



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

Materi pornografi bukan suatu hal yang rahasia pada era masa kini (Moustafa, 2015). Berbagai media sosial sudah sejak lama melakukan penutupan konten dewasa baik berupa gambar maupun video dengan menggunakan *Intelligent Image Analysis*. Namun, metode ini membutuhkan banyak sekali data untuk meningkatkan kinerja klasifikasi yang menjadikan kata kuncinya (Moustafa, 2015). Beberapa cara yang telah dilakukan untuk melakukan identifikasi konten pornografi pada suatu gambar adalah dengan melakukan klasifikasi berdasarkan deteksi kulit manusia (Moustafa, 2015). Dengan kata lain, semakin sedikit kulit yang tertutup oleh busana, peluang gambar tersebut dikatakan porno semakin besar (Srisa-An, 2016).

Pada penelitian tersebut, tingkat akurasi yang didapatkan sebesar 85% dengan rata-rata waktu 213.14 msec per gambar (Srisa-An, 2016). Namun, cara investigasi gambar dan video porno ini hanya berbasis warna kulit yang akan menjadi indikator, sedangkan jika gambar yang digunakan contohnya adalah wajah yang didekatkan akan tergolong porno padahal tidak sama sekali (Moustafa, 2015). Dengan alasan ini, *computer vision model* yang dirancang tidak hanya mendeteksi dari kulit pada gambar atau video tetapi juga suatu bentuk dari objek tersebut (Moustafa, 2015).

Bentuk dari objek yang dihasilkan dari deteksi merupakan hasil data yang dipelajari oleh komputer melalui *convolution layer* (Defferrard, Bresson, & Vandergheynst, 2016). *Convolution layer* tersebut terbentuk dari proses ekstraksi layer dan klasifikasi yang terjadi secara otomatis (Moustafa, 2015).

Convolution layer adalah teknik *Deep Learning* dengan sebutan *Convolutional Neural Network* dan mulai digunakan untuk klasifikasi gambar atau video porno (Moustafa, 2015). Hal ini menunjukkan suatu peningkatan dibanding metode klasifikasi yang menggunakan pendekatan deteksi kulit manusia dengan akurasi 92% menggunakan struktur AlexNet (Moustafa, 2015).

Penggunaan *Deep Learning* yang terjadi kurang lebih dalam sepuluh tahun terakhir telah meningkatkan penggunaan *Machine Learning* dalam berbagai hal (Rusk, 2016). Disisi lain, *Deep Learning* dikenal sebagai *black box* (Alain & Bengio, 2016). Model prediksi yang dihasilkan sulit untuk dijelaskan terlebih cara kerja *black box* model tersebut (Koh & Liang, 2017). Dengan memahami *black box*, kita mampu mengartikan model tersebut, mengetahui peran dari masing-masing layer, serta memahami sisi kebodohan pada model yang dibuat (Alain & Bengio, 2016). Kebodohan *Deep Learning* ini sering kali merusak akurasi dan mengacaukan klasifikasi yang ada. Proses perusakan ini dilakukan pada proses *training* oleh sebab itu dikenal dengan *training-set attack* (Koh & Liang, 2017). *Black box* bisa menjadi suatu hal yang berbahaya jika tanpa kemampuan interpretasi yang baik, sisi ini adalah sisi yang tidak mampu ditebak oleh manusia dari *deep learning* (Guidotti et al., 2018).

Black box didefinisikan sebagai sebuah hal internal yang dimiliki model yang tidak bisa diartikan oleh manusia. Pentingnya pengungkapan *black box* yakni agar model mampu mengungkap temuan dalam data yang menjelaskan keputusan atau menjelaskan bagaimana model tersebut bekerja (Guidotti et al., 2018).

Berangkat dari permasalahan ini, meskipun *Deep Learning* mendapatkan akurasi yang lebih baik pada prediksi dan klasifikasi gambar maupun video porno dibandingkan dengan metode deteksi kulit manusia, tetapi penelitian sebelumnya tidak mengungkap *black box* yang ada pada model *Deep Learning* tersebut. Target penelitian ini adalah mengungkap *black box* pada model, sehingga model tersebut mampu diinterpretasikan. Dengan memahami *black box* tersebut, akurasi dan klasifikasi pada model dapat ditingkatkan.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, disusunlah rumusan masalah sebagai berikut:

- a. Hal apa yang mengindikasikan bahwa suatu gambar memiliki konten porno dari model *Deep Learning* yang diuji? (*black box*)
- b. Bagaimana pembelajaran model pada penelitian berlangsung antar layer?

1.3. Batasan Masalah

Dalam penelitian ini, peneliti memberikan beberapa batasan masalah agar dalam pelaksanaan dapat lebih mudah dicakup lingkungan yang diinginkan serta sesuai dengan yang diharapkan. Pembuatan penelitian ini dibatasi hanya pada batasan berikut:

1. Penggunaan data berdasarkan data yang diperoleh dari NPDI *Dataset*.
2. AlexNet model yang digunakan dengan input berukuran 224 x 224 dengan 5 layer konvolusi dan 3 *fully-connected* layer.
3. Labelisasi didasarkan pada kategori yang sesuai dengan pengelompokan data dari NPDI *Dataset*.
4. Proses eksplorasi dilakukan tanpa terikat dengan tingkat akurasi yang didapatkan.

1.4. Tujuan dan Manfaat Penelitian

Adapun tujuan dan manfaat dilakukannya penelitian ini adalah:

- a. Mengetahui apa yang dipelajari model, sehingga model mampu mengindikasikan suatu gambar memiliki konten porno (*black box*).
- b. Dengan mengetahui pembelajaran model, mampu dilakukan peningkatan akurasi pada model agar mendapatkan akurasi lebih baik.