ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

KLASIFIKASI TIGA JENIS MACAN TUTUL BERBASIS PENGOLAHAN CITRA DIGITAL MENGGUNAKAN KOMBINASI FITUR RGB, GLCM, PCA DAN SVM

Joanne Olivia, Adriano Deshon✉, Khenichi Kuolimpo
Teknik Informatika, Universitas Tarumanagara, Jakarta, Indonesia
Email: adriano.535230017@stu.untar.ac.id

DOI: https://doi.org/10.46880/jmika.Vol10No2.pp855-862

ABSTRACT

Classification of leopard species based on digital images is one of the applications of image processing and data mining that can support the automatic identification of wildlife. This study aims to develop a classification model for three leopard species, namely African Leopard, Clouded Leopard, and Snow Leopard, based on the color and texture characteristics of images. The research process includes image pre-processing, color feature extraction using the Red, Green, Blue (RGB) channels, and texture feature extraction using the Gray-Level Co-occurrence Matrix (GLCM), which produces contrast, correlation, energy, and homogeneity features. The extracted features are then reduced using Principal Component Analysis (PCA) while retaining 95% of the information variance, before being classified using Support Vector Machine (SVM) with the multi-class Error-Correcting Output Codes (ECOC) approach. Evaluation is conducted using three data-splitting scenarios for training and testing sets. The results show that the scenario with 70% training data and 30% testing data achieves the highest accuracy of 87.06%, with class accuracies of 87.5% for African Leopard, 79.3% for Clouded Leopard, and 94.6% for Snow Leopard. These results indicate that the combination of RGB, GLCM, PCA, and SVM can classify the three leopard species with reasonably good performance.

Keywords: Citra Classification, Leopard, RGB, GLCM.

ABSTRAK

Klasifikasi spesies macan tutul berdasarkan citra digital merupakan salah satu penerapan pengolahan citra dan data mining yang dapat membantu proses identifikasi satwa secara otomatis. Penelitian ini bertujuan mengembangkan model klasifikasi tiga jenis macan tutul, yaitu African Leopard, Clouded Leopard, dan Snow Leopard berdasarkan karakteristik warna dan tekstur citra. Tahapan penelitian meliputi pre-processing citra, ekstraksi fitur warna menggunakan kanal Red, Green, Blue (RGB), dan juga ekstraksi fitur tekstur digunakan Gray-Level Co-occurrence Matrix (GLCM) dan menghasilkan fitur contrast, correlation, energy, dan homogeneity. Fitur yang diperoleh kemudian direduksi menggunakan Principal Component Analysis (PCA) dengan mempertahankan 95% variansi informasi, sebelum diklasifikasikan menggunakan Support Vector Machine (SVM) dengan pendekatan multi-kelas Error-Correcting Output Codes (ECOC). Evaluasi dilakukan menggunakan tiga skenario pembagian data latih dan data uji. Hasil menunjukkan bahwa skenario 70% data latih dan 30% data uji memberikan akurasi tertinggi sebesar 87,06%, dengan akurasi kelas African Leopard sebesar 87,5%, Clouded Leopard 79,3%, dan Snow Leopard 94,6%. Hasil tersebut menyimpulkan bahwa kombinasi RGB, GLCM, PCA, dan SVM dapat mengklasifikasikan tiga jenis macan tutul dengan performa yang cukup baik.

Kata Kunci: Klasifikasi Citra, Macan Tutul, RGB, GLCM.

PENDAHULUAN

Macan tutul (*leopard*) merupakan salah satu spesies kucing besar dari famili Felidae yang tersebar luas di Asia dan Afrika. Tiga jenis yang paling dikenal masyarakat adalah *African Leopard, Snow Leopard*, dan *Clouded Leopard*. Meski sama-sama disebut "leopard", *Snow Leopard* (*Panthera uncia*) dan *Clouded Leopard* (*Neofelis nebulosa*) sebenarnya tergolong spesies berbeda dari *African Leopard* (*Panthera pardus*), namun kerap disamakan karena kemiripan morfologis. Perbedaan pola tubuh dan tekstur bulu di antara ketiganya justru membuka peluang untuk membedakannya melalui analisis citra visual.

Pemantauan populasi macan tutul di alam liar umumnya mengandalkan kamera jebak, di mana metode ini memungkinkan pencatatan perilaku satwa tanpa campur tangan langsung manusia. Namun, di balik kepraktisannya, metode ini menyisakan persoalan baru, yaitu proses identifikasi jenis macan tutul dari tumpukan foto yang dihasilkan masih dilakukan secara manual, sehingga menyita banyak waktu dan tenaga,

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

sekaligus rawan terhadap kekeliruan penilaian oleh manusia (Bohnett et al., 2023). Persoalan inilah yang mendorong perlunya sistem identifikasi otomatis yang lebih cepat dan konsisten.

Pengolahan citra digital yang dikombinasikan dengan pendekatan data mining menawarkan jalan keluar atas persoalan tersebut. Karakteristik visual seperti pola tekstur dan corak warna bulu dapat diolah menjadi fitur numerik yang kemudian dipelajari oleh sistem untuk membedakan ketiga jenis macan tutul. Fitur tekstur citra dapat diperoleh melalui *Gray Level Cooccurrence Matrix* (GLCM). Metode ini telah banyak diadopsi dalam riset klasifikasi citra karena kemampuannya menggambarkan variasi, keteraturan, dan keseragaman pola melalui parameter kontras, korelasi, homogenitas, dan energi (Muhathir et al., 2021; Pariyandani et al., 2020). Di sisi lain, informasi warna dapat digali dari distribusi kanal Red Green Blue (RGB) untuk menangkap proporsi warna yang mendominasi setiap citra (Khomiak & Benndorf, 2021). Gabungan kedua jenis fitur ini kemudian dapat dipersempit dimensinya lewat *Principal Component Analysis* (PCA), sehingga kompleksitas data berkurang tanpa mengorbankan informasi yang esensial (Lu et al., 2024). Fitur hasil reduksi tersebut selanjutnya menjadi masukan bagi algoritma *Support Vector Machine* (SVM) yang dikenal efektif untuk menangani data berdimensi tinggi serta mampu memisahkan kelas data baik secara linear maupun nonlinear melalui pendekatan kernel (Abdullah & Abdulazeez, 2021; Pariyandani et al., 2020).

Berdasarkan latar belakang tersebut, maka penelitian ini memiliki tujuan untuk mengembangkan sistem klasifikasi tiga jenis macan tutul menggunakan kombinasi GLCM, RGB, PCA, dan SVM, serta mengevaluasi tingkat akurasinya. Data citra bersumber dari dataset cats-in-the-wild di Kaggle dan Google Images. Cakupan analisis ini dibatasi hanya pada fitur tekstur dan warna bulu. Secara akademis, penelitian ini diharapkan menjadi acuan pengembangan metode klasifikasi citra berbasis data mining dan dapat menjadi solusi otomatis untuk mengurangi ketergantungan pada identifikasi manual yang lambat dan rawan bias.

.

TINJAUAN PUSTAKA

Penelitian klasifikasi tiga jenis macan tutul berbasis pengolahan citra digital akan menggunakan empat metode. Proses diawali dengan menggunakan model warna *Red Green Blue* (RGB) yang digunakan untuk mengekstraksi fitur warna dari citra, dilanjutkan dengan metode *Gray Level Co-occurrence Matrix* (GLCM) untuk mengekstraksi fitur tekstur. Selanjutnya, dibahas *Principal Component Analysis* (PCA) sebagai teknik reduksi dimensi yang digunakan untuk menyederhanakan fitur hasil ekstraksi tanpa menghilangkan informasi penting, serta *Support Vector Machine* (SVM) sebagai algoritma klasifikasi yang digunakan untuk membedakan jenis macan tutul berdasarkan fitur-fitur yang telah diperoleh. Keempat metode ini saling berkaitan dan menjadi kerangka utama dalam proses eksperimen yang dilakukan pada penelitian ini.

Ekstraksi Fitur Warna *Red Green Blue* (RGB)

RGB adalah salah satu model warna yang paling umum diterapkan dalam pengolahan citra digital. Pada model ini, setiap piksel direpresentasikan melalui kombinasi tiga komponen warna dasar, yaitu merah

(*red*), hijau (*green*), dan biru (*blue*). Nilai intensitas dari ketiga kanal tersebut dapat diolah menjadi fitur warna yang berguna dalam proses klasifikasi citra. Citra berformat RGB mengandung beragam informasi visual, mulai dari bentuk objek, transformasi geometris, hingga tekstur permukaan (Khomiak & Benndorf, 2021). Kombinasi ketiga kanal warna ini memungkinkan sistem menangkap variasi nuansa dan tingkat saturasi warna secara lebih detail, sehingga proporsi dominasi masing-masing kanal pada suatu citra dapat dijadikan indikator pembeda antar kelas objek.

Gray Level Co-occurrence Matrix (GLCM)

GLCM merupakan pendekatan statistik yang sering digunakan untuk menganalisis pola tekstur pada citra digital dengan mengolah hubungan antar-piksel berdasarkan tingkat intensitasnya, kemudian menghitung nilai histogram tingkat kedua (Muhathir et al., 2021). Metode ini pertama kali diperkenalkan oleh Robert Haralick pada tahun 1973 dan banyak dimanfaatkan dalam analisis tekstur karena mampu mengungkap informasi mendalam seperti kontras, korelasi, homogenitas, dan energi dari suatu citra (Pariyandani et al., 2020). Kelebihan GLCM terletak pada kemudahannya dalam mengukur pola permukaan yang tidak seragam serta tingkat kekasaran visual pada gambar.

Secara teknis, GLCM menghitung seberapa sering pasangan nilai kecerahan piksel tertentu muncul berdampingan dalam citra, lalu merepresentasikannya sebagai matriks berukuran sesuai jumlah tingkat keabuan gambar. Matriks ini dibangun berdasarkan pasangan piksel pada arah tertentu (0°,45°,90°, dan 135°). Proses dimulai dengan mengonversi citra RGB menjadi skala abu-abu, kemudian dilanjutkan dengan penentuan nilai qwd intensitas maksimum. Setelah matriks *co-occurrence* terbentuk, setiap elemennya

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

dinormalisasi dengan membagi nilai elemen terhadap jumlah total nilai dalam matriks, sehingga hasilnya tidak bergantung pada ukuran citra. Hasil normalisasi ini selanjutnya digunakan sebagai dasar perhitungan probabilistik untuk mengidentifikasi empat fitur tekstur utama berikut (Pariyandani et al., 2020):

1. **Contrast**
   Mengukur besarnya variasi atau keragaman intensitas keabuan pada citra dengan rumus berikut:

$$\sum_{x,y}(x-y)^2 p(x,y)$$

2. **Correlation**
   Mengukur derajat hubungan linier antar pasangan piksel dalam citra dengan rumus berikut:

$$\sum_{x,y}\frac{(x-\mu_x)(y-\mu_y)p(x,y)}{\sigma_x\sigma_y}$$

3. **Homogeneity**
   Mengukur tingkat keseragaman level keabuan yang sejenis dalam citra dengan rumus berikut:

$$\sum_{x,y}\frac{p(x,y)}{1+|x-y|}$$

4. **Energy**
   Mengukur tingkat keseragaman (*uniformity*) atau konsentrasi nilai dalam matriks, di mana semakin tinggi nilainya, semakin seragam pola teksturnya. Berikut rumus yang digunakan:

$$\sum_{x,y}p(x,y)^2$$

Keempat fitur di atas dipilih karena telah terbukti efektif digunakan pada berbagai penelitian klasifikasi citra digital, salah satunya pada klasifikasi citra jenis wayang (Muhathir et al., 2021). Kombinasi *contrast, correlation, homogeneity*, dan *energy* dinilai mampu mewakili dimensi utama karakteristik tekstur, yakni variasi, pola, keseragaman, dan keteraturan citra.

***Principal Component Analysis* (PCA)**

PCA merupakan metode statistik yang dipakai untuk mereduksi dimensi data citra. Penerapan metode ini bertujuan untuk menekan kompleksitas data, menyaring fitur-fitur yang paling signifikan, serta mempercepat dan meningkatkan efisiensi proses klasifikasi (Lu et al., 2024). Tahapan pengolahan data dengan PCA dapat dimulai dengan membentuk matriks kovarians dari data latih untuk mengidentifikasi hubungan antar fitur dengan persamaan berikut:

$$\sum = \frac{1}{n}X^T X$$

di mana X adalah matriks data yang telah dinormalisasi dan Σ adalah matriks kovarians. Kemudian dilakukan dekomposisi eigen untuk memperoleh eigenvalue dan eigenvector yang merepresentasikan variansi maksimum data dengan persamaan berikut:

$$\sum e_i = \lambda_i e_i$$

Komponen utama (*principal components*) akan dipilih berdasarkan eigenvalue terbesar. Langkah terakhir adalah memproyeksi data ke ruang komponen utama untuk memperoleh fitur-fitur utama dengan persamaan berikut:

$$X_{proj} = XE$$

di mana X adalah matriks data asli, E adalah matriks vector-vektor eigen terpilih (*principal components*), dan $X_{\%\&'(}$ adalah data hasil proyeksi ke ruang PCA. Vektor-vektor utama tersebut disusun berdasarkan urutan variansi:

$$Var(XU_1) \geq VAR(XU_2) \geq VAR(XU_p)$$

***Support Vector Machine* (SVM)**

SVM adalah algoritma *supervised learning* yang pertama kali diperkenalkan oleh Vapnik dan digunakan baik untuk keperluan klasifikasi maupun regresi (Pariyandani et al., 2020). Prinsip dasar metode ini adalah mencari *hyperplane optimal* dengan margin maksimum yang mampu memisahkan kelas-kelas data secara jelas. Apabila data tidak dapat dipisahkan secara linear, SVM menerapkan *kernel trick* untuk memetakan data ke ruang berdimensi lebih tinggi agar pemisahan antar kelas dapat dilakukan. Bentuk umum persamaan *hyperplane* dinyatakan sebagai berikut (Muhathir et al., 2021):

$$w \cdot \phi(x) + b = 0$$

di mana w adalah vektor bobot, $\phi(x)$ adalah fungsi pemetaan (kernel), dan b adalah nilai bias. Adapun fungsi optimisasi yang digunakan dalam SVM adalah:

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

$$\min_{\alpha}\psi(\alpha) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j K(x_i, x_j)\alpha_i\alpha_j - \sum_{i=1}^{n}\alpha_i$$

dengan batasan $\sum y_i\alpha_i = 0{,}0 \leq \alpha_i \leq C$. SVM memiliki keunggulan dalam menangani data berdimensi tinggi serta fleksibel digunakan pada kasus data linear maupun non-linear melalui penerapan kernel (Muhathir et al., 2021; Pariyandani et al., 2020). Namun demikian, metode ini juga memiliki keterbatasan berupa beban komputasi yang cukup besar ketika diterapkan pada dataset berskala besar.

METODE PENELITIAN

Secara umum, alur penelitian meliputi pengumpulan dataset, pra-pemrosesan data, hingga tahapan eksperimen yang mencakup ekstraksi fitur, reduksi dimensi, klasifikasi, dan evaluasi performa model.

Pengumpulan Dan Pembagian Dataset

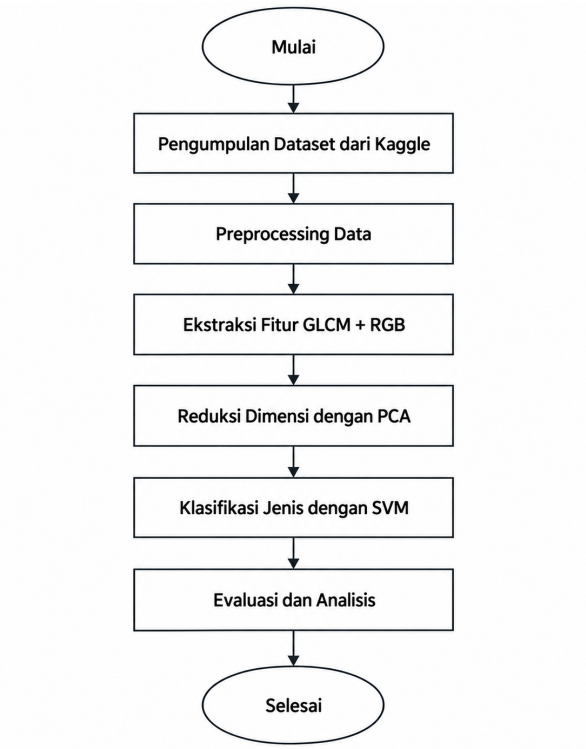
Pengumpulan data dilakukan dengan mengambil citra digital dari tiga spesies *big cats*, yaitu *African Leopard* (*Panthera pardus*), *Snow Leopard* (*Panthera uncia*), dan *Clouded Leopard* (*Neofelis nebulosa*). Data dikumpulkan dari *platform* Kaggle (*Cats in the Wild Image Classification*) serta gambar tambahan yang bersumber dari Google Images. Pemilihan ketiga spesies ini didasarkan pada keunikan karakteristik visual serta pola tutul pada tubuh hewan yang khas, sehingga sangat relevan untuk diidentifikasi secara otomatis berbasis pengolahan citra. Dataset yang digunakan pada penelitian ini disediakan dalam tiga folder skenario pembagian data latih (*train*) dan data uji (*test*), yaitu 60:40, 70:30, dan 80:20, dengan seluruh citra pada masing-masing folder skenario digunakan apa adanya untuk pelatihan dan pengujian model. Total citra (data latih dan data uji) pada ketiga folder skenario tersebut berturut-turut berjumlah 607, 602, dan 588 citra, dengan rincian jumlah citra per kelas pada setiap skenario disajikan pada Tabel 2.

Pra-Pemrosesan Data

Tahap pra-pemrosesan data dilakukan untuk menyelaraskan kualitas dan struktur citra sebelum memasuki tahap ekstraksi fitur. Langkah awal dimulai dengan pembersihan latar belakang (*background removal*) menggunakan alat bantu otomatis *remove.bg*. Proses ini bertujuan untuk menghilangkan elemen visual non-objek yang tidak relevan di sekitar hewan, sehingga analisis dapat berfokus penuh pada karakteristik morfologi dan tekstur tubuh macan tutul. Selanjutnya, seluruh citra diseragamkan ukurannya menjadi 224 X 224 piksel untuk menyesuaikan dengan standar masukan pada model serta mencegah terjadinya kesalahan komputasi saat proses pelatihan. Sebagai penutup tahap ini, citra yang telah dibersihkan dan diubah ukurannya disusun secara terstruktur ke dalam direktori subfolder sesuai dengan label kelas masing-masing pada folder *train* dan *test*.

Diagram Alir Proses Eksperimen

Tahapan proses klasifikasi jenis macan tutul secara keseluruhan dijelaskan melalui alur pada diagram alir eksperimen. Proses dimulai dari tahapan pengumpulan dataset yang diambil dari *platform* Kaggle dan Google Images untuk tiga jenis macan tutul, yaitu *African Leopard*, *Snow Leopard*, dan *Clouded Leopard*. Kemudian, dilakukan tahapan *pre-processing* data untuk menyiapkan citra sebelum diproses dan dianalisis lebih lanjut, yang mencakup pengorganisasian data, penghapusan latar belakang, serta penyesuaian ukuran citra menjadi 224 × 224 piksel.



Gambar 1. Diagram Alir Proses Klasifikasi Macan Tutul

Setelah data melalui tahapan *pre-processing*, tahap selanjutnya adalah ekstraksi fitur dengan menggunakan dua pendekatan, yaitu fitur warna dari kanal RGB dan fitur tekstur menggunakan metode *Gray Level Cooccurrence Matrix* (GLCM). Fitur-fitur

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

tersebut kemudian dimasukkan ke dalam tahap reduksi dimensi menggunakan metode *Principal Component Analysis* (PCA) untuk menyederhanakan jumlah fitur sekaligus mempertahankan informasi penting dari data. Fitur yang telah direduksi kemudian digunakan dalam proses klasifikasi jenis. macan tutul menggunakan algoritma *Support Vector Machine* (SVM). Tahapan terakhir yang dilakukan adalah melakukan evaluasi dan analisis terhadap hasil klasifikasi untuk menilai performa model berdasarkan metrik-metrik seperti akurasi, presisi, dan *recall*. Proses berakhir setelah seluruh tahapan analisis dan evaluasi selesai dilakukan.

Konfigurasi Parameter Eksperimen

Untuk mendukung replikasi penelitian, Tabel 1 berikut merangkum konfigurasi parameter yang digunakan pada eksperimen.

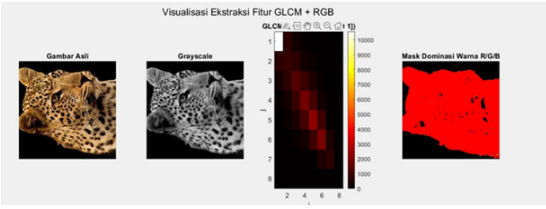
Tabel 1. Konfigurasi Parameter Eksperimen

Komponen	**Parameter**	**Nilai yang digunakan**
Ekstraksi warna RGB	Statistik yang dihitung per kanal (R, G, B)	Rata-rata dan standar deviasi intensitas piksel per kanal, dihitung hanya pada piksel foreground (non-transparan) hasil background removal
GLCM	Jarak piksel (d)	d = 1
GLCM	Sudut arah (θ)	0°, 45°, 90°, 135° (dirata-ratakan)
GLCM	Jumlah level keabuan	32 level (citra grayscale dikuantisasi ulang ke 32 level sebelum GLCM dihitung)
PCA	Variansi informasi yang dipertahankan	95%
PCA	Jumlah principal component hasil reduksi	5 komponen utama (konsisten pada ketiga skenario pembagian data)
SVM	Jenis kernel	Linear
SVM	Nilai parameter C	0,1
SVM	Parameter kernel (γ, degree, dsb.)	gamma = scale (default scikit-learn; parameter γ tidak berpengaruh pada kernel linear)
SVM multi-kelas	Skema Error-Correcting Output Codes (ECOC)	One-vs-One (setara dengan skema default fitcecoc MATLAB untuk kasus 3 kelas), diimplementasikan melalui OneVsOneClassifier
Pembagian dataset	Metode pembagian data latih-uji	Menggunakan pembagian data latih-uji yang telah disediakan per skenario (folder terpisah), tanpa pengacakan ulang
Pembagian dataset	Random seed (untuk PCA, SVM, dan validasi internal)	42

HASIL DAN PEMBAHASAN

Ekstraksi Fitur Dan Reduksi Dimensi

Proses ekstraksi fitur dilakukan melalui dua tahap utama. Pertama, ekstraksi fitur warna menggunakan model *Red Green Blue* (RGB) untuk mengambil proporsi dan distribusi intensitas warna dominan dan citra. Kedua, ekstraksi fitur tekstur menggunakan *Gray Level Cooccurrence Matrix* (GLCM) untuk mengekstrak empat parameter statistik utama, yaitu contrast, correlation, homogeneity, dan energy.



Gambar 2. Visualisasi Ekstraksi Fitur GLCM dan RGB

Seluruh fitur visual yang diperoleh dari RGB dan GLCM kemudian digabungkan dan diolah menggunakan *Principal Component Analysis* (PCA). Penerapan PCA bertujuan untuk mereduksi dimensi data citra, menekan kompleksitas komputasi, serta menyaring fitur-fitur paling signifikan sebelum dimasukkan ke dalam algoritma *Support Vector Machine* (SVM).

Pengujian Dan Evaluasi Model

Pengujian dilakukan untuk mengukur performa sistem dalam mengklasifikasikan tiga spesies macan tutul (*African Leopard, Clouded Leopard*, dan *Snow Leopard*). Evaluasi dilakukan dengan membandingkan tiga skenario pembagian data latih dan data uji, yaitu 60:40, 70:30, dan 80:20. Kinerja model dinilai berdasarkan metrik *precision, recall, F1-score,* dan *overall accuracy*. Ketiga proporsi ini dipilih untuk menguji seberapa sensitif performa model terhadap

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ukuran data latih: skenario 60:40 dan 80:20 masing-masing merepresentasikan kondisi data latih paling sedikit dan paling banyak yang diuji, sedangkan 70:30 merupakan proporsi umum yang banyak digunakan pada penelitian klasifikasi citra sejenis. Hasil pengujian lengkap (jumlah citra uji, precision, recall, dan F1-score per kelas) untuk ketiga skenario disajikan pada Tabel 2.

Tabel 2. Hasil Pengujian Klasifikasi 3 Jenis Macan Tutul

Skenario	Kelas	CU	Benar	P(%)	R(%)	F1(%)
60:40	African Leopard	83	53	79,10	63,86	70,67
60:40	Clouded Leopard	78	66	83,54	84,62	84,08
60:40	Snow Leopard	84	81	81,82	96,43	88,52
70:30	African Leopard	56	46	80,70	82,14	81,42
70:30	Clouded Leopard	58	46	85,19	79,31	82,14
70:30	Snow Leopard	56	55	93,22	98,21	95,65
80:20	African Leopard	56	46	80,70	82,14	81,42
80:20	Clouded Leopard	58	46	85,19	79,31	82,14
80:20	Snow Leopard	43	42	91,30	97,67	94,38

Sebagai pelengkap, Tabel 3 menyajikan confusion matrix lengkap untuk skenario terbaik (70:30), menunjukkan sebaran prediksi model secara rinci antar ketiga kelas.

Tabel 3. Confusion Matrix Skenario 70:30 (Skenario Terbaik)

Aktual/Prediksi	African Leopard	Clouded Leopard	Snow Leopard
African Leopard	46	7	3
Clouded Leopard	11	46	1
Snow Leopard	0	1	55

Analisis Dan Pembahasan

Berdasarkan hasil eksperimen pada Tabel 2, skenario pembagian 70% data latih dan 30% data uji menghasilkan performa terbaik dengan tingkat akurasi keseluruhan sebesar 86,47% (147 dari 170 citra uji diklasifikasikan dengan benar), sedikit lebih tinggi dibandingkan skenario 80:20 (85,35%) dan skenario 60:40 (81,63%). Pembahasan mengenai kinerja pengujian, termasuk analisis penyebab kesalahan klasifikasi berdasarkan confusion matrix pada Tabel 3, dijabarkan sebagai berikut:

1. **Skenario 70:30 (Akurasi Terbaik: 86,47%):** Pembagian ini memberikan titik data latih yang relatif seimbang untuk pelatihan model SVM. Kelas *African Leopard* menghasilkan *recall* sebesar 82,14% (*precision* 80,70%, *F1-score* 81,42%), dengan sebagian besar kesalahan klasifikasi mengarah ke kelas *Clouded Leopard* akibat kemiripan corak tutul dan rentang warna bulu kedua spesies (lihat confusion matrix pada Tabel 3). *Clouded Leopard* mencatat *recall* 79,31% (*F1-score* 82,14%), di mana kesalahan prediksi juga didominasi kekeliruan menjadi *African Leopard*. Di sisi lain, *Snow Leopard* berhasil dikenali hampir sempurna (*recall* 98,21%, *F1-score* 95,65%) berkat karakteristik visual dan warna bulu yang jauh lebih terang dibandingkan dua kelas lainnya.
2. **Skenario 60:40 (Akurasi: 81,63%):** Performa pada skenario ini merupakan yang terendah di antara ketiga skenario. Kelas *African Leopard* mencatatkan *recall* terendah (63,86%; 53 dari 83 citra uji), dengan kesalahan klasifikasi didominasi oleh kekeliruan menjadi *Snow Leopard* (18 dari 30 kesalahan) diikuti *Clouded Leopard* (12 dari 30 kesalahan). *Clouded Leopard* mencapai *recall* 84,62% (*F1-score* 84,08%) dengan seluruh kesalahannya terklasifikasi sebagai *African Leopard*, sedangkan *Snow Leopard* kembali paling mudah dikenali dengan *recall* 96,43% (*F1-score* 88,52%).
3. **Skenario 80:20 (Akurasi: 85,35%):** Akurasi keseluruhan pada skenario ini (85,35%) sedikit lebih rendah dibandingkan skenario 70:30, namun lebih tinggi dibandingkan skenario 60:40. Pola kesalahan pada kelas *African Leopard* dan *Clouded Leopard* identik dengan skenario 70:30 (*recall* masing-masing 82,14% dan 79,31%), sedangkan *Snow Leopard* kembali dikenali hampir sempurna dengan *recall* 97,67% (42 dari 43 citra uji). Hal ini menunjukkan bahwa performa model relatif stabil pada proporsi data latih yang lebih besar, tanpa indikasi penurunan akurasi akibat *overfitting*.

Secara keseluruhan, integrasi fitur warna RGB dan tekstur GLCM yang disederhanakan oleh PCA terbukti membantu algoritma SVM memisahkan ketiga kelas macan tutul, dengan *Clouded Leopard* secara konsisten menjadi kelas paling sulit dibedakan pada seluruh skenario akibat variasi pola dan warna tutulnya yang lebih tinggi dibandingkan dua kelas lainnya, serta kerap tertukar dengan *African Leopard*. *Recall Snow Leopard* yang secara konsisten tinggi (96-98%) pada seluruh skenario sejalan dengan proses *background removal* yang diterapkan secara konsisten pada seluruh citra, baik data latih maupun data uji (lihat Bagian 3.2), sehingga performa tersebut diinterpretasikan berasal dari karakteristik morfologi dan warna bulu *Snow*

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Leopard yang cukup khas (jauh lebih terang), bukan dari kontras latar belakang lingkungan bersalju yang sudah dihilangkan pada tahap pra-pemrosesan.

Studi Ablasi, Perbandingan Model, dan Validasi Tambahan

Untuk menjawab sejauh mana masing-masing komponen fitur berkontribusi terhadap performa yang dilaporkan pada Bagian 4.3, serta untuk menempatkan hasil penelitian ini secara objektif terhadap metode lain, Tabel 4 menyajikan studi ablasi kontribusi komponen fitur sekaligus perbandingan dengan model pembanding lain, seluruhnya pada skenario 70:30 (602 citra) agar perbandingan adil.

Tabel 4. Studi Ablasi Komponen Fitur dan Perbandingan dengan Model Pembanding (Skenario 70:30)

Kombinasi Fitur / Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
RGB + SVM	71,76	71,59	71,82	71,68
GLCM + SVM	62,94	63,47	63,12	63,09
RGB + GLCM + SVM (tanpa PCA)	87,06	87,11	87,15	87,08
RGB + GLCM + PCA + SVM (diusulkan)	86,47	86,37	86,56	86,40
Random Forest	84,71	85,25	84,85	84,76
K-Nearest Neighbor (KNN)	88,24	88,77	88,24	88,38
Convolutional Neural Network (CNN)	80,59	80,98	80,64	80,75

Sebagaimana ditunjukkan pada Tabel 4, model K-Nearest Neighbor (KNN, k = 3) mencatatkan akurasi tertinggi (88,24%) dibandingkan seluruh model lain pada skenario 70:30, termasuk metode yang diusulkan (RGB + GLCM + PCA + SVM, akurasi 86,47%); varian tanpa PCA (RGB + GLCM + SVM, akurasi 87,06%) juga sedikit lebih tinggi dibandingkan dengan PCA, mengindikasikan bahwa reduksi dimensi melalui PCA pada dataset ini berpotensi menghilangkan sebagian variansi fitur yang masih cukup diskriminatif, meskipun tetap mempertahankan 95% variansi total. Random Forest (84,71%) dan Convolutional Neural Network yang dilatih dari awal tanpa transfer learning (80,59%) mencatatkan akurasi di bawah metode yang diusulkan. Kontribusi utama penelitian ini terletak pada kombinasi fitur RGB dan GLCM secara terintegrasi, yang secara konsisten mengungguli penggunaan salah satu fitur saja (RGB atau GLCM secara terpisah, lihat Tabel 4), sehingga integrasi kedua jenis fitur tersebut terbukti lebih berperan terhadap performa akhir dibandingkan pemilihan algoritma klasifikasi tertentu.

Untuk memastikan bahwa akurasi terbaik pada skenario 70:30 tidak semata-mata dipengaruhi oleh komposisi sampel dari satu kali pembagian data, dilakukan evaluasi *stratified 5-fold cross-validation* pada seluruh 602 citra skenario 70:30 (data latih dan data uji digabung) menggunakan konfigurasi SVM yang sama (kernel linear, C = 0,1, PCA 95% variansi), menghasilkan akurasi per fold berturut-turut sebesar 96,69%; 90,91%; 93,33%; 85,83%; dan 91,67% (rata-rata 91,69% ± standar deviasi 3,54%). Rentang variasi antar fold yang cukup lebar ini mengindikasikan bahwa performa model cukup sensitif terhadap komposisi data latih-uji pada pembagian tertentu, sehingga akurasi pada satu skenario tunggal (Tabel 2) sebaiknya diinterpretasikan bersama rentang variasi ini, bukan sebagai estimasi tunggal yang pasti.

KESIMPULAN DAN SARAN

Sistem klasifikasi tiga spesies macan tutul (*African Leopard*, *Clouded Leopard*, dan *Snow Leopard*) berbasis kombinasi fitur tekstur GLCM, fitur warna RGB, reduksi dimensi PCA, dan algoritma klasifikasi SVM berhasil dikembangkan dan dievaluasi menggunakan dataset sebagaimana disediakan pada masing-masing folder skenario, dengan performa terbaik pada skenario 70:30 (akurasi 86,47%, precision makro 86,37%, recall makro 86,56%, F1-score makro 86,40%). *Snow Leopard* secara konsisten paling mudah dikenali (*recall* 96-98% pada seluruh skenario), sedangkan *Clouded Leopard* paling sulit dibedakan dan kerap tertukar dengan *African Leopard*. Studi ablasi menunjukkan bahwa integrasi fitur RGB dan GLCM secara konsisten mengungguli penggunaan salah satu fitur saja, sementara K-Nearest Neighbor (88,24%) dan varian tanpa PCA (87,06%) sedikit mengungguli metode yang diusulkan pada skenario 70:30, dan validasi 5-fold cross-validation menunjukkan variasi akurasi antar fold yang cukup besar (85,83%-96,69%; rata-rata 91,69% ± 3,54%), mengindikasikan bahwa performa model cukup sensitif terhadap komposisi data latih-uji pada pembagian tertentu. Untuk pengembangan penelitian selanjutnya, disarankan untuk memperluas variasi dataset dengan memperhitungkan faktor variasi pencahayaan dan latar belakang citra jebakan kamera (*camera trap*), mengeksplorasi ruang warna lain seperti HSV atau LAB untuk memperkuat deskripsi fitur warna,

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

membandingkan hasil dengan penelitian klasifikasi satwa liar lain menggunakan dataset yang lebih besar dan terstandardisasi, serta mempertimbangkan implementasi algoritma berbasis *Deep Learning* seperti *Convolutional Neural Network* (CNN) dengan dataset yang lebih besar.

DAFTAR PUSTAKA

Abdullah, D. M., & Abdulazeez, A. M. (2021). Machine learning applications based on SVM classification: A review. *Qubahan Academic Journal*, *1*(2), 81–90. https://doi.org/10.48161/qaj.v1n2a50

Bohnett, E., Holmberg, J., Poya Faryabi, S., An, L., Ahmad, B., Rashid, W., & Ostrowski, S. (2023). Comparison of two individual identification algorithms for snow leopards (*Panthera uncia*) after automated detection. *Ecological Informatics*, *77*, Article 102214. https://doi.org/10.1016/j.ecoinf.2023.102214

Khomiak, O., & Benndorf, J. (2021). Image segmentation methods for quick characterization of ore chip using RGB images. *IOP Conference Series: Earth and Environmental Science*, *942*(1), Article 012033. https://doi.org/10.1088/1755-1315/942/1/012033

Kızrak Yurtsever, A., Arıkan, A., & Kaçar, C. (2023). Gray-level co-occurrence matrix uniformity correction algorithm in positron emission tomographic image: A phantom study. *Applied Sciences*, *13*(10), Article 6036. https://doi.org/10.3390/app13106036

Lu, R., Cheng, Z., Zhou, K., & Wu, Y. (2024). Optimized image classification techniques: Comparing orthogonal basis and logistic regression with entropy and PCA in cat and dog image recognition. In *International Conference on Optics, Electronics, and Communication Engineering (OECE 2024)* (Vol. 13395, Article 133953U). SPIE.

Muhathir, M., Santoso, M. H., & Larasati, D. (2021). Wayang image classification using SVM method and GLCM feature extraction. *Journal of Informatics and Telecommunication Engineering*, *4*(2), 373–382. https://doi.org/10.31289/jite.v4i2.4524

Ovirianti, N. H., Zarlis, M., & Mawengkang, H. (2022). Support vector machine using a classification algorithm. *SinkrOn*, *6*(3), 2103–2107. https://doi.org/10.33395/sinkron.v7i3.11597

Pariyandani, A., Wanti, E. P., & Muhathir, M. (2020). Analysis of the Naive Bayes method in classifying formalin fish images using GLCM feature extraction. *Journal of Computer Science, Information Technology and Telecommunication Engineering*, *1*(2), 120–128. https://doi.org/10.30596/jcositte.v1i2.5171

Teng, D., Li, F., & Zhang, W. (2021). Using comprehensive machine-learning models to classify complex morphological characters. *Ecology and Evolution*, *11*(15), 10421–10431. https://doi.org/10.1002/ece3.7845

Tian, R., Sun, G., Liu, X., & Zheng, B. (2021). Sobel edge detection based on weighted nuclear norm minimization image denoising. *Electronics*, *10*(6). https://doi.org/10.3390/electronics10060655

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.