LAPORAN PRAKTIKUM PEMBELAJARAN MESIN

PENERAPAN CONVOLUTIONAL NEURAL NETWORK (CNN) UNTUK IDENTIFIKASI HEWAN PELIHARAAN: ANJING DAN KUCING DENGAN DATASET DOG VS CAT CLASSIFICATION



Disusun Oleh:

KELOMPOK 4 RA

1.	Muhammad Nabil Azizi	120450090
2.	Rika Ajeng Finatih	121450036
3.	A Rafi Paringgom Iwari	121450039
4.	Anissa Luthfi Alifia	121450093
5.	Elsyah Sapyrah	121450096

PROGRAM STUDI SAINS DATA FAKULTAS SAINS INSTITUT TEKNOLOGI SUMATERA 2023/2024

LAPORAN PRAKTIKUM PEMBELAJARAN MESIN

PENERAPAN CONVOLUTIONAL NEURAL NETWORK (CNN) UNTUK IDENTIFIKASI HEWAN PELIHARAAN: ANJING DAN KUCING DENGAN DATASET DOG VS CAT CLASSIFICATION

Muhammad Nabil Azizi1), Rika Ajeng Finatih2), A Rafi Paringgom Iwari3), Anissa Luthfi Alifia4), Elsyah Sapyrah5).

Program Studi Sains Data, Jurusan Sains, Institut Teknologi Sumatera *Email: muhammad.120450090@student.itera.ac.id 1)*, rika.121450036@student.itera.ac.id 2), arafi.121450039@student.itera.ac.id 3), anissa.121450093@student.itera.ac.id 4), elsyah.121450096@student.itera.ac.id 5)

Abstract

Keeping animals especially dogs and cats has become an integral part of modern life. In this context, the development of an automatic identification system using Convolutional Neural Network (CNN) becomes important. This research builds and evaluates a CNN model to classify dog and cat images with a dataset containing 1,494 images using a training and test data ratio of 80:20. Preprocessing involves TensorFlow, Keras, and EBImage to prepare the data. The CNN model consists of three convolutional layers with 128 filters, using ReLu as the activation function, 3x3 kernels, and 2x2 maxpooling. With 15 epochs, batch_size 38, and validation_split 0.1, the model achieved approximately 87.99% accuracy on the test data. Evaluation shows the model's ability to classify dogs (716/952) and cats (183/244) on test data. These results mark the effectiveness of CNNs in identifying pets, indicating the potential for broader applications in image classification tasks.

Keywords: Convolutional Neural Network, Image Recognition, Pet Classification, Automatic Identification, Data Preprocessing.

Abstrak

Pemeliharaan hewan khususnya anjing dan kucing telah menjadi bagian integral dari kehidupan modern. Dalam konteks ini, pengembangan sistem identifikasi otomatis menggunakan Convolutional Neural Network (CNN) menjadi penting. Penelitian ini membangun dan mengevaluasi model CNN untuk mengklasifikasikan gambar anjing dan kucing dengan dataset berisi 1,494 gambar menggunakan perbandingan data latih dan uji 80:20. Preprocessing melibatkan TensorFlow, Keras, dan EBImage untuk mempersiapkan data. Model CNN terdiri dari tiga lapisan konvolusi dengan 128 filter, menggunakan ReLu sebagai fungsi aktivasi, kernel 3x3, dan maxpooling 2x2. Dengan 15 epoch, batch_size 38, dan validation_split 0,1, model mencapai akurasi sekitar 87,99% pada data uji. Evaluasi menunjukkan kemampuan model dalam mengklasifikasikan anjing (716/952) dan kucing (183/244) pada data uji. Hasil ini menandai efektivitas CNN dalam mengidentifikasi hewan peliharaan, menunjukkan potensi aplikasi yang lebih luas dalam tugas pengklasifikasian citra.

Kata Kunci: Convolutional Neural Network, Pengenalan Gambar, Klasifikasi Hewan Peliharaan, Identifikasi Otomatis, Preprocessing Data.

1. PENDAHULUAN

Dengan meningkatnya popularitas pemeliharaan hewan terutama anjing dan kucing, kebutuhan akan sistem identifikasi otomatis untuk mengklasifikasikan gambar hewan

peliharaan menjadi semakin penting. Identifikasi hewan peliharaan ini dapat memiliki dampak besar dalam konteks manajemen dan pemantauan keberadaan hewan peliharaan.

Pemeliharaan hewan terutama anjing dan kucing telah menjadi bagian tak terpisahkan dari kehidupan manusia dan kebutuhan akan sistem identifikasi otomatis untuk mengklasifikasikan gambar hewan peliharaan semakin mendesak. Dalam konteks ini, *Convolutional Neural Network* (CNN) muncul sebagai metode unggulan dalam pengenalan objek pada citra digital.

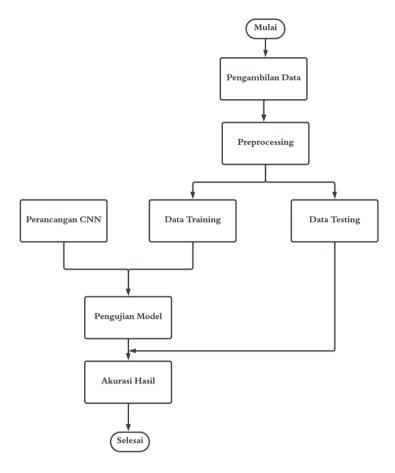
Sebelum dilakukan pemodelan CNN, data yang telah dikumpulkan dilakukan *preprocessing* terlebih dahulu. *Preprocessing* adalah tahapan untuk melakukan pembersihan data (*data cleaning*) dan menggabungkan data yang sudah ada (*integrasi data*) [1]. Lalu, dilakukan split data menjadi data train dan data test. Kemudian, dilakukan perancangan model CNN, pengujian model dan menghitung akurasi.

Convolutional Neural Network (CNN) adalah salah satu algoritma dari Deep Learning yang merupakan pengembangan dari Multi Layer Perceptron (MLP) yang dirancang untuk mengolah data dalam bentuk grid, salah satunya citra dua dimensi, misalnya gambar atau suara. Convolutional Neural Network digunkaan untuk mengklasifikasikan data yang terlabel dengan menggunakan metode supervised learning, yang mana cara kerja dari supervised learning adalah terdapat data yang dilatih dan terdapat variabel yang ditargetkan sehingga tujuan dari metode ini adalah mengelompokan suatu data ke data yang sudah ada. CNN sering digunkan untuk mengenali benda atau pemandangan, dan melakukan deteksi dan segmentasi objek. Convolutional Neural Network menggabungkan tiga pokok arsitektur, yaitu local receptive fields, Shared weight yang berupa filter, dan spatial subsampling yang berupa Pooling [2].

Tujuan utama percobaan ini adalah mengembangkan model CNN yang dapat secara akurat membedakan antara gambar anjing dan kucing. Dengan mencapai tujuan ini, diharapkan hasil percobaan dapat memberikan kontribusi pada pengembangan sistem identifikasi hewan peliharaan yang lebih canggih dan efisien. Keberhasilan implementasi CNN dalam klasifikasi hewan peliharaan akan memberikan dampak positif terhadap pemilik hewan peliharaan, penyedia layanan veteriner, dan upaya mereka untuk meningkatkan manajemen dan perawatan hewan peliharaan. Selain itu, percobaan ini juga dapat membuka jalan untuk aplikasi lebih lanjut dalam bidang pengenalan objek pada citra digital menggunakan teknologi *deep learning*.

2. METODE

Pada percobaan digunakannya pemilihan pengklasifikasian hewan peliharaan yaitu Dog vs Cat. Hal ini dilakukan dengan menggunakan metode *Convolutional Neural Network*. Proses perancangan dan implementasi terdiri atas beberapa tahapan. Tahap dalam percobaan disajikan pada Gambar 1.



Gambar 1: Tahap Pengklasifikasian

Hewan Peliharaan Dog vs Cat dengan Metode CNN

Pada Gambar 1 pengambilan data dilakukan, data yang didapatkan pada percobaan ini berasal dari sebuah situs yang bernama *kaggle*, dimana di dalam situs ini berisikan banyak sekali dataset yang dapat digunakan dalam bahan percobaan. Dataset ini akan digunakan pada saat proses pelatihan dan pengujian model arsitektur CNN. Pada tahap *preprocessing* dilakukan persiapan data agar dapat digunakan pada tahap selanjutnya. Setelah data diproses, tahapan berikutnya adalah membangun arsitektur model *Convolution Neural Network* (CNN). Model ini akan digunakan untuk mengetahui tingkat akurasi dari klasifikasi hewan peliharaan yakni anjing dan kucing. Selanjutnya adalah tahap pelatihan model yang sudah dirancang menggunakan data yang telah diolah.

```
# Mount Google Drive
drive.mount('/content/drive')
# Definisi path untuk dataset gambar anjing dan kucing
fig_dog = '/content/drive/MyDrive/Kelompok04_PM_RA/Kel04_RA_PM_Modul04/PetImages/Dog'
fig_cat = '/content/drive/MyDrive/Kelompok04_PM_RA/Kel04_RA_PM_Modul04/PetImages/Cat'
# Membaca nama file gambar anjing dan kucing
dog_files = daftar_file(fig_dog)
cat_files = daftar_file(fig_cat)
# Menentukan label untuk gambar anjing dan kucing
label_anjing = tetapkan_label(dog_files, 1)
label_kucing = tetapkan_label(cat_files, 0)
# Menggabungkan file dan label
semua_file = gabung_list(dog_files, cat_files)
semua_label = gabung_list(label_anjing, label_kucing)
# Memuat dan mengolah gambar
data = muat_dan_proses_gambar(semua_file)
# Konversi label dan gambar ke dalam bentuk array NumPy
x, y = konversi_ke_array_numpy(data, semua_label)
# Pembagian dataset menjadi data pelatihan dan pengujian
x_pelatihan, x_pengujian, y_pelatihan, y_pengujian = bagi_dataset(x, y, ukuran_pengujian=0.2, seed_acak=42)
# Membuat model jaringan saraf
model = buat_model_jaringan_saraf()
# Kompilasi model
kompilasi_model(model, pengoptimal='adam', kerugian='sparse_categorical_crossentropy', metrik=['akurasi'])
 # Pelatihan model
 latih model(model, x_pelatihan, y_pelatihan, pembagian_validasi=0.1, epoch=15)
 # Evaluasi model pada dataset pengujian
 evaluasi_model(model, x_pengujian, y_pengujian)
 # Prediksi gambar
 gambar_input = baca_gambar_input(fig_dog + '/10226.jpg')
 gambar_diubah = ubah_ukuran_gambar(gambar_input, (128, 128))
gambar_diskalakan = skalakan_gambar(gambar_diubah)
gambar_direshape = bentuk_ulang_gambar(gambar_diskalakan)
 prediksi = prediksi_gambar(model, gambar_direshape)
# Menampilkan hasil prediksi
 tampilkan_hasil_prediksi(prediksi)
```

Gambar 2: Pseudocode Pembuatan Model CNN



Gambar 3: Flowchart Pembuatan Model CNN

Pada tampilan Gambar 2 dan Gambar 3, disajikan *Pseudocode* dan *Flowchart* dari kode program yang telah dibuat dalam pembuatan model CNN. Tujuan dari penyajian *Pseudocode* dan *Flowchart* dalam laporan ini adalah untuk memberikan visualisasi yang jelas dan sistematis

terhadap langkah-langkah utama yang diimplementasikan dalam pemodelan CNN. *Pseudocode* memberikan representasi deskriptif mengenai algoritma dan logika pemrograman yang digunakan, sementara *Flowchart* memberikan gambaran visual mengenai alur kerja program dan hubungan antar bagian dalam model CNN. Dengan demikian, *pseudocode* dan *Flowchart* bertindak sebagai alat bantu yang memudahkan pemahaman dan interpretasi terhadap implementasi kode program secara keseluruhan.

2.1 Pengumpulan Data

Pengumpulan data yang dilakukan ini didapatkan dari sebuah situs web yang bernama *kaggle*. Dimana didalam *kaggle* [3] terdapat banyak sekali dataset yang dapat digunakan dalam klasifikasi. Di Dalam situs ini kita dapat dengan mudah mendapatkan data dan mengunduhnya. Di dalam dataset *dog vs cat classification* ini terdapat 1,494 gambar klasifikasi dari dua kelas hewan peliharaan anjing dan kucing. Banyaknya gambar yang diambil sebanyak 1196 sebagai data latih dan 298 sebagai data uji. Dari keseluruhan gambar tersebut terdiri gambar hewan peliharaan anjing sebanyak 1189, dan sebanyak 205 gambar hewan peliharaan kucing.

2.2 Preprocessing Data

Preprocessing gambar dengan mengubah ukuran gambar hewan peliharaan anjing dan kucing menjadi ukuran yang sama. Langkah selanjutnya dilakukan pembagian data yang telah dikumpulkan menjadi dua, yaitu data latih (training) dan data uji (testing). Langkah selanjutnya dilakukan pelabelan pada masing-masing gambar sebagai pelabelan dari jenis hewan peliharaan yaitu anjing dan kucing.

2.3 Perancangan CNN

Perancangan CNN merupakan fase kunci dalam perancangan model yang bertujuan untuk melatih data agar dapat mengenali objek yang diinginkan. Proses perancangan model melibatkan penentuan jumlah lapisan yang digunakan, pengaturan filter, menentukan ukuran kernel, serta memilih fungsi aktivasi dan ukuran *pooling* yang sesuai.

2.4 Pelatihan Model

Setelah merancang model CNN, langkah berikutnya adalah melatih model menggunakan data latih yang telah didapatkan sebelumnya. Selama Proses pengujian model, jumlah epoch (iterasi) digunakan untuk mengatur seberapa sering jaringan akan menjalani pelatihan. Pada tahap ini, peran penting dimainkan oleh fungsi loss dalam mengevaluasi performa dari model CNN yang di kembangkan.

2.5 Pengujian Model

Langkah berikutnya adalah melakukan uji coba model pada data uji, sebuah tahap yang bertujuan untuk mengevaluasi tingkat akurasi dari *Convolution Neural Network* (CNN) yang telah dibangun. Tingkat akurasi ini mencerminkan sejauh mana model mampu mengklasifikasikan jenis hewan peliharaan dengan benar. Tingkat akurasi yang tinggi

menandakan bahwa model mampu melakukan klasifikasi yang baik terhadap kedua jenis hewan peliharaan yang diuji.

3. HASIL DAN PEMBAHASAN

Pada hasil dan pembahasan ini, akan dibahas mengenai hasil yang diperoleh setelah dilakukannya percobaan pembuatan model *Convolutional Neural Network* (CNN) dengan menggunakan dataset untuk pengklasifikasian hewan peliharaan yakni, anjing dan kucing. Analisis hasil ini disajikan secara terperinci dengan tujuan untuk memberikan pemahaman mendalam tentang kinerja model CNN dalam mengenali dan memisahkan gambar-gambar hewan peliharaan tersebut.

3.1 Preprocessing Data

Setelah data terkumpul, langkah selanjutnya adalah memasukan data dan membuat sistem untuk pengenalan hewan peliharaan yaitu anjing dan kucing. Proses pengenalan ini menggunakan metode deep learning, yakni *Convolutional Neural Network* (CNN). Dataset yang digunakan berasal dari *Kaggle* dengan nama *Dog vs Cat Classification*.

Dalam Percobaan ini, langkah awal sebelum memasukan dataset agar dapat dibaca oleh sistem adalah melakukan instalasi *packages* yang akan digunakan dalam pengolahan data. Packages yang di install meliputi: *packages TensorFlow, packages Keras,* dan *packages EBImage*. Setelah berhasil menginstall *packages*, langkah selanjutnya adalah memanggil packages tersebut agar dapat digunakan dalam pembuatan model *Convolutional Neural Network* (CNN).

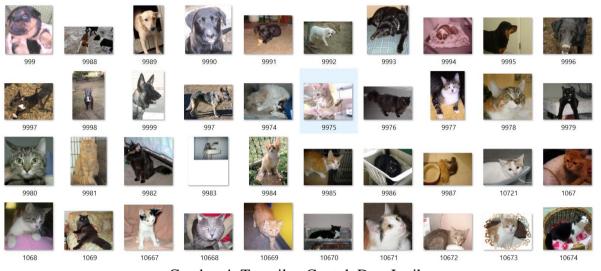
Metode pengenalan menggunakan CNN dimulai dengan tahap klasifikasi gambar melalui proses pelatihan. Pelatihan dilakukan untuk memastikan bahwa model yang dibuat memiliki kemampuan mengenali objek yang diinginkan. Pada percobaan ini, pembagian antara data latih dan uji dapat ditemukan dalam Tabel 1.

Berdasarkan Tabel 1 yang digunakan terdiri dari 1,494 gambar yang dikategorikan sebagai gambar Anjing dan Kucing. Gambar anjing terdiri dari 1,189 data, sedangkan data gambar kucing terdiri dari 305 data. Sebelumnya, kedua dataset ini telah dikelompokkan secara manual ke dalam kelas Anjing dan Kucing. Proses pembagian antara data latih dan data uji menggunakan perbandingan 80:20. Detail pembagian dari masing-masing kategori dapat dilihat pada Tabel 1.

Jenis Hewan	Jumlah	Jumlah	
Peliharaan	Data Latih	Data Uji	
Anjing	952	237	
Kucing	244	61	
Total	1196	298	

Tabel 1: Pembagian Dataset Untuk Pengujian

Contoh dari data latih yang diterapkan pada percobaan ini tersaji dalam Gambar 4. Sementara itu, data uji terdiri dari 237 gambar hewan peliharaan anjing dan 61 gambar hewan peliharaan kucing, yang dapat dilihat pada Gambar 5. Langkah selanjutnya adalah melakukan pelabelan pada kedua jenis data tersebut. Pelabelan ini dilakukan dengan mengasosiasikan bilangan biner yaitu 0 dan 1. Label angka 0 diberikan untuk gambar kucing, sedangkan label angka 1 diberikan untuk gambar hewan peliharaan anjing.



Gambar 4: Tampilan Contoh Data Latih



Gambar 5: Tampilan Contoh Data Uji

3.2 Pembuatan Model

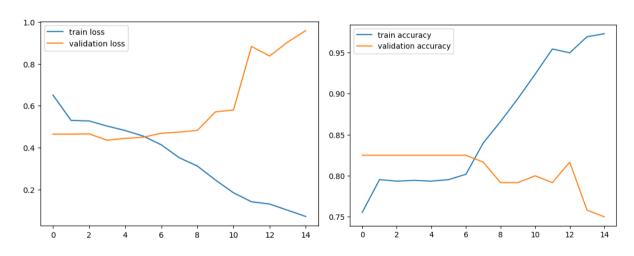
Dalam pembuatan model, langkah umum melibatkan proses konvolusi yang ditemani oleh fungsi aktivasi dan proses *pooling*. Jumlah proses ini biasanya disesuaikan dengan kebutuhan. Model konstruksi dengan mengintegrasikan beberapa jenis lapisan yang berbeda, termasuk lapisan konvolusi, lapisan *pooling*, lapisan *dropout*, lapisan *flatten*, dan lapisan *dense*. Proses konvolusi pada percobaan ini dilakukan sebanyak 3 kali, sebagaimana umumnya tercermin dari penggunaan beberapa *convolution layer*. Alasan digunakannya konvolusi

sebanyak 3 kali karena dengan pendekatan ini dirasa dengan melihat peneliti sebelumnya sudah terbukti efektif dalam mengekstraksi fitur hirarki dari data gambar.

Dalam proses CNN ini juga, digunakannya fungsi aktivasi ReLu (*Rectified Linear Unit*) untuk mempercepat tahap pelatihan. Ukuran kernel yang diterapkan pada setiap lapisan konvolusi adalah 3x3. Selama proses *pooling*, metode *maxpooling* digunakan dengan filter berukuran 2x2. Hal ini menyebabkan adanya pemilihan nilai maksimum dari setiap area 2x2 piksel pada setiap pergeseran, dan memperkuat representasi fitur yang diambil.

Jumlah filter *convolution* pada ketiga layer yang digunakan sebesar 128. Pemilihan jumlah filter sebesar 128 dipertimbangkan sebagai ukuran yang memadai untuk memenuhi karakteristik dataset *dog vs cat classification*. Keputusan ini didasarkan pada evaluasi terhadap kompleksitas fitur dalam gambar dan kebutuhan model untuk dapat mengekstrak representasi yang memadai. Dengan menggunakan jumlah filter yang cukup, diharapkan model dapat mengenali pola dan fitur dengan baik, dan meningkatkan kinerja klasifikasi jenis hewan peliharaan.

Langkah berikutnya adalah melatih data gambar jenis hewan peliharaan menggunakan model dengan fungsi *fit*. Saat proses *fit model*, ditentukan epoch sebanyak 15, *batch_size* sebanyak 38, dan *validation_split* sebesar 0,1. Epoch digunakan untuk menentukan seberapa banyak kali keseluruhan data akan diperiksa oleh jaringan. Nilai *Batch_size* mencerminkan jumlah contoh pelatihan dalam satu iterasi maju atau bahkan mundur. Semakin besar nilai *batch_size*, semakin banyak memori yang dibutuhkan. Hasil loss dan akurasi dari model yang terbentuk, dapat dilihat pada Gambar 6.



Gambar 6: Grafik Hasil Nilai Loss dan *Accuracy*

Pada Gambar 6, terlihat bahwa pada iterasi ke-15 nilai loss untuk data latih mencapai 1,1144, sementara akurasi pada data latih mencapai 0,75250834. Kemudian, saat model di uji pada data pengujian, diperoleh hasil loss sebesar 0,48852372 dengan akurasi 0,979933. Hasil ini menunjukkan bahwa model yang dibuat sangat efektif dalam mengklasifikasikan data latih dengan loss yang mendekati nol dan akurasi sebesar 75,25%. Pada data pengujian, model juga

memberikan hasil yang sangat baik, dengan loss yang rendah dan akurasi yang tinggi (sekitar 97,99%). Sehingga, dari pemodelan ini bisa dikatakan dan dianggap berhasil dan mampu menggeneralisasikan dengan baik pada data baru.

3.3 Uji Coba dan Evaluasi

Hasil klasifikasi untuk data latih dan data uji terdokumentasikan dalam Gambar 4 dan Gambar 5. Pada Gambar 5, terlihat hasil prediksi untuk data latih. Semua gambar hewan peliharaan anjing sebanyak 716 gambar berhasil diklasifikasikan, demikian pula dengan gambar hewan peliharaan kucing yang mencapai 183 dari total 244 gambar kucing yang digunakan.

	Actual		
		0	1
Predicted	0	183	61
	1	237	716

Gambar 7: Hasil Klasifikasi Data Latih

Hasil klasifikasi data uji pada Gambar 6, terlihat bahwa klasifikasi untuk data uji terdiri dari 237 gambar untuk hewan peliharaan anjing, dan 61 gambar dengan hewan peliharaan kucing. Hasil prediksi model pada data uji, jumlah gambar hewan peliharaan anjing yang diklasifikasikan adalah 232. Selain itu, gambar hewan peliharaan kucing yang diklasifikasikan sebanyak 59 dari total 61 gambar yang di klasifikasi.

	Actual		
		0	1
Predicted	0	59	2
	1	5	232

Gambar 8: Hasil Klasifikasi Data Uji

4. KESIMPULAN

Dalam percobaan ini, pembuat model CNN untuk mengklasifikasikan hewan peliharaan (anjing vs kucing) telah berhasil memberikan hasil yang positif. Dataset yang terdiri dari 1,494 gambar anjing dan kucing telah dipersiapkan dengan pembagian data latih dan data uji menggunakan perbandingan 80:20. Proses preprocessing melibatkan instalasi *packages TensorFlow, Keras*, dan *EBImage* untuk memastikan kesiapan data sebelum dimasukkan ke dalam model.

Arsitektur model CNN terdiri dari tiga lapisan konvolusi dengan jumlah filter sebanyak 128, menggunakan fungsi aktivasi ReLu, dan kerbel berukuran 3x3. Selama proses *pooling*, metode *maxpooling* diterapkan dengan filter berukuran 2x2. Proses pelatihan model dilakukan sebanyak 15 epoch, dengan batch_size 38, dan validastion_split 0,1. Hasil akhir pelatihan menunjukkan loss sekitar 1,1144 dan akurasi sekitar 75,25% pada data latih, sedangkan pada data pengujian model memberikan hasil loss sebesar 0,48882372 dan akurasi sekitar 87,99%.

Pada uji coba dan evaluasi, model berhasil mengklasifikasikan gambar hewan peliharaan pada data latih dengan akurasi yang baik. Pada data uji, model tetap mampu memberikan hasil yang memuaskan, dengan kemampuan mengklasifikasikan gambar anjing sebanyak 716 dari 952, dan gambar kucing sebanyak 183 dari 244. Kesimpulan akhir menunjukkan bahwa model CNN ini efektif dalam mengidentifikasi dan memisahkan gambar hewan peliharaan. Hal ini menunjukkan bahwa potensi untuk diterapkan pada tugas pengklasifikasian yang lebih luas.

REFERENSI

- [1] A. Adji Dwi Alfarizi, "Pemanfaatan Data Mining Dalam Memprediksi Produksi Pada PTPupuk Sriwidjaja Palembang Menggunakan Algoritma Regresi Liniear Berganda," *Jurnal Nasional Ilmu Komputer*, p. 56, 2021.
- [2] "IMPLEMENTASI CONVOLUTION NEURAL NETWORK (CNN) UNTUK KLASIFIKASI JAMUR KONSUMSI DI INDONESIA MENGGUNAKAN KERAS," *Tutut Furi Kusumaningrum*, p. 34, 2018.
- [3] K. PADHIAR, Petfinder.com & Microsoft, [Online]. Available: https://www.kaggle.com/datasets/karakaggle/kaggle-cat-vs-dog-dataset. [Diakses 23 11 2023].

LAMPIRAN

Berikut merupakan lampiran dari hasil kode program, dataset yang digunakan, dan juga video presentasi dari kelompok 4 RA Praktikum Pembelajaran Mesin: https://drive.google.com/drive/folders/11ZlxTwbyM3AtAExn5-9SspzdSkgbJKE6?usp=sharing