluasi yang digunakan adalah tabel matriks konfusi untuk mengukur performansi model dari segi akurasi, presisi, recall, dan f1-score. Pada [4] meski MCC lebih baik digunakan untuk pengukuran evaluasi, namun perhitungan akurasi, presisi, recall, dan f1-score masih umum digunakan dan perhitungan matematis lebih

mudah diterapkan dan juga cocok pada label multiclass. Sehingga metode evaluasi ini digunakan karena model yang dibuat merupakan hasil dari metode klasifikasi SVM. Rumus yang diimplementasi pada metode evaluasi ditampilkan pada persamaan 4, 5, 6, dan 7.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

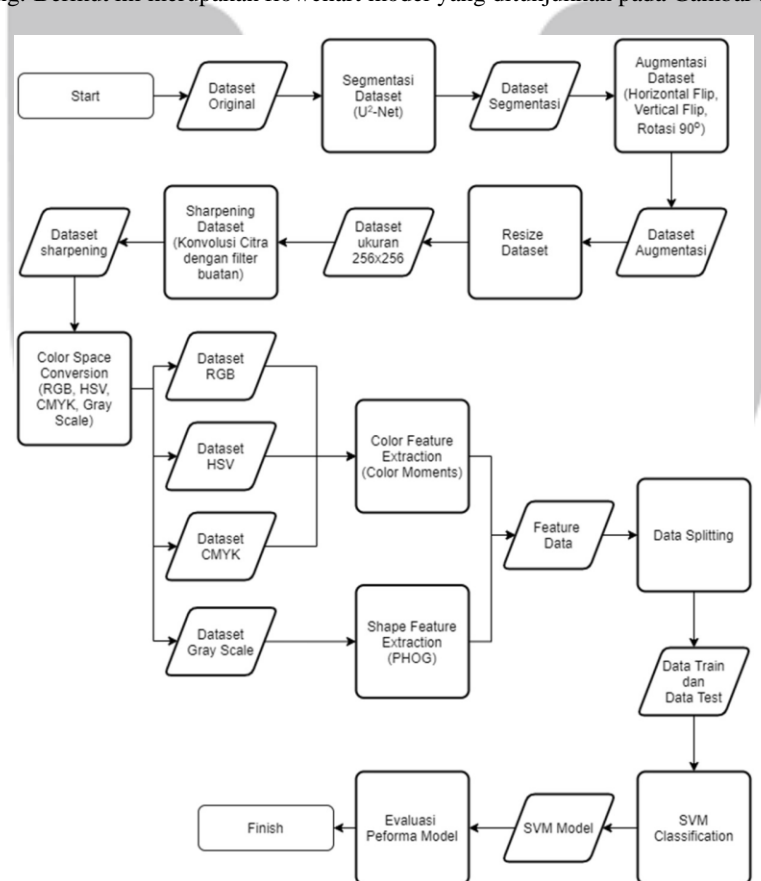
$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \tag{7}$$

3. Sistem yang Dibangun

Sistem yang dibuat dapat mengklasifikasi jenis ikan cupang dengan mempertimbangkan keseluruhan bentuk dan warna ikan cupang dengan menerapkan metode Support Vector Machine dengan menghitung ciri yang dihasilkan oleh ekstraksi fitur PHOG dan color moments. Sebelum fitur ciri diambil, citra diolah terlebih dahulu pada tahap preprocessing. Berikut ini merupakan flowchart model yang ditunjukkan pada Gambar 5.



Gambar 5. Flowchart model

Model yang dibangun menggunakan 3082 citra ikan cupang yang kemudian diolah pada tahap preprocessing.

Tahap preprocessing data merupakan proses yang dilakukan untuk mempersiapkan data sebelum memasuki tahap ekstraksi fitur. Pada tahap preprocessing, data diolah sehingga informasi yang dimiliki data lebih maksimal dan menghilangkan informasi yang tidak digunakan. Tahap preprocessing mencakup dari segmentasi sampai konversi ruang warna yang ditunjukkan pada flowchart Gambar 5.

Setelah melalui tahap preprocessing, data yang telah diolah diekstrak informasinya pada tahap ekstraksi fitur. Tahap ekstraksi fitur mengambil informasi yang ada pada citra berdasarkan informasi bentuk dan warna. Ekstraksi fitur berdasarkan warna menggunakan metode color moment pada tiap ruang warna yang telah dikonversi sebelumnya. Ekstraksi fitur berdasarkan bentuk menggunakan metode PHOG pada level 1, 2, dan 3 pada warna ruang grayscale. Pada tahap ini nilai vektor ciri didapat dari tiap ekstraksi fitur yang akan digunakan sebagai data train dan data test pada klasifikasi SVM. Tahap ekstraksi fitur pada flowchart ditunjukkan pada bagian color feature extraction dan shape feature extraction pada Gambar 5.

Setelah didapatkan fitur data berbentuk vektor ciri dari tahap ekstraksi fitur, data dipisah menjadi data test dan data train dengan rasio 80% : 20%. Data yang telah dipisah digunakan untuk membuat beberapa model dengan klasifikasi multiclass SVM dengan kernel yang berbeda. Model SVM yang telah mengklasifikasi enam jenis ikan cupang dengan kernel yang berbeda kemudian diuji performa model dengan menerapkan matriks konfusi. Matriks konfusi memberikan hasil performa akurasi, presisi, recall, dan f1-score. Tahap ini ditunjukkan pada data splitting sampai evaluasi performa model pada flowchart Gambar 5.

4. Evaluasi

4.1 Hasil Pengujian

Dataset yang digunakan dan hasil pengolahan citra yang telah diterapkan ditampilkan pada gambar 1-3. Pada Gambar 1 terdapat contoh data original dari enam jenis citra ikan cupang. Dari 3,082 citra original ikan cupang yang digunakan terlebih dahulu diolah menggunakan segmentasi U²-Net untuk menghilangkan background yang ada pada citra. Hasil segmentasi dapat dilihat pada Gambar 2. Setelah dilakukan segmentasi dilakukan kembali augmentasi untuk memperbanyak dataset dengan melakukan horizontal flip, vertical flip, dan rotasi 90°. Hasil augmentasi ditampilkan pada Gambar 3. Total dataset citra ikan cupang yang digunakan berjumlah 12,328 yang berasal dari gabungan data original dengan data yang telah diaugmentasi.

Model yang telah dibuat untuk identifikasi jenis ikan cupang dievaluasi berdasarkan nilai akurasi, presisi, recall, dan f1-score. Semakin besar nilai yang dihasilkan performa model yang diberikan akan semakin baik. Performa model yang diberikan juga dipengaruhi dari parameter yang digunakan pada tiap metode, detail parameter dari tiap model yang dibuat ditampilkan pada Tabel 1.

Tabel 1. Performa model SVM dengan segmentasi U²-Net

Model No.	Color Moments Color Space	PHOG Level	Parameter SVM			Perform Model			
			Kernel	C	gamma	Akurasi	Presisi	Recall	F1-Score
1	RGB, HSV, CMYK	1	Linear	0.1	1	73%	73%	73%	73%
2	RGB, HSV, CMYK	1	Polynomial	1000	0.0001	79%	79%	79%	79%
3	RGB, HSV, CMYK	1	RBF	1000	0.0001	86%	86%	86%	86%
4	RGB, HSV, CMYK	2	Linear	0.1	1	74%	74%	74%	74%
5	RGB, HSV, CMYK	2	Polynomial	1000	0.0001	80%	81%	80%	81%
6	RGB, HSV, CMYK	2	RBF	1000	0.0001	86%	87%	87%	87%
7	RGB, HSV, CMYK	3	Linear	0.1	1	74%	74%	74%	74%
8	RGB, HSV, CMYK	3	Polynomial	1000	0.0001	81%	80%	80%	80%
9	RGB, HSV, CMYK	3	RBF	1000	0.0001	87%	87%	87%	87%

Dari hasil eksperimen yang telah dilakukan dan ditampilkan pada Tabel 1. Penggunaan color space yang berbeda diterapkan secara bersamaan untuk mendapatkan informasi yang lebih detail dan eksperimen dilakukan dengan menerapkan kernel yang berbeda pada klasifikasi SVM. Performansi model paling optimal dihasilkan pada model ke-9 yang menerapkan Kernel RBF dengan rata-rata performa sebesar 87%.

4.2 Analisis Hasil Pengujian

Penggunaan ekstraksi fitur Color Moments yang digabung dengan PHOG mampu menghasilkan model dengan performa optimal di 87% dengan menerapkan klasifikasi SVM. Ekstraksi fitur PHOG yang digunakan adalah level 1, level 2, dan level 3 pada ruang warna abu-abu (*grayscale*). Level pada PHOG ini menunjukkan berapa banyak resolusi yang dibagi oleh PHOG dalam citra. Seiring meningkatnya level PHOG detail resolusi akan semakin detail. Parameter yang diterapkan pada PHOG juga mendukung performa yang dihasilkan. Nilai orientasi yang

bernilai 11 memberikan informasi nilai gradient yang lebih spesifik. Nilai orientasi berfungsi untuk menampung nilai gradient yang telah dihitung kedalam kumpulan bin arah. Berbagai nilai cell per block dalam mencari bentuk gradient juga diterapkan pada PHOG untuk mendapatkan perspektif informasi bentuk yang lebih banyak. Hasil dari proses PHOG dengan menerapkan parameter ini kemudian menghasilkan informasi berbentuk gradient dalam berbagai resolusi. Seiring meningkatnya level PHOG performa yang dihasilkan seharusnya juga meningkat. Secara umum performansi level PHOG pada setiap kernel mengalami peningkatan namun tidak signifikan. Hal ini dapat terjadi karena informasi yang didapat pada PHOG level yang tinggi tidak terlalu berbeda dengan level sebelumnya.

Penerapan Color Moments di ruang warna RGB, HSV, dan CMYK secara bersamaan bertujuan untuk mendapatkan informasi berbasis warna pada citra yang lebih detail. Perbandingan performa ruang warna telah dilakukan terhadap model ke-9 dengan hasil yang ditunjukkan pada Tabel 2.

Tabel 2. Perbandingan performa ruang warna pada model ke-9

Model ke-9 versi	Color Space	Perform Model			
		Akurasi	Presisi	Recall	F1-Score
1	RGB, HSV, CMYK	87%	87%	87%	87%
2	RGB	76%	76%	76%	76%
3	HSV, CMYK	81%	81%	81%	81%

Pengaruh Color Moments pada model yang menerapkan ruang warna RGB dan model yang menerapkan di ruang warna HSV dan CMYK memiliki performa yang kurang optimal dibandingkan model yang menerapkan ruang warna RGB, HSV, dan CMYK secara bersamaan. Meskipun nilai HSV dan CMYK didapat berdasarkan turunan nilai RGB, namun informasi yang diberikan tetaplah berbeda. Informasi warna ini sangat mempengaruhi kualitas performa model dikarenakan dataset ikan cupang memiliki salah satu ciri keunikan pada warnanya. Hal ini diperkuat dengan adanya perbandingan performa ekstraksi fitur terhadap dataset ikan cupang yang ditunjukkan pada Tabel 3.

Tabel 3. Perbandingan performa ekstraksi fitur pada model ke-9

Model ke-9 versi	Ekstraksi Fitur	Perform Model			
		Akurasi	Presisi	Recall	F1-Score
1	PHOG Level 3	65%	65%	65%	65%
2	Color Moments	78%	78%	78%	78%

Dari model ke-9 yang hanya menerapkan PHOG level 3 memberikan hasil performa rata-rata di 65%, sedangkan model ke-9 yang hanya menerapkan Color Moments memberikan hasil performa rata-rata di 78%. Dengan demikian informasi warna yang dimiliki oleh ikan cupang sangat berpengaruh berdasarkan data yang ditampilkan pada Tabel 2 dan Tabel 3.

Dari ketiga kernel yang diuji, performa model yang menerapkan kernel linear memiliki performa lebih kecil dibandingkan model yang menerapkan kernel polynomial maupun RBF. Model yang menerapkan kernel RBF memiliki performa yang paling optimal. Performa yang diberikan dari tiap kernel berbeda bergantung dari bentuk hyperplane. Bentuk hyperplane RBF yang ditunjukkan pada Gambar 4. lebih kompleks dibandingkan bentuk linear. Dari hasil performa model yang paling optimal menerapkan kernel RBF, kompleksitas bentuk hyperplane pada RBF cocok dalam membagi dataset ikan cupang yang memiliki informasi yang kompleks juga. Berdasarkan hasil performa model ke-9 kompatibilitas dari metode dan dataset yang digunakan sangat mempengaruhi performa yang dihasilkan.

Selain penerapan kernel yang berbeda pada tiap model, nilai parameter koefisien dan gamma yang berbeda diterapkan juga pada tiap model SVM. Nilai koefisien dan gamma yang digunakan merupakan hasil hypertuning parameter SVM, dimana nilai tersebut dicari oleh SVM secara *brute force* dengan kecocokan terhadap kernel yang digunakan. Nilai koefisien yang tinggi pada kernel polinomial dan RBF menunjukkan bahwa peranan hyperplane dalam kontrol error label sangat rendah, namun sebaliknya pada kernel linear hyperplane dalam pembagian klasifikasi memiliki pengaruh yang lebih besar. Sebanding dengan nilai gamma yang kecil pada kernel polinomial dan RBF peranan dari tiap titik label pada proses klasifikasi saling mempengaruhi label, sedangkan tidak pada kernel linear dengan nilai gamma yang lebih besar.

Setiap model dilihat performanya berdasarkan nilai akurasi, presisi, recall, dan f1-score yang dihasilkan. Penggunaan nilai f1-score yang didapat dari presisi dan recall berguna untuk memberikan persepsi performa yang lebih detail dibanding hanya dengan menggunakan nilai akurasi. Mayoritas model menghasilkan performa dengan rata-rata nilai yang sama. Setiap model berhasil secara akurat mengklasifikasikan dengan benar berdasarkan performa akurasi yang dihasilkan. Dari hasil presisi setiap model berhasil memberikan informasi yang diminta saat proses testing. Hasil recall setiap model juga berhasil menemukan kembali sebuah informasi saat proses evaluasi. Dengan

seimbang nilai presisi dan recall pada setiap model tersebut, nilai f1-score yang dihasilkan juga seimbang.

Dari model yang telah dibuat, model ke-9 menghasilkan performa paling optimal dibandingkan model yang lain. Model ke-9 menerapkan kernel RBF yang dimana bentuk hyperplane paling kompleks dibandingkan kernel polinomial dan linear. Model ke-9 juga menerapkan PHOG level 3 yang mengambil informasi bentuk citra secara lebih detail berdasarkan level resolusi yang lebih banyak. Dari model ke-9 ini menunjukkan bahwa kecocokan dari setiap metode yang digunakan terhadap dataset ikan cupang memiliki pengaruh yang sangat signifikan. Informasi yang kompleks berdasarkan keunikan ikan cupang dapat diolah menggunakan metode yang mengekstrak informasi secara mendetail dan maksimal untuk mendapatkan model yang optimal. Meski model ke-9 memiliki performa paling optimal di 87%, masih terdapat performa yang hilang sebesar 13%. Berdasarkan Tabel 3 performa yang diberikan oleh ekstraksi fitur PHOG masih belum maksimal. PHOG yang menerapkan parameter yang sama pada tiap levelnya dapat menyebabkan tidak kompatibel antara parameter yang digunakan dengan dataset citra. Hilangnya performa ini juga karena dataset menerapkan augmentasi rotasi 90° dan penerapan ekstraksi fitur PHOG sendiri menghitung arah gradient dan orientasi, sehingga informasi yang diolah menjadi redundan.

5. Kesimpulan

Multiclass SVM dapat memberikan hasil performa yang cukup tinggi dalam klasifikasi terhadap dataset ikan cupang. Kompabilitas antara metode yang digunakan dengan bentuk dataset ikan cupang sangat mempengaruhi performa model. Pembuatan model untuk identifikasi jenis ikan cupang dapat dilakukan dengan data yang cukup. Beberapa model telah dibuat dan menghasilkan performa paling kecil 73%. Model yang dibuat dengan menerapkan SVM pada kernel RBF dengan ekstraksi fitur Color Moments dan PHOG pada ruang warna yang beragam mampu menghasilkan performa yang optimal di 87%. Meski begitu masih ada ruang pengembangan dengan memperkaya color space yang berbeda ataupun penambahan nilai color moment yang lainnya. Penambahan fitur ekstraksi berbasis bentuk yang berbeda juga dapat dilakukan untuk mencari komparabilitas yang lebih terhadap dataset ikan cupang. Penggunaan metode preprocessing yang lain juga bisa diterapkan untuk mencari kecocokan terhadap metode ekstraksi fitur yang digunakan sehingga performa yang diberikan lebih baik lagi. Dengan model yang berhasil dibuat diharapkan penelitian terkait ikan cupang yang memiliki bentuk dan warna kompleks dapat menginspirasi para peneliti lain untuk meneliti terhadap ikan hias maupun ikan laut yang memiliki informasi yang lebih kompleks lagi.

REFERENSI

- [1] M. K. Alsmadi, K. B. Omar, S. A. Noah, et al. Fish classification based on robust features extraction from color signature using back-propagation classifier. *Journal of Computer Science*, 7(1):52, 2011.
- [2] A. Bosch, A. Zisserman, and X. Munoz. Representing shape with a spatial pyramid kernel. In *Proceedings of the 6th ACM international conference on Image and video retrieval*, pages 401–408, 2007.
- [3] E. Cernadas, M. Fernández-Delgado, E. González-Rufino, and P. Carrión. Influence of normalization and color space to color texture classification. *Pattern Recognition*, 61:120–138, 2017.
- [4] D. Chicco and G. Jurman. The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. *BMC genomics*, 21(1):1–13, 2020.
- [5] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 886–893. Ieee, 2005.
- [6] M. Ebrahimi, M.-H. Khoshtaghaza, S. Minaei, and B. Jamshidi. Vision-based pest detection based on svm classification method. *Computers and Electronics in Agriculture*, 137:52–58, 2017.
- [7] M. Hamiane and F. Saeed. Svm classification of mri brain images for computer-assisted diagnosis. *International Journal of Electrical and Computer Engineering*, 7(5):2555, 2017.
- [8] A. Hernández-Serna and L. F. Jiménez-Segura. Automatic identification of species with neural networks. *PeerJ*, 2:e563, 2014.
- [9] J. Hu, D. Li, Q. Duan, Y. Han, G. Chen, and X. Si. Fish species classification by color, texture and multi-class support vector machine using computer vision. *Computers and electronics in agriculture*, 88:133–140, 2012.
- [10] N. Keen. Color moments. *School of informatics, University of Edinburgh*, pages 3–6, 2005.

- [11] K. Kumar and P. Kishore. Indian classical dance mudra classification using hog features and svm classifier. *International Journal of Electrical & Computer Engineering (2088-8708)*, 7(5), 2017.
- [12] S. Ogunlana, O. Olabode, S. Oluwadare, and G. Iwasokun. Fish classification using support vector machine. *African Journal of Computing & ICT*, 8(2):75–82, 2015.
- [13] X. E. Pantazi, D. Moshou, and A. A. Tamouridou. Automated leaf disease detection in different crop species through image features analysis and one class classifiers. *Computers and electronics in agriculture*, 156:96–104, 2019.
- [14] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand. U2-net: Going deeper with nested u-structure for salient object detection. *Pattern Recognition*, 106:107404, 2020.
- [15] D. Rathi, S. Jain, and S. Indu. Underwater fish species classification using convolutional neural network and deep learning. In *2017 Ninth international conference on advances in pattern recognition (ICAPR)*, pages 1–6. IEEE, 2017.
- [16] V. K. Vishnoi, K. Kumar, and B. Kumar. Plant disease detection using computational intelligence and image processing. *Journal of Plant Diseases and Protection*, pages 1–35, 2020.
- [17] S. Zhang, Y. Zhu, Z. You, and X. Wu. Fusion of superpixel, expectation maximization and phog for recognizing cucumber diseases. *Computers and Electronics in Agriculture*, 140:338–347, 2017.

Lampiran

