

BAB I: PENDAHULUAN

1.1 Latar Belakang

Pendeteksian pada wajah (*face detection*) merupakan sebuah usungan teknologi *biometrik* yang banyak sekali dikembangkan dan terus berjalan hingga dewasa ini, yang di mana terdapat suatu perangkat teknologi dengan algoritma pendeteksian pada wajah untuk melakukan sebuah identifikasi kecocokan data yang diambil dari individu tertentu untuk menentukan suatu kesamaan dengan individu yang terdapat di dalam *database* untuk proses verifikasi.

Dengan bantuan *face detection* para pihak kepolisian juga akan sangat terbantu untuk mencari para perilaku kriminal yang sedang diburu. Namun, apakah teknologi ini dapat melakukan identifikasi wajah para kriminal yang sedang menyamar dan bersatu dengan riuh ramainya kota? Sebut saja, perubahan yang terjadi dari para kriminal seperti mulai tumbuhnya kumis dan brewok, atau penangkapan citra para kriminal dengan *cctv* yang menggunakan aksesoris pada wajah, apakah sistem ini masih akurat? Apakah sistem ini masih bisa mengenali mereka dengan pencocokan pada *record* terakhir yang ada pada *database*?

Penelitian ini banyak sekali menarik perhatian pengamat *computer vision* untuk menjadi sebuah bahan diskusi di antara para praktisi kecerdasan buatan (*artificial intelligence*) dan *machine learning*. Dengan muncul berbagai variasi algoritma pendeteksian pada wajah manusia, diharapkan pendeteksian pada wajah dapat terus berkembang menjadi semakin akurat dan mampu menjawab semua permasalahan yang telah dijabarkan di atas.

Sistem pendeteksian pada wajah (*face-detection*) sudah banyak diaplikasikan dengan berbagai macam metode, di antaranya; Metode Haar Feature [1], Metode Cascade Classifier [1], Metode 3D Morphable [2], Metode 3D *Face Recognize* [3], Metode Bayesian Framework [4], Metode SVM [5], Metode EP [6], Metode EBGM [7], Metode HMM [8], Metode Kernel [9].

Juga ada sejumlah teknik yang berhasil dalam pendeteksian wajah dengan citra tegak atau frontal dari depan [10]-[14], walaupun dengan istilah *frontal* atau *tegak*

berbeda-beda pada berbagai perpektif dari satu sistem ke sistem lainnya, dalam kenyataannya ada banyak sekali algoritma yang diterapkan dalam pendeteksian pada wajah namun tidak dari banyaknya algoritma dapat berfungsi optimal ketika citra diputar. Pada *framework* Paul Viola dan Michael Jones [1], merupakan sebuah *framework* yang dapat mendeteksi wajah dengan berbagai pose secara real-time yang ditentukan dari menggabungkan detektor-detektor yang masing-masing mengambil alih satu sudut pandang. *Framework* dapat menangani berbagai masalah yang *face detection* lain alami seperti: *non-frontal*, *non-upright*. Dengan menyajikan beberapa metode umum untuk memilih di antara sekumpulan detektor-detektor saat dilakukannya pemindain untuk melakukan masukan pada citra.

Adapun arsitektur biometrik yang bekerja pada fitur wajah seperti FaceNet. FaceNet merupakan salah satu arsitektur yang berguna dalam proses biometrik, namun alasan penulis tidak menggunakan arsitekturnya karena FaceNet bekerja pada "face-recognize" [15] sedangkan penulis hanya sampai pada proses "face-detection" dan "facial expression recognize", penggunaan dari FaceNet akan dibutuhkan ketika ingin membangun sebuah arsitektur seperti *face unlock* karena melakukan pengkategorian citra digital dengan sampel beberapa wajah manusia untuk dilakukan proses *triple loss* pada *face-id* terdaftar. Contoh: terdapat 4 *face-id* dengan berbagai gambar berbeda dari ke-4 wajah manusia, setelah proses training wajah dalam proses pengujian akan dibandingkan ke dalam 4 *face-id* yang akan dilakukan proses perbandingan dengan nama arsitektur *triple loss* dari salah satu dari *face-id* pemilik smartphone, apabila ditemukan kecocokan maka kunci pada layar smartphone akan terbuka begitu juga sebaliknya. Sehingga penulis rasa arsitektur ini tidak cocok pada proses training terhadap kelas-kelas ekspresi wajah dari tiap objek manusia yang berbeda. Karena proses pengkategorian pada "facial expression recognize" melakukan perbandingan pada arsitektur kontur wajah saat objek wajah berekspresi.

Dari metode-metode yang sudah disebutkan di atas, di sini akan dicoba implementasi dari pendeteksian wajah (*Face Detection*) menggunakan *framework Viola-Jones*, sehingga tugas akhir ini akan dikembangkan sebuah program terpadu terhadap pendeteksian pada wajah dengan memanfaatkan metode *Haar-like*

Features dan *Cascade Classifier* untuk menentukan proses pendeteksian pada wajah secara tepat. Setelah wajah terdeteksi akan dilakukan prediksi dari ekspresi wajah dari citra digital terdeteksi wajah, menggunakan metode *Transfer Learning* pre-trained model MobileNetV2 dengan dataset training berupa citra digital *grayscale* ukuran 48 x 48. Pembangunan arsitektur dari *deep learning* membutuhkan dataset yang sangat banyak untuk proses data training, maka penulis memanfaatkan arsitektur *deep learning* pada MobileNetV2, karena kebutuhan dataset training yang paling sedikit dari arsitektur *deep learning* pada metode *Transfer Learning* lainnya [26]. Adapun pada pre-trained terdapat proses *tuning* yang dilakukan oleh para peneliti untuk menyesuaikan kelas-kelas kategori yang ada ke dalam kebutuhan para peneliti karena pada setiap pre-trained model memiliki kapabilitas dalam melakukan proses identifikasi ke dalam 1000 kelas-kelas kategori [16]. Pada penelitian ini penulis membagi kelas-kelas ekspresi ke dalam 7 kategori.

1.2 Perumusan Masalah

Berdasarkan pemaparan di atas, didapatkan beberapa poin permasalahan yang ditemui yaitu sebagai berikut:

1. Bagaimana kapabilitas dari *haar-features* dalam pendeteksian terhadap wajah pada citra digital one person on frame?
2. Bagaimana kapabilitas dari *haar-features* dalam pendeteksian terhadap wajah pada citra digital many person on frame?
3. Bagaimana hasil pengujian citra-citra digital terhadap penentuan ekspresi wajah?
4. Bagaimana tingkat akurasi pengujian pada ekspresi wajah?
5. Bagaimana resposibilitas dari proses deteksi wajah saat live demo webcam?

1.3 Tujuan Penelitian

1. Mengetahui cara kerja dari *Haar-like Features* dan *Cascade Classifier* pada implementasi algoritma Viola-Jones.

2. Pembangunan sebuah sistem pengujian pendeteksian wajah dan prediksi ekspresi wajah.
3. Menentukan kelayakan sistem pengujian prediksi ekspresi wajah dengan dataset proses training data yang sedikit.

1.4 Batasan Masalah

1. Dataset berupa citra digital dengan ukuran piksel 48 x 48.
2. Dataset untuk proses training data sebanyak 2.450 data.
3. Proses training dataset menggunakan pre-trained model MobileNetV2.
4. Proses training dataset dilakukan untuk proses prediksi ekspresi wajah.
5. Pengujian dilakukan dengan gambar bebas lisensi dari internet.

1.5 Manfaat Penelitian

Dapat membantu peneliti selanjutnya yang akan menggunakan metode *Viola-Jonex* dalam mengembangkan program *face detection* untuk mendapatkan tingkat akurasi yang lebih akurat

1.6 Metode Penelitian

1.6.1 Studi Literatur

Penulis sudah melalui tahapan pembelajaran terhadap berbagai jurnal-jurnal terkait yang sudah ada diinternet. Dengan menggunakan *framework Viola-Jones* dan *Open CV python* penulis merasa bahwa metode ini dirasa mampu memenuhi kriteria yang penulis inginkan dalam pembuatan program pada proyek akhir ini.

1.6.2 Pengujian Sistem

Setelah program sudah siap, penulis akan melakukan pengujian terhadap citra-citra uji, untuk menentukan dimana dominasi ekspresi yang ada pada citra-citra uji terhadap ketujuh kelas-kelas ekspresi wajah.

1.6.3 Pengumpulan Data

Pada penelitian ini penulis menggunakan sampel-sampel citra yang sudah lebih dulu di *grayscale* pada *kaggle.com: "FER 2013"* dengan total 28.709. Namun penulis hanya mengambil sekitar 2.450 untuk proses data training dengan

pembagian 350 dataset untuk setiap kelas ekspresi wajah, yang penulis rasa cukup akurat untuk melakukan proses pengujian sampel-sampel uji coba. Yang dimana proses validasi data penulis menggunakan data "test" pada "FER 2013" dengan pembagian 110 dataset untuk setiap kelas ekspresi wajah. Adapun pembagian dari dataset "pengujian" merupakan gambar RGB, penulis membaginya ke dalam poin-poin berikut ini:

1. Pengujian face detection *One Person in Frame*, menggunakan dataset yang penulis unduh dari situs gambar bebas lisensi Unsplash.com sebanyak 20 sampel data dan 5 sampel data merupakan gambar wajah dari penulis sendiri dengan objek wajah tegak lurus pada gambar (citra digital).
2. Pengujian face detection *Many persons in frame*, menggunakan dataset yang penulis unduh dari situs gambar bebas lisensi Unsplash.com sebanyak 20 sampel data dengan pembagian: objek wajah tegak lurus pada gambar (citra digital) sebanyak 13 sampel dan objek wajah tidak tegak lurus pada gambar sebanyak 7 sampel data.
3. Pengujian facial expression recognize, menggunakan 20 dataset pengunduhan pada proses face detection *One Person in Frame* yang lulus tahap uji pendeteksian wajah sebanyak 15 data dan 5 data pengunduhan tambahan.
4. Pengujian face detection live webcam demo dengan proses pendeteksian yang dilakukan secara langsung dari webcam lokal laptop penulis dengan spesifikasi sebaga berikut: 0.3 MP camera with single mic.

1.6.4 Penulisan Hasil Penelitian

Hasil dari pengujian terhadap tugas akhir ini berupa persentase dan akurasi dari arsitektur *face detection*, *facial expression recognition*, serta tingkat responsibilitas dari arsitektur *face detection* terhadap demo yang dilakukan secara real-time pada webcam lokal penulis.