



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB III

METODE PENELITIAN

3.1 Objek Penelitian

Objek penelitian pada penelitian ini adalah *Mask R-CNN*. *Mask R-CNN* merupakan sebuah *framework* untuk melakukan segmentasi objek. Pendekatan yang ingin di capai oleh *Mask R-CNN* adalah bagaimana cara untuk melakukan deteksi objek dalam sebuah gambar secara efisien serta pada saat yang bersamaan dapat menghasilkan *mask* atau lapisan yang dapat menutupi objek yang di deteksi (He, Gkioxari, Dollar, & Girshick, 2017).

Mask R-CNN merupakan perkembangan dari metode sebelumnya, yaitu *Faster R-CNN*, dengan cara menambahkan fitur yang mampu melakukan prediksi terhadap dimana kemungkinan sebuah objek berada didalam *bounding box* atau kotak pembatas dari sebuah objek dan memberikan *mask* atau lapisan pada objek tersebut (He, Gkioxari, Dollar, & Girshick, 2017).

Mask R-CNN adalah sebuah *framework* yang dapat diadaptasi untuk mendeteksi objek tertentu, sesuai dengan kebutuhan yang diperlukan, oleh karena itu *Mask R-CNN* dapat digunakan untuk melakukan deteksi terhadap manusia pada sebuah gambar (He, Gkioxari, Dollar, & Girshick, 2017).

3.2 Pengumpulan Data

3.2.1 Sumber Data

Sumber data di dapatkan sumber yang menyediakan *dataset* yang mengandung gambar manusia. Sumber gambar didapatkan dari:

1. *INRIA Person Dataset* (Dalal, 2019)
2. *MPII Human Pose Dataset* (Andriluka, Pishchulin, Gehler, & Schiele, Bernt, 2014)

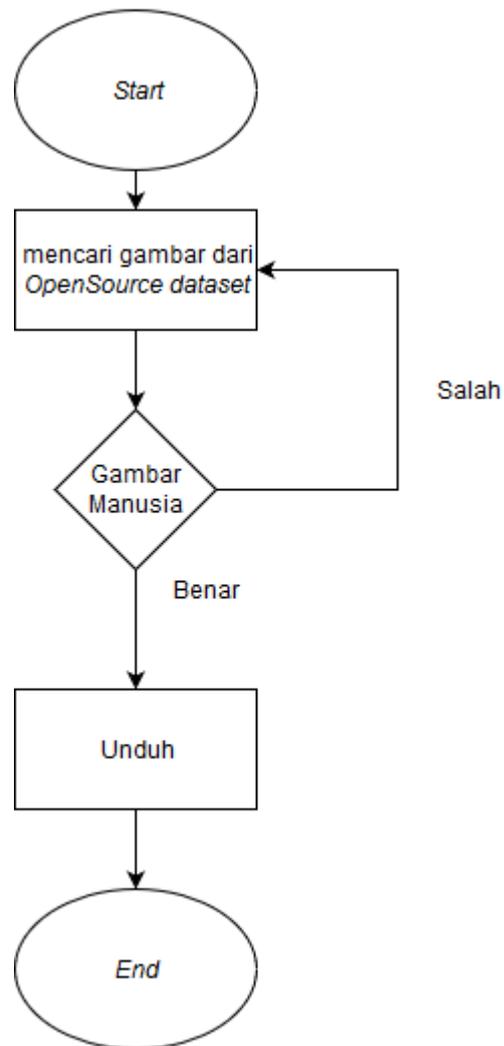
Dari keseluruhan data yang terkumpul, terdapat 30140 buah gambar yang mengandung manusia, dalam berbagai posisi, pencahayaan, visibilitas, serta kualitas yang beragam yang dapat digunakan untuk *training*. Akan tetapi, tidak keseluruhan data yang terkumpul akan digunakan untuk *training*.

Pada penelitian ini, 740 gambar yang dipilih yang akan digunakan untuk keperluan *training*. 740 gambar di ambil berdasarkan *paper* “*ADAPTING MASK-RCNN FOR AUTOMATIC NUCLEUS SEGMENTATION*” (Johnson, 2018). *Paper* tersebut di gunakan karena memiliki *model* yang serupa dengan yang digunakan pada penelitian ini, dan memiliki kesamaan tujuan, yaitu mencoba melakukan segmentasi pada sebuah objek. Pengumpulan data sebanyak 30140 buah, memberikan kesempatan untuk melihat, dan memilih, data manusia dalam berbagai posisi, pencahayaan, dan visibilitas yang diperkirakan cocok untuk ditambahkan ke dalam proses *training*.

3.1.2 Proses Pengumpulan Data

Tahapan dalam melakukan pengumpulan data dapat dilihat pada Gambar

3.1



Gambar 3.0.1 Flowchart Pengumpulan Data

3.2 Metode Penelitian

3.2.1 Hipotesis Penelitian

Hipotesis penelitian dibangun berdasarkan penelitian terdahulu. Pada penelitian ini, terdapat satu buah hipotesis, yaitu “Apakah *Mask R-CNN* yang menggunakan *weights* yang sudah di *trained* dengan gambar kualitas rendah lebih

baik daripada *pretrained weights mask_rcnn_coco* dalam melakukan deteksi manusia pada gambar dengan kualitas rendah”.

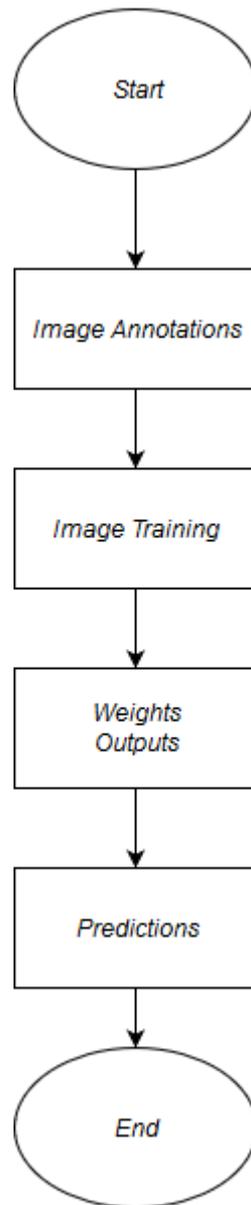
Hipotesis akan diuji dengan melakukan perbandingan deteksi dengan dataset yang sama, menggunakan model *Mask R-CNN* dengan *pretrained weights mask_rcnn_coco*

3.2.2 Proses Penelitian

Penelitian ini akan menggunakan *Mask R-CNN* (He, Gkioxari, Dollar, & Girshick, 2017) untuk melatih model menggunakan *dataset* yang telah disediakan. *Mask R-CNN* digunakan karena *Mask R-CNN* dapat memberikan *instance segmentation*, sehingga dapat menampilkan objek dengan *binary highlight* pada object untuk memudahkan klasifikasi, dan sejalan dengan tujuan penelitian, yaitu melakukan *tracking* pada manusia pada gambar dengan kualitas rendah.

3.2.3 Alur Penelitian

Alur penelitian dari penelitian ini dapat dilihat pada Gambar 3.2



Gambar 3.2 Alur Penelitian

Berikut penjelasan mengenai alur yang ditampilkan pada Gambar 3.2

Sebelum model dapat melakukan *training*, pertama-tama data yang telah di kumpulkan harus di proses terlebih dahulu. Proses tersebut di sebut dengan anotasi, atau penambahan *mask* terhadap objek manusia yang ingin kita deteksi.

A. *Image Annotations*

Proses labelisasi data dilakukan dengan melakukan *masking* terhadap gambar manusia menggunakan bantuan alat bernama *VGG Image Annotations (VIA)*. *Mask RCNN* membutuhkan data yang telah di berikan anotasi terlebih dahulu sehingga program dapat lebih cepat dalam mencari objek yang dimaksud karena proses anotasi sebelumnya telah menandakan *region of interest* yang mengandung objek manusia, yaitu objek yang ingin dipelajari dari sebuah gambar.

VIA digunakan karena mampu memberikan efisiensi terhadap pengerjaan proses data. Hal ini disebabkan oleh karena output yang dihasilkan oleh *VGG* merupakan satu buah file *JSON* di mana program serupa mengeluarkan file berbentuk gambar dengan *masking* yang di berikan, sehingga menggharuskan proses lebih lanjut terhadap hasil output yang di berikan oleh program tersebut.

Proses labelisasi data yang digunakan dalam penelitian ini dilakukan secara manual, menggunakan bantuan alat *VIA* yang dilakukan seorang diri, tanpa bantuan dari orang lain.

Selanjutnya, setelah proses anotasi selesai di lakukan, data yang telah di berikan *mask* tersebut di kelompokkan untuk kemudian di bagi menjadi dua buah kategori, yaitu gambar untuk keperluan validasi, dan gambar untuk keperluan training, dimana masing-masing kategori, terbagi kembali menjadi 2 bagian, yaitu,

bagian yang mengalami penyuntingan, serta bagian yang tidak di berikan suntingan. Untuk keterangan lebih jelas dapat melihat Gambar 4.2 pada bab selanjutnya.

B. *Image Training*

Dalam proses *training* data, data yang ada akan *input* ke dalam program dan program akan melakukan analisis data menggunakan *keras* sebagai antarmukanya, sehingga pembobotan dalam layer-layer di dalam *keras* menghasilkan *output* berupa model yang selanjutnya dapat digunakan untuk melakukan prediksi.

Training dilakukan dengan cara mengolah dataset yang telah disediakan menggunakan model *Mask R-CNN*. Pertama, gambar akan melewati *layer CNN*, sebuah layer yang akan menentukan fitur, atau hal apa yang dirasa penting didalam sebuah gambar. Hal ini dilakukan dengan memberikan filter (perkalian matrix) 3x3 (Gambar 2.9) terhadap gambar yang diberikan (Gambar 2.10 kiri) . Dimana hasilnya adalah *feature map* (Gambar 2.10 kanan). *Layer* ini menghasilkan *feature maps* dan *region proposal*, atau proposal / perkiraan mengenai dimana sebuah objek berada, untuk selanjutnya di terima oleh *layer RoI Align*. Pada layer ini, *feature maps* beserta *region proposal*, digabungkan menjadi satu buah data untuk selanjutnya dilakukan perhitungan menggunakan dan menghasilkan daerah yang memiliki objek. *Layer* selanjutnya, sekaligus *layer* terakhir dalam proses training, terbagi menjadi 3 buah bagian, yaitu *softmax classifier*, *boundary box*, dan *mask prediction*. *Softmax* merupakan sebuah fungsi yang menerima banyak input, dan menampilkannya dalam rentang 0 – 1. Fungsi

dapat dilihat pada Gambar 4.7. Fungsi ini akan menampilkan probabilitas keyakinan program terhadap sebuah objek. Semakin yakin program terhadap sebuah objek merupakan jawaban yang benar, semakin mendekati hasil angka 1, begitu pula sebaliknya.

Selanjutnya, setelah proses *training* selesai, model akan mengeluarkan *weights* yang berfungsi sebagai persentase, atau sebagai sebuah fungsi yang akan menentukan *output* dari setiap neuron yang ada di dalam layer, dengan pembobotannya, sehingga menghasilkan prediksi yang sesuai dengan data yang diberikan.

C. *Weights Output*

Setelah proses *training* selesai maka program akan mengeluarkan *file* berupa hasil pembelajaran atas gambar-gambar yang telah di berikan. Hasil analisa ini di sebut sebagai *weights*. *Weights* merupakan persentase untuk menentukan *neuron* mana yang harus digunakan untuk menghubungkan *neuron* pada sebuah *layer* ke *neuron* pada *layer* berikutnya.

Hasil dari *weight* oleh model digunakan untuk melakukan prediksi atas data yang diberikan.

D. *Predictions*

Proses ini menggunakan *weights* yang sudah dihasilkan untuk melakukan prediksi terhadap gambar ataupun video dengan kualitas rendah

dan berusaha melakukan deteksi terhadap objek yang menyerupai dengan data yang sudah dilatih.