

## Model Klasifikasi Jenis Hewan Dengan SVM, KNN, *Logistic Regression* Menggunakan *Pre-Trained VGG 16*

Jaka Tirta Samudra<sup>1</sup>, Rika Rosnelly<sup>2</sup>, Zakarias Situmorang<sup>3</sup>, Puji Sari Ramadhan<sup>4</sup>

<sup>1,2,3</sup>Magister Ilmu Komputer, Universitas Potensi Utama, Kota Medan, Indonesia

<sup>4</sup>Sistem Informasi, STMIK Triguna Dharma, Kota Medan, Indonesia

Email: <sup>1</sup>jakartirta135@gmail.com, <sup>2</sup>rikarosnelly@gmail.com, <sup>3</sup>zakarias65@yahoo.com, <sup>4</sup>pujisariramadhan@gmail.com

Email Penulis Korespondensi: jakartirta135@gmail.com

### Article History:

Received Aug 01<sup>th</sup>, 2023

Revised Aug 10<sup>th</sup>, 2023

Accepted Aug 25<sup>th</sup>, 2023

### Abstrak

Proses komputasi serta *filtering* pada komputer untuk melaksanakan suatu tugas yang diinginkan untuk melakukan kegiatan tertentu tentunya tidak lepas dari sebuah metode pada pembelajaran. Dalam proses pembelajaran tersebut ada beberapa dari berbagai metode dapat dilakukan untuk dapat memenuhi periode training dan uji tersebut untuk memberikan komputer suatu keahlian tertentu. Salah satu cara tujuan untuk melakukan penunjang pada periode tersebut adalah dengan menggunakan algoritma *support vector machine*, *k-nearest neighbor*, dan *logistic regression*. Dimana pada algoritma ini mampu memuat keseluruhan skala informasi klasifikasi objek tanpa kehilangan dari pengetahuan keakuratannya. Tujuan dari penelitian ini adalah memberikan pada komputer dalam menggali kemampuan untuk mengenali jenis binatang dan memprediksi jenis binatang berdasarkan gambar yang dimasukkan. Penelitian ini juga bertujuan untuk menilai keakuratan hasil training metode pembelajaran dibandingkan dengan hasil keluaran dari pembelajaran. Metode yang digunakan dalam penelitian ini adalah mentraining secara komputasi, sejumlah gambar dari beberapa hewan yang memiliki 10 jenis hewan yang dekat kepada manusia salah satunya hewan ternak, peliharaan, dan buas. Kemudian test akan dilakukan dengan cara yang sama setelah melalui tahapan *konvulasi training*. Hasil dari penelitian ini keakuratan hasil training mencapai 84%.

**Kata Kunci :** Jenis Hewan, Klasifikasi, Gambar, Prediksi, Akurasi.

### Abstract

*The process of computing and filtering on a computer to carry out a task that is desired to carry out certain activities cannot be separated from a method of learning. In the learning process there are several of various methods that can be done to be able to fulfill the training and test period to give the computer a certain expertise. One way to support this period is to use support vector machine, k-nearest neighbor, and logistic regression algorithms. Where in this algorithm is able to load the entire scale of object classification information without losing knowledge of its accuracy. The purpose of this study is to provide the computer with the ability to recognize the type of animal and predict the type of animal based on the image entered. This study also aims to assess the accuracy of the training results of learning methods compared to the outputs of learning. The method used in this study is computational training, a number of images of several animals that have 10 types of animals that are close to humans, one of which is livestock, pets, and wild animals. Then the test will be carried out in the same way after going through the convolution training stages. The results of this study the accuracy of the training results reached 84%.*

**Keyword :** *Types of Animals, Classification, Images, Predictions, Accuracy.*

## 1. PENDAHULUAN

Hewan merupakan salah satu makhluk hidup yang memiliki revolusi yang cukup banyak bahkan sudah ada sejak manusia belum mulai berevolusi yang dimana bisa dikatakan dari zaman dinosaurus hingga saat ini zaman perabadan. Bahkan hewan juga merupakan makhluk yang sangat dekat pada manusia dari beberapa hewan sesuai dengan jenisnya seperti hewan ternak, hewan peliharaan, dan bahkan hewan liar pun bisa dekat sama manusia sesuai bagaimana cara merawat dan berperilaku hewan tersebut [1], [2].

Pada kiatannya seperti contohnya kepada hewan seperti apa saja yang dekat kepada manusia seperti hewan ternak yang terdiri seperti sapi, kerbau, kambing, domba, ayam, dan lainnya. Pada contoh lainnya untuk hewan peliharaan yang sangat dekat kepada manusia seperti kucing dan anjing pada umumnya yang sering di pelihara oleh manusia dan untuk jenis yang tidak umum atau ada ketertarikan kepada hewan seperti tupai, laba-laba, musang, ular, jenis serangga bahkan reptile seperti buaya, harimau, singa, dan lainnya [3]. Teruntuk hewan lainnya ada juga hanya saja sebagai hewan eksotis yang diperlihara pada kalangan tertentu seperti pada pemerintah dan penguasa yang memiliki tujuan menciptakan pemberitahuan dan wawasan kepada hewan dengan maksud membuat taman hewan yang berisikan beranekaragaman hewan yang terdiri dari seperti banteng, kuda, monyet, ular, kupu-kupu, gajah, jerapa, landak, burung, dan lainnya [4].

Beberapa karestik hewan yang ada yang memiliki banyak ciri serta kepada karakteristik sehingga jenis hewan ini begitu sulit untuk dibedakan apalagi bagi sedikit orang yang baru tau hewan apa dan juga anak kecil seabgai bahan ajar dalam pengenalan nama hewan berdasarkan jenisnya [5]. Beberapa karakteristik pada hewan mamalia dapat menjadi fitur sebagai pembeda dalam klasifikasi. Citra jejak kaki yang merepresentasikan jumlah dan ukuran gumpalan bentuk kaki telah digunakan sebagai fitur pada klasifikasi jenis hewan [6]. Tidak hanya jejak kaki yang dapat menjadi fitur klasifikasi hewan, citra wajah hewan tampak depan juga mampu menjadi fitur utama. Citra wajah hewan tampak depan menjadi pilihan fitur klasifikasi karena dapat langsung mewakili dalam representasi hewan yang dimaksud. Pada penelitian [7], [8], [9], citra wajah hewan telah digunakan sebagai fitur dalam klasifikasi hewan. Beberapa penggunaan metode ekstraksi fitur dan metode klasifikasi telah digunakan untuk klasifikasi hewan. Salah satu metode klasifikasi yang telah digunakan adalah *support vector machine*, *k-nearest neighbor*, dan *logistic regression* [10]. Metode *support vector machine*, *k-nearest neighbor*, dan *logistic regression* telah banyak digunakan pada kasus klasifikasi seperti pada citra wajah [11], daun [12], batik, dan khususnya pada jenis hewan [13], [14].

Penelitian yang menggunakan metode Logistic regression, k-nearest neighbor, ANN, SVM dengan pembagian data 90%:10% dan akurasi terbaik sebesar 86% didapatkan ketika menggunakan metode SVM. Penelitian selanjutnya oleh beberapa penelitian tersebut menunjukkan bahwa penelitian pembelajaran mesin melakukan pembagian data latih dan data uji bervariasi. Penelitian [15] menghasilkan akurasi menggunakan metode SVM sebesar 92.30%, Logistic Regression sebesar 93.40%, dan ANN sebesar 92,30% dapat disimpulkan hasil akurasi dari ketiga metode tersebut tidak terlalu signifikan dengan pengolahan pembagian data latih dan uji tidak bervariasi [16]. Berdasarkan permasalahan tersebut, penelitian selanjutnya akan melakukan eksperimen untuk membandingkan metode SVM, Logistic Regression, dan ANN berdasarkan pembagian data testing dan data training bervariasi bertujuan untuk mendapatkan tingkat akurasi pembagian data yang lebih baik dengan menggunakan tools pengujian yang berbeda dari [16] penelitian sebelumnya. Pemilihan ketiga metode bertujuan untuk mengetahui metode terbaik dari ketiga metode tersebut [17].

Fitur prediksi jenis gambar tersebut nantinya akan menjadi salah satu alat penting yang sangat dibutuhkan bagi dunia penelitian karena permasalahan yang ditemukan oleh penelitian terdahulu adalah jumlah binatang yang ada di dunia ini sangat banyak, sebut saja sekitar 8,7 Juta spesies [5], dimana satu spesies ada lebih dari beberapa ribu hingga beberapa juta jenis [18].

## 2. METODOLOGI PENELITIAN

Pada penelitian ini, *dataset* yang digunakan adalah LHI-Animal-Faces. *Dataset* ini merupakan data citra wajah hewan tampak depan yang terdiri dari 10 jenis hewan yang mudah kita ketahui dan sangat dekat dengan kita. Sesuai dengan kebutuhan penelitian ini maka diambil 10 jenis hewan yang tiap jenisnya terdiri dari 50 citra dengan ukuran 150x150 *pixel*. Citra hewan mamalia tampak depan pada *dataset* dapat dilihat pada Gambar 1.

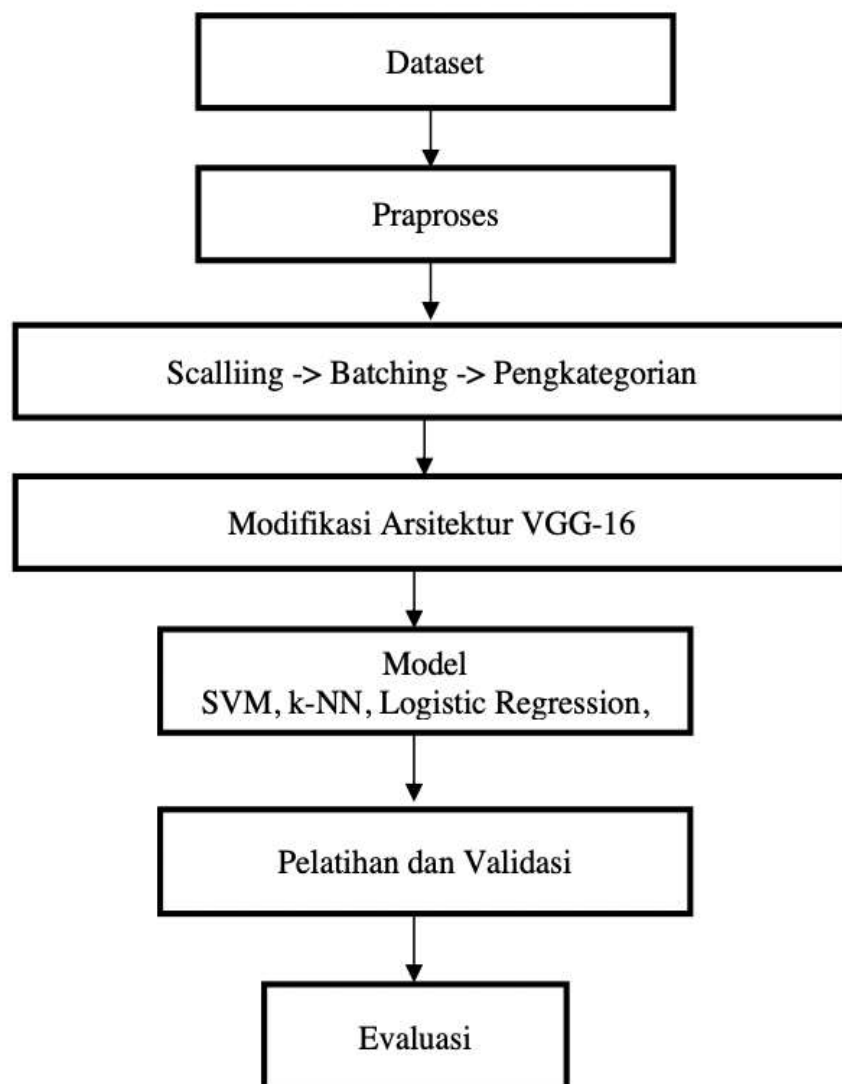


Gambar 1. Sampel Data Gambar Hewan

Kategori jenis keterangan hewan :

- Anjing disingkat (Aj).
- Ayam disingkat (Ay).
- Sapi disingkat (Sp).
- Domba disingkat (Db).
- Gajah disingkat (Gj).
- Kucing disingkat (Kg).
- Kuda disingkat (Kd).
- Kupu-kupu disingkat (Ku).
- Laba-laba disingkat (La).
- Tupai disingkat (Tp).

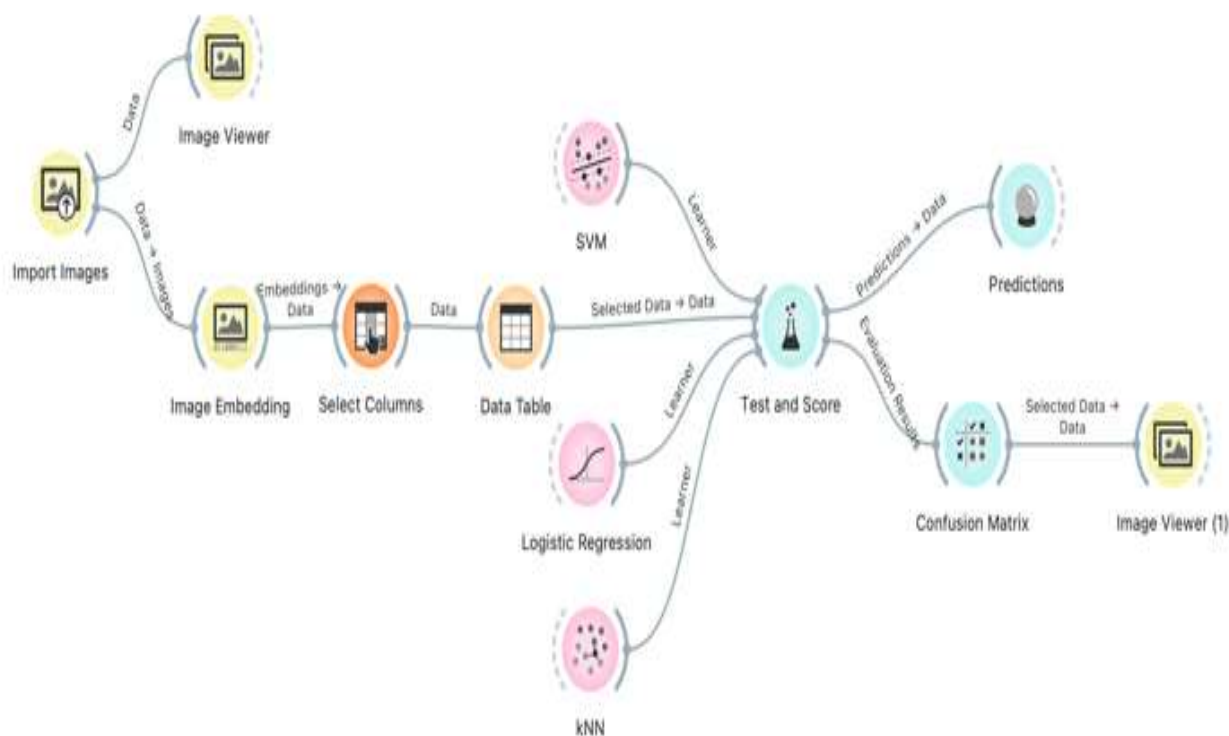
Tahapan yang dilakukan pada penelitian ini dimulai dengan membagi *dataset* hewan mamalia menjadi data latih (*data training*) dan data uji (*data testing*). Data latih dan data uji yang digunakan masing-masing terdiri dari 50 citra dan 10 citra pada setiap jenis hewan mamalia. Setiap citra baik pada data latih dan data uji dilakukan *saliency map* untuk mengambil bagian wilayah yang dianggap penting yang pada sebuah citra. Selanjutnya, pada citra hasil *saliency* dilakukan ekstraksi fitur bentuk. Semua fitur yang dihasilkan pada data latih dan data uji akan digunakan untuk proses klasifikasi menggunakan metode *support vector machine*, *k-nearest neighbor*, *logistic regression*, dan *neural network*. Metode *support vector machine*, *k-nearest neighbor*, dan *logistic regression*. Secara keseluruhan, tahapan metode yang diusulkan pada penelitian ini dapat dilihat pada Gambar 2.



Gambar 2. Tahapan Pemodelan Penelitian

1. Dataset, memiliki data dari yang digunakan adalah LHI-Animal-Faces. *Dataset* ini merupakan data citra wajah hewan tampak depan yang terdiri dari 10 jenis hewan yang mudah kita ketahui dan sangat dekat dengan kita sesuai kepada dengan kebutuhan penelitian ini maka diambil 10 jenis hewan yang tiap jenisnya terdiri dari 50 citra seperti pada contoh keterangan dari Gambar 1.
2. Praproses, tahapan praproses yang dilakukan pada penelitian ini yakni menentukan ukuran dan dimensi citra, kemudian menentukan direktori data citra uji dan data citra latih sesuai dengan kelas label yang telah ditentukan sebelum masuk ke tahap selanjutnya.
3. Arsitektur VGG-16 digunakan sebagai metode transfer learning dimana model tersebut sudah dilatih terhadap suatu dataset untuk menyelesaikan masalah yang serupa dengan cara digunakan sebagai titik permulaan dengan memodifikasi atau mengubah parameternya sehingga dapat sesuai dengan dataset baru.
4. Model, untuk melakukan perbandingan antara algoritma yang digunakan seperti :
  - a. KNN, merupakan metode untuk mengklasifikasikan data atau kasus baru berdasarkan kesamaan. Hal ini digunakan untuk mengklasifikasikan titik data berdasarkan klasifikasi data yang berdekatan [13].
  - b. Logistic Regression, adalah metode yang umum digunakan untuk melakukan analisis data dengan satu atau lebih variabel prediktor yang menggambarkan variabel respons. Variabel respon regresi logistik bernilai hanya 1 (ya) dan 0 (tidak), sehingga menghasilkan variabel respon Bernoulli [11].
  - c. SVM, merupakan singkatan dari support vector machine, dimana metode ini menggunakan fungsi linier untuk membagi dua kelompok kelas data untuk menemukan ukuran maksimum sisi sudut yang terdapat dalam ruang fitur dan ruang input [17].
5. Evaluasi, Pada tahapan ini dilakukan evaluasi model dari klasifikasi algoritma yang dilakukan dengan ekstraksi fitur yang digunakan untuk melihat hasil dari *accuracy precision, recall*, dan *f1 - Score* yang dihasilkan.

Pada pemodelan yang digunakan menggunakan aplikasi *orange* untuk melakukan proses klasifikasi gambar seperti pada gambar 3 berikut :



Gambar 3. Pemodelan Klasifikasi Gambar

### 3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk melakukan klasifikasi dalam menentukan jenis pada hewan dengan menggunakan fitur dari VGG-16 dari struktur gambar hewan yang setiap jenisnya terdiri dari 10 jenis dan pada jenis hewannya memiliki 50 data gambar perjenis hewan dan sedangkan algoritma klasifikasi yang digunakan pada penelitian ini diantaranya LR, SVM dan k-NN.



Pengujian dilakukan dengan seluruh gambar jenis hewan yang ada pada dataset gambar hewan yang sudah dikategorikan berdasarkan jenis hewannya sesuai pada intruksi sampel gambar 1 yang berjumlah 500 gambar hewan dari 10 jenis hewan, dengan dimensi 150x150 pixel berdasarkan pengujian tersebut diperoleh hasil sebagai berikut :

Tabel 1. Hasil Klasifikasi Gambar Jenis Hewan

Model	CA	F1	Precision	Recall
Logistic Regression	0.84	0.84	0.84	0.84
SVM	0.76	0.75	0.75	0.76
k-NN	0.72	0.72	0.73	0.72

Dari hasil klasifikasi algoritma LR, SVM dan k-NN pada Gambar 1, terlihat bahwa untuk nilai CA dari nilai tertinggi sebesar 0.84 pada *logistic regression*, untuk nilai dari F1 nilai tertinggi sebesar 0.84 pada *logistic regression*, untuk nilai dari *precision* nilai tertinggi sebesar 0.84 pada *logistic regression*, untuk nilai dari *recall* nilai tertinggi sebesar 0.84 pada *logistic regression*, dengan begitu dipastikan nilai tertinggi rata-rata didapatkan kepada *logistic regression* sedangkan untuk model SVM peringkat kedua dan k-NN ketiga sesuai pada Tabel 1.

Dari hasil yang didapatkan juga dapat juga disimpulkan kepada hasil pada Gambar 4 untuk melihat *confusion matrix* SVM pada algoritma sebagai berikut :

	Anjing	Ayam	Banteng	Domba	Gajah	Kucing	Kuda	Kupu-kupu	Laba-laba	Tupai	Σ
Anjing	25	0	1	6	1	9	4	2	0	2	50
Ayam	0	49	0	0	0	0	0	0	0	1	50
Banteng	2	0	22	10	5	0	10	1	0	0	50
Domba	3	0	0	35	7	1	0	0	1	3	50
Gajah	0	0	2	2	44	0	1	0	1	0	50
Kucing	4	0	0	0	0	42	0	1	2	1	50
Kuda	1	1	4	2	3	0	37	0	0	2	50
Kupu-kupu	0	0	0	0	1	0	0	48	1	0	50
Laba-laba	0	0	0	0	0	0	0	5	41	4	50
Tupai	0	0	0	4	2	2	0	1	6	35	50
Σ	35	50	29	59	63	54	52	58	52	48	500

Gambar 4. Hasil Prediksi Klasifikasi Jenis Hewan Pada SVM

Dari hasil yang didapatkan juga dapat juga disimpulkan kepada hasil pada Gambar 5 untuk melihat *confusion matrix* *logistic regression* pada algoritma sebagai berikut :

	Anjing	Ayam	Banteng	Domba	Gajah	Kucing	Kuda	Kupu-kupu	Laba-laba	Tupai	Σ
Anjing	36	0	3	1	2	2	2	2	0	2	50
Ayam	1	49	0	0	0	0	0	0	0	0	50
Banteng	0	0	35	3	2	1	9	0	0	0	50
Domba	4	0	2	37	3	1	0	0	1	2	50
Gajah	0	0	0	2	47	0	1	0	0	0	50
Kucing	3	0	0	0	0	45	0	0	1	1	50
Kuda	2	0	3	1	1	1	41	0	1	0	50
Kupu-kupu	0	0	0	0	1	0	0	48	0	1	50
Laba-laba	1	0	0	0	0	0	0	3	44	2	50
Tupai	0	1	0	4	1	2	0	0	3	39	50
Σ	47	50	43	48	57	52	53	53	50	47	500

Gambar 5. Hasil Prediksi Klasifikasi Jenis Hewan Pada LR

Dari hasil yang didapatkan juga dapat disimpulkan kepada hasil pada Gambar 6 untuk melihat *confusion matrix* k-NN pada algoritma sebagai berikut :

	Anjing	Ayam	Banteng	Domba	Gajah	Kucing	Kuda	Kupu-kupu	Laba-laba	Tupai	Σ
Anjing	32	1	1	4	1	4	4	1	0	2	50
Ayam	0	44	0	0	0	1	1	1	0	3	50
Banteng	3	1	34	6	2	0	4	0	0	0	50
Domba	2	0	7	26	12	0	0	1	1	1	50
Gajah	0	0	5	4	40	0	0	1	0	0	50
Kucing	5	0	1	0	0	41	0	0	2	1	50
Kuda	2	1	10	3	2	0	31	0	0	1	50
Kupu-kupu	0	0	0	0	0	0	0	46	3	1	50
Laba-laba	2	0	0	0	2	0	0	9	36	1	50
Tupai	1	1	1	8	2	1	1	3	1	31	50
Σ	47	48	59	51	61	47	41	62	43	41	500

Gambar 6. Hasil Prediksi Klasifikasi Jenis Hewan Pada k-NN

Dari hasil yang didapatkan juga dapat disimpulkan kepada hasil pada Gambar 1 yang sudah dimodelkan dapat dilihat salah satu model pada *logistic regression* yang salah dalam klasifikasi jenis hewan yang dari tupai adalah sebagai berikut :



Gambar 6. Hasil Prediksi Klasifikasi Jenis Hewan Pada LR yang salah

#### 4. KESIMPULAN

Berdasarkan dari kesimpulannya untuk melakukan klasifikasi pada gambar yang menggunakan data sebanyak 500 gambar hewan dengan 10 jenis hewan yang berbeda dari setiap jenisnya memiliki 50 data gambar hewan yang menentukan jenis kategori hewan yang menggunakan model *support vector machine*, *k-nearest neighbor*, *logistic regression* didapatkan bahwasanya tingkatan akurasi tertinggi adalah *logistic regression* dan berikutnya kepada *support vector machine* lalu *k-nearest neighbor*.

#### UCAPAN TERIMA KASIH

Terima kasih kepada Ibu Dr. Rika Rosnelly, S.Kom., M.Kom. serta kepada Bapak Dr. Zakarias Situmorang, MT. dan juga Bapak Puji Sari Ramadhan, S.Kom., M.Kom. yang bersama-sama untuk menyelesaikan karya ilmiah ini dengan sebaik-baiknya.

#### DAFTAR PUSTAKA

- [1] K. O. Lauw *et al.*, "Identifikasi Jenis Anjing Berdasarkan Gambar Menggunakan Convolutional Neural Network Berbasis Android," *J. Infra*, vol. 8, no. 2, pp. 37–43, 2020.
- [2] H. Dhika, N. R. Kurnianda, P. Irfansyah, and W. Ananta, "Model Prediksi Jenis Hewan dengan Metode Convolution Neural Network," *J. fORMAT*, vol. 9, no. 1, pp. 31–40, 2020.
- [3] M. Afif, A. Fawwaz, K. N. Ramadhani, and F. Sthevanie, "Klasifikasi Ras pada Kucing menggunakan Algoritma Convolutional Neural Network(CNN)," *J. Tugas Akhir Fak. Inform.*, vol. 8, no. 1, pp. 715–730, 2020.
- [4] E. D. Suryanto, "Ekstraksi Fitur Haralick Menggunakan Citra Mikroskop Digital Trinocular Untuk Proses Identifikasi Cacing Penyakit Kaki Gajah," no. October, pp. 1–85, 2015.

- [5] D. Ricardo and G. Gasim, "Perbandingan Akurasi Pengenalan Jenis Beras dengan Algoritma Propagasi Balik pada Beberapa Resolusi Kamera," *J. RESTI (Rekayasa Sist. dan Teknol. Informatika)*, vol. 3, no. 2, pp. 131–140, 2019.
- [6] Siti Khotimatul Wildah, S. Agustiani, Ali Mustopa, Nanik Wuryani, Hendri Mahmud Nawawi, and Rizky Ade Safitri, "Pengenalan Wajah Menggunakan Pembelajaran Mesin Berdasarkan Ekstraksi Fitur Pada Gambar Wajah Berkualitas Rendah," *INFOTECH J. Inform. Teknol.*, vol. 2, no. 2, pp. 95–103, 2021.
- [7] E. Zangeneh, M. Rahmati, and Y. Mohsenzadeh, "Low resolution face recognition using a two-branch deep convolutional neural network architecture," *Expert Syst. Appl.*, vol. 139, pp. 1–11, 2020.
- [8] F. Setiawan and D. A. R., "Sistem Pengenalan Wajah Dengan Metode Local Binary Pattern Histogram Pada Firebase," *SeNTIK*, vol. 4, no. 1, pp. 19–25, 2020.
- [9] M. Athoillah, "Pengenalan Wajah Menggunakan SVM Multi Kernel dengan Pembelajaran yang Bertambah," *J. Online Inform.*, vol. 2, no. 2, p. 84, 2018.
- [10] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, R. Sun, and I. Garcíá-Magarinõ, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," *Mob. Inf. Syst.*, vol. 2018, 2018.
- [11] D. Devito, R. C. Wihandika, and A. W. Widodo, "Ekstraksi Ciri Untuk Klasifikasi Gender Berbasis Citra Wajah Menggunakan Metode Histogram of Oriented Gradients," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 8, pp. 8002–8011, 2019.
- [12] "Face Recognition Untuk Sistem Pengaman Rumah Menggunakan Metode HOG dan KNN Berbasis Embedded."
- [13] A. Chatterjee, J. Saha, J. Mukherjee, S. Aikat, and A. Misra, "Unsupervised Land Cover Classification of Hybrid and Dual-Polarized Images Using Deep Convolutional Neural Network," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 6, pp. 969–973, 2021.
- [14] Y. Yohannes, M. R. Pribadi, and L. Chandra, "Klasifikasi Jenis Buah dan Sayuran Menggunakan SVM Dengan Fitur Saliency-HOG dan Color Moments," *Elkha*, vol. 12, no. 2, p. 125, 2020.
- [15] D. Alita, Y. Fernando, and H. Sulistiani, "Implementasi Algoritma Multiclass Svm Pada Opini Publik Berbahasa Indonesia Di Twitter," *J. Tekno Kompak*, vol. 14, no. 2, p. 86, 2020.
- [16] F. Handayani, "Komparasi Support Vector Machine, Logistic Regression Dan Artificial Neural Network Dalam Prediksi Penyakit Jantung," *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 3, p. 329, 2021.
- [17] A. Bimantara and T. A. Dina, "Klasifikasi Web Berbahaya Menggunakan Metode Logistic Regression," *Annu. Res. Semin.*, vol. 4, no. 1, pp. 173–177, 2019.
- [18] T. Purwaningsih, T. Nurhikmat, and P. B. Utami, "Image classification of Golek puppet images using convolutional neural networks algorithm," *Int. J. Adv. Soft Comput. its Appl.*, vol. 11, no. 1, pp. 34–45, 2019.