

KLASIFIKASI RAS ANJING MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK DENGAN ARSITEKTUR VGG-16

Sandy Danish Arkansa¹⁾ Chairisni Lubis²⁾

¹⁾²⁾ Teknik Informatika Universitas Tarumanagara
Jl. Letjen S Parman no 1, Jakarta 11440 Indonesia

¹⁾sandydanish123@gmail.com ²⁾chairisnil@fti.untar.ac.id

Corresponding Author : chairisnil@fti.untar.ac.id

ABSTRACT

Dogs are lovable animals that bring companionship and fun to any household, but things to consider before getting a dog are their daily needs including food, shelter, and animal care, as well as affection and physical and mental stimulation. This research was conducted to introduce dog breeds and how to care for each breed. This system is built using Mask R-CNN and Convolutional Neural Network (CNN), a deep learning architecture.

Mask R-CNN model is used for detecting and cropping dog in images, trained using Microsoft Common Object of Context (MS COCO) dataset. CNN is used for classification of dog breeds in images and is trained using 17.513 images of 17 different breeds. Result for Mask R-CNN show the detection accuracy for dogs has 74% using test images, and CNN show the identification accuracy using test images has 82% accuracy, and for CNN using cropped images has 87% accuracy.

Key words

Convolutional Neural Network, Dogs, Identification, Mask R-CNN

1. Pendahuluan

Anjing adalah hewan yang menyenangkan yang membawa persahabatan dan kesenangan ke dalam suatu rumah tangga, tetapi hal yang harus diperhatikan sebelum memelihara anjing adalah kebutuhan sehari-hari termasuk makanan, tempat tinggal, dan perawatan hewan, begitu juga untuk kasih sayang dan stimulasi fisik dan mental. Hal lain yang perlu diperhatikan adalah makanan untuk anjing tersebut, vaksinasi kesehatan, biaya untuk dokter hewan lainnya, asuransi juga perlu diperhatikan untuk perawatan anjing [1].

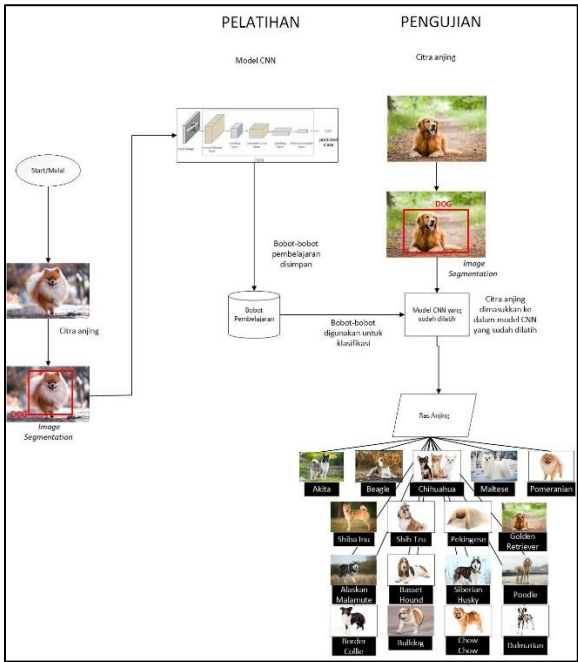
Dengan adanya berbagai ras anjing yang berbeda, maka perawatan yang dilakukan terhadap anjing juga berbeda-beda sesuai dengan rasnya. Banyaknya jenis variasi anjing yang muncul akibat pembiakan yang direncanakan mempersulit pemilik anjing untuk mengetahui ras anjing yang dimiliki. Perbedaan ciri setiap anjing dapat digunakan untuk melakukan klasifikasi jenis anjing.

Tujuan dari perancangan klasifikasi ras anjing ini adalah untuk membuat sistem yang dapat mendeteksi anjing dengan menggunakan metode Mask R-CNN yang merupakan kerangka kerja yang secara konseptual sederhana, fleksibel, dan umum untuk *instance segmentation* objek dan melakukan klasifikasi ras anjing berdasarkan bentuk wajah dan bentuk telinga dengan menggunakan metode *Convolutional Neural Network (CNN)* yang merupakan salah satu metode dari *Deep Learning (DL)* digunakan untuk pengenalan dan klasifikasi pada suatu gambar/objek. Arsitektur pada perancangan ini menggunakan arsitektur VGG-16 yang merupakan salah satu arsitektur terbaik untuk model CNN.

Pada perancangan sebelumnya dengan judul "Identifikasi Jenis Anjing Berdasarkan Gambar Menggunakan Convolutional Neural Network Berbasis Android" oleh Kevin Oktovio Lauw, Leo Willyanto Santoso, dan Rolly Intan mendapatkan hasil klasifikasi mendapatkan tingkat akurasi sekitar 56% [2]. Pada penelitian ini, akan dikembangkan sistem untuk meningkatkan tingkat akurasi berdasarkan 17 ras anjing terbaik untuk dipelihara di rumah yaitu Akita, Beagle, Chihuahua, Maltese, Pomeranian, Shiba Inu, Shih Tzu, Pekingese, Golden Retriever, Alaskan Malamute, Basset Hound, Siberian Husky, Poodle, Border Collie, Bulldog, Chow Chow, dan Dalmatian.

2. Metode Perancangan

Pada perancangan klasifikasi ras anjing ini, sistem dibangun menggunakan metode *Mask R-CNN* untuk mendeteksi anjing pada suatu gambar, dan menggunakan metode CNN untuk melakukan klasifikasi ras anjing berdasarkan gambar anjing tersebut. Model *Mask R-CNN* menggunakan *pre-trained* model sehingga dapat langsung digunakan untuk mendeteksi anjing, sedangkan untuk model CNN akan dilakukan pelatihan menggunakan data latih yang menghasilkan model untuk diimplementasikan pada program. Tahapan pada sistem dapat dilihat pada Gambar 1



Gambar 1. Flow diagram Program

2.1. Dataset

Pada percobaan ini, dataset yang digunakan untuk mendeteksi anjing dengan Mask R-CNN adalah *Microsoft Common Object of Context (MS COCO)* yang merupakan kumpulan data berskala besar yang berisi 328.000 gambar untuk objek sehari-hari dan manusia, sedangkan dataset yang digunakan untuk klasifikasi ras anjing dengan metode CNN adalah citra anjing dari 17 ras anjing yang terdiri dari kurang lebih 1000 citra yang berasal dari *images.cv*. Contoh data citra dapat dilihat pada Tabel 1.

Tabel 1. Citra Ras Anjing

Ras	Citra
-----	-------

Akita	
Alaskan Malamute	
Basset	
Beagle	
Border Collie	
Bulldog	
Chihuahua	

Chow	
Dalmatian	
Golden Retriever	
Maltese	
Pekingese	
Pomeranian	

Poodle	
Shiba Inu	
Shih Tzu	
Siberian Husky	

Tabel untuk jumlah dataset dapat dilihat pada Tabel 2

Tabel 2. Dataset Citra Anjing

Ras	Jumlah
Akita	648
Alaskan Malamute	1368
Basset	1175
Beagle	1434
Border Collie	948
Bulldog	880
Chihuahua	1041
Chow	1050
Dalmatian	807
Golden Retriever	1065
Maltese	1368
Pekingese	1124
Pomeranian	980
Poodle	866
Shiba Inu	672
Shih Tzu	849
Siberian Husky	1238

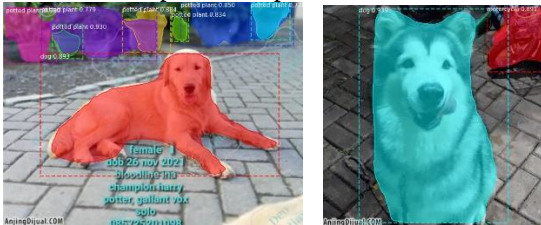
Selain itu, terdapat dataset untuk pengujian yang dapat dilihat pada Tabel 3.

Tabel 3. Data Uji Citra Anjing

Ras	Jumlah
Akita	5
Alaskan Malamute	5
Basset	5
Beagle	5
Border Collie	5
Bulldog	5
Chihuahua	5
Chow	5
Dalmatian	5
Golden Retriever	5
Maltese	5
Pekingese	5
Pomeranian	5
Poodle	5
Shiba Inu	5
Shih Tzu	5
Siberian Husky	5

2.2. Mask R-CNN

Mask R-CNN adalah kerangka kerja yang secara konseptual sederhana, fleksibel, dan umum untuk *instance segmentation* objek. Pendekatan secara efisien mendeteksi objek dalam gambar sekaligus menghasilkan masker segmentasi berkualitas tinggi untuk setiap instansi. Mask R-CNN memperluas Faster R-CNN dengan menambahkan cabang untuk memprediksi topeng objek secara paralel dengan cabang yang ada untuk pengenalan *bounding box* [3]. Mask R-CNN disini menggunakan pre-trained model sehingga tidak perlu melalui proses pelatihan dan dapat langsung digunakan. Contoh hasil penggunaan Mask R-CNN dapat dilihat pada Gambar 3.



Gambar 3. Contoh proses *cropping* dengan *bounding box* pada ras anjing Golden Retriever (Kiri) dan Alaskan Malamute (Kanan)

2.3. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) adalah subtepe dari *deep learning* yang terutama digunakan

untuk aplikasi dalam pengenalan gambar dan ucapan. Lapisan konvolusi mengurangi dimensi gambar yang tinggi tanpa kehilangan informasi yang ada di dalamnya. Itulah mengapa CNN sangat cocok digunakan pada kasus pengenalan gambar dan ucapan.

CNN terdiri atas satu lapisan masukan (input layer), suatu lapisan keluaran (output layer) dan sejumlah lapisan tersembunyi (hidden layer) [4]. Hidden layer CNN terdiri dari *convolutional layer*, *pooling layer*, dan *fully-connected layer*.

Convolution layer adalah adalah blok bangunan inti dari CNN, dimana sebagian besar komputasi terjadi. *Convolution Layer* memerlukan beberapa komponen yaitu *input data*, *filter*, dan *feature map*. *Pooling Layer* atau yang biasa dikenal dengan *downsampling*, melakukan proses untuk pengurangan dimensi, dan mengurangi jumlah parameter dalam input. *Fully-Connected Layer* adalah *layer* yang dimana setiap *node* pada *output layer* sudah saling terhubung dengan *node* yang ada pada *layer* sebelumnya. Di dalam lapisan ini, setiap neuron terhubung ke semua neuron dari lapisan sebelumnya, yang disebut pendekatan Fully Connected (FC).

Konfigurasi untuk model CNN dengan arsitektur VGG-16 dapat dilihat pada Tabel 4.

Tabel 4. Konfigurasi Model

Layer	Input Shape	Output Shape
Input	[(None, None, None, 3)]	[(None, None, None, 3)]
block1_conv_1	(None, None, None, 3)	(None, None, None, 64)
block1_conv_2	(None, None, None, 64)	(None, None, None, 64)
block1_pool (Max Pooling)	(None, None, None, 64)	(None, None, None, 64)
block2_conv_1	(None, None, None, 64)	(None, None, None, 128)
block2_conv_2	(None, None, None, 128)	(None, None, None, 128)
block2_pool (Max Pooling)	(None, None, None, 128)	(None, None, None, 128)
block3_conv_1	(None, None, None, 128)	(None, None, None, 256)

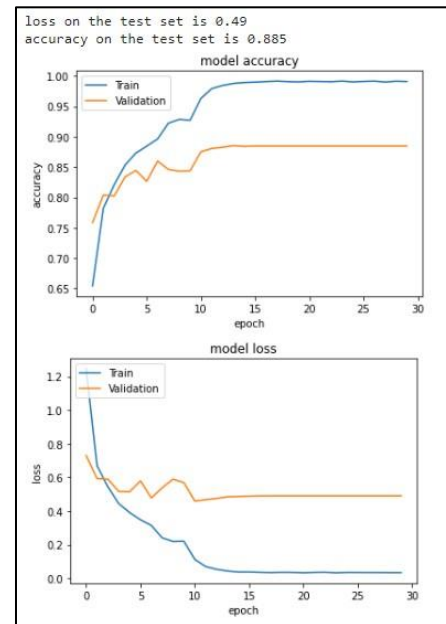
Tabel 4. Lanjutan

block3_conv_2	(None, None, None, 256)	(None, None, None, 256)
block3_conv_3	(None, None, None, 256)	(None, None, None, 256)
block3_pool (Max Pooling)	(None, None, None, 256)	(None, None, None, 256)
block4_conv_1	(None, None, None, 256)	(None, None, None, 512)
block4_conv_2	(None, None, None, 512)	(None, None, None, 512)
block4_conv_3	(None, None, None, 512)	(None, None, None, 512)
block4_pool (Max Pooling)	(None, None, None, 512)	(None, None, None, 512)
block5_conv_1	(None, None, None, 512)	(None, None, None, 512)
block5_conv_2	(None, None, None, 512)	(None, None, None, 512)
block5_conv_3	(None, None, None, 512)	(None, None, None, 512)
block5_pool (Max Pooling)	(None, None, None, 512)	(None, None, None, 512)
global_average_pooling	(None, None, None, 512)	(None, 512)
dense_3	(None, 512)	(None, 1024)
dense_4	(None, 1024)	(None, 1024)
dropout_5	(None, 1024)	(None, 1024)
dense_5	(None, 1024)	(None, 512)
dense_6	(None, 512)	(None, 17)

3. Hasil Pengujian

Pelatihan model CNN dilakukan dengan epoch 30 dan *batch size* sebesar 64 menggunakan data gambar

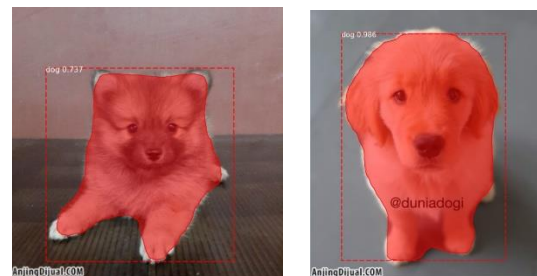
hasil *cropping* dari Mask R-CNN dengan menggunakan arsitektur VGG-16 yang terdapat 16 lapisan konvolusi yang digunakan.. Hasil pelatihan dari model CNN dapat dilihat pada Gambar 4.



Gambar 4. Hasil Pelatihan Model CNN dengan epoch 30

Pada hasil pelatihan model CNN didapatkan hasil pelatihan dengan tingkat akurasi sebesar 88,5%, dengan *loss* sebesar 0.49.

Pengujian untuk Mask R-CNN dilakukan dengan cara melihat *bounding box* yang mendapatkan label 'dog' pada gambar anjing. Contoh pendeteksian dengan hasil yang benar dapat dilihat pada Gambar 5.



Gambar 5. Hasil pengujian Mask R-CNN yang benar

Pengujian model Mask R-CNN dilakukan menggunakan lima citra anjing dari setiap ras anjing dengan jumlah total 85 citra. Hasil pengujian citra anjing dapat dilihat pada Tabel 5.

Tabel 5. Hasil pengujian Mask R-CNN

Ras	Jumlah	Hasil <i>cropping</i> yang benar
Akita	5	2
Alaskan Malamute	5	5
Basset	5	5
Beagle	5	5
Border Collie	5	3
Bulldog	5	5
Chihuahua	5	4
Chow	5	2
Dalmatian	5	5
Golden Retriever	5	5
Maltese	5	5
Pekingese	5	3
Pomeranian	5	2
Poodle	5	3
Shiba Inu	5	3
Shih Tzu	5	3
Siberian Husky	5	3

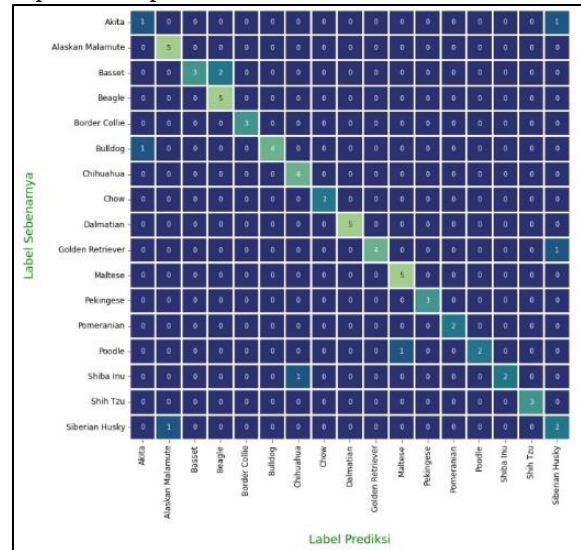
Model Mask R-CNN mendapatkan hasil benar sebesar 63 citra dari 85 citra yang menghasilkan tingkat akurasi sebesar 74%. Hasil *cropping* citra ini akan digunakan untuk pengujian model CNN.

Pengujian pada model CNN dilakukan menggunakan dua jenis data, yaitu data yang merupakan hasil *cropping* yang berhasil dari data uji, dan data tanpa *cropping*. Pengujian data uji dengan citra hasil *cropping* menggunakan data sebanyak 63 citra, sedangkan untuk pengujian data uji dengan data tanpa *cropping* menggunakan 85 citra. Metode pengujian dilakukan dengan menggunakan *Confusion Matrix*. *Confusion Matrix* adalah salah satu *tools* analitik prediktif yang menampilkan dan membandingkan nilai *actual* atau nilai sebenarnya dengan nilai hasil prediksi model yang dapat digunakan untuk menghasilkan metrik evaluasi seperti *Accuracy* (akurasi), *Precision*, *Recall*, dan *F1-Score* atau *F-Measure*.

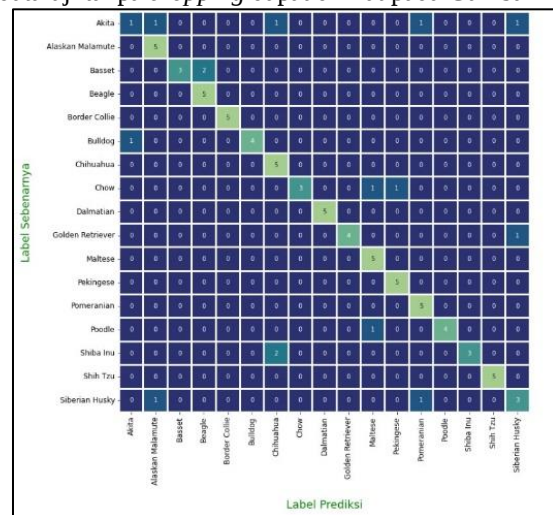
Nilai tersebut didapatkan melalui nilai pada *Confusion Matrix* yaitu *True Positive*, *False Positive*, *True Negative*, *False Negative* dengan keterangan sebagai berikut :

1. *True Positive* : Jumlah data bernilai positif dan prediksi bernilai positif.
2. *False Positive* : Jumlah data bernilai negatif, tetapi prediksi bernilai positif.
3. *False Negative* : Jumlah data bernilai positif, tetapi prediksi bernilai negatif.
4. *True Negative* : Jumlah data bernilai negatif dan prediksi bernilai negatif.

Hasil *Confusion Matrix* dari pengujian menggunakan data uji dengan citra hasil *cropping* dapat dilihat pada Gambar 6.

Gambar 6. Hasil *Confusion Matrix* dari data uji dengan citra hasil *cropping*

Hasil *Confusion Matrix* dari pengujian menggunakan data uji tanpa *cropping* dapat dilihat pada Gambar 7.

Gambar 7. Hasil *Confusion Matrix* dari data uji tanpa *cropping*

Perhitungan untuk *Confusion Matrix* mendapatkan nilai akurasi, presisi, recall, dan F1-score dengan menggunakan rumus sebagai berikut [5] :

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

$$F1\ Score = 2x \frac{Presisi \times Recall}{Presisi + Recall} \times 100\% \quad (4)$$

Hasil akurasi, presisi, recall, dan F1-score model CNN dengan menggunakan data hasil *cropping* dan menggunakan data tanpa *cropping* dapat dilihat pada Tabel 6.

Tabel 6. Hasil *Confusion Matrix*

	Akurasi	Presisi	Recall	F1-Score
Data Hasil <i>Cropping</i>	87%	89%	86%	87%
Data Tanpa <i>Cropping</i>	82%	85%	82%	81%

4. Kesimpulan

Kesimpulan yang didapatkan dari percobaan yang telah dilakukan pada citra anjing dengan menggunakan *Mask R-CNN* dan CNN adalah sebagai berikut :

- Model Mask R-CNN dapat digunakan untuk melakukan proses *cropping* dengan baik dengan tingkat akurasi sebesar 74%.
- Hasil *cropping* dari Mask R-CNN memengaruhi hasil pengenalan ras anjing pada CNN, sehingga hasil *cropping* yang semakin baik akan mendapatkan hasil pengenalan ras anjing yang lebih akurat.
- Metode CNN dengan menggunakan arsitektur VGG-16 dapat melakukan pengenalan ras anjing dengan baik dan akurat dengan tingkat akurasi mencapai 87% dengan data hasil *cropping*, dan 82% dengan menggunakan data tanpa *cropping*.

Dalam pengembangan untuk model Mask R-CNN dapat ditingkatkan lagi agar sistem dapat mendeteksi gambar anjing lebih akurat. Jumlah untuk ras anjing dapat diperbanyak agar jangkauan pemakaian sistem ini juga semakin luas.

REFERENSI

- [1] Tombesi-Walton, David; Tombesi-Walton, Sylvia, 2015, "101 Essential Tips: Dog Care", New York: DK Publishing.
- [2] Lauw, Kevin, Oktovio; Santoso, Leo, Willyanto; Intan, Rolly, 2020 "Identifikasi Jenis Anjing Berdasarkan Gambar Menggunakan Convolutional

Neural Network Berbasis Android", JURNAL INFRA, Vol 8, Nomor 2.

[3] K. He, G. Gkioxari, P. Dollar, and R. Girshick, 2020, "Mask R-CNN", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no.2

[4] Gerald, Calvin; Lubis, Chairisni, 2020, "Pendeteksian dan Pengenalan Jenis Mobil Menggunakan Algoritma You Only Look Once dan Convolutional Neural Network", Jurnal Ilmu Komputer dan Sistem Informasi, Vol. 8, Nomor 2.

[5] Danny; Lina, 2021, "Pendeteksian Sel Darah Putih dari Citra Preparat dengan Convolutional Neural Network", Jurnal Ilmu Komputer dan Sistem Informasi, Vol. 9, Nomor 1.

Sandy Danish Arkansa, saat ini sebagai mahasiswa Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara

Chairisni Lubis Dra., M.Kom, memperoleh gelar Dra dari Universitas Indonesia. Kemudian memperoleh gelar M.Kom dari Universitas Indonesia. Saat ini sebagai Dosen program studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara