

Machine Learning Pengenalan Anura Berdasarkan Corak dan Warna

Hery Sunandar

Fakultas Ilmu Komputer Dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan

ARTICLE INFORMATION

Received: September, 2023
Revised: September 2023
Available online: Oktober 2023

KEYWORDS

Five words maximum, comma separated

CORRESPONDENCE

Phone: +62 (0751) 12345678
E-mail: herysun1975@gmail.com

ABSTRACT

The Identification of Anura (frog) based on pattern and color is a complex problem that takes a long time and costs quite a lot. Therefore, this study uses Machine Learning techniques to develop a frog species recognition model based on pattern and color. Image data of different species, patterns, and colors were taken from reliable sources and then divided into training and test data sets. Deep learning techniques were used to study the visual patterns in frog images and develop classification models that can predict frog species based on their patterns and colors. The test results show that the developed model is a fairly high accuracy and can correctly identify frog species based on their patterns and colors. However, the identification of frogs based on pattern and color may not always be accurate in some cases of the images tested. therefore, Machine Learning techniques must be combined with other frog identification methods. The methods used are Contour, Hough Line Transform, K-Means, and Logistic Regression. This research can assist in the conservation of endangered frog species by being able to identify frog species quickly to monitor frog populations in the wild.

PENDAHULUAN

Identifikasi Anura (katak) berdasarkan corak dan warna merupakan tantangan yang sulit dan memakan waktu yang lama. Meskipun para ahli telah mengembangkan metode identifikasi katak yang berbeda, seperti identifikasi suara dan morfologi fisik, namun metode ini masih memiliki keterbatasan dan mungkin tidak selalu akurat [1]. Katak adalah spesies hewan yang penting dalam ekosistem dan memiliki peran yang signifikan dalam rantai makanan. Namun, populasi berbagai jenis katak di beberapa daerah Indonesia telah menurun dikarenakan kerusakan habitat dan perubahan iklim. Oleh karena itu, penting untuk memantau populasi katak yang diketahui berkurang untuk dilestarikan agar dapat bertahan dan mampu berkembang pada masa yang akan datang. Anura dapat ditemukan di seluruh dunia, kecuali di daerah-daerah yang sangat dingin seperti kutub utara dan selatan. Anura memiliki berbagai jenis dan spesies, yang dibedakan oleh ciri-ciri fisik seperti bentuk tubuh, warna, dan corak pada kulit[2]

Machine Learning membangun model matematika berdasarkan data sampel dan dikenal sebagai *data-train* [3], [4]. *Machine Learning* ditujukan sebagai sistem yang mampu belajar dengan sendirinya dan memutuskan sesuatu tanpa harus berulang kali diprogram oleh manusia sehingga komputer menjadi semakin cerdas dengan cara belajar dari pengalaman data yang dimiliki komputer tersebut. Terdapat dua teknik pembelajarannya yaitu *supervised learning* menggunakan dataset (*data training*) yang sudah berlabel, sedangkan *unsupervised learning* menarik kesimpulan berdasarkan dataset [5], [6]. Input berupa dataset digunakan pembelajaran mesin untuk menghasilkan analisis yang benar [7].

Teknik *Machine Learning* telah menjadi populer dalam bidang identifikasi spesies hewan berdasarkan citra visual. *Machine Learning* dapat digunakan untuk mengenali pola visual yang berbeda pada gambar atau video hewan dan mempelajari fitur-fitur yang khas dari setiap spesies [8], [9]. Oleh karena itu, teknik *Machine Learning* dapat menjadi solusi yang potensial untuk mengatasi tantangan dalam identifikasi Anura berdasarkan corak dan warna [2]. Penggunaan teknik *Machine Learning* untuk klasifikasi dengan mengembangkan model pengenalan spesies katak berdasarkan corak dan warna. Penggunaan *dataset* gambar katak yang berbeda spesies, corak, dan warna dan kemudian membaginya menjadi set data latih dan uji. Teknik *deep learning* digunakan untuk mempelajari pola visual pada gambar katak dan mengembangkan model klasifikasi yang dapat memprediksi spesies katak berdasarkan corak dan warna [8], [9].

Penelitian sebelumnya telah menunjukkan bahwa teknik *Machine Learning* dapat digunakan untuk mengenali spesies hewan lainnya dengan akurasi yang tinggi berdasarkan citra visual. Namun, belum ada banyak penelitian yang dilakukan pada identifikasi Anura berdasarkan corak dan warna menggunakan teknik *Machine Learning*. Beberapa karya ilmiah ilmu komputer lokal melakukan penelitian tentang Anura dengan konsep sistem pendukung keputusan ataupun datamining secara konvensional, namun tidak melakukan identifikasi Anura dengan teknik klasifikasi atau pengelompokan.

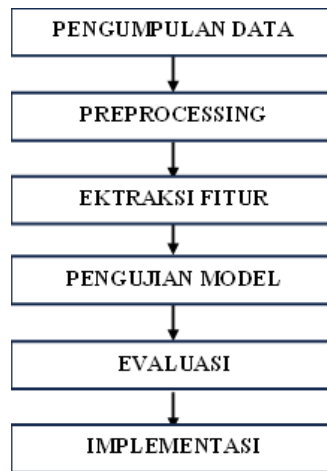
Penelitian ini memiliki implikasi penting dalam konservasi dan pengawetan spesies katak yang terancam punah dengan mampu mengenali spesies katak dengan cepat dan akurat berdasarkan corak dan warna, melakukan pemantauan populasi katak di alam liar dan mengambil tindakan konservasi yang diperlukan untuk memastikan kelangsungan hidup mereka [2]. Penelitian ini didasari oleh pengetahuan ilmu komputer yang dikembangkan dengan mempelajari dan memanfaatkan pengetahuan-pengetahuan bersumber dari penelitian biologi. Adapun sumber-sumber utama dari pengetahuan biologi yang pertama diambil dari buku panduan Keanekaragaman Amfibi di Sipirok, Tapanuli Selatan oleh Rahmat Aulan, S. Hut dan kawan-kawan. Buku panduan ini membahas bagaimana mengidentifikasi Anura yang ada di Sipirok. Sipirok berlokasinya berdekatan dengan area konservasi, yaitu Cagar Alam Dolok Sibual-buali. Sipirok diketahui mempunyai ekosistem hutan pegunungan rendah dan ekosistem pegunungan. Sumber pengetahuan kedua oleh Toni Febri Qurniawan dengan Ulasan singkat: Polimorfisme corak dan warna Anura. Dengan informasi minimnya tentang peristiwa polimorfisme corak dan warna pada Anura di Indonesia, maka dilakukan observasi terhadap variasi corak dan warna Anura dengan lokasi penelitian di Yogyakarta [1].

Tindakan ini mengambil tindakan konservasi yang tepat untuk menjaga kelangsungan hidup mereka. Identifikasi spesies katak berdasarkan morfologi fisik dan suara memerlukan waktu dan biaya yang cukup besar. Selain itu, metode ini mungkin tidak selalu akurat karena variasi dalam penampilan dan suara antara individu yang sama spesies. Dengan demikian penggunaan teknik *Machine Learning* dalam identifikasi spesies katak berdasarkan corak dan warna dapat memberikan solusi yang cepat dan efektif

untuk mengetahui jenis katak. Teknik ini dapat mempercepat proses identifikasi dan menghasilkan akurasi yang lebih tinggi dibandingkan dengan metode identifikasi tradisional [10]. Identifikasi spesies katak dengan akurasi yang tinggi sangat penting untuk konservasi spesies katak yang terancam punah. Dengan menggunakan teknik *Machine Learning* untuk mengenali spesies katak dengan cepat dan akurat berdasarkan corak dan warna katak.

METODE PENELITIAN

Penelitian ini menggunakan beberapa teknik atau model untuk disiapkan menjadi suatu sistem yang disebut *Machine Learning*. Arsitektur jaringan saraf tiruan yang dalam memungkinkan model untuk secara otomatis belajar fitur-fitur hierarkis dari data tanpa perlu ekstraksi manual. Ini dicapai melalui propagasi maju (*forward pass*) dan propagasi balik (*backward pass*) di seluruh lapisan jaringan untuk menyesuaikan bobot agar model dapat memahami pola-pola kompleks dalam data [4], [11]. Sistem yang dirancang ini untuk mampu mengidentifikasi Anura berdasarkan corak dan warna yang terdapat pada Anura tersebut. Untuk itu diperlukan ketentuan dalam mempelajari corak dan warna Anura secara spesifik. Tahapan penelitian diilustrasikan pada gambar di bawah.



Gambar 1. Tahapan penelitian

Berikut adalah keterangan dari alur penelitian dari gambar di atas untuk penelitian *Machine Learning* pengenalan Anura berdasarkan corak dan warna:

1. Pengumpulan data gambar corak dan warna pada kulit Anura dari berbagai sumber seperti buku, artikel ilmiah, atau sumber *online* yang terpercaya.
2. *Preprocessing* data untuk memastikan kualitas gambar yang baik dan Segmentasi gambar Anura untuk memisahkan gambar Anura dari latar belakang atau elemen lain yang tidak penting pada gambar dengan *Region Mask*.
3. Ekstraksi fitur corak dan warna pada gambar Anura menggunakan algoritma *Computer Vision* dan *Machine Learning* seperti algoritma K-Means, algoritma *Hough Line Transform* dan *Contour*.
4. Pengujian model dengan menggunakan data pengujian. Pada tahap ini, model menghitung data uji setiap fitur pada setiap jenis Anura, kemudian memilih jenis Anura dengan probabilitas tertinggi.
5. Evaluasi hasil pengujian untuk menentukan akurasi dan kinerja model. Jika hasil evaluasi tidak tersedia, model dapat dioptimalkan dengan melakukan menambah data baru untuk pelatihan. Namun pelabelan masih berbentuk data baru yang diinisialisasikan. Dibutuhkan peran pengguna untuk memberikan label pada data baru tersebut.
6. Implementasi model dalam aplikasi atau sistem yang digunakan untuk pengenalan jenis-jenis Anura berdasarkan corak dan warna dengan menggunakan algoritma *Hough Line Transform*, *Contour*, dan K-Means.

3.1 Anura

Anura adalah ordo hewan yang terdiri dari katak dan katak-katak kecil yang dikenal sebagai kodok atau katak berudu. Anura adalah salah satu dari tiga ordo dalam kelas *amphibia* (amfibi) bersama dengan *urodela* (salamander dan sirip) dan *apoda* (*caecilian*) [1]. Anura dapat ditemukan di seluruh dunia, kecuali di daerah-daerah yang sangat dingin seperti kutub utara dan selatan. Anura memiliki berbagai jenis dan spesies, yang dibedakan oleh ciri-ciri fisik seperti bentuk tubuh, warna, dan corak pada kulit[2].

Berikut adalah beberapa contoh variasi corak dan warna pada enam jenis Anura [1].:

1. *Duttaphrynus Melanostictus*
Duttaphrynus Melanostictus memiliki variasi tiga corak dan warna yang terdiri dari kuning, merah bata, dan coklat/krem. Sementara variasi yang banyak ditemukan yaitu merah dan coklat
2. *Ferjervarya Cancrivorra*
Ferjervarya Cancrivorra memiliki variasi empat corak dan warna yang terdiri dari hijau, hijau bergaris lurus, coklat, dan coklat bergaris lurus. Sementara variasi yang banyak ditemukan hijau bergaris lurus.
3. *Ferjervarya Limnocharis*
Ferjervarya Limnocharis memiliki variasi enam corak dan warna yang terdiri dari hijau bergaris lurus, hijau bergaris bengkok, coklat, coklat bergaris lurus, coklat bergaris bengkok, dan coklat kemerahan. Sementara variasi yang banyak ditemukan hijau bergaris lurus dan hijau bergaris bengkok.
4. *Oocydoziga Sumatrana*

Occydoziga Sumatrana memiliki variasi lima corak dan warna yang terdiri dari coklat, hitam, coklat bergaris lurus oranye, merah bergaris lurus merah, dan coklat bergaris lurus merah. Sementara variasi yang banyak ditemukan coklat.

5. *Occydoziga Lima*

Occydoziga Lima memiliki variasi lima corak dan warna yang terdiri dari coklat bertotol gelap, coklat bermoncong hijau, coklat bergaris lurus warna kuning, coklat bergaris lurus warna merah bata, dan coklat bergaris lurus warna hijau. Sementara variasi yang banyak ditemukan coklat bertotol gelap.

6. *Philautus Aurifasciatus*

Philautus Aurifasciatus memiliki variasi tiga belas corak dan warna yang terdiri dari coklat bergaris lurus, coklat corak huruf V, coklat corak huruf L, coklat corak huruf H, coklat corak huruf X, hijau corak huruf H, coklat/krem, hijau bercak coklat, hijau bercak gelap, coklat gelap tekstur kulit kasar, gelap bercak hijau kemerahan, dan coklat dengan lipatan timpanik warna oranye. Sementara variasi yang banyak ditemukan coklat corak huruf H dan hijau corak huruf H

3.2 Ekstraksi Tekstur Corak dan Warna





Fitur pertama yang digunakan untuk pendefinisian adalah corak Anura, fitur ini mendefinisikan ciri tekstur kulit tubuh belakang (punggung) Anura. Tekstur kulit Anura sangat berbeda baik satu jenis kelompok ataupun tidak. *Hough Line Transform* digunakan untuk mendeteksi garis lurus dalam gambar. *Hough Line Transform* bekerja dengan mengubah titik-titik dalam ruang koordinat menjadi garis-garis dalam ruang parameter [12]. *Hough Line Transform* mampu mengidentifikasi garis-garis pada gambar walaupun garis tersebut tidak sempurna atau terputus-putus. Pelacakan *Contour* digunakan pada tahap ini untuk mendeteksi tepi atau segmentasi corak pada tubuh Anura, sementara *Hough Line Transform* untuk lebih jelas menerapkan deteksi tepi dari tekstur. Berikut ini Anura yang memiliki corak untuk pendefinisian tekstur.















Tabel 1. Data corak untuk pendefinisian tekstur

No.	Nama Anura	Corak / motif tekstur
1.	<i>Ferjervarya Cancrivorra</i>	hijau bergaris lurus coklat bergaris lurus
2.	<i>Ferjervarya Limnocharis</i>	hijau bergaris lurus hijau bergaris bengkok coklat bergaris lurus coklat bergaris bengkok
3.	<i>Occydoziga Sumatrana</i>	coklat bergaris lurus oranye merah bergaris lurus merah coklat bergaris lurus merah
4.	<i>Occydoziga Lima</i>	coklat bertotol gelap coklat bertotol hijau coklat bergaris lurus kuning coklat bergaris lurus merah coklat bergaris lurus hijau
5.	<i>Philautus Aurifasciatus</i>	coklat bergaris lurus coklat corak huruf V coklat corak huruf L coklat corak huruf H coklat corak huruf X hijau corak huruf H hijau bercak coklat hijau bercak hitam coklat gelap tekstur kulit kasar coklat gelap bercak hijau kemerahan coklat dengan lipatan timpanik oranye

Ekstraksi fitur warna menggunakan algoritma K-Means. K-Means digunakan untuk pengelompokan (*clustering*) dan fungsi utamanya digunakan untuk mengelompokkan data ke dalam beberapa kelompok atau kluster berdasarkan kesamaan fitur. K-Means mencoba mempartisi data ke dalam kelompok-kelompok yang memiliki varian rendah dalam kluster dan varian tinggi antara kluster [13]. Algoritma ini difungsikan untuk kluster warna yang dominan yang terdapat pada gambar Anura. Kluster tiga warna dari enam jenis Anura untuk pertama kali dipersiapkan sebanyak sembilan data gambar.

Tabel 2. Data warna untuk pendefinisian tekstur

No.	Jenis Anura	ID Jenis	Gambar	Nilai Warna Kluster (W)	Warna Kluster
1.	<i>Duttaphrynus Melanostictus</i>	DM01		W1=137, 123, 95 W2=249, 248, 244 W3=58, 55, 34	
2.	<i>Duttaphrynus Melanostictus</i>	DM02		W1=180, 177, 81 W2=229, 229, 135 W3=130, 125, 39	

3.	<i>Duttaphrynus Melanostictus</i>	DM03		W1=199, 157, 100 W2=149, 113, 59 W3=230, 185, 123	
4.	<i>Ferjervarya Cancrivorra</i>	FC01		W1=181, 169, 169 W2=145, 130, 122 W3=98, 83, 68	
5.	<i>Ferjervarya Cancrivorra</i>	FC02		W1=76, 75, 70 W2=179, 157, 150 W3=111, 108, 107	
6.	<i>Ferjervarya Cancrivorra</i>	FC03		W1= 82, 78, 43 W2=172, 153, 130 W3=111, 107, 70	
7.	<i>Ferjervarya Cancrivorra</i>	FC04		W1=175, 163, 84 W2=140, 126, 49 W3=212, 204, 138	
8.	<i>Ferjervarya Limnocharis</i>	FL01		W1=114, 107, 103 W2=180, 164, 141 W3=78, 72, 70	
9.	<i>Ferjervarya Limnocharis</i>	FL02		W1=122, 111, 117, W2=172, 165, 192, W3=95, 83, 77	

Langkah-langkah untuk mengidentifikasi 3 warna dominan pada tabel di atas dari gambar RGB menggunakan algoritma K-Means adalah sebagai berikut:

1. Konversi Gambar ke Format Data yang Sesuai.
Ubah gambar RGB menjadi format data yang cocok untuk digunakan dengan algoritma K-Means. Setiap piksel akan direpresentasikan sebagai vektor dalam ruang berdimensi tiga (R, G, B).
2. Inisialisasi K-Means
Pilih nilai K (jumlah kluster) yang sama dengan 3 karena Anda ingin mengidentifikasi 3 warna dominan.
3. Inisialisasi Pusat Kluster
Pilih tiga titik awal secara acak sebagai pusat kluster awal. Ini bisa dilakukan dengan memilih tiga piksel acak dari gambar atau dengan metode lainnya.
4. Iterasi K-Means
Ulangi langkah-langkah berikut hingga konvergensi:
 - a. Penugasan Data: Hitung jarak antara setiap piksel dan tiga pusat kluster. Masukkan setiap piksel ke dalam kluster dengan pusat terdekat.
 - b. Pembaruan Pusat Kluster: Hitung ulang pusat kluster dengan mengambil rata-rata dari semua piksel dalam kluster.
5. Hasil Akhir:
Setelah konvergensi, tiga pusat kluster akan mewakili tiga warna dominan dalam gambar.
Penggunaan K-Means pada gambar RGB dapat digambarkan dengan rumus matematis sebagai berikut:
 1. Data gambar:
 $D = \{(R_1, G_1, B_1), (R_2, G_2, B_2), \dots, (R_n, G_n, B_n)\}$, dengan n adalah jumlah piksel dalam gambar.
Inisialisasi pusat kluster: C1, C2, C3.
 2. Langkah iterasi:
 - a. Penugasan Data:
Setiap piksel (Ri, Gi, Bi) diberikan ke kluster dengan pusat terdekat yaitu $\text{argmin}_j \|(R_i, G_i, B_i) - C_j\|^2$.
 - b. Pembaruan Pusat Kluster:
 $C_j = (\sum (R_i, G_i, B_i) \text{ dalam kluster } j) / (\text{jumlah piksel dalam kluster } j)$.
 3. Ulangi langkah iterasi hingga konvergensi (pusat kluster tidak berubah secara signifikan).

Setelah konvergensi, C1, C2, dan C3 akan mewakili tiga warna dominan dalam gambar.

Langkah-langkah di atas untuk mengidentifikasi 3 warna dominan dalam gambar RGB menggunakan K-Means. Implementasi praktisnya dapat dilakukan dengan bantuan pustaka pemrosesan citra OpenCV atau pustaka pemrosesan data seperti scikit-learn dalam bahasa pemrograman seperti Python.

HASIL DAN PEMBAHASAN

4.1 Matrik X dan Y

Matrik x adalah pengabungan dari matrik fitur corak dan fitur warna. Untuk pengolahan awal berikutnya akan menggunakan data fitur warna sebanyak 9 warna, sedangkan untuk matrik fitur corak sudah dimiliki untuk keseluruhan data Anura sebanyak 37 corak. Namun untuk matrik x awal disesuaikan jumlah data fitur corak sebanyak 9 corak.

No	ID Anura	Corak R	Corak G	Corak B	Garis R	Garis G	Garis B	Garis Lurus	Garis Bengkok	Bertotol	V	L	H	X	Kasar	Bercak	Lipatan
1	DM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	DM2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	DM3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	FC1	181	169	169	0	0	0	1	0	0	0	0	0	0	0	0	0
5	FC2	76	75	70	0	0	0	1	0	0	0	0	0	0	0	0	0
6	FC3	82	78	43	172	153	130	1	0	0	0	0	0	0	0	0	0
7	FC4	175	163	84	140	126	49	1	0	0	0	0	0	0	0	0	0
8	FL1	114	107	103	180	164	141	1	0	0	0	0	0	0	0	0	0
9	FL2	122	111	117	172	165	192	0	1	0	0	0	0	0	0	0	0

Gambar 2. Daftar Matrik fitur Corak

Pada tabel 3 di atas didapatkan corak Anura sebanyak 16 fitur corak. Untuk jenis Anura *Duttaphrynus Melanostictus* tidak memiliki corak khusus maka nilai fitur coraknya 0 semua dan Anura jenis ini lebih dibedakan pada fitur warna. Sedangkan untuk *Ferjervarya Cancrivorra* semuanya memiliki corak garis lurus bernilai 1 dan beberapa memiliki nilai warna R, G, dan B corak dan garis. Jenis *Ferjervarya Limnocharis* baru memiliki dua data corak dengan rincian 1 bergaris lurus dan 1 bergaris bengkok dan keduanya memiliki nilai warna R, G, dan B corak dan garis seperti pada tabel di atas.

No	ID Anura	D1 R	D1 G	D1 B	D2 R	D2 G	D2 B	D3 R	D3 G	D3 B
1	DM1	137	123	95	249	248	244	58	55	34
2	DM2	180	177	81	229	229	135	130	125	39
3	DM3	199	157	100	149	113	59	230	185	123
4	FC1	181	169	169	145	130	122	98	83	68
5	FC2	76	75	70	179	157	150	111	108	107
6	FC3	82	78	43	172	153	130	111	107	70
7	FC4	175	163	84	140	126	49	212	204	138
8	FL1	114	107	103	180	164	141	78	72	70
9	FL2	122	111	117	172	165	192	95	83	77

Gambar 2. Daftar Matrik fitur warna

Tabel 4 di atas merupakan nilai dominan fitur warna yang terbagi atas 3 klaster, untuk masing-masing klaster memiliki nilai R, G, dan B. Nilai fitur warna dominan 1 digunakan untuk nilai warna corak dan nilai warna dominan 2 digunakan untuk nilai warna garis. Namun, nilai ini digunakan jika jenis Anura tersebut memiliki corak, jika tidak maka nilai warna corak dan warna garis bernilai 0. Nilai warna R, G, B pada tersebut hanya sebagai penetapan warna awal karena untuk identifikasi warna R, G, dan B kurang tepat, dalam penerapannya akan dikonversi ke warna HSV dalam bentuk jarak warna.

Matrik y disebutkan sebagai kelas yang diambil dari jumlah jenis Anura. Kelas untuk matrik y terdiri dari 6 yaitu jumlah Anura sebanyak 6 jenis. Maka matrik y = (1, 2, 3, 4, 5, 6) untuk 9 sampel gambar awal dan seterusnya nanti untuk keseluruhan sampel gambar.

4.2 Proses klasifikasi *Logistic Regression*

Logistic Regression berfungsi sebagai statistik dan Machine Learning yang digunakan untuk melakukan klasifikasi pada data dengan dua kelas atau lebih. Meskipun namanya mengandung "*regression*," namun ini untuk klasifikasi dan bukan regresi [14], [15]. *Logistic Regression* memodelkan probabilitas bahwa suatu data masuk ke dalam salah satu kelas. [16], [17]. Berikut langkah-langkah dalam penyelesaian algoritma *Logistic Regression* untuk data matriks X dan label kelas y dari dataset Anura di atas, yaitu:

1. Inisialisasi Parameter
Inisialisasi bobot (*weights*) dan bias dengan nilai acak kecil atau nol.
2. Transformasi Fungsi Logit
Hitung nilai logit (*log-odds*) untuk setiap sampel gambar dengan menggunakan matriks fitur-fitur X, bobot, dan bias. Logit adalah hasil perkalian matriks X dengan bobot ditambah dengan nilai bias.
3. Fungsi *Sigmoid*
Hitung nilai *sigmoid* untuk setiap nilai logit. *Sigmoid* adalah fungsi non-linear yang mengubah nilai logit menjadi probabilitas kelas positif (1).
4. Fungsi *Loss* (Fungsi Biaya)
Hitung fungsi *loss* untuk seluruh sampel gambar berdasarkan perbedaan antara nilai prediksi *sigmoid* dan label kelas sebenarnya. Fungsi *loss* umumnya menggunakan *log-loss* (atau *cross-entropy loss*) yang mengukur kesalahan prediksi.
5. Penurunan Gradien
Hitung gradien (turunan parsial) dari fungsi *loss* terhadap bobot dan bias. *Update* bobot dan bias menggunakan gradien descent (metode penurunan gradien) dengan langkah pembelajaran (*learning rate*) tertentu. Proses ini membantu menemukan nilai bobot dan bias yang mengurangi fungsi *loss*.
6. Iterasi dan Konvergensi
Ulangi langkah 2-5 untuk sejumlah iterasi tertentu atau hingga tercapai kondisi konvergensi.
7. Prediksi
Setelah model dilatih, gunakan nilai bobot dan bias yang telah dipelajari untuk memprediksi label kelas pada data baru. Hitung nilai logit dan *sigmoid* untuk data baru, lalu tentukan kelas berdasarkan probabilitas *sigmoid*.

Penting untuk memahami setiap langkah dalam algoritma ini secara mendalam, termasuk perhitungan matematika yang lebih rinci. Langkah-langkah ini melibatkan operasi matriks dan vektor yang lebih kompleks. Regularisasi sering digunakan untuk menghindari *overfitting* dan dapat ditambahkan pada langkah 4 atau 5. Algoritma *Logistic Regression* dalam implementasinya melibatkan kerangka kerja pemrosesan data dan optimasi numerik yang telah ada, seperti NumPy, TensorFlow, atau scikit-learn dalam Python.

4.3 Labeling

Tahap labeling ini menampilkan informasi langsung pada gambar setelah dilakukan pemindaian gambar sesuai konsep ekstraksi motif dan warna. Hasil dari proses ekstraksi motif dan warna untuk pertama kali dilakukan sebagai awal dasar pengetahuan yang dimiliki sistem. Hasil tersebut dapat disimpan dalam format CSV sebagai dataset Anura yang telah dimiliki. Pemberian label pada gambar input uji merupakan bentuk informasi hasil deteksi Anura yang telah dimiliki sebagai bentuk pengetahuan dalam sistem.

Saat pemberian label, jika diketahui informasi hasil tidak terdapat dalam dataset, maka sistem akan menyimpan dalam bentuk nama Anura dan id Anura yang dibuat secara indeks yang belum terdefinisi namun serta menyimpan ekstraksi corak dan warnanya. Pada bagian ini dibutuhkan upaya *user* pengelola sistem untuk mengubah nama Anura dan id Anura secara manual dan menentukan perlu atau tidaknya pembuatan kelas baru. Jika *user* pengelola sistem tidak mengubah nama Anura dan id Anura, serta penetapan kelas maka, sistem tetap akan mengidentifikasi gambar Anura baru sebagai gambar input yang berlaku secara indeks sebelumnya dan kelas baru. Inilah bentuk sistem yang belajar dari pengetahuan Anura yang dimilikinya.

4.4 Identifikasi

Pengujian *Machine Learning* pengenalan Anura berdasarkan corak dan warna memiliki tampilan GUI yang sederhana, namun memiliki penerapan algoritma-algoritma untuk region mask, deteksi tepi, segmentasi, kluster, dan klasifikasi untuk mengolah gambar. Adapun input gambar berfungsi untuk pengetahuan sebagai dataset dan/atau sebagai input gambar yang diuji untuk melakukan identifikasi. Tampilan GUI pengenalan Anura beserta hasil identifikasi dapat dilihat pada gambar 1 di bawah. Pada aplikasi tahapan awal dilakukan mengambil gambar dengan menekan tombol *browse* dan setelah itu menekan tombol proses. Aplikasi akan bekerja sesuai dengan algoritma-algoritma yang diterapkan dan menggunakan fungsi OpenCV dan librari-librari Python proses tersebut dikerjakan dengan singkat dan baik.



Gambar 3. Hasil pengujian identifikasi Anura

KESIMPULAN

Kesimpulan dari penelitian yang telah dilakukan mengungkapkan pencapaian penelitian ini, yaitu:

1. Tahapan *Preprocessing* dilakukan dengan algoritma *Hough Line Transform*, *Contour*, dan K-Means untuk menghasilkan matrik x dan y. Matrik x dan y tersebut untuk proses identifikasi Anura (katak) berdasarkan corak dan warna melalui *Machine Learning* dengan algoritma *Logistic Regression* dengan menerapkan nilai x dari data fitur warna dan corak, sedangkan matrik y adalah jumlah kelas berasal dari jumlah jenis anura.
2. Hasil pengujian gambar input dengan data latih gambar memberikan hasil yang baik namun terdapat beberapa hasil gambar yang tidak akurat, hal ini terdapat pada Anura *Philautus Aurifasciatus* yang memiliki banyak corak unik pada gambar latihnya.
3. Perancangan sistem untuk *Machine Learning* menerapkan OpenCV dan Bahasa Python 2.0 sebagai fungsi pembacaan dan proses data latih dan data uji gambar pengenalan Anura.

DAFTAR PUSTAKA

- [1] T. Febri Q, "Ulasan Singkat Polimorfisme Corak Dan Warna Anura," *Fauna Indonesia : Masyarakat Zoologi Indonesia*, vol. 13, no. 2, 2014.
- [2] D. Satrla Yudha, R. Eprilurhman, M. Falsal ALawI, and A. Tarekat, "KEANEKARAGAMAN JENIS KATAK DAN KODOK (ORDO ANURA) DI SEPANJANG SUNGAI OPAK PROPINSI DAERAH ISTIMEWA YOGYAKARTA THE DIVERSITY OF FROGS AND TOADS (ANURA) ON ALONG THE OPAK RIVER AT THE PROVINCE OF DAERAH ISTIMEWA YOGYAKARTA," 2014.
- [3] A. S. Diantika and A. Yuki Firmanto, SE., MSA., CA., "Implementasi Machine Learning Pada Aplikasi Penjualan Produk Digital (Studi Pada Grabkios)," vol. 53, no. 9, pp. 1689–1699, 2013.

- [4] R. R. Pratama, "Analisis Model Machine Learning Terhadap Pengenalan Aktifitas Manusia," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 19, no. 2, pp. 302–311, 2020, doi: 10.30812/matrik.v19i2.688.
- [5] M. H. Rifqo and A. Wijaya, "IMPLEMENTASI ALGORITMA NAIVE BAYES DALAM PENENTUAN PEMBERIAN KREDIT," 2017. [Online]. Available: www.ejournal.unib.ac.id/index.php/pseudocode
- [6] Y. Sitoto, T. Allo, V. Sofica, N. Hasan, and M. Septiani, "Penggunaan Metode Naïve Bayes Dalam Mengklasifikasi Pengangguran Pada Desa Bojong Kulur," *Bianglala Informatika*, vol. 10, no. 1, p. 2022, 2022.
- [7] E. Retnoningsih and R. Pramudita, "Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python," *Bina Insani Ict Journal*, vol. 7, no. 2, p. 156, 2020, doi: 10.51211/biict.v7i2.1422.
- [8] D. Sarkar, R. Bali, and T. Sharma, *Practical Machine Learning with Rust*. 2020. doi: 10.1007/978-1-4842-5121-8.
- [9] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian Journal on Computer and Information Technology)*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
- [10] Y. Hasan and K. Siregar, "COMPUTER VISION IDENTIFICATION OF SPECIES, SEX, AND AGE OF INDONESIAN MARINE LOBSTERS," *JURNAL INFOKUM*, vol. 9, no. 2, pp. 478–489, 2021, [Online]. Available: <http://infor.seaninstitute.org/index.php/infokum/index>
- [11] M. Kümmerer, M. Bethge, and T. S. A. Wallis, "DeepGaze III: Modeling free-viewing human scanpaths with deep learning," *J Vis*, vol. 22, no. 5, pp. 1–27, Apr. 2022, doi: 10.1167/jov.22.5.7.
- [12] C. R. Qi, O. Litany, K. He, and L. J. Guibas, "Deep Hough Voting for 3D Object Detection in Point Clouds," 2019. doi: <https://doi.org/10.48550/arXiv.1904.09664>.
- [13] M. R. Larijani, E. A. Asli-Ardeh, E. Kozegar, and R. Loni, "Evaluation of image processing technique in identifying rice blast disease in field conditions based on KNN algorithm improvement by K-means," *Food Sci Nutr*, vol. 7, no. 12, pp. 3922–3930, Dec. 2019, doi: 10.1002/fsn3.1251.
- [14] J. M. Hilbe, "Practical Guide to Logistic Regression."
- [15] S. Annas, A. Aswi, M. Abdy, and B. Poerwanto, "Stroke Classification Model using Logistic Regression," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Dec. 2021. doi: 10.1088/1742-6596/2123/1/012016.
- [16] W. H. Nugroho, S. Handoyo, Y. J. Akri, and A. D. Sulistyono, "Building Multiclass Classification Model of Logistic Regression and Decision Tree Using the Chi-Square Test for Variable Selection Method," *Journal of Hunan University Natural Sciences*, vol. 49, no. 4, pp. 172–181, Apr. 2022, doi: 10.55463/issn.1674-2974.49.4.17.
- [17] C. L. Tsien A ' B, H. S. F. Fraser, W. J. Long, and R. L. Kennedy, "Using Classification Tree and Logistic Regression Methods to Diagnose Myocardial Infarction," 1998.