



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 I

PENDAHULUAN

1.1. Latar Belakang

Deteksi terhadap manusia menggunakan *computer vision* merupakan salah satu bidang yang cukup banyak mendapatkan perhatian dewasa ini. Namun hampir keseluruhan penelitian yang dilakukan dalam bidang ini menggunakan gambar beresolusi tinggi, dengan peralatan yang canggih (Hinduja, 2016) (Vasiljevic , Chakrabarti, & Shakhnarovich, 2017).

Evaluasi terhadap sebuah kualitas gambar, cenderung didapatkan dari sudut pandang seorang manusia. Namun belum tentu apa yang dilihat manusia sama dengan apa yang dimengerti oleh sebuah program. Sebuah gambar yang sama di mata seorang manusia, bisa jadi tidak sama pada saat di analisa oleh sebuah program (Dodge & Karam, 2016).

Beberapa faktor yang dapat mempengaruhi kualitas sebuah gambar berdampak pada hasil akhir analisa dari sebuah program adalah *blur*, *noise*, *obstacle*, pencahayaan, dan lain sebagainya. Dimana faktor tersebut dapat muncul dikarenakan ketidakpastian kondisi lapangan yang mungkin terjadi pada saat gambar didapatkan. Seperti pada kamera yang di pasang pada *dashboard* mobil, atau *cctv* yang di pasang pada tempat yang tinggi, yang meliputi permukaan yang luas (Vasiljevic , Chakrabarti, & Shakhnarovich, 2017) (Dodge & Karam, 2016) (Al-Azzeh, Zahran, & Alqadi, 2018) (Hinduja, 2016).

Tujuan dari penelitian ini adalah untuk mencari cara untuk membuat program yang mampu melakukan deteksi terhadap manusia, dalam gambar yang memiliki kualitas rendah, sehingga dapat melakukan deteksi terhadap objek manusia pada gambar yang memiliki kualitas rendah. Salah satu contoh penerapan yang dapat di aplikasikan adalah pada tempat-tempat yang memerlukan pengawasan secara *remote*, dimana tidak terdapat banyak aktifitas manusia, seperti pada *data storage center*, *server center*, ataupun tempat-tempat lain yang jarang ada aktifitas manusia, namun memerlukan pengawasan terhadap intrusi dari manusia.

Gambar dengan kualitas rendah dapat menjadi masalah bagi *model* yang di *train* dengan menggunakan data yang memiliki kualitas tinggi, tanpa memiliki distorsi, ataupun hal lain yang mempengaruhi kualitas dari sebuah gambar. Ketika *model* diberikan gambar yang memiliki kualitas yang berbeda dengan gambar yang diberikan untuk dipelajari, ditemukan adanya degradasi terhadap kemampuan model melakukan deteksi terhadap objek yang terdapat di dalam gambar (Dodge & Karam, 2016) (Vasiljevic , Chakrabarti, & Shakhnarovich, 2017). Hal ini dapat ditemukan pada sumber gambar dari rekaman *CCTV* pada tempat-tempat yang memerlukan pengawasan, dan cenderung gelap, seperti *data center* di mana lampu untuk menerangi ruangan *server* tidak akan terlalu terang, karena akan memakan daya listrik yang cukup besar, sehingga rekaman yang di hasilkan oleh *CCTV* di tempat tersebut cenderung memiliki kualitas gambar yang rendah.

Penelitian ini akan menggunakan metode *Mask R-CNN*, sebuah metode yang dapat melakukan segmentasi objek, sebuah metode untuk melakukan klasifikasi,

semantic segmentation, deteksi objek, dan melakukan *instance segmentations*. (He, Gkioxari, Dollar, & Girshick, 2017)

Metode ini dipilih karena dapat memberikan *output* berupa gambar yang sudah diberikan area mengenai di mana objek yang kita cari dari sebuah gambar, sehingga dapat membantu tujuan penelitian ini, yaitu melakukan deteksi terhadap manusia pada gambar dengan kualitas gambar yang rendah.

Mask R-CNN dipilih karena merupakan salah satu metode, yang dapat memberikan *instance segmentation* pada sebuah gambar. Sebuah tantangan bagi sebuah program untuk melakukan *instance segmentation* (Abdulla, 2019). Untuk dapat mencapai *instance segmentation*, diperlukan beberapa langkah yang harus dilalui, yaitu *classification*, mengetahui ada objek apa saja yang ada pada data yang diberikan (Abdulla, 2019), *semantic segmentation*, yaitu melakukan identifikasi pada sebuah gambar menggunakan *pixel* dari gambar tersebut (Wang, 2019). *Object Detection*, melakukan deteksi terhadap objek dari data yang disediakan, lalu *Instance Segmentation*, melakukan segmentasi terhadap sebuah objek, beserta dengan keterangan mengenai *pixel* yang dimiliki oleh setiap objek yang dideteksi.

Selain daripada itu, *Mask R-CNN* juga di pilih karena mampu memberikan deteksi berupa campuran antara *bounding box* dengan *mask* yang dapat memberikan secara gamblang dimana objek tersebut berada, dan *pixel*¹ mana saja yang menjadi bagian dari objek tersebut, jika di bandingkan dengan *Faster R-CNN* yang hanya dapat memberikan *bounding box* terhadap objek yang berhasil di

¹ Bagian terkecil dari sebuah gambar.

deteksi (He, Gkioxari, Dollar, & Girshick, 2017) (Ren, He, Girshick, & Sun, 2015). Deteksi sampai kepada level pixel ini memungkinkan proses identifikasi lebih lanjut terhadap objek manusia yang dideteksi tersebut, seperti perkiraan postur tubuh, perkiraan tinggi, dan lain sebagainya, yang dapat digunakan untuk melakukan pemetaan ataupun perkiraan data mengenai orang tersebut yang dapat kemudian dilakukan pencarian mengenai siapa tersangka yang paling mendekati dengan postur tubuh tersebut.

1.2. Rumusan Masalah

Banyak penelitian terdahulu yang telah dilakukan yang mampu melakukan deteksi manusia pada sebuah gambar. Namun seluruh penelitian tersebut mengandalkan sumber data gambar yang memiliki tingkat kualitas yang sangat tinggi dengan peralatan yang canggih. Di mana sebaliknya, sedikit penelitian yang membahas bagaimana cara melakukan deteksi manusia dalam kondisi yang tidak sempurna, seperti kualitas gambar yang rendah, terdapat distorsi, dan lain sebagainya. Faktor kualitas sebuah gambar sangatlah mempengaruhi sebuah program untuk melakukan deteksi terhadap manusia dalam sebuah gambar.

Berdasarkan rumusan masalah di atas, berikut adalah rumusan masalah yang ada di dalam pembahasan ini adalah “Apakah *Mask R-CNN* dapat melakukan deteksi manusia pada gambar yang memiliki kualitas gambar rendah yang memiliki *blur* dan *noise*.”

1.3. Batasan Masalah

Batasan masalah pada penelitian ini ditetapkan dengan tujuan mempermudah pelaksanaan penelitian serta mempersempit jangkauan deteksi dari program dalam melakukan deteksi terhadap gambar.

Pembuatan penelitian ini memiliki batasan sebagai berikut:

1. Data gambar diambil dari internet, yaitu:
 - a. *InriaPerson Dataset* (Dalal, 2019)
 - b. *MPIHumanPose* (Andriluka, Pishchulin, Gehler, & Schiele, Bernt, 2014)
2. Menggunakan *library mask R-CNN* dengan *tensorflow*
3. Deteksi hanya di batasi kepada deteksi terhadap manusia
4. Kualitas rendah di batasi kepada gambar atau video yang diberikan distorsi menggunakan beberapa jenis metode seperti:
 - a. *Gaussian Blur* (Dodge & Karam, 2016)
 - b. *Salt and Pepper Noise* (Al-Azzeh, Zahran, & Alqadi, 2018)

1.4. Manfaat Penelitian

1.4.1 Tujuan Penelitian

Penelitian ini dilakukan dengan beberapa tujuan sebagai berikut:

- Mengetahui seberapa baik metode *Mask R-CNN*. dalam melakukan deteksi manusia pada gambar yang memiliki kualitas rendah
- Mengetahui kekurangan yang dimiliki dalam melakukan deteksi manusia pada gambar yang memiliki kualitas rendah menggunakan metode *Mask R-CNN*.

1.4.2 Manfaat Penelitian

Manfaat yang ingin di capai dalam melakukan penelitian ini adalah:

1. Untuk memberikan manfaat secara teoritis dan berguna sebagai sumbangan untuk pendidikan dan penelitian selanjutnya yang akan mengembangkan hasil dari penelitian ini pada bidang *computer vision*
2. Melakukan klasifikasi manusia pada gambar yang memiliki kualitas rendah, sehingga dapat membantu pihak-pihak atau instansi yang membutuhkan untuk melakukan deteksi terhadap manusia pada gambar yang memiliki kualitas rendah.