

Implementasi Metode Faster Region Convolutional Neural Network (Faster R-CNN) Untuk Pengenalan Jenis Burung Lovebird

Fino Charli¹, Hadi Syaputra^{*2}, Muhammad Akbar³, Siti Sauda⁴, Febriyanti Panjaitan⁵

^{1,2,3,4,5}Informatics Departement , Bina Darma University, Palembang, Indonesia

Email: ¹finocharli77@gmail.com, ²hadisyaputra@binadarma.ac.id,

³muhammadakbar@binadarma.ac.id, ⁴sitisauda@binadarma.ac.id,

⁵febriyantipanjaitan@binadarma.ac.id

Abstrak

Lovebird merupakan burung sosial yang hidup berkelompok di alam bebas. Burung dari genus Agapornis ini ukuran tubuhnya relative mungil bila dibanding burung berparuh bengkok lainnya. Panjangnya sekitar 13-17 cm dengan bobot 30-60 gram. Ragam dan jenis burung lovebird sangat banyak namun bila dilihat dengan pendekatan taksonomi, hanya terdapat 9 spesies burung lovebird. Penelitian metode Faster Region Convolutional Neural Network ini digunakan untuk mengenali jenis burung lovebird guna mencari tahu motif atau warna asli dari jenis burung lovebird tertentu berdasarkan tingkat akurasi dari hasil deteksi objek. Faster R-CNN merupakan pengembangan Convolutional Neural Network (CNN) yang ditemukan oleh Ross Girshick. Sampai saat ini metode ini masih terus dikembangkan dan digunakan untuk pengenalan objek atau objek recognition. Penelitian ini menggunakan data set burung lovebird dengan total 808 gambar dan 8 (kelas) yaitu Agapornis Personata, Agapornis Fischeri, Agapornis Liliana, Agapornis Nigrigenis, Agapornis Roseicollis, Agapornis Taranta, Agapornis Pullarius, Agapornis Canus. Hasil dari proses learning didapatkan akurasi sebesar 96% dengan loss sebesar 0,2% . sehingga dapat disimpulkan bahwa metode Faster R-CNN dapat diterapkan pada pengenalan jenis burung lovebird.

Kata Kunci: Faster Regional Neural Network, object detection, convolutional neural network, image, deep learning.lovebird

1. PENDAHULUAN

Deep learning merupakan suatu kelas dari teknik pembelajaran mesin yang memanfaatkan banyak lapisan dalam pemrosesan informasi untuk ekstraksi dan informasi. *Deep learning* lahir pada tahun 2006 diperkenalkan oleh geoffrey hinton untuk menjelaskan algoritma baru yang dapat membuat komputer melihat(Purnama, 2019). Deep learning mulai dikenal pada kompetisi pengenalan citra *ImageNet Large Scale Visual Recognition Competition* (ILSVRC) tahun 2012(Suyanto, Kurniawan Nur Ramadhani, Satria Mandala, 2019), teknik ini menjadi pemenang pertama di

(ILSVRC) dengan akurasi yang jauh lebih tinggi dibanding *shallow learning* (machine learning) konvensional.

Ada banyak Teknik pengenalan objek dengan *deep learning* ,diantaranya Region based Convolutional Neural Network (R-CNN), Fast R-CNN, Faster R-CNN, dan You Only Look Once (YOLO) (Gandhi, 2018). Faster R-CNN merupakan Teknik pengenalan objek oleh mesin yang dikembangkan oleh Ross Girshick dkk pada tahun 2015 (Gandhi, 2018). Faster R-CNN memiliki 3 modul utama yang pertama *deep fully convolutional network* yang mengusulkan region, modul kedua adalah detector Fast R-CNN yang menggunakan region yang diusulkan dan RPN yang berfungsi mengarahkan Fast R-CNN untuk melihat(Ren et al., 2016).

Untuk melakukan pengenalan objek terlebih dahulu kita mengumpulkan dataset, dataset merupakan kumpulan citra yang kita kumpulkan dan diberi label oleh manusia yang kemudian di proses oleh mesin untuk sebagai pembelajaran sebelum melakukan klasifikasi(Purnama, 2019). burung lovebird merupakan hewan asli daratan Afrika dan memiliki 9 (sembilan) jenis diantaranya : Agapornis personata, Agapornis fischeri, Agapornis lilianae, Agapornis nigrigenis, Agapornis roseicollis, Agapornis taranta, Agapornis pullarius, Agapornis canus, dan Agapornis canus. Lovebird merupakan burung social berparuh bengkok asli daratan Afrika. Dialam bebas burung lovebird dapat terdiri dari 5 hingga 20 ekor per kelompok. Sedangkan burung dewasa akan hidup berpasangan dengan sangkar tak jauh dari koloninya, burung dari genus Agapornis ini ukuran tubuhnya relative mungil dengan bobot 30- 60 gram. Mutasi warna yang banayak terjadi pada burung Lovebird yang ada di Indonesia adalah sebuah hasil rekayas genetik dengan mengawin silangkan burung Lovebird yang tidak sejenis sehingga banyak dijumpai warna yang berbeda dalam satu jenis burung. Pada penelitian ini penulis akan menggunakan Algoritma Faster Region Convolutional Neural Network (Faster R-CNN) untuk mengenali jenis burung Lovebird pada sebuah citra input.

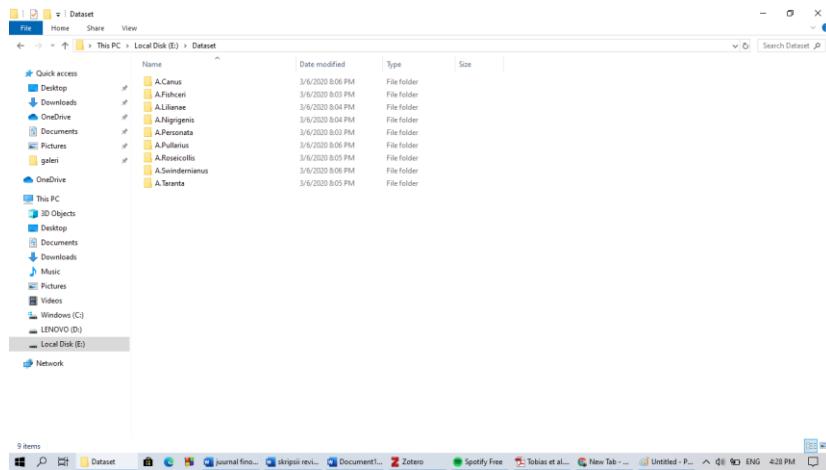
2. METODOLOGI PENELITIAN

2.1. Tahapan Penelitian

1) Pengumpulan Gambar

Data yang diambil dan dikumpulkan hanyalah 8 (delapan) jenis dari 9 (Sembilan) jenis lovebird yang ada. Masing- masing kelas berjumlah 100 gambar berformat jpg. hal ini dilakukan agar data uji dan data latih mendapat bobot training yang sama sehingga menghasilkan bias rendah pada data latih dan data uji. Mengingat jenis Lovebird jenis Agapornis swindernianus susah

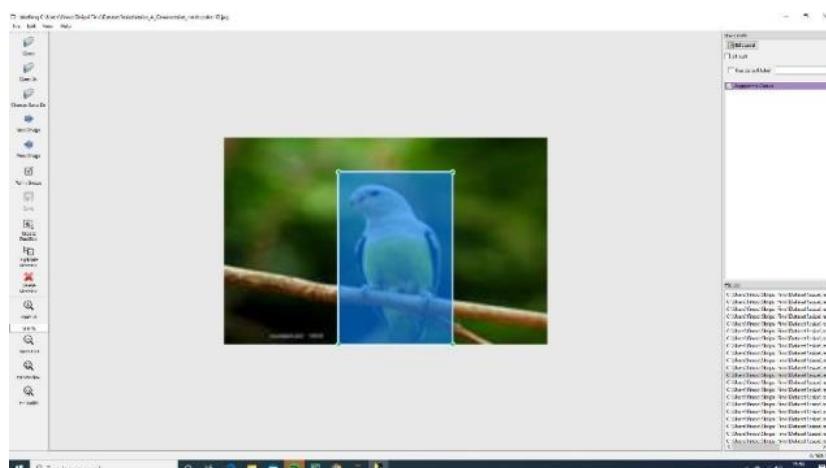
untuk ditangkarkan dan sangat Langkah sehingga sulit untuk mendapatkan 100 buah gambar jenis burung ini.



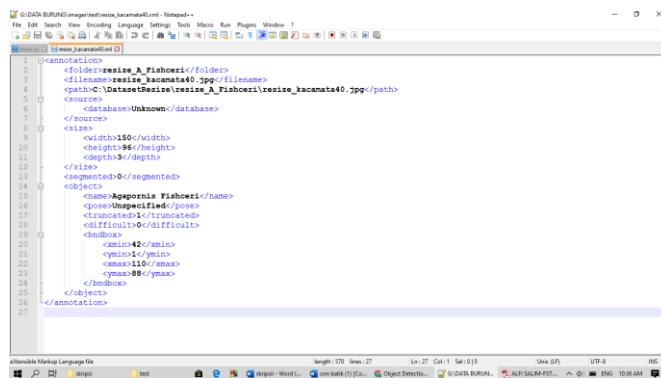
Gambar 2. folder per serratus gambar

2) Pelabelan data

Pemberian label pada gambar atau anotasi ini bertujuan untuk mendapatkan ciri khusus yang merepresentasikan masing-masing objek yang terdapat dalam gambar. Ciri khusus tersebut nantinya akan dijadikan learning dalam proses training.



Gambar 3. Pelabelan data



```

<annotation>
  <id>1</id>
  <filename>resize_A_Fishceri40.jpg</filename>
  <path>C:\Dataset\Resize\resize_A_Fishceri\resize_kacanata40.jpg</path>
  <source>
    <database>Unknown</database>
  </source>
  <size>
    <width>150</width>
    <height>96</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>Apogonis_Fishceri</name>
    <pose>Unspecified</pose>
    <truncated>1</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>42</xmin>
      <ymin>79</ymin>
      <xmax>116</xmax>
      <ymax>98</ymax>
    </bndbox>
  </object>
</annotation>

```

Gambar 4. Hasil Anotasi

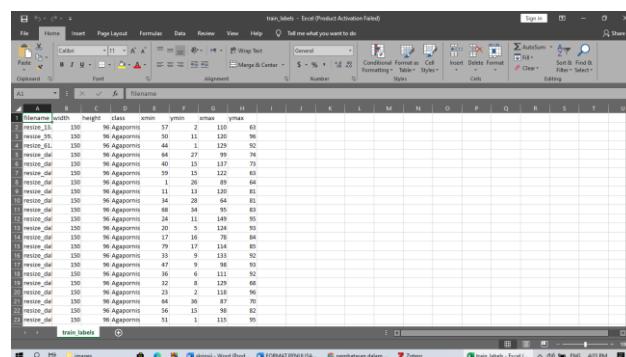
Hasil anotasi akan berupa koordinat objek dalam bentuk file xml seperti gambar 4.

3) Membagi Dataset

Pada tahap ini dataset dibagi menjadi dua bagian yaitu data train dan data tes. Sesuai dengan konsep split dataset, data train harus lebih banyak dari data test. Pada penelitian ini pembagian dataset merujuk pada penelitian yang pernah dilakukan sebelumnya (Susanto et al., 2019). Dataset dibagi menjadi 80% untuk data train dan 20% untuk data test.

4) Konversi file xml ke .csv

Konversi file xml menjadi file .csv bertujuan untuk membuat himpunan data yang semula tidak rapi menjadi bentuk tabel yang berisi fitur yang akan dipelajari dalam proses konvolusi. Berikut hasil konversi file xml hasil anotasi yang sudah dikonversi kedalam bentuk .csv



resize_33	150	96	Apogonis	37	2	110	60
resize_35	150	96	Apogonis	50	11	130	96
resize_36	150	96	Apogonis	44	11	120	52
resize_37	150	96	Apogonis	64	27	95	36
resize_38	150	96	Apogonis	40	15	117	79
resize_39	150	96	Apogonis	79	13	122	60
resize_40	150	96	Apogonis	1	26	89	64
resize_41	150	96	Apogonis	11	11	127	81
resize_42	150	96	Apogonis	34	28	64	81
resize_43	150	96	Apogonis	68	34	95	83
resize_44	150	96	Apogonis	24	11	147	95
resize_45	150	96	Apogonis	20	5	124	93
resize_46	150	96	Apogonis	17	16	78	84
resize_47	150	96	Apogonis	79	17	118	86
resize_48	150	96	Apogonis	33	9	135	52
resize_49	150	96	Apogonis	47	9	98	93
resize_50	150	96	Apogonis	36	4	111	56
resize_51	150	96	Apogonis	32	8	129	68
resize_52	150	96	Apogonis	23	2	118	96
resize_53	150	96	Apogonis	54	34	87	70
resize_54	150	96	Apogonis	56	15	98	82
resize_55	150	96	Apogonis	31	1	115	95

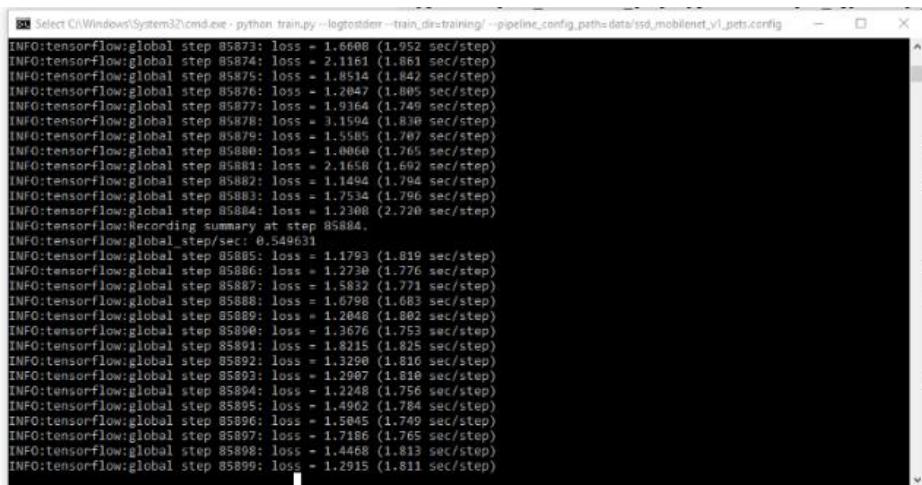
Gambar 5. Hasil konversi file xml

a. Mengenerate file ke bentuk TFRecord

Setelah mengumpulkan semua fitur dalam bentuk csv maka Langkah selanjutnya ialah mengubah file .csv yang berisi fitur2 tadi kedalam bentuk biner agar dapat diolah menggunakan tensorflow

b. Training Data

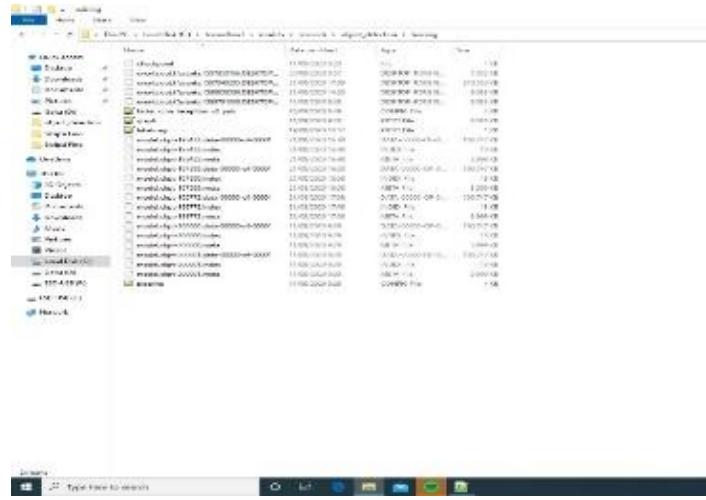
Training data adalah proses yang memakan waktu lama. Training data dilakukan dengan cara menjalankan sebuah script yang berisi source code dan konfigurasi yang telah kita setting pada komputer kita. Training dilakukan sebanyak 2000000 step yang mana setiap step ditampilkan hasil dari pembelajaran (learning()) sistem. Setiap step akan menampilkan sebuah variabel yang disebut loss. Loss menunjukkan apakah model berlatih dengan baik atau tidak. Untuk lebih jelasnya kita dapat melihatnya dalam Inference graph yang merupakan catatan atau rekaman selama kita melakukan training.



Gambar 6. Proses training

c. Menyimpan Hasil frozen graph

selama proses training, tensorflow akan merekam semua proses yang terjadi pada saat proses berlangsung, file terekam dalam bentuk .ckpt (cekpoint) dalam folder training. File yang tersimpan dan merekam step terakhir akan dikonversi menjadi sebuah model hasil pelatihan dengan format protobuf dengan ekstensi file ".pb"



Gambar 7. model training tensorflow dan checkpoint

d. Testing

Setelah menghasilkan model yang dilatih berdasarkan chechkpoint terakhir dan frozen graph. Maka model sudah siap untuk diuji untuk menunjukkan keakuratan model dalam pengklasikiasian jenis burung Lovebird. Pengujian dilakukan dengan menjalankan `test_video.py` di Pycharm yang sudah dimasukkan video berisi gambar burung Lovebird. Hasilnya adalah video yang sama namun dengan bounding box dan nama jenis burung lovebird dan tingkat akurasinya

2.2 Burung Lovebird

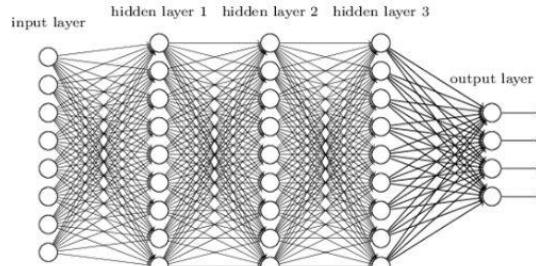
Burung Lovebird yang nama latinnya Agapornis, masih satu family dengan burung kaka tua. Dari 9 (Sembilan) jenis spesies 8 diantaranya berasal dari daratan afrika. Hanya 1 Agapornis Canus yang berasal dari pulau madagaskar sebuah pulau dekat didekat benua Afrika (Tim Karya Tani Mandiri, 2018).

Burung Lovebird bereproduksi dengan bertelur. Dalam setiap kelahiran menghasilkan 3-6 telur dengan lama pengeraman 22 hari. Seiring berjalanannya waktu, lovebird sering disilangkan untuk menghasilkan warna mutasi baru. Kondisi alam yang disukai burung lovebird adalah lahan kering dan iklim yang terik. Pada dfasarnya lovebird hanya terdiri dari sembilan jenis saja, namun seiring berjalanannya waktu, lovebird sering disilangkan untuk menghasilkan warna mutasi bulu. Dapatkah metode Faster Region Convolutional Neural Network (faster R-CNN) dapat memprediksi jenis burung Lovebird dalam kasus ini.

2.3 Deep Learning

Deep learning adalah suatu teknik pada machine learning yang memanfaatkan banyak layer pada pengolahan informasi non-linier yang digunakan untuk melakukan ekstraksi fitur pengenalan pola dan klasifikasi (Deng dan Yu, 2014).

Deep neural network

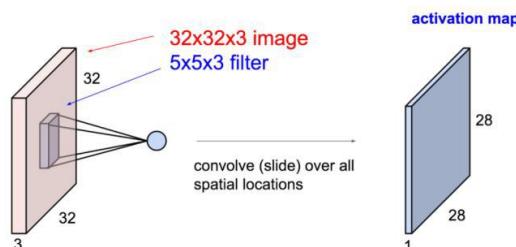


Gambar 8. Deep Neural Network

2.4 Convolutional Neural Network

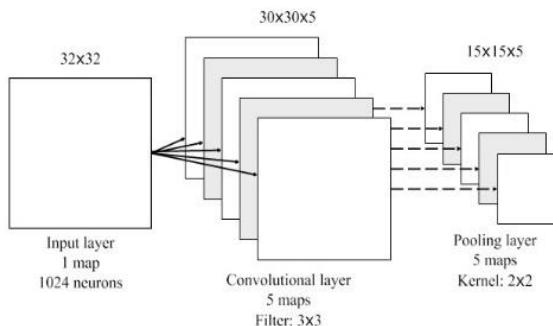
Convolutional Neural Network merupakan salah satu jenis neural network yang biasanya digunakan dalam pengolahan data image (Santoso & Ariyanto, 2018). Arsitektur CNN terdiri atas satu lapisan (input layer), satu lapis keluaran (output layer), dan lapis tersembunyi (hidden layer). Lapis tersembunyi umumnya berisi *convolutional layer*, *pooling layer*, *normalization layer*, *ReLU layer*, *fully connected layer*, dan *loss layer* (Suyanto, Kurniawan Nur Ramadhan, Satria Mandala, 2019). Pada proses training terdapat 3 tahapan yaitu convolutional layer, pooling layer, dan fully connected layer (Santoso & Ariyanto, 2018). Berikut ulasan 3 layer tersebut:

- 1) Convolutional layer, terdiri dari neuron yang tersusun sedemikian rupa sehingga membentuk sebuah filter dengan tinggi dan panjang (pixels). Tiap filter memiliki nilai dan nilai dari filter inilah yang menjadi parameter yang akan di-update dalam proses learning.



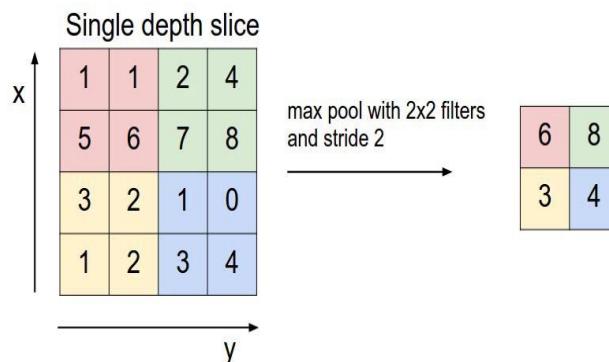
Gambar 9. Convolutional Layer

- 2) Pooling layer, pooling layer berfungsi menjaga ukuran data Ketika convolution dilakukan, yaitu dengan melakukan reduksi sampel (down sampling) contoh gambar 2.2 dengan pooling, kita dapat merepresentasikan data menjadi lebih kecil, mudah dikelola, dan mudah mengontrol overfitting.



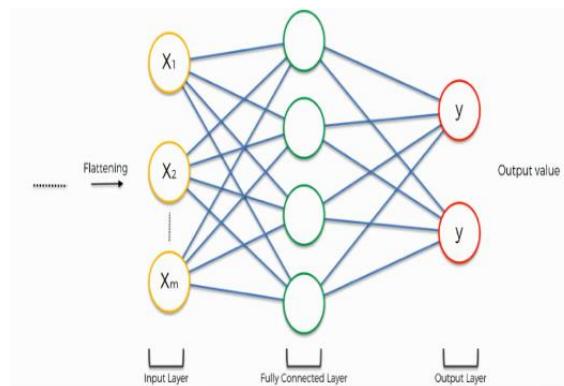
Gambar 10. proses pooling untuk mereduksi data

Pada umumnya proses pooling dilakukan menggunakan max pooling atau average pooling. Caranya adalah dengan memilih nilai maksimum atau nilai rata-rata dalam suatu area tertentu contoh gambar 10.



Gambar 2.4 Teknik Max pooling

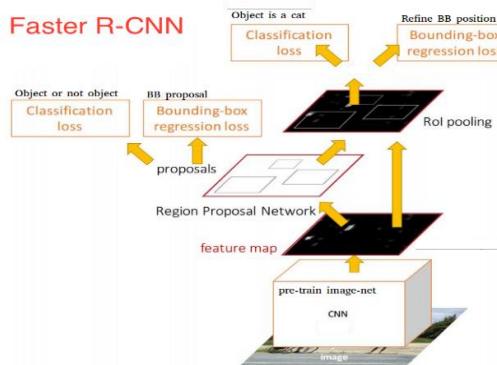
- 3) Fully Connected layer, pada lapisan yang terhubung secara penuh (fully connected layer), setiap neurons memiliki koneksi penuh ke semua aktivasi dalam lapisan sebelunya. Hal ini sama persis dengan MLP, yaitu komputasi menggunakan suatu perkalian matriks yang diikuti dengan bias offset



Gambar 11. Fully Connected Layer

2.4 Faster Region Convolutional Neural Network

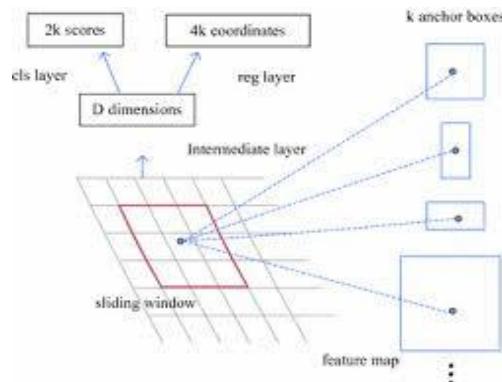
Faster R-CNN merupakan pengembangan dari Fast R-CNN. Sebagai generasi terbaru dari metode objek setection menunjukkan hasil yang mengesankan pada objek deteksi



Gambar 12. Arsitektur dari Faster R-CNN (Le, 2020)

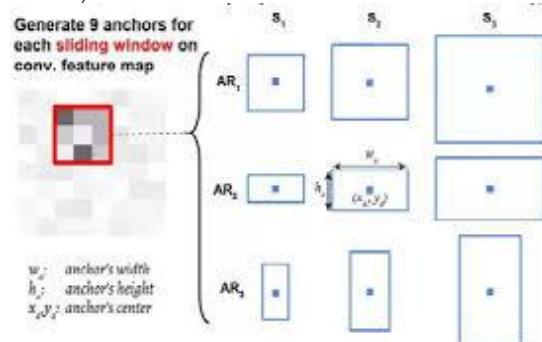
Gambar 12. menunjukkan arsitektur dari faster R-CNN. Faster R-CNN memiliki Langkah-langkah sebagai berikut :

1. Region Proposal Network (RPN) yaitu tugas cepat yang berguna untuk mencari pada input gambar kemungkinan lokasi dari objek. Posisi dari objek yang ada pada gambar memiliki kemungkinan objek dibatasi dari wilayah yang diketahui sebagai region of interest (ROI). RPN mengambil gambar dari berbagai ukuran sebagai input dan output sekumpulan proposal objek persegi panjang, masing -masing dengan skor objektivitas.



Gambar 13 Region Proposal Network (RPN)

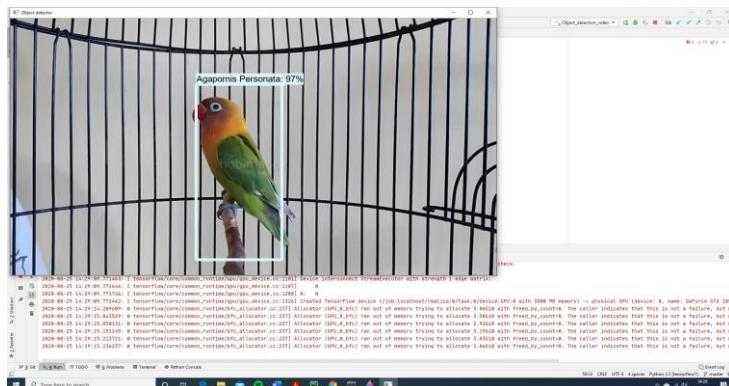
Pada Region Proposal Network awalnya gambar dimasukkan ke dalam jaringan Convolutional Neural Network. Gambar input diteruskan ke jaringan convolutional layer terakhir yang menampilkan feature map (Gilbert, 2019). Sliding window ditempatkan pada setiap bagian dari feature map. Sliding window mask biasanya diambil dari ukuran mask $n \times n$. Sesuai dengan setiap sliding window, secara bersamaan memprediksi beberapa region proposal, dimana jumlah proposal maksimum yang mungkin untuk setiap lokasi dilambangkan sebagai k . Layer reg memiliki output $4k$ yang mengkodekan koordinat kotak k , dan layer cls menghasilkan $2k$ yang memperkirakan probabilitas objek atau tidak objek untuk setiap proposal. Setiap anchor diposisikan di tengah dari sliding windows. Secara standar menggunakan 3 skala dan 3 rasio aspek yang menghasilkan $k=9$ anchor pada setiap sliding windows (Gilbert, 2019). Klasifikasi menunjukkan probabilitas 0 atau 1 yang menunjukkan apakah wilayah tersebut berisi objek atau tidak dengan $p^* = 1$ jika $IoU > 0,7$, $p^* = .1$ jika $IoU < 0,3$, $p^* = 0$ jika selain itu (Gilbert, 2019).



Gambar 14 Tiga Anchor dengan Aspect ratio dan Scaling

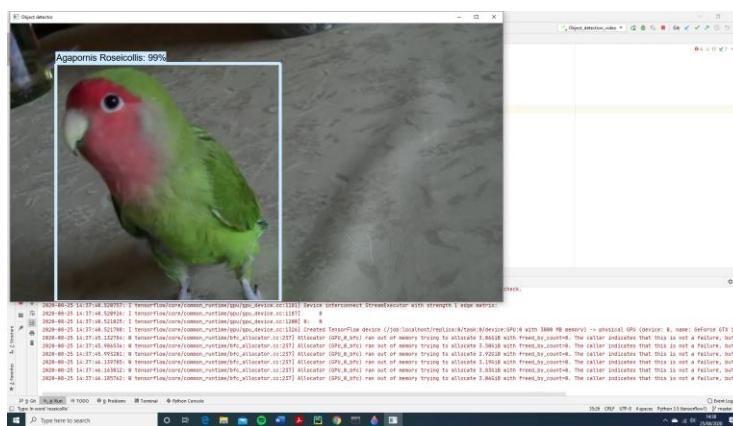
3. HASIL DAN PEMBAHASAN

Setelah melakukan proses training menggunakan algoritma deep learning dengan model Faster R-CNN dapat kita ketahui bahwa training dilakukan hingga nilai loss konsisten dibawah 0,05. Sehingga mampu menghasilkan akurasi yang cukup tinggi, hal tersebut dapat dilihat dari gambar berikut



Gambar 15 Agapornis Personata

Gambar diatas merupakan hasil capture dari video yang dijalankan dengan test_video.py menggunakan pycharm. Hasil prediksi bertujuan untuk menentukan dua hal, apakah ada burung Lovebird dalam suatu citra yang diberikan ? jika ada jenis burung Lovebird apa yang ada pada citra ?.



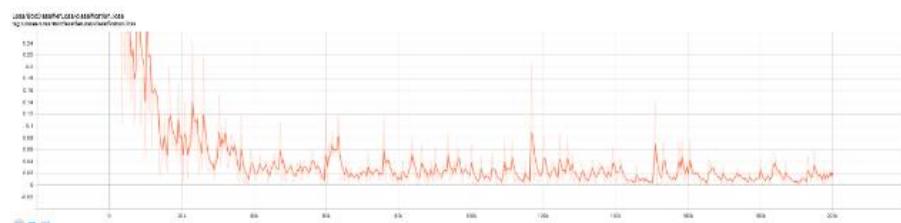
Gambar 16. Agapornis Roseicollis

Dari gambar 15 dan 16, dapat dilihat bounding box mengenali jenis burung lovebird hingga mampu memberikan persentase dari deteksi objek dari video

yang diberikan. Terlihat hasil akurasi sebesar 97% pada Agapornis personata dan 99% untuk video yang bergambar lovebird jenis Agapornis Roseicollis.

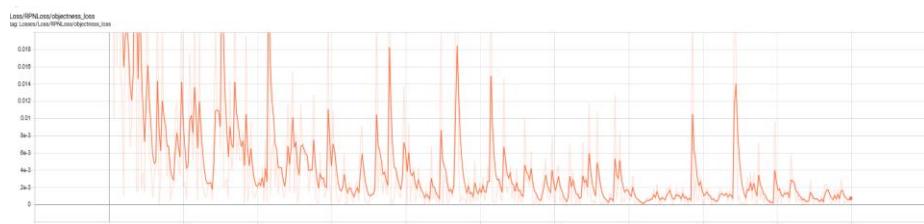
3.1 Inference graph

Seperti yang telah kita bahas sebelumnya proses training dapat diamati jika kita menggunakan framework tensorflow. Hal ini dikarenakan tools tensorboard yang memuat inference graph. Adapun beberapa grafik utama yang menjadi acuan penulis dalam mengamati proses training diantaranya grafik box classifier loss dan grafik RPN loss.



Gambar 17. Grafix box classifier loss/classification

Grafik ini menunjukkan kesalahan yang dilakukan model dalam mengklasifikasi objek yang terdeteksi menjadi berbagai macam kelas lain. Dengan demikian model mendapatkan nilai kesalahan rata-rata dibawah 0,03 dalam mendeteksi objek menjadi kelas yang seharusnya selama proses training



Gambar 18 Grafik RPN loss/objectness loss

Pada grafik ini tensorboard menginformasikan kesalahan classifier dalam mengklasifikasikan sebuah objek atau sebuah latar belakang. Selama proses trainig tensorboard akan merekam semua proses yang terjadi pada saat proses berlangsung. File terekam dalam file bernama .ckpt (checkpoint) dalam folder training. File yang tersimpan dan merekam step step terakhir akan dikonversi menjadi sebuah model hasil pelatihan dengan format protobuf dengan ekstensi file .pb

4. KESIMPULAN

Berdasarkan hasil analisis yang telah dilakukan, maka dapat diperoleh beberapa kesimpulan sebagai berikut :

- 1) Ragam warna yang terdapat pada jenis burung Lovebird sudah dapat menjadi bahan poembeda antara satu jennies ke jenis lainnya.
- 2) Metode Faster Region Convolutional Neural Network dapat memprediksi jenis burung lovebird.
- 3) Tingkat akurasi model yang didapatkan dari hasil pendektsian jenis burung lovebird pada suatu citra digital menggunakan Faster region convolutional Neural Network berkisar 78% hingga 99%.

DAFTAR PUSTAKA

- Gandhi, R. (2018, July 9). *R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms*. Medium. <https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e>
- Gilbert, N. (2019). *Implementasi Faster Region-Based Convolutional Neural Network pada Klasifikasi Bentuk Diagram iStar 2.0 untuk Requirements Modeling* [Bachelor_thesis, Universitas Multimedia Nusantara]. <http://kc.umn.ac.id/10245/>
- Le, J. (2020, January 29). *The 5 Computer Vision Techniques That Will Change How You See The World*. Medium. <https://heartbeat.fritz.ai/the-5-computer-vision-techniques-that-will-change-how-you-see-the-world-1ee19334354b>
- Purnama, B. (2019). *Pengantar Machine Learning Konsep dan Praktikum dengan Contoh Latihan Berbasis R dan Python*. Informatika.
- Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *ArXiv:1506.01497 [Cs]*. <http://arxiv.org/abs/1506.01497>
- Santoso, A., & Ariyanto, G. (2018). IMPLEMENTASI DEEP LEARNING BERBASIS KERAS UNTUK PENGENALAN WAJAH. *Emitor: Jurnal Teknik Elektro*, 18(01), 15–21. <https://doi.org/10.23917/emitor.v18i01.6235>
- Susanto, K. N., Gunadi, K., & Setyati, E. (2019). Pengenalan Karakter pada Plat Nomor Indonesia dengan Tilt Correction dan Metode Faster R-CNN. *Jurnal Infra*, 7(1), 1–7.
- Suyanto, Kurniawan Nur Ramadhani, Satria Mandala. (2019). *Deep Learning Modernisasi Machine Learning Untuk Big Data*. Informatika.
- Tim Karya Tani Mandiri. (2018). *Rahasia Sukses Budidaya Burung Berkicau. Nuansa Aulia*.