

Identifikasi Jenis Ras pada Kucing Menggunakan Algoritma Support Vector Machine (SVM)

Prasetyo Defantara Hadi ^{a,1,*}, Diah Arie Widhining K ^{a,2}, Farrady Alif Fiolana ^{a,3}

^a Universitas Islam Kadiri, Jalan Sersan Suharmaji No.38, Kediri, Indonesia

¹ prasetyodefan@gmail.com*; ² diahariwkw@uniska-kediri.ac.id; ³ farradyalif@uniska-kediri.ac.id

* Penulis Koresponden

INFO ARTIKEL

Histori Artikel

Pengajuan 2023-09-04

Diperbaiki 2023-09-08

Diterima 2023-12-30

Kata Kunci

Ras Kucing

Klasifikasi

Machine Learning

Support Vector Machine

ABSTRAK

Kucing adalah mamalia karnivora kecil yang dikenal sebagai satu-satunya spesies jinak dalam keluarga *Felidae*. Di Indonesia, beberapa jenis ras kucing populer sebagai hewan peliharaan, antara lain Bengal, Ragdoll, Russian Blue, Siamese, dan Persia. Namun, keberagaman ras kucing ini seringkali membuat pemilik sulit mengidentifikasi jenis ras yang dimiliki oleh kucing mereka. Penelitian ini bertujuan untuk melakukan klasifikasi pada lima jenis ras kucing tersebut menggunakan metode *Support Vector Machine* (SVM) dengan kernel *Radial Basis Function* (RBF). Data yang digunakan terdiri dari tiga jenis dataset, yaitu data latih sebanyak 1400 sampel, data validasi sebanyak 600 sampel dan data uji sebanyak 250 sampel. Hasil penelitian menunjukkan bahwa model yang dibangun menggunakan metode SVM dengan kernel RBF berhasil mencapai tingkat akurasi sebesar 86%, presisi sebesar 87%, recall sebesar 86%, dan f1 score sebesar 86%. Hasil tersebut menandakan bahwa model klasifikasi ini mampu melakukan prediksi dengan tingkat keakuratan yang baik.

Ini adalah artikel akses terbuka di bawah lisensi [CC-BY-SA](#).



1. Pendahuluan

Selain berfungsi untuk mengisi waktu luang, peliharaan juga dapat menemani manusia di berbagai aktivitas, salah satu hewan yang paling banyak dipelihara adalah kucing. Kucing memiliki nama ilmiah *Felis silvestris catus* atau *Felis catus*. Kucing merupakan mamalia karnivora kecil dimana mereka satu-satunya spesies jinak dalam keluarga *Felidae*. Pada umumnya ada beberapa jenis ras yang sering dijadikan peliharaan orang Indonesia yaitu Bengal, Ragdoll, Russian Blue, Siamese dan Persia. Saat ini terdapat 135 ras kucing yang terdata, karena banyaknya jenis ras pada kucing bisa saja membuat pemilik sulit untuk mengidentifikasi jenis ras pada kucingnya. Untuk itu, perlu adanya metode klasifikasi untuk mengatasi hal tersebut, identifikasi jenis ras bermanfaat dalam memberikan perawatan yang tepat untuk kucing, karena setiap jenis ras memiliki kebutuhan kesehatan yang berbeda-beda. Penelitian tentang pengenalan gambar mencakup analisis visual dan pengklasifikasian objek dalam gambar digital. Dalam pengenalan citra, metode ekstraksi fitur dengan *Local Binary Patterns* (LBP)[1][2] dan klasifikasi dengan metode *Support Vector Machine* (SVM) [3][4][5][6] adalah dua pendekatan yang umum digunakan. Metode tersebut digunakan dalam penelitian untuk membangun sistem pengenalan ras kucing dengan SVM.

Untuk mencapai tujuan ini, beberapa langkah harus dilakukan. Pertama, dataset yang terdiri dari foto ras kucing yang sudah dilabeli menggunakan platform seperti Roboflow. Setelah dataset lengkap, proses pra-pemrosesan gambar yang melibatkan *resize* citra, dan *unsharp masking*. Model SVM dilatih dengan data fitur dari LBP[7], dan hasil pelatihan disimpan dalam format *pickle* untuk digunakan di masa depan, metode menyimpan model dengan *pickle* ini mempermudah mengklasifikasikan gambar ras kucing yang belum dikenal tanpa perlu melakukan proses pelatihan model ulang. Untuk mengetahui berapa nilai tingkat akurasi dari data yang telah dilakukan training dan testing menggunakan algoritma SVM maka dilakukan pengecekan menggunakan *confussion Matrix*[4][8] dan *Receiver Operating Characteristic* (ROC)[9][10].

2. Metode penelitian

Penelitian ini melibatkan serangkaian tahapan yang perlu dilakukan guna mencapai hasil akhir yang diinginkan. Tahapan – tahapan tersebut dapat diuraikan lebih lanjut sebagai berikut :

1. Pengumpulan Data

Pada tahap pertama ini bertujuan untuk menjelaskan bagaimana proses pembuatan *dataset* yang akan digunakan dalam penelitian ini

2. Pengolahan Data

Pada tahapan kedua ini menjelaskan bagaimana melakukan *pre - process* data pada *dataset* jenis ras kucing sebelum memasuki proses klasifikasi

3. Pemodelan SVM

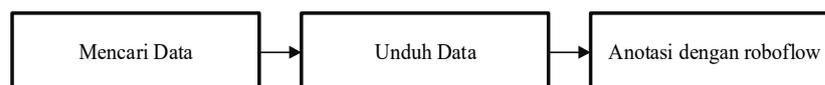
Pada tahapan ketiga ini bertujuan untuk melakukan pembuatan model svm, training model, dan proses menyimpan model

4. Pengujian Model

Pada tahapan terakhir ini berisi mengenai tahapan tahapan pengujian yang akan dilakukan dalam penelitian ini

2.1 Pengumpulan Data

Pada penelitian ini menggunakan *dataset* yang dibuat menggunakan sejumlah citra yang diperoleh dari Google Image, Bing Image, dan Kaggle. Selanjutnya, citra tersebut akan diberi label dan dilakukan proses anotasi dengan menentukan kotak pembatas (*bounding box*).



Gambar 1. Alur Pengumpulan Data

Tahapan-tahapan dari Teknik Pengolahan Data Citra, yaitu:

- a. Mencari Data

Pada tahap ini, dilakukan pencarian data jenis ras kucing menggunakan beberapa mesin pencari seperti Google Image dan Bing Image, serta melalui website penyedia dataset Kaggle. Dengan memanfaatkan berbagai sumber data ini, akan memungkinkan untuk memperoleh dataset yang bervariasi mengenai jenis ras kucing yang diinginkan.

- b. Unduh Data

Pada tahapan ini, dilakukan pengunduhan data citra dari internet ke penyimpanan lokal. Setelah data citra berhasil diunduh, langkah selanjutnya adalah memproses data menggunakan website Roboflow.

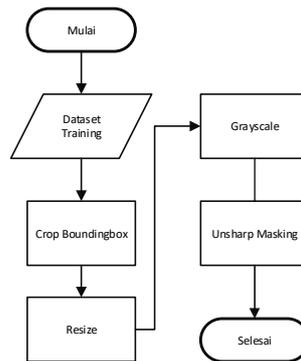
- c. Anotasi dengan roboflow

Pada proses ini, terlebih dahulu dilakukan pengunggahan citra dari penyimpanan lokal ke website roboflow. Setelah itu, citra-citra tersebut akan dilakukan pelabelan dan penentuan bounding box. Tahap selanjutnya, citra-citra tersebut diekspor dalam bentuk pascal voc (citra dan file xml) untuk dilakukan proses selanjutnya.

2.2 Pengolahan Data

Pada tahapan ini, dilakukan pengolahan data pada dataset jenis ras kucing yang telah dikumpulkan. Pengolahan data bertujuan untuk mempersiapkan data sebelum memasuki proses klasifikasi menggunakan metode SVM. Berikut adalah penjelasan lebih lanjut mengenai pengolahan data:

1. Pre-Process Data



Gambar 2. Alur Preprocess

Pada sub-tahapan ini, dilakukan serangkaian proses *pre-process* untuk mempersiapkan dataset jenis ras kucing sebelum dilakukan proses klasifikasi menggunakan metode SVM. Tahapan *pre-processing* meliputi *input dataset*, *crop bounding box*, *resize citra*, *grayscale*, dan *unsharp masking* yang akan dijelaskan lebih lanjut :

- *Crop` bounding box* :

Gambar yang telah diberi anotasi dan label dilakukan *cropping* dengan menggunakan data *bounding box* yang diperoleh dari *file xml* untuk memfokuskan pada objek kucing dalam gambar.

- *Resize citra*:

Gambar - gambar yang telah dilakukan proses *crop* kemudian dilakukan *resize* agar memiliki dimensi yang lebih kecil. Hal ini bertujuan untuk mengurangi kompleksitas komputasi dan mempercepat proses klasifikasi.

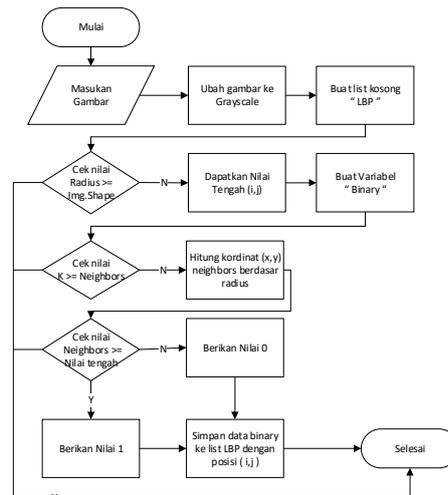
- *Convert ke grayscale* :

Gambar - gambar diubah dari citra berwarna menjadi citra *grayscale*. Menggunakan citra *grayscale* dapat mengurangi dimensi data dan menghilangkan informasi warna yang tidak relevan dalam proses klasifikasi ras kucing.

- *Unsharp Masking*:

Menggunakan teknik *unsharp masking* untuk memperjelas *detail* pada gambar. *Unsharp masking* menghasilkan gambar yang tajam dengan memperkuat detail pada tepi objek.

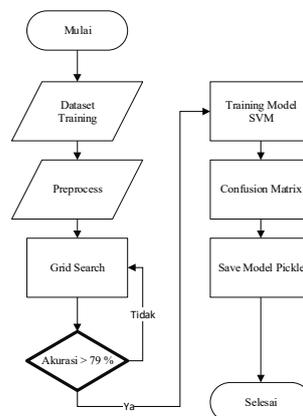
2. Ekstraksi Data



Gambar 3. Flowchart Ekstraksi LBP

Pada tahapan ini, dilakukan ekstraksi fitur dari gambar-gambar yang telah melewati tahapan *pre-processing* menggunakan *Local Binary Pattern* (LBP). Ekstraksi fitur bertujuan untuk mengubah data gambar menjadi representasi yang lebih sederhana dan informatif yang dapat digunakan oleh metode klasifikasi SVM. Proses ekstraksi fitur ini menghasilkan vektor fitur yang mewakili setiap gambar dalam dataset. Vektor fitur ini akan digunakan sebagai input dalam proses klasifikasi menggunakan metode SVM.

3. Pemodelan SVM



Gambar 4. Alur Pemodelan SVM

Pada tahapan ini, dilakukan pembuatan model SVM untuk melakukan klasifikasi ras kucing berdasarkan vektor fitur yang telah diekstraksi dari gambar-gambar dalam dataset. Proses pemodelan SVM meliputi:

1. Pembuatan model SVM:

Model SVM dibentuk dengan menggunakan vektor fitur sebagai input dan label jenis ras kucing sebagai target. Model SVM bertujuan untuk mempelajari pola dan hubungan antara vektor fitur dan label kelas dalam dataset.

2. Tuning model SVM:

Setelah pembuatan model awal, dilakukan proses tuning untuk memperoleh konfigurasi model SVM yang optimal. Beberapa parameter yang dapat dituning antara lain:

a. Pemilihan parameter C:

parameter yang mengatur *trade-off* antara margin dan kesalahan klasifikasi pada model SVM. Nilai yang lebih tinggi akan memberikan penekanan yang lebih besar pada kesalahan klasifikasi pada data pelatihan, sedangkan nilai yang lebih rendah akan memberikan penekanan yang lebih besar pada margin yang lebih lebar.

b. Pemilihan kernel:

SVM menggunakan kernel untuk memetakan data ke dalam ruang fitur yang lebih tinggi dimensi, di mana dapat dilakukan pemisahan linier. Beberapa jenis kernel yang umum digunakan adalah kernel linier, kernel polinomial, dan kernel fungsi basis radial (RBF). Pemilihan kernel yang tepat akan bergantung pada karakteristik data dan kompleksitas masalah klasifikasi.

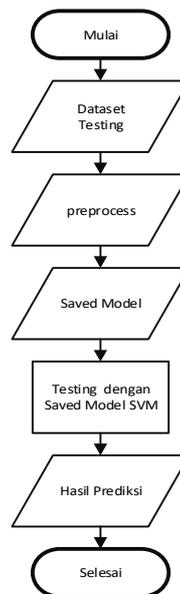
Proses *tuning svm* pada penelitian ini menggunakan metode *grid search* dengan memberikan *input* parameter (*C* dan *Kernel*) yang akan digunakan untuk klasifikasi *svm*

3. Proses penyimpanan model:

Setelah model SVM dilakukan *tuning*, dilakukan proses penyimpanan model untuk penggunaan selanjutnya. Model SVM yang telah dilatih akan disimpan dalam format file *pickle* agar dapat digunakan untuk melakukan klasifikasi pada data uji atau data baru

Dalam tahapan ini, penting untuk melakukan evaluasi model SVM untuk memastikan kinerja yang baik sebelum memasuki tahapan pengujian. Evaluasi model SVM dapat dilakukan dengan menggunakan metrik evaluasi yang relevan, seperti *akurasi*, *presisi*, *recall*, *F1-score*, atau *confusion matrix*.

4. Pengujian Model



Gambar 5. Alur Pengujian Model

Pada tahapan ini, dilakukan pengujian model SVM yang telah dilatih pada dataset pengujian. Pengujian model bertujuan untuk mengevaluasi kinerja model SVM dalam melakukan klasifikasi ras kucing pada data yang belum pernah dilihat sebelumnya. Berikut adalah penjelasan lebih lanjut mengenai tahapan pengujian model :

1. Dataset pengujian:

Dataset pengujian yang digunakan terdiri dari 5 kelas ras kucing yang berbeda. Totalnya terdapat 1400 data latih, 600 data validasi, dan 250 data uji. Dataset ini bertujuan untuk menguji kinerja model SVM dalam melakukan klasifikasi ras kucing pada data yang belum pernah dilihat sebelumnya.. Pemisahan dataset dilakukan secara acak dengan memastikan bahwa

2. Hasil pengujian model SVM:

Setelah melakukan pengujian model SVM pada dataset pengujian, diperoleh hasil berupa metrik evaluasi seperti *akurasi*, *presisi*, *recall*, dan *F1-score*. Metrik-metrik ini memberikan gambaran tentang kinerja model dalam melakukan klasifikasi ras kucing. Selain itu, juga didapatkan nilai support untuk setiap kelas ras kucing, yang menunjukkan jumlah data yang termasuk dalam setiap kelas.

3. Hasil dan Analisis

Pada bab ini, akan dibahas hasil dari pra-pemrosesan data, ekstraksi fitur, pemodelan SVM, dan pengujian model. Pra-pemrosesan data dilakukan untuk mempersiapkan dataset jenis ras kucing sebelum klasifikasi. Selanjutnya, fitur-fitur yang relevan diekstraksi untuk digunakan dalam SVM. Model SVM dibangun dan disesuaikan untuk mencapai performa yang optimal. Pengujian dilakukan untuk mengevaluasi performa model berdasarkan metrik evaluasi yang digunakan.

3.1 Hasil Pengolahan Data

Sebelum dilakukan proses klasifikasi dataset perlu diolah terlebih dahulu (*pre-process*) supaya didapati hasil akhir klasifikasi dengan tingkat akurasi yang baik. Pada penelitian ini, dilakukan pengolahan data atau *pre-processing* yang melibatkan beberapa langkah penting.



Gambar 6. Perbandingan Gambar Original dengan Crop

Pertama, dilakukan proses *cropping* untuk memfokuskan pada area yang paling relevan dalam gambar dan menghilangkan area yang tidak penting.



Gambar 7. Hasil Resize 32 x 32

Kemudian, dilakukan proses *resize* untuk mengubah ukuran dimensi gambar menjadi 32 x 32 piksel sesuai dengan kebutuhan analisis. *Resize* dilakukan untuk memperkecil ukuran gambar guna untuk mempercepat proses komputasi. Ukuran piksel tersebut dipilih setelah dilakukan beberapa eksperimen ukuran, dijelaskan pada tabel berikut :

Tabel 1. Percobaan ukuran piksel

Ukuran Piksel	Waktu Proses (Detik)	Akurasi
8	205.59	0.53
16	127.18	0.82
32	162.07	0.83
64	196.82	0.80
128	770.32	0.78
256	2721.31	0.75



Gambar 8. Hasil Grayscale

Langkah selanjutnya adalah mengkonversi gambar ke *grayscale*, yang mengubah gambar berwarna menjadi gambar dengan tingkat keabuan. Hal ini dilakukan untuk mengurangi kompleksitas data dan mempertahankan informasi penting dalam bentuk intensitas keabuan tunggal.



Gambar 9. Hasil Unsharp Masking

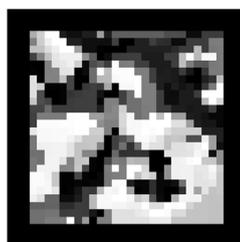
Kemudian, dilakukan proses *unsharp masking*, yang bertujuan untuk meningkatkan ketajaman gambar dengan meningkatkan kontras pada tepi objek. Dengan menghapus komponen halus atau tidak relevan dalam gambar, teknik unsharp masking dapat mempertajam detail penting dalam gambar.

Setelah melalui langkah-langkah preprocessing yang telah dijelaskan sebelumnya, data gambar yang telah diproses siap untuk dilakukan ekstraksi fitur menggunakan metode *Local Binary Patterns* (LBP).



Gambar 10. Hasil Ekstraksi LBP

Gambar diatas merupakan data representasi biner yang dihasilkan dari proses pengolahan gambar menggunakan ekstraksi LBP, apabila kode numerik diatas diterjemahkan atau divisualisasikan dalam bentuk gambar, maka akan menghasilkan gambar sebagai berikut



Gambar 11. Hasil Visualisasi Ekstraksi LBP

Dalam penelitian ini, ekstraksi fitur LBP menjadi langkah penting dalam memperoleh representasi biner yang merepresentasikan pola tekstur pada gambar yang telah diproses sebelumnya. Fitur LBP ini kemudian dapat digunakan sebagai masukan dalam algoritma klasifikasi SVM.

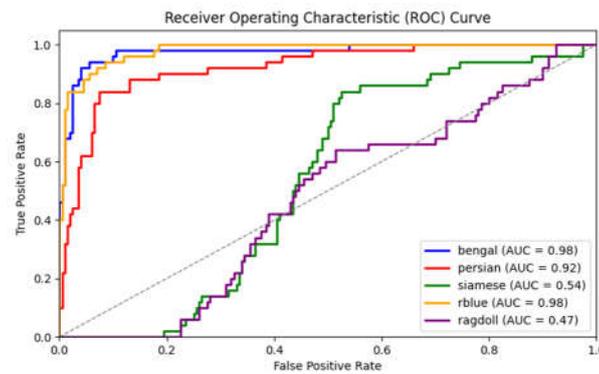
3.2 Hasil Klasifikasi

Setelah melakukan penyesuaian model berdasarkan hasil grid search dan menemukan parameter terbaik, langkah selanjutnya dalam proses ini adalah menyimpan model yang telah dioptimalkan. Untuk melakukan hal tersebut, model klasifikasi yang telah disesuaikan dengan parameter $C : 10$ dan kernel *radial basis function* (RBF) disimpan menggunakan format pickle. Dengan menggunakan format pickle, model dapat dilakukan *serialize* dan disimpan dalam bentuk file yang dapat dengan mudah diambil kembali saat diperlukan. Proses penyimpanan model dengan format *pickle* memungkinkan untuk menghemat waktu dan usaha yang diperlukan untuk melakukan penyesuaian ulang pada setiap penggunaan model, sehingga dapat dengan cepat dan efisien memanfaatkan model yang telah dioptimalkan untuk melakukan prediksi pada data baru.

Tabel 2. Confusion Matrix Pengujian Model

True label	Predicted label					Confusion Matriks		
	bengal	persian	siamese	rblue	ragdoll	Precision	Recall	F1-Score
bengal	46	3	0	1	0	0.82	0.92	0.87
persian	2	42	2	2	2			
siamese	5	2	42	1	0			
rblue	3	4	1	42	0			
ragdoll	0	6	1	0	43			
Rata - Rata						0.87	0.86	0.86
Akurasi : 86%								

Hasil dari eksperimen pengujian SVM yang melibatkan lima jenis ras kucing menunjukkan bahwa tingkat akurasi mencapai 86%, *presisi* 87%, *recall* 86%, *f1 score* 86%.



Gambar 12. Grafik ROC Pengujian Model

Gambar grafik dari hasil penelitian menunjukkan bahwa model memiliki AUC yang tinggi untuk kelas "bengal" (0.98) dan "russian blue" (0.98), serta kelas "persian" dengan AUC 0.92. Namun, terdapat kinerja yang kurang memuaskan untuk kelas "siamese" dengan AUC 0.54 dan kelas "ragdoll" dengan AUC 0.42. AUC yang mendekati 0.5 menunjukkan bahwa model memiliki kinerja yang hampir sebanding dengan pemilihan acak. Dengan demikian, hasil penelitian ini memberikan pemahaman yang lebih mendalam tentang kemampuan model dalam membedakan kelas-kelas yang berbeda.

Pada penelitian ini dilakukan eksperimen menggunakan kombinasi metode untuk melakukan klasifikasi yaitu menggunakan kombinasi *Pyramid Histogram of Adaptive Gradient* (PHOG) + *Support Vector Machine* (SVM) dan *Local Binary Pattern* (LBP) + *Support Vector Machine* (SVM) adapun hasil dari eksperimen tersebut ditampilkan pada table dibawah

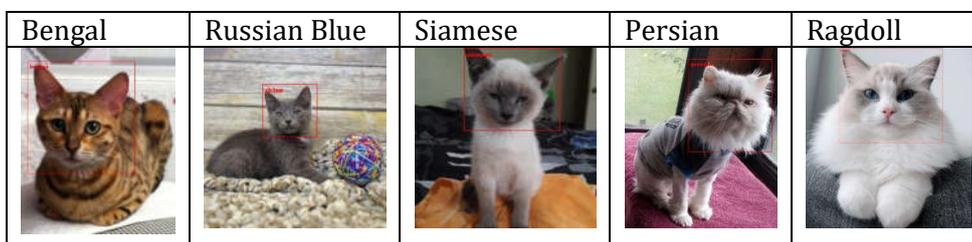
Prasetyo Defantara Hadi (Identifikasi Jenis Ras Pada Kucing Menggunakan Algoritma Support Vector Machine (SVM))

Tabel 3. Eksperimen Kombinasi Metode

METODE	LBP + SVM	PHOG + SVM
Accuracy Score	0.83	0.58
Precision Score	0.84	0.58
Recall Score	0.83	0.58
F1 Score	0.83	0.57

Tabel diatas merupakan data dari eksperimen percobaan kombinasi metode yang telah dilakukan, metode LBP + SVM menunjukkan hasil yang lebih baik dalam hal Precision Score, Recall Score, dan F1 Score, yaitu 0.84, 0.83, dan 0.83 masing-masing. Sementara itu, metode PHOG + SVM memiliki nilai yang lebih rendah, yaitu 0.58, 0.58, dan 0.57 masing-masing. Ini menunjukkan bahwa metode LBP + SVM memiliki kinerja yang lebih baik dalam mengklasifikasikan dan mengenali pola dalam citra dibandingkan dengan metode PHOG + SVM berdasarkan dataset yang dievaluasi.

Tabel 4. Hasil Deteksi



Pada tabel diatas ditunjukkan hasil prediksi dari model svm yang telah dibuat pada proses sebelumnya, citra yang telah dilakukan prediksi diberikan *boundingbox* pada bagian wajah dan diberikan *labeling* ras sesuai prediksi dari model

4. Kesimpulan

Berdasarkan pada proses pengujian, pengamatan dan analisis hasil yang telah dilakukan dapat disimpulkan bahwa implementasi klasifikasi jenis ras kucing menggunakan data fitur dari hasil ekstraksi *Local Binary Pattern* (LBP) yang kemudian diklasifikasi menggunakan *Support Vector Machine* (SVM) telah membuktikan keefektifan SVM dalam mengklasifikasikan ras kucing dengan tingkat akurasi yang cukup baik mencapai. Melalui eksperimen dengan SVM menggunakan kernel RBF dengan dataset berjumlah 2250 mampu mencapai akurasi 86%, recall 87%, presisi 86%, dan F1 Score 86% dalam pengujian prediksi. Penelitian ini akan memberikan kontribusi untuk pengembangan metode klasifikasi jenis ras kucing menggunakan *Support Vector Machine* (SVM), dan menjadikan landasan untuk penelitian selanjutnya dalam meningkatkan akurasi untuk memenuhi kebutuhan praktis dalam bidang pengolahan citra

Referensi

- [1] A. A. Bardekar dan V. M. Deshmukh, "An Empirical Study : Musical Influence on Face Using the Local Binary Pattern (LBP) Approach," 2012.
- [2] Y. Li, H. Tang, W. Xie, dan W. Luo, "Multidimensional Local Binary Pattern for Hyperspectral Image Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, hal. 1–13, 2022.
- [3] D. Ayon, "Machine Learning Algorithms : A Review," *Int. J. Comput. Sci. Inf. Technol.*, vol. 7, no. 3, hal. 1174–1179, 2016, doi: 10.21275/ART20203995.
- [4] R. R. Jaka Kusuma¹, Abwabul Jinan², Muhammad Zulkarnain Lubis³, Rubianto⁴, "Komparasi Algoritma Support Vector Machine Dan Naive Bayes Pada Klasifikasi Ras Prasetyo Defantara Hadi (Identifikasi Jenis Ras Pada Kucing Menggunakan Algoritma Support Vector Machine (SVM))

- Kucing," *J. Generic*, vol. 14, no. 1, hal. 8–12, 2022, [Daring]. Tersedia pada: <http://generic.ilkom.unsri.ac.id/index.php/generic/article/view/122>
- [5] D. A. Pisner dan D. M. Schnyer, *Support vector machine*. Elsevier Inc., 2019. doi: 10.1016/B978-0-12-815739-8.00006-7.
- [6] J. Mase, M. T. Furqon, dan B. Rahayudi, "Penerapan Algoritme Support Vector Machine (SVM) Pada Pengklasifikasian Penyakit Kucing," *J. Pengembangan Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 10, hal. 3648–3654, 2018.
- [7] I. Muslihah dan M. Muqorobin, "Texture Characteristic of Local Binary Pattern on Face Recognition with Probabilistic Linear Discriminant Analysis," *Int. J. Coop. Inf. Syst.*, vol. 1, hal. 22–26, 2020.
- [8] A. Suryadibrata dan S. D. Salim, "Klasifikasi Anjing dan Kucing menggunakan Algoritma Linear Discriminant Analysis dan Support Vector Machine," *Ultim. J. Tek. Inform.*, vol. 11, no. 1, hal. 46–51, 2019, doi: 10.31937/ti.v11i1.1076.
- [9] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognit.*, vol. 30, no. 7, hal. 1145–1159, 1997, doi: 10.1016/S0031-3203(96)00142-2.
- [10] Fernanda Januar Pratama, Wikky Fawwaz Al Maki, dan Febryanti Sthevanie, "Big Cats Classification Based on Body Covering," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 5, hal. 984–991, 2021, doi: 10.29207/resti.v5i5.3328.