

BAB I. PENDAHULUAN

1.1. Latar Belakang

Saat ini, kecerdasan buatan telah masuk ke berbagai bidang. Kecerdasan buatan sudah masuk ke bidang forensik digital dan bidang keamanan. Beberapa jenis forensik yang dapat diterapkan kecerdasan buatan di antaranya: *Network Forensics*, *Cloud Investigations*, dan *Big vs. Large Data* [1]. Pada bidang keamanan, kerap kali menerapkan kecerdasan buatan berbasis pengolahan citra(kamera). Dalam tulisan A. Pazhampilly Sreedevi, kecerdasan buatan digunakan untuk mendeteksi pencurian mobil melalui pengolahan citra dengan menggunakan *Principal Component Analysis* (PCA) Algorithm [2].

Pengenalan citra untuk identifikasi objek termasuk wajah saat ini sudah sangat berkembang dalam beberapa tahun terakhir. Perkembangan *Convolution Neural Networks* (CNNs) menjadi penyebab utama perkembangan pesat pengenalan objek berbasis citra. Penerapan CNNs dalam kasus identifikasi wajah pada umumnya memerlukan jaringan saraf tiruan yang dalam, sehingga mempengaruhi kecepatan pada saat *training* ataupun *prediction*.

Suatu model jaringan saraf tiruan biasanya hanya menyelesaikan masalah yang spesifik, apabila ingin menyelesaikan beberapa masalah, diperlukan beberapa model sehingga saling bekerja membentuk sebuah sistem yang terintegrasi.

Kecerdasan buatan dapat digunakan dalam membantu manusia dengan kondisi kesehatan tertentu seperti tuna netra. Mengutip pertuni.or.id, paling tidak sekurang-kurangnya 3,750,000 penduduk Indonesia mengalami kondisi kebutaan ataupun melemahnya penglihatan [3]. Menggunakan Kecerdasan buatan, kita dapat membantu tuna netra untuk melihat dunia seperti yang dilakukan NoonGil. NoonGil ialah sebuah *software* fiktif dari drama korea berjudul 'STARTUP'. NoonGil mampu mengenali objek berbasis visual dan membacakan objek apa saja yang ada dalam visual tersebut. Tujuan utama NoonGil ialah membantu tuna netra agar tetap melihat dunia melalui bantuan kecerdasan buatan. Tulisan ini akan

meningkatkan kemampuan NoonGil agar dapat mengenali orang yang ada dalam sebuah gambar. Agar memudahkan penyebutan, maka sistem yang diajukan penulis disebut NoonGil Lens⁺.

Membuat sistem kecerdasan buatan yang dapat mendeteksi objek dan mengenali manusia dalam sebuah model *deep learning* sangat sulit dilakukan. Model yang dibutuhkan akan lebih kompleks, di samping itu rentan terhadap masalah *underfitting* apabila dataset kurang mendukung ataupun *loss function* yang kurang tepat. Satu model yang kompleks juga berpotensi *overfitting* sehingga model tidak dapat bekerja dengan baik pada *validation set*. Umumnya, masalah tersebut diselesaikan dengan membentuk dua buah model yang terpisah yaitu model untuk mendeteksi objek dan model untuk mengenali manusia. Apabila kedua model ini tidak dilakukan optimasi, maka akan meningkatkan waktu komputasi ataupun mengurangi performa dari model-model tersebut.

YOLO (*You Look Only Once*) mempunyai 26 *layers* yang terdiri dari 24-layer konvolusi dan 2 *fully connected layer*. Arsitektur YOLO mampu mendeteksi objek hingga kecepatan 55 FPS pada GPU Titan X [4]. Arsitektur FaceNet mempunyai 17 *layers*, terdiri dari 11 *Inception layer* yang mempunyai 2 *layer*, secara implisit FaceNet mempunyai 28 *layer* [5]. Kecepatan eksekusi FaceNet sekitar 0.42 dari waktu yang diperlukan YOLO. Artinya, kedua model harus cukup kompleks agar mempunyai performa yang baik dalam menjalankan tugasnya.

Meninjau masalah di atas, maka penulis mengajukan sebuah pendekatan algoritma untuk mendeteksi wajah dalam sebuah citra, tidak terbatas pada citra bergerak(video). Pendekatan dilakukan dengan *2-Level Classification*, di mana tahap pertama mendeteksi objek pada gambar, kemudian dilanjutkan dengan identifikasi wajah pada objek yang dikenali manusia. Penulis mengajukan pendekatan *Region Selection* untuk menengahi antara YOLOv3 dan FaceNet agar sistem menghasilkan performa yang baik dengan waktu eksekusi secepat mungkin.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka permasalahan yang dapat dirumuskan adalah sebagai berikut:

1. Bagaimana mendesain pendekatan algoritma yang dapat menyelesaikan dua masalah (objek deteksi dan identifikasi wajah seorang) dalam satu proses.
2. Bagaimana cara meminimalisir *trade off* antara kecepatan dan akurasi pada pendekatan algoritma yang digunakan.
3. Bagaimana cara agar FaceNet dapat mengenali wajah yang ada pada suatu gambar cuplikan dari luaran YOLOv3 untuk kelas 'person'.

1.3. Batasan Masalah

1. Penelitian ini dilakukan di atas layer *You Look Only Once* (YOLOv3) [6].
2. Penelitian tidak mengembangkan/mengimplementasi YOLOv3 karena alasan waktu, namun menggunakan implementasi yang tersedia secara *open source* dari *github*.
3. Sistem yang dikembangkan tidak diimplementasikan ke dalam *web service* ataupun platform lainnya.

1.4. Tujuan Penelitian

1. Mengajukan pendekatan algoritma mendeteksi objek dan pengenalan wajah seseorang pada gambar.
2. Meminimalisir *trade off* kecepatan dan akurasi CNNs pada kasus multi-klasifikasi, dalam kasus ini deteksi objek dan pengenalan wajah manusia.

1.5. Metode Penelitian

Tugas Akhir dilaksanakan dengan metode Penelitian dan Pengembangan (R&D). Tulisan berfokus kepada penelitian dan mengembangkan suatu sistem dalam hal ini NoonGil Lens⁺ untuk mengurangi *trade-off* akurasi dan kecepatan.

Pelaksanaan penelitian akan dilaksanakan mulai dari studi literatur, pengumpulan dataset dan uji coba YOLOv3, pengembangan model FaceNet untuk *Face Verification*, *training* FaceNet, pengujian dan evaluasi model FaceNet, pengembangan sistem NoonGil Lens⁺, pengujian sistem dan pengumpulan data hasil pengujian, dan dilanjutkan dengan penulisan hasil penelitian. Setelah dataset terkumpul dan pengujian YOLOv3 berhasil, dilanjutkan dengan proses pengembangan model FaceNet, proses ini mengembangkan model dari nol sesuai tulisannya [5]. Seluruh implementasi dilakukan menggunakan Tensorflow dan Pyhon.

1.6. Sistematika Penulisan

Tulisan ini terdiri atas 3 bagian utama yaitu, bagian awal, bagian isi, dan bagian akhir. Bagian awal berisi tentang dokumen-dokumen pendukung Tugas Akhir, sampul, kata pengantar, daftar isi, daftar gambar, daftar tabel, dan intisari. Pada bagian isi, bab 1 berisi tentang latar belakang kenapa penelitian dilakukan, rumusan masalah pada penelitian, batasan masalah dalam penelitian, tujuan dilakukannya penelitian, metode penelitian yang digunakan, dan sistematika penulisan. Bab 2 berisi tentang penelitian terdahulu yang melandasi penelitian yang dilakukan penulis. Bab 3 membahas landasan teori yang terdiri atas *Deep Learning dan Neural Network*, *Konvolusi Citra*, *Image Processing*, dan *Tensorflow*. Bab 4 berisi tentang pembahasan dataset, dan pengembangan model/metode yang digunakan serta evaluasi model yang dipakai. Untuk bab 5, penulis menyampaikan dan membahas hasil penelitian yang ditemukan, di antaranya hasil pengembangan sistem, tingkat akurasi sistem, efisiensi sistem, akurasi model FaceNet yang dikembangkan, dan pembahasan tentang *Region Selection*, salah satu pendekatan yang diajukan penulis. Bab 6, berisi tentang kesimpulan dan saran untuk penelitian selanjutnya. Bagian akhir tulisan berisi daftar pustaka.